Mengenlehre

Von

Dr. F. Hausdorff
em. o. Professor der Mathematik
an der Universität Bonn

Dritte Auflage

Mit 12 Figuren

Walter de Gruyter & Co.
vormals G. J. Göschen'sche Verlagshandlung
J. Guttentag, Verlagsbuchhandlung — Georg
Reimer — Karl J. Trübner — Veit & Comp.
Berlin W 10 und Leipzig
1935

47

Aus dem Vorwort zur 2. Auflage.

Das vorliegende Werk versucht, die wichtigsten Theoreme der Mengenlehre mit vollständig ausgeführten Beweisen darzustellen, so daß seine Lektüre nirgends der Ergänzung durch fremde Hilfsmittel bedarf, wohl aber ihrerseits zum tieferen Eindringen in die umfangreiche Literatur befähigt. Es setzt beim Leser keine höheren mathematischen Kenntnisse als etwa die Anfangsgründe der Differential- und Integralrechnung, allerdings aber eine gewisse Schärfe des abstrakten Denkens voraus und wird von Studierenden in mittleren Semestern mit Erfolg gelesen werden können. Schwierigere Gegenstände am Ende der einzelnen Kapitel mögen beim ersten Studium überschlagen werden; der Leser, der nur die einfachsten Tatsachen der Punktmengentheorie kennen lernen will, kann nach flüchtiger Durchsicht der ersten beiden Kapitel sogleich das sechste in Angriff nehmen. — Den Fachgenossen hoffe ich wenigstens in formaler Hinsicht, insbesondere durch Verschärfung der Sätze, Vereinfachung der Beweise und Befreiung von unnötigen Voraussetzungen, einiges Neue zu bieten.

Die Stoffauswahl muß bei einem so ausgedehnten, sich täglich noch erweiternden Gebiet immer einen etwas subjektiven Charakter haben und manche Wünsche (auch die des Verfassers) unerfüllt lassen: ein Lehrbuch kann nicht die Vollständigkeit eines Referats anstreben. In diesem Falle kam noch hinzu, daß der äußere Rahmen, in dem das Buch jetzt erscheint, eine starke Einschränkung des Umfangs gegenüber der ersten Auflage (Grundzüge der Mengenlehre, Leipzig 1914) erforderte, wodurch eine Umarbeitung in den kleinsten Teilen notwendig wurde, der ich schließlich eine völlige Neubearbeitung vorgezogen habe. Von den damals behandelten Gegenständen glaubte ich am ehesten die etwas isoliert für sich stehende Theorie der geordneten Mengen (bis auf einen geringen Rest) und sodann die Einführung in die Lebesguesche Maß- und Integraltheorie opfern zu können, für die es ja an sonstigen Darstellungen nicht mangelt. Mehr als diese Streichungen wird vielleicht bedauert werden, daß ich zu weiterer Raumersparnis in der Punktmengenlehre den *topologischen* Standpunkt, durch den sich die erste Auflage anscheinend viele Freunde erworben hat, aufgegeben und mich auf die einfachere Theorie der *metri-*

schen Räume beschränkt habe, wofür ein flüchtiger Überblick (§ 40) über die topologischen Räume kein genügender Ersatz ist. Schließlich habe ich die Allgemeinheit wie nach oben, so auch nach unten begrenzt und die spezielle Theorie der Euklidischen Räume (z. B. den Jordanschen Satz über ebene Kurven) weggelassen, d. h. etwa alles, was auf approximierenden Polygonen und Polyedern beruht; man wird also zwar eine Menge von Sätzen über den Euklidischen Raum finden, aber nur solche, die für ihn als Sonderfall eines separablen oder vollständigen oder lokal zusammenhängenden Raumes o. dgl. gelten. — Diesen Abstrichen steht als Zugang gegenüber eine vollständigere Behandlung der Borelschen und der 1917 entdeckten Suslinschen Mengen, sowie der Baireschen Funktionen; auch stetige Abbildung und Homöomorphie ist eingehender als damals berücksichtigt. Zu einer Diskussion über Antinomien und Grundlagenkritik habe ich mich jetzt ebensowenig wie damals entschließen können.

Vorwort zur 3. Auflage.

Die immer weiter anhaltende lebhafte Entwicklung der Mengenlehre hätte an sich wieder eine Neubearbeitung dieses Buches wünschenswert gemacht; aus äußeren Gründen mußte davon abgesehen werden. Demgemäß sind die ersten neun Kapitel ein fast unveränderter Abdruck der 2. Auflage. Um aber den inzwischen erzielten Fortschritten wenigstens teilweise gerecht zu werden, habe ich in einem neu hinzugefügten zehnten Kapitel zwei Gegenstände, die mir das besonders zu verdienen scheinen, ausführlich dargestellt und auf drei weitere in kleineren Nachträgen, allerdings ohne Beweise, hingedeutet; der Kreis hätte, wäre nicht der Raummangel hinderlich gewesen, erheblich weiter ausgedehnt werden können.

Inhaltsverzeichnis.

Vorbemerkungen.

Intervalle reeller Zahlen werden mit eckigen oder runden Klammern bezeichnet, je nachdem die Endpunkte mitgezählt werden sollen oder nicht. Für $a < b$ ist also

$$[a, b] \quad [a, b) \quad (a, b] \quad (a, b)$$

die Menge der Zahlen x, die den Bedingungen

$$a \leqq x \leqq b \quad a \leqq x < b \quad a < x \leqq b \quad a < x < b$$

genügen. $[a, b]$ heißt ein *abgeschlossenes*, (a, b) ein *offenes* Intervall, die beiden andern halb abgeschlossen oder halb offen. Für einseitig unendliche Intervalle (abgeschlossene und offene Halbgerade) verwenden wir die uneigentlichen Endpunkte $+\infty$, $-\infty$, die dem Intervall nicht zuzurechnen sind; es ist also

$$[a, +\infty) \quad (a, +\infty) \quad (-\infty, b] \quad (-\infty, b)$$

die Menge der Zahlen x, die den Bedingungen

$$x \geqq a \quad x > a \quad x \leqq b \quad x < b$$

genügen. $(-\infty, +\infty)$ ist die Menge aller reellen Zahlen (die ganze Gerade).

Die größte und die kleinste von endlich vielen reellen Zahlen $x_1, x_2, \ldots, x_n$ heißt ihr *Maximum* und *Minimum*, in Zeichen

$$\max [x_1, x_2, \ldots, x_n], \quad \min [x_1, x_2, \ldots, x_n].$$

Z. B. $\max [2, -3] = 2$, $\min [2, -3] = -3$, $\max [2, 2] = \min [2, 2] = 2$. Ebenso wird die größte und kleinste von unendlich vielen Zahlen bezeichnet, falls eine solche vorhanden ist.

Wenn eine Folge reeller Zahlen $x_1, x_2, \ldots$ nach oben beschränkt ist, d. h. wenn es Zahlen v gibt derart, daß $v \geqq x_n$ für jedes n, so gibt es unter diesen Zahlen v eine kleinste v_1. Sie heißt fast allgemein die *obere Grenze* (Weierstraß; bei einigen auch die obere Schranke) der Folge x_n; wir übersetzen dies mit *Supremum* und schreiben

$$v_1 = \sup [x_1, x_2, \ldots] = \sup x_n.$$

Z. B. $\sup [0, \frac{1}{2}, \frac{2}{3}, \ldots] = \sup \frac{n-1}{n} = 1$. Wenn die Folge ein Maximum hat, fällt es mit dem Supremum zusammen. Genau so wird die *untere Grenze* oder das *Infimum*

$$u_1 = \inf [x_1, x_2, \ldots] = \inf x_n$$

einer nach unten beschränkten Folge erklärt. Entsprechende Bezeichnungen gelten für Zahlenmengen, die nicht in Form einer Folge gegeben sind, etwa

$$\sup_{a \leq x \leq b} f(x).$$

Für eine nach oben beschränkte Folge existieren die sämtlichen Suprema

$$v_n = \sup [x_n, x_{n+1}, x_{n+2}, \ldots]$$

mit $v_1 \geqq v_2 \geqq \ldots$; wenn diese v_n nach unten beschränkt sind und also einen Grenzwert v haben, so heißt dieser der *obere Limes* oder *Limes superior* der Folge, in Zeichen

$$v = \lim\sup x_n = \overline{\lim}\, x_n.$$

Entsprechend ist der *untere Limes* oder *Limes inferior*

$$u = \lim\inf x_n = \underline{\lim}\, x_n$$

zu erklären.

Falls die über Beschränktheit gemachten Voraussetzungen nicht zutreffen, werden die Zeichen $\pm\infty$ verwendet. Z. B. für eine nach oben nicht beschränkte Folge x_n wird $\sup x_n = +\infty$, $\overline{\lim}\, x_n = +\infty$ gesetzt; für eine nach oben beschränkte Folge x_n, für welche die obigen Suprema v_n nicht nach unten beschränkt sind, $\overline{\lim}\, x_n = -\infty$.

Für Konvergenz von Folgen und Funktionen verwenden wir in der Regel das Zeichen des Pfeiles, z. B.

$$x_n \to x \quad \text{soviel wie} \quad \lim x_n = x,$$

ebenso für eigentliche Divergenz ($x_n \to +\infty$, $x_n \to -\infty$).

Eine Behauptung über die natürliche Zahl n ($= 1, 2, 3, \ldots$) gilt *schließlich* oder *für fast alle* n (G. Kowalewski), wenn sie von einem bestimmten n ab ($n \geq n_0$) oder für alle n mit höchstens endlich vielen Ausnahmen gilt; sie gilt *unendlich oft*, wenn sie *für unendlich viele* n gilt (z. B. für $n = 2, 4, 6, 8, \ldots$). Wenn wir von fast allen oder unendlich vielen Gliedern einer Folge x_n sprechen, so meinen wir die zu fast allen oder unendlich vielen Werten n gehörigen Glieder x_n, mögen diese Glieder verschieden sein oder nicht.

Erstes Kapitel.

Mengen und ihre Verknüpfungen.

§ 1. Mengen. [1]

Eine Menge entsteht durch Zusammenfassung von Einzeldingen zu einem Ganzen. Eine Menge ist eine Vielheit, als Einheit gedacht. Wenn diese oder ähnliche Sätze Definitionen sein wollten, so würde man mit Recht einwenden, daß sie idem per idem oder gar obscurum per obscurius definieren. Wir können sie aber als Demonstrationen gelten lassen, als Verweisungen auf einen primitiven, allen Menschen vertrauten Denkakt, der einer Auflösung in noch ursprünglichere Akte vielleicht weder fähig noch bedürftig ist. Wir wollen uns mit dieser Auffassung begnügen und es als Grundtatsache hinnehmen, daß ein Ding M in eigentümlicher, nicht definierbarer Weise gewisse andere Dinge $a, b, c, \ldots$ und diese wiederum jenes bestimmen; eine Beziehung, die wir mit den Worten ausdrücken: die Menge M besteht aus den Dingen $a, b, c, \ldots$

Eine Menge kann aus einer natürlichen Zahl von Dingen bestehen oder nicht; je nachdem heißt sie *endlich* oder *unendlich*. Beispiele sind [2]
einerseits die Menge der Einwohner einer Stadt, der Wasserstoffatome in der Sonne, der natürlichen Zahlen von 1 bis 1000, andererseits die Menge aller natürlichen Zahlen, aller Punkte einer Geraden, aller Kreise in einer Ebene. Es ist das unsterbliche Verdienst Georg Cantors (1845—1918), diesen Schritt in die Unendlichkeit gewagt zu haben, unter inneren wie äußeren Kämpfen gegen scheinbare Paradoxien, populäre Vorurteile, philosophische Machtsprüche (infinitum actu non datur), aber auch gegen Bedenken, die selbst von den größten Mathematikern ausgesprochen worden waren. Er ist dadurch der Schöpfer einer neuen Wissenschaft, der Mengenlehre geworden — denn die Betrachtung endlicher Mengen ist ja nichts weiter als elementare Arithmetik und Kombinatorik —, die heute das Fundament der gesamten Mathematik bildet. An diesem Triumph der Cantorschen Ideen ändert es nach unserer Ansicht nichts, daß noch eine bei allzu uferloser Freiheit der Mengenbildung auftretende Antinomie [3]
der vollständigen Aufklärung und Beseitigung bedarf.

Die fundamentale Beziehung eines Dinges a zu einer Menge A, der es angehört, bezeichnen wir mit G. Peano in Wort und Formel folgendermaßen:

$$a \text{ ist Element von } A\colon\ a \varepsilon A\,.$$

Das Gegenteil dieser Aussage lautet:

a ist nicht Element von A: $a \bar{\varepsilon} A$.

[4] Zwei Mengen werden dann und nur dann als gleich, in Formel

$$A = B,$$

definiert, wenn jedes Element der einen auch Element der andern ist (also wenn beide dieselben Elemente enthalten). Hiernach ist eine Menge durch ihre Elemente eindeutig bestimmt; wir bringen dies so zum Ausdruck, daß wir die Menge durch die zwischen geschwungene Klammern gesetzten Elemente bezeichnen, wobei nicht angegebene Elemente durch Punkte angedeutet werden. So ist

$$A = \{a\}, \quad A = \{a, b\}, \quad A = \{a, b, c\}$$

die Menge, die aus dem einen Element a, den beiden Elementen a, b, den drei Elementen a, b, c besteht;

$$A = \{a, b, c, \ldots\}$$

ist eine Menge, die aus den Elementen a, b, c und (möglicherweise) noch andern besteht. Welches diese andern, durch Punkte bezeichneten Elemente sein sollen, muß natürlich irgendwie angegeben werden, z. B.

die Menge der natürlichen Zahlen $\{1, 2, 3, \ldots\}$
die Menge der geraden natürlichen Zahlen $\{2, 4, 6, \ldots\}$
die Menge der Quadratzahlen $\{1, 4, 9, \ldots\}$
die Menge der Potenzen von 2 $\{1, 2, 4, 8, \ldots\}$
die Menge der Primzahlen $\{2, 3, 5, 7, \ldots\}$.

Eine Unterscheidung zwischen dem Ding a und der Menge $\{a\}$, die nur dies eine Element hat, ist begrifflich jedenfalls notwendig (wenn auch praktisch oft belanglos), schon weil wir auch Mengen (Systeme) zulassen werden, deren Elemente selbst wieder Mengen sind. Die Menge $a = \{1, 2\}$ besteht aus den zwei Elementen 1, 2, die Menge $\{a\}$ aus dem einen Element a.

Wir lassen aus Zweckmäßigkeitsgründen auch eine Menge 0, die *Nullmenge* oder *leere Menge*, zu, die kein Element enthält[1]). Nach der Definition der Gleichheit von Mengen gibt es nur eine Nullmenge. $A = 0$ bedeutet, daß die Menge A kein Element hat, leer ist, „verschwindet". Wollten wir die Nullmenge nicht als Menge zulassen, so würden wir in zahllosen Fällen, wo wir von einer Menge sprechen, zu dem Zusatz genötigt sein: falls diese Menge existiert. Denn die Definition der Elemente einer Menge sagt uns häufig noch gar nicht, ob solche Elemente vorhanden sind; z. B. ist bis jetzt nicht bekannt, ob die Menge der natürlichen Zahlen n, für welche die Gleichung $x^{n+2} + y^{n+2} = z^{n+2}$ in natürlichen Zahlen x, y, z lösbar ist,

---

[1]) Ob das Zeichen 0 die Nullmenge oder die Zahl Null bedeutet, geht aus dem jeweiligen Zusammenhang unzweideutig hervor.

leer ist oder nicht (d. h. ob der berühmte Fermatsche Satz richtig ist oder [5]
nicht). Die Aussage $A = 0$ kann also eine wirkliche Erkenntnis — natürlich in andern Fällen auch eine Trivialität — bedeuten; viele oder, wenn man vor Künsteleien nicht zurückschreckt, alle mathematischen Behauptungen lassen sich auf die Gestalt $A = 0$ bringen. Die Einführung der Nullmenge ist demnach, wie die der Zahl Null, durch Zweckmäßigkeitsgründe geboten; andererseits zwingt sie auch manchmal, das Nichtverschwinden einer Menge (wie das einer Zahl) unter den Voraussetzungen eines Satzes ausdrücklich namhaft zu machen.

Sind A und B zwei Mengen, so entsteht die Frage, ob die Elemente der einen vielleicht auch der andern angehören. Bedeuten a und b Elemente von A und B, so wollen wir zunächst die beiden Alternativen bilden:

$$\begin{array}{ll} \text{jedes } a \varepsilon B, & \text{nicht jedes } a \varepsilon B, \\ \text{jedes } b \varepsilon A, & \text{nicht jedes } b \varepsilon A. \end{array}$$

Durch deren Verbindung erhalten wir vier mögliche Fälle, von denen die drei ersten durch eine beigefügte Formel bezeichnet werden:

(1) Jedes $a \varepsilon B$, jedes $b \varepsilon A$: $A = B$
(2) Jedes $a \varepsilon B$, nicht jedes $b \varepsilon A$: $A < B$
(3) Nicht jedes $a \varepsilon B$, jedes $b \varepsilon A$: $A > B$
(4) Nicht jedes $a \varepsilon B$, nicht jedes $b \varepsilon A$.

Im Fall (1) sind in der Tat beide Mengen gleich, nach der früheren Erklärung. Im Fall (2) enthält A nur Elemente von B, aber nicht alle, wodurch sich A als die kleinere, B als die größere Menge charakterisiert;
dies wird durch die Bezeichnung $A < B$ zum Ausdruck gebracht, die an [6]
die Zahlenbeziehung $\alpha < \beta$ erinnern soll. Im Fall (3) steht es umgekehrt, sodaß $A > B$ so viel wie $B < A$ ist. Im allgemeinen wird keiner dieser Fälle, sondern der Fall (4) eintreten, zu dessen besonderer Bezeichnung kein Anlaß besteht.

Die Relation „kleiner als" ist *transitiv*, d. h. aus $A < B$, $B < C$ folgt $A < C$ (ebenso natürlich die Relationen „größer als" und „gleich").

Ist jedes $a \varepsilon B$, so daß einer der Fälle (1)(2) vorliegt, so bezeichnen wir dies zusammenfassend [1]) als

$$A \leqq B, \quad A \text{ ist } \textit{Teilmenge} \text{ von } B$$

(auch Teil oder Untermenge von B); wenn schärfer $A < B$ ist, wird A gelegentlich *echte Teilmenge* von B genannt. *Zu den Teilmengen von B wird also B selbst sowie auch die Nullmenge mitgerechnet*; für $A = 0$ ist ja die Beziehung: jedes $a \varepsilon B$, dadurch — in allerdings trivialer Weise — erfüllt,

[1]) Das Gleichheitszeichen wird von vielen für entbehrlich gehalten. Statt der runden Ungleichheitszeichen sind auch eckige verschiedener Formen im Gebrauch.

daß es gar kein a gibt [1]). Dabei bedeutet $0 < B$ oder $B > 0$, daß die Menge B nicht leer ist. Die Zweckmäßigkeit jener Verabredung tritt z. B. hervor, wenn wir die Teilmengen einer endlichen Menge zählen. Die Teilmengen von $\{1, 2, 3\}$ sind

$$0,\ \{1\},\ \{2\},\ \{3\},\ \{1, 2\},\ \{1, 3\},\ \{2, 3\},\ \{1, 2, 3\},$$

es gibt ihrer $8 = 2^3$. Eine Menge aus n Elementen besitzt $\binom{n}{m}$ Teilmengen, die aus m Elementen bestehen, wobei $\binom{n}{m} = \frac{n!}{m!\,(n-m)!}$ der Binomialkoeffizient und $\binom{n}{0} = \binom{n}{n} = 1$ zu setzen ist; die Anzahl der Teilmengen ist

$$\binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{n-1} + \binom{n}{n} = 2^n.$$

Schon dies einfache Ergebnis spricht für den oben erklärten Begriff der Teilmenge.

Ist $A \leqq B$, so bedeute

$$B - A$$

die Menge der Elemente von B, die nicht Elemente von A sind; diese *Differenz* wird auch das *Komplement von A in B* genannt. Offenbar ist

$$B - 0 = B, \quad B - B = 0,$$
$$B - (B - A) = A.$$

Es sei ausdrücklich hervorgehoben, weil andere Autoren anders verfahren, daß wir bei der Differenzbildung den Subtrahenden stets als Teilmenge des Minuenden annehmen. Z. B. setzt die Schreibweise $C - (B - A)$ zunächst $A \leqq B$, sodann $B - A \leqq C$ voraus. Beispiel: $A = \{5, 6, 7, \ldots\}$ sei die Menge der natürlichen Zahlen von 5 ab, $B = \{1, 2, 3, \ldots\}$ die Menge aller natürlichen Zahlen, $C = \{1, 2, 3, 4, 5, 6\}$ die Menge der ersten sechs natürlichen Zahlen. Hier ist $B - A = \{1, 2, 3, 4\}$, $C - (B - A) = \{5, 6\}$.

§ 2. Funktionen.

Der Funktionsbegriff ist fast ebenso fundamental und ursprünglich wie der Mengenbegriff. Eine funktionale Beziehung baut sich aus *Elementpaaren* ebenso auf wie eine Menge aus einzelnen *Elementen*.

Statt eines einzelnen Elementes betrachten wir eine Zusammenstellung von zwei Elementen *in einer bestimmten Reihenfolge* oder, wie wir es nennen,
[7] ein *geordnetes Paar* (a, b) von Elementen, worin a das erste, b das zweite

[1]) Ausführlicher: die Aussage „wenn $a \varepsilon A$, so ist auch $a \varepsilon B$“ ist richtig, weil die Voraussetzung $a \varepsilon A$ nie erfüllt ist. Sind p, q Urteile, so ist die Aussage „wenn p richtig ist, so ist auch q richtig“ (aus p folgt q) gewiß richtig, wenn p falsch ist. Aus einem falschen Urteil folgt jedes beliebige Urteil; wenn $2 \times 2 = 5$, so gibt es Hexen.

Element ist. Zwei solche geordneten Paare gelten dann und nur dann als gleich, wenn sie dasselbe erste und dasselbe zweite Element haben:

$$(a^*, b^*) = (a, b) \text{ soviel wie } a^* = a, \quad b^* = b.$$

Die Paare (a, b) und (b, a) sind hiernach verschieden, falls $a \neq b$; andererseits hindert nichts, auch aus zwei gleichen Elementen ein geordnetes Paar (a, a) zu bilden. Z. B. entstehen durch Kombination natürlicher Zahlen die geordneten Paare

$$(1, 1)\ (1, 2)\ (2, 1)\ (1, 3)\ (2, 2)\ (3, 1) \ldots;$$

solcher Art sind die Doppelindizes an den Elementen einer Matrix oder Determinante. Durch Kombination reeller Zahlen entstehen die geordneten Zahlenpaare (x, y); durch solche Paare Cartesischer Koordinaten, bei denen die Abszisse x und die Ordinate y nicht vertauscht werden dürfen, lassen sich die Punkte der Ebene darstellen.

Das geordnete Paar (a, b) ist eine andere Begriffsbildung als die Menge $\{a, b\}$; bei dieser sind a und b als verschieden angenommen, und es kommt auf die Reihenfolge nicht an.

Die geordneten Paare ermöglichen die Einführung des Funktionsbegriffes, wie sie uns auch zur Multiplikation (§ 4) und zur Ordnung (§ 9) von Mengen dienen werden. Sei P eine Menge geordneter Paare $p = (a, b)$; für jedes in P vorkommende Paar p $(p \varepsilon P)$ wollen wir b *ein Bild* von a, a *ein Urbild* von b nennen, und es sei A die Menge aller Urbilder a (d. h. aller ersten Elemente von Paaren $p \varepsilon P$), B die Menge aller Bilder b (d. h. aller zweiten Elemente von Paaren $p \varepsilon P$). Hiernach bestimmt jedes a seine Bilder, jedes b seine Urbilder; das ist der Zusammenhang, der zwischen den Mengen A und B durch die Paarmenge P hergestellt wird: es findet, wie man sagt, eine *Abbildung* der einen Menge auf die andere statt.

In dem besonderen Falle, wo jedes a nur ein einziges Bild b besitzt, bezeichnen wir dieses durch a bestimmte, von a abhängige Element b mit

$$b = f(a)$$

und sagen, daß dies eine *in der Menge A definierte eindeutige Funktion von a* sei. Z. B. definiert die Menge der Paare $(1, 2)$ $(2, 1)$ $(3, 2)$ eine Abbildung der Menge $A = \{1, 2, 3\}$ auf die Menge $B = \{1, 2\}$, und zwar eine in der Menge A eindeutige Funktion, nämlich

$$f(1) = 2, \quad f(2) = 1, \quad f(3) = 2.$$

Wenn überdies auch jedes b nur ein einziges Urbild a hat, so bezeichnen wir dieses durch b bestimmte Element a mit

$$a = g(b)$$

und haben damit eine in B definierte eindeutige Funktion von b. Jede dieser Funktionen heißt zur andern *invers* oder die *Umkehrung* der andern; beide werden als *eindeutig umkehrbar* oder *eineindeutig* bezeichnet; man nennt die zwischen A und B bestehende Abbildung *eineindeutig* oder *schlicht* und sagt, daß zwei Mengen A, B, die in einer solchen Beziehung stehen können, *äquivalent* seien, in Zeichen

$$A \sim B, \quad B \sim A.$$

Dieser fundamentale Begriff der Äquivalenz wird die Grundlage des zweiten Kapitels bilden; begnügen wir uns hier mit dem Beispiel der Äquivalenz zwischen der Menge A der natürlichen und der Menge B der geraden (positiven) Zahlen. Die Menge der geordneten Paare

$$(1, 2) \ (2, 4) \ (3, 6) \ \ldots$$

stellt die eineindeutige Beziehung her, indem sie jeder natürlichen Zahl a das Bild $b = 2a$, jeder geraden Zahl b das Urbild $a = \frac{1}{2} b$ zuordnet.

Wenn das Element a mehrere Bilder hat, so wird die Bezeichnung $b = f(a)$ in dem Sinne beibehalten werden können, daß $f(a)$ nicht ein einziges, sondern mehrere (vielleicht unendlich viele) Elemente b bedeutet; wir haben dann eine *mehrdeutige* Funktion $f(a)$. Das Entsprechende gilt von $g(b)$; beide, im allgemeinen mehrdeutige Funktionen heißen auch jetzt noch zueinander invers. Ein häufig zu betrachtender Fall ist der, daß zwar $f(a)$ eindeutig, die Umkehrungsfunktion $g(b)$ aber mehrdeutig ist. Z. B. definiert die Menge der geordneten Paare $(a, \sin a)$, wenn a alle reellen Zahlen durchläuft, eine Abbildung der Menge A aller reellen Zahlen auf die Menge B der Zahlen $-1 \leqq b \leqq 1$ von der Art, daß zwar $b = \sin a$ eindeutig, $a = \arcsin b$ aber in der bekannten Weise mehrdeutig ist, nämlich nicht nur eine Zahl a_0 mit $b = \sin a_0$, sondern zugleich alle Zahlen $2k\pi + a_0$ und $(2k + 1)\pi - a_0$ (k ganzzahlig) bedeutet. Beschränkt man statt dessen a auf das Intervall $-\frac{\pi}{2} \leqq a \leqq \frac{\pi}{2}$, so wird die Beziehung zwischen beiden Mengen eindeutig umkehrbar.

Die hier gegebene Definition des Funktionsbegriffs wird dem Anfänger, der noch im Bann der elementaren, allenfalls der stetigen Funktionen steht, etwas abstrakt erscheinen; sie ist aber notwendig, um diesem fundamentalen Begriff seine Freiheit und Allgemeinheit zu lassen. Es ist, um bei der eindeutigen Funktion $f(a)$ zu bleiben, nur wesentlich, daß $f(a)$ nach irgendeiner Vorschrift (die hier durch die Paarmenge P gegeben war) ein durch a wohlbestimmtes Element ist, unwesentlich, ob sich diese Vorschrift durch „analytische Ausdrücke" oder sonstwie fixieren läßt, unwesentlich auch, ob unsere Kenntnisse und Hilfsmittel uns gestatten, auch nur für ein einziges a die tatsächliche Bestimmung von $f(a)$ durchzuführen. Dasselbe,
[8] was hier über den allgemeinen, von Dirichlet formulierten Funktionsbegriff gesagt wurde, wäre auch schon über den Cantorschen Mengenbegriff zu sagen gewesen. Die Menge der rationalen Zahlen ist wohl-

definiert, obwohl wir nicht wissen, ob π^{π} zu ihr gehört oder nicht, und die Funktion $f(a)$, die für rationales a gleich 1, für irrationales gleich 0 sein soll, ist wohldefiniert, obwohl wir den Wert von $f(\pi^{\pi})$ nicht kennen.

§ 3. Summe und Durchschnitt.

Sind A, B zwei Mengen, so verstehen wir unter ihrer *Summe*

$$S = A \dot{+} B$$

die Menge der Elemente, die zu A *oder* zu B (oder auch zu beiden) gehören, unter ihrem *Durchschnitt*

$$D = AB$$

die Menge der Elemente, die zugleich zu A *und* zu B gehören. Wenn $D = 0$, die beiden Mengen A, B also kein gemeinsames Element haben, so heißen sie zueinander *fremd* oder *disjunkt*; nur in diesem Falle bezeichnen wir die Summe auch mit

$$S = A + B$$

und bemerken, daß dann offenbar $S - A = B$, $S - B = A$ ist[1]).

Beispiel. A sei das Intervall[2]) $[1, 3]$, d. h. die Menge der reellen Zahlen x mit $1 \leqq x \leqq 3$, B ebenso das Intervall $[2, 4]$. Dann ist S das Intervall $[1, 4]$, D das Intervall $[2, 3]$.

Sind A, B endliche Mengen und zueinander fremd, A aus m, B aus n Elementen bestehend, so besteht $A + B$ aus $m + n$ Elementen.

Es ist

$$S - A = B - D, \quad S - B = A - D,$$

ersteres nämlich die Menge der Elemente, die nur zu B, nicht zu A gehören. Also

$$D = B - (S - A) = A - (S - B),$$

der Durchschnitt läßt sich mit Hilfe von Summe und Differenz bilden.

Die Summen- und Durchschnittsbildung läßt sich ohne weiteres auf beliebig, endlich oder unendlich viele Mengen ausdehnen. Als abkürzendes Summenzeichen verwenden wir das deutsche $\mathfrak{S}$, bei disjunkten Summanden auch das griechische Σ, als abkürzendes Durchschnittszeichen das deutsche $\mathfrak{D}$. Denken wir uns zunächst einmal den natürlichen Zahlen $1, 2, \ldots, k$ oder auch allen natürlichen Zahlen $1, 2, \ldots$ Mengen $A_1, A_2, \ldots$ zugeordnet (wobei schon der Begriff der eindeutigen Funktion aus § 2 zur Anwendung kommt), die übrigens durchaus nicht paarweise verschieden zu sein brauchen, so ist ihre Summe

[1]) Manche nennen die Summe auch Vereinigung oder Vereinigungsmenge; andererseits wird vielfach die Addition auch nicht fremder Summanden mit dem einfachen Pluszeichen (ohne Punkt) geschrieben. Das Zeichen $\dot{+}$ wurde von C. Carathéodory eingeführt.

[2]) Wegen der Bezeichnung der Intervalle vgl. die Vorbemerkungen.

$$S = A_1 \dotplus A_2 \dotplus \cdots = \mathfrak{S}_m A_m$$

die Menge der Elemente, die mindestens einem A_m angehören, ihr Durchschnitt

$$D = A_1 A_2 \cdots = \mathfrak{D}_m A_m$$

die Menge der Elemente, die allen A_m zugleich angehören. Nur wenn die Summanden *disjunkt*, d. h. *paarweise fremd* sind, also

$$A_m A_n = 0 \quad \text{für} \quad m \neq n,$$

schreiben wir die Summe auch

$$S = A_1 + A_2 + \cdots = \sum_m A_m.$$

Endlich der allgemeine Fall: den Elementen m einer Menge $M = \{m, n, p, \ldots\}$ seien Mengen A_m zugeordnet; dann ist ihre Summe

$$S = A_m \dotplus A_n \dotplus A_p \dotplus \cdots = \mathfrak{S}_m^M A_m$$

die Menge der Elemente, die mindestens einem A_m angehören, ihr Durchschnitt

$$D = A_m A_n A_p \cdots = \mathfrak{D}_m^M A_m$$

die Menge der Elemente, die allen A_m zugleich angehören; im Falle disjunkter (paarweise fremder) Summanden schreiben wir auch

$$S = A_m + A_n + A_p + \cdots = \sum_m^M A_m.$$

Beispiele: $M = \{0, 1, 2, \ldots\}$ sei die Menge der ganzen Zahlen $\geqq 0$ und A_m die Menge der natürlichen Zahlen, die genau durch 2^m teilbar sind, also

$$\begin{aligned} A_0 &= \{1, \ 3, \ 5, \ 7, \ldots\} \\ A_1 &= \{2, \ 6, 10, 14, \ldots\} \\ A_2 &= \{4, 12, 20, 28, \ldots\} \\ A_3 &= \{8, 24, 40, 56, \ldots\} \\ &\cdots\cdots\cdots \end{aligned}$$

dann ist $$S = A_0 + A_1 + A_2 + \cdots = \sum_m^M A_m$$

die Menge aller natürlichen Zahlen.

M sei die Menge der reellen Zahlen $m > 1$ und $A_m = [0, m)$ das Intervall der Zahlen $0 \leqq x < m$. Dann ist

$$D = \mathfrak{D}_m^M A_m = [0, 1]$$

das Intervall der Zahlen $0 \leqq x \leqq 1$.

Die beiden Prozesse, Summen- und Durchschnittsbildung, sind *kommutativ*, *assoziativ* und *jeder zum andern distributiv*, d. h. es ist, um nur die einfachsten Fälle anzugeben,

$$A \dotplus B = B \dotplus A, \qquad AB = BA,$$
$$(A \dotplus B) \dotplus C = A \dotplus (B \dotplus C) = A \dotplus B \dotplus C, \quad (AB)C = A(BC) = ABC,$$
$$(A \dotplus B)\,C = AC \dotplus BC,$$
$$AB \dotplus C = (A \dotplus C)(B \dotplus C).$$

Von den beiden letzten Formeln, den distributiven, ist die erste evident und auch die zweite leicht einzusehen; wir bemerken noch den nächst allgemeineren Fall

$$C \cdot \mathfrak{S}_m A_m = \mathfrak{S}_m C A_m,$$
$$C \dotplus \mathfrak{D}_m A_m = \mathfrak{D}_m (C \dotplus A_m).$$

Sind sämtliche A_m Teilmengen einer umfassenden Menge E und $B_m = E - A_m$ ihre Komplemente, so ist

$$E = \mathfrak{S}_m A_m + \mathfrak{D}_m B_m = \mathfrak{D}_m A_m + \mathfrak{S}_m B_m.$$

Denn jedes Element von E gehört entweder mindestens einem A_m an (also zu $\mathfrak{S} A_m$) oder keinem, im letzteren Falle gehört es allen B_m an (also zu $\mathfrak{D} B_m$). Wir können diese wichtige Formel kurz so aussprechen: *das Komplement der Summe ist der Durchschnitt der Komplemente, das Komplement des Durchschnitts ist die Summe der Komplemente.* Wenn also eine Menge P aus den Mengen A_m durch wiederholte Summen- und Durchschnittsbildung hervorgeht, so erhält man ihr Komplement $Q = E - P$, indem man die A_m durch ihre Komplemente B_m ersetzt und zugleich die Operationen $\mathfrak{S}$, $\mathfrak{D}$ vertauscht. So erhält man z. B. aus der einen Formel des distributiven Gesetzes

$$A \cdot \mathfrak{S} A_m = \mathfrak{S} A A_m$$

unmittelbar die andere

$$B \dotplus \mathfrak{D} B_m = \mathfrak{D}(B \dotplus B_m).$$

Wenn man diesen Prozeß, den Übergang zu den Komplementen, auf eine Mengen*ungleichung* anwendet, muß man außerdem die Zeichen $<$ und $>$ vertauschen (weil aus $P < P^*$ für die Komplemente $Q > Q^*$ folgt).

Es sei $A_1, A_2, A_3, \ldots$ eine Mengenfolge, d. h. jeder natürlichen Zahl n eine Menge A_n zugeordnet. Als ihren *oberen Limes*

$$\overline{A} = \overline{\operatorname{Lim}}\, A_n$$

bezeichnen wir die Menge der Elemente x, die *unendlich vielen* A_n angehören ($x \,\varepsilon\, A_n$ für unendlich viele n); als ihren *unteren Limes*

$$\underline{A} = \underline{\operatorname{Lim}}\, A_n$$

die Menge der Elemente x, die *fast allen* A_n angehören ($x \,\varepsilon\, A_n$ für fast alle n). Da die zweite Forderung schärfer ist als die erste, ist immer $\overline{A} \geqq \underline{A}$; gilt hier insbesondere das Gleichheitszeichen, so heißt die Menge $\overline{A} = \underline{A} = A$ der *Limes* der Mengenfolge

$$A = \mathrm{Lim}\, A_n,$$

und diese Folge wird *konvergent* genannt.

Beispiele. Für die Folge $M, N, M, N, \ldots$ ist $M \dotplus N$ der obere, MN der untere Limes; sie konvergiert nur für $M = N$. Eine aufsteigende Folge $A_1 \leqq A_2 \leqq \cdots$ konvergiert nach der Summe $\mathfrak{S} A_n$, eine absteigende $A_1 \geqq A_2 \geqq \cdots$ nach dem Durchschnitt $\mathfrak{D} A_n$, eine Folge disjunkter Mengen nach der Nullmenge. — Man bezeichne mit A_1 das Intervall $[0, 1]$, mit A_2, A_3 die beiden Intervalle $[0, \frac{1}{2}]$, $[\frac{1}{2}, 1]$, mit A_4, A_5, A_6 die drei Intervalle $[0, \frac{1}{3}]$, $[\frac{1}{3}, \frac{2}{3}]$, $[\frac{2}{3}, 1]$ und fahre so fort. Der obere Limes der Folge A_n ist das Intervall $[0, 1]$, der untere die Nullmenge.

Ist wieder E eine alle A_n umfassende Menge und $B_n = E - A_n$, so ist

$$E = \underline{A} + \overline{B} = \overline{A} + \underline{B}.$$

Denn ein $x \varepsilon E$ gehört *entweder* fast allen A_n, d. h. nur endlich vielen B_n, *oder* unendlich vielen B_n an.

Die Mengen $\overline{A}, \underline{A}$ entstehen aus den A_n durch wiederholte Summen- und Durchschnittsbildung, und zwar ist

$$\begin{aligned} \underline{A} &= D_1 \dotplus D_2 \dotplus D_3 \dotplus \cdots, & D_n &= A_n A_{n+1} A_{n+2} \cdots, \\ \overline{A} &= S_1 S_2 S_3 \ldots & , \quad S_n &= A_n \dotplus A_{n+1} \dotplus A_{n+2} \dotplus \cdots. \end{aligned}$$

Die erste Formel folgt unmittelbar aus der Definition, die zweite am einfachsten durch Komplementbildung.

Die Mengen $\overline{A}, \underline{A}$ bleiben ungeändert, wenn man in der Folge A_n endlich viele Mengen wegläßt, hinzufügt oder abändert. Ferner ist für eine Teilfolge A_p (wo p eine Folge natürlicher Zahlen durchläuft) offenbar

$$\underline{\mathrm{Lim}}\, A_n \leqq \underline{\mathrm{Lim}}\, A_p \leqq \overline{\mathrm{Lim}}\, A_p \leqq \overline{\mathrm{Lim}}\, A_n.$$

Wenn die ganze Folge nach A konvergiert, so auch jede Teilfolge.

Charakteristische Funktionen (C. de la Vallée Poussin). Jeder Teilmenge A einer festen Menge E läßt sich umkehrbar eindeutig eine in E definierte Funktion $f(x)$ zuordnen, die nur die Werte 0, 1 annimmt, nämlich

$$f(x) = 1 \quad \text{für} \quad x \varepsilon A, \qquad f(x) = 0 \quad \text{für} \quad x \varepsilon E - A.$$

Man nennt das die charakteristische Funktion zur Menge A; wir bezeichnen sie, indem wir nur die Abhängigkeit von A hervorheben und das Argument x weglassen, einfach mit $[A]$. Der ganzen Menge E entspricht die konstante Funktion $[E] = 1$, der Nullmenge 0 die konstante Funktion $[0] = 0$.

Den Verknüpfungen von Mengen entsprechen einfache Verknüpfungen der charakteristischen Funktionen. So ist

$$[B - A] = [B] - [A] \qquad (A \leqq B)$$

eine Gleichung, die, wie die folgenden, für jede Stelle x gilt; denn man hat für die drei möglichen Fälle

$$\begin{array}{rccc} x\,\varepsilon & E-B, & B-A, & A: \\ [A]= & 0\ , & 0\ , & 1 \\ [B]= & 0\ , & 1\ , & 1 \\ [B-A]= & 0\ , & 1\ , & 0. \end{array}$$

Insbesondere ist

$$[E-A] = 1 - [A].$$

Ferner ist

$$[AB] = [A][B]$$

und durch Komplementbildung

$$[A \dot{+} B] = 1 - (1-[A])(1-[B]),$$
$$[A \dot{+} B] + [AB] = [A] + [B],$$

insbesondere für disjunkte Summanden

$$[A + B] = [A] + [B].$$

Andererseits ist auch

$$[A \dot{+} B] = \max\,[[A], [B]], \qquad [AB] = \min\,[[A], [B]]$$

und allgemein für die Summe $S = \mathfrak{S} A_m$ und den Durchschnitt $D = \mathfrak{D} A_m$ beliebig vieler Mengen

$$[S] = \max\,[A_m], \qquad [D] = \min\,[A_m].$$

Für den oberen Limes $\overline{A} = \overline{\mathrm{Lim}}\, A_n$ und den unteren $\underline{A} = \underline{\mathrm{Lim}}\, A_n$ einer Mengenfolge ist

$$[\overline{A}] = \overline{\lim}\,[A_n], \qquad [\underline{A}] = \underline{\lim}\,[A_n];$$

denn $\overline{\lim}\,[A_n]$ ist dann und nur dann $= 1$ (sonst $= 0$), wenn *unendlich oft* $[A_n] = 1$, d. h. $x\,\varepsilon\,A_n$, $x\,\varepsilon\,\overline{A}$, und $\underline{\lim}\,[A_n]$ ist dann und nur dann $= 1$, wenn *schließlich* $x\,\varepsilon\,A_n$, $x\,\varepsilon\,\underline{A}$. Für den Limes $A = \mathrm{Lim}\, A_n$ einer konvergenten Mengenfolge ist $[A] = \lim\,[A_n]$.

§ 4. Produkt und Potenz.

Es bleibt uns noch übrig, das *Produkt* von Mengen zu definieren. Bilden wir aus zwei Mengen A, B die Menge P der *geordneten Paare* (§ 2) $p = (a, b)$, worin a alle Elemente von A, b alle Elemente von B durchläuft, also

$$p\,\varepsilon\,P \text{ soviel wie } a\,\varepsilon\,A,\ b\,\varepsilon\,B.$$

Falls die Mengen[1]) endlich sind und A aus m, B aus n Elementen besteht, besteht P aus mn Elementen, so daß diese Paarmenge tatsächlich den Charakter eines Produkts hat. Es sei bemerkt, daß diese *vollständige* Paar-

[1]) Diese müssen hier nicht disjunkt und können sogar identisch sein.

menge, die *alle* Paare mit $a \varepsilon A$, $b \varepsilon B$ enthält, alle jene Paarmengen, die nach § 2 eine funktionale Beziehung zwischen A und B vermitteln, zu Teilmengen hat; sie selbst liefert die „meistdeutige", die in stärkstem Maße mehrdeutige Abbildung, indem sie jedem a alle b als Bilder, jedem b alle a als Urbilder zuordnet. — Zur Bezeichnung des Produkts steht uns die einfache Schreibweise AB, die bereits an den Durchschnitt vergeben ist, nicht mehr zur Verfügung. Nun können wir ja aus den Mengen A, B selbst ein geord-
[9] netes Paar (A, B) bilden, und es steht uns frei, diesem Gebilde selbst wieder die Bedeutung einer Menge zu geben, nämlich das *geordnete Mengenpaar* als *Menge der geordneten Elementpaare* zu definieren. D. h. für geordnete Elementpaare und Mengenpaare wird die ε-Beziehung

$$(a, b)\, \varepsilon\, (A, B) \quad \text{soviel wie} \quad a \varepsilon A,\ b \varepsilon B$$

erklärt, während andere Dinge als Elementpaare überhaupt nicht als Elemente einer Menge (A, B) auftreten sollen. Dann ist also

$$P = (A, B)$$

das Produkt der beiden Mengen.

Beispiel. A und B seien Mengen reeller Zahlen. Wird das Zahlenpaar (a, b) durch den Punkt einer Ebene mit den rechtwinkligen Koordinaten a, b oder, kurz, durch den Punkt (a, b) dargestellt, so ziehe man durch alle Punkte $(a, 0)$ $(a \varepsilon A)$ der X-Achse Parallelen zur Y-Achse, durch alle Punkte $(0, b)$ $(b \varepsilon B)$ der Y-Achse Parallelen zur X-Achse; die Schnittpunkte (a, b) dieser Parallelen bilden das Produkt (A, B).

Das Produkt läßt sich als Summe äquivalenter Summanden auffassen (§ 6).

Die Ausdehnung des Produkts auf mehr als zwei Faktoren bietet keine Schwierigkeit. Analog wie die geordneten Paare (a, b) lassen sich *geordnete*
[10] *Tripel* (a, b, c) erklären, das sind Zusammenstellungen von drei Elementen in bestimmter Reihenfolge, also mit der Gleichheitsdefinition

$$(a^*, b^*, c^*) = (a, b, c) \quad \text{soviel wie} \quad a^* = a,\ b^* = b,\ c^* = c\,;$$

übrigens brauchen diese drei Elemente nicht verschieden zu sein. Die geordneten Tripel (a, b, c), worin a alle Elemente von A, b alle Elemente von B, c alle Elemente von C durchläuft, bilden per definitionem die Menge (A, B, C), das Produkt der drei Mengen A, B, C. Ebenso ist für jede endliche Faktorenzahl zu verfahren.

Das kommutative und assoziative Gesetz gelten bei den Produkten zwar nicht im Sinne der Gleichheit, wohl aber der Äquivalenz (§ 2). Die Elemente (a, b) und (b, a) der Mengen (A, B) und (B, A) sind ja nicht dieselben, stehen aber in umkehrbar eindeutiger Beziehung, so daß $(A, B) \sim (B, A)$. Auch die Elemente $((a, b), c)$ von $((A, B), C)$, d. h. die geordneten Paare, deren erstes Element ein geordnetes Paar (a, b), deren zweites ein

Element c ist, sind mit den Tripeln (a, b, c), den Elementen von (A, B, C) nicht identisch, entsprechen ihnen aber eineindeutig, und es ist

$$((A, B), C) \sim (A, (B, C)) \sim (A, B, C).$$

Um Produkte mit beliebiger Faktorenmenge zu erklären, verallgemeinern wir zuerst den Begriff des geordneten Paares oder Tripels, indem wir jedem Element m einer Menge $M = \{m, n, q, \ldots\}$ irgendein Element a_m zuordnen oder, anders gesagt, eine eindeutige Funktion $f(m) = a_m$ in M definieren; die Zusammenstellung dieser Elemente liefert das, was man einen *Elementkomplex*

$$p = (a_m, a_n, a_q, \ldots)$$

oder auch (Cantor) eine *Belegung* der Elemente m mit den Elementen a_m nennt. Beide Namen bezeichnen keine neue Konstruktion, sondern sind eben Synonyme für jene in M definierte Funktion $f(m)$. Zwei Elementkomplexe gelten dann und nur dann als gleich, wenn sie jedem m dasselbe Element a_m (als Bild) zuordnen, d. h.

$$(a_m^*, a_n^*, a_q^*, \ldots) = (a_m, a_n, a_q, \ldots) \text{ so viel wie } a_m^* = a_m, a_n^* = a_n, a_q^* = a_q, \ldots,$$

oder zwei Funktionen gelten dann und nur dann als gleich, wenn sie für jedes m denselben „Wert" haben: $f^*(m) = f(m)$; diese Vorschrift tritt an Stelle jener bestimmten Reihenfolge bei den geordneten Paaren, Tripeln usw. Übrigens hängen die Elementkomplexe, was in der obigen Schreibweise unter Umständen nicht völlig zum Ausdruck kommt, von der zugrunde liegenden Menge M ab; ein dreigliedriger Elementkomplex, der den Elementen 1, 2, 3 die Elemente a, b, c als Bilder zuordnet, ist nicht dasselbe wie einer, der den Elementen 4, 5, 6 dieselben Elemente a, b, c zuordnet, also auch nicht dasselbe wie ein geordnetes Tripel (a, b, c), das dieselben Elemente drei bestimmten Plätzen eines geschriebenen oder gedruckten Schemas zuordnet. Wohl aber lassen sich — und nur darauf kommt es schließlich an — die Tripel (a, b, c) und die Elementkomplexe, die einer beliebigen Menge aus drei Elementen die Bilder a, b, c zuordnen, in eineindeutige Beziehung bringen.

Ordnen wir jetzt jedem $m \varepsilon M$ eine Menge A_m zu, so entsteht ein *Mengenkomplex*

$$P = (A_m, A_n, A_q, \ldots)$$

und diesen erklären wir zugleich als Komplexmenge, nämlich als *Menge der Elementkomplexe* $p = (a_m, a_n, a_q, \ldots)$, in denen a_m alle Elemente von A_m, a_n alle von A_n, a_q alle von $A_q, \ldots$ durchläuft. P wird als *Produkt* (auch als Verbindungsmenge) der Mengen $A_m, A_n, A_q, \ldots$ definiert. Mit Verwendung des griechischen Π als abkürzenden Produktzeichens schreiben wir auch

$$P = \prod_m^M A_m.$$

Die Zuordnung der A_m zu den m ist so viel wie Definition einer eindeutigen Funktion $F(m) = A_m$ in M, einer Funktion, deren „Werte" nicht Elemente, sondern Mengen sind. Und das Produkt P ist die Menge aller eindeutigen Funktionen $f(m)$, für die

$$f(m) \varepsilon F(m) \quad \text{für jedes} \quad m \varepsilon M.$$

Lassen wir endlich alle $A_m = A$ zusammenfallen, so definieren wir mit

$$P = A^M$$

die *Potenz* (mit der Basis A und dem Exponenten M) als Produkt von lauter gleichen Faktoren. Dies ist also die Menge aller Elementkomplexe $p = (a_m, a_n, \ldots)$, deren Elemente zu A gehören ($a_m \varepsilon A$, $a_n \varepsilon A, \ldots$), oder die Menge aller eindeutigen Funktionen $a = f(m)$, die jedem $m \varepsilon M$ als Bild ein Element $a \varepsilon A$ zuordnen.

Ein wichtiges Beispiel: $A = \{a, b\}$ bestehe aus zwei Elementen; A^M ist die Menge der Funktionen $f(m)$, für die

$$f(m) = a \quad \text{oder} \quad b.$$

Jede solche Funktion $f(m)$ bestimmt [1]) die Menge M_a derjenigen m, für die $f(m) = a$, sowie die Menge M_b derjenigen m, für die $f(m) = b$, wobei

$$M = M_a + M_b$$

eine Spaltung von M in zwei komplementäre Teilmengen ist. Ist umgekehrt M_a eine beliebige Teilmenge von M, $M_b = M - M_a$ ihr Komplement in M, und definieren wir

$$f(m) = a \quad \text{für} \quad m \varepsilon M_a, \qquad f(m) = b \quad \text{für} \quad m \varepsilon M_b,$$

so haben wir eine unserer Funktionen $f(m)$. Die Funktionen $f(m)$ und die Mengen $M_a \leqq M$ stehen also in umkehrbar eindeutiger Beziehung, d. h. *A^M ist mit der Menge aller Teilmengen von M äquivalent.*

Man sieht ebenso, daß für eine Menge $A = \{a, b, c\}$ mit drei Elementen A^M äquivalent ist mit der Menge aller Zerlegungen

$$M = M_a + M_b + M_c$$

in drei disjunkte Summanden, wobei auf die Reihenfolge zu achten, d. h. die Zerlegung $M = M_a^* + M_b^* + M_c^*$ mit der obigen dann und nur dann als gleich anzusehen ist, wenn $M_a^* = M_a$, $M_b^* = M_b$, $M_c^* = M_c$.

Was über die bisher eingeführten Grundbegriffe in diesem Kapitel gesagt wurde, ist nur als vorläufige Orientierung zu betrachten. Insbesondere den Aufbau von Mengen mittels Summen- und Durchschnittsbildung werden wir künftig (in Kap. V) gründlicher behandeln müssen. In den nächsten drei Kapiteln wird der Durchschnitt, in den späteren das Produkt gegenüber den andern Verknüpfungen eine Nebenrolle spielen.

[1]) Für $a = 1, b = 0$ ist $f(m)$ die charakteristische Funktion (S. 20) zu M_a.

Zweites Kapitel.

Kardinalzahlen.

§ 5. Mengenvergleichung.

Wir nannten (§ 2) zwei Mengen äquivalent, in Formel

$$A \sim B,$$

wenn sich eine umkehrbar eindeutige Zuordnung ihrer Elemente herstellen läßt, derart, daß jedem a ein einziges $b = f(a)$, jedem b ein einziges $a = g(b)$ entspricht. Offenbar gilt:

$$A \sim A;$$

$$\text{wenn } A \sim B, \quad \text{ist } B \sim A;$$

$$\text{wenn } A \sim B, \quad B \sim C, \quad \text{ist } A \sim C;$$

man sagt, daß die Äquivalenzbeziehung *reflexiv*, *symmetrisch* und *transitiv* sei.

Endliche Mengen[1]) sind ersichtlich dann und nur dann äquivalent, wenn sie dieselbe Anzahl von Elementen haben. Demgemäß sagen wir allgemein, daß äquivalente Mengen dieselbe *Kardinalzahl* oder *Mächtigkeit* [11]
haben. D. h. wir ordnen jeder Menge A ein Ding $\mathfrak{a}$ zu derart, daß äquivalenten Mengen und nur solchen dasselbe Ding entspricht:

$$\mathfrak{a} = \mathfrak{b} \quad \text{soviel wie } A \sim B.$$

Diese neuen Dinge nennen wir Kardinalzahlen oder Mächtigkeiten; wir sagen: A hat die Mächtigkeit $\mathfrak{a}$, $\mathfrak{a}$ ist die Mächtigkeit von A, wohl auch (indem wir $\mathfrak{a}$ als Zahlwort verwenden) A hat $\mathfrak{a}$ Elemente.

Diese formale Erklärung sagt, was die Kardinalzahlen sollen, nicht was sie sind. Prägnantere Bestimmungen sind versucht worden, aber sie befriedigen nicht und sind auch entbehrlich. Relationen zwischen Kardinalzahlen sind uns nur ein bequemer Ausdruck für Relationen zwischen Mengen: das „Wesen" der Kardinalzahl zu ergründen, müssen wir der Philosophie überlassen.

Einer endlichen Menge aus n Elementen ($n = 1, 2, 3, \ldots$) wird als Mächtigkeit die Zahl n zugeordnet, der Nullmenge die Zahl 0.

Die Mächtigkeit der Menge $\{1, 2, 3, \ldots\}$ der natürlichen Zahlen wird $\aleph_0$ (Alef-Null) genannt[2]). Mengen dieser Mächtigkeit, die sich also in Gestalt einer *Folge*

---

[1]) Eine vollständige Mengenlehre hat auch die Theorie der endlichen Mengen und der natürlichen Zahlen exakt zu begründen; wir wollen hier diese Präliminarien als erledigt ansehen.

[2]) $\aleph$ ist der erste Buchstabe des hebräischen Alphabets.

$$\{a_1, a_2, a_3, \ldots\} \qquad (a_m \neq a_n \text{ für } m \neq n)$$

setzen lassen, heißen *abzählbar*[1]).

Der Menge der reellen Zahlen oder der damit äquivalenten Menge der Punkte einer geraden Linie geben wir die Mächtigkeit $\aleph$ (Alef); sie heißt die *Mächtigkeit des Kontinuums.*

Die Mächtigkeit der abzählbaren Mengen wird vielfach mit $\mathfrak{a}$, die des Kontinuums mit $\mathfrak{c}$ bezeichnet.

Wir haben schon bei der ersten Einführung des Begriffs die Äquivalenz der Menge der natürlichen Zahlen mit der der geraden Zahlen erwähnt. Wie die folgende Zusammenstellung lehrt, in der die übereinanderstehenden Zahlen einander zuzuordnen sind:

1	2	3	$\ldots\; n$	$\ldots$	
2	4	6	$\ldots\; 2n$	$\ldots$	
1	3	5	$\ldots\; 2n-1$	$\ldots$	
1	4	9	$\ldots\; n^2$	$\ldots$	
10	100	1000	$\ldots\; 10^n$	$\ldots$	
2	3	5	$\ldots\; p_n$	$\ldots$	(p_n die $n$te Primzahl)

sind die Menge der natürlichen, der geraden, der ungeraden, der Quadratzahlen, der Potenzen von 10, der Primzahlen allesamt einander äquivalent, haben also dieselbe Mächtigkeit $\aleph_0$. Man kann die Liste weiterführen und Zahlenfolgen immer rapideren Wachstums bilden (man denke etwa an 10, 10^{10}, $10^{10^{10}}$, ...), die dennoch, wie dünn sie auch in der Menge aller natürlichen Zahlen gesät sein mögen, keine geringere Mächtigkeit als die Gesamtmenge haben. Diese Verletzung des Axioms „totum parte majus" ist eine jener „Paradoxien des Unendlichen", an die man sich gewöhnen muß und gewöhnt hat; zwischen den Gesetzen endlicher und unendlicher Mengen gibt es naturgemäß Abweichungen, die selbstverständlich gar keinen Einwand gegen die unendlichen Mengen begründen.

Die Zahlen- oder Punktintervalle $0 \leqq x \leqq 1$, $0 \leqq y \leqq 1000$ haben als Zahlen- oder Punktmengen dieselbe Mächtigkeit, obwohl das zweite tausendmal so lang ist wie das erste; die eineindeutige Zuordnung wird durch $y = 1000\,x$ oder irgendeine Projektion vermittelt. Das Intervall $-\frac{\pi}{2} < x < \frac{\pi}{2}$ und die Menge aller reellen Zahlen y haben dieselbe Mächtigkeit vermöge der Zuordnung $y = \operatorname{tg} x$.

Daß dieses Verhalten — die Äquivalenz mit einer echten Teilmenge — für die unendlichen Mengen charakteristisch ist, werden wir sofort sehen.

[1]) Ihre Elemente können mit Hilfe aller natürlichen Zahlen „abgezählt" (numeriert) werden. Mengen, die endlich ($\geqq 0$) oder abzählbar sind, nennen wir *höchstens abzählbar*; unendliche, nicht abzählbare Mengen heißen *unabzählbar*.

I. *Jede unendliche Menge enthält eine abzählbare Teilmenge.*

Es sei a_1 ein Element der unendlichen Menge A, a_2 ein Element der (immer noch) unendlichen Menge $A - \{a_1\}$, a_3 ein Element der (immer noch) unendlichen Menge $A - \{a_1, a_2\}$ usw. Die paarweise verschiedenen Elemente $a_1, a_2, a_3, \ldots$ bilden eine abzählbare Teilmenge von A.

II. *Jede unendliche Menge ist mit einer echten Teilmenge äquivalent.* [12]

Indem man aus der unendlichen Menge A eine abzählbare Teilmenge herauszieht und deren Komplement in A mit B bezeichnet, erhält man

$$A = \{a_1, a_2, \ldots, a_n \quad , \ldots\} + B;$$

dies ist z. B. mit

$$\{a_2, a_3, \ldots, a_{n+1}, \ldots\} + B$$

äquivalent (man ordne jedem a_n als Bild a_{n+1}, jedem $b \varepsilon B$ als Bild b selbst zu).

Umgekehrt kann eine mit einer echten Teilmenge äquivalente Menge nicht endlich sein. Diese Eigenschaft charakterisiert also die unendlichen Mengen (und ist von R. Dedekind zur Definition der unendlichen Mengen verwendet worden).

III. (Äquivalenzsatz von F. Bernstein.) *Zwei Mengen, deren jede mit einer Teilmenge der andern äquivalent ist, sind selbst äquivalent.* [13]

Es sei $A \sim B_1$, $B \sim A_1$, A_1 Teilmenge von A, und zwar echte, da sonst nichts zu beweisen ist, also $A_1 < A$, $B_1 < B$. Durch die eineindeutige Abbildung von B auf A_1 wird auch B_1 auf eine (echte) Teilmenge A_2 von A_1 abgebildet, also

$$A > A_1 > A_2, \quad A \sim B_1 \sim A_2.$$

Demnach ist der Satz auf folgenden zurückgeführt:

$$\text{wenn } A > A_1 > A_2, \quad A \sim A_2, \quad \text{so ist } A \sim A_1,$$

d. h. *wenn eine Menge zwischen zwei äquivalenten Mengen liegt, so ist sie mit diesen äquivalent.*

Bei der eineindeutigen Abbildung von A auf A_2 werde A_1 auf A_3, A_2 auf A_4, A_3 auf A_5 usw. abgebildet, wobei also

$$A > A_1 > A_2 > A_3 > A_4 > A_5 > \cdots.$$

Sei

$$D = A A_1 A_2 \ldots$$

der Durchschnitt der Mengen A_n, dann ist

$$\begin{aligned} A &= D + (A - A_1) + (A_1 - A_2) + (A_2 - A_3) + (A_3 - A_4) + \cdots \\ A_1 &= D + (A_1 - A_2) + (A_2 - A_3) + (A_3 - A_4) + \cdots \end{aligned}$$

Denn, um z. B. die erste Formel zu beweisen: ein Element $a \varepsilon A$ gehört entweder allen A_n an, also zu D, oder es gibt ein erstes A_n, dem es nicht angehört, während es noch zu A_{n-1} gehört, also $a \varepsilon A_{n-1} - A_n$ ($A_0 = A$ gesetzt). Da nun

$$A - A_1 \sim A_2 - A_3 \sim A_4 - A_5 \sim \cdots,$$

so ist auch

$$(A - A_1) + (A_2 - A_3) + \cdots \sim (A_2 - A_3) + (A_4 - A_5) + \cdots,$$

die links stehende Menge wird auf die rechts stehende abgebildet. Indem man die übrigen Elemente von A sich selbst als Bilder zuordnet, erhält man eine eineindeutige Abbildung von A auf A_1, also $A \sim A_1$.

Bis jetzt haben wir nur von Äquivalenz, also von Gleichheit zweier Kardinalzahlen gesprochen. Die nächste Frage wäre nun: wenn zwei Mengen nicht äquivalent, also ihre Kardinalzahlen ungleich sind, läßt sich dann in natürlicher Weise die eine Kardinalzahl als die größere, die andere als die kleinere definieren? Kurz gesagt: haben die Kardinalzahlen Größencharakter? Sind sie vergleichbar?

Die Antwort scheint vorläufig verneinend auszufallen. Bedeuten A und B zwei Mengen, A_1 und B_1 irgendwelche Teilmengen von ihnen, so bestehen vier Möglichkeiten:

(1) Es gibt ein $A_1 \sim B$ und ein $B_1 \sim A$.
(2) Es gibt kein $A_1 \sim B$, aber ein $B_1 \sim A$.
(3) Es gibt ein $A_1 \sim B$, aber kein $B_1 \sim A$.
(4) Es gibt kein $A_1 \sim B$ und kein $B_1 \sim A$.

Im Fall (1) ist nach dem Äquivalenzsatz $\mathfrak{a} = \mathfrak{b}$; im Fall (2) wird man naturgemäß $\mathfrak{a} < \mathfrak{b}$, im Fall (3) $\mathfrak{a} > \mathfrak{b}$ zu definieren haben. Im Fall (4) können wir offenbar, wegen seiner Symmetrie in bezug auf beide Mengen, weder $\mathfrak{a} < \mathfrak{b}$ noch $\mathfrak{a} > \mathfrak{b}$ definieren, da sonst beides zugleich gelten müßte; ebensowenig ist $\mathfrak{a} = \mathfrak{b}$ zulässig, was mit der bisherigen natürlichen Definition der Gleichheit in Widerspruch treten würde. Wir haben hier also eine vierte Relation, die wir $\mathfrak{a} \parallel \mathfrak{b}$ schreiben und als *Unvergleichbarkeit* von $\mathfrak{a}$ mit $\mathfrak{b}$ bezeichnen, während in den drei ersten Fällen $\mathfrak{a}$ und $\mathfrak{b}$ vergleichbar heißen sollen. Also.

$$\left.\begin{array}{l}(1)\ \mathfrak{a} = \mathfrak{b} \\ (2)\ \mathfrak{a} < \mathfrak{b} \\ (3)\ \mathfrak{a} > \mathfrak{b}\end{array}\right\} \mathfrak{a}, \mathfrak{b} \text{ vergleichbar}$$

$$(4)\ \mathfrak{a} \parallel \mathfrak{b} \quad \mathfrak{a}, \mathfrak{b} \text{ unvergleichbar.}$$

Die Vergleichbarkeit wäre also nur zu retten, wenn man zeigen könnte, daß der vierte Fall tatsächlich nicht eintreten kann. Bei zwei endlichen Mengen ist dies in der Tat der Fall; denn ist, mit Numerierung der Elemente, $A = \{a_1, a_2, \ldots, a_m\}$, $B = \{b_1, b_2, \ldots, b_n\}$ und bildet man behufs eineindeutiger Zuordnung die Paare (a_1, b_1), (a_2, b_2) usw., so kommt man zu Ende, sobald eine der Mengen verbraucht ist. (Auch eine endliche und eine unendliche Mächtigkeit sind stets vergleichbar, jene ist die kleinere.) Auf eine ähnliche, aber erst später exakt zu begründende Art (mittels Wohlordnung der Mengen, § 13) werden wir künftig zeigen, daß der Fall (4) niemals eintreten kann, zwei Kardinalzahlen also stets vergleichbar sind.

Ist A einer Teilmenge von B äquivalent (Fall (1) oder (2) tritt ein), so ist $\mathfrak{a} = \mathfrak{b}$ oder $\mathfrak{a} < \mathfrak{b}$, was wir in $\mathfrak{a} \leqq \mathfrak{b}$ zusammenfassen. Der Äquivalenzsatz schreibt sich dann in der einleuchtenden Form: wenn $\mathfrak{a} \leqq \mathfrak{b}$ und $\mathfrak{a} \geqq \mathfrak{b}$, so ist $\mathfrak{a} = \mathfrak{b}$.

Ist	$\mathfrak{a} = \mathfrak{b}, \quad \mathfrak{a} < \mathfrak{b}, \quad \mathfrak{a} > \mathfrak{b}, \quad \mathfrak{a} \parallel \mathfrak{b},$
so ist	$\mathfrak{b} = \mathfrak{a}, \quad \mathfrak{b} > \mathfrak{a}, \quad \mathfrak{b} < \mathfrak{a}, \quad \mathfrak{b} \parallel \mathfrak{a}.$

Ist $\mathfrak{a} = \mathfrak{b}$, $\mathfrak{b}\,\varrho\,\mathfrak{c}$, so ist $\mathfrak{a}\,\varrho\,\mathfrak{c}$, wenn ϱ eine der vier Relationen bedeutet.

Ist $\mathfrak{a} < \mathfrak{b}$, $\mathfrak{b} < \mathfrak{c}$, so ist $\mathfrak{a} < \mathfrak{c}$; die Relation $<$ ist transitiv.

Jede unendliche Mächtigkeit ist $\geqq \aleph_0$ (Satz I); $\aleph_0$ ist die kleinste unendliche Mächtigkeit.

Jede unendliche Teilmenge einer abzählbaren Menge (z. B. die Menge der Primzahlen) ist abzählbar, da ihre Mächtigkeit $\leqq \aleph_0$ und zugleich $\geqq \aleph_0$, also nach dem Äquivalenzsatz $= \aleph_0$ ist.

§ 6. Summe, Produkt, Potenz. [14]

Die Mengen A, B seien fremd. Sind beide endlich und besteht A aus m, B aus n Elementen, so besteht $A + B$ aus $m + n$ Elementen. Ferner: ist $A \sim A_1$, $B \sim B_1$ und auch A_1, B_1 fremd, so ist $A + B \sim A_1 + B_1$. Diese Bemerkungen berechtigen zu folgender Definition:

Die *Summe* $\mathfrak{a} + \mathfrak{b}$ zweier Kardinalzahlen ist die Mächtigkeit der Mengensumme $A + B$, wenn A, B irgend zwei fremde Mengen mit den Mächtigkeiten $\mathfrak{a}, \mathfrak{b}$ sind.

Und allgemein: sind den Elementen m einer Menge

$$M = \{m, n, p, \ldots\}$$

Mächtigkeiten $\mathfrak{a}_m$ zugeordnet, so ist

$$\sum_m^M \mathfrak{a}_m = \mathfrak{a}_m + \mathfrak{a}_n + \mathfrak{a}_p + \cdots$$

die Mächtigkeit der Mengensumme

$$\sum_m^M A_m = A_m + A_n + A_p + \cdots,$$

wenn die A_m disjunkte (paarweise fremde) Mengen von den Mächtigkeiten $\mathfrak{a}_m$ sind.

Beispiele. Die Menge der natürlichen Zahlen läßt sich in $\{1, 2, \ldots, n\} + \{n + 1, n + 2, \ldots\}$ zerlegen, wo der zweite Summand abzählbar ist. Das gibt

$$n + \aleph_0 = \aleph_0 + n = \aleph_0.$$

Die Menge der natürlichen Zahlen zerfällt in die der geraden und der ungeraden, beide abzählbar, also

$$\aleph_0 + \aleph_0 = \aleph_0;$$

oder sie zerfällt in die drei abzählbaren Mengen der Zahlen, die durch 3 dividiert die Reste 0, 1, 2 geben, also

$$\aleph_0 + \aleph_0 + \aleph_0 = \aleph_0;$$

dasselbe folgt aus dem assoziativen Gesetz

$$\aleph_0 + \aleph_0 + \aleph_0 = \aleph_0 + (\aleph_0 + \aleph_0) = \aleph_0 + \aleph_0 = \aleph_0.$$

Ebenso ist für jede endliche Anzahl von Summanden

$$\aleph_0 + \aleph_0 + \cdots + \aleph_0 = \aleph_0.$$

Man kann die Menge der natürlichen Zahlen aber auch in abzählbar viele abzählbare Mengen spalten, wie u. a. die Schemata

$$\begin{array}{llll} 1 & 3 & 5 & 7\ldots \\ 2 & 6 & 10 & 14\ldots \\ 4 & 12 & 20 & 28\ldots \\ 8 & 24 & 40 & 56\ldots \\ \ldots & \ldots & \ldots & \ldots \end{array} \qquad \begin{array}{llll} 1 & 2 & 4 & 7\ldots \\ 3 & 5 & 8 & \ldots. \\ 6 & 9 & \ldots & \ldots \\ 10 & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \end{array}$$

(dyadisches Schema) (Diagonalschema)

zeigen; im ersten stehen in der n-ten Zeile die genau durch 2^{n-1} teilbaren Zahlen; im zweiten sind die Zahlen der Reihe nach in Diagonalen (von rechts oben nach links unten) angeordnet. So erhält man die Gleichung

$$\aleph_0 + \aleph_0 + \aleph_0 + \cdots = \aleph_0$$

oder präziser

$$\mathfrak{a}_1 + \mathfrak{a}_2 + \mathfrak{a}_3 + \cdots = \aleph_0 \quad \text{für } \mathfrak{a}_1 = \mathfrak{a}_2 = \mathfrak{a}_3 = \cdots = \aleph_0.$$

Es ist aber auch

$$1 + 1 + 1 + \cdots = \aleph_0,$$

wie man durch Spaltung einer abzählbaren Menge in ihre einzelnen Elemente erkennt, und daher nach dem Äquivalenzsatz

$$\mathfrak{a}_1 + \mathfrak{a}_2 + \mathfrak{a}_3 + \cdots = \aleph_0 \quad \text{für } 1 \leqq \mathfrak{a}_n \leqq \aleph_0,$$

z. B.

$$2 + 2 + 2 + \cdots = \aleph_0,$$
$$1 + 2 + 3 + \cdots = \aleph_0.$$

Das Intervall $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ ist mit der Menge der reellen Zahlen äquivalent (S. 26) und durch lineare Transformation überträgt sich dies auf jedes Intervall (α, β). Danach ist jede Zahlenmenge, die ein Intervall enthält, nach dem Äquivalenzsatz von der Mächtigkeit $\aleph$ des Kontinuums. Die Aneinandersetzung von zwei Intervallen, etwa $[0, 1) + [1, 2) = [0, 2)$, gibt

$$\aleph + \aleph = \aleph;$$

danach ist (n endlich)

$$\aleph \leqq n + \aleph \leqq \aleph_0 + \aleph \leqq \aleph + \aleph = \aleph,$$

also

$$n + \aleph = \aleph_0 + \aleph = \aleph + \aleph = \aleph.$$

Die Vereinigung der abzählbar vielen Intervalle $[n-1, n)$ $(n = 1, 2, \ldots)$ zur Halbgeraden gibt

$$\aleph + \aleph + \aleph + \cdots = \aleph,$$

präziser $\mathfrak{a}_1 + \mathfrak{a}_2 + \mathfrak{a}_3 + \cdots = \aleph$ für $\mathfrak{a}_1 = \mathfrak{a}_2 = \mathfrak{a}_3 = \cdots = \aleph$.

Für jede unendliche Mächtigkeit $\mathfrak{a}$ ist $\mathfrak{a} + \aleph_0 = \mathfrak{a}$. Denn nach § 5, I läßt sich setzen $\mathfrak{a} = \mathfrak{b} + \aleph_0$, also

$$\mathfrak{a} + \aleph_0 = (\mathfrak{b} + \aleph_0) + \aleph_0 = \mathfrak{b} + (\aleph_0 + \aleph_0) = \mathfrak{b} + \aleph_0 = \mathfrak{a}.$$

Die Mächtigkeit einer unendlichen Menge wird durch Tilgung endlich vieler Elemente nicht verringert, d. h. wenn $\mathfrak{a} = \mathfrak{b} + n$, $\mathfrak{a}$ unendlich, n endlich, so ist $\mathfrak{a} = \mathfrak{b}$. In der Tat ist

$$\mathfrak{a} = \mathfrak{a} + \aleph_0 = \mathfrak{b} + (n + \aleph_0) = \mathfrak{b} + \aleph_0 = \mathfrak{b},$$

da auch $\mathfrak{b}$ unendlich ist.

Besteht A aus m, B aus n Elementen, so hat das Produkt (A, B) mn Elemente (§ 4). Ferner gilt allgemein: ist $A_1 \sim A$, $B_1 \sim B$, so ist $(A_1, B_1) \sim (A, B)$. Dies berechtigt zu der Definition:

Das *Produkt* $\mathfrak{ab}$ zweier Kardinalzahlen ist die Mächtigkeit des Mengenprodukts (A, B), wenn A, B irgend zwei Mengen mit den Mächtigkeiten $\mathfrak{a}, \mathfrak{b}$ sind.

(A, B brauchen hier nicht fremd zu sein.)

Und allgemein: sind den Elementen m einer Menge

$$M = \{m, n, p, \ldots\}$$

Mächtigkeiten $\mathfrak{a}_m$ zugeordnet, so ist

$$\prod_m^M \mathfrak{a}_m = \mathfrak{a}_m \mathfrak{a}_n \mathfrak{a}_p \ldots$$

die Mächtigkeit des Mengenprodukts

$$\prod_m^M A_m = (A_m, A_n, A_p, \ldots),$$

wenn die A_m irgendwelche Mengen von den Mächtigkeiten $\mathfrak{a}_m$ sind.

Endlich definieren wir:

Sind $\mathfrak{a}, \mathfrak{m}$ zwei Kardinalzahlen, so ist die *Potenz* $\mathfrak{a}^{\mathfrak{m}}$ die Mächtigkeit der Potenz A^M, wenn A, M irgend zwei Mengen mit den Mächtigkeiten $\mathfrak{a}, \mathfrak{m}$ sind.

Wird allen $m \varepsilon M$ dieselbe Kardinalzahl $\mathfrak{a}_m = \mathfrak{a}$ zugeordnet, so ist

$$\sum_m^M \mathfrak{a}_m = \mathfrak{am}, \quad \prod_m^M \mathfrak{a}_m = \mathfrak{a}^{\mathfrak{m}},$$

wobei $\mathfrak{m}$ die Mächtigkeit von M bedeutet; es führt also auch im Gebiet des Unendlichen Addition gleicher Summanden zur Multiplikation, Multiplikation gleicher Faktoren zur Potenzierung. Um die erste Formel zu beweisen, greifen wir aus der Menge (A, M) der Paare (a, m) mit $a \varepsilon A$, $m \varepsilon M$ die mit festem m heraus, sie bilden eine mit A äquivalente Menge A_m, und aus der Mengengleichung

$$(A, M) = \sum_m^M A_m$$

folgt die behauptete Zahlengleichung. Zum Beweis der zweiten Formel kann man in $\prod_m^M A_m$ direkt $A_m = A$ setzen und erhält A^M.

Für die definierten Prozesse gelten nun kommutative, assoziative und distributive Gesetze, mit deren vollständiger Ausführung wir den Leser nicht plagen wollen; begnügen wir uns, kurz den Beweis der *Potenzregeln* zu skizzieren.

Ist $M = M_1 + M_2$, so erhält man eine in M definierte Funktion $f(m)$, indem man eine in M_1 definierte $f(m_1)$ und eine in M_2 definierte $f(m_2)$ zu einem Eunktionenpaar verbindet. Das liefert die Äquivalenz

$$A^{M_1+M_2} \sim (A^{M_1}, A^{M_2})$$

und allgemein für eine beliebige Summe $\sum_n M_n$ disjunkter Summanden

$$A^{M_1+M_2+\cdots} \sim (A^{M_1}, A^{M_2}, \ldots).$$

Läßt man den Index n von M_n eine Menge N durchlaufen und nimmt dann alle M_n äquivalent einer und derselben Menge M, so folgt

$$A^{(M,N)} \sim (A^M)^N.$$

Andererseits sei $A = (A_1, A_2)$; man definiert das Paar (a_1, a_2) als Funktion von m, indem man zwei Funktionen $a_1 = f_1(m)$, $a_2 = f_2(m)$ zu einem Paare verbindet; daraus folgt

$$(A_1, A_2)^M \sim (A_1^M, A_2^M)$$

und für ein beliebiges Produkt

$$(A_1, A_2, \ldots)^M \sim (A_1^M, A_2^M, \ldots),$$

sowie abermals, wenn der Index n von A_n eine Menge N durchläuft und alle A_n gleich A angenommen werden,

$$(A^N)^M \sim (A^M)^N.$$

In Kardinalzahlen haben wir also folgende Potenzregeln [1]):

$$\mathfrak{a}^{\mathfrak{m}_1+\mathfrak{m}_2+\cdots} = \mathfrak{a}^{\mathfrak{m}_1}\mathfrak{a}^{\mathfrak{m}_2}\ldots$$
$$(\mathfrak{a}_1\,\mathfrak{a}_2\ldots)^{\mathfrak{m}} = \mathfrak{a}_1^{\mathfrak{m}}\,\mathfrak{a}_2^{\mathfrak{m}}\ldots$$
$$(\mathfrak{a}^{\mathfrak{m}})^{\mathfrak{n}} = \mathfrak{a}^{\mathfrak{m}\mathfrak{n}}$$

Beispiele für Produkt und Potenz:

$$2\aleph_0 = \aleph_0 + \aleph_0 = \aleph_0$$
$$n\aleph_0 = \aleph_0 + \cdots + \aleph_0 = \aleph_0$$

und aus der Summe über abzählbar viele Summanden $\aleph_0$

$$\aleph_0\aleph_0 = \aleph_0 + \aleph_0 + \cdots = \aleph_0,$$
$$\aleph_0^2 = \aleph_0\aleph_0 = \aleph_0,$$
$$\aleph_0^n = \aleph_0\aleph_0\ldots\aleph_0 = \aleph_0.$$

---

[1]) $\mathfrak{a}^0$ ist $= 1$, $1^{\mathfrak{m}} = 1$, $0^{\mathfrak{m}} = 0$ zu setzen; jedes Produkt verschwindet, falls einer der Faktoren verschwindet. 0^0 zu definieren ist zwecklos.

$\mathfrak{a}^{\aleph_0}$ ist die Mächtigkeit der Menge aller Funktionen $a = f(n)$ (n natürliche Zahl, $a \varepsilon A$) oder aller *Folgen*

$$(a_1, a_2, \ldots) \qquad (a_n \varepsilon A),$$

wo A eine Menge der Mächtigkeit $\mathfrak{a}$ bedeutet; z. B. $2^{\aleph_0}$ die der Menge aller *dyadischen*, aus den Ziffern 0, 1 gebildeten Folgen, $10^{\aleph_0}$ die der Menge aller *dekadischen*, aus den Ziffern 0, 1, ..., 9 gebildeten Folgen.

Die Menge aller Teilmengen einer Menge M von der Mächtigkeit $\mathfrak{m}$ hat die Mächtigkeit $2^{\mathfrak{m}}$.

Denn wir sahen (§ 4), daß diese Menge mit A^M äquivalent ist, wo A aus zwei Elementen besteht.

§ 7. Die Skala der Mächtigkeiten.

Bis jetzt wissen wir noch nicht, ob unendliche Mächtigkeiten verschieden sein können oder ob vielleicht das populäre Vorurteil berechtigt ist, das im Unendlichen die bloße indifferente, der Abstufung unfähige Negation des Endlichen erblickt.

I. *Es ist stets $2^{\mathfrak{m}} > \mathfrak{m}$, d. h. die Menge der Teilmengen von M hat größere Mächtigkeit als M selbst.*

Zunächst gibt es ein mit M äquivalentes System von Teilmengen, nämlich das der einelementigen $\{m\}$, $m \varepsilon M$. Also ist $2^{\mathfrak{m}} \geqq \mathfrak{m}$.

Andererseits denken wir uns ein mit M äquivalentes System von Teilmengen, indem wir jedem Element $m \varepsilon M$ umkehrbar eindeutig eine Teilmenge $M_m \leqq M$ entsprechen lassen, und wollen zeigen, daß es dann immer noch eine weitere, von allen M_m verschiedene Menge $N \leqq M$ gibt, womit die Gleichung $2^{\mathfrak{m}} = \mathfrak{m}$ ausgeschlossen ist. Das Element m kann der ihm zugeordneten Menge M_m angehören oder nicht:

$$\text{entweder } m \varepsilon M_m \text{ oder } m \bar{\varepsilon} M_m;$$

die von den Elementen der zweiten Art gebildete Menge N ist dann gewiß von allen M_m verschieden. Denn da

$$m \varepsilon N \text{ soviel wie } m \bar{\varepsilon} M_m$$

ist, so würde, wenn für irgendein m $N = M_m$ wäre, der Widerspruch $m \varepsilon N$, $m \bar{\varepsilon} N$ herauskommen. Oder anders ausgedrückt: gehört m zu M_m, so gehört es nicht zu N; gehört m nicht zu M_m, so gehört es zu N; M_m und N unterscheiden sich also jedenfalls, da m der einen, aber nicht der andern Menge angehört.

Mit I ist also die Existenz verschiedener und zwar unendlich vieler verschiedener unendlicher Kardinalzahlen gesichert; gehen wir von einem unendlichen $\mathfrak{m}$ (z. B. $\aleph_0$ oder $\aleph$) aus, so können wir bilden

$$\begin{aligned} \mathfrak{m}_1 &= 2^{\mathfrak{m}} > \mathfrak{m}, \\ \mathfrak{m}_2 &= 2^{\mathfrak{m}_1} > \mathfrak{m}_1, \\ \mathfrak{m}_3 &= 2^{\mathfrak{m}_2} > \mathfrak{m}_2, \\ &\cdots \end{aligned}$$

(übrigens gilt I auch für endliche $\mathfrak{m}$: $1 = 2^0 > 0$, $2 = 2^1 > 1, \ldots$).

Ferner ist dann $\mathfrak{m} + \mathfrak{m}_1 + \mathfrak{m}_2 + \cdots$ eine noch größere Kardinalzahl, und mit dieser kann der Prozeß fortgesetzt werden; allgemein gilt:

II. *Ist jedem $m \varepsilon M$ eine Kardinalzahl $\mathfrak{a}_m$ zugeordnet und befindet sich unter diesen keine größte, so ist die Summe*

$$\mathfrak{a} = \sum_m^M \mathfrak{a}_m$$

größer als jedes $\mathfrak{a}_m$.

Denn zunächst ist sicher $\mathfrak{a} \geqq \mathfrak{a}_m$ für jedes m, wobei aber die Gleichheit ausgeschlossen ist, da sonst $\mathfrak{a}_m$ die größte unter den gegebenen Kardinalzahlen wäre.

Beispielsweise würde, wenn es unvergleichbare Kardinalzahlen $\mathfrak{a}, \mathfrak{b}$ gibt, $\mathfrak{a} + \mathfrak{b}$ größer als jede von beiden sein.

Die Sätze I, II ermöglichen ein unbegrenztes Aufsteigen zu höheren Mächtigkeiten. Freilich bieten sie auch den Angriffspunkt für eine Antinomie (S. 11): da es zu jeder Menge von Kardinalzahlen eine noch größere gibt, so kann keine solche Menge *alle* Kardinalzahlen umfassen, und die
[15] „Menge aller Kardinalzahlen" ist undenkbar. Wir werden hier also vor die Tatsache gestellt, daß die Forderung, „alle" Dinge einer gewissen Art zu sammeln, nicht immer vollziehbar ist: wenn man sie alle zu haben glaubt, sind es doch nicht alle. Das Beunruhigende dieser Antinomie liegt nicht darin, daß sich ein Widerspruch ergibt, sondern daß man auf einen Widerspruch nicht gefaßt war: die Menge aller Kardinalzahlen scheint a priori so unverdächtig wie die Menge aller natürlichen Zahlen. Daraus entsteht nun die Unsicherheit, ob nicht auch andere, vielleicht alle unendlichen Mengen solche widerspruchbehafteten Scheinmengen, „Unmengen" sein mögen, und sodann die Aufgabe, diese Unsicherheit wieder zu beseitigen; die Mengenlehre ist auf neuer (axiomatischer) Grundlage so aufzubauen, daß Antinomien ausgeschlossen sind. Wir können auf die dahin zielenden, von E. Zermelo begonnenen und sicheren Erfolg versprechenden Untersuchungen in diesem Buche nicht eingehen und müssen unseren „naiven" Mengenbegriff festhalten.

[16] III. (Satz von J. König). *Sind jedem $m \varepsilon M$ zwei Kardinalzahlen $\mathfrak{a}_m$, $\mathfrak{b}_m$ zugeordnet und ist stets $\mathfrak{a}_m < \mathfrak{b}_m$, so ist*

$$\sum_m^M \mathfrak{a}_m < \prod_m^M \mathfrak{b}_m .$$

Dies ist eine Verallgemeinerung des Satzes I, der für $\mathfrak{a}_m = 1$, $\mathfrak{b}_m = 2$ daraus entsteht.

Beweis. Wir bezeichnen der Einfachheit wegen einige Elemente von M mit $1, 2, 3, \ldots$. Die A_m seien disjunkte Mengen von den Mächtigkeiten $\mathfrak{a}_m$, die B_m Mengen von den Mächtigkeiten $\mathfrak{b}_m$; wegen der Ersetzbarkeit der B_m durch äquivalente Mengen kann A_m als Teilmenge von B_m angenommen werden, also

$$A_m < B_m, \quad B_m = A_m + C_m, \quad C_m > 0.$$

Mit
$$A = \sum_m^M A_m = A_1 + A_2 + A_3 + \cdots + A_m + \cdots$$
$$B = \prod_m^M B_m = (B_1, B_2, B_3, \ldots, B_m, \ldots)$$

soll $\mathfrak{a} < \mathfrak{b}$ nachgewiesen werden. B ist die Menge der Komplexe

$$p = (b_1, b_2, b_3, \ldots, b_m, \ldots) \quad (b_m \varepsilon B_m).$$

Zunächst ist $\mathfrak{a} \leqq \mathfrak{b}$. Denn sei c_m ein festes Element aus $C_m = B_m - A_m$, so bilden die Komplexe

$$\begin{array}{ll} (a_1, c_2, c_3, \ldots, c_m, \ldots) & (a_1 \varepsilon A_1) \\ (c_1, a_2, c_3, \ldots, c_m, \ldots) & (a_2 \varepsilon A_2) \\ \quad\cdot\;\cdot\;\cdot & \\ (c_1, c_2, c_3, \ldots, a_m, \ldots) & (a_m \varepsilon A_m), \\ \quad\cdot\;\cdot\;\cdot & \end{array}$$

die nur je ein a_m (das die Menge A_m durchläuft) und sonst lauter Elemente c enthalten, Teilmengen von B, die paarweise fremd und mit $A_1, A_2, \ldots, A_m, \ldots$ äquivalent sind; also ist A mit einer Teilmenge von B äquivalent.

Andererseits sei

$$P = \sum_m^M P_m = P_1 + P_2 + P_3 + \cdots + P_m + \cdots$$

eine mit A äquivalente Teilmenge von B ($P_m \sim A_m$); wir zeigen, daß sie nicht mit der ganzen Menge B identisch sein kann, womit also die Gleichung $\mathfrak{a} = \mathfrak{b}$ ausgeschlossen ist und nur $\mathfrak{a} < \mathfrak{b}$ übrig bleibt.

Betrachten wir die zu P_m gehörigen Komplexe

$$p_m = (b_{m1}, b_{m2}, \ldots, b_{mm}, \ldots)$$

und insbesondere die zugehörigen Elemente b_{mm}; sie bilden in B_m eine Menge D_m von einer Mächtigkeit $\leqq \mathfrak{a}_m$ (denn es gibt ja nur $\mathfrak{a}_m$ Komplexe p_m, also, da zu verschiedenen p_m nicht einmal verschiedene b_{mm} gehören müssen, höchstens $\mathfrak{a}_m$ verschiedene b_{mm}). Demnach ist $D_m < B_m$ oder $B_m = D_m + E_m$, $E_m > 0$. Wählt man nun ein beliebiges Element e_m von E_m, für jedes $m \varepsilon M$, so ist der Komplex

$$p = (e_1, e_2, e_3, \ldots, e_m, \ldots)$$

von allen p_m verschieden ($e_m \neq b_{mm}$) und zwar für jedes m, er gehört also nicht zu P und es kann nicht $P = B$ gewesen sein. Damit ist der Königsche Satz bewiesen.

3*

Sei speziell $0 < \mathfrak{a}_1 < \mathfrak{a}_2 < \cdots$ eine Folge wachsender Mächtigkeiten, so ist

$$\mathfrak{a}_1 + \mathfrak{a}_2 + \mathfrak{a}_3 + \cdots < \mathfrak{a}_2 \mathfrak{a}_3 \mathfrak{a}_4 \ldots$$

oder wenn wir jetzt $\mathfrak{a} = \mathfrak{a}_1 + \mathfrak{a}_2 + \mathfrak{a}_3 + \cdots$, $\mathfrak{b} = \mathfrak{a}_1 \mathfrak{a}_2 \mathfrak{a}_3 \ldots$ setzen:

$$\begin{aligned} \mathfrak{a} = \mathfrak{a}_1 + \mathfrak{a}_2 + \mathfrak{a}_3 + \cdots &< 1 \cdot \mathfrak{a}_2 \mathfrak{a}_3 \cdots \leqq \mathfrak{a}_1 \mathfrak{a}_2 \mathfrak{a}_3 \cdots \\ &= \mathfrak{b} \leqq \mathfrak{a}\mathfrak{a}\mathfrak{a} \cdots = \mathfrak{a}^{\aleph_0}, \quad \mathfrak{a} < \mathfrak{b} \leqq \mathfrak{a}^{\aleph_0}. \end{aligned}$$

Es gibt also Mächtigkeiten $\mathfrak{a}$ mit $\mathfrak{a} < \mathfrak{a}^{\aleph_0}$, und zwar unendlich viele, da man ja mit einem beliebig großen $\mathfrak{a}_1$ anfangen kann. Es gibt aber ebenso unendlich viele Mächtigkeiten $\mathfrak{c}$ mit $\mathfrak{c} = \mathfrak{c}^{\aleph_0}$; es sind das eben die Mächtigkeiten der Form $\mathfrak{b}^{\aleph_0}$, denn $(\mathfrak{b}^{\aleph_0})^{\aleph_0} = \mathfrak{b}^{\aleph_0 \aleph_0} = \mathfrak{b}^{\aleph_0}$.

Unter den Voraussetzungen von III gilt natürlich auch

$$\sum_m^M \mathfrak{a}_m \leqq \sum_m^M \mathfrak{b}_m, \quad \prod_m^M \mathfrak{a}_m \leqq \prod_m^M \mathfrak{b}_m,$$

jedoch kann hier das Gleichheitszeichen auftreten. So ist z. B.

$$\begin{gathered} 1 + 2 + 3 + \cdots = 2 + 3 + 4 + \cdots = \aleph_0, \\ 1 \cdot 2 \cdot 3 \cdots = 2 \cdot 3 \cdot 4 \cdots = 2^{\aleph_0}; \end{gathered}$$

das Produkt ist nämlich $\geqq 2 \cdot 2 \cdot 2 \cdots = 2^{\aleph_0}$ und

$$\leqq \aleph_0 \aleph_0 \aleph_0 \cdots = \aleph_0^{\aleph_0} \leqq (2^{\aleph_0})^{\aleph_0} = 2^{\aleph_0}.$$

§ 8. Die elementaren Mächtigkeiten.

Als elementar wollen wir die folgenden drei Mächtigkeiten bezeichnen:

die Mächtigkeit $\aleph_0$ der abzählbaren Mengen;

die Mächtigkeit $\aleph$ des Kontinuums, die wir mit $2^{\aleph_0}$ als identisch erkennen werden;

die Mächtigkeit $2^{\aleph}$.

[17] *Die Mächtigkeit* $\aleph_0$ (die kleinste unendliche Mächtigkeit). Wir wissen schon, daß eine Summe von endlich oder abzählbar vielen Summanden $\aleph_0$ und ein Produkt von endlich vielen Faktoren $\aleph_0$ wieder gleich $\aleph_0$ ist. Nennen wir, außer den schon vorgekommenen, noch einige abzählbare Mengen.

(α) Die Menge aller ganzen Zahlen $\gtreqless 0$.

Die negativen, die Zahl 0 und die positiven ganzen Zahlen bilden zusammen eine Menge von der Mächtigkeit $\aleph_0 + 1 + \aleph_0 = \aleph_0$. Man bringt die ganzen Zahlen in eine einfache Folge $\{a_1, a_2, \ldots\}$ etwa durch die Anordnung

$$0, 1, -1, 2, -2, 3, -3, \ldots.$$

(β) Die Menge aller Paare natürlicher Zahlen.

Ihre Mächtigkeit ist $\aleph_0 \aleph_0 = \aleph_0$. Man kann die Paare natürlicher Zahlen den natürlichen Zahlen selbst eineindeutig zuordnen oder eine Doppelfolge in eine einfache verwandeln in der Weise, daß man die Zahlenpaare (p, q) zu einer Matrix (wo (p, q) in der p-ten Zeile, q-ten Spalte steht) zusammenstellt

$$\begin{array}{llll} (1,1) & (1,2) & (1,3) & \dots \\ (2,1) & (2,2) & (2,3) & \dots \\ (3,1) & (3,2) & (3,3) & \dots \\ . & . & . & \dots \\ . & . & . & \dots, \end{array}$$

andererseits die sämtlichen natürlichen Zahlen ebenfalls in eine solche Matrix anordnet (zwei Beispiele S. 30) und die auf gleichen Plätzen stehenden Elemente einander entsprechen läßt. Das Diagonalschema liefert folgende Korrespondenz:

$$\begin{array}{llllllll} n: & 1 & 2 & 3 & 4 & 5 & 6 & \dots \\ (p,q): & (1,1) & (1,2) & (2,1) & (1,3) & (2,2) & (3,1) & \dots, \end{array}$$

wobei die Zahlenpaare nach wachsender Summe $p+q$ und bei gleicher Summe nach wachsendem p geordnet sind. Das dyadische Schema liefert die Beziehung

$$n = 2^{p-1}(2q-1),$$

die nach p, q eindeutig „auflösbar“ ist.

(γ) Die Menge aller endlichen Komplexe natürlicher Zahlen (p_1), (p_1, p_2), $(p_1, p_2, p_3), \dots, (p_1, p_2, \dots, p_k), \dots$, wo also k und die p_k alle natürlichen Zahlen durchlaufen. Ihre Mächtigkeit ist

$$\aleph_0 + \aleph_0^2 + \aleph_0^3 + \cdots = \sum_k \aleph_0^k = \sum_k \aleph_0 = \aleph_0 \aleph_0 = \aleph_0.$$

Die Zuordnung zwischen diesen Komplexen und den natürlichen Zahlen n wird etwa durch die dyadische Entwicklung

$$n = 2^{p_1-1} + 2^{p_1+p_2-1} + \dots + 2^{p_1+p_2+\cdots+p_k-1}$$

vermittelt (z. B. wird der Zahl $27 = 1 + 2 + 2^3 + 2^4$ der Komplex $(1, 1, 2, 1)$ zugeordnet).

(δ) Die Menge der rationalen Zahlen.

Läßt man der positiven rationalen Zahl $\frac{p}{q}$, wo p und q teilerfremde natürliche Zahlen sind, das Zahlenpaar (p, q) entsprechen (Beispiel (β)), so wird die Menge der $\frac{p}{q}$ mit einem Teil der Menge der (p, q) äquivalent, hat also eine Mächtigkeit $\leqq \aleph_0$, als unendliche Menge aber die Mächtigkeit $\aleph_0$. Das Diagonalschema liefert eine Anordnung der $\frac{p}{q}$ nach wachsender Summe $p+q$, bei Gleichheit dieser nach wachsendem Zähler:

$$\tfrac{1}{1}, \tfrac{1}{2}, \tfrac{2}{1}, \tfrac{1}{3}, \tfrac{3}{1}, \tfrac{1}{4}, \tfrac{2}{3}, \tfrac{3}{2}, \tfrac{4}{1}, \tfrac{1}{5}, \tfrac{5}{1}, \dots,$$

wobei also die reduziblen Brüche wegzulassen sind.

Auch die Menge aller rationalen Zahlen ist abzählbar, ebenso die der rationalen Zahlen eines Intervalls. Die des Intervalls (0, 1) lassen sich z. B. auch so in eine einfache Folge bringen, daß man nach wachsenden Nennern und bei Gleichheit dieser nach wachsenden Zählern ordnet:

$\frac{1}{2}, \frac{1}{3}, \frac{2}{3}, \frac{1}{4}, \frac{3}{4}, \frac{1}{5}, \frac{2}{5}, \frac{3}{5}, \frac{4}{5}, \ldots$

Daß die Menge der rationalen Zahlen, die doch (geometrisch gesprochen) die gerade Linie überall dicht erfüllen, keine höhere Mächtigkeit hat als die der ganzen, gehört wie vieles andere zu den Tatsachen der Mengenlehre, die überraschend und sogar paradox wirken.

[18] (ε) Die Menge der (reellen) algebraischen Zahlen.

Auch diese Menge, wiederum viel feiner und dichter als die der rationalen Zahlen, ist nur abzählbar. Beweisen wir es gleich für die reellen und komplexen algebraischen Zahlen; eine algebraische Zahl k-ten Grades ist Wurzel einer (durch sie eindeutig bestimmten) irreduziblen Gleichung

$$x^k + r_1 x^{k-1} + \cdots + r_{k-1} x + r_k = 0$$

mit rationalen Koeffizienten. Solcher Gleichungen gibt es unendlich viele, aber gewiß nicht mehr, als es Komplexe $(r_1, r_2, \ldots, r_k)$ rationaler Zahlen gibt (da nicht alle Komplexe irreduzible Gleichungen liefern), also höchstens $\aleph_0^k = \aleph_0$. Also, es gibt $\aleph_0$ irreduzible Gleichungen k-ten Grades, jede hat k Wurzeln, es gibt $k\aleph_0 = \aleph_0$ algebraische Zahlen k-ten Grades und $\sum_k \aleph_0 = \aleph_0 \aleph_0 = \aleph_0$ algebraische Zahlen überhaupt.

Die Menge aller rationalen oder algebraischen Zahlenpaare, Zahlentripel usw. ist natürlich auch nur abzählbar. Die Menge aller Elemente, die mit endlich vielen Zeichen eines endlichen oder abzählbaren Zeichensystems dargestellt werden können, ist abzählbar, wenn man im Fall eines endlichen Zeichensystems beliebig lange Zeichenkomplexe zuläßt. Denn es ist

$$\begin{aligned} \aleph_0 + \aleph_0^2 + \cdots + \aleph_0^k &= \aleph_0, \\ \aleph_0 + \aleph_0^2 + \cdots &= \aleph_0, \\ n + n^2 + \cdots &= \aleph_0. \end{aligned}$$

Z. B. ist die Menge aller aus einem endlichen Alphabet herstellbaren, beliebig langen „Worte" (d. h. Buchstabenkomplexe, mit oder ohne Sinn) abzählbar, ebenso die Menge aller Bücher, Symphonien u. dgl.

Die Mächtigkeit $\aleph$ des Kontinuums. Wir fanden bereits (S. 30, 31), daß die Summe von endlich oder abzählbar vielen Summanden $\aleph$ gleich $\aleph$ ist:

$$n\aleph = \aleph_0 \aleph = \aleph.$$

Durch die Dezimalbrüche $0, a_1 a_2 a_3 \ldots$ wird jede Zahl des Intervalls $[0, 1]$ mindestens einmal und höchstens zweimal dargestellt. Daher

$$\aleph \leqq 10^{\aleph_0} \leqq 2\aleph = \aleph,$$

also $\aleph = 10^{\aleph_0}$. Dasselbe gilt natürlich, wenn man statt dekadischer Brüche solche mit einer andern Basis > 1 benutzt, also

$$\aleph = 2^{\aleph_0} = 3^{\aleph_0} = \cdots$$

(aber $1^{\aleph_0} = 1$). Hieraus folgt nun

$$\aleph^{\aleph_0} = (2^{\aleph_0})^{\aleph_0} = 2^{\aleph_0 \aleph_0} = 2^{\aleph_0} = \aleph,$$

d. h. $\aleph$ gehört jener bereits (S. 36) betrachteten Kategorie von Kardinalzahlen an, die ihrer $\aleph_0$-ten Potenz gleich sind, und ist sicher nicht Summe einer aufsteigenden Folge von Kardinalzahlen. Nach dem Äquivalenzsatz ist dann erst recht

$$\aleph = \aleph^2 = \aleph^3 = \cdots = \aleph^{\aleph_0}$$

(oder direkt: $\quad \aleph^2 = \aleph\aleph = 2^{\aleph_0} \cdot 2^{\aleph_0} = 2^{\aleph_0+\aleph_0} = 2^{\aleph_0} = \aleph$)

wie auch $\quad \aleph = 2^{\aleph_0} = 3^{\aleph_0} = \cdots = \aleph_0^{\aleph_0} = \aleph^{\aleph_0}.$

In diesen Formeln ist wieder eine Überraschung verborgen: $\aleph^2 = \aleph\aleph$ ist doch die Mächtigkeit der Menge aller reellen Zahlenpaare (x, y) oder aller Punkte der Ebene, $\aleph^3$ ist die Mächtigkeit des dreidimensionalen Raumes oder der Menge aller Zahlentripel (x, y, z) usw. bis zu $\aleph^{\aleph_0}$, welches die Mächtigkeit des $\aleph_0$-dimensionalen Raumes, d. h. der Menge aller reellen Zahlenfolgen $(x_1, x_2, x_3, \ldots)$ ist. Also: *alle Räume von endlich oder abzählbar vielen Dimensionen haben dieselbe Mächtigkeit* $\aleph$. Auch dies wirkt paradox genug und scheint den Dimensionsbegriff umzustürzen, den später erst eine ganz andersartige Betrachtung (die Forderung nicht nur eineindeutiger, sondern auch beiderseits stetiger Abbildung) wiederherstellen wird; und doch ist die Tatsache, daß Gerade und Ebene sozusagen gleichviel Punkte haben, nicht rätselhafter als die, daß man die natürlichen Zahlen in gerade und ungerade einteilen kann, d. h. $\aleph\aleph = \aleph$ im Grunde nichts anderes als $\aleph_0 + \aleph_0 = \aleph_0$. Denn aus einem Dezimalbruch kann man durch Spaltung zwei machen und aus zweien durch Ineinanderschlingen wieder einen: das ist das ganze Geheimnis. Um dabei noch einen Augenblick zu verweilen und dem Leser die Äquivalenz zwischen Linie und Fläche recht deutlich zu machen, stellen wir (mit einer Modifikation der systematischen Bruchentwicklungen, die deren Zweideutigkeit vermeidet) jede reelle Zahl des Intervalls $J = (0, 1]$ (d. h. also $0 < x \leq 1$) als dyadischen Bruch mit unendlich vielen Einsen, d. h. in der Form

$$x = (\tfrac{1}{2})^{x_1} + (\tfrac{1}{2})^{x_1+x_2} + (\tfrac{1}{2})^{x_1+x_2+x_3} + \cdots$$

dar, wo die x_n natürliche Zahlen sind; dies ist stets und nur auf eine Weise möglich und illustriert übrigens die Gleichung $\aleph = \aleph_0^{\aleph_0}$. Schreiben wir dafür abgekürzt

$$x = [x_1, x_2, x_3, \ldots].$$

Aus zwei solchen Zahlen x und $y = [y_1, y_2, y_3, \ldots]$ erhalten wir dann eine einzige

$$t = [x_1, y_1, x_2, y_2, \ldots]$$

und umgekehrt aus $\quad t = [t_1, t_2, t_3, t_4, \ldots]$

wiederum $\quad x = [t_1, t_3, t_5, \ldots], \quad y = [t_2, t_4, t_6, \ldots],$

womit die geordneten Zahlenpaare (x, y) und die Zahlen t in eindeutig umkehrbare Beziehung gebracht sind, also das Quadrat $0 < \begin{matrix} x \\ y \end{matrix} \leq 1$ auf die Strecke $0 < t \leq 1$ oder das Intervallprodukt (J, J) auf das eindimensionale Intervall J eineindeutig abgebildet ist. Ebenso ist mit Zahlentripeln (x, y, z) und endlichen Komplexen zu verfahren; für Zahlenfolgen $(x, y, z, \ldots)$ stelle man die Abbildung etwa mittels des Diagonalschemas her:

$$t = [x_1, x_2, y_1, x_3, y_2, z_1, \ldots]$$

und umgekehrt

$$\begin{aligned} x &= [t_1, t_2, t_4, \ldots] \\ y &= [t_3, t_5, t_8, \ldots] \\ z &= [t_6, t_9, t_{13}, \ldots] \\ &\ldots\ldots\ldots\ldots \end{aligned}$$

Aus $\aleph = 2^{\aleph_0}$ folgt noch $\aleph > \aleph_0$, *das Kontinuum ist nicht abzählbar.* Eine gleich zu Beginn seiner Forschungen von Cantor aufgestellte, aber bis heute noch unbewiesene Vermutung, die *Kontinuumhypothese*, besagt, daß $\aleph$ die *nächstgrößere* Mächtigkeit nach $\aleph_0$ sei; die Frage, ob dies stimmt
[19] oder nicht, ist das *Kontinuumproblem.* — Die Unabzählbarkeit des Kontinuums ergibt sich auch so: keine *abzählbare* Menge reeller Zahlen (des Intervalls J)

$$\begin{aligned} a &= [a_1, a_2, a_3, \ldots] \\ b &= [b_1, b_2, b_3, \ldots] \\ c &= [c_1, c_2, c_3, \ldots] \\ &\ldots \end{aligned}$$

kann *alle* diese Zahlen umfassen, da es noch weitere

$$x = [x_1, x_2, x_3, \ldots] \qquad (x_1 \neq a_1, x_2 \neq b_2, x_3 \neq c_3, \ldots)$$

gibt. Dieses „Diagonalverfahren" ist das einfachste Modell für den Beweis des Satzes § 7, I.

Da es nur $\aleph_0$ rationale und algebraische Zahlen gibt, so gibt es also sicher irrationale und transzendente Zahlen, und zwar sogar ebensoviele ($\aleph$) wie reelle Zahlen. Denn die Mächtigkeit einer *unabzählbaren* (d. h. unendlichen, nicht abzählbaren) Menge wird durch Tilgung abzählbar vieler Elemente nicht verringert (Beweis wie der der entsprechenden Bemerkung S. 31). Man kann die reellen Zahlen $0 < x \leqq 1$ auf die irrationalen $0 < y < 1$ z. B. vermöge

$$\begin{aligned} x &= [x_1, x_2, x_3, \ldots] \\ y &= \frac{1|}{|x_1} + \frac{1|}{|x_2} + \frac{1|}{|x_3} + \cdots \end{aligned}$$

abbilden, wo x durch die obige dyadische Darstellung, y durch einen Kettenbruch der Folge natürlicher Zahlen $x_1, x_2, \ldots$ zugeordnet ist.

Die Mächtigkeit $2^{\aleph}$.

Diese Mächtigkeit, die wieder $> \aleph$ ist, kommt der Menge aller Teilmengen des Kontinuums (Mengen reeller Zahlen) oder aller linearen, ebenen, räumlichen Punktmengen zu, ferner der Menge aller eindeutigen Funktionen $f(x)$ einer reellen Variablen, wo $f(x)$ zwei Werte annehmen kann; aus

$$\aleph^{\aleph} = (2^{\aleph_0})^{\aleph} = 2^{\aleph_0 \aleph} = 2^{\aleph}$$

folgt aber sofort, daß die Menge der Funktionen $f(x)$, die alle reellen Werte durchlaufen können, auch nur von der Mächtigkeit $2^{\aleph}$ ist. Die Mächtigkeit spezieller Funktionenklassen kann natürlich geringer sein, z. B. die Menge aller *stetigen* Funktionen $f(x)$ hat nur die Mächtigkeit $\aleph$. Denn $f(x)$ ist dann durch seine Werte $f(r)$ an den rationalen Stellen r bestimmt; die Menge aller $f(r)$ hat die Mächtigkeit $\aleph^{\aleph_0} = \aleph$ und die Menge aller stetigen $f(x)$ (da nicht jedes $f(r)$ ein stetiges $f(x)$ erzeugt) eine Mächtigkeit $\leqq \aleph$, aber auch $\geqq \aleph$ (die Funktionen $f(x)$ = constans bilden schon eine Menge der Mächtigkeit $\aleph$), also genau $= \aleph$.

Wir schließen mit einer Zusammenstellung von Formeln für $\mathfrak{a} + \mathfrak{b}$, $\mathfrak{a}\mathfrak{b}$, $\mathfrak{a}^{\mathfrak{b}}$ für Kombinationen aus $\aleph, \aleph_0$ und natürlichen Zahlen n; soweit sie noch nicht bewiesen sind, lassen sie sich aus dem Äquivalenzsatz ablesen.

$$
\begin{array}{ll}
(\alpha) & 1^{\aleph_0} = 1^{\aleph} = 1 \\
(\beta) & n + \aleph_0 = n\aleph_0 = \aleph_0^n = \aleph_0 + \aleph_0 = \aleph_0\aleph_0 = \aleph_0 \\
(\gamma) & (n+1)^{\aleph_0} = \aleph_0^{\aleph_0} = n + \aleph = n\aleph = \aleph^n = \\
 & = \aleph_0 + \aleph = \aleph_0\aleph = \aleph^{\aleph_0} = \aleph + \aleph = \aleph\aleph = \aleph \\
(\delta) & (n+1)^{\aleph} = \aleph_0^{\aleph} = \aleph^{\aleph} = 2^{\aleph}.
\end{array}
$$

Drittes Kapitel.

Ordnungstypen.

§ 9. Ordnung.

Viele Mengen stellen sich von vornherein in einer natürlichen *Ordnung* dar, in der von zwei verschiedenen Elementen das eine vorangeht, das andere nachfolgt. Wir bringen eine solche Ordnung durch die übliche Schreibweise zum Ausdruck, in der das vorangehende Element links vom nachfolgenden steht. So erscheinen die Buchstaben in der Ordnung des Alphabets, die natürlichen Zahlen der Größe nach in der Ordnung

$$1, 2, 3, \ldots;$$

ebenso die reellen Zahlen. Wir können aber solche Ordnungen auch willkürlich vorschreiben, z. B. die natürlichen Zahlen abnehmend ordnen

$$\ldots, 3, 2, 1$$

oder die ungeraden vor die geraden stellen, beide Klassen für sich wachsend oder abnehmend geordnet:

$$
\begin{array}{l}
1, 3, 5, \ldots, 2, 4, 6, \ldots \\
1, 3, 5, \ldots, \ldots, 6, 4, 2 \\
\ldots, 5, 3, 1, 2, 4, 6, \ldots \\
\ldots, 5, 3, 1, \ldots, 6, 4, 2.
\end{array}
$$

Wir können eine Menschenmenge nach Gewicht, Länge, Alter, nach der alphabetischen Reihenfolge der Namen oder nach ihren Garderobenummern im Theater ordnen.

[20] Eine Menge wird also geordnet, indem eine Vorschrift gegeben wird, nach der von zwei verschiedenen Elementen a, b das eine vorangeht, das andere nachfolgt. Soll a vor b, b nach a stehen, so schreiben wir[1])

$$a < b, \quad b > a.$$

Die Beziehung $<$ soll transitiv sein, d. h.

$$\text{wenn } a < b, \quad b < c, \quad \text{so ist } a < c.$$

Der räumliche oder zeitliche Charakter, der dieser Erklärung durch die Präpositionen vor und nach anzuhaften scheint, ist natürlich belanglos, und wir wollen uns durch andere Definitionen davon überzeugen, daß wir hier nichts anderes als eine Anwendung des Funktionsbegriffs vor uns haben.

Bilden wir die geordneten Paare $p = (a, b)$ aus *verschiedenen* Elementen der Menge A (die natürlich mindestens zwei Elemente haben soll); die Paare p und $p^* = (b, a)$ sind verschieden und mögen zueinander invers heißen. Die Menge aller geordneten Paare werde dann in zwei komplementäre Teilmengen $P + P^*$ gespalten mit der besonderen Vorschrift:

(α) Von zwei inversen Paaren gehört das eine zu P, das andere zu P^*.

(β) Wenn $p = (a, b)$ und $q = (b, c)$ zu P gehören, so auch $r = (a, c)$.

Schreibt man dann für $(a, b) = p \varepsilon P$

$$a < b \quad \text{oder} \quad b > a,$$

so ist eine Ordnung im obigen Sinne definiert (wie umgekehrt aus jeder Ordnung eine Spaltung $P + P^*$ entsteht, indem man diejenigen und nur diejenigen Paare $p = (a, b)$ zu P rechnet, für die $a < b$). Hier ist es wohl klar, daß das Vorangehen und Nachfolgen keine neuen Geheimnisse einführt, sondern daß es sich nur darum handelt, von zwei inversen Paaren das eine — oder von zwei verschiedenen Elementen das eine — gegenüber
[21] dem andern auszuzeichnen. P heiße die *ordnende Paarmenge.*

Noch einfacher ist folgende Erklärung: man ordne jedem Element a von A eindeutig umkehrbar eine Menge $M(a)$ zu, derart, daß diese Mengen (im Sinne der Disjunktion S. 13) immer vergleichbar sind, also, wegen der vorausgesetzten Eineindeutigkeit,

$$M(a) \lesseqgtr M(b) \quad \text{für} \quad a \neq b.$$

Schreibt man dann $a < b$ statt $M(a) < M(b)$, so ist eine Ordnung der Menge A definiert; umgekehrt läßt sich aus jeder Ordnung ein System von Mengen $M(a)$ der angegebenen Art herleiten, z. B.

---

[1]) Es scheint uns entbehrlich, statt $<>$ andere Formen dieser Zeichen zu wählen.

$M(a) =$ Menge der Elemente von A, die $< a$ sind.

Ob es Ordnungen von A oder Paarmengen P oder Systeme von Mengen $M(a)$ der vorgeschriebenen Art wirklich gibt, geht aus dieser Betrachtung, in der nur die Gleichwertigkeit der verschiedenen Erklärungen gezeigt ist, noch nicht hervor; später (§ 12) werden wir sehen, daß sich jede Menge nicht nur ordnen, sondern auf spezielle Art ordnen (wohlordnen) läßt. Für die Anzahl der verschiedenen Ordnungen, d. h. die Mächtigkeit der Menge der verschiedenen Ordnungen erhalten wir $2^{\mathfrak{a}\mathfrak{a}}$ als obere Grenze, wo $\mathfrak{a}$ die Mächtigkeit der zu ordnenden Menge A ist. Denn (A, A), von der Mächtigkeit $\mathfrak{a}\mathfrak{a} = \mathfrak{a}^2$, ist die Menge aller Paare (a, b) aus (gleichen oder verschiedenen) Elementen von A; hiervon ist jene ordnende Paarmenge P eine Teilmenge. Es gibt $2^{\mathfrak{a}\mathfrak{a}}$ Teilmengen von (A, A), also höchstens $2^{\mathfrak{a}\mathfrak{a}}$ ordnende Paarmengen oder verschiedene Ordnungen von A. Z. B. für eine endliche Menge aus n Elementen ist die Anzahl der verschiedenen Ordnungen

$$n! = 1 \cdot 2 \cdots n < 2^{nn}.$$

Die großen lateinischen Buchstaben sollen jetzt geordnete Mengen bedeuten, und eine Gleichung $A = B$ soll besagen, daß A und B nicht nur dieselben Elemente, sondern auch dieselbe Ordnung haben. Bei Angabe einzelner Elemente soll die Schreibweise von links nach rechts zugleich die Ordnung bezeichnen; es ist also

$$A = \{\ldots, a, \ldots, b, \ldots, c, \ldots\}$$

eine Menge, in der $a < b < c$, während die Punkte das Vorhandensein weiterer Elemente vor a, zwischen a und b, zwischen b und c, nach c andeuten.

$$A = \{a, \ldots, b, \ldots\}$$

bedeutet demgemäß eine Menge, in der kein Element vor a steht, a das *erste* Element ist;

$$A = \{\ldots, b, \ldots, c\}$$

eine solche, in der kein Element nach c steht, c das *letzte* Element ist;

$$A = \{\ldots, a, b, \ldots\}$$

eine solche, in der kein Element zwischen a und b steht, a und b *benachbarte* Elemente sind.

Aus A entsteht, indem man die Zeichen $<$ $>$ durchweg vertauscht, die *invers* geordnete Menge

$$A^* = \{\ldots, c, \ldots, b, \ldots, a, \ldots\},$$

von der P^* die ordnende Paarmenge ist.

Zwei geordnete Mengen heißen *ähnlich*, in Formel

$$A \simeq B,$$

wenn es eine umkehrbar eindeutige Beziehung $b = f(a)$, $a = g(b)$ zwischen ihnen gibt, bei der die Ordnung erhalten bleibt, d. h. mit $a < a_1$ zugleich $b < b_1$ ist.

Z. B. ist die Menge $\{1, 2, 3, 4, \ldots\}$ mit $\{2, 3, 4, \ldots\}$. aber nicht mit $\{2, 3, 4, \ldots, 1\}$ ähnlich.

Die Ähnlichkeit ist wie die Äquivalenz reflexiv, symmetrisch, transitiv.
[22] Von zwei ähnlichen Mengen sagt man, daß sie denselben *Ordnungstypus* haben; d. h. man weist jeder geordneten Menge A ein Zeichen α, ihren Ordnungstypus (kurz Typus), derart zu, daß ähnliche Mengen und nur solche denselben Ordnungstypus haben:

$$\alpha = \beta \quad \text{so viel wie} \quad A \simeq B.$$

Der Typus der zu A inversen Menge A^* wird α^* genannt.

Ähnlichkeit schließt Äquivalenz ein, aber im allgemeinen nicht umgekehrt; aus $A \simeq B$ folgt $A \sim B$, aus $\alpha = \beta$ folgt $\mathfrak{a} = \mathfrak{b}$. Wir dürfen daher sagen, daß ein Typus α eine bestimmte Mächtigkeit $\mathfrak{a}$ habe.

Eine endliche Menge aus n Elementen kann auf $n!$ Arten geordnet (permutiert) werden, aber die entstehenden geordneten Mengen sind immer mit der Menge $\{1, 2, \ldots, n\}$ ähnlich; wir nennen diesen Typus wieder n, da die Verwechselung mit der Kardinalzahl n ungefährlich ist. Eine Menge aus einem Element habe den Typus 1, die Nullmenge den Typus 0.

Die Menge $\{1, 2, 3, \ldots\}$ der natürlichen Zahlen in natürlicher Ordnung habe den Typus ω, die invers geordnete $\{\ldots, 3, 2, 1\}$ also den Typus ω^*.

§ 10. Summe und Produkt.

Es seien A, B disjunkte geordnete Mengen. Dann soll

$$S = A + B$$

die Summe beider Mengen in folgender Ordnung bedeuten: die Ordnung der Elemente a unter sich und die der Elemente b unter sich bleibt bestehen, und es wird jedes a vor jedes b gesetzt ($a < b$). Das ist also von $B + A$ zu unterscheiden, welche Menge zwar dieselben Elemente, aber in anderer Ordnung besitzt: *die Addition geordneter Mengen ist nicht kommutativ.*

Beispiel. $A = \{1, 3, 5, \ldots\}$, $B = \{2, 4, 6, \ldots\}$.

$$A + B = \{1, 3, 5, \ldots, \; 2, 4, 6, \ldots\}$$
$$B + A = \{2, 4, 6, \ldots, \; 1, 3, 5, \ldots\}.$$

Man kann auch sagen: in $S = A + B$ ist $A < B$, die ganze Menge A wird vor die ganze Menge B gesetzt. Allgemein möge, wenn A, B Teilmengen einer geordneten Menge sind, $A < B$ so viel wie

$$a < b \quad \text{für} \quad a \varepsilon A, \; b \varepsilon B$$

bedeuten; analog sind die Relationen $a < B$, $A < b$ zu verstehen.

Sind auch A_1, B_1 disjunkte geordnete Mengen und $A_1 \simeq A$, $B_1 \simeq B$, so ist $A_1 + B_1 \simeq A + B$. Das berechtigt zu der Definition:

Die Typen*summe* $\alpha + \beta$ ist der Typus der Menge $A + B$, wenn A, B irgendzwei disjunkte Mengen mit den Typen α, β sind.

Allgemein sei jedem Element m einer geordneten Menge

$$M = \{\ldots, m, \ldots, n, \ldots, p, \ldots\}$$

eine geordnete Menge A_m zugewiesen, und diese Mengen seien disjunkt (paarweise fremd). Als Summe

$$S = \sum_m^M A_m = \cdots + A_m + \cdots + A_n + \cdots + A_p + \cdots$$

definieren wir die aus den Elementen aller A_m gebildete Menge in folgender Ordnung: die Ordnung der Elemente $a_m \varepsilon A_m$ unter sich bleibt unverändert, für jedes m, während für $m < n$ die ganze Menge A_m vor die ganze Menge A_n gesetzt wird: $A_m < A_n$.

Ersetzt man die Summanden durch *ähnliche* (natürlich wieder disjunkte), so geht auch die Summe in eine ähnliche über, und dies berechtigt zu der Definition: sind den Elementen $m \varepsilon M$ Typen α_m zugeordnet, so wird unter der Typensumme [23]

$$\sigma = \sum_m^M \alpha_m = \cdots + \alpha_m + \cdots + \alpha_n + \cdots + \alpha_p + \cdots$$

der Typus der obigen Mengensumme S verstanden.

Es gilt ein allgemeines assoziatives Gesetz, von dem wir nur den einfachsten Fall

$$(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma) = \alpha + \beta + \gamma$$

anführen wollen, aber kein kommutatives: ändert man die Ordnung von M, so ändert man S und im allgemeinen auch den Typus σ.

Beispiele. Die Spaltung der natürlichen Zahlenreihe (Typus ω) in

$$\{1, 2, \ldots, n\} + \{n + 1, n + 2, \ldots\},$$

wo der zweite Summand wieder vom Typus ω ist, gibt

$$n + \omega = \omega.$$

Dagegen ist $\omega + n$ der Typus von

$$\{n + 1, n + 2, \ldots,\ 1, 2, \ldots, n\},$$

und diese Menge hat, da sie ein letztes Element hat, gewiß nicht den Typus ω, also

$$\omega + n \neq n + \omega;$$

offenbar sind die Typen $\omega, \omega + 1, \omega + 2, \ldots$ paarweise verschieden. Die vier Anordnungen (S. 41) der natürlichen Zahlen bei Voranstellung der ungeraden vor die geraden haben die Typen

$$\omega + \omega, \quad \omega + \omega^*, \quad \omega^* + \omega, \quad \omega^* + \omega^*,$$

die wieder, wie leicht erkennbar, voneinander und von den Typen $\omega + n$ und deren Inversen $n + \omega^*$ verschieden sind. $\omega^* + \omega$ ist auch der Typus der Menge aller ganzen Zahlen

$$\{\ldots, -2, -1, 0, 1, 2, \ldots\}.$$

in natürlicher Reihenfolge.

Wird jeder natürlichen Zahl m eine natürliche Zahl a_m zugeordnet, so liefert die Spaltung der natürlichen Zahlenreihe in Gruppen von je a_m Gliedern die Typengleichung

$$\sum_m a_m = a_1 + a_2 + a_3 + \cdots = \omega,$$

z. B.[1]

$$1 + 1 + 1 + \cdots = \omega$$
$$1 + 2 + 3 + \cdots = \omega.$$

Verteilt man hingegen die natürlichen Zahlen auf abzählbar viele Reihen, etwa (Diagonalschema)

$$\{1, 2, 4, \ldots\} + \{3, 5, 8, \ldots\} + \{6, 9, 13, \ldots\} + \cdots,$$

so entsteht wieder ein neuer Typus

$$\sum_m \omega = \omega + \omega + \omega + \cdots.$$

Bei der Inversion einer Summe ist jeder Summand und die Ordnung der Summanden umzukehren, d. h.

$$\text{mit } S = \sum_m^M A_m \text{ ist } S^* = \sum_m^{M^*} A_m^*$$

und entsprechend für Typen. Z. B.

$$(\alpha + \beta)^* = \beta^* + \alpha^*.$$

Produkte endlich vieler Faktoren. Die geordneten Paare (a, b), $a \varepsilon A$, $b \varepsilon B$, wo A, B geordnete (nicht notwendig disjunkte) Mengen sind, lassen sich in eine Ordnung bringen, die man treffend die *lexikographische* oder die Ordnung *nach ersten Differenzen* nennt. Wir definieren nämlich

$$(a, b) < (a_1, b_1), \text{ wenn entweder } a < a_1$$
$$\text{oder } a = a_1, \ b < b_1.$$

Die so geordnete Menge soll wieder (A, B) genannt werden; es ist das geordnete Produkt der geordneten Mengen und von (B, A) zu unterscheiden, mit welcher Menge es äquivalent, aber im allgemeinen nicht ähnlich ist. Ist $A_1 \simeq A$, $B_1 \simeq B$, so ist $(A_1, B_1) \simeq (A, B)$, und dies berechtigt wieder, den Typus von (A, B) als Produkt der Typen α, β zu definieren. Leider kommen wir hier nicht um eine historisch gegebene Unbequemlichkeit
[24] herum: *der Typus von (A, B) wird $\beta\alpha$, nicht $\alpha\beta$ genannt*[2].

[1]) Da wir gleichzeitig $1 + 1 + 1 + \cdots = \aleph_0$ geschrieben haben, könnte die unterschiedslose Verwendung der endlichen Zahlen als Mächtigkeiten und als Ordnungstypen Bedenken erregen; andererseits sieht man es jeder Gleichung mit unendlichen Symbolen ja sofort an, ob sie Zahlen- oder Typengleichung sein soll.

[2]) Cantor hat ursprünglich $\alpha\beta$, später $\beta\alpha$ geschrieben, und dies hat sich überwiegend durchgesetzt; ein nochmaliger Wechsel würde die Verwirrung in Permanenz erklären. Man könnte die Unstimmigkeit beseitigen, indem man antilexikographisch, nach letzten Differenzen ordnet; das ist auch wieder unbequem.

Beispiel.

$$A = \{1, 2\}, \quad B = \{1, 2, 3, \ldots\}, \quad \alpha = 2, \quad \beta = \omega.$$

(A, B) ist die lexikographisch geordnete Menge der Paare (a, b), die also die Ordnung

$$(1, 1)(1, 2)(1, 3)\ldots(2, 1)(2, 2)(2, 3)\ldots$$

erhalten; ihr Typus ist $\beta\alpha = \omega 2 = \omega + \omega$. (B, A) ist die Menge der Paare (b, a) in der Ordnung

$$(1, 1)(1, 2)(2, 1)(2, 2)(3, 1)(3, 2)\ldots,$$

ihr Typus $\alpha\beta = 2\omega = \omega$.

Die Addition gleicher Summanden führt auch hier wieder zur Multiplikation, d. h. wenn alle $\alpha_m = \alpha$ sind und μ den Typus von M bedeutet, so ist

$$\sum_m^M \alpha = \alpha\mu.$$

In der Tat ist $\alpha\mu$ der Typus von (M, A), also der Menge der lexikographisch geordneten Paare (m, a). Die Menge dieser Paare bei festem m sei A_m, dann ist

$$(M, A) = \sum_m^M A_m,$$

woraus wegen $A_m \simeq A$ die behauptete Typengleichung folgt. $\alpha\mu$ entsteht, indem man „α in μ einsetzt", d. h. in eine Menge vom Typus μ für jedes Element eine Menge vom Typus α einsetzt.

Beispiele. $\omega + \omega + \omega = \omega 3$, $3 + 3 + 3 + \cdots = 3\omega = \omega$. Allgemein ist $n\omega = n + n + n + \cdots = \omega$, $\omega n = \omega + \omega + \cdots + \omega$ mit n Summanden.

Das distributive Gesetz besteht nur in bezug auf den zweiten Faktor, d. h. es ist

$$\beta \cdot \sum_m^M \alpha_m = \sum_m^M \beta\alpha_m,$$

während bei Nachsetzung des Faktors β diese Gleichung nicht zu gelten braucht.

In der Tat: ist $A = \sum_m^M A_m$, so ist für jede Menge B

$$(A, B) = \sum_m^M (A_m, B),$$

wie aus der lexikographischen Ordnung der Paare (a, b) unmittelbar folgt; hieraus ergibt sich die behauptete Typengleichung, die man auch durch „Einsetzung" von β in $\alpha = \sum_m^M \alpha_m$ unmittelbar erhält. Speziell ist

$$\gamma(\alpha + \beta) = \gamma\alpha + \gamma\beta,$$

aber nicht notwendig $(\alpha + \beta)\gamma = \alpha\gamma + \beta\gamma$.

Beispiel. $2(\omega + 1) = 2\omega + 2 = \omega + 2$, in der Tat ist (Einsetzung)

$$2(\omega + 1) = 2 + 2 + 2 + \cdots + 2 = \omega + 2.$$

Dagegen

$$(\omega + 1)\,2 = (\omega + 1) + (\omega + 1) = \omega 2 + 1 \neq \omega 2 + 2 = \omega 2 + 1 \cdot 2.$$

[25] Die Inversion von (A, B) ergibt die lexikographische Anordnung der Paare (a, b) von (A^*, B^*), also

$$(A, B)^* = (A^*, B^*),$$
$$(\beta\alpha)^* = \beta^* \alpha^*,$$

bei der Inversion eines Produkts sind die Faktoren, nicht ihre Reihenfolge umzukehren (anders als bei der Addition).

Die Ausdehnung der Multiplikation auf drei und mehr Faktoren ist evident. So sei (A, B, C) die lexikographisch geordnete Menge der Tripel (a, b, c), wo also $(a, b, c) < (a_1, b_1, c_1)$, wenn

$$\begin{array}{ll} \text{entweder} & a < a_1 \\ \text{oder} & a = a_1,\ \ b < b_1 \\ \text{oder} & a = a_1,\ \ b = b_1,\ \ c < c_1; \end{array}$$

ihr Typus heißt $\gamma\beta\alpha$. Es gilt ersichtlich das assoziative Gesetz

$$\gamma(\beta\alpha) = (\gamma\beta)\alpha = \gamma\beta\alpha.$$

In gleicher Weise kann jede beliebige endliche Faktorenzahl behandelt werden, und weiter: ist $M = \{1, 2, 3, \ldots\}$ die Menge der natürlichen Zahlen, so lassen sich die Komplexe (Folgen)

$$p = (a_1, a_2, a_3, \ldots) \quad (a_m \varepsilon A_m)$$

auch noch lexikographisch ordnen, denn zwei verschiedene Komplexe p und $q = (b_1, b_2, b_3, \ldots)$ haben eine *erste Differenzstelle* m, d. h.

$$a_1 = b_1, \ldots, a_{m-1} = b_{m-1}, \quad a_m \neq b_m,$$

und dann definiere man $p < q$, falls $a_m < b_m$. Das so geordnete Produkt werde $(A_1, A_2, A_3, \ldots)$ und sein Typus $\ldots \alpha_3 \alpha_2 \alpha_1$ genannt.

Beispiel. Ist jedes A_m die Menge der natürlichen Zahlen, so erhalten wir $\ldots \omega\omega\omega$ als Typus der lexikographisch geordneten Menge der *natürlichen Zahlenfolgen* $p = (a_1, a_2, a_3, \ldots)$. Ordnen wir jedem p die reelle Zahl

$$x = \left(\frac{1}{2}\right)^{a_1} + \left(\frac{1}{2}\right)^{a_1 + a_2} + \left(\frac{1}{2}\right)^{a_1 + a_2 + a_3} + \cdots$$

zu (umkehrbar eindeutig), so sieht man, daß der lexikographischen Ordnung der p die *umgekehrte* Ordnung der x nach der Größe entspricht, d. h. $p < q$ so viel wie $x > y$. Da x das Intervall $(0, 1]$ durchlief, so ist also $\ldots \omega\omega\omega$ der Typus der Zahlen $1 - x$ in natürlicher Ordnung, d. h. des Intervalls $[0, 1)$.

Offenbar wird man die lexikographische Ordnung auf jedes Produkt $\prod\limits_m^M A_m$ ausdehnen können, wenn man nur sicher ist, daß zwei verschiedene seiner Komplexe stets eine erste Differenzstelle haben, also allgemein, *wenn* *jede Teilmenge von M ein erstes Element hat* (d. h. M wohlgeordnet ist, Kap. IV).

Auf diese und weitere Ausdehnungen des Produktbegriffs soll später (§ 16) eingegangen werden. Auch ein allgemeiner Potenzbegriff kann erst dann erklärt werden, indessen werden wir natürlich $\alpha\alpha = \alpha^2$, $\alpha\alpha\alpha = \alpha^3, \ldots$ setzen. (Das vorhin genannte Produkt $\ldots\, \omega\omega\omega$ wäre ω^{ω^*} zu schreiben.) So ist

$$\omega^2 = \omega\omega = \omega + \omega + \omega + \cdots$$

der Typus einer Folge von Folgen oder einer Doppelfolge, ω^3 der einer Folge von Doppelfolgen oder einer dreifachen Folge usw. Es ist

$$\omega + \omega^2 = \omega(1 + \omega) = \omega\omega = \omega^2,$$

$$\omega^2 + \omega = \omega(\omega + 1)$$

ein davon verschiedener Typus;

$$(\omega + \omega)\,\omega = (\omega 2)\,\omega = \omega(2\omega) = \omega\omega = \omega^2,$$

verschieden von

$$\omega(\omega + \omega) = \omega(\omega 2) = \omega^2 2 = \omega^2 + \omega^2.$$

Die aus endlichen Typen und ω durch endlich viele Additionen und Multiplikationen entstehenden Typen können ganze rationale Funktionen von ω oder Polynome genannt werden; sie lassen sich, und zwar eindeutig (wie wir noch sehen werden) auf die Form

$$\omega^n a + \omega^{n_1} a_1 + \cdots + \omega^{n_k} a_k$$

bringen ($n > n_1 > \cdots > n_k \geqq 0$, $a, a_1, \ldots, a_k$ natürliche Zahlen).

§ 11. Typen der Mächtigkeit $\aleph_0$ und $\aleph$.

Jeder Typus α hat eine bestimmte Mächtigkeit $\mathfrak{a}$. Die verschiedenen Typen der Mächtigkeit $\mathfrak{a}$ bilden eine *Typenklasse* $T(\mathfrak{a})$; um sie zu er- [26]
halten, braucht man nur eine feste Menge A von der Mächtigkeit $\mathfrak{a}$ auf alle möglichen Weisen zu ordnen, wobei verschiedene Ordnungen natürlich nicht notwendig verschiedene Typen liefern.

Die Typenklassen $T(n)$ endlicher Mächtigkeit ($n = 0, 1, 2, \ldots$) enthalten immer nur einen Typus n. Aber von $T(\aleph_0)$ haben wir schon unendlich viele Typen kennengelernt: ω, $\omega + 1$, $\omega + 2, \ldots, \omega + \omega$, ω^* u. a.

I. *Die Menge der abzählbaren Typen ist von der Mächtigkeit des Kon-* [27]
tinuums.

D. h. $T(\aleph_0)$ hat die Mächtigkeit $\aleph$. Die eine Hälfte dieses Satzes ist uns bereits bekannt: es gibt (S. 43) höchstens $2^{\aleph_0\aleph_0} = 2^{\aleph_0} = \aleph$ verschiedene abzählbare Typen.

Andererseits sei $\zeta = \omega^* + \omega$ der Typus der Menge $\{\ldots -2, -1, 0, 1, 2, \ldots\}$ der ganzen Zahlen in natürlicher Ordnung,

$$a = (a_1, a_2, a_3, \ldots)$$

eine Folge natürlicher Zahlen und

$$\alpha = a_1 + \zeta + a_2 + \zeta + a_3 + \zeta + \cdots$$

ein durch a bestimmter, offenbar abzählbarer Typus. Wenn wir zeigen können, daß α auch seinerseits die Folge a bestimmt, d. h. zwischen den a und α eineindeutige Zuordnung besteht, so ist bewiesen, daß es auch mindestens $\aleph_0^{\aleph_0} = \aleph$ verschiedene abzählbare Typen gibt und, nach dem Äquivalenzsatz, daß es deren genau $\aleph$ gibt.

Wir haben also zu zeigen: ist $\beta = b_1 + \zeta + b_2 + \zeta + \cdots$ und $\alpha = \beta$, so ist $a_1 = b_1$, $a_2 = b_2, \ldots$. Das beruht auf folgenden Schlüssen:

(α) Ist $A_1 + A_2 \simeq B_1 + B_2$, A_1 und B_1 endlich, A_2 und B_2 ohne erstes Element, so ist $A_1 \simeq B_1$, $A_2 \simeq B_2$.

Denn bei der ähnlichen Abbildung kann ein Element $b_1 \varepsilon B_1$ nicht Bild eines Elements $a_2 \varepsilon A_2$ sein, weil b_1 nur endlich viele Vorgänger (eventuell keinen) hat, a_2 unendlich viele. Ebensowenig kann ein b_2 einem a_1 entsprechen. Also müssen den a_1 die b_1 und den a_2 die b_2 entsprechen.

(β) Ist $A_1 + A_2 \simeq B_1 + B_2$, A_1 und B_1 vom Typus ζ, so ist $A_2 \simeq B_2$.

Auch hier kann einem a_2 kein b_1 entsprechen, denn die Menge der Vorgänger von a_2 enthält eine Teilmenge (A_1) ohne letztes Element, während die der Vorgänger von b_1 vom Typus ω^* ist und keine Teilmenge ohne letztes Element (natürlich die Nullmenge ausgenommen) enthält. Wieder also muß A_1 auf B_1 und A_2 auf B_2 abgebildet werden.

Aus: $a_1 + \zeta + \cdots = b_1 + \zeta + \cdots$ folgt daher nach (α):

$$a_1 = b_1, \quad \zeta + a_2 + \zeta + \cdots = \zeta + b_2 + \zeta + \cdots,$$

hieraus nach (β) $a_2 + \zeta + \cdots = b_2 + \zeta + \cdots$, dann wieder $a_2 = b_2$ usw.

Eine (unendliche) Menge ohne erstes und letztes Element heiße *unbegrenzt*, eine solche ohne benachbarte Elemente *dicht*; wir übertragen diese Prädikate auch auf ihren Typus. Es wird also verlangt, daß vor wie nach jedem Element bzw. zwischen zwei Elementen immer noch weitere Elemente der Menge existieren. Solche Mengen sind die der rationalen Zahlen und die der reellen Zahlen in natürlicher Ordnung; ihre Typen werden η und λ genannt (λ der Typus des Kontinuums).

II. *Ist A abzählbar, B unbegrenzt und dicht, so ist A mit einer Teilmenge von B ähnlich.*

[28] III. *Zwei unbegrenzte dichte abzählbare Mengen sind ähnlich.*

Beweis von II. Sei $A = \{a_1, a_2, \ldots\}$ [1]), es wird behauptet, daß man A auf eine Teilmenge von B unter Erhaltung der Ordnung abbilden kann. Indem wir dem a_1 ein beliebiges Element von B zuordnen, wenden wir sodann Induktion an: den Elementen von $A_n = \{a_1, a_2, \ldots, a_n\}$ seien bereits

[1]) Die Ordnung der Elemente von links nach rechts in $\{a_1, a_2, \ldots\}$ bedeutet hier nicht die Ordnung von A.

Bilder in der verlangten Weise zugeordnet, und dasselbe soll nun für a_{n+1} geschehen. Entweder fällt a_{n+1} zwischen zwei Elemente von A_n oder es ist $a_{n+1} < A_n$ oder endlich $a_{n+1} > A_n$. Wir haben dann also nur zwischen zwei Elementen von B oder vor einem solchen (dem Bilde des ersten Elements von A_n) oder nach einem solchen ein weiteres Element von B zu suchen, und das ist möglich, weil B unbegrenzt und dicht ist. Es gibt also ein passendes Bild von a_{n+1}, und man kann der Reihe nach allen a_n Bilder zuordnen mit Erhaltung der Ordnung.

Beweis von III. Seien $A = \{a_1, a_2, \ldots\}$ und $B = \{b_1, b_2, \ldots\}$ unbegrenzt und dicht; man kann also A auf eine Teilmenge von B, aber auch B auf eine Teilmenge von A abbilden [1]), und indem man die einzelnen Zuordnungsakte abwechselnd in der einen und andern Richtung ausführt, kann man auch A auf B abbilden. Ordnen wir dem a_1 das b_1 zu und schreiben $a^1 = a_1$, $b^1 = b_1$. Nun wenden wir Induktion an: es seien bereits die Paare $(a^1, b^1), \ldots, (a^n, b^n)$ unter Erhaltung der Ordnung gebildet und die Mengen $A^n = \{a^1, \ldots, a^n\}$, $B^n = \{b^1, \ldots, b^n\}$ also ähnlich aufeinander bezogen; es soll nun das Paar (a^{n+1}, b^{n+1}) hinzugefügt werden. Für gerades n sei a^{n+1} das a_k mit niedrigstem Index k, das nicht zu A^n gehört, und b^{n+1} das b_k mit niedrigstem Index k, das zu B^n dieselbe Lage hat wie a^{n+1} zu A^n. Für ungerades n sei umgekehrt b^{n+1} das niedrigste noch unabgebildete b_k und a^{n+1} das niedrigste mit den Ordnungsbeziehungen verträgliche a_k. Die Existenz passender Bilder ist durch die Unbegrenztheit und Dichtigkeit beider Mengen gesichert, während zugleich kein Element bei der Abbildung übergangen werden kann; also $A \simeq B$.

Aus II (worin man sowohl für A wie für B die Menge der rationalen Zahlen, vom Typus η, setzen kann) und III folgt demnach:

IV. *Jede unbegrenzte dichte Menge enthält eine Teilmenge vom Typus η. Eine Menge vom Typus η enthält abzählbare Teilmengen von jedem Typus. Jede unbegrenzte dichte abzählbare Menge ist vom Typus η.* [29]

Beispiele. $\sum_m^M \eta = \eta\mu$ ist, wenn M (vom Typus μ) endlich oder abzählbar ist, stets ein unbegrenzter dichter abzählbarer Typus, also $\eta\mu = \eta$, z. B.

$$\eta + \eta = \eta 2 = \eta n = \eta\omega = \eta^2 = \eta .$$

Dichte abzählbare Typen können sich nur durch Existenz oder Nichtexistenz von ersten und letzten Elementen unterscheiden, es gibt deren also vier: η, $1 + \eta$, $\eta + 1$, $1 + \eta + 1$.

$\sum_m^M (1 + \eta) = (1 + \eta)\mu$ (M endlich oder abzählbar) ist entweder $= 1 + \eta$ oder $= \eta$, je nachdem M ein erstes Element hat oder nicht, z. B.

---

[1]) Der dem Äquivalenzsatz analoge Ähnlichkeitssatz (zwei Mengen, deren jede einer Teilmenge der andern ähnlich ist, sind selbst ähnlich) ist falsch! Beispiel: ein Intervall mit und eins ohne Endpunkte.

$$(1+\eta)2 = (1+\eta)n = (1+\eta)\omega = 1+\eta,$$
$$(1+\eta)\omega^* = (1+\eta)\eta = \eta.$$

Die Menge der rationalen Zahlen $> a$ ist vom Typus η; es gibt also z. B. eine die Größenordnung erhaltende, d. h. mit r monoton wachsende Funktion $s = f(r)$, die jeder rationalen Zahl $r > 0$ eine rationale Zahl $s > a$ und umgekehrt zuordnet. Sie läßt sich offenbar zu einer stetigen, mit x monoton wachsenden Funktion $y = f(x)$ erweitern, die die Halbgerade $x > 0$ auf die Halbgerade $y > a$ abbildet und dabei jedem rationalen x ein rationales y, jedem irrationalen x ein irrationales y zuordnet. Für rationales a ist das natürlich trivial, da $y = x + a$ eine solche Zuordnung bewirkt.

Die Menge der rationalen Zahlen des Intervalls $(0, 1)$ und die der dyadisch rationalen (Brüche mit einer Potenz von 2 als Nenner) desselben Intervalls sind ähnlich, beide vom Typus η. Die Zuordnung, die wir sogleich zu einer das ganze Intervall betreffenden

$$y = f(x) \qquad (0 < x \leqq 1,\ 0 < y \leqq 1)$$

erweitern wollen, die einem dyadisch rationalen x ein rationales y zuordnet und umgekehrt, kann folgendermaßen geschehen. Wir zerlegen das Intervall $0 < x \leqq 1$ durch Einschiebung von $\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \ldots$ und das Intervall $0 < y \leqq 1$ durch Einschiebung von $\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots$ in die Teilintervalle

$$\left(\frac{1}{2}\right)^{n_1} < x \leqq \left(\frac{1}{2}\right)^{n_1 - 1}$$
$$\frac{1}{n_1+1} < y \leqq \frac{1}{n_1},$$

wo n_1 eine natürliche Zahl ist. Dafür können wir schreiben

$$x = \left(\frac{1}{2}\right)^{n_1}(1 + x_1) \qquad (0 < x_1 \leqq 1)$$
$$y = \frac{1}{n_1 + 1 - y_1} \qquad (0 < y_1 \leqq 1).$$

Indem wir dies Verfahren für x_1, y_1 wiederholen und so unbegrenzt fortfahren, gelangen wir für x zu dem dyadischen Bruch

$$x = \left(\frac{1}{2}\right)^{n_1} + \left(\frac{1}{2}\right)^{n_1+n_2} + \left(\frac{1}{2}\right)^{n_1+n_2+n_3} + \cdots,$$

für y zu dem Kettenbruch

$$y = \frac{1\,|}{|\,n_1+1} - \frac{1\,|}{|\,n_2+1} - \frac{1\,|}{|\,n_3+1} - \cdots,$$

wo $n = (n_1, n_2, n_3, \ldots)$ eine Folge natürlicher Zahlen ist. Die durch gleiches n vermittelte Zuordnung zwischen x, y hat die gewünschte Eigenschaft, einem dyadisch rationalen x ein rationales y zuzuordnen und umgekehrt; überdies ordnet sie einem rationalen (aber nicht dyadisch rationalen) x ein quadratisch irrationales y zu und umgekehrt (H. Minkowski). Die genauere Ausführung dieser Angaben müssen wir dem in den Elementen der Kettenbruchtheorie bewanderten Leser überlassen.

Der Typus λ der Menge aller reellen Zahlen ist auch der eines Intervalls ohne Endpunkte oder einer Halbgeraden ohne Endpunkt ($x > a$, $x < a$).

Die vier Intervalle

	(a, b)	$[a, b)$	$(a, b]$	$[a, b]$
haben die Typen	λ	$1 + \lambda$	$\lambda + 1$	$1 + \lambda + 1$.

Aus $(a, b) + [b, c) = (a, c)$ $(a < b < c)$ erhält man $\lambda + 1 + \lambda = \lambda$, während $\lambda + \lambda \neq \lambda$ ist; ebenso liefert die Aneinanderreihung passender Intervalle

$$1 + \lambda = (1 + \lambda)\,2 = (1 + \lambda)\,n = (1 + \lambda)\,\omega, \quad \lambda = (\lambda + 1)\,\omega \quad \text{u. a.}$$

Um λ genauer zu charakterisieren, stellen wir folgende allgemeine Betrachtung an. Eine Zerlegung

$$A = P + Q \quad (P < Q)$$

der geordneten Menge A in ein „Anfangsstück" P und ein „Endstück" Q kann folgende vier Fälle darbieten:

P hat ein letztes, Q ein erstes Element: *Sprung*
P hat ein letztes, Q kein erstes Element: } *Schnitt*
P hat kein letztes, Q ein erstes Element: }
P hat kein letztes, Q kein erstes Element: *Lücke*.

Eine Menge ohne Sprünge ist dicht, z. B. die Menge der rationalen Zahlen, die aber Lücken hat. Eine Menge ohne Sprünge und Lücken heißt im Dedekindschen Sinne *stetig*; z. B. die Menge der reellen Zahlen (daher die Ausdrücke Zahlenkontinuum, Mächtigkeit des Kontinuums). Sie entsteht aus der Menge der rationalen Zahlen durch Ausfüllung der Lücken (mittels der irrationalen Zahlen); die bekannte Art, wie Dedekind das gemacht hat, läßt sich auf eine beliebige dichte Menge A übertragen, die wir überdies unbegrenzt annehmen wollen. Betrachten wir die Zerlegungen

$$A = P + Q,$$

wo P kein letztes Element hat; Q kann ein erstes haben oder nicht. Diese Anfangsstücke P bilden selbst eine geordnete Menge $\mathfrak{P}$, wenn man nämlich $P < P_1$ für $P < P_1$ definiert, und wir behaupten: $\mathfrak{P}$ *ist stetig*. Denn erstens ist $\mathfrak{P}$ dicht, d. h. zwischen P_1 und $P_2 > P_1$ liegt immer noch ein weiteres P; da nämlich $P_2 - P_1$ kein letztes Element, also unendlich viele hat, so sei a ein Element von $P_2 - P_1$, das nicht das erste ist, und P die Menge aller Elemente $< a$ von A, dann ist $P_1 < P < P_2$. Zweitens hat $\mathfrak{P}$ auch keine Lücken. Denn sei $\mathfrak{P} = \mathfrak{P}_1 + \mathfrak{P}_2$ Zerlegung von $\mathfrak{P}$ in eine Anfangsklasse und Endklasse, d. h. in zwei Klassen von Anfangsstücken P_1 und P_2, wo durchweg $P_1 < P_2$. Bilden wir die Summe aller P_1, $P = \mathfrak{S} P_1$; sie liefert mit dem Durchschnitt $Q = \mathfrak{D} Q_1$ der Komplemente ($A = P_1 + Q_1$) wieder eine Zerlegung $A = P + Q$, wo P also ein Anfangsstück und offenbar ohne letztes Element ist. Aus $P_1 < P_2$ folgt nun $P \leqq P_2$, d. h. durchweg $P_1 \leqq P \leqq P_2$, und also ist P entweder das letzte Element von $\mathfrak{P}_1$ oder das erste von $\mathfrak{P}_2$; $\mathfrak{P}_1 + \mathfrak{P}_2$ ist ein Schnitt, weder Sprung noch Lücke. Also $\mathfrak{P}$ ist stetig.

Bedienen wir uns noch der Ausdrucksweise:

$B \leqq A$ ist *in A dicht*, wenn zwischen zwei Elementen von A immer mindestens ein Element von B liegt,

wobei also A, B selber dicht sind, so können wir den Typus λ des Kontinuums durch folgenden Satz charakterisieren:

V. *Jede stetige Menge enthält eine Teilmenge vom Typus λ. Jede unbegrenzte stetige Menge, in der eine abzählbare Menge dicht ist, ist vom Typus λ.*

Beweis. A sei stetig; wir können sie überdies als unbegrenzt annehmen. Sie enthält (nach IV) eine Teilmenge B vom Typus η. Ist $B = P + Q$ eine Lücke von B, so muß zwischen P und Q mindestens ein Element von A liegen, da andernfalls, wie leicht zu sehen, auch A eine Lücke hätte. Also enthält A eine Teilmenge C, die aus B durch Ausfüllung der Lücken entsteht, d. h. vom Typus λ ist. Andererseits, wenn B zugleich in A dicht ist (jede in A dichte Menge ist wieder unbegrenzt und dicht, also, falls abzählbar, vom Typus η), so kann zwischen P und Q auch nur ein einziges Element von A liegen, d. h. $A = C$. Damit ist der Satz bewiesen.

Es gibt unendlich viele verschiedene stetige Typen von der Mächtigkeit des Kontinuums. Ist $\vartheta = 1 + \lambda + 1$ der Typus des abgeschlossenen Intervalls $J = [0, 1]$, so sind die Potenzen $\vartheta, \vartheta^2, \vartheta^3, \ldots$ stetig und voneinander verschieden. Sei nämlich $J_2 = (J, J)$, $J_3 = (J, J, J), \ldots, J_m$ die lexikographisch geordnete Menge der Zahlenkomplexe $x = (x_1, x_2, \ldots, x_m)$, wo jedes x_k das Intervall J durchläuft. Die Dichtigkeit von J_m ist evident; die Lückenlosigkeit von J_m wird auf die von J_{m-1} folgendermaßen zurückgeführt. Es sei $H_m(a)$ die mit J_{m-1} ähnliche Menge der Komplexe $(a, x_2, \ldots, x_m)$ bei festem $x_1 = a$, also

$$J_m = \sum_a^J H_m(a).$$

Bei einer Zerlegung $J_m = P_m + Q_m$ wird entweder einer der Summanden $H_m(a)$ mit zerlegt und damit der Fall auf eine Zerlegung $J_{m-1} = P_{m-1} + Q_{m-1}$ zurückgeführt, oder die Zerlegung ist von der Form

$$P_m = \sum_a^P H_m(a), \quad Q_m = \sum_b^Q H_m(b), \quad J = P + Q,$$

und dann hat, wenn a_1 das größte a ist, P_m das letzte Element $(a_1, 1, \ldots, 1)$, wenn b_1 das kleinste b ist, Q_m das erste Element $(b_1, 0, \ldots, 0)$. — Um die

Verschiedenheit der ϑ^m zu beweisen, zeigen wir: *ist $m > 1$ und J_m einer Teilmenge von J_n ähnlich, so ist $n > 1$ und J_{m-1} einer Teilmenge von J_{n-1} ähnlich.* Die Menge J_m der Komplexe $x = (x_1, \ldots, x_m)$ sei auf eine Teilmenge von J_n, der Menge der Komplexe $y = (y_1, \ldots, y_n)$ ähnlich abgebildet. Den beiden Komplexen $(x_1, 0, \ldots, 0)$ und $(x_1, 1, \ldots, 1)$ mögen als Bilder $y = (y_1, y_2, \ldots, y_n)$ und $\eta = (\eta_1, \eta_2, \ldots, \eta_n)$ entsprechen mit $y < \eta$, also $y_1 \leqq \eta_1$. Es kann dann nicht durchweg (für jedes x_1) $y_1 < \eta_1$ sein (also insbesondere nicht $n = 1$), denn die offenen Intervalle (y_1, η_1) wären dann disjunkt und ihre Menge wie die der x_1 von der Mächtigkeit $\aleph$, während es doch solcher Intervalle, in deren jedem gewiß eine rationale Zahl liegt, nur abzählbar viele geben kann. Es gibt also ein $x_1 = a$ derart, daß den Komplexen $(a, 0, \ldots, 0)$ und $(a, 1, \ldots, 1)$ als Bilder Komplexe $(b, y_2, \ldots, y_n)$ und $(b, \eta_2, \ldots, \eta_n)$ mit demselben $y_1 = \eta_1 = b$ entsprechen; die vorhin angeführte Menge $H_m(a)$ ist dann mit einer Teilmenge von $H_n(b)$, also J_{m-1} mit einer Teilmenge von J_{n-1} ähnlich. — Endlich folgt nun, daß für $m > n$ nicht J_m mit einer Teilmenge von J_n (auch nicht mit J_n selbst) ähnlich sein kann, denn es wäre J_{m-1} mit einer Teilmenge von J_{n-1} und, so fortfahrend, J_{m-n+1} mit einer Teilmenge von J_1 ähnlich, im Widerspruch zum Obigen.

Viertes Kapitel.

Ordnungszahlen.

§ 12. Der Wohlordnungssatz. [30]

Wir definieren: *eine geordnete Menge heißt wohlgeordnet, wenn jede* (nicht leere) *Teilmenge ein erstes Element hat.* Der Ordnungstypus einer wohlgeordneten Menge heißt eine *Ordnungszahl* oder Ordinalzahl.

In einer wohlgeordneten Menge gibt es keine Teilmenge vom Typus ω^*, jede fallende Reihe von Elementen $a > b > c > \cdots$ enthält nur endlich viele Glieder. Diese Eigenschaft könnte auch zur Definition dienen.

Auf jedes Element, wenn es nicht das letzte ist, folgt unmittelbar ein nächstes; bei jeder Zerlegung $A = P + Q$ hat das Endstück Q ein erstes Element, mag das Anfangsstück P ein letztes haben oder nicht. Umgekehrt, wenn bei jeder Zerlegung (die uneigentliche $0 + A$ mitgerechnet) das Endstück Q ein erstes Element hat, so ist A wohlgeordnet; denn ist $B > 0$ eine beliebige Teilmenge von A, P die Menge der Elemente $< B$ und $A = P + Q$, so ist das erste Element von Q auch das erste von B.

Die endlichen Mengen $\{1, 2, \ldots, n\}$, die Menge $\{1, 2, 3, \ldots\}$ der natürlichen Zahlen, die Menge $\{1, 3, 5, \ldots, 2, 4, 6, \ldots\}$ sind wohlgeordnet, ihre Typen $n, \omega, \omega + \omega$ Ordnungszahlen. Die Ordnungstypen ω^*, η, λ sind keine Ordnungszahlen.

Eine unendliche wohlgeordnete Menge A hat ein erstes Element a_0, dann ein zweites a_1, ein drittes a_2 usw.; wenn sie außer dieser Folge noch weitere Elemente hat, so ist unter diesen ein erstes a_ω, dann wieder ein nächstes $a_{\omega+1}$ usw.

Es ist also

(1) $$A = \{a_0, a_1, a_2, \ldots, a_\omega, a_{\omega+1}, \ldots\}.$$

Das hier angedeutete, nachher durchzuführende Bezeichnungsprinzip ist, daß jedem Element als Index der Typus der Menge der vorangehenden Elemente zugeordnet wird. Damit dies auch für die endlichen Indizes zutreffe, haben wir mit 0 begonnen; a_n trägt als Index den Typus n der Menge $\{a_0, \ldots, a_{n-1}\}$, a_0 den Typus 0 der Nullmenge.

Nach dieser vorläufigen Orientierung beweisen wir den

[31] *Wohlordnungssatz von E. Zermelo*: **Jede Menge kann wohlgeordnet werden.**

Wenn wir zu einer Wohlordnung in der Gestalt (1) gelangen wollen, müssen wir z. B. jeder Menge $P_n = \{a_0, a_1, \ldots, a_{n-1}\}$ ein bestimmtes Element a_n als das zunächst folgende zuordnen oder aus der Menge $Q_n = A - P_n$ der noch ungeordneten Elemente ein bestimmtes a_n gewählt denken. Wenn man so verfährt, erscheinen die Wahlakte selbst in einer gewissen Ordnung: a_n kann und muß erst dann gewählt werden, wenn seine Vorgänger $a_0, \ldots, a_{n-1}$ gewählt sind. Der Beweis des Wohlordnungssatzes kann auch mittels dieser *sukzessiven* Wahlakte geführt werden, dann aber erst an späterer Stelle, nach genauerer Untersuchung der Ordnungszahlen. Um ihn schon jetzt zu erbringen, verwenden wir ein System *simultaner*, voneinander unabhängiger Wahlakte: wir ordnen *jeder von A verschiedenen Teilmenge P von A ein ihr nicht angehöriges Element $a = f(P) \varepsilon A - P$* zu, oder wir wählen *aus jeder von 0 verschiedenen Teilmenge Q von A ein ihr angehöriges Element $a = \varphi(Q) \varepsilon Q$*. Beide Ausdrucksweisen bezeichnen dasselbe: es ist $f(P) = \varphi(A - P)$ oder $\varphi(Q) = f(A - Q)$. Wir wollen die erste Form bevorzugen und nennen $a = f(P)$ das *Ansatzelement* von P und die aus P durch Hinzufügung des Ansatzelementes entstehende Menge $P_+ = P + \{a\}$ die *Nachfolgerin* von P. Bei diesem Verfahren werden mehr Wahlakte vollzogen, als unbedingt notwendig sind, denn bei der Wohlordnung (1) würde ja z. B. die Menge $P = \{a_0, a_2\}$ und ihr Ansatzelement gar nicht gebraucht werden. Dafür aber sind die Wahlakte, wie gesagt, voneinander unabhängig, anders ausgedrückt: die Funktion $a = f(P)$ hat einen von vornherein feststehenden Definitionsbereich, nämlich die Menge aller $P < A$.

Die Art, wie nun aus dieser Zuordnung $a = f(P)$ zwangsläufig eine Wohlordnung von A hervorgeht, ist im Grunde sehr einfach, obwohl sie an das abstrakte Denken des Lesers einige Ansprüche stellt. Wir betrachten ein System von Mengen $\leqq A$, das

(α) die Nullmenge enthält,

(β) mit beliebig vielen Mengen auch deren Summe enthält,

(γ) mit jeder Menge $P < A$ auch deren Nachfolgerin P_+ enthält.

Ein solches System heiße eine *Kette*. Es gibt Ketten, z. B. die umfassendste: das System aller Mengen $\leqq A$. Der Durchschnitt beliebig vieler Ketten ist ersichtlich wieder eine Kette; es gibt also eine *kleinste Kette* $\mathfrak{K}$, die der Durchschnitt *aller* Ketten ist. Diese untersuchen wir jetzt und verstehen unter den im Folgenden vorkommenden Mengen wie P, X immer Mengen aus $\mathfrak{K}$.

Das Wesentliche ist, die Vergleichbarkeit aller Mengen von $\mathfrak{K}$ zu zeigen (im Sinne von S. 13), d. h. daß für zwei Mengen P, X immer eine der drei Relationen $X \lesseqgtr P$ besteht. Nennt man eine Menge P, die mit allen Mengen X vergleichbar ist, *normal*, so ist zu beweisen, daß alle Mengen normal sind. Der erste Schritt besteht im Nachweis des Satzes:

I. *Ist $P < A$ normal, so sind alle Mengen entweder $\leqq P$ oder $\geqq P_+$.*

Wir zeigen, daß diese Mengen X ($\leqq P$ oder $\geqq P_+$) eine Kette bilden; diese muß mit $\mathfrak{K}$ identisch sein, da $\mathfrak{K}$ die kleinste Kette war. Wir haben also die Ketteneigenschaften festzustellen:

(α) Die Nullmenge ist ein X.

Evident, weil $0 \leqq P$.

(β) Die Summe beliebig vieler X ist ein X.

Sei $S = \mathfrak{S} X_m$; entweder ist jedes $X_m \leqq P$ und $S \leqq P$, oder mindestens ein $X_m \geqq P_+$ und $S \geqq P_+$.

(γ) Die Nachfolgerin jedes $X < A$ ist ein X.

Für $X \geqq P_+$ ist $X_+ > P_+$; für $X = P$ ist $X_+ = P_+$. Für $X < P$ muß $X_+ \leqq P$ sein; denn X_+ ist jedenfalls mit dem normalen P vergleichbar, wäre aber $X_+ > P$, so müßte $X_+ - X = (X_+ - P) + (P - X)$ mindestens zwei Elemente enthalten, während diese Menge doch nur ein Element $f(X)$ enthält.

Nun können wir schließen:

II. *Alle Mengen sind normal.*

Wir zeigen wieder, daß die normalen Mengen eine Kette bilden, die folglich mit $\mathfrak{K}$ identisch sein muß.

(α) Die Nullmenge ist normal.

(β) Die Summe beliebig vieler normaler Mengen ist normal.

Sei $P = \mathfrak{S} P_m$ Summe normaler Mengen, X eine beliebige Menge, also $P_m \lesseqgtr X$. Entweder ist jedes $P_m \leqq X$ und $P \leqq X$, oder mindestens ein $P_m > X$ und $P > X$. P ist also mit jedem X vergleichbar.

(γ) Die Nachfolgerin P_+ jeder normalen Menge $P < A$ ist normal. Das ist durch den Satz I bewiesen.

Durch die Vergleichbarkeit aller Mengen wird nun $\mathfrak{K}$ geordnet, indem man von zwei verschiedenen Mengen die kleinere vor die größere setzt ($P_1 < P_2$ so viel wie $P_1 < P_2$). *Diese Ordnung ist eine Wohlordnung.* Da $\mathfrak{K}$ selbst ein erstes Element hat (die Nullmenge), so ist nur zu zeigen, daß bei einer Zerlegung $\mathfrak{K} = \mathfrak{K}_1 + \mathfrak{K}_2$ in ein Anfangs- und Endstück das letztere ein erstes Element hat. Sei $P_1 \varepsilon \mathfrak{K}_1$, $P_2 \varepsilon \mathfrak{K}_2$, $P_1 < P_2$; P die Summe aller P_1. Dann ist $P_1 \leqq P \leqq P_2$, also P entweder das erste P_2 oder das letzte P_1. Im zweiten Falle ist aber P_+ das erste P_2; wir haben die Spaltung des Satzes I vor uns. In beiden Fällen gibt es ein erstes P_2.

Endlich lassen sich die Mengen $P < A$ den Elementen $a \varepsilon A$ eineindeutig zuordnen, und damit wird auch A wohlgeordnet. Das geschieht durch die Relation $a = f(P)$, die das Ansatzelement von P bestimmt. Verschiedene Mengen $P_1 < P_2$ haben verschiedene Ansatzelemente a_1, a_2; denn P_2 enthält die Nachfolgerin von P_1, also $a_1 \varepsilon P_2$, $a_2 \bar{\varepsilon} P_2$. Andererseits ist jedes a Ansatzelement einer, und also nur einer, Menge P. Denn sei $P = F(a)$ die Summe aller Mengen P_1, die a nicht enthalten (zu denen z. B. die Nullmenge gehört); dann muß $a = f(P)$ sein, da andernfalls auch $P_+ > P$ noch das Element a nicht enthielte. Durch

$$a = f(P), \quad P = F(a)$$

werden also die Elemente $a \varepsilon A$ und die Mengen $P < A$ (d. h. also alle Mengen von $\mathfrak{K}$ bis auf die letzte A) eineindeutig aufeinander bezogen; die Übertragung der Ordnung von den P auf die a ($a_1 < a$ so viel wie $P_1 < P$) macht also A wohlgeordnet. Da $P_1 < P$ so viel wie $a_1 \varepsilon P$ bedeutet, so ist $P = F(a)$ die Menge der Elemente $a_1 < a$ und umgekehrt $a = f(P)$ das nächste in der Wohlordnung auf P folgende Element (dies gilt aber nur für die in der Kette $\mathfrak{K}$ auftretenden Mengen $P < A$).

Ein Beispiel zu diesem Wohlordnungsverfahren: aus jeder Menge Q natürlicher Zahlen werde diejenige Zahl $a = \varphi(Q)$ gewählt, die die wenigsten Primfaktoren hat und unter denen mit gleicher Zahl der Primfaktoren die kleinste ist. Es entsteht folgende Wohlordnung der natürlichen Zahlen: zuerst kommt die Zahl 1, dann die Primzahlen der Größe nach, dann die Produkte aus zwei Primfaktoren, wieder der Größe nach, usw.; sie hat den Typus $\omega + \omega + \omega + \cdots = \omega^2$.

§ 13. Die Vergleichbarkeit der Ordnungszahlen.

Jedes Element a der wohlgeordneten Menge A bestimmt

den *Abschnitt* P = Menge aller Elemente $< a$,
den *Rest* Q = Menge aller Elemente $\geqq a$

und damit die Zerlegung $A = P + Q \quad (P < A, \quad Q > 0)$.

Umgekehrt ist jede Zerlegung von A in ein Anfangsstück P und ein Endstück Q von dieser Form, da Q ein erstes Element a hat. Ist a das erste Element von A, so ist $P = 0$, $Q = A$ zu setzen.

Nun gilt:

I. *Ist $b = f(a)$ eine ähnliche Abbildung der wohlgeordneten Menge A auf eine Teilmenge B, so ist stets $f(a) \geqq a$.*

D. h. also, bei einer solchen Abbildung kann ein Element nie ein früheres zum Bilde haben. In der Tat: gäbe es Elemente a mit $f(a) < a$, so gäbe es unter diesen ein erstes; es sei dies a und $b = f(a)$ sein Bild, also $b < a$ und wegen der Ähnlichkeit $f(b) < f(a)$, d. h. $f(b) < b$; b wäre also doch noch ein früheres Element jener Eigenschaft.

Insbesondere:

II. *Eine wohlgeordnete Menge ist keinem ihrer Abschnitte ähnlich.*

Denn wäre A dem durch a bestimmten Abschnitt B ähnlich, so würde $f(a) \varepsilon B$, also $f(a) < a$ sein.

Die Beziehung, daß A einem Abschnitt von B ähnlich ist, bleibt bei Ersetzung dieser Mengen durch ähnliche erhalten. Dies berechtigt zu der Definition:

Sind α und β zwei Ordnungszahlen, A und B wohlgeordnete Mengen dieser Typen, so werde

$$\alpha < \beta \text{ oder } \beta > \alpha$$

(α kleiner als β, β größer als α) erklärt, falls A einem Abschnitt von B ähnlich ist.

Offenbar gilt das transitive Gesetz:

$$\text{wenn } \alpha < \beta, \ \beta < \gamma, \text{ so ist } \alpha < \gamma$$

(A ist einem Abschnitt eines Abschnitts von C, d. h. einem Abschnitt von C ähnlich).

Nach II ist niemals $\alpha < \alpha$; d. h. die Relationen $\alpha < \beta$ und $\alpha = \beta$ schließen einander aus, ebenso $\alpha > \beta$ und $\alpha = \beta$. Aber auch $\alpha < \beta$ und $\alpha > \beta$, denn aus $\alpha < \beta$, $\beta < \alpha$ würde nach dem transitiven Gesetz $\alpha < \alpha$ folgen. Von den Relationen $\alpha \substack{< \\ = \\ >} \beta$ kann also höchstens eine bestehen; wir zeigen, daß auch immer eine besteht, d. h. daß zwei Ordnungszahlen stets vergleichbar sind.

Wir führen dauernd die folgende Bezeichnung ein: jede Ordnungszahl α bestimmt die Menge

$$W(\alpha) = \text{Menge der Ordnungszahlen} < \alpha,$$ [32]

die etwa ein *Zahlenabschnitt* heiße. *Die Zahlen von $W(\alpha)$ sind vergleichbar und $W(\alpha)$ ist, der Größe nach geordnet, vom Typus α.* Denn ist

$$A = \{\ldots, a, \ldots, b, \ldots\}$$

eine wohlgeordnete Menge vom Typus α, so sind ja nach Definition die Zahlen $< \alpha$ den Abschnitten von A, also den Elementen von A, eineindeutig und ähnlich zugeordnet: jedes a bestimmt seinen Abschnitt P_a vom Typus π_a, und ist $a < b$, so ist P_a Abschnitt von P_b, $\pi_a < \pi_b$. Also ist

$$W(\alpha) = \{\ldots, \pi_a, \ldots, \pi_b, \ldots\},$$

womit die letzte Behauptung bewiesen ist. Umgekehrt ermöglicht dies, die in § 12, (1) begonnene Bezeichnung allgemein durchzuführen und die Elemente einer wohlgeordneten Menge vom Typus α durch ihre Zuordnung zu den Zahlen von

$$W(\alpha) = \{0, 1, \ldots, \xi, \ldots\} \qquad (\xi < \alpha)$$

derart zu numerieren, daß in

$$A = \{a_0, a_1, \ldots, a_\xi, \ldots\} \qquad (\xi < \alpha)$$

der Index jedes Elements der Typus des zugehörigen Abschnitts ist. Wir müssen dabei nur die Ordnungszahl 0 als Typus der Nullmenge mitrechnen, die dem ersten Element als Abschnitt zugehört. Also z. B.

$$W(1) = \{0\}, \quad W(2) = \{0, 1\}, \quad W(n) = \{0, 1, \ldots, n-1\}$$

für endliches $n > 0$, während $W(0) = 0$ die Nullmenge ist.

Nun seien α, β zwei Ordnungszahlen, $A = W(\alpha)$, $B = W(\beta)$ und $D = AB$ ihr Durchschnitt, also die Menge der Ordnungszahlen, die gleichzeitig $< \alpha$ und $< \beta$ sind. D ist wohlgeordnet, ihr Typus δ eine Ordnungszahl; wir behaupten, daß $\delta \leqq \alpha$. Für $D = A$ ist $\delta = \alpha$; für $D < A$ ist aber in

$$A = D + (A - D)$$

D ein *Anfangsstück*, $A - D$ ein Endstück von A. Denn für $\xi \varepsilon D$, $\eta \varepsilon A - D$ sind ξ, η als Elemente von A vergleichbar, also $\xi \lessgtr \eta$; nun kann nicht $\eta < \xi < \alpha, \beta$ sein, weil sonst η zu D gehören würde, also $\xi < \eta$. Dann ist aber D ein *Abschnitt* von A und $\delta < \alpha$; übrigens ist dann offenbar δ das erste Element von $A - D$ und $D = W(\delta)$. Somit haben wir

$$\delta \leqq \alpha, \quad \delta \leqq \beta.$$

Hier ist noch die Kombination $\delta < \alpha$, $\delta < \beta$ auszuschließen, weil dann doch $\delta \varepsilon D$ wäre, und es bleiben also nur die Fälle

$$\begin{aligned} \delta = \alpha, \quad \delta = \beta &: \alpha = \beta, \\ \delta = \alpha, \quad \delta < \beta &: \alpha < \beta, \\ \delta < \alpha, \quad \delta = \beta &: \alpha > \beta. \end{aligned}$$

Damit ist bewiesen:

III. (Vergleichbarkeitssatz.) *Zwei Ordnungszahlen α, β sind stets vergleichbar, d. h. es besteht eine und nur eine der drei Relationen*

$$\alpha < \beta, \quad \alpha = \beta, \quad \alpha > \beta.$$

Ist insbesondere $A \leqq B$, so ist $\alpha \leqq \beta$. Denn für $\alpha > \beta$ wäre B einem (durch $a \varepsilon A$ bestimmten) Abschnitt P von A ähnlich, und bei der Abbildung von B auf P würde das Bild von a in P fallen, also $< a$ sein, im Widerspruch zu Satz I. Es kann aber für $A < B$ trotzdem $\alpha = \beta$ sein, natürlich nur wenn A kein Abschnitt von B ist; z. B. haben die unendlichen Teilmengen der natürlichen Zahlenreihe alle den Typus ω.

Mit dem Wohlordnungs- und Vergleichbarkeitssatz ist nun auch die Lücke in der Theorie der Kardinalzahlen ausgefüllt: *zwei Kardinalzahlen* [33] *sind stets vergleichbar.* Denn $\mathfrak{a}$, $\mathfrak{b}$ können als Mächtigkeiten *wohlgeordneter* Mengen A, B mit den Ordnungstypen α, β aufgefaßt werden, und dann ist

$$\begin{array}{rll} \text{entweder} & \alpha = \beta, & \mathfrak{a} = \mathfrak{b} \\ \text{oder} & \alpha < \beta, & \mathfrak{a} \leqq \mathfrak{b} \\ \text{oder} & \alpha > \beta, & \mathfrak{a} \geqq \mathfrak{b}; \end{array}$$

denn $\alpha < \beta$ heißt: A ist einem Abschnitt von B ähnlich, A ist einer Teilmenge von B äquivalent. Die Umkehrung gibt

$$\begin{array}{rll} \text{entweder} & \mathfrak{a} = \mathfrak{b}, & \alpha \lesseqgtr \beta \\ \text{oder} & \mathfrak{a} < \mathfrak{b}, & \alpha < \beta \\ \text{oder} & \mathfrak{a} > \mathfrak{b}, & \alpha > \beta, \end{array}$$

wovon die erste Gleichung sagt, daß bei gleicher Mächtigkeit noch verschiedene Wohlordnungen möglich sind (z. B. bei $\mathfrak{a} = \aleph_0 : \alpha = \omega$, $\omega + 1, \ldots$).

IV. *In jeder* (nicht leeren) *Menge von Ordnungszahlen gibt es eine kleinste, jede Menge von Ordnungszahlen ist also der Größe nach wohlgeordnet.*

Denn ist W eine solche Menge, α eine Zahl von W und noch nicht die kleinste, so ist der Durchschnitt $W\,W(\alpha)$ als Teilmenge von $W(\alpha)$ wohlgeordnet und seine kleinste Zahl auch die kleinste in W.

Ist die Menge W vom Typus β, so kann sie also in der Gestalt

$$W = \{\alpha_0, \alpha_1, \ldots, \alpha_\eta, \ldots\} \qquad (\eta < \beta)$$

geschrieben werden, wo für $\xi < \eta$ auch $\alpha_\xi < \alpha_\eta$.

V. *Zu jeder Menge von Ordnungszahlen gibt es größere, insbesondere eine bestimmte nächstgrößere Ordnungszahl.*

Denn man wähle (S. 34) eine Mächtigkeit $\mathfrak{a}$ größer als die Mächtigkeiten aller Ordnungszahlen in W; ist α eine Ordnungszahl von der Mächtigkeit $\mathfrak{a}$, so ist α größer als alle Ordnungszahlen in W, kurz $\alpha > W$. Die kleinste Zahl $> W$ ist dann entweder α selbst oder eine gewisse Zahl aus $W(\alpha)$.

Nach V ist der Begriff „Menge aller Ordnungszahlen" undenkbar (vgl. S. 34).

Die Zahlen $> \alpha$ sind die Zahlen $\alpha + \beta$ $(\beta > 0)$ und umgekehrt (A einem Abschnitt von $A + B$ ähnlich); die kleinste Zahl $> \alpha$ ist $\alpha + 1$.

Eine Zahl $\lambda > 0$, die keinen unmittelbaren Vorgänger hat, d. h. für die $W(\lambda)$ kein letztes Element hat, heißt eine *Limeszahl*; die niedrigsten Limeszahlen sind ω, $\omega + \omega = \omega 2$, $\omega 3, \ldots$ Eine Zahl, die nicht Limeszahl ist, heißt *isoliert*; isolierte Zahlen sind außer 0 die Zahlen der Form $\alpha + 1$. Hat die Menge

$$W = \{\alpha_0, \ldots, \alpha_\eta, \ldots\} \qquad (\eta < \beta)$$

von Ordnungszahlen kein letztes Element, d. h. ist β Limeszahl, so wird die *nächstgrößere* Zahl $\lambda > W$, die offenbar Limeszahl ist, der Limes von W genannt und mit

$$\lambda = \lim W$$

oder auch mit

$$\lambda = \lim \alpha_\eta$$

bezeichnet. Z. B. ist ω der Limes von $\{0, 1, 2, \ldots\}$ oder von jeder wachsenden Folge $\{\alpha_0, \alpha_1, \alpha_2, \ldots\}$ endlicher Zahlen α_ν, $\omega = \lim \nu = \lim \alpha_\nu$.

Transfinite Induktion. An Stelle des Schlusses von n auf $n+1$ im endlichen Zahlengebiet tritt jetzt:

Eine Aussage $f(\alpha)$ über die Ordnungszahl α ist für jedes α richtig, sobald $f(0)$ richtig ist und sobald aus der Richtigkeit aller $f(\xi)$ für $\xi < \alpha$ die Richtigkeit von $f(\alpha)$ folgt.

Denn wäre $f(\beta)$ falsch und $\alpha (\leqq \beta)$ die kleinste Zahl, für die $f(\alpha)$ falsch ist, so kommen wir sowohl für $\alpha = 0$ als für $\alpha > 0$ auf einen Widerspruch.

Wie zu Beweisen, so findet diese transfinite Induktion auch zu Definitionen Verwendung:

Eine Funktion $f(\alpha)$ der Ordnungszahl α ist für jedes α definiert, sobald $f(0)$ definiert ist und sobald vermöge der Definition aller $f(\xi)$ für $\xi < \alpha$ auch $f(\alpha)$ definiert ist.

Ersetzt man hierin $f(0)$ durch irgendein $f(\alpha_0)$, so ist eine kleine Abänderung des Wortlauts nötig, aus der hervorgeht, daß $f(\alpha)$ für $\alpha \geqq \alpha_0$ richtig resp. definiert ist.

§ 14. Verknüpfungen von Ordnungszahlen.

Summe und Produkt ist bereits für Ordnungstypen definiert. Man erkennt leicht, daß eine Summe geordneter Mengen

$$S = \sum_m^M A_m$$

wohlgeordnet ist, sobald M und die Summanden A_m wohlgeordnet sind.

Also: *eine wohlgeordnete Summe von Ordnungszahlen und ein Produkt endlich vieler Ordnungszahlen ist selbst eine Ordnungszahl.*

Z. B. sind ω^2, ω^3, $\omega + \omega^2 + \omega^3 + \cdots$ Ordnungszahlen.

Hier läßt sich nun in gewissem Umfang auch Subtraktion und Division erklären. Wir schicken folgende Ungleichungen voraus:

Aus $\alpha < \beta$ folgt

$$(1)\qquad \begin{cases} \mu + \alpha < \mu + \beta, & \alpha + \mu \leqq \beta + \mu \\ \mu\alpha < \mu\beta\,(\mu > 0), & \alpha\mu \leqq \beta\mu. \end{cases}$$

Denn $\alpha < \beta$ (A Abschnitt von B, wie angenommen werden kann) ist so viel wie $\beta = \alpha + \gamma$ $(\gamma > 0)$ und vice versa. Demnach:

$$\mu + \beta = \mu + (\alpha + \gamma) = (\mu + \alpha) + \gamma > \mu + \alpha,$$
$$\mu\beta = \mu(\alpha + \gamma) = \mu\alpha + \mu\gamma > \mu\alpha \quad \text{für} \quad \mu > 0.$$

Bei Nachsetzung des Summanden oder Faktors μ kann das Gleichheitszeichen auftreten. Z. B.

$$\omega + 1 < \omega + 2, \qquad 1 + \omega = 2 + \omega = \omega,$$
$$\omega 1 < \omega 2, \qquad 1\,\omega = 2\,\omega = \omega.$$

Die Umkehrung von (1) liefert folgende Schlüsse:

$$(2)\qquad \begin{cases} \text{Aus } \mu + \alpha < \mu + \beta \text{ oder } \alpha + \mu < \beta + \mu & \text{folgt } \alpha < \beta \\ \text{Aus } \mu + \alpha = \mu + \beta & \text{folgt } \alpha = \beta \\ \text{Aus } \mu\alpha < \mu\beta \text{ oder } \alpha\mu < \beta\mu & \text{folgt } \alpha < \beta \\ \text{Aus } \mu\alpha = \mu\beta \text{ und } \mu > 0 & \text{folgt } \alpha = \beta. \end{cases}$$

Aus $\alpha + \mu = \beta + \mu$ oder $\alpha\mu = \beta\mu$ folgt nicht $\alpha = \beta$.

Subtraktion. Durch α und $\beta > \alpha$ ist, wie wir soeben sahen, eine der Gleichung

$$\alpha + \xi = \beta$$

genügende Zahl ξ stets und eindeutig bestimmbar, wir bezeichnen sie mit

$$\xi = -\alpha + \beta,$$

so daß $\alpha + (-\alpha + \beta) = \beta$. ξ ist der Typus von $W(\beta) - W(\alpha)$, also eines *Restes* (S. 58) von $W(\beta)$, kurz ein *Resttypus* von β. Bei festem β ist übrigens für $\alpha < \alpha_1 < \beta$ offenbar $\xi \geqq \xi_1$; die verschiedenen Resttypen bilden also eine *invers* wohlgeordnete, zugleich aber als Ordnungszahlen eine wohlgeordnete Menge, d. h. *es gibt nur endlich viele verschiedene Resttypen einer Ordnungszahl.*

Z. B. hat ω nur den einen Resttypus ω; $\omega + 3$ hat die Resttypen $\omega + 3$, 3, 2, 1, den Zerlegungen (ν endlich)

$$\omega + 3 = \nu + (\omega + 3) = \omega + 3 = (\omega + 1) + 2 = (\omega + 2) + 1$$

entsprechend.

Dagegen ist für $\beta > \alpha$ die Gleichung

$$\eta + \alpha = \beta$$

nach η nicht immer auflösbar (α muß ein Resttypus von β sein): z. B. ist $\eta + \omega = \omega + 1$ unlösbar, da der linksstehende Typus einer Menge ohne letztes Element, der rechtsstehende einer Menge mit letztem Element entspricht. Wenn die Gleichung auflösbar ist, so hat sie für $\alpha \geqq \omega$ immer unendlich viele Lösungen $\eta = \eta_0, \eta_0 + 1, \eta_0 + 2, \ldots$, dagegen für endliches α nur eine einzige; sie besagt dann nämlich, daß $\eta + (\alpha - 1)$ der unmittelbare Vorgänger von β, $\eta + (\alpha - 2)$ der Vorgänger von $\eta + (\alpha - 1)$ sein soll usw., wodurch nach endlich vielen Schritten η eindeutig bestimmt ist. Nur in diesem Falle bezeichnen wir die Lösung mit

$$\eta = \beta - \alpha,$$

so daß $(\beta - \alpha) + \alpha = \beta$ ist; diese Gleichung bedeutet also, daß α eine natürliche Zahl ist und η aus β durch Weglassung der α letzten Elemente entsteht. Z. B. ist $\beta - 1$ der Vorgänger von β (β als isolierte Zahl > 0 vorausgesetzt).

Division. Jede Zahl $\zeta < \alpha\beta$ ist in der Form

$$(3) \qquad \zeta = \alpha\eta + \xi \quad (\xi < \alpha,\ \eta < \beta)$$

darstellbar, wo ξ und η durch α, β, ζ eindeutig bestimmt sind.

Denn sind A, B wohlgeordnet von den Typen α, β, so ist $\alpha\beta$ der Typus des Produkts (B, A), der lexikographisch geordneten Menge der Paare (b, a). ζ ist Typus eines Abschnittes von (B, A), der durch das Element (b, a) bestimmt sei; dieser Abschnitt wird gebildet durch die Paare (y, x) mit $y < b$, $x \varepsilon A$ und (b, x) mit $x < a$. Sind also ξ, η die Typen der durch a, b bestimmten Abschnitte von A, B, so folgt die obige Formel; man sieht zugleich, daß (b, a) und ξ, η durch ζ, α, β bestimmt sind. Läßt man β beliebig, so kann man sagen:

Jede Zahl ist, $\alpha > 0$ vorausgesetzt, in der Form

$$(4) \qquad \zeta = \alpha\eta + \xi \quad (\xi < \alpha)$$

darstellbar, wo ξ und η durch α, ζ eindeutig bestimmt sind.

Denn man kann β so groß wählen (z. B. $\beta = \zeta + 1$), daß $\zeta < \alpha\beta$, und dann (3) anwenden; es können auch nicht die zu zwei verschiedenen β gehörigen Darstellungen (4) verschieden sein, da es sonst für das größere dieser beiden β zwei verschiedene Darstellungen (3) gäbe.

Wir haben hier also ein (einseitiges) Analogon der Verhältnisse im Endlichen: η ist sozusagen der Quotient, ξ der Rest der Division von ζ durch α; für $\xi = 0$ ist ζ durch α als linken Faktor teilbar. Auch der Euklidische Algorithmus läßt sich übertragen:

$$\begin{aligned} \alpha &= \alpha_1\eta_1 + \alpha_2 \quad (\alpha_1 > \alpha_2) \\ \alpha_1 &= \alpha_2\eta_2 + \alpha_3 \quad (\alpha_2 > \alpha_3) \\ &\dots \end{aligned}$$

mit abnehmenden, daher nur endlich vielen Resten, unter denen also schließlich die Null auftreten muß. Das führt dazu, die *geordneten Paare* von Ordnungszahlen in Kettenbrüche zu entwickeln und analog wie die gewöhnlichen Rationalzahlen zu ordnen; wir gehen darauf nicht ein.

Ausdehnung der Multiplikation. Wir haben schon S. 48 von Produkten mit unendlich vielen Faktoren gesprochen, etwa $\dots \alpha_3\,\alpha_2\,\alpha_1$; diese sind aber nicht gleichzeitig mit den Faktoren immer Ordnungszahlen.

z. B. war $\ldots\omega\omega\omega = 1 + \lambda$ der Typus des Intervalls $[0, 1)$. Wir werden jetzt Produkte mit wohlgeordneter Reihenfolge der Faktoren, wie $\alpha_1\alpha_2\alpha_3\ldots$, erklären; die jetzige Definition hat mit der damaligen zunächst gar nichts zu schaffen, und erst eine spätere Betrachtung (§ 16) wird uns lehren, daß beide doch Spezialfälle eines allgemeinen Produktbegriffs sind.

Unsere gegenwärtige Erklärung beruht auf transfiniter Induktion (S. 62); der Symmetrie wegen wollen wir auch die — schon einfacher und in weiterem Umfang, nämlich für Ordnungs*typen*, erklärte — Addition noch einmal auf diese Weise definieren, d. h. auf Addition von zwei Summanden zurückführen.

Jeder Ordnungszahl α sei eine Ordnungszahl μ_α zugeordnet. Wir definieren die — bei festgewählten Summanden nur noch als Funktion von α anzusehende — Summe

$$f(\alpha) = \sum_{\xi}^{W(\alpha)} \mu_\xi = \mu_0 + \mu_1 + \cdots + \mu_\xi + \cdots$$

als Ordnungszahl durch folgende induktive Vorschrift:

$$(5)\quad \begin{cases} f(0) = 0\,; \\ \text{für } \alpha > 0 \text{ sei } f(\alpha) \text{ die kleinste Zahl} \geqq f(\xi) + \mu_\xi \quad (\text{für alle } \xi < \alpha). \end{cases}$$

Nehmen wir der Einfachheit wegen alle $\mu_\alpha > 0$ (Summanden $= 0$ wären wegzulassen), so ist für $\alpha > \xi$

$$f(\alpha) + \mu_\alpha > f(\alpha) \geqq f(\xi) + \mu_\xi > f(\xi),$$

die Zahlen $f(\alpha)$ wie $f(\alpha) + \mu_\alpha$ haben die Ordnung ihrer Argumente α. Speziell ist also

$$(6)\qquad f(\alpha + 1) = f(\alpha) + \mu_\alpha.$$

Ist ferner α eine Limeszahl, so ist

$$(7)\qquad f(\alpha) = \lim f(\xi) \quad (\xi < \alpha);$$

es ist dann nämlich mit $\xi < \alpha$ auch $\xi + 1 < \alpha$ und die erste Zahl $\geqq f(\xi + 1)$ mit der ersten Zahl $> f(\xi)$ identisch, also $f(\alpha)$ mit $\lim f(\xi)$ identisch (es sind (5) und (6) zu vergleichen).

Daß diese jetzige Erklärung der Summe mit der früheren identisch ist, d. h. daß der frühere Summenbegriff die Eigenschaften (5) oder (6)(7) hat, ist leicht einzusehen. Nach (7) ist die Summe der Limes der Partialsummen, wie bei den konvergenten Reihen der Analysis, z. B.

$$f(\omega) = \lim f(\nu), \qquad \mu_0 + \mu_1 + \mu_2 + \cdots = \lim\,(\mu_0 + \mu_1 + \cdots + \mu_{\nu-1}).$$

In gleicher Weise verfahren wir für das Produkt, das wir auf Produkte zweier Faktoren zurückführen. Jeder Ordnungszahl α sei eine Ordnungszahl $\mu_\alpha > 0$ zugeordnet (Produkte mit einer Null als Faktor sollen selbst gleich Null sein); wir definieren das als Funktion von α anzusehende Produkt

$$f(\alpha) = \prod_{\xi}^{W(\alpha)} = \mu_0\mu_1\ldots\mu_\xi\ldots$$

als Ordnungszahl durch die Vorschrift:

$$(8)\quad \begin{cases} f(0) = 1\,; \\ \text{für } \alpha > 0 \text{ sei } f(\alpha) \text{ die kleinste Zahl } \geqq f(\xi)\mu_\xi \quad (\text{für alle } \xi < \alpha). \end{cases}$$

Nehmen wir der Einfachheit wegen alle Faktoren > 1 an (Einsen wären wegzulassen), so ist zunächst klar, daß $f(\alpha) > 0$,

$$f(\alpha)\,\mu_\alpha > f(\alpha) \geqq f(\xi)\,\mu_\xi > f(\xi)$$

und man erhält wie zuvor

$$(9)\qquad f(\alpha+1) = f(\alpha)\,\mu_\alpha$$

und für eine Limeszahl α

$$(10)\qquad f(\alpha) = \lim f(\xi) \qquad (\xi < \alpha).$$

Die Produkte mit endlichem α stimmen mit unseren früheren überein,

$$f(1) = \mu_0, \quad f(2) = \mu_0\mu_1, \ldots,$$

sodann ist $\quad f(\omega) = \lim f(\nu), \quad \mu_0\mu_1\mu_2 \ldots = \lim \mu_0\mu_1 \ldots \mu_{\nu-1}$,

z. B. [1] $\quad 2.\,3.\,4 \ldots = \lim\{2, 6, 24, \ldots\} = \omega$.

Speziell wenn alle Faktoren $\mu_\alpha = \mu > 1$, definieren wir das Produkt als die *Potenz* $f(\alpha) = \mu^\alpha$. Es ist also

$$(11)\qquad \mu^{\alpha+1} = \mu^\alpha \,.\, \mu$$

und für eine Limeszahl α

$$(12)\qquad \mu^\alpha = \lim \mu^\xi \qquad (\xi < \alpha)$$

z. B. $\quad 2^\omega = \lim 2^\nu = \lim\{2, 4, 8, \ldots\} = \omega$,

allgemein $\quad 2^\omega = 3^\omega = 4^\omega = \cdots = \omega$,

dagegen $\quad \omega^\omega = \lim \omega^\nu = \lim\{\omega, \omega^2, \omega^3, \ldots\}$,

wofür man wegen $1 + \omega = \omega$, $1 + \omega + \omega^2 = \omega(1+\omega) = \omega^2$ usw. auch schreiben kann

$$\omega^\omega = 1 + \omega + \omega^2 + \cdots = \Sigma\omega^\nu.$$

Es gelten die Potenzregeln

$$(13)\qquad \mu^\alpha\mu^\beta = \mu^{\alpha+\beta}, \quad (\mu^\alpha)^\beta = \mu^{\alpha\beta},$$

die man am schnellsten durch Induktion beweist (sie sind für β richtig, wenn sie für $\eta < \beta$ richtig sind). Ein kommutatives Gesetz kann natürlich nicht gelten; $(\mu\nu)^2 = \mu\nu\,\mu\nu$ ist von $\mu^2\nu^2 = \mu\mu\,\nu\nu$ im allgemeinen verschieden.

Wir betonen nochmals, daß die hier definierten Produkte und Potenzen mit den früheren zunächst gar keinen Zusammenhang haben; es sind im allgemeinen nicht Typen von Mengenprodukten. Daher hat z. B. α^β nicht notwendig die Mächtigkeit $\mathfrak{a}^\mathfrak{b}$ (während $\alpha + \beta$, $\alpha\beta$ die Mächtigkeiten $\mathfrak{a} + \mathfrak{b}$, $\mathfrak{a}\mathfrak{b}$ haben); $2^\omega = \omega$ hat nur die Mächtigkeit $\aleph_0$, nicht $2^{\aleph_0}$.

[1]) In Kardinalzahlen ist $2\,.\,3\,.\,4 \ldots = 2^{\aleph_0}$; die Verwendung der endlichen Zahlen in beiden Bedeutungen könnte hier besonders bedenklich erscheinen und ist es doch nicht.

Wie eine natürliche Zahl durch die Potenzen von 10, so läßt sich jede Ordnungszahl durch die Potenzen einer beliebigen *Basis* $\beta > 1$ ausdrücken. Sei $\zeta > 0$ eine Ordnungszahl und β^γ die niedrigste Potenz von β, die $> \zeta$ ist (daß es solche gibt, erkennt man aus der induktiv leicht zu beweisenden Ungleichung $\beta^\gamma \geqq \gamma$, wonach $\beta^{\zeta+1} > \zeta$ ist). Dann ist γ keine Limeszahl, sonst wäre für jedes $\xi < \gamma$ auch $\xi + 1 < \gamma$, also

$$\beta^{\xi+1} \leqq \zeta, \quad \beta^\xi < \zeta, \quad \beta^\gamma = \lim \beta^\xi \leqq \zeta.$$

Es hat also $\gamma (> 0)$ einen unmittelbaren Vorgänger α und es ist

$$\beta^\alpha \leqq \zeta < \beta^{\alpha+1},$$

wobei α durch ζ (bei fester Basis) eindeutig bestimmt ist. Die Zahl $\zeta < \beta^\alpha \cdot \beta$ läßt sich nach (3) in der Form

$$\zeta = \beta^\alpha \eta + \zeta_1 \qquad (\eta < \beta, \ \zeta_1 < \beta^\alpha)$$

darstellen, η und ζ_1 durch ζ bestimmt. Ist noch $\zeta_1 > 0$, so erhalten wir weiter

$$\zeta_1 = \beta^{\alpha_1} \eta_1 + \zeta_2 \qquad (\eta_1 < \beta, \ \zeta_2 < \beta^{\alpha_1})$$

usw. Da aber $\zeta \geqq \beta^\alpha > \zeta_1 \geqq \beta^{\alpha_1} > \zeta_2 \geqq \cdots$, also $\zeta > \zeta_1 > \zeta_2 > \cdots$, $\alpha > \alpha_1 > \alpha_2 > \cdots$, so muß das Verfahren einmal mit dem Rest 0

$$\zeta_n = \beta^{\alpha_n} \eta_n$$

endigen und wir haben die Darstellung

$$(14) \qquad \zeta = \beta^\alpha \eta + \beta^{\alpha_1} \eta_1 + \cdots + \beta^{\alpha_n} \eta_n \qquad \begin{pmatrix} \alpha > \alpha_1 > \cdots > \alpha_n \geqq 0 \\ 0 < \eta, \eta_1, \ldots, \eta_n < \beta \end{pmatrix},$$

worin alles durch ζ eindeutig bestimmt ist: die Gliederanzahl $n + 1$ (für $n = 0 : \zeta = \beta^\alpha \eta$), die Exponenten α und die Koeffizienten η. Und zwar ist nicht nur die soeben konstruierte, sondern jede, gleichviel auf welchem Wege gewonnene, Darstellung von der Form (14) eindeutig bestimmt. Denn der Ausdruck (14) ist $< \beta^{\alpha+1}$, wie man durch den Schluß von n auf $n + 1$ erkennt: es ist dann nämlich $\zeta < \beta^\alpha \eta + \beta^{\alpha_1+1} \leqq \beta^\alpha(\eta + 1) \leqq \beta^{\alpha+1}$. Also muß $\beta^\alpha \leqq \zeta < \beta^{\alpha+1}$ sein und die Exponenten wie die Koeffizienten bestimmen sich genau wie oben.

Beispiele. $\quad \beta = 2: \quad \zeta = 2^\alpha + 2^{\alpha_1} + \cdots + 2^{\alpha_n}.$

$$\beta = \omega: \quad (15) \quad \zeta = \omega^\alpha \nu + \omega^{\alpha_1} \nu_1 + \cdots + \omega^{\alpha_n} \nu_n$$
$$(\nu, \ldots, \nu_n \text{ natürliche Zahlen}).$$

Insbesondere läßt sich jede Zahl $\zeta < \beta^\omega$ als *Polynom* in β (mit endlichen Exponenten) darstellen, und zwar in der Form (14) nur auf eine Weise.

Bei der Darstellung (14) kann es übrigens vorkommen, daß ζ gar nicht durch kleinere Zahlen ausgedrückt wird, sondern die Gleichung

$$(16) \qquad \zeta = \beta^\zeta$$

5*

besteht (die im endlichen Zahlengebiet für $\beta > 1$ ja unmöglich ist); so hatten wir $\omega = 2^\omega$ gefunden. Definiert man für $\nu = 0, 1, 2, \ldots$ die Zahlen ζ_ν durch $\zeta_0 = 1$, $\zeta_{\nu+1} = \beta^{\zeta_\nu}$, so ist $(\beta > 1)$

$$\begin{aligned} &\zeta_0 < \zeta_1, \quad \text{also } \beta^{\zeta_0} < \beta^{\zeta_1}, \\ &\zeta_1 < \zeta_2, \quad \text{also } \beta^{\zeta_1} < \beta^{\zeta_2}, \\ &\zeta_2 < \zeta_3 \quad \text{usw.} \end{aligned}$$

Für $\zeta = \lim \zeta_\nu$ ist

$$\zeta = \lim \zeta_{\nu+1} = \lim \beta^{\zeta_\nu} = \beta^\zeta,$$

also ζ eine Zahl mit der Eigenschaft (16); sie ist der Limes von $1, \beta, \beta^\beta, \beta^{\beta^\beta}, \ldots$ (Zahlen, für die $\zeta = \omega^\zeta$, heißen bei Cantor ε-Zahlen).

Die Potenzen ω^α von ω sind durch die Eigenschaft charakterisiert, daß sie sich selbst als einzigen Resttypus haben (S. 63). In der Tat geht aus (15) hervor: wenn ζ keinen Resttypus $< \zeta$ hat, so muß die Gliederanzahl der Entwicklung 1 sein, da sonst $\omega^{\alpha_n} < \zeta$ ein Resttypus wäre, also $\zeta = \omega^\alpha \nu$, und zwar $\nu = 1$, da sonst $\omega^\alpha < \zeta$ ein Resttypus wäre, mithin $\zeta = \omega^\alpha$. Umgekehrt: ω^α hat sich selbst als einzigen Resttypus, d. h. es ist

$$\eta + \omega^\alpha = \omega^\alpha \qquad (\eta < \omega^\alpha).$$

Denn ist $(\eta > 0)$ $\eta = \omega^\beta \nu + \eta_1$ (ν natürliche Zahl, $\eta_1 < \omega^\beta$) der Anfang der Entwicklung von η, wobei $\beta < \alpha$ ist, und setzt man $\omega^\alpha = \omega^{\beta+1} + \varrho$ (für $\beta + 1 = \alpha$ also $\varrho = 0$), so ist

$$\begin{aligned} \omega^\alpha &\leqq \eta + \omega^\alpha \leqq \omega^\beta(\nu + 1) + \omega^\alpha \\ &= \omega^\beta(\nu + 1 + \omega) + \varrho = \omega^{\beta+1} + \varrho = \omega^\alpha, \end{aligned}$$

also $\eta + \omega^\alpha = \omega^\alpha$.

Da ein Resttypus eines Resttypus von ζ wieder ein Resttypus von ζ ist, so folgt noch, daß der kleinste Resttypus einer Zahl ζ immer eine Potenz von ω ist (eventuell $\omega^0 = 1$); in der Entwicklung (15) ist ω^{α_n} der kleinste Resttypus von ζ.

Natürliche Summen und Produkte. Die Entwicklung (15) stellt eine Ordnungszahl als eine Art Polynom in ω dar, im allgemeinen mit unendlichen Exponenten. Wenn man mit diesen Polynomen wie mit gewöhnlichen Polynomen rechnet, so erhält man mit G. Hessenberg *natürliche Summen* $\sigma(\xi, \eta)$ und *natürliche Produkte* $\pi(\xi, \eta)$ von Ordnungszahlen: Bildungen, die denen für endliche Zahlen wesentlich näher stehen als $\xi + \eta$ und $\xi\eta$.

Schreiben wir die Entwicklung nach Potenzen von ω in der einfachen Form

$$\xi = \sum_\alpha \omega^\alpha x_\alpha = \cdots + \omega^\omega x_\omega + \cdots + \omega^2 x_2 + \omega x_1 + x_0,$$

indem wir den Exponenten α *alle* Ordnungszahlen unterhalb einer genügend groß gewählten Schranke durchlaufen lassen und die Koeffizienten x_α als endliche (ganze) Zahlen $\geqq 0$ annehmen; in Wahrheit sind nur endlich viele

Koeffizienten von 0 verschieden. (Für $\xi = 0$ sind alle $x_\alpha = 0$.) Die Darstellung ist eindeutig bestimmt. Für zwei Ordnungszahlen

$$\xi = \sum_\alpha \omega^\alpha x_\alpha, \qquad \eta = \sum_\alpha \omega^\alpha y_\alpha$$

definieren wir dann

$$\sigma(\xi, \eta) = \sum_\alpha \omega^\alpha (x_\alpha + y_\alpha) = \sigma(\eta, \xi).$$

Dies braucht weder mit $\xi + \eta$ noch mit $\eta + \xi$ übereinzustimmen; z. B. ist $\sigma(\omega, \omega^2 + 1) = \omega^2 + \omega + 1$ von $\omega + (\omega^2 + 1) = \omega^2 + 1$ und von $(\omega^2 + 1) + \omega = \omega^2 + \omega$ verschieden.

Bei gegebenem ζ hat die Gleichung $\sigma(\xi, \eta) = \zeta$ nur endlich viele Lösungen ξ, η. Denn es muß $x_\alpha + y_\alpha = z_\alpha$ sein, und für x_α sind nur die Werte $0, 1, \ldots, z_\alpha$, zulässig; die Anzahl aller Lösungen ist das Produkt aller Faktoren $1 + z_\alpha$, von denen nur endlich viele > 1 sind.

Die Ungleichung $\xi < \eta$ bedeutet, daß die *erste* nicht verschwindende Differenz $y_\alpha - x_\alpha$ positiv ist (die erste in der Ordnung der Summe nach fallenden Exponenten, also die zum höchsten Exponenten gehörige). D. h. es gibt eine Ordnungszahl $\beta \geqq 0$ mit $x_\beta < y_\beta$, $x_\gamma = y_\gamma$ für $\gamma > \beta$. Hieraus geht hervor, daß $\sigma(\xi, \eta)$ mit jedem seiner Summanden wächst: für $\xi_0 < \xi$ ist $\sigma(\xi_0, \eta) < \sigma(\xi, \eta)$. Ist also $\sigma(\xi_0, \eta_0) = \sigma(\xi, \eta)$ und $\xi_0 < \xi$, so ist zugleich $\eta_0 > \eta$.

Wenn $\zeta_0 < \zeta = \sigma(\xi, \eta)$, so hat die Gleichung $\sigma(\xi_0, \eta_0) = \zeta_0$ eine Lösung mit $\xi_0 \leqq \xi$, $\eta_0 \leqq \eta$ (unter Ausschluß mindestens eines Gleichheitszeichens). Schreiben wir nämlich, unter Hervorhebung der ersten Differenzstelle β für ζ_0 und ζ,

$$\begin{aligned}
\xi &= \sum_\gamma \omega^\gamma x_\gamma + \omega^\beta x_\beta + \sum_\alpha \omega^\alpha x_\alpha, \\
\eta &= \sum_\gamma \omega^\gamma y_\gamma + \omega^\beta y_\beta + \sum_\alpha \omega^\alpha y_\alpha, \\
\zeta_0 &= \sum_\gamma \omega^\gamma c_\gamma + \omega^\beta c_\beta + \sum_\alpha \omega^\alpha c_\alpha
\end{aligned}$$

mit $\gamma > \beta > \alpha$, $x_\gamma + y_\gamma = c_\gamma$, $x_\beta + y_\beta > c_\beta$. Wir bestimmen dann zwei ganze Zahlen a_β, b_β mit $0 \leqq a_\beta \leqq x_\beta$, $0 \leqq b_\beta \leqq y_\beta$, $a_\beta + b_\beta = c_\beta$; etwa $a_\beta = \min [x_\beta, c_\beta]$ und $b_\beta = c_\beta - a_\beta$. Dabei ist mindestens eine der Ungleichungen $a_\beta < x_\beta$, $b_\beta < y_\beta$ erfüllt; wenn etwa die erste, so setze man

$$\begin{aligned}
\xi_0 &= \sum_\gamma \omega^\gamma x_\gamma + \omega^\beta a_\beta + \sum_\alpha \omega_\alpha c_\alpha, \\
\eta_0 &= \sum_\gamma \omega^\gamma y_\gamma + \omega^\beta b_\beta,
\end{aligned}$$

dann ist $\sigma(\xi_0, \eta_0) = \zeta_0$, $\xi_0 < \xi$, $\eta_0 \leqq \eta$.

Die natürlichen Summen verhalten sich also ganz wie endliche; wir werden davon mehrfach Gebrauch machen.

Das natürliche Produkt erhalten wir, indem wir

$$\xi = \sum_\alpha \omega^\alpha x_\alpha, \qquad \eta = \sum_\beta \omega^\beta y_\beta$$

wie Polynome multiplizieren und dabei die Exponenten *natürlich* addieren,

also $$\pi(\xi, \eta) = \sum_{\alpha\beta} \omega^{\sigma(\alpha,\beta)} x_\alpha y_\beta = \pi(\eta, \xi)$$

oder $$\pi(\xi, \eta) = \sum_\gamma \omega^\gamma z_\gamma, \quad z_\gamma = \overset{\gamma}{\Sigma} x_\alpha y_\beta,$$

$\overset{\gamma}{\Sigma}$ über die (endlich vielen) Paare mit $\sigma(\alpha, \beta) = \gamma$ erstreckt. Wir wollen dies nur erwähnen und dem Leser überlassen, auch hier die Analogie mit endlichen Produkten festzustellen.

§ 15. Die Alefs.

Alle Kardinalzahlen können als Mächtigkeiten wohlgeordneter Mengen aufgefaßt werden (Wohlordnungssatz) und sind daher vergleichbar. Als *Zahlenklasse* $Z(\mathfrak{a})$ bezeichnen wir die Menge aller *Ordnungszahlen* α, die die Mächtigkeit $\mathfrak{a}$ haben; sie ist Teilmenge der entsprechenden *Typenklasse* $T(\mathfrak{a})$ (S. 49). Für endliches $n = 0, 1, 2, \ldots$ besteht $Z(n)$ nur aus der einen Ordnungszahl n; aus $Z(\aleph_0)$ kennen wir schon unendlich viele Vertreter

$$\omega, \omega + 1, \omega + 2, \ldots, \omega 2, \ldots, \omega 3, \ldots, \omega^2, \ldots, \omega^3, \ldots, \omega^\omega, \ldots$$

Ist $\mathfrak{a} < \mathfrak{b}$ und sind α, β irgend zwei Ordnungszahlen aus den Zahlenklassen $Z(\mathfrak{a})$, $Z(\mathfrak{b})$, so ist $\alpha < \beta$.

I. *Jede Menge von Kardinalzahlen ist der Größe nach wohlgeordnet.*

Denn ordnen wir jeder Kardinalzahl eine Ordnungszahl aus der zugehörigen Zahlenklasse zu, so wird die Menge der Kardinalzahlen mit einer Menge von Ordnungszahlen ähnlich, also wohlgeordnet (§ 13, IV).

II. *Über jeder Menge von Kardinalzahlen gibt es größere, insbesondere eine bestimmte nächstgrößere Kardinalzahl.*

Das erste wissen wir aus § 7, das andere folgt aus I; denn wenn es überhaupt eine Mächtigkeit $\mathfrak{a} > \mathfrak{K}$ gibt ($\mathfrak{K}$ eine Menge von Kardinalzahlen), so ist entweder $\mathfrak{a}$ die kleinste oder in der wohlgeordneten Menge der Kardinalzahlen, die $> \mathfrak{K}$ und $< \mathfrak{a}$ sind, befindet sich eine kleinste. (Wir sagen nicht, daß in der Menge der Kardinalzahlen $> \mathfrak{K}$ eine kleinste sei, weil diese Menge ebenso undenkbar ist wie die Menge aller Kardinalzahlen.)

Die zu einer Kardinalzahl $\mathfrak{a}$ nächstgrößere $\mathfrak{b}$ wird, wie leicht einzusehen, so erhalten, daß man zur Zahlenklasse $Z(\mathfrak{a})$ die nächstgrößere Ordnungszahl β sucht; deren Mächtigkeit ist $\mathfrak{b}$. Ob die Kardinalzahl $2^\mathfrak{a} > \mathfrak{a}$ die *nächstgrößere* nach $\mathfrak{a}$ ist, ist noch für kein unendliches $\mathfrak{a}$ bekannt; für $\mathfrak{a} = \aleph_0$ ist diese Frage das Kontinuumproblem (S. 40).

Die unendlichen Kardinalzahlen ($\geq \aleph_0$) heißen *Alefs.* Das erste unter ihnen ist $\aleph_0$, die Mächtigkeit der Menge der natürlichen Zahlen; das nächstgrößere wird $\aleph_1$, das hierauf folgende $\aleph_2$ genannt usw., das kleinste, alle $\aleph_\nu$ mit endlichem Index übertreffende wird mit $\aleph_\omega$, das nächstgrößere mit $\aleph_{\omega+1}$ bezeichnet usf. D. h. *jedes Alef* $\aleph_\alpha$ *erhält als Index den Typus der Menge aller vorangehenden Alefs.*

Z. B. ist die Mächtigkeit $\aleph$ des Kontinuums $> \aleph_0$, also $\aleph \geqq \aleph_1$; die Frage, ob hier das Gleichheitszeichen gilt oder nicht, ist das Kontinuumproblem.

Die kleinste zur Zahlenklasse $Z(\aleph_\alpha)$ gehörige Ordnungszahl heißt die *Anfangszahl* ω_α dieser Zahlenklasse. Die kleinste Anfangszahl ist $\omega_0 = \omega$, [34]
sodann folgen $\omega_1, \omega_2, \ldots, \omega_\omega, \omega_{\omega+1}, \ldots$; jede Anfangszahl ω_α hat als Index den Typus der Menge aller vorangehenden Anfangszahlen. Es ist also

$$\begin{aligned} Z(\aleph_\alpha) &= \text{Menge der Zahlen } \omega_\alpha \leqq \mu < \omega_{\alpha+1} \\ &= W(\omega_{\alpha+1}) - W(\omega_\alpha) \end{aligned}$$

oder, im Sinne der Addition geordneter Mengen

$$W(\omega_{\alpha+1}) = W(\omega_\alpha) + Z(\aleph_\alpha), \tag{1}$$

$$W(\omega_\alpha) = W(\omega_0) + \sum_{\xi<\alpha} Z(\aleph_\xi), \tag{2}$$

wo für $\alpha = 0$ die letzte Summe $= 0$ zu setzen ist und $W(\alpha)$ wie früher die Menge der Ordnungszahlen $< \alpha$ bedeutet. $W(\omega_0) = W(\omega)$ repräsentiert die Vereinigung der endlichen Zahlenklassen $Z(0), Z(1), \ldots$; übrigens wird vielfach (nach Cantor) $W(\omega)$ die erste, $Z(\aleph_0)$ die zweite, $Z(\aleph_1)$ die dritte Zahlenklasse genannt.

Jede Ordnungszahl ist von der Form $\omega\mu + \nu$ ($\nu < \omega$, d. h. ν endlich; § 14, (4)); eine Limeszahl speziell von der Form $\lambda = \omega\mu$, so daß für ihre Mächtigkeit gilt

$$\mathfrak{l} = \aleph_0 \mathfrak{m}, \quad \aleph_0 \mathfrak{l} = \aleph_0^2 \mathfrak{m} = \aleph_0 \mathfrak{m} = \mathfrak{l}.$$

Da jede Anfangszahl offenbar Limeszahl ist (eine unendliche Mächtigkeit kann sich durch Hinzufügung eines Elements nicht ändern), so ist also

$$\aleph_0 \aleph_\alpha = \aleph_\alpha$$

für jedes Alef $\aleph_\alpha$, also nach dem Äquivalenzsatz erst recht

$$\begin{aligned} 2\aleph_\alpha &= \aleph_\alpha + \aleph_\alpha = \aleph_\alpha, \\ \mathfrak{x} + \aleph_\alpha &= \aleph_\alpha \qquad \text{für } \mathfrak{x} < \aleph_\alpha. \end{aligned}$$

Ist $\mathfrak{x} < \aleph_\alpha$, $\mathfrak{y} < \aleph_\alpha$, so ist auch noch $\mathfrak{x} + \mathfrak{y} < \aleph_\alpha$; denn sei $\mathfrak{x} \leqq \mathfrak{y}$ und $\mathfrak{y} = \aleph_\eta$ mit $\eta < \alpha$ (für endliches $\mathfrak{y}$ ist nichts zu beweisen), so ist $\mathfrak{x} + \mathfrak{y} = \mathfrak{x} + \aleph_\eta = \aleph_\eta < \aleph_\alpha$. Daraus folgt weiter, daß jeder Rest von ω_α den Typus ω_α hat, denn andernfalls wäre in $\omega_\alpha = \xi + \eta$ ($\aleph_\alpha = \mathfrak{x} + \mathfrak{y}$) sowohl $\mathfrak{x}$ wie $\mathfrak{y} < \aleph_\alpha$. Aus (1) ergibt sich also:

(3) $Z(\aleph_\alpha)$ hat den Typus $\omega_{\alpha+1}$ und die Mächtigkeit $\aleph_{\alpha+1}$.

Z. B. hat $Z(\aleph_0)$ die Mächtigkeit $\aleph_1$, während die entsprechende Typenklasse $T(\aleph_0)$ die Mächtigkeit $\aleph$ des Kontinuums hatte, was wieder die Ungleichung $\aleph \geqq \aleph_1$ und das Kontinuumproblem illustriert. Nach (2) ist demgemäß

$$\left.\begin{aligned}\omega_\alpha &= \omega_0 + \sum_\xi \omega_{\xi+1} = \omega_0 + \omega_1 + \cdots + \omega_{\xi+1} + \cdots \\ \aleph_\alpha &= \aleph_0 + \sum_\xi \aleph_{\xi+1} = \aleph_0 + \aleph_1 + \cdots + \aleph_{\xi+1} + \cdots\end{aligned}\right\} \quad (\xi < \alpha)$$

z. B. $\aleph_\omega = \aleph_0 + \aleph_1 + \aleph_2 + \cdots$; $\aleph_\omega$ erfüllt, im Gegensatz zu $\aleph$, die Ungleichung $\aleph_\omega < \aleph_\omega^{\aleph_0}$ (S. 36).

Jede Potenz eines Alef mit endlichem Exponenten ist diesem Alef selbst gleich, d. h. es gilt

$$(4) \qquad \aleph_\alpha^2 = \aleph_\alpha .$$

Ordnet man nämlich die Paare (ξ, η) mit $\xi < \omega_\alpha$, $\eta < \omega_\alpha$, die eine Menge von der Mächtigkeit $\aleph_\alpha^2$ bilden, nach der natürlichen Summe (S. 69) $\sigma(\xi, \eta) = \zeta$, so ist [1] $\zeta < \omega_\alpha$, und da zu jedem ζ nur endlich viele Paare (ξ, η) gehören, so erhalten wir eine Menge von der Mächtigkeit $\leqq \aleph_0 \aleph_\alpha = \aleph_\alpha$.

Danach ist $\aleph_\alpha = \aleph_\alpha^2 = \aleph_\alpha^3 = \cdots$. Bezüglich des Exponenten $\aleph_0$ wissen wir, daß es Alefs zweier Kategorien gibt, solche mit $\aleph_\alpha^{\aleph_0} > \aleph_\alpha$ (wie $\aleph_0, \aleph_\omega$) und solche mit $\aleph_\alpha^{\aleph_0} = \aleph_\alpha$ (die Alefs der Form $\mathfrak{a}^{\aleph_0}$, z. B. $2^{\aleph_0} = \aleph$). Die Frage, ob $\aleph_1$ der einen oder andern Klasse angehört, welches Zeichen also in $\aleph_1^{\aleph_0} \geqq \aleph_1$ gilt, ist wieder das Kontinuumproblem, da $\aleph_1^{\aleph_0} = \aleph$ (aus $2 < \aleph_1 \leqq \aleph$ folgt $2^{\aleph_0} \leqq \aleph_1^{\aleph_0} \leqq \aleph^{\aleph_0} = 2^{\aleph_0}$).

Aus (4) folgt:

III. *Eine Summe, vom Typus* $< \omega_{\alpha+1}$, *von Ordnungszahlen* $< \omega_{\alpha+1}$ *ist selbst noch* $< \omega_{\alpha+1}$.

Denn sei

$$\sigma = \sum_\eta \alpha_\eta = \alpha_0 + \alpha_1 + \cdots + \alpha_\eta + \cdots (\eta < \beta;\ \beta < \omega_{\alpha+1},\ \alpha_\eta < \omega_{\alpha+1}).$$

Jeder Summand hat eine Mächtigkeit $< \aleph_{\alpha+1}$ oder $\leqq \aleph_\alpha$, dasselbe gilt von β, also hat σ eine Mächtigkeit $\leqq \aleph_\alpha \aleph_\alpha = \aleph_\alpha < \aleph_{\alpha+1}$, $\sigma < \omega_{\alpha+1}$.

Eine andere Form des Satzes ist:

IV. *Ist eine Menge von Ordnungszahlen* $< \omega_{\alpha+1}$ *vom Typus* $< \omega_{\alpha+1}$, *so ist die nächstgrößere Ordnungszahl immer noch* $< \omega_{\alpha+1}$.

Denn ist $W = \{\alpha_0, \alpha_1, \ldots, \alpha_\eta, \ldots\}$ vom Typus β $(\eta < \beta)$ unter denselben Voraussetzungen wie soeben und $\sigma = \sum_\eta \alpha_\eta$, so ist $W < \sigma + 1 < \omega_{\alpha+1}$, also die kleinste Zahl $> W$ immer noch $< \omega_{\alpha+1}$.

Diese Sätze regeln den Umfang der Zahlenabschnitte $W(\omega_{\alpha+1})$ oder der Zahlenklassen $Z(\aleph_\alpha)$. Z. B. gehört zu $Z(\aleph_0)$ mit einer Zahl α auch deren Nachfolger $\alpha + 1$ und mit einer ω-Folge von Zahlen $\alpha_0 < \alpha_1 < \alpha_2 < \cdots$ auch deren Limes; zu $Z(\aleph_1)$ gehört der Nachfolger jeder Zahl, mit einer ω-Folge deren Limes und mit einer ω_1-Folge $\alpha_0 < \alpha_1 < \cdots < \alpha_\omega < \alpha_{\omega+1} < \cdots$ auch deren Limes.

---

[1] Ist $\omega^\gamma z = \omega^\gamma (x + y)$ das höchste vorkommende Glied von ζ und etwa $x > 0$, so ist $\xi = \omega^\gamma x + \cdots < \omega_\alpha$, also $\omega^\gamma < \omega_\alpha$ und jedes Vielfache von ω^γ auch noch $< \omega_\alpha$, also $\zeta < \omega^\gamma (z + 1) < \omega_\alpha$.

Für Anfangszahlen, deren Index eine Limeszahl ist, brauchen die Sätze III IV nicht zu gelten; z. B. gehören zu $W(\omega_\omega)$ die Zahlen $\omega_0, \omega_1, \omega_2, \cdots$, aber nicht der Limes ω_ω dieser ω-Folge. Die Anfangszahlen, für die IV gilt, heißen *regulär*; zu ihnen gehören also die $\omega_{\alpha+1}$ und $\omega_0 = \omega$. [35]
Reguläre Anfangszahlen mit Limeszahl-Index sind bisher nicht bekannt; sie müßten von exorbitanter Größe sein.

§ 16. Der allgemeine Produktbegriff. [36]

Es sei $M = \{\ldots, m, \ldots, n, \ldots, p, \ldots\}$ eine geordnete Menge, deren Elementen m geordnete Mengen A_m zugewiesen sind; wir erhalten damit das zunächst ungeordnete Produkt

$$A = \prod_m^M A_m = (\ldots, A_m, \ldots, A_n, \ldots, A_p, \ldots)$$

als die Menge der Komplexe

$$a = (\ldots, a_m, \ldots, a_n, \ldots, a_p, \ldots) \qquad (a_m \varepsilon A_m).$$

Zwei solche Komplexe a und

$$b = (\ldots, b_m, \ldots, b_n, \ldots, b_p, \ldots)$$

bestimmen die Menge $M(a, b)$ derjenigen m, für die $a_m \neq b_m$; sie ist > 0 dann und nur dann, wenn die beiden Komplexe verschieden sind, und in diesem Falle eine geordnete Teilmenge von M. Nennen wir der Kürze halber M das *Argument*, die Elemente von $M(a, b)$ die Differenzstellen zwischen a, b.

Für drei Komplexe a, b, c ist offenbar

$$(1) \qquad M(a, c) \leqq M(a, b) \dot{+} M(b, c),$$

denn wenn $a_m \neq c_m$, so muß mindestens eine der Ungleichungen $a_m \neq b_m$, $b_m \neq c_m$ bestehen.

Wir wiesen bereits am Schlusse von § 10 darauf hin, daß im Falle eines *wohlgeordneten* M eine *lexikographische* Ordnung des Produkts A möglich ist. Hier hat nämlich die Menge $M(a, b)$ für $a \neq b$ stets ein erstes Element m und wir können dann $a \lesseqgtr b$ definieren, je nachdem $a_m \lesseqgtr b_m$; daß dies wirklich eine Ordnung, d. h. das Zeichen $<$ transitiv ist, werden wir gleich sehen. Wir verstehen dann unter dem obigen A das lexikographisch geordnete Produkt; seinen Typus müssen wir allerdings mit

$$\alpha = \prod_m^{M^*} \alpha_m = \cdots \alpha_p \cdots \alpha_n \cdots \alpha_m \cdots$$

bezeichnen, wo die Reihenfolge der Faktoren (α_m Typus von A_m) die umgekehrte ist wie im Mengenprodukt und in M. Das umgekehrte Argument $M^* = N$ kann passenderweise der *Exponent* genannt werden. Bei gleichen Faktoren $A_m = B$ entsteht aus dem Produkt die *Potenz* B^M mit dem Typus $\beta^{\mu^*} = \beta^\nu$ (β, μ, ν die Typen von B, M, N). Wir hatten z. B. die Potenz $\omega^{\omega^*} = 1 + \lambda$ gefunden (S. 48).

Im Falle eines *beliebigen* M, das nicht wohlgeordnet zu sein braucht, sind wir darauf angewiesen, die lexikographische Ordnung soweit zu definieren, wie sie definierbar ist. D. h. *wenn* $M(a, b)$ ein erstes Element m hat, das wir übrigens mit $m(a, b)$ bezeichnen wollen, und $a_m \lessgtr b_m$, so sei $a \lessgtr b$. Mit $a < b$ ist $b > a$ und es gilt das transitive Gesetz: mit $a < b$, $b < c$ ist $a < c$. Denn sei $m = m(a, b)$, $n = m(b, c)$ und $p = \min[m, n]$ (nämlich $p = m$ für $m \leqq n$, $p = n$ für $n \leqq m$), so ist

$$\text{für } l < p: a_l = b_l,\ b_l = c_l, \quad \text{also } a_l = c_l,$$

hingegen

$$a_p \leqq b_p, \quad b_p \leqq c_p$$

mit mindestens einem Ungleichungszeichen: $a_p < c_p$. D. h. $M(a, c)$ hat das erste Element p und es ist $a < c$. Wir können also das transitive Gesetz so aussprechen:

Für $a < b$, $b < c$ ist $a < c$ und

$$(2) \qquad m(a, c) = \min[m(a, b), \quad m(b, c)].$$

Hat, bei $a \neq b$, $M(a, b)$ kein erstes Element, so werden wir a, b lexikographisch *unvergleichbar* nennen und dies durch das Zeichen

$$a \,||\, b, \quad b \,||\, a$$

andeuten. Das Mengenprodukt A wird also im allgemeinen nur teilweise geordnet sein; zwei verschiedene Elemente stehen in einer und nur einer der drei Relationen

$$a < b, \quad a > b, \quad a \,||\, b,$$

die mit

$$b > a, \quad b < a, \quad b \,||\, a$$

gleichbedeutend sind. Man bemerke übrigens, daß die Vergleichbarkeit nicht transitiv zu sein braucht; es kann $a < b$, $b > c$, $a \,||\, c$ sein.

Um bei diesem Sachverhalt etwas für die Ordnungstheorie zu retten, werden wir uns an die *geordneten Teilmengen* von A halten müssen, und zwar, da es deren viele gibt, an solche, die möglichst einfach und willkürfrei definiert sind und möglichst viele von den Eigenschaften eines Mengenprodukts bewahren. Dazu verhilft uns wieder die Theorie der Wohlordnung.

Zwei Komplexe, für welche die Menge $M(a, b)$ der Differenzstellen *wohlgeordnet* ist, wollen wir etwa mit einem der Zahlentheorie entlehnten Ausdruck *kongruent* nennen:

$$a \equiv b \quad \text{oder} \quad b \equiv a.$$

Auch der Fall $M(a, b) = 0$ soll hierher gerechnet werden: $a \equiv a$. Wegen (1) gilt das transitive Gesetz: mit $a \equiv b$, $b \equiv c$ ist $a \equiv c$. Danach ist eine Einteilung von A in Klassen möglich derart, daß kongruente Komplexe derselben Klasse, inkongruente verschiedenen Klassen angehören; eine solche Klasse $A(a)$, nämlich die Menge der mit a und daher untereinander kongruenten Komplexe, ist mit der Klasse $A(b)$ entweder identisch (wenn $b \equiv a$) oder hat kein Element mit ihr gemein.

Die Komplexe einer Klasse sind lexikographisch vergleichbar und $A(a)$ ist also eine geordnete Menge. Auf die Eigenschaften (z. B. das assoziative Gesetz), die ihr Produktcharakter verleihen, wollen wir nicht eingehen und nur noch feststellen, daß sie eine *größte*, d. h. nicht mehr erweiterungsfähige geordnete Teilmenge von A ist (natürlich ist lexikographische Ordnung gemeint). Ist nämlich $c \not\equiv a$, so spalte man die nicht wohlgeordnete Menge $M(a, c)$ in zwei Komplemente P, Q, wo P wohl- [37]
geordnet und $Q > 0$ ohne erstes Element ist (z. B. sei Q die Summe aller Teilmengen von $M(a, c)$, die kein erstes Element haben, dann hat auch Q keins und das Komplement P kann keine von Null verschiedene Teilmenge ohne erstes Element haben, ist also wohlgeordnet). Definiert man dann einen Komplex b durch

$$b_m = c_m \text{ für } m \varepsilon P, \qquad b_m = a_m \text{ für } m \varepsilon M - P,$$

so ist $M(a, b) = P$, $M(b, c) = Q$, also $b \equiv a$, $b \| c$, folglich: wenn $c \not\equiv a$, so ist c mit mindestens einem Komplex b der Klasse $A(a)$ unvergleichbar, $A(a)$ ist nicht erweiterungsfähig.

Ein wichtiges Beispiel liefert der Fall, daß der *Exponent* $N = M^*$ wohlgeordnet ist (nicht mehr, wie ursprünglich, das Argument M). Hier ist jede Menge $M(a, b)$ invers wohlgeordnet; soll sie auch wohlgeordnet sein, so ist sie endlich, d. h. $a \equiv b$ bedeutet, daß sich die Komplexe a, b nur an endlich vielen Stellen unterscheiden.

Nehmen wir etwa den Fall

$$N = \{0, 1, 2, \ldots\}, \qquad M = \{\ldots, 2, 1, 0\},$$

N vom Typus ω; unsere Komplexe sind

$$a = (\ldots, a_2, a_1, a_0), \qquad a_m \varepsilon A_m.$$

Um die Klasse $A(a)$ und ihren Typus zu untersuchen, sei

$$A_m = B_m + \{a_m\} + C_m, \qquad \alpha_m = \beta_m + 1 + \gamma_m$$

die durch a_m bewirkte Zerlegung der Menge A_m und ihres Typus; es mögen ferner x_m, y_m, z_m die Mengen A_m, B_m, C_m durchlaufen. Die Komplexe $\equiv a$ erscheinen dann in folgender Ordnung:

$$\begin{array}{c}
\vdots \\
(\ldots, a_4, y_3, x_2, x_1, x_0) \\
(\ldots, a_4, a_3, y_2, x_1, x_0) \\
(\ldots, a_4, a_3, a_2, y_1, x_0) \\
(\ldots, a_4, a_3, a_2, a_1, y_0) \\
(\ldots, a_4, a_3, a_2, a_1, a_0) \\
(\ldots, a_4, a_3, a_2, a_1, z_0) \\
(\ldots, a_4, a_3, a_2, z_1, x_0) \\
(\ldots, a_4, a_3, z_2, x_1, x_0) \\
(\ldots, a_4, z_3, x_2, x_1, x_0) \\
\vdots
\end{array}$$

Die Punkte oben und unten bedeuten die weitere Fortsetzung des Schemas; die Punkte innerhalb der Komplexe bedeuten, daß dort Elemente a_m stehen; jeder Komplex außer dem mittelsten repräsentiert (für $x_m \varepsilon A_m$, $y_m \varepsilon B_m$, $z_m \varepsilon C_m$) eine ganze Menge solcher, die lexikographisch zu ordnen sind, und diese Mengen bilden als Summanden, von oben nach unten geordnet, die ganze Menge $A(a)$. Für den Typus von $A(a)$ erhält man demgemäß

$$\alpha(a) = \cdots + \alpha_0\alpha_1\alpha_2\beta_3 + \alpha_0\alpha_1\beta_2 + \alpha_0\beta_1 + \beta_0 + 1$$
$$+ \gamma_0 + \alpha_0\gamma_1 + \alpha_0\alpha_1\gamma_2 + \alpha_0\alpha_1\alpha_2\gamma_3 + \cdots;$$

man sieht, wie er durch eine *nach links und rechts* fortschreitende Summation gewissermaßen als Limes der Partialprodukte

$$\begin{aligned}
\alpha_0 &= \beta_0 + 1 + \gamma_0 \\
\alpha_0\alpha_1 &= \alpha_0(\beta_1 + 1 + \gamma_1) = \alpha_0\beta_1 + \alpha_0 + \alpha_0\gamma_1 \\
&= \alpha_0\beta_1 + \beta_0 + 1 + \gamma_0 + \alpha_0\gamma_1 \\
\alpha_0\alpha_1\alpha_2 &= \alpha_0\alpha_1\beta_2 + \alpha_0\beta_1 + \beta_0 + 1 + \gamma_0 + \alpha_0\gamma_1 + \alpha_0\alpha_1\gamma_2 \\
&\cdots
\end{aligned}$$

erscheint. Die Ähnlichkeit mit der Entstehung der Cantorschen Produkte von Ordnungszahlen (§ 14) ist unverkennbar; ja diese sind tatsächlich als Spezialfälle in unserem allgemeinen Produktbegriff enthalten. Nimmt man nämlich alle A_m als wohlgeordnet und die a_m als ihre ersten Elemente, so ist $\beta_m = 0$, $\alpha_m = 1 + \gamma_m$ zu setzen, also

$$\alpha(a) = 1 + \gamma_0 + \alpha_0\gamma_1 + \alpha_0\alpha_1\gamma_2 + \cdots$$

und das ist (alle $\alpha_m > 1$, $\gamma_m > 0$ vorausgesetzt) im Cantorschen Sinne eben

$$\alpha_0\alpha_1\alpha_2 \ldots = \lim \alpha_0\alpha_1 \ldots \alpha_m.$$

Allgemein erhält man die Cantorschen Produkte als Typen unserer Klassen $A(a)$, wenn man den Exponenten N und die Mengen A_m wohlgeordnet und den Komplex a aus den ersten Elementen a_m der A_m bestehend annimmt; mit dieser nicht schwer zu beweisenden Tatsache ist nun auch der Umstand (S. 66) aufgeklärt, daß die Cantorschen Produkte nicht die Mächtigkeit der vollen Produkte A — von denen ja die $A(a)$ nur Teilmengen sind — zu haben brauchen.

Man kann jene Klassen $A(a)$ noch weiter spalten und damit neue produktartige Mengen gewinnen, wenn man auf die Mächtigkeit der Menge $M(a, b)$ achtet. Schreiben wir

$$a \equiv b\,(\omega_\xi),$$

wenn die Menge $M(a, b)$ wohlgeordnet und von einem Typus $< \omega_\xi$ (oder einer Mächtigkeit $< \aleph_\xi$) ist, wo ω_ξ eine Anfangszahl bedeutet; auch diese verschärfte Kongruenz ist nach (1) transitiv und liefert eine Einteilung in Klassen $A_\xi(a)$, wobei, für $\xi < \eta$, $A_\xi(a)$ Teilmenge von $A_\eta(a)$ ist; für hinlänglich großes ξ fallen diese Klassen mit den $A(a)$ zusammen. Die kleinste

Klasse $A_0(a)$, durch $a \equiv b(\omega)$ definiert, besteht aus den Komplexen, die sich von a nur an endlich vielen Stellen unterscheiden; für wohlgeordneten Exponenten fällt bereits sie mit $A(a)$ zusammen. Ist das Argument M wohlgeordnet, so gibt es nur eine Klasse $A(a) = A$; aber A kann auch in diesem Fall in kleinere Klassen $A_\xi(a)$ zerfallen. — Bei gleichen Faktoren $A_m = B$ erhält man entsprechende Mengen, deren *Potenzcharakter* aber nur dann ausreichend gewahrt bleibt, wenn man den Komplex a auch aus lauter gleichen Elementen $a_m = b$ bestehen läßt.

Näher auf den allgemeinen Produktbegriff einzugehen verbietet der Raum; wir wollten aber doch dem Leser die Aufklärung nicht schuldig bleiben, daß und wie sich die verschiedenen Produktbildungen, die er kennengelernt hat, unter einen umfassenden Begriff subsumieren.

Fünftes Kapitel.

Mengensysteme.

§ 17. Ringe und Körper. [38]

Eine Menge von Mengen wollen wir der Deutlichkeit wegen ein *System* von Mengen nennen; wir bezeichnen die Mengensysteme mit großen deutschen Buchstaben. $M \varepsilon \mathfrak{M}$ bedeutet also, daß die Menge M dem System $\mathfrak{M}$ angehört. Die betrachteten Mengen M sind reine Mengen, ohne Relationen (Ordnung) zwischen ihren Elementen, so daß wir in dieser Hinsicht zum Standpunkt der ersten beiden Kapitel zurückkehren; die inzwischen erlangte Kenntnis der Ordnungszahlen wird uns aber doch nützlich und bisweilen unentbehrlich sein. Wir richten unser Augenmerk insbesondere auf *größte* und *kleinste* Mengensysteme gewisser Art: Systeme, denen man, nach einer bestimmten Vorschrift, kein Element hinzufügen oder keins wegnehmen kann.

1. Ringe. *Ein Mengensystem heiße ein Ring*[1]*, wenn Summe und Durchschnitt von zwei Mengen des Systems wieder dem System angehören.* Dasselbe ist dann für endlich viele Mengen des Systems der Fall. Ein Ring ist also eine Art größten Mengensystems, ein System, das durch die Operationen Summe und Durchschnitt (an endlich vielen Elementen) nicht erweitert werden kann.

Beispiele. Das System aller Teilmengen einer gegebenen Menge ist ein Ring. — Bedeutet $I = [\alpha, \beta)$ das Intervall der Zahlen $\alpha \leqq x < \beta$, so bilden die Summen

$$S = I_1 + I_2 + \cdots + I_n$$

---

[1]) Die Namen Ring und Körper beruhen auf einer schwachen Analogie zu den gleichbenannten Begriffen der Zahlentheorie.

aus endlich vielen disjunkten I, mit Hinzurechnung der Nullmenge, einen Ring. Denn der Durchschnitt zweier I ist ein I oder 0, also nach dem distributiven Gesetz der Durchschnitt zweier S ein S. Schließt man S in ein I ein, so ist offenbar $I - S$ ein S; demnach ist $I - (S_1 \dot{+} S_2) = (I - S_1)(I - S_2)$ ein S, und das Komplement davon in I, also $S_1 \dot{+} S_2$, wieder ein S. Übrigens ist auch die Differenz $S - S_1$ zweier S ein S, denn schließt man S in ein I ein, so ist $S(I - S_1) = S - S_1$ ein S.

Zu einem beliebigen Mengensystem $\mathfrak{M}$ gibt es einen eindeutig bestimmten kleinsten Ring $\geqq \mathfrak{M}$.

Wir werden diesen Ring, d. h. die ihm angehörigen Mengen, hier zwar sofort sehr einfach angeben können, wollen aber als Vorbild für andere, weniger elementare Fälle einen allgemeinen Existenzbeweis geben. Es gibt überhaupt Ringe $\geqq \mathfrak{M}$; denn ist S die Summe aller Mengen $\varepsilon\, \mathfrak{M}$, so ist das System $\mathfrak{S}$ aller Teilmengen von S ein Ring über $\mathfrak{M}$. Da der Durchschnitt beliebig vieler Ringe offenbar wieder ein Ring ist, so ist der Durchschnitt $\mathfrak{R}_0$ aller Ringe $\mathfrak{R}$, für die $\mathfrak{M} \leqq \mathfrak{R} \leqq \mathfrak{S}$, ein Ring über $\mathfrak{M}$. Er ist der kleinstmögliche, d. h. in *jedem* Ringe $\mathfrak{R}$ über $\mathfrak{M}$ enthalten, da ja $\mathfrak{M} \leqq \mathfrak{R}\mathfrak{S} \leqq \mathfrak{S}$, also $\mathfrak{R}_0 \leqq \mathfrak{R}\mathfrak{S} \leqq \mathfrak{R}$.

Dieser kleinste Ring über $\mathfrak{M}$, nennen wir ihn jetzt $\mathfrak{R}$, läßt sich offenbar so bilden: er besteht aus den endlichen Summen

$$R = D_1 \dot{+} D_2 \dot{+} \cdots \dot{+} D_n,$$

deren Summanden ihrerseits von der Form

$$D = M_1 M_2 \ldots M_m$$

d. h. Durchschnitte aus endlich vielen Mengen $M (\varepsilon\, \mathfrak{M})$ sind. Daß die genannten Mengen R zu $\mathfrak{R}$ gehören müssen, ist klar; sie bilden aber schon selbst einen Ring (also $\mathfrak{R}$), da nach dem assoziativen Gesetz die Summe, nach dem distributiven der Durchschnitt zweier R wieder ein R ist.

Offenbar kann man die Operationen Summe und Durchschnitt auch in umgekehrter Reihenfolge anwenden, d. h. $\mathfrak{R}$ besteht aus den Durchschnitten

$$R = S_1 S_2 \ldots S_n$$

aus endlich vielen Mengen S, die ihrerseits Summen

$$S = M_1 \dot{+} M_2 \dot{+} \cdots \dot{+} M_m$$

aus endlich vielen Mengen M sind.

2. Körper. *Ein Mengensystem heiße ein Körper, wenn Summe, Durchschnitt und Differenz von zwei Mengen des Systems wieder dem System angehören.*

Bei der Differenz ist, wie immer, der Subtrahend als Teilmenge des Minuenden anzunehmen. Es würde übrigens genügen, die Forderung nur für Summe und Differenz zu stellen, da der Durchschnitt auf diese beiden Operationen zurückführbar ist (S. 17). Ein Körper ist a fortiori ein Ring.

Beispiel. Die obigen Intervallsummen S bilden einen Körper, wie vorhin bereits bewiesen wurde. Hätte man statt der halboffenen Intervalle $I = [\alpha, \beta)$ entweder offene (α, β) oder abgeschlossene $[\alpha, \beta]$ genommen, so würden die S zwar noch einen Ring, aber keinen Körper bilden.

Daß über einem beliebigen Mengensystem $\mathfrak{M}$ ein kleinster Körper $\mathfrak{K}$ existiert, erkennt man genau wie im Fall des Ringes; aber die Darstellung seiner Mengen ist hier nicht ganz so trivial. Da $\mathfrak{K}$, wenn überhaupt eine Menge, so jedenfalls die Nullmenge enthält, so wollen wir diese bereits in $\mathfrak{M}$ aufnehmen; ferner enthält $\mathfrak{K}$ den kleinsten Ring $\mathfrak{R} \geqq \mathfrak{M}$ und ist auch der kleinste Körper über $\mathfrak{R}$. Wir setzen daher alsbald $\mathfrak{M}$ als Ring voraus, dem die Nullmenge angehört. Betrachten wir dann endlich viele Mengen $M_1 \geqq M_2 \geqq \cdots \geqq M_n$ oder, zur Vereinfachung der Schreibweise, eine absteigende Mengenfolge von Mengen M

$$(1) \qquad M_1 \geqq M_2 \geqq M_3 \geqq \cdots$$

mit schließlich verschwindenden Gliedern. Die Differenzen $M_1 - M_2$, $M_2 - M_3, \ldots$ sind disjunkt. Die Menge

$$(2) \qquad A = (M_1 - M_2) + (M_3 - M_4) + (M_5 - M_6) + \cdots$$

heiße eine (endliche) *Differenzenkette* aus dem System $\mathfrak{M}$. Diese Mengen A sind offenbar sämtlich in $\mathfrak{K}$ aufzunehmen; wir werden zeigen, daß sie selbst schon einen Körper bilden, der folglich der gesuchte Körper $\mathfrak{K}$ ist.

Die Komplemente $M - A$ sind wieder Mengen A. Schreibt man

$$A = MA = (MM_1 - MM_2) + (MM_3 - MM_4) + \cdots,$$

was wieder eine Darstellung der Form (2) ist, so sieht man, daß man nur zu beweisen braucht: $M_0 - A$ ist für $M_0 \geqq M_1$ ein A. Dann ist aber

$$(3) \qquad M_0 = (M_0 - M_1) + (M_1 - M_2) + (M_2 - M_3) + \cdots,$$

$$(4) \qquad M_0 - A = (M_0 - M_1) + (M_2 - M_3) + \cdots$$

und die Behauptung ist bewiesen. Man beachte noch, daß zu jedem A ein $M \geqq A$ existiert, z. B. M_1.

Der Durchschnitt zweier A ist ein A. Hier ist eine veränderte Bezeichnung zweckmäßig: wir setzen

$$(5) \qquad \left\{\begin{aligned} A &= (M_0 - M_0') + (M_1 - M_1') + \cdots \\ B &= (N_0 - N_0') + (N_1 - N_1') + \cdots \\ &M_0 \geqq M_0' \geqq M_1 \geqq M_1' \geqq \cdots \\ &N_0 \geqq N_0' \geqq N_1 \geqq N_1' \geqq \cdots \end{aligned}\right.$$

mit schließlich verschwindenden[1]) M_i, M_i', N_k, N_k', die dem System $\mathfrak{M}$ angehören.

Definieren wir nun (die Indizes i, k, l durchlaufen die Zahlen $0, 1, 2, \cdots$)

[1]) Diese Voraussetzung ist übrigens hier nicht wesentlich; auch die unendlichen Differenzenketten (2) haben die Durchschnittseigenschaft.

$$(6) \qquad \begin{cases} P_l = \mathfrak{S}_{i+k=l} \; M_i N_k \\ P'_l = \mathfrak{S}_{i+k=l} \; (M'_i N_k \dotplus M_i N'_k), \end{cases}$$

also
$$P_0 = M_0 N_0, \quad P_1 = M_0 N_1 \dotplus M_1 N_0, \ldots$$
$$P'_0 = M'_0 N_0 \dotplus M_0 N'_0, \ldots$$

so gehören auch die P_l, P'_l dem Ringe $\mathfrak{M}$ an und verschwinden schließlich. Dabei ist $P_l \geqq P'_l$ und, wenn man

$$P_{l+1} = \mathfrak{S}_{i+k=l}(M_{i+1} N_k \dotplus M_i N_{k+1})$$

beachtet, $P'_l \geqq P_{l+1}$, also

$$P_0 \geqq P'_0 \geqq P_1 \geqq P'_1 \geqq \ldots .$$

Wir zeigen nun, daß

$$(7) \qquad AB = (P_0 - P'_0) + (P_1 - P'_1) + \cdots$$

und sogar einzeln

$$(8). \qquad \sum_{i+k=l} (M_i - M'_i)(N_k - N'_k) = P_l - P'_l.$$

Nennen wir die hier links stehende Summe C_l. Wenn $x \varepsilon C_l$, so ist etwa $x \varepsilon (M_{i_0} - M'_{i_0})(N_{k_0} - N'_{k_0})$, also $x \varepsilon M_{i_0} N_{k_0} \leqq P_l$, aber zugleich $x \,\bar{\varepsilon}\, P'_l$. Denn für $i + k = i_0 + k_0 = l$ ist entweder $i \geqq i_0$, $M'_i N_k \leqq M'_{i_0}$ oder $i < i_0$, $k > k_0$, $M'_i N_k \leqq N_{k_0+1} \leqq N'_{k_0}$; also $x \,\bar{\varepsilon}\, M'_i N_k$, ebenso $x \,\bar{\varepsilon}\, M_i N'_k$. Also ist $C_l \leqq P_l - P'_l$. Umgekehrt, ist $x \varepsilon P_l - P'_l \leqq P_l$ und etwa $x \varepsilon M_i N_k$, so ist $x \,\bar{\varepsilon}\, M'_i$, da sonst $x \varepsilon M'_i N_k \leqq P'_l$ wäre, ebenso $x \,\bar{\varepsilon}\, N'_k$, $x \varepsilon (M_i - M'_i)(N_k - N'_k) \leqq C_l$, also $P_l - P'_l \leqq C_l$. Damit ist (8) und (7) bewiesen: AB ist ein A.

Die Differenz zweier A ist ein A. Zu $A \geqq A_1$ wähle man $M \geqq A$, dann ist $A - A_1 = A(M - A_1)$ ein A.

Die Summe zweier A ist ein A. Zu A_1, A_2 wähle man umfassende Mengen M_1, M_2 und $M = M_1 \dotplus M_2$; dann ist $M - (M - A_1)(M - A_2) = A_1 \dotplus A_2$ ein A.

Damit ist bewiesen, daß die A einen Körper bilden.

3. Erweiterte Körper. Einer späteren Anwendung (§ 30, II III) wegen wollen wir die Differenzenketten ins Unendliche verlängern. Die griechischen Buchstaben sollen die Ordnungszahlen $< \omega_\mu$ durchlaufen, wo ω_μ eine festgewählte *Anfangszahl*, die erste Ordnungszahl von der Mächtigkeit $\aleph_\mu$ ist. Aus dem System $\mathfrak{M}$, das wieder ein Ring sein und die Nullmenge enthalten soll, bilden wir eine absteigend wohlgeordnete Folge vom Typus ω_μ

$$(9) \qquad M_1 \geqq M_2 \geqq \cdots \geqq M_{\omega+1} \geqq M_{\omega+2} \geqq \cdots,$$

deren Glieder also mit $M_{\xi+1}$ bezeichnet sind, und damit die Menge

$$(10)\quad A = (M_1 - M_2) + (M_3 - M_4) + \cdots + (M_{\omega+1} - M_{\omega+2}) + \cdots$$
$$= \sum_{\xi}(M_{2\xi+1} - M_{2\xi+2}),$$

wobei daran zu erinnern ist, daß die Ordnungszahlen entweder gerade (2ξ) oder ungerade ($2\xi+1$) sind, die Limeszahlen insbesondere gerade. Wir nennen A wieder eine *Differenzenkette* aus $\mathfrak{M}$, und zwar vom Typus ω_μ, falls alle Mengen (9) von Null verschieden sind, dagegen vom Typus $\eta < \omega_\mu$, falls $M_{\eta+1}$ die erste verschwindende Menge in (9) ist; natürlich hängt der Typus nicht nur von der Menge A selbst, sondern auch von der gewählten Darstellung ab. Wir behaupten:

I. *Wenn das System* $\mathfrak{M}$ (das ein Ring ist und die Nullmenge enthält) *auch noch so beschaffen ist, daß der Durchschnitt aus weniger als* $\aleph_\mu$ *Mengen* M *ein* M *ist, so bilden die Differenzenketten* (*aus* $\mathfrak{M}$) *vom Typus* $< \omega_\mu$ *einen Körper.*

Beweis. $M - A$ *ist wieder ein* A. Wie oben genügt es, zu zeigen, daß $M_0 - A$ für $M_0 \geqq M_1$ ein A ist. Wir definieren noch für Limeszahlen $\eta < \omega_\mu$ die Mengen

$$M_\eta = \mathfrak{D}_{\xi<\eta} M_{\xi+1},$$

die nach der Durchschnitts-Voraussetzung wieder Mengen M sind. Es bilden jetzt die sämtlichen Mengen M_ξ ein absteigend wohlgeordnetes System

$$M_0 \geqq M_1 \geqq M_2 \geqq \cdots \geqq M_\omega \geqq M_{\omega+1} \geqq \cdots \geqq M_{\omega 2} \geqq \cdots,$$

worin jede Menge mit Limesindex der Durchschnitt aller vorangehenden Mengen ist. Hierbei ist

$$(11)\qquad M_0 = \sum_{\xi}(M_\xi - M_{\xi+1}),$$

denn für jedes $x \varepsilon M_0$ gibt es, da die M_ξ schließlich $= 0$ sind, ein erstes M_ζ mit $x \bar{\varepsilon} M_\zeta$, wobei $\zeta > 0$ und keine Limeszahl, also $\zeta = \xi + 1$ und $x \varepsilon M_\xi - M_{\xi+1}$ ist. Durch Trennung der geraden und ungeraden Indizes folgt

$$(12)\quad M_0 - A = \sum_{\xi}(M_{2\xi} - M_{2\xi+1})$$
$$= (M_0 - M_1) + (M_2 - M_3) + \cdots + (M_\omega - M_{\omega+1}) + \cdots.$$

Der Durchschnitt zweier A *ist wieder ein* A [1]). Wir schreiben jetzt wieder wie in (5)

$$(13)\qquad \left\{\begin{aligned} A &= \sum_{\xi}(M_\xi - M'_\xi) \\ B &= \sum_{\xi}(N_\xi - N'_\xi) \\ M_0 \geqq M'_0 &\geqq \cdots \geqq M_\omega \geqq M'_\omega \geqq \cdots \\ N_0 \geqq N'_0 &\geqq \cdots \geqq N_\omega \geqq N'_\omega \geqq \cdots. \end{aligned}\right.$$

---

[1]) Hierbei genügt, daß $\mathfrak{M}$ ein Ring sei; auch die Voraussetzung des schließlichen Verschwindens der Mengen (9) ist entbehrlich, und der Durchschnitt zweier Differenzenketten vom Typus $\leqq \omega_\mu$ ist wieder eine solche.

Um (6) zu übertragen, müssen wir die *natürliche Summe* verwenden: wir definieren

$$(14)\qquad \begin{cases} P_\zeta = \overset{\zeta}{\mathfrak{S}} M_\xi N_\eta \\ P'_\zeta = \overset{\zeta}{\mathfrak{S}} (M'_\xi N_\eta \dotplus M_\xi N'_\eta), \end{cases}$$

wo $\overset{\zeta}{\mathfrak{S}}$ bedeutet, daß über die Paare ξ, η mit $\sigma(\xi, \eta) = \zeta$ summiert werden soll. Da es deren nur endlich viele gibt, gehören die Mengen P, P' dem Ringe $\mathfrak{M}$ an. Mit $\xi < \omega_\mu$, $\eta < \omega_\mu$ ist auch $\zeta < \omega_\mu$ (S. 72 Anm.) und man sieht leicht, daß die P, P' schließlich verschwinden, wenn die M, M' und N, N' schließlich verschwinden. Ferner ist

$$P_0 \geqq P'_0 \geqq P_1 \geqq P'_1 \geqq \cdots \geqq P_\omega \geqq P'_\omega \geqq \cdots.$$

In der Tat folgt $P_\zeta \geqq P'_\zeta$ unmittelbar; überdies ist $P'_{\zeta_0} \geqq P_\zeta$ für $\zeta_0 < \zeta$. Denn ist $x \varepsilon P_\zeta$ und etwa $x \varepsilon M_\xi N_\eta$ mit $\sigma(\xi, \eta) = \zeta$, so gibt es (S. 69) ein Zahlenpaar ξ_0, η_0 mit $\sigma(\xi_0, \eta_0) = \zeta_0$ und $\xi_0 \leqq \xi$, $\eta_0 \leqq \eta$; ist dann etwa $\xi_0 < \xi$, so ist $M_\xi \leqq M'_{\xi_0}$, $N_\eta \leqq N_{\eta_0}$ und $x \varepsilon P'_{\zeta_0}$, ebenso für $\eta_0 < \eta$.

Nun ist wiederum

$$(15)\qquad AB = \mathfrak{\Sigma}_\zeta (P_\zeta - P'_\zeta)$$

und sogar einzeln

$$(16)\qquad \overset{\zeta}{\Sigma} (M_\xi - M'_\xi)(N_\eta - N'_\eta) = P_\zeta - P'_\zeta;$$

der Beweis gestaltet sich genau wie der von (8), weil die natürliche Summe alle notwendigen Eigenschaften einer gewöhnlichen hat (insbesondere: wenn $\sigma(\xi, \eta) = \sigma(\xi_0, \eta_0)$ und $\xi < \xi_0$, so ist $\eta > \eta_0$).

Auch daß Summe und Differenz zweier A wieder ein A ist, folgt wie bei den endlichen Differenzenketten. Damit ist I bewiesen.

Für $\omega_\mu = \omega_0 = \omega$, wo $\mathfrak{M}$ nur ein Ring zu sein braucht, erhalten wir wieder das Ergebnis der vorigen Nummer: die endlichen Differenzenketten bilden einen Körper (den kleinsten über $\mathfrak{M}$). Für $\omega_\mu = \omega_1 = \Omega$ ergibt sich, daß die höchstens abzählbaren Differenzenketten (vom Typus $< \Omega$) einen Körper bilden, falls der Durchschnitt abzählbar vieler M ein M ist.

Wenn der Durchschnitt beliebig vieler M ein M ist, so bilden alle Differenzenketten aus $\mathfrak{M}$ einen Körper. Denn alle diese Differenzenketten (jede in einer bestimmten Darstellung) sind vom Typus $< \omega_\mu$, wenn ω_μ hinlänglich groß gewählt wird.

§ 18. Borelsche Systeme.

1. Die Prozesse σ, δ. Ein Mengensystem, dem Summe und Durchschnitt von zwei (oder endlich vielen) Mengen des Systems angehört, hieß ein Ring. Dehnen wir diese Forderung auf Summe und Durchschnitt von *abzählbar* vielen Mengen aus, so gelangen wir zum Begriff eines *Borelschen Systems.* Die einfache Art, wie wir den kleinsten Ring über einem gegebenen

Mengensystem $\mathfrak{M}$ bilden konnten, läßt sich aber auf den gegenwärtigen Fall nicht übertragen; wollten wir etwa wie damals (S. 78) die Durchschnitte

$$D = M_1 M_2 M_3 \cdots$$

aus Folgen von Mengen $M (\varepsilon\, \mathfrak{M})$ und damit die Summen

$$R = D_1 \dot{+} D_2 \dot{+} D_3 \dot{+} \cdots$$

aus Folgen von Mengen D bilden, so bilden diese R noch durchaus kein Borelsches System: die Summe abzählbar vieler R ist zwar nach dem assoziativen Gesetz wieder ein R, aber der Durchschnitt abzählbar vieler R gibt, nach dem distributiven Gesetz entwickelt, eine Summe *unabzählbar* vieler D und im allgemeinen also kein R.

Diese Verwicklung läßt es ratsam erscheinen, die Forderungen für Summe und Durchschnitt zunächst einmal zu trennen.

Ein Mengensystem heiße ein $\left\{\begin{matrix}\sigma\textit{-System}\\ \delta\textit{-System}\end{matrix}\right\}$, *wenn* $\left\{\begin{matrix}\textit{die Summe}\\ \textit{der Durchschnitt}\end{matrix}\right\}$ *jeder Folge von Mengen des Systems wieder dem System angehört.*

Wir machen einige Bemerkungen über σ-Systeme, die sich dann ohne weiteres auf δ-Systeme übertragen. In einem σ-System gehört auch die Summe

$$A \dot{+} B \dot{+} A \dot{+} B \dot{+} \cdots = A \dot{+} B$$

zweier und endlich vieler Mengen des Systems wieder dem System an.

Das kleinste σ-System über einem gegebenen Mengensystem $\mathfrak{M}$ (seine Existenz ist wie die des kleinsten Ringes zu erschließen) heiße $\mathfrak{M}_\sigma$. Es wird von den Mengen

$$(1) \qquad M_\sigma = M_1 \dot{+} M_2 \dot{+} M_3 \dot{+} \cdots = \mathfrak{S} M_n \qquad (M_n \,\varepsilon\, \mathfrak{M})$$

gebildet, d. h. von den Summen aus Folgen von Mengen $M \varepsilon \mathfrak{M}$. Denn die M_σ müssen zu $\mathfrak{M}_\sigma$ gehören, bilden aber bereits selbst ein σ-System (Verwandlung einer Doppelfolge in eine einfache).

Ist $\mathfrak{M}$ ein Ring, so ist auch $\mathfrak{M}_\sigma$ ein Ring; denn der Durchschnitt von zwei M_σ ist, nach dem distributiven Gesetz, ein M_σ. Ferner können in diesem Fall die M_σ als Summen *aufsteigender* Folgen von Mengen M dargestellt werden, denn aus (1) folgt

$$M_\sigma = \mathfrak{S} S_n, \quad S_n = M_1 \dot{+} M_2 \dot{+} \cdots \dot{+} M_n,$$

wo die S_n wieder Mengen M und $S_1 \leqq S_2 \leqq \cdots$ ist.

In einem δ-System gehört auch der Durchschnitt von endlich vielen Mengen des Systems dem System an. Das kleinste δ-System $\mathfrak{M}_\delta$ über $\mathfrak{M}$ wird von den Mengen

$$(2) \qquad M_\delta = M_1 M_2 M_3 \ldots = \mathfrak{D} M_n \qquad (M_n \,\varepsilon\, \mathfrak{M})$$

gebildet: die M_δ sind die Durchschnitte aus Folgen von Mengen $M \varepsilon \mathfrak{M}$. Ist $\mathfrak{M}$ ein Ring, so ist $\mathfrak{M}_\delta$ ein Ring, und die M_δ können als Durchschnitte *absteigender* Folgen von Mengen M dargestellt werden:

$$M_\delta = \mathfrak{D} D_n, \qquad D_n = M_1 M_2 \cdots M_n.$$

Ein Mengensystem, das zugleich σ-System und δ-System ist, heiße ein $(\sigma\delta)$-System oder ein Borelsches System.

[39] Über einem Mengensystem $\mathfrak{M}$ existiert wieder ein kleinstes Borelsches System $\mathfrak{B} = \mathfrak{M}_{(\sigma\delta)}$. Die ihm angehörigen Mengen B heißen die von $\mathfrak{M}$ oder von den Mengen $M \varepsilon \mathfrak{M}$ erzeugten *Borelschen Mengen.*

Wir können diese Mengen in folgender Weise darstellen. Lassen wir $i, k, l, m, \ldots$ die natürlichen Zahlen durchlaufen und setzen

$$S = \underset{i}{\mathfrak{S}} D_i, \quad D_i = \underset{k}{\mathfrak{D}} S_{ik}, \quad S_{ik} = \underset{l}{\mathfrak{S}} D_{ikl}, \quad D_{ikl} = \underset{m}{\mathfrak{D}} S_{iklm}, \ldots$$

mit der Bestimmung, daß für jede Folge natürlicher Zahlen $i, k, l, m, \ldots$ in der Folge der Mengen $D_i, S_{ik}, D_{ikl}, S_{iklm} \ldots$ schließlich lauter Mengen M auftreten sollen. Ebenso sei, indem man die abwechselnde Summen- und Durchschnittsbildung mit der letzteren beginnt,

$$D = \underset{i}{\mathfrak{D}} S_i, \quad S_i = \underset{k}{\mathfrak{S}} D_{ik}, \quad D_{ik} = \underset{l}{\mathfrak{D}} S_{ikl}, \quad S_{ikl} = \underset{m}{\mathfrak{S}} D_{iklm}, \ldots$$

und in jeder Folge $S_i, D_{ik}, S_{ikl}, D_{iklm}, \ldots$ sollen schließlich lauter Mengen M auftreten. Dann bilden sowohl die S als auch die D die von $\mathfrak{M}$ erzeugten Borelschen Mengen.

Da nämlich die D_i der ersten Formelgruppe beliebige Mengen D der zweiten Gruppe sind, so ist jedes D_σ (Summe einer Folge von Mengen D) ein S und vice versa, insbesondere ist jedes $D = D \dotplus D \dotplus \cdots$ selbst ein S. Ebenso sind die S_δ mit den D identisch und jedes S ein D. Danach sind die S mit den D identisch und bilden ein Borelsches System, das die M enthält ($S = D_i = S_{ik} = \cdots = M$); das kleinste Borelsche System $\mathfrak{B}$ ist darin enthalten, d. h. jede Menge B ist ein S.

Umgekehrt ist auch jedes S ein B. Wenn nämlich S kein B wäre, so müßte mindestens ein D_i kein B sein, dann mindestens ein S_{ik}, ein D_{ikl} usw., es gäbe also mindestens eine Zahlenfolge $i, k, l, \ldots$ derart, daß in der Folge $D_i, S_{ik}, D_{ikl}, \cdots$ kein B vorkäme, im Widerspruch zu der Festsetzung, daß die Folge schließlich aus Mengen M bestehen soll.

Die gewonnene Darstellung, die freilich alle Borelschen Mengen mit einem Schlage liefert, ist aber doch nicht durchsichtig genug, und wir versuchen einen sukzessiven Aufbau des Borelschen Systems $\mathfrak{B}$. Wir müssen in dieses, wenn die Bezeichnungen (1)(2) naturgemäß weitergeführt werden, jedenfalls die folgenden Mengen aufnehmen:

$$(3) \quad \begin{cases} \text{die Mengen } M\,(\varepsilon\,\mathfrak{M}), \\ \text{die Summen } M_\sigma \text{ aus Folgen von Mengen } M, \\ \text{die Durchschnitte } M_{\sigma\delta} \text{ aus Folgen von Mengen } M_\sigma, \\ \text{die Summen } M_{\sigma\delta\sigma} \text{ aus Folgen von Mengen } M_{\sigma\delta} \\ \qquad\qquad\qquad \text{usw.} \end{cases}$$

wobei wir Summen- und Durchschnittsbildung abwechseln lassen, da z. B. die Mengen $M_{\sigma\sigma}$, Summen aus Folgen von Mengen M_σ, mit den Mengen

M_σ identisch sind. Ebenso hätten wir, mit Durchschnitten statt mit Summen beginnend, die folgenden Mengen aufzunehmen:

$$(4)\quad \left\{\begin{array}{l} \text{die Mengen } M\,(\varepsilon\,\mathfrak{M}), \\ \text{die Durchschnitte } M_\delta \text{ aus Folgen von Mengen } M, \\ \text{die Summen } M_{\delta\sigma} \text{ aus Folgen von Mengen } M_\delta, \\ \text{die Durchschnitte } M_{\delta\sigma\delta} \text{ aus Folgen von Mengen } M_{\delta\sigma} \\ \qquad\qquad\qquad \text{usw.} \end{array}\right.$$

Die von diesen Mengen gebildeten Systeme wären mit $\mathfrak{M}, \mathfrak{M}_\sigma, \mathfrak{M}_{\sigma\delta}, \mathfrak{M}_{\sigma\delta\sigma}, \ldots$ oder $\mathfrak{M}, \mathfrak{M}_\delta, \mathfrak{M}_{\delta\sigma}, \mathfrak{M}_{\delta\sigma\delta}, \ldots$ zu bezeichnen (bei der Bezeichnung $\mathfrak{B} = \mathfrak{M}_{(\sigma\delta)}$ dürfen also die Klammern nicht fehlen).

Beispielsweise sind der obere und der untere Limes (S. 20)

$$\overline{M} = \overline{\text{Lim}}\; M_n = S_1 S_2 S_3 \ldots, \qquad S_n = M_n \dot{+} M_{n+1} \dot{+} M_{n+2} \dot{+} \cdots$$

$$\underline{M} = \underline{\text{Lim}}\; M_n = D_1 \dot{+} D_2 \dot{+} D_3 \dot{+} \cdots, \qquad D_n = M_n\, M_{n+1}\, M_{n+2} \cdots$$

von den Mengen M_n ($\varepsilon\mathfrak{M}$) erzeugte Borelsche Mengen; die S_n, D_n sind Mengen M_σ, M_δ, der obere Limes $\overline{M}$ ein $M_{\sigma\delta}$, der untere $\underline{M}$ ein $M_{\delta\sigma}$; der Limes einer konvergenten Folge beides zugleich.

Wenn wir nun aber auch sämtliche Mengen (3) oder (4) mit endlich vielen Indices σ, δ gebildet haben, so ist das Borelsche System $\mathfrak{B}$ noch nicht fertig. Wir müssen etwa in Fortsetzung des Verfahrens (3) die Durchschnitte

$$N = M\, M_\sigma\, M_{\sigma\delta}\, M_{\sigma\delta\sigma} \cdots$$

adjungieren, dann die $N_\sigma, N_{\sigma\delta}, N_{\sigma\delta\sigma}, \ldots$, schließlich wieder

$$P = N\, N_\sigma\, N_{\sigma\delta}\, N_{\sigma\delta\sigma} \cdots$$

und so weiter. Dieses „und so weiter" zu präzisieren, haben wir aber ein Mittel, nämlich die *Ordnungszahlen.*

2. Aufbau des Borelschen Systems $\mathfrak{B}$. Es seien ξ, η Ordnungszahlen $< \Omega$, wo $\Omega = \omega_1$ die Anfangszahl der Zahlenklasse $Z(\aleph_1)$ ist; wir unterscheiden wieder gerade Ordnungszahlen 2ξ und ungerade $2\xi + 1$. Wir ordnen nun jedem ξ ein System $\mathfrak{A}^\xi$ von Mengen A^ξ zu durch die *induktive* Vorschrift: [40]

$$(\alpha)\quad \left\{\begin{array}{l} \textit{Die Mengen } A^0 \textit{ sind die Mengen } M(\mathfrak{A}^0 = \mathfrak{M}); \\ \textit{die Mengen } A^\eta \textit{ sind für ungerades } \eta \textit{ die Summen, für gerades } \eta > 0 \\ \textit{die Durchschnitte aus Folgen von Mengen } A^\xi (\xi < \eta). \end{array}\right.$$

Dann ist das System aller Mengen A^ξ das kleinste Borelsche System $\mathfrak{B}$ über $\mathfrak{M}$.

Daß in der Tat alle Mengen A^ξ zu $\mathfrak{B}$ gehören müssen, ist klar, also nur noch zu zeigen, daß sie selbst schon ein Borelsches System bilden. Ist nun eine Folge von Mengen $A^{\xi_1}, A^{\xi_2}, \ldots$ gegeben, so sei ξ die nächstgrößere Ordnungszahl nach $\xi_1, \xi_2, \ldots$ (wegen § 15, IV ist $\xi < \Omega$). Eine der Zahlen $\xi, \xi + 1$ ist gerade, die andere ungerade, und unter den Mengen

A^ξ, $A^{\xi+1}$ kommt also Summe und Durchschnitt der gegebenen Folge vor. Man sieht daraus, daß es keinen Zweck gehabt hätte, die Definition auf $\xi \geqq \Omega$ auszudehnen (es würde $\mathfrak{A}^\Omega = \mathfrak{A}^{\Omega+1} = \cdots = \mathfrak{B}$ sein).

Die Systeme $\mathfrak{A}^\xi$ wachsen mit dem Index, d. h. $\mathfrak{A}^\xi \leqq \mathfrak{A}^\eta$ für $\xi < \eta$. Denn wegen $M = M \dot{+} M \dot{+} \cdots = MM\ldots$ ist jedes A^ξ auch ein $A^{\xi+1}$, $A^{\xi+2}, \ldots$ Daraus folgt, daß man die $A^{\xi+1}$ auch als Summen oder Durchschnitte (je nachdem ξ gerade oder ungerade ist) aus Folgen von Mengen A^ξ erklären kann; für eine Limeszahl η sind die A^η die Durchschnitte aus Folgen von Mengen A^ξ $(\xi < \eta)$. Den Anfang bilden also die Mengen (3); die

$$A^0, \quad A^1, \quad A^2, \quad A^3, \ldots$$

sind mit den

$$M, \quad M_\sigma, \quad M_{\sigma\delta}, \quad M_{\sigma\delta\sigma}, \ldots$$

identisch, dann kommen die A^ω als Durchschnitte aus Folgen vorangehender Mengen, die $A^{\omega+1}$ als Summen aus Folgen von Mengen A^ω usw. Wir bezeichnen die $\mathfrak{A}^\xi$ als *Borelsche Klassen* und sagen, daß eine Menge *genau* zur Klasse $\mathfrak{A}^\eta$ gehört, wenn sie dieser, aber keiner früheren Klasse angehört, wenn sie also ein A^η, aber kein A^ξ (für $\xi < \eta$) ist. (Es ist sonst üblich, aber weniger bequem, die Klassen disjunkt anzunehmen, also $\mathfrak{A}^\eta - \mathop{\mathfrak{S}}\limits_{\xi<\eta} \mathfrak{A}^\xi$ eine Klasse zu nennen.)

Eine modifizierte Erzeugung von $\mathfrak{B}$ erhält man aus der obigen, wenn man Durchschnitt und Summe die Rollen tauschen läßt, also die Systeme $\mathfrak{B}^\xi$ von Mengen B^ξ so definiert:

(β) *Die Mengen B^0 sind die Mengen M $(\mathfrak{B}^0 = \mathfrak{M})$;*
die Mengen B^η sind für ungerades η die Durchschnitte, für gerades $\eta > 0$ die Summen aus Folgen von Mengen B^ξ $(\xi < \eta)$.

Alle Mengen B^ξ bilden wieder das kleinste Borelsche System $\mathfrak{B}$ über $\mathfrak{M}$. Die $B^{\xi+1}$ sind (für gerades oder ungerades ξ) die Durchschnitte oder Summen aus Folgen von Mengen B^ξ. Den Anfang bilden die Mengen (4), d. h. die

$$B^0, \quad B^1, \quad B^2, \quad B^3, \ldots$$

sind mit den

$$M, \quad M_\delta, \quad M_{\delta\sigma}, \quad M_{\delta\sigma\delta}, \ldots$$

identisch. Auch diese Mengensysteme werden Borelsche Klassen genannt. Man erkennt leicht, daß jedes A^ξ ein $B^{\xi+1}$ und jedes B^ξ ein $A^{\xi+1}$ ist; z. B. erscheinen die $A^1 = M_\sigma$ mit unter den $B^2 = M_{\delta\sigma}$ und die $B^1 = M_\delta$ mit unter den $A^2 = M_{\sigma\delta}$.

[41] Natürlich kann es passieren, daß man in Wahrheit nicht *alle* Mengen A^ξ oder B^ξ wirklich braucht, um das Borelsche System zu vollenden, d. h. es kann sein, daß bereits ein $\mathfrak{A}^\xi$ das ganze Borelsche System ist. Das hängt vom Ausgangssystem $\mathfrak{M}$ ab; wenn dieses etwa selbst ein Borelsches System ist, so ist bereits $\mathfrak{A}^0 = \mathfrak{B}$ und jede Erweiterung überflüssig. Wir werden später (§ 33, I) sehen, daß gerade in den wichtigsten Fällen jeder Schritt neue Mengen liefert und alle A^ξ zur Bildung von $\mathfrak{B}$ erforderlich sind.

Bemerkt sei noch: wenn, für irgend ein $\xi \geqq 1$, $\mathfrak{A}^\xi = \mathfrak{A}^{\xi+1}$ sein sollte, so ist $\mathfrak{A}^\xi = \mathfrak{B}$. Denn von den beiden Mengensystemen $\mathfrak{A}^\xi, \mathfrak{A}^{\xi+1}$ ist das eine ein σ-System, das andere ein δ-System; fallen sie zusammen, so haben wir ein Borelsches System vor uns.

3. Darstellung der Borelschen Mengen. Wir sahen (um etwa an die zweite Form anzuknüpfen), daß jedes $B^{\xi+1}$ für gerades ξ Durchschnitt, für ungerades ξ Summe einer Folge von Mengen B^ξ ist. Ist η eine Limeszahl, so ist

$$(5) \qquad B^\eta = C_1 \dot{+} C_2 \dot{+} \cdots \qquad (C_n \text{ ein } B^{\xi_n} \text{ mit } \xi_n < \eta);$$

diese Form leidet an einer gewissen Willkürlichkeit, indem die ξ_n von der darzustellenden Menge B^η abhängen können, und wir wollen sie in eine solche überführen, wo die ξ_n eine feste, nur von η abhängige Folge bilden (wie dies für $\eta = \xi + 1$ der Fall war, wo $\xi_n = \xi = \eta - 1$ gewählt werden konnte). Beispielsweise: wir wollen eine Menge B^ω, die sich zunächst in der Form $B^\omega = B^{\xi_1} \dot{+} B^{\xi_2} \dot{+} \cdots$ mit beliebigen endlichen Indices ξ_n darstellt, in die spezielle Form $B^\omega = B^1 \dot{+} B^2 \dot{+} \cdots$ überführen, wo $\xi_n = n$ ist. Wählen wir eine feste, nur von der Limeszahl η abhängige Folge von Ordnungszahlen

$$1 \leqq \eta_1 < \eta_2 < \eta_3 < \cdots, \quad \lim \eta_n = \eta;$$

es soll dann gezeigt werden, daß jedes B^η in der speziellen Form

$$(6) \qquad B^\eta = B_1 \dot{+} B_2 \dot{+} \cdots \qquad (B_n \text{ ein } B^{\eta_n})$$

darstellbar ist. Um (5) in (6) überzuführen, können wir zunächst die Indices ξ_n vergrößern, da B^ξ ein $B^{\xi+1}$ usw. ist; und da für jedes $\xi < \eta$ ein n mit $\xi \leqq \eta_n$ existiert, so können wir die natürlichen Zahlen $p < q < r < \cdots$ so wählen, daß $\xi_1 \leqq \eta_p$, $\xi_2 \leqq \eta_q$, $\xi_3 \leqq \eta_r, \ldots$ Setzen wir dann $C_1 = B_p$, $C_2 = B_q$, $C_3 = B_r, \ldots$, so erhalten wir statt (5)

$$B^\eta = B_p \dot{+} B_q \dot{+} B_r \dot{+} \cdots,$$

wo für die hier vertretenen n jedes B_n ein B^{η_n} ist. Um auf die Form (6) zu kommen, hätten wir für die noch fehlenden $n = 1, \ldots, p-1, p+1, \ldots$ geeignete Summanden B_n einzuschalten. Nun enthält jedes B^ξ mit $\xi \geqq 1$ gewiß eine Teilmenge der Form $M_\delta = B^1$. Denn die Eigenschaft, eine solche Teilmenge zu enthalten, überträgt sich von irgendwelchen Mengen N stets auf die N_σ und N_δ (N_δ enthält eine Teilmenge der Form $M_{\delta\delta} = M_\delta$) und also von den B^1 auf die späteren B^ξ. Ist also B eine Teilmenge von B^η, von der Form B^1, so können wir die einzuschaltenden B_n sämtlich gleich B wählen; denn B_n ist als B^1 auch ein B^{η_n}. Damit ist die Form (6) hergestellt.

Denken wir uns jeder Mengenfolge $M_1, M_2, \ldots$ durch irgendeine Vorschrift eine Menge

$$(7) \qquad X = \Phi(M_1, M_2, M_3, \ldots)$$

eindeutig zugeordnet (z. B. $X = M_1 M_2 M_3 \ldots$ oder $X = M_1 \dotplus M_2 \dotplus M_3 \dotplus \cdots$), so daß also Φ eine eindeutige Funktion der Mengen M_n bedeutet. Wählt man auf alle möglichen Weisen die Mengen M_n aus einem gegebenen Mengensystem $\mathfrak{M}$, so durchläuft X ein Mengensystem $\mathfrak{X}$. Wir behaupten dann:

[42] I. *Für jedes $\xi \geqq 1$ läßt sich eine nur von ξ abhängige Funktion Φ_ξ so bestimmen, daß*

$$X = \Phi_\xi(M_1, M_2, M_3, \ldots) \qquad (M_n \varepsilon \mathfrak{M}) \tag{8}$$

genau die von $\mathfrak{M}$ erzeugten Borelschen Mengen B^ξ darstellt.

Daß X genau die B^ξ darstellt, soll natürlich heißen, daß X alle B^ξ und keine andern Mengen darstellt, oder daß X das Mengensystem $\mathfrak{X} = \mathfrak{B}^\xi$ durchläuft.

Der Beweis ist sehr einfach. Setzen wir

$$\Phi_1(M_1, M_2, M_3, \ldots) = M_1 M_2 M_3 \ldots, \tag{9}$$

so stellt $X = \Phi_1$ genau die $M_\delta = B^1$ dar.

Wenn ferner die Funktion Φ_ξ bereits definiert ist, so erhalten wir, gemäß der Bildung der $B^{\xi+1}$ aus den B^ξ, eine geeignete Funktion $\Phi_{\xi+1}$ auf folgende Weise. Wir spalten die Menge der natürlichen Zahlen in abzählbar viele abzählbare Teilmengen, etwa nach dem dyadischen Schema S. 30. Für ungerades ξ setzen wir dann

$$\left\{\begin{aligned} &\Phi_{\xi+1}(M_1, M_2, M_3, \ldots) \\ &= \Phi_\xi(M_1, M_3, M_5, \ldots) \dotplus \Phi_\xi(M_2, M_6, M_{10}, \ldots) \\ &\qquad \dotplus \Phi_\xi(M_4, M_{12}, M_{20}, \ldots) \dotplus \cdots \end{aligned}\right. \tag{10}$$

und für gerades ξ

$$\left\{\begin{aligned} &\Phi_{\xi+1}(M_1, M_2, M_3, \ldots) \\ &= \Phi_\xi(M_1, M_3, M_5, \ldots)\, \Phi_\xi(M_2, M_6, M_{10}, \ldots)\, \Phi_\xi(M_4, M_{12}, M_{20}, \ldots) \ldots \end{aligned}\right. \tag{11}$$

Für eine Limeszahl η endlich setzen wir, gemäß der Darstellung (6),

$$\left\{\begin{aligned} &\Phi_\eta(M_1, M_2, M_3, \cdots) \\ &= \Phi_{\eta_1}(M_1, M_3, M_5, \ldots) \dotplus \Phi_{\eta_2}(M_2, M_6, M_{10}, \ldots) \\ &\qquad \dotplus \Phi_{\eta_3}(M_4, M_{12}, M_{20}, \ldots) \dotplus \cdots. \end{aligned}\right. \tag{12}$$

Damit sind die Funktionen Φ_ξ induktiv definiert und es ist klar, daß sie das Verlangte leisten, denn die Trennung der Indices auf der rechten Seite der drei letzten Formeln gestattet, z. B. in (10) und (11) jedes Φ_ξ unabhängig von den übrigen zu einem beliebig vorgeschriebenen B^ξ zu machen, so daß $\Phi_{\xi+1}$ wirklich *alle* Mengen $B^{\xi+1}$ darstellt; das Gleiche gilt von (12). Z. B. liefert

$$\begin{aligned} &\Phi_2(M_1, M_2, M_3, \ldots) \\ &= (M_1 M_3 M_5 \ldots) \dotplus (M_2 M_6 M_{10} \ldots) \dotplus (M_4 M_{12} M_{20} \ldots) \dotplus \cdots \end{aligned}$$

für $M_n \varepsilon \mathfrak{M}$ genau die Mengen $M_{\delta\sigma} = B^2$.

Die hier auftretenden, durch (9) bis (12) induktiv erklärten Funktionen Φ_ξ sind nun alle [1]) von einer besonderen Gestalt, nämlich *Summen von Durchschnitten aus Teilfolgen der Mengenfolge* $M_1, M_2, \ldots$. Diese besonderen Funktionen Φ werden also so gebildet: jeder Folge

$$\nu = (n_1, n_2, n_3, \ldots)$$

wachsender natürlicher Zahlen ordnen wir den Durchschnitt

$$M_\nu = M_{n_1} M_{n_2} M_{n_3} \ldots$$

zu; sodann sei N eine Menge solcher Folgen und

$$X = \mathfrak{S}_{\nu}^{N} M_\nu = \Phi(M_1, M_2, M_3, \ldots), \tag{13}$$

wobei das Funktionszeichen Φ und die Menge N einander entsprechen. Wir wollen diese öfters auftretenden Funktionen als *δs-Funktionen* be- [43]
zeichnen. Wenn wir nämlich, wie bisher, Summen- und Durchschnittsbildung an *abzählbar* vielen Mengen mit σ und δ, an *beliebig* vielen Mengen mit s und d bezeichnen, so ist in (13) jeder Summand M_ν ein M_δ, die Summe selbst ein $M_{\delta s}$ (falls N abzählbar ist, ein $M_{\delta\sigma}$; falls N nur ein Element hat, ein M_δ). Durch Vertauschung der Rollen von Summe und Durchschnitt würde man in analoger Weise *σd-Funktionen* erhalten, Durchschnitte von Summen aus Teilfolgen einer Mengenfolge; die Komplemente der Mengen (13) liefern ein Beispiel.

Wir haben nun behauptet, daß die Funktionen Φ_ξ solche δs-Funktionen (13) sind, müssen also zeigen: es gibt für jedes $\xi \geqq 1$ eine nur von ξ abhängige Menge N_ξ derart, daß

$$\Phi_\xi(M_1, M_2, M_3, \ldots) = \mathfrak{S}_{\nu}^{N_\xi} M_\nu. \tag{14}$$

Z. B. besteht N_1 aus der einzigen Folge $(1, 2, 3, \ldots)$; N_2 aus den Folgen $(1, 3, 5, \ldots)$ $(2, 6, 10, \ldots)$ $(4, 12, 20, \ldots)$ usw. Sei nun (14) für ein bestimmtes ξ richtig; dann ist

$$\Phi_\xi(M_1, \ M_3, \ M_5, \ldots) = \mathfrak{S}_{\alpha}^{A_\xi} M_\alpha$$

$$\Phi_\xi(M_2, \ M_6, \ M_{10}, \ldots) = \mathfrak{S}_{\beta}^{B_\xi} M_\beta$$

$$\Phi_\xi(M_4, \ M_{12}, \ M_{20}, \ldots) = \mathfrak{S}_{\gamma}^{\Gamma_\xi} M_\gamma,$$

$$\ldots$$

wobei jedem $\nu = (n_1, n_2, n_3, \ldots)$ eine Folge $\alpha = (2n_1 - 1, 2n_2 - 1, 2n_3 - 1, \ldots)$ aus lauter ungeraden Zahlen und damit der Menge N_ξ eine bestimmte Menge A_ξ entspricht; analog sind $\beta, B_\xi, \gamma, \Gamma_\xi, \ldots$ aufzufassen. Gilt dann (10), so ist

---

[1]) In den Satz I könnte man, mit $\Phi_0(M_1, M_2, M_3, \ldots) = M_1$, auch die Mengen B^0 einbeziehen, aber für diese gilt das Folgende nicht.

$$\Phi_{\xi+1}(M_1, M_2, M_3, \ldots) = \mathop{\mathfrak{S}}_{\nu}^{N_{\xi+1}} M_\nu, \qquad N_{\xi+1} = A_\xi + B_\xi + \Gamma_\xi + \cdots,$$

und analog wird im Falle der Formel (12) die Menge N_η gefunden. Bei der Durchschnittsbildung (11) tritt zunächst, nach dem distributiven Gesetz, die Summe [1]

$$\mathfrak{S}\, M_\alpha M_\beta M_\gamma \ldots$$

auf, erstreckt über $\alpha \varepsilon A_\xi, \beta \varepsilon B_\xi, \gamma \varepsilon \Gamma_\xi, \ldots$; d. h. die Folge $(\alpha, \beta, \gamma, \ldots)$ von Folgen durchläuft das Mengenprodukt $(A_\xi, B_\xi, \Gamma_\xi, \ldots)$. Jene Folge von Folgen bestimmt, einfach durch Vereinigung der disjunkten Folgen $\alpha, \beta, \gamma, \ldots$, eine einzige Folge ν, und das Mengenprodukt bestimmt in diesem Sinne eine Menge $N_{\xi+1}$ von Folgen ν, wonach

$$\Phi_{\xi+1}(M_1, M_2, M_3, \ldots) = \mathop{\mathfrak{S}}_{\nu}^{N_{\xi+1}} M_\nu$$

wird.

Damit ist also folgende präzisere Form des Satzes I bewiesen:

II. *Für jedes $\xi \geq 1$ läßt sich eine nur von ξ abhängige Menge N_ξ, deren Elemente $\nu = (n_1, n_2, n_3, \ldots)$ Folgen wachsender natürlicher Zahlen sind, so bestimmen, daß die δs-Funktion*

$$(15) \qquad X = \mathop{\mathfrak{S}}_{\nu}^{N_\xi} M_\nu = \mathop{\mathfrak{S}}_{\nu}^{N_\xi} M_{n_1} M_{n_2} M_{n_3} \ldots$$
$$= \Phi_\xi(M_1, M_2, M_3, \ldots) \qquad (M_n \varepsilon \mathfrak{M})$$

genau die von $\mathfrak{M}$ erzeugten Borelschen Mengen B^ξ darstellt.

Eine analoge Behandlung gestatten natürlich auch die Borelschen Mengen A^ξ (S. 85); die Φ_ξ würden dann nicht Summen von Durchschnitten, sondern Durchschnitte von Summen aus Teilfolgen der Folge M_n sein, d. h. σd-Funktionen.

§ 19. Die Suslinschen Mengen.

Beim Anblick des Satzes II im vorigen Paragraphen liegt die Frage nahe, ob es vielleicht eine *feste* δs-Funktion oder eine feste Menge N von Folgen $\nu = (n_1, n_2, \ldots)$ wachsender natürlicher Zahlen gibt, derart, daß die Menge

$$(1) \quad X = \mathop{\mathfrak{S}}_{\nu}^{N} M_\nu = \mathop{\mathfrak{S}}_{\nu}^{N} M_{n_1} M_{n_2} M_{n_3} \ldots = \Phi(M_1, M_2, M_3, \ldots) \qquad (M_n \varepsilon \mathfrak{M})$$

[44] *genau alle* von $\mathfrak{M}$ erzeugten Borelschen Mengen durchläuft. Diese Frage ist übrigens zu verneinen. Wenn man aber den Zusatz „genau" streicht, ist sie bejahend zu beantworten: es gibt eine Darstellung (1), die alle Borelschen Mengen, darüber hinaus aber im allgemeinen noch andere liefert. Wir kommen darauf am Schluß dieses Paragraphen zurück und behandeln

[1]) Die Menge der Summanden ist schon bei den B^3 unabzählbar.

die fraglichen Mengen zunächst in einer andern Bezeichnung, in der statt der Mengen M_n Mengen zwar mit mehrfachen, aber unabhängig voneinander variierenden Indices auftreten.

Wir ordnen einmal nicht den natürlichen Zahlen, sondern den (ebenfalls in abzählbarer Menge vorhandenen) *endlichen Komplexen* (n_1), (n_1, n_2), (n_1, n_2, n_3), ... natürlicher Zahlen Mengen $M_{n_1}, M_{n_1 n_2}, M_{n_1 n_2 n_3}, \ldots$ zu; bis auf die äußere Form ist das auch nichts anderes als eine Mengenfolge. Geben wir (ohne Punkte zur Andeutung weiterer Mengen, um das Schema nicht unübersichtlich zu machen) einige dieser Mengen an:

$$(2) \qquad \begin{cases} M_1 & & & & M_2 & & & \\ M_{11} & & M_{12} & & M_{21} & & M_{22} & \\ M_{111} & M_{112} & M_{121} & M_{122} & M_{211} & M_{212} & M_{221} & M_{222}. \end{cases}$$

Hiernach ordnen wir jeder Folge

$$(3) \qquad \nu = (n_1, n_2, n_3, \ldots)$$

natürlicher Zahlen, die jetzt nicht paarweise verschieden zu sein brauchen, den Durchschnitt

$$(4) \qquad M_\nu = M_{n_1} M_{n_1 n_2} M_{n_1 n_2 n_3} \cdots$$

zu, gebildet aus den Mengen, die den endlichen Abschnitten der Folge ν entsprechen, und bilden die über *alle* Zahlenfolgen ν erstreckte Summe

$$(5) \qquad X = \underset{\nu}{\mathfrak{S}} M_\nu .$$

Es ist das also eine Funktion $\Phi(M_1, M_{11}, M_2, \ldots)$ der erzeugenden Mengen (2). Werden diese auf alle möglichen Weisen einem Mengensystem $\mathfrak{M}$ entnommen, so durchläuft X ein Mengensystem $\mathfrak{X}$, und diese Mengen X heißen die von $\mathfrak{M}$ (oder von den Mengen $M \varepsilon \mathfrak{M}$) erzeugten *Suslinschen* [45]
Mengen. Die Summanden M_ν sind dann zwar Mengen $M_{\mathfrak{d}}$, aber ihre Menge ist unabzählbar, wie dies auch bei den Borelschen Mengen B^ξ für $\xi \geqq 3$ der Fall war. Man beachte übrigens, daß X nicht etwa mit dem Durchschnitt

$$\mathfrak{S} M_{n_1} \cdot \mathfrak{S} M_{n_1 n_2} \cdot \mathfrak{S} M_{n_1 n_2 n_3} \cdots$$

(der ein $M_{\sigma\mathfrak{d}}$ ist) übereinstimmt, denn dieser ist, nach dem distributiven Gesetz entwickelt, gleich

$$\mathfrak{S} M_{a_1} M_{b_1 b_2} M_{c_1 c_2 c_3} \cdots,$$

wo alle Indices unabhängig voneinander alle natürlichen Zahlen durchlaufen, während in unserem Falle $a_1 = b_1 = c_1 = \cdots = n_1$ usw. sein soll.

Nennen wir die von $\mathfrak{M}$ erzeugten Suslinschen Mengen etwa Mengen M_S, das von ihnen gebildete System $\mathfrak{M}_S$.

Jedes M_σ und $M_{\mathfrak{d}}$ ist ein M_S; die Suslinsche Formel umfaßt also Summen- und Durchschnittsbildung aus Folgen. Denn ist eine Folge von Mengen $M^1, M^2, \ldots$ gegeben und setzt man $M_{n_1 n_2 \ldots n_k} = M^{n_1}$, so ist $M_\nu = M^{n_1}$ und $X = \mathfrak{S} M^{n_1}$; setzt man andererseits $M_{n_1 n_2 \ldots n_k} = M^k$, so ist jedes M_ν und demnach X selbst gleich $\mathfrak{D} M^k$.

Die fundamentalste Eigenschaft der Suslinschen Mengen spricht aber der folgende Satz aus:

I. *Die von Suslinschen Mengen M_S erzeugten Suslinschen Mengen sind wieder Mengen M_S.*

Oder kurz: jede Menge M_{SS} ist nur eine Menge M_S (wie jedes $M_{\sigma\sigma}$ ein M_σ ist). Der Suslinsche Prozeß ist derart umfassend, daß seine Iteration nichts Neues liefert.

Der Beweis sieht nur wegen der vielen Indizes schwierig aus, ist aber ganz einfach. Es sei

$$P = \underset{\nu}{\mathfrak{S}} N^{n_1} N^{n_1 n_2} N^{n_1 n_2 n_3} \ldots$$

eine Suslinsche Menge, deren erzeugende Mengen selbst wieder Suslinsche Mengen sind:

$$\begin{aligned} N^{n_1} &= \underset{\alpha}{\mathfrak{S}} M_{a_1}^{n_1} \quad M_{a_1 a_2}^{n_1} \quad M_{a_1 a_2 a_3}^{n_1} \ldots \\ N^{n_1 n_2} &= \underset{\beta}{\mathfrak{S}} M_{b_1}^{n_1 n_2} \quad M_{b_1 b_2}^{n_1 n_2} \quad M_{b_1 b_2 b_3}^{n_1 n_2} \ldots \\ N^{n_1 n_2 n_3} &= \underset{\gamma}{\mathfrak{S}} M_{c_1}^{n_1 n_2 n_3} M_{c_1 c_2}^{n_1 n_2 n_3} M_{c_1 c_2 c_3}^{n_1 n_2 n_3} \ldots \\ &\ldots \end{aligned}$$

wobei $\nu = (n_1, n_2, \ldots)$, $\alpha = (a_1, a_2, \ldots)$, $\beta = (b_1, b_2, \ldots)$, ... unabhängig voneinander alle natürlichen Zahlenfolgen durchlaufen. Es ist zu zeigen, daß P als Suslinsche Menge aus den M selber gebildet werden kann. Nach dem distributiven Gesetz ist P die Summe aller Durchschnitte

$$(6) \qquad M_{a_1}^{n_1} M_{a_1 a_2}^{n_1} M_{a_1 a_2 a_3}^{n_1} \ldots M_{b_1}^{n_1 n_2} M_{b_1 b_2}^{n_1 n_2} \ldots M_{c_1}^{n_1 n_2 n_3} \ldots,$$

die Summe erstreckt über alle Folgen $\nu, \alpha, \beta, \ldots$ oder alle natürlichen Zahlen $n_k, a_k, b_k, \ldots$. Wir verwandeln, etwa wieder nach dem dyadischen Schema, die Folge der Folgen $\nu, \alpha, \beta, \gamma, \ldots$ in eine einzige Folge $\mu = (m_1, m_2, \ldots)$, d. h. so, daß die Matrizen

$$\begin{array}{cccc|cccc} n_1 & n_2 & n_3 & \ldots & m_1 & m_3 & m_5 & \ldots \\ a_1 & a_2 & a_3 & \ldots & m_2 & m_6 & m_{10} & \ldots \\ b_1 & b_2 & b_3 & \ldots & m_4 & m_{12} & m_{20} & \ldots \\ c_1 & c_2 & c_3 & \ldots & m_8 & m_{24} & m_{40} & \ldots \\ . & . & . & \ldots & . & . & . & \ldots \end{array}$$

einander gleichgesetzt werden. Hierdurch geht (6) in

$$(7) \qquad \begin{aligned} &M_{m_2}^{m_1} M_{m_2 m_6}^{m_1} M_{m_2 m_6 m_{10}}^{m_1} \ldots M_{m_4}^{m_1 m_3} M_{m_4 m_{12}}^{m_1 m_3} \ldots M_{m_8}^{m_1 m_3 m_5} \ldots \\ &= M(1)\, M(3) \quad M(5) \quad \ldots M(2) \quad M(6) \quad \ldots \quad M(4) \quad \ldots \end{aligned}$$

über, wobei $M(k)$ eine nur von $m_1, m_2, \ldots, m_{2k}$ (im allgemeinen aber nicht von allen diesen Zahlen) abhängige Menge bedeutet und k alle natürlichen Zahlen einmal durchläuft; denn es ist leicht zu sehen, daß die oberen Zahlen $m_1, m_3, \ldots, m_{2h-1}$, von denen eine Menge in (7) abhängt, kleinere Indices haben als die unteren, die mit m_{2h} beginnen, und daß also das höchste m_k, von dem die Menge abhängt, geraden Index hat. Ordnen wir nun schließlich die Zahlenpaare (m_{2k-1}, m_{2k}) den natürlichen Zahlen p_k eineindeutig zu und definieren die von $p_1, \ldots, p_k$ oder $m_1, \ldots, m_{2k}$ abhängige Menge

$$\dot{M}_{p_1 p_2 \ldots p_k} = M(k),$$

so geht (7) in $M_{p_1} M_{p_1 p_2} M_{p_1 p_2 p_3} \ldots$ über und P ist die Summe dieser Durchschnitte, über alle Folgen $(p_1, p_2, p_3, \ldots)$ natürlicher Zahlen erstreckt, d. h. eine aus den M gebildete Suslinsche Menge.

Damit ist I bewiesen. Die Mengen $N = M_S$ erzeugen keine neuen Suslinschen Mengen; jedes N_S ist ein N. Insbesondere ist also jedes N_σ und N_δ ein N, die Mengen N bilden ein Borelsches System $\mathfrak{N}$ über $\mathfrak{M}$; das kleinste Borelsche System $\mathfrak{B}$ ist demnach in $\mathfrak{N}$ enthalten: *alle von $\mathfrak{M}$ erzeugten Borelschen Mengen sind auch Suslinsche Mengen.* Das Umgekehrte ist im Allgemeinen nicht richtig, wie wir später (§ 33, I) sehen werden.

Schließlich können wir von der Suslinschen Formel (5)(4) leicht zu einer Darstellung der Gestalt (1) zurückgehen und damit die Frage zu Beginn dieses Paragraphen beantworten. Wir brauchen nur eine eineindeutige Zuordnung zwischen den natürlichen Zahlen p und den endlichen Komplexen natürlicher Zahlen $(n_1, n_2, \ldots, n_k)$ herzustellen, etwa die aus der dyadischen Schreibweise entspringende

$$(8) \qquad p = 2^{n_1 - 1} + 2^{n_1 + n_2 - 1} + \cdots + 2^{n_1 + n_2 + \cdots + n_k - 1};$$

sodann setzen wir

$$M^p = M_{n_1 n_2 \ldots n_k}.$$

Dann entspricht jeder Folge $\nu = (n_1, n_2, n_3, \ldots)$ beliebiger natürlicher Zahlen eine Folge $\pi = (p_1, p_2, p_3, \ldots)$ wachsender natürlicher Zahlen der Gestalt

$$p_1 = 2^{n_1 - 1}, \quad p_2 = 2^{n_1 - 1} + 2^{n_1 + n_2 - 1},$$
$$p_3 = 2^{n_1 - 1} + 2^{n_1 + n_2 - 1} + 2^{n_1 + n_2 + n_3 - 1}, \ldots$$

(d. h. die Zahlen $p_1, p_2 - p_1, p_3 - p_2, \ldots$ bilden eine Teilfolge von 1, 2, 4, 8, ...); ist Π die Menge aller dieser Folgen π, so geht (5) in

$$X = \underset{\pi}{\overset{\Pi}{\mathfrak{S}}} M^{p_1} M^{p_2} M^{p_3} \ldots$$

über. Damit ist, wenn wir zur Bezeichnung (1) zurückkehren, der Satz bewiesen:

II *Es gibt eine feste Menge N von Folgen $\nu = (n_1, n_2, n_3, \ldots)$ wachsender natürlicher Zahlen derart, daß die δs-Funktion*

$$(1) \qquad X = \underset{\nu}{\overset{N}{\mathfrak{S}}} M_\nu = \underset{\nu}{\overset{N}{\mathfrak{S}}} M_{n_1} M_{n_2} M_{n_3} \ldots$$
$$= \Phi(M_1, M_2, M_3, \ldots) \qquad (M_n \varepsilon \mathfrak{M})$$

genau die von $\mathfrak{M}$ erzeugten Suslinschen Mengen darstellt.

Sechstes Kapitel.
Punktmengen.
§ 20. Entfernung.

Wir haben bisher teils reine Mengen, teils solche mit Relationen zwischen den Elementen (geordnete Mengen) betrachtet. Der zweiten Klasse gehören auch die Mengen an, die wir nun untersuchen wollen: Mengen, in denen zwei Elemente eine Entfernung haben.

[46] **1. Metrische Räume.** Es sei E eine Menge, deren Elemente wir jetzt *Punkte* nennen. Jedem Punktpaar (x, y) sei eine reelle Zahl xy, die *Entfernung* beider Punkte, zugeordnet [1]), also eine reelle Funktion $xy = f(x, y)$ in (E, E) definiert. Hierfür sollen die folgenden *Entfernungsaxiome* oder *-postulate* gelten:

(α) $xx = 0$
(β) $xy = yx > 0$ für $x \neq y$
(γ) $xy + yz \geqq xz.$

Das Axiom (γ), das wichtigste unter ihnen, heiße das *Dreiecksaxiom* oder die Dreiecksungleichung (die Summe von zwei Dreiecksseiten ist mindestens gleich der dritten). Die Menge E wird als metrische Menge, *Punktmenge* oder *metrischer Raum* bezeichnet; bei der letzten Bezeichnung wird vorwiegend an das Verhältnis von E zu seinen Teilmengen und Punkten, bei den ersten beiden an das Enthaltensein von E in einem umfassenden Raume gedacht.

Das nächstliegende Beispiel ist die Menge E_1 der reellen Zahlen, wo als Entfernung der absolute *Betrag* $|x - y|$ des Unterschiedes beider Zahlen definiert wird. Sodann der *n-dimensionale Euklidische Zahlenraum* E_n; seine Elemente sind die Komplexe

$$x = (x_1, x_2, \ldots, x_n)$$

aus n reellen Zahlen; die Gleichheit ist wie üblich ($x = y$ so viel wie $x_1 = y_1, \ldots, x_n = y_n$) und die Entfernung durch

$$xy = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2}$$

definiert, die Wurzel $\geqq 0$ genommen. Die Postulate $(\alpha)(\beta)$ sind erfüllt, auf den Beweis von (γ) kommen wir zurück.

Wenn zwei Räume *eineindeutig* und *entfernungstreu* aufeinander abgebildet werden können (d. h. die Punkte x den Punkten ξ eineindeutig so entsprechen, daß $xy = \xi\eta$), so werden sie *isometrisch* genannt. Z. B. ist der E_1 mit der Menge der Punkte $(x, 0)$ des E_2 isometrisch. Isometrie, im

[1]) Wenn Verwechselung mit einem Produkt in Frage kommt, wird statt xy eine andere Bezeichnung zu wählen sein.

Grunde nichts anderes als die Kongruenz in der Elementargeometrie, ist für metrische Mengen ein analoger Begriff wie Äquivalenz für reine, Ähnlichkeit für geordnete Mengen (ein anderes, wichtigeres Analogon ist Homöomorphie, § 38); indessen ist es nicht nötig, einen der Kardinalzahl und dem Ordnungstypus entsprechenden Namen einzuführen. Zwei isometrische Räume, jeder im Verhältnis zu seinen Punkten und Teilmengen betrachtet, können einfach als identisch angesehen werden (nicht aber zwei isometrische Mengen in einem umfassenden Raum).

Ein dem E_n isometrischer Raum heißt ein *n-dimensionaler Euklidischer Raum*; $x_1, \ldots, x_n$ heißen rechtwinklige Cartesische Koordinaten des Punktes ξ, der dem Komplex x entspricht. Die isometrische Abbildung ist auf unendlich viele Weisen möglich (Koordinatentransformation, orthogonale Substitution). Unser idealisierter Anschauungs- und Erfahrungsraum ist ein dreidimensionaler Euklidischer Raum, seine Ebenen und Geraden sind zwei- und eindimensionale Euklidische Räume; hieraus entspringen bekannte geometrische Ausdrucksweisen auch für den E_n und andere Räume.

2. Lineare Räume. Die Punkte des Euklidischen Zahlenraums waren endliche Komplexe reeller Zahlen. Dies Verfahren läßt sich sofort verallgemeinern: wir ordnen jedem Element m einer beliebigen Menge

$$M = \{m, n, \ldots\}$$

eine reelle Zahl x_m zu (oder definieren in M eine reelle Funktion), wodurch wir den Komplex oder Punkt

$$x = (x_m, x_n, \ldots)$$

mit den „Koordinaten" $x_m, x_n, \ldots$ erhalten. Gleichheit der Punkte bedeutet Gleichheit sämtlicher Koordinaten ($x = y$ so viel wie $x_m = y_m$ für jedes $m \varepsilon M$). Der Punkt

$$0 = (0, 0, \ldots)$$

heißt der Nullpunkt. Wir definieren Multiplikation eines Punktes mit einer reellen Zahl α und Addition von Punkten durch

$$\alpha x = (\alpha x_m, \ \alpha x_n, \ldots)$$
$$x + y = (x_m + y_m, \ x_n + y_n, \ldots),$$

wonach die Bedeutung von Zeichen wie $-x$, $x - y$, $\alpha x + \beta y$, $\alpha x + \beta y + \gamma z$ klar ist.

Wir sagen, drei Punkte x, y, z liegen in gerader Linie oder sind kollinear, wenn es drei reelle, nicht sämtlich verschwindende Zahlen α, β, γ gibt mit
$$\alpha x + \beta y + \gamma z = 0, \quad \alpha + \beta + \gamma = 0.$$
Dabei kann γ nicht 0 sein außer für $x = y$ (dann liegt *jeder* Punkt z mit x, y in gerader Linie); für $x \neq y$ kann, da nur die Verhältnisse $\alpha : \beta : \gamma$ in Betracht kommen, $\gamma = -1$ gewählt werden, und es liefert also

$$(1) \qquad z = \alpha x + \beta y, \quad \alpha + \beta = 1$$

alle mit x, y kollinearen Punkte z, deren Menge die durch x, y bestimmte *Gerade* heißt; z hängt von einer reellen Variablen α oder β ab. Wird zu (1) die Beschränkung $\alpha \geqq 0, \beta \geqq 0$ hinzugefügt, so erhält man die durch x, y bestimmte (abgeschlossene) *Strecke* $[x, y]$.

Man kann dies fortsetzen. Nennen wir $k+1$ Punkte $x_0, x_1, \ldots, x_k$ *linear abhängig*, wenn es eine Beziehung

$$\alpha_0 x_0 + \alpha_1 x_1 + \cdots + \alpha_k x_k = 0, \quad \alpha_0 + \alpha_1 + \cdots + \alpha_k = 0$$

mit reellen Zahlen $\alpha_0, \alpha_1, \ldots, \alpha_k$ gibt, die nicht sämtlich verschwinden; andernfalls *linear unabhängig*. Sind $x_0, x_1, \ldots, x_k$ linear abhängig, jedoch $x_1, \ldots, x_k$ linear unabhängig, so kann α_0 nicht 0 sein, und man erhält, indem man $\alpha_0 = -1$ setzt, alle von $x_1, \ldots, x_k$ linear abhängigen Punkte x_0 in der Form

$$x_0 = \alpha_1 x_1 + \ldots + \alpha_k x_k, \quad \alpha_1 + \cdots + \alpha_k = 1;$$

sie hängen von $k-1$ reellen Parametern ab und bilden, wie man sich ausdrückt, den durch $x_1, \ldots, x_k$ bestimmten $(k-1)$-*dimensionalen Raum* R_{k-1}, bei Beschränkung auf $\alpha_1 \geqq 0, \ldots, \alpha_k \geqq 0$ das durch jene Punkte bestimmte $(k-1)$-*dimensionale Simplum* S_{k-1}. (R_1 Gerade, R_2 Ebene, S_1 Strecke, S_2 Dreieck, S_3 Tetraeder.)

Eine aus unseren Punkten x gebildete Menge E heiße *linear*[1]), wenn sie mit zwei verschiedenen Punkten x, y auch die ganze durch sie bestimmte Gerade enthält; sie heiße *konvex*, wenn sie mit x, y auch die Strecke $[x, y]$ enthält. So ist die oben erklärte Menge R_{k-1} linear, S_{k-1} konvex; einen linearen Raum erhält man auch, wenn man alle Punkte nimmt, deren Koordinaten ein System linearer Gleichungen erfüllen. Der Durchschnitt beliebig vieler linearer oder konvexer Mengen ist offenbar wieder linear oder konvex. Ist E ein linearer oder konvexer Raum, $A \leqq E$, so ist der Durchschnitt K aller konvexen Mengen $\geqq A$ die *kleinste konvexe Menge* $\geqq A$ und wird als *konvexe Hülle* von A bezeichnet [2]). K enthält mit irgendwelchen linear unabhängigen Punkten von A das ganze durch sie bestimmte Simplum und ist einfach die Summe aller dieser Simpla.

Ist a ein fester Punkt, so werden durch

$$\xi = x + a, \quad x = \xi - a \tag{2}$$

die Punkte x, ξ eineindeutig aufeinander abgebildet; man nennt diese Beziehung eine *Schiebung* oder Translation. Dabei geht jeder lineare Raum wieder in einen solchen über und kann insbesondere in einen solchen (homo-

---

[1]) Bei Cantor u. a. heißen die Mengen auf der Geraden E_1 linear.

[2]) Kleinste Mengen $\geqq A$ werden *Hüllen*, größte Mengen $\leqq A$ *Kerne* von A genannt, mit einem kennzeichnenden Adjektiv (konvexe, vollständige, abgeschlossene Hülle; offener, insichdichter Kern). Die Existenz solcher größter und kleinster Mengen von irgendwelcher Eigenschaft muß natürlich bewiesen werden.

gen linearen) übergehen, *der den Nullpunkt enthält*, was wir nun voraussetzen wollen. Ein solcher Raum E enthält zu jedem Punkt x auch alle Punkte $\alpha x = \alpha x + (1-\alpha) . 0$, und zu zwei Punkten x, y alle linearen Kombinationen $\alpha x + \beta y$ (nicht nur die mit $\alpha + \beta = 1$). Er wird durch die Schiebungen (2), wenn a ein Punkt von E selbst ist, in sich transformiert. Wir wollen E nun in der Weise zu einem metrischen Raum machen, daß die Schiebungen isometrische Abbildungen werden, d. h. daß die Punkte x, y dieselbe Entfernung haben wie $x + a$, $y + a$, insbesondere wie $x - y$, 0 oder 0, $y - x$. Dadurch werden alle Entfernungen auf Entfernungen vom Nullpunkte zurückgeführt; bezeichnet man die Entfernung des Punktes x vom Nullpunkt mit $|x|$ und nennt sie den *Betrag* von x, so ist $xy = |x - y|$ die Entfernung zweier beliebiger Punkte. Man nennt ein geordnetes Punktpaar (x, y) auch einen *Vektor* und definiert die Gleichheit zweier Vektoren (x, y), (ξ, η) durch $y - x = \eta - \xi$; Punktpaare, die gleiche Vektoren liefern, haben also auch gleiche Entfernung (den Betrag oder die Länge des Vektors).

Die reelle Funktion $|x|$ von x muß nun den drei *Betragsaxiomen* genügen, die den Entfernungsaxiomen entsprechen:

(α_0) $|0| = 0$,

(β_0) $|x| = |-x| > 0$ für $x \neq 0$,

(γ_0) $|x + y| \leqq |x| + |y|$.

Das letzte, das wir das *Summenaxiom* nennen, ergibt sich aus dem Dreiecksaxiom (γ), indem man x, y, z durch x, 0, $-y$ ersetzt. Umgekehrt, wenn $|x|$ diesen Betragspostulaten genügt, so genügt

$$xy = |x - y| \tag{3}$$

den Entfernungspostulaten; (γ) folgt aus (γ_0) wegen

$$|x - z| = |(x - y) + (y - z)| \leqq |x - y| + |y - z|.$$

Die in der folgenden Ziffer 3 vorkommenden Beträge werden übrigens noch die Bedingung

$$|\alpha x| = |\alpha| |x|$$

erfüllen (α eine reelle Zahl), d. h. positiv-homogene Funktionen ersten Grades der Koordinaten sein. In diesem Falle werden, wie auch sonst der Betrag erklärt sei, die Geraden unseres Raumes mit E_1 oder mit Euklidischen Geraden isometrisch. Denn für zwei Punkte

$$z = x + \beta (y - x)$$

der Geraden (1) ist

$$z_1 - z = (\beta_1 - \beta)(y - x), \quad |z_1 - z| = |\beta_1 - \beta| |y - x|,$$

die Punkte z stehen also mit den Zahlen $c \beta$ ($c = |y - x|$ konstant) in isometrischer Beziehung.

3. Beispiele linearer Räume. Der Euklidische Raum E_n entsteht, wenn man den Betrag des Punktes

$$x = (x_1, x_2, \ldots, x_n) \tag{4}$$

durch

$$|x| = (x_1^2 + x_2^2 + \cdots + x_n^2)^{\frac{1}{2}} \tag{5}$$

erklärt. Der Beweis des Summenpostulats ergibt sich am kürzesten durch den bekannten Schluß: da die quadratische Form der reellen Variablen u, v

$$\Sigma(x_k u + y_k v)^2 = a u^2 + 2 b u v + c v^2$$

mit
$$a = \Sigma x_k^2, \quad b = \Sigma x_k y_k, \quad c = \Sigma y_k^2$$

nicht negativ ist, ist ihre Determinante $\geqq 0$,

$$\begin{aligned} b^2 &\leqq a c, \\ b &\leqq a^{\frac{1}{2}} c^{\frac{1}{2}}, \\ a + 2b + c &\leqq (a^{\frac{1}{2}} + c^{\frac{1}{2}})^2, \\ (a + 2b + c)^{\frac{1}{2}} &\leqq a^{\frac{1}{2}} + c^{\frac{1}{2}}, \end{aligned}$$

was eben die zu beweisende Ungleichung ist.

Wenn man n unbegrenzt wachsen läßt, so besagt die Summenungleichung zunächst, daß mit $x_1^2 + x_2^2 + \cdots$, $y_1^2 + y_2^2 + \cdots$ auch $(x_1 + y_1)^2 + (x_2 + y_2)^2 + \cdots$ konvergiert. Betrachten wir also jetzt reelle Zahlenfolgen

$$x = (x_1, x_2, \ldots), \tag{6}$$

d. h. Komplexe mit dem Träger $M = \{1, 2, \ldots\}$, und ist H die Menge derer, die *konvergente Quadratsumme* haben, so ist H ein linearer Raum (neben x, y gehören ihm auch $x + y$, $\alpha x + \beta y$ an), der durch die Betragsdefinition

$$|x| = (x_1^2 + x_2^2 + \cdots)^{\frac{1}{2}} \tag{7}$$

und die entsprechende Entfernungsdefinition (3) metrisch wird. Er heißt der *Hilbertsche Raum* und ist in gewissem Sinne ein $\aleph_0$-dimensionaler Euklidischer Raum. Übrigens läßt sich zwischen die Euklidischen Räume E_n und den Hilbertschen noch ein anderer einschieben, gebildet von den Zahlenfolgen x mit nur *endlich vielen* von Null verschiedenen Zahlen $x_1, x_2, \ldots$; man kann diesen Raum als $E_1 \dot{+} E_2 \dot{+} \cdots$ bezeichnen, wenn man nämlich die mit E_n isometrische Menge der Zahlenfolgen

$$(x_1, \ldots, x_n, 0, 0, \ldots)$$

mit E_n identifiziert.

Dies Verfahren läßt sich auf jede beliebige Menge M, als Träger der Komplexe, ausdehnen, wobei es übrigens, wenn man isometrische Räume nicht unterscheidet, nur auf die Mächtigkeit von M ankommt. Man behalte nur diejenigen $x = (x_m, \ldots)$ bei, in denen nur *endlich viele* bzw. *höchstens abzählbar viele* x_m, letzterenfalls *mit konvergenter Quadratsumme*, von Null verschieden sind, und definiere den Betrag $|x|$ wie oben. Diese beiden Räume können als Euklidische Räume, deren Dimensionenzahl die Mächtigkeit von M ist, bezeichnet werden.

Wenn man nicht an der Quadratsumme $x_1^2 + \cdots + x_n^2$, sondern an dem Quadratmittel $\frac{1}{n}(x_1^2 + \cdots + x_n^2)$ den Grenzübergang $n \to \infty$ vollzieht, so kommt man auf *Integrale*. Wenn z. B. M das Intervall $a \leqq t \leqq b$ bedeutet, so wird die Menge der in diesem Intervall stetigen Funktionen

$$(8) \qquad x = x(t)$$

durch die Betragsdefinition

$$(9) \qquad |x| = \left[\int_a^b x(t)^2\, dt\right]^{\frac{1}{2}}$$

und die entsprechende Entfernungsdefinition (3) zu einem metrischen linearen Raume. Dehnt man dies auf Integrale unstetiger Funktionen aus, etwa zunächst auf eigentliche Riemannsche Integrale, so tritt eine Erscheinung auf, die uns auch später gelegentlich wieder begegnen wird: es kann $|x| = 0$ sein, ohne daß im ursprünglich (S. 95) definierten Sinn $x = 0$ ($x(t) = 0$ im ganzen Intervall) ist. Um die Axiome $(\beta)(\beta_0)$ aufrechtzuerhalten, muß man die ursprüngliche Erklärung modifizieren und die [47] Gleichheit $x = y$ eben durch $xy = 0$ definieren. Anders gesagt: nennt man eine Funktion x mit $|x| = 0$ eine Nullfunktion, so bedeuten zwei Funktionen, die sich nur um eine Nullfunktion unterscheiden, denselben Punkt des metrischen Funktionenraumes.

Der Euklidische Raum, dessen Spielarten und Grenzfälle wir bisher betrachtet haben, beruhte auf der Betragsdefinition (5). In dieser kann man den Exponenten 2, der dem Pythagoreischen Lehrsatz entspricht, durch irgendeinen Exponenten $p > 1$ (der nicht ganzzahlig zu sein braucht) ersetzen, also den Betrag des Komplexes

$$(4) \qquad x = (x_1, x_2, \ldots, x_n)$$

durch

$$(10) \qquad |x| = (|x_1|^p + |x_2|^p + \cdots + |x_n|^p)^{\frac{1}{p}} \qquad (p > 1)$$

definieren. Die Summenungleichung, hier *Minkowskische Ungleichung* genannt, beweisen wir so. Für positive x, ξ gibt die Taylorsche Formel

$$x^p = \xi^p + p\,\xi^{p-1}(x - \xi) + \frac{p(p-1)}{2}\,\xi_1^{p-2}(x - \xi)^2$$

(ξ_1 zwischen x, ξ), also bei Weglassung des nichtnegativen Restgliedes

$$\frac{x^p}{\xi^{p-1}} \geqq \xi + p(x - \xi) = px - (p - 1)\,\xi.$$

Ersetzt man darin, für positive y, η, x und ξ durch $\frac{x}{x+y}$, $\frac{\xi}{\xi+\eta}$, so kommt

$$\frac{x^p}{\xi^{p-1}} \geqq \frac{(x+y)^p}{(\xi+\eta)^{p-1}}\left[p\,\frac{x}{x+y} - (p-1)\,\frac{\xi}{\xi+\eta}\right].$$

7*

Vertauscht man x, ξ mit y, η und addiert, so wird

$$\frac{x^p}{\xi^{p-1}} + \frac{y^p}{\eta^{p-1}} \geqq \frac{(x+y)^p}{(\xi+\eta)^{p-1}}.$$

Hierin setze man x_k, y_k $(k = 1, 2, \ldots, n)$ für x, y, ferner

$$\xi^p = \Sigma x_k^p, \quad \eta^p = \Sigma y_k^p$$

und summiere nach k, dann folgt

$$(\xi + \eta)^p \geqq \Sigma (x_k + y_k)^p$$

oder

$$(\Sigma x_k^p)^{\frac{1}{p}} + (\Sigma y_k^p)^{\frac{1}{p}} \geqq (\Sigma (x_k + y_k)^p)^{\frac{1}{p}}.$$

Dies gilt für positive, natürlich auch für nichtnegative x_k, y_k; für beliebige erhält man, wegen $|x_k| + |y_k| \geqq |x_k + y_k|$,

$$(\Sigma |x_k|^p)^{\frac{1}{p}} + (\Sigma |y_k|^p)^{\frac{1}{p}} \geqq (\Sigma |x_k + y_k|^p)^{\frac{1}{p}},$$

also die zu beweisende Ungleichung.

[48] Der mit dem Betrag (10) definierte Raum mag etwa der *pseudo-Euklidische* n-dimensionale Raum E_n^p genannt werden.

Die bisher > 1 angenommene Zahl p kann übrigens auch $= 1$ gesetzt werden, also die Betragsdefinition

$$(11) \qquad |x| = |x_1| + |x_2| + \cdots + |x_n|.$$

Andererseits kann man sagen, daß dem Grenzfall $p = \infty$ die Betragsdefinition

$$(12) \qquad |x| = \max[|x_1|, \ |x_2|, \ldots, |x_n|]$$

entspricht; denn wird dieser größte unter den Beträgen $|x_k|$ mit ξ bezeichnet, so folgt aus (10) $\xi \leqq |x| \leqq n^{\frac{1}{p}} \cdot \xi$, also für $p \to \infty$ $|x| \to \xi$. Für diese beiden letzten Betragsdefinitionen ist übrigens die Summenungleichung trivial.

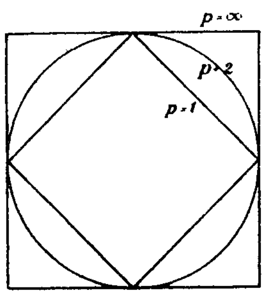

Fig. 1.

Deutet man x_1, x_2 als rechtwinklige Koordinaten in der Euklidischen Ebene, wodurch man den E_2^p eineindeutig (aber nur für $p = 2$ isometrisch) auf diese abbildet, so ist es nicht ohne Interesse, sich die „Eichkurve" $|x_1|^p + |x_2|^p = 1$, das Bild des „Einheitskreises" $|x| = 1$, vorzustellen. Für $p = 2$ ist es der Euklidische Einheitskreis, für $p = 1$ das ihm eingeschriebene Quadrat mit den Ecken $(\pm 1, 0)$ und $(0, \pm 1)$, für $p = \infty$ das ihm umschriebene Quadrat mit den Ecken $(\pm 1, \pm 1)$. Für die übrigen p verläuft die Kurve in den entsprechenden Zwischengebieten. — Für $n = 3$ liefert $p = 2$ die Euklidische Einheitskugel, $p = 1$ ein ihr eingeschriebenes Oktaeder, $p = \infty$ einen ihr umschriebenen Würfel.

Die dem Grenzübergang $n \to \infty$ entsprechenden Räume ergeben sich wie im Falle $p = 2$; so der pseudo-Hilbertsche Raum H^p als Raum der Zahlenfolgen, für die $|x_1|^p + |x_2|^p + \cdots$ konvergiert, mit der Betragsdefinition

$$(13) \qquad |x| = (|x_1|^p + |x_2|^p + \cdots)^{\frac{1}{p}},$$

der Raum der stetigen Funktionen $x = x(t)$ mit

$$(14) \qquad |x| = \left(\int_a^b |x(t)|^p \, dt\right)^{\frac{1}{p}}$$

usw. Für $p = \infty$ ist die sachgemäße Übertragung von (12): Raum der *beschränkten* Zahlenfolgen mit

$$(15) \qquad |x| = \sup |x_n|,$$

was sich ohne weiteres auf Komplexe mit beliebigem Träger ausdehnen läßt.

Bisher haben wir als Betrag von x stets eine symmetrische Funktion der Koordinaten gewählt; wenn man darauf verzichtet, läßt sich der Kreis der Beispiele sehr erweitern. So kann man den beschränkten Zahlenfolgen $x = (x_1, x_2, \ldots)$ die Beträge

$$|x| = c_1 |x_1| + c_2 |x_2| + \cdots$$

zuordnen, wo $c_1 + c_2 + \cdots$ eine feste konvergente Reihe positiver Zahlen ist. Oder: wir definieren den Betrag einer (reellen) Matrix

$$x = \begin{pmatrix} x_{11} & \ldots & x_{1n} \\ \cdot & \cdots & \cdot \\ x_{m1} & \ldots & x_{mn} \end{pmatrix}$$

folgendermaßen: es sei

$$v_i = \sum_k x_{ik} u_k \qquad (i = 1, \ldots, m; \; k = 1, \ldots, n)$$

und $|x|$ das Maximum von $(\Sigma v_i^2)^{\frac{1}{2}}$ unter der Bedingung $\Sigma u_k^2 = 1$. Man sieht leicht, daß $|x|$ die Betragsaxiome erfüllt, also die Entfernung $|x - y|$ die Menge der Matrizen (aus mn Elementen) zu einem metrischen Raum macht. Man kann den Exponenten 2 wieder durch p ersetzen und den Übergang zu unendlichen Matrizen machen.

4. Bairesche Räume. Ist $x = (x_1, x_2, \ldots)$ eine Folge reeller Zahlen oder auch von Elementen beliebiger Mengen $(x_n \varepsilon A_n)$, und ist $m(x, y)$ die erste Differenzstelle für zwei Folgen $x \neq y$, d. h. die durch [49]

$$x_m \neq y_m, \quad x_n = y_n \text{ für } n < m$$

definierte natürliche Zahl, so erhält man eine zulässige Entfernungsdefinition durch

$$(16) \qquad xx = 0, \quad xy = 1 : m(x, y) \text{ für } x \neq y.$$

Hier gilt das Dreiecksaxiom sogar in der schärferen Gestalt

$$xz \leqq \max [xy, yz].$$

Denn unter der Voraussetzung, daß alle drei Folgen verschieden sind (andernfalls ist nichts zu beweisen), hat man

$$m(x, z) \geqq \min [m(x, y), \quad m(y, z)],$$

weil, die rechte Seite $= m$ gesetzt, für $n < m$ $x_n = y_n$ und $y_n = z_n$, also $x_n = z_n$ und daher $m(x, z) \geqq m$ ist. Diese Räume heißen *Bairesche Räume*; insbesondere entsteht der *Bairesche Nullraum*, wenn man die x_n alle natürlichen Zahlen durchlaufen läßt.

5. Der p-adische Raum. p sei eine natürliche Primzahl, ϑ eine Zahl des Intervalls $0 < \vartheta < 1$, beide fest gewählt. Man kann dann jeder rationalen Zahl x einen Betrag oder besser, da man hier eine Verwechslung vermeiden muß, eine „Bewertung" $\|x\|$ in folgender Weise zuordnen:

$$(17) \qquad \|0\| = 0, \quad \|x\| = \vartheta^{m(x)} \text{ für } x \neq 0,$$

wo $m(x)$ diejenige ganze Zahl m ($\gtreqless 0$) bedeutet, für die x genau durch p^m teilbar, d. h. $x = p^m x_0$ und x_0 Quotient von zwei nicht durch p teilbaren ganzen Zahlen ist. Das Summenaxiom gilt hier wieder in der schärferen Form

$$\|x + y\| \leqq \max [\|x\|, \|y\|].$$

Denn sind $x, y, x + y \neq 0$ (sonst ist nichts zu beweisen), so ist

$$m(x + y) \geqq \min [m(x), m(y)];$$

ist nämlich $x = p^m x_0$, $y = p^n y_0$ und etwa $m \leqq n$, so ist $x + y = p^m (x_0 + p^{n-m} y_0)$, wo der zweite Faktor mit einem durch p nicht teilbaren Nenner (dem Produkt der Nenner von x_0, y_0) darstellbar und $x + y$ also genau durch p^m oder eine höhere Potenz teilbar ist. Die Entfernung $\|x - y\|$ macht die Menge der rationalen Zahlen zu einem metrischen Raum; die Vervollständigung (§ 21) dieses Raumes führt zu den von K. Hensel geschaffenen p-adischen Zahlen, die für die Theorie der algebraischen Zahlkörper von hoher Bedeutung sind.

6. Produkt metrischer Räume. Man lasse x einen metrischen Raum A, y einen metrischen Raum B durchlaufen; das Punktpaar $z = (x, y)$ durchläuft dann (§ 4) die Menge $C = (A, B)$, die man auf unendlich viele Weisen wieder zu einem metrischen Raum machen kann, etwa durch

$$(18) \qquad z_1 z_2 = [(x_1 x_2)^p + (y_1 y_2)^p]^{\frac{1}{p}} \qquad (p \geqq 1)$$

oder

$$(19) \qquad z_1 z_2 = \max [x_1 x_2, y_1 y_2].$$

Nimmt man z. B. $p = 2$, so wird das Produkt (E_m, E_n) von zwei Euklidischen Räumen mit dem Euklidischen Raum E_{m+n} isometrisch. Das zu einem $z = (x, y)$ gehörige x heißt die *Projektion des Punktes z auf den Raum A*, und die Projektionen aller Punkte z einer Menge $Z \leqq C$ bilden die Projektion X von Z. Die Entfernung zweier Punkte ist mindestens gleich der ihrer Projektionen: $z_1 z_2 \geqq x_1 x_2$. — Man kann diese Produktbildung ohne weiteres auf jede endliche Zahl von Faktoren ausdehnen.

§ 21. Konvergenz.

1. Im metrischen Raume E sei $(x_1, x_2, \ldots)$ eine Punktfolge. Wenn es in E einen Punkt x gibt derart, daß die Entfernung $x x_n$ mit $n \to \infty$ nach 0 konvergiert:

$$\lim x x_n = 0 \quad \text{oder} \quad x x_n \to 0,$$

so heißt die Folge *in E konvergent* und x ihr *Limes* (Grenzpunkt), in Zeichen

$$\lim x_n = x \quad \text{oder} \quad x_n \to x.$$

Es kann natürlich nur einen solchen Punkt geben, denn aus $x x_n \to 0$, $y x_n \to 0$ folgt nach dem Dreiecksaxiom $xy = 0$, $x = y$. Allgemeiner: für zwei in E konvergente Folgen $x_n \to x$, $y_n \to y$ ist

$$x_n y_n \to xy. \tag{1}$$

Denn nach der Dreiecksungleichung ist

$$\begin{aligned} x_n y_n &\leqq x_n x + xy + y y_n, \\ xy &\leqq x x_n + x_n y_n + y_n y, \\ |x_n y_n - xy| &\leqq x x_n + y y_n, \end{aligned}$$

woraus für $n \to \infty$ (1) folgt. Mit einer dem Leser schon jetzt verständlichen Sprechweise drücken wir (1) so aus: xy ist eine *stetige Funktion* von x und y.

Die Punktfolge $(x_1, x_2, \ldots)$ heißt eine *Fundamentalfolge* oder Cauchysche Folge, wenn es in ihr für jedes (noch so kleine) positive ε einen Punkt x_m gibt, der von allen folgenden eine Entfernung $< \varepsilon$ hat:

$$x_m x_n < \varepsilon \quad \text{für} \quad n > m.$$

Jede konvergente Folge ist ersichtlich eine Fundamentalfolge. Das Umgekehrte kann nicht allgemein richtig sein, da ja die Definition der Fundamentalfolge nur aus der Folge selbst geschöpft ist, die Konvergenz in E aber vom Raume E abhängt. Wenn z. B. E die Menge der rationalen Zahlen ist, so ist eine Folge rationaler Zahlen, die nach einer irrationalen Zahl konvergiert, zwar eine Fundamentalfolge, aber in E nicht konvergent. *Wenn in E jede Fundamentalfolge konvergiert, heißt E ein vollständiger Raum* (eine vollständige Punktmenge). Die Menge der rationalen Zahlen ist also nicht vollständig, wohl aber die der reellen Zahlen: daß jede Fundamentalfolge reeller Zahlen konvergiert, ist ja genau die Aussage des sogenannten *allgemeinen Konvergenzkriteriums* von Cauchy, das dem Leser aus den Elementen bekannt ist.

2. Beispiele vollständiger und unvollständiger Räume. Jeder Euklidische Raum, aber auch der Hilbertsche ist vollständig; beweisen wir es für diesen. Der Betrag eines Punktes x

$$|x| = (x_1^2 + x_2^2 + \cdots)^{\frac{1}{2}}$$

ist mindestens gleich dem Betrag einer Koordinate, $|x| \geqq |x_k|$; also die Entfernung $|x - y| \geqq |x_k - y_k|$. Bilden also die Punkte [1])

$$x^n = (x_1^n, x_2^n, \ldots, x_k^n, \ldots)$$

eine Fundamentalfolge, so bilden ihre k-ten Koordinaten erst recht eine und haben einen Limes

$$x_k = \lim_n x_k^n;$$

setzen wir dann $\qquad x = (x_1, x_2, \ldots, x_k, \ldots).$

Bei vorgeschriebenem $\varepsilon > 0$ ist $|x^n - x^m| < \varepsilon$ für geeignetes m und jedes $n > m$, also

$$\sum_{k=1}^{\infty} (x_k^n - x_k^m)^2 < \varepsilon^2,$$

um so mehr für jede natürliche Zahl l

$$\sum_{k=1}^{l} (x_k^n - x_k^m)^2 < \varepsilon^2,$$

daraus für $n \to \infty$

$$\sum_{k=1}^{l} (x_k - x_k^m)^2 \leqq \varepsilon^2;$$

hieraus folgt wieder für $l \to \infty$

$$\sum_{k=1}^{\infty} (x_k - x_k^m)^2 \leqq \varepsilon^2;$$

aus der Konvergenz der links stehenden Reihe ergibt sich, daß $x - x^m$ und also x dem Hilbertschen Raum angehört. Weiter ist $|x - x^m| \leqq \varepsilon$, mit $|x^n - x^m| < \varepsilon$ vereinigt also $|x - x^n| < 2\varepsilon$ für $n > m$, und dies besagt $|x - x^n| \to 0$ oder $x^n \to x$.

Jener zwischen den Euklidischen Räumen und dem Hilbertschen
[50] stehende Raum $E = \mathfrak{S} E_n$ (S. 98) ist unvollständig; denn ist $x = (x_1, x_2, \ldots)$ ein Punkt des Hilbertschen Raumes, so konvergieren die zu E gehörigen Punkte

$$x^n = (x_1, x_2, \ldots, x_n, 0, 0, \ldots)$$

nach x und bilden eine Fundamentalfolge, die aber, falls unendlich viele Koordinaten $x_n \neq 0$, in E nicht konvergiert.

Der Raum der in $[a, b]$ stetigen Funktionen $x = x(t)$ mit der Betragsdefinition

$$|x| = \left(\int_a^b x(t)^2 \, dt\right)^{\frac{1}{2}}$$

ist nicht vollständig. Zu jeder (beschränkten, im Riemannschen Sinne) integrablen Funktion ξ läßt sich eine Folge stetiger Funktionen x_n mit $|x_n - \xi| \to 0$ angeben; wenn zugleich eine stetige Funktion x mit

[1]) n und m sind obere Indizes, nicht Exponenten.

$|x_n - x| \to 0$, also $|x - \xi| \doteq 0$ vorhanden sein soll, so muß die Nullfunktion $x - \xi$ an ihren Stetigkeitsstellen verschwinden, x an den Stetigkeitsstellen von ξ mit ξ übereinstimmen. Die einfachsten Beispiele lehren bereits, daß eine solche stetige Funktion x nicht zu existieren braucht: man lasse etwa $\xi(t)$ an einer einzigen Stelle c zwischen a und b einen Sprung mit $\xi(c-0) \neq \xi(c+0)$ machen und sonst stetig sein. Auch der Raum der im Riemannschen Sinne integrablen Funktionen ist aber noch nicht vollständig (erst der Raum der Funktionen, für die $x(t)$ und sein Quadrat im Lebesgueschen [1]) Sinne integrabel ist).

Auch die pseudo-Euklidischen und der pseudo-Hilbertsche Raum sind vollständig; insbesondere dem Grenzfall $p = \infty$ entsprechend der Raum der beschränkten Zahlenfolgen mit der Betragsdefinition

$$|x| = \sup |x_n|.$$

Hier ist auch der Raum der stetigen Funktionen $x(t)$ mit der Betragsdefinition

$$|x| = \max_{a \leqq t \leqq b} |x(t)|$$

vollständig; denn eine Fundamentalfolge stetiger Funktionen konvergiert alsdann *gleichmäßig* gegen eine, mithin ebenfalls stetige Grenzfunktion $x(t)$.

Die reellen Zahlenfolgen $x = (x_1, x_2, \ldots)$ bilden auf Grund der Betragsdefinition

$$|x| = \inf_k \left(x_1^2 + x_2^2 + \cdots + x_k^2 + \frac{1}{k^2}\right)^{\frac{1}{2}},$$

deren Zulässigkeit der Leser beweisen möge, einen vollständigen Raum, worin die Konvergenz $x^n \to x$ mit Konvergenz sämtlicher Koordinaten $(x_1^n \to x_1, x_2^n \to x_2, \ldots)$ gleichbedeutend ist.

Der Bairesche Raum (S. 101) der Elementfolgen

$$x = (x_1, x_2, \ldots)$$

ist vollständig, wenn die x_k unabhängig voneinander alle Elemente gegebener Mengen A_k durchlaufen. Denn bilden die Punkte

$$x^n = (x_1^n, x_2^n, \ldots)$$

eine Fundamentalfolge, so ist, für jede natürliche Zahl k, $x^n x^m < \frac{1}{k}$ für geeignetes m und jedes $n > m$, was aber nach der Definition der Entfernung zur Folge hat, daß $x_k^m = x_k^{m+1} = x_k^{m+2} = \cdots$. Nennen wir dies Element x_k und bilden die Folge

$$x = (x_1, x_2, \ldots),$$

so stimmt x^n in einer, mit n gleichzeitig nach ∞ strebenden Zahl k von Anfangselementen mit x überein, d. h. $x x^n \to 0$.

[1]) Die Lebesguesche Maß- und Integraltheorie wird in diesem Buche nicht behandelt und die Bekanntschaft mit ihr nicht vorausgesetzt. Der Leser, der die obige Anspielung auf den Riesz-Fischerschen Satz nicht versteht, betrachte sie als nicht vorhanden.

3. Vervollständigung eines Raumes. Wie wir in § 11 die Dedekindsche Theorie der Irrationalzahlen als Vorbild zur Ausfüllung der Lücken geordneter Mengen genommen haben, so können wir nach dem Muster der Cantor-Mérayschen Theorie, welche die Irrationalzahlen durch Fundamentalfolgen rationaler Zahlen definiert, jeden metrischen Raum E zu einem vollständigen Raum E erweitern, dessen Elemente die Fundamentalfolgen

$$\xi = (x_1, x_2, \ldots, x_n, \ldots)$$

aus Punkten von E sind. Bemerken wir zunächst:

(α) Für zwei Fundamentalfolgen ξ, η existiert immer $\lim x_n y_n$.

In der Tat ist

$$x_m y_m \leqq x_m x_n + x_n y_n + y_n y_m,$$

$$|x_m y_m - x_n y_n| \leqq x_m x_n + y_m y_n,$$

woraus unmittelbar folgt, daß die reellen Zahlen $x_n y_n$ eine Fundamentalfolge bilden und konvergieren.

(β) Von zwei Folgen ξ, η mit $x_n y_n \to 0$ ist mit der einen zugleich auch die andere eine Fundamentalfolge.

Dies folgt aus

$$|y_m y_n - x_m x_n| \leqq x_m y_m + x_n y_n.$$

Gemäß (α) definieren wir als Entfernung von ξ, η

$$\xi\eta = \lim x_n y_n.$$

Aus $x_n y_n + y_n z_n \geqq x_n z_n$ folgt die Dreiecksungleichung $\xi\eta + \eta\zeta \geqq \xi\zeta$. Um das Entfernungsaxiom (β) (S. 94) aufrechtzuerhalten, definieren wir zwei Fundamentalfolgen ξ, η dann und nur dann als gleich, wenn $\xi\eta = 0$, abweichend von der sonstigen Gleichheitserklärung für Komplexe (vgl. die Bemerkung über unstetige Funktionen S. 99).

Zu den Fundamentalfolgen gehören gewiß die konstanten Folgen

$$(x, x, x, \ldots),$$

die Entfernung von zwei solchen ist xy, so daß E isometrisch auf eine Teilmenge von E bezogen ist. Wir können diese konstanten Folgen ohne Gefahr mit den Punkten x selbst identifizieren, so daß E geradezu Teilmenge von E wird. Die Entfernung zwischen $x = (x, x, \ldots)$ und $\xi = (x_1, x_2, \ldots)$ ist $x\xi = \lim x\, x_n$. Insbesondere ist die Entfernung zwischen ξ und einem seiner eigenen Punkte x_m

$$x_m \xi = \lim_n x_m x_n.$$

Diese konvergiert für $m \to \infty$ nach 0; denn wählt man m so, daß für $n > m$ stets $x_m x_n < \varepsilon$, so ist $x_m \xi \leqq \varepsilon$, $x_n \xi < 2\varepsilon$ für $n > m$. Daher kann man zu jedem ξ ein x mit beliebig kleinem $x\xi$ bestimmen (E ist in E dicht, § 25).

Nennen wir für den Augenblick, der Deutlichkeit wegen, die Folgen von Elementen ξ Sequenzen. Ist $(\xi_1, \xi_2, \ldots)$ eine Fundamentalsequenz, so bestimmen wir zu jedem ξ_n ein x_n mit $x_n \xi_n < \frac{1}{n}$; nach der Bemerkung

(β) bilden die x_n dann eine Fundamentalsequenz in E oder eine Fundamentalfolge $\xi = (x_1, x_2, \ldots)$ in E. Demnach ist $\xi x_n \to 0$, außerdem $x_n \xi_n \to 0$, also $\xi \xi_n \to 0$ nach dem Dreiecksaxiom, d. h. ξ_n konvergiert nach ξ und der Raum E ist vollständig.

Nennen wir die Menge E in ihrer Abhängigkeit von E jetzt $\overline{E}$ (Menge der Fundamentalfolgen aus E). Für einen vollständigen Raum V ist $V = \overline{V}$. Ist E in dem vollständigen Raum V enthalten, so ist $\overline{E} \leqq \overline{V} = V$, also auch $\overline{E}$ in V enthalten: $\overline{E}$ ist *der kleinste vollständige Raum, in dem E enthalten ist*, und soll *die vollständige Hülle*[1]) von E heißen. Allerdings haben wir dabei die Punkte x von E mit den konstanten Folgen $(x, x, \ldots)$ oder, wegen der Gleichheitsdefinition für Fundamentalfolgen, mit den nach x konvergenten Folgen $(x_1, x_2, \ldots)$ identifiziert. Will man diese Verwechslung isometrischer Räume vermeiden, so kann man sagen: ist V ein vollständiger Raum $\geqq E$, so konvergieren alle Fundamentalfolgen aus E nach Punkten x von V, und die Menge $\overline{E}$ dieser Punkte x heißt *eine vollständige Hülle* von E; alle vollständigen Hüllen von E sind isometrisch und zwar derart, daß bei der isometrischen Beziehung die Punkte von E sich selbst entsprechen.

4. Kompakte Mengen. Es sei x_p eine Teilfolge der Folge x_n, ausführ- [51]
licher: $(x_{p_1}, x_{p_2}, \ldots)$ eine Teilfolge von $(x_1, x_2, \ldots)$, wo $p_1 < p_2 < \cdots$ eine Folge wachsender natürlicher Zahlen ist. Wenn x_n eine konvergente Teilfolge $x_p \to x$ hat, so heißt x ein *Häufungspunkt* der ganzen Folge x_n. Dafür ist notwendig und hinreichend, daß für jedes $\varepsilon > 0$ *unendlich oft* $x x_n < \varepsilon$ sei (während für $x_n \to x$ die Bedingung lautet: *schließlich* $x x_n < \varepsilon$). Notwendig: es ist schließlich $x x_p < \varepsilon$, also unendlich oft $x x_n < \varepsilon$. Hinreichend: es gibt ein p_1 mit $x x_{p_1} < 1$, ein $p_2 > p_1$ mit $x x_{p_2} < \frac{1}{2}$, ein $p_3 > p_2$ mit $x x_{p_3} < \frac{1}{3}$ usw., also $x_p \to x$.

Wir definieren nun:

Eine Teilmenge A des Raumes E heißt in E kompakt, wenn jede Punktfolge $x_n \varepsilon A$ eine in E konvergente Teilfolge, d. h. einen Häufungspunkt in E hat.

Eine Menge A heißt bedingt kompakt, wenn jede Punktfolge $x_n \varepsilon A$ eine Fundamentalfolge als Teilfolge hat.

Eine in sich kompakte Menge E ist vollständig. E sei in sich kompakt, d. h. in E; jede Fundamentalfolge aus E hat in E einen Häufungspunkt, der aber in diesem Falle der Limes der ganzen Folge ist.

Eine im Raume E kompakte Menge ist bedingt kompakt. Umgekehrt, eine bedingt kompakte Menge A ist in einem geeigneten Raume $E \geqq A$ kompakt, z. B. in einem vollständigen Raum E (etwa in der vollständigen

[1]) Vgl. die Anmerkung S. 96.

Hülle $\bar{A}$); denn dort hat jede Fundamentalfolge einen Limes, also jede Punktfolge $x_n \varepsilon A$ einen Häufungspunkt. Genauer gilt:

I. *Dann und nur dann, wenn E vollständig ist, ist jede bedingt kompakte Teilmenge von E auch in E kompakt.*

Denn wenn E nicht vollständig und x_n eine Fundamentalfolge ohne Limes ist, wobei wir offenbar die x_n paarweise verschieden annehmen können, so ist die Menge $A = \{x_1, x_2, \ldots\}$ zwar bedingt kompakt, aber nicht in E kompakt.

Die bedingt kompakten Mengen sind also die, die in geeigneten Räumen kompakt sind.

Um die bedingt kompakten Mengen zu charakterisieren, stellen wir noch folgende Definitionen auf.

Eine Menge A heißt beschränkt, wenn die Entfernungen ihrer Punktpaare eine endliche obere Grenze haben; diese wird der Durchmesser $d(A)$ von A genannt.

Eine Menge A heißt total beschränkt, wenn sie sich für jedes $\delta > 0$ als Summe endlich vieler Mengen von Durchmessern $\leqq \delta$ darstellen läßt.

Jede total beschränkte Menge ist erst recht beschränkt. Im Euklidischen Raume E_p sind beschränkte Mengen auch total beschränkt, denn ein Würfel läßt sich, indem man seine Kanten in n gleiche Teile teilt, in n^p Teilwürfel teilen, deren Durchmesser mit wachsendem n beliebig klein werden. Im Hilbertschen Raum ist z. B. die Menge der Punkte

$$\begin{gathered}(1, 0, 0, 0, \ldots)\\(0, 1, 0, 0, \ldots)\\(0, 0, 1, 0, \ldots)\\\cdots\end{gathered}$$

zwar beschränkt, aber nicht total beschränkt; denn die Entfernung zweier Punkte ist $\sqrt{2}$ und für $\delta < \sqrt{2}$ läßt sich die Menge nicht als Summe endlich vieler Mengen von Durchmessern $\leqq \delta$ darstellen.

II. *Bedingt kompakte und total beschränkte Mengen sind identisch.*

Sei A bedingt kompakt und ϱ eine beliebig kleine positive Zahl. Wir wählen einen Punkt a_1 von A, dann einen zweiten a_2 mit $a_1 a_2 \geqq \varrho$, einen dritten a_3 mit $a_1 a_3$, $a_2 a_3 \geqq \varrho$ usw., kurz, wir suchen, *so lange es geht*, Punkte, die paarweise Entfernungen $\geqq \varrho$ haben. Solcher Punkte kann es nur endlich viele geben, etwa $a_1, a_2, \ldots, a_n$, da wir andernfalls eine Folge erhielten, von der gewiß keine Teilfolge eine Fundamentalfolge ist. Nunmehr hat jeder Punkt $x \varepsilon A$ von mindestens einem dieser Punkte a_k eine Entfernung $< \varrho$, da man sonst den Punkt x noch den Punkten a_k hätte hinzufügen können. Ist also A_k die Menge der Punkte $x \varepsilon A$ mit $x a_k < \varrho$, so ist $A = A_1 \dot{+} A_2 \dot{+} \cdots \dot{+} A_n$ und jede Menge A_k hat einen Durchmesser $\leqq 2\varrho$. A ist also total beschränkt.

Umgekehrt sei A total beschränkt. Stellt man es als Summe endlich vieler Mengen A_k mit Durchmessern $\leqq \delta$ dar und ist x_n eine Punktfolge aus A, so muß mindestens ein Summand A_k unendlich viele x_n enthalten; wir können sagen: *jede Punktfolge x_n hat eine Teilfolge x_p von beliebig kleinem Durchmesser* $\leqq \delta$ (d. h. in der zwei Punkte eine Entfernung $\leqq \delta$ haben). Danach bilden wir von den x_n eine Teilfolge

$$x_{p_1} x_{p_2} x_{p_3} \ldots$$

vom Durchmesser $\leqq 1$, von diesen x_p eine Teilfolge

$$x_{q_1} x_{q_2} x_{q_3} \ldots$$

vom Durchmesser $\leqq \frac{1}{2}$, von diesen x_q eine Teilfolge

$$x_{r_1} x_{r_2} x_{r_3} \ldots$$

vom Durchmesser $\leqq \frac{1}{3}$ usw. Da nun $p_1 \leqq q_1 < q_2 \leqq r_2 < r_3 \leqq \cdots$, so können wir die *Diagonalfolge*

$$x_{p_1} x_{q_2} x_{r_3} \ldots$$

bilden, die vom k-ten Gliede an Teilfolge der k-ten unter den vorangehenden Folgen ist und in der also das k-te Glied von allen folgenden eine Entfernung $\leqq \frac{1}{k}$ hat; sie ist also eine Fundamentalfolge, die in der willkürlich gewählten Teilfolge x_n enthalten ist, und A ist bedingt kompakt. Damit ist II bewiesen.

Im Euklidischen Raum E (der ja vollständig ist) sind beschränkte, total beschränkte, bedingt kompakte und in E kompakte Mengen identisch. Der Leser kennt dies als das Bolzano-Weierstraßsche Theorem: *jede beschränkte Punktfolge des Euklidischen Raumes hat einen Häufungspunkt.* Im allgemeinen ist aber nicht Beschränktheit, sondern totale Beschränktheit das echte Äquivalent der bedingten Kompaktheit; zu dieser, die eine Eigenschaft der Menge A selbst ist, muß noch eine geeignete Beschaffenheit des Raumes E, als welche Vollständigkeit jedenfalls hinreicht, hinzukommen, um A zu einer in E kompakten Menge zu machen.

§ 22. Innere Punkte und Randpunkte.

Wir betrachten jetzt Punkte und Teilmengen eines festen metrischen Raumes E. Ist x ein Punkt, ϱ eine positive Zahl, so heißt die Menge der Punkte y, die von x eine Entfernung $< \varrho$ haben, *eine Umgebung* von x, ϱ ihr *Radius*. Sie wird mit U_x oder genauer $U_x(\varrho)$ bezeichnet.

Daß ein Punkt x *Limes* (*Häufungspunkt*) der Folge x_n ist, drückt sich mit dem Umgebungsbegriff so aus: *jede Umgebung U_x enthält fast alle* (*unendlich viele*) *Punkte der Folge x_n.* In der Tat ist doch für jedes $\varrho > 0$ *schließlich* (*unendlich oft*) $x x_n < \varrho$ oder $x_n \varepsilon U_x(\varrho)$.

A sei eine Punktmenge. Wenn ein Punkt x von A so beschaffen ist, daß es eine Umgebung $U_x \leqq A$ gibt, so heißt er ein *innerer Punkt*, andernfalls ein *Randpunkt* von A. Die Menge der inneren Punkte sei A_i, die der Randpunkte A_r, womit die Spaltung in disjunkte Summanden

$$A = A_i + A_r$$

gegeben ist. A_i heiße der *offene Kern* (aus einem nachher ersichtlichen Grunde), A_r der *Rand* von A. *Eine Menge, die aus lauter inneren Punkten*
[52] *besteht* ($A_r = 0$), *heiße eine offene Menge*; eine, die aus lauter Randpunkten besteht ($A_i = 0$), eine *Randmenge* [1]).

Wir geben einige Beispiele aus der Euklidischen Ebene E mit rechtwinkligen Koordinaten x_1, x_2.

A sei eine Kreisfläche mit Peripherie, z. B. $x_1^2 + x_2^2 \leqq 1$. Die Peripheriepunkte ($x_1^2 + x_2^2 = 1$) sind Randpunkte, die übrigen innere Punkte.

A sei eine Kreisfläche ohne Peripherie, $x_1^2 + x_2^2 < 1$; sie ist offen.

A sei eine Kreislinie, $x_1^2 + x_2^2 = 1$; das ist eine Randmenge.

A sei die Menge der „rationalen Punkte", d. h. derer mit zwei rationalen Koordinaten, $E - A$ die Menge der „irrationalen Punkte", d. h. mit mindestens einer irrationalen Koordinate; beide sind Randmengen.

$f(x_1, x_2)$ sei eine stetige Funktion. Die Menge $[f > 0]$, soll heißen die Menge der Punkte x wo $f(x_1, x_2) > 0$, ist offen. Z. B. das Innere einer Ellipse: $1 - \frac{x_1^2}{a_1^2} - \frac{x_2^2}{a_2^2} > 0$.

Das gilt für jeden Raum. Ist jedem Punkt x von E eindeutig ein Punkt $y = f(x)$ eines andern oder desselben metrischen Raumes zugeordnet, so heißt diese Funktion *im Punkte x stetig*, wenn eine der beiden gleichwertigen Bedingungen erfüllt ist:

(1) Wenn x_n nach x konvergiert, konvergiert $y_n = f(x_n)$ nach $y = f(x)$.

(2) Ist $y = f(x)$, $\eta = f(\xi)$, so wird, wenn $x\xi$ *hinlänglich* klein ist, $y\eta$ *beliebig* klein, d. h. zu jedem $\sigma > 0$ läßt sich ein $\varrho > 0$ so bestimmen, daß mit $x\xi < \varrho$ zugleich $y\eta < \sigma$ ist. Dabei ist x festzuhalten.

Ist $f(x)$ in jedem Punkte x stetig, so heißt sie *stetig* schlechthin. Wir kommen auf die stetigen Funktionen ausführlich zurück. Einstweilen aber können wir sagen:

I. *Ist $f(x)$ eine reelle stetige Funktion, so ist die Menge $[f > 0]$ offen.*

Denn ist $y = f(x) > 0$, so läßt sich zu gegebenem $\sigma > 0$ ein passendes $\varrho > 0$ finden derart, daß für $x\xi < \varrho$ zugleich $|y - \eta| < \sigma$, also (wenn $\sigma < y$ gewählt wird) $\eta > y - \sigma > 0$ ist. D. h. es gibt eine Umgebung $U_x(\varrho)$, für deren Punkte ξ auch noch $f(\xi) > 0$; x ist innerer Punkt der Menge $[f > 0]$.

Wir wollen uns überzeugen, daß auf diese Weise alle offenen Mengen erhalten werden können. Bezeichnen wir für einen Punkt x und eine Menge $B > 0$ als *untere Entfernung*

[1]) In beiden Fällen ist die Nullmenge mitzuzählen.

$$\delta(x, B) = \inf_{y \varepsilon B} xy$$

die untere Grenze der Entfernungen des Punktes x von den Punkten y der Menge B. Aus $xy \leqq \xi y + x\xi$ folgt, indem man beiderseits die untere Grenze für $y \varepsilon B$ nimmt,

$$\delta(x, B) \leqq \delta(\xi, B) + x\xi$$

und durch Vertauschung von x, ξ

$$|\,\delta(x, B) - \delta(\xi, B)\,| \leqq x\xi.$$

dies besagt: $\delta(x, B)$ *ist stetige Funktion von* x. Offenbar ist nun, wenn $A = E - B$ das Komplement von B bedeutet, die Ungleichung $\delta(x, B) > 0$ damit gleichbedeutend, daß x eine von Punkten von B freie, also zu A gehörige Umgebung (mit dem Radius $\varrho = \delta(x, B)$ nämlich) hat, d. h. innerer Punkt von A ist. Also für die stetige Funktion $f(x) = \delta(x, B)$ ist die Menge $[f > 0]$ mit A_i, bei offenem $A = A_i$ mit A identisch. Im Ausnahmefall $B = 0$, $A = E$ können wir etwa die konstante Funktion $f(x) = 1$ verwenden. Also gilt:

II. *Zu jeder offenen Menge* A *existiert eine in* E *stetige Funktion* $f(x)$ *derart, daß die Menge* $[f > 0]$ *mit* A *identisch ist.*

Aus I folgt, daß die *Umgebungen offene Mengen sind*; denn $\varrho - ax$ ist stetige Funktion von x und positiv in der Menge $U_a(\varrho)$.

Mit $A \leqq B$ ist natürlich $A_i \leqq B_i$; A_i ist eine „*monotone*" Funktion von A.

Die Menge A_i *ist stets offen*, da sie oben als Menge $[f > 0]$ mit stetigem $f(x)$ dargestellt wurde. Ist B offene Teilmenge von A, so ist $B = B_i \leqq A_i$, also A_i die *größte offene Teilmenge* von A, daher der Name *offener Kern* von A im Einklang mit der S. 96 getroffenen Verabredung.

Die Menge A_r ist stets eine Randmenge.

Sind $A, B, \ldots$ beliebig (auch unendlich) viele Mengen,

$$S = A \dot{+} B \dot{+} \cdots, \quad D = AB \cdots$$

deren Summe und Durchschnitt, so folgt aus der Monotonie jedenfalls

$$S_i \geqq A_i \dot{+} B_i \dot{+} \cdots, \quad D_i \leqq A_i B_i \cdots.$$

Dagegen gilt für endlich viele, z. B. zwei Mengen, schärfer:

$$\text{mit } D = AB \text{ ist } D_i = A_i B_i.$$

Denn $x \varepsilon A_i B_i$ hat eine Umgebung $U \leqq A$ und eine Umgebung $V \leqq B$; die kleinere der beiden Umgebungen gehört dann zu D, $x \varepsilon D_i$. Also $A_i B_i \leqq D_i$; andererseits war $D_i \leqq A_i B_i$.

Hieraus folgt ohne weiteres:

III. *Die Summe beliebig vieler und der Durchschnitt endlich vieler offener Mengen ist wieder eine offene Menge.*

Denn mit den obigen Bezeichnungen ist $S_i \geqq S$, also $S_i = S$, und $D_i = D$. — Der Durchschnitt unendlich vieler offener Mengen braucht nicht offen zu sein. In der Euklidischen Ebene ist der Durchschnitt konzentrischer offener Kreisflächen mit den Radien $\varrho + \frac{1}{n}$ $(n = 1, 2, \ldots)$ die Kreisfläche vom Radius ϱ mit Peripherie; jeder Punkt ist der Durchschnitt seiner Umgebungen mit den Radien $\frac{1}{n}$.

Eine Summe beliebig vieler Umgebungen ist offen, und jede offene Menge > 0 kann so dargestellt werden, etwa als Summe aller in ihr enthaltenen Umgebungen.

Sei B das Komplement von A, $E = A + B$. Die inneren Punkte von B heißen auch *äußere Punkte* von[1] A, und umgekehrt. Die Ränder beider Mengen vereinigt ergeben die *Begrenzung* (Grenze) von A und von B:

$$A_g = A_r + B_r = B_g.$$

Ist z. B. A eine Kreisfläche in der Ebene, gleichviel ob mit oder ohne Peripherie oder mit einem Teil der Peripherie, so ist die Begrenzung in jedem Fall die Kreisperipherie. Ist A die Menge der rationalen Punkte, so ist A_g gleich der ganzen Ebene.

Die Begrenzung einer offenen Menge reduziert sich auf den Rand des Komplements ($A_g = B_r$), ist also eine Randmenge. Die Begrenzung des ganzen Raumes oder der Nullmenge ist die Nullmenge[2]).

[53] § 23. Die α-, β-, γ-Punkte.

A sei eine Punktmenge im Raum E, x Punkt von E (nicht notwendig von A): Wir definieren: x heißt ein

α-Punkt , β-Punkt , γ-Punkt

von A, wenn jede Umgebung U_x

mindestens einen Punkt, unendlich viele, unabzählbar viele Punkte

von A enthält. Die Mengen der α-, β-, γ-Punkte seien A_α, A_β, A_γ.

Die β-Punkte heißen auch *Häufungspunkte*, die γ-Punkte *Verdichtungspunkte*, die Menge A_β die *Ableitung* von A.

Es wird also verlangt, daß der Durchschnitt $A\,U_x$, für jedes U_x, mindestens die Mächtigkeit $1, \aleph_0, \aleph_1$ habe; natürlich könnte man diese Skala weiterführen. Um im folgenden gelegentlich abzukürzen, setzen wir

$$k_\alpha = 1, \quad k_\beta = \aleph_0, \quad k_\gamma = \aleph_1$$

---

[1]) Die Präposition „von" wird also hier und in andern Fällen auch für Punkte verwendet, die nicht zu A gehören. Das ist vielleicht sprachlich anfechtbar, aber nun einmal üblich.

[2]) Der Sprachgebrauch bezüglich Rand und Begrenzung ist leider sehr schwankend.

und sagen: x ist ein λ-Punkt von A, wenn jedes $A\,U_x$ mindestens die Mächtigkeit k_λ hat; die Menge der λ-Punkte sei A_λ ($\lambda = \alpha, \beta, \gamma$).

Ein Häufungspunkt x von A kann auch so erklärt werden: jede Umgebung U_x enthält mindestens einen von x verschiedenen Punkt von A.

Beispiele aus der Euklidischen Ebene E:

A Kreisfläche ohne Peripherie: $A_\alpha = A_\beta = A_\gamma =$ Kreisfläche mit Peripherie.

A Menge der rationalen Punkte: $A_\alpha = A_\beta = E$, $A_\gamma = 0$.

A Menge der irrationalen Punkte: $A_\alpha = A_\beta = A_\gamma = E$.

A Menge der ganzzahligen Punkte: $A_\alpha = A$, $A_\beta = A_\gamma = 0$.

A sei (Fig. 2) die Menge der Punkte $(x_1, x_2) = \left(\frac{m}{n}, \frac{1}{n}\right)$, wo n die Reihe der Zahlen 1, 2, 4, 8, ... und m alle ganzen Zahlen $\gtreqless 0$ durchläuft. A_β ist die Gerade $x_2 = 0$, $A_\alpha = A + A_\beta$, $A_\gamma = 0$.

Mit dem Limesbegriff, statt mit dem Umgebungsbegriff, lassen sich die λ-Punkte folgendermaßen erklären. Die α-Punkte sind offenbar genau die Limespunkte $x = \lim a_n$ der (in E) konvergenten Folgen von Punkten aus A ($a_n \varepsilon A$); die β-Punkte diejenigen unter ihnen, die sich mit $x \neq a_n$ darstellen lassen. Oder auch: die β-Punkte sind α-Punkte jeder Menge, die aus A durch Tilgung endlich vieler Punkte entsteht, und vice versa; die γ-Punkte sind α-Punkte jeder Menge, die aus A durch Tilgung von endlich oder abzählbar vielen Punkten besteht, und vice versa. Häufungspunkte von Folgen aus A (S. 107) sind nicht notwendig Häufungspunkte von A, aber jedenfalls α-Punkte.

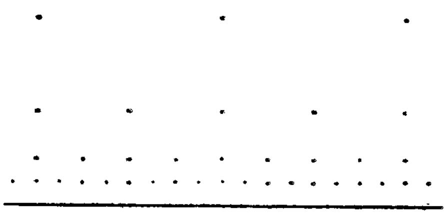

Fig. 2.

Wir wollen im Anschluß an die Mengen A_λ noch einige Bezeichnungen einführen, deren Fülle den Leser vielleicht am wenigsten verwirren wird, wenn wir radikal verfahren und auch die erst später mehr hervortretenden Glieder des Systems schon jetzt bringen. Es handelt sich um die Durchschnitte der A_λ mit A und die Komplemente dieser Durchschnitte in A, wobei wir A_α wegen $A\,A_\alpha = A$ außer Spiel lassen können. Wir setzen

$$(1) \quad \begin{cases} A_h = A\,A_\beta, & A_v = A\,A_\gamma \\ A = A_h + A_j = A_v + A_u. \end{cases}$$

A_h, A_v sind also die Mengen der zu A selbst gehörigen Häufungs- und Verdichtungspunkte; A_j, A_u ihre Komplemente. Ein Punkt a von A_j hat, da er nicht Häufungspunkt ist, eine Umgebung, die nur endlich viele Punkte von A enthält, also auch eine, die gar keinen Punkt von A außer a

selbst enthält; er heißt daher ein *isolierter* Punkt von A und die Menge A_j der isolierten Punkte der *isolierte Teil* von A. Cantor bezeichnet auch A_j
[54] als *Adhärenz*, A_h als *Kohärenz* von A. Ein Punkt von A_u hat, da er nicht Verdichtungspunkt ist, eine Umgebung, die höchstens abzählbar viele Punkte von A enthält; er heißt ein *unverdichteter* Punkt von A und A_u der *unverdichtete* Teil von A. Auf die Mengen A_u, A_v kommen wir erst später (§ 25) zurück.

Außer den trivialen Relationen

$$A_\alpha \geqq A, \quad A_\alpha \geqq A_\beta \geqq A_\gamma, \quad A \geqq A_h \geqq A_v$$

gilt, ebenfalls ganz leicht einzusehen,

$$(2) \qquad A_\alpha = A \dotplus A_\beta,$$

woraus man mit disjunkten Summanden

$$(3) \qquad A_\alpha = (A - A_h) + A_\beta = A + (A_\beta - A_h)$$

erhält; der isolierte Teil von A ist also

$$(4) \qquad A_j = A - A_h = A_\alpha - A_\beta,$$

während
$$A_\alpha - A = A_\beta - A_h$$

die Menge der nicht zu A gehörigen Häufungspunkte von A ist.

Die Menge A heißt

isoliert, wenn $A_h = 0$, $A_j = A$

insichdicht, wenn $A_j = 0$, $A_h = A$,

wenn sie also aus lauter isolierten oder lauter Häufungspunkten besteht. (Wir schreiben insichdicht in einem Wort, da nach einer späteren Definition, § 25, jede Menge in sich dicht ist.)

Beispiele. In der Ebene ist die Menge der rationalen Punkte insichdicht, die der ganzzahligen isoliert; jede Menge, die überhaupt keinen Häufungspunkt hat, ist isoliert, insbesondere jede endliche Menge. Die Zahlenmenge $\{1, \frac{1}{2}, \frac{1}{3}, \ldots\}$ ist isoliert, da ihr einziger Häufungspunkt 0 ihr nicht angehört; ebenso die Menge A in Fig. 2.

Die Menge A_j kann keinen Häufungspunkt ihrer selbst (der ja auch Häufungspunkt von A wäre) enthalten, ist also selbst eine isolierte Menge. Wohl aber kann A_h isolierte Punkte ihrer selbst (die dann nicht isolierte Punkte von A sind) enthalten, braucht also nicht insichdicht zu sein. Für die Zahlenmenge $A = \{1, \frac{1}{2}, \frac{1}{3}, \ldots, 0\}$ ist $A_h = \{0\}$ sogar isoliert. Mit Verwendung mehrerer Indices, die von links nach rechts zu lesen sind, so daß z. B. $A_{jh} = (A_j)_h$ und $A_{hj} = (A_h)_j$ bedeutet, ist also

$$A_{jj} = A_j, \quad A_{jh} = 0,$$

während nicht $A_{hh} = A_h$, $A_{hj} = 0$ zu sein braucht.

Die Bedingung der insichdichten Menge kann $A \leqq A_\beta$ geschrieben werden. Das veranlaßt zu weiteren Definitionen: die Menge A heißt

$$
\begin{aligned}
&\textit{insichdicht}, && \text{wenn } A \leqq A_\beta \\
&\textit{abgeschlossen}, && \text{wenn } A \geqq A_\beta \\
&\textit{perfekt}, && \text{wenn } A = A_\beta,
\end{aligned}
$$ [55]

wenn also jeder Punkt von A Häufungspunkt ist, bzw. wenn jeder Häufungspunkt Punkt von A selber ist, bzw. wenn beides zugleich gilt. Man kann dies nach (2) auch durch Gleichungen ausdrücken, indem nämlich A_α einem der Summanden A, A_β oder beiden gleich wird: die Menge A ist

$$
\begin{aligned}
&\textit{insichdicht}, && \text{wenn } A_\alpha = A_\beta \\
&\textit{abgeschlossen}, && \text{wenn } A_\alpha = A \\
&\textit{perfekt}, && \text{wenn } A_\alpha = A_\beta = A.
\end{aligned}
$$

Zitieren wir wieder die einfachsten Beispiele aus der Ebene:

A Kreisfläche ohne Peripherie: insichdicht.

A Kreisfläche mit Peripherie und einem außerhalb liegenden Punkt: abgeschlossen.

A Kreisfläche mit Peripherie: perfekt.

Mengen ohne Häufungspunkte ($A_\beta = 0$), insbesondere endliche, zählen natürlich zu den abgeschlossenen. Die Nullmenge ist alles: isoliert, insichdicht, abgeschlossen, perfekt (offen, Randmenge).

Mit $A \leqq B$ ist natürlich $A_\lambda \leqq B_\lambda$ $(\lambda = \alpha, \beta, \gamma)$, $A_h \leqq B_h$, $A_v \leqq B_v$; diese Mengen verhalten sich monoton, aber nicht die Komplemente A_j, A_u.

Ferner ist (über die Bedeutung mehrfacher Indices s. oben)

$$A_{\lambda\alpha} = A_\lambda \qquad (\lambda = \alpha, \beta, \gamma) \tag{5}$$

$$A_{\alpha\beta} = A_\beta. \tag{6}$$

Beweis von (5). Ist $x \varepsilon A_{\lambda\alpha}$, so enthält jedes U_x einen Punkt $y \varepsilon A_\lambda$, es gibt aber, weil U_x offen ist, ein $U_y \leqq U_x$, U_y und demnach U_x enthält mindestens k_λ Punkte von A, also $x \varepsilon A_\lambda$. Demnach ist $A_{\lambda\alpha} \leqq A_\lambda$, andererseits war $A_{\lambda\alpha} \geqq A_\lambda$.

(5) besagt: *Die Mengen A_λ sind abgeschlossen.* Für jede abgeschlossene Menge $B \geqq A$ gilt $B = B_\alpha \geqq A_\alpha$: *$A_\alpha$ ist also die kleinste abgeschlossene Menge über A, die abgeschlossene Hülle von A* (S. 96).

Beweis von (6). Ist $x \varepsilon A_{\alpha\beta}$, so enthält jedes U_x einen von x verschiedenen Punkt $y \varepsilon A_\alpha$. Man wähle eine Umgebung $U_y < U_x$, die x nicht enthält; U_y enthält einen Punkt $z \varepsilon A$. Also enthält jedes U_x einen von x verschiedenen Punkt $z \varepsilon A$, x ist β-Punkt von A. Also $A_{\alpha\beta} \leqq A_\beta$, andererseits ist $A_\beta \leqq A_{\alpha\beta}$.

Aus (6) folgt: *Die Mengen A und A_α haben dieselben isolierten Punkte.*

Denn $$A_j = A_\alpha - A_\beta = A_{\alpha\alpha} - A_{\alpha\beta} = A_{\alpha j}.$$

Ist A insichdicht, so ist A_α perfekt, und vice versa.

Sind wieder $A, B, \ldots$ beliebig (auch unendlich) viele Mengen,

$$S = A \dot{+} B \dot{+} \cdots, \quad D = AB \ldots$$

8*

deren Summe und Durchschnitt, so folgt aus der Monotonie

$$S_\lambda \geqq A_\lambda \dotplus B_\lambda \dotplus \cdots, \quad D_\lambda \leqq A_\lambda B_\lambda \ldots, \qquad (\lambda = \alpha, \beta, \gamma).$$

Schärfer gilt für endlich viele, z. B. zwei Mengen:

(7) mit $S = A \dotplus B$ ist $S_\lambda = A_\lambda \dotplus B_\lambda$. $\quad (\lambda = \alpha, \beta, \gamma)$.

Es braucht nur bewiesen zu werden, daß $S_\lambda \leqq A_\lambda \dotplus B_\lambda$ oder, daß ein Punkt x, der nicht zu $A_\lambda \dotplus B_\lambda$ gehört, auch nicht zu S_λ gehört. Nun gibt es, wenn x weder von A noch von B λ-Punkt ist, Umgebungen U_x, V_x derart, daß $A\,U_x$, $B\,V_x$ von Mächtigkeiten $< k_\lambda$ sind, und ist etwa U_x die mit kleinerem Radius, so ist auch $S\,U_x = A\,U_x \dotplus B\,U_x$ von einer Mächtigkeit $< k_\lambda$, x auch von S kein λ-Punkt.

Daraus folgt:

I. *Die Summe endlich vieler und der Durchschnitt beliebig vieler abgeschlossener Mengen ist abgeschlossen.*

Dieser Satz ist mit dem Satze § 22, III ganz gleichbedeutend, denn es gilt:

II. *Die offenen und die abgeschlossenen Mengen sind Komplemente voneinander.*

Denn ist $E = A + B$ und x ein Punkt, so enthält *entweder* jedes U_x einen Punkt von A ($x \varepsilon A_\alpha$) *oder* es gibt ein U_x mit $A\,U_x = 0$, $U_x \leqq B$ ($x \varepsilon B_i$). Also:

(8) mit $E = A + B$ ist $E = A_\alpha + B_i = A_i + B_\alpha$.

Daraus folgt II. *Die Begrenzung jeder Menge ist abgeschlossen*; denn $A_g = A_r + B_r$ ist Komplement der offenen Menge $A_i + B_i$, oder $A_g = A_\alpha B_\alpha$ ist Durchschnitt von zwei abgeschlossenen Mengen.

Die abgeschlossenen Mengen können danach wie die offenen (§ 22, I II) durch reelle, im Raum *E stetige Funktionen* $f(x)$ charakterisiert werden. Die Mengen $[f > 0]$, $[f < 0]$ und ihre Summe $[f \neq 0]$ sind offen; ihre Komplemente, die Mengen $[f \leqq 0]$, $[f \geqq 0]$ und deren Durchschnitt $[f = 0]$ sind abgeschlossen; alle offenen und abgeschlossenen Mengen können so erhalten werden. Für $f(x) = \delta(x, A)$ ist $[f = 0] = A_\alpha$, also $= A$, falls A abgeschlossen ist. Die Kurven und Flächen der analytischen Geometrie sind abgeschlossene Mengen, insoweit sie durch Nullsetzen von (einer oder mehreren) stetigen Funktionen der Koordinaten definiert sind. Nicht dasselbe gilt bei *Parameterdarstellungen* (die in den Zusammenhang des achten Kapitels gehören). Eine in der Form

$$x_1 = \varphi_1(t), \qquad x_2 = \varphi_2(t)$$

oder allgemeiner $\quad f(x_1, x_2, t) = 0, \quad g(x_1, x_2, t) = 0$

mit stetigen Funktionen dargestellte Kurve A der $x_1 x_2$-Ebene (wobei t alle reellen Zahlen durchläuft) ist nicht notwendig abgeschlossen. Das ist kein Widerspruch zum Vorangehenden: wohl definieren diese Gleichungen eine abgeschlossene Menge B im $x_1 x_2 t$-Raum, aber deren Projektion A auf die $x_1 x_2$-Ebene braucht nicht abgeschlossen zu sein. Ein Beispiel für nicht abgeschlossenes A liefert schon der einfache Fall $x_1 = \frac{t}{1+|t|}$, $x_2 = 0$, wo die Kurve das offene Intervall $(-1, 1)$ der Abszissenachse wird.

Ein zweiter, nicht weniger wichtiger Zusammenhang zwischen offenen und abgeschlossenen Mengen ist dieser:

III. *Jede abgeschlossene Menge ist Durchschnitt einer Folge offener, jede* [56]
offene Menge ist Summe einer Folge abgeschlossener Mengen.

Sei nämlich A eine beliebige Menge; wir legen um jeden Punkt $x \varepsilon A$ die Umgebung $U_x(\varrho)$ mit festem Radius ϱ; die Summe dieser Punktumgebungen

$$U(A, \varrho) = \mathfrak{S}_x^A U_x(\varrho) \tag{9}$$

kann passend als *Umgebung der Menge* A mit dem Radius ϱ bezeichnet werden. Sie ist eine offene Menge, bestehend aus den Punkten y, zu denen mindestens ein Punkt $x \varepsilon A$ mit $xy < \varrho$ vorhanden ist. Wir behaupten nun

$$A_\alpha = U(A, 1)\, U(A, \tfrac{1}{2})\, U(A, \tfrac{1}{3}) \ldots, \tag{10}$$

womit A_α, also jede abgeschlossene Menge, als Durchschnitt einer Folge offener dargestellt ist. Denn der Durchschnitt rechterhand, der übrigens auch der Durchschnitt aller $U(A, \varrho)$ mit $\varrho > 0$ ist, ist ja gerade die Menge aller Punkte y, zu denen für jedes $\varrho > 0$ ein Punkt $x \varepsilon A$ mit $xy < \varrho$ vorhanden ist, d. h. aller α-Punkte von A. Die zweite Hälfte des Satzes III folgt aus der ersten durch Komplementbildung.

Der Satz ergibt sich auch unmittelbar aus den für jede reelle Funktion $f(x)$ gültigen Identitäten $(n = 1, 2, 3, \ldots)$

$$[f > 0] = \mathfrak{S}_n \left[f \geqq \frac{1}{n}\right]$$

$$[f \leqq 0] = \mathfrak{D}_n \left[f < \frac{1}{n}\right],$$

wenn man $f(x)$ stetig annimmt. Die zweite Gleichung ist nichts anderes als (10), wenn $f(x) = \delta(x, A)$ gesetzt wird.

Aus der früheren Formel

$$S = A \dot{+} B \dot{+} \cdots, \quad S_\lambda \geqq A_\lambda \dot{+} B_\lambda \dot{+} \cdots$$

folgt für $\lambda = \beta$:

IV. *Die Summe beliebig vieler insichdichter Mengen ist insichdicht.*

Dies gestattet, die Summe aller insichdichten Teilmengen von A (zu denen ja jedenfalls die Nullmenge gehört) zu bilden, die nun wieder insichdicht, also die *größte insichdichte Teilmenge* oder der *insichdichte Kern* von A ist. Wir nennen sie A_k, die nicht zum insichdichten Kern gehörigen

Punkte von A *separierte* Punkte und ihre Menge A_s den *separierten Teil* von A, womit wir eine neue, letzte Spaltung von A

$$(11) \qquad A = A_k + A_s$$

in disjunkte Summanden erhalten. Eine nur aus separierten Punkten
[57] bestehende Menge ($A_k = 0$) heißt eine *separierte* Menge, womit wir also ein neues Gegenstück zu insichdicht erhalten: A ist

$$\begin{aligned} &\textit{insichdicht}, \text{ wenn } A_s = 0, \quad A_k = A, \\ &\textit{separiert}, \text{ wenn } A_k = 0, \quad A_s = A. \end{aligned}$$

Die Menge A_s der separierten Punkte ist offenbar selbst eine separierte Menge, wie die Menge A_j der isolierten Punkte eine isolierte Menge war (womit die Namen separierter Teil, isolierter Teil sprachlich gerechtfertigt sind). Offenbar ist

$$(12) \qquad A_j \leqq A_s, \quad A_h \geqq A_k,$$

da ein isolierter Punkt von A gewiß keiner insichdichten Teilmenge von A angehört; jede insichdichte Teilmenge von A gehört zur Kohärenz A_h. Aber dann gehört sie auch zur zweiten Kohärenz A_{hh}, zur dritten Kohärenz A_{hhh} usw.; wir werden später (§ 30) sehen, wie man durch transfinite Wiederholung der Kohärenzbildung schließlich zum insichdichten Kern gelangt. Man kann auch sagen: dem separierten Teil A_s gehört nicht nur der isolierte Teil A_j von A an, sondern auch der isolierte Teil A_{hj} von A_h, der isolierte Teil A_{hhj} von A_{hh} (die erste, zweite, dritte Adhärenz von A) usw.

Beispiel. A sei die Menge der Zahlen $\frac{1}{p} + \frac{1}{q} + \frac{1}{r}$, wo p, q, r die natürlichen Zahlen durchlaufen. Dann ist A_h die Menge der Zahlen $\frac{1}{p} + \frac{1}{q}$ (die ja in der Form $\frac{1}{p} + \frac{1}{2q} + \frac{1}{2q}$ zu A gehören; außerdem ist nur noch 0 Häufungszahl von A); A_{hh} die Menge der Zahlen $\frac{1}{p}$, $A_{hhh} = 0$, also $A_k = 0$, die Menge A ist separiert.

Als Ergänzung zu den Formeln, die sich auf das Verhalten der λ-Punkte bei Summen- und Durchschnittsbildung beziehen, geben wir noch folgendes über den Durchschnitt von A mit einer *offenen* Menge an.

Es ist

$$(13) \qquad (A\,G)_\lambda \geqq A_\lambda G \qquad (G \text{ offen}) \; (\lambda = \alpha, \beta, \gamma),$$

während außerdem natürlich die triviale Formel $(A\,G)_\lambda \leqq A_\lambda G_\lambda$ gilt. Denn ist $x \,\varepsilon\, A_\lambda G$, U_x eine Umgebung, so ist $G\,U_x$ offen und es gibt eine Umgebung $V_x \leqq G\,U_x$; diese enthält mindestens k_λ Punkte von A, die aber auch zu G, also zu $A\,G$ gehören. Also enthält jedes U_x mindestens k_λ Punkte von $A\,G$, $x \,\varepsilon\, (A\,G)_\lambda$.

Anwendungen hiervon sind:

V. *Der Durchschnitt einer offenen mit einer insichdichten Menge ist insichdicht.*

Denn aus $A_\beta \geqq A$ folgt nach (13) $(AG)_\beta \geqq A_\beta G \geqq AG$.

Als Gegensatz zu IV und V sei bemerkt, daß schon der Durchschnitt von zwei insichdichten Mengen nicht insichdicht zu sein braucht. Beispiel: zwei sich schneidende Gerade der Euklidischen Ebene.

VI. *Eine offene Menge, die keinen isolierten Punkt des Raumes enthält, ist insichdicht.*

Denn ist G offen und $\leqq E_\beta$, so ist $G_\beta = (EG)_\beta \geqq E_\beta G = G$. Die für jede insichdichte Menge notwendige Bedingung, in E_β enthalten zu sein, ist also für offene Mengen auch hinreichend.

Ist A eine Menge, die keinen isolierten Punkt des Raumes enthält, so ist also A_i insichdicht, folglich

(14) $$A_i \leqq A_k, \quad A_r \geqq A_s.$$

Ist A wieder beliebig, so besteht zwischen isoliertem und separiertem Teil die Beziehung

(15) $$A_j \leqq A_s \leqq A_{j\alpha},$$

deren erste Hälfte aus (12) schon bekannt ist. Die zweite besagt, daß jeder Punkt $x \varepsilon A_s$ α-Punkt von A_j ist. Andernfalls hätte er eine Umgebung U mit $A_j U = 0$; da U offen ist, wäre nach (13)

$$AU = A_h U \leqq A_\beta U \leqq (AU)_\beta$$

und AU insichdicht, also $AU \leqq A_k$ im Widerspruch zu $x \varepsilon A_s$. A_s besteht also aus dem isolierten Teil und (gewissen) Häufungspunkten desselben; das ist die schärfere Fassung der Tatsache, daß mit A_j auch A_s verschwindet (weil dann A insichdicht ist).

Als Gegenstück zu VI beweisen wir noch den Satz

VII. *Eine abgeschlossene Menge, die keinen isolierten Punkt des Raumes enthält, ist eine Ableitung.*

Das ist also eine notwendige und hinreichende Bedingung, denn jede Ableitung A_β ist abgeschlossen und $\leqq E_\beta$. Wir gründen den Beweis auf den Hilfssatz

VIII. *Ist H die Begrenzung der offenen Menge G, so läßt sich H als Ableitung A_β einer isolierten Menge $A \leqq G$ darstellen*[1].

Wir setzen $H > 0$ voraus (für $H = 0$ kann man $A = 0$ wählen). $H = G_\alpha - G$ ist die abgeschlossene Menge der nicht zu G gehörigen Häufungspunkte von G (oder der Rand $F - F_i$ des Komplements $F = E - G$).

Für jeden Punkt $x \varepsilon G$ ist die untere Entfernung von H positiv:

$$\delta(x) = \delta(x, H) > 0.$$

[1]) Beispiel in der Euklidischen Ebene: G die Halbebene $x_2 > 0$, H die Gerade $x_2 = 0$, A die Menge der Fig. 2.

Es sei A eine *größte*, nicht erweiterungsfähige Teilmenge[1]) von G derart, daß für je zwei verschiedene Punkte von ihr

$$(16) \qquad xy \geqq \tfrac{1}{2}\,\delta(x) + \tfrac{1}{2}\,\delta(y).$$

Dann ist $A_\beta \leqq H$. Denn sei x ein Häufungspunkt von A; man kann ihn als Limes $x = \lim x_n$ einer Folge paarweise verschiedener Punkte $x_n \varepsilon A$ darstellen. Dann ist, wegen der Stetigkeit der Funktion $\delta(x)$, $\delta(x_n) \to \delta(x)$, und da in

$$x_n x_{n+1} \geqq \tfrac{1}{2}\,\delta(x_n) + \tfrac{1}{2}\,\delta(x_{n+1})$$

die linke Seite nach 0, die rechte nach $\delta(x)$ konvergiert, $\delta(x) = 0$. x ist also Häufungspunkt von G, ohne zu G zu gehören, d. h. Punkt von H: $A_\beta \leqq H$.

Andererseits ist $H \leqq A_\beta$. Sei $h \varepsilon H$ und $y \varepsilon G$ mit beliebig kleinem $hy < \delta$. Ist $y \varepsilon A$, so ist es gut; wenn nicht, so gibt es mindestens ein $x \varepsilon A$, das mit y die Relation (16) nicht erfüllt, da sonst A durch Aufnahme von y erweiterungsfähig wäre. Nun schließt man:

$$\begin{aligned} xy &< \tfrac{1}{2}\,\delta(x) + \tfrac{1}{2}\,\delta(y) \leqq \tfrac{1}{2}\,hx + \tfrac{1}{2}\,hy, \\ hx &\leqq hy + xy < \tfrac{1}{2}\,hx + \tfrac{3}{2}\,hy, \\ hx &< 3\,hy < 3\,\delta. \end{aligned}$$

In jedem Falle gibt es also zu h Punkte $x \varepsilon A$ mit beliebig kleiner Entfernung ($< 3\,\delta$), womit $H \leqq A_\beta$ und $H = A_\beta$ bewiesen ist.

Der Beweis von VII verläuft nun so. Es sei $F = F_r + F_i = H + F_i$ abgeschlossen, H also die Begrenzung der offenen Menge $G = E - F$ und demnach $H = A_\beta$ mit $A \leqq G$. Wegen $F \leqq E_\beta$ ist andererseits, nach VI, F_i insichdicht, $F_i \leqq F_\beta$. Also hat man

$$\begin{aligned} F &= F_i + H \leqq F_\beta \dotplus H \leqq F \dotplus H = F, \\ F &= F_\beta \dotplus H = F_\beta \dotplus A_\beta = (F + A)_\beta, \end{aligned}$$

F ist die Ableitung von $F + A$.

IX. *Die Summe von endlich vielen separierten Mengen ist separiert.*

Es genügt, zwei separierte Mengen zu betrachten und zu zeigen, daß ihre Summe $S = A \dotplus B$, wenn insichdicht, Null ist (dasselbe gilt für jede Teilmenge von S, S ist separiert). Da $G = E - A_\alpha$ offen ist, ist nach V $SG = BG$ insichdicht, also Null; $S \leqq A_\alpha$, $S_\alpha = A_\alpha$, $A \leqq S \leqq S_\beta = A_\beta$, A insichdicht, also $A = 0$, ebenso $B = 0$.

[58] § 24. Relative und absolute Begriffe.

Die in den letzten beiden Paragraphen definierten Mengen und Eigenschaften (A_α, A_i, abgeschlossen, offen usw.) hängen natürlich nicht nur von

[1]) Die Existenz einer solchen kann man kaum anders als durch Wohlordnung von G beweisen. Man ordne (§ 12) jeder Menge $B < G$ als Ansatzelement, wenn möglich, einen Punkt von $G - B$ zu, der zu allen Punkten von B die Relation (16) hat. Ein Abschnitt der resultierenden Wohlordnung liefert eine Menge A der verlangten Beschaffenheit.

der Menge A, sondern auch von dem sie umfassenden Raum E ab; sie sind *relativen* Charakters. Dies muß unter Umständen durch eine vollständigere Bezeichnung zum Ausdruck gebracht werden, indem wir z. B. $A_\alpha(E)$ statt A_α schreiben und im Falle $A = A_\alpha(E)$ sagen: *A ist in E abgeschlossen.* (So hieß es ja auch in § 21: eine Folge ist in E konvergent, A ist in E kompakt.) Dieselbe Punktmenge A kann in einem Raum E abgeschlossen, offen, Randmenge o. dergl. sein, ohne es in einem andern Raum $\overline{E}$ zu sein. Die Menge der irrationalen Zahlen zwischen 0 und 1 ist im Raume der irrationalen Zahlen abgeschlossen, im Raume der reellen Zahlen nicht. Eine Kreisscheibe (ohne Peripherie) ist in ihrer Ebene offen, im dreidimensionalen Raum Randmenge. Jede Menge ist in sich selbst zugleich abgeschlossen und offen.

Solche Begriffe, Eigenschaften, Beziehungen, die vom umgebenden Raum unabhängig sind, heißen *absolut.* So ist eine Menge A *absolut abgeschlossen* zu nennen, wenn sie in *jedem* Raum $E \geqq A$ abgeschlossen ist. Die Aussage „A ist insichdicht" hat absoluten Charakter, denn sie besagt: für jeden Punkt $a \varepsilon A$ und jedes $\varrho > 0$ gibt es unendlich viele Punkte von A, die von a eine Entfernung $< \varrho$ haben, und hierin ist nur von A, nicht von einem umfassenden Raum die Rede.

Die Aussage „x ist λ-Punkt von A" bedeutet, daß es für jedes $\varrho > 0$ mindestens k_λ Punkte $a \varepsilon A$ mit $ax < \varrho$ gibt (S. 112); sie setzt zwar voraus, daß x und A einem metrischen Raume E angehören, hängt aber von diesem nicht ab und ist absoluten Charakters. Sind A und D Teilmengen von E, so ist die Menge

$$A_\lambda(D)$$

der zu D gehörigen λ-Punkte von A absolut, von E unabhängig definiert. Dabei ist offenbar

$$A_\lambda(D) = D A_\lambda(E), \tag{1}$$

wobei also die rechte Seite nur scheinbar von E abhängt. Lassen wir nun, zur ursprünglichen Bezeichnung zurückkehrend, das Argument E wieder fort, so daß sich die Mengen A_λ und die Begriffe abgeschlossen, offen auf den Raum E beziehen, so wird

$$A_\lambda(D) = D A_\lambda. \tag{2}$$

Die Menge A ist *in D abgeschlossen*, wenn $A = A_\alpha(D)$, also

$$A = D A_\alpha. \tag{3}$$

Die in D abgeschlossenen Mengen sind die Durchschnitte von D mit abgeschlossenen Mengen. Denn einerseits ist (3) ein solcher Durchschnitt; andererseits, wenn

$$A = D F \tag{4}$$

Durchschnitt von A mit einer abgeschlossenen Menge F ist, so folgt aus [59]
$A \leqq A_\alpha \leqq F$ durch Schneiden mit D: $A \leqq D A_\alpha \leqq A$, also (3).

Ist $A + B = D$, so ist der offene Kern $B_i(D)$ von B im Raume D als Komplement $D - A_\alpha(D)$ zu erklären. Das gibt nach (2) die Menge $D(E - A_\alpha)$ oder

$$(5) \qquad B_i(D) = DB_i^*,$$

wo $B^* = E - A$ das Komplement von A in E und B_i^* dessen auf E bezüglicher offener Kern ist[1]). Die Menge B ist *in D offen*, wenn

$$(6) \qquad B = DB_i^*.$$

Die in D offenen Mengen sind die Komplemente (in D) der in D abgeschlossenen oder die Durchschnitte

$$(7) \qquad B = DG$$

von D mit offenen Mengen. (7) geht aus (4) hervor, wenn $G = E - F$ das offene Komplement von F bezeichnet.

Daß A in D abgeschlossen ist, ist damit gleichbedeutend, daß $A \leqq D$ ist und $D - A$ keinen Häufungspunkt von A enthält: $A_\alpha(D - A) = A_\beta(D - A) = 0$.

Der insichdichte Kern A_k von A ist in A abgeschlossen. Denn enthielte $A - A_k$ einen Häufungspunkt h von A_k, so wäre $A_k + \{h\}$ insichdicht und A_k nicht die größte insichdichte Menge $\leqq A$. *Der insichdichte Kern einer abgeschlossenen Menge ist perfekt.*

Da (2) nur von A, D und nicht vom umgebenden Raum E abhängt, so hängen insbesondere die Mengen

$$(8) \qquad A_\lambda(A) = A A_\lambda$$

nur von A ab, sie haben absoluten Charakter. Die Mengen

$$A_h = A A_\beta, \quad A_v = A A_\gamma$$

und ihre Komplemente

$$A_j = A - A_h, \quad A_u = A - A_v$$

sind durch A allein bestimmt. Demnach sind *insichdicht* und *isoliert* Absolutbegriffe, und auch der insichdichte Kern A_k wie sein Komplement $A_s = A - A_k$ hängen nur von A ab. Es sind also, wenn wir sämtliche bisher definierten Mengen noch einmal zusammenstellen,

$A_i, A_r, A_g, A_\alpha, A_\beta, A_\gamma$ vom Raume E abhängig,
$A_h, A_j, A_v, A_u, A_k, A_s$ vom Raume E unabhängig.

Natürlich ist A *in D perfekt* zu nennen, wenn A insichdicht und in D abgeschlossen ist. Dann ist $A = DA_\alpha$ der Durchschnitt von D mit einer perfekten Menge; das ist aber nicht umkehrbar: ein solcher Durchschnitt braucht nicht insichdicht zu sein. A_k ist in A perfekt.

Ist E ein *vollständiger* Raum, so ist offenbar

---

[1]) In (5) rechts heißt es also nicht DB_i, wie man analog zu (3) vermuten könnte.

(9) $$A_\alpha = A_\alpha(E) = \bar{A}$$

eine *vollständige Hülle* (S. 107) von A. *Die abgeschlossene Hülle von A in einem vollständigen Raum ist eine vollständige Hülle von A; die in einem vollständigen Raum abgeschlossenen Mengen sind vollständig.* Dieses $\bar{A}$ hängt im wesentlichen (bis auf die Bezeichnung der Elemente von $\bar{A} - A$) nur von A ab, da alle vollständigen Hüllen in der S. 107 angegebenen Weise isometrisch sind. In diesem Sinne ist $\bar{A}$ die *größte abgeschlossene Hülle* von A; denn ist $A \leqq D$ und E ein vollständiger Raum $\geqq D$ (den man z. B. durch Vervollständigung von D erhalten kann), so ist $A_\alpha(D)$ Teilmenge von $A_\alpha(E) = \bar{A}$. Ist hierbei A selbst vollständig, so ist $A = \bar{A} = A_\alpha$, A in E und erst recht in D abgeschlossen ($A = DA = DA_\alpha$); d. h. *vollständige Mengen sind absolut abgeschlossen* (in jedem Raume $D \geqq A$ abgeschlossen) *und umgekehrt*, denn eine absolut abgeschlossene Menge ist insbesondere in jedem vollständigen Raume abgeschlossen, also vollständig.

Auch die beiden andern Mengen $A_\lambda = A_\lambda(E)$ $(\lambda = \beta, \gamma)$ hängen, wenn E vollständig ist, im wesentlichen nur von A ab und sind die größtmöglichen ihrer Art.

Das Seitenstück zu den absolut abgeschlossenen wären die absolut offenen, d. h. in jedem sie enthaltenden Raum offenen Mengen; indessen ist leicht zu sehen, daß es solche Mengen (abgesehen von der Nullmenge) nicht geben kann. Denn jede Menge A kann in einen Raum E so eingelagert werden, daß sie Randpunkte, sogar lauter Randpunkte enthält, etwa so wie die Euklidische Ebene in den dreidimensionalen Raum eingebettet ist: man braucht nur (S. 102) das Produkt (A, B) von A mit der Menge B der reellen Zahlen zu bilden; in diesem Raum der Punkte (x, y) ist jede „Schicht" $y =$ constans mit A isometrisch und besteht nur aus Randpunkten.

Auch Kompaktheit ist ein Relativbegriff; A war *in E kompakt* genannt worden, wenn jede Punktfolge aus A eine in E konvergente Teilfolge enthält. Man kann auch sagen: *A ist dann und nur dann in E kompakt, wenn jede unendliche Teilmenge von A mindestens einen Häufungspunkt in E hat.*

Die *in sich kompakten* Mengen sind vollständig oder absolut abgeschlossen (S. 107); sie können auch als *absolut kompakt* (in jedem umfassenden Raum) bezeichnet werden. Ist A in irgendeinem Raum E kompakt, so ist A bereits in seiner abgeschlossenen Hülle A_α kompakt und diese in sich kompakt; ist A in E zugleich abgeschlossen und kompakt, so ist A in sich kompakt.

Die Begriffe vollständig und kompakt einerseits, abgeschlossen andererseits sind völlig verschiedenen Charakters, wie gegenüber dem auch hier leider schwankenden Sprachgebrauch betont werden muß; jene fordern in gewissen Fällen *Existenz* von α-Punkten oder Häufungspunkten, dieser verlangt, *wenn* sie existieren, ihre Zugehörigkeit zu der betrachteten Menge.

Bei den Begriffen abgeschlossen und kompakt zeigt sich der Unterschied sehr prägnant in den für $A \leqq B \leqq C$ gültigen Schlußweisen: ist A in C abgeschlossen, so auch in B; ist A in B kompakt, so auch in C.

Die Erörterungen dieses Paragraphen waren zur Klärung notwendig; sie werden uns aber nicht hindern, auch im folgenden Relativbegriffe wie abgeschlossen, offen, kompakt und die Bezeichnungen A_α usw. ohne Zusatz zu gebrauchen, wo sie sich dann auf den jeweilig betrachteten Raum beziehen. Insbesondere versteht es sich von selbst, daß zwei zugleich vorkommende Relativbegriffe bezüglich *desselben* Raumes gelten sollen; wenn wir z. B. von einer abgeschlossenen kompakten Menge sprechen, so heißt das: sie soll in E abgeschlossen und in E kompakt sein.

§ 25. Separable Räume.

In diesem und dem folgenden Paragraphen werden Mächtigkeitsfragen die Hauptrolle spielen, und zwar werden wir die Mächtigkeiten (des Raumes und gewisser Teilmengen) zunächst nach oben, sodann nach unten begrenzen. Wie dies geschieht, darüber orientiert als vorläufiges Beispiel die Art, wie die Mächtigkeit der Menge E der reellen Zahlen x bestimmt wird. Sie ist *höchstens* gleich $\aleph = \aleph_0^{\aleph_0}$, weil jedes x als $\lim r_n$ von rationalen Zahlen darstellbar, die abzählbare Menge R der rationalen Zahlen „in E dicht" ist; sie ist *mindestens* gleich $\aleph$, weil umgekehrt jeder $\lim r_n$ auch eine reelle Zahl x, d. h. E die vollständige Hülle von R oder selbst vollständig ist.

Betrachten wir wieder nur Punkte und Teilmengen eines festen metrischen Raumes E. Wir sagen, *A ist zu B dicht*, wenn

$$A_\alpha \geqq B.$$

Das besagt: jeder Punkt $b \,\varepsilon\, B$ ist α-Punkt von A oder als $b = \lim a_n$ $(a_n \,\varepsilon\, A)$ darstellbar, oder jede Umgebung U_b enthält mindestens einen Punkt von A, oder zu jedem b gibt es ein a mit beliebig kleiner Entfernung. Die letzte Fassung oder die Schreibweise $B = B A_\alpha = A_\alpha(B)$ besagt nebenbei, daß diese Eigenschaft vom umgebenden Raume E unabhängig ist. Die Dichtigkeit von A zu B liefert für die Mächtigkeit $\mathfrak{b}$ von B eine obere Schranke vermöge der Mächtigkeit $\mathfrak{a}$ von A, nämlich

$$\mathfrak{a}^{\aleph_0} \geqq \mathfrak{b}.$$

Denn es gibt nur $\mathfrak{a}^{\aleph_0}$ Punktfolgen $(a_1, a_2, \ldots)$ aus A, also höchstens so viele (in E) konvergente, und A_α hat eine Mächtigkeit $\leqq \mathfrak{a}^{\aleph_0}$.

Die Ungleichungen $A_\alpha \geqq B$ und $A_\alpha \geqq B_\alpha$ sind gleichbedeutend. Zwei *zueinander dichte* Mengen $(A_\alpha = B_\alpha)$ rechnen wir zu derselben *Dichtigkeitsklasse*; in jeder solchen Klasse gibt es genau eine abgeschlossene Menge $(F = A_\alpha = B_\alpha = \cdots = F_\alpha)$, die größte Menge der Klasse.

Wenn A zu B dicht und gleichzeitig Teilmenge von B ist, sagen wir: *A ist in B dicht.* Dann ist $A_\alpha = B_\alpha$, beide Mengen gehören zur selben Dichtigkeitsklasse.

Jede Menge A ist in A_α dicht, insbesondere in ihrer vollständigen Hülle $\bar{A}$. Der isolierte Teil A_j ist im separierten Teil A_s dicht, wegen § 23, (15). Im Euklidischen Raum E ist die Menge R der rationalen wie die Menge J der irrationalen Punkte dicht; beide sind zueinander dicht ($R_\alpha = J_\alpha = E$).

Betrachten wir alle *im Raum E dichten* Mengen A: $A_\alpha = E$. (Das bedeutet: das Komplement $B = E - A$ ist eine Randmenge, $B_i = 0$.) Lassen wir die endlichen, d. h. aus nur endlich vielen Punkten bestehenden Räume beiseite, so ist A unendlich. Unter den Mächtigkeiten aller dieser A ist eine *kleinste* $\aleph_\mu$ vorhanden; der Raum hat dann eine Mächtigkeit $\leqq \aleph_\mu^{\aleph_0}$.

Der einfachste und wichtigste Fall ist, daß in E eine abzählbare Menge R dicht ist; $E = R_\alpha$ ist also höchstens von der Mächtigkeit des Kontinuums $\aleph_0^{\aleph_0} = \aleph$. *Eine Menge, in der eine abzählbare Menge dicht ist, heißt separabel*, mit einem nicht gerade sehr suggestiven, aber nun einmal eingebürgerten Ausdruck von M. Fréchet. Eine endliche oder separable Menge heißt *höchstens separabel.*

Beispiele. Wir erwähnten schon, daß der Euklidische Raum E_n separabel ist; in ihm ist die Menge R der Punkte $r = (r_1, r_2, \ldots, r_n)$ mit rationalen Koordinaten dicht, denn zu jedem $x = (x_1, x_2, \ldots, x_n)$ gibt es ein r mit beliebig kleiner Entfernung. R ist abzählbar, von der Mächtigkeit $\aleph_0^n = \aleph_0$. — Auch der Hilbertsche Raum ist separabel; in ihm ist die Menge R der Punkte

$$r = (r_1, r_2, \ldots, r_n, 0, 0, \ldots)$$

mit nur endlich vielen von Null verschiedenen, rationalen Koordinaten dicht. Denn zu jedem $x = (x_1, x_2, \ldots)$ mit konvergenter Quadratsumme Σx_n^2 gibt es ein $(x_1, x_2, \ldots, x_n, 0, 0, \ldots)$ und zu diesem ein r mit beliebig kleiner Entfernung. — Der Raum der für $a \leqq t \leqq b$ stetigen Funktionen $x(t)$ mit der Betragsdefinition $|x| = \max |x(t)|$ und der entsprechenden Entfernungsdefinition $|x - y|$ ist separabel; in ihm ist die Menge R der Polynome

$$r(t) = r_0 + r_1 t + \cdots + r_n t^n$$

mit rationalen Koeffizienten dicht. Denn nach dem Weierstraßschen Approximationssatz gibt es zu $x(t)$ ein Polynom, und zu diesem eins mit rationalen Koeffizienten, mit beliebig kleiner Entfernung. Das gilt um so mehr bei der auf $|x| = \left(\int_a^b x(t)^2\, dt\right)^{\frac{1}{2}}$ beruhenden Entfernung. R hat, wie auch beim Hilbertschen Raume, die Mächtigkeit $\aleph_0 + \aleph_0^2 + \aleph_0^3 + \cdots = \aleph_0$.

Auch nichtseparable Räume lassen sich leicht bilden. Man mache eine beliebige Menge $M = \{m, n, \ldots\}$ zum metrischen Raum, indem man je zwei verschiedenen Elementen die Entfernung 1 zuschreibt; in M ist offenbar keine Menge $< M$ dicht und wir erhalten, wenn M eine Mächtigkeit $\mathfrak{m} > \aleph_0$ hat, einen nichtseparablen Raum. Ist dieses M, oder allgemeiner eine Menge M, in der zwei verschiedene Punkte eine Entfernung $\geqq \varrho$ haben (ϱ eine feste positive Zahl), in einem Raum E gelegen, so hat eine in E dichte Menge niemals eine Mächtigkeit $< \mathfrak{m}$. Denn sie muß zu jedem $m \varepsilon M$ einen Punkt x_m mit $m x_m < \frac{1}{2} \varrho$ enthalten, und diese Punkte sind paarweise verschieden, da, für $m \neq n$, $x_m x_n \geqq mn - m x_m - n x_n > \varrho - \frac{1}{2}\varrho - \frac{1}{2}\varrho = 0$.

Man kann diese Bemerkung noch benutzen, um in jedem Raume E eine dichte Menge geringster Mächtigkeit zu bilden. Nennen wir eine Menge $E(\varrho)$, in der je zwei verschiedene Punkte eine Entfernung $\geqq \varrho$ haben und die *nicht erweiterungsfähig* ist [1]), ein *Netz* des Raumes E, so hat jede in E dichte Menge mindestens die Mächtigkeit jedes Netzes. Ist $E(\varrho)$ ein Netz, so kann man für $\sigma < \varrho$ offenbar ein Netz $E(\sigma)$ bilden, das $E(\varrho)$ als Teilmenge enthält. Zu jedem Punkt x von E existiert ein Punkt y von $E(\varrho)$ mit $xy < \varrho$ (sonst wäre das Netz durch x erweiterungsfähig). Bildet man eine Folge von Netzen $E(1), E(\frac{1}{2}), \ldots$, so ist also

$$(1) \qquad R = E(1) \dotplus E(\tfrac{1}{2}) \dotplus \cdots$$

in E dicht, und wählt man überdies jedes Netz so, daß es das vorangehende als Teilmenge enthält, so ist die Mächtigkeit einer beliebigen in E dichten Menge mindestens gleich der von R. Denn ist $\mathfrak{a}_n$ die Mächtigkeit von $E\left(\frac{1}{n}\right) - E\left(\frac{1}{n-1}\right)$, $\mathfrak{a}_1$ die von $E(1)$, $\mathfrak{m}_n = \mathfrak{a}_1 + \cdots + \mathfrak{a}_n$ die von $E\left(\frac{1}{n}\right)$, so ist $\mathfrak{a}_1 + \mathfrak{a}_2 + \cdots$ die Mächtigkeit von R und die kleinste Mächtigkeit, die $\geqq \mathfrak{m}_n$ für alle n.

Wir haben die Netze schon S. 108 in total beschränkten Mengen verwendet, wo sie endlich waren und daher R höchstens abzählbar ist, also: *jede total beschränkte, d. h. jede bedingt kompakte Menge ist endlich oder separabel.*

Wir nehmen nun E als separabel an; die abzählbare Menge R sei in E dicht. Für jeden Punkt $r \varepsilon R$ betrachten wir die Umgebungen $U_r(\varrho)$ mit *rationalen* Radien ϱ und nennen sie *spezielle Umgebungen.* Die Menge aller speziellen Umgebungen ist abzählbar. Wir bezeichnen diese Umgebungen daher mit $V_1, V_2, \ldots$, eine einzelne spezielle Umgebung auch mit V und, wenn sie den Punkt x enthält, mit V_x.

---

[1]) Die Bildung einer solchen Menge geschieht durch Wohlordnung von E ähnlich wie S. 120.

Zu jeder Umgebung U_x gibt es ein $V_x \leqq U_x$. Denn hat U_x den Radius σ, so gibt es ein r mit $xr < \frac{1}{2}\sigma$ und eine rationale Zahl ϱ mit $xr < \varrho < \frac{1}{2}\sigma$. Dann ist $U_r(\varrho)$ ein V_x und ist in U_x enthalten, denn aus $ry < \varrho$ folgt $xy \leqq xr + ry < 2\varrho < \sigma$.

Zu jedem Punkt x einer offenen Menge G gibt es ein $V_x \leqq G$.

Jede Teilmenge eines separablen Raumes ist höchstens separabel.

Es seien V_p $(p = p_1, p_2, \ldots)$ diejenigen unter den V_n, die Punkte von A enthalten; man wähle einen Punkt $a_p \varepsilon A V_p$. Die Punkte a_p bilden eine höchstens abzählbare, in A dichte Menge. Denn ist $x \varepsilon A$, U_x eine Umgebung, $V_x \leqq U_x$, so ist V_x ein V_p und U_x enthält den Punkt a_p. — Die bewiesene Behauptung ist nicht ganz trivial, da eine in E dichte Menge R zunächst nur *zu* A dicht ist.

Erinnern wir uns an die früheren Bezeichnungen § 23, (1)

$$(2) \qquad A_v = A A_\gamma, \quad A = A_v + A_u,$$

die Spaltung von A in verdichtete und unverdichtete Punkte betreffend. Ist A höchstens abzählbar, so ist $A_\gamma = 0$ in jedem Raum. Davon gilt hier die verschärfte Umkehrung:

I. *Ist der Raum separabel und $A_v = 0$, so ist A höchstens abzählbar.*

Denn ein Punkt $x \varepsilon A$ hat, da er unverdichtet ist, eine Umgebung U_x mit höchstens abzählbarem $A U_x$; es sei $V_x \leqq U_x$. Dann ist $\underset{x}{\overset{A}{\mathfrak{S}}} A V_x = A$ höchstens abzählbar, da die Summe höchstens abzählbar viele verschiedene Summanden hat, deren jeder höchstens abzählbar ist.

Dieser Satz oder sogar einer der aus ihm folgenden und daher formal weniger besagenden Sätze:

Wenn $A_\gamma = 0$, ist A höchstens abzählbar,

Wenn $A_\beta = 0$, ist A höchstens abzählbar

ist für die (endlichen oder) separablen Räume charakteristisch; wenn etwa der letzte für jedes $A \leqq E$ gilt, ist E höchstens separabel. Das folgt unmittelbar aus (1), wo jedes Netz $E(\varrho)$, als Menge ohne Häufungspunkt, höchstens abzählbar ist.

Betrachten wir nun die drei Spaltungen

$$A = A_h + A_i = A_k + A_s = A_v + A_u$$

von A in Häufungs- und isolierte Punkte, Kern- und separierte Punkte, verdichtete und unverdichtete Punkte. A_u ist nach I höchstens abzählbar, da es keinen Verdichtungspunkt von A, geschweige von A_u enthält. Dann ist aber A_v insichdicht, denn

$$A_{u\gamma} = 0, \quad A_\gamma = A_{u\gamma} \dot{+} A_{v\gamma} = A_{v\gamma},$$
$$A_v \leqq A_\gamma = A_{v\gamma} \leqq A_{v\beta}.$$

Demnach ist

$$(3) \qquad A_v \leqq A_k \leqq A_h, \quad A_u \geqq A_s \geqq A_i$$

und mit A_u sind auch A_s, A_i höchstens abzählbar. Also:

II. *Der isolierte, der separierte, der unverdichtete Teil jeder Menge eines separablen Raumes ist höchstens abzählbar.*

Insbesondere sind isolierte, separierte, unverdichtete Mengen höch-
[60] stens abzählbar. Ist A unabzählbar, so sind auch $A_h, A_k, A_v, A_\alpha, A_\beta, A_\gamma$ unabzählbar. Eine Menge mit höchstens abzählbarer Ableitung A_β ist höchstens abzählbar; ist A_γ höchstens abzählbar, so ist A höchstens abzählbar und daher in Wahrheit $A_\gamma = 0$.

Außer den uns aus § 23, (5) (6) bekannten, allgemeingültigen Iterationsformeln

$$(4) \qquad \begin{cases} A_{\alpha\alpha} = A_\alpha, \quad A_{\beta\alpha} = A_\beta, \quad A_{\gamma\alpha} = A_\gamma, \\ \qquad\qquad A_{\alpha\beta} = A_\beta \end{cases}$$

erhalten wir hier noch

$$(5) \qquad \begin{cases} A_{\gamma\alpha} = A_{\gamma\beta} = A_{\gamma\gamma} = A_\gamma \\ \qquad A_{\alpha\gamma} = A_{\beta\gamma}. \end{cases}$$

Denn wir hatten $A_\gamma = A_{v\gamma} \leqq A_{\gamma\gamma} \leqq A_{\gamma\beta} \leqq A_{\gamma\alpha} = A_\gamma$ und die zweite Zeile folgt daraus, daß $A_\alpha - A_\beta = A_j$ höchstens abzählbar ist. In der ersten ist enthalten, daß A_γ perfekt ist.

III. *Ein System disjunkter offener Mengen des separablen Raumes ist höchstens abzählbar.*

Denn zu jeder dieser offenen Mengen $G (> 0)$ wählen wir einen Punkt $x \varepsilon G$ und ein $V_x \leqq G$; die den verschiedenen G zugeordneten V sind verschieden und es gibt deren höchstens abzählbar viele.

IV. *Es gibt im separablen Raume genau $\aleph$ offene (und ebensoviele abgeschlossene) Mengen.*

Eine offene Menge $G (> 0)$ läßt sich als Summe von speziellen Umgebungen $V_x \leqq G (x \varepsilon G)$ darstellen, also in der Form

$$G = V_{m_1} \dotplus V_{m_2} \dotplus \cdots$$

mit einer passend gewählten Menge natürlicher Zahlen $m_1, m_2, \ldots$. Es gibt also höchstens so viele G wie Mengen natürlicher Zahlen $(2^{\aleph_0} = \aleph)$. Aber es gibt auch mindestens so viele (was z.B. im Euklidischen Raum trivial ist, da schon die Umgebungen eines Punktes eineindeutig vom Radius abhängen). Sei etwa A eine abzählbare isolierte Menge; es gibt solche, denn wenn nicht etwa der ganze Raum eine ist, so sei $x \varepsilon E^\beta$ und $A = \{a_1, a_2, \ldots\}$ eine abzählbare Menge von Punkten $a_n \neq x$, die nach x konvergieren. Legt man um jeden Punkt a_n eine Umgebung U_n, die keinen andern Punkt von A enthält, so entspricht jeder Teilmenge B von A eineindeutig eine offene Menge $G = \underset{a_n}{\overset{B}{\mathfrak{S}}} U_n$, deren Durchschnitt mit A genau B ist, und die Menge dieser G hat die Mächtigkeit $2^{\aleph_0} = \aleph$.

V. *Wenn der insichdichte Kern des separablen Raumes E nicht verschwindet, gibt es genau $\aleph$ perfekte Mengen.*

Wir dürfen E durch E_k ersetzen (die in E perfekten Mengen sind mit den in E_k perfekten identisch, da E_k in E abgeschlossen ist), also annehmen, daß E selbst insichdicht (perfekt) ist. Es sei $x \varepsilon E$ und wie soeben $A = \{a_1, a_2, \ldots\}$ eine Menge von Punkten $a_n \neq x$, $a_n \to x$. Die oben verwendeten U_n können wir, indem wir ihre Radien halbieren, *disjunkt* annehmen, dann hat nicht nur $G = \sum\limits_{a_n}^{B} U_n$, sondern auch G_α mit A genau den Durchschnitt B, denn wenn z. B. a_1 nicht zu B gehört, so ist $G\,U_1 = 0$ und a_1 auch nicht α-Punkt von G. Die G sind aber insichdicht, die G_α perfekt. — Die den abzählbaren Mengen B entsprechenden G_α enthalten überdies den Punkt x; es gibt also auch genau $\aleph$ perfekte Mengen, die einen festen Punkt des perfekten Raumes enthalten.

VI. *Ein System offener Mengen im separablen Raume kann, ohne* [61]
Änderung seiner Summe, durch ein höchstens abzählbares Teilsystem ersetzt werden.

Denn ist $G = \mathfrak{S} G_m$, die Summe über eine beliebige Menge erstreckt, so suche man die (höchstens abzählbar vielen) verschiedenen speziellen Umgebungen V_p, die in den Summanden G_m enthalten sind, und wähle für jedes p ein bestimmtes $G_{m_p} \geqq V_p$ aus, dann ist

$$G = \mathfrak{S}_p V_p \leqq \mathfrak{S}_p G_{m_p} \leqq \mathfrak{S}_m G_m = G,$$

also $G = \mathfrak{S}_p G_{m_p}$. Durch Übergang zu den Komplementen folgt:

VII. *Ein System abgeschlossener Mengen im separablen Raume kann, ohne Änderung seines Durchschnitts, durch ein höchstens abzählbares Teilsystem ersetzt werden.*

§ 26. Vollständige Räume.

1. Die Durchschnittssätze. Wie die Mächtigkeit des Kontinuums im vorigen Paragraphen als obere, so wird sie in diesem für gewisse Mengen als untere Grenze erscheinen. Die Grundlage dafür bildet der nachher folgende zweite Durchschnittssatz, dem wir als Gegenstück einen ersten nebst Folgerungen vorausschicken. Dieser bezieht sich auf in sich kompakte, jener auf beschränkte vollständige Mengen.

I (Erster Durchschnittssatz). *Eine absteigende Folge $A_1 \geqq A_2 \geqq \cdots$ kompakter, abgeschlossener, von Null verschiedener Mengen hat einen von Null verschiedenen Durchschnitt.*

Wir wählen aus jeder Menge A_n einen Punkt a_n. Die Folge $(a_1, a_2, \ldots)$ hat als Punktfolge aus A_1 eine konvergente Teilfolge mit dem Limes x, der auch Häufungspunkt jeder Restfolge $(a_n, a_{n+1}, \ldots)$ ist. Also ist x ein α-Punkt jeder Menge A_n, d. h. Punkt von A_n; $x \varepsilon A_1 A_2 \ldots$.

Der Raum E ist hierbei ganz beliebig; wir hätten ja auch, statt von kompakten abgeschlossenen[1]), von in sich kompakten Mengen sprechen können.

II (Satz von E. Borel). *Ist eine kompakte, abgeschlossene Menge in der Summe einer Folge offener Mengen enthalten, so ist sie bereits in einer Summe von endlich vielen dieser offenen Mengen enthalten.*

Es sei $A \leqq S = G_1 \dot{+} G_2 \dot{+} \cdots$, A kompakt abgeschlossen, G_n offen. Die Behauptung lautet: es gibt ein n mit $A \leqq S_n = G_1 \dot{+} \cdots \dot{+} G_n$. In der Tat bilden die kompakten abgeschlossenen Mengen $A(E - S_n)$ eine absteigende Folge mit dem Durchschnitt $A(E - S) = 0$; sie können also nach I nicht alle von Null verschieden sein, und es gibt ein n mit $A(E - S_n) = 0$, d. h. $A \leqq S_n$.

Aus der Verbindung mit § 25, VI folgt eine Verschärfung des Borelschen Satzes:

[62] III. *Ist eine kompakte abgeschlossene Menge des separablen Raumes in der Summe eines beliebigen Systems offener Mengen enthalten, so ist sie bereits in einer Summe von endlich vielen dieser offenen Mengen enthalten.*

Auch der erste Durchschnittssatz läßt sich mit Hilfe von § 25, VII verschärfen:

IV. *Ist der Raum separabel, so hat ein System kompakter abgeschlossener Mengen einen von Null verschiedenen Durchschnitt, falls endlich viele Mengen des Systems stets einen von Null verschiedenen Durchschnitt haben.*

Denn der Durchschnitt des ganzen Systems läßt sich als Durchschnitt $A_1 A_2 \ldots$ einer Folge von Mengen des Systems darstellen, d. h. als Durchschnitt der absteigenden, von Null verschiedenen, kompakten abgeschlossenen Mengen $A_1 \geqq A_1 A_2 \geqq \cdots$.

V (Zweiter Durchschnittssatz). *In einem vollständigen Raum hat eine absteigende Folge $A_1 \geqq A_2 \geqq \cdots$ beschränkter, abgeschlossener, von Null verschiedener Mengen, deren Durchmesser nach 0 konvergieren, genau einen gemeinsamen Punkt.*

Wir wählen wieder $a_n \varepsilon A_n$; die Folge $(a_1, a_2, \ldots)$ ist dann eine Fundamentalfolge, da a_n von allen folgenden (auch zu A_n gehörigen) Punkten einen Abstand $\leqq d(A_n)$ hat. Also existiert $x = \lim a_n$ und wie bei I sieht man, daß $x \varepsilon A_1 A_2 \ldots$, welcher Durchschnitt hier wegen $d(A_n) \to 0$ nur einen Punkt haben kann.

Man hätte natürlich auch hier die Beziehung auf den Raum E beseitigen und einfach von vollständigen Mengen A_n sprechen können. Keiner der Durchschnittssätze ist Folge des andern; der zweite läßt zwar beschränkte statt kompakter Mengen zu, hat aber dafür die Durchmesser-

[1]) Das soll natürlich heißen: in E kompakten und in E abgeschlossenen; vgl. die Schlußbemerkung von § 24.

bedingung. Diese ist nicht zu entbehren, wie das Beispiel $A_n = \{x_n, x_{n+1}, \ldots\}$ zeigt, worin $x_1, x_2, \ldots$ Punkte bedeuten, die paarweise die Entfernung 1 haben. In Euklidischen Räumen, wo beschränkte und kompakte Mengen identisch sind, ist der erste Durchschnittssatz inhaltreicher; der einfachste Fall, daß eine absteigende Folge abgeschlossener Intervalle $[a_n, b_n]$, also mit $a_n \leqq a_{n+1} < b_{n+1} \leqq b_n$, als Durchschnitt ein abgeschlossenes Intervall oder einen Punkt hat, ist dem Leser aus den Elementen bekannt.

2. Dyadische Mengen. In einem vollständigen Raume E sei ein System *abgeschlossener*, *beschränkter*, *von Null verschiedener* Mengen $V_p, V_{pq}, V_{pqr}, \ldots$ gegeben, deren Indices die beiden Werte 1, 2 durchlaufen, so daß wir also zwei Mengen V_1, V_2 mit einem Index, vier Mengen $V_{11}, V_{12}, V_{21}, V_{22}$ mit zwei Indices, allgemein 2^n Mengen mit n Indices vor uns haben. Für jede dyadische (aus den Ziffern 1, 2 gebildete) Ziffernfolge $(p, q, r, \ldots)$ soll überdies

$$V_p \geqq V_{pq} \geqq V_{pqr} \geqq \cdots \tag{1}$$

gelten und die Durchmesser dieser Mengen sollen nach 0 konvergieren; nach V besteht der Durchschnitt $V_p V_{pq} V_{pqr} \ldots$ aus einem einzigen Punkt x. Die Menge aller dieser Punkte

$$D = \mathfrak{S} V_p V_{pq} V_{pqr} \cdots, \tag{2}$$

die Summe über alle dyadischen Ziffernfolgen erstreckt, soll eine *dyadische Menge* heißen. Der Name bezeichnet nur die Darstellungsform; was die Natur dieser Mengen selbst betrifft, werden wir gleich sehen, daß dyadische und in sich kompakte Mengen identisch sind.

Es gilt

$$D = \mathfrak{S} V_p \cdot \mathfrak{S} V_{pq} \cdot \mathfrak{S} V_{pqr} \cdots, \tag{3}$$

so daß sich die Menge D als abgeschlossen (im vollständigen Raum E), also als vollständig erweist. In der Tat ist die Menge (2) zunächst in (3) enthalten. Umgekehrt sei x Punkt der Menge (3), etwa $x \varepsilon V_{p_1} V_{p_2 q_2} V_{p_3 q_3 r_3} \cdots$. Bei den hier auftretenden Mengen V muß aber unendlich oft der gleiche erste Index $p_n = p$ vorkommen, bei diesen V wieder unendlich oft der gleiche zweite Index $q_n = q$, bei diesen V unendlich oft der gleiche dritte Index $r_n = r$ usw. Dann ist aber $x \varepsilon V_p V_{pq} V_{pqr} \ldots$, x Punkt der Menge (2).

Die für die Mengen (1) geforderte Konvergenz der Durchmesser nach 0 ist für alle dyadischen Ziffernfolgen gleichmäßig. D. h. wenn δ_n der größte Durchmesser der Mengen V mit n Indices ist, wobei offenbar $\delta_1 \geqq \delta_2 \geqq \cdots$ gilt, so ist $\delta_n \to 0$. Denn seien $V_{p_1}, V_{p_2 q_2}, V_{p_3 q_3 r_3}, \ldots$ Mengen mit den Durchmessern $\delta_1, \delta_2, \delta_3, \ldots$; wie soeben schließt man, daß unendlich oft $p_n = p$, dann $q_n = q$, dann $r_n = r$ usw. sein muß. Also kann nicht $\delta_n \geqq \delta > 0$ sein, da sonst $V_p, V_{pq}, V_{pqr}, \ldots$ Durchmesser $\geqq \delta$ hätten.

Nach § 21, 4 ist demnach die Menge D total beschränkt; überdies war sie vollständig, also ist sie in sich kompakt.

9*

Bevor wir umgekehrt die in sich kompakten Mengen als dyadische darstellen, verallgemeinern wir die letztgenannten folgendermaßen. Im vollständigen Raum E sei ein System abgeschlossener, beschränkter, von Null verschiedener Mengen $W_i, W_{ik}, W_{ikl}, \ldots$ gegeben, wo nun jeder Index eine endliche, eventuell von den vorangehenden Indices abhängige Menge von mindestens zwei Ziffern durchläuft. In W_i sei also $i = 1, 2, \ldots, m$; in W_{ik} sei $k = 1, 2, \ldots, m_i$; in W_{ikl} sei $l = 1, 2, \ldots, m_{ik}$ usw. ($m, m_i, m_{ik}, \ldots \geqq 2$). Für jede in dieser Weise gebildete Ziffernfolge $(i, k, l, \ldots)$ sei

$$W_i \geqq W_{ik} \geqq W_{ikl} \geqq \cdots$$

mit Durchmessern, die nach 0 konvergieren, so daß der Durchschnitt dieser Mengen einpunktig ist; die Summe

$$P = \mathfrak{S}\, W_i W_{ik} W_{ikl} \cdots \tag{4}$$

heiße eine *polyadische* Menge. Man schließt wie oben, daß

$$P = \mathfrak{S}\, W_i \,.\, \mathfrak{S}\, W_{ik} \,.\, \mathfrak{S}\, W_{ikl} \cdots$$

und die Konvergenz der Durchmesser nach 0 gleichmäßig ist. In Wahrheit ist sogar P nichts anderes als eine dyadische Menge, wie wir folgendermaßen erkennen.

Wir bezeichnen die Mengen $W_i, W_{ik}, \ldots$ allgemein mit W, die aus einem W durch Anhängung eines weiteren Index entstehenden Mengen mit $W^1, W^2, \ldots$ (z. B. für $W = W_{ik}$: $W^l = W_{ikl}$). Sodann werden wir Mengen $V_p, V_{pq}, \ldots$ mit dyadischen Indices (1 oder 2) bilden, die wieder allgemein mit V bezeichnet werden; die aus V durch Anhängung eines weiteren Index entstehenden Mengen seien V^1, V^2. Dabei soll entweder V gleich einem einzelnen W

$$V = W, \tag{α}$$

oder V die Summe mehrerer W mit *gleichvielen* Indices

$$V = W_{\mathrm{I}} \dot{+} W_{\mathrm{II}} \dot{+} \cdots \tag{β}$$

sein, wobei die Summanden in lexikographischer Ordnung der Indices geschrieben sein mögen. Wir beginnen die Definition mit

$$V_1 = W_1 \dot{+} W_3 \dot{+} \cdots, \quad V_2 = W_2 \dot{+} W_4 \dot{+} \cdots$$

und setzen sie durch Induktion fort, indem wir V^1, V^2 definieren, falls V bereits definiert ist: im Falle (α) sei

$$V^1 = W^1 \dot{+} W^3 \dot{+} \cdots, \quad V^2 = W^2 \dot{+} W^4 \dot{+} \cdots;$$

im Falle (β) sei

$$V^1 = W_{\mathrm{I}} \dot{+} W_{\mathrm{III}} \dot{+} \cdots, \quad V^2 = W_{\mathrm{II}} \dot{+} W_{\mathrm{IV}} \dot{+} \cdots.$$

Beispielsweise würden sich aus den ersten Mengen W

$$\begin{array}{lll} W_1 & W_2 & W_3 \\ W_{11} W_{12} W_{13} & W_{21} W_{22} W_{23} W_{24} & W_{31} W_{32} \end{array}$$

die ersten Mengen V folgendermaßen ergeben:

$$\begin{array}{llll}
V_1 \;\;\;= W_1 \dot{+} W_3 & & V_2 \;\;\;= W_2 & \\
V_{11} \;\;= W_1 & V_{12} \;\;= W_3 & V_{21} \;\;= W_{21} \dot{+} W_{23} & V_{22} \;\;= W_{22} \dot{+} W_{24} \\
V_{111} \;= W_{11} \dot{+} W_{13} & V_{112} \;= W_{12} & V_{121} = W_{31} & V_{122} = W_{32} \\
V_{211} \;= W_{21} & V_{212} \;= W_{23} & V_{221} = W_{22} & V_{222} = W_{24} \\
V_{1111} = W_{11}, & V_{1112} = W_{13}. & &
\end{array}$$

Man sieht nun unmittelbar: jedes V ist ein W oder Summe endlich vieler W, also abgeschlossen, beschränkt, von Null verschieden. Es ist $V \geqq V^1$, $V \geqq V^2$, also sind für jede dyadische Ziffernfolge die Ungleichungen (1) erfüllt. Ist V im Falle (β), so sind V^1, V^2 Summen mit weniger Gliedern als V, und nach hinreichend oftmaliger Anhängung von Indices kommt immer ein V des Falles (α) wieder; jede Folge $V_p, V_{pq}, V_{pqr}, \ldots$ hat eine Teilfolge $W_i, W_{ik}, W_{ikl}, \ldots$ und demnach Durchmesser, die nach 0 konvergieren, wobei zugleich die Durchschnitte

$$V_p V_{pq} V_{pqr} \ldots = W_i W_{ik} W_{ikl} \ldots$$

identisch sind. Umgekehrt ist jedes W einem bestimmten V gleich und jede Folge $W_i, W_{ik}, W_{ikl}, \ldots$ bestimmt eine Folge $V_p, V_{pq}, V_{pqr}, \ldots$, von der sie Teilfolge ist (z. B. $W_1, W_{13}, \ldots$ die Folge $V_1, V_{11}, V_{111}, V_{1112}, \ldots$). Demnach ist die von den W erzeugte polyadische Menge P gleich der von den V erzeugten dyadischen Menge D.

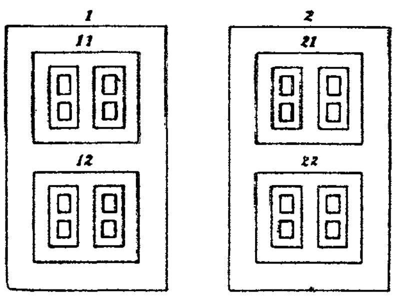

Fig. 3.

Eine in sich kompakte Menge P ist als polyadische Menge darstellbar. Denn sie ist Summe endlich vieler Mengen W_i von beliebig kleinen Durchmessern $\leqq \delta_1$, wobei jede dieser Mengen abgeschlossen (in P), also wieder in sich kompakt und ihre Anzahl $\geqq 2$ angenommen werden kann. Indem man dies fortsetzt, erhält man

$$P = \mathfrak{S}_i W_i, \quad W_i = \mathfrak{S}_k W_{ik}, \quad W_{ik} = \mathfrak{S}_l W_{ikl}, \ldots$$

wobei die W in sich kompakt (also vollständig und beschränkt), die Durchmesser der Mengen mit n Indices $\leqq \delta_n$ sind und $\delta_n \to 0$; P ist die von den W erzeugte polyadische Menge (4).

Dyadische Mengen und in sich kompakte Mengen sind also identisch.

Wir betrachten jetzt insbesondere eine dyadische Menge, wo V_1 und V_2 *disjunkt* sind[1]), ebenso V_{p1} und V_{p2}, V_{pq1} und $V_{pq2}, \ldots$; allgemein

[1]) Fig. 3 veranschaulicht den Fall abgeschlossener Rechtecksflächen V, von denen nur die Umfänge (ohne Schraffierung des Innern) gezeichnet sind.

sollen, für jedes n, die Mengen mit n Indices paarweise disjunkt sein. Eine solche Menge

$$D = \Sigma V_p V_{pq} V_{pqr} \ldots = \Sigma V_p \cdot \Sigma V_{pq} \cdot \Sigma V_{pqr} \ldots$$

heiße ein *dyadisches Diskontinuum* [1]). Sie ist äquivalent mit der Menge der dyadischen Ziffernfolgen, also von der Mächtigkeit $2^{\aleph_0} = \aleph$ des Kontinuums; ferner ist sie *perfekt* und sogar genauer

$$D = D_\alpha = D_\beta = D_\gamma,$$

jeder Punkt von D Verdichtungspunkt. Denn bei Festhaltung endlich vieler Anfangsziffern erhält man die Mengen $D V_p$, $D V_{pq}$, $D V_{pqr}, \ldots$ (z. B. $D V_1$ = Menge der Punkte $x \varepsilon D$ mit $p = 1$), die immer noch von der Mächtigkeit $\aleph$ sind und nach 0 konvergente Durchmesser haben; ist x der Durchschnittspunkt dieser Mengen, so enthält jede Umgebung U_x $\aleph$ Punkte von D.

Ein polyadisches Diskontinuum (wo die Mengen W mit n Indices disjunkt sind, für jedes n) ist mit einem dyadischen identisch.

Das klassische Beispiel eines dyadischen Diskontinuums ist das folgende, das wegen seiner Beziehung zu den triadischen Brüchen auch als *triadische Menge von G. Cantor* bezeichnet wird. In der Menge E der reellen Zahlen sei V das abgeschlossene Intervall $[0, 1]$, ferner (wenn wir aus einem sogleich ersichtlichen Grunde die dyadischen Ziffernfolgen mit den Ziffern 0, 2 statt 1, 2 bilden) $V_0 = \left[0, \frac{1}{3}\right]$ das linke und $V_2 = \left[\frac{2}{3}, 1\right]$ das rechte Drittel von V, ebenso V_{p0} das linke und V_{p2} das rechte Drittel von V_p usw. Hierbei besteht, wie eine leichte Überlegung lehrt, der Durchschnitt $V_p V_{pq} V_{pqr} \ldots$ aus der Zahl

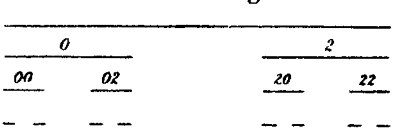

Fig. 4.

$$x = \frac{p}{3} + \frac{q}{3^2} + \frac{r}{3^3} + \cdots \qquad (p, q, r, \ldots = 0, 2)$$

und D ist die Menge der Zahlen x, die *wenigstens* eine solche triadische Entwicklung mit lauter Nullen und Zweien, ohne Einsen, zulassen. Der Zusatz *wenigstens* bezieht sich auf die (triadisch rationalen) Zahlen zwischen 0 und 1, die zwei triadische Entwicklungen haben; wenn eine davon frei von Einsen ist, so soll die Zahl zu D gehören. Z. B. gehören die Zahlen

$$\frac{1}{3} = \frac{0}{3} + \frac{2}{3^2} + \frac{2}{3^3} + \cdots \quad \left(= \frac{1}{3} + \frac{0}{3^2} + \frac{0}{3^3} + \cdots\right)$$

$$\frac{2}{3} = \frac{2}{3} + \frac{0}{3^2} + \frac{0}{3^3} + \cdots \quad \left(= \frac{1}{3} + \frac{2}{3^2} + \frac{2}{3^3} + \cdots\right)$$

zu D, nicht aber eine Zahl zwischen beiden, da deren erste triadische Ziffer

[1]) Der Name erklärt sich aus § 29, XVI. In § 36, 1 werden wir als Gegenstück das dyadische Kontinuum einführen.

sicher $p = 1$ ist. Das offene Komplement $E - D$ besteht, außer den Halbgeraden $(-\infty, 0)$ und $(1, \infty)$, aus den mittleren Dritteln von $V, V_p, V_{pq}, \ldots$, also aus den offenen Intervallen $\left(\frac{1}{3}, \frac{2}{3}\right), \left(\frac{1}{9}, \frac{2}{9}\right), \left(\frac{7}{9}, \frac{8}{9}\right), \ldots$. Es ist, wie leicht zu sehen, in E dicht, so daß D kein noch so kleines Intervall enthält (D ist in E nirgendsdicht, § 27). Der Grenzwert, dem die Längen der Intervallsummen $\Sigma V_p, \Sigma V_{pq}, \ldots$ zustreben, wird das (Lebesguesche) Längenmaß von D genannt. Hier sind diese Längen $\frac{2}{3}, \frac{4}{9}, \ldots$, also das Längenmaß Null; es erscheint paradox, daß eine Menge vom Längenmaß 0 doch die Mächtigkeit $\aleph$ des ganzen Intervalls haben kann. Man kann aber die Intervallängen anders, insbesondere die offenen Intervalle von $E - D$ kleiner und die Längensummen $\lambda_1, \lambda_2, \ldots$ von $\Sigma V_p, \Sigma V_{pq}, \ldots$ so langsam abnehmend wählen, daß das Längenmaß $\lambda = \lim \lambda_n$ von D positiv wird und sogar beliebig nahe an 1 (wenn auch natürlich < 1) liegt; dann ist es wieder paradox, daß eine Menge positiven Längenmaßes kein noch so kleines Intervall enthält.

Die Übertragung dieser einfachen Konstruktion auf die Ebene ist evident. Man teilt z. B. ein Quadrat V (abgeschlossene Quadratfläche, definiert etwa durch $0 \leqq x_1 \leqq 1$, $0 \leqq x_2 \leqq 1$, also Produkt zweier Intervalle $[0, 1]$) durch Drittelung der Seiten in neun Teilquadrate, von denen V_1, V_2, V_3, V_4 die an die Ecken von V anstoßenden seien; $V - \Sigma V_p$ ist ein Kreuz, gebildet aus den fünf übrigen Teilquadraten mit teilweise fehlenden Rändern. In gleicher Weise werden aus V_p die Teilquadrate $V_{p1}, V_{p2}, V_{p3}, V_{p4}$ gebildet usw. Das von den V erzeugte polyadische Diskontinuum

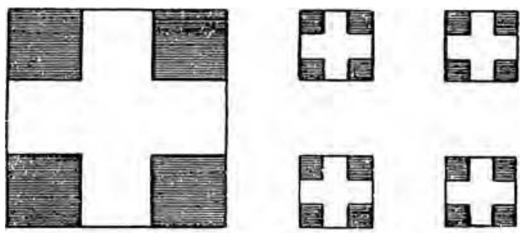

Fig. 5.

$$D = \Sigma V_p \cdot \Sigma V_{pq} \cdot \Sigma V_{pqr} \cdots$$

ist Produkt von zwei Cantorschen triadischen Mengen. Sein Flächenmaß (Grenzwert der Flächeninhalte von $\Sigma V_p, \Sigma V_{pq}, \ldots$) ist 0, kann aber durch abgeänderte Konstruktion auch positiv < 1 gemacht werden.

3. Die Mächtigkeitssätze. [63]

VI (G. Cantor). *In einem vollständigen Raum hat jede perfekte Menge* > 0 *mindestens die Mächtigkeit* $\aleph$.

VII (G. Cantor). *In einem vollständigen Raum hat jede abgeschlossene Menge, deren insichdichter Kern nicht verschwindet, mindestens die Mächtigkeit* $\aleph$.

VIII (W. H. Young). *In einem vollständigen Raum hat jede Menge $Y = G_\delta$ (Durchschnitt einer Folge offener Mengen), deren insichdichter Kern nicht verschwindet, mindestens die Mächtigkeit* $\aleph$.

Dem Beweise mögen einige erläuternde Bemerkungen vorausgehen. Jeder dieser Sätze ist Sonderfall des folgenden (eine abgeschlossene Menge F ist Durchschnitt einer Folge offener Mengen G). Andererseits folgt aus VI auch VII; und später wird sich (wegen der Homöomorphie der G_δ in vollständigen Räumen mit vollständigen Räumen selbst; § 38, III) auch VIII als Folgerung aus VII ergeben. Die Mengen

$$G_\delta = \text{Durchschnitt einer Folge offener Mengen,}$$
$$F_\sigma = \text{Summe einer Folge abgeschlossener Mengen}$$

bilden nach den offenen Mengen G und abgeschlossenen Mengen F die nächste Klasse der Borelschen Mengen (§ 32), auf die wir künftig den Mächtigkeitssatz ausdehnen werden. Diese Mengencharaktere sind natürlich relativ, vom Raume E abhängig, und man hätte ausführlich $F(E)$ usw. zu schreiben. Für $D \leqq E$ ist aber, wenn das Argument E wieder weggelassen wird,

$$F(D) = DF, \quad G(D) = DG, \quad F_\sigma(D) = DF_\sigma, \quad G_\delta(D) = DG_\delta,$$

die auf D bezüglichen Mengen sind die Durchschnitte von D mit den entsprechenden auf E bezüglichen. Es gibt auch *absolute* F_σ, G_δ, Mengen, die diesen Charakter in *jedem* sie umfassenden Raume haben. Von den F_σ ist das evident; ein absolutes F_σ ist (z. B. in seiner vollständigen Hülle betrachtet) Summe einer Folge absolut abgeschlossener (vollständiger) Mengen, die einem metrischen Raume angehören, und vice versa. Gegen die Existenz absoluter G_δ scheint zunächst die Nichtexistenz absolut
[64] offener Mengen zu sprechen, aber es verhält sich doch anders: die drei Aussagen

A ist ein G_δ in seiner vollständigen Hülle,
A ist ein G_δ in einem vollständigen Raum,
A ist ein G_δ in jedem umfassenden Raum, ein absolutes G_δ

sind gleichbedeutend. Denn ist A ein G_δ in E, so auch in jedem kleineren Raume D ($A \leqq D \leqq E$); ist andererseits A ein G_δ in seiner abgeschlossenen Hülle A_α, so ist $A = A_\alpha G_\delta$ auch im Raum E ein G_δ, weil A_α abgeschlossen, also ein G_δ und der Durchschnitt zweier G_δ wieder ein G_δ ist. Für vollständiges E geht hieraus die Gleichwertigkeit der drei Aussagen hervor. Die absoluten G_δ werden wir mit Rücksicht auf den Youngschen
[65] Satz VIII häufig *Youngsche Mengen* nennen. — Danach können wir in den drei Mächtigkeitssätzen die Beziehung auf den Raum tilgen und sagen:

IX. *Jede absolut perfekte Menge, jede absolut abgeschlossene (vollständige) Menge, jedes absolute G_δ (Youngsche Menge) hat, wenn ihr insichdichter Kern nicht verschwindet, mindestens die Mächtigkeit* $\aleph$.

Gehen wir nun zum Beweise über; wir brauchen nur den dritten Satz zu beweisen, werden aber die beiden ersten ohnehin unterwegs vorfinden. Wir werden zeigen, daß die fraglichen Mengen ein dyadisches Diskontinuum D enthalten, und werden die erzeugenden abgeschlossenen (vollständigen) Mengen V hier speziell als abgeschlossene Kugeln wählen. Wir unterscheiden nämlich

$$\textit{offene Kugel } U_x(\varrho) = \text{Menge der Punkte } y \text{ mit } xy < \varrho,$$
$$\textit{abgeschlossene Kugel } V_x(\varrho) = \text{Menge der Punkte } y \text{ mit } xy \leqq \varrho$$

(die offenen Kugeln sind unsere alten Umgebungen); sie mögen, bei gleichem Radius und Mittelpunkt, zusammengehörig oder einander zugehörig heißen. Zusammengehörige U, V bezeichnen wir mit denselben Indices. (Im Euklidischen Raum ist dann U der offene Kern von V, V die abgeschlossene Hülle von U.)

Nun sei $A > 0$ eine insichdichte Menge im vollständigen Raum E. Wir wählen in A zwei Punkte a_1, a_2 und machen sie zu Mittelpunkten disjunkter abgeschlossener Kugeln V_1, V_2, denen die offenen Kugeln U_1, U_2 zugehören. Jede der beiden Mengen $A\,U_p$ ist > 0 und insichdicht (§ 23, V); wir können also die Konstruktion wiederholen und zwei ihrer Punkte a_{p1}, a_{p2} mit disjunkten abgeschlossenen Kugeln V_{p1}, $V_{p2} \leqq U_p$ umgeben. Jede der vier Mengen $A\,U_{pq}$ ist > 0 und insichdicht; wir umgeben zwei ihrer Punkte a_{pq1}, a_{pq2} mit disjunkten abgeschlossenen Kugeln V_{pq1}, $V_{pq2} \leqq U_{pq}$ und fahren so fort.

Nichts hindert dabei, die Kugelradien beliebig klein zu wählen, etwa die der V_p, V_{pq}, $V_{pqr}, \ldots$ kleiner als $1, \frac{1}{2}, \frac{1}{3}, \ldots$. Wir erhalten so ein dyadisches Diskontinuum D, und dieses ist in A_α enthalten; denn ein Punkt x von D, der zu $V_p V_{pq} V_{pqr} \ldots$ gehört, ist Limes der Kugelmittelpunkte $a_p, a_{pq}, a_{pqr}, \ldots$. Also:

X. *Ist A eine insichdichte, von Null verschiedene Menge des vollständigen Raumes E, so enthält A_α ein dyadisches Diskontinuum D.*

Damit ist VI bewiesen (ist A insichdicht, so ist A_α perfekt, und vice versa), ferner VII: wenn F abgeschlossen ist und A enthält, so ist $F \geqq A_\alpha \geqq D$. Zum Beweis von VIII nehmen wir

$$A \leqq Y = G_1 G_2 G_3 \ldots \quad (G_n \text{ offen})$$

an und können die Kugelradien dann auch noch so klein wählen, daß $V_p \leqq G_1$ (da doch $a_p \varepsilon G_1$), $V_{pq} \leqq G_2$, $V_{pqr} \leqq G_3, \ldots$ also $D \leqq Y$.

Damit sind unsere Sätze bewiesen; wir haben sogar noch etwas mehr gezeigt, nämlich daß die betreffenden Mengen eine perfekte Teilmenge D enthalten. Ferner, um bei dem allgemeinsten Fall (VIII) einer Youngschen Menge Y zu bleiben: da $D V_1$ von der Mächtigkeit $\aleph$ ist und hierbei V_1 eine beliebig kleine Kugel um einen beliebigen Punkt a_1 von A bedeuten kann, so ist jeder Punkt von A Verdichtungspunkt von Y, und da A eine beliebige insichdichte Teilmenge von Y war, so heißt dies: *jeder Punkt des insichdichten Kerns von Y ist Verdichtungspunkt von Y,*

$$(5) \qquad Y_k \leqq Y_v.$$

Kombinieren wir dies alles schließlich mit den Tatsachen, die in einem separablen Raume gelten (§ 25): hier hat jede Punktmenge höchstens die Mächtigkeit $\aleph$, ferner war dort stets $A_v \leqq A_k$, also für eine Youngsche Menge $Y_k = Y_v$, und $A_u \geqq A_s$ höchstens abzählbar. Also:

XI. *In einem vollständigen separablen Raum ist jede Youngsche Menge $Y = G_\delta$ entweder höchstens abzählbar oder von der Mächtigkeit $\aleph$, je nachdem ihr insichdichter Kern $Y_k = Y_v$ verschwindet oder nicht.*

Für eine abgeschlossene Menge ist $Y_k = Y_\gamma$, für eine perfekte $Y = Y_\gamma$. (Daß jede Menge A_γ perfekt ist, gilt nach S. 128 schon in jedem separablen Raum; hier ergibt sich auch die Umkehrung, daß jede perfekte Menge ein A_γ ist.)

In einem separablen Raum läßt sich jede abgeschlossene Menge F vermöge $F = F_k + F_s$ in einen *perfekten Teil* F_k und einen *höchstens abzählbaren Teil* F_s spalten (der insichdichte Kern einer abgeschlossenen Menge ist perfekt, S. 122): eine Tatsache, die als Cantor-Bendixsonscher Satz bezeichnet wird. Die Zerlegung $A = P + Q$ einer Menge in einen perfekten Teil P und einen höchstens abzählbaren Teil Q ist aber im allgemeinen nicht eindeutig; ist z. B. E der Raum der rationalen Zahlen, P die Menge der rationalen Zahlen $\geqq a$ (a beliebig), so ist $E = P + (E - P)$ eine solche Zerlegung, denn P ist in E perfekt. Im vollständigen separablen Raum ist aber diese Spaltung, wenn überhaupt, so nur auf eine Weise möglich; man erhält nämlich $A_\gamma = P_\gamma \dotplus Q_\gamma = P + 0 = P$; die Spaltung ist also nur dann ausführbar, wenn $A_\gamma \leqq A$ (also insbesondere für die abgeschlossenen Mengen) und dann nur in der Weise $A = A_\gamma + (A - A_\gamma) = A_v + A_u$; für Youngsche Mengen dieser Gattung ist das mit $A = A_k + A_s$ identisch.

Eine abzählbare Menge A des vollständigen Raumes E, deren insichdichter Kern nicht verschwindet, ist gewiß kein G_δ, denn $A_k > 0$ und $A_v = 0$ widerspricht der Bedingung $Y_k \leqq Y_v$. Als abzählbare Menge ist sie aber ein F_σ. Beispiel: die Menge der rationalen Punkte im Euklidischen Raum. Ihr Komplement $E - A$ ist ein G_δ, aber kein F_σ.

§ 27. Mengen erster und zweiter Kategorie.

Wir betrachten einen beliebigen Raum E. *Die abgeschlossene Menge F heiße in E nirgendsdicht, wenn ihr offenes Komplement G in E dicht ist.* Das ist mit $G_\alpha = E$, also $F_i = 0$ gleichbedeutend: *abgeschlossene nirgendsdichte Mengen sind mit abgeschlossenen Randmengen identisch.* Sie sind auch mit den *Begrenzungen offener Mengen* identisch; denn ist F nirgendsdicht, so ist $F = F_r$ die Begrenzung seines Komplements G, und andererseits, ist G eine beliebige offene Menge und F ihr Komplement, so ist die Begrenzung F_r von G eine Randmenge und überdies, wie jede Begrenzung, abgeschlossen.

Eine beliebige Menge A heißt in E nirgendsdicht, wenn ihre abgeschlossene Hülle A_α in E nirgendsdicht ist. Dafür ist notwendig und hinreichend, daß der offene Kern B_i ihres Komplements $B = E - A$ in E dicht sei.

Die gewöhnlichen Kurven in der Euklidischen Ebene E_2, die Kurven und Flächen im Raum E_3 usw. sind nirgendsdichte Mengen. Genauer: ist $f(x_1, x_2)$ eine reelle stetige Funktion, die niemals in einer ganzen Umgebung eines Punktes verschwinden kann, z. B. ein Polynom mit nicht sämtlich verschwindenden Koeffizienten, so stellt $f = 0$ eine nirgendsdichte Menge in E_2 dar, denn sie ist abgeschlossen und hat keinen inneren Punkt.

Ein dyadisches Diskontinuum D (§ 26, 2) ist im vollständigen Raume E nirgendsdicht, wenn man noch die schärferen Ungleichungen

$$V_p > V_{p1} + V_{p2}, \qquad V_{pq} > V_{pq1} + V_{pq2}, \ldots$$

(mit Ausschluß der Gleichheitszeichen) hinzufügt. Denn in beliebiger Nähe jedes Punktes $x \varepsilon D$, welcher dem Durchschnitt $V_p V_{pq} \ldots$ angehören möge, liegen Punkte von

$$(V_p - V_{p1} - V_{p2}) + (V_{pq} - V_{pq1} - V_{pq2}) + \cdots$$

und diese gehören nicht zu D, x ist nicht innerer Punkt von D. Bei der Kugelkonstruktion S. 137 kann man offenbar diese Verschärfung herbeiführen; es gibt also in jedem vollständigen Raum, dessen insichdichter Kern nicht verschwindet, nirgendsdichte perfekte Mengen von der Mächtigkeit $\aleph$. Das einfachste Beispiel ist die triadische Menge Cantors (S. 134) nebst den durch Abänderung der Intervallängen daraus entstehenden Mengen.

Gehen wir zur Erklärung der im Raume E nirgends dichten Mengen A zurück. Setzen wir zur Abkürzung

$$E = A_\alpha + A_e,$$

wo also $A_e = E - A_\alpha$ die Menge der *äußeren Punkte* von A (der inneren des Komplements $E - A$) bedeutet, so sind die Aussagen

$$A \text{ in } E \text{ nirgendsdicht}, \quad A_e \text{ in } E \text{ dicht}$$

identisch. Wir wollen nun statt E irgendeine Teilmenge M des Raumes setzen, wobei man statt $A_\alpha(E) = A_\alpha$ natürlich $A_\alpha(M) = M A_\alpha$ zu setzen hat; dabei soll nicht notwendig $A \leqq M$, sondern nur A und M als Mengen im Raum E angenommen werden, so daß wir die allgemeinere Beziehung „*A zu M* nirgendsdicht" als Gegenstück der uns schon bekannten „A zu M dicht" erhalten. Kurzum, mit der Bezeichnung

$$(\alpha) \qquad P = M A_\alpha, \quad Q = M A_e, \quad M = P + Q$$

definieren wir die Aussagen

A zu M nirgendsdicht, Q in M dicht

als identisch; nach der früheren Erklärung ($M \leqq A_\alpha$) waren die Aussagen

A zu M dicht, $\quad Q = 0, \quad P = M$

identisch. (Das volle Gegenstück zu dieser wäre $Q = M$, $P = 0$, $M \leqq A_e$, M liegt ganz außerhalb von A; dies hat kein besonderes Interesse). Im Falle $A \leqq M$ wählen wir die Präposition „in": A in M nirgendsdicht, A in M dicht. Diese Relationen, wie allgemein die Mengen P und Q, hängen nur von den Mengen A und M, nicht vom umgebenden Raum E ab. P ist in M abgeschlossen, Q in M offen.

Ist A in M abgeschlossen, so ist $P = A$, und wir haben die Identität der Aussagen:

A in M nirgendsdicht, $M - A$ in M dicht;

A in M dicht, $A = M$.

Wir stellen einige formale Eigenschaften dieser Begriffe zusammen, deren Beweis, wenn nichts dazu gesagt wird, trivial ist.

I. (Transitives Gesetz). *Ist A zu B, B zu C dicht, so ist A zu C dicht.*

II. *Ist A zu M dicht (nirgendsdicht), so ist A^* zu M^* dicht (nirgendsdicht), wenn A^*, M^* Mengen derselben Dichtigkeitsklasse wie A, M sind* [1]).

In der Tat hängen P, Q nur von A_α, nicht von A ab. Was M betrifft, so wissen wir, daß, für eine offene Menge G, $(MG)_\alpha \geqq M_\alpha G \geqq MG$, also $(MG)_\alpha = (M_\alpha G)_\alpha$ ist; setzt man darin $G = A_e$, so folgt $Q_\alpha = Q^*_\alpha$, wenn $Q = M A_e$ und $Q^* = M^* A_e$ gemäß der Zerlegung (α) zu M und M^* gehören. Die Aussagen $Q = 0$ bzw. $Q_\alpha = M_\alpha$ sind also mit denselben für Q^*, M^* gleichbedeutend.

III. *Ist A zu M dicht, so ist jede Menge $\geqq A$ zu jeder Menge $\leqq M$ dicht. Ist A zu M nirgendsdicht, so ist jede Menge $\leqq A$ zu M nirgendsdicht; ist A in M nirgendsdicht, so ist A in jeder Menge $\geqq M$ nirgendsdicht.*

Hiervon bedarf nur die letzte Aussage eines Beweises; wir können die Menge $\geqq M$ dabei als Raum E wählen. Es ist

$$A_\alpha \leqq M_\alpha = Q_\alpha \leqq A_{e\alpha}$$

und zugleich $A_e \leqq A_{e\alpha}$, also $E \leqq A_{e\alpha}$, A_e in E dicht. Man darf in dieser letzten Behauptung nicht „zu" statt „in" setzen (allerdings genügt statt der Zusatzbedingung $A \leqq M$ die mildere $A_\alpha \leqq M_\alpha$); das geht schon daraus hervor, daß nach unserer Terminologie jede Menge A zur Nullmenge $M = 0$ nirgendsdicht ist, ohne doch zu jeder Menge $\geqq 0$ nirgendsdicht zu sein!

IV. *G sei offen. Ist A zu M dicht, so AG zu MG. Ist A zu M nirgendsdicht, so A zu MG.*

[1]) D. h. $A^*_\alpha = A_\alpha$, $M^*_\alpha = M_\alpha$ (S. 124).

Der erste Teil folgt aus $(AG)_\alpha \geqq A_\alpha G \geqq MG$. Der zweite stützt sich auf den ersten und (α): Q ist in M dicht, also $QG = MGA_e$ in MG dicht, A zu MG nirgendsdicht.

V. *Sind A_1, A_2 in M offen und dicht, so ist auch ihr Durchschnitt in M (offen und) dicht.*

Denn sei $A_1 = MG_1$, $A_2 = MG_2$ (G_1, G_2 offen). Es ist dicht: A_1 in M, nach IV $A_1 G_2$ in MG_2, d. h. $A_1 A_2$ in A_2, weiter A_2 in M, also (transitives Gesetz) $A_1 A_2$ in M.

VI. *Sind A_1, A_2 zu M nirgendsdicht, so ist auch ihre Summe zu M nirgendsdicht.*

Man mache für A_1, A_2 und $A = A_1 \dot{+} A_2$ die Zerlegungen (α): $M = P_1 + Q_1 = P_2 + Q_2 = P + Q$. Dann ist $P = P_1 \dot{+} P_2$, $Q = Q_1 Q_2$. Q_1 und Q_2 sind in M offen und dicht, also nach V auch Q.

Die letzten beiden Sätze gelten natürlich auch für endlich viele Mengen $A_1, A_2, \dots, A_n$. In einem *vollständigen* Raume lassen sie sich, in eingeschränkter Weise, auf abzählbar viele übertragen.

VII. *In einem vollständigen Raume sei eine Folge von Mengen A_n ($n = 1, 2, 3, \dots$) mit $A_{1\alpha} \leqq A_{2\alpha} \leqq \cdots$ und eine Folge offener Mengen $G_n \geqq A_n$ gegeben. Dann ist $Y = G_1 G_2 \dots$ zu A_1 dicht.*

Wir umgeben einen Punkt $a_1 \varepsilon A_1$ mit einer abgeschlossenen Kugel V_1 (d. h. a_1 sei ihr Mittelpunkt); in der zugehörigen offenen Kugel U_1 liegt, wegen $a_1 \varepsilon A_{2\alpha}$, ein Punkt $a_2 \varepsilon A_2$, den wir mit einer abgeschlossenen Kugel $V_2 \leqq U_1$ umgeben; in U_2 liegt ein Punkt $a_3 \varepsilon A_3$, den wir mit einer Kugel $V_3 \leqq U_2$ umgeben usf. Wir lassen dabei die Kugelradien nach 0 konvergieren und nehmen überdies $V_n \leqq G_n$, so daß der Punkt x von $V_1 V_2 \dots$ zu Y gehört. Da V_1 beliebig klein sein kann, ist $a_1 \varepsilon Y_\alpha$, also $A_1 \leqq Y_\alpha$.

VIII. *In einem vollständigen Raume sei eine Folge Youngscher Mengen $Y_1, Y_2, \dots$ derselben Dichtigkeitsklasse ($Y_{1\alpha} = Y_{2\alpha} = \cdots$) gegeben; dann gehört ihr Durchschnitt $Y = Y_1 Y_2 \dots$ noch zur selben Klasse.*

Sei $Y_m = G_{m1} G_{m2} \dots$ (G_{mn} offen); setzen wir $Y_{mn} = Y_m \leqq G_{mn}$ und wenden VII auf die Folge $Y_{11}, Y_{12}, Y_{21}, \dots$ an, so ist $\underset{mn}{\mathfrak{D}} G_{mn} = \underset{m}{\mathfrak{D}} Y_m = Y$ zu $Y_{11} = Y_1$, d. h. in Y_1 dicht: $Y_\alpha = Y_{1\alpha}$. Insbesondere:

IX. *Ist $A_1, A_2, \dots$ eine Folge von Mengen, die in der Youngschen Menge M dicht und offen sind, so ist auch ihr Durchschnitt in M dicht.*

Denn die Mengen A_n sind (als MG) selbst Youngsche Mengen und gehören mit M zu einer Dichtigkeitsklasse, also auch ihr Durchschnitt.

X. *Ist $A_1, A_2, \dots$ eine Folge von Mengen, die zu der Youngschen* [66]
Menge M nirgendsdicht sind, und A ihre Summe, so ist $M - MA$ in M dicht.

Machen wir für die A_n die Zerlegungen (α), $M = P_n + Q_n$, und setzen $P_0 = P_1 \dot{+} P_2 \dot{+} \cdots$, $Q_0 = Q_1 Q_2 \dots$, so sind die Q_n in M offen und dicht, also nach IX Q_0 in M dicht. Überdies ist $MA_n \leqq P_n$, $MA \leqq P_0$, $M - MA \geqq Q_0$, also ist $M - MA$ erst recht in M dicht.

Wir können aber nicht behaupten, daß A noch zu M nirgendsdicht sei; für die zu A gehörige Zerlegung $M = P + Q$ würde man ja nur $P \geqq P_0$, $Q \leqq Q_0$ schließen können, so daß Q nicht mehr in M dicht zu sein braucht. Bringt man die rationalen Zahlen in eine Folge $r_1, r_2, \ldots$ und läßt A_n aus der einzigen Zahl r_n bestehen, so ist A_n in der Menge M der reellen Zahlen nirgendsdicht, A aber sogar dicht; der Satz X bleibt trotzdem gültig, denn $M - A$ ist in M dicht. Versteht man aber unter M die Menge der rationalen Zahlen selbst, die keine Youngsche Menge ist (sie ist in ihrer vollständigen Hülle, d. h. in der Menge der reellen Zahlen kein G_δ, S. 138), so ist der Satz X nicht anwendbar: $M - A$ ist dann Null und nicht in M dicht.

Man nennt, mit einer farblosen Bezeichnung (R. Baire), eine Menge A
[67] *von erster Kategorie oder von zweiter Kategorie zu M*, je nachdem sie *Summe einer Folge zu M nirgendsdichter Mengen* ist *oder nicht.* Ist $A \leqq M$, so heißt sie von erster oder zweiter Kategorie *in M*; diesen wichtigsten Fall wollen wir abkürzend bezeichnen:

M_{I} eine Menge erster Kategorie in M,

M_{II} eine Menge zweiter Kategorie in M.

Der Satz X, auf $A \leqq M$ angewendet, lautet dann:

XI. *Ist M eine Youngsche Menge, so ist $M - M_{\mathrm{I}}$ in M dicht.*

Aus der Definition sowie aus II, III, IV folgt:

XII. *Sind die endlich oder abzählbar vielen Mengen A_n zu M von erster Kategorie, so auch ihre Summe.*

XIII. *Ist A zu M von erster Kategorie, so auch zu M^*, wenn M^* derselben Dichtigkeitsklasse wie M angehört.*

XIV. *Ist A zu M von erster Kategorie, so auch jede Menge $\leqq A$; ist A in M von erster Kategorie, so auch in jeder Menge $\geqq M$.*

XV. *Ist A zu M von erster Kategorie, so auch zu $M G$ (G offen).*

Diese Sätze lassen sich natürlich in solche über Mengen zweiter Kategorie umformen, z. B. statt XIV: ist A zu M von zweiter Kategorie, so auch jede Menge $\geqq A$ usw.

Es kann sein, daß M selbst ein M_{I} ist; dann gibt es überhaupt kein M_{II}, da nach XIV alle Teilmengen von M ebenfalls M_{I} sind. Ist aber M selbst ein M_{II}, so kann es nach XII nicht die Summe zweier M_{I} sein, und das Komplement $M - M_{\mathrm{I}}$ eines M_{I} ist jedenfalls ein M_{II} (wogegen das Komplement eines M_{II} sowohl ein M_{I} als ein M_{II} sein kann). *Eine Youngsche Menge $M > 0$ ist gewiß ein M_{II}.* Denn nach XI ist $M - M_{\mathrm{I}}$ in M dicht und kann nicht Null sein.

Auf Grund der Eigenschaft, daß von einer Menge $M = M_{\mathrm{II}}$ nach Tilgung eines M_{I} immer noch ein M_{II} zurückbleibt, kann man sich die Vorstellung bilden, daß die Teilmengen M_{I} sozusagen entbehrliche Akzidentien von dünner Struktur sind und ihre Komplemente $M - M_{\mathrm{I}}$ die eigentliche Substanz der Menge M enthalten. Diese Vorstellung ist zwar insofern zweckmäßig, als sie den nichtssagenden Bezeichnungen „erste und zweite Kategorie" etwas Farbe verleiht; sie kann aber mit ähnlichen Vorstellungen, die auf Grund anderer Prämissen gebildet sind, kollidieren. Bei der Zerlegung $M = M_s + M_k$ in separierten Teil und insichdichten Kern würde man doch diesen als Substanz und jenen als Akzidens betrachten, zumal wenn M_k von der Mächtigkeit des Kontinuums und M_s nur abzählbar ist; trotzdem kann M_k ein M_{I} und M_s ein M_{II} sein. Ist etwa J eine isolierte Menge mit perfekter Ableitung J_β (z. B. J die Menge der Punkte $\left(\frac{m}{n}, \frac{1}{n}\right)$ in Fig. 2 S. 113), so ist $M = J_\alpha = J + J_\beta$ eine solche Zerlegung, wo J_β ein M_{I} und J ein M_{II} ist (denn $M - J_\beta$ ist in M dicht; überdies stellt jeder einzelne Punkt von J bereits eine Menge M_{II} dar). — Ein einzelner Punkt von M gibt eine Menge M_{II} oder M_{I}, je nachdem er isoliert oder Häufungspunkt ist. Jede Menge M_{I} ist also in der Kohärenz M_h enthalten, und jede endliche oder abzählbare Teilmenge von M_h ist ein M_{I}. Z. B. ist in der Menge E der reellen Zahlen die Menge R der rationalen Zahlen ein E_{I}, sogar ein R_{I}; die Menge J der irrationalen ist (da E als Youngsche Menge ein E_{II} ist) ein E_{II} und erst recht ein J_{II}. — Ist $M = M_s$ separiert, so ist M_h ein M_{I}, da $M - M_h = M_j$ in $M_s = M$ dicht ist (§ 23, (15)). Bilden wir also die mit $M, M_h, M_{hh}, \ldots$ beginnende Folge der Kohärenzen $M_0, M_1, M_2, \ldots, M_\omega, \ldots$, so ist jede dieser Mengen in den vorangehenden von erster Kategorie, aber (solange sie > 0 ist) in sich von zweiter Kategorie, weil sie isolierte Punkte hat.

Ist die Menge E, die wir jetzt als Raum annehmen, eine Youngsche Menge, so ist auch jede (in E) abgeschlossene Menge F und jede offene Menge G (sogar jede Menge G_δ) eine Youngsche Menge, also, falls > 0, in sich von zweiter Kategorie. Auf Grund dieser zwei Eigenschaften wollen wir zwei Klassen von Mengen definieren, die als *Verallgemeinerung der Youngschen Mengen* vielfach eine Rolle spielen. Wir sagen:

E heißt eine F_{II}-Menge, wenn jede abgeschlossene Menge $F > 0$ in sich von zweiter Kategorie (ein F_{II}) ist [1]*). E heißt eine G_{II}-Menge, wenn jede offene Menge $G > 0$ in sich von zweiter Kategorie (ein G_{II}) ist.* [68]

Die Bedingung der F_{II}-Menge kann dahin eingeschränkt werden, daß nur jede *perfekte* Menge $P > 0$ ein P_{II} zu sein braucht; denn eine Menge F mit isolierten Punkten ist ohnehin ein F_{II}.

Die Bedingung der G_{II}-Menge kann durch diese ersetzt werden, daß jede offene Menge $G > 0$ ein E_{II} ist. Denn: ist G ein G_{I}, so auch ein E_{I} (nach XIV); ist G ein E_{I}, so auch ein G_{I} (nach XV). G ist also in sich und in E von derselben Kategorie (wogegen eine abgeschlossene Menge F zugleich ein F_{II} und E_{I} sein kann, z. B. jede im Euklidischen Raum E nirgends-

[1]) Hierzu vgl. Nachtrag A.

dichte abgeschlossene Menge > 0). Ferner lassen sich die G_{II}-Mengen auch so charakterisieren:

E ist dann und nur dann eine G_{II}-Menge, wenn $E - E_{\mathrm{I}}$ stets in E dicht ist.

Denn ist $E - E_{\mathrm{I}}$ stets in E dicht, so ist $G > 0$ kein E_{I}, da $E - G = F$ nicht in E dicht ist ($F_\alpha = F < E$). Ist umgekehrt E eine G_{II}-Menge und A ein E_{I}, $G > 0$, so ist in $G = (E - A)G + AG$ der zweite Summand ein E_{I}, also der erste ein E_{II} und jedenfalls > 0. Aus $G > 0$ folgt also $(E - A)G > 0$, d. h. $E - A$ ist in E dicht.

Mit E ist auch $E - E_{\mathrm{I}}$ eine G_{II}-Menge. Denn sei A in E, B in $E - A = D$ von erster Kategorie, so sind B und $A + B$ auch in E von erster Kategorie, also $E - (A + B) = D - B$ in E und erst recht in D dicht; D ist eine Menge, für die $D - D_{\mathrm{I}}$ in D dicht ist, also eine G_{II}-Menge.

Jede Youngsche Menge ist eine F_{II}-Menge; es gibt aber, wie wir später (§ 43, 2) sehen werden, auch F_{II}-Mengen, die keine Youngschen Mengen sind.

Jede F_{II}-Menge ist eine G_{II}-Menge. E sei eine F_{II}-Menge, $G > 0$ offen, $F = G_\alpha$, ferner $F - G = G_\alpha - G = H$ die Begrenzung von G. Sie ist in F nirgendsdicht (weil sie abgeschlossen und ihr Komplement $F - H = G$ in F dicht ist), also ein F_{I}; demnach ist G ein F_{II} und erst recht ein G_{II}.

Es gibt aber auch G_{II}-Mengen, die keine F_{II}-Mengen sind. Sei E die Euklidische Ebene, A die Menge der irrationalen Punkte $(i, 0)$ auf der x-Achse X; dann ist A ein E_{I} und $D = E - A$ wieder eine G_{II}-Menge. Sie ist aber keine F_{II}-Menge, da die in ihr abgeschlossene Menge $D\,X$, die Menge der rationalen Punkte $(r, 0)$ auf der x-Achse, in sich von erster Kategorie ist.

Eine Bemerkung verdienen noch die Mengen F_σ, Summen von Folgen abgeschlossener Mengen. Jede in E nirgendsdichte Menge A ist Teilmenge einer in E abgeschlossenen nirgendsdichten Menge A_α, also jede Menge E_{I} Teilmenge einer Menge $F_\sigma = E_{\mathrm{I}}$; die Mengen F_σ, die in E von erster Kategorie sind, sind also die größten Mengen erster Kategorie in dem Sinne, daß sie und ihre Teilmengen alle E_{I} liefern. Es gilt noch:

Damit $A = F_\sigma$ ein E_{I} sei, ist hinreichend und im Fall einer G_{II}-Menge E auch notwendig, daß $E - A$ in E dicht sei.

Denn ist $A = \mathfrak{S} F_n$ und $E - A$ in E dicht, so ist erst recht $E - F_n$ in E dicht, F_n in E nirgendsdicht und A ein E_{I}. Die andere Hälfte der Behauptung folgt aus der Definition der G_{II}-Mengen.

Ist der Raum E eine F_{II}-Menge, so ist jedes $A = G_\delta > 0$ in sich von zweiter Kategorie. Denn $F = A_\alpha$ ist ein F_{II}; $F - A$ ist ein F_σ. also in F von erster Kategorie, da sein Komplement A in F dicht ist; also ist A in F und erst recht in sich von zweiter Kategorie. *In einer*

F_{II}-Menge ist jedes G_δ, insbesondere jede abgeschlossene oder offene Menge eine F_{II}-Menge. Jede in einer G_{II}-Menge offene Menge ist eine G_{II}-Menge.

§ 28. Mengenräume. [69]

1. Entfernungen zwischen Mengen. Wir betrachten wieder Punkte und Teilmengen eines Raumes E. Die Entfernungen ab der Punktpaare zweier Mengen $A, B (> 0)$ haben stets eine *untere Grenze* $\delta(A, B) = \delta(B, A)$ und, falls beide Mengen beschränkt sind, eine (endliche) *obere Grenze* $d(A, B) = d(B, A)$; wir nennen dies die *untere* und *obere Entfernung* zwischen A und B. Insbesondere ist $d(A, A) = d(A)$ der *Durchmesser* von A. Ist der Punkt b fest, so ist $\delta(A, b) = \delta(b, A)$ die untere, $d(A, b) = d(b, A)$ die obere Entfernung des Punktes b von der Menge A.

Aus $xa \leqq ya + xy$ folgt, wenn wir beiderseits die obere Grenze für $a \varepsilon A$ nehmen (A beschränkt): $d(x, A) \leqq d(y, A) + xy$ und durch Vertauschung von x, y

$$| d(x, A) - d(y, A) | \leqq xy ;$$

auch für die unteren Entfernungen gilt dasselbe (S. 111). $\delta(x, A)$ und $d(x, A)$ sind stetige Funktionen von x. $\delta(x, A) = 0$ ist mit $x \varepsilon A_\alpha$ gleichbedeutend. Für $\varrho > 0$ ist $\delta(x, A) < \varrho$ damit gleichbedeutend, daß zu x ein $a \varepsilon A$ mit $ax < \varrho$ existiert oder daß x der Umgebung $U(A, \varrho)$ von A mit dem Radius ϱ angehört (S. 117).

Wir versuchen nun, zwischen zwei Mengen eine den Entfernungsaxiomen genügende *Entfernung* zu erklären und die Mengen dadurch zu Elementen eines neuen metrischen Raumes zu machen. Die unteren und oberen Entfernungen sind, wie man von vornherein vermuten kann und ein Versuch sofort bestätigen würde, dazu nicht geeignet.

Wir ziehen nur *beschränkte* Mengen $A, B > 0$ in Betracht. Die Relation $(\varrho > 0)$

$$(\varrho) \quad \left\{ \begin{array}{c} B \leqq U(A, \varrho) \\ \text{für jedes } b \varepsilon B \text{ ist } \delta(A, b) < \varrho \\ \text{zu jedem } b \varepsilon B \text{ gibt es ein } a \varepsilon A \text{ mit } ab < \varrho \end{array} \right.$$

deren drei Formen gleichbedeutend sind, besteht jedenfalls für $\varrho > d(A, B)$ (wo alle $ab < \varrho$) und besteht nicht für $\varrho \leqq \delta(A, B)$ (wo alle $ab \geqq \varrho$). Die untere Grenze der ϱ, für die sie besteht, ist eine Zahl $\geqq 0$ und werde mit $\varrho(A, B)$ bezeichnet, was von $\varrho(B, A)$ zu unterscheiden ist; nach der zweiten Form der Relation (ϱ) ist

$$(1) \qquad \varrho(A, B) = \sup_{b \varepsilon B} \delta(A, b),$$

ferner

$$(2) \qquad \delta(A, B) \leqq \varrho(A, B) \leqq d(A, B).$$

Für $\varrho > \varrho(A, B)$ besteht die Relation (ϱ), für $\varrho < \varrho(A, B)$ besteht sie nicht (für $\varrho = \varrho(A, B)$ besteht sie dann und nur dann, wenn das Supremum in (1) für kein b erreicht wird). $\varrho(A, B) = 0$ heißt so viel wie $B \leqq A_\alpha$, A zu B dicht; in jedem andern Falle ist $\varrho(A, B) > 0$. Ferner gilt

$$(3) \qquad \varrho(A, B) + \varrho(B, C) \geqq \varrho(A, C).$$

Denn für $B \leqq U(A, \varrho)$, $C \leqq U(B, \sigma)$ ist $C \leqq U(A, \varrho + \sigma)$, woraus der Schluß leicht zu ziehen ist. Besteht eine der Mengen aus einem einzigen Punkt, so ist nach (1)

$$(4) \qquad \varrho(a, B) = \sup_{b \varepsilon B} ab = d(a, B), \quad \varrho(A, b) = \delta(A, b).$$

Dieses $\varrho(A, B)$ ist wegen seiner Asymmetrie noch nicht als Entfernung brauchbar; wir bilden daher

$$(5) \qquad \overline{AB} = \max\,[\varrho(A, B),\ \varrho(B, A)] = \overline{BA}$$

und schließen aus (3), daß damit das Dreiecksaxiom $\overline{AB} + \overline{BC} \geqq \overline{AC}$ erfüllt ist. Damit ist also eine *Entfernung* und ein metrischer Mengenraum definiert, in dem man nur, analog zu früheren Fällen, zwei Mengen mit $\overline{AB} = 0$, d. h. zwei Mengen *derselben Dichtigkeitsklasse* als identisch anzusehen hat; für $\overline{AB} = 0$ ist auch $\varrho(A, C) = \varrho(B, C)$ und $\varrho(C, A) = \varrho(C, B)$ sowie $\overline{AC} = \overline{BC}$.

Nach (2) liegt die Entfernung $\overline{AB}$ zwischen der unteren und oberen Entfernung; nach (4) ist, wenn die eine Menge einpunktig ist,

$$\overline{Ab} = \max\,[\delta(A, b),\ d(A, b)] = d(A, b).$$

Wir wollen nachher der präziseren Ausdrucksweise halber von jeder Dichtigkeitsklasse nur die größte (abgeschlossene) Menge beibehalten, verengern also den Definitionsbereich von $\varrho(A, B)$ und $\overline{AB}$ auf die *abgeschlossenen beschränkten Mengen* > 0 des Raumes E. In diesem metrischen Mengenraum ist die Konvergenz $A = \lim A_n$ oder $A_n \to A$ natürlich durch $\overline{AA_n} \to 0$ zu erklären; die Folge heiße dann *metrisch konvergent*, A ihr *metrischer Limes*; der Limes einer metrisch konvergenten Teilfolge von A_n ein metrisches Häufungselement.

[70] **2. Abgeschlossener und offener Limes.** Außerdem kommen, zunächst für beliebige Mengenfolgen A_n des metrischen Raumes E, folgende Grenzmengen in Betracht, die denen für reine Mengen (oberer und unterer Limes, § 3) analog sind: der *obere* und *untere abgeschlossene Limes*

$$(6) \qquad \overline{F} = \overline{Fl}\, A_n, \quad \underline{F} = \underline{Fl}\, A_n,$$

im Gleichheitsfalle der *abgeschlossene Limes*

$$(7) \qquad F = Fl\, A_n;$$

sowie der *obere* und *untere offene Limes*

(8) $$\overline{G} = \overline{Gl}\, A_n, \quad \underline{G} = \underline{Gl}\, A_n,$$
im Gleichheitsfalle der *offene Limes*

(9) $$G = Gl\, A_n$$
(die Abkürzungen Fl, Gl mögen etwa F-Limes, G-Limes gelesen werden). Die Punkte dieser Mengen sollen durch folgende Eigenschaften definiert sein:

$x \varepsilon \overline{F}$: jede Umgebung U_x hat mit unendlich vielen A_n Punkte gemein.
$x \varepsilon \underline{F}$: jede Umgebung U_x hat mit fast allen A_n Punkte gemein.
$x \varepsilon \overline{G}$: es gibt eine Umgebung U_x, die unendlich vielen A_n angehört.
$x \varepsilon \underline{G}$: es gibt eine Umgebung U_x, die fast allen A_n angehört.

Es ist $\overline{F} \geqq \underline{F}$, $\overline{G} \geqq \underline{G}$. Sind $B_n = E - A_n$ die Komplemente der A_n, so ist, wie eine leichte Überlegung lehrt,

(10) $$E = \overline{Fl}\, A_n + \underline{Gl}\, B_n = \underline{Fl}\, A_n + \overline{Gl}\, B_n.$$

Daß $\overline{G}$, $\underline{G}$ offen sind, folgt unmittelbar aus der Definition: es ist $\overline{G}$ bzw. $\underline{G}$ die Summe aller Umgebungen, die unendlich vielen bzw. fast allen A_n angehören; fast ebenso unmittelbar oder aus (10) folgt, daß $\overline{F}$, $\underline{F}$ abgeschlossen sind. Abänderung, Hinzufügung, Weglassung endlich vieler Mengen ist auf die Grenzmengen ohne Einfluß; beim Übergang zu einer Teilfolge hat man, wenn sich die griechischen Buchstaben auf diese beziehen,

(11) $$\underline{F} \leqq \underline{\Phi} \leqq \overline{\Phi} \leqq \overline{F}, \quad \underline{G} \leqq \underline{\Gamma} \leqq \overline{\Gamma} \leqq \overline{G},$$
so daß mit der ganzen Folge auch jede Teilfolge einen abgeschlossenen (offenen) Limes hat.

Beispiele. Für die Folge $A, B, A, B, \ldots$ ist
$$\overline{F} = A_\alpha \dot{+} B_\alpha, \quad \underline{F} = A_\alpha B_\alpha,$$
$$\overline{G} = A_i \dot{+} B_i, \quad \underline{G} = A_i B_i.$$

Werden die Folgen A_n mit den Grenzmengen $\overline{F}, \underline{F}, \overline{G}, \underline{G}$ und B_n mit den Grenzmengen $\overline{\Phi}, \underline{\Phi}, \overline{\Gamma}, \underline{\Gamma}$ zu einer Folge $A_1, B_1, A_2, B_2, \ldots$ vereinigt, so hat diese die Grenzmengen
$$\overline{F} \dot{+} \overline{\Phi}, \quad \underline{F}\underline{\Phi}, \quad \overline{G} \dot{+} \overline{\Gamma}, \quad \underline{G}\underline{\Gamma}.$$

Ist in der Euklidischen Ebene, für $t > 0$, $A(t)$ die durch $x_1^2 + \left(\frac{x_2}{t}\right)^2 = 1$ definierte Ellipse mit den Scheiteln $(\pm 1, 0)$ und $(0, \pm t)$, so ist der abgeschlossene Limes für die Folge $A(n)$ mit $n = 1, 2, 3, \ldots$ das Geradenpaar $x_1 = \pm 1$, für die Folge $A\left(\frac{1}{n}\right)$ die Strecke $-1 \leqq x_1 \leqq 1$, $x_2 = 0$.

I (Auswahlsatz). *Ist der Raum separabel, so enthält jede Mengenfolge eine Teilfolge, für die der abgeschlossene (und eine, für die der offene) Limes existiert.*

10*

Wegen (10) genügt es, den Fall des offenen Limes zu beweisen (übrigens liefert zweimalige Anwendung des Satzes eine Teilfolge, für die sowohl der abgeschlossene als auch der offene Limes existiert). Statt der Umgebungen U benutzen wir wieder die abzählbar vielen speziellen Umgebungen V (§ 25); die in unendlich vielen A_n enthaltenen V (deren Summe $\overline{G}$ ist), mögen *obere* V heißen, die in fast allen A_n enthaltenen (deren Summe $\underline{G}$ ist) *untere* V, und diejenigen oberen, die keine unteren sind, *diskrepante* V. Beim Übergang zu einer Teilfolge wird die Menge der oberen, unteren, diskrepanten V kleiner, größer, kleiner (im weiteren Sinn, d. h. Gleichbleiben nicht ausgeschlossen); wählt man die Teilfolge aber geeignet, so kann man ein beliebiges diskrepantes V wirklich tilgen. Denn wenn V den unendlich vielen A_p, aber nicht fast allen A_n angehört, so genügt der Übergang von A_n zur Teilfolge A_p, um V zu einem unteren V zu machen und demnach aus der Reihe der diskrepanten V zu streichen. Danach verfahre man, falls $\overline{G} > \underline{G}$, also diskrepante V vorhanden sind, folgendermaßen: man bringe die (höchstens abzählbar vielen) diskrepanten V in Form einer Folge $V_1, V_2, \ldots$ und tilge V_1 durch Übergang von A_n zu einer Teilfolge A_p: falls für diese noch diskrepante V vorhanden sind, etwa $V_m, \ldots$, so tilge man das niedrigste V_m durch Übergang von A_p zu einer Teilfolge A_q usf. Entweder gelangt man schon nach endlich vielen Schritten zu einer Teilfolge, für die keine diskrepanten V mehr vorhanden sind, also der offene Limes existiert, oder das Verfahren ist unbegrenzt fortsetzbar. In diesem Fall bilde man aus den (wachsend geordneten) Zahlenfolgen $p, q, r, \ldots$ die Diagonalfolge t, d. h. aus

$$\begin{aligned} p &= p_1, p_2, p_3, \ldots \\ q &= q_1, q_2, q_3, \ldots \\ r &= r_1, r_2, r_3, \ldots \\ &\ldots \end{aligned}$$

die Folge
$$t = p_1, q_2, r_3, \ldots.$$

Diese liefert, da sie, von endlich vielen Gliedern abgesehen, Teilfolge jeder vorangehenden ist, eine Mengenfolge A_t, für die kein diskrepantes V mehr vorhanden ist, also der offene Limes existiert.

3. Beziehung zwischen abgeschlossenem und metrischem Limes. Lassen wir jetzt die offenen Grenzmengen beiseite und fragen, ob zwischen dem abgeschlossenen und dem metrischen Limes, $Fl\,A_n$ und $\lim A_n$, oder allgemeiner zwischen den beiden abgeschlossenen Grenzmengen und den metrischen Häufungselementen Beziehungen bestehen. Bemerken wir, daß $\overline{F}$ und $\underline{F}$ unverändert bleiben, wenn wir die A_n durch ihre abgeschlossenen Hüllen ersetzen, denn $A\,U_x > 0$ ist gleichbedeutend mit $A_\alpha U_x > 0$. Wir nehmen daher künftig die A_n als abgeschlossen an, überdies als beschränkt und >0. Bedeutet a_n einen Punkt von A_n, so kann man auf Grund einer leichten Überlegung auch sagen:

$x \varepsilon \overline{F}$ bedeutet: es gibt eine Folge a_n mit dem Häufungspunkt x (d. h. mit einer nach x konvergenten Teilfolge).

$x \varepsilon \underline{F}$ bedeutet: es gibt eine Folge $a_n \to x$.

Wir erhalten dann folgenden Satz, wo auch X, Y abgeschlossene beschränkte Mengen >0 bedeuten:

II. *Wenn* $\varrho(A_n, X) \to 0$, *so ist* $X \leqq \underline{F}$; *wenn* $\varrho(Y, A_n) \to 0$, *so ist* $Y \geqq \overline{F}$.

Beweis. Sei $\varrho_n = \varrho(A_n, X) \to 0$, $x \varepsilon X$; es gibt (weil für $\varrho > \varrho(A, B)$ die Relation (ϱ) besteht) einen Punkt $a_n \varepsilon A_n$ mit $x a_n < \varrho_n + \frac{1}{n}$, also $a_n \to x$, $x \varepsilon \underline{F}$.

Sei $\varrho_n = \varrho(Y, A_n) \to 0$, $x \varepsilon \overline{F}$; es gibt eine Folge $a_n \varepsilon A_n$ mit einer konvergenten Teilfolge $a_p \to x$ und zu jedem a_n einen Punkt $y_n \varepsilon Y$ mit $a_n y_n < \varrho_n + \frac{1}{n}$. Also $y_p \to x$, x ist α-Punkt von Y, $x \varepsilon Y_\alpha = Y$.

Damit ist II bewiesen. Aus $X \leqq \underline{F} \leqq \overline{F} \leqq Y$ folgt nun, daß für $X = Y$ diese Menge auch mit $\underline{F} = \overline{F}$ identisch ist, d. h.

III. *Wenn der metrische Limes* $X = \lim A_n$ *existiert, so ist* $X = F l A_n$ *zugleich der abgeschlossene Limes der Mengenfolge.*

Daß eine Umkehrung dieser Sätze nicht ohne weiteres möglich ist zeigt schon ein triviales Beispiel: sei im Raume E der reellen Zahlen A_n das Intervall $[-n, n]$, so existiert der abgeschlossene Limes dieser Folge, nämlich E selbst, aber E ist nicht beschränkt und $\overline{A_n E}$ hat keinen Sinn. Diese Folge hat kein Häufungselement, da die Entfernungen $\overline{A_m A_n} = |n - m|$ von zwei verschiedenen Mengen $\geqq 1$ sind. — Ein anderes Beispiel erhält man mit

$$A_{2n} = \left\{0, \frac{1}{n}\right\}, \quad A_{2n-1} = \{0, n\} \quad (n = 1, 2, \ldots),$$

wo also jedes A_n aus zwei Zahlen besteht. Hier existiert der abgeschlossene Limes F, aus der einen Zahl 0 bestehend, er ist zugleich das (einzige) metrische Häufungselement $\lim A_{2n}$, die ganze Folge ist aber nicht metrisch konvergent.

Allgemein kann man nur sagen: wenn X metrisches Häufungselement von A_n ist, d. h. eine Teilfolge mit $\overline{A_p X} \to 0$ existiert, so ist nach III und (11) $\underline{F} \leqq X \leqq \overline{F}$; alle etwaigen Häufungselemente der Folge liegen zwischen dem unteren und oberen abgeschlossenen Limes. Ist insbesondere $F = F l A_n$ vorhanden, so hat die Folge kein von F verschiedenes Häufungselement; sie braucht aber, wie die Beispiele gezeigt haben, überhaupt keins zu haben, oder sie kann das einzige Häufungselement F haben, ohne metrisch zu konvergieren.

Wenn aber der Raum E *in sich kompakt* ist[1]), lassen sich unsere Sätze umkehren. Es gilt zunächst als Gegenstück zu II:

IV. *Wenn E in sich kompakt ist, so ist $\varrho(A_n, \underline{F}) \to 0$ (falls $\underline{F}$ von Null verschieden ist) und $\varrho(\overline{F}, A_n) \to 0$.*

Beweis. Für jeden Punkt $x \varepsilon \underline{F}$, der ja als $\lim a_n$ dargestellt werden kann, ist $\delta(A_n, x) \leqq a_n x \to 0$. Wäre nun, für $\underline{F} > 0$, nicht $\varrho(A_n, \underline{F}) \to 0$, also für eine geeignete Teilfolge $\varrho(A_p, \underline{F}) > \varrho > 0$, so gäbe es nach der Definition (1) einen Punkt $x_p \varepsilon \underline{F}$ mit $\delta(A_p, x_p) > \varrho$. Die x_p haben, wegen der Kompaktheit, eine konvergente Teilfolge $x_q \to x$, wobei x der abgeschlossenen Menge $\underline{F}$ angehört und $\delta(A_q, x) \to 0$. Aus $|\delta(A_q, x_q) - \delta(A_q, x)| \leqq x x_q$ folgt demnach $\delta(A_q, x_q) \to 0$ im Widerspruch zu $\delta(A_p, x_p) > \varrho$.

Zum Beweis des zweiten Teils stellen wir zunächst $\overline{F} > 0$ fest, da jede Folge $a_n \varepsilon A_n$ mindestens einen Häufungspunkt $x \varepsilon \overline{F}$ hat. Wäre nicht $\varrho(\overline{F}, A_n) \to 0$, also für eine gewisse Teilfolge $\varrho(\overline{F}, A_p) > \varrho > 0$, so gäbe es einen Punkt $a_p \varepsilon A_p$ mit $\delta(\overline{F}, a_p) > \varrho$, was damit in Widerspruch steht, daß die a_p einen Häufungspunkt $x \varepsilon \overline{F}$ haben müssen.

Aus IV folgt als Umkehrung von III:

[71] V. *Wenn E in sich kompakt ist und der abgeschlossene Limes $X = Fl\, A_n$ existiert, so ist $X = \lim A_n$ zugleich der metrische Limes der Mengenfolge.*

Da ein kompakter Raum höchstens separabel ist (S. 126), so liefert der Auswahlsatz I in Verbindung mit V:

VI. *Ist der Raum E in sich kompakt, so hat jede Mengenfolge $A_n > 0$ eine metrisch konvergente Teilfolge. D. h. der metrische Raum der abgeschlossenen Mengen $F (0 < F \leqq E)$ ist wieder in sich kompakt.*

§ 29. Zusammenhang.

1. Grundlagen. Eine Zerlegung

$$(1) \qquad A = A_1 + A_2 = A F_1 + A F_2$$

der Menge A in zwei disjunkte, *in A abgeschlossene* Summanden heißt eine *Zerstückelung* von A, falls beide Summanden > 0 sind. Eine Menge, die sich zerstückeln läßt, heißt unzusammenhängend; *eine Menge > 0, die sich nicht zerstückeln läßt, heißt zusammenhängend.* Die Summanden A_1, A_2 sind übrigens gleichzeitig *in A offen.* Insbesondere ist also eine abgeschlossene Menge > 0 zusammenhängend, wenn sie sich nicht in zwei disjunkte abgeschlossene Mengen > 0 spalten läßt, und eine offene Menge > 0 ist zusammenhängend, wenn sie sich nicht in zwei disjunkte offene Mengen > 0 spalten läßt. *Eine abgeschlossene zusammenhängende Menge heißt ein Kon-*

---

[1]) Es genügt, daß die sämtlichen A_n einer in E kompakten Menge angehören oder daß $S = \mathfrak{S} A_n$ in E kompakt ist; man kann dann S_α als Raum annehmen.

tinuum, eine offene zusammenhängende Menge ein Gebiet. Ob eine Menge A zusammenhängend ist oder nicht, hängt nur von ihr selbst, nicht vom umgebenden Raume E ab (während die Begriffe Kontinuum und Gebiet Relativbegriffe bezüglich E sind).

Beispiele. Eine Hyperbel ist unzusammenhängend, die Spaltung in ihre beiden Zweige ist eine Zerstückelung [1]. Die Menge der positiven und negativen reellen Zahlen, ohne Null, ist unzusammenhängend. Eine einpunktige Menge gilt als zusammenhängend (Kontinuum). Ein Intervall $[a, b]$ reeller Zahlen ist zusammenhängend; denn sei eine Spaltung $A_1 + A_2$ in disjunkte Mengen gegeben, $a \varepsilon A_1$ und c die untere Grenze der Zahlen von A_2. Ist nun A_2 abgeschlossen, so ist c die kleinste Zahl von A_2, also $c > a$, und A_1 ist nicht abgeschlossen, da das halboffene Intervall $[a, c)$ zu A_1 gehört, c selbst aber nicht.

Um eine Menge $A\,(>0)$ als zusammenhängend nachzuweisen, ist zu zeigen, daß bei jeder Spaltung (1) einer der Summanden Null ist. Man beachte hierbei, daß diese Spaltung auch für jede Teilmenge B von A eine entsprechende Spaltung

$$B = B_1 + B_2 = BA_1 + BA_2 = BF_1 + BF_2$$

hervorruft.

I. *Wenn je zwei Punkte von A „sich in A verbinden lassen", d. h. einer zusammenhängenden Teilmenge von A angehören, so ist A zusammenhängend.*

Denn gäbe es eine Zerstückelung (1), $x_1 \varepsilon A_1$, $x_2 \varepsilon A_2$, so verbinde man x_1, x_2 durch eine zusammenhängende Teilmenge B; dann wäre aber auch $B = B_1 + B_2$ eine Zerstückelung.

II. *Die Summe von zwei zusammenhängenden Mengen, deren Durchschnitt nicht verschwindet, ist zusammenhängend.*

Sei $S = A \dotplus B$, $D = AB > 0$. Einer Spaltung $S = S_1 + S_2$ entsprechen $A = A_1 + A_2$, $B = B_1 + B_2$, $D = D_1 + D_2$. Weil A zusammenhängt, ist einer der Summanden 0, etwa $A_2 = 0$, also $D_2 = 0$, $D_1 > 0$, $B_1 > 0$; weil B zusammenhängt, ist also $B_2 = 0$, $S_2 = 0$: S ist zusammenhängend.

Eine Art Umkehrung von II ist der Satz:

III. *Zwei abgeschlossene Mengen, deren Summe und Durchschnitt zusammenhängend ist, sind selbst zusammenhängend.*

Sei wieder $S = A \dotplus B$, $D = AB$. Einer Spaltung $A = A_1 + A_2$ entspricht $D = D_1 + D_2$, also ist etwa $D_2 = 0$, $A_2 B = 0$. Dann ist $S = (A_1 \dotplus B) + A_2$; damit dies keine Zerstückelung sei, muß $A_2 = 0$ sein: A ist zusammenhängend. — Der Satz gilt bereits, wenn A und B in S (das man als Raum wählen kann) abgeschlossen sind.

[1]) Wir sind in der Euklidischen Ebene E_2, ohne unendlich ferne oder uneigentliche Punkte!

Sagen wir von endlich vielen, in bestimmter Reihenfolge geschriebenen Mengen $A, B, C, \ldots$, daß sie eine *Kette* bilden, falls je zwei benachbarte Mengen gemeinsame Punkte haben ($AB > 0$, $BC > 0, \ldots$). Sind diese Mengen zusammenhängend, so ist nach II auch $A \dot{+} B$, $(A \dot{+} B) \dot{+} C, \ldots$ zusammenhängend. Nach I folgt:

IV. *Die Summe beliebig vieler zusammenhängender Mengen, von denen je zwei einer Kette angehören, ist zusammenhängend.*

Insbesondere gilt dies, wenn je zwei Mengen gemeinsame Punkte haben.

Eine *größte* (in keiner andern enthaltene) *zusammenhängende* Teilmenge von A heißt eine *Komponente* von A. Eine solche erhält man, wenn man die Menge $A(x)$ aller Punkte, die sich in A mit einem Punkt $x \varepsilon A$ verbinden lassen, oder die (nach IV zusammenhängende) Summe aller den Punkt x enthaltenden zusammenhängenden Teilmengen von A bildet. $A(x)$ kann sich auf die einpunktige Menge $\{x\}$ reduzieren. Zwei Komponenten $A(x)$, $A(y)$ sind entweder disjunkt oder identisch. Man erhält so die Spaltung von A in Komponenten

$$A = A(x) + A(y) + \cdots, \tag{2}$$

deren es eine (falls A zusammenhängt) oder mehrere (endlich oder unendlich viele) geben kann. Eine Menge, deren sämtliche Komponenten einpunktig sind, die also *keine mehrpunktige zusammenhängende Teilmenge* enthält, heiße *punkthaft* (total unzusammenhängend), während eine Menge, die *kein mehrpunktiges Kontinuum* enthält, *diskontinuierlich* genannt werde [1]). Die erste Bedingung ist also schärfer, und eine Menge kann diskontinuierlich sein, ohne punkthaft zu sein, sogar diskontinuierlich und zusammenhängend (Beispiele § 31, 2 und § 42, 5). Punkthaft ist ein absoluter, diskontinuierlich ein relativer Begriff. A ist in E diskontinuierlich heißt: A hat keine mehrpunktige, zusammenhängende, in E abgeschlossene Teilmenge. Eine in sich diskontinuierliche Menge ist auch punkthaft (wegen VI).

V. *Mit A ist auch jede Menge zwischen A und A_α ($A \leqq B \leqq A_\alpha$) zusammenhängend.*

Einer Spaltung $B = B_1 + B_2$ in relativ abgeschlossene Mengen entspricht $A = A_1 + A_2$. Hier ist etwa $A_2 = 0$, $A_1 = A$, $A \leqq B_1 \leqq A_\alpha$, $B_{1\alpha} = A_\alpha$, und da B_1 in B abgeschlossen ist: $B_1 = BB_{1\alpha} = BA_\alpha = B$, $B_2 = 0$.

VI. *Die Komponenten von A sind in A abgeschlossen.*

Für eine Komponente P ist AP_α (zwischen P und P_α gelegen) nach V zusammenhängend, also, da P eine größte zusammenhängende Menge $\leqq A$ ist, $AP_\alpha = P$.

---

[1]) Auch hier herrscht kein einheitlicher Sprachgebrauch.

VII. *Eine zusammenhängende Menge C, die zwei Punkte a, b von Komplementärmengen $A, B\,(A + B = E)$ verbindet, trifft die Begrenzung dieser Mengen. Eine mehrpunktige zusammenhängende Menge ist insichdicht und mindestens von der Mächtigkeit des Kontinuums.*

Denn wäre $CA_g = C\,(A_r + B_r) = 0$, so hätte man eine Zerstückelung $C = CA_i + CB_i$. Wählt man speziell A als Menge der Punkte $ax \leqq \varrho$, B als Menge der Punkte $ax > \varrho$, wo $0 < \varrho < ab$, so folgt, daß C für jedes dieser ϱ mindestens einen Punkt mit $ax = \varrho$ enthalten muß, und damit die zweite Hälfte des Satzes.

Insbesondere ist ein mehrpunktiges Kontinuum perfekt. Eine endliche oder abzählbare Menge ist punkthaft.

2. Lineare und lokal zusammenhängende Räume. In den linearen Räumen (§ 20, 2), wo die Entfernung $xy = |x - y|$ auf einem Betrage $|x|$ beruhte, der eine positiv homogene Funktion ersten Grades der Koordinaten ist, sind die Strecken $[x, y]$ mit reellen Zahlenintervallen $[a, b]$ isometrisch, also zusammenhängend. Nach I sind also die *konvexen* Mengen (die mit zwei Punkten x, y auch ihre Verbindungsstrecke $[x, y]$ enthalten) zusammenhängend, so der ganze Raum, eine offene Kugel (Umgebung), eine abgeschlossene Kugel u. dgl. Weiter ist ein *Streckenzug*

$$(3)\quad [x_0, x_1, x_2, \ldots, x_{n-1}, x_n] = [x_0, x_1] \dotplus [x_1, x_2] \dotplus \cdots \dotplus [x_{n-1}, x_n]$$

als Spezialfall einer Kette (S. 152) zusammenhängend und damit jede Menge, in der sich zwei Punkte durch einen (der Menge angehörigen) Streckenzug verbinden lassen; so ist in der Euklidischen Ebene (oder im E_m für $m \geqq 2$) die Menge J der irrationalen Punkte zusammenhängend, denn sind x, y zwei Punkte von J und z ein dritter Punkt, der auf keiner der (abzählbar vielen) Geraden liegt, die x oder y mit einem rationalen Punkte verbinden, so sind x, y in J durch den Streckenzug $[x, z, y]$ verbunden. — Für eine *offene* Menge im linearen Raum oder in einem konvexen Teilraum ist diese Zusammenhangsbedingung (Verbindung durch einen Streckenzug) nicht nur hinreichend, sondern auch notwendig. Denn ist G offen und $G(x)$ die Menge der Punkte y, die sich mit $x \varepsilon G$ durch einen Streckenzug $[x, \ldots, y]$ in G verbinden lassen, so ist $G(x)$ offen, weil für $U_y \leqq G$ jeder Punkt $z \varepsilon U_y$ mit y durch $[y, z] \leqq U_y$, also mit x durch $[x, \ldots, y, z]$ verbunden ist. Zwei Mengen $G(x)$, $G(y)$ sind entweder disjunkt oder identisch. Demnach hat man wie bei der Zerlegung in Komponenten

$$(4)\qquad G = G(x) + G(y) + G(z) + \cdots.$$

Ist nur ein Summand vorhanden, so ist $G = G(x)$ zusammenhängend; sind zwei oder mehrere vorhanden, so liefert

$$G = G(x) + [G(y) + G(z) + \cdots]$$

eine Zerstückelung. Bildet man hier den Durchschnitt mit einer Teilmenge von G, so wird auch diese zerstückelt, wenn sie mit mindestens zwei Summanden, etwa $G(x)$ und $G(y)$, gemeinsame Punkte hat; eine zusammenhängende Teilmenge von G liegt also ganz in einem einzigen Summanden, d. h. die $G(x)$ sind die Komponenten von G. Wir haben also gezeigt:

VIII. *Ist der Raum konvex, so sind die Komponenten einer offenen Menge selbst offene Mengen; eine offene Menge ist dann und nur dann zusammenhängend (ein Gebiet), wenn sich je zwei Punkte durch einen Streckenzug in ihr verbinden lassen.*

In den separablen Räumen dieser Art, z. B. im Euklidischen, ist die Menge der Komponenten einer offenen Menge höchstens abzählbar.

Ist E die gerade Linie oder die Menge der reellen Zahlen, so enthält nach VII eine zusammenhängende Menge nebst zwei Punkten a, b auch das Intervall $[a, b]$, woraus leicht folgt, daß es hier keine andern zusammenhängenden Mengen gibt als einzelne Punkte, Intervalle, Halbgerade und die ganze Gerade selbst (die Intervalle können offen, halboffen oder abgeschlossen sein, die Halbgeraden offen oder abgeschlossen). Gebiete sind die offenen Intervalle (a, b), die offenen Halbgeraden (a, ∞) und $(-\infty, b)$ sowie die ganze Gerade. Die allgemeinste offene Menge G entsteht durch Addition von endlich oder abzählbar vielen disjunkten Gebieten. Wenn dabei aneinandergrenzende Intervalle wie (a, b) und (b, c) vorkommen, so wird b ein isolierter Punkt des abgeschlossenen Komplements $F = E - G$; trifft dies niemals zu, so wird F perfekt.

Die in VIII zuerst genannte Eigenschaft kommt einer allgemeineren Klasse von Räumen zu, wie wir nun zeigen wollen.

Nennen wir — zunächst in einem beliebigen Raume — eine Spaltung

$$A = P + Q + R + \cdots \tag{5}$$

der Menge A in disjunkte Summanden > 0 eine *natürliche* Spaltung, wenn sie die zusammenhängenden Teilmengen von A nicht zerreißt, die Summanden also Komponentensummen oder, wenn zusammenhängend, Komponenten sind. Natürliche Spaltungen sind also u. a. die Spaltung in Komponenten, ferner in endlich viele relativ abgeschlossene oder in beliebig viele relativ offene Mengen (P und $A - P$ sind dann gleichzeitig in A abgeschlossen oder offen, woraus die Behauptung leicht folgt). Wenn sich z. B. eine offene Menge G in offene Mengen spalten läßt, so sind diese, insoweit sie zusammenhängend sind, Komponenten von G. Ein Gebiet G mit der Begrenzung H ist Komponente von $E - H$, denn $G_\alpha = G + H$ ist abgeschlossen, $E - (G + H) = G_1$ offen, $E - H = G + G_1$ eine natürliche Spaltung. Ein Gebiet ist durch seine Begrenzung und einen seiner Punkte eindeutig bestimmt.

Der Raum E heiße *lokal zusammenhängend*[1]), *wenn in jeder Umgebung U_x eines beliebigen Punktes x ein den Punkt x enthaltendes Gebiet G_x enthalten ist.* [72] Er braucht darum (als Ganzes) nicht zusammenhängend zu sein [2]). Ein konvexer Raum ist lokal zusammenhängend, da hier sogar jede Umgebung U_x selbst zusammenhängend ist. Das einfachste Beispiel eines nicht lokal zusammenhängenden Raumes bildet die Menge $E = \{1, \frac{1}{2}, \frac{1}{3}, \ldots, 0\}$; die einzige, den Punkt 0 enthaltende zusammenhängende Menge ist die einpunktige Menge $\{0\}$ selbst und diese ist in E nicht offen. Ein Beispiel für einen zusammenhängenden, aber nicht lokal zusammenhängenden Raum folgt nach Satz XII. Nun gilt:

IX. *Ist der Raum lokal zusammenhängend und $A = P + Q + R + \cdots$ eine natürliche Spaltung von A, so ist*

$$(6) \qquad A_i = P_i + Q_i + R_i + \cdots, \qquad A_r = P_r + Q_r + R_r + \cdots,$$

$$(7) \qquad A_g = S_\alpha, \qquad S = P_g \dot{+} Q_g \dot{+} R_g \dot{+} \cdots.$$

In dieser einfachen Weise hängen also offener Kern, Rand und Begrenzung der ganzen Menge von denen der Summanden ab. Zum Beweise von (6) erinnern wir uns, daß allgemein

$$A_i \geqq P_i + Q_i + \cdots, \qquad A_r \leqq P_r + Q_r + \cdots$$

gilt. Andrerseits sei $x \varepsilon A_i$, $U_x \leqq A$ und G_x ein Gebiet mit $x \varepsilon G_x \leqq U_x$, dann ist G_x in einem einzigen Summanden der Spaltung enthalten, etwa $G_x \leqq P$, $x \varepsilon P_i$, also $A_i \leqq P_i + Q_i + \cdots$.

Zum Beweise von (7) betrachten wir das Komplement $B = E - A$, dann ist

$$\begin{aligned} E - P &= B + Q + R + \cdots \\ (E - P)_r &\leqq B_r + Q_r + R_r + \cdots \\ P_g = P_r + (E - P)_r &\leqq B_r + P_r + Q_r + R_r + \cdots = B_r + A_r = A_g, \end{aligned}$$

also $A_g \geqq P_g \dot{+} Q_g \dot{+} \cdots = S$ und $A_g \geqq S_\alpha$, weil A_g abgeschlossen ist. Andererseits ist

$$(8) \qquad A_g \leqq S_\alpha$$

und diese Hälfte der Behauptung gilt sogar (in einem lokal zusammenhängenden Raum) für jede Zerlegung $A = P \dot{+} Q \dot{+} R \dot{+} \cdots$ (die keine natürliche Spaltung zu sein und nicht einmal disjunkte Summanden zu haben braucht). Denn sei $x \varepsilon A_g$, U_x beliebig, $G_x \leqq U_x$; da x α-Punkt von A und von B ist, enthält die Menge G_x Punkte beider Mengen; sie enthalte etwa einen Punkt von P und einen von $B \leqq E - P$, also (nach VII) auch einen von P_g oder S, folglich $x \varepsilon S_\alpha$.

---

[1]) oder *im kleinen zusammenhängend* (H. Hahn).

[2]) Ein isolierter Raum ist lokal zusammenhängend; jede einpunktige Menge $\{x\}$ ist ein G_x.

Bei einer natürlichen Spaltung in endlich viele Summanden vereinfacht sich (7) zu $A_g = S = P_g \dot{+} Q_g \dot{+} \cdots$.

X. *Ist der Raum lokal zusammenhängend, so sind die Komponenten offener Mengen stets Gebiete, und umgekehrt.*

Denn ist A offen, so sind nach (6) auch $P, Q, R, \ldots$ offen, was also insbesondere von den Komponenten gilt. Wenn umgekehrt die Komponenten offener Menge Gebiete sind, so ist die x enthaltende Komponente von U_x ein Gebiet G_x, der Raum lokal zusammenhängend.

Eine Anwendung von (8) ist der Satz:

Ein zusammenhängender und lokal zusammenhängender Raum, der eine F_{II}-Menge ist, läßt sich nicht in abzählbar viele disjunkte abgeschlossene Mengen > 0 spalten.

Angenommen, es sei $E = F_1 + F_2 + \cdots$ ($F_n > 0$ abgeschlossen); weil E zusammenhängend ist, hat F_n eine Begrenzung $H_n = F_{nr} > 0$; die Summe $H = H_1 + H_2 + \cdots$ ist ebenfalls abgeschlossen (als Komplement von ΣF_{ni}). Für die Begrenzung von

$$E - F_1 = F_2 + F_3 + \cdots$$

ergibt sich nach (8)

$$H_1 \leqq (H_2 + H_3 + \cdots)_\alpha,$$

d. h. $H_2 + H_3 + \cdots$ ist in H dicht, H_1 und jedes H_n in H nirgends dicht, also H in sich von erster Kategorie im Widerspruch zur Annahme, daß E eine F_{II}-Menge sein soll. — Diese Verschärfung des Zusammenhangs, die eine Zerstückelung nicht nur in endlich, sondern auch in abzählbar viele abgeschlossene Summanden ausschließt, gilt z. B. für Euklidische Räume; wir werden sie nachher auf anderem Wege auch für kompakte Kontinua beweisen (S. 162).

Wir haben bisher den lokalen Zusammenhang schlechthin, für alle Punkte des Raumes, definiert. Es läge nahe, E in dem einzelnen Punkt x lokal zusammenhängend zu nennen, wenn jedes U_x ein G_x (Gebiet, dem x angehört) enthält; indessen ist es zweckmäßig, diese Bedingung etwas zu lockern und so zu definieren:

Der Raum E heißt *im Punkte x lokal zusammenhängend, wenn jede Umgebung U_x eine zusammenhängende Menge C enthält, der x als innerer Punkt angehört* ($x \varepsilon C_i$); man kann dabei unter C die x enthaltende Komponente von U_x verstehen.

XI. *Wenn der Raum E in jedem Punkt x lokal zusammenhängend ist, so ist er* (schlechthin) *lokal zusammenhängend; und umgekehrt.*

Denn ist C_x die x enthaltende Komponente von U_x, $y \varepsilon C_x$, $U_y \leqq U_x$ und C_y die y enthaltende Komponente von U_y, so ist $C_y \leqq C_x$, weil $C_x \dot{+} C_y$ zusammenhängend und $< U_x$, also mit C_x identisch ist. Nun ist y innerer Punkt von C_y, also auch von C_x, C_x besteht nur aus inneren

Punkten und ist ein x enthaltendes Gebiet $\leqq U_x$. Die Umkehrung ist trivial.

Man kann den lokalen Zusammenhang in x noch etwas anders formulieren (St. Mazurkiewicz). Es sei $\widehat{xy}$ die untere Grenze der Durchmesser $d(C)$ aller *beschränkten zusammenhängenden Mengen* C, die x und y enthalten; wenn es keine solche gibt, werde $\widehat{xy}$ nicht definiert. Es ist $\widehat{xy} \geqq xy$; wenn ferner $\widehat{xy}$ und $\widehat{yz}$ existieren, so auch $\widehat{xz} \leqq \widehat{xy} + \widehat{yz}$, auf Grund der leicht ersichtlichen Tatsache, daß $d(A \dot{+} B) \leqq d(A) + d(B)$, falls $AB > 0$. Dieses $\widehat{xy}$ hat also, soweit es existiert, die Eigenschaften einer Entfernung. Wir können dann sagen:

XII. *Der Raum ist in x lokal zusammenhängend dann und nur dann, wenn mit xy auch $\widehat{xy}$ nach Null konvergiert.*

Die Bedingung soll natürlich besagen: für beliebiges $\varrho > 0$ und geeignetes $\sigma > 0$ ist $\widehat{xy}$ definiert und $< \varrho$, sobald $xy < \sigma$.

Ist diese Bedingung erfüllt, so seien U_x, V_x die Umgebungen von x mit den Radien ϱ, σ. Jedes $y \varepsilon V_x$ läßt sich mit x durch eine zusammenhängende Menge C_y mit $d(C_y) < \varrho$ verbinden. Die Menge $C = \overset{V_x}{\underset{y}{\mathfrak{S}}} C_y$ ist nach IV zusammenhängend und in U_x enthalten, da jeder ihrer Punkte von x eine Entfernung $< \varrho$ hat; sie enthält V_x und hat demnach x zum inneren Punkt; E ist in x lokal zusammenhängend.

Ist umgekehrt E in x lokal zusammenhängend, so habe U_x den beliebig vorgeschriebenen Radius ϱ; die x enthaltende Komponente C von U_x hat x zum inneren Punkt, enthält also eine gewisse Umgebung V_x vom Radius σ. Für $xy < \sigma$ ist dann $\widehat{xy} \leqq d(C) \leqq 2\varrho$.

Beispiel. In der Euklidischen $\xi\eta$-Ebene sei r_n der Punkt $\xi = n^{-1}$, $\eta = (-1)^n$ für $n = 1, 2, 3, \ldots$. Die Summe der Strecken $[r_1, r_2]$, $[r_2, r_3], \ldots$ ist der unendliche Streckenzug

$$R = [r_1, r_2, r_3, \ldots],$$

der in beliebiger Nähe der Ordinatenachse unendlich viele Hin- und Rückwege zwischen den Geraden $\eta = \pm 1$ macht, so daß jeder Punkt der Strecke

$$S: \ \xi = 0, \ -1 \leqq \eta \leqq 1$$

Häufungspunkt von R ist. Die Menge $E = R + S$ ist, als Raum gedacht, in den Punkten von S nicht lokal zusammenhängend. Denn will man einen Punkt $s \varepsilon S$ mit einem noch so nahen Punkte $r \varepsilon R$ durch eine zusammenhängende Menge $C \leqq E$ verbinden, so muß C

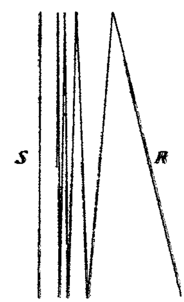

Fig. 6.

von jeder zwischen s und r verlaufenden Geraden $\xi = \text{const.}$ in mindestens einem Punkte getroffen werden; dann enthält aber C einen Rest $[r_n, r_{n+1}, \ldots]$ des Streckenzuges und hat einen Durchmesser > 2, so daß $\widehat{sr}$ mit sr nicht nach 0 konvergiert. Übrigens ist aber R und E zusammenhängend.

3. Distanzen. Ein endlicher Punktkomplex $(x_1, x_2, \ldots, x_n)$, worin die Entfernungen $x_1x_2, \ldots, x_{n-1}x_n$ benachbarter Punkte $\leqq \varrho$ sind, heiße eine *ϱ-Kette* $(\varrho > 0)$; wir sagen dann, x_1 und x_n lassen sich durch eine ϱ-Kette verbinden, insbesondere durch eine ϱ-Kette in der Menge A, wenn alle Punkte der Kette zu A gehören. Suchen wir für zwei Punkte x, y von A alle sie in A verbindenden ϱ-Ketten; die untere Grenze der dabei auftretenden Zahlen ϱ heiße die *Distanz* der Punkte x, y und werde mit $\overline{xy}$ bezeichnet (sie hängt aber nicht nur von den beiden Punkten, sondern auch von A ab). Für $\varrho > \overline{xy}$ lassen sich also x, y durch eine ϱ-Kette in A verbinden, für $\varrho < \overline{xy}$ nicht; für $\varrho = \overline{xy}$ hängt es vom Einzelfall ab. Offenbar ist $\overline{xy} \leqq xy$. Die Distanzen erfüllen das Dreiecksaxiom in der verschärften Fassung

$$\overline{xz} \leqq \max\,[\overline{xy}, \overline{yz}]; \tag{9}$$

denn wenn sich x und y durch eine ϱ-Kette, y und z durch eine σ-Kette verbinden lassen, so x und z durch eine τ-Kette mit $\tau = \max\,[\varrho, \sigma]$. Wenn $x \to \xi, y \to \eta$, so gilt $\overline{x\xi} \to 0$, $\overline{y\eta} \to 0$, $\overline{xy} \to \overline{\xi\eta}$; die Distanz $\overline{xy}$ ist *stetige Funktion* von x, y. Natürlich ist $\overline{xx} = 0$ zu setzen; aber $\overline{xy} = 0$ besagt nicht etwa, daß $x = y$, sondern nur, daß sich x, y für *jedes* $\varrho > 0$ durch eine ϱ-Kette verbinden lassen. In der Menge der reellen oder der rationalen Zahlen haben je zwei Punkte die Distanz 0. Nun gilt (betreffs der unteren Entfernung $\delta(P, Q)$ vgl. S. 145):

XIII. *Ist $A = P + Q$ mit $\delta(P, Q) = \delta > 0$, so hat jeder Punkt $p \varepsilon P$ von jedem Punkt $q \varepsilon Q$ eine Distanz $\overline{pq} \geqq \delta$. Haben zwei Punkte p, q von A eine Distanz $\overline{pq} = \varrho > 0$, so ist eine Zerlegung $A = P + Q$ mit $p \varepsilon P$, $q \varepsilon Q$ und $\delta(P, Q) \geqq \varrho$ möglich.*

Das erste ist evident. Das zweite beweist sich so: P sei die Menge der Punkte $x \varepsilon A$ mit $\overline{px} < \varrho$ und $Q = A - P$. Es ist $p \varepsilon P$, $q \varepsilon Q$. Wäre nun $\delta(P, Q) < \varrho$, so gäbe es zwei Punkte $x \varepsilon P$, $y \varepsilon Q$ mit $xy < \varrho$; dann wäre erst recht $\overline{xy} < \varrho$ und, nach (9), $\overline{py} \leqq \max\,[\overline{px}, \overline{xy}] < \varrho$; gegen die Voraussetzung wäre also $y \varepsilon P$.

Da die Zerlegung $A = P + Q$ des Satzes XIII offenbar eine Zerstückelung ist (P und Q in A abgeschlossen), so folgt: *in einer zusammenhängenden Menge haben je zwei Punkte die Distanz Null.* Das ist aber nicht umkehrbar [1]), wie die Menge der rationalen Zahlen zeigt, die nicht zusammenhängend ist, obwohl zwei ihrer Punkte die Distanz Null haben. Die Zerstückelung von A in zwei Teile mit positiver unterer Entfernung

[1]) Nach einer Definition Cantors, von der wir abgewichen sind, heißt eine Menge dann schon zusammenhängend, wenn je zwei ihrer Punkte die Distanz Null haben.

ist eben nur ein sehr grober, sozusagen mit bloßem Auge sichtbarer Mangel an Zusammenhang, während die Menge auch in feinerer, mikroskopischer Weise unzusammenhängend sein, nämlich eine Zerstückelung $A = P + Q$ mit $\delta(P, Q) = 0$ gestatten kann (wie die Menge der rationalen Zahlen in die der Zahlen $< \sqrt{2}$ und $> \sqrt{2}$).

In einem wichtigen Falle ist indessen die obige Behauptung umkehrbar. Wenn die *in sich kompakte* (also kompakte, abgeschlossene) Menge $A = P + Q$ zerstückelt, d. h. in zwei abgeschlossene Mengen > 0 gespalten wird, so ist deren untere Entfernung $\delta(P, Q) = \delta > 0$. In der Tat wird hier die untere Grenze δ der Entfernungen pq wirklich erreicht, das Infimum zum Minimum; denn ist $p_n q_n \to \delta$, so gibt es eine Teilfolge mit $p_n \to p$, von dieser eine Teilfolge mit $q_n \to q$, und dann ist $p \varepsilon P$, $q \varepsilon Q$, $\delta = pq > 0$. Hier ist also jede Zerstückelung „mit bloßem Auge sichtbar", und wenn je zwei Punkte in A die Distanz 0 haben, so ist A zusammenhängend. Ferner ist, wenn A nicht zusammenhängend ist, die den Punkt p enthaltende Komponente P mit der Menge $P(0)$ der Punkte $\overline{px} = 0$ identisch. Denn einerseits ist immer $P \leqq P(0)$; andererseits ist $P(0)$ abgeschlossen, weil die Distanz $\overline{px}$ eine stetige Funktion von x ist, also in sich kompakt und daher zusammenhängend, da je zwei ihrer Punkte die Distanz 0 haben[1]), demnach $P(0) \leqq P$. Wir haben also gezeigt:

XIV. *Eine in sich kompakte Menge A ist dann (und nur dann) zusammenhängend, wenn je zwei ihrer Punkte die Distanz 0 haben. Die den Punkt p enthaltende Komponente ist die Menge der Punkte, die von p die Distanz 0 haben.*

Beim Beweise von XIII zeigten wir, daß die Menge der Punkte x mit $\overline{px} < \varrho$ ebenso wie ihr Komplement abgeschlossen ist; dasselbe gilt auch von der Menge $P(\varrho)$ der Punkte $\overline{px} \leqq \varrho$ und ihrem Komplement $Q(\varrho) = A - P(\varrho)$ für $\varrho > 0$; man sieht wie dort, daß die untere Entfernung der beiden Mengen $> \varrho$ ist ($Q(\varrho) > 0$ vorausgesetzt). Für $\varrho = 0$ gilt dies nicht mehr; wohl ist noch die p enthaltende Komponente

$$(10) \qquad P(0) = \mathfrak{D}_{\varrho > 0} P(\varrho) = P(1)\, P(\tfrac{1}{2})\, P(\tfrac{1}{3}) \ldots$$

abgeschlossen, aber $Q(0)$ braucht es nicht zu sein und die untere Entfernung beider kann 0 sein.

Die folgenden drei Sätze sind Anwendungen der Distanzen auf kompakte Mengen.

XV. *Eine in sich kompakte Menge bleibt in sich kompakt, wenn man statt der Entfernungen die Distanzen zugrunde legt, wobei sie in eine punkthafte Menge übergeht.*

Die Menge sei, in Komponenten zerlegt,

$$A = P + Q + R + \cdots;$$

[1]) Hierzu vgl. Nachtrag B.

[73] macht man sie auf Grund der Distanzen in A zum metrischen Raum $\bar{A}$, so fallen alle Punkte einer Komponente von A in einen einzigen Punkt zusammen; man kann, indem man aus jeder Komponente je einen Punkt wählt,

$$\bar{A} = \{p, q, r, \ldots\}$$

setzen[1]). Eine Punktfolge x_n hat in A einen Häufungspunkt x, d. h. eine Teilfolge mit $xx_\nu \to 0$, und dann ist erst recht $\overline{xx_\nu} \to 0$, x_n hat auch in $\bar{A}$ den Häufungspunkt x. $\bar{A}$ ist also in sich kompakt, und ebenso entspricht jeder in sich kompakten Teilmenge B von A eine in sich kompakte Teilmenge $\bar{B}$ von $\bar{A}$. Sind p, q Punkte verschiedener Komponenten von A. ϱ positiv und $< \overline{pq}$, so ist die Menge $P(\varrho)$ der Punkte $\overline{px} \leqq \varrho$ wie ihr Komplement $Q(\varrho) = A - P(\varrho)$ in A abgeschlossen; die ihnen entsprechenden, von denen die eine den Punkt p und die andere den Punkt q enthält, sind es in $\bar{A}$, und zwei verschiedene Punkte von $\bar{A}$ gehören also niemals einer zusammenhängenden Teilmenge von $\bar{A}$ an: $\bar{A}$ ist punkthaft (S. 152).

Da die Menge $\bar{A}$ vollständig und höchstens separabel ist, so folgt: *die Menge der Komponenten einer in sich kompakten Menge A ist entweder höchstens abzählbar oder von der Mächtigkeit des Kontinuums.*

XVI. *Die kompakten perfekten punkthaften Mengen sind mit den dyadischen Diskontinuen identisch.*

Sei A zunächst in sich kompakt. Wir bilden mit $\varrho > 0$ eine größte Menge von Punkten, die paarweise in A Distanzen $> \varrho$ (also erst recht Entfernungen $> \varrho$) haben; diese Menge ist endlich und bestehe aus $x_1, x_2, \ldots, x_n$. Zu jedem Punkt $x \varepsilon A$ gibt es dann mindestens ein x_i $(i = 1, 2, \ldots, n)$ mit $\overline{xx_i} \leqq \varrho$ (sonst wäre jene größte Menge noch erweiterungsfähig) und auch nur ein einziges (sonst wäre $\overline{x_i x_k} \leqq \varrho$). Ist also A_i die Menge der Punkte $\overline{xx_i} \leqq \varrho$, so ist

$$(11) \qquad A = A_1 + A_2 + \cdots + A_n$$

in disjunkte, abgeschlossene kompakte Mengen zerlegt, deren jede einen „Distanzdurchmesser“ $\leqq \varrho$ hat, d. h. zwei Punkte von A_i haben eine Distanz $\leqq \varrho$.

Ist A zweitens zugleich punkthaft, so konvergiert mit der Distanz $\overline{xy}$ auch die Entfernung xy $(\geqq \overline{xy})$ nach Null, und zwar gleichmäßig, d. h. jedem $\sigma > 0$ entspricht ein $\varrho > 0$ derart, daß mit $\overline{xy} < \varrho$ auch $xy < \sigma$. Denn andernfalls gäbe es eine Folge von Punktpaaren mit $\overline{x_n y_n} \to 0$, $x_n y_n \geqq \sigma$, wobei wir mit Beschränkung auf Teilfolgen $x_n \to x$, $y_n \to y$ annehmen können, also $\overline{xy} = 0$, $xy \geqq \sigma$; die verschiedenen Punkte x, y würden derselben Komponente angehören. A läßt sich also für jedes $\delta > 0$ in endlich viele disjunkte, in sich kompakte Mengen von Durchmessern $\leqq \delta$ spalten.

---

[1]) $\bar{A}$ ist übrigens stetiges Bild von A und XV eine Folge von § 35, III.

Ist drittens A auch noch perfekt, so sind die Summanden einer solchen Spaltung (11) wieder perfekt, da sie in A auch offen, also insichdicht sind. Die A_i sind also Mengen wie A und das Verfahren kann wiederholt werden:

$$A = \sum_i A_i, \quad A_i = \sum_k A_{ik}, \quad A_{ik} = \sum_l A_{ikl}, \ldots$$

wobei etwa die Durchmesser der Mengen mit n Indizes kleiner als $\frac{1}{n}$ gewählt werden können, natürlich auch so klein, daß jede Summe aus mindestens zwei Summanden besteht. Dann ist

$$A = \Sigma A_i \,.\, \Sigma A_{ik} \,.\, \Sigma A_{ikl} \ldots$$

ein polyadisches Diskontinuum (S. 134) und läßt sich in ein dyadisches verwandeln.

Umgekehrt ist ein dyadisches Diskontinuum

$$A = \Sigma V_p \,.\, \Sigma V_{pq} \,.\, \Sigma V_{pqr} \ldots$$

punkthaft. Denn ist $C \leqq A$ zusammenhängend, so kann, damit $C = CV_1 + CV_2$ keine Zerstückelung sei, C nur mit einem V_p, dann nur mit einem V_{pq}, mit einem V_{pqr} usw. gemeinsame Punkte haben, d. h. $C = V_p V_{pq} V_{pqr} \ldots$ ist einpunktig. Damit ist XVI bewiesen.

XVII (Randsatz von Z. Janiszewski). *F sei abgeschlossen, $G = E - F$ ihr offenes Komplement, $H = F_r$ die Begrenzung beider Mengen, C ein kompaktes Kontinuum mit $CH > 0$* [1]. *Dann hat (α) jede Komponente von CF Punkte mit H gemein, (β) für $CG > 0$ jede Komponente von CG Häufungspunkte in H, (γ) für $CF_i > 0$ jede Komponente von CF_i Häufungspunkte in H.*

Die Figur veranschaulicht den Fall, daß F eine abgeschlossene Kreisfläche in der Euklidischen Ebene E, G das Äußere des Kreises, H die Kreisperipherie ist.

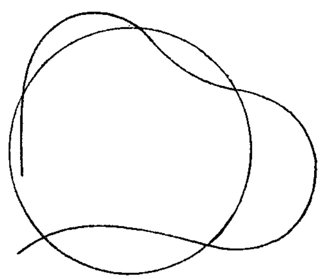

Fig. 7.

Beweis. (α) Ist $CF = P + Q$ in zwei abgeschlossene Mengen gespalten und $P > 0$, so ist auch $PH > 0$. Denn wegen $C = CF \dotplus CG_\alpha$ ist $C = P \dotplus (Q \dotplus CG_\alpha)$; damit letzteres keine Zerstückelung sei, muß der Durchschnitt beider Summanden $PCG_\alpha = PFG_\alpha = PH$ von Null verschieden sein. Sei nun $p \varepsilon CF$ und $P(\varrho)$ die Menge der Punkte x, die in der abgeschlossenen kompakten Menge CF eine Distanz $\overline{px} \leq \varrho$ haben, so ist für $\varrho > 0$, wie wir wissen, $P(\varrho)$ und das Komplement $Q(\varrho) = CF - P(\varrho)$ abgeschlossen, also $P(\varrho)H > 0$. Nach dem ersten Durchschnittssatz § 26, I ist der Durchschnitt der Mengen

[1]) Diese Voraussetzung ist wegen VII insbesondere erfüllt, wenn $CF > 0$, $CG > 0$.

$P(\varrho)H$ für $\varrho = 1, \frac{1}{2}, \frac{1}{3}, \ldots$ auch noch > 0, nach (10) also $P(0)H > 0$, wobei $P(0)$ eine beliebige Komponente von CF bedeuten kann.

(β) Sei F_n die abgeschlossene Menge der Punkte $\delta(x, F) \geqq \frac{1}{n}$ ($n = 1, 2, 3, \ldots$); es ist $G = F_1 \dot{+} F_2 \dot{+} \cdots$ und für $p \varepsilon CG$ *schließlich* $p \varepsilon CF_n$. Nach dem unter (α) Bewiesenen hat die p enthaltende Komponente von CF_n (C trifft sowohl F_n als $E - F_n \geqq F$, nach VII also auch den Rand von F_n), um so mehr die von CG, einen Punkt x_n auf dem Rande von F_n, d. h. mit $\delta(x_n, F) = \frac{1}{n}$, und diese Punkte haben in C einen Häufungspunkt x mit $\delta(x, F) = 0$, $x \varepsilon F G_\alpha = H$.

(γ) C trifft sowohl F_i als $E - F_i \geqq H$, nach VII also auch die Begrenzung von F_i; nach (β), auf F_i statt G angewandt, hat jede Komponente von CF_i Häufungspunkte in dieser Begrenzung, die gleich $F_{i\alpha} - F_i \leqq F - F_i = H$ ist.

Anwendungen. *Ist C ein kompaktes Kontinuum und A ein Teilkontinuum $< C$, so gibt es ein Kontinuum B mit $A < B < C$.* Man wähle nämlich einen Punkt $y \varepsilon C - A$, eine positive Zahl $\varrho < \delta(y, A)$ und setze für F die Menge der Punkte $\delta(x, A) \leqq \varrho$, für B die A enthaltende Komponente von CF. C trifft sowohl F als $E - F$ (im Punkte y), also den Rand von F; demnach enthält B einen Randpunkt von F mit $\delta(x, A) = \varrho$, also $B > A$; andererseits ist $y \bar{\varepsilon} B$, also $B < C$.

Ein kompaktes Kontinuum läßt sich nicht in abzählbar viele disjunkte abgeschlossene Mengen > 0 spalten[1] (W. Sierpiński).

Angenommen, das kompakte Kontinuum A sei als Summe

$$A = A_1 + A_2 + \cdots$$

disjunkter abgeschlossener Mengen $A_n > 0$ darstellbar. Wir zeigen, daß A ein Kontinuum B enthält, das mit unendlich vielen A_n, jedoch nicht mit A_1, Punkte gemein hat:

$$B = BA_{p_1} + BA_{p_2} + \cdots = B_{p_1} + B_{p_2} + \cdots$$

($1 < p_1 < p_2 < \cdots$, $B_p > 0$). Damit ist schon alles bewiesen; denn die Wiederholung des Verfahrens gibt ein Kontinuum

$$C = CB_{q_1} + CB_{q_2} + \cdots = C_{q_1} + C_{q_2} + \cdots$$

mit $C_q > 0$, wo die q eine Teilfolge der p bilden und $p_1 < q_1 < q_2 < \cdots$ usw.; hierbei würde also $ABC \cdots = 0$ sein im Widerspruch zum ersten Durchschnittssatz. — Um die auf B bezügliche Behauptung zu beweisen, verstehen wir unter $2\varrho = \delta(A_1, A_2) > 0$ die untere Entfernung von A_1, A_2 und (etwa im Raume A selbst) unter F die abgeschlossene Menge der Punkte $\delta(x, A_1) \geqq \varrho$, sodaß $A_2 \leqq F \leqq A - A_1$. Sei dann B die irgend-

[1]) Vgl. den ähnlichen Satz S. 156.

einen Punkt von A_2 enthaltende Komponente von F, so daß $BA_1 = 0$, $BA_2 > 0$. Nun enthält aber nach XVII B einen Randpunkt von F, d. h. mit $\delta(x, A_1) = \varrho$, und dieser gehört nicht zu A_2; B hat also mit mindestens zwei A_n, als Kontinuum folglich mit unendlich vielen A_n Punkte gemein, wie behauptet.

4. Folgen zusammenhängender Mengen.

XVIII. *Eine absteigende Folge $A_1 \geqq A_2 \geqq \cdots$ kompakter Kontinua hat als Durchschnitt ein Kontinuum.* [74]

Bemerken wir voraus: zwei disjunkte abgeschlossene Mengen F_1, F_2 lassen sich stets in zwei disjunkte offene Mengen G_1, G_2 einschließen; man braucht nur jeden Punkt $x_1 \varepsilon F_1$ mit einer Umgebung vom Radius $\frac{1}{2}\delta(x_1, F_2)$ und ebenso jeden Punkt $x_2 \varepsilon F_2$ mit einer Umgebung vom Radius $\frac{1}{2}\delta(x_2, F_1)$ zu versehen. — Wäre nun $A = A_1 A_2 \ldots$ (eine kompakte abgeschlossene Menge > 0) in zwei abgeschlossene Mengen zerstückelbar, so schließen wir diese in disjunkte offene Mengen G_1, G_2 ein, so daß $A = AG_1 + AG_2$; mit $F = E - (G_1 + G_2)$ ist $AF = 0$. Nach dem ersten Durchschnittssatz müßte dann aber bereits, für ein geeignetes n, $A_n F = 0$, also $A_n = A_n G_1 + A_n G_2$ zerstückelbar sein.

XIX (Satz von L. Zoretti). *Ist der Raum in sich kompakt, so ist der obere abgeschlossene Limes einer Folge zusammenhängender Mengen wieder zusammenhängend, falls der untere abgeschlossene Limes nicht Null ist.*

Die Mengen A_n seien zusammenhängend, $\underline{F} = \underline{Fl}\, A_n > 0$, es soll $\overline{F} = \overline{Fl}\, A_n$ als zusammenhängend erwiesen werden. Es sei $x \varepsilon \underline{F}$, $y \varepsilon \overline{F}$; danach (S. 149) gibt es eine Folge $a_n \to x$ mit $a_n \varepsilon A_n$, eine Teilfolge $b_p \to y$ mit $b_p \varepsilon A_p$, und endlich ist (§ 28, IV) $\varrho(\overline{F}, A_n) \to 0$, woraus zusammen folgt: zu beliebigem $\varrho > 0$ gibt es eine Menge $A_p = A$ und in ihr zwei Punkte a, b derart, daß

$$ax < \varrho, \; by < \varrho, \; \varrho(\overline{F}, A) < \varrho.$$

Die letzte Relation besagt, daß es zu jedem $c \varepsilon A$ ein $z \varepsilon \overline{F}$ mit $cz < \varrho$ gibt. Nun lassen sich a und b in A durch eine ϱ-Kette $(c_0, c_1, \ldots, c_n)$ mit $c_0 = a$, $c_n = b$ verbinden und zu jedem c_i läßt sich ein $z_i \varepsilon \overline{F}$ mit $c_i z_i < \varrho$ angeben, wobei insbesondere $z_0 = x$ und $z_n = y$ gewählt werden kann. Dann ist $(z_0, z_1, \ldots, z_n)$ eine (3ϱ)-Kette, die x und y in $\overline{F}$ verbindet. Da dies für jedes ϱ möglich ist, haben x, y in $\overline{F}$ die Distanz 0; da dies für jedes $y \varepsilon \overline{F}$ gilt, so haben je zwei Punkte von $\overline{F}$ die Distanz 0; da $\overline{F}$ abgeschlossen, also in sich kompakt ist, ist $\overline{F}$ zusammenhängend.

Das Beispiel der Folge $A, B, A, B, \ldots$ mit $\overline{F} = A_\alpha + B_\alpha$, $\underline{F} = A_\alpha B_\alpha$ zeigt, daß die Voraussetzung $\underline{F} > 0$ nicht entbehrlich ist und daß $\underline{F}$ nicht zusammenhängend zu sein braucht. Auch die Kompaktheit des Raumes ist wesentlich: ist in der Euklidischen Ebene $A_n = [a, c_n, b]$ der die Punkte

11*

$a = (-1, 0)$, $c_n = (0, n)$, $b_n = (1, 0)$ verbindende Streckenzug, so besteht $\overline{F} = \underline{F}$ aus den beiden Halbgeraden $x = \pm 1$, $y \geqq 0$ und ist nicht zusammenhängend.

Ist der Raum in sich kompakt und hat die Folge der zusammenhängenden Mengen A_n den abgeschlossenen Limes A oder, was nach § 28, III V dasselbe ist, den metrischen Limes A, so ist A ein Kontinuum. Hiervon ist XVIII ein Sonderfall. Ist der Raum in sich kompakt, so hat nach § 28, VI jede Folge zusammenhängender Mengen eine metrisch konvergente Teilfolge, deren Limes also wieder ein Kontinuum ist. (Man erinnere sich, daß der metrische Limes abgeschlossen sein sollte.) Die Kontinua eines in sich kompakten Raumes bilden also wieder einen in sich kompakten Raum.

Siebentes Kapitel.

Punktmengen und Ordnungszahlen.

§ 30. Hüllen und Kerne.

Die Theorie der Wohlordnung, ursprünglich von Cantor gerade für Zwecke der Punktmengenlehre ausgebildet, ist später von dieser Mission etwas zurückgedrängt worden (nicht immer aus billigenswerten Motiven), und wir haben selbst im vorigen Kapitel gesehen, was sich ohne sie erreichen läßt. Indessen gibt es doch Fälle, wo die Ordnungszahlen zurzeit unentbehrlich, und andere, wo sie zur feineren Ausgestaltung eines ohne sie erzielbaren Ergebnisses mindestens sehr willkommen sind. In der einen oder andern Rolle treten sie besonders in Erscheinung, wenn es sich um Bildung von Mengen oder Mengensystemen handelt, die irgendwie die *kleinsten* oder *größten* ihrer Art sind.

1. Schema für Hüllen und Kerne. In einem Raume E (der zunächst eine reine Menge sein kann) sei jeder Menge A eindeutig eine Menge A_φ zugeordnet (beide also $\leqq E$; φ ist ein Funktionszeichen). Diese Mengenfunktion sei *monoton*, d. h.

(1) $$\text{mit } A < B \text{ ist } A_\varphi \leqq B_\varphi.$$

Ist dann $S = A \dot{+} B \dot{+} \cdots$, $D = AB \ldots$

Summe und Durchschnitt beliebig vieler Mengen, so ist

$$S_\varphi \geqq A_\varphi \dot{+} B_\varphi \dot{+} \cdots, \quad D_\varphi \leqq A_\varphi B_\varphi \cdots.$$

Es gibt Mengen A, für die $A \leqq A_\varphi$, z. B. die Nullmenge $A = 0$; es gibt auch Mengen A, für die $A \geqq A_\varphi$, z. B. der ganze Raum $A = E$. Man erkennt nun sofort:

I. *Die Summe beliebig vieler Mengen mit $A \leqq A_\varphi$ ist wieder eine solche; der Durchschnitt beliebig vieler Mengen mit $A \geqq A_\varphi$ ist wieder eine solche.*

Danach läßt sich für eine beliebige Menge M definieren:

die Summe $\underline{M}$ aller Mengen $A \leqq M$ mit $A \leqq A_\varphi$ (zu denen gewiß $A = 0$ gehört) oder die größte Menge $A \leqq M$ mit $A \leqq A_\varphi$: *der φ-Kern* [75] *von* M;

der Durchschnitt $\overline{M}$ aller Mengen $A \geqq M$ mit $A \geqq A_\varphi$ (zu denen gewiß $A = E$ gehört) oder die kleinste Menge $A \geqq M$ mit $A \geqq A_\varphi$: *die φ-Hülle von M.*

Das typische Beispiel liefert $A_\varphi = A_\beta$, die Menge der Häufungspunkte von A in einem (metrischen) Raum. Hier bedeutet

$$A \leqq A_\beta : A \text{ insichdicht,}$$
$$A \geqq A_\beta : A \text{ abgeschlossen,}$$

und wir gelangen zum insichdichten Kern $\underline{M} = M_k$ sowie zur abgeschlossenen Hülle $\overline{M} = M_\alpha$.

Es kann sein, daß stets $A \leqq A_\varphi$ ist; dann ist $\underline{M} = M$ und $\overline{M}$ die kleinste Menge $A \geqq M$ mit $A = A_\varphi$. Beispiel $A_\varphi = A_\alpha$, Menge der α-Punkte von A, wobei $\overline{M}$ wieder die abgeschlossene Hülle von M wird

Es kann sein, daß stets $A \geqq A_\varphi$ ist; dann ist $\overline{M} = M$ und $\underline{M}$ die größte Menge $A \leqq M$ mit $A = A_\varphi$. Beispiel $A_\varphi = A_i$, Menge der inneren Punkte von A, wobei $\underline{M}$ der offene Kern M_i wird.

Man kann diese Spezialfälle herbeiführen, indem man die mit A_φ zugleich monotonen Funktionen

$$(2) \qquad A_s = A \dotplus A_\varphi, \quad A_d = A\,A_\varphi$$

betrachtet[1]. Hier ist stets $A \leqq A_s$, und $A = A_s$ mit $A \geqq A_\varphi$ gleichbedeutend; ferner stets $A \geqq A_d$, und $A = A_d$ mit $A \leqq A_\varphi$ gleichbedeutend. Für Bildung des Kerns kann also A_φ durch A_d, für Bildung der Hülle A_φ durch A_s ersetzt werden, und wir haben:

Der φ-Kern $\underline{M}$ ist die größte Menge $A \leqq M$ mit $A = A_d$; die φ-Hülle $\overline{M}$ ist die kleinste Menge $A \geqq M$ mit $A = A_s$.

Beispiel. Für $A_\varphi = A_\beta$ ist $A_s = A_\alpha$, $A_d = A_h$ (Kohärenz von A); der insichdichte Kern M_k ist die größte Menge $A \leqq M$ mit $A = A_h$; die abgeschlossene Hülle M_α die kleinste Menge $A \geqq M$ mit $A = A_\alpha$.

2. Eingreifen der Ordnungszahlen. Betrachten wir zuerst $\underline{M}$, die größte Menge $A \leqq M$ mit $A = A_d$. Diese Menge A erfüllt dann, wegen der Monotonie der Funktion A_d, auch die Ungleichung $A = A_d \leqq M_d$, sie ist nicht nur in M, sondern in der (im allgemeinen) kleineren Menge M_d enthalten. Sie ist dann folglich auch in $M_{dd}, M_{ddd}, \ldots$ enthalten, womit natürlich wiederholte Bildung der Mengenfunktion A_d gemeint ist: $M_{dd} = (M_d)_d$ usw. Sie ist dann auch in dem Durchschnitt $M M_d M_{dd} \ldots$ der bisher gebildeten Mengen und darüber hinaus in allen Mengen ent-

[1]) A_s bedeutet hier nicht den separierten Teil von A.

halten, die durch weitere Anwendung des Prozesses d entstehen. D. h.
[76] wenn wir induktiv jeder Ordnungszahl ξ eine Menge M_ξ zuordnen durch die Vorschrift:

$$(3)\qquad \begin{cases} M_0 = M \\ M_{\xi+1} = (M_\xi)_d \\ M_\eta = \mathfrak{D}_{\xi<\eta} M_\xi \quad (\eta \text{ Limeszahl}), \end{cases}$$

so ist unser $A = \underline{M}$ in allen M_ξ enthalten.

Für jede Ordnungszahl $\eta > 0$ gilt die Zerlegung [1])

$$(4)\qquad M = \sum_{\xi<\eta} (M_\xi - M_{\xi+1}) + M_\eta.$$

Denn wenn $x \varepsilon M$ nicht zu M_η gehört, so sei $M_\zeta (0 < \zeta \leqq \eta)$ die Menge mit niedrigstem Index, der x nicht angehört; dieses ζ kann wegen der dritten Vorschrift (3) keine Limeszahl sein, wir können demnach setzen $\zeta = \xi + 1$ $(0 \leqq \xi < \eta)$, und $x \varepsilon M_\xi - M_{\xi+1}$.

Die disjunkten Summanden $M_\xi - M_{\xi+1}$ können aber nicht unaufhörlich > 0 sein, da ihre Summe die Mächtigkeit von M nicht überschreiten darf. Es sei η die kleinste Ordnungszahl, für die $M_\eta = M_{\eta+1} = M_{\eta d}$. Da diese Menge M_η also auch der Bedingung $A = A_d$ genügt, so ist sie in der größten Menge $\underline{M}$ dieser Art enthalten, andererseits war aber $\underline{M}$ in jedem M_ξ enthalten, also ist $\underline{M} = M_\eta$.

Also: man erhält den φ-Kern von M in der Weise, daß man in dem nach (3) gebildeten System absteigender Mengen

$$(5)\qquad M_0, M_1, M_2, \ldots, M_\omega, M_{\omega+1}, \ldots,$$

das mit M, M_d, M_{dd}, ... beginnt, die erste Menge M_η sucht, die mit $M_{\eta+1}$ übereinstimmt. Sie stimmt dann auch mit allen folgenden überein und ist die *kleinste* Menge des Systems. Es ist bemerkenswert, daß der φ-Kern $\underline{M}$, ursprünglich als größte Menge $A \leqq M A_d$ oder als Summe aller dieser Mengen definiert, jetzt als kleinste Menge des Systems (5) oder Durchschnitt aller Mengen dieses Systems erscheint, also nicht von unten, sondern von oben her erreicht wird. Der „Kern" kommt hier wirklich durch Ablösung der „Schale" zum Vorschein.

Im Beispiel $A_\varphi = A_\beta$, $A_d = A_h$ sind die Mengen (5) die *Kohärenzen* M, M_h, M_{hh}, ... von M, wobei der Prozeß $A_h = A - A_j$ in der Abspaltung der isolierten Punkte besteht; hier sind also die Summanden $M_\xi - M_{\xi+1}$ der Formel (4) isolierte Mengen und man erhält den *insichdichten Kern* oder die *kleinste Kohärenz*, indem man den Prozeß der Abspaltung isolierter Punkte bis zum Stillstand fortsetzt. — Ist insbesondere M abgeschlossen, so ist (5) die Reihe der Ableitungen M, M_β, $M_{\beta\beta}$, ... und der insichdichte, in diesem Fall *perfekte Kern* ist die *kleinste Ableitung*.

---

[1]) Auch für $\eta = 0$, wenn statt $\sum_{\xi<\eta}$ dann die Nullmenge gesetzt wird.

Bevor wir ein weiteres Beispiel für Kernbildung bringen, ist noch die φ-Hülle $\overline{M}$, die kleinste Menge $A \geqq M$ mit $A = A_s$, analog zu behandeln. Wegen der Monotonie von A_s hat man dann $A = A_s \geqq M_s$, A enthält mit M auch die (im allgemeinen) größere Menge M_s, also weiter M_{ss}, M_{sss} usf. Definieren wir die Mengen M^ξ durch die induktive Vorschrift

$$(6) \qquad \begin{cases} M^0 &= M \\ M^{\xi+1} &= (M^\xi)_s \\ M^\eta &= \underset{\xi<\eta}{\mathfrak{S}} M^\xi \quad (\eta \text{ Limeszahl}), \end{cases}$$

wobei die Zerlegung

$$(7) \qquad M^\eta = M^0 + \sum_{\xi<\eta} (M^{\xi+1} - M^\xi)$$

gilt und die Summanden $M^{\xi+1} - M^\xi$ einmal Null werden müssen, da ihre Summe die Mächtigkeit des Raumes nicht überschreiten darf, so finden wir die φ-Hülle von M als größte Menge des aufsteigenden Systems

$$(8) \qquad M^0, M^1, M^2, \ldots, M^\omega, M^{\omega+1}, \ldots,$$

d. h. als erste, die mit ihrer Nachfolgerin (und allen folgenden Mengen) übereinstimmt.

In dem Beispiel $A_\varphi = A_\beta$, $A_s = A_\alpha$ beginnt das System (8) mit $M, M_\alpha, M_{\alpha\alpha}, \ldots$ und schon die zweite Menge ist die größte, $M_\alpha = M_{\alpha\alpha} = \cdots$ die abgeschlossene Hülle von M.

Sind die Elemente des Raumes etwa reelle Funktionen $x = x(t)$ einer Variablen t, die ihrerseits die Menge der reellen Zahlen oder irgendeine Menge T durchläuft, versteht man sodann unter A_φ die Menge der Funktionen $y = y(t)$, die als Grenzfunktionen $y(t) = \lim x_n(t)$ von überall (in T) konvergenten Folgen von Funktionen $x_n \varepsilon A$ darstellbar sind, so ist $A_\varphi = A_s$; eine Funktionenmenge oder ein Funktionensystem $A = A_s$, das also durch Limesbildung nicht erweitert wird, heißt ein *Bairesches Funktionensystem*[1]). Das kleinste Bairesche System $\overline{M}$ über einem gegebenen System M wird als größtes Glied der Reihe (8) gefunden. Hier hat dieses größte Glied M^η gewiß einen Index $\eta \leqq \Omega$, wo $\Omega = \omega_1$ die Anfangszahl von $Z(\aleph_1)$ ist; denn sind $x_1, x_2, \ldots$ Funktionen aus $M^\Omega = \underset{\xi<\Omega}{\mathfrak{S}} M^\xi$ und etwa $x_n \varepsilon M^{\xi_n}$, so sei ξ $(< \Omega)$ die erste, alle ξ_n übertreffende Ordnungszahl, und dann gehören alle x_n zu M^ξ, also $\lim x_n$ zu $M^{\xi+1} \leqq M^\Omega$, M^Ω ist ein Bairesches System.

Auch das kleinste Borelsche System (§ 18) über einem gegebenen Mengensystem fällt unter dies Schema der φ-Hüllen; man hat nur die jetzigen Mengen A durch Mengensysteme $\mathfrak{A}$ zu ersetzen und etwa $\mathfrak{A}_\varphi = \mathfrak{A}_{\sigma\delta}$ (kleinstes δ-System über dem kleinsten σ-System über $\mathfrak{A}$) zu wählen, also je zwei Stufen des damaligen Aufbaus zu einer zusammenzufassen.

---

[1]) Genaueres darüber § 43.

[77] **3. Residuen.** Ein sehr bemerkenswertes Beispiel für Kernbildung erhält man auf folgende Weise. Es sei

$$A_\varrho = A_\alpha - A \tag{9}$$

die Menge der nicht zu A gehörigen Häufungspunkte von A (der Rand des Komplements $E - A$) und

$$A_\psi = A_{\varrho\varrho} \tag{10}$$

die durch zweimalige Anwendung dieses Prozesses entstehende Menge. Es folgt

$$A_\alpha \geqq A_{\varrho\alpha} \geqq A_{\psi\alpha} \tag{11}$$

und die Vergleichung der Formeln

$$\begin{aligned} A_\alpha &= A_\varrho + A \\ A_{\varrho\alpha} &= A_\varrho + A_\psi \end{aligned}$$

ergibt, daß $A \geqq A_\psi$, daß $A_\psi = A A_{\varrho\alpha}$ in A abgeschlossen ist und daß

$$A - A_\psi = A_\alpha - A_{\varrho\alpha} \tag{12}$$

Differenz abgeschlossener Mengen ist.

Die Funktion A_ψ ist nicht monoton; z. B. ist für den ganzen Raum $E_\varrho = 0$, $E_\psi = 0$, während etwa für eine Menge A, die samt ihrem Komplement in E dicht ist (wie die Menge der rationalen Punkte im Euklidischen Raum), $A_\varrho = E - A$ und $A_\psi = A$ wird. Wenn wir aber nicht das System aller Punktmengen, sondern nur *das System $\mathfrak{P}$ der in einer festen Menge P abgeschlossenen Mengen* betrachten, so ist mit $A \,\varepsilon\, \mathfrak{P}$ auch $A_\psi \,\varepsilon\, \mathfrak{P}$ und A_ψ in $\mathfrak{P}$ eine monotone Funktion. Denn sei $E = P + Q$; es ist

$$A_\varphi = P(Q A_\alpha)_\alpha$$

eine (für alle $A \leqq E$ definierte) monotone Funktion von A; schreiben wir

$$B = Q A_\alpha, \quad A_\varphi = P B_\alpha.$$

Da B in Q abgeschlossen ist, ist jedenfalls $A_\varphi = B_\alpha - QB_\alpha = B_\alpha - B = B_\varrho$, und wenn A in P abgeschlossen ist, ist ebenso $B = A_\varrho$, $A_\varphi = A_\psi$; dabei ist A_ψ in A, also wieder in P abgeschlossen, und $A_d = A_\varphi = A_\psi$. Bilden wir also für eine in P abgeschlossene Menge M die Reihe (5), so folgt induktiv, daß auch alle M_ξ in P abgeschlossen sind und $M_{\xi+1} = M_{\xi\psi}$ ist; wir nennen diese Mengen M_ξ, deren Reihe mit $M, M_\psi, M_{\psi\psi}, \ldots$ beginnt,
[78] *Residuen* von M. Das *kleinste Residuum* ist der φ-Kern $\underline{M}$ von M, d. h. (da man $P = M$ wählen kann), *die größte in M abgeschlossene Menge A, für die $A = A_\psi$.* Eine Menge, deren kleinstes Residuum verschwindet, heiße *reduzibel.*

Nach § 17, 3 ergibt sich:

II. *Die reduziblen Mengen sind die aus abgeschlossenen Mengen gebildeten (endlichen oder unendlichen) Differenzenketten.*

Zunächst folgt nämlich aus (12)

$$(13) \qquad \begin{cases} M_\xi - M_{\xi+1} = F_\xi - F'_\xi, \\ \quad F_\xi \;\; = M_{\xi\alpha}, \quad F'_\xi = M_{\xi_{\varrho\alpha}}. \end{cases}$$

Wegen (11), angewendet auf $A = M_\xi$, ist $F_\xi \geqq F'_\xi \geqq F_{\xi+1}$, und aus $\xi < \eta$ folgt $\xi + 1 \leqq \eta$, $M_{\xi+1} \geqq M_\eta$, $F_{\xi+1} \geqq F_\eta$. Es ist also

$$(14) \qquad F_0 \geqq F'_0 \geqq F_1 \geqq F'_1 \geqq \cdots \geqq F_\omega \geqq F'_\omega \geqq \cdots$$

und für eine reduzible Menge M liefert (4), worin M_η und F_η schließlich verschwindet,

$$(15) \qquad M = \sum_\xi (M_\xi - M_{\xi+1}) = \sum_\xi (F_\xi - F'_\xi),$$

M ist eine Differenzenkette aus abgeschlossenen Mengen.

Dem umgekehrten Schluß schicken wir voraus: ist A in $M = F - N$ abgeschlossen (F abgeschlossen), so ist A_ϱ in N abgeschlossen; denn wegen $A_\alpha \leqq F$ ist $A_\alpha - A = (F - M) A_\alpha = N A_\alpha$. Eine zweimalige Anwendung gibt: ist A in $M = (F - F') + P$ abgeschlossen (F und F' abgeschlossen, $F \geqq F' \geqq P$), so ist A_ϱ in $F' - P$ und $A_{\varrho\varrho} = A_\psi$ in P abgeschlossen. Sei demnach aus den abgeschlossenen, schließlich verschwindenden Mengen (14) eine Differenzenkette

$$M = \sum_\xi (F_\xi - F'_\xi)$$

gebildet und setzen wir

$$P_\xi = \sum_{\eta \geqq \xi} (F_\eta - F'_\eta);$$

es ist

$$P_0 = M, \quad P_\xi = (F_\xi - F'_\xi) + P_{\xi+1},$$
$$P_\eta = \mathfrak{D}_{\xi < \eta} P_\xi \quad (\eta \text{ Limeszahl}).$$

Mit einer in M abgeschlossenen Menge A bilden wir die Residuen A_ξ ($A_0 = A$, $A_1 = A_\psi$, $A_2 = A_{\psi\psi}, \ldots$). Dann ist A_0 in P_0 abgeschlossen; ist A_ξ in P_ξ abgeschlossen, so folgt nach der vorausgeschickten Betrachtung, daß $A_{\xi+1}$ in $P_{\xi+1}$ abgeschlossen ist; ist η Limeszahl und, für $\xi < \eta$, A_ξ in P_ξ abgeschlossen, so ist A_η in P_η abgeschlossen. Damit ist bewiesen, daß jedes A_ξ in P_ξ abgeschlossen ist und mit P_ξ schließlich Null wird, A ist reduzibel. Jede in der Differenzenkette M abgeschlossene Menge, insbesondere M selbst, ist reduzibel.

III. *Die reduziblen Mengen bilden einen Körper.*

Dies folgt aus der Schlußbemerkung von § 17, da der Durchschnitt beliebig vieler abgeschlossener Mengen wieder abgeschlossen ist.

Zu den reduziblen Mengen gehören unter vielen andern die *separierten*. Denn aus (9)(10) folgt $A_\varrho \leqq A_\beta$, $A_{\varrho\varrho} \leqq A_{\varrho\beta} \leqq A_{\beta\beta}$ und jedenfalls $A_\psi \leqq A_\beta$, wonach das kleinste Residuum jeder Menge insichdicht ist ($A = A_\psi \leqq A_\beta$) und für eine separierte Menge verschwindet. Man erhält auch aus (4), wenn die M_ξ Kohärenzen bedeuten, eine separierte Menge (deren kleinste Kohärenz 0 ist) unmittelbar als Differenzenkette aus abgeschlossenen Mengen, weil $A_\alpha \geqq A_\beta \geqq A_{h\alpha}$ und $A - A_h = A_\alpha - A_\beta$ ist, also

$$M_\xi - M_{\xi+1} = F_\xi - F'_\xi, \quad F_\xi = M_{\xi\alpha}, \quad F'_\xi = M_{\xi\beta}.$$

4. Fall eines separablen Raumes. Hier tritt die Vereinfachung ein, daß die kleinste Kohärenz oder das kleinste Residuum immer nach höchstens abzählbar vielen Schritten erreicht wird. Es gilt nämlich:

[79] IV. *Ist der Raum separabel, so ist ein auf- oder absteigend wohlgeordnetes System offener (oder abgeschlossener) Mengen höchstens abzählbar.*

Sei $$\{G_0, G_1, \ldots, G_\xi, \ldots\} \qquad (\xi < \mu)$$
ein aufsteigendes System offener Mengen, $G_\xi < G_\eta$ für $\xi < \eta$; wir können seinen Typus μ als Limeszahl annehmen. Unter den in $G_{\xi+1}$ enthaltenen speziellen Umgebungen V (S. 126) gibt es gewiß eine, V_ξ, die in G_ξ nicht enthalten ist. Dann ist $V_\xi \neq V_\eta$ für $\xi < \eta$ (V_ξ ist in G_η enthalten, V_η nicht) und die Menge der V_ξ wie die der G_ξ höchstens abzählbar. Ebenso verläuft der Beweis für ein absteigendes System; für abgeschlossene Mengen erhält man den Satz durch Komplementbildung.

Man wendet ihn z. B. auf absteigende Folgen abgeschlossener Mengen in der Form an: es sei Ω die Anfangszahl von $Z(\aleph_1)$ und jedem $\xi < \Omega$ eine abgeschlossene Menge F_ξ zugeordnet; mit $\xi < \eta$ sei $F_\xi \geqq F_\eta$. Dann können höchstens abzählbar viele dieser Mengen verschieden sein, d. h. es gibt ein (erstes) η derart, daß $F_\eta = F_{\eta+1} = \cdots$ mit allen folgenden Mengen identisch ist.

Natürlich gilt dies alles auch von den in einer festen Menge $M \leqq E$ offenen und abgeschlossenen Mengen, da M auch ein (höchstens) separabler Raum ist.

In dem System (5) wird also, falls seine Mengen alle in M abgeschlossen sind (z. B. wenn es sich um Kohärenzen oder Residuen handelt), die kleinste Menge, d. h. die erste M_η, für die $M_\eta = M_{\eta+1}$, sicher für $\eta < \Omega$ erreicht.

Die reduziblen Mengen sind jetzt also höchstens abzählbare Differenzenketten (vom Typus $< \Omega$) oder Summen von höchstens abzählbar vielen Mengen der Form $F - F'$. Jede solche Menge, die auch in der Form $F(E - F') = FG'$ als Durchschnitt einer abgeschlossenen und einer offenen Menge geschrieben werden kann, ist ein F_σ, Summe einer Folge abgeschlossener Mengen (§ 23, III). Demnach ist eine reduzible Menge ein $F_{\sigma\sigma} = F_\sigma$; da ihr Komplement wegen der Körpereigenschaft auch reduzibel ist, ist sie zugleich ein G_δ, Durchschnitt einer Folge offener Mengen. Also:

V. *Die reduziblen Mengen eines separablen Raumes sind gleichzeitig Mengen F_σ und G_δ.*

Als Gegenstück bemerken wir:

VI. *Ist der Raum eine F_{II}-Menge, so sind die Mengen, die gleichzeitig F_σ und G_δ sind, reduzibel.*

Sei M gleichzeitig F_σ und G_δ, A ihr kleinstes Residuum, also $A = A_{\varrho\varrho}$ oder mit $B = A_\varrho$

$$A_\alpha = A + B = B_\alpha = F.$$

A ist in M abgeschlossen, also gleichzeitig F_σ und G_δ; also ist auch $E - A$ und $B = F(E - A)$ gleichzeitig F_σ und G_δ. A und B sind beide in F dicht und, als F_σ mit dichtem Komplement, in F von erster Kategorie; danach ist auch F in sich von erster Kategorie, folglich $F = 0$ ($F > 0$ wäre ein F_{II}). M ist reduzibel.

In einem separablen F_{II}-Raum, z. B. im separablen vollständigen Raum sind die reduziblen Mengen identisch mit denen, die gleichzeitig F_σ und G_δ sind. Es ist merkwürdig, daß man die Frage, ob eine Menge [80]
zugleich F_σ und G_δ ist, hier durch ein wohlbestimmtes (wenn auch unendliches) Verfahren, nämlich Residuenbildung, entscheiden kann, während z. B. für die Mengen F_σ nichts Ähnliches bekannt ist.

Der Satz IV gestattet folgende Verallgemeinerung:

VII. *Ist der Raum separabel, so ist ein auf- oder absteigend wohl-* [81]
geordnetes System reduzibler Mengen höchstens abzählbar.

Es genügt, ein aufsteigendes System zu betrachten (Komplementbildung). Jeder Ordnungszahl $\beta < \Omega$ sei eine reduzible Menge M_β zugeordnet und für $\beta < \gamma$ sei $M_\beta \leqq M_\gamma$; es ist zu zeigen, daß die M_β schließlich gleich werden ($M_\beta = M_{\beta+1} = \cdots$ für geeignetes β). Es sei $M_{\beta\xi}$ das ξ-te Residuum von $M_\beta = M_{\beta 0}$ ($\xi < \Omega$, wie auch alle noch auftretenden Ordnungszahlen $< \Omega$ sind). Nach (13)(15) ist

$$M_\beta = \sum_\xi (M_{\beta\xi} - M_{\beta\xi+1}) = \sum_\xi (F_{\beta\xi} - F'_{\beta\xi}),$$
$$F_{\beta\xi} = M_{\beta\xi\alpha}, \quad F'_{\beta\xi} = M_{\beta\xi\varrho\alpha}.$$

Nun behaupten wir: es gibt eine feste, von β unabhängige Folge abgeschlossener Mengen

$$(16) \qquad F_0 \geqq F'_0 \geqq \cdots \geqq F_\xi \geqq F'_\xi \geqq \cdots$$

derart, daß schließlich $F_{\beta\xi} = F_\xi$ und $F'_{\beta\xi} = F'_\xi$ wird (für $\beta \geqq \beta_\xi$, wo man überdies die β_ξ mit ξ wachsend annehmen kann). Wenn dies bereits für alle $\xi < \eta$ feststeht und γ_η die erste Ordnungszahl nach allen β_ξ ist, so ist für $\beta \geqq \gamma_\eta$

$$M_\beta = \sum_{\xi<\eta} (F_\xi - F'_\xi) + M_{\beta\eta};$$

die $M_{\beta\eta}$ wachsen dann mit β, und ihre abgeschlossenen Hüllen $F_{\beta\eta}$, von denen das Gleiche gilt, werden nach IV schließlich identisch: $F_{\beta\eta} = F_\eta$ für $\beta \geqq \delta_\eta$. Dann ist aber

$$M_{\beta\eta\varrho} = F_\eta - M_{\beta\eta},$$

diese Mengen und ihre abgeschlossenen Hüllen $F'_{\beta\eta}$ nehmen mit wachsendem β ab und es ist, wieder nach IV, $F'_{\beta\eta} = F'_\eta$ für $\beta \geqq \beta_\eta (\geqq \delta_\eta \geqq \gamma_\eta)$. Damit ist die Behauptung für η bewiesen (auch der Anfang des induktiven Verfahrens, $\eta = 0$, ist derselbe). Nun werden die Mengen (16) selbst wieder schließlich identisch: es sei $F_\eta = F'_\eta = F_{\eta+1} = \cdots = F$. Für $\beta \geqq \beta_\eta$ ist dann

$$M_{\beta\eta} - M_{\beta\eta+1} = F_\eta - F'_\eta = 0,$$

$M_{\beta\eta}$ das kleinste Residuum von M_β, also $M_{\beta\eta} = 0$ und

$$M_\beta = \sum_{\xi<\eta} (F_\xi - F'_\xi) = M$$

von β unabhängig, womit VII bewiesen ist.

Es ist von Interesse, festzustellen, daß man Mengen mit *beliebig großem* Index ($< \Omega$) der kleinsten Kohärenz oder des kleinsten Residuums bilden kann. Betrachten wir z. B. beschränkte abgeschlossene, der Größe nach wohlgeordnete Mengen reeller Zahlen [1])

$$M = \{m_1, m_2, \ldots, m_\mu\},$$

wo also der Index von m_α die Werte $\alpha = 1, 2, \ldots, \mu$ $(1 \leqq \alpha \leqq \mu)$ durchläuft, für $\alpha < \beta$ zugleich $m_\alpha < m_\beta$ und, für eine Limeszahl β, m_β die obere Grenze der Zahlen $m_\alpha (\alpha < \beta)$ ist. Man kann solche Mengen mit vorgeschriebenem $\mu < \Omega$ sofort bilden, z. B. (μ unendlich angenommen), indem man eine konvergente Folge positiver Zahlen nach dem Typus μ ordnet, also mit $c_\alpha (0 \leqq \alpha < \mu)$ bezeichnet und für $1 \leqq \beta \leqq \mu$

$$m_\beta = \sum_{\alpha<\beta} c_\alpha$$

setzt.

Der Leser wird unschwer erkennen, daß die erste Ableitung M_1 aus denjenigen m_α besteht, deren Indizes $\alpha = \omega\,\alpha_1$ Limeszahlen sind, die zweite Ableitung M_2 wieder aus denjenigen ebengenannten m_α, wo auch $\alpha_1 = \omega\alpha_2$ Limeszahl, also $\alpha = \omega^2 \alpha_2$ ein Vielfaches von ω^2 ist, und so induktiv fortschließend, daß die ξ-te Ableitung M_ξ aus denjenigen m_α besteht, deren Index $\alpha = \omega^\xi \alpha_\xi$ ein Vielfaches von ω^ξ ist; für hinlänglich großes ξ sind solche Vielfache in der Reihe $1, 2, \ldots, \mu$ natürlich nicht mehr vorhanden und M_ξ verschwindet. Z. B. für $\mu = \omega^\eta$ und $\xi \leqq \eta$ besteht M_ξ aus den m_α mit

$$\alpha = \omega^\xi \cdot 1,\ \omega^\xi \cdot 2, \ldots,\ \omega^\xi \cdot \omega^{-\xi+\eta};$$

M_η besteht aus dem einen m_α mit $\alpha = \omega^\eta$ und $M_{\eta+1} = 0$ ist die kleinste Ableitung.

Um auch Mengen mit beliebig hohem Index des kleinsten Residuums nachzuweisen, bilden wir mit einer abgeschlossenen separierten Menge M (z. B. von der soeben behandelten Art) und ihren Ableitungen M_ξ die Differenzenkette

$$A = (M_0 - M_1) + (M_2 - M_3) + \cdots + (M_\omega - M_{\omega+1}) + \cdots$$

und setzen

[1]) Wohlgeordnete Mengen reeller Zahlen m sind nach Satz IV höchstens abzählbar, da die zugehörigen Halbgeraden $(-\infty, m]$ ein wohlgeordnetes System abgeschlossener Mengen bilden.

$$A_\xi = \sum_{\eta \geqq \xi} (M_{2\eta} - M_{2\eta+1}) = (M_{2\xi} - M_{2\xi+1}) + (M_{2\xi+2} - M_{2\xi+3}) + \cdots.$$

Es ist $A = (M_0 - M_1) + A_1$. Da $M_0 - M_1 = M_i$ der isolierte Teil von M und im separierten Teil (§ 23, (15)), in diesem Fall in M selbst dicht ist, so ist A in M dicht, $A_\alpha = M_0$ und $A_\varrho = M_0 - A = M_1 - A_1$. Hier ist $A_1 \leqq M_2$, $A_\varrho \geqq M_1 - M_2$ in M_1 dicht, und die Wiederholung des Schlusses liefert $A_\psi = A_1$. Das erste Residuum von A ist also A_1, und die induktive Fortsetzung zeigt, daß A_ξ das ξ-te Residuum von A ist; es verschwindet schließlich, aber erst dann, wenn $M_{2\xi}$ die kleinste Ableitung ($= 0$) von M ist, und mit dieser erhält also das kleinste Residuum von A einen beliebig großen Index.

§ 31. Sonstige Anwendungen der Ordnungszahlen.

1. Maximal- und Minimalmengen. In den bisherigen Fällen konnten die Ordnungszahlen zur Not entbehrt werden, da die Existenz der Hüllen und Kerne von vornherein feststand; immerhin dienten sie zur Erforschung feinerer Merkmale, z. B. des Index der kleinsten Kohärenz. In andern Fällen aber läßt sich sogar der Existenzbeweis für gewisse Objekte nicht ohne sie führen. Wenn in einem Mengensystem $\mathfrak{A}$ keine größte Menge (die Summe aller Mengen des Systems) existiert, so kann es nichtsdestoweniger (verschiedene) *Maximalmengen* A in dem Sinne geben, daß im System keine Menge $> A$ vorhanden ist. Zu einer solchen Maximalmenge wird man im allgemeinen nur folgendermaßen gelangen. Man wählt im System eine beliebige Menge A_0, sodann, wenn Mengen $> A_0$ existieren, unter diesen eine Menge A_1, wenn noch Mengen $> A_1$ existieren, unter diesen eine Menge A_2 usw. Gelangt man für endliches ν zu einer Menge A_ν derart, daß keine Menge $> A_\nu$ existiert, so hat man eine Maximalmenge. Andernfalls findet man zunächst eine Mengenfolge $A_0 < A_1 < A_2 < \cdots$ vom Typus ω. Es kann nun sein, daß keine Menge A im System existiert, die alle A_ν umfaßt; dann ist das Unternehmen gescheitert [1]). Wir machen daher die Annahme, daß eine aufsteigende Mengenfolge ohne letztes Element stets fortsetzbar sei, also:

(α) Ist η eine Limeszahl und

$$A_0 < A_1 < \cdots < A_\xi < A_{\xi+1} < \cdots \quad (\xi < \eta)$$

eine Folge wachsender Mengen des Systems $\mathfrak{A}$ vom Typus η, so gibt es in $\mathfrak{A}$ eine Menge A_η, die alle A_ξ enthält ($A_\xi < A_\eta$).

Dann läßt sich das obige Verfahren fortsetzen und muß schließlich zu einer Maximalmenge führen. Denn definieren wir induktiv A_η ($\eta > 0$) als eine alle A_ξ ($\xi < \eta$) übertreffende Menge ($A_\xi < A_\eta$), falls nämlich alle A_ξ bereits definiert sind und eine solche Menge A_η vorhanden ist, während

[1]) In einem abzählbaren Raum gibt es unter den endlichen Teilmengen keine Maximalmenge, unter den abzählbaren keine Minimalmenge.

andernfalls A_η nicht definiert wird, so kann die Menge der definierten A_η die Mächtigkeit des Systems $\mathfrak{A}$ nicht überschreiten, und es gibt eine kleinste Ordnungszahl $\zeta > 0$, für die A_ζ nicht definiert ist. Nach (α) ist ζ keine Limeszahl, also $\zeta = \eta + 1$; A_η ist noch definiert und die Mengen $A_0, \ldots, A_\eta$ wachsen mit ihren Indizes. Es gibt im System keine Menge $> A_\eta$, da diese sonst als $A_{\eta+1}$ gewählt werden könnte; A_η ist eine Maximalmenge.

Der hier geschilderte Prozeß unterscheidet sich von denen in § 30 hauptsächlich durch die mangelnde Zwangläufigkeit: man hat im allgemeinen für jedes A_ξ die Wahl unter mehreren möglichen Mengen und gelangt so auch zu verschiedenen Maximalmengen. Man kann diese Willkür nicht beseitigen, nur zurückverlegen, indem man das System $\mathfrak{A}$ von vornherein zu $\mathfrak{A}^*$ wohlordnet und dann für jedes A_ξ unter den verfügbaren Mengen die in $\mathfrak{A}^*$ erste wählt; eine bestimmte Wohlordnung liefert eine bestimmte Maximalmenge, verschiedene Wohlordnungen können verschiedene Maximalmengen liefern. Wenn man das Zermelosche Wohlordnungsverfahren so anwendet, daß man als Ansatzelement (§ 12) von $\{A, B, \ldots\}$, *wenn möglich*, eine Menge wählt, die $A, B, \ldots$ als echte Teilmengen enthält, so beginnt $\mathfrak{A}^*$ mit wachsenden Mengen $A_0 < A_1 < \cdots < A_\eta$, deren letzte eine Maximalmenge ist; das ist nichts wesentlich anderes als das ursprünglich angegebene Verfahren, nur daß dort $A_{\eta+1}$ nicht mehr definiert war, während die Wohlordnung $\mathfrak{A}^*$ noch weitergeht.

Häufig wird es möglich sein, die wachsenden Mengen A_ξ so zu bilden, daß jede durch Hinzufügung eines einzigen Punktes x_ξ in ihre Nachfolgerin $A_{\xi+1}$ übergeht. Ein einfaches Beispiel bilden die früher (S. 126) gelegentlich verwendeten *Netze* $E(\varrho)$, d. h. Maximalmengen im System derjenigen Mengen A, in denen je zwei Punkte eine Entfernung $\geqq \varrho > 0$ haben. Man wählt im Raume einen beliebigen Punkt x_0; sind für $\eta > 0$ die Punkte x_ξ (mit $\xi < \eta$) bereits definiert, so nenne man die Menge dieser Punkte A_η und definiere x_η als Punkt, der von allen x_ξ eine Entfernung $\geqq \varrho$ hat, wenn es einen solchen gibt; andernfalls werde x_η nicht definiert. Die Menge der definierten Punkte kann die Mächtigkeit des Raumes nicht überschreiten; ist η die erste Ordnungszahl, für die x_η nicht definiert ist, so ist A_η ein Netz $E(\varrho)$.

Ähnlich konstruiert man mit G. Hamel eine *Basis A* für die Menge E der reellen Zahlen [1]). Ist A eine Menge reeller Zahlen, so sei $[A]$ die Menge der reellen Zahlen

$$(1) \qquad y = \sum_x^A r x,$$

die sich als lineare Verbindungen *endlich* vieler Zahlen $x \varepsilon A$ mit *rationalen* Koeffizienten r darstellen lassen (in der Summe sind also nur endlich viele $r \neq 0$ zugelassen). Wenn mindestens eine Gleichung

[1]) oder allgemeiner für einen linearen Raum E.

$$(2) \qquad 0 = \sum_x^A r\,x$$

mit nicht sämtlich verschwindenden r besteht, heißt A rational *abhängig*, andernfalls rational *unabhängig* oder eine *Basis* für $[A]$; wenn A rational unabhängig ist, läßt sich jede Zahl y von $[A]$ nur auf eine einzige Weise in der Gestalt (1) darstellen. Eine Basis A für E, die als Maximalmenge im System aller Basen (übrigens auch als Minimalmenge im System der Mengen A mit $[A] = E$) aufzufassen ist, erhält man so: man wähle eine reelle Zahl $x_0 \neq 0$, dann eine x_1, die nicht rationales Vielfaches $r_0 x_0$ von x_0 ist, dann eine x_2, die nicht rationale Verbindung $r_0 x_0 + r_1 x_1$ von x_0, x_1 ist; allgemein: ist $\eta > 0$, sind für $\xi < \eta$ die Zahlen x_ξ, deren Menge A_η heiße, bereits definiert und ist noch $[A_\eta] < E$, so sei x_η eine Zahl aus $E - [A_\eta]$; andernfalls werde x_η nicht definiert. Ist η die kleinste Ordnungszahl, für die x_η nicht definiert ist, so ist A_η eine Basis für E.

Ist A eine Basis für E, wird für $x \varepsilon A$ die reelle Funktion $f(x)$ ganz willkürlich definiert und sodann für jedes reelle y gemäß der (einzigen) Darstellung (1)

$$f(y) = \sum_x^A r\,f(x)$$

erklärt, so ist stets $f(y + z) = f(y) + f(z)$. Diese Funktionalgleichung [82]
hat also (da die Basis A offenbar von der Mächtigkeit des Kontinuums ist) $\aleph^\aleph$ Lösungen; setzt man voraus, daß $f(y)$ (an einer einzigen Stelle und infolgedessen überall) stetig ist, so hat sie bekanntlich nur die Lösungen $f(y) = c\,y$ mit konstantem c.

Für *Minimalmengen* (Mengen A, zu denen im System $\mathfrak{A}$ keine Menge $< A$ vorkommt) ist der Existenzbeweis natürlich analog zu führen, unter einer zu (α) entsprechenden Annahme (β) über Fortsetzbarkeit absteigender Mengenfolgen mit Limestypus. Als Beispiel betrachten wir im Raum E alle Kontinua A (zusammenhängende abgeschlossene Mengen), die zwei feste, verschiedene Punkte x, y enthalten. Eine Minimalmenge A, die also kein kleineres, die Punkte x, y verbindendes Kontinuum als Teilmenge hat, heißt nach L. Zoretti ein *zwischen x und y irreduzibles Kontinuum* [1]. Eine Kreisperipherie, auf der x, y liegen, ist reduzibel, während jeder der beiden Kreisbögen, die x und y zu Endpunkten haben, irreduzibel ist.

Die Existenz solcher Minimalmengen ist natürlich nicht in allen Fällen sicher. Jedoch gilt (Z. Janiszewski):

Ein kompaktes Kontinuum E, das die Punkte x, y enthält, hat (mindestens) *ein zwischen x, y irreduzibles Teilkontinuum.*

[1]) Dieser Begriff hat mit dem in § 30, 3 erklärten Begriff reduzibel nichts zu schaffen. Vgl. § 39.

Wir definieren für $\xi < \Omega$ (Ω Anfangszahl von $Z(\aleph_1)$, wie immer) die Mengen A_ξ folgendermaßen:

$A_0 = E$;

$A_{\xi+1}$ ein x, y enthaltendes Kontinuum $< A_\xi$, wenn es ein solches gibt, andernfalls $A_{\xi+1} = A_\xi$;

$A_\eta = \mathfrak{D}_\xi A_\xi$ ($\xi < \eta$, η Limeszahl).

Man sieht, daß die A_ξ mit wachsendem Index abnehmen (Gleichheit nicht ausgeschlossen) und x, y enthaltende Kontinua sind; für die A_η mit Limesindex folgt das aus dem Satze § 29, XVIII, indem man η als Limes einer Folge $\xi_1 < \xi_2 < \cdots$ darstellt und $A_\eta = A_{\xi_1} A_{\xi_2} \ldots$ beachtet. Da E separabel ist, gilt der Satz § 30, IV und es muß einmal $A_\eta = A_{\eta+1}$, A_η zwischen x und y irreduzibel sein.

Man kann übrigens so verfahren, daß spätestens A_ω irreduzibel wird. Man betrachte die (im Raum E) offenen Komplemente G der Mengen A und suche eine Maximalmenge im System der G. Es seien $V_1, V_2, \ldots$ die speziellen Umgebungen (§ 25) im Raume E. Wenn E reduzibel ist, also ein $G_1 > 0$ existiert, so wähle man dieses so, daß das erste $V_p \leqq G_1$ möglichst kleinen Index hat; ist $A_1 = E - G_1$ reduzibel, also ein $G_2 > G_1$ vorhanden, so wähle man dieses so, daß das erste V_q ($q > p$), das in G_2, aber nicht in G_1 enthalten ist, einen möglichst kleinen Index hat, und fahre so fort. Entweder ist ein G_n mit endlichem Index oder spätestens $G_\omega = G_1 \dot{+} G_2 \dot{+} \cdots$ ein maximales G, also $A_n = E - G_n$ oder $A_\omega = E - G_\omega$ irreduzibel.

Wir geben noch ein weiteres Beispiel für Anwendung der Ordnungszahlen, wo es sich nicht mehr um Maximal- oder Minimalmengen handelt.

2. Total imperfekte Mengen. E sei ein perfekter, separabler, vollständiger Raum (z. B. ein Euklidischer). Das System der perfekten Mengen $P > 0$ hat die Mächtigkeit $\aleph$ (§ 25, V) und jede Menge P auch E selbst, hat die Mächtigkeit $\aleph$ (§ 26, XI). Eine Menge, die kein P als Teilmenge enthält, heiße nach F. Bernstein *total imperfekt*, so z. B. jede höchstens abzählbare Menge oder jede Menge von der Mächtigkeit $< \aleph$. Es gibt aber auch *total imperfekte Mengen von der Mächtigkeit* $\aleph$, und dies scheint wiederum nicht ohne Ordnungszahlen beweisbar zu sein. Es sei $\aleph = \aleph_\mu$, wo μ eine unbekannte Ordnungszahl $\geqq 1$ ist, ω_μ die Anfangszahl der Zahlenklasse $Z(\aleph_\mu)$. Wir können die sämtlichen perfekten Mengen $P > 0$ in eine Folge vom Typus ω_μ bringen

$$P_0, P_1, \ldots, P_\omega, \ldots, P_\xi, \ldots \quad (\xi < \omega_\mu).$$

Wir wählen zwei verschiedene Punkte $x_0, y_0 \,\varepsilon\, P_0$. Sind, für $0 < \eta < \omega_\mu$, alle Punkte x_ξ, y_ξ ($\xi < \eta$) bereits definiert und werden die von ihnen gebildeten Mengen A_η, B_η genannt, so ist $A_\eta \dot{+} B_\eta$ von einer Mächtigkeit $< \aleph$ und es gibt also noch unendlich viele ($\aleph$) Punkte von P_η, die der Menge

$A_\eta \dot{+} B_\eta$ nicht angehören; zwei verschiedene solche mögen als x_η, y_η definiert werden. Aus dieser Erklärung folgt, daß für $\xi < \eta$ die vier Punkte $x_\xi, y_\xi, x_\eta, y_\eta$ verschieden sind. Sei A die Menge aller Punkte x_ξ; ihr Komplement $B = E - A$ enthält alle Punkte y_ξ (wobei man es übrigens so einrichten kann, daß B genau die Menge der Punkte y_ξ wird). Beide Mengen A, B sind von der Mächtigkeit $\aleph$ und total imperfekt, denn da $A\,P_\xi$ den Punkt x_ξ und $B\,P_\xi$ den Punkt y_ξ enthält, ist niemals $P_\xi \leqq B$ oder $P_\xi \leqq A$.

Wir haben also E in zwei total imperfekte Mengen A, B gespalten. Das ist auch für die Theorie des Zusammenhangs interessant; denn A und B sind *diskontinuierlich* (S. 152) und doch, falls E ein Euklidischer Raum von mindestens zwei Dimensionen ist, *zusammenhängend*. Es gilt nämlich: *In einem Euklidischen Raum von mindestens zwei Dimensionen ist das Komplement einer total imperfekten Menge zusammenhängend* (W. Sierpiński).

Nehmen wir etwa E als Euklidische Ebene an; es ist die Unmöglichkeit einer Zerlegung $E = A + B$ zu zeigen, in der A total imperfekt und B unzusammenhängend ist. Sei $B = P + Q$ eine Zerstückelung in zwei relativ abgeschlossene Mengen $P = B\,P_\alpha$, $Q = B\,Q_\alpha$; wegen $P\,Q = B\,P_\alpha Q_\alpha = 0$ ist die abgeschlossene Menge $F = P_\alpha Q_\alpha$ in A enthalten. Verbinden wir zwei Punkte $p \varepsilon P$, $q \varepsilon Q$ durch einen Streckenzug, etwa $C = [p, m, q,]$, wo m auf der Mittelsenkrechten von p, q liegt. Jede (abgeschlossene) Teilstrecke von C ist perfekt, also nicht in A enthalten, enthält demnach Punkte von B; $C\,B$ ist in C dicht, $C \leqq B_\alpha$, $C = C\,B_\alpha = C\,P_\alpha \dot{+} C\,Q_\alpha$, wo nun, damit dies keine Zerstückelung des Kontinuums C sei, beide Summanden (die > 0 sind) gemeinsame Punkte haben müssen: $C\,P_\alpha Q_\alpha = C\,F > 0$. Die abgeschlossene Menge F wird also von C getroffen, und zwar in Punkten $\neq p, q$, da diese zu B gehören und $F \leqq A$ ist. Läßt man m die Mittellinie durchwandern und damit C variieren, so ergibt sich, daß F von der Mächtigkeit des Kontinuums und sein perfekter Kern in A enthalten ist, im Widerspruch zur Annahme eines total imperfekten A.

§ 32. Die Borelschen und Suslinschen Mengen.

Die von den *abgeschlossenen* Mengen F erzeugten Borelschen und [83]
Suslinschen Mengen (§ 18, 19) heißen die *Borelschen und Suslinschen* [84]
Mengen des Raumes E. [1])

Da der Durchschnitt F_δ einer Folge abgeschlossener Mengen (sogar beliebig vieler) auch noch abgeschlossen ist, so wird man bei Erzeugung

[1]) Die übliche Bezeichnung ist: B-Mengen und A-Mengen (ensembles mesurables B, ensembles A).

der Borelschen Mengen mit den F_σ beginnen, also wie in § 18, (3) oder gemäß der dortigen Vorschrift (α); man hat also folgende Definition für die Ordnungszahlen $\xi < \Omega$:

[85] (α) Die Mengen F^0 sind die Mengen F.
Die Mengen F^η sind für ungerades η die Summen, für gerades $\eta > 0$ die Durchschnitte aus Folgen von Mengen F^ξ $(\xi < \eta)$.

Die Mengen $F^0, F^1, F^2, F^3, \ldots$ sind die $F, F_\sigma, F_{\sigma\delta}, F_{\sigma\delta\sigma}, \ldots$, dann kommen die F^ω als Durchschnitte aus Folgen früherer Mengen usw.

Betrachten wir auch noch die von den *offenen* Mengen G erzeugten Borelschen Mengen, wobei man mit den G_δ beginnen wird, also wie in § 18, (4) oder gemäß der Vorschrift (β); man hat dann folgende Definition:

(β) Die Mengen G^0 sind die Mengen G.
Die Mengen G^η sind für ungerades η die Durchschnitte, für gerades $\eta > 0$ die Summen aus Folgen von Mengen G^ξ $(\xi < \eta)$.

Die $G^0, G^1, G^2, G^3, \ldots$ sind die $G, G_\delta, G_{\delta\sigma}, G_{\delta\sigma\delta}, \ldots$, dann kommen die G^ω als Summen aus Folgen früherer Mengen usw.

Offenbar sind die F^ξ und G^ξ Komplemente voneinander. Überdies aber sind beide Borelschen Systeme identisch, also auch die G^ξ bilden in ihrer Gesamtheit die Borelschen Mengen des Raumes, nämlich:

Jedes F^ξ ist ein $G^{\xi+1}$, jedes G^ξ ein $F^{\xi+1}$.

Dies gilt für $\xi = 0$ (jedes F ist ein G_δ, jedes G ein F_σ, § 23, III) und überträgt sich induktiv auf $\eta > 0$, wenn es für $\xi < \eta$ gilt: ist η ungerade, so ist F^η Summe einer Folge von Mengen F^ξ, diese sind Mengen $G^{\xi+1}$, also jedenfalls auch Mengen G^η, und die Summe einer Folge von G^η ist ein $G^{\eta+1}$. Analog für gerades η.

Auch folgendes ist induktiv leicht einzusehen. Die F^ξ bilden einen Ring [1]), außerdem für ungerades ξ ein σ-System, für gerades ξ ein δ-System. Die G^ξ bilden einen Ring und für ungerades ξ ein δ-System, für gerades ξ ein σ-System. Ist η Limeszahl, so ist das System aller $F^\xi (\xi < \eta)$ mit dem aller G^ξ identisch und ein Körper, aber im allgemeinen weder ein σ- noch ein δ-System; die Summen aus Folgen von Mengen dieses Systems sind die G^η, die Durchschnitte die F^η. Das ganze Borelsche System ist ein Körper und gleichzeitig σ- und δ-System. Die Mengen, die gleichzeitig F^ξ und G^ξ sind (*zweiseitige* oder *ambige* Borelsche Mengen), bilden ebenfalls einen Körper: so die Mengen, die zugleich offen und abgeschlossen sind (was in einem zusammenhängenden Raume nur für die beiden Mengen E und 0 zutrifft), sodann die Mengen, die zugleich F_σ und G_δ sind (im separablen vollständigen Raum sind das die reduziblen Mengen). Die Differenz zweier F^ξ, oder die zweier G^ξ, oder der Durchschnitt $F^\xi G^\xi$ eines F^ξ mit einem G^ξ, ist gleichzeitig $F^{\xi+1}$ und $G^{\xi+1}$.

---

1) Über die Begriffe Ring, Körper, σ-System, δ-System vgl. § 17, 18.

Die *Suslinschen Mengen* des Raumes sind die von abgeschlossenen Mengen F_{n_1} usw. erzeugten Mengen

$$F_S = \mathfrak{S}\, F_{n_1} F_{n_1 n_2} F_{n_1 n_2 n_3} \ldots$$

in der aus § 19 bekannten Bezeichnung. Wir erinnern uns, daß alle Borelschen Mengen auch Suslinsche sind und daß die Iteration des Suslinschen Prozesses nichts Neues liefert: jedes F_{SS} ist ein F_S. Die von den offenen Mengen erzeugten Suslinschen Mengen

$$G_S = \mathfrak{S}\, G_{n_1} G_{n_1 n_2} G_{n_1 n_2 n_3} \ldots$$

sind in ihrer Gesamtheit mit den F_S identisch; denn jedes G ist ein F_σ, also ein F_S, und G_S ein $F_{SS} = F_S$, ebenso F_S ein G_S. Die G_S sind aber nicht die Komplemente der F_S, die ja vielmehr so aussehen würden:

$$E - F_S = \mathfrak{D}(G_{n_1} \dotplus G_{n_1 n_2} \dotplus G_{n_1 n_2 n_3} \dotplus \cdots),$$

und die Suslinschen Mengen bilden, wie wir noch sehen werden, im allgemeinen keinen Körper.

Der Charakter einer Menge als Borelscher oder Suslinscher Menge ist relativ und bezieht sich auf den Raum E; man hätte also ausführlicher $F^\xi(E)$ usw. zu schreiben. Wir haben für die F, G, F_σ, G_δ davon schon gesprochen (§ 26, 3); ist $D \leqq E$ und wird das Argument E wieder weggelassen, so ist

$$F^\xi(D) = D\,F^\xi, \quad G^\xi(D) = D\,G^\xi,$$

d. h. die Borelschen Mengen des Raumes D sind die Durchschnitte von D mit den entsprechenden Borelschen Mengen des Raumes E. Der induktive Beweis liegt auf der Hand. Dasselbe gilt von den Suslinschen Mengen:

$$F_S(D) = D\,F_S, \quad G_S(D) = D\,G_S.$$

Es gibt auch *absolut* Borelsche und *absolut* Suslinsche Mengen, d. h. solche, [86]
die in jedem umgebenden Raume diesen Charakter haben, und zwar gibt es absolute F^ξ, ferner für $\xi \geqq 1$ absolute G^ξ (nicht für $\xi = 0$, absolut offene Mengen gab es ja nicht), und absolute F_S oder G_S, und der Leser wird leicht (wie S. 136) erkennen, daß eine Menge A einen dieser Charaktere absolut besitzt, wenn sie ihn in irgendeinem vollständigen Raume, speziell in ihrer vollständigen Hülle $\overline{A}$ besitzt. Daß die Nichtexistenz absoluter G der Existenz absoluter $G^\xi (\xi \geqq 1)$ nicht widerspricht, haben wir uns damals für $\xi = 1$ schon klargemacht; auch in G_S treten ja nur Summanden der Form $G_\delta = G^1$ auf.

Wir beweisen nun den *Mächtigkeitssatz*: [87]

I. *In einem vollständigen separablen Raum ist jede Suslinsche, also insbesondere jede Borelsche Menge entweder höchstens abzählbar oder von der Mächtigkeit* $\aleph$.

Das ist eine Verallgemeinerung von § 26, XI; jedoch entscheidet über die Mächtigkeit nicht mehr wie dort das Verschwinden oder Nichtverschwinden des insichdichten Kerns. Die Menge der rationalen Zahlen

12*

(ein F_σ) ist insichdicht und doch nur abzählbar. Wir werden wie damals zeigen, daß eine unabzählbare Suslinsche Menge ein dyadisches Diskontinuum D als Teilmenge enthält, müssen aber bei der Kugelkonstruktion statt der Häufungspunkte jetzt Verdichtungspunkte benutzen. Es sei

$$A = \underset{ikl\ldots}{\mathfrak{S}} F(i)\, F(i,k)\, F(i,k,l) \ldots$$

eine von abgeschlossenen Mengen F erzeugte Suslinsche Menge, die Summe über alle natürlichen Zahlenfolgen $(i, k, l, \ldots)$ erstreckt, während, wie wir sogleich bemerken wollen, $(p, q, r, \ldots)$ eine aus den Ziffern 1, 2 gebildete dyadische Zahlenfolge bedeuten soll. Zur Vereinfachung können wir für jede Folge natürlicher Zahlen

$$F(i) \geqq F(i,k) \geqq F(i,k,l) \geqq \cdots$$

annehmen (indem wir statt der ursprünglichen F die Durchschnitte mit den vorangehenden einführen). Hält man Anfangsindizes fest, so entstehen die Mengen

$$A(i) = \underset{kl\ldots}{\mathfrak{S}} F(i)\, F(i,k)\, F(i,k,l) \ldots,$$

$$A(i,k) = \underset{l\ldots}{\mathfrak{S}} F(i)\, F(i,k)\, F(i,k,l) \ldots$$

usw., wobei

$$A = \underset{i}{\mathfrak{S}} A(i), \quad A(i) = \underset{k}{\mathfrak{S}} A(i,k), \quad A(i,k) = \underset{l}{\mathfrak{S}} A(i,k,l), \ldots .$$

Nun sei A und folglich auch $A_v = A\, A_\gamma$ unabzählbar; wir wählen zwei Punkte a_1, a_2 dieser Menge (Verdichtungspunkte von und in A) und machen sie zu Mittelpunkten disjunkter abgeschlossener Kugeln V_1, V_2; für die zugehörigen offenen Kugeln U_1, U_2 ist also jede der beiden Mengen $A\, U_p$ unabzählbar. Dann ist mindestens ein Summand von $A\, U_p = \underset{i}{\mathfrak{S}} A(i)\, U_p$ unabzählbar, sagen wir $A(i_p)\, U_p$. Von und in dieser Menge wählen wir zwei Verdichtungspunkte a_{p1}, a_{p2} und umgeben sie mit disjunkten abgeschlossenen Kugeln $V_{p1}, V_{p2} \leqq U_p$; für die zugehörigen offenen Kugeln U_{p1}, U_{p2} ist jede der vier Mengen $A(i_p)\, U_{pq}$ unabzählbar. Dann ist mindestens ein Summand von $A(i_p)\, U_{pq} = \underset{k}{\mathfrak{S}} A(i_p, k)\, U_{pq}$ unabzählbar, sagen wir $A(i_p, k_{pq})\, U_{pq}$. In dieser Weise wird fortgefahren, und es entspricht also jeder dyadischen Zahlenfolge $(p, q, r, \ldots)$ eine natürliche Zahlenfolge $(i_p, k_{pq}, l_{pqr}, \ldots)$ derart, daß die Mengen

$$A(i_p)\, U_p, \quad A(i_p, k_{pq})\, U_{pq}, \quad A(i_p, k_{pq}, l_{pqr})\, U_{pqr}, \ldots$$

unabzählbar sind, um so mehr also die abgeschlossenen Mengen

$$F(i_p)\, V_p, \quad F(i_p, k_{pq})\, V_{pq}, \quad F(i_p, k_{pq}, l_{pqr})\, V_{pqr}, \ldots$$

(da $A(i) \leqq F(i)$ usw. ist). Hat man nun, woran nichts hindert, die Kugeln $V_p, V_{pq}, V_{pqr}, \ldots$ mit Radien $< 1, \frac{1}{2}, \frac{1}{3}, \ldots$ gewählt, so bilden die zuletzt genannten abgeschlossenen Mengen eine absteigende Folge mit Durchmessern $\to 0$, haben also einen einzigen Durchschnittspunkt x, der sowohl dem dyadischen Diskontinuum

$$D = \Sigma V_p \,.\, \Sigma V_{pq} \,.\, \Sigma V_{pqr} \ldots$$

als der Menge A angehört. Dies gilt für jede dyadische Folge, also jeden Punkt $x \varepsilon D$, und D ist Teilmenge von A.

Der Satz I ist der umfassendste Mächtigkeitssatz, den wir kennen. Aber er bezieht sich doch nur auf ein verschwindend kleines Teilsystem im System aller Punktmengen. Es gibt im separablen Raum $\aleph$ abgeschlossene Mengen (§ 25, IV), also $\aleph^{\aleph_0} = \aleph$ Folgen abgeschlossener Mengen oder Komplexe $(F_1, F_2, F_{11}, F_3, F_{12}, F_{21}, F_{111}, \ldots)$, wie sie zur Definition einer Suslinschen Menge dienen, also $\aleph$ Suslinsche Mengen (auch $\aleph$ Borelsche Mengen). Der separable vollständige Raum ist, wenn sein perfekter Kern nicht verschwindet [1]), von der Mächtigkeit $\aleph$ und hat $2^{\aleph} > \aleph$ Teilmengen. Für die überwiegende Mehrheit der Punktmengen bleibt also die Mächtigkeitsfrage und damit das Kontinuumproblem ungeklärt.

§ 33. Existenzbeweise.

Bei der Definition der Borelschen Mengen B in der Form F^ξ oder G^ξ erhebt sich die Frage, ob hier alle Indizes $\xi < \Omega$ wirklich notwendig sind, um das ganze Borelsche System $\mathfrak{B}$ zu erhalten. Das hängt natürlich vom Raume E ab. Besteht E aus lauter isolierten Punkten, so ist jede Punktmenge zugleich abgeschlossen und offen, und schon der erste Schritt des Verfahrens ist überflüssig. Ist der Raum abzählbar, so ist jede Punktmenge zugleich ein F_σ und G_δ, und die $F_{\sigma\delta}$ usw. liefern nichts Neues. Wenn wir also sagen: eine Menge G^η ist *genau* ein G^η oder gehört *genau* der Borelschen Klasse $\mathfrak{G}^\eta$ an, wenn sie nicht schon ein $G^\xi (\xi < \eta)$ ist, so bleibt die Frage, unter welchen Umständen es für jeden Index $\eta < \Omega$ Mengen gibt, die genau Mengen G^η sind. Dazu ist zu bemerken, daß in

$$\mathfrak{G}^0 \leqq \mathfrak{G}^1 \leqq \cdots \leqq \mathfrak{G}^\omega \leqq \mathfrak{G}^{\omega+1} \leqq \cdots$$

entweder lauter Ungleichungszeichen oder von einem Index an lauter Gleichheitszeichen gelten, denn wenn $\mathfrak{G}^\eta = \mathfrak{G}^{\eta+1}$, so ist $\mathfrak{G}^\eta$ bereits das ganze Borelsche System (S. 87).

Ferner war jede Borelsche Menge B auch eine Suslinsche Menge S, und man wird fragen, ob es Mengen S gibt, die keine B sind. In den obigen trivialen Räumen waren ja alle Punktmengen Borelsche Mengen, und jedes S war auch ein B. Damit ein S ein B sei, ist jedenfalls notwendig, daß ihr Komplement $E - S$ auch (ein B, also) ein S sei, weil die B einen Körper bilden, und ein S, dessen Komplement kein S ist, ist also gewiß kein B.

Beide Fragen lassen sich zusammen behandeln und finden ihre Antwort in dem Existenzsatz: [88]

[1]) Andernfalls ist er abzählbar und alle seine Punktmengen sind Mengen F_σ.

I. *In einem vollständigen Raum, dessen perfekter Kern nicht verschwindet, gibt es Suslinsche Mengen, die keine Borelschen sind, und für jede Ordnungszahl $\xi < \Omega$ Borelsche Mengen, die genau Mengen G^ξ (oder F^ξ) und nicht von niederer Klasse sind.*

Zum Beweise bilden wir im vollständigen Raum E mit $E_k > 0$ ein dyadisches Diskontinuum

$$D = \Sigma V_{p_1} \cdot \Sigma V_{p_1 p_2} \cdot \Sigma V_{p_1 p_2 p_3} \cdots$$

in der bekannten Weise (S. 131); die dyadischen Folgen $(p_1, p_2, \ldots)$ sind aus den Ziffern 1, 2 gebildet, und jeder solchen Folge entspricht umkehrbar eindeutig ein Punkt x von D, nämlich der Durchschnittspunkt der Mengen

$$V_{p_1} > V_{p_1 p_2} > \cdots,$$

deren Durchmesser nach 0 konvergieren.

Es gibt abzählbar viele Ziffernfolgen, die nur endlich viele Einsen enthalten; tilgen wir die zugehörigen Punkte, so bleibt die Menge C der x mit unendlich vielen $p_n = 1$ übrig, die aus der perfekten Menge D durch Weglassung einer Menge F_σ (einer abzählbaren) entsteht und daher ein G_δ ist. Wir wollen die Punkte $x \varepsilon C$ etwas anders, nämlich nicht mit dyadischen, sondern mit natürlichen Zahlenfolgen $(x_1, x_2, \ldots)$ darstellen. Für eine natürliche Zahl x sei unter $[x]$ ein dyadischer Komplex verstanden, der erst $x - 1$ Ziffern 2 und zuletzt eine Ziffer 1 hat, z. B.

$$[1] = (1), \quad [2] = (2, 1), \quad [3] = (2, 2, 1), \ldots,$$

ferner bedeute $[x_1, x_2, \ldots, x_k]$ den durch Aneinanderreihung von $[x_1]$, $[x_2], \ldots, [x_k]$ in dieser Reihenfolge entstehenden dyadischen Komplex, z. B. $[2, 3, 1] = (2, 1, 2, 2, 1, 1)$, und ebenso $[x_1, x_2, x_3, \ldots]$ die in gleicher Weise entstehende dyadische Ziffernfolge mit unendlich vielen Einsen (nämlich $p_{x_1} = p_{x_1+x_2} = p_{x_1+x_2+x_3} = \cdots = 1$, die übrigen $p_n = 2$). Setzt man dann

$$F_{x_1 x_2 \ldots x_k} = V_{[x_1, x_2, \ldots, x_k]},$$

so ist für jede Folge natürlicher Zahlen

$$F_{x_1} > F_{x_1 x_2} > F_{x_1 x_2 x_3} > \cdots$$

mit Durchmessern $\to 0$, die Mengen F mit k Indizes sind disjunkt (z. B. $F_2 = V_{21}$ und $F_3 = V_{221} < V_{22}$), und es ist

$$C = \Sigma F_{x_1} \cdot \Sigma F_{x_1 x_2} \cdot \Sigma F_{x_1 x_2 x_3} \cdots, \tag{1}$$

wobei jeder Punkt x eineindeutig einer Folge natürlicher Zahlen $(x_1, x_2, \ldots)$ entspricht (derart, daß x der einzige Punkt des Durchschnitts $F_{x_1} F_{x_1 x_2} \ldots$ ist).

[89] Es genügt nun, den Satz I für den (nicht vollständigen) Raum C zu beweisen. Denn Suslinsche Mengen in C und Borelsche Mengen G^ξ in $C (\xi \geqq 1)$ sind ebensolche in E, weil C in E ein $G_\delta = G^1$ ist, und umgekehrt, Mengen dieses Charakters in E haben ihn, falls sie Teilmengen von C sind, auch in C.

Die Mengen D und C sind separabel (z. B. ist die abzählbare Menge $D - C$ in D dicht). Wir betrachten in dem separablen Raume C die schon so häufig verwendeten speziellen Umgebungen (S. 126 mit V bezeichnet); nennen wir sie hier $U_1, U_2, \ldots$ und bilden mit ihnen das Mengensystem [90]

$$(2) \qquad \mathfrak{M} = \{U_1, U_2, \ldots\}.$$

Die von $\mathfrak{M}$ erzeugten Borelschen Mengen $B^\xi (B^0 = M,\ B^1 = M_\delta,\ B^2 = M_{\delta\sigma}, \ldots)$ und Suslinschen Mengen S sind mit denen des Raumes C identisch, da bereits die Mengen M_σ (also jedenfalls die B^2) alle offenen Mengen G dieses Raumes umfassen. Jedes B^ξ ist ein G^ξ, und jedes G^ξ jedenfalls ein $B^{2+\xi}$ (für $\xi \geqq \omega$ ein B^ξ).

Nun sei $\nu = (n_1, n_2, \ldots)$ eine Folge wachsender natürlicher Zahlen, N eine Menge solcher Folgen und

$$(3) \qquad X = \mathop{\mathfrak{S}}_{\nu}^{N} M_{n_1} M_{n_2} \ldots = \Phi(M_1, M_2, \ldots) \quad (M_n \,\varepsilon\, \mathfrak{M}),$$

wo die δs-Funktion Φ ihrer Argumente (die beliebige Mengen aus $\mathfrak{M}$ sind) durch die Menge N bestimmt ist. Wir wissen, daß bei geeignetem N oder Φ die Menge X genau die Mengen S oder auch die Mengen $B^\xi (1 \leqq \xi < \Omega)$ darstellt, nach den Sätzen § 18, II und § 19, II. Eine bestimmte Funktion Φ, gleichviel welche, werde gewählt und festgehalten; die damit in der Form (3) darstellbaren Mengen mögen *die Mengen Φ* heißen.

Jedem Punkt $x \,\varepsilon\, C$ ordnen wir nun, vermöge der durch ihn bestimmten Folge natürlicher Zahlen $x_1, x_2, \ldots$, die von x abhängige Menge

$$(4) \qquad \Phi(x) = \Phi(U_{x_1}, U_{x_2}, \ldots)$$

eindeutig zu. Das ist also eine Menge Φ, und *jede* Menge Φ ist so darstellbar, denn $M_n \,\varepsilon\, \mathfrak{M}$ heißt doch eben, daß $M_n = U_{x_n}$ eins der U sein soll. (Die natürlichen Zahlen $x_1, x_2, \ldots$ sind ganz beliebig, nicht etwa wachsend oder paarweise verschieden.) Jedes Φ ist also ein $\Phi(x)$, wobei übrigens recht wohl $\Phi(x) = \Phi(y)$ sein kann trotz $x \neq y$.

Der Punkt x kann nun der ihm zugeordneten Menge $\Phi(x)$ angehören oder nicht. Es sei A die Menge der $x \,\varepsilon\, \Phi(x)$, B die Menge der $x \,\bar{\varepsilon}\, \Phi(x)$; [91]
$A + B = C$. Dann ist B von allen $\Phi(x)$ verschieden, da von den beiden Mengen B und $\Phi(x)$ eine, und nur eine, den Punkt x enthält. *B gehört also nicht zu den Mengen Φ.*

Für eine natürliche Zahl n sei A_n die Menge der x, für die $x \,\varepsilon\, U_{x_n}$. Dann ist

$$(5) \qquad A = \Phi(A_1, A_2, \ldots)$$

mit derselben Funktion Φ wie zuvor. Denn $x \,\varepsilon\, A$ oder $x \,\varepsilon\, \Phi(x)$ besagt: für eine in N vorkommende Zahlenfolge $\nu = (n_1, n_2, \ldots)$ gehört x dem Summanden $M_{n_1} M_{n_2} \ldots = U_{x_{n_1}} U_{x_{n_2}} \ldots$ von $\Phi(M_1, M_2, \ldots) = \Phi(U_{x_1}, U_{x_2}, \ldots)$ an, dann ist aber $x \,\varepsilon\, A_{n_1} A_{n_2} \ldots$ und x gehört dem entsprechenden Summanden von $\Phi(A_1, A_2, \ldots)$ an. Und umgekehrt.

Wir haben schließlich noch zu erkennen, daß die A_n Borelsche Mengen sind. Ist k eine natürliche Zahl, so sei A_{nk} die Menge der x, für die $x_n = k$. Das heißt so viel, daß x einer der (in C) abgeschlossenen Mengen

$$C\, F_{x_1 \ldots x_{n-1} k}$$

(für $n = 1$ der Menge $C\, F_k$) angehört, mit beliebigen natürlichen Zahlen $x_1, \ldots, x_{n-1}$; A_{nk} ist die Summe dieser Mengen nach $x_1, \ldots, x_{n-1}$, also ein F_σ (in C). Endlich ist

$$A_n = \sum_k A_{nk}\, U_k\,,$$

also ebenfalls eine Borelsche Menge; denn $x \,\varepsilon\, U_{x_n}$ besagt doch, daß es eine natürliche Zahl k gibt mit $x \,\varepsilon\, U_k$, $x_n = k$, d. h. $x \,\varepsilon\, A_{nk}\, U_k$.

Damit ist unser Ziel erreicht, nämlich:

Stellt Φ genau die Suslinschen Mengen S dar, so ist nach (5) A eine von Suslinschen (sogar Borelschen) Mengen A_n erzeugte Suslinsche Menge, also selbst eine Menge S. Ihr Komplement B ist aber kein S, und danach ist A keine Borelsche Menge.

Stellt Φ genau die Borelschen Mengen B^ξ dar, so ist nach (5) A eine von Borelschen Mengen A_n erzeugte Borelsche Menge, also selbst eine Borelsche Menge des Raumes C. Ihr Komplement B ist dann auch eine Borelsche Menge, aber kein B^ξ, sondern also genau ein B^η mit $\eta > \xi$. Es gibt also in C Borelsche Mengen, die genau B^η sind mit *beliebig hohem* Index, und das gilt dann auch für die G^η. Dann gibt es aber für *jeden Index* η Mengen, die genau G^η sind. Das gilt zunächst für Indizes $\xi + 1$: wäre jedes $G^{\xi+1}$ ein G^ξ, so wäre $\mathfrak{G}^\xi = \mathfrak{G}^{\xi+1}$ das ganze Borelsche System und alle G^η wären schon Mengen G^ξ. Es gilt aber auch für Limeszahlen η: wäre jedes G^η ein $G^\xi (\xi < \eta)$, also ein $F^{\xi+1}$, so wäre jedes $G^{\eta+1}$ ein F^η, jedes $F^{\eta+1}$ ein G^η und jedes $G^{\eta+2}$ ein G^η. Damit ist der Existenzsatz bewiesen.

[92] § 34. Kriterien für Borelsche Mengen.

1. Notwendige Bedingungen. Der Existenzsatz hat uns belehrt, daß es in einem geeigneten Raume Suslinsche Mengen S gibt, die keine Borelschen Mengen B sind. Wir müssen demgemäß nach Kriterien dafür fragen, daß ein S ein B sei. Zwei Bedingungen sind dafür aufgestellt worden: die erste von M. Suslin, daß das Komplement $E - S$ wieder ein S sei, die zweite von N. Lusin, daß S *mit disjunkten Summanden darstellbar* sei, daß es nämlich eine (d. h. unter den verschiedenen möglichen Darstellungen wenigstens eine) Darstellung

$$S = \Sigma\, F_{n_1} F_{n_1 n_2} \ldots$$

gebe, bei der die zu verschiedenen Zahlenfolgen $(n_1, n_2, \ldots)$ und $(\nu_1, \nu_2, \ldots)$ gehörigen Summanden $F_{n_1} F_{n_1 n_2} \ldots$ und $F_{\nu_1} F_{\nu_1 \nu_2} \ldots$ stets disjunkt sind (weswegen wir das Summenzeichen Σ statt $\mathfrak{S}$ gesetzt haben). Wir zeigen zunächst, daß beide Bedingungen *notwendig* sind, gleichviel wie der Raum beschaffen sei.

I. *Wenn die Suslinsche Menge S eine Borelsche ist, so ist ihr Komplement $E - S$ wieder eine Suslinsche Menge.*

Denn $E - S$ ist ja sogar eine Borelsche Menge. Wir haben diese evidente Tatsache schon im Existenzsatz verwendet.

II. *Wenn die Suslinsche Menge S eine Borelsche ist, so ist sie mit disjunkten Summanden darstellbar.*

Der Beweis beruht auf folgenden Schlüssen, bei denen wir der Kürze halber eine mit disjunkten Summanden darstellbare Suslinsche Menge als L-Menge und eine Menge A, die nebst ihrem Komplement $E - A$ eine L-Menge ist, als B-Menge bezeichnen wollen.

(A) *Die Summe einer Folge disjunkter L-Mengen ist eine L-Menge.*

Man kann diese L-Mengen in der Form

$$A_{n_1} = \sum_{n_2 n_3 \ldots} F_{n_1 n_2} F_{n_1 n_2 n_3} \ldots$$

darstellen; setzt man noch $F_{n_1} = E$ (der ganze Raum), so ist

$$\sum_{n_1} A_{n_1} = \sum_{n_1 n_2 n_3 \ldots} F_{n_1} F_{n_1 n_2} F_{n_1 n_2 n_3} \ldots$$

eine L-Menge.

(B) *Der Durchschnitt einer Folge von L-Mengen ist eine L-Menge.*

Es seien

$$\begin{aligned} A &= \Sigma A_{a_1} A_{a_1 a_2} A_{a_1 a_2 a_3} \ldots \\ B &= \Sigma B_{b_1} B_{b_1 b_2} B_{b_1 b_2 b_3} \ldots \\ C &= \Sigma C_{c_1} C_{c_1 c_2} C_{c_1 c_2 c_3} \ldots \\ &\ldots \end{aligned}$$

abzählbar viele L-Mengen (A_{a_1} usw. abgeschlossen). Man vereinige, etwa nach dem Diagonalschema, die sämtlichen Indizesfolgen zu einer einzigen

$$\begin{aligned} & n_1\ n_2\ n_3\ n_4\ n_5\ n_6 \ldots \\ = {} & a_1\ a_2\ b_1\ a_3\ b_2\ c_1 \ldots, \end{aligned}$$

dann wird

$$\begin{aligned} A\,B\,C \ldots &= \Sigma A_{n_1} A_{n_1 n_2} B_{n_3} A_{n_1 n_2 n_4} B_{n_3 n_5} C_{n_6} \ldots \\ &= \Sigma F_{n_1} F_{n_1 n_2} F_{n_1 n_2 n_3} \ldots, \end{aligned}$$

wo $F_{n_1 \ldots n_k}$ diejenige von den darüberstehenden Mengen bedeutet, deren höchster Index n_k ist (so daß $F_{n_1 \ldots n_k}$ nur von $n_1, \ldots, n_k$, wenn auch nicht von allen diesen Zahlen abhängt, z. B. $F_{n_1 n_2 n_3} = B_{n_3}$). Wir haben $A\,B\,C \ldots$ damit als L-Menge dargestellt. Natürlich ist auch der Durchschnitt endlich vieler L-Mengen eine L-Menge.

(C) *Summe und Durchschnitt einer Folge von B-Mengen ist eine B-Menge.*

Die Mengen $A_1, A_2, \ldots$ und ihre Komplemente $B_1, B_2, \ldots$ seien L-Mengen. Nach (B) ist der Durchschnitt $A_1 A_2 \ldots$, unter Zuziehung von (A) aber auch die Summe

$$A_1 \dot{+} A_2 \dot{+} A_3 \dot{+} \cdots = A_1 + B_1 A_2 + B_1 B_2 A_3 + \cdots$$

eine L-Menge, und zwar eine B-Menge, da von den B_n dasselbe gilt.

[93] (D) *Die Mengen F_σ sind L-Mengen.*

Eine Menge F_σ

$$A = F_1 \dot{+} F_2 \dot{+} F_3 \dot{+} \cdots$$

kann mit aufsteigenden abgeschlossenen Summanden angenommen werden, dann ist

$$A = F_1 + (F_2 - F_1) + (F_3 - F_2) + \cdots$$

mit disjunkten Summanden, die als Differenzen abgeschlossener Mengen wieder (spezielle) Mengen F_σ sind. So ergibt sich

$$A = \underset{n_1}{\mathfrak{S}} F_{n_1} = \sum_{n_1} A_{n_1}, \quad A_{n_1} = F_{n_1} - F_{n_1-1},$$

$$A_{n_1} = \underset{n_2}{\mathfrak{S}} F_{n_1 n_2} = \sum_{n_2} A_{n_1 n_2}, \quad A_{n_1 n_2} = F_{n_1 n_2} - F_{n_1,\, n_2-1}$$

usw. ($F_0 = F_{n_1 0} = \cdots = 0$). Wegen

$$F_{n_1} \geqq A_{n_1} \geqq F_{n_1 n_2} \geqq A_{n_1 n_2} \geqq \cdots$$

ist dann
$$A = \Sigma A_{n_1} A_{n_1 n_2} \ldots = \Sigma F_{n_1} F_{n_1 n_2} \ldots$$

als L-Menge dargestellt.

Hierzu werde noch die augenblicklich überflüssige, später zu verwendende Bemerkung gemacht: wenn der Raum separabel ist, kann die Darstellung von A mit disjunkten Summanden so eingerichtet werden, daß die Durchmesser der Mengen $F_{n_1 \ldots n_k}$ mit $k \to \infty$ nach 0 konvergieren. Denn E kann als Summe einer Folge abgeschlossener Mengen mit beliebig kleinen Durchmessern $\leqq \delta$ dargestellt werden (z. B. von abgeschlossenen Kugeln mit Radien $\frac{1}{2}\delta$, deren Mittelpunkte eine in E dichte Menge bilden); das gleiche gilt von jeder abgeschlossenen Menge und jeder Menge F_σ. Schreibt man demnach

$$A = V_1 \dot{+} V_2 \dot{+} \cdots, \quad F_n = V_1 \dot{+} \cdots \dot{+} V_n,$$

wo die V_n abgeschlossen sind und Durchmesser $\leqq \delta$ haben, so hat auch $A_n = F_n - F_{n-1} \leqq V_n$ einen Durchmesser $\leqq \delta$. Setzt man dies fort, so kann man also etwa erreichen, daß die Mengen $A_{n_1 \ldots n_k}$ Durchmesser $< \frac{1}{k}$ bekommen. Der Durchmesser der etwa auftretenden Nullmengen ist natürlich $= 0$ zu setzen.

Nach (D) sind insbesondere die abgeschlossenen und die offenen Mengen L-Mengen, d. h. B-Mengen, und da nach (C) die B-Mengen ein Borelsches System bilden, so sind die Borelschen Mengen des Raumes sicherlich B-Mengen, womit II bewiesen ist.

Wesentlich tiefer liegt der Beweis, daß die beiden Bedingungen unter Umständen auch *hinreichend* sind. Wir müssen dazu eine merkwürdige Darstellung der Suslinschen Mengen als Summen und Durchschnitte von $\aleph_1$ Borelschen Mengen vorausschicken.

2. Die Indizes. Es sei [94]

(1) $$A = \mathfrak{S}\, F(n_1)\, F(n_1, n_2)\, F(n_1, n_2, n_3) \ldots$$

eine Suslinsche Menge des Raumes E, $B = E - A$ ihr Komplement. Hierbei ist also jedem endlichen Komplex natürlicher Zahlen, den wir abkürzend mit

(2) $$r = (n_1, n_2, \ldots, n_k)$$

bezeichnen, eine abgeschlossene Menge

(3) $$F(r) = F(n_1, n_2, \ldots, n_k)$$

zugeordnet, und die Summe in (1) erstreckt sich über alle Folgen $(n_1, n_2, \ldots)$ natürlicher Zahlen. Wir dürfen dabei wie schon früher

(4) $$F(n_1) \geqq F(n_1, n_2) \geqq F(n_1, n_2, n_3) \geqq \cdots$$

voraussetzen.

Wenn wir an den Komplex (2) eine weitere natürliche Zahl anhängen, so entsteht ein Komplex

$$(r, n) = (n_1, \ldots, n_k, n),$$

den wir *einen Nachfolger* von r nennen; r heißt *der Vorgänger* von (r, n).

Sei R_0 eine beliebige Menge von Komplexen r. Wenn r, aber kein Nachfolger von r zu R_0 gehört, heiße r ein *Endelement* von R_0. Durch Weglassung der Endelemente entsteht aus R_0 eine neue Komplexmenge R_1, und indem wir diese Streichung der Endelemente wiederholen, gelangen wir zu einer absteigenden Folge

$$R_0, R_1, R_2, \ldots, R_\omega, R_{\omega+1}, \ldots$$

von Komplexmengen, die induktiv dadurch definiert sind, daß $R_{\xi+1}$ aus R_ξ durch Tilgung der Endelemente entsteht und für eine Limeszahl η

$$R_\eta = \underset{\xi<\eta}{\mathfrak{D}}\, R_\xi.$$

Da R_0 höchstens abzählbar ist, können von den disjunkten Mengen $R_\xi - R_{\xi+1}$ (dies ist die Menge der Endelemente von R_ξ) höchstens abzählbar viele von Null verschieden sein und wir gelangen für einen *ersten* Index η $(0 \leqq \eta < \Omega)$ zu einer Menge $R_\eta = R_{\eta+1}$, die also keine Endelemente mehr hat, aber natürlich Null sein kann. Nennen wir dies η den *Index* der Menge R_0; für $\xi < \eta$ ist also noch $R_\xi > R_{\xi+1}$. Die Menge $R = R_\eta = R_{\eta+1} = \cdots$ werde etwa der *Kern* von R_0 genannt (er ist die größte Menge $\leqq R_0$, die keine Endelemente hat). Prozesse dieser Art sind uns aus § 30 geläufig; der gegenwärtige hat große Ähnlichkeit mit der Abspaltung der isolierten Punkte und der Bildung der kleinsten Kohärenz.

Nun bestimmt jeder Punkt x des Raumes die Menge $R_0(x)$ derjenigen r, für die $x \,\varepsilon\, F(r)$, und wenn wir mit dieser Menge den eben geschilderten Prozeß ausführen, so gelangen wir zu den Mengen $R_\xi(x)$, endigend mit dem Kern $R(x) = R_\eta(x) = R_{\eta+1}(x) = \cdots$; der von x abhängige Index $\eta = \eta(x)$

der Menge $R_0(x)$ werde auch der *Index des Punktes* x genannt. Wir haben also für die Mengen A, B eine Spaltung nach den Indizes ihrer Punkte:

(5) $$A = \sum_\eta A_\eta, \quad B = \sum_\eta B_\eta \quad (\eta < \Omega),$$

wo A_η, B_η die Mengen der Punkte von A, B sind, deren Index $\eta(x) = \eta$. Wir werden nun sehen, daß diese Mengen Borelsche Mengen sind.

Jede Menge $R_\xi(x)$ *enthält die Vorgänger ihrer Elemente*, oder

(6) $$\text{mit } (r, n) \,\varepsilon\, R_\xi(x) \text{ ist } r \,\varepsilon\, R_\xi(x)\,.$$

Dies gilt nämlich für $\xi = 0$, als andere Form der Aussage $F(r, n) \leqq F(r)$ oder (4). Wenn (6) aber für ξ gilt, so auch für $\xi + 1$; denn wenn (r, n) in $R_{\xi+1}(x)$ und also erst recht in $R_\xi(x)$ vorkommt, so ist r Element, aber nicht Endelement von $R_\xi(x)$ und bleibt also Element in $R_{\xi+1}(x)$. Endlich, nach der Definition von $R_\eta(x)$ für eine Limeszahl η, gilt (6) für η, wenn es für alle $\xi < \eta$ gilt; womit (6) induktiv für alle ξ bewiesen ist.

Wir hatten von den $F(r)$ aus die Komplexmengen $R_0(x)$ und daraus durch Abspaltung der Endelemente die $R_\xi(x)$ definiert; nun drehen wir wieder um und nennen $F_\xi(r)$ die Menge der x, für welche $r \,\varepsilon\, R_\xi(x)$, so daß die Relationen

(7) $$r \,\varepsilon\, R_\xi(x), \quad x \,\varepsilon\, F_\xi(r)$$

gleichbedeutend sind. Wie die R_ξ nehmen auch die F_ξ mit wachsendem Index ab, denn das heißt ja, daß (7) für ξ gilt, wenn es für $\eta > \xi$ gilt. Beiläufig werden aber im Allgemeinen die absteigenden Mengen

$$F_0(r) \geqq F_1(r) \geqq \cdots \qquad \geqq F_\omega(r) \geqq \cdots$$

nicht etwa schließlich identisch, denn die Relation $R_\eta(x) = R_{\eta+1}(x)$ gilt ja nicht „gleichmäßig" für irgendein festes η, sondern für ein von x abhängiges $\eta = \eta(x)$ oder $\eta \geqq \eta(x)$. Aus (6) folgt ferner

(8) $$F_\xi(r, n) \leqq F_\xi(r)\,.$$

Die Mengen F_ξ stellen sich nun folgendermaßen dar:

(9) $$\begin{cases} F_0(r) = F(r) \\ F_{\xi+1}(r) = \mathfrak{S}_n\, F_\xi(r, n) \\ F_\eta(r) = \mathfrak{D}_{\xi<\eta}\, F_\xi(r) \quad (\eta \text{ Limeszahl}). \end{cases}$$

Hiervon entspringt die erste Gleichung aus der Definition von $R_0(x)$, die dritte aus der von $R_\eta(x)$, wonach (7) dann und nur dann für η gilt, wenn es für alle $\xi < \eta$ gilt. Die mittlere Gleichung (9) folgt so: $x \,\varepsilon\, F_{\xi+1}(r)$ oder $r \,\varepsilon\, R_{\xi+1}(x)$ bedeutet, daß r Element, aber nicht Endelement von $R_\xi(x)$ ist, also einen Nachfolger $(r, n) \,\varepsilon\, R_\xi(x)$ hat; oder, daß ein $(r, n) \,\varepsilon\, R_\xi(x)$ vorhanden ist (womit die Aussage $r \,\varepsilon\, R_\xi(x)$ nach (6) überflüssig wird); oder, daß x einer der Mengen $F_\xi(r, n)$ angehört.

Weiter ist

(10) $$S_\xi = \mathfrak{S}_r\, F_\xi(r)$$

die Menge der Punkte x mit $R_\xi(x) > 0$; denn das heißt ja, daß es ein r gibt, wofür (7) gilt. Und es ist

$$T_\xi = \underset{r}{\mathfrak{S}}[F_\xi(r) - F_{\xi+1}(r)] \tag{11}$$

die Menge der Punkte x mit $R_\xi(x) > R_{\xi+1}(x)$; denn dies bedeutet, daß es ein r gibt, das Endelement von $R_\xi(x)$ ist, wofür also

$$r \,\varepsilon\, R_\xi(x), \quad r \,\bar{\varepsilon}\, R_{\xi+1}(x),$$
$$\text{d. h.} \qquad x \,\varepsilon\, F_\xi(r), \quad x \,\bar{\varepsilon}\, F_{\xi+1}(r)$$

oder $x \,\varepsilon\, F_\xi(r) - F_{\xi+1}(r)$.

Für diese Mengen gilt aber nun

$$\left\{\begin{aligned} S_\xi &= A + \sum_{\eta>\xi} B_\eta \\ T_\xi &= \sum_{\eta>\xi} A_\eta + \sum_{\eta>\xi} B_\eta . \end{aligned}\right. \tag{12}$$

Zunächst ist nämlich der Kern $R(x)$ von $R_0(x)$ dann und nur dann > 0, wenn $x \,\varepsilon\, A$. Ist $x \,\varepsilon\, A$ und etwa $x \,\varepsilon\, F(n_1)\, F(n_1, n_2) \ldots$, so gehören die sämtlichen Komplexe (n_1), (n_1, n_2), (n_1, n_2, n_3), ... zu $R_0(x)$; sie sind nicht Endelemente und gehören also zu $R_1(x)$; so weiterschließend erkennt man, daß sie zu jedem $R_\xi(x)$, also zum Kern $R(x)$ gehören, der demnach > 0 ist. Wenn umgekehrt der Kern $R(x) > 0$ ist und daher nach (6) einen eingliedrigen Komplex (n_1) enthält, so enthält er, als Menge ohne Endelemente, auch einen zweigliedrigen (n_1, n_2), einen dreigliedrigen (n_1, n_2, n_3) usf., alle diese Komplexe gehören auch zu $R_0(x)$, d. h.

$$x \,\varepsilon\, F(n_1)\, F(n_1, n_2)\, F(n_1, n_2, n_3) \ldots \leqq A .$$

Für $x \,\varepsilon\, B$ und nur in diesem Falle ist $R(x) = 0$.

Nach der Definition der Indizes $\eta(x)$ ist nun

$$R_\xi(x) > R_{\xi+1}(x) \quad \text{mit} \quad \xi < \eta(x),$$
$$R_\xi(x) = R_{\xi+1}(x) = R(x) \quad \text{mit} \quad \xi \geqq \eta(x)$$

gleichbedeutend. Danach ist T_ξ die Menge der Punkte, deren Index $> \xi$ ist, während S_ξ (durch $R_\xi(x) > 0$ definiert) genau alle Punkte von A und diejenigen Punkte von B, deren Index $> \xi$, enthält. Damit ist (12) bewiesen. Man erhält daraus noch

$$\left\{\begin{aligned} S_\xi - T_\xi &= A_0 + A_1 + \cdots + A_\xi \\ E - S_\xi &= B_0 + B_1 + \cdots + B_\xi . \end{aligned}\right. \tag{13}$$

Nun ergibt sich aus (9), von den abgeschlossenen Mengen $F_0(r)$ ausgehend, durch Induktion, daß alle Mengen $F_\xi(r)$ Borelsche Mengen des Raumes sind, aus (10) und (11), daß die S_ξ, T_ξ Borelsche Mengen sind; dasselbe folgt für die A_ξ, B_ξ, etwa aus (13) durch Induktion. Nach (5) ist also *die Suslinsche Menge A und ihr Komplement B als Summe von* $\aleph_1$

Borelschen Mengen dargestellt[1]; beide Mengen sind dann auch Durchschnitte von $\aleph_1$ Borelschen Mengen, insbesondere nach (12) einfach

$$(14) \qquad A = \mathfrak{D}_{\xi} S_{\xi}.$$

Die S_ξ und T_ξ nehmen mit wachsendem Index ab, die Mengen (13) nehmen zu; der Durchschnitt aller T_ξ ist 0. Wir bemerken noch, daß man für (10) auf Grund von (8) auch

$$(15) \qquad S_\xi = \mathfrak{S}_{n} F_\xi(n)$$

schreiben, also die sämtlichen Komplexe r durch die eingliedrigen ersetzen kann; für T_ξ ist dies nicht zulässig. Überdies ist

$$(16) \qquad T_{\xi+\omega} \leqq \mathfrak{S}_{r\varrho} F_\xi(r) F_\xi(\varrho),$$

wo $\xi + \omega$ die erste auf $\xi, \xi + 1, \xi + 2, \ldots$ folgende Ordnungszahl und die Summe rechts über alle Paare *verschiedener* Komplexe *gleicher* Zifferzahl

$$(17) \qquad r = (n_1, \ldots, n_k), \quad \varrho = (\nu_1, \ldots, \nu_k) \quad (r \neq \varrho)$$

zu erstrecken ist. Sei in der Tat $x \varepsilon T_{\xi+\omega}$, also

$$(18) \qquad x \varepsilon F_{\xi+\omega}(r) - F_{\xi+\omega+1}(r)$$

für ein gewisses r. Für $m = 0, 1, 2, \ldots$ ist demnach wegen (9)

$$x \varepsilon F_{\xi+m+1}(r), \quad x \varepsilon F_{\xi+m}(r, n_m)$$

mit geeignetem n_m; diese Zahlen $n_0, n_1, \ldots$ können aber nicht sämtlich gleich sein, denn aus $x \varepsilon F_{\xi+m}(r, n)$ bei festem n würde folgen $x \varepsilon F_{\xi+\omega}(r, n) \leqq F_{\xi+\omega+1}(r)$ im Widerspruch zu (18). Demnach gibt es mindestens zwei Zahlen m, μ mit $n_m \neq n_\mu$ und es ist

$$x \varepsilon F_{\xi+m}(r, n_m) F_{\xi+\mu}(r, n_\mu) \leqq F_\xi(r, n_m) F_\xi(r, n_\mu),$$

x gehört also der in (16) rechtsstehenden Menge an.

3. Hinreichende Bedingungen[2]. Schicken wir zwei Hilfsbetrachtungen voraus.

(α) Sind jedem $\xi < \Omega$ abzählbar viele Mengen D_ξ^n (etwa für $n = 1, 2, \ldots$) zugeordnet, die mit wachsendem ξ abnehmen ($D_\xi^n \geqq D_\eta^n$ für $\xi < \eta$) und ist für jedes ξ $\quad \mathfrak{S}_{n} D_\xi^n > 0$,

so gibt es ein n derart, daß für jedes ξ auch $D_\xi^n > 0$.

Denn zu jedem ξ gibt es zunächst ein $n(\xi)$ mit $D_\xi^{n(\xi)} > 0$. Die Funktion $n(\xi)$, die nur abzählbar vieler Werte fähig ist, muß mindestens einen Wert n unabzählbar oft annehmen; dann ist $D_\xi^n > 0$ für unabzählbar viele ξ, also (wegen der monotonen Abnahme mit wachsendem ξ) für alle ξ.

---

[95] [1] Im separablen vollständigen Raume schließt man daraus, daß die Mengen B, die Komplemente Suslinscher Mengen, entweder höchstens abzählbar oder von einer der Mächtigkeiten $\aleph_1$, $\aleph$ sind: das ist wegen des ungelösten Kontinuumproblems ein weniger präzises Resultat als der für die Suslinschen Mengen selbst gültige Satz § 32, I.

[2] Hierzu vgl. § 46, 1.

(β) Ist der Raum separabel, so kann in der Darstellung (1) der Suslinschen Mengen angenommen werden, daß die Durchmesser der Mengen $F(n_1, \ldots, n_k)$ mit $k \to \infty$ nach 0 konvergieren. (Der Durchmesser der Nullmenge ist natürlich $= 0$ zu setzen.)

Wir hatten oben (S. 186) bei Punkt (D) bemerkt, daß die Mengen F_σ, also insbesondere der Raum selbst, einer solchen Darstellung und zwar mit disjunkten Summanden fähig sind. Setzt man demgemäß

$$E = \Sigma E(m_1)\, E(m_1, m_2) \ldots,$$

wo die $E(m_1, \ldots, m_k)$ abgeschlossen sind und mit $k \to \infty$ nach 0 konvergente Durchmesser haben [1]), bringt dies mit der Menge (1) zum Durchschnitt, ordnet den Zahlenpaaren (m_k, n_k) die natürlichen Zahlen p_k eineindeutig zu und setzt

$$F_{p_1 \ldots p_k} = E(m_1, \ldots, m_k)\, F(n_1, \ldots, n_k),$$

so wird
$$A = \mathfrak{S}\, F_{p_1}\, F_{p_1 p_2} \ldots$$

eine Darstellung der verlangten Form, die übrigens, wie wir beachten wollen, disjunkte Summanden hat, sobald dies für (1) der Fall ist. Auch die Nebenbedingung $F_{p_1} \geqq F_{p_1 p_2} \geqq \cdots$ läßt sich aufrechterhalten, wenn man $E(m_1) \geqq E(m_1, m_2) \geqq \cdots$ voraussetzt.

Dies vorausgeschickt, wollen wir nun die Sätze I, II umkehren.

III. *Ist der Raum vollständig und separabel, so ist eine Suslinsche Menge, deren Komplement eine Suslinsche Menge ist, eine Borelsche. Zwei disjunkte Suslinsche Mengen sind in ihrer Summe Borelsche Mengen.* [96]

Zum Beweise betrachten wir zwei Suslinsche Mengen $A, \bar{A}$ mit den entsprechenden Borelschen Mengen $F_\xi(r)$, $\bar{F}_\xi(r)$ und $S_\xi, \bar{S}_\xi$. Die Darstellungen (1) von A und $\bar{A}$ mögen die Durchmesserbedingung (β) erfüllen. Wir zeigen dann: *wenn* $S_\xi \bar{S}_\xi > 0$ *für jedes* ξ, *so ist* $A\bar{A} > 0$. [97]

In der Tat ist, nach (15), *für jedes* ξ (so auch im folgenden)

$$\mathfrak{S}_{n\nu}\, F_\xi(n)\, \bar{F}_\xi(\nu) > 0,$$

die Summe über alle Paare natürlicher Zahlen n, ν erstreckt. Nach der Hilfsbetrachtung (α) existiert ein Summand

$$F_\xi(n_1)\, \bar{F}_\xi(\nu_1) > 0.$$

Ersetzen wir hierin ξ durch $\xi + 1$, so ist

$$\mathfrak{S}_{n\nu}\, F_\xi(n_1, n)\, \bar{F}_\xi(\nu_1, \nu) > 0,$$

es existiert ein Summand

$$F_\xi(n_1, n_2)\, \bar{F}_\xi(\nu_1, \nu_2) > 0$$

[1]) Umgekehrt ist ein solcher Raum höchstens separabel; denn wählt man aus jeder Menge $E(m_1, \ldots, m_k) > 0$ einen Punkt, so ist die Menge dieser Punkte in E dicht.

usf. Wir erhalten also zwei Zahlenfolgen $(n_1, n_2, \ldots)$, $(\nu_1, \nu_2, \ldots)$ derart, daß für jedes k ($\xi = 0$ gesetzt)

$$F(n_1, \ldots, n_k)\, \bar{F}(\nu_1, \ldots, \nu_k) > 0\,.$$

Diese abgeschlossenen, mit wachsendem k abnehmenden Mengen mit Durchmessern $\to 0$ haben nach dem zweiten Durchschnittssatz einen (einzigen) gemeinsamen Durchschnittspunkt x, der sowohl zu A als auch zu $\bar{A}$ gehört; also $A\,\bar{A} > 0$.

Wenn, für jedes ξ, $S_\xi\,\bar{A} > 0$, so ist nach (12) oder (14) erst recht $S_\xi\,\bar{S}_\xi > 0$ und daher $A\,\bar{A} > 0$. Also umgekehrt:

Wenn $A\bar{A} = 0$, so muß einmal für ein ξ (und alle folgenden) $S_\xi\,\bar{A} = 0$ sein, also

$$A = A\,S_\xi = (A + \bar{A})\,S_\xi\,,$$

A ist der Durchschnitt von $A + \bar{A}$ mit einer Borelschen Menge. Zwei disjunkte Suslinsche Mengen sind also in ihrer Summe Borelsche Mengen. Insbesondere, wenn das Komplement $B = E - A$ einer Suslinschen Menge auch eine ist, so sind beide Borelsche Mengen (in $A + B = E$), und zwar von einem gewissen Index an $A = S_\xi$.

Damit ist III bewiesen; die Suslinsche Bedingung ist hinreichend, falls der Raum vollständig und separabel ist. Der Nachsatz von III zeigt übrigens, daß der Satz auch gilt, wenn der Raum eine Suslinsche Menge M in einem vollständigen separablen Raum E ist.

[98] Ebenso läßt sich das Lusinsche Kriterium als hinreichend erweisen:

[99] IV. *Ist der Raum vollständig und separabel, so ist jede mit disjunkten Summanden darstellbare Suslinsche Menge eine Borelsche.*

Nach (12) und (13) ist $S_\xi - T_\xi \leqq A \leqq S_\xi$; wenn $T_\xi = 0$, so ist $A = S_\xi$ eine Borelsche Menge. Ist also A keine Borelsche Menge, so ist $T_\xi > 0$ *für jedes* ξ (so auch im folgenden), also nach (16)

$$\mathfrak{S}_{r\varrho}\, F_\xi(r)\, F_\xi(\varrho) > 0\,,$$

die Summe über die (abzählbar vielen) Paare von verschiedenen Komplexen gleicher Zifferzahl erstreckt. Nach der Schlußweise (α) existiert ein Summand

$$F_\xi(r)\, F_\xi(\varrho) > 0$$

mit einem Komplexpaar (17). Ersetzt man ξ durch $\xi + 1$, so ist

$$\mathfrak{S}_{n\nu}\, F_\xi(r, n)\, F_\xi(\varrho, \nu) > 0$$

und es existiert ein Summand

$$F_\xi(r, n_{k+1})\, F_\xi(\varrho, \nu_{k+1}) > 0\,.$$

So fortfahrend erhält man zwei (wegen $r \neq \varrho$) *verschiedene* Zahlenfolgen $(n_1, n_2, \ldots)$ und $(\nu_1, \nu_2, \ldots)$ derart, daß ($\xi = 0$)

$$F(n_1, \ldots, n_h)\, F(\nu_1, \ldots, \nu_h) > 0$$

für jedes h; die Anwendung des zweiten Durchschnittsatzes lehrt, daß die beiden verschiedenen Summanden $F(n_1)\,F(n_1, n_2)\ldots$ und $F(\nu_1)\,F(\nu_1, \nu_2)\ldots$ denselben Punkt liefern. Wenn also A keine Borelsche Menge ist, so ist sie gewiß nicht mit disjunkten Summanden darstellbar (weder mit noch ohne Durchmesserbedingung, wie wir bei (β) bemerkt haben), und damit ist IV bewiesen.

Achtes Kapitel.

Abbildung zweier Räume.

§ 35. Stetige Abbildung.

1. Grundlagen.

Wie schon in § 2 gehen wir von einer vorgegebenen Menge geordneter Paare (x, y) aus; die Mengen der hierin auftretenden Elemente x, y seien A, B. Jedem $x \varepsilon A$ sind hierdurch ein oder mehrere *Bilder* (Bildpunkte) $y = \varphi(x)$ zugeordnet, nämlich diejenigen y, die mit x ein Paar (x, y) jener Paarmenge bilden; ebenso sind jedem $y \varepsilon B$ ein oder mehrere *Urbilder* $x = \psi(y)$ zugeordnet. Wir haben hiermit zwei zueinander inverse, im allgemeinen mehrdeutige Funktionen φ, ψ und eine Abbildung zwischen den Räumen A, B vor uns, wobei die durch die geordneten Paare hineingetragene und in den Bezeichnungen (Bild, Urbild) festgehaltene Asymmetrie zunächst unwesentlich ist. Von den Punktfunktionen φ, ψ gelangen wir auf natürlichem Wege zu Mengenfunktionen Φ, Ψ. Eine Menge $P \leqq A$ bestimmt ihr *Bild* (Bildmenge) $\Phi(P)$, nämlich die Menge aller Bilder aller Punkte von P, wobei $\Phi(0) = 0$ und $\Phi(A) = B$ ist, und eine Menge $Q \leqq B$ bestimmt in gleicher Weise ihr *Urbild* $\Psi(Q)$. Hat jedes x nur ein einziges Bild $y = \varphi(x)$, so heißt diese Funktion *eindeutig*, *B eindeutiges Bild von A*; sind beide Funktionen $y = \varphi(x)$, $x = \psi(y)$ eindeutig, so heißen sie *umkehrbar eindeutig* oder *eineindeutig* oder *schlicht*, *B schlichtes Bild von A*.

Eine leichte Überlegung lehrt:

$$
(1)\qquad \left\{\begin{array}{c}
\Psi(\Phi(P)) \geqq P \\
\Phi(P_1 \dotplus P_2 \dotplus \cdots) = \Phi(P_1) \dotplus \Phi(P_2) \dotplus \cdots \\
\Phi(P_1\, P_2 \cdots) \leqq \Phi(P_1)\, \Phi(P_2) \ldots
\end{array}\right.
$$

nebst den durch Vertauschung der beiden Funktionen entsprechenden Beziehungen; die Summen und Durchschnitte in (1) beziehen sich auf beliebig viele Mengen. Die Ungleichheitszeichen sind im allgemeinen nicht entbehrlich, denn — um etwa die erste Formel zu deuten — zu den Urbildern der Bilder der Punkte von P gehören zwar die Punkte von P selbst, aber im allgemeinen noch weitere Punkte. Wenn jedoch die

Funktion $y = \varphi(x)$ eindeutig ist, so gelten in den aus (1) durch Vertauschung hervorgehenden Formeln die Gleichheitszeichen:

$$(2) \qquad \begin{cases} \Phi(\Psi(Q)) = Q \\ \Psi(Q_1 \dotplus Q_2 \dotplus \cdots) = \Psi(Q_1) \dotplus \Psi(Q_2) \dotplus \cdots \\ \Psi(Q_1 Q_2 \ldots) = \Psi(Q_1)\, \Psi(Q_2) \ldots, \end{cases}$$

was man am einfachsten dadurch erkennt, daß die Menge A jetzt in disjunkte Teilmengen $\Psi(y)$, die Urbilder der einzelnen Punkte y von B, zerfällt. (Präziser wäre $\Psi(\{y\})$, Urbild der einpunktigen Menge $\{y\}$, zu schreiben.) *Wir betrachten künftig nur den Fall, daß $\varphi(x)$ eindeutig ist,* ohne die inverse Funktion als eindeutig vorauszusetzen; damit ist also wirkliche Asymmetrie zwischen A, B eingetreten.

Nun seien A, B metrische Räume. *Die Funktion $\varphi(x)$ heißt im Punkte x stetig, wenn für jede Folge $x_n \to x$ auch $\varphi(x_n) \to \varphi(x)$.*

In einem isolierten Punkt von A ist jede Funktion stetig. Eine stetige Funktion von einer stetigen Funktion ist wieder stetig. Ausführlicher: $y = \varphi(x)$ und $z = \chi(y)$ seien eindeutige Funktionen, wobei x, y, z die Mengen A, B, C durchlaufen, und $\varphi(x)$ sei in x_0, $\chi(y)$ in $y_0 = \varphi(x_0)$ stetig: dann ist $\chi(\varphi(x))$ in x_0 stetig. Der Beweis ist evident: aus $x_n \to x_0$ folgt $y_n \to y_0$, daraus $z_n \to z_0$.

Eine in jedem Punkte von A stetige Funktion heißt in A stetig oder (schlechthin) *stetig*; in diesem Falle heißt B *stetiges Bild* von A.

Eine in A stetige Funktion ist bekannt, wenn man die Bilder der Punkte einer in A dichten Menge R kennt. Denn jedes $x \varepsilon A$ ist als $x = \lim r_n$ $(r_n \varepsilon R)$ darstellbar, also $\varphi(x) = \lim \varphi(r_n)$. Mit einer leichtverständlichen Bezeichnung kann man sagen: die Gesamtfunktion $\varphi(x) = \varphi(x \mid A)$ ist durch die *Teilfunktion* $\varphi(x \mid R)$ bestimmt. Daß dies die Mächtigkeit des Systems der stetigen Funktionen gegenüber dem aller Funktionen herabdrücken kann, ist uns bekannt (S. 41).

Die Stetigkeit an der Stelle a besagt, daß mit der Entfernung $a\,x$ auch $b\,y$ nach 0 konvergieren soll, wobei $b = \varphi(a)$, $y = \varphi(x)$. Dies läßt sich nach einer aus den Elementen bekannten Schlußweise auch so ausdrücken: zu jedem (beliebig kleinen) $\sigma > 0$ läßt sich ein (hinlänglich kleines) $\varrho > 0$ angeben derart, daß mit $a\,x < \varrho$ auch $b\,y < \sigma$. Oder: *jeder Umgebung V_b entspricht eine Umgebung U_a, deren Bild in V_b liegt*; hierbei sind, jetzt wie künftig (wenn nicht ausdrücklich das Gegenteil gesagt ist), die U_a auf den Raum A, die V_b auf den Raum B bezogen. Endlich kann man auch sagen: *das Urbild jeder Umgebung V_b hat a als inneren Punkt.* Daraus folgt:

I. *B ist dann und nur dann stetiges Bild von A, wenn jede in B offene (abgeschlossene) Menge als Urbild eine in A offene (abgeschlossene) Menge hat.*

Der Zusatz, den die eingeklammerten Worte aussprechen, wird durch Komplementbildung bewiesen, weil $\Psi(B-Q) = A - \Psi(Q)$.

Ein spezieller Fall des Satzes ist uns aus § 22, I bekannt; ist $\varphi(x)$ eine reelle stetige Funktion, B also eine Menge reeller Zahlen, so ist die dort mit $[\varphi > 0]$ bezeichnete Menge ja nichts anderes als das Urbild der Menge Q, die der Durchschnitt von B mit der Halbgeraden $y > 0$, also eine in B offene Menge ist.

Aus I folgt wegen der Formeln (2) noch mehr:

II. *Ist B stetiges Bild von A, so haben die Borelschen und Suslinschen Mengen des Raumes B als Urbilder Borelsche und Suslinsche Mengen des Raumes A, und zwar entsprechen den Mengen $F^\xi(B)$, $G^\xi(B)$ als Urbilder Mengen $F^\xi(A)$, $G^\xi(A)$.*

In der Tat geht dies eben daraus hervor, daß die Urbilder von Summen und Durchschnitten die Summen und Durchschnitte der Urbilder sind. Hat man also z. B. schon bewiesen, daß für $\xi < \eta$ jedes $F^\xi(B)$ als Urbild ein $F^\xi(A)$ hat, so ist $F^\eta(B)$ Summe oder Durchschnitt (je nachdem η ungerade oder gerade ist) von abzählbar vielen $F^\xi(B)$, und sein Urbild Summe oder Durchschnitt von abzählbar vielen $F^\xi(A)$, d. h. ein $F^\eta(A)$. Ebenso ist für die G^ξ und die Suslinschen Mengen zu verfahren.

Wenn B stetiges Bild von A, also die Funktion $\varphi(x)$ (eindeutig und) stetig ist, so kann die inverse Funktion $\psi(y)$ mehrdeutig sein. Ist sie eindeutig, so wird B als *schlichtes stetiges Bild* von A zu bezeichnen sein. Die einfachsten Beispiele lehren, daß $\psi(y)$ dann noch nicht stetig zu sein braucht. Wenn die Menge B der rationalen Zahlen schlicht auf die Menge A der natürlichen Zahlen abgebildet wird, so ist B stetiges Bild von A (weil jeder Punkt von A isoliert ist), während die inverse Funktion $x = \psi(y)$ an keiner Stelle stetig ist. *Allgemein kann jede beliebige Menge B als schlichtes stetiges Bild einer isolierten, sogar einer absolut abgeschlossenen* [100]
Menge A aufgefaßt werden: man braucht ja nur eine mit B äquivalente Menge A zu nehmen und je zwei verschiedenen Punkten von A die Entfernung 1 zuzuschreiben. Die Funktion $x = \psi(y)$ ist aber nur in den isolierten Punkten von B stetig, also nirgends, wenn B insichdicht ist. — Noch ein anderes Beispiel: in der Euklidischen Ebene der Punkte $x = (x_1, x_2)$ sei A eine Menge, die von jeder Geraden $x_1 = \text{const.}$ in genau einem Punkte geschnitten wird, d. h. die durch $x_2 = f(x_1)$ mit eindeutiger, sonst völlig willkürlicher Funktion f dargestellte „Kurve". B sei die senkrechte Projektion von A auf die Gerade $x_2 = 0$, d. h. jedem $x = (x_1, x_2)$ entspreche $y = (x_1, 0) = \varphi(x)$. Dann ist B schlichtes stetiges Bild von A; das Umgekehrte gilt nur, wenn $f(x_1)$ stetig ist.

Wenn aber beide Funktionen $y = \varphi(x)$, $x = \psi(y)$ eindeutig und stetig sind, so wird jede von ihnen als *umkehrbar stetig* oder beiderseits stetig oder doppeltstetig bezeichnet (fonction bicontinue). Man nennt auch B *homöo-*

morphes Bild von A (ebenso A von B) und die zwischen beiden Mengen bestehende schlichte Abbildung eine *Homöomorphie*; ferner sagt man, daß A und B zueinander homöomorph seien, was wir durch das Zeichen

$$B \approx A \quad \text{oder} \quad A \approx B$$

ausdrücken wollen. Um die Begriffe also nochmals zusammenzustellen, so hat man je nach der Beschaffenheit der Funktion $x = \psi(y)$ folgende Fälle ($y = \varphi(x)$ als stetig vorausgesetzt):

B stetiges Bild von A: $\psi(y)$ kann mehrdeutig sein.

B schlichtes stetiges Bild von A: $\psi(y)$ eindeutig.

B homöomorphes Bild von A: $\psi(y)$ stetig.

Der dritte Fall ist als Sonderfall des zweiten, dieser als Sonderfall des ersten anzusehen. Im dritten Fall ist die Asymmetrie zwischen A, B wieder verschwunden.

[101] **2. In sich kompakte Mengen.** Die Sätze I, II gelten nur für die Urbilder, nicht für die Bilder. Ist B stetiges Bild von A, so braucht das Bild einer in A abgeschlossenen Menge keine in B abgeschlossene zu sein; wenn A isoliert ist und demnach jede mit A äquivalente Menge B als schlichtes stetiges Bild von A angesehen werden kann, so ist jede Teilmenge von A in A abgeschlossen, *jede* Teilmenge von B also Bild einer in A abgeschlossenen. Wir sahen, daß sogar das schlichte stetige Bild einer *absolut abgeschlossenen* (vollständigen) Menge A noch ganz beliebig sein kann, bis auf die Äquivalenz mit A. Dagegen gilt ein einfaches und präzises Resultat, wenn A eine in sich kompakte (also kompakte, absolut abgeschlossene) Menge ist:

III. *Das stetige Bild B einer in sich kompakten Menge A ist wieder in sich kompakt; ist die Abbildung schlicht, so ist sie eine Homöomorphie.*

Die Voraussetzung besagt, daß jede Punktfolge aus A eine konvergente Teilfolge hat, deren Limes zu A gehört. Dasselbe ist für B zu zeigen. Ist $y_n \varepsilon B$, $x_n = \psi(y_n)$ ein Urbild von y_n, so gibt es eine konvergente Teilfolge $x_p \to x \varepsilon A$, demnach $\varphi(x_p) \to \varphi(x)$ oder $y_p \to y \varepsilon B$, B ist in sich kompakt. Der zweite Teil der Behauptung folgt nun am einfachsten so: jede in A abgeschlossene Menge P ist in sich kompakt, ihr Bild $Q = \Phi(P)$ in sich kompakt, also absolut und daher in B abgeschlossen, dies aber besagt nach I, daß die als eindeutig vorausgesetzte Funktion $x = \psi(y)$ stetig ist.

Im Euklidischen Raume ist eine Menge A dann und nur dann in sich kompakt, wenn sie beschränkt und abgeschlossen ist; ihr — wieder in einem Euklidischen Raum angenommenes — stetiges Bild B ist wieder beschränkt und abgeschlossen. Insbesondere durchläuft eine in A stetige reelle Funktion eine abgeschlossene und beschränkte Zahlenmenge, hat also einen größten und kleinsten Wert (d. h. die obere Grenze ist ein wirklich erreichtes Maximum, die untere ein Minimum).

Nennt man eine in sich kompakte Menge K, die Summe einer Folge solcher Mengen K_σ, so ist das stetige Bild eines K wieder ein K, das eines K_σ wieder ein K_σ. (Man vergleiche die mittlere Formel (1)). Der Euklidische Raum ist ein K_σ (z. B. Summe konzentrischer abgeschlossener Kugeln mit Radien $n = 1, 2, 3, \ldots$), also auch jede in ihm abgeschlossene Menge; deren stetiges Bild ist also ein K_σ, braucht aber, wenn in einem Euklidischen Raum gelegen, nicht abgeschlossen zu sein. Jede eindimensionale Menge F_σ, $B = B_1 \dotplus B_2 \dotplus \ldots$, ist Projektion einer ebenen abgeschlossenen Menge $A = A_1 \dotplus A_2 \dotplus \cdots$; man lasse einfach dem Punkte $(x_1, 0)$ von B_n den Punkt (x_1, n) von A_n entsprechen, d. h. übertrage die Menge B_n von der Geraden $x_2 = 0$ auf $x_2 = n$. Auch das stetige Bild eines Euklidischen F_σ ist ein K_σ, insbesondere das einer offenen Menge. Es sei bemerkt, daß aber der Hilbertsche Raum kein K_σ, vielmehr in ihm jedes K_σ von erster Kategorie ist, während er selbst, als vollständiger Raum, in sich von zweiter Kategorie ist. Denn jede Umgebung des Nullpunktes $(0, 0, 0, \ldots)$ enthält eine Menge von Punkten $(\varrho, 0, 0, \ldots)$, $(0, \varrho, 0, \ldots)$, $\ldots$, die paarweise die Entfernung $\sqrt{2}\,\varrho$ haben und also keinem K angehören; dasselbe gilt von jedem Punkt, und das Komplement eines K ist also im Hilbertschen Raume dicht, K nirgendsdicht.

Es sei hier der Begriff der gleichmäßigen Stetigkeit erwähnt; die eindeutige Funktion $y = \varphi(x)$ heißt in A *gleichmäßig stetig*, wenn für jede Folge von Punktpaaren mit $x_n \xi_n \to 0$ zugleich $y_n \eta_n \to 0$. (Bei festem $\xi_n = \xi$ ist das die Stetigkeitsforderung für die Stelle ξ). Diese Bedingung ist, wie in üblicher Weise zu erkennen, damit gleichbedeutend: jedem $\sigma > 0$ entspricht ein (nur von σ abhängiges) $\varrho > 0$ derart, daß mit $x\,\xi < \varrho$ zugleich $y\,\eta < \sigma$ ist. Es gilt dann:

IV. *Ist A in sich kompakt, so ist jede in A stetige Funktion gleichmäßig stetig.*

Sei $x_n \xi_n \to 0$; x_n hat eine konvergente Teilfolge $x_p \to x$, also auch $\xi_p \to x$; für die Bilder gilt dann $y_p \to y$, $\eta_p \to y$, $y_p \eta_p \to 0$. D. h. $y_n \eta_n$ hat eine nach 0 konvergente Teilfolge. Das Gleiche gilt von jeder Teilfolge $y_\nu \eta_\nu$, also muß die ganze Folge $y_n \eta_n$ nach 0 konvergieren (sonst hätte sie eine Teilfolge $y_\nu \eta_\nu \geqq \sigma > 0$).

V. *Jede in sich kompakte Menge ist stetiges Bild einer kompakten perfekten punkthaften Menge. Zwei kompakte perfekte punkthafte Mengen sind stets homöomorph.*

Eine kompakte perfekte punkthafte Menge A läßt sich nach § 29, XVI als dyadisches Diskontinuum

$$A = \Sigma V_p V_{pq} V_{pqr} \cdots$$

darstellen (§ 26, 2), wobei die Mengen V in A abgeschlossen, also in sich kompakt angenommen werden können (indem man V durch AV ersetzt). Bei festem n haben die disjunkten Mengen V mit n Indizes paarweise positive untere Entfernungen, unter denen ε_n die kleinste sei. Sind x, ξ verschiedene Punkte von A und unterscheiden sich die zugehörigen dyadischen Ziffernfolgen in der n-ten Ziffer, so ist $x\xi \geqq \varepsilon_n$; hieraus folgt, daß für $x\xi \to 0$ die beiden zugehörigen Ziffernfolgen in einer über alle Grenzen wachsenden Zahl von Anfangsziffern übereinstimmen. — Eine in sich kompakte Menge B läßt sich (S. 133) als dyadische Menge

$$B = \mathfrak{S}\, W_p W_{pq} W_{pqr} \cdots$$

darstellen. Man ordne dem einzigen Punkt x des Durchschnitts $V_p V_{pq} V_{pqr} \ldots$ den einzigen Punkt y des (zur gleichen dyadischen Folge gehörigen) Durchschnitts $W_p W_{pq} W_{pqr} \ldots$ zu. Durch diese Zuordnung $y = \varphi(x)$ wird B stetiges Bild von A. Denn ist δ_n der größte Durchmesser der Mengen W mit n Indizes, so ist $y\eta \leqq \delta_n$, falls y und η zu zwei dyadischen Ziffernfolgen mit n übereinstimmenden Anfangsziffern gehören; aus $\delta_n \to 0$ und dem zuvor Gesagten folgt dann, daß für $x\xi \to 0$ auch $y\eta \to 0$. — Ist schließlich auch B kompakt, perfekt und punkthaft, also als dyadisches Diskontinuum darstellbar, so ist die besprochene Abbildung schlicht, also eine Homöomorphie.

Wenn wir also ein bestimmtes dyadisches Diskontinuum A wählen, z. B. die Cantorsche triadische Menge oder auch die Menge der dyadischen Zifferfolgen $(p, q, r, \ldots)$ selbst mit der § 20, 4 erklärten Entfernung (den dyadischen Baireschen Raum), so erhalten wir alle in sich kompakten Mengen als stetige, alle kompakten perfekten punkthaften Mengen als homöomorphe Bilder von A. Umgekehrt ist jedes stetige Bild von A in sich kompakt, jedes homöomorphe überdies insichdicht (perfekt) und, wie sich sofort zeigen wird, auch punkthaft.

3. Erhaltung des Zusammenhanges.

VI. *Das stetige Bild einer zusammenhängenden Menge ist wieder zusammenhängend. Das stetige Bild einer beliebigen Menge hat nicht mehr Komponenten als diese.*

Ist B stetiges Bild von A und $B = B_1 + B_2$ eine Zerstückelung von B in zwei (in B) abgeschlossene Mengen, so entspricht dieser eine Zerstückelung von $A = \psi(B_1) + \psi(B_2) = A_1 + A_2$ in zwei (in A) abgeschlossene Mengen. Mit B ist also A unzusammenhängend, mit A ist B zusammenhängend. — Da mit $\varphi(x) = \varphi(x \mid A)$ auch jede Teilfunktion $\varphi(x \mid P)$ stetig ist, so entspricht jeder zusammenhängenden Teilmenge von A auch eine zusammenhängende von B; das Bild einer Komponente von A fällt ganz in eine Komponente von B, und das System der Komponenten von B hat höchstens dieselbe Mächtigkeit wie das der Komponenten von A. — Sind A und B homöomorph, so entsprechen einander die Komponenten beider Mengen; ist die eine punkthaft, so auch die andere.

Insbesondere hat eine in der zusammenhängenden Menge A (z. B. in einem reellen Zahlenintervall) stetige *reelle* Funktion eine zusammenhängende Wertmenge B, nimmt also zugleich mit zwei Werten $y_1 < y_2$ auch jeden Wert y zwischen ihnen $(y_1 < y < y_2)$ an. Diese „Mittelwertseigenschaft" ist übrigens keineswegs für die stetigen Funktionen charakteristisch, sondern kann auch unstetigen Funktionen zukommen, z. B. eignet sie den zu differenzierbaren Funktionen $f(x)$ gehörigen Ableitungen $f'(x)$, die ja (allerdings, wie wir sehen werden, in eingeschränkter Weise) unstetig sein können. Denn ist etwa $x_1 < x_2$, $f'(x_1) < 0$, $f'(x_2) > 0$, so hat $f(x)$ im abgeschlossenen Intervall $[x_1, x_2]$ ein Minimum, das sicher nicht an den Intervallgrenzen eintritt, und dort ist $f'(x) = 0$; aus diesem Spezialfall läßt sich der allgemeine leicht herleiten. Ein radikaleres Beispiel gibt H. Lebesgue: es sei $x = 0, x_1 x_2 \ldots$ die Dezimalbruchentwicklung (in dubio mit unendlich vielen Nullen) einer Zahl x mit $0 \leq x < 1$; falls die Ziffernfolge $x_1, x_3, x_5, \ldots$ schließlich periodisch ist und die (kürzeste) Periode aus n Ziffern besteht, sei

$$f(x) = 0, x_{2n}\, x_{2n+2}\, x_{2n+4} \cdots,$$

andernfalls etwa $f(x) = x$. Der Leser überzeuge sich, daß $f(x)$ überall unstetig ist, aber in jedem noch so kleinen Intervall alle Werte $0 \leq y \leq 1$ annimmt.

Der lokale Zusammenhang (§ 29, 2) bleibt bei stetiger Abbildung im allgemeinen nicht erhalten; wir verweisen immer wieder auf das triviale Beispiel, daß jede Menge stetiges Bild einer isolierten, also lokal zusammenhängenden Menge ist. Jedoch gilt:

VII. *Ist die Menge A in sich kompakt und in jedem Urbildpunkte $x = \psi(y)$ von y lokal zusammenhängend, so ist ihr stetiges Bild B in y lokal zusammenhängend.*

B ist wie A in sich kompakt, also jedenfalls beschränkt. Sei $y_n \to y$; wir haben nach § 29, XII zu beweisen, daß die dort erklärten Zahlen $\widehat{y y_n}$ nach 0 konvergieren. Seien $x_n = \psi(y_n)$ irgendwelche Urbilder der y_n; die x_n haben eine konvergente Teilfolge $x_p \to x$, dann ist $y_p = \varphi(x_p) \to \varphi(x)$, also $y = \varphi(x)$, $x = \psi(y)$ ein Urbild von y. A ist in x lokal zusammenhängend, $\widehat{x x_p} \to 0$; das heißt, daß (von endlich vielen p abgesehen) x und x_p durch eine zusammenhängende Menge $A_p \leq A$ verbunden werden, deren Durchmesser mit $p \to \infty$ nach 0 konvergiert. Dann sind y und y_p durch das zusammenhängende Bild $B_p = \Phi(A_p)$ verbunden, dessen Durchmesser wegen der *gleichmäßigen* Stetigkeit (Satz IV) ebenfalls nach 0 konvergiert, also $\widehat{y y_p} \to 0$. Also: die (teilweise vielleicht nicht definierten) $\widehat{y y_n}$ haben eine nach 0 konvergente Teilfolge; das Gleiche gilt von jeder

Teilfolge $\widehat{yy}_\nu$, und daher muß die ganze Folge nach 0 konvergieren, da es sonst eine Teilfolge $\widehat{yy}_\nu$ gäbe, deren Glieder entweder undefiniert oder $\geqq \sigma > 0$ wären.

Aus III und VII folgt:

VIII. *Das stetige Bild einer in sich kompakten und* (in allen Punkten) *lokal zusammenhängenden Menge ist wieder eine solche.*

§ 36. Streckenbilder.

1. Dyadische Kontinua. Ein stetiges Bild einer abgeschlossenen geradlinigen Strecke, etwa des Intervalls $T = [0, 1]$ der reellen Zahlen $0 \leqq t \leqq 1$, wird eine *stetige Kurve*, ein schlichtes stetiges und daher (§ 35, III) homöomorphes Bild von T eine *einfache Kurve* genannt. Stetige Kurven wollen wir indes lieber *Streckenbilder* nennen, da sie, wie wir sehen
[102] werden, mit dem, was der Anschauung als Kurve vorschwebt, wenig Ähnlichkeit zu haben brauchen. T ist in sich kompakt, zusammenhängend und lokal zusammenhängend; nach den Sätzen VI VIII des vorigen Paragraphen müssen diese Eigenschaften auf jedes Streckenbild übergehen, und wir werden zeigen, daß ihr Besitz auch hinreicht.

Zuvor wollen wir eine einfache Erzeugung der Streckenbilder kennen lernen. Wir bilden, wie schon oft, eine dyadische Menge, indem wir jeder dyadischen, aus den Ziffern 1, 2 gebildeten Folge $(p, q, r, \ldots)$ abgeschlossene beschränkte, von Null verschiedene Mengen

$$V_p \geqq V_{pq} \geqq V_{pqr} > \cdots \tag{1}$$

eines vollständigen Raumes E entsprechen lassen mit Durchmessern, die nach 0 konvergieren; wie wir wissen (S. 131), ist diese Konvergenz gleichmäßig, d. h. der größte Durchmesser δ_n der Mengen V mit n Indizes konvergiert nach 0 für $n \to \infty$. Wenn wir die Mengen mit n Indizes disjunkt annahmen, erhielten wir ein dyadisches Diskontinuum. Jetzt wollen wir das andere Extrem voraussetzen: die genannten Mengen sollen in lexikographischer Anordnung eine Kette bilden, d. h. in

$$\mathfrak{S} V_p = V_1 \dotplus V_2, \quad \mathfrak{S} V_{pq} = V_{11} \dotplus V_{12} \dotplus V_{21} \dotplus V_{22}, \; \ldots \tag{2}$$

haben benachbarte Mengen gemeinsame Punkte. Die von den V erzeugte dyadische Menge

$$C = \mathfrak{S}\, V_p V_{pq} V_{pqr} \cdots, \tag{3}$$

wofür wir, wie bekannt, auch

$$C = \mathfrak{S}\, V_p \,.\, \mathfrak{S}\, V_{pq} \,.\, \mathfrak{S}\, V_{pqr} \,.\, \cdots \tag{4}$$

schreiben können, heiße dann ein *dyadisches Kontinuum*; die Berechtigung des Namens wird sich sofort ergeben.

I. *Streckenbilder und dyadische Kontinua sind identisch.*

Es sei $T_1 = [0, \frac{1}{2}]$ die linke und $T_2 = [\frac{1}{2}, 1]$ die rechte Hälfte des Intervalls T, ebenso T_{p1} die linke und T_{p2} die rechte Hälfte von T_p, T_{pq1} die linke und T_{pq2} die rechte Hälfte von T_{pq} usw. Ist C stetiges Bild von T, wobei den Teilintervallen $T_p, T_{pq}, \ldots$ die Bilder $C_p, C_{pq}, \ldots$ entsprechen, so sind dies kompakte abgeschlossene Mengen (übrigens Kontinua), deren Durchmesser, wegen der gleichmäßigen Stetigkeit der Abbildungsfunktion, mit wachsender Zahl der Indizes nach 0 konvergieren, und die Summen

$$C = \mathfrak{S} C_p = \mathfrak{S} C_{pq} = \cdots$$

bilden in lexikographischer Anordnung Ketten. Ein Streckenbild ist also ein dyadisches Kontinuum.

Umgekehrt sei C ein dyadisches Kontinuum (3). Wir lassen für jede dyadische Zifferfolge den einzigen Punkt x von $V_p V_{pq} V_{pqr} \ldots$ dem einzigen Punkt t von $T_p T_{pq} T_{pqr} \ldots$ entsprechen. Diese Funktion $x = \varphi(t)$ ist *eindeutig*, obwohl gewisse t (die bei den Halbierungen der Intervalle auftretenden Teilpunkte) zu zwei dyadischen Ziffernfolgen gehören. So ist $t = \frac{1}{2}$ Punkt der beiden (und nur der beiden) Durchschnitte $T_1 T_{12} T_{122} \ldots$, $T_2 T_{21} T_{211} \ldots$. Aber auch die entsprechenden Durchschnitte $V_1 V_{12} V_{122} \ldots$, $V_2 V_{21} V_{211} \ldots$ sind identisch; denn da $V_1 V_2 > 0$, $V_{12} V_{21} > 0, \ldots$, so haben die Mengen $V_1 \dotplus V_2$, $V_{12} \dotplus V_{21}, \ldots$ Durchmesser $\leqq 2\delta_1, 2\delta_2, \ldots$ und ihr Durchschnitt ist einpunktig. — Die Funktion $x = \varphi(t)$ ist ferner *stetig*: mit $t - \tau$ konvergiert $x\xi$ nach 0. Denn sobald $|t - \tau| < \frac{1}{2^n}$, so gehören t und τ entweder zu einem oder zu zwei benachbarten T^n (Teilintervallen mit n Indizes), also x und ξ zu einem oder zu zwei lexikographisch benachbarten V^n (Mengen V mit n Indizes), und es ist $x\xi \leqq 2\,\delta_n$. Ein dyadisches Kontinuum ist also Streckenbild.

Bemerken wir folgenden *Zusatz* zu I. Wenn in sämtlichen Ketten (2) die Durchschnitte lexikographisch benachbarter V *und nur diese* von Null verschieden sind, so ist C *schlichtes, also homöomorphes Bild* von T. Ein Punkt x kann dann nämlich nur dann zu zwei dyadischen Ziffernfolgen gehören, wenn diese lexikographisch benachbart sind, und hat also auch in diesem Falle nur ein einziges Urbild $t = \psi(x)$. Wenn z. B. x zu den Zifferfolgen $(1, q_1, r_1, \ldots)$ und $(2, q_2, r_2, \ldots)$ gehört, so ist $x \varepsilon V_{1q_1} V_{2q_2}$ und das ist jetzt nur für $q_1 = 2$, $q_2 = 1$ möglich; sodann $x \varepsilon V_{12r_1} V_{21r_2}$, was nur für $r_1 = 2$, $r_2 = 1$ möglich ist, usf.; x gehört zu $(1, 2, 2, \ldots)$ und $(2, 1, 1, \ldots)$, sein einziges Urbild ist $t = \frac{1}{2}$.

Man kann die dyadische Konstruktion dahin abändern, daß jede der Ziffern $p, q, r, \ldots$ nicht zwei, sondern irgendeine endliche, vielleicht auch von den vorangehenden Ziffern abhängige Zahl ($\geqq 2$) von Werten durchläuft; durch die entsprechende Einteilung von T ergibt sich auch dann noch das „polyadische" Kontinuum (3) als Streckenbild.

Unter den Beispielen zum Satz I sind einige sehr überraschend: eine abgeschlossene Dreiecks-, Quadrat-, Kreis*fläche* ist Streckenbild, ebenso ein dreidimensionaler Tetraeder-, Würfel-, Kugel*körper*; d. h. es gibt stetige *Kurven*, die in einem Euklidischen E_n $(n \geqq 2)$ gelegen sind und innere Punkte haben, also in diesem Sinne n-dimensionale Gebilde sind. Man nennt sie nach ihrem Entdecker *Peanosche Kurven.* Um etwa mit K. Knopp ein Dreieck V (es ist die abgeschlossene Dreiecksfläche gemeint), das wir der bestimmten Vorstellung wegen rechtwinklig

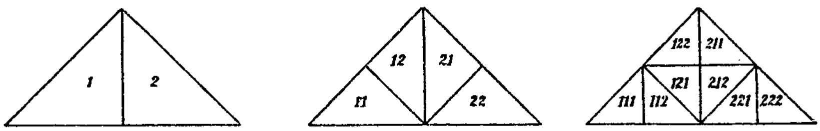

Fig. 8

und gleichschenklig annehmen, als Streckenbild darzustellen, fällen wir von der Spitze das Lot auf die Grundlinie und erhalten zwei wiederum rechtwinklige und gleichschenklige Dreiecke V_1, V_2; jedes V_p wird auf die gleiche Weise in V_{p1}, V_{p2} usw. zerlegt, wobei die Numerierung durch die lexikographische Kettenforderung bestimmt ist. Es ist

$$C = V = \mathfrak{S}\, V_p = \mathfrak{S}\, V_{pq} = \mathfrak{S}\, V_{pqr} = \cdots$$

und, da die Durchmesser bei jeder Operation im Verhältnis $\sqrt{2} : 1$ verkleinert werden, C ein dyadisches Kontinuum.

Ein Quadrat V kann man (D. Hilbert) durch Halbierung der Seiten in vier Quadrate V_1, V_2, V_3, V_4, jedes dieser V_p wieder ebenso in V_{p1}, V_{p2}, V_{p3}, V_{p4} zerlegen usw.; die Numerierung läßt sich so einrichten, daß die lexikographische Bedingung erfüllt ist. Die von G. Peano zuerst gegebene arithmetische Darstellung ist geometrisch mit Seitendrittelung und Teilung in neun Teilquadrate gleichbedeutend.

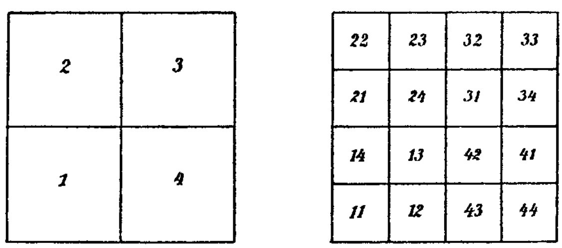

Fig. 9.

Eine arithmetisch sehr durchsichtige Abbildung

$$x_1 = \varphi_1(t), \quad x_2 = \varphi_2(t)$$

der Strecke T auf das Quadrat $(0 \leqq x_1 \leqq 1,\ 0 \leqq x_2 \leqq 1)$ stammt von H. Lebesgue. Es sei P die Cantorsche triadische Menge der Zahlen $t \in T$, die als triadische Brüche mit lauter Nullen und Zweien darstellbar sind:

$$t = 2\left(\frac{t_1}{3} + \frac{t_2}{3^2} + \frac{t_3}{3^3} + \cdots\right) \quad (t_n = 0, 1)\,.$$

Diese Darstellung ist einzig; wenn zwei solche Zahlen t, τ zuerst in der n-ten Ziffer differieren, so ist $|t-\tau| \geqq \frac{1}{3^n}$, und wenn also t nach τ konvergiert, so stimmt t in einer über alle Grenzen wachsenden Zahl von Anfangsziffern mit τ überein. Hieraus erkennt man, daß

$$x_1 = \frac{t_1}{2} + \frac{t_3}{2^2} + \cdots = \varphi_1(t), \qquad x_2 = \frac{t_2}{2} + \frac{t_4}{2^2} + \cdots = \varphi_2(t)$$

in P stetige Funktionen von t sind; schon hierbei durchläuft $x = (x_1, x_2)$ das ganze Quadrat. Endlich dehnt man diese Funktionen auf das ganze Intervall T aus, indem man sie in den offenen Intervallen (α, β) (den Komponenten) von $T - P = Q$ linear verlaufen läßt, d. h.

$$\varphi_1(t) = \frac{\beta - t}{\beta - \alpha}\,\varphi_1(\alpha) + \frac{t - \alpha}{\beta - \alpha}\,\varphi_1(\beta)$$

und ebenso $\varphi_2(t)$ definiert (α, β gehören zu P). Diese erweiterten Funktionen sind auch noch stetig, da durch die lineare Interpolation nur Mittelwerte zwischen den schon vorhandenen eingeführt werden. Wenn $t \,\varepsilon\, Q$ monoton nach $\tau \,\varepsilon\, P$ konvergiert (der einzige Fall, der einer Überlegung bedarf) und (α, β) das t einschließende Intervall ist, so konvergieren auch α und β nach τ und $\varphi_1(\alpha), \varphi_1(\beta), \varphi_1(t)$ nach $\varphi_1(\tau)$: es sei denn, daß t schließlich in einem festen Intervall (α, β) bleibt und nach einem der Endpunkte konvergiert, wobei $\varphi_1(t) \to \varphi_1(\tau)$ trivial ist.

Man kann, indem man die Folge der t_n in drei oder mehr Teilfolgen spaltet, auch den drei- oder mehrdimensionalen Würfel als Streckenbild erhalten, ja sogar bis zu $\aleph_0$ Dimensionen gehen. Spaltet man etwa nach dem dyadischen Schema S. 30, so werden die Funktionen $x_n = \varphi_n(t)$, die in P durch die dyadischen Brüche

$$x_1 = 0, t_1 t_3 t_5 \ldots, \quad x_2 = 0, t_2 t_6 t_{10} \ldots, \quad x_3 = 0, t_4 t_{12} t_{20} \ldots, \quad \ldots$$

erklärt und wie oben auf das Intervall T auszudehnen sind, allesamt stetig und durchlaufen unabhängig voneinander alle Werte der Intervalle $0 \leqq x_n \leqq 1$; die Zahlenfolge $x = (x_1, x_2, x_3, \ldots)$ durchläuft den ganzen, durch diese Ungleichungen $0 \leqq x_n \leqq 1$ definierten $\aleph_0$-dimensionalen Würfel W. Die so definierte Funktion $x = \varphi(t)$ wird ihrerseits stetig und W also Streckenbild, falls man eine geeignete Entfernung im Raum der beschränkten Zahlenfolgen einführt, etwa die dem Betrage $|x| = \Sigma c_n |x_n|$ entsprechende

$$|x - \xi| = \Sigma c_n |x_n - \xi_n|$$

(Σc_n eine festgewählte konvergente Reihe mit positiven Gliedern). Denn für $x = \varphi(t)$, $\xi = \varphi(\tau)$ konvergiert die Reihe $\Sigma c_n |x_n - \xi_n|$ gleichmäßig und ist stetige Funktion von t, für $t \to \tau$ also $|x - \xi| \to 0$.

Mit den Peanoschen Kurven erhält der Begriff *Dimension* einen zweiten Stoß; nachdem schon das Quadrat *schlichtes* Bild der Strecke sein konnte (wegen der Äquivalenz), kann es nun auch *stetiges* Bild sein. Erst die Forderung der *Homöomorphie* setzt den Dimensionsbegriff wieder in seine Rechte ein; denn es gilt der (in voller Allgemeinheit zuerst von L. E. J. Brouwer bewiesene) Satz von der *Invarianz der Dimensionenzahl*:

Eine Menge A im Euklidischen E_m und eine Menge B im Euklidischen E_n sind, wenn $n > m$ und B innere Punkte hat, niemals homöomorph.

Da wir den allgemeinen Beweis in diesem Buche nicht bringen können, sei wenigstens der ganz einfache Fall $m = 1$ besprochen: nehmen wir, indem wir B durch eine Teilmenge ersetzen, B als (abgeschlossenen) Würfel im E_n an, so bleibt B nach Tilgung eines Punktes zusammenhängend, während A durch Tilgung jedes zwischen zwei anderen gelegenen Punktes unzusammenhängend wird. Daher kann A nicht schlichtes stetiges Bild von B sein. (Wohl aber B von A! Denn B ist Streckenbild, stetiges Bild von T; behält man von den Urbildern jedes Punktes von B nur ein einziges bei, so erhält man eine Menge $A < T$, von der B schlichtes stetiges Bild ist.)

Insbesondere kann eine *einfache Kurve*, im E_n für $n > 1$ gelegen, keine inneren Punkte haben; dies kommt der anschaulichen Vorstellung einer „Kurve" schon näher als die stetige Kurve.

Auch *einfache Kurven* merkwürdiger Art lassen sich als dyadische Kontinua leicht erzeugen. Gehen wir zur Knoppschen Dreieckskonstruktion zurück. V sei ein gleichschenkliges Dreieck mit den Basiswinkeln $\frac{\pi}{5}$ und dem Winkel an der Spitze $\frac{3\pi}{5}$; durch Drittelung des letzteren entstehen drei Dreiecke, von denen wir das mittlere weglassen und die beiden äußeren (abgeschlossen) mit V_1, V_2 bezeichnen; sie sind mit V ähnlich und haben nur einen Punkt gemein. Mit V_p verfahren wir wie mit V und erhalten zwei neue Dreiecke V_{p1}, V_{p2} usw. Hier sind bei passender Numerierung die Zusatzbedingungen erfüllt, die das dyadische Kontinuum C

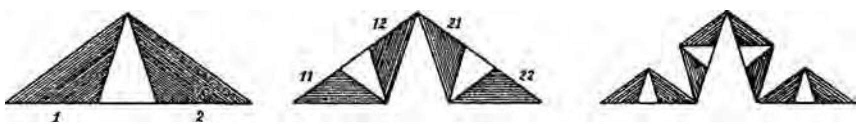

Fig. 10.

als *homöomorphes* Streckenbild sichern. Diese einfache Kurve C hat nirgends eine Tangente, da sich in beliebiger Nachbarschaft eines Punktes $x \varepsilon V_p V_{pq} \ldots$ Dreiecke mit festen Winkeln ($V_p, V_{pq}, \ldots$) befinden, in denen x liegt und deren Ecken zu C gehören; gäbe es eine Tangente in x, so müssten die Winkel dieser Dreiecke nach 0 oder π konvergieren. — Ändert man die Konstruktion unter Verzicht auf feste Winkel und gleichschenklige Dreiecke ab, mit Beibehaltung des Prinzips, daß jedes Dreieck von einer passenden Ecke aus in drei Dreiecke zerlegt und davon das mittelste weggelassen wird, so kann man auch erreichen, daß die Flächeninhalte der Dreieckssummen $\mathfrak{S} V_p$, $\mathfrak{S} V_{pq}$, ... beliebig langsam abnehmen und ihr Limes, das Flächenmaß von C, positiv wird. Es gibt also *einfache Kurven positiven Flächenmaßes*, was uns übrigens insofern nicht mehr überraschen wird, als wir ja sogar *punkthafte* ebene Mengen positiven Flächenmaßes kennen (S. 135).

2. Bedingungen für Streckenbilder.

II (Satz von W. Sierpiński). *Die Menge A ist dann und nur dann ein Streckenbild, wenn sie ein kompaktes Kontinuum ist und sich für jedes $\delta > 0$ als Summe endlich vieler kompakter Kontinua mit Durchmessern $\leqq \delta$ darstellen läßt.*

Der Ausdruck kompaktes Kontinuum könnte übrigens an beiden Stellen des Satzes durch vollständiges oder absolutes Kontinuum (zusammenhängende Menge, die vollständig oder absolut abgeschlossen ist) ersetzt werden, da von A ja zugleich totale Beschränktheit (S. 108) gefordert wird.

Daß die Bedingung notwendig ist, ist evident (Zerlegung von T in n gleiche Teilintervalle; die Durchmesser ihrer Bilder konvergieren mit $n \to \infty$ nach 0). Es ist zu zeigen, daß sie hinreicht.

Nennen wir zur Abkürzung eine der Bedingung des Satzes genügende Menge A ein S-Kontinuum. Die Hauptsache ist, zu zeigen, daß bei einer Zerlegung [1])

$$A = \mathfrak{S}_p A_p \tag{5}$$

von A in Kontinua A_p von Durchmessern $\leqq \delta$ diese selbst wieder als S-Kontinua angenommen werden können. Und dazu wieder muß bewiesen werden:

Ist C ein Kontinuum $< A$, so läßt sich C in ein S-Kontinuum $\leqq A$ einschließen, dessen Durchmesser beliebig wenig größer ist als der von C.

Der Durchmesser von C sei $d(C) = \delta$, und $\delta_1 + \delta_2 + \cdots$ sei eine konvergente Reihe positiver Zahlen. Wir können C in Kontinua C_p mit Durchmessern $\leqq \delta_1$ einschließen, die sämtlich mit C Punkte gemein haben; wir brauchen ja nur eine Darstellung (5) mit $d(A_p) \leqq \delta_1$ zu wählen und unter den C_p diejenigen A_p zu verstehen, für die $C A_p > 0$. In gleicher Weise können wir jedes C_p in Kontinua C_{pq} mit Durchmessern $\leqq \delta_2$, jedes C_{pq} in Kontinua C_{pqr} mit Durchmessern $\leqq \delta_3$ einschließen usw. Also

$$C \leqq \mathfrak{S}_p C_p, \quad C_p \leqq \mathfrak{S}_q C_{pq}, \quad C_{pq} \leqq \mathfrak{S}_r C_{pqr}, \ldots \tag{6}$$

$$C C_p > 0, \quad C_p C_{pq} > 0, \quad C_{pq} C_{pqr} > 0, \ldots \tag{7}$$

---

[1]) Alle hier auftretenden Summen sollen nur endlich viele Summanden haben.

Jede Menge C mit n Indices hat einen Durchmesser $\leqq \delta_n$. Die Menge

$$B = C \dotplus \underset{p}{\mathfrak{S}} C_p \dotplus \underset{pq}{\mathfrak{S}} C_{pq} \dotplus \cdots$$

ist zusammenhängend und hat einen Durchmesser $d(B) \leqq \delta + 2\,\delta_1 + 2\,\delta_2 + \cdots$ In der Tat sind zwei Punkte von B, etwa $x \,\varepsilon\, C_{pqr}$ und $\xi \,\varepsilon\, C_{\pi\chi}$, durch die Menge

$$C_{\pi\chi} \dotplus C_\pi \dotplus C \dotplus C_p \dotplus C_{pq} \dotplus C_{pqr}$$

verbunden, die wegen (7) als Kette zusammenhängender Mengen zusammenhängend ist und deren Durchmesser höchstens gleich der Summe der Durchmesser der Kettenglieder, also $\leqq \delta_2 + \delta_1 + \delta + \delta_1 + \delta_2 + \delta_3 < \delta + 2\,\delta_1 + 2\,\delta_2 + \cdots$ ist. Also ist B_α ein Kontinuum, dessen Durchmesser $d(B_\alpha) = d(B)$ bei hinlänglicher Kleinheit von $\delta_1 + \delta_2 + \cdots$ beliebig wenig den Durchmesser $d(C) = \delta$ übertrifft. B_α ist aber ein S-Kontinuum. Denn setzen wir weiter

$$B_p = C_p \dotplus \underset{q}{\mathfrak{S}} C_{pq} \dotplus \underset{qr}{\mathfrak{S}} C_{pqr} \dotplus \cdots,$$

$$B_{pq} = C_{pq} \dotplus \underset{r}{\mathfrak{S}} C_{pqr} \dotplus \cdots$$

usw., so sind diese genau wie B gebildeten Mengen wiederum zusammenhängend, ihre Durchmesser $d(B_p) \leqq \delta_1 + 2\,\delta_2 + 2\,\delta_3 + \cdots$, $d(B_{pq}) \leqq \delta_2 + 2\,\delta_3 + \cdots$ usw. Man hat dann wegen (6)

$$B = \underset{p}{\mathfrak{S}} B_p = \underset{pq}{\mathfrak{S}} B_{pq} = \cdots$$

$$B_\alpha = \underset{p}{\mathfrak{S}} B_{p\alpha} = \underset{pq}{\mathfrak{S}} B_{pq\alpha} = \cdots,$$

und das gibt Zerlegungen von B_α in endlich viele Kontinua, deren Durchmesser im Falle von n Indices $\leqq \delta_n + 2\,\delta_{n+1} + \cdots$ sind.

Damit ist also unsere Behauptung, die Einschließung von C in ein S-Kontinuum betreffend, bewiesen und gezeigt, daß man A in endlich viele S-Kontinua von beliebig kleinen Durchmessern zerlegen und also mit diesen das Verfahren fortsetzen kann:

$$(8) \qquad A = \underset{p}{\mathfrak{S}} A_p, \quad A_p = \underset{q}{\mathfrak{S}} A_{pq}, \quad A_{pq} = \underset{r}{\mathfrak{S}} A_{pqr}, \ldots,$$

wo alle diese Mengen S-Kontinua und die Durchmesser der A^n mit n Indices $\leqq \delta_n$ sind ($\delta_n \to 0$). Wir zeigen noch, daß bei passender Numerierung die Mengen A^n in lexikographischer Anordnung eine Kette bilden. Bei der ersten Zerlegung (5) betrachten wir alle Ketten, die aus den A_p, wiederholte Verwendung derselben Menge zugelassen, gebildet werden können, z. B. $A_1, A_2, A_1, A_3, A_3, A_1, A_2, A_4$ (wenn $A_1 A_2 > 0$, $A_1 A_3 > 0$, $A_2 A_4 > 0$). Wenn A_i mit A_k, A_k mit A_l durch eine solche Kette verbunden werden kann, so auch A_i mit A_l. Auf Grund dieses transitiven Verhaltens läßt sich jedes A_p mit jedem andern durch eine Kette verbinden; denn wenn etwa A_1 mit den A_i verbunden werden könnte, mit den A_k nicht, so kann kein

A mit keinem A_k verbunden werden, was insbesondere $A_i A_k = 0$ bedingt, und dann wäre $A = \mathfrak{S} A_i + \mathfrak{S} A_k$ eine Zerstückelung. Sind nun x, y zwei beliebige Punkte von A, $x \varepsilon A_i$, $y \varepsilon A_k$ (wobei auch $i = k$ sein darf), so kann man, indem man etwa A_i mit A_1, A_1 mit A_2, A_2 mit A_3 usw., schließlich das letzte A_p mit A_k verbindet, eine Kette $A_i \dotplus \cdots \dotplus A_k$ herstellen, die alle A_p mindestens einmal enthält, also $= A$ ist, und deren Anfangsglied x, deren Endglied y enthält; eine solche Kette (d. h. die Summe ihrer Glieder) heiße $K(x, y)$. Nebenbei sei bemerkt, daß man die Gliederzahl der Kette durch wiederholtes Setzen eines Gliedes beliebig vergrößern und demnach in (8) annehmen kann, daß p von 1 bis P, q von 1 bis Q (unabhängig von p), r von 1 bis R (unabhängig von p, q) läuft usw. Man kann dann

$$A = A_1 \dotplus \cdots \dotplus A_P = K(x, y)$$

annehmen (x, y beliebige Punkte von A; $x \varepsilon A_1, y \varepsilon A_P$), sodann

$$A_p = A_{p1} \dotplus \cdots \dotplus A_{pQ} = K(x_p, y_p)$$

(x_p, y_p beliebige Punkte von A_p; $x_p \varepsilon A_{p1}, y_p \varepsilon A_{pQ}$) usf. Wählt man hier

$$x_1 = x \varepsilon A_1, \quad y_1 = x_2 \varepsilon A_1 A_2, \quad y_2 = x_3 \varepsilon A_2 A_3, \ldots, \quad y_{P-1} = x_P \varepsilon A_{P-1} A_P,$$
$$y_P = y \varepsilon A_P,$$

so bilden die Mengen A_{pq} in lexikographischer Ordnung eine Kette, da z. B. $A_{1Q} A_{21}$ den Punkt $y_1 = x_2$ enthält, und es ist evident, daß das Verfahren unbegrenzt fortsetzbar ist. Dann ist A polyadisches Kontinuum (S. 201), also Streckenbild und II bewiesen.

III (Satz von H. Hahn und St. Mazurkiewicz). *Die Menge A ist dann und nur dann ein Streckenbild, wenn sie ein kompaktes, lokal zusammenhängendes Kontinuum ist.*

Um die Bedingung als hinreichend zu erweisen, zeigen wir, daß ein kompaktes, lokal zusammenhängendes Kontinuum A die Sierpińskische Bedingung des Satzes II erfüllt (ein S-Kontinuum ist). Umgeben wir jeden Punkt $x \varepsilon A$ mit einer abgeschlossenen Kugel V_x (auf A als Raum bezogen) vom Radius ϱ, so ist die x enthaltende Komponente A_x von V_x ein kompaktes Kontinuum und hat x als inneren Punkt (wegen des lokalen Zusammenhangs), enthält also eine Umgebung U_x. Nach dem Borelschen Satz § 26, II ist A bereits in einer Summe endlich vieler U_x enthalten, also für eine geeignete endliche Teilmenge B von A

$$A = \overset{B}{\underset{x}{\mathfrak{S}}} U_x = \overset{B}{\underset{x}{\mathfrak{S}}} A_x,$$

A als Summe endlich vieler Kontinua mit Durchmessern $\leqq 2\varrho$ darstellbar.

Daß die Bedingung in III notwendig ist, folgt aus § 35, VI, VIII; übrigens läßt sich auch unmittelbar zeigen, daß ein S-Kontinuum lokal zusammenhängend ist. Denn ist (5) eine Zerlegung von A in endlich viele

Kontinua mit Durchmessern $\leqq \delta$, und ist x in den Kontinuen A_i enthalten, in den A_k nicht, so ist $C = \mathfrak{S} A_i$ ein x enthaltendes Kontinuum vom Durchmesser $\leqq 2\delta$, das x als inneren Punkt hat, weil x in der abgeschlossenen Menge $\mathfrak{S} A_k$ nicht enthalten ist; A ist in x lokal zusammenhängend.

§ 37. Bilder Suslinscher Mengen.

Bei den Versuchen, aus der Art einer Menge A auf die Art ihres stetigen Bildes B zu schließen, haben wir bisher eigentlich nur das eine Ergebnis erzielt: das stetige Bild einer *in sich kompakten* Menge K ist wieder ein K; abgesehen von den Zusätzen, die sich auf Erhaltung des Zusammenhanges und des lokalen Zusammenhanges beziehen. Und es scheint auch zunächst gar keine Aussicht zu bestehen, über dies eine Ergebnis hinauszugelangen, da schon eine *absolut abgeschlossene* Menge F eine beliebige Menge (natürlich von höchstens derselben Mächtigkeit wie F) zum stetigen Bilde haben konnte. Wenn wir indessen Mengen in *separablen vollständigen* Räumen betrachten, so läßt sich doch ein einigermaßen scharfes Resultat gewinnen: *die stetigen Bilder Suslinscher Mengen sind Suslinsche, und die schlichten stetigen Bilder Borelscher Mengen sind Borelsche Mengen.* Suslinsche und Borelsche Mengen eines vollständigen Raumes haben ja übrigens diesen Charakter in jedem Raume, sind *absolut* Suslinsche und Borelsche Mengen, so daß wir also auch hier von *separablen absoluten* S und B sprechen können.

Bevor wir die ausgesprochene Behauptung präzisieren und beweisen, seien noch einige Bemerkungen über Raumprodukte sowie über die Projektion (S. 102) vorausgeschickt, die wir als Beispiel einer stetigen Abbildung schon öfter verwendet haben und verwenden werden. Das Produkt $Z = (X, Y)$ zweier metrischer Räume, also die Menge der Paare $z = (x, y)$ mit $x \varepsilon X$, $y \varepsilon Y$, wurde durch eine Festsetzung wie § 20, (18) oder (19) zum metrischen Raum gemacht; es kommt nicht sehr darauf an, welche Wahl wir treffen, da alle diese Räume Z homöomorph sind. Jedes Paar $z = (x, y)$ bestimmt sein erstes Element $x = \varphi(z)$, die Projektion von z auf X; das ist eine stetige, sogar gleichmäßig stetige Funktion. Das Bild $A = \Phi(C)$ einer Teilmenge von Z ist die Projektion von C auf X. Zu den Teilmengen des Produkts gehören insbesondere die Produkte von Teilmengen, $C = (A, B)$ mit $A \leqq X$, $B \leqq Y$; diese können sich auf einzelne Punkte reduzieren, und es sei z. B. (A, y) die (mit A isometrische) Menge der Punkte (x, y) mit festem y und $x \varepsilon A$. Das Produkt $C = (A, B)$ ist gegenüber Summe und Durchschnitt distributiv; es ist

$$(\mathfrak{S} A_m, B) = \mathfrak{S}(A_m, B), \quad (\mathfrak{D} A_m, B) = \mathfrak{D}(A_m, B),$$

natürlich ebenso für den zweiten Faktor. Sind X, Y vollständig, so ist Z vollständig, und umgekehrt. Ist A in X abgeschlossen, so ist (A, Y)

in (X, Y) abgeschlossen und umgekehrt; durch Komplementbildung und distributives Gesetz erkennt man: ist A in X offen oder eine Borelsche Menge F^ξ, G^ξ oder eine Suslinsche Menge S, so ist (A, Y) in (X, Y) offen oder ein F^ξ, G^ξ, S, und umgekehrt. Ist A in X und B in Y ein F^ξ, so ist in Z sowohl (A, Y) als (X, B) als auch ihr Durchschnitt (A, B) ein F^ξ, und umgekehrt: ist (A, B) ein F^ξ in $Z = (X, Y)$, so auch in (X, B), und A ist ein F^ξ in X; auch dies gilt für die G^ξ und S.

Nun sei B stetiges Bild von A, A liege in einem Raume X, B in einem Raume Y. Durch die abbildende Funktion $y = \varphi(x)$ können wir jeder Menge $P \leqq X$ eine in Y abgeschlossene Menge

$$Q = (\Phi(A\,P))_\alpha \tag{1}$$

zuordnen; d. h. wir suchen das Bild von $A\,P$ und erweitern es in Y zur abgeschlossenen Hülle. Drücken wir diese Beziehung durch $Q = \Delta(P)$ aus. Dann gilt:

I. *Wenn die Mengen $P_1 \geqq P_2 \geqq \cdots$ eine absteigende Folge bilden, nach 0 konvergente Durchmesser und einen (einzigen) zu A gehörigen Durchschnittspunkt x haben, so haben die Mengen $Q_n = \Delta(P_n)$ einen einzigen Durchschnittspunkt $y = \varphi(x)$, das Bild von x.*

Da $x \,\varepsilon\, A\,P_n$, ist $y \,\varepsilon\, \Phi(A\,P_n) \leqq Q_n$, für jedes n, also $y \,\varepsilon\, Q = Q_1 Q_2 \ldots$. Es ist nur zu zeigen, daß Q keinen von y verschiedenen Punkt hat. Sei $z \,\varepsilon\, Q$, $z \,\varepsilon\, Q_n$, also z ein α-Punkt von $\Phi(A\,P_n)$; es gibt demnach einen Punkt $y_n \,\varepsilon\, \Phi(A\,P_n)$ mit $z\,y_n < \frac{1}{n}$. Dies wieder bedeutet Vorhandensein eines Punktes $x_n \,\varepsilon\, A\,P_n$ mit $y_n = \varphi(x_n)$. Da x_n und x zu P_n gehören, ist $x x_n \leqq d(P_n)$, also $x_n \to x$, $\varphi(x_n) \to \varphi(x)$, $y_n \to y$; da zugleich $y_n \to z$, ist $z = y$.

Nun können wir zeigen:

II. *Das stetige Bild einer separablen, absolut Suslinschen Menge ist* [103]
wieder eine solche (wenn es nicht endlich ist) [1]). *Das schlichte stetige Bild einer separablen, absolut Borelschen Menge ist wieder eine solche.*

Wir bringen die Behauptung auf die Form: A liege im separablen vollständigen Raum X, der z. B. als vollständige Hülle von A gedacht werden kann, B in einem beliebigen Raum Y, B sei stetiges (oder schlichtes stetiges) Bild von A; wenn dann A eine Suslinsche (oder Borelsche) Menge in X ist, so auch B in Y.

Sei
$$A = \mathfrak{S}\, P_{n_1} P_{n_1 n_2} \ldots$$
eine Suslinsche Menge; die in X abgeschlossenen Mengen $P_{n_1 \ldots n_k}$ können so angenommen werden, daß ihre Durchmesser mit $k \to \infty$ nach 0 konvergieren; ferner sei $P_{n_1} \geqq P_{n_1 n_2} \geqq \cdots$. Vermöge der Abbildung $y = \varphi(x)$

1) Das stetige Bild B einer separablen Menge A ist höchstens separabel; denn das Bild einer in A dichten Menge ist in B dicht.

ordnen wir gemäß (1) jeder dieser Mengen die in Y abgeschlossene Menge zu

$$Q_{n_1 \ldots n_k} = \Delta(P_{n_1 \ldots n_k}).$$

Dann ist
$$Q_{n_1} Q_{n_1 n_2} \ldots = \Phi(P_{n_1} P_{n_1 n_2} \ldots).$$

Denn seien P, Q diese Durchschnitte; *entweder* besteht P aus einem einzigen Punkt $x \varepsilon A$, und nach I besteht dann Q aus dem einzigen Punkt $y = \varphi(x)$; *oder* $P = 0$, dann muß aber nach den Eigenschaften eines vollständigen Raumes (zweiter Durchschnittssatz) $P_{n_1 \ldots n_k}$ schließlich 0 sein, also auch $Q_{n_1 \ldots n_k} = 0$ und $Q = 0$. Also ist

$$B = \mathfrak{S}\, Q_{n_1} Q_{n_1 n_2} \ldots$$

eine Suslinsche Menge.

Ist A eine Borelsche Menge, also (§ 34, II) mit disjunkten Summanden

$$A = \Sigma P_{n_1} P_{n_1 n_2} \ldots$$

darstellbar und B *schlichtes* stetiges Bild von A, so ist auch

$$B = \Sigma Q_{n_1} Q_{n_1 n_2} \ldots$$

mit disjunkten Summanden darstellbar; da B wie A separabel ist und der Raum Y demnach als vollständig und separabel angenommen werden kann, so ist B eine Borelsche Menge.

Die Unterschiede des Satzes II von § 35, II sind sehr zu beachten. Damals handelte es sich um relativ Borelsche und Suslinsche Mengen (in A, in B), es wurde von Mengen $\leqq B$ auf ihre Urbilder $\leqq A$ geschlossen, die Borelschen Klassencharaktere F^ξ, G^ξ blieben erhalten. Jetzt ist von absolut Borelschen und Suslinschen Mengen die Rede, der Schluß geht von der Menge A auf ihr Bild B, und die Borelschen Klassencharaktere brauchen nicht erhalten zu bleiben. In dieser Beziehung gilt vielmehr das extreme Gegenteil: alle separablen, absolut Suslinschen Mengen sind stetige Bilder
[104] separabler, *absolut abgeschlossener* Mengen, ja sogar einer einzigen solchen Menge, nämlich des *Baireschen Nullraums.* Hierunter verstanden wir die Menge A der Folgen natürlicher Zahlen

$$x = (n_1, n_2, \ldots)$$

mit der in § 20, 4 erklärten Entfernungsdefinition; daß A vollständig oder absolut abgeschlossen ist, haben wir schon S. 105 bemerkt, und A ist separabel, da z. B. die abzählbare Menge der x, in denen schließlich lauter Einsen stehen, in A dicht ist. *Der Bairesche Nullraum ist mit der Menge der irrationalen Zahlen homöomorph* (die ein absolutes G_δ ist). Denn ordnen wir der Zahlenfolge x eineindeutig die irrationale Zahl

$$i = \frac{1|}{|n_1} + \frac{1|}{|n_2} + \cdots$$

zwischen 0 und 1 zu, so wird die Menge dieser Zahlen stetiges Bild von A und umgekehrt: die Konvergenz $x_m \to x$ bedeutet nämlich, daß x_m in einer mit m nach ∞ strebenden Zahl von Anfangsziffern mit x übereinstimmt, und daraus folgt $i_m \to i$, ebenso umgekehrt. Das offene Intervall $(0, 1)$ läßt sich nun homöomorph und mit Erhaltung des rationalen oder irrationalen Charakters auf die Menge aller reellen Zahlen abbilden, etwa durch

$$v = \frac{2u-1}{1-|2u-1|}, \quad 2u-1 = \frac{v}{1+|v|} \quad (0 < u < 1),$$

wodurch die Menge der irrationalen Zahlen zwischen 0 und 1 homöomorph auf die Menge aller irrationalen Zahlen abgebildet wird. Nach dieser Vorbemerkung beweisen wir den Satz

III. *Jede separable, absolut Suslinsche Menge ist stetiges Bild des Baireschen Nullraums A (oder der Menge I aller irrationalen Zahlen). Jede separable, absolut Borelsche Menge ist schlichtes stetiges Bild einer in A (oder I) abgeschlossenen Menge.* [105]

Es sei wie oben A der Bairesche Raum der Folgen $x = (n_1, n_2, \ldots)$ und

$$B = \mathfrak{S} F_{n_1} F_{n_1 n_2} \cdots$$

eine Suslinsche Menge in einem separablen vollständigen Raum; für jede Folge x mögen die abgeschlossenen Mengen $F_{n_1} \geqq F_{n_1 n_2} \geqq \cdots$ eine absteigende Folge mit Durchmessern $\to 0$ bilden; der Durchschnitt

$$F(x) = F_{n_1} F_{n_1 n_2} \ldots$$

ist entweder einpunktig oder leer. Im ersten Fall ordnen wir seinen einzigen Punkt $y = \varphi(x)$ dem x als Bild zu, womit bereits alle Punkte von B erhalten werden. Um auch im zweiten Fall (wo die Mengen $F_{n_1}, \ldots$ schließlich verschwinden) einen solchen Punkt $\varphi(x)$ zu definieren, wählen wir aus B einen festen Punkt y_0 und aus jeder Menge $B F_{n_1 \ldots n_k} > 0$ einen festen Punkt $y_{n_1 \ldots n_k}$; wenn dann in der Folge der Mengen

$$B F_{n_1}, \quad B F_{n_1 n_2}, \ldots$$

bereits die erste Menge verschwindet, sei $\varphi(x) = y_0$; wenn $B F_{n_1} > 0$ und $B F_{n_1 \ldots n_k}$ die letzte nicht verschwindende Menge ist, sei $\varphi(x) = y_{n_1 \ldots n_k}$. Hiermit ist $\varphi(x)$ in A eindeutig erklärt und B das Bild von A; wir behaupten, daß $\varphi(x)$ stetig ist. Sei $\xi \to x$, $\eta = \varphi(\xi)$, $y = \varphi(x)$. Die Anzahl k der Anfangsstellen, in denen $\xi = (\nu_1, \nu_2, \ldots)$ mit $x = (n_1, n_2, \ldots)$ übereinstimmt, strebt nach $+\infty$. Ist $F(x) > 0$, so gehören η und y derselben Menge $F_{n_1 \ldots n_k}$ an, deren Durchmesser mit $k \to \infty$ nach 0 konvergiert: $\eta y \to 0$. Ist $F(x) = 0$, so ist einfach $\eta = y$ für $k > h$ oder schon für $k > 0$, je nachdem $B F_{n_1 \ldots n_h}$ die letzte nichtverschwindende Menge oder bereits $B F_{n_1} = 0$ ist. Damit ist die erste Hälfte von III bewiesen. — Zum Beweise der zweiten bemerken wir, daß die Menge der x mit $F(x) > 0$ in A abgeschlossen oder die der x mit $F(x) = 0$ offen ist. Denn ist $F(x) = 0$ und $F_{n_1 \ldots n_k}$ die erste verschwindende Menge, so haben alle ξ mit $\xi x < \frac{1}{k}$ $(F_{\nu_1 \ldots \nu_k} = F_{n_1 \ldots n_k})$ ebenfalls die Eigenschaft $F(\xi) = 0$. Wir stellen sodann die Borelsche Menge B wie zuvor als Suslinsche Menge mit *disjunkten* Summanden dar und be-

14*

schränken die Definition von $y = \varphi(x)$ auf $F(x) > 0$; dann wird B schlichtes stetiges Bild einer in A abgeschlossenen Menge. — Daß man in beiden Behauptungen von III die Menge A durch die homöomorphe Menge I ersetzen kann, ist evident.

Die Suslinschen (Borelschen) Mengen in separablen vollständigen Räumen sind also stetige (schlichte stetige) Bilder von absolut abgeschlossenen Mengen oder reellen G_δ-Mengen — nicht aber von reellen oder Euklidischen abgeschlossenen Mengen (deren stetige Bilder immer nur absolute F_σ sind). Man kann sie in entsprechender Weise auch als *Projektionen* darstellen. Die Suslinsche Menge B im separablen vollständigen Raum Y sei vermöge $y = \varphi(x)$ stetiges, die Borelsche Menge B schlichtes stetiges Bild der Menge A im separablen vollständigen Raum X. Die Punktpaare(x, y) mit $x \varepsilon X$, $y \varepsilon Y$ bilden den separablen vollständigen Raum $Z = (X, Y)$, etwa mit der Entfernungsdefinition $\sqrt{x\xi^2 + y\eta^2}$. Die Menge C der Punkte $(x, \varphi(x))$, also durch $x \varepsilon A$, $y = \varphi(x)$ definiert, ist wegen der Stetigkeit von $\varphi(x)$ im Raum (A, Y) abgeschlossen; B ist ihre Projektion (bei schlichtem $\varphi(x)$ schlichte Projektion) auf Y. Wir konnten nun A als absolut abgeschlossen (Baireschen Nullraum oder Teilmenge davon) wählen; dann ist auch (A, Y) und C absolut abgeschlossen. Oder wir konnten A als G_δ in der Menge X der reellen Zahlen wählen; dann ist (A, Y) ein G_δ in (X, Y), also absolutes G_δ, ebenso C. B ist also (schlichte) Projektion eines separablen absoluten F oder G_δ. — Interessant ist besonders der Fall, daß B eine *Euklidische* Menge im Raume $Y = E_n$ ist; sie kann, wenn wir für A die zweite Wahl treffen, als Projektion von C dargestellt werden, wo C ein G_δ in $(X, Y) = (E_1, E_n)$, d. h. im Raume E_{n+1} ist. Wir haben also, wenn wir nur Euklidische Mengen und als stetige Abbildungen nur Projektionen von Räumen höherer auf Räume niederer Dimension in Betracht ziehen:

[106] IV. *Die (schlichte) Projektion einer Suslinschen (Borelschen) Menge ist eine Suslinsche (Borelsche); jede Suslinsche (Borelsche) Menge im E_n kann als (schlichte) Projektion einer Menge G_δ im E_{n+1} dargestellt werden.*

Wir wissen, daß es im Euklidischen Raum Suslinsche Mengen gibt, die keine Borelschen sind (§ 33). Die (nicht schlichte) Projektion einer Borelschen Menge ist also wohl eine Suslinsche, braucht aber keine Borelsche zu sein; z. B. liefern die ebenen G_δ als Projektionen bereits alle eindimensionalen Suslinschen Mengen, also auch die nicht-Borelschen. Die schlichte Projektion einer Borelschen Menge ist zwar eine Borelsche, aber die Klasse bleibt nicht notwendig erhalten; die ebenen G_δ liefern als schlichte Projektionen eindimensionale Borelsche Mengen beliebig hoher Klasse. Schon durch das einfache Problem, die Natur der Projektionen von Borelschen Mengen aufzuklären, wird man also in natürlicher Weise über die Borelschen Mengen hinaus und zu den Suslinschen Mengen geführt.

§ 38. Homöomorphie.

Eine Homöomorphie zwischen A und B ($A \approx B$) war als schlichte, in beiden Richtungen stetige Abbildung $y = \varphi(x)$, $x = \psi(y)$ zwischen beiden Mengen erklärt worden. Wir halten, trotz der Symmetrie dieser Beziehung, an den Bezeichnungen Bild und Urbild fest. An die Betrachtungen des vorigen Paragraphen anknüpfend, die sich hier erheblich umgestalten werden, beweisen wir zunächst einen dem dortigen Satz I ungefähr analogen Satz:

I. *Es sei $A \approx B$, A in einem vollständigen Raume X, B in einem beliebigen Raume Y gelegen. In X sei eine Folge offener Mengen $P_1, P_2, \ldots$ gegeben. Diesen kann man offene Mengen $Q_1, Q_2, \ldots$ des Raumes Y so zuordnen, daß für jede Folge $\nu = (n_1, n_2, \ldots)$ wachsender natürlicher Zahlen, für die $P_\nu = P_{n_1} P_{n_2} \ldots \leqq A$, zugleich $Q_\nu = Q_{n_1} Q_{n_2} \ldots$ das Bild von P_ν ist.* [107]

Um zu erkennen, was eigentlich bewiesen werden soll, bemerken wir: die Mengen $A_n = A\, P_n$ sind in A offen, ihre Bilder B_n in B offen (weil die Funktion $x = \psi(y)$ stetig ist!); und weil hier das Bild des Durchschnitts der Durchschnitt der Bilder ist, so ist $B_\nu = B_{n_1} B_{n_2} \ldots$ das Bild von $A_\nu = A_{n_1} A_{n_2} \ldots$. Wenn wir also irgendwie $B_n = B\, Q_n$ als Durchschnitt von B mit einer offenen Menge Q_n darstellen, so ist $B Q_\nu$ das Bild von $A\, P_\nu$; es ist zu zeigen, daß man diese offenen Mengen Q_n so „klein" wählen kann [1]), daß mit $P_\nu \leqq A$ auch $Q_\nu \leqq B$. Wir verfahren so: zu jedem Punkt $x \,\varepsilon\, A_n$ wählen wir eine Umgebung $U_n(x)$, die mitsamt ihrer abgeschlossenen Hülle in der offenen Menge P_n liegt, und zu dem Bildpunkt $y = \varphi(x) \,\varepsilon\, B_n$ eine Umgebung $V_n(y)$ so, daß das Urbild von $B\, V_n(y)$ in $U_n(x)$ liegt (was wegen der Stetigkeit der Funktion $x = \psi(y)$ möglich ist); überdies sollen die Radien beider Umgebungen $< \frac{1}{n}$ sein. Unter Q_n verstehen wir dann die Summe der Umgebungen $V_n(y)$, erstreckt über $y \,\varepsilon\, B_n$; es ist klar, daß $B\, Q_n = B_n$, denn die Menge $B\, V_n(y)$ hat ein in $A\, U_n(x) \leqq A_n$ gelegenes Urbild, ist also selbst $\leqq B_n$, daher $B\, Q_n \leqq B_n$, während zugleich $B\, Q_n$ alle Punkte $y \,\varepsilon\, B_n$ enthält. — Nunmehr ist in der Tat $Q_\nu \leqq B$ für $P_\nu \leqq A$. Seien $m < n$ Zahlen der Folge ν. Wenn $z \,\varepsilon\, Q_\nu$, so ist, für jedes n, $z \,\varepsilon\, Q_n$, also $z \,\varepsilon\, V_n(y_n)$ für geeignetes $y_n \,\varepsilon\, B_n$, $z y_n < \frac{1}{n}$, $y_n \to z$. Bei festgehaltenem m liegt also y_n wie z *schließlich* in $V_m(y_m)$. Dann liegt das Urbild $x_n = \psi(y_n) \,\varepsilon\, A_n$ schließlich in $U_m(x_m)$, also $x_m x_n < \frac{1}{m}$ für hinreichend großes n, woraus hervorgeht, daß die x_n eine Fundamentalfolge bilden, deren Limes $x \,\varepsilon\, X$ in der abgeschlossenen Hülle von $U_m(x_m)$, also in P_m liegt. Da dies

[1]) In dem analogen Fall des Satzes § 37, I war Q_n die abgeschlossene Hülle von B_n; aber eine offene Hülle gibt es nicht.

für jedes m gilt, so ist $x \varepsilon P_\nu \leqq A_\nu$. x hat also ein Bild $y = \varphi(x) \varepsilon B$, und aus $x_n \to x$ folgt $y_n \to y$, also $z = y \varepsilon B$, was zu beweisen war.

Betrachten wir nun, wie schon mehrmals, eine Menge N von Folgen
[108] $\nu = (n_1, n_2, \ldots)$ wachsender natürlicher Zahlen und bilden mit den in X offenen Mengen P_n die Menge (δs-Funktion)

$$A = \overset{N}{\underset{\nu}{\mathfrak{S}}} P_\nu = \overset{N}{\underset{\nu}{\mathfrak{S}}} P_{n_1} P_{n_2} \ldots = \Phi(P_1, P_2, \ldots),$$

so läßt sich jede mit A homöomorphe Menge B nach I in der Gestalt

$$B = \overset{N}{\underset{\nu}{\mathfrak{S}}} Q_\nu = \overset{N}{\underset{\nu}{\mathfrak{S}}} Q_{n_1} Q_{n_2} \ldots = \Phi(Q_1, Q_2, \ldots)$$

darstellen mit offenen Mengen Q_n des Raums $Y \geqq B$ (X soll vollständig, Y kann beliebig sein). *Die Funktion Φ bleibt also bei Homöomorphie erhalten.* Bei geeignetem N stellt nun A alle Borelschen Mengen G^ξ ($\xi \geqq 1$ fest) des Raumes X dar, oder auch alle Suslinschen Mengen. Wir erhalten demgemäß:

[109] II. *Das homöomorphe Bild einer absolut Suslinschen Menge ist wieder eine solche; das homöomorphe Bild einer absolut Borelschen Menge G^ξ ($\xi \geqq 1$) ist wiederum eine solche (mit demselben Index ξ).*

Gegenüber dem Satz II des vorigen Paragraphen bemerke man den Verzicht auf die Voraussetzung der Separabilität und beachte die Erhaltung der Borelschen Klasse.

Bezüglich der absoluten Mengen F^ξ lehrt II nur, daß sie, als spezielle Mengen $G^{\xi+1}$, zu homöomorphen Bildern wieder Mengen $G^{\xi+1}$ haben. Wir werden dies Ergebnis nachher verbessern und die Erhaltung der Mengen F^ξ für $\xi \geqq 2$ erkennen; auch für $\xi = 1$ wird sich eine gewisse Verschärfung ergeben. Für $\xi = 0$ ist aber über den Satz II nicht hinauszukommen; es gilt vielmehr:

[110] III. *Das homöomorphe Bild einer absolut abgeschlossenen Menge ist ein absolutes G_δ, und jedes absolute G_δ ist mit einer absolut abgeschlossenen Menge homöomorph.*

Die erste Hälfte folgt aus II, da jedes absolute F ein absolutes $G_\delta = G^1$ ist. Die noch zu beweisende zweite können wir so ausdrücken: eine Youngsche Menge A, d. h. ein G_δ in einem vollständigen Raume E, ist mit einem vollständigen Raume homöomorph.

Zunächst sei E nur ein metrischer Raum, F abgeschlossen ($0 < F < E$); $\delta(x, F)$ die untere Entfernung des Punktes x von der Menge F. Wir betrachten Punkte x, y, z des Komplements $G = E - F$ und definieren

$$(1) \qquad \varphi(x, y \mid F) = \frac{xy}{xy + \delta(x, F) + \delta(y, F)}.$$

Dieser in x, y symmetrische Ausdruck hat den Charakter einer Entfernung; er verschwindet für $x = y$, ist andernfalls positiv (< 1) und *erfüllt das Dreiecksaxiom*, denn schreibt man kurz δ_x für $\delta(x, F)$ und beachtet $\delta_y \leqq \delta_z + yz$, $\delta_y \leqq \delta_x + xy$, so ist

$$\frac{xy}{xy + \delta_x + \delta_y} + \frac{yz}{yz + \delta_y + \delta_z} \geqq \frac{xy + yz}{xy + yz + \delta_x + \delta_z} \geqq \frac{xz}{xz + \delta_x + \delta_z}.$$

Indem man noch $\delta_y \leqq \delta_x + xy$ in (1) anwendet, erhält man

$$(2) \qquad \frac{xy}{xy + \delta(x, F)} \geqq \varphi(x, y \mid F) \geqq \tfrac{1}{2} \frac{xy}{xy + \delta(x, F)}.$$

Nun sei $A = G_1 G_2 \ldots$ ein G_δ, das Komplement $B = F_1 \dotplus F_2 \dotplus \cdots$ ein F_σ; man wähle eine konvergente Reihe positiver Zahlen $c_1 + c_2 + \cdots$ und definiere für Punkte von A

$$\overline{xy} = \Sigma c_n \varphi(x, y \mid F_n).$$

Dies hat wieder Entfernungscharakter, und durch Beilegung dieser Entfernungen wird die Menge A ein (von A verschiedener) metrischer Raum $\bar{A}$, der schlicht auf A abgebildet ist (jedem Punkt $x \varepsilon A$ entspricht derselbe Punkt $x \varepsilon \bar{A}$). Wir wollen zeigen, daß A und $\bar{A}$ hierdurch homöomorph werden, d. h. daß, bei festem x, aus $xy \to 0$ zugleich $\overline{xy} \to 0$ folgt und umgekehrt. In der Tat folgt aus (2)

$$\overline{xy} \leqq \Sigma c_n \frac{xy}{xy + \delta(x, F_n)};$$

für $xy \to 0$ konvergiert, wegen $\delta(x, F_n) > 0$, jedes Glied der Reihe nach 0 und wegen der gleichmäßigen Konvergenz auch die Reihensumme. Andererseits ist nach (2)

$$(3) \qquad \overline{xy} \geqq \frac{c_1}{2} \frac{xy}{xy + \delta(x, F_1)}$$

und mit $\overline{xy}$ konvergiert xy nach 0.

Endlich sei
$$\widehat{xy} = \max[xy, \overline{xy}];$$
das ist wieder eine Entfernung und erzeugt einen dritten, mit A und $\bar{A}$ homöomorphen Raum $\hat{A}$. Dieser ist nun vollständig, wenn E vollständig ist, was den Beweis von III vollendet. Sei nämlich x_n eine Fundamentalfolge in $\hat{A}$, also (wegen $\widehat{xy} \geqq xy$) auch eine in A und in $\bar{A}$; sie hat also, im Sinne der ursprünglichen Entfernungen xy, einen Limes in E. Dieser kann kein Punkt t von B sein; denn wäre etwa $t \varepsilon F_1$, so würde aus (3) für $m < n$

$$\overline{x_m x_n} \geqq \frac{c_1}{2} \frac{x_m x_n}{x_m x_n + \delta(x_n, F_1)},$$

also für $n \to \infty$ ($x_m x_n \to x_m t > 0$, $\delta(x_n, F_1) \to 0$)

$$\liminf_n \overline{x_m x_n} \geqq \frac{c_1}{2}$$

folgen, und dann ist x_n in A, also erst recht in $\widehat{A}$, keine Fundamentalfolge, da doch für eine solche sogar $\limsup\limits_n \overline{x_m x_n}$ mit wachsendem m nach 0 konvergieren müßte. Also gibt es einen Punkt $x \varepsilon A$ mit $xx_n \to 0$, $\widehat{xx_n} \to 0$, der Raum $\widehat{A}$ ist vollständig.

Ein Beispiel ist uns schon bekannt (S. 210): die Menge der irrationalen Zahlen, ein absolutes G_δ, ist mit dem absolut abgeschlossenen Baireschen Nullraum homöomorph.

Sind zwei in Euklidischen Räumen gelegene Mengen A, B homöomorph und A abgeschlossen, so ist B ein G_δ, aber zugleich ein F_σ (ein K_σ, S. 197), also reduzibel. Die nicht reduzible Menge der irrationalen Zahlen kann also mit keiner Euklidischen abgeschlossenen Menge homöomorph sein.

Wir wollen den Satz II noch auf einem andern, etwas weiter führenden Wege beweisen.

[111] IV (Satz von M. Lavrentieff). *Eine Homöomorphie zwischen A und B läßt sich zu einer Homöomorphie zwischen zwei absoluten G_δ-Mengen $X \geqq A$ und $Y \geqq B$ erweitern.*

Es sei zunächst nur B stetiges Bild von A, vermöge der stetigen Funktion $y = \varphi(x)$; A_0 und B_0 seien vollständige Hüllen von A und B. Betrachten wir einen Punkt $x \varepsilon A_0$, der also Limes mindestens einer Folge von Punkten $a_n \varepsilon A$ ist; $b_n = \varphi(a_n)$ seien deren Bilder. Die b_n können eine Fundamentalfolge bilden oder nicht; im ersten Fall konvergieren sie nach einem Punkt $y \varepsilon B_0$. Es kann nun sein, daß *allen* nach x konvergenten Folgen a_n Fundamentalfolgen b_n entsprechen (wie dies z. B. für $x \varepsilon A$ der Fall ist, wo stets $b_n \to y = \varphi(x)$ ist); dann konvergieren alle diese Fundamentalfolgen b_n nach *einem und demselben* Punkte $y \varepsilon B_0$; denn wenn den beiden Folgen $a_n \to x$, $\bar{a}_n \to x$ zwei Folgen $b_n \to y$, $\bar{b}_n \to \bar{y}$ mit $\bar{y} \neq y$ entsprächen, so würde die Folge $a_1, \bar{a}_1, a_2, \bar{a}_2, \ldots$ nach x konvergieren und ihre Bilder $b_1, \bar{b}_1, b_2, \bar{b}_2, \ldots$ keine Fundamentalfolge liefern. Nennen wir A_1 die Menge der Punkte $x \varepsilon A_0$, welche die angegebene Eigenschaft haben (allen Folgen $a_n \to x$ entsprechen Fundamentalfolgen b_n), so entspricht also jedem Punkte $x \varepsilon A_1$ ein ganz bestimmter Punkt $y = \varphi_1(x)$ von B_0, nämlich so, daß für $a_n \to x$ stets $b_n \to y$; wir haben hier eine eindeutige Funktion $\varphi_1(x)$ in A_1 definiert, wobei $A \leqq A_1$ und für $x \varepsilon A$ $\varphi_1(x) = \varphi(x)$ ist. Diese erweiterte Funktion, die wir nun wieder $\varphi(x)$ nennen können, ist in A_1 stetig; denn wenn $x_n \to x$ (x_n, $x \varepsilon A_1$) und $y_n = \varphi(x_n)$, $y = \varphi(x)$, so kann man ein $a_n \varepsilon A$ mit $a_n x_n < \frac{1}{n}$ so bestimmen, daß auch $b_n y_n < \frac{1}{n}$, dann ist $a_n \to x$, also $b_n \to y$ und $y_n \to y$. — Die Menge A_1 ist in A_0 ein G_δ, d. h. ein absolutes G_δ. In der Tat läßt sich die Forderung, daß bei beliebiger Annäherung von a_n an x das Bild b_n eine Fundamentalfolge beschreiben soll, auch so ausdrücken: es gibt für jedes $\sigma > 0$ ein $\varrho > 0$ derart, daß für $x a < \varrho, \bar{x} \bar{a} < \varrho$ die Entfernung $b \bar{b}$ der Bilder $\leqq \sigma$ ist, oder es gibt für jedes σ eine (auf A_0 als Raum bezogene) Umgebung U_x derart, daß das Bild $\Phi(A U_x)$ einen Durchmesser $\leqq \sigma$ hat. Halten wir zunächst σ fest, so ist die Menge $G(\sigma)$ der Punkte x, die eine solche Umgebung haben, offen (in A_0); denn jeder Punkt von U_x hat selbst eine solche Umgebung ($\leqq U_x$). Die Menge der x, die für jedes σ, oder für $\sigma = 1, \frac{1}{2}, \frac{1}{3}, \ldots$ eine solche Umgebung haben, ist $A_1 = G(1)G(\frac{1}{2})\ldots$, also ein G_δ[1]).

Wir haben also bewiesen: eine in A stetige Funktion $\varphi(x)$ läßt sich zu einer in A_1 stetigen Funktion erweitern (derart, daß in A beide übereinstimmen); A_1 ist ein absolutes G_δ, in der vollständigen Hülle A_0 von A enthalten.

Sind nun A, B homöomorph, so erweitern wir auch die in B stetige Funktion $\psi(y)$ zu einer in B_1 stetigen, wo B_1 ein G_δ in B_0 ist.

Die Funktion $\varphi(x)$ liefert von A_1 das stetige Bild $B_2 \leqq B_0$, das nicht mit B_1 übereinzustimmen braucht (vgl. das unten folgende Beispiel). D. h. wenn x die Menge A_1 durchläuft, durchläuft $\varphi(x)$ die Menge B_2. Ebenso sei A_2 das stetige Bild von B_1 vermöge der Funktion $\psi(y)$; d. h. wenn y die Menge B_1 durchläuft, durchläuft $\psi(y)$ die Menge A_2.

Endlich sei A_3 die Menge der Punkte $x \varepsilon A_1$, für die $\varphi(x) \varepsilon B_1$, also das Urbild von $B_1 B_2$ bei der stetigen Abbildung von A_1 auf B_2; offenbar ist $A \leqq A_3 \leqq A_1$. Da $B_1 B_2$ ein G_δ in B_2 ist, so ist nach § 35, II A_3 ein G_δ in A_1, also wieder ein absolutes G_δ. Desgleichen sei B_3 die Menge der Punkte $y \varepsilon B_1$, für die $\psi(y) \varepsilon A_1$; es ist $B \leqq B_3 \leqq B_1$ und B_3 ein absolutes G_δ.

Die Mengen A_3, B_3 werden durch $y = \varphi(x), x = \psi(y)$ homöomorph aufeinander abgebildet. Sei nämlich $x \varepsilon A_3$, $y = \varphi(x) \varepsilon B_1$; nach der Definition von B_1 besagt das: für eine beliebige Folge $b_n \to y$ $(b_n \varepsilon B)$ konvergiert $a_n = \psi(b_n)$ nach $\psi(y)$. Nehmen wir aber eine beliebige Folge $a_n \to x (a_n \varepsilon A)$, so konvergiert $b_n = \varphi(a_n)$ nach $y = \varphi(x)$, demnach muß a_n nach $\psi(y)$ konvergieren, d. h. $x = \psi(y)$ sein. Wegen $\psi(y) \varepsilon A_1$ ist überdies $y \varepsilon B_3$. Also: aus $x \varepsilon A_3, y = \varphi(x)$ folgt $y \varepsilon B_3, x = \psi(y)$ und umgekehrt; B_3 ist vermöge $y = \varphi(x)$ stetiges Bild von A_3 und ebenso A_3 vermöge $x = \psi(y)$ stetiges Bild von B_3. Damit ist IV bewiesen ($X = A_3$, $Y = B_3$). Übrigens ist $A_3 = A_1 A_2$, $B_3 = B_1 B_2$. Denn aus $A_3 \leqq A_1$ folgt vermöge der Abbildung $y = \varphi(x): B_3 \leqq B_2$, also $B_3 \leqq B_1 B_2$, $A_3 \leqq A_1 A_2$. Umgekehrt ist auch $A_1 A_2 \leqq A_3$; denn aus $x \varepsilon A_1 A_2$ folgt (wegen $x \varepsilon A_2$), daß $x = \psi(y)$ mit $y \varepsilon B_1$, sodann (wegen $\psi(y) \varepsilon A_1$), daß $y \varepsilon B_3$ und $x \varepsilon A_3$ ist. Schließlich bemerke man noch, daß A_3, B_3 die *größten* Mengen X, Y ($\leqq A_0, B_0$) sind, auf welche die Homöomorphie zwischen A, B ausgedehnt werden kann; denn wenn x, y einander entsprechende Punkte von X, Y sind, so muß $x \varepsilon A_1$, $y = \varphi(x)$ und $y \varepsilon B_1$, $x = \psi(y)$, demnach $x \varepsilon A_3$, $y \varepsilon B_3$ sein.

[1]) Das würde übrigens auch ohne die Voraussetzung der Stetigkeit von $\varphi(x)$ gelten, nur brauchte dann A nicht in A_1 enthalten zu sein.

Beispiel. Es sei

$$A = (0, 1) + (1, 2) + (2, 3)$$

die Summe der drei offenen Zahlenintervalle $(0, 1)$, $(1, 2)$, $(2, 3)$; in diesen werde die Funktion $y = \varphi(x)$ durch

$$\varphi(x) = 1 - x, 3 - x, x$$

definiert. Hierdurch wird A homöomorph auf sich selbst ($B = A$) derart abgebildet, daß

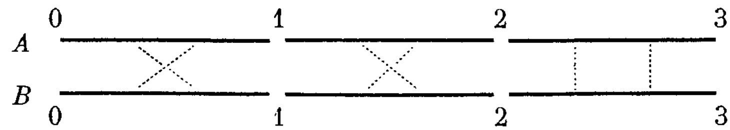

Fig. 11.

die beiden ersten Intervalle an ihren Mittelpunkten gespiegelt werden, das dritte Punkt für Punkt in Ruhe bleibt. Da die Abbildung involutorisch, d. h. die inverse Funktion $\psi(y)$ mit $\varphi(y)$ identisch ist, so stimmen A_0, A_1, A_2, A_3 mit B_0, B_1, B_2, B_3 überein. Man hat

$$\begin{aligned} A_0 &= A + \{0, 1, 2, 3\} \\ A_1 &= A + \{0, 3\} \\ A_2 &= A + \{1, 3\} \\ A_3 &= A + \{3\}. \end{aligned}$$

Die zweite Formel ergibt sich so: für $x \to 0$ ist $y \to 1$ und für $x \to 3$ ist $y \to 3$; die Punkte $0, 3$ sind also in A_1 und ihre Bilder $1, 3$ in A_2 aufzunehmen. Dagegen sind 1 und 2 nicht in A_1 aufzunehmen, denn wenn x von links oder rechts nach 1 konvergiert, so konvergiert y nach 0 oder 2, so daß nicht jeder Folge $x \to 1$ eine Fundamentalfolge entspricht; analog verhält sich der Punkt 2.

Aus IV folgt nun der Satz II noch einmal mit dem Zusatz:

[112] II*. *Das homöomorphe Bild einer absolut Borelschen Menge $F^\xi (\xi \geqq 2)$ ist wieder eine solche (mit demselben Index ξ).*

Nachdem nämlich die Homöomorphie auf die absoluten G_δ-Mengen X, Y ausgedehnt ist, können wir den Satz § 35, II anwenden. Ist A ein absolutes $G^\xi (\xi \geqq 1)$, so auch in X; dann ist, wegen der Stetigkeit von $x = \psi(y)$, B ein G^ξ in Y, also ein absolutes G^ξ, da Y ein absolutes G^1 ist. Ebenso ist, falls A ein absolutes $F^\xi (\xi \geqq 2)$ ist, B ein F^ξ in Y, also ein absolutes F^ξ, da Y ein absolutes G^1 und F^2 ist. Analog ist der Beweis für die absolut Suslinschen Mengen zu führen.

Einige weitere, bei Homöomorphie invariante Mengencharaktere kann man aus IV durch Betrachtung der Komplemente $X - A$, $Y - B$ erhalten Z. B. gehen die Komplemente absolut Suslinscher Mengen (in einem vollständigen Raum) durch Homöomorphie wieder in solche über. Die Differenzen von zwei absoluten G_δ gehen wieder in solche über; denn ist $A = A_1 - A_2$ Differenz von zwei absoluten G_δ, die man wegen $XA = XA_1 - XA_2$ als Teilmengen von X annehmen kann, so ist $B = B_1 - B_2$ auch eine solche Differenz. Das trifft insbesondere auf die absoluten Mengen $F^1 = F_\sigma$ zu.

Wir stellen in einer kleinen Tabelle die wichtigsten Ergebnisse zusammen, die sich auf stetige, schlichte stetige und homöomorphe Bilder von Mengen beziehen; K bedeutet eine in sich kompakte Menge; die Charaktere F (abgeschlossen), F^ξ, G^ξ, B (Borelsche Menge), S (Suslinsche Menge) sind absolut zu verstehen; der Zusatz sep. bedeutet separabel; ein leerer Platz bedeutet, daß sich nichts aussagen läßt.

A	Bild von A		
	stetig	schlicht stetig	homöomorph
K	K	K	K
F			G_δ
F^ξ			$F^\xi (\xi \geqq 2)$
G^ξ			$G^\xi (\xi \geqq 1)$
S			S
sep. F	S	B	G_δ
sep. B	S	B	B
sep. S	S	S	S

§ 39. Einfache Kurven. [113]

1. Bedingungen für einfache Kurven.

Unser Streckenbild, d. h. das stetige Bild einer abgeschlossenen Strecke oder des Zahlenintervalls $T = [0, 1]$ brauchte mit dem, was man sich anschaulich unter einer Kurve vorstellt, wenig Ähnlichkeit zu haben. Etwas näher kommt dieser Vorstellung die *einfache Kurve*[1]), nämlich das *homöomorphe Bild* von T. Damit C ein solches sei, sind jedenfalls folgende Eigenschaften notwendig, deren Verzeichnis leicht noch weiterzuführen wäre und bei deren Fassung wir die Begriffe abgeschlossen, Kontinuum usw. auf C selbst als Raum beziehen:

(α) C ist in sich kompakt.

(β) C hat zwei Punkte a, b, zwischen denen es *irreduzibles Kontinuum* ist; d. h. C selbst ist Kontinuum, aber kein Kontinuum $< C$ enthält die Punkte a, b.

---

[1]) Auch einfacher Kurvenbogen, Jordanscher Kurvenbogen genannt. Als einfache geschlossene Kurve oder Jordansche Kurve bezeichnet man das homöomorphe Bild einer Kreisperipherie (in den Fundamenta Mathematicae bedeutet ligne de Jordan u. dgl. so viel wie Streckenbild).

(γ) C hat zwei Punkte a, b, zwischen denen es *irreduzibel zusammenhängend* ist; d. h. C selbst ist zusammenhängend, aber keine zusammenhängende Menge $< C$ enthält die Punkte a, b.

(δ) C ist zusammenhängend und hat zwei Punkte a, b folgender Art: für jeden Punkt $x \varepsilon C$ gibt es eine Zerlegung $C = A \dotplus B$, wo A, B abgeschlossen sind, nur den Punkt x gemein haben und $a \varepsilon A$, $b \varepsilon B$.

(ε) C ist lokal zusammenhängend.

In der Tat folgt die Notwendigkeit von (α)(ε) aus § 36, III. Die übrigen Bedingungen ergeben sich, wenn man unter a, b die Bilder der Endpunkte 0, 1 von T versteht; keine zusammenhängende Menge $< T$ enthält die Punkte 0, 1, woraus (γ) und um so mehr (β) folgt; und für jedes $t \varepsilon T$ gibt $T = [0, t] \dotplus [t, 1]$ eine Zerlegung, die auf C übertragen die Eigenschaft (δ) hat (unter $[0, 0]$ und $[1, 1]$ sind die einpunktigen Mengen $\{0\}$ und $\{1\}$ zu verstehen).

Wir werden nun zeigen, daß die Bedingungen ($\alpha, \beta, \varepsilon$) oder ($\alpha, \gamma$) oder ($\alpha, \delta$) auch hinreichend sind.

I. *Die Bedingungen* (γ) *und* (δ) *sind gleichbedeutend.*

Daß (γ) aus (δ) folgt, ist sehr einfach einzusehen. Wenn $D < C$ die Punkte a, b enthält, sei $x \varepsilon C - D$ und $C = A \dotplus B$ eine gemäß (δ) zu x gehörige Zerlegung; dann ist $DAB = 0$ und $D = DA + DB$ eine Zerstückelung, D nicht zusammenhängend.

Umgekehrt, sei (γ) erfüllt oder auch nur folgende Teilannahme: C sei zwischen a, b irreduzibles Kontinuum im Sinne von (β) und werde durch Tilgung eines von a, b verschiedenen Punktes unzusammenhängend. Um daraus (δ) zu beweisen, können wir x von a, b verschieden annehmen, da für $x = a$ und $x = b$ die Zerlegungen $C = a \dotplus C = C \dotplus b$ den gestellten Forderungen genügen (wir bezeichnen jetzt einpunktige Mengen $\{x\}$ einfach mit x). Dann ist $C - x$ unzusammenhängend, also in $C - x = P + Q$ zerstückelbar, wo P, Q in $C - x$ abgeschlossen sind. Dann sind die Mengen $A = P + x$, $B = Q + x$ (in C) abgeschlossen, überdies sind sie nach § 29, III zusammenhängend, da ihre Summe C und ihr Durchschnitt x zusammenhängend sind. Keines der Kontinua $A < C$, $B < C$ kann beide Punkte a, b enthalten; es ist also etwa $a \varepsilon A$, $b \varepsilon B$ und $C = A \dotplus B$ ist eine Zerlegung gemäß (δ). Damit ist I bewiesen.

II. *Wenn C separabel ist und der Bedingung* (γ) *oder* (δ) *genügt, so ist das reelle Zahlenintervall $T = [0, 1]$ schlichtes stetiges Bild von C.*

Daß C separabel ist, wird erst am Schluß eine Rolle spielen; sehen wir von dieser Bedingung zunächst ab. Die Mengen A, B in (δ) sind, wie wir soeben sahen, zusammenhängend (Kontinua).

Sind x_1, x_2 zwei (nicht notwendig verschiedene) Punkte und $C = A_1 \dotplus B_1 = A_2 \dotplus B_2$ zu ihnen gehörige Zerlegungen, so gilt:

(1) $\qquad$ für $x_1 \varepsilon A_2$ ist $A_1 \leqq A_2$.

Es ist nämlich $A_2 + B_1$, als Summe zusammenhängender Mengen mit dem gemeinsamen Punkt x_1, zusammenhängend; da sie a, b enthält, ist nach (γ) $C = A_2 \dotplus B_1$, mit A_1 geschnitten $A_1 = A_1 A_2 \dotplus x_1 = A_1 A_2$.

Für $x_1 = x_2$ folgt hieraus $A_1 = A_2$, ebenso natürlich $B_1 = B_2$, d. h. die zu x gehörige Zerlegung $C = A \dotplus B$ ist durch x *eindeutig bestimmt*. Nennen wir A das zu x gehörige *Anfangsstück* von C (B das Endstück). Zu $x = a$ gehört das Anfangsstück a, zu $x = b$ das Anfangsstück C.

Wenn aber in (1) $x_1 \neq x_2$, so ist

(2) $\qquad A_1 < A_2, \quad x_2 \varepsilon A_2 - A_1$.

Wäre nämlich zugleich $x_1 \varepsilon A_2$ und $x_2 \varepsilon A_1$, so würde $A_1 = A_2$ folgen, also $B_1 - x_1 = B_2 - x_2$, $B_1 + x_2 = B_2 + x_1$. In dieser letzten Gleichung, die eine Zerlegung in Komponenten ausspricht, müßte $B_1 = x_1$, $B_2 = x_2$ sein; aber B_1 und B_2 enthalten b, und man gelangt zum Widerspruch $x_1 = x_2 = b$. Also: sind x_1, x_2 verschieden und $x_1 \varepsilon A_2$, so ist nicht $x_2 \varepsilon A_1$, womit (2) bewiesen ist.

Natürlich können für $x_1 \neq x_2$ auch die Relationen $x_1 \varepsilon B_2$, $x_2 \varepsilon B_1$ nicht gleichzeitig bestehen. Es besteht also eine und nur eine der Relationen

$$x_1 \varepsilon A_2, \quad A_1 < A_2, \quad x_2 \varepsilon A_2 - A_1,$$

oder

$$x_2 \varepsilon A_1, \quad A_2 < A_1, \quad x_1 \varepsilon A_1 - A_2.$$

Die zu verschiedenen Punkten gehörigen Anfangsstücke stehen also immer in der Beziehung $A_1 \lesseqgtr A_2$. Wir machen C zu einer *geordneten Menge*, indem wir

$$x_1 < x_2 \quad \text{für} \quad A_1 < A_2$$

definieren. Dies ist mit $x_1 \varepsilon A_2$, $x_1 \neq x_2$ gleichbedeutend; außerdem enthält A_2 noch den Punkt x_2.

Nach geschehener Ordnung ist also A die Menge der Punkte $\leqq x$, B die Menge der Punkte $\geqq x$; wir schreiben dafür, analog zur Bezeichnung reeller Zahlenintervalle,

$$A = [a, x], \quad B = [x, b].$$

Diese Mengen sind abgeschlossen, ebenso allgemein die Menge $[x_1, x_2]$ der Punkte $x_1 \leqq x \leqq x_2$ (der Durchschnitt $A_2 B_1$). Mengen wie $C - B = A - x = [a, x)$ (die Menge der Punkte $< x$), $(x, b]$, (x_1, x_2) sind offen (in C). Natürlich ist $[a, a] = a$, $[a, a) = 0$ zu setzen.

Die geordnete Menge C hat das erste Element a und das letzte b. Sie ist *dicht* (S. 50), d. h. für $x_1 < x_2$ gibt es immer noch ein Element x mit $x_1 < x < x_2$. Denn die zusammenhängende Menge A_2 kann nicht $= A_1 + x_2$ sein, also enthält $A_2 - A_1$ gewiß noch einen von x_2 verschiedenen Punkt x. Sie ist ferner *stetig* (S. 53). Denn sei $C = C_1 + C_2$ in zwei disjunkte Mengen > 0 gespalten, für deren Elemente stets $x_1 < x_2$. Wenn C_1 kein letztes Element hat, so ist sie gleich der Summe (über $x_1 \varepsilon C_1$) der Mengen

$[a, x_1)$, also offen; ebenso ist C_2 offen, falls sie kein erstes Element hat. Beides zugleich würde eine Zerstückelung von C bedeuten, also hat entweder C_1 ein letztes oder C_2 ein erstes Element.

Wenn ferner C separabel ist, also eine abzählbare Menge R in C dicht ist (im metrischen Sinn), so ist R auch im ordinalen Sinn (S. 54) in C dicht. Denn (x_1, x_2) ür $x_1 < x_2$ ist offen, enthält also einen Punkt r von R, d. h. $x_1 < r < x_2$. Nach § 11, V ist $C - \{a, b\}$ also vom Typus λ des offenen Intervalls $(0, 1)$, C selbst vom Typus des abgeschlossenen Intervalls $T = [0, 1]$. Es existiert demnach eine *ähnliche* Abbildung

$$x = \varphi(t), \quad t = \psi(x)$$

von C auf $T(0 \leqq t \leqq 1)$ mit $a = \varphi(0)$, $b = \varphi(1)$. Hier ist die Funktion $t = \psi(x)$ stetig. Denn jedem in T offenen Intervall (Durchschnitt von T mit einem offenen Intervall), nämlich $[0, t)$, $(t, 1]$, (t_1, t_2) entspricht eine in C offene Menge, nämlich $[a, x)$, $(x, b]$, (x_1, x_2), also jeder in T offenen Menge eine in C offene Menge, was eben nach § 35, I die Stetigkeit von $\psi(x)$ besagt. Damit ist II bewiesen.

Der Satz läßt sich übrigens dahin umkehren: *wenn C zusammenhängend und T schlichtes stetiges Bild von C ist, so erfüllt C die Bedingung (γ) oder (δ).* Denn die Zerlegung $T = [0, t] \dot{+} [t, 1]$ gibt eine durch (δ) geforderte Zerlegung $C = A \dot{+} B$. Aus der bloßen Tatsache, daß T schlichtes stetiges Bild von C ist, folgt weder der Zusammenhang noch die Separabilität von C; C kann ja (S. 195) eine beliebige Menge der Mächtigkeit $\aleph$ sein, in der je zwei Punkte die Entfernung 1 haben.

Ist nun C in sich kompakt und T schlichtes stetiges Bild von C, so ist die Abbildung eine Homöomorphie (§ 35, III). Also:

III. *Damit C eine einfache Kurve sei, sind die Bedingungen (α, γ) oder (α, δ) notwendig und hinreichend.*

Die Bedingungen (α, γ) sind von N. J. Lennes, die (α, δ) von W. Sierpinski gegeben worden.

Wenn man nun die Bedingung (γ) zu (β) abschwächt, so ist eine anderweitige Verstärkung notwendig: diese wird durch die Bedingung des lokalen Zusammenhangs geliefert, d. h. es gilt:

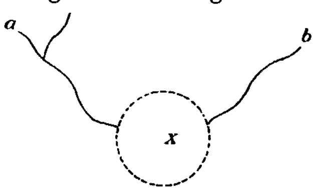

Fig. 12.

IV. *Damit C eine einfache Kurve sei, sind die Bedingungen $(\alpha, \beta, \varepsilon)$ notwendig und hinreichend.*

Es braucht nur gezeigt zu werden, daß für jeden Punkt x eine Zerlegung $C = A \dot{+} B$ gemäß (δ) möglich ist, wobei wieder x von a, b verschieden angenommen sei. Sei $(n = 1, 2, 3, \ldots) U_n$ die Umgebung von x mit dem Radius $\frac{1}{n}$, $F_n = C - U_n$ abgeschlossen. Für hinlänglich großes n gehören a, b zu F_n, aber wegen (β) nicht zu derselben Komponente von F_n;

es sei P_n die a enthaltende, Q_n die b enthaltende Komponente von F_n (solange noch $a \varepsilon U_n$, werde $P_n = 0$ gesetzt; analog Q_n). Die Mengen

$$P = \mathfrak{S} P_n, \quad Q = \mathfrak{S} Q_n$$

(wobei die Summanden mit n zunehmen) sind disjunkt und zusammenhängend. Hierbei enthält nach dem Randsatz § 29, XVII P_n einen Punkt auf dem Rande von F_n, also in der Entfernung $\frac{1}{n}$ von x; x ist Häufungspunkt von P und Q; $P + Q + x$ zusammenhängend, als Summe der zusammenhängenden Mengen $A = P + x$, $B = Q + x$. Wenn nun noch gezeigt wird, daß A und B abgeschlossen sind, so ist der Beweis erbracht; denn aus der Irreduzibilität (β) folgt dann $C = P + Q + x = A \dotplus B$, die gewünschte Zerlegung. Es ist also zu zeigen: wenn $c \neq x$ α-Punkt von P ist, so ist $c \varepsilon P$. Sei R_n die c enthaltende Komponente von F_n und $R = \mathfrak{S} R_n$ (schließlich ist $c \varepsilon F_n$; vorher sei wieder $R_n = 0$). Nun ist schließlich c innerer Punkt von F_n und wegen des lokalen Zusammenhangs innerer Punkt von R_n (schon die c enthaltende Komponente von F_{ni} hat c zum inneren Punkt); R_n enthält also Punkte von P. Danach ist $R_n P > 0$, $RP > 0$ und für hinlänglich großes n auch $R_n P_n > 0$, d. h. $R_n = P_n$ und $R = P$, also $c \varepsilon P$.

Damit ist IV bewiesen. Ein kompaktes, lokal zusammenhängendes Kontinuum war immer stetiges Bild einer Strecke; ist es überdies zwischen zwei gewissen Punkten irreduzibel, so ist es homöomorphes Streckenbild.

2. Primteile eines Kontinuums [1]). C sei mehrpunktig, zusammenhängend und werde wieder als Raum angenommen, auf den sich die Relativbegriffe beziehen. Die Punkte r, in denen C lokal zusammenhängend ist, mögen *reguläre* Punkte, die übrigen s *singuläre* Punkte heißen; ist R die Menge der regulären, S die der singulären Punkte, so ist

$$C = R + S = R_i + S_\alpha.$$

Die Punkte (einpunktigen Mengen) *von R_i und die Komponenten von S_α heißen die Primteile von C* (H. Hahn). Falls $R_i = 0$, S in C dicht ist, hat C sich selbst als einzigen Primteil und heiße ein *Primkontinuum.* Ist $R_i > 0$, so ist diese insichdichte Menge lokal zusammenhängend, da einer ihrer Punkte mit einem hinlänglich benachbarten durch eine zusammenhängende Menge $\leqq C$ von beliebig kleinem Durchmesser, also $\leqq R_i$, verbunden werden kann; ihre Komponenten sind in R_i offen und mehrpunktig, R_i also (§ 29, VII) mindestens von der Mächtigkeit $\aleph$. In diesem Fall hat also C mindestens $\aleph$ Primteile und heiße ein *zusammengesetztes Kontinuum*; die Primteile sind selbst Kontinua.

Führen wir für zwei Punkte x, y einen der (immer verschwindenden) Distanz in C analogen Begriff, die *singuläre Distanz* ein, die wir wieder mit $\overline{xy}$ bezeichnen wollen. Eine ϱ-Kette (S. 158)

[1]) Hierzu vgl. Nachtrag C.

(3) $$(x, s_1, s_2, \ldots, s_{n-1}, y),$$

deren Zwischenpunkte $s_1, s_2, \ldots, s_{n-1}$ singulär sind (wozu wir auch für $xy \leqq \varrho$ die Kette (x, y) ohne Zwischenpunkte rechnen), heiße eine *singuläre ϱ-Kette.* Verbinden wir x, y durch alle möglichen singulären ϱ-Ketten; die untere Grenze der hierbei auftretenden Zahlen ϱ sei als die singuläre Distanz $\overline{xy}$ definiert. Es ist $\overline{xy} \leqq xy$ und es gilt das Dreiecksaxiom $\overline{xy} + \overline{yz} \geqq \overline{xz}$, denn wenn sich x und y durch eine singuläre ϱ-Kette, y und z durch eine singuläre σ-Kette verbinden lassen, so x und z durch eine singuläre $\varrho + \sigma$-Kette. Wenn wir statt der singulären ϱ-Ketten solche

(4) $$(x, t_1, t_2, \ldots, t_{n-1}, y)$$

verwenden, deren Zwischenpunkte aus S_α entnommen sind, so ändert sich die untere Grenze der ϱ nicht; denn eine ϱ-Kette (3) ist eine spezielle ϱ-Kette (4), und andererseits gehört zu einer ϱ-Kette (4) eine $\varrho + \delta$-Kette (3), indem man zu jedem t_k ein s_k mit $s_k t_k < \frac{\delta}{2}$ wählt ($\delta > 0$ beliebig klein).

Die singuläre Distanz $\overline{xy}$ verschwindet natürlich für $x = y$; für $x \neq y$ dann und nur dann, wenn beide Punkte zu S_α gehören und in dieser Menge die Distanz 0 haben. Nehmen wir von jetzt an C als *kompaktes* Kontinuum an, so ist $\overline{xy} = 0$ damit gleichbedeutend, daß x, y demselben Primteil angehören.

Durch diese Distanzen wird C zu einem neuen metrischen Raum Γ, worin die Punkte eines Primteils als identisch anzusehen sind. Man kann dies auch so ausdrücken: jedem Punkt $x \varepsilon C$ wird eindeutig ein Punkt $\xi = \varphi(x)$ zugeordnet, mit der Vorschrift, daß dann und nur dann $\varphi(x) = \varphi(y)$ sein soll, wenn x, y demselben Primteil von C angehören; die Menge Γ der Punkte ξ wird eindeutiges Bild von C, während umgekehrt das Urbild von ξ ein ganzer Primteil ist. Durch die Entfernung $\xi\eta = \overline{xy}$ wird Γ zum metrischen Raum und, wegen $\overline{xy} \leqq xy$, *stetiges* Bild von C, also wieder ein kompaktes Kontinuum: nennen wir es das *Kontinuum der Primteile von C.*

Das Urbild eines Kontinuums $\Delta \leqq \Gamma$ ist ein Kontinuum $D \leqq C$. Denn D ist abgeschlossen; wäre $D = D_1 + D_2$ zerstückelbar, so kann kein Primteil (weil er Kontinuum ist) mit beiden Summanden Punkte gemein haben; dann sind die Bilder Δ_1 und Δ_2 von D_1 und D_2 disjunkt und kompakte abgeschlossene Mengen, $\Delta = \Delta_1 + \Delta_2$ zerstückelbar. — Es folgt daraus: *wenn C zwischen x, y irreduzibel ist, so ist Γ zwischen den Bildpunkten ξ, η irreduzibel* (irreduzibel im Sinne von (β), also irreduzibles Kontinuum). Dabei setzen wir $\xi \neq \eta$, x und y verschiedenen Primteilen angehörig voraus; im ausgeschlossenen Fall wäre C Primkontinuum, Γ einpunktig.

Das Bild $\Phi(G)$ einer in C offenen Menge G braucht in Γ nicht offen zu sein. Wenn jedoch G den ganzen Primteil A, das Urbild des Punktes ξ, enthält, so hat $\Phi(G)$ den Punkt ξ zum inneren Punkt. Denn ist $F = C - G$, so ist das Bild $\Phi(F)$ abgeschlossen (als stetiges Bild einer kompakten abgeschlossenen Menge) und enthält ξ nicht, hat also von ξ eine positive untere Entfernung ϱ; alle Punkte η mit $\xi\eta < \varrho$ gehören demnach zu $\Phi(G)$.

Wir wollen nun das eigentliche Ziel dieser Untersuchung erreichen, nämlich beweisen, daß das Kontinuum der Primteile *lokal zusammenhängend* ist.

V. *Ist C kompaktes Kontinuum, F (in C) abgeschlossen, P eine Komponente von F, so ist jeder Punkt x von $F_i P(F - P)_\alpha$ singulär, und der x enthaltende Primteil A trifft den Rand F_r von F.*

Folgendes sei vorausbemerkt. Damit der Satz einen Inhalt habe, d. h. ein Punkt x vorhanden sei, muß $F_i > 0$ und $F > P$ sein, also F unzusammenhängend und insbesondere $F < C$. Da der Raum C zusammenhängend ist, kann F nicht zugleich offen sein, also $F_r > 0$. Der Randsatz § 29, XVII, auf dem hauptsächlich der Beweis des jetzigen Satzes beruht, sagt aus, daß jede Komponente von $CF = F$ den Rand F_r trifft, ferner: wenn Q ein Kontinuum ist, das F_r und F_i trifft, so hat jede Komponente von QF_i Häufungspunkte in F_r.

Nun kann x nicht regulär sein; denn ist $U_x \leqq F_i$, so müßte x innerer Punkt der ihn enthaltenden Komponente von U_x, also erst recht von P sein und wäre nicht Häufungspunkt von $F - P$.

Sei weiter $F = P + \Sigma Q_m$ die Zerlegung von F in Komponenten. Wegen $\delta(x, Q_m) > 0$ muß es, weil x Häufungspunkt von $F - P$ sein soll, eine Folge von Komponenten Q_n mit $\delta(x, Q_n) \to 0$ geben; nach dem Auswahlsatz können wir annehmen, daß der *abgeschlossene Limes* $Q = Fl\, Q_n$ existiert. Er ist wieder ein Kontinuum $\leqq F$ (Satz von Zoretti) und enthält x, ist also in P enthalten, überdies natürlich in $(F - P)_\alpha$; nach dem bereits Bewiesenen sind also alle Punkte (zu denen x gehört) von QF_i singulär, $QF_i \leqq S$. Wie oben bereits vorausbemerkt wurde, treffen alle Q_m den Rand F_r, also ist auch $QF_r > 0$, und die x enthaltende Komponente X von QF_i hat Häufungspunkte in F_r, $X_\alpha F_r > 0$. Da nun $X \leqq S$, $X_\alpha \leqq S_\alpha$ und X_α zusammenhängend ist, ist X_α in dem zu x gehörigen Primteil A enthalten und $AF_r > 0$.

VI. *Ist A ein Primteil des kompakten Kontinuums C, F die Menge der Punkte mit $\delta(x, A) \leqq \varrho\ (\varrho > 0)$, P die A enthaltende Komponente von F, so ist $A \leqq P_i$.*

Denn ist $x \varepsilon A \leqq F_i P$, so darf x nicht zu $(F - P)_\alpha$ gehören, da sonst, nach V, A den Rand von F treffen müßte; es gibt also eine Umgebung $U_x \leqq F_i$, die zu $F - P$ disjunkt ist, d. h. $U_x \leqq P$, x ist innerer Punkt von P.

VII. (Satz von R. L. Moore). *Das Kontinuum der Primteile eines kompakten Kontinuums ist lokal zusammenhängend.*

Sei $\xi \varepsilon \Gamma$, das Urbild von ξ der Primteil A von C; Γ_0 eine Umgebung von ξ, ihr Urbild eine in C offene Menge $G_0 \geqq A$. Ist ϱ positiv und kleiner als die untere Entfernung zwischen A und der Begrenzung[1]) von G_0 (beides sind kompakte Kontinua), so ist die Menge F des Satzes VI in G_0 enthalten; P sei wie dort die A enthaltende Komponente von F und $A \leqq G \leqq P \leqq G_0$, wo $G = P_i$ offen ist. Für die Bilder folgt $\Phi(G) \leqq \Phi(P) \leqq \Phi(G_0) = \Gamma_0$. Hierbei enthält, wie wir oben sahen, $\Phi(G)$ den Punkt ξ als inneren Punkt, und $\Phi(P)$ ist ein Kontinuum $\leqq \Gamma_0$, das ξ als inneren Punkt enthält, womit also der lokale Zusammenhang von Γ in ξ bewiesen ist.

VIII. (Satz von H. Hahn). *Ist C ein kompaktes, zusammengesetztes, zwischen zwei gewissen Punkten irreduzibles Kontinuum, so ist das Kontinuum seiner Primteile mit dem Intervall $T(0 \leqq t \leqq 1)$ homöomorph. T ist stetiges Bild von C in der Weise, daß die Urbilder der Punkte t die Primteile von C sind.*

Denn Γ, das Kontinuum der Primteile, ist kompakt, mehrpunktig, zwischen zwei Punkten irreduzibel und lokal zusammenhängend, also nach IV mit T homöomorph. T ist stetiges Bild von Γ, Γ von C, also T von C, wobei jedem t genau ein Punkt von Γ und ein Primteil von C entspricht.

Ein einfaches Beispiel in der (x_1, x_2)-Ebene ist die aus den beiden Mengen

$$R: \ 0 < x_1 \leqq 1, \quad x_2 = \sin\frac{\pi}{x_1}$$
$$S: \ x_1 = 0, \quad -1 \leqq x_2 \leqq 1$$

gebildete Menge $C = R + S$, wobei zugleich R die Menge der regulären, S die der singulären Punkte ist[2]). Die Primteile sind S und die Punkte von R; die Projektion auf die x_1-Achse liefert $[0, 1]$ als stetiges Bild von C. Zwischen jedem Punkt $a \varepsilon S$ und dem Punkt $b(x_1 = 1, x_2 = 0)$ ist C irreduzibles Kontinuum, zwischen allen andern Punktpaaren ist es reduzibel. Irreduzibel zusammenhängend im Sinne von (γ) ist es auch zwischen a und b nicht, da $a + R$ zusammenhängend ist; dagegen ist die Menge $a + R$ zwischen a und b irreduzibel zusammenhängend und nach II T schlichtes stetiges Bild von ihr (wieder durch Projektion).

[114] § 40. Topologische Räume.

Homöomorphien werden auch *topologische* Abbildungen und homöomorphe Mengen topologisch äquivalent genannt; Eigenschaften, die homöomorphen Mengen gemeinsam sind, heißen topologische Invarianten. So ist die Eigenschaft, ein absolutes $G^{\xi}(\xi \geqq 1)$ zu sein, topologisch invariant, nicht aber absolute Abgeschlossenheit (Vollständigkeit); relative

[1]) Wenn diese verschwindet ($G_0 = C$), kann $\varrho > 0$ beliebig sein.

[2]) Dasselbe leistet die Menge in Fig. 6 (S. 157).

Abgeschlossenheit und Offenheit sind hingegen wieder topologisch invariant, d. h. ist A in E abgeschlossen (offen) und E mit $\bar{E}$, A mit einer entsprechenden Teilmenge $\bar{A}$ homöomorph, so ist $\bar{A}$ in $\bar{E}$ abgeschlossen (offen). Die mathematische Disziplin, die sich mit diesen Dingen beschäftigt, heißt Topologie oder (mit einem von Riemann wieder aufgegriffenen Leibnizschen Ausdruck) Analysis situs. Wir haben in sie nicht tiefer eindringen und nur wenige, allgemein gehaltene Sätze daraus beweisen können. Es ist dies aber der passende Anlaß, in aller Kürze diejenigen Punktmengentheorien zu streifen, die den topologischen Standpunkt von vornherein zur Geltung bringen und nur mit topologisch invarianten Begriffen arbeiten; ein so definierter *topologischer Raum* ist von homöomorphen Räumen in demselben Sinne ununterscheidbar, wie in unserer Theorie ein metrischer Raum von isometrischen Räumen. Dabei handelt es sich nicht bloß um eine formale Umgestaltung der metrischen Theorie, sondern um einen neuen, weiteren Raumbegriff; die *metrisierbaren*, d. h. mit metrischen Räumen homöomorphen Räume bilden nur einen Spezialfall unter den topologischen Räumen. Das Primäre im topologischen Raum E sind die (in E) *abgeschlossenen* und ihre Komplemente, die *offenen* Mengen; hierauf stützt sich der *Stetigkeitsbegriff*, nämlich (§ 35, I): die eindeutige Funktion $\bar{x} = f(x)$, die den Raum E auf den Raum $\bar{E}$ abbildet, heißt stetig, wenn jeder in $\bar{E}$ abgeschlossenen (offenen) Menge als Urbild eine in E abgeschlossene (offene) Menge entspricht. Eine schlichte, beiderseits stetige Abbildung heißt Homöomorphie. Die abgeschlossenen oder offenen Mengen können unerklärt an die Spitze gestellt oder auf verwandte Begriffe (abgeschlossene Hülle, Limespunkt, Häufungspunkt; offener Kern, Umgebung) zurückgeführt werden, immer unter Wahrung des topologisch invarianten Charakters; die nähere Beschaffenheit des Raumes wird dann durch eine passende Auswahl von *Axiomen* geregelt, von denen wir nur die drei Hauptgruppen in Betracht ziehen: *Summen- und Durchschnittsaxiome*, *Trennungsaxiome*, *Mächtigkeitsaxiome.* Solche Axiome können z. B. der metrischen Raumtheorie entlehnt werden, wo sie als beweisbare Theoreme auftreten.

I. *Summen- und Durchschnittsaxiome.* Die abgeschlossenen Mengen sollen unter allen Umständen den Forderungen genügen:

(1) Der Raum E und die Nullmenge 0 ist abgeschlossen.

(2) Die Summe von zwei abgeschlossenen Mengen ist abgeschlossen.

(3) Der Durchschnitt von beliebig vielen abgeschlossenen Mengen ist abgeschlossen.

Hiernach kann die abgeschlossene Hülle A_α als Durchschnitt aller abgeschlossenen Mengen $\geqq A$ (zu denen jedenfalls E gehört) definiert werden; sie hat folgende Eigenschaften:

15*

(a) $0_\alpha = 0$

(b) $A_\alpha \geqq A$

(c) $A_{\alpha\alpha} = A_\alpha$

(d) $(A \dotplus B)_\alpha = A_\alpha \dotplus B_\alpha$.

Man könnte auch diese Eigenschaften als Axiome über die Mengenfunktion A_α an die Spitze stellen (K. Kuratowski), abgeschlossene Mengen durch $A_\alpha = A$ definieren und damit die Sätze über abgeschlossene Mengen beweisen; z. B. folgt (3) daraus, daß wegen (d) A_α eine monotone Mengenfunktion, d. h. mit $A < B$ auch $A_\alpha \leqq B_\alpha$ ist.

Für die offenen Mengen gelten die entsprechenden Sätze:

(1) Der Raum E und die Nullmenge 0 sind offen.

(2) Der Durchschnitt von zwei offenen Mengen ist offen.

(3) Die Summe von beliebig vielen offenen Mengen ist offen.

Die Summe aller offenen Mengen $\leqq A$ heißt der offene Kern A_i und hat analoge Eigenschaften (a, b, c, d) wie A_α. Man kann natürlich auch diese Eigenschaften oder die Sätze über offene Mengen als Axiome an die Spitze stellen. Eine Modifikation davon bilden die Axiome über *Umgebungen.* Bezeichnen wir die den Punkt x enthaltenden offenen Mengen (zu denen jedenfalls E gehört) mit G_x und greifen aus ihnen ein Teilsystem von Mengen U_x derart heraus, daß jedes G_x ein U_x enthält; diese U_x heißen Umgebungen von x, und die Umgebungen aller Punkte bilden ein *volles Umgebungssystem* für den Raum E. Solcher Systeme wird es im allgemeinen verschiedene geben; das größte besteht aus allen offenen Mengen > 0; im metrischen Raume bilden die in § 22 eingeführten „sphärischen" Umgebungen $U_x(\varrho)$ mit positiven Radien ϱ ein solches, aber auch schon dann, wenn man sich auf rationale ϱ oder $\varrho = 1, \frac{1}{2}, \frac{1}{3}, \ldots$ beschränkt; im separablen Raume bilden die „speziellen" Umgebungen V von § 25 ein volles Umgebungssystem; im lokal zusammenhängenden Raum gibt es ein volles System aus zusammenhängenden Umgebungen. Zwei volle Systeme, mit den Umgebungen U_x und V_x, stehen in der Beziehung: jedes U_x enthält ein V_x und jedes V_x ein U_x. Ist der Raum E, mit den Umgebungen U_x, homöomorph mit $\overline{E}$, so geben die Bilder der U_x ein volles Umgebungssystem für $\overline{E}$. Aus den Sätzen über offene Mengen ergeben sich folgende Eigenschaften der Umgebungen:

(A) Jeder Punkt x hat mindestens eine Umgebung U_x; es ist stets $x \varepsilon U_x$.

(B) Zu zwei Umgebungen U_x, V_x desselben Punktes gibt es eine dritte $W_x \leqq U_x V_x$.

(C) Jeder Punkt $y \varepsilon U_x$ hat eine Umgebung $U_y \leqq U_x$.

Nun kann man wieder die Umgebungen als unerklärten Begriff an die Spitze stellen und die Sätze (A, B, C) als *Umgebungsaxiome* postulieren [1]). Offene Mengen G sind dann als Summen von Umgebungen oder als solche Mengen zu definieren, wo jeder Punkt $x \varepsilon G$ eine Umgebung $U_x \leqq G$ hat (hierzu die Nullmenge mitgerechnet). Dann werden die Sätze (1, 2, 3) über offene Mengen beweisbar.

Mit den bisherigen Axiomen, mag man diese oder jene der erwähnten Formen vorziehen, ist der Raum natürlich noch sehr arm an Eigenschaften. Es ist sogar noch nicht einmal dafür gesorgt, daß einpunktige (und damit nach (2) auch endliche) Mengen abgeschlossen sind. Dies kann man entweder durch ein Axiom eben dieses Inhalts oder lieber durch folgende Forderung erreichen:

Zu zwei verschiedenen Punkten x, y gibt es eine abgeschlossene Menge, die x, aber nicht y enthält.

Denn die abgeschlossene Hülle der einpunktigen Menge $\{x\}$ ist dann diese Menge selbst. — In dieser Forderung kann man das Wort abgeschlossen auch durch offen ersetzen; dann geht sie in das erste der folgenden Trennungsaxiome über.

II. *Trennungsaxiome.*

(4) Ist $x_1 \neq x_2$, so gibt es eine offene Menge G_1, die x_1, aber nicht x_2 enthält.

(5) Ist $x_1 \neq x_2$, so gibt es zwei disjunkte offene Mengen G_1, G_2 mit $x_1 \varepsilon G_1$, $x_2 \varepsilon G_2$.

(6) Ist F_2 eine den Punkt x_1 nicht enthaltende abgeschlossene Menge, so gibt es zwei disjunkte offene Mengen G_1, G_2 mit $x_1 \varepsilon G_1$, $F_2 \leqq G_2$.

(7) Sind F_1, F_2 zwei disjunkte abgeschlossene Mengen, so gibt es zwei disjunkte offene Mengen G_1, G_2 mit $F_1 \leqq G_1$, $F_2 \leqq G_2$.

(8) Sind A_1, A_2 zwei disjunkte, in ihrer Summe abgeschlossene Mengen, so gibt es zwei disjunkte offene Mengen G_1, G_2 mit $A_1 \leqq G_1$, $A_2 < G_2$.

Jedes dieser Axiome, dabei (4) festgehalten, ist eine Verschärfung des vorangehenden, und man erhält auf diese Weise topologische Räume zunehmender Spezialisierung. In metrischen Räumen gilt das letzte Trennungsaxiom, also alle (vgl. S. 163; der Beweis reicht auch für (8) aus).

III. *Mächtigkeitsaxiome.* Hier würde es sich vor allem (wenn man die trivialen endlichen Räume ausschließt) um die Beziehungen zum Abzählbaren handeln; wir sprechen sie mit Hilfe der Umgebungen in den beiden *Abzählbarkeitsaxiomen* aus:

(9) Es gibt ein volles Umgebungssystem, wo jeder Punkt höchstens abzählbar viele Umgebungen hat.

(10) Es gibt ein abzählbares volles Umgebungssystem.

[1]) Eine solche Umgebungstheorie wurde in der ersten Auflage dieses Buches durchgeführt.

Das erste ist in jedem metrischen Raume, das zweite in jedem separablen Raume erfüllt.

Das hier gegebene Verzeichnis von Axiomen kann natürlich ergänzt werden; aber die Axiome der Gruppe I und wenigstens eins der Trennungsaxiome sind wohl das Minimum dessen, was man von einem topologischen Raume fordern muß, wenn er nicht gar zu abnorm ausfallen soll. Mit dem zweiten Trennungs- und dem ersten Abzählbarkeitsaxiom, also mit (1, 2, 3, 5, 9) läßt sich schon eine der metrischen annähernd ähnliche Raumtheorie begründen. Zur Metrisierbarkeit des Raumes ist aber das schärfste Trennungsaxiom, also (1, 2, 3, 8, 9) jedenfalls notwendig; es sei hier ohne Beweis erwähnt, daß zur Homöomorphie mit einem separablen metrischen Raume die Gültigkeit der Axiome (1, 2, 3, 6, 10) notwendig *und hinreichend* ist (Satz von P. Urysohn, etwas verschärft von A. Tychonoff).

Eine interessante Kategorie topologischer Räume hat M. Fréchet aufgestellt: sie legt den Begriff der konvergenten Folge oder des Limes (Grenzpunkts) zugrunde. Gewissen Punktfolgen $(x_1, x_2, \ldots)$ des Raumes E sollen eindeutig Punkte x von E zugeordnet sein; eine solche Punktfolge heißt konvergent (nach x) und der zugeordnete Punkt $x = \lim x_n$ ihr Limes. Hierbei sollen die beiden *Limesaxiome* gelten:

(α) Jede konstante Folge $(x, x, x, \ldots)$ konvergiert nach x.

(β) Jede Teilfolge einer nach x konvergenten Folge konvergiert nach x.

Der mit einer solchen Limesdefinition ausgestattete Raum heiße ein *L-Raum.*

Ist $A \leqq E$ und konvergiert eine Punktfolge aus A nach x, so heiße x ein Limespunkt[1]) von A; die Menge der Limespunkte von A sei A_λ. Nach dem Axiom (α) ist $A_\lambda \geqq A$; die Mengen mit $A_\lambda = A$ werden als *abgeschlossen* definiert. Dann gelten die Sätze (1, 2, 3) über abgeschlossene Mengen; das folgt leicht daraus, daß A_λ eine monotone Mengenfunktion und wegen (β) offenbar $(A \dotplus B)_\lambda = A_\lambda \dotplus B_\lambda$ ist. Die abgeschlossene Hülle A_α von A ergibt sich dann nach § 30, 2 als größte Menge der Folge $A^0, A^1, \ldots, A^\omega, \ldots$, die durch

$$A^0 = A, \quad A^{\xi+1} = A^\xi_\lambda, \quad A^\eta = \underset{\xi<\eta}{\mathfrak{S}} A^\xi \quad (\eta \text{ Limeszahl})$$

induktiv definiert ist und mit $A, A_\lambda, A_{\lambda\lambda}, \ldots$ beginnt; übrigens ist, vermöge derselben Überlegung wie im Fall eines Baireschen Funktionensystems (S. 167), jedenfalls $A_\alpha = A^\Omega$, d. h. die größte Menge $A_\alpha = A^\eta$ wird schon für einen Index $\eta \leqq \Omega$ erreicht.

---

[1]) Die Limespunkte sind hier, wie wir sogleich sehen werden, nicht notwendig mit den α-Punkten (Punkten der abgeschlossenen Hülle) identisch, sondern bilden nur einen Teil von ihnen.

Übrigens bilden die *Funktionensysteme*, auf die eben zurückverwiesen wurde, das einfachste Beispiel für L-Räume; ein Beispiel, aus dem zugleich hervorgeht, daß im allgemeinen, anders als in metrischen Räumen, A_λ noch nicht abgeschlossen ist und die Limespunkte nur einen Teil der durch unbegrenzt wiederholte Limesbildung entstehenden α-Punkte bilden. Die Elemente von E („Punkte") seien die in einem Raume T definierten reellen Funktionen $x = x(t)$, und die Konvergenz $x = \lim x_n$ werde durch $x(t) = \lim x_n(t)$ erklärt, d. h. es soll im gewöhnlichen Sinne überall, für jedes $t \varepsilon T$, $x_n(t)$ nach $x(t)$ konvergieren. Eine Menge $A \leqq E$ ist also ein Funktionensystem, eine abgeschlossene Menge ein Bairesches Funktionensystem (S. 167 und § 43), die abgeschlossene Hülle das kleinste Bairesche System über A. Ist nun z. B. T die Menge der reellen Zahlen, A das System der *stetigen* Funktionen x, A_λ das der Funktionen $y = \lim x_n$, $A_{\lambda\lambda}$ das der Funktionen $z = \lim y_n$, so ist $A_\lambda < A_{\lambda\lambda}$; die Funktionen y haben nämlich, wie wir in § 42 sehen werden, immer noch Stetigkeitspunkte, während die Funktionen z schon überall unstetig sein können. Die abgeschlossene Hülle $A_\alpha = A^\eta$ wird hier erst für $\eta = \Omega$ erreicht. Man halte sich das metrische Gegenstück vor Augen: wenn wir das System E der beschränkten Funktionen durch die Entfernung $xy = \sup|x(t) - y(t)|$ zum *metrischen* Raum machen, so bedeutet $\lim x_n = x$ *gleichmässige* Konvergenz von $x_n(t)$ nach $x(t)$ für alle t; dann ist A_λ stets abgeschlossen und $A_\lambda = A_{\lambda\lambda} = \cdots = A_\alpha$.

Eine weitere Abweichung der L-Räume von den metrischen ist, daß in ihnen schon das *zweite Trennungsaxiom* (5) nicht zu gelten braucht (das erste (4) gilt, da die einpunktigen Mengen wegen des Limesaxioms (α) abgeschlossen sind). Es kann sogar in einer ganz flagranten Weise verletzt werden. Nennen wir den Raum E *zerlegbar*, wenn er als Summe $F_1 \dotplus F_2$ von zwei abgeschlossenen Mengen $< E$ darstellbar ist (falls F_1, F_2 überdies disjunkt sind, handelt es sich um eine Zerstückelung, so daß unzusammenhängende Räume gewiß zerlegbar, unzerlegbare gewiß zusammenhängend sind, jedoch nicht umgekehrt). Wir hatten in unserer Theorie niemals Anlaß, diesen Begriff zu erwähnen, weil, abgesehen von einpunktigen Räumen, die natürlich unzerlegbar sind, jeder metrische Raum zerlegbar ist. Sogar ist jeder Raum, in dem das zweite Trennungsaxiom (5) gilt, zwischen zwei beliebigen seiner Punkte $x_1 \neq x_2$ zerlegbar, d. h. so, daß x_1 nur zu F_1 und nicht zu F_2, x_2 nur zu F_2 und nicht zu F_1 gehört: man braucht unter F_1, F_2 nur die Komplemente ($F_1 = E - G_2$, $F_2 = E - G_1$) der in (5) genannten offenen Mengen zu verstehen. — Da offenbar das stetige Bild eines unzerlegbaren Raumes wieder unzerlegbar ist (wie das eines zusammenhängenden wieder zusammenhängend war), so muß dieses Bild, falls es ein metrischer Raum ist, einpunktig sein: insbesondere ist *in einem unzerlegbaren Raum jede reelle stetige Funktion* *konstant.* Bei dieser paradoxen Beschaffenheit der unzerlegbaren Räume ist es besonders merkwürdig, daß ein (mehrpunktiger) *L-Raum unzerlegbar sein kann*, wie folgendes Beispiel zeigt. E sei eine Kreisperipherie, für deren Drehungen um den Mittelpunkt wir einen positiven Sinn festsetzen; das Zeichen φ bedeute eine Drehung um den festen Winkel $2\pi\delta$, wo δ irrational ist; der Punkt x von E gehe hierdurch in x_φ, die Menge $A \leqq E$ in A_φ über. Als konvergent definieren wir erstens die konstanten Folgen $(x, x, x, \ldots)$ mit dem Limes x, zweitens die aus lauter verschiedenen Punkten bestehenden, im gewöhnlichen Sinn (auf Grund der elementargeometrischen Entfernungen) nach einem Punkt x konvergenten Folgen, diesen geben wir aber den Limes x_φ. Die beiden Limesaxiome sind erfüllt. Dann ist $A_\lambda = A \dotplus (A')_\varphi$, wo A' die Menge der Häufungspunkte von A im gewöhnlichen Sinn bedeutet. Wir behaupten nun, daß für $E = A \dotplus B$ mindestens eine der Gleichungen $A_\alpha = E$, $B_\alpha = E$ besteht, womit die Unzerlegbarkeit von E gezeigt ist. Liegen auf jedem Kreisbogen Punkte von A, so ist $A' = E$, $A_\lambda = E$, also erst recht $A_\alpha = E$. Ist das Gegenteil der Fall, so enthält B einen Kreisbogen C (mit Endpunkten). Dann ist $C' = C, C_\lambda = C \dotplus C_\varphi, C_{\lambda\lambda} = C \dotplus C_\varphi \dotplus C_{\varphi\varphi}$ usw., $C_\alpha > C \dotplus C_\varphi \dotplus C_{\varphi\varphi} \dotplus \cdots$ und diese Menge ist die ganze Kreisperipherie, denn wenn x der mittelste Punkt von C ist, so ist die Menge $\{x, x_\varphi, x_{\varphi\varphi}, \ldots\}$ in E dicht (im gewöhnlichen Sinn). Also ist $C_\alpha = E$, $B_\alpha = E$.

Neuntes Kapitel.

Reelle Funktionen.

[115] ## § 41. Funktionen und Urbildmengen.

1. Urbildmengen. Im Raume A (der zunächst übrigens nur eine reine Menge, kein metrischer Raum zu sein braucht) sei eine eindeutige reelle Funktion $f(x)$ definiert, d. h. jedem Punkt $x \varepsilon A$ eine reelle Zahl $f(x)$ zugeordnet. Die Menge der Punkte x, wo $f(x) > y$ (y eine gegebene reelle Zahl), werde wie in § 22 kurz mit $[f > y]$ bezeichnet; ebenso sind Mengen wie $[f \geqq y]$, $[y < f < z]$ u. dgl. zu erklären. Ist B die *Wertmenge* von $f(x)$, d. h. die von $f(x)$ durchlaufene Zahlenmenge, so sind diese Mengen nichts anderes als Urbilder gewisser Teilmengen von B, nach der Ausdrucksweise in § 35. Da reelle Funktionen nach oben und unten verschiedenes Verhalten zeigen [1]), z. B. nach oben beschränkt und nach unten unbeschränkt sein können, so empfiehlt es sich, gleichzeitig etwa die Mengen $[f > y]$ und $[f < y]$ zu betrachten; statt der letzteren nehmen wir lieber

[1]) Ist auch x eine reelle Variable, so kommt noch Unterscheidung zwischen rechts und links in Frage.

ihre Komplemente und nennen die Mengen

$$[f > y], \quad [f \geqq y]$$

die zur Funktion f gehörigen *Urbildmengen* (oder *Lebesgueschen Mengen*). [116]
Zwischen beiden Arten bestehen aber Relationen; es ist

$$(1) \qquad \begin{cases} [f \geqq y] = \mathfrak{D}_n \left[f > y - \dfrac{1}{n}\right] \\ [f > y] = \mathfrak{S}_n \left[f \geqq y + \dfrac{1}{n}\right] \end{cases} \qquad (n = 1, 2, 3, \ldots).$$

Eine Funktion bestimmt ihre Urbildmengen, aber auch umgekehrt; und zwar sieht man aus (1), daß bereits die Mengen $[f > y]$ zur Bestimmung von f ausreichen, denn dann sind auch die Mengen $[f \geqq y]$ und, als Differenzen, die Mengen $[f = y]$ bekannt. Auch die Mengen $[f > y]$ brauchen nicht sämtlich bekannt zu sein, da zwischen ihnen weitere Relationen bestehen; es genügen z. B. die Mengen $[f > r]$ für rationales r, da

$$[f > y] = \mathfrak{S}_{r > y} [f > r].$$

Zwischen den Mengen $[f > r]$ bestehen noch die Relationen

$$[f > r] = \mathfrak{S}_{\varrho > r} [f > \varrho], \quad A = \mathfrak{S}_r [f > r], \quad 0 = \mathfrak{D}_r [f > r],$$

und der Leser wird sich leicht überzeugen, daß ein diesen Bedingungen genügendes System von Mengen $M(r)$ wirklich eine (und nur eine) Funktion f mit $[f > r] = M(r)$ definiert.

Der Zusammenhang zwischen den Eigenschaften einer Funktion und denen ihrer Urbildmengen wird der Hauptgegenstand dieses Kapitels sein. Ein Ergebnis dieser Art ist uns bereits bekannt: ist f eine stetige Funktion, so ist jede Menge $[f > y]$ offen (in A), jede Menge $[f \geqq y]$ abgeschlossen. Hiervon gilt auch die Umkehrung: wenn alle Mengen $[f > y]$ und $[f < z]$ offen sind, so auch ihre Durchschnitte $[y < f < z]$, und da jede eindimensionale offene Menge G Summe von offenen Intervallen ist, so ist das Urbild von BG offen, nach § 35, I also $f(x)$ stetig (B wieder die Wertmenge von f).

Wir betrachten nun ein System von Funktionen f, *die alle in demselben Raum A definiert sein sollen.*

Zu zwei Funktionen f_1, f_2 erhält man, indem man an jeder Stelle x das Maximum und das Minimum der beiden Zahlen $f_1(x), f_2(x)$ betrachtet, die beiden weiteren Funktionen

$$\bar{f} = \max [f_1, f_2], \quad \underline{f} = \min [f_1, f_2].$$

Dann ist offenbar

$$(2) \qquad \begin{cases} [\bar{f} > y] = [f_1 > y] \dotplus [f_2 > y] \\ [\bar{f} \geqq y] = [f_1 \geqq y] \dotplus [f_2 \geqq y] \\ [\underline{f} > y] = [f_1 > y]\,[f_2 > y] \\ [\underline{f} \geqq y] = [f_1 \geqq y]\,[f_2 \geqq y]. \end{cases}$$

Für die Summe

$$f = f_1 + f_2$$

besagt $f > y$ oder $f_1 > y - f_2$ die Existenz einer (von der betrachteten Stelle x abhängigen) rationalen Zahl r mit $f_1 > r > y - f_2$ oder $f_1 > r$, $f_2 > y - r$, woraus man die erste der beiden Formeln

$$(3) \qquad \begin{cases} [f > y] = \mathfrak{S}_r \, [f_1 > r]\,[f_2 > y - r] \\ [f \geqq y] = \mathfrak{D}_r \, [f_1 \geqq r] \dotplus [f_2 \geqq y - r] \end{cases}$$

erhält; die zweite ergibt sich durch entsprechende Behandlung von $[f < y]$ und Komplementbildung.

Ferner sei $f_1, f_2, \ldots$ eine *Funktionenfolge*, ihre *obere* und *untere Grenze* sei

$$(4) \qquad \begin{cases} g = \sup\,[f_1, f_2, \ldots] = \sup f_n \\ h = \inf\,[f_1, f_2, \ldots] = \inf f_n \end{cases}$$

wobei für jedes x die Folge der Zahlen $f_n(x)$ im ersten Fall nach oben, im zweiten nach unten beschränkt sein soll. (Wir betrachten nur endliche
[117] Funktionen, schließen also die uneigentlichen Werte $\pm\infty$ aus, deren Berücksichtigung, wo sie notwendig ist, einer gesonderten Betrachtung vorbehalten bleibt.) Hier gilt

$$(5) \qquad \begin{cases} [g > y] = \mathfrak{S}_n \, [f_n > y] \\ [h \geqq y] = \mathfrak{D}_n \, [f_n \geqq y], \end{cases}$$

während sich die mittleren Formeln (2) nicht auf unendlich viele Funktionen übertragen lassen. Es erscheint nach (5) nicht unpassend, die Funktionen $\sup f_n$ mit f_σ, die Funktionen $\inf f_n$ mit f_δ zu bezeichnen. Die obere Grenze $f_{\sigma\sigma}$ einer Folge von Funktionen f_σ ist wieder ein f_σ, denn statt

$$g = \sup_m g_m, \quad g_m = \sup_n f_{mn}$$

kann man schreiben:

$$g = \sup_{mn} f_{mn} = \sup\,[f_{11}, f_{12}, f_{21}, \ldots]$$

mit Verwandlung der Doppelfolge in eine einfache; ebenso ist $f_{\delta\delta}$ ein f_δ.

[118] Der *obere* und *untere Limes* einer Folge ist bekanntlich so definiert:

$$(6) \qquad \begin{cases} \overline{\lim} f_n = \lim g_n, \quad g_n = \sup\,[f_n, f_{n+1}, f_{n+2}, \ldots] \\ \underline{\lim} f_n = \lim h_n, \quad h_n = \inf\,[f_n, f_{n+1}, f_{n+2}, \ldots]. \end{cases}$$

Im ersten Fall ist f_n nach oben, g_n nach unten beschränkt anzunehmen, im zweiten Fall f_n nach unten, h_n nach oben (diese Voraussetzungen über Beschränktheit sollen künftig nicht immer wiederholt werden). Es ist $g_1 \geqq g_2 \geqq \cdots$, $h_1 \leqq h_2 \leqq \cdots$, so daß man auch

$$\overline{\lim} f_n = \inf g_n, \quad \underline{\lim} f_n = \sup h_n$$

schreiben kann; die g_n sind Funktionen f_σ, $\overline{\lim} f_n$ also ein $f_{\sigma\delta}$, ebenso [119]
$\underline{\lim} f_n$ ein $f_{\delta\sigma}$. Die Grenzfunktion $\lim f_n$ einer konvergenten Folge ist beides zugleich.

Im Falle gleichmäßiger Konvergenz tritt eine wesentliche Vereinfachung ein: *die Grenzfunktion* $\lim f_n$ *einer gleichmäßig konvergenten Folge ist zugleich ein* f_σ *und* f_δ, vorausgesetzt, daß wir die Funktionen f einem System entnehmen, dem zugleich mit f auch $f +$ constans angehört. Denn wenn $|\varphi - f_n| \leqq \varepsilon_n$, $\varepsilon_n \to 0$, so ist $\varphi = \inf (f_n + \varepsilon_n) = \sup (f_n - \varepsilon_n)$.

Um den Aussagen über Urbildmengen eine bequeme Form zu geben, treffen wir folgende Verabredung. Die Mengen M, N sollen gegebene Mengensysteme $\mathfrak{M}$, $\mathfrak{N}$ durchlaufen. Wenn dann $[f >^* y]$ für jedes y ein M ist, so sagen wir: *die Funktion f ist von der Klasse* $(M, *)$. Wenn $[f \geqq y]$ stets ein N ist, sagen wir: *f ist von der Klasse* $(*, N)$. Wenn beides gilt:
f ist von der Klasse (M, N). Z. B. sind die stetigen Funktionen des metri- [120]
schen Raumes A von der Klasse (G, F), wo G die offenen, F die abgeschlossenen Mengen des Raumes A durchläuft, und vice versa. Wenn insbesondere die N die Komplemente $A - M$ der M sind, so sind die Aussagen gleichbedeutend: f ist von der Klasse $(M, *)$, $-f$ ist von der Klasse $(*, N)$.

Nehmen wir jetzt an, daß die M wie die N einen *Ring* bilden (Summe und Durchschnitt zweier M ist ein M, Summe und Durchschnitt zweier N ist ein N). Dann können wir folgende Sätze aussprechen:

I. *Wenn die Funktionen f von der Klasse* $(M, *)$ *sind, so ist*

$$\max[f_1, f_2] \textit{ und } \min[f_1, f_2] \textit{ ebenfalls von der Klasse } (M, *),$$
$$f_\delta = \inf f_n \textit{ von der Klasse } (*, M_\delta),$$
$$f_\sigma = \sup f_n \textit{ und } f_1 + f_2 \textit{ von der Klasse } (M_\sigma, *).$$

II. *Wenn die Funktionen f von der Klasse* $(*, N)$ *sind, so ist*

$$\max[f_1, f_2] \textit{ und } \min[f_1, f_2] \textit{ ebenfalls von der Klasse } (*, N),$$
$$f_\sigma = \sup f_n \textit{ von der Klasse } (N_\sigma, *),$$
$$f_\delta = \inf f_n \textit{ und } f_1 + f_2 \textit{ von der Klasse } (*, N_\delta).$$

Alle diese Behauptungen folgen unmittelbar aus (2) (3) (5); nur bei den mittleren hat man noch (1) heranzuziehen. Sind die f von der Klasse $(M, *)$, so sind sie zugleich von der Klasse $(*, M_\delta)$ und f_δ ist von der Klasse $(*, M_{\delta\delta}) = (*, M_\delta)$. Sind die f von der Klasse $(*, N)$, so auch von der Klasse $(N_\sigma, *)$ und f_σ von der Klasse $(N_{\sigma\sigma}, *) = (N_\sigma, *)$.

Wir wollen ein System von Funktionen f ein *gewöhnliches Funktionensystem* nennen, wenn es folgenden Postulaten genügt:

(α) *Jede konstante Funktion ist ein f.*

(β) *Maximum und Minimum zweier f ist ein f.*

(γ) *Summe, Differenz, Produkt und Quotient* (mit nirgends verschwindendem Divisor) *zweier f ist ein f.*

Wegen der Identitäten

$$|f| = \max[f, -f]$$

$$\frac{\max}{\min}[f_1, f_2] = \tfrac{1}{2}(f_1 + f_2) \pm \tfrac{1}{2}|f_1 - f_2|$$

kann die Forderung (β) durch diese ersetzt werden: *der Betrag jedes f ist ein f.*

Ein gewöhnliches Funktionensystem heiße *vollständig*, wenn es auch noch dem Postulat genügt:

(δ) *Der Limes einer gleichmäßig konvergenten Folge von Funktionen f ist ein f.*

Wir haben dann folgenden Satz:

III. *Die Mengen M, zu denen der ganze Raum A und die Nullmenge gehört, sollen einen σ-Ring*[1]) *bilden, ihre Komplemente $N = A - M$ demgemäß einen δ-Ring. Dann bilden alle Funktionen f der Klasse (M, N) ein vollständiges System.*

Denn (α) ist erfüllt, da A und 0 Mengen M, N sind; (β) gilt, weil nach I II $\max[f_1, f_2]$ und $\min[f_1, f_2]$ von der Klasse (M, N) sind. Ferner ist nunmehr f_σ und $f_1 + f_2$ von der Klasse $(M, *)$, f_δ und $f_1 + f_2$ von der Klasse $(*, N)$, also $f_1 + f_2$ und der gleichmäßige Limes von Funktionen f von der Klasse (M, N); (δ) ist erfüllt. Mit der Summe ist auch die Differenz zweier f ein f, da $-f$ ein f ist. Das Quadrat eines f ist ein f; denn $[f^2 > y]$ ist für $y < 0$ der ganze Raum, für $y \geqq 0$ Summe der beiden Mengen $[f > \sqrt{y}]$ und $[f < -\sqrt{y}]$, also ein M; ebenso $[f^2 \geqq y]$ ein N. Danach ist das Produkt zweier f

$$f_1 f_2 = \left(\frac{f_1 + f_2}{2}\right)^2 - \left(\frac{f_1 - f_2}{2}\right)^2$$

ein f. Mit $f \neq 0$ ist $\frac{1}{f}$ ein f; denn $\left[\frac{1}{f} > y\right]$ ist für $y > 0$ mit $\left[0 < f < \frac{1}{y}\right]$, welches Durchschnitt zweier M ist, identisch, für $y = 0$ mit $[f > 0]$, für $y < 0$ mit der Summe der Mengen $[f > 0]$ und $\left[f < \frac{1}{y}\right]$. Also $\frac{1}{f}$ und $-\frac{1}{f}$ sind von der Klasse $(M, *)$, $\frac{1}{f}$ von der Klasse (M, N). Damit ist auch die Gültigkeit von (γ) bewiesen.

[121] **2. Erweiterung gewöhnlicher Systeme.** Es durchlaufe f ein gewöhnliches Funktionensystem. Sodann bedeute f^* den Limes einer (überall) konvergenten Folge von Funktionen f, speziell g den einer aufsteigenden $(f_1 \leqq f_2 \leqq \cdots)$, h den einer absteigenden $(f_1 \geqq f_2 \geqq \cdots)$, k eine Funktion, die zugleich ein g und ein h ist. In dem Schema

$$f \quad k \quad \begin{matrix} g \\ h \end{matrix} \quad f^*$$

[1]) d. h. einen Ring und ein σ-System (jedes M_σ ist ein M).

ist jede Funktion Spezialfall der rechts folgenden; es ist übrigens mit

$$-f \quad -k \quad \begin{matrix} -h \\ -g \end{matrix} \quad -f^*$$

identisch. Von den auf g und h bezüglichen Tatsachen werden wir künftig meist nur die eine Hälfte beweisen.

Jedes g ist ein $f_\sigma = \sup f_n$, aber auch umgekehrt, weil

$$\sup f_n = \lim \max [f_1, f_2, \ldots, f_n].$$

Zu den k (die gleichzeitig f_σ und f_δ sind) gehören insbesondere die Limites gleichmäßig konvergenter Folgen f_n. Der Limes gleichmäßig konvergenter g_n, h_n, k_n ist ein g, h, k ($\lim g_n$ ist ein $g_\sigma = g$, $\lim h_n$ ein $h_\delta = h$).

Maximum, Minimum, Summe zweier g ist ein g. Denn wenn f_n, f_n' aufsteigend nach g, g' konvergieren, so konvergieren $\max [f_n, f_n']$, $\min [f_n, f_n']$, $f_n + f_n'$ aufsteigend nach $\max [g, g']$, $\min [g, g']$, $g + g'$.

Maximum, Minimum, Summe, Differenz, Produkt zweier f^* ist natürlich ein f^*, aber auch der Quotient (mit nirgends verschwindendem Divisor). Es genügt zu zeigen, daß mit $\varphi \neq 0$ auch $\frac{1}{\varphi}$ ein f^* ist. Ist überall $\varphi > 0$ und $\varphi = \lim f_n$, so kann man auch $f_n > 0$ annehmen, indem man es durch $\max \left[f_n, \frac{1}{n}\right]$ ersetzt, welches nach $\max [\varphi, 0] = \varphi$ konvergiert, und dann ist $\varphi = \lim \frac{1}{f_n}$. Ist $\varphi \gtrless 0$, so ist $\frac{1}{\varphi} = \varphi \cdot \frac{1}{\varphi^2}$ Produkt zweier f^*, also [122]
ein f^*.

Die f^* bilden also ein gewöhnliches System, aber darüber hinaus gilt:

IV. *Der Limes einer gleichmäßig konvergenten Folge von Funktionen f^* ist ein f^*; die f^* bilden ein vollständiges System.*

Zunächst bemerke man: wenn $f^* = \lim f_n$ vom Betrage $\leqq \varepsilon$ ist, so kann man auch $|f_n| \leqq \varepsilon$ voraussetzen, denn $f_n' = \max [f_n, -\varepsilon]$ und $f_n'' = \min [f_n', \varepsilon]$ konvergieren ebenfalls nach f^*. — Nun sei F gleichmäßiger Limes von Funktionen $F_0, F_1, \ldots$, die Funktionen f^* sind; man kann mit Beschränkung auf eine Teilfolge voraussetzen, daß F_m von allen folgenden um höchstens ε_{m+1} abweicht ($m = 0, 1, \ldots$), wo $\varepsilon_1 + \varepsilon_2 + \cdots$ eine konvergente Reihe positiver Zahlen ist, und demnach schreiben

$$F - F_0 = (F_1 - F_0) + (F_2 - F_1) + \cdots$$

oder

$$\varphi = \varphi_1 + \varphi_2 + \cdots$$

wo die φ_m ($m = 1, 2, \ldots$) Funktionen f^* vom Betrage $\leqq \varepsilon_m$ sind; zu zeigen ist, daß φ ein f^* ist. Nach dem oben Gesagten kann man

$$\varphi_m = \lim_n f_{mn}, \quad |f_{mn}| \leqq \varepsilon_m$$

annehmen. Dann konvergiert aber

$$f_n = f_{1n} + f_{2n} + \cdots + f_{nn}$$

nach φ, womit die Behauptung bewiesen ist. In der Tat ist für $n > m$

$$f_n - (f_{1n} + \cdots + f_{mn}) = f_{m+1,n} + \cdots + f_{nn}$$

vom Betrage $\leqq \varepsilon_{m+1} + \cdots + \varepsilon_n < \delta_m$, wenn $\delta_m = \varepsilon_{m+1} + \varepsilon_{m+2} + \cdots$ gesetzt wird; aus

$$f_{1n} + \cdots + f_{mn} - \delta_m < f_n < f_{1n} + \cdots + f_{mn} + \delta_m$$

folgt für $n \to \infty$

$$\varphi_1 + \cdots + \varphi_m - \delta_m \leqq \underline{\lim}\, f_n \leqq \overline{\lim}\, f_n \leqq \varphi_1 + \cdots + \varphi_m + \delta_m$$

und daraus für $m \to \infty$

$$\lim f_n = \varphi_1 + \varphi_2 + \cdots = \varphi.$$

Wir bezeichnen nun die Urbildmengen

$$\begin{array}{lcccccc} & [f > y] & [f \geqq y] & [g > y] & [h \geqq y] & [f^* > y] & [f^* \geqq y] \\ \text{mit} & M & N & P & Q & M^* & N^* \;; \end{array}$$

also bedeutet z. B. M alle Mengen, die als $[f > y]$ erscheinen, wenn f das betrachtete Funktionensystem und y die reellen Zahlen durchläuft, wobei man übrigens y fest, etwa $y = 0$ annehmen kann, da $f - y$ wieder ein f ist.

Definitionsgemäß sind die Funktionen f, g, h, k, f^* von den Klassen (M, N), $(P, *)$, $(*, Q)$, (P, Q), (M^*, N^*). *Es wird sich darum handeln, diese Aussagen, soweit es möglich ist, umzukehren*, zuvor aber die P, Q, M^*, N^* durch die M, N auszudrücken.

Die M, N sind Komplemente voneinander, ebenso die P, Q, ebenso die M^*, N^*. Alle sechs Mengensysteme sind *Ringe*, z. B. ist Summe und Durchschnitt zweier P nach (2) wieder ein P, weil Maximum und Minimum zweier g ein g ist.

Weiter folgt aus I II:

V. *Die Mengen P, Q, M^*, N^* sind Mengen $M_\sigma, N_\delta, Q_\sigma, P_\delta$.*

Denn $g = f_\sigma$ ist von der Klasse $(M_\sigma, *)$, also jedes P ein M_σ; f_σ ist von der Klasse $(P, *)$, also $f_{\sigma\delta}$ von der Klasse $(*, P_\delta)$, letzteres gilt insbesondere von $\overline{\lim}\, f_n$ und erst recht von f^*, also ist jedes N^* ein P_δ. Ebenso sind die andern Behauptungen zu beweisen.

Zur Umkehrung von V schalten wir vier einfache Hilfssätze ein:

(A) *Zu jedem M gibt es ein f, das in M positiv ist und sonst* (in $A - M$) *verschwindet.*

Denn es gibt eine Funktion f mit $M = [f > 0]$; dann erfüllt $f' = \max[f, 0]$ die Forderung. Dasselbe tut übrigens $f'' = \min[f', \varepsilon]$ für $\varepsilon > 0$, d. h. die in (A) geforderte Funktion kann noch beliebig klein $(0 \leqq f \leqq \varepsilon)$ angenommen werden.

(B) *Zu jedem M_σ gibt es eine Funktion F, die Limes einer gleichmäßig konvergenten Folge von Funktionen f, ferner in M_σ positiv ist und sonst verschwindet.*

Sei $M_\sigma = M_1 \dot{+} M_2 \dot{+} \cdots$, $\varepsilon_1 + \varepsilon_2 + \cdots$ eine konvergente Reihe positiver Zahlen; man bestimme nach (A) eine Funktion f_n mit $0 \leqq f_n \leqq \varepsilon_n$, die in M_n positiv ist und sonst verschwindet. Dann ist

$$F = f_1 + f_2 + \cdots$$

eine Funktion der verlangten Art.

(C) *Zu jedem M gibt es ein g, das in M gleich 1 und sonst gleich 0 ist.*

Wir wählen f wie in (A), dann sind z. B.

$$g = \lim_n \frac{nf}{1+nf}, \quad g = \lim_n \min [nf, 1]$$

Funktionen der verlangten Art.

(D) *Zu jedem M_σ gibt es ein g, das in M_σ gleich 1 und sonst gleich 0 ist.*

Ist $M_\sigma = M_1 \dot{+} M_2 \dot{+} \cdots$, so sei gemäß (C) g_n eine Funktion, die in M_n gleich 1 und sonst 0 ist; dann ist

$$g = \sup g_n$$

(welches noch ein g ist) eine Funktion der verlangten Art.

Nunmehr können wir V umkehren:

VI. *Die Mengen $M_\sigma, N_\delta, Q_\sigma, P_\delta$ sind Mengen P, Q, M^*, N^*.*

Nach (D) ist jedes $M_\sigma = [g > 0]$ ein P; natürlich auch jedes N_δ ein Q. Zugleich ist aber $M_\sigma = [g \geqq 1]$ ein N^* (da g ein f^* ist), also jedes P ein N^*, jedes Q ein M^*, jedes Q_σ ein M^*_σ. Wendet man dann (B) an, aber nicht auf die f, sondern auf die f^*, wobei nun die in (B) genannte Funktion F wegen IV ein f^* ist, so folgt, daß jedes $M^*_\sigma = [F > 0]$ ein M^* ist. Also ist jedes Q_σ ein M^*, ebenso jedes P_δ ein N^*.

Es sind also die $\quad P, Q, M^*, N^*$
identisch mit den $\quad M_\sigma, N_\delta, Q_\sigma, P_\delta$

oder (bei Elimination der P, Q) die M^*, N^* mit den $N_{\delta\sigma}$, $M_{\sigma\delta}$; der Übergang von den f zu den f^* induziert für die Urbildmengen den σ- und δ-Prozeß.

3. Umkehrung der Klassensätze. Wir kommen nun zum Hauptsatz der ganzen Theorie, nämlich:

VII. *Die Funktionen der Klasse (P, Q) sind identisch mit den Funktionen v, die das kleinste vollständige System über dem System der f bilden.*

Eine Hälfte der Behauptung folgt aus III. Die P bilden einen Ring, dem (wegen des Postulats (α) für gewöhnliche Systeme) der Raum A und die Nullmenge angehört; sie bilden überdies wegen ihrer Identität mit den M_σ ein σ-System (P_σ ist ein P). Das Entsprechende gilt von ihren Komplementen $Q = A - P$. Die Funktionen der Klasse (P, Q)

bilden demnach ein vollständiges System über dem System der f; das kleinste vollständige System dieser Art muß in ihm enthalten sein, d. h. jede Funktion v ist von der Klasse (P, Q).

Um zu zeigen, daß auch umgekehrt jede Funktion φ der Klasse (P, Q) ein v ist, schicken wir voraus: falls $y_1 < y_2$, so gibt es stets eine Funktion v derart, daß für

$$\begin{array}{lll} \varphi \leqq y_1, & y_1 < \varphi < y_2, & \varphi \geqq y_2 \\ v = 0, & 0 < v < 1, & v = 1 \end{array}$$

ist. Sei nämlich $P_1 = [\varphi > y_1]$, $P_2 = [\varphi < y_2]$. Zu diesen Mengen $P = M_\sigma$ gibt es nach dem Hilfssatz (B) Funktionen v_1, v_2 (Limites gleichmäßig konvergenter Folgen f_n), die in diesen Mengen positiv sind und in ihren Komplementen $Q_1 = [\varphi \leqq y_1]$, $Q_2 = [\varphi \geqq y_2]$ verschwinden. Da $Q_1 Q_2 = 0$, so verschwinden v_1, v_2 nicht zugleich, es ist $v_1 + v_2 > 0$ und die Funktion $v = \frac{v_1}{v_1 + v_2}$ erfüllt die gestellte Forderung [1]). In der Tat ist in den obigen drei Mengen

$$\begin{array}{ccc} Q_1 P_2 & P_1 P_2 & P_1 Q_2 \\ v_1 = 0 & v_1 > 0 & v_1 > 0 \\ v_2 > 0 & v_2 > 0 & v_2 = 0 \\ v = 0 & 0 < v < 1 & v = 1\,. \end{array}$$

Ist nun zunächst φ *beschränkt*, etwa $0 \leqq \varphi \leqq 1$, so wähle man eine (beliebig große) natürliche Zahl n, bestimme für $m = 1, 2, \ldots, n$ eine Funktion v_m so, daß für

$$\begin{array}{lll} \varphi \leqq \frac{m-1}{n}, & \frac{m-1}{n} < \varphi < \frac{m}{n}, & \varphi \geqq \frac{m}{n} \\ v_m = 0, & 0 < v_m < 1, & v_m = 1 \end{array}$$

und setze $v = \frac{1}{n}(v_1 + \cdots + v_n)$ (die Bedeutung von v_1, v_2, v ist anders als oben). Ist, in einem Punkt x, $\frac{m-1}{n} \leqq \varphi \leqq \frac{m}{n}$, so ist dort $v_1 = \cdots = v_{m-1} = 1$, $0 \leqq v_m \leqq 1$, $v_{m+1} = \cdots = v_n = 0$, also $\frac{m-1}{n} \leqq v \leqq \frac{m}{n}$, demnach überall $|\varphi - v| \leqq \frac{1}{n}$. Man kann also φ durch ein v gleichmäßig approximieren, wonach φ selbst ein v ist.

Ist Φ von der Klasse (P, Q), aber nicht beschränkt, so ist

$$\varphi = \frac{\Phi}{1 + |\Phi|} \tag{7}$$

beschränkt $(-1 < \varphi < 1)$ und von derselben Klasse. Denn durch die Beziehung

---

[1]) v braucht nicht mehr gleichmäßiger Limes von Funktionen f zu sein, gehört aber dem kleinsten vollständigen System über den f an.

$$y = \frac{Y}{1+|Y|}, \quad Y = \frac{y}{1-|y|}$$

wird das offene Intervall $-1 < y < 1$ auf die Menge aller reellen Zahlen Y ähnlich (im ordinalen Sinn, also stetig und monoton) abgebildet. Jede Menge $[\varphi > y]$ ist, für $-1 < y < 1$, eine Menge $[\Phi > Y]$ und umgekehrt; zu jenen tritt, für $y \leqq -1$ und $y \geqq 1$, noch der ganze Raum und die Nullmenge hinzu, die aber, wie schon gesagt wurde, ebenfalls Mengen P sind. Mit Φ ist also auch φ von der Klasse $(P, *)$, und umgekehrt; das nämliche gilt von $(*, Q)$. Danach ist φ wieder ein v, folglich aber auch

$$\Phi = \frac{\varphi}{1-|\varphi|} \tag{8}$$

ein v, und der Satz VII ist bewiesen. Diese „Beschränkungstransformation“, die Ersetzung der unbeschränkten Funktion Φ durch die beschränkte φ, wird noch weiterhin anzuwenden sein.

VIII. *Bilden die Funktionen f ein vollständiges System, so bilden die M einen σ-Ring, die N einen δ-Ring, und die Funktionen der Klasse (M, N) sind mit den f identisch.*

Denn Hilfssatz (B), wo F nunmehr ein f ist, zeigt, daß jedes M_σ ein M ist. Die P, Q werden dann mit den M, N identisch, die Funktionen v mit den f, und der Rest der Behauptung folgt aus VII. Der Satz VIII ist gleichzeitig die Umkehrung von III.

Die Voraussetzung der Vollständigkeit trifft nun nach IV jedenfalls für die f^* zu, also:

IX. *Die Funktionen der Klasse (M^*, N^*) sind mit den f^* identisch.*

Suchen wir nun die Umkehrung der Tatsache, daß jedes g von der Klasse $(P, *)$ ist. Definitionsgemäß gibt es zu jedem g ein $f \leqq g$, eine *Minorante* f. Umgekehrt gilt nun auch:

X. *Jede Funktion der Klasse $(P, *)$, die eine Minorante f hat, ist ein g.*

Sei φ von der Klasse $(P, *)$ und $\varphi \geqq f$ oder, wenn wir $1 - f$ statt f schreiben, $\varphi + f > 0$. Auch $\varphi + f$ ist von der Klasse $(P, *)$, wie aus (3) vermöge $P = M_\sigma$ folgt, und wir können also mit veränderter Schreibweise $\varphi > 0$ annehmen. Für ein $\delta > 0$ und $n = 1, 2, \ldots$ suchen wir zu der Menge $[\varphi > n\delta]$, die ein $P = M_\sigma$ ist, nach dem Hilfssatz (D) ein g_n, das in dieser Menge 1 und sonst 0 ist. Die überall konvergente Reihe

$$g = \delta(g_1 + g_2 + \cdots),$$

deren Glieder ja für jedes x schließlich verschwinden, stellt als Limes aufsteigender g ein g dar. Ist in einem Punkt

$$(n-1)\delta < \varphi \leqq n\delta,$$

so ist daselbst

$$g_1 = \cdots = g_{n-1} = 1, \quad g_n = g_{n+1} = \cdots = 0, \quad g = (n-1)\delta$$

und also überall $\qquad 0 < \varphi - g \leqq \delta.$

φ läßt sich also durch Funktionen g gleichmäßig approximieren und ist selbst ein g.

XI. *Jede Funktion der Klasse* $(P, *)$ *ist Limes einer aufsteigenden Folge von Funktionen* v *der Klasse* (P, Q).

Für eine beschränkte Funktion ist dies in X enthalten. Ist Φ unbeschränkt und von der Klasse $(P, *)$, so machen wir wieder die Beschränkungstransformation (7); φ ist von der Klasse $(P, *)$ und beschränkt $(-1 < \varphi < 1)$, also ein g; $\varphi = \lim f_n$ mit $f_1 \leqq f_2 \leqq \cdots$, wobei man $-1 \leqq f_n \leqq \varphi < 1$ annehmen kann, indem man f_n durch $\max\,[f_n, -1]$ ersetzt. Um aber zu Φ zurückzukommen, muß man statt der Funktionen f_n solche haben, welche die untere Grenze -1 nicht erreichen. Setzen wir

$$v_n = \tfrac{1}{2} f_n + \tfrac{1}{4} f_{n+1} + \tfrac{1}{8} f_{n+2} + \cdots,$$

dies ist als Summe einer gleichmäßig konvergenten Reihe von Funktionen f ein v (sogar ein k). Offenbar ist $f_n \leqq v_n \leqq v_{n+1} \leqq \varphi$, v_n konvergiert aufsteigend nach φ. In $v_n \geqq f_n$ kann das Gleichheitszeichen nur gelten, wenn $f_n = f_{n+1} = f_{n+2} = \cdots$, also $v_n = \varphi > -1$; andernfalls ist $v_n > f_n \geqq -1$. Da also stets $-1 < v_n < 1$, so können wir die Funktionen

$$V_n = \frac{v_n}{1 - |v_n|}$$

bilden, die wieder Funktionen v sind und aufsteigend nach $\dfrac{\varphi}{1 - |\varphi|} = \Phi$ konvergieren.

Die entsprechenden Sätze gelten über die Funktionen der Klasse $(*, Q)$; sie sind, wenn sie eine Majorante f haben, Funktionen h, jedenfalls aber Limites absteigender Folgen von Funktionen v. Jede zwischen zwei Funktionen f verlaufende Funktion v der Klasse (P, Q) ist zugleich ein g und h, also ein k. Da also insbesondere jede beschränkte Funktion v ein k ist, so folgt aus der Beschränkungstransformation, daß man jede Funktion v in der Gestalt $\dfrac{k}{1 - |k|}$ mit $|k| < 1$ darstellen kann. Hier ist übrigens $|k| = \max\,[k, -k]$ als Maximum zweier k wieder ein k, ebenso $1 - |k|$, d. h. *die Funktionen* v *sind als Quotienten zweier* k *darstellbar.* Umgekehrt ist jeder Quotient zweier k als Quotient zweier v wieder ein v.

Wenn das System der f vollständig ist, so folgt aus XI, daß jede Funktion der Klasse $(M, *)$ ein g ist. Wir stellen alle Vereinfachungen, die in diesem Falle eintreten, nochmals zusammen:

XII. *Wenn die* f *ein vollständiges System bilden, so sind die* f, g, h, f^* *identisch mit den Funktionen der Klassen* (M, N), $(M, *)$, $(*, N)$, (N_σ, M_δ).

Die k und v sind hier mit den f, die Mengen P, Q, M^*, N^* mit den M, N, N_σ, M_δ identisch.

Bilden die f nur ein gewöhnliches System, so ist XII auf die v anwendbar: die $v, v_\sigma, v_\delta, v^*$ sind identisch mit den Funktionen der Klassen (P, Q), $(P, *)$, $(*, Q)$, (Q_σ, P_δ). Dabei sind nach IX die f^* mit den v^* identisch, während die f, g, h (oder f, f_σ, f_δ) nur einen Teil der v, v_σ, v_δ bilden; die g sind identisch mit den v_σ, die eine Minorante f haben; die h mit den v_δ, die eine Majorante f haben; die k (von denen wieder die f nur ein Teil sind) mit den v, die zwischen zwei Funktionen f verlaufen.

Der Leser möge die bisherigen Betrachtungen noch einmal überblicken und sich etwa folgendes einfaches Beispiel vor Augen halten: die f seien die Funktionen, die nur endlich viele Werte annehmen; dann sind die g die nach unten beschränkten, die h die nach oben beschränkten, die k die beiderseits beschränkten, die v und f^* die völlig willkürlichen Funktionen. Daß in der Tat jede nach unten beschränkte Funktion, etwa $\varphi \geqq 0$, ein g ist (das Umgekehrte ist trivial), sieht man folgendermaßen. In der Menge $R_n = \left\{\frac{0}{n}, \frac{1}{n}, \frac{2}{n}, \ldots, \frac{n^2-1}{n}, \frac{n^2}{n}\right\}$ sei $f_n(x)$ die größte Zahl $\leqq \varphi(x)$; f_n ist ein f und konvergiert nach φ, denn für $n > \varphi$ ist $f_n \leqq \varphi < f_n + \frac{1}{n}$. Da $R_n < R_{2n}$, ist $f_n \leqq f_{2n}$, die Funktionen $f_1, f_2, f_4, f_8, \ldots$ konvergieren aufsteigend nach φ. Die übrigen Behauptungen sind ganz leicht einzusehen. Schon die Mengen M, N und um so mehr die P, Q, M^*, N^* sind die willkürlichen Teilmengen von A. Hiernach verfolge man die Umkehrung der Klassensätze, z. B. daß nicht jede Funktion der Klasse (M, N) ein f ist oder daß in X die Einschränkung, wonach die fragliche Funktion eine Minorante f haben (d. h. nach unten beschränkt sein) soll, nicht wegbleiben darf.

4. Einschiebungs- und Erweiterungssatz. [123]

XIII (Einschiebungssatz). *Wenn g eine Funktion g, h eine Funktion h und überall $g \geqq h$ ist, so gibt es eine Funktion k mit $g \geqq k \geqq h$.*

Wir setzen zur Abkürzung für reelles t

$$\{t\} = \max\,[t, 0] = \tfrac{1}{2}\,|\,t\,| + \tfrac{1}{2}\,t\,; \tag{9}$$

das ist eine stetige Funktion von t, die nichtnegativ ist und mit wachsendem t zunimmt (genauer: nicht abnimmt). Sodann stellen wir folgende Betrachtung an, auf die wir nachher (beim Beweise von XVI) noch einmal zurückgreifen werden. Sei

$$\begin{aligned} \varphi &= \lim \varphi_n, \quad \varphi_1 \leqq \varphi_2 \leqq \cdots \\ \psi &= \lim \psi_n, \quad \psi_1 \geqq \psi_2 \geqq \cdots \end{aligned}$$

und $\varphi \geqq \psi$. Es ist

$$\psi_1 - \varphi_1 \geqq \psi_1 - \varphi_2 \geqq \psi_2 - \varphi_2 \geqq \psi_2 - \varphi_3 \geqq \cdots, \tag{10}$$

daher auch

$$\{\psi_1 - \varphi_1\} \geqq \{\psi_1 - \varphi_2\} \geqq \{\psi_2 - \varphi_2\} \geqq \{\psi_2 - \varphi_3\} \geqq \cdots$$

16*

und diese Funktionen konvergieren nach $\{\psi - \varphi\} = 0$. Die alternierende Reihe

$$(11) \quad \omega = \varphi_1 + \{\psi_1 - \varphi_1\} - \{\psi_1 - \varphi_2\} + \{\psi_2 - \varphi_2\} - \{\psi_2 - \varphi_3\} + \cdots$$

konvergiert daher überall; wir behaupten, daß $\varphi \geqq \omega \geqq \psi$. Unterscheiden wir die Punkte mit $\varphi = \psi$ und $\varphi > \psi$. Ist $\varphi = \psi$, so sind alle Glieder der Folge (10) $\geqq 0$, die geschwungenen Klammern können weggelassen werden und es ist

$$\begin{aligned} \omega &= \varphi_1 + (\psi_1 - \varphi_1) - (\psi_1 - \varphi_2) + (\psi_2 - \varphi_2) - (\psi_2 - \varphi_3) + \cdots \\ &= \lim \varphi_n = \lim \psi_n = \varphi = \psi. \end{aligned}$$

Ist $\varphi > \psi$, so werden die Glieder (10) schließlich negativ. Ist das erste negative Glied $\psi_n - \varphi_n$, also $\varphi \geqq \varphi_n > \psi_n \geqq \psi$, so ist

$$\omega = \varphi_1 + (\psi_1 - \varphi_1) - \cdots - (\psi_{n-1} - \varphi_n) = \varphi_n, \; \varphi \geqq \omega > \psi.$$

Ist das erste negative Glied $\psi_n - \varphi_{n+1}$, also $\varphi \geqq \varphi_{n+1} > \psi_n \geqq \psi$, so ist

$$\omega = \varphi_1 + (\psi_1 - \varphi_1) - \cdots + (\psi_n - \varphi_n) = \psi_n, \; \varphi > \omega \geqq \psi.$$

Um nun XIII zu beweisen, sei $\varphi = g$, $\psi = h$; die Funktionen φ_n, ψ_n seien Funktionen f. Auch die Glieder und Partialsummen der Reihe ω sind Funktionen f; da die Partialsummen mit ungerader Gliederzahl eine aufsteigende, mit gerader eine absteigende Folge bilden, ist ω gleichzeitig ein g und ein h, also ein k.

Bei dem folgenden Satz handelt es sich darum, eine in der Menge $B < A$ definierte Funktion φ zu einer in A definierten ψ zu „erweitern" (was natürlich bedeutet, daß in B $\psi = \varphi$ sein soll). Wir sagen, daß φ von der Klasse $(M, *)$ sei, wenn jede Menge $[\varphi > y]$ der Durchschnitt von B mit einem M ist, und entsprechend für die Klassen $(*, N)$ und (M, N).

XIV (Erweiterungssatz). *Ist Q_0 eine Menge Q, so läßt sich eine in Q_0 definierte Funktion von der Klasse (P, Q) zu einer im ganzen Raum A definierten Funktion von der Klasse (P, Q), d. h. zu einer Funktion v erweitern.*

Es sei $P_0 = A - Q_0$; φ sei in Q_0 definiert, von der Klasse (P, Q) und zunächst beschränkt, etwa $-1 \leqq \varphi \leqq 1$. Wir definieren die Funktion h:

$$h = \varphi \text{ in } Q_0, \quad h = -1 \text{ in } P_0$$

und behaupten, daß dies ein h ist. Sie ist nämlich von der Klasse $(*, Q)$, da $[h \geqq y]$ für $y > -1$ mit $[\varphi \geqq y]$ (welches ein $Q_0 Q$, also ein Q ist) übereinstimmt, für $y \leqq -1$ aber der ganze Raum ist; überdies ist h beschränkt, also nach X ein h. Ebenso schließt man, daß die Funktion g:

$$g = \varphi \text{ in } Q_0, \; g = 1 \text{ in } P_0$$

ein g ist ($-g$ ist ein h). Da $g \geqq h$, läßt sich eine Funktion k einschieben, $g \geqq k \geqq h$; in Q_0 ist $k = \varphi$, φ läßt sich hier speziell zu einem k erweitern.

Wir wollen noch zeigen, daß für $|\varphi| < 1$ auch $|k| < 1$ erreicht werden kann. Bestimmen wir k wie soeben, so ist $|k| \leqq 1$; das Gleichheitszeichen kann nur in Punkten von P_0 gelten. Wir bestimmen nun nach

Hilfssatz (B) zu P_0, welches ein M_σ ist, eine Funktion k_0 (wieder ein k), die in P_0 positiv ist und in Q_0 verschwindet. Die Funktion $v = \frac{k}{1+k_0}$ ist jedenfalls ein v (wegen der Beschränktheit aber immer noch ein k); in Q_0 ist $v = k = \varphi$, in P_0 $|v| < |k|$ oder $v = 0$, also durchweg $|v| < 1$.

Ist endlich Φ in Q_0 definiert, von der Klasse (P, Q) und nicht beschränkt, so machen wir die Beschränkungstransformation (7). Auf φ treffen die vorigen Voraussetzungen zu, es läßt sich zu einem v mit $|v| < 1$ erweitern, und Φ läßt sich zu $V = \frac{v}{1-|v|}$ erweitern, d. h. zu einer Funktion v.

5. Absolute Konvergenz. Die Funktionen f^* lassen sich als Summen konvergenter *Reihen* von Funktionen f darstellen:

$$f^* = \lim f_n = f_1 + (f_2 - f_1) + (f_3 - f_2) + \cdots = \varphi_1 + \varphi_2 + \varphi_3 + \cdots$$

Wenn wir hier statt der einfachen *absolute Konvergenz* fordern (d. h. $|\varphi_1| + |\varphi_2| + \cdots$ soll überall konvergieren), so werden wir nur einen Teil der Funktionen f^* erhalten: bezeichnen wir diese Funktionen mit d. *Sie bilden wieder ein gewöhnliches System.* Denn zunächst ist Summe, Differenz und Produkt zweier d wieder ein d. Sodann ist der Betrag $|d|$ eines d wieder ein d; denn wegen $||\beta| - |\alpha|| \leqq |\beta - \alpha|$ ist mit der Reihe $f_1 + (f_2 - f_1) + \cdots$ auch diejenige absolut konvergent, die aus ihr durch Vertauschung von f_n mit $|f_n|$ entsteht. Es ist nur noch zu beweisen, daß mit $d \neq 0$ auch $\frac{1}{d}$ ein d ist. Sei zunächst $d = \lim f_n > 0$; setzt man $f_n' = \max\left[f_n, \frac{1}{n}\right]$, welche Funktion ebenfalls nach d konvergiert, so erhält man, am raschesten aus der evidenten Ungleichung

$$\begin{aligned}\max[\alpha_1, \beta_1] - \max[\alpha, \beta] &\leqq \max[\alpha_1 - \alpha, \beta_1 - \beta]\\ &\leqq |\alpha_1 - \alpha| + |\beta_1 - \beta|,\end{aligned}$$

die Abschätzung $|f_{n+1}' - f_n'| \leqq |f_{n+1} - f_n| + \left(\frac{1}{n} - \frac{1}{n+1}\right)$; man kann daher f_n durch f_n' ersetzen, also von vornherein $f_n > 0$ annehmen. Dann ist $\frac{1}{d} = \lim \frac{1}{f_n}$ und $\frac{1}{f_{n+1}} - \frac{1}{f_n} = \frac{f_n - f_{n+1}}{f_n f_{n+1}}$ zugleich mit $f_{n+1} - f_n$ das allgemeine Glied einer absolut konvergenten Reihe; $\frac{1}{d}$ ist ein d. Ist endlich $d \gtrless 0$, so ist $\frac{1}{d} = d \cdot \frac{1}{d^2}$ als Produkt zweier d wieder ein d.

Aus der Zerlegung

$$t = \frac{|t| + t}{2} - \frac{|t| - t}{2} = \{t\} - \{-t\}$$

folgt, daß bei absoluter Konvergenz

$$d = \Sigma\varphi_n = \Sigma\{\varphi_n\} - \Sigma\{-\varphi_n\}$$

Differenz von zwei konvergenten Reihen mit Gliedern $\geqq 0$ ist; d. h. jede Funktion d ist als *Differenz* $g - g'$ *zweier Funktionen* g darstellbar, wofür man (mit $g = -h'$, $g' = -h$) auch $h - h'$ oder $g + h$ setzen kann: Differenz zweier Funktionen h oder Summe eines g und eines h. (Wogegen $g - h = g + g'$ wieder ein g, $h - g = h + h'$ wieder ein h ist.) Umgekehrt ist klar, daß die Funktionen g, h und ihre Summen und Differenzen wieder Funktionen d sind. Bei Beschränkung auf absolute Konvergenz entstehen also aus den f die Funktionen

(12) $$d = g - g' = h - h' = g + h.$$

Nennen wir eine reelle Funktion, deren Wertmenge *isoliert* ist, eine *Treppenfunktion*, so gilt:

XV. *Jede Funktion f^*, die zugleich eine Treppenfunktion ist, ist eine Funktion d.*

Dem Beweise schicken wir folgende Bemerkung voraus. Bezeichnen wir diejenigen Mengen, die gleichzeitig M^* und N^* sind, mit R, so ist jede Funktion φ, die eine Treppenfunktion f^* ist, von der Klasse (R, R). Denn die Menge $[\varphi > y]$ ist, wenn φ den Wert y nicht annimmt, mit $[\varphi \geqq y]$, andernfalls für hinreichend kleines $\delta > 0$ mit $[\varphi \geqq y + \delta]$ identisch; ebenso die Menge $[\varphi \geqq y]$ mit $[\varphi > y]$ oder mit $[\varphi > y - \delta]$; beide Mengen sind zugleich M^* und N^*. Die Mengen R bilden offenbar einen Körper. Bei jeder Funktion der Klasse (R, R) sind auch die Mengen $[\varphi = y]$ Mengen R. — Sei nun φ eine Treppenfunktion f^* oder, allgemeiner, eine Funktion der Klasse (R, R) mit höchstens abzählbarer Wertmenge; die verschiedenen Werte, die sie annimmt, seien $c_1, c_2, \ldots$. Die Mengen $[\varphi = c_m]$ sind Mengen R, insbesondere Mengen $M^* = Q_\sigma$, lassen sich also mit aufsteigenden Summanden Q folgendermaßen darstellen:

$$\begin{aligned}[\varphi = c_m] &= Q_{m1} \dot{+} Q_{m2} \dot{+} Q_{m3} \dot{+} \cdots \\ &= Q_{m1} + (Q_{m2} - Q_{m1}) + (Q_{m3} - Q_{m2}) + \cdots \\ &= D_{m1} + D_{m2} + D_{m3} + \cdots.\end{aligned}$$

Man setze nun $c_m = a_m - b_m$, wo die positiven a_m, b_m mit m über alle Grenzen wachsen, z. B.

$$a_m = \tfrac{1}{2}\,|\,c_m\,| + \tfrac{1}{2}\,c_m + m, \quad b_m = \tfrac{1}{2}\,|\,c_m\,| - \tfrac{1}{2}\,c_m + m$$

und definiere die Funktionen g, g' durch

$$g(x) = a_m + n, \quad g'(x) = b_m + n \quad \text{für} \quad x \varepsilon D_{mn},$$

so daß überall $\varphi = g - g'$. Diese Funktionen sind aber Funktionen g. Denn bei gegebenem y kann die Ungleichung $g \leqq y$ nur für endliche viele m $(a_m \leqq y - 1)$ und dann nur für endliche viele $n \leqq y - a_m$ bestehen; es ist also, wenn n_m die größte ganze Zahl $\leqq y - a_m$ bedeutet,

$$[g \leqq y] = \sum_{a_m \leqq y - 1} \sum_{n \leqq n_m} D_{mn} = \sum_{a_m \leqq y - 1} Q_{m n_m}$$

Summe endlich vieler Q, also selbst ein Q (bzw. 0, wenn kein $a_m \leqq y - 1$ existiert). g ist also von der Klasse $(P, *)$ und > 0, also nach X eine Funktion g; dasselbe gilt von g', womit XV bewiesen ist.

Die Funktionen g, h, d lassen sich durch Treppenfunktionen derselben Art beliebig genau approximieren. Sei g eine Funktion g, $\delta > 0$ und $m = 0$, ± 1, $\pm 2, \cdots$; die Menge $[g > m\delta] = P_m$ ist ein P. Definieren wir die Funktion g_0 gleich $m\delta$ in der Menge $P_{m-1} - P_m = [(m-1)\delta < g \leqq m\delta]$. Man erkennt sofort, daß jede Menge $[g_0 > y]$ ein P_m, also g_0 von der Klasse $(P, *)$ ist; da mit g auch $g_0 \geqq g$ eine Minorante f hat, ist g_0 ein g (Satz X). Ferner ist $0 \leqq g_0 - g < \delta$ und g_0 eine spezielle Treppenfunktion, die nur Werte $m\delta$ annimmt. Durch ebensolche Treppenfunktionen h_0 und $d_0 = g_0 + h_0$ kann man die Funktionen h und $d = g + h$ approximieren.

Andererseits lassen sich die allgemeinen Funktionen f^* durch Funktionen d approximieren. Es gilt nämlich wieder ein Einschiebungssatz:

XVI. *Sind φ, ψ zwei Funktionen f^* und überall $\varphi > \psi$, so gibt es eine Funktion $\omega = d$ mit $\varphi \geqq \omega \geqq \psi$.*

Wir können auf den Beweis von XIII zurückgreifen. Nach (6) läßt sich jedes f^* als Limes absteigender g oder aufsteigender h darstellen; nehmen wir die ψ_n des genannten Beweises also als Funktionen g, die φ_n als Funktionen h an. Die Funktionen (10) sind wieder Funktionen g und bleiben es bei Einschließung in geschwungene Klammern, da max $[g, 0]$ ein g ist. Weil wir jetzt die schärfere Voraussetzung $\varphi > \psi$ (nicht $\varphi \geqq \psi$) gemacht haben, werden die Glieder der Folge (10) an jeder Stelle schließlich negativ und die der Reihe (11) schließlich 0; daher konvergieren die beiden Reihen

$$\begin{aligned} g &= \{\psi_1 - \varphi_1\} + \{\psi_2 - \varphi_2\} + \cdots \\ g' &= -\varphi_1 + \{\psi_1 - \varphi_2\} + \{\psi_2 - \varphi_3\} + \cdots \end{aligned}$$

für sich und stellen, als Grenzfunktionen aufsteigender g, wieder Funktionen g dar; dann ist $\omega = g - g'$ und $\varphi \geqq \omega \geqq \psi$, der Satz bewiesen.

Wenden wir ihn insbesondere auf den Fall $\psi = \varphi - \delta$ ($\delta > 0$ konstant) an, so folgt: *jede Funktion f^* läßt sich durch Funktionen d und daher auch durch Treppenfunktionen d beliebig genau approximieren.* Sie läßt sich daher auch durch eine *absolut konvergente* Reihe mit Gliedern d darstellen, d. h. bei Beschränkung auf absolute Konvergenz erreichen wir von den f aus die f^* in zwei Schritten, über die Zwischenstufe der d.

§ 42. Funktionen erster Klasse.

1. Einführung. Wir nehmen den Raum A als metrisch an und identifizieren die Funktionen f des vorigen Paragraphen mit den *stetigen* Funktionen; da sie ein vollständiges System bilden (der gleichmäßige Limes einer Folge stetiger Funktionen ist wieder stetig), so tritt der dortige Satz XII in Kraft. Die M sind mit den *offenen* Mengen G, die N mit den *abgeschlossenen* Mengen F des Raumes identisch (§ 22, I II). Also:

I. *Die stetigen Funktionen f und die Grenzfunktionen g, h, f^* aufsteigender, absteigender, konvergenter Folgen von stetigen Funktionen sind mit den Funktionen der Klassen (G, F), $(G, *)$, $(*, F)$, (F_σ, G_δ) identisch.*

Die f^* werden als *Funktionen der ersten* (B a i r e schen) *Klasse* bezeichnet, was in § 43 auf höhere Klassen ausgedehnt werden wird. Die Funktionen
[124] der Klassen $(G, *)$, $(*, F)$ heißen *halbstetig*, und zwar die der Klasse $(G, *)$ *unterhalb stetig*, die der Klasse $(*, F)$ *oberhalb stetig*; diese Namen sollen sogleich erläutert werden. Mit Rücksicht darauf, daß hier die Mengen P, Q mit den M, N und die Funktionen k mit den f zusammenfallen, nehmen Einschiebungs- und Erweiterungssatz die Form an:

II. *Ist g unterhalb stetig, h oberhalb stetig und überall $g \geqq h$, so gibt es eine stetige Funktion f mit $g \geqq f \geqq h$.*

III. *Eine in der abgeschlossenen Menge F definierte stetige Funktion läßt sich zu einer im ganzen Raume A stetigen Funktion erweitern.*

Die Bezeichnungen „oberhalb und unterhalb stetig" (nach oben, nach unten halbstetig) rühren davon her, daß die Stetigkeitsbedingung hier in zwei Hälften gespalten wird. Eine Funktion $f(x)$ ist an der Stelle a stetig, wenn es für jedes $\sigma > 0$ eine Umgebung U_a gibt, für deren Punkte $x \varepsilon U_a$

$$| f(x) - f(a) | < \sigma .$$

Wenn zu jedem $\sigma > 0$ eine Umgebung U_a existiert, in der

$$f(x) - f(a) < \sigma ,$$

so heißt $f(x)$ im Punkte a *oberhalb stetig*; gibt es stets eine, in der

$$f(x) - f(a) > - \sigma ,$$

so heißt $f(x)$ im Punkte a *unterhalb stetig*. Zur Veranschaulichung bemerken wir, daß man z. B. aus einer stetigen Funktion durch Vergrößerung (Verkleinerung) des Funktionswertes $f(a)$ allein, ohne Änderung der umgebenden Werte $f(x)$, eine oberhalb (unterhalb) stetige Funktion erhält. Ist $f(a) = \pm 1$, sonst überall $f(x) = 0$, so ist $f(x)$ im Punkte $a \begin{cases}\text{oberhalb}\\ \text{unterhalb}\end{cases}$ stetig. Ist $f(x)$ oberhalb stetig, so ist $-f(x)$ unterhalb stetig.

Ist $f(x)$ an der Stelle a unterhalb stetig, so ist a innerer Punkt jeder Menge $[f > y]$, der er angehört, und vice versa. Denn ist $f(a) > y$, und wird $0 < \sigma < f(a) - y$ gewählt, so ist in einer gewissen Umgebung U_a noch $f(x) > f(a) - \sigma > y$, a innerer Punkt von $[f > y]$. Ist umgekehrt die genannte Bedingung erfüllt, so ist a innerer Punkt von $[f > f(a) - \sigma]$ für jedes $\sigma > 0$ und es gibt eine Umgebung U_a, in der $f(x) > f(a) - \sigma$, $f(x)$ ist in a unterhalb stetig. Danach folgt: damit $f(x)$ an jeder Stelle unterhalb stetig sei, ist notwendig und hinreichend, daß jede Menge $[f > y]$

offen sei. Und für eine oberhalb stetige Funktion f, d. h. eine unterhalb stetige $-f$, ist notwendig und hinreichend, daß $[f < y]$ offen, $[f \geqq y]$ abgeschlossen sei. Damit ist die Bezeichnung „unterhalb, oberhalb stetig" für die Funktionen der Klassen $(G, *)$, $(*, F)$ erklärt. Diese Funktionen sind von R. Baire eingeführt worden, der auch zuerst bewiesen hat, daß jede unterhalb stetige Funktion Limes einer aufsteigenden Folge stetiger Funktionen, d. h. in unserer Bezeichnung jede Funktion der Klasse $(G, *)$ ein g ist. Der umgekehrte Schluß, daß jedes g von der Klasse $(G, *)$ ist, gilt hier übrigens in viel weiterem Umfange, als es allgemein der Fall ist; nicht nur die obere Grenze abzählbar vieler, sondern *beliebig* vieler stetiger oder unterhalb stetiger Funktionen

$$g = \sup g_n$$

ist noch unterhalb stetig, da ja

$$[g > y] = \mathfrak{S}\,[g_n > y]$$

als Summe offener Mengen wieder offen ist (an jeder Stelle x sollen natürlich die $g_n(x)$ nach oben beschränkt sein). Man betrachte z. B. zu einer willkürlichen Funktion φ die unterhalb stetigen Minoranten $g \leqq \varphi$; wenn es überhaupt solche gibt, so ist unter ihnen eine größte, nämlich die obere Grenze aller. Entsprechendes gilt für oberhalb stetige Funktionen.

Man wird den Satz I vielleicht durch eine Konstruktionsvorschrift ergänzt wünschen, nach der die Funktionen der Klassen $(G, *)$, $(*, F)$, (F_σ, G_δ) wirklich als Grenzfunktionen aufsteigender, absteigender, konvergenter Folgen von stetigen Funktionen darstellbar sind. Solche Vorschriften, gültig für den allgemeinen Fall beliebiger Ausgangsfunktionen f, sind natürlich in den Beweisen der umgekehrten Klassensätze (§ 41, 3) enthalten; es fragt sich nur, ob sie sich hier nicht vereinfachen lassen. Für den Baireschen Satz, daß eine unterhalb stetige Funktion φ Limes aufsteigender stetiger Funktionen ist, läßt sich in der Tat ein überaus einfacher Beweis geben, wenigstens wenn φ nach unten beschränkt, etwa $\varphi \geqq 0$ ist. Es sei t eine positive Zahl und

$$f(x) = \inf_z\,[\varphi(z) + t \cdot xz],$$

die untere Grenze für alle Punkte z des Raumes genommen; offenbar ist $0 \leqq f \leqq \varphi$ (man erhält $f(x) \leqq \varphi(x)$, indem man $z = x$ wählt). Für zwei Raumpunkte x, y ist

$$\varphi(z) + t \cdot xz \leqq \varphi(z) + t \cdot yz + t \,.\, xy,$$

also, wenn man die unteren Grenzen nach z nimmt,

$$f(x) \leqq f(y) + t \,.\, xy$$

oder, beide Punkte vertauscht, $|f(x) - f(y)| \leqq t \,.\, xy$; die Funktion f ist also stetig. Lassen wir nun t die natürlichen Zahlen durchlaufen, so bilden die Funktionen

$$f_n(x) = \inf_z [\varphi(z) + n \cdot xz]$$

eine aufsteigende Folge mit $f_n \leqq \varphi$, also $g = \lim f_n \leqq \varphi$; wir zeigen, daß auch $g \geqq \varphi$, also $g = \varphi$ ist. Wählen wir bei vorgeschriebenem $\sigma > 0$ den Punkt z_n so, daß

$$f_n(x) > \varphi(z_n) + n \cdot xz_n - \sigma,$$

so ist (wegen $\varphi \geqq f_n$, $\varphi \geqq 0$), $\varphi(x) > n \cdot xz_n - \sigma$, also $xz_n \to 0$; dann ist aber schließlich, weil φ unterhalb stetig ist,

$$\varphi(z_n) > \varphi(x) - \sigma, \quad f_n(x) > \varphi(x) - 2\sigma,$$

also $g(x) \geqq \varphi(x)$.

2. Beispiele von Funktionen erster Klasse. Nächst den unterhalb stetigen Funktionen g und den oberhalb stetigen Funktionen h sind die Funktionen

$$d = g + h = g - g' = h - h'$$

bemerkenswert, die wir allgemein in § 41, 5 besprochen haben; sie entstehen durch Summation absolut konvergenter Reihen von stetigen Funktionen. Jede Funktion erster Klasse läßt sich durch Funktionen d, insbesondere durch Treppenfunktionen d beliebig approximieren.

Jede Funktion mit höchstens abzählbar vielen Unstetigkeitspunkten ist von erster Klasse. Sei C die Menge der Stetigkeitspunkte, D die Menge der Unstetigkeitspunkte von f; M sei die Menge $[f > y]$[1]. Wenn nun $x \varepsilon MC$, so besteht die Ungleichung $f > y$ auch noch in einer gewissen Umgebung von x, also $MC \leqq M_i$, $MD \geqq M_r$. Ist D höchstens abzählbar, so auch M_r, und $M = M_i + M_r$ setzt sich aus einer offenen und einer höchstens abzählbaren Menge zusammen, ist also ein F_σ. Da für $[f < y]$ dasselbe gilt, ist f von der Klasse (F_σ, G_δ).

Unter den Funktionen erster Klasse einer reellen Variablen x sind einerseits die Ableitungen differenzierbarer Funktionen $\varphi'(x) = \lim n\,[\varphi\left(x + \frac{1}{n}\right) - \varphi(x)]$ erwähnenswert; andererseits die Funktionen $f(x)$, bei denen überall die rechts- und linksseitigen Grenzwerte $f(x+0)$, $f(x-0)$ vorhanden sind (so insbesondere die monotonen Funktionen und deren Summen und Differenzen: die Funktionen beschränkter Variation). Diese Funktionen mit einseitigen Grenzwerten haben, wie eine leichte Überlegung lehrt, höchstens abzählbar viele Unstetigkeitspunkte, sind also in der Tat von erster Klasse: übrigens sind sie sogar Funktionen d. Da nämlich $f(x+0)$ als Funktion von x dieselben einseitigen Grenzwerte hat wie $f(x)$, und $\max[f(x), g(x)]$ den rechtsseitigen Grenzwert $\max[f(x+0),\ g(x+0)]$ hat, so gilt für die Funktion

[1]) Wir lassen die Einschränkung, daß der Buchstabe f immer eine stetige Funktion bedeuten soll, jetzt fallen.

$$\varphi(x) = \max\,[f(x),\quad f(x+0),\quad f(x-0)]$$

$\varphi(x \pm 0) = f(x \pm 0) \leqq \varphi(x)$, $\varphi(x)$ ist oberhalb stetig; ebenso ist

$$\psi(x) = \min\,[f(x),\quad f(x+0),\quad f(x-0)]$$

unterhalb stetig; $\varphi + \psi$ ist eine Funktion d. Ersetzt man hierin $f(x)$ durch

$$f(x+0),\quad f(x+0) - f(x),\quad f(x-0) - f(x),$$

so ergeben sich

$$f(x+0) + f(x-0),\quad f(x+0) - f(x),\quad f(x-0) - f(x)$$

als Funktionen d, durch lineare Kombination also auch $f(x)$ selbst.

3. Stetigkeitspunkte. Ist zunächst f eine beliebige im Raume A definierte reelle Funktion, so sei C die Menge ihrer Stetigkeitspunkte, $D = A - C$ die ihrer Unstetigkeitspunkte. *C ist ein G_δ, D ein F_σ* (bezogen auf den Raum A). Denn die Stetigkeit läßt sich auch so charakterisieren: $f(x)$ ist an der Stelle a dann und nur dann stetig, wenn es zu jedem $\sigma > 0$ eine Umgebung U_a gibt, für deren Punkte x, y immer

$$|f(x) - f(y)| < \sigma\,.$$

Für festes σ sei $C(\sigma)$ die Menge der Punkte a, die eine solche Umgebung U_a haben; jeder Punkt $b \varepsilon U_a$ hat dann auch eine solche Umgebung $U_b (\leqq U_a)$, also $U_a \leqq C(\sigma)$, $C(\sigma)$ ist offen. Dann ist C der Durchschnitt aller $C(\sigma)$ für $\sigma > 0$ oder auch

$$C = C(1)\,C(\tfrac{1}{2})\,C(\tfrac{1}{3}) \cdots,$$

C ein G_δ (W. H. Young).

Das eine Extrem bilden die (überall) stetigen Funktionen mit $C = A$, $D = 0$, das andere die überall unstetigen mit $C = 0$, $D = A$; eine solche ist z. B. die sogenannte Dirichletsche Funktion der reellen Variablen x, die für rationales x gleich 1, für irrationales gleich 0 ist. Den stetigen Funktionen am nächsten stehen die Funktionen, für welche *C in A dicht* ist; man nennt sie nach H. Hankel *punktweise unstetig* (oder besser: höchstens punktweise unstetig, unter Hinzurechnung auch der stetigen Funktionen). Nach S. 144 ist dann D, als F_σ mit dichtem Komplement, *von erster Kategorie* in A (ein A_{I}); und wenn also A selbst in sich von zweiter Kategorie (ein A_{II}) ist, z. B. eine Youngsche Menge, so ist C ein A_{II}. Dabei kann auch D in A dicht sein, aber beide Mengen sind von verschiedener Kategorie und können nicht etwa die Rollen tauschen. So kann z. B., wenn A die Menge der reellen Zahlen ist, C die Menge der irrationalen, D die Menge der rationalen Zahlen sein (ist $D = \{r_1, r_2, \ldots\}$ und Σc_n eine konvergente Reihe positiver Zahlen, so ist

$$f(x) = \sum_{r_n < x} c_n$$

eine monotone Funktion dieser Art), nicht aber umgekehrt.

Weiter sei $f_n \to f$ eine überall (in A) konvergente Funktionenfolge; fragen wir, unter welchen Umständen man aus der Stetigkeit der f_n auf die der Grenzfunktion schließen kann. Wir sagen, daß die Folge im Punkte a *uniform* konvergiert, wenn es zu jedem $\sigma > 0$ eine natürliche Zahl m und eine Umgebung U_a gibt derart, daß

(1) $$|f_m(x) - f(x)| \leqq \sigma \qquad (x \varepsilon U_a).$$

Dann gilt:

IV. *Ist $f_n \to f$ und jedes f_n im Punkte a stetig, so ist f dann und nur dann im Punkte a stetig, wenn a ein Punkt uniformer Konvergenz ist.*

Beweis. Ist a ein Punkt uniformer Konvergenz, so erfülle man (1) durch geeignetes m und U_a, wobei man U_a wegen der Stetigkeit von f_m in a zugleich so klein wählen kann, daß

$$|f_m(x) - f_m(a)| \leqq \sigma \qquad (x \varepsilon U_a).$$

Da nach (1) insbesondere

$$|f_m(a) - f(a)| \leqq \sigma$$

ist, so folgt aus den drei Ungleichungen

$$|f(x) - f(a)| \leqq 3\sigma \qquad (x \varepsilon U_a),$$

d. h. die Stetigkeit von f in a.

Ist umgekehrt f in a stetig, so bestimme man zuerst ein m gemäß

$$|f_m(a) - f(a)| \leqq \sigma,$$

dann, auf Grund der Stetigkeit beider Funktionen, ein U_a so, daß

$$|f_m(x) - f_m(a)| \leqq \sigma, \quad |f(x) - f(a)| \leqq \sigma \qquad (x \varepsilon U_a);$$

aus diesen drei Ungleichungen folgt

$$|f_m(x) - f(x)| \leqq 3\sigma \qquad (x \varepsilon U_a),$$

die uniforme Konvergenz in a.

Die Menge K der Punkte uniformer Konvergenz läßt sich so bilden. *Es sei $G(\sigma)$ die Menge der Punkte a, für die sich* (1) *durch geeignetes m und U_a erfüllen läßt.* Man erkennt wieder, daß $G(\sigma)$ offen ($U_a \leqq G(\sigma)$) und K der Durchschnitt aller $G(\sigma)$ für $\sigma > 0$ oder

$$K = G(1)\; G(\tfrac{1}{2})\; G(\tfrac{1}{3}) \cdots$$

ist; die Menge der Punkte uniformer Konvergenz ist wieder ein G_δ.

4. Das Bairesche Theorem. Sind die Funktionen f_n überall stetig, $f \doteq \lim f_n$ Funktion erster Klasse, so fallen die Stetigkeitspunkte von f mit den Punkten uniformer Konvergenz zusammen: $C = K$. Diese Annahme werde nun festgehalten. Sodann sei, bei einem festen $\sigma > 0$, $F_m(\sigma)$ oder kürzer F_m die Menge der Punkte x, die den sämtlichen Ungleichungen

(2) $$|f_n(x) - f_m(x)| \leqq \sigma \qquad \text{für } n > m$$

genügen; für jedes einzelne n definiert diese Ungleichung eine abgeschlossene

Menge, F_m ist deren Durchschnitt für $n = m + 1, m + 2, \ldots$, also wieder abgeschlossen. Wegen der Konvergenz der Folge erfüllt jeder Punkt x die Ungleichungen (2) für hinreichend großes m, und es ist also

$$(3) \qquad A = F_1 \dotplus F_2 \dotplus \cdots$$

Andererseits, wenn a *innerer* Punkt von F_m ist, also die Ungleichungen (2) und die für $n \to \infty$ daraus folgende

$$|f(x) - f_m(x)| \leqq \sigma$$

nicht nur für $x = a$, sondern in einer gewissen Umgebung ($x \varepsilon U_a$) bestehen, so gehört a der vorhin genannten Menge $G(\sigma)$ an, also

$$(4) \qquad G(\sigma) \geqq F_{1i} \dotplus F_{2i} \dotplus \cdots$$

Hieraus folgt: wenn $G(\sigma) = 0$, so ist jedes $F_{mi} = 0$, F_m als abgeschlossene Menge ohne innere Punkte in A *nirgends dicht* und nach (3) A von erster Kategorie in sich. Umgekehrt also: ist A ein A_{II}, so ist jedes $G(\sigma) > 0$.

Weiter: ist A eine G_{II}-Menge (S. 143), also jede (in A) offene Menge $G > 0$ in sich von zweiter Kategorie, so betrachten wir die auf G eingeschränkten Teilfunktionen $f_n(x \mid G)$, $f(x \mid G)$; das vorhin genannte $G(\sigma)$ ist dann durch $GG(\sigma)$ zu ersetzen. Wir finden also: für $G > 0$ ist $GG(\sigma) > 0$, d. h. *$G(\sigma)$ ist in A dicht.* Das abgeschlossene Komplement $F(\sigma)$ dieser Menge ist also in A nirgendsdicht, die Menge

$$D = F(1) \dotplus F(\tfrac{1}{2}) \dotplus F(\tfrac{1}{3}) \dotplus \cdots$$

der Unstetigkeitspunkte von f ein A_{I}, die Menge C der Stetigkeitspunkte in A dicht und ein A_{II}. So haben wir den Satz [1]) gefunden:

V. *Ist der Raum A eine G_{II}-Menge* (d. h. jede offene Menge $G > 0$ von zweiter Kategorie in sich oder in A), *so ist jede Funktion erster Klasse höchstens punktweise unstetig, die Menge ihrer Stetigkeitspunkte in A dicht.*

Das trifft also insbesondere zu, wenn A eine Youngsche Menge, speziell ein vollständiger Raum ist. So kann, wenn A die Menge der reellen Zahlen ist, eine Funktion erster Klasse nur punktweise unstetig sein; die Dirichletsche Funktion $f(x) \left(= \begin{matrix} 1 \\ 0 \end{matrix} \text{ für } \begin{matrix} \text{rationales} \\ \text{irrationales} \end{matrix} \ x \right)$ ist nicht von erster Klasse, wohl aber ist sie von zweiter Klasse (§ 43), d. h. Limes von Funktionen erster Klasse, wie die Formel

$$f(x) = \lim_m \lim_n (\cos m!\, \pi x)^{2n}$$

zeigt.

Der Satz V versagt, wenn die Voraussetzung über A nicht zutrifft. In der abzählbaren Menge $A = \{a_1, a_2, \ldots\}$ von reellen Zahlen ist jede Funktion von erster Klasse, da man, bei willkürlich vorgeschriebenen $b_n = f(a_n)$, immer eine stetige Funktion $f_n(x)$, z. B. ein Polynom, so bestimmen kann, daß $f_n(a_1) = b_1, \ldots, f_n(a_n) = b_n$, also $f_n \to f$. Ist A

[1]) Vgl. hierzu § 45, 3.

die Menge der rationalen Zahlen und etwa $f(x) = 1$ oder 0, je nachdem x dyadisch rational ist oder nicht, so ist f überall unstetig.

Von einer glatten Umkehrung des Satzes V in dem Sinne, daß punktweise unstetige Funktionen auch von erster Klasse seien, kann schon aus Mächtigkeitsgründen keine Rede sein. Ist A etwa die Menge der reellen Zahlen, so gibt es nur $\aleph$ stetige Funktionen und $\aleph^{\aleph_0} = \aleph$ Funktionen erster Klasse; es gibt aber $\aleph^{\aleph} = 2^{\aleph}$ punktweise unstetige Funktionen. Denn ist D eine perfekte nirgendsdichte Menge (von der Mächtigkeit $\aleph$), C ihr offenes, in A dichtes Komplement, so ist jede Funktion, die in C verschwindet, in D nicht verschwindet (also in $x \varepsilon C$ stetig, in $x \varepsilon D$ unstetig ist), punktweise unstetig. — Wenden wir aber V auf die in $f(x) = f(x \mid A)$ enthaltenen Teilfunktionen $f(x \mid B)$ an, die ja wieder in ihrem Raume B Funktionen erster Klasse sind; ist B eine G_{II}-Menge, so liegen die Stetigkeitspunkte der Teilfunktion (die nicht Stetigkeitspunkte der Gesamtfunktion zu sein brauchen) in B dicht. Dies gilt insbesondere, wenn A eine F_{II}-Menge (S. 143) und B in A abgeschlossen, also wieder eine F_{II}-Menge ist, so daß wir aus V folgern können:

VI. *Ist der Raum A eine F_{II}-Menge* (d. h. jede abgeschlossene Menge $F > 0$ in sich von zweiter Kategorie), *f eine Funktion erster Klasse, so enthält jede (in A) abgeschlossene Menge $F > 0$ mindestens einen Stetigkeitspunkt der Teilfunktion $f(x \mid F)$.*

Hierzu gilt nun folgendes Gegenstück:

VII. *Ist der Raum A separabel, f eine in ihm definierte Funktion, und enthält jede (in A) abgeschlossene Menge $F > 0$ mindestens einen Stetigkeitspunkt der Teilfunktion $f(x \mid F)$, so ist f von erster Klasse.*

Wir haben zu beweisen, daß f von der Klasse (F_σ, G_δ) ist, d. h. daß alle Mengen $B = [f > y]$ und $C = [f < z]$ Mengen F_σ sind. Betrachten wir zwei solche mit $y < z$, so daß $A = B \dot{+} C$. Nun sei $F > 0$ abgeschlossen, a ein Stetigkeitspunkt von $f(x \mid F)$. Wenn $a \varepsilon B$, $f(a) > y$, so gibt es also eine Umgebung U_a derart, daß in $F U_a$ auch noch $f > y$, d. h. $F U_a \leqq B$ ist; ebenso gibt es für $a \varepsilon C$ eine Umgebung mit $F U_a \leqq C$, und mindestens einer dieser Fälle tritt ein (sogar beide, wenn $a \varepsilon BC$). Setzen wir $F - F U_a = F_1$, so folgt also: *jede abgeschlossene Menge $F > 0$ enthält eine kleinere abgeschlossene Menge $F_1 < F$ derart, daß die Differenz $F - F_1$ entweder in B oder in C enthalten ist.* Ist noch $F_1 > 0$, so erhalten wir ebenso eine Menge $F_2 < F_1$, und mit transfiniter Fortsetzung dieses Verfahrens definieren wir für alle Ordnungszahlen $\xi < \Omega$ die abgeschlossenen Mengen F_ξ in der uns geläufigen Weise, nämlich (außer $F_0 = F$): ist $F_\xi > 0$, so sei $F_{\xi+1}$ eine Menge $< F_\xi$ derart, daß die Differenz $D_\xi = F_\xi - F_{\xi+1}$ entweder in B oder in C (möglicherweise in beiden Mengen) enthalten ist; für $F_\xi = 0$ sei auch $F_{\xi+1} = 0$; für ine Limeszahl η sei $F_\eta = \mathfrak{D}_{\xi < \eta} F_\xi$.

Da der Raum separabel sein soll, so gibt es (§ 30, IV) eine kleinste Menge $F_\eta = F_{\eta+1} = F_{\eta+2} = \cdots$; sie muß der Konstruktion zufolge Null sein, und damit ist $F = \sum_{\xi<\eta} D_\xi$ in höchstens abzählbar viele Summanden gespalten, die $\leqq B$ oder $\leqq C$ und überdies, als Differenzen abgeschlossener Mengen, spezielle Mengen F_σ sind. Vereinigt man wieder die zu B und die zu C gehörigen Summanden, so erhält man eine Spaltung $F = Y + Z$ in zwei disjunkte Mengen F_σ mit $Y \leqq B$, $Z \leqq C$; insbesondere gestattet der ganze Raum eine solche Spaltung $A = Y + Z$.

Nun endlich halten wir y fest, während z eine Folge $z_1 > z_2 > \cdots$ mit $z_n \to y$ beschreiben möge. Zu $B = [f > y]$ und $C_n = [f < z_n]$ bestimmen wir eine Spaltung des Raumes $A = Y_n + Z_n$ in zwei disjunkte Mengen F_σ mit $Y_n \leqq B$, $Z_n \leqq C_n$. Sei

$$Y = \mathfrak{S} Y_n, \quad Z = \mathfrak{D} Z_n, \quad C = \mathfrak{D} C_n;$$

man findet $C = [f \leqq y] = A - B$. Aus $A = B + C = Y + Z$ und $Y \leqq B$, $Z \leqq C$ folgt dann $Y = B$, $Z = C$; d. h. die Menge $B = [f > y]$ ist ein F_σ. Dasselbe gilt natürlich von den Mengen $[f < z]$ und damit ist VII bewiesen.

Aus VI und VII ergibt sich:

VIII. (Theorem von R. Baire). *Ist der Raum A eine separable* [125]
F_{II}-Menge, so ist die in ihm definierte Funktion f dann und nur dann von erster Klasse, wenn jede (in A) abgeschlossene Menge $F > 0$ mindestens einen Stetigkeitspunkt der Teilfunktion $f(x \mid F)$ enthält.

Man kann hierin das Wort *abgeschlossen* ohne weiteres durch *perfekt* ersetzen, da ein isolierter Punkt von F eo ipso Stetigkeitspunkt von $f(x \mid F)$ ist.

5. Anwendung des Baireschen Theorems. Es sei A die Menge der reellen Zahlen, $f(x)$ also reelle Funktion der reellen Variablen x; in der Ebene E mit rechtwinkligen Koordinaten x, y betrachten wir die durch $y = f(x)$ definierte Menge („Kurve") C. Wir bezeichnen die Punkte der Ebene mit $z = (x, y)$, die von C mit $z = (x, f(x)) = \varphi(x)$. Fragen wir, wann C *zusammenhängend* ist. Dazu ist folgende Bedingung jedenfalls *notwendig*:

(α) *Jeder Punkt von C ist Häufungspunkt von links und rechts*, d. h. zu jedem x gibt es eine Folge $x_n < x$ mit $\varphi(x_n) \to \varphi(x)$ und eine ebensolche Folge $x_n > x$.

Denn spalten wir durch $x < x_0$, $x \geqq x_0$ die Menge C in C_1, C_2, wo C_2 in C abgeschlossen ist, so muß, damit keine Zerstückelung eintrete, C_1 einen Häufungspunkt in C_2 haben, der offenbar kein anderer als $\varphi(x_0)$ sein kann. Daß die Bedingung (α) für sich allein nicht hinreicht, zeigt eine Funktion $f(x)$ wie die Dirichletsche, die nur die Werte 0, 1 annimmt, jeden auf einer in A dichten Menge.

Ist aber $f(x)$ von erster Klasse, so ist (α) *auch hinreichend.*

Nehmen wir an, $C = C_1 + C_2$ sei in zwei relativ abgeschlossene Mengen zerstückelbar; das liefert auf die x-Achse projiziert

$$A = A_1 + A_2 = G_1 + G_2 + F,$$

wo $G_1 = A_{1i}$, $G_2 = A_{2i}$, $F = A_{1\alpha} A_{2\alpha} > 0$ gesetzt ist [1]). Nach VI hat die Teilfunktion $f(x \mid F)$ einen Stetigkeitspunkt x, etwa $x \varepsilon A_1$. Dann kann x nicht Limes von Punkten $x_n \varepsilon A_2 F$ sein, da sonst $\varphi(x) \varepsilon C_1$ Limes von Punkten $\varphi(x_n) \varepsilon C_2$ wäre; es gibt also eine zu $A_2 F$ disjunkte Umgebung U von x. Dann kann U nicht auch noch zu G_2 disjunkt sein, da doch x α-Punkt von $A_2 = A_2 F + G_2$ ist, also ist $U G_2 > 0$. Dies aber geht wegen der Bedingung (α) nicht. Denn ist das offene Intervall (a, b) eine Komponente der offenen Menge $U G_2 < U$, so liegt mindestens einer der Endpunkte, etwa a, in U; da nun zufolge (α) der Punkt $\varphi(a)$ von rechts her Häufungspunkt von C, also von C_2 ist, wäre $a \varepsilon A_2$, $a \varepsilon A_2 - G_2 = A_2 F$ und U doch nicht zu $A_2 F$ disjunkt. Demnach läßt sich C nicht zerstückeln.

Wir können hier noch ein Beispiel (außer § 31, 2) einer zusammenhängenden diskontinuierlichen Menge bilden. Fragen wir, unabhängig vom Vorangehenden, wann C ein mehrpunktiges Kontinuum (bezüglich der Ebene E, also eine zusammenhängende, in E perfekte Menge) enthält. Die Antwort lautet: dann und nur dann, wenn die Stetigkeitspunkte von $f(x)$ ein ganzes Intervall erfüllen. Der eine Teil dieser Aussage ist trivial: ist $f(x)$ für $a \leqq x \leqq b$ stetig, so enthält C eine einfache Kurve. Den andern formulieren wir so: ist K ein mehrpunktiges Kontinuum $\leqq C$ und sind $\varphi(a)$, $\varphi(b)$ zwei Punkte von K ($a < b$), so ist $f(x)$ in $[a, b]$ stetig (in a nach rechts, in b nach links, dazwischen beiderseits).

Seien nämlich K^a, K_a^b, K_b die durch $x \leqq a$, $a \leqq x \leqq b$, $x \geqq b$ definierten Teilmengen von K. Sie sind (in K, also in E) abgeschlossen und nach § 29, III Kontinua, weil K^a und K_a die Summe K und einen einpunktigen Durchschnitt, ebenso K^a und K_a^b die Summe K^b und einen einpunktigen Durchschnitt haben. Die Projektion von K_a^b auf die y-Achse ist zusammenhängend, $f(x)$ nimmt also jeden Wert zwischen $f(a)$ und $f(b)$, falls diese beiden verschieden sind, in (a, b) an; dasselbe gilt von jedem Teilintervall. Dann muß aber $f(x)$ in $[a, b]$ *beschränkt* sein. Denn wäre, wie wir voraussetzen können, $f(x_n) \to + \infty$, $x_n \to x$, $f(x_n) > y > f(x)$, so müßte f zwischen x_n und x an einer Stelle ξ_n den Wert y annehmen; das gäbe $f(\xi_n) = y$, $\xi_n \to x$ und der von $\varphi(x)$ verschiedene Punkt (x, y) müßte der abgeschlossenen Menge K angehören, was nicht stimmt. Also ist K_a^b ein *beschränktes* (kompaktes) Kontinuum. Nun betrachten wir seine Projektion auf die x-Achse; sie ist wieder zusammenhängend, also mit dem ganzen Intervall $[a, b]$ identisch; sie ist schlichtes stetiges Bild

[1]) Die Relativbegriffe ohne Zusatz beziehen sich auf den Raum A.

von K_a^b, also umgekehrt (§ 35, III) auch K_a^b schlichtes stetiges Bild von $[a, b]$, d. h. eben: $f(x)$ ist in $[a, b]$ stetig, im oben angegebenen Sinne.

Demnach ist C dann und nur dann diskontinuierlich (in E), wenn die Unstetigkeitspunkte von $f(x)$ in A dicht liegen. Und C ist zugleich zusammenhängend und diskontinuierlich, wenn $f(x)$ Funktion erster Klasse ist, die Bedingung (α) erfüllt und ihre Unstetigkeitspunkte in A dicht liegen. Eine solche Funktion ist leicht zu bilden. Die durch

$$\sigma(x) = \sin\frac{\pi}{x} \quad (x \neq 0), \quad \sigma(0) = 0$$

erklärte Funktion ist nur an der Stelle $x = 0$ unstetig, erfüllt aber auch dort die Bedingung (α), da $\sigma\left(\frac{1}{n}\right) = \sigma\left(-\frac{1}{n}\right) = 0$ für $n = 1, 2, 3, \ldots$. Ist $\{r_1, r_2, r_3 \ldots\}$ die Menge der rationalen Zahlen, $c_1 + c_2 + c_3 + \cdots$ eine konvergente Reihe positiver Zahlen, so ist, wie aus der gleichmäßigen Konvergenz leicht folgt, die Funktion

$$f(x) = \Sigma c_n \sigma(x - r_n)$$

nur an den rationalen Stellen unstetig, erfüllt aber auch dort die Bedingung (α) und ist Funktion erster Klasse (weil sie nur abzählbar viele Unstetigkeitsstellen hat).

§ 43. Bairesche Funktionen.

1. Bairesche Systeme. Wir betrachten reelle Funktionen, die sämtlich in demselben Raume A (der zunächst eine reine Menge sein kann) definiert sind. Ein System solcher Funktionen f heißt ein *Bairesches System, wenn der Limes jeder konvergenten Folge von Funktionen f wieder dem System angehört.* Über einem gegebenen System Φ von Funktionen f gibt es Bairesche Systeme (z. B. das aller in A definierten Funktionen) und ein kleinstes solches (den Durchschnitt aller Baireschen Systeme $\geqq \Phi$); die Funktionen dieses Systems heißen die von den f erzeugten *Baireschen Funktionen.* Wir können sie zunächst in folgender Form darstellen, der S. 84 bemerkten Darstellung Borelscher Mengen entsprechend. Wir betrachten, indem wir $i, k, l, \ldots$ die natürlichen Zahlen durchlaufen lassen, alle Funktionen der Gestalt

$$g = \lim_i g_i, \quad g_i = \lim_k g_{ik}, \quad g_{ik} = \lim_l g_{ikl}, \ldots$$ [126]

mit der Vorschrift, daß für jede Folge natürlicher Zahlen $i, k, l, \ldots$ in der Folge der Funktionen $g_i, g_{ik}, g_{ikl}, \ldots$ schließlich lauter Funktionen f auftreten sollen. Diese Funktionen g sind mit den von den f erzeugten Baireschen Funktionen b identisch. Denn einerseits bilden die g offenbar ein Bairesches System über Φ, so daß jedes b ein g ist. Andererseits ist auch jedes g ein b; wäre g kein b, so wäre auch mindestens ein g_i, dann ein g_{ik}, ein g_{ikl} usw. vorhanden, das kein b ist, im Widerspruch dazu, daß in der Folge dieser Funktionen schließlich lauter Glieder f auftreten.

Dieser Darstellung ist wiederum eine sukzessive Konstruktion vorzuziehen. Zu den Baireschen Funktionen gehören die Funktionen f selbst, die Funktionen $f^1 = \lim f_n$, die Funktionen $f^2 = \lim f_n^1$ usw. mit endlicher oder unendlicher Wiederholung, die aber nicht über die Ordnungszahlen $\xi < \Omega$ hinaus erstreckt zu werden braucht. Bei der induktiven Definition der f^ξ empfiehlt es sich, um ein möglichst glattes Schlußergebnis zu erzielen, folgendermaßen vorzugehen (ein vollständiges Funktionensystem ist S. 236 erklärt):

Die f^0 sind die Funktionen eines vollständigen Anfangsystems Φ^0;
die $f^{\xi+1}$ sind die Limites konvergenter Folgen von Funktionen f^ξ;
die f^η (η Limeszahl) sind die Funktionen des kleinsten vollständigen Systems, dem alle Funktionen $f^\xi (\xi < \eta)$ angehören.

Demnach folgen z. B. auf die Funktionen $f^0, f^1, f^2, \ldots$ mit endlichen Indizes ξ als Funktionen f^ω noch nicht die Limites konvergenter Folgen von Funktionen f^ξ, sondern erst die Funktionen des kleinsten vollständigen Systems über den f^ξ; diese Abweichung von der sonst üblichen, auch in § 30, 2 provisorisch angenommenen Bezeichnung rechtfertigt sich durch die bessere Übereinstimmung der Funktionen- und Mengenindizes, wie wir sehen werden. Denn die f^ξ bilden nun für *jeden* Index ein vollständiges System (nicht nur die $f^{\xi+1}$, § 41, IV) und damit sind, wie wir wissen, Vereinfachungen verbunden.

Jedes f^ξ ist ein spezielles f^η für $\xi < \eta$. Die f^ξ aller Indizes bilden das kleinste Bairesche System über Φ^0; bei festem ξ bilden sie eine
[127] *Bairesche Funktionenklasse* Φ^ξ, die auch kurz als *Klasse* ξ bezeichnet wird. Die Klassen wachsen mit den Indizes, Gleichheit nicht ausgeschlossen; wenn aber $\Phi^\xi = \Phi^{\xi+1}$, so ist Φ^ξ bereits das ganze Bairesche System und $\Phi^\xi = \Phi^\eta$ für $\eta > \xi$. Eine Funktion gehört *genau* zur Klasse ξ, wenn sie dieser, aber keiner früheren Klasse angehört.

Wir definieren ferner:

g^ξ = Limes einer aufsteigenden, h^ξ = Limes einer absteigenden Folge von Funktionen f^ξ;

außerdem für eine Limeszahl η

g_η = Limes einer aufsteigenden, h_η = Limes einer absteigenden Folge von Funktionen $f^\xi (\xi < \eta)$.

Die Urbildmengen sollen mit

$$M^\xi = [f^\xi > y], \quad N^\xi = [f^\xi \geqq y] \tag{1}$$

bezeichnet werden; wir werden sofort sehen, daß wir keine anderen brauchen.

Identifizieren wir die f des § 41 mit den f^ξ, die ja stets ein vollständiges System bilden, so liefert der dortige Satz XII:

I. *Die Funktionen* f^ξ, g^ξ, h^ξ *sind mit denen der Klassen* (M^ξ, N^ξ), $(M^\xi, *)$, $(*, N^\xi)$ *identisch.*

Außerdem sind die Funktionen $f^{\xi+1}$ (die damaligen f^*) mit denen der Klasse $(N^\xi_\sigma, M^\xi_\delta)$ identisch; das gibt also

$$M^{\xi+1} = N^\xi_\sigma, \quad N^{\xi+1} = M^\xi_\delta. \tag{2}$$

Ferner sei η Limeszahl; wir identifizieren die damaligen f mit den Funktionen f^ξ aller Klassen $\xi < \eta$. Die M sind jetzt die M^ξ, die N sind die N^ξ; die P, Q aber sind die M^η, N^η, da ja die Funktionen f^η jetzt die Rolle der damaligen v spielen. Wir schließen also aus $P = M_\sigma$, $Q = N_\delta$:

$$\left\{\begin{array}{ll} M^\eta = \text{Summe einer Folge von Mengen } M^\xi & (\xi < \eta) \\ N^\eta = \text{Durchschnitt einer Folge von Mengen } N^\xi & (\eta \text{ Limeszahl}). \end{array}\right. \tag{3}$$

Dabei sind, wie aus den Bemerkungen zum Satz XII in § 41, 3 hervorgeht, die Limites konvergenter Folgen von Funktionen f^ξ mit den $f^{\eta+1}$ identisch; die g_η sind diejenigen g^η, die eine Minorante f^ξ haben; die h_η diejenigen h^η, die eine Majorante f^ξ haben; die k_η (die zugleich g_η und h_η sind) diejenigen f^η, die zwischen zwei Funktionen f^ξ verlaufen. Die f^η sind die Quotienten zweier k_η. — Dagegen sind die k^ξ (die zugleich g^ξ und h^ξ sind) stets mit den f^ξ identisch.

Die Formeln (2)(3) bestimmen die M^ξ, N^ξ induktiv, ausgehend von den Mengen $M^0 = [f^0 > y]$, $N^0 = [f^0 \geqq y]$ des Ausgangssystems. Lassen wir den oberen Index 0 weg, so hat man also

$$\left\{\begin{array}{lll} M^0 = M & N^0 = N & \\ M^1 = N^0_\sigma = N_\sigma & N^1 = M^0_\delta = M_\delta & \\ M^2 = N^1_\sigma = M_{\delta\sigma} & N^2 = M^1_\delta = N_{\sigma\delta} & \\ M^3 = N^2_\sigma = N_{\sigma\delta\sigma} & N^3 = M^2_\delta = M_{\delta\sigma\delta} & \\ \quad\cdots & \quad\cdots & \\ M^\omega = \mathfrak{S}\, M^{\xi_n} & N^\omega = \mathfrak{D}\, N^{\xi_n} & (\xi_n < \omega) \\ M^{\omega+1} = N^\omega_\sigma & N^{\omega+1} = M^\omega_\delta & \\ \quad\cdots & \quad\cdots & \end{array}\right. \tag{4}$$

Wir notieren noch den Einschiebungs- und Erweiterungssatz:

II. *Ist* $g^\xi \geqq h^\xi$, *so gibt es eine Funktion* f^ξ *mit* $g^\xi \geqq f^\xi \geqq h^\xi$.

Für Limeszahlen η tritt noch hinzu: ist $g_\eta \geqq h_\eta$, so gibt es eine Funktion k_η mit $g_\eta \geqq k_\eta \geqq h_\eta$.

III. *Ist* N *eine Menge* N^ξ, *so läßt sich eine in* N *definierte Funktion von der Klasse* (M^ξ, N^ξ) *zu einer im ganzen Raume definierten Funktion* f^ξ *erweitern.*

2. Die Baireschen Funktionen des Raumes.

Wählt man insbesondere als Ausgangsfunktionen $f^0 = f$ die im (metrischen) Raume A *stetigen* Funktionen, so heißen die von ihnen erzeugten Baireschen Funktionen die *Baireschen Funktionen* (oder die *analytisch darstellbaren Funktionen*) *des Raumes* A. Die Funktionen M, N sind hier [128]

17*

die G, F, und die M^ξ, N^ξ werden die *Borelschen Mengen* des Raumes A; nach der Definition der G^ξ, F^ξ (§ 32 $(\alpha)(\beta)$) folgt aus (4)

$$(5)\quad \begin{cases} M^0 = G = G^0 & N^0 = F = F^0 \\ M^1 = F_\sigma = F^1 & N^1 = G_\delta = G^1 \\ M^2 = G_{\delta\sigma} = G^2 & N^2 = F_{\sigma\delta} = F^2 \\ M^3 = F_{\sigma\delta\sigma} = F^3 & N^3 = G_{\delta\sigma\delta} = G^3 \\ \dots & \dots \\ M^\omega = G^\omega & N^\omega = F^\omega \\ M^{\omega+1} = F^{\omega+1} & N^{\omega+1} = G^{\omega+1} \\ \dots & \dots \end{cases}$$

und allgemein

$$(6)\quad \begin{cases} M^\xi = G^\xi, \quad N^\xi = F^\xi & (\xi \text{ gerade}) \\ M^\xi = F^\xi, \quad N^\xi = G^\xi & (\xi \text{ ungerade}). \end{cases}$$

Man bemerke hierbei, daß es für eine Limeszahl η ganz gleichgültig ist, ob man M^η als Summe von Mengen M^ξ oder N^ξ oder G^ξ oder $F^\xi (\xi < \eta)$ auffaßt, da ja jedes M^ξ ein spezielles $N^{\xi+1}$ usw. ist.

[129] IV. *Die Baireschen Funktionen des Raumes A sind identisch mit den Funktionen der Klasse (B, B), wo B die Borelschen Mengen des Raumes A durchläuft.*

Denn jede Bairesche Funktion ist von der Klasse (B, B). Ist umgekehrt f von der Klasse (B, B), so betrachte man für rationales r die abzählbar vielen Borelschen Mengen $[f > r]$, $[f \geqq r]$ und wähle die Ordnungszahl ξ so groß, daß alle diese Mengen zugleich G^ξ und F^ξ, also zugleich M^ξ und N^ξ sind. Dann ist für jedes y

$$[f > y] = \mathfrak{S}_{r > y} [f > r]$$

ein $M^\xi_\sigma = M^\xi$ und

$$[f \geqq y] = \mathfrak{D}_{r < y} [f \geqq r]$$

ein $N^\xi_\delta = N^\xi$, also f von der Klasse (M^ξ, N^ξ) oder ein f^ξ.

Aus § 33, I folgt unmittelbar der *Existenzsatz* für Bairesche Funktionen:

V. *In einem vollständigen Raume, dessen insichdichter Kern nicht verschwindet, gibt es für jedes $\xi < \Omega$ Bairesche Funktionen f^ξ, die genau der Klasse ξ (keiner früheren) angehören.*

In einem separablen Raum gibt es $\aleph$ stetige Funktionen (mindestens $\aleph$, weil schon die konstanten Funktionen in dieser Mächtigkeit vorhanden sind, und höchstens $\aleph$, weil eine stetige Funktion bereits durch ihre Werte in einer abzählbaren dichten Teilmenge bestimmt ist). Es gibt dann auch nur $\aleph$ Funktionen f^ξ für jedes ξ, wie induktiv folgt, und $\aleph$ (nämlich höchstens $\aleph_1 \aleph \leqq \aleph\aleph = \aleph$) Bairesche Funktionen überhaupt. Die Baireschen Funktionen bilden also, falls der Raum die Mächtigkeit $\aleph$ hat

(z. B. vollständig ist und einen insichdichten Kern > 0 hat), ein verschwindend kleines Teilsystem im System aller ($\aleph^\aleph = 2^\aleph$) Funktionen.

Daß auch von der Mächtigkeit abgesehen die Baireschen Funktionen den stetigen Funktionen noch recht nahe stehen, nämlich bei „Vernachlässigung“ von Mengen erster Kategorie selbst stetig sind, zeigt der folgende Satz[1]):

VI. *Ist der Raum A eine G_{II}-Menge, so gibt es zu jeder Baireschen Funktion f eine in A dichte Menge $C = G_\delta$ derart, daß die Teilfunktion $f(x \mid C)$ stetig ist.*

Das gilt für die stetigen Funktionen selbst ($C = A$) und wird auf sämtliche Baireschen Funktionen ausgedehnt, indem man es von den Funktionen f_n einer konvergenten Folge auf ihren Limes f überträgt. Sei $f_n(x \mid C_n)$ stetig, C_n ein in A dichtes G_δ, also $D_n = A - C_n$ ein F_σ von erster Kategorie in A; dann ist auch $D_0 = D_1 + D_2 + \cdots$ ein F_σ und A_{I}, $C_0 = A - D_0 = C_1 C_2 \ldots$ ein in A dichtes G_δ und wieder eine G_{II}-Menge (S. 144). Nun sind alle Funktionen $f_n(x \mid C_0)$ stetig, ihr Limes $f(x \mid C_0)$ also nach § 42, V höchstens punktweise unstetig, d. h. die Menge C seiner Stetigkeitspunkte in C_0 dicht und ein G_δ in C_0, also in A dicht und ein G_δ in A; $f(x \mid C)$ ist stetig.

Z. B. ist die Dirichletsche Funktion f (S. 253) in der Menge C der irrationalen Zahlen konstant $= 0$, also stetig, d. h. die irrationalen Zahlen sind Stetigkeitsstellen der Teilfunktion $f(x \mid C)$, nicht der Gesamtfunktion.

Wir drücken VI kurz so aus: jede Bairesche Funktion ist *bis auf Mengen erster Kategorie stetig*; hierunter verstehen wir, daß der Raum eine Spaltung $A = C + D$ zuläßt, bei der D ein A_{I} und $f(x \mid C)$ stetig ist. Man kann dann eo ipso C als G_δ und D als F_σ annehmen, indem man die nirgendsdichten Mengen, deren Summe D ist, durch ihre abgeschlossenen Hüllen ersetzt.

Der Satz VI hat eine gewisse Ähnlichkeit mit dem Satz § 42, V über Funktionen erster Klasse und gestattet auch ein Analogon zu § 42, VI:

Ist der Raum A eine F_{II}-Menge, f eine Bairesche Funktion, so ist für jede (in A) abgeschlossene Menge $F > 0$ auch die Teilfunktion $f(x \mid F)$ bis auf Mengen erster Kategorie stetig.

D. h. es gibt eine Spaltung $F = C + D$, wo D ein F_{I} und $f(x \mid C)$ stetig ist. Das folgt daraus, daß F wieder eine F_{II}-Menge, also (S. 144) eine G_{II}-Menge und $f(x \mid F)$ eine Bairesche Funktion des Raumes F ist.

Man könnte nun, dem Satze § 42, VII entsprechend, vermuten, daß etwa im separablen Raum die obige Stetigkeitseigenschaft auch hinreichend sei: diese Vermutung ist aber falsch. Zunächst bemerke man, daß die Stetigkeitseigenschaft nicht abgeschwächt wird, wenn man sie, statt für alle *abgeschlossenen*, nur für alle *perfekten* Mengen fordert: d. h. sie besteht

[1]) Vgl. hierzu § 45, 3.

dann auch für alle abgeschlossenen. Sei nämlich

$$F = F_j + F_\beta = F_j + (F_s - F_j) + F_k$$

abgeschlossen, $F_j > 0$ der isolierte, F_s der separierte Teil, F_k der insichdichte (perfekte) Kern, F_β die Ableitung. Die Ableitung $F_{j\beta} \leqq F_\beta$ des isolierten Teils ist in F nirgends dicht (ihr Komplement $F - F_{j\beta} \geqq F_j$ ist in F dicht, da seine abgeschlossene Hülle $\geqq F_{j\alpha} \dotplus (F - F_{j\beta}) = F$ ist), also ein F_{I}; sie besteht aus $F_s - F_j$ und denjenigen Punkten von F_k, die Häufungspunkte von F_j sind. Da nun vorausgesetztermaßen $f(x \mid F_k)$ bis auf Mengen erster Kategorie stetig ist, so gibt es eine Menge D, von erster Kategorie in F_k und erst recht in F, derart, daß $f(x \mid F_k - D)$ stetig ist. Entfernen wir nun aus F die Menge erster Kategorie $F_{j\beta} \dotplus D$, so bleibt $F_j + C$ übrig, wo $f(x \mid C)$ stetig ist und C keinen Häufungspunkt von F_j enthält; infolgedessen ist auch $f(x \mid F_j + C)$ in den Punkten von C und natürlich auch in den (isolierten) Punkten von F_j stetig, d. h. $f(x \mid F)$ bis auf Mengen erster Kategorie stetig. (Im Falle $F_k = 0$ ist $f(x \mid F_j)$ stetig, F_β ein F_{I}).

[130] Wir zeigen nun nach N. Lusin: es kann (im Raum A der reellen Zahlen) für jede perfekte Menge $P > 0$ $f(x \mid P)$ bis auf Mengen erster Kategorie stetig sein, während f keine Bairesche Funktion ist. Wir konstruieren nämlich eine *unabzählbare* Menge L derart, daß LP stets ein P_{I} ist. Dies als geschehen vorausgesetzt, sehen wir: L ist keine Borelsche (auch keine Suslinsche) Menge, sonst müßte sie eine perfekte Menge $P > 0$ enthalten und $LP = P$ wäre ein P_{II}. Die charakteristische Funktion f von L ($= 1$ in L, $= 0$ in $A - L$) ist also keine Bairesche Funktion, während doch stets $f(x \mid P)$ bis auf Mengen erster Kategorie stetig, nämlich $f(x \mid P - LP) = 0$ stetig ist.

Um eine solche Menge L zu bilden, betrachten wir Folgen natürlicher Zahlen

$$X = (x_1, x_2, \ldots, x_n, \ldots)$$

und definieren $X < Y$ (X „final" kleiner als Y), wenn schließlich, für $n \geqq n_0$, $x_n < y_n$ ist; die kleinste Zahl n_0, wofür dies der Fall ist, werde mit $n(X, Y)$ bezeichnet. Diese Relation ist transitiv: wenn $X < Y$, $Y < Z$, ist auch $X < Z$ und zwar offenbar

$$n(X, Z) \leqq \max\,[n(X, Y),\ n(Y, Z)].$$

Es gibt zu jeder endlichen oder abzählbaren Menge von Zahlenfolgen $X, Y, Z, \ldots$ eine final größere U; man braucht ja nur

$$u_1 > x_1, \quad u_2 > \max\,[x_2, y_2], \quad u_3 > \max\,[x_3, y_3, z_3], \ldots$$

zu wählen. Demgemäß kann man eine Menge von Zahlenfolgen $X_0 < X_1 < \cdots < X_\omega < \cdots$ vom Typus Ω konstruieren.

Andererseits ordnen wir jeder Zahlenfolge X den Kettenbruch

$$t = \frac{1|}{|x_1} + \frac{1|}{|x_2} + \frac{1|}{|x_3} + \cdots$$

zu (t irrationale Zahl zwischen 0 und 1), den obigen Folgen X_ξ die Zahlen t_ξ ($\xi < \Omega$). Dann hat die Menge $L = \{t_0, t_1, \ldots, t_\omega, \ldots\}$ die geforderten Eigenschaften.

Schreiben wir kurz $n_{\xi\eta}$ für $n(X_\xi, X_\eta)(\xi < \eta)$ und bemerken: wenn eine Folge von Zahlen t_η nach t_ξ konvergiert, so strebt die Anzahl der gemeinsamen Anfangsziffern beider Kettenbrüche und folglich auch $n_{\xi\eta}$ nach ∞. Wir wählen nun ein $\eta > 0$ und setzen

$$M = \{t_0, \ldots, t_\eta\}, \quad N = \{t_{\eta+1}, t_{\eta+2}, \ldots\},$$

$L = M + N$; ferner spalten wir die Menge N der Zahlen t_ζ ($\zeta > \eta$) in $N = N_1 + N_2 + \cdots$, wo N_n die Menge der t_ζ mit $n_{\eta\zeta} = n$ ist. *Dann ist kein Punkt von M Häufungspunkt von N_n.* Denn für $\xi < \eta$ ist

$$n_{\xi\zeta} \leqq \max [n_{\xi\eta}, n_{\eta\zeta}],$$

also für $t_\zeta \varepsilon N_n$

$$n_{\xi\zeta} \leqq \max [n_{\xi\eta}, n],$$

$n_{\xi\zeta}$ beschränkt, so daß t_ζ nicht nach t_ξ konvergieren kann. Dasselbe gilt für $\xi = \eta$. — Wir können nun (L separabel) η so groß wählen, daß M in L dicht ist, dann ist N_n in L nirgendsdicht. Denn es ist $MN_{n\alpha} = 0$ und $L - LN_{n\alpha} \geqq M$ in L dicht. Also ist N von erster Kategorie in L, d. h. es gibt eine Spaltung $L = M + N$, wo M abzählbar und N ein L_{I} ist.

Endlich sei $P > 0$ perfekt. Ist LP unabzählbar, so ist dies eine Menge wie L selbst (einer Menge von Zahlenfolgen $X_{\xi_0} < X_{\xi_1} < \cdots < X_{\xi_\omega} < \cdots$ entsprechend) und gestattet eine Spaltung $LP = M + N$, wo N von erster Kategorie in LP, also in P, und M abzählbar, also auch von erster Kategorie in P ist (jede abzählbare Menge $\leqq Q_h$ ist ein Q_{I}, S. 143): LP ist ein P_{I}. Dasselbe gilt, wenn LP höchstens abzählbar ist. Damit ist die Betrachtung zu Ende geführt: eine Funktion f braucht keine Bairesche Funktion zu sein, auch wenn stets $f(x \mid F)$ bis auf Mengen erster Kategorie stetig ist.

Es sei noch erwähnt, daß das Komplement $K = A - L$ der Lusinschen Menge eine F_{II}-Menge ist, d. h. (S. 143) jede in K perfekte Menge > 0 ist in sich von zweiter Kategorie. Eine solche Menge, die also insichdicht und in K abgeschlossen ist, ist von der Form KP, wo P (in der Menge A der reellen Zahlen) perfekt ist; wegen $P = KP + LP$ ist KP in P, also erst recht in sich selbst von zweiter Kategerie. Es gibt also F_{II}-Mengen, die keine absolut Borelschen Mengen sind, womit eine früher (S. 144) ausgesprochene Behauptung, daß eine F_{II}-Menge keine Youngsche Menge zu sein braucht, in viel weiterem Umfang bewiesen ist.

Wenn eine Funktion f (mag sie eine Bairesche Funktion sein oder nicht) im Raume A, der als G_{II}-Menge gedacht ist, bis auf Mengen erster Kategorie stetig ist, d. h. mindestens eine Spaltung $A = C + D$ existiert, wo D ein A_{I} und $f(x \mid C)$ stetig ist, so gibt es (W. Sierpiński) unter allen solchen Spaltungen $A = C_1 + D_1 = C_2 + D_2 = \cdots$ eine mit größtem C und kleinstem D, nämlich

$$C = C_1 \dotplus C_2 \dotplus \cdots, \quad D = D_1 D_2 \ldots .$$

Es braucht nur gezeigt zu werden, daß $f(x \mid C)$ stetig ist (D ist natürlich ein A_{I}). Sei $x \varepsilon C$ und etwa $x \varepsilon C_1$; da $f(x \mid C_1)$ stetig ist, so gibt es bei beliebigem $\sigma > 0$ eine Umgebung U_x mit

$$(7) \qquad | f(z) - f(x) | < \sigma \quad \text{für} \quad z \varepsilon C_1 U_x .$$

Sodann sei $y \varepsilon C U_x$; ist etwa $y \varepsilon C_2$ (wo auch $C_2 = C_1$ sein kann), so gibt es eine Umgebung U_y mit

$$(8) \qquad | f(z) - f(y) | < \sigma \quad \text{für} \quad z \varepsilon C_2 U_y .$$

Die offene Menge $U_x U_y > 0$ (sie enthält y) ist ein A_{II}, also nicht in $D_1 \dotplus D_2$ enthalten; es ist also $U_x U_y C_1 C_2 > 0$, es gibt mindestens ein z, das (7) und (8) zugleich erfüllt, danach

$$| f(x) - f(y) | < 2\sigma \quad \text{für} \quad y \varepsilon C U_x,$$

und da x ein beliebiger Punkt von C war, so ist $f(x \mid C)$ stetig.

3. Raumerweiterung. Man kann die Baireschen Funktionen einer reellen Veränderlichen, statt durch das System ihrer eindimensionalen Urbildmengen, durch nur zwei ebene Mengen und allgemein die Baireschen Funktionen des Raumes A durch zwei Mengen des erweiterten Raumes $A^* = (A, Y)$ charakterisieren, der das Produkt von A mit der Menge Y der reellen Zahlen ist und die Entfernungen $\sqrt{x\xi^2 + (y - \eta)^2}$ haben mag. Die Borelschen Mengen (6) des *erweiterten* Raumes mögen mit M^ξ, N^ξ bezeichnet werden. Die bisher betrachteten Urbildmengen

$$(9) \qquad M(\beta) = [f > \beta], \quad N(\beta) = [f \geqq \beta]$$

waren die Mengen der Punkte x, die den eingeklammerten Ungleichungen genügen; unter

$$(10) \qquad M = [f > y], \quad N = [f \geqq y]$$

verstehen wir jetzt die (von keinem Parameter mehr abhängigen) Mengen der Punkte (x, y), die den fraglichen Ungleichungen genügen. In demselben Sinn sei

[131] $$(11) \qquad C = N - M = [f = y] .$$

Geometrisch interpretiert, im Falle einer ebenfalls reellen Variablen x, ist C die durch $y = f(x)$ dargestellte „Kurve“, M der Teil der xy-Ebene unterhalb der Kurve, $M(\beta)$ die Projektion auf die x-Achse des Teiles von C, der in der Halbebene $y > \beta$ liegt (oder des Teiles von M, der auf der Geraden $y = \beta$ liegt); entsprechendes ist von N und $N(\beta)$ zu sagen. Nun gilt der Satz:

VII. *$f(x)$ ist dann und nur dann eine Bairesche Funktion $f^\xi(x)$, wenn die Menge M ein M^ξ und die Menge N ein N^ξ ist. Ist $f(x)$ ein $f^\xi(x)$, so ist auch die Menge C ein N^ξ.*

Zum Beweise betrachten wir die Baireschen Funktionen $f^\xi(x)$ des

Raumes A und die Baireschen Funktionen $f^\xi(x, y)$ des erweiterten Raumes A^*. Es gilt:

Jedes $f^\xi(x, 0)$ *ist ein* $f^\xi(x)$. Das ist für $\xi = 0$ richtig, da man aus einer in beiden Variablen x, y stetigen Funktion durch die Substitution $y = 0$ eine in x stetige Funktion erhält, und wird sodann durch Induktion bewiesen. Der Schluß von ξ auf $\xi + 1$ ist trivial; ist ferner η Limeszahl und die Behauptung für alle $\xi < \eta$ bewiesen, so schließt man auf η vermöge der Darstellung von f^η als Quotient zweier k_η (S. 259). Ebenso folgt:

Jedes $f^\xi(x)$ *ist ein* $f^\xi(x, y)$; $f^\xi(x) - y$ *ist ein* $f^\xi(x, y)$, letzteres, weil y stetige Funktion von x, y, also ein $f^\xi(x, y)$ jeder Klasse ist.

Ist danach $f(x)$ ein $f^\xi(x)$, so ist $f(x) - y$ ein $f^\xi(x, y)$, also von der Klasse (M^ξ, N^ξ), die Menge M ist ein M^ξ, N ein N^ξ, C ein N^ξ. Ist umgekehrt M ein M^ξ, N ein N^ξ, so ist $f(x, y) = f(x) - y$ von der Klasse (M^ξ, N^ξ), denn diese Funktion hat die Besonderheit, daß ihre Urbildmengen (in A^*) aus M, N durch Translationen $\eta = y + \beta$ entstehen, die den ganzen Raum A^* isometrisch in sich transformieren. Also ist $f(x, y)$ ein $f^\xi(x, y)$, $f(x, 0) = f(x)$ ein $f^\xi(x)$.

Der letzte Teil des Satzes VII ist im allgemeinen nicht umkehrbar. Jedoch gibt der folgende Satz eine bedingte Umkehrung, die bemerkenswert ist, weil die Mengen $C(\beta) = [f = \beta]$ des Raumes A keinen Schluß auf den Charakter der Funktion $f(x)$ erlauben:

VIII. *Ist der Raum A eine separable, absolut Suslinsche Menge und C eine Suslinsche Menge im erweiterten Raum (A, Y), so ist $f(x)$ eine Bairesche Funktion.* [132]

A sei Suslinsche Menge im separablen vollständigen Raum X, dann ist $A^* = (A, Y)$ Suslinsche Menge im separablen vollständigen Raum (X, Y), also separables S (S = absolut Suslinsche Menge). C und der in A^* offene Halbraum $y > \beta$ sind Suslinsche Mengen in A^*, ihr Durchschnitt also ein höchstens separables S; hiervon ist $M(\beta)$ die Projektion auf den Raum A, also stetiges Bild, nach § 37, II also wieder ein S. Durch Betrachtung des Halbraums $y \leqq \beta$ ergibt sich auch das Komplement $A - M(\beta)$ als S, beide Mengen sind in ihrer Summe A Borelsche Mengen (§ 34, III). Wie $M(\beta)$ ergibt sich auch $N(\beta)$ als Borelsche Menge in A, $f(x)$ ist Bairesche Funktion. In Wahrheit ist dann, wie wir aus VII wissen, C eine Borelsche Menge in A^* gewesen, nicht nur eine Suslinsche. Übrigens kann man aus der Borelschen Klasse von C nicht auf die Bairesche Klasse von $f(x)$ schließen; wenn C ein N^ξ ist, kann $f(x)$ ein $f^\eta(x)$ mit $\eta \geqq \xi$ sein.

Die Projektion von C auf Y ist die von $f(x)$ durchlaufene Wertmenge B, das (reelle, eindeutige) Bild von A vermöge $y = f(x)$; wir bezeichnen diese Menge, falls $f(x)$ eine Bairesche Funktion ist, als *Baireschès Bild* von A.

Hierüber gilt:

[133] IX. *Das Bairesche Bild einer separablen, absolut Suslinschen Menge ist eine Suslinsche Menge. Das schlichte Bairesche Bild einer separablen, absolut Borelschen Menge ist eine Borelsche Menge. Ist B schlichtes Baireschesches Bild einer reellen Suslinschen Menge A, so ist auch A schlichtes Bairesches Bild von B.*

Der Beweis ist größtenteils schon in dem von VIII enthalten. Mit A war auch A^* separables S, ebenso C, ebenso B (§ 37, II). Ist A separable, absolut Borelsche Menge oder Borelsche Menge in X, so (A, Y) in (X, Y), also A^* separable absolut Borelsche Menge, ebenso C (als Borelsche Menge in A^*), ebenso B, falls die Projektion schlicht ist. Bei der dritten Behauptung sei $x = g(y)$ die in B definierte Umkehrfunktion von $y = f(x)$; nun war B ein separables S, C ein S und daher Suslinsche Menge im Raum $B^* = (X, B)$, wo X jetzt die Menge der reellen Zahlen ist; es liegt also, mit Vertauschung der Variablen $x\,y$, wieder die Bedingung des Satzes VIII vor und $g(y)$ ist Bairesche Funktion (auf deren Klasse man aus der von $f(x)$ nicht schließen kann).

Der Satz IX ist eine Verallgemeinerung von § 37, II, allerdings zunächst nur für reelle Funktionen (vgl. XIII). Übrigens geben bereits (§ 37, III) die im Baireschen Nullraum oder in der Menge I der irrationalen Zahlen definierten (reellen) *stetigen* Funktionen als Wertmengen alle (reellen) Suslinschen Mengen. Wenn aber x die Menge A *aller* reellen Zahlen durchlaufen soll, so genügen die stetigen Funktionen $f(x)$ nicht, um als Bilder alle Suslinschen Mengen zu liefern (das stetige Bild von A ist ein F_σ); wohl aber genügen die Funktionen erster Klasse, nämlich:

X. *Jede reelle Suslinsche Menge ist Wertmenge einer Funktion erster Klasse $f(x)$ der reellen Variablen x* (sogar einer Funktion mit höchstens abzählbar vielen Unstetigkeitsstellen).

Wir konnten nämlich die reelle Suslinsche Menge B als stetiges Bild von I, also vermittelst einer stetigen Funktion $f(i)$ darstellen, wo i die irrationalen Zahlen durchläuft. Wir erweitern diese Funktion (ohne Vergrößerung ihrer Wertmenge) auf alle reellen Zahlen, indem wir die rationalen Zahlen in eine Folge $\{r_1, r_2, \ldots\}$ bringen, zu jedem r_n eine irrationale Zahl i_n mit $|i_n - r_n| < \frac{1}{n}$ wählen und $f(r_n) = f(i_n)$ setzen. Dann ist auch noch die Gesamtfunktion $f(x)$ in jedem irrationalen Punkte i stetig, denn für $r_p \to i$ ist auch $i_p \to i$ und $f(r_p) = f(i_p) \to f(i)$. $f(x)$ ist also höchstens in den rationalen Punkten unstetig und nach S. 250 von erster Klasse.

4. Nichtreelle Bairesche Funktionen. A und Y seien metrische Räume; wir betrachten alle in A definierten eindeutigen Funktionen $y = \varphi(x)$ mit $y \varepsilon Y$, wobei also das Bild $B = \Phi(A)$ Teilmenge von Y ist. A heiße der *Raum*, Y der *Bildraum* dieses Funktionensystems. (Wenn Y die Menge der reellen Zahlen ist, handelt es sich um die reellen Funktionen $\varphi(x)$, denen dies Kapitel gewidmet war.) Eine Funktionenfolge $y_n = \varphi_n(x)$, die für alle x konvergiert, erzeugt eine neue Funktion $y = \varphi(x)$. Wir können danach wieder für $\xi < \Omega$ die Baireschen Funktionen φ^ξ definieren, nur mit Ausschluß der Limeszahlen, solange wir die auf reelle Funktionen zugeschnittenen Begriffe eines gewöhnlichen und vollständigen Funktionensystems (§ 41, 1) nicht auf einen beliebigen Bildraum übertragen können. Die φ^0 seien die stetigen Funktionen; sind die φ^ξ bereits definiert, so seien die $\varphi^{\xi+1}$ die Grenzfunktionen konvergenter Folgen von Funktionen φ^ξ; ist η Limeszahl, so definieren wir (keine φ^η, sondern) die $\varphi^{\eta+1}$ als die Grenzfunktionen konvergenter Folgen von Funktionen $\varphi^\xi (\xi < \eta)$. Für reelle Funktionen stimmen die Indizes dieser Klassen mit den früher definierten überein, nur lassen sich dann noch die φ^η einschieben.

Die Lebesgueschen Urbildmengen sind hier, wo keine Unterscheidung von oben und unten in Frage kommt, zweckmäßig so zu erklären. Zunächst werde für $Q \leqq Y$ unter $[y \varepsilon Q]$ die Menge der x verstanden, für die $y = \varphi(x) \varepsilon Q$, also das Urbild von BQ (B Bild von A). Die Mengen

$$[y \varepsilon G], \quad [y \varepsilon F],$$

wo G alle in Y *offenen*, F alle in Y *abgeschlossenen* Mengen durchläuft, sollen dann als Urbildmengen der Funktion $y = \varphi(x)$ bezeichnet werden; sie sind übrigens Komplemente von einander und es würde genügen, die eine Art in Betracht zu ziehen. Wenn M ein gegebenes System von Mengen $\leqq A$ und $N = A - M$ deren Komplemente durchläuft, so heiße die Funktion φ von der Klasse (M, N), falls jede Menge $[y \varepsilon G]$ ein M und jede Menge [134]
$[y \varepsilon F]$ ein N ist. Es gilt dann folgender Satz, bei dem unter verschiedenen möglichen Klassenaussagen diejenige bevorzugt ist, die für Borelsche Mengen die einfachste ist:

XI. *Die Funktionen $\varphi_n \to \varphi$ seien von der Klasse (M, N). Dann ist φ von der Klasse $(N_{\delta\sigma}, M_{\sigma\delta})$, im Falle gleichmäßiger Konvergenz von der Klasse (M_σ, N_δ). Wenn die M ein σ-System, die N ein δ-System bilden, so ist φ von der Klasse (N_σ, M_δ), im Fall gleichmäßiger Konvergenz von der Klasse (M, N).*

Es sei $F > 0$ in Y abgeschlossen; für $\nu = 1, 2, 3, \ldots$ sei F_ν die abgeschlossene Menge der Punkte y mit $\delta(y, F) \leqq \frac{1}{\nu}$, G_ν die offene Menge der Punkte $\delta(y, F) < \frac{1}{\nu}$, H_ν eine beliebige Menge zwischen beiden $(G_\nu \leqq H_\nu \leqq F_\nu)$. Also

$$F = \mathfrak{D} G_\nu = \mathfrak{D} H_\nu = \mathfrak{D} F_\nu .$$

Ferner sei $y_n = \varphi_n(x) \to y = \varphi(x)$. Dann sind folgende Behauptungen über x gleichbedeutend:

(α) $y \varepsilon F$,
(β) für jedes ν ist $y_n \varepsilon H_\nu$ für fast alle n,
(γ) für jedes ν ist $y_n \varepsilon H_\nu$ für unendlich viele n.

Aus (α) folgt (β). Wenn $y \varepsilon F$, so ist, für jedes ν, $y \varepsilon G_\nu$, also (weil G_ν offen ist) $y_n \varepsilon G_\nu \leqq H_\nu$ für fast alle n.

Aus (β) folgt (γ).

Aus (γ) folgt (α). Für jedes ν ist $y_n \varepsilon H_\nu \leqq F_\nu$ für unendlich viele n, also (weil F_ν abgeschlossen ist) $y \varepsilon F_\nu$, folglich $y \varepsilon F$.

Hiernach erhält man mit Benutzung von (α)(γ)

$$[y \varepsilon F] = \mathfrak{D}_\nu \overline{\mathrm{Lim}}_n \, [y_n \varepsilon H_\nu].$$

Wählt man darin $H_\nu = G_\nu$, so ist die Menge $[y_n \varepsilon H_\nu]$ ein M, ihr oberer Limes für $n \to \infty$ ein $M_{\sigma\delta}$ und $[y \varepsilon F]$ ein $M_{\sigma\delta}$. Die Funktion φ ist von der Klasse $(*, M_{\sigma\delta}) = (N_{\delta\sigma}, M_{\sigma\delta})$.

Ist die Konvergenz gleichmäßig, so ist (α) mit der verschärften Behauptung

(β^*) für jedes ν ist $y_n \varepsilon H_\nu$ für $n \geqq n_\nu$

(wo n_ν nur von ν, nicht von x abhängt) gleichbedeutend. Denn wählt man n_ν so, daß $y y_n < \frac{1}{\nu}$ für $n \geqq n_\nu$, so folgt aus (α), für jedes ν und $n \geqq n_\nu$, $\delta(y_n, F) \leqq y y_n < \frac{1}{\nu}$, $y_n \varepsilon G_\nu \leqq H_\nu$, also ($\beta^*$); umgekehrt folgt ($\alpha$) ja schon aus ($\beta$). Danach ist

$$[y \varepsilon F] = \mathfrak{D}_\nu \mathfrak{D}_{n \geqq n_\nu} [y_n \varepsilon H_\nu];$$

wählt man jetzt $H_\nu = F_\nu$, so ergibt sich $[y \varepsilon F]$ als N_δ und φ ist von der Klasse (M_σ, N_δ). Damit ist XI bewiesen.

Definieren wir nun wieder die Borelschen Mengen (6) des Raumes A, so gilt:

[135] XII. *Die Baireschen Funktionen* φ^ξ *sind von der Klasse* (M^ξ, N^ξ).

Das folgt aus XI durch Induktion, weil die M^ξ ein σ-System und ihre Komplemente N^ξ ein δ-System bilden. Der Schluß von ξ auf $\xi + 1$ ist trivial; der von $\xi < \eta$ auf $\eta + 1$ (η Limeszahl) beruht darauf, daß alle Funktionen φ^ξ von der Klasse (M^η, N^η) sind.

Bei reellen Funktionen ließ sich XII umkehren: die Funktionen der Klasse (M^ξ, N^ξ) sind Bairesche Funktionen φ^ξ. Bei beliebigem Bildraum versagen nicht nur die Beweismethoden von § 41, 3, sondern die Behauptung selbst kann falsch sein (abgesehen von $\xi = 0$; die Funktionen der Klasse (G, F) sind nach § 35, I immer stetig, wobei hier G und F die in A offenen und abgeschlossenen Mengen bedeutet). Wenn A zusammenhängend (mehrpunktig) und Y die aus nur zwei Zahlen 0, 1 bestehende Menge ist, so sind die beiden konstanten Funktionen $\varphi = 0$, $\varphi = 1$ die einzigen stetigen und die einzigen Baireschen Funktionen; eine Funktion φ, die an einer einzigen Stelle $= 1$ und sonst überall $= 0$ ist, ist von der Klasse (F_σ, G_δ), ohne ein φ^1 zu sein. Die Frage, in welchen Bildräumen Y oder für welche Paare von Räumen A, Y XII umkehrbar ist, verdient Untersuchung [1]). Jedenfalls gilt die Umkehrung bei beliebigem A, falls Y ein *Euklidischer* Raum ist. Denn hier ist Konvergenz von Punkten mit Konvergenz ihrer sämtlichen Koordinaten gleichbedeutend; $y = (y_1, \ldots, y_m)$ als Funktion $\varphi(x)$ erklären heißt jede Koordinate y_k als reelle Funktion $f_k(x)$ erklären; φ ist dann und nur dann ein φ^ξ, wenn alle f_k reelle f^ξ sind. Ist nun φ von der Klasse (M^ξ, N^ξ), so ist jedes f_k von derselben Klasse, weil der Halbraum $y_k > \beta (y_k \geqq \beta)$ in Y offen (abgeschlossen) ist; also ist jedes f_k ein reelles f^ξ und φ ein φ^ξ.

Die beiden ersten Sätze von IX, die sich damals auf reelle Bairesche Bilder bezogen, gelten für jeden Bildraum Y, gleichviel ob die Klassensätze umkehrbar sind oder nicht. Also:

XIII. *Das Bairesche Bild (reell oder nicht) einer separablen, absolut Suslinschen Menge ist eine höchstens separable, absolut Suslinsche Menge. Das schlichte Bairesche Bild (reell oder nicht) einer separablen, absolut Borelschen Menge ist eine separable, absolut Borelsche Menge.*

Zunächst ist zu sagen: wenn $\varphi_n \to \varphi$ und jedes φ_n von A ein höchstens separables Bild B_n entwirft, so liefert auch φ ein höchstens separables Bild B, weil B offenbar Teilmenge von $(\mathfrak{S} B_n)_\alpha$ ist (B ist übrigens schon im unteren abgeschlossenen Limes $F\underline{l}B_n$ enthalten). Ist also A separabel, so sind ihre stetigen (S. 209 Anm.) und ihre Baireschen Bilder allesamt höchstens separabel. Infolgedessen kann man den Bildraum Y einer Baireschen Funktion $y = \varphi(x)$ durch einen höchstens separablen Y_0 ersetzen, nämlich so, daß nicht nur das Bild B von A in Y_0 liegt, sondern auch φ Bairesche Funktion im Bildraum Y_0 bleibt. Dazu benutze man etwa die S. 257 erwähnte Darstellung

$$\varphi = \lim_p \varphi_p, \quad \varphi_p = \lim_q \varphi_{pq}, \quad \varphi = \lim_r \varphi_{pqr}, \ldots$$ [136]

in der für jede natürliche Zahlenfolge die Funktionen $\varphi_p, \varphi_{pq}, \varphi_{pqr}, \ldots$ schließlich stetig sind (diese Form von φ ist, wie wir sahen, für Bairesche Funktionen notwendig und hinreichend). Sind $B, B_p, B_{pq}, B_{pqr}, \ldots$ die durch diese Funktionen entworfenen Baireschen (schließlich stetigen) Bilder, so kann man offenbar Y durch

$$Y_0 = B \dotplus \mathfrak{S} B_p \dotplus \mathfrak{S} B_{pq} \dotplus \mathfrak{S} B_{pqr} \dotplus \cdots$$

ersetzen. Wir können demgemäß den Bildraum Y als separabel, überdies als vollständig annehmen.

Ferner: sind A und Y beliebig, $\varphi(x)$ Bairesche Funktion, so ist die Menge C der Punkte (x, y) mit $y = \varphi(x)$ im erweiterten Raum $A^* = (A, Y)$ eine Borelsche Menge (vgl. VII, VIII). Bezeichnen wir nämlich die Ent-

[1]) Hierzu vgl. Nachtrag D.

ernungen $y\eta$ im Raume Y wie in linearen Räumen mit $|y - \eta|$, ohne damit Y als linearen Raum vorauszusetzen, und betrachten für irgend eine in A definierte Funktion $\varphi(x)$ des Bildraums Y die Entfernung

$$f(x, y) = |y - \varphi(x)|,$$

das ist eine reelle Funktion im erweiterten Raum A^*. Ist $\varphi(x)$ stetig, so ist $f(x, y)$ stetig; da ferner aus $\varphi_n \to \varphi$

$$|y - \varphi_n(x)| \to |y - \varphi(x)|$$

folgt, so erkennt man sofort, daß, wenn φ eine Bairesche Funktion φ^ξ ist, f eine reelle Bairesche Funktion f^ξ ist. Die Menge $C = [f = 0]$ ist demnach ein N^ξ des erweiterten Raumes A^*.

Der Rest des Beweises verläuft nun wie für IX. A sei Suslinsche Menge im separablen vollständigen Raum X und Y sei separabel und vollständig; (A, Y) ist Suslinsche Menge in (X, Y), also separables S, ebenso C, ebenso B als Projektion von C. Ist A speziell Borelsche Menge, so auch (A, Y), ebenso C und, falls die Projektion schlicht ist, auch B. Hierzu vgl. § 46.

§ 44. Konvergenzmengen.

Im Raum A sei eine Folge stetiger, jetzt wieder reeller Funktionen $f_n = f_n(x)$ definiert; wir fragen nach der *Konvergenzmenge* C dieser Folge, d. h. nach der Menge der Punkte x (Konvergenzpunkte, Konvergenzstellen), wo $\lim f_n(x)$ existiert. Hierzu ist notwendig und hinreichend: es gibt für jedes $\sigma > 0$ ein Glied $f_m(x)$ der Folge, das sich von allen nachfolgenden um höchstens σ unterscheidet, wofür also

$$(1) \qquad |f_n(x) - f_m(x)| \leqq \sigma \text{ für } n > m.$$

Daraus folgt: es sei (wie S. 252) $F_m(\sigma)$ die Menge der Punkte x, in denen (1) bei festem σ, m gilt,

$$C(\sigma) = F_1(\sigma) \dotplus F_2(\sigma) \dotplus F_3(\sigma) \dotplus \cdots$$

die Menge der x, wo (1) bei festem σ für irgend ein m gilt; dann ist C der Durchschnitt aller $C(\sigma)$ oder, da die Mengen $F_m(\sigma)$ und $C(\sigma)$ mit σ gleichzeitig abnehmen,

$$C = C(1)\, C(\tfrac{1}{2})\, C(\tfrac{1}{3}) \cdots$$

$F_m(\sigma)$ ist abgeschlossen, $C(\sigma)$ ein F_σ, *die Konvergenzmenge C ist ein $F_{\sigma\delta}$.*

Man kann dies auch so erkennen. Setzt man voraus, daß

$$\overline{f} = \overline{\lim} f_n, \quad \underline{f} = \underline{\lim}\, f_n$$

überall existieren (d. h. endlich sind), so sind dies Funktionen zweiter Klasse, und zwar nach § 41, (6) $\overline{f}$ ein h^1 von der Klasse $(*, G_\delta)$ und $\underline{f}$ ein g^1 von der Klasse $(F_\sigma, *)$. Ihre Differenz

$$\omega(f) = \overline{f} - \underline{f},$$

die man als *Oscillation* von f_n für $n \to \infty$ bezeichnet, ist als Funktion zweiter Klasse[1]) von der Klasse $(G_{\delta\sigma}, F_{\sigma\delta})$; die Menge der Divergenzpunkte, $\omega(f) > 0$, also ein $G_{\delta\sigma}$; die Menge der Konvergenzpunkte, $\omega(f) = 0$, ein $F_{\sigma\delta}$. Die Menge der Stellen, wo $\lambda < \lim f_n < \mu$, ist auch ein $F_{\sigma\delta}$, nämlich der Durchschnitt $[\overline{f} = \underline{f}][\underline{f} > \lambda][\overline{f} < \mu] = F_{\sigma\delta} F_\sigma F_\sigma$. Danach kann man sich von der Voraussetzung endlicher $\overline{f}, \underline{f}$ durch die Beschränkungstransformation $\varphi_n = \frac{f_n}{1+|f_n|}$ befreien; die Menge der Stellen, wo f_n konvergiert, ist mit der Menge der Stellen identisch, wo $-1 < \lim \varphi_n < 1$ (während $\varphi_n \to 1$ mit $f_n \to \infty$, $\varphi_n \to -1$ mit $f_n \to -\infty$ gleichbedeutend ist).

Die Menge $[\overline{f} = \infty] = [\overline{\varphi} = 1] = [\overline{\varphi} \geqq 1]$ ist übrigens ein G_δ, ebenso die Menge $[\underline{f} = -\infty]$, und auch die Summe beider, d. h. *die Menge* D_∞ *der Stellen, wo* $f_n(x)$ *nicht beschränkt ist, ist ein* G_δ. Wenn man z. B. (wie in der Theorie der Fourierreihen) eine Folge stetiger Funktionen einer reellen Variablen konstruiert, die an allen rationalen Stellen in dieser Weise (so, daß $f_n(x)$ nicht beschränkt ist) divergieren, so findet dasselbe Verhalten automatisch auch noch an $\aleph$ irrationalen Stellen statt, denn D_∞ hat als dichtes G_δ nach § 26, XI die Mächtigkeit des Kontinuums.

Die Menge der Stellen, wo f_n nach einem vorgeschriebenen Werte konvergiert, ist auch ein $F_{\sigma\delta}$, z. B.

$$[f_n \to 0] = [\varphi_n \to 0] = [\underline{\varphi} \geqq 0]\,[\overline{\varphi} \leqq 0].$$

Aber auch die Menge, wo f_n nach $+\infty$ oder $-\infty$ divergiert[2]), ist ein $F_{\sigma\delta}$, da

$$[f_n \to \infty] = [\underline{f} = \infty] = [\underline{\varphi} \geqq 1].$$

Wir haben bisher die erste Hälfte des Satzes bewiesen:

I (H. Hahn). *Die Konvergenzmenge einer Folge reeller stetiger Funktionen ist stets ein* $F_{\sigma\delta}$, *und zu jeder Menge* $C = F_{\sigma\delta}$ *läßt sich eine Folge reeller stetiger Funktionen angeben, für welche* C *die Konvergenzmenge ist.*

Zum Beweis der zweiten Hälfte nehmen wir C zunächst als F_σ

$$C = F_1 \dotplus F_2 \dotplus F_3 \dotplus \cdots, \quad F_1 \leqq F_2 \leqq F_3 \leqq \cdots$$

an (F_n abgeschlossen); $D = A - C$ sei das Komplement. Die Funktion h, die in

$$F_1, \quad F_2 - F_1, \quad F_3 - F_2, \ldots, \quad D$$

gleich

$$1, \quad \tfrac{1}{2}, \quad \tfrac{1}{3}, \quad \ldots, \quad 0$$

definiert wird; ist oberhalb stetig; denn $[h \geqq y]$ ist, wenn nicht die Nullmenge oder der ganze Raum, eine der Mengen F_n. Sie ist also Limes einer absteigenden Folge stetiger Funktionen $\psi_1 \geqq \psi_2 \geqq \cdots$; da die Funktionen $\psi_n' = \min[\psi_n, 1]$ und $\psi_n'' = \max\left[\psi_n', \frac{1}{n}\right]$ ebenfalls absteigend

[1]) Sie ist selbst ein h^1 von der Klasse $(G_{\delta\sigma}, G_\delta)$, woraus aber nicht mehr folgt als im Text.

[2]) Dies wird als *eigentliche Divergenz* bezeichnet, sollte aber besser *uneigentliche Konvergenz* heißen (H. Hahn).

nach h konvergieren, kann man alsbald $\frac{1}{n} \leqq \psi_n \leqq 1$ annehmen. Die stetigen Funktionen $\varphi_n = \frac{1}{\psi_n}$, für die $1 \leqq \varphi_n \leqq n$ gilt, bilden eine aufsteigende Folge und konvergieren in C nach $\frac{1}{h}$, d. h. nach ganzzahligen Grenzwerten, in D divergieren sie nach $+\infty$. Das wäre also schon eine Lösung des Problems für den Fall $C = F_\sigma$; zur Weiterführung transformieren wir sie folgendermaßen. Es ist $0 \leqq \varphi_{n+1} - \varphi_n \leqq n$; schieben wir zwischen φ_n und φ_{n+1} noch Funktionen mit gebrochenem Index

$$\varphi_{n+t} = \varphi_n + t(\varphi_{n+1} - \varphi_n) \qquad \left(t = \frac{1}{2n}, \frac{2}{2n}, \ldots, \frac{2n-1}{2n}\right)$$

ein und ändern die Bezeichnung, indem wir ursprüngliche und eingeschaltete Funktionen durchlaufend mit natürlichen Zahlen numerieren, so haben wir jetzt eine Folge stetiger Funktionen $1 = \varphi_1 \leqq \varphi_2 \leqq \cdots$, $\varphi_{n+1} - \varphi_n \leqq \frac{1}{2}$; in C ist $\lim \varphi_n$ eine natürliche Zahl, in D ist $\lim \varphi_n = +\infty$. Die Funktionen

$$f_n = a \sin \pi \varphi_n$$

(a eine positive Konstante) konvergieren in C nach 0 und divergieren in D; denn $(x \varepsilon D)$ in jedes der Intervalle $[k + \frac{1}{4}, k + \frac{3}{4}]$ von der Länge $\frac{1}{2}$ $(k = 1, 2, 3, \ldots)$ fällt mindestens ein φ_n und dann ist $(-1)^k \sin \pi \varphi_n \geqq \sin \frac{\pi}{4}$, also unendlich oft $f_n \geqq \frac{a}{\sqrt{2}}$ und unendlich oft $f_n \leqq -\frac{a}{\sqrt{2}}$.

Es gibt also zu jedem $C = F_\sigma$ eine Folge stetiger, gleichmäßig beschränkter Funktionen f_n (vom Betrage $\leqq a$), *die in C nach 0 konvergieren, in $D = A - C$ divergieren* (W. Sierpiński).

Nun sei $C = C_1 C_2 \ldots$ Durchschnitt einer Folge von Mengen $C_m = F_\sigma$; D_m sei das Komplement von C_m, $D = D_1 \dotplus D_2 \dotplus \cdots$ das von C. Wir bestimmen zu jedem $m = 1, 2, \ldots$ eine Folge stetiger Funktionen f_{mn} vom Betrage $\leqq \frac{1}{m}$ derart, daß der Grenzwert $\lim\limits_n f_{mn}$ in C_m vorhanden und $= 0$, in D_m nicht vorhanden ist. Die Doppelfolge dieser Funktionen f_{mn} werde als einfache Folge f_p geschrieben (etwa $f_1, f_2, f_3, \ldots = f_{11}, f_{12}, f_{21}, \ldots$). Dann ist der Grenzwert $\lim f_p$ in C vorhanden und $= 0$, in D nicht vorhanden. Denn für $x \varepsilon C$ kann die Ungleichung $|f_{mn}| \geqq \varepsilon > 0$ wegen $|f_{mn}| \leqq \frac{1}{m}$ nur für endlich viele m und dann, wegen $\lim\limits_n f_{mn} = 0$, jedesmal nur für endlich viele n gelten; für $x \varepsilon D$ und etwa $x \varepsilon D_m$ bilden schon die Funktionen $f_{m1}, f_{m2}, \ldots$ eine divergente Teilfolge der f_p. Damit ist I bewiesen.

Wenn die Funktionen f_n, statt stetig zu sein, dem System aller Funktionen der Klasse (M, N) entnommen sind, wo die M einen σ-Ring und ihre Komplemente N einen δ-Ring bilden, so ist ihre Konvergenzmenge ein $N_{\sigma\delta}$; der Beweis ist wie zu Beginn dieses Paragraphen. Das überträgt sich auch auf nichtreelle Funktionen (§ 43, 4), falls der Bildraum vollständig ist.

Kehren wir wieder zu reellen stetigen Funktionen zurück und betrachten statt einer *Folge* von Funktionen $f_n(x)$ eine *Schar* von Funktionen $f(x, y)$, die von einem reellen, positiven Parameter y abhängen (bei festem y also stetige Funktionen von $x \varepsilon A$ sind) und deren Verhalten für $y \to 0$ untersucht werden soll. *Die Konvergenzmenge C der Punkte x,* wo $\lim\limits_{y \to 0} f(x, y)$ *existiert, ist wieder ein $F_{\sigma\delta}$.* Denn hierzu ist notwendig und hinreichend: es gibt für jedes $\sigma > 0$ ein $\eta > 0$ derart, daß

$$|f(x, y_1) - f(x, y_2)| \leqq \sigma \quad \text{für } y_1 < y_2 < \eta. \tag{2}$$

Bei festen σ, η ist die Menge $F(\sigma, \eta)$ der Punkte x, in denen (2) gilt, abgeschlossen (als Durchschnitt der Mengen, wo die Ungleichung (2) für ein bestimmtes Paar y_1, y_2 gilt). Die Menge der Punkte x, wo (2) für festes σ und irgendein η gilt, ist $C(\sigma) = \mathfrak{S}_\eta F(\sigma, \eta)$, wofür wir

$$C(\sigma) = F(\sigma, 1) \dotplus F(\sigma, \tfrac{1}{2}) \dotplus F(\sigma, \tfrac{1}{3}) \dotplus \cdots$$

schreiben können, da die $F(\sigma, \eta)$ mit abnehmendem η wachsen; endlich ist wieder C der Durchschnitt der $C(\sigma)$ oder

$$C = C(1)\, C(\tfrac{1}{2})\, C(\tfrac{1}{3}) \ldots,$$

womit C als $F_{\sigma\delta}$ erkannt ist.

Der obere Limes

$$\bar{f}(x) = \overline{\lim_{y \to 0}} f(x, y)$$

wird durch

$$\bar{f}(x) = \lim_{\eta \to 0} g(x, \eta), \quad g(x, \eta) = \sup_{y \leqq \eta} f(x, y) \quad (\eta > 0)$$

definiert; soll er (bei festem x) existieren, d. h. endlich sein, so muß, für hinlänglich kleines y, $f(x, y)$ nach oben und $g(x, y)$ nach unten beschränkt sein; $g(x, y)$ nimmt zugleich mit y ab. Existiert $\bar{f}(x)$ für jedes x, so ist $\bar{f}(x)$ wieder eine Funktion h^1 der Klasse $(*, G_\delta)$. Um das zu erkennen, können wir $f(x, y)$, bei festem x, für *alle* $y > 0$ nach oben beschränkt annehmen, indem wir statt seiner $\min\left[f(x, y), \frac{1}{y}\right]$ betrachten, welche Funktion für hinlänglich kleines y, etwa $y \leqq \eta$, mit $f(x, y)$ übereinstimmt und sonst, für $y \geqq \eta$, jedenfalls $\leqq \frac{1}{\eta}$ ist. Bei dieser Annahme existiert $g(x, y)$ für jedes y und ist, als obere Grenze stetiger Funktionen, unterhalb stetig (S. 249); ferner ist dann

$$\bar{f}(x) = \lim g\left(x, \frac{1}{n}\right)$$

als Limes absteigender Funktionen erster Klasse ein h^1. Der analog zu definierende

$$\underline{f}(x) = \underline{\lim}_{y \to 0} f(x, y)$$

ist, wenn überall endlich, ein g^1 von der Klasse $(F_\sigma, {}^*)$.

Z. B. für

$$f(x, y) = \frac{\varphi(x+y) - \varphi(x)}{y},$$

wo $\varphi(x)$ stetige Funktion der reellen Variablen x ist, ergibt sich: die Menge der Stellen, wo $\varphi(x)$ nach rechts differenzierbar ist, ist ein $F_{\sigma\delta}$; die rechte obere Derivierte (überall endlich angenommen) ist ein h^1, die rechte untere ein g^1. Dasselbe gilt natürlich nach links, übrigens auch beiderseits (wenn man rechts und links nicht unterscheidet).

Bei einer Funktionenschar sind diese Schlüsse (auf die Klasse von $\bar{f}$ und $\underline{f}$ sowie der Konvergenzmenge) von stetigen Funktionen *nicht* auf Funktionen der Klasse (M, N) übertragbar; sie beruhen ja wesentlich darauf, daß der Durchschnitt *beliebig* (nicht bloß abzählbar) vieler F ein F, die Summe beliebig vieler G ein G ist. So hat die Tatsache, daß

[137] $$g(x) = \sup_y f(x, y)$$

($f(x, y)$ sei bei festem x nach oben beschränkt) unterhalb stetig ist, falls $f(x, y)$ bei festem y nach x stetig ist, kein Analogon für Bairesche Funktionen f höherer Klasse. Wenn $f(x, y)$ Bairesche Funktion beider Variablen im erweiterten Raum (A, Y) und A separables S (S = absolut Suslinsche Menge) ist, so ergibt sich $g(x)$ als Funktion der Klasse (S, S), denn offenbar ist die Menge $[g > c]$ Projektion der Menge $[f > c]$ auf den Raum A, also ein S, ebenso $[g \geqq c] = \mathfrak{D}_n \left[g > c - \frac{1}{n}\right]$. Die Mengen $[g < c]$, $[g \leqq c]$ sind Komplemente $A - S$. (Wenn speziell der Raum A Euklidisch und $f(x, y)$ von erster Klasse ist, so ist $g(x)$ von der Klasse $(F_\sigma, {}^*)$, also von zweiter Klasse. Denn hier ist die Projektion eines F_σ wieder ein F_σ.)

Es ist lehrreich, noch einige andere Konvergenzmengen zu betrachten. Wir nehmen x reell, y positiv an; die in der oberen Halbebene $y > 0$ definierte reelle Funktion $f(x, y)$ sei bei festem y stetig in x. Wir lassen (x, y) aus der oberen Halbebene nach einem Punkt $(\xi, 0)$ der x-Achse konvergieren; je nach der Art, wie diese Konvergenz

$$(3) \qquad (x, y) \to (\xi, 0)$$

vorgeschrieben ist, kann die Menge C der Zahlen ξ oder der Punkte $(\xi, 0)$, für die $\lim f(x, y)$ existiert, verschiedenen Charakter haben.

Geradlinige Annäherung. Bei Annäherung parallel der y-Achse, wo es sich also um $\lim_{y \to 0} f(\xi, y)$ handelt, wissen wir, daß C ein $F_{\sigma\delta}$ ist. Dasselbe gilt bei Annäherung auf der festen Geraden

$$(4) \qquad x - \xi = ty,$$

die mit der y-Achse den Winkel α bildet $\left(-\frac{\pi}{2} < \alpha < \frac{\pi}{2},\quad t = \operatorname{tg}\alpha\right)$; hier handelt es sich um den Grenzwert

$$(5)\qquad \lim_{y\to 0} f(\xi + ty, y) = g(\xi, t)$$

einer, bei festem y und t, stetigen Funktion von ξ, und die Menge $C(t)$ der ξ, wo dieser Grenzwert existiert, ist ein $F_{\sigma\delta}$. Dies ist also das Ergebnis bei *Annäherung in fester Richtung.*

Die Summe $\overline{C} = \mathfrak{S}C(t)$ ist die Menge der Punkte, wo bei *Annäherung in mindestens einer Richtung* Konvergenz stattfindet. $\overline{C}$ ist eine *Suslinsche Menge.* Denn $f(\xi + ty, y)$ ist bei festem y stetige Funktion der Variablen ξ, t; die Menge Γ der Punkte (ξ, t), wo der Grenzwert (5) existiert, ist ein $F_{\sigma\delta}$ in der ξt-Ebene, und $\overline{C}$ die Projektion von Γ auf die ξ-Achse.

Der Durchschnitt $\underline{C} = \mathfrak{D}C(t)$ ist die Menge der Punkte, wo bei *Annäherung in jeder Richtung* Konvergenz stattfindet (wobei der Grenzwert (5) aber von der Richtung t abhängen kann). $\underline{C}$ ist *das Komplement einer Suslinschen Menge*, nämlich der Projektion von $E - \Gamma$ auf die ξ-Achse, wo Γ die eben erwähnte Menge und E die ξt-Ebene ist.

Bei *Annäherung in einem festen Winkelraum*

$$|x - \xi| \leqq ty \qquad (t > 0)$$

ist die Konvergenzmenge ein $F_{\sigma\delta}$, bei *Annäherung in mindestens einem* (hinlänglich schmalen) *Winkelraum* ein $F_{\sigma\delta\sigma}$, bei *Annäherung in jedem* (beliebig breiten) *Winkelraum* ein $F_{\sigma\delta}$; diese Behauptungen möge der Leser selbst beweisen.

Bei ganz *willkürlicher Annäherung* (3) ist die Konvergenzmenge C ein G_δ, und zwar ohne jede Voraussetzung über $f(x, y)$, das also nicht mehr nach x stetig zu sein braucht. Denn zur Konvergenz ist notwendig und hinreichend, daß für jedes $\sigma > 0$ der Punkt $(\xi, 0)$ eine ebene Umgebung U habe, in deren oberer Hälfte $(y > 0)$ für je zwei Punkte

$$|f(x_1, y_1) - f(x_2, y_2)| \leqq \sigma$$

ist. Die Menge $G(\sigma)$ der Punkte $(\xi, 0)$, die bei festem σ eine solche Umgebung haben, ist in der ξ-Achse offen, denn jeder in U liegende Punkt $(\eta, 0)$ hat ebenfalls eine solche Umgebung $(\leqq U)$; und $C = \mathfrak{D}G\left(\frac{1}{n}\right)$ ist ein G_δ. — Diese Betrachtung ist im Grunde dieselbe, aus der sich die Menge der Stetigkeitspunkte einer beliebigen Funktion als G_δ ergab; man vergleiche auch die Anmerkung zu S. 217.

Zehntes Kapitel.

Ergänzungen.

§ 45. Die Bairesche Bedingung.

1. Moduln und Kongruenzen. Alle Mengen, die wir betrachten, seien Teilmengen einer festen Menge E, die wir als „Raum" und deren
[138] Elemente wir als „Punkte" bezeichnen. Als *Diskrepanz* zweier Mengen A, B erklären wir die Menge

$$\begin{aligned}[A, B] &= (A \dot{+} B) - AB \\ &= (A - AB) + (B - AB) = [B, A]\end{aligned}$$

der Punkte, die einer dieser Mengen, aber nicht der anderen angehören. Sind $a(x)$, $b(x)$ die charakteristischen Funktionen (S. 20) von A, B, so ist $|a(x) - b(x)|$ die charakteristische Funktion von $[A, B]$.

Für drei Mengen A, B, C gilt

$$(1) \qquad [A, C] \leqq [A, B] \dot{+} [B, C].$$

Denn sei $x \varepsilon [A, C]$, etwa $x \varepsilon A$, $x \bar{\varepsilon} C$; je nachdem $x \bar{\varepsilon} B$ oder $x \varepsilon B$, gehört x zu $[A, B]$ oder zu $[B, C]$.

Ferner ist offenbar

$$(2) \qquad [E - A, E - B] = [A, B].$$

Ist jedem Element (Index) n einer beliebigen Menge[1]) N ein Paar von Mengen A_n, B_n zugeordnet, so ist

$$(3) \qquad \begin{cases} [\mathfrak{S} A_n, \mathfrak{S} B_n] \leqq \mathfrak{S}[A_n, B_n] \\ [\mathfrak{D} A_n, \mathfrak{D} B_n] \leqq \mathfrak{S}[A_n, B_n] \end{cases}$$

Um die erste Formel zu beweisen, sei x Punkt der links stehenden Menge und etwa $x \varepsilon \mathfrak{S} A_n$, $x \bar{\varepsilon} \mathfrak{S} B_n$; es gibt also ein n mit $x \varepsilon A_n$, zugleich ist $x \bar{\varepsilon} B_n$, $x \varepsilon [A_n, B_n]$. Die zweite Formel (3) beweist man ebenso oder mit Hilfe von (2).

Bei dem Suslinschen Konstruktionsverfahren (S. 91) sei

$$\begin{aligned} A &= \mathfrak{S}_{\nu} A_{\nu}, \quad A_{\nu} = A_{n_1} A_{n_1 n_2} \cdots \\ B &= \mathfrak{S}_{\nu} B_{\nu}, \quad B_{\nu} = B_{n_1} B_{n_1 n_2} \cdots \end{aligned}$$

Dann ist

$$(4) \qquad [A, B] \leqq \mathfrak{S}[A_{n_1}, B_{n_1}] \dot{+} \mathfrak{S}[A_{n_1 n_2}, B_{n_1 n_2}] \dot{+} \cdots$$

Das folgt aus (3), weil

$$\begin{gathered} [A, B] \leqq \mathfrak{S}_{\nu}[A_{\nu}, B_{\nu}], \\ [A_{\nu}, B_{\nu}] \leqq [A_{n_1}, B_{n_1}] \dot{+} [A_{n_1 n_2}, B_{n_1 n_2}] \dot{+} \cdots \end{gathered}$$

---

[1]) Diese braucht natürlich nicht Teilmenge von E zu sein.

Ein nichtleeres System $\mathfrak{M}$ von Teilmengen M des Raumes E heißt unter folgenden Bedingungen ein *Modul*: [139]

(α) Jede Teilmenge einer Menge aus $\mathfrak{M}$ gehört wieder zu $\mathfrak{M}$,

(β) Die Summe endlich vieler Mengen aus $\mathfrak{M}$ gehört wieder zu $\mathfrak{M}$.

Wird statt (β) die schärfere Forderung gestellt:

(β_σ) Die Summe (endlich oder) abzählbar vieler Mengen aus $\mathfrak{M}$ gehört wieder zu $\mathfrak{M}$,

so heißt ein System $\mathfrak{M}$, das (α) und (β_σ) erfüllt, ein *σ-Modul.* Jeder Modul enthält die leere Menge 0.

Die endlichen oder (falls E ein metrischer Raum ist) die in E nirgends dichten Mengen bilden einen Modul, die höchstens abzählbaren oder die Mengen erster Kategorie in E (die E_{I}) einen σ-Modul.

Nach dem Vorbild der Zahlentheorie und Algebra nennen wir zwei Mengen A, B *kongruent nach dem Modul* $\mathfrak{M}$,

$$A \equiv B \ (\mathfrak{M}),$$

wenn sie sich „nur um Mengen $M \varepsilon \mathfrak{M}$ unterscheiden", genauer gesprochen: *wenn ihre Diskrepanz* $[A, B]$ *zu* $\mathfrak{M}$ *gehört.* Man kann das so ausdrücken: die beiden Summanden

$$\begin{aligned} M_1 &= A - AB = (A \dot{+} B) - B, \\ M_2 &= B - AB = (A \dot{+} B) - A, \end{aligned}$$

aus denen sich die Diskrepanz zusammensetzt, sind Mengen M des Moduls $\mathfrak{M}$, und gemäß

$$\begin{aligned} B &= AB + M_2 = (A - M_1) + M_2 \\ &= (A \dot{+} B) - M_1 = (A + M_2) - M_1 \end{aligned}$$

entsteht von den kongruenten Mengen die eine aus der andern, indem man eine Menge M abzieht und dafür eine Menge M hinzufügt. Man kann auch sagen, daß A und B „bis auf Mengen M" übereinstimmen oder „mit Vernachlässigung von Mengen M" identisch sind. $A \equiv 0$ $(\mathfrak{M})$ heißt, daß A zu $\mathfrak{M}$ gehört. Wir lassen im folgenden bei den Kongruenzen die Angabe des festen Moduls fort und schreiben statt $A \equiv B$ $(\mathfrak{M})$ einfach $A \equiv B$.

Daß $A \equiv A$ und mit $A \equiv B$ auch $B \equiv A$, d. h. daß die Kongruenzbeziehung reflexiv und symmetrisch ist, ist trivial; nach (1) ist sie auch transitiv, d. h. mit $A \equiv B$, $B \equiv C$ ist $A \equiv C$. Sie gestattet also die Einteilung der Teilmengen von E in Klassen kongruenter Mengen, wobei zwei Klassen entweder identisch oder disjunkt sind. Ist $E \equiv 0$, so sind alle Mengen $\equiv 0$ und es gibt nur eine Klasse; das andere Extrem wäre, daß $\mathfrak{M}$ nur aus der leeren Menge 0 besteht und jede einzelne Menge für sich eine Klasse bildet.

Aus (2) folgt: mit $A \equiv B$ ist $E - A \equiv E - B$.

Aus (3) folgt: mit $A_n \equiv B_n$ ist

$$\mathfrak{S}A_n \equiv \mathfrak{S}B_n, \quad \mathfrak{D}A_n \equiv \mathfrak{D}B_n$$

für eine endliche, im Fall eines σ-Moduls auch für eine abzählbare Menge von Indizes n.

Aus (4) ergibt sich im Fall eines σ-Moduls:

$$\text{mit } A_{n_1 \ldots n_k} \equiv B_{n_1 \ldots n_k} \text{ ist } A \equiv B,$$

der Suslinsche Prozeß läßt die Kongruenz bestehen.

E sei jetzt ein metrischer Raum; als Umgebung U_x von x bezeichnen wir jede offene, den Punkt x enthaltende Menge. $\mathfrak{M}$ sei wieder ein fester Modul und A eine beliebige Menge des Raumes. Für das lokale Verhalten von A zu $\mathfrak{M}$ in einem Raumpunkt x machen wir folgende Unterscheidung:

x heiße ein *Nullpunkt* von A (für $\mathfrak{M}$), wenn er eine Umgebung U_x mit

$$(5) \qquad A U_x \equiv 0$$

hat; andernfalls, wenn also für jede Umgebung U_x

$$A U_x \not\equiv 0$$

ist, ein *positiver* Punkt. Die Menge der positiven Punkte werde A_p, die der Nullpunkte A_q genannt, also $E = A_p + A_q$. *A_q ist offen, A_p abgeschlossen.* Denn wenn $x \varepsilon A_q$ und also eine Umgebung U_x mit (5) hat, so ist $U_x \leqq A_q$, weil jeder Punkt von U_x ebenfalls eine solche Umgebung (nämlich U_x) hat. Es ist also

$$A_q = \mathfrak{S}_x^{A_q} U_x,$$

wenn jedem Nullpunkt x ein U_x gemäß (5) zugeordnet wird, und

$$A A_q = \mathfrak{S}_x^{A_q} A U_x,$$

wobei hier jeder Summand $\equiv 0$ ist. Wenn $\mathfrak{M}$ ein σ-Modul und E (oder wenigstens A) separabel ist, so folgt daraus

$$(6) \qquad A A_q \equiv 0,$$

da man gemäß § 25, VI die Summe über die U_x (oder $A U_x$) auf abzählbar viele dieser Summanden beschränken kann. Die Gleichung (6) kann aber auch gelten, wenn E nicht separabel ist (vgl. den folgenden Satz I).

Ein Punkt $x \varepsilon E - A_\alpha$ ist Nullpunkt von A, da er eine Umgebung $U_x = E - A_\alpha$ mit $A U_x = 0$ hat; hieraus folgt

$$(7) \qquad A_p \leqq A_\alpha$$

Wenn alle endlichen Teilmengen von A zu $\mathfrak{M}$ gehören, ist $A_p \leqq A_\beta$; wenn alle abzählbaren, so ist $A_p \leqq A_\gamma$ (vgl. § 23).

Wenn $A \equiv B$, so ist auch $A U_x \equiv B U_x$ und jeder Nullpunkt der einen Menge auch Nullpunkt der andern, also $A_p = B_p$, $A_q = B_q$ (dies läßt sich nicht umkehren).

2. Die Bairesche Bedingung für Mengen. $\mathfrak{M}$ sei jetzt der σ-Modul der Mengen erster Kategorie (in E). Die Nullpunkte und positiven Punkte von A werden hier *Punkte erster und zweiter Kategorie* von A genannt. Wir zeigen zunächst, daß die Formel (6) in jedem Falle, auch bei nichtseparablem Raum E gilt:

I (Satz von S. Banach). *Die in A liegenden Punkte erster Kategorie von A bilden eine Menge erster Kategorie.*

Wir betrachten zuerst den Spezialfall $A \leqq A_q$, wollen also beweisen: *wenn alle Punkte von A Punkte erster Kategorie sind, so ist A von erster Kategorie.* Wir bilden ein maximales, d. h. nicht erweiterungsfähiges System von *disjunkten* offenen Mengen U_x mit $AU_x \equiv 0$; die Existenz eines solchen Systems folgt wie in § 31 durch Wohlordnung und aus Mächtigkeitsgründen. x durchläuft eine geeignete Teilmenge von A_q; $U = \sum\limits_x U_x$ ist offen, $V = E - U$ abgeschlossen. Dann ist V nirgendsdicht, der offene Kern $V_i = 0$; denn wenn $V_i > 0$ wäre, so könnte, je nachdem $AV_i = 0$ oder a ein Punkt von AV_i ist, entweder V_i selbst oder (da A in a von erster Kategorie ist) ein geeignetes $U_a \leqq V_i$ mit $AU_a \equiv 0$ den Mengen U_x hinzugefügt werden. Wir haben jetzt $V \equiv 0$,

$$A = AU + AV \equiv AU$$

und es bleibt $AU \equiv 0$ zu zeigen. Als Menge erster Kategorie ist

$$AU_x = A_{x1} + A_{x2} + \cdots = \sum_n A_{xn}$$

Summe abzählbar vieler nirgendsdichter Mengen A_{xn}, die (bei festem x) disjunkt angenommen werden können. Dann sind die sämtlichen Mengen A_{xn} paarweise disjunkt, und wenn wir zeigen können, daß jedes

$$A_n = \sum_x A_{xn}$$

nirgendsdicht ist, so folgt aus

$$AU = \sum_{xn} A_{xn} = \sum_n A_n$$

die Behauptung $AU \equiv 0$. Nun ist offenbar

$$A_{xn} = A_n U_x;$$

hieraus folgt für die abgeschlossenen Hüllen (S. 118, (13))

$$A_{xn\alpha} \geqq A_{n\alpha} U_x$$

und wenn G den offenen Kern $A_{n\alpha i}$ von $A_{n\alpha}$ bedeutet: $A_{xn\alpha} \geqq GU_x$, also $GU_x = 0$, da A_{xn} nirgendsdicht ist. Demnach ist $GU = 0, G = GV \leqq V_i = 0$, $G = 0$, A_n nirgendsdicht.

Hiermit ist der besondere Fall $A \leqq A_q$ des Satzes I bewiesen; im allgemeinen Fall beachten wir, daß für jede Teilmenge B von A ersichtlich $A_q \leqq B_q$ ist, insbesondere für $B = AA_q$: $B \leqq B_q$, also $B \equiv 0$.

Die Begrenzung jeder offenen Menge G oder jeder abgeschlossenen

Menge F ist nirgendsdicht, also $G_\alpha - G \equiv F - F_i \equiv 0$; jede offene Menge ist mit einer abgeschlossenen und umgekehrt kongruent ($G \equiv G_\alpha$, $F \equiv F_i$).

[140] Wir sagen, eine Menge A sei eine *β-Menge* oder genüge der *Baireschen Bedingung* (im weiteren Sinn, C. Kuratowski), wenn sie mit einer offenen oder auch mit einer abgeschlossenen Menge kongruent ist.

Das Komplement einer β-Menge ist eine β-Menge. Denn $A \equiv G$ gibt $E - A \equiv E - G = F$.

Die β-Mengen bilden ein Borelsches System [1]). Denn ist $A_n \equiv G_n \equiv F_n$ für $n = 1, 2, 3, \ldots$, so folgt

$$\mathfrak{S}A_n \equiv \mathfrak{S}G_n = G, \quad \mathfrak{D}A_n \equiv \mathfrak{D}F_n = F.$$

Alle Borelschen Mengen des Raumes sind also β-Mengen.

Jede β-Menge entsteht aus einer offenen, indem man ein E_{I} wegläßt und hinzufügt. Will man nur die eine dieser beiden Operationen zulassen, so ergibt sich: die β-Mengen sind die Mengen der Form

$$(8) \qquad G_\delta + E_{\mathrm{I}} \text{ oder } F_\sigma - E_{\mathrm{I}}.$$

Denn sei $A \equiv G$ eine β Menge; die Diskrepanz $[A, G]$ ist als Menge erster Kategorie in einer Menge $D = F_\sigma$ von erster Kategorie enthalten; setzt man $C = E - D = G_\delta$, so ist $C\,[A, G] = [CA, CG] = 0$, $CA = CG$ und

$$A = CA + DA = CG + DA,$$

wobei CG ein G_δ und DA ein E_{I} ist. Durch Komplementbildung erhält man die zweite Form (8) [2]).

A ist dann und nur dann β-Menge, wenn es eine Menge $C \equiv E$ ($E - C$ ein E_{I}) gibt derart, daß CA in C offen, oder abgeschlossen, oder zugleich offen und abgeschlossen ist.

Daß dies hinreicht, ist evident; ist z. B. $CA = CG$ und $C \equiv E$, so ist $A = EA \equiv CA = CG \equiv EG = G$. Die Notwendigkeit folgt so: ist $A \equiv H$, $C = E - [A, H] \equiv E$, so ist wie oben $C[A, H] = [CA, CH] = 0$, $CA = CH$. Setzt man für H eine offene oder abgeschlossene Menge, so gibt es also ein $C_1 \equiv E$ mit $C_1A = C_1G$, ein $C_2 \equiv E$ mit $C_2A = C_2F$ und ein $C = C_1C_2 \equiv E$ mit $CA = CG = CF$.

Weitere notwendige und hinreichende Bedingungen für β-Mengen ergeben sich mit Hilfe der beiden Mengen A_q, A_p der Punkte erster und zweiter Kategorie von A. Setzen wir $B = E - A$. Für jede Menge A ist

---

[141] [1]) Sogar ein Suslinsches (O. Nikodym).

[142] [2]) Der Leser, der die Lebesguesche Maßtheorie kennt, sei darauf hingewiesen, daß zwischen den meßbaren Mengen und den β-Mengen, insbesondere zwischen den Mengen vom Maße 0 und den Mengen erster Kategorie eine weitgehende Analogie besteht (W. Sierpiński). Die meßbaren Mengen sind die Mengen der Form

$$G_\delta - N \text{ oder } F_\sigma + N,$$

wo N eine Menge vom Maße 0 ist; man vergleiche dies mit (8).

nach (6)

(6*) $$A = A(A_p + A_q) \equiv A A_p = A_p - B A_p.$$

Hiernach sind die Kongruenzen

(9) $BA_p \equiv 0$, (10) $A \equiv A_p$

gleichbedeutend; sie sind für eine β-Menge A hinreichend (A ist mit einer abgeschlossenen Menge A_p kongruent), aber auch notwendig, denn für eine abgeschlossene Menge F ist nach (7) und (6*) $F_p \leqq F$, also $F \equiv FF_p = F_p$, und wenn $A \equiv F$ β-Menge ist, so ist $A_p = F_p$, also $A \equiv A_p$. (9) besagt, daß die nicht zu A gehörigen Punkte, in denen A von zweiter Kategorie ist, nur eine Menge erster Kategorie bilden.

Aus (10) gehen durch Komplementbildung sowie durch Vertauschung der Mengen A, B die Kongruenzen

(11) $$A \equiv A_p \equiv B_q, \quad B \equiv B_p \equiv A_q$$

hervor; jede einzelne davon ist für β-Mengen notwendig und hinreichend, z. B. auch $$A_p \equiv B_q,$$

denn hieraus folgt $BA_p \equiv BB_q \equiv 0$, die Kongruenz (9). Die Formeln (11) besagen, daß eine β-Menge A bis auf Mengen erster Kategorie aus den Punkten zweiter Kategorie von A oder den Punkten erster Kategorie von $B = E - A$ besteht. Endlich erwähnen wir noch die folgende Charakterisierung der β-Mengen:

(12) $A_q \dot{+} B_q$ dicht, $A_p B_p$ nirgendsdicht (im Raum E),

d. h. die Punkte, wo wenigstens eine der beiden Mengen A, B von erster Kategorie ist, bilden eine dichte (offene) Menge; diejenigen, wo A und B beide von zweiter Kategorie sind, eine nirgendsdichte (abgeschlossene). In der Tat ist (12) hinreichend, denn aus der allgemein gültigen Kongruenz $B \equiv BB_p$ folgt $BA_p \equiv BA_pB_p$ und im Falle (12) $BA_p \equiv 0$, also (9). Andererseits ist nach (7) $A_pB_p \leqq A_\alpha B_\alpha = A_g$ (Begrenzung von A); ist $A \equiv F$, $B \equiv G = E - F$, so ist $A_pB_p = F_pG_p \leqq F_g$; die Begrenzung einer abgeschlossenen Menge ist aber nirgendsdicht und (12) auch notwendig.

Wir nennen eine Menge A *α-abgeschlossen*, wenn sie mit ihrer abgeschlossenen Hülle kongruent ist ($A \equiv A_\alpha$); *α-offen*, wenn sie mit ihrem offenen Kern kongruent ist ($A \equiv A_i$); eine *α-Menge*, wenn sie beides ist, d. h. $A_g = A_\alpha - A_i \equiv 0$, wenn also die Begrenzung von A von erster Kategorie ist. Dies alles sind spezielle β-Mengen. Die α-abgeschlossenen und α-offenen Mengen sind Komplemente von einander.

Die Summe endlich vieler und der Durchschnitt abzählbar vieler α-abgeschlossener Mengen ist α-abgeschlossen; der Durchschnitt endlich vieler und die Summe abzählbar vieler α-offener Mengen ist α-offen.

Es genügt, die erste Behauptung zu beweisen. Für $A = \mathfrak{S} A_n$ mit endlich vielen Summanden ist $A_\alpha = \mathfrak{S} A_{n\alpha} \equiv \mathfrak{S} A_n = A$. Für $A = \mathfrak{D} A_n$

mit endlich oder abzählbar vielen Mengen ist $A_\alpha \leqq \mathfrak{D} A_{n\alpha} = A^*$ und $A^* \equiv \mathfrak{D} A_n = A$; aus $A \leqq A_\alpha \leqq A^*$ und $A \equiv A^*$ folgt $A \equiv A_\alpha$.

Wenn wir statt des σ-Moduls der Mengen erster Kategorie den Modul
[143] $\mathfrak{M}_0$ der *nirgendsdichten* Mengen zugrunde legen und Kongruenzen nach $\mathfrak{M}_0$ mit $A \sim B$ bezeichnen (d. h. $[A, B]$ ist nirgendsdicht), wobei wieder $E - A \sim E - B$ und aus $A_n \sim B_n$ für *endlich* viele n

$$\mathfrak{S} A_n \sim \mathfrak{S} B_n, \quad \mathfrak{D} A_n \sim \mathfrak{D} B_n$$

folgt, so kann man Analoga der β-Mengen und α-Mengen erklären, nämlich β_0-Mengen durch $A \sim G$ oder $A \sim F$ (es ist wieder $G \sim G_\alpha$, $F \sim F_i$) und α_0-Mengen durch $A_g \sim 0$, d. h. als Mengen mit nirgendsdichter Begrenzung definieren. Indessen fallen die β_0-Mengen hier mit den α_0-Mengen zusammen. Denn da mit D auch D_α nirgendsdicht ist, so folgt aus $B = A + D$: $B_\alpha = A_\alpha \dotplus D_\alpha \sim A_\alpha$ und aus $B \sim A$ oder $B = (A - D_1) + D_2$: $B_\alpha \sim (A - D_1)_\alpha \sim A_\alpha$, also ist mit $A \sim B$ zugleich $A_\alpha \sim B_\alpha$ und mittels zweimaliger Komplementbildung auch $A_i \sim B_i$. Ist also $A \sim F \sim G$, so ist $A_\alpha \sim F_\alpha = F$ und $A_i \sim G_i = G$, $A_\alpha \sim A_i$.

Nennen wir die *Mengen mit nirgendsdichter Begrenzung Mengen* N; aus ihnen entstehen die zuvor betrachteten allgemeineren Mengen auf folgende Weise:

II. *Die α-abgeschlossenen Mengen sind die N_δ, die α-offenen die N_σ; die β-Mengen sind die Mengen* Lim N_n (also gleichzeitig $N_{\sigma\delta}$ und $N_{\delta\sigma}$).

Zunächst ist jedes N_δ eine α-abgeschlossene Menge, da die N-Mengen α-abgeschlossen sind und die α-abgeschlossenen Mengen ein δ-System bilden. Andererseits sei A α-abgeschlossen, $A_\alpha - A = E_{\mathrm{I}}$ oder

$$A_\alpha = A + D_1 + D_2 + \cdots,$$

wo die D_n nirgendsdicht sind und disjunkt angenommen werden können. Mit

$$A_n = A + D_n + D_{n+1} + \cdots$$

folgt $A \leqq A_n \leqq A_\alpha$, $A_{n\alpha} = A_\alpha$, $A_{n\alpha} - A_n = D_1 + \cdots + D_{n-1}$ nirgendsdicht, also $A_n \sim A_{n\alpha}$ eine N-Menge, wonach $A = \mathfrak{D} A_n$ ein N_δ ist. Durch Komplementbildung ($E - N$ ist ein N) folgt die Behauptung über α-offene Mengen. Was die β-Mengen betrifft, so sei an die Definition von $\overline{\mathrm{Lim}}$, $\underline{\mathrm{Lim}}$ und Lim (S. 19, 20) erinnert; da nun offenbar

$$\underline{\mathrm{Lim}}\, A_n \cdot \underline{\mathrm{Lim}}\, B_n = \underline{\mathrm{Lim}}\, A_n B_n \leqq \overline{\mathrm{Lim}}\, A_n B_n \leqq \overline{\mathrm{Lim}}\, A_n \cdot \overline{\mathrm{Lim}}\, B_n,$$

so folgt aus der Existenz von Lim A_n, Lim B_n auch die von Lim $A_n B_n$ $=$ Lim $A_n \cdot$ Lim B_n, ebenso $\mathrm{Lim}(A_n \dotplus B_n) = \mathrm{Lim}\, A_n \dotplus \mathrm{Lim}\, B_n$. Wir stellen nun nach (8) eine β-Menge in der Form $A = G_\delta + E_{\mathrm{I}}$ oder

$$A = C + D, \quad C = \mathfrak{D} C_n, \quad D = \mathfrak{S} D_n$$

dar mit offenen Mengen $C_1 \geqq C_2 \geqq \ldots$ und nirgendsdichten Mengen $D_1 \leqq D_2 \leqq \ldots$, sodaß Lim $C_n = C$, Lim $D_n = D$; dann ist $A_n = C_n + D_n \sim C_n$ eine N-Menge und Lim $A_n = A$.

3. Die Bairesche Bedingung für Funktionen. Jedem Punkt x des metrischen Raumes E sei durch die Funktion

$$y = \varphi(x)$$

ein Punkt y des metrischen Raumes H eindeutig zugeordnet (vgl. S. 194); y heißt das Bild von x; das Bild einer Menge $A \leqq E$ (d. h. die Menge der Bilder der Punkte von A) werde mit $\varphi(A)$ bezeichnet. Das Bild $H_0 = \varphi(E)$ des ganzen Raumes E kann der ganze Raum H oder eine Teilmenge davon sein; wir nennen $y = \varphi(x)$ eine Abbildung von E *auf* H_0 oder von E *in* H (im Falle $H_0 = H$ also eine Abbildung von E auf H). Als das *Urbild* $\psi(B)$ einer Menge $B \leqq H$ bezeichnen wir die Menge aller Punkte x von E mit $\varphi(x) \varepsilon B$; offenbar ist $\psi(B) = \psi(H_0 B)$.

Der Punkt x heißt ein Stetigkeitspunkt von $\varphi(x)$, wenn für jede Folge $x_n \to x$ auch $\varphi(x_n) \to \varphi(x)$. *Die Menge C der Stetigkeitspunkte ist immer ein G_δ, die Menge D der Unstetigkeitspunkte ein F_σ* (Beweis wie S. 251 für reelle Funktionen). Wir erinnern noch an die topologische Charakterisierung der Stetigkeitspunkte x (S. 194): *für $y = \varphi(x)$ hat das Urbild jeder Umgebung V_y den Punkt x als inneren Punkt.* Dabei kann man sich auf die V_y beschränken, die einem *vollen Umgebungssystem* (S. 228) oder, wie wir kürzer sagen wollen, einer *Basis* $\mathfrak{B}$ des Raumes H angehören: *eine Basis ist ein System offener Mengen, aus denen durch Summenbildung alle offenen Mengen erhalten werden.* Läßt man also V alle Basismengen, $U = \psi(V)$ deren Urbilder durchlaufen, so ist x dann und nur dann Stetigkeitspunkt, wenn er innerer Punkt jeder Menge U ist, der er angehört, folglich dann und nur dann Unstetigkeitspunkt, wenn es mindestens ein U mit $x \varepsilon U - U_i = U_r$ (U_r Rand von U) gibt, und wir erhalten für die Menge D der Unstetigkeitspunkte die einfache Darstellung

$$D = \mathfrak{S} U_r, \tag{13}$$

wobei die Summe, wie gesagt, über die Urbilder $U = \psi(V)$ aller Mengen V einer Basis von H (oder, wegen $\psi(V) = \psi(H_0 V)$, von H_0) zu erstrecken ist. Für eine stetige Abbildung ($D = 0$) folgt daraus die uns bekannte Charakterisierung: die Urbilder aller offenen Mengen sind offen (§ 35, I).

Die Abbildung $y = \varphi(x)$ liefert, wenn man x auf eine Menge $A \leqq E$ beschränkt, eine Teilfunktion $\varphi(x|A)$ oder eine Teilabbildung von A in H (auf $\varphi(A) \leqq H$), bei der das Urbild einer Menge $B \leqq H$ offenbar $A\,\psi(B)$ ist.

Wir nennen $\varphi(x)$ eine *α-Funktion*, wenn die Menge D ihrer Unstetigkeitspunkte von erster Kategorie ($D \equiv 0$, $C \equiv E$) ist. Jede punktweise unstetige Funktion (S. 251) ist α-Funktion, und wenn $E - E_{\mathrm{I}}$ stets in E dicht, E ein G_{II}-Raum ist (S. 143), z. B. ein vollständiger, so ist jede α-Funktion auch punktweise unstetig.

Wir sagen, daß $\varphi(x)$ eine *β-Funktion* ist oder der *Baireschen Be-*

dingung (im weiteren Sinne) genügt, wenn es eine Menge $C \equiv E$ gibt, wofür $\varphi(x \mid C)$ stetig ist. Diese Funktionen sind also „bis auf Mengen erster Kategorie stetig" im Sinne von S. 261. Es sei nochmals daran erinnert, daß die Punkte von C hier Stetigkeitspunkte der Teilfunktion $\varphi(x \mid C)$, aber nicht notwendig der Gesamtfunktion $\varphi(x)$ sind.

Den Zusammenhang zwischen α-Funktionen und α-Mengen, β-Funktionen und β-Mengen vermitteln folgende Sätze.

III. *Für eine α-Funktion ist notwendig und bei höchstens separablem H auch hinreichend, daß das Urbild jeder offenen* (abgeschlossenen) *Menge α-offen* (α-abgeschlossen) *sei.*

Das folgt aus (13). Wird die Basis $\mathfrak{B}$ von allen offenen Mengen gebildet, so ist mit $D \equiv 0$ jedes $U_r \equiv 0$, $U \equiv U_i$, das Urbild jeder offenen Menge V ist α-offen und das Urbild $E - U = \psi(H - V)$ jeder abgeschlossenen Menge $H - V$ ist α-abgeschlossen. Umgekehrt folgt aus $U_r \equiv 0$ bei höchstens separablem H, indem man eine höchstens abzählbare Basis wählt, $D \equiv 0$. (Es genügt natürlich, daß H_0 höchstens separabel sei.)

[144] IV. *Für eine β-Funktion ist notwendig und bei höchstens separablem H auch hinreichend, daß das Urbild jeder offenen* (oder jeder abgeschlossenen oder jeder Borelschen) *Menge eine β-Menge sei.*

Bei der Teilabbildung $\varphi(x \mid C)$ ist das Urbild von V, wie wir oben bemerkten, $C\psi(V) = CU$ und, falls $\varphi(x \mid C)$ stetig ist, in C offen: $CU = CG$; hieraus folgt für $C \equiv E$: $U \equiv G$, U ist β-Menge, falls $\varphi(x)$ β-Funktion ist. Die eingeklammerten Behauptungen folgen durch Komplement-, Summen- und Durchschnittsbildung. — Umgekehrt, ist jedes $U = \psi(V)$ eine β-Menge und H höchstens separabel, so durchlaufe V eine höchstens abzählbare Basis; wir ordnen jedem U eine offene Menge $G \equiv U$ zu und bilden die Summe der Diskrepanzen $D = \mathfrak{S}[U, G] \equiv 0$, sowie $C = E - D \equiv E$; dann ist $CU = CG$, d. h. bei der Teilfunktion $\varphi(x \mid C)$ ist das Urbild CU jeder Basismenge V in C offen und folglich das Urbild jeder offenen Menge in C offen: $\varphi(x \mid C)$ ist stetig, $\varphi(x)$ eine β-Funktion.

Die charakteristische Funktion $\varphi(x)$ einer Menge A ist dann und nur dann α-Funktion (β-Funktion), wenn A α-Menge (β-Menge) ist.

Dies folgt aus III IV, wenn man den Raum $H = \{0, 1\}$ nur aus den beiden Zahlen 0 und 1 bestehen läßt; alle vier Teilmengen von H sind zugleich offen und abgeschlossen und ihre Urbilder sind E, A, $E - A$ und 0. — Übrigens folgt die Behauptung für α-Funktionen unmittelbar daraus, daß A_g die Menge der Unstetigkeitspunkte von $\varphi(x)$ ist; insbesondere ist $\varphi(x)$ dann und nur dann stetig, wenn A zugleich offen und abgeschlossen ist, $\varphi(x \mid C)$ dann und nur dann stetig, wenn CA in C zugleich offen und abgeschlossen ist; die Behauptung über β-Funktionen folgt dann aus der S. 280 angegebenen charakteristischen Eigenschaft der β-Mengen.

Wir betrachten jetzt eine Folge von Abbildungen $y_n = \varphi_n(x)$ des Raumes E in den Raum H, die für jedes x konvergiert, sodaß

$$y = \lim y_n = \lim \varphi_n(x) = \varphi(x)$$

wieder eine Abbildung von E in H ergibt; $\varphi(x)$ heiße die Grenzfunktion der konvergenten Folge $\varphi_n(x)$.

V. *Die Grenzfunktion einer konvergenten Folge stetiger Funktionen ist eine α-Funktion.*

VI. *Die Grenzfunktion einer konvergenten Folge von β-Funktionen ist eine β-Funktion.*

Diese Sätze von Kuratowski sind den Sätzen § 42, V und § 43, VI ähnlich (auch in den Beweisen, wie sich sofort zeigen wird), haben aber den Vorzug der Allgemeingültigkeit für jeden Raum E. Sie sagen aus, daß eine gewisse Menge D von erster Kategorie, ihr Komplement $C \equiv E$ ist; ist *insbesondere* E eine G_{II}-Menge, so folgt daraus nachträglich daß C in E dicht ist.

Beweis von V. Die $\varphi_n(x)$ seien stetig. Die Menge C der Stetigkeitspunkte von $\varphi(x)$ bestimmt sich dann folgendermaßen (vgl. § 42, 3): es sei, für $\sigma > 0$, $P_m(\sigma)$ die Menge der Punkte x, wo

$$(\alpha) \qquad yy_m \leqq \sigma,$$

und

$$G(\sigma) = \mathfrak{S}_m P_{mi}(\sigma)$$

die Summe der offenen Kerne dieser Mengen für $m = 1, 2, \ldots$; dann ist

$$C = \mathfrak{D}_{\sigma > 0} G(\sigma),$$

wofür man wegen der monotonen Abnahme von $G(\sigma)$ mit σ auch

$$C = \mathfrak{D}_n G\left(\frac{1}{n}\right)$$

setzen kann. Nämlich: ist a Stetigkeitspunkt von $\varphi(x)$, so wählen wir auf Grund von $b_n = \varphi_n(a) \to \varphi(a) = b$ ein m so groß, daß $bb_m \leqq \frac{\sigma}{3}$ und dann, wegen der Stetigkeit von $\varphi(x)$ und $\varphi_m(x)$ an der Stelle a, eine Umgebung U_a, in der $by \leqq \frac{\sigma}{3}$, $b_m y_m \leqq \frac{\sigma}{3}$; dann ist in U_a $yy_m \leqq \sigma$, also $a \varepsilon P_{mi}(\sigma) \leqq G(\sigma)$; da dies für jedes σ gilt, ist $C \leqq \mathfrak{D} G(\sigma)$. — Umgekehrt sei $a \varepsilon \mathfrak{D} G(\sigma)$ und, für ein bestimmtes σ und zugehöriges m, $a \varepsilon P_{mi}\left(\frac{\sigma}{3}\right)$, sodaß es eine Umgebung U_a gibt, in der $yy_m \leqq \frac{\sigma}{3}$, insbesondere auch $bb_m \leqq \frac{\sigma}{3}$; wegen der Stetigkeit von $\varphi_m(x)$ kann man U_a überdies so wählen, daß $b_m y_m \leqq \frac{\sigma}{3}$.

Also ist $by \leqq \sigma$ in U_a; da dies für jedes σ gilt, ist a Stetigkeitspunkt von $\varphi(x)$.

Ferner sei $F_m(\sigma)$ die Menge der x, wo

$$(\beta) \quad y_m y_n \leqq \sigma \quad \text{für } n = m+1, \quad m+2, \ldots;$$

diese Menge ist abgeschlossen (für jedes einzelne n definiert die Ungleichung (β) eine abgeschlossene Menge wegen der Stetigkeit von $\varphi_m(x)$ und $\varphi_n(x)$; für die sämtlichen $n \geqq m+1$ den Durchschnitt dieser abgeschlossenen Mengen). Wegen der Konvergenz der Folge $\varphi_n(x)$ ist, für jedes x, (β) für ein geeignetes m erfüllt; hieraus folgt

$$E = \underset{m}{\mathfrak{S}} F_m(\sigma).$$

Andererseits folgt aber (α) aus (β), also $F_m(\sigma) \leqq P_m(\sigma), F_{mi}(\sigma) \leqq P_{mi}(\sigma)$,

$$\underset{m}{\mathfrak{S}} F_{mi}(\sigma) \leqq G(\sigma).$$

Da nun stets $F_i = F$, so ist $\underset{m}{\mathfrak{S}} F_{mi}(\sigma) \equiv E$ und erst recht $G(\sigma) \equiv E$, also $G\left(\frac{1}{n}\right) \equiv E$ und $C = \mathfrak{D} G\left(\frac{1}{n}\right) \equiv E$, $\varphi(x)$ α-Funktion.

Beweis von VI. $\varphi_n(x)$ sei β-Funktion, $\varphi_n(x \mid C_n)$ mit $C_n \equiv E$ stetig, $C_0 = \mathfrak{D} C_n \equiv E$. Die Funktionen $\varphi_n(x \mid C_0)$ sind stetig, ihre Grenzfunktion $\varphi(x \mid C_0)$ also in C_0 eine α-Funktion, d. h. ihre Stetigkeitspunkte bilden eine Menge C, für die $C_0 - C$ in C_0 und erst recht in E von erster Kategorie ist: $C \equiv C_0 \equiv E$. $\varphi(x \mid C)$ ist stetig, $\varphi(x)$ β-Funktion.

Die β-Funktionen bilden also ein Bairesches System, dem die stetigen Funktionen und alle Baireschen Funktionen des Raumes E angehören.

VII. *Jede β-Funktion ist Grenzfunktion einer Folge von α-Funktionen.*

$\varphi(x)$ sei β-Funktion, $\varphi(x \mid C)$ stetig, $E - C = D = E_{\mathrm{I}}$; D ist Summe einer Folge nirgendsdichter Mengen. Wenn man diese durch ihre abgeschlossenen Hüllen ersetzt, vergrößert sich D, bleibt aber ein E_{I}, C verkleinert sich und die verkleinerte Teilfunktion bleibt stetig: wir können mithin alsbald $D = \mathfrak{S} F_k$ als Summe abgeschlossener nirgendsdichter Mengen F_k und diese als aufsteigend, also

$$D = F_1 + (F_2 - F_1) + (F_3 - F_2) + \cdots = \Sigma(F_k - F_{k-1})$$

mit disjunkten Summanden $F_k - F_{k-1}$ voraussetzen ($F_0 = 0$). Jedem Punkt $x \varepsilon D$ entspricht also eine bestimmte natürliche Zahl $k(x)$ mit $x \varepsilon F_{k(x)} - F_{k(x)-1}$; wenn eine Folge von Punkten $x_n \varepsilon D$ nach einem Punkt $c \varepsilon C$ konvergiert, so ist für jedes k schließlich $x_n \bar{\varepsilon} F_k$ (da sonst auch c zu der abgeschlossenen Menge F_k gehören würde), also $k(x_n) > k$, mithin $k(x_n) \to \infty$. Wir definieren nun eine Abbildung $\alpha(x)$ von E in H (sogar auf $\varphi(C)$), indem wir für $x \varepsilon C$ $\alpha(x) = \varphi(x)$ setzen, hingegen für $x \varepsilon D$ einen Punkt $c(x) \varepsilon C$ wählen, der von x eine Entfernung $< \delta(x, C) + \frac{1}{k(x)}$ hat ($\delta(x, C) \geqq 0$ die untere Entfernung des x von C), und dann $\alpha(x) = \varphi(c(x))$ erklären. Diese Funktion, behaupten wir, ist in jedem Punkt $c\varepsilon C$ stetig, also eine α-Funktion. Es ist, für $x_n \to c$, $\alpha(x_n) \to \alpha(c)$ zu zeigen, wobei man sich auf die beiden Fälle beschränken kann, daß alle x_n zu C oder alle zu D gehören. Für $x_n \varepsilon C$ ist $\alpha(x_n) = \varphi(x_n) \to \varphi(x) = \alpha(x)$ wegen der Stetigkeit von $\varphi(x \mid C)$. Für $x_n \varepsilon D$ ist $\delta(x_n, C) \to 0, k(x_n) \to \infty$, die Entfernung zwischen x_n und $c(x_n)$ konvergiert nach 0, also $c(x_n) \to c$ und $\alpha(x_n) = \varphi(c(x_n)) \to \varphi(c) = \alpha(c)$, wieder wegen der Stetigkeit von $\varphi(x \mid C)$. — Nunmehr sei

$$\varphi_k(x) = \varphi(x) \text{ für } x\varepsilon C + F_k$$
$$\varphi_k(x) = \alpha(x) \text{ für } x\varepsilon D - F_k.$$

Auch diese Funktionen sind in $c\varepsilon C$ stetig, denn für $x_n \to c$, $x_n\varepsilon D$ ist schließlich $x_n\varepsilon D - F_k$, $\varphi_k(x_n) = \alpha(x_n) \to \alpha(c) = \varphi(c) = \varphi_k(c)$. Andererseits ist $\lim\limits_k \varphi_k(x) = \varphi(x)$, sogar $\varphi_k(x) = \varphi(x)$ für alle k, falls $x\varepsilon C$, und für $k \geqq k(x)$, falls $x\varepsilon D$. Damit ist $\varphi(x)$ als Grenzfunktion der α-Funktionen $\varphi_k(x)$ dargestellt.

In gewissen Fällen, z. B. wenn es sich um reelle Funktionen $\varphi(x)$ handelt (H Raum der reellen Zahlen), ist jede β-Funktion sogar Grenzfunktion von solchen Funktionen, bei denen die Menge der Unstetigkeitspunkte *nirgendsdicht* (nicht nur von erster Kategorie) ist, also von *N-Funktionen*, wie wir diese Funktionen analog zu den N-Mengen mit nirgendsdichter Begrenzung nennen wollen: die charakteristische Funktion einer N-Menge ist eine N-Funktion und umgekehrt. Ist $\varphi(x)$ eine β-Funktion, $\varphi(x \mid C)$ stetig, wobei wir wie beim Beweis von VII $D = E - C = \mathfrak{S}F_k$ mit aufsteigenden nirgendsdichten abgeschlossenen Mengen F_k annehmen, so läßt sich nach dem Erweiterungssatz § 43, III die in $C = G_\delta$ stetige Funktion $\varphi(x \mid C)$ zu einer in E definierten Funktion erster Klasse

$$f(x) = \lim f_k(x)$$

erweitern, d. h. zur Grenzfunktion einer Folge (in E) stetiger Funktionen, wobei also $f(x) = \varphi(x)$ in C. Setzen wir dann

$$\varphi_k(x) = f_k(x) \quad \text{für} \quad x\varepsilon E - F_k$$
$$\varphi_k(x) = \varphi(x) \quad \text{für} \quad x\varepsilon F_k,$$

so ist $\varphi_k(x)$ in den Punkten der offenen Menge $E - F_k$ stetig, die Menge ihrer Unstetigkeitspunkte nirgendsdicht ($\leqq F_k$). Es ist $\lim\limits_k \varphi_k(x) = \varphi(x)$, nämlich in C $\varphi_k(x) = f_k(x) \to f(x) = \varphi(x)$, und in jedem Punkt von D schließlich $\varphi_k(x) = \varphi(x)$. — Allgemeiner gilt, wie wir ohne Beweis angeben: wenn H oder H_0 höchstens separabel ist, so ist jede β-Funktion Grenzfunktion von N-Funktionen, wie jede β-Menge Limes von N-Mengen (Satz II).

4. Die engere Bairesche Bedingung. Ist M eine Teilmenge von E, so heiße $\varphi(x)$ eine β_M-Funktion, wenn $\varphi(x \mid M)$ bis auf Mengen erster

Kategorie (in M) stetig oder im Raum M eine β-Funktion ist, d. h. wenn eine Spaltung $M = C + D$ möglich ist mit $D = M_{\mathrm{I}}$ und stetigem $\varphi(x \mid C)$. Man kann hierbei sowohl von der abgeschlossenen Hülle M_α als auch vom insichdichten Kern M_k auf M selbst schließen, d. h. es gilt:

Ist $F = M_\alpha$, so ist eine β_F-Funktion zugleich β_M-Funktion.

Denn ist $F = C + D$, D in F von erster Kategorie, $\varphi(x \mid C)$ stetig, so ist $\varphi(x \mid MC)$ stetig und MD in F, also (§ 27, XIII) auch in M von erster Kategorie.

Ist $K = M_k$, so ist eine β_K-Funktion zugleich β_M-Funktion.

Denn (vgl. S. 262 oben) ist

$$M = J + (S - J) + K,$$

wo J der isolierte, S der separierte Bestandteil von M ist, so ist MJ_β in MJ_α abgeschlossen und nirgendsdicht, da sein Komplement $MJ = J$ in MJ_α dicht ist, um so mehr MJ_β ein M_{I}. Nach § 23, (15) ist $S \leqq J_\alpha$, $S - J \leqq J_\beta$. Ferner gibt es ein D, von erster Kategorie in K und um so mehr in M, wofür $\varphi(x \mid K - D)$ stetig ist. Nach Tilgung von $MJ_\beta \dot{+} D$ aus M bleibt $J + C$, wo $\varphi(x \mid C)$ stetig und $CJ_\beta = 0$; daher ist auch $\varphi(x \mid J + C)$ in den Punkten von C und in den (isolierten) Punkten von J stetig, $\varphi(x)$ eine β_M-Funktion.

Wir sagen, $\varphi(x)$ genüge der *engeren Baireschen Bedingung* oder sei eine *β^*-Funktion*, wenn sie für jedes $M \leqq E$ eine β_M-Funktion ist. Dazu genügt bereits, daß sie für jedes *perfekte* P eine β_P-Funktion sei; denn man kann für $M_\alpha = F$ und $F_k = P$ von P auf F, von F auf M schließen. Die β_M-Funktionen bilden bei festem M ein Bairesches System, dem die Baireschen Funktionen des Raumes E angehören, und dasselbe gilt vom Durchschnitt aller dieser Systeme, d. h. vom System der β^*-Funktionen.

Die Menge A heißt eine β_M-Menge, wenn MA im Raume M eine β-Menge ist, wenn also eine in M offene Menge MG (G in E offen) existiert, wofür $[MA, MG] = M[A, G]$ in M von erster Kategorie ist. Das ist damit gleichbedeutend, daß die charakteristische Funktion $\varphi(x)$ von A eine β_M-Funktion ist, denn $\varphi(x \mid M)$ ist die charakteristische Funktion von MA für den Raum M. A genügt der engeren Baireschen Bedingung oder ist
[145] β^*-Menge, wenn sie für jedes M eine β_M-Menge ist; hierzu reicht aus, daß sie für jedes perfekte P eine β_P-Menge ist. Wenn insbesondere PA für jedes perfekte P in P von erster Kategorie ist, so ist A immer β_P-Menge und daher β^*-Menge (jedoch nicht etwa stets MA in M von erster Kategorie, z. B. sicher nicht, wenn M aus einem einzigen Punkt von A besteht). Das ist der Fall der S. 262 angegebenen Lusinschen Menge L.

§ 46. Halbschlichte Abbildungen. [146]

1. Trennbarkeit.

Zwei Mengen A, B des Raumes E heißen *trennbar*, wenn sie sich in zwei disjunkte Borelsche Mengen P, Q einschließen lassen:

$$A \leqq P, \quad B \leqq Q, \quad PQ = 0.$$

Das ist damit gleichbedeutend, daß A, B *disjunkt und in ihrer Summe Borelsche Mengen* sind. Denn aus der Trennbarkeit folgt

$$AB = 0, \quad A = (A + B)\, P, \quad B = (A + B)\, Q;$$

umgekehrt folgt aus diesen Gleichungen $(A + B)\, PQ = 0$ und

$$A \leqq P - PQ, \quad B \leqq Q - PQ.$$

In unsymmetrischer Form kann die Trennbarkeit von A, B auch so ausgedrückt werden: es gibt eine Borelsche Menge P, die A einschließt und B ausschließt: $A \leqq P$, $BP = 0$. Die Trennbarkeit von A und $E - B$ ist damit gleichbedeutend: es gibt eine Borelsche Menge P mit $A \leqq P \leqq B$.

Wenn für $m = 1, 2, 3, \ldots$ die Mengen A_m und B trennbar sind, so sind auch $A = \mathfrak{S} A_m$ und B trennbar. Denn aus $A_m \leqq P_m$, $BP_m = 0$ folgt mit $P = \mathfrak{S} P_m$: $A \leqq P$, $BP = 0$. — Wenn für $m, n = 1, 2, 3, \ldots$ die Mengen A_m und B_n trennbar sind, so sind auch $A = \mathfrak{S} A_m$ und $B = \mathfrak{S} B_n$ trennbar, umgekehrt also: wenn A, B nicht trennbar sind, so gibt es mindestens ein Paar nicht trennbarer Mengen A_m, B_n. — Wenn die Mengen A_m paarweise trennbar sind, so sind sie „simultan" trennbar, d. h. in paarweise disjunkte Borelsche Mengen einschließbar. Denn A_m und $\sum\limits_{n \neq m} A_n$ sind trennbar und in disjunkten Borelschen Mengen P_m, Q_m enthalten; dann ist A_m in $P_m \mathop{\mathfrak{D}}\limits_{n \neq m} Q_n$ enthalten, und diese Mengen sind disjunkt.

Bevor wir weitergehen, wollen wir nach N. Lusin zeigen, wie sich die beiden Hauptsätze § 34, III IV über Suslinsche Mengen mit dem Begriff der Trennbarkeit ohne Benutzung der unendlichen Ordnungszahlen (§ 34, 2) beweisen lassen. Es sei

$$A = \mathop{\mathfrak{S}}_{ikl\,\cdots} C_i C_{ik} C_{ikl} \ldots$$

eine Suslinsche Menge in dem zunächst noch beliebigen Räume E, gebildet von den abgeschlossenen Mengen $C_i \geqq C_{ik} \geqq C_{ikl} \geqq \ldots$; die Indizes durchlaufen alle natürlichen Zahlen. Hält man Anfangsindizes fest, so entstehen die Mengen

$$A_i = \mathop{\mathfrak{S}}_{kl\,\cdots}, \quad A_{ik} = \mathop{\mathfrak{S}}_{l\ldots}, \quad \ldots$$

mit

$$A = \mathop{\mathfrak{S}}_{i} A_i, \quad A_i = \mathop{\mathfrak{S}}_{k} A_{ik}, \quad A_{ik} = \mathop{\mathfrak{S}}_{l} A_{ikl}, \ldots$$

Ebenso sei

$$B = \mathop{\mathfrak{S}}_{pqr\cdots} D_p D_{pq} D_{pqr} \ldots$$

mit den entsprechenden Voraussetzungen und Bezeichnungen. Nehmen wir an, A und B seien nicht trennbar; wegen $A = \mathop{\mathfrak{S}}\limits_{i} A_i$, $B = \mathop{\mathfrak{S}}\limits_{p} B_p$ gibt es dann ein nicht trennbares Mengenpaar A_i, B_p, sodann ein nicht trennbares A_{ik}, B_{pq}, ein nicht trennbares A_{ikl}, B_{pqr} usw. Wegen $A_i \leqq C_i$, $B_p \leqq D_p$ (C_i, D_p abgeschlossen) ist dann $C_i D_p > 0$, ebenso $C_{ik} D_{pq} > 0$, $C_{ikl} D_{pqr} > 0, \ldots$ Ist nun E separabel, so können wir annehmen (S. 191), daß die Durchmesser der C, D mit wachsender Zahl der Indizes nach 0 konvergieren; ist er überdies vollständig, so hat die Folge der $C_i D_p$, $C_{ik} D_{pq}, \ldots$ einen Durchschnittspunkt, der offenbar zu AB gehört. D. h. wenn A, B nicht trennbar sind, so sind sie nicht disjunkt, oder:

[147] *Im separablen vollständigen Raum sind zwei disjunkte Suslinsche Mengen trennbar, d. h. in ihrer Summe Borelsche Mengen* (§ 34, III).

Ferner sei

$$A = \sum C_i C_{ik} C_{ikl} \ldots$$

mit disjunkten Summanden darstellbar; es ist dann

$$A = \sum_i A_i, \quad A_i = \sum_k A_{ik}, \quad A_{ik} = \sum_l A_{ikl}, \ldots$$

mit disjunkten Summanden. Ist E separabel und vollständig, so sind die disjunkten Suslinschen Mengen A_i paarweise trennbar, also simultan trennbar: es gibt disjunkte Borelsche Mengen $P_i \geqq A_i$, die man überdies durch $C_i P_i$ ersetzen, also $\leqq C_i$ annehmen kann. Ebenso gibt es disjunkte Borelsche Mengen $P_{ik} \geqq A_{ik}$, $P_{ik} \leqq C_{ik}$, wobei man noch P_{ik} durch $P_i P_{ik}$ ersetzen, d. h. $P_{ik} \leqq P_i$ annehmen darf. So erhält man

$$A_i \leqq P_i \leqq C_i, \quad A_{ik} \leqq P_{ik} \leqq C_{ik}, \quad A_{ikl} \leqq P_{ikl} \leqq C_{ikl}, \ldots$$

$$P_i \geqq P_{ik} \geqq P_{ikl} \geqq \ldots$$

Da aber $A_i A_{ik} A_{ikl} \ldots = C_i C_{ik} C_{ikl} \ldots$, so ist dies auch $= P_i P_{ik} P_{ikl} \ldots$,

$$A = \sum P_i P_{ik} P_{ikl} \ldots$$

und da die Mengen P mit gleichvielen Indizes paarweise disjunkt sind, so ist dies mit

$$A = \sum P_i \cdot \sum P_{ik} \cdot \sum P_{ikl} \ldots$$

gleichbedeutend, A ist Borelsch. *Im separablen vollständigen Raum ist jede mit disjunkten Summanden darstellbare Suslinsche Menge Borelsch* (§ 34, IV).

2. Erhaltung der *B*-Mengen. Die (eindeutige) Abbildung $y = \varphi(x)$ von A auf B hieß schlicht, wenn sie eindeutig umkehrbar, also das Urbild $\psi(y)$ jedes Punktes $y \varepsilon B$ einpunktig ist; sie heiße *halbschlicht*, B halbschlichtes Bild von A, wenn jedes Urbild $\psi(y)$ *höchstens abzählbar* ist.

Die Borelschen und Suslinschen Mengen eines *separablen vollständigen*
[148] Raumes mögen kurz B-Mengen und S-Mengen heißen. Wir wissen aus § 43, XIII:

Das Bairesche Bild einer S-Menge ist eine S-Menge; das schlichte Bairesche Bild einer B-Menge ist eine B-Menge.

Der Sonderfall *stetiger* Bilder ist bereits in § 37, II behandelt. Wir wollen zeigen, daß auch das halbschlichte Bairesche Bild einer B-Menge noch eine B-Menge ist und daß sich eine solche Abbildung in höchstens abzählbare viele schlichte Bairesche Abbildungen von B-Mengen spalten läßt. Wir stellen einen Satz voran, auf dem die Beweise beider Behauptungen beruhen werden:

I. *Die Abbildung $y = \varphi(x)$ des vollständigen Raumes A sei stetig. Jeder dyadischen* (aus den Ziffern 1, 2 gebildeten) *Folge $p, q, r, \ldots$ sei eine Folge beschränkter, in A offener Mengen* [1])

$$U_p > U_{pq} > U_{pqr} > \cdots$$

mit Durchmessern $\to 0$ und den Bildern

$$V_p = \varphi(U_p), \quad V_{pq} = \varphi(U_{pq}), \quad V_{pqr} = \varphi(U_{pqr}), \ldots$$

derart zugeordnet, daß die 2^n Mengen U mit n Indizes paarweise disjunkt sind, hingegen ihre Bilder V gemeinsame Punkte haben:

$$(1) \qquad V_1 V_2 > 0, \quad V_{11} V_{12} V_{21} V_{22} > 0, \ldots$$

Dann hat die unabzählbare Menge

$$D = \Sigma U_p \dot{U}_{pq} U_{pqr} \cdots$$

ein einpunktiges Bild, die Abbildung ist also nicht halbschlicht.

Der Beweis ist äußerst einfach. Der Durchschnitt $U_p U_{pq} U_{pqr} \ldots$ kann durch den der abgeschlossenen Hüllen ersetzt werden und ist also einpunktig; verschiedenen dyadischen Folgen entsprechen verschiedene Punkte $x = U_p U_{pq} U_{pqr} \ldots$, D ist von der Mächtigkeit des Kontinuums [2]). Ist $y = \varphi(x)$, so konvergieren wegen der Stetigkeit auch die Durchmesser der Bilder $V_p, V_{pq}, V_{pqr}, \ldots$ nach 0. Ist $\xi = U_\pi U_{\pi\varkappa} U_{\pi\varkappa\varrho} \ldots$, so ist $V_p V_\pi > 0$, $V_{pq} V_{\pi\varkappa} > 0$, $V_{pqr} V_{\pi\varkappa\varrho} > 0, \ldots$ und da die Summe von zwei Mengen mit nichtleerem Durchschnitt einen Durchmesser hat, der höchstens der Summe der beiden Einzeldurchmesser gleich ist, so konvergieren auch die Durchmesser von $V_p \dot{+} V_\pi$, $V_{pq} \dot{+} V_{\pi\varkappa}$, $V_{pqr} \dot{+} V_{\pi\varkappa\varrho}, \ldots$ nach 0, d. h. $\eta = \varphi(\xi)$ fällt mit y zusammen. Also ist $\varphi(D) = y$ einpunktig, $\psi(y) \geqq D$ unabzählbar.

Die Schwierigkeit besteht hauptsächlich in der Verwirklichung der Bedingungen (1). In dem zunächst zu behandelnden Fall (Satz II) wird das Nichtverschwinden der Mengen (1) durch ihre Nicht-Trennbarkeit von einer gewissen anderen Menge bewiesen werden, in dem späteren (Satz IV) dadurch, daß die Mengen U mit n Indizes „kohärente" Systeme bilden.

---

[1]) Zur Abkürzung bedeute $P > Q$ (oder $Q < P$) soviel wie $P_i \geqq Q_\alpha$, der offene Kern von P enthält die abgeschlossene Hülle von Q. Für eine Folge $P > Q > R > \ldots$ ist offenbar $PQR \ldots = P_\alpha Q_\alpha R_\alpha \ldots = P_i Q_i R_i \ldots$

[2]) D ist übrigens dyadisches Diskontinuum (S. 134).

19*

Bei der Abbildung $y = \varphi(x)$ von A auf B bezeichnen wir den Punkt y von B als *einfaches* oder *mehrfaches Bild*, jenachdem sein Urbild $\psi(y)$ einpunktig oder mehrpunktig ist. Die Menge M der mehrfachen Bilder läßt sich offenbar in der Gestalt

$$M = \mathfrak{S}\, \varphi(U)\, \varphi(U') \tag{2}$$

darstellen, wo U, U' alle Paare disjunkter Mengen einer Basis (S. 283) von A durchläuft. Es folgt daraus:

(A) *Bei der stetigen oder Baireschen Abbildung einer S-Menge ist die Menge der mehrfachen Bilder eine S-Menge.*

Denn die in A offenen Mengen U, U' und ihre Bilder sind S-Mengen, und die Basis kann (höchstens) abzählbar gewählt werden.

(B) *Ist $y = \varphi(x)$ stetige Abbildung der B-Menge A in den separablen vollständigen Raum Y, $B = \varphi(A)$ das Bild von A und M die Menge der mehrfachen Bilder, so ist jede in $B - M$ enthaltene S-Menge S von $Y - B$ trennbar.*

Die disjunkten S-Mengen S und M sind (in Y) trennbar, es gibt eine B-Menge Q mit $S \leqq Q$, $QM = 0$. Da BQ in B Borelsch ist, so ist ihr Urbild $P = \psi(BQ)$ in A Borelsch (§ 35, II), also B-Menge. Alle Punkte von BQ sind wegen $QM = 0$ einfache Bilder und BQ ist als schlichtes stetiges Bild von P wieder B-Menge (also in Y, nicht nur in B Borelsch). Wir haben $S \leqq BQ \leqq B$, zwischen S und B liegt die B-Menge BQ, d. h. S ist von $Y - B$ trennbar.

(C) *Für $k = 1, \ldots, n$ seien $y_k = \varphi_k(x_k)$ stetige Abbildungen beliebiger B-Mengen A_k in den separablen vollständigen Raum Y, $B_k = \varphi_k(A_k)$ das Bild von A_k und M_k die Menge der mehrfachen Bilder für φ_k, endlich B eine die B_k enthaltende beliebige Menge $\leqq Y$. Wenn dann $M_1 \ldots M_n$ von $Y - B$ trennbar ist, so ist auch $B_1 \ldots B_n$ von $Y - B$ trennbar.*

Sei $D = B_1 \ldots B_n$. Zwischen $D M_1 \ldots M_n$ $(= M_1 \ldots M_n)$ und B läßt sich nach Annahme eine B-Menge einschieben. Wenn bereits für ein k der Reihe $n, n-1, \ldots, 1$ feststeht, daß sich zwischen $DM_1 \ldots M_k$ und B eine B-Menge Q einschalten läßt, so gilt dies, behaupten wir, auch für $DM_1 \ldots M_{k-1}$ (worunter für $k = 1$ die Menge D zu verstehen ist). In der Tat: sei

$$DM_1 \ldots M_k \leqq Q \leqq B;$$

dann ist $DM_1 \ldots M_{k-1}(Y - Q)$ eine S-Menge, enthalten in B_k (wegen $D \leqq B_k$) und disjunkt zu M_k (weil $DM_1 \ldots M_{k-1} M_k$ zu $Y - Q$ disjunkt ist), also enthalten in $B_k - M_k$, folglich nach (B) von $Y - B_k$ und erst recht von der kleineren Menge $Y - B$ trennbar. Wir haben also mit einer B-Menge Q'

$$DM_1 \ldots M_{k-1}(Y - Q) \leqq Q' \leqq B$$

sowie

$$DM_1 \ldots M_{k-1} Q \leqq Q \leqq B,$$

addiert

$$DM_1 \ldots M_{k-1} \leqq Q \dot{+} Q' \leqq B.$$

Damit ist der Schluß von k auf $k-1$ gerechtfertigt, D von $Y-B$ trennbar.

II. *Das halbschlichte stetige Bild eines separablen vollständigen Raumes A ist eine B-Menge.*

Vorausbemerkt sei: bei beliebigem $\delta > 0$ bilden diejenigen Mengen einer Basis des Raumes A, die Durchmesser $< \delta$ haben, wieder eine Basis; ein separabler Raum A hat also für jedes $n = 1, 2, \ldots$ eine abzählbare Basis $\mathfrak{B}_n$, die aus Mengen U von Durchmessern $< \frac{1}{n}$ besteht. Ferner: ist $G > 0$ in A offen, so bilden die einer Basis von A angehörigen Mengen $U < G$ (d. h. wie gesagt, $U_\alpha \leqq G$) eine Basis für den Raum G[1]). Bei separablem A hat also G eine höchstens abzählbare Basis $\mathfrak{B}_n(G)$, bestehend aus Mengen $U < G$ mit Durchmessern $< \frac{1}{n}$. — Demgemäß lassen wir $p, q, r, \ldots$ die natürlichen Zahlen durchlaufen und verstehen unter U_p die Mengen von $\mathfrak{B}_1$, unter U_{pq} die Mengen von $\mathfrak{B}_2(U_p)$, unter U_{pqr} die Mengen von $\mathfrak{B}_3(U_{pq})$ usw.[2]). Es ist also

$$U_p > U_{pq} > U_{pqr} > \ldots$$

und die Mengen U mit n Indices haben Durchmesser $< \frac{1}{n}$. Bei der Abbildung $y = \varphi(x)$ seien $V_p, V_{pq}, V_{pqr}, \ldots$ die Bilder von $U_p, U_{pq}, U_{pqr}, \ldots$

Wir beweisen II in der Form: wenn das stetige Bild $B = \varphi(A)$ von A keine B-Menge ist, so ist die Abbildung nicht halbschlicht. B liege im separablen vollständigen Raum Y.

Die Menge der mehrfachen Bilder in B ist nach (2)

$$M = \underset{p\pi}{\mathfrak{S}}\, V_p V_\pi,$$

die Summe über alle disjunkten Mengenpaare U_p, U_π der Basis $\mathfrak{B}_1$ erstreckt. Nun ist B von $Y-B$ nicht trennbar (d. h. eben: B keine Borelsche Menge in Y), nach (C) (für $n = 1$) auch M von $Y-B$ nicht trennbar, also einer der (höchstens abzählbar vielen) Summanden von M nicht von $Y-B$ trennbar; bei passender Numerierung

$$V_1 V_2 \text{ von } Y-B \text{ nicht trennbar.}$$

Die Menge der mehrfachen Bilder bei der Abbildung $\varphi(x \mid U_p)$ von U_p auf V_p sei M_p $(p = 1, 2)$; es ist

$$M_1 M_2 = \underset{\substack{q\varkappa \\ r\varrho}}{\mathfrak{S}} V_{1q} V_{1\varkappa} V_{2r} V_{2\varrho},$$

die Summe über die disjunkten Paare $U_{1q}, U_{1\varkappa}$ und $U_{2r}, U_{2\varrho}$ der Basen

[1]) Dies folgt daraus, daß nach dem Trennungsaxiom (6), S. 229 jede Umgebung U_1 von x eine Umgebung $U < U_1$ enthält.

[2]) Ist eine dieser Basen endlich, so denken wir uns eine Basismenge unendlich oft geschrieben.

$\mathfrak{B}_2(U_1)$ und $\mathfrak{B}_2(U_2)$ erstreckt. Mit $V_1 V_2$ ist nach (C) (für $n = 2$) auch $M_1 M_2$ und einer der Summanden, bei geeigneter Numerierung

$$V_{11} V_{12} V_{21} V_{22} \text{ von } Y - B \text{ nicht trennbar.}$$

So fortfahrend verwirklichen wir die Voraussetzungen von I, insbesondere die Ungleichungen (1) dadurch, daß die Mengen linker Hand von $Y - B$ nicht trennbar sind. Hiernach ist $\varphi(x)$ nicht halbschlicht und II bewiesen.

[149] III. *Das halbschlichte Bairesche Bild einer B-Menge ist wieder eine B-Menge.*

Sei $y = \varphi(x)$ halbschlichte Abbildung der B-Menge A auf B. Ist $\varphi(x)$ zunächst stetig, so beachten wir (§ 37, III), daß A vermöge einer schlichten stetigen Abbildung $x = \xi(t)$ Bild eines separablen vollständigen Raumes T ist, z. B. einer im Baireschen Nullraum abgeschlossenen Menge (den trivialen Fall, daß A und T nur endlich sind, können wir ausschließen). Dann ist $y = \varphi(\xi(t)) = \eta(t)$ halbschlichte stetige Abbildung von T auf B, also B nach II eine B-Menge.

Im allgemeinen Fall, daß $y = \varphi(x)$ Bairesche Abbildung von A in $Y (\geqq B)$ ist, verfahren wir wie beim Beweis von § 43, XIII (S. 269, 270). Im Produktraum (A, Y) ist die Menge C der Punkte $(x, \varphi(x))$ Borelsch, also eine B-Menge, da man Y als separabel und vollständig annehmen kann; B ist Projektion von C auf Y, also stetiges, in unserem Fall halbschlichtes Bild der B-Menge C und, nach dem bereits Bewiesenen, selbst B-Menge.

3. Spaltbarkeit.

Bei der Abbildung $y = \varphi(x)$ des Raumes A nennen wir eine Menge $P \leqq A$ *spaltbar*, wenn P die Summe einer Folge von *B-Mengen* ist, die durch φ schlicht abgebildet werden, also

$$P = \mathfrak{S} P_n, \quad P_n \text{ } B\text{-Menge}, \quad \varphi(x \mid P_n) \text{ schlicht.}$$

Dann ist natürlich P selbst B-Menge und $\varphi(x \mid P)$ halbschlicht. Übrigens können die Summanden P_n, wenn man will, auch disjunkt angenommen werden, indem man P_n durch $P_n - P_n \underset{m<n}{\mathfrak{S}} P_m$ ersetzt. (Das Wesentliche der Forderung ist, daß die P_n B-Mengen sein sollen; ohne diese Bestimmung wäre für halbschlichtes $\varphi(x)$ jede Menge P spaltbar).

Eine Summe abzählbar vieler spaltbarer Mengen ist spaltbar. Jede in einer spaltbaren Menge enthaltene B-Menge ist wieder spaltbar.

(D) *Bei der halbschlichten Baireschen Abbildung einer B-Menge ist die Menge der mehrfachen Bilder eine B-Menge.*

Das geht aus (2) hervor, wo die U und ihre Bilder nach III B-Mengen sind und die Basis (höchstens) abzählbar gewählt werden kann.

(E) $y = \varphi(x)$ *sei halbschlichte stetige Abbildung der B-Menge A in den Raum Y; für* $y \varepsilon Y$ *sei* $\psi(y)$ *das Urbild von y und* $\psi_j(y)$ *die Menge der isolierten Punkte von* $\psi(y)$. *Dann ist die Menge* $P = \sum_y \psi_j(y)$ *spaltbar* [1].

[1]) Die Summe nach y soll sich hier und im folgenden über $y \varepsilon Y$ erstrecken;

Es seien $U_1, U_2, \ldots$ die Mengen einer Basis von A, $V_n = \varphi(U_n)$ das Bild von U_n und bei der Abbildung $\varphi(x \mid U_n)$ sei M_n die Menge der mehrfachen, $Q_n = V_n - M_n$ die der einfachen Bilder. Die Mengen U_n, V_n, M_n, Q_n sind B-Mengen, ebenso das Urbild $P_n = U_n \psi(Q_n)$ von Q_n bei der Abbildung $\varphi(x \mid U_n)$ (denn Q_n ist in V_n, also P_n in U_n Borelsch). Dann ist $\varphi(x \mid P_n)$ schlicht und demgemäß die Menge $\mathfrak{S} P_n$ spaltbar. Das ist aber die Menge $P = \sum_y \psi_j(y)$. Denn ist $x \varepsilon \psi_j(y)$, so gibt es ein U_n mit $\psi(y)\, U_n = x$, also ist $y = \varphi(x)$ einfaches Bild bei $\varphi(x \mid U_n)$ und $x \varepsilon P_n$ — und vice versa.

Um die hiermit begonnene Abtrennung eines spaltbaren Bestandteils von A fortzusetzen, verstehen wir (vgl. S. 166 und 170) für die Ordnungszahlen $\xi < \Omega$ unter $\psi_\xi(y)$ die *Kohärenzen* von $\psi(y)$; d. h. $\psi_0(y) = \psi(y)$, $\psi_{\xi+1}(y)$ ist die Menge der in $\psi_\xi(y)$ liegenden Häufungspunkte von $\psi_\xi(y)$, also $\psi_\xi(y) - \psi_{\xi+1}(y)$ die Menge der isolierten Punkte von $\psi_\xi(y)$; für eine Limeszahl η ist $\psi_\eta(y) = \mathfrak{D}_{\xi<\eta} \psi_\xi(y)$. Entsprechend setzen wir

$$A_\xi = \sum_y \psi_\xi(y) \qquad (A_0 = A),$$

dann wird für $\xi < \eta$

$$A_\xi - A_\eta = \sum_y [\psi_\xi(y) - \psi_\eta(y)]$$

und die Gleichung

$$\psi(y) - \psi_\eta(y) = \sum_{\xi<\eta} [\psi_\xi(y) - \psi_{\xi+1}(y)]$$

gibt nach y summiert

$$(3) \qquad A - A_\eta = \sum_{\xi<\eta} (A_\xi - A_{\xi+1}).$$

Es gilt dann unter den Voraussetzungen von (E):

(F) *Die Mengen $A_\xi - A_{\xi+1}$ sind spaltbar.*

Dies ist für $\xi = 0$ nach (E) richtig. Ist es bereits für $\xi < \eta$ bewiesen, so ist nach (3) $A - A_\eta$ spaltbar, also B-Menge, A_η ebenfalls B-Menge. Bei der halbschlichten stetigen Abbildung $\varphi(x|A_\eta)$ ist $A_\eta \psi(y) = \psi_\eta(y)$ das Urbild von y und $\psi_\eta(y) - \psi_{\eta+1}(y)$ die Menge seiner isolierten Punkte; die Anwendung von (E) ergibt Spaltbarkeit von $A_\eta - A_{\eta+1}$.

Nach (3) ist nun jedes $A - A_\eta$ spaltbar und zur Spaltbarkeit der ganzen Menge A ist hinreichend, daß es einen Index η mit $A_\eta = 0$ gebe, so daß also $\psi_\eta(y) = 0$ für jedes y ist: man kann das so ausdrücken, daß die Mengen $\psi(y)$ „gleichmäßig" separiert sind, d. h. ihre kleinsten Kohärenzen (insichdichten Kerne) sind nicht nur leer, sondern werden allesamt bereits bei Indizes $\leqq \eta$ erreicht.

Bei einer beliebigen Abbildung $\varphi(x)$ des beliebigen Raumes A wollen wir sagen, daß die in A offenen Mengen $U_1, \ldots, U_n$ ein *kohärentes System*

---

die Summanden können teilweise leer sein, z. B. ist schon $\psi(y) = 0$, wenn y nicht zu $B = \varphi(A)$ gehört.

bilden, wenn der Durchschnitt

$$D_\xi = \varphi(A_\xi U_1) \ldots \varphi(A_\xi U_n)$$

der Bilder von $A_\xi U_1, \ldots, A_\xi U_n$ für alle $\xi < \Omega$ nichtleer ist; insbesondere ist dann der Durchschnitt $D_0 = \varphi(U_1) \ldots \varphi(U_n)$ nichtleer.

(G) *Wenn bei der beliebigen Abbildung $\varphi(x)$ des separablen Raumes A die n Mengen U_k $(k = 1, \ldots, n)$ ein kohärentes System bilden und $\mathfrak{B}(U_k)$ eine höchstens abzählbare Basis von U_k ist, so gibt es in dieser Basis zwei disjunkte Mengen U_{k1}, U_{k2} derart, daß die $2n$ Mengen U_{k1}, U_{k2} ein kohärentes System bilden.*

Sei $y_\xi \varepsilon D_{\xi+1}$, also für jedes k in $A_{\xi+1} U_k$ ein Urbildpunkt von y_ξ vorhanden: $\psi(y_\xi) A_{\xi+1} U_k = \psi_{\xi+1}(y_\xi) U_k > 0$; da $\psi_{\xi+1}(y)$ die Kohärenz von $\psi_\xi(y)$ ist, ist $\psi_\xi(y_\xi) U_k$ unendlich [1]). Es gibt also in $\mathfrak{B}(U_k)$ zwei von ξ abhängige disjunkte Mengen $U_{k1}(\xi)$, $U_{k2}(\xi)$, die mit $\psi_\xi(y_\xi)$ nichtleeren Durchschnitt haben. Aber von solchen Systemen von $2n$ Mengen U_{k1}, U_{k2} gibt es höchstens abzählbar viele verschiedene; mindestens eines muß unabzählbar oft vorkommen, und für unabzählbar viele ξ ist also

$$U_{k1}(\xi) = U_{k1}, \quad U_{k2}(\xi) = U_{k2},$$
$$\psi_\xi(y_\xi) U_{k1} > 0, \quad \psi_\xi(y_\xi) U_{k2} > 0,$$
$$\varphi(A_\xi U_{11}) \varphi(A_\xi U_{12}) \ldots \varphi(A_\xi U_{n1}) \varphi(A_\xi U_{n2}) > 0.$$

Diese letzte für unabzählbar viele ξ richtige Ungleichung gilt aber für alle ξ, da die A_ξ mit wachsendem ξ abnehmen, also bilden die $2n$ Mengen $U_{11}, \ldots, U_{n2}$ ein kohärentes System.

IV. *Bei einer halbschlichten stetigen Abbildung des separablen vollständigen Raumes A gibt es ein ξ mit $A_\xi = 0$; A ist also spaltbar.*

Wir beweisen: wenn alle $A_\xi > 0$, so ist die stetige Abbildung $y = \varphi(x)$ nicht halbschlicht. A selbst bildet ein kohärentes System, also gibt es in $\mathfrak{B}_1$ (mit den Bezeichnungen wie beim Beweise von II) zwei disjunkte U_1, U_2, die ein kohärentes System bilden, sodann in $\mathfrak{B}_2(U_p)$ zwei disjunkte U_{p1}, U_{p2} derart, daß U_{11}, U_{12}, U_{21}, U_{22} ein kohärentes System bilden usf. Damit tritt I in Kraft; die Bedingungen (1) sind jetzt dadurch verwirklicht, daß die Bilder der Mengen eines kohärenten Systems nichtleeren Durchschnitt haben.

[150] V. *Eine B-Menge ist bei jeder halbschlichten Baireschen Abbildung spaltbar, d. h. Summe abzählbar vieler B-Mengen, die schlicht abgebildet werden.*

Dies wird auf IV so zurückgeführt wie III auf II. Ist $y = \varphi(x)$ halbschlichte Abbildung der B-Menge A und zunächst stetig, so ist bei der halbschlichten stetigen Abbildung $y = \varphi(\xi(t)) = \eta(t)$ (vgl. den Beweis

---

[1]) Ist P_β die Ableitung, $P_h = PP_\beta$ die Kohärenz von P und U offen, so ist nach § 23, (13) $(PU)_\beta \geqq P_\beta U \geqq P_h U$; ist $P_h U > 0$, so kann PU nicht endlich sein.

von III) $T = \Sigma\, T_n$ in B-Mengen T_n spaltbar, die schlichte Bilder $B_n = \eta(T_n)$ haben; da auch T_n und $A_n = \xi(T_n)$ in schlichter Beziehung stehen, so auch A_n und $B_n = \varphi(A_n)$, $A = \Sigma\, A_n$ ist bei der Abbildung $\varphi(x)$ spaltbar.

Im allgemeinen Fall der Baireschen halbschlichten Abbildung $y = \varphi(x)$ von A auf B betrachten wir die Punkte $z = (x, y)$ des Produktraums (A, B) und die stetigen Funktionen $x = \xi(z)$, $y = \eta(z)$ (Projektionen von z auf A und B). Die B-Menge C der Punkte $(x, \varphi(x))$ wird durch $\xi(z)$ schlicht stetig auf A, durch $\eta(z)$ halbschlicht stetig auf B abgebildet; hierbei ist $C = \Sigma\, C_n$ in B-Mengen C_n spaltbar, die schlichte Bilder $B_n = \eta(C_n)$ und $A_n = \xi(C_n)$ haben; die B-Menge A_n wird durch $\varphi(x)$ schlicht auf B_n abgebildet und $A = \Sigma\, A_n$ ist bei der Abbildung $\varphi(x)$ spaltbar.

Nachträge.

[151] (A) Zu S. 143.

Der Raum E hieß eine F_{II}-Menge, wenn jede abgeschlossene Menge $F > 0$ in sich von zweiter Kategorie ist. Er heiße eine F_{I}-Menge, wenn er keine F_{II}-Menge ist, wenn es also eine abgeschlossene Menge $F > 0$ gibt, die in sich von erster Kategorie ist. Hierüber hat W. Hurewicz bemerkenswerte Ergebnisse erzielt:

Der Raum E ist dann und nur dann F_{I}-Menge, wenn er eine perfekte *abzählbare* Menge P enthält; also dann und nur dann F_{II}-Menge, wenn jede perfekte Menge $P > 0$ unabzählbar ist.

In einem separablen Raum E sind die Borelschen Mengen (und sogar die Komplemente Suslinscher Mengen) sämtlich F_{I}-Mengen mit eventueller Ausnahme der G_δ. Diese können F_{II}-Mengen sein; andererseits sind ja die G_δ eines vollständigen, auch eines nicht separablen, Raumes — die Youngschen Mengen — wirklich F_{II}-Mengen (S. 144).

Demselben Gedankenkreis gehören noch folgende beiden Sätze an, aus denen im Fall eines vollständigen Raumes E die Mächtigkeitssätze § 26, VIII und § 32, I entspringen:

Jede Menge G_δ im beliebigen Raum E, deren insichdichter Kern nicht verschwindet, enthält eine (in E) perfekte Teilmenge $P > 0$. Jede unabzählbare Suslinsche Menge im separablen Raum E enthält eine (in E) perfekte Teilmenge $P > 0$.

(B) Zu S. 159.

Zu der Schlußweise im Text wäre erforderlich, daß zwei Punkte von $P(0)$ in $P(0)$ selbst, nicht nur in A, die Distanz 0 haben. Der Zusammenhang von $P(0)$ ist etwa so zu beweisen. Nehmen wir an, es sei $P(0) = Q + R$ in die beiden in sich kompakten Mengen Q und R zerstückelt; verbinden wir einen Punkt $q \varepsilon Q$ und einen Punkt $r \varepsilon R$ durch eine $\frac{1}{n}$-Kette in A, was wegen $\overline{qr} = 0$ für jede natürliche Zahl n möglich ist. Achten wir für die Punkte x der Kette darauf, ob die untere Entfernung $\delta(x, Q)$ kleiner als oder mindestens gleich $\delta(x, R)$ ist; für den ersten Punkt q tritt der erste, für den letzten r der zweite Fall ein, und es muß also in der Kette einen Punkt x_n geben, für den noch $\delta(x_n, Q) < \delta(x_n, R)$, während für den darauf folgenden Punkt y_n schon $\delta(y_n, Q) \geqq \delta(y_n, R)$ ist; hierbei ist $\overline{qx_n} \leqq \frac{1}{n}$ und $x_n y_n \leqq \frac{1}{n}$. Solche Punktpaare gibt es also für jedes $n = 1, 2, 3, \ldots$ Die Folge x_n hat eine konvergente Teilfolge $x_\nu \to x$, wo $x \varepsilon A$, dann ist auch $y_\nu \to x$ und wegen der Stetigkeit der Distanzen und Entfernungen $\overline{qx} = 0$, $\delta(x, Q) = \delta(x, R)$. Also würde x zu $P(0) = Q + R$ gehören und das ergibt einen Widerspruch, da dann von den beiden Zahlen $\delta(x, Q)$ und $\delta(x, R)$ die eine 0 und die andere positiv ist.

(C) Zu § 39, 2.

Es sei $t = \varphi(x)$ eine stetige Abbildung des kompakten, zwischen zwei gewissen Punkten a, b irreduziblen Kontinuums C auf das reelle Zahlenintervall T $(0 \leqq t \leqq 1)$. Die Urbilder der Zahlen t mögen die *Schichten* der Abbildung φ heißen; $C(x, \varphi)$ sei die Schicht, der der Punkt x angehört, also die Menge aller y mit $\varphi(y) = \varphi(x)$. Die Abbildung werde *monoton* genannt, wenn alle Schichten *Kontinua* (ein- oder mehrpunktige) sind; z. B. ist in dem Hahnschen Satz VIII für ein zusammengesetztes Kontinuum C die Existenz einer monotonen Abbildung bewiesen, deren Schichten die Primteile sind. Nun geht aus einer Untersuchung von C. Kuratowski hervor: wenn C überhaupt monotone Abbildungen zuläßt und $C(x)$ den Durchschnitt aller $C(x, \varphi)$ für sämtliche monotone Abbildungen φ mit $\varphi(a) = 0$, $\varphi(b) = 1$ bezeichnet, so sind diese $C(x)$ wiederum die Schichten einer monotonen Abbildung, die also eine Spaltung von C in die kleinstmöglichen Schichten bewirkt. Diese Kuratowskischen Minimalschichten $C(x)$ verdienen den Namen Primteile eher als die in § 39, 2 so genannten Mengen, die möglicherweise noch in Minimalschichten gespalten werden können.

(D) Zu S. 269.

Die Frage ist von St. Banach behandelt worden mit folgendem Hauptergebnis. Man definiere für $0 \leqq \eta < \Omega$ die „modifizierten" Baireschen Funktionen $f^{\eta+1}$ durch die Induktionsvorschrift:

Die f^1 sind die Funktionen der Klasse $(M^1, N^1) = (F_\sigma, G_\delta)$.

Für $\eta > 0$ sind die $f^{\eta+1}$ die Grenzfunktionen konvergenter Folgen von Funktionen $f^{\xi+1}$ $(\xi < \eta)$.

Dann sind bei höchstens separablem Bildraum Y die Funktionen $f^{\eta+1}$ mit denen der Klasse $(M^{\eta+1}, N^{\eta+1})$ identisch.

Dadurch also, daß man als Ausgangspunkt der Limesbildung nicht die stetigen Funktionen, sondern die f^1 nimmt (die im Allgemeinen nicht Grenzfunktionen von Folgen stetiger Funktionen zu sein brauchen), wird der Satz XII bei höchstens separablem Y für die modifizierten Baireschen Funktionen umkehrbar. Die Frage bleibt offen, wie für eine Limeszahl η die Funktionen der Klasse (M^η, N^η) aus konvergenten Folgen von Funktionen niederer Klassen entstehen, oder welche Bedingung im allgemeinen Falle der Vollständigkeit von Systemen reeller Funktionen (S. 236, 258) entspricht.

Literatur.

R. Baire, Lecons sur les fonctions discontinues (Paris 1905).

St. Banach, Théorie des opérations linéaires (Warszawa 1932).

E. Borel, Leçons sur la théorie des fonctions (Paris 1898, 3. éd. 1928), zitiert als: Leçons (1898).

E. Borel, Leçons sur les fonctions de variables réelles (Paris 1905), zitiert als Leçons (1905).

E. Borel, Méthodes et problèmes de théorie des fonctions (Paris 1922).

C. Carathéodory, Vorlesungen über reelle Funktionen (Leipzig und Berlin 1918), zitiert als: Vorlesungen.

A. Fraenkel, Einleitung in die Mengenlehre (3. Aufl. Berlin 1928).

M. Fréchet, Les espaces abstraits (Paris 1928).

H. Hahn, Theorie der reellen Funktionen I (Berlin 1921), zitiert als: Theorie.

H. Hahn, Reelle Funktionen I (Leipzig 1932).

G. Hessenberg, Grundbegriffe der Mengenlehre (Göttingen 1906).

E. W. Hobson, The theory of functions of a real variable (Cambridge 1907, 3. ed. I 1927, 2. ed. II 1926).

E. Kamke, Mengenlehre (Berlin und Leipzig 1928).

C. Kuratowski, Topologie I (Warszawa-Lwów 1933).

N. Lusin, Leçons sur les ensembles analytiques (Paris 1930).

K. Menger, Dimensionstheorie (Leipzig und Berlin 1928).

J. Pierpont, Lectures on the theory of functions of real variables I II (Boston 1905, 1912).

A. Rosenthal, Neuere Untersuchungen über Funktionen reeller Veränderlichen (Math. Enzykl. II C 9).

A. Schoenflies, Die Entwickelung der Lehre von den Punktmannigfaltigkeiten I II (Leipzig 1900, 1908).

A. Schoenflies, Entwickelung der Mengenlehre und ihrer Anwendungen (Leipzig und Berlin 1913).

W. Sierpiński, Leçons sur les nombres transfinis (Paris 1928).

W. Sierpiński, Hypothèse du continu (Warszawa-Lwów 1934).

W. Sierpiński, Introduction to general topology (Toronto 1934).

H. Tietze und L. Vietoris, Beziehungen zwischen den verschiedenen Zweigen der Topologie (Math. Enzykl. III A B 13).

C. de la Vallée Poussin, Intégrales de Lebesgue, Fonctions d'ensemble, Classes de Baire (Paris 1916).

W. H. Young and Grace Chisholm Young, The theory of sets of points (Cambridge 1906).

Die von St. Mazurkiewicz und W. Sierpinski herausgegebene Zeitschrift Fundamenta Mathematicae (Warschau, seit 1920) ist wesentlich der Mengenlehre gewidmet.

Quellenangaben.

Die folgenden Zitate, zu deren Ergänzung wir auf die Berichte von A. Schoenflies und das Referat von A. Rosenthal verweisen, sollen nur über den Ursprung der wichtigsten Begriffe und Theoreme Rechenschaft geben. Übrigens sind die meisten Sätze über Punktmengen zunächst nur für den eindimensionalen oder höchstens Euklidischen Raum ausgesprochen und hier erheblich umgestaltet worden. Bei Zitaten aus Büchern ist die Seitenzahl angegeben, bei Zitaten aus Zeitschriften Bandzahl (zuweilen eingeklammert Serienzahl), Jahr und Seitenzahl.

S. 10. G. Kowalewski, Einführung in die Infinitesimalrechnung (Leipzig 1908), 14.

§ 1. Von Zitaten Cantors wurde im allgemeinen abgesehen; der Inhalt der ersten vier Kapitel und die Grundbegriffe der Punktmengentheorie gehen fast ausschließlich auf ihn zurück. Von seinen zahlreichen Abhandlungen kommen hauptsächlich in Betracht: Über unendliche lineare Punktmannigfaltigkeiten I—VI: Math. Ann. 15 (1879), 1; 17 (1880), 355; 20 (1882), 113; 21 (1883), 51, 545; 23 (1884), 453. Beiträge zur Begründung der transfiniten Mengenlehre I II: Math. Ann. 46 (1895), 481; 49 (1897), 207.

S. 11. A. Genocchi-G. Peano, Differentialrechnung und Grundzüge der Integralrechnung (Leipzig 1899), 340.

§ 3. S. 17. C. Carathéodory, Vorlesungen, 23.

S. 19. Oberer und unterer Limes als ensemble limite complet und ensemble limite restreint bei E. Borel, Leçons (1905), 18.

S. 20. C. de la Vallée Poussin, Intégrales, 7.

§ 5. III. F. Bernstein in E. Borel, Leçons (1898), 103.

§ 7. S. 34. E. Zermelo, Math. Ann. 65 (1908), 261. Über die Antinomienfrage findet der Leser Näheres bei A. Fraenkel, Einleitung in die Mengenlehre (1928).

III. J. König, Math. Ann. 60 (1904), 177.

§ 11. I. F. Bernstein, Untersuchungen aus der Mengenlehre (Diss. Halle 1901), 34.

S. 53. H. Minkowski, Verh. d. Heidelb. Kongr. (Leipzig 1905), 171.

S. 53. R. Dedekind, Stetigkeit und irrationale Zahlen (Braunschweig 1872).

§ 12. E. Zermelo, Math. Ann. 59 (1904), 514 und 65 (1908), 107.

§ 13. Zur Wohlordnungstheorie hat G. Hessenberg, Grundbegriffe der Mengenlehre (1906) Erhebliches beigetragen.

§ 14. S. 68. G. Hessenberg, Grundbegriffe, § 75.

§ 15. S. 71. Zur Frage $\aleph \geqq \aleph_1$ vgl. D. Hilbert, Math. Ann. 95 (1925), 181.

§ 16. F. Hausdorff, Leipz. Ber. 58 (1906), 108.

§ 18. E. Borel, Leçons (1898), 46.

§ 19. M. Souslin[1]), Compt. rend. 164 (1917), 88. N. Lusin, ib. 91. Vgl. § 32.

§ 20. 1. Bei M. Fréchet, Rend. Pal. 22 (1906), 17, 30 heißt ein metrischer Raum classe (E) (Entfernung = écart), in späteren Arbeiten classe (D) (distance).

---

[1]) Michael Suslin (1894—1919) hat nur diese eine Arbeit selbst veröffentlicht. Der russische Name ist deutsch Suslin auszusprechen.

§ 20, 3. S. 98. D. Hilbert, Götting. Nachr. 8 (1906), 157.
S. 99. H. Minkowski, Geometrie der Zahlen (Leipzig und Berlin 1910), 116.

§ 20, 4. R. Baire, Acta math. 32 (1909), 105.

§ 20, 5. K. Hensel, Theorie der algebraischen Zahlen I (Leipzig und Berlin 1908). J. Kürschák, Journ. für Math. 142 (1913), 212.

§ 21, 1. Vollständiger Raum der Sache nach bei M. Fréchet, Rend. Pal. 22 (1906), 23.

§ 21, 3. S. 106. G. Cantor, Math. Ann. 5 (1872), 123; Ch. Méray, Nouveau précis d'analyse infinitésimale (Paris 1872).

§ 21, 4. M. Fréchet, Rend. Pal. 22 (1906), 6.

§ 22. Offene Menge: H. Lebesgue, Ann. di mat. (3) 7 (1902), 242; C. Carathéodory, Vorlesungen, 40. In der ersten Auflage wurden die offenen Mengen als Gebiete bezeichnet, worunter wir jetzt zusammenhängende offene Mengen verstehen.

§ 23. Die meisten Begriffe dieses § stammen von G. Cantor. Verdichtungspunkt: E. Lindelöf, Acta math. 29 (1905), 184. Insichdichter Kern: H. Hahn, Theorie, 76.

§ 25. Die Sätze dieses § stammen meist von G. Cantor.
S. 125. Separabel (in etwas anderem Sinne): M. Fréchet, Rend. Pal. 22 (1906), 23.
IV. F. Bernstein, Diss. (Halle 1901), 44.

§ 26, 1. II. E. Borel, Ann. éc. norm. (3) 12 (1895), 51.
III. E. Lindelöf, Compt. rend. 137 (1903), 697; W. H. Young, Lond. math. soc. proc. 35 (1903), 384.
IV. W. Sierpiński, Bull. Ac. Cracovie (1918), 49.

§ 26, 3. VIII. W. H. Young, Leipz. Ber. 55 (1903), 287; Theory, 64.

§ 27. S. 142. R. Baire, Ann. di Mat. (3) 3 (1899), 67.

§ 28, 1. Entfernung $\overline{AB}$: D. Pompéju, Ann. Fac. Toulouse (2) 7 (1905), 281.

§ 28, 2. Die Mengen $\overline{F}, \underline{F}$ der Sache nach bei P. Painlevé; siehe L. Zoretti, Journ. de math. (6) 1 (1905), 8; Bull. soc. math. France 37 (1909), 116.

§ 29, 1. Zusammenhang: N. J. Lennes, Amer. J. of Math. 33 (1911), 303. Kontinuum: C. Jordan, Cours d'analyse I (2. éd. Paris 1893), 25. Gebiet (auf Weierstraß zurückgehend) C. Carathéodory, Vorlesungen, 208, 222.
III. Z. Janiszewski und C. Kuratowski, Fund. math. 1 (1920), 211.

§ 29, 2. S. 155. H. Hahn, Jahresber. D. Math. V. 23 (1914), 319; Wiener Ber. 123 (1914), 2433.
X. H. Hahn, Fund. math. 2 (1921), 189. C. Kuratowski, Fund. math. 1 (1920), 43.
S. 157. St. Mazurkiewicz, Fund. math. 1 (1920), 167.

§ 29, 3. S. 158. G. Cantor, Math. Ann. 21 (1883), 576.
XV. L. E. J. Brouwer, Amst. Ak. Proc. 12 (1910), 785.
XVII. Z. Janiszewski, Thèse (Paris 1911) = Journ. éc. polyt. (2) 16 (1912), 100.
S. 162. W. Sierpiński, Tôhoku Math. J. 13 (1918), 300.

§ 29, 4. XIX. L. Zoretti, Journ. de math. (6) 1 (1905), 8.

§ 30, 4. VII. Z. Zalcwasser, Fund. math. 3 (1922), 44.

§ 31, 1. S. 174. G. Hamel, Math. Ann. 60 (1905), 459.
S. 175. L. Zoretti, Ann. éc. norm. (3) 26 (1909), 487.
S. 175. Z. Janiszewski, Compt. rend. 151 (1910), 198.

§ 31, 2. S. 176. F. Bernstein, Leipz. Ber. 60 (1908), 325.
S. 177. W. Sierpiński, Fund. math. 1 (1920), 7.

§ 32. I. Für Borelsche Mengen: P. Alexandroff, Compt. rend. 162 (1916), 323; F. Hausdorff, Math. Ann. 77 (1916), 430.
Zur Theorie der Suslinschen Mengen vgl.:
M. Souslin, Compt. rend. 164 (1917), 88.
N. Lusin, Ensembles analytiques (1930).
N. Lusin und W. Sierpiński, Bull. Ac. Crac. (1918), 35; Journ. de Math. (7) 2 (1923), 53.
W. Sierpiński, Bull. Ac. Crac. (1919), 161, 179.

§ 33. I. W. Sierpiński, Compt. rend. 175 (1922), 859.

§ 34, 2. N. Lusin und W. Sierpiński, J. de Math. (7) 2 (1923), 53.

§ 35, 2. III. C. Jordan, Cours d'Analyse I (2. éd. Paris 1893), 53.
V. L. E. J. Brouwer, Amst. Ak. Proc. 12 (1910), 785.

§ 35, 3. S. 199. H. Lebesgue, Leçons sur l'intégration (Paris 1904), 105.

§ 36, 1. Peanosche Kurven: G. Peano, Math. Ann. 36 (1890), 157. D. Hilbert, Math. Ann. 38 (1891), 459. K. Knopp, Archiv d. Math. u. Ph. (3) 26 (1918), 103. H. Lebesgue, Leçons sur l'intégration (Paris 1904), 44.
S. 204. L. E. J. Brouwer, Math. Ann. 70 (1911), 161. Über die Brouwer-Urysohn-Mengersche Dimensionentheorie vgl. K. Menger, Dimensionstheorie (1928).

§ 36, 2. II. W. Sierpiński, Fund. math. 1 (1920), 44.
III. H. Hahn, Jahresb. D. Math. V. 23 (1914), 319; Wiener Ber. 123 (1914), 2433; St. Mazurkiewicz, Compt. r. Varsovie (III) 6 (1913), 305 (polnisch); Fund. math. 1 (1920), 191.

§ 37. II. III. W. Sierpiński, Bull. Ac. Crac. (1918), 29 und (1919), 161; Fund. math. 5 (1924), 155.
IV. M. Souslin, Compt. rend. 164 (1917), 88.

§ 38. I. II. W. Sierpiński, Compt. rend. 171 (1920), 24.
III. St. Mazurkiewicz, Bull. Ac. Crac. (1916), 490. P. Alexandroff, Compt. rend. 178 (1924), 185. F. Hausdorff, Fund. math. 6 (1924), 146.
IV. M. Lavrentieff, Compt. rend. 178 (1924), 187; Fund. math. 6 (1924), 149.

§ 39, 1. II. L. Vietoris, Monatshefte Math. Phys. 31 (1921), 179.
III. N. J. Lennes Amer. J. of Math. 33 (1911), 308; W. Sierpiński, Ann. di Mat. (3) 26 (1916), 131.

§ 39, 2. H. Hahn, Wiener Ber. 130 (1921), 217.
VII. R. L. Moore, Math. Ztschr. 22 (1925), 307.

§ 40. Genaueres A. Rosenthal (Enzykl. II C 9) Nr. 26; H. Tietze und L. Vietoris (Enzykl. III A B 13); M. Fréchet, Les espaces abstraits (1928); C. Kuratowski, Topologie I (1933); W. Sierpiński, General topology (1934).
S. 228. C. Kuratowski, Fund. math. 3 (1922), 182.
S. 229. Trennungsaxiome: H. Tietze, Math. Ann. 88 (1923), 290.
S. 230. P. Urysohn, Math. Ann. 94 (1925), 309. A. Tychonoff, Math. Ann. 95 (1925), 139.
S. 230. M. Fréchet, Rend. Pal. 22 (1906).

§ 41. Diese Theorie stammt hauptsächlich von H. Lebesgue, Journ. de Math.

(6) 1 (1905), 139—216. Vgl. W. H. Young, Lond. math. soc. proc. (2) 12 (1912), 260. F. Hausdorff, Math. Ztschr. 5 (1919), 292.

§ 41, 4. XIII, XIV. H. Hahn, Wien. Ber. 126 (1917), 103. H. Tietze, Journ. f. Math. 145 (1914), 9. F. Hausdorff, Math. Ztschr. 5 (1919), 295.

§ 41, 5. W. Sierpiński, Fund. math. 2 (1921), 15, 37; St. Mazurkiewicz, ib. 28; St. Kempisty, ib. 64, 131.

§ 42. Die Sätze dieses § stammen meistens von R. Baire, Ann. di mat. (3) 3 (1899), 1—122; Leçons sur les fonctions discontinues (Paris 1905).

§ 42, 1. S. 249. R. Baire, Bull. soc. math. 32 (1904), 125; H. Tietze, Journ. f. Math. 145 (1914), 9; F. Hausdorff, Math. Ztschr. 5 (1919), 293.

§ 42, 3. S. 251. W. H. Young, Wien. Ber. 112 (1903), 1307; H. Lebesgue, Bull. soc. math. 32 (1904), 235.

S. 251. H. Hankel, Math. Ann. 20 (1887), 89.

S. 251 (C, D nicht vertauschbar): V. Volterra, Giorn. di mat. 19 (1881), 76.

§ 42, 4. VII. H. Lebesgue in E. Borel, Leçons (1905), 149 und Journ. de math. (6) 1 (1905), 182. C. de la Vallée Poussin, Intégrales, 121.

VIII. R. Baire, Ann. di Mat. (3) 3 (1899), 16, 30.

§ 42, 5. C. Kuratowski u. W. Sierpiński, Fund. math. 3 (1922), 303.

§ 43. Vgl. R. Baire, Ann. di mat. (3) 3 (1899), 68; H. Lebesgue, Journ. de math. (6) 1 (1905).

§ 43, 2. V. H. Lebesgue, Journ. de math. (6) 1 (1905), 205. C. de la Vallée Poussin, Intégrales, 145.

VI. R. Baire, Ann. di mat. (3) 3 (1899), 81.

S. 262. N. Lusin, Fund. math. 2 (1921), 155.

S. 263. W. Sierpiński, Fund. math. 5 (1924), 20.

§ 43, 3. W. Sierpiński, Compt. rend. 170 (1920), 919; Fund. Math. 2 (1921), 74; Fund. Math. 3 (1922), 26; Bull. Ac. Crac. (1919), 161, 179.

§ 44. I. H. Hahn, Archiv Math. Ph. 28 (1919), 34. W. Sierpiński, Fund. math. 2 (1921), 41.

§ 45, 2. I. S. Banach, Fund. Math. 16 (1930), 395.

S. 280. C. Kuratowski, Fund. Math. 16 (1930), 390.

Erst durch diese beiden Arbeiten ist die Bairesche Bedingung klar und einfach geworden. Als Bairesche Bedingung (im engeren Sinn, vgl. § 45, 4) wurden früher Eigenschaften recht verschiedenen, teilweise komplizierten Wortlauts bezeichnet, die vor Kenntnis des Banachschen Satzes nur im separablen Raum gleichwertig waren; über ihren Zusammenhang im beliebigen Raum vgl. G. Steinbach, Beiträge zur Mengenlehre (Diss. Bonn 1930), § 8.

S. 280. O. Nikodym, Fund. Math. 7 (1925), 149.

S. 280. W. Sierpiński, Hypothèse du continu (1934).

§ 45, 3. C. Kuratowski, Fund. Math. 5 (1924), 75. Die dortigen Bezeichnungen fonction α-continue, β-continue habe ich in α-Funktion, β-Funktion abgekürzt und danach die α-Mengen und β-Mengen benannt.

§ 46. N. Lusin, Ensembles analytiques (1930).

§ 46, 1. N. Lusin, Compt. rend. soc. polon. math. (1926), 104.

§ 46, 2. Vgl. auch H. Hahn, Reelle Funktionen (1932), § 42, 4 und § 42, 5.

Nachtrag A. W. Hurewicz, Fund. Math. 12 (1928), 78.

Nachtrag C. C. Kuratowski, Fund. Math. 3 (1922), 200, ib. 10 (1927), 225

Nachtrag D. St. Banach, Fund. Math. 17 (1931), 283.

Register.

(Verweisung auf Seitenzahlen.)