amount of shielding is introduced into the ionic potential. This can be achieved by changing the denominator q^2 in (20) to $q^2+\Delta^2$.

Substituting this new expression for v_q into (21), and using the GAUSSian trial function

$$\psi(r)=(\beta/\pi)^{3/4}\,e^{-\beta r^2/2}\,, \tag{22}$$

we find, after assuming Δ^2/β to be a small quantity and that the main contributions to $E(\beta)$ come from the region of small q,

$$E'_{\min}=\tfrac{1}{2}\pi\,\Delta-\tfrac{1}{8}\pi\,\alpha^2, \tag{23}$$

where

$$E'=\frac{E}{3\,\hbar^2/4\,m\,a^2}\,,\qquad \alpha=\frac{4\,m\,e^2\,Z^2}{3\,g\,\hbar^2\,a^2}\,\frac{m}{M}\,. \tag{24}$$

From this we see that a discontinuity in the self-energy may occur if Δ is sufficiently small.

Our conclusion is tentative, but the analysis shows that self-trapping is critically dependent on the amount of screening involved.

Acknowledgements

We express thanks to the D. S. I. R. (now the S. R. C.) for a maintenance grant to one of us (AEKD) during the tenure of which this work was carried out.

Hauptsätze über das Messen als Grundlage der Hilbert-Raumstruktur der Quantenmechanik

GÜNTHER LUDWIG

Institut für Theoretische Physik der Universität Marburg

(Z. Naturforschg. **22 a**, 1303—1323 [1967]; eingegangen am 24. Mai 1967)

Herrn Professor Dr. PASCUAL JORDAN *zum 65. Geburtstag gewidmet*

Ähnlich wie bei den Hauptsätzen der Thermodynamik werden heuristische Gesichtspunkte diskutiert, die zur Aufstellung von allgemeinen Hauptsätzen über die Wirkung von Mikroobjekten auf makroskopische (d. h. dem thermodynamischen Gesetz der Entropiezunahme genügenden) Systeme führen. Das Ziel ist, durch solche allgemeinen „Hauptsätze des Messens" die Struktur des HILBERT-Raumes der Quantenmechanik zu charakterisieren. Die hier angegebenen Hauptsätze (Axiome 1 bis 3) führen zwar zu mathematischen Strukturen, die sehr viel von den Strukturen des HILBERT-Raumes enthalten, aber noch allgemeiner als der HILBERT-Raum sein können. Vielleicht enthalten sie neue interessante Möglichkeiten für neue physikalische Theorien.

I. Grundbegriffe

Die Quantenmechanik ist durch eine merkwürdige „Grundstruktur" charakterisiert, den HILBERT-Raum. Deshalb ließ es schon gleich zu Beginn vielen Physikern keine Ruhe, ein tieferes Verständnis für diese Grundstruktur zu gewinnen [1]. Trotz mancher Teilerfolge [1, 2] erwiesen sich aber alle bisherigen Versuche als entweder physikalisch nicht ganz durchschaubar oder konnten das Ziel, den HILBERT-Raum zu charakterisieren, nicht voll erreichen. Wir wollen hier versuchen, auf jeden Fall physikalisch interpretierbare Sätze als Ausgangspunkt der Theorie zu benutzen; aber auch wir werden nicht ganz das Endziel, eine *vollständige* Charakterisierung des HILBERT-Raumes, erreichen. Vielleicht aber sollte man hierüber nicht unzufrieden sein, falls man hierbei andere interessante Basistrukturen für physikalische Theo-

[1] W. HEISENBERG, Z. Physik **43**, 172 [1927]. — N. BOHR, Naturwiss. **16**, 245 [1928]. — W. HEISENBERG, Die physikalischen Prinzipien der Quantentheorie, Hirzel, Berlin 1930. — G. BIRKHOFF u. J. v. NEUMANN, Ann. Math. **37**, 823 [1936]. — P. JORDAN, Anschauliche Quantentheorie, Springer, Berlin 1936. — I. E. SEGAL, Ann. Math. **48**, 930 [1947].

[2] D. BOHM, Phys. Rev. **85**, 166 [1952]. — P. JORDAN, Z. Physik **133**, 21 [1952]. — G. W. MACKEY, Amer. Math. Monthly **64**, No. 8 (II) 45 [1957]. — N. ZIERLER, Pac. J. Math. **11**, 1151 [1961]. — G. EMCH u. C. PIRON, J. Math. Phys. **4**, 469 [1963]. — G. EMCH, Helv. Phys. Acta **36**, 739 [1963]. — J M. JAUCH u. C. PIRON, Helv. Phys. Acta **36**, 827 [1963]. — C. PIRON, Helv. Phys. Acta **37**, 439 [1964]. — G. LUDWIG, Z. Physik **181**, 233 [1964]. — M. D. MACLAREN, ANL 7065 Argonne National Laboratory, Argonne, Illinois 1965. — I. AMEMIYA u. H. ARAKI, Research Inst. Math. Sci. RIMS-12 [1966], Kyoto, Japan. — G. LUDWIG, Comm. Math. Phys. **4**, 331 [1967].

rien entdeckt, was aber noch nicht geklärt ist. Natürlich bliebe auch dann noch die Aufgabe bestehen, diese verschiedenen Basistrukturen durch physikalisch interpretierbare Aussagen sauber von einander zu trennen. Nur die Abtrennung des „klassischen" Falles von dem „quantenmechanischen" wird uns auf diese Weise gelingen (Axiom 3 k).

Die Struktur des HILBERT-Raumes unterliegt allen quantenmechanischen Systemen, ob man z. B. Wasserstoff oder Helium oder Eisenatome betrachtet. Die einzelnen Arten von Systemen unterscheiden sich erst durch andere Bestimmungsstücke, wie den HAMILTON-Operator. In dieser Weise hat unser Problem durchaus eine Ähnlichkeit zur Thermodynamik: Die Hauptsätze der Thermodynamik gelten für alle Systeme; erst durch andere Bestimmungsstücke, wie die speziellen Zustandsgleichungen, können verschiedene Arten von Systemen unterschieden werden. Unser hier zu behandelndes Problem ist es also nicht, die Quantenmechanik z. B. des Heliumatoms herzuleiten, sondern die aller Quantenmechanik gemeinsam zugrunde liegende Struktur des HILBERT-Raumes zu verstehen. Sie ist das mathematische Bild sehr allgemeiner physikalischer Strukturen, aber welcher?

Was aber ist die Struktur des HILBERT-Raumes? Die Quantenmechanik beschreibt hierdurch den *Erwartungswert* einer *Observablen* A im Zustand V der untersuchten Systeme:

$$\text{Erwartungswert von } A \text{ im Zustand } V = \mathrm{Sp}(VA). \tag{1}$$

Hierbei ist A ein HERMITEscher Operator, der der Observablen zugeordnet ist, V ein positiv semidefiniter HERMITEscher Operator (mit $\mathrm{Sp}(V) = 1$), der dem Zustand zugeordnet ist. (1) ist die kurze Formulierung der „statistischen Deutung" der Quantentheorie. In (1) gehen drei physikalische Begriffe ein: Erwartungswert, Observable, Zustand, abgebildet auf die mathematischen Symbole Sp (= Spur), A und V.

Wenn wir nun beabsichtigen, die Formel (1) aus sehr allgemeinen „Hauptsätzen über das Messen" herzuleiten, müssen wir uns erst genauer über diese drei Grundbegriffe: Erwartungswert, Observable, Zustand, klar werden. Da diese Begriffe selbst sehr allgemein sind, stehen wir vor einer ähnlich schweren Aufgabe wie etwa der analogen, den Begriff des thermodynamischen (Gleichgewichts-) Zustandes als Grundbegriff der Thermodynamik einzuführen. Mathematisch ist die Aufgabe immer einfach, wie wir sehen werden, da die Bezeichnung irgendwelcher Elemente einer Menge z. B. als „Punkte" in der Geometrie oder als „Zustände" in der Thermodynamik oder als „Zustände" in der Quantenmechanik erst implizit durch die Axiome charakterisiert wird, denen die mathematischen Symbole genügen. Physikalisch aber ist vorher zu sagen, zu beschreiben, wovon wir sprechen, d. h. zu klären, welche *realen* Sachverhalte mit diesen Begriffen gemeint sind.

Als Ausgangspunkt für diese Aufgabe scheint uns (1) mit den zugrundegelegten drei Begriffen nicht besonders geeignet zu sein. Deshalb wollen wir zunächst eine zu (1) äquivalente Darstellung geben, um auf die Begriffe hingeleitet zu werden, die wir dann schließlich als Ausgangspunkt unserer allgemeinen Theorie benutzen wollen.

a) Entscheidungsmessungen

Der Begriff der Observablen ist für unsere Zwecke zu komplex. Beginnen wir zur Illustration wieder mit der Quantenmechanik, wo eine Observable durch einen HERMITEschen Operator A charakterisiert wird. Bekanntlich läßt A die Spektraldarstellung

$$A = \int_{-\infty}^{+\infty} \alpha \, \mathrm{d}E_\alpha \tag{2}$$

zu mit einer Spektralschar von Projektionsoperatoren E_α mit $E_{-\infty} = 0$, $E_{+\infty} = 1$ und $E_{\alpha_1} \geqq E_{\alpha_2}$ für $\alpha_1 \geqq \alpha_2$. Der Projektionsoperator $E_{\alpha_1} - E_{\alpha_2}$ charakterisiert nach der Quantenmechanik folgende Entscheidungsmessung: Fällt der Meßwert α von A in das Intervall $\alpha_2 \ldots . \alpha_1$? Der „Meßwert" von $E_{\alpha_1} - E_{\alpha_2}$ ist gleich 1, wenn der Meßwert von A in das Intervall $\alpha_2 \ldots \alpha_1$ fällt, sonst gleich 0. Der Erwartungswert von $E_{\alpha_1} - E_{\alpha_2}$ ist also die *Wahrscheinlichkeit* dafür, daß der Meßwert α von A in das Intervall $\alpha_2 \ldots \alpha_1$ fällt:

Wahrscheinlichkeit für

$$(E_{\alpha_1} - E_{\alpha_2}) = \mathrm{Sp}(V(E_{\alpha_1} - E_{\alpha_2})).$$

Diesen Gedankengang können wir auch umkehren: Statt von beliebigen „Observablen" gehen wir von „Entscheidungsmessungen" aus. Den Entscheidungsmessungen ordnen wir nach der Quantenmechanik Projektionsoperatoren P zu. Die Wahrscheinlichkeit für P ist nach der – aus (1) folgenden – spezielleren Formel

$$\text{Wahrscheinlichkeit für } P = \mathrm{Sp}(VP) \tag{3}$$

zu berechnen. Soweit wäre der Ausgangspunkt der Entscheidungsmessungen und der Formel (3) eine

Spezialisierung des Begriffes ‚Observable' und der Formel (1). Um von diesem spezialisierten Ausgangspunkt wieder zu (1) zurückzukommen, muß man nun allerdings als neuen Begriff den der „Kommensurabilität" von Entscheidungsmessungen einführen. Entscheidungsmessungen P_i heißen kommensurabel, wenn alle P_i mit Hilfe eines einzigen Apparates zusammen an einem Objekt gemessen werden können. Nach der Quantenmechanik sind die P_i genau dann kommensurabel, wenn sie paarweise kommutieren (d. h. $P_i P_j = P_j P_i$). Eine Observable ist dann nichts anderes als eine durch eine „Skala" α geordnete Schar kommensurabler Entscheidungsmessungen E_α (d. h. $E_{-\infty}=0$, $E_{+\infty}=1$, $E_{\alpha_1} \geqq E_{\alpha_2}$ für $\alpha_1 \geqq \alpha_2$).

$E_{\alpha_1} - E_{\alpha_2}$ ist dann per definitionem eine Entscheidungsmessung, die man durch: „Der Skalenwert von α liegt im Intervall $\alpha_2 \ldots \alpha_1$" charakterisieren kann. Der Erwartungswert des Skalenwertes ist dann durch

$$\text{Erwartungswert von } \alpha = \int_{-\infty}^{+\infty} \alpha \, \mathrm{d}\mu(\alpha) \qquad (4)$$

gegeben mit $\mu(\alpha)$ = Wahrscheinlichkeit für E_α, d. h. dafür, daß der Skalenwert kleiner als α ist. Mit $P = E_\alpha$ folgt dann aus (3) und (4) rückwärts sofort (1), wenn man A nach (2) *definiert.* Wir sehen also, daß sich der Begriff der „Observablen" in einer physikalisch *sehr sinnvollen* Weise einführen läßt, wenn man die Begriffe „Entscheidungsmessung" und „kommensurabel" hat.

Statt des Begriffes *Entscheidungsmessung* wurden häufig auch andere Worte, wie *Eigenschaften, Ja-Nein-Messungen, Aussagen* (propositions)[1–3] (um nur die häufigsten zu nennen) benutzt. Alle diese Worte enthalten irgendeine Vorstellung über das, was eben in der Quantenmechanik durch die Projektionsoperatoren gekennzeichnet wird. Diese Vorstellungen reichen von dem Wort ‚Aussagen' (wobei die Projektionsoperatoren die Elemente eines logischen Aussagekalküls werden) über das Wort ‚Eigenschaften' (wobei die Projektionsoperatoren Eigenschaften der Objekte symbolisieren) bis zum Wort ‚Entscheidungs*messung*' (wobei die Projektionsoperatoren die Anwendung eines Meßapparates auf die Objekte charakterisieren). Es würde hier zu weit führen, diese Auffassungen einer jeweiligen Kritik zu unterwerfen. Kurz und prägnant soll daher zunächst nur unsere Stellungnahme zu diesen Auffassungen formuliert werden. Erst nachträglich und rückwärts wollen wir am Schluß dieser Arbeit einige kritische Bemerkungen zu den anderen Auffassungen geben.

1) Die Auffassung der Projektionsoperatoren als Elemente eines echten Aussagekalküls lehnen wir grundsätzlich ab. Dementsprechend wird eine Änderung der Logik nicht in Erwägung gezogen. Die Logik der Aussagen, auf der die Physik basiert, bleibt weiterhin die zweiwertige Logik. Einer *nachträglichen* Bezeichnung der Projektionsoperatoren als „symbolische Aussagen" steht natürlich nichts im Wege, da die *Bedeutung* einer symbolischen Aussage erst auf übliche bedeutungsvolle Aussagen im Rahmen einer zweiwertigen Logik zurückgeführt werden muß. Sogenannte Aussagen innerhalb einer „mehrwertigen" Logik sind bedeutungsleer, solange sie eben nicht auf andere schon in ihrem Aussageinhalt bekannte Aussagen zurückgeführt werden, ein Verfahren, das eben als *Basis* eine zweiwertige Logik voraussetzt.

2) Die Auffassung der Projektionsoperatoren als Eigenschaften von Objekten lehnen wir weder ab, noch aber nehmen wir sie zum Ausgangspunkt unserer Darstellung. Erst *nachträglich* stellen wir die Frage, ob auf Grund der physikalischen Strukturen den Objekten wie Atomen etwas zugeschrieben werden kann ähnlich den unmittelbar feststellbaren Eigenschaften makroskopischer Objekte. Anders ausgedrückt: vor Anwendung des Wortes ‚Eigenschaft' wird erst eine genauere Definition dieses Wortes versucht, um dann zu sehen, ob und inwieweit es anwendbar ist. Die Benutzung des Wortes Eigenschaft gleich zu Beginn der Theorie kann zu leicht dazu führen, mit diesem Wort unbewußt Voraussetzungen einzuführen, deren Gültigkeit entweder fraglich ist oder aber erst hergeleitet werden sollte.

3) Als Basis bleibt also nur die dem realen Umgehen der Physik mit Mikroobjekten entsprechende Tatsache des „Messens" der Objekte mit einem Apparat. Zwar hat man hiermit die Kompliziertheit des Messens von vornherein mit in die Theorie hineingenommen, aber eben *noch nicht vorausgesetzt,* daß dieses Messen ganz elementar eliminierbar wäre, indem man eben nur von den gemessenen „Eigenschaften" der Objekte oder gar nur von den auf Grund einer Messung zu machenden „Aussagen" über die Objekte zu sprechen hätte. Die Tatsache (bewiesen in vielen Arbeiten über den Meßprozeß in

[3] C. F. v. Weizsäcker, Komplementarität und Logik, I: Naturwiss. **42**, 521, 545 [1955]; II, III: Z. Naturforschg. **13 a**, 245, 705 [1958].

der Quantenmechanik [4]), daß eine Meßanordnung nur bei raffinierter Konstruktion wirklich ein P (P Projektionsoperator) mißt, hat wahrscheinlich bisher davor zurückschrecken lassen, Strukturgesetze über das Messen als Ausgangspunkt einer Theorie zu wählen; denn wie soll man „von Anfang an" die richtigen Messungen (die eben ein P messen) von *unzureichenden Meßversuchen* unterscheiden? Dieser Einwand besteht zu recht; aber er besagt noch lange nicht, daß das Messen als Ausgangspunkt einer Theorie unbrauchbar ist, sondern vielmehr nur, daß das, was bisher als Entscheidungsmessung bezeichnet wurde, noch nicht allgemein genug ist, um als Basis für eine physikalische Theorie zu dienen. Warum also nur diejenigen Wechselwirkungsanordnungen zwischen Objekt und Aparatur betrachten, die zu den idealen Entscheidungsmessungen führen? Wenn wir zu Anfang *alle* möglichen Apparaturen – und mögen sie noch so schlecht zum Messen (im idealen Sinn) geeignet sein – zulassen, so haben wir also nichts vorweggenommen. Die *wirklichen* Aparate der Experimentalphysiker sind meist auch alles andere als die „idealen Meßapparate" der Theoretiker. Daher enthält ein solcher allgemeinerer Ausgangspunkt nicht nur weniger Voraussetzungen, sondern ist zu allem noch realistischer. Zwar würde ein Theoretiker das Wort „Messen" gerne nur auf die idealen Messungen eingeschränkt sehen; da uns aber kein besseres Wort für die Tätigkeit der Experimentatoren eingefallen ist, behalten wir das Wort *messen* auch für die allgemeinsten Wechselwirkungsanordnungen zwischen Objekten und beliebigen Apparaturen mit an den Apparaten registrierbaren Ergebnissen der Wechselwirkung bei.

b) Effekte

Für diese eben angegebene allgemeinste Wechselwirkungsanordnung mit registrierbaren Ergebnissen wollen wir als kurzes Wort *Effekt* einführen. Der Effekt ist also eine Verallgemeinerung der Entscheidungsmessung, etwas ungenau ausgedrückt, eine *schlechte* Entscheidungsmessung. Wenn wir jetzt versuchen, den Effekt als Grundbegriff unserer Theorie einzuführen, stehen wir vor einer ähnlichen Schwierigkeit wie bei der Einführung des Begriffes ‚Gleichgewichtszustand' als Grundlagenbegriff der Thermodynamik. Da wir Effekt als Grundbegriff einführen, kann er also nicht durch andere definiert werden. Da er sehr allgemein anwendbar sein soll, kann er nicht durch eine scharfe und enge physikalische Beschreibung erfaßt werden. Wir können daher nur an Beispielen klarzustellen versuchen, was wir damit meinen und wo der Begriff anwendbar ist, ohne aber damit endgültig Fälle ausschließen zu können, bei denen seine Anwendbarkeit fraglich wird. Dasselbe gilt ja auch bei dem Begriff des Gleichgewichtszustandes der Thermodynamik.

Mit dem Wort ‚Effekt' bezeichnen wir erstens die Apparatur, z. B. eine Nebelkammer, und zweitens einen bestimmten, an dieser Apparatur auftretenden und registrierbaren Effekt (im engeren Sinne), wie z. B. ob ein Nebeltropfen in einem bestimmten Raumgebiet der Nebelkammer aufgetreten ist. Mit Apparatur ist ein technisch herstellbarer Gegenstand gemeint, aber natürlich auch (als primitivstes technisches Verfahren!) ein in der Natur einfach vorgefundener Gegenstand. Entscheidend ist, daß dieser Gegenstand als objektiv realer Gegenstand vorweisbar ist wie ein Weinglas auf dem Tisch. Atome sind in dieser Weise nicht vorweisbar, da man erst andere Apparate braucht, um ihre Wirkungen vorweisen zu können. Typisch für Apparaturen ist es, daß sie auch dem Entropiesatz genügen, während sich für einzelne Atome im allgemeinen keine Entropie definieren läßt. Unter einer Apparatur wollen wir aber nicht nur ihre technische Konstruktion verstehen, sondern auch die Zeitspanne miteinbeziehen, in der sie in bezug auf die zu untersuchenden Objekte eingesetzt wird. Diesen letzten Punkt können wir erst etwas genauer schildern, wenn wir davon sprechen, was wir unter einer Gesamtheit von Objekten verstehen; denn die Zeitspanne von der wir sprechen, läßt sich nur sinnvoll relativ zur Gesamtheit der Objekte einführen, die mit dieser Apparatur in Wechselwirkung treten. Nur wollten wir gleich darauf hinweisen, daß Zählrohre gleicher Konstruktion, aber zu verschiedenen Zeiten geöffnet, als verschiedene Apparaturen zählen sollen.

Unter Effekt (im engeren Sinne) verstehen wir eine Veränderung an der eben geschilderten Apparatur, die nur auftreten kann (aber nicht muß), wenn irgendwelche Einwirkungen auf die Apparatur vorliegen, d. h. es ist dem Experimentalphysiker mög-

[4] J. v. NEUMANN, Mathem. Grundlagen der Quantenmechanik, Springer-Verlag, Berlin 1932 (insbes. Abschn. VI, Nr. 3). — G. LUDWIG, Gelöste und ungelöste Probleme des Meßprozesses in der Quantenmechanik, in „W. Heisenberg und die Physik unserer Zeit", Vieweg & Sohn, Braunschweig 1961 (insbes. Abschn. V, e). — E. P. WIGNER, Z. Physik **133**, 101 [1952]. — Vgl. auch die in M. M. YANASE, Amer. J. Phys. **32**, 208 [1964] zitierten Arbeiten.

lich, die Apparatur so zu isolieren, daß der Effekt *nicht* auftritt. Das Auftreten der Veränderung muß objektiv vorweisbar, registrierbar sein. Das Auftreten des Effektes (im engeren Sinne) an der Apparatur wird im folgenden (wenn keine Mißverständnisse möglich sind) kurz auch als Effekt bezeichnet.

An ein und derselben Apparatur können bei ein und demselben Experiment mehrere Effekte (im engeren Sinne) auftreten; z. B. können bei einem System von Zählrohren mehrere Zählrohre ansprechen. Effekte, die an ein und derselben Apparatur bei ein und demselben Experiment auftreten können, nennen wir kurz *koexistent* (den Begriff *kommensurabel* wollen wir für einen anderen Zusammenhang aufheben). Jedes Experiment beruht nach unserer Meinung in der Registrierung von Effekten, d. h., unsere Begriffsbildung des *Effektes* und *koexistenter* Effekte ist breit genug um als Basis einer physikalischen Theorie zu dienen.

Die Bedeutung und die Aussagen der später einzuführenden Hauptsätze über das Messen werden wir an Gedankenexperimenten (ähnlich wie in der Thermodynamik) erläutern. Dazu wollen wir hier noch kurz einige spezielle Möglichkeiten von Effekten schildern. Wir werden etwas später sehen, daß es darauf ankommt, bei einer großen Zahl von Experimenten das Auftreten eines Effektes F zu zählen, d. h. die Häufigkeit von F bei einem Experiment zu bestimmen. Nun wendet der Experimentator zu diesem Zweck häufig Apparaturen an, die dieses Zählen selbst übernehmen oder diese Häufigkeit gleich „abzulesen" gestatten. Betrachten wir z. B. den Effekt, daß ein Photon in einer Apparatur ein Korn einer Photoplatte (nach der Entwicklung der Platte) *schwärzt.* Als Effekt können wir z. B. betrachten, daß ein Korn in einem bestimmten Gebiet der Platte geschwärzt wird. Statt diese Effekte zu zählen, kann man auch einfach summarisch den Schwärzungsgrad der Platte untersuchen, natürlich unter der Voraussetzung, daß diese Schwärzung nicht so hoch ist, daß später eintreffende Photonen nicht mehr praktisch dieselbe Chance zum Schwärzen eines Kornes hatten wie die ersten. In dem so geschilderten Fall vermittelt also der Schwärzungsgrad direkt die gesuchte Häufigkeit des Effektes F.

Ähnlich wirkt beispielsweise ein Filter. Als Effekt kann man die Absorption eines Lichtquants im Filter betrachten. Aber statt von jedem Lichtquant einzeln zu registrieren, ob es absorbiert wurde oder nicht, kann man auch einfach den Absorptionskoeffizient summarisch messen. Dieses letzte Beispiel des Filters wird uns als anschauliches und lehrreiches Beispiel dienen, um die Bedeutung der aufzustellenden Hauptsätze zu demonstrieren.

Ebenso wie in der Thermodynamik mathematisch die Zustände einfach als Elemente einer Menge eingeführt werden, führen wir die Effekte F als Elemente einer Menge $\mathbf{L}$ ein [5].

Definition 1: $\mathbf{L}$ ist die Menge aller Effekte F.

Die mathematische Formulierung des Begriffes „koexistent" geschieht an späterer Stelle.

Zur Veranschaulichung des Begriffes Effekt soll die Frage beantwortet werden, wie ein Effekt in der bekannten Quantenmechanik zu beschreiben ist. Dazu hat man ein Mikroobjekt mit einer Apparatur zu koppeln und die Wahrscheinlichkeit für das Auftreten eines Effektes (im engeren Sinn) an der Apparatur auf Grund der Wechselwirkung mit dem Mikroobjekt zu berechnen. Das Ergebnis [6] ist, daß man diese Wahrscheinlichkeit *immer* (wie auch die Apparatur konstruiert sein mag und welchen Effekt man an der Apparatur auch feststellen mag) *allein* mit Hilfe mathematischer Größen im HILBERT-*Raum des Mikroobjektes* ausdrücken kann in der Form $\mathrm{Sp}(VF)$, wobei V den Zustand (die Statistik) der Mikroobjekte charakterisiert und F ein HERMITEscher Operator im HILBERT-Raum der Mikroobjekte ist, dessen Spektrum zwischen 0 und 1 liegt. F hängt natürlich von der Apparatur und dem an der Apparatur beobachteten Effekt (im engeren Sinn) ab. Wenn F nur die Eigenwerte 0 und 1 hat, d. h., wenn F gerade ein Projektionsoperator ist, dann liegt genau der Fall vor, daß man von einer Entscheidungsmessung im zu Anfang geschilderten Sinn sprechen kann. So kommt man bei der Diskussion des Meßprozesses in der Quantenmechanik in ganz natürlicher Weise zu der Menge der Effekte (statt nur zu der kleineren Menge der Entscheidungsmessungen), wobei Effekte durch HERMITEsche Operatoren F mit $0 \leqq F \leqq 1$ zu charakterisieren sind. Um Konfusion

[5] G. LUDWIG, a) Z. Physik **181**, 233 [1964]; b) Comm. Math. Phys. **4**, 331 [1967]; c) „Versuch einer axiomatischen Grundlegung der Quantenmechanik und allgemeinerer physikalischer Theorien I, II", Preprints Marburg 1964 und 1966.

[6] K.-E. HELLWIG, Makroskopische Effekte und quantenmechanische Messung, Preprint Marburg 1967.

zu vermeiden, sei ausdrücklich darauf hingewiesen, daß F hierbei keine *Observable* symbolisiert, genausowenig wie der HERMITEsche Operator V, der den Zustand bestimmt.

c) Gesamtheiten

Wenn wir eben von den Effekten sprachen, so war auch schon davon die Rede, daß diese durch Einwirkungen auf die Apparatur hervorgerufen werden. Wir müssen jetzt also noch einen Grundbegriff für diese Einwirkungsmöglichkeiten prägen. In der Quantenmechanik ist es der durch einen positiv-semidefiniten HERMITEschen Operator V [mit $\mathrm{Sp}(V) = 1$] charakterisierte „Zustand der Systeme", der für das Hervorrufen der Effekte verantwortlich ist. Was mit „Zustand der Systeme" bezeichnet wird, wollen wir versuchen, etwas näher zu schildern, und als zweiten Grundbegriff unserer Theorie einführen. Für dieselbe Größe V sind auch noch andere Worte in Gebrauch. Entsprechend der Deutung der Projektionsoperatoren als Aussagen wird V als Symbol für die Kenntnisse angesehen, auf Grund derer ein Subjekt Aussagen einen bestimmten Wahrscheinlichkeitswert zumißt. Entsprechend unserer grundsätzlichen Ablehnung der Deutung der Projektionsoperatoren als Aussagen lehnen wir auch die Deutung von V als Symbol für die Kenntnisse ab. Das viel benutzte Wort „Zustand" für V ist uns ebenso wie „Eigenschaft" für die Projektionsoperatoren schon beladen mit einer vorweggenommenen Vorstellung, als ob jedes Objekt etwas zu eigen hat, etwas besitzt (nämlich seinen Zustand), was durch V charakterisiert wird. Uns scheint deshalb das Wort ‚Zustand' nicht besonders geeignet, als *Ausgangspunkt* einer Theorie zu dienen. Diese Auffassung steht natürlich nicht dem entgegen, *nachträglich* etwas zu definieren, was in gewisser Weise den Objekten zugeschrieben und deshalb besser mit ‚Zustand' bezeichnet werden kann.

Wir wollen hier für die (in der Quantenmechanik durch V) zu beschreibende Situation in der Natur das Wort *Gesamtheit* wählen. Es ist auch nur ein Wort, dem wir erst durch Beschreibung einen Inhalt geben müssen. Im Wort ‚Gesamtheit' fassen wir (wieder wie oben schon beim Wort ‚Effekt') zwei Dinge zusammen: 1) eine Versuchsanordnung zum Produzieren von Wirkungen und 2) eine häufige Wiederholung, eine häufige Anwendung dieser Versuchsanordnung.

Um nun näher zu beschreiben, was wir unter solchen Versuchsanordnungen verstehen, seien zunächst einige Beispiele angegeben. Eine solche Anordnung ist z. B. ein Stück Uran bestimmter Größe. Dieses kann aber auch in einer Bleikammer mit einem zeitweise geöffneten Fenster sitzen. Die eine bestimmte Zeitspanne eingeschaltete Glühbirne ist eine andere solche Anordnung; aber auch ein Zyklotron oder Synchrotron. Wichtig hierbei ist nur, daß ein technisches Verfahren angebbar ist, um eine solche Apparatur als wohl-definierbares Objekt herzustellen, wobei man natürlich auch vorgegebene Dinge aus der Natur wählen kann. Unbrauchbar sind solche Stücke in der Natur, bei denen sich nicht festlegen läßt, was zu dem gewählten Objekt nun wirklich dazugehört und was nicht.

Wir sehen jetzt im Zusammenhang mit dem obigen Begriff ‚Effekt', daß wir mit dem Begriffspaar Gesamtheit – Effekt jede Situation meinen, wo ein wohlgekennzeichneter Gegenstand an einem anderen eine Wirkung (Effekt) hervorbringt (oder nicht hervorbringt). Entscheidend dabei ist, daß beide Gegenstände nach einem definierbaren Verfahren hergestellt werden können und damit der Versuch sehr oft wiederholt werden kann. Wir haben jetzt eben allerdings bei der Kennzeichnung des Zusammenwirkens der beiden Apparaturen (z. B. eines Stückes Uran und einer Nebelkammer) einen zusätzlichen noch nicht erläuterten Begriff benutzt, nämlich der *Wirkung* der einen Apparatur auf die andere. Das taten wir, um die Situation erst einmal zu veranschaulichen. Bei der Einführung des dritten Grundbegriffes der Wahrscheinlichkeit werden wir dann näher darauf eingehen, was wir unter dem Wirkungszusammenhang der beiden Apparaturen verstehen wollen.

Da wir gleich die Wiederholbarkeit der Versuche, d. h. der Anwendung der Wirkungen produzierender Apparatur mit in den Grundbegriff aufnehmen wollen, haben wir das Wort *Gesamtheit* gewählt; nicht nur eine konkrete, einmalige Versuchsanordnung wird betrachtet, sondern die Gesamtheit der Wiederholungen mit einer solche Wirkungen produzierenden Versuchsanordnung bestimmter Art.

Mathematisch werden solche Gesamtheiten wieder einfach als Elemente einer Menge eingeführt [5].

Definition 2: **K** ist die Menge aller Gesamtheiten V.

Die Situation (V, F_1) bedeutet also wiederholte Versuche, wobei eine durch V charakterisierte Sorte von Apparaturen an einer anderen den Effekt F_1 hervorbringt; (V, F_2) bedeutet, daß einerseits dieselbe Sorte V benutzt wird, aber daß V auf eine andersartige Apparatur einwirkt und daher einen anderen Effekt F_2 hervorbringen kann. Die Zerlegbarkeit der experimentellen Situation in die zwei Teile V und F ist also charakteristisch für die von uns zu beschreibende Grundsituation; überall wo dies möglich ist und auch andere Kombinationen desselben V mit anderen F und desselben F mit anderen V zustande gebracht werden können, ist die hier darzulegende Theorie anwendbar.

Zur Veranschaulichung von Gedankenexperimenten ist es nützlich, besonders einfache Beispiele für V zu wählen. Wie bei der Absorption von *Filtern* als Beispiel für F kann man sehr einfach experimentieren, wenn man auch für die Wiederholung der Experimente die Apparatur V nicht dauernd neu „ein- und ausschaltet", sonderen „laufen" läßt, natürlich unter der Voraussetzung, daß während der Arbeitszeit der Apparatur V keine merkliche Veränderung an ihr auftritt. Wenn man also z. B. mit „Wasserstoffatomen" experimentieren will, wird man sich eine Apparatur V herstellen, die solche Wasserstoffatome eins nach dem anderen liefert. Statt aber nun jedesmal, wenn ein Wasserstoffatom von V emittiert wurde, die Apparatur abzuschalten und dann den Versuch in dieser Weise zu wiederholen, läßt man V *laufend* ein Wasserstoffatom nach dem anderen emittieren (natürlich nur unter der Bedingung, daß die Wasserstoffatome so *getrennt* von einander emittiert werden, daß man sie auch hätte einzeln zählen können und sie sich nicht gegenseitig beeinflussen; und der weiteren Bedingung, daß die Emissionsstärke der Apparatur sich nicht zeitlich merklich ändert). Statt N Versuche einzeln durchzuführen, mißt man im zweiten Falle einfach die Emissionsintensität der Apparatur V und vergleicht damit die Absorption von Wasserstoffatomen auf einem Auffänger einer Effektapparatur, die z. B. aus einem Target und einer Auffängerfläche besteht. So können wir uns später in Gedanken als V eine *Emissionsapparatur* und für F eine *Absorptionsapparatur* (Filter) vorstellen.

In dem eben geschilderten Beispiel war von der Emission von „Wasserstoffatomen" die Rede. Dies war nur zur Veranschaulichung erwähnt worden; wie man das bezeichnet, *was* V emittiert, ist nicht von vornherein klar, es sei denn, daß man es eben *nur* durch die Emissionsapparatur, d. h. mit V, bezeichnet. Nachträglich kann man aber durchaus auf Grund der von V erzeugten F andere Gesichtspunkte zur Bezeichnung der Emissionsprodukte von V angeben (Kapitel VIII).

Um sich nicht von vornherein etwas zu Spezielles unter V vorzustellen, wollen wir quantenmechanisch den Fall betrachten, wo V zum Bruchteil α Wasserstoffatome und zum Bruchteil $(1-\alpha)$ Heliumatome emittiert. Ein solcher Fall wird meist in der Theorie nicht ausführlich behandelt, da er sich sehr einfach auf die beiden Fälle *nur* Wasserstoffatome und nur Heliumatome reduzieren läßt. Da wir hier aber die Absicht haben, die Grundlagen einer Theorie zu entwickeln, können wir diesen Fall nicht von vornherein ausschließen. Wir wollen daher seine quantenmechanische Beschreibungsweise angeben:

Wir haben zwei Hilbert-Räume $\mathfrak{H}_1$ und $\mathfrak{H}_2$ ($\mathfrak{H}_1$ für die Wasserstoffatome, $\mathfrak{H}_2$ für die Heliumatome) zu benutzen. Die Entscheidungsmessungen sind durch ein Paar von Projektionsoperatoren P_1 (aus $\mathfrak{H}_1$), P_2 (aus $\mathfrak{H}_2$) zu charakterisieren, die Gesamtheit V durch ein Paar von positiv-semidefiniten Hermiteschen Operatoren V_1 (aus $\mathfrak{H}_1$) und V_2 (aus $\mathfrak{H}_2$) mit $\mathrm{Sp}(V_1)=1$ und $\mathrm{Sp}(V_2)=1$. Die Wahrscheinlichkeit für die Entscheidungsmessung (P_1, P_2) ist dann gleich

$$\alpha\,\mathrm{Sp}(V_1 P_1) + (1-\alpha)\,\mathrm{Sp}(V_2 P_2).$$

Zum Beispiel ist die Entscheidungsmessung „Wasserstoffatom" (und nicht Helium) durch $P_1=1$ und $P_2=0$ charakterisiert. Der eben geschilderte Fall ist auch unter dem Schlagwort „Superauswahlregel" bekannt: Zwischen Wasserstoffatomen und Heliumatomen gilt eben eine Superauswahlregel.

d) *Wahrscheinlichkeit*

Unter dem Wort ‚Wahrscheinlichkeit', das auch wir als Bezeichnung eines dritten Grundbegriffes unserer Theorie benutzen wollen, verstehen durchaus nicht alle dasselbe. Da wir grundsätzlich die Projektionsoperatoren P als *Aussagen*, die statistischen Operatoren V als *Kenntnis* abgelehnt haben, weisen wir auch die Wahrscheinlichkeit als subjektives Maß für die Richtigkeit einer Aussage als für die Physik unbrauchbar zurück. Wahrscheinlichkeit kann im Rahmen unserer Theorie hier nur einen objektiven Sachverhalt zwischen V und F charakterisieren, den man kurz mit dem Satz: „V ruft F mit der Wahrscheinlichkeit μ hervor" auszudrücken pflegt. Genau diesem Satz müssen wir einen physikalisch nachprüfbaren Inhalt geben.

Genauso wie wir vorher die Deutung der Projektionsoperatoren P als Eigenschaften, des statistischen Operators V als Zustand deshalb nicht zum Ausgangspunkt einer Theorie wählen wollten, weil uns darin schon gewisse Momente vorweggenommen scheinen, genauso wollen wir nicht *von vornherein* die Wahrscheinlichkeit als ein dem Einzelsystem (Einzelversuch) innewohnende Struktur auffassen, sondern ganz entsprechend dem experimentellen Vorgehen als Häufigkeit bei vielen Versuchen, als Häufigkeit des Auftretens des Effektes F in der Gesamtheit V. Führen wir also N Versuche durch und tritt N_+-mal dabei der Effekt F auf, so soll sich also N_+/N für große Werte von N als reproduzierbar erweisen, d. h.: Wird nach demselben Verfahren die Versuchsreihe (V, F) wiederholt, indem man z. B. statt N einmal N' Versuche durchführt, wobei N_+'-mal der Effekt F auftritt, so soll sich annähernd $N_+/N \approx N_+'/N'$ erweisen, wenn nur N und N' groß genug sind. Diese Reproduzierbarkeit des Häufigkeitsverhältnisses ist durchaus nicht selbstverständlich und gar nicht immer realisiert. Daß wir so wenig von den nicht reproduzierbaren Fällen sprechen, liegt einzig und allein an unserer Übung, solche Fälle von vornherein als uninteressant auszuschließen; *unvergleichlich* viel mehr unbrauchbare Versuche würden in der Experimentalphysik durchgeführt werden, wenn man nicht schon von vornherein etwa wüßte, wie ein nützlicher Versuch aufzubauen ist. Von unserem Standpunkt der Grundlegung einer Theorie müssen wir eben doch auch an die Menge der unnützen Experimente denken, um eben dann formulieren zu können, was ein nützlicher Versuch ist.

Um zu zeigen, was wir meinen, nehmen wir ein extremes Beispiel: Ein Physiker hat in einem Labor ein Zählrohr. F ist der Effekt, daß dieses Zählrohr anspricht. In einem anderen Gebäude weit davon entfernt hat er ein Stück Uran. Dieses Stück Uran ist dann V. Mißt er jetzt die Häufigkeit von F, so kann es ihm passieren, daß (z. B. in einer anderen Umgebung als in seinem ersten Labor) die Häufigkeit des Effektes F nicht reproduzierbar ist. Er sagt dann, die gemessenen Effekte waren gar nicht von V, sondern von irgendwelchen anderen Umständen aus der Umgebung hervorgerufen worden. Jeder Experimentalphysiker weiß, welche Mühe es macht, alle „Dreckeffekte" zu eliminieren, d. h. eben die gewünschte reproduzierbare Häufigkeit für (V, F) zu bekommen.

Ist es dem Experimentator gelungen, eine reproduzierbare Häufigkeit für F bei V herzustellen, so hat er einen nützlichen Versuch durchgeführt. Die Tatsache, daß sich eine reproduzierbare Häufigkeit ergeben hat, ist für uns das Zeichen, daß ein struktureller Zusammenhang zwischen V und F besteht, der die Ursache für diese Reproduzierbarkeit der Häufigkeiten ist. Genau diesen strukturellen Zusammenhang meinen wir, wenn wir von der Wahrscheinlichkeit $\mu(V, F)$ sprechen, wobei μ approximativ durch die Häufigkeit N_+/N „gemessen" wird. Ist $\mu(V, F) \neq 0$, so drücken wir den Zusammenhang noch prägnanter durch Hinzufügen von Wörtern wie „hervorrufen", „bewirken" aus: V ruft F mit der Wahrscheinlichkeit μ hervor. Ist $\mu(V, F) = 0$, so sagen wir, daß V den Effekt F nicht hervorruft.

Die Einführung des physikalischen Begriffes Wahrscheinlichkeit hat also methodisch eine gewisse Ähnlichkeit mit der Einführung des Begriffes der Temperatur in der Thermodynamik. Beide Begriffe setzen gewisse Strukturen in der Natur voraus. Zur Einführung des Begriffes der Temperatur benutzt man folgende *Experimentiermöglichkeiten:* Gewisse Paare von Zuständen Z_1 und Z_2 von zwei Systemen können in thermischen Kontakt gebracht werden, ohne daß an ihnen Veränderungen auftreten; diese Relation zwischen Z_1 und Z_2 stellt sich *experimentell* als Äquivalenzrelation heraus. Zur Einführung des Begriffes der Wahrscheinlichkeit benutzt man die Experimentiermöglichkeiten: Ein Versuch V, F läßt sich so „isolieren", daß man reproduzierbare Häufigkeiten bei sehr häufiger Wiederholung des Versuches erhält.

Gegen die hier erläuterte Art und Weise des Einführens des physikalischen Begriffes der Wahrscheinlichkeit wird häufig eingewandt, daß ein Vergleich der Häufigkeit N_+/N mit der Wahrscheinlichkeit nur mit „Wahrscheinlichkeit" möglich ist. Leider ist eine ausführliche Diskussion dieses Einwandes aus Platzgründen hier nicht möglich. Es seien nur kurz einige Hinweise gegeben: Zunächst ist es durchaus nicht klar, was man bei diesem Einwand mit dem Wort „Wahrscheinlichkeit" meint. Die Benutzung desselben Wortes wie bei der durch Häufigkeiten N_+/N zu testenden Wahrscheinlichkeit μ besagt durchaus nicht, daß dasselbe gemeint ist. Für kleinere Werte von N kann man nun durchaus N-Tupel von Versuchen als *einen* Versuch ansehen und diese „N-Tupel-Versuche" *sehr häufig* wiederholen, um eine Häufigkeitsverteilung von N_+/N aufzunehmen; dann sind wir aber immer noch bei dem physikalischen Begriff der Wahrscheinlichkeit, wie wir ihn oben eingeführt haben. Es zeigt sich nun, daß die Häufigkeitsverteilung der N_+/N sehr gut durch

$$w\left(\frac{N_+}{N}\right) = \binom{N}{N_+} \cdot \mu^{N_+} \tag{5}$$

wiedergegeben werden kann. *Extrapoliert* man diese Formel auch für große N (wo sie sich vielleicht gar *nicht mehr experimentell* prüfen läßt!), so kann man sie trotzdem benutzen, um sich *subjektiv* zu entschließen, N_+/N mit einer gewissen Genauigkeit als Test für μ anzuerkennen, auch *ohne* die annähernde Reproduzierbarkeit von N_+/N experimentell nachzuprüfen. Aber sogar dann, wenn sich die Formel (5) *experimentell* für größere Werte von N als *falsch* erweisen würde, z. B. in der Art, daß für noch so große Werte von N immer eine endliche Schwankung von etwa 1% für die Häufigkeiten N_+/N übrig bliebe, wäre der von uns eingeführte physikalische Begriff der Wahrscheinlichkeit nicht unbrauchbar. Er wäre dann vielmehr (was er vermutlich auch tatsächlich ist, so wie die meisten physikalischen Begriffe) ein nur approximativer Begriff, mit einem in seiner Größe von etwa 1% bekannten Fehler behaftet, eben ein Begriff, der die Wirklichkeit nur in einer gewissen Näherung darzustellen gestattet.

Geht man von der Tatsache aus, daß sich in der Natur durch die reproduzierbare Häufigkeit eine Struktur kundtut, dann kann man natürlich (aber in einem nicht mehr rein physikalischen Schritt) diese Struktur als eine jedem Einzelversuch innewohnende, d. h. eine dem Einzelversuch zukommende Wahrscheinlichkeit deuten. Es könnte aber auch eine Struktur sein, die in den äußeren Umständen liegt, auf Grund derer die Einzelversuche wiederholt werden können. Um aber eben alle solche Fragen nicht von vornherein zu entscheiden, haben wir die von uns angegebene Deutung des Wahrscheinlichkeitsbegriffes als Grundbegriff unserer Theorie gewählt.

Die mathematische Erfassung des geschilderten Wahrscheinlichkeitsbegriffes geschieht, wie in jeder physikalischen Theorie, durch Idealisierungen. Als erste Idealisierung sehen wir also von der (oben geschilderten) Möglichkeit eines endlichen Fehlers ab, d. h. einer Versuchsreihe V, F wird genau ein (und nur ein) reeller Zahlenwert μ als Wahrscheinlichkeit zugeordnet. Da wir keinen experimentellen Grund haben, bestimmte Versuchsreihen auszuschließen, wird also *jeder* Versuchsreihe ein μ zugeordnet [5]:

Axiom 1: Auf der Menge $\mathbf{K} \times \mathbf{L}$ (Menge aller Paare (V, F) mit $V \in \mathbf{K}$, $F \in \mathbf{L}$) ist eine reellwertige Funktion $\mu(V, F)$ definiert mit
- α) $0 \leqq \mu(V, F) \leqq 1$.
- β) Zu jedem $V \in \mathbf{K}$ gibt es ein $F \in \mathbf{L}$ mit $\mu(V, F) = 1$.
- γ) Es gibt ein $F \in \mathbf{L}$, das mit 0 bezeichnet wird, mit $\mu(V, 0) = 0$ für alle $V \in \mathbf{K}$.

α) gibt wieder, daß μ mit Hilfe von N_+/N getestet wird und $0 \leqq N_+/N \leqq 1$ per definitionem gilt. β) drückt aus, daß man die Zahl der Versuche zählen kann, d. h. für das Zählen der Versuche ist $N_+ = N$. γ) ist nur formal; man kann es auch als die Versuchsreihe deuten, wo mit V kein Effektapparat gekoppelt wird.

II. Klassen von Effekten und Gesamtheiten

Durch $V_1 \sim V_2$ für $[\mu(V_1, F) = \mu(V_2, F)$ für alle $F \in \mathbf{L}]$ wird eine Äquivalenzrelation in $\mathbf{K}$ definiert. Die Menge der Äquivalenzklassen bezeichnen wir mit $\boldsymbol{K}$. Ebenso kann man durch $F_1 \sim F_2$ für $[\mu(V, F_1) = \mu(V, F_2)$ für alle $V \in \mathbf{K}]$ eine Äquivalenzrelation in $\mathbf{L}$ definieren. Die Menge dieser Äquivalenzklassen bezeichnen wir mit $\boldsymbol{L}$. Die Elemente von $\boldsymbol{K}$ nennen wir Gesamtheiten (zweiter Art), die Elemente von $\boldsymbol{L}$ Effekte (zweiter Art). Auf Grund von $\mu(V, F)$ mit $V \in \mathbf{K}$, $F \in \mathbf{L}$ läßt sich sofort eine reellwertige Funktion $\mu(V, F)$ mit $V \in \boldsymbol{K}$ und $F \in \boldsymbol{L}$ definieren durch $\mu(V, F) = \mu(V_1, F_1)$, wobei V_1 aus der Klasse V und F_1 aus der Klasse F ist. Aus Axiom 1 α) β) γ) folgt sofort:

Satz 1: Auf $\boldsymbol{K} \times \boldsymbol{L}$ ist eine reellwertige Funktion $\mu(V, F)$ $(V \in \boldsymbol{K}, F \in \boldsymbol{L})$ definiert, für die gilt:
- α) $0 \leqq \mu(V, F) \leqq 1$.
- β) Für jedes $V \in \boldsymbol{K}$ gibt es ein $F \in \boldsymbol{L}$ mit $\mu(V, F) = 1$.
- γ) Es gibt in $\boldsymbol{L}$ ein $F(=0)$ mit $\mu(V, 0) = 0$.
- δ) Aus $\mu(V_1, F) = \mu(V_2, F)$ für alle $F \in \boldsymbol{L}$ folgt $V_1 = V_2$.
- ε) Aus $\mu(V, F_1) = \mu(V, F_2)$ für alle $V \in \boldsymbol{K}$ folgt $F_1 = F_2$.

δ) und ε) gelten, weil wir eben zu den Mengen der Klassen übergegangen sind. Dieser Übergang zu den Klassen wäre nur eine mathematische Spielerei, wenn nicht (was gleich vorweggenommen sei) alle weiteren Axiome der Theorie immer nur die Elemente von $\boldsymbol{K}$ und $\boldsymbol{L}$ enthalten. Diesen Sachverhalt wollen wir in seiner physikalischen Bedeutung etwas näher untersuchen.

Sicher ist diese Einteilung in Klassen ohne jede Bedeutung, wenn jede Klasse aus nur einem einzigen

Element besteht. Die Erfahrung zeigt, daß das gerade nicht der Fall ist. Beginnen wir mit den verschiedenen Gesamtheiten $V_i \in \mathbf{K}$ aus einer Klasse $V \in \boldsymbol{K}$. Wir haben also verschiedene technische Konstruktionsvorschriften V_i, die aber alle zu „denselben" Wirkungen führen, denselben, d. h., alle Effekte werden von den V_i mit denselben Häufigkeiten $[\mu(V_i, F)$ unabhängig von $i]$ erzeugt. Diese Tatsache drücken wir dadurch aus, daß wir sagen: Die Wirkungen der zwar *verschiedenen* Apparate V_i werden durch *dieselben* Objekte (dieselben Wirkungsträger) ausgeübt, wobei wir uns die Wirkung in *einem* Experiment von der V_1-Apparatur auf die F-Apparatur durch *einen* Wirkungsträger vermittelt denken. In diesem Sinne sprechen wir dann bei vielen Versuchen statt von der Gesamtheit der V_i von der Gesamtheit der Wirkungsträger. So können wir also sagen: Die verschiedenen V_i einer Klasse V erzeugen dieselbe Gesamtheit von Wirkungsträgern, eben die Gesamtheit (zweiter Art) V. Da nun die einzelnen V_i niemals mehr weiter in Erscheinung treten, spricht man dann einfach nur noch von der Gesamtheit (zweiter Art) V von Wirkungsträgern, wobei man dann noch die Worte „von zweiter Art" fortläßt und statt Wirkungsträger auch *Mikroobjekt* sagt: Die $V \in \boldsymbol{K}$ charakterisieren also herstellbare Gesamtheiten von Mikroobjekten. Die Bezeichnung ‚Mikroobjekt' soll dabei einerseits durch „Mikro" den Unterschied zu der technisch baubaren Apparatur $V_i \in \mathbf{K}$ charakterisieren und durch „-objekt" auf die eben geschilderte *Unabhängigkeit* (= Objektivität) der Wirkungsträger von der speziellen Form V_i der Apparatur hinweisen.

Auch bei den Effekten zeigt die Erfahrung, daß viele $F_i \in \mathbf{L}$ zu einer Klasse $F \in \boldsymbol{L}$ gehören. In diesem Fall ist es sogar sehr simpel, verschiedene F_i einer Klasse F zu konstruieren: Man ändere nur die technische Art der Fixierung des Effektes; z. B. nehme man ein Zählrohr, an das man *verschiedene* Registrierapparate für das „Ansprechen" des Zählrohres anschließt. Hierbei bezeichnet man schon von vornherein auf Grund unseres Verständnisses der Funktionsweise des Zählrohres den Effekt mit dem Wort „ansprechen", obwohl die eigentlichen Effekte je nach den benutzten Registrierapparaten ganz verschieden aussehen können. Aber die verschiedenen F_i einer Klasse F können sich auch grundsätzlicher als in diesem Beispiel unterscheiden; aber auch dann faßt schon der Experimentalphysiker solche verschiedenen Apparate unter einem Wort, wie z. B. Zähler von geladenen Teilchen, oder Neutronenzähler usw., zusammen. Genau das ist gemeint, wenn wir die F_i zu Klassen zusammenfassen.

Es zeigt sich, daß die F_i einer Klasse auch in bezug auf weitere Strukturbeziehungen wie die später zu besprechende Koexistenz nicht unterscheidbar sind. Daher spricht man kurz von den $F \in \boldsymbol{L}$ auch als Effekten (zweiter Art), wobei man dann schließlich den Zusatz „zweiter Art" noch fortläßt. Diese eben geschilderte gewisse Unabhängigkeit der Wirkungen $F \in \boldsymbol{L}$ von der speziellen Apparatur F_i kann es nahelegen, F nicht nur als Effekt, sondern versuchsweise als „Eigenschaft" der Mikroobjekte zu bezeichnen. Wenn man nur ein neues Wort wählt und weiß, was man mit dem neuen Wort bezeichnet, wäre das ein durchaus erlaubtes Verfahren. Wir wollen aber das Wort ‚Eigenschaft' für etwas anderes aufheben. Daher behalten wir das Wort ‚Effekt' (auch ohne den eigentlich notwendigen Zusatz „zweiter Art") bei, weil Mißverständnisse im folgenden nicht zu befürchten sind, da fast nur noch von Gesamtheiten und Effekten (beide von zweiter Art) die Rede sein wird. Mathematisch kann man also die ganze Theorie statt mit Axiom 1 $\alpha)\beta)\gamma)$ auch mit Satz 1 als erstem Axiom beginnen, so daß die Mengen $\mathbf{L}$ und $\mathbf{K}$ nicht in Erscheinung getreten wären. Uns kam es aber hier darauf an, etwas näher die physikalische Bedeutung zu diskutieren.

Da wir zur Illustration als Beispiel die Quantenmechanik wählen, sei darauf hingewiesen, daß man $\boldsymbol{K}$ (und nicht $\mathbf{K}$) mit einer Teilmenge der Menge aller positiv semidefiniten Hermiteschen Operatoren V mit $\mathrm{Sp}(V) = 1$ zu identifizieren hat und $\boldsymbol{L}$ (nicht $\mathbf{L}$) mit einer Teilmenge der Menge aller positiv semidefiniten Hermiteschen Operatoren F mit $F \leqq 1$. $\mu(V, F) = \mathrm{Sp}(V, F)$ erfüllt dann Satz 1 α bis ε, wenn z. B. in $\boldsymbol{L}$ der 1- und der Null-Operator liegen und $\boldsymbol{K}$ und $\boldsymbol{L}$ so viele V bzw. F enthalten, daß aus $(\varphi_1, F\varphi_1) = (\varphi_2, F\varphi_2)$ für alle $F \in \boldsymbol{L}$ $\varphi_1 = \varphi_2$ und aus $(\varphi_1, V\varphi_1) = (\varphi_2, V\varphi_2)$ für alle $V \in \boldsymbol{K}$ ebenfalls $\varphi_1 = \varphi_2$ folgt. Diese Voraussetzungen wollen wir für $\boldsymbol{K}$ bzw. $\boldsymbol{L}$ als erfüllt ansehen.

Alle Physik hat einen gewissen „Endlichkeitscharakter". Damit soll folgendes gemeint sein: Durch endlich viele Experimente und endlich viele Tatsachen läßt sich die Struktur eines Bereiches der Wirklichkeit (wenigstens approximativ) erkennen. Eine mathematische Menge der Mächtigkeit des Kontinuums ohne topologische Charakterisierung (d. h. mit der diskreten Topologie, bei der jede Teilmenge eine offene Menge ist) ist aber nicht von dieser Art. Dagegen kann ein endliches Gebiet eines metrischen dreidimensionalen Raumes approximativ (eben in bezug auf die Metrik) durch endlich viele Punkte charakterisiert werden. Wenn wir also durch unsere Axiome physikalische Strukturen darstellen wollen, müssen wir noch diesen durch endlich viele Schritte approximierbaren Charakter mathematisch formulieren. Deswegen führen wir erst einige Definitionen ein:

Definition 3: $\mathbf{B}$ sei die Menge aller Funktionen $X(F)$ auf $\mathbf{L}$ mit

$$X(F) = \sum_{i=1}^{n} a_i\, \mu(V_i, F)\,; \qquad V_i \in \mathbf{K}$$

a_i reelle Zahlen, n beliebige ganze Zahlen.

$\mathbf{B}$ ist damit ein linearer reller Raum. Durch $V \longleftrightarrow X(F) = \mu(V, F)$ kann man die Menge $\mathbf{K}$ mit einer Teilmenge von $\mathbf{B}$ identifizieren. Durch

$$\| X \| = \sup_{F \in \mathbf{L}} | X(F) |$$

ist in $\mathbf{B}$ eine Norm definiert. Durch Abschließen in bezug auf diese Norm erhält man aus $\mathbf{B}$ einen Banach-Raum B. $\mathbf{K}$ ist dann auch Teilmenge von B.

Definition 4: K sei die abgeschlossene konvexe Hülle von $\mathbf{K}$ in B.

Den Elementen von K kann man eine gewisse physikalische Bedeutung geben. Jedes Element von K ist approximierbar durch eine endliche Summe:

$$\sum_{i=1}^{m} \lambda_i V_i \quad \text{mit} \sum_{i=1}^{m} \lambda_i = 1, \qquad \lambda_i > 0\,.$$

Mit einer *großen* Zahl N und Zahlen N_i lassen sich die λ_i durch N_i/N approximieren, so daß jedes $V \in K$ durch eine Summe

$$\sum_{i=1}^{m} (N_i/N)\, V_i \quad \text{mit} \quad \sum_{i=1}^{m} N_i = N$$

approximierbar ist. $\sum_i (N_i/N)\, V_i$ kann man aber als die Gesamtheit deuten, wo N Versuche so durchgeführt werden, daß je N_i der Sorte V_i ausgeführt werden. Deshalb bezeichnet man $\sum_{i=1}^{m} \lambda_i V_i$ als Gemisch der V_i mit den Gewichten λ_i.

Die Funktion $\mu(V, F)$ kann mit Hilfe von Definition 3 $\mu(X, F) = X(F)$ und der Norm für alle $X \in B$ definiert werden. Wenn wir ein $V \in \mathbf{K}$ haben und es ein $F_1 \neq 0$ gibt mit $\mu(V, F_1) = 0$, so kann man nach Satz 1 ein V_1 wählen mit $\mu(V_1, F_1) \neq 0$; für $\frac{1}{2} V + \frac{1}{2} V_1$ ist dann die Menge der F mit $\mu(\frac{1}{2} V + \frac{1}{2} V_1, F) = 0$ kleiner als für V. Unser „Endlichkeitspostulat“ soll besagen, daß man mit endlich vielen Schritten des Mischens für „fast alle“ $F \neq 0$ auch $\mu(V, F) \neq 0$ erreichen kann, was so gemeint sei, daß im Limes der Normtopologie $\mu(V, F) \neq 0$ für alle $F \neq 0$ erreichbar sei:

Axiom 1 η: Es gibt ein $V \in K$ (also nicht unbedingt $V \in \mathbf{K}$!) mit $\mu(V, F) \neq 0$ für alle $F \neq 0$.

Da es uns bisher nicht gelungen ist, mathematisch im selben Umfang die Struktur unserer Theorie zu klären, wie in dem Sonderfall, wo Axiom 1 η durch ein schärferes ersetzt wurde, wollen wir hier dieses schärfere anführen, um später darauf hinweisen zu können, welche Ergebnisse nur im Falle des schärferen Axioms bewiesen werden konnten. Das schärfere Axiom lautet:

Axiom 1 η *(endlich):* Der Banach-Raum B ist endlichdimensional.

Allgemein ist zu B der duale Banach-Raum B' aller normstetigen linearen Funktionale über B definiert. Durch $\mu(X, F)$ kann man kanonisch $\mathbf{L}$ als Teilmenge von B' auffassen. Neben der Normtopologie des Banach-Raumes B' führt man in bekannter Weise in B' noch die schwache Topologie ein (die Menge der schwachstetigen Linearformen über B' ist mit B identisch). Für später geben wir noch folgende Definition an:

Definition 5: L ist die schwach abgeschlossene Hülle von $\mathbf{L}$ in B', $\hat{L}$ die schwach abgeschlossene konvexe Hülle von $\mathbf{L}$.

Im Falle der Quantenmechanik ist K gleich der Menge *aller* positiv semidefiniten Hermiteschen Operatoren V mit $\mathrm{Sp}(V) = 1$. B ist die Menge aller Hermiteschen Operatoren der sogenannten „Trace-Klasse“: Jedes $W \in B$ hat die Form $W = \alpha V_1 - \beta V_2$ mit $\alpha \geqq 0$, $\beta \geqq 0$ und $V_1, V_2 \in K$. B' ist die Menge aller (beschränkten) Hermiteschen Operatoren, $\hat{L}$ die Menge aller positiv semidefiniten Hermiteschen Operatoren F mit $F \leqq 1$.

III. Hauptsatz über die Empfindlichkeitssteigerung von Effekten

Die bisherigen Untersuchungen führten nur zur Definition einer Reihe physikalischer Begriffe. Um aber Bedingungen für diese Begriffe zu finden, brauchen wir ähnlich wie in der Thermodynamik *Hauptsätze.* In diesem III. Kapitel und im V. sollen zwei solche Hauptsätze besprochen werden.

Diese Hauptsätze können als sehr allgemeine Aussagen über Strukturen in der Natur nicht bewiesen, sondern nur in ihrer Bedeutung an Beispielen demonstriert werden. Die physikalische Rechtfertigung dieser Hauptsätze besteht darin, daß man sie erstens an Beispielen eventuell direkt bestätigen kann und daß sich zweitens die aus ihnen gezogenen Folgerungen als in der Natur erfüllt herausstellen. Wir wollen in diesem Kapitel für den ersten der beiden Hauptsätze in Gedankenexperimenten seine Bedeutung aufzeigen. Ein diesem Hauptsatz widersprechendes Experiment ist nicht bekannt.

Während in der Thermodynamik die Hauptsätze Aussagen über die *Unmöglichkeit* der Konstruierbarkeit gewisser Apparate machen, machen die Hauptsätze des Messens umgekehrt Aussagen über Konstruktionsmöglichkeiten von Apparaten. Zur Formulierung des ersten Hauptsatzes definieren wir einen sehr anschaulichen Begriff:

Definition 7: Wirs chreiben für zwei Elemente Y_1 und Y_2 aus B': $Y_1 \leqq Y_2$, wenn $\mu(V, Y_1) \leqq (V, Y_2)$ für alle $V \in \mathbf{K}$ (und damit auch für alle $V \in K$) ist. Ist für zwei Effekte F_1 und F_2 (d. h. $F_1 \in \mathbf{L}$, $F_2 \in \mathbf{L}$) $F_1 \leqq F_2$, so sagen wir, das F_2 *empfindlicher* oder *stärker* als F_1 ist, F_1 unempfindlicher oder schwächer als F_2.

Nun der erste Hauptsatz über das Messen [5]:

Axiom 2: Zu jedem Paar F_1, $F_2 \in L$ gibt es ein $F_3 \in L$, so daß $F_1 \leqq F_3$, $F_2 \leqq F_3$ und für jedes $V \in K$ mit $\mu(V, F_1) = 0$ und $\mu(V, F_2) = 0$ auch $\mu(V, F_3) = 0$ ist.

In der Formulierung des Axioms 2 wird die ganze Menge L (Def. 5) und nicht nur $\mathbf{L}$ benutzt. Die Forderung des Axioms 2 für $\mathbf{L}$ wäre unnötig und unphysikalisch scharf, wie aus den späteren Ausführungen hervorgehen wird. Durch die Einführung von L läßt sich eben Axiom 2 leichter und mathematisch streng formulieren. Um aber den physikalischen Inhalt von Axiom 2 zu veranschaulichen, müssen wir natürlich auf die Menge $\mathbf{L}$ der Effekte zurückgehen.

Die Elemente von L, die nicht zu $\mathbf{L}$ gehören, sollen kurz ideale Effekte genannt werden. Ein idealer Effekt ist aber physikalisch approximierbar durch einen realen Effekt aus $\mathbf{L}$. Um dem Wort „physikalisch approximierbar“ einen Sinn zu geben, wurde die schwache Topologie in B' eingeführt. Diese Topologie besagt, daß es zu einem idealen Effekt F und endlich vielen (es gibt immer nur endlich viele Apparate und Experimente!) Gesamtheiten $V_1, \ldots, V_n$ und einer Zahl $\varepsilon > 0$ immer einen realen Effekt F' gibt mit $|\mu(V_i, F) - \mu(V_i, F')| < \varepsilon$. „Es gibt einen realen Effekt F'“ heißt aber genau, es läßt sich eine reale Effektapparatur bauen. Natürlich kann ε physikalisch nie „beliebig klein“ gewählt werden. Da aber eine endliche untere Grenze für den „Fehler“ ε physikalisch unbekannt ist, ist die mathematische Definition der schwachen Topologie gerade die vernünftigste Idealisierung des Begriffes einer physikalischen Approximierbarkeit.

Auch für die idealen Effekte werden wir die in Definition 6 erklärten Worte: empfindlicher, stärker (bzw. unempfindlicher, schwächer) benutzen. Der erste Hauptsatz besagt also nach Axiom 2 etwas über die Möglichkeiten der Empfindlichkeitssteigerung von idealen Effekten und damit eben auch etwas über die Konstruktionsmöglichkeiten empfindlicherer Effekt-Apparaturen. Nun sieht man auch sofort, wie schwierig eine Formulierung von Axiom 2 würde, wollte man sich unmittelbar nur auf die realen Effekte beziehen. Zum Beispiel wäre eine Gleichung $\mu(V, F_1) = 0$ durch eine approximative: $\mu(V, F_1)$ ungefähr, d. h. praktisch physikalisch so gut wie Null, zu ersetzen, wobei dann auch V nur aus $\mathbf{K}$ (und nicht aus K) zu sein braucht. Auch die Ungleichung $F_1 \leqq F_3$ müßte man durch eine approximative ersetzen. Aber gerade *nur* in diesem physikalisch approximativen (und nicht exakt erfüllbaren!) Sinn stellt Axiom 2 eine Forderung an die Konstruktionsmöglichkeit realer Effekte dar; und nur so approximativ scheint uns dieser erste Hauptsatz physikalisch sinnvoll zu sein. In diesem Sinne werden auch die folgenden Beispiele den Inhalt von Axiom 2 erläutern.

Seien F_1 und F_2 zwei reale Zählrohre mit allen ihren experimentellen Mängeln (d. h., sie brauchen nicht mit 100%iger Sicherheit anzusprechen). Schaltet man die beiden Zählrohre so zusammen, daß man einen registrierten Effekt erhält, sobald wenigstens eines von beiden anspricht, so hat man schon die Apparatur F_3. Hier war es also besonders leicht, F_3 zu konstruieren; dies ist immer so einfach, wenn F_1 und F_2 wie in diesem Beispiel *koexistent* sind; bei der Definition koexistenter Effekte wird gerade von dieser einfachen Konstruktionsmöglichkeit Gebrauch gemacht (siehe Kapitel VI).

Seien F_1 und F_2 zwei Filter. Wir wollen annehmen, daß sie so konstruiert sind, daß zwei Filter derselben Sorte F_1 (bzw. F_2) hintereinander geschaltet praktisch auch alles *durchlassen*, was ein Filter F_1 (bzw. F_2) allein durchläßt. Das, was ein Filter F *durchläßt*, sind gerade diejenigen V mit $\mu(V, F) \approx 0$. Wir suchen nun ein Filter F_3, das stärker als F_1 und stärker als F_2 absorbiert, aber alles das praktisch durchläßt, was sowohl F_1 als auch F_2 durchläßt. Schalten wir nun direkt das Filter F_2 *hinter* F_1, so läßt dieses Filterpaar $F_1 F_2$ alles durch, was sowohl F_1 als auch F_2 durchläßt, und es absorbiert sicher stärker als F_1, aber durchaus nicht immer (!) stärker als F_2. Als Beispiel dafür, daß tatsächlich das Filterpaar $F_1 F_2$ nicht immer stärker absorbiert als F_2, betrachte man zwei Polarisationsfilter für Licht, die zueinander schief eingestellt sind. Licht, das von F_2 vollständig absorbiert wird, wird nicht mehr vollständig vom Filterpaar $F_1 F_2$ absorbiert! Nun zeigt aber die Erfahrung, daß man sich durch wiederholtes Hintereinanderschalten von vielen Paaren $F_1 F_2$, d. h. durch einen Filterblock $F_1 F_2 F_1 F_2 F_1 E_2$ ... tatsächlich allmählich approximativ einem Filter F_3 nähert, das sowohl stärker als F_1 als auch stärker als F_2 absorbiert, aber praktisch noch alles durchläßt, was sowohl F_1 als auch F_2 durchläßt.

Dieses Filterbeispiel ist auch sehr geeignet, die Charakterisierung des „klassischen Falles" zu verdeutlichen. Ohne jetzt schon auf die Namengebung „klassisch" einzugehen, wollen wir diesen Fall zunächst durch eine Verschärfung des Axioms 2 festlegen [5].

Axiom 2k: Die Menge aller F_3, die die Bedingungen von Axiom 2 (bei festem F_1, F_2) erfüllen, hat eine untere Grenze F_3', d. h. es gibt ein solches F_3', das die Bedingungen von Axiom 2 erfüllt, so daß für alle anderen F_3, die auch die Bedingungen von Axiom 2 erfüllen, $F_3' \leqq F_3$ gilt.

Betrachten wir diesen Fall in unserem Filterbeispiel. Als Minimalfilter in einer Menge M von Filtern wollen wir ein solches Filter bezeichnen, zu dem es in M kein schwächer absorbierendes Filter gibt. Es kann also in M *mehrere* Minimalfilter geben! Ist das Axiom 2k erfüllt, so gibt es zu jedem Paar F_1, F_2 in der Menge M der Filter F_3, die stärker als F_1 und F_2 absorbieren und alles durchlassen, was sowohl F_1 als auch F_2 durchlassen, genau ein und nur ein Minimalfilter. Alle anderen Filter aus M absorbieren dann stärker als dieses eine Minimalfilter. Ist Axiom 2k nicht erfüllt, so gibt es ein Paar F_1, F_2, so daß es in der eben definierten Menge mindestens zwei Minimalfilter F_3' und F_3'' gibt, für die dann also weder $F_3' \leqq F_3''$ noch $F_3'' \leqq F_3'$ gelten kann, die also nicht „vergleichbar" sind.

Die Bedeutung des Axioms 2k werden wir erst später (Kapitel VIII) diskutieren können, weil wir dazu noch die Begriffe ‚koexistent' und ‚kommensurabel' haben müssen.

Da wir hier nicht in mathematische Einzelheiten einsteigen können, sei nur kurz dargestellt, daß der quantenmechanische Fall die hier vorliegenden Bedingungen erfüllt: Sind F_1 und F_2 zwei Effekte (d. h. HERMITEsche Operatoren mit $0 \leqq F \leqq 1$), P_1 bzw. P_2 die Projektoren auf die Eigenräume zum Eigenwert 0 von F_1 bzw. F_2, P der Projektor auf den Durchschnitt der zu P_1 und P_2 gehörigen Teilräume des HILBERT-Raumes, so sind die V mit $\mu(V, F_1) = \mu(V, F_2) = 0$ durch PVP charakterisiert. Es gibt dann immer einen HERMITEschen Operator F_3 mit $0 \leqq F_3 \leqq 1$, $F_1 \leqq F_3$, $F_2 \leqq F_3$ und mit P als Projektor auf den Eigenraum zum Eigenwert 0 von F_3. Ein solches F_3 ist z. B. $1 - P$.

IV. Entscheidungseffekte

Der in Axiom 2 formulierte Hauptsatz über das Messen hat zur Folge [5], daß es ein Analogon zu den sogenannten Ja-Nein-Messungen der Quantenmechanik auch in dem von uns allgemein betrachteten Fall gibt. Dazu betrachten wir irgendeine Menge k von Gesamtheiten (d. h. k ist eine Teilmenge von K). Die Menge aller idealen Effekte F aus L mit $\mu(V, F) = 0$ für $V \in k$ nennen wir kurz $L_-(k)$. Man kann nun zeigen, daß $L_-(k)$ ein Supremum hat, d. h., es gibt einen (idealen) Effekt $E \in L_-(k)$, so daß $F \leqq E$ für alle $F \in L_-(k)$ gilt. Diese so (für alle möglichen Teilmengen k von K) eingeführten Effekte nennen wir *Entscheidungseffekte.* Die Entscheidungseffekte sind ideale Effekte, soweit sie nicht **L** angehören.

Die physikalische Bedeutung können wir uns am Beispiel von Filtern wieder klar machen. $L_-(k)$ ist die Menge aller Filter, die eine gewisse vorgegebene Sorte von Mikroobjekten hindurchläßt, nämlich alle Gesamtheiten V aus einer vorgegebenen Menge k. Alle anderen Mikroobjekte sollen so stark als möglich absorbiert werden. Der erste Hauptsatz über das Messen hat zur Folge, daß es genau ein und nur ein solches stärkstes Filter E gibt. E ist also eine optimale Apparatur, die alle Mikroobjekte so weit als möglich anzeigt unter der Nebenbedingung, daß sie irgendeine durch k vorgegebene Sorte von Mikroobjekten nicht anzeigen darf. E zählt also so weit als irgend möglich alles, was nicht zur Sorte k gehört.

Oben haben wir das Beispiel zweier Filter F_1 und F_2 untersucht und fragten nach einem F_3, das Axiom 2 erfüllt. Betrachten wir als k die Menge aller V mit $\mu(V, F_1) = \mu(V, F_2) = 0$, so gibt es also nicht nur ein F_3 mit $\mu(V, F_3) = 0$ für $V \in k$ und $F_1 \leqq F_3$, $F_2 \leqq F_3$, sondern sogar ein Entscheidungsfilter E mit $\mu(V, E) = 0$ für $V \in k$ und $F_1 \leqq F_3 \leqq E$, $F_2 \leqq F_3 \leqq E$. Bei der obigen Diskussion betrachteten wir Filterpakete der Form $F_1 F_2 F_1 F_2 F_1 F_2 \ldots$. Die Erfahrung zeigt nun, daß ein solches Filterpaket aus vielen F_1, F_2 gerade das eben angegebene E physikalisch approximiert. Insbesondere erhält man durch ein Filterpaket $F F F \ldots$ aus einem Filter F ein Entscheidungsfilter $E \geqq F$ mit $\mu(V, E) = 0$ für alle $V \in K$ mit $\mu(V, F) = 0$; eine dem Experimentalphysiker durchaus geläufige Methode, um die „Flankensteilheit" von Filtern zu verbessern.

Es läge nahe, das Verhalten der Entscheidungseffekte so zu interpretieren, daß die Mikroobjekte, die eine ganz bestimmte „Eigenschaft" haben, den Entscheidungseffekt E hervorrufen, die anderen nicht. Dann könnte man die Entscheidungseffekte den Eigenschaften der Mikroobjekte eindeutig zuordnen. Da wir aber vermeiden wollen, Vorwegdeutungen mit in die Theorie aufzunehmen, benutzen wir nicht das Wort Eigenschaften, zumal wir es an einer späteren Stelle in einem anderen Sinn als hier einführen wollen.

Sei G die Menge aller Entscheidungseffekte. G ist also eine Teilmenge von L. In bezug auf die in Definition 6 angegebene Ordnungsrelation erweist sich G als *vollständiger Verband* [5]. Die Verbandsstruktur von G hat also nichts mit Logik zu tun, sondern ist eine Folge von Axiom 2. Die Verbandsoperationen sind dabei definiert durch: $E_1 \cup E_2$ als Supremum von E_1 und E_2 in G, $E_1 \cap E_2$ als Infinum von E_1 und E_2 in G. Es gilt für die Gesamtheiten $V \in K$, daß aus $\mu(V, E_1) = 0$ und $\mu(V, E_2) = 0$ auch immer $\mu(V, E_1 \cup E_2) = 0$ folgt. Das Supremum aller Elemente von G wird mit 1 bezeichnet. Es ist $\mu(V, 1) = 1$.

In der Quantenmechanik ist G die Menge aller Projektionsoperatoren E, von denen man weiß, daß sie einen vollständigen Verband bilden. $E_1 \cap E_2$ ist dabei der Projektionsoperator auf den Durchschnitt der zu E_1 und E_2 gehörigen Teilräume. $E_1 \cup E_2$ ist der Projektionsoperator auf den von den beiden zu E_1 und E_2 gehörigen Teilräumen aufgespannten Teilraum.

V. Hauptsatz über Zerlegbarkeit und Verwandtschaft von Gesamtheiten

Als extremale Teilmenge k von K bezeichnen wir eine solche, für die gilt:

1) Aus $V_1 \in k$, $V_2 \in k$ folgt $\lambda V_1 + (1-\lambda) V_1 \in k$ für jedes λ mit $0 \leqq \lambda \leqq 1$.
2) Aus $V \in k$ und $V = \lambda V_1 + (1-\lambda) V_2$ für ein λ mit $0 < \lambda < 1$ folgt $V_1 \in k$ und $V_2 \in k$.
3) k ist abgeschlossen im Banach-Raum B.

Der Durchschnitt beliebig vieler extremaler Teilmengen ist wieder eine solche. Zu jedem $V \in K$ gibt es also eine kleinste extremale Teilmenge, die V enthält; sie sei mit $C(V)$ bezeichnet. $C(V)$ hat einen sehr anschaulichen physikalischen Sinn: V kann als Gemisch von (näherungsweise) jedem Element von $C(V)$ mit einem passenden anderen Element von $C(V)$ angesehen werden. Jedes andere Element von K, das nicht zu $C(V)$ gehört, kann nicht als Mischungskomponente von V auftreten. Wir sagen für diesen Sachverhalt kurz: $C(V)$ ist die Menge aller Komponenten von V, in die V zerlegt werden kann.

Zwei Gesamtheiten V_1 und V_2 sollen *verwandt* heißen, wenn $L_-(V_1) = L_-(V_2)$. Hierbei war, wie oben definiert, $L_-(V)$ die Menge aller F aus L mit $\mu(V, F) = 0$. Zwei Gesamtheiten sind also verwandt, wenn die Menge aller Effekte, die sie nicht hervorrufen können, gleich ist. Auch diese Beziehung hat eine sehr anschauliche physikalische Bedeutung. Auf Grund von Satz 1 (Seite 1311) ist ein $V \in K$ eindeutig durch alle Wahrscheinlichkeiten $\mu(V, F)$ für die verschiedenen Effekte F bestimmt. Die Menge $L_-(V)$ aller F mit $\mu(V, F) = 0$ stellt die Menge aller Effekte dar, die die Mikroobjekte von V nicht hervorrufen können, weil die Gesamtheit V so beschaffen ist, daß (was natürlich eine Interpretation ist) die zugehörigen Mikroobjekte nicht in eine Wechselwirkung oder zumindest geeignete Wechselwirkung mit den Effekt-Apparaturen aus $L_-(V)$ treten können. Zwei V sind also verwandt, wenn ihre Mikroobjekte genau den gleichen Einschränkungen der Wechselwirkungsmöglichkeiten unterliegen. Da $L_-(V) = L_-(C(V))$ ist, sind also zwei Gesamtheiten, die dieselben Komponenten $C(V_1) = C(V_2)$ enthalten, immer auch verwandt, was anschaulich nichts anderes sagt, als daß eben die Komponenten von V ebenfalls aus Mikroobjekten bestehen, die keine geeignete Wechselwirkung mit den F aus $L_-(V)$ hervorbringen können. Wenn zwei verwandte V_1, V_2 nicht in ihren Komponenten übereinstimmen, so könnte man also eine Gesamtheit V' von Mikroobjekten herstellen, die man als Teil, als Komponente von z. B. V_1 ansehen kann, die aber nicht als Komponente von V_2 denkbar ist. Sollten diese Mikroobjekte von V' nicht aber dann auch einige der Apparate aus

$L_-(V_2)$ zum Ansprechen bringen können? Der zweite Hauptsatz über das Messen sagt nun aus, daß diese Frage zu bejahen ist. *Nur* solche V_1, V_2 können verwandt sein, die sich in ihren Komponenten nicht unterscheiden.

Axiom 3: Sind zwei Gesamtheiten $V_1 \in K$, $V_2 \in K$ verwandt, so ist auch $C(V_1) = C(V_2)$.

Um uns diese Aussage in ihrer physikalischen Bedeutung noch näherzubringen, denken wir uns zwei verwandte Gesamtheiten V_1 und V_2 gegeben. In $L_-(V_1) = L_-(V_2)$ gibt es als Supremum einen Entscheidungseffekt E. Sprechen wir von E in der Weise eines Filters: E läßt also V_1 und V_2 durch, aber jedes Filter, was stärker als E absorbiert, läßt weder V_1 noch V_2 vollständig durch. Neben wir nun einmal $C(V_1) \neq C(V_2)$ an. Welche physikalische experimentell nachprüfbare Konsequenz würde das haben?

Aus $L_-(V_1) = L_-(V_2)$ folgt auch $L_-(V_1) = L_-(\frac{1}{2}V_1 + \frac{1}{2}V_2)$, d. h., V_1, V_2 und auch $V = \frac{1}{2}V_1 + \frac{1}{2}V_2$ sind verwandt. Wegen $V_1 \in C(V)$ und $V_2 \in C(V)$, ist also $C(V) \supseteq C(V_1)$ und $C(V) \supseteq C(V_2)$. Da $C(V_1) \neq C(V_2)$, kann nicht $C(V) = C(V_1)$ und $C(V) = C(V_2)$ sein. Nehmen wir $C(V) \neq C(V_1)$ an. Wir haben also zwei verwandte Gesamtheiten V, V_1 mit $C(V) \supset C(V_1)$. Das Filter E ist genau dadurch festgelegt, daß es alle in $C(V_1)$ liegenden Gesamtheiten durchlassen muß, aber sonst so stark als möglich absorbiert. Trotzdem gäbe es aber in $C(V)$ liegende Gesamtheiten, die nicht als Komponenten von Gesamtheiten aus $C(V_1)$ in Frage kommen, und trotzdem ebenfalls durchgelassen werden. Die Durchlässigkeit des Filters E wäre also größer als es *unbedingt* auf Grund der Forderung $\mu(V_1, E) = 0$ nötig wäre, in dem Sinne *unbedingt,* als mit V_1 auch alle möglichen Komponenten von V_1 unbedingt durchgelassen werden müssen. Der zweite in Axiom 3 formulierte Hauptsatz sagt also aus, daß ein Entscheidungsfilter E, das unter der Nebenbedingung $\mu(V, E) = 0$ so stark als möglich absorbiert, dann nur die Sorten V' von Mikroobjekten durchläßt, die man sich als Komponenten in V denken kann.

Dieser zweite Hauptsatz ist für die Quantenmechanik erfüllt. $C(V)$ ist hier die Menge aller V' mit $V'\varphi = 0$ für alle Eigenvektoren φ von V zum Eigenwert 0. In diesem Fall ist $L_-(V)$ die Menge aller F mit $F\psi = 0$ für alle ψ, die zum Eigenraum von V zum Eigenwert 0 senkrecht stehen. Axiom 3 ist also erfüllt.

Aus Axiom 3 folgt speziell im Falle von Axiom 1 η (endlich), daß es zu jedem $V \in K$ mit $C(V) \neq K$ ein $F \neq 0$ gibt mit $\mu(V, F) = 0$. Dies wurde in [5] als Axiom postuliert. Ob aber dieses letztere ausreicht, um daraus die stärkere Formulierung des in dieser Arbeit aufgestellten Axioms 3 abzuleiten, ist fraglich. Zur Herleitung der Sätze in [5] genügt das dortige Axiom. Es sei noch vermerkt, daß wir hier der Formulierung des ersten Hauptsatzes in dem hier angegebenen Axiom 2 den Vorzug vor den Formulierungen in [5] gegeben haben. Die hier gegebene Formulierung scheint uns physikalisch weniger einschränkend zu sein; alle mathematischen Konsequenzen bleiben aber unverändert.

VI. Koexistenz und Kommensurabilität von Effekten

Ein sehr wichtiger Punkt des Experimentierens mit Gesamtheiten und Effekten ist zwar schon mehrfach berührt worden, muß aber nun auch einer mathematischen Formulierung zugänglich gemacht werden, nämlich die Tatsache, daß bei einem einzigen Versuch (nicht bei der Wiederholung der Versuche zur Aufnahme der Häufigkeitsverteilung!) an dem Effekt-Apparat mehrere verschiedene Effekte hervorgerufen werden können. Man denke als Beispiele an: 1) eine Batterie von Zählrohren, 2) eine Nebelkammerspur, bei der man jedes Nebeltröpfchen als Effekt bezeichnen kann, 3) eine Photoplatte, bei der jedes geschwärzte Silberkorn einen Effekt darstellt.

Wir wollen solche Effekte, die an einer Apparatur bei einem Versuch auftreten können, als koexistent bezeichnen (von „gleichzeitig“ auftreten ist nicht die Rede!). Physikalisch ist also die Bedeutung von koexistent für die Elemente von $\mathbf{L}$ unmittelbar als experimentell gegeben anzusehen. Läßt sich dieser Begriff auf die Elemente von $\mathbf{L}$ übertragen? Zunächst wollen wir aus ganz geläufigen experimentellen Tatsachen einige Schlußfolgerungen ziehen:

Sind F_1' und F_2' zwei Effekte erster Art, also $F_1' \in \mathbf{L}$, $F_2' \in \mathbf{L}$, die an einer Apparatur auftreten können, so kann man durch „Schaltungen“ sofort leicht folgende Effekte erzeugen:

$F_1' \dot{+} F_2'$ = dem Effekt, der anspricht (registriert wird), sobald F_1' oder F_2', aber nicht beide, hervorgerufen werden. $F_1' \cdot F_2'$ = dem Effekt, der anspricht, wenn sowohl F_1' als auch F_2' hervorgerufen werden. Es ist nun bekannt, daß man durch wiederholtes Anwenden der beiden Operationen $\dot{+}$ und $\cdot$ *alle* möglichen Schaltkombinationen erhalten kann. Auf Grund der Definition ergibt sich, daß

$$\mu(V, F_1') + \mu(V, F_2') = 2\,\mu(V, F_1' \cdot F_2') + \mu(V, F_1' \dot{+} F_2')$$

für alle V sein muß und daß die Operationen $\dot{+}$ und $\cdot$ aus allen zusammen auftretenden Effekten einen kommutativen Ring machen, die „Schaltalgebra" zusammen auftretender Effekte.

Wenn nun F_1', F_2' zwei zusammen an einer Apparatur auftretende Effekte sind und F_1'', F_2'' zwei solche an einer anderen Apparatur, aber so, daß F_1' und F_1'' (bzw. F_2', F_2'') zur selben Klasse $F_1 \in \mathbf{L}$ (bzw. $F_2 \in \mathbf{L}$) gehören (d. h. z. B. $\mu(V, F_1') = \mu(V, F_1'')$ für alle V), so ist es *nicht* notwendig, daß auch $F_1' \dot{+} F_2'$ und $F_1'' \dot{+} F_2''$ zur selben Klasse gehören, d. h., daß

$$\mu(V, F_1 \dot{+} F_2) = \mu(V, F_1'' \dot{+} F_2'')$$

für alle V ist. Trotzdem aber kann man die Operationen $\dot{+}$ und $\cdot$ (zwar nicht notwendig eindeutig!) auf gewisse Elemente von $\mathbf{L}$ übertragen, indem man von einem speziellen Apparat (z. B. dem mit F_1' und F_2') ausgeht. Auf diese Weise läßt sich eine Schaltalgebra für eine gewisse Gruppe von Elementen aus $\mathbf{L}$ definieren, nämlich für solche Elemente, d. h. Klassen, in denen es Effekte (erster Art) gibt, die an einer Apparatur zusammen auftreten können. Eine solche Gruppe von Elementen aus $\mathbf{L}$ bildet dann eine Boolesche Algebra (nicht notwendig mit Eiselement). Als Boolesche Algebra bezeichnen wir einen kommutativen Ring (Operationen $\dot{+}$ und $\cdot$), für den $a \cdot a = a$ für jedes Element des Ringes gilt. Es folgt dann auch $a \dot{+} a = 0$. Durch diese Überlegungen werden wir zu folgender Definition geleitet [5]:

Definition 7: Eine Teilmenge $A \subseteq \hat{L}$ heißt *koexistent*, wenn es eine Menge R so gibt, daß $A \subseteq R \subseteq \hat{L}$ ist, und wenn sich in R zwei Operationen $\dot{+}$ und $\cdot$ so definieren lassen, daß R zu einer Booleschen Algebra wird und für je zwei Elemente $F_1 \in R$, $F_2 \in R$ für alle $V \in$ K gilt:

$$\mu(V, F_1) + \mu(V, F_2) = 2\,\mu(V, F_1 \cdot F_2) + \mu(V, F_1 \dot{+} F_2).$$

Das Wort koexistent ist nach Definition 7 eine rein mathematische Eigenschaft gewisser Teilmengen $A \subseteq \hat{L}$. (Durch Definition 7 ist eine Teilmenge der Potenzmenge von $\hat{L}$ ausgezeichnet, nämlich die Menge aller koexistenten Teilmengen von $\hat{L}$.) Zunächst hat damit Definition 7 *keine* physikalische Bedeutung. Ohne mathematisch ein neues Axiom, das Einfluß auf die Struktur von $\hat{L}$ hat, einzuführen, bekommen wir aber neue physikalische Aussagen, wenn wir an die Definition 7 ein neues physikalisches Prinzip (das eben nur *experimentell* bestätigt oder widerlegt werden kann) anknüpfen:

Ist R eine Definition 7 erfüllende Boolsche Algebra, so gibt es eventuell in R Boolesche Teilalgebren $\boldsymbol{r}$, für die auch noch $A \subseteq r$ gilt. Da der Durchschnitt Boolescher Algebren wieder eine solche ist, gibt es in R eine kleinste Boolesche Algebra, die noch A umfaßt; wir wollen diese A_r nennen. (Es ist nicht gesagt, daß man ausgehend von einer anderen Booleschen Algebra R' mit anders definierten $\dot{+}$ und $\cdot$ zur selben Algebra A_r kommt!) In die Definition 7 geht nur L ein. L erzeugt $\hat{L}$. Wir postulieren nun für L (relativ zu $\hat{L}$):

Axiom 4: Ist A koexistent und ist $A \subseteq L$, so gibt es wenigstens ein A_r mit $A_r \subseteq L$.

Nun das entscheidende physikalische Prinzip:

Prinzip der Koexistenz: Ist eine Menge $A \subseteq L$ nach Definition 7 koexistent, so läßt sich ein Apparat so konstruieren, daß an ihm Effekte (erster Art) $F_i'' \in \mathbf{L}$ auftreten, die folgende Bedingungen erfüllen: Zu jedem $F_i \in A$ gibt es ein physikalisch approximierendes Element $F_i' \in \mathbf{L}$ so, daß F_i'' der Klasse F_i' angehört.

Das Prinzip der Koexistenz setzt also eine mathematische Möglichkeit als Bild einer physikalischen Möglichkeit ein. Axiom 4 ist dann nur eine Konsequenz des Prinzips der Koexistenz. Wir haben es als Axiom 4 extra formuliert, weil nur dieses Stück des Prinzips der Koexistenz einen mathematischen Sinn hat. Aber auch mathematisch besagt Axiom 4 nicht viel. Es ändert nichts an der Struktur von $\hat{L}$, sondern legt nur dem L als Teilmenge von $\hat{L}$ eine gewisse Bedingung auf.

Nachdem wir das Prinzip der Koexistenz formuliert haben, wollen wir seine physikalische Bedeutung noch etwas näher erläutern, insbesondere die Bedeutung der Menge $\hat{L}$; denn man wäre sicher versucht, in Definition 7 die Menge L statt $\hat{L}$ zu benutzen. Beginnen wir deshalb mit der Menge $\hat{L}$. Zunächst wären wir aller Überlegungen enthoben, wenn $L = \hat{L}$ ist. In diesem Fall ist auch Axiom 4 unnötig. Zunächst haben wir die Menge $\mathbf{L}$ eingeführt als die Menge der physikalisch realisierbaren Effekte (zweiter Art). Daß wir nicht gleich $\mathbf{L} = L$ voraussetzen, hat einen guten Grund. Auf diesen können wir allerdings in dieser Arbeit noch nicht eingehen, doch sei auf einige Gesichtspunkte hingewiesen, damit es nicht zu künstlich aussieht, wenn wir $\mathbf{L}$ nicht mit L identifizieren. An einer späteren Stelle, bei weiterer Konkretisierung der Theorie, ist es notwendig, gewisse Stetigkeitseigenschaften der realisierbaren Effekte aus $\mathbf{L}$ zu verlangen, die nicht von allen Elementen aus L erfüllt werden können. Die Situation ist ähnlich der, die man erhält, wenn man eine Menge $\mathbf{L}$ von stetigen Funktionen in einem dreidimensionalen Raum hat, aber in bezug auf eine aus einer Maßtheorie folgenden Topologie zu der größeren Menge L der meßbaren Funktionen kommt. In bezug auf unsere Wahrscheinlichkeitsmaße befinden wir uns sozusagen erst im allgemeinen Rahmen der „meßbaren" Effekte aus L und müssen später gewisse „stetige" Effekte $\mathbf{L}$ durch neue Axiome auszeichnen. Um diese Möglichkeit von vornherein nicht auszuschließen, haben wir $\mathbf{L}$ neben L eingeführt. Wenn wir von dem eben diskutierten Abschließen absehen, so unterscheidet sich $\hat{L}$ von L und damit von $\mathbf{L}$ im wesentlichen dadurch, daß man in $\hat{L}$ auch alle Elemente $\sum_{i=1}^{n} \lambda_i F_i$ mit $F_i \in L$, $\lambda_i > 0$ und $\sum_i \lambda_i = 1$ hinzugenommen hat. Per definitionem von $\sum_{i=1}^{n} \lambda_i F_i$ ist

$$\mu(V, \textstyle\sum_i \lambda_i F_i) = \sum_i \lambda_i \mu(V, F_i).$$

Wahrscheinlichkeiten für die Elemente von L folgen also sehr einfach als Mischung der Werte für die einzelnen F_i. Ob man aber direkt einen Apparat bauen kann, an dem ein Effekt F gerade so erzeugt wird, daß man ihn als $\sum_i \lambda_i F_i$ interpretieren kann, wollen wir nicht unbdingt voraussetzen. Hier handelt es sich erst einmal tatsächlich um eine Vorsichtsmaßnahme, weil man eben mit der geringeren Annahme (im Axiom 4) als mit der Annahme $\hat{L} = L$ auskommt. Axiom 4 ist aber nur eine Folge des physikalischen Prinzips der Koexistenz. Bei diesem geht nun wirklich ein, daß koexistent mit Hilfe von $\hat{L}$ und und nicht nur L definiert (Definition 7) wurde.

Versuchen wir, uns die Bedeutung dessen an dem Beispiel zweier koexistenter Effekte F_1 und F_2 klarzumachen, d. h. für $A = \{F_1, F_2\}$. Es gibt also andere Elemente $F_1 \dot{+} F_2$, $F_1 \cdot F_2$, $F_1 \dot{+} F_1 \cdot F_2$, $F_2 \dot{+} F_1 \cdot F_2$, $F_1 \dot{+} F_2 \dot{+} F_1 \cdot F_2$ aus $\hat{L}$, die zusammen mit 0, F_1, F_2 gerade A_r bilden. Bezeichnen wir kurz

$$F_1 \cdot F_2 = F,$$
$$F_1 \dot{+} F_1 \cdot F_2 = F_1',$$
$$F_2 \dot{+} F_1 \cdot F_2 = F_2',$$

so besagt die Definition 7 gerade, daß

$$\mu(V, F_1) = \mu(V, F) + \mu(V, F_1'),$$
$$\mu(V, F_2) = \mu(V, F) + \mu(V, F_2'),$$
$$\mu(V, F_1 \dot{+} F_2 \dot{+} F_1 \cdot F_2) = \mu(V, F_1') + \mu(V, F_2') + \mu(V, F)$$

ist, was man kurz mit $+$ als Addition im Vektorraum B' auch

$$F_1 = F \dot{+} F_1', \qquad F_2 = F + F_2',$$
$$F_1 \dot{+} F_2 \dot{+} F_1 \cdot F_2 = F_1' + F_2' + F$$

schreiben kann. Gibt es umgekehrt neben F_1 und F_2 drei Elemente F, F_1', F_2' aus $\hat{L}$, so daß

$$F_1 = F + F_1', \quad F_2 = F + F_2' \quad \text{und} \quad F_1' + F_2' + F \in \hat{L}$$

gilt, so erhält man durch die Definitionen $F_1 \cdot F_2 = F$, $F_1 \dot{+} F_1 \cdot F_2 = F_1'$, $F_2 \dot{+} F_1 \cdot F_2 = F_2'$, $F_1 \dot{+} F_2 = F_1' + F_2'$ und $F_1 \dot{+} F_2 \dot{+} F_1 \cdot F_2 = F_1' + F_2' + F$ sofort den zu $A = \{F_1, F_2\}$ nach Definition 7 notwendigen Ring R. Damit haben wir folgenden Satz:

$A = \{F_1, F_2\}$ ist koexistent (wofür wir auch kurz F_1 und F_2 sind koexistent sagen), wenn es drei andere Elemente F, F_1', F_2' aus $\hat{L}$ gibt mit

$$F_1 = F + F_1', \quad F_2 = F + F_2', \quad F_1' + F_2' + F \in \hat{L}. \tag{6}$$

Ist dies umgekehrt der Fall, so kann man einen Booleschen Ring R, der $\{F_1, F_2\}$ umfaßt und die Bedingungen von Definition 7 erfüllt, definieren durch F, F_1', F_2' mit $F \cdot F_1' = 0$, $F \cdot F_2' = 0$, $F_1' \cdot F_2' = 0$. Es folgt aus (6) $F_1 = F \dot{+} F_1'$, $F_2 = F + F_2'$.

Unser Prinzip der Koexistenz besagt also, daß es zu zwei Effekten F_1, F_2, zu denen es $F, F_1', F_2' \in \hat{L}$ gibt, die (6) erfüllen, auch gerade solche F, F_1', F_2' (im Falle mehrerer solcher Möglichkeiten) gibt, zu denen man einen Apparat so bauen kann, daß an ihm Effekte F, F_1', F_2' so zusammen auftreten, daß die an diesem Apparat durch Schalttechnik erzeugten (d. h. $\dot{+}$ und $\cdot$ sind jetzt technisch – nicht nur mathematisch – definiert) Effekte $F \cdot F_1' = 0$, $F \cdot F_2' = 0$, $F_1' \cdot F_2' = 0$ und $F_1 = F + F_1'$, $F_2 = F \dot{+} F_2'$ werden.

Wenn wir in (6) die strengere Forderung, daß alle F_1', F_2', F und $F_1' + F_2' + F$ Elemente von L sind, gestellt hätten, so hätten wir indirekt zum *zweiten Male* die Forderung hineingesteckt, daß eben die F, F_1', F_2' konstruierbar sind. Das Prinzip der Koexistenz enthielte dann in sich einen Zirkel. So wie es aber hier eingeführt wurde, macht es eben eine Aussage über die Konstruktionsmöglichkeit von Effektapparaten. Diese letzte Argumentation ist natürlich kein Beweis; denn wie jedes physikalische Prinzip kann auch dieses nur an der Erfahrung getestet werden. Sollte aber dieses Prinzip in der angegebenen Form der Erfahrung widersprechen, so ist es auf keinen Fall etwa dadurch zu retten, daß man in der Definition von koexistent L statt $\hat{L}$ benutzt, es sei denn, man führt noch irgendeine neue, bisher noch nicht mit den Begriffen Gesamtheit, Effekt und Wahrscheinlichkeit erfaßbare physikalische Struktur ein, durch die $\hat{L}$ gegenüber L ausgezeichnet wird. In unserem Aufbau hier wird aber gerade in umgekehrter Richtung des Ablaufs der Folgerungen Axiom 4 zu einem Postulat an L.

Um spätere Aussagen kurz formulieren zu können, führen wir noch folgende Definition ein [5]:

Definition 8: Eine Teilmenge A von G (d. h. von Entscheidungseffekten) heißt *kommensurabel*, wenn es eine Menge R so gibt, daß $A \subseteq R \subseteq G$ ist, und wenn sich in R zwei Operationen $\dot{+}$ und $\cdot$ so definieren lassen, daß R zu einer BOOLEschen Algebra wird und für je zwei Elemente $E_1 \in R$, $E_2 \in R$ für alle $V \in K$ gilt:

$$\mu(V, E_1) + \mu(V, E_2) = 2\,\mu(V, E_1 \cdot E_2) + \mu(V, E_1 \dot{+} E_2).$$

Da $G \subseteq L$ gilt, geht also in die Definition 8 die Menge $\hat{L}$ nicht ein und es entfällt daher eine ähnliche Diskussion über die Bedeutung von $\hat{L}$. Wer das oben allgemein aufgestellte Prinzip der Koexistenz als zu weitehend empfindet, kann auch erst einmal Definition 7 nur als mathematische Definition einführen und Axiom 4 fortlassen. Dann müssen wir aber an einer anderen Stelle ein Prinzip der Kommensurabilität einführen, das sich bei unserer Darstellung als Folge des Prinzips der Koexistenz *ergibt* (Kapitel VII).

VI. Strukturen der Effekte und Gesamtheiten

In diesem Kapitel wollen wir kurz über einige wichtige Folgerungen aus den Axiomen 1 bis 3 berichten, die allerdings bisher [5b] nur im Fall des „Axioms 1 η (endlich)" abgeleitet werden konnten.

Zunächst ergibt sich[5], daß zwei Entscheidungseffekte E_1 und E_2 mit $E_1 \leqq E_2$ kommensurabel, also erst recht koexistent sind. Dies hat weiter zur Folge, daß man im Verband G der Entscheidungseffekte eine Orthokomplementation einführen kann: Jedem Entscheidungseffekt E kann eindeutig ein Entscheidungseffekt E^* durch

$$\mu(V, E) + \mu(V, E^*) = 1$$

für alle Gesamtheiten V aus K zugeordnet werden. Man bezeichnet deshalb oft E^* mit „nicht E". Wenn man noch $E_1 \cap E_2$ mit „E_1 und E_2" und $E_1 \cup E_2$ mit „E_1 oder E_2" bezeichnet, bekommt der Verband G die Form einer *formalen* Logik. In unserer Darstellung hat diese *nur formale* Logik nicht den Sinn einer Logik als Gesetze über Aussagen, sondern nur als abgekürzte, symbolische Redeweise, deren Sinn erst genau nach den Definitionen von E^*, $E_1 \cap E_2$, $E_1 \cup E_2$ zu eruieren ist.

Der Verband G wird mit der eingeführten Orthokomplementation zu einem orthokomplementären Verband. Im Falle der Quantenmechanik, wo die Elemente von G Projektionsoperatoren sind, ist E^* der zu E orthogonale Projektionsoperator, d. h. die zu E und E^* gehörigen Teilräume des HILBERT-Raumes sind orthogonal und spannen zusammen den ganzen HILBERT-Raum auf. Auch in G können wir *allgemein* eine „Orthogonalität" zweier Elemente definieren: $E_1 \perp E_2$ genau dann, wenn $E_1 \leqq E_2^*$ (dann ist auch $E_2 \leqq E_1^*$!).

Der Verband G ist auch noch *orthomodular*. Die Orthomodularität kann auf viele Arten und Weisen eingeführt werden, z. B. dadurch, daß für zwei orthogonale $E_1 \perp E_2$ und ein drittes $E_3 \geqq E_1$ die *modulare* Bedingung

$$E_3 \cap (E_1 \cup E_2) = E_1 \cup (E_3 \cap E_2)$$

gilt. Äquivalent damit ist folgende Bedingung:

Aus $E_2 \perp E_1$, $E_2' \perp E_1$ und $E_1 \cup E_2 = E_1 \cup E_2'$ folgt $E_2 = E_2'$.

Die Wahrscheinlichkeiten sind auf G orthoadditiv, d. h., für paarweise orthogonale E_i gilt für alle $V \in K$ [5]:

$$\mu(V, \bigcup_i E_i) = \sum_i \mu(V, E_i).$$

Dies ist eine im HILBERT-Raum der Quantenmechanik sehr bekannte Relation; und umgekehrt hat diese Relation für eine Maßfunktion $m(E)$ über den Projektionsoperatoren E (mit $0 \leqq m(E) \leqq 1$) nach GLEASON [7] zur Folge, daß sie die Form $\mu(V, E) = \mathrm{Sp}(VE)$ mit V als positiv semidefinitem HERMITEschen Operator mit $\mathrm{Sp}(V) = 1$ hat.

[7] A. M. GLEASON, J. Math. Mech. 6, 885 [1957].

Mit Hilfe der Operationen $*$, $\cap$, $\cup$ läßt sich leicht eine hinreichende und notwendige Bedingung dafür angeben, daß zwei Entscheidungseffekte E_1 und E_2 kommensurabel sind (im Falle der Quantenmechanik siehe z. B. [8]):

$$E_1 = (E_1 \cap E_2^*) \cup (E_1 \cap E_2).$$

Jedes Element $F \in \hat{L}$ läßt sich in der Form $F = \sum_\nu \lambda_\nu E_\nu$ mit $0 \leqq \lambda_\nu \leqq 1$, E_ν paarweise orthogonale Elemente aus G darstellen ($\sum_\nu$ ist hierbei wieder im Sinne der Vektoraddition in B' verstanden, d. h. durch $\mu(V, F) = \sum_\nu \lambda_\nu \mu(V, E_\nu)$ gegeben). Dies ist im Falle der Quantenmechanik (in einem endlich dimensionalen HILBERT-Raum, Axiom 1 η endlich!) die Spektralzerlegung des HERMITEschen Operators F, wobei λ_ν seine Eigenwerte und E_ν die Projektionsoperatoren auf die Eigenräume sind.

Wir sehen also, wie viele bekannte Strukturen der Quantenmechanik allein schon aus den hier angegebenen Axiomen 1 bis 3 folgen.

Wir wollen dieses Kapitel damit beenden, daß wir die physikalische Bedeutung des Prinzips der Koexistenz näher diskutieren. Die aufgeführten Aussagen in diesem Kapitel benutzen nur rein formal mathematisch die Definitionen 7 und 8. Wir hatten uns im vorigen Kapitel die physikalische Bedeutung der Koexistenz für zwei Effekte F_1 und F_2 klar gemacht. Was ergibt sich speziell, wenn F_1 und F_2 zwei Entscheidungseffekte E_1 und E_2 sind? Zwei koexistente Entscheidungseffekte sind auch immer kommensurabel und die Zerlegung (6) (notwendig für die Koexistenz) ist im Falle zweier Entscheidungseffekte *nur* in der Form

$$\begin{aligned} E_1 &= (E_1 \cap E_2) + (E_1 \cup E_2^*), \\ E_2 &= (E_1 \cap E_2) + (E_2 \cap E_1^*), \\ (E_1 \cap E_2) &+ (E_1 \cup E_2^*) + (E_2 \cap E_1^*) = E_1 \cup E_2 \end{aligned} \tag{7}$$

möglich. Aus der Bedingung $\{E_1, E_2\} \subseteq R \subseteq \hat{L}$ mit R als kleinstem, $\{E_1, E_2\}$ umfassenden Ring folgt also auch $\{E_1, E_2\} \subseteq R \subseteq G$. Für Entscheidungseffekte ist also wegen $G \subseteq L$ immer von selbst $R \subseteq L$, d. h., die im vorigen Kapitel durchgeführte Diskussion zur Einführung der Menge $\hat{L}$ hat nur für Effekte F, die keine Entscheidungseffekte sind, eine Bedeutung. Für Entscheidungseffekte wird [wegen der Eindeutigkeit der Zerlegung (7)] die (physikalisch approximative) Konstruktion eines Apparates eindeutig, was sich aus der Diskussion im Anschluß an (6) ergibt. Aus dem Prinzip der Koexistenz folgt also das

Prinzip der Kommensurabilität: Kommensurable Entscheidungseffekte können zusammen von einer Apparatur gemessen werden.

Man hätte auch dieses Prinzip als Basis zur Beschreibung des „Zusammenmessens" machen können, könnte dann aber ohne weitere Prinzipien nicht zurück zum allgemeinen Begriff der Koexistenz kommen. Daher haben wir den hier angegebenen Weg vorgezogen und das stärkere Prinzip der Koexistenz postuliert.

[8] G. LUDWIG, Grundlagen der Quantenmechanik, Springer-Verlag, Berlin–Göttingen–Heidelberg 1954.

VIII. Zerlegung der Gesamtheiten und Effekte in irreduzible Teile

Die dargestellte Theorie enthält die Quantenmechanik (worauf wir immer wieder hinwiesen), aber auch jede klassische Theorie: z. B. K die Menge aller meßbaren Funktionen $\varrho(x)$ (x kann z. B. Punkt im Phasenraum der klassischen Mechanik sein) mit $0 \leqq \varrho(x)$ und $\int \varrho(x) \, dm(x) = 1$, $\hat{L}$ die Menge aller meßbaren Funktionen $f(x)$ mit $0 \leqq f(x) \leqq 1$,

$$\mu(V, F) = \int \varrho(x) f(x) \, dm(x).$$

Nun läßt jede Theorie, die die Axiome 1 bis 3 erfüllt, sich in irreduzible Teile zerlegen, was sich einfach im Falle des Axioms 1 η (endlich) zeigen läßt [5]. Eine „klassische" Theorie, die das Axiom 1 η (endlich) erfüllt, ist z. B. der sehr bekannte, einfache Fall: Gegeben eine Menge M von *endlich* vielen Elementen $x_1 \ldots x_n$; K = Menge aller $\varrho(x)$ mit $0 \leqq \varrho(x)$ und $\sum_{i=1}^{n} \varrho(x_i) = 1$; $\hat{L}$ = Menge aller $f(x)$ mit $0 \leqq f(x) \leqq 1$; $\mu(\varrho, f) = \sum_{i=n}^{n} \varrho(x_i) f(x_i)$. Die Menge G der Entscheidungseffekte ist die Menge der $f(x)$ mit $f(x) = 0$ oder 1; damit kann G auch mit der Menge aller Teilmengen von M identifiziert werden.

Im allgemeinen Fall der Axiome 1 bis 3 führe man folgende, auch physikalisch sehr bedeutungsvolle Begriffe ein:

Als Zentrum Z von G bezeichne man die Menge der Entscheidungseffekte E', wobei jedes E' mit allen Elementen von G kommensurabel ist, d. h. $E' \in Z$ ist äquivalent zu: $\{E', E\}$ kommensurabel für alle E aus G. 0 und 1 liegen in Z. G heißt irreduzibel, wenn Z *nur* aus 0, 1 besteht.

Ist G nicht irreduzibel, so kann man [im Falle des Axioms 1 η (endlich)] folgende Ausreduktion durch-führen: E_α seien die kleinsten Elemente ungleich Null aus Z. Die E_α sind dann paarweise orthogonal. Es gilt $\sum_\alpha E_\alpha = 1$ ($\sum_\alpha$ ist endlich, da es wegen Axiom 1 η (endlich) nur endlich viele E_α gibt). Jedes $F \in \hat{L}$ hat die Form $\sum_\alpha F_\alpha$ mit $F_\alpha \leqq E_\alpha$. Jedes V kann in der Form

$$V = \sum_\alpha w_\alpha V_\alpha$$

geschrieben werden mit

$$\mu(V, F) = \sum_\alpha w_\alpha \mu(V_\alpha, F_\alpha), 0 \leqq w_\alpha, \sum_\alpha w_\alpha = 1,$$
$$\mu(V_\alpha, E_\alpha) = 1.$$

Für festes α erfüllen die Mengen K_α der Komponenten V_α und $\hat{L}_\alpha$ der Komponenten F_α alle Axiome 1 bis 3.

Diese mathematische Zerlegung legt folgenden Sprachgebrauch nahe: Die Elemente von Z bezeichnen wir als „Eigenschaften“, die kleinsten Elemente $E_\alpha \neq 0$ von Z als „elementare Eigenschaften“. Jede Eigenschaft aus Z ist dann eine Zusammenfassung $\sum'_\alpha E_\alpha$ einiger der elementaren Eigenschaften E_α ($\sum'_\alpha$ deutet an, daß die $\sum'_\alpha$ nicht über alle E_α, sondern nur über einen Teil der E_α zu erstrecken ist). Es ist weiterhin die Vorstellung möglich, daß jedes Mikroobjekt genau eine und nur eine dieser elementaren Eigenschaften E_α besitzt, denn *jede* Gesamtheit V von Mikroobjekten ist ein Gemisch $V = \sum_\alpha w_\alpha V_\alpha$ von Gesamtheiten V_α, wobei wegen $\mu(V_\alpha, E_\alpha) = 1$ [und $\mu(V_\alpha, E_{\alpha'}) = 0$ für $\alpha' \neq \alpha$!] alle Mikroobjekte der Gesamtheit V_α die Eigenschaft E_α haben. Wir wollen die Objekte einer solchen Gesamtheit von gleicher *Art* nennen. Zum Verständnis dieser Sachlage hatten wir am Ende von Abschnitt I c als Beispiel aus der Quantenmechanik Gesamtheiten von Wasserstoffatomen und Heliumatomen betrachtet. In diesem Beispiel ist die eine elementare Eigenschaft: Wasserstoff, die andere: Helium. Ist $Z = G$, so sind alle Entscheidungseffekte miteinander kommensurabel und dann auch alle Effekte miteinander koexistent. Diesen Fall wollen wir den *klassischen* Fall nennen. K und $\hat{L}$ ist dann genau mit dem oben diskutierten Fall der Funktionen über der endlichen Menge M der $x_1 \ldots x_n$ identisch.

Eine notwendige und hinreichende Bedingung für den klassischen Fall wird aber gerade durch das Axiom 2k (Kapitel III) gegeben.

Im allgemeinen Fall sind die Mengen K_α, $\hat{L}_\alpha$ nicht triviale [9] irreduzible Lösungen des Axiomensystems 1 bis 3. So ist z. B. die Quantenmechanik von Wasserstoffatomen irreduzibel.

Ein irreduzibles System ist, wie wir sahen, der Struktur des HILBERT-Raumes der Quantenmechanik sehr ähnlich. Gibt es noch andere Lösungen? Dies scheint der Fall zu sein. Bei der Klassifizierung der verschiedenen irreduziblen Lösungen scheint die Automorphismengruppe (d. h. die Gruppe aller linearen Transformationen in B', die $\hat{L}$ umkehrbar eindeutig auf sich abbilden) eine wichtige Rolle zu spielen.

IX. Schlußbemerkungen über die Objektivierbarkeit physikalischer Theorien

Am Ausgangspunkt unserer physikalischen Theorien stehen „objektive“, durch Symbole der Form V, F charakterisierte Situationen. Das Wort ‚objektiv‘ heißt dabei, daß die Situation (V, F) als eine unabhängig vom Subjekt vorgegebene angesehen wird. Der ganze physiologisch-psychologische Prozeß des Erkennens der Situation (V, F) wurde ausgeklammert. Nun ist die Menge der so objektiv erkennbaren Situationen durchaus nicht vom Beginn der Physik an wohl-definiert; d. h., es kann sein, daß man eine bisher als objektiv gegeben betrachtete Situation S in Frage stellt. Dieses Infragestellen hat aber nur dann einen physikalischen Sinn, wenn man die Hypothese, daß eine Situation S als objektiv gegeben anzusehen ist, mit den *erkannten* physikalischen Sätzen und anderen als objektiv gegebene Situation vergleicht. Entsteht kein Widerspruch, kann die Hypothese der „Objektivität“ von S beibehalten werden.

Die hier kurz geschilderte Methode, Physik als Wissenschaft zu betreiben (und genau so wurde bisher immer Physik betrieben) würde zusammenbrechen, wenn nicht ein Bereich von solchen Situationen übrig bliebe, die ohne „physikalische Widersprüche in sich“ als objektiv real gegeben angesehen werden dürfen und auf denen die Physik aufgebaut werden kann. Diese Konsistenz in sich ist kein „Beweis“ für die Richtigkeit des Ausgangspunktes, sondern nur eine Rechtfertigung der Methode, ein Beweis für die Unangreifbarkeit der Physik als Wissenschaft. In einer Erfahrungswissenchaft kann diese Konsistenz (Widerspruchsfreiheit) natürlich nicht endgültig bewiesen, sondern nur approximativ erhärtet werden. (Heute weiß man, daß auch und sogar für die Mathematik ein solcher Beweis der Widerspruchsfreiheit unmöglich ist.)

[9] Als triviale Lösung sehen wir an: K=der Menge, die nur aus der Zahl 1 besteht, L=Menge aller Zahlen f mit $0 \leqq f \leqq 1$ und $\mu(V, F) = f$.

Es wird nun die Meinung vertreten, daß die Quantenmechanik gezeigt hätte, daß die eben geschilderte Methode, Physik als Wissenschaft zu betreiben, d. h. von objektiv gegebenen Situationen auszugehen, unmöglich ist. Gerade deshalb müsse man einen Schritt weitergehen und in die Physik schon die Sprache über Physik mit hineinnehmen: Die „formale“ Logik der Entscheidungseffekte sei gar keine formale Logik, sondern vielmehr die für die Sprache der Physik allgemein gültige Logik. *Nur so* sei die Konsistenz der Physik als Wissenschaft zu retten.

Dieser vermeintliche, zur Rettung der Physik notwendige Schritt wird dadurch erkauft, daß der Inhalt physikalischer Aussagen unklar bleibt und die ganze Physik zu einem leeren (zwar in sich konsistenten) Formalismus wird; es sei denn, daß man doch an irgendeiner Stelle den Inhalt physikalischer Sätze auf in ihrer Bedeutung unmittelbar sinnvolle und nicht nur formale Sätze zurückführt, und sei es auch, daß man dies erst bei der subjektiven Feststellung einer Wahrnehmung tut.

Wir sind aber der Meinung, daß überhaupt kein physikalischer Grund besteht, von objektiv gegebenen realen Situationen abzugehen. Trotz aller modernen Physik steht ein Glas genau so objektiv auf einem Tisch wie vorher. Ist es aber nicht *notwendig* von der oben geschilderten Methode, Physik als Wissenschaft zu betreiben (nämlich von einem Bereich objektiv gegebener Sachverhalte auszugehen), abzugehen, so sollte man dies auch *nicht tun,* so interessant auch nachträglich physiologisch-psychologische Probleme der Wahrnehmung sind, die aber eben nur dann als solche behandelt werden können, wenn die Physiologie als Teil der Physik betrachtet wird. Innerhalb des physiologischen Bereiches ist *nur* die physikalische Konsistenz in sich zu verlangen, die aber zu erwarten ist, wenn man die physikalische Konsistenz ganz allgemein (approximativ) hat erhärten können. Durch die Einbeziehung der Psychologie der Wahrnehmungen kommt man dann zu einer Untersuchung der Konsistenz zwischen der Voraussetzung objektiver Situationen und ihrer sinnlichen Wahrnehmung, einer Konsistenz auf einer höheren Ebene. Von einer genaueren Lösung dieser physiologisch-psychologischen Probleme kann aber die Physik absehen, weil sie sich als in *sich* konsistente Wissenchaft formulieren läßt.

Gerade auch deswegen, um diese Tatsache herauszukehren, haben wir diesen Aufbau zur Begründung auch quantenmechanischer Theorien gestartet. Natürlich kann erst eine nachträgliche Theorie des Meßprozesses auf Grund einzelner physikalischer Gesetze die gewünschte Konsistenz erhärten. Dieses Meßproblem ist gerade so physikalisch sinnvoll und lösbar (obwohl manche auf Grund von Vorentscheidungen oder vermeintlicher Beweise auf Grund physikalisch unzulässiger Verallgemeinerungen anderer Meinung sind). Hier können wir aber nicht mehr auf diese Fragen eingehen.

Eine ganz andere Frage ist es, inwieweit die hier dargestellte Theorie es erlaubt, von objektiv gegebenen Mikrosystemen mit objektiven Eigenschaften und objektiv an ihnen haftenden Wahrscheinlichkeiten zu sprechen. Wenn in einer solchen Weise die Physik interpretiert wird, so ist mehr gemeint, als die von uns hier gegebenen Definitionen z. B. von Effekten, Entscheidungseffekten, Eigenschaften und elementaren Eigenschaften. Gerade aber dadurch, daß man *mehr* meint, offenbaren sich diese Überlegungen als nicht mehr rein physikalische Überlegungen. Sie wollen vielmehr eine in irgendeiner Weise ontologische Interpretation der physikalischen Strukturen geben. Dagegen ist nichts einzuwenden. Mir scheint aber, daß auch hierfür der hier gegebene Aufbau der Theorie klärend wirken kann, weil er besser erlaubt, die einzelnen Schritte sauber zu trennen.