Endlich-dimensionale Dichtematrizen I[1]

Von A. UHLMANN, Leipzig

1. Dichtematrizen

Die Dichtematrizen (Dichteoperatoren) wurden durch *J. von Neumann* in die Quantenphysik eingeführt, der auch eine Reihe ihrer wichtigsten allgemeinen Eigenschaften bewies.

Im Abschnitt 1.1. werden die endlich-dimensionalen Dichtematrizen definiert und ihre einfachsten Eigenschaften angegeben. Jede Matrix kann auch als linearer Operator aufgefaßt werden, der in einem endlich-dimensionalen linearen Raum (Vektorraum) wirkt. Daher ist auch das Wort „Dichteoperator" sehr gebräuchlich. Im Abschnitt 1.1. sind alle Beweise nur angedeutet, da sie praktisch in jedem Lehrbuch über Matrizen oder lineare Operatoren zu finden sind. Daher sollte dieser Abschnitt als eine Erinnerung an wohlbekannte Dinge und eine Einführung in einige Bezeichnungsweisen angesehen werden.

Im Abschnitt 1.2. wird gezeigt, weshalb die Dichtematrizen (-operatoren) eindeutig physikalischen Zuständen eines (hier endlichen) physikalischen Systems zugeordnet sind. Wegen der „Endlichkeit" und der „Irreduzibilität" der von uns betrachteten Klasse physikalischer Systeme ist diese Zuordnung sogar eindeutig. Daher bezeichnen wir die Dichtematrizen gelegentlich auch als „Zustände" und die Gesamtheit dieser Zustände auch als „Zustandsraum" unseres physikalischen Systems. Die letztere ist *keine* Standardbezeichnung, da meist die Gesamtheit der sog. „reinen" Zustände als Zustandsraum bezeichnet wird und nicht die Gesamtheit „aller", also auch der „gemischten" Zustände.

Der Abschnitt 1.3. kann beim ersten Studium überschlagen werden. Er enthält eine gewisse Rechtfertigung der hier benutzten Sprechweise durch den sog. algebraischen Zugang zur Quantentheorie mit Hilfe der C*-Algebren oder allgemeiner der *-Algebren. In diesen Theorien tritt der allgemeine Begriff des Zustands im Sinne des *von Neumannschen* Zustandsgemisches als primärer und leicht definierbarer (natürlicher) Begriff auf, während „reine" Zustände entweder als Extremalpunkte (siehe Abschnitt 2.) des Zustandsraumes oder durch die Irreduzibilität der von ihnen induzierten Darstellung (Gelfand-Neumark-Segal-Darstellung) ausgezeichnet sind. Auf den letztgenannten, unmittelbar mit dem Überlagerungsprinzip verbundenen Zusammenhang gehen wir jedoch in dieser Arbeit nicht explizit ein.

[1] Die nachstehende Arbeit ist Teil einer Vorlesung, die der Autor im Frühjahrssemester 1971 gehalten und auf Grund der Ausarbeitungen von *W. Laßner* und *P. Alberti* vervollständigt hat. Sie gehört zu einem Vorlesungszyklus „Observable und Zustände", der von *G. Laßner* (Sektion Mathematik der Karl-Marx-Universität) und dem Autor geleitet wird.

1.1. Definitionen und elementare Eigenschaften

Wir betrachten zunächst $n \times n$-Matrizen A mit komplexen Matrixelementen. Die zu A hermitisch konjugierte Matrix wird mit A^* bezeichnet. A heißt positiv semidefinit, wenn

$$\Sigma A_{ik}\xi^i\bar{\xi}^k \geq 0 \tag{1.1}$$

für alle komplexen Vektoren ξ mit Komponenten $\xi^1, \ldots, \xi^n$ ist. Dabei bezeichnet $\bar{\xi}^k$ die zu ξ^k konjugierte komplexe Zahl. Wir schreiben genau dann

$$A \geq B, \tag{1.2}$$

wenn $A - B$ positiv semidefinit ist. Also bedeutet $A \geq 0$, daß A positiv semidefinit ist. Geht man in (1.1) zum Konjugiert-Komplexen über und vertauscht die Summationsindizes, so bekommt man leicht, daß $A = A^*$, A also hermitisch ist. Einsetzen der Eigenvektoren von A zeigt dann, daß kein Eigenwert negativ sein darf. Wegen der Diagonalisierbarkeit der hermitischen Matrizen kann man ebenso einfach aus der Nicht-Negativität der Eigenwerte auf die positive Semidefinitheit schließen:

Lemma 1.1.

Eine Matrix ist genau dann positiv-semidefinit, wenn sie hermitisch ist und keinen negativen Eigenwert besitzt.

Als nächstes erinnern wir an die Definition der Spur

$$\mathrm{Tr.}\, A = \Sigma A_{ii} \quad \text{für} \quad A = (A_{ik}). \tag{1.3}$$

Ehe wir jedoch die wichtigsten Eigenschaften der Spur referieren, wollen wir die $n \times n$-Matrizen noch wie üblich den linearen Operatoren eines n-dimensionalen linearen Raumes zuordnen. In der Tat bilden die Vektoren

$$\xi = (\xi^1, \ldots, \xi^n)$$

einen komplexen linearen Raum der Dimension n, den wir mit $\mathfrak{H}$ bezeichnen wollen:

$$\mathfrak{H} = \{\xi : \xi = (\xi^1, \ldots, \xi^n), \xi^i \text{ bel. komplex}. \tag{1.4}$$

Es ist dann die Addition und die skalare Multiplikation durch

$$\begin{aligned}(\xi + \eta)^i &= \xi^i + \eta^i, \\ (\lambda\xi)^i &= \lambda\xi^i, \quad \lambda \text{ komplexe Zahl}\end{aligned} \tag{1.5}$$

definiert. Eine der möglichen Basen dieses Raumes ist durch die Vektoren

$$(1, 0, 0, \ldots, 0),\ (0, 1, 0, 0, \ldots), \ldots, (0, 0, \ldots, 0, 1) \tag{1.6}$$

gegeben. Durch Einführen des Skalarproduktes

$$(\xi, \eta) = \sum_{i=1}^{n} \bar{\xi}^i \eta^i \tag{1.7}$$

wird $\mathfrak{H}$ zu einem Hilbert-Raum der Dimension n; denn

$$(\xi, \xi) \geq 0 \tag{1.8}$$

$$\begin{aligned}(\xi, \eta) &= \overline{(\eta, \xi)} \\ (\xi, \lambda\eta) &= \lambda(\xi, \eta), \quad \lambda \text{ komplexe Zahl}\end{aligned} \tag{1.9}$$

und

$$(\xi, \xi) = 0 \quad \text{genau dann, wenn} \quad \xi = 0. \tag{1.10}$$

Die Matrix $A = (A_{ik})$ identifizieren wir nun mit dem linearen Operator A, der durch

$$(A\xi)^i = \sum_{k=1}^{n} A_{ik}\xi^k \tag{1.11}$$

definiert ist.
Die Einheitsmatrix bzw. den Einsoperator (Identität) bezeichnen wir mit

$\underline{1}$ bzw. $\underline{1}_n$,

wenn die Dimension des Raumes hervorgehoben werden soll:

$$\underline{1} = (\delta_{ik}). \tag{1.12}$$

Es ist A^* auch durch die Forderung

$$(\xi, A^*\eta) = (A\xi, \eta) \quad \text{für alle} \quad \xi, \eta \in \mathfrak{H} \tag{1.13}$$

definierbar. Weiter ist

$$A \geq 0 \quad \text{genau dann, wenn} \quad (\xi, A\xi) \geq 0 \quad \text{für alle} \quad \xi \tag{1.14}$$

gilt.
Eine vollständige Orthonormalbasis $\eta_1, \ldots, \eta_n$ besteht aus genau n Vektoren η_j mit

$$(\eta_i, \eta_k) = \delta_{ik},$$

wobei wie üblich $\delta_{ii} = 1$ und $\delta_{ik} = 0$ für $i \neq k$ gesetzt wird. (1.6) ist eine vollständige Orthonormalbasis. Sind $\xi_1, \ldots, \xi_n$ Vektoren und $\eta_1, \ldots, \eta_n$ eine vollständige Orthonormalbasis, so ist der durch

$$U\eta_i = \xi_i, \quad i = 1, 2, \ldots, n$$

definierte lineare Operator (und die ihm nach (1.11) zugeordnete Matrix) genau dann unitär, d.h.

$$UU^* = U^*U = \underline{1},$$

wenn auch $\xi_1, \ldots, \xi_n$ eine vollständige Orthonormalbasis ist.
Nun wollen wir zu (1.3) zurückkehren.

Lemma 1.2.

Ist $\eta_1, \eta_2, \ldots, \eta_n$ eine beliebige vollständige Orthonormalbasis, so ist

$$\text{Tr.}\, A = \sum_{i=1}^{n} (\eta_i, A\eta_i). \tag{1.15}$$

Beweis: Indem wir das vollständige Orthonormalsystem (1.6) einsetzen, sehen wir die Richtigkeit von (1.15) für die Basis $\xi_1, \ldots, \xi_n$ mit $(\xi_i)^k = \delta_{ik}$. Es ist aber

$$\text{Tr.}\, A = \sum (\xi_i, A\xi_i) = \sum_{i,j,k} (\eta_j, A\eta_k)\, (\xi_i, \eta_j)\, (\eta_k, \xi_i) = \sum_{j,k} (\eta_j, A\eta_k)\, \delta_{jk} = \sum_k (\eta_k, A\eta_k),$$

wobei wir die für zwei vollständige Orthonormalbasen geltende Formel

$$\sum_{i=1}^{n} (\eta_k, \xi_i)(\xi_i, \eta_j) = \delta_{kj} \tag{1.16}$$

benutzt haben. |=|

Aus Lemma 1.2. leiten sich leicht weitere wichtige Eigenschaften der Spur ab.

Eigenschaft a

Ist A hermitisch und sind $\lambda_1, \lambda_2, \ldots, \lambda_n$ sämtliche Eigenwerte von A, so ist

$$\sum_{i=1}^{n} \lambda_i = \mathrm{Tr.}\, A. \tag{1.17}$$

Beweis: In (1.15) wird ein vollständiges Orthonormalsystem von Eigenvektoren von A gewählt.

Eigenschaft b

Für unitäres U ist

$$\mathrm{Tr.}\, A = \mathrm{Tr.}\, UAU^{-1}. \tag{1.18}$$

Beweis: Nach Lemma 1.2. ist

$$\mathrm{Tr.}\, A = \Sigma\, (U\xi_i, AU\xi_i) = \Sigma\, (\xi_i, U^{-1}AU\xi_i).$$

Eigenschaft c

Es ist

$$\begin{aligned} \mathrm{Tr.}\, \lambda A &= \lambda\, \mathrm{Tr.}\, A, \quad \lambda \text{ komplexe Zahl} \\ \mathrm{Tr.}\, (A + B) &= \mathrm{Tr.}\, A + \mathrm{Tr.}\, B \\ \mathrm{Tr.}\, A^* &= \overline{\mathrm{Tr.}\, A}. \end{aligned} \tag{1.19}$$

Eigenschaft d

Es ist

$$\mathrm{Tr.}\, AB = \mathrm{Tr.}\, BA. \tag{1.20}$$

Beweis: Seien zunächst A und B hermitisch und $\xi_1, \ldots, \xi_n$ eine Orthonormalbasis von Eigenvektoren von A; dann ist

$$\mathrm{Tr.}\, AB = \Sigma\, (\xi_i, AB\xi_i) = \Sigma \lambda_i\, (\xi_i, B\xi_i)$$

$$\mathrm{Tr.}\, BA = \Sigma\, (\xi_i, BA\xi_i) = \Sigma\, (\xi_i, B\xi_i)\, \lambda_i.$$

Also gilt (1.20) für hermitische Operatoren (Matrizen). Sind nun A und B beliebig, so betrachten wir ihre Zerlegung

$$A = A_1 + iA_2, \quad B = B_1 + iB_2$$

in hermitische Operatoren A_1, A_2, B_1, B_2. Durch Einsetzen dieser Zerlegungen in Tr. AB kann die allgemeine Gültigkeit von (1.20) auf die von (1.20) für hermitische Operatoren zurückgeführt werden. |=|

Eigenschaft e

Ist $A \geq 0$ und $B \geq 0$, so ist

$$\text{Tr. } AB \geq 0 \tag{1.21}$$

und gilt in (1.21) das Gleichheitszeichen, so ist $AB = 0$.

Beweis: Ist $A \geq 0$, so ist mit den Bezeichnungen des Beweises von Eigenschaft d einmal $\lambda_i \geq 0$ und zum anderen $(\xi_i, B\xi_i) \geq 0$. Also folgt aus der Voraussetzung

$$\text{Tr. } AB = \Sigma \lambda_i (\xi_i, B\xi_i) \geqslant 0. \tag{1.22}$$

Sei nun Tr. $AB = 0$. Wir ordnen die λ_1 so, daß $\lambda_j \geq \cdots \geq \lambda_m > 0$, aber $\lambda_{m+1} = \cdots = \lambda_n = 0$ ist. Dann folgt aus (1.22) wegen Tr. $AB = 0$

$$(\xi_i, B\xi_i) = 0 \quad \text{für} \quad i = 1, \ldots, m$$

und, da $B \geq 0$ sein sollte, sogar $B\xi_i = 0$. Mit

$$\xi = \sum_{i=1}^{n} \alpha_i \xi_i, \quad \alpha_j \text{ beliebig}$$

folgt somit

$$(\xi, AB\xi) = (A\xi, B\xi) = \Sigma\, \alpha_k \bar{\alpha}_i \lambda_i (\xi_i, B\xi_k) = \sum_{k=1}^{n} \alpha_k \sum_{i=1}^{m} \bar{\alpha}_i \lambda_i (B\xi_i, \xi_k) = 0.$$

Eigenschaft f

Sei B hermitisch. Genau dann ist

$$\text{Tr. } AB \geq 0 \quad \text{für alle} \quad A \geq 0, \quad \text{wenn} \quad B \geq 0. \tag{1.23}$$

Beweis: Bei Eigenschaft e wurde schon bewiesen, daß $B \geq 0$ hinreichend für (1.23) ist. Gelte daher (1.23). Sei $\eta_1, \ldots, \eta_n$ eine vollständige Orthonormalbasis von B. Die Operatoren

$$A_i \xi = (\eta_i, \xi)\, \eta_i \tag{1.24}$$

sind die Projektionsoperatoren auf die η_i, und wegen

$$(\xi, A_i \xi) = (\eta_i, \xi)(\xi, \eta_i) = |(\eta_i, \xi)|^2 \geq 0 \tag{1.25}$$

sind diese Operatoren sämtlich positiv-semidefinit. Daher ist

$$0 \geq \text{Tr. } A_i B = \Sigma (\eta_j, A_i B \eta_j) = \Sigma \lambda_j (\eta_j, A_i \eta_j) = \lambda_i$$

und die Eigenwerte von B können nicht negativ sein. Das zeigt wegen Lemma 1.1., daß $B \geq 0$ auch notwendig ist. |=|

Nach diesen Vorbereitungen kommen wir zur Definition der Dichtematrizen (Dichteoperatoren).

Definition 1.1.

Eine Matrix ϱ (ein linearer Operator ϱ) heißt Dichtematrix (Dichteoperator) genau dann, wenn

a) $\varrho \geq 0$
b) Tr. $\varrho = 1$

gilt.

Nach dem bisherigen können wir auch sagen:

Lemma 1.3.

ϱ ist genau dann Dichtematrix (Dichteoperator), wenn ϱ hermitisch ist und für die Eigenwerte $\lambda_1, \lambda_2, \ldots, \lambda_n$ von ϱ gilt

$$\lambda_i \geq 0, \quad \sum_{j=1}^{n} \lambda_j = 1. \tag{1.26}$$

Bemerkung: Dichtematrizen (Dichteoperatoren) bezeichnen wir mit den Buchstaben $\varrho, \varrho', \omega, \ldots$ Wie im nächsten Abschnitt erläutert wird, spielen gewisse Matrizen physikalisch eine Doppelrolle. Einmal haben wir Matrizen, die den Zustand eines Systems beschreiben (Dichtematrizen $\varrho, \omega, \ldots$), und zum anderen hermitische Operatoren (Matrizen $A, B, \ldots$), die Observable des Systems beschreiben. Die Unterscheidung der Schreibweise ist also mathematisch ohne Bedeutung und soll nur auf die unterschiedliche physikalische Stellung des Objektes hinweisen. Im Abschnitt 1.3. werden wir jedoch andeuten, daß man Zuständen und Observablen auch getrennte mathematische Objekte zuordnen kann (und in Systemen mit unendlich vielen Freiheitsgraden oft auch muß).

1.2. Physikalische Deutung

Einer der wichtigsten Begriffe der Physik ist der des Zustandes. Dieser Begriff ist so tief und von solch struktureller Vielfalt, daß eine abschließende, alle Fälle überdeckende Definition wohl kaum gegeben werden kann. Es erscheint am natürlichsten, den Zustandsbegriff mit dem nicht weniger fundamentalen der Observable zu verknüpfen: Der Zustandsbegriff wäre leer, gäbe es nicht Observable, die verschiedene Zustände zu unterscheiden gestatten.

Wir wollen hier diejenigen physikalischen Systeme betrachten, deren Zustände durch endlich-dimensionale Dichtematrizen einer festen Dimension n und deren Observable durch die hermitischen Matrizen (Operatoren) derselben Dimension beschreibbar sind. Ein solches System wollen wir kurz ein „endliches" physikalisches System nennen bzw. genauer ein „irreduzibles endliches physikalisches System", da wir voraussetzen wollen, daß *alle* hermitischen Operatoren Observable beschreiben.

Ist nun A ein hermitischer Operator und ϱ eine Dichtematrix, so ist der Erwartungswert $\varrho(A)$ der Observablen A im Zustand ϱ durch

$$\varrho(A) = \mathrm{Tr.}\, \varrho A \tag{1.27}$$

definiert. Da der Erwartungswert linear in A ist und jede Matrix Linearkombination aus hermitischen Matrizen, kann (1.27) formal auch als Definition des „Erwartungswertes" einer beliebigen Matrix (eines beliebigen Operators) A dienen. Dieser Regelung wollen wir uns im folgenden stets anschließen.

Lemma 1.4.

Die Erwartungswerte der hermitischen Operatoren A bestimmen die Dichtematrix eindeutig:

Aus $\varrho_1(A) = \varrho_2(A)$ für alle hermitischen A
folgt $\varrho_1 = \varrho_2$.

Aus der Voraussetzung folgt nämlich die Gleichheit des Erwartungswertes auch für die Matrix $A = \varrho_1 - \varrho_2$, und es folgt, daß die Spur des positiv semidefiniten Operators $(\varrho_1 - \varrho_2)^2$ verschwindet. Daher muß $\varrho_1 = \varrho_2$ sein. |=|

Nun wenden wir uns der weiteren Rechtfertigung von (1.27) zu. Nach den Axiomen der Quantentheorie ist der Erwartungswert das arithmetische Mittel der Meßwerte, die (bis auf den Meßfehler) wiederum mit Eigenwerten des zugeordneten hermitischen Operators identisch sein müssen. Hat daher ein hermitischer Operator keine negativen Eigenwerte und kann er als eine Observable des Systems aufgefaßt werden, so sollte sein Erwartungswert nicht negativ sein – unabhängig von dem Zustand des Systems, der bei der Messung vorlag.

Satz 1.1.

Sei jedem hermitischen Operator A eine reelle Zahl $\Phi(A)$ zugeordnet.

Genau dann existiert eine Dichtematrix ϱ mit

$$\Phi(A) = \mathrm{Tr.}\, A\varrho,$$

wenn die drei folgenden Bedingungen erfüllt sind:

a) $A \to \Phi(A)$ ist linear in A.

b) Es ist $\Phi(A) \geq 0$ für alle $A \geq 0$.

c) Φ ist normiert, d.h., $\Phi(\underline{1}) = 1$.

Bemerkung: Ist das physikalische System nicht in dem oben beschriebenen Sinn endlich, so stellen a) bis c) keine hinreichenden Bedingungen für die Existenz einer Dichtematrix dar.

Beweis: Ist ϱ eine Dichtematrix, so erfüllt der Erwartungswert (1.27) die Bedingungen a) bis c): a) folgt aus Eigenschaft c der Spur. b) folgt aus Eigenschaft e der Spur und der Voraussetzung $\varrho \geq 0$. Ebenso ist c) die Aussage $\mathrm{Tr.}\, \varrho = 1$ der Definition der Dichtematrizen. Wir zeigen nun, daß aus a) bis c) die Darstellung (1.27) folgt. In dem linearen Raum der hermitischen Operatoren wird durch

$$\langle A, B \rangle = \mathrm{Tr.}\, AB \tag{1.28}$$

ein Skalarprodukt eingeführt, so daß ein reeller Hilbert-Raum entsteht; denn $\langle A, B\rangle$ ist linear in beiden Variablen, und nach Eigenschaft d der Spur ist (1.28) symmetrisch in A und B. Schließlich ist $\langle A, A\rangle$ niemals negativ und genau dann gleich Null, wenn $A = 0$ ist (siehe Eigenschaft a der Spur). Da der Hilbert-Raum endlich dimensional ist, ist jedes lineare Funktional beschränkt. Nach dem Satz von *Riesz* gibt es daher genau eine hermitische Matrix σ mit

$$\Phi(A) = \langle A, \sigma \rangle \equiv \mathrm{Tr.}\, A\sigma. \tag{1.29}$$

Aus der Bedingung c) folgt, daß $\mathrm{Tr.}\, \sigma = 1$ sein muß. Aus der Eigenschaft der Spur folgt unmittelbar, daß die Bedingung b) mit $\sigma \geq 0$ identisch ist. Also ist σ eine Dichtematrix. |=|

Wir betrachten jetzt eine leichte Modifikation des eben bewiesenen Satzes. Hierzu bemerken wir, daß $A \geq 0$ genau dann gilt, wenn eine (mindestens eine) Darstellung $A = BB^*$ existiert. Daher ist die Bedingung b) des Satzes 1.1. mit

$$\Phi\,(BB^*) \geq 0 \quad \text{für alle} \quad B \tag{1.30}$$

identisch. Damit ist schon fast alles zur Begründung des nächsten Satzes gesagt:

Satz 1.2.

Jedem Operator A sei eine komplexe Zahl $\Phi(A)$ zugeordnet.
Genau dann existiert eine Dichtematrix ϱ mit

$$\Phi(A) = \text{Tr. } A\varrho,$$

wenn die folgenden drei Bedingungen erfüllt sind:

a) $A \to \Phi(A)$ ist linear.
b) $\Phi\,(BB^*) \geq 0$ für alle B.
c) $\Phi(1) = 1$.

Beweis: Aus der Linearität folgt, daß wir nach Beschränkung auf die hermitischen Operatoren wieder die Voraussetzungen von Satz 1.1. antreffen, falls Φ(A) für hermitische Operatoren reell ist. Nun ist Φ nicht-negativ (und somit trivialerweise reell) für die Matrix $(1 + \lambda A)(1 + \lambda A)^*$. Aus der Linearität, der Reellität von $\Phi\,(AA^*)$ und wegen $\Phi(1) = 1$ folgt, daß

$$\lambda\Phi\,(A) + \bar{\lambda}\Phi\,(A^*)$$

reell sein muß. Da λ beliebig komplex war, ist

$$\overline{\Phi(A)} = \Phi(A^*). \tag{1.31}$$

Insbesondere ist somit Φ reell für hermitische Argumente. |=|

Die Sätze 1.1. und 1.2. zeigen hinreichend deutlich, daß für *endliche* und *irreduzible* physikalische Systeme die Dichtematrizen eineindeutig den Zuständen des Systems zugeordnet sind. Daher identifizieren wir gelegentlich die Zustände mit den Dichtematrizen und sehen Operationen mit Dichtematrizen auch als Operationen mit Zuständen an, denen unmittelbare physikalische Bedeutung zukommt.
Die eben angegebene Festlegung des Zustandsbegriffs für endliche irreduzible Systeme umfaßt sowohl die sog. reinen als auch die sog. gemischten Zustände (Zustandsgemische). Historisch gesehen bildeten die „reinen Zustände" das Untersuchungsobjekt der Quantenmechanik, die gemischten Zustände dasjenige der Thermodynamik. Die reinen Zustände korrespondieren im klassischen Grenzfall zu den Punkten des Phasenraumes, die gemischten Zustände haben als klassischen Grenzfall die Wahrscheinlichkeitsverteilungen über dem Phasenraum. Eine solche einfache Aufteilung ist jedoch nicht immer sachlich gerechtfertigt (besonders bei Systemen mit unendlich vielen Freiheitsgraden). Beim sog. algebraischen Zugang zur Quantentheorie erscheint es im Gegenteil angebracht, zuerst – wie eben geschehen – den allgemeinen Zustandsbegriff zu geben und danach die reinen Zustände als Grenzfälle (siehe 2.) zu betrachten.
In Einklang mit diesen Bemerkungen geben wir folgende

Definition 1.2.

Die Gesamtheit der Dichtematrizen eines endlichen irreduziblen physikalischen Systems heißt *Zustandsraum* des Systems. Wir bezeichnen ihn mit Ω:

$$\Omega = \{\varrho : \varrho \geq 0,\ \text{Tr. } \varrho = 1\}. \tag{1.32}$$

Wollen wir hervorheben, daß es sich um $n \times n$-Matrizen handelt, das System also die Dimension n hat, so schreiben wir Ω_n.

1.3. Zusammenhang mit einem allgemeineren Zustandsbegriff

Um ein physikalisches System zu beschreiben, benötigt man zunächst zwei Begriffe, die sich gegenseitig bedingen: den Begriff des Zustandes und den der Observablen. Ein physikalisches System hat die Möglichkeit, verschiedene Zustände einzunehmen, und wir gewinnen so einen ersten Eindruck von seiner strukturellen Vielfalt. Andererseits setzt der Zustandsbegriff notwendigerweise den der Observablen voraus; denn die verschiedenen Resultate von Messungen observabler Größen unterscheiden die verschiedenen Zustände. Man kann jeden dieser beiden Begriffe zum Ausgangspunkt nehmen. Zum Beispiel kann man, wie es u.a. *Schrödinger*, *von Neumann* und *Dirac* taten, einen Hilbert-Raum $\mathfrak{H}$ einführen, dessen eindimensionale Teilräume zu den reinen Zuständen des in Rede stehenden physikalischen Systems eindeutig korrespondieren. Dann haben wir die selbstadjungierten Operatoren (oder eine geeignete Auswahl dieser Menge) als Observable zu interpretieren, und schließlich erweist es sich als notwendig, Zustandsgemische mit Hilfe von Dichtematrizen oder allgemeineren Objekten einzuführen.

Wir können jedoch auch den anderen Weg verfolgen und zunächst Observable einführen, zwischen denen algebraische Relationen (z.B. Vertauschungsregeln) bestehen, einen Weg, den zuerst *Heisenberg* einschlug. Dieser Weg ist einfacher bei der Handhabung des folgenden Problems: Wie besonders *Dirac* in seinen "Principles of Quantum Mechanics" herausstellte, ist es eine besondere Annahme, daß *jedem* selbstadjungierten Operator ein Meßverfahren zugeordnet werden kann, welches ihn als Observable ausweist. Dieses „Abschlußpostulat" wurde lange als eine Art Verzierung betrachtet, die den Axiomen der Quantenmechanik die innere Geschlossenheit gibt, ohne wesentliche praktische Konsequenzen zu besitzen. Eine solche Auffassung ist jedoch nicht berechtigt; denn wenn eine wesentlich „kleinere" Menge von Operatoren als Observable legitim ist, so können zwei (und mehr) linear unabhängige Vektoren des Hilbert-Raumes bezüglich dieser Observablenmenge stets die gleichen Erwartungswerte liefern. Sie sind dann durch diese Observablen nicht unterscheidbar und repräsentieren daher den gleichen Zustand! Es ist offensichtlich, daß solche Verallgemeinerungen leichter zu handhaben sind, wenn wir von den Observablen ausgehen und mit ihrer Hilfe beschreiben, was unter einem Zustand verstanden werden soll.

Eine sehr allgemeine Beschreibung von mathematischen Objekten, die geeignet sind, Observable darzustellen, hat *Segal* gegeben. Die bisherigen Erfahrungen zeigen jedoch, daß wesentlich stärkere Postulate notwendig sind, um den mathematischen Apparat aussagekräftig zu machen. Im folgenden besprechen wir nur kurz die algebraischen Postulate und zeigen ihre Verträglichkeit mit dem im Abschnitt 1.2. dargelegten.

Unter einer Algebra R verstehen wir einen komplexen linearen Raum R, in dem noch eine Multiplikation

$$a, b \to a \cdot b; \quad a, b \in R, \quad a \cdot b \in R \tag{1.33}$$

erklärt ist, für die die assoziativen und distributiven Gesetze

$$\begin{gathered}(ab)\, c = a\,(bc)\\ a\,(b_1 + b_2) = ab_1 + ab_2, \quad (b_1 + b_2)\, a = b_1 a + b_2 a\\ \lambda\,(ab) = (\lambda a)\, b = a\,(\lambda b), \quad \lambda \text{ komplexe Zahl}\end{gathered} \tag{1.34}$$

gelten. Wir wollen weiter voraussetzen, daß R ein Einselement $\underline{1}$ besitzt, für das gilt

$$\underline{1}a = a = a\underline{1} \quad \text{alle} \quad a \in R. \tag{1.35}$$

Eine solche Struktur erlaubt also eine Addition und eine Multiplikation, die wir hier als eine Verallgemeinerung der Addition und Multiplikation von Operatoren auffassen wollen. Bei der eingangs erwähnten Formulierung der Quantenmechanik über den Begriff des Hilbert-Raumes konnten wir nicht jeden, sondern nur hermitische (besser selbstadjungierte) Operatoren als observabel betrachten. Auch in der Algebra R benötigen wir eine solche Unterscheidung, wenn wir sie zur Charakterisierung von Observablen benutzen wollen. Diese Notwendigkeit führt auf die *-Algebren.
Eine Algebra R wird zu einer *-Algebra, wenn in ihr eine „Involution"

$$a \to a^* \tag{1.36}$$

definiert ist, die folgenden (nach dem Vorbild der hermitischen Konjugation aufgestellten) Regeln genügt:

$$\begin{aligned} (\lambda a)^* &= \bar{\lambda} a, \quad \lambda \text{ komplexe Zahl} \\ (a + b)^* &= a^* + b^* \\ (ab)^* &= b^* a^* \\ (1)^* &= 1. \end{aligned} \tag{1.37}$$

Ein Element a einer *-Algebra R heißt hermitisch, wenn es gleich a^* ist:

$$a \text{ hermitisch:} \quad a = a^* \tag{1.37a}$$

Postulat 1.

Die Observablen eines physikalischen Systems sind die hermitischen Elemente einer geeigneten *-Algebra.

Nun kommen wir zu den Zuständen. Ein Zustand wird charakterisiert durch den Ausgang möglicher Messungen, die man am System vornehmen kann. Einen Zustand kennen heißt, den Erwartungswert jeder Observablen in diesem Zustand angeben zu können (bzw. ein Verfahren, daß bei Meisterung aller mathematischen Schwierigkeiten auf das gesuchte Resultat führt). Wir können dies so ausdrücken: Ein Zustand Φ induziert eine Abbildung der Menge aller Observablen in die reellen Zahlen

$$\Phi : a \to \Phi(a), \quad a \text{ hermitisch,} \tag{1.38}$$

wobei $\Phi(a)$ der Erwartungswert von a im Zustand Φ ist. (1.38) charakterisiert den Zustand vollständig, da sonst zwei Zustände durch Observable nicht unterscheidbar wären. Wir können daher Zustände mit gewissen Funktionalen (1.38) auf der Menge der Observablen identifizieren. Natürlich kann nicht jedes Funktional auf der Menge der Observablen einem Zustand entsprechen. Bei der Aufgabe, diese Menge einzuschränken, helfen uns wieder die Verhältnisse, die wir aus der Quantenmechanik kennen und besonders auch Satz 1.2. des vorangegangenen Abschnittes. Ist z. B. b ein beliebiges Element der in Postulat 1. erwähnten *-Algebra R, so ist b^*b hermitisch und daher eine Observable. In Analogie zu den Operatoren eines Hilbert-Raumes betrachten wir b^*b darüber hinaus als „positiv semidefinit“ und nehmen an, daß die zugehörigen Meßwerte niemals negativ sein können, so daß der Erwartungswert von b^*b ebenfalls nicht negativ sein kann. Da wir auch die Linearität von (1.38) in den Observablen übernehmen, kann dieses Funktional eindeutig auf die ganze Algebra R ausgedehnt werden. Das alles legt uns die folgende mathematische Definition nahe:

Definition 1.3.

R sei eine *-Algebra mit Einselement.
Ein Zustand von R ist ein *lineares* Funktional über (dem linearen Raum) R, das folgende Bedingungen erfüllt:

a) Normierung: $\Phi(1) = 1$.
b) Positivität: $\Phi(a^*a) \geq 0$ für alle $a \in R$.

Bemerkung: Aus b) und der Existenz des Einselements folgt, daß Φ für hermitische Elemente reelle Werte annimmt.

Postulat 2.

Die Zustände eines physikalischen Systems werden durch die Zustände der in Postulat 1. geforderten *-Algebra beschrieben.

Bemerkung 1: Die Menge aller Operatoren unseres n-dimensionalen Hilbert-Raumes $\mathfrak{H}$ bzw. die Menge aller $n \times n$-Matrizen ist auf natürliche Weise eine *-Algebra $\mathfrak{L}_n$. Satz 1.2. zeigt, daß in diesem Fall die durch Definition 1.3. und Postulat 2. definierten Zustände mit den Erwartungswerten (1.27) von Dichtematrizen übereinstimmen. Die Verhältnisse im nichtendlichen Fall sind weit vielschichtiger.

Bemerkung 2: Das Postulat 2. gibt sofort „alle“ Zustände, d.h. die reinen und die gemischten. Die reinen Zustände können durch eine Verallgemeinerung des im Abschnitt 2. gesagten bestimmt werden.

Bemerkung 3: In den Postulaten fehlt die explizite Erwähnung des fundamentalen Überlagerungsprinzips. Dies wird erst durch die Darstellungen der *-Algebra als Algebra von Operatoren in einem Hilbert-Raum sichtbar. Die reinen Zustände sind in natürlicher Weise mit irreduziblen Darstellungen verknüpft.

Bemerkung 4: Die Berücksichtigung von Symmetrien, der Positivität von Massen- und Energiespektren und andere offensichtliche physikalische Forderungen wählen im allgemeinen aus der Gesamtheit der Zustände der *-Algebra eine Teilmenge als „physikalisch“ aus. Um den Grundgedanken auszudrücken, sind die Postulate vereinfacht dargestellt worden.

2. Die Konvexität des Zustandsraumes

Im Abschnitt 2.1. stellen wir einige mathematische Definitionen und Sätze bereit, die zum Verständnis der weiteren Überlegungen notwendig sind. Im Abschnitt 2.2. wird nämlich gezeigt, daß die Menge aller Dichtematrizen eine konvexe Menge bildet, der Zustandsraum Ω also konvex ist. Ω enthält neben den sog. reinen auch alle gemischten Zustände, eine Unterscheidung, die im Abschnitt 2.2. behandelt wird. Dieser Ab-

schnitt sagt somit etwas über die „Geometrie“ des Zustandsraumes aus und über das Problem, eine Dichtematrix (einen Zustand) nicht nur als spezielle Matrix, sondern als Element, als „Punkt“ mit einer speziellen Lage in diesem Raum Ω zu begreifen.
Für ein erstes Verständnis langt es, die mathematische Einführung bis zur Formel (2.7) zu kennen.

2.1. Mathematische Einführung

Wir betrachten im folgenden einen reellen linearen Raum $\mathfrak{L}$. (Für die weiteren Anwendungen denken wir z. B. an die Menge aller hermitischen Matrizen einer gegebenen Zeilen- und Spaltenzahl, $\mathfrak{L} = \{A : A = A^*\}$.)

Definition 2.1.

Sei $\mathfrak{M}$ eine beliebige Teilmenge von $\mathfrak{L}$. Ein Element x aus $\mathfrak{L}$ heißt genau dann konvex abhängig von $\mathfrak{M}$, wenn es eine Darstellung

$$x = \sum_{i=1}^{m} p_i x_i, \quad x_j \in \mathfrak{M} \tag{2.1}$$

besitzt, wobei $x_1, \ldots, x_m$ endlich viele Elemente aus $\mathfrak{M}$ sind und die reellen Zahlen $p_1, \ldots, p_m$ den Nebenbedingungen

$$p_j \geq 0, \quad \sum_i p_i = 1 \tag{2.2}$$

genügen. |=|

Eine Linearkombination (2.1), für deren Koeffizienten p_j die Nebenbedingungen (2.2) erfüllt sind, nennen wir eine „konvexe Linearkombination“. Das Element x in (2.1) kann als gewichtetes arithmetisches Mittel der x_i mit den Gewichten p_i gedeutet werden.
An einigen wenigen Stellen spielt die Zahl der Summanden in (2.1), für die $p_i \neq 0$ gilt, eine Rolle. Diese Zahl wird dann als Länge der konvexen Linearkombination bezeichnet.

△ Ein von einer Menge $\mathfrak{M}$ konvex abhängiges Element x hat in der Regel mehrere verschiedene Darstellungen der Gestalt (2.1) mit Nebenbedingung (2.2).

Definition 2.2.

Wir bezeichnen mit $[\mathfrak{M}]$ die Menge aller Elemente x aus $\mathfrak{L}$, die konvex von $\mathfrak{M}$ abhängen:

$$[\mathfrak{M}] = \{x \in \mathfrak{L} : x \quad \text{konvex abhängig von } \mathfrak{M}\}. \tag{2.3}$$

$[\mathfrak{M}]$ heißt konvexe Hülle von $\mathfrak{M}$. |=|

Die konvexe Hülle besitzt folgende Eigenschaft

$$[[\mathfrak{M}]] = [\mathfrak{M}]; \qquad \text{denn} \tag{2.4}$$

ist $x \in [[\mathfrak{M}]]$, so gibt es konvexe Linearkombinationen $x = \Sigma p_i y_i$ und $y_i = \sum_j p_{ij} x_i$ mit $y_i \in [\mathfrak{M}]$, $x_{ij} \in \mathfrak{M}$. Folglich ist $x = \sum p_i p_{ij} x_{ij}$ mit $\sum_i \sum_j p_i p_{ij} = \sum p_i = 1$, und daher ist $x \in [\mathfrak{M}]$. Aus $x \in [\mathfrak{M}]$ aber folgt $x \in [[\mathfrak{M}]]$ trivialerweise, denn x ist selbst konvexe Linearkombination der Länge 1 mit Gewicht 1. |=|

Definition 2.3.

Eine Menge $\mathfrak{M}$ heißt konvex, wenn $\mathfrak{M} = [\mathfrak{M}]$ ist.

Eine konvexe Menge ist also dadurch gekennzeichnet, daß sie alle von ihr konvex abhängigen Elemente enthält.
Bei der Untersuchung konvexer Mengen spielen die Extremalpunkte eine große Rolle, die wie folgt definiert sind.

Definition 2.4.

Sei $\mathfrak{K}$ eine konvexe Menge. Ein Element x heißt Extremalpunkt von $\mathfrak{K}$ genau dann, wenn es in $\mathfrak{K}$ liegt und wenn aus

$$x = px_1 + (1 - p)\, x_2, \quad x_1, x_2 \in \mathfrak{K}, \quad 0 < p < 1 \tag{2.5}$$

notwendig $x_1 = x_2$ folgt.

Die Extremalpunkte sind daher diejenigen Punkte (Elemente) einer konvexen Menge, die von ihr nur in trivialer Weise konvex abhängen. Aus der Definition folgt nämlich leicht, daß aus

$$x = \sum_{i=1}^{m} p_i x_i, \quad p_i > 0, \quad \sum p_i = 1, \quad x_i \in \mathfrak{K} \tag{2.6}$$

stets $x_1 = x_2 = \cdots = x_m$ folgt, wenn x ein Extremalpunkt ist. Wir bezeichnen die Menge aller Extremalpunkte mit ex $\mathfrak{K}$:

$$\text{ex}\, \mathfrak{K} = \{x \in \mathfrak{L} : x \text{ ist Extremalpunkt von } \mathfrak{K}\}. \tag{2.7}$$

Für eine im Abschnitt 7. besprochene Verallgemeinerung des Entropiebegriffes benötigen wir das

Lemma 2.1.

Für jede konvexe Menge $\mathfrak{K}$ ist

$$\text{ex}\, \mathfrak{K} = \cap\, \mathfrak{M} \quad \text{mit} \quad [\mathfrak{M}] = \mathfrak{K}. \tag{2.8}$$

Beweis: Ist $x \in \text{ex}\, \mathfrak{K}$, so ist $x \in \mathfrak{M}$; denn sonst wäre x eine nicht-triviale konvexe Linearkombination sogar von Elementen aus $\mathfrak{M}$ und damit aus $\mathfrak{K}$. Ist aber $x \notin \text{ex}\, \mathfrak{K}$, so ist x von der Menge

$$\mathfrak{M}_1 = \{y : y \in \mathfrak{K}, \quad y \neq x\}$$

konvex abhängig und $[\mathfrak{M}_1] = \mathfrak{K}$. Also kann x nicht im Durchschnitt aller Mengen $\mathfrak{M}$ mit $[\mathfrak{M}] = \mathfrak{K}$ liegen. |=|
Wir wollen nun unsere Aufmerksamkeit auf eine etwas engere Klasse konvexer Mengen lenken. In einem endlich-dimensionalen linearen Raum $\mathfrak{L}$ ist auf natürliche Weise ein Konvergenzbegriff (eine Topologie) erklärt: Eine Folge $x_1, x_2, x_3, \ldots$ von Elementen heißt konvergent mit Limes x, wenn für jedes lineare Funktional φ über $\mathfrak{L}$ die Zahlen-

folge $\varphi(x_i)$, $i = 1, 2, \dots$ gegen $\varphi(x)$ konvergiert. Ist $y_1, \dots, y_N$ eine linear unabhängige Basis von $\mathfrak{L}$ und ist

$$x_i = \sum_k \alpha_{ik} y_k; \quad x = \sum_k \alpha_k y_k, \tag{2.9}$$

so konvergiert $\{x_j\}, j = 1, 2, \dots$ genau dann gegen x, wenn

$$\lim_i \alpha_{ik} = \alpha_k \quad \text{für jedes } k \tag{2.10}$$

gilt.
Eine Teilmenge von $\mathfrak{L}$ heißt abgeschlossen, wenn sie mit jeder in ihr enthaltenen konvergenten Folge auch den Limes dieser Folge enthält.
Eine Teilmenge heißt beschränkt, wenn auf ihr jedes lineare Funktional beschränkt ist. $\mathfrak{M}$ ist genau dann beschränkt, wenn es zu einer Basis $y_1, y_2, \dots, y_N$ von $\mathfrak{L}$ eine Zahl $a > 0$ so gibt, daß aus

$$x = \Sigma\, \alpha_k y_k, \quad x \in \mathfrak{M}$$

stets $|\alpha_k| < a$ folgt.
Schließlich wollen wir noch vermerken, daß eine Menge $\mathfrak{M}$ eines endlich-dimensionalen linearen Raumes kompakt genannt wird, wenn sie sowohl abgeschlossen als auch beschränkt ist.

△ Die hier angegebene Methode, die Begriffe konvergent, abgeschlossen, beschränkt und kompakt zu definieren, ist nur im endlich-dimensionalen Fall sinnvoll. Für unendlich-dimensionale lineare Räume siehe z. B. [4].

Wir kommen nun zu einigen Eigenschaften konvexer Mengen, für deren Beweis wir auf die Literatur verweisen [14].

Lemma 2.2. (Carathéodory)

Ist dim $\mathfrak{L} = n$ und $\mathfrak{M} \subset \mathfrak{L}$ beliebig, so gibt es zu jedem von $\mathfrak{M}$ konvex abhängigen Element x

$$x \in [\mathfrak{M}] \tag{2.11}$$

$n + 1$ Elemente $x_1, \dots, x_{n+1}$ aus $\mathfrak{M}$ so, daß x bereits von $\{x_1, \dots, x_{n+1}\}$ konvex abhängt:

$$x \in [x_1, \dots, x_{n+1}] \subseteq [\mathfrak{M}]. \tag{2.12}$$

Die Aussagen von Lemma 2.2. können mit Hilfe des Kompaktheitsbegriffes noch etwas verschärft werden.

Lemma 2.3. (Minkowski, Carathéodory)

Ist $\mathfrak{K}$ eine kompakte konvexe Menge des n-dimensionalen reellen linearen Raumes $\mathfrak{L}$, so ist

$$\mathfrak{K} = [\mathrm{ex}\, \mathfrak{K}] \tag{2.13}$$

und zu jedem x aus $\mathfrak{K}$ gibt es $n + 1$ Extremalpunkte $x_1, \dots, x_{n+1}$ von $\mathfrak{K}$, von denen x konvex abhängt.

Einen eleganten Beweis findet man z.B. in [15].
In einem endlich-dimensionalen Raum gilt schließlich auch

Lemma 2.4. (Bourbaki)

Ist $\mathfrak{M}$ eine kompakte Menge eines endlich-dimensionalen linearen Raumes, so ist auch $[\mathfrak{M}]$ kompakt.

Hier ist noch eine Warnung angebracht:

⚠ Ist die konvexe Menge $\mathfrak{K}$ kompakt, so ist ex $\mathfrak{K}$ nicht notwendig kompakt. Ein einfaches Gegenbeispiel findet man bei [4 (I, §3)].

2.2. Die Konvexität der Menge aller Dichteoperatoren

Wir werden nun den Zustandsraum, d.h. die Menge Ω aller Dichtematrizen

$$\Omega \equiv \Omega_n = \{\omega : \omega \text{ ist Dichtematrix}\} \tag{2.14}$$

als Teilmenge des reellen linearen Raumes aller hermitischen Matrizen von n Zeilen und n Spalten betrachten. Dann gilt

Satz 2.1.

a) Ω ist konvex und kompakt.
b) ex Ω besteht aus allen Projektionsoperatoren der Dimension 1.
c) Es ist $[\text{ex}\, \Omega] = \Omega$.

Jedes Element $\omega \in \Omega$ ist als eine konvexe Linearkombination von Extremalpunkten von Ω darstellbar, deren Länge die Zahl n nicht übersteigt.

Beweis: 1. Aus $\varrho_1, \varrho_2 \in \Omega$, $1 \geq p \geq 0$ folgt $p\varrho_1 + (1-p)\varrho_2 = \varrho \geq 0$ und tr. $\varrho = 1$. Also ist Ω konvex. Ω ist auch abgeschlossen; denn konvergieren die Dichtematrizen $\omega_1, \omega_2, \ldots$ gegen ω, so ist für jedes $A \geq 0$ auch lim tr. $A\omega_i =$ tr. $A\omega \geq 0$ und daher ist $\omega \geq 0$. Setzen wir $A = 1$, so folgt auch die Normierungsbedingung für ω. Ω ist auch beschränkt; denn da die Eigenwerte einer Dichtematrix kleiner oder gleich Eins sind, gilt $(\eta, \eta) \geq (\eta, \omega\eta)$. Hieraus aber folgt mittels der Schwarzschen Ungleichung $|(\eta_1, \omega\eta_2)|^2 \leq (\eta_1, \eta_1)(\eta_2, \eta_2)$, und daher sind die Matrixelemente der Dichtematrizen gleichmäßig beschränkt. Als abgeschlossene und beschränkte Menge eines endlich-dimensionalen Raumes ist Ω somit kompakt.

2. Wir kommen nun zur Aussage b): Ein „eindimensionaler Projektionsoperator" ist eine hermitische Matrix ω, deren Wirkung auf einen beliebigen Vektor η mit Hilfe eines festen normierten Hilfsvektors η_0 wie folgt beschrieben werden kann:

$$\omega\eta = (\eta_0, \eta)\,\eta_0, \quad (\eta_0, \eta_0) = 1. \tag{2.15}$$

Wir sagen dann, ω projeziere auf η_0 oder genauer projeziere auf den von η_0 aufgespannten eindimensionalen Teilraum. Eine Darstellung (2.15) existiert genau dann, wenn $\omega = \omega^*$ und $\omega = \omega^2$ sowie tr. $\omega = 1$ ist. Wegen $(\eta, \omega\eta) = |(\eta_0, \eta)|^2 \geq 0$ ist dann auch $\omega \geq 0$ und somit $\omega \in \Omega$.
Sind $\lambda_1, \ldots, \lambda_n$ die Eigenwerte der Dichtematrix ω und sind $\omega_1, \ldots, \omega_n$ die Projektionsoperatoren auf die zugehörigen Eigenvektoren $\eta_1, \eta_2, \ldots, \eta_n$, für die wir

$$(\eta_i, \eta_k) = \delta_{ik} \tag{2.16}$$

voraussetzen können, so ist

$$\omega = \sum_{i=1}^{n} \lambda_i \omega_i . \tag{2.17}$$

Wegen $\lambda_i \geq 0$ und $\Sigma \lambda_j = 1$ ist (2.17) eine Darstellung von ω als konvexe Linearkombination, deren Länge höchstens gleich n ist. Die Darstellung (2.17) heißt „Spektraldarstellung" von ω. Satz 2.1. ist nun vollständig bewiesen, wenn wir noch zeigen können, daß ex Ω mit der Menge aller eindimensionalen Projektionsoperatoren zusammenfällt.

3. Nehmen wir nun an, $\omega \in \Omega$ sei ein eindimensionaler Projektor, der als konvexe Linearkombination

$$\omega = p\omega_1 + (1-p)\,\omega_2, \quad 0 < p < 1$$

zweier Dichtematrizen dargestellt sei. Wir zeigen $\omega_1 = \omega_2$. Ist nämlich $\omega\eta_0 = \eta_0$ und η zu η_0 orthogonal, so ist $\omega\eta = 0$. Es folgt daher wegen

$$p\,(\eta, \omega_1\eta) + (1-p)\,(\eta, \omega_2\eta) = 0$$

und der Positivität der Summanden $(\eta, \omega_i\eta) = 0$ und somit wegen $\omega_i \geq 0$ auch $\omega_i\eta = 0$. Daher ist $\omega_i = \mu_i\omega$, und die Bildung der Spur liefert $\mu_i = 1$, d.h. $\omega_1 = \omega_2 = \omega$. Daher liegt jeder eindimensionale Projektionsoperator in ex Ω. Nun zeigen wir noch die Umkehrung. Sei $\omega \in$ ex Ω und $\omega = \Sigma \lambda_i \omega_i$ die Spektralzerlegung von ω. Diese konvexe Linearkombination aber muß trivial sein, d.h. aus $\lambda_i \neq 0$ und $\lambda_k \neq 0$ muß $\omega_i = \omega_k$ folgen, sonst wäre ω kein Extremalpunkt. Da aber $\omega_i\omega_k = 0$ für $i \neq k$ nach Voraussetzung ist, darf ω nur einen von Null verschiedenen Eigenwert besitzen, der dann wegen tr. $\omega = 1$ gleich 1 sein muß. Also ist ω Projektionsoperator der Dimension 1. |=|

Bemerkung: Die Dimension des reell-linearen Raumes der hermitischen Matrizen ist gleich η^2. Aussage c) des Satzes 2.1. ist daher nicht aus Lemma 2.3. gewinnbar.

Definition 2.5.

Wenn eine Dichtematrix ϱ als konvexe Summe

$$\varrho = \sum_{i=1}^{m} p_i \varrho_i \tag{2.18}$$

anderer Dichtematrizen $\varrho_1, \varrho_2, \ldots, \varrho_m$ darstellbar ist, so sagen wir, ϱ sei eine Mischung der Dichtematrizen $\varrho_1, \ldots, \varrho_m$ mit Gewichten $p_1, \ldots, p_m$.

Aus (2.18) folgt sofort, daß der nach (1.27) definierte Erwartungswert einer Observablen A im Zustand ϱ das mit den Gewichten p_j genommene arithmetische Mittel der Erwartungswerte in den Zuständen ϱ_j ist:

$$\varrho(A) = \Sigma\, p_i \varrho_i\,(A), \quad \varrho_i(A) = \text{tr.}\, \varrho_i A . \tag{2.19}$$

Betrachten wir nun die Unschärfe (Schwankung, Fluktuation) der Observablen A im Zustand ϱ:

$$(\Delta_\varrho A)^2 = \varrho(A^2) - \varrho(A)^2 \equiv \text{Tr.}\, \varrho A^2 - (\text{Tr.}\, \varrho A)^2 . \tag{2.20}$$

Einfaches Nachrechnen ergibt die Regel

$$\Delta_\varrho (A + \underline{1} \cdot \lambda) = \Delta_\varrho A. \tag{2.21}$$

Vergleichen wir nun $\Delta_\varrho A$ mit den Schwankungen $\Delta_{\varrho_i} A$ falls (2.18) gilt. Wir haben dann

$$(\Delta_\varrho A)^2 = \Sigma\, p_i \,\mathrm{Tr.}\, A^2 \varrho_i - (\Sigma\, p_j \,\mathrm{Tr.}\, A\varrho_j)^2 = \Sigma\, p_i \{\mathrm{Tr.}\, A^2 \varrho_i - (\mathrm{Tr.}\, A\varrho_i)^2\}$$
$$+ \Sigma\, p_i (\mathrm{Tr.}\, A\varrho_i)^2 - (\Sigma\, p_i \,\mathrm{Tr.}\, \varrho_i A)^2$$

und hieraus folgt

$$(\Delta_\varrho A)^2 \geq \Sigma\, p_i (\Delta_{\varrho_i} A)^2. \tag{2.22}$$

Sind überdies alle p_j von Null verschieden, so tritt das Gleichheitszeichen genau dann ein, wenn Tr. ϱA = Tr. $\varrho_j A$ für alle j ist. Dies ergibt sich als Spezialfall der Ungleichung von *Jensen* für die konvexe Funktion x^2 (siehe Abschnitt 4.). Da weiter $\sqrt{x}$ eine konkave Funktion ist (siehe auch hier Abschnitt 4.), gilt $\sqrt{\Sigma p_i x_i} \geq \Sigma p_i \sqrt{x_i}$ und daher nach (2.22) die etwas schwächere Aussage

$$\Delta_\varrho A \geq \Sigma\, p_i \Delta_{\varrho_i} A. \tag{2.23}$$

Dabei gilt das Gleichheitszeichen nur dann, wenn es in (2.22) gilt und alle Schwankungen gleich sind ($p_i \neq 0$ vorausgesetzt). Wegen der physikalischen Bedeutung des Resultates wollen wir es nochmals zusammenfassen.

Satz 2.2.

Ist $\varrho = \Sigma p_i \varrho_i$ mit $p_i \neq 0$ eine Mischung von Dichtematrizen, so gilt für eine beliebige Observable $A = A^*$

a) $\Delta_\varrho A \geq \Sigma p_i \Delta_{\varrho_i} A$.
b) $(\Delta_\varrho A)^2 \geq \Sigma p_i (\Delta_{\varrho_i} A)^2$.

In b) gilt genau dann das Gleichheitszeichen, wenn

c) $\varrho(A) = \varrho_i(A)$ für alle i

gilt. In a) gilt genau dann das Gleichheitszeichen, wenn außer der Bedingung c) noch

d) $\Delta_\varrho A = \Delta_{\varrho_i} A$ für alle i

erfüllt ist.

Eine besondere Rolle spielen diejenigen Zustände ω, die nur triviale Darstellungen als konvexe Linearkombinationen (Mischungen) anderer Zustände besitzen. Nach Definition 2.4. sind diese gerade die Extremalpunkte des Zustandsraumes. Die Extremalpunkte heißen auch „reine Zustände" oder „reine Dichtematrizen bzw. -operatoren". Ist $\omega \in \mathrm{ex}\, \Omega$, so gibt es nach (2.15) und gemäß Satz 2.1. einen normierten Vektor ξ, $(\xi, \xi) = 1$, so daß

$$\omega\eta = (\xi, \eta)\, \xi \tag{2.24}$$

die Wirkung von ω auf den beliebig gewählten Vektor η angibt. Ist A eine Observable (ein beliebiger Operator), so ist der Erwartungswert von A im Zustand ω durch

$$\omega(A) = (\xi, A\xi) \tag{2.25}$$

gegeben. Da (2.24) gegen eine Phasenänderung von ξ invariant ist, entsprechen die Elemente von ex Ω gerade den eindimensionalen Teilräumen des Hilbert-Raumes $\mathfrak{H}$ und somit den durch $\mathfrak{H}$ darstellbaren Zuständen des physikalischen Systems.
Bemerkung 1: Die reinen Zustände wurden historisch als die primären Objekte der Quantenphysik angesehen, aus denen sich nach *J. von Neumann* die Zustandsgemische (durch Dichtematrizen beschreibbar) aufbauen. Vom Standpunkt des sog. algebraischen Zuganges zur Quantentheorie erscheint es natürlicher, die Gesamtheit der Zustände (reine und gemischte) als durch die Observablenalgebra definiert anzusehen. Die „geometrische" Struktur dieses Zustandsraumes bestimmt dann ganz zwangsläufig auch die reinen Zustände als „ungemischt" bzw. „nur in trivialem Sinne gemischt".

Bemerkung 2: Jede Dichtematrix (Zustand) ϱ kann als eine Mischung von reinen Dichtematrizen (reinen Zuständen) angesehen werden; denn Ω ist die konvexe Hülle von ex Ω. (2.17) gibt eine solche, physikalisch offenbar ausgezeichnete konvexe Linearkombination (Mischung) an, nämlich die Spektralzerlegung von ϱ. Man kann sich aber sehr leicht überzeugen, daß sich – mit Ausnahme der Extremalpunkte – jede Dichtematrix ϱ auf unendlich viele wesentlich verschiedene Weisen als konvexe Summe

$$\varrho = \Sigma\, p_i \omega_i \quad \text{mit} \quad \omega_i \in \text{ex}\, \Omega \tag{2.26}$$

schreiben läßt. Auf diesen Umstand hat besonders *Fano* [17] aufmerksam gemacht. Die ω_i sind dann natürlich Projektoren auf Vektoren, die zueinander *nicht* orthogonal sind.

△ Man kann also in der Quantentheorie einen Zustand im allgemeinen auch als eine Mischung von reinen Zuständen ansehen, die in der Spektralzerlegung überhaupt nicht vorkommen.

Dies trifft insbesondere für endliche irreduzible Systeme zu. Eine Übersicht über die bei einer solchen beliebigen Zerlegung auftretenden Gewichte p_i findet man im Teil II dieser Vorlesungen (Abschnitt 7.), wo auch die Spektralzerlegung einer Dichtematrix vor allen anderen Zerlegungen (2.26) durch eine Extremaleigenschaft charakterisiert wird.

3. Die Mischung von Zuständen

In diesem Kapitel wollen wir uns der Frage zuwenden, unter welchen Umständen eine Dichtematrix ϱ' „gemischter" („chaotischer", „weniger Information tragend") ist als eine andere ϱ. Es ist sehr bemerkenswert, daß dieses Problem eine eindeutige, nur von den inneren Eigenschaften des Zustandsraumes Ω abhängige Lösung besitzt, der wir uns im Abschnitt 3.1. zuwenden. Im Abschnitt 3.2. zeigen wir dann an Beispielen etwas vom physikalischen Sinn dieser Halbordnung: Ein Gibbssches Ensemble ist um so gemischter, je höher seine Temperatur ist. Weiter: Jeder Meßprozeß führt einen Zustand ϱ in einen Zustand ϱ' über, der gemischter ist als ϱ. Später werden wir u.a. sehen, daß die Entropie eines Zustandes ϱ' größer ist als die des Zustandes ϱ, wenn ϱ' gemischter ist als ϱ (ohne daß eine Umkehrung dieser Aussage gilt). Wir werden auch weitere Prozesse betrachten, bei denen der Mischungsgrad von Dichtematrizen erhöht wird. Im Abschnitt 3.3. wird die Definition „gemischter als“ dadurch verallgemeinert, daß an die Stelle der Gruppe aller unitären Transformationen kompakte Untergruppen treten. Der Inhalt dieses Abschnittes wird vor allem im Zusammenhang mit Abschnitt 8. benötigt und kann zunächst überschlagen werden.

3.1. Die Relation „gemischter als“

Wir wollen zunächst bemerken, daß eine Mischung

$$\varrho = p_1\varrho_1 + p_2\varrho_2 + \cdots, \quad p_1 > 0, \tag{3.1}$$

durchaus nicht im Widerspruch stehen muß mit einer konvexen Linearkombination (Mischung)

$$\varrho_1 = q_1\varrho + q_2\varrho_2' + \cdots, \quad q_1 > 0. \tag{3.2}$$

Daher ist es nicht richtig, aus einer beliebigen Darstellung (3.1) schon schließen zu wollen, ϱ sei in irgendeinem Sinne gemischter als ϱ_1.
Zu einer sinnvollen Festlegung kommt man aber dann, wenn man unitär äquivalente Dichtematrizen als „gleich stark gemischt“ ansieht und annimmt, daß durch Mischung unitär äquivalenter Dichtematrizen der „Mischungsgrad“ zunimmt. Wir werden im folgenden zeigen, daß auf diese Weise eine mathematisch und physikalisch sinnvolle Ordnungsrelation entsteht.

Definition 3.1.

ϱ' und ϱ seien zwei Dichtematrizen. ϱ' heißt gemischter als ϱ, in Zeichen

$$\varrho' \succ \varrho \tag{3.3}$$

genau dann, wenn ϱ' eine Mischung zu ϱ unitär äquivalenter Dichtematrizen ist, d.h., wenn

$$\varrho' = \Sigma\, p_j U_j \varrho U_j^{-1} \tag{3.4}$$

für gewisse unitäre Matrizen U_j und Zahlen $p_j > 0$ mit $\Sigma\, p_j = 1$ gilt.

Bemerkung: Anstelle von „ϱ' ist gemischter als ϱ“ sagen wir auch gelegentlich „ϱ' ist chaotischer als ϱ“ oder auch „ϱ trägt mehr Information als ϱ'“. Diese Sprechweise soll heuristisch andeuten, daß durch Mischung unitär äquivalenter Zustände die Ordnung des Systems, sein Gehalt an Struktur bzw. Information abnimmt und einem ungeordneteren, chaotischeren Zustand zustrebt.

Satz 3.1.

Die Relation „gemischter als“ ist eine Halbordnung der Klassen unitär äquivalenter Dichtematrizen:

a) Aus $\varrho_1 \succ \varrho_2$ und $\varrho_2 \succ \varrho_3$ folgt $\varrho_1 \succ \varrho_3$.
b) Es gilt $\varrho' \succ \varrho$ und $\varrho \succ \varrho'$ genau dann gemeinsam, wenn die Matrizen ϱ und ϱ' unitär äquivalent sind.

Beweis: Aus

$$\varrho_1 = \Sigma\, p_i U_i \varrho_2 U_i^{-1} \quad \text{und} \quad \varrho_2 = \Sigma\, q_j V_j \varrho_3 V_j^{-1}$$

folgt

$$\varrho_1 = \sum_{i,j} p_i q_j\, (U_i V_j)\, \varrho_3\, (U_i V_j)^{-1}$$

und dies ist wegen

$$\Sigma\, p_i q_j = 1, \quad p_i q_j \geq 0$$

und der Unitarität von $U_i V_j$ eine Mischung von zu ϱ_3 unitär äquivalenten Matrizen. Zum Beweis von b) ist klar, daß aus $\varrho_1 = U \varrho_2 U^{-1}$ mit unitärem U sowohl $\varrho_1 \succ \varrho_2$ als auch $\varrho_2 \succ \varrho_1$ folgt. Der Beweis der Umkehrung wird nach Satz 3.2. erbracht. Einen zweiten, völlig anderen Beweis findet man im Abschnitt 3.3. als Beweis von Satz 3.7.

Bemerkung: Der Begriff „Halbordnung" deutet an, daß zwischen zwei beliebigen Dichtematrizen ϱ_1 und ϱ_2 weder eine Relation $\varrho_1 \succ \varrho_2$ noch eine Relation $\varrho_2 \succ \varrho_1$ bestehen muß.

Folgerung: Ist ϱ_1 gemischter als ϱ_2, so ist jede zu ϱ_1 unitär äquivalente Dichtematrix gemischter als jede zu ϱ_2 unitär äquivalente Dichtematrix.

Satz 3.2.

Seien

$$\lambda_1 \geqslant \lambda_2 \geq \cdots \geq \lambda_n \quad \text{bzw.} \quad \mu_1 \geq \mu_2 \geq \quad \geq \mu_n \tag{3.5}$$

die sämtlichen Eigenwerte der Dichtematrizen ϱ bzw. ω. Es gilt $\varrho \succ \omega$ genau dann, wenn die Ungleichung

$$\sum_{i=1}^{j} \lambda_i \leq \sum_{i=1}^{j} \mu_i \tag{3.6}$$

für jedes j, $1 \leq j \leq n$ gilt.

Bemerkung 1: Aus Satz 3.2. folgt die Aussage b) des Satzes 3.1.; denn $\varrho \succ \omega$ und $\omega \succ \varrho$ zieht die Gültigkeit des Gleichheitszeichens in (3.6) und daher $\lambda_i = \mu_i$ für alle i nach sich. Matrizen mit gleichen Eigenwerten sind aber unitär äquivalent.

Bemerkung 2: Wegen $\mathrm{Tr.}\,\varrho = \mathrm{Tr.}\,\omega = 1$ gilt in (3.7) für $j = n$ das Gleichheitszeichen.
Um Satz 3.2. beweisen zu können, benötigen wir einige neue Hilfsmittel, die über die Zwecke des Beweises hinaus wichtig sind.
In Anlehnung an den (noch unbewiesenen) Satz 3.2. wollen wir folgende Sprechweise einführen:

Definition 3.2.

$$a_1 \geq a_2 \geq \cdots \geq a_n \quad \text{und} \quad b_1 \geq b_2 \geq \cdots \geq b_n \tag{3.7}$$

seien zwei Folgen reeller Zahlen mit

$$\sum_i a_i = \sum_i b_i. \tag{3.8}$$

Genau dann schreiben wir

$$\{a_i\} \succ \{b_j\} \tag{3.9}$$

und sagen $\{a_i\}$ sei gemischter als $\{b_j\}$, wenn gilt

$$\sum_{i=1}^{j} a_i \leqslant \sum_{i=1}^{j} b_i \quad \text{für} \quad j = 1, 2, \ldots, n. \tag{3.10}$$

Auch hier sagen wir gelegentlich „chaotischer“ oder „weniger Information tragend“ anstelle von „gemischter“.
Für den uns interessierenden Fall, in dem $a_i \geq 0, b_j \geq 0$ und $\Sigma\, a_i = \Sigma\, b_j = 1$ ist, kann man die a_i und b_j als Wahrscheinlichkeiten z. B. für den Ausgang gewisser Experimente deuten. Ist dann $\{a_i\} \succ \{b_j\}$, so liefert das auf $\{b_j\}$ führende Experiment nach jedem vernünftigen Informationsmaß eine größere Information als das auf $\{a_i\}$ führende. In diesem Sinne könnte man auch sagen, daß die Folge $\{b_j\}$ „absolut“ mehr Information trägt als $\{a_i\}$. Die Berechtigung für diese Behauptung liefert im wesentlichen Lemma 3.3.

Beispiel 3.1.
Ist $p_i \geq 0$ und $\Sigma p_i = 1$, so ist

$$\{1, 0, 0, \ldots, 0\} \prec \{p_1, p_2, \ldots, p_n\} \prec \left\{\frac{1}{n}, \frac{1}{n}, \ldots, \frac{1}{n}\right\}. \tag{3.11}$$

Bei der Folge $\{1, 0, 0, \ldots, 0\}$ tritt ein Ereignis mit Sicherheit ein, bei der Folge $\{1/n, \ldots, 1/n\}$ jedes mit der gleichen Wahrscheinlichkeit und ein Informationsgewinn findet nicht statt. |=|
Als nächstes führen wir den Begriff der bistochastischen Matrizen, die in der Literatur gelegentlich auch „Schursche Transformationen“ genannt werden, ein.

Definition 3.3.
Eine Matrix (t_{ij}) heißt genau dann bistochastisch, wenn

a) $t_{ij} \geq 0$ für alle i, j,
b) $\sum_i t_{ij} = 1$,
c) $\sum_j t_{ij} = 1$,

gilt.
Eine bistochastische Matrix ist stets quadratisch; denn die Gesamtsumme ihrer Elemente liefert sowohl die Zeilen- als auch die Spaltenanzahl. Wir sehen weiter, daß das Produkt

$$\sum_k t_{ik}\tilde{t}_{ks} = \hat{t}_{is}$$

zweier bistochastischer Matrizen wieder bistochastisch ist. Spezielle bistochastische Matrizen sind die Permutationsmatrizen. Eine Permutationsmatrix ist eine quadratische Matrix, bei der in jeder Zeile und in jeder Spalte genau eine 1 steht und sonst Nullen. Bei der Anwendung einer Permutationsmatrix auf einen Vektor werden dessen Komponenten permutiert. Ein einfaches Beispiel ist

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} b \\ a \\ c \end{pmatrix}.$$

Lemma 3.1.

Jede konvexe Linearkombination bistochastischer Matrizen ist wieder eine bistochastische Matrix.

Es ist nämlich

$$t_{rs} = \sum_i p_i t_{rs}^{(i)} \geq 0$$

und z. B.

$$\sum_s t_{rs} = \sum_i p_i \sum_s t_{rs}^{(i)} = \sum p_i = 1 .$$

Lemma 3.2. (Birkhoff)

Jede bistochastische Matrix ist konvexe Linearkombination von Permutationsmatrizen.

Wegen Lemma 2.3. kann dem Satz von *Birkhoff* auch folgende Form gegeben werden: „Die Extremalpunkte der konvexen Menge aller bistochastischen Matrizen gegebener Zeilenzahl sind die Permutationsmatrizen."
Wir können nun eine erste Verbindung mit dem Anliegen von Satz 3.2. herstellen.

Lemma 3.3. (Ostrowski, Karamata, Hardy, Littlewood, Polya)

Sei

$$a_1 \geq a_2 \geq \cdots \geq a_n, \quad b_1 \geq b_2 \geq \cdots \geq b_n \tag{3.12}$$

und

$$\Sigma\, a_i = \Sigma\, b_j . \tag{3.13}$$

Genau dann gilt

$$\{a_i\} \succ \{b_j\}, \tag{3.14}$$

wenn eine bistochastische Matrix (t_{ij}) mit

$$a_i = \sum_j t_{ij} b_j \tag{3.15}$$

existiert.

Nun kommen wir zum
Beweis von Satz 3.2.: Seien also ϱ und ω zwei Dichtematrizen, und ihre Eigenwerte seien wie in (3.6) angegeben geordnet.

a) (3.7) ist hinreichend für $\varrho \succ \omega$.

Nach Lemma 3.3. gibt es eine bistochastische Matrix (t_{ik}) mit

$$\lambda_i = \sum_k t_{ik}\, \mu_k$$

und nach Lemma 3.2. gibt es Permutationsmatrizen $t^j = (t^j_{ik})$ so, daß

$$t_{ik} = \Sigma\, p_j t^j_{ik}, \quad \Sigma\, p_j = 1, \quad p_j \geq 0$$

gilt.

Sei

$$\sum_k t^j_{ik}\mu_k = \mu_{j_i},$$

wobei $i \to j_i$ die durch t^j_{ik} bestimmte Permutation ist. Dann gilt

$$\lambda_i = \Sigma\, p_j \mu_{j_i}.$$

Es gibt daher Matrizen U_j mit

$$U_j \omega U_j^{-1} = \begin{pmatrix} \mu_{j_1} & & 0 \\ & \mu_{j_2} & \\ 0 & & \ddots \end{pmatrix}$$

und daher ist

$$\Sigma\, p_j U_j \omega U_j^{-1} = \begin{pmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & & \ddots \end{pmatrix}$$

eine Matrix mit Eigenwerten $\lambda_1, \lambda_2, \ldots, \lambda_n$, die daher zu ϱ unitär äquivalent ist. Also ist $\varrho \succ \omega$.

b) (3.6) ist notwendig für $\varrho \succ \omega$.
Sei $\xi_1, \ldots, \xi_n$ ein Orthonormalsystem von Eigenvektoren von ϱ mit $\varrho\xi_i = \lambda_i \xi_i$. Aus

$$\varrho = \Sigma\, p_j U_j \omega U_j^{-1},$$

(s. u.) folgt dann

$$\lambda_1 + \lambda_2 + \cdots + \lambda_k = \sum_j p_j \sum_{i=1}^{k} (U_j^{-1}\xi_i, \omega U_j^{-1}\xi_i).$$

$U_j^{-1}\xi_1, \ldots, U_j^{-1}\xi_n$ ist wieder ein Orthonormalsystem und nach dem Satz von *Ky Fan* folgt

$$\mu_1 + \mu_2 + \cdots + \mu_k \geq \sum_{i=1}^{k} (U_j^{-1}\xi_i, \omega U_j^{-1}\xi_i).$$

Wegen $\Sigma\, p_i = 1$ ist daher

$$\lambda_1 + \cdots + \lambda_k \leq \mu_1 + \cdots + \mu_k$$

und der Beweis ist beendet. |=|
Wir tragen noch das Resultat von *Ky Fan* nach:

Lemma 3.4. (Ky Fan)

A sei eine hermitische Matrix mit Eigenwerten $a_1 \geq a_2 \geq \cdots \geq a_n$. Ist dann $(\xi_i, \xi_j) = \delta_{ij}$ für $i, j = 1, \ldots, k$, so gilt

$$a_1 + a_2 + \cdots + a_k \geq \sum_{i=1}^{k} (\xi_i, A\xi_i). \tag{3.16}$$

Zum Schluß dieses Abschnittes geben wir noch eine aus den obigen Beweisen folgende äquivalente Form zu Satz 3.2. an:

Satz 3.2.a.

Seien $\{\lambda_i\}$ die (irgendwie geordneten) Eigenwerte der Dichtematrix ϱ und $\{\mu_i\}$ die Eigenwerte von ω. Genau dann ist

$$\varrho \succ \omega,$$

wenn es eine bistochastische Matrix (t_{ij}) mit

$$\lambda_i = \sum_j t_{ij}\,\mu_j$$

gibt.

3.2. Beispiele

Beispiel 3.2.

Die Beziehung (3.11) können wir auf Grund von Satz 3.2. als Aussage über Dichtematrizen auffassen.
Für $\varrho \in \Omega$ gilt

$$\frac{1}{n}\,1 \succ \varrho \succ \omega \quad \text{alle} \quad \omega \in \text{ex}\,\Omega. \tag{3.17}$$

Jeder Zustand ϱ ist also gemischter als jeder reine Zustand. Andererseits ist der durch die Dichtematrix $(1/n)\,\underline{1}$ gegebene Zustand gemischter als jeder beliebige Zustand. $(1/n)\,\underline{1}$ können wir daher als den „vollständig chaotischen Zustand" oder als „totales Chaos" bezeichnen. (Ein solcher Zustand existiert bei nicht-endlichen Systemen i.a. nicht.) Das totale Chaos ist vollständig strukturlos, da der Erwartungswert jeder Observablen in diesem Zustand das arithmetische Mittel seiner Eigenwerte, also seiner möglichen Meßwerte ist. Jeder beliebige Meßwert einer jeden Observablen besitzt dann die gleiche a priori Wahrscheinlichkeit. In diesem Sinne ist $(1/n)\,\underline{1}$ auch der „homogenste, gleichförmigste" aller Zustände und, da $U\,\underline{1}(1/n)\,U^{-1} = (1/n)\,\underline{1}$ für jede unitäre Matrix ist, auch der „maximal symmetrischste".

Beispiel 3.3.

Seien A und $\bar{A}$ zwei hermitische Matrizen mit Eigenwerten

$$A_1 \geq A_2 \geq \cdots \quad \text{bzw.} \quad \bar{A}_1 \geq \bar{A}_2 \geq \cdots \bar{A}_n. \tag{3.18}$$

Wir betrachten dann die Dichtematrizen

$$\varrho = \frac{e^A}{\text{tr.}\, e^A} \quad \text{und} \quad \bar{\varrho} = \frac{e^{\bar{A}}}{\text{tr.}\, e^{\bar{A}}} \tag{3.19}$$

und fragen nach Kriterien für $\varrho \succ \bar{\varrho}$. Zu diesem Zweck betrachten wir die Größen

$$a_m = \sum_{i=1}^{m} e^{A_i} \tag{3.20}$$

$$b_m = \sum_{i=m+1}^{n} e^{A_i} \tag{3.21}$$

und die analogen Bildungen $\bar{a}_m$, $\bar{b}_m$ bezüglich $\bar{A}$. Die Summe der m größten Eigenwerte von ϱ ist offenbar

$$(\text{Tr. } e^A)^{-1} \sum_{i=1}^{m} e^{A_i} = \frac{a_m}{a_m + b_m}. \tag{3.22}$$

Nehmen wir nun

$$\Delta_1 \leq \Delta_2 \leq \cdots \leq \Delta_n \quad \text{für} \quad A_i = \bar{A}_i + \Delta_i \tag{3.23}$$

an. Dann gilt

$$a_m = \sum_{i=1}^{m} e^{\bar{A}_i + \Delta_i} \leq e^{\Delta_m} \bar{a}_m \tag{3.24}$$

und

$$b_m = \sum_{i=m+1}^{n} e^{\bar{A}_i + \Delta_i} \geq e^{\Delta_{m+1}} \bar{b}_m \geq e^{\Delta_m} \bar{b}_m. \tag{3.25}$$

Also ist

$$b_m / a_m \geq \bar{b}_m / \bar{a}_m$$

und somit

$$\left(\frac{a_m}{a_m + b_m} \right) = \left(1 + \frac{b_m}{a_m} \right)^{-1} \leq \left(1 + \frac{\bar{b}_m}{\bar{a}_m} \right)^{-1} = \left(\frac{\bar{a}_m}{\bar{a}_m + \bar{b}_m} \right). \tag{3.26}$$

Nach Satz 3.2. bedeutet dies wegen (3.22), daß ϱ gemischter ist als $\bar{\varrho}$. Damit ist bewiesen

Satz 3.3.

Seien A und B hermitische Matrizen mit Eigenwerten

$$\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_n \quad \text{und} \quad \beta_1 \geq \beta_2 \geq \cdots \geq \beta_n. \tag{3.27}$$

Gilt dann

$$\alpha_i - \alpha_{i+1} \leq \beta_i - \beta_{i+1} \text{ für alle } i, \tag{3.28}$$

so ist

$$\frac{e^A}{\text{Tr. } e^A} \succ \frac{e^B}{\text{Tr. } e^B}. \tag{3.29}$$

Als Anwendung betrachten wir eine Gibbssche mikrokanonische Gesamtheit, die durch die Ersetzung

$$A = -\frac{1}{kT_1} H \quad \text{bzw.} \quad \bar{A} = -\frac{1}{kT_2} H \tag{3.30}$$

entsteht.

Hat H die Eigenwerte $E_1 \leq E_2 \leq \ldots$, so ist für $T_1 > 0$

$$A_i = -\frac{1}{kT_1} E_i \quad \text{und} \quad A_i - A_{i+1} = \frac{1}{kT_1} (E_{i+1} - E_i).$$

Daher tritt (3.28) für $T_1 > T_2 > 0$ ein. Analoges findet man für negative Temperaturen $0 > T_2 > T_1$. Der Zustand mit der betragsmäßig größeren Temperatur erweist sich in beiden Fällen als der gemischtere.

Satz 3.4.

Sei mit hermitischem H

$$\varrho(T) = (\text{Tr. } e^{-\beta H})^{-1}\, e^{-\beta H} \quad \text{mit} \quad \beta = \frac{1}{kT}. \tag{3.31}$$

Dann folgt aus

$$+0 \leq T_1 \leq T_2 \quad \text{oder} \quad -0 \geq T_1 \geq T_2 \tag{3.32}$$

stets

$$\varrho(T_1) \prec \varrho(T_2). \tag{3.33}$$

Durch die obigen Abschätzungen können Zustände positiver Temperatur mit Zuständen negativer Temperatur *nicht* bezüglich ihres Mischungsgrades verglichen werden; denn dann stehen die Folgen $\{-\beta_1 E_i\}$ und $\{-\beta_2 E_i\}$ stets zur Voraussetzung (3.27) im Widerspruch.

Zur Veranschaulichung können wir Ω durch einen Kreis in der Ebene symbolisieren, dessen Rand gerade die reinen Zustände sind. Für $T = -0$ und $T = +0$ liefert dann die Kurve $T \to \varrho(T)$ reine Zustände (zum größten und zum kleinsten Eigenwert, die wir als nicht-entartet voraussetzen). $T = \infty$ entspricht dem totalen Chaos (Abb. 1).

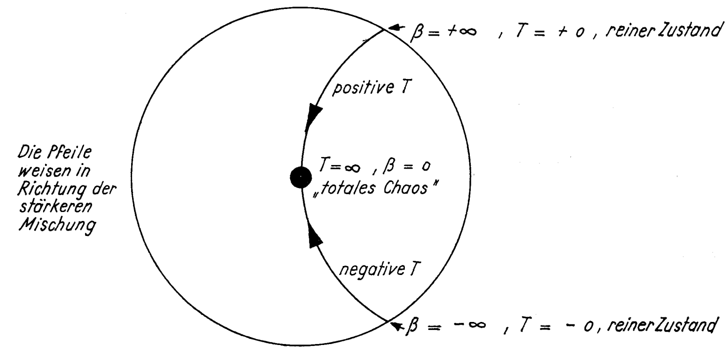

Abb. 1. Symbolische Darstellung der „Kurve" $T \to \varrho(T) \in \Omega$

Beispiel 3.4.

Nach *J. von Neumann* ist das Resultat der Ausführung einer vollständigen Messung mit kommutierenden Observablen $A_1, A_2, \ldots, A_s$ an einem Zustand ϱ ein neuer Zustand ϱ', der sich wie folgt beschreiben läßt: Ist $\xi_1, \ldots, \xi_n$ ein vollständiges Orthonormalsystem von gemeinsamen Eigenvektoren der Operatoren $A_1, \ldots, A_s$, so ist ϱ' durch

$$(\xi_i, \varrho' \xi_j) = \delta_{ij}\, (\xi_i, \varrho \xi_j) \tag{3.34}$$

gegeben. Bezogen auf die Basis $\xi_1, \ldots, \xi_n$ werden also in der Dichtematrix alle nicht in der Diagonale stehenden Elemente gleich Null gesetzt. Heuristisch ist klar, daß gilt

Satz 3.5.

$$\varrho' = \begin{pmatrix} \alpha_{11} & & 0 \\ & \alpha_{22} & \\ 0 & & \ddots \end{pmatrix} \succ \varrho = (\alpha_{ik}) \tag{3.35}$$

für jede Dichtematrix ϱ.

Denn ϱ' trägt offenbar weniger Information (Struktur) als ϱ, da u.a. eine Reihe von relativen Phasen „gelöscht" sind. ϱ' sollte also chaotischer, gemischter sein als ϱ. Wegen Satz 3.5. kann man obigen Satz 3.2. auch so formulieren:

Sei $\varrho = (\alpha_{ik})$ eine Dichtematrix mit Eigenwerten $\lambda_1 \geq \lambda_2 \geq \ldots$, dann gilt

$$\lambda_1 + \lambda_2 + \cdots + \lambda_k \geq \alpha_{11} + \alpha_{22} + \cdots + \alpha_{kn} \text{ für alle } k. \tag{3.36}$$

Das ist nichts als eine andere Leseart des Lemmas 3.4. von *Ky Fan*, wodurch bereits die Richtigkeit der Behauptung folgt. Wir wollen jedoch noch eine etwas allgemeinere Aussage beweisen und bemerken zuvor, daß

$$\varrho' = \sum_{i=1}^{n} (\mathrm{Tr.}\, \varrho\pi_i)\, \pi_i \tag{3.37}$$

gilt, falls $\pi_1, \pi_2, \ldots, \pi_n$ die Projektoren auf die gemeinsamen Eigenvektoren der Observablen $A_1, A_2, \ldots, A_s$ sind.
Wegen (3.37) enthält die folgende Aussage Satz 3.5. als einen speziellen Fall.

Satz 3.6.
Seien $B_1, \ldots, B_m$ positiv halb-definite Matrizen

$$B_i \geq 0 \quad \text{mit} \quad \sum_{j=1}^{m} B_j = 1, \tag{3.38}$$

dann ist mit ϱ auch

$$\varrho' = \sum_{j=1}^{m} (\mathrm{Tr.}\, \varrho B_j) \frac{B_j}{\mathrm{Tr.}\, B_j} \tag{3.39}$$

eine Dichtematrix und

$$\varrho' \succ \varrho. \tag{3.40}$$

Beweis: Wegen $B_i \geq 0$ und der Positivität der Koeffizienten in (3.39) folgt $\varrho' \geq 0$. Wegen

$$\mathrm{Tr.}\, \varrho' = \Sigma\, \mathrm{Tr.}\, (\varrho B_i) = \mathrm{Tr.}\, \varrho = 1$$

ist somit ϱ' eine Dichtematrix. Seien nun $\xi_1, \ldots, \xi_n$ und $\eta_1, \ldots, \eta_n$ vollständige Orthonormalsysteme von Eigenvektoren von ϱ und ϱ'

$$\varrho\xi_i = \lambda_i\xi_i, \quad \varrho'\eta_j = \lambda'_j\eta_j.$$

Wegen

$$\mathrm{Tr.}\, \varrho B_k = \sum_{i=1}^{n} \lambda_i\, (\xi_i, B_k\xi_i)$$

folgt aus (3.39)

$$\lambda'_j = \sum_{i=1}^{m} (\mathrm{Tr.}\, \varrho B_i) \frac{(\eta_j, B_i, \eta_j)}{\mathrm{Tr.}\, B_i} = \sum_{i,k} \lambda_k \frac{(\xi_k, B_i\xi_k)(\eta_j, B_i\eta_j)}{\mathrm{Tr.}\, B_i}.$$

Nach Satz 3.2.a. ist es hinreichend zu zeigen, daß

$$\alpha_{kj} = \sum_i \frac{(\xi_k, B_i\xi_i)(\eta_j, B_i\eta_j)}{\mathrm{Tr.}\, B_i} \tag{3.41}$$

eine bistochastische Matrix (α_{kj}) definiert.
In der Tat sind wegen $B_i \geq 0$ die Matrixelemente α_{kj} reell und nicht-negativ. Die Summation über k ist analog der über j:

$$\sum_k \alpha_{kj} = \sum_i \frac{(\mathrm{Tr.}\, B_i)(\eta_j B_i \eta_j)}{\mathrm{Tr.}\, B_i} = (\eta_j, \eta_j) = 1.$$

|=|

3.3. Eine Verallgemeinerung

Eine natürliche Verallgemeinerung der Definition 3.1. von „gemischter als" entsteht dadurch, daß wir in der Gleichung (3.4) nicht mehr beliebige unitäre Matrizen zulassen, sondern fordern, daß sie einer gewissen Gruppe angehören. Eine weitere, wegen des linearen Charakters in ϱ der Zuordnung (3.4) nicht so schwerwiegende Verallgemeinerung entsteht noch dadurch, daß wir die Beschränkung auf Dichtematrizen aufheben und beliebige Matrizen betrachten. Wir beschreiben in diesem Abschnitt weiter die Mittelung von Matrizen über einer kompakten Gruppe unitärer Operatoren, auf die wir im Teil II (Abschnitt 8.) wieder zu sprechen kommen. Insgesamt dienen die Betrachtungen dieses Abschnittes der Vorbereitung auf die Symmetrien und die „Teilsysteme" von physikalischen Systemen der hier betrachteten Art.
Sei G eine kompakte Gruppe unitärer Matrizen.

Definition 3.4.

Seien A und B zwei Matrizen. Wir sagen, daß A G-gemischter als B sei und schreiben

$$A \overset{G}{\succsim} B \tag{3.42}$$

genau dann, wenn A darstellbar ist als konvexe Summe

$$A = \Sigma\, p_i U_i^{-1} B U_i \quad \text{mit} \quad U_j \in G \tag{3.43}$$

mit $p_i \geq 0$ und $\Sigma p_j = 1$.

Bemerkung: Aus $A \overset{G}{\succsim} B$ folgt natürlich $A \succ B$, wobei $A \succ B$ den Fall bezeichnet, daß die zugrunde gelegte Gruppe *alle* unitären Matrizen umfaßt.

Satz 3.7.

Aus $A \overset{G}{\succsim} B$ und $B \overset{G}{\succsim} C$ folgt $A \overset{G}{\succsim} C$.

Es ist $A \overset{G}{\succsim} B$ und $B \overset{G}{\succsim} A$ gleichzeitig genau dann, wenn $A = U^{-1} BU$ mit $U \in G$ ist.

So wie Definition 3.4. die Definition 3.1. verallgemeinert, ist Satz 3.7. eine Verallgemeinerung von Satz 3.1.

Beweis von Satz 3.7.: Der Beweis der ersten Behauptung benutzt nur die Gruppeneigenschaft von G und ist dem von Satz 3.1. genau analog. Für die zweite Behauptung geben wir einen anderen Beweis. Wir stellen zunächst fest, daß $A \overset{G}{\succ} B$ genau dann gilt, wenn A von der Menge

$$N = \{C : C = U^{-1} BU, \quad U \in G\} \tag{3.44}$$

konvex abhängt:

$$A \subseteq [N] \quad \text{genau dann, wenn} \quad A \overset{G}{\succ} B. \tag{3.45}$$

Da G kompakt ist, ist N kompakt und daher nach Lemma 2.4. auch die konvexe Hülle $[N]$. Es folgt, daß $[N]$ von ex $[N]$ konvex erzeugt wird (Lemma 2.3.). Nach Lemma 2.1. ist ex $[N] \subseteq N$. Nun ist aber $U^{-1}[N]U = [N]$ für alle $U \in G$, und daher ist auch ex $[N]$ gegen die unitären Transformationen aus G invariant. Die einzige nicht-leere, bezüglich G invariante Teilmenge von N ist aber offensichtlich N selbst. Also ist ex $[N] = N$. Dieses Resultat ist für sich interessant, und wir fassen es zusammen in

Lemma 3.5.

Die Menge

$$\{A : A \overset{G}{\succ} B\}$$

ist kompakt und konvex. Ihre Extremalpunkte sind genau die Matrizen

$$U^{-1}BU \quad \text{mit} \quad U \in G.$$

Ist nun analog

$$M = \{C : C \overset{G}{\succ} A\},$$

so folgt, daß $M \subseteq [N]$ und $[M] \subseteq [N]$ ist. Also folgt aus dem gleichzeitigen Bestehen der Relationen $A \overset{G}{\succ} B$ und $B \overset{G}{\succ} A$ die Beziehung $[M] = [N]$. Nach Lemma 3.5. ist dann wegen ex $[M]$ = ex $[N]$ auch $M = N$ und somit $A = U^{-1}BU$ für ein $U \in G$. Die Umkehrung dieser Aussage aber ist trivial, und somit ist Satz 3.7. bewiesen. |=|

Lemma 3.6.

Ist

$$p_i \geq 0, \quad \sum_{i=1}^{\infty} p_i = 1, \tag{3.46}$$

so ist für beliebige $U_i \in G$

$$A = \sum_{i=1}^{\infty} p_i U_i B U_i^{-1} \succ B. \tag{3.47}$$

Beweis: Die Folge der durch

$$\left(\sum_{i=1}^{m} p_i\right) A_m = \sum_{i=1}^{m} p_i U_i B U_i^{-1}$$

definierten Matrizen A_m konvergiert gegen A, und es ist $A_m \overset{G}{\succ} B$. Nach Lemma 3.5. ist dann auch der Limes dieser Folge G-gemischter als B. |=|

Lemma 3.7.
Ist dv ein Wahrscheinlichkeitsmaß über der Gruppe G, so gilt für alle Matrizen B

$$\int_G U^{-1}BU\, dv(U) \succ B. \tag{3.48}$$

Beweis: Ein Wahrscheinlichkeitsmaß über G ist ein Maß, das auf 1 normiert ist und nirgends negativ ist: Ist $Q \subseteq G$ eine Borelsche Teilmenge, so gilt

$$\int_Q dv(U) \geq 0 \quad \text{und besonders} \quad \int_G dv(U) = 1. \tag{3.49}$$

Das Integral (3.48) ist ein Limes von Summen G der Gestalt

$$\Sigma\, U_j^{-1}BU_j \cdot \int_{Q_j} dv(U) \quad \text{mit} \quad U_j \in Q_j,$$

wobei Q_1, Q_2, ... eine Unterteilung von G in Borelsche Teilmengen ist. Nun folgt wie bei Lemma 3.6. aus Lemma 3.5. die Behauptung. |=|
Jetzt machen wir Gebrauch von der Existenz und Eindeutigkeit des Haarschen Maßes auf G: Da G eine kompakte Gruppe ist, gibt es auf ihr ein eindeutig bestimmtes invariantes Wahrscheinlichkeitsmaß

$$d_G U,$$

das (3.49) erfüllt.

Für eine beliebige Matrix B definieren wir dann B^G durch

$$B^G = \int_G U^{-1}BU d_G\, U \tag{3.50}$$

oder, was dasselbe ist, durch

$$(\xi, B^G\eta) = \int_G (U\xi, BU\eta)\, d_G U. \tag{3.51}$$

Es ist

$$(\lambda_1 B_1 + \lambda_2 B_2)^G = \lambda_1 B_1^G + \lambda_2 B_2^G \tag{3.52}$$

$$(U^{-1}BU)^G = UB^G U^{-1} = B^G, \quad U \in G, \tag{3.53}$$

wobei die letztere Beziehung aus der Invarianz

$$d_G U = d_G VU = d_G UV, \quad V \in G \tag{3.54}$$

des Maßes $d_G U$ folgt.
Ein Spezialfall von Lemma 3.7. ist die wichtige Relation

$$B^G \overset{G}{\succ} B. \tag{3.55}$$

Nach (3.53) ist B^G mit allen $U \in G$ vertauschbar. Wir wollen nun zeigen, daß B^G die einzige Matrix ist, die sowohl G-gemischter als B als auch vertauschbar mit allen $U \in G$ ist. Die konvexe Hülle der UBU^{-1}, $U \in G$, enthält daher genau ein mit allen U aus G vertauschbares Element.

Satz 3.8.

Bei gegebenen B und gegebener kompakter Gruppe G unitärer Matrizen gibt es genau ein A mit

$$A \overset{G}{\succ} B$$

und

$$AU = UA \quad \text{für alle} \quad U \in G.$$

Dieses durch B und G eindeutig bestimmte Element ist B^G.

Beweis: Wegen (3.55) braucht nur die Eindeutigkeit gezeigt werden. Aus (3.50) folgt zunächst $A^G = A$. Ist

$$A = \Sigma\, p_i U_i B U_i^{-1}, \quad U_i \in G,$$

so folgt nach (3.53) und (3.52)

$$A = A^G = \Sigma\, p_i\, (U_i B U_i^{-1})^G = \Sigma\, p_i B^G = B^G,$$

und damit ist die Behauptung gezeigt. |=|

Zusammenfassung

Zunächst werden die endlich-dimensionalen Dichtematrizen definiert und ihre physikalische Bedeutung sowie ihr Zusammenhang mit dem allgemeinen Zustandsbegriff erläutert. Mit Hilfe der Konvexität wird dann die Mischung von Zuständen betrachtet und das Problem gestellt, Eigenschaften einer Dichtematrix aus ihrer Lage in der Gesamtheit aller Dichtematrizen herzuleiten. Diese Zielstellung führt im letzten Kapitel zur Definition der Relation „gemischter als", deren Eigenschaften zusammen mit einigen Anwendungen hergeleitet werden. Die Verbindung zum Verhalten solcher konkaver Funktionen wie der Entropie wird im noch folgenden Teil II hergestellt.

Bemerkungen zur Literatur

Im Abschnitt 1.1. werden Tatsachen aus der linearen Algebra benutzt, die als allgemein bekannt gelten dürfen und die in fast jedem Lehrbuch der linearen Algebra oder – für die Verallgemeinerungen auf den unendlich-dimensionalen Fall – der Funktionalanalysis zu finden sind. Für weitergehende Studien gibt das Buch

[1] *Глазман, И. М., и Ю. И. Любич,* Конечномерный линейный яналиэ, Москва 1969,

viele Anregungen für den endlich-dimensionalen Fall. Für die Klasse der Operatoren mit Spur in einem Hilbert-Raum kann außer den Standardwerken der Funktionalanalysis auch

[2] *Гохберг, И. Ц., и М. Г. Крейн,* Введение в теорию линейних несамосопряженнх операторов, Москва 1956,

als Referenzbeispiel genannt werden. Für Abschnitt 1.2. kommt vor allem

[3] *von Neumann, J.*, Mathematische Grundlagen der Quantenmechanik, Berlin 1932,

in Betracht. Zum mathematischen Verständnis von Abschnitt 1.3. kann eines der Bücher

[4] *Neumark, M.A.*, Normierte Algebren, Berlin 1959;
[5] *Dixmier, J.*, Les algebres d'operateurs dans l'espace Hilbertien, Paris 1957;
[6] *Dixmier, J.*, Les C*-algèbres, Paris 1964;
[7] *Sakai, S.*, C*-Algèbras and W*-Algèbras, Berlin–Heidelberg–New York 1970;

herangezogen werden. Die in diesem Abschnitt angedeutete Verallgemeinerung hat ihren Ursprung offenbar in Überlegungen von *J. von Neumann*, der zunächst die nach ihm benannten Algebren (W*-Algebren) als Observablen-Algebren in Betracht zog und untersuchte. Die entscheidenden Impulse kamen jedoch aus der relativistischen Quantenfeldtheorie, wobei zu nennen sind

[8] *Wightman, A.S.*, Phys. Rev., *101*, 860, 1956.
[9] *Streater, R.F.*, und *A.S. Wightman*, PCT, Spin and Statistics and all that, New York–Amsterdam 1964;
[10] *Haag, R.*, Nuovo Cim., Suppl. *14*, 131, 1959;
[11] *Haag, R.*, und *B.Schroer*, J. Math. Phys., *3*, 248, 1962;
[12] *Segal, I.E.*, Mathematical Problems of Relativistic Physics, Providence, Rhode Island 1963;
und die Formulierung der Wightmanschen Theorie mit Hilfe von *-Algebren durch *Borchers* und *Uhlmann* (siehe [9]). Die ersten Anwendungen der im Abschnitt 1.3. skizzierten Gedanken auf Fragen der Statistischen Physik verdankt man *Ruelle*. Eine gute Orientierung gibt das Buch
[13] *Ruelle, D.*, Statistical Mechanics, Rigorous Results, New York–Amsterdam 1969.
Eine Einführung in die Theorie der konvexen Mengen kann man ebenfalls an sehr vielen Stellen finden. Es sei hier nur auf [1], [4] und [18] verwiesen. Eine Zusammenstellung wichtiger Resultate findet man in
[14] *Bonnesen, T.*, und *W.Fenchel*, Theorie der konvexen Körper, Ergebnisse der Mathematik, Berlin 1934.
Neuere Entwicklungen auf diesem Gebiet sind dargestellt in
[15] *Phelps, L.*, und *R.Robert*, Lectures on Choquet's Theorem, Princeton 1966;
[16] *Alfsen, E.M.*, Compact Convex Sets and Boundary Integrals, Berlin–Heidelberg–New York 1970.
Abschnitt 2.2. geht im wesentlichen auf *J.Neumann* zurück [3]. Auf den bemerkenswerten Unterschied zwischen einem klassischen und einem quantalen Gibbschen Ensemble (Nichteindeutigkeit der Zerlegung in „reine Zustände") wies besonders *Fano* in seinem Übersichtsartikel
[17] *Fano, U.*, Rev. Mod. Phys., *29*, 74, 1957,
hin. Für Abschnitt 3. kommen vor allem [1] sowie einige Abschnitte aus
[18] *Hardy, G.H.*, *J.E.Littlewood* und *G.Pólya*, Inequalities, London 1951;
[19] *Beckenbach, E.F.*, und *R.Bellmann*, Inequalities, Berlin–Göttingen–Heidelberg 1961;
in Betracht, wo auch weitere Literaturhinweise bezüglich der Lemmata 3.1. bis 3.4. zu finden sind. Die Definition 3.1. sowie die Sätze 3.2., 3.4., 3.6. und 3.7. scheinen neu zu sein. Eine Anwendung von Lemma 3.3. (bzw. eines kontinuierlichen Analogons) zeigten
[20] *Okubo, S.*, und *A.Isihara*, Some considerations of Entropy Change, University of Rochester, UR-875–343.
Den Hauptsatz von Abschnitt 3.1. findet man auch in
[21] *Uhlmann, A.*, Wiss. Z. Karl-Marx-Univ. Leipzig, Math.-Naturwiss. R., *20*, 633, 1971.

Anschrift des Verfassers:
Prof. Dr. sc. A. Uhlmann,
Sektion Physik, Karl-Marx-Universität,
DDR – 701 Leipzig, Linnéstr. 5