Quantum States with Maximum Information Entropy. II

W. Ochs and W. Bayer

Sektion Physik der Universität München

(Z. Naturforsch. **28 a**, 1571–1585 [1973]; received 8 June 1973)

We examine the conditions which a given information of the form "$\mathrm{Tr}(WA_r) = m_r$; $r = 1, \ldots, h$" must satisfy in order to determine a unique *quantum state with maximum information entropy*. Special consideration is given to the case of commuting A_r which is most important for statistical thermodynamics.

1. Introduction

In a previous paper[1,1a] we began the examination of one of the basic mathematical problems in the information theory approach to quantum statistics. The mean values

$$\mathrm{Tr}(WA_r) = m_r; \quad r = 1, \ldots, h \tag{1.1}$$

of h self-adjoint (s. a.) operators A_r in a separable Hilbert space $\mathscr{H}$ do not in general determine the state operator (density operator) W uniquely, but they define a whole class of state operators satisfying (1.1):

$$\mathfrak{W}_{\boldsymbol{A}}^{\boldsymbol{m}} \equiv \{W \in \mathfrak{W} \colon \mathrm{Tr}(W\boldsymbol{A}) = \boldsymbol{m}\},$$

where $\mathfrak{W}$ is the set of all state operators and the vector notation

$$\boldsymbol{m} = \{m_1, \ldots, m_h\}, \qquad \boldsymbol{A} = \{A_1, \ldots, A_h\}$$

is used. The problem at hand is to find necessary and/or sufficient conditions under which the h mean values (1.1) determine a unique *quantum state with maximum information entropy* (QME) $\hat{W} \in \mathfrak{W}_{\boldsymbol{A}}^{\boldsymbol{m}}$ defined by

$$(\forall W \in \mathfrak{W}_{\boldsymbol{A}}^{\boldsymbol{m}}) \quad W \neq \hat{W} \Leftrightarrow H(W) < H(\hat{W}) < \infty . \tag{1.2}$$

Here, the *information entropy* $H(W)$ of a state operator W is given by $H(W) = -\mathrm{Tr}(W \ln W)$ and the *mean value* of a s.a. operator $A = \int \lambda \, \mathrm{d}E_A(\lambda)$ in the state W is defined by the (generalized) mean value functional[2,3]

$$\mathrm{Tr}(WA) \equiv \int_{-\infty}^{\infty} \lambda \, \mathrm{d} \, \mathrm{Tr}[WE_A(\lambda)] . \tag{1.3}$$

In case of an operator with a pure point spectrum and the spectral representation

$$A = \sum_n a_n P_{a_n},$$

the definition (1.3) takes the form

$$\mathrm{Tr}(WA) = \sum_n a_n \mathrm{Tr}(WP_{a_n}) . \tag{1.4}$$

The above problem has hitherto been solved only in special cases: (I) For arbitrary $h \in \mathbf{N}$ and $\dim \mathscr{H} < \infty$, Wichmann[4] has shown that every (consistent) information of the form (1.1) determines a unique QME. (II) For $\dim \mathscr{H} = \infty$ and $h = 1$, Ingarden and Urbanik[5] and Bayer and Ochs[1] have shown that, apart from a few additional exceptional cases, the information (1.1) determines a unique QME if and only if m_1 is limited to a certain (A_1-dependent) interval and if a real number β exists with $\mathrm{Tr}[\exp(-\beta A_1)] < \infty$. In these cases, the problem can be solved completely because of its strong specialization.

If one examines, however, the possibility of QMEs under the general condition (1.1), then one encounters considerable mathematical difficulties: Given two s.a. operators A, B, little is known about the conditions under which $A + B$ is also s. a., and equally little is known about when e^{A+B} is of trace class or even when it exists. And the generalization of the mean value definition entails additional difficulties: Obviously, the functional $\mathrm{Tr}(W \cdot)$ of (1.3) is not linear since the sum of two s. a. operators is not necessarily s. a. again; but, even worse, the additivity of $\mathrm{Tr}(W \cdot)$ is not certain for s. a. sums of s. a. operators[2]. As a consequence, the general problem allows only of much weaker results then the particular cases mentioned above.

In Section 2 we examine the general conditions under which the information (1.1) determines a unique QME; this analysis turns out to be limited to state operators whose mean value functionals are linear in the A_r at least with respect to the linear combinations relevant for our problem. In Section 3 we then treat in more detail the special case of commuting A_r which is of particular importance for

Reprint requests to Dr. W. Ochs, Sektion Physik, Lehrstuhl Prof. Süßmann, *D-8000 München 2*, Theresienstraße 37, Germany.

statistical physics. In this case, all essential results of the case $h = 1$ can be retained under very general assumptions on the A_r which also cover the *macrocanonical ensemble*.

2. The General Case

The central notion in the solution[1,5] of the special problem "$h = 1$, $\dim\mathscr{H} = \infty$" was that of a *regular operator*: An operator X is called *(thermodynamically) regular* if X is s. a. and if there exists a real number β such that $\mathrm{Tr}[\exp(-\beta X)] < \infty$. In the present paper, we generalize the notion of regularity in order to apply it to sets of operators [and by an "operator" we understand henceforth a linear operator acting in a given separable Hilbert space $\mathscr{H}$ of infinite dimensions]. To avoid trivial complications we need some sort of independence among operators. We call n operators $A_1, \ldots, A_n$ *1-independent* if the $n+1$ operators $1, A_1, \ldots, A_n$ are linearly independent on the intersection of their domains.

Definition 1. A set $\boldsymbol{A} = \{A_1, \ldots, A_h\}$ of h s. a. operators is called $\boldsymbol{\alpha}$-*regular* for an $\boldsymbol{\alpha} \in \mathbf{R}^h$ if the linear combination

$$\boldsymbol{\alpha}\boldsymbol{A} = \sum_{r=1}^{h} \alpha_r A_r$$

is a regular operator. By $\mathfrak{R}_{\boldsymbol{A}}$ we denote the set of all $\boldsymbol{\alpha} \in \mathbf{R}^h$ for which $\boldsymbol{A}$ is $\boldsymbol{\alpha}$-regular. The operator set $\boldsymbol{A}$ is called *regular*[6], if the $A_r \in \boldsymbol{A}$ are s. a. and 1-independent and if $\mathfrak{R}_{\boldsymbol{A}} \neq \emptyset$. $\boldsymbol{A}$ is called *strongly regular* if $\boldsymbol{A}$ is regular and if $\mathfrak{R}_{\boldsymbol{A}}$ contains an element $\boldsymbol{\beta}$ with

$$\prod_{r=1}^{h} \beta_r \neq 0 .$$

For $h = 1$, regularity and strong regularity of an operator set are equivalent to the regularity of its element. For $h > 1$, regularity is a much weaker property of operator sets than strong regularity since every set of 1-independent, s. a. operators including at least one regular operator is itself a regular set. Strong regularity of a set of 1-independent, s. a. operators, however, is logically independent of the regularity of its elements: A set of 1-independent regular operators needs not to be strongly regular; on the other hand, two extremely nonregular operators can form a strongly regular set as is shown by the Hamiltonian of the harmonic oscillator.

For the following we repeat some notations from paper I[1]: The spectrum of an operator A is denoted by $\sigma(A)$ and, for s. a. operators, we set

$$\overline{\sigma}(A) = \sup\{x \in \sigma(A)\}, \qquad \underline{\sigma}(A) = \inf\{x \in \sigma(A)\}.$$

The domain of an operator A is denoted by $\mathscr{D}(A)$ and a complete orthonormal set of eigenvectors of a s. a. operator A with a pure point spectrum is called an *eigenbasis of A* or simply an *A-basis*. $P(\psi)$ is the projection operator which projects upon the one-dimensional subspace generated by the vector ψ. In analogy with the sets $\mathfrak{W}_{\boldsymbol{A}}^{\boldsymbol{m}}$ of the introduction we define

$$\mathfrak{W}_X^{\boldsymbol{m}} \equiv \{W \in \mathfrak{W}\colon \ \mathrm{Tr}(XW) = m\}, \quad \mathfrak{W}_{\boldsymbol{A}} \equiv \bigcup_{m \in \mathbf{R}^h} \mathfrak{W}_{\boldsymbol{A}}^{\boldsymbol{m}}.$$

In order to develop an applicable theory, we have to confine ourselves to state operators W whose mean value functionals $\mathrm{Tr}(W\,\cdot)$ are linear in the A_r with respect to the linear combinations occuring in our problem. Thus we define

$$\mathfrak{L}_{\boldsymbol{A}} \equiv \{W \in \mathfrak{W}_{\boldsymbol{A}}\colon \ (\forall \boldsymbol{\alpha} \in \mathfrak{R}_{\boldsymbol{A}})\ \boldsymbol{\alpha}\,\mathrm{Tr}(W\boldsymbol{A}) = \mathrm{Tr}(W\boldsymbol{\alpha}\boldsymbol{A})$$
$$\mathfrak{L}_{\boldsymbol{A}}^{\boldsymbol{m}} \equiv \mathfrak{W}_{\boldsymbol{A}}^{\boldsymbol{m}} \cap \mathfrak{L}_{\boldsymbol{A}} .$$

For simplicity we will often drop the index $\boldsymbol{A}$ if no confusion results. With the help of these notions, Corollary 6 and parts of Theorem II of [1] can be extended to the case of arbitrary h[1a]. For a simple formulation of the first lemma, we define the supremum of a real function on the empty set as zero.

Lemma 1. Let $\boldsymbol{A} = \{A_1, \ldots, A_h\}$ be a set of s.a. operators with $\mathfrak{R}_{\boldsymbol{A}} \neq \emptyset$. Then

$$\sup\{H(W)\colon \ W \in \mathfrak{L}_{\boldsymbol{A}}^{\boldsymbol{m}}\} < \infty \tag{2.1}$$

for all $\boldsymbol{m} \in \mathbf{R}^h$.

Proof. For $\boldsymbol{\alpha} \in \mathfrak{R}$ we obtain

$$\mathfrak{L}^{\boldsymbol{m}} \subset (\mathfrak{L} \cap \mathfrak{W})_{\boldsymbol{\alpha}\boldsymbol{A}}^{\boldsymbol{\alpha}\boldsymbol{m}} \subset \mathfrak{W}_{\boldsymbol{\alpha}\boldsymbol{m}}^{\boldsymbol{\alpha}\boldsymbol{m}}. \tag{2.2}$$

If $\mathfrak{L}^{\boldsymbol{m}} \neq \emptyset$, then there exists a $W \in \mathfrak{W}$ with

$$\mathrm{Tr}(W\boldsymbol{\alpha}\boldsymbol{A}) = \boldsymbol{\alpha}\boldsymbol{m}$$

and this implies $\underline{\sigma}(\boldsymbol{\alpha}\boldsymbol{A}) \leqq \boldsymbol{\alpha}\boldsymbol{m} \leqq \overline{\sigma}(\boldsymbol{\alpha}\boldsymbol{A})$. According to Corollary 6 of [1], this yields

$$\sup\{H(W)\colon \ W \in \mathfrak{W}_{\boldsymbol{\alpha}\boldsymbol{A}}^{\boldsymbol{\alpha}\boldsymbol{m}}\} < \infty \tag{2.3}$$

and from (2.2) and (2.3) we arrive at (2.1). In case of $\mathfrak{L}^{\boldsymbol{m}} = \emptyset$, the assertion holds by definition. □

To generalize Theorem II of [1] we need the following lemmata.

Lemma 2. The mean value functional $\mathrm{Tr}(W\,\cdot)$ introduced in (1.3) has the following properties:

$(\iota)^2$ $(\forall c \in \mathbf{R})$ $\mathrm{Tr}(W c A) = c\,\mathrm{Tr}(WA)$;

$(\iota\iota)^2$ If the s. a. operators $A_1, \ldots, A_h$ mutually commute, then the functional $\mathrm{Tr}(W\,\cdot)$ is linear in the A_r for all $W \in \mathfrak{W}_A$;

$(\iota\iota\iota)$ Let $\boldsymbol{A} = \{A_1, \ldots, A_h\}$ be a set of h s. a. operators and let V be a state operator of $\mathfrak{W}_A$ with a diagonal representation

$$V = \sum_{i=1}^{\infty} v_i\, P(\varphi_i)$$

such that $v_i \neq 0$ implies $\varphi_i \in \bigcap_r \mathscr{D}(A_r)$; then $V \in \mathfrak{L}_A$.

Proof of $(\iota\iota\iota)$: In case of $\mathfrak{R} = \emptyset$, the assertion is trivially valid. Accordingly we assume in the following that $\boldsymbol{\alpha} \in \mathfrak{R}$ (which implies $\boldsymbol{\alpha} \neq 0$). Let X be a positive s. a. operator and let W be a state operator with the properties

$$W = \sum_{i=1}^{\infty} w_i\, P(\psi_i), \qquad w_i \neq 0 \Rightarrow \psi_i \in \mathscr{D}(X)\,.$$

Then, according to Langerholc[2],

$$\mathrm{Tr}(WX) = \sum_{i=1}^{\infty} w_i (\psi_i, X \psi_i)\,. \tag{2.4}$$

We now decompose the operators $A_r \in \boldsymbol{A}$ into their positive and negative parts, $A_r = A_r^+ - A_r^-$ with $A_r^\pm \geqq \mathbf{0}$. Because of $V \in \mathfrak{W}_A$, all expressions $\mathrm{Tr}(V A_r^\pm)$ are finite and we have

$$\sum_{r=1}^{h} |\alpha_r| \{\mathrm{Tr}(V A_r^+) + \mathrm{Tr}(V A_r^-)\} < \infty\,. \tag{2.5}$$

Setting $C \equiv \boldsymbol{\alpha} \boldsymbol{A}$ we can assume without loss of generality that the regular operator C is bounded from below and hence it follows from (2.4) that

$$\mathrm{Tr}(VC^-) = \sum_{i=1}^{\infty} v_i (\varphi_i, C^- \varphi_i) \leqq \| C^- \| < \infty\,. \tag{2.6}$$

From (2.4) to (2.6) we obtain

$$\begin{aligned} &\sum_{r=1}^{h} \alpha_r\, \mathrm{Tr}(V A_r) \\ &= \sum_{r=1}^{h} \alpha_r \left\{ \sum_{i=1}^{\infty} v_i (\varphi_i, A_r^+ \varphi_i) \right\} - \sum_{r=1}^{h} \alpha_r \sum_{i=1}^{\infty} v_i (\varphi_i, A_r^- \varphi_i) \\ &= \sum_{i=1}^{\infty} v_i \sum_{r=1}^{h} \alpha_r \{(\varphi_i, A_r^+ \varphi_i) - (\varphi_i, A_r^- \varphi_i)\} \\ &= \sum_{i=1}^{\infty} v_i \{(\varphi_i, C^+ \varphi_i) - (\varphi_i, C^- \varphi_i)\}\,. \end{aligned} \tag{2.7}$$

From (2.5) to (2.7) it follows that

$$\begin{aligned} \infty &> \sum_{r=1}^{h} \alpha_r\, \mathrm{Tr}(V A_r) + 2 \| C^- \| \\ &\geqq \sum_{i=1}^{\infty} v_i \{(\varphi_i, C^+ \varphi_i) - (\varphi_i, C^- \varphi_i)\} \\ &\quad + 2 \sum_{i=1}^{\infty} v_i (\varphi_i, C^- \varphi_i) \\ &= \sum_{i=1}^{\infty} v_i \{(\varphi_i, C^+ \varphi_i) + (\varphi_i, C^- \varphi_i)\} \\ &= \mathrm{Tr}(VC^+) + \mathrm{Tr}(VC^-) \end{aligned} \tag{2.8}$$

and the Eqs. (2.7) and (2.8) finally yield

$$\begin{aligned} \sum_{r=1}^{h} \alpha_r\, \mathrm{Tr}(V A_r) &= \mathrm{Tr}(VC^+) - \mathrm{Tr}(VC^-) \\ &= \mathrm{Tr}(V \boldsymbol{\alpha} \boldsymbol{A}). \end{aligned}$$

□

Lemma 3. The inequality

$$\mathrm{Tr}(W \ln V) \leqq \mathrm{Tr}(W \ln W) \tag{2.9}$$

holds for all $V, W \in \mathfrak{W}$. In case of finite $\mathrm{Tr}(W \ln W)$, equality holds in (2.9) if and only if $V = W$.

Proofs of Lemma 3 are found in[7, 8]. These proofs, however, should be supplemented by a comment on the meaning of the expression $\mathrm{Tr}(W \ln V)$ which is not defined by the usual definitions for arbitrary $V, W \in \mathfrak{W}$. Let us consider, e.g., the case $W \lneqq [V]$, $\mathrm{Tr}([V]) < \infty$ where $[V]$ denotes the projection operator on the subspace generated by the range of V. In this case neither $W \ln V$ nor $(\ln V) W$ are trace class operators and also the generalized mean value definition (1.3) does not cover this case since $\ln V$ is not s. a. But the definition (1.4) suggests a unique extension of the domain of $\mathrm{Tr}(W\,\cdot)$ to all operators $\ln V$ with $V \in \mathfrak{W}$: If V has the spectral representation

$$V = \sum_{i \in I} v_i\, R_i\,,$$

then we define

$$\mathrm{Tr}(W \ln V) \equiv \sum^{i \in I.} \ln v_i\, \mathrm{Tr}(W R_i) \tag{2.10}$$

with $\qquad I_0 \equiv \{i \in I\colon \mathrm{Tr}(W R_i) \neq 0\}\,.$

This formula defines $\mathrm{Tr}(W \ln V)$ for arbitrary $V, W \in \mathfrak{W}$ and yields $\mathrm{Tr}(W \ln V) = -\infty$ in the case $W \lneqq [V]$.

Lemma 4. Let A, B be s. a. operators with the properties
$|\mathrm{Tr}(X e^Y)| < \infty$, $\mathrm{Tr}(e^X) < \infty$ for all $X, Y \in \{A, B\}$.
Then the equations

$$\mathrm{Tr}(A e^A)\, \mathrm{Tr}(e^B) = \mathrm{Tr}(A e^B)\, \mathrm{Tr}(e^A)\,,$$
$$\mathrm{Tr}(B e^A)\, \mathrm{Tr}(e^B) = \mathrm{Tr}(B e^B)\, \mathrm{Tr}(e^A)$$

are equivalent to the relation $A = B + c\,\mathbf{1}$ for some $c \in \mathbf{R}$.

With the help of Theorem II (c) of [1], the proof of Lemma 4 can be accomplished in the same way as in the case of a finite dimensional Hilbert space (see Lemma IV of [4])[1a].

To facilitate the formulation of Theorem I we introduce the notations

$$\begin{aligned} Z_A(\boldsymbol{\alpha}) &\equiv \mathrm{Tr}[\exp(-\boldsymbol{\alpha} \boldsymbol{A})] \\ V_A(\boldsymbol{\alpha}) &\equiv Z_A(\boldsymbol{\alpha})^{-1} \exp(-\boldsymbol{\alpha} \boldsymbol{A}) \qquad (2.11) \\ \langle A_r \rangle(\boldsymbol{\alpha}) &\equiv \mathrm{Tr}[V_A(\boldsymbol{\alpha}) A_r] \end{aligned}$$

and

$$\mathfrak{M}_A \equiv \{\boldsymbol{m} \in \mathbf{R}^h\colon (\exists \boldsymbol{\alpha} \in \mathfrak{R}_A)\; \boldsymbol{m} = \langle \boldsymbol{A} \rangle(\boldsymbol{\alpha})\}.$$

For simplicity we will often drop the index $\boldsymbol{A}$ if no confusion results.

Theorem I. Let $\boldsymbol{A} = \{A_1, \ldots, A_h\}$ be a regular operator set. Then:

(a) $(\forall \boldsymbol{\alpha} \in \mathfrak{R}_A)\quad V_A(\boldsymbol{\alpha}) \in \mathfrak{W}_A \Rightarrow V_A(\boldsymbol{\alpha}) \in \mathfrak{L}_A$;

(b) $\langle \boldsymbol{A} \rangle^{-1}$ is an injective map of $\mathfrak{M}_A$ into $\mathfrak{R}_A$;

(c) For every point $\boldsymbol{\alpha} \in \mathfrak{R}_A$ there exists a real number $u \in (0, 1]$ such that the half-line $\{x\boldsymbol{\alpha}\colon x \in \mathbf{R},\ x > u\}$ is contained in $\mathfrak{R}_A$ whereas the half-line $\{x\boldsymbol{\alpha}\colon x \in \mathbf{R},\ x < u\}$ has no point in common with $\mathfrak{R}_A$;

(d) For every $\boldsymbol{m} \in \mathfrak{M}_A$, the information entropy H has a unique maximum in $\mathfrak{L}_A^{\boldsymbol{m}}$ which is assumed at the QME $\hat{W}[\boldsymbol{m}] \equiv V\{\langle \boldsymbol{A} \rangle^{-1}(\boldsymbol{m})\}$:

$$W \neq \hat{W}[\boldsymbol{m}] \Leftrightarrow H(W) < H(\hat{W}[\boldsymbol{m}]) < \infty$$

for all $\boldsymbol{m} \in \mathfrak{M}_A$, $W \in \mathfrak{L}_A^{\boldsymbol{m}}$ and

$$\begin{aligned} H(\hat{W}[\boldsymbol{m}]) &= \sup\{H(W)\colon W \in \mathfrak{L}_A^{\boldsymbol{m}}\} \\ &= \ln Z_A\{\langle \boldsymbol{A} \rangle^{-1}(\boldsymbol{m})\} + \boldsymbol{m} \langle \boldsymbol{A} \rangle^{-1}(\boldsymbol{m}). \end{aligned}$$

Proof. ad (a): According to Definition (2.11), $V(\boldsymbol{\alpha})$ is contained in $\mathfrak{W}$ for all $\boldsymbol{\alpha} \in \mathfrak{R}$ and has the same eigenbasis as $\boldsymbol{\alpha A}$. Hence every $V(\boldsymbol{\alpha})$-basis is contained in $\mathscr{D}(\boldsymbol{\alpha A})$ and from

$$\mathscr{D}(\boldsymbol{\alpha A}) = \bigcap_r \mathscr{D}(\alpha_r A_r) = \bigcap_r \mathscr{D}(A_r)$$

and Lemma 2 ($\iota\iota$), the assertion follows.

ad (b): Let $\mathrm{Tr}[V(\boldsymbol{\alpha}) A_r] = \mathrm{Tr}[V(\boldsymbol{\beta}) A_r]$ for $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathfrak{R}$ and $r = 1, \ldots, h$. By multiplying the r-th equation first with $-\alpha_r$ and then with $-\beta_r$ we get

$$\mathrm{Tr}[-\alpha_r A_r V(\boldsymbol{\alpha})] = \mathrm{Tr}[-\alpha_r A_r V(\boldsymbol{\beta})], \qquad (2.12\text{a})$$

$$\mathrm{Tr}[-\beta_r A_r V(\boldsymbol{\alpha})] = \mathrm{Tr}[-\beta_r A_r V(\boldsymbol{\beta})]. \qquad (2.12\text{b})$$

If we add the h Eqs. (2.12a) and (2.12b), respectively, and use statement (a), then we find

$$\begin{aligned} &\mathrm{Tr}(-\boldsymbol{\alpha A}\, e^{-\boldsymbol{\alpha A}})\,\mathrm{Tr}(e^{-\boldsymbol{\beta A}}) \\ &\qquad = \mathrm{Tr}(-\boldsymbol{\alpha A}\, e^{-\boldsymbol{\beta A}})\,\mathrm{Tr}(e^{-\boldsymbol{\alpha A}}), \\ &\mathrm{Tr}(-\boldsymbol{\beta A}\, e^{-\boldsymbol{\alpha A}})\,\mathrm{Tr}(e^{-\boldsymbol{\beta A}}) \qquad (2.13) \\ &\qquad = \mathrm{Tr}(-\boldsymbol{\beta A}\, e^{-\boldsymbol{\beta A}})\,\mathrm{Tr}(e^{-\boldsymbol{\alpha A}}). \end{aligned}$$

The Eqs. (2.13) and Lemma 4 imply $\boldsymbol{\alpha A} = \boldsymbol{\beta A} + c\,\mathbf{1}$ and the assumed 1-independence of the operators $A_1, \ldots, A_h$ finally yields $\boldsymbol{\alpha} = \boldsymbol{\beta}$. The assertion (c) results immediately from Theorem II of [1].

ad (d): Let $\boldsymbol{m}$ be an arbitrary element of $\mathfrak{M}$ with $\boldsymbol{\alpha} = \langle \boldsymbol{A} \rangle^{-1}(\boldsymbol{m})$ and let W be an arbitrary element of $\mathfrak{L}^{\boldsymbol{m}}$. Then it follows from Lemma 3 that

$$H(W) = -\mathrm{Tr}(W \ln W) \leqq -\mathrm{Tr}[W \ln V(\boldsymbol{\alpha})]. \qquad (2.14)$$

Because $\boldsymbol{\alpha} \in \mathfrak{R}$, $\boldsymbol{\alpha A}$ has a spectral representation

$$\boldsymbol{\alpha A} = \sum_{i=1}^{\infty} d_i D_i$$

and this implies

$$V(\boldsymbol{\alpha}) = Z(\boldsymbol{\alpha})^{-1} \sum_{i=1}^{\infty} \exp\{-d_i\} D_i .$$

From (2.10) it follows that

$$\begin{aligned} \mathrm{Tr}[W \ln V(\boldsymbol{\alpha})] &= \sum_{i=1}^{\infty} \{-\ln Z(\boldsymbol{\alpha}) - d_i\}\,\mathrm{Tr}(W D_i) \\ &= -\ln Z(\boldsymbol{\alpha}) - \mathrm{Tr}(W \boldsymbol{\alpha A}). \qquad (2.15) \end{aligned}$$

From (2.14) and (2.15) we obtain for all $W \in \mathfrak{L}^{\boldsymbol{m}}$

$$\begin{aligned} H(W) &\leqq \ln Z(\boldsymbol{\alpha}) + \mathrm{Tr}(W \boldsymbol{\alpha A}) \\ &= \ln Z(\boldsymbol{\alpha}) + \sum_{r=1}^{h} \alpha_r \mathrm{Tr}(W A_r) \\ &= \ln Z(\boldsymbol{\alpha}) + \boldsymbol{m\alpha} \qquad (2.16) \\ &= \ln Z\{\langle \boldsymbol{A} \rangle^{-1}(\boldsymbol{m})\} + \boldsymbol{m} \langle \boldsymbol{A} \rangle^{-1}(\boldsymbol{m}). \end{aligned}$$

If, in particular, we set $W = V(\boldsymbol{\alpha})$ in (2.15) and transform the right side exactly as in (2.16), then we find

$$\begin{aligned} H[V(\boldsymbol{\alpha})] &= H(\hat{W}[\boldsymbol{m}]) \\ &= \ln Z\{\langle \boldsymbol{A} \rangle^{-1}(\boldsymbol{m})\} + \boldsymbol{m} \langle \boldsymbol{A} \rangle^{-1}(\boldsymbol{m}). \end{aligned} \qquad (2.17)$$

The Eqs. (2.16), (2.17) and Lemma 3 finally yield

$$(\forall W \in \mathfrak{L}^{\boldsymbol{m}}) \quad W \neq \hat{W}[\boldsymbol{m}] \Leftrightarrow H(W) < H(\hat{W}) < \infty .$$

□

A comparison of this theorem with the analogous theorem in the case $h = 1$ shows that Theorem I contains a new assumption, viz. the 1-independence of the A_r, as well as a new limitation of its assertions, viz. the confinement to state operators $W \in \mathfrak{L}_A$. Neither point, however, seems grave to us from the physical point of view. The assumption of the 1-independence of the A_r imports no loss of generality; it can always be satisfied by eliminating the redundant information and reducing the set $\boldsymbol{A}$ in a suitable way. The confinement to $W \in \mathfrak{L}_A$, on the other hand, is certainly a restriction of the generality of Theorem I; it can hardly be avoided as Eq. (2.16) does not exclude that $\mathfrak{W}_A^m \backslash \mathfrak{L}_A^m$ contains elements with an information entropy greater than $H(\hat{W}[\boldsymbol{m}])$. But this very confinement became necessary only because we introduced the generalized mean value functional (1.3) in order to increase the set $\mathfrak{W}_A^m$ of admissible state operators. If we had defined the mean value of an observable, as usual, by the *matrix trace*

$$\operatorname{tr}(AW) \equiv \sum_{\{\varphi_i\}} (\varphi_i, AW\varphi_i),$$

then from the very beginning only those state operators W would have been admitted for which the h operators [9] $A_1 W, \ldots, A_h W$ are of trace class. But for these state operators, the (restricted) mean value functional $\operatorname{tr}(\cdot\, W)$ is trivially linear in the A_r since

$$\sum_{r=1}^{h} \alpha_r (A_r W) = (\boldsymbol{\alpha} \boldsymbol{A}) W$$

and since the trace class is a vector space over $\mathbf{C}$. In other words, the usual definition of mean values would have obviated the confinement to W of $\mathfrak{L}_A$. This raises the question whether the generalization (1.3) of the mean value definition was at all worthwile in view of the fact that this generalization not only poses additional mathematical difficulties but had to be reduced again in Lemma 1 and Theorem I. However, even with the restriction to the state operators of $\mathfrak{L}_A$, the generalization is meaningful because the set $\{W \in \mathfrak{W}\colon \operatorname{tr}(\boldsymbol{A}W) = \boldsymbol{m}\}$ is in general a true subset of $\mathfrak{L}_A$.

In addition to these minor limitations of generality, Theorem I also lacks of some essential results which have been deduced in the special cases (I) and (II), viz. the statements about extension and shape of $\mathfrak{R}_A$ and $\mathfrak{M}_A$ as well as statements about the differentiability of the functions Z, $\langle A_r \rangle$ and H with respect to the α_s. This deficiency is unavoidable and due to the extreme variety of the class of all regular operator sets. If we consider, e.g., a $\boldsymbol{\beta}$-regular operator set $\boldsymbol{X} = \{X_1, \ldots, X_h\}$, then we do not even know whether $\boldsymbol{X}$ is $\boldsymbol{\alpha}$-regular for all $\boldsymbol{\alpha}$ in a sufficiently small neighborhood of $\boldsymbol{\beta}$. And just as little can be said about $\mathfrak{M}_A$; even for strongly regular operator sets $\boldsymbol{A}$, $\mathfrak{M}_A$ may be empty as the following example shows.

Example (2.18): Let $\{\varphi_i\colon i \in \mathbf{N}\}$ be an arbitrary basis of $\mathscr{H}$; then the operators

$$A = \sum_{i=1}^{\infty} a_i P(\varphi_i), \qquad a_i = \ln i - (-2)^i,$$

$$B = \sum_{i=1}^{\infty} b_i P(\varphi_i), \qquad b_i = (-2)^i$$

are s. a., 1-independent and commuting. From

$$Z(\alpha, \beta) = \sum_{i=1}^{\infty} \exp\{-(\alpha a_i + \beta b_i)\} = \sum_{i=1}^{\infty} i^{-\alpha} \exp\{(\alpha - \beta)(-2)^i\}$$

and

$$\langle B \rangle (\alpha, \alpha) = Z(\alpha, \alpha)^{-1} \sum_{i=1}^{\infty} b_i \exp\{-(\alpha a_i + \alpha b_i)\} = (1/Z) \sum_{i=1}^{\infty} i^{-\alpha} (-2)^i$$

it follows that the operator set $\{A, B\}$ is strongly regular and we find

$$\mathfrak{R}_{\{A,B\}} = \{(\alpha, \beta) \in \mathbf{R}^2\colon \alpha = \beta > 1\}, \quad \mathfrak{M}_{\{A,B\}} = \emptyset.$$

For similar reasons, Theorem I contains no statement about the differentiability of the functions Z, $\langle A_r \rangle$ and H. These stronger statements can only be extended to the case "$h > 1$, $\dim \mathscr{H} = \infty$" under additional restricting conditions on the operator set $\boldsymbol{A}$.[10]

Theorem II provides concrete examples of strongly regular operator sets with an h-dimensional $\mathfrak{R}_A$ and non-vanishing $\mathfrak{M}_A$ without assuming commuting A_r.

Notations: A s. a. operator is called *semibounded* if it is bounded from at least one side. On the set of all s. a., semibounded operators we define a function ν indicating the direction of the (possible) unboundedness:

$$\nu(A) = -1 \quad \text{if } A \text{ is not bounded from below}$$
$$\nu(A) = +1 \quad \text{otherwise.}$$

For a regular operator X we define[1]

$$\lambda(X) \equiv \nu(X) \inf\{\alpha\, \nu(X)\colon \alpha \in \mathbf{R},\ Z_X(\alpha) < \infty\}.$$

For the given operators $A_r \in \boldsymbol{A}$ and their regular linear combinations we introduce the special abbreviations

$$\nu_r = \nu(A_r), \qquad \nu_\alpha = \nu(\boldsymbol{\alpha A}),$$
$$\boldsymbol{\nu} \equiv \{\nu_1, \ldots, \nu_h\}, \qquad \lambda_r = \lambda(A_r), \quad \lambda_\alpha = \lambda(\boldsymbol{\alpha A}).$$

By removing all isolated eigenvalues with finite multiplicity from the spectrum of a s. a. operator A one obtains the *essential spectrum* $\sigma_{\text{ess}}(A)$. An operator A is called *B-bounded* if

$$\mathscr{D}(B) \subset \mathscr{D}(A)$$

and

$$\|A\varphi\| \leqq a\|\varphi\| + b\|B\varphi\| \quad \text{for all } \varphi \in \mathscr{D}(B);$$

the infimum of all possible values b is called the *B-bound of A*. By $K(\boldsymbol{\alpha}, \varepsilon)$ we denote the closed ball $\{\boldsymbol{x} \in \mathbf{R}^h: |\boldsymbol{x} - \boldsymbol{\alpha}| \leqq \varepsilon\}$ with centre $\boldsymbol{\alpha}$ and radius ε.

Theorem 2. (ι) Let $\boldsymbol{A} = \{A_1, \ldots, A_h\}$ be a set of h s. a., 1-independent operators such that exactly one of them, say A_k, is regular whereas the remaining operators are completely continuous. Then the operator set $\boldsymbol{A}$ is strongly regular and we find

$$\{\boldsymbol{\alpha} \in \mathbf{R}^h: \nu_k \alpha_k > \nu_k \lambda_k\} \subseteqq \Re_{\boldsymbol{A}}, \tag{2.19}$$

$$(\forall \boldsymbol{\alpha} \in \text{int}\, \Re_{\boldsymbol{A}}) \quad |\langle \boldsymbol{A} \rangle(\boldsymbol{\alpha})| < \infty. \tag{2.20}$$

($\iota\iota$) Let A, B be two regular, 1-independent operators and let A be B-bounded with B-bound b. Then $\{A, B\}$ is a strongly regular operator set and we find

$$\left\{(\alpha, \beta) \in \mathbf{R}^2: \begin{array}{c} \mathsf{b}|\alpha| < |\beta| \\ \nu(A)\alpha > \nu(A)\lambda(A) \\ \nu(B)\beta > \nu(B)\lambda(B) \end{array}\right\} \subseteqq \Re_{\{A,B\}}, \tag{2.21}$$

$$(\forall(\alpha, \beta) \in \text{int}\, \Re_{\{A,B\}}) \qquad \begin{array}{l} |\langle A \rangle(\alpha, \beta)| < \infty \\ |\langle B \rangle(\alpha, \beta)| < \infty. \end{array} \tag{2.22}$$

Proof of (ι): In Lemma 4 of [1] we have shown that a regular operator is semibounded and has an empty essential spectrum. As all $A_r \in \boldsymbol{A}$ are s. a., all A_r except for A_k are completely continuous and A_k is regular, we infer from standard theorems in the perturbation theory of linear operators[11] that the operators $\boldsymbol{\alpha A}$ are s. a. for all $\boldsymbol{\alpha} \in \mathbf{R}^h$ and have the additional properties $\mathscr{D}(\boldsymbol{\alpha A}) = \mathscr{D}(A_k)$ and

$$\sigma_{\text{ess}}(\boldsymbol{\alpha A}) = \sigma_{\text{ess}}(A_k) = \emptyset.$$

As A_k is semibounded, $\boldsymbol{\alpha A}$ is also semibounded[11] with $\nu_\alpha = \nu_k \operatorname{sgn} \alpha_k$. Hence the operator $\exp(-\boldsymbol{\alpha A})$ is s. a., positive and bounded for all $\boldsymbol{\alpha} \in \mathbf{R}^h$ with $\nu_k \operatorname{sgn} \alpha_k = +1$ and has a pure point spectrum, with zero as the only possible limit point. These properties guarantee that the matrix trace

$$\operatorname{tr}(e^{-\boldsymbol{\alpha A}}) = \sum_{i=1}^{\infty} (\varphi_i, e^{-\boldsymbol{\alpha A}} \varphi_i)$$

is independent of the basis $\{\varphi_i\}$ (but not necessarily finite)[12] and satisfies, by reason of *Peierls inequality*[8], the relation

$$\operatorname{tr}(e^{-\boldsymbol{\alpha A}}) = \sup \sum_{i=1}^{\infty} \exp\{-(\varphi_i, \boldsymbol{\alpha A} \varphi_i)\} \tag{2.23}$$

in which the supremum is taken over all bases $\{\varphi_i: i \in \mathbf{N}\} \subset \mathscr{D}(A_k)$. From

$$\Big\| \sum_{\substack{r=1 \\ r \neq k}}^{h} \alpha_r A_r \Big\| \equiv C(\boldsymbol{\alpha}) < \infty$$

for all $\boldsymbol{\alpha} \in \mathbf{R}^h$ we obtain

$$\begin{aligned} &\exp\{-(\psi, \boldsymbol{\alpha A} \psi)\} \\ &\quad = \exp\{-(\psi, \alpha_k A_k \psi)\} \exp\{-(\psi, \sum_{r \neq k} \alpha_r A_r \psi)\} \\ &\quad \leqq e^{C(\boldsymbol{\alpha})} \exp\{-(\psi, \alpha_k A_k \psi)\} \end{aligned} \tag{2.24}$$

for all $\psi \in \mathscr{D}(A_k)$. And from (2.23) and (2.24) it follows

$$\begin{aligned} \operatorname{tr}(e^{-\boldsymbol{\alpha A}}) &= \sup_{\{\varphi_i\}} \sum_{i=1}^{\infty} \exp\{-(\varphi_i, \boldsymbol{\alpha A} \varphi_i)\} \\ &\leqq \sup_{\{\varphi_i\}} \sum_{i=1}^{\infty} e^{C(\boldsymbol{\alpha})} \exp\{-(\varphi_i, \alpha_k A_k \varphi_i)\} \\ &= e^{C(\boldsymbol{\alpha})} \operatorname{tr}[\exp(-\alpha_k A_k)]. \end{aligned} \tag{2.25}$$

If we further restrict α_k to the half-line $\{\nu_k \alpha_k > \nu_k \lambda_k\}$, then from (2.25) and Theorem II of [1] it follows

$$\operatorname{Tr}(e^{-\boldsymbol{\alpha A}}) \leqq e^{C(\boldsymbol{\alpha})} \operatorname{Tr}\{\exp(-\alpha_k A_k)\} < \infty$$

which proves Equation (2.19).

Next we choose an element $\boldsymbol{\alpha} \in \text{int}\, \Re$. Then there exists a real number $\varepsilon > 0$ such that $K(\boldsymbol{\alpha}, \varepsilon)$ is also contained in int $\Re$. Since $\boldsymbol{A}$ is $\boldsymbol{\alpha}$-regular, $\boldsymbol{\alpha A}$ has a diagonal representation

$$\boldsymbol{\alpha A} = \sum_{i=1}^{\infty} d_i P(\psi_i). \tag{2.26}$$

Without loss of generality we can assume

$$\operatorname{sgn} \alpha_k = \nu_k = +1$$

and the d_i as monotonically increasing. According to our assumptions, $\|A_r\|$ is finite for all $r \neq k$ and this implies

$$(\forall r \in \{1, \ldots, h\};\ r \neq k) \quad |\langle A_r \rangle(\boldsymbol{\alpha})| < \infty. \tag{2.27}$$

Dividing the regular operator A_k into its positive and negative part, $A_k = A_k^+ - A_k^-$, we infer from $\nu_k = +1$ that $\|A_k^-\| < \infty$ which implies

$$\langle A_k^- \rangle (\boldsymbol{\alpha}) < \infty . \qquad (2.28)$$

For the unbounded, positive operator A_k^+ we obtain

$$\alpha_k A_k^+ = \boldsymbol{\alpha} \boldsymbol{A} + \alpha_k A_k^- - \sum_{r \neq k} \alpha_r A_r ,$$

and hence it follows from (2.26) that

$$\alpha_k (\psi_i, A_k^+ \psi_i) \leqq |d_i| + \alpha_k \|A_k^-\| + \sum_{r \neq k} |\alpha_r| \cdot \|A_r\| \equiv |d_i| + F \quad \text{with} \quad F < \infty;$$

$$\sum_{i=1}^{\infty} (\psi_i, A_k^+ \psi_i) e^{-d_i} \leqq Z(\boldsymbol{\alpha}) F + \alpha_k^{-1} \sum_{i=1}^{\infty} |d_i| e^{-d_i} . \qquad (2.29)$$

If we denote the greater of the two solutions of the equation $x = e^{yx}$ (with $0 < x$, $0 < y < 1/e$) by $g(y)$ and set $G \equiv \max\{1, |\underline{\sigma}(A_k)|\}$, then we find

$$\sum_{i=1}^{\infty} |d_i| e^{-d_i} \leqq G g(y) \sum_{d_i < g(y)} e^{-d_i} + \sum_{d_i \geqq g(y)} e^{-(1-y) d_i} < G g(y) Z(\boldsymbol{\alpha}) + \sum_{i=1}^{\infty} e^{-(1-y) d_i} . \qquad (2.30)$$

If we now choose y such that $(1-y)\boldsymbol{\alpha} \in K(\boldsymbol{\alpha}, \varepsilon)$, then it follows

$$\sum_{i=1}^{\infty} e^{-(1-y) d_i} = \sum_{i=1}^{\infty} (\psi_i, \exp[-(1-y)\boldsymbol{\alpha} \boldsymbol{A}] \psi_i) = Z[(1-y)\boldsymbol{\alpha}] < \infty \qquad (2.31)$$

and the Eqs. (2.29) to (2.31) yield

$$\sum_{i=1}^{\infty} (\psi_i, A_k^+ \psi_i) e^{-d_i} < \infty . \qquad (2.32)$$

From $A_k^+ \geqq \mathbf{0}$ and the Eqs. (2.4) and (2.32) we nfer that

$$\langle A_k^+ \rangle (\boldsymbol{\alpha}) = Z(\boldsymbol{\alpha})^{-1} \sum_{i=1}^{\infty} (\psi_i, A_k^+ \psi_i) e^{-d_i} < \infty .$$

The relation $\langle A_k \rangle = \langle A_k^+ \rangle - \langle A_k^- \rangle$ and the above equations finally yield $|\langle \boldsymbol{A} \rangle (\boldsymbol{\alpha})| < \infty$ for all $\boldsymbol{\alpha} \in \operatorname{int} \mathfrak{R}$.

Proof of (ii): Let α, β be two real numbers with $\mathsf{b}|\alpha| < |\beta|$ and set $\delta = \frac{1}{2}(|\beta| \cdot |\alpha|^{-1} - \mathsf{b})$. Then, according to our assumption, there exists a real number a such that

$$(\forall f \in \mathscr{D}(B)) \quad \|A f\| \leqq a \|f\| + (\mathsf{b} + \delta) \|B f\|$$

and this yields

$$(\forall f \in \mathscr{D}(B)) \quad \|\alpha A f\| \leqq c \|f\| + q \|\beta B f\| \qquad (2.33)$$

with $c = |\alpha| a$ and

$$q = |\alpha| (\mathsf{b} + \delta) |\beta|^{-1} = \tfrac{1}{2}(|\alpha| \mathsf{b} |\beta|^{-1} + 1) < 1 .$$

If we demand in addition that $\nu(A)\alpha > \nu(A)\lambda(A)$ and $\nu(B)\beta > \nu(B)\lambda(B)$, then it follows from the assumed regularity of A and B and from Theorem II of [1] that $\operatorname{Tr}(e^{-\alpha A}) < \infty$ and $\operatorname{Tr}(e^{-\beta B}) < \infty$, and hence $e^{-\alpha A} e^{-\beta B}$ is of trace class, too. By virtue of Eq. (2.33) and $q < 1$, $\alpha A + \beta B$ is also s. a.[11] and the *Golden-Thompson inequality*[13]

$$\operatorname{Tr}[\exp(-\alpha A - \beta B)] \leqq \operatorname{Tr}[\exp(-\alpha A)\exp(-\beta B)]$$

yields $\operatorname{Tr}[\exp(-\alpha A - \beta B)] < \infty$. Thus the operator set $\{A, B\}$ is (α, β)-regular for all $(\alpha, \beta) \in \mathbf{R}^2$ with $\mathsf{b}|\alpha| < |\beta|$ and

$$\nu(A)\alpha > \nu(A)\lambda(A), \quad \nu(B)\beta > \nu(B)\lambda(B) .$$

Next, we choose a point $(\alpha, \beta) \in \operatorname{int} \mathfrak{R}_{\{A, B\}}$. Then there exists a real number $\varepsilon > 0$ such that

$$K\{(\alpha, \beta), \varepsilon\}$$

is also contained in $\operatorname{int} \mathfrak{R}_{\{A, B\}}$, and there exists an $(\alpha A + \beta B)$-basis $\{\chi_i : i \in \mathbf{N}\}$ so that

$$\exp\{-\alpha A - \beta B\} = \sum_{i=1}^{\infty} \exp\{-x_i\} P(\chi_i) \qquad (2.34)$$

with

$$x_i = \alpha(\chi_i, A \chi_i) + \beta(\chi_i, B \chi_i) . \qquad (2.34)$$

Without loss of generality we assume $\nu(A) = +1$ and set $y_i = |(\chi_i, A \chi_i)|$ and

$$Q = \max\{g(\varepsilon), |\underline{\sigma}(A)|\} .$$

Then we obtain

$$\langle A^- \rangle (\alpha, \beta) \leqq \|A^-\| < \infty ;$$

$$\sum_{i=1}^{\infty} (\chi_i, A^- \chi_i) e^{-x_i} \leqq \|A^-\| Z(\alpha, \beta) < \infty ;$$

$$\sum_{i=1}^{\infty} y_i e^{-x_i} = \sum_{y_i \leqq Q} y_i e^{-x_i} + \sum_{y_i > Q} y_i e^{-x_i} \equiv \Sigma_1 + \Sigma_2 ,$$

$$\Sigma_1 \leqq Q Z(\alpha, \beta) < \infty$$

$$\Sigma_2 \leqq \sum_{i=1}^{\infty} \exp\{-(\alpha - \varepsilon)(\chi_i, A \chi_i) - \beta(\chi_i, B \chi_i)\} = \sum_{i=1}^{\infty} \exp\{\chi_i, -[(\alpha - \varepsilon) A + \beta B] \chi_i\} \leqq \operatorname{Tr}[\exp\{-(\alpha - \varepsilon) A - \beta B\}] \quad \text{because of Peierls inequality}$$

$= Z(\alpha - \varepsilon, \beta) < \infty$ because

$(\alpha - \varepsilon, \beta) \in K\{(\alpha, \beta), \varepsilon\}$.

These equations yield

$$\sum_{i=1}^{\infty} (\chi_i, A^+ \chi_i)\, e^{-x_i} \leqq \sum_{i=1}^{\infty} |(\chi_i, A\chi_i)|\, e^{-x_i} + \sum_{i=1}^{\infty} (\chi_i, A^- \chi_i)\, e^{-x_i} < \infty . \tag{2.35}$$

From $A^+ \geqq \mathbf{0}$ and the Eqs. (2.4), (2.34) and (2.35) we infer that $\langle A^+ \rangle(\alpha, \beta) < \infty$ and thus we finally obtain $|\langle A \rangle(\alpha, \beta)| < \infty$. In complete analogy one shows $|\langle B \rangle(\alpha, \beta)_r| < \infty$.

□

3. The Case of Commuting A_r

In this section we assume that all operators $A_r \in \boldsymbol{A}$ mutually commute. This assumption allows of much stronger results than those of the preceding section. Theorem I and Lemma 2(*u*) immediately yield

Corollary 5. Let $\boldsymbol{A}$ be a regular set of mutually commuting, s. a. operators. Then $\mathfrak{W}_{\boldsymbol{A}}^{\boldsymbol{m}} \subset \mathfrak{L}_{\boldsymbol{A}}$, and the assertion (d) of Theorem I takes the form:

(d*) $\quad W \neq \hat{W}[\boldsymbol{m}] \Leftrightarrow H(W) < H(\hat{W}[\boldsymbol{m}]) < \infty$

for all $\boldsymbol{m} \in \mathfrak{M}_{\boldsymbol{A}}$, $W \in \mathfrak{W}_{\boldsymbol{A}}^{\boldsymbol{m}}$ and

$$\begin{aligned} H(\hat{W}[\boldsymbol{m}]) &= \sup\{H(W)\colon W \in \mathfrak{W}_{\boldsymbol{A}}^{\boldsymbol{m}}\} \\ &= \ln Z(\langle \boldsymbol{A} \rangle^{-1}(\boldsymbol{m})) + \boldsymbol{m} \langle \boldsymbol{A} \rangle^{-1}(\boldsymbol{m}) \end{aligned}$$

for all $\boldsymbol{m} \in \mathfrak{M}_{\boldsymbol{A}}$.

To establish more detailed results we must impose a further condition on the operators A_r. In the first theorem of this section we will show that in the case of a regular set of commuting and *semibounded* operators, the essential results of the case $h = 1$ can be extended to the case of arbitrary h. For this purpose we prove some relations between the spectra of commuting operators which are also of interest on their own merits. Following Kato[11], we define the *continuous spectrum* $\sigma_{\text{con}}(A)$ of a s.a. operator A as the spectrum of the continuous part of A. In addition, we introduce the unit elements $\boldsymbol{e}_r = \{0, \ldots, \underset{(r)}{1}, 0, \ldots, 0\}$ of $\mathbf{R}^h$ for $r = 1, \ldots, h$.

Lemma 6. Let X be a s. a. operator on a separable Hilbert space $\mathscr{H}$ and let f be a real-valued Borel function on $\mathbf{R}$ such that $Y = f(X)$ is also a s.a. operator on $\mathscr{H}$. Then

$$\sigma_{\text{ess}}(Y) = \emptyset \Rightarrow \sigma_{\text{con}}(X) = \emptyset , \tag{3.1}$$

$$\sigma_{\text{con}}(X) = \emptyset \Rightarrow \sigma_{\text{con}}(Y) = \emptyset . \tag{3.2}$$

Proof: If $X = \int \lambda \,\mathrm{d}E_\lambda$, $Y = \int \lambda \,\mathrm{d}F_\lambda$ are the spectral representations of the s.a. operators X, Y and E, F the corresponding spectral measures on the set $\mathscr{B}$ of all Borel sets of $\mathbf{R}$, then $Y = f(X)$ implies the relation[14]

$$(\forall M \in \mathscr{B}) \quad F(M) = E\{f^{-1}(M)\} . \tag{3.3}$$

Setting

$$M_{2n+1} = (n, n+1] , \qquad M_{2n} = (-n-1, -n]$$

and

$$B_n = f^{-1}(M_n) \quad \text{for} \quad n = 0, 1, 2, \ldots$$

we obtain

$$(\forall i, j) \quad B_i \in \mathscr{B}, \quad i \neq j \Leftrightarrow B_i \cap B_j = \emptyset, \quad \bigcup_{i=0}^{\infty} B_i = \mathbf{R} . \tag{3.4}$$

If we now assume $\sigma_{\text{ess}}(Y) = \emptyset$, then we obtain $\operatorname{Tr}\{F(M_i)\} < \infty$ and from (3.3) it follows

$$\operatorname{Tr}\{E(B_i)\} < \infty \quad \text{for} \quad i = 0, 1, 2, \ldots . \tag{3.5}$$

By reason of $[E(B_i), X] = \mathbf{0}$, X is reduced by all projection operators $E(B_i)$ and can be represented as the orthogonal sum[15]

$$X = \sum_{i=0}^{\infty} \oplus X_i \tag{3.6}$$

of its s.a. parts X_i which are defined on

$$\mathscr{H}_i \equiv E(B_i)\mathscr{H}$$

by

$$X_i \varphi = X\varphi \quad \text{for all} \quad \varphi \in \mathscr{D}(X_i) = \mathscr{D}(X) \cap \mathscr{H}_i .$$

Equation (3.5) implies that all the s. a. operators X_i on $\mathscr{H}_i$ have a pure point spectrum, since continuous parts of the spectrum are possible only in infinite dimensional Hilbert spaces. Thus, according to (3.6), X also has a pure point spectrum and (3.1) is proven. Equation (3.2) is trivially satisfied. □

Lemma 7. Let $\boldsymbol{A} = \{A_1, \ldots, A_h\}$ be a set of mutually commuting, s.a. operators on a separable Hilbert space $\mathscr{H}$ such that at least one linear combination $\boldsymbol{c}\boldsymbol{A}$ forms a s.a. operator on $\mathscr{H}$ with $\sigma_{\text{ess}}(\boldsymbol{c}\boldsymbol{A}) = \emptyset$. Then all operators $A_r \in \boldsymbol{A}$ have a pure point spectrum.

Proof. Owing to the commutativity of the A_r, there exists a s.a. operator X and appropriate

Borel functions[16] f_r such that

$$A_r = f_r(X) \quad \text{for} \quad r = 1, \ldots, h\,. \tag{3.7}$$

This implies

$$\boldsymbol{c}\,\boldsymbol{A} = \sum_{r=1}^{h} c_r\, f_r(X) \equiv t(X) \tag{3.8}$$

in which t is a Borel function, too. According to Lemma 6 and (3.8), X has a pure point spectrum and, because of (3.2) and (3.7), this proves the assertion. □

Corollary 8. Let $A_1, \ldots, A_n$ be n s.a., mutually commuting operators on a separable Hilbert space at least one of which has an empty essential spectrum. Then all A_r have a pure point spectrum.

Theorem III. Let $\boldsymbol{A} = \{A_1, \ldots, A_h\}$ be a regular set of mutually commuting, semibounded operators. Then the following statements hold in addition to Theorem I(a)—(c) and Corollary 5:

(e) $\mathfrak{R}_{\boldsymbol{A}}$ is a convex, h-dimensional, unbounded subset of $\mathbf{R}^h$ and $Z_{\boldsymbol{A}}$ is a convex function on $\mathfrak{R}_{\boldsymbol{A}}$. Together with a point $\boldsymbol{\alpha}$, $\mathfrak{R}_{\boldsymbol{A}}$ contains also the 2^h*-tant* with vertex $\boldsymbol{\alpha}$ and the edges

$$\{\boldsymbol{\alpha} + c\,\nu_r\,\boldsymbol{e}_r \colon c \geqq 0\}\,.$$

(f) All elements $\boldsymbol{m} \in \mathfrak{M}_A$ satisfy the equations

$$\underline{\sigma}(A_r) < m_r < \overline{\sigma}(A_r) \quad \text{for} \quad r = 1, \ldots, h\,.$$

(g) For $n_r = 0, 1, 2, \ldots$ and $\downarrow X \downarrow_r$ equal to X or $|X|$, the functions

$$Z, \left\langle \prod_{r=1}^{h} \downarrow A_r \downarrow_r^{n_r} \right\rangle \quad \text{and} \quad H[V(\cdot)]$$

are finite and arbitrarily often differentiable in the interior of $\mathfrak{R}_A$. One obtains

$$\langle A_r \rangle = -\,\partial \ln Z / \partial \alpha_r$$
$$\partial \langle A_r \rangle / \partial \alpha_j = \partial \langle A_j \rangle / \partial \alpha_r = -\,\partial^2 \ln Z / \partial \alpha_j\, \partial \alpha_r = -\,\langle (A_r - \langle A_r \rangle)(A_j - \langle A_j \rangle) \rangle\,.$$

(h) According to Theorem I, the information entropy can be expressed also as a function of the "mean values" m_r:

$$\tilde{H}(\boldsymbol{m}) \equiv H(\hat{W}[\boldsymbol{m}]) = \ln Z(\langle \boldsymbol{A} \rangle^{-1}(\boldsymbol{m})) + \boldsymbol{m} \langle \boldsymbol{A} \rangle^{-1}(\boldsymbol{m})\,. \tag{3.9}$$

Under the additional assumptions of the present theorem, $\hat{\mathfrak{M}} \equiv \langle \boldsymbol{A} \rangle (\mathrm{int}\, \mathfrak{R}_{\boldsymbol{A}})$ is an open subset of $\mathbf{R}^h$ and $\tilde{H}$ is continuously differentiable at least twice in the entire $\hat{\mathfrak{M}}$ yielding

$$\partial \tilde{H} / \partial m_k = \langle A_k \rangle^{-1}(\boldsymbol{m}) = \alpha_k\,.$$

Proof ad (e): According to Lemma 4 of [1], the essential spectrum of a regular operator is empty. Hence it follows from the regularity of $\boldsymbol{A}$ and Lemma 7 that all $A_r \in \boldsymbol{A}$ have a pure point spectrum. And this again implies the existence of a (not necessarily unique) basis $\{\varphi_i \colon i \in \mathbf{N}\}$ by means of which the operators A_r can be represented in the form

$$A_r = \sum_{i=1}^{\infty} a_{ri}\, P(\varphi_i) \tag{3.10}$$

where the first index of the eigenvalues a_{ri} indicates the operator A_r to which a_{ri} belongs whereas the second index labels the eigenvalues of A_r. From this common diagonal representation of the A_r it follows

$$Z(\boldsymbol{\alpha}) = \mathrm{Tr}[\exp(-\,\boldsymbol{\alpha} \boldsymbol{A})] = \sum_{i=1}^{\infty} \exp\left(-\sum_{r=1}^{h} \alpha_r\, a_{ri}\right). \tag{3.11}$$

The convexity of the exponential implies

$$\exp\left\{-\,\lambda \sum_{r=1}^{h} \alpha_r\, a_{ri} - (1-\lambda) \sum_{r=1}^{h} \beta_r\, a_{ri}\right\} \leqq \lambda \exp\left(-\sum_{r=1}^{h} \alpha_r\, a_{ri}\right) + (1-\lambda) \exp\left(-\sum_{r=1}^{h} \beta_r\, a_{ri}\right) \tag{3.12}$$

for all $i \in \mathbf{N}$ and $\lambda \in [0, 1]$ and this yields

$$Z\{\lambda\, \boldsymbol{\alpha} + (1-\lambda)\, \boldsymbol{\beta}\} \leqq \lambda\, Z(\boldsymbol{\alpha}) + (1-\lambda)\, Z(\boldsymbol{\beta}) \tag{3.13}$$

for all $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathfrak{R}$ and all $\lambda \in [0, 1]$. Hence $\mathfrak{R}$ is a convex set and Z is a convex function on $\mathfrak{R}$.

Next we choose an arbitrary element $\boldsymbol{\alpha} \in \mathfrak{R}$ and an arbitrary index $k \in \{1, \ldots, h\}$ and consider the points $\boldsymbol{\delta} = \boldsymbol{\alpha} + x\, \nu_k\, \boldsymbol{e}_k$ with $x \geqq 0$ and the corresponding partition function

$$Z(\boldsymbol{\delta}) = \sum_{i=1}^{\infty} \exp\left(-\sum_{r=1}^{h} \alpha_r\, a_{ri}\right) \exp(-\,x\, \nu_k\, a_{ki})\,. \tag{3.14}$$

According to our assumption, A_k is semibounded. If A_k is even bounded, then it follows from (3.14) that

$$Z(\boldsymbol{\delta}) \leqq \exp(x\, \| A_k \|)\, Z(\boldsymbol{\alpha}) < \infty\,.$$

If A_k is unbounded, then we can assume without loss of generality that A_k is bounded from below (i.e. $\nu_k = +\,1$) with the bound c, and we obtain

$$Z(\boldsymbol{\delta}) = \sum_{i=1}^{\infty} \exp\left(-\sum_{r=1}^{h} \alpha_r\, a_{ri}\right) \exp(-\,x\, a_{ki}) < (1 + e^{x|c|})\, Z(\boldsymbol{\alpha}) < \infty\,.$$

Thus $\boldsymbol{\delta}$ lies in $\mathfrak{R}$ for all $1 \leqq k \leqq h$, $x \geqq 0$ and $\boldsymbol{\alpha} \in \mathfrak{R}$. The rest of (e) follows from the convexity of $\mathfrak{R}$.

ad (f): The relation $\underline{\sigma}(A_r) \leqq m_r \leqq \overline{\sigma}(A_r)$ is a trivial consequence of definition (1.3). Let us suppose that there exists an element $\boldsymbol{\alpha} \in \mathfrak{R}$ with $\langle A_k \rangle (\boldsymbol{\alpha}) = m_k = \underline{\sigma}(A_k)$. Because $\mathfrak{M} \subset \mathbf{R}^h$ this implies

$$- \infty < \inf \{a_{ki} \colon i \in \mathbf{N}\} = Z(\boldsymbol{\alpha})^{-1} \sum_{i=1}^{\infty} a_{ki} \exp \left(- \sum_{r=1}^{h} \alpha_r a_{ri} \right)$$

which is obviously wrong. Hence $m_k \neq \underline{\sigma}(A_k)$. Analogously $m_k \neq \overline{\sigma}(A_k)$.

ad $(g)_s$: We define

$$\tilde{\mathfrak{R}} \equiv \left\{ \boldsymbol{\alpha} \in \mathfrak{R}_A : \prod_{r=1}^{h} \alpha_r \neq 0 \right\}.$$

For every $\boldsymbol{\alpha} \in \operatorname{int} \tilde{\mathfrak{R}}$ there exists a real number $\varepsilon \in (0, 1/e]$ such that the closed ball $K(\boldsymbol{\alpha}, \varepsilon)$ is also contained in int $\tilde{\mathfrak{R}}$. If one denotes the greater of the two positive solutions of the equation $x^u = e^{xy}$ (with $u > 0$, $0 < y < u/e$) by $g(u, y)$ and sets $g(0, y) \equiv 1$, then it is easily seen that

$$|x|^u \leqq C^u g(u, \delta) e^{\delta C} e^{\delta x} \tag{3.15}$$

for all real numbers $C \geqq 1$, $x \geqq -C$, $u \geqq 0$, $\delta > 0$ with $\delta < u/e$ in case of $u \neq 0$. Hence, choosing h arbitrary non-negative integers $n_1, \ldots, n_h$, we obtain

$$\begin{aligned} Z(\boldsymbol{\alpha}) & \left\langle \prod_{r=1}^{h} |A_r|^{n_r} \right\rangle (\boldsymbol{\alpha}) \\ &= \sum_{i=1}^{\infty} \left(\prod_{r=1}^{h} |a_{ri}|^{n_r} \right) \exp \left(- \sum_{r=1}^{h} \alpha_r a_{ri} \right) \\ &= \sum_{i=1}^{\infty} \prod_{r=1}^{h} |a_{ri}|^{n_r} \exp(-\alpha_r a_{ri}) \\ &\leqq \sum_{i=1}^{\infty} \prod_{r=1}^{h} S_r^{n_r} g(n_r, \varepsilon)^{n_r} \exp(\varepsilon S_r) \\ &\qquad \cdot \exp\{-a_{ri}(\alpha_r - \varepsilon \nu_r)\} \\ &= \left\{ \prod_{r=1}^{h} S_r^{n_r} g(n_r, \varepsilon)^{n_r} \exp(\varepsilon S_r) \right\} Z(\boldsymbol{\alpha} - \varepsilon \boldsymbol{\nu}) \end{aligned} \tag{3.16}$$

where we introduced the constants

$$S_r \equiv \begin{cases} \max(1, |\overline{\sigma}(A_r)|) & \text{if } \nu_r = -1 \text{ or } \|A_r\| < \infty, \\ \max(1, |\underline{\sigma}(A_r)|) & \text{if } \nu_r = +1 \text{ and } \|A_r\| = \infty. \end{cases}$$

According to our choice of $\boldsymbol{\alpha}$ and ε, the point $\boldsymbol{\alpha} - \varepsilon \boldsymbol{\nu}$ is also contained in int $\tilde{\mathfrak{R}}$ and thus we find that

$$\left\langle \prod_{r=1}^{h} |A_r|^{n_r} \right\rangle (\boldsymbol{\alpha}) < \infty \quad \text{for all} \quad \boldsymbol{\alpha} \in \operatorname{int} \tilde{\mathfrak{R}}$$

and arbitrary non-negative integers $n_1, \ldots, n_h$. If we multiply the inequality (3.12) by

$$\prod_{r=1}^{h} |a_{ri}|^{n_r}$$

and sum over all i, then it follows that

$$F_{\boldsymbol{n}}(\boldsymbol{\alpha}) \equiv Z(\boldsymbol{\alpha}) \left\langle \prod_{r=1}^{h} |A_r|^{n_r} \right\rangle (\boldsymbol{\alpha}) = \sum_{i=1}^{\infty} \prod_{r=1}^{h} |a_{ri}|^{n_r} \exp(-\alpha_r a_{ri})$$

is convex and hence finite in the entire int $\mathfrak{R}$. As a convex function, $F_{\boldsymbol{n}}$ is also continuous on int $\mathfrak{R}$[17]; and as a continuous convergent series of non-negative terms, the serie; $F_{\boldsymbol{n}}$ even converges uniformly on all $K(\boldsymbol{\alpha}, \varepsilon) \subset \operatorname{int} \mathfrak{R}$[18]. This implies that all series $\langle \prod_r \downarrow A_r \downarrow_r^{n_r} \rangle$ are uniformly convergent on all $K(\boldsymbol{\alpha}, \varepsilon) \subset \operatorname{int} \mathfrak{R}$. Finally, it follows from the uniform convergence of all series $\langle \prod_r \downarrow A_r \downarrow_r^{n_r} \rangle$ on all $K(\boldsymbol{\alpha}, \varepsilon) \subset \operatorname{int} \mathfrak{R}$ and from the relation

$$(\partial / \partial \beta_r)^{n_r} \exp(-\alpha_r a_{ri}) = (-1)^{n_r} (a_{ri})^{n_r} \exp(-\alpha_r a_{ri})$$

that Z $(= F_{\boldsymbol{0}})$ as well as all functions $\langle \prod_r \downarrow A_r \downarrow_r^{n_r} \rangle$ have partial derivatives of arbitrary order with respect to the α_r in the entire int $\mathfrak{R}$ and that these derivatives are obtained by differentiating the respective series term by term. Because

$$H[V(\boldsymbol{\alpha})] = \ln Z(\boldsymbol{\alpha}) + \boldsymbol{\alpha} \langle \mathbf{A} \rangle (\boldsymbol{\alpha}),$$

the function $H[V(\cdot)]$ can also be differentiated arbitrarily often with respect to the α_r. As the lowest derivatives we obtain

$$\begin{aligned} \frac{\partial}{\partial \alpha_k} Z(\boldsymbol{\alpha}) &= \sum_{i=1}^{\infty} \frac{\partial}{\partial \alpha_k} \exp \left(- \sum_{r=1}^{h} \alpha_r a_{ri} \right) \\ &= - \sum_{i=1}^{\infty} a_{ki} \exp \left(- \sum_{r=1}^{h} \alpha_r a_{ri} \right) = - Z(\boldsymbol{\alpha}) \langle A_k \rangle (\boldsymbol{\alpha}), \end{aligned}$$

$$\begin{aligned} - \frac{\partial^2}{\partial \alpha_k \, \partial \alpha_j} \ln Z(\boldsymbol{\alpha}) &= \frac{\partial}{\partial \alpha_k} \langle A_j \rangle (\boldsymbol{\alpha}) \\ &= \sum_{i=1}^{\infty} a_{ji} \frac{\partial}{\partial \alpha_k} Z(\boldsymbol{\alpha})^{-1} \exp \left(- \sum_{r=1}^{h} \alpha_r a_{ri} \right) \\ &= Z(\boldsymbol{\alpha})^{-1} \sum_{i=1}^{\infty} \{-a_{ki} a_{ji} + a_{ji} \langle A_k \rangle\} \exp \left(- \sum_{r=1}^{h} \alpha_r a_{ri} \right) \end{aligned}$$

$$= - Z(\boldsymbol{\alpha})^{-1} \sum_{i=1}^{\infty} \{(a_{ki} - \langle A_k \rangle)(a_{ji} - \langle A_j \rangle)\} \cdot \exp\left(- \sum_{r=1}^{h} \alpha_r a_{ri}\right)$$
$$= - \langle (A_k - \langle A_k \rangle)(A_j - \langle A_j \rangle) \qquad (3.17)$$
$$= \langle A_k \rangle \langle A_j \rangle - \langle A_k A_j \rangle .$$

ad (h): We consider the measurable space $(\mathbf{N}, \mathscr{P}(\mathbf{N}))$ where $\mathscr{P}(\mathbf{N})$ is the power set of the set $\mathbf{N}$ of all natural numbers. To every $\boldsymbol{\alpha} \in \Re$ we define a probability $\mathfrak{p}_\alpha$ by

$$(\forall M \in \mathscr{P}(\mathbf{N})) \quad \mathfrak{p}_\alpha(M) \equiv \sum_{i \in M} p_i(\boldsymbol{\alpha}) \qquad (3.18)$$

with

$$p_i(\boldsymbol{\alpha}) = Z(\boldsymbol{\alpha})^{-1} \exp\left\{- \sum_{r=1}^{h} \alpha_r a_{ri}\right\}.$$

Hence the expectation value $E(f)$ of a *random variable* (r.v.) f on the probability space

$$(\mathbf{N}, \mathscr{P}(\mathbf{N}), \mathfrak{p}_\alpha)$$

has the form

$$E_\alpha(f) = \sum_{i=1}^{\infty} f(i)\, p_i(\boldsymbol{\alpha}) . \qquad (3.19)$$

To every polynomial $X = P[A_1, \ldots, A_h]$ in the operators $A_r \in \boldsymbol{A}$ we associate a r.v. $\tilde{X}$ on

$$(\mathbf{N}, \mathscr{P}(\mathbf{N}), \mathfrak{p}_\alpha) \quad \text{by} \quad \tilde{X}(i) = P[a_{1i}, \ldots, a_{hi}].$$

For the expectation values of these r.v.'s we obtain

$$E_\alpha(P[\tilde{A}_1, \ldots, \tilde{A}_h]) = Z(\boldsymbol{\alpha})^{-1} \sum_{i=1}^{\infty} P[a_{1i}, \ldots, a_{hi}] \exp\left\{- \sum_{r=1}^{h} \alpha_r a_{ri}\right\}$$
$$= \operatorname{Tr}\{P[A_1, \ldots, A_h]\, V(\boldsymbol{\alpha})\} \qquad (3.20)$$
$$= \langle P[A_1, \ldots, A_h] \rangle (\boldsymbol{\alpha}) .$$

According to (3.20) and III(g), the *covariance matrix* (Δ_{rs}) of the h r.v.'s $\tilde{A}_1, \ldots, \tilde{A}_h$,

$$\Delta_{rs}(\boldsymbol{\alpha}) \equiv E_\alpha(\tilde{A}_r \tilde{A}_s) - E_\alpha(\tilde{A}_r)\, E_\alpha(\tilde{A}_s) = \Delta_{sr},$$

exists for all $\boldsymbol{\alpha} \in \operatorname{int} \Re$, and the Eqs. (3.17) and (3.20) imply

$$- \Delta_{rs}(\boldsymbol{\alpha}) = \frac{\partial \langle A_r \rangle}{\partial \alpha_s} = \frac{\partial \langle A_s \rangle}{\partial \alpha_r} . \qquad (3.21)$$

Now, the 1-independence of the operators of $\boldsymbol{A}$ is obviously equivalent to the linear independence of the $h+1$ functions $1, \tilde{A}_1, \ldots, \tilde{A}_h$ which *a fortiori* implies the linear independence of the $h+1$ r.v.'s $1, \tilde{A}_1, \ldots, \tilde{A}_h$ $\mathfrak{p}_\alpha$-almost everywhere. According to a standard theorem of probability theory [19], this yields $\| \Delta_{rs} \| > 0$, and from the above considerations we infer

$$(\forall \boldsymbol{\alpha} \in \operatorname{int} \Re) \quad (-1)^h \| \partial \langle A_r \rangle / \partial \alpha_s \| > 0 . \qquad (3.22)$$

From Eq. (3.22) it follows [18] that $\widehat{\mathfrak{M}}$ is an open subset of $\mathbf{R}^h$ and that the functions $\langle A_r \rangle^{-1}$ are continuously differentiable with respect to the m_k in the entire $\widehat{\mathfrak{M}}$. The Eqs. (3.9) and (3.17) finally yield

$$\frac{\partial \tilde{H}}{\partial m_k} = \frac{1}{Z} \sum_{r=1}^{h} \frac{\partial Z}{\partial \alpha_r} \frac{\partial \alpha_r}{\partial m_k} + \alpha_k + \sum_{r=1}^{h} \langle A_r \rangle \frac{\partial \alpha_r}{\partial m_k}$$
$$= \alpha_k = \langle A_k \rangle^{-1}(\boldsymbol{m}) . \qquad \square$$

In Theorem III we have assumed, mainly for simplicity, that all A_r are semibounded. In fact, this assumption can be weakened without injuring the results considerably. In the following theorem, we consider a weakened assumption which is of particular importance for statistical thermodynamics.

Theorem IV. Let $\boldsymbol{A} = \{A_1, \ldots, A_h\}$ be a regular set of mutually commuting operators with the additional property that

(#) the operators $A_2, \ldots, A_h$ are semibounded whereas A_1 is majorized from (at least) one side by a semibounded operator of the form

$$x_1 \mathbf{1} + \sum_{r=2}^{h} x_r A_r \quad \text{with} \quad \boldsymbol{x} \in \mathbf{R}^h .$$

Then, in addition to Theorem I(a)—(c) and Corollary 5, the statements III(f)—(h) remain valid whereas III(e) has to be slightly modified as follows:

($e^{\#}$) $\Re_{\boldsymbol{A}}$ is a convex, h-dimensional, unbounded subset of $\mathbf{R}^h$ and $Z_{\boldsymbol{A}}$ is a convex function on $\Re_{\boldsymbol{A}}$. Together with a point $\boldsymbol{\alpha}$, $\Re_{\boldsymbol{A}}$ contains also the convex hull of the h half- ines

$$\{\boldsymbol{\alpha} + c\, v_r\, \boldsymbol{e}_r \colon c \geqq 0\} \quad \text{for} \quad r = 2, \ldots, h$$

and $\{c\,\boldsymbol{\alpha} \colon c \geqq 1\}$ emanating from $\boldsymbol{\alpha}$.

Proof. ad ($e^{\#}$). The convexity of $\Re$ and the convexity of Z on $\Re$ can be proved just as in Theorem III(e). Let $\boldsymbol{\alpha}$ be an element of $\Re$. Then, according to I(c), the half-line $\{c\,\boldsymbol{\alpha} \colon c \geqq 1\}$ is also contained in $\Re$. For $k \in \{2, \ldots, h\}$ we obtain

$$Z(\boldsymbol{\alpha} + c\, v_k\, \boldsymbol{e}_k) = \sum_{i=1}^{\infty} \exp(- c\, v_k\, a_{ki}) \cdot \exp\left\{- \sum_{r=1}^{h} \alpha_r a_{ri}\right\} < e^{c S_k} Z(\boldsymbol{\alpha}) < \infty$$

for all $c \geqq 0$ and $2 \geqq k \geqq h$. Thus all assertions of ($e^{\#}$) are proved except for dim $\mathfrak{R} = h$. Obviously, dim $\mathfrak{R} = h$ if and only if $\boldsymbol{A}$ is strongly regular. By virtue of the assumed regularity of $\boldsymbol{A}$, $\mathfrak{R}$ contains at least one $\boldsymbol{\alpha} \neq 0$. Setting $z_r \equiv 2 \max(|\alpha_r|, |x_r|) + 1$ we obtain *in case of* $\alpha_1 \neq 0$

$$Z(\boldsymbol{\beta}) \equiv Z\left(\boldsymbol{\alpha} + \sum_{r=2}^{h} \nu_r z_r\right)$$
$$= \sum_{i=1}^{\infty} \exp(-\alpha_1 a_{1i}) \prod_{r=2}^{h} \exp(-\alpha_r a_{ri} - \nu_r z_r a_{ri})$$
$$\leqq \left\{\prod_{r=2}^{h} \exp(z_r S_r)\right\} Z(\boldsymbol{\alpha}) < \infty \quad \text{with} \quad \prod_{r=1}^{h} \beta_r \neq 0.$$

In the case $\alpha_1 = 0$ we need the assumption (#). Without loss of generality, we take

$$A_1 \geqq -(x_1 \mathbf{1} + Y)$$
with (3.23)
$$Y \equiv \sum_{r=2}^{h} x_r A_r, \quad x_1 \geqq 2 \quad \text{and} \quad x_1 \mathbf{1} + Y \geqq \mathbf{0}.$$

Because of (3.10), this implies

$$A_1 \geqq -u\mathbf{1} - \sum_{r=2}^{h} \nu_r z_r A_r$$
with (3.24)
$$u \equiv 2x_1 + \sum_{r=2}^{h} z_r S_r.$$

From $\alpha_1 = 0$ and (3.24) we finally obtain

$$Z(\boldsymbol{\gamma}) \equiv Z\left(\boldsymbol{\alpha} + \boldsymbol{e}_1 + \sum_{r=2}^{h} \nu_r z_r \boldsymbol{e}_r\right)$$
$$= \sum_{i=1}^{\infty} e^{-a_{1i}} \exp\left\{-\sum_{r=2}^{h} \alpha_r a_{ri}\right\} \exp\left\{-\sum_{r=2}^{h} \nu_r z_r a_{ri}\right\}$$
$$\leqq \sum_{i=1}^{\infty} e^{u} \exp\left\{\sum_{r=2}^{h} \nu_r z_r a_{ri}\right\}$$
$$\cdot \exp\left\{-\sum_{r=2}^{h} (\alpha_r a_{ri} + \nu_r z_r a_{ri})\right\}$$
$$= e^{u} Z(\boldsymbol{\alpha}) < \infty \quad \text{with} \quad \prod_{r=1}^{h} \gamma_r \neq 0.$$

This completes the proof of ($e^{\#}$).

ad (f)–(h): In the proofs of III(f)–(h), the stronger assumption that all $A_r \in \boldsymbol{A}$ are semi-bounded has been used at only one point, viz., to prove that all functions $\langle \prod_r |A_r|^{n_r} \rangle$ are finite in the entire int $\tilde{\mathfrak{R}}$ [see Equation (3.16)]. Hence, in order to complete the proof of Theorem IV, we have merely to show that the finiteness of these functions can also be deduced from the assumptions of the present theorem.

From (3.15) it follows that

$$|a_{ri}|^{n_r} \leqq S_r^{n_r} g(n_r, \delta_r)^{n_r} e^{\delta_r S_r} \exp(\delta_r \nu_r a_{ri})$$
$$\equiv D_r(\delta_r) \exp(\delta_r \nu_r a_{ri}) \tag{3.25}$$

for $r = 2, \ldots, h$ and $0 < \delta_r < 1/e$. From

$$x_1 \mathbf{1} + Y \geqq \mathbf{0}$$

we obtain

$$\left|\sum_{r=2}^{h} x_r a_{ri}\right|^{n_1} = |y_i|^{n_1} < u^{n_1} g(n_1, \varepsilon)^{n_1} e^{\varepsilon u} e^{\varepsilon y_i},$$
$$(u + y_i)^{n_1} \leqq u^{n_1} |y_i|^{n_1} + (2 + u)^{n_1} \tag{3.26}$$
$$< (2 + u)^{n_1} + u^{2n_1} g(n_1, \varepsilon)^{n_1} e^{\varepsilon u} e^{\varepsilon y_i}.$$

From (3.15), (3.23) and (3.26) we obtain

$$|a_{1i}|^{n_1} \leqq (u + y_i)^{n_1} g(n_1, \delta_1)^{n_1} \exp\{\delta_1 (u + y_i + a_{1i})\}$$
$$< C_1(\delta_1) \exp(\delta_1 y_i + \delta_1 a_{1i}) + C_2(\delta_1, \varepsilon)$$
$$\cdot \exp(\varepsilon y_i + \delta_1 y_i + \delta_1 a_{1i}) \tag{3.27}$$

for all $0 < \delta_1 < 1/e$, $0 < \varepsilon < 1/e$, where $C_1(\delta_1)$ and $C_2(\delta_1, \varepsilon)$ are finite for $\delta_1, \varepsilon \neq 0$. The Eqs. (3.25) and (3.27) yield

$$Z(\boldsymbol{\alpha}) \left\langle \prod_{r=1}^{h} |A_r|^{n_r} \right\rangle (\boldsymbol{\alpha}) = \sum_{i=1}^{\infty} \left(\prod_{r=1}^{h} |a_{ri}|^{n_r}\right) \exp\left\{-\sum_{r=1}^{h} \alpha_r a_{ri}\right\}$$
$$= C_1(\delta_1) \left(\prod_{r=2}^{h} D_r(\delta_r)\right) \sum_{i=1}^{\infty} \exp\{-(\alpha_1 - \delta_1) a_{1i}\} \exp\left\{-\sum_{r=2}^{h} a_{ri}(\alpha_r - \delta_r \nu_r - \delta_1 x_r)\right\}$$
$$+ C_2(\delta_1, \varepsilon) \left(\prod_{r=2}^{h} D_r(\delta_r)\right) \sum_{i=1}^{\infty} \exp\{-(\alpha_1 - \delta_1) a_{1i}\} \exp\left\{-\sum_{r=2}^{h} a_{ri}(\alpha_r - \delta_r \nu_r - \delta_1 x_r - \varepsilon x_r)\right\}$$

Defining the vectors

$$\boldsymbol{\chi} \equiv \{\delta_1, \nu_2 \delta_2 + x_2 \delta_1, \ldots, \nu_h \delta_h + x_h \delta_1\},$$
$$\boldsymbol{\xi} \equiv \{\delta_1, \nu_2 \delta_2 + x_2 \delta_1 + x_2 \varepsilon, \ldots, \nu_h \delta_h + x_h \delta_1 + x_h \varepsilon\},$$

we finally obtain

$$Z(\boldsymbol{\alpha})\left\langle\prod_{r=1}^{h}|A_r|^{n_r}\right\rangle(\boldsymbol{\alpha}) \leqq C_1(\delta_1)\left(\prod_{r=2}^{h}D_r(\delta_r)\right)Z(\boldsymbol{\alpha}-\boldsymbol{\chi})+C_2(\delta_1,\varepsilon)\left(\prod_{r=2}^{h}D_r(\delta_r)\right)Z(\boldsymbol{\alpha}-\boldsymbol{\xi})\,. \qquad (3.28)$$

Let us now consider a point $\boldsymbol{\alpha}\in \operatorname{int}\mathfrak{R}$. Then there exists also a closed ball $K(\boldsymbol{\alpha},\varrho)$ in $\operatorname{int}\mathfrak{R}$ with $\varrho>0$, and we can choose the parameters $\varepsilon,\delta_1,\ldots,\delta_h$ such that

$$\varepsilon\prod_{r=1}^{h}\delta_r>0 \quad\text{and}\quad |\boldsymbol{\chi}|<\varrho,\quad |\boldsymbol{\xi}|<\varrho\,.$$

Hence all expressions on the right side of (3.28) are finite and we obtain

$$\left\langle\prod_{r=1}^{h}|A_r|^{n_r}\right\rangle(\boldsymbol{\alpha})<\infty$$

for all $\boldsymbol{\alpha}\in\operatorname{int}\mathfrak{R}$ and arbitrary $n_r\in\mathbf{N}\cup\{0\}$.

□

In the Theorems III and IV, the regularity of $\boldsymbol{A}$ has been presupposed. But in general it is difficult to ascertain that a given operator set is regular. Hence in many cases, the following statement may be useful which results immediately from Definition 1 and the above theorems.

Corollary 9. Let $\boldsymbol{A}$ be a set of mutually commuting, s.a. operators which are either all semibounded or satisfy the property (#) of Theorem IV. Then $\boldsymbol{A}$ is regular if and only if $\boldsymbol{A}$ is strongly regular. In particular, $\boldsymbol{A}$ is (strongly) regular, if the operators of $\boldsymbol{A}$ are 1-independent and if at least one of them is regular itself.

The case of commuting A_r is of particular importance for statistical thermodynamics: In the frame of the information theory approach to quantum statistical thermodynamics, *equilibrium* may be characterized[20, 21] as a situation in which the given information refers only to mutually compatible *constants of the motion*. This means: In order for the information (1.1) to characterize an equilibrium state, the operators A_r of (1.1) must not only form a regular set but, in addition, must commute with one another as well as with the Hamiltonian of the system considered. The most important example of an equilibrium state with $h>1$ is furnished by the *macrocanonical ensemble*; it describes the equilibrium state of an open many-particle system with fixed mean values of the energy and the number of particles by means of the "intensive parameters" *temperature* and *chemical potential*. In this case we have the operator set $\boldsymbol{A}_{\text{mac.ens.}}=\{\hat{\mathsf{N}},\hat{\mathsf{H}}\}$ where $\hat{\mathsf{N}}$ denotes the particle number operator and $\hat{\mathsf{H}}$ the Hamiltonian of the open system. $\hat{\mathsf{N}}$ and $\hat{\mathsf{H}}$ commute and, under very general conditions on the interaction between the particles[8], fulfil the property (#) of Theorem IV. Accordingly, the macrocanonical ensemble is covered by Theorem IV provided that $\{\hat{\mathsf{N}},\hat{\mathsf{H}}\}$ is regular. Unfortunately, Corollary 9 does not guarantee the regularity of $\{\hat{\mathsf{N}},\hat{\mathsf{H}}\}$, since for realistic particle interactions neither $\hat{\mathsf{N}}$ nor $\hat{\mathsf{H}}$ is regular. We will thus conclude this paper by specifying a group of physically reasonable conditions under which $\{\hat{\mathsf{N}},\hat{\mathsf{H}}\}$ satisfies all assumptions of Theorem IV.

Notations. We consider an open system consisting of an undetermined number of identical particles. A system with a definite number n of particles is described in the separable Hilbert space $\mathscr{H}_n^{\varepsilon}$ which is the subspace of all symmetric or of all antisymmetric vectors of the n-fold tensor product $\mathscr{H}_1\otimes\cdots\otimes\mathscr{H}_1$ depending on whether the particles are *bosons* or *fermions*[22]. The open system with an undetermined number of particles is accordingly described in the *Fock space*

$$\hat{\mathscr{H}}^{\varepsilon}=\bigoplus_{n=0}^{\infty}\mathscr{H}_n^{\varepsilon};$$

Here we subjoined also the trivial case of particle number zero with $\mathscr{H}_0^{\varepsilon}$ being defined as a one-dimensional Hilbert space, i.e. $\mathscr{H}_0^{\varepsilon}\triangleq\mathbf{C}$. If H_n is the Hamiltonian of the n-particle system, $\mathbf{1}_n$ the identity operator of $\mathscr{H}_n^{\varepsilon}$ and if we set $\mathsf{H}_0=\mathbf{0}_0$, then the Hamiltonian $\hat{\mathsf{H}}$ and the particle number operator $\hat{\mathsf{N}}$ of the open system have the form

$$\hat{\mathsf{H}}=\bigoplus_{n=0}^{\infty}\mathsf{H}_n,\qquad \hat{\mathsf{N}}=\bigoplus_{n=0}^{\infty}n\,\mathbf{1}_n\,. \qquad (3.29)$$

By the *spectral set* $\mathbf{S}(A)$ of a s.a. operator A with a pure point spectrum we understand the set of real numbers which contains only eigenvalues of A and which contains every eigenvalue of A exactly as often as indicated by its multiplicity.

Lemma 10. Let the Hamiltonians H_n satisfy the following conditions:

(1) all H_n have a pure point spectrum;

(2) $(\exists a \in \mathbf{R})$ $(\forall n \in \mathbf{N})$ $\mathsf{H}_n \geqq -a n \mathbf{1}_n$;

(3) the increasingly ordered elements $h_1^{(n)}, h_2^{(n)}, \ldots$ of $\mathbf{S}(\mathsf{H}_n)$ satisfy the equation

$$h_i^{(n)} \geqq b \ln (i/n^q) + c \tag{3.30}$$

for all $i > M_n \equiv d n^p + v$ with $b > 0$, $n \in \mathbf{N}$ and $p, q, d \geqq 0$.

Then $Z_{\hat{\mathsf{N}}, \hat{\mathsf{H}}}(\alpha, \beta) \equiv \Xi(\alpha, \beta)$ is finite for all $\alpha > a\beta > 0$, $\beta > 1/b$, and $\{\hat{\mathsf{N}}, \hat{\mathsf{H}}\}$ has the property (#).

Proof.

$$\begin{aligned}\Xi(\alpha, \beta) - 1 &= \sum_{n=1}^{\infty} \sum_{i=1}^{\infty} e^{-\alpha n} \exp\{-\beta h_i^{(n)}\} \\ &= \sum_{n=1}^{\infty} e^{-\alpha n} \sum_{i=1}^{M_n} \exp\{-\beta h_i^{(n)}\} \\ &\quad + \sum_{n=1}^{\infty} e^{-\alpha n} \sum_{i=M_n+1}^{\infty} \exp\{-\beta h_i^{(n)}\} \\ &\equiv \Sigma_1 + \Sigma_2 .\end{aligned}$$

$$\begin{aligned}\Sigma_1 &< \sum_{n=1}^{\infty} e^{-\alpha n} M_n e^{\beta a n} = \sum_{n=1}^{\infty} e^{-n(\alpha - a\beta)} \{d n^p + v\} \\ &\leqq v \sum_{n=1}^{\infty} [e^{-(\alpha - a\beta)}]^n + d \sum_{n=1}^{\infty} g(p, \varepsilon)^p e^{\varepsilon n} e^{-n(\alpha - a\beta)} \\ &= v \sum_{n=1}^{\infty} [e^{-(\alpha - a\beta)}]^n + d\, g(p, \varepsilon)^p \sum_{n=1}^{\infty} [e^{-\alpha - a\beta - \varepsilon)}]^n\end{aligned}$$

for all $0 \leqq \varepsilon$ if $p = 0$ and for $0 < \varepsilon < p/e$ if $p > 0$. This yields $\Sigma_1 < \infty$ for $\alpha > a\beta$.

$$\begin{aligned}\Sigma_2 &\leqq \sum_{n=1}^{\infty} e^{-\alpha n} \sum_{i=M_n+1}^{\infty} e^{-\beta c} \exp\{-\beta b \ln (i/n^q)\} \\ &\leqq e^{-\beta c} \sum_{n=1}^{\infty} e^{-\alpha n} n^{q b \beta} \sum_{i=1}^{\infty} (1/i)^{b\beta} .\end{aligned}$$

If we now assume $b\beta > 1$, then we obtain

$$Q(\beta) \equiv \sum_{i=1}^{\infty} (1/i)^{b\beta} < \infty \quad \text{and}$$

$$\Sigma_2 \leqq Q(\beta)\, e^{-\beta c} g(q b \beta, \delta)^{q b \beta} \sum_{n=1}^{\infty} [e^{-(\alpha - \delta)}]^n$$

for all $\beta > 1/b$, $0 \leqq \delta$ if $q = 0$ and for $\beta > 1/b$, $0 < \delta < q b \beta/e$ if $q > 0$. This yields $\Sigma_2 < \infty$ for $\alpha > 0$ and $\beta > 1/b$ and completes the proof. □

Acknowledgements

The authors would like to thank Mr. H. Spohn and Mr. J. Voigt for valuable comments and they are, in particular, indebted to Mr. J. Voigt for his assistance in proving Lemma 6.

---

1 W. Bayer and W. Ochs, Z. Naturforsch. 28a, 693 [1973].

1a We take the opportunity to correct a printing error which confused the succession of statements in Theorem II of [1] (p. 698, col. 2): Statement IIb) ends with the third line whereas the lines 4 to 8 form the second part of II c). The passage "II e)" should be inserted between the third and the forth line as the first part of statement II c). The passages "II c)" and "II d)" following below should then be relabelled d) and e).

2 J. Langerholc, J. Math. Phys. **6**, 1210 [1965].

3 In accordance with [2] we use the trace sign "Tr" in two different meanings: (1) If X is a trace class operator, then we define

$$(*) \qquad \mathrm{Tr}(X) \equiv \mathrm{tr}(X) \equiv \sum_i (\varphi_i, X \varphi_i)$$

where the sum runs over an arbitrary basis $\{\varphi_i\}$ of $\mathscr{H}$ and does not depend on the particular basis chosen. (2) If W is a state operator and $A = \int \lambda \, \mathrm{d}E_A(\lambda)$ an arbitrary s.a. operator, then we define

$$(**) \qquad \mathrm{Tr}(WA) \equiv \mathrm{Tr}(AW) \equiv \int_{-\infty}^{\infty} \lambda \, \mathrm{dTr}[W E_A(\lambda)]$$

where the trace in the integrand is defined by (*). The expression (**) exists (with an infinite value admitted) if and only if at least one of the integrals

$$\mathrm{Tr}(WA^+) = \int_0^{\infty} \lambda \, \mathrm{dTr}[W E_A(\lambda)],$$

$$\mathrm{Tr}(WA^-) = \int_{-\infty}^{0} (-\lambda) \, \mathrm{dTr}[W E_A(\lambda)]$$

is finite. This twofold usage is consistent and reasonable as the two functionals (*) and (**) are equal on the intersection of their domains [2].

4 E. H. Wichmann, J. Math. Phys. **4**, 884 [1963].

5 R. S. Ingarden and K. Urbanik, Acta Phys. Polon. **21**, 281 [1962].

6 A. Kossakowski, Bull. Acad. Polon. Sci., Ser. sci. math. astron. phys. **17**, 263 [1969].

7 S. Watanabe, Knowing and Guessing, § 9.5. Wiley, New York-London-Sydney-Toronto 1968.

[8] D. Ruelle, Statistical Mechanics. Benjamin, New York 1969.

[9] If A is bounded, then AW as well as WA are of trace class for all $W \in \mathfrak{W}$ and one has $\mathrm{tr}(WA) = \mathrm{tr}(AW) < \infty$. For an unbounded A, on the other hand, WA is never of trace class because of $\mathscr{D}(WA) = \mathscr{D}(A) \neq \mathscr{H}$ whereas the product AW can still be a trace class operator. Hence if the mean values are to be defined by means of the matrix trace, then one should employ the product AW (and not WA).

[10] For results about the differentiability of Z in special cases of noncommuting operators, we refer to N. N. Bogolubov Jr., Physica **41**, 601 [1969]. — H. D. Maison, Commun. math. Phys. **22**, 166 [1971].

[11] T. Kato, Perturbation Theory for Linear Operators. Springer-Verlag, Berlin-Heidelberg-New York 1966.

[12] I. C. Gohberg and M. G. Krein, Introduction to the Theory of Linear Nonselfadjoint Operators, § III.8. Am. Math. Soc., Providence 1969.

[13] M. Breitenbecker and H.-R. Grümm, Commun. math. Phys. **26**, 276 [1972].

[14] N. Dunford and J. T. Schwartz, Linear Operators, Part II, § XII.2.9. Interscience Publishers, New York 1963.

[15] A. I. Plesner, Spectral Theory of Linear Operators, Vol. II, § 6.7. Frederic Ungar Publishing Co., New York 1969.

[16] F. Riesz and B. Sz.-Nagy, Vorlesungen über Funktionalanalysis, § 130. Deutscher Verlag der Wissenschaften, Berlin 1956.

[17] H. G. Eggleston, Convexity, Chapt. 3. University Press, Cambridge (U.K.) 1958.

[18] A. Ostrowski, Vorlesungen über Differential- und Integralrechnung, Band II (2. Aufl.). Birkhäuser-Verlag, Basel 1961.

[19] H. Richter, Wahrscheinlichkeitstheorie, § V.4. Springer-Verlag, Berlin-Heidelberg-New York 1956.

[20] A. Katz, Principles of Statistical Mechanics. Freeman and Co., San Francisco-London 1967.

[21] A. Hobson, Concepts in Statistical Mechanics. Gordon and Breach, New York-London-Paris 1971.

[22] J. M. Jauch, Foundations of Quantum Mechanics. Addison-Wesley, Reading (Mass.) 1968.