International Journal of Theoretical Physics, Vol. 7, No. 6 (1973), pp. 431–442

Expectation Values of Unbounded Observables

K. KRAUS

Physikalisches Institut der Universität Würzburg, Würzburg, Germany

and

J. SCHRÖTER

Institut für Theoretische Physik der Universität Marburg, Marburg, Germany

Received: 13 *June* 1972

Abstract

For an unbounded quantum mechanical observable A, the expectation value $\langle A\rangle_f$ and the mean square deviation $\Delta_f A$ cannot be defined for all (pure) states f by $\langle A\rangle_f = (f, Af)$ and $(\Delta_f A)^2 = (f, A^2 f) - (f, Af)^2$, respectively. More general definitions are given here, which are also valid for state mixtures (density matrices). A general uncertainty relation for unbounded observables is derived.

1. *Introduction*

The definitions

$$\langle A\rangle_f = (f, Af) \tag{1.1}$$

and

$$\langle A\rangle_W = \operatorname{tr}(AW) \tag{1.2}$$

for expectation values of a quantum mechanical observable A in states described by a unit vector f (pure states) and a density matrix W (state mixtures), respectively, are valid for all f and W only if A is bounded. We will investigate here the modifications of (1.1) and (1.2) which are necessary if A is an unbounded self-adjoint operator.

Although this problem is scarcely relevant for practical applications of quantum theory, it should not be entirely ignored if one wants to formulate the foundations of quantum theory with some mathematical rigor. Moreover, the effort required for a satisfactory generalization of (1.1) and (1.2) is moderate. In order to make the paper more easily readable, we have included a short review of some mathematical results which, although well-known to mathematicians, do not yet belong to the standard machinery of quantum mechanics.

Consider the spectral representation

$$A = \int \lambda\, dE_\lambda \tag{1.3}$$

of an unbounded self-adjoint operator A, corresponding to some physical observable, on a Hilbert space $\mathscr{H}$. From a physical point of view, the unbounded observable A is the idealized limit of bounded observables

$$A_{mn} = \int_{-m}^{n} \lambda \, dE_\lambda + n(1 - E_n) - mE_{-m} \tag{1.4}$$

because each real measuring instrument for A has a finite scale ranging, say, from $-m$ to $+n$, with m and n positive. We thus define

$$\langle A \rangle_f \underset{\text{d.f.}}{=} \lim_{m,n\to\infty} \langle A_{mn} \rangle_f = \lim_{m,n\to\infty} (f, A_{mn} f) \tag{1.5}$$

for all unit vectors $f \in \mathscr{H}$ for which this double limit exists. An alternative choice would be

$$A_{mn} = \int_{-m}^{n} \lambda \, dE_\lambda$$

This modification of (1.4) would not change any of our conclusions.

Now consider, besides A, the operators

$$B_+ = \int_0^\infty \lambda^{1/2} \, dE_\lambda, \qquad B_- = \int_{-\infty}^{0} |\lambda|^{1/2} \, dE_\lambda \tag{1.6}$$

and

$$A_+ = \int_0^\infty \lambda \, dE_\lambda, \qquad A_- = \int_{-\infty}^{0} \lambda \, dE_\lambda \tag{1.7}$$

such that

$$A_+ + A_- = A, \qquad B_+ + B_- = |A|^{1/2}$$

If D_O denotes the domain of definition of an operator O, we have

$$\tilde{D}_A \underset{\text{d.f.}}{=} D_{B_+} \cap D_{B_-} \equiv D_{|A|^{1/2}} \supseteq D_A \tag{1.8}$$

and

$$f \in \tilde{D}_A \leftrightarrow \int |\lambda| \, d(f, E_\lambda f) < \infty \tag{1.9}$$

For non-negative A, $\tilde{D}_A$ is known to be the domain to which the sesquilinear form (f, Ag), $f, g \in D_A$ may be extended by closure (Kato, 1966). A similar result is obtained here:

Lemma 1. The expectation value (1.5) exists if and only if $f \in \tilde{D}_A$. For such f,

$$\langle A \rangle_f = \|B_+ f\|^2 - \|B_- f\|^2 = \int \lambda d(f, E_\lambda f) \tag{1.10}$$

If, in particular, $f \in D_A$, then

$$\langle A \rangle_f = (f, Af) \tag{1.11}$$

Proof. By (1.4) and (1.5), $\langle A \rangle_f$ exists if and only if both

$$a_+ = \lim_{n\to\infty} \left\{ \int_0^n \lambda \, d(f, E_\lambda f) + n(f, [1 - E_n] f) \right\}$$

and

$$a_- = \lim_{m\to\infty}\left\{\int_{-m}^{0} |\lambda|\, d(f, E_\lambda f) + m(f, E_{-m} f)\right\}$$

are finite, $\langle A\rangle_f$ being equal to $a_+ - a_-$. Moreover, (1.6) and the estimates

$$0 \leqslant n(1 - E_n) \leqslant \int_n^\infty \lambda\, dE_\lambda, \qquad 0 \leqslant mE_{-m} \leqslant \int_{-\infty}^{-m} |\lambda|\, dE_\lambda$$

imply that $a_\pm$ is finite and equal to $\|B_\pm f\|^2$ if and only if $f \in D_{B_\pm}$.

In the next section, we will describe a formalism which allows the generalization of this result to state mixtures (density matrices) W.

2. *Description of State Mixtures by Unit Vectors*

Usually, the description of pure states f and mixtures W is unified by using equation (1.2) throughout, with $W = |f\rangle\langle f|$ for a pure state. For our purposes another formalism is more appropriate, which describes all states, pure or mixed, by unit vectors of a suitable Hilbert space.†

Denote by $\mathscr{B}_0$, $\mathscr{B}_1$ and $\mathscr{B}_2$ the set of all bounded operators, the trace class and the Hilbert-Schmidt class of operators on $\mathscr{H}$, respectively.‡ For $B \in \mathscr{B}_0$, $T \in \mathscr{B}_1$ and $S \in \mathscr{B}_2$, $\|B\|$ is the norm, $\operatorname{tr} T$ the trace, $\|T\|_1 = \operatorname{tr}((T^*T)^{1/2})$ the trace norm and $\|S\|_2 = (\operatorname{tr}(S^*S))^{1/2}$ the Hilbert-Schmidt norm. $\mathscr{B}_2$ is a Hilbert space with the inner product

$$(S, S')_2 = \operatorname{tr}(S^* S')$$

Density matrices are positive semidefinite self-adjoint $W \in \mathscr{B}_1$ with $\|W\|_1 = \operatorname{tr} W = 1$. Then $W^{1/2} \in \mathscr{B}_2$ with $\|W^{1/2}\|_2 = 1$, and for a self-adjoint $A \in \mathscr{B}_0$ (a bounded observable) equation (1.2) is equivalent to

$$\langle A\rangle_W = (W^{1/2}, AW^{1/2})_2 \tag{2.1}$$

Thus states may equally well be described by unit vectors $W^{1/2}$ in the Hilbert space $\mathscr{B}_2$, whereas bounded observables $A \in \mathscr{B}_0$ are represented in $\mathscr{B}_2$ by operator multiplication, and expectation values are given by (2.1).

Consider, more generally, the multiplication operator $\mathbf{B}$ on $\mathscr{B}_2$ corresponding to an arbitrary $B \in \mathscr{B}_0$, i.e.,

$$\mathbf{B}S = BS \qquad \text{for all } S \in \mathscr{B}_2 \tag{2.2}$$

$\mathbf{B}$ is bounded on $\mathscr{B}_2$ with norm $\|B\|$. There exists an isomorphism $\mathscr{V}$ of $\mathscr{B}_2$ onto the product Hilbert space $\mathscr{H} \otimes \mathscr{H}$ such that (Dixmier, 1957)

$$\mathscr{V}\mathbf{B}\mathscr{V}^{-1} = B \otimes 1 \tag{2.3}$$

† This formalism has been used before in quantum statistical mechanics (Haag *et al.*, 1967; Wehrl, 1971).

‡ An extensive discussion of $\mathscr{B}_1$ and $\mathscr{B}_2$ may be found in Schatten's book (Schatten, 1960).

Thus, by (2.3), instead of the mapping $B \to \mathbf{B}$ the more familiar mapping $B \to B \otimes 1$ may be studied. The isomorphism $\mathscr{V}$ may be constructed explicitly as follows. Take an arbitrary antiunitary operator V on $\mathscr{H}$, and define

$$\mathscr{V} S(f, g) = f \otimes Vg \tag{2.4}$$

with

$$S(f, g) = |f\rangle\langle g|$$

If extended by linearity and continuity, this $\mathscr{V}$ has the required properties, i.e., it maps $\mathscr{B}_2$ onto $\mathscr{H} \otimes \mathscr{H}$, preserving linear combinations and inner products. Equation (2.3) is easily verified if applied to $f \otimes g$, and the equality then extends to $\mathscr{H} \otimes \mathscr{H}$ because both sides of (2.3) define bounded operators.

For the mapping $B \to B \otimes 1$ we will need only the following very familiar result. If $U(t)$ is a strongly continuous one-parameter group of unitary operators on $\mathscr{H}$ and E_λ the corresponding spectral family of projection operators, i.e.,

$$U(t) = \int \exp(i\lambda t)\, dE_\lambda \tag{2.5}$$

then $U(t) \otimes 1$ and $E_\lambda \otimes 1$ have corresponding properties, and in particular

$$U(t) \otimes 1 = \int \exp(i\lambda t)\, d(E_\lambda \otimes 1) \tag{2.6}$$

This follows immediately from the fact that the mapping $B \to B \otimes 1$ is a sufficiently continuous *-isomorphism (Dixmier, 1957).

Now consider a self-adjoint operator A on $\mathscr{H}$, not necessarily bounded, with spectral representation (1.3), and the corresponding unitary one-parameter group $U(t)$ given by (2.5). We will generalize the definition (2.2) to such A. This may be done in several different ways. By (2.3) and (2.6), $\mathbf{U}(t)$ and $\mathbf{E}_\lambda$ are a unitary one-parameter group and a spectral family, respectively, on $\mathscr{B}_2$, and

$$\mathbf{U}(t) = \int \exp(i\lambda t)\, d\mathbf{E}_\lambda \tag{2.7}$$

This leads to:

Definition 1

$$\mathbf{A} = \int \lambda\, d\mathbf{E}_\lambda \tag{2.8}$$

Equivalently, $\mathbf{A}$ may be defined as the self-adjoint generator of $\mathbf{U}(t)$,

$$\mathbf{U}(t) = \exp(i\mathbf{A}t) \tag{2.9}$$

Another definition, which reduces to (2.2) for bounded A, is:

Definition 2. Let $D_{\mathbf{A}}$ be the set of all $S \in \mathscr{B}_2$ such that $AS \in \mathscr{B}_2$, and

$$\mathbf{A}S = AS \qquad \text{for all } S \in D_{\mathbf{A}} \tag{2.10}$$

Lemma 2. The definitions (2.8) and (2.10) are equivalent.

Proof. Denote by D_1 and D_2 the domains $D_{\mathbf{A}}$ corresponding to (2.8) and (2.10), respectively, and consider an arbitrary complete orthonormal system $\{f_i \mid i = 1, 2 \ldots\}$ in $\mathscr{H}$.

$S \in D_1$ is equivalent to

$$\int \lambda^2 \, d(S, E_\lambda S)_2 = \int \lambda^2 \, d(\mathrm{tr}(S^* E_\lambda S)) = \int \lambda^2 \, d\left[\sum_i (Sf_i, E_\lambda Sf_i)\right] < \infty$$

In this expression, the sum over i may be interchanged with the integration, due to the positivity of the integrand.† Thus $S \in D_1$ is equivalent to

$$\sum_i \int \lambda^2 \, d(Sf_i, E_\lambda Sf_i) < \infty \tag{2.11}$$

for all complete orthonormal systems $\{f_i\}$.

$S \in D_2$ is equivalent to

$$Sf_i \in D_A \tag{2.12}$$

and

$$\sum_i \|ASf_i\|^2 < \infty \tag{2.13}$$

for all complete orthonormal systems $\{f_i\}$. Obviously (2.12) and (2.13) are equivalent to (2.11), thus $D_1 = D_2$.

Finally, $AS \in \mathscr{B}_2$ implies $A(SB) = (AS)B \in \mathscr{B}_2$ for all $B \in \mathscr{B}_0$, and therefore $S \in D_{\mathbf{A}}$ implies $SB \in D_{\mathbf{A}}$. Take f arbitrary and $B = |f\rangle\langle f|$. Then, for $\mathbf{A}$ given by (2.8), we get

$$\begin{aligned}\|f\|^2 (f, \mathbf{A}Sf) &= \mathrm{tr}(B^* \mathbf{A}SB) = (B, \mathbf{A}SB)_2 = \int \lambda \, d(B, E_\lambda SB)_2 \\ &= \int \lambda \, d(\mathrm{tr}(B^* E_\lambda SB)) = \|f\|^2 \int \lambda \, d(f, E_\lambda Sf) = \|f\|^2 (f, ASf)\end{aligned}$$

which implies $\mathbf{A}S = AS$. ∎

Remark 1. If $S \in \mathscr{B}_2$ is self-adjoint, with eigenvalues $s_i \neq 0$, orthonormal eigenvectors f_i and spectral representation

$$S = \sum_i s_i |f_i\rangle\langle f_i|$$

then $S \in D_{\mathbf{A}}$ if and only if $f_i \in D_A$ for all i and

$$\sum_i s_i^2 \|Af_i\|^2 < \infty \tag{2.14}$$

† It is easily shown that, for an ascending sequence of measures μ_n converging to a measure μ and a non-negative function $f(\lambda)$, one has

$$\lim_{n \to \infty} \int f(\lambda) \, d\mu_n(\lambda) = \int f(\lambda) \, d\mu(\lambda)$$

In our case, take

$$f(\lambda) = \lambda^2 \qquad \text{and} \qquad d\mu_n(\lambda) = d\Big(\sum_{i \leqslant n} (Sf_i, E_\lambda Sf_i)\Big)$$

Alternatively, one may decompose the infinite λ interval into finite ones, and apply Helly's theorem (Dunford & Schwartz, 1958) to each of them.

3. *Expectation Values*

The foregoing considerations allow an immediate generalization of Lemma 1. We define, in analogy to (1.5),

$$\langle A\rangle_W \underset{\text{d.f.}}{=} \lim_{m,n\to\infty} \langle A_{mn}\rangle_W = \lim_{m,n\to\infty} \operatorname{tr}(A_{mn} W) = \lim_{m,n\to\infty} (W^{1/2}, A_{mn} W^{1/2})_2 \tag{3.1}$$

Theorem 1. With a given self-adjoint operator A, the following conditions are equivalent for a density matrix

$$W = \sum_i w_i |f_i\rangle\langle f_i|, \qquad w_i > 0, \qquad \sum_i w_i = 1 \tag{3.2}$$

(1) The limit (3.1) exists.

(2) $B_\pm W^{1/2} \in \mathscr{B}_2$, with $B_\pm$ defined by (1.6).

(3) $\int |\lambda|\, d(\operatorname{tr}(E_\lambda W)) < \infty$

(4) $f_i \in \tilde{D}_A$, with $\tilde{D}_A$ defined by (1.8), for all i, and $\sum_i w_i \|B_\pm f_i\|^2 < \infty$

For such W, the expectation value (3.1) of A is also given by

$$\langle A\rangle_W = \|B_+ W^{1/2}\|_2{}^2 - \|B_- W^{1/2}\|_2{}^2 = \int \lambda\, d(\operatorname{tr}(E_\lambda W)) = \sum_i w_i \langle A\rangle_{f_i} \tag{3.3}$$

This follows from Lemma 1 if we replace $\mathscr{H}$ by $\mathscr{B}_2$, A by $\mathbf{A}$, $f \in \mathscr{H}$ by $W^{1/2} \in \mathscr{B}_2$, and use Lemma 2 and Remark 1 of Section 2.

Remark 2. If $AW \in \mathscr{B}_1$, then

$$\langle A\rangle_W = \operatorname{tr}(AW) = \sum_i w_i (f_i, Af_i) \tag{3.4}$$

Proof. According to (1.7), $A_+ \supseteq (1 - E_0)A$, $A_- \supseteq E_0 A$. This implies that $AW \in \mathscr{B}_1$ if and only if $A_\pm W \in \mathscr{B}_1$. $A_\pm W \in \mathscr{B}_1$ implies, with (3.2), $f_i \in D_{A_\pm} \subseteq D_{B_\pm}$ and

$$\pm\operatorname{tr}(A_\pm W) = \pm \sum_i w_i (f_i, A_\pm f_i) = \sum_i w_i \|B_\pm f_i\|^2 < \infty$$

Therefore, $\langle A\rangle_W$ exists, and is equal to

$$\sum_i w_i (\|B_+ f_i\|^2 - \|B_- f_i\|^2) = \operatorname{tr}(A_+ W) + \operatorname{tr}(A_- W) = \operatorname{tr}(AW)$$

Remark 3. If $AW^{1/2} \in \mathscr{B}_2$, then

$$\langle A\rangle_W = (W^{1/2}, AW^{1/2})_2 \tag{3.5}$$

[This is a particular case of equation (3.4).]

For pure states $W = |f\rangle\langle f|$, the particular cases (3.4) and (3.5) both correspond to $f \in D_A$ with equation (1.11).

Remark 4. If $B_\pm W^{1/2}$ belong to $\mathscr{B}_2$, the same is true for the closures $(W^{1/2} B_\pm)^{**}$ of $W^{1/2} B_\pm$, and

$$\langle A\rangle_W = \|(W^{1/2} B_+)^{**}\|_2{}^2 - \|(W^{1/2} B_-)^{**}\|_2{}^2$$

This follows because $(W^{1/2} B_\pm)^* = B_\pm W^{1/2}$, and $\|S^*\|_2 = \|S\|_2$ for all $S \in \mathscr{B}_2$. Similar considerations apply to the cases $AW \in \mathscr{B}_1$ and $AW^{1/2} \in \mathscr{B}_2$, thus leading to the formulae

$$\langle A\rangle_W = \mathrm{tr}\,((WA)^{**})$$

and

$$\langle A\rangle_W = (W^{1/2}, (W^{1/2} A)^{**})_2$$

respectively.

Another property of expectation values for pure states f is given by:

Lemma 3. For $f \in \tilde{D}_A$ and $U(t) = \exp(iAt)$, the function

$$G_f(t) = (f, U(t) f) \tag{3.6}$$

is differentiable for all t, and

$$\frac{1}{i} G_f{}'(0) = \langle A\rangle_f \tag{3.7}$$

Proof. Since $f \in \tilde{D}_A$ means

$$\int |\lambda|\, d(f, E_\lambda f) < \infty$$

the estimate

$$\left|\frac{1}{i\tau}\{\exp[i\lambda(t+\tau)] - \exp(i\lambda t)\}\right| \leqslant |\lambda|$$

and the dominated convergence theorem of Lebesgue (Hewitt & Stromberg, 1965, p. 174) allow us to interchange the limit $\tau \to 0$ with the integration in the expression

$$\lim_{\tau\to 0} \int \frac{1}{i\tau}\{\exp[i\lambda(t+\tau)] - \exp(i\lambda t)\}\, d(f, E_\lambda f) \;\blacksquare$$

If, as above, we replace $\mathscr{H}$ by $\mathscr{B}_2$, A by $\mathbf{A}$ and f by $W^{1/2}$, Lemma 3 leads to:

Theorem 2. For W satisfying the conditions of Theorem 1, the function

$$G_W(t) = \mathrm{tr}\,(U(t)\, W) \tag{3.8}$$

is differentiable for all t, and

$$\frac{1}{i} G_W{}'(0) = \langle A\rangle_W \tag{3.9}$$

The converse of Lemma 3 is not true, i.e., $G_f'(t)$ may exist for some $f \notin \tilde{D}_A$. This is illustrated by the following example. Take $\mathscr{H} = L_2(\mathbb{R}^1)$, $A =$ multiplication by $x \in \mathbb{R}^1$ and

$$f(x) = \begin{cases} \dfrac{1}{\sqrt{2}x} & \text{for } |x| \geqslant 1 \\ 0 & \text{for } |x| < 1 \end{cases}$$

With $(E_\lambda f)(x) = \theta(\lambda - x) f(x)$ we get

$$(f, E_\lambda f) = \begin{cases} -\dfrac{1}{2\lambda} & \text{for } \lambda \leqslant -1 \\ \frac{1}{2} & \text{for } |\lambda| \leqslant 1 \\ 1 - \dfrac{1}{2\lambda} & \text{for } \lambda \geqslant 1 \end{cases}$$

and thus $\int |\lambda| \, d(f, E_\lambda f)$ diverges, whereas

$$G_f(t) = \int \exp(i\lambda t) \, d(f, E_\lambda f) = \cos t + t \left(\int_0^t \frac{\sin s}{s} \, ds - \frac{\pi}{2} \right)$$

is differentiable for all t. One might be tempted to define $\langle A \rangle_W$ by equation (3.9) for all W for which this definition makes sense. However, this more general definition would violate the physical interpretation given in the Introduction.

4. *Mean Square Deviations*

With the same physical motivation as for expectation values, we define the mean square deviation $\Delta_W A$ of an observable A in a state W by

$$(\Delta_W A)^2 \underset{\text{d.f.}}{=} \lim_{m,n \to \infty} (\Delta_W A_{mn})^2 \tag{4.1}$$

with the usual definition

$$(\Delta_W A_{mn})^2 = \langle (A_{mn} - \langle A_{mn} \rangle_W)^2 \rangle_W = \langle A_{mn}^2 \rangle_W - \langle A_{mn} \rangle_W{}^2$$

for the bounded observables A_{mn}.

It is quite plausible that the existence of the limit (4.1) which contains A_{mn} quadratically implies the existence of $\langle A \rangle_W$ as defined by (3.1) which contains A_{mn} only linearly. If this is taken for granted, then instead of (4.1) we need only investigate the existence of

$$\lim_{m,n \to \infty} \langle A_{mn}^2 \rangle_W \tag{4.2}$$

By a slight modification of the methods applied above, we then immediately get the following results.

Lemma 4. The mean square deviation

$$\Delta_f A = \lim_{m,n\to\infty} \Delta_f A_{mn} \tag{4.3}$$

of an observable A in a pure state f exists if and only if $f \in D_A$, and

$$(\Delta_f A)^2 = \|Af\|^2 - \langle A\rangle_f{}^2 = \int \lambda^2 d(f, E_\lambda f) - (\int \lambda\, d(f, E_\lambda f))^2 \tag{4.4}$$

Theorem 3. The following conditions for $W = \sum_i w_i |f_i\rangle\langle f_i|$ are equivalent.

(1) The limit (4.1) exists.
(2) $AW^{1/2} \in \mathscr{B}_2$
(3) $\int \lambda^2 d(\mathrm{tr}(E_\lambda W)) < \infty$
(4) $f_i \in D_A$ for all i, and $\sum_i w_i \|Af_i\|^2 < \infty$

For such W,

$$\begin{aligned}(\Delta_W A)^2 &= \|AW^{1/2}\|_2{}^2 - \langle A\rangle_W{}^2\\ &= \int \lambda^2 d(\mathrm{tr}(E_\lambda W)) - [\int \lambda\, d(\mathrm{tr}(E_\lambda W))]^2\\ &= \sum_i w_i\|Af_i\|^2 - \left(\sum_i w_i (f_i, Af_i)\right)^2 \end{aligned}\tag{4.5}$$

A more rigorous proof of Lemma 4, which avoids the above-mentioned plausibility argument, is given in the Appendix.

Remark 5. Because $\langle A_{mn}^2\rangle_W$ increases monotonically with m and n, we may equally well consider the single limit

$$\lim_{n\to\infty} \langle A_{nn}^2\rangle_W \tag{4.6}$$

instead of the double limit (4.2). The same simplification is allowed in equation (4.1), as shown in the Appendix.

Remark 6. We may also write

$$(\Delta_f A)^2 = \langle A^2\rangle_f - \langle A\rangle_f{}^2 \tag{4.7}$$

and

$$(\Delta_W A)^2 = \langle A^2\rangle_W - \langle A\rangle_W{}^2 \tag{4.8}$$

as suggested by (4.4) and (4.5).

This is due to the fact that the limit (4.6) coincides with $\langle A^2\rangle_W$ as defined by an equation analogous to (3.1). Because A^2 is non-negative, we need in this case an upper cutoff only, say n^2, and obtain

$$\langle A^2\rangle_W = \lim_{n^2\to\infty} \langle [A^2]_{n^2}\rangle$$

with

$$[A^2]_{n^2} = \int_{-n}^{n} \lambda^2 dE_\lambda + n^2(1 - E_n + E_{-n}) \equiv A_{nn}^2$$

Another useful relation is obtained, for $AW^{1/2} \in \mathscr{B}_2$, from

$$\|(A - \langle A \rangle_W)\, W^{1/2}\|_2{}^2 = ((A - \langle A \rangle_W)\, W^{1/2}, (A - \langle A \rangle_W)\, W^{1/2})_2$$
$$= \|AW^{1/2}\|_2{}^2 + \langle A \rangle_W{}^2 - \langle A \rangle_W \{(W^{1/2}, AW^{1/2})_2 + (AW^{1/2}, W^{1/2})_2\}$$

Because **A** is self-adjoint, the term in curly brackets becomes

$$2(W^{1/2}, AW^{1/2})_2 = 2\langle A \rangle_W$$

by (3.5). Thus

$$\Delta_W A = \|(A - \langle A \rangle_W)\, W^{1/2}\|_2 \tag{4.9}$$

The corresponding formula for pure states is

$$\Delta_f A = \|(A - \langle A \rangle_f) f\| \tag{4.10}$$

Equations (4.9) and (4.10) lead to general uncertainty relations. Consider two observables A and B and states W such that both $\Delta_W A$ and $\Delta_W B$ are defined, i.e., $W^{1/2} \in D_{\mathbf{A}} \cap D_{\mathbf{B}}$. For $S, S' \in \mathscr{B}_2$, the Schwarz inequality yields

$$\|S\|_2 \|S'\|_2 \geqslant |(S, S')_2| \geqslant |\mathrm{Im}\,(S, S')_2| = \tfrac{1}{2}|(S, S')_2 - (S', S)_2|.$$

If applied to

$$S = (A - \langle A \rangle_W)\, W^{1/2}, \qquad S' = (B - \langle B \rangle_W)\, W^{1/2}$$

this estimate and (4.9) yield after a short calculation:

Theorem 4. For any two observables A and B, the uncertainty relation

$$\Delta_W A \,.\, \Delta_W B \geqslant \tfrac{1}{2}|(AW^{1/2}, BW^{1/2})_2 - (BW^{1/2}, AW^{1/2})_2| \tag{4.11}$$

holds true for all W, for which both $\Delta_W A$ and $\Delta_W B$ are defined.

The corresponding relation for pure states reads

$$\Delta_f A \,.\, \Delta_f B \geqslant \tfrac{1}{2}|(Af, Bf) - (Bf, Af)| \tag{4.12}$$

for $f \in D_A \cap D_B$. This relation is a generalization of the estimate

$$\Delta_f A \,.\, \Delta_f B \geqslant \tfrac{1}{2}|(f, [A, B]f)|, \qquad [A, B] \equiv AB - BA$$

which can be found in all textbooks but which may be misleading in some cases, because $[A, B]f$ need not be defined for all $f \in D_A \cap D_B$.

For $W = \sum_i w_i |f_i\rangle\langle f_i|$, (4.11) reads

$$\Delta_W A \,.\, \Delta_W B \geqslant \tfrac{1}{2}\left|\sum_i w_i((Af_i, Bf_i) - (Bf_i, Af_i))\right| \tag{4.13}$$

With this formula, previous discussions of uncertainty relations for pure states (Kraus, 1965, 1967, 1970) can now be easily generalized to state mixtures W.

Appendix: Proof of Lemma 4

We introduce the abbreviation

$$[\lambda]_m^n = \begin{cases} n & \text{if } \lambda > n \\ \lambda & \text{if } -m \leqslant \lambda \leqslant n \\ -m & \text{if } \lambda < -m \end{cases}$$

Then

$$A_{mn} = \int [\lambda]_m^n \, dE_\lambda$$

and

$$(\Delta_f A_{mn})^2 = \int ([\lambda]_m^n)^2 \, d\mu(\lambda) - [\int [\lambda]_m^n \, d\mu(\lambda)]^2 \tag{A.1}$$

$$= \tfrac{1}{2} \int\int f_{mn}(\lambda, \lambda') \, d\mu(\lambda) \, d\mu(\lambda') = a_{mn} \tag{A.2}$$

with

$$\mu(\lambda) = (f, E_\lambda f)$$

and

$$f_{mn}(\lambda, \lambda') = ([\lambda]_m^n - [\lambda']_m^n)^2$$

By Fubini's theorem (Hewitt & Stromberg, 1965, p. 385), the double integral in (A.2) may be considered as an integral with the product measure $\mu(\lambda) \times \mu(\lambda')$.

A straightforward but somewhat lengthy calculation leads to

$$f_{mn}(\lambda, \lambda') \leqslant f_{m'n'}(\lambda, \lambda')$$

for $n \leqslant n'$, $m \leqslant m'$, which implies

$$a_{mn} \leqslant a_{m'n'}$$

Therefore the double limit

$$\lim_{m,n \to \infty} a_{mn}$$

which occurs in the definition (4.1) of $(\Delta_f A)^2$ may be replaced by a single limit, i.e.,

$$(\Delta_f A)^2 = \lim_{n \to \infty} a_{nn} \tag{A.3}$$

[This is also true with $\mu(\lambda) = \operatorname{tr}(E_\lambda W)$, i.e., for (4.1) with a general state W.]

Moreover, the non-decreasing sequence of functions $f_{nn}(\lambda, \lambda')$ approaches $(\lambda - \lambda')^2$ for $n \to \infty$. By the theorem of B. Levi (Hewitt & Stromberg, 1965, p. 172) we have

$$\lim_{n \to \infty} a_{nn} = \tfrac{1}{2} \int (\lambda - \lambda')^2 \, d(\mu(\lambda) \times \mu(\lambda')) < \infty \tag{A.4}$$

if $(\Delta_f A)^2$ exists.

Again by Fubini's theorem, (A.4) may be evaluated as a double integral. Thus $g(\lambda) = \lambda - \lambda'$ is square-integrable, for almost all λ', with respect to $\mu(\lambda)$. Because $h(\lambda) = \lambda' = \text{const.}$ is square-integrable too, the same is true for $g(\lambda) + h(\lambda) = \lambda$, i.e., $f \in D_A$.

Vice versa, for $f \in D_A$ both terms in (A.1) converge separately for $m = n \to \infty$, thus leading immediately to equation (4.4).

Acknowledgement

We are very grateful to Dr. L. J. Landau (Zürich) for useful comments about a previous version of this paper.

References

Dixmier, J. (1957). *Les algèbres d'opérateurs dans l'espace Hilbertien* (*Algèbres de von Neumann*). Gauthier-Villars, Paris.

Dunford, N. and Schwartz, J. T. (1958). *Linear Operators*, Vol. I. Interscience, New York, p. 391.

Haag, R., Hugenholtz, N. M. and Winnink, M. (1967). *Communications in Mathematical Physics*, **5**, 215.

Hewitt, E. and Stromberg, K. (1965). *Real and Abstract Analysis*. Springer, Berlin, Heidelberg and New York.

Kato, T. (1966). *Perturbation Theory for Linear Operators*. Springer, Berlin, Heidelberg and New York.

Kraus, K. (1965). *Zeitschrift für Physik*, **188**, 374.

Kraus, K. (1967). *Zeitschrift für Physik*, **201**, 134.

Kraus, K. (1970). *American Journal of Physics*, **38**, 1489.

Schatten, R. (1960). *Norm Ideals of Completely Continuous Operators*. Springer, Berlin, Göttingen and Heidelberg.

Wehrl, A. (1971). *Communications in Mathematical Physics*, **23**, 319.