International Journal of Theoretical Physics (2023) 62:94
https://doi.org/10.1007/s10773-023-05342-8

# Real-Valued Observables and Quantum Uncertainty

**Stan Gudder**[1]

Received: 3 January 2023 / Accepted: 24 March 2023 / Published online: 28 April 2023


**Abstract**
We first present a generalization of the Robertson-Heisenberg uncertainty principle. This generalization applies to mixed states and contains a covariance term. For faithful states, we characterize when the uncertainty inequality is an equality. We next present an uncertainty principle version for real-valued observables. Sharp versions and conjugates of real-valued observables are considered. The theory is illustrated with examples of dichotomic observables. We close with a discussion of real-valued coarse graining.

**Keywords** Uncertainty principle · Real-valued observables · Coarse-grain

## 1 Introduction

One of the basic principles of quantum theory is the Robertson-Heisenberg uncertainty inequality [4, 7]

$$\Delta_\psi(A)\Delta_\psi(B) \geq \frac{1}{4}|\langle\psi, [A, B]\psi\rangle|^2 \tag{1.1}$$

where $A, B$ are self-adjoint operators and $\psi$ is a vector state on a Hilbert space. The inequality (1.1) is usually applied to position and momentum operators $A, B$ in which case $|\langle\psi, [A, B]\psi\rangle|^2 = \hbar^2$ where $\hbar$ is Planck's constant. In this situation, $A$ and $B$ are unbounded operators, but for mathematical rigor we shall only deal with bounded operators. However, our results can be extended to the unbounded case by considering a dense subspace common to the domains of $A$ and $B$. In this paper, we derive a generalization of (1.1). This generalization applies to mixed states and contains an additional covariance term that results in a stronger inequality.

The main result in Section 2 is an uncertainty principle for observable operators. This principle contains four parts: a commutator term, a covariance term, a correlation term and a product of variances term. This last term is sometimes called a product of uncertainties. In Section 2 we also characterize, for faithful states, when the uncertainty inequality is an

Dedicated to the memory of Richard Greechie (1941–2022).
The author's cherished friend, long time colleague and collaborator.

✉ Stan Gudder
sgudder@du.edu

1 Department of Mathematics, University of Denver, Denver 80208, CO, USA

equality. Section 3 introduces the concept of a real-valued observable. If $\rho$ is a state and $A$ is a real-valued observable, we define the $\rho$-average, $\rho$-deviation and $\rho$-variance of $A$. If $B$ is another real-valued observable, we define the $\rho$-correlation and $\rho$-covariance of $A, B$. An uncertainty principle for real-valued observables is given in terms of these concepts. An important role is played by the stochastic operator $\widetilde{A}$ for $A$. In Section 3 we also define the sharp version of a real-valued observable and characterize when two real-valued observables have the same sharp version

Section 4 illustrates the theory presented in Section 3 with two examples. The first example considers two dichotomic arbitrary real-valued observables. The second example considers the special case of two noisy spin observables. In this case, the uncertainty inequality becomes very simple. Section 5 discusses real-values coarse graining of observables.

## 2 Quantum Uncertainty Principle

For a complex Hilbert space $H$, we denote the set of bounded linear operators by $\mathcal{L}(H)$ and the set of bounded self-adjoint operators by $\mathcal{L}_S(H)$. A positive trace-class operator with trace one is a *state* and the set of states on $H$ is denoted by $\mathcal{S}(H)$. A state $\rho$ is *faithful* if $\mathrm{tr}\,(\rho C^*C) = 0$ for $C \in \mathcal{L}(H)$ implies that $C = 0$. For $\rho \in \mathcal{S}(H)$ and $C, D \in \mathcal{L}(H)$ we define the sesquilinear form $\langle C, D\rangle_\rho = \mathrm{tr}\,(\rho C^*D)$.

**Lemma 2.1** (i) If $C \in \mathcal{L}(H)$, $\rho \in \mathcal{S}(H)$, then $\mathrm{tr}\,(\rho C^*) = \overline{\mathrm{tr}\,(\rho C)}$. (ii) The form $\langle \bullet, \bullet\rangle_\rho$ is a positive semi-definite inner product. (iii) A state $\rho$ is faithful if and only if $\langle \bullet, \bullet\rangle_\rho$ is an inner product

***Proof*** (i) If $D$ is a trace-class operator and $\{\phi_i\}$ is an orthonormal basis for $H$, we have

$$\mathrm{tr}\,(D^*) = \sum_i \langle \phi_i, D^*\phi_i\rangle = \sum_i \overline{\langle D^*\phi_i, \phi_i\rangle} = \sum_i \overline{\langle \phi_i, D\phi_i\rangle} = \overline{\mathrm{tr}\,(D)}$$

Hence,

$$\mathrm{tr}\,(\rho C^*) = \mathrm{tr}\,\left[(C\rho)^*\right] = \overline{\mathrm{tr}\,(C\rho)} = \overline{\mathrm{tr}\,(\rho C)}$$

(ii) Applying (i), we have

$$\overline{\langle C, D\rangle_\rho} = \overline{\mathrm{tr}\,(\rho C^*D)} = \mathrm{tr}\,\left[\rho(C^*D)^*\right] = \mathrm{tr}\,(\rho D^*C) = \langle D, C\rangle_\rho$$

Moreover, since $C^*C \geq 0$ we have $\langle C, C\rangle_\rho = \mathrm{tr}\,(\rho C^*C) \geq 0$. Hence, $\langle \bullet, \bullet\rangle_\rho$ is a positive semi-definite inner product. (iii) If $\langle \bullet, \bullet\rangle_\rho$ is an inner product, then

$$\langle C, C\rangle_\rho = \mathrm{tr}\,(\rho C^*C) = 0$$

implies $C = 0$ so $\rho$ is faithful. Conversely, if $\rho$ is faithful, then

$$\mathrm{tr}\,(\rho C^*C) = \langle C, C\rangle_\rho = 0$$

implies $C = 0$ so $\langle \bullet, \bullet\rangle_\rho$ is an inner product □

For $A \in \mathcal{L}_S(H)$ and $\rho \in \mathcal{S}(H)$, the $\rho$-*average* (or $\rho$-*expectation*) of $A$ is $\langle A\rangle_\rho = \mathrm{tr}\,(\rho A)$ and $\rho$-*deviation* of $A$ is $D_\rho(A) = A - \langle A\rangle_\rho I$ where $I$ is the identity map on $H$. If $A, B \in \mathcal{L}_S(H)$, the $\rho$-*correlation* of $A, B$ is

$$\mathrm{Cor}_\rho(A, B) = \mathrm{tr}\,\left[\rho D_\rho(A)D_\rho(B)\right]$$

Although $\mathrm{Cor}_\rho(A, B)$ need not be a real number, it is easy to check that $\overline{\mathrm{Cor}_\rho(A, B)} = \mathrm{Cor}_\rho(B, A)$. We say that $A$ and $B$ are *uncorrelated* if $\mathrm{Cor}_\rho(A, B) = 0$. The *$\rho$-covariance* of $A$, $B$ is $\Delta_\rho(A, B) = \mathrm{Re}\,\mathrm{Cor}_\rho(A, B)$ and the *$\rho$-variance* of $A$ is

$$\Delta_\rho(A) = \Delta_\rho(A, A) = \mathrm{Cor}_\rho(A, A) = \mathrm{tr}\left[\rho D_\rho(A)^2\right]$$

It is straightforward to show that

$$\mathrm{Cor}_\rho(A, B) = \mathrm{tr}\,(\rho AB) - \langle A\rangle_\rho \langle B\rangle_\rho \tag{2.1}$$

$$\Delta_\rho(A, B) = \mathrm{Re}\,\mathrm{tr}\,(\rho AB) - \langle A\rangle_\rho \langle B\rangle_\rho \tag{2.2}$$

$$\Delta_\rho(A) = \langle A^2\rangle_\rho - \langle A\rangle_\rho^2 \tag{2.3}$$

We see from (2.1) that $A$ and $B$ are $\rho$-uncorrelated if and only if $\mathrm{tr}\,(\rho AB) = \langle A\rangle_\rho \langle B\rangle_\rho$. We say that $A$ and $B$ *commute* if their commutant $[A, B] = AB - BA = 0$.

**Example 1** *In the tensor product $H_1 \otimes H_2$ let $\rho = \rho_1 \otimes \rho_2 \in \mathcal{S}(H_1 \otimes H_2)$ be a product state and let $A_1 \in \mathcal{L}_S(H_1)$, $A_2 \in \mathcal{L}_S(H_2)$. Then $A = A_1 \otimes I_2$, $B = I_1 \otimes A_2 \in \mathcal{L}_S(H_1 \otimes H_2)$ are $\rho$-uncorrelated because*

$$\begin{aligned}\mathrm{tr}\,(\rho AB) &= \mathrm{tr}\,[\rho_1 \otimes \rho_2 (A_1 \otimes I_2)(I_2 \otimes A_2)] = \mathrm{tr}\,[\rho_1 \otimes \rho_2 (A_1 \otimes A_2)]\\ &= \mathrm{tr}\,(\rho_1 A_1 \otimes \rho_2 A_2) = \mathrm{tr}\,(\rho_1 A_1)\mathrm{tr}\,(\rho_2 A_2)\\ &= \mathrm{tr}\,(\rho_1 \otimes \rho_2 A_1 \otimes I_2)\mathrm{tr}\,(\rho_1 \otimes \rho_2 I_1 \otimes A_2) = \langle A\rangle_\rho \langle B\rangle_\rho\end{aligned}$$

*This shows that $A$, $B$ are $\rho$-uncorrelated for any product state $\rho$. Of course, $[A, B] = 0$ in this case. However, there are examples of noncommuting operators that are uncorrelated. For instance, on $H = \mathbb{C}^2$ let $\alpha = \begin{bmatrix}1\\0\end{bmatrix}$, $\phi = \begin{bmatrix}0\\1\end{bmatrix}$, $\psi = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix}$. With $\rho = |\alpha\rangle\langle\alpha|$, $A = |\phi\rangle\langle\phi|$, $B = |\psi\rangle\langle\psi|$ we have*

$$\mathrm{tr}\,(\rho AB) = \langle A\rangle_\rho \langle B\rangle_\rho = 0$$

*Hence, $A$, $B$ are $\rho$-uncorrelated. However,*

$$\begin{aligned}AB &= \langle\phi, \psi\rangle |\phi\rangle\langle\psi| = \tfrac{1}{\sqrt{2}}|\phi\rangle\langle\psi|\\ BA &= \langle\psi, \phi\rangle |\psi\rangle\langle\phi| = \tfrac{1}{\sqrt{2}}|\psi\rangle\langle\phi|\end{aligned}$$

*so* $[A, B] \neq 0$. □

We now present our main result.

**Theorem 2.2** *If $A, B \in \mathcal{L}_S(H)$ and $\rho \in \mathcal{S}(H)$, then (i) $\frac{1}{4}|\mathrm{tr}\,(\rho[A, B])|^2 + \left[\Delta_\rho(A, B)\right]^2 = |\mathrm{Cor}_\rho(A, B)|^2$*

*(ii) $\frac{1}{4}|\mathrm{tr}\,(\rho[A, B])|^2 + \left[\Delta_\rho(A, B)\right]^2 \leq \Delta_\rho(A)\Delta_\rho(B)$*

***Proof*** (i) Applying Lemma 2.1 we have

$$\begin{aligned}\mathrm{tr}\,([A, B]) &= \mathrm{tr}\,(\rho AB) - \mathrm{tr}\,(\rho BA) = \mathrm{tr}\,(\rho AB) - \overline{\mathrm{tr}\,[\rho(BA)^*]}\\ &= \mathrm{tr}\,(\rho AB) - \overline{\mathrm{tr}\,(\rho A^* B^*)} = \mathrm{tr}\,(\rho AB) - \overline{\mathrm{tr}\,(\rho AB)}\\ &= 2i\,\mathrm{Im}\,[\mathrm{tr}\,(\rho AB)]\end{aligned} \tag{2.4}$$

From (2.2) and (2.4) we obtain

$$\begin{aligned}\frac{1}{4}|\text{tr }(\rho[A,B])|^2+\left[\Delta_\rho(A,B)\right]^2&=[\text{Im }(\rho AB)]^2+\left[\text{Re tr }(\rho AB)-\langle A\rangle_\rho\langle B\rangle_\rho\right]^2\\&=|\text{Re tr }(\rho AB)-\langle A\rangle_\rho\langle B\rangle_\rho+i\text{ Im tr }(\rho AB)|^2\\&=|\text{tr }(\rho AB)-\langle A\rangle)_\rho\langle B\rangle_\rho|^2=|\text{Cor}_\rho(A,B)|^2\end{aligned}$$

(ii) Applying Lemma 2.1(ii), the form $\langle C,D\rangle_\rho=\text{tr }(\rho C^*D)$ is a positive semi-definite inner product. Hence, Schwarz's inequality holds and we have

$$\begin{aligned}|\text{Cor}_\rho(A,B)|^2&=|\text{tr }\left[\rho D_\rho(A)D_\rho(B)\right]|^2=|\langle D_\rho(A),D_\rho(B)\rangle_\rho|^2\\&\le\langle D_\rho(A),D_\rho(A)\rangle_\rho\langle D_\rho(B),D_\rho(B)\rangle_\rho=\text{tr }\left[\rho D_\rho(A)^2\right]\text{tr }\left[\rho D_\rho(B)^2\right]\\&=\Delta_\rho(A)\Delta_\rho(B)\end{aligned}$$

□

We call Theorem 2.2(i) the *uncertainty equation* and Theorem 2.2(ii) the *uncertainty inequality*. Together, they are called the *uncertainty principle*. Notice that Theorem 2.2(ii) is a considerable strengthening of the usual Robertson-Heisenberg inequality (1.1) since it contains the term $\left[\Delta_\rho(A,B)\right]^2$ and it applies to arbitrary states. Thus, even when $[A,B]=0$ we still have an uncertainty relation

$$\left[\Delta_\rho(A,B)\right]^2=|\text{tr }\left[\rho D_\rho(A)D_\rho(B)\right]|^2\le\Delta_\rho(A)\Delta_\rho(B)$$

**Lemma 2.3** A state $\rho$ is faithful if and only if the eigenvalues of $\rho$ are positive.

***Proof*** Suppose the eigenvalues $\lambda_i$ of $\rho$ are positive with corresponding normalized eigenvectors $\phi_i$. Then we can write $\rho=\sum\lambda_i|\phi_i\rangle\langle\phi_i|$ for the orthonormal basis $\{\phi_i\}$. For any $A\in\mathcal{L}(H)$ we obtain

$$\text{tr }(\rho A^*A)=\sum\lambda_i\text{tr }\left(|\phi_i\rangle\langle\phi_i|A^*A\right)=\sum\lambda_i\langle A\phi_i,A\phi_i\rangle=\sum\lambda_i\|A\phi_i\|^2$$

Hence, $\text{tr }(\rho A^*A)=0$ implies $A\phi_i=0$ for all $i$. It follows that $A=0$. Conversely, if 0 is an eigenvalue of $\rho$ and $\phi$ is a corresponding unit eigenvector, then setting $P_\phi=|\phi\rangle\langle\phi|$ we have

$$\text{tr }(\rho P_\phi^*P_\phi)=\text{tr }(\rho P_\phi)=\langle\phi,\rho\phi\rangle=0$$

But $P_\phi\ne0$ so $\rho$ is not faithful. □

**Theorem 2.4** *If $\rho$ is faithful. then the following statements are equivalent. (i) The uncertainty inequality of Theorem 2.2(ii) is an equality. (ii) $D_\rho(B)=\alpha D_\rho(A)$ for $\alpha\in\mathbb{R}$. (iii) $B=\alpha A+\beta I$ for $\alpha,\beta\in\mathbb{R}$. If one of the conditions holds, then*

$$\left[\Delta_\rho(A,B)\right]^2=|\text{Cor}_\rho(A,B)|^2=\Delta_\rho(A)\Delta_\rho(B)\tag{2.5}$$

***Proof*** (i)⇒(ii) If the uncertainty inequality is an equality, then

$$|\text{tr }\left[\rho D_\rho(A)D_\rho(B)\right]|^2=\Delta_\rho(A)\Delta_\rho(B)\tag{2.6}$$

We can rewrite (2.6) as

$$|\langle D_\rho(A),D_\rho(B)\rangle_\rho|^2=\langle D_\rho(A),D_\rho(A)\rangle_\rho\langle D_\rho(B),D_\rho(B)\rangle_\rho$$

Since we have equality in Schwarz's inequality and $\langle \bullet, \bullet \rangle_\rho$ is an inner product, it follows that $D_\rho(B) = \alpha D_\rho(A)$ for some $\alpha \in \mathbb{C}$. Since $D_\rho(B)^* = D_\rho(B)$ and $D_\rho(A)^* = D_\rho(A)$ we conclude that $\alpha \in \mathbb{R}$. (ii)⇒(iii) If $D_\rho(B) = \alpha D_\rho(A)$ for $\alpha \in \mathbb{R}$, we have

$$B - \langle B \rangle_\rho I = \alpha \left( A - \langle A \rangle_\rho I \right)$$

Hence, letting $\beta = \langle B \rangle_\rho - \alpha \langle A \rangle_\rho$ we have $B = \alpha A + \beta I$. Since $A, B \in \mathcal{L}_S(H)$ and $\alpha \in \mathbb{R}$, we have that $\beta \in \mathbb{R}$. (iii)⇒(i) If (iii) holds, then

$$\langle B \rangle_\rho = \mathrm{tr}\,(\rho B) = \alpha \mathrm{tr}\,(\rho A) + \beta = \alpha \langle A \rangle_\rho + \beta$$

Hence, $\beta = \langle B \rangle_\rho - \alpha \langle A \rangle_\rho$ so that

$$\begin{aligned} D_\rho(B) &= B - \langle B \rangle_\rho I = \alpha A + \beta I - \langle B \rangle_\rho I \\ &= \alpha A + \langle B \rangle_\rho I - \alpha \langle A \rangle_\rho I - \langle B \rangle_\rho I = \alpha D_\rho(A) \end{aligned}$$

Thus, (ii) holds and it follows that (2.6) holds and this implies (i). Equation (2.5) holds because (2.6) holds. □

**Example 2** *The simplest faithful state when* $\dim H = n < \infty$ *is* $\rho = I/n$. *Then* $\langle A, B \rangle_\rho = \frac{1}{n}\mathrm{tr}\,(A^* B)$ *which is essentially the Hilbert-Schmidt inner product* $\langle A, B \rangle_{HS} = \mathrm{tr}\,(A^* B)$. *In this case for* $A, B \in \mathcal{L}_S(H)$ *we have* $\langle A \rangle_\rho = \frac{1}{n}\mathrm{tr}\,(A)$, $D_\rho(A) = A - \frac{1}{n}\mathrm{tr}\,(A) I$. *The other statistical concepts become:*

$$\begin{aligned} \mathrm{Cor}_\rho(A, B) &= \mathrm{tr}\,\left[\rho D_\rho(A) D_\rho(B)\right] = \frac{1}{n}\mathrm{tr}\,(AB) - \frac{1}{n^2}\mathrm{tr}\,(A)\mathrm{tr}\,(B) \\ \Delta_\rho(A, B) &= \frac{1}{n}\mathrm{Re}\,\mathrm{tr}\,(AB) - \frac{1}{n^2}\mathrm{tr}\,(A)\mathrm{tr}\,(B) \\ \Delta_\rho(A) &= \frac{1}{n}\mathrm{tr}\,(A^2) - \left[\frac{1}{n}\mathrm{tr}\,(A)\right]^2 \\ \mathrm{tr}\,(\rho\,[A, B]) &= \frac{2i}{n}\mathrm{Im}\,\mathrm{tr}\,(AB) \end{aligned}$$

*The uncertainty principle is given by:*

$$\begin{aligned} [\mathrm{Im}\,\mathrm{tr}\,(AB)]^2 + \left[\mathrm{Re}\,\mathrm{tr}\,(AB) - \frac{1}{n}\mathrm{tr}\,(A)\mathrm{tr}\,(B)\right]^2 &= \left|\mathrm{tr}\,(AB) - \frac{1}{n}\mathrm{tr}\,(A)\mathrm{tr}\,(B)\right|^2 \\ &\le \left[\mathrm{tr}\,(A^2) - \frac{1}{n}\mathrm{tr}\,(A)^2\right]\left[\mathrm{tr}\,(B^2) - \frac{1}{n}\mathrm{tr}\,(B)^2\right] \end{aligned}$$ □

## 3 Real-Valued Observables

An *effect* is an operator $C \in \mathcal{L}_S(H)$ that satisfies $0 \le C \le I$ [1, 4–6]. Effects are thought of as two outcomes *yes-no* measurements. When the result of measuring $C$ is *yes* , we say that $C$ *occurs* and when the result is *no* , then $C$ *does not occur*. A *real-valued observable* is a finite set of effects $A = \{A_x : x \in \Omega_A\}$ where $\sum_{x \in \Omega_A} A_x = I$ and $\Omega_A \subseteq \mathbb{R}$ is the *outcome space* for $A$. The effect $A_x$ occurs when the result of measuring $A$ is the outcome $x$. The condition $\sum_{x \in \Omega_A} A_x = I$ specifies that one of the possible outcomes of $A$ must occur. An observable is also called a *positive operator-valued measure* (POVM). We say $A$ is *sharp* if $A_x$ is a projection for all $x \in \Omega_A$ and in this case, $A$ is a *projection-valued measure* [4, 7].

Corresponding to $A$ we have the *stochastic operator* $\widetilde{A} \in \mathcal{L}(H)$ given by $\widetilde{A} = \sum_{x \in \Omega_A} xA_x$. Notice that we need $A$ to be real-valued in order for $\widetilde{A}$ to exist.

We now apply the theory presented in Section 2 to real-valued observables. For $\rho \in \mathcal{S}(H)$, the $\rho$-*average* (or $\rho$-*expectation*) of $A$ is defined by

$$\langle A\rangle_\rho = \langle \widetilde{A}\rangle_\rho = \mathrm{tr}\,(\rho\widetilde{A}) = \sum_{x\in\Omega_A} x\mathrm{tr}\,(\rho A_x) \tag{3.1}$$

We interpret $\mathrm{tr}\,(\rho A_x)$ as the probability that a measurement of $A$ results in the outcome $x$ when the system is in state $\rho$. Thus, (3.1) says that the $\rho$-average of $A$ is the sum of its outcomes times the probabilities these outcomes occur. We define the $\rho$-*deviation* of $A$ by

$$\begin{aligned} D_\rho(A) = D_\rho(\widetilde{A}) = \widetilde{A} - \langle A\rangle_\rho I &= \sum_{x\in\Omega_A} xA_x - \sum_{x\in\Omega_A} x\mathrm{tr}\,(\rho A_x)I \\ &= \sum_{x\in\Omega_A} x\left[A_x - \mathrm{tr}\,(\rho A_x)I\right] \end{aligned}$$

If $A, B$ are real-valued observables, the $\rho$-*correlation* of $A, B$ is $\mathrm{Cor}_\rho(A,B) = \mathrm{Cor}_\rho(\widetilde{A},\widetilde{B})$, $\rho$-*covariance* of $A, B$ is $\Delta_\rho(A,B) = \Delta_\rho(\widetilde{A},\widetilde{B})$ and the $\rho$-*variance* of $A$ is $\Delta_\rho(A) = \Delta_\rho(\widetilde{A})$. Applying (2.1) we obtain

$$\begin{aligned} \mathrm{Cor}_\rho(A,B) &= \mathrm{tr}\,(\rho\widetilde{A}\widetilde{B}) - \langle\widetilde{A}\rangle_\rho\langle\widetilde{B}\rangle_\rho = \mathrm{tr}\left(\rho\sum_{x,y} xyA_xB_y\right) - \langle\widetilde{A}\rangle_\rho\langle\widetilde{B}\rangle_\rho \\ &= \sum_{x,y} xy\left[\mathrm{tr}\,(\rho A_xB_y) - \mathrm{tr}\,(\rho A_x)\mathrm{tr}\,(\rho B_y)\right] \end{aligned} \tag{3.2}$$

It follows that

$$\Delta_\rho(A,B) = \sum_{x,y} xy\left[\mathrm{Re}\,\mathrm{tr}\,(\rho A_xB_y) - \mathrm{tr}\,(\rho A_x)\mathrm{tr}\,(\rho B_y)\right] \tag{3.3}$$

and

$$\Delta_\rho(A) = \sum_{x,y} xy\left[\mathrm{tr}\,(\rho A_xA_y) - \mathrm{tr}\,(\rho A_x)\mathrm{tr}\,(\rho A_y)\right] \tag{3.4}$$

We also have by (2.4) that

$$\begin{aligned} \mathrm{tr}\,\left(\rho\left[\widetilde{A},\widetilde{B}\right]\right) = 2i\,\mathrm{Im}\,\mathrm{tr}\,(\rho\widetilde{A}\widetilde{B}) &= 2i\,\mathrm{Im}\,\mathrm{tr}\left(\rho\sum_{x,y} xyA_xB_y)\right) \\ &= 2i\sum_{x,y} xy\,\mathrm{Im}\,\mathrm{tr}\,(\rho A_xB_y) \end{aligned} \tag{3.5}$$

Substituting $\widetilde{A}, \widetilde{B}$ for $A, B$ in Theorem 2.2 gives an uncertainty principle for real-valued observables.

Two observables $A, B$ are *compatible* (or *jointly measurable*) if there exists a *joint observable* $C_{(x,y)}$, $(x,y) \in \Omega_A \times \Omega_B$, such that $A_x = \sum_y C_{(x,y)}$, $B_y = \sum_x C_{(x,y)}$ for all $x \in \Omega_A$, $y \in \Omega_B$. If $\left[A_x, B_y\right] = 0$ for all $x, y$, then $A, B$ are compatible with $C_{(x,y)} = A_xB_y$ for all

$(x, y) \in \Omega_A \times \Omega_B$. However, if $A, B$ are compatible, they need not commute [4]. If $A, B$ are compatible real-valued observables, then

$$\widetilde{A} = \sum_x x A_x = \sum_{x,y} x C_{(x,y)}$$
$$\widetilde{B} = \sum_y y B_y = \sum_{x,y} y C_{(x,y)}$$

Using (3.2), (3.3), (3.4), (3.5) we can write $\mathrm{Cor}_\rho(A, B)$, $\Delta_\rho(A, B)$, $\Delta_\rho(A)$, $\Delta_\rho(B)$ and $\mathrm{tr}\left(\rho\left[\widetilde{A}, \widetilde{B}\right]\right)$ in terms of $C_{(x,y)}$. Hence, we can express the uncertainty principle in terms of $C_{(x,y)}$.

If $A = \{A_x : x \in \Omega_A\}$ is a real-valued observable, then $\widetilde{A}$ has spectral decomposition $\widetilde{A} = \sum_{i=1}^{n} \lambda_i P_i$ where $\lambda_i \in \mathbb{R}$ are the distinct eigenvalues of $\widetilde{A}$ and $P_i$ are projections with $\sum P_i = I$. We call $\widehat{A} = \{P_i : i = 1, 2, \ldots, n\}$ the *sharp version* of $A$. Then $\widehat{A}$ is a real-valued observable with outcome space $\Omega_{\widehat{A}} = \{\lambda_i : i = 1, 2, \ldots, n\}$ and $P_{\lambda_i} = P_i$. Since $(\widehat{A})^{\sim} = \widetilde{A}$, $A$ and $\widehat{A}$ have the same stochastic operator. It follows that $\langle A\rangle_\rho = \langle \widehat{A}\rangle_\rho$, $\Delta_\rho(A) = \Delta_\rho(\widehat{A})$ and if $B$ is another real-valued observable, then $\mathrm{Cor}_\rho(A, B) = \mathrm{Cor}_\rho(\widehat{A}, \widehat{B})$ and $\Delta_\rho(A, B) = \Delta_\rho(\widehat{A}, \widehat{B})$.

**Lemma 3.1** The following statements are equivalent. (i) $\widehat{A} = \widehat{B}$. (ii) $\widetilde{A} = \widetilde{B}$. (iii) $\langle A\rangle_\rho = \langle B\rangle_\rho$ for all $\rho \in \mathcal{S}(H)$.

***Proof*** (i)⇒(ii) If $\widehat{A} = \widehat{B}$ then

$$\widetilde{A} = (\widehat{A})^{\sim} = (\widehat{B})^{\sim} = \widetilde{B}$$

(ii)⇒(iii) If $\widetilde{A} = \widetilde{B}$ then

$$\langle A\rangle_\rho = \langle \widetilde{A}\rangle_\rho = \langle \widetilde{B}\rangle_\rho = \langle B\rangle_\rho$$

(iii)⇒(i) If $\langle A\rangle_\rho = \langle B\rangle_\rho$ for all $\rho \in \mathcal{S}(H)$, then $\langle \widetilde{A}\rangle_\rho = \langle \widetilde{B}\rangle_\rho$ for all $\rho \in \mathcal{S}(H)$. It follows that $\widehat{A} = \widehat{B}$. □

Let $\widetilde{A} = \sum x A_x = \sum \lambda_i P_i$ so $\widehat{A} = \{P_i : i = 1, 2, \ldots, n\}$ is a sharp version of $A$. Let $B = \{B_x : x \in \Omega_A\}$ be the real-valued observable given by $B_x = \sum_{i=1}^{n} P_i A_x P_i$. We conclude that $A$ and $B$ have the same sharp version because

$$\widetilde{B} = \sum_x x B_x = \sum_i P_i \sum_x x A_x P_i = \sum_i P_i \widetilde{A} P_i = \sum_i P_i \sum_j \lambda_j P_j P_i$$
$$= \sum_{i,j} \lambda_i P_i P_j P_i = \sum_i \lambda_i P_i = \widetilde{A}$$

so by Lemma 3.1, $\widehat{A} = \widehat{B}$. We say that $B$ is a *conjugate* of $A$. Letting $C_{ix} = P_i A_x P_i$, we have that

$$\{C_{ix} : i = 1, 2, \ldots, n, x \in \Omega_A\}$$

is an observable and $\sum_i C_{ix} = B_x$, $\sum_x C_{ix} = P_i$. It follows that $B$ and $\widehat{A}$ are compatible with joint observable $\{C_{ix}\}$. We say that an observable $A = \{A_x : x \in \Omega_A\}$ is *commutative* if $\left[A_x, A_y\right] = 0$ for all $x, y \in \Omega_A$. Notice that if $A$ is sharp, then $A$ is commutative. However, there are many unsharp observables that are commutative.

**Theorem 3.2** *If $A$ is commutative, then $B$ is conjugate to $A$ if and only if $B = A$.*

***Proof*** If $A$ is commutative, we show that $A$ is conjugate to $A$. Since

$$\widehat{A} = \sum x A_x = \sum \lambda_i P_i$$

we have that $\left[\widehat{A}, A_x\right] = 0$ for all $x \in \Omega_A$. By the spectral theorem, $[A_x, P_i] = 0$ for all $x, i$ so $A_x = \sum P_i A_x P_i$. Therefore, $A$ is conjugate to $A$. Conversely, suppose $A$ is commutative and $B$ is conjugate to $A$. Then $B_x = \sum\limits_i P_i A_x P_i$ for all $x \in \Omega_A$. As before, we have that $\left[\widehat{A}_x, A_x\right] = 0$ for all $x \in \Omega_A$ so $[A_x, P_i] = 0$ for all $x, i$. Hence,

$$B_x = \sum_i P_i A_x P_i = A_x \sum_i P_i = A_x$$

for all $x \in \Omega_B = \Omega_A$ so $B = A$. □

Thus, nontrivial conjugates only occur in the nonclassical case where $A$ is noncommutative.

## 4 More Examples

This section illustrates the theory in Sections 2 and 3 with two examples.

**Example 3** *A two outcome observable is called a dichotomic observable. Of course, a dichotomic observable is commutative but it need not be sharp. Let $A = \{A_1, I - A_1\}$ be a dichotomic observable with $\Omega_A = \{1, -1\}$. Then*

$$\begin{aligned}
\widetilde{A} &= A_1 - (I - A_1) = 2A_1 - I \\
\langle A\rangle_\rho &= \mathrm{tr}\,(\rho\widetilde{A}\,) = \mathrm{tr}\,[\rho(2A_1 - I)] = 2\,\mathrm{tr}\,(\rho A_1) - 1 \\
D_\rho(A) &= \widetilde{A} - \langle A\rangle_\rho I = 2A_1 - I - 2\,\mathrm{tr}\,(\rho A_1)I + I = 2\,[A_1 - \mathrm{tr}\,(\rho A_1)I]
\end{aligned}$$

*If $B = \{B_1, I - B_1\}$ is another dichotomic observable with $\Omega_B = \{1, -1\}$, then*

$$\begin{aligned}
\mathrm{Cor}_\rho(A, B) &= \mathrm{tr}\,(\rho\widetilde{A}\widetilde{B}\,) - \langle A\rangle_\rho\langle B\rangle_\rho \\
&= \mathrm{tr}\,[\rho(2A_1 - I)(2B_1 - I)] - [2\,\mathrm{tr}\,(\rho A_1 - 1)]\,[2\,\mathrm{tr}\,(\rho B_1 - 1)] \\
&= \mathrm{tr}\,[\rho(4A_1B_1 - 2A_1 - 2B_1 + I)] - 4\,\mathrm{tr}\,(\rho A_1)\mathrm{tr}\,(\rho B_1) \\
&\quad +2\,\mathrm{tr}\,(\rho A_1) + 2\mathrm{tr}\,(\rho B_1) - 1 \\
&= 4\,[\mathrm{tr}\,(\rho A_1B_1) - \mathrm{tr}\,(\rho A_1)\mathrm{tr}\,(\rho B_1)]
\end{aligned} \tag{4.1}$$

*Hence,*

$$\Delta_\rho(A, B) = 4\,[\mathrm{Re}\,\mathrm{tr}\,(\rho A_1B_1) - \mathrm{tr}\,(\rho A_1)\mathrm{tr}\,(\rho B_1)]$$

*and*

$$\Delta_\rho(A) = \Delta_\rho(A, A) = 4\left[\mathrm{tr}\,(\rho A_1^2) - (\mathrm{tr}\,(\rho A_1))^2\right]$$

*We also have*

$$\begin{aligned}
\left[\widetilde{A}, \widetilde{B}\,\right] &= [2A_1 - I, 2B_1 - I] = (2A_1 - I)(2B_1 - I) - (2B_1 - I)(2A_1 - I) \\
&= 4\,[A_1, B_1]
\end{aligned}$$

*We conclude that* $\left[\widetilde{A}, \widetilde{B}\right] = 0$ *if and only if* $[A_1, B_1] = 0$ *and this does not hold in general so* $\widetilde{A}, \widetilde{B}$ *need not commute. The uncertainty principle becomes*

$$\begin{aligned}
&[\operatorname{Im}\operatorname{tr}(\rho A_1 B_1)]^2 + [\operatorname{Re}\operatorname{tr}(\rho A_1 B_1) - \operatorname{tr}(\rho A_1)\operatorname{tr}(\rho A_2)]^2 \\
&= |\operatorname{tr}(\rho A_1 B_1) - \operatorname{tr}(\rho A_1)\operatorname{tr}(\rho B_1)|^2 \\
&\leq \left[\operatorname{tr}(\rho A_1^2) - (\operatorname{tr}(\rho A_1))^2\right]\left[\operatorname{tr}(\rho B_1^2) - (\operatorname{tr}(\rho B_1))^2\right]
\end{aligned} \tag{4.2}$$

□

**Example 4** *We now consider a special case of Example 3. For* $H \in \mathbb{C}^2$ *we define the Pauli matrices*

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_y = \begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix}, \quad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

*Let* $\mu \in [0, 1]$ *and define the dichotomic observable* $A = \{A_1, I - A_1\}$*, where*

$$A_1 = \frac{1}{2}(I + \mu\sigma_x) = \frac{1}{2}\begin{bmatrix} 1 & \mu \\ \mu & 1 \end{bmatrix}$$

*and* $\Omega_A = \{1, -1\}$*. Similarly, let* $B = \{B_1, I - B_1\}$*, where*

$$B_1 = \frac{1}{2}(I + \mu\sigma_y) = \frac{1}{2}\begin{bmatrix} 1 & i\mu \\ -i\mu & 1 \end{bmatrix}$$

*and* $\Omega_B = \{1, -1\}$*. We call A and B noisy spin observables along the x and y directions, respectively, with noise parameter* $1 - \mu$ [7].

*Any state* $\rho \in \mathcal{S}(H)$ *has the form* $\rho = \frac{I}{2}(I + \overrightarrow{r} \bullet \overrightarrow{\sigma})$ *where* $\overrightarrow{r} \in \mathbb{R}^3$ *with* $\|\overrightarrow{r}\| \leq 1$ [1, 2]*. This is called the Block sphere representation of* $\rho$ [4, 7]*. The eigenvalues of* $\rho$ *are* $\lambda_\pm = \frac{1}{2}\left(1 \pm \|\overrightarrow{r}\|\right)$*. Then* $\lambda_+ = 1$, $\lambda_- = 0$ *if and only if* $\|\overrightarrow{r}\| = 1$ *and these are precisely the pure states. Letting* $\sigma_1 = \sigma_x$, $\sigma_2 = \sigma_y$, $\sigma_3 = \sigma_z$ *we obtain*

$$\rho = \frac{1}{2}\begin{bmatrix} 1 + r_3 & r_1 - ir_2 \\ r_1 + ir_2 & 1 - r_3 \end{bmatrix}$$

*and*

$$\begin{aligned}
\rho A_1 &= \frac{1}{4}\begin{bmatrix} 1 + r_3 & r_1 - ir_2 \\ r_1 + ir_2 & 1 - r_3 \end{bmatrix}\begin{bmatrix} 1 & \mu \\ \mu & 1 \end{bmatrix} \\
&= \begin{bmatrix} 1 + r_3 + (r_1 - ir_2)\mu & (1 + r_3)\mu + r_1 - ir_2 \\ (1 - r_3)\mu + r_1 + ir_2 & 1 - r_3 + (r_1 + ir_2)\mu \end{bmatrix}
\end{aligned}$$

*Hence,* $\operatorname{tr}(\rho A_1) = \frac{1}{2}(1 + r_1\mu)$ *and as in Example 3,* $\langle A\rangle_\rho = r_1\mu$*. Similarly,* $\operatorname{tr}(\rho B_1) = \frac{1}{2}(1 + r_2\mu)$ *and* $\langle B\rangle_\rho = r_2\mu$*. We also obtain*

$$\operatorname{tr}(\rho A_1 B_1) = \frac{1}{4}\left[1 + (r_1 + r_2)\mu + ir_2\mu^2\right]$$

*and it follows from* (4.1) *that*

$$\begin{aligned}
\operatorname{Cor}_\rho(A, B) &= 4\left[\operatorname{tr}(\rho A_1 B_1) - \operatorname{tr}(\rho A_1)\operatorname{tr}(\rho B_1)\right] \\
&= 1 + (r_1 + r_2)\mu + ir_3\mu^2 - (1 + r_1\mu)(1 + r_2\mu) = -r_1 r_2\mu^2 + ir_3\mu^2
\end{aligned}$$

*Therefore,* $\Delta_\rho(A,B)=-r_1r_2\mu^2$. *A straightforward calculation shows that*

$$\operatorname{tr}(\rho A_1^2)=\frac{1}{4}(1+\mu^2)+\frac{1}{2}\mu r_1$$
$$\operatorname{tr}(\rho B_1^2)=\frac{1}{4}(1+\mu^2)+\frac{1}{2}\mu r_2$$

*It follows that*

$$\Delta_\rho(A)=4\left[\operatorname{tr}(\rho A_1^2)-(\operatorname{tr}(\rho A_1))^2\right]=\mu^2(1-r_1^2)$$

*and similarly,* $\Delta_\rho(B)=\mu^2(1-r_2^2)$.

*The commutator term in* (4.2) *becomes*

$$[\operatorname{Im}\operatorname{tr}(\rho A_1B_1)]^2=\frac{1}{16}r_3^2\mu^4$$

*The covariance term in* (4.2) *is*

$$[\operatorname{Re}(\rho A_1B_1)-\operatorname{tr}(\rho A_1)\operatorname{tr}(\rho B_1)]^2=\frac{1}{16}r_1^2r_2^2\mu^4$$

*and the correlation term in* (4.2) *is*

$$|\operatorname{tr}(\rho A_1B_1)-\operatorname{tr}(\rho A_1)\operatorname{tr}(\rho B_1)|^2=\frac{1}{16}(r_3^2+r_1^2r_2^2)\mu^4$$

*Finally, the variance term in* (4.2) *is given by*

$$\Delta_\rho(A_1)\Delta_\rho(B_1)=\frac{1}{16}(1-r_1^2)(1-r_2^2)\mu^4$$

*The inequality in* (4.2) *reduces to*

$$\frac{1}{16}(r_3^2+r_1^2+r_2^2)\mu^4\le\frac{1}{16}(1-r_1^2)(1-r_2^2)\mu^4 \tag{4.3}$$

*If* $\mu\ne 0$, (4.3) *is equivalent to the inequality*

$$\|\overrightarrow{r}\|^2=r_1^2+r_2^2+r_3^2\le 1$$

*If the commutator term vanishes and* $\mu\ne 0$, *the uncertainty inequality becomes*

$$r_1^2r_2^2\le(1-r_1^2)(1-r_2^2) \tag{4.4}$$

*which is equivalent to* $r_1^2+r_2^2\le 1$. *If A and B are* $\rho$*-uncorrelated and* $\mu\ne 0$, *the uncertainty inequality becomes* $r_3^2\le(1-r_1^2)(1-r_2^2)$ *which is equivalent to* $\|\overrightarrow{r}\|^2\le 1+r_1^2r_2^2$. *This inequality and* (4.4) *are weaker than* (4.3). □

## 5 Real-Valued Coarse Graining

Let $A=\{A_x: x\in\Omega_A\}$ be an arbitrary observable. We assume that $A$ is not necessarily real-valued so the outcome space $\Omega_A$ is an arbitrary finite set. For $f:\Omega_A\to\mathbb{R}$ with range $\mathcal{R}(f)$ we define the real-valued observable $f(A)$ by $\Omega_{f(A)}=\mathcal{R}(f)$ and for all $z\in\Omega_{f(A)}$

$$f(A)_z=A_{f^{-1}(z)}=\sum\{A_x: f(x)=z\}$$

We call $f(A)$ a *real-valued coarse graining* of $A$ [2–4]. Then $f(A)$ has stochastic operator

$$f(A)^{\sim}=\sum_z zf(A)_z=\sum_z zA_{f^{-1}(z)}=\sum_z\sum_{x\in f^{-1}(z)}zA_x=\sum_x f(x)A_x$$

It follows that $\langle f(A)\rangle_\rho=\sum\limits_x f(x)\mathrm{tr}\,(\rho A_x)$ for all $\rho\in\mathcal{S}(H)$. If $B$ is another observable and $g:\Omega_B\to\mathbb{R}$ we have

$$\begin{aligned}\mathrm{Cor}_\rho\,[f(A),g(B)]&=\sum_{x,y}f(x)g(y)\mathrm{tr}\,(\rho A_xB_y)-\langle f(A)\rangle_\rho\langle g(B)\rangle_\rho\\ \Delta_\rho\,[f(A),g(B)]&=\sum_{x,y}f(x)g(y)\mathrm{Re}\,\mathrm{tr}\,(\rho A_xB_y)-\langle f(A)\rangle_\rho\langle g(B)\rangle_\rho\\ \Delta_\rho\,[f(A)]&=\sum_{x,y}f(x)f(y)\mathrm{tr}\,(\rho A_xA_y)-\langle f(A)\rangle_\rho^2\end{aligned}$$

Moreover, we have the uncertainty inequality

$$|\mathrm{Cor}_\rho\,[f(A),g(B)]|^2\le\Delta_\rho\,[f(A)]\,\Delta_\rho\,[g(B)]$$

We denote the set of trace-class operators on $H$ by $\mathcal{T}(H)$. An *operation* on $H$ is a completely positive, trace reducing, linear map $\mathcal{O}:\mathcal{T}(H)\to\mathcal{T}(H)$ [1–4]. If $\mathcal{O}$ preserves the trace, then $\mathcal{O}$ is called a *channel*. A (finite) *instrument* is a finite set of operators $\mathcal{I}=\{\mathcal{I}_x:x\in\Omega_\mathcal{I}\}$ such that $\overline{\mathcal{I}}=\sum\{\mathcal{I}_x:x\in\Omega_\mathcal{I}\}$ is a channel [1–4]. We say that $\mathcal{I}$ *measures* an observable $A$ if $\Omega_\mathcal{I}=\Omega_A$ and $\mathrm{tr}\,[\mathcal{I}_x(\rho)]=\mathrm{tr}\,(\rho A_x)$ for all $x\in\Omega_\mathcal{I}$. It can be shown that $\mathcal{I}$ measures a unique observable which we denote by $J(\mathcal{I})$ [2, 3]. Conversely, any observable is measured by many instruments [1–4]. Corresponding to an operation $\mathcal{O}$ we have its *dual-operation* $\mathcal{O}^*:\mathcal{L}(H)\to\mathcal{L}(H)$ defined by $\mathrm{tr}\,\left[\rho\mathcal{O}^*(C)\right]=\mathrm{tr}\,[\mathcal{O}(\rho)C]$ for all $\rho\in\mathcal{S}(H)$ [2, 3]. It can be shown that $J(\mathcal{I})_x=\mathcal{I}_x^*(I)$ for all $x\in\Omega_\mathcal{I}$ where $I$ is the identity operator [2, 3].

As with observables, if $\mathcal{I}$ is an instrument, and $f:\Omega_\mathcal{I}\to\mathbb{R}$ we define the real-valued instrument $f(\mathcal{I})$ such that $\Omega_{f(\mathcal{I})}=\mathcal{R}(f)$ and

$$f(\mathcal{I})_z=\sum\{\mathcal{I}_x:f(x)=z\}$$

If $J(\mathcal{I})=A$, then $J\,[f(\mathcal{I})]=f(A)$ because

$$\begin{aligned}\mathrm{tr}\,\left[f(\mathcal{I})_z(\rho)\right]&=\mathrm{tr}\,\left[\sum\{\mathcal{I}_x(\rho):f(x)=z\}\right]=\sum\{\mathrm{tr}\,[\mathcal{I}_x(\rho)]:f(x)=z\}\\&=\sum\{\mathrm{tr}\,(\rho A_x):f(x)=z\}=\mathrm{tr}\,\left[\rho\sum\{A_x:f(x)=z\}\right]\\&=\mathrm{tr}\,\left[\rho f(A)_z\right]\end{aligned}$$

for all $z\in\Omega_{f(A)}=\Omega_{f(\mathcal{I})}$. If $\mathcal{I}$ is real-valued, we define $\widetilde{\mathcal{I}}$ on $\mathcal{L}(H)$ by $\widetilde{\mathcal{I}}(C)=\sum x\mathcal{I}_x(C)$ and $\langle\mathcal{I}\rangle_\rho=\mathrm{tr}\,\left[\widetilde{\mathcal{I}}(\rho)\right]$. If $J(\mathcal{I})=A$, then

$$\langle\mathcal{I}\rangle_\rho=\mathrm{tr}\,\left[\sum x\mathcal{I}_x(\rho)\right]=\sum x\mathrm{tr}\,[\mathcal{I}_x(\rho)]=\sum x\mathrm{tr}\,(\rho A_x)=\langle A\rangle_\rho$$

for all $\rho\in\mathcal{S}(H)$. We also define $\Delta_\rho(\mathcal{I})=\Delta_\rho(A)$. It follows that $\langle f(\mathcal{I})\rangle_\rho=\langle f(A)\rangle_\rho$, $\Delta_\rho\,[f(\mathcal{I})]=\Delta_\rho\,[f(A)]$ and $f(\mathcal{I})^{\sim}=\sum f(x)\mathcal{I}_x$.

Let $A=\{A_x : x\in\Omega_A\}$, $B=\left\{B_y : y\in\Omega_B\right\}$ be arbitrary observables and suppose $\mathcal{I}$ is an instrument with $J(\mathcal{I})=A$. Define the $\mathcal{I}$-*product* observable $A\circ B$ with $\Omega_{A\circ B}=\Omega_A\times\Omega_B$ given by $(A\circ B)_{(x,y)}=\mathcal{I}_x(B_y)$ [2, 3]. Then $A\circ B$ is indeed an observable because

$$\sum_{x,y}(A\circ B)_{(x,y)}=\sum_{x,y}\mathcal{I}_x^*(B_y)=\sum_x\mathcal{I}_x^*\left(\sum_y B_y\right)=\sum_x\mathcal{I}_x^*(I)=\sum_x A_x=I$$

Although $A\circ B$ depends on $\mathcal{I}$, we shall not indicate this for simplicity. We interpret $A\circ B$ as the observable obtained by first measuring $A$ using $\mathcal{I}$ and then measuring $B$. If $f:\Omega_A\times\Omega_B\to\mathbb{R}$ we obtain the real-valued observable $f(A,B)=f(A\circ B)$. We then have

$$\begin{aligned}
f(A,B)_z&=(A\circ B)_{f^{-1}(z)}=\sum\left\{(A\circ B)_{(x,y)}:f(x,y)=z\right\}\\
&=\sum\left\{\mathcal{I}_x^*(B_y):f(x,y)=z\right\}\\
f(A,B)^{\sim}&=\sum_{x,y}f(x,y)(A\circ B)_{(x,y)}=\sum_{x,y}f(x,y)\mathcal{I}_x^*(B_y)\\
\langle f(A,B)\rangle_\rho&=\sum_{x,y}f(x,y)\mathrm{tr}\left[\rho(A\circ B)_{(x,y)}\right]=\sum_{x,y}f(x,y)\mathrm{tr}\left[\rho\mathcal{I}_x^*(B_y)\right]\\
\Delta_\rho\left[f(A,B)\right]&=\sum_{x,y,x',y'}f(x,y)f(x',y')\mathrm{tr}\left[\rho(A\circ B)_{(x,y)}(A\circ B)_{(x',y')}\right]-\langle f(A,B)\rangle_\rho^2\\
&=\mathrm{tr}\left\{\rho\left[\sum_{x,y}f(x,y)\mathcal{I}_x^*(B_y)\right]^2\right\}-\langle f(A,B)\rangle_\rho^2
\end{aligned}$$

If $f$ is a product function $f(x,y)=g(x)h(y)$ we obtain

$$f(A,B)_z=\sum_z\left\{\mathcal{I}_x^*(B_y):g(x)h(y)=z\right\}$$

We then have the simplification

$$\begin{aligned}
f(A,B)^{\sim}&=\sum_{x,y}g(x)h(y)\mathcal{I}_x^*(B_y)=\sum_x g_x\mathcal{I}_x^*\left(\sum_y h(y)B_y\right)\\
&=\sum_x g(x)\mathcal{I}_x^*\left[h(B)^{\sim}\right]
\end{aligned}$$

Hence,

$$\begin{aligned}
\langle f(A,B)\rangle_\rho&=\mathrm{tr}\left[\rho f(A,B)^{\sim}\right]=\mathrm{tr}\left\{\rho\sum_x g(x)\mathcal{I}_x^*\left[h(B)^{\sim}\right]\right\}\\
&=\sum_x g(x)\mathrm{tr}\left\{\rho\mathcal{I}_x^*\left[h(B)^{\sim}\right]\right\}=\sum_x g(x)\mathrm{tr}\left\{\mathcal{I}_x(\rho)\left[h(B)^{\sim}\right]\right\}\\
&=\mathrm{tr}\left\{\sum_x g(x)\mathcal{I}_x(\rho)\left[h(B)^{\sim}\right]\right\}=\mathrm{tr}\left\{g(\mathcal{I})^{\sim}(\rho)\left[h(B)^{\sim}\right]\right\}
\end{aligned}$$

In a similar way we obtain

$$\Delta_\rho\left[f(A,B)\right] = \text{tr}\left\{\left(g(\mathcal{I})^\sim(\rho)\left[h(B)^\sim\right]\right)^2\right\} - \langle f(A,B)\rangle_\rho^2$$

If $A$ and $B$ are arbitrary observables, we define the observable $B$ *conditioned* by $A$ to be

$$(B \mid A)_y = \mathcal{I}_{\Omega_A}^*(B_y) = \sum_{x\in\Omega_A}\mathcal{I}_x^*(B_y)$$

where $\Omega_{B|A} = \Omega_B$ [2, 3]. We interpret $(B \mid A)$ as the observable obtained by first measuring $A$ without taking the outcome into account and then measuring $B$. If $B$ is real-valued we have

$$(B \mid A)^\sim = \sum_y y(B \mid A)_y = \sum_{x,y} y\mathcal{I}_x^*(B_y) = \mathcal{I}_{\Omega(A)}^*(\widetilde{B})$$

$$\langle(B \mid A)\rangle_\rho = \sum_y y\text{tr}\left[\rho\mathcal{I}_{\Omega(A)}^*(B_y)\right] = \sum y\text{tr}\left[\overline{\mathcal{I}}(\rho)B_y\right] = \text{tr}\left[\overline{\mathcal{I}}(\rho)\widetilde{B}\right] = \langle B\rangle_{\overline{\mathcal{I}}(\rho)}$$

$$\Delta_\rho\left[(B \mid A)\right] = \Delta_\rho\left[(B \mid A)^\sim\right] = \Delta_\rho\left[\mathcal{I}_{\Omega(A)}^*(\widetilde{B})\right] = \text{tr}\left\{\left[\mathcal{I}_{\Omega(A)}^*(\widetilde{B})\right]^2\right\} - \left[\langle B\rangle_{\overline{\mathcal{I}}(\rho)}\right]^2$$

We now illustrate the theory of this section with some examples.

**Example 5** *The simplest example of an instrument is a trivial instrument $\mathcal{I}_x(\rho) = \omega(x)\rho$ where $\omega$ is a probability measure on the finite set $\Omega_{\mathcal{I}}$. It is clear that $\mathcal{I}$ measures the trivial observable $A_x = \omega(x)I$. Let $B$ be an arbitrary observable and let $f : \Omega_A \times \Omega_B \to \mathbb{R}$. We then have*

$$(A \circ B)_{(x,y)} = \mathcal{I}_x^*(B_y) = \omega(x)B_y$$

$$f(A,B)_z = f(A\circ B)_z = \sum\left\{\omega(x)B_y : f(x,y) = z\right\}$$

*We conclude that*

$$f(A,B)^\sim = \sum_{x,y} f(x,y)\omega(x)B_y$$

$$\langle f(A,B)\rangle_\rho = \sum_{x,y} f(x,y)\omega(x)\text{tr}\,(\rho B_y)$$

$$\Delta_\rho\left[f(A,B))\right] = \text{tr}\left\{\rho\left[\sum_{x,y} f(x,y)\omega(x)B_y\right]^2\right\} - \langle f(A,B)\rangle_\rho^2$$

*Moreover, since*

$$(B \mid A)_y = \sum_x \mathcal{I}_x^*(B_y) = \sum_x \omega(x)(B_y) = B_y$$

*we have that $(B \mid A) = B$.* □

**Example 6** *Let $A = \{A_x : x \in \Omega_A\}$ and $B = \left\{B_y : y \in \Omega_B\right\}$ be arbitrary observables and let $\mathcal{H}_x(\rho) = \text{tr}\,(\rho A_x)\alpha_x$, $\alpha_x \in \mathcal{S}(H)$ be a Holevo instrument [2, 3]. Then $\mathcal{H}$ measure $A$ because*

$$\text{tr}\left[\mathcal{H}_x(\rho)\right] = \text{tr}\left[\text{tr}\,(\rho A_x)\alpha_x\right] = \text{tr}\,(\rho A_x)$$

*Since $\mathcal{H}_x^*(a) = \text{tr}\,(\alpha_x a)A_x$ for all $x \in \Omega_A$ [2, 3], we have*

$$(A \circ B)_{(x,y)} = \mathcal{H}_x^*(B_y) = \text{tr}\,(\alpha_x B_y)A_x$$

*If $f : \Omega_A \times \Omega_B \to \mathbb{R}$, we obtain the real-valued observable*

$$f(A,B)_z = \sum \left\{ \text{tr}\,(\alpha_x B_y) A_x : f(x,y) = z \right\}$$

*We conclude that*

$$\begin{aligned} f(A,B)_z &= \sum_{x,y} f(x,y) \mathcal{H}_x^*(B_y) = \sum_{x,y} f(x,y) \text{tr}\,(\alpha_x B_y) A_x \\ \langle f(A,B) \rangle_\rho &= \sum_{x,y} f(x,y) \text{tr}\,(\alpha_x B_y) \text{tr}\,(\rho A_x) \\ \Delta_\rho \left[ f(A,B) \right] &= \sum_{x,y,x',y'} f(x,y) f(x',y') \text{tr}\left[ \rho \text{tr}\,(\alpha_x B_y) A_x \text{tr}\,(\alpha_{x'} B_{y'}) A_{x'} \right] \\ &\quad - \langle f(A,B) \rangle_\rho^2 \\ &= \text{tr}\left\{ \rho \left[ \sum_{x,y} f(x,y) \text{tr}\,(\alpha_x B_y) A_x \right]^2 \right\} - \langle f(A,B) \rangle_\rho^2 \end{aligned}$$

*Moreover, we have*

$$(B \mid A)_y = \sum_x \mathcal{H}_x^*(B_y) = \sum_x \text{tr}\,(\alpha_x B_y) A_x \qquad \square$$

**Example 7** *Let $A, B$ be arbitrary observables and let $\mathcal{L}$ be the Lüders instrument given by $\mathcal{L}_x(\rho) = A_x^{1/2} \rho A_x^{1/2}$* [2, 3, 6]. *Then*

$$\text{tr}\,[\mathcal{L}_x(\rho)] = \text{tr}\,(A_x^{1/2} \rho A_x^{1/2}) = \text{tr}\,(\rho A_x)$$

*so $\mathcal{L}$ measures $A$. Since $\mathcal{L}_x^*(a) = A_x^{1/2} a A_x^{1/2}$* [2, 3] *we have*

$$(A \circ B)_{(x,y)} = A_x^{1/2} B_y A_x^{1/2}$$

*If $f : \Omega_A \times \Omega_B \to \mathbb{R}$, we obtain the real-valued observable*

$$f(A,B)_z = \sum \left\{ A_x^{1/2} B_y A_x^{1/2} : f(x,y) = z \right\}$$

*We conclude that*

$$\begin{aligned} f(A,B)^\sim &= \sum_{x,y} f(x,y) A_x^{1/2} B_y A_x^{1/2} \\ \langle f(A,B) \rangle_\rho &= \sum_{x,y} f(x,y) \text{tr}\,(\rho A_x^{1/2} B_y A_x^{1/2}) = \sum_{x,y} f(x,y) \text{tr}\,(A_x^{1/2} \rho A_x^{1/2} B_y) \\ \Delta_\rho \left[ f(A,B) \right] &= \text{tr}\left\{ \rho \left[ \sum_{x,y} f(x,y) A_x^{1/2} B_y A_x^{1/2} \right]^2 \right\} - \langle f(A,B) \rangle_\rho^2 \end{aligned}$$

*Moreover, we have*

$$(B \mid A)_y = \sum_x \mathcal{L}_x^*(B_y) = \sum_x A_x^{1/2} B_y A_x^{1/2} \qquad \square$$

## References

1. Busch, P., Grabowski, M., Lahti, P.: Operational Quantum Physics. Springer-Verlag, Berlin (1995)
2. Gudder, S.: Conditional observables in quantum mechanics. arXiv:quant-ph **2005**, 04775 (2020)
3. Gudder, S.: Quantum instruments and conditioned observables (2020). arXiv preprint arXiv:2005.08117
4. Heinosaari, T., Ziman, M.: The Mathematical Language of Quantum Theory. Cambridge University Press, Cambridge (2012)
5. Kraus, K.: States. Effects and Operations. Springer-Verlag, Berlin (1983)
6. Lüders, G.: Über due Zustandsänderung durch den Messprozess. Ann. Physik. **6**, 322–328 (1951)
7. Nielson, M., Chuang, I.: Quantum Computation and Quantum Information. Cambridge University Press, Cambridge (2000)