Heisenberg's uncertainty relation: Violation and reformulation



2014 J. Phys.: Conf. Ser. 504 012024

(http://iopscience.iop.org/1742-6596/504/1/012024)

Journal of Physics: Conference Series **504** (2014) 012024 doi:10.1088/1742-6596/504/1/012024

# Heisenberg's uncertainty relation: Violation and reformulation

**Masanao Ozawa**

Graduate School of Information Science, Nagoya University, Chikusa-ku, Nagoya, 464-8601, Japan

E-mail: ozawa@is.nagoya-u.ac.jp

**Abstract.** The uncertainty relation formulated by Heisenberg in 1927 describes a trade-off between the error of a measurement of one observable and the disturbance caused on another complementary observable so that their product should be no less than a limit set by Planck's constant. In 1980, Braginsky, Vorontsov, and Thorne claimed that this relation leads to a sensitivity limit for gravitational wave detectors. However, in 1988 a model of position measurement was constructed that breaks both this limit and Heisenberg's relation. Here, we discuss the problems as to how we reformulate Heisenberg's relation to be universally valid and how we experimentally quantify the error and the disturbance to refute the old relation and to confirm the new relation.

## 1. Heisenberg's error-disturbance relation (EDR)

The discovery of quantum mechanics introduced non-commutativity in algebraic calculus of observables; the *canonical commutation relation* (CCR)

$$[Q, P] = i\hbar \tag{1}$$

is required to hold between a coordinate $Q$ of a particle and its momentum $P$, where the commutator $[Q, P]$ is defined by $[Q, P] = QP - PQ$. In 1927, Heisenberg proposed an operational meaning of the non-commutativity: "the more precisely the position is determined, the less precisely the momentum is known, and conversely" [1, p. 64].

By the famous $\gamma$ ray microscope thought experiment he derived the relation

$$\varepsilon(Q)\eta(P) \geq \frac{\hbar}{2}, \tag{2}$$

where $\varepsilon(Q)$ is the "mean error" of a position measurement and $\eta(P)$ is the thereby caused "discontinuous change" in the momentum $P$:

> Let $\varepsilon(Q)$ be the precision with which the value $Q$ is known ($\varepsilon(Q)$ is, say, the mean error of $Q$), therefore here the wavelength of the light. Let $\eta(P)$ be the precision with which the value $P$ is determinable; that is, here, the discontinuous change of $P$ in the Compton effect [1, p. 64].

Here, "mean error" is naturally understood to be "root-mean-square (rms) error" as introduced by Gauss [2], and "discontinuous change" is often called "mean disturbance." Heisenberg claimed that Eq. (2) is a "straightforward mathematical consequence" of Eq. (1) [1, p. 65] and gave its mathematical justification [1, p. 69].

## 2. Heisenberg's derivation

In his mathematical justification of Eq. (2), Heisenberg firstly derived the relation

$$\sigma(Q)\sigma(P) = \frac{\hbar}{2} \tag{3}$$

for the standard deviations $\sigma(Q)$ and $\sigma(P)$ of the position $Q$ and the momentum $P$ in the state described by a Gaussian wave function [1, p. 69], which Kennard [3] subsequently generalized as the relation

$$\sigma(Q)\sigma(P) \geq \frac{\hbar}{2} \tag{4}$$

for arbitrary wave functions. Here, the standard deviation is defined for any observable $A$ by $\sigma(A)^2 = \langle A^2\rangle - \langle A\rangle^2$, where $\langle\cdots\rangle$ stands for the mean value in a given state. Note that in Ref. [1, p. 69] Heisenberg actually derived the relation

$$\tilde{\sigma}(Q)\tilde{\sigma}(P) = \hbar \tag{5}$$

for $\tilde{\sigma}(Q) = \sqrt{2}\sigma(Q)$ and $\tilde{\sigma}(P) = \sqrt{2}\sigma(P)$ in Gaussian wave functions, and Kennard [3] actually derived the relation

$$\tilde{\sigma}(Q)\tilde{\sigma}(P) \geq \hbar \tag{6}$$

that generalizes Heisenberg's relation to arbitrary wave functions.

Heisenberg secondly applied Eq. (4) to the state just after the measurement assuming:

(H1) *Any measurement with rms error $\varepsilon(A)$ of an observable $A$ leaves the state satisfying $\sigma(A) \leq \varepsilon(A)$.*

(H2) *If an observable $A$ can be measured with the rms error $\varepsilon(A)$ and the rms disturbance $\eta(B)$ of another observable $B$, then $A$ and $B$ can be jointly measured with the rms errors $\varepsilon(A)$ and $\varepsilon(B) = \eta(B)$, respectively.*

Then, it can be easily seen that Eq. (2) can be derived from Eq. (4) under assumptions (H1) and (H2). In fact, if $Q$ can be measured with $\varepsilon(Q) = \alpha$ and $\eta(P) = \beta$, then by (H2) $Q$ and $P$ can be measured jointly with $\varepsilon(Q) = \alpha$ and $\varepsilon(P) = \beta$, so that by (H1) the state after the joint measurement satisfies $\sigma(Q) \leq \alpha$ and $\sigma(P) \leq \beta$, and hence Eq. (4) concludes Eq. (2).

## 3. Heisenberg's unsupported assumption

Assumption (H2) is considered to hold in general; see Ref. [4] for a detailed discussion. However, assumption (H1) is not, whereas Heisenberg's contemporaries, including von Neumann, supported assumption (H1):

> We are then to show that if $Q, P$ are two canonically conjugate quantities, and a system is in a state in which the value of $Q$ can be given with the accuracy $\varepsilon[= \sigma(Q)]$ (i.e., by a $Q$ measurement with an error range $\varepsilon[= \varepsilon(Q)]$), then $P$ can be known with no greater accuracy than $\eta[= \sigma(P)] = \hbar/(2\varepsilon)$. Or: a measurement of $Q$ with the accuracy $\varepsilon[= \varepsilon(Q)]$ must bring about an indeterminancy $\eta[= \eta(P)] = \hbar/(2\varepsilon)$ in the value of $P$ [5, pp. 238–239]. (Terms in [...] are supplemented by the present author.)

In those days the repeatability hypothesis was considered a natural requirement for all the precise measurements of an observable $A$ and (H1) is considered as a natural generalization to arbitrary approximate measurements of $A$. Here, the *repeatability hypothesis* is formulated as follows.

(RH) *If an observable $A$ is measured twice in succession in a system, then we get the same value each time* [5, pp. 335].

Under (RH), any precise measurement of $A$ with $\varepsilon(A) = 0$ changes the state to be an eigenstate of the measured observable $A$, which satisfies $\sigma(A) = 0$. However, in the light of modern theory of quantum measurement, (RH) has been abandoned as proposed by Davies and Lewis [6]:

> One of the crucial notions is that of repeatability which we show is implicitly assumed in most of the axiomatic treatments of quantum mechanics, but whose abandonment leads to a much more flexible approach to measurement theory [6, p.239].

In fact, in Ref. [7] we have mathematically characterized all the physically possible quantum measurements, shown that (RH) is no longer universally valid, and even more that no precise measurements of continuous observables satisfy (RH). Thus, (H1) does not hold even in the case where $\varepsilon(A) = 0$.

Therefore, in the light of the modern theory of quantum measurement, assumption (H1) cannot be accepted, so that Eq. (2) cannot be considered as an immediate consequence of Eq. (4), although their meanings have often been confused even in standard text books [5, 8, 9, 10]. As above, the original justification of Eq. (2) was limited, but its universal validity has not been refuted in theory until 1980's.

## 4. Von Neumann's model of position measurement

Until 1980's the only solvable model of position measurement had been given by von Neumann [5]. In what follows, we discuss the von Neumann model and show that this long-standing standard model satisfies the Heisenberg error-disturbance relation (EDR) (2). Thus, the model analysis of position measurement did not lead to refuting the Heisenberg EDR but rather enforced the belief that good position measurements satisfy the Heisenberg EDR (2).

Consider a one-dimensional mass, called an *object*, with position $Q$ and momentum $P$, described by a Hilbert space $\mathcal{H}$. The measurement of $Q$ is carried out by a coupling between the object $\mathbf{S}$ and a probe $\mathbf{P}$ from time $t = 0$ to $t = \Delta t$. The probe $\mathbf{P}$ is another one-dimensional mass with position $\overline{Q}$ and momentum $\overline{P}$, described by a Hilbert space $\mathcal{K}$. The outcome of the measurement is obtained by measuring the probe position $\overline{Q}$, called the *meter observable*, at time $t = \Delta t$. The total Hamiltonian for the object and the probe is taken to be

$$H_{\mathbf{S}+\mathbf{P}} = H_{\mathbf{S}} + H_{\mathbf{P}} + KH, \tag{7}$$

where $H_{\mathbf{S}}$ and $H_{\mathbf{P}}$ are the free Hamiltonians of $\mathbf{S}$ and $\mathbf{P}$, respectively, $H$ represents the measuring interaction, and $K$ is the coupling constant. We assume that the coupling is so strong ($K \gg 1$) that $H_{\mathbf{S}}$ and $H_{\mathbf{P}}$ can be neglected. We choose $\Delta t$ as $K\Delta t = 1$. In the von Neumann model the measuring interaction is given by

$$H = Q \otimes \overline{P}. \tag{8}$$

Then, the unitary operator of the time evolution of $\mathbf{S} + \mathbf{P}$ from $t = 0$ to $t = \tau \le \Delta t$ is given by

$$U(\tau) = \exp\left(\frac{-iK\tau}{\hbar} Q \otimes \overline{P}\right). \tag{9}$$

Suppose that the object $\mathbf{S}$ and the probe $\mathbf{P}$ are in the state $|\psi\rangle$ and $|\xi\rangle$, respectively, just before the measurement; we assume that the wave functions $\psi(x) = \langle x|\psi\rangle$ and $\xi(y) = \langle y|\xi\rangle$ are Schwartz rapidly decreasing functions [11], where $|x\rangle$ and $|y\rangle$ are the position bases of $\mathbf{S}$ and $\mathbf{P}$, respectively.

Then, the state of the composite system $\mathbf{S} + \mathbf{P}$ just after the measurement is $U(\Delta t)|\psi, \xi\rangle$. By solving the Schrödinger equation, we have

$$\langle x, y|U(\Delta t)|\psi, \xi\rangle = \langle x|\psi\rangle\langle y - x|\xi\rangle. \tag{10}$$

If the observer observes the meter observable $\overline{Q}(\Delta t)$ just after the measuring interaction, the probability distribution of the outcome is given by

$$\Pr\{a < \mathbf{x} \le b\} = \int_a^b dy \int_{-\infty}^{+\infty} |\langle x|\psi\rangle|^2 \, |\langle y-x|\xi\rangle|^2 \, dx. \tag{11}$$

This shows that if the probe initial wave function $\xi(y)$ approaches the Dirac delta function $\delta(x)$, the output probability distribution approaches the correct Born probability distribution for the observable $Q$ at time $t = 0$.

In the Heisenberg picture, we denote

$$\begin{array}{ll} Q(\tau) = U(\tau)^\dagger (Q \otimes I) U(\tau), & P(\tau) = U(\tau)^\dagger (P \otimes I) U(\tau), \\ \overline{Q}(\tau) = U(\tau)^\dagger (I \otimes \overline{Q}) U(\tau), & \overline{P}(\tau) = U(\tau)^\dagger (I \otimes \overline{P}) U(\tau). \end{array}$$

Solving the Heisenberg equations of motion, we have

$$\begin{aligned} Q(\Delta t) &= Q(0), && (12)\\ \overline{Q}(\Delta t) &= Q(0) + \overline{Q}(0), && (13)\\ P(\Delta t) &= P(0) - \overline{P}(0), && (14)\\ \overline{P}(\Delta t) &= \overline{P}(0). && (15) \end{aligned}$$

## 5. Root-mean-square error and disturbance

In order to define "root-mean-square error" of this measurement, we recall classical definitions. Suppose that the true value is given by $X = x$ and its measured value is given by $Y = y$. For each pair of values $(X, Y) = (x, y)$, the error is defined as $y - x$. To define the "mean error" with respect to the joint probability distribution $\mu^{X,Y}(dx, dy)$ of $X$ and $Y$, Gauss [2] introduced the *root-mean-square error* $\varepsilon_G(X, Y)$ of $Y$ for $X$ as

$$\varepsilon_G(X,Y) = \left( \iint_{\mathbf{R}^2} (y-x)^2 \mu^{X,Y}(dx, dy) \right)^{1/2}, \tag{16}$$

which Gauss [2] called the "mean error" or the "mean error to be feared," and has long been accepted as a standard definition for the "mean error."

In the von Neumann model, the value of the observable $Q(0)$ is measured by the value of the meter observable $\overline{Q}(\Delta t)$. Since $Q(0)$ and $\overline{Q}(\Delta t)$ commute, as seen from Eq. (13), we have the joint probability distribution $\mu^{Q(0),\overline{Q}(\Delta t)}(dx, dy)$ of the values of $Q(0)$ and $\overline{Q}(\Delta t)$ as

$$\mu^{Q(0),\overline{Q}(\Delta t)}(dx, dy) = \langle E^{Q(0)}(dx) E^{\overline{Q}(\Delta t)}(dy) \rangle, \tag{17}$$

where $E^A$ stands for the spectral measure of an observable $A$ [12], and $\langle \cdots \rangle$ stands for the mean value in the state $|\psi, \xi\rangle$. Then, from Eq. (16) the *root-mean-square error* $\varepsilon(Q)$ of $\overline{Q}(\Delta t)$ for $Q(0)$ in $|\psi\rangle$ is given by

$$\begin{aligned} \varepsilon(Q) &= \varepsilon_G(Q(0), \overline{Q}(\Delta t)) \\ &= \left( \iint_{\mathbf{R}^2} (y-x)^2 \mu^{Q(0),\overline{Q}(\Delta t)}(dx, dy) \right)^{1/2} \\ &= \langle (\overline{Q}(\Delta t) - Q(0))^2 \rangle^{1/2} \\ &= \langle \overline{Q}(0)^2 \rangle^{1/2}. && (18) \end{aligned}$$

Since $P(0)$ and $P(\Delta t)$ commute, as seen from Eq. (14), we have the joint probability distribution $\mu^{P(0),P(\Delta t)}(dx, dy)$ of the values of $P(0)$ and $P(\Delta t)$ as

$$\mu^{P(0),P(\Delta t)}(dx, dy) = \langle E^{P(0)}(dx)E^{P(\Delta t)}(dy)\rangle. \tag{19}$$

The *root-mean-square disturbance* $\eta(P)$ of $P$ from $t = 0$ to $t = \Delta t$ is defined as the root-mean-square error of $P(\Delta t)$ for $P(0)$ given by

$$\begin{aligned} \eta(P) &= \varepsilon_G(P(0), P(\Delta t)) \\ &= \left(\iint_{\mathbf{R}^2} (y - x)^2 \mu^{P(0),P(\Delta t)}(dx, dy)\right)^{1/2} \\ &= \langle (P(\Delta t) - P(0))^2\rangle^{1/2} \\ &= \langle \overline{P}(0)^2\rangle^{1/2}. \end{aligned} \tag{20}$$

Then, by the Kennard inequality (4) we have

$$\begin{aligned} \varepsilon(Q)\eta(P) &= \langle \overline{Q}(0)^2\rangle^{1/2}\langle \overline{P}(0)^2\rangle^{1/2} \\ &\geq \sigma(\overline{Q}(0))\sigma(\overline{P}(0)) \geq \frac{\hbar}{2}. \end{aligned} \tag{21}$$

Thus, the von Neumann model satisfies the Heisenberg EDR (2).

Since only the von Neumann model is available as a mathematically solvable mode of position measurement until 1980's, model analysis of position measurement did not lead to refuting the Heisenberg EDR but rather enforced the belief that good position measurements satisfy the Heisenberg EDR (2); see for example Refs. [13, 14].

This belief was also enforced by Arthurs and Kelly [15] suggesting that all the joint unbiased measurement of position and momentum satisfy the Heisenberg error tradeoff relation. A *joint position-momentum measurement* can be modeled by a triple $(M_Q, M_P, |\xi\rangle)$ consisting of commuting meter observables $M_Q$ and $M_P$ in the composite system $\mathbf{S} + \mathbf{P}$ described by $\mathcal{H} \otimes \mathcal{K}$ with the initial state $|\xi\rangle \in \mathcal{K}$ of the probe $\mathbf{P}$. Then, the joint measurement $(M_Q, M_P, |\xi\rangle)$ is called *unbiased* if the mean values of meter observables $M_Q$ and $M_P$ coincides with the mean values of $Q$ and $P$, respectively, i.e., $\langle M_Q\rangle = \langle Q\rangle$ and $\langle M_P\rangle = \langle P\rangle$, in any input states of $\mathbf{S}$. Then, Arthurs and Kelly [15] showed that the relation

$$\sigma(M_Q)\sigma(M_P) \geq \hbar \tag{22}$$

holds for any unbiased joint position-momentum measurement $(M_Q, M_P, |\xi\rangle)$ in any input state of $\mathbf{S}$. It is explained that the lower bound is twice as much as the lower bound for $\sigma(Q)\sigma(P)$ because of inevitable errors included in the value of $M_Q$ and $M_P$. In Ref. [16], their result was reformulated so that the Heisenberg error tradeoff relation

$$\varepsilon(Q)\varepsilon(P) \geq \frac{\hbar}{2} \tag{23}$$

holds for any unbiased joint position-momentum measurement $(M_Q, M_P, |\xi\rangle)$ in any input state of $\mathbf{S}$, where $\varepsilon(Q)$ and $\varepsilon(P)$ are the rms errors of joint measurement of position $Q$ and momentum $P$ defined through the joint probability distribution of $Q$ and $M_Q$ and that of $P$ and $M_P$, which always exist in unbiased case; see for another approach Ref. [17]. Thus, if we consider unbiased joint measurements as good joint measurements, we can say that every good joint measurement satisfies the Heisenberg error tradeoff relation (23).

Journal of Physics: Conference Series **504** (2014) 012024 doi:10.1088/1742-6596/504/1/012024

However, this does not imply that every good position measurement satisfies the Heisenberg EDR (2). Given a position measurement with the meter observable $M$ in the probe, if the momentum is measured just after the position measurement, we have a joint position-momentum measurement $(M_Q, M_P, |\xi\rangle)$ with $M_Q = M(\Delta t)$ and $M_P = P(\Delta t)$. In this case, the error tradeoff relation (23) is equivalent with the EDR (2) with $\varepsilon(P) = \eta(P)$, and we can consider unbiased position measurements as good position measurements. Nevertheless, we cannot conclude that every good position measurement satisfies the Heisenberg EDR (2), since an unbiased position measurement followed by a precise momentum measurement not necessarily satisfies $\langle P(\Delta t)\rangle = \langle P(0)\rangle$ or $\langle M_P\rangle = \langle P\rangle$.

## 6. Measurement violating the Heisenberg EDR

In 1980, Braginsky, Vorontsov, and Thorne [18] claimed that the Heisenberg EDR (2) leads to a sensitivity limit, called the *standard quantum limit* (SQL), for gravitational wave detectors of interferometer type, which make use of free-mass position monitoring. Subsequently, Yuen [19] questioned the validity of the SQL, and then Caves [14] defended the SQL by giving a new proof of the SQL without a direct appeal to Eq. (2). Eventually, the conflict was reconciled in Refs. [20, 21] by pointing out that Caves [14] used (unfounded) assumption (H1) in his derivation of the SQL, and a solvable model of an error-free position measurement was constructed that breaks the SQL (see also Ref. [22]); later this model was shown to break the Heisenberg EDR (2) [23].

In what follows, we introduce this model by modifying the measuring interaction of the von Neumann model. In this new model, the object, the probe, and the probe observables are the same systems and the same observable as the von Neumann model. The measuring interaction is taken to be [20]

$$H = \frac{\pi}{3\sqrt{3}}(2Q \otimes \overline{P} - 2P \otimes \overline{Q} + QP \otimes I - I \otimes \overline{QP}). \tag{24}$$

The coupling constant $K$ and the time duration $\Delta t$ are chosen as before so that $K \gg 1$ and $K\Delta t = 1$. Then, the unitary operator $U(\tau)$ for $0 \le \tau \le \Delta t$ is given by

$$U(\tau) = \exp\left[\frac{-i\pi K\tau}{3\sqrt{3}\hbar}(2Q \otimes \overline{P} - 2P \otimes \overline{Q} + QP \otimes I - I \otimes \overline{QP})\right]. \tag{25}$$

Solving the Heisenberg equations of motion for $t < t + \tau < t + \Delta t$, we obtain

$$\begin{aligned}
Q(\tau) &= \frac{2}{\sqrt{3}}Q(0)\sin\frac{(1+K\tau)\pi}{3} + \frac{-2}{\sqrt{3}}\overline{Q}(0)\sin\frac{K\tau\pi}{3}, && (26)\\
\overline{Q}(\tau) &= \frac{2}{\sqrt{3}}Q(0)\sin\frac{K\tau\pi}{3} + \frac{-2}{\sqrt{3}}\overline{Q}(0)\sin\frac{(1-K\tau)\pi}{3}, && (27)\\
P(\tau) &= \frac{-2}{\sqrt{3}}P(0)\sin\frac{(1-K\tau)\pi}{3} + \frac{-2}{\sqrt{3}}\overline{P}(0)\sin\frac{K\tau\pi}{3}, && (28)\\
\overline{P}(\tau) &= \frac{2}{\sqrt{3}}P(0)\sin\frac{K\tau\pi}{3} + \frac{2}{\sqrt{3}}\overline{P}(0)\sin\frac{(1+K\tau)\pi}{3}. && (29)
\end{aligned}$$

For $\tau = \Delta t = 1/K$, we have

$$\begin{aligned}
Q(\Delta t) &= Q(0) - \overline{Q}(0), && (30)\\
\overline{Q}(\Delta t) &= Q(0), && (31)\\
P(\Delta t) &= -\overline{P}(0), && (32)\\
\overline{P}(\Delta t) &= P(0) + \overline{P}(0). && (33)
\end{aligned}$$

Journal of Physics: Conference Series **504** (2014) 012024 doi:10.1088/1742-6596/504/1/012024

As in the von Neumann model, the value of the observable $Q(0)$ is measured by the value of the meter observable $\overline{Q}(\Delta t)$. Since $Q(0)$ and $\overline{Q}(\Delta t)$ commute, as seen from Eq. (31), we have the joint probability distribution $\mu^{Q(0),\overline{Q}(\Delta t)}(dx,dy)$ of the values of $Q(0)$ and $\overline{Q}(\Delta t)$ by Eq. (17). Then, from Eq. (16) the rms error $\varepsilon(Q)$ of $\overline{Q}(\Delta t)$ for $Q(0)$ in $|\psi\rangle$ is given by

$$\begin{aligned}\varepsilon(Q) &= \left(\iint_{\mathbf{R}^2}(y-x)^2\mu^{Q(0),\overline{Q}(\Delta t)}(dx,dy)\right)^{1/2}\\ &= \langle(\overline{Q}(\Delta t)-Q(0))^2\rangle^{1/2}\\ &= 0.\end{aligned} \tag{34}$$

Since $P(0)$ and $P(\Delta t)$ commute, as seen from Eq. (32), we have the joint probability distribution $\mu^{P(0),P(\Delta t)}(dx,dy)$ of the values of $P(0)$ and $P(\Delta t)$ by Eq. (19). The rms disturbance $\eta(P)$ of $P$ from $t=0$ to $t=\Delta t$ is given by

$$\begin{aligned}\eta(P) &= \left(\iint_{\mathbf{R}^2}(y-x)^2\mu^{P(0),P(\Delta t)}(dx,dy)\right)^{1/2}\\ &= \langle(P(\Delta t)-P(0))^2\rangle^{1/2}\\ &= \langle(\overline{P}(0)+P(0))^2\rangle^{1/2}<\infty.\end{aligned} \tag{35}$$

Consequently, we have

$$\varepsilon(Q)\eta(P)=0. \tag{36}$$

Therefore, our model obviously violates the Heisenberg EDR (2).

Taking advantage of the above model, the argument was refuted that the uncertainty principle generally leads to the SQL claimed in Ref. [18] for monitoring free-mass position [19, 20].

If $\langle P(0)^2\rangle\to 0$ and $\langle\overline{P}(0)^2\rangle\to 0$ (i.e., $|\psi\rangle$ and $|\xi\rangle$ tend to the momentum eigenstate with zero momentum) then we have even $\eta(P(t))\to 0$ with $\varepsilon(Q)=0$. Thus, we can measure position precisely without effectively disturbing momentum in a near momentum eigenstate; see Ref. [24] for detailed discussion on the quantum state reduction caused by the above model.

As shown above, the Heisenberg EDR (2) is taken to be a breakable limit [25], but then the problem remains: what is the unbreakable constraint between error and disturbance, which Heisenberg originally intended?

## 7. Universally valid EDR

In 2003, the present author [26, 4, 27] showed the relation

$$\varepsilon(A)\eta(B)+|\langle[n(A),B]\rangle+\langle[A,d(B)]\rangle|\ge\frac{1}{2}\,|\langle[A,B]\rangle|\,, \tag{37}$$

which is universally valid for any observables $A,B$, any system state, and any measuring apparatus, where $n(A)$ and $d(B)$ are system observables representing the first moments of the error and the disturbance for $A$ and $B$, respectively. ¿From Eq. (37), it is concluded that if the error and the disturbance are statistically independent from system state, then the Heisenberg EDR

$$\varepsilon(A)\eta(B)\ge\frac{1}{2}|\langle[A,B]\rangle| \tag{38}$$

holds, extending the previous results [28, 29, 16, 17]. The additional correlation term in Eq. (37) allows the error-disturbance product $\epsilon(A)\eta(B)$ to violate the Heisenberg EDR (38). In general, the relation

$$\epsilon(A)\eta(B)+\epsilon(A)\sigma(B)+\sigma(A)\eta(B)\ge\frac{1}{2}\,|\langle[A,B]\rangle| \tag{39}$$

holds for any observables $A, B$, any system state, and any measuring apparatus [26, 4, 27, 30, 31, 32].

The new relation (39) leads to the following new constraints for error-free measurements and non-disturbing measurements: if $\varepsilon(A) = 0$ then

$$\sigma(A)\eta(B) \geq \frac{1}{2} |\langle [A, B] \rangle| , \tag{40}$$

and if $\eta(B) = 0$ then

$$\epsilon(A)\sigma(B) \geq \frac{1}{2} |\langle [A, B] \rangle| . \tag{41}$$

Note that if $\langle [A, B] \rangle \neq 0$, Heisenberg EDR (38) leads to divergences in both cases. The new error bound Eq. (41) was used to derive conservation-law-induced limits for measurements [27, 33] (see also [34, 35]) quantitatively generalizing the Wigner-Araki-Yanase theorem [36, 37, 38, 39] and was used to derive an accuracy limit for quantum computing induced by conservation laws [27] (see also [40, 41, 42, 43, 44, 45, 46]).

## 8. Operator formalism for error and disturbance

To derive the above relations, consider a measuring process $\mathbf{M} = (\mathcal{K}, |\xi\rangle, U, M)$ determined by the probe system $\mathbf{P}$ described by a Hilbert space $\mathcal{K}$, the initial probe state $|\xi\rangle$, the unitary evolution $U$ of the composite system $\mathbf{S} + \mathbf{P}$ during the measuring interaction, and the meter observable $M$ of the probe $\mathbf{P}$ to be directly observed [4]. We assume that the measuring interaction turns on at time $t = 0$ and turns off at time $t = \Delta t$. In the Heisenberg picture, we write

$$A_1(0) = A_1 \otimes I, \quad A_2(0) = I \otimes A_2, \quad A_{12}(\Delta t) = U^\dagger A_{12}(0) U,$$

for an observable $A_1$ of $\mathbf{S}$, an observable $A_2$ of $\mathbf{P}$, and an observable $A_{12}(0)$ of $\mathbf{S} + \mathbf{P}$.

The *error observable* $N(A)$ representing the difference between the measured observable $A(0)$ and the meter observable $M(\Delta t)$ to be read and the *disturbance observable* $D(A)$ representing the change in $B$ caused by the measuring interaction are defined by

$$N(A) = M(\Delta t) - A(0), \tag{42}$$
$$D(B) = B(\Delta t) - B(0). \tag{43}$$

The *mean error operator* $n(A)$ and the *mean disturbance operator* $d(B)$ in Eq. (37) are defined by

$$n(A) = \langle \xi | N(A) | \xi \rangle, \tag{44}$$
$$d(B) = \langle \xi | D(B) | \xi \rangle. \tag{45}$$

The *(root-mean-square) error* $\varepsilon(A, \rho)$ and the *(root-mean-square) disturbance* $\eta(B, \rho)$ for observables $A, B$ and state (density operator) $\rho$ on a Hilbert space $\mathcal{H}$ were defined by

$$\varepsilon(A, \rho)^2 = \mathrm{Tr}[N(A)^2 \rho \otimes |\xi\rangle\langle\xi|], \tag{46}$$
$$\eta(B, \rho)^2 = \mathrm{Tr}[D(B)^2 \rho \otimes |\xi\rangle\langle\xi|]. \tag{47}$$

The definition of $\varepsilon(A) = \varepsilon(A, \rho)$ is uniquely derived from the classical notion of root-mean-square error if $M(\Delta t)$ and $A(0)$ commute [47], as in the models discussed in the previous sections. Otherwise, it is considered as a natural quantization of the notion of classical root-mean-square error. It is also pointed out that $\varepsilon(A)$ coincides with the root-mean-square error of a quantum estimator for an orthogonal pure state estimation problem with the uniform prior distribution

[48]. The definition of $\eta(B) = \eta(B, \rho)$ is derived analogously, although there are recent debates on alternative approaches [49, 50, 51, 47].

In particular, Busch, Heinonen, and Lahti [49] pointed out that there is a case in which $\varepsilon(A) = 0$ holds but $A$ cannot be considered to be measured precisely. In response to this, we characterized the case where $A$ is measured precisely as follows [52, 53].

We say that the measuring process $\mathbf{M}$ *precisely measures* an observable $A$ in a state $\rho$ if observables $A(0)$ and $M(\Delta t)$ commute in the state $\rho\otimes|\xi\rangle\langle\xi|$ and the joint probability distribution $\mu^{A(0),M(\Delta t)}$ of $A(0)$ and $M(\Delta t)$ concentrates on the diagonal, i.e.,

$$\mu^{A(0),M(\Delta t)}(\{(x, y) \in \mathbf{R}^2 \mid x \neq y\} = 0. \tag{48}$$

The *weak joint distribution* $\mu_W^{A(0),M(\Delta t)}$ of $A(0)$ and $M(\Delta t)$ in a state $\rho$ is defined by

$$\mu_W^{A(0),M(\Delta t)}(dx, dy) = \langle E^{A(0)}(dx)E^{M(\Delta t)}(dy)\rangle. \tag{49}$$

The joint probability distribution $\mu^{A(0),M(\Delta t)}$ exists only when $A(0)$ and $M(\Delta t)$ commute in the state $\rho \otimes |\xi\rangle\langle\xi|$, while the weak joint distribution $\mu_W^{A(0),M(\Delta t)}$ always exists. The *cyclic subspace* $\mathcal{C}(A, \rho)$ generated by $A$ and $\rho$ is defined as the closed subspace of $\mathcal{H}$ generated by $\{E^A(\Delta)|\phi\rangle \mid \Delta \in \mathcal{B}(\mathbf{R}), |\phi\rangle \in \rho\mathcal{H}\}$, where $\mathcal{B}(\mathbf{R})$ is the Borel $\sigma$-field of the real line $\mathbf{R}$. A *generating subset* of $\mathcal{C}(A, \rho)$ is a set $\mathcal{S}$ of vector states $|\phi\rangle \in \mathcal{H}$ such that $\mathcal{S}^{\perp\perp} = \mathcal{C}(A, |\phi\rangle)$, where $\perp$ stands for the orthogonal complement. Then, the following theorem holds [52, 53].

**Theorem 1** *Let $\mathbf{M} = (\mathcal{K}, |\xi\rangle, U, M)$ be a measuring process for the system $\mathbf{S}$ described by a Hilbert space $\mathcal{H}$. Let $A$ be an observable of $\mathbf{S}$ and $\rho$ be a state of $\mathbf{S}$. Then, the following conditions are equivalent.*

*(i) The measuring process $\mathbf{M}$ precisely measures observable $A$ in state $\rho$.*

*(ii) The weak joint distribution $\mu_W^{A(0),M(\Delta t)}$ in state $\rho$ concentrates on the diagonal, i.e.,*

$$\langle E^{A(0)}(\Delta)E^{M(\Delta t)}(\Gamma)\rangle = 0$$

*if $\Delta \cap \Gamma = \emptyset$.*

*(iii) $\varepsilon(A, |\phi\rangle) = 0$ for all $|\phi\rangle \in \mathcal{C}(A, \rho)$.*

*(iv) There exists a generating subset $\mathcal{S}$ of $\mathcal{C}(A, \rho)$ such that $\varepsilon(A, |\phi\rangle) = 0$ for all $|\phi\rangle \in S$.*

We say that the measuring process $\mathbf{M}$ *does not disturb* an observable $B$ in a state $\rho$ if observables $B(0)$ and $B(\Delta t)$ commute in the state $\rho\otimes|\xi\rangle\langle\xi|$ and the joint probability distribution $\mu^{B(0),B(\Delta t)}$ of $B(0)$ and $B(\Delta t)$ concentrates on the diagonal. The non-disturbing measuring processes defined above can be characterized analogously.

¿From the above theorem, we can conclude that $\varepsilon(A)$ is negatively biased in the sense that positive error $\varepsilon(A) > 0$ always implies that the $A$ cannot be measured precisely. Thus, a non-zero lower bound for $\varepsilon(A)$ indicates a limitation for precise measurements. For $\eta(B) > 0$ we have an analogous conclusion. Moreover, the above characterizations of precise and non-disturbing measurements lead to the following definitions of the *locally uniform root-mean-squre error* $\overline{\varepsilon}(A, \rho)$ and the *locally uniform root-mean-squre disturbance* $\overline{\eta}(B, \rho)$ [54]:

$$\overline{\varepsilon}(A, \rho) = \sup_{|\phi\rangle\in\mathcal{S}} \varepsilon(A, |\phi\rangle), \tag{50}$$

$$\overline{\eta}(B, \rho) = \sup_{|\phi\rangle\in\mathcal{S}} \eta(B, |\phi\rangle), \tag{51}$$

where $\mathcal{S} = \mathcal{C}(A,\rho)$ or $\mathcal{S}$ is a generating subset of $\mathcal{C}(A,\rho)$. Then, it is shown that $\overline{\varepsilon}(A,\rho) = 0$ if and only if the measurement precisely measures $A$ in $\rho$, and that $\overline{\eta}(B,\rho) = 0$ if and only if the measurement does not disturb $B$ in $\rho$. For those quantities, Heisenberg's EDR

$$\overline{\varepsilon}(Q,\rho)\overline{\eta}(P,\rho) \geq \frac{\hbar}{2} \tag{52}$$

is still violated by a linear position measurement [54], and the relation

$$\overline{\varepsilon}(A)\overline{\eta}(B) + \overline{\varepsilon}(A)\sigma(B) + \sigma(A)\overline{\eta}(B) \geq \frac{1}{2}|\langle[A,B]\rangle| \tag{53}$$

holds universally [54], where $\overline{\varepsilon}(A) = \overline{\varepsilon}(A,\rho)$ and $\overline{\eta}(B) = \overline{\eta}(B,\rho)$.

## 9. Experimental tests

There has been a controversy [55, 56] on the question as to whether the rms error and rms disturbance are experimentally accessible without knowing the details of measuring process $(\mathcal{K}, |\xi\rangle, U, M)$ and the state $\rho$. To clear this question two methods have been proposed so far: the "three-state method" proposed by the present author [31] and the "weak-measurement method" proposed by Lund-Wiseman [57] based on the relation between the rms error/distrubance and the weak joint distribution [20, 16, 52]. The three-state method was demonstrated for qubit systems: projective measurement of a neutron-spin qubit [58, 59] and generalized measurement of a photon-polarization qubit [60]. The weak-measurement method was demonstrated for generalized measurement of a photon-polarization qubit by Rozema *et al* [61], Baek *et al* [62], and Ringbauer *et al.* [63]. All the above experiments observed that the Heisenberg EDR (38) does not hold, while the universally valid EDR (39) and a new stronger universally valid EDR recently proposed by Branciard [64] hold. Very recently, we have proposed the third method, the "two-point quantum correlator method," for measuring rms error and disturbance with an experimental proposal for quit measurements in Ref. [65].

## Acknowledgments

This work was supported by MIC SCOPE, No. 121806010, and the John Templeton Foundation, ID #35771.

## References

[1] Heisenberg W 1983 *Quantum Theory and Measurement* ed Wheeler J A and Zurek W H (Princeton, NJ: Princeton University Press) pp 62–84 [Translation into English by J. A. Wheeler and W. H .Zurek, 1981. Originally published under the title, "Über den anschaulichen Inhalt der quantemtentheoretischen Kinematik und Mechanik," Zeitschrift für Physik, 43, 172-98 (1927).]

[2] Gauss C F 1995 *Theoria Combinationis Observationum Erroribus Miinimis Obnoxiae, Pars Prior, Pars Posterior, Supplementum (Societati Regiae Scientiarum Exhibita, Feb. 15, 1821) [Theory of the Combination of Observations Least Subject to Errors: Part One, Part Two, Supplement, translated by G. W. Stewart]* (Philadelphia, USA: Society for Industrial and Applied Mathematics)

[3] Kennard E H 1927 *Z. Phys.* **44** 326–352

[4] Ozawa M 2003 *Phys. Rev. A* **67** 042105

[5] von Neumann J 1955 *Mathematical Foundations of Quantum Mechanics* (Princeton, NJ: Princeton University Press) [English translation of *Mathematische Grundlagen der Quantenmechanik* (Springer, Berlin, 1932)]

[6] Davies E B and Lewis J T 1970 *Commun. Math. Phys.* **17** 239–260

[7] Ozawa M 1984 *J. Math. Phys.* **25** 79–87

[8] Bohm D 1951 *Quantum Theory* (New York: Prentice-Hall)

[9] Messiah A 1959 *Mécanique Quantique* vol I (Paris: Dunod) [*Quantum Mechanics*, Vol. I (North-Holland, Amsterdam, 1959)]

[10] Schiff L I 1968 *Quantum Mechanics* (New York: MacGraw-Hill)
[11] Reed M and Simon B 1980 *Methods of Modern Mathematical Physics, I: Functional Analysis (Reivsed and Enlarged Edition)* (New York: Academic)
[12] Halmos P R 1951 *Introduction to Hilbert Space and the Theory of Spectral Multiplicity* (New York: Chelsea)
[13] Caves C M, Thorne K S, Drever R W P, Sandberg V D and Zimmermann M 1980 *Rev. Mod. Phys.* **52** 341–392
[14] Caves C M 1985 *Phys. Rev. Lett.* **54** 2465–2468
[15] Arthurs E and Kelly, Jr J L 1965 *Bell. Syst. Tech. J.* **44** 725–729
[16] Ozawa M 1991 *Quantum Aspects of Optical Communications* ed Bendjaballah C, Hirota O and Reynaud S (Berlin: Springer) pp 3–17
[17] Ishikawa S 1991 *Rep. Math. Phys.* **29** 257–273
[18] Braginsky V B, Vorontsov Y I and Thorne K S 1980 *Science* **209** 547–557
[19] Yuen H P 1983 *Phys. Rev. Lett.* **51** 719–722 [see also *ibid.* p. 1603]
[20] Ozawa M 1988 *Phys. Rev. Lett.* **60** 385–388
[21] Ozawa M 1989 *Squeezed and Nonclassical Light* ed Tombesi P and Pike E R (New York: Plenum) pp 263–286
[22] Maddox J 1988 *Nature* **331** 559
[23] Ozawa M 2002 *Phys. Lett. A* **299** 1–7
[24] Ozawa M 2001 *Phys. Lett. A* **282** 336–342
[25] Giovannetti V, Lloyd S and Maccone L 2004 *Science* **306** 1330–1336
[26] Ozawa M 2003 *Phys. Lett. A* **318** 21–29
[27] Ozawa M 2003 *Int. J. Quant. Inf.* **1** 569–588
[28] Arthurs E and Goodman M S 1988 *Phys. Rev. Lett.* **60** 2447–2449
[29] Raymer M G 1994 *Am. J. Phys.* **62** 986–993
[30] Ozawa M 2004 *Phys. Lett. A* **320** 367–374
[31] Ozawa M 2004 *Ann. Phys. (N.Y.)* **311** 350–416
[32] Ozawa M 2005 *J. Opt. B: Quantum Semiclass. Opt.* **7** S672–S681
[33] Ozawa M 2004 *AIP Conf. Proc.* **734** 95–98
[34] Ozawa M 2002 *Phys. Rev. Lett.* **88** 050402
[35] Busch P and Loveridge L 2011 *Phys. Rev. Lett.* **106** 110406
[36] Wigner E P 1952 *Z. Phys.* **133** 101–108
[37] Araki H and Yanase M M 1960 *Phys. Rev.* **120** 622–626
[38] Yanase M M 1961 *Phys. Rev.* **123** 666–668
[39] Ozawa M 1991 *Phys. Rev. Lett.* **67** 1956–1959
[40] Ozawa M 2002 *Phys. Rev. Lett.* **89** 057902
[41] Gea-Banacloche J 2002 *Phys. Rev. Lett.* **89** 022308
[42] Ozawa M 2003 *Proceedings of the Sixth International Conference on Quantum Communication, Measurement and Computing* ed Shappiro J H and Hirota O (Princeton: Rinton Press) pp 175–180
[43] Gea-Banacloche J and Ozawa M 2005 *J. Opt. B: Quantum Semiclass. Opt.* **7** S326–S332
[44] Gea-Banacloche J and Ozawa M 2006 *Phys. Rev. A* **74** 060301(R)
[45] Karasawa T and Ozawa M 2007 *Phys. Rev. A* **75** 032324
[46] Karasawa T, Gea-Banacloche J and Ozawa M 2009 *J. Phys. A: Math. Theor.* **42** 225303
[47] Ozawa M 2013 Disproving Heisenberg's error-disturbance relation arXiv:1308.3540v1 [quant-ph]
[48] Buscemi F, Hall M J, Ozawa M and Wilde M M 2014 *Phys. Rev. Lett.* **112** 050401
[49] Busch P, Heinonen T and Lahti P 2004 *Phys. Lett. A* **320** 261–270
[50] Busch P, Lahti P and Werner R F 2013 *Phys. Rev. Lett.* **111** 160405
[51] Rozema L A, Mahler D H, Hayat A and Steinberg A M 2013 A note on different definitions of momentum disturbance arXiv:1307.3604v1 [quant-ph]
[52] Ozawa M 2005 *Phys. Lett. A* **335** 11–19
[53] Ozawa M 2006 *Ann. Phys. (N.Y.)* **321** 744–769
[54] Ozawa M 2006 *Proc. SPIE* **6244** 62440Q
[55] Werner R F 2004 *Quantum Inf. Comput.* **4** 546–562
[56] Koshino K and Shimizu A 2005 *Phys. Rep.* **412** 191–275
[57] Lund A P and Wiseman H M 2010 *New J. Phys.* **12** 093011
[58] Erhart J, Sponar S, Sulyok G, Badurek G, Ozawa M and Hasegawa Y 2012 *Nature Phys.* **8** 185–189
[59] Sulyok G, Sponar S, Erhart J, Badurek G, Ozawa M and Hasegawa Y 2013 *Phys. Rev. A* **88** 022110
[60] Baek S Y, Kaneda F, Ozawa M and Edamatsu K 2013 *Sci. Rep.* **3** 2221
[61] Rozema L A, Darabi A, Mahler D H, Hayat A, Soudagar Y and Steinberg A M 2012 *Phys. Rev. Lett.* **109** 100404
[62] Kaneda F, Baek S Y, Ozawa M and Edamatsu K 2014 *Phys. Rev. Lett.* **112** 020402

[63] Ringbauer M, Biggerstaff D N, Broome M A, Fedrizzi A, Branciard C and White A G 2014 *Phys. Rev. Lett.* **112** 020401

[64] Branciard C 2013 *Proc. Natl. Acad. Sci. USA* **110** 6742–6747

[65] Buscemi F, Dall'Arno M, Ozawa M and Vedral V 2013 Direct observation of any two-point quantum correlation function arXiv:1312.4240 [quant-ph]