Error-Disturbance Relation in Stern-Gerlach Measurements

Doctoral Thesis

YUKI INOUE

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

January 22, 2021

Contents

Abstract

Heisenberg's uncertainty principle is represented by a rigorously proven relation about intrinsic uncertainties in quantum states. On the other hand, Heisenberg's error-disturbance-relation (EDR) has been commonly believed to be another aspect of the principle. However, recent studies of quantum measurements have revealed the violation of Heisenberg's EDR. Furthermore, a universally valid error-disturbance relation was obtained and experimentally tested with neutrons and with photons, respectively. These results indicate that Heisenberg's EDR is violated by other measurements.

We investigate the error and disturbance of Stern-Gerlach measurements of a spin-1/2 particle. Here, we determine the range of the possible values of the error and disturbance for arbitrary Stern-Gerlach apparatuses with the orbital degree prepared in an arbitrary Gaussian state. We show that their error-disturbance region is close to the theoretical optimal and actually violates Heisenberg's EDR in a broad range of experimental parameters. We also show the existence of orbital states in which the error is minimized by the screen at a finite distance from the magnet, in contrast to the standard assumption. We further report that even the original Stern-Gerlach experiment in 1922, the available experimental data show, violates Heisenberg's EDR. The results suggest that Heisenberg's EDR is more ubiquitously violated than it has long been supposed.

Acknowledgments

First and foremost, I would like to express my deepest appreciation to my supervisor, Professor Masanao Ozawa. Without his sincere guidance and persistent help, this thesis would have never been accomplished. I would also like to show my appreciation to Professor Francesco Buscemi, Professor Harumichi Nishimura and other professors in Department of Computer Science and Mathematical Informatics for thTeir sincere guidance and encouragement. Furthermore, I would like to express special thanks to Dr. Kazuya Okamura for helpful discussions in our research.

Chapter 1

Introduction

A fundamental feature of quantum measurement is nontrivial error-disturbance relations (EDRs), first found by Heisenberg [22], who, using the famous γ-ray microscope thought experiment, derived the relation

$$\varepsilon(Q)\eta(P) \geq \frac{\hbar}{2} \tag{1.1}$$

between the position measurement error $\varepsilon(Q)$ and the momentum disturbance $\eta(P)$ thereby caused. His formal derivation of this relation from the well-established relation

$$\sigma(Q)\sigma(P) \geq \frac{\hbar}{2} \tag{1.2}$$

for standard deviations $\sigma(Q)$ and $\sigma(P)$, due to Heisenberg [22] for the minimum uncertainty wave packets and Kennard [28] for arbitrary wave functions, needs an additional assumption on the state change caused by the measurement [44].

Nowadays, the state change caused by a measurement is generally described by a completely positive (CP) instrument, a family of CP maps summing to a trace-preserving CP map [32]. In such a general description of quantum measurements, Heisenberg's EDR (1.1) loses its universal validity, as revealed in the debate in the 1980s on the sensitivity limit for gravitational wave detection derived by Heisenberg's EDR (1.1), but settled questioning the validity of Heisenberg's EDR [4, 12, 58, 11, 33, 34]. A universally valid error-disturbance relation for arbitrary pairs of observables

$$\varepsilon(A)\eta(B) + \varepsilon(A)\sigma(B) + \sigma(A)\eta(B) \geq \frac{1}{2}|\langle[A, B]\rangle|, \tag{1.3}$$

where $\sigma(A)$ and $\sigma(B)$ are the standard deviations of A and B just before the measurement, was derived by Ozawa [37, 36, 38] and has recently received considerable attention. The validity of this relation, as well as a stronger version of this relation [5, 6, 43, 45], was experimentally tested with neutrons [31, 16, 54, 14] and with photons [48, 1, 56, 27, 47]. Other approaches generalizing Heisenberg's original relation (1.1) can be found, for example, in [9, 10, 30], apart from the information-theoretic approach [7, 53].

Stern-Gerlach measurements [18, 19, 20] are among the most important quantum measurements, and a number of theoretical analyses are available from many authors. In his famous textbook (see [3], p. 596), Bohm derived the wave function of a spin-$1/2$ particle that has passed through the Stern-Gerlach apparatus. In his argument, he assumed that the magnetic field points in the same direction everywhere and varies in strength linearly with the z coordinate of the position as

$$\mathbf{B} = \begin{pmatrix} 0 \\ 0 \\ B_0 + B_1 z \end{pmatrix}. \tag{1.4}$$

However, as Bohm pointed out (see [3], p. 594), such a magnetic field does not satisfy Maxwell's equations. Theoretical studies [51, 13, 46] of Stern-Gerlach measurements with the magnetic field

$$\mathbf{B} = \begin{pmatrix} -B_1 x \\ 0 \\ B_0 + B_1 z \end{pmatrix} \tag{1.5}$$

satisfying Maxwell's equations were performed only recently. According to these studies, if the magnetic field in the center of the beam is sufficiently strong, the precession of the spin component to be measured becomes small, and hence Bohm's approximation (1.4) holds.

Home *et al.* [23] investigated the error of Stern-Gerlach measurements with respect to the distinguishability of apparatus states. As an indicator of the operational distinguishability of apparatus states, they used the error integral, which is equal to the probability of finding the particle in the spin-up state on the lower half of the screen. They analyzed the error integral in the case where the spin state of the particle just before the measurement is the eigenstate $|\uparrow\rangle_z$ of σ_z corresponding to the eigenvalue $+1$. Nevertheless, the trade-off between the error and disturbance in Stern-Gerlach measurements has not been studied in the literature, even though the subject would elucidate the fundamental limitations of measurements in quantum theory, as Heisenberg did with the γ-ray microscope thought experiment.

In this thesis, we determine the range of the possible values of the error and disturbance for arbitrary Stern-Gerlach apparatuses, based on the general theory of the error and disturbance, which has recently been developed to establish universally valid reformulations of Heisenberg's uncertainty relation. Throughout this thesis, we consider an electrically neutral particle with spin $1/2$. Following Bohm [3], we assume that the magnetic field of a Stern-Gerlach apparatus is represented by Eq. (1.4), which is assumed to be sufficiently strong. The particle is assumed to stay in the magnet from time 0 to time Δt. Only the one-dimensional orbital degree of freedom along the z axis is considered. The kinetic energy is not neglected. The particle having passed through the magnetic field is assumed to evolve freely from time Δt to $\Delta t + \tau$. The initial state of the spin of the particle is assumed to be arbitrary. The initial state of the orbital degree of freedom is such that mean values of the position and momentum are both 0.

We study in detail the error $\varepsilon(\sigma_z)$ in measuring σ_z with a Stern-Gerlach apparatus and the disturbance $\eta(\sigma_x)$ caused thereby on σ_x for the orbital degree of freedom to be prepared in a Gaussian pure state [50]. We obtain the EDR

$$\left|\frac{\eta(\sigma_x)^2 - 2}{2}\right| \leq \exp\left\{-\left[\mathrm{erf}^{-1}\left(\frac{\varepsilon(\sigma_z)^2 - 2}{2}\right)\right]^2\right\} \tag{1.6}$$

for Stern-Gerlach measurements, where erf^{-1} represents the inverse of the error function $\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x \exp(-s^2)ds$. We compare the above EDR with Heisenberg's EDR for spin measurements

$$\varepsilon(\sigma_z)^2\eta(\sigma_z)^2 \geq 1, \tag{1.7}$$

which holds for measurements with statistically independent error and disturbance [37, 38]. We show that Stern-Gerlach measurements violate Heisenberg's EDR in a broad range of experimental parameters. We also compare it with the EDR

$$\left|\frac{\eta(\sigma_x)^2 - 2}{2}\right| \leq 1 - \left(\frac{\varepsilon(\sigma_z)^2 - 2}{2}\right)^2, \tag{1.8}$$

which holds for improperly directed projective measurements experimentally tested with neutron spin measurements conducted by Hasegawa and co-workers [16, 54], and the tight EDR

$$\left|\frac{\eta(\sigma_x)^2 - 2}{2}\right| \leq \sqrt{1 - \left(\frac{\varepsilon(\sigma_z)^2 - 2}{2}\right)^2} \tag{1.9}$$

for the range of $(\varepsilon(\sigma_z), \eta(\sigma_x))$ values of arbitrary qubit measurements obtained by Branciard and Ozawa [5, 6, 43] [see also Eq. (2.34) below].

We further show that according to the available experimental data, even the original Stern-Gerlach experiment performed in 1922 [18, 19, 20] violates Heisenberg's EDR. The results suggest that Heisenberg's EDR is more ubiquitously violated than it has been supposed for a long time.

In Chapter 2, the general theory of the error and disturbance is reviewed and Stern-Gerlach measurements are investigated in the Heisenberg picture in detail. In Chapter 3, the error and disturbance of Stern-Gerlach measurements and their EDR are derived. In Chapter 4, we show that the original Stern-Gerlach experiment violates the Heisenberg's error-disturbance relation. Chapter 5 is devoted to the conclusion.

Chapter 2

Preliminaries

2.1 Error and disturbance in quantum measurements

In this section, we review the general theory of error and disturbance in quantum measurements developed in [38, 45].

2.1.1 Classical root-mean-square error

Let us consider the classical case first. Recall the root-mean-square (rms) error introduced by Gauss [17]. Consider a measurement of the value x of a quantity X by actually observing the value y of a meter quantity Y. Then the error of this measurement is given by $y - x$. If these quantities obey a joint probability distribution $\mu(x, y)$, then the rms error $\varepsilon_G(\mu)$ is defined as

$$\varepsilon_G(\mu) = \left(\sum_{x,y} (y - x)^2 \, \mu(x, y) \right)^{1/2}. \tag{2.1}$$

2.1.2 Quantum measuring processes

We consider a quantum system $\mathbf{S}$ described by a finite-dimensional Hilbert space $\mathcal{H}$. We assume that every measuring apparatus for the system $\mathbf{S}$ has its own output variable $\mathbf{x}$. The statistical properties of the apparatus $\mathbf{A}(\mathbf{x})$ having the output variable $\mathbf{x}$ are determined by (i) the probability distribution $\Pr\{\mathbf{x} = m \| \rho\}$ of $\mathbf{x}$ for the input state ρ, and (ii) the output state $\rho_{\{\mathbf{x}=m\}}$ given the outcome $\mathbf{x} = m$.

A measuring process of the apparatus $\mathbf{A}(\mathbf{x})$ measuring $\mathbf{S}$ is specified by a quadruple $\mathbf{M} = (\mathcal{K}, |\xi\rangle, U, M)$ consisting of a Hilbert space $\mathcal{K}$ describing the probe system $\mathbf{P}$, a state vector $|\xi\rangle$ in $\mathcal{K}$ describing the initial state of $\mathbf{P}$, a unitary operator U on $\mathcal{H} \otimes \mathcal{K}$ describing the time evolution of the composite system $\mathbf{S}+\mathbf{P}$ during the measuring interaction, and an observable, M, called the meter observable, of $\mathbf{P}$ describing the meter of the apparatus.

The instrument of the measuring process $\mathbf{M}$ is defined as a completely positive map valued function $\mathcal{I}$ given by

$$\mathcal{I}(m)\rho = \mathrm{Tr}_{\mathcal{K}}[(\mathbb{1} \otimes P^M(m))U(\rho \otimes |\xi\rangle\langle\xi|)U^\dagger] \tag{2.2}$$

for any state ρ and real number m. The statistical properties of the apparatus $\mathbf{A}(\mathbf{x})$ are determined by the instrument $\mathcal{I}$ of $\mathbf{M}$ as

$$\Pr\{\mathbf{x} = m \| \rho\} = \mathrm{Tr}[\mathcal{I}(m)\rho], \tag{2.3}$$

$$\rho_{\{\mathbf{x}=m\}} = \frac{\mathcal{I}(m)\rho}{\mathrm{Tr}[\mathcal{I}(m)\rho]}. \tag{2.4}$$

The non-selective operation T of $\mathbf{M}$ is defined by

$$T = \sum_{m \in \mathbb{R}} \mathcal{I}(m). \tag{2.5}$$

Then we have

$$T(\rho) = \mathrm{Tr}_{\mathcal{K}}[U(\rho \otimes |\xi\rangle\langle\xi|)U^{\dagger}]. \tag{2.6}$$

See Refs. [32, 34, 38] for detailed descriptions of measuring processes and instruments.

2.1.3 Heisenberg picture

In the measuring process $\mathbf{M}$, we suppose that the measuring interaction is turned on from time $t = 0$ to time $t = t_0$. Then, the outcome $\mathbf{x} = m$ of the apparatus $\mathbf{A}(\mathbf{x})$ described by the measuring process $\mathbf{M}$ is defined as the outcome m of the meter measurement at time $t = t_0$. To describe the time evolution of the composite system $\mathbf{S} + \mathbf{P}$ in the Heisenberg picture, let

$$\begin{array}{rclrcl} A(0) & = & A \otimes \mathbb{1}, & A(t_0) & = & U^{\dagger}A(0)U, \\ B(0) & = & B \otimes \mathbb{1}, & B(t_0) & = & U^{\dagger}B(0)U, \\ M(0) & = & \mathbb{1} \otimes M, & M(t_0) & = & U^{\dagger}M(0)U, \end{array} \tag{2.7}$$

where A and B are observables of $\mathbf{S}$.

Then, the POVM Π of $\mathbf{M}$ is defined as

$$\Pi(m) = \langle\xi|P^{M(t_0)}(m)|\xi\rangle \tag{2.8}$$

and satisfies

$$\Pr\{\mathbf{x} = m \| \rho\} = \mathrm{Tr}[\Pi(m)\rho]. \tag{2.9}$$

The n-th moment operator of Π for $n = 1, \ldots, n$ is defined by

$$\hat{\Pi}^{(n)} = \langle \xi \,| M(t_0)^n |\, \xi \rangle \,. \tag{2.10}$$

The dual non-selective operation T^* of $\mathbf{M}$ is defined by

$$T^*(B) = \langle\xi|B(t_0)|\xi\rangle \tag{2.11}$$

for any observable B of $\mathbf{S}$ and satisfies

$$\mathrm{Tr}\left\{[T^*(B)]\,\rho\right\} = \mathrm{Tr}\left\{B\,[T(\rho)]\right\} \tag{2.12}$$

for any observable B and state ρ.

2.1.4 Measurement of observables

If the observables $A(0)$ and $M(t_0)$ commute in the initial state $\rho \otimes |\xi\rangle\langle\xi|$, that is,

$$[P^{A(0)}(a), P^{M(t_0)}(m)](\rho \otimes |\xi\rangle\langle\xi|) = 0 \tag{2.13}$$

for all $a, m \in \mathbb{R}$, then their joint probability distribution $\mu(a, m)$ is defined as

$$\mu(a, m) = \mathrm{Tr}[P^{A(0)}(a)P^{M(t_0)}(m)\,(\rho \otimes |\xi\rangle\langle\xi|)] \tag{2.14}$$

and satisfies

$$\mathrm{Tr}[f(A(0), M(t_0))(\rho \otimes |\xi\rangle\langle\xi|)] = \sum_{a,m} f(a, m)\,\mu(a, m) \tag{2.15}$$

for any polynomial $f(A(0), M(t_0))$ of $A(0)$ and $M(t_0)$.

We say that the measuring process $\mathbf{M}$ accurately measures the observable A in a state ρ if $A(0)$ and $M(t_0)$ are perfectly correlated in the state $\rho \otimes |\xi\rangle\langle\xi|$ [39, 42, 45], namely, one of the following two equivalent conditions holds: (i) $A(0)$ and $M(t_0)$ commute in $\rho \otimes |\xi\rangle\langle\xi|$ and their joint probability distribution μ satisfies

$$\sum_{a,m:a=m} \mu(a, m) = 1 \tag{2.16}$$

or (ii) for any $a, m \in \mathbb{R}$ with $a \neq m$,

$$\mathrm{Tr}\left[\Pi(m)P^{A}(a)\,\rho\right] = 0. \tag{2.17}$$

Note that $\nu(a, m) := \mathrm{Tr}\left[\Pi(m)P^{A}(a)\,\rho\right]$, called the weak joint distribution of $A(0)$ and $M(t_0)$, always exists and is operationally accessible by weak measurement and post-selection [26, 31], but possibly takes negative or complex values. Since $\nu(a, m)$ is operationally accessible, our definition of accurate measurements is operationally accessible.

2.1.5 Quantum root-mean-square error

The noise operator $N(A, \mathbf{M})$ of the measuring process $\mathbf{M}$ for measuring A is defined as

$$N(A, \mathbf{M}) = M(t_0) - A(0). \tag{2.18}$$

The (noise-operator based) quantum rms error $\varepsilon_{\mathrm{NO}}(A, \mathbf{M}, \rho)$ for measuring A in ρ by $\mathbf{M}$ is defined as the root mean square of the noise operator, i.e.,

$$\varepsilon_{\mathrm{NO}}(A, \mathbf{M}, \rho) = \left\{\mathrm{Tr}\left[N(A, \mathbf{M})^2(\rho \otimes |\xi\rangle\langle\xi|)\right]\right\}^{1/2}. \tag{2.19}$$

To argue the reliability of the error measure $\varepsilon_{\mathrm{NO}}$ defined above, we consider the following requirements for any reliable error measures ε generalizing the classical root-mean-square error ε_G to quantify the mean error $\varepsilon(A, \mathbf{M}, \rho)$ of the measurement of an observable A in a state ρ described by a measuring process $\mathbf{M}$ [45].

(i) Operational definability. The error measure ε should be definable by the POVM Π of the measuring process $\mathbf{M}$ with the observable A to be measured and the initial state ρ of the measured system $\mathbf{S}$.

(ii) *Correspondence principle.* In the case where $A(0)$ and $M(t_0)$ commute in $\rho \otimes |\xi\rangle\langle\xi|$, the relation

$$\varepsilon(A, \mathbf{M}, \rho) = \varepsilon_G(\mu) \tag{2.20}$$

holds for the joint probability distribution μ of $A(0)$ and $M(t_0)$ in $\rho \otimes |\xi\rangle\langle\xi|$.

(iii) *Soundness.* If $\mathbf{M}$ accurately measures A in ρ, then ε vanishes, i.e., $\varepsilon(A, \mathbf{M}, \rho) = 0$.

(iv) *Completeness.* If ε vanishes, then $\mathbf{M}$ accurately measures A in ρ.

It was shown in [45] that the noise-operator-based quantum rms error $\varepsilon = \varepsilon_{\mathrm{NO}}$ satisfies requirements (i)–(iii), so it is a sound generalization of the classical rms error. However, as pointed out by Busch *et al.* [8], $\varepsilon = \varepsilon_{\mathrm{NO}}$ may not satisfy the completeness requirement (iv) in general. To improve this point, in Ref. [45] a modification of the noise-operator-based quantum rms error $\varepsilon_{\mathrm{NO}}$ was introduced to satisfy all the requirements (i)–(iv) as follows. The locally uniform quantum rms error $\overline{\varepsilon}$ is defined by

$$\overline{\varepsilon}(A, \mathbf{M}, \rho) = \sup_{t \in \mathbb{R}} \varepsilon_{\mathrm{NO}}(A, \mathbf{M}, e^{-itA} \rho e^{itA}). \tag{2.21}$$

Then $\varepsilon = \overline{\varepsilon}$ satisfies all the requirements (i)–(iv) including completeness. In addition to (i)–(iv), the new error measure $\overline{\varepsilon}$ has the following two properties.

(v) *Dominating property.* The error measure $\overline{\varepsilon}$ dominates $\varepsilon_{\mathrm{NO}}$, i.e., $\varepsilon_{\mathrm{NO}}(A, \mathbf{M}, \rho) \leq \overline{\varepsilon}(A, \mathbf{M}, \rho)$.

(vi) *Conservation property for dichotomic measurements.* The error measure $\overline{\varepsilon}$ coincides with $\varepsilon_{\mathrm{NO}}$ for dichotomic measurements, i.e., $\overline{\varepsilon}(A, \mathbf{M}, \rho) = \varepsilon_{\mathrm{NO}}(A, \mathbf{M}, \rho)$ if $A(0)^2 = M(t_0)^2 = \mathbb{1}$.

By property (v) the new error measure $\overline{\varepsilon}$ maintains the previously obtained universally valid EDRs [37, 5, 43]. In this thesis we consider the measurement of a spin component σ_z of a spin-$1/2$ particle using a dichotomic meter observable M, i.e., $M^2 = \mathbb{1}$, so by property (vi) of $\overline{\varepsilon}$ we conclude that the noise-operator-based quantum rms error $\varepsilon_{\mathrm{NO}}$ satisfies all the requirements (i)–(iv) for our measurements under consideration without modifying it to be $\overline{\varepsilon}$.

As shown in Eq. (3.62) in Chapter 3, in our model of the Stern-Gerlach measurement, the Heisenberg observables $A(0)$ and $M(t_0)$ commute, so the error measure satisfying (i) and (ii) is uniquely determined as the (noise-operator-based) quantum rms error.

Busch *et al.* [10] criticized the use of the noise-operator-based quantum rms error, by comparing it with the error measure based on the Wasserstein 2-distance, another error measure defined as the Wasserstein 2-distance between the probability distributions of $A(0)$ and $M(t_0)$. As shown in Ref. [45], the error measure based on the Wasserstein 2-distance or based on any distance between the probability distributions of $A(0)$ and $M(t_0)$ satisfies (i) and (iii) but does not satisfy (ii) or (iv), so the discrepancies between those two measures do not lead to the conclusion that the noise-operator-based quantum rms error is less reliable than the error measured based on the Wasserstein 2-distance or based on any distance between probability distributions of $A(0)$ and $M(t_0)$.

In what follows, where no confusion may occur, we will write $\varepsilon(A) = \varepsilon_{NO}(A)$ for brevity.

2.1.6 Disturbance of observables

We say that the measuring process $\mathbf{M}$ does not disturb the observable B in a state ρ if $B(0)$ and $B(t_0)$ are perfectly correlated in the state $\rho \otimes |\xi\rangle\langle\xi|$ [39, 42, 41], namely, one of the following two equivalent conditions holds: (i) $B(0)$ and $B(t_0)$ commute in $\rho \otimes |\xi\rangle\langle\xi|$ and their joint probability distribution μ satisfies

$$\sum_{b,b':b=b'} \mu(b, b') = 1 \tag{2.22}$$

or (ii) for any $b, b' \in \mathbb{R}$ with $b \neq b'$,

$$\mathrm{Tr}\left[P^{B(t_0)}(b')P^{B(0)}(b)\rho \otimes |\xi\rangle\langle\xi|\right] = 0. \tag{2.23}$$

Note that the left-hand side of Eq. (2.23) is called the weak joint distribution of $B(0)$ and $B(t_0)$ and always exists, possibly taking negative or complex values. The weak joint distribution is operationally accessible by weak measurement of $B(0)$ and post selection for $B(t_0)$ [26, 31]. Thus, our definition of non disturbing measurement is operationally accessible.

2.1.7 Quantum root-mean-square disturbance

For any observable B of the system $\mathbf{S}$, the disturbance operator $D(B, \mathbf{M})$ for the measuring process $\mathbf{M}$ causing the observable B is defined as the change of the observable B during the measurement, i.e.,

$$D(B, \mathbf{M}) = B(t_0) - B(0). \tag{2.24}$$

Similarly to the quantum rms error, the quantum rms disturbance $\eta(B, \mathbf{M}, \rho)$ of B in ρ caused by $\mathbf{M}$ is defined as the rms of the disturbance operator, i.e.,

$$\eta(B, \mathbf{M}) = \left\{\mathrm{Tr}[D(B, \mathbf{M})^2(\rho \otimes |\xi\rangle\langle\xi|)]\right\}^{1/2}. \tag{2.25}$$

The quantum rms disturbance η has properties analogous to the (noise-operator-based) quantum rms error as follows.

(i) Operational definability. The quantum rms disturbance η is definable by the non selective operation T of the measuring process $\mathbf{M}$, the observable B to be disturbed, and the initial state ρ of the measured system $\mathbf{S}$.

(ii) Correspondence principle. In the case where $B(0)$ and $B(t_0)$ commute in $\rho \otimes |\xi\rangle\langle\xi|$, the relation

$$\eta(B, \mathbf{M}, \rho) = \varepsilon_G(\mu) \tag{2.26}$$

holds for the joint probability distribution μ of $B(0)$ and $B(t_0)$ in $\rho \otimes |\xi\rangle\langle\xi|$.

(iii) Soundness. If $\mathbf{M}$ does not disturb B in ρ, then η vanishes.

(iv) Completeness for dichotomic observables. In the case where $B^2 = \mathbb{1}$, if η vanishes, then $\mathbf{M}$ does not disturb B in ρ.

Korzekwa *et al.* [29] criticized the use of the operator-based quantum rms disturbance relying on their definition of non disturbing measurements. They define non disturbing measurements in a system state ρ as measurements satisfying that $B(0)$ and $B(t_0)$ have identical probability distributions for the initial state $\rho \otimes |\xi\rangle\langle\xi|$. They claimed that the operator-based quantum rms disturbance does not satisfy the soundness requirement based on their definition of non disturbing measurements. However, the conflict can be easily reconciled, since their definition of non disturbing measurement is not strong enough, i.e., they call a measurement non disturbing even when the disturbance is operationally detectable. In fact, they supposed that the projective measurement of $A = \sigma_z$ of a spin-1/2 particle in the state $|\sigma_z = +1\rangle$ does not disturb the observable $B = \sigma_x$. However, this measurement really disturbs the observable $B = \sigma_x$. In fact, we have

$$\begin{aligned} &\langle\psi,\xi|P^{B(t_0)}(b')P^{B(0)}(b)|\psi,\xi\rangle \\ &\quad = |\langle\sigma_z = +1|\sigma_x = b'\rangle|^2 |\langle\sigma_z = +1|\sigma_x = b\rangle|^2. \end{aligned}$$

Thus, $B(0)$ and $B(t_0)$ have the same probability distribution, i.e.,

$$\langle\psi,\xi|P^{B(t_0)}(b)|\psi,\xi\rangle = \langle\psi,\xi|P^{B(0)}(b)|\psi,\xi\rangle, \tag{2.27}$$

but the weak joint distribution operationally detects the disturbance on B, i.e.,

$$\langle\psi,\xi|P^{B(t_0)}(-1)P^{B(0)}(+1)|\psi,\xi\rangle = 1/4. \tag{2.28}$$

In this case, we have $\eta(B, \mathbf{M}, \rho) = \sqrt{2} \neq 0$ (see [40]p. S680). However, this does not mean that η does not satisfy the soundness requirement, since $\mathbf{M}$ disturbs B in ρ according to Eq. (2.28). The detail will be discussed elsewhere.

2.1.8 Universally valid error-disturbance relations

In the following, where no confusion may occur, we abbreviate $\varepsilon(A, \mathbf{M}, \rho)$ as $\varepsilon(A)$ and $\eta(B, \mathbf{M}, \rho)$ as $\eta(B)$.

In Ref. [37] Ozawa derived the relation

$$\varepsilon(A)\eta(B)+\varepsilon(A)\sigma(B)+\sigma(A)\eta(B) \geq \frac{1}{2}|\mathrm{Tr}([A,B]\rho)|, \tag{2.29}$$

holding for any pair of observables A and B, state $|\psi\rangle$, and measuring process $\mathbf{M}$. subsequently, Brancirard [5] and Ozawa [43] obtained a stronger EDR given by

$$\begin{aligned} &\varepsilon(A)^2\sigma(B)^2 + \sigma(A)^2\eta(B)^2 \\ &\qquad +2\varepsilon(A)\eta(B)\sqrt{\sigma(A)^2\sigma(B)^2 - D_{AB}^2} \geq D_{AB}^2, \end{aligned} \tag{2.30}$$

where

$$D_{AB} = \frac{1}{2}\mathrm{Tr}(|\sqrt{\rho}[A,B]\sqrt{\rho}|). \tag{2.31}$$

In the case where $A^2 = B^2 = \mathbb{1}$ and $M^2 = \mathbb{1}$, the relation (2.30) can be strengthened as [5, 43]

$$\hat{\varepsilon}(A)^2 + \hat{\eta}(B)^2 + 2\hat{\varepsilon}(A)\hat{\eta}(B)\sqrt{1 - D_{AB}^2} \geq D_{AB}^2, \tag{2.32}$$

where $\hat{\varepsilon}(A) = \epsilon(A)\sqrt{1 - \frac{\epsilon(A)^2}{4}}$ and $\hat{\eta}(B) = \eta(B)\sqrt{1 - \frac{\eta(B)^2}{4}}$. In the case where

$$A = \sigma_z, \quad B = \sigma_x, \quad \langle \sigma_z(0) \rangle_\rho = \langle \sigma_x(0) \rangle_\rho = 0, \tag{2.33}$$

the inequality (2.32) is reduced to the tight relation [5, 43]

$$\left[\varepsilon(\sigma_z)^2 - 2\right]^2 + \left[\eta(\sigma_x)^2 - 2\right]^2 \leq 4, \tag{2.34}$$

as depicted in FIG 2.1.

Lund and Wiseman [31] proposed a measurement model $\mathbf{M}(\theta)$ measuring σ_z of the system $\mathbf{S}$ with another q-bit system as the probe $\mathbf{P}$ prepared in the state $|\xi(\theta)\rangle = \cos\theta|0\rangle + \sin\theta|1\rangle$ with the meter observable $M = \sigma_z$ of the probe $\mathbf{P}$. The measuring interaction is described by the controlled-NOT (CNOT) operation $U_{\text{CNOT}} = |0\rangle\langle 0| \otimes \mathbb{1} + |1\rangle\langle 1| \otimes \sigma_x$. For any state ρ the error $\varepsilon(\sigma_z)$ and the disturbance $\eta(\sigma_x)$ of $\mathbf{M}(\theta)$ satisfy $\varepsilon(\sigma_z) = 2|\sin\theta|$ and $\eta(\sigma_x) = \sqrt{2}|\cos\theta - \sin\theta|$. Thus, they attain the bound

$$[\varepsilon(\sigma_z)^2 - 2]^2 + [\eta(\sigma_x)^2 - 2]^2 = 4 \tag{2.35}$$

for the tight EDR (2.34). Experimental realizations of this model were reported by Rozema et al. [48] and Refs. [1, 56, 27, 47, 53].

In this study, we consider another type of measurement model measuring σ_z, known as Stern-Gerlach measurements, and investigate the admissible region of the error $\varepsilon(\sigma_z)$ for σ_z measurement and the disturbance $\varepsilon(\sigma_x)$ on σ_x, obtained from Gaussian orbital states.

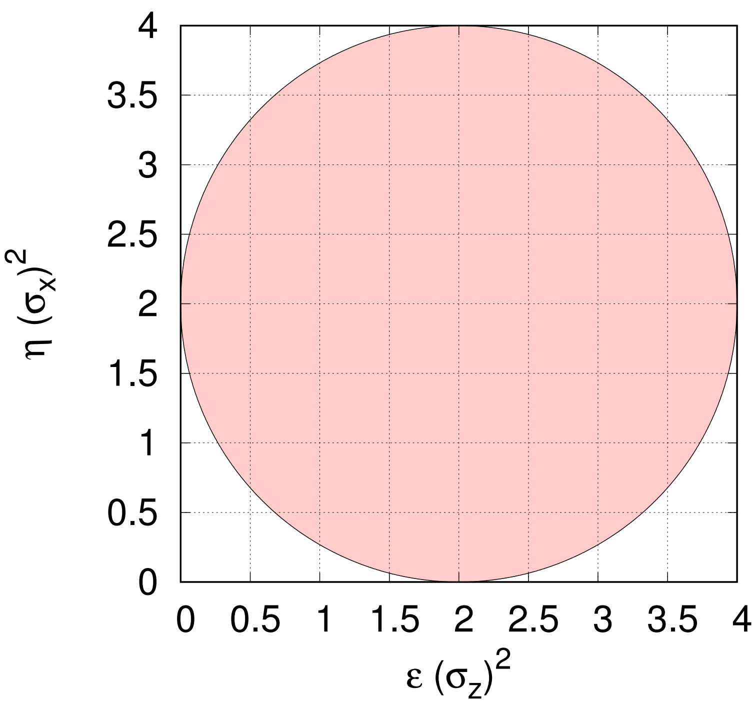

Figure 2.1: $\varepsilon(\sigma_z)^2$ - $\eta(\sigma_x)^2$ plot of tight EDR (2.34) for spin measurements in the state satisfying Eq. (2.33).

2.2 Gaussian wave packets

In this section, we review the relations between Gaussian states and inequalities. Let Z and P be the canonical position and momentum observables, respectively, of a one-dimensional quantum system. These observables satisfy the usual canonical commutation relation $[Z, P] = i\hbar$. Here we consider only a vector state denoted by ψ. However, some of the results in this section can easily be generalized to mixed states.

2.2.1 Schrödinger inequality

For the variances of the position and momentum, the following inequality holds [49]:

$$\mathrm{Var}_\psi(Z)\mathrm{Var}_\psi(P) \geq \frac{(\langle\{Z, P\}\rangle_\psi - 2\langle Z\rangle_\psi\langle P\rangle_\psi)^2 + \hbar^2}{4}. \tag{2.36}$$

The inequality (2.36) is known as the Schrödinger inequality. The proof proceeds as follows. First, we consider the case $\langle Z\rangle_\psi = \langle P\rangle_\psi = 0$. Then we have

$$\mathrm{Im}\,\langle Z\psi, P\psi\rangle = \frac{1}{2i}\langle[Z, P]\rangle_\psi = \hbar/2, \tag{2.37}$$

$$\mathrm{Re}\,\langle Z\psi, P\psi\rangle = \frac{1}{2}\langle\{Z, P\}\rangle_\psi. \tag{2.38}$$

Consequently, we have

$$|\langle Z\psi, P\psi\rangle|^2 = \frac{(\langle\{Z, P\}\rangle_\psi)^2 + \hbar^2}{4}. \tag{2.39}$$

On the other hand, according to the Cauchy-Schwarz inequality,

$$|\langle Z\psi, P\psi\rangle|^2 \leq \langle Z^2\rangle_\psi\langle P^2\rangle_\psi = \mathrm{Var}_\psi(Z)\mathrm{Var}_\psi(P). \tag{2.40}$$

Hence, the Schrödinger inequality (2.36) holds if $\langle Z\rangle_\psi = \langle P\rangle_\psi = 0$ holds. We can obtain the proof for the general case by substituting Z and P into $Z - \langle Z\rangle_\psi$ and $P - \langle P\rangle_\psi$, respectively. This concludes the proof.

The equation in this inequality holds if and only if

$$(Z - \langle Z\rangle_\psi)\,\psi = c\,(P - \langle P\rangle_\psi)\,\psi \tag{2.41}$$

for some complex number c. From the condition above, we obtain the differential equation for the wave function as

$$\frac{d}{dz}\psi(z) = -2k\left[z - \left(\langle Z\rangle_\psi + \frac{i}{2\hbar k}\langle P\rangle_\psi\right)\right]\psi(z), \tag{2.42}$$

where k is a complex number. Therefore, we have

$$\psi(z) = A\exp\left(-k\left[z - \left(\langle Z\rangle_\psi + \frac{i}{2\hbar k}\langle P\rangle_\psi\right)\right]^2\right), \tag{2.43}$$

where A is a constant. Since the wave function should be normalizable, the constant k must satisfy $\mathrm{Re}\,k > 0$.

2.2.2 Kennard inequality

The inequality, which is known as the Kennard inequality [28]

$$\mathrm{Var}_\psi(Z)\mathrm{Var}_\psi(P) \geq \hbar^2/4, \tag{2.44}$$

can be derived from the Schrödinger inequality (2.36). The equality in Eq. (2.44) holds if and only if $2i\hbar k\left(Z - \langle Z\rangle_\psi\right)\psi = \left(P - \langle P\rangle_\psi\right)\psi$ for some positive real number k. A wave function ψ satisfies the equality in the Kennard inequality (2.44) if and only if ψ has the form

$$\psi(z) = A\exp\left(-k\left[z - \left(\langle Z\rangle_\psi + \frac{i}{2\hbar k}\langle P\rangle_\psi\right)\right]^2\right) \tag{2.45}$$

for some positive real number k. This wave function has the same form as that of Eq. (2.43) except for the condition of the constant k, i.e., the constant k in Eq. (2.43) is a complex number with a positive real part whereas the constant k in Eq. (2.45) is a positive real number. The state in Eq. (2.45) is known as the minimum-uncertainty state.

2.2.3 Squeezed state

For any two complex numbers μ and ν satisfying $|\mu|^2 - |\nu|^2 = 1$, the squeezed operator $c_{\mu,\nu}$ is defined as

$$c_{\mu,\nu} := \mu a + \nu a^\dagger, \tag{2.46}$$

where a and $a^\dagger$ are the annihilation and creation operators, respectively.

$$a := \sqrt{\frac{m\omega}{2\hbar}}Z + i\sqrt{\frac{1}{2\hbar m\omega}}P. \tag{2.47}$$

Here$m$ and ω are the mass and angular frequency of the corresponding harmonic oscillator, respectively. A coherent state [21] is defined as the eigenstate of the annihilation operator a in Eq. (2.47). A squeezed state [57] is defined as the eigenstate of squeezed operator $c_{\mu,\nu}$,

$$c_{\mu,\nu}\psi = \lambda\psi. \tag{2.48}$$

By this definition, the wave function of every squeezed state satisfies the differential equation

$$\left[(\mu+\nu)\sqrt{\frac{m\omega}{2\hbar}}z + (\mu-\nu)\sqrt{\frac{\hbar}{2m\omega}}\frac{d}{dz}\right]\psi(z) = \lambda\psi(z). \tag{2.49}$$

The solution of this differential equation is

$$\psi(z) := A\exp\left[-\frac{m\omega}{2\hbar}\frac{\mu+\nu}{\mu-\nu}\left(z - \sqrt{\frac{2\hbar}{m\omega}}\frac{\lambda}{\mu-\nu}\right)^2\right]. \tag{2.50}$$

Hence, the equality in the Schrödinger inequality (2.36) holds for squeezed states.

Next let us consider the relation between these parameters and the mean values of the position and momentum. By comparing the two formulas, (2.43) and (2.50), we have

$$\langle Z\rangle_\psi + \frac{i}{m\omega}\frac{\mu-\nu}{\mu+\nu}\langle P\rangle_\psi = \sqrt{\frac{2\hbar}{m\omega}}\frac{\lambda}{\mu-\nu}. \tag{2.51}$$

Taking the imaginary part, we have

$$\langle P\rangle_\psi = \sqrt{2\hbar m\omega}|\mu+\nu|^2 \mathrm{Im}\left(\frac{\lambda}{\mu-\nu}\right), \tag{2.52}$$

$$\langle Z\rangle_\psi = \sqrt{\frac{2\hbar}{m\omega}} \mathrm{Re}\left(\frac{(\mu+\nu)(\mu^*-\nu^*)}{\mu-\nu}\lambda\right). \tag{2.53}$$

Next, let us calculate the variances of the position and momentum and the correlation $\langle\{Z,P\}\rangle_\psi$. Setting $\tilde{z} = z - \langle Z\rangle_\psi$, we have

$$\begin{aligned}
&\mathrm{Var}(Z) \\
&= |A|^2 \int_{-\infty}^{\infty} \tilde{z}^2 \exp\Bigg(-\frac{m\omega}{\hbar} \\
&\times \mathrm{Re}\left[\frac{\mu-\nu}{\mu+\nu}\left(\frac{\mu+\nu}{\mu-\nu}\tilde{z} + \frac{i}{m\omega}\langle P\rangle_\psi\right)^2\right]\Bigg) d\tilde{z} \\
&= \frac{\hbar}{2m\omega}|\mu-\nu|^2.
\end{aligned} \tag{2.54}$$

To calculate the variance of the momentum, it is convenient to obtain the Fourier transform of the wave function $\tilde{\psi}(\tilde{z}) := \psi(\tilde{z} + \langle Z\rangle_\psi)$,

$$\begin{aligned}
\hat{\psi}(p) &= \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} \tilde{\psi}(\tilde{z}) \exp(ip\tilde{z}/\hbar) d\tilde{z} \\
&= \hat{A} \exp\left[-\frac{1}{2\hbar m\omega}\frac{\mu-\nu}{\mu+\nu}(p - \langle P\rangle_\psi)^2\right],
\end{aligned} \tag{2.55}$$

where $\hat{A}$ is the normalization constant. Consequently, we have

$$\begin{aligned}
&\mathrm{Var}(P) = \langle (P - \langle P\rangle_\psi)^2\rangle_\psi \\
&= |\hat{A}|^2 \int_{-\infty}^{\infty} \tilde{p}^2 \exp\left[-\frac{1}{\hbar m\omega}\mathrm{Re}\left(\frac{\mu-\nu}{\mu+\nu}\right)\tilde{p}^2\right] d\tilde{p} \\
&= \frac{\hbar m\omega}{2}|\mu+\nu|^2.
\end{aligned} \tag{2.56}$$

Finally, we calculate the correlation term

$$\begin{aligned}
&\langle\{Z - \langle Z\rangle_\psi, P - \langle P\rangle_\psi\}\rangle_\psi \\
&= \langle\{Z - \langle Z\rangle_\psi, P\}\rangle_\psi \\
&= 2\mathrm{Re}\,\langle \tilde{Z}\psi, P\psi\rangle \\
&= 2\mathrm{Re}\Bigg(|A|^2 im\omega \\
&\times \int_{-\infty}^{\infty} \frac{\mu+\nu}{\mu-\nu}\tilde{z}^2 \exp\left[-\frac{m\omega}{\hbar}\mathrm{Re}\left(\frac{\mu+\nu}{\mu-\nu}\tilde{z}^2\right)\right] d\tilde{z}\Bigg) \\
&= 2\hbar \mathrm{Im}\,(\mu^*\nu).
\end{aligned} \tag{2.57}$$

The coherent state is defined as the eigenstate of the annihilation operator. Using the results of the calculation above, the corresponding wave function is

$$\psi(z) = A\exp\left[-\frac{m\omega}{2\hbar}\left(z - \sqrt{\frac{2\hbar}{m\omega}}\lambda\right)^2\right], \tag{2.58}$$

where λ is the corresponding eigenvalue of the annihilation operator. Thus, every coherent state satisfies the equation in the Schrödinger inequality (2.36) and the Kennard inequality (2.44).

Since $\dfrac{\mu+\nu}{\mu-\nu}$ moves all over the right half plane of the complex plane as μ and ν move all over the complex plane satisfying $|\mu|^2 - |\mu|^2 = 1$, the union of all squeezed states and coherent states coincides with the states that satisfy the Schrödinger inequality (2.36), namely, $\mathcal{G}$.

2.2.4 Contractive state

The contractive state was introduced by Yuen [58] as a squeezed state whose correlation term is negative. This state contracts during some period of time if it evolves freely. To see this, let us calculate the variance of the position in the Heisenberg picture. The position operator $Z(t)$ at time t in the Heisenberg picture is

$$\begin{aligned} Z(t) &= \exp\left[-\frac{t}{2i\hbar m}P(t)^2\right] Z(0)\exp\left[\frac{t}{2i\hbar m}P(t)^2\right] \\ &= Z(0) + \frac{t}{m}P(0). \end{aligned} \tag{2.59}$$

Hence, we have

$$\begin{aligned} &\mathrm{Var}_\psi[Z(t)] \\ &= \left\langle\left(Z(0) + \frac{t}{m}P(0) - \langle Z(0) + \frac{t}{m}P(0)\rangle_\psi\right)^2\right\rangle_\psi \\ &= \frac{t^2}{m^2}\mathrm{Var}_\psi[P(0)] + \mathrm{Var}_\psi[Z(0)] \\ &+ \frac{t}{m}\left\langle\{Z(0) - \langle Z(0)\rangle_\psi, P(0) - \langle P(0)\rangle_\psi\}\right\rangle_\psi. \end{aligned} \tag{2.60}$$

Therefore, if the state is a contractive state, the variance of the position contracts until the time

$$t = -\frac{m\langle\{Z(0) - \langle Z(0)\rangle_\psi, P(0) - \langle P(0)\rangle_\psi\}\rangle_\psi}{2\langle P(0)^2\rangle_\psi}. \tag{2.61}$$

2.2.5 Covariance matrix formalism

Recently, the covariance matrix was used to characterize Gaussian states [55]. For a single-mode Gaussian state,

$$\psi(z) = A\exp\left(-k\left[z - \left(\langle Z\rangle_\psi + \frac{i}{2\hbar k}\langle P\rangle_\psi\right)\right]^2\right), \tag{2.62}$$

the covariance matrix V is defined as

$$V = \begin{pmatrix} \mathrm{Var}_{\psi}(Z) & \mathrm{Cor}_{\psi}(Z,P) \\ \mathrm{Cor}_{\psi}(Z,P) & \mathrm{Var}_{\psi}(P) \end{pmatrix} = \begin{pmatrix} [4\mathrm{Re}(k)]^{-1} & -\frac{\hbar \mathrm{Im}(k)}{\mathrm{Re}(k)} \\ -\frac{\hbar \mathrm{Im}(k)}{\mathrm{Re}(k)} & \frac{\hbar^2 |k|^2}{\mathrm{Re}(k)} \end{pmatrix}. \tag{2.63}$$

Here, we used the abbreviation,

$$\mathrm{Cor}_{\psi}(Z,P) = \langle \{Z - \langle Z \rangle_{\psi}, P - \langle P \rangle_{\psi}\} \rangle_{\psi}. \tag{2.64}$$

2.2.6 Summary

We have discussed the relation between the inequalities and the subclasses of Gaussian states whose wave functions are of the form

$$\psi(z) = A \exp\left(-k \left[z - (\langle Z \rangle_{\psi} + \frac{i}{2\hbar k} \langle P \rangle_{\psi}) \right]^2 \right) \tag{2.65}$$

and obtained the relations shown in Table. 2.1. Figure. 2.2 represents the inclusion relation between the subsets of the set of Gaussian wave packets.

Table 2.1: Classification of Gaussian states in terms of the parameter k.

k	Type of state	Inequality whose equality holds
$\mathrm{Re}\, k > 0$	Squeezed	Schrödinger
$\mathrm{Re}\, k > 0$ and $\mathrm{Im}\, k > 0$	Contractive	Schrödinger
$\mathrm{Re}\, k > 0$ and $\mathrm{Im}\, k = 0$	Minimum uncertainty	Kennard
$k = \hbar$	Coherent	Kennard

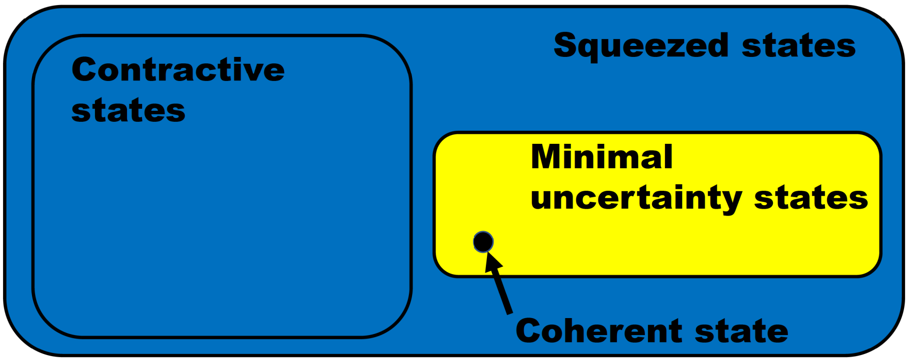

Figure 2.2: Inclusion relation of the subsets of wave functions. A wave function is in the yellow region if and only if the equality in the Kennard inequality holds. A wave function is in the blue or yellow region if and only if the equality in the Schrödinger inequality holds.

2.3 Time evolution of Gaussian wave packets

In this section we discuss the time evolution of the probability density of a Gaussian wave packet during free evolution. The wave function under consideration is the Gaussian wave packet derived in Sec. 2.2,

$$\psi(z) := A\exp\left(-kz^2\right), \tag{2.66}$$

where k is a complex number with a positive real part. For simplicity, we consider only the case in which the mean values of the position and momentum are zero. Applying the Fourier transform $\mathfrak{F}$ successively, we obtain

$$\begin{aligned}
&\exp\left(\frac{t}{2i\hbar m}P^2\right)\psi(z)\\
&=\mathfrak{F}^{-1}\exp\left(\frac{t}{2i\hbar m}p^2\right)\hat{A}\exp\left(-\frac{p^2}{4k\hbar^2}\right)\\
&=\frac{\hat{A}}{\sqrt{2\pi\hbar}}\int_{-\infty}^{\infty}\exp\left[\left(\frac{t}{2i\hbar m}-\frac{1}{k\hbar^2}\right)p^2-ipz/\hbar\right]dp\\
&=N\exp\left(-\frac{z^2}{k^{-1}-\frac{2\hbar t}{im}}\right),
\end{aligned}\tag{2.67}$$

where N is the normalization constant. Thus, the probability density $\Pr(z)$ at time t has the form

$$\Pr(z)=|N|^2\exp\left(-rz^2\right) \tag{2.68}$$

for some positive real number r, that is, we have again obtained a Gaussian distribution. Since the variance of the Gaussian distribution is

$$\left\langle Z(t)^2\right\rangle_\psi=\left\langle\left[Z(0)+\frac{t}{m}P(0)\right]^2\right\rangle_\psi, \tag{2.69}$$

we have

$$\Pr(z)=|N|^2\exp\left(-\frac{z^2}{2\langle\left(Z(0)+\frac{t}{m}P(0)\right)^2\rangle_\psi}\right). \tag{2.70}$$

2.4 Relationship between the Heisenberg picture and the Schrödinger picture

Let us consider the relation between the Heisenberg picture and the Schrödinger picture. Consider the time evolution of quantum system S described by $\mathcal{H}$. Let A be an observable of system S and state ψ. Denote by $E(A,\psi,t)$ the expectation value of the outcome of the measurement of observable A at time t, provided system S is in state ψ at time 0. In the Schrödinger picture, state $\psi(t)$ evolves in time t as a solution of the Schrödinger equation by the time evolution operator $U(t)$ as $\psi(t)=U(t)\psi$ with the initial condition $U(0)=\mathbb{1}$, so $E(A,\psi,t)=\langle\psi(t),A\psi(t)\rangle$ holds. The unitary operator $U^{\mathrm{S}}(t_2,t_1)$ describing the time evolution from time $t=t_1$ to $t=t_2$ ($t_1\le t_2$) in the Schrödinger picture is defined by

$$U^{\mathrm{S}}(t_2,t_1)=U(t_2)U^\dagger(t_1). \tag{2.71}$$

Then we have

$$U^{\mathrm{S}}(t_2, t_1)\psi(t_1) = \psi(t_2), \tag{2.72}$$

$$U^{\mathrm{S}}(t_3, t_2)U^{\mathrm{S}}(t_2, t_1) = U^{\mathrm{S}}(t_3, t_1). \tag{2.73}$$

In the Heisenberg picture, observable $A(t)$ evolves in time t by the time evolution operator $U(t)$ as $A(t) = U(t)^\dagger AU(t)$, so $E(A, \psi, t) = \langle\psi, A(t)\psi\rangle$ holds. The unitary operator $U^{\mathrm{H}}(t_2, t_1)$ describing the time evolution from time $t = t_1$ to $t = t_2$ $(t_1 \le t_2)$ in the Heisenberg picture is defined by

$$U^{\mathrm{H}}(t_2, t_1) = U^\dagger(t_1)U(t_2). \tag{2.74}$$

Then we have

$$U^{\mathrm{H}}(t_2, t_1)^\dagger A(t_1)U^{\mathrm{H}}(t_2, t_1) = A(t_2), \tag{2.75}$$

$$\alpha^{\mathrm{H}}(t_3, t_2)\alpha^{\mathrm{H}}(t_2, t_1) = \alpha^{\mathrm{H}}(t_3, t_1), \tag{2.76}$$

where

$$\alpha^{\mathrm{H}}(t_2, t_1)A = U^{\mathrm{H}}(t_2, t_1)^\dagger AU^{\mathrm{H}}(t_2, t_1). \tag{2.77}$$

We have the following relations between the Schrödinger picture and the Heisenberg picture:

$$U(t) = U^{\mathrm{S}}(t, 0) = U^{\mathrm{H}}(t, 0). \tag{2.78}$$

$$U^{\mathrm{H}}(t_2, t_1) = U(t_1)^\dagger U^{\mathrm{S}}(t_2, t_1)U(t_1). \tag{2.79}$$

Let $f(A_1, \ldots, A_n, t, s)$ be a function of observables $A_1, \ldots, A_n$ and real numbers t and s. If

$$U^{\mathrm{S}}(t_2, t_1) = f(A_1, \ldots, A_n, t_1, t_2), \tag{2.80}$$

then

$$U^{\mathrm{H}}(t_2, t_1) = f(A_1(t_1), \ldots, A_n(t_1), t_1, t_2). \tag{2.81}$$

Chapter 3

Error-disturbance relation in Stern-Gerlach measurements

3.1 Stern-Gerlach Measurements

Let us consider the setting of a Stern-Gerlach measurement as depicted in Figure 3.1. A particle with spin $1/2$ goes through the inhomogeneous magnetic field and then evolves freely. The inhomogeneous magnetic field is approximated to be $\mathbf{B} \simeq \begin{pmatrix} 0, & 0, & B_0 + B_1 z \end{pmatrix}$. The state of the spin degree of freedom $\mathbf{S}$ is supposed to be an arbitrary mixed state satisfying $\langle \sigma_z \rangle_\rho = \langle \sigma_x \rangle_\rho = 0$, e.g., $\rho = |\sigma_y = \pm 1\rangle\langle \sigma_y = \pm 1|$.

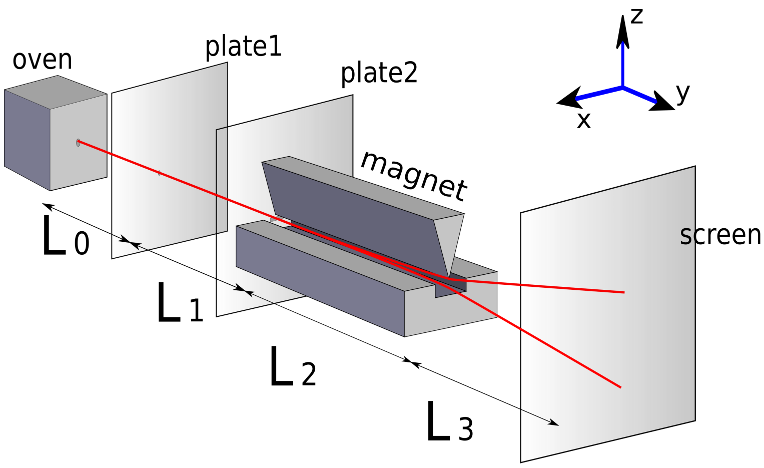

Figure 3.1: Illustration of the experimental setup for a Stern-Gerlach measurement. The relations between the length and the time interval are $L_2 = v_y \Delta t$, $L_3 = v_y \tau$.

The measuring process of this Stern-Gerlach measurement is given as follows. The probe system $\mathbf{P}$ is the z-component of the orbital degree of freedom of the particle. We assume that the initial state of the probe system $\mathbf{P}$ is a general Gaussian state given by $\xi_\lambda(z) = A \exp(-\lambda z^2)$, where $\lambda \in \mathbb{C}$ and $\mathrm{Re}\,\lambda > 0$. The Hamiltonian of the composite system $\mathbf{S}+\mathbf{P}$ is given by

$$H(t)=\begin{cases}\mu\sigma_z\otimes(B_0+B_1Z)+\dfrac{1}{2m}\mathbb{1}\otimes P^2 & (0\le t\le\Delta t),\\ \dfrac{1}{2m}\mathbb{1}\otimes P^2 & (\Delta t\le t\le\Delta t+\tau),\end{cases} \tag{3.1}$$

where μ denotes the magnetic moment of the particle and m denotes the mass of the particle. The meter observable is $M=f(Z)$, where

$$f(z)=\begin{cases}-1 & (\text{if } z\ge 0),\\ +1 & (\text{if } z<0).\end{cases}$$

where μ denotes the magnetic moment of the particle and m denotes the mass of the particle. By solving the Schrödinger equation, we obtain the time evolution operator $U(t)$ of $\mathbf{S}+\mathbf{P}$ for $0\le t\le\Delta t+\tau$ by

$$U(t)=\begin{cases}\exp\left\{\dfrac{t}{i\hbar}\left[\mu\sigma_z\otimes(B_0+B_1Z)+\dfrac{1}{2m}\mathbb{1}\otimes P^2\right]\right\}\\ \qquad\qquad\qquad\qquad\qquad\qquad(0\le t\le\Delta t),\\ \exp\left[\dfrac{t-\Delta t}{2i\hbar m}\mathbb{1}\otimes P^2\right]\\ \times\exp\left\{\dfrac{\Delta t}{i\hbar}\left[\mu\sigma_z\otimes(B_0+B_1Z)+\dfrac{1}{2m}\mathbb{1}\otimes P^2\right]\right\}\\ \qquad\qquad\qquad\qquad\qquad\qquad(\Delta t\le t\le\Delta t+\tau).\end{cases} \tag{3.2}$$

To describe the time evolution of the composite system $\mathbf{S}+\mathbf{P}$ in the Heisenberg picture, we introduce Heisenberg operators for $0\le t\le\Delta t+\tau$ as

$$Z(0)=\mathbb{1}\otimes Z,\quad Z(t)=U(t)^\dagger Z(0)U(t), \tag{3.3}$$

$$P(0)=\mathbb{1}\otimes P,\quad P(t)=U(t)^\dagger P(0)U(t), \tag{3.4}$$

$$\sigma_j(0)=\sigma_j\otimes\mathbb{1},\quad \sigma_j(t)=U(t)^\dagger\sigma_j(0)U(t), \tag{3.5}$$

where $j=x,y,z$. By solving Heisenberg equations of motion for $Z(t)$, $P(t)$, $\sigma_x(t)$, $\sigma_y(t)$, and $\sigma_z(t)$, we have

$$Z(\Delta t+\tau)=Z(0)+\frac{\Delta t+\tau}{m}P(0) -\frac{\mu B_1\Delta t}{m}\left(\tau+\frac{\Delta t}{2}\right)\sigma_z(0), \tag{3.6}$$

$$P(\Delta t+\tau)=P(0)-\mu B_1\Delta t\sigma_z(0), \tag{3.7}$$

$$\sigma_x(\Delta t+\tau)=\begin{pmatrix}0 & \exp[iS(\Delta t)]\\ \exp[-iS(\Delta t)] & 0\end{pmatrix}, \tag{3.8}$$

$$\sigma_y(\Delta t+\tau)=\begin{pmatrix}0 & -i\exp[iS(\Delta t)]\\ i\exp[-iS(\Delta t)] & 0\end{pmatrix}, \tag{3.9}$$

$$\sigma_z(\Delta t+\tau)=\sigma_z(0), \tag{3.10}$$

where

$$S(\Delta t) \quad = \quad \frac{2\mu\Delta t}{\hbar}\left[B_0 + B_1\left(Z(0) + \frac{\Delta t}{2m}P(0)\right)\right]. \tag{3.11}$$

To obtain them, we proceed as follows. To consider the time evolution from time $t = \Delta t$ to time $\Delta t + \tau$, suppose $\Delta t \leq t \leq \Delta t + \tau$. By the Heisenberg equation of motion, the position operator $Z(t)$ satisfies

$$\frac{d}{dt}Z(t) = \frac{1}{i\hbar}[Z(t), \frac{1}{2m}P(t)^2] = \frac{1}{m}P(t). \tag{3.12}$$

Thus, we have

$$Z(t) = Z(\Delta t) + \frac{1}{m}\int_{\Delta t}^{t} P(t')dt'. \tag{3.13}$$

In contrast, $P(t)$ does not change since $[P(t), H(t)] = 0$. Consequently, we have

$$Z(t) \quad = \quad Z(\Delta t) + \frac{t - \Delta t}{m}P(\Delta t), \tag{3.14}$$

$$P(t) \quad = \quad P(\Delta t). \tag{3.15}$$

Since $\sigma_z(t)$ and $\sigma_x(t)$ commute with $H(t)$, we have

$$\sigma_z(t) = \sigma_z(\Delta t), \quad \sigma_x(t) = \sigma_x(\Delta t). \tag{3.16}$$

To describe the observables at time $t = \Delta t$ in terms of the observables at time $t = 0$, suppose that $0 \leq t \leq \Delta t$. With the Heisenberg equations of motion, we obtain

$$\frac{d}{dt}Z(t) = \frac{1}{i\hbar}[Z(t), H(t)] = \frac{1}{m}P(t) \tag{3.17}$$

and

$$Z(\Delta t) = Z(0) + \frac{1}{m}\int_{0}^{\Delta t} P(t)dt. \tag{3.18}$$

On the other hand, we have

$$\frac{d}{dt}P(t) = \frac{1}{i\hbar}[P(t), H(t)] = -\mu B_1\sigma_z(t). \tag{3.19}$$

Now $\sigma_z(t)$ commutes with Hamiltonian $H(t)$. Hence, we have

$$\sigma_z(t) = \sigma_z(0). \tag{3.20}$$

Consequently, we have

$$P(t) \quad = \quad P(0) - \mu B_1 t\sigma_z(0), \tag{3.21}$$

$$Z(t) \quad = \quad Z(0) + \frac{t}{m}P(0) - \frac{\mu B_1 t^2}{2m}\sigma_z(0). \tag{3.22}$$

Therefore, we have

$$\begin{aligned}Z(\Delta t+\tau) &= Z(0)+\frac{\Delta t+\tau}{m}P(0)\\ &\qquad -\frac{\mu B_1\Delta t}{m}\left(\tau+\frac{\Delta t}{2}\right)\sigma_z(0), \quad (3.23)\\ P(\Delta t+\tau) &= P(0)-\mu B_1\Delta t\sigma_z(0), \quad (3.24)\\ \sigma_z(\Delta t+\tau) &= \sigma_z(0). \quad (3.25)\end{aligned}$$

Next we calculate the x and y components of the spin of the particle at time $t=\Delta t+\tau$. Since the Hamiltonian $H(t)$ from time $t=\Delta t$ to time $\Delta t+\tau$ commutes with $\sigma_x(t)$ and $\sigma_y(t)$, we have

$$\begin{aligned}\sigma_x(t) &= \sigma_x(\Delta t), \quad (3.26)\\ \sigma_y(t) &= \sigma_y(\Delta t) \quad (3.27)\end{aligned}$$

if $\Delta t \le t \le \Delta t+\tau$, and it suffices to calculate $\sigma_x(\Delta t)$ and $\sigma_y(\Delta t)$.

Suppose $0\le t\le \Delta t$. By the Heisenberg equations of motion we have

$$\begin{aligned}\frac{d}{dt}\sigma_x(t) &= \frac{1}{i\hbar}\left[\sigma_x(t),H(t)\right]\\ &= \frac{1}{i\hbar}\left[\sigma_x(t),\frac{P(t)^2}{2m}+\mu[B_0+B_1Z(t)]\sigma_z(t)\right]\\ &= \frac{\mu}{i\hbar}[B_0+B_1Z(t)]\left[-2i\sigma_y(t)\right]\\ &= -\frac{2\mu}{\hbar}[B_0+B_1Z(t)]\sigma_y(t). \quad (3.28)\end{aligned}$$

Similarly, we have

$$\begin{aligned}\frac{d}{dt}\sigma_y(t) &= \frac{1}{i\hbar}\left[\sigma_y(t),H(t)\right]\\ &= \frac{1}{i\hbar}\left[\sigma_y(t),\frac{P(t)^2}{2m}+\mu[B_0+B_1Z(t)]\sigma_z(t)\right]\\ &= \frac{\mu}{i\hbar}[B_0+B_1Z(t)]\left[2i\sigma_x(t)\right]\\ &= \frac{2\mu}{\hbar}[B_0+B_1Z(t)]\sigma_x(t). \quad (3.29)\end{aligned}$$

Now let us introduce σ_+ and σ_- by

$$\begin{aligned}\sigma_+(t) &= \frac{1}{\sqrt{2}}[\sigma_x(t)+i\sigma_y(t)], \quad (3.30)\\ \sigma_-(t) &= \frac{1}{\sqrt{2}}[\sigma_x(t)-i\sigma_y(t)]. \quad (3.31)\end{aligned}$$

From Eqs. (3.28) and (3.29), we have

$$\frac{d}{dt}\sigma_\pm(t)=\pm\frac{2\mu i}{\hbar}\left[B_0+B_1\left(U^\dagger(t)Z(0)U(t)\right)\right]\sigma_\pm(t). \quad (3.32)$$

Let

$$\gamma_{\pm}(t) = U(t)\sigma_{\pm}(t) = \exp\left[\frac{H(0)}{i\hbar}t\right]\sigma_{\pm}(t). \tag{3.33}$$

The left-hand side (LHS) and right-hand side (RHS) of Eq. (3.32) satisfy

$$\begin{aligned}\text{LHS} &= \frac{d}{dt}U(-t)\gamma_{\pm}(t)\\ &= -\frac{H(0)}{i\hbar}U(-t)\gamma_{\pm}(t) + U(-t)\frac{d}{dt}\gamma_{\pm}(t),\end{aligned} \tag{3.34}$$

$$\begin{aligned}\text{RHS} &= \pm\frac{2\mu i}{\hbar}U^{\dagger}(t)[B_0 + B_1 Z(0)]U(t)U^{\dagger}(t)\gamma_{\pm}(t)\\ &= \pm\frac{2\mu i}{\hbar}U(-t)[B_0 + B_1 Z(0)]\gamma_{\pm}(t).\end{aligned} \tag{3.35}$$

Hence, we have

$$\frac{d}{dt}\gamma_{\pm}(t) = \left(\frac{H(0)}{i\hbar} \pm \frac{2\mu i}{\hbar}\left[B_0 + B_1 Z(0)\right]\right)\gamma_{\pm}(t). \tag{3.36}$$

The solution of the above differential equation is given by

$$\gamma_{\pm}(t) = \exp\left(\frac{it}{\hbar}\left\{-H(0) \pm 2\mu\left[B_0 + B_1 Z(0)\right]\right\}\right)\gamma_{\pm}(0). \tag{3.37}$$

Since $\gamma_{\pm}(0) = \sigma_{\pm}(0)$, we have

$$\sigma_{\pm}(t) = \exp\left(\frac{it}{\hbar}H(0)\right)\exp\left(\frac{it}{\hbar}\left\{-H(0) \pm 2\mu\left[B_0 + B_1 Z(0)\right]\right\}\right)\sigma_{\pm}(0). \tag{3.38}$$

Using the Baker-Campbell-Hausdorff formula [2] we have

$$\begin{aligned}\exp(A)\exp(B) = \exp\Big\{&(A+B) + \frac{1}{2}[A,B]\\ &+ \frac{1}{12}\left([[A,B],B] - [[A,B],A]\right) + \cdots\Big\}.\end{aligned} \tag{3.39}$$

Hence, for

$$A = \frac{it}{\hbar}H(0), \tag{3.40}$$

$$B = \frac{it}{\hbar}\{-H(0) \pm 2\mu\left[B_0 + B_1 Z(0)\right]\}, \tag{3.41}$$

we have

$$\begin{aligned}[A,B] &= \left[\frac{it}{\hbar}H(0), \frac{it}{\hbar}\{-H(0) \pm 2\mu\left[B_0 + B_1 Z(0)\right]\}\right]\\ &= -\frac{t^2}{\hbar^2}\left[\frac{1}{2m}P(0)^2, \pm 2\mu[B_0 + B_1 Z(0)]\right]\\ &= \pm\frac{2i\mu B_1 t^2}{m\hbar}P(0),\end{aligned} \tag{3.42}$$

$$
\begin{aligned}
[[A,B],A] &= \left[\pm\frac{2i\mu B_1 t^2}{m\hbar}P(0), \frac{it}{\hbar}H(0)\right] \\
&= \mp\frac{2\mu B_1 t^3}{m\hbar^2}\left[P(0), \mu[B_0 + B_1 Z(0)]\sigma_z(0)\right] \\
&= \pm\frac{2i\mu^2 B_1^2 t^3}{m\hbar}\sigma_z(0),
\end{aligned} \tag{3.43}
$$

$$
\begin{aligned}
&[[A,B],B] \\
&= \left[\pm\frac{2i\mu B_1 t^2}{m\hbar}P(0), \frac{it}{\hbar}\{-H(0) \pm 2\mu\,[B_0 + B_1 Z(0)]\}\right] \\
&= \mp\frac{2i\mu^2 B_1^2 t^3}{m\hbar}\sigma_z(0) \\
&\quad \mp\frac{2\mu B_1 t^3}{m\hbar^2}\left[P(0), \pm 2\mu[B_0 + B_1 Z(0)]\sigma_z(0)\right] \\
&= \frac{2i\mu^2 B_1^2 t^3}{m\hbar}[2 \mp \sigma_z(0)].
\end{aligned} \tag{3.44}
$$

The commutators of the higher orders , denoted by an ellipsis in Eq. (3.39), are 0 since the third commutators $[[A,B],A]$ and $[[A,B],B]$ commute with A and B, respectively.

Let

$$
R(t) = \frac{\mu^2 B_1^2 t^3}{3m\hbar}, \tag{3.45}
$$

$$
S(t) = \frac{2\mu t}{\hbar}\left[B_0 + B_1\left(Z + \frac{t}{2m}P\right)\right]. \tag{3.46}
$$

We have

$$
\sigma_\pm(t) = \exp i\{[R(t) \pm S(t)]\mathbb{1} \mp R(t)\sigma_z(0)\}\sigma_\pm(0). \tag{3.47}
$$

Since

$$
\sigma_+(0) = \frac{1}{\sqrt{2}}[\sigma_z(0) + i\sigma_y(0)] = \begin{pmatrix} 0 & \sqrt{2} \\ 0 & 0 \end{pmatrix}, \tag{3.48}
$$

$$
\sigma_-(0) = \frac{1}{\sqrt{2}}[\sigma_z(0) - i\sigma_y(0)] = \begin{pmatrix} 0 & 0 \\ \sqrt{2} & 0 \end{pmatrix}, \tag{3.49}
$$

we have

$$
\begin{aligned}
&\sigma_+(t) \\
&= \begin{pmatrix} \exp[iS(t)] & 0 \\ 0 & \exp i[S(t) + 2R(t)] \end{pmatrix}\begin{pmatrix} 0 & \sqrt{2} \\ 0 & 0 \end{pmatrix} \\
&= \exp[iS(t)]\sigma_+(0),
\end{aligned} \tag{3.50}
$$

$$
\begin{aligned}
&\sigma_-(t) \\
&= \begin{pmatrix} \exp\{i[-S(t) + 2R(t)]\} & 0 \\ 0 & \exp[-iS(t)] \end{pmatrix}\begin{pmatrix} 0 & 0 \\ \sqrt{2} & 0 \end{pmatrix} \\
&= \exp[-iS(t)]\sigma_-(0).
\end{aligned} \tag{3.51}
$$

Therefore, $\sigma_x(t)$ and $\sigma_y(t)$ from time $t = 0$ to time $t = \Delta t$ are

$$\begin{aligned}\sigma_x(t) &= \frac{1}{\sqrt{2}}[\sigma_+(t) + \sigma_-(t)] \\ &= \begin{pmatrix} 0 & \exp[iS(t)] \\ \exp[-iS(t)] & 0 \end{pmatrix},\end{aligned} \tag{3.52}$$

$$\begin{aligned}\sigma_y(t) &= -\frac{i}{\sqrt{2}}[\sigma_+(t) - \sigma_-(t)] \\ &= \begin{pmatrix} 0 & -i\exp[iS(t)] \\ i\exp[-iS(t)] & 0 \end{pmatrix}.\end{aligned} \tag{3.53}$$

3.2 Error

Let us consider the quantum rms error of a Stern-Gerlach measurement $\mathbf{M}$ of the z component $\sigma_z(0)$ of the spin at time 0 using the meter observable

$$M(\Delta t + \tau) = f(Z(\Delta t + \tau)), \tag{3.54}$$

introduced in Sec. 3.1. The noise operator N of this measurement is given by

$$N = M(\Delta t + \tau) - \sigma_z(0). \tag{3.55}$$

The initial state ρ of the spin $\mathbf{S}$ is supposed to be an arbitrary state with the matrix

$$\rho = \frac{1}{2}(\mathbb{1} + n_x\sigma_x + n_y\sigma_y + n_z\sigma_z) \tag{3.56}$$

where $n_x, n_y, n_z \in \mathbb{R}$ and $n_x^2 + n_y^2 + n_z^2 \leq 1$, so that the initial state of the composite system $\mathbf{S} + \mathbf{P}$ is given by $\rho \otimes |\xi\rangle\langle\xi|$, where $|\xi\rangle$ is a fixed but arbitrary wave function describing the initial state of the orbital degree of freedom $\mathbf{P}$. Then the error, namely, the quantum rms error, of this measurement of σ_z is given by

$$\varepsilon(\sigma_z) = \sqrt{\langle N^2\rangle_{\rho\otimes|\xi\rangle\langle\xi|}}, \tag{3.57}$$

where we abbreviate $\mathrm{Tr}(A\rho)$ as $\langle A\rangle_\rho$ for observable A and density operator ρ. We will give an explicit formula for $\varepsilon(\sigma_z)$, which eventually shows that the error depends only on the parameter n_z in Eq. (3.56).

Let

$$U_t = \exp\left[\frac{t}{2i\hbar m}P^2\right], \tag{3.58}$$

$$\tilde{U}_t = \mathbb{1}_{\mathbf{S}} \otimes U_t, \tag{3.59}$$

$$g_0 = \frac{\mu B_1 \Delta t}{m}\left(\tau + \frac{\Delta t}{2}\right). \tag{3.60}$$

From Eq. (3.6) we have

$$\begin{aligned} &Z(\Delta t+\tau)\\ &\quad=\tilde{U}^{\dagger}_{\Delta t+\tau}\begin{pmatrix} Z-g_0 & 0\\ 0 & Z+g_0\end{pmatrix}\tilde{U}_{\Delta t+\tau}. \end{aligned} \tag{3.61}$$

Thus, we have

$$\begin{aligned} N&=f(Z(\Delta t+\tau))-\sigma_z(0)\\ &=2\tilde{U}^{\dagger}_{\Delta t+\tau}\begin{pmatrix} -\chi_+(Z-g_0) & 0\\ 0 & \chi_-(Z+g_0)\end{pmatrix}\tilde{U}_{\Delta t+\tau}, \end{aligned} \tag{3.62}$$

where

$$\chi_+(z)=\begin{cases}1 & \text{if } z\geq 0,\\ 0 & \text{otherwise},\end{cases} \tag{3.63}$$

$$\chi_-(z)=1-\chi_+(z), \tag{3.64}$$

$$f(z)=1-2\chi_+(z). \tag{3.65}$$

It follows that

$$N^2=4\tilde{U}^{\dagger}_{\Delta t+\tau}\begin{pmatrix} \chi_+(Z-g_0) & 0\\ 0 & \chi_-(Z+g_0)\end{pmatrix}\tilde{U}_{\Delta t+\tau}. \tag{3.66}$$

Therefore, we have

$$\begin{aligned} \varepsilon(\sigma_z)^2&=\langle N^2\rangle_{\rho\otimes|\xi\rangle\langle\xi|}\\ &=\left\langle\xi|\mathrm{Tr}_{\mathbf{S}}[N^2\rho]|\xi\right\rangle\\ &=2(1+n_z)\langle\xi|U^{\dagger}_{\Delta t+\tau}\chi_+(Z-g_0)U^{\dagger}_{\Delta t+\tau}|\xi\rangle\\ &\quad+2(1-n_z)\langle\xi|U^{\dagger}_{\Delta t+\tau}\chi_-(Z-g_0)U^{\dagger}_{\Delta t+\tau}|\xi\rangle. \end{aligned} \tag{3.67}$$

Consequently, we have

$$\begin{aligned} \varepsilon(\sigma_z)^2=2(1+n_z)&\int_{g_0}^{\infty}|U_{\Delta t+\tau}\xi(z)|^2dz\\ &+2(1-n_z)\int_{-\infty}^{-g_0}|U_{\Delta t+\tau}\xi(z)|^2dz. \end{aligned} \tag{3.68}$$

3.3 Disturbance

Let us consider the quantum rms disturbance, $\eta(\sigma_x)$, for the x-component of the spin in Stern-Gerlach measurements. The disturbance operator, σ_x, is given by

$$D=\sigma_x(\Delta t+\tau)-\sigma_x(0). \tag{3.69}$$

From Eq. (3.8) we have

$$D=\begin{pmatrix} 0 & \exp[iS(\Delta t)]-1\\ \exp[-iS(\Delta t)]-1 & 0\end{pmatrix}. \tag{3.70}$$

Consequently, we have

$$D^2 = \mathbb{1} \otimes [2 - 2\cos S(\Delta t)]\,, \tag{3.71}$$

and thus

$$\begin{aligned}&\eta(\sigma_x)^2\\ &= 2-2\left\langle \cos\left\{\frac{2\mu\Delta t}{\hbar}\left[B_0 + B_1\left(Z + \frac{\Delta t}{2m}P\right)\right]\right\}\right\rangle_\xi.\end{aligned} \tag{3.72}$$

3.4 Error and disturbance for Gaussian states

Let us consider the error and disturbance in Stern-Gerlach measurements under the condition that the orbital state of the particle is in the family $\mathcal{G}$ of Gaussian states given by

$$\mathcal{G} = \left\{ \xi_\lambda \in L^2(\mathbb{R}) \;\middle|\; \begin{array}{l} \xi_\lambda(z) = A\exp\left(-\lambda z^2\right) \\ \int_{-\infty}^{\infty} |\xi_\lambda(z)|^2 dz = 1 \\ \lambda \in \mathbb{C}, \mathrm{Re}\,(\lambda) > 0 \end{array} \right\}. \tag{3.73}$$

This family of states consists of all Gaussian pure states [50], whose mean values of the position and momentum are both 0. For simplicity, it is assumed that the spin state of the particle is in the eigenstate of the spin component σ_y. It is easy to minimize the error of the measurement with respect to the mean values of the position and momentum. In particular, $\mathcal{G}$ is the family of optimal states for the measurement among the Gaussian pure states if the spin state of the particle is the eigenstate of σ_y. We remark that the equality in the Schrödinger inequality [see Eq. (2.36)] holds for any state ξ in $\mathcal{G}$, i.e.,

$$\langle Z^2\rangle_\xi \langle P^2\rangle_\xi - \frac{1}{4}\langle\{Z,P\}\rangle_\xi^2 = \frac{\hbar^2}{4}. \tag{3.74}$$

Here we use the abbreviation $\langle A\rangle_\xi = \langle\xi|A|\xi\rangle$. The converse also holds, that is, any state ξ satisfying $\langle P\rangle_\xi = \langle Z\rangle_\xi = 0$ and Eq. (3.74) belongs to $\mathcal{G}$.

Let us consider the range of the error and disturbance of Stern-Gerlach measurements. Let

$$V(\psi, t) = \left\langle \left(Z + \frac{t}{m}P\right)^2 \right\rangle_\psi \tag{3.75}$$

for any orbital state ψ. For the disturbance $\eta(\sigma_x)$, from Eq. (3.72) we have

$$\begin{aligned}&\eta(\sigma_x)^2\\ &= 2 - 2\left\langle \cos\left[\frac{2\mu\Delta t}{\hbar}\left(B_0 + B_1 Z\right)\right]\right\rangle_{U_{\Delta t/2}\xi_\lambda}\\ &= 2 - \frac{2}{\sqrt{2\pi V(\xi_\lambda, \Delta t/2)}}\\ &\times \int_{-\infty}^{\infty} \exp\left(-\frac{z^2}{2V(\xi_\lambda,\Delta t/2)}\right)\cos\left[\frac{2\mu\Delta t}{\hbar}\left(B_0 + B_1 z\right)\right] dz\\ &= 2 - 2\exp\left(-\frac{2\mu^2 B_1^2\Delta t^2}{\hbar^2}V(\xi_\lambda, \Delta t/2)\right)\cos\frac{2\mu\Delta t B_0}{\hbar}.\end{aligned} \tag{3.76}$$

From the above formula, the disturbance is determined by $V(\xi_\lambda, \Delta t/2)$ and the parameters of the magnet if the orbital state is in $\mathcal{G}$. Now, for a fixed constant v let us find the error for state ξ_λ in $\mathcal{G}$ and time interval Δt satisfying $V(\xi_\lambda, \Delta t/2) = v$. In the following, we fix the time interval Δt.

From Eq. (3.68) we have

$$\begin{aligned}\varepsilon(\sigma_z)^2 &= 4\int_{g_0}^{\infty} |U_{\Delta t+\tau}\xi_\lambda(z)|^2\, dz \\ &= \frac{4}{\sqrt{\pi}}\int_{g_0/\sqrt{2V(\xi_\lambda,\Delta t+\tau)}}^{\infty} \exp(-w^2)dw. \end{aligned} \tag{3.77}$$

Here we use the relation $n_z = 0$, which is obtained from the assumption that the mean value of the z component of the spin of the particle is 0. Eq. (3.77) shows that the error is minimized by maximizing the lower limit of the integration $g_0/\sqrt{2V(\xi_\lambda, \Delta t+\tau)}$. First, we fix the state ξ_λ and focus on the time interval τ. Let $W_{\xi_\lambda}(\tau) = g_0/\sqrt{2V(\xi_\lambda, \Delta t+\tau)}$.

Putting $\sigma(t) = V(\xi_\lambda, t)^{1/2}$, the parameter $\sigma(\Delta t/2)$ appears in the formula of the disturbance, because the disturbance of the spin along the x-axis is caused by this uncontrollable precession around z-axis. On the other hand, the error in the Stern-Gerlach setup comes from the non-zero dispersion $\sigma(\Delta t+\tau)$ of the particle position on the screen. By the uncertainty relation

$$\sigma\left(\frac{\Delta t}{2}\right)\sigma(\Delta t+\tau) \geq \frac{\hbar}{2m}\left(\frac{\Delta t}{2}+\tau\right), \tag{3.78}$$

the smaller the dispersion $\sigma(\Delta t+\tau)$ of the particle position on the screen, the greater the dispersion $\sigma(\Delta t/2)$ of the particle position in the Stern-Gerlach magnet. This is why $\sigma(\Delta t+\tau)$ appears in the formula of the error, and this yields a tradeoff between $\varepsilon(\sigma_z)$ and $\eta(\sigma_x)$.

From now on, we suppose $B_1 \leq 0$ and $b^2 < \dfrac{8\hbar^2}{m^2}$. If

$$m\langle\{Z,P\}\rangle_{\xi_\lambda} + \langle P^2\rangle_{\xi_\lambda}\Delta t < 0 \tag{3.79}$$

holds, then $W_{\xi_\lambda}(\tau)$ assumes the maximum value

$$W_{\xi_\lambda}(\tau_0) = \frac{\sqrt{2V(\xi_\lambda, \Delta t/2)}\mu B_1\Delta t}{\hbar} \tag{3.80}$$

at

$$\begin{aligned}\tau =& \tau_0 \\ =& -\frac{4m^2\langle Z^2\rangle_{\xi_\lambda} + 3m\langle\{Z,P\}\rangle_{\xi_\lambda}\Delta t + 2\langle P^2\rangle_{\xi_\lambda}\Delta t^2}{2\left(m\langle\{Z,P\}\rangle_{\xi_\lambda} + \langle P^2\rangle_{\xi_\lambda}\Delta t\right)}.\end{aligned} \tag{3.81}$$

In fact, setting

$$W_{\xi_\lambda}(\tau) = \alpha\left(\tau+\frac{\Delta t}{2}\right)\left[a+b(\Delta t+\tau)+c(\Delta t+\tau)^2\right]^{-1/2}, \tag{3.82}$$

where $\alpha = \frac{\mu B_1 \Delta t}{\sqrt{2}m}$, $a = \langle Z^2 \rangle$, $b = \frac{\langle \{Z, P\} \rangle}{m}$, and $c = \frac{\langle P^2 \rangle}{m^2}$, the derivative of function $W_{\xi_\lambda}(\tau)$ is

$$\begin{aligned}
&\frac{d}{d\tau} W_\lambda(\tau) \\
&= \frac{\alpha}{4} \left[a + b(\Delta t + \tau) + c(\Delta t + \tau)^2 \right]^{-3/2} \\
&\times \left[2(b + c\Delta t)(\Delta t + \tau) + 4a + b\Delta t \right].
\end{aligned} \tag{3.83}$$

Hence, $W_{\xi_\lambda}(t)$ assumes the maximum value at $\tau = \tau_0 = -\frac{4a + 3b\Delta t + 2c\Delta t^2}{2(b + c\Delta t)} \geq 0$ if the following conditions hold: (i) $W'(0) > 0$ and (ii) $2b + 2c\Delta t < 0$. Condition (i) holds automatically. In fact, (i) is equivalent to the condition

$$4a + 3b\Delta t + 2c\Delta t^2 \geq 0. \tag{3.84}$$

Now let us consider the function

$$f(t) = 4a + 3bt + 2ct^2. \tag{3.85}$$

This function assumes the minimum value at $t = -\frac{3b}{4c}$,

$$\begin{aligned}
f(t) &\geq f\left(-\frac{3b}{4c} \right) \\
&= \frac{32ac - 9b^2}{8c} \\
&= \frac{9}{8c}(4ac - b^2) - \frac{4ac}{8c} \\
&= \frac{9\hbar^2}{8cm^2} - \frac{\hbar^2}{8cm^2} - \frac{b^2}{8c} \\
&> \frac{8\hbar^2}{8cm^2} - \frac{8\hbar^2}{8cm^2} \\
&= 0.
\end{aligned} \tag{3.86}$$

Therefore, condition (i) is satisfied automatically. Here we use the Schrödinger inequality (2.36). Hence, if condition (ii) holds, the function $W_\lambda(\tau)$ assumes the maximum value at $\tau = \tau_0 \geq 0$. The maximum value of $W_{\xi_\lambda}(\tau)$ for $\tau \geq 0$ is

$$\begin{aligned}
W_{\xi_\lambda}(\tau_0) &= -\alpha \frac{4a + 2b\Delta t + c\Delta t^2}{2(b + c\Delta t)} \\
&\times \left[a + b(\Delta t + \tau_0) + c(\Delta t + \tau_0)^2 \right]^{-1/2} \\
&= \alpha \left(4a + 2b\Delta t + c\Delta t^2 \right)^{1/2} (4ac - b^2)^{-1/2} \\
&= \frac{2\alpha m}{\hbar} \left[a + b\frac{\Delta t}{2} + c\left(\frac{\Delta t}{2} \right)^2 \right]^{1/2} \\
&= \frac{\sqrt{2}\mu B_1 \Delta t}{\hbar} \left\langle \left(Z + \frac{\Delta t}{2m} P \right)^2 \right\rangle_{\xi_\lambda}^{1/2}.
\end{aligned} \tag{3.87}$$

If condition (ii) does not hold, the function $W_{\xi_\lambda}(\tau)$ increases monotonically and we have

$$\sup_{\tau\geq 0} W_{\xi_\lambda}(\tau) = \lim_{\tau\to\infty} W_{\xi_\lambda}(\tau) = \frac{\mu B_1 \Delta t}{\sqrt{2\langle P^2\rangle_{\xi_\lambda}}}. \tag{3.88}$$

Now let us consider the maximization of $W_{\xi_\lambda}(\tau)$ with respect to the state ξ_λ. For any pair of orbital states ψ and ϕ in $\mathcal{G}$ satisfying $V(\psi, \Delta t/2) = v$ and $V(\phi, \Delta t/2) = v$, respectively, if ψ satisfies the condition (3.79), then

$$W_\psi(\tau_0) \geq \lim_{\tau\to\infty} W_\phi(\tau) \tag{3.89}$$

holds, since $W_\psi(\tau_0)/\lim_{\tau\to\infty} W_\phi(\tau) \geq 1$ by the Kennard inequality (1.2). Therefore, we obtain the supremum of $W_{\xi_\lambda}(\tau)$ with respect to the state ξ_λ and time interval τ as

$$\sup_{\mathrm{Re}(\lambda)>0,\tau\geq 0} W_{\xi_\lambda}(\tau) = \frac{\sqrt{2v}\mu B_1 \Delta t}{\hbar}. \tag{3.90}$$

Although the above argument is for finding the range of the error and disturbance that Stern-Gerlach measurements can assume, it contains one more important assertion. That is, the calculation suggests that the error of Stern-Gerlach measurements is minimized by placing the screen at a finite distance from the magnet under the condition represented by (3.79), in contrast to the conventional assumption that the error is minimized by placing the screen at infinity. If a state in $\mathcal{G}$ satisfies the condition (3.79), then the correlation term [58] $\left\langle\left\{Z - \langle Z\rangle_{\xi_\lambda}, P - \langle P\rangle_{\xi_\lambda}\right\}\right\rangle_{\xi_\lambda}$ is negative, and this leads to a narrowing of the standard deviation of the position of the particle during the free evolution (see Sec. 2.2.4.) Such a class of states was introduced by Yuen [58] and they are known as contractive states.

Let us return to the problem of finding the range of values of the error and disturbance that Stern-Gerlach measurements can assume. Now setting $W_0 = \sqrt{2v}\mu B_1 \Delta t/\hbar$, the disturbance and the infimum of the error under the condition that $V(\lambda, \Delta t/2) = v$ for fixed Δt and v are

$$\eta(\sigma_x)^2 = 2 - 2\exp\left(-W_0^2\right)\cos\frac{2\mu\Delta t B_0}{\hbar}, \tag{3.91}$$

$$\inf_{\lambda,\tau} \varepsilon(\sigma_z)^2 = \frac{4}{\sqrt{\pi}}\int_{W_0}^{\infty} \exp(-w^2)dw, \tag{3.92}$$

respectively. By varying the parameter of the magnet B_0, we obtain the range of the disturbance as

$$2 - 2\exp\left(-W_0^2\right) \leq \eta(\sigma_x)^2 \leq 2 + 2\exp\left(-W_0^2\right). \tag{3.93}$$

We obtain the range of the disturbance and the infimum of the error of Stern-Gerlach measurements for each constant v. By varying B_1, we obtain the range of the error and disturbance as the inequalities

$$\left|\frac{\eta(\sigma_x)^2 - 2}{2}\right| \leq \exp\left\{-\left[\mathrm{erf}^{-1}\left(\frac{\varepsilon(\sigma_z)^2 - 2}{2}\right)\right]^2\right\}, \tag{3.94}$$

$$0 \leq \varepsilon(\sigma_z)^2 \leq 2, \tag{3.95}$$

where erf^{-1} represents the inverse of the error function $\mathrm{erf}(x) = (2/\sqrt{\pi}) \int_0^x \exp(-s^2) ds$. The square of the error varies from 0 to 2 since W_0 is positive.

We now remove the constraint $B_1 \leq 0$. For $B_1 \geq 0$, similarly to the above discussion, we have

$$\left| \frac{\eta(\sigma_x)^2 - 2}{2} \right| \leq \exp\left\{ - \left[\mathrm{erf}^{-1}\left(\frac{\varepsilon(\sigma_z)^2 - 2}{2} \right) \right]^2 \right\}, \tag{3.96}$$

$$2 \leq \varepsilon(\sigma_z)^2 \leq 4. \tag{3.97}$$

Therefore, we have

$$\left| \frac{\eta(\sigma_x)^2 - 2}{2} \right| \leq \exp\left\{ - \left[\mathrm{erf}^{-1}\left(\frac{\varepsilon(\sigma_z)^2 - 2}{2} \right) \right]^2 \right\}. \tag{3.98}$$

the plot of this region is shown in Fig. 3.2.

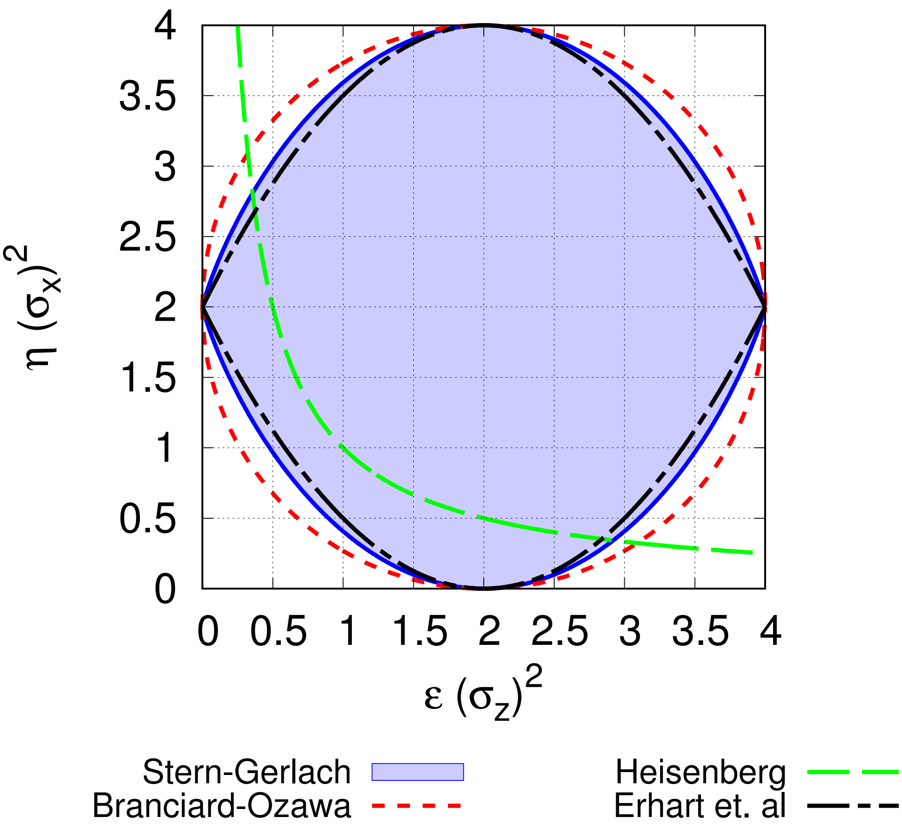

Figure 3.2: Range of error and disturbance for Stern-Gerlach measurements. Range of error and disturbance for Stern-Gerlach measurements. The blue region is the region (3.98) that Stern-Gerlach measurements can achieve. The red dotted line is the boundary of the Branciard-Ozawa tight EDR (2.34). The green dashed line is the boundary of Heisenberg's EDR (1.7). The black dash-dotted line is the theoretical boundary (1.8) of the EDR of the experiment conducted by Erhart and co-workers [16, 54]. The error-disturbance region of Stern-Gerlach measurements is close to the theoretical optimum given by the Branciard-Ozawa tight EDR (2.34) and actually violates Heisenberg's EDR (1.7) in a broad range of experimental parameters.

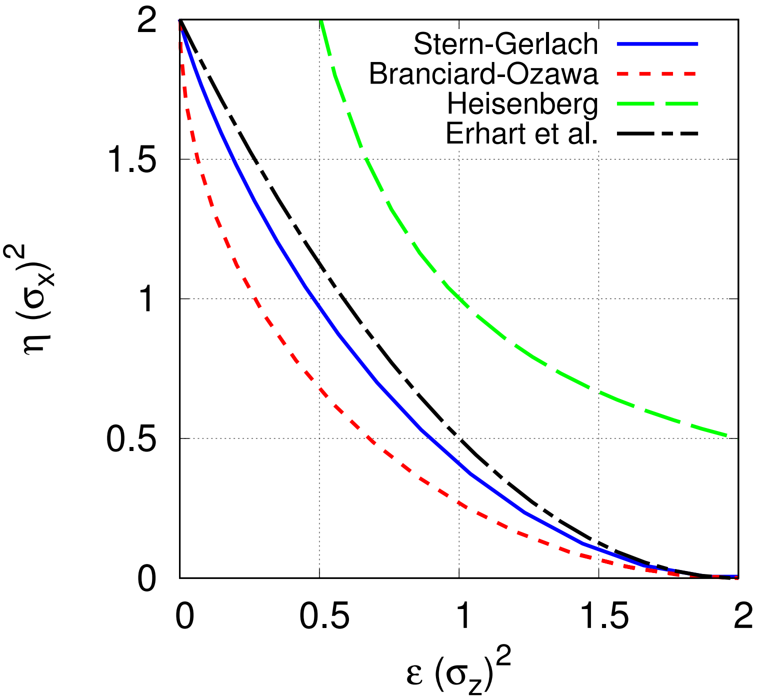

Figure 3.3: Enlarged plot for the part $[0, 2] \times [0, 2]$ of Fig. 3.2 .

For comparison, the figure shows the plot of the boundary of the Branciard-Ozawa tight EDR (2.34) for general spin measurement. From this plot, we conclude that the range of the error and disturbance for Stern-Gerlach measurements considered in this thesis is close to the theoretical optimal given by the Branciard-Ozawa tight EDR (2.34). Here the range of the error and disturbance for Stern-Gerlach measurements is also compared with Heisenberg's EDR (1.7) (green line) and the EDR (1.8) for the neutron experiment [16, 54] (black line). We conclude that Stern-Gerlach measurements actually violate Heisenberg's EDR (1.7) in a broad range of experimental parameters.

Roughly speaking, the parameter v represents the spread of the wave packet of the particle in the Stern-Gerlach magnet. The reason why v appears in the formula of the disturbance is that the particle in the Stern-Gerlach magnet is exposed to the inhomogeneous magnetic field and its spin is processed in an uncontrollable way. This uncontrollable precession occurs because the position of the particle is uncertain while the magnetic field is inhomogeneous and hence depends on the position. The disturbance of the spin along the x axis is caused by this uncontrollable precession around the z axis. This is why v appears in the formula of the disturbance. On the other hand, the error in our Stern-Gerlach setup comes from the non zero dispersion of the z component of the particle position when the particle has reached the screen. The smaller the dispersion of the particle position when the particle has reached the screen, the greater the dispersion of the z component of the particle position in the Stern-Gerlach magnet. This is why v appears in the formula of the error.

3.5 Comparison with "Aspects of nonideal Stern-Gerlach experiment and testable ramifications"

Home *et al.* [23] discussed the same error of Stern-Gerlach measurements as we do for similar conditions. We consider in what sense their paper is related to ours and we compare its results with ours. They derived the wave function of a particle in the Stern-Gerlach apparatus under the following conditions.

(i) The magnetic field is oriented along the z axis everywhere and the gradient of the z component of the magnetic field is non zero only in the z direction.

(ii) The initial orbital state is a Gaussian state whose mean values of the position and momentum, and the correlation term of the particle in the wave function are all zero.

(iii) Unlike Bohm's discussion [3], the kinetic energy of the particle in the magnetic field is not neglected.

Based on their argument, they discussed the distinguishability of the value of the measured observable by observing the probe system directly in Stern-Gerlach measurements. To consider this problem, they introduced the two indices,

$$I := \left| \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \psi_+^*(\mathbf{x}, \tau) \psi_-(\mathbf{x}, \tau) d\mathbf{x} \right|, \tag{3.99}$$

$$E(t) := \int_{-\infty}^{0} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\psi_+(\mathbf{x}, t)|^2 \, dxdydz, \tag{3.100}$$

where $\psi_\pm$ are the wave functions of the particle in the Schrödinger picture whose spin z components are $\pm 1/2$, respectively. The origin of time is taken to be the moment when the particle enters the Stern-Gerlach magnet. In addition, τ is the time at which the particle emerges from the Stern-Gerlach magnet (τ corresponds to Δt in our notation) and t is any time after emerging from the Stern-Gerlach magnet (t corresponds to $\Delta t + \tau$ in our notation). Namely, they adopted the inner product I of the two wave functions with different spin directions, and the probability $E(t)$ of finding the particle with the spin z components of $+1/2$ and $-1/2$ within the lower and upper half planes, respectively, at time t. They concluded that I always vanishes whenever $E(t)$ vanishes, but that $E(t)$ does not necessarily vanish even when I vanishes.

We discuss the relation between their paper and ours. The relation between the quantities $E(t)$ and $\varepsilon(\sigma_z)$ is

$$\varepsilon(\sigma_z)^2 = 4E(t). \tag{3.101}$$

Although this relation is model dependent, it bridges the two approaches and will enforce a theoretical background for our definition of a sound and complete quantum generalization of the classical root-mean-square error [45].

We compare their research with ours as follows.

(i) Their setup and approximation are the same as ours and they used the same Hamiltonian as in our research.

(ii) In both papers, the orbital state of the particle is assumed to be the pure state where the mean values of its position and momentum are zero. We assume that the correlation term of a Gaussian pure state is not necessarily zero, whereas they assumed that the orbital state is a Gaussian pure state with no correlation.

(iii) We evaluate the tradeoff between the error and disturbance, whereas they compared the error with the inner product I of the emerging wave functions expressing formal distinguishability. In addition, we obtain the range of error and disturbance under the condition that the orbital state is a Gaussian pure state whose correlation term is not necessarily zero.

Chapter 4

Violation of Heisenberg's error-disrurbance relation in Stern-Gerlach measurements

4.1 Original Stern-Gerlach measurement

Here, we estimate the error and disturbance of the original Stern-Gerlach experiment conducted by Stern and Gerlach [18, 19, 20] by our theoretical model. We summarize the set up of their experiment (cf. Figure 3.1). A beam of silver atoms emerging from a small hole of a lid of an oven heated to $1500\,[\mathrm{K}]$ was collimated by two plates made of platinum. The atoms passed a pinhole with an area of $3 \times 10^{-3}[\mathrm{mm}]$ (or $d_1 = 6.2 \times 10^{-2}\,[\mathrm{mm}]$ in diameter) in the first plate P_1 and then passed the slit $d_2 = 3.0$ to $4.0 \times 10^{-2}\,[\mathrm{mm}]$ in width in the second plate P_2. The slit was parallel to the x-axis. These plates were arranged perpendicular to the orbit of the atoms and the distance between them was $L_1 = 3.3\,[\mathrm{cm}]$. An $L_2 = 3.5\,[\mathrm{cm}]$ long knife edged magnetic pole was arranged parallel to the orbit of atoms just after the plate P_2. The z-component of the gradient of the magnetic field around the orbit of atoms was $B_1 = -1.35 \times 10^3\,[\mathrm{T} \cdot \mathrm{m}^{-1}]$. A glass plate was arranged immediately after the magnetic pole, in which the atoms are deposited. These conditions of the experiment is summarized in Table 4.1.

After the 8 hours of the operation of the system and developing, they obtained a lip-shaped pattern. The maximum width of the opening of the lip shaped pattern was $1.1 \times 10^{-1}\,[\mathrm{mm}]$. The distance between the centers of the two arc-shaped pattern was $2.0 \times 10^{-1}\,[\mathrm{mm}]$. The velocity distribution of atoms in the oven is assumed to be the Maxwell distribution. Thus, the atoms emerging from the small hole of the lid of the oven are estimated to have the well-known distribution of flux [52]:

$$f_{\text{flux}}(v) = \text{Const.} \times v^3 \exp\left(-\frac{mv^2}{2k_B T}\right). \tag{4.1}$$

The root-mean-square v_y of the y-component of the velocity of atoms is given by [52]

$$v_y = \sqrt{\frac{4k_B T}{m}}. \tag{4.2}$$

Let us estimate the z-component $|\xi_\lambda\rangle$ of the orbital state of an atom in the beam just before entering the magnetic field. We assume the orbital state arriving at plate 1 to be

Table 4.1: The data for the experiment conducted by Gerlach and Stern [18, 19, 20] in 1922.

Experimental Parameters	Values	Related Variables
Temperature T of Oven	$1500\,[\mathrm{K}]$	Δt, τ
Gradient B_1 of Magnetic Field	$-1.35 \times 10^3\,[\mathrm{T/m}]$	B_1
L_1	$3.3 \times 10^{-2}\,[\mathrm{m}]$	ξ
L_2	$3.5 \times 10^{-2}\,[\mathrm{m}]$	Δt
L_3	$0\,[\mathrm{m}]$	τ
Diameter d_1 of Hole of Plate1	$6.2 \times 10^{-5}\,[\mathrm{m}]$	ξ
Width d_2 of Slit of Plate2	$4.0 \times 10^{-5}\,[\mathrm{m}]$	ξ

$\xi_a(z) = (2a/\pi)^{1/4}\exp(-az^2)$ with $a > 0$. We model the operations of the collimator and the slit as approximate momentum-position successive measurements by the canonical D_p-approximate momentum measurement and the canonical D_z-approximate position measurement introduced in [35, Eq. (75)], so that for the outcomes $(P, Z) = (0, 0)$ the posteriori (output) state $|\xi_\lambda\rangle$ for the prior (input) state $|\xi_a\rangle$ is given by

$$|\xi_\lambda\rangle \propto \exp\left(-\frac{Z^2}{4D_z^2}\right)\exp\left(-\frac{P^2}{4D_p^2}\right)|\xi_a\rangle, \tag{4.3}$$

where $\propto$ stands for the equality up to a constant factor. The parameters D_p and D_z will later be determined relative to the structure of the collimator and the slit. Then, we have

$$\xi_\lambda(z) \propto \exp\left\{-\left[\left(\frac{1}{a} + \frac{\hbar^2}{D_p^2}\right)^{-1} + \frac{1}{4D_z^2}\right]z^2\right\}. \tag{4.4}$$

We naturally assume $\sigma(P)_{\xi_a} \gg D_p$, so that we have

$$\frac{1}{a} = 4\sigma(Z)^2_{\xi_a} = \frac{\hbar^2}{\sigma(P)^2_{\xi_a}} \ll \frac{\hbar^2}{D_p^2} \tag{4.5}$$

and we have

$$\xi_\lambda(z) \propto \exp\left[-\left(\frac{D_p^2}{\hbar^2} + \frac{1}{4D_z^2}\right)z^2\right] \tag{4.6}$$

up to arbitrary order.

As depicted in Figure 4.1 the parameters D_p and D_z are estimated by taking into account the half width δP of the possible classical momentum after passing through the collimator (with plates 1 and 2) and the half width δZ of the possible classical position after passing through the slit (on plate 2) as

$$D_p \sim \delta P = \frac{d_1 + d_2}{2L_1} m v_y, \tag{4.7}$$

$$D_z \sim \delta Z = \frac{d_2}{2}. \tag{4.8}$$

To make unambiguous estimates, we suppose that

$$0.75\,\delta P \le D_p \le 1.25\,\delta P, \tag{4.9}$$

$$0.75\,\delta Z \le D_z \le 1.25\,\delta Z. \tag{4.10}$$

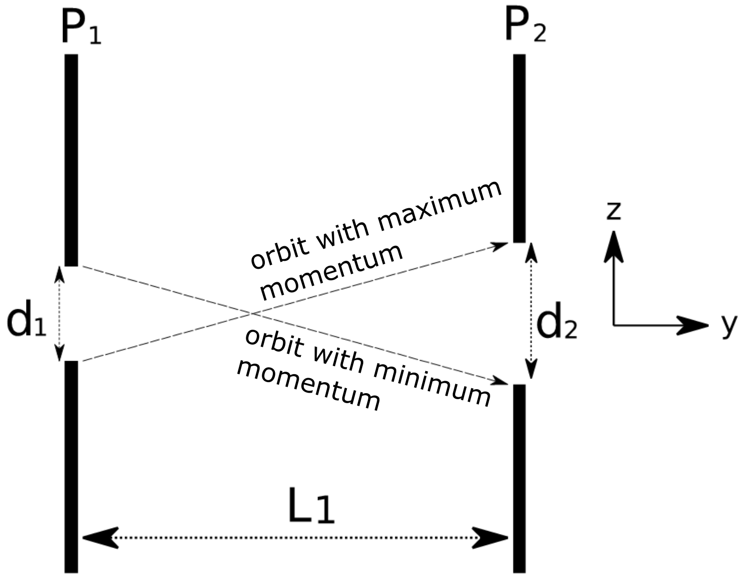

Figure 4.1: Geometry of the collimator and the slit.

From Eq. (3.77) the error $\varepsilon(\sigma_z)$ of the original Stern-Gerlach measurement is given by

$$\varepsilon(\sigma_z)^2 = 2\,\mathrm{erfc}\left(\frac{g_0}{\sqrt{2}\sigma(\Delta t)}\right). \tag{4.11}$$

Then, according to the parameter values given in Table 4.1, we have

$$0.972 \le \frac{g_0}{\sqrt{2}\sigma(\Delta t)} \le 1.62, \tag{4.12}$$

and, therefore, we conclude

$$4.38 \times 10^{-2} \le \varepsilon(\sigma_z)^2 \le 3.38 \times 10^{-1}. \tag{4.13}$$

For the disturbance $\varepsilon(\sigma_x)$ of the original Stern-Gerlach measurement, from Eq. (3.76) we have

$$\eta(\sigma_x)^2 = 2. \tag{4.14}$$

See Section 4.2 for the detailed calculations.

From the above we conclude that the error probability $\varepsilon(\sigma_z)^2/4$ of the experiment is at most 8.5%. This appears to be consistent with Stern-Gerlach's original estimate of the error to be 10% based on the agreement between the observed deflection and the theoretical prediction [20].

As depicted in Figure 4.2, the estimated error-disturbance region clearly violates Heisenberg's EDR.

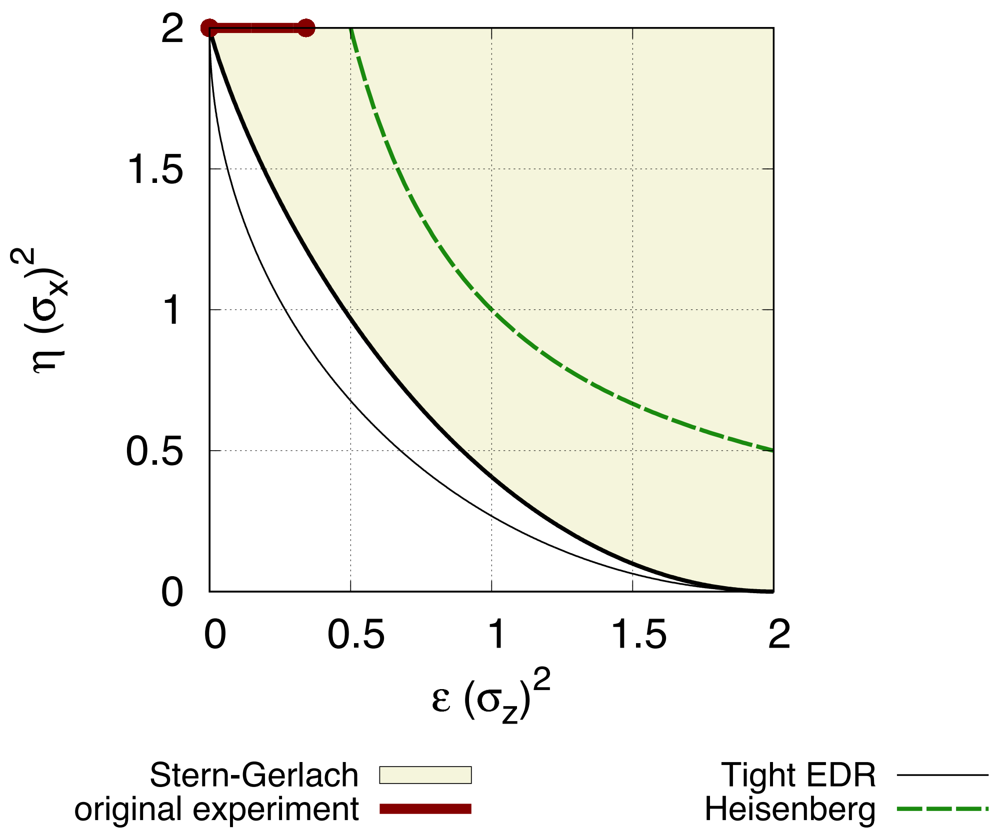

Figure 4.2: The estimated error-disturbance region for the original experiment performed by Gerlach and Stern [18, 19, 20] in 1922. Beige region: the region Eq. (3.98) that Stern-Gerlach measurements can achieve. Deep red line: the estimated error-disturbance region (4.13), (4.14) for the original Stern-Gerlach experiment in 1922. Black thine line: the boundary of the tight EDR (2.34). Green dashed line: the boundary of Heisenberg's EDR Eq. (1.1).

4.2 Derivations of Eq. (4.13) and Eq. (4.14)

From 2018 CODATA, the Boltzmann constant, the Avogadro constant N_A, the electron magnetic moment μ, and the reduced Planck constant $\hbar$ are given by

$$\begin{aligned} k_B &= 1.380649 \times 10^{-23}[\mathrm{J/K}], \\ N_A &= 6.02214076 \times 10^{23} \left[\mathrm{mol}^{-1}\right], \\ \mu &= -9.2847647043 \times 10^{-24}[\mathrm{J/T}], \\ \hbar &= 1.054571817 \times 10^{-34}[\mathrm{J \cdot s}]. \end{aligned}$$

The mass m of the silver atom with the standard atomic weight $107.86822[\mathrm{g/mol}]$ is given by

$$\begin{aligned} m &= \frac{1.0786822 \times 10^{-1}[\mathrm{kg/mol}]}{6.02214076 \times 10^{23}\left[\mathrm{mol}^{-1}\right]} \\ &= 1.7911939 \times 10^{-25}[\mathrm{kg}]. \end{aligned}$$

From Eq. (4.2) and Table 1 we obtain

$$\begin{aligned} v_y &= \sqrt{\frac{4k_BT}{m}} = \sqrt{\frac{4 \times 1.380 \times 10^{-23} \times 1500}{1.791 \times 10^{-25}}} \\ &= 6.80 \times 10^2[\mathrm{m/s}]. \end{aligned}$$

From Table 1 we obtain

$$\begin{aligned} \Delta t &= \frac{L_2}{v_y} \\ &= \frac{3.5 \times 10^{-2}}{6.80 \times 10^2} \\ &= 5.14 \times 10^{-5}[\mathrm{s}]. \end{aligned}$$

As depicted in Figure 4.1, the parameters δP and δZ are introduced as

$$\begin{aligned} \delta P &= \frac{d_1 + d_2}{2L_1} m v_y, \\ \delta Z &= \frac{d_2}{2}. \end{aligned}$$

We obtain

$$\begin{aligned} 1.25\delta Z &= \frac{5d_2}{8} = 2.50 \times 10^{-6}[\mathrm{m}], \\ 1.25\delta P &= \frac{5(d_1 + d_2)}{8L_1} m v_y \\ &= \frac{3.1 \times 10^{-4} + 2.0 \times 10^{-4}}{8 \times 3.3 \times 10^{-2}} \times 1.791 \times 10^{-25} \\ &\quad \times 6.80 \times 10^2 \\ &= 2.35 \times 10^{-25}[\mathrm{kg \cdot m/s}]. \end{aligned}$$

The parameters D_p and D_z are assumed to satisfy

$$D_p = 1.25\, K\delta P,$$
$$D_z = 1.25\, K\delta Z$$

for $0.6 \leq K \leq 1$. We obtain

$$\begin{aligned}
\mathrm{Var}(Z,\xi_\lambda) &= \frac{1}{4}\left(\frac{D_p^2}{\hbar^2} + \frac{1}{4D_z^2}\right)^{-1} \\
&= \frac{1}{4}\Bigg(\frac{(K \times 2.35 \times 10^{-25}[\mathrm{kg}\cdot\mathrm{m/s}])^2}{(1.054 \times 10^{-34}[\mathrm{J}\cdot\mathrm{s}])^2} \\
&\qquad + \frac{1}{4(K \times 2.50 \times 10^{-6}[\mathrm{m}])^2}\Bigg)^{-1} \\
&= \frac{1}{4}\big(K^2 \times 4.97 \times 10^{18}[\mathrm{m}^{-2}] \\
&\qquad + K^{-2} \times 4.00 \times 10^{10}[\mathrm{m}^{-2}]\big)^{-1} \\
&= K^{-2} \times 5.03 \times 10^{-20}[\mathrm{m}^2], \\
\frac{\Delta t^2}{m^2}\mathrm{Var}(P,\xi_\lambda) &= \frac{\Delta t^2}{m^2}\frac{\hbar^2}{4\mathrm{Var}(Z,\xi_\lambda)} \\
&= \frac{(5.14 \times 10^{-5})^2}{(1.791 \times 10^{-25})^2} \\
&\qquad \times \frac{(1.054 \times 10^{-34})^2}{4 \times K^{-2} \times 5.03 \times 10^{-20}} \\
&= K^2 \times 4.54 \times 10^{-9}[\mathrm{m}^2], \\
\sigma(\Delta t)^2 &= \mathrm{Var}(Z,\xi_\lambda) + \frac{\Delta t^2}{m^2}\mathrm{Var}(P,\xi_\lambda) \\
&= \frac{\Delta t^2}{m^2}\mathrm{Var}(P,\xi_\lambda) \\
&= K^2 \times 4.54 \times 10^{-9}[\mathrm{m}^2],
\end{aligned}$$

$$\begin{aligned}
g_0 &= \frac{\mu B_1 \Delta t^2}{2m} \\
&= \frac{(-9.28 \times 10^{-24}[\mathrm{J/T}])}{2 \times (1.791 \times 10^{-25}[\mathrm{kg}])} \\
&\qquad \times (-1.35 \times 10^3[\mathrm{T/m}]) \\
&\qquad \times (5.14 \times 10^{-5}[\mathrm{s}])^2 \\
&= 9.26 \times 10^{-5}[\mathrm{m}], \\
\frac{g_0}{\sqrt{2}\sigma(\Delta t)} &= \frac{9.26 \times 10^{-5}}{K\sqrt{2 \times 4.54 \times 10^{-9}}} \\
&= K^{-1} \times 0.972.
\end{aligned}$$

From Eq. (4.11) we have

$$\begin{aligned}\varepsilon(\sigma_z)^2 &= 2\,\mathrm{erfc}\left(\frac{g_0}{\sqrt{2}\sigma(\Delta t)}\right)\\ &= 2\,\mathrm{erfc}\left(K^{-1}\times 0.972\right).\end{aligned}$$

For $K = 1$, we obtain

$$2\,\mathrm{erfc}(0.972) = 2\times 0.1692 = 3.38\times 10^{-1}.$$

For $K = 0.6$, we obtain

$$\begin{aligned}2\,\mathrm{erfc}(0.972/0.6) &= 2\,\mathrm{erfc}(1.620) = 2\times 0.0219\\ &= 4.38\times 10^{-2}.\end{aligned}$$

Thus, we conclude

$$0.972 \le \frac{g_0}{\sqrt{2}\sigma(\Delta t)} \le 1.620,$$
$$4.38\times 10^{-2} \le \varepsilon(\sigma_z)^2 \le 3.38\times 10^{-1}.$$

To calculate the disturbance $\eta(\sigma_x)$, we have

$$\begin{aligned}\sigma\left(\frac{\Delta t}{2}\right)^2 &= \frac{1}{4}\sigma\,(\Delta t)^2\\ &= K^2\times 1.135\times 10^{-9}[\mathrm{m}^2],\\ \frac{\mu B_1\Delta t}{\hbar} &= \frac{(-9.28\times 10^{-24}[\mathrm{J/T}])}{1.054\times 10^{-34}[\mathrm{J\cdot s}]}\\ &\quad\times(-1.35\times 10^{3}[\mathrm{T/m}])\\ &\quad\times(5.14\times 10^{-5}[\mathrm{s}])\\ &= 6.10\times 10^{9}[\mathrm{m}^{-1}],\\ \frac{2\mu^2B_1^2\Delta t^2}{\hbar^2}\sigma\left(\frac{\Delta t}{2}\right)^2 &= 2\times(6.10\times 10^{9}[\mathrm{m}^{-1}])^2\\ &\quad\times K^2\times 1.135\times 10^{-9}[\mathrm{m}^2]\\ &= K^2\times 8.44\times 10^{10},\\ 2\exp\left[-\frac{2\mu^2B_1^2\Delta t^2}{\hbar^2}\sigma\left(\frac{\Delta t}{2}\right)^2\right]&\\ &= 2\exp(-K^2\times 8.44\times 10^{10})\\ &= 0.\end{aligned}$$

Thus, from Eq. (3.76) we conclude

$$\begin{aligned}\eta(\sigma_x)^2 &= 2 - 2\exp\left[-\frac{2\mu^2B_1^2\Delta t^2}{\hbar^2}\sigma\left(\frac{\Delta t}{2}\right)^2\right]\cos\frac{2\mu\Delta t B_0}{\hbar}\\ &= 2.\end{aligned}$$

Chapter 5

Conclusion

Stern-Gerlach measurements, originally performed by Gerlach and Stern [18, 19, 20], have been discussed for a long tims as a typical model or a paradigm of quantum measurement [3]. As Heisenberg's uncertainty principle suggests, Stern-Gerlach measurements of one spin component inevitably disturb its orthogonal component, and Heisenberg's EDR (1.7) has been commonly believed to be its precise quantitative expression. However, general quantitative relations between error and disturbance in arbitrary quantum measurements have been extensively investigated over the past two decades and universally valid EDRs have been obtained to reform Heisenberg's original EDR (see, e.g., [37, 16, 5, 10, 45] and references therein).

Here we investigated the EDR for this familiar class of measurements in light of the general theory leading to the universally valid EDR relations. We have determined the range of possible values of the error and disturbance achievable by arbitrary Stern-Gerlach apparatuses, assuming that the orbital state is a Gaussian state. Our result is depicted in Fig. 3.2 and the boundary of the error-disturbance region is given in Eq. (3.98) as a closed formula. The result shows that the error-disturbance region of Stern-Gerlach measurements occupies a near-optimal subregion of the universally valid error-disturbance region for arbitrary measurements. It can be seen that one of the earliest methods of quantum measurement violates Heisenberg's EDR (1.7) in a broad range of experimental parameters. Furthermore, we found a class of initial orbital states in which the error can be minimized an arbitrarily small amount by the screen at a finite distance from the magnet in contrast to the conventional assumption that the error decreases asymptotically.

Based on the above theoretical results, we have estimated the error and disturbance of the original Stern-Gerlach experiment performed in 1922, and concluded that the original Stern-Gerlach experiment violates Heisenberg's EDR.

The relation for the general class of states beyond Gaussian states is left to future study. In addition, we also leave it to future research to analyze more realistic models, for example, a model described by the magnetic field satisfying Maxwell's equations [13, 46] or a model considering the decoherence of the particle during the measuring process [15].

Our results will contribute to answer the question as to how various experimental parameters can be controlled to achieve the ultimate limit. These results also suggest that Heisenberg's EDR is more ubiquitously violated than we have believed for a long time, and it opens a new research interest exploring violations of Heisenberg's EDR in more common measurement setups to deepen our understanding of Heisenberg's uncertainty principle. It will contribute to new developments in precision measurements such as optomechanical metrology and multi-messenger astronomy. We expect that the present study will provoke further experimental studies.

Bibliography

[1] S-Y. Baek, F. Kaneda, M. Ozawa, and K. Edamatsu, Experimental violation and reformulation of the Heisenberg's error-disturbance uncertainty relation, *Scientific Reports* **3**, 2221 (2013).

[2] H. F. Baker, Alternants and continuous groups, *Proceedings of the London Mathematical Society* **s2-3**, 24–47 (1905).

[3] D. Bohm, *Quantum theory*, (Prentice-Hall, Englewood Cliffs, NJ, 1951).

[4] V. B. Braginsky, Y. I. Vorontsov, and K. S. Thorne, Quantum nondemolition measurements, *Science* **209**, 547–557 (1980).

[5] C. Branciard, Error-tradeoff and error-disturbance relations for incompatible quantum measurements, *Proceedings of the National Academy of Sciences of the United States of America* **110**, 6742–6747 (2013).

[6] C. Branciard, Deriving tight error-trade-off relations for approximate joint measurements of incompatible quantum observables, *Physical Review A* **89**, 022124 (2014).

[7] F. Buscemi, M. J. W. Hall, M. Ozawa, and M. M. Wilde, Noise and disturbance in quantum measurements: an information-theoretic approach, *Physical Review Letters* **112**, 050401 (2014).

[8] P. Busch, T. Heinonen, and P. Lahti, Noise and disturbance in quantum measurement, *Physics Letters A* **320**, 261–270 (2003).

[9] P. Busch, P. Lahti, and R. F. Werner, Proof of Heisenberg's error-disturbance relation, *Physical Review Letters* **111**, 160405 (2013).

[10] P. Busch, P. Lahti, and R. F. Werner, Quantum rms error and Heisenberg's error-disturbance relation, *Reviews of Modern Physics* **86**, 1261–1281 (2014).

[11] C. M. Caves, Defense of the standard quantum limit for free-mass position, *Physical Review Letters* **54**, 2465–2468 (1985).

[12] C. M. Caves, K. S. Thorne, R. W. P. Drever, V. D. Sandberg, and M. Zimmermann, On the measurement of a weak classical force coupled to a quantum-mechanical oscillator. I. Issues of principle, *Reviews of Modern Physics* **52**, 341–392 (1980).

[13] S. Cruz-Barrios and J. Gómez-Camacho, Semiclassical description of Stern-Gerlach experiments, *Physical Review A* **63**, 012101 (2000).

[14] B. Demirel, S. Sponar, G. Sulyok, M. Ozawa, and Y. Hasegawa, Experimental test of residual error-disturbance uncertainty relations for mixed spin-$1/2$ states, *Physical Review Letters* **117**, 140402 (2016).

[15] M. Devereux, Reduction of the atomic wavefunction in the Stern-Gerlach magnetic field, *Canadian Journal of Physics* **93**, 1382–1390 (2015).

[16] J. Erhart, S. Sponar, G. Sulyok, G. Badurek, M. Ozawa, and Y. Hasegawa, Experimental demonstration of a universally valid error-disturbance uncertainty relation in spin measurements, *Nature Physics* **8**, 185–189 (2012).

[17] C. F. Gauss, *Theory of the Combination of Observations Least Subject to Errors, Part One, Part Two, Supplement*, (Society for Industrial and Applied Mathematics1995). [Originally published: *Theoria Combinationis Observationum: Erroribus Minimis Obnoxiae*, Pars prior, Pars posterior, Supplementum, (Societati Regiae Exhibita, 1821)].

[18] W. Gerlach and O. Stern, Der experimentelle Nachweis des magnetischen Moments des Silberatoms, *Zeitschrift für Physik* **8**, 110–111 (1922).

[19] W. Gerlach and O. Stern, Der experimentelle Nachweis der Richtungsquantelung im Magnetfeld, *Zeitschrift für Physik* **9**, 349–352 (1922).

[20] W. Gerlach and O. Stern, Das magnetische Moment des Silberatoms, *Zeitschrift für Physik* **9**, 353–355 (1922).

[21] R. J. Glauber, Coherent and incoherent states of the radiation field, *Physical Review Series II* **131**, 2766–2788 (1963).

[22] W. Heisenberg, Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik, *Zeitschrift für Physik* **43**, 172–198 (1927). [English translation by J. A. Wheeler and W. H. Zurek, in *Quantum Theory and Measurement*, edited by J. A. Wheeler and W. H. Zurek (Princeton UP, Princeton, NJ, 1983), pp. 62–84].

[23] D. Home, A. K. Pan, M. M. Ali, and A. S. Majumdar, Aspects of nonideal Stern-Gerlach experiment and testable ramifications, *Journal of Physics A: Mathematical and Theoretical* **40**, 13975–13988 (2007).

[24] Y. Inoue and M. Ozawa, Error-disturbance relation in Stern-Gerlach measurements, *Physical Review A* **102**, 042226 (2020).

[25] Y. Inoue and M. Ozawa, Violation of Heisenberg's error-disturbance relation by Stern-Gerlach measurements, (2020) *EPL (Europhysics Letters)* (in press).

[26] R. Jozsa, Complex weak values in quantum measurement, *Physical Review A* **76**, 044103 (2007).

[27] F. Kaneda, S-Y. Baek, M. Ozawa, and K. Edamatsu, Experimental test of error-disturbance uncertainty relations by weak measurement, *Physical Review Letters* **112**, 020402 (2014).

[28] E. Kennard. Zur Quantenmechanik einfacher Bewegungstypen, *Zeitschrift für Physik* **44**, 326–352 (1927).

[29] K. Korzekwa, D. Jennings, and T. Rudolph, Operational constraints on state-dependent formulations of quantum error-disturbance trade-off relations, *Physical Review A* **89**, 052108 (2014).

[30] X.-M. Lu, S. Yu, K. Fujikawa, and C. H. Oh, Improved error-tradeoff and error-disturbance relations in terms of measurement error components, *Physical Review A* **90**, 042113 (2014).

[31] A. P. Lund and H. M. Wiseman, Measuring measurement-disturbance relationships with weak values, *New Journal of Physics* **12**, 093011 (2010).

[32] M. Ozawa, Quantum measuring processes of continuous observables, *Journal of Mathematical Physics* **25**, 79–87 (1984).

[33] M. Ozawa, Measurement breaking the standard quantum limit for free-mass position, *Physical Review Letters* **60**, 385–388 (1988).

[34] M. Ozawa, Realization of measurement and standard quantum limit, In P. Tombesi and E. R. Pike, *Squeezed and nonclassical light*, 263–286 (Plenum Press, New York, 1989).

[35] M. Ozawa, Canonical approximate quantum measurements, *Journal of Mathematical Physics* **34**, 5596–5624 (1993).

[36] M. Ozawa, Physical content of heisenberg's uncertainty relation: limitation and reformulation, *Physics Letters A* **318**, 21–29 (2003).

[37] M. Ozawa, Universally valid reformulation of the Heisenberg uncertainty principle on noise and disturbance in measurement, *Physical Review A* **67**, 042105 (2003).

[38] M. Ozawa, Uncertainty relations for noise and disturbance in generalized quantum measurements, *Annals of Physics* **311**, 350–416 (2004).

[39] M. Ozawa, Perfect correlations between noncommuting observables, *Physics Letters A* **335**, 11–19 (2005).

[40] M. Ozawa, Universal uncertainty principle in the measurement operator formalism, *Journal of Optics B* **7**, s672–s681 (2005).

[41] M. Ozawa, Noise and disturbance in quantum measurements and operations, *Proceedings of SPIE* **6244**, 190–198 (2006).

[42] M. Ozawa, Quantum perfect correlations, *Annals of Physics* **321**, 744–769 (2006).

[43] M. Ozawa, Error-disturbance relations in mixed states, arXiv:1404.3388 (2014).

[44] M. Ozawa, Heisenberg's original derivation of the uncertainty principle and its universally valid reformulations, *Current Science* **109**, 2006–2016 (2015).

[45] M. Ozawa, Soundness and completeness of quantum root-mean-square errors, *npj quantum information* **5**, 1 (2019).

[46] G. Potel, F. Barranco, S. Cruz-Barrios, and J. Gómez-Camacho, Quantum mechanical description of Stern-Gerlach experiments, *Physical Review A* **71**, 052106 (2005).

[47] M. Ringbauer, D. N. Biggerstaff, M. A. Broome, A. Fedrizzi, C. Branciard, and A. G. White, Experimental joint quantum measurements with minimum uncertainty, *Physical Review Letters* **112**, 020401 (2014).

[48] L. A. Rozema, A. Darabi, D. H. Mahler, A. Hayat, Y. Soudagar, and A. M. Steinberg, Violation of Heisenberg's measurement-disturbance relationship by weak measurements, *Physical Review Letters* **109**, 100404 (2012).

[49] E. Schrödinger, Zum Heisenbergschen Unschärfeprinzip, *Proceedings of The Prussian Academy of Sciences Physics-Mathematical Section* **19**, 296–303 (1930).

[50] B. L. Schumaker, Quantum mechanical pure states with Gaussian wave functions, *Physics Reports* **135**, 317–408 (1986).

[51] M. O. Scully, W. E. Lamb, and A. Barut, On the theory of the Stern-Gerlach apparatus, *Foundations of Physics* **17**, 575–583 (1987).

[52] O. Stern, Eine direkte Messung der thermischen Molekulargeschwindigkeit, *Zeitschrift für Physik* **3**, 417–421 (1920).

[53] G. Sulyok, S. Sponar, B. Demirel, F. Buscemi, M. J. W. Hall, M. Ozawa, and Y. Hasegawa, Experimental test of entropic noise-disturbance uncertainty relations for spin-$1/2$ measurements, *Physical Review Letters* **115**, 030401 (2015).

[54] G. Sulyok, S. Sponar, J. Erhart, G. Badurek, M. Ozawa, and Y. Hasegawa, Violation of Heisenberg's error-disturbance uncertainty relation in neutron-spin measurements, *Physical Review A* **88**, 022110 (2013).

[55] C. Weedbrook, S. Pirandola, R. G. Patrón, N. J. Cerf, T. C. Ralph, J. H. Shapiro, and S. Lloyd, Gaussian quantum information, *Reviews of Modern Physics* **84**, 621–669 (2012).

[56] M. M. Weston, M. J. W. Hall, M. S. Palsson, H. M. Wiseman, and G. J. Pryde, Experimental test of universal complementarity relations, *Physical Review Letters* **110**, 220402 (2013).

[57] H. P. Yuen, Two-photon coherent states of the radiation field, *Physical Review A* **13**, 2226–2243 (1976).

[58] H. P. Yuen, Contractive states and the standard quantum limit for monitoring free-mass positions, *Physical Review Letters* **51**, 719–722 (1983).