A reexamination of the Wigner and Araki-Yanase theorem, 2

Ryukyu Math. J., 7(1994), 45-64

A REEXAMINATION OF THE WIGNER AND ARAKI-YANASE THEOREM, II

Shuichi Matsumoto

Abstract. Measurement of a spin component of a spin-$1/2$ particle is reconsidered under the constraint that all three components of total angular momentum are conserved in the measuring process. We obtain an inequality by which we can estimate the amount of disturbance of the measured system caused by the measurement. Our formula has more information than the inequalities previously obtained by Yanase and others in the sense that it shows us how to choose the initial state of the apparatus which will make the measurement closest to the ideal. It is emphasized that the conventional interpretation of the Wigner-Araki-Yanase theorem, in which much attention is focussed only on the size of the apparatus, fails to take into account the initial state of the appratus.

1. Introduction

It is well-known [1)-3)] that any quantum mechanical observable which does not commute with an additive conserved quantity cannot be measured according to the ideal scheme proposed by von Neumann [4)]. The proof of this claim can be sketched as follows.

Let $\mathcal{H}_S$ and $\mathcal{H}_A$ be the Hilbert spaces of the observed system and the measuring apparatus, respectively. Let M be the operator of the space $\mathcal{H}_S$ corresponding to the measured observable and $\{\phi_m\}$ be complete orthonormal eigenstates of M $(M\phi_m = m\phi_m)$. (In this ar-

Received November 30, 1994

ticle, we assume that M has a non-degenerate discrete spectrum. For an observable with a continuous spectrum, see Ref. 5)). The ideal measurement scheme assumes the existence of a normalized state A in the Hilbert space $\mathcal{H}_A$ and of a unitary operator U_t on $\mathcal{H}_S \otimes \mathcal{H}_A$ representing the time evolution of the measured system plus apparatus, such that

$$U_t(\phi_m \otimes A) = \phi_m \otimes A_m \quad \text{and} \quad \langle A_m | A_{m'} \rangle = \delta_{m,m'}, \tag{1.1}$$

where A_m are the final states of the apparatus.

Now let $L_S + L_A$ be the additive conserved quantity of the measured system plus apparatus. Then we have

$$\begin{aligned} &\langle \phi_m \otimes A | U_t^\dagger (L_S + L_A) U_t | \phi_{m'} \otimes A \rangle \\ &\qquad = \langle \phi_m \otimes A | L_S + L_A | \phi_{m'} \otimes A \rangle \end{aligned} \tag{1.2}$$

for all m and m'. Substituting (1.1) into (1.2), we have

$$\begin{aligned} &\langle \phi_m | L_S | \phi_{m'} \rangle \delta_{m,m'} + \delta_{m,m'} \langle A_m | L_A | A_{m'} \rangle \\ &\qquad = \langle \phi_m | L_S | \phi_{m'} \rangle + \delta_{m,m'} \langle A | L_A | A \rangle, \end{aligned} \tag{1.3}$$

and therefore

$$\langle \phi_m | L_S | \phi_{m'} \rangle = 0 \qquad (m \neq m'). \tag{1.4}$$

Eq. (1.4) indicates that $[M, L_S] = 0$, because it claims that

$$\langle \phi_m | \ [M, L_S] \ | \phi_{m'} \rangle = (m - m') \langle \phi_m | L_S | \phi_{m'} \rangle = 0 \tag{1.5}$$

for all m and m'.

Thus if $[M, L_S] \neq 0$ then M cannot be measured according to the ideal scheme (1.1). We must then resort to a nonideal measurement scheme. To this end, we start by writing an equation expressing the most general type of time evolution :

$$U_t(\phi_m \otimes A) = \sum_{m'} \phi_{m'} \otimes A_{m',m}. \tag{1.6}$$

When the equations

$$A_{m',m} = 0 \quad (m \neq m') \tag{1.7}$$

and

$$\langle A_{m,m} | A_{m',m'} \rangle = 0 \quad (m \neq m') \tag{1.8}$$

are both satisfied, the evolution (1.6) reduces to the ideal one (1.1). The presence of nonzero states $A_{m',m}$ $(m \neq m')$ corresponds to some disturbance of the measured system which is caused by the measuring process and so to an erroneous value for the measurement. The non-orthogonality of the states $A_{m,m}$ corresponds likewise to the same effects. Therefore the problem arises of minimizing the norms of the states $A_{m',m}$ $(m \neq m')$, and of making the states $A_{m,m}$ as nearly orthogonal as possible.

This has been the subject of many investigations, initially [6] for the case in which the measured observable is a spin component of a particle with spin 1/2 and the conserved quantity is a component of the total angular momentum. It has been subsequently discussed in more general cases [7)–10)]. In these investigations the condition (1.8) is assumed in principle and the lower bound of $\sum_{m' \neq m} ||A_{m',m}||^2$ is estimated. Their results take essentially the following form :

$$\sum_{m' \neq m} ||A_{m',m}||^2 \geq \frac{\text{const.}}{\langle A | {L_A}^2 | A \rangle}. \tag{1.9}$$

The inequality (1.9), however, is somewhat unsatisfactory in the following senses : Firstly, as we will show in Sec. 3 under some reasonable constraints, it is possible to construct some initial state A for which the left-hand side of (1.9) cannot be small although the right-hand side can be made arbitrarily small. In this sense, the inequality (1.9) contains no information about how to choose an initial state

A such that the value of $\sum_{m'\neq m} ||A_{m',m}||^2$ is sufficiently small. Secondly, it has been derived by consideration of only one conservation law involved in the measurement process.

The objective of this article is to present a new method to resolve the above unsatisfactory points. We consider the process in which a component of the spin angular momentum of a particle is measured. By taking into account all three components of angular momentum, each of which should be conserved, we arrive at a more desirable inequality than (1.9), in the sense that it indicates which initial state of the apparatus we should select in order to lower the value of $\sum_{m'\neq m} ||A_{m',m}||^2$.

2. Formal preliminaries

In this section, we consider a measuring process in which the $z-$component of the spin angular momentum of a particle with spin j is measured. We denote the spin angular momentum vector of the observed system and of the measuring apparatus by (s_x, s_y, s_z) and (S_x, S_y, S_z), respectively. Then $(s_x + S_x, s_y + S_y, s_z + S_z)$ is of course the spin angular momentum vector of the combined system $\mathcal{H}_S \otimes \mathcal{H}_A$. Let σ_x, σ_y and σ_z be the Pauli matrices. They then generate the Lie algebra of the special unitary group $SU(2)$, and there exists an action of $SU(2)$ on $\mathcal{H}_S$ such that the actions of $\sigma_x/2, \sigma_y/2$ and $\sigma_z/2$ correspond to s_x, s_y and s_z, respectively. In the same manner we have some action of $SU(2)$ on $\mathcal{H}_A$ $(\mathcal{H}_S \otimes \mathcal{H}_A)$, in which the actions of $\sigma_x/2, \sigma_y/2$ and $\sigma_z/2$ coincide with S_x, S_y and S_z $(s_x + S_x, s_y + S_y$ and $s_z + S_z)$, respectively.

In this article we neglect the spatial dimensions of the measured system, the measuring apparatus and the combined system. Under this assumption s_x+S_x, s_y+S_y and s_z+S_z are all conserved quantities, and therefore they have to commute with the time evolution U_t :

$$[s_x + S_x, U_t] = [s_y + S_y, U_t] = [s_z + S_z, U_t] = 0. \tag{2.1}$$

Hence, if we denote the Hamiltonian of the combined system $\mathcal{H}_S \otimes \mathcal{H}_A$ by H, we have

$$[s_x + S_x, H] = [s_y + S_y, H] = [s_z + S_z, H] = 0, \tag{2.2}$$

i.e., the operator H has to commute with the action of each element of $SU(2)$.

We define the Casimir operator of $SU(2)$ by

$$\Omega \equiv \frac{1}{4}({\sigma_x}^2 + {\sigma_y}^2 + {\sigma_z}^2). \tag{2.3}$$

The actions of Ω on $\mathcal{H}_S, \mathcal{H}_A$ and $\mathcal{H}_S \otimes \mathcal{H}_A$ are given by

$$\mathrm{s}^2, \quad \mathbf{S}^2 \quad \text{and} \quad (\mathrm{s} + \mathbf{S})^2 = \mathrm{s}^2 + \mathbf{S}^2 + 2\, \mathrm{s} \cdot \mathbf{S}, \tag{2.4}$$

respectively, where we set $\mathrm{s} = (s_x, s_y, s_z)$ and $\mathbf{S} = (S_x, S_y, S_z)$. Since the action of the Casimir operator is always commutable with all actions of the elements of $SU(2)$, if we define the Hamiltonian H by

$$H = \alpha\, \mathrm{s}^2 + \beta\, \mathbf{S}^2 + \gamma\, \mathrm{s} \cdot \mathbf{S} \tag{2.5}$$

for some real constants α, β and γ, then H commutes with the action of each element of $SU(2)$. The author believes that this selection of H is reasonable.

For each non-negative half integer $k = 0, 1/2, 1, \cdots$, there exists [11] an irreducible unitary representation (say, V_k) of $SU(2)$ such that

$$\dim V_k = 2k + 1. \tag{2.6}$$

It would be natural to think that

$$\mathcal{H}_S = V_j. \tag{2.7}$$

Let

$$\mathcal{H}_A = \sum_p V_{k(p)} \tag{2.8}$$

be an irreducible decomposition of $\mathcal{H}_A$ with respect to the action of $SU(2)$. Since a decomposition

$$V_j \otimes V_k = \sum_{l=-j}^{j} V_{k+l} \tag{2.9}$$

exists for each k, we have

$$\mathcal{H}_S \otimes \mathcal{H}_A = \sum_p \sum_{l=-j}^{j} V_{k(p)+l}. \tag{2.10}$$

We can conclude from the Schur's lemma [12] that there are real constants $\{a_{p,l}\}$ such that the operator H coincides with $a_{p,l}I$ on the space $V_{k(p)+l}$, i.e.,

$$H|_{V_{k(p)+l}} = a_{p,l}I \tag{2.11}$$

where I is the identity operator on $V_{k(p)+l}$, because (a) the Hamiltonian H given by Eq. (2.5) remains $V_j \otimes V_{k(p)}$ invariant, (b) H commutes with the action of each element of $SU(2)$, and (c) the multiplicity of an irreducible representation in the decomposition $V_j \otimes V_{k(p)} = \sum_{l=-j}^{j} V_{k(p)+l}$ is at most one.

There is, for each k, an orthonormal system ϕ_n^k $(n = k, k-1, \cdots, -k)$ in V_k such that

$$\frac{1}{2}\sigma_z \phi_n^k = n\phi_n^k, \tag{2.12}$$

$$\begin{aligned} \sigma_+ \phi_n^k &= \sqrt{(k-n)(k+n+1)}\phi_{n+1}^k, \\ \sigma_- \phi_n^k &= \sqrt{(k-n+1)(k+n)}\phi_{n-1}^k \end{aligned} \tag{2.13}$$

where $\sigma_\pm \equiv (\sigma_x \pm i\sigma_y)/2$. In our measuring process ϕ_m^j $(m = j, j-1, \cdots, -j)$ corresponds to ϕ_m in Sec. 1. For the sake of brevity we shall sometimes use the following notations :

$$|k, n\rangle \equiv \phi_n^k, \qquad |j, m; k, n\rangle \equiv \phi_m^j \otimes \phi_n^k. \tag{2.14}$$

The space

$$V_j \otimes V_k = \sum_{l=-j}^{j} V_{k+l} \tag{2.15}$$

is spanned by the vectors $|j,m;k,n\rangle$ $(m = j, j-1, \cdots, -j\ ;\ n = k, k-1, \cdots, -k)$, and the Clebsch-Gordan coefficients $c_{m,n}^{k+l}$ $(l = j, j-1, \cdots, -j)$ are defined [11)] by the equation

$$|k+l,\mu\rangle = \sum_{m+n=\mu} c_{m,n}^{k+l}\ |j,m;k,n\rangle. \tag{2.16}$$

These coefficients satisfy the following equations :

$$\begin{aligned} c_{m,n}^{k+l} &= \langle j,m;k,n|k+l,m+n\rangle, \\ |j,m;k,n\rangle &= \sum_l c_{m,n}^{k+l\,*}|k+l,m+n\rangle \end{aligned} \tag{2.17}$$

and

$$\sum_l |c_{m,n}^{k+l}|^2 = 1 \tag{2.18}$$

for all m and n, where $c_{m,n}^{k+l\,*}$ is the complex conjugate of $c_{m,n}^{k+l}$.

Now we can express the initial state $A \in \mathcal{H}_A = \sum_p V_{k(p)}$ as follows

$$A = \sum_p \lambda_p A_p, \quad A_p = \sum_{n=-k(p)}^{k(p)} c_{p,n}|k(p),n\rangle \in V_{k(p)} \tag{2.19}$$

where we may assume that

$$\sum_p |\lambda_p|^2 = 1 \ \text{ and } \ ||A_p||^2 = \sum_n |c_{p,n}|^2 = 1. \tag{2.20}$$

Then, using (2.19), (2.17), (2.11) and (2.16), we can calculate the vectors $A_{m',m}$ in (1.6) :

$$\begin{aligned}
&U_t(\phi^j_m \otimes A)\\
&= \sum_p \lambda_p \sum_{n=-k(p)}^{k(p)} c_{p,n}\\
&\qquad \times \left[\sum_l c^{k(p)+l^*}_{m,n} e^{-ita_{p,l}} |k(p)+l, m+n\rangle \right] \qquad (2.21)\\
&= \sum_{m'} \phi^j_{m'} \otimes \left[\sum_p \lambda_p \sum_n c_{p,n} \right.\\
&\qquad \left. \times \left(\sum_l e^{-ita_{p,l}} c^{k(p)+l^*}_{m,n} c^{k(p)+l}_{m',m+n-m'} \right) \phi^{k(p)}_{m+n-m'} \right],
\end{aligned}$$

i.e., we have

$$\begin{aligned}
A_{m',m} &= \sum_p \lambda_p \sum_n c_{p,n}\\
&\quad \times \left(\sum_l e^{-ita_{p,l}} c^{k(p)+l^*}_{m,n} c^{k(p)+l}_{m',m+n-m'} \right) \phi^{k(p)}_{m+n-m'}.
\end{aligned} \qquad (2.22)$$

For the sake of convenience we here introduce "the p-th component $A_{p;m',m}$" of $A_{m',m}$ by

$$\begin{aligned}
A_{m',m} &= \sum_p \lambda_p A_{p;m',m},\\
A_{p;m',m} & \qquad (2.23)\\
&= \sum_n c_{p,n} \left(\sum_l e^{-ita_{p,l}} c^{k(p)+l^*}_{m,n} c^{k(p)+l}_{m',m+n-m'} \right) \phi^{k(p)}_{m+n-m'}.
\end{aligned}$$

Under the above settings, we now try to estimate the value of $\sum_{m' \neq m} ||A_{m',m}||^2$. Firstly, we have that

$$\sum_{m'} ||A_{m',m}||^2 = ||U_t(\phi^j_m \otimes A)||^2 = 1 \qquad (2.24)$$

for each m, and hence that

$$\sum_{m' \neq m} ||A_{m',m}||^2 = 2j + 1 - \sum_m ||A_{m,m}||^2. \qquad (2.25)$$

Secondly, using (2.22) we have

$$
\begin{aligned}
&\langle A_{m,m} | A_{m',m'} \rangle \\
&= \sum_p |\lambda_p|^2 \sum_{n=-k(p)}^{k(p)} |c_{p,n}|^2 \left(\sum_l e^{ita_{p,l}} |c_{m,n}^{k(p)+l}|^2 \right) \\
&\qquad\qquad \times \left(\sum_{l'} e^{-ita_{p,l'}} |c_{m',n}^{k(p)+l'}|^2 \right) \\
&= \sum_p |\lambda_p|^2 \sum_{n=-k(p)}^{k(p)} |c_{p,n}|^2 \\
&\qquad\qquad \times \left[\sum_{l,l'} e^{it(a_{p,l}-a_{p,l'})} |c_{m,n}^{k(p)+l}|^2 |c_{m',n}^{k(p)+l'}|^2 \right] .
\end{aligned}
\tag{2.26}
$$

Therefore if we denote the matrix whose (l,m)–component is $|c_{m,n}^{k+l}|^2$ by $C_{k,n}$, and the column vector whose l–component is $e^{ita_{p,l}}$ by α_p, then we have

$$
\begin{aligned}
&\langle A_{m,m} | A_{m',m'} \rangle \\
&= \sum_p |\lambda_p|^2 \sum_{n=-k(p)}^{k(p)} |c_{p,n}|^2 \left(C_{k(p),n}^\dagger \alpha_p \alpha_p^\dagger C_{k(p),n} \right)_{m,m'} .
\end{aligned}
\tag{2.27}
$$

Combining (2.25) and (2.27), we get

$$
\begin{aligned}
&\sum_{m' \neq m} \| A_{m',m} \|^2 \\
&= \sum_p |\lambda_p|^2 \sum_{n=-k(p)}^{k(p)} |c_{p,n}|^2 \left[2j + 1 - \alpha_p^\dagger C_{k(p),n} C_{k(p),n}^\dagger \alpha_p \right] .
\end{aligned}
\tag{2.28}
$$

3. The case of $j = 1/2$

In order to make our argument as clear as possible, we restrict ourselves to the case of $j = 1/2$ in the rest of this article. Under this restriction we can calculate explicitly the value of (2.28). The Racah's formula [11),13)] tells us that

$$
\begin{aligned}
c^{k+1/2}_{1/2,n} &= \sqrt{\frac{k+n+1}{2k+1}}, & c^{k+1/2}_{-1/2,n} &= \sqrt{\frac{k-n+1}{2k+1}}, \\
c^{k-1/2}_{1/2,n} &= \sqrt{\frac{k-n}{2k+1}}, & c^{k-1/2}_{-1/2,n} &= -\sqrt{\frac{k+n}{2k+1}},
\end{aligned}
\tag{3.1}
$$

and hence that

$$
C_{k,n} = \begin{pmatrix} \frac{k+n+1}{2k+1} & \frac{k-n+1}{2k+1} \\ \frac{k-n}{2k+1} & \frac{k+n}{2k+1} \end{pmatrix}. \tag{3.2}
$$

Let us define the subspaces V and W of the 2–dimensional complex vector space $\mathbf{C}^2$ by

$$
V = \mathbf{C}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{and} \quad W = \mathbf{C}\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \tag{3.3}
$$

and denote the $V-$ and $W-$component of the vector α_p by β_p and γ_p, respectively ;

$$
\alpha_p = \beta_p + \gamma_p, \tag{3.4}
$$

$$
\begin{aligned}
\beta_p &= \frac{1}{2}(e^{ita_{p,1/2}} + e^{ita_{p,-1/2}})\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \\
\gamma_p &= \frac{1}{2}(e^{ita_{p,1/2}} - e^{ita_{p,-1/2}})\begin{pmatrix} 1 \\ -1 \end{pmatrix}.
\end{aligned}
\tag{3.5}
$$

Direct calculations show us that

$$
\beta_p^\dagger C_{k(p),n} = \beta_p^\dagger, \quad C_{k(p),n}^\dagger \beta_p = \beta_p \tag{3.6}
$$

and that

$$
\begin{aligned}
&2 - \alpha_p^\dagger C_{k(p),n} C_{k(p),n}^\dagger \alpha_p \\
&= 2 - \Big[\beta_p^\dagger \beta_p + \beta_p^\dagger C_{k(p),n}^\dagger \gamma_p \\
&\qquad + \gamma_p^\dagger C_{k(p),n}\beta_p + \gamma_p^\dagger C_{k(p),n} C_{k(p),n}^\dagger \gamma_p\Big].
\end{aligned}
\tag{3.7}
$$

Using (3.5) and (3.2) we can easily obtain that

$$
\begin{aligned}
&2-\beta_p^\dagger\beta_p = 1-\cos t(a_{p,1/2}-a_{p,-1/2}),\\
&\beta_p^\dagger C_{k(p),n}^\dagger\gamma_p+\gamma_p^\dagger C_{k(p),n}\beta_p = 0
\end{aligned}
\tag{3.8}
$$

and that

$$
\begin{aligned}
&\gamma_p^\dagger C_{k(p),n}C_{k(p),n}^\dagger\gamma_p\\
&=\left[1-\cos t(a_{p,1/2}-a_{p,-1/2})\right]\frac{4n^2+1}{(2k(p)+1)^2}.
\end{aligned}
\tag{3.9}
$$

Substituting (3.8) and (3.9) into (3.7) we have

$$
\begin{aligned}
&2-\alpha_p^\dagger C_{k(p),n}C_{k(p),n}^\dagger\alpha_p\\
&=\left[1-\cos t(a_{p,1/2}-a_{p,-1/2})\right]\left(1-\frac{4n^2+1}{(2k(p)+1)^2}\right).
\end{aligned}
\tag{3.10}
$$

Therefore we finally have from (2.28) and (3.10) that

$$
\begin{aligned}
&\sum_{m'\neq m}\|A_{m',m}\|^2\\
&=\sum_p|\lambda_p|^2\left[1-\cos t(a_{p,1/2}-a_{p,-1/2})\right]\\
&\qquad\times\sum_{n=-k(p)}^{k(p)}|c_{p,n}|^2\left(1-\frac{4n^2+1}{(2k(p)+1)^2}\right).
\end{aligned}
\tag{3.11}
$$

On the other hand (2.23) tells us that

$$
\begin{aligned}
&\|A_{p;1/2,1/2}-A_{p;-1/2,-1/2}\|^2\\
&=\|A_{p;1/2,1/2}\|^2+\|A_{p;-1/2,-1/2}\|^2\\
&\qquad-2\mathrm{Re}\langle A_{p;1/2,1/2}|A_{p;-1/2,-1/2}\rangle\\
&=\|A_{p;1/2,1/2}\|^2+\|A_{p;-1/2,-1/2}\|^2-2\sum_{n=-k(p)}^{k(p)}|c_{p,n}|^2\\
&\qquad\times\left[\sum_{l,l'}|c_{1/2,n}^{k(p)+l}|^2|c_{-1/2,n}^{k(p)+l'}|^2\cos t(a_{p,l}-a_{p,l'})\right]\\
&\leq 2\left[1-\cos t(a_{p,1/2}-a_{p,-1/2})\right]
\end{aligned}
\tag{3.12}
$$

where we have used the formulae (2.18) and (2.20). Substituting (3.12) into (3.11), we get an important inequality

$$
\begin{aligned}
&\sum_{m' \neq m} \|A_{m',m}\|^2 \\
&\geq \frac{1}{2} \sum_p |\lambda_p|^2 \|A_{p;1/2,1/2} - A_{p;-1/2,-1/2}\|^2 \\
&\qquad \times \sum_{n=-k(p)}^{k(p)} |c_{p,n}|^2 \left(1 - \frac{4n^2+1}{(2k(p)+1)^2}\right).
\end{aligned}
\tag{3.13}
$$

The author believes that the formulae (3.11) and (3.13) have some advantages over the inequalities of type (1.9), which would become clear in the following.

Firstly, let us look at one point of agreement between the two approaches. Since it is clear that

$$
\sum_{n=-k(p)}^{k(p)} |c_{p,n}|^2 n^2 \leq k(p)^2, \tag{3.14}
$$

we have from (3.13) that

$$
\begin{aligned}
&\sum_{m' \neq m} \|A_{m',m}\|^2 \\
&\geq \sum_p |\lambda_p|^2 \|A_{p;1/2,1/2} - A_{p;-1/2,-1/2}\|^2 \frac{2k(p)}{(2k(p)+1)^2}.
\end{aligned}
\tag{3.15}
$$

This formula indicates that we have to make the value of $k(p)$ sufficiently large for each p to make the measurement close to the ideal, because the value of

$$
\begin{aligned}
&\frac{1}{2} \sum_p |\lambda_p|^2 \|A_{p;1/2,1/2} - A_{p;-1/2,-1/2}\|^2 \\
&\quad = \frac{1}{2} \|A_{1/2,1/2} - A_{-1/2,-1/2}\|^2
\end{aligned}
\tag{3.16}
$$

should be almost unity for a close-to-ideal measurement process, i.e., for a process which almost satisfies the conditions (1.7) and (1.8). Now let us calculate the right-hand side of the inequality (1.9) for the case $L_A = S_x$:

$$
\begin{aligned}
&\langle A|{S_x}^2|A\rangle \\
&= \frac{1}{4}\sum_p |\lambda_p|^2 \| \sum_n c_{p,n}\sqrt{(k(p)-n)(k(p)+n+1)}\phi_{n+1}^{k(p)} \\
&\qquad + \sum_n c_{p,n}\sqrt{(k(p)-n+1)(k(p)+n)}\phi_{n-1}^{k(p)}\|^2 \\
&= \frac{1}{2}\sum_p |\lambda_p|^2 \Bigg[k(p)^2 + k(p) - \sum_n |c_{p,n}|^2 n^2 \\
&\quad + \mathrm{Re}\sum_n {c_{p,n-1}}^* c_{p,n+1}\sqrt{(k(p)+n)(k(p)-n)} \\
&\qquad\qquad \times \sqrt{(k(p)+n+1)(k(p)-n+1)}\Bigg].
\end{aligned}
\tag{3.17}
$$

Hence we can derive the same condition on $k(p)$ from inequality (1.9) ; both inequalities (1.9) and (3.13) contain the same information on this point.

Next we consider an initial state A given by

$$
A = \sum_p \lambda_p |k(p), 0\rangle, \tag{3.18}
$$

i.e., such that $c_{p,n} = 0$ for $n \neq 0$ in the expression (2.19). In this case Eq. (3.17) tells us that

$$
\langle A|{S_x}^2|A\rangle = \frac{1}{2}\sum_p |\lambda_p|^2 \left[k(p)^2 + k(p)\right], \tag{3.19}
$$

and therefore that the right-hand side of the inequality (1.9) can be made arbitrarily small by taking large $k(p)$. Furthermore, Eqs. (2.26) and (3.2) show us that

$$
\begin{aligned}
&\langle A_{1/2,1/2}|A_{-1/2,-1/2}\rangle \\
&\approx \frac{1}{2}\sum_p |\lambda_p|^2 \left[1 + \cos t(a_{p,1/2} - a_{p,-1/2})\right]
\end{aligned}
\tag{3.20}
$$

where the approximation is valid for sufficiently large $k(p)$. Since the right-hand side of (3.20) can be 0 for some constants $a_{p,l}$ and t, we can therefore conclude that there are some initial states A having the form (3.18) for which both the right-hand side of (1.9) and the value of $\langle A_{1/2,1/2}|A_{-1/2,-1/2}\rangle$ can be made arbitrarily small by taking large $k(p)$.

The formula (3.13), however, shows us that

$$\begin{aligned}
&\sum_{m'\neq m} \|A_{m',m}\|^2 \\
&\geq \frac{1}{2}\sum_{p} |\lambda_p|^2 \|A_{p;1/2,1/2} - A_{p;-1/2,-1/2}\|^2 \\
&\qquad\qquad \times \left(1 - \frac{1}{(2k(p)+1)^2}\right) \\
&\approx \frac{1}{2}\|A_{1/2,1/2} - A_{-1/2,-1/2}\|^2
\end{aligned} \tag{3.21}$$

for any initial state A having the form (3.18), where the approximation is valid for sufficiently large $k(p)$. Since the value of the last line of (3.21) has to be almost unity for a close-to-ideal measurement process, the formula (3.21) insists that we cannot make the measurement close to the ideal for the initial state A taken in (3.18) even if we make all values of $k(p)$ very large.

That is to say, there exist some initial states A for which we cannot make the measurement close to the ideal although both the right-hand side of (1.9) and the value of $\langle A_{1/2,1/2}|A_{-1/2,-1/2}\rangle$ can be made arbitrarily small. In this sense the inequality (1.9) gives us no insight into which initial state A we should choose to attain a close-to-ideal measurement process.

In contrast, if we fix a constant a such that $0 \leq a < 1$, then we can write (3.13) as :

$$
\begin{aligned}
&\sum_{m'\neq m} \|A_{m',m}\|^2 \\
&\geq \frac{1}{2}\sum_p |\lambda_p|^2 \|A_{p;1/2,1/2} - A_{p;-1/2,-1/2}\|^2 \\
&\qquad \times \sum_{|n|\leq ak(p)} |c_{p,n}|^2 \left(1 - \frac{4n^2+1}{(2k(p)+1)^2}\right) \\
&\geq \frac{1}{2}\sum_p |\lambda_p|^2 \|A_{p;1/2,1/2} - A_{p;-1/2,-1/2}\|^2 \\
&\qquad \times \left(1 - \frac{4a^2k(p)^2+1}{(2k(p)+1)^2}\right) \sum_{|n|\leq ak(p)} |c_{p,n}|^2 \\
&\approx \frac{1}{2}\left(1-a^2\right)\sum_p |\lambda_p|^2 \|A_{p;1/2,1/2} - A_{p;-1/2,-1/2}\|^2 \\
&\qquad \times \sum_{|n|\leq ak(p)} |c_{p,n}|^2
\end{aligned}
\tag{3.22}
$$

where again the approximation holds when all values of $k(p)$ are sufficiently large. The inequality (3.22) implies that we have to reduce the value of

$$\sum_{|n|\leq ak(p)} |c_{p,n}|^2$$

to almost 0 for each p to approximate an ideal measurement. Since the constant a is arbitrary, we can conclude the following : The initial state A must be prepared to be close to a state having the form

$$\tilde{A} = \sum_p \lambda_p \left[c_{p,k(p)}|k(p),k(p)\rangle + c_{p,-k(p)}|k(p),-k(p)\rangle\right] \tag{3.23}$$

(i.e., a state which satisfies the condition $c_{p,n} = 0$ for $n \neq \pm k(p)$ in the expression (2.19)) if we wish to make the measurement process close to the ideal. On the other hand, Eq. (3.11) shows us that this state $\tilde{A}$ is sufficient to make the value of $\sum_{m'\neq m} \|A_{m',m}\|^2$ arbitrarily small if we take all $k(p)$ very large.

That is to say, if we choose such an initial state A that both the right-hand side of (3.13) and the value of $\langle A_{1/2,1/2}|A_{-1/2,-1/2}\rangle$ are sufficiently small, then we can always attain a close-to-ideal measurement. In this sense the formulae (3.11) and (3.13) contain more information than the inequality (1.9) about which initial state A we should select to make an almost ideal measurement. This is the main distinguishing difference to the altenative inequality (1.9).

4. Concluding remarks

A measuring process has been considered in which a component of the spin angular momentum of a particle with spin 1/2 is measured under the constraint that all three components of the total angular momentum are conserved in the process. We have obtained some formulae for the lower bound of the term which corresponds to the disturbance the measurement causes to the measured system. Our formulae, in contrast with the traditional ones, yield information about how to select the initial state of the apparatus in order to make the measurement close to the ideal.

In this section we comment on the physical meaning of our results. We should note here that our calculations resort to the restricted form (2.5) of the Hamiltonian H to a great extent. If the Hamiltonian H can be expressed as a polynomial of the operators s_x, s_y, s_z, S_x, S_y, and S_z, then our calculations are still applicable. (Of course, H has to commute with all of the operators $s_x + S_x, s_y + S_y$ and $s_z + S_z$). But if it cannot be so expressed, then we are unable to guarantee the existence of the constants $a_{p,l}$ in (2.11), which play an essential role in our calculations.

The author's opinion is however that the assumption of (2.5) is physically reasonable. If we accept this assumption to be reasonable then we have to pay some attention to the difference between the conventional physical interpretation of the inequalities of type (1.9), which has been accepted by several authors, and the physical meaning of our formulae (3.11) and (3.13).

The conventional physical interpretation of the inequality (1.9) is as follows [14] : "For many observables, the ideal measurement process is merely an ideal concept in the sense that it can never be realized by any real process. But we can make a real measurement as close to the ideal as we want by taking a sufficiently large and complicated apparatus." Araki and Yanase [2] constructed an initial state by which they were able to make the measurement close to the ideal. The initial state constructed by them is a superposition of many eigenstates of the conserved quantity L_A, and seems to be an acceptable macroscopic state. It appears to be an easily prepared physically normal state if we assume a sufficiently large and complicated apparatus. This would appear to be the major reason why the factor of the initial state has been overlooked and only the factor of the size of apparatus has been emphasized in the conventional interpretation of (1.9).

But our formulae (3.11) and (3.13) suggest that we would have to prepare a rather restrictive initial state if we want to make the measurement process close to the ideal. The inequality (3.22) can be interpreted that : When the initial state of the apparatus is not like $\tilde{A}$ in (3.23), we cannot reduce the value of $\sum_{m' \neq m} ||A_{m',m}||^2$ to 0 even if the values of $k(p)$ are all very large. (Araki and Yanase discussed the measurement process in terms of the restriction imposed on it by the existence of *one* conserved quantity. In contrast, we discuss it under more restrictive condition (2.5). This would be the major reason why we have to prepare a rather restrictive initial state than the state constructed by Araki and Yanase). If such a state as $\tilde{A}$ is a physically normal state and if it can be easily prepared, then the results of this article pose no new problem and the conventional interpretation of (1.9) mentioned above survives without any change.

But it seems to the author that the actual preparation of such a state $\tilde{A}$ becomes more and more difficult as the values of $k(p)$ increase. We discussed, in Ref. 15), the case in which the Hilbert space $\mathcal{H}_A$ of the apparatus is identified with an irreducible representation space, i.e., the case of

$$\mathcal{H}_A = V_k \tag{4.1}$$

for some k. We proved that we have to prepare the initial state

$$\tilde{A} = 2^{-1/2}\left(e^{i\theta}|k,k\rangle + e^{i\theta'}|k,-k\rangle\right) \tag{4.2}$$

with very large k for a near-ideal measurement process. Here $|k,k\rangle = \phi^k_k$ is an extreme state in the sense that S_z takes its maximum value k, and $|k,-k\rangle = \phi^k_{-k}$ is another extreme state in the sense that S_z takes its minimum value $-k$. The state $\tilde{A}$ in (4.2) is seen to be a superposition of these extremes with the same probabilities. The author presumes that this would depart from a physically normal state as k increases, and that its actual preparation would become increasingly difficult. Relaxing the condition (4.1) to the general condition (2.8), we have obtained the expression (3.23) for the initial state, which is the generalization of (4.2).

So far, the author has no physical evidence to support the view that the state $\tilde{A}$ in (3.23) is rather peculiar and that the actual preparation of such a state becomes more and more difficult as the values of $k(p)$ increase. But if we accept this view, then we can summarize the results of this article as follows : We can make a measurement process as close to the ideal as we want by taking a sufficiently large and complicated apparatus, but this will be balanced by increasingly difficult preparation of the initial state of the apparatus.

Acknowledgements

I am very grateful to Drs. Kiyotaka Kakazu and Shoju Kudaka for many stimulating discussions.

References

1. E. P. Wigner, Z. Phys. **131** (1952), 101.
2. H. Araki and M. M. Yanase, Phys. Rev. **120** (1960), 622.
3. H. Stein and A. Shimony, in *Foundations of Quantum Mechanics, Proceedings of the Enrico Fermi International Summer School*, edited by B. d'Espagnat (Academic, New York, 1971).
4. J. von Neumann, *Die Mathematische Grundlagen der Quantenmechanik* (Springer Verlag, Berlin, 1932).
5. M. Ozawa, Phys. Rev. Lett. **67** (1991), 1956 ; in *Classical and Quantum Systems - Foundations and Symmetries, Proceedings of the Second International Wigner Symposium*, edited by H. D. Doebner, W. Scherer, F. Schroeck, Jr. (World Scientific, Singapore, 1992).
6. M. M. Yanase, Phys. Rev. **123** (1961), 666.
7. G. C. Ghirardi, T. Weber, F. Miglietta, and A. Rimini, *Lecture Notes in Physics,* **135** (Springer-Verlag, Berlin, 1980).
8. G. C. Ghirardi, F. Miglietta, A. Rimini, and T. Weber, Phys. Rev. D **24** (1981), 347, 353.
9. F. Crisciani, G. C. Ghirardi, A. Rimini, and T. Weber, Nuovo Cimento B **64** (1981), 338.
10. G. C. Ghirardi, A. Rimini, and T. Weber, J. Math. Phys. **23** (1982), 1792; **24** (1983), 2454.
11. A. Messiah, *Mécanique Quantique* (Dunod, Paris, 1959).
12. E. P. Wigner, *Group Theory and Its Application to the Quantum Mechanics of Atomic Spectra* (Academic Press, New York, London, 1959).
13. G. Racah, Phys. Rev. **61** (1942), 186; **62** (1942), 438.

14. B. d'Espagnat, *Conceptual Foundation of Quantum Mechanics* (W.A.Benjamin, Inc., 1976).
15. S. Matsumoto, Prog. Theor. Phys. **90** (1993), 35.

Department of Mathematics
College of Education
University of the Ryukyus
Nishihara-cho, Okinawa 903-01
Japan