Moments of the Wigner Distribution and a Generalized Uncertainty Principle

R. Simon\textsuperscript{1} and N. Mukunda\textsuperscript{2}

\textsuperscript{1}The Institute of Mathematical Sciences, C.I.T. Campus, Madras 600 113, India
\textsuperscript{2}Centre for Theoretical Studies, Indian Institute of Science, Bangalore 560 012, India and Jawaharlal Nehru Centre for Advanced Scientific Research, Jakkur, Bangalore 560 064, India

(June 22, 2004)

Abstract

The nonnegativity of the density operator of a state is faithfully coded in its Wigner distribution, and this places constraints on the moments of the Wigner distribution. These constraints are presented in a canonically invariant form which is both concise and explicit. Since the conventional uncertainty principle is such a constraint on the first and second moments, our result constitutes a generalization of the same to all orders. Possible application in quantum state reconstruction using optical homodyne tomography is noted.

PACS numbers: 03.65.Bz, 42.50.Ar, 42.50.Wm, 42.50.Lc
The uncertainty principle exhibits a fundamental manner in which the quantum description of nature departs from the classical one. For the canonical pair of variables $\hat{q}, \hat{p}$ the Heisenberg commutation relation $[\hat{q}, \hat{p}] = i\hbar$ leads, for any state $|\psi\rangle$, to the unbeatable limitation

$$\langle (\Delta \hat{q})^2 \rangle \langle (\Delta \hat{p})^2 \rangle - \left\langle \frac{\Delta \hat{q} \Delta \hat{p} + \Delta \hat{p} \Delta \hat{q}}{2} \right\rangle \geq \frac{\hbar^2}{4},$$

(1)

where $\langle \hat{q} \rangle = \langle \psi | \hat{q} | \psi \rangle$, $\Delta \hat{q} = \hat{q} - \langle \hat{q} \rangle$, and so on. Every Gaussian pure state saturates this inequality. An important attribute of the uncertainty principle (1) is that it is invariant under all real linear canonical transformations, just as the canonical commutation relation is.

This inequality can be generalized in a naive manner to higher orders in $\hat{q}, \hat{p}$. For any pair of hermitian operators $\hat{A}, \hat{B}$ and state $|\psi\rangle$ we have the Schwartz inequality

$$\langle \hat{A}^2 \rangle \langle \hat{B}^2 \rangle \geq \left\langle \frac{\hat{A} \hat{B} + \hat{B} \hat{A}}{2} \right\rangle^2 + \left\langle \frac{\hat{A} \hat{B} - \hat{B} \hat{A}}{2i} \right\rangle^2.$$  

(2)

It is saturated if an only if $\hat{A} |\psi\rangle$ and $\hat{B} |\psi\rangle$ are linearly dependent as vectors. Clearly, (1) is a particular case of (2) corresponding to $\hat{A} = \hat{q} - \langle \hat{q} \rangle$, $\hat{B} = \hat{p} - \langle \hat{p} \rangle$. Clearly, the choice $\hat{A} = \hat{q}^2 - \langle \hat{q}^2 \rangle$, $\hat{B} = \hat{p}^2 - \langle \hat{p}^2 \rangle$ will lead to a higher order uncertainty principle involving $\langle \hat{q}^4 \rangle$, $\langle \hat{p}^4 \rangle$; the Fock states $|n\rangle$, being eigenstates of $\hat{q}^2 + \hat{p}^2$, will be expected to saturate this higher order uncertainty principle. That they indeed do so can be explicitly verified.

One may indeed produce any number of such naive generalized uncertainty principles by making various choices for $\hat{A}, \hat{B}$ in (2). But every one of them will suffer from the deficiency of not being invariant under linear canonical transformations. Further, there seems to be no reasonable sense in which the set of all such generalizations based on (2) can be considered to be complete.

The purpose of this Letter is to present a generalization of the uncertainty principle which largely overcomes these difficulties. This is achieved by applying to the Wigner quasiprobability concepts and results from the classical problem of moments. The final result is a nested sequence of constraints on the moments of the Wigner distribution. These constraints are tailored to capture the positivity of the density operator of a quantum state. Equivalently, a given real phase space distribution has to necessarily meet these constraints in order to qualify to be a bonafide Wigner distribution.

It should be appreciated that the higher moments of the Wigner distribution are no more objects of purely academic interest. An enormous progress in quantum state reconstruction using optical homodyne tomography has been achieved in the last few years: the Wigner distribution of a state can now be fully mapped out, as has been demonstrated by several groups.

There exist rigorous and mathematically sophisticated approaches to the quantum mechanical moment problem. But our considerations here are explicit and take full advantage of the canonical invariance underlying the Heisenberg commutation relation.

Details of a classical probability density $\rho(x)$ are coded in its moments $\gamma_n = \int dx x^n \rho(x)$. An important result in the problem of moments is this: given a sequence of numbers it qualifies to be the moment sequence of a bonafide probability distribution if and only if the symmetric matrix defined below is nonnegative:
\[ \Gamma = \begin{pmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \cdots \\ \gamma_1 & \gamma_2 & \gamma_3 & \cdots \\ \gamma_2 & \gamma_3 & \gamma_4 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \geq 0. \tag{3} \]

This can be broken into a sequence of positivity conditions on the determinants of the submatrices of \( \Gamma \), which in turn can be viewed as a nested sequence of constraints on the moments \( \gamma_n \); and these constraints are tailored to capture the pointwise nonnegativity of \( \rho(x) \). Reconstruction of \( \rho(x) \) from its moment sequence is the other part of the classical problem of moments [2].

In quantum mechanics, the state is described not by a true probability density in phase space, but by one of several possible quasiprobabilities [1]. The earliest, and probably the most prominent, quasiprobability is the one introduced by Wigner [1]. It is intimately related to phase space variables and the algebra \( \hat{A} \) of operator valued functions \( \hat{F}(\hat{q}, \hat{p}) \) of the canonical operators. The rule is specified first through the one to one correspondence \( e^{i\theta \hat{q} + \tau \hat{p}} \leftrightarrow e^{i\theta \hat{q} + \tau \hat{p}} \) for plane waves, and then extended linearly to the entire algebra using Fourier techniques.

The Weyl rule could equally well be specified in the monomial basis instead of the plane wave basis through the association \( q^m p^n \leftrightarrow \hat{T}_{m,n} \) for \( m, n = 0, 1, 2, \cdots \) where the Weyl ordered monomial \( \hat{T}_{m,n} \) is the coefficient of \((m!n!)^{-1} \theta^m \tau^n \) in the Taylor expansion of \( e^{i\theta \hat{q} + \tau \hat{p}} \). This is an isomorphism between \( \mathcal{A} \) and \( \hat{\mathcal{A}} \) only at the level of vector spaces but not at the level of algebras. In particular, the product of two \( \hat{T}_{m,n} \)'s is not another monomial but a linear combination of monomials [1]:

\[
\hat{T}_{m,n} \hat{T}_{m',n'} = \sum_{r,s} d_{r,s} \hat{T}_{m+m'-r-s,n+n'-r-s},
\]

\[
d_{r,s} = \frac{(-1)^r \left(\frac{i\theta}{2}\right)^{s+r} m! n!}{(m-s)! (n-r)!} \binom{m'}{r} \binom{n'}{s}. \tag{4}\]

The intimate connection between Weyl ordering and Wigner distribution is this:

\[
\text{tr}(\hat{\rho} \hat{T}_{m,n}) = \int dq \, dp \, q^m p^n W(q, p). \tag{5}\]

That is, the quantum mechanical expectation of the Weyl ordered monomial \( \hat{T}_{m,n} \) is precisely the \( mn \)-th moment of the Wigner function. By linearity, similar relation holds for any pair \( f(q, p), \hat{F}(\hat{q}, \hat{p}) \) related by Weyl ordering.

The monomials \( \hat{T}_{m,n} \) are hermitian, and transform in a simple manner under the group \( \text{Sp}(2, \mathbb{R}) \) of real linear canonical transformations. This group can be identified with \( \text{SL}(2, \mathbb{R}) \), the group of \( 2 \times 2 \) real matrices with unit determinant. \( \text{Sp}(2, \mathbb{R}) \) acts identically on the pairs \( (q, p) \) and \( (\hat{q}, \hat{p}) \), and this action induces linear transformation in the algebras \( \mathcal{A} \) and \( \hat{\mathcal{A}} \) in the natural manner.

The set of homogeneous polynomials of order \( 2j \) in \( q \) and \( p \) (being linear combinations of \( q^{-s} p^{j-s} \) for \( s = -j, -j + 1, \cdots, j \) ) transform linearly among themselves under this transformation, leading to the spin-\( j \) representation of \( \text{Sp}(2, \mathbb{R}) \) in \( \mathcal{A} \). The \( \hat{T}_{m,n} \)'s in \( \hat{\mathcal{A}} \) transform in the same manner as the \( q^m p^n \)'s in \( \mathcal{A} \), and thus the vector space \( \hat{\mathcal{A}} \) decouples.
into a direct sum of invariant subspaces under $\text{Sp}(2, \mathbb{R})$: $\hat{A} = \hat{V}^{(0)} \oplus \hat{V}^{(\frac{j}{2})} \oplus \hat{V}^{(1)} \oplus \cdots$. Clearly, $\hat{V}^{(j)}$ is of dimension $2j + 1$, and is spanned by $\hat{\xi}_{js} = \hat{T}_{j-s,j+s}$ with $s$ running over the range $s = -j, -j+1, \cdots, j$. It acts as the carrier space for the spin-$j$ representation of $\text{Sp}(2, \mathbb{R})$ in $\hat{A}$. Thus, every spin-$j$ representation of $\text{Sp}(2, \mathbb{R})$ occurs in $\hat{A}$ once and only once.

It is convenient to arrange the $\hat{\xi}_{js}$’s for fixed $j$ into a $2j + 1$ dimensional column vector $\hat{\xi}^{(j)}$ and then, for any chosen $J$, arrange these columns into a grand column vector $\hat{\xi}_J$ of dimension $(J+1)(2J+1)$.

Let the $(2j + 1) \times (2j + 1)$ matrix $K^{(j)}(S)$ denote the spin-$j$ representation for $S \in \text{Sp}(2, \mathbb{R})$. Since the defining representation of $\text{Sp}(2, \mathbb{R})$ is the spin-$\frac{1}{2}$ representation, we have $K^{(\frac{1}{2})}(S) = S$. Let $\hat{K}_J(S)$ be the block diagonal matrix of order $(J + 1) \times (2J + 1)$ with diagonal blocks $K^{(0)}(S) = 1, K^{(\frac{1}{2})}(S), \cdots, K^{(J)}(S)$. Then the action of $\text{Sp}(2, \mathbb{R})$ in $\hat{A}$ has the concise description

$$\hat{\xi}_J \longrightarrow \hat{K}_J(S)\hat{\xi}_J, \quad \hat{\xi}^{(j)} = K^{(j)}(S)\hat{\xi}^{(j)}.$$

We are now in a position to present the generalized uncertainty principle. For each $J = 0, \frac{1}{2}, 1, \cdots$ form the square matrix $\hat{\Omega}_J$, of order $(J + 1)(2J + 1)$, with operator entries, through the definition ($\hat{\xi}_J^\dagger$ is a row vector with the same entries as the column vector $\hat{\xi}_J$)

$$\hat{\Omega}_J = \hat{\xi}_J \hat{\xi}_J^\dagger, \quad (\hat{\Omega}_J)_{js,j's'} = \hat{\xi}_{js} \hat{\xi}_{j's'}.$$

We may write $\hat{\Omega}_J$ in more detail in the block form

$$\hat{\Omega}_J = \begin{pmatrix} 1 & \hat{\xi}^{(\frac{1}{2})\dagger} & \cdots & \hat{\xi}^{(J)\dagger} \\ \hat{\xi}^{(\frac{1}{2})} & \hat{\xi}^{(\frac{1}{2})\dagger} & \cdots & \hat{\xi}^{(\frac{1}{2})(J)\dagger} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\xi}^{(J)} & \hat{\xi}^{(J)\dagger} & \cdots & \hat{\xi}^{(J)(J)\dagger} \end{pmatrix}.$$

It is to be understood that each element of $\hat{\Omega}_J$ is written as a linear combination of the $\hat{T}_{m,n}$’s using (4). For purpose of illustration, we detail one of these blocks:

$$\hat{\xi}^{(1)} \hat{\xi}^{(\frac{1}{2})\dagger} = \begin{pmatrix} \hat{T}_{3,0} & \hat{T}_{2,1} + i\hbar \hat{T}_{1,0} \\ \hat{T}_{2,1} - i\hbar \hat{T}_{1,0} & \hat{T}_{1,2} + i\hbar \frac{1}{2} \hat{T}_{0,1} \\ \hat{T}_{1,2} - i\hbar \frac{1}{2} \hat{T}_{0,1} & \hat{T}_{0,3} \end{pmatrix}.$$

Let $M_J = \langle \hat{\Omega} \rangle$ be the hermitian $c$-number matrix obtained from $\hat{\Omega}_J$ by taking (entrywise) quantum mechanical expectation value in the given state $\hat{\rho}$:

$$M_J = \text{tr}(\rho \hat{\Omega}_J) = \langle \hat{\xi}_J^\dagger \hat{\xi}_J \rangle; \quad (M_J)_{js,j's'} = \text{tr}(\hat{\rho} \hat{\xi}_{js} \hat{\xi}_{j's'})$$

(8)
It will prove useful to write $M_J$ in the block form

$$
M_J = \begin{pmatrix}
1 & M^{0,J} & \cdots & M^{0,J} \\
M^{1,0} & M^{1,1} & \cdots & M^{1,J} \\
\vdots & \vdots & \ddots & \vdots \\
M^{J,0} & M^{J,1} & \cdots & M^{J,J} 
\end{pmatrix},
$$

(9)

where $M^{j,j'} = \langle \tilde{\xi}^{(j)}(\tilde{\xi}^{(j')})^\dagger \rangle$ is a $(2j+1) \times (2j'+1)$ dimensional block, and $M^{j',j} = (M^{j,j'})^\dagger$.

Since $\tilde{\xi}^{(0)} = 1$, $M^{0,0} = 1$ for all states. For purpose of illustration, we write out a few leading blocks of $M_J$ explicitly: it is clear that the row vectors $M^{0,\frac{1}{2}}$ and $M^{0,1}$ have entries $(\bar{q}, \bar{p})$ and $(\bar{q}^2, \bar{q}\bar{p}, \bar{p}^2)$ respectively; further

$$
M^{\frac{1}{2},\frac{1}{2}} = \begin{pmatrix}
\bar{q}^2 & \bar{q}\bar{p} + \frac{i\hbar}{2} \\
\bar{q}\bar{p} - \frac{i\hbar}{2} & \bar{p}^2
\end{pmatrix},
$$

and finally, the $3 \times 3$ hermitian block $M^{1,1} = (M^{1,1})^\dagger$ has the form

$$
\begin{pmatrix}
\bar{q}^4 & \bar{q}^3\bar{p} + \frac{i\hbar}{2} & \bar{q}^2\bar{p}^2 + 2i\hbar\bar{q}\bar{p} - \frac{\hbar^2}{2} \\
\bar{q}^3\bar{p} - \frac{i\hbar}{2} & \bar{q}^2\bar{p}^2 + \frac{\hbar^2}{4} & \bar{q}\bar{p}^3 + i\hbar\bar{p}^2 \\
\bar{q}^2\bar{p}^2 - 2i\hbar\bar{q}\bar{p} - \frac{\hbar^2}{2} & \bar{q}\bar{p}^3 - i\hbar\bar{p}^2 & \bar{p}^4
\end{pmatrix}.
$$

Here, $\langle \bar{q}^m\bar{p}^n \rangle$ stands for the average of $q^m p^n$ with the Wigner distribution as the weight as in (8). In other words, $M_J$ is the matrix formed out of the moments of the Wigner distribution function, of order atmost $2J$.

We now prove the important fact that the nonnegativity of the density operator $\hat{\rho}$ forces the hermitian matrix $M_J$ to be a nonnegative matrix, for every $J$. For a given fixed value of $J$ consider the operator

$$
\hat{\eta} = \sum_{j=0}^{J} \sum_{s=-j}^{j} c_{js} \hat{\xi}_{js},
$$

where $c_{js}$ are arbitrary $c$-number expansion coefficients which can be arranged into a $(J+1)(2J+1)$ dimensional column vector $C$. Now form the operator

$$
\hat{\zeta} = \hat{\eta}^\dagger \hat{\eta} = \sum_{j,s} \sum_{j',s'} c_{js}^* c_{j's'} \hat{\xi}_{js} \hat{\xi}_{j's'},
$$

(10)
which is hermitian nonnegative by construction. Since \( \hat{\rho} \geq 0 \), we necessarily have \( \text{tr}(\hat{\rho} \hat{\zeta}) \geq 0 \), for every choice of the coefficients \( \{c_{j,s}\} \). But from (9), (10) we find

\[
\text{tr}(\hat{\rho} \hat{\zeta}) = \sum_{j,s} \sum_{j',s'} c_{j,s}^* c_{j',s'} M_{j,s,j',s'}.
\]

That is, \( \text{tr}(\hat{\rho} \hat{\zeta}) = C^\dagger M C \) for every \( C \). This completes the proof that \( \hat{\rho} \geq 0 \) implies \( M_J \geq 0 \) for every \( J \).

A little reflection should convince the reader that this is the generalized form of the uncertainty principle we have been after, and we state it as follows:

**Generalized Uncertainty Principle:** Let \( M_J \) be the hermitian \( \epsilon \)-number matrix formed out of the moments of the Wigner distribution of a state \( \hat{\rho} \) in accordance with the prescription (8). Then

\[
M_J \geq 0, \quad J = 0, \frac{1}{2}, 1, \cdots
\]  

(11)

For a given state not all moments will exist in general. It is clear that in such a case where \( M_J \) is finite only for all \( J \leq J_{\text{max}} \), our generalized uncertainty principle should be modified to read \( M_J \geq 0, \quad J = 0, \frac{1}{2}, \cdots, J_{\text{max}} \).

While the hermiticity and unit trace properties of \( \hat{\rho} \) are reflected in the reality and normalization of the Wigner distribution, the generalized uncertainty principle presented in the concise matrix form (11) exhibits the constraints on the moments \( q_m \) of the Wigner distribution resulting from the nonnegativity of \( \hat{\rho} \). While the conventional uncertainty principle is such a constraint on the first and second moments, ours is a generalization to all orders. It should be appreciated that the canonical commutation relation enters \( M_J \) in (11) through (4).

The following mathematical lemma is helpful in analyzing the content of this generalized uncertainty principle: A hermitian matrix \( Q \) of the block form

\[
Q = \begin{pmatrix} A & C^\dagger \\ C & B \end{pmatrix}
\]

is positive definite if and only if \( A \) and \( B - CA^{-1}C^\dagger \) are positive definite. The proof simply consists in recognizing the congruence

\[
Q \sim Q' = LL^\dagger, \quad L = \begin{pmatrix} 1 & 0 \\ -CA^{-1} & 1 \end{pmatrix},
\]

where \( Q' \) is a block diagonal matrix with diagonal blocks \( A \) and \( B - CA^{-1}C^\dagger \).

The usual uncertainty principle (1) is contained in (11) as a particular case: it is equivalent to the condition \( \det M = 0 \geq 0 \). Next consider the case \( J = 1 \). Use of the lemma with \( C^\dagger = (M^{0,\frac{1}{2}} M^{0,1}) \) renders \( M_J \sim M'_J \), where

\[
M'_J = \begin{pmatrix} 1 & 0 & 0 \\ 0 & M^{\frac{1}{2},\frac{1}{2}} - M^{\frac{1}{2},0} M^{0,\frac{1}{2}} & M^{\frac{1}{2},1} - M^{\frac{1}{2},0} M^{0,1} \\ 0 & M^{1,\frac{1}{2}} - M^{1,0} M^{0,\frac{1}{2}} & M^{1,1} - M^{1,0} M^{0,1} \end{pmatrix}.
\]
Now $M_I \geq 0$ implies $M'_I \geq 0$ which in turn implies that its diagonal block $M^{I,\frac{1}{2}} - M^{I,0} M^{0,\frac{1}{2}} \geq 0$. Written in terms of the moments, the last condition reads

$$\begin{pmatrix}
  \overline{q^2} - \overline{q}^2 & \overline{qp} - \overline{q} \overline{p} + \frac{i\hbar}{2} \\
  \overline{qp} - \overline{q} \overline{p} - \frac{i\hbar}{2} & \overline{p^2} - \overline{p}^2
\end{pmatrix} \geq 0,$$

(12)

which is precisely the usual uncertainty principle (1).

One more application of the lemma on the nontrivial part of $M'_I$ further strengthens the positivity requirement on the other diagonal block $M^{1,1} - M^{1,0} M^{0,1}$ to

$$M^{1,1} - M^{1,0} M^{0,1} \geq C \left( M^{1,\frac{1}{2}} - M^{I,0} M^{0,\frac{1}{2}} \right)^{-1} C^\dagger,$$

$$C = \left( M^{1,\frac{1}{2}} - M^{1,0} M^{0,\frac{1}{2}} \right).$$

(13)

This $3 \times 3$ matrix condition, together with the $2 \times 2$ matrix condition (12), constitutes a complete statement of the generalised uncertainty principle involving moments of all order up to and including the fourth.

It is clear that yet another application of the lemma, starting with $M_I = \frac{3}{2}$, will lead to a positivity statement on a $4 \times 4$ matrix which, together with (12) and (13), will constitute a complete statement of our uncertainty principle on moments of all orders up to and including the sixth. Evidently, this reduction algorithm based on the above lemma can be continued to any desired value of $J$, and hence up to any desired (even) order of the moments, eventually rendering $M_I$ block diagonal.

We see from (6), (8) that $M_J$ transforms in the following manner under $S \in \text{Sp}(2,\mathbb{R})$:

$$S : \quad M_J \longrightarrow K_J(S) M_J K_J(S)^T.$$  

(14)

The nonnegativity of $M_J$ is manifestly preserved under this transformation. Thus, our generalized uncertainty principle is invariant under linear canonical transformations. Further, the reduction algorithm suggested by the lemma is invariant under linear canonical transformations, for it follows from (6) and (14) that $M^{i,j}$ transforms to $K^{i}(S) M^{i,j} K^{j}(S)^T$ under $S \in \text{Sp}(2,\mathbb{R})$.

An evidently useful way of reading (14) is that the components of $M_J$, just as the $\hat{T}_{m,n}$’s, transform as tensors under Sp(2, $\mathbb{R}$). And the fact that our generalized uncertainty principle is invariant under Sp(2, $\mathbb{R}$) means that it is implicitly stated in terms of the invariants of these tensors. These invariants, in the classical case, have been studied in great detail by Dragt and coworkers [10].

While the nonnegativity of $\hat{\rho}$ implies the nonnegativity of $M_J$ for all $J$, it is of interest to know if nonnegativity of $M_J$ for all $J$ implies nonnegativity of $\hat{\rho}$. Phrasing it somewhat differently, we may ask: Given a real normalized phase space distribution whose moments satisfy the condition $M_J \geq 0$, for all $J$, does it follow that the phase space distribution is a bonafide Wigner distribution?

From the very construction of $M_J$, it is clear that $\text{tr}(\hat{\rho} \hat{O}) \geq 0$ when $\hat{O}$ is of the form $\hat{\zeta}$ in (11). By linearity, this is true also when $\hat{O}$ is a (convex) linear combination of operators of this type (with nonnegative coefficients). Thus, (11) will be sufficient to characterise the Wigner distribution if the set of all such convex combinations is dense in the space of
nonnegative operators. Intuitively, this may appear to be the case. However, the monomials $T_{m,n}$ are generically noncompact, and hence a careful analysis of the issue of convergence should be made before one can make any claim in this direction.

We have already referred to the intensity with which current experimental research dealing with measurement of the Wigner distribution is being pursued [4–7]. Since measurements are always accompanied by errors of various origins, it will be of interest to see to what extent the Wigner distribution reconstructed in a real experiment respects the generalised uncertainty principle. Further, it may be of interest to examine the possibility of incorporating these fundamental inequalities in the algorithm for tomographically reconstructing the Wigner distribution from measured data, in such a way as to improve the reconstruction itself. Finally, our analysis applies equally well to any other quasiprobability, provided we choose suitably ordered monomials and modify the product formula (4) accordingly.
REFERENCES


