About Heisenberg Uncertainty Relation

By E. Schrödinger

Proceedings of The Prussian Academy of Sciences
Physics-Mathematical Section. 1930. XIX, pp.296-303

Abstract

The original Schrödinger’s paper is translated and annotated in honour of the 70-th anniversary of his Uncertainty Relation [will appear in Bulg.J.Phys.]. In the annotation it is shown that the Uncertainty Relation can be written in a complete compact canonical form.

ANNOTATION by A. Angelow ++, M.-C. Batoni +++

The main reason to publish the original Schrödinger’s paper in English, is the fact that no one of the books on Quantum Mechanics cites it (see for example [1-15]). Actually, Schrödinger’s paper is chiefly based on the notes of the seminars of Physics-Mathematical Section of The Prussian Academy, where many famous physicists worked to establish the underlying basis of Quantum Theory. Being a kind of internal report, this work remained, for many years at a certain marginal distance from the physicist scientific awareness.

Another argument in favour of its oblivion concerns the enthusiastic discussions, mostly about the physical interpretation of the uncertainty principle, rather than its mathematical straightforward derivation. After Schrödinger, the very first appearance of the new uncertainty relation occurs in the book of Merzbacher. However, he did not pay any attention to the new term (the covariance) and directly derives the Heisenberg inequality [11]. The same embarrassment appears in [12] and [13].

Fortunately, over the last years the scenario changed: for example, in the field of Quantum Optics, which was advocated to demonstrate the fundamental limit of Quantum Theory with understanding underlying on that new term, two monographs [14,15] were published, but the authors missed to cite the original Schrödinger’s paper.

Schrödinger’s work, originally written in German, was translated only in russian by A. Rogali [16], in 1976. However, we would like to emphasize that the sentence (*), written in the eleventh line after equation (11), is wrongly translated and it contradicts to the end of the same paragraph [16]. In that paragraph Schrödinger established classes of states with non-vanishing covariance.

Finally, we would like to demonstrate the essential contribution of Schrödinger’s inequality presenting it in a new compact form, using a modern terminology from mathematics, rather than that used in 1930. Let us take the three independent second-order central moments of the joint quantum distribution of two variables A and B, which are of special interest and warrant a special notation (see, for ex. [17,18]):

\[
E[(A - E[A])^2] \equiv (A - \overline{A})^2 \equiv \text{Var}[A] \equiv \Delta(A)^2 \\
E[(B - E[B])^2] \equiv (B - \overline{B})^2 \equiv \text{Var}[B] \equiv \Delta(B)^2
\]  

(1)

(2)
\[
E \left[ \frac{(A - E[A])(B - E[B]) + (B - E[B])(A - E[A])}{2} \right] \equiv \text{Cov}[A, B] \equiv \Delta(A, B) 
\]

where \(E[\cdot]\) means the expectation value. Obviously, the fourth second-order moment is \(\text{Cov}[B, A] = \text{Cov}[A, B]\), (respectively \(\Delta(B, A) = \Delta(A, B)\)). Note that when two observables \(X\) and \(Y\) don’t commute the correct expression for their product is not \(XY\), but the symmetrized one \(\frac{XY + YX}{2}\), and we conclude that the covariance is equivalent to the new term in the Schrödinger’s inequality:

\[
\text{Cov}[A, B] = \frac{AB + BA}{2} - AB 
\]

Let us construct the so-called \([19]\) covariance matrix (keeping in mind the non-commutativity of the observables in contrast to \([19]\)):

\[
\sigma[A, B] \equiv \begin{pmatrix} \text{Var}[A] & \text{Cov}[A, B] \\ \text{Cov}[B, A] & \text{Var}[B] \end{pmatrix} \equiv \begin{pmatrix} \Delta(A)^2 & \Delta(A, B) \\ \Delta(B, A) & \Delta(B)^2 \end{pmatrix} 
\]

Now, we can write the Schrödinger Uncertainty Principle (see below-eq.(9)) in the canonical form

\[
\det(\sigma[A, B]) \geq \frac{1}{4} |[A, B]|^2, \quad \text{or, for position and momentum} \quad \det(\sigma[q, p]) \geq \frac{\hbar^2}{4}, 
\]

and it is easy to see that the uncertainty relation is invariant under the rotation transformation in the phase space, while the Heisenberg one is not. We would like to emphasize that the new term in the inequality also plays an important role in the method of linear invariants in Quantum Mechanics, where the covariance is expressed in terms of the solution of the equation of a non-stationary two-dimensional harmonic oscilator \([20]\).

Translated and annotated in honour of 70th anniversary of Schrödinger Uncertainty Relation.
Sofía, January 1999

PACS: 03.65.-w

\[\text{§1. Recently E. U. Condon and H. P. Robertson [1] took into consideration the generalization of the fundamental principle of the quantum mechanics - that of the uncertainty - over an arbitrary canonical non-conjugate couple of physical variables. Trying to reach the same, I arrived at a slightly wider generalization than the Robertson’s one, which is, in fact stronger than the original Heisenberg inequality.}
\]

First of all, let us set out what is well known. The state-of-the-art of the “interpretation question” is the following: the test domain is a single specific physical system. The bases for the system knowledge that we dispose of - the catalogue of all that we can assert about the system - is equivalent to a complex function \(\Psi\) in the coordinate space of the system (it changes in a regular manner in time, but is not important at the moment). The mathematical correlate of a “physical variable”, i.e. of a very specific measurement that one might apply to the system, is a very specific linear Hermitean operator that from each \(\Psi\)-function produces an other such a \(\Psi\)-function. One can calculate the expectation
value of the respective measurement from the measure operator, say $A$, and the given $\Psi$-function:

$$\overline{A} = \int \Psi^* A \Psi dx$$

(1)

($\Psi^*$ is the complex conjugate, the integration goes over the whole coordinate space; given that $\Psi$ is constantly normalized, i.e. $\int \Psi^* \Psi dx = 1$).

The meaning of the expectation value is: mean value by unlimited number of measurements, while one must be sure that the system state is the same before each measurement, not changed by the measurement itself. In general, all possible statements one can make about the system are encoded in the expectation values. Moreover, one should keep in mind that it is up to us to choose the marking of reference scale of our measurement instrument. We can, for example, to set a value one only to one scale division and zero to all other. There is a specific operator attached to this “measurement” – one could name it as operator in blinkers$^{11}$, V. Neumann named it unitary operator. The respective expectation is obviously nothing else than the probability of the corresponding measurement value or measurement value intervals. The $\Psi$-function determines also the total measurement statistics.

The average error or the mean uncertainty of the value, which belongs to the operator $A$, is defined as

$$\Delta A = \sqrt{(A - \overline{A})^2} = \sqrt{A^2 - (\overline{A})^2}$$

(2)

(where in the first of the two expressions $\overline{A}$ should be more precise: $\overline{A}$ multiplied by the identical operator.) It may be proven, that this definition is not only formally constructed according to the theory of errors, but $\Delta A$ is really the average error of the variable $A$, when the statistics is defined in the above given way.

To prove now, that the product of the uncertainties of two random variables $A$ and $B$ satisfies the Heisenberg or yet more precise inequality we need to denote the following mathematical statements:

1. the Hermitean character of $A$ implies that the expectation value (1) is constantly real;

2. for each Hermitean operator it holds

$$\int f A g dx = \int g A^* f dx,$$

(3)
i.e., it could be rolled over on the other factor in such an integral, in such case the operator transforms into its conjugate form [2];

3. the product of two Hermitean operators is in general not Hermitean, but it could be split into “symmetrical product” and its (half) commutator:

$$AB = \frac{AB + BA}{2} + \frac{AB - BA}{2}$$

(4)
The first term is Hermitean, the last one is “skew Hermitean”, i.e. it becomes Hermitean multiplying by $i = \sqrt{-1}$. The splitting in many aspects corresponds to the splitting of a random (complex)$^{11}$ number into real and imaginary part. Immediately from this one
might extract the splitting of the expectation value into real and imaginary parts. The “expectation value” of every commutator is pure imaginary.

4. Finally, we need the so-called Schwartz inequality [3]

\[ (a_1 a_1^* + a_2 a_2^* + \ldots + a_n a_n^*)(b_1 b_1^* + b_2 b_2^* + \ldots + b_n b_n^*) \geq |a_1 b_1 + a_2 b_2 + \ldots + a_n b_n|^2, \] (5)

that we will apply in a limiting case on the continuous range of values of both functions \(f\) and \(g\) in the coordinate space:

\[ \int f f^* dx \cdot \int g g^* dx \geq \left| \int f g dx \right|^2. \] (5')

We assume here that specially

\[ f = B \Psi \quad g = A^* \Psi^*, \] (6)

where \(A\) and \(B\) are some Hermitean operators and \(\Psi\) is an arbitrary wave function, i.e. an arbitrary continuous and normalized function in the coordinate space. Using the equation (3) one obtains

\[ \int \Psi^* B^2 \Psi dx \cdot \int \Psi^* A^2 \Psi dx \geq \left| \int \Psi^* A B \Psi dx \right|^2, \] (7)

i.e., in terms of the notation (1)

\[ \overline{A^2} \cdot \overline{B^2} \geq |\overline{A B}|^2. \] (7')

When we decompose the right hand side according to (4), because of the linked note with this splitting we get

\[ \overline{A^2} \cdot \overline{B^2} \geq \left( \frac{AB + BA}{2} \right)^2 + \left( \frac{AB - BA}{2} \right)^2. \] (8)

This is already the inequality that we need to proof, but only in the special case when \(\overline{A}\) and \(\overline{B}\) vanish. In order to arrive at the general case, one should apply (8) and instead of the operators \(A\) and \(B\) rather use the following

\[ A - \alpha 1 \quad \text{and} \quad B - \alpha 1. \]

First of all \(\alpha\) and \(\beta\) must be arbitrary real constants, \(\alpha 1\) is the (unitary)† operator multiplied by \(\alpha\). The resulting inequality is therefore valid: 1. for an arbitrary \(\Psi\) 2. for every real pair of constants \(\alpha, \beta\). Therefore, there is no limitation on the \(\Psi\)-function to influence the choice of the pair of constants and especially to set

\[ \alpha = \overline{A}, \quad \beta = \overline{B}. \]

Finally, we end up with:

\[ (\Delta A)^2 (\Delta B)^2 \geq \left( \frac{AB + BA}{2} - \overline{A B} \right)^2 + \left( \frac{AB - BA}{2} \right)^2. \] (9)

This is the final form. The first from the two addends on the right hand side is new (to the best of my knowledge). (Without that term the inequality stands as the one of H. P.
Robertson.) So, the inequality links together three quantities: 1. the product of the mean deviations squared, 2. the absolute value squared of half of the mean value of the commutator, 3. a quantity which could be defined as a square of the mean deviations-product (the covariance)\(^\dagger\) in the condition that non-commutability is taken into account, i.e. the mean deviations-product must be define as the arithmetic mean of

\[
(A - \overline{A})(B - \overline{B}) \quad \text{and} \quad (B - \overline{B})(A - \overline{A})
\]  

which are the “mixed” expressions (≡ covariances, see eq.(3\(^\dagger\)))\(^\dagger\), completely analogous to \((\Delta A)^2\) and \((\Delta B)^2\) \(^\dagger\dagger\).

One is led to the Heisenberg inequality when the last mentioned quantities are stricken out in order to make stronger the inequality and \(A, B\) are chosen to be canonically conjugate:

\[
AB - BA = \frac{\hbar}{2\pi i}.
\]

Then it results in

\[
\Delta A \cdot \Delta B \geq \frac{\hbar}{4\pi}.
\]  

On the other hand, it is known that the Heisenberg limit is not really too low, but for some special \(\Psi\)-functions achieves even higher value [4]. This implies that at least for these special \(\Psi\)-functions the (central)\(^\dagger\) mean deviations-product of the canonical conjugate operators vanishes. This will be used in §2.

In the classical theory of errors or fluctuation theory it is well known that the vanishing of the mean deviations-product is a necessary (but not sufficient) condition for two values to fluctuate totally independent one from an other. While canonically conjugate quantum variables have some “independence” that could mean that some precise knowledge about one excludes such a knowledge about the other, so one could perhaps suppose that their mean deviations-product, i.e. for each \(\Psi\)-function, has vanishing expectation value. But this is not the case\(^\ast\). Let us consider the two canonically conjugate operators

\[
A = x \quad B = \frac{\hbar}{2\pi i} \frac{\partial}{\partial x},
\]

so we get [5]

\[
\frac{2\pi i}{\hbar} \frac{AB + BA}{2} = \frac{1}{2} \int \Psi^* \left[ x \frac{\partial \Psi}{\partial x} + \frac{\partial}{\partial x} (x \Psi) \right] dx = \frac{1}{2} \int x \left( \Psi^* \frac{\partial \Psi}{\partial x} - \Psi \frac{\partial \Psi^*}{\partial x} \right) dx
\]

\[
\overline{A} = \int x \Psi^* \Psi dx
\]

\[
\frac{2\pi i}{\hbar} \overline{B} = \int \Psi^* \frac{\partial \Psi}{\partial x} dx = \frac{1}{2} \int \left( \Psi^* \frac{\partial \Psi}{\partial x} - \Psi \frac{\partial \Psi^*}{\partial x} \right) dx
\]

\[
= \frac{1}{2} \int \Psi \Psi^* \frac{\partial}{\partial x} \left( \ln \frac{\Psi}{\Psi^*} \right) dx.
\]
Let now \( \Psi = re^{i\phi} \) with real \( \Phi \) and real, non-negative \( r \), which must satisfy the normalizing condition
\[
\int r^2 dx = 1
\]
Then we get:
\[
\frac{2\pi}{\hbar} \left( \frac{AB + BA}{2} - AB \right) = \int x r^2 \frac{\partial \phi}{\partial x} dx - \int x r^2 dx \cdot \int r^2 \frac{\partial \phi}{\partial x} dx.
\]
As \( \frac{\partial \phi}{\partial x} \) is any real function and \( r^2 \) is an absolute non-negative function (not taking into account the normalizing condition), so we get that in general the right hand side does not vanish. One needs for example to choose \( r^2 \) to be even and \( \frac{\partial \phi}{\partial x} \) to be odd (and not identically vanishing), so the deviation product is surely positive.

As known, the canonically conjugate quantum variable is not unambiguously defined. If \( B \) is conjugate to \( A \) this implies that \( B + \varepsilon A \) is as well (\( \varepsilon \) is any real number). With this change the mean deviations-product changes as well, and becomes, as one could easily calculate, \( \varepsilon (\Delta A)^2 \). In the same manner, the result would be \( \varepsilon (\Delta B)^2 \), if \( A \) was changed to \( A + \varepsilon B \). This can always make the deviations-product equal to zero by changing one of the operators, without changing their canonical relation. The change depends on the above shown special \( \Psi \)-function of course. One can not reach an identical vanishing of the deviations-product in such a manner.

\( \S 2. \) To the discovery of the complete inequality (9) we are led, by a chance, to the following question, which is interesting by itself. Let us consider a force free mass point, mass \( m \), coordinate \( q \), momentum \( p \), Hamilton-function \( H = \frac{p^2}{2m} \). I must undertake simultaneous measurements of the coordinate and momentum at "time zero", with highest possible precision, i.e. so that
\[
\Delta q_0 \Delta p_0 = \frac{\hbar}{4\pi}.
\]
Further I must distribute the error on \( q_0 \) and \( p_0 \) so that for a given later time point \( t \), it could be achieved the most precise place. This means \( \Delta q \) to become the least possible. We use for this purpose the very convenient "\( q \)-number-method", which is in a methodical manner opposing to the wave mechanics. I would like to elucidate shortly on it here, repeating what is well known. For the theorist working on a wave mechanics the operator, which corresponds to a specific physical variable, does not change in time. If one wants to know the mathematical expectation value for this variable, one calculates the \( \Psi \)-function for this later moment from the "time-dependent wave equation". Then one applies the corresponding operator, which as already mentioned, is the same for every moment. On the other side the \( q \)-number-theorist has to operate with one single \( \Psi \)-function at one single chosen moment, once and for all. However it is unnecessary to express any statement for it, once the moment is totally arbitrary chosen. One assumes, instead that the operators are time dependent and we may ask: how does the operator change itself in time, i.e. which operator should be applied on the original \( \Psi \)-function, in order to calculate the mathematical expectation of the respective value at the time \( t \)?

Here we point out, one may calculate the operators (or \( q \)-numbers, or matrices) almost as the usual numbers, and indeed, their change in time is determined by the equation
of motion of the classical mechanics. The only difference is that, occasionally, when it is the case, one should pay special attention to an eventual non-commutability of the operator multiplication.

So, in this present simple case, the integration of the equation of motion reads:

\[ q = q_0 + \frac{t}{m}p_0. \]

One can directly make from it the mean square deviation of the coordinate, \((\Delta q)^2\), for every moment \(t\):

\[
(\Delta q)^2 = \left( q_0 + \frac{t}{m}p_0 \right)^2 - \left( q_0 + \frac{t}{m}p_0 \right)^2
= (\Delta q_0)^2 + \frac{2t}{m} \left( \frac{q_0 p_0 + p_0 q_0}{2} - \frac{q_0 p_0}{2} \right)^2 + (\Delta p_0)^2.
\]

The middle term above is essentially the mean deviations-product of \(q_0\) and \(p_0\), which vanishes, in accordance with the prediction, when \(q_0\) and \(p_0\) are determined with optimal precision. Then we simply have

\[
(\Delta q)^2 = (\Delta q_0)^2 + \frac{t^2}{m^2}(\Delta p_0)^2
\]

or using (12)

\[
(\Delta q)^2 = (\Delta q_0)^2 + \left( \frac{\hbar}{4\pi} \right)^2 \frac{t^2}{m^2} \frac{1}{(\Delta q_0)^2}, \quad \text{see Ref. [6]}.\]

This expression becomes a minimum for that value of \((\Delta q_0)^2\), which makes both addends on the right hand side equal, i.e. for

\[
(\Delta q_0)^2 = \frac{\hbar t}{4\pi m};
\]

\((\Delta q)^2\) is then exactly twice the value of \((\Delta q_0)^2\), i.e.

\[
\Delta q = \sqrt{\frac{\hbar t}{2\pi m}}. \quad (13)
\]

It seems to me, that this final result is likely to have two points of interest. First, the proportional relation with the square root from the time, which makes allusion to well known classical deviation principles. Secondly, that the statement has an remarkable absolute character, namely, the precision attainable in a later moment depends only on intermediate time and not on the initial momentum. For example, for a free electron one might give a place prognosis for the end of the first second on the bases of already taken measurements of position and momentum, in the most favorable case with a precision of \(1cm\), quite independent of whether the electron is fast or slow [7].

Of course, at a very high speed this will be changed as it should taken into consideration the relativity theory. I believe that this could occur by the following simple considerations.
The equation (13) is applied to the rest reference system of a point mass. Let $m$ be the rest mass, $t_r$, the internal time:

$$(\Delta q)_r = \sqrt{\frac{ht_r}{2\pi m_r}}. \quad (14)$$

This is the precision that is attainable for a moving observer when calculating the position of the point mass of the co-moving system for the moment, called “$t_r$ seconds later”. When the observer shows his knowledge through signs in the space, to the “rest observer” those signs are seeming to be nearer to each other in relation of $\sqrt{1-\beta^2} : 1$; further more he must say looking from his standpoint that the prognosis were made for a time interval

$$t = \frac{t_r}{\sqrt{1-\beta^2}}. \quad (15)$$

because for him the clock, with which all the statements of the moving observer were made, run slower than his own clock. From his point of view, the mean error decreases

$$\Delta q = \sqrt{1-\beta^2} \sqrt{\frac{ht\sqrt{1-\beta^2}}{2\pi m_r}} = \sqrt{1-\beta^2} \sqrt{\frac{ht}{2\pi m}}. \quad (16)$$

It becomes smaller and comes nearer to zero when the velocity approaches is nearing the speed of the light. This happens not only when the mass $m$ goes to infinity, but also for a series of point masses moving with an ever growing speed and an ever smaller rest mass such that all the moving masses $m$ keep the same value $m$. Even in this case the maximum precision grows unlimited with the velocity approaching the speed of the light. This is indeed satisfying, since this is a boundary process that gives a hope to obtain an accurate statement for a light quantum. And this is really true for light quantum because the Maxwell waves exhibit no dispersion. They preserve indefinitely long the place precision that they got in the beginning and it could indefinitely grow, since the strong momentum dispersion, which is connected with it does not have a bad influence.

Reported on the 5th June 1930
Joint-Staff Meeting on the 19th June 1930
Distributed on the 16th July 1930

References

   H. P. Robertson, Phys. Rev. 34 (1929) 163.
   Mr. A. Sommerfeld was so kind to point out to me these two notes, when I told him about the following considerations.

[2] $A^*$ is so defined that evidently $A^*\Psi^* = (A\Psi)^*$, and therefore $A^*\Psi = (A\Psi^*)^*$.

   This proof is closely connected to the proof of the Heisenberg inequality given there.

[5] All integrals are from $-\infty$ to $+\infty$.

[6] This equation was already developed by Heisenberg in his first work on the uncertainty principle (Zeitschr. f. Phys. 43 (1927) 188), but for the special case of pseudo arbitrary character $\Psi$-function.

[7] Of course, there are wave functions that, in a precise given later moment, can determine the position with an arbitrary given precision. But one needs simply the chance to be able to trace back such an approximation of a “maximal function” by means of a wave equation over the corresponding time interval, and the resulting function to be taken as original state. But it should be not less than “optimal”.

\footnote{– Note added from the translators.}
\footnote{††) – By this Schrödinger means that the operator does not change the direction and modulus of the vector, as the horse in blinkers does not change the direction and the speed until this is not required by the driver.}
\footnote{†††) – Indeed, if one put $B = A$ in (3'), then $\text{Cov}[A, A] = \text{Var}[A] \equiv \Delta(A)^2$.}

+\footnote{Reprinted in: ERVIN SCHRODINGER, Gesammelte Abhandlungen, Band 3, Wien, Verlag der Österreichischen Akademie der Wissenschaften, pp.348-356 (1984)}\footnote{++ \footnote{Permanent address: Institute of Solid State Physics, Bulgarian Academy of Sciences, 72 Trackia Blvd., 1784 Sofia, Bulgaria. \footnote{+++ \footnote{Permanent address: Instituto de Fisica Teorica, Universidade Estadual Paulista, Rua Pamplona, 145, CEP 01405-900 Bela Vista, São Paulo, Brazil.}}} References added from translators.

[1'] W. Heisenberg, The physical principles of the quantum theory, New York, Dover, (1930);
[2'] J. Von Neumann, Mathematical foundations of quantum mechanics, Princeton, NJ, University Press, (1955);
[7'] A. Bohm, M. Loewe, Quantum mechanics: foundations and applications. 2nd rev.
and enl.ed., New York, NY, Springer-Verlag (1986);
[10] A. Messiah, Quantum Mechanics, v.1, North-Holland Publishing Company, Amsterdam (1961) p.300, eq.(VIII.9); actually the author had been very close to the general relation, only that he had not taken the centralized covariance, see the proof below eq.(VIII.9);
[16] About Heisenberg uncertainty relation, translated in russian by A.Rogali, pp.210-217, in “E.Schrödinger, Izbrannie trudi po kvantovoi mehanike (Collected papers on quantum mechanics)” Moscow, Nauka (1976)” ; the russian translation of the sentence (*) is: “However, this never happens.”;