A historical derivation of Heisenberg's uncertainty relation is flawed

John H. Marburger III[a)]
Office of Science and Technology Policy, Executive Office of the President, New Executive Office Building, Washington, District of Columbia 20502

(Received 17 August 2007; accepted 26 January 2008)

Kennard's 1927 demonstration of the uncertainty relation, cited by Heisenberg as the earliest derivation using the formalism of quantum theory, invokes a trial function that severely limits the class of wave functions for which the uncertainty relation is shown to be valid.
[DOI: 10.1119/1.2870628]

I. HISTORICAL BACKGROUND

Heisenberg proposed his famous uncertainty relation 80 years ago in the form $\Delta q \Delta p \sim h$ and asserted, but did not demonstrate, that it could be derived directly from the mathematics of the new matrix mechanics, then all of 18 months old.[1] The quantities Δq and Δp are uncertainty measures of the position q and momentum p, and h is Planck's constant. A derivation appeared immediately thereafter in a fascinating paper by Kennard, which Heisenberg cited in connection with his own proof in his 1929 University of Chicago lectures.[3,4] At about the same time Weyl sketched a proof based on the Schwarz inequality in his book on applications of group theory in quantum mechanics, citing a remark by Pauli.[5] Historians credit Kennard with the first proof.[6] However, Kennard's proof is flawed, and we must look to one of the other actors for the first unambiguously correct demonstration of this famous inequality.

Bohr appears to have been the first to point out the path to a proper derivation by linking it (in his 1927 Como lecture) to a theorem of physical optics.[7] But he gave no details and in any case his method of derivation was not what Heisenberg had in mind when he said the relation followed directly from quantum theory. By juxtaposing the uncertainty relation with the commutator $\mathbf{qp}-\mathbf{pq}=i(h/2\pi)\mathbf{I}$ Heisenberg gave the clear impression that the latter directly implies the former. Here $\mathbf{q}$ and $\mathbf{p}$ are matrices representing position and momentum and $\mathbf{I}$ is the identity matrix. The general link between commutators and uncertainty products like $\Delta q \Delta p$ was established by Robertson in 1929 and extended by Schrödinger the following year,[8,9] but most derivations for the position-momentum uncertainty product do not explicitly invoke the commutator. One of these is Kennard's.

Kennard's paper consists of two long parts—an overview of the new quantum mechanical formalism, and several applications to "simple motions," between which the uncertainty principle derivation appears beginning on p. 337 of Ref. 3. The paper has not been translated into English or included in any of the useful compilations of historical papers from this period,[2,10,11] but Nieto has remarked that Kennard's result for the motion of harmonic oscillator wave packets was a discovery of "squeezed states" that was "too far ahead of [its] time."[12] Bohr also cited Kennard's work on the motion of wave packets in some simple cases.[7] The difficulty with Kennard's approach to the uncertainty relation seems to have passed undetected for eight decades.

The uncertainty measure in these early papers often differs by $\sqrt{2}$ from the now-conventional root-mean-square deviation from the mean. This factor multiplies the right-hand side of the relations by a factor of two so, for example, Eq. (14) in Heisenberg's Chicago lectures,[4]

$$\delta p \, \delta q \geqslant \frac{h}{2\pi}, \tag{1}$$

is correct using the definition

$$(\delta q)^2 = 2 \int (q-q_m)^2 |\psi(q)|^2 dq, \tag{2}$$

where q_m is the expected value of q with the distribution $|\psi|^2$. Kennard set $q_i=\delta q$, $p_i=\delta p$, and I will use this notation and also set the means q_m, p_m to zero, which amounts to resetting the origins of the coordinate and momentum frames. I also reserve Δq and Δp for the conventional root-mean-square deviations, that is, omitting the factor of two on the right of Eq. (2), and replace $h/2\pi$ everywhere with $\hbar$. In the following $\psi(q)$ is the probability amplitude (wave function) in the position representation and $\phi(p)$ is its corresponding momentum space representation.

II. KENNARD'S DERIVATION

Kennard begins by examining a wave function that has a Gaussian amplitude and a quadratic term in its phase. The position and momentum space forms (in our notation),

$$\psi(q) = C \exp\left(-\frac{q^2}{2q_i^2} + ib^2q^2\right), \tag{3a}$$

$$\phi(p) = C_1 \exp\left(-\frac{p^2}{2p_i^2} + ib_1^2p^2\right), \tag{3b}$$

are related by a Fourier transform, which leads to the relations $C^2=1/\sqrt{\pi}q_i$, $C_1^2=1/\sqrt{\pi}p_i$, $b_1^2=b^2q_i^2/\hbar^2(1+4b^4q_i^4)$, and

$$p_i^2 = \frac{\hbar^2}{q_i^2}(1+4b^4q_i^4). \tag{4}$$

Here q_i and b are arbitrary parameters. For this particular class of functions

$$q_ip_i \geqslant \hbar \text{ for Gaussian wave packets with quadratic phase,} \tag{5}$$

and the minimum uncertainty product for these functions is achieved when the quadratic part of the phase vanishes: $b=0$.

To establish the lower bound generally, it is necessary to consider an arbitrary wave packet, which Kennard represented as

$$\psi(q) = f(q)\psi_0(q) = f(q)C_0 \exp\left(-\frac{q^2}{2q_i^2}\right). \tag{6}$$

Here C_0 is a real normalization factor and $\psi(q)$ is an arbitrary normalizable wave function whose uncertainty measure is q_i determined from Eq. (2) with $\delta q = q_i$. The function $f(q)$ is defined by this relation. Kennard substituted this form into the definition of p_i^2, which he wrote in the position representation in the form (after an integration by parts)

$$p_i^2 = 2\int \psi^*(q)\left[\frac{\hbar}{i}\frac{\partial}{\partial q}\right]^2 \psi(q)dq = 2\hbar^2 \int \psi^{*\prime}(q)\psi'(q)dq. \tag{7}$$

Primes here denote differentiation with respect to q; the reader will recognize $(\hbar/i)(\partial/\partial q)$ as the momentum operator in the position representation. After simple manipulations the reader may verify using Eq. (6) that Eq. (7) can be expressed as

$$p_i^2 = \frac{\hbar^2}{q_i^2}\left(1 + 2q_i^2 \int f^{*\prime} f' \psi_0^*(q)\psi_0(q)dq\right). \tag{8}$$

Compare Eq. (8) with Eq. (4). In Kennard's words "since the integral on the right cannot be negative, [Eq. (5)] holds in all generality" ("gilt ganz allgemein").

III. ANALYSIS OF KENNARD'S TRIAL FUNCTION

This argument, however, is circular. Equation (8) is an identity contrived to look like Eq. (4) by the particular choice of trial function in Eq. (6). Kennard's derivation can be executed with any trial function represented by an amplitude $f(q)$ times a factor $\psi_0(q)$ whose second derivative is proportional to itself, with the coefficient of proportionality at most quadratic in q (a linear term is excluded by our convention $q_m = 0$). Thus we let

$$\frac{d^2\psi_0(y)}{dy^2} = (\alpha + \beta y^2)\psi_0(y), \tag{9}$$

where $y = q/q_i$ and α and β are arbitrary real constants. We evaluate the derivatives in Eq. (7) with $\psi_0(q)$ satisfying Eq. (9) and find

$$p_i^2 = \frac{\hbar^2}{q_i^2}\left(-(2\alpha + \beta) + 2q_i^2 \int f^{*\prime} f' \psi_0^*(q)\psi_0(q)dq\right). \tag{10}$$

Because α and β are arbitrary, the lower limit achieved by setting the non-negative integral on the right to zero can be adjusted at will. For Kennard's $\psi_0(q)$ of Eq. (6), $\alpha = -1$ and $\beta = 1$. But why choose these values? Replacing q_i in Eq. (6) by q_i/γ, for example, gives $\alpha = -\gamma^2$ and $\beta = \gamma^4$. Choosing $\gamma^2 = 1/2$ would then suggest a minimum uncertainty product of $3\hbar/4$. Other forms for $\psi_0(q)$ that satisfy Eq. (9), such as $\cos(\gamma q/q_i)$ permit similar flexibility in choosing the lower limit. The point is that this reasoning does not establish a definite lower limit for the uncertainty product, nor tell us which function will achieve it. Kennard's trial function "stacks the deck" in favor of the Gaussian wave function. The arbitrariness of $\psi(q)$ is irrelevant because the expression that adds to the non-negative integral in Eq. (10) depends entirely on Eq. (9) for $\psi_0(q)$. We know from the proofs based on the Schwarz inequality that Eq. (5) is generally valid, and the minimum is achieved by a Gaussian wave packet. But Kennard's derivation is incapable of giving this information.

Is there a way to save this proof? Rather than defining $f(q)$ in terms of the arbitrary function $\psi(q)$, which makes Eq. (8) an identity, we can imagine choosing $f(q)$ freely and defining $\psi(q)$ through Eq. (6). Then the parameters C_0 and q_i would have to be determined self-consistently from the normalization condition and Eq. (2). The same mathematical steps would be valid, but now $f(q)$ would be an independent arbitrary function. Unfortunately, the class of functions for which this approach can be made to work is limited.

The problem appears when we attempt to determine the parameter q_i. We insert Eq. (6) for $\psi(q)$ into Eq. (2) with $q_m = 0$, and find

$$q_i^2 = 2\int q^2 C_0^2 |f(q)|^2 e^{-q^2/q_i^2} dq. \tag{11}$$

The normalization condition may be used to evaluate C_0^2 whose inverse is a function $G(q_i)$ of q_i:

$$\int |f(q)|^2 e^{-q^2/q_i^2} dq = \frac{1}{C_0^2} \equiv G(q_i). \tag{12}$$

We next rewrite Eq. (11) in terms of $G(q_i)$ to find

$$q_i^2 G = 2q_i^4 \frac{dG}{d(q_i^2)} \tag{13}$$

or

$$\frac{dG}{dq_i} = \frac{G}{q_i}. \tag{14}$$

Equation (14) is not a differential equation for $G(q_i)$, but an implicit equation for q_i. Solutions occur whenever a ray from the origin $G = q_i = 0$ is tangent to the (everywhere positive) graph of $G(q_i)$. Unfortunately, it does not possess physically meaningful solutions for arbitrary choices of $f(q)$. It is a nonlinear eigenvalue equation that serves not only to determine q_i, but also to restrict the class of amplitudes $|f|$ that permit real, positive, nonzero values of q_i. For example, if $|f(q)|^2$ remains finite and smooth for small q, then $G(q_i)$ for small q_i becomes

$$G(q_i) \overset{q_i\text{small}}{\rightarrow} \sqrt{\pi} q_i |f(0)|^2. \tag{15}$$

Although $G(q_i)$ automatically satisfies Eq. (14), it leads to a trial function ψ proportional to a Dirac delta function with a vanishing coefficient:

$$\psi(q) = \frac{1}{\sqrt{G(q_i)}} f(q) \exp\left(-\frac{q^2}{2q_i^2}\right) \overset{q_i\text{small}}{\rightarrow} (2\sqrt{\pi})^{1/2}\sqrt{q_i} e^{i\theta(q)} \delta(q), \tag{16}$$

where θ is the phase of f. For these functions well-behaved near zero, Eq. (14) may possess solutions other than $q_i = 0$ for some $f(q)$, but that is not necessary. For example, if f is a simple Gaussian with an arbitrary width q_f,

$$f(q) = \exp\left(-\frac{q^2}{2q_f^2}\right), \tag{17}$$

we find

$$G(q_i) = \frac{\sqrt{\pi}q_i}{\sqrt{1 + q_i^2/q_f^2}}, \tag{18}$$

for which $q_i=0$ is the only solution of Eq. (14). The same is true for $f=\cos(kq)$. The function whose squared amplitude is $q\,\sinh(kq)$ leads to the imaginary solution $q_i=2i/k$. All these functional forms are excluded from the class of wave functions for which this approach to Kennard's derivation can be implemented.

Note that a nonzero solution of Eq. (14) can occur only if $G(q_i)$ possesses an inflection point. As the width q_i of the Gaussian factor in the integral of Eq. (12) increases, spreading like an umbrella over the amplitude factor $|f(q)|^2$, the integral can increase abruptly as the Gaussian overlaps a peak in the amplitude. Thus, a function $f(q)$ whose squared amplitude has peaks displaced from the origin can possess real, positive, nonzero values of q_i. It is precisely among wave functions that do possess a single central peak, however, that we might expect to find one that reduces the uncertainty product below the lower limit already found in Eq. (5) for a Gaussian, and a correct proof needs to consider these functions.

IV. CONCLUSION

No equation in Kennard's proof is incorrect, only the statement that the proof is general, or that the minimum uncertainty product has been demonstrated. Depending on whether $\psi(q)$ or $f(q)$ in Eq. (6) is regarded as the given arbitrary function, the argument is either a tautology or entails a condition that restricts the range of functions for which the proof is valid.

This conclusion raises the question of who deserves priority for a rigorous demonstration of the correct mathematical uncertainty relation. Many renowned physicists and mathematicians were active during this period, and it is difficult to know from their published work who first conceived the proof that Heisenberg used in his Chicago lectures.[4] My favorite statement of the proof that is close to the modern version appears in Fock's 1931 text.[13] Fock cites Weyl, whose derivation is rigorous if not elegant, and Weyl acknowledges Pauli, with whom Heisenberg had shared his manuscript for Ref. 1 even before he showed it to Bohr. Pauli's derivation in his lectures is very similar to Heisenberg's.[14] Perhaps Pauli was the ultimate source of inspiration for the first rigorous derivation of this important consequence of quantum theory.

ACKNOWLEDGMENT

The author is indebted to Robert P. Crease for stimulating discussions about the historical context of the early development of quantum mechanics and the uncertainty relation.

[a]On leave from State University of New York at Stony Brook, Department of Physics, Stony Brook, New York 11794. Electronic mail: jhmiii@jhmarburger.com

[1]W. Heisenberg, "Über den anschaulichen inhalt der quanten theoretischen Kinematik und Mechanik," Zeit. f. Phys. **43**, 172–198 (1927). A translation is reprinted in Ref. 2. In this paper, Heisenberg calculates the width of Gaussian wave packets in position and momentum space and finds that their product is $h/2\pi$.

[2]*Quantum Theory and Measurement*, edited by J. Wheeler and W. Zurek (Princeton University Press, NJ, 1983).

[3]E. H. Kennard, "Zur Quantenmechanik einfacher Bewegungstypen," Zeit. f. Phys. **44**, 326–352 (1927).

[4]W. Heisenberg, *The Physical Principles of the Quantum Theory*, translated by C. Eckart and F. C. Hoyt (Dover, NY, 1930), p. 18.

[5]H. Weyl, *Group Theory and Quantum Mechanics*, translated by H. P. Robertson (Dover, NY, 1928).

[6]M. Jammer, *The Conceptual Development of Quantum Mechanics* (McGraw–Hill, NY, 1966), p. 333.

[7]N. Bohr, "The quantum postulate and the recent development of atomic theory," Nature (London) **121**, 580–590 (1928). Reprinted in Wheeler and Zureck, Ref. 2. Bohr wrote that "These relations—well known from the theory of optical instruments, especially from Rayleigh's investigation of the resolving power of spectral apparatus—express the condition that the wave-trains extinguish each other by interference at the space-time boundary of the wave field."

[8]H. P. Robertson, "The uncertainty principle," Phys. Rev. **34**, 163–164 (1929).

[9]E. Schrödinger, "Zum Heisenbergschen Unschärfeprinzip," Abh. Preuss. Akad. Wiss., Phys.-Math. Kl. **19**, 296–303 (1930).

[10]*Sources of Quantum Mechanics*, edited by B. L. van der Waerden (Dover, NY, 1968).

[11]E. Schrödinger, *Collected Papers on Wave Mechanics* (Chelsea, NY, 1982), 3rd English ed.

[12]M. M. Nieto, "The discovery of squeezed states – In 1927," in *Proceedings of the Fifth International Conference on Squeezed States and Uncertainty Relations*, NASA Conference Publication NASA/CP-1998-206855, edited by D. Han, J. Janszky, Y. S. Kim, and V. I. Man'ko (NASA, Washington, DC, 1998), pp. 175–180, ArXiv:quant-ph/9708012.

[13]V. A. Fock, *Fundamentals of Quantum Mechanics*, translated by E. Yankovsky (MIR, Moscow, 1978), pp. 111–112.

[14]W. Pauli, *Pauli Lectures on Physics: Wave Mechanics*, edited by Charles P. Enz, translated by H. R. Lewis and S. Margulies (MIT Press, Cambridge, MA, 1973), p. 7. These lectures were delivered in 1956–57, but drew on material in Pauli's famous article, "Die allgemeinen Prinzipien der Wellenmechanik," Handbuch der Physik, Band 24/1 (Springer, Berlin, 1933).