UM–P–95/51

$[p,q] \neq i\hbar$

John P. Costella*

School of Physics, The University of Melbourne, Parkville, Victoria 3052, Australia

(22 May 1995)

Abstract

In this short note, I point out that $[p,q] \neq i\hbar$, contrary to the original claims of Born and Jordan, and Dirac. Rather, $[p,q]$ is equal to something that is *infinitesimally different* from $i\hbar$. While this difference is usually harmless, it does provide the solution of the Born–Jordan "trace paradox" of $[p,q]$. More recently, subtleties of a very similar form have been found to be of fundamental importance in quantum field theory.

When Born and Jordan [1] and Dirac [2] discovered the relationship $[p,q] = i\hbar$, it was a turning point in physics. Classically, physical quantities had always been assumed to *commute*; quantum mechanics was born when this assumption was discarded. Matrix mechanics reflects this non-commutativity by representing quantities such as p and q by *matrices*; wave mechanics does likewise by considering them to be *operators*; and Dirac's c-number and q-number formulation simply takes non-commutativity as the starting point.

Born and Jordan obtained $[p,q] = i\hbar$ by means of arguments based on the correspondence principle; and Dirac obtained it by his Poisson bracket ansatz. Let us review its standard wave-mechanical derivation. In the q-representation, the state vector $|\psi\rangle$ is a function $\psi(q)$, the operator q is simply multiplication by q, and the operator p is defined as $i\hbar\,\partial/\partial q$. Thus the identity $[p,q] = i\hbar$ is just a scaling by the factor $i\hbar$ of the identity

$$[\partial_q, q] = 1, \tag{1}$$

where I am using the notation ∂_q to denote $\partial/\partial q$. The meaning of (1) is made more explicit if we write in the implied function $\psi(q)$ on both sides:

$$[\partial_q, q]\psi(q) = \psi(q).$$

It is straightforward to multiply out the left-hand side, and use the product rule to obtain the right-hand side:

$$\begin{aligned}[\partial_q, q]\psi(q) &\equiv \partial_q q\psi(q) - q\partial_q\psi(q)\\ &= \psi(q) + q\partial_q\psi(q) - q\partial_q\psi(q)\\ &= \psi(q).\end{aligned}$$

A problem, however, arises when we want to make the transition to the *matrix mechanical* formulation of quantum mechanics. In this formulation, the state vector $|\psi\rangle$ is written as a *column vector*. Physically observable quantities, such as p and q, must be represented by *Hermitian matrices*. The equivalent of the q-representation of wave mechanics is obtained by taking the vertical position in the column vector as a linear function of the value q. Since q is a continuous variable, and the rows of a column vector are discrete, we must consider the *limit* of a sequence of discrete matrix representations, of ever increasing dimension, such that the positions in the column vector for $|\psi\rangle$ "fill in" the domain of q more and more densely, so that in the limit of an infinite-dimensional matrix they form a continuum. Let us label the rows in such a way that the "middle one" has the index value $n = 0$; the rows above are rows $n = -1, -2, -3, \ldots$, and those below are $n = +1, +2, +3, \ldots$. Let us then deem that row 0 is to represent the origin of the q coordinate, $q = 0$. Then the relationship between q and n is of the form

$$q = \ell n,$$

where ℓ is some length scale, that will shrink as the dimension of the matrix representation is increased. (The precise mathematical form of this "shrinking rate" does not need to be known for our purposes). In other words, if we denote the column vector representing $|\psi\rangle$ by the boldface symbol $\boldsymbol{\psi}$, then we have

$$\begin{pmatrix} \vdots \\ \boldsymbol{\psi}_{-2} \\ \boldsymbol{\psi}_{-1} \\ \boldsymbol{\psi}_{0} \\ \boldsymbol{\psi}_{+1} \\ \boldsymbol{\psi}_{+2} \\ \vdots \end{pmatrix} \equiv \begin{pmatrix} \vdots \\ \psi(-2\ell) \\ \psi(-\ell) \\ \psi(0) \\ \psi(+\ell) \\ \psi(+2\ell) \\ \vdots \end{pmatrix}. \tag{2}$$

Let us now construct the matrix $\boldsymbol{q}$ that represents the operator q. Clearly, the quantity $q\psi(q)$ is given in the matrix representation by

$$\boldsymbol{q}\boldsymbol{\psi} = \begin{pmatrix} \vdots \\ -2\ell\,\psi(-2\ell) \\ -\ell\,\psi(-\ell) \\ 0 \\ +\ell\,\psi(+\ell) \\ +2\ell\,\psi(+2\ell) \\ \vdots \end{pmatrix};$$

from the identity (2), it is then clear that

$$\boldsymbol{q} = \ell \begin{pmatrix} \ddots & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & -2 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & -1 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 0 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & +1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & +2 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \ddots \end{pmatrix}$$

is the matrix corresponding to the operator $\boldsymbol{q}$, where dots indicate zero entries in the matrix.

Constructing a matrix $\boldsymbol{p}$ to represent the operator p is a little more subtle. Clearly, it relies on us devising a suitable matrix operator that is equivalent to the operator ∂_q, in the limit of an infinite dimensional matrix. Now, since p is postulated to be an observable quantity, the matrix $\boldsymbol{p}$ must be Hermitian; and since by definition $p \equiv i\hbar\partial_q$, then it follows that the matrix representation $\boldsymbol{\partial}_q$ of ∂_q must be *anti*-Hermitian. Furthermore, since the derivative ∂_q of any *real* function $\psi(q)$ must itself be real, and since the definition of $\boldsymbol{\partial}_q$ cannot depend on whether the function we apply it to is real or complex, then it follows that $\boldsymbol{\partial}_q$ must in full generality be a real matrix. Taken together, these two considerations already tell us that $\boldsymbol{\partial}_q$ must be a *real, antisymmetric* matrix. To find its exact form, let us consider the meaning of the derivative ∂_q from first principles: for a function $\psi(q)$,

$$\partial_{q'}\psi(q')\Big|_{q'=q} \equiv \lim_{\varepsilon,\varepsilon'\to 0} \frac{\psi(q+\varepsilon)-\psi(q-\varepsilon')}{\varepsilon+\varepsilon'}, \tag{3}$$

where ε and ε' are real numbers greater than zero. Now, in the matrix representation, for a finite dimension, we do not have positions that are infinitesimally close to a given $q_n \equiv n\ell$; rather, the closest we can get are the two points $q_{n+1} = (n+1)\ell$ and $q_{n-1} = (n-1)\ell$. However, in the limit of an infinite-dimensional matrix representation, these two points will shrink around the point q_n in the way we desire. Moreover, we already know that we cannot use the point q_n itself in the definition of ∂_q, since the matrix $\boldsymbol{\partial}_q$ must be antisymmetric, which means that the diagonal elements must vanish. The best that we can do is therefore

$$(\boldsymbol{\partial}_q\boldsymbol{\psi})_n \equiv \frac{\boldsymbol{\psi}_{n+1}-\boldsymbol{\psi}_{n-1}}{2\ell}, \tag{4}$$

which is equivalent to (3), with $\varepsilon = \varepsilon'$, in the limit of an infinite-dimensional matrix. This then imples that

$$\boldsymbol{\partial}_q = \frac{1}{2\ell}\begin{pmatrix} \ddots & \ddots & \cdot & \cdot & \cdot & \cdot & \cdot \\ \ddots & 0 & +1 & \cdot & \cdot & \cdot & \cdot \\ \cdot & -1 & 0 & +1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & -1 & 0 & +1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & -1 & 0 & +1 & \cdot \\ \cdot & \cdot & \cdot & \cdot & -1 & 0 & \ddots \\ \cdot & \cdot & \cdot & \cdot & \cdot & \ddots & \ddots \end{pmatrix}, \tag{5}$$

as can be verified directly by multiplying (5) by (2). In other words, the matrix $\boldsymbol{p}$ is given by

$$\boldsymbol{p} = \frac{\hbar}{2\ell}\begin{pmatrix} \ddots & \ddots & \cdot & \cdot & \cdot & \cdot & \cdot \\ \ddots & 0 & +i & \cdot & \cdot & \cdot & \cdot \\ \cdot & -i & 0 & +i & \cdot & \cdot & \cdot \\ \cdot & \cdot & -i & 0 & +i & \cdot & \cdot \\ \cdot & \cdot & \cdot & -i & 0 & +i & \cdot \\ \cdot & \cdot & \cdot & \cdot & -i & 0 & \ddots \\ \cdot & \cdot & \cdot & \cdot & \cdot & \ddots & \ddots \end{pmatrix}, \tag{6}$$

which is clearly Hermitian, as required.

We can now turn immediately to the issue raised by the title of this note, by computing $[p, q]$ in the matrix representation above—namely, by computing the matrix commutator $[\boldsymbol{p}, \boldsymbol{q}]$. By multiplying out the matrices, it is easily seen that

$$\boldsymbol{pq} = \frac{i\hbar}{2}\begin{pmatrix} \ddots & \ddots & \cdot & \cdot & \cdot & \cdot & \cdot \\ \ddots & 0 & -1 & \cdot & \cdot & \cdot & \cdot \\ \cdot & +2 & 0 & 0 & \cdot & \cdot & \cdot \\ \cdot & \cdot & +1 & 0 & +1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & 0 & 0 & +2 & \cdot \\ \cdot & \cdot & \cdot & \cdot & -1 & 0 & \ddots \\ \cdot & \cdot & \cdot & \cdot & \cdot & \ddots & \ddots \end{pmatrix}$$

and

$$\boldsymbol{qp} = \frac{i\hbar}{2}\begin{pmatrix} \ddots & \ddots & \cdot & \cdot & \cdot & \cdot & \cdot \\ \ddots & 0 & -2 & \cdot & \cdot & \cdot & \cdot \\ \cdot & +1 & 0 & -1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & 0 & 0 & 0 & \cdot & \cdot \\ \cdot & \cdot & \cdot & -1 & 0 & +1 & \cdot \\ \cdot & \cdot & \cdot & \cdot & -2 & 0 & \ddots \\ \cdot & \cdot & \cdot & \cdot & \cdot & \ddots & \ddots \end{pmatrix};$$

we therefore find that

$$[\boldsymbol{p}, \boldsymbol{q}] = \frac{i\hbar}{2}\begin{pmatrix} \ddots & \ddots & \cdot & \cdot & \cdot & \cdot & \cdot \\ \ddots & 0 & 1 & \cdot & \cdot & \cdot & \cdot \\ \cdot & 1 & 0 & 1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & 1 & 0 & 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & 1 & 0 & 1 & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1 & 0 & \ddots \\ \cdot & \cdot & \cdot & \cdot & \cdot & \ddots & \ddots \end{pmatrix}. \tag{7}$$

Here is the subtlety. The problem is that the matrix (7) is *not* equal to $i\hbar$ times the unit matrix $\mathbf{1}$,

$$i\hbar\,\mathbf{1} \equiv \frac{i\hbar}{2}\begin{pmatrix} \ddots & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & 2 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 2 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 2 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 2 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & 2 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \ddots \end{pmatrix}. \tag{8}$$

Rather, the matrix (7) is effectively obtained by taking each diagonal element of 2 and "splitting it" between the off-diagonals above and below. Thus we have proved the relation

$$[\boldsymbol{p}, \boldsymbol{q}] \neq i\hbar\mathbf{1} \tag{9}$$

in the matrix representation of quantum mechanics, and hence in full generality

$$[p,q] \neq i\hbar, \tag{10}$$

as I have claimed in the title of this note.

The result (10) might be somewhat disturbing. However, in almost all cases, it is of academic interest only. The reason is that using the matrix (8) rather than (7), in any practical calculation, corresponds to the replacement

$$\lim_{\ell\to 0} \frac{\psi(q-\ell)+\psi(q+\ell)}{2} \to \psi(q),$$

which is arguably harmless for any reasonable $\psi(q)$. In fact, we can obtain exactly the same result $[p,q] \neq i\hbar$ using the *wave mechanical* representation, if we treat the operation of differentiation more carefully, rather than by simply using the product rule. If we write down the wave-mechanical equivalent of the Hermitian (*i.e.*, symmetrical) definition (4), namely, the symmetrical version of (3),

$$\partial_{q'}\psi(q')\Big|_{q'=q} \equiv \lim_{\varepsilon\to 0} \frac{\psi(q+\varepsilon)-\psi(q-\varepsilon)}{2\varepsilon},$$

then we find that

$$\begin{aligned}
[\partial_q, q]\psi(q) &\equiv \partial_q q\psi(q) - q\partial_q\psi(q) \\
&\equiv \lim_{\varepsilon\to 0}\left\{ \frac{(q+\varepsilon)\,\psi(q+\varepsilon)-(q-\varepsilon)\,\psi(q-\varepsilon)}{2\varepsilon} - q\,\frac{\psi(q+\varepsilon)-\psi(q-\varepsilon)}{2\varepsilon}\right\} \\
&\equiv \lim_{\varepsilon\to 0} \frac{\psi(q+\varepsilon)+\psi(q-\varepsilon)}{2},
\end{aligned}$$

in agreement with the matrix mechanical result.

It might seem that claiming that $[p,q] \neq i\hbar$ is a pedantry. After all, when would shifting the argument q by an infinitesimal amount, or shifting by one row or column in an infinite-dimensional matrix representation, make any difference? There is at least one situation that I am aware of in which this change *does* make a difference: whenever the *trace* of the matrix is taken. For example, the Born and Jordan's [1] well-known "trace paradox" of $[p,q]$ points out the following: since

$$\mathrm{Tr}(\boldsymbol{AB}) \equiv \mathrm{Tr}(\boldsymbol{BA})$$

for any finite matrices $\boldsymbol{A}$ and $\boldsymbol{B}$, then in the finite-dimensional case we must have

$$\mathrm{Tr}[\boldsymbol{p},\boldsymbol{q}] \equiv \mathrm{Tr}(\boldsymbol{pq}-\boldsymbol{qp}) \equiv 0.$$

But if $[p,q] = i\hbar$ were to hold true, then we would need to have

$$\mathrm{Tr}[\boldsymbol{p},\boldsymbol{q}] = i\hbar\,\mathrm{Tr}(\boldsymbol{1}) = i\hbar\, D,$$

where D is the dimension of the matrix representation, which, rather than vanishing, approaches infinity in the infinite-dimensional limit! I emphasise that *this is a fallacy*; it is the matrix (7) that must be used, *not* the identity matrix (8). And of course the matrix (7) is identically *traceless*; hence, the Born–Jordan "trace paradox" of $[p, q]$ is due to the incorrect assumption that $[p, q] = i\hbar$, whereas at the level of individual rows and columns of the matrix representation it fails.

It might be claimed that this simply shows that one cannot take infinite-dimensional matrix mechanics to be the infinite-dimension limit of finite-dimensional matrix mechanics. But then what would this "matrix mechanics" have to do with matrices as we know them? Moreover, it is generally believed that the correct way of dealing with infinities, or infinitesimals, in physical problems is to take them to be the limit of *finite* quantities. Surely, then, it is better to modify the *postulate* of $[p, q] = i\hbar$ by an infinitesimal amount, rather than remove all chance of using a well-defined limiting procedure?

Furthermore, this ability of the trace—to be able to yield an answer that is either zero or infinite, depending on how carelessly one defines one's matrix quantities—turns out to be more important to real-world calculations than one might naïvely think. In quantum field theory, the effect of effectively "including the diagonal terms" in the time-ordered product operation, when in fact they should *not* be included, leads to a drastic and fundamental change in the predictions of calculations involving loop diagrams. This has has been pointed out several times in the past two decades, but has not gained much attention; we shall be providing a full and explicit description of these developments shortly [3].

Acknowledgments

Helpful discussions with T. D Kieu, B. H. J. McKellar, A. A. Rawlinson, M. J. Thomson and G. J. Stephenson, Jr. are gratefully acknowledged. This work was supported in part by the Australian Research Council.

References

* jpc@physics.unimelb.edu.au; http://www.ph.unimelb.edu.au/~jpc/homepage.htm.

[1] M. Born and P. Jordan, *Z. Phys.* **34** (1925) 858.

[2] P. A. M. Dirac, *Proc. Roy. Soc. London A* **109** (1926) 642.

[3] J. P. Costella, T. D. Kieu, B. H. J. McKellar, A. A. Rawlinson, M. J. Thomson and G. J. Stephenson, Jr., in preparation.