QUANTUM MECHANICS AS AN INTUITIONISTIC FORM OF CLASSICAL MECHANICS

JOHN V. CORBETT AND MURRAY ADELMAN

Abstract. Intuitionistic real numbers are constructed as sheaves on the state space of the Schrödinger representation of a CCR-algebra with a finite number of degrees of freedom. These numbers are used as the values of position and momentum variables that obey Newton's equations of motion. Heisenberg's operator equations of motion are shown to give rise to numerical equations that, on a family of open subsets of state space, are local approximations to Newton's equations of motion for the intuitionistically valued variables.

Introduction

Do sub-atomic particles have positions and momenta at all times? Can the numerical value of the position of such a particle always be given a triplet of real numbers? What real numbers should be used?

In this paper we argue that there is a class of real numbers that can be used to label the positions and momenta of sub-atomic particles at all instances of time. Furthermore if these numbers, as values of positions and momenta, are assumed to satisfy the equations of motion of classical mechanics, then these classical equations are approximated locally by the quantum mechanical operator equations of motion, restricted to act on certain subsets of state space.

The real numbers that we use are real numbers in the topos of sheaves on the state space of the Schrödinger representation of a CCR-algebra with a finite number of degrees of freedom. In the standard Hilbert space framework for quantum mechanics[1] the algebra of the canonical commutation relations (CCR) is generated by the operators $Q_j = Q_j^*$, $P_j = P_j^*$, $j = 1, \ldots, n$, and the identity operator I which satisfy the relations $[P_j, Q_j] = -iI$, $[P_j, Q_k] = 0$ if k is different from j, $[Q_j, Q_k] = 0$, $[P_j, P_k] = 0$ for all j, k. We have put Planck's constant divided by 2π equal to one.

We will further simplify the notation by taking $n = 1$.

1. The Schrödinger representation

The Schrödinger representation of the CCR-algebra M is the representation in which the Hilbert space is $L^2(R)$. Q is represented by multiplication by the real variable x and P by $(1/i)$ times the operator of differentiation with respect to x. Let $\mathcal{S}(\mathbb{R})$ denote the Schwartz space of infinitely differentiable functions of rapid decrease on $\mathbb{R}$. Then the physical quantities are represented by self-adjoint elements in the closure $\bar{M}$ of the CCR-algebra M, where $\bar{M}$ is the smallest closed extension of M. $\bar{M} = \{X^- \mid X \in M$ and X^- is the restriction to $\mathcal{S}(\mathbb{R})$ of the Hilbert space closure of $X\}$. Following the definitions of Powers[7], M is essentially self-adjoint because the adjoint M^* of M equals the closure $\bar{M}$ of M.

2. The State Space $\mathcal{E}_{\mathcal{S}}$

We give only a resume of the results, for more details see Inoue[8]. A linear functional f on M is strongly positive iff $f(X) \geq 0$ for all $X \geq 0$ in M.

Definition 1. *The states on M are the strongly positive linear functionals on M that are normalised to take the value 1 on the element I of M.*

Theorem 1 (Inoue[8]). *Every strongly positive linear functional on M is given by a trace functional.*

Definition 2. *The state space $\mathcal{E}_{\mathcal{S}}$ of the Schrödinger representation of M is the set of all strongly positive linear functionals on M that are normalised.*

The state space $\mathcal{E}_{\mathcal{S}}$ is contained in the convex hull of projections $\mathcal{P}$ onto one-dimensional subspaces spanned by unit vectors $u \in \mathcal{S}(\mathbb{R}^3)$. Recall that each state $\rho \in \mathcal{E}_{\mathcal{S}}$ is a trace class positive bounded operator with trace 1, which can be written as $\rho = \sum \lambda_n \mathcal{P}_n$, where the sum over n may go from 1 to ∞. For all n, $\lambda_n \geq 0$, $\sum \lambda_n = 1$ and the $\mathcal{P}_n$ are orthogonal projections. The following result demonstrates that all states in $\mathcal{E}_{\mathcal{S}}$ must satisfy the further condition that as n approaches infinity the sequence $\{\lambda_n\}$ converges to zero faster than any power of $1/n$.

Theorem 2. *Given $\rho = \sum \lambda_n \mathcal{P}_n \in \mathcal{E}_{\mathcal{S}}$, where the projections $\mathcal{P}_n$ project onto one-dimensional subspaces spanned by unit vectors $u_n \in \mathcal{S}(\mathbb{R}^3)$. $\mathrm{tr}\,\rho A$ is finite for any self-adjoint operator A in the Schrödinger representation of the CCR-algebra M if and only if $\lim n^k \lambda_n = 0$, for all integers $k > 0$.*

Proof. Let A be a self-adjoint operator in M and ρ is a state on M which belongs to $\mathcal{E}_{\mathcal{S}}$. Then for any orthonormal basis of $\mathcal{H}$, $\{u_m\}$, $m = 1$ to ∞, which is composed from elements in $\mathcal{S}$, $\operatorname{tr} \rho A = \sum (u_m\,, \rho A u_m)$.

Now using $\rho = \sum \lambda_n \mathcal{P}_n$ the trace becomes a double sum

$$\operatorname{tr} \rho A = \sum \big(u_m\,, \sum \lambda_n \mathcal{P}_n A u_m\big)\ .$$

Because the trace is independent of the basis used to calculate it, we can choose the ortho-normal basis $\{u_m\}$ to be in the ranges of the $\mathcal{P}_n$ so that the double sum reduces to

$$\operatorname{tr} \rho A = \sum \lambda_m \big(u_m\,, A u_m\big)\ .$$

Since A is in the algebra M, it is a polynomial of some finite degree k in the self-adjoint operators P, Q. By Lemma 18 in Jaffe[10], A is majorized by a polynomial of degree k in the self-adjoint operator $H = \frac{1}{2}\sum(P^2 + Q^2)$,

$$\big|\big(u_m, A u_m\big)\big|^2 \leq \big(u_m, p_k(H) u_m\big)\ .$$

When the $\{u_m\}$ are the eigenfunctions of H, $p_k(H)u_m = p_k(\mu_m)u_m$, where $\mu_m = \frac{1}{2}(2m+1)$ are the eigenvalues of H, so that the absolute value of the m^{th} term of the series $\sum \lambda_m(u_m, A u_m)$ satisfies

$$\begin{aligned} 0 &\leq \lambda_m\big|(u_m, A u_m)\big| \\ &\leq \lambda_m\big(u_m, H(\mu_m)u_m\big)^{\frac{1}{2}} = \lambda_m p_k(\mu_m)^{\frac{1}{2}}. \end{aligned}$$

Therefore if $\lim n^k \lambda_n = 0$, for all $k > 0$, then $\exists$ a positive integer N and a positive constant C such that, for all $m > N$,

$$\lambda_m p_k(\mu_m)^{\frac{1}{2}} < \frac{C}{m^2}$$

and so the series $\sum \lambda_m(u_m, A u_m)$ converges absolutely by the comparison test. On the other hand, if $\operatorname{tr} \rho A$ is finite for all self-adjoint operators A then the series $\sum \lambda_m(u_m, A u_m)$ must converge. If A is positive then the convergence is absolute so that when A is a monomial of degree k in the Hamiltonian operator H,then $\lim \lambda_m(\mu_m)^{\frac{1}{2}(k)} = 0$ for all positive k which shows the converse is true.

The same results hold even when the $\{u_m\}$ are not the eigenfunctions of H. The minimax principle for eigenvalues of symmetric operators, see for example, Kato[6] section I.6.10, implies that

$$\big(u_m, p_k(H) u_m\big) \leq p_k(\mu_m)$$

because it states that the maximum of $\big(v, p_k(H)v\big)$, when v is a unit vector and $(v, e_i) = 0$, $i = 1, \cdots, n-1$, for any orthonormal set of vectors $\{e_i\}$, is the eigenvalue $p_k(\mu_m)$ obtained when the $\{e_i\}$ are the first $n-1$ eigenvalues of H. This result is applicable here to the set of vectors $\{u_i\}$, $i = 1, \cdots, m$, with $v = u_m$.

We have shown that $\operatorname{tr} \rho A$ is finite for any self-adjoint operator A in this representation of the CCR-algebra M if and only if $\lim n^k \lambda_n = 0$, for all integers $k > 0$. □

The topology on $\mathcal{E}_\mathcal{S}$ is chosen to make the functions $\hat{A}$, where $\hat{A}(\rho) = \operatorname{tr} \rho A$, continuous when A is an essentially self-adjoint continuous operator on $\mathcal{S}(\mathbb{R})$.

3. The topos Shv($\mathcal{E}_\mathcal{S}$)

A *sheaf* Y on a topological space X can be described[3] by a rule which assigns to each point x of X a set $Y(x)$ consisting of the germs of a prescribed class of functions, where the germs of the functions are defined in neighborhoods of the point x. The collection of sets $Y(x)$ which are labelled by points x in X can be glued together to form a space Y in such a way that the projection from Y onto X is a local homeomorphism; that is, for each x in X and each y on the fiber above x (i.e. for each such y the projection of y onto X is x) there is a neighborhood N of y such that the projection of N onto X is a neighborhood of x.

A *section* of the sheaf Y over the open subset U of X is a function s from U to Y that belongs to the prescribed class of functions and satisfies the condition that, for all x in U, the projection of $s(x)$ onto X is x. The sheaf construction allows a section f defined on the open set U to be restricted to sections $f\big|_V$ on open subsets V contained in the open set U and, conversely, the section f on U can be recovered by patching together the sections $f\big|_{V'}$ where V' belongs to an open cover of U.

A *spatial topos* is a category of sheaves on a topological space. The objects of this category are sheaves over the topological space and the arrows are sheaf morphisms, that is, an arrow is a continuous function that maps a sheaf Y to a sheaf Y' in such a way that it sends fibers in Y to fibers in Y', equivalently, sections of Y over U to sections of Y' over U, where U is an open subset of X.

The topos Shv($\mathcal{E}_\mathcal{S}$) of sheaves on the topological space $\mathcal{E}_\mathcal{S}$ is constructed in this way.

In 1970, Lawvere[2] showed that toposes can be viewed as a "variable" set theory whose internal logic is intuitionistic. The propositional calculus of the logic of the spatial topos of sheaves over X is the Heyting Algebra[3] of the open subsets of X. This means that as well as being true or false, propositions in this logic can be true to intermediate extents which are given by open subsets of X. True corresponds to the whole set X. False corresponds to the empty set. There exist Boolean algebras in which propositions can be true to varying extents but, in addition, the Heyting algebra of open sets does not satisfy all the laws of classical logic. Two of the most striking differences between classical and intuitionistic logics are that the law of the excluded middle and the Axiom of Choice do not hold for intuitionistic logic. We have argued elsewhere[5] that aspects of quantum mechanics, such as the two slit experiment,that are difficult to understand with Boolean logic are better described using intuitionistic logic.

There is an analogy between the language of toposes and that of sets which makes it easier to work in toposes. Sheaves in a topos correspond to sets, subsheaves of a sheaf to subsets and local sections to elements of a set. Then as long as a proof in set theory does not use the law of excluded middle or the Axiom of Choice then it can be translated into a proof in topos theory.

4. Real Numbers in Spatial Toposes

Dedekind numbers are defined to be the completion of the rational numbers obtained by using cuts, and *Cauchy numbers* are defined as the completion of the rationals obtained by using Cauchy sequences. These different constructions can only be shown to be equivalent by using either the Axiom of Choice or the law of the excluded middle.[3] Therefore, when intuitionistic logic is assumed, we expect that these two types of real numbers are not equivalent.

It has been shown[4] that in a spatial topos the sheaf of rational numbers $\mathbb{Q}$ is the sheaf whose sections over an open set U are given by locally constant functions from U with values in the standard rationals while the sheaf of Cauchy reals $\mathbb{R}_{\mathrm{C}}$ is the sheaf whose sections over an open set U are given by locally constant functions from U with values in the standard reals. A function is locally constant if it is constant on each connected open subset of its domain.

On the other hand the sheaf of Dedekind reals $\mathbb{R}_{\mathrm{D}}$ is the sheaf whose sections over U are given by continuous functions from U to the standard reals.

The Cauchy reals form a proper sub-sheaf of the Dedekind reals unless the underlying topological space X is the one point space.

5. The Quantum Reals

By the construction of the topology on $\mathcal{E}_{\mathcal{S}}$ for any self-adjoint operator A in the Schrödinger representation of the CCR, the function $\mathrm{tr}\,\rho A$ is a globally defined continuous function and therefore in $\mathbb{R}_{\mathrm{D}}$. We interpret the functions $A(\rho) = \mathrm{tr}\,\rho A$, with domains given by open subsets of state space, to be the numerical values of the physical quantity that is represented by the self-adjoint operator A. Not every Dedekind real number is of this form. Real numbers of this form are a proper subsheaf of the Dedekind real numbers in the spatial topos of sheaves on the state space $\mathcal{E}_{\mathcal{S}}$. We will call them the quantum reals, they belong to the sheaf of locally linear functions.[5]

Definition 3. *If U is an open subset of $\mathcal{E}_{\mathcal{S}}$ then the function f from U to the standard reals is* locally linear *at ρ in U and there is an open neighborhood U' of ρ, with U' inside U, and a bounded self-adjoint operator A such that $f\big|_{U'} = \hat{A}\big|_{U'}$.*

Definition 4. *The sheaf of locally linear functions, $\mathbb{A}$, is defined by its sections over any open subset U of $\mathcal{E}_{\mathcal{S}}$ as the set of all locally linear functions on U with the requirement that if the open set V is contained in U then the sheaf of locally linear functions over V is obtained by restricting the locally linear functions over U.*

The global elements of the locally linear functions are given by the functions $\hat{A}$, where A is a self-adjoint operator, that are continuous on $\mathcal{S}$. It suffices to define algebraic relations between elements of $\mathbb{A}$ globally, because $\mathcal{E}_{\mathcal{S}}$ is locally connected and so we can treat functions which are defined on disjoint connected components as if they were globally defined.

When U is an open neighbourhood of the state ρ then the quantum real numbers belonging to the sections of $\mathbb{A}$ over U can be thought of as real numbers tangent to those Dedekind reals that have a tangent space at ρ.

6. The Dedekind Reals $\mathbb{R}_{\mathrm{D}}$

Stout[4] has shown that the usual order on the rational numbers $\mathbb{Q}$ can be extended to the following order on $\mathbb{R}_{\mathrm{D}}$.

Definition 5. *The order relation $<$ on the Dedekind reals, $\mathbb{R}_{\mathrm{D}}$, is given by the definition:*

$$x < y \text{ if and only if } \exists\, q \in \mathbb{Q}\ \big((q \in x^{+}) \wedge (q \in y_{-})\big)$$

where x^{+} is the upper cut of x and y_{-} is the lower cut of y. The relation $<$ is the subobject of $\mathbb{R}_{\mathrm{D}} \otimes \mathbb{R}_{\mathrm{D}}$ consisting of such pairs (x, y).

Trichotomy does not hold universally for the order $<$ on $\mathbb{R}_D$.

The order $\leq$ on $\mathbb{R}_D$ has the property that $x \leq y$ is not the same as $(x < y) \vee (x = y)$.

Definition 6. *The order relation $x \leq y$ is the subobject of $\mathbb{R}_D \otimes \mathbb{R}_D$ consisting of the pair (x, y) with $x^+ \subset y^+$ and $y_- \subset x_-$, where x^+ is the upper cut of x and x_- is the lower cut of x, and similarly for y.*

Stout[4] also showed that the statement $(x \leq y)$ is equivalent to the statement $\neg(y < x)$.

Definition 7. *The open interval (x, y) for $x < y$ is the subobject of $\mathbb{R}_D$ consisting of those z in $\mathbb{R}_D$ that satisfy $x < z < y$.*

The closed interval $[x, y]$ for $x < y$ is the subobject of $\mathbb{R}_D$ consisting of those z in $\mathbb{R}_D$ that satisfy $x \leq z \leq y$.

The open intervals can be used to construct an interval topology T on $\mathbb{R}_D$ analogously to the interval topology on the standard real numbers that is generated from the open intervals by finite intersection and arbitrary union.

The topology T on $\mathbb{R}_D$ is such that $\mathbb{Q}$ is dense in $\mathbb{R}_D$ with respect to T.

If the max function is defined using the order $\leq$ by the conditions: (i) $x \leq \max(x, y)$ and $y \leq \max(x, y)$, and (ii) if $z \leq x$ and $z \leq y$ then $z \leq \max(x, y)$ and the norm function: $|\ \ | : \mathbb{R}_D \to \mathbb{R}_D$ is defined by $|x| = \max(x, -x)$, then the norm $|\ \ |$ satisfies the usual conditions of non-negativity, that only 0 has norm zero and that the triangle inequality holds.

$(\mathbb{R}_D, T)$ is a metric space with the metric $d(x, y) = \big\|x - y\big\|$. It is both complete and separable[4].

$\mathbb{R}_D$ is a field in the sense that for all a in $\mathbb{R}_D$,if a does not belong to the sheaf of germs of invertible functions, Unit($\mathbb{R}_D$),then $a = 0$.

7. Properties of the Quantum Reals

Theorem 3. *$\mathbb{A}$ is a proper sub-sheaf of the sheaf of Dedekind numbers $\mathbb{R}_D$ and is dense in $\mathbb{R}_D$ in the metric topology T.*

The sheaf $\mathbb{A}$ inherits the orders $\leq$ and $<$ from $\mathbb{R}_D$. On the other hand $\mathbb{A}$ can be ordered as a consequence of the orders on the self-adjoint operators:

1. A is strictly positive, $A > 0$, if $(Au, u) > 0$ for $u \neq 0$, $u \in D(A)$.
2. A is non-negative, $A \geq 0$, if $(Au, u) \geq 0$ for all $u \in D(A)$.

Lemma 8. *The orders $\leq$ and $<$ on $\mathbb{A}$ inherited from $\mathbb{R}_D$ are equivalent to those obtained from those on continuous self-adjoint operators.*

Proof. A is a non-negative self-adjoint operator iff $\operatorname{tr} \rho A \geq 0$ for all ρ in $\mathcal{E}_{\mathcal{S}}$, i.e. $\hat{A} \geq 0$ globally. When, as a Dedekind real number, $a = \hat{A} \geq 0$ then $0^+ \subset a^+$ and $a_- \subset 0_-$. Globally, $0^+ = \{q \in \mathbb{Q} \mid q > 0\}$ and $0^- = \{q \in \mathbb{Q} \mid q > 0\}$ so that, if $a = \hat{A} \geq 0$ then A is a non-negative operator and if A is a non-negative operator then $a = \hat{A} \geq 0$.

The positivity order for a continuous self-adjoint operator A is equivalent to A being bounded away from zero, i.e. there exists a rational number $q > 0$ such that $(u, Au) > q$ for all $u \neq 0$. This gives the equivalence of the operator $>$ with the $>$ for Dedekind reals restricted to $\mathbb{A}$, because for the latter $a > 0$ means globally that $\exists q \in \mathbb{Q}$ with $q \in 0^+$ and $q \in a_-$. □

If $a = \hat{A}$ and $b = \hat{B}$ then the $\mathbb{R}_{\mathrm{D}}$ distance between a and b is given by the metric $|a - b|$ on $\mathbb{R}_{\mathrm{D}}$. There is another metric on the quantum numbers $\mathbb{A}$, it is given by the number $\widehat{|A - B|}$, where $|C|$ is the operator $|C| = (C^2)^{\frac{1}{2}}$, when C is self-adjoint.

Proposition 9. *The two metrics coincide on* $\mathbb{A}$*, that is,* $\widehat{|A - B|} = |a - b|$*, for all pairs of quantum numbers* $a = \hat{A}$ *and* $b = \hat{B}$*.*

Proof. It is sufficient to let $B = 0$. We will only consider the global sections.

It is well-known, see for example section VI.2.7 of Kato[6], that $\big|(u, Au)\big| \leq (u, |A|u)$ for all u in the domain $D(A) = D(|A|)$, whence $\big|\operatorname{tr} \rho A\big| \leq \operatorname{tr} \rho |A|$ for all ρ in $\mathcal{E}_{\mathcal{S}}$, i.e. $|\hat{A}| \leq \widehat{|A|}$.

As on elements of $\mathbb{R}_{\mathrm{D}}$, $|\hat{A}| = \max(\hat{A}, -\hat{A})$ and $\big|\widehat{|A|}\big| = \widehat{|A|}$. The lower cut of $|\hat{A}|$ is the union of the lower cuts of $\hat{A}$ and of $-\hat{A}$, that is $|\hat{A}|_- = (\hat{A})_- \cup (-\hat{A})_-$, which means that $|\hat{A}|_- \subset \big(\widehat{|A|}\big)_-$. The upper cut of $|\hat{A}|$ is the intersection of the upper cuts of $\hat{A}$ and of $-\hat{A}$, that is $q \in \mathbb{Q}$ belongs to $|\hat{A}|^+$ if q is greater than or equal to both $\hat{A}$ and of $-\hat{A}$. Thus if $q \in \big(\widehat{|A|}\big)^+$ then $q \in |\hat{A}|^+$, therefore $|\hat{A}|^+ \subset \big(\widehat{|A|}\big)^+$.

This shows that $\widehat{|A|} \leq \big|\hat{A}\big|$ and therefore $\widehat{|A|} = \big|\hat{A}\big|$. □

The metric is used to define Cauchy sequences in $\mathbb{R}_{\mathrm{D}}$, this result means that we can define Cauchy Sequences in $\mathbb{A}$ in the same way. In order to ensure uniqueness of the limits of a Cauchy sequence we need a concept of *apartness*, $><$, that is stronger than that of *not equal to*, $\neq$.

Definition 10. *The pair* (a, b) *of Dedekind reals are apart,* $a >< b$*, iff* $(a > b) \vee (a < b)$*.*

Proposition 11. $a >< b$ *iff* $|a - b| > 0$*.*

8. Calculus of Functions

The concept of limit is available for functions: $\mathbb{R}_{\mathrm{D}} \to \mathbb{R}_{\mathrm{D}}$ and differential calculus for such functions can be developed. The definition of the limit of a function $G \in \mathbb{R}_{\mathrm{D}}$ is modeled on the standard definition; the limit of G as x tends to b is L iff $\forall n > 0, \exists\, m > 0, \forall x \in \mathbb{R}_{\mathrm{D}}$; $0 < |x-b| < 1/m \implies \big|G(x) - L\big| < 1/n$.

In this definition both m and n are in $\mathbb{Q}$. The uniqueness of the limit L holds in the sense that if L and M are two limits of G as x tends to b then L and M are apart, $\big|L - M\big| = 0$.

For the requirements of this paper it is enough to consider only polynomial functions. For any $b \in \mathbb{R}_{\mathrm{D}}$ all powers b^n of b are in $\mathbb{R}_{\mathrm{D}}$ for n a natural number because products of continuous functions are continuous. Let both b and $c \in \mathbb{R}_{\mathrm{D}}$, then $bc = cb \in \mathbb{R}_{\mathrm{D}}$ for the same reason. The product, however, is defined only on the intersection of the domains of the continuous functions so that care has to be taken with the extent to which the product exists. Sums of numbers $b + c$ exist to the extents given by the intersection of the extents of b and c. Hence we can construct polynomials

$$F(b) = \sum a_m b^m$$

where the sum is over finitely many terms and the coefficients $a_m \in \mathbb{R}_{\mathrm{D}}$. We can construct power series by defining the convergence of the sequence of partial sums in the metric on $\mathbb{R}_{\mathrm{D}}$.

The derivative of $b^m = mb^{m-1}$ which implies that the derivative of a monomial is defined to the same extent as the monomial. This allows us to obtain the derivatives of polynomials and power series.

We define continuity of a function $F : \mathbb{R}_{\mathrm{D}} \to \mathbb{R}_{\mathrm{D}}$ at b in its domain by the requirement that

$$\text{(1)} \quad \forall n > 0, \exists m > 0, \forall x \in \mathbb{R}_{\mathrm{D}}; \quad 0 < |x - b| < \frac{1}{m} \implies \big|F(x) - F(b)\big| < \frac{1}{n}\,.$$

In this definition we have taken m and n to be in $\mathbb{Q}$.

A function is continuous on an interval I in $\mathbb{R}_{\mathrm{D}}$ iff it is continuous at each number in I.

Interesting phenomena occur in this calculus and they deserve more study. However we have enough structure now to return to the dynamics of quantum mechanics.

9. Comparing the Dynamics

Consider the example of a non-relativistic quantum particle of positive mass μ that moves in a central force field which is derived from a potential function V. We assume that the quantum values $\hat{X}$ of the position of the particle satisfy Newton's equations of motion globally, that is,

$$\mathrm{D}\hat{Q} = \frac{1}{\mu}\hat{P} \tag{2}$$

and

$$\mathrm{D}\hat{P} = F(\hat{Q}) \tag{3}$$

or

$$\mu\,\mathrm{D}^2(\hat{Q}) = F(\hat{Q}) \tag{4}$$

where $\hat{P}$ represents the momentum of the particle and D denotes differentiation with respect to time. F represents the force, it is the negative gradient of the potential function V as a function: $\mathbb{R}_{\mathrm{D}} \to \mathbb{R}_{\mathrm{D}}$.

We will now prove that standard quantum mechanics is a local approximation to a global classical mechanics. More precisely, the theorem states that if the quantum values of the components of position and momentum of a particle are assumed to satisfy Newton's equations of motion globally then the self-adjoint operators corresponding to these values locally satisfy equations that well approximate Heisenberg's equations of motion.

The theorem relates a set of operator equations (Heisenberg's equations) to a set of numerical equations (Newton's equations). The straightforward way to get a numerical equation from an operator equation is to multiply each side of the operator equation by a suitable trace class operator (a state) and then take the trace of each side. The original operator equation has become a family of numerical equations which is labelled by the states.

Recall that Heisenberg's equations for an operator A are

$$\mathrm{D}A = -i[A, H] \tag{5}$$

where H is the Hamiltonian operator of the system and the square bracket denotes the operator commutator.

Therefore, we get Heisenberg's equations of motion, when Q and P represent operators and the Hamiltonian operator is $H = (1/2\mu)P^2 + V(Q)$, to be

$$\mathrm{D}Q = \frac{1}{\mu}P \tag{6}$$

and

$$\mathrm{D}P = i[P,H] = F(Q)\ . \tag{7}$$

If we multiply each side of equations (6) and (7) by the operator ρ and then take the trace of each side we get Heisenberg's numerical equations:

$$\mathrm{D}(\mathrm{tr}\,\rho Q) = \frac{1}{\mu}\,\mathrm{tr}\,\rho P \tag{8}$$

and

$$\mathrm{D}(\mathrm{tr}\,\rho P) = \mathrm{tr}\,\rho(\mathrm{D}P) = \mathrm{tr}\,\rho\big(i[P,H]\big) = \mathrm{tr}\,\rho\big(F(Q)\big)\ . \tag{9}$$

Equation (8) is just Newton's equation (2) at the state ρ. Equation (9) is similar to Newton's equation (3) at the state ρ. The difference between equations (3) and (9) is the same as that which restricts the validity of the result known as Ehrenfest's Theorem, namely, that, in general,

$$\mathrm{tr}\,\rho F(Q) \neq F(\mathrm{tr}\,\rho Q). \tag{10}$$

To simplify the discussion we remove the explicit dependence of the equations (2), (3), (8), and (9) on the operators P and use equations of motion in the form of the second order differential equations:

Newton's equations to the extent W are

$$\mu D^2\big(\hat{Q}\big|_W\big) = F\big(\hat{Q}\big|_W\big)\ . \tag{11}$$

Heisenberg's numerical equations to the extent W are

$$\mu D^2\big(\hat{Q}\big|_W\big) = \widehat{F(Q)}\big|_W\ . \tag{12}$$

It is possible, however, that the difference between the two sides of (10) is small at some state ρ_a and remains small for all states in an open neighborhood of ρ_a. In this case the equations (11) and (12) are approximately the same in that open set. We deduce that Heisenberg's equations approximate Newton's equation in that open set because if (11) holds in an open set W then it holds at each ρ in W.

We claim that for a suitable class of functions F, Heisenberg's equations approximate Newton's equations locally. The localness of the assertion is twofold, we mean that for every standard real number r we can find an open set,W, on which both the number $\hat{Q}\big|_W$ is arbitrarily close to r, and $\widehat{F(Q)}\big|_W$ is arbitrarily close to $F\big(\hat{Q}\big|_W\big)$. The physical interpretation is that if an observer's measurement apparatus is located in a neighbourhood of the position r (in this world with one spatial dimension) then the difference between the accelerations of the particle due to the two forces cannot be distinguished with this apparatus.

The class of suitable functions is defined through the concept of $\mathcal{S}$-continuity.

Definition 12. *We call functions G, $\mathcal{S}$-continuous, if they are real-valued continuous functions on $\mathbb{R}$ such that,for the position operator Q, $G(Q) : \mathcal{S} \to \mathcal{S}$, is continuous in the standard countably normed topology on the Schwartz space $\mathcal{S}$.*

Theorem 4. *If the force F is $\mathcal{S}$-continuous, then given $\varepsilon > 0$, Heisenberg's equations of motion approximate Newton's equations of motion to within ε on each member of a collection of open sets $\{W(r, \varepsilon)\}$ of state space $\mathcal{E}_{\mathcal{S}}$, indexed by the standard real numbers r and $\varepsilon > 0$. That is, on each $\{W(r, \varepsilon)\}$,*

$$\left| \widehat{F(Q)}\big|_W - F\big(\hat{Q}\big|_W\big) \right| < \varepsilon \ . \tag{13}$$

Proof. The idea behind the proof is to find states ρ_r on which $F(\operatorname{tr}\rho_r Q)$ closely approximates $\operatorname{tr}\rho_r F(Q)$, then $F(\hat{Q})$ will be close to $\widehat{F(Q)}$ when $F(\hat{Q})$ is close to $F\big(\hat{Q}(\rho_r)\big)\hat{1\!\!1}$ and $\widehat{F(Q)}$ is close to $\operatorname{tr}\rho_r F(Q)\hat{1\!\!1}$ as Dedekind numbers.

The following pair of lemmas complete the proof. □

Lemma 13. *If Q is a self-adjoint operator which has only absolutely continuous spectrum, then for any real number r in its spectrum we can construct a sequence of pure states $\{\rho_n\}$ such that for the given $\mathcal{S}$-continuous function F,*

$$\lim \operatorname{tr}\rho_n F(Q) = F(r) \ .$$

Proof. Weyl's criterion for the spectrum of self-adjoint operators[9] implies that, for any number r in the spectrum of Q, there exists a sequence of unit vectors $\{u_n\}$, in the domain of Q, such that

$$\lim \big\|(Q - r)u_n\big\| = 0 \ .$$

Take ρ_n to be the projection onto the one dimensional subspace spanned by u_n. It is easy to check that

$$\lim \operatorname{tr}\rho_n Q = r \ .$$

The vectors $\{u_n\}$ can be chosen to be in $\mathcal{S}$. For example, for any positive integer n, let

$$u_n(x) = n^{\frac{1}{2}}\pi^{-\frac{1}{4}} \exp\big(-\tfrac{1}{2}n^2(x-r)^2\big)$$

The sequence $\{u_n\}$ satisfies the requirements of Weyl's lemma for the operator Q and the number r.

Furthermore we can find a sequence of vectors $\{u_n\}$ in $\mathcal{S}$ so that for n large enough the support of $\{u_n\}$ lies in a narrow interval centred on r. Then, by the spectral theorem for Q, the corresponding pure states ρ_n form a sequence so that both

$$\lim \operatorname{tr} \rho_n F(Q) = F(r)$$

and

$$\lim \operatorname{tr} \rho_n Q = r \ .$$

□

From Lemma 13, given a real number r in the spectrum of Q, the $\mathcal{S}$-continuous function F and a real number $\varepsilon > 0$, there exists an integer N such that, for all $j > N$, both

$$\left|\operatorname{tr} \rho_j F(Q) - F(r)\right| < \tfrac{1}{6}\varepsilon$$

and

$$\left|(\operatorname{tr} \rho_j Q) - r\right| < \delta$$

where δ is such that

$$\left|F(r) - F(x)\right| < \tfrac{1}{6}\varepsilon \qquad \text{when} \quad |r - x| < \delta \ . \tag{14}$$

We choose $\rho_r = \rho_j$, for some $j > N$, and deduce that

$$\left|\operatorname{tr} \rho_r F(Q) - F(\operatorname{tr} \rho_r Q)\right| < \tfrac{1}{3}\varepsilon \tag{15}$$

because

$$\begin{aligned} &\left|\operatorname{tr} \rho_r F(Q) - F(\operatorname{tr} \rho_r Q)\right| \\ &\quad = \left|\operatorname{tr} \rho_r F(Q) - F(r) + F(r) - F(\operatorname{tr} \rho_r Q)\right| \\ &\quad \leq \left|\operatorname{tr} \rho_r F(Q) - F(r)\right| + \left|F(r) - F(\operatorname{tr} \rho_r Q)\right| \ . \end{aligned}$$

With this choice of ρ_r,the open set $W(r,\varepsilon)$ can be defined as

$$W(r,\varepsilon) = N(\rho_r, Q, \delta) \cap N(\rho_r, F(Q), \tfrac{1}{3}\varepsilon)$$

where δ satisfies the requirements (14), and for any $\mathcal{S}$-continuous function F and $\varepsilon > 0$ we have that $N(\rho_r, F(Q), \varepsilon)$ is given by

$$N\big(\rho_r, F(Q), \varepsilon\big) = \Big\{\rho\, ; \left|\operatorname{tr} \rho F(Q) - \operatorname{tr} \rho_r F(Q)\right|\Big\} < \varepsilon \ .$$

Lemma 14. *When ρ_r is chosen so that equation* (15) *holds, then for all ρ in $W(r,\varepsilon)$,*

$$\left|\operatorname{tr} \rho F(Q) - F(\operatorname{tr} \rho Q)\right| < \varepsilon \ . \tag{16}$$

That is, for $W = W(r, \varepsilon)$,

$$\left|\widehat{F(Q)}\big|_W - F(\hat{Q}\big|_W)\right| < \varepsilon \ .$$

In the definition of the open neighborhood W, δ may depend upon ρ_r, as well as on F and ε.

Proof. For any pair of states, ρ and ρ_r, we have

$$\begin{aligned}\left|\operatorname{tr}\rho F(Q) - F(\operatorname{tr}\rho Q)\right| &\leq \left|\operatorname{tr}\rho F(Q) - \operatorname{tr}\rho_r F(Q)\right| \\ &+ \left|\operatorname{tr}\rho_r F(Q) - F(\operatorname{tr}\rho_r Q)\right| + \left|F(\operatorname{tr}\rho_r Q) - F(\operatorname{tr}\rho Q)\right| \ .\end{aligned}$$

If ρ is in $N(\rho_r, F(Q), \frac{1}{3}\varepsilon)$ the first summand is $< \frac{1}{3}\varepsilon$, as is the second by choice of ρ_r. The final summand is also because by assumption the function $F : \mathbb{R} \to \mathbb{R}$ is continuous everywhere in the usual topology on $\mathbb{R}$. Because given $\varepsilon > 0$, there exists a $\delta(x_a) > 0$, such that $\left|F(x) - F(x_a)\right| < \frac{1}{3}\varepsilon$, whenever $|x - x_a| < \delta$. Apply this to $x = (\operatorname{tr}\rho Q)$ and $x_a = (\operatorname{tr}\rho_r Q)$.

Therefore, given ρ_r, for any ρ in $N(\rho_r, Q, \delta) \cap N(\rho_r, F(Q), \frac{1}{3}\varepsilon)$, the inequality (16) holds. □

The question remains whether we can construct sufficiently many of these open sets. In general, for a given smooth function F, the family of open sets $\{W(r, \varepsilon)\}$, does not form an open cover of state space $\mathcal{E}_\mathcal{S}$.

However for every standard real number r, and hence for every point in classical coordinate space, all the standard real numbers that lie within δ of r in the standard norm topology on $\mathbb{R}$ also lie in $\{W(r, \varepsilon)\}$ as Cauchy real numbers $\mathbb{R}_\mathrm{C}$. In this sense, the family of open sets $\{W(r, \varepsilon)\}$ covers the classical coordinate space of the physical system.

10. Conclusion

It is important to be clear about the meaning of this result. It involves the comparison of two different theories. Assume that Newton's equations of motion hold for the position and momentum variables of a non-relativistic massive quantum particle when they are expressed in terms of the real numbers $\mathbb{R}_\mathrm{D}$. Assume also that these real numbers $\mathbb{R}_\mathrm{D}$ are given by sheaves of continuous functions over the state space $\mathcal{E}_\mathcal{S}$ and include quantum numbers like $\hat{Q} = \operatorname{tr}\rho Q$ and $\hat{P} = \operatorname{tr}\rho P$ for $\rho \in \{\text{open subsets of } \mathcal{E}_\mathcal{S}\}$, where P and Q are self-adjoint operators on an underlying Hilbert space. Then we can find open subsets of $\mathcal{E}_\mathcal{S}$ such that on each open subset, the restrictions of the functions $\hat{P}$ and $\hat{Q}$ can be reinterpreted as the average values of the operators P and Q which almost satisfy the Heisenberg's equations of motion with the analogous Hamiltonian operator. This result only involves a comparison of the dynamical equations of motion. Comparison of the trajectories requires that the initial data be compatible which leads to further constraints on the allowable trajectories. Nevertheless, there are trajectories that can be described with these real numbers.

References

[1] J. von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, 1955), p. 212.

[2] W. Lawvere, "Quantifiers as Sheaves", Actes Congres Intern. Math. **1**, (1970).

[3] S. MacLane and I. Moerdijk, *Sheaves in Geometry and Logic* (Springer–Verlag, New York, 1994).

[4] L. N. Stout, Cahiers Top. et Geom. Diff. **XVII**, 295 (1976); C. Mulvey, "Intuitionistic Algebra and Representation of Rings" in Memoirs of the AMS **148** (1974).

[5] M. Adelman and J. V. Corbett, Applied Categorical Structures **3**, 79 (1995).

[6] T. Kato, *Perturbation Theory for Linear Operators* (Springer-Verlag, New York, 1966).

[7] R.T. Powers, Commun. Math. Phys. **21**, 85–124 (1971).

[8] A. Inoue, *Tomita–Takesaki Theory in Algebras of Unbounded Operators* (Lecture Notes in Mathematics,1699, Springer, 1998)

[9] W. Reed and B. Simon, *Methods of Mathematical Physics I: Functional Analysis* (Academic Press, New York, 1972).

[10] A. Jaffe, "Dynamics of a cut-off $\lambda\phi^4$-field theory", Ph.D. Thesis, Princeton University, 1965; see also B. Simon, J. Math. Phys. **12**, 140 (1971).

E-mail address: `jvc@maths.mq.edu.au`

E-mail address: `murray@maths.mq.edu.au`

Department of Mathematics,, Macquarie University, Sydney, N.S.W. 2109, Australia