

Article

Contextuality in Classical Physics and Its Impact on the Foundations of Quantum Mechanics

Fritiof Wallentin 

International Center for Mathematical Modeling in Physics and Cognitive Sciences, Linnaeus University, 351 95 Växjö, Sweden; fritiof.persson@lnu.se

Abstract: It is shown that the hallmark quantum phenomenon of contextuality is present in classical statistical mechanics (CSM). It is first shown that the occurrence of contextuality is equivalent to there being observables that can differentiate between pure and mixed states. CSM is formulated in the formalism of quantum mechanics (FQM), a formulation commonly known as the Koopman–von Neumann formulation (KvN). In KvN, one can then show that such a differentiation between mixed and pure states is possible. As contextuality is a probabilistic phenomenon and as it is exhibited in both classical physics and ordinary quantum mechanics (OQM), it is concluded that the foundational issues regarding quantum mechanics are really issues regarding the foundations of probability.

Keywords: Koopman–von Neumann formulation of classical mechanics; contextuality; double-slit experiment; quantum interference; foundations of probability



Citation: Wallentin, F. Contextuality in Classical Physics and Its Impact on the Foundations of Quantum Mechanics. *Entropy* **2021**, *23*, 968. <https://doi.org/10.3390/e23080968>

Academic Editors: Karl Svozil and Andrei Khrennikov

Received: 29 April 2021
Accepted: 21 July 2021
Published: 27 July 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Feynman [1] pointed out that the interference phenomenon as observed in the double-slit experiment (DSE) can be seen as a purely probabilistic phenomenon, as a violation of the formula of total probability (FTP). As one is in the DSE still measuring probabilities in the same way, i.e., as frequencies of occurrence, Feynman concluded that FQM constitutes a probability theory on par with the classical one but only with different rules of computation. By 'classical' Feynman meant that the computational rules equivalent to those in the measure theoretic formulation due to Kolmogorov (KPM) [2] were satisfied. Hence Feynman referred to the result of DSE as FQM violating classical probability theory. However, Ballentine [3] and Koopman [4], though not dismissing the probabilistic origin of the interference phenomenon in the DSE, disagreed with Feynman regarding his claim about it allegedly corresponding to a violation of FTP and classical probability theory. They correctly pointed out that the violation was only seemingly as such, resulting only because of a naive application of FTP. To briefly explain, we consider the DSE in a more generic form referred to as double-slit-type experiments (DSTE). In a DSTE, two observables A and B are considered. For simplicity, A is assumed dichotomous with possible outcomes being a_i , $i = 1, 2$. These observables are measured in three different contexts of measurement $C_{1,2}$, C_1 and C_2 , where a *context of measurement* of an observable is a specification of the physical conditions under which this observable is measured. By construction, each context C_i , $i = 1, 2$ is such that the probability distribution over A , as measured under the circumstances that defines C_i , satisfies

$$\mathbb{P}(A = a_j | C_i) = \delta_{i,j}. \quad (1)$$

Based on the measured probability distributions one then calculates the following function/distribution over b ,

$$\mathbb{P}(B = b | C_{1,2}) - \mathbb{P}(B = b | C_1)\mathbb{P}(A = a_1 | C_{1,2}) + \mathbb{P}(B = b | C_2)\mathbb{P}(A = a_2 | C_{1,2}), \quad (2)$$

where the analogous notation as in (1) has been applied for the different probability distributions. Now, DSE corresponds to the particular DSTE for which A is the observable of the slit passage and B the observable of where on the screen the particle hits, with the respective contexts being:

$C_{1,2}$ Both slits are open.

C_1 Slit 1 is open while slit 2 is closed.

C_2 Slit 2 is open while slit 1 is closed.

In DSE the actual result is a non-zero (2), which is what constitutes the interference phenomenon. If the context is ignored in the notation, this looks like a violation of FTP. However, as Koopman and Ballentine correctly pointed out, there is in reality no violation of FTP, Feynman had only failed to take the contexts into account. Indeed, as Feynman had shown [1], classical mechanics cannot account for the DSE. So DSE is 'unclassical' in that sense. However, in the purely probabilistic sense there is nothing unclassical about it. In this article it will be shown that this typical quantum phenomenon is present in CSM as well. The above argument demonstrates that there is not inherent contradiction in this result.

Note that Koopman and Ballentine only disagreed with Feynman regarding his use of the term 'classical'. Their argument did not dismiss that FQM provided computational rules for calculating the term (2). Hence their views can be combined, which will be done in this article. FQM will be seen as a framework for computing probabilities in which the context dependence of observables is explicitly taken into account. This view of FQM will be made manifest by considering in the framework of *contextual probability*, as developed by Khrennikov in [5]. In it, violations of the type (2) are viewed as measures of contextuality. For a certain class of such violation FQM serves as a particular framework into which they can be represented. More specifically, this is done by considering the generic DSTE together with the *generalized formula of total probability* (generalized FTP), defined as

$$\begin{aligned} \mathbb{P}(B = b|C_{1,2}) &= \mathbb{P}(B = b|C_1)\mathbb{P}(A = a_1|C_{1,2}) + \mathbb{P}(B = b|C_2)\mathbb{P}(A = a_2|C_{1,2}) \\ &+ 2\lambda_b \sqrt{\mathbb{P}(B = b|C_1)\mathbb{P}(A = a_1|C_{1,2})\mathbb{P}(B = b|C_2)\mathbb{P}(A = a_2|C_{1,2})}, \end{aligned} \quad (3)$$

where the *interference coefficient* λ_b serves as a measure of contextuality. A trivial interference coefficient means no contextuality is present in the particular DSTE. On the other hand, if

$$|\lambda_b| \leq 1, \quad (4)$$

then we have contextuality. In [5], it is shown that if a DSTE demonstrates such interference coefficients (4), then it can be represented in FQM. Indeed, in this case, the observables A and B are represented as mutually non-commuting self-adjoint operators. This will also be demonstrated here in Section 2. The purpose of doing so is to show that the occurrence of contextuality is, in FQM, equivalent to being able to physically tell the difference between pure and mixed states. The main result of Section 2 is that the way in which probability distributions transform under time evolution can be used to distinguish between mixed and pure states. It is in this sense that contextuality will be shown to occur in CSM.

In Section 3, KvN [6–8] will be presented and be shown to correspond to CSM. This will be in the sense of all solutions of the Schrödinger equation of KvN via Born's rule also being solutions to the classical Liouville equation. As such, we can apply the statistical-mechanical notion of an equilibrium state as states that are stationary in time and non-equilibrium as states not being stationary. In particular, the states of equilibrium are identified as corresponding to eigenvectors ψ_n of the Liouvillian \hat{T} , i.e., the KvN generator of time, as the associated probability distributions to these transform trivially under time evolution. The associated probability distributions of non-trivial superpositions of such eigenstates,

$$\sum_n c_n \psi_n, \quad (5)$$

are non-stationary. Hence, they can be identified as non-equilibrium states. By applying the principle of maximum Gibbs entropy, we will in fact be able to the main equilibrium states—the microcanonical and canonical ensembles—as corresponding to such eigenstates. Hence, it will be shown that KvN, when applying the principles of statistical mechanics, corresponds to CSM. As non-equilibrium states (4) transform non-trivially under time evolution while a corresponding mixed state

$$\sum_n |c_n|^2 |\psi_n\rangle\langle\psi_n| \quad (6)$$

does not, we are able to tell the difference between pure and mixed states. Hence, contextuality will be shown to be present in CSM, and is hence not only a phenomenon confined to OQM, where ‘OQM’ refers to what is obtained through some variant of canonical quantization. In Section 4, it is shown what impacts on the foundation of quantum mechanics this has. In short, as contextuality is not restricted solely to OQM and as it is quantifiable as a purely probabilistic phenomenon, it is concluded that the issues regarding the foundations of quantum mechanics are really about the foundations of probability.

2. Time Evolution as an Indicator of Contextuality

Let A and B be two quantum observables on some Hilbert space \mathcal{H} . For the sake of the argument, it suffices to assume that A is non-degenerate with an orthonormal eigenbasis $\{\phi_a\}_{a \in \mathbb{A}}$, where the eigenvectors are labeled by their respective eigenvalue a . We impose no such restriction on B . We let ψ_b denote an arbitrary (generalized) eigenvector of B with associated eigenvalue b . By Born’s rule, the probability distribution over B , given initial state ψ , is

$$\mathbb{P}(B = b|\psi) := |\langle\psi_b, \psi\rangle|^2. \quad (7)$$

Analogously, for A ,

$$\mathbb{P}(A = a|\psi) := |\langle\phi_a, \psi\rangle|^2. \quad (8)$$

By utilizing the completeness relation in terms of $\{\phi_a\}_{a \in \mathbb{A}}$, we obtain

$$\mathbb{P}(B = b|\psi) = \left| \sum_{a \in \mathbb{A}} \langle\psi_b, \phi_a\rangle \langle\phi_a, \psi\rangle \right|^2 \quad (9)$$

$$= \sum_{a, a' \in \mathbb{A}} \langle\psi, \phi_{a'}\rangle \langle\phi_{a'}, \psi_b\rangle \langle\psi_b, \phi_a\rangle \langle\phi_a, \psi\rangle \quad (10)$$

$$= \sum_{a \in \mathbb{A}} \mathbb{P}(B = b|\phi_a) \mathbb{P}(A = a|\psi) \quad (11)$$

$$+ \sum_{a \neq a' \in \mathbb{A}} \langle\psi, \phi_{a'}\rangle \langle\phi_{a'}, \psi_b\rangle \langle\psi_b, \phi_a\rangle \langle\phi_a, \psi\rangle \quad (12)$$

$$= \sum_{a \in \mathbb{A}} \mathbb{P}(B = b|\phi_a) \mathbb{P}(B = b|\psi) \quad (13)$$

$$+ 2 \sum_{a < a' \in \mathbb{A}} \Re\{\langle\psi, \phi_{a'}\rangle \langle\phi_{a'}, \psi_b\rangle \langle\psi_b, \phi_a\rangle \langle\phi_a, \psi\rangle\}. \quad (14)$$

Now, there exists a unique number $\theta \in [0, 2\pi)$ such that

$$\begin{aligned} & \langle\psi, \phi_{a'}\rangle \langle\phi_{a'}, \psi_b\rangle \langle\psi_b, \phi_a\rangle \langle\phi_a, \psi\rangle \\ &= \sqrt{\mathbb{P}(A = a'|\psi) \mathbb{P}(B = b|\phi_{a'}) \mathbb{P}(B = a'|\phi_a) \mathbb{P}(A = a|\psi)} e^{i\theta}. \end{aligned} \quad (15)$$

From which it follows that

$$\begin{aligned} & \Re\{\langle\psi, \phi_{b'}\rangle \langle\phi_{b'}, \psi_a\rangle \langle\psi_a, \phi_b\rangle \langle\phi_b, \psi\rangle\} \\ &= \sqrt{\mathbb{P}(A = a'|\psi) \mathbb{P}(B = b|\phi_{a'}) \mathbb{P}(B = a'|\phi_a) \mathbb{P}(A = a|\psi)} \cos(\theta_{b,a,a'}), \end{aligned} \quad (16)$$

with θ being unique if we restrict it to values in $[0, \pi]$. Hence, we end up with

$$\begin{aligned} \mathbb{P}(B = b|\psi) &= \sum_{a \in \mathbb{A}} \mathbb{P}(B = b|\phi_a)\mathbb{P}(A = a|\psi) \\ &+ \sum_{a < a'} \sqrt{\mathbb{P}(A = a'|\psi)\mathbb{P}(B = b|\phi_{a'})\mathbb{P}(B = a'|\phi_a)\mathbb{P}(A = a|\psi)} \cos(\theta_{b,a,a'}) \end{aligned} \quad (17)$$

Hence, it has been shown that FQM indeed satisfies the generalized FTP (3).

Before moving on to the main point of this section, it is worth pointing out that we have implicitly interpreted quantum states as corresponding to contexts of measurement here. This is already evident from the choice of notation in (7) and (8). Note also that the meaning of a conditional probability—as the probability of outcome $A = a$ given the conditions C —does by itself not necessitate that it must satisfy

$$\mathbb{P}(A = a|C) = \frac{\mathbb{P}(A = a \text{ and } C)}{\mathbb{P}(C)}, \quad (18)$$

as it is defined in KPM, i.e., that C can be attributed to some random variable on the same measure space as A . Indeed, Kolmogorov himself [2] pointed out that probability measures

$$\mathbb{P}(\cdot|C) \quad (19)$$

are based on the complex of all experimental conditions and that not all observables are representable as random variables on the same measure space. Hence, neither the notation in (7) and (8) nor the interpretation of quantum states as contexts cause any contradictions.

Now, assume that we have performed some DSTE from which we obtained the probability distributions

$$\mathbb{P}(B = b|C_{1,2}), \quad \mathbb{P}(B = b|C_1), \quad \mathbb{P}(B = b|C_2) \quad \text{and} \quad \mathbb{P}(A = a|C_{1,2}) \quad (20)$$

We wish to represent this DSTE in FQM. This means that A and B are to correspond to self-adjoint operators, which we, for simplicity, assume to be non-degenerate. As (1) holds by construction of the experiment it also means that C_1 and C_2 are to be represented as eigenvectors ϕ_{a_1} and ϕ_{a_2} of A , respectively. In addition, we must have

$$\mathbb{P}(B = b|C_i) = |\langle \psi_b, \phi_{a_i} \rangle|^2. \quad (21)$$

Furthermore, on the basis of $\mathbb{P}(B = b|C_{1,2})$ alone, $C_{1,2}$ may be represented as

$$\psi = \sum_{a \in \mathbb{A}} e^{i\theta_a} \sqrt{\mathbb{P}(A = a|C_{1,2})} \phi_a, \quad (22)$$

up to some choices of the θ_a in $[0, 2\pi)$, or as the mixed state

$$\rho = \sum_{a \in \mathbb{A}} \mathbb{P}(A = a|C_{1,2}) |\phi_a\rangle \langle \phi_a|. \quad (23)$$

However, as

$$\begin{cases} \mathbb{P}(B = b|\psi) = |\langle \psi_b, \psi \rangle|^2 \\ \mathbb{P}(B = b|\rho) = \langle \psi_b, \rho \psi_b \rangle \end{cases}, \quad (24)$$

we obtain via a straightforward calculation that

$$\begin{aligned} &\mathbb{P}(B = b|\psi) - \mathbb{P}(B = b|\rho) \\ &= \cos(\theta_{a_1} - \theta_{a_2} + \theta_{b,a_1,a_2}) \sqrt{\mathbb{P}(B = b|C_1)\mathbb{P}(A = a_1|C_{1,2})\mathbb{P}(B = b|C_2)\mathbb{P}(A = a|C_{1,2})}, \end{aligned} \quad (25)$$

where θ_{b,a_1,a_2} is the number such that

$$\begin{aligned} & \langle \psi_b, \phi_{a_1} \rangle \langle \phi_{a_1}, \psi \rangle \langle \psi_b, \phi_{a_2} \rangle \langle \phi_{a_2}, \psi \rangle \\ & = e^{i\theta_{b,a_1,a_2}} \sqrt{\mathbb{P}(B = b|C_1)\mathbb{P}(A = a_1|C_{1,2})\mathbb{P}(B = b|C_2)\mathbb{P}(A = a_2|C_{1,2})}. \end{aligned} \tag{26}$$

As

$$\mathbb{P}_\rho(B = b|\rho) = \sum_{i=1,2} \mathbb{P}(B = b|C_i)\mathbb{P}(A = a_2|C_{1,2}), \tag{27}$$

by comparing (25) with (3), we see that contextuality means that we are able to differentiate between pure and mixed states.

Recall that in an identification of $C_{1,2}$ as ψ in (22), there is the ambiguity of the complex phases θ_b . This ambiguity can be thought of as the identification being only up to a unitary transformation W of the form

$$W\psi = \sum_{a \in \mathbb{A}} e^{i(\theta_a + \omega_a)} \sqrt{\mathbb{P}(A = a|C_{1,2})} \phi_a. \tag{28}$$

As ρ transforms trivially under all such transformations, we obtain from (25) that

$$\begin{aligned} & \mathbb{P}(B = b|W\psi) - \mathbb{P}(B = b|\rho) \\ & = \cos(\theta'_{a_1} - \theta'_{a_2} + \theta_{b,a_1,a_2}) \sqrt{\mathbb{P}(B = b|C_1)\mathbb{P}(A = a_1|C_{1,2})\mathbb{P}(B = b|C_2)\mathbb{P}(A = a_2|C_{1,2})}, \end{aligned} \tag{29}$$

where we have defined

$$\theta'_{a_i} := \theta_{a_i} + \omega_{a_i}. \tag{30}$$

That is, this difference between how pure and mixed transform under such unitary transformations is equivalent to contextuality. The relevance of this here is that in the case where A corresponds to the generator of time evolution, then the time evolution e^{iAt} is such a transformation W . The equilibrium states of CSM will be shown to in KVN correspond to eigenvectors of the generator of time evolution \hat{T} , and as there certainly exist states in CSM that transform non-trivially under time evolution, it will have been shown that contextuality is not just an OQM phenomenon but occurs in classical physics as well.

Note, it was not discussed above whether there exists mixed states diagonal with respect to some other basis than $\{|a\rangle\}$ such that their respective probability distributions with respect to B transform identically under time evolution. In Appendix A it is shown that there are no such states.

As a side note, the ambiguity of the complex phases in representing contexts as quantum states. Has crucial foundational importance when it comes to FQM and should be seen as a feature rather as something redundant. For instance, in a previous article [9], the author demonstrated that Born's rule can be proved rather than postulated by enforcing the probability to be such that it is invariant under certain such unitary transformations W . Its foundational importance as it relates to Born's rule has also for instance been demonstrated elsewhere, e.g., [10].

3. The Koopman-von Neumann Representation of Classical Mechanics

Let $N \in \mathbb{N}$ be the number of considered particles. We are considering a phase space $\mathcal{P} \simeq \mathbb{R}^{6N}$ together with a set of fixed global canonical coordinates (p, q) . Typically, these canonical coordinates are chosen such that q corresponds to the observable of the position in Cartesian coordinates and p , corresponding to its the conjugate momentum [11] associated with q , defined as

$$p := \frac{\partial L}{\partial \dot{q}}, \tag{31}$$

where L is the given Lagrangian function. We note that p defined in this way does not always correspond to linear momentum, e.g., the charged particle moving in a magnetic field [12]. However, for all Hamiltonian functions of the typical form

$$H(p, q) = \frac{p^2}{2m} + V(q) \tag{32}$$

p does correspond to linear momentum. Furthermore, solely in terms of the formalism of Hamiltonian mechanics, there is nothing special about this particular choice of canonical coordinates. This choice of a coordinate q and its conjugate momentum does, however, play a distinguished role in going from the Lagrangian picture of classical mechanics to the Hamiltonian one via a Legendre transformation [11]. Note that it is not even a necessity to interpret q as a coordinate of position. We will, however, stick to this conventional interpretation of p and q here and consider a Hamiltonian function H which in terms of them has the typical form (31).

To be explicit with the notation, we have defined

$$(p, q) := (p_1^1, p_2^1, p_3^1, \dots, p_1^N, q_2^N, p_3^N, q_1^1, q_2^1, q_3^1, \dots, q_1^N, q_2^N, q_3^N). \tag{33}$$

As we are assuming these to be global coordinates, we can without loss of generality assume that each p_i^n corresponds to the canonical projection $\pi_{3(n-1)+1}$, where

$$\pi_i(x_1, \dots, x_K) := x_k. \tag{34}$$

Similarly, each q_i^n is defined as the projection $\pi_{3(N+n-1)+1}$. The time evolution of this Hamiltonian dynamical structure is given by a Hamiltonian flow

$$U : t \in \mathbb{R} \mapsto U_t \in \text{Symp}(\mathcal{P}), \tag{35}$$

satisfying

$$(-\partial_q H, \partial_q H,) = \frac{d}{d\tau} \Big|_{\tau=0} U_\tau. \tag{36}$$

Now, KvN corresponds to the unitary representation on $L^2(\mathcal{P})$ corresponding to the action

$$\widehat{U}_t \psi = \psi \circ U_{-t}. \tag{37}$$

for every ψ . This representation is constructed by first defining it according to (37) on the space of Schwartz functions ψ on \mathcal{P} . This action (37) acts bijectively on the Schwartz space as all U_{-t} are smooth diffeomorphisms. As the Schwartz space is dense in $L^2(\mathcal{P})$ we may uniquely linearly extend this action (37) to all of $L^2(\mathcal{P})$. Now, by Liouville’s theorem [13], we have

$$dx_t dy_t = dx dy, \tag{38}$$

where

$$(x_t, y_t) := U_{-t}(x, y). \tag{39}$$

Hence,

$$\int \overline{(\widehat{U}_t \psi)}(x, y) (\widehat{U}_t \phi)(x, y) dx dy = \int \overline{\psi}(U_{-t}(x, y)) \phi(U_{-t}(x, y)) dx dy \tag{40}$$

$$= \int \overline{\psi}(x_t, y_t) \phi(x_t, y_t) dx_t dy_t \tag{41}$$

$$= \int \overline{\psi}(x, y) \phi(x, y) dx dy. \tag{42}$$

Thus, each \widehat{U}_t preserves the inner product. \widehat{U} hence defines a unitary representation of the Hamiltonian flow U . Given this representation, we can make some natural identifications.

Consider the multiplication operators

$$\begin{cases} (\hat{p}_i^n \psi)(x, y) = x_i^n \psi(x, y) \\ (\hat{q}_i^n \psi)(x, y) = y_i^n \psi(x, y) \end{cases}, \quad (43)$$

where we, similarly as in (33), have applied the notation

$$(x, y) := (x_1^1, x_2^1, x_3^1, \dots, x_1^N, x_2^N, x_3^N, y_1^1, y_2^1, y_3^1, \dots, y_1^N, y_2^N, y_3^N). \quad (44)$$

Since we have assumed that U acts on points (p, q) having the interpretation of kinetic momentum p and Cartesian position q , we naturally interpret the operator \hat{p}_i^n as corresponding to the observable of the i th component of the kinetic momentum of the n th particle and \hat{q}_i^n as corresponding to the observable of the i th component of the position of the n th particle. Based on this, the operator

$$\hat{E} := H(\hat{p}, \hat{q}) \quad (45)$$

is interpreted as the observable of energy. We have here, for the sake of simplicity, dropped the sub-/supscripts, as will be done generally from now on unless their inclusion is a necessity.

As \hat{U} defines a unitary representation of the Lie group \mathbb{R} through U , we may apply Stone's theorem to obtain a generator of time evolution \hat{T} . From (37), we see that \hat{T} acts as

$$(\hat{T}\psi)(x, y) := i \frac{d}{d\tau} \Big|_{\tau=0} (\hat{U}_\tau \psi)(x, y) \quad (46)$$

$$= i \partial_p H|_{(x,y)} \partial_y \psi(x, y) - i \partial_q H|_{(x,y)} \partial_x \psi(x, y), \quad (47)$$

on all $\psi \in L^2(\mathcal{P})$ in its domain $D(T)$. Notice that we may also write \hat{T} more concisely as

$$\hat{T}\psi = i\{H, \psi\}, \quad (48)$$

with $\{\cdot, \cdot\}$ denoting the Poisson bracket on \mathcal{P} , which in the coordinates (p, q) , read

$$\{f, g\} = \partial_p f \cdot \partial_q g - \partial_q f \cdot \partial_p g. \quad (49)$$

The corresponding Schrödinger equation is, hence,

$$i\partial_t \psi_t = \hat{T}\psi_t. \quad (50)$$

In (48), \hat{T} acts as a derivative. Hence, as \hat{p} and \hat{q} are multiplication operators, it follows that

$$[\hat{T}, \hat{p}]\psi = i\{H, p\psi\} - i\hat{p}\{H, \psi\} \quad (51)$$

$$= i\{H, p\}(\hat{p}, \hat{q})\psi + i\hat{p}\{H, \psi\} - i\hat{p}\{H, \psi\} \quad (52)$$

$$= i\{H, p\}(\hat{p}, \hat{q})\psi \quad (53)$$

$$= -i\partial_q V(\hat{p}, \hat{q})\psi. \quad (54)$$

That is,

$$[\hat{T}, \hat{p}] = -i\partial_q V(\hat{p}, \hat{q}), \quad (55)$$

and similarly that

$$[\hat{T}, \hat{q}] = -i\frac{\hat{p}}{m}. \quad (56)$$

Indeed, these are merely the equations of motion in operator form. Hence, \hat{p} and \hat{q} satisfy the expected dynamical relations in line with their interpretation.

Next, we move on to the relation between the KvN Schrödinger equation and the Liouville equation. By Born's rule,

$$\rho_t := |\psi_t|^2 \quad (57)$$

is a probability distribution over phase space. Assuming that ψ_t is a solution to (50), it follows that

$$\partial_t \rho_t = \psi_t \overline{\partial_t \psi_t} + \overline{\psi_t} \partial_t \psi_t \quad (58)$$

$$= i\psi_t \overline{i\partial_t \psi_t} - i\overline{\psi_t} i\partial_t \psi_t \quad (59)$$

$$= i\psi_t \overline{\widehat{T}\psi_t} - i\overline{\psi_t} \widehat{T}\psi_t \quad (60)$$

$$= i\psi_t \overline{i\{H, \psi_t\}} - i\overline{\psi_t} i\{H, \psi_t\} \quad (61)$$

$$= \psi_t \overline{\{H, \psi_t\}} + \overline{\psi_t} \{H, \psi_t\} \quad (62)$$

$$= \{H, \rho_t\}. \quad (63)$$

This means that ρ_t solves the Liouville equation. We, furthermore, note that if ψ_λ is an eigenvector of \widehat{T} , i.e.,

$$\widehat{T}\psi_\lambda = \lambda\psi_\lambda, \quad (64)$$

then the probability distribution

$$\rho_\lambda := |\psi_\lambda|^2 \quad (65)$$

satisfies

$$\{H, \rho_\lambda\} = \psi_\lambda \overline{\{H, \psi_\lambda\}} + \overline{\psi_\lambda} \{H, \psi_\lambda\} \quad (66)$$

$$= i\psi_\lambda \overline{\widehat{T}\psi_\lambda} - i\overline{\psi_\lambda} \widehat{T}\psi_\lambda \quad (67)$$

$$= i\psi_\lambda \overline{\lambda\psi_\lambda} - i\overline{\psi_\lambda} \lambda\psi_\lambda \quad (68)$$

$$= 0. \quad (69)$$

We hence see that eigenvectors of \widehat{T} correspond to statistical equilibria, while non-trivial superpositions of them correspond to statistical non-equilibria as their induced probability distribution, via Born's rule, transforms non-trivially under \widehat{T} .

To recap, abstractly speaking, we have unitarily represented the Hamiltonian flow U as \widehat{U} such that there exist (non-trivial) operators \widehat{p} and \widehat{q} satisfying (55) and (56), with \widehat{T} being the associated generator of time evolution of \widehat{U} . Note that OQM is also such a representation of U . As such, KvN and OQM are, at this level, equivalent. This equivalence is, however, broken when considering a specific representation of U . This representation theoretic view of the difference/similarity between KvN and OQM will, however, not be formalized further here. However, as KvN itself contains subrepresentations of \widehat{U} in which \widehat{p} and \widehat{q} still act invariantly and non-trivially, we will apply the essence of this representation theoretic view to identify these subrepresentations as 'proper' quantum theories.

For instance, for every $E > 0$, we can construct such a subrepresentation on the Hilbert space $L^2(\mathcal{P}, \nu_E)$, where

$$d\nu_E(x, y) = \delta(E - H(x, y)) dx dy, \quad (70)$$

which corresponds to a subspace of $L^2(\mathcal{P})$ in the sense of the direct integral [14], i.e.,

$$L^2(\mathcal{P}) \simeq \int_{[0, \infty)}^{\oplus} L^2(\mathcal{P}, \nu_E) dE. \quad (71)$$

As the measure ν_E is invariant under U , \widehat{U} is unitarily represented on each $L^2(\mathcal{P}, \nu_E)$ in the same fashion as on $L^2(\mathcal{P})$. Of course, these representations $L^2(\mathcal{P}, \nu_E)$ have the natural interpretation of corresponding to subrepresentations of KvN with fixed total energy E . The measure ν_E corresponds to the microcanonical measure as known from CSM.

Note that ν_E has not yet been identified as a quantum state and so cannot at this point be interpreted as a microcanonical ensemble for CSM. This identification will, however, come next. To do this, we first need to find eigenvectors to \hat{T} in a ‘generic enough’ fashion. We take this to mean that U can be assumed to be periodic with a period τ and that there exists a smooth function Ω on phase space such that

$$\Omega \circ U_{-t} = \Omega + t. \tag{72}$$

Then

$$i\{H, \Omega\} = it. \tag{73}$$

Assuming, furthermore, that all conserved quantities are of the form $g \circ H$, it then follows that a generic eigenvector of \hat{T} is of the form

$$\psi_{n,f} = e^{i\frac{2\pi}{\tau}n\Omega} f \circ H, \tag{74}$$

where $n \in \mathbb{Z}$. Now, in $L^2(\mathcal{P})$, the only restriction of f is that it must be such that ψ_n is in $L^2(\mathcal{P})$. Note, however, that f in the subrepresentations $L^2(\mathcal{P}, \nu_E)$ simply corresponds to a constant, a normalization factor, and hence this degeneracy is removed. However, the degeneracy it represents regains physical meaning in the formalism of direct integrals. The degeneracy induced by the ambiguity of the choice of f can be considered as a section in the sense of direct integrals. That is, $\psi_{n,f}$ is identified as the element in

$$\int_{[0,\infty)}^{\oplus} L^2(\mathcal{P}, \nu_E) dE \tag{75}$$

defined as

$$\psi_{n,f} : E \mapsto \psi_{n,f(E)} := e^{i\frac{2\pi}{\tau}n\Omega} f(E) \in L^2(\mathcal{P}, \nu_E). \tag{76}$$

Now, the amplitude of $\psi_{n,f}$ as a section is given as

$$\langle \psi_{n,f(E)}, \psi_{n,f(E)} \rangle_E. \tag{77}$$

Therefore, Born’s rule tells us that given the equilibrium state $\psi_{n,f}$, the probability distribution over the energy E is

$$\mathbb{P}(E|\psi_{n,f}) = \langle \psi_{n,f(E)}, \psi_{n,f(E)} \rangle_E = |f(E)|^2 \nu_E(\mathcal{P}). \tag{78}$$

We can simply apply the principles of statistical mechanics to find the desired statistical equilibrium state here. For example, by maximizing the Gibbs entropy given a fixed mean energy, thus concluding that (78) must correspond to the canonical measure, i.e.,

$$|f(E)|^2 \nu_E(\mathcal{P}) = \beta e^{-\beta E}. \tag{79}$$

We can also maximize the Gibbs entropy given a fixed energy E . As

$$\mathbb{P}(p, q|\psi_{n,f(E)}) = \frac{\delta(E - H(x, y))}{\nu_E(\mathcal{P})} \tag{80}$$

is the only statistical equilibrium given a fixed energy, it is trivially the distribution that maximizes the Gibbs entropy. As such, we have applied the tools of statistical mechanics to provide the eigenstates of \hat{T} physical interpretation as statistical equilibria. Now, as has already been pointed out, the statistical non-equilibria correspond to non-trivial superpositions of these equilibria. Hence, in accordance with Section 2, we have demonstrated that contextuality exists in CSM as well.

Note, if \hat{T} commuted with both \hat{p} and \hat{q} , then indeed no contextuality would be demonstrable by measuring their associated probability distributions or average as these

would then be invariant under time evolution for all states. However, because of (55) and (56), this is not so for the relevant cases.

4. Conclusions

In Section 2, it was demonstrated that in FQM, the occurrence of contextuality is equivalent to being able to differentiate between pure and mixed states. In turn, it was shown that mixed states that are diagonal in the eigenbasis of the considered generator of time evolution transform trivially under time evolution while the corresponding pure states do not. Hence, one can, through time evolution, demonstrate contextuality.

In Section 3, KvN was developed as a unitary representation of the Hamiltonian flow U on $L^2(\mathcal{P})$ through the action (37). Similarly to OQM, it was noted that in KvN U is represented such that there exist self-adjoint operators \hat{p} and \hat{q} satisfying the operator version of Hamilton's equations of motion (55) and (56), which can be seen as a necessity for interpreting \hat{p} and \hat{q} as, respectively, corresponding to quantum observables of momentum and position. As such, KvN and OQM can be seen as equivalent at this pre-representation level. This representation theoretic view was not formalized further. However, the heuristics of it were used to motivate us to consider subrepresentations of KvN for which this still held. In particular, the subrepresentations considered were those corresponding to constant energy, these were subrepresentations in the sense of the direct integral (71). It was also shown that solutions of the KvN Schrödinger equation via Born's rule are also solutions of the Liouville equation. In particular, the eigenstates of the generator of time evolution of KvN were shown to, in the sense of statistical mechanics, correspond to statistical equilibria, while statistical non-equilibria correspond to non-trivial superpositions of them. The principle maximum entropy was applied to show that the equilibrium states in the subrepresentations of constant energy correspond to the microcanonical ensemble. It was also shown that maximizing the entropy given a fixed average energy gives the equilibrium states in the direct integral (71) that correspond to the canonical ensemble. It was shown that the eigenstates of the KvN generator of time evolution permit a physical interpretation as equilibria in the sense of CSM, and hence the generator of time evolution itself corresponds to an observable. As the difference between equilibrium and non-equilibrium is physically observable, it followed that when the principles of statistical mechanics are applied to KvN, KvN exhibits contextuality.

Now, a possible counterargument for this claim about CSM in the form of KvN exhibiting contextuality is that CSM does not need to be done in FQM, i.e., in the form of KvN. That is, however, besides the point. As explained in the introduction and in Section 2, FQM is merely a computational framework for calculating probabilities of outcomes. In FQM, contextuality takes the form of being able to distinguish between pure and mixed states. Contextuality is, however, still a more general probabilistic phenomena which is quantifiable via generalized FTP. Therefore, it does not matter which formalism is used as long as one is dealing with probability.

It is worth pointing out that it has not been claimed that KvN can replace OQM. KvN cannot, for instance, account for the DSE [7]. Hence, no such replacement can be done. However, no such replacement is needed either. What has been shown is that contextuality, which is often considered as a hallmark of the OQM phenomenon, exists in classical physics as well. As contextuality really is a probabilistic phenomenon, it is, in this article, concluded that the issues regarding the foundations of quantum mechanics are really about the foundations of probability. It is here worth pointing out that Kolmogorov himself pointed out that probabilities are inherently contextual [2].

One might argue against this conclusion of this article by claiming that all classical physics really is reducible to OQM, and that OQM is the 'more fundamental theory'. Based on this, one would then conclude that any contextuality demonstrated in CSM would in reality only be a result of, say, the contextuality induced by \hat{p} and \hat{q} satisfying the canonical commutation relations. However, classical physics does not so simply reduce to OQM through, say, some $\hbar \rightarrow 0$ limit [15]. Therefore, if one still insists on this reductionism, one

would have to consider another form of reduction of a different mathematical form but still carrying the essence of the physical interpretation of $\hbar \rightarrow 0$, the existence of which it is far from obvious and quite possibly impossible. Therefore one cannot so simply disregard the contextuality in CSM demonstrated here as not a 'real' quantum phenomenon. Moreover, this reductionistic view of classical mechanics as being reducible to OQM is common. It is often even applied on the larger landscape of physics, and the natural sciences in general. It is for instance claimed that thermodynamics is just statistical mechanics and that chemistry is just physics. However, as shown in [15,16], and also pointed out by Dyson [17], this type of reductionistic relation does not hold in general. In [15,16] concrete demonstrations of the failure of reductionism are presented. In these references and in [18] it is furthermore discussed what this non-reductionism implies for the hierarchical ordering of theories in terms of a contextified view of emergence versus reduction and ontic versus epistemic.

Lastly we note that Bell-violations may be seen as being induced by contextuality. For in the proof of Bell's theorem [19] the validity of FTP is crucial, something which is even more obvious when considering it in its Wigner form [20]. As such, Bell violations are also ways of demonstrating contextuality [21]. Contextuality is really the hallmark phenomenon of quantum mechanics and, contrary to popular belief, it is present in classical physics as well.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Acknowledgments: I would like to thank Andrei Khrennikov for stimulating discussions and comments. I would also like to thank my wife, this article would not have been possible without her support when writing got tough.

Conflicts of Interest: The author declares no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

| | |
|------|--|
| CSM | Classical statistical mechanics |
| DSE | Double-slit experiment |
| DSTE | Double-slit-type experiment |
| FTP | Formula of total probability |
| FQM | Formalism of quantum mechanics |
| KvN | Koopman–von Neumann formalism of classical mechanics |
| OQM | Ordinary quantum mechanics |

Appendix A

It is here shown that in a pure state

$$\psi := \sum_a \sqrt{\rho_{a,a}} e^{i\theta_a} |a\rangle \quad (\text{A1})$$

there exists no associated mixed state

$$\rho := \sum_{a,a'} \rho_{a,a'} |a\rangle \langle a'| \quad (\text{A2})$$

whose respective probability distribution with respect to B transforms identically under time evolution.

As ρ is a state, its trace is equal to one and it is self-adjoint, i.e.,

$$\sum_a \rho_{a,a} = 1, \tag{A3}$$

and for all a, a' ,

$$\rho_{a,a'}^* = \rho_{a',a}. \tag{A4}$$

As it, in addition, is mixed,

$$\text{Tr}\{\rho^2\} \neq 1. \tag{A5}$$

As

$$\text{Tr}\{\rho^2\} = \sum_{a,a'} |\rho_{a,a'}|^2 \tag{A6}$$

$$= \underbrace{\left(\sum_a \rho_{a,a}\right)^2}_{=1} - 2 \sum_{a < a'} \left(\rho_{a,a} \rho_{a',a'} - |\rho_{a,a'}|^2\right) \tag{A7}$$

$$= 1 - 2 \sum_{a < a'} \left(\rho_{a,a} \rho_{a',a'} - |\rho_{a,a'}|^2\right), \tag{A8}$$

ρ is mixed if and only if

$$\sum_{a < a'} \left(\rho_{a,a} \rho_{a',a'} - |\rho_{a,a'}|^2\right) \neq 0. \tag{A9}$$

Under a time evolution of duration t , ρ transforms to

$$\rho_t = \sum_{a,a'} \rho_{a,a'} e^{i(a-a')t} |a\rangle\langle a'| \tag{A10}$$

and a corresponding pure state ψ to

$$\psi_t = \sum_a \sqrt{\rho_{a,a}} e^{i(\theta_a + at)} |a\rangle. \tag{A11}$$

A straightforward calculation shows that

$$\begin{aligned} &\mathbb{P}(b|\rho_t) - \mathbb{P}(b|\psi_t) \\ &= 2 \sum_{a < a'} \text{Re}\left\{ \left(\rho_{a,a'} - \sqrt{\rho_{a,a} \rho_{a',a'}} e^{i(\theta_a - \theta_{a'})}\right) e^{i(a-a')t} \langle b|a\rangle \langle a'|b\rangle \right\} \end{aligned} \tag{A12}$$

As

$$\mathbb{P}(b|\rho_t) - \mathbb{P}(b|\psi_t) = 0 \tag{A13}$$

if and only if

$$\sum_b f(b) (\mathbb{P}(b|\rho_t) - \mathbb{P}(b|\psi_t)) = 0 \tag{A14}$$

for all functions f , it follows that this can only hold if

$$\text{Re}\left\{ \left(\rho_{a,a'} - \sqrt{\rho_{a,a} \rho_{a',a'}} e^{i(\theta_a - \theta_{a'})}\right) e^{i(a-a')t} \langle b|a\rangle \langle a'|b\rangle \right\} = 0. \tag{A15}$$

As this must be true for all t , we can pick a t such that

$$e^{i(a-a')t} \langle b|a\rangle \langle a'|b\rangle \tag{A16}$$

is real valued. Hence, (A15) implies that

$$\operatorname{Re}\left\{\rho_{a,a'} - \sqrt{\rho_{a,a}\rho_{a',a'}}e^{i(\theta_a - \theta_{a'})}\right\} = 0. \quad (\text{A17})$$

In turn, this implies that

$$|\rho_{a,a'}|^2 = \rho_{a,a}\rho_{a',a'}, \quad (\text{A18})$$

which contradicts (A9). Hence, ρ and ψ cannot transform similarly under time evolution.

References

1. Feynman, R. The Concept of Probability in Quantum Mechanics. *Berkeley Symp. Math. Stat. Probabil.* **1951**, *2*, 533–542.
2. Kolmogorov, A. *Foundations of Theoru of Probability*; Chelsea Publishing Company: New York, NY, USA, 1956.
3. Ballentine, L. Probability Theory in Quantum Mechanics. *FTPH* **1989**, *24*, 31–42.
4. Koopman, B. Quantum Theory and the Foundations of Probability. *Appl. Probabil.* **1955**, 97–102.
5. Khrennikov, A. *Contextual Approach to Quantum Formalism*; Springer Science: Berlin/Heidelberg, Germany, 2009.
6. Koopman, B.O. Hamiltonian Systems and Transformations in Hilbert Space. *Proc. Natl. Acad. Sci. USA* **1931**, *17*, 315–318. [[CrossRef](#)]
7. Mauro, D. Topics in Koopman-von Neumann Theory. Ph.D. Thesis, University of Trieste, Trieste, Italy, 2002.
8. Frank Wilczek's Website. Available online: <http://frankwilczek.com/2015/koopmanVonNeumann02.pdf> (accessed on 16 April 2021).
9. Wallentin, F. Born's Rule from the Principle of Unitary Equivalence. *Rep. Math. Phys.* **2020**, *85*, 209–225. [[CrossRef](#)]
10. Skilling, J.; Knuth, K. The Arithmetic of Uncertainty unifies Quantum Formalism and Relativistic Spacetime. *arXiv* **2020**, arXiv:2104.05395v1.
11. Goldstein, H.; Poole C.; Safko, J. *Classical Mechanics*, 3rd ed.; Addison Wesley: Boston, MA, USA, 2000.
12. Ballentine, L. *Quantum Mechanics: A Modern Development*; World Scientific Publishing Co. Pte. Ltd.: Singapore, 1998.
13. Arnold, V.I. *Mathematical Methods of Classical Mechanics*; Springer: Berlin/Heidelberg, Germany, 1989.
14. Hall, B.C. *Quantum Theory for Mathematicians*; Springer: Berlin/Heidelberg, Germany, 2013.
15. Primas, H. Emergence in Exact Natural Sciences. *Acta Polytechn. Scand.* **1998**, *97*, 83–98.
16. Bishop, R.; Atmanspacher, A. Contextual Emergence in the Description of Properties. *Found. Phys.* **2006**, *36*, 1753–1777. [[CrossRef](#)]
17. Dyson, F. *The Scientist as Rebel*; The New York Review of Books: New York, NY, USA, 2016.
18. Atmanspacher, H.; Kronz, F. Relative Onticity. In *On Quanta, Mind and Matter. Fundamental Theories of Physics (An International Book Series on The Fundamental Theories of Physics: Their Clarification, Development and Application)*; Springer: Berlin/Heidelberg, Germany, 1999; Volume 102, pp. 273–297.
19. Bell, J. On the Einstein Podolsky Rosen Paradox. *Physics* **1964**, *1*, 195–200. [[CrossRef](#)]
20. Khrennikov, A. *Probability and Randomness. Quantum versus Classical*; Imperial College Press: London, UK, 2016.
21. Khrennikov, A. Get Rid of Nonlocality from Quantum Physics. *Entropy* **2019**, *1*, 195. [[CrossRef](#)] [[PubMed](#)]