

Framework for generalized Hamiltonian systems through reasonable postulates

Libo Jiang,^{1,*} Daniel R. Terno,^{2,†} and Oscar Dahlsten^{3,4,5,‡}

¹*Southern University of Science and Technology (SUSTech), Shenzhen 518055, China*

²*Department of Physics and Astronomy, Macquarie University, Sydney, New South Wales 2109, Australia*

³*Department of Physics, City University of Hong Kong, Tat Chee Avenue, Kowloon, Hong Kong SAR, China*

⁴*Shenzhen Institute for Quantum Science and Engineering and Department of Physics, Southern University of Science and Technology, Shenzhen 518055, China*

⁵*Institute of Nanoscience and Applications, Southern University of Science and Technology, Shenzhen 518055, China*



(Received 13 November 2023; accepted 23 February 2024; published 21 March 2024)

Hamiltonian mechanics describes the evolution of a system through its Hamiltonian. The Hamiltonian typically also represents the energy observable, a Noether-conserved quantity associated with the time invariance of the law of evolution. In both quantum and classical mechanics, Hamiltonian mechanics demands a precise relationship between time evolution and observable energy, albeit using slightly different terminology. We distill basic conditions satisfied in both quantum and classical mechanics, including canonical coordinate symmetries and inner product invariance. We express these conditions in the framework of generalized probabilistic theories, which includes generalizing the definition of energy eigenstates in terms of time-invariant properties of the Hamiltonian system. By postulating these conditions to hold, we derive a unified Hamiltonian system model. This unified framework describes quantum and classical mechanics in a consistent language, facilitating their comparison. We moreover discuss alternative theories: an equation of motion given by a mixture of commutation relations, an information-restricted version of quantum theory, and Spekken's toy theory. The findings give a deeper understanding of the Hamiltonian in quantum and classical theories and point to several potential research topics.

DOI: [10.1103/PhysRevA.109.032218](https://doi.org/10.1103/PhysRevA.109.032218)

I. INTRODUCTION

Hamiltonian mechanics, whether in classical or quantum cases, describes the time evolution of systems through their Hamiltonian [1–3],

$$\frac{\partial \rho(q, p)}{\partial t} = \{H, \rho\}, \quad \frac{\partial \hat{\rho}}{\partial t} = \{\hat{H}, \hat{\rho}\}_{\hbar}. \quad (1)$$

Here H is the classical Hamiltonian and \hat{H} its quantum counterpart. The classical Liouville density ρ and the quantum mechanical density operator $\hat{\rho}$ evolve via their Poisson bracket with the Hamiltonian. In the quantum case the bracket is defined as $\{\hat{A}, \hat{B}\}_{\hbar} := [\hat{A}, \hat{B}]/(i\hbar)$ [3]. The time-independent Hamiltonian represents energy, which is a conserved dynamical quantity or quantum observable corresponding to time translation symmetry according to Noether's theorem [4]. Therefore, the Hamiltonian formalism establishes a connection between energy and the time evolution.

Work in the foundations of quantum mechanics led to the development of the theoretical framework of generalized probability theories (GPTs). Systems' states in a GPT are represented as probability vectors (or other real vectors) from which probabilities can be extracted [5–9]. The

state depends on the system preparation and any subsequent dynamical transformations. A key goal of GPTs is to understand the structure of quantum theory, particularly which elements necessarily follow from its probabilistic nature, and to elucidate the relations between classical and quantum mechanics [5–7,10]. Classical and quantum theories, as well as classical-quantum hybrid models [11,12], appear as special cases.

Using the GPT framework, the notion of a Hamiltonian has been generalized in finite dimensions [13,14]. There is also a long-running interest in unifying and comparing quantum and classical mechanics [12,15], as well as efforts to explore potential new theories via the creation of toy theories that are not classical or quantum [16–18]. Taken together, these results give hope that a more fully generalized Hamiltonian mechanics can be created, encompassing classical and quantum mechanics and more.

Here we accordingly aim to create a framework for generalized Hamiltonian systems, giving full details of the results in Ref. [19]. Figure 1 depicts the interrelations between the concepts we discuss in this work. These are foundational efforts, strengthening our understanding of quantum and classical mechanics and what may lie beyond.

We generalize the phase space representation of quantum mechanics [20,21], specifically the Wigner function [21–24]. The phase space formalism models quantum and classical mechanics in a similar manner. Quantum states are described by real quasiprobability distributions of position and momentum

*11930020@mail.sustech.edu.cn

†daniel.terno@mq.edu.au

‡oscar.dahlsten@cityu.edu.hk

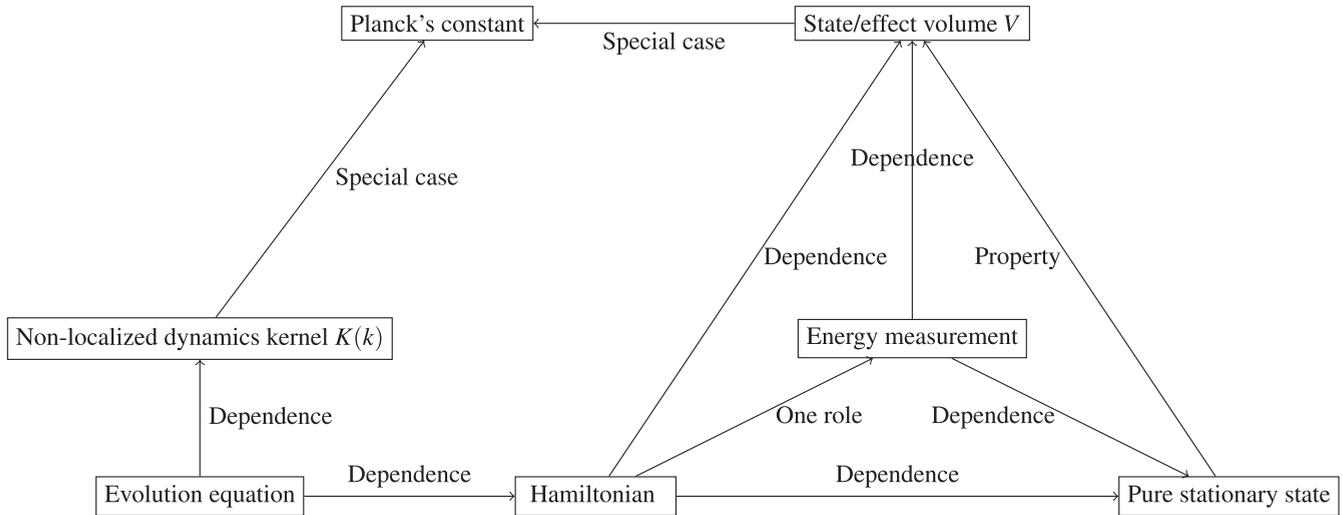


FIG. 1. Relations between our key concepts.

$W(q, p)$. We make use of the fact that the Wigner function framework may be viewed as a continuous variable generalized probabilistic theory [25].

We establish a Hamiltonian formalism for GPTs based on postulates that are satisfied by both quantum and classical theories. Through these postulates, we obtain a generalized measurement of energy and a generalized equation of motion. A specific theory, such as quantum or classical mechanics, is obtained by completing the system of axioms in a way that is described below.

A fundamental aspect of this framework is the observation that both quantum and classical evolutions can be generated by pure stationary states, which then serve as the generalized energy eigenstates within our framework. We define these to be the most pure stationary states, a definition that leads to simple expressions for the generalized mechanics. The set of stationary states is a time-independent characteristic of the system and encodes the key part of the time-independent evolution rule.

The Planck constant plays a crucial role in distinguishing between quantum and classical theories. In the generalized framework, we find that \hbar has two distinct roles. One role pertains to the uncertainty of the state, which we refer to as the *state or effect volume*, while the other role appears in the equation of motion via a nonlocalized dynamics integration kernel. Although these quantities coincide in quantum and classical mechanics, they may have different values in generalized theories, such as in epistemically restricted classical theory and quantum mechanics with a particular information restriction.

Furthermore, an intriguing finding beyond quantum and classical mechanics is an equation of motion. It is given by a series of commutators with the Hamiltonian for which each commutator can be different, resulting in a nonassociative algebra. This evolution rule, derived here from reasonable postulates, happens to provide a restricted version of the “generalized Moyal bracket” proposed in Ref. [18].

As an application of our model, we demonstrate how the concept of state volume helps to understand the possibility of chaos in the sense of strong sensitivity to initial state pertur-

bations [26,27]. The contrasting chaotic behaviours observed in quantum and classical cases can be attributed to differences in the volumes of pure states. As well, despite developing the model within the framework of continuous phase space, we discover that certain concepts, such as state or effect volume, can also be extended to discrete systems, such as Spekken’s toy model [16].

II. PRELIMINARIES

In this section we summarize some key results central to this paper from generalized probabilistic theories, the phase space formalism, and the action-angle formalism of classical mechanics.

A. Phase space representation

Classical models that we consider describe nonconstrained systems with a finite number of n degrees of freedom. Its states and (the algebra of) observables are smooth functions on the phase space \mathfrak{P} [1–3], which is then a $2n$ -dimensional symplectic manifold that is a cotangent bundle of the configuration space. (Some mathematical aspects of the phase space formalism as summarized in Appendix A). The local coordinates on \mathfrak{P} are $z = (q, p)$, where q are the generalized coordinates and the p are the canonical conjugate momenta. Pure states represent the perfect knowledge of position and momentum and are thus δ distributions, $\rho_{z_0} = \delta(q - q_0)\delta(p - p_0)$. In situations of incomplete knowledge about a system’s state, like in statistical mechanics, the state is represented as a probability (Liouville) density $\rho(q, p)$.¹

The evolution of the state is generated by the Hamiltonian $H(q, p)$ according to the Hamilton equations $\dot{q} = \frac{\partial H}{\partial p}$ and $\dot{p} = -\frac{\partial H}{\partial q}$. A probability density $\rho(q, p)$ then evolves according to

¹Notice that the density matrix $\hat{\rho}$ has the unit of probability instead of density, while we choose ρ to represent probability densities.

the Liouville equation,²

$$\frac{\partial \rho(q, p)}{\partial t} = \frac{\partial \rho}{\partial p} \frac{\partial H}{\partial q} - \frac{\partial \rho}{\partial q} \frac{\partial H}{\partial p} = \{H, \rho\} = -H \Lambda \rho, \quad (2)$$

where H is the Hamiltonian and $\{, \}$ is the Poisson bracket.

The operator (symplectic matrix) $\Lambda := \overleftarrow{\frac{\partial}{\partial p}} \overrightarrow{\frac{\partial}{\partial q}} - \overleftarrow{\frac{\partial}{\partial q}} \overrightarrow{\frac{\partial}{\partial p}}$ provides an alternative form of the Poisson bracket. The right and left arrows on the operators mean that the derivative will act on the right and left side's function: $f \overleftarrow{\frac{\partial}{\partial p}} \overrightarrow{\frac{\partial}{\partial q}} g = \frac{\partial f}{\partial p} \frac{\partial g}{\partial q}$.

The phase space formalism can be generalized to quantum states. There are different approaches to quantum phase spaces [20]. One of the most common versions is the Wigner function $W(q, p)$ [21–24,28]. Wigner functions are real functions of canonical coordinates (q, p) . They are obtained via the Wigner transform of density matrices

$$W(q, p) = \frac{1}{\pi \hbar} \int dy \langle q - y | \hat{\rho} | q + y \rangle e^{2ipy/\hbar}, \quad (3)$$

with an obvious extension to n degrees of freedom. Consider the eigenstates of a simple harmonic oscillator with $H = q^2 + p^2$ as an example. Their Wigner functions are given by

$$W_{E_n}(q, p) = \frac{(-1)^n}{\pi \hbar} L_n \left[\frac{2(q^2 + p^2)}{\hbar} \right] e^{-(q^2 + p^2)/\hbar}, \quad (4)$$

where L_n are the Laguerre polynomials. Wigner functions are normalized,

$$\int dq dp W(q, p) = 1, \quad (5)$$

but not necessarily positive. Some example distributions are depicted in Fig. 2.

The Born rule is reproduced by the following inner product:

$$\text{Tr}(\hat{\rho}_1 \hat{\rho}_2) = h \int W_1 W_2 dq dp, \quad (6)$$

where W_1, W_2 are Wigner functions corresponding to $\hat{\rho}_1, \hat{\rho}_2$ and h is Planck's constant.

Unlike classical probability densities, many Wigner functions have some small areas with negative values and are thus called *quasi* probability distribution. The probabilities of any allowed measurement outcome are nevertheless positive, which can be understood as the uncertainty principle rescuing positivity by banning measurements that would single out small phase space regions.

The inverse map that takes a phase space function to an operator is called the Weyl transform [29]. The Weyl-Wigner transforms provide a mathematical method that connects phase space functions and noncommutative operators:

$$\text{Wigner}\{\hat{A}\}(q, p) = 2 \int dz e^{i \frac{2pz}{\hbar}} \langle q - z | \hat{A} | q + z \rangle. \quad (7)$$

²Consider time $t \rightarrow t + dt$, where dt is sufficiently small. Then, under the tangential approximation, $\rho(q, p, t) \rightarrow \rho(q + \frac{\partial q}{\partial t} dt, p + \frac{\partial p}{\partial t} dt, t + dt)$, $d\rho = \frac{\partial \rho}{\partial q} \frac{\partial q}{\partial t} dt + \frac{\partial \rho}{\partial p} \frac{\partial p}{\partial t} dt + \frac{\partial \rho}{\partial t} dt$. By Liouville's theorem, $d\rho = 0$. Combining this with Hamilton's equations gives the Liouville equation.

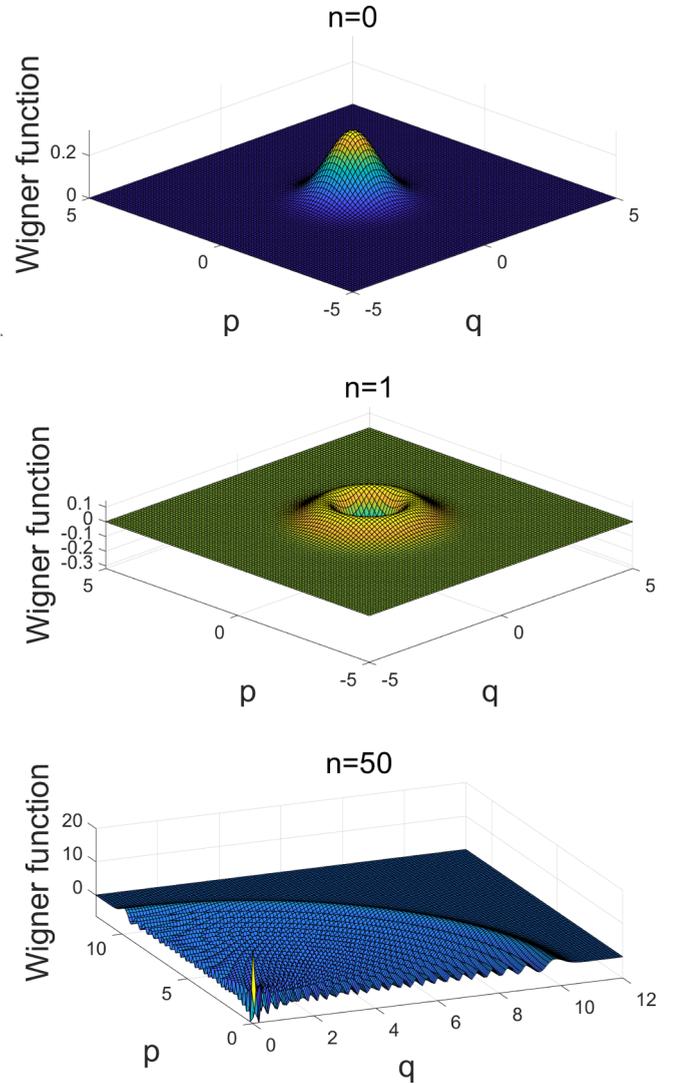


FIG. 2. Wigner functions for the $n = 0, 1, 50$ eigenstates of a simple harmonic oscillator. $H = k(q^2 + p^2)$, where k is an arbitrary positive constant with the dimension of $[\frac{1}{t}]$. Both q and p are plotted in units of $\hbar^{\frac{1}{2}}$, and the Wigner function is plotted in units of $\frac{1}{\hbar}$.

In the other direction,

$$\hat{A} = \frac{1}{4\pi^2 \hbar^2} \int \text{Wigner}\{\hat{A}\}(q, p) e^{i \frac{a(q-\hat{q}) + b(p-\hat{p})}{\hbar}} dq dp da db, \quad (8)$$

where $\text{Wigner}\{\cdot\}$ labels the Wigner transform.³

One sees that the Wigner function of Eq. (7) is the Wigner transformation of density matrices of Eq. (3) with an extra factor: $W(q, p) = \frac{1}{h} \text{Wigner}\{\hat{\rho}\}$. This is consistent with the Wigner function having the unit of probability density (probability per phase space area), whereas the Wigner transformation does not change the dimensionless unit of the density matrix.

³The Weyl-Wigner transform implies a one-to-one correspondence between real power functions $q^n p^m$ with the same power function of operators in a certain order. This order is called the Weyl ordering [23]: $q^n p^m \leftrightarrow \frac{1}{2^n} \sum_{i=0}^n C_n^i \hat{q}^{n-i} \hat{p}^m \hat{q}^i = \frac{1}{2^m} \sum_{j=0}^m C_m^j \hat{p}^{m-j} \hat{q}^n \hat{p}^j$.

The (noncommutative) product of operators appears in the Wigner function representation as

$$\text{Wigner}\{\hat{A}\hat{B}\} = \text{Wigner}\{\hat{A}\} \star \text{Wigner}\{\hat{B}\},$$

where

$$\star := \exp\left(-\frac{i\hbar}{2}\Lambda\right), \quad (9)$$

and \star is the Moyal (star) product.

One can deduce the evolution of the Wigner function [23] by applying the Wigner transform and Moyal product to $i\hbar\frac{\partial\rho}{\partial t} = [H, \hat{\rho}]$, which gives

$$\frac{\partial W}{\partial t} = \frac{2}{\hbar}W(q, p) \sin\left(\frac{\hbar}{2}\Lambda\right)H(q, p), \quad (10)$$

where $H(q, p)$ is the Hamiltonian in phase space obtained by the Wigner transform. When $\hbar \rightarrow 0$, Eq. (10) transforms to the classical Poisson bracket of Eq. (2): $\frac{\partial\rho}{\partial t} = \rho\Lambda H = -\{\rho, H\}_{P.B.}$.

The time evolution can be written in another way (the Wigner transport equation) for the case of $H = \frac{p^2}{2m} + V(q)$ [23]:

$$\frac{\partial W}{\partial t} = -\frac{p}{m}\frac{\partial W}{\partial q} + \int djW(q, p+j)J(q, j),$$

where $J(q, j) = \frac{i}{\pi\hbar^2} \int dy[V(q+y) - V(q-y)]e^{-2ijy/\hbar}$.

$$(11)$$

The $-\frac{p}{m}\frac{\partial W}{\partial q}$ term is contributed by the kinetic energy term of the Hamiltonian, while $\int djW(q, p+j)J(q, j)$ is contributed by the potential energy term. The kinetic term is the same as in classical mechanics, but the potential term contains an integral over all momenta, implying that the distribution ‘‘jumps’’ in the momentum direction. The jumping in Eq. (11) is associated with the infinite orders of derivatives in Eq. (10), since an infinite-order Taylor expansion enables the expansion to an arbitrary distance with arbitrary precision. We will return to this point in Sec. VI B and Appendix B.

B. Action-angle variables

Action-angle variables are useful in the analysis of classical systems [1,2]. The *action*, also called abbreviated action I is a number associated with an orbit defined as [1]

$$I = \frac{1}{2\pi} \oint p dq. \quad (12)$$

This quantity gives the phase space volume (enclosed by the orbit up to $\frac{1}{2\pi}$).⁴

The *angle* variable θ specifies where the phase space point is along the orbit, as illustrated by the simple harmonic oscillator case depicted in Fig. 3. (θ, I) can be obtained, at least in certain cases, for a given Hamiltonian via a canonical transformation from (q, p) [1]. More specifically, the

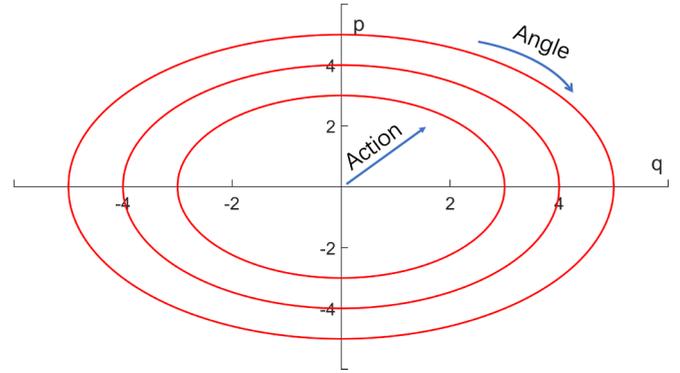


FIG. 3. Action contour lines of a harmonic oscillator whose $H = k(q^2 + p^2)$. Both q and p are plotted in units of $\hbar^{\frac{1}{2}}$. Their actions from outside to inside are $\frac{25\hbar}{2\pi}$, $\frac{16\hbar}{2\pi}$, $\frac{9\hbar}{2\pi}$. The action contour lines correspond to orbits in phase space. The angle alone changes during the evolution.

Liouville-Arnold theorem says that an action-angle coordinate system exists for all completely integrable systems [1]. Harmonic oscillators are prominent examples of completely integrable systems.

The canonical equation for action-angle coordinates is

$$\frac{\partial I}{\partial t} = -\frac{\partial H}{\partial \theta} = 0, \quad \frac{\partial \theta}{\partial t} = \frac{\partial H}{\partial I}, \quad (13)$$

where H depends only on action. While the action is invariant for the time-independent Hamiltonian, the angle θ evolves at a constant speed: $\dot{\theta}(I, t) = \frac{\partial H}{\partial I}$, indicating the phase of the periodic motion.

By the Liouville equation, Eq. (2), the evolution of the distribution f under these coordinates is given by

$$\frac{\partial f}{\partial t} = -\frac{\partial f}{\partial \theta} \frac{\partial H}{\partial I}. \quad (14)$$

C. Generalized probabilistic theories

Generalized probabilistic theories (GPTs) express the idea that at the operational level, only statistics of measurement outcomes conditional on preparations and measurement procedures form the empirical basis of a theory, in contrast with indirect concepts like force. Therefore, GPTs are also called operational probabilistic theories. GPTs associate experiments on a system with real vectors, e.g., the probability vectors \vec{f} corresponding to the individual measurements and outcomes [6,7,30], such as

$$\vec{f} = \begin{bmatrix} \vdots \\ P(j|f, M_i) \\ P(j+1|f, M_i) \\ \vdots \\ P(k|f, M_{i+1}) \\ P(k+1|f, M_{i+1}) \\ \vdots \end{bmatrix}, \quad (15)$$

where $P(j|f, M_i)$ represents the probability of the j th outcome of the i th measurement on a state f .

⁴The historically significant ‘‘Sommerfeld’’ quantization condition is that $2\pi I = nh$ where n is an integer.

Mixed states are represented by the linear combinations of state vectors. For example, $\vec{f}' = p\vec{f}_1 + (1-p)\vec{f}_2$, $p \in [0, 1]$ represents a probabilistic mixture of state \vec{f}_1 and \vec{f}_2 with weights p and $1-p$ respectively. As a result state spaces are always convex sets. Given a set of states, the states that cannot be obtained as a mixture of other states are, as in quantum and classical theories, called *pure* states.

The states are assumed to transform linearly. The transformations are modeled as real matrices T such that they respect probabilistic mixtures: $T(p\vec{f}_1 + (1-p)\vec{f}_2) = pT\vec{f}_1 + (1-p)T\vec{f}_2$. The set of allowed transformations must be such that allowed states are taken to allowed states by them. If the inverse matrix T^{-1} exists and is an allowed transform one says the transformation T is *reversible*.

The measurements are represented by the dual elements of state space in a certain sense. One can introduce linear operators $e_{M,j}$ so that $e_{M,j}(f) = P(j|f, M)$ gives the probability of the j th outcome in measurement M . These operators are called *effects* in GPTs. For instance, the Born rule $P = \text{Tr}(\hat{E}_j \hat{\rho})$ in quantum mechanics can be understood as an effect $\text{Tr}(\hat{E}_j \dots)$ applied on the state $\hat{\rho}$. Every measurement always ends with a result, which requires the set of effects corresponding to the measurement to be *complete*. The completeness of effects means that for arbitrary states f and measurement M ,

$$\sum_j e_{M,j}(f) = 1. \quad (16)$$

This work treats the phase space formalism as a continuous variable case of a general GPT formalism, following the connection between these two frameworks established in Ref. [25]. A generalized phase-space-like formalism was also employed recently in Ref. [17].

In the following we assume that in a given GPT valid states are normalized functions (or distributions) on \mathfrak{P} . Convex combinations of states are also states. As part of specifying a given theory the state space will in general have further restrictions.

In GPTs, effects are linear functionals of states, such that probabilistic mixtures of states lead to corresponding probabilistic mixtures of measurement outcomes. In phase space, effects are described as

$$P(i|f) = e_i(f) = \int h_i f dq dp, \quad (17)$$

where h_i is some function of q, p .

The completeness condition $\sum_i e_i(f) = 1$ for arbitrary f requires

$$\sum_i h_i = 1. \quad (18)$$

If only a finite region $\mathfrak{D} \subset \mathfrak{P}$ is of concern, the completeness condition becomes

$$\sum_i h_i = 1_{\mathfrak{D}}, \quad (19)$$

where $1_{\mathfrak{D}}$ is a function that equals one when $(q, p) \in \mathfrak{D}$ and zero otherwise.

For continuous effects labeled by a continuous variable μ , the probability of the outcome falling into a continuous

interval $(\mu, \mu + d\mu)$ is $dP(\mu; d\mu|f) = \rho(\mu|f)d\mu$, where $\rho(\mu|f)$ is the probability density for the outcome μ given the state f . The most general expression is

$$\rho(\mu|f) = \int f(q, p) h_{\mu}(z) dq dp. \quad (20)$$

For example, classical (sharp) phase space localization has $\mu = (q, p) \in \mathfrak{P}$. The state is given by the Liouville density $f = \rho(q, p)$. The probability of being within the volume $dq_0 dp_0$ around the point (q_0, p_0) in \mathfrak{P} is $dP = \rho(q_0, p_0) dq_0 dp_0$. In this case, $h_{(q_0, p_0)}(q, p) = \delta(q - q_0) \delta(p - p_0)$.

GPTs also include the conditional update rule after measurements

$$f \xrightarrow{i} g_{(f,i)} \quad (21)$$

of the measured state f if the outcome i was registered. For example, the sharp classical measurement with the outcome (q_0, p_0) leads to the update $\rho \rightarrow \delta(q - q_0) \delta(p - p_0)$. In quantum mechanics, von Neumann measurement collapses the wave function to the projector that describes the effect, while the most general state transformer is given by the Kraus matrices [31,32]. The state update rule after measurements will not be discussed in this paper.

III. GENERALIZED CANONICAL COORDINATE SYMMETRIES WITH A UNIQUE INNER PRODUCT

In this section, we demand certain symmetries on \mathfrak{P} and derive an inner product for quasiprobability distributions on \mathfrak{P} . The inner product will provide a generalization of the Born rule (up to a constant which is determined in the subsequent section). The rule gives the operational meaning to states in terms of probabilities of measurement outcomes for given system preparations. The symmetries restrict the Born rule to a natural mathematical generalization of the classical and quantum cases.

We shall demand certain elementary symmetries to narrow both the Born rule and the time evolution. For simplicity, we focus on the case of a two-dimensional phase space with the coordinates (q, p) . At a minimum, valid states need to be normalized functions in state space. Convex combinations of valid states are also valid states. As part of specifying a given theory the state space will in general have further restrictions.

We demand that there exists a coordinate system of generalized position and momentum, (q, p) , that satisfies the following ‘‘canonical’’ coordinate symmetries.

Postulate 1 (Canonical coordinate symmetries). There exists such a coordinate system (q, p) where the physical laws manifested by equations of motion and measurement are invariant under the following coordinate transformations:

(1) Translation: $(q, p, t) \mapsto (q + a, p + b, t + c)$, for any $a, b, c \in \mathbb{R}$. We represent its action on functions via $(\hat{T}_{a,b,c} f)(q, p, t) = f(q + a, p + b, t + c)$.

(2) Switch: $(q, p, t) \mapsto (Cp, q/C, -t)$, where C is an arbitrary constant with units $[C] = [q/p]$.

(3) Time reversal: $(q, p, t) \mapsto (q, -p, -t)$ [equivalent to $(q, p, t) \mapsto (-q, p, -t)$ by switch].

Invariance under spacetime translations and boosts are one of the basic symmetries of nature, and time reversal symmetry,

while approximately correct in low-energy physics, is a useful computational tool [33]. The switch symmetry is not usually explicitly presented, though it exists in classical mechanics (via canonical transformations [1,34]) and quantum mechanics (via corresponding unitary and antiunitary transformations [32,33]). It can be viewed as placing position and momentum on an equal footing. Together, the symmetries physically define a canonical coordinate system. These symmetries generate a group that includes other symmetries. For example, we will utilize parity symmetry implied by time reversal and switch symmetry in Appendix C.

In a coordinate system obeying these canonical symmetries, we can get a unique inner product by introducing an additional natural restriction, namely, the inner product between two states goes to zero with increasing separation between the configurations they describe.

Postulate 2 (Local inner product). The inner product is local which means for two arbitrary quasiprobability distributions f_1 and f_2 ,

$$\lim_{a \rightarrow \infty} \langle f_1, \hat{T}_{a,0,0} f_2 \rangle = 0. \quad (22)$$

These postulates identify the inner product.

Theorem 1 (Generalized inner product). Consider an inner product of two arbitrary quasiprobability distributions f_1, f_2 of an elementary two-dimensional system. If the inner product is local and the generalized coordinates obey the canonical symmetries, then

$$\langle f_1, f_2 \rangle \propto \int f_1(q, p) f_2(q, p) dq dp. \quad (23)$$

Proof. An inner product is a bilinear symmetric function of two states. For phase space distributions f_1 and f_2 , a general form of a bilinear function is

$$\int M(q, p, \Delta q, \Delta p) f_1(q, p) f_2(q + \Delta q, p + \Delta p) \times dq dp d\Delta q d\Delta p, \quad (24)$$

where M is an arbitrary function. The symmetric condition on the inner product $\langle f_1, f_2 \rangle = \langle f_2, f_1 \rangle$ further requires

$$M(q, p, \Delta q, \Delta p) = M(q, p, -\Delta q, -\Delta p) \quad (25)$$

for arbitrary $a, b, c, d \in \mathbb{R}$.

Translation symmetry requires $\langle f_1(q, p), f_2(q, p) \rangle = \langle f_1(q + a, p + b), f_2(q + a, p + b) \rangle$ such that

$$\begin{aligned} & \int M(q, p, \Delta q, \Delta p) f_1(q, p) f_2(q + \Delta q, p + \Delta p) d\Omega \\ &= \int M(q, p, \Delta q, \Delta p) f_1(q + a, p + b) \\ & \quad \times f_2(q + a + \Delta q, p + b + \Delta p) d\Omega, \end{aligned} \quad (26)$$

where $d\Omega = dq dp d\Delta q d\Delta p$. Equation (26) holds for arbitrary f_1, f_2 , so

$$M(q, p, \Delta q, \Delta p) = M(q - a, p - b, \Delta q, \Delta p) \quad (27)$$

for all $a, b \in \mathbb{R}$. Therefore, M depends on only the relative distance $\Delta q, \Delta p$,

$$M(q, p, \Delta q, \Delta p) = M(\Delta q, \Delta p). \quad (28)$$

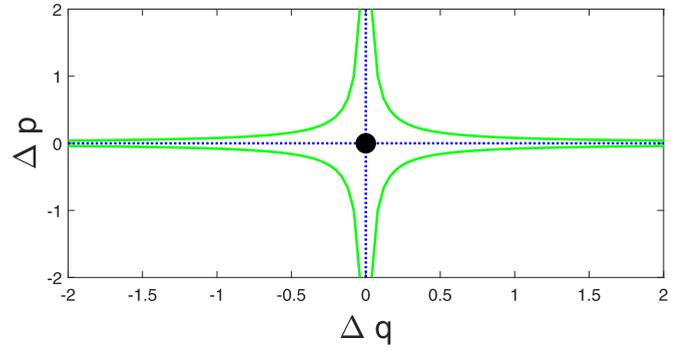


FIG. 4. Contour plot of $M(\Delta q, \Delta p)$, a function that appears in the initially general form of the inner product as in Eq. (24). Both q and p are plotted in units of $\hbar^{1/2}$, but they can have different units as long as $[qp] = [\hbar]$. The points in a pair of hyperbolas ($|\Delta q \Delta p| = c, c > 0$) share the same value of M , and so does the $\Delta q, \Delta p$ axis (except the origin). The origin is an isolated point in the contour plot because it represents the inner product contributed by the point itself, unaffected by any symmetry operations.

Similarly, switch symmetry with dimensional constant C requires $\langle f_1(q, p), f_2(q, p) \rangle = \langle f_1(p/C, Cq), f_2(p/C, Cq) \rangle$, which leads to

$$M(\Delta q, \Delta p) = M(\Delta p/C, C\Delta q). \quad (29)$$

Time reversal symmetry requires $\langle f_1(q, p), f_2(q, p) \rangle = \langle f_1(q, -p), f_2(q, -p) \rangle$, which leads to

$$M(\Delta q, \Delta p) = M(\Delta q, -\Delta p). \quad (30)$$

Equations (25), (28), (29), and (30) imply that $M(q, p, \Delta q, \Delta p)$ is constant when $|\Delta p \Delta q| = c$ for arbitrary $c \geq 0$, except at the origin ($\Delta q = \Delta p = 0$). All these contour lines extend to infinity, as illustrated in Fig. 4. Nevertheless, under Postulate 2, $\langle f_1, f_2 \rangle = 0$ for infinitely separated states, $M(\Delta q, \Delta p)$ must go to 0 when $\Delta q \rightarrow \infty$. This implies that $M(\Delta q, \Delta p) = 0$ except for at the origin ($\Delta q = \Delta p = 0$). Thus, $M(\Delta q, \Delta p) \propto \delta(\Delta q)\delta(\Delta p)$, and the inner product must have the form

$$\langle f_1, f_2 \rangle \propto \int f_1(q, p) f_2(q, p) dq dp. \quad (31)$$

The inner product of two density matrices or Wigner functions gives the measurement probability under the standard Born rule. We present a similar measurement formula in the generalized framework in Sec. IV, which determines the constant factor in the generalized Born rule and its relation to the Planck constant h .

IV. GENERALIZED PLANCK CONSTANT OF UNCERTAINTY: EFFECT AND STATE VOLUME

In this section, we introduce a property of effects that can be interpreted as the phase space volume they effectively occupy, which is called “effect volume.” Then, we introduce the generalized Born rule for a type of state-associated measurement, which is fundamental in quantum and classical theories. In this case, the effect volume can also be called state volume. The volume equals the Planck constant h for quantum pure

states and is zero (in a suitable limiting sense) for classical pure states.

A. Effects and their phase space volume

Recall that the completeness of a measurement inside a region \mathcal{D} gives

$$\sum_i h_i = 1_{\mathcal{D}}. \quad (32)$$

Integrating both sides results in

$$\int \sum_i h_i dq dp = \int 1_{\mathcal{D}} dq dp = V_{\mathcal{D}}, \quad (33)$$

where $V_{\mathcal{D}}$ is the volume of the phase space region \mathcal{D} .

We can factorize the h_i as

$$h_i = c_i g_i, \quad (34)$$

where g_i is a normalized quasiprobability distribution, $\int g_i dq dp = 1$, and c_i is some constant weight. Then interchanging summation and integration in Eq. (33) results in

$$\sum_i c_i = V_{\mathcal{D}}. \quad (35)$$

Equation (35) identifies the sum of the weights c_i with the phase space volume, inspiring the following.

Definition 1 (Effect volume). For the phase space representation of a discrete effect $e_i = (g_i, c_i)$ the *effect volume* is defined as its weight,

$$V_i := c_i. \quad (36)$$

Hence the probability of the outcome i can be written as

$$e_i(f) = V_i \int f g_i dq dp. \quad (37)$$

Note that, despite its normalization, g_i does not necessarily represent a valid state in a GPT.

Aggregating different outcomes and thus combining different effects we count the total probability of the aggregate, arriving at a *coarse-grained measurement*. We can check that the effect volume is additive like ‘‘volume’’ in the sense that the coarse-grained effect volumes are the sum of aggregated effects’ volumes.

Theorem 2 (Coarse-grained measurement). Consider a coarse-grained measurement effect, whose outcome probability is given by

$$e_{C.G.}(f) = \sum_{i \in \mathbb{K}} e_i(f), \quad (38)$$

where \mathbb{K} is a set of undistinguished results. Then its effect volume $V_{C.G.}$ is given by $V_{C.G.} = \sum_{i \in \mathbb{K}} V_i$.

Proof.

$$e_{C.G.}(f) = \int h_{C.G.} f dq dp = \sum_{i \in \mathbb{K}} V_i \int g_i f dq dp, \quad (39)$$

where we defined $h_{C.G.} = \sum_{i \in \mathbb{K}} V_i g_i$. We factorize $h_{C.G.}$:

$$h_{C.G.} = \sum_{j \in \mathbb{K}} V_j \sum_{i \in \mathbb{K}} \frac{V_i g_i}{\sum_{j \in \mathbb{K}} V_j}. \quad (40)$$

By Eq. (34), $h_{C.G.} = V_{C.G.} g_{C.G.}$ where $\int g_{C.G.} dq dp = 1$. Notice

$$\int \sum_{i \in \mathbb{K}} \frac{V_i g_i}{\sum_{j \in \mathbb{K}} V_j} dq dp = 1, \quad (41)$$

so the integrated function is $g_{C.G.}$. Consequently, $V_{C.G.} = \sum_{i \in \mathbb{K}} V_i$. ■

B. Effect volume in state-dual measurements: State volume

We now consider an important special case of effect volume, concerning effects that are associated with valid states. In this case, the effect volume can also be termed a state volume. We show that such a state volume is given by the inverse 2-norm of the distribution. We discuss the implications, including how the minimal state volume can be viewed as a generalization of the Planck constant associated with uncertainty.

Definition 2 (State-dual measurements). If all the effects $\{e_i\}$ of a measurement satisfy

- (1) $e_i(f) \propto \langle f, g_i \rangle$, where g_i is a valid state
- (2) $e_i(g_i) = 1$

we call it a *state-dual measurement*.

Definition 2 leaves the measurement update rule of self-dual measurements general, e.g., it is not required that the measurement is repeatable (such that an iterated application always leads to the same result [19,31]).

State-dual measurements are particularly fundamental in *self-dual* theories, which roughly means a one-to-one correspondence between states and effects [6,7] up to normalization. Both quantum and classical theories are self-dual. Another motivation for state-dual measurement comes from the dual role of Hamiltonian. While the Hamiltonian describes the evolution of states (property of states), it is also an observable (with effects), so it must bridge the states and effects. and we shall later use the generalized energy eigenstates to form a state-dual energy measurement just like the quantum case.

Physically, state-dual measurement effects can be understood as determining ‘‘Is the system in a state g_i ?’’ More specifically, the definite outcome $e_i(g_i) = 1$ implies that other outcomes must have zero probability: $e_i(g_j) = 0$ when $i \neq j$. Consequently, states g_i associated with a state-dual measurement must be orthogonal to each other:

$$\langle g_i, g_j \rangle = 0, \text{ when } i \neq j. \quad (42)$$

Definition 2 restricts the effect volume of state-dual effects. First, note that, since *Theorem 1* has derived a unique inner product, condition (1) in *Definition 2* means $e_i(f) \propto \int f g_i dq dp$. The probability of an outcome i is given by Eq. (37) but its volume V_i is still undefined. Condition (2) identifies V_i .

Theorem 3 (Volume of state-dual effects). The volume of a state-dual measurement effect associated with state g_i is given by $V_i = 1 / \int g_i^2 dq dp$.

Proof. By condition (2) of *Definition 2*,

$$e_i(g_i) = V_i \int g_i^2 dq dp = 1. \quad (43)$$

Therefore,

$$V_i = 1 / \int g_i^2 dq dp. \quad (44)$$

The V_i of Eq. (44) depends on only the state g_i , therefore we similarly term V_i the *state volume* of g_i . For normalized distributions, the larger the 2-norm, the more peaked the function is. Thus the volume of a state, the inverse of the 2-norm, intuitively reflects the uncertainty of outcomes given that state. We shall later show how the state volume, minimized over states, becomes the Planck constant in the case of quantum theory.

The state volume is closely connected to the number of distinguishable states. Consider a complete state-dual measurement (or complete inside region $\mathcal{D} \subset \mathfrak{P}$) with the associated set of states $\{g_i\}$. This measurement can distinguish these $\{g_i\}$ in a single shot. States and effects are symmetric in state-dual measurements, so we can apply the completeness condition of effects to states,

$$\sum_i V_i g_i = 1 \quad (\text{or } 1_{\mathcal{D}}), \quad (45)$$

when g_i are associated with a complete set of state-dual measurements (or complete inside region \mathcal{D}). [The functions 1 or $1_{\mathcal{D}}$ on the right-hand side represent the (unnormalized) *maximally mixed state*. When we know nothing about a system, it is in the maximally mixed state, which is the probabilistic mixture of all possible results by Eq. (45)].

Repeating the analysis of Sec. IV, we know $\sum_i V_{g_i} = V_{\mathcal{D}}$ when completeness is inside region \mathcal{D} . If V_{g_i} equals a constant V_g for all i , then we can count the total number N_g of states inside $\{g_i\}$ by

$$N_g = \frac{V_{\mathcal{D}}}{V_g}. \quad (46)$$

Thus, the state volume determines how many orthogonal states can be stored in a finite phase space region.

These arguments can be extended to the case where the completeness condition is approximately defined ($\sum_i V_i g_i \simeq 1_{\mathcal{D}}$).

C. Examples

In quantum mechanics, projective measurements are state-dual measurements. Their probabilities are given by

$$P(\phi|\psi) = e_{\phi}(\psi) = |\langle \phi | \psi \rangle|^2 = h \int W_{\phi} W_{\psi} \quad (47)$$

or the coarse graining of the above outcomes.

The state or effect volume is associated with the purity of quantum states by

$$\text{tr}(\hat{\rho}^2) = h \int W^2 dq dp = h/V. \quad (48)$$

Therefore, pure states have a minimal volume h . All the mixed states have a larger state volume. This also provides a reflection of the uncertainty principle in the Wigner function formalism.

The Eq. (47) implements the so-called *reciprocity law* of quantum mechanics [28], which states that for two pure states

ϕ and ψ the probability of observing outcome ϕ in a maximal test following preparation of state ψ equals to the probability of observing outcome ψ in a maximal test following a preparation of state ϕ . If a GPT has two or more sets of state-dual effects, $e = (g_i, V_i)$, $e' = (g'_i, V'_i)$, ... that satisfy the reciprocity relation in the form

$$P(i|j') = P(j'|i), \quad (49)$$

then all effects of all these measurements have the same volume. This follows from the application of Eq. (37) twice, reversing the roles of the state and the effect for the effects of two measurements.

Position and momentum eigenstates have state volumes different from “regular” pure quantum states consistent with the fact that they are actually outside the Hilbert space. They are unnormalizable delta functions. One may approximate them by Gaussians with finite width, in which case the state volume is still h .

The classical sharp phase space localization has $\mu = (q_0, p_0) \in \mathfrak{P}$. As we have seen in Sec. IV the effect (q_0, p_0) (with the “uncertainty” $dq_0 dp_0$) is represented by a normalized distribution $g_{(q_0, p_0)}(q, p) = \delta(q - q_0)\delta(p - p_0)$ on \mathfrak{P} with $c_{(q_0, p_0)} = 1$. As a result,

$$dV_{(q_0, p_0)} = dq_0 dp_0. \quad (50)$$

The coarse-grained version of such sharp effects provides another example. As a simple example consider the box mixed state:

$$g_{\square}(q, p) = \begin{cases} (\epsilon\delta)^{-1}, & q \in (q_0, q_0 + \epsilon) \text{ and } p \in (p_0, p_0 + \delta) \\ 0, & \text{otherwise} \end{cases}. \quad (51)$$

The state volume that is occupied by each of these states is $V_g = \epsilon\delta \rightarrow 0$, and a domain \mathcal{D} contains

$$N_g = V_{\mathcal{D}}/(\epsilon\delta) \rightarrow \infty \quad (52)$$

orthogonal states.

V. GENERALIZED ENERGY EIGENSTATES: PURE STATIONARY STATES

The concept of eigenstate appeared with the foundations of quantum mechanics; it also appears in the quantum-like Koopman formalism of classical mechanics [12,28,35–37]. However, here we will define the energy eigenstates from a slightly different perspective, which also works for generalized theories. We define the generalized eigenstates via the dynamic features of states: the purest stationary states. The Noether theorem inspires this: the energy directly represents the time-invariant feature. These states are a cornerstone to put quantum and classical mechanics in the same framework.

A. Definition

Definition 3 (Stationary states). Given a time evolution rule, *stationary states* are states represented by time-independent phase space functions.

Probabilistic mixtures of stationary states are, by inspection, also stationary, so there is a convex set of stationary states.

Definition 4 (Pure stationary states). Pure stationary states

are states in the set of stationary states that cannot be represented as nontrivial probabilistic mixtures of other stationary states.

Note that pure *stationary states* are not necessarily pure states of the convex set of *all* allowed states of a GPT.

We will identify pure stationary states with the generalized energy eigenstates. We show in subsequent sections that they satisfy three natural desiderata:

- (1) Pure stationary states can be assigned sharp energy values, always giving the same value in energy measurement.
- (2) They describe the time evolution of the system.
- (3) They coincide with the standard quantum energy states in the case of quantum mechanics.

B. Examples

In the classical case orbits in phase space describe the time evolution, and uniform distributions over orbits are the only stationary states. More specifically, stationary states by definition obey $\frac{\partial f}{\partial t} = 0$. In the action-angle coordinates described in Sec. II B,

$$\frac{\partial f}{\partial t} = \frac{\partial f}{\partial \theta} \frac{\partial H}{\partial I}. \quad (53)$$

Consider the case of $\frac{\partial H}{\partial I} \neq 0$ first. Then a stationary state requires $\frac{\partial f}{\partial \theta} = 0$ for all θ , so

$$f(I, \theta) = f(I) = \int \rho(I_i) f_i dI_i, \quad (54)$$

where

$$f_i(I) = \frac{1}{2\pi} \delta(I - I_i) \quad (55)$$

are thus (normalized) pure states of the set of stationary states. We see that these states are not pure in the set of all states. As an illustrative example, the delta functions in Fig. 3 are examples of pure stationary states in harmonic oscillators. The state-dual measurement of I is given by

$$P(I_0, dI_0|f) = p(I_0|f) dI_0 = 2\pi dI_0 \int f \frac{1}{2\pi} \delta(I - I_0) dq dp. \quad (56)$$

Therefore, pure stationary states have infinitesimal volume,

$$dV_{I_0} = 2\pi dI_0, \quad (57)$$

in agreement with the usual geometric interpretation of the action variables [1,2].

Finally, consider the case when $\frac{\partial H}{\partial I} = 0$, for some values of I , such that any distribution over θ for those values of I is stationary. Hence, the associated classical pure stationary states become pure states $\delta(I - I_0)\delta(\theta - \theta_0)$, where $I_0 \in \{I' | \frac{\partial H}{\partial I'} = 0\}$, $\theta_0 \in [0, 2\pi)$. (The state volume is $V = dq_0 dp_0 = dI_0 d\theta_0$ by the same reasoning as above.)

Classical pure stationary states correspond to “eigen-wave functions” in the Koopman–von Neumann formalism [35,38],

where the classical states are described by wave functions like quantum mechanics. Koopman mechanics describes classical states by probability amplitudes $\phi(q, p, t)$, and the probability density $\rho(q, p) = |\phi(q, p)|^2$. The evolution of $\phi(q, p, t)$ is given by $i \frac{\partial \phi}{\partial t} = \hat{L} \phi$, where $\hat{L} = -i \frac{\partial H}{\partial p} \frac{\partial}{\partial x} + i \frac{\partial H}{\partial x} \frac{\partial}{\partial p}$, named the Liouvillian operator, is a generator of time translations, analogous to the quantum Hamiltonian operator. The classical Hamiltonian H that enters the expression for \hat{L} is the energy observable. For example, (unnormalized) eigenfunctions of a free particle Liouvillian are $\phi_{p_0, \lambda} = e^{i\lambda q} \delta(p - p_0)$, where λ_q and p_0 can be arbitrary real numbers. However, the corresponding probability distributions containing $\delta^2(p - p_0)$ are ill-defined. Instead, pure stationary states are effectively $|\phi|$ up to normalization in this case. Moreover, a conceptual separation between the energy observable and the time translations generator puts the Koopmanian formalism outside the GPTs we consider.

For quantum mechanics, the time evolution can be described by the commutator of the density matrix and the Hamiltonian. Therefore, stationary states must have diagonal density matrices in the energy basis. Pure states among stationary states are just energy eigenstates, including in cases of degeneracy. Hence, the pure stationary state is just another name for the energy eigenstate in the quantum case.

By inspection, both quantum and classical pure stationary states have sharp energy values. This provides a vehicle to define energy measurement, which associates measurements with states. Whether the pure stationary states can provide an energy measurement in generalized theories will be established in Sec. VII.

The next section shows how pure stationary states relate to time evolution in generalized theories. The dynamical symmetry represented by pure stationary states is connected to energy as a conserved quantity. This is consistent with Noether’s theorem, as will be explained later.

VI. GENERALIZED EQUATION OF MOTION FROM POSTULATES

We will now derive the equation of motion in terms of generalized eigenstates.

A. Evolution based on pure stationary states

As an analogy of quantum energy eigenstates, we introduce the following nontrivial postulates.

Postulate 3 (Evolution dependence). The time evolution of a state only linearly depends on the *pure stationary states*, up to some dimensional factors \mathcal{E}_i to keep the dimensions identical,

$$\frac{\partial f}{\partial t} = G \left(f, \sum_i \mathcal{E}_i g_i \right) = \sum_i \mathcal{E}_i G(f, g_i), \quad (58)$$

where g_i is a set of pure stationary states and \mathcal{E}_i are corresponding parameters and G is some bilinear functional.

There are several reasons for Postulate 3. First, stationary states are the only choice of states, under time translation symmetry, to depict the time evolution without external factors because they are the only states independent of time. Second, using *pure* stationary states gives the evolution as

much freedom as possible. If we used some mixtures of stationary states instead of pure stationary states, the evolution is equivalent to a specific linear combination of pure stationary states, which is just a special case of this postulate. Finally, the pure stationary states contain only time-independent information, while $\frac{\partial f}{\partial t}$ contains the dimension of time. Hence, we must have some parameters \mathcal{E}_i containing the dimension of the inverse time (which will later turn out to be proportional to the energy value).

Postulate 4 (Independence of stationary states). The pure stationary states are independent in the sense that $G(g_i, g_j) = 0$ holds for arbitrary i, j .

When you change the value of \mathcal{E}_i , in principle, the original pure stationary states may not be stationary anymore. Postulate 4 aims to avoid such a complex situation. All the pure stationary states are stationary under other pure stationary states' impact so that their stationarity is independent of \mathcal{E}_i (energy values). Both quantum and classical mechanics satisfy this postulate by inspection.

B. Localized and nonlocalized evolution expressions

Before deriving the equation of motion, we will introduce a different perspective to describe the time evolution, which will benefit later work. We will discuss the idea of the evolution of the phase space state distribution being localized in phase space.

Often time evolution is written as an *apparently localized expression*. For example, in the continuity equation $\partial_t f + \vec{\nabla} \cdot (\vec{v}f) = 0$ for conserved distributions evolution of f at a point being determined locally depends on the velocity field v at this point. In general, an apparently localized expression (of evolution) means the state update at one point depends only on quantities (derivatives) at this point. The Liouville equation (2) is an apparently localized expression in classical mechanics. On the other side, Eq. (10) is an apparently local expression to describe quantum evolution.

In contrast, the Wigner transport equation (Eq. (11)) associates the evolution of one phase-space point evolution with nonadjacent Wigner function points. When the evolution equation at one point contains terms at other points, we call the equations *apparently nonlocalized expressions* (of evolution). An important example is Eq. (11). The original Eq. (11) works for the special case $H = p^2/2m + V(q)$. We may generalize it to arbitrary Hamiltonians:

$$\begin{aligned} \frac{\partial W}{\partial t}(q, p) &= \int W(q+l, p+j)J(q, p, l, j) dj dl, \\ J(q, p, l, j) &= \frac{i}{\pi^2 \hbar^3} \int [H(q-y, p+z) - H(q+y, p-z)] \\ &\quad \times e^{-\frac{2i}{\hbar}(jy+lz)} dy dz. \end{aligned} \quad (59)$$

Equation (59) is equivalent to the quantum Liouville equation (10) (as shown in Appendix B). From the Eq. (59) representation of time evolution, we find the evolution at (q, p) to depend on the distribution at $(q+l, p+j)$ (and thus everywhere). The probability conservation is guaranteed by $J(q, p, l, j) = -J(q+2l, p+2j, -l, -j)$, so the term $W(q+l, p+j)J(q, p, l, j)$ actually contributes to $\frac{\partial W(q,p)}{\partial t}$ and $-\frac{\partial W(q+2l, p+2j)}{\partial t}$. The distribution $W(q+l, p+j)$ in that

sense plays the role of a porter, transferring other distributions from $(q+2l, p+2j)$ to (q, p) . This differs from the standard classical stochastic evolution wherein the probability current from A to B is always proportional to $P(A)$.

A general apparently nonlocalized expression (only assuming $\frac{\partial f}{\partial t}$ is linear in f) is as follows:

$$\frac{\partial f}{\partial t}(q, p) = \int f(q+l, p+j)J(q, p, l, j) dl dj, \quad (60)$$

where the state update is described by a nonlocalized generator J . In quantum and classical theories, J depends on the Hamiltonian. We will apply the apparently nonlocalized expression in the following derivations. By inspection, Eq. (60) assumes only that the evolution is linear in f , respecting the property of (quasi)probability distributions.

Apparently localized and nonlocalized expressions themselves do not imply any physical difference, they are just two different languages to express the evolution rule. The apparently nonlocalized Eq. (60) can transform to an apparently localized expression like Eq. (10) and vice versa (for well-behaved functions). Consider classical free particles as an example, $H = \frac{p^2}{2m}$, $\frac{\partial f}{\partial t} = -\frac{\partial f}{\partial q} \frac{p}{m}$. We can express the same evolution by $J(q, p, l, j) = \delta'(l)\delta(j)\frac{p}{m}$ in the nonlocalized expression [$\delta'(\cdot)$ is the derivative of the delta function; Appendix B shows the derivation].

We introduce the apparently nonlocalized expression because it is more convenient for describing *physically nonlocalized evolution*. The physically nonlocalized evolution here denotes the appearance of infinite-order derivative [39,40] in the apparently localized expression. Quantum mechanics has a physically nonlocalized evolution, which can be seen from the $\sin(\Lambda)$ term of Eq. (10).

On the other hand, if the evolution can be described by finite-order derivatives in the apparently localized expression, then we call it *physically localized evolution*. Classical mechanics has physically localized evolution since the Liouville equation (2) is an apparently localized expression that only contains first-order derivatives.

When describing physically nonlocalized evolution by apparently localized expressions, we have to deal with infinite-order derivatives and their physical meaning is abstract. However, beginning with an apparently nonlocalized expression like Eq. (60) can explicitly demonstrate the ‘‘jumping’’ in phase space, which has a clear physical picture. Since our generalized framework may contain physically nonlocalized evolution, we choose apparently nonlocalized expressions to derive the generalized equation of motion. We also find it eases the derivation process.

C. Conditions to derive the equation of motion

Before deriving the equation of motion, we will present several physical requirements and see how these restrict $J(q, p, l, j)$ in the equation of motion.

Postulate 3 says the evolution linearly depends on pure stationary states, which is the core of our framework. We also have the canonical coordinate symmetries introduced in Postulate 1. We further require the following postulate.

Postulate 5 (Inner product invariance). The time derivative of inner products $\frac{\partial}{\partial t} \int f_1(t)f_2(t) dq dp = 0$ for arbitrary states f_1, f_2 and time point t .

The postulate has several consequences. First, combined with Definition 2, it implies that applying a state-dual measurement on another state has a time-independent probability when both states evolve under a given evolution. For example, if initially you know a system has a 50% probability of being state $f(t=0)$, then during the evolution, we expect the system to still have a 50% probability of being $f(t)$. This is a physical motivation of Postulate 5. Second, $\frac{\partial}{\partial t} \int f^2 dq dp = 0$ for any f , which means the state volume does not change. Third, the maximally mixed state is the state that maximizes the state volume, so it has to be stationary. Thus it is a convex combination of pure stationary states and accordingly the pure stationary states satisfy the completeness condition. Finally, any state f 's inner product with the maximally mixed state is invariant under time evolution, $\frac{\partial}{\partial t} \int f dq dp = 0$, which means the total probability is conserved. The inner product invariance is the only use of the inner product in the derivation of the equation of motion.

D. Restrictions to the nonlocalized generator J

Next, we will restrict the nonlocalized generator $J(q, p, l, j)$ by the above requirements.

Postulate 3 says the evolution linearly depends on pure stationary states, therefore J linearly depends on pure stationary states g . It gives us

$$J_g(q, p, l, j) = \int g(q+x, p+y)A(q, p, x, y, l, j) dx dy, \quad (61)$$

where A is some unsettled function. What appears in the final equation of motion is $J_{\sum_i \mathcal{E}_i g_i}$, but here we focus only on the functional J . This is similar to Eq. (60) for the same reason: this is the most general form of linear dependence. However, we step further and argue the following.

Lemma 1 (J 's dependence on g).

$$J_g(q, p, l, j) = \int g(q+x, p+y)A(x, y, l, j) dx dy. \quad (62)$$

A is some unsettled function and cannot depend on q, p and state g .

The reason that A cannot depend on q, p is the following: The coordinates themselves are physically meaningless; all the dependence on (q, p) means dependence on some state. We have already assumed that the dynamics are independent of time (Postulate 1), and the only states independent of time are pure stationary states or linear combinations of them. If A depends on pure stationary states, it breaks the linearity relation between J and g . Therefore, there is no space for A to depend on q, p .⁵ Equation (62) gives a general form of J under

⁵Whenever we utilize symmetries to restrict some function, we always implicitly assume the function depends only on the elements that we are aware of in this issue. Otherwise, arbitrary symmetry can be achieved by introducing a new degree of freedom, just like gauge fields.

Postulate 3. An equivalent statement is that if we translate all the pure stationary states, $g'(q, p) = g(q + \Delta q, p + \Delta p)$ without changing anything else, then $J_{g'}(q, p, l, j) = J_g(q + \Delta q, p + \Delta p, l, j)$.

Then, to keep the inner product invariance of Postulate 5,

$$\begin{aligned} 0 &= \int \frac{\partial f_1}{\partial t} f_2 + \frac{\partial f_2}{\partial t} f_1 dq dp \\ &= \int f_1(q+l, p+j)J(q, p, l, j)f_2(q, p) \\ &\quad + f_2(q+l, p+j)J(q, p, l, j) \\ &\quad \times f_1(q, p) dq dp dj dl \\ &= \int f_1(q+l, p+j)J(q, p, l, j)f_2(q, p) \\ &\quad + f_2(q-l, p-j)J(q, p, -l, -j) \\ &\quad \times f_1(q, p) dq dp dj dl \\ &= \int f_1(q+l, p+j)J(q, p, l, j)f_2(q, p) \\ &\quad + f_2(q, p)J(q+l, p+j, -l, -j) \\ &\quad \times f_1(q+l, p+j) dq dp dj dl. \end{aligned} \quad (63)$$

[When the dependence on g is not emphasized, we write $J(q, p, l, j)$ instead of $J_g(q, p, l, j)$]. We require the following.

Lemma 2 (Requirement from inner product invariance).

$$J(q, p, l, j) = -J(q+l, p+j, -j, -l). \quad (64)$$

This is consistent with the known fact that the generator of an orthogonal group, i.e., the group keeping the dot product invariant, is antisymmetric.

Subsequently, we utilize the canonical coordinate symmetries introduced in Postulate 1. Due to translation symmetry, we can focus, without loss of generality, on the evolution of one point ($q=0, p=0$) defining a functional:

$$J_g^0(l, j) = J_g(q=0, p=0, l, j). \quad (65)$$

For the equation of motion

$$\frac{\partial f}{\partial t}(0, 0) = \int J_g^0(l, j)f(l, j) dl dj, \quad (66)$$

we apply a transformation to it,⁶

$$\frac{\partial f'}{\partial t}(0, 0) = \int J_{g'}^0(l, j)f'(l, j) dl dj. \quad (67)$$

A symmetry means that the functional is invariant under a transformation: $J^{0'} = J^0$. The canonical symmetries will restrict J in such a way. The concrete derivation is in Appendix C. In the end, we find the canonical symmetries (Postulate 1) require the following.

⁶We can write $(q, p, r) \mapsto (q', p', r')$, or equivalently $f(q, p, r), g(q, p, r) \mapsto f'(q, p, r), g'(q, p, r)$. We choose the latter one here.

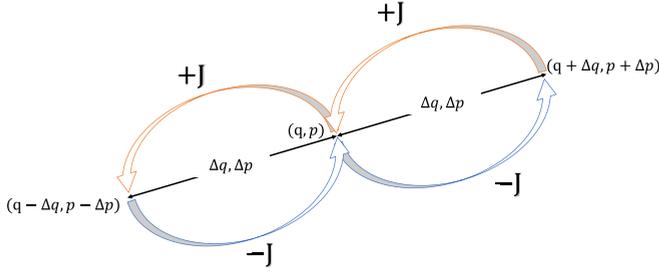


FIG. 5. The symmetry restrictions that Lemmas 2 and 3 impose on the nonlocalized generator J . There are three phase space points: (q, p) , $(q + \Delta q, p + \Delta p)$, $(q - \Delta q, p - \Delta p)$. The restriction is that for any $q, p, \Delta q$ and Δp , $J(q, p, \Delta q, \Delta p) = -J(q, p, -\Delta q, -\Delta p) = -J(q + \Delta q, p + \Delta p, -\Delta q, -\Delta p)$.

Lemma 3 (Requirement from canonical coordinate symmetries).

$$J(q, p, l, j) = -J(q, p, -l, -j). \quad (68)$$

Lemmas 2 and 3 (the requirements of J from the inner product invariance and the canonical symmetries) are illustrated by Fig. 5.

E. Derivation of the equation of motion

We already have all the ingredients on the table to derive the functional $J_g(q, p, l, j)$. By the end of this section, we will find that the evolution is similar to the quantum evolution but with extra degrees of freedom.

Figure 5, which combines Lemmas 2 and 3, shows that

$$J(q, p, l, j) = J(q + l, p + j, l, j), \quad (69)$$

which means that J is periodic in q, p space. Therefore, J has only the components that satisfy $k_q l + k_p j = 2\pi n$ ($n \in \mathbb{Z}$) in the frequency domain (k_q, k_p are the angular frequency of q, p , respectively).

An observation will simplify the calculation: Eq. (62) is a convolution between g and A . The convolution theorem gives

$$J(q, p, l, j) = \text{Re} \int \tilde{g}(k_q, k_p) A''''(k_q, k_p, l, j) \times e^{i(k_q q + k_p p)} dk_q dk_p, \quad (70)$$

where \tilde{g} is the Fourier transform of g . (All the A with primes are unsettled functions, which help to absorb unimportant parameters.)

The periodic property leads to a $\sum_n \delta(k_q l + k_p j - 2\pi n)$ term in the frequency domain:

$$\begin{aligned} J(q, p, l, j) &= \text{Re} \int \tilde{g}(q_k, p_k) A''''(k_q, k_p, l, j) \sum_n \delta(k_q l + k_p j - 2\pi n) e^{i(k_q q + k_p p)} dk_q dk_p \\ &= \text{Re} \int g(q + y, p + z) \sum_n A''\left(\frac{2\pi n - k_p j}{l}, k_p, l, j\right) e^{i\left(\frac{2\pi n - k_p j}{l} y + k_p z\right)} dk_p dy dz \\ &= \text{Re} \int g(q + y, p + z) \sum_n A'(n, k', l, j) e^{i\frac{2\pi n y}{l}} e^{-ik'(jy - lz)} dk' dy dz, \end{aligned} \quad (71)$$

where we relabeled $\frac{k_p}{l}$ by k' . The term $e^{i\frac{2\pi n y}{l}}$ is not well defined when $l = 0$, but this term will vanish later.

The antisymmetric condition Lemma 3 requires that J is an odd function of l, j , so

$$J_g = \text{Im} \int g_i(q + y, p + z) \sum_n A'(n, k', l, j) e^{i\frac{2\pi n y}{l}} e^{-ik'(jy - lz)} dk' dy dz, \quad (72)$$

and $A'(n, k', l, j) = A'^*(n, k', -l, -j)$.

One more requirement is that J_g must keep g itself stationary (Postulate 4), i.e., that

$$\begin{aligned} \int g(q + l, p + j) J_g(q, p, l, j) dl dj &= \text{Im} \int g(q + l, p + j) g(q + y, p + z) \sum_n A'(n, k', l, j) e^{i\frac{2\pi n y}{l}} e^{-ik'(jy - lz)} dk' dy dz dl dj \\ &= 0. \end{aligned} \quad (73)$$

Observe that the equation can be written in the matrix form

$$g_{lj} M^{ljyz} g_{yz} = 0, \quad (74)$$

where $M^{ljyz} = \text{Im} \int \sum_n A'(n, k', l, j) e^{i\frac{2\pi n y}{l}} e^{-ik'(jy - lz)} dk'$. This relation holds for arbitrary vector g , the matrix M turns every vector into an orthogonal vector. It means M must be a generator of the orthogonal group, which is antisymmetric, $M^{ljyz} = -M^{yzlj}$, swapping yz with lz changes its sign. Therefore, we require that n can only equal zero and

$A'(n, k', l, j) = A'(k')$. Now the form of J is

$$J_g(q, p, l, j) = \text{Im} \int g(q + y, p + z) A'(k') \times e^{-ik'(jy - lz)} dk' dy dz dl dj. \quad (75)$$

The equation is already very similar to the generalized Wigner equation (59) (taking the imaginary part of the whole equation is equivalent to taking the odd part of the integrand). To harmonize the notation with the standard formulation quantum equation, we relabel $2/k' = k$ and replace $A'(k')$ by $K(k)$ such that $K(k)dk = A'(k')dk'$.

Theorem 4 (Generalized equation of motion). Given that (i) the evolution depends linearly on the generalized energy eigenstates (Postulates 3 and 4), (ii) inner product invariance (Postulate 5), and (iii) the symmetries of phase space (Postulate 1), the equation of motion has the following form:

$$\begin{aligned} \frac{\partial f}{\partial t} &= \sum_i \mathcal{E}_i \text{Im} \left\{ \int f(q+l, p+j) g_i(q+y, p+z) \right. \\ &\quad \left. \times K(k) e^{-i\frac{2(jy-lz)}{k}} d\Omega \right\} \\ &= -i\frac{\pi^2}{2} \sum_i \mathcal{E}_i \int k^2 K(k) \text{Wigner}_k\{\hat{f}_k, (\hat{g}_i)_k\} dk, \quad (76) \end{aligned}$$

where $d\Omega = dk dy dz dl dj$, the \mathcal{E}_i are temporary parameters mentioned in Postulate 3, $K(k)$ is a theory-specific distribution, $\text{Wigner}_k\{\}$ represents the $\hbar = k$ Wigner transform, and $\hat{f}_k, (\hat{g}_i)_k$ are operators corresponding to f, g under the $\hbar = k$ Weyl transform (8) (their units are different from density matrices). (Appendix D derives the operator form of the equation of motion).

In the second line of Eq. (76), we use the Wigner-Weyl transform [Eqs. (7) and (8)] to highlight the similarity to the common quantum expression in terms of commutator $\frac{\partial \hat{\rho}}{\partial t} = \frac{i}{\hbar}[\hat{\rho}, H]$ for which there is no integral over k .

A key generalization relative to the quantum and classical cases is the extra factor $K(k)$. In the quantum case, $K(k) = \delta(k - \hbar)$, and in the classical case $K(k) \rightarrow \delta(k)$. A nontrivial $K(k)$ can thus be interpreted as a linear combination of commutation relations instead of a single one in quantum mechanics. In quantum mechanics, this commutator gives \hbar in quantum mechanics, so one may interpret this loosely as a linear combination of different \hbar s. To understand the qualitative meaning of k , notice that k appears only in $e^{-i\frac{2(jy-lz)}{k}}$ in Eq. (76), namely, k is proportional to the jumping distance (l, j) in phase space, so $K(k)$ represents the jumping ability during evolution; hence we call it the *nonlocalized dynamics kernel*.

F. Examples

Now we will discuss the dynamics corresponding to different nonlocalized kernels $K(k)$. For the case that $K(k)$ is a single delta function $K(k) = \delta(k - \kappa)$, the second line of Eq. (76) is just a commutator in phase space, and the first line is, up to the appearance of the Hamiltonian, the quantum expression of Eq. (59) with $\hbar = \kappa$:

$$\begin{aligned} \frac{\partial f}{\partial t} &= \int f(q+l, p+j) J(q, p, l, j) dj dl, \\ J(q, p, l, j) &= \int \sum_i \mathcal{E}_i g_i(q-y, p+z) e^{-2\frac{i}{\kappa}(jy+lz)} dy dz. \end{aligned} \quad (77)$$

(The Fourier transform of the odd part functions is equivalent to the imaginary part of the Fourier transform.) The Hamiltonian will be defined from \mathcal{E}_i, g_i later.

When $\kappa \rightarrow 0$, like in path integrals, the phase oscillates so fast that only the first-order derivatives of $f(q, p)$ and $g_i(q, p)$

contribute to the imaginary part of the integral. When $K(k) \rightarrow \delta(k)$ the equation returns to the classical case with physically localized evolution [Eq. (2)].

$$\frac{\partial f}{\partial t} \propto \left\{ \sum_i \mathcal{E}_i g_i, f \right\}. \quad (78)$$

However, even with the same equation of motion, quantum/classical mechanics are not the only possible theories because the state space could differ. In Sec. IX we will introduce information-restricted quantum mechanics as an example.

The structure of stationary states can determine the value of κ . The Weyl transform under κ can simultaneously diagonalize all the stationary states. However, κ could differ from the state volume of stationary states divided by 2π , like in information-restricted quantum mechanics. It can neither be understood as the “theoretically” (all the states without negative probability are allowed) minimal volume of stationary states.⁷ Take quantum harmonic oscillators as an example. The evolution of the Wigner function is completely classical and independent of \hbar . It means that the theoretically minimal stationary state can be arbitrarily small.

Another unexpected possibility is that $K(k)$ is not a delta function.⁸ A distribution of hybrid commutators describes the evolution. In such a case, the Wigner function’s associative Moyal product (star product) $A \star B = Ae^{i\frac{\hbar}{2}\Lambda} B$, is replaced by a nonassociative hybrid Moyal product $A \star_H B = \int AK(k)e^{i\frac{k}{2}\Lambda} B dk$ (see Appendix E for the proof). Consequently, when we do a Weyl transform to rewrite everything in the operator form, the operator product is nonassociative. Moreover, the dynamics do not allow for the decomposition of density matrices by wave functions. We can give a trivial example of where such dynamics work. In harmonic oscillators, the evolution for arbitrary $K(k)$ is completely classical; there can be an orthogonal and complete set of pure stationary states without negative probability.

It is interesting but less rigorous to apply Eq. (76) directly to the $I - \theta$ coordinates (which do not rigorously satisfy our canonical symmetries). This equation of motion implies that if one canonical coordinate θ is periodic, then jumping in the I direction has to be *discrete*. This strongly suggests that I itself is discrete, consistent with the result of the discrete action-angle Wigner function in [41] and the Sommerfeld quantization condition that $2\pi I = nh$ where n is an integer.

VII. GENERALIZED HAMILTONIAN

After setting down the equation of motion, we would like to reconstruct the Hamiltonian. We will then have Hamiltonian mechanics that can apply to quantum and classical theories

⁷Actually, $2\pi\kappa$ is indeed the theoretically minimal average volume of stationary states for finite-dimensional cases, but we cannot apply this statement to phase space.

⁸Reference [18] has proposed a similar but weaker (less restricted) equation independently, called the generalized Moyal bracket by the authors.

and, in principle, more. Besides the generator of time evolution, the Hamiltonian also represents an observable, which contains the *measurement effects* and the *energy value*. We discuss each concept in turn before defining the Hamiltonian.

A. Energy measurement effects

Energy eigenstates provide the energy measurement effects in quantum mechanics. Generalizing this idea, we wish to define the measurement effects by the generalized energy eigenstates. However, pure stationary states do not always, for any GPT, provide a set of effects for state-dual measurement. Section IX will give an example: Spekkens' toy model [42]. Therefore, we have to postulate the following.

Postulate 6 (Existence of energy measurement). There exists a state-dual measurement whose effects all correspond to pure stationary states.

The postulate guarantees that pure stationary states can construct a complete and orthogonal set and the corresponding state-dual measurement exists. The energy measurement constructed from pure stationary states provides all the information we can learn that is irrelevant to the time point and induces a conserved observable. The time translation symmetry, along with the inner product invariance, ensures that there always exists stationary states. Postulate 6 ensures that there is always a state-dual measurement corresponding to pure stationary states. When stationary states give the set of measurement effects and any state's inner product with these stationary states is invariant (Postulate 5), then this set

of effects always gives a conserved observable. Therefore, the energy measurement here agrees with the definition given by the Noether theorem: it is the conserved observable promised by the time translation symmetry.

B. Energy value

Next, we will determine the value of energy. Energy is usually associated with the timescale in the case of natural units ($\hbar = c = 1$). This is also true for the \mathcal{E}_i , which appears in Postulate 3 and the equation of motion (76) representing the speed of evolution. We, therefore, anticipate that the energy associated with stationary state g_i , which we shall call E_i , is proportional to \mathcal{E}_i .

The probabilities of energy eigenstates are always invariant by inner product invariance, so any time-independent function of \mathcal{E}_i is conserved. To further constrain the definition of the energy value E_i , we demand that it is an extensive quantity (it adds under an independent composition of systems). To construct an extensive quantity, consider a composite system with two independent subsystems (labeled by 1,2) without any interaction and correlation, which means

$$f_{12} = f_1 f_2, \quad g_{12ij} = g_{1i} g_{2j} \quad (79)$$

and

$$\frac{\partial f_{12}}{\partial t} = \frac{\partial f_1}{\partial t} f_2 + \frac{\partial f_2}{\partial t} f_1. \quad (80)$$

We expect the energy in the composite system's equation of motion to be the sum of the subsystems' energy. Substituting the equation of motion Eq. (76) into Eq. (80) gives

$$\frac{\partial f_{12}}{\partial t} = \sum_i \mathcal{E}_i \text{Im} \int f_2(q_2, p_2) f_1(q_1 + l_1, p_1 + j_1) g_{1i}(q_1 + y_1, p_1 + z_1) K_1(k_1) e^{i2(j_1 y_1 - l_1 z_1)/k} dy_1 dz_1 dl_1 dj_1 + 1 \Leftrightarrow 2, \quad (81)$$

where the $1 \Leftrightarrow 2$ term is the same as the first term on the r.h.s. except that indices 1 and 2 are interchanged. We next introduce the following identity:

$$f(q, p) = \int f(q + l, p + j) \delta(l) \delta(j) dl dj = 1/(\pi^2 \bar{k}^2) \int f(q + l, p + j) \sum_i V_i g_i K(k) e^{i2(jy - lz)/k} dl dj dy dz dk, \quad (82)$$

where $\bar{k}^2 = \int k^2 K(k) dk$, and we have used $\delta(x) = \frac{1}{2\pi} \int e^{ikx} dk$, $\sum_i V_i g_i = 1$. Replacing f_2 in Eq. (81) by Eq. (82) gives

$$\begin{aligned} \frac{\partial f_{12}}{\partial t} &= \sum_{ij} \frac{\mathcal{E}_{1i}}{\bar{k}^2 2\pi^2} V_{2j} \text{Im} \int f_2 f_1 g_{1i} g_{2j} K_1(k_1) K_2(k_2) e^{i2(j_1 y_1 - l_1 z_1)/k_1 + (j_2 y_2 - l_2 z_2)/k_2} (dy dz dl dj)_{1,2} + 1 \Leftrightarrow 2 \\ &= \sum_{ij} \left(\frac{\mathcal{E}_{1i}}{\bar{k}^2 2\pi^2} (V_{2j}) + \frac{\mathcal{E}_{2i}}{\bar{k}^2 2\pi^2} (V_{1j}) \right) \text{Im} \int f_2 f_1 g_{1i} g_{2j} K_1(k_1) K_2(k_2) e^{i2[(j_1 y_1 - l_1 z_1)/k_1 + (j_2 y_2 - l_2 z_2)/k_2]} (dy dz dl dj)_{1,2}, \end{aligned} \quad (83)$$

where all the f depends on $(q + l, p + j)$ and all the g depend on $(q + y, q + z)$ with the corresponding subscript. On the other hand, the equation of motion for the composite system should have a consistent form

$$\frac{\partial f_{12}}{\partial t} = \sum_{ij} \mathcal{E}_{ij}^{\text{composite}} \text{Im} \int f_1 f_2 g_{1i} g_{2j} K_1(k_1) K_2(k_2) e^{i2[(j_1 y_1 - l_1 z_1)/k_1 + (j_2 y_2 - l_2 z_2)/k_2]} (dy dz dl dj)_{1,2}. \quad (84)$$

Comparing Eqs. (83) and (84), we find that if $\mathcal{E}_i \propto \frac{E_i V_i}{\bar{k}^2}$ (we omit the subscript for subsystems), then $\mathcal{E}_{ij}^{\text{composite}} \propto (E_{1i} + E_{2j}) \frac{V_{1i} V_{2j}}{\bar{k}^2 2\pi^2}$, i.e., E is an additive quantity.

However, $\frac{\mathcal{E}_i \bar{k}^2}{V_i}$ has the unit (not natural units) $[\frac{1}{t}]$ ($[\mathcal{E}_i] = [\frac{1}{qpt}]$, $[\bar{k}^2] = [q^2 p^2]$, $[V] = [qp]$) instead of the conventional quantum and classical unit for energy of $[\frac{qp}{t}]$. Of course, we

may multiply an arbitrary constant with the unit $[qp]$ without breaking conservation. However, there may not exist a simple choice of such parameters in generalized theories because V_i can be different for different states and $K(k)$ can be different for different subsystems; there is no unique constant like \hbar in our framework.

We need an extra assumption to construct energy with $[\frac{\hbar}{t}]$, for example: *the pure states share the same state volume V_p in a theory.*⁹ This assumption is plausible though *a priori* not necessary for a theory to be self-consistent. Otherwise, we can also accept the generalized energy with dimension $[\frac{1}{t}]$.

To restore the classical and quantum energy value, we define the energy of the i th pure stationary state E_i to be related to \mathcal{E}_i via

$$E_i = \frac{\pi \mathcal{E}_i V_p \bar{k}^2}{4V_i}, \quad (85)$$

where $\bar{k}_n^2 = \int k^2 K(k) dk$ as above.

In the quantum case, $\bar{k}^2 = \hbar^2$, $V_p = V_i = 2\pi\hbar$. Thus, the relation between \mathcal{E}_i and E_i is: $E_i = \frac{\pi\hbar^2\mathcal{E}_i}{4}$. In the classical case, the energy eigenstates are continuous and the label i can naturally be replaced by the action I of the phase space orbit. The energy $E(I)$ of the given eigenstate also obeys the Eq. (85) relation with the corresponding $\mathcal{E}(I)$. The $\mathcal{E}(I)$, which do not have a clear meaning in classical mechanics, usually diverge. Otherwise, the corresponding $E(I)$ equals zero, which does not contribute to the evolution. We conclude that the definition of energy eigenvalue E_i from quantum theory is consistent with the definition of classical mechanical energy $E(I)$ within this framework. They are both special cases of Eq. (85).

C. Definition of Hamiltonian

Finally, we can define a function in phase space corresponding to a generalized Hamiltonian.

Definition 5 (Hamiltonian). The Hamiltonian is a phase space function $H(q, p)$. For a specific system it is given by

$$H(q, p) := \sum_i E_i V_{g_i} g_i, \quad (86)$$

where g_i are pure stationary states, with state volume V_{g_i} , and E_i are energy values corresponding to g_i .¹⁰ $V_i g_i$ is dimensionless so H has the dimension of E_i as expected.

Applying the Hamiltonian definition to the equation of motion (76) gives the following.

⁹When considering the interaction between quantum and classical systems, people tried to assign different Planck constants to different subsystems, but this has been argued to be impossible [43,44]. Nevertheless, there is also work pointing out special cases that avoid the prohibition [45]. Our model allows a nonassociative algebra which is seldom considered in the previous works, so we cannot directly apply the above results.

¹⁰Why is the quantum Hamiltonian only the combination of eigenvalues and eigenstates, but here we have an extra V_{g_i} ? Recall that the Wigner function is the Wigner transformation of density matrices with an extra factor, $W(q, p) = \frac{1}{\hbar} \text{Wigner}\{\hat{\rho}\}$. We always have an extra factor in phase space.

Theorem 5 (Equation of motion in terms of Hamiltonian). The time evolution of a state $f(q, p)$ can be represented by

$$\frac{\partial f}{\partial t} = \left(\frac{4}{\pi V_p \bar{k}^2} \right) \text{Im} \int f(q+l, p+j) H(q+y, p+z) \times K(k) e^{i2(jy-lz)/k} d\Omega, \quad (87)$$

where $d\Omega = dydzdl djdk$ and the nonlocalized dynamics kernel $K(k)$ is a theory-specific function specifying the jumping in phase space, as we discussed below Eq. (76). Again $\bar{k}^2 = \int k^2 K(k) dk$. An equivalent expression is

$$\frac{\partial f}{\partial t} = \frac{4\pi}{V_p \bar{k}^2} \int K(k) k^2 f \sin\left(\frac{\Lambda k}{2}\right) H dk. \quad (88)$$

We can compute the expectation value of energy by $\langle E \rangle = \sum_i P(i|f) E_i$, where $P(i|f)$ is the probability of measuring the pure stationary state effect g_i given that the system is in the state f . By Definition 2 of state-dual measurements, $P(i|f) = \int V_{g_i} g_i f dq dp$. Combining that with the definition of the Hamiltonian (Definition 5), we have the following.

Theorem 6 (Energy measurement in terms of Hamiltonian). The expectation value of energy for state f is given by

$$\langle E \rangle = \int H f dq dp. \quad (89)$$

D. Examples

In quantum mechanics, the $K(k)$ is a single delta function $K(k) = \delta(\hbar - k)$, and $V_p = V_i = 2\pi\hbar$. The Hamiltonian is given by $H = \sum_i \hbar E_i g_i$, where g_i are the Wigner functions of energy eigenstates with eigenvalues E_i .

Equation (87) becomes

$$\frac{\partial f}{\partial t} = \left(\frac{2}{\pi^2 \hbar^3} \right) \text{Im} \int f(q+l, p+j) \times H(q+y, p+z) e^{i2(jy-lz)/\hbar} d\Omega. \quad (90)$$

This is equivalent to the original quantum equation of motion (59). (Taking the imaginary part is equivalent to taking the odd part inside the Fourier transform.)

Similarly, in the quantum case, Eq. (88) becomes

$$\frac{\partial f}{\partial t} = \frac{2}{\hbar} f \sin\left(\frac{\hbar}{2}\Lambda\right) H, \quad (91)$$

which is exactly the quantum equation (10).

In classical mechanics, there are infinite pure stationary states. $H(q, p) = \int E_i \left(\frac{1}{2\pi}\delta(I(q, p) - I_i)\right) 2\pi dI_i$, where $[\frac{1}{2\pi}\delta(I(q, p) - I_i)]$ are the normalized pure stationary states with state volume $2\pi dI_i$. $K(k) = \delta(k)$ in the classical limit. We can take the limit of quantum equation (91),

$$\frac{\partial f}{\partial t} = \lim_{\hbar \rightarrow 0} \frac{2}{\hbar} f \sin\left(\frac{\hbar}{2}\Lambda\right) H, \quad (92)$$

it is obvious that only the $\frac{\hbar}{2}\Lambda$ term in the $\sin(\frac{\hbar}{2}\Lambda)$ has a nonzero contribution. We get

$$\frac{\partial f}{\partial t} = f \Lambda H, \quad (93)$$

which is the classical Liouville equation.

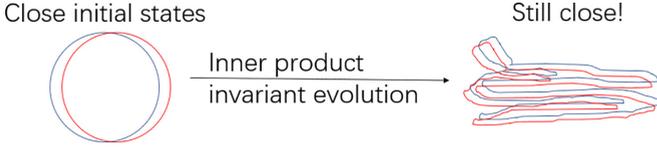


FIG. 6. A sketch of how inner product invariance together with finite state volume restricts the evolution of perturbed states. When there is a finite state volume, there is a nonzero overlap of a state and a perturbed version of it. The demand that the inner product of the two states is invariant then restricts the motion significantly, counteracting the possibility of the final state being strongly sensitive to initial perturbations.

The limit of the apparently nonlocalized expression Eq. (90) is more tricky. Like in path integrals, when the phase of the integrand is so fast, only the stagnation point will contribute. Since the apparently localized and nonlocalized expressions are equivalent, one finds $\frac{\partial f}{\partial t} = f \wedge H$.

VIII. STATE VOLUME AND CHAOS

One example where quantum and classical theories appear to differ qualitatively is *chaos*. A common qualitative definition of (classical) chaos is as follows. If two close states' distance grows exponentially during evolution, we say the system is chaotic because a small error in the initial condition will destroy predictability [26,27]. Classical chaos is known to be quite universal, e.g., in turbulence and weather systems. However, although quantum chaos' definition is still being debated [26,27,46], it must be different from the classical case because the Schrödinger equation is a linear function that cannot exponentially magnify the perturbation [46]. Here we explain the different behaviors in the chaos aspect by our framework.

The state volume explains the lack of sensitivity to initial state perturbation in quantum systems, as was noticed early [47,48]. If states have finite state volume, the inner product invariance protects the perturbation from exponential growth (as depicted in Fig. 6). Similarly, if we consider a nonzero state volume state, like a Gaussian distribution, in a chaotic classical system, a perturbation to it will not lead to “chaos.” Because the perturbed distribution value changes negligibly at any point, you can easily estimate the evolution of the perturbed Gaussian distribution by the known evolution of the unperturbed state. Instead, simulating the evolution of a single Gaussian distribution requires heavy computation when it contains the chaotic evolution of infinite points. The pure quantum states themselves have nonzero state volume. Therefore we conclude that reversible quantum evolution is stable under small perturbations to states.¹¹ On the other hand, when the perturbation is much larger compared with the state volume, the restriction given by inner product invariance is negligible.

¹¹For a classically chaotic Hamiltonian, will it (always) be difficult to simulate the evolution of a single quantum state? If not, how does the quantum evolution ease the simulation? This question deserves further study.

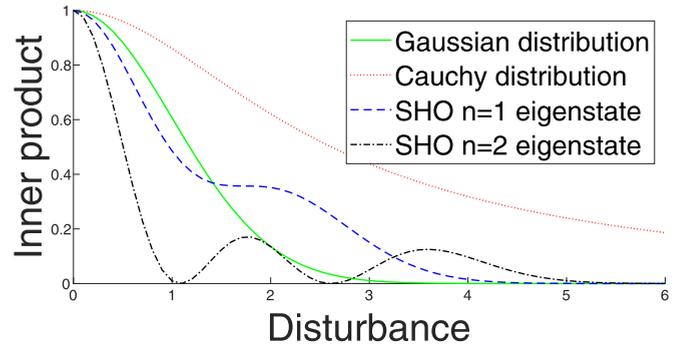


FIG. 7. The function between the inner product with the original state and the disturbance Δq for Gaussian distributions $[\exp(-\frac{q^2-p^2}{\hbar})]$, Cauchy distribution $(\frac{\hbar}{q^2+p^2+\hbar})$, the n th eigenstates of the simple harmonic oscillator [SHO, $H = k(q^2 + p^2)$]. We normalized the inner products such that the inner products with themselves equal 1. The disturbance is plotted in units of $\hbar^{\frac{1}{2}}$. The nonmonotonicity and possible zero points are interesting phenomena for future studies.

We construct such a model: an initial state f_I will evolve to f_F under some inner-product-invariant evolution. If we perturb the initial state to f'_I , how much do we know about the final state f'_F ? When the evolution is chaotic, we assume all the information about f'_F is given by the inner product invariance, which says f'_F should be as close to f_F as f'_I close to f_I in terms of inner product (Fig. 6).

For simplicity, consider a Gaussian wave packet $f_I = e^{-\frac{q^2-p^2}{2V_I/\pi}}$ and the perturbed one $f'_I = e^{-\frac{(q+\Delta x)^2-p^2}{2V_I/\pi}}$. The Δx is a random perturbation, we assume $\Delta x \in [-\sigma, \sigma]$, where σ represents the strength of perturbation. The perturbation Δx may produce a set of states f'_I , whose inner products with f_I are bounded by

$$\langle f'_I, f_I \rangle \geq I \left(\frac{\sigma}{\sqrt{V_I}} \right). \quad (94)$$

The function I for Gaussian states can be seen in Fig. 7.

The only information of the final state f'_F comes from the inner product bound in Eq. (94). The information can be reflected by entropy $S = -\sum_i p_i \log_2(\frac{1}{p_i})$, which can be estimated by $S \sim \log_2(N)$, where N means the number of possible perturbed states. Since the evolution is a one-to-one map, we can count the number of possible initial states that satisfy Eq. (94), which is equal to the number of possible final states that satisfy Eq. (94).

Only an area $A \sim \sigma^2$ of states can satisfy inequality (94), which can be counted as $N = \frac{A}{V_p} \sim \frac{\sigma^2}{V_p}$ pure states. Here we count the states by the pure state's volume V_p and assume it is a constant for different pure states. Notice the initial state f_I might not be a pure state; therefore, we distinguish its state volume V_I from V_p .

Now we can estimate the entropy

$$S \sim \log_2 \left(\frac{V_I}{V_p} + N \right) \sim \log_2 \left(\frac{\sigma^2 + V_I}{V_p} \right), \quad (95)$$

where we have added a V_I term because when there is no perturbation at all, $\sigma = 0$, we expect the entropy $S \sim \log_2(\frac{V_I}{V_p})$,

which is the entropy of f_I . the above equation shows that, for a certain initial state with V_I , the ratio $\frac{\sigma^2}{V_p}$ determines the upper bound of the perturbed evolution's uncertainty.

Above, we considered the case of finite state volume. What is the situation in the classical case? First, as $V_p \rightarrow 0$, the entropy blows up, which is common in the classical state count. However, there is another issue: originally, we estimate the area of possible states $A \sim \sigma^2$ under the inner product bound Eq. (94), but Eq. (94) becomes trivial for classical pure states. All the inner products between different delta functions $\delta(q - q_0)\delta(p - p_0)$ are 0. Therefore, we cannot learn anything from inner product invariance, which means

$$A \begin{cases} \sim \sigma^2 & \text{for } V_I > 0, \\ = \infty & \text{for } V_I = 0. \end{cases} \quad (96)$$

This relationship does capture the sensitivity to the initial state in classical mechanics. However, the discontinuity of A from a very small V_I to $V_I = 0$ is disturbing. We can modify our model to make it smooth. We introduce an uncertainty in the inner product ΔI , which is a constant that might be caused by experimental limitations. The ΔI causes an extra term $\frac{d\Delta x}{dt} \Delta I$ in the perturbation. The relation between I and Δx is shown in Fig. 7. When $V_I \rightarrow 0$, $\frac{\Delta x}{\sqrt{V_I}}$ and $\frac{d\Delta x}{dt}$ go to infinity. A small uncertainty in inner product ΔI will cause a large uncertainty area A in the classical limit.

After introducing ΔI , the modified model shows

$$S \sim \log_2 \left(\frac{V_I + \sigma^2 + (\Delta I \frac{d\Delta x}{dt})^2}{V_p} \right), \quad (97)$$

where S is the upper bound entropy of the perturbed final state constrained by inner product variance, V_I is the state volume of the initial state (can be mixed), σ represents the strength of perturbation ($\Delta x \sim \sigma$), ΔI is an uncertainty in the inner product ($\frac{\partial I}{\partial \Delta x}$ is a function of Δx , but we roughly take it as a constant here), and V_p is state volume of pure state, which is used to count how many pure states are possible final states. By this relation, we can see the state volume determines the restriction given by inner product invariance. Although choosing this specific model to do this semiquantitative analysis, we believe a similar tendency exists in the general case: the nonzero state volume restricts the system from evolving chaotically with the help of inner product invariance. Quantum states have finite state volume, while classical pure states have zero state volume, this difference explains why classical systems can be sensitive to the initial state, but quantum systems are not.

IX. TWO EXAMPLES OTHER THAN CLASSICAL AND QUANTUM MECHANICS

A. Case of information-restricted quantum mechanics

We now create another example theory and analyze how our results apply there. In this theory, there is a limit to our knowledge about what pure quantum state a quantum system is in. This leads to a different relation between state volume and the evolution-related nonlocalized dynamics kernel, i.e., splitting between the generalized Planck constants associated with states and evolution, respectively. This theory can be called information-restricted quantum mechanics.

The information-restricted quantum mechanics is one of the postquantum theories that can also be contained in our framework. In this theory, pure and pure stationary states have uniform state volumes but differ from κ in $K(k) = \delta(k - \kappa)$. Consider modified quantum mechanics where all the pure states have the following form of the density matrix (in a particular basis):

$$\frac{1}{2} \begin{bmatrix} \dots & & & & & \\ & |a|^2 & 0 & b^*a & 0 & \\ & 0 & |a|^2 & 0 & b^*a & \\ & a^*b & 0 & |b|^2 & 0 & \\ & 0 & a^*b & 0 & |b|^2 & \\ & & & & & \dots \end{bmatrix}, \quad (98)$$

where a, b are arbitrary complex numbers as long as the density matrix is normalized. Repeated numbers like $|a|^2, |b|^2$ are the restriction to density matrices ("..." contains c, d, \dots terms in a similar form). In quantum mechanics, such a density matrix is a uniform mixture of two orthogonal pure states,

$$\begin{bmatrix} \dots & & & & & \\ & |a|^2 & 0 & b^*a & 0 & \\ & 0 & 0 & 0 & 0 & \\ & a^*b & 0 & |b|^2 & 0 & \\ & 0 & 0 & 0 & 0 & \\ & & & & & \dots \end{bmatrix}, \quad (99)$$

$$\begin{bmatrix} \dots & & & & & \\ & 0 & 0 & 0 & 0 & \\ & 0 & |a|^2 & 0 & b^*a & \\ & 0 & 0 & 0 & 0 & \\ & 0 & a^*b & 0 & |b|^2 & \\ & & & & & \dots \end{bmatrix}.$$

(We can always find a basis where arbitrary two orthogonal states' density matrices have such a form.) Such a coarse graining doubles the state volume of pure states, $V = 2V_{\text{quantum pure state}} = 2h$ by Theorem 2. We assume the evolution is the same as in the standard quantum mechanics. Then the state volumes of pure stationary states have similarly been doubled, while the nonlocalized dynamics kernel and equation of motion remain the same.

B. Spekkens' toy model

Beyond the phase space formalism, the idea of generalized energy eigenstates has broader application. Here we will discuss a discrete system Spekkens' toy model [42], which has also been discussed in a GPT context [16,49,50]. Spekkens' toy model assumes four ontic states, labeled 1,2,3,4. Additionally, there is an information restriction: all the measurements can confirm a state is in only one pair of ontic states or the other pair, for example, $1 \vee 2$ or $3 \vee 4$. Every measurement inevitably disturbs the ontic states, ensuring that consecutive measurements cannot gain better knowledge. Consequently, there are six pure epistemic states, $1 \vee 2, 1 \vee 3, 1 \vee 4, 2 \vee 3, 2 \vee 4, 3 \vee 4$. The transformation of these epistemic states depends on the permutation of ontic states, which can be classified by the following:

(1) All the ontic states are stationary. All the pure epistemic states are pure stationary states.

(2) Two ontic states are stationary. For example, only 1,2 are stationary, pure stationary states are $1 \vee 2$, $3 \vee 4$, $1 \vee 3 + 1 \vee 4$, $2 \vee 3 + 2 \vee 4$. (+ here represents a uniform probabilistic mixture.)¹²

(3) One ontic state is stationary. For example, 1 is stationary, pure stationary states are $1 \vee 2 + 1 \vee 3 + 1 \vee 4$, $2 \vee 3 + 2 \vee 4 + 3 \vee 4$.

(4) Permutation in pairs. For example, 1,2 permute, 3,4 permute, pure stationary states are $1 \vee 2$, $3 \vee 4$.

(5) Cyclic permutation of four ontic states. The only pure stationary state is the maximally mixed state, $1 \vee 2 \vee 3 \vee 4$.

Pure stationary states can determine the evolution except for the cyclic permutation of four ontic states. The only pure stationary state is the maximally mixed state. We cannot distinguish whether it is $1 \mapsto 2 \mapsto 3 \mapsto 4 \mapsto 1$ or $1 \mapsto 3 \mapsto 2 \mapsto 4 \mapsto 1$ or other possibilities. As we have mentioned, deriving an equation of motion from pure stationary states is generally a nontrivial task. For Spekkens' toy model, we need more conditions to derive a unique equation of motion.

Meanwhile, we can also find that the pure stationary states cannot provide a state-dual measurement. For example, the permutation of 1,2,3 has the pure stationary states $1 \vee 2 + 2 \vee 3 + 3 \vee 1$ and $1 \vee 4 + 2 \vee 4 + 3 \vee 4$, which are nonorthogonal, so there is no measurement can satisfy Definition 2. The model violates the Postulate 6, measurements are so limited that a state-dual energy measurement does not always exist.

We can also generalize the effect or state volume to Spekkens' toy model. It still gives a proper normalization factor in measurement. Following Theorem 3 which gives volume, we can choose a form of the inner product to define the state volume. The dot product is a natural choice if we hope all the ontic states are on equal footing. (Other definitions of the inner product also work.) Then the state volumes in Spekkens' toy model can be defined by

$$V_f = \frac{1}{\sum_{i=1}^4 f_i^2}, \quad (100)$$

where f_i represent the probabilities of the i th ontic state. All the pure epistemic states share the same state volume 2, exactly the number of possible ontic states. Likewise, the state volume of the maximally mixed state is 4. These particular state volumes are the consequences of the inner product, which depends on the symmetries we demand. Finally, we test the corresponding state-dual measurement. Consider the probability from $1 \vee 2$ to $1 \vee 3$:

$$P = V_{1 \vee 3} \vec{f}_{1 \vee 2} \cdot \vec{f}_{1 \vee 3} = \frac{1}{2}. \quad (101)$$

The effect volume which equals state volume gives the outcome as expected. The example of Spekkens' toy model shows that our framework also, at least partially, works on discrete systems.

X. SUMMARY AND OUTLOOK

We built a framework describing a generalized energy concept and time evolution rule, which describes quantum

and classical in a unified way. We introduced six postulates: (1) canonical symmetries; (2) local inner product; (3) pure stationary states decide the evolution; (4) pure stationary states are independent; (5) inner product invariance; and (6) existence of energy measurement. Postulates (1) and (2) provide a unique inner product in phase space, which helps to define state-dual measurements. Postulates (1) and (3)–(5) derive our generalized equation of motion (76) in phase space. Based on the above results, (6) further guarantees the existence of a state-dual measurement of the Hamiltonian, i.e., the conserved observable describing time evolution. Rather than taking an algebraic approach, we endeavored to make every postulate have a clear physical or operational meaning. We derived a generalized Hamiltonian system in phase space that encompasses quantum and classical theories but also generalizes the original ideas. This includes generalizing Planck's constant.

In our framework, the Planck constant provides the state or effect volume of pure states and corresponding measurements. It also appears in the equation of motion. In general, there is no good limit when taking a quantum state's volume to zero to get a classical state. Still, it is always possible to take the limit of the quantum equation of motion to get the corresponding classical evolution. The two roles are related. For example, physically nonlocalized evolution causes a negative distribution, so we need a nonzero state volume to avoid negative probability. However, they can have different values in general theories. This framework possesses the potential for application and advancement in various directions. First, there exists an intriguing connection between contextuality and evolution. Specifically, when $K(k)$ is not directly proportional to $\delta(k)$, the evolution associated with infinite-order derivatives [39,40] emerges alongside a negative distribution, which could relate to contextuality [51].

Second, it is possible to construct alternative forms of mechanics that do not conform to classical or quantum paradigms. One approach involves considering different values of the Planck constant in the equation of motion compared to its value in uncertainty. Another possibility entails selecting a non-delta function for $K(k)$, which can be interpreted as a probabilistic combination of diverse commutation relations during the process of evolution.

Third, this framework facilitates clear analogies and comparisons between quantum and classical dynamics. For instance, it can be employed to elucidate the apparent acceleration exhibited by quantum walks in contrast to classical walks [52].

Last, it is natural to utilize this framework to formulate a theory of thermodynamics that is not reliant on the underlying choice of mechanics.

ACKNOWLEDGMENTS

We thank Meng Fei, Chelvannithilan Sivapalan, Andrew Garner, and Dario Egloff for discussions. L.J. and O.D. acknowledge support from the National Natural Science Foundation of China (Grants No.12050410246, No.1200509, No.12050410245) and the City University of Hong Kong (Project No. 9610623).

¹²Strictly speaking, we are discussing a generalized Spekkens' toy model with convex state space.

APPENDIX A: SOME MATHEMATICAL ASPECTS OF THE PHASE SPACE FORMALISM

In the settings we consider the $2n$ local coordinates $z = (q, p)$ on \mathfrak{P} are given by the generalized coordinates of the configuration space $q = \{q^a\}$ and the generalized momenta $p = \{p_a\}$, $a = 1, \dots, n$. These momenta are related to the coordinates q and velocities \dot{q} via $q_a := \partial L / \partial \dot{x}^a$, where $L(q, \dot{q})$ is the system's Lagrangian. As the system is unconstrained, the n equations for momenta can be inverted to express the velocities as functions of positions and momenta.

\mathfrak{P} is a symplectic manifold as a nondegenerate closed 2-form is defined on it. It can always be written in local coordinates as

$$\omega^{(2)} = dp_a \wedge dq^a \equiv dp \wedge dq, \quad (\text{A1})$$

and in these coordinates, it is represented as a $2n \times 2n$ matrix

$$J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}. \quad (\text{A2})$$

The symplectic form establishes the isomorphism between vectors and 1-forms (covectors) on \mathfrak{P} by matching a vector η with the form $\omega_\eta^{(1)}$ via $\omega_\eta^{(1)}(\xi) := \omega^{(2)}(\eta, \xi)$. Hence the Hamilton (canonical) equations of motion are written as

$$\dot{z}(t) = \xi_H(z(t)), \quad (\text{A3})$$

representing the Hamiltonian phase flow

$$\xi_H = J \vec{\nabla}_z H. \quad (\text{A4})$$

The Poisson brackets of Eq. (2) are then

$$\{f, g\} = -\omega^{(2)}(\vec{\nabla}_z f, \vec{\nabla}_z g) = (\vec{\nabla}_z f)^T \cdot J \cdot \vec{\nabla}_z g. \quad (\text{A5})$$

Classical observables are smooth functions on the phase space and form the algebra. The Poisson bracket can be defined more abstractly as a Lie bracket on the underlying manifold: it is linear and antisymmetric and satisfies the Jacobi identity

$$\{f, \{g, h\}\} = \{\{f, g\}, h\} + \{g, \{f, h\}\}. \quad (\text{A6})$$

In addition, it satisfies the Leibnitz rule with respect to the product defining the algebra, $f \circ g(q, p) := f(q, p)g(q, p)$,

$$\{f, gh\} = \{f, g\}h + g\{f, h\}. \quad (\text{A7})$$

Technically this is the Jordan-Lee algebra with associative multiplication, i.e., the Poisson algebra.

APPENDIX B: WIGNER TRANSPORT EQUATION IN QUANTUM THEORY FOR GENERAL HAMILTONIAN

In this Appendix, we describe how to convert between different expressions of the time evolution in the quantum case. The original Wigner transportation function (11)

provides jumping only in the momentum direction. In that case, $J(q, j)$ can be viewed as a Fourier transform of the odd part of the potential, so strictly speaking, it is well defined only for restricted Hamiltonian. However, if you accept derivatives of delta functions as the Fourier transform of polynomials, this formula can give a reasonable description for general Hamiltonians.

1. Case of $H = p^2/2m$

As an example, we can first generalize this idea to the kinetic term of Hamiltonian $H = p^2/2m$. We attempt to find an apparently nonlocalized expression that

$$\frac{\partial W}{\partial t} = \int W(q + j, p) J_q(p, j) dj = -\frac{p}{m} \frac{\partial W}{\partial q}. \quad (\text{B1})$$

(It is also the equation of motion for classical free particles.) The delta function's derivatives can be defined through partial integral,

$$\int \delta'(x - x_0) f(x) dx = -f'(x_0). \quad (\text{B2})$$

Therefore, the kinetic effect can be expressed by

$$\frac{\partial W}{\partial t} = \frac{p}{m} \int W(q + l, p) \delta'(l) dl. \quad (\text{B3})$$

After this, we check if the Wigner transportation function also works for the kinetic term:

$$J_q(p, l) = \frac{p}{m} \delta'(l) = \frac{p}{2\pi m} \int (-2iy/\hbar) e^{-2ily/\hbar} d(-2y/\hbar) \quad (\text{B4})$$

$$= \frac{-i}{2\pi \hbar} \int \frac{4py}{2m} e^{-2ily/\hbar} d(-2y/\hbar) \quad (\text{B5})$$

$$= \frac{i}{\pi \hbar^2} \int [T(p + y) - T(p - y)] e^{-2ily/\hbar} dy \quad (\text{B6})$$

$$\left[\delta(x) = \frac{1}{2\pi} \int e^{ikx} dk, \delta'(x) = \frac{1}{2\pi} \int ik e^{ikx} dk \right].$$

Therefore, we can see that the Wigner transportation function also works for the power series once we introduce the derivatives of the delta function as the Fourier transform of the power series.

2. General H

Based on the above idea, we check the generalized apparently nonlocalized Eq. (59) to see if it can convert to the original apparently localized Eq. (10). Use the Taylor expansion of multivariate function to expand $H(q - y, p + z) - H(q + y, p - z)$ terms in Eq. (59):

$$\begin{aligned} & H(q - y, p + z) - H(q + y, p - z) \\ &= -2y \frac{\partial H}{\partial q} + 2z \frac{\partial H}{\partial p} + \dots \\ &= \sum_{m, n | m, n \in \mathbb{N}, m+n \text{ odd}} 2((-y)^m z^n) \frac{C_{m+n}^m}{(m+n)!} \frac{\partial^{m+n} H}{\partial q^m \partial p^n}, \quad (\text{B7}) \end{aligned}$$

$$J = \sum_{m,n|m,n \in \mathbb{N}, m+n \in \text{odd}} 2 \frac{i}{\hbar} \left(\frac{i\hbar}{2} \right)^{m+n} (-1)^m \delta^{(m)}(j) \times \delta^{(n)}(l) \frac{C_{m+n}^m}{(m+n)!} \frac{\partial^{m+n} H}{\partial q^m \partial p^n}, \quad (\text{B8})$$

$$\frac{\partial W}{\partial t} = \sum_{m,n|m,n \in \mathbb{N}, m+n \in \text{odd}} 2 \frac{i}{\hbar} \left(\frac{i\hbar}{2} \right)^{m+n} (-1)^m (-1)^{m+n} \times \frac{\partial^{n+m} W}{\partial q^n \partial p^m} \frac{C_{m+n}^m}{(m+n)!} \frac{\partial^{m+n} H}{\partial q^m \partial p^n}, \quad (\text{B9})$$

$$\frac{\partial W}{\partial t} = \frac{-2}{\hbar} \sum_{m,n|m,n \in \mathbb{N}, m+n \in \text{odd}} \left(\frac{\hbar}{2} \right)^{m+n} (-1)^{\frac{m+n+1}{2}} (-1)^m \times \frac{\partial^{n+m} W}{\partial q^n \partial p^m} \frac{C_{m+n}^m}{(m+n)!} \frac{\partial^{m+n} H}{\partial q^m \partial p^n}. \quad (\text{B10})$$

On the other side, expand Eq. (10)

$$\begin{aligned} \frac{\partial W}{\partial t} &= \frac{-2}{\hbar} H \sin \left(\frac{\hbar \Lambda}{2} \right) W(q, p) \\ &= \frac{-2}{\hbar} \sum_{a,b|a,b \in \mathbb{N}, a+b \in \text{odd}} \frac{C_{a+b}^a}{(a+b)!} \left(\frac{\hbar}{2} \right)^{a+b} (-1)^{\frac{a+b-1}{2}} \\ &\quad \times (-1)^b \frac{\partial^{a+b} H}{\partial q^b \partial p^a} \frac{\partial^{a+b} W}{\partial q^a \partial p^b}. \end{aligned} \quad (\text{B11})$$

Notice that $m+n$ is odd, so $(-1)^m = -(-1)^n$, two equations [(10) and (59)] are completely equivalent.

APPENDIX C: SYMMETRY IN THE EQUATION OF MOTION

We first take time-reversal symmetry as an example to show how it restricts J^0 . The reversal symmetry operation changes the following elements: $\frac{\partial f'}{\partial t}(0, 0) = -\frac{\partial f}{\partial t}(0, 0)$, $g'(q, p) = g(q, -p)$, $f'(l, j) = f(l, -j)$. Substitute them into Eq. (67), $\frac{\partial f'}{\partial t}(0, 0) = \int J^0_{g'}(l, j) f'(l, j) dl dj$:

$$-\frac{\partial f}{\partial t}(0, 0) = \int J^0_{g'}(l, j) f(l, -j) dl dj, \quad (\text{C1})$$

$$\frac{\partial f}{\partial t}(0, 0) = \int (-J^0_{g'}(l, -j)) f(l, j) dl dj. \quad (\text{C2})$$

Since f can be an arbitrary function,

$$-J^0_{g'}(l, -j) = J^0_g(l, j). \quad (\text{C3})$$

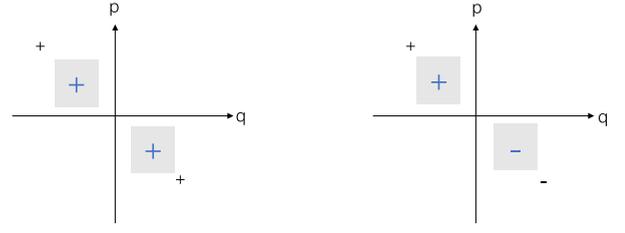
The symmetry operation has transformed $g \rightarrow g'$. We can decompose g by eigenstates of the operation $g_1(q, p) = g_1(q, -p)$ and $g_2(q, p) = -g_2(q, -p)$, and the corresponding J^0 has the symmetry

$$J^0_{g_1}(l, -j) = -J^0_{g_1}(l, j), \quad J^0_{g_2}(l, -j) = J^0_{g_2}(l, j). \quad (\text{C4})$$

We have learned the symmetry of J corresponding to eigenstates of symmetry operation.

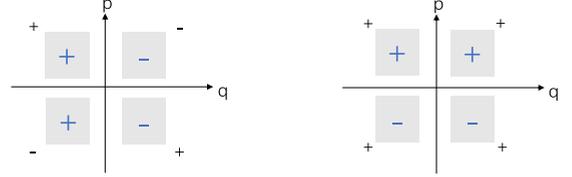
We consider three symmetry operations in total: time reversal, parity (the composition of time reversals and switches, $(q, p, t) \mapsto (-q, -p, t)$), and switch+time reversal+switch

Parity

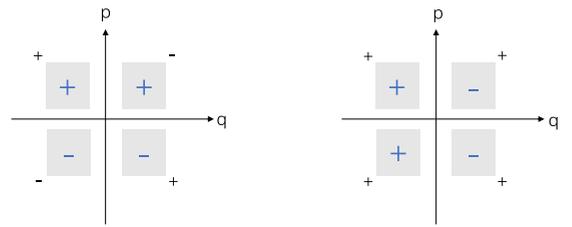


Even pure stationary state

Time reversal

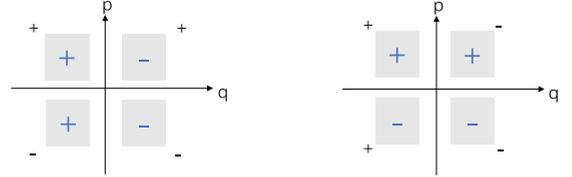


Switch + Time reversal + Switch



Odd pure stationary state

Time reversal



Switch + Time reversal + Switch

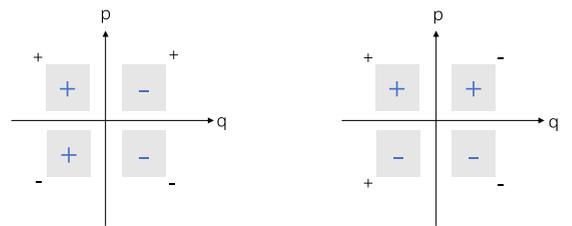


FIG. 8. The black +, - notations denote the delta functions as components of the pure stationary state, and blue +, - with gray shading denote the required symmetry of $J^0(l, j)$. The first picture shows the symmetry of \tilde{J} required by parity, and the following pictures show the requirement by time reversal and switch+time reversal+switch given even and odd pure stationary states. The left and right figures show different eigenstates with eigenvalue ± 1 , respectively. We find for an even pure stationary state, the time-reversal symmetry contradicts the switch+time-reversal+switch symmetry unless $J^0_{g_{\text{even}}} = 0$.

$(q, p, t) \mapsto (-q, p, -t)$. We decompose pure stationary states g by joint eigenstates of symmetry operations with eigenvalue ± 1 . First, even and odd functions (eigenstates of parity):

$$g = g_{\text{odd}} + g_{\text{even}},$$

where

$$g_{\text{odd}}(a, b) = -g_{\text{odd}}(-a, -b), \quad (\text{C5})$$

$$g_{\text{even}}(a, b) = g_{\text{even}}(-a, -b).$$

We can further decompose the odd and even parts by the joint eigenstates of parity, time reversal, and switch+time reversal+switch. The even part can be decomposed by delta functions:

$$g_{\text{even}} = \int \rho_1(q_0, p_0) g_{e1} + \rho_2(q_0, p_0) g_{e2} dq_0 dp_0,$$

where $g_{e1}(q, p, q_0, p_0)$

$$= \delta(q - q_0)\delta(p - p_0) + \delta(q - q_0)\delta(p + p_0) + \delta(q + q_0)\delta(p - p_0) + \delta(q + q_0)\delta(p + p_0),$$

$g_{e2}(q, p, q_0, p_0)$

$$= \delta(q - q_0)\delta(p - p_0) - \delta(q - q_0)\delta(p + p_0) - \delta(q + q_0)\delta(p - p_0) + \delta(q + q_0)\delta(p + p_0), \quad (\text{C6})$$

ρ_1 and ρ_2 are corresponding weight of g_{e1} and g_{e2} . The odd part can be decomposed by similar delta functions.

Figure 8 illustrates the eigenstates of parity, time reversal, and switch+time reversal+switch symmetry, and how their J^0 is restricted from symmetries. For an even pure stationary state, the time-reversal symmetry contradicts the switch+time-reversal+switch symmetry unless $J^0_{g_{\text{even}}} = 0$. For an odd pure stationary state, the requirements are consistent. $J^0_{g_{\text{odd}}}(l, j)$ only has to be an odd function due to the parity symmetry.

Therefore, only the odd part of the pure stationary state contributes to the J , and

$$J^0(l, j) = -J^0(-l, -j) \quad (\text{C7})$$

for arbitrary pure stationary states. We get Lemma 3.

APPENDIX D: OPERATOR FORM OF THE EQUATION OF MOTION

The Wigner transform is

$$A(q, p) = \int dz e^{i\frac{2}{\hbar}pz} \langle q - z | \hat{A} | q + z \rangle. \quad (\text{D1})$$

However, we can use an arbitrary factor labeled by μ instead of \hbar to do the Wigner transform. Consider the following expression:

$$\int A(q + l, p + j) B(q + y, p + z) K(k) \times e^{-i\frac{2}{k}(jy-lz)} dk dy dz dl dj, \quad (\text{D2})$$

We replace phase space functions A and B by their operator forms:

$$= \int \langle q + l - m | \hat{A}_\mu | q + l + m \rangle e^{im\frac{2}{\mu}(p+j)} \langle q + y - n | \hat{B}_\mu | q + y + n \rangle e^{in\frac{2}{\mu}(p+z)} K(k) e^{-i\frac{2}{k}(jy-lz)} dk dy dz dl dj dn dm, \quad (\text{D3})$$

$$= \int \langle q + l - m | \hat{A}_\mu | q + l + m \rangle \langle q + y - n | \hat{B}_\mu | q + y + n \rangle e^{i\frac{2}{\mu}(m+n)p} K(k) e^{i(\frac{2}{\mu}m - \frac{2}{k}y)j} e^{i(\frac{2}{\mu}n + \frac{2}{k}l)z} dk dy dz dl dj dn dm, \quad (\text{D4})$$

$$= 4\pi^2 \int \langle q + l - m | \hat{A}_\mu | q + l + m \rangle \langle q + y - n | \hat{B}_\mu | q + y + n \rangle e^{i\frac{2}{\mu}(m+n)p} K(k) \delta\left(\frac{2}{\mu}m - \frac{2}{k}y\right) \delta\left(\frac{2}{\mu}n + \frac{2}{k}l\right) dk dy dl dn dm, \quad (\text{D5})$$

$$= 4\pi^2 \int \langle q - \frac{kn}{\mu} - m | \hat{A}_\mu | q - \frac{kn}{\mu} + m \rangle \langle q + \frac{km}{\mu} - n | \hat{B}_\mu | q + \frac{km}{\mu} + n \rangle e^{i\frac{2}{\mu}(m+n)p} K(k) \left(\frac{k}{2}\right)^2 dk dn dm. \quad (\text{D6})$$

The notations like \hat{A}_μ mean $A = \int dz e^{i\frac{2}{\mu}pz} \langle r - z/2 | \hat{A}_\mu | r + z/2 \rangle$, μ take the place of \hbar in the original Weyl transform. To the same phase space function, different μ lead to different operators. We are free to choose the value of μ , let $\mu = k$:

$$= 4\pi^2 \int \langle q - n - m | \hat{A}_k | q - n + m \rangle \langle q + m - n | \hat{B}_k | q + m + n \rangle e^{i\frac{2}{k}(m+n)p} K(k) \left(\frac{k}{2}\right)^2 dm dn dk. \quad (\text{D7})$$

Let $m + n = a$, $m - n = b$, and then replace $q + b$ by r' :

$$= 4\pi^2 \int \langle q - a | \hat{A}_k | q + b \rangle \langle q + b | \hat{B}_k | q + a \rangle e^{i\frac{2}{k}bp} K(k) \left(\frac{k}{2}\right)^2 da db dk$$

$$= 4\pi^2 \int \langle q - a | \hat{A}_k | r' \rangle \langle r' | \hat{B}_k | q + a \rangle e^{i\frac{2}{k}bp} K(k) \left(\frac{k}{2}\right)^2 da dr' dk. \quad (\text{D8})$$

Compare the result with the Wigner transform of the product $\hat{A}\hat{B}$:

$$AB = \int dz dr' e^{i\frac{2}{\hbar}pz} \langle q - z | \hat{A} | r' \rangle \langle r' | \hat{B} | q + z \rangle, \quad (\text{D9})$$

we find

$$\int A(q+l, p+j)B(q+y, p+z)K(k)e^{-i\frac{2}{k}(jy-lz)} dk dy dz dl dj = \pi^2 \int k^2 K(k) \text{Wigner}_k\{\hat{A}_k \hat{B}_k\} dk. \quad (\text{D10})$$

The l.h.s. of Eq. (76) takes only the imaginary part, which can be computed by $\text{Im}f = \frac{f-f^*}{2i}$. Therefore, the overall result leads to a commutator-like equation, where the commutation relation is given by the distribution $K(k)$:

$$\begin{aligned} & \text{Im} \int f(q+l, p+j)g(q+y, p+z)K(k)e^{-i\frac{2(jy-lz)}{k}} dk dy dz dl dj \\ &= \frac{1}{2i} \int [f(q+l, p+j)g(q+y, p+z) - f(q+y, p+z)g(q+l, p+j)]K(k)e^{-i\frac{2(jy-lz)}{k}} dk dy dz dl dj \\ &= -i\frac{\pi^2}{2} \int k^2 K(k) \text{Wigner}_k\{[\hat{f}_k, \hat{g}_k]\} dk. \end{aligned} \quad (\text{D11})$$

APPENDIX E: ASSOCIATIVE MOYAL PRODUCT AND NONASSOCIATIVE HYBRID MOYAL PRODUCT

Observe that

$$\partial_x(f(x, y)|_{x=y}) = ((\partial_x + \partial_y)f(x, y))|_{x=y}.$$

We introduce an notation $E(\partial_x, \partial_y)f(x)g(y)|_{x=y} := f(q, p)e^{i\frac{1}{2}\Lambda}g(q, p)$, where x, y represent the vector (q, p) .

We consider an example of the hybrid Moyal product, $K(k) = \delta(k - k_1) + \delta(k - k_2)$, $f \star_H g := E_1(\partial_x, \partial_y) + E_2(\partial_x, \partial_y)f(x)g(y)$.

Compute

$$(f \star_H g) \star_H h = [E_1(\partial_x, \partial_z) + E_2(\partial_x, \partial_z)][\{[E_1(\partial_x, \partial_y) + E_2(\partial_x, \partial_y)]f(x)g(y)|_{x=y}\}h(z)|_x=z] \quad (\text{E1})$$

$$= [E_1(\partial_x + \partial_y, \partial_z) + E_2(\partial_x + \partial_y, \partial_z)][E_1(\partial_x, \partial_y) + E_2(\partial_x, \partial_y)]f(x)g(y)h(z)|_{x=y=z} \quad (\text{E2})$$

$$= [E_1(\partial_x, \partial_z)E_1(\partial_y, \partial_z) + E_2(\partial_x, \partial_z)E_2(\partial_y, \partial_z)][E_1(\partial_x, \partial_y) + E_2(\partial_x, \partial_y)]f(x)g(y)h(z)|_{x=y=z} \quad (\text{E3})$$

$$= [E_1(\partial_x, \partial_z)E_1(\partial_y, \partial_z)E_1(\partial_x, \partial_y) + E_2(\partial_x, \partial_z)E_2(\partial_y, \partial_z)E_2(\partial_x, \partial_y) \quad (\text{E4})$$

$$+ E_1(\partial_x, \partial_z)E_1(\partial_y, \partial_z)E_2(\partial_x, \partial_y) + E_2(\partial_x, \partial_z)E_2(\partial_y, \partial_z)E_1(\partial_x, \partial_y)]f(x)g(y)h(z)|_{x=y=z}. \quad (\text{E5})$$

Similarly,

$$\begin{aligned} f \star_H (g \star_H h) &= [E_1(\partial_x, \partial_z)E_1(\partial_y, \partial_z)E_1(\partial_x, \partial_y) + E_2(\partial_x, \partial_z)E_2(\partial_y, \partial_z)E_2(\partial_x, \partial_y) \\ &+ E_1(\partial_x, \partial_y)E_1(\partial_x, \partial_z)E_2(\partial_y, \partial_z) + E_2(\partial_x, \partial_y)E_2(\partial_x, \partial_z)E_1(\partial_y, \partial_z)]f(x)g(y)h(z)|_{x=y=z}. \end{aligned} \quad (\text{E6})$$

If $E_1 \neq 0, E_2 = 0$, namely, a conventional Moyal product, we find the two results coincide, so the Moyal product is associative. However, the hybrid Moyal product is not associative in general.

-
- [1] V. I. Arnold, *Mathematical Methods of Classical Mechanics*, Graduate Texts in Mathematics, Vol. 60 (Springer Science & Business Media, New York, Heidelberg, Berlin, 2013).
- [2] V. I. Arnold, V. V. Kozlov, and A. I. Neishtadt, *Mathematical Aspects of Classical and Celestial Mechanics* (Springer, Berlin, 2006).
- [3] N. P. Landsman, *Foundations of Quantum Theory: From Classical Concepts to Operator Algebras* (Springer, Cham, Switzerland, 2017).
- [4] E. Noether, Nachrichten der Akademie der Wissenschaften in Göttingen, Mathematisch-Physikalische Klasse **2**, 235 (1918).
- [5] L. Hardy, [arXiv:quant-ph/0101012](https://arxiv.org/abs/quant-ph/0101012).
- [6] J. Barrett, *Phys. Rev. A* **75**, 032304 (2007).
- [7] P. Janotta and H. Hinrichsen, *J. Phys. A: Math. Theor.* **47**, 323001 (2014).
- [8] M. P. Müller and C. Ududec, *Phys. Rev. Lett.* **108**, 130401 (2012).
- [9] O. C. O. Dahlsten, A. J. P. Garner, and V. Vedral, *Nat. Commun.* **5**, 4592 (2014).
- [10] M. Plavála, *Phys. Rep.* **1033**, 1 (2023).
- [11] H.-T. Elze, *J. Phys.: Conf. Ser.* **442**, 012007 (2013).
- [12] D. R. Terno, [arXiv:2309.05014](https://arxiv.org/abs/2309.05014).
- [13] H. Barnum, M. P. Müller, and C. Ududec, *New J. Phys.* **16**, 123029 (2014).
- [14] D. Branford, O. C. O. Dahlsten, and A. J. P. Garner, *Found. Phys.* **48**, 982 (2018).
- [15] D. Ristè, M. P. da Silva, C. A. Ryan, A. W. Cross, A. D. Córcoles, J. A. Smolin, J. M. Gambetta, J. M. Chow, and B. R. Johnson, *npj Quantum Inf.* **3**, 16 (2017).
- [16] M. F. Pusey, *Found. Phys.* **42**, 688 (2012).
- [17] M. Plávala and M. Kleinmann, *Phys. Rev. Lett.* **128**, 040405 (2022).
- [18] M. Plávala and M. Kleinmann, *Phys. Rev. A* **108**, 052212 (2023).

- [19] L. Jiang, D. R. Terno, and O. Dahlsten, companion paper, *Phys. Rev. Lett.* **132**, 120201 (2024).
- [20] H. W. Lee, *Phys. Rep.* **259**, 147 (1995).
- [21] C. K. Zachos, D. B. Fairlie, and T. L. Curtright, editors, *Quantum Mechanics in Phase Space* (World Scientific, Singapore, 2005).
- [22] E. Wigner, *Phys. Rev.* **40**, 749 (1932).
- [23] M. Hillery, R. F. O'Connell, M. O. Scully, and E. P. Wigner, *Phys. Rep.* **106**, 121 (1984).
- [24] A. Royer, *Phys. Rev. A* **15**, 449 (1977).
- [25] Y. L. Lin and O. C. O. Dahlsten, *Phys. Rev. A* **102**, 062210 (2020).
- [26] S. Wimberger, *Nonlinear Dynamics and Quantum Chaos*, Graduate Texts in Physics, Vol. 10 (Springer, Cham, 2014).
- [27] L. D'Alessio, Y. Kafri, A. Polkovnikov, and M. Rigol, *Adv. Phys.* **65**, 239 (2016).
- [28] A. Peres, *Quantum Theory: Concepts and Methods* (Kluwer, 1995).
- [29] H. Weyl, *The Theory of Groups and Quantum Mechanics* (Courier Corporation, 1950).
- [30] H. Barnum and A. Wilce, *Electron. Notes Theor. Comput. Sci.* **270**, 3 (2011).
- [31] P. Busch, M. Grabowski, and P. Lahti, *Operational Quantum Physics* (Springer, Berlin, 1995).
- [32] P. Busch, P. Lahti, J.-P. Pellonpaa, and K. Ylino, *Quantum Measurement* (Springer, Cham, Switzerland, 2016).
- [33] R. Haag, *Local Quantum Physics: Fields, Particles, Algebras* (Springer, Berlin, 1996).
- [34] H. Goldstein, *Classical Mechanics* (Addison-Wesley, Reading MA, 1980).
- [35] B. O. Koopman, *Proc. Natl. Acad. Sci. USA* **17**, 315 (1931).
- [36] A. Peres and D. R. Terno, *Phys. Rev. A* **63**, 022101 (2001).
- [37] L. L. Salcedo, *Phys. Rev. A* **85**, 022127 (2012).
- [38] M. Reed and B. Simon, *I: Functional Analysis, Methods of Modern Mathematical Physics*, Vol. 1 (Academic Press, New York, 1981).
- [39] P. A. M. Dirac, *Phys. Rev.* **73**, 1092 (1948).
- [40] A. Pais and G. E. Uhlenbeck, *Phys. Rev.* **79**, 145 (1950).
- [41] T. Hakioglu and E. Tepedelenlioglu, *J. Phys. A: Math. Gen.* **33**, 6357 (2000).
- [42] R. W. Spekkens, *Phys. Rev. A* **75**, 032110 (2007).
- [43] J. Caro and L. L. Salcedo, *Phys. Rev. A* **60**, 842 (1999).
- [44] D. Sahoo, *J. Phys. A: Math. Gen.* **37**, 997 (2004).
- [45] M. J. Lake, [arXiv:2008.13183](https://arxiv.org/abs/2008.13183).
- [46] M. Berry, *Phys. Scr.* **40**, 335 (1989).
- [47] M. V. Berry, N. L. Balazs, M. Tabor, and A. Voros, *Ann. Phys.* **122**, 26 (1979).
- [48] H. J. Korsch and M. V. Berry, *Physica D* **3**, 627 (1981).
- [49] L. Catani and D. E. Browne, *New J. Phys.* **19**, 073035 (2017).
- [50] A. J. Garner, O. C. Dahlsten, Y. Nakata, M. Muraio, and V. Vedral, *New J. Phys.* **15**, 093044 (2013).
- [51] R. W. Spekkens, *Phys. Rev. Lett.* **101**, 020401 (2008).
- [52] S. E. Venegas-Andraca, *Quantum Info. Proc.* **11**, 1015 (2012).