Coherent-state path-integral calculation of the Wigner function

2000 J. Phys. A: Math. Gen. 33 5219

(http://iopscience.iop.org/0305-4470/33/29/306)

J. Phys. A: Math. Gen. **33** (2000) 5219–5229. Printed in the UK PII: S0305-4470(00)12959-2

Coherent-state path-integral calculation of the Wigner function

J H Samson

Department of Physics, Loughborough University, Loughborough, Leicestershire LE11 3TU, UK

E-mail: j.h.samson@lboro.ac.uk

Received 29 March 2000

Abstract. We consider a set of operators $\hat{\boldsymbol{x}} = (\hat{x}_1, \ldots, \hat{x}_N)$ with diagonal representatives $\boldsymbol{P}(\boldsymbol{n})$ in the space of generalized coherent states $|\boldsymbol{n}\rangle$: $\hat{\boldsymbol{x}} = \int \mathrm{d}\mu(\boldsymbol{n})\, \boldsymbol{P}(\boldsymbol{n})|\boldsymbol{n}\rangle\langle\boldsymbol{n}|$. We regularize the coherent-state path integral as a limit of a sequence of averages $\langle\ \rangle_L$ over polygonal paths with L vertices $\boldsymbol{n}_{1,\ldots,L}$. The distribution of the path centroid $\bar{\boldsymbol{P}} = \frac{1}{L}\sum_{l=1}^{L} \boldsymbol{P}(\boldsymbol{n}_l)$ tends to the Wigner function $W(\boldsymbol{x})$, the joint distribution for the operators: $W(\boldsymbol{x}) = \lim_{L\to\infty}\langle\delta_N(\boldsymbol{x} - \bar{\boldsymbol{P}})\rangle_L$. This result is proved in the case where the Hamiltonian commutes with $\hat{\boldsymbol{x}}$. The Wigner function is non-positive if the dominant paths with path centroid in a certain region have Berry phases close to odd multiples of π. For finite L the path centroid distribution is a Wigner function convolved with a Gaussian of variance inversely proportional to L. The results are illustrated by numerical calculations of the spin Wigner function from $SU(2)$ coherent states. The relevance to the quantum Monte Carlo sign problem is also discussed.

1. Introduction

It is possible, therefore, that a closer study of the relation of classical and quantum theory might involve us in negative probabilities, and so it does. R P Feynman [1]

The application of classical concepts to quantum systems comes at a cost. It is indeed possible to represent operators in many ways by functions f of commuting variables. Expectation values take a form reminiscent of classical statistical mechanics: they are averages with respect to a distribution $W(\boldsymbol{x})$, where $\boldsymbol{x}$ takes values in some (as yet unspecified) space Γ:

$$\langle\hat{f}\rangle = \mathrm{Tr}(\hat{f}\hat{\rho}) = \int_\Gamma \mathrm{d}\boldsymbol{x}\, f(\boldsymbol{x})W(\boldsymbol{x}). \tag{1}$$

Here f is a function that depends only on the operator $\hat{f}$ of interest, and W is a normalized distribution that depends only on the density matrix $\hat{\rho}$ [2]. Although *quasiprobability* distributions $W(\boldsymbol{x})$ satisfying (1) indeed exist [3, 4], it is often the case that no *positive-definite* $W(\boldsymbol{x})$ exists for a given $\hat{\rho}$. The distribution that is the principal subject of this paper is the Wigner function, best known as a joint distribution of position and momentum [3]. Another example is the distribution of local hidden variables that predetermine the outcome of measurements on components of an entangled state. Here the assumption $W(\boldsymbol{x}) \geqslant 0$ implies certain inequalities between correlations that are violated by the quantum mechanical result, demonstrating the non-existence of a positive distribution [5].

Equation (1) also represents path-integral methods, where $\boldsymbol{x}$ is a path in a coherent-state manifold or a time-dependent auxiliary field. There is no difficulty, in principle, in working with non-positive distributions, provided one is sufficiently wary when applying the axioms of probability theory. In practice, however, non-positivity can be a serious hindrance to numerical computation of the integrals. Monte Carlo techniques typically evaluate integrals of the form (1) by sampling $f(\boldsymbol{x})$ from a distribution W. If the average sign $\int_\Gamma W(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}/\int_\Gamma |W(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x}$ is small, convergence with sample size becomes intolerably slow. This notorious *sign problem*, a frequent hindrance to quantum Monte Carlo calculations, was first noted in the context of fermion simulations [6] but appears in other contexts [7]. The coherent-state path integral presents the system perhaps least amenable to Monte Carlo simulation: the weight is complex and contains a rapidly varying Berry phase. For this reason, such integrals have rarely [8, 9] been tackled by Monte Carlo techniques. A more widely used approach for Monte Carlo simulation of many-body systems replaces the interaction by a Gaussian average over an auxiliary field. The auxiliary field transports the system around a path in state space, with the state relaxing towards the instantaneous ground state in the field [7, 10]. The distribution of the auxiliary field is a thermally broadened distribution of the operators to which it is coupled [11]. Thus the results presented here for coherent states can cast light on the sign problem in the auxiliary-field Monte Carlo method.

An alternative to direct evaluation of the path integral is to integrate all non-zero frequency modes out of the path integral, thereby mapping the system to an effective *classical* system, usually determined variationally [12–14]. Paths $\boldsymbol{x}(\tau)$, $0 \leqslant \tau < \beta$ are classified according to their *path centroid* $\bar{\boldsymbol{x}} = \frac{1}{\beta}\int_0^\beta \boldsymbol{x}(\tau)\,\mathrm{d}\tau$. The path integral with action $S[\boldsymbol{x}]$ reduces to an ordinary integral over c-number variables with classical effective Hamiltonian $H_{\mathrm{eff}}(\bar{\boldsymbol{x}})$. Excursions of the path from $\bar{\boldsymbol{x}}$ provide quantum corrections to the potential in the effective Hamiltonian. The function $W(\bar{\boldsymbol{x}}) = \exp(-\beta H_{\mathrm{eff}}(\bar{\boldsymbol{x}}))$, regarded as a Boltzmann distribution of the variables $\bar{\boldsymbol{x}}$, forms the basis of a classical statistical mechanics in phase space [15, 16]. This distribution is an intuitive interpretation of the path-centroid distribution *if the observables represented by* $\boldsymbol{x}$ *are compatible*. However, this is not always the case: the distribution resulting from a phase-space path integral over canonical coordinates (q, p) would be a Wigner function. It is such distributions that form the subject of this paper.

The present author has previously shown a formal correspondence between the path-centroid spin distribution in the coherent-state path integral and the spin Wigner function [11]. This work made the assumption of continuous paths, implicit in field-theoretical treatments [17], leading to a representation of the spin by its matrix element $\boldsymbol{Q} = s\boldsymbol{n}$. (Here the coherent-state label $\boldsymbol{n}$ is a unit vector in $\mathbb{R}^3$.) However, because of the conditionally convergent nature of the path integral, the correspondence between the spin representation in the path integral and the Wigner function depends on the class of paths appearing in the path integral. While some forms of the measure require use of the matrix element [18], Brownian paths in the limit of a divergent diffusion coefficient are non-differentiable, and require the diagonal representative $\boldsymbol{P} = (s+1)\boldsymbol{n}$ [19]. The apparent ambiguity of the functional integral representations of the Hubbard model has similarly been ascribed to questions of continuity of paths [20]. Subsequent work reported briefly [21] demonstrated that, if the path integral for spin is defined on a sequence of spherical polygons, the distribution of the path centroid of the diagonal representative of the spin converges onto the spin Wigner function. The aim of this paper is to generalize this result to coherent-state representations in arbitrary finite-dimensional Hilbert spaces, and discuss the convergence with path discretization.

The next section defines the Wigner function and coherent state formalism used here. Section 3 derives the main result of this work, that the Wigner function of a set of operators can be computed as a histogram of their time-averaged diagonal representatives:

$$W(\boldsymbol{x}) \equiv \langle \delta_N(\boldsymbol{x} - \hat{\boldsymbol{x}}) \rangle = \lim_{L\to\infty} \langle \delta_N(\boldsymbol{x} - \bar{\boldsymbol{P}}) \rangle_L. \tag{2}$$

Here $\hat{\boldsymbol{x}}$ are operators commuting with the Hamiltonian, $\boldsymbol{x}$ are c-number variables and $\boldsymbol{P}$ are the diagonal representatives of the operators $\hat{\boldsymbol{x}} = \int \mathrm{d}\mu(\boldsymbol{n})\, \boldsymbol{P}(\boldsymbol{n}) |\boldsymbol{n}\rangle\langle \boldsymbol{n}|$. $\bar{\boldsymbol{P}}$ is the path centroid, the time average of the diagonal representative. The first set of angle brackets represents a thermal expectation value, and the second an average over L-vertex polygonal paths in the coherent-state path integral; δ_N is the delta function in $\mathbb{R}^N$. For finite L the distribution is broadened, typically by a Gaussian of variance inversely proportional to L. An application to a spin-s particle is presented in section 4. Finally, section 5 discusses the interpretation and wider applicability of these results. Non-positive regions of the Wigner function tend to emerge if the dominant path for a given value of the path centroid has a Berry phase that is an odd multiple of π.

2. Definitions

2.1. The Wigner function

The state of a system is specified by its density matrix $\hat{\rho} = Z^{-1} \exp(-\beta \hat{H})$, a positive Hermitian operator with trace $\mathrm{Tr}\,\hat{\rho} = 1$. This is written in the form of a canonical density matrix of a system in thermal equilibrium at an inverse temperature β; the foregoing equation can be taken as defining a Hamiltonian (up to an additive constant) for any non-singular $\hat{\rho}$. We map the system onto the statistical mechanics of c-number variables $\boldsymbol{x} \equiv (x_1, \ldots, x_N) \in \mathbb{R}^N$ corresponding to N linearly independent operators $\hat{\boldsymbol{x}} \equiv (\hat{x}_1, \ldots, \hat{x}_N)$. The quasiprobability distribution $W(\boldsymbol{x})$ is the *Wigner function* [3] of the corresponding operators. This function could be considered as a Boltzmann distribution with a (not necessarily real or bounded) classical effective Hamiltonian $H_{\mathrm{eff}}(\boldsymbol{x})$. We require $W(\boldsymbol{x})$ to be a linear function of $\hat{\rho}$, defined as

$$W(\boldsymbol{x}) = \mathrm{Tr}\left(\hat{\rho}\, \delta_N(\boldsymbol{x} - \hat{\boldsymbol{x}})\right) \tag{3}$$

where the N-dimensional delta function δ_N is defined with symmetrical operator ordering [22]

$$\delta_N(\boldsymbol{x} - \hat{\boldsymbol{x}}) \equiv \int \frac{\mathrm{d}^N \boldsymbol{\lambda}}{(2\pi)^N} \exp\left(\mathrm{i}\boldsymbol{\lambda} \cdot (\boldsymbol{x} - \hat{\boldsymbol{x}})\right). \tag{4}$$

The Wigner function is then the Fourier transform of the *characteristic function*

$$\chi(\boldsymbol{\lambda}) = \mathrm{Tr}\left(\hat{\rho}\, \mathrm{e}^{-\mathrm{i}\boldsymbol{\lambda}\cdot\hat{\boldsymbol{x}}}\right) \tag{5}$$

$$W(\boldsymbol{x}) = \int \frac{\mathrm{d}^N \boldsymbol{\lambda}}{(2\pi)^N} \mathrm{e}^{\mathrm{i}\boldsymbol{\lambda}\cdot\boldsymbol{x}} \chi(\boldsymbol{\lambda}). \tag{6}$$

It can easily be shown that the Wigner function (6) reduces to the correct positive marginal distribution in an M-dimensional commuting subspace of operators on integration over the remaining $N - M$ variables. The Wigner function can therefore be used to find expectation values of linear combinations of arbitrary functions of commuting operators. The correlation between two spins, $\langle \hat{\boldsymbol{S}}_1 \cdot \hat{\boldsymbol{S}}_2 \rangle = \sum_{i=x,y,z} \langle \hat{S}_1^{(i)} \hat{S}_2^{(i)} \rangle$, is of this form, as are the correlations $\langle \boldsymbol{a} \cdot \hat{\boldsymbol{S}}_1 \boldsymbol{b} \cdot \hat{\boldsymbol{S}}_2 \rangle$ measured in the Einstein–Podolsky–Rosen experiment [5, 23].

If $\hat{\boldsymbol{x}} = (\hat{q}, \hat{p})$ are the canonical position and momentum operators, equation (6) reduces to the well known form of the Wigner function. If $\hat{\boldsymbol{x}} = \hat{\boldsymbol{S}}$ is the spin operator for a spin-s

particle, we obtain a rather singular form consisting of derivatives of δ-functions supported on spheres of quantized radius. In zero field this is [11, 22, 24]

$$W_s(\boldsymbol{S}) = \begin{cases} \dfrac{-1}{2s+1} \displaystyle\sum_{m=1/2}^{s} \dfrac{1}{2\pi S}\delta'(S-m) & s \text{ half-odd integer} \\ \dfrac{\delta_3(\boldsymbol{S})}{2s+1} - \dfrac{1}{2s+1} \displaystyle\sum_{m=1}^{s} \dfrac{1}{2\pi S}\delta'(S-m) & s \text{ integer.} \end{cases} \tag{7}$$

(Here $S = |\boldsymbol{S}|$ is the magnitude of the classical spin vector, and s is the spin quantum number.)

The representation of the spin distribution as a function of a vector $\boldsymbol{S} \in \mathbb{R}^3$ demands comment. This is appropriate in the context of the statistical mechanics of composite spins, where the state is not restricted to a single spin-s representation. The total spin has a distribution in the same space, which allows for dispersion in the magnitude, and the formalism is explicitly isotropic. A more natural choice for a fixed spin s is the distribution of directions of a vector of magnitude $\sqrt{s(s+1)}$; the space is a sphere S^2, and a family of distributions on the sphere can be derived from the coherent-state representation of the spin [22, 23]. We shall relate the distributions by embedding the sphere in $\mathbb{R}^3$: the correspondence is between a set of points on S^2 and their vector average.

2.2. *Coherent states*

Before relating these Wigner functions to coherent-state path integrals, we review the properties of generalized coherent states relevant to this paper [25–27].

Coherent states $|\boldsymbol{n}\rangle$, labelled by a variable taking values in some manifold, $\boldsymbol{n} \in \mathcal{M}$, provide a continuous, overcomplete and non-orthogonal basis for a finite-dimensional Hilbert space $\mathcal{H}$. The completeness is expressed by the resolution of the identity

$$\hat{1} = \int_{\mathcal{M}} \mathrm{d}\mu(\boldsymbol{n})\, |\boldsymbol{n}\rangle\langle\boldsymbol{n}| \tag{8}$$

where $\mathrm{d}\mu(\boldsymbol{n})$ is a measure on the manifold. There exists a family of representations of operators by functions on the manifold, of which the matrix element and diagonal representative are of relevance here [3, 28]. The *matrix element* Q (also known as the antinormal or upper symbol) of an operator $\hat{A}$ is

$$Q(\boldsymbol{n}) = \langle\boldsymbol{n}|\hat{A}|\boldsymbol{n}\rangle \tag{9}$$

and the *diagonal representative* P (also known as the normal or lower symbol) obeys

$$\hat{A} = \int_{\mathcal{M}} \mathrm{d}\mu(\boldsymbol{n})\, P(\boldsymbol{n})|\boldsymbol{n}\rangle\langle\boldsymbol{n}|. \tag{10}$$

This does not uniquely define the diagonal representative; we only require that it exist, be bounded and be a linear function of the operators. This holds in common cases for suitable choices of the fiducial vector [29], and can be constructed for $SU(2)$ coherent states [30, 31].

For use in section 4, we recall the form of $SU(2)$ coherent states for spin-s as defined by Radcliffe [32]. Hilbert space $\mathcal{H} = \mathbb{C}^{2s+1}$ is $(2s+1)$-dimensional and the coherent-state manifold is the Bloch sphere, $\mathcal{M} = S^2$, with $\boldsymbol{n} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$ a unit vector. The states are obtained by rotating the highest-weight eigenstate $|s\rangle$ about an axis in the xy-plane:

$$|\boldsymbol{n}\rangle = \cos^{2s}\frac{\theta}{2}\exp\left[\tan\frac{\theta}{2}\mathrm{e}^{\mathrm{i}\phi}\hat{S}_-\right]|s\rangle. \tag{11}$$

This has the useful property that the spin is 'pointing' in the direction $\boldsymbol{n}$, i.e. $\boldsymbol{n}\cdot\hat{\boldsymbol{S}}|\boldsymbol{n}\rangle = s|\boldsymbol{n}\rangle$. Such states are not orthogonal:

$$\langle \boldsymbol{n}_1|\boldsymbol{n}_2\rangle = \left(\tfrac{1}{2}(1+\boldsymbol{n}_1\cdot\boldsymbol{n}_2)\right)^s \mathrm{e}^{\mathrm{i}s\Omega} \tag{12}$$

where (with the present gauge) Ω is the area of the spherical triangle formed by the z-axis, $\boldsymbol{n}_1$ and $\boldsymbol{n}_2$. The measure is the element of area $\mathrm{d}\mu(\boldsymbol{n}) = \frac{2s+1}{4\pi}\,\mathrm{d}\cos\theta\,\mathrm{d}\phi$. The matrix element and diagonal representatives of the spin operator are $\boldsymbol{Q} = s\boldsymbol{n}$ and $\boldsymbol{P} = (s+1)\boldsymbol{n}$, respectively [33].

3. Path-integral calculation of Wigner function

3.1. Path-centroid distributions

We now derive the main result, relating the Wigner function of an operator to the distribution of its diagonal representative, in the case of conserved operators $\hat{\boldsymbol{x}}$, i.e.

$$[\hat{H}, \hat{x}_\mu] = 0 \qquad \mu = 1,\ldots,N. \tag{13}$$

Thus we can compute the joint distribution of the symmetry generators of the Hamiltonian, such as the total spin of an isotropic ferromagnet. We also require $\hat{H}$ and $\hat{\boldsymbol{x}}$ to be bounded operators, and their diagonal representatives to be bounded functions.

We apply the Suzuki–Trotter decomposition to the characteristic function (5),

$$\chi(\boldsymbol{\lambda}) = \lim_{L\to\infty} Z^{-1}\,\mathrm{Tr}\left(1 - \frac{\beta\hat{H} + \mathrm{i}\boldsymbol{\lambda}\cdot\hat{\boldsymbol{x}}}{L}\right)^L. \tag{14}$$

Let P_H be the diagonal representative of the Hamiltonian and $P_{\mu\nu\cdots}$ be the diagonal representative of the operator product

$$\hat{x}_\mu\hat{x}_\nu\cdots = \int_{\mathcal{M}} \mathrm{d}\mu(\boldsymbol{n})\, P_{\mu\nu\cdots}(\boldsymbol{n})|\boldsymbol{n}\rangle\langle\boldsymbol{n}| \tag{15}$$

so that $\boldsymbol{P}(\boldsymbol{n}) = (P_1(\boldsymbol{n}),\ldots,P_N(\boldsymbol{n}))$ is the diagonal representative of $\hat{\boldsymbol{x}}$. The matrix element clearly exists and is unique for a bounded operator. Replacing the operators in each of the L factors in (14) by their diagonal representatives gives

$$\chi(\boldsymbol{\lambda}) = \lim_{L\to\infty} Z^{-1}\,\mathrm{Tr}\int_{\mathcal{M}^L} \mathrm{d}\mu_L[\boldsymbol{n}] \prod_{l=1}^{L}\left[\left(1 - \frac{\beta P_H(\boldsymbol{n}_l) + \mathrm{i}\boldsymbol{\lambda}\cdot\boldsymbol{P}(\boldsymbol{n}_l)}{L}\right)|\boldsymbol{n}_l\rangle\langle\boldsymbol{n}_l|\right] \tag{16}$$

$$= \lim_{L\to\infty} Z^{-1}\int_{\mathcal{M}^L} \mathrm{d}\mu_L[\boldsymbol{n}]\,\mathrm{e}^{-\mathcal{S}_L[\boldsymbol{n}]} \prod_{l=1}^{L}\left(1 - \frac{\beta P_H(\boldsymbol{n}_l) + \mathrm{i}\boldsymbol{\lambda}\cdot\boldsymbol{P}(\boldsymbol{n}_l)}{L}\right). \tag{17}$$

Here $\mathcal{M}^L$ comprises ordered sets of L points $\boldsymbol{n}_l \in \mathcal{M}$, with measure $\mathrm{d}\mu_L[\boldsymbol{n}] = \prod_{l=1}^L \mathrm{d}\mu(\boldsymbol{n}_l)$. The gauge-invariant action $\mathcal{S}_L[\boldsymbol{n}]$ is defined by

$$\exp(-\mathcal{S}_L[\boldsymbol{n}]) = \langle\boldsymbol{n}_1|\boldsymbol{n}_2\rangle\langle\boldsymbol{n}_2|\cdots|\boldsymbol{n}_L\rangle\langle\boldsymbol{n}_L|\boldsymbol{n}_1\rangle. \tag{18}$$

We can consider these paths as a truncation of the function space $\{\boldsymbol{n} : [0,\beta) \to \mathcal{M}\}$ to piecewise-constant paths

$$\boldsymbol{n}(\tau) = \boldsymbol{n}_l \qquad (l-1)\beta/L \leqslant \tau < l\beta/L. \tag{19}$$

If we consider successive points as linked by a geodesic, with $\boldsymbol{n}_L$ linked to $\boldsymbol{n}_1$, these paths represent geodesic polygons. The amplitude of this factor (18) decreases with increasing path length, and its (Berry) phase is related to the enclosed area [34].

Equation (17) can be re-exponentiated to give

$$\lim_{L\to\infty} \chi_L(\boldsymbol{\lambda}) = \chi(\boldsymbol{\lambda}) \tag{20}$$

where

$$\begin{aligned}\chi_L(\boldsymbol{\lambda}) &= \left\langle \exp\left(-\mathrm{i}L^{-1}\sum_{l=1}^{L} \boldsymbol{\lambda}\cdot\boldsymbol{P}(\boldsymbol{n}_l)\right)\right\rangle_L \\ &= \left\langle \exp(-\mathrm{i}\boldsymbol{\lambda}\cdot\overline{\boldsymbol{P}}[\boldsymbol{n}])\right\rangle_L .\end{aligned} \tag{21}$$

Here the functional average is denoted by

$$\langle f\rangle_L \equiv Z^{-1}\int_{\mathcal{M}^L} \mathrm{d}\mu_L[\boldsymbol{n}]\, f[\boldsymbol{n}] \exp\left(-\mathcal{S}_L[\boldsymbol{n}] - \beta\bar{P}_H[\boldsymbol{n}]\right) \tag{22}$$

and the *path centroid* is

$$\overline{\boldsymbol{P}}[\boldsymbol{n}] = \frac{1}{L}\sum_{l=1}^{L} \boldsymbol{P}(\boldsymbol{n}_l). \tag{23}$$

The path-centroid distribution (PCD) $W_L(\boldsymbol{x})$, the approximant to the Wigner function obtained from polygonal paths with L vertices, is the Fourier transform of χ_L,

$$W_L(\boldsymbol{x}) = \left\langle \delta_N(\boldsymbol{x} - \overline{\boldsymbol{P}}[\boldsymbol{n}])\right\rangle_L . \tag{24}$$

From (20) we obtain our main result,

$$\lim_{L\to\infty} W_L(\boldsymbol{x}) = W(\boldsymbol{x}). \tag{25}$$

This limit exists in the sense of distributions; averages of polynomials in $\boldsymbol{x}$ with weight W_L converge with L. Since the Wigner function may be singular, as in equation (7), convergence of the PCD is not necessarily pointwise.

3.2. Finite-L corrections

By cumulant expansion of the characteristic function (14) we obtain finite-L corrections to the Wigner function (24) for the case of a vanishing Hamiltonian. Inclusion of the leading large-L correction in (14) gives

$$\chi(\boldsymbol{\lambda}) = Z^{-1}\,\mathrm{Tr}\left(1 - \frac{1}{L}\mathrm{i}\lambda_\mu\hat{x}_\mu - \frac{1}{2L^2}\lambda_\mu\lambda_\nu\hat{x}_\mu\hat{x}_\nu + \mathrm{O}(L^{-3})\right)^L . \tag{26}$$

(Summation over repeated indices is assumed.) Replacing the factors by their diagonal representations, as in section 3.1, and performing a cumulant expansion gives

$$\chi(\boldsymbol{\lambda}) = \left\langle \exp\left(\frac{-\mathrm{i}\lambda_\mu}{L}\sum_{l=1}^{L} P_\mu(\boldsymbol{n}_l) - \frac{\lambda_\mu\lambda_\nu}{2L^2}\sum_{l=1}^{L}(P_{\mu\nu}(\boldsymbol{n}_l) - P_\mu(\boldsymbol{n}_l)P_\nu(\boldsymbol{n}_l)) + \mathrm{O}(L^{-2})\right)\right\rangle_L . \tag{27}$$

Defining the average

$$C_{\mu\nu} = \left\langle \frac{1}{L}\sum_{l=1}^{L}\left(P_{\mu\nu}(\boldsymbol{n}_l) - P_\mu(\boldsymbol{n}_l)P_\nu(\boldsymbol{n}_l)\right)\right\rangle_L \tag{28}$$

and comparing (27) with (21) gives

$$\chi_L(\boldsymbol{\lambda}) = \chi(\boldsymbol{\lambda})\exp\left(\frac{1}{2L}\lambda_\mu\lambda_\nu C_{\mu\nu} + \mathrm{O}(L^{-2})\right) \tag{29}$$

provided that the quantity in angle brackets in (28) is uncorrelated to the linear term. In that case, if C is negative-definite, we can compute the PCD as the exact Wigner function convolved (to leading order) with a Gaussian of variance inversely proportional to L:

$$W_L(\boldsymbol{x}) \approx (\det(-2\pi C/L))^{-1/2} \int \mathrm{d}^N \boldsymbol{y}\; W(\boldsymbol{x}-\boldsymbol{y}) \exp\left(LC^{-1}_{\mu\nu} y_\mu y_\nu/2\right). \qquad (30)$$

It is worth noting that the distribution of auxiliary fields coupled to these operators $\hat{\boldsymbol{x}}$ is similarly the Wigner function convolved with a Gaussian of variance proportional to temperature [11].

4. Illustrative example: spin s

We conclude this section with the example of a spin-s particle with vanishing Hamiltonian, both to illustrate the theory and to recover and extend previous results [11, 21].

The path integration is now over spherical-polygon paths on the unit (Bloch) sphere. Following (12), the action of such spherical-polygon paths is

$$\mathcal{S}_L = -s \sum_{l=1}^{L} \ln\left(\tfrac{1}{2}(1 + \boldsymbol{n}_{l-1} \cdot \boldsymbol{n}_l)\right) - \mathrm{i}s\Omega \qquad (31)$$

where Ω is the solid angle enclosed and $\boldsymbol{n}_0 \equiv \boldsymbol{n}_L$. The imaginary part $s\Omega$ is the Berry phase. The resulting PCD for a free spin s is

$$W_L(\boldsymbol{S}) = \langle \delta_3(\boldsymbol{S} - (s+1)\bar{\boldsymbol{n}})\rangle_L\,. \qquad (32)$$

From the previous argument, this is the spin Wigner function (7), broadened by a Gaussian for large L. To find the variance of this Gaussian, we insert into (28) Lieb's expressions [33] for the diagonal representatives of $\hat{S}_z$ and $\hat{S}_z^2$, which are $P_z = (s+1)\cos\theta$ and $P_{zz} = (s+1)\left(s+\frac{3}{2}\right)\cos^2\theta - (s+1)/2$, respectively:

$$C_{\mu\nu} = -\tfrac{1}{3}(s+1)\delta_{\mu\nu}. \qquad (33)$$

The PCD (30) is therefore the exact Wigner function (7) broadened by a Gaussian of variance $(s+1)/(3L)$:

$$W_L(\boldsymbol{S}) \approx \frac{1}{2s+1}\left(\frac{3L}{2\pi(s+1)}\right)^{3/2} \sum_{m=-s}^{s} \left(1 - \frac{m}{S}\right) \exp\left(-\frac{3L(S-m)^2}{2(s+1)}\right). \qquad (34)$$

Equation (32) implies the vanishing of the exact PCD for $S > s+1$ (i.e. $|\bar{\boldsymbol{n}}| > 1$). The approximation (34) has Gaussian tails in this region, vanishing in the large-L limit for any fixed $S > s$.

Expression (32) suggests a means of calculation of the PCD by accumulation of a histogram of values of $\bar{\boldsymbol{n}}$. While closed forms for the PCD are obtainable for small L and an asymptotic form (34) for large L, it is nevertheless informative to compute the PCD by Monte Carlo integration. For each Monte Carlo step, an ordered set of L independent unit vectors $\boldsymbol{n}_l$ is drawn from a uniform distribution on the sphere and their vector average $\bar{\boldsymbol{n}}$ is taken. The real part of the weight $\exp(-\mathcal{S}_L)$ is added to the corresponding bin in the histogram. As the time-reversed path appears with equal probability, the imaginary part is discarded. There is substantial cancellation of the real part of the weights (the sign problem), which hinders convergence.

Since the PCD is spherically symmetric, we only require the radial distribution

$$w_L(S) = 4\pi S^2 W_L(\boldsymbol{S}). \qquad (35)$$

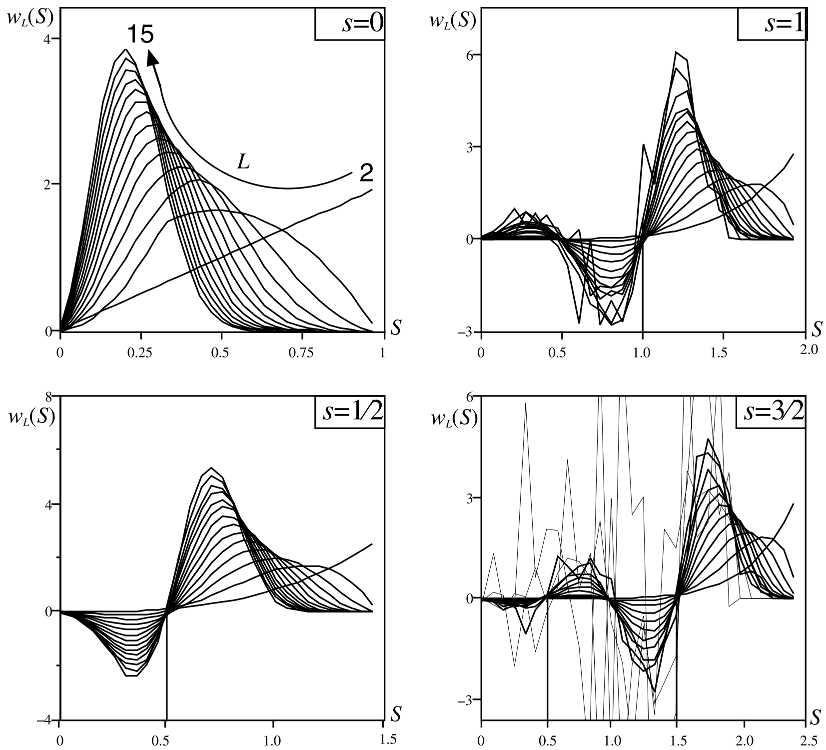

Figure 1. Radial path-centroid distribution (35) for spherical polygons with $L = 2, \ldots, 15$ vertices, computed for spins $s = 0, \ldots, \frac{3}{2}$ with 10^7 Monte Carlo steps. The vertical bars show the expected positions of the δ' functions as $L \to \infty$. To aid the eye, plots are shown as thin curves where convergence is poor.

Because the coordinate transformation is nonlinear, this distribution is not the Wigner function of the operator $\sqrt{\hat{S} \cdot \hat{S}}$; in particular, it is not a positive distribution. Figure 1 shows the numerically accumulated histograms for the radial distribution and their convergence with L to the Wigner function. The plots are shown for a relatively small number 10^7 of Monte Carlo steps to highlight the influence of the sign problem on convergence.

Knowledge of the phase distribution, and of the correlation between the phase and the path centroid, is of importance both to determine the feasibility of convergence and to develop methods of accelerating convergence [35]. The paths contributing to $|\bar{n}| \approx 1$ enclose a small area, so that the phase distribution is peaked near zero and the spin distribution is positive. The action is real for $L = 1$ (the static approximation to the path integral), and for $L = 2$, where the path encloses no area. Non-zero phases appear from $L = 3$, where the spherical polygons enclose non-zero area. This is the smallest value that retains information about spin quantization, and distinguishes the spectra of ferromagnetic and antiferromagnetic spin systems [36]. In this case the weights are complex, but there is still correlation between the time-averaged spin and the phase, as figure 2 shows. For $L \gg 3$ the Berry phases (modulo 2π) are nearly uniformly distributed between 0 and 2π [37], and weakly correlated to the path centroid, so that convergence is poor. Statistics are poor for the Wigner function for large path averages ($|\bar{n}| \gg 1/\sqrt{L}$) due to uniform sampling, and for small path averages ($|\bar{n}| \ll \frac{s}{s+1}$) due to destructive interference between paths of positive and negative weights. These ranges

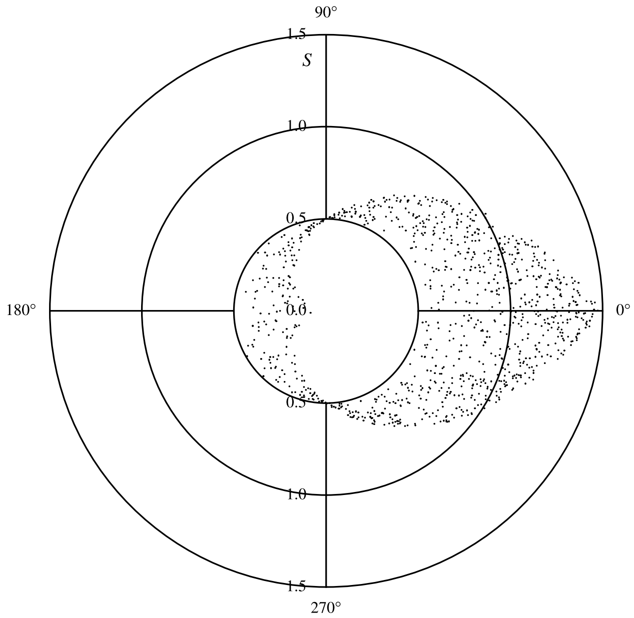

Figure 2. Polar plot of magnitude of the path-centroid spin $S = 3|\bar{\boldsymbol{n}}|/2$ against the Berry phase for spin $\frac{1}{2}$ for a random sample of 1000 spherical triangles. The phase is less (greater) than $\pi/2$ in magnitude for $S > (<)\frac{1}{2}$.

overlap for large s and L, giving poor statistics for the full interval. The histograms would converge to the Wigner function if the limit of infinitely many Monte Carlo steps were taken *before* the limit $L \to \infty$.

5. Discussion

The main result of this investigation is proved in section 3.1: the Wigner function of a set of operators is obtained in terms of the path centroid distribution of the diagonal representatives of the operators. The Wigner function is obtained in the limit of geodesic polygonal paths as the number of vertices $L \to \infty$. Quantization is apparent when $L \geqslant 3$ and becomes exact as $L \to \infty$. The distribution corresponding to commuting operators, measured at a single time, is a non-negative effective Boltzmann distribution in configuration space [12–15]. Interference between paths occurs as the result of the superposition of amplitudes, but the probabilities are positive. If the operators do not commute, the resulting distribution need not be positive. Should one interpret the exponentiated action of a path not just as the weight by which functionals of the path are to be averaged, but as a (complex) quasiprobability that a spin take a particular closed path, one can see the quasiprobabilities themselves interfering destructively. The question of whether incompatible observables each take given values is impermissible in quantum mechanics; we should not expect the answer to be drawn from a positive distribution. However, the question of whether the point $\boldsymbol{x}$ lies in a subset $\gamma \subset \Gamma$ is well defined, provided that γ is sufficiently large that the coarse-grained Wigner function $\int_\gamma W(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$ is positive. Thus the sum of the weights of all paths with centroid in γ will be positive.

A heuristic argument helps one to understand the oscillations in sign of the Wigner function in terms of the Berry phases of paths. Because long paths are suppressed by an overlap factor (18), the dominant paths $\{\boldsymbol{P}(\boldsymbol{n}_l), l = 1, \ldots, L\}$ contributing to the path centroid $\bar{\boldsymbol{P}} = \sum_l \boldsymbol{P}(\boldsymbol{n}_l)/L$ are the shortest ones exploring a region of low energy. The Wigner function will be negative where these paths have a Berry phase near an odd multiple of π. Of course the entropy gain in extending the paths leads to a broad distribution of paths [37], but the minimum may still be visible as a caustic. There are two ways in which this may arise. Firstly, if the coherent state manifold is curved the path centroid does not in general lie in the manifold; this implies a lower bound on the length of the path. Such is the case for $SU(2)$ coherent states discussed in section 4. The shortest paths on the unit sphere with centroid $|\bar{\boldsymbol{n}}| = \cos\theta$ are small circles enclosing solid angle $2\pi(1-\cos\theta)$; a spin s performing such a conical path will acquire a Berry phase $2\pi s(1-\cos\theta)$. The PCD should therefore be positive for $\cos\theta = 1$ and cross zero whenever the phase is an odd multiple of $\pi/2$. Such an argument predicts zeros with positive gradient at

$$\cos\theta = \frac{4s-1}{4s}, \frac{4s-5}{4s}, \ldots, \frac{3}{4s}\left[\frac{1}{4s}\right] \tag{36}$$

to be compared with the exact result (7)

$$\cos\theta = \frac{s}{s+1}, \frac{s-1}{s+1}, \ldots, \frac{1}{s+1}\left[\frac{1}{2(s+1)}\right] \tag{37}$$

for integer (half-integer) spins, respectively. This (admittedly crude) argument gives the correct spacing of the zeros to $\mathrm{O}(1/s)$. It is also possible for the dominant paths to have non-zero phase in a flat coherent state manifold if the Hamiltonian has a local maximum near the position of the path centroid. This is left for a future investigation.

These results provide insight into the origin of the Monte Carlo sign problem. If the variables sampled in a quantum Monte Carlo simulation correspond to non-commuting operators, the joint distribution of their time average will be a Wigner function, indicating the presence of a non-positive integrand. Such an example might be the total spin of an interacting system or, more generally, a vector order parameter generating a non-Abelian symmetry group. Much more frequently, the Hubbard–Stratonovich transformation replaces the interaction by an auxiliary field, with Monte Carlo integration over time-dependent configurations of this field. The distribution of the time-averaged auxiliary field is a Gaussian convolution of the distribution of the operators to which it is coupled, with variance proportional to temperature [11]. Thus the same considerations apply to the auxiliary-field Monte Carlo method at low temperatures; however, as the path space differs from the geodesic polygons discussed here, the results will be quantitatively different.

References

[1] Feynman R P 1987 *Quantum Implications* ed B J Hiley and F D Peat (London: Routledge) p 235
[2] Stratonovich R L 1956 *Zh. Eksp. Teor. Fiz.* **31** 1012 (Engl. transl. 1957 *Sov. Phys.–JETP* **4** 891)
[3] Hillery M, O'Connell R F, Scully M O and Wigner E P 1984 *Phys. Rep.* **106** 121
[4] Mückenheim W *et al* 1986 *Phys. Rep.* **133** 337
[5] Bell J S 1964 *Physics* **1** 195
[6] Blankenbecler R, Scalapino D J and Sugar R L 1981 *Phys. Rev.* D **24** 2278
[7] Samson J H 1995 *Int. J. Mod. Phys.* C **6** 427
[8] Takano H 1987 *Quantum Monte Carlo Methods in Equilibrium and Non-equilibrium Systems* ed M Suzuki (Berlin: Springer) p 144
[9] Vieira V R and Sacramento P D 1994 *Physica* A **207** 584
[10] Fahy S and Hamann D R 1991 *Phys. Rev.* B **43** 765

[11] Samson J H 1995 *Phys. Rev.* B **51** 223
[12] Feynman R P and Hibbs A R 1965 *Quantum Mechanics and Path Integrals* (New York: McGraw-Hill) p 280
[13] Giachetti R and Tognetti V 1985 *Phys. Rev. Lett.* **55** 912
[14] Feynman R P and Kleinert H 1986 *Phys. Rev.* A **34** 5080
[15] Cao J and Voth G A 1994 *J. Chem. Phys.* **100** 5093
Cao J and Voth G A 1994 *J. Chem. Phys.* **100** 5106
[16] Cuccoli A *et al* 1995 *J. Phys.: Condens. Matter* **7** 7891
[17] Fradkin E and Stone M 1988 *Phys. Rev.* B **38** 7215
[18] Klauder J R 1982 *J. Math. Phys.* **23** 1797
[19] Klauder J R and Daubechies I 1984 *Phys. Rev. Lett.* **52** 1161
Daubechies I and Klauder J R 1985 *J. Math. Phys.* **26** 2239
[20] Prange R E 1981 *Electron Correlation and Magnetism in Narrow-band Systems* ed T Moriya (Berlin: Springer) p 69
[21] Samson J H 1999 *Proc. 6th Int. Conf. on Path Integrals from peV to TeV (Firenze)* ed R Casalbuoni *et al* (Singapore: World Scientific) p 550
[22] Scully M O and Wódkiewicz K 1994 *Found. Phys.* **24** 85
[23] Agarwal G S, Home D and Schleich W 1992 *Phys. Lett.* A **170** 359
Agarwal G S 1993 *Phys. Rev.* A **47** 4608
[24] Chandler C *et al* 1992 *Found. Phys.* **22** 867
[25] Klauder J R and Skagerstam B S 1994 *Coherent States: Applications in Physics and Mathematical Physics* (Singapore: World Scientific)
[26] Perelomov A 1986 *Generalized Coherent States and their Applications* (Berlin: Springer) p 59
[27] Zhang W M, Feng D H and Gilmore R 1990 *Rev. Mod. Phys.* **62** 867
[28] Brif C and Mann A 1999 *Phys. Rev.* A **59** 971
[29] Mukunda N *et al* 2000 *Preprint* quant-ph/0002070
[30] Vieira V R and Sacramento P D 1994 *J. Phys. A: Math. Gen.* **27** L783
[31] Gilmore R 1976 *J. Phys. A: Math. Gen.* **9** L65
[32] Radcliffe J M 1971 *J. Phys. A: Math. Gen.* **4** 313
[33] Lieb E H 1973 *Commun. Math. Phys.* **31** 327
[34] Mukunda N and Simon R 1993 *Ann. Phys., NY* **228** 205
[35] Deutsch J M 1994 *Phys. Rev.* E **50** R2411
[36] Samson J H 2000 *J. Phys. A: Math. Gen.* **33** 3111
[37] Sinha S and Samuel J 1994 *Phys. Rev.* B **50** 13 871