Husimi transform of an operator product

2000 J. Phys. A: Math. Gen. 33 3903

(http://iopscience.iop.org/0305-4470/33/21/304)

J. Phys. A: Math. Gen. **33** (2000) 3903–3915. Printed in the UK PII: S0305-4470(00)11541-0

Husimi transform of an operator product

D M Appleby

Department of Physics, Queen Mary and Westfield College, Mile End Road, London E1 4NS, UK

E-mail: D.M.Appleby@qmw.ac.uk

Received 31 January 2000

Abstract. It is shown that the series derived by Mizrahi, giving the Husimi transform (or covariant symbol) of an operator product, is absolutely convergent for a large class of operators. In particular, the generalized Liouville equation, describing the time evolution of the Husimi function, is absolutely convergent for a large class of Hamiltonians. In contrast, the series derived by Groenewold, giving the Weyl transform of an operator product, is often only asymptotic, or even undefined. The result is used to derive an alternative way of expressing expectation values in terms of the Husimi function. The advantage of this formula is that it applies in many of the cases where the anti-Husimi transform (or contravariant symbol) is so highly singular that it fails to exist as a tempered distribution.

1. Introduction

A particularly useful and illuminating way of studying the classical limit is to formulate quantum mechanics in terms of phase space distributions [1–3]. The advantage of such a formulation as compared with the standard Hilbert space formulation is that it puts quantum mechanics into a form which is similar to the probabilistic phase space formulation of classical mechanics. At least from a formal, mathematical point of view it thus allows one to regard quantum mechanics as a kind of generalized version of classical mechanics.

There are, of course, many different phase space formulations of quantum mechanics. The one which was discovered first is the formulation based on the Wigner function [1–4]. In the case of a system having one degree of freedom with position $\hat{x}$, momentum $\hat{p}$ and density matrix $\hat{\rho}$ the Wigner function is defined by

$$W(x, p) = \frac{1}{2\pi} \int \mathrm{d}y\, \mathrm{e}^{\mathrm{i}py} \langle x - \tfrac{1}{2}y | \, \hat{\rho} \, | x + \tfrac{1}{2}y \rangle$$

(in units chosen such that $\hbar = 1$). The Wigner function continues to find many important applications (in quantum tomography [3], for example). However, if the aim is specifically to represent quantum mechanics in a manner which resembles classical mechanics as closely as possible, then the Wigner function suffers from the serious disadvantage that it is not strictly non-negative (except in special cases [5]), which makes the analogy with the classical phase space probability distribution somewhat strained.

There has accordingly been some interest in the problem of constructing alternative distributions, which are strictly non-negative, and which can be interpreted as probability density functions. There are, in fact, infinitely many such functions [6–11]. The one which was discovered first, and which is the focus of this paper, is the Husimi, or Q-function [1–3, 12–15], which is obtained from the Wigner function by smearing it with a Gaussian convolution:

$$Q(x, p) = \frac{1}{\pi} \int \mathrm{d}x' \, \mathrm{d}p' \exp\left[-(x - x')^2 - (p - p')^2\right] W(x', p')$$

(in units such that $\hbar = 1$, and where we assume that x and p have been made dimensionless by choosing a suitable length scale λ, and making the replacements $x \to x/\lambda$, $p \to \lambda p$). It should be emphasized that it is not simply that the Husimi function has the mathematical significance of a probability density function. It also has this significance physically. It has been shown that the Husimi function is the probability distribution describing the outcome of a joint measurement of position and momentum in a number of particular cases [3, 8, 16–18]. More generally it can be shown [19, 20] that the Husimi function has a universal significance: namely, it is the probability density function describing the outcome of *any* retrodictively optimal joint measurement process. In Appleby [20] it is argued that this means that the Husimi function may be regarded as the canonical quantum mechanical phase space probability distribution, which plays the same role in relation to joint measurements of x and p as does the function $|\langle x \mid \psi\rangle|^2$ in relation to single measurements of x only.

If one wants to construct a systematic procedure for investigating the transition from quantal to classical it is not enough simply to find an analogue for the classical phase space probability distribution. One also needs an analogue of the classical Liouville equation, giving the time evolution of the probability distribution. In the formulation based on the Husimi function this is accomplished by means of Mizrahi's formula [15], giving the Husimi transform of an operator product (also see Lee [2], Cohen [21], Prugovečki [22] and O'Connell and Wigner [23]).

Let A_{W} denote the Weyl transform of the operator $\hat{A}$, defined by [1, 2, 24]

$$A_{\mathrm{W}}(x, p) = \int \mathrm{d}y \, \mathrm{e}^{\mathrm{i}py} \left\langle x - \tfrac{1}{2}y \right| \hat{A} \left| x + \tfrac{1}{2}y \right\rangle.$$

The Husimi transform (or covariant symbol) A_{H} is then given by

$$A_{\mathrm{H}}(x, p) = \frac{1}{\pi} \int \mathrm{d}x' \, \mathrm{d}p' \exp\left[-(x - x')^2 - (p - p')^2\right] A_{\mathrm{W}}(x', p'). \tag{1}$$

Mizrahi [15] has derived the following formula for the Husimi transform of the product of two operators $\hat{A}$, $\hat{B}$:

$$(\hat{A}\hat{B})_{\mathrm{H}} = A_{\mathrm{H}} \mathrm{e}^{\overleftarrow{\partial}_{+}\overrightarrow{\partial}_{-}} B_{\mathrm{H}} \tag{2}$$

where $\partial_{\pm} = 2^{-1/2}(\partial_x \mp \mathrm{i}\partial_p)$. Using this formula, and the fact that the Husimi function is just the Husimi transform of the density matrix scaled by a factor of $1/(2\pi)$, it is straightforward to derive the following generalization of the Liouville equation:

$$\frac{\partial}{\partial t} Q = \left\{H_{\mathrm{H}}, Q\right\}_{\mathrm{H}} \tag{3}$$

where H_{H} is the Husimi transform of the Hamiltonian, and $\left\{H_{\mathrm{H}}, Q\right\}_{\mathrm{H}}$ is the generalized Poisson bracket

$$\left\{H_{\mathrm{H}}, Q\right\}_{\mathrm{H}} = \sum_{n=0}^{\infty} \frac{2}{n!} \operatorname{Im}\left(\partial_{+}^{n} H_{\mathrm{H}} \, \partial_{-}^{n} Q\right). \tag{4}$$

The first term in the sum on the right-hand side is just the ordinary Poisson bracket. The remaining terms represent quantum mechanical corrections.

It is not apparent from Mizrahi's derivation, whether these expressions are exact, or whether they are only asymptotic. In section 2 we will show that there is a large class of operators for which the series in equation (2) (and consequently the series in equation (4)) is absolutely convergent. This property is closely connected with the complex analytic properties of the Husimi transform, as discussed by Mehta and Sudarshan [25] and Appleby [26].

The significance of the result proved in section 2 is best appreciated if one compares equations (2)–(4) with the corresponding formulae in the Wigner–Weyl formalism [1, 2, 27–29]:

$$(\hat{A}\hat{B})_{\mathrm{W}} = A_{\mathrm{W}} \exp\left[\tfrac{1}{2}\mathrm{i}\left(\overleftarrow{\partial_x}\,\overrightarrow{\partial_p} - \overleftarrow{\partial_p}\,\overrightarrow{\partial_x}\right)\right] B_{\mathrm{W}} \tag{5}$$

$$\frac{\partial}{\partial t} W = \{H_{\mathrm{W}}, W\}_{\mathrm{W}} \tag{6}$$

$$\{H_{\mathrm{W}}, W\}_{W} = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!2^{2n}} H_W \left(\overleftarrow{\partial_x}\,\overrightarrow{\partial_p} - \overleftarrow{\partial_p}\,\overrightarrow{\partial_x}\right)^{2n+1} W \tag{7}$$

where H_{W} is the Weyl transform of the Hamiltonian, and where $\{H_{\mathrm{W}}, W\}_{\mathrm{W}}$ denotes the Moyal bracket [28].

It can be seen that equations (2)–(4) and equations (5)–(7) are formally very similar. However, this formal resemblance is somewhat deceptive, as it turns out that the two sets of equations have quite different convergence properties.

The formula for the Weyl transform of an operator product, equation (5), is exact if either $\hat{A}$ or $\hat{B}$ is a polynomial in $\hat{x}$ and $\hat{p}$ (in which case the series terminates after a finite number of terms). More generally, if A_{W}, B_{W} are C^{∞} functions satisfying appropriate conditions on their growth at infinity, then it can be shown that the series is asymptotic [30]. However, there are many operators of physical interest for which the Weyl transform is only defined in a distributional sense, and for operators such as this the series can be highly singular. Consider, for example, the parity operator $\hat{V}$, whose action in the x-representation is given by

$$\langle x | \hat{V} | \psi \rangle = \langle -x \mid \psi \rangle.$$

We have

$$V_{\mathrm{W}}(x, p) = \pi \delta(x)\delta(p).$$

Substituting this expression into equations (5) gives

$$(\hat{V}^2)_{\mathrm{W}}(x, p) = \pi^2 \delta(x)\delta(p) \exp\left[\tfrac{1}{2}\mathrm{i}\left(\overleftarrow{\partial_x}\,\overrightarrow{\partial_p} - \overleftarrow{\partial_p}\,\overrightarrow{\partial_x}\right)\right]\delta(x)\delta(p).$$

The left-hand side of this equation is equal to 1, whereas the expression on the right-hand side is an infinite sum, each individual term of which is ill-defined (being a product of distributions concentrated at the origin).

Mizrahi's [15] derivation of the formula for the Husimi transform of an operator product, equation (2), depends on the same kind of formal manipulation that is used in Groenewold's [27] derivation of equation (5), and so it might be supposed that the validity of the formula is similarly restricted. However, it turns out that the sum in equation (2) is actually much better behaved. In fact, it will be shown in section 2 that, subject to certain not very restrictive conditions on the operators $\hat{A}$ and $\hat{B}$, the sum on the right-hand side of equation (2) is not only defined and asymptotic; it is even absolutely convergent for all x and p. This is essentially because A_{H} is typically a much less singular object than A_{W} (due to the Gaussian convolution in equation (1)).

In section 3 we apply the result just described to the problem of expressing expectation values in terms of the Husimi function.

The expectation value of an operator $\hat{A}$ can be obtained from the Wigner function using the formula [1, 2]

$$\mathrm{Tr}(\hat{\rho}\hat{A}) = \int \mathrm{d}x\, \mathrm{d}p\, A_{\mathrm{W}}(x, p) W(x, p). \tag{8}$$

In certain cases we can also express the expectation value in terms of the Husimi function using [1, 2, 15]

$$\mathrm{Tr}(\hat{\rho}\hat{A}) = \int \mathrm{d}x\, \mathrm{d}p\, A_{\overline{\mathrm{H}}}(x, p)\, Q(x, p) \tag{9}$$

where $A_{\overline{\mathrm{H}}}$ is the anti-Husimi transform (or contravariant symbol) of $\hat{A}$, defined by

$$A_{\overline{\mathrm{H}}} = \mathrm{e}^{-\partial_+\partial_-} A_{\mathrm{H}} \tag{10}$$

(with $\partial_\pm = 2^{-1/2}(\partial_x \mp \mathrm{i}\partial_p)$, as before). Equation (9) is valid (for example) whenever [31] $A_{\overline{\mathrm{H}}}$ exists as a tempered distribution and Q belongs to the corresponding space of test functions (i.e. the C^∞ functions of rapid decrease). However, we have the problem that $A_{\overline{\mathrm{H}}}$ is often so highly singular that it is not defined as a tempered distribution, which means that the usefulness of equation (9) is somewhat limited. This is often seen as a serious drawback of the Husimi formalism.

However, it turns out that it is often possible to circumvent this difficulty. Suppose we substitute the series given by equation (10) into the right-hand side of equation (9), and suppose we then reverse the order of the sum and integral. This gives

$$\mathrm{Tr}(\hat{\rho}\hat{A}) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \int \mathrm{d}x\, \mathrm{d}p\, \left(\partial_+^n \partial_-^n A_{\mathrm{H}}(x, p)\right) Q(x, p). \tag{11}$$

In section 3 we show that it often happens that the sum on the right-hand side of this equation is absolutely convergent, even in many of the cases where $A_{\overline{\mathrm{H}}}$ fails to exist as a tempered distribution.

2. Convergence of the product formula

We will find it convenient to work in terms of coherent states. Define

$$\hat{a} = \frac{1}{\sqrt{2}}(\hat{x} + \mathrm{i}\hat{p}) \qquad \hat{a}^\dagger = \frac{1}{\sqrt{2}}(\hat{x} - \mathrm{i}\hat{p})$$

and let ϕ_n denote the nth (normalized) eigenstate of the number operator $\hat{a}^\dagger\hat{a}$:

$$\hat{a}\,\phi_0 = 0 \qquad \phi_n = \frac{1}{\sqrt{n!}}(\hat{a}^\dagger)^n \phi_0.$$

Let $\hat{D}_{xp}$ be the displacement operator

$$\hat{D}_{xp} = \mathrm{e}^{\mathrm{i}(p\hat{x} - x\hat{p})}$$

and define

$$\phi_{n;xp} = \hat{D}_{xp}\phi_n \qquad \phi_{xp} = \phi_{0;xp}.$$

The ϕ_{xp} are the coherent states. Let $\hat{A}$ be any operator (not necessarily bounded) with domain of definition $\mathscr{D}_{\hat{A}}$, and suppose that $\phi_{xp} \in \mathscr{D}_{\hat{A}}$ for all x, p. It is then straightforward to show that

$$A_{\mathrm{H}}(x, p) = \langle \phi_{xp}, \ \hat{A}\phi_{xp} \rangle$$

(we no longer use the Dirac bra–ket notation, because the existence of $\langle \psi, \ \hat{A}\chi \rangle$ does not, in general, imply the existence of $\langle \hat{A}^{\dagger}\psi, \ \chi \rangle$).

If $\hat{A}$, $\hat{B}$ are both bounded then the proof of equation (2) is comparatively straightforward. However, we want to make the proof as general as possible. We then have the difficulty that the sum in equation (2) will only be defined if A_{H} and B_{H} are both C^{∞}; whereas functions of the form $\langle \hat{D}_{xp}\psi, \ \hat{A}\hat{D}_{xp}\psi \rangle$ are, in general, not even once-differentiable, let alone C^{∞}. We are thus faced with the question: what conditions must we impose on the operator $\hat{A}$ in order to ensure that the function A_{H} is C^{∞}? One answer to this question is given by the following theorem.

Theorem 1. *Let $\mathscr{D}_{\hat{A}}$, $\mathscr{D}_{\hat{A}^{\dagger}}$ be the domains of definition of $\hat{A}$, $\hat{A}^{\dagger}$, respectively. Suppose that $\phi_{xp} \in \mathscr{D}_{\hat{A}} \cap \mathscr{D}_{\hat{A}^{\dagger}}$ for all x, p. Suppose, also, that $\langle \phi_{x_1 p_1}, \ \hat{A}\phi_{x_2 p_2} \rangle$, $\langle \phi_{1;x_1 p_1}, \ \hat{A}\phi_{x_2 p_2} \rangle$ and $\langle \phi_{1;x_1 p_1}, \ \hat{A}^{\dagger}\phi_{x_2 p_2} \rangle$ are continuous functions on $\mathbb{R}^4$. Then A_{H} is an analytic function, which uniquely continues to a holomorphic function defined on the whole of $\mathbb{C}^2$.*

The continuation is given by

$$A_{\mathrm{H}}(x, p) = \frac{\langle \phi_{x_- p_-}, \ \hat{A}\phi_{x_+ p_+} \rangle}{\langle \phi_{x_- p_-}, \ \phi_{x_+ p_+} \rangle} \tag{12}$$

where x, p are arbitrary complex, and where $x_{\pm}$, $p_{\pm}$ are the real variables defined by

$$x_{\pm} = \tfrac{1}{2}(x + x^*) \pm \tfrac{1}{2}\mathrm{i}(p - p^*) \tag{13}$$

$$p_{\pm} = \tfrac{1}{2}(p + p^*) \mp \tfrac{1}{2}\mathrm{i}(x - x^*). \tag{14}$$

This theorem is a strengthened version of results proved by Mehta and Sudarshan [25] and Appleby [26]. The proof is given in appendix A.

It is worth noting that the condition in the statement of this theorem is quite weak. If the three functions listed exist and are continuous then, without making any explicit assumption regarding the differentiability of these functions, it automatically follows that A_{H} must be complex analytic.

We also have the following lemma.

Lemma 2. *Suppose that $\hat{A}$ satisfies the conditions of theorem 1. Then*

$$\frac{\langle \phi_{n;x_- p_-}, \ \hat{A}\phi_{x_+ p_+} \rangle}{\langle \phi_{x_- p_-}, \ \phi_{x_+ p_+} \rangle} = \frac{1}{\sqrt{n!}} \sum_{r=0}^{n} \binom{n}{r} (z_+ - z_-^*)^{n-r} \frac{\partial^r}{\partial z_-^r} A_{\mathrm{H}}(x, p) \tag{15}$$

$$\frac{\langle \hat{A}^{\dagger}\phi_{x_- p_-}, \ \phi_{n;x_+ p_+} \rangle}{\langle \phi_{x_- p_-}, \ \phi_{x_+ p_+} \rangle} = \frac{1}{\sqrt{n!}} \sum_{r=0}^{n} \binom{n}{r} (z_- - z_+^*)^{n-r} \frac{\partial^r}{\partial z_+^r} A_{\mathrm{H}}(x, p) \tag{16}$$

where $x_{\pm}$, $p_{\pm}$ are the variables defined by equations (13) and (14), and where

$$z_{\pm} = \frac{1}{\sqrt{2}}(x \pm \mathrm{i}p).$$

The proof of this lemma is given in appendix B.

If x, p are both real (so that $z_- = z_+^*$) equations (15) and (16) become

$$\langle \phi_{n;xp}, \hat{A}\phi_{xp} \rangle = \frac{1}{\sqrt{n!}} \partial_-^n A_{\mathrm{H}}(x, p) \tag{17}$$

$$\langle \hat{A}^\dagger \phi_{xp}, \phi_{n;xp} \rangle = \frac{1}{\sqrt{n!}} \partial_+^n A_{\mathrm{H}}(x, p) \tag{18}$$

where

$$\partial_\pm = \frac{\partial}{\partial z_\pm} = \frac{1}{\sqrt{2}} \left(\frac{\partial}{\partial x} \mp \mathrm{i} \frac{\partial}{\partial p} \right).$$

Using these results the proof of the product formula becomes very straightforward. Let $\hat{A}$, $\hat{B}$ be any pair of operators satisfying the conditions of theorem 1. Suppose also that $\phi_{xp} \in \mathscr{D}_{\hat{A}\hat{B}}$ for all $x, p \in \mathbb{R}$. Then, for real x, p,

$$\begin{aligned} (\hat{A}\hat{B})_{\mathrm{H}}(x, p) &= \langle \phi_{xp}, \hat{A}\hat{B}\phi_{xp} \rangle = \langle \hat{A}^\dagger \phi_{xp}, \hat{B}\phi_{xp} \rangle \\ &= \sum_{n=0}^{\infty} \langle \hat{A}^\dagger \phi_{xp}, \phi_{n;xp} \rangle \langle \phi_{n;xp}, \hat{B}\phi_{xp} \rangle \end{aligned} \tag{19}$$

where we have used the fact that the $\phi_{n;xp}$ constitute an orthonormal basis. The absolute convergence of this sum is an immediate consequence of basic Hilbert space theory. Using equations (17) and (18) we deduce

$$(\hat{A}\hat{B})_{\mathrm{H}}(x, p) = \sum_{n=0}^{\infty} \frac{1}{n!} \partial_+^n A_{\mathrm{H}}(x, p) \partial_-^n B_{\mathrm{H}}(x, p) = A_{\mathrm{H}}(x, p)\, \mathrm{e}^{\overleftarrow{\partial}_+ \overrightarrow{\partial}_-} B_{\mathrm{H}}(x, p)$$

which is the product formula. The absolute convergence of this sum follows from the absolute convergence of the sum in equation (19).

3. Expectation values

We now discuss the implications that the result just proved has for the convergence of equation (11), giving the expectation value of $\hat{A}$ in terms of the Husimi function.

Of course, one does not expect the right-hand side of equation (11) to converge for arbitrary $\hat{A}$ and $\hat{\rho}$, since, apart from anything else, an unbounded operator does not have a well defined expectation value for every state $\hat{\rho}$. We therefore need to place some kind of restriction on the class of operators $\hat{A}$ and density matrices $\hat{\rho}$ considered. The result we prove is probably not the most general possible. However, it will serve to illustrate the point, that the sum on the right-hand side of equation (11) is often absolutely convergent, even in many of the cases where the anti-Husimi transform fails to exist as a tempered distribution.

We accordingly confine ourselves to the case of density matrices for which the Husimi function $Q \in \mathscr{S}(\mathbb{R}^2)$, where $\mathscr{S}(\mathbb{R}^2)$ is the space of C^∞ functions which are rapidly decreasing at infinity [31] (i.e. the space of test functions for the space of tempered distributions). In other words, we assume that

$$\sup_{(x,p)\in\mathbb{R}^2} \left| (1 + x^2 + p^2)^l \partial_x^m \partial_p^n Q(x, p) \right| < \infty$$

for every triplet of non-negative integers l, m, n.

We assume that $\hat{A}$ has the properties

(a) $\phi_{xp} \in \mathscr{D}_{\hat{A}} \cap \mathscr{D}_{\hat{A}^\dagger}$ for all $x, p \in \mathbb{R}$.
(b) There exist positive constants $K_\pm$ and non-negative integers $N_\pm$ such that

$$\|\hat{A}\phi_{xp}\| \leqslant K_-(1+x^2+p^2)^{N_-} \tag{20}$$

$$\|\hat{A}^\dagger\phi_{xp}\| \leqslant K_+(1+x^2+p^2)^{N_+} \tag{21}$$

for all $x, p \in \mathbb{R}$.

We will say that an operator satisfying these two conditions is *polynomially bounded*. The following lemma gives two properties of such operators which will be needed in the following.

Lemma 3. *Suppose that $\hat{A}$ is polynomially bounded. Then*

(a) $\hat{A}$ satisfies the conditions of theorem 1. In particular, $A_{\rm H}$ is analytic.
(b) For every pair of non-negative integers m, n there exists a positive constant K_{mn} and a non-negative integer N_{mn} such that

$$\left|\partial_x^m\partial_p^n A_{\rm H}(x,p)\right| \leqslant K_{mn}(1+x^2+p^2)^{N_{mn}} \tag{22}$$

for all $x, p \in \mathbb{R}$.

The proof is given in appendix C.

We are now ready to prove the main result of this section.

Theorem 4. *Suppose that $\hat{A}$ is polynomially bounded, and suppose that the density matrix $\hat{\rho}$ is such that the corresponding Husimi function is rapidly decreasing at infinity. Suppose, also, that $\hat{A}\hat{\rho}$ is of trace-class. Then*

$$\mathrm{Tr}(\hat{A}\hat{\rho}) = \sum_{n=0}^{\infty}\frac{(-1)^n}{n!}\int \mathrm{d}x\,\mathrm{d}p\,\left(\partial_+^n\partial_-^n A_{\rm H}(x,p)\right)Q(x,p) \tag{23}$$

where the sum on the right-hand side is absolutely convergent.

Proof. We have

$$\begin{aligned}\mathrm{Tr}(\hat{A}\hat{\rho}) &= \frac{1}{2\pi}\int \mathrm{d}x\,\mathrm{d}p\,\langle\phi_{xp},\ \hat{A}\hat{\rho}\phi_{xp}\rangle \\ &= \frac{1}{2\pi}\int \mathrm{d}x\,\mathrm{d}p\left(\sum_{n=0}^{\infty}\langle\hat{A}^\dagger\phi_{xp},\ \phi_{n;xp}\rangle\langle\phi_{n;xp},\ \hat{\rho}\phi_{xp}\rangle\right).\end{aligned}$$

We now use Lebesgue's dominated convergence theorem [31] to show that we may reverse the order of sum and integral. In fact, it follows from the Schwartz inequality that

$$\begin{aligned}\left|\sum_{n=0}^{m}\langle\hat{A}^\dagger\phi_{xp},\ \phi_{n;xp}\rangle\langle\phi_{n;xp},\ \hat{\rho}\phi_{xp}\rangle\right| &\leqslant \left(\left(\sum_{n=0}^{m}\left|\langle\hat{A}^\dagger\phi_{xp},\ \phi_{n;xp}\rangle\right|^2\right)\left(\sum_{n=0}^{m}\left|\langle\phi_{n;xp},\ \hat{\rho}\phi_{xp}\rangle\right|^2\right)\right)^{\frac{1}{2}} \\ &\leqslant \|\hat{A}^\dagger\phi_{xp}\|\|\hat{\rho}\phi_{xp}\|.\end{aligned}$$

We have

$$\|\hat{\rho}\phi_{xp}\| = \left(\langle\phi_{xp},\ \hat{\rho}^2\phi_{xp}\rangle\right)^{\frac{1}{2}} \leqslant \left(\langle\phi_{xp},\ \hat{\rho}\phi_{xp}\rangle\right)^{\frac{1}{2}} = \sqrt{2\pi\,Q(x,p)}$$

which, together with the inequality (21), implies

$$\|\hat{A}^\dagger\phi_{xp}\|\|\hat{\rho}\phi_{xp}\| \leqslant \sqrt{2\pi}\,K_+(1+x^2+p^2)^{N_+}\sqrt{Q(x,p)}.$$

By assumption, $Q(x, p) \in \mathscr{S}(\mathbb{R}^2)$. It follows that $\|\hat{A}^\dagger \phi_{xp}\|\|\hat{\rho}\phi_{xp}\|$ is integrable. We may therefore use Lebesgue's dominated convergence theorem [31] to deduce

$$\mathrm{Tr}(\hat{A}\hat{\rho}) = \frac{1}{2\pi} \sum_{n=0}^{\infty} \int \mathrm{d}x\, \mathrm{d}p\, \langle \hat{A}^\dagger \phi_{xp},\, \phi_{n;xp} \rangle \langle \phi_{n;xp},\, \hat{\rho}\phi_{xp} \rangle \tag{24}$$

where the sum is absolutely convergent, since

$$\sum_{n=0}^{\infty} \left| \int \mathrm{d}x\, \mathrm{d}p\, \langle \hat{A}^\dagger \phi_{xp},\, \phi_{n;xp} \rangle \langle \phi_{n;xp},\, \hat{\rho}\phi_{xp} \rangle \right| \leqslant \int \mathrm{d}x\, \mathrm{d}p\, \|\hat{A}^\dagger \phi_{xp}\|\|\hat{\rho}\phi_{xp}\| < \infty.$$

We know from lemma 3 that $\hat{A}$ satisfies the conditions of theorem 1. We may therefore use the results proved in the last section to rewrite equation (24) in the form

$$\mathrm{Tr}(\hat{A}\hat{\rho}) = \sum_{n=0}^{\infty} \frac{1}{n!} \int \mathrm{d}x\, \mathrm{d}p\, \partial_+^n A_{\mathrm{H}}(x, p) \partial_-^n Q(x, p).$$

Finally, it follows from the inequality (22), together with the fact that $Q \in \mathscr{S}(\mathbb{R}^2)$, that we may partially integrate term-by-term to obtain

$$\mathrm{Tr}(\hat{A}\hat{\rho}) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \int \mathrm{d}x\, \mathrm{d}p\, \big(\partial_+^n \partial_-^n A_{\mathrm{H}}(x, p)\big) Q(x, p). \qquad \square$$

The right-hand side of equation (8) (expressing $\langle \hat{A} \rangle$ in terms of the Wigner function) is defined whenever $W \in \mathscr{S}(\mathbb{R}^2)$ and A_{W} exists as a tempered distribution. On the other hand, although it is true that $Q \in \mathscr{S}(\mathbb{R}^2)$ whenever $W \in \mathscr{S}(\mathbb{R}^2)$ (see theorem IX.3 of [31]), the fact that A_{W} exists as a tempered distribution is not evidently sufficient to ensure that $\hat{A}$ is polynomially bounded. So we have not shown that equation (23) has the same range of validity as equation (8). However, it can be shown that $\hat{A}$ is polynomially bounded if $\phi_{xp} \in \mathscr{D}_{\hat{A}} \cap \mathscr{D}_{\hat{A}^\dagger}$, and if $(\hat{A}^\dagger \hat{A})_{\mathrm{W}}$ and $(\hat{A}\hat{A}^\dagger)_{\mathrm{W}}$ exist as tempered distributions (see theorem IX.4 of [31]). In applications one meets operators satisfying these conditions much more commonly than one meets operators for which $A_{\overline{\mathrm{H}}}$ exists as a tempered distribution. For instance, every bounded operator is polynomially bounded, whereas there are many bounded operators of physical interest for which $A_{\overline{\mathrm{H}}}$ fails to exist as a tempered distribution. The above result consequently represents a significant improvement on the results that were previously known.

Of course, just from the fact that equation (23) is convergent, it does not necessarily follow that the convergence is sufficiently rapid to make the formula useful in practical, numerical work. This question requires further investigation.

4. Conclusion

As has been stressed by Mizrahi [15], Lalović *et al* [9], Davidović and Lalović [33] and others, the Husimi formalism provides an especially perspicuous method for studying the relationship between quantum and classical mechanics. It establishes a one-to-one correspondence between the basic equations of the two theories, so that one can start with a classical formula, and then turn it into the corresponding quantum formula by adding successive correction terms. Moreover, the fact that $Q(x, p)$ describes the outcome of a retrodictively optimal joint measurement of x and p [19, 20], means that one could reasonably argue that the Husimi function is the most natural choice for a quantum mechanical analogue of the classical probability distribution.

In this paper we have investigated the convergence properties of two of the key formulae in the Husimi formalism. We have shown that the formula giving the Husimi transform of an operator product has much better convergence properties than the corresponding formula in the Wigner function formalism. In particular, the Husimi formalism leads to a convergent generalization of the Liouville equation for a very large class of Hamiltonians. We have also shown that the convergence properties of the formula expressing the expectation value $\langle \hat{A} \rangle$ in terms of the Husimi function, although seemingly not as good as those of the corresponding formula in the Wigner function formalism, are significantly better than the often highly singular character of $A_{\overline{\mathrm{H}}}$ would suggest.

These results lend additional support to the suggestion that, in so far as the aim is specifically to formulate quantum mechanics as a kind of generalized version of classical mechanics, then the formalism based on the Husimi function has some significant advantages.

Appendix A. Proof of theorem 1

For arbitrary complex x, p define

$$F(x, p) = \frac{\langle \phi_{x_- p_-}, \hat{A}\, \phi_{x_+ p_+} \rangle}{\langle \phi_{x_- p_-}, \phi_{x_+ p_+} \rangle}$$

where $x_\pm$, $p_\pm$ are the (real) variables defined by equations (13) and (14).

It is easily seen that, if x, p are both real, then

$$F(x, p) = \langle \phi_{xp}, \hat{A}\, \phi_{xp} \rangle = A_{\mathrm{H}}(x, p).$$

The problem thus reduces to that of showing that, if $\hat{A}$ has the properties stipulated, then F is holomorphic. We will do this by showing that F satisfies the Cauchy–Riemann equations with respect to the complex variables

$$z_\pm = \frac{1}{\sqrt{2}}(x_\pm \pm \mathrm{i} p_\pm) = \frac{1}{\sqrt{2}}(x \pm \mathrm{i} p). \tag{A1}$$

In fact, it is straightforward to show that ϕ_{xp}, regarded as a vector-valued function of two real variables, is differentiable in the norm topology; the derivatives being given by

$$\frac{\partial}{\partial x}\phi_{xp} = \frac{1}{\sqrt{2}}\phi_{1;\,xp} - \frac{\mathrm{i}}{2} p \phi_{xp}$$

$$\frac{\partial}{\partial p}\phi_{xp} = \frac{\mathrm{i}}{\sqrt{2}}\phi_{1;\,xp} + \frac{\mathrm{i}}{2} x \phi_{xp}.$$

Also,

$$\langle \phi_{x_- p_-}, \phi_{x_+ p_+} \rangle = \exp\left[-\tfrac{1}{4}(x_+ - x_-)^2 - \tfrac{1}{4}(p_+ - p_-)^2 + \tfrac{1}{2}\mathrm{i}(p_+ x_- - p_- x_+)\right].$$

Consequently, F is differentiable with respect to the variables x_-, p_-. Moreover,

$$\begin{aligned}\frac{\partial}{\partial x_-}F(x, p) &= \frac{1}{\sqrt{2}}\frac{\langle \phi_{1;\,x_- p_-}, \hat{A}\phi_{x_+ p_+} \rangle}{\langle \phi_{x_- p_-}, \phi_{x_+ p_+} \rangle} + \frac{1}{\sqrt{2}}(z_-^* - z_+)F(x, p) \\ &= \mathrm{i}\frac{\partial}{\partial p_-}F(x, p)\end{aligned} \tag{A2}$$

from which it follows that F satisfies the Cauchy–Riemann equations with respect to the complex variable $z_- = (x_- - \mathrm{i}p_-)/\sqrt{2}$.

We can alternatively write

$$F(x,\,p)=\frac{\langle \hat{A}^{\dagger}\,\phi_{x_-p_-},\ \phi_{x_+p_+}\rangle}{\langle \phi_{x_-p_-},\ \phi_{x_+p_+}\rangle}.$$

Consequently, F is also differentiable with respect to the real variables x_+, p_+. Moreover,

$$\begin{aligned}\frac{\partial}{\partial x_+}F(x,\,p)&=\frac{1}{\sqrt{2}}\frac{\langle \hat{A}^{\dagger}\,\phi_{x_-p_-},\ \phi_{1;\,x_+p_+}\rangle}{\langle \phi_{x_-p_-},\ \phi_{x_+p_+}\rangle}+\frac{1}{\sqrt{2}}(z_+^*-z_-)F(x,\,p)\\&=-\mathrm{i}\frac{\partial}{\partial p_+}F(x,\,p)\end{aligned}\qquad\text{(A3)}$$

from which it follows that F satisfies the Cauchy–Riemann equations with respect to the complex variable $z_+=(x_++\mathrm{i}p_+)/\sqrt{2}$.

If $\hat{A}$ has the properties specified in the statement of theorem, then we see from equations (A2) and (A3) that the partial derivatives $\partial F/\partial x_\pm$, $\partial F/\partial p_\pm$, are continuous functions on $\mathbb{R}^4$. It follows [32] that F is a holomorphic function of the complex variables $z_\pm$. Referring to equation (A1) it can be seen that the variables $z_\pm$ are linear combinations of x, p. We conclude that F is a holomorphic function of x, p.

Appendix B. Proof of lemma 2

It is straightforward to show that $\phi_{n;x_-p_-}$, regarded as a vector-valued function of two real variables, is differentiable in the norm topology. Moreover,

$$\frac{1}{\sqrt{2}}\left(\frac{\partial}{\partial x_-}-\mathrm{i}\frac{\partial}{\partial p_-}\right)\phi_{n;x_-p_-}=\sqrt{n+1}\,\phi_{(n+1);x_-p_-}+\tfrac{1}{2}z_-\,\phi_{n;x_-p_-}.$$

Hence

$$\begin{aligned}&\frac{1}{\sqrt{2}}\left(\frac{\partial}{\partial x_-}+\mathrm{i}\frac{\partial}{\partial p_-}\right)\frac{\langle \phi_{n;x_-p_-},\ \hat{A}\phi_{x_+p_+}\rangle}{\langle \phi_{x_-p_-},\ \phi_{x_+p_+}\rangle}\\&\qquad=\sqrt{n+1}\frac{\langle \phi_{(n+1);x_-p_-},\ \hat{A}\phi_{x_+p_+}\rangle}{\langle \phi_{x_-p_-},\ \phi_{x_+p_+}\rangle}+\left(z_-^*-z_+\right)\frac{\langle \phi_{n;x_-p_-},\ \hat{A}\phi_{x_+p_+}\rangle}{\langle \phi_{x_-p_-},\ \phi_{x_+p_+}\rangle}.\end{aligned}$$

Iterating this result, and using

$$\frac{1}{2^{r/2}}\left(\frac{\partial}{\partial x_-}+\mathrm{i}\frac{\partial}{\partial p_-}\right)^r A_{\mathrm{H}}(x,\,p)=\frac{\partial^r}{\partial z_-^r}A_{\mathrm{H}}(x,\,p)$$

we obtain equation (15).

The proof of equation (16) is similar.

Appendix C. Proof of lemma 3

Proof of (1). We need to show that, if $\hat{A}$ is polynomially bounded, then the functions $\langle \phi_{x_1p_1},\ \hat{A}\phi_{x_2p_2}\rangle$, $\langle \phi_{1;x_1p_1},\ \hat{A}\phi_{x_2p_2}\rangle$ and $\langle \phi_{1;x_1p_1},\ \hat{A}^{\dagger}\phi_{x_2p_2}\rangle$ are continuous.

Consider the function $\langle \phi_{1;x_1p_1},\ \hat{A}\phi_{x_2p_2}\rangle$. We have

$$\begin{aligned}&\left|\left\langle \phi_{1;x_1'p_1'},\ \hat{A}\phi_{x_2'p_2'}\right\rangle-\left\langle \phi_{1;x_1p_1},\ \hat{A}\phi_{x_2p_2}\right\rangle\right|\\&\qquad\leqslant\left|\left\langle (\phi_{1;x_1'p_1'}-\phi_{1;x_1p_1}),\ \hat{A}\phi_{x_2'p_2'}\right\rangle\right|+\left|\left\langle \phi_{1;x_1p_1},\ \hat{A}(\phi_{x_2'p_2'}-\phi_{x_2p_2})\right\rangle\right|.\end{aligned}$$

In view of equation (20) we have

$$\left|\left\langle(\phi_{1;x_1'p_1'}-\phi_{1;x_1p_1}),\ \hat{A}\phi_{x_2'p_2'}\right\rangle\right| \leqslant K_-(1+x_2'^2+p_2'^2)^{N_-}\|\phi_{1;x_1'p_1'}-\phi_{1;x_1p_1}\|.$$

Also, using the completeness relation for coherent states, together with equation (21), we find

$$\left|\left\langle\phi_{1;x_1p_1},\ \hat{A}(\phi_{x_2'p_2'}-\phi_{x_2p_2})\right\rangle\right| = \left|\frac{1}{2\pi}\int \mathrm{d}x_3\,\mathrm{d}p_3\left\langle\phi_{1;x_1p_1},\ \phi_{x_3p_3}\right\rangle\left\langle\hat{A}^\dagger\phi_{x_3p_3},\ (\phi_{x_2'p_2'}-\phi_{x_2p_2})\right\rangle\right|$$
$$\leqslant f(x_1,p_1)\|\phi_{x_2'p_2'}-\phi_{x_2p_2}\|$$

where f is the polynomial

$$f(x_1,p_1) = \frac{K_+}{2\pi}\int \mathrm{d}x_3\,\mathrm{d}p_3\left|\left\langle\phi_{1;x_1p_1},\ \phi_{x_3p_3}\right\rangle\right|(1+x_3^2+p_3^2)^{N_+}$$
$$= \frac{K_+}{2^{\frac{3}{2}}\pi}\int \mathrm{d}x_3'\,\mathrm{d}p_3'\sqrt{1+x_3'^2+p_3'^2}\left(1+(x_3'+x_1)^2+(p_3'+p_1)^2\right)^{N_+}$$
$$\times\exp\left[-\tfrac{1}{4}\left(x_3'^2+p_3'^2\right)\right].$$

Putting these results together we find

$$\left|\left\langle\phi_{1;x_1'p_1'},\ \hat{A}\phi_{x_2'p_2'}\right\rangle-\left\langle\phi_{1;x_1p_1},\ \hat{A}\phi_{x_2p_2}\right\rangle\right|$$
$$\leqslant K_-(1+x_2'^2+p_2'^2)^{N_-}\|\phi_{1;x_1'p_1'}-\phi_{1;x_1p_1}\|+f(x_1,p_1)\|\phi_{x_2'p_2'}-\phi_{x_2p_2}\|$$

and ϕ_{xp} and $\phi_{1;xp}$, regarded as vector-valued functions on $\mathbb{R}^2$, are continuous in the norm topology. Consequently,

$$\left\langle\phi_{1;x_1'p_1'},\ \hat{A}\phi_{x_2'p_1'}\right\rangle\to\left\langle\phi_{1;x_1p_1},\ \hat{A}\phi_{x_2p_2}\right\rangle$$

as $(x_1',p_1',x_2',p_2')\to(x_1,p_1,x_2,p_2)$. It follows that $\left\langle\phi_{1;x_1p_1},\ \hat{A}\phi_{x_2p_2}\right\rangle$ is continuous. Continuity of the functions $\left\langle\phi_{x_1p_1},\ \hat{A}\phi_{x_2p_2}\right\rangle$ and $\left\langle\phi_{1;x_1p_1},\ \hat{A}^\dagger\phi_{x_2p_2}\right\rangle$ is proved in the same way.

Proof of (2). The completeness relation for coherent states implies

$$A_{\mathrm{H}}(x,p) = \langle\phi_{xp},\ \hat{A}\phi_{xp}\rangle = \frac{1}{2\pi}\int \mathrm{d}x'\,\mathrm{d}p'\,\langle\phi_{xp},\ \phi_{x'p'}\rangle\langle\hat{A}^\dagger\phi_{x'p'},\ \phi_{xp}\rangle.$$

Using

$$\partial_+\phi_{n;xp} = \sqrt{n+1}\phi_{(n+1);xp}+\tfrac{1}{2}z^*\phi_{n;xp}$$
$$\partial_-\phi_{n;xp} = \begin{cases}-\frac{1}{2}z\phi_{xp} & \text{if}\quad n=0\\ -\sqrt{n}\phi_{(n-1);xp}-\frac{1}{2}z\phi_{n;xp} & \text{if}\quad n>0\end{cases}$$

(where $\partial_\pm = 2^{-1/2}(\partial_x\mp\mathrm{i}\partial_p)$ and $z=2^{-1/2}(x+\mathrm{i}p)$)), and differentiating under the integral sign, it is not difficult to show that

$$\partial_x^m\partial_p^n A_{\mathrm{H}}(x,p) = \sum_{r,s=0}^{n+m} c_{rs}\int \mathrm{d}x'\,\mathrm{d}p'\,\langle\phi_{r;xp},\ \phi_{x'p'}\rangle\langle\hat{A}^\dagger\phi_{x'p'},\ \phi_{s;xp}\rangle$$

for suitable constants c_{rs}. We have

$$\left|\langle\phi_{r;xp},\ \phi_{x'p'}\rangle\right| = \frac{1}{2^{r/2}\sqrt{r!}}\left((x'-x)^2+(p'-p)^2\right)^{r/2}\exp\left[-\tfrac{1}{4}(x'-x)^2-\tfrac{1}{4}(p'-p)^2\right].$$

In view of the inequality (21) it follows that

$$\left|\int \mathrm{d}x'\,\mathrm{d}p'\,\langle\phi_{r;xp},\,\phi_{x'p'}\rangle\langle\hat{A}^{\dagger}\phi_{x'p'},\,\phi_{s;xp}\rangle\right| \leqslant \frac{K_+}{2^{r/2}\sqrt{r!}}\int \mathrm{d}x''\,\mathrm{d}p''\,\left(1+(x''+x)^2+(p''+p)^2\right)^{N_+} \times\left(x''^2+p''^2\right)^{r/2}\exp\left[-\tfrac{1}{4}\left(x''^2+p''^2\right)\right].$$

It can be seen that the expression on the right-hand side of this inequality is a polynomial in x and p. Consequently,

$$\left|\partial_x^m\partial_p^n A_{\mathrm{H}}(x,\,p)\right| \leqslant f(x,\,p)$$

for some polynomial $f(x,\,p)$. The claim is now immediate.

References

[1] Hillery M, O'Connell R F, Scully M O and Wigner E P 1984 *Phys. Rep.* **106** 121
[2] Lee H W 1995 *Phys. Rep.* **259** 147
[3] Leonhardt U 1997 *Measuring the Quantum State of Light* (Cambridge: Cambridge University Press)
[4] Wigner E P 1932 *Phys. Rev.* **40** 749
[5] Hudson R L 1974 *Rep. Math. Phys.* **6** 249
Soto F and Claverie P 1983 *J. Math. Phys.* **24** 97
Jagannathan R, Simon R, Sudarshan E C G and Vasudevan R 1987 *Phys. Lett.* A **120** 161
Narcowich F J 1988 *J. Math. Phys.* **29** 2036
Bröcker T and Werner R F 1995 *J. Math. Phys.* **36** 62
[6] Davies E B 1976 *Quantum Theory of Open Systems* (London: Academic)
[7] Cartwright N D 1976 *Physica* A **83** 210
[8] Wódkiewicz K 1984 *Phys. Rev. Lett.* **52** 1064
Wódkiewicz K 1986 *Phys. Lett.* A **115** 304
Wódkiewicz K 1987 *Phys. Lett.* A **124** 207
[9] Lalović D, Davidović D M and Bijedić N 1992 *Phys. Rev.* A **46** 1206
Lalović D, Davidović D M and Bijedić N 1992 *Physica* A **184** 231
Lalović D, Davidović D M and Bijedić N 1992 *Phys. Lett.* A **166** 99
[10] Halliwell J J 1992 *Phys. Rev.* D **46** 1610
[11] Wünsche A 1996 *Quantum Semiclass. Opt.* **8** 343
Wünsche A and Bužek V 1997 *Quantum Semiclass. Opt.* **9** 631
[12] Husimi K 1940 *Proc. Phys. Math. Soc. Japan* **22** 264
[13] Kano Y 1965 *J. Math. Phys.* **6** 1913
[14] Glauber R J 1965 *Quantum Optics and Electronics* ed C de Witt, A Blandin and C Cohen-Tannoudji (New York: Gordon and Breach)
[15] Mizrahi S S 1984 *Physica* A **127** 241
Mizrahi S S 1986 *Physica* A **135** 237
Mizrahi S S 1988 *Physica* A **150** 541
[16] Arthurs E and Kelly J L 1965 *Bell Syst. Tech. J.* **44** 725
Busch P 1985 *Int. J. Theor. Phys.* **24** 63
Braunstein S L, Caves C M and Milburn G J 1991 *Phys. Rev.* A **43** 1153
Stenholm S 1992 *Ann. Phys., NY* **218** 233
Appleby D M 1998 *J. Phys. A: Math. Gen.* **31** 6419
[17] Raymer M G 1994 *Am. J. Phys.* **62** 986
[18] Leonhardt U and Paul H 1993 *J. Mod. Opt.* **40** 1745
Leonhardt U and Paul H 1993 *Phys. Rev.* A **48** 4598
[19] Ali S T and Prugovečki E 1977 *J. Math. Phys.* **18** 219
[20] Appleby D M 1999 *Int. J. Theor. Phys.* **38** 807
[21] Cohen L 1966 *J. Math. Phys.* **7** 781
[22] Prugovečki E 1978 *Ann. Phys., NY* **110** 102
[23] O'Connell R F and Wigner E P 1981 *Phys. Lett.* A **85** 121
[24] Weyl H 1927 *Z. Phys.* **46** 1
[25] Mehta C L and Sudarshan E C G 1965 *Phys. Rev.* **138** B274

[26] Appleby D M 1999 *J. Mod. Opt.* **46** 825
[27] Groenewold H J 1946 *Physica* **12** 405
[28] Moyal J E 1949 *Camb. Phil. Soc.* **45** 99
[29] Baker G A 1958 *Phys. Rev.* **109** 2198
Imre K, Ozizmir E, Rosenbaum M and Zweifel P F 1967 *J. Math. Phys.* **8** 1097
[30] Omnès R 1994 *The Interpretation of Quantum Mechanics* (Princeton NJ: Princeton University Press)
[31] Reed M and Simon B 1980 *Methods of Modern Mathematical Physics* vol 1–4 (New York: Academic)
[32] Grauert H and Fritzsche K 1976 *Several Complex Variables (Graduate Texts in Mathematics no 38)* (New York: Springer)
[33] Davidović D M and Lalović D 1993 *J. Phys. A: Math. Gen.* **26** 5099