New Journal of Physics
The open-access journal for physics

Quantum Fourier transform, Heisenberg groups and quasi-probability distributions

Manas K Patra[1] **and Samuel L Braunstein**[1]
Department of Computer Science, University of York, York YO10 5DD, UK
E-mail: manas@cs.york.ac.uk and schmuel@cs.york.ac.uk

New Journal of Physics **13** (2011) 063013 (36pp)
Received 22 November 2010
Published 9 June 2011
Online at http://www.njp.org/
doi:10.1088/1367-2630/13/6/063013

Abstract. This paper aims to explore the inherent connection between Heisenberg groups, quantum Fourier transform (QFT) and (quasi-probability) distribution functions. Distribution functions for continuous and finite quantum systems are examined from three perspectives and all of them lead to Weyl–Gabor–Heisenberg groups. The QFT appears as the intertwining operator of two equivalent representations arising out of an automorphism of the group. Distribution functions correspond to certain distinguished sets in the group algebra. The marginal properties of a particular class of distribution functions (Wigner distributions) arise from a class of automorphisms of the group algebra of the Heisenberg group. We then study the reconstruction of the Wigner function from the marginal distributions via inverse Radon transform giving explicit formulae. We consider some applications of our approach to quantum information processing and quantum process tomography.

[1] Authors to whom any correspondence should be addressed.

Contents

1. Introduction

Quasi-probability distribution functions (or simply distribution functions) on a quantum system provide an alternative and equivalent description of quantum states. We will discuss three possible approaches to distribution functions. The first approach is essentially Wigner's original approach [1] and it attempts to give a 'phase-space' description of quantum states. The state of a quantum system determines the probability distributions of its observables. It is possible to completely specify the state by giving the distributions of functions of a pair of *conjugate* nondegenerate observables $\hat{f}$ and $\hat{g}$. The most well-known example is the position–momentum pair. Thus corresponding to a (mixed) state ρ, we associate a *real* function $W(p, q : \rho)$ of 'c-number' variables that have the same information content as the state and give the correct marginals. Now for conjugate observables, the expectation values of the functions $\phi(\hat{f}, \hat{g})$ will specify the state completely (ignoring the questions of operator ordering). In classical probability theory, these expectation values are generated by the *characteristic function*. The characteristic function of (classical) random variables $X_1, X_2, \ldots, X_n$ with joint probability distribution $F(x_1, \ldots, x_n) \equiv F(\mathbf{x})$ is given by [2]

$$\widetilde{F}(\mathbf{t}) = \int \mathrm{e}^{\mathrm{i}\mathbf{t}\cdot\mathbf{x}}\, \mathrm{d}F(\mathbf{x}),$$

where $\mathbf{t}\cdot\mathbf{x}$ is the usual Euclidean scalar product of two real vectors $\mathbf{x} = (x_1, \ldots, x_n)^{\mathrm{T}}$ and $\mathbf{t} = (t_1, \ldots, t_n)^{\mathrm{T}}$, where T denotes transpose. If the probability distribution is given by a probability density $p(\mathrm{x})$, then the characteristic function is simply the Fourier transform of $p(\mathrm{x})$. Another way of viewing the characteristic function is to note that $\tilde{F}(\mathbf{t}) = \langle \mathrm{e}^{\mathrm{i}\mathbf{t}\cdot\mathbf{X}} \rangle$, the expectation value of the complex random variable $\mathrm{e}^{\mathrm{i}\mathbf{t}\cdot\mathbf{X}}$. Then assuming the existence of a probability density, it is given by the inverse Fourier transformation

$$p(\mathbf{x}) = \int \mathrm{e}^{-\mathrm{i}\mathbf{t}\cdot\mathbf{x}} \langle \mathrm{e}^{\mathrm{i}\mathbf{t}\cdot\mathbf{X}} \rangle \, \mathrm{d}\mathbf{t}. \tag{1}$$

In the form (1), it is suitable for a 'quantum' extension [3]. Of course, in the quantum case, the observables (or quantum random variables) X_i will *not* commute in general and we have the problem of interpreting the function p as a joint probability distribution. However, for a set of *compatible* or commuting observables, a joint distribution is unambiguously defined. For incompatible observables, we may take (1) as the definition of joint probability distribution. The values that are obtained on joint measurement of these observables constitute the joint spectrum that, in general, may have both continuous and discrete segments. In the finite-dimensional case, we have a finite spectrum and hence the integral has to be replaced by a sum. But there are problems in interpreting this as a function in a classical phase space [4]. Alternatively, we can work with *finite* Fourier transform. Such a transform is defined over a finite abelian group. From the structure of such groups, we may focus on the group $Z_N = \mathbb{Z}/N\mathbb{Z}$, the additive group of integers modulo N. We can thus take our 'configuration space' to be Z_N. This in turn forces us to consider observables that take 'values' in the set $\{0, 1, \ldots, N-1\}$, where N must now be identified with the dimension of the Hilbert space. Since we are concerned only with probability distributions of the possible outcomes, this is really not a restriction. Operationally, we can always calibrate our instruments to yield these outcomes. The 'position' and 'momentum' variables are both identified with Z_N and the corresponding unitary representations, respectively, act multiplicatively and additively on the 'position space'. Hence we have two representations of Z_N, but they do not constitute a representation of the 'phase space' $Z_N \times Z_N$ because the latter is a commutative group. The simplest possible *noncommutative* extension is a central extension [5]. After some restrictions due to finiteness of the dimension, we obtain a Heisenberg group.

Now let us approach the problem of the distribution function from another perspective. The state of the system, ρ, is a positive semi-definite operator with trace 1. In infinite dimensions, this belongs to a special class called *trace class* operators. In finite dimensions, every operator is clearly trace class. Trace class operators admit a Hilbert space structure. Thus, for two such operators A and B define $(A, B) = \mathrm{Tr}(A^\dagger B)$. In finite dimensions, this introduces the familiar Frobenius or Hilbert–Schmidt norm. If we restrict ourselves to Hermitian operators, we obtain a real Hilbert space $\mathcal{K}$. Pick an orthonormal basis $\{A_i : i = 1, 2, \ldots\}$ in $\mathcal{K}$. We can write the state

$$\rho = \sum_i W_i(\rho) A_i \quad \text{with} \quad W_i(\rho) = \mathrm{Tr}(\rho A_i).$$

Clearly, W_i are real. If we also demand that $\sum_i W_i = 1$, then we have a quasi-probability distribution over the index set $\mathcal{I}$ ($i \in \mathcal{I}$). We want this index set to have a classical interpretation and a natural choice is the *phase space*. Then, $i = (x, z)$ is a *pair* of indices[2]. Henceforth, we assume this and write $A(x, z)$ instead of i. The works [6, 7] follow this approach to distribution function (see also [8] for a review in the finite-dimensional case). Thus the choice of distribution is equivalent to the choice of special bases. The operators $A(x, z)$ are called phase point operators. Actually, they have to satisfy some extra conditions. We will see that the phase-point operators correspond to certain sets (called Wigner sets) in the group algebra of the Heisenberg group.

There is yet another view of distribution function that has its origins in signal analysis. A signal may be represented in the time domain as $f(t)$ or in the *frequency* domain as $\tilde{f}(\omega)$ (Fourier transform). Thus, we represent the signal in terms of 'elementary' harmonic

[2] We often use z in place of p for the momentum variable to avoid confusion with probability density. Further, x and z can be vectors.

signals and the coefficients give the representation in the frequency domain. But it can also be represented by other elementary nonharmonic signals with minimum uncertainty. This was Gabor's seminal idea [9]. Unlike harmonic signals, Gabor's elementary signals are localized in time *and* frequency domains. This joint time–frequency domain is the analogue of phase space. How do we generate these elementary signals? Starting with a 'reasonable' initial signal, say a Gaussian function in the time domain, we apply a sequence of two operators, multiplication *and* translation, which are 'diagonal' in the time and frequency domains, respectively. The resulting sequence of functions are used to represent an arbitrary signal. Now to represent a vector in some space, we do not need a basis; any set of vectors that *span* the space will do. In finite-dimensional Hilbert space, such sets define *frames* [10]. An example of such overcomplete sets is the set of coherent states in quantum optics. The frame-theoretic approach to distribution functions was recently proposed in [11]. We will not go into the details of frame theory but mention that frames are a generalization of orthonormal bases in Hilbert space. In this context, Gabor's elementary functions constitute the Gabor–Weyl–Heisenberg (GWH) frames. Most distribution functions are examples of such frames, although there are some exceptions. The GWH frames are generated by applying a sequence of translation and multiplication operators to (continuous) signals creating a function in the time–frequency domain (the phase space!). These operators generate a discrete Heisenberg group.

Finally, we come to another significant property of the Wigner function, a particularly important distribution function. Let $W(x, p)$ be the continuous Wigner function with x and p representing the classical position and momentum variables. Then the marginals $\sum_x W(x, p)$ and $\sum_p W(x, p)$ are *probability* distributions of the *quantum* momentum and position operators, respectively. Further, if we sum $W(x, p)$ along some line $ax + bp = 0$, then the resulting function is the probability distribution of a quantum operator 'orthogonal' to $c\hat{x} + \mathrm{d}\hat{p}$, where (c, d) is a vector orthogonal to (a, b) in the x–p plane. We will make these definitions precise later. Call this the Radon property. This is an important property and can be used to invert the transform. We will see that the Radon property is related to the transformation properties of the distribution function under some automorphisms of the Heisenberg group. We note that the Radon property is very useful in reconstructing states and processes. We prove a general Radon property that gives us a lot of freedom in our choice of possible measurements.

The brief (and incomplete) survey of the approaches to distribution functions in the preceding paragraphs indicates that much work has been performed in this area[3]. In addition to their theoretical significance, distribution functions have applications in state tomography [13], statistical mechanics and quantum optics [14]. It is also intimately connected with the theory of coherent states. The GBH-type operators after complexification and some algebra give rise to the familiar *displacement* operators. The coherent states are the orbits of the Weyl–Heisenberg group (henceforth only the Heisenberg group) [15]. In this work, we mainly focus on the finite-dimensional case. This case presents some difficulties absent in the continuous case. The finite-dimensional case is also significant for quantum information processing (QIP) [16, 17].

We have considered $Z_N \times Z_N$ as the basic model of finite 'phase space'. In the literature, other phase spaces have been considered (see [11] for a discussion and references). The intrinsic structure of these phase spaces may have an interesting bearing on the corresponding distribution functions. In particular, some authors have considered finite field F_N with N

[3] See [11, 12] and [8] for long lists of references relevant to the present work.

elements instead of Z_N (see e.g. [18]). This is only possible if $N = p^n$ is a power of some prime p. Now Z_p and F_p coincide. In general, the additive groups of F_{p^n} and $Z_p \times \cdots \times Z_p$ (n factors) are isomorphic. We can consider the Heisenberg groups over the latter (by central extension). We do not follow this here, as the paper is already quite lengthy. However, note the analogy between what the authors in [18] call quantum nets and Wigner sets defined in this paper. More precisely, quantum nets correspond to those Wigner sets that are permuted by the action of the automorphism group $SL(2, Z_N)$. The paper is quite self-contained. We give most of the proofs. Some of the results are known but they were derived using different approaches. Let us first note some of the main contributions of the present work.

1. We use the Heisenberg groups as the basic approach to distribution functions. As we have seen in the preceding paragraphs, this is *the* unifying thread tying the different approaches and perspectives on the distribution function.
2. The Heisenberg group has been used in the literature in the context of distribution functions. But here we use discrete Heisenberg groups and a family of finite quotient groups thereof. We define these groups abstractly in terms of generators and relations. Thus, we can consider different representations (irreducible and reducible) and operators between representations.
3. Our treatment of the Heisenberg groups is defined in terms of generators and relations. This makes the computations and proofs easier. Further, we do not need the language of (pseudo) phase spaces.
4. We show that the existence of distribution functions is equivalent to certain sets in the group *algebras* of the Heisenberg groups. This provides us with powerful methods of representation theory. We list some of the outcomes by the use of these methods:
   (a) The analysis of marginal distributions becomes transparent. They correspond to an invariance up to permutations under certain groups of automorphisms (e.g. $SL(2, Z_N)$) of the Heisenberg groups.
   (b) Demanding this invariance, we obtain unique distribution functions in odd dimensions.
   (c) We also infer that it is impossible to retain full invariance *and* other properties such as hermiticity and linear independence in even dimension. We therefore have three possible strategies: (i) drop the requirement of invariance, (ii) drop the requirement of independence or hermiticity or (iii) demand invariance under a smaller set of transformations. We discuss all three and give some alternative candidates for distribution functions in even dimensions.
   (d) Our analysis via the automorphism groups obviates the need for ad hoc hypothesis and guess work.

   We note again that although some results mentioned in (a), (b) and parts of (c) are known, our approach via automorphisms is different.
5. We give explicit formulae in most cases. The close connection with finite Radon transform is made clear. It is used to derive the formulae for state reconstruction. The inversion formulae in the case where dimension $= 2^n$ appear to be new.
6. We explore applications to quantum computing and information. The fact that the Weyl–Heisenberg groups describe the kinematics of quantum systems is known [19, 20]. We show that the *dynamics* are described by (unitary) automorphisms of the *group algebra*.

The case of unitary automorphisms of the *group* itself is analysed in [21]. The latter correspond to the Clifford group, and to go beyond it ('nonclassical' dynamics) we have to consider the group algebra. We also illustrate the utility of the Heisenberg groups in quantum circuits. We present an application to quantum process tomography.

This paper is organized as follows. In section 2, we discuss the quantum Fourier operators. The quantum Fourier transform (QFT) is another form of finite Fourier transform [22]. Consider the two representations of Z_N acting as multiplication and translation, respectively. The QFT connects the two. The generators of the two representations give us the basic operators: Z and X. We can consider these as the finite-dimensional analogue of unitary operators generated by 'position' and 'momentum', respectively. In section 3, we review (continuous) quasi-probability distribution functions or simply distribution functions. The continuous distribution functions are somewhat easier to deal with because the 'infinitesimal' generators satisfy simple commutation rules (the Heisenberg relations). In section 4, we come to one of our main themes: the finite distribution functions. We list a set of properties, satisfied by the continuous Wigner function, and demand that *any* distribution function must satisfy them. In particular, we give examples of discrete Wigner functions. Here, we encounter the difficulties when the dimension is even. We also derive explicit formulae for the phase-point operators. The odd-dimensional case (apart from some constants) is essentially the same as Wootters' [7] operators in prime dimension.

In section 5, we study the Heisenberg groups. We start with the continuous version as has been studied well in connection with Fourier transforms [23]. We then look at discrete and finite Heisenberg groups, their structure, representation and automorphisms, all of which play an important role in our study of the finite distribution functions. We show that there is a one-to-one correspondence between distribution functions in dimension N and certain sets $\{A(x, z)\}$ (called Wigner sets) in the group algebra of the Heisenberg group $\mathbf{H}$ in that dimension. The representations of these sets are the phase-point operators. A slight generalization of the Wigner sets may be used to define Weyl–Heisenberg frames. The group $SL(2, Z_N)$ of 2×2 matrices in Z_N with determinant 1 induces automorphisms on the Heisenberg group $\mathbf{H}$. Thus, for each $M \in SL(2, Z_N)$, we define an automorphism σ_M of $\mathbf{H}$. These automorphisms, in turn, determine the marginal properties of the Wigner functions. Thus, if $W(x, z) \equiv W(\boldsymbol{\zeta})$ is a distribution function, then the functions

$$Q(z) = \sum_x W(M^{-1}\boldsymbol{\zeta}) \quad \text{and} \quad P(x) = \sum_z W(M^{-1}\boldsymbol{\zeta})$$

are the marginals. $Q(z)$ is called a simple marginal if it is the probability distribution (in the given state) of an observable $\hat{z}_M$ defined by $\mathrm{e}^{\mathrm{i}\hat{z}_M} = \sigma_M(Z)$. A similar definition can be given for $P(x)$. We show the necessary and sufficient condition for the existence of simple marginals for all members of $SL(2, Z_N)$. Thus the requirement that all marginals be simple determine the distribution function up to isomorphism. In the case of odd dimensions for the Wigner function, all marginals are simple. This can be neatly expressed as follows. Let $A_M(x, z) = A(M^{-1}x, M^{-1}z)$. Then $\{A_M(x, z)\}$ is also a Wigner set. An analogous result is called Clifford invariance in [24]. This is not true in even dimensions. We investigate three alternatives by weakening our requirements. First, we do not demand that the phase-point operators form a basis. Now they constitute a *frame*. This is the most common approach (see, for example, [16, 25]). The marginal conditions are simple but at the expense of losing orthogonality of bases. We show that this is similar to the case of spin-1/2 representation in the

sense that a complete 'rotation' does not preserve the values of the functions involved. More precisely, we obtain functions that are not periodic on Z_N. However, they have period $2N$. So, we go over to $Z_{2N} \times Z_{2N}$ as the phase space. Next we drop the requirement that the marginals be simple in the above sense. It is still possible to compute the marginals in terms of the probability distribution of the observable $\hat{z}_M$. We indicate how explicit formulae can be derived to compute this. Finally, since it is not possible to satisfy simple marginal conditions on all of $SL(2, Z_N)$, we consider certain subsets adequate for inversion, i.e. computing the state from the marginal data. We give such a subset in dimension $N = 2^K$. A similar construction from a different perspective was done in [26], but we present our formulae in an explicit functional form.

In section 5.1, we explore the fact that the definition of marginals is a Radon transform of W in the sense of [27–29]. We then give the inversion formulae in several cases. The inversion formulae for odd dimensions have been given in [13]. Our derivation, however, is more general and applicable to *any* finite distribution function. The important point is that we can invert these transforms and recover the Wigner function and hence the quantum state. In the next section, we discuss some applications to QIP. We provide some simple relations between standard quantum gates and operators representing the Heisenberg groups that will prove useful for implementing the state and process determination schemes using phase-point operators. We give formulae for quantum process tomography using phase-point operators. In the final section, we discuss other possible applications and directions for future work.

2. Quantum Fourier operators

Let $G = \mathbb{Z}/N\mathbb{Z}$ be the additive group of integers modulo-N. There are two obvious representations of G on an N-dimensional Hilbert space H. Let g be a generator of G. Suppose that $\phi : G \to \mathcal{U}(H)$ is a faithful representation of G by unitary operators, where $\mathcal{U}(H)$ is the set of unitary operators on H. If $\phi(g) = Z$, then we must have $Z^N = 1$ since the order of G is N and the representation is faithful. The eigenvalues of Z are Nth roots of unity. Let $\{|i\rangle : i = 0, \ldots, N-1\}$ be the corresponding eigenvectors such that $Z|i\rangle = \omega^i |i\rangle$, where ω is a primitive Nth root. Call it the *computational basis* $\mathcal{B}_c$. There is another representation ϕ' of G defined by $\phi'(g) = X$, where $X|i\rangle = |i + 1 \pmod N\rangle$. We can think of ϕ as the multiplicative and ϕ' as the additive representations. Z and X represent multiplication and translation operators, respectively. Clearly, X is unitary and there is basis $\mathcal{B}_f$ in which it is diagonal. The QFT is the unitary map connecting the two representations taking $\mathcal{B}_c \to \mathcal{B}_f$. The eigenvalues of X are also roots of unity as $X^N = I$. Since ϕ' is also faithful, the diagonalization of X yields Z fixing the ordering. Hence, there exists a unitary operator Ω such that

$$\Omega^\dagger X \Omega = Z. \tag{2}$$

The explicit form of Ω in the computational basis is easy to compute. Thus, if $\alpha = \sum_i x_i |i\rangle$ is an eigenvector, then $X\alpha = u\alpha$ implies $x_0 = ux_1, \ldots, x_{N-2} = ux_{N-1}$, and $x_{N-1} = ux_0$. This yields, after normalization,

$$(\Omega)_{ij} = \frac{1}{\sqrt{N}} \omega^{-ij}. \tag{3}$$

So the QFT is the map

$$|k\rangle \to \frac{1}{\sqrt{N}} \sum_j \omega^{-jk} |j\rangle = \frac{1}{\sqrt{N}} \sum_j \mathrm{e}^{-2\pi i jk/N} |j\rangle .$$

We note that we follow the convention of mathematicians in the definition of discrete or finite Fourier transform. In the quantum information literature, the usual definition is with a positive sign in the exponent, which is our *inverse* Fourier transform. Now X and Z can be expressed as

$$X = \mathrm{e}^{\mathrm{i}\hat{x}} \quad \text{and} \quad Z = \mathrm{e}^{\mathrm{i}\hat{z}}, \tag{4}$$

where $\hat{x}$ and $\hat{z}$ are the Hermitian generators of the respective unitary rotation. Moreover, their eigenstates are (discrete) Fourier transforms of each other. This is reminiscent of position and momentum observables that also have the property that their (generalized) eigenstates are (continuous) Fourier transforms of each other. We may therefore regard the observables $\hat{x}$ and $\hat{z}$ as conjugate 'dynamical variables'. This terminology is further justified by the following observation, which is crucial for our calculations of quasi-probability distributions,

$$XZ = \omega ZX. \tag{5}$$

This is most easily derived by applying both sides to vectors in the computational basis $\mathcal{B}_c$. We observe that the unitary operators $\mathrm{e}^{\mathrm{i}a\hat{p}}$ and $\mathrm{e}^{\mathrm{i}b\hat{q}}$ corresponding to translations in (continuous) position and and momentum space, respectively, obey a similar relation.

Suppose now that H is a product space, that is, $H = \otimes^m H_d$, where H_d is a d-dimensional space and $N = d^m$. As a simple application of the basic relation (2), we show that the Fourier transform of product states in the computational basis are also product states and generalize a computationally useful formula.

Lemma 1. *If the basis $\mathcal{B}_c$, the eigenvectors of Z, consists of m-fold product states, then their Fourier transforms are also product states given by*

$$\Omega(|j\rangle) = \left(\sum_{r=0}^{d-1} \omega^{-jd^{m-1}r} |r\rangle\right) \otimes \cdots \otimes \left(\sum_{r=0}^{d-1} \omega^{-jdr} |r\rangle\right) \otimes \left(\sum_{r=0}^{d-1} \omega^{-jr} |r\rangle\right)$$

Proof. Observe that there is an implicit ordering of the product states. Thus, if $j = \sum_{r=0}^{m-1} d^r j_r$ is the representation of a positive integer $0 \leqslant j \leqslant d^m - 1$, then the state $|j\rangle = |j_{m-1}\rangle \otimes \cdots \otimes |j_1\rangle \otimes |j_0\rangle \equiv |j_{m-1}\rangle \cdots |j_1\rangle|j_0\rangle \equiv |j_{m-1} \cdots j_0\rangle$, where we suppress the tensor product symbol in the last two relations. Further, we write $|0\rangle$ for $|0 \cdots 0\rangle$.

From the definition of QFT,

$$\begin{aligned}\langle k_{m-1} \cdots k_0|\Omega|j_{m-1} \cdots j_0\rangle &= \langle 0|X^{-k}\Omega|j_{m-1} \cdots j_0\rangle \\ &= \langle 0|\Omega\Omega^\dagger X^{-k}\Omega|j_{m-1} \cdots j_0\rangle = \langle 0|\Omega Z^{-k}|j_{m-1} \cdots j_0\rangle \\ &= \omega^{-kj}\langle 0|\Omega|j_{m-1} \cdots j_0\rangle = \frac{\omega^{-kj}}{\sqrt{N}}.\end{aligned}$$

A direct computation shows that for the state

$$|\psi_j\rangle = \frac{1}{\sqrt{N}}\left(\sum_{r=0}^{d-1} \omega^{-jd^{m-1}r} |r\rangle\right) \otimes \cdots \otimes \left(\sum_{r=0}^{d-1} \omega^{-jdr} |r\rangle\right) \otimes \left(\sum_{r=0}^{d-1} \omega^{-jr} |r\rangle\right),$$

$\langle k|\psi_j\rangle = \omega^{-kj}$. Since this is true for all $0 \leqslant k \leqslant N-1$, $\Omega|j\rangle = |\psi_j\rangle$. □

3. Distribution functions in quantum systems

One of the motivating factors for distribution functions in Wigner's work [1] was the construction of a quantum analogue of the Liouville density in classical phase space. Following this approach, suppose that we want a 'joint' distribution function of the operators X and Z. More precisely, we seek Hermitian operators $\hat{x}$ and and $\hat{z}$ such that

$$X = \mathrm{e}^{\mathrm{i}\hat{x}} \quad \text{and} \quad Z = \mathrm{e}^{\mathrm{i}\hat{z}} \tag{6}$$

and then try to find distribution functions associated with the observables $\hat{x}$ and $\hat{z}$. We will do our computations in the computational basis $\mathcal{B}_{\mathrm{c}}$ in which Z is diagonal. It is easy to find $\hat{z}$. Thus

$$\hat{z} = \frac{2\pi}{N} \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ & & & & \\ \vdots & & & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & N-1 \end{pmatrix}. \tag{7}$$

Of course, $\hat{z}$ is only determined modulo $2\pi k$. From (2) and (3), we have

$$(\hat{x})_{ij} = (\Omega^{\dagger}\hat{z}\Omega)_{ij} = \sum_{k=0}^{N-1} k\omega^{k(i-j)} = \begin{cases} \pi(N-1) & \text{if} \quad i = j, \\ \dfrac{2\pi}{N}\dfrac{N-1}{\omega^{i-j}-1}, & i \neq j. \end{cases} \tag{8}$$

These entries of the matrix $\hat{x}$ imply that it is a Hermitian *circulant* matrix. But a general linear combination $u\hat{x} + v\hat{z}$ is not a circulant but a Toeplitz matrix. There are efficient algorithms for finding the eigenvalues and eigenvectors of such matrices. So, in principle, we can compute expressions such as $\langle \mathrm{e}^{\mathrm{i}(u\hat{x}+v\hat{z})} \rangle$ for real or integer u, v using the standard diagonalization procedure. It is feasible to find analytic expressions, however, in low dimensions only. We will tackle the problem by a different approach. Let us briefly review the continuous case first.

3.1. The Wigner distribution

In the case of canonically conjugate variables, such as position and momentum, a number of quasi-probability distributions are possible, each corresponding to an operator ordering prescription. This is facilitated by the fundamental commutation relation

$$[\hat{q}, \hat{p}] = \mathrm{i}\hbar$$

between the position and momentum operators. Taking traces, it is clear that such a relation is not possible in finite dimensions. So in finite dimension it is not clear how to prescribe ordering of operators. Moreover, there is some ambiguity in defining Hermitian generators themselves. For example, for integers a and b, $\hat{x}' = \hat{x} + 2\pi a I$ and $\hat{z}' = \hat{z} + 2\pi b I$ are also infinitesimal generators for X and Z, respectively, but their linear combinations give rise to a different set of unitary operators. The problem of nonuniqueness is essentially the same as the one that arises in defining roots and logarithms of complex numbers. Hence, we restrict ourselves to the principal branch of the logarithm, as evident in the definition of $\hat{x}$ and $\hat{z}$.

3.1.1. Continuous Wigner distribution. The inversion formula of a characteristic function of classical probability is different for continuous and discrete probability distributions. In finite-dimensional quantum systems or more generally in the discrete part of the spectrum of a quantum observable, we should use a formula analogous to that for discrete distributions [3]. But this gives a quasi-probability distribution that may *not* have the desired properties [4]. The problem seems to be rooted in the noncommutativity of quantum observables. The continuous Wigner distribution is defined by

$$W_{\mathrm{c}}(x, z) \equiv \frac{1}{(2\pi)^2} \iint \langle \mathrm{e}^{\mathrm{i}(u\hat{x}+v\hat{z})} \rangle \, \mathrm{e}^{-\mathrm{i}(ux+vz)} \, \mathrm{d}u \, \mathrm{d}v. \tag{9}$$

In this equation and the rest of the paper, unless the limits are explicitly stated, the real integrals are over the whole real line. Further, we use the notation $\mathbf{r} = (x, y)^T$ for a real vector in two dimensions. The following theorem gives some of the important properties of the continuous Wigner distribution. First, we make the dependence on the state ρ (mixed state, in general) explicit when necessary: $W_{\mathrm{c}}(x, z : \rho)$. We give a simple proof of well-known results in the appendix.

Theorem 1. *The function $W_c(x, z : \rho)$ is real. Moreover, it gives the correct marginal distributions.*

$$\int_a^b \mathrm{d}z \int W_c(x, z : \rho)\mathrm{d}x = \int_a^b \langle z|\rho|z\rangle \, \mathrm{d}z, \tag{10}$$

where $|z\rangle$ are generalized eigenvectors of $\hat{z}$. We have a similar relation for the other marginal. We also have the following results on general marginal distribution. Let R be an orthogonal matrix of order 2 representing a rotation. Let

$$\mathbf{r}' = \begin{pmatrix} x' \\ z' \end{pmatrix} = R\mathbf{r} \quad \text{with} \quad R = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}.$$

Similarly, we define orthogonal transformation in the 'noncommutative' space. If

$$\begin{pmatrix} \hat{x}' \\ \hat{z}' \end{pmatrix} = R \begin{pmatrix} \hat{x} \\ \hat{z} \end{pmatrix} = \begin{pmatrix} \cos\theta\hat{x} + \sin\theta\hat{z} \\ -\sin\theta\hat{x} + \cos\theta\hat{z} \end{pmatrix},$$

then

$$\int_a^b \mathrm{d}z' \int W_c(x, z : \rho) \, \mathrm{d}x' = \int_a^b \mathrm{d}z' \, \langle z'|\rho|z'\rangle \, , \tag{11}$$

where $|z'\rangle$ are generalized eigenvectors of $\hat{z}'$ and the variables x and z are considered as functions of x', z'. A similar result holds for the conjugate observable $\hat{x}$.

The continuous version of the distribution function of discrete observables is problematic. First, we want the marginals to resemble classical discrete probability distributions, so we have delta functions at the isolated points. To justify the latter, we have to integrate over some domain of the continuous variable and this causes some problems in interpreting these as probability distributions. Some authors have attempted to tackle these problems by focusing on continuous families of discrete observables, such as spin direction. These approaches seem somewhat unnatural to us. Discrete distributions must be characterized by a discrete measure (e.g. the counting measure) and thus the integrals must be replaced by sums. In particular, for finite

systems, we must have finite sums. This is the avenue we will explore in this paper. Finally, let us mention an important point. The Wigner distribution and some other related probability and quasi-probability distributions are sometimes interpreted as *joint* probability distributions of incompatible observables. Clearly, any measurement of such a distribution must give unsharp values of these observables, consistent with the uncertainty principle. The Arthrus–Kelley scheme [30, 31] is an example. For discrete observables, the concept of joint distribution of noncommuting observables is difficult even for fuzzy measurements.

4. Discrete quasi-probability distributions

The Wigner and other distribution functions are an alternative to the density matrix formulation of quantum theory and are given by the distribution function $W(\mathbf{y}:\rho)$ with $\mathbf{y}$ representing classical parameters. Expectation values of any quantum mechanical quantity that can be computed in a given state ρ can be computed from $W(\mathbf{y}:\rho)$. Hence we have a correspondence between quantum observables and 'classical' observables along with an ordering prescription. Since the density matrix provides a maximal description of a quantum system, so does $W(\mathbf{y}:\rho)$. Thus we have an alternative semiclassical picture. In some situations, the latter may be easier to determine experimentally. In any case, such quasi-probability distributions provide a useful tool for visualization.

4.1. Properties of distribution functions

Let us make precise the requirements we impose on distribution functions. Let $W(\mathbf{y}:\rho)$ be a distribution function associated with a quantum state ρ of a quantum system S and $\mathbf{y}$ is a real vector representing a finite set of 'classical' parameters. Let H be the system Hilbert space and $\mathcal{S}(H)$ the set of states, i.e., the *convex* set of normalized positive trace-class operators.

R1. $W(\mathbf{y}:\rho)$ is a continuous real function on $\mathcal{S}(H)$ that preserves convex combinations: if $\rho_1, \rho_2 \in \mathcal{S}(H)$ and $0 \leqslant s \leqslant 1$, then

$$W(\mathbf{y}: s\rho_1 + (1-s)\rho_2) = sW(\mathbf{y}:\rho_1) + (1-s)W(\mathbf{y}:\rho_2).$$

It is nondegenerate in the sense that at no point in phase space is $W(\mathbf{y}:\rho)$ identically 0 for all ρ.

R2. For two states ρ and ρ',

$$\mathrm{Tr}(\rho\rho') = K\int W(\mathbf{y}:\rho)W(\mathbf{y}:\rho')\mathrm{d}\mathbf{y}. \tag{12}$$

Part of the above requirement is that we define the appropriate measure $\mathrm{d}\mathbf{y}$, which also fixes the constant K. The constant $K = 2\pi\hbar$ for continuous systems and $K = N$ for a finite system of dimension N. This constant equals the volume of a 'phase-space cell'. With respect to this measure, we also demand the normalization condition

$$\int W(\mathbf{y}:\rho)\mathrm{d}\mathbf{y} = 1.$$

Note that this is a nontrivial requirement as this implies that the left-hand side of the above equation must be independent of the quantum state.

R3. For any observable A on S, there is a real function $\tilde{A}(\mathbf{y})$ such that the expectation value (in state ρ)

$$\langle A\rangle = \mathrm{Tr}(\rho A) = \int W(\mathbf{y}:\rho)\tilde{A}(\mathbf{y})\mathrm{d}\mathbf{y}. \tag{13}$$

The first requirement is that W must be real. If we try to impose non-negativity, however, it becomes too stringent. As $W(\mathbf{y}:\rho)$ is convex linear on states in the finite-dimensional spaces, it has a unique extension to a liner functional (for fixed $\mathbf{y}$) on $\mathcal{K}(H)$ the set of bounded Hermitian operators (observables). In infinite dimensions, we need some delicate continuity arguments. Henceforth, we will assume linearity of $W(\mathbf{y}:\rho)$. In these specifications for the distribution function, we have not mentioned marginals. We will discuss them shortly. What are the characteristics of the parametric vector $\mathbf{y}$? If it is to be somehow identified with generators of classical observables, its dimension must be related to degrees of freedom. The third item in the list gives a clue. A physical system, whether classical or quantum, is completely characterized by the set of observables $\mathcal{O}$. Often $\mathcal{O}$ has more structure; in particular, it is an algebra. The difference between quantum and classical algebra of observables is that the former is *noncommutative*. These algebras have minimal sets of generators. For example, the observable algebra of a classical system with N degrees of freedom is generated by generalized coordinates $\{q_i : i = 1,\ldots,N\}$ and the conjugate momenta $\{p_i : i = 1,\ldots,N\}$. The corresponding quantum algebra is also generated by the *operators* $\hat{q}_i$ and $\hat{p}_i$, which do not commute. In the finite-dimensional case, we have no classical analogue. But we will be guided by this example. We have already discussed the close analogy between the finite-dimensional unitary operators X and Z and the continuous operators $\mathrm{e}^{\mathrm{i}\hat{p}}$ and $\mathrm{e}^{\mathrm{i}\hat{q}}$. We show next that $\hat{x}$ and $\hat{z}$ are actually generators of the *complex* algebra of observables in the appropriate dimension.

Proposition 1. *Let the dimension of the system Hilbert space be N and $\hat{x}$ and $\hat{z}$ be as given in (8) and (7), respectively. The completion of the complex algebra generated by $\hat{x}$ and $\hat{z}$ equals $M_n(\mathbb{C})$, the algebra of complex matrices of order N.*

Proof. The completion of algebra means that we include the limits of convergent sequences. In particular, $X = \mathrm{e}^{\mathrm{i}\hat{x}}$ and $Z = \mathrm{e}^{\mathrm{i}\hat{z}}$ are in completion. We show that X and Z generate $M_n(\mathbb{C})$. Let $\omega = \mathrm{e}^{2\pi i/N}$ and $Z_k = \omega^{-k}Z$. It is easy to see that

$$(I + Z_k + Z_k^2 + \cdots + Z_k^{N-1})/N = D_k,$$

where $D_k(ij) = \delta_{ij}\delta_{jk}$ is the diagonal matrix with 1 in the kth row (and column) and 0s everywhere. We also see that $X^j D_k = E_{j+k,k}$, where E_{ij} are the elementary matrices with 1 in the ijth place and 0s everywhere else. Note that $j+k$ is to be considered mod N. Thus every elementary matrix is generated by $\hat{x}$ and $\hat{z}$. As the elementary matrices constitute a basis for $M_n(\mathbb{C})$, the proof is complete. □

We mention that the assertion of the proposition was essentially proved by Schwinger [20] in a different way. The continuous quasi-probability distribution functions can be written as

$$W(\mathbf{q},\mathbf{p}) = \int f(\mathbf{u},\mathbf{v})\langle \mathrm{e}^{\mathrm{i}\hat{\mathbf{q}}\cdot\mathbf{u}}\mathrm{e}^{\mathrm{i}\hat{\mathbf{p}}\cdot\mathbf{v}}\rangle \mathrm{e}^{-\mathrm{i}(\mathbf{u}\cdot\mathbf{q}+\mathbf{v}\cdot\mathbf{p})}\,\mathrm{d}\mathbf{u}\,\mathrm{d}\mathbf{v}. \tag{14}$$

Here, f is a scalar or c-number function that is usually interpreted as an operator ordering prescription. The Wigner distribution function is a special case corresponding to Weyl ordering. All of this is possible because of the simple commutation properties of the observables $\hat{q}_i$ and $\hat{p}_i$. We have observed that the unitary operators X and Z have multiplicative relations very similar to $e^{i\hat{p}}$ and $e^{i\hat{q}}$ (see equation (5)). This analogy extends further,

$$Z^a X^b = \omega^{ab} X^b Z^a; \quad e^{ia\hat{q}}\, e^{ib\hat{p}} = e^{iab}\, e^{ib\hat{p}}\, e^{ia\hat{q}}, \tag{15}$$

provided that a and b are integers. Of course, the second formula is valid for all real a and b but the first fails if *both* are noninteger. This provides another reason to construct a discrete version of distribution functions. Henceforth, we will restrict ourselves to finite-dimensional spaces mostly. Since the operators $\hat{x}$ and $\hat{z}$ can be used as generators, we will assume that the 'phase space' spanned by $\mathbf{y} = (x, z)$ is two-dimensional (2D). Now we can state the marginal conditions corresponding to the 'axes'.

R4. The quasi-probability distribution function $W(x, z)$ has marginal distributions that coincide with probability distributions of the *quantum* observables $\hat{x}$ and $\hat{z}$,

$$\sum_x W(x, z : \rho) = \delta_{xj} \mathrm{Tr}(|j\rangle\langle j|\,\rho) \quad \text{and} \quad \sum_z W(x, z : \rho) = \delta_{zj} \mathrm{Tr}(|\tilde{j}\rangle\langle\tilde{j}|\,\rho), \tag{16}$$

where $|\tilde{j}\rangle\langle\tilde{j}|$ (resp. $|\tilde{j}\rangle\langle\tilde{j}|$) are eigenvectors of z (resp. x) with eigenvalue $2\pi j/N$.

We seek a finite distribution function similar to the form (14) above. For a state ρ in a finite quantum system of dimension N, define

$$\begin{aligned} &W(x, z : \rho) = \textstyle\sum_{m,n=0}^{N-1} f(m, n)\langle X^m Z^n\rangle \omega^{-(mx+nz)} \\ &\text{with } \omega = e^{2\pi i/N} \quad \text{and} \quad 0 \leqslant j,\ k \leqslant N-1 \text{ integers.} \end{aligned} \tag{17}$$

Call the functions f in the above expression *ordering* functions. To compute the expectation values, we need the following matrix elements in computational basis,

$$\langle k|X^m Z^n|j\rangle = \begin{cases} \delta_{m,k-j}\omega^{jn} & \text{if} \quad k \geqslant j, \\ \delta_{m,N+k-j}\omega^{jn} & \text{if} \quad k < j. \end{cases} \tag{18}$$

To see the implications of the reality condition R1, it is sufficient to verify it for pure states. Hence, for $\rho = |\alpha\rangle\langle\alpha|$,

$$\begin{aligned} \overline{W(x, z : \rho)} &= \sum_{m,n=0}^{N-1} \overline{f(m, n)}\, \langle\alpha|Z^{-n}X^{-m}|\alpha\rangle\, \omega^{mx+nz} \\ &= \sum_{m,n=0}^{N-1} \overline{f(m, n)}\omega^{mn}\, \langle\alpha|X^{-m}Z^{-n}|\alpha\rangle\, \omega^{mx+nz} \\ &= \sum_{m,n=1}^{N} \overline{f(N-m, N-n)}\omega^{mn}\, \langle\alpha|X^m Z^n|\alpha\rangle\, \omega^{-(mx+nz)} \\ &= W(x, z : \rho). \end{aligned}$$

In the last step, we use $X^N = Z^N = I$. Since this must hold for all state vectors α, we have

$$\begin{aligned} &\overline{f(N-m,0)} = f(m,0), \quad \overline{f(0,N-n)} = f(0,n) \quad \text{and} \\ &\overline{f(N-m,N-n)}\omega^{mn} = f(m,n), \quad 1 < m,n < N. \end{aligned} \tag{19}$$

We will see later that the condition of nondegeneracy is automatically satisfied. Next, we consider R2. Let $\rho = \sum \rho_{jk} |j\rangle \langle j|k$ and $\rho' = \sum \rho'_{jk} |j\rangle \langle j|k$. Then using (18)

$$\begin{aligned} &\sum_{xz} W(x,z:\rho)W(x,z:\rho') \\ &\quad = \sum_{xz} \sum_{\substack{jmn \\ j'm'n'}} f(m,n)f(m',n')\rho_{j,j+m}\rho_{j',j'+m'}\omega^{jn+j'n'}\omega^{-((m+m')x+(n+n')z)} \\ &\quad = N^2 \sum_{jmn}\sum_{j'} f(m,n)f(N-m,N-n)\rho_{j,j+m}\rho_{j',j'+N-m}\omega^{(j-j')n} \\ &\quad = N^2 \sum_{jmn}\sum_{j'} f(m,n)\overline{f(m,n)}\rho_{j,j+m}\rho_{j',j'+N-m}\omega^{(j-j'+m)n}, \end{aligned} \tag{20}$$

where we have used (19) in the last step. According to R2, this should be equal to $\mathrm{Tr}(\rho\rho')/N = (\sum_{jk} \rho_{jk}\rho'_{kj})/N$ for all choices of density matrices ρ and ρ'. This is possible if $|f(m,n)|^2$ is a constant independent of m and n. A straightforward computation yields $|f(m,n)| = 1/N^2$. Setting $f(m,n) = g(m,n)/N^2$, we may write $g(m,n) = \omega^{\beta(m,n)}$. We now prove the existence and properties of distribution functions satisfying the conditions R1–R4.

Theorem 2. *For every density matrix ρ in an N-dimensional Hilbert space and $\omega = \mathrm{e}^{2\pi i/N}$, let*

$$W(x,z:\rho) = \sum_{m,n=0}^{N-1} f(m,n)\langle X^m Z^n\rangle \omega^{-(mx+nz)}.$$

Then there exist functions $f(m,n)$ such that the corresponding W satisfies R1, R2 and R4. Moreover, for any W satisfying these conditions, there are unique Hermitian operators $\hat{a}(x,z)$ such that the following hold,

$$W(x,z:\rho) = \mathrm{Tr}(\rho\hat{a}(x,z)) \quad \text{and} \quad \rho = N\sum_{xz} W(x,z:\rho)\hat{a}(x,z), \tag{21a}$$

$$\mathrm{Tr}(\hat{a}(x,z)\hat{a}(x',z')) = \frac{1}{N}\delta_{xx'}\delta_{zz'} \quad \text{and} \quad \sum_{xz}\hat{a}(x,z) = I. \tag{21b}$$

Given a Hermitian operator T, let $t(x,z) = \mathrm{Tr}(T\hat{a}(x,z))$, then

$$\langle T\rangle = \sum_{xz} W(x,z:\rho)t(x,z). \tag{22}$$

Thus R3 is also satisfied.

Proof. We have observed that functions $f(m,n) = \omega^{\beta(m,n)}/N^2$ satisfying relations (19) provide a distribution function $W(x,z:\rho)$ that satisfies conditions R1 and R2. To see the implications of condition R4 on marginals, we observe that

$$\begin{aligned}\sum_x W(x,z:\rho) &= \sum_{mn} f(m,n)\langle X^m Z^n\rangle \omega^{-nz} \sum_x \omega^{-mx} \\ &= \sum_{mn} f(m,n)\langle X^m Z^n\rangle \omega^{-nz}\delta_{m0} = N\sum_n f(0,n)\langle Z^n\rangle\omega^{-nz} \\ &= N\sum_{jkn} f(0,n)\rho_{jk}\,\langle k|Z^n|j\rangle\,\omega^{-nz} = N\sum_{jn} f(0,n)\rho_{jj}\omega^{(j-z)n}. \end{aligned} \tag{23}$$

For the last expression to be equal to $\mathrm{Tr}(\rho|z\rangle\langle z|)$, the probability of finding the system in an eigenstate of $\hat{z}$ with eigenvalue $2\pi z/N$, we must have $f(0,n) = 1/N^2$ for all n. Computing the trace in the Fourier transformed basis $|\tilde{j}\rangle = \Omega\,|j\rangle$, we conclude that the second condition in R3 yields $f(m,0) = 1/N^2$. Assuming $\beta(m,n)$ to be a real polynomial in m and n, we conclude that it must be of the form $\beta(m,n) = mn\alpha(m,n)$. More generally, we may take $\beta(m,n) = \gamma(m,n) + mn\alpha(m,n)$ with $\gamma(0,n) = \gamma(m,0) = 0$. With this choice of $\beta(m,n)$, the first set of equations in (19) are satisfied. The second set gives the following requirement on the function α,

$$\begin{aligned} &mn(\alpha(m,n)-1) + (N-m)(N-n)\alpha(N-m,N-n) + \gamma(m,n) \\ &\qquad +\gamma(N-m,N-m) = 0 \bmod N. \end{aligned} \tag{24}$$

Note that we do not require α or γ to be integer-valued or symmetric. There exist (real) functions satisfying equation (24) for all $0 \leqslant m, n \leqslant N-1$. Simple solutions to these equations are given below.

$$f_0(m,n) = \begin{cases} \dfrac{\omega^{mn(N+1)/2}}{N^2} & \text{if } N \text{ is odd,} \\ \nu_{mn}\dfrac{\omega^{mn/2}}{N^2} & N \text{ even,} \end{cases} \tag{25}$$

where ν_{mn} satisfies

$$|\nu_{mn}| = 1 \quad \text{and} \quad \overline{\nu_{N-m,M-m}} = (-1)^{m+n}\nu_{mn}. \tag{26}$$

A particular choice of ν satisfying (26) is

$$\nu_{mn} = \omega^{(1-\delta_{m0})(1-\delta_{n0})(m+n)^2 N/4}. \tag{27}$$

Other choices of ν_{mn} will be given later when we impose more conditions on the distribution functions. Finally, suppose the functions f in the definition of $W(x,z:\rho)$ satisfy the reality conditions (19) and the marginal condition $f(m,0) = f(0,n) = 1/N^2$. Then it is easy to see that for the incoherent state I/N,

$$W\left(x,z:\frac{I}{N}\right) = \frac{1}{N^2}. \tag{28}$$

The distribution function is *nondegenerate* at each point in phase space. This proves the existence of solutions to equations (24) and hence quasi-probability distributions satisfying R1, R2 and R4 in all finite dimensions.

Observe that the map $\Xi(x, z) : \rho \to W(x, z : \rho)$ is real and can be uniquely extended to a linear map on all Hermitian operators. That is, Ξ is a linear functional on K_{H}, the linear space of Hermitian operators on the system Hilbert space H. K_{H} is a real Hilbert space with respect to the scalar product $\langle A, B\rangle = \mathrm{Tr}(AB),\ A, B \in K_{\mathrm{H}}$. Since $\Xi(x, z)$ is nondegenerate at each point, there exists a unique nonzero $\hat{a}(x, z) \in K_H$ such that $W(x, z : \rho) = \langle \hat{a}(x, z), \rho\rangle = \mathrm{Tr}(\hat{a}(x, z)\rho)$. So the first of the equations in (21*a*) holds. The condition R2 and (21*a*) together imply

$$\begin{aligned}\langle \rho, \rho'\rangle = \mathrm{Tr}(\rho\rho') &= \sum_{xz} W(x, z : \rho)W(x, z : \rho') \\ &= \sum_{xz} W(x, z : \rho)\mathrm{Tr}(\hat{a}(x, z)\rho') = \mathrm{Tr}\left(\left(\sum_{xz} W(x, z : \rho)\hat{a}(x, z)\right)\rho'\right) \\ &= \left\langle \sum_{xz} W(x, z : \rho)\hat{a}(x, z), \rho'\right\rangle.\end{aligned}$$

Since this is true for all positive definite operators ρ' with trace 1, we conclude that the second of the equations in (21*a*) must hold. Now using this expansion of ρ in the operators $\hat{a}(x, z)$ and that fact R2 again, we conclude that equations (21*b*) hold. Finally, to prove that condition R3 also holds, observe that any Hermitian operator T can be written in the form $T = b_1\rho_1 - b_2\rho_2$, with $b_1, b_2 > 0$ and ρ_1, ρ_2 density matrices. Then,

$$\begin{aligned}\langle T\rangle = \mathrm{Tr}(\rho T) = \langle \rho, T\rangle &= b_1\langle \rho, \rho_1\rangle - b_2\langle \rho, \rho_2\rangle \\ &= b_1\mathrm{Tr}\left(\left(\sum_{xz} W(x, z : \rho)\hat{a}(x, z)\right)\rho_1\right) - b_2\mathrm{Tr}\left(\left(\sum_{xz} W(x, z : \rho)\hat{a}(x, z)\right)\rho_2\right) \\ &= \sum_{xz} W(x, z : \rho)\mathrm{Tr}(\hat{a}(x, z)(b_1\rho_1 - b_2\rho_2)) = \sum_{xz} W(x, z : \rho)\mathrm{Tr}(\hat{a}(x, z)T) \\ &= \sum_{xz} W(x, z : \rho)t(x, z).\end{aligned}$$

□

The distribution functions corresponding to f_0 will be called (finite) Wigner functions. In the case of odd dimensions, there is one such function. But in even dimensions, we have to be more careful in our choices.

4.2. Explicit formulae

The mere existence of 'orthonormal' Hermitian operators like $\hat{a}(x, z)$, which span the (real) space of observables, is simply a statement about the existence of orthonormal bases in any Hilbert space. Two characteristics distinguish $\hat{a}(x, z)$: firstly, the marginal distributions associated with them (R4), and secondly, the way they were derived via the QFT. Our next task is to find explicit forms for these operators. Let

$$W(x, z : \rho, f) = \sum_{m,n} f(m, n)\langle X^m Z^n\rangle\omega^{-(mx+nz)} \tag{29}$$

be a quasi-probability distribution satisfying R1–R4. We have indicated explicit dependence on the ordering function f. From this it follows that the phase-point operators are given by

$$\hat{a}(x, z : f) = \sum_{m,n} f(m, n) X^m Z^n \omega^{-(mx+nz)}. \tag{30}$$

The fact that $\hat{a}(x, z : f)$ form an orthonormal operator basis can be verified directly. The name 'phase-point operator' derives from the fact that (x, z) may be considered as a 'point' in a *finite* phase space. The quasi-probability distribution $W(x, z : \rho, f)$ are simply the coefficients in the expansion of ρ in the basis $\{\hat{a}(x, z : f)\}$. We will compute these operators in the 'computational' basis $\{|j\rangle = |j \bmod N\rangle\}$, i.e., the eigenbasis of the operator Z. Then $X^m = \sum_j |j+m\rangle\langle j|$ and $X^m Z^n = \sum_j \omega^{jn} |j+m\rangle\langle j|$. A straightforward calculation then gives

$$\hat{a}(x, z : f)_{kl} \equiv \langle k|\hat{a}(x, z : f)|l\rangle = \omega^{-(k-l)x} \sum_n f(k-l, n)\omega^{n(l-z)}. \tag{31}$$

In particular, the diagonal terms are easy,

$$\hat{a}(x, z : f)_{kk} = \delta_{k,z}/N. \tag{32}$$

Now, using formulae (25) for $f(m, n)$ in the formula, we obtain the following two cases for N. First for N odd,

$$\hat{a}(x, z : f_0)_{kl} = \frac{\omega^{-(k-l)x}}{N^2} \sum_n \omega^{n(k-l)(N+1)/2}\omega^{n(l-z)} = \frac{\omega^{-(k-l)x}\delta_{k+l,2z}}{N}. \tag{33}$$

Apart from ordering and normalization, these are precisely the phase-point operators found in [7] for prime dimensions. Note that we do not require the dimension N to be prime. If $N = 2r$ is even, the calculation is a bit more involved because the corresponding expression for $f_0(m, n)$ in (25) is not 'homogeneous'. We now have

$$\hat{a}(x, z : f_0)_{kl} \equiv \mathrm{Tr}(|l\rangle\langle k|\,\hat{a}(x, z : f)) = \omega^{-(k-l)x} \sum_n \nu_{mn}\omega^{n(k+l-z)/2}. \tag{34}$$

Evaluating these sums is not difficult but one has to be careful about the signs. For the choice of ν_{mn} given in (27), we obtain

$$\hat{a}(x, z)_{kl} = \begin{cases} \dfrac{\omega^{-(k-l)x}(1 \pm i)}{2N}\delta_{k+l,2z}, & k-l \text{ even}, \\ \dfrac{\omega^{-(k-l)x}\cot(\pi(k+l-2z)/N) \pm \mathrm{icsc}(\pi(k+l-2z)/N)}{N^2}, & k-l \text{ odd}. \end{cases} \tag{35}$$

So we see that the quasi-probability functions given above are much more complicated in even dimensions. More importantly, the phase-point operators given by (33) are more *sparse* than the one (35) for even dimensions. This, in turn, implies that in general quasi-probability distributions are sparser in odd dimensions and 'computationally simpler'. Let us illustrate this with an example.

Suppose that a quantum circuit or protocol is supposed to produce a state $|b\rangle$ in the computational basis. Because of noise and imperfections, we actually get a state (possibly mixed) that lies in the state space corresponding to the subspace K spanned by $\{|b \pm i\rangle : i \leqslant a\}$. From formulae (33) and (35), it is easy to see that the number of nonzero entries $W(x, z)$ in the odd case is $O(a)$ and in the even case it is $O(a^2)$. From the duality between X and Z, this is

also true if the computational basis is replaced by its Fourier transform. Since finding the quasi-probability distribution is equivalent to determining the state, does it mean that odd dimensions are tomographically 'better'? Should we look at qutrits too?

5. Heisenberg groups

In this section, we turn to our main theme: the Heisenberg groups and their close connections with Fourier transforms and distribution functions (see [23] for this connection in the continuous case). There are families of continuous and discrete Heisenberg groups. Although our primary focus will be on the discrete Heisenberg groups, we first take a look at the continuous Wigner function from a different perspective. We start with the (continuous) n-dimensional Heisenberg group $\mathbf{H}^n$ whose group manifold is $\mathbb{R}^{2n+1}$. Using vector notation, we write the elements as (p, q, t), where p and q are vectors in $\mathbb{R}^n$ and t is a real number. The reader can easily recognize the 'phase space' behind this notation. The group multiplication is defined by

$$(p, q, t)(p', q', t') = (p + p', q + q', t + t' + (p \cdot q' - q \cdot p')/2),$$

where the $\cdot$ denotes the usual scalar product. The symplectic structure is apparent in the above definition. By changing the parameterization of the group $(p, q, t) \to (p, q, t - pq/2) = (p', q', t')$, we obtain the multiplication law of the (polar) Heisenberg group [23]

$$(p'_1, q'_1, t'_1)(p'_2, q'_2, t'_2) = (p'_1 + p'_2, q'_1 + q'_2, t'_1 + t'_2 + p'_1 q'_2). \tag{36}$$

Note that the element $(0, 0, t)$ is in the centre of the group. Let us restrict ourselves to $n = 1$ for simplicity. The Lie group $\mathbf{H}^1$ is generated by the Lie algebra $\mathfrak{h}_1$ with generators $\{p, q, \lambda\}$ with brackets $[p, q] = \lambda$ and $[\lambda, p] = [\lambda, q] = 0$. One constructs the Poisson structure on the dual space $\mathfrak{h}_1^*$ in a natural way. The Heisenberg group plays a fundamental role in quantum mechanics. The Stone–von Neumann theorem asserts that the standard representations of position and momentum are essentially unique. In other words, the Schroedinger picture,

$$(p, q, t) \to \gamma(p, q, t) \equiv \mathrm{e}^{2\pi \mathrm{i} t}\, \mathrm{e}^{2\pi \mathrm{i}(p\hat{p} + q\hat{q})},$$

with $\hat{q}\psi(q') = q'\psi(q')$ and $\hat{p}\psi(q) = -\mathrm{i}\frac{\partial \psi(q)}{\partial x}$ is a unique representation of the Heisenberg group under some conditions of continuity. Here, ψ is the wave function in one dimension. Mathematically, it lives in the space $H = L^2(\mathbb{R})$ of complex square integrable functions (we ignore the technical difficulties arising due to the unboundedness of the operators). Since the elements $(0, 0, t)$ are in the centre, it is often sufficient to consider only elements of the form $\gamma(p, q) = \gamma(p, q, 0) = \mathrm{e}^{2\pi \mathrm{i}(p\hat{p} + q\hat{q})}$. This is the reduced Heisenberg group. Let $\{\psi_\alpha(x)\}$ be a basis in H. The matrix elements in this basis are given by

$$\begin{aligned} V_{\alpha\alpha'}(p, q) &= \langle \psi_\alpha | \gamma(p, q) | \psi_{\alpha'} \rangle = \langle \psi_\alpha | \mathrm{e}^{2\pi \mathrm{i}(p\hat{p} + q\hat{q})} | \psi_{\alpha'} \rangle \\ &= \int \overline{\psi_\alpha(u)} \mathrm{e}^{2\pi \mathrm{i}(p\hat{p} + q\hat{q})} \psi_{\alpha'}(v)\, \mathrm{d}u\, \mathrm{d}v. \end{aligned}$$

These are precisely matrix elements of the Fourier transform of the phase-point operators in the continuous case. In particular, $V_{\alpha\alpha}(0, q)$ yields Fourier transforms of the position probability density corresponding to the state ψ_α. Similarly, using the momentum representation, we obtain the other marginal for $V_{\alpha\alpha'}(p, 0)$. Since the basis was arbitrary, we conclude that the Wigner function,

$$W(p, q) = \int \langle \gamma(u, v) \rangle \mathrm{e}^{-2\pi \mathrm{i}(pu + qv)}\, \mathrm{d}u\, \mathrm{d}v,$$

when integrated over the strip between $q = c_1$ and $q = c_2$, gives the probability of the particle in (pure) state ψ to have its *position* between c_1 and c_2. Explicitly,

$$\int_{c_1}^{c_2} \mathrm{d}q \int_{-\infty}^{\infty} W(p, q : |\psi\rangle)\mathrm{d}p$$

yields the probability that the position observable has a value between c_1 and c_2 and similarly for the momentum observable. This is easily seen by expanding $|\psi\rangle$ in the position basis. What do we obtain if we integrate over an arbitrary strip, not necessarily parallel to the p or q axes, say the lines $ap + bq = c_1$ and $ap + bq = c_2$? The answer is well known and is discussed in [7] and [32]. But we look at it from a different perspective. First, put $ap + bq = p'$. This defines a family of parallel lines $p' = c$ in the p–q plane. Another line $cp + dq = q'$ does *not* belong to this family if and only if $ad - bc \neq 0$. Thus, the matrix $\zeta = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ is invertible and defines a change of coordinate in the phase plane. Then the form $up + vq = u'p' + v'q'$ where $(u', v') = (u, v)\zeta$. This in turn defines a transformation on the Lie algebra generated by $\hat{p}, \hat{q}$,

$$\gamma(u, v) = \mathrm{e}^{2\pi\mathrm{i}(u\hat{p}+v\hat{q})} = \mathrm{e}^{2\pi\mathrm{i}(u'\hat{p}'+v\hat{q}')}, \qquad \begin{pmatrix} \hat{p}' \\ \hat{q}' \end{pmatrix} = \zeta \begin{pmatrix} \hat{p} \\ \hat{q} \end{pmatrix}.$$

If the transformation $\hat{p} \to \hat{p}', \hat{q} \to \hat{q}'$ is an *automorphism*, then $\hat{p}'$ and $\hat{q}'$ have the same commutation relation as $\hat{p}$ and $\hat{q}$. This will happen if and only if $\det \zeta = ad - bc = 1$. But then if we change the variable of integration to p', q', the measure remains unchanged ($|\det \zeta| = 1$). We can now carry over the argument from the case of axis marginals and conclude that the integration of $W(p, q : |\psi\rangle)$ over a strip between $ap + bq = c_1$ and $ap + bq = c_2$ gives the probability that the observable $\hat{p}' = a\hat{p} + b\hat{q}$ will have a value lying between c_1 and c_2. Let us observe that $\zeta \in SL(2, \mathbb{R}) = Sp(1, \mathbb{R})$, where $SL(n, \mathbb{R})$ is the group of $n \times n$ real matrices with determinant 1, and $Sp(n, \mathbb{R})$ is the real symplectic group of order n. In general, $Sp(n, \mathbb{R})$ is a subgroup of the automorphism group of $\mathbf{H}^n$ and is different from $SL(2n, \mathbb{R})$.

Now we turn to the discrete Heisenberg group $\mathbf{H}$. We define a *presentation* of the group in terms of generators and defining relations [5]. $\mathbf{H}$ is generated by $\{\mathbf{x}, \mathbf{z}, \gamma\}$. The defining relations are

$$\mathbf{zx} = \gamma\mathbf{xz}, \quad \gamma\mathbf{x} = \mathbf{x}\gamma \quad \text{and} \quad \gamma\mathbf{z} = \mathbf{z}\gamma. \tag{37}$$

The advantage of this approach is that any map ϕ from the generators of a group $\mathbf{H}$ to another group K that satisfies the same defining relations as above can be uniquely extended to a group *homomorphisms* $\mathbf{H} \to K$. A simple realization of the group over integers is given by the set $\mathbf{Z}^3$. The multiplication is defined by

$$(j_1, k_1, t_1)(j_2, k_2, t_2) = (j_1 + j_2, k_1 + k_2, t_1 + t_2 + j_1k_2).$$

The generators are $\mathbf{z} = (1, 0, 0)$, $\mathbf{x} = (0, 1, 0)$, and $\gamma = (0, 0, 1)$. If we specialize to $\mathbf{Z}_N$, the integers modulo N, we get corresponding Heisenberg group $\mathbf{H}_N$ with generators X, Z and γ and the relations

$$X^N = Z^N = \gamma^N = e \text{ (identity)}, \quad \gamma X = X\gamma, \quad \gamma Z = Z\gamma \quad \text{and} \quad ZX = \gamma XZ. \tag{38}$$

Since $\mathbf{H}_N$ is a finite group, its finite-dimensional representations are unitary and completely reducible. Let ϕ be a representation of $\mathbf{H}$ or $\mathbf{H}_N$ on a vector space V of finite dimensions. We say that the central element γ acts *maximally* if the order of $\phi(\gamma)$ is $\dim(V)$. The following theorem characterizes representation of $\mathbf{H}$ ($\mathbf{H}_N$) and their relation to QFT.

Theorem 3. *Let ϕ be a (unitary) irreducible representation of $\mathbf{H}_N$ on a finite-dimensional space V. Let τ be the automorphism of $\mathbf{H}_N$ (and $\mathbf{H}$) given by $X \to Z$, $Z \to X$ and $\gamma \to \gamma^{-1}$ and ϕ' the representation defined by $\phi'(g) \equiv \phi(\tau g)$. Then the following statements are true:*

1. *$\phi(\gamma) = \omega$, a primitive Nth root of 1 and the eigenvalues of $\phi(Z)$ and $\phi(X)$ are $\{\omega^k : 0 \leqslant N-1\}$. γ acts maximally if and only if ϕ is faithful (one-to-one).*
2. *ϕ and ϕ' are unitarily equivalent: $\phi' = \Omega\phi\Omega^\dagger$ and Ω is the quantum Fourier operator.*
3. *Any unitary irreducible representation ψ of the full discrete Heisenberg group $\mathbf{H}$ in V in which the order of γ, $o(\gamma) = \dim(V) = K$ is equivalent to an irreducible faithful representation of $\mathbf{H}_K$.*

Proof. Assume first that γ acts maximally. Since ϕ is irreducible and γ is in the centre, it must act as a constant (Schur's lemma). As the order of γ is N, $\gamma = \omega I$, where ω is a primitive Nth root of unity. Since $\mathbf{H}_N$ is finite, we may assume the representations to be unitary. Let α be an eigenvector of $\phi(Z)$ with eigenvalue c. As $Z^N = e$, c must be an Nth root of 1. Consider the set $S = \{\alpha, \phi(X)\alpha, \ldots, \phi(X^{N-1})\alpha\}$. As

$$\phi(Z)\phi(X^k)\alpha = \phi(\gamma^k)\phi(X^k)\phi(Z)\alpha = c\omega^k\phi(X^k)\alpha,$$

$\phi(X^k)\alpha = \phi(X)^k\alpha$, $k = 0, 1, \ldots, N-1$ are eigenvector of $\phi(Z)$ with eigenvalue $c\omega^k$. These eigenvalues are distinct roots of 1 and hence S is linearly independent and a basis of V. We can reason similarly for $\phi(X)$. The converse is trivial. If $\gamma^k = I$ for $k < N$, then ϕ cannot be faithful.

Next we recall some facts from the theory of characters associated with representation of a group [33]. If ρ is a representation of a finite group G on a finite-dimensional vector space V, the character χ_ρ is a scalar function on G defined by $\chi_\rho(g) = \mathrm{Tr}(\rho(g))$. It is constant on conjugacy classes. If we have two characters χ_ρ and $\chi_{\rho'}$ corresponding to representations ρ and ρ', then their scalar product is defined as $(\chi_\rho, \chi_{\rho'}) = (1/N)\sum_{g\in G}\overline{\chi_\rho(g)}\chi_{\rho'}(g)$. It is a fundamental result that two *irreducible* representations ρ and ρ' are (unitarily) equivalent if and only if $(\chi_\rho, \chi_{\rho'}) \neq 0$. We apply this to the representations ϕ and ϕ' of $\mathbf{H}_N$. First, observe that since $ZX^mZ^nZ^{-1} = \gamma^m X^m Z^n$, $\chi_\phi(X^mZ^n) = \omega^m\chi_\phi(X^mZ^n)$, which is possible iff either $m = 0$ or $\chi_\phi(X^mZ^n) = 0$. Conjugating with X, we conclude that χ_ϕ is nonzero only on the centre of $\mathbf{H}_N$. Hence, to prove equivalence of ϕ and ϕ' it suffices to show that the scalar product of χ_ϕ and $\chi_{\phi'}$ is nonzero. But ϕ and ϕ' have the same effect on the centre (generated by γ) of $\mathbf{H}_N$. Hence $(\chi_\rho, \chi_{\rho'}) = (1/N)\sum_k \overline{\chi_\phi(\gamma^k)}\chi_{\phi'}(\gamma^k) = 1$. Since ϕ and ϕ' are equivalent, there exists a unitary map $\Omega : V \to V$ such that $\phi'(g) = \Omega^\dagger\phi(g)\Omega$. In particular, $\phi'(Z) = \phi(\tau Z) = \phi(X) = \Omega^\dagger\phi(Z)\Omega$. Now let $\{|j\rangle : j = 0, \ldots, \}$ be a complete set of eigenvectors of $\phi(Z)$ with $\phi(Z)|j\rangle = \omega^j|j\rangle$, and similarly let $\{|\tilde{j}\rangle\}$ be an eigenbasis of $\phi(X)$. Then $\phi(Z) = \sum_j \omega^j\,|j\rangle\langle j|$ and $\phi(X) = \sum_j \omega^j\,|\tilde{j}\rangle\langle\tilde{j}|$. Observing that $\{|j\rangle\langle\tilde{k}| : j,\ \ k = 0, \ldots, N-1\}$ form a basis the space of operators on V, it is easy to check that $\Omega = \sum_j |j\rangle\langle\tilde{j}|$. We have also seen that $\phi(X)|j\rangle = |j+1\rangle$. From these and the normalization $\langle\tilde{j}|0\rangle = 1$, we obtain $\langle j|\Omega|k\rangle = \omega^{-jk}/\sqrt{n}$. We have proved item 2.

To prove the last assertion, we again start with an eigenvector α with eigenvalue a of $\psi(\mathbf{z})$. Note that we can no longer assume that a is a Kth root of 1. However, the hypothesis that the order of $\psi(\gamma)$ is K implies that $\alpha, \psi(\mathbf{x})\alpha, \ldots, \psi(\mathbf{x}^{K-1})\alpha$ are eigenvectors of $\psi(\mathbf{z})$ with *distinct* eigenvalues $a\psi(\gamma)^j$, $j = 1, \ldots, K-1$. They must then be independent. This implies $\mathbf{x}^K\alpha = \alpha$

and hence $\mathbf{x}^K = 1$. Hence the eigenvalues of $\mathbf{x}$ must be Kth roots of 1. Interchanging the role of $\mathbf{x}$ and $\mathbf{z}$, we conclude that a must be a primitive Kth root of 1 and the assertion follows. □

Note that the condition on γ (maximal order) is necessary in the case of the groups $\mathbf{H}$ and $\mathbf{H}_N$. For example, let $N = 3$, $\rho(\gamma) = -I$, $\rho(Z)|0\rangle = |0\rangle$, $\rho(Z)|1\rangle = -|1\rangle$, $\rho(Z)|2\rangle = |2\rangle$ and $\rho(X)$ is the cyclic permutation. Then ρ is an irreducible representation of $\mathbf{H}_9$. We see the connection between representations of the Heisenberg groups and the QFT. For a vector $\tilde{\alpha}$, we write $\tilde{\alpha} = \Omega\alpha$ for its Fourier transform. The Plancherel formula $\|\alpha\|^2 = \|\alpha\|^2$ is simply stating that the Fourier operator Ω is unitary. We also note that since the representations of the group $\mathbf{H}$ is equivalent to $\mathbf{H}_N$ when γ acts maximally, it will be sufficient to consider $\mathbf{H}_N$ in a fixed representation space. However, when we are dealing with different representations (e.g., taking tensor products), we have to deal with the full Heisenberg group. The condition that $o(\gamma) = \dim(V)$ is a special case of general irreducible representations of $\mathbf{H}$. It is sufficient for our purposes and will be implicitly assumed. Henceforth, for a fixed representation ρ, we simply write the action of a group element g as $g\alpha$ instead of $\rho(g)\alpha$ if the context is clear.

Now we turn our attention to distribution functions. We have seen that the distribution functions can be given an alternative characterization in terms of phase-point operators. The formula (30) for these operators implies that they are a linear combination of the unitary operators of the group. Thus, we look for them in the *group algebra*. Recall that for a group G the group algebra $C(G)$ over complex numbers is the set of formal finite linear combinations $\sum_i c_i g_i$, $g_i \in G$ and $c_i \in \mathbb{C}$. The algebra product is defined as

$$\sum_i c_i g_i \sum_j d_j g_j = \sum_{ij} c_i d_j g_i g_j.$$

Any representation of the group is a representation of the group algebra and vice versa. Now for a *unitary* representation ρ of G on a finite-dimensional vector space, the character χ_ρ of the representation induces a scalar product on $C(G)$. Thus

$$(\mu, \nu) = \chi_\rho(\mu^* \nu), \quad \text{where} \quad \mu = \sum c_i g_i, \quad \nu = \sum c'_j g'_j \quad \text{and} \quad \mu^* = \sum \bar{c}_i g_i^{-1}. \tag{39}$$

This is indeed a scalar product on $C(G)$. The resulting norm coincides with the Hilbert–Schmidt norm on the corresponding operators on V. Call an element $\mu \in C(G)$ self-adjoint if $\mu^* = \mu$. Let $G = \mathbf{H}_N$ or $\mathbf{H}$. Since the central element γ acts as a scalar, we write elements of $C(G)$ in the form $\sum_{m,n} c_{mn} X^m Z^n$. For a representation ϕ of $\mathbf{H}_N$ with γ acting as ωI, consider the following elements,

$$A(x, z) = \sum_{mn} c_{mn} \omega^{-mx+nz} X^m Z^n, \quad x, z = 0, \ldots, N-1 \text{ in } C(\mathbf{H}_N).$$

We demand that the set $\mathcal{G} = \{A(x, z)\}$ be mutually orthogonal, self-adjoint and satisfy the following: the elements $P(x) = \sum_z A(x, z)$ and $Q(z) = \sum_x W(x, z)$ are projections, that is; $P(x)^2 = P(x)$ and $P(z)^2 = P(z)$. We call such a set of elements in $C(G)$ a Wigner set. We have the following theorem.

Theorem 4. *For a representation ϕ of $\mathbf{H}_N$ on N-dimensional space V, Wigner sets exist in $C(\mathbf{H}_N)$. If $\mathcal{G}$ is a Wigner set, then for $A(x, z) \in \mathcal{G}$, $A(x, z)/N$ are phase-point operators. In other words, given a quantum state ρ, the function $W(x, z : \rho) \equiv Tr(\phi(A(x, z))\rho)/N$ is a distribution function. Conversely, given a distribution function $W(x, z : \rho)$ on V, there is a unique Wigner set $A(x, z)$ in $C(\mathbf{H}_N)$ such that $W(x, z : \rho) \equiv Tr(\phi(A(x, z))\rho)/N$. Wigner*

sets are translation invariant in the sense that the transformation $c_{mn} \to c_{mn}\omega^{am+bn}$, $a, b \in Z_N$ *permutes the operators* $A(x, z)$ *in a Wigner set.*

Proof. The proof is similar to that of theorem 2. We only sketch some of the basic arguments because we are dealing with group algebras. First, the self-adjoint property implies conditions like (19) with c_{mn} in place of $f(m, n)$ because the $\{X^m Z^n\}$ are independent in $C(\mathbf{H}_N)$. Let us compute the scalar product of two elements from $\mathcal{G}$. Assuming now self-adjointness, we have

$$(A(x', z'), A(x, z)) = \mathrm{Tr}\left(\sum_{\substack{m,n \\ m',n'}} \bar{c}_{m'n'} c_{mn} \omega^{mn'} X^{m+m'} Z^{n+n'} \omega^{-(m'x'+n'z')} \omega^{-(mx+nz)} \right)$$

$$= \mathrm{Tr}\left(\sum_{\substack{m,n \\ m',n'}} c_{mn} c_{N-m,N-n} \omega^{mn} \omega^{-m(x-x')-n(z-z')} \right)$$

$$= N \sum |c_{mn}|^2 \omega^{-m(x-x')-n(z-z')}.$$

In deriving the second step, we use the fact that $\mathrm{Tr}(\phi(X^j Z^k)) = 0$ unless $j = k = 0 \bmod N$. The last expression will be proportional to $\delta_{xx'}\delta_{zz'}$, if $|c_{mn}|^2 = K$, a constant. We will fix K shortly. Hence, we assume that $c_{mn} = K\omega^{b_{mn}}$. For the last requirement, we have

$$Q(z) = \sum_x A(x, z) = N \sum_n c_{0n} Z^n \omega^{-nz} = Q(z)^2 = N^2 \sum_{mn} c_{0m} c_{0n} Z^{m+n} \omega^{-z(m+n)}.$$

This would be possible if $c_{0n} = K = 1/N$ for all n. We have already proved the existence of functions satisfying these conditions in theorem 2. The fact that $W(x, z : \rho) = \mathrm{Tr}(A(x, z)\rho)$ is real follows from self-adjointness. The orthogonality property implies R2 in section 4.1:

$$\mathrm{Tr}(\rho\rho') = \sum_{xz} W(x, z : \rho) W(x, z : \rho').$$

Finally, the property about marginals is equivalent to showing that $P(z)$ and $P(x)$ represent projections on $|z\rangle$ and $|\tilde{x}\rangle$, respectively. We can deduce this directly from the fact that

$$\phi(Z) = \sum_j \omega^j |j\rangle \langle j| \quad \text{and} \quad \phi(X) = \sum_j \omega^j |\tilde{j}\rangle \langle \tilde{j}|.$$

The proof of the converse is straightforward.

To prove the last statement let $c'_{mn} = c_{mn}\omega^{am+bn}$. Then

$$A'(x, z) \equiv \sum_{mn} c'_{mn} \omega^{-mx+nz} X^m Z^n$$

$$= \sum_{mn} c_{mn} \omega^{-m(x-a)+n(z-b)} X^m Z^n$$

$$= A(x', z') \text{ where } \quad x' = x - a \quad \text{and} \quad z' = z - b.$$

The assertion follows from this and the proof is complete. □

We see the correspondence between orthogonal sets in the group algebra $C(\mathbf{H}_N)$ and distribution functions. We have seen that the QFT arises out of a particular automorphism τ of the Heisenberg group. So we expect the general automorphisms of $\mathbf{H}_N$ and $\mathbf{H}$ to contain more structure and information relating to QFT and distribution functions. It is easy to see that any two representations of $\mathbf{H}_N$ in which γ acts maximally and has the same value are equivalent. In particular, if σ is an automorphism of $\mathbf{H}_N$ that fixes γ and ϕ is an arbitrary representation, then ϕ and $\phi \cdot \sigma$ are equivalent. If $\{A(x, z) : 0 \leqslant x, z \leqslant N-1\}$ is a Wigner set, then $\{\sigma A(x, z)\}$ is also a Wigner set. So we can generate new Wigner sets by automorphisms. As the value of σ on X and Z determines it on $C(\mathbf{H}_N)$, let $\sigma(X) = X^a Z^b$ and $\sigma(Z) = X^c Z^d$. We must have

$$(X^a Z^b)^N = \gamma^{abN(N-1)/2} = (X^c Z^d)^N = \gamma^{cdN(N-1)/2} = e \quad \text{and}$$
$$X^c Z^d X^a Z^b = \gamma^{ad-bc} X^a Z^b X^c Z^d = \gamma X^a Z^b X^c Z^d$$

for σ to be an automorphism. The second condition implies that the matrix

$$M^\sigma \equiv \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

has determinant 1. That is, $M^\sigma \in SL(2, Z_N)$, the set of matrices with entries in the ring Z_N and determinant $1 \bmod N$. Conversely, given $M \in SL(2, Z_N)$, N odd, we can define an automorphism σ_M as above. If N is even this simple definition of σ_M does not work in general. For example, if ab is *odd*, then $(X^a Z^b)^N = \gamma^{N/2} = -1$. This is reminiscent of half-integral representation of rotation group ($SU(2)$ actually). Hence for even N, we have an automorphism of $\mathbf{H}$ rather than $\mathbf{H}_N$. Note that in this case for any $M \in SL(2, Z_N)$, $(\sigma_M(g))^{2N} = 1$, $g \in \mathbf{H}_N$. There is, however, a proper subgroup of $SL(2, Z_N)$ which induces an automorphism of $\mathbf{H}_N$. Alternatively, for even N define the function

$$\operatorname{sgn}(u) = \begin{cases} 0, & u \in Z \text{ and } u \text{ even}, \\ 1, & u \text{ odd}. \end{cases}$$

Now, we can define the automorphism σ_M on $\mathbf{H}_N$, $M \in SL(2, Z_N)$ for odd N and for even N, we define it on the representation space.

$$\sigma_M(X) = \begin{cases} X^a Z^b, & N \text{ odd}, \\ \omega^{\operatorname{sgn}(ab)/2} X^a Z^b, & N \text{ even}. \end{cases} \tag{40}$$

Let $A(x, z)$ be a Wigner set and ϕ a representation of $\mathbf{H}_N$. Then we have seen that a distribution function is defined by

$$W(x, z : \rho) = \operatorname{Tr}(\phi(A(x, z))\rho)/N.$$

This means that if ϕ and ϕ' are equivalent representations connected by a unitary operator U, and W and W' are the corresponding distribution functions, then

$$W(x, z : \rho) = W'(x, z : U^\dagger \rho U). \tag{41}$$

In particular, if $M \in SL(2, Z_N)$, then it induces an equivalent representation ϕ_M. In case N is odd, ϕ_M is the representation that is given via the automorphism generated by M. In even dimension, ϕ_M is defined by (40) above. Now let us look at other 'marginals' of a distribution function $W(x, z : \rho)$. One way of constructing such marginals is via a finite Radon transform [27, 28]. Thus for $f : Z_N \times Z_N \to \mathbb{C}$ define 'lines'

$$\begin{aligned} &S_{ab} = \{(x, z) \in Z_N \times Z_N : ax + bz = 0 \bmod N, \quad \gcd(a, b, N) = 1\}, \\ &\hat{W}(x', z' : \rho) = \textstyle\sum_{x,z \in S_{ab}+(x',z')} W(x, z, : \rho). \end{aligned} \tag{42}$$

The condition $\gcd(a, b, N) = 1$ ensures that the 'line' $ax + bz = t$ has a solution for all $t \in Z_N$. The 'coordinate axes', for example, correspond to the sets S_{10} and S_{01}. The function $\hat{W}(z')$ in (42) is a *Radon transform* [28] of the distribution function $W(x, z)$ and each pair $(a, b) \in Z_N \times Z_N$ such that $\gcd(a, b)$ is invertible in Z_N, defines such a transform. Let $x' = ax + bz$. Let $(c, d) \in Z_N \times Z_N$ such that the matrix

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2, Z_N)$$

and set $z' = cx + \mathrm{d}z$. In the following, it will be convenient to use vector notation. Thus $\boldsymbol{\xi} = (x, z)^T \in Z_n \times Z_n$ is a 2D 'vector'[4]. We will also occasionally use the component notation: $\boldsymbol{\xi} = (\xi_1, \xi_2)^T$. So the distribution functions and phase-point operators may be written as $W(\boldsymbol{\xi} : \rho)$ and $A(\boldsymbol{\xi})$, respectively. Then a marginal with respect to the second component is given by

$$\hat{W}(z' : \rho) = \sum_{M\boldsymbol{\xi}_1 \in Z_N} W(M\boldsymbol{\xi} : \rho).$$

Here, $z' = M\boldsymbol{\xi}_2$ should be replaced by $z' = (M\boldsymbol{\xi})_2$. In analogy with the continuous case, we require that $\hat{W}(z' : \rho)$ is a probability distribution with respect to z'. More precisely, in the representation $\phi_{M^{-1}}$ of the Heisenberg group corresponding to the automorphism induced by M^{-1}, $\hat{W}(z'; \rho)$ gives probability distribution of the *quantum observable* $-\mathrm{i} \ln \phi_{M^{-1}}(Z)$ in the state ρ. However, there is a sharp difference between the distribution functions in even and odd dimensions. The general marginal condition holds in odd dimensions for the Wigner distribution function defined by (25) but not for even dimensions. For even dimension, we have more complicated formulae for the marginals of the Wigner function. In fact, we will show that in this case no distribution function satisfying conditions R1–R4 in section 4.1 will satisfy the general marginal condition for all $M \in SL(2, Z_N)$.

Theorem 5. *Let ϕ be a representation of $\mathbf{H}_N$ on V. Let*

$$A(x, z) = \frac{1}{N^2} \sum_{m,n} f(m, n) X^m Z^n \omega^{-mx+nz} \tag{43}$$

be a Wigner set and $W(x, z : \rho) = \mathrm{Tr}(A(x, z)\rho)$ be the corresponding distribution function.

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2, Z_N).$$

We say that simple marginal condition (with respect to M) is satisfied if the marginals

$$\hat{W}(z' : \rho) \equiv \sum_{M^{-1}\xi_1 \in Z_N} W(M^{-1}\boldsymbol{\xi} : \rho) \tag{44}$$

are probability distribution in the eigenbasis of the operator $\phi_M(Z) \equiv \phi(\sigma_M(Z)) = \phi(X^c Z^d)$. Then the following statements hold:

1. *The (simple) marginal condition is satisfied if and only if*

$$A_M(x, z) \equiv A(M^{-1}\boldsymbol{\xi})$$

is a Wigner set for the representation ϕ_M.

[4] To be accurate, $Z_N \times Z_N$ is a *module* over the ring Z_N.

2. *If marginal conditions are satisfied for all* $M \in SL(2, Z_N)$*, then the Wigner sets (and distribution functions) are determined uniquely up to translations.*
3. *If the dimension* $N > 2$ *is even, it is not possible to satisfy the marginal condition for all* $M \in SL(2, Z_N)$*.*
4. *In odd dimensions, Wigner functions are given by*

$$W(x, z) = \sum_{m,n} \omega^{mn(N+1)/2} X^m Z^n \omega^{-(mx+nz)} \tag{45}$$

up to translations. In that case, $\hat{W}(z' : \rho) = \langle \alpha_{z'+cd/2} | \rho | \alpha_{z'+cd/2} \rangle$*, where* $|\alpha_j\rangle$ *are the eigenvectors of* $\sigma_M(Z)$*.*

Proof. Since $\gcd(a, b, N) = 1$, there exist integers c, d and k such that $ad - bc + kN = 1$ and hence $ad - bc = 1 \bmod N$. For a matrix M, let $M' = (M^{\mathrm{T}})^{-1}$. The first assertion in the list is relatively straightforward. If the equation (44) is satisfied, then $W_M(x, z) = W(M^{-1}x, M^{-1}z)$ is a distribution function. The condition R1 (reality) is clear, R2 follows from the state transformation equation (41) and (44) gives the marginal condition R4. From the correspondence between distribution functions and Wigner sets, the first assertion is clear.

Suppose the marginal conditions are satisfied for some $M \in SL(2, Z_N)$ given above. Using the formula

$$(X^u Z^v)^m = \omega^{uvm(m-1)/2} X^{um} Z^{vm}, \tag{46}$$

we obtain setting $\boldsymbol{\xi} = (m, n)^T \in Z_n \times Z_n$.

$$\begin{aligned}
N^2 \sum_{\xi_1} A(M^{-1}\boldsymbol{\xi} : \rho) &= \sum_{M\xi_1} \sum_{m,n} f(m, n) X^m Z^n \omega^{-\boldsymbol{\zeta}\cdot M^{-1}\boldsymbol{\xi}} \\
&= \sum_{\xi_1} \sum_{m,n} f(m, n) X^m Z^n \omega^{-M'\boldsymbol{\zeta}\cdot\boldsymbol{\xi}} \\
&= \sum_{M\xi_1} \sum_{\zeta'=M'\zeta} f(M^{\mathrm{T}}\boldsymbol{\zeta}') X^m Z^n \omega^{-\boldsymbol{\zeta}'\cdot\boldsymbol{\xi}} \\
&= \sum_{x\in Z_N} \sum_{m,n} f(am + cn, bm + dn) \omega^{-(abm(m-1)/2 + cdn(n-1)/2 + bcmn)} \\
&\quad \times \sigma_M(X)^m \sigma_M(Z)^n \omega^{-(mx+nz)}, \quad (\xi_1 = x) \\
&= \sum_n f(cn, dn) \omega^{-(cdn(n-1)/2)} \sigma_M(Z)^n \omega^{-nz} \\
&\equiv \sum_n g(n) Z'^n, \quad g(n) = f(cn, dn) \omega^{-(cdn(n-1)/2)} \quad \text{and} \quad Z' = \sigma_M(Z) \omega^{-z}.
\end{aligned}$$

We require that the operator $T = \sum_n g(n) Z'^n$ be a projection: it must be Hermitian and satisfy $T^2 = T$. Hence, we must have

$$\sum_{m,n} g(m) g(n) Z'^{n+m} = \sum_k \sum_m g(m) g(k - m) Z'^k = \sum_k g(k) Z'^k.$$

Since Z' like Z has no repeated eigenvalue, its minimum polynomial is the characteristic polynomial $\lambda^N - 1$. Hence, the operators $I, Z', Z'^2, \ldots, Z'^{N-1}$ are linearly independent and we

must have

$$\sum_m g(m)g(k-m) = g(k).$$

But the left-hand side is the *convolution* of the function g with itself. Taking (finite) Fourier transform of both sides, we obtain $\tilde{g}^2 = \tilde{g}$. Then $\tilde{g}(m) = 1$ or 0. This implies that

$$g(n) = \sum_i \omega^{t_i n},$$

where $t_i \in Z_N$ are the values at which $\tilde{g} = 1$. But from condition R2, we infer that $|g(x)| = 1$ and since R1 implies that $g(0) = 1$, we conclude that there must be exactly one term in the above sum,

$$g(n) = f(cn, dn)\omega^{-cdn(n-1)/2} = \omega^{t(c,d)n}. \tag{47}$$

Putting $n = 1$, this implies that $f(c, d) = \omega^{t(c,d)}$ whenever $\gcd(c, d, N) = 1$. Hence, we rewrite the above equation as

$$f(cn, dn) = f(c, d)^n \omega^{cdn(n-1)/2} \text{ and so } f(cn, n) = f(c, 1)^n \omega^{cn(n-1)/2}. \tag{48}$$

Now suppose N is even. By the definition of Wigner sets, they must be independent since the operators are mutually orthogonal. Consequently, the function f must be periodic with period N and since $f(c, 1) = \omega^{t(c,1)}$, $t(c, 1)$ must be an integer. Putting $n = N$ in the second equation in (48) and noting that $f(x, 0) = 1$, $\forall x \in \mathbb{Z}$, we get a contradiction when c is odd for the right-hand side is -1. Hence, it is not possible to have Wigner sets satisfying *all* simple marginal conditions.

Next suppose that N is odd. Then 2 has an inverse $(N+1)/2$ in Z_N. It is an easy verification that the function $f(m, n) = \omega^{mn(N+1)/2}$ satisfies the functional relation (19). To prove uniqueness, we assume that $t(m, n)$ can be extended to all Z and that it can be expressed as a polynomial in m and n with integer coefficients (which may depend on N). Since $f(m, 0) = f(n, 0) = 1$, we may assume that the polynomial is of the form $t(m, n) = \omega^{mn[a_0+g(m,n)]}$, where a_0 is a constant and $g(m, n)$ is a polynomial without constant term. Then, we have

$$f(cn, n) = \omega^{cn^2(a_0+g(cn,n))} = \omega^{nc(a_0+g(c,1))}\omega^{cn(n-1)/2}.$$

Since this must be satisfied for all n, we must have $a_0 = (N+1)/2$ and $g = 0 \bmod N$. This proves uniqueness up to linear terms.

The last statement is easily derived from the above proof of the existence and uniqueness of distribution function satisfying all of the marginal conditions for odd N. □

We note that a similar relation holds for the marginal distribution over x when we average over the variable z. In fact, satisfaction of marginal conditions under the full $SL(2, Z_N)$ for one variable implies the same for others. In even dimensions, there exists no distribution function satisfying all marginal conditions. Therefore, we have to relax some of the conditions of the theorem to obtain the marginal distributions. Let us recall why the marginal conditions are desirable. One of the main reasons is that by determining a sufficient number of marginal distributions, we can reconstruct the state if the simple marginal condition stated in theorem 5 is satisfied (see (44) and the statement that follows it) the marginal distribution corresponds to probabilities for a complete projective measurement in a suitable basis. In even dimension, we

have three options:

1. We do not require that the Wigner set be *independent*. Then the representation of the Heisenberg group $\mathbf{H}_N$ need not be irreducible. This was the approach adopted in [25].
2. We drop the conditions that the marginals are of simple type. As will be shown next, we can still determine the 'marginal' distributions from the measurement probabilities.
3. We do not demand that the marginal condition is satisfied for the full $SL(2, Z_N)$ but only for a subset. We show that in case $N = 2^K$ there is such a subset and the marginal distribution for it is sufficient to reconstruct the distribution function.

We start with the first option [16, 25]. Since the operators $A(x, z)$ are no longer independent, the function f (as a function on $\mathbb{Z}$) is not required to be periodic and the labels (x, z) can take any integer values. A minimal extension is obtained by looking at the basic recurrence relations (48). The problematic factor $\omega^{cn(n-1)/2}$ is periodic with a period $2N$ (as function of $\mathbb{Z}$). The same relations then suggest that we take $f(m, n) = \omega^{mn/2}$, where $\omega^{1/2}$ is a primitive $2N$th root of 1. Hermiticity of phase-point operators then requires that we now define them as

$$A(x, z) = \sum_{m,n \in Z_{2N}} \omega^{mn/2} X^m Z^n \omega^{-(mx+nz)/2}.$$

Because of redundancy, these operators are not uniquely determined (up to linear factors). But we can modify the proof in theorem 5 for odd dimension to determine the possible solutions in this case.

Next we look at option 2. We defined a family of distribution functions, say $W(x, z : \rho, \nu)$, in the even case in (25) depending on some function ν. The function ν is arbitrary apart from the condition (26). Let W_0 denote the special case when ν is given by (27). Of course, W_0 does not satisfy the marginal condition but the results below show how it may be computed from the measurement probabilities.

Proposition 2. *Let V be an irreducible representation space of $\mathbf{H}_N$, with N being even. Let*

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2, Z_N).$$

Let $u = gcd(t, N)$, where $t = c$ if c is even and d otherwise. Suppose that N/u is even. Then

$$\begin{aligned} \hat{W}(z' : \rho, \nu) &\equiv \sum_{x' \in Z_n} W(M^{-1}x', M^{-1}z' : \rho) \\ &= \sum_j \rho_{jj} \sum_n \nu_{cn,dn} (-1)^{dn\lfloor \frac{cn}{N} \rfloor + cn \lfloor \frac{dn}{N} \rfloor} \omega^{((cd+\mathrm{sgn}(cd))/2 + j - z')n}, \end{aligned} \tag{49}$$

where $\lfloor x \rfloor$ is the greatest integer $\leqslant x$. In particular, for W_0 with ν_{mn} given by (27), we have

$$\begin{aligned} \hat{W}_0(z' : \rho) = {} & (\langle \alpha_{z' - \frac{cd-\mathrm{sgn}(cd)}{2}} | \rho | \alpha_{z' - \frac{cd-\mathrm{sgn}(cd)}{2}} \rangle + \langle \alpha_{\frac{N}{2} + z' - \frac{cd-\mathrm{sgn}(cd)}{2}} | \rho | \alpha_{\frac{N}{2} + z' - \frac{cd-\mathrm{sgn}(cd)}{2}} \rangle)/2 \\ & - \frac{2}{N} \sum_j \rho_{jj} \sum_{n \text{ odd}} (-1)^{d\lfloor \frac{cn}{N} \rfloor + c \lfloor \frac{dn}{N} \rfloor} h(j, z'), \end{aligned} \tag{50}$$

$$\text{where} \quad h(j, z') = \begin{cases} \cos \dfrac{2\pi(\frac{cd-1}{2} + j - z')}{N}, & \text{if } cd \text{ odd}, \\ -\sin \dfrac{2\pi(\frac{cd-1}{2} + j - z')}{N}, & \text{if } cd \text{ even}. \end{cases}$$

The proof is given in the appendix. Observe that if $N = 2^k$, $K > 1$, then $N_1 = N/u$ is always even. We can also write the appropriate formulae for the case N_1 odd. We avoid doing so as they are even more complicated. We can also simplify the trigonometric sums in (49). However, note that if we know the probabilities $\langle j|\rho|j\rangle$, then in principle the Radon transform $\hat{W}(z:\rho, M)$ can be computed by evaluating these sums. In the case of odd dimensions, the expressions for the marginals are simpler but we still have to estimate the probability distribution in the basis $\{|\alpha_j\rangle\}$ defined above. If we have these probabilities, doing the sums in the even case is routine. Hence, is there a deeper reason for imposing the marginal conditions on the distribution function? Two possible reasons could be simplicity and some theoretical insight.

We consider the third option listed above for dimension $N = 2^k$ only. Thus, we aim to construct a distribution function that satisfies the marginal conditions for only a subset of $SL(2, Z_N)$. The theorem below gives an explicit formula for this important case. Thus, let $L_1 \subset SL(2, Z_N)$ be a subset consisting of the following matrices. If $M \in L_1$, then each row has at least one entry $= 1$, and if the diagonal entry is $\neq 1$, it is even.

Theorem 6. *Let $N = 2^k$. Define*

$$W_1(x,z) = \frac{1}{N^2}\left(\sum_{\substack{m,n\\ \text{even}}} \langle X^m Z^n\rangle \omega^{mn/2-(mx+nz)} + \sum_m \langle X^m\rangle \omega^{-mx} + \sum_n \langle Z^n\rangle \omega^{-nz} + \sum_{\substack{m>0\\ n \text{ odd}}} (-1)^{\left\lfloor \frac{(mn^{-1})n}{N}\right\rfloor} \langle X^m Z^n\rangle \omega^{mn/2-(mx+nz)} + \sum_{\substack{n \text{ even} > 0\\ m \text{ odd}}} (-1)^{\left\lfloor \frac{(nm^{-1})m}{N}\right\rfloor} \langle X^m Z^n\rangle \omega^{mn/2-(mx+nz)}\right), \tag{51}$$

where the expressions like (mn^{-1}) are first computed modulo N in the residue class $\{0,\ldots,N-1\}$ and then treated as an integer. $\lfloor x\rfloor$ denotes the largest integer less than or equal to x. Then W_1 satisfies the conditions R1–R4 and for every $M \in L_1$, W_1 satisfies a simple marginal condition with respect to the variable x,

$$\hat{W}_1(z':\rho,v) \equiv \sum_{x'\in Z_n} W_1(M^{-1}x', M^{-1}z':\rho) = \begin{cases} \langle \alpha_{z'-\frac{c+\mathrm{sgn}(c)}{2}}|\rho|\alpha_{z'-\frac{c+\mathrm{sgn}(c)}{2}}\rangle, & d=1,\\ \langle \alpha_{z'-\frac{d}{2}}|\rho|\alpha_{z'-\frac{d}{2}}\rangle, & c=1 \quad \text{and} \quad d \text{ even}. \end{cases} \tag{52}$$

Proof. We first note that the notation n^{-1} makes sense in the ring Z_N since every odd n is invertible. The reality condition R_1 is seen from the following simple observation. For $0 \leqslant m, n < N$ let $mn^{-1} = k_1 N + n_1$, $k_1 \geqslant 0$ and $0 \leqslant n_1 < N$. Since $(N-m)(N-n)^{-1} = (mn^{-1}) \bmod N$, we obtain

$$((N-m)(N-n)^{-1})(N-n) = (m-1-k_1)N + n_1.$$

This implies that k_1 has same (opposite) parity as $\lfloor((N-m)(N-n)^{-1})(N-n)\rfloor$ if m is odd (even). Hence

$$(-1)^{\lfloor(mn^{-1})n\rfloor} = (-1)^{m+n}(-1)^{\lfloor((N-m)(N-n)^{-1})n\rfloor}.$$

We can argue similarly for m odd and n even. Hence, the reality condition (19) is satisfied. It is clear that W_1 is normalized. The other conditions easily follow from the definition and the analysis of these conditions in section 4.1. Finally, the simple marginal condition with respect to x is seen to be satisfied as follows. From theorem 5 and proposition 2, we note that we have to consider pairs of the form (cn, dn), where (c, d) is the second row of M, in the calculation of the marginals. Using the notation of proposition 2, we set

$$\nu_{mn} = \begin{cases} 1 & m, n \text{ even}, \\ \left\lfloor \dfrac{(mn^{-1})n}{N} \right\rfloor & n \text{ odd}, \\ \left\lfloor \dfrac{(nm^{-1})m}{N} \right\rfloor & m \text{ odd}, n \text{ even}. \end{cases}$$

As the matrices belong to L_1, we consider two cases. If the diagonal element $d = 1$, then the only terms in the sum yielding W_1 that contribute to the marginal are indexed by $((cn), n)$, where n runs through Z_N and (cn) is calculated mod N. The case $c = 0$ is already covered. If $c \neq 0$, then from (49),

$$\begin{aligned} \hat{W}_1(z' : \rho, \nu) &\equiv \sum_{x' \in Z_n} W_1(M^{-1}x', M^{-1}z' : \rho) \\ &= \sum_j \rho_{jj} \sum_n \nu_{cn,dn} (-1)^{dn\lfloor \frac{cn}{N} \rfloor + cn \lfloor \frac{dn}{N} \rfloor} \omega^{((cd+\mathrm{sgn}(cd))/2+j-z')n} \\ &= \sum_j \rho_{jj} \left(\sum_{n \text{odd}} \nu_{cn,n} \left\lfloor \frac{cn}{N} \right\rfloor \omega^{((c+\mathrm{sgn}(c))/2+j-z')n} + \sum_{n \text{ even}} \nu_{cn,n} \omega^{((c+\mathrm{sgn}(c))/2+j-z')n} \right) \\ &= \sum_j \rho_{jj} \sum_n \omega^{((c+\mathrm{sgn}(c))/2+j-z')n} = \left\langle \alpha_{z'-\frac{c+\mathrm{sgn}(c)}{2}} | \rho | \alpha_{z'-\frac{c+\mathrm{sgn}(c)}{2}} \right\rangle. \end{aligned}$$

For the case $c = 1$ and d even, the terms in which n is odd drop out from the sum for $\hat{W}_1(z' : \rho, \nu)$ and the proof is similar to the first case. □

We note that the subset L_1 of matrices from $SL(2, Z_N)$ cannot be extended arbitrarily, preserving the property of simple marginals. For example, if we admit matrices with $c = 1$ and d *odd*, then we get a factor of $\mathrm{sgn}(\lfloor \frac{(d^{-1})n}{N} \rfloor)$ instead of $\mathrm{sgn}(\lfloor \frac{dn}{N} \rfloor)$. The two need not be equal. However, as we will see below, the set L_1 is sufficient to determine W_1.

5.1. Inverse Radon transform and state determination

In the previous section, we saw that the finite Wigner distribution function enjoys a rich variety of marginal properties. We can use this to determine the former. This is equivalent to inverting a finite set of Radon transforms. From the distribution function, we can determine the state. The invertibility of the Radon transforms also shows that the Wigner distribution function is unique up to a translation. In the rest of the section, $W(x, z)$ will denote the Wigner distribution

function. Replacing the matrix M^{-1} by

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad \det M = 1 \bmod N,$$

in theorem 5, we rewrite the basic Radon property stated in (44) and the statement following. For example, for odd N,

$$\hat{W}(z : \rho, M) \equiv \sum_{x \in Z_N} W(Mx, Mz : \rho) = \mathrm{Tr}(|\alpha_{z-ac}\rangle \langle \alpha_{z-ac}| \rho). \tag{53}$$

The problem is to reconstruct $W(x, z)$ from $\hat{W}(z : \rho, M)$. Call the latter the Radon transform of W with respect to the matrix M. The idea is that $\hat{W}(z : \rho, M)$ is the probability distribution of the observable $-\mathrm{i} \ln(X^{-c} Z^a)$ in the odd case. In the case of even dimensions, it can be computed from the distributions. Assuming that these distributions can be approximately determined experimentally, we can reconstruct W and hence ρ. We have seen that in odd dimension N, there is a distribution function satisfying simple marginal conditions for every $M \in SL(2, Z_N)$ and in dimension $N = 2^k$ we have only a subset of $SL(2, Z_N)$ with simple marginal conditions. We give explicit formulae for these two cases. First some notation. For a subset S of some set, let χ_S denote the indicator function: $\chi_S(x) = 1$ if $x \in S$ and 0 otherwise. In the rest of the section, we use the boldface vector notation to denote a member of $Z_N \times Z_N$ and other nonbold letters to denote 'scalars' belonging to Z_N. For example,

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \quad \mu_1, \mu_2 \in Z_N.$$

Given $M \in SL(2, Z_N)$, let $\mathbf{C}_i(M)$, $i = 1, 2$, denote the column vectors of M. Let

$$S_i(M) = \{\mathbf{C}_i(M)x : x \in Z_N\} \subset Z_N \times Z_N, \quad i = 1, 2.$$

Theorem 7. *Any distribution function $W(x, z)$ can be uniquely determined from the (finite) set of Radon transforms $\hat{W}(z : \rho, M)$, where*

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2, Z_N).$$

In particular, for odd dimensions the Wigner function given in (45) and any M, we have

$$\widetilde{W}(ct, -at) = \frac{1}{N} \sum_z \langle \alpha_{z-ac/2}(M)|\rho|\alpha_{z-ac/2}(M)\rangle \, \omega^{zt} \tag{54}$$

and for $N = 2^k$ and $M \in L_1$,

$$\widetilde{W}(ct, -at) = \begin{cases} \dfrac{1}{N} \displaystyle\sum_z \langle \alpha_{z-\frac{c+\mathrm{sgn}(c)}{2}}(M)|\rho|\alpha_{z-\frac{c+\mathrm{sgn}(c)}{2}}(M)\rangle \, \omega^{zt}, & a = 1, \\ \langle \alpha_{z-\frac{a}{2}}(M)|\rho|\alpha_{z-\frac{a}{2}}(M)\rangle \, \omega^{zt}, & c = -1 \quad \text{and} \quad a \text{ even}, \end{cases} \tag{55}$$

where $\widetilde{W}$ is the Fourier transform of W in $Z_N \times Z_N$ and $|\alpha_j(M)\rangle$ are the eigenvectors of $\sigma_{M^{-1}}(Z) = X^{-c} Z^a$.[5] *In either of the cases, the Wigner function W or W_1 can be reconstructed from the marginal distributions.*

[5] Recall that here M replaces M^{-1} of theorem 5.

Proof. Write the Radon transform (53) as

$$\hat{W}(z:\rho, M) = \sum_{x\in Z_N} W(Mx, Mz:\rho) = \sum_{\mathbf{x}\in S_1'(M)} W(\mathbf{C}_2(M)z - \mathbf{x}:\rho),$$

where $S_1' = -S_1$. We can write this as a *convolution*. Thus

$$\begin{aligned}\hat{W}(\mathbf{u}:\rho, M) &= \chi_{S_z}(\mathbf{u}) \sum_{\mathbf{x}} W(\mathbf{u}-\mathbf{x}:\rho)\chi_{S_1'(M)}(\mathbf{x}) \\ &= \chi_{S_z}(\mathbf{u})(W \star \chi_{S_1'(M)}(\mathbf{u})), \quad S_z = \{\mathbf{C}_2(M)z\}.\end{aligned}$$

Now we take the finite Fourier transform of the above equation in the group $Z_N \times Z_N$ [22]. Recall that the Fourier transform of a complex function $f(\mathbf{u})$ on $Z_N \times Z_N$ by $\tilde{f}$ is a function on the dual group $(Z_N \times Z_N)^*$,

$$\tilde{f}(\boldsymbol{\mu}) = \frac{1}{N} \sum_{u_1,u_2} \omega^{-(\mu_1 u_1 + \mu_2 u_2)} f(\mathbf{u}).$$

Using the fact that $\hat{W} = \hat{W}\chi_{S_z}$ and that the Fourier transform of a convolution is a product and vice versa, we have (suppressing ρ and M)

$$\begin{aligned}\widetilde{\hat{W}}(\boldsymbol{\mu}) &= \sum_{u_1,u_2} \hat{W}(\mathbf{u})\omega^{-\boldsymbol{\mu}\cdot\mathbf{u}} \\ &= \widetilde{\chi}_{S_z} \star (\widetilde{W}\widetilde{\chi}_{S_1'(M)})(\boldsymbol{\mu}) = \sum_{\nu} \widetilde{\chi}_{S_z}(\boldsymbol{\mu}-\boldsymbol{\nu})\widetilde{W}(\boldsymbol{\nu})\widetilde{\chi}_{S_1'(M)}(\boldsymbol{\nu}) \\ &= \sum_{\{\boldsymbol{\nu}: a\nu_1 + c\nu_2 = 0\}} \omega^{-[(\mu_1-\nu_1)b + (\mu_2-\nu_2)d]z}\widetilde{W}(\boldsymbol{\nu}) \\ &= \omega^{-(\mu_1 b + \mu_2 d)z} \sum_t \omega^{((ct)b-(at)d)z}\widetilde{W}(ct, -at) \\ &= \omega^{-(\mu_1 b + \mu_2 d)z} \sum_t \omega^{-tz} F(t) = \sqrt{N}\omega^{-(\mu_1 b + \mu_2 d)z}\widetilde{F}(z),\end{aligned}$$

where $F(t) = \widetilde{W}(ct, -at)$ and $\widetilde{F}$ is its Fourier transform in Z_N. In proving the above, we use the following facts: $\widetilde{\chi}_{S_1'(M)}(\boldsymbol{\nu}) \neq 0$ iff $a\nu_1 + c\nu_2 = 0$ and the solution to the congruence equation $a\nu_1 + c\nu_2 = 0 \bmod N$ is given by the set $\{(ct, -at) : t \in Z_N$. This follows from a similar result for linear Diophantine equations [34] and the fact that $\gcd(a, c, N) = 1$. We also use $\det M = ad - bc = 1$ in the last but one step. The factor $\sqrt{N}$ appears because of the normalization used in our definition of finite Fourier transform. It now follows that

$$\tilde{W}(ct, -at) = \frac{1}{N} \sum_z \widehat{W}(z)\omega^{zt}. \tag{56}$$

This formula is valid for *any* distribution function. Let now N be odd. Combining this with equation (44) in theorem 5, we obtain (54). Similarly, when $N = 2^K$ and $M^{-1} \in L_1$, we obtain (55). Note that the formulae in (52) are valid under the assumption that $M \in L_1$ (see footnote 4 above).

We next show that it is always possible to find $a, c \in Z_N$ such that $\gcd(a, c, N) = 1$ and the 'lines' $\{(ct, -at) : t \in Z_N\}$ cover the 'plane' $Z_N \times Z_N$ in the above two cases. When N is odd, this is obvious. If $N = 2^k$, consider $(x, y) \in Z_N \times Z_N$. For $0 < j < N$, let h_j denote the

highest power of 2 that divides j, that is, $j/2^{h_j}$ is an odd integer. If $h_x \geqslant h_y$, then we put $a = 1$ and $c = -2^{h_x - h_y}(y/2^{h_y})^{-1}$ where the inverse is evaluated in Z_N and we assume that $y \neq 0$. Then $(x, y) = (ct, -at)$ for $t = -y$. If $h_x < h_y$, then put $c = 1$ and $a = -2^{h_y - h_x}(x/h_x)^{-1}$. We have therefore shown that in all of these cases the Radon transforms together can be inverted, for from the values $\hat{W}(\mu_1, \mu_2)$ so obtained, we can take the inverse Fourier transform and the last assertion of the theorem is proved. □

We can thus recover any distribution function $W(x, z : \rho)$ and consequently the state ρ from the Radon transform data that are in turn probability distributions of measurement in appropriate bases (see (53)). The theorem shows the existence of an inverse transform corresponding to the set of Radon transforms of W, each corresponding to an element M in the group $SL(2, Z_N)$. But we do not need all of the Radon transforms. What is an optimal subset $Q \subset SL(2, Z_N)$ that suffices to determine the state uniquely from probability distributions corresponding to measurements in appropriate bases? This question can only be satisfactorily answered in the context of prior information about the state. One can show that without any such information the cardinality of Q is $O(N)$. Even then, we have a lot of freedom. We can use our choices so as to ensure optimal measurement. Recall from theorem 5 that the Radon transforms are given by probability distribution (corresponding to a state ρ) in the basis that diagonalizes the unitary operator $X^c Z^d$. The only condition imposed on the pair $(c, d) \in Z_N \times Z_N$ is that $\gcd(c, d, N) = 1$. We can often compute this basis explicitly. Then we can use quantum circuits to transform our original 'computational basis' to the required basis. A criterion for the choice of (c, d) could be those that minimize the size of the circuit. For example, if $N = 6$, the choice $c = 3, d = 2$ leads to a particularly simple basis. The analysis becomes simpler if the dimension N is a prime power. We aim to address these issues in future.

5.2. Distribution functions and quantum information

In this section, we discuss some potential applications of distribution functions in QIP. This is a developing area and we only sketch how our formalism may prove useful in various areas in QIP. For this it is best to view the distribution function as coefficients in the expansion of the state in some orthonormal basis in the space of operators, in particular, the basis consisting of phase-point operators. First we generalize to automorphism groups of the *group algebra* $C(\mathbf{H}_N)$: a linear isomorphism $T : C(\mathbf{H}_N) \to C(\mathbf{H}_N)$ such that $T(xy) = T(x)T(y)$ is bijective. It is sufficient to check the last condition for the generators X, Z and γ. We will consider only those automorphism for which $T(\gamma) = \gamma$. Then $T(X)$, $T(Z)$ and γ generate a group isomorphic to $\mathbf{H}_N$ provided $T(X)^N = T(Z)^N = 1$. In particular, if $c \in C(\mathbf{H}_N)$ is invertible then the map $T(x) = cxc^{-1}$ is an automorphisms satisfying these conditions. Such automorphisms are called *inner*. Further call an inner automorphism unitary if $c^{-1} = c^*$ (see (39) for the definition of the * operation). We can prove the following.

Proposition 3. *If $T(x) = cxc^{-1}$ is a unitary inner automorphism and ϕ is representation of $\mathbf{H}_N$, then there is a unitary operator U_c such that $\phi(T(x)) = U_c \phi(x) U_c^{-1}$. Conversely, for any unitary operator U on the representation space of $\mathbf{H}_N$, there is a unitary inner automorphism T_U such that $U\phi(x)U^{-1} = \phi(T_U(x))$. Thus there is a one-to-one correspondence between the set $\mathcal{U}(\mathbf{H}_N)$ of unitary inner automorphisms on $C(\mathbf{H}_N)$ and quantum dynamics on the representative Hilbert space.*

This result is neither difficult nor surprising given the fact that the $\mathbf{H}_N$ completely characterizes the kinematics of the system. It does, however, give us an alternative description and algebraic tools to study the dynamics. Thus, we can study the effect of unitary operations on distribution functions [16] using these transformations. Note, however, that we allow *reducible* representations now. The set of automorphisms of the *group* $\mathbf{H}_N$ is a subgroup of $\mathcal{U}(\mathbf{H}_N)$.

In this work, we have concentrated on irreducible representations of $\mathbf{H}_N$ in which γ acts maximally. By dropping the last assumption, we can obtain all finite-dimensional representations. The order of $\phi(\gamma)$ in the representation ϕ is the dimension. We can then use the products of these representations (actually we need some extra structures) for studying unitary gates. We aim to explore this in future. Let us note some interesting relations in the case $N = 2^n$. If $u \in \mathbf{H}_N$, we will denote by ϕ_k the representation in which $\gamma^{2^k} = 1$. Let σ_i, $i = 1, 2, 3$, denote the Pauli matrices and I_r the identity matrix of order r. Then

$$\phi_1(X) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \sigma_1, \quad \phi_1(Z) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \sigma_3,$$
$$\phi_2(X) = C\sigma_1 \otimes \sigma_1, \quad \phi_2(Z) = \sigma_3 \otimes S,$$
$$\text{where } C = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \quad \text{and} \quad S = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}$$

are a CNOT gate (C) and the phase gate (S), respectively [35]. We also note that in the general case, $\phi_n(X)$ is the cyclic shift operator. It can be efficiently constructed, for example, using full adder circuits with $n+1$ ancillary qubits. Similarly, $\phi_n(Z)$ can be constructed using appropriate controlled phase gates as in the QFT. We also observe that iterating the simple relations $\phi_k(X^2) = \phi_{k-1}(X) \otimes I_2$ and $\phi_k(Z^2) = I_2 \otimes \phi_{k-1}(X)$, we obtain the interesting relations

$$\phi_n(X^{2^k}) = \phi_{n-k}(X) \otimes I_{2^k} \quad \text{and} \quad \phi_n(Z^{2^k}) = I_{2^k} \otimes \phi_{n-k}(Z). \tag{57}$$

These relations can be used to devise more efficient implementations.

We conclude this section with a discussion of potential applications of these constructions to *quantum process tomography* [36]. A quantum process is characterized by a completely positive map T acting on the operators on the system Hilbert space. If we have a complete set of phase-point operators $\{A(x, z)\}$, then T is determined by its action on these. Let us assume that the dimension is odd so that we have a set of phase-point operators satisfying the full set of marginal conditions. Using theorem 5, we can prove the following.

Proposition 4. *Let T be a quantum process (a CP map) given by*

$$T(A(x, z)) = \sum_{x', z'} T(x', z' : x, z) A(x', z').$$

Here, $T(x', z' : x, z)$ is the 'matrix' of T in the basis $\{A(x, z)\}$ of phase-point operators.

$$M^{-1} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{and} \quad M'^{-1} = \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} \in SL(2, Z_N).$$

Then writing $\mathbf{u} = (u_1, u_2) = (x, z)$ *and* $\mathbf{u}' = (u_1', u_2') = (x', z')$,

$$\sum_{(M'^{-1}\mathbf{u}')_1,(M^{-1}\mathbf{u})_1} T(M'^{-1}\mathbf{u}' : M^{-1}\mathbf{u}) = \left\langle z' + \frac{c'd'}{2} \middle| T\left(\left| z + \frac{cd}{2} \right\rangle \left\langle z + \frac{cd}{2} \right| \right) \middle| z' + \frac{c'd'}{2} \right\rangle.$$

Here, $\{|z\rangle\}$ *and* $\{|z'\rangle\}$ *denote the ordered basis of eigenvectors of* $\sigma_M(Z)$ *and* $\sigma_{M'}(Z)$, *respectively.*

We do not prove it here since it is similar to the proof given in theorem 5. Note that we are averaging over two indices now. To use the theorem, we apply T to the projections $|z\rangle\langle z|$ and measure the result in the basis $\{|z'\rangle\}$. These transition probabilities yield the right-hand side in the above equations. In principle, these equations can be inverted using the inverse Radon transforms (see section 5.1) to yield the coefficients $T(x', z' : x, z)$ and thus determining T (see [37] for a different perspective on phase-space tomography). Several optimizations are possible, especially if we have some prior knowledge of the process. But we do not discuss these issues here as they merit a separate investigation.

6. Discussion

In this work, we have analysed quasi-probability distribution functions corresponding to quantum states. Our viewpoint is that these are the real coefficients of bases (or generally frames) in the space of Hermitian operators. The choice of these bases is dictated by certain conditions we impose. This leads to expressing these bases or collection of phase-point operators in terms of operators representing the Weyl–Heisenberg groups. In the language of the frame theory [10, 11], these operators generate the Weyl–Heisenberg frames. We do not go into the intricacies of frame theory approach here. Our approach is more group-theoretic, emphasizing the role of Weyl–Heisenberg groups in quantum kinematics. The other groups that play an important role are $SL(2, Z_N)$, which yield the marginals. Conversely, we can use distribution functions to study these groups. We have given explicit formulae for the Radon transforms and their inversions. These can be used to solve the problem state or operator reconstruction. Even when we do not have sufficient data on marginals to invert the transforms, we can obtain partial information about the state by taking *generalized* inverses [28, 29]. We aim to address these and other issues on state and process estimation and reconstruction, including the practical and computational aspects, in the future.

Appendix

Proof. [Proof of proposition 1] We prove only (11). The invariance of scalar product and the standard measure on $\mathbb{R}^2$ under rotation implies that

$$\begin{aligned}
\int W_c(x, z : \rho)\mathrm{d}x' &= \int \mathrm{d}x' \int \langle \mathrm{e}^{\mathrm{i}(u\hat{x}+v\hat{z})} \rangle \, \mathrm{e}^{-\mathrm{i}(ux+vz)} \, \mathrm{d}u \, \mathrm{d}v \\
&= \int \mathrm{d}x' \int \langle \mathrm{e}^{\mathrm{i}(u'\hat{x}'+v\hat{z}')} \rangle \, \mathrm{e}^{-\mathrm{i}(u'x'+v'z')} \, \mathrm{d}u' \, \mathrm{d}v' \\
&= \int \langle \mathrm{e}^{\mathrm{i}(u'\hat{x}'+v\hat{z}')} \rangle \, \mathrm{e}^{-\mathrm{i}(u'x'+v'z')} \delta(u') \, \mathrm{d}u' \, \mathrm{d}v' \\
&= \int \langle \mathrm{e}^{\mathrm{i}v'\hat{z}'} \rangle \, \mathrm{e}^{-\mathrm{i}v'z'} \, \mathrm{d}v'.
\end{aligned}$$

Let $|z'\rangle$ be the eigenvectors of $\hat{z}'$ with eigenvalues z'. Then

$$\int_a^b \mathrm{d}z' \langle \mathrm{e}^{\mathrm{i}v\hat{z}'} \rangle = \int_a^b \mathrm{d}z'\, \mathrm{e}^{\mathrm{i}vz'} \langle z'|\rho|z'\rangle$$

and equation (11) follows. □

Proof. [Proof of proposition 2] We will prove the second formula only. The proof is similar for the first formula. Using the induced automorphism given in (40), we obtain

$$\begin{aligned}
&\sum_{x'\in Z_n} W_0(M^{-1}x', M^{-1}z' : \rho) \\
&\quad = \frac{1}{N^2} \sum_{x'\in Z_N} \sum_{m,n} \omega^{(1-\delta_{am+cn,0})(1-\delta_{bm+dn,0})((a+b)m+(c+d)n)^2 N/4} \omega^{(am+cn)(bm+dn)/2} \\
&\qquad \times \omega^{-(abm(m-1)/2+cdn(n-1)/2+cdmn)} \langle \sigma_M(X)^m \sigma_M(Z)^n \rangle \omega^{-(mx'+nz')} \\
&\quad = \frac{1}{N} \sum_{x'\in Z_N} \sum_{m,n} \omega^{(1-\delta_{am+cn,0})(1-\delta_{bm+dn,0})((a+b)m+(c+d)n)^2 N/4} \omega^{(am+cn)(bm+dn)/2} \\
&\qquad \times \omega^{-(abm(m-1)/2+cdn(n-1)/2+(m+n)(cd+\mathrm{sgn}(cd)/2))} \langle \sigma_M(X)^m \sigma_M(Z)^n \rangle \omega^{-nz'} \delta_{m0} \\
&\quad = \frac{1}{N} \sum_{n} \omega^{((c+d)n)^2 N/4} \omega^{(cn)(dn)/2} \omega^{-\mathrm{sgn}(cd)n/2-cdn(n-1)/2} \langle \sigma_M(Z)^n \rangle \omega^{-nz'}.
\end{aligned}$$

Here, we use the fact that $cn, dn \neq 0 \bmod N$ for any odd n since N/u is even. We have to consider the two cases separately; suppose first that cd is odd. Then

$$\begin{aligned}
\sum_{x'\in Z_n} W_0(M^{-1}x', M^{-1}z' : \rho) &= \frac{1}{N} \sum_j \rho_{jj} \left(\sum_{r=0}^{N/2-1} \omega^{(cd/2+j-z')(2r)} \right. \\
&\quad \left. + \sum_{r=0}^{N/4-1} (-1)^{d\lfloor \frac{c(2r+1)}{N} \rfloor + c\lfloor \frac{d(2r+1)}{N} \rfloor} \omega^{(cd/2+j-z')(2r+1)} + \text{comp. conj.} \right) \\
&= (\langle z' - cd/2|\rho|z' - cd/2\rangle + \langle N/2 + z' - cd/2|\rho|N/2 + z' - cd/2\rangle)/2 \\
&\quad + \frac{2}{N} \sum_j \rho_{jj} \sum_{\substack{n \text{ odd} \\ n < N/2}} (-1)^{d\lfloor \frac{cn}{N} \rfloor + c\lfloor \frac{dn}{N} \rfloor} \cos \frac{2\pi(cd/2 + j - z')}{N}.
\end{aligned}$$

We can prove the second case (cd even) similarly. □

References

[1] Wigner E P 1932 *Phys. Rev.* **40** 749
[2] Shiryayev A N 1984 *Probability* (Berlin: Springer)
[3] Moyal J E 1949 *Proc. Camb. Phil. Soc.* **45** 99
[4] Stratonovich R L 1957 *Sov. Phys. JETP* **4** 891
[5] Rotman J J 1994 *An Introduction to the Theory of Groups* 4th edn (Berlin: Springer)
[6] Baker G A 1958 *Phys. Rev.* **109** 2198

[7] Wootters W K 1987 *Ann. Phys.* **176** 1
[8] Vourdas A 2004 *Rep. Prog. Phys.* **67** 267
[9] Gabor D 1946 *J. IEE* **93** 429
[10] Chistensen O 2003 *An Introduction to Frames and Riesz Bases* (Basel: Birkhäser)
[11] Ferrie C and Emerson J 2009 *New. J. Phys.* **11** 1
[12] Durt T, Englert B G, Bengtsson I and Życzkowski K 2010 *Int. J. Quantum Inf.* **8** 535
[13] Leonhardt U 1995 *Phys. Rev. Lett.* **74** 4101
[14] Leonhardt U 1997 *Measuring the Quantum State of Light* (Cambridge: Cambridge University Press)
[15] Perelomov A 1984 *Generalized Coherent States and Applications* (Berlin: Springer)
[16] Miquel C, Paz J P and Saraceno M 2002 *Phys. Rev.* A **65** 062309
[17] Paz J P, Roncaglia A J and Saraceno M 2005 *Phys. Rev.* A **72** 012309
[18] Gibbons K S, Hoffman M J and Wootters W K 2004 *Phys. Rev.* A **70** 062101
[19] Weyl H 1950 *The Theory of Groups and Quantum Mechanics* (New York: Dover)
[20] Schwinger J 1970 *Quantum Kinematics and Dynamics* (New York: Benjamin)
[21] Appleby D M 2005 *J. Math. Phys.* **46** 052107
[22] Terras A 1999 *Fourier Analysis on Finite Groups* (Cambridge: Cambridge University Press)
[23] Folland G B 1989 *Harmonic Analysis in Phase Space* (Princeton, NJ: Princeton University Press)
[24] Gross D 2006 *J. Math. Phys.* **47** 122107
[25] Leonhardt U 1996 *Phys. Rev.* A **53** 2998
[26] Chaturvedi S, Mukunda N and Simon R 2010 *J. Phys.* A **43** 075302
[27] Diaconis P and Graham R L 1985 *Pac. J. Math.* **118** 323
[28] Fill J A 1989 *SIAM. J. Discrete Math.* **2** 262
[29] Velasquez E 1997 *Pac. J. Math.* **177** 369
[30] Arthurs E, J L and Kelley J 1965 *Bell Syst. Tech. J.* **44** 725
[31] Braunstein S L, Caves C M and Milburn G J 1991 *Phys. Rev.* A **43** 1153
[32] Hillery M, O'Connell R, Scully M O and Wigner E P 1984 *Phys. Rep.* **106** 121
[33] Serre J P 1977 *Linear Representations of Finite Groups* (Berlin: Springer)
[34] Mordell L J 1969 *Diophantine Equations* (New York: Academic)
[35] Nielsen M and Chuang I 2000 *Quantum Computation and Quantum Information* (Cambridge: Cambridge University Press)
[36] Mohseni M, Rezakhani T and Lidar D A 2008 *Phys. Rev.* A **77** 032322
[37] Paz J P, Roncaglia A J and Saraceno M 2004 *Phys. Rev.* A **69** 032312