

Gaussian Quantum States in Discrete Phase Space

by

Zachary Zanussi

A thesis submitted to
the Faculty of Graduate and Postdoctoral Affairs
in partial fulfillment of
the requirements for the degree of
Master of Science

School of Mathematics and Statistics
Ottawa-Carleton Institute for Mathematics and Statistics
Carleton University
Ottawa, Ontario, Canada

©Copyright
2019, Zachary Zanussi

Abstract

Gaussian states play a central role in continuous quantum mechanics, and an appropriate generalization might play a similar role in discrete quantum mechanics. A common-sense discretization of these turns out to be trivial. In this work, we describe two new discretizations based on recent developments in discrete phase space methods. We also extend existing notions of convergence between continuous and discrete systems and apply some of them to our constructions. Accompanying this work is a numerical software package `qWeyl` which we use to give examples and provide visualizations.

Acknowledgements

First and foremost I thank my supervisor Jason Crann for providing guidance and support throughout the project; you have certainly been an excellent teacher who always made sure to make time for me. You are surely at the beginning of a long and productive career, and I'm grateful to have been a part of it. Thanks for sending me to Corsica, an amazing experience I will never forget, and to Regina, where I met brilliant mathematicians from all over Canada. I would also like to thank my committee members, Michel Barbeau and Anne Broadbent. Thanks to everyone at the Ottawa Quantum Group for being so welcoming.

I am very fortunate to have a wide support network of friends and family. I am grateful to the grads and undergrads at both Carleton and uOttawa Math for stimulating conversations and help unwinding with Tetris or Duck Game. To my friends back home in Red Deer and here in Ottawa, thanks for listening to me blather on about whatever cool math thing I read that week. Thanks to the Ravens Fencing Team for letting me blow off steam, and thanks to Eli Sukunda for teaching me patience. Thanks to Sharon for always helping me keep sight of the bigger picture. Thanks to my parents (and Alexandra) for always encouraging me to be my best, and supporting me through almost two decades of school.

Contents

Abstract	ii
Acknowledgements	iii
1 Introduction	1
1.1 Introduction	1
1.2 qWeyl	2
2 Preliminaries	5
2.1 Hilbert Spaces	5
2.2 Locally Compact Abelian Groups	6
2.3 Gaussianity in \mathbb{R}	8
2.3.1 Density and Characteristic Functions	9
2.3.2 Bernstein's Characterization for Gaussianity	10
2.4 Probability in a Locally Compact Abelian Group	11
3 Harmonic Analysis on Phase Space	15
3.1 Hamiltonian Mechanics	15
3.2 Heisenberg Groups	16
3.2.1 The Heisenberg Group of \mathbb{R}^n	16
3.2.2 The Heisenberg Group of a Locally Compact Abelian Group	21
3.2.3 The Weyl Representation is Irreducible	22
3.2.4 The Stone-von Neumann Theorem	23
3.3 Metaplectic Unitaries	23
3.3.1 The Symplectic Group of \mathbb{R}^n	23
3.3.2 The Metaplectic Representation of $\mathrm{Sp}(\mathbb{R}^n)$	25
3.3.3 Construction of Metaplectics in a Finite Abelian Group	26
3.4 Continuous Gaussian States	31
3.4.1 Coherent States	31
3.4.2 Gibbs Formalism	36
3.4.3 The Wigner Function	44
4 Approaches to Discrete Gaussian States	47
4.1 Hudson states	47
4.2 Bernstein States	55

4.3	The Harper Functions	61
4.4	Discrete Gibbs Formalism	63
4.5	Comparisons	68
4.5.1	Hudson States are not in the Gibbs Framework	68
5	Convergence	74
5.1	Inductive Resolutions	74
5.2	The Harper Operator	78
5.2.1	Harper Convergence	78
6	Conclusion	81
	Appendices	81
A	Gaussian Integrals	82
B	Code to Generate Figures	83

List of Figures

- 2.1 Plot of Gaussian Probability Distributions 8
- 2.2 Plot of the characteristic function of a Gaussian distribution 10

- 3.1 Plot of some continuous minimal uncertainty states 33
- 3.2 Plot of some continuous number states 34
- 3.3 Plot of the characteristic function of a Gaussian state 37
- 3.4 The Wigner function of a sum of two minimal uncertainty states and its marginal distributions 45
- 3.5 The nonnegative Wigner function of a mixture of minimal uncertainty states and its marginal distributions 46

- 4.1 Plot of some discrete minimal uncertainty states 58

- 5.1 Plot that displays how the Harper functions induce the Hermite-Gaussians 76

Chapter 1

Introduction

1.1 Introduction

Gaussianity plays a crucial role throughout classical information theory and signal processing. For instance, Gaussian distributions are fundamentally linked to the central limit theorem and maximum entropy principles, and arguably the most important continuous alphabet channel in signal processing is additive Gaussian noise, a.k.a, the Gaussian channel.

Analogous to the classical setting, one of the most important classes of quantum channels is the class of Gaussian channels [34]. These are often used in quantum optics as noise models, for instance, attenuation/amplification of photonic signals through optical fibres, and additive Gaussian noise. Their underlying mathematical structure is governed by Heisenberg groups and metaplectic representations of symplectic groups, which allows phase space techniques to simplify their information theoretic analysis.

By definition, a Gaussian channel is a metaplectic interaction with a Gaussian *state* in the environment. Gaussian states are well studied in continuous quantum mechanics, but despite considerable development of discrete quantum phase space techniques (see, for example, [18]), there is a gap in the literature around Gaussian states in this setting. In this work we aim to fill this gap by providing a general framework for Gaussian states on a phase space defined over a finite abelian group.

Like the distribution, a continuous Gaussian state is typically defined to be one whose characteristic function is the exponential of a positive semidefinite form. We show that imposing this on a finite abelian group is too restrictive; all such distributions are necessarily point masses. To get around this, we discuss two equivalent notions in the continuous setting and introduce new discretizations of these. The first leads to our class of Bernstein Gaussian states, which may be viewed as mixed state analogues of the familiar stabilizer pure states in quantum error correction [19]. The second leads to our class of Gibbs Gaussian states, defined through Gibbs states of Harper's Hamiltonian, which is a discrete analogue of the harmonic oscillator Hamiltonian. Both classes contain, as special cases, the recently introduced minimal uncertainty states for finite-dimensional quantum systems [21], and are related to their continuous cousins through an acceptable notion of convergence.

Another contribution of our work is the creation of the Mathematica package `qWeyl`. It was designed to be a testing ground for computations surrounding the discrete phase space approach to finite-dimensional quantum systems. As examples, explicit Mathematica implementations of our concepts and results are included throughout the thesis. Several new algorithms are also introduced, including the computation of arbitrary metaplectic unitaries in finite dimensions.

Ultimately, this thesis lays the groundwork for a larger program on Gaussianity in quantum information over discrete phase space. In particular, our classes of Gaussian states yield new classes of Gaussian quantum channels, which open the door to several new lines of investigation. It is the author's hope that other researchers will use and expand on our general framework and find it helpful for experimenting and proving results related to quantum mechanics over discrete phase space.

The structure of the thesis is as follows. In Chapter 2, we discuss some basics of probability theory over locally compact abelian groups and different notions of Gaussianity in that setting. In Chapter 3, we review the phase space approach to quantum mechanics, the Stone-von Neumann Uniqueness theorem, the metaplectic representation, the Wigner function, and finally a number of equivalent formulations of Gaussian states. In Chapter 4, we develop our analogues for these over discrete phase space and make a first attempt at comparing them. Last, in Chapter 5, we discuss ways that these analogues can be viewed as approximations of their corresponding continuous objects.

The author's contributions are as follows. In Chapter 4, we provide two new discretizations of Gaussian states, and prove that at least one of them is different from an existing discretization. In Chapter 5, we provide a new characterization of convergence that we prove is equivalent to one in the literature.

We assume the reader is familiar with the basics of functional analysis and Hilbert space theory.

1.2 `qWeyl`

Much of finite-dimensional quantum mechanics can be modeled using linear algebra over the complex numbers, which makes it amenable to computations using computer algebra software such as Wolfram's `Mathematica`. Most of the material discussed in this thesis has been coded up in a `Mathematica` package, called `qWeyl`. This package is an extension of one released by Timo Felbinger in 2001, called `qmatrix`¹. This package keeps track of the subsystem each object lives in, and automates standard operations including outer product, tensor product, and inner product. This package was written in `Mathematica` Version 4.0; as we are now using Version 10.0, there are some functions that run with warnings that can mostly be ignored.

Accompanying each section will be a short display of the code corresponding to the topics in that section. In this way, the thesis could also be considered partial documentation for the package, including both theoretical ideas and how to implement them. It is the author's hope that other researchers will find this code useful, and moreover, continue to develop it

¹Timo Felbinger's "qmatrix" package: <https://library.wolfram.com/infocenter/MathSource/1893>

into a robust quantum information software suite that can be used to test hypotheses and aid in teaching and understanding. To this end, the package `qWeyl` has been released on the software repository `github`²; anyone interested is encouraged to clone the repository and continue to develop it.

Mathematica is a functional programming language, which gives it a natural feel for pure mathematicians. The basics are easy enough to teach oneself in an afternoon, and one can find great interactive tutorials on the Wolfram site³.

We outline the basics of working with the `qWeyl` package. Once installed, the package must be loaded at the beginning of every session using the line

```
In[1]:= <<qWeyl‘
```

This line will also load in the `qmatrix` package automatically. Once loaded, one can initialize the quantum system they will be working in. The relevant code has been wrapped up into the function `qInit`; the following initializes a tripartite system with dimensions 3, 6, and 9, respectively.

```
In[2]:= qInit[3,{3,6,9}]
```

```
Out[2]= {q1,q2,q3}
```

The size of each subsystem is stored in the global variable `nlist`; one obtains, for example, the size of the second subsystem using

```
In[3]:= nlist[[2]]
```

```
Out[3]= 6
```

We can define state vectors in each system; they are represented as a `matrix` object, which stores an array in the first coordinate and the subsystem in which the object lives in the second. Here we have a few functions from the package; `unitKet[q,n]` generates the n th standard basis vector on the q th subsystem, and `vjKet[q,n]` gives the Fourier transform of the same. The function `randomKet[sys]` takes a pointer to the subsystem `sys`; we can use the wrapper `q[i]` to point to the i th subsystem.

```
In[4]:= unit1 = unitKet[1,1]
        vec2 = vjKet[2,2]
        ran3 = randomKet[q[3]]
```

```
Out[4]=  $\begin{pmatrix} 0. \\ 1. \\ 0. \end{pmatrix}$ 
        {ket[q1]}
```

²“qWeyl” package repository: <https://github.com/zaczanussi/qWeyl>

³Getting Started with Mathematica: <https://www.wolfram.com/language/fast-introduction-for-math-students/en/get-started/>

```
Out[5]=
  (
    0.408248
   -0.204124+0.353553 i
   -0.204124-0.353553 i
    0.408248
   -0.204124+0.353553 i
   -0.204124-0.353553 i
  )
{ket[q2]}
```

```
Out[6]=
  (
    0.147949-0.370189 i
    0.28976-0.19919 i
   -0.166233-0.339271 i
   -0.0200751+0.202744 i
   -0.0587958+0.193107 i
   -0.324968-0.0433902 i
    0.227235+0.26594 i
    0.295425-0.313997 i
   -0.264932+0.0808351 i
  )
{ket[q3]}
```

The operator `**` automatically detects the systems for each `matrix` object and applies the appropriate operation, such as tensor product, inner product, or outer product;

```
In[7]:= unit1**vec2
        unit1**hc[unit1]
        hc[ran3]**ran3
```

```
Out[7]=
  (
    0.
    0.
    0.
    0.
    0.
    0.
    0.408248
   -0.204124+0.353553 i
   -0.204124-0.353553 i
    0.408248
   -0.204124+0.353553 i
   -0.204124-0.353553 i
    0.
    0.
    0.
    0.
    0.
    0.
  )
{ket[q1], ket[q2]}
```

```
Out[8]=
  (0. 0. 0.)
  (0. 1. 0.)
  (0. 0. 0.)
{ket[q1], bra[q1]}
```

```
Out[9]=
  1.+0. i
  {}
```

For more examples of the above, one should see the documentation and tutorials for `qmatrix`.

Chapter 2

Preliminaries

2.1 Hilbert Spaces

We assume basic familiarity with a Hilbert space \mathcal{H} , but we will mention some of the more intricate details, especially those which arise when the dimension of \mathcal{H} is infinite.

An *operator* on \mathcal{H} is a linear map $A : \mathcal{H} \rightarrow \mathcal{H}$. We can define a norm on the set of operators, which is called the *operator norm*,

$$\|A\| = \sup_{\psi} \|A\psi\|_2. \quad (2.1)$$

Here, the supremum is taken over all unit vectors $\psi \in \mathcal{H}$. We denote the norm in \mathcal{H} by $\|\cdot\|_2$ because this is the norm that we will use in our applications. Any operator A with $\|A\| < \infty$ is called *bounded*, and the set of bounded operators on \mathcal{H} is denoted $\mathcal{B}(\mathcal{H})$.

If \mathcal{H} is separable then we may fix an orthonormal basis $\{|e_k\rangle\}_{k \in \mathbb{N}}$. Here we use *Dirac's Bra-ket notation* for vectors in \mathcal{H} . Given an operator A on \mathcal{H} , we may define another norm, called the *trace norm*, by

$$\|A\|_1 = \sum_{k=1}^{\infty} \langle e_k | (A^*A)^{1/2} | e_k \rangle. \quad (2.2)$$

If this quantity exists and is finite, then we say that A is *trace-class*. Note that in this case, the value $\|A\|_1$ will be independent of choice of basis. The set of trace-class operators is denoted by $\mathcal{T}(\mathcal{H})$. In a finite dimensional \mathcal{H} , this class of operators coincides with $\mathcal{B}(\mathcal{H})$, but in infinite dimensions they are not the same. A positive semidefinite operator ρ with trace norm 1 is called a *state* or a *density operator*. These two conditions on ρ imply that its eigenvalues will be nonnegative and sum to 1, hence the terminology “density.”

An operator U is called *unitary* if it is surjective and preserves the inner product on \mathcal{H} : that is, for all $\psi, \varphi \in \mathcal{H}$,

$$\langle U\psi | U\varphi \rangle = \langle \psi | \varphi \rangle. \quad (2.3)$$

It follows that $UU^* = U^*U = I$, where $*$ denotes the *adjoint*, and I is the identity on \mathcal{H} . We denote the set of unitaries on \mathcal{H} by $U(\mathcal{H})$. This space has a natural topology called the *strong operator topology*; a sequence $\{U_k\}$ in $U(\mathcal{H})$ converges to U if and only if $\|U\psi - U_k\psi\| \rightarrow 0$ for every $\psi \in \mathcal{H}$.

A special type of state is of the form $\rho = |\psi\rangle\langle\psi|$ for some unit vector $\psi \in \mathcal{H}$. This ρ defines a rank-one projection onto the subspace of \mathcal{H} spanned by ψ . We often call a unit vector ψ a *state vector*, with the understanding that it can define a state ρ . We call a state of the form ρ *pure*, otherwise we say it is *mixed*.

Denote by $D(A)$ the domain of an operator A . An operator $A : D(A) \subseteq L^2(\mathbb{R}^n) \rightarrow L^2(\mathbb{R}^n)$ is called *essentially self-adjoint* if A has a unique self-adjoint extension to $L^2(\mathbb{R}^n)$.

Quantum mechanics models a quantum system as a Hilbert space \mathcal{H} and quantum states as states ρ inside it. To measure a state, we act on it with an *observable* \hat{o} , which is a self-adjoint operator on \mathcal{H} . In general, \hat{o} may be unbounded.

As mechanics is the study of how systems change with time, given an initial state we need to be able to compute how it evolves. We define an operator \hat{H} , called the *Hamiltonian*, that encodes the evolution of a closed system. Here, *closed* means that the system does not interact with the environment around it; one can always “zoom out” from an open system until we can see the entire, closed, system. There are different ways to consider how a system evolves; one, called the *Schroedinger picture*, involves assuming that all observables are fixed, and that the states change with time. The other, called the *Heisenberg picture*, assumes that the states stay fixed and the observables evolve. Both are equivalent, and it is convenient to consider either in different cases. We will only need to use the Heisenberg picture, however, so we will review it here.

Suppose \hat{H} is the Hamiltonian for the system we are considering. We demand that it be self-adjoint; this is a requirement for Stone’s Theorem which we will see in section 3.2.1. If \hat{o} is an observable, then its dynamics over time in system governed by \hat{H} is given by

$$\frac{d}{dt}\hat{o} = i[H, \hat{o}]. \quad (2.4)$$

The operator \hat{o} on a closed system evolves unitarily according to $\hat{o} \mapsto e^{i\hat{H}t}\hat{o}e^{-i\hat{H}t}$.

2.2 Locally Compact Abelian Groups

Definition 2.2.1. A *topological group* is a set G that has the structure of a group and of a Hausdorff topological space such that the group operations $(x, y) \mapsto xy$ and $x \mapsto x^{-1}$ are continuous (where the topology on $G \times G$ is the product topology). If G is locally compact, it is called a *locally compact group*.

Definition 2.2.2. Let G be a topological group. A left (respectively, right) *Haar measure* on G is a nonzero regular Borel measure μ that satisfies $\mu(xA) = \mu(A)$ (respectively, $\mu(Ax) = \mu(A)$) for every Borel set $A \subseteq G$ and every $x \in G$.

Natural examples of locally compact groups include the real numbers \mathbb{R} with the Lebesgue measure, or any discrete group with the counting measure. An important subclass of the latter are *finite abelian groups*, which are necessarily isomorphic to a product of cyclic groups.

Theorem 2.2.3. *Every locally compact group possesses a left and a right Haar measure which are unique up to positive multiplicative constants.*

For $1 \leq p < \infty$, we denote by $L^p(G)$ the *Lebesgue space* of (equivalence classes of) p -integrable complex-valued functions on G . Denote by $L^\infty(G)$ the space of equivalence classes of essentially bounded complex-valued functions on G .

Definition 2.2.4. Let G be a locally compact abelian group and let \mathbb{T} denote the group of complex numbers of modulus one. A *character* on G is a continuous homomorphism $\chi : G \rightarrow \mathbb{T}$.

The set of characters on a locally compact abelian group G , denoted \widehat{G} , forms an abelian group under pointwise multiplication, and is called the *dual group* of G . Given the topology of compact convergence, the dual group becomes a locally compact abelian group such that each $y \in G$ induces a character $\Phi(y)$ on \widehat{G} via $\Phi(y)(\chi) = \chi(y)$. In fact, these are the only characters:

Theorem 2.2.5 (Pontryagin Duality Theorem). *Let G be a locally compact abelian group. The mapping $G \ni y \mapsto \Phi(y) \in \widehat{\widehat{G}}$ is a homeomorphism of topological groups.*

We give a some examples of dual groups. If $G = \mathbb{R}$, then $\chi \in \widehat{\mathbb{R}}$ has $\chi(1) = \exp(ix)$ for some $x \in \mathbb{R}$, and the homomorphism property forces $\chi(t) = \exp(itx)$, for all $t \in \mathbb{R}$. Thus we can identify χ with x , so that $\widehat{G} \cong \mathbb{R}$. We say that \mathbb{R} is *self-dual*. One can use a similar argument to show that \mathbb{R}^n is self-dual for all $n \in \mathbb{N}$. If $X = \mathbb{Z}$, we have $\chi(1) = z$ for some $z \in \mathbb{T}$, which forces $\chi(n) = z^n$ for all $n \in \mathbb{Z}$. Thus, we can identify χ with z and say that $\widehat{\mathbb{Z}} \cong \mathbb{T}$. Similarly, we can see that $\widehat{\mathbb{Z}_m} \cong \mathbb{Z}_m$ so that any $\chi \in \widehat{\mathbb{Z}_m}$ is of the form $\chi(l) = \exp(2\pi ikl/m)$ for some $k \in \mathbb{Z}_m$. Thus \mathbb{Z}_m is also self-dual.

If X is a locally compact Hausdorff space, we denote by $C_0(X)$ the space of continuous functions on X vanishing at infinity. For a locally compact abelian G , fix a left Haar measure $dy := d\mu_G(y)$ on G and consider $L^1(G)$. The *Fourier transform* is the mapping $\mathcal{F} : L^1(G) \rightarrow C_0(\widehat{G})$ given by

$$\mathcal{F}(f)(\chi) = \int_G \overline{\chi(y)} f(y) dy, \quad f \in L^1(G), \quad \chi \in \widehat{G}. \quad (2.5)$$

There is a natural product on $L^1(G)$ called *convolution*: if $f, g \in L^1(G)$, then we define

$$(f * g)(x) = \int_G f(y)g(x - y)dy. \quad (2.6)$$

The Fourier transform sends the convolution product in $L^1(G)$ to the pointwise product in $C_0(\widehat{G})$: indeed,

$$\begin{aligned} (F)(f * g)(\chi) &= \int_G \overline{\chi(x)}(f * g)(x)dx \\ &= \int_G \overline{\chi(x)} \int_G f(y)g(x - y)dydx. \end{aligned}$$

Making the substitution $z = x - y$, so that $dz = dx$, we have

$$= \int_G \int_G f(y)\overline{\chi(y + z)}g(z)dydz$$

$$\begin{aligned}
&= \int_G f(y)\bar{\chi}(y)dy \int_G \bar{\chi}(z)g(z)dz \\
&= \mathcal{F}(f)(\chi)\mathcal{F}(g)(\chi).
\end{aligned}$$

Let G be a locally compact group. A *representation* of G is a homomorphism π from G into the group $U(\mathcal{H})$ of unitary operators on some nonzero Hilbert space \mathcal{H} that is continuous with respect to the strong operator topology. That is, a map $\pi : G \rightarrow U(\mathcal{H})$ that satisfies $\pi(xy) = \pi(x)\pi(y)$ and $\pi(x)^{-1} = \pi(x^{-1}) = \pi(x)^*$ for all $x, y \in G$, and for which $x \mapsto \pi(x)\psi$ is a continuous map from G to \mathcal{H} for any $\psi \in \mathcal{H}$. A representation π is called *faithful* if the map $\pi : G \rightarrow U(\mathcal{H})$ is injective.

Given two representations π_1 and π_2 of G , an *intertwining operator* for π_1 and π_2 is a bounded linear map $T : \mathcal{H}_1 \rightarrow \mathcal{H}_2$ such that $T\pi_1(x) = \pi_2(x)T$ for all $x \in G$. The set of all such operators is denoted by $C(\pi_1, \pi_2)$. Two representations π_1 and π_2 are called *equivalent* if $C(\pi_1, \pi_2)$ contains a unitary operator U , so that $\pi_2(x) = U\pi_1(x)U^{-1}$. We write $C(\pi)$ for $C(\pi, \pi)$.

A closed subspace V of \mathcal{H} is called *invariant* for π if $\pi(x)V = \{\pi(x)\psi \mid \psi \in V\} \subseteq V$ for all $x \in G$. If π admits a nontrivial (that is, proper or nonzero) invariant subspace, then π is called *reducible*, otherwise π is *irreducible*. The following lemma provides a necessary and sufficient condition for irreducibility.

Lemma 2.2.6 (Schur's Lemma). *A representation π of G is irreducible if and only if $C(\pi)$ contains only scalar multiples of the identity.*

2.3 Gaussianity in \mathbb{R}

The *Gaussian distribution*, often called *normal*, is ubiquitous in every field that utilizes statistical methods, including physics (Gaussian wave packets), engineering (signal noise), sociology (many characteristics of a population, such as height, are assumed to be normally distributed), and more. It is commonly referred to as “the bell curve” due to its distinctive shape: a few of these are plotted in Figure 2.1. If X is a random variable whose probability density function is a Gaussian distribution, then we call X *Gaussian* or a *normal deviate*.

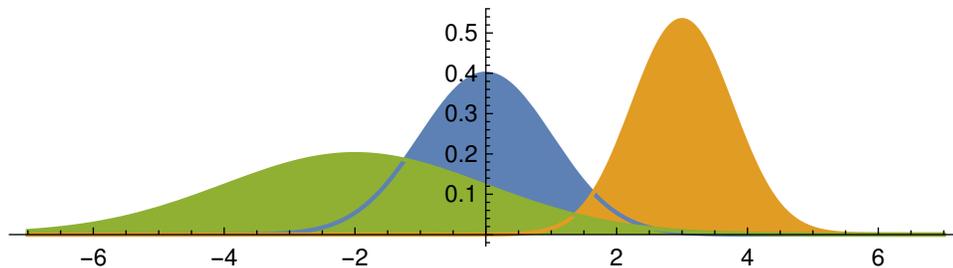


Figure 2.1: Three “bell curves,” or Gaussian probability density functions. In blue is the standard normal, with mean $\mu = 0$ and variance $\sigma^2 = 1$. In green is a normal distribution with $\mu = -2$, $\sigma^2 = 2$, and in orange is $\mu = 2$, $\sigma^2 = 0.75$. In all three cases, the shaded area under the curve is equal to 1.

Many equivalent characterizations of Gaussianity have been studied. We will review each of the notions relevant to the present thesis.

2.3.1 Density and Characteristic Functions

A density function γ on \mathbb{R} is said to be *Gaussian* with mean μ and variance σ^2 if it is of the form

$$\gamma(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right). \quad (2.7)$$

The mean μ determines where the top of the *bell* lies; if $\mu = 0$ we say that γ is *centered*. The variance σ^2 determines the width of the curve; as σ^2 grows, the curve gets wider and shorter, while if it gets small, the curve grows taller and leaner.

One can also consider the limiting case as the variance σ^2 goes to zero. In this case, the probability density $\gamma(x)$ tends to zero for any $x \neq \mu$ but grows ever larger at $x = \mu$ as its integral must remain equal to one. We can define “loosely” the Dirac delta function δ_μ as follows:

$$\delta_\mu(x) = \begin{cases} +\infty, & x = \mu \\ 0, & x \neq \mu. \end{cases} \quad (2.8)$$

That is, δ_μ is supported only at the mean μ , where it is a very large spike. We say this is “loosely” defined because there is no function that has this property. One can define δ_μ in a rigorous way as a distribution or a measure, in such a way that it makes sense to say “the integral of the delta function is one.” However we will not need this level of rigor; for us it is only important to note that as the variance approaches zero the Gaussian distribution “approaches” a point mass.

One can easily compute the characteristic function φ of a Gaussian distribution γ with mean μ and variance σ^2 directly, using the well-known formula A.1:

$$\begin{aligned} \varphi(t) &= \hat{\gamma}(t) \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{itx} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \exp\left(\frac{-x^2 + (2\mu + 2it\sigma^2)x - \mu^2}{2\sigma^2}\right) dx \\ &= \frac{\sqrt{2\pi\sigma^2}}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{2\sigma^2(2\mu - 2it\sigma^2)^2}{4(2\sigma^2)^2} - \frac{\mu^2}{2\sigma^2}\right) \\ &= \exp(it\mu - \frac{1}{2}t^2\sigma^2). \end{aligned}$$

We now have our second equivalent notion of Gaussianity in \mathbb{R} : any distribution with characteristic function of the form

$$\varphi(t) = \exp(it\mu - \frac{1}{2}t^2\sigma^2) \text{ for some } \mu, \sigma^2. \quad (2.9)$$

Such a characteristic function is plotted in Figure 2.2.

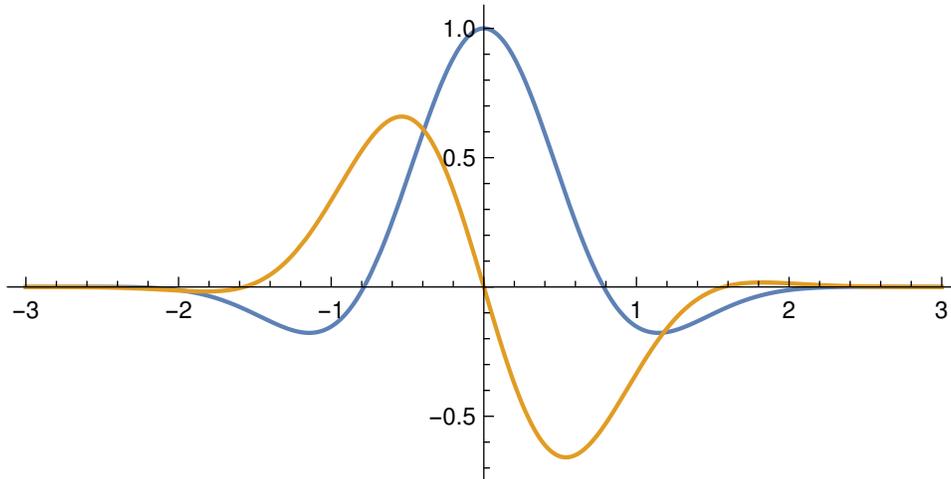


Figure 2.2: The characteristic function of a Gaussian distribution with mean $\mu = -2$ and variance $\sigma^2 = 2$. The blue line is the real part, and the yellow is the imaginary.

We now pause to mention a useful simplification especially powerful for theoretical work. Suppose X is a Gaussian deviate with mean μ and variance σ^2 . If we define a new random variable Z by $Z = \frac{X-\mu}{\sigma}$, then Z is a Gaussian deviate with mean 0 and variance 1. We call such a random variable Z the *standard normal deviate*. In particular, given a statement involving a normal deviate X with mean μ we can often instead consider $X' = X - \mu$ and thus prove the result for a centered distribution X' .

2.3.2 Bernstein's Characterization for Gaussianity

The ubiquity of the Gaussian distribution makes it one of the most well-studied distributions on the real line. While many of the characterizations are natural, some are rather surprising at first. Bernstein's Theorem is an example of such a characterization.

Theorem 2.3.1 (Bernstein's Theorem). *Suppose X_1 and X_2 are independent real-valued random variables with finite means. Set $Y_1 = X_1 + X_2$ and $Y_2 = X_1 - X_2$. If Y_1 and Y_2 are independent, then X_1 and X_2 are normally distributed with the same variance.*

This was first proved by Sergei Bernstein in 1941 ([7]), however one struggles to find a copy of this publication that is translated out of its native Russian. One can see a more modern treatment of this at [29].

Bernstein's theorem states that, given two independent random variables whose sum and difference are also independent, then the random variables must be Gaussian. The converse of this statement is more well known:

Proposition 2.3.2. *Suppose X_1 and X_2 are independent Gaussian random variables with means μ_1, μ_2 and variances σ_1^2, σ_2^2 , respectively, and suppose $\sigma_1^2 \geq \sigma_2^2$. Then $Y_1 = X_1 + X_2$ and $Y_2 = X_1 - X_2$ are also Gaussian.*

Proof. By independence, the joint characteristic function ϕ_{Y_1} is the product of the characteristic functions ϕ_{X_1} and ϕ_{X_2} . Then

$$\phi_{Y_1}(t) = \phi_{X_1}(t)\phi_{X_2}(t) = \exp\left(it(\mu_1 + \mu_2) - \frac{(\sigma_1^2 + \sigma_2^2)t^2}{2}\right).$$

By uniqueness, Y_1 is Gaussian with mean $\mu_1 + \mu_2$ and variance $\sigma_1^2 + \sigma_2^2$.

For Y_2 , we recall that we have not defined a notion of *negative* variance; thus the requirement $\sigma_1^2 \geq \sigma_2^2$. The proof in this case proceeds identically. \square

By virtue of these two results, we may consider Bernstein's characterization an equivalent condition for Gaussianity: two independent identically distributed random variables are Gaussian if and only if their sum and difference are independent. Thus we have a third equivalent condition for Gaussianity, which we call *Gaussian in the sense of Bernstein*.

2.4 Probability in a Locally Compact Abelian Group

Let G be a locally compact abelian group. We would like to study probability distributions on G . Let $\mathcal{B}(G)$ be the σ -algebra of Borel sets on G , that is, the smallest σ -algebra of subsets of G which contains all the open subsets of G . We call a nonnegative countable additive set function $\mu : \mathcal{B}(G) \rightarrow \mathbb{R}$ a *measure* on G . We call a measure μ a *distribution* if $\mu(G) = 1$. The *support* $\text{supp}(\mu)$ of a distribution μ is the smallest closed subset $A \subseteq G$ such that $\mu(A) = \mu(G)$.

We define the *convolution* $\mu * \nu$ of two distributions μ, ν by

$$(\mu * \nu)(B) = \int_G \mu(B - x)d\nu(x), \quad B \in \mathcal{B}(G), \quad (2.10)$$

where $B - x = \{g - x | g \in B\}$ is the set B *shifted* by $x \in G$. Denote by δ_x the *degenerate distribution* concentrated at x , that is,

$$\delta_x(A) = \begin{cases} 1, & x \in A \\ 0, & \text{otherwise.} \end{cases}$$

We note that the convolution $\mu * \delta_x$ of a distribution μ implements the *shift* of μ : that is, for $B \in \mathcal{B}(G)$,

$$(\mu * \delta_x)(B) = \int_G \mu(B - y)d\delta_x(y) = \mu(B - x).$$

Let Ω be a set, \mathcal{A} be a σ -algebra of subsets of Ω , and p be a probability measure on \mathcal{A} , that is, a countably additive function $p : \mathcal{A} \rightarrow [0, 1]$ with $p(\Omega) = 1$. The triple (Ω, \mathcal{A}, p) is called a *probability space*. By a *random variable* on (Ω, \mathcal{A}, p) with values in G we understand a function $\xi : \Omega \rightarrow G$ such that $\xi^{-1}(B) \in \mathcal{A}$ for any Borel subset $B \subseteq G$. The random variable ξ defines a distribution μ_ξ on $\mathcal{B}(G)$ via the formula

$$\mu_\xi(B) = p(\{\omega \in \Omega | \xi(\omega) \in B\}), \quad \text{for } B \in \mathcal{B}(G).$$

Two random variables ξ and η with values in G are called *independent* if the equality

$$p(\{\omega|\xi(\omega) \in A \text{ and } \eta(\omega) \in B\}) = p(\{\omega|\xi(\omega) \in A\})p(\{\omega|\eta(\omega) \in B\})$$

holds for all $A, B \in \mathcal{B}(G)$.

For a distribution μ on G , define the *characteristic function* $\widehat{\mu} : \widehat{G} \rightarrow \mathbb{C}$ to be the Fourier transform of μ , that is,

$$\widehat{\mu}(\chi) = \int_G \overline{\chi(x)} d\mu(x). \quad (2.11)$$

We are now ready to discuss Gaussian distributions on G :

Definition 2.4.1 (Gaussian Distribution). A distribution γ on a group G is called *Gaussian* if its characteristic function can be written in the form

$$\widehat{\gamma}(\chi) = \chi(x) \exp(-\varphi(\chi)), \quad \chi \in \widehat{G}, \quad (2.12)$$

where $x \in G$ and $\varphi(\chi)$ is a continuous nonnegative function on the group \widehat{G} satisfying the equation

$$\varphi(\chi_1 + \chi_2) + \varphi(\chi_1 - \chi_2) = 2(\varphi(\chi_1) + \varphi(\chi_2)), \text{ for } \chi_1, \chi_2 \in \widehat{G}. \quad (2.13)$$

We denote by $\Gamma(G)$ the set of Gaussian distributions on G . A Gaussian distribution is called *centered* if $x = 0$ in Equation 2.12. We denote by $\Gamma^C(G)$ the set of centered distributions on G .

As an example, if we set $G = \mathbb{R}^n$, then $\widehat{G} \cong \mathbb{R}^n$, and one can see that a continuous nonnegative function φ on \mathbb{R}^n satisfying 2.13 is of the form

$$\varphi(y) = y^\top A y, \quad y = (y_1, \dots, y_n) \in \mathbb{R}^n,$$

where A is a symmetric positive definite matrix. In particular, if $\gamma \in \Gamma(\mathbb{R})$, then the characteristic function can be written

$$\widehat{\gamma}(t) = \exp(it\mu - \frac{1}{2}t^2\sigma^2),$$

for $\mu, t \in \mathbb{R}$, $\sigma \geq 0$. This agrees with the characteristic function we calculated in 2.9.

We have seen that, on the real line, a distribution γ is Gaussian if and only if its characteristic function has the form given in 2.9. It is clear by properties of the exponential that the support of γ is the entire real line, that is, $\text{supp}(\gamma) = \mathbb{R}$. It turns out that Gaussianity imposes a fairly strict condition on the support of a distribution, as described in the following proposition, adapted from [15, Prop 3.6]. Recall that, in general, we may restrict our attention to centered distributions.

Proposition 2.4.2. *Let G be a locally compact abelian group and let γ be a centered distribution with characteristic function φ of the form given in 2.12. Then the support of γ is contained in some connected subset A of G , that is, $\text{supp}(\gamma) \subseteq A$.*

This proposition is weaker than the one given by Feldman but it encodes all that we are interested in. If G is a finite abelian group then it is discrete, so that the only connected subsets are singletons. If γ is a distribution with a Gaussian characteristic function, then by Proposition 2.4.2 its support is connected, which means that there exists $x \in G$ such that $\gamma = \delta_x$. Thus, the characteristic function definition for Gaussianity is trivial in this setting.

In a locally compact abelian group G , a *Haar measure* m_G has the properties that

$$m_G(B + x) = m_G(B) \text{ for all } x \in G, B \in \mathcal{B}(G), \text{ and } m_G(-B) = m_G(B) \text{ for all } B \in \mathcal{B}(G).$$

If G is compact, then $m_G(G) < \infty$, so in this case we may assume that m_G is a distribution. If G is finite abelian, the translation invariance condition implies that m_G is uniform: that is, $m_G(x) = 1/|G|$ for all $x \in G$.

Suppose K is a compact subgroup of G . If m_K is the Haar distribution on K , then the characteristic function of m_K is of the form

$$\widehat{m_K}(\chi) = \begin{cases} 1, & \text{if } \chi(x) = 1 \text{ for all } x \in K \\ 0, & \text{otherwise.} \end{cases} \quad (2.14)$$

A distribution μ on G is called *idempotent* if $\mu * \mu = \mu$. In this case, $\mu = m_K$ for some compact K in G [15, Section 2.14].

We now turn to the Bernstein characterization. We have seen that this is another equivalent notion in the real case, but we now seek to understand what this condition imposes on a general locally compact abelian group G . This has been studied in [15, Chapter 7]. In this context, Proposition 2.3.2 takes the following form: for a Gaussian distribution γ , if ξ_1 and ξ_2 are two independent random variables with values in G and distribution γ , then $\xi_1 + \xi_2$ and $\xi_1 - \xi_2$ are independent. We denote the set of all idempotent distributions with this property by $I_B(G)$. That is, $m_K \in I_B(G)$ if ξ_1 and ξ_2 are independent random variables with values in G and distribution m_K , then $\xi_1 + \xi_2$ and $\xi_1 - \xi_2$ are independent.

Set $G^{(2)} = \{2x | x \in G\}$. A group G is said to be a *Corwin group* if $G^{(2)} = G$. That is, a Corwin group is one where, for every $x \in G$, there is a $y \in G$ such that $x = 2y$: that is, x is a multiple of 2. Some examples of Corwin groups include \mathbb{R} and \mathbb{Z}_{2k+1} for every $k \in \mathbb{N}$. Indeed, any subgroup of \mathbb{Z}_{2k+1} is Corwin. Non-examples of Corwin groups include \mathbb{Z} and \mathbb{Z}_{2k} for any k ; in both of these cases, the element 1 is not a multiple of 2. One can show [15, Proposition 7.4] that for a compact subgroup K of G , K is Corwin if and only if $m_K \in I_B(G)$. We quote the following [15, Theorem 7.10]:

Theorem 2.4.3. *Let G be a locally compact abelian group with no elements of order 2. Let ξ_1 and ξ_2 be independent random variables with values in G and distributions μ_1 and μ_2 such that $\xi_1 + \xi_2$ and $\xi_1 - \xi_2$ are independent. Then $\mu_j \in \Gamma(G) * I_B(G)$ and $\mu_1 = \mu_2 * \delta_x$ for some $x \in G$.*

We have already seen that if G is finite abelian then $\gamma \in \Gamma(G)$ implies $\gamma = \delta_x$ for some $x \in G$. We have also seen that every element of $I_B(G)$ is of the form m_K for some Corwin subgroup K of G , and since G is finite, $m_K(x) = \frac{\chi_K(x)}{|K|}$, where χ_K is the indicator function on K . Thus, we can classify the set of distributions in $\Gamma(G) * I_B(G)$ as all μ of the form $\mu = \delta_x * m_K$ for a Corwin subgroup K . That is, μ is a shift of a uniform distribution over a

Corwin subgroup. Since convolution by a degenerate distribution corresponds to a shift, we may as well write

$$\mu = \frac{\chi_{K-x}}{|K|}, \quad (2.15)$$

We phrase this as a definition:

Definition 2.4.4 (Gaussian in the Sense of Bernstein). Let G be a finite abelian group. A distribution μ is *Gaussian in the sense of Bernstein* if it is of the form

$$\mu = \frac{\chi_{K+x}}{|K|}$$

for some Corwin subgroup K of G .

We give an example. Let $G = \mathbb{Z}_d \times \mathbb{Z}_d$ for some odd d . Since \mathbb{Z}_d is Corwin, G and every subgroup of G is Corwin. In particular, we define the *line* with slope $m \in \mathbb{Z}_d$ to be the set $L_m = \{(q, qm) \mid q \in \mathbb{Z}_d\}$. It is clear that L_m is a subgroup of G . The line with slope m and intercept (a, b) , defined by

$$L_m^{(a,b)} = \{(q + a, qm + b) \mid q \in \mathbb{Z}_d\}.$$

is a shift of L_m , so that a distribution of the form

$$\mu = \frac{1}{d} \chi_{L_m^{(a,b)}}$$

is Gaussian in the sense of Bernstein. This will be relevant later in Section 4.2.

Chapter 3

Harmonic Analysis on Phase Space

In this chapter we review the fundamentals of the Heisenberg and symplectic groups and Weil’s framework of metaplectic unitaries. We begin with a review of Hamiltonian mechanics.

3.1 Hamiltonian Mechanics

Many are familiar with Newtonian mechanics, or at least have heard of Newton’s second law, which essentially relates the forces on a particle to its motion using the second-order ordinary differential equation $\vec{F} = m \frac{d^2}{dt^2} \vec{d}$. Using this, once the forces on the particle are known, one can completely determine the position and velocity of the particle at a given time. In this framework, one focuses on *forces*, and characterizes the “state” of the particle using its position and velocity. One can instead focus on the total *energy* of the system, and consider how the system evolves with time without considering explicitly the forces acting on the particle. This change of mindset is the cornerstone of *Hamiltonian mechanics*.

Let us consider a single particle with mass m traveling in n dimensions. Rather than considering position and velocity, we find it preferable to specify the state of a particle using its position $q = (q_1, \dots, q_n)$ and its *momentum* $p = (p_1, \dots, p_n)$, which we view as points in \mathbb{R}^n , henceforth referred to as the *configuration space*. We then characterize the state of the particle at any time using a *phase space* \mathbb{R}^{2n} with coordinates

$$(p, q) = (p_1, \dots, p_n, q_1, \dots, q_n).$$

We can then model the trajectory of the particle as a curve $\phi : \mathbb{R} \rightarrow \mathbb{R}^{2n}$, so that for any time t we get a vector $\phi(t)$ consisting of the particle’s current position and momentum.

The central object of Hamiltonian mechanics is the *Hamiltonian operator* $\hat{H} = \hat{H}(p, q)$, which represents the total energy in a system. In a closed system, it is calculated as the sum of the potential and kinetic energy at any time, denoted V and T , respectively. Let us consider a pet example, the *harmonic oscillator*. The harmonic oscillator is a model for any system that experiences a restoring force proportional to its displacement from its equilibrium point, such as a pendulum, or a mass on a spring. Let us consider this latter example, a mass m on a frictionless tabletop connected to a spring and displaced so that it oscillates horizontally in one dimension about an equilibrium $q = 0$. At any time, it has a kinetic energy $T = 1/2m\dot{q}^2$, where $\dot{q} = \frac{d}{dt}q$ is the time derivative of position q , and its potential

energy is given by Hooke's law, so that if the spring has constant k , we have $V = 1/2kq^2$. The Hamiltonian \hat{H} of this system is the sum of these, however it must be given as a function of q and p , whereas T as we have written it is a function of \dot{q} . We can fix this by noting that $p = m\dot{q}$. So the Hamiltonian of this very simple system is given by

$$\hat{H} = T + V = \frac{p^2}{2m} + \frac{kq^2}{2}. \quad (3.1)$$

The Hamiltonian of a system is all one needs to solve for its motion. This motion in phase space is governed by *Hamilton's equations*:

$$\frac{dp}{dt} = -\frac{\partial H}{\partial q} \quad \text{and} \quad \frac{dq}{dt} = \frac{\partial H}{\partial p}. \quad (3.2)$$

If n is the number of degrees of freedom (we have been working with $n = 1$ in this example) then Hamilton's equations are a system of $2n$ equations that give the trajectory of the curve ϕ .

Associated to a Hamiltonian system is a *Poisson bracket* $\{\cdot, \cdot\}$, an anti-symmetric, bilinear form. In many cases, such as the configuration space \mathbb{R}^n , the Poisson bracket is simply the *symplectic form*,

$$[(p, q), (p', q')] = pq' - qp',$$

where $pq = \sum_{i=1}^n p_i q_i$ is the dot product. Any set of coordinates that satisfies the following relations is called *canonical*:

$$\{q_i, p_j\} = \delta_{ij}, \quad \{q_i, q_j\} = 0, \quad \{p_i, p_j\} = 0. \quad (3.3)$$

Thus, any transformation that preserves these relations is a permissible coordinate change. Such transformations may also be called *canonical*; we will call them *symplectic*.

Quantities that one might hope to observe about the state of the system are called *observables*, modeled as real-valued functions on the phase space. The simplest observables one might consider involve sampling the position or momentum of the particle at a given time; these observables are modeled as projections from the phase space onto the corresponding coordinate in configuration space. In fact, every observable may be modeled as a function of position and momentum.

3.2 Heisenberg Groups

3.2.1 The Heisenberg Group of \mathbb{R}^n

The Heisenberg group of \mathbb{R}^n arises from investigating the Poisson bracket relations 3.3 from an abstract algebra perspective. Indeed, we fix a natural n and consider \mathbb{R}^{2n+1} with coordinates

$$(p_1, \dots, p_n, q_1, \dots, q_n, t) = (p, q, t).$$

Define a Lie bracket on \mathbb{R}^{2n+1} by

$$[(p, q, t), (p', q', t')] = (0, 0, t + t' + [(p, q), (p', q')]), \quad (3.4)$$

where the bracket on the right hand side is the usual symplectic form on \mathbb{R}^{2n} . This makes \mathbb{R}^{2n+1} into a Lie algebra known as the *Heisenberg algebra*, denoted \mathfrak{h}_n . Letting $P_1, \dots, P_n, Q_1, \dots, Q_n, T$ be the standard basis for \mathbb{R}^{2n+1} , one can witness the Lie algebra structure as

$$[P_j, P_k] = [Q_j, Q_k] = [P_j, T] = [Q_j, T] = 0, \quad [P_j, Q_k] = \delta_{jk}T, \quad (3.5)$$

which is exactly the structure we have seen in the Poisson bracket.

It is most convenient to identify \mathfrak{h}_n and to view its corresponding Lie group using the following matrix representations. For $(p, q, t) \in \mathbb{R}^{2n+1}$, we define the matrix $m(p, q, t) \in M_{n+2}(\mathbb{R})$ by

$$m(p, q, t) = \begin{pmatrix} 0 & p_1 & \dots & p_n & t \\ 0 & 0 & \dots & 0 & q_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & q_n \\ 0 & 0 & \dots & 0 & 0 \end{pmatrix}, \quad (3.6)$$

and set

$$M(p, q, t) = I + m(p, q, t).$$

It is easy to see the following composition laws:

$$\begin{aligned} m(p, q, t)m(p', q', t') &= m(0, 0, pq') \\ M(p, q, t)M(p', q', t') &= M(p + p', q + q', t + t' + pq'), \end{aligned}$$

from which we can verify that the commutator satisfies

$$[m(p, q, t), m(p', q', t')] = m(0, 0, pq' - qp').$$

So indeed, the map $X \mapsto m(X)$ is a Lie algebra isomorphism from \mathfrak{h}_n to $\{m(X) \mid X \in \mathbb{R}^{2n+1}\}$, and thus to obtain the corresponding Lie group we may simply apply the matrix exponential map. Since $m(p, q, t)^k = 0$ for $k \geq 3$, we have

$$\exp(m(p, q, t)) = I + m(p, q, t) + \frac{1}{2}m(0, 0, pq) = M(p, q, t + \frac{1}{2}pq). \quad (3.7)$$

Hence,

$$\exp(m(p, q, t)) \exp(m(p', q', t')) = \exp\left(m(p + p', q + q', t + t' + \frac{1}{2}[(p, q), (p', q')])\right), \quad (3.8)$$

so if we identify $X \in \mathbb{R}^{2n+1}$ with the matrix $\exp(m(X))$, we have that \mathbb{R}^{2n+1} is a group with law

$$(p, q, t)(p', q', t') = (p + p', q + q', t + t' + \frac{1}{2}[(p, q), (p', q')]). \quad (3.9)$$

We call this group the *Heisenberg group* and denote it by $H(\mathbb{R}^n)$. One can see from the group relation that the inverse of an element (p, q, t) is $(-p, -q, -t)$.

Sometimes it is favorable to identify (p, q, t) with the matrix $M(p, q, t)$, which yields the group law

$$(p, q, t)(p', q', t') = (p + p', q + q', t + t' + pq'). \quad (3.10)$$

\mathbb{R}^{2n+1} with this group law is called the *polarized Heisenberg group*, denoted $H^{pol}(\mathbb{R}^n)$. The inverse of an element (p, q, t) in $H^{pol}(\mathbb{R}^n)$ is $(-p, -q, -t + pq)$.

From either group relation one observes that

$$\mathcal{Z} = \{(0, 0, t) \mid t \in \mathbb{R}\} \quad (3.11)$$

is the center of $H(\mathbb{R}^n)$ and $H^{pol}(\mathbb{R}^n)$.

The Heisenberg group $H(\mathbb{R}^n)$ is faithfully represented on $L^2(\mathbb{R}^n)$ by the Weyl representation, which we now outline. For $j = 1, \dots, n$, $\xi = (\xi_1, \dots, \xi_n) \in \mathbb{R}^n$, and $\psi \in L^2(\mathbb{R}^n)$, we consider the operators \hat{q}_j, \hat{p}_j given by

$$(\hat{q}_j \psi)(\xi) = \xi_j \psi(\xi), \quad (\hat{p}_j \psi)(\xi) = \frac{\hbar}{i} \frac{\partial}{\partial \xi_j} \psi(\xi), \quad (3.12)$$

called *position* and *momentum* respectively. These are the *canonical observables*, completely analogous to those given in classical mechanics. The positive \hbar is *Planck's constant*, which is necessary in applications but can be set to 1, as is common in mathematical physics. In general these operators are unbounded, however they behave well if we restrict their domain.

A natural restriction is to the Schwartz space $\mathcal{S}(\mathbb{R}^n)$, the subspace of infinitely differentiable functions $C^\infty(\mathbb{R}^n)$ which are rapidly decreasing in the sense that the function and all of its derivatives decrease faster than any power of ξ . As a set, we write

$$\mathcal{S}(\mathbb{R}^n) = \{f \in C^\infty(\mathbb{R}^n) \mid \|f\|_{\alpha, \beta} := \sup_{\xi \in \mathbb{R}^n} |\xi^\alpha \frac{\partial^\beta}{\partial \xi^\beta} f(\xi)| < \infty, \text{ for all multi-indices } \alpha, \beta\}. \quad (3.13)$$

An example of such a function is $f(x) = \xi^k e^{-a\|\xi\|^2}$ for any multi-index k and positive a : one can see this either by direct computation or by noting that any derivative of f contains a factor of $e^{-a\|\xi\|^2}$ which will clearly decrease faster than any power of ξ .

Since by definition $\|f\|_{\alpha, 0} < \infty$ for $f \in \mathcal{S}(\mathbb{R}^n)$ and any multi-index α , we have that $\xi^\alpha f(\xi)$ is always bounded. Further, for any $1 \leq p \leq \infty$, $f \in \mathcal{S}(\mathbb{R}^n)$ implies $f \in L^p(\mathbb{R}^n)$. Indeed, given such an f , we see that

$$\int |f(\xi)|^p d\xi = \int ((1 + |\xi|^2)|f(\xi)|)^p \frac{1}{(1 + |\xi|^2)^p} d\xi \quad (3.14)$$

Now $(1 + |\xi|^2)|f(\xi)|$ is bounded since f is Schwartz, and for $p \geq 1$ $\frac{1}{(1 + |\xi|^2)^p} \leq \frac{1}{1 + |\xi|^2} \in L^1(\mathbb{R})$, so

$$\int |f(\xi)|^p dx \leq \|(1 + |\cdot|^2)|f(\cdot)|\|_\infty^p \int \frac{1}{1 + |\xi|^2} d\xi < \infty \quad (3.15)$$

Thus $\|f\|_p$ is finite so $f \in L^p(\mathbb{R}^n)$. This justifies that $\mathcal{S}(\mathbb{R}^n)$ is a subspace of $L^2(\mathbb{R}^n)$, and in fact it is dense [17, Page 241].

It is immediately clear that \hat{q}_j, \hat{p}_j map $\mathcal{S}(\mathbb{R}^n)$ to itself, and on this space they are self-adjoint. On the Schwartz space these operators satisfy the Heisenberg Canonical Commutation Relations (CCR):

$$[\hat{q}_j, \hat{p}_k] = i\hbar \delta_{jk}, \quad [\hat{q}_j, \hat{q}_k] = 0, \quad [\hat{p}_j, \hat{p}_k] = 0, \quad (3.16)$$

where the right-hand side of the first equation is viewed as a scalar multiple of the identity operator. Indeed, for any $\psi \in \mathcal{S}(\mathbb{R}^n)$, if $j \neq k$ these operators clearly commute, and if $j = k$ then

$$\begin{aligned} ([\hat{q}_j, \hat{p}_j]\psi)(\xi) &= \hat{q}_j(\hat{p}_j\psi)(\xi) - \hat{p}_j(\hat{q}_j\psi)(\xi) \\ &= \xi(\hat{p}_j\psi)(\xi) - \frac{\hbar}{i} \frac{d}{d\xi}(\hat{q}_j\psi)(\xi) \\ &= \xi \frac{\hbar}{i} \frac{d}{d\xi} \psi(\xi) - \frac{\hbar}{i} \frac{d}{d\xi} \xi \psi(\xi) \\ &= \xi \frac{\hbar}{i} \frac{d}{d\xi} \psi(\xi) - \frac{\hbar}{i} (\psi(\xi) + \xi \frac{d}{d\xi} \psi(\xi)) = (i\hbar I)\psi(\xi) \end{aligned}$$

These relations should remind the reader of the Heisenberg algebra relations 3.5, and indeed we will use these operators to construct a representation of the group $H(\mathbb{R}^n)$. We first recall a fundamental result of Stone on the one-to-one correspondence between representations of \mathbb{R}^n and collections of essentially self-adjoint operators [23, Theorem 12.2].

Theorem 3.2.1 (Stone's Theorem). *Given a representation $\{U(\xi)\}_{\xi \in \mathbb{R}^n}$ on $L^2(\mathbb{R}^n)$, there exists essentially self-adjoint operators A_1, \dots, A_n such that*

$$U(\xi) = \exp\left(i \sum_{j=1}^n \xi_j A_j\right). \quad (3.17)$$

Conversely, given essentially self-adjoint operators A_1, \dots, A_n , Equation 3.17 defines a representation on $L^2(\mathbb{R}^n)$.

One can view this as a generalization of the following basic fact: a real number a can be associated with a point u on the unit circle via $u = e^{ia}$, and vice versa, given an exponential function $u(t) = e^{iat}$, one can recover a by taking i times the derivative at the identity, $a = \frac{1}{i} \frac{du}{dt}(0)$. Even further, it is well-known that a self-adjoint matrix has real eigenvalues, and that a unitary matrix has eigenvalues on the unit circle. Then if A has spectral decomposition $A = \sum_{j=1}^n a_j |j\rangle\langle j|$, then $U(t) = \sum_{j=1}^n \exp(ia_j t) |j\rangle\langle j|$. So Stone's theorem is a non-commutative generalization of this relation between the real numbers and the unit circle.

We now compute the representations corresponding to the single variable position and momentum operators \hat{q} and \hat{p} ; we will need to use the Taylor expansion of \exp to interpret what these operators do to vectors. Let $\psi \in \mathcal{S}(\mathbb{R})$, then

$$\begin{aligned} \exp(it\hat{q})\psi(\xi) &= \left(1 + it\hat{q} + \frac{i^2 t^2}{2!} \hat{q}^2 + \dots\right)\psi(\xi) \\ &= \psi(\xi) + it\xi\psi(\xi) + \frac{i^2 t^2}{2!} \xi^2 \psi(\xi) + \dots \\ &= \exp(it\xi)\psi(\xi) \end{aligned}$$

and

$$\exp(it\hat{p})\psi(\xi) = \left(1 + it\hat{p} + \frac{i^2 t^2}{2!} \hat{p}^2 + \dots\right)\psi(\xi)$$

$$\begin{aligned}
&= \psi(\xi) + it \frac{1}{i} \frac{d}{dt} \psi(\xi) + \frac{i^2 t^2}{2!} \frac{1}{i^2} \frac{d^2}{dt^2} \psi(\xi) + \dots \\
&= \psi(\xi + t)
\end{aligned}$$

where the last equality follows from the Taylor expansion of ψ about the point t . We define

$$\hat{Z}_t \psi(\xi) = \exp(it\xi) \psi(\xi) \quad \text{and} \quad \hat{X}_t \psi(\xi) = \psi(\xi + t) \quad (3.18)$$

and call them the *boost* and *shift* operators, respectively. If we now consider \hat{q}_j and \hat{p}_j on $\mathcal{S}(\mathbb{R}^n)$ and define

$$\hat{Z}_v \psi(\xi) = \exp(i\xi v) \psi(\xi), \quad \hat{X}_u \psi(\xi) = \psi(\xi + u),$$

for $u, v, \xi \in \mathbb{R}^n$, we can use a nearly identical calculation as above to show that

$$\hat{Z}_v = \exp\left(i \sum_{j=1}^n v_j \hat{q}_j\right) \quad \hat{X}_u = \exp\left(i \sum_{j=1}^n u_j \hat{p}_j\right).$$

Recall the commutation relations 3.16 (here we set $\hbar = 1$). We would like an equivalent relation between \hat{Z} and \hat{X} . Letting $u, v, \xi \in \mathbb{R}^n$, we have

$$\begin{aligned}
\hat{X}_u \hat{Z}_v \psi(\xi) &= \hat{Z}_v \psi(\xi + u) = \exp(iv(\xi + u)) \psi(\xi + u) = \exp(ivu) \exp(iv\xi) \hat{X}_u \psi(\xi) \\
&= \exp(ivu) \hat{Z}_v \hat{X}_u \psi(\xi).
\end{aligned}$$

It turns out that this commutation relation is equivalent to the one for \hat{q} and \hat{p} : to show this, we need the Baker-Campbell-Hausdorff Formula for \mathfrak{h}_n .

In general, if two elements $A, B \in \mathfrak{h}_n$ do not commute, then the product of their exponentials is not the exponent of their sum, that is, $\exp(A) \exp(B) \neq \exp(A + B)$. In the special case that $[A, B] = c$ for some $c \in \mathbb{C}$, the Baker-Campbell-Hausdorff Formula gives

$$\exp(A) \exp(B) = \exp\left(A + B + \frac{1}{2}[A, B]\right). \quad (3.19)$$

For a more detailed statement and its proof, see [33, Section 7.6].

Returning to the equivalence of the two commutation relations, suppose that the single-variable operators satisfy $[\hat{q}, \hat{p}] = i$. Then for $u, v \in \mathbb{R}$,

$$\begin{aligned}
\hat{X}_u \hat{Z}_v &= \exp(iu\hat{p}) \exp(iv\hat{q}) \\
&= \exp\left(iu\hat{p} + iv\hat{q} + \frac{1}{2}[iu\hat{p}, iv\hat{q}] + \frac{1}{2}([iu\hat{p}, iv\hat{q}] + [iv\hat{q}, iu\hat{p}])\right) \\
&= \exp([iu\hat{p}, iv\hat{q}] + iv\hat{q} + iu\hat{p} + \frac{1}{2}[iv\hat{q}, iu\hat{p}]) \\
&= \exp([iu\hat{p}, iv\hat{q}]) \exp(iv\hat{q}) \exp(iu\hat{p}) \\
&= \exp(iuv) \hat{Z}_v \hat{X}_u.
\end{aligned}$$

Indeed, the other direction is proven in the same way. Since the observables on different variables commute, an identical calculation shows the equivalence in \mathbb{R}^n .

Now we have an alternative framework for continuous phase space; rather than considering self-adjoint operators, we may work with unitaries. This unitary paradigm will be necessary in the next chapter.

The *Weyl Representation* of $H(\mathbb{R}^n)$ is the map $w : H(\mathbb{R}^n) \rightarrow U(L^2(\mathbb{R}^n))$ given by

$$w(u, v, t) = e^{it} \exp\left(\frac{i}{2}uv\right) \hat{Z}_v \hat{X}_u, \quad (3.20)$$

where $u, v \in \mathbb{R}^n$. This is a faithful representation of $H(\mathbb{R}^n)$, which is irreducible (see Section 3.2.3). The restriction of w to $\mathbb{R}^{2n} \cong \mathbb{R}^{2n} \times \{0\}$ gives a projective representation $w : \mathbb{R}^{2n} \rightarrow U(L^2(\mathbb{R}^n))$ defined by

$$w(u, v) = \exp\left(\frac{i}{2}uv\right) \hat{Z}_v \hat{X}_u. \quad (3.21)$$

Note the phase factor on $w(u, v)$; this ensures that it is *involutive*, that is, $w(u, v)^{-1} = w(-u, -v) = w(u, v)^*$.

One can define a version of the Weyl representation on the polarized Heisenberg group $H^{pol}(\mathbb{R}^n)$ by

$$w(u, v, t) = e^{it} \hat{Z}_v \hat{X}_u, \quad (3.22)$$

and similarly for its restriction to \mathbb{R}^{2n} . This representation fails to be involutive, but is still considered in the literature.

3.2.2 The Heisenberg Group of a Locally Compact Abelian Group

We now outline a more general framework for Heisenberg groups on locally compact abelian groups, developed by André Weil [35].

Let G be a locally compact abelian group with Pontryagin dual \hat{G} . We define the (polarized) *Heisenberg group of G* to be $H^{pol}(G) = G \times \hat{G} \times \mathbb{T}$ with multiplication

$$(x_1, \chi_1, \lambda_1) \cdot (x_2, \chi_2, \lambda_2) = (x_1 x_2, \chi_1 \chi_2, \chi_2(x_1) \lambda_1 \lambda_2). \quad (3.23)$$

The inverse of an element (x, χ, λ) is given by $(x^{-1}, \chi^{-1}, \lambda^{-1} \chi(x))$.

We define the *boost* and *shift* operators on $L^2(G)$ as follows: for $\psi \in L^2(G)$,

$$\hat{Z}_\chi \psi(t) = \chi(t) \psi(t), \quad \hat{X}_x \psi(t) = \psi(tx), \quad (3.24)$$

for $t \in G$. The Weyl representation is the map $w : H^{pol}(G) \rightarrow U(L^2(G))$ given by

$$w(x, \chi, \lambda) \psi(t) = \lambda \hat{Z}_\chi \hat{X}_x \psi(t) = \lambda \chi(x) \psi(tx), \quad \psi \in L^2(G), t \in G. \quad (3.25)$$

Identifying the *phase space* $Z_G := G \times \hat{G}$ with the corresponding subgroup of $H^{pol}(G)$, w restricts to a projective representation of the abelian group Z_G , with commutation relation

$$w(z_1) w(z_2) = \Omega(z_1, z_2) w(z_2) w(z_1), \quad z_i = (x_i, \chi_i) \in Z_G, \quad (3.26)$$

where $\Omega : Z_G \times Z_G \rightarrow \mathbb{T}$ is the *symplectic form* given by

$$\Omega(z_1, z_2) = \frac{\chi_2(x_1)}{\chi_1(x_2)}, \quad z_i = (x_i, \chi_i) \in Z_G, \quad (3.27)$$

An automorphism $T \in \text{Aut}(Z_G)$ is said to be *symplectic* if it leaves the symplectic form invariant, that is

$$\Omega(Tz_1, Tz_2) = \Omega(z_1, z_2). \quad (3.28)$$

We denote by $\text{Sp}(G)$ the set of symplectic automorphisms of the phase space Z_G . The group $\text{Sp}(G)$ is intrinsically related to $\text{Aut}_0(H(G))$, the subgroup of $\text{Aut}(H(G))$ which acts trivially on the center $\mathcal{Z}(H(G)) = \mathbb{T}$. Indeed, any such automorphism is given by

$$H(G) \ni (z, \lambda) \mapsto (Tz, \varphi(z)\lambda) \in H(G), \quad (3.29)$$

where $T \in \text{Sp}(G)$ and $\varphi : Z_G \rightarrow \mathbb{T}$ is a *second-degree character* of Z_G , meaning that the function

$$Z_G \times Z_G \ni (z_1, z_2) \mapsto \varphi(z_1 + z_2)\varphi(z_1)^{-1}\varphi(z_2)^{-1} \in \mathbb{T} \quad (3.30)$$

is a bi-character on Z_G .

It follows from [35, Section 10] that for all $S \in \text{Sp}(G)$ there exists a unitary U_S on $L^2(G)$ such that

$$U_S w(z) U_S^* = \varphi(z) w(Sz), \quad (3.31)$$

for some second-degree character φ of Z_G .

3.2.3 The Weyl Representation is Irreducible

An important feature of the Weyl representation is that it is *irreducible*. There are multiple proofs of this fact, and indeed some are more computational and require none of the advanced machinery we are about to introduce. However, this result is particularly slick when placed in the context of operator algebra, so we include the proof using these tools.

First we need to provide the context. We consider the Weyl representation as a map $w : H^{\text{pol}}(G) \rightarrow \mathcal{B}(L^2(G))$. One can embed the space $L^\infty(G)$ into $\mathcal{B}(L^2(G))$ via the weak*-weak* continuous map $f \mapsto M_f$ where $M_f \psi(\xi) = f(\xi)\psi(\xi)$. This is an isometry and thus an injection. We call M_f a *multiplication operator*.

We will need the following fact. Suppose G is a group. For a subset \mathcal{S} of $\mathcal{B}(L^2(G))$, define the *commutant* \mathcal{S}' to be the subspace of $\mathcal{B}(L^2(G))$ consisting of the elements that commute with all elements of \mathcal{S} , i.e.

$$\mathcal{S}' = \{A \in \mathcal{B}(L^2(G)) \mid AT = TA, \forall T \in \mathcal{S}\}.$$

It is well known that $L^\infty(G)' = L^\infty(G)$ for any locally compact abelian G . One can see a proof of this in [8, Theorem 5.3.13].

Proposition 3.2.2 (Irreducibility of the Weyl Representation). *The Weyl representation w of $H^{\text{pol}}(G)$ is irreducible.*

Proof. We use Schur's lemma; suppose $T \in C(w)$, so that T commutes with, in particular, every operator $\hat{Z}_\chi = M_\chi$ for $\chi \in \hat{G}$. Viewing $L^\infty(G)$ as the set of multiplication operators on $L^2(G)$, it follows that the (weak-*) closed linear span of $\{\hat{Z}_\chi \mid \chi \in \hat{G}\}$ is equal to $L^\infty(G)$ [17, Theorem 3.2.7]. Thus, if T commutes with everything in $\{\hat{Z}_\chi \mid \chi \in \hat{G}\}$ then $T \in L^\infty(G)' = L^\infty(G)$, so T is of the form M_f for some $f \in L^\infty(G)$. But T also commutes

with \hat{X}_x for every $x \in G$. Denote by r_x the automorphism of $L^\infty(G)$ given by $r_x f(t) = f(tx)$. We claim that $\hat{X}_x M_f \hat{X}_x^* = M_{r_x f}$. Indeed, if $\psi \in L^2(G)$,

$$(\hat{X}_x M_f \hat{X}_x^*)\psi(t) = M_f \hat{X}_x^* \psi(tx) = \hat{X}_x^* f(tx) \psi(tx) = f(tx) \psi(tx) = r_x f(t) \psi(t) = M_{r_x f} \psi(t).$$

Thus, $M_f = M_{r_x f}$ for all $x \in G$. But we mentioned that the multiplication operators are injective, which means that $f = r_x f$ for all $x \in G$. This clearly implies that f is constant, so $T \in \mathbb{C}I$, and thus we have that the Weyl representation is irreducible. \square

3.2.4 The Stone-von Neumann Theorem

We are now ready to state the Stone-von Neumann theorem, which shows that certain irreducible representations of the Heisenberg group are equivalent.

Theorem 3.2.3 (Stone-von Neumann Uniqueness Theorem). *Let π be a representation of $H(\mathbb{R}^n)$ on a Hilbert space \mathcal{H} such that $\pi(0, 0, t) = e^{it}I$ for all $t \in \mathbb{R}$. If π is irreducible, then π is unitarily equivalent to w .*

Most proofs of this result utilize notions from representation theory that are beyond the scope of this work. One can find a readable version in modern notation based on Folland's [16] exposition of von Neumann's original proof at [28]. In fact, Theorem 3.2.3 holds verbatim when replacing \mathbb{R}^n with any locally compact abelian group G .

The generalized theorem gives us the following: if π is an irreducible representation of $H^{pol}(G)$ such that $\pi(0, 0, \lambda) = \lambda I$ for all $\lambda \in \mathbb{T}$, then there exists a unitary $U \in U(L^2(G))$ such that for every $(x, \chi, \lambda) \in H^{pol}(G)$,

$$U\pi(x, \chi, \lambda)U^{-1} = w(x, \chi, \lambda). \quad (3.32)$$

In particular, let S be some automorphism of $H^{pol}(G)$ that fixes the center. One can show that the map $w \circ S$ is again an irreducible unitary representation of $H^{pol}(G)$. Thus, we are guaranteed the existence of a unitary U_S such that

$$U_S w(x, \chi, \lambda) U_S^{-1} = w(S(x, \chi, \lambda)) \quad (3.33)$$

for all $(x, \chi, \lambda) \in H^{pol}(G)$. In this way, we get a representation of the automorphism S on $L^2(G)$, which is the subject of the next section.

3.3 Metaplectic Unitaries

Equation 3.33 gives an assignment $\text{Aut}_0(H^{pol}(G)) \ni S \mapsto U_S \in U(L^2(G))$ which generates the so-called *metaplectic representation* of $\text{Sp}(G)$.

3.3.1 The Symplectic Group of \mathbb{R}^n

Here we collect some useful facts about the symplectic group $\text{Sp}(\mathbb{R}^n)$, which is a subset of the $2n \times 2n$ real-valued matrices. For more general exposition of symplectic vector spaces

see [12]. Recall that a symplectic matrix is one that preserves the symplectic form, and thus it represents a permissible coordinate change in phase space.

Define the matrix

$$\Omega = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}. \quad (3.34)$$

which implements the *symplectic form* on \mathbb{R}^{2n} :

$$[w_1, w_2] = w_1^\top \Omega w_2. \quad (3.35)$$

Observe that $\Omega^\top = -\Omega = \Omega^{-1}$, where A^\top denotes the transpose throughout this section. We shall frequently refer to the matrix Ω as the *symplectic form* for simplicity.

We typically write symplectic matrices S in block form,

$$S = \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \quad (3.36)$$

where A, B, C and D are $n \times n$ matrices. The *symplectic group* $\text{Sp}(\mathbb{R}^n)$ is the group of $2n \times 2n$ real matrices which preserve the symplectic form: that is,

$$S \in \text{Sp}(\mathbb{R}, n) \iff [Sw_1, Sw_2] = [w_1, w_2] \text{ for all } w_1, w_2 \in \mathbb{R}^{2n}. \quad (3.37)$$

There are several equivalent conditions for S to be symplectic, which we list below:

1. Since

$$[Sw_1, Sw_2] = (Sw_1)^\top \Omega Sw_2 = w_1^\top (S^\top \Omega S) w_2, \quad (3.38)$$

we have that $S^\top \Omega S = \Omega$.

2. Taking the determinant of both sides above, we get

$$\det(\Omega) = \det(S^\top) \det(\Omega) \det(S) \Rightarrow 1 = \det(S^\top) \det(S) \quad (3.39)$$

which shows that $\det(S) = \pm 1$ and, in particular, S is invertible.

3. Taking the transpose of the formula in 1 and rearranging for S^{-1} , we get

$$S^{-1} = \Omega S^\top \Omega^{-1} = \begin{pmatrix} D^\top & -B^\top \\ -C^\top & A^\top \end{pmatrix}. \quad (3.40)$$

4. Writing out the formula in 1 in block matrix form, we get

$$A^\top C = C^\top A, \quad B^\top D = D^\top B, \quad \text{and } A^\top D - C^\top B = I. \quad (3.41)$$

We will end this section by describing how $\text{Sp}(\mathbb{R}^n)$ is generated by the union of three sets of symplectic matrices. The proof of this decomposition is well treated in the literature (see e.g. [16, Proposition 4.10] or [12, Section 2.2.2]) so we will simply state it in the form that is most convenient for our uses:

Theorem 3.3.1 (Generators of the Symplectic Group). *Consider the following subgroups of $\text{Sp}(\mathbb{R}, n)$:*

$$L = \left\{ \begin{pmatrix} A & 0 \\ 0 & (A^\top)^{-1} \end{pmatrix} \mid A\text{-invertible} \right\}, \quad R = \left\{ \begin{pmatrix} I & 0 \\ C & I \end{pmatrix} \mid C = C^\top \right\}. \quad (3.42)$$

Then $\text{Sp}(\mathbb{R}, n)$ is algebraically generated by $L \cup R \cup \{\Omega\}$.

3.3.2 The Metaplectic Representation of $\mathrm{Sp}(\mathbb{R}^n)$

We begin by noting that any symplectic map S determines an automorphism of the Heisenberg group $H(\mathbb{R}^n)$ in the following way:

$$(p, q, t) \mapsto (S(p, q), t). \quad (3.43)$$

This is a homomorphism because S preserves the symplectic form, and invertible because symplectic maps are invertible. Further, by the equation above this automorphism clearly fixes the center of $H(\mathbb{R}^n)$.

Let w be the Weyl representation, where we let $\hbar = 1$. Then the composition $w \circ S$ is an irreducible, unitary representation of $H(\mathbb{R}^n)$ that fixes the center. Thus, by Stone-von Neumann there must exist a unitary U_S that *intertwines* these two representations: that is,

$$U_S w(p, q, t) U_S^{-1} = w(S(p, q), t) \quad (3.44)$$

In quantum mechanics, we always consider unit vectors up to a global scalar phase factor, so that we would not differentiate between a state ψ and, say, $e^{it}\psi$. For this reason it suffices to consider a projective representation of $\mathrm{Sp}(\mathbb{R}^n)$, that is, a homomorphism from $\mathrm{Sp}(\mathbb{R}^n)$ to the quotient group $U(L^2(G))/\mathcal{Z}$, where $\mathcal{Z} = \{cI \mid |c| = 1\}$ is the center of $U(L^2(\mathbb{R}^n))$. To create a bona fide representation, one can pass to the double cover of $\mathrm{Sp}(\mathbb{R}^n)$, the so-called *metaplectic group*, but this is beyond the scope of this thesis.

In general, given a symplectic S it is difficult to calculate its metaplectic U_S . However, for a few basic classes of symplectics it is simple. These classes correspond to those given in the generating set from Theorem 3.3.1. We will compute the metaplectics for a general element in each set.

First, let $S = \begin{pmatrix} A & 0 \\ 0 & (A^\top)^{-1} \end{pmatrix}$ for some invertible A . Then

$$\begin{aligned} [w \circ S(p, q)]f(x) &= w(Ap, (A^\top)^{-1}q)f(x) = \exp(i((A^\top)^{-1}q)x + \frac{i}{2}(Ap)^\top (A^\top)^{-1}q)f(x + Ap) \\ &= \exp(iq^\top A^{-1}x + \frac{i}{2}pq)(f \circ A)(A^{-1}x + p) = U w(p, q) U^{-1} f(x), \end{aligned} \quad (3.45)$$

where $U f(x) = |\det A|^{-\frac{1}{2}} f(A^{-1}x)$. Note that the factor of the determinant ensures that U is unitary.

Next, let $S = \begin{pmatrix} I & 0 \\ C & I \end{pmatrix}$ with $C = C^\top$. Then

$$\begin{aligned} [w \circ S(p, q)]f(x) &= w(p, Cp + q)f(x) = \exp(i(Cp + q)^\top x + \frac{i}{2}p^\top (Cp + q))f(x + p) \\ &= e^{-\frac{i}{2}x^\top Cx} e^{iqx + \frac{i}{2}pq} e^{\frac{i}{2}(x+p)^\top C(x+p)} f(x + p) = U w(p, q) U^{-1} f(x), \end{aligned} \quad (3.46)$$

where $U f(x) = \exp(-\frac{i}{2}x^\top Cx) f(x)$.

Last, let $S = \Omega$. It turns out that the inverse Fourier transform works: define

$$F f(x) = \frac{1}{\sqrt{2\pi}} \int \exp(-ixy) f(y) dy.$$

Then,

$$\begin{aligned}
w(p, q)Ff(x) &= \frac{1}{\sqrt{2\pi}} \exp(iqx + \frac{i}{2}pq) \int \exp(-i(x+p)y) f(y) dy \\
&= \frac{1}{\sqrt{2\pi}} \exp(iqx + \frac{i}{2}pq) \int \exp(-i(x+p)(y+q)) f(y+q) dy \\
&= \frac{1}{\sqrt{2\pi}} \int \exp(-ixy) \exp(-ipy - \frac{i}{2}pq) f(y+q) dy \\
&= Fw(q, -p)f(x)
\end{aligned}$$

We note that this choice of Fourier transform, without 2π in the exponent, is to match the phase factor of the Weyl operators. Indeed, some authors choose to include a factor of 2π in their Weyl operators, in which case they must use a different definition for F . In the next section, our discrete operators will have this factor, so we will introduce the slightly different Fourier transform then.

So, for a given symplectic, the task becomes to decompose it into the generators given in Theorem 3.3.1, and use that to compose the corresponding unitaries. We describe a way to do this in a finite abelian group in the next section, but we will see that this process also generalizes to a general locally compact group (such as \mathbb{R}^n).

3.3.3 Construction of Metaplectics in a Finite Abelian Group

We quote a well-known characterization of finite abelian groups:

Theorem 3.3.2. *Let G be a finite abelian group. Then G can be written in the following generic form:*

$$G \cong \mathbb{Z}_{n_1} \times \cdots \times \mathbb{Z}_{n_d}, \quad \text{where } n_1 | n_2 | \cdots | n_d$$

Let G be a finite abelian group. Given $S \in \text{Sp}(G)$, we are interested in computing its corresponding metaplectic operator U_S . We follow the construction given in [25]. We let G be of the form given in Theorem 3.3.2. We encode the dimensions of G in the operator $N = \text{diag}(n_1, \dots, n_d)$, so that $|G| = \det(N)$, and define an inner product $\langle \cdot, \cdot \rangle : G^2 \rightarrow \mathbb{T}$ by

$$\langle m, k \rangle = \exp(2\pi i m^\top N k). \quad (3.47)$$

We note that this pairing encodes the duality of G : indeed, since a finite abelian group is self-dual, it is possible to find, for any $\chi \in \widehat{G}$, an element $m \in G$ such that $\chi(k) = \langle m, k \rangle$. In this way, we can identify the *phase space* $Z_G = G \times \widehat{G}$ as G^2 .

Let $v \in L^2(G)$ be a state vector. Given $\lambda \in G^2$, we define the Weyl operator $\pi(\lambda)$ for a function $v : G \rightarrow \mathbb{C}$ by

$$\pi(\lambda)v(k) = \langle m, k \rangle v(k-l), \quad \lambda = (l, m) \in G^2, k \in G.$$

One recognizes this as the Weyl representation on the *polarized* Heisenberg group $H^{\text{pol}}(G)$. Because we will later consider the Weyl operators on the Heisenberg group $H(G)$, we choose to denote these operators by π . Note that these operators fail to be involutive.

We must now discuss the structure of the symplectic group on G . One can refer to [22] for a discussion of endomorphisms on G . We summarize the necessary notions here. Given G as above, we can represent the endomorphism group $\text{End}(G)$ of G by equivalence classes of integer matrices. A representative $[A] = (a_{r,s})$ of an endomorphism A satisfies the condition that

$$\frac{n_r}{n_s} \text{ divides } a_{r,s} \text{ if } s < r, \quad r, s = 1, \dots, d.$$

This condition can be written in words as follows: the elements below the diagonal must be divisible by the quotient of the appropriate group factors. This preserves the fact that, for example, if an element g of G is written in the form $g = (g_1, g_2, \dots, g_d)$, the first component g_1 is in \mathbb{Z}_{n_1} ; multiplying it by $\frac{n_j}{n_1}$ gives an equivalent element in \mathbb{Z}_{n_j} . Any other representative $(a'_{r,s})$ satisfies

$$a'_{r,s} = a_{r,s} \pmod{n_r}, \quad r, s = 1, \dots, d.$$

So we can reduce each row r of a representative by n_r . Under these constraints, the endomorphism ring structure is given by the usual matrix operations.

Given an $A \in \text{End}(G)$ with representative $[A]$, we can give a representative for the adjoint A^* as follows:

$$[A]^* = N[A]^\top N^{-1}.$$

Indeed, we see that

$$\begin{aligned} \langle m, Ak \rangle &= \exp(2\pi i m^\top N^{-1} Ak) \\ &= \exp(2\pi i (m^\top N^{-1} AN) N^{-1} k) \\ &= \exp(2\pi i (NA^\top N^{-1} m)^\top N^{-1} k) = \langle A^* m, k \rangle, \quad k, m \in G \end{aligned}$$

Because formulae such as this one do not depend on the particular representative $[A]$, we will no longer distinguish between an endomorphism A and its representative $[A]$.

We denote by $\text{Aut}(G) \subset \text{End}(G)$ the group of automorphisms of G . They have the same form as the endomorphisms, but in order to be invertible we have the additional requirement that the automorphisms be invertible modulo every prime in the decomposition of G . More precisely, if $A \in \text{End}(G)$, then $A \in \text{Aut}(G)$ if and only if $A \pmod{p}$ is invertible as an element of $\text{GL}_n(\mathbb{Z}_p)$ for every p with $p \mid |G|$.

The symplectic group $\text{Sp}(G)$ of G is described by $2|G| \times 2|G|$ matrices in block form. If $S \in \text{Sp}(G)$, we can write

$$S = \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \quad A, B, C, D \in \text{End}(G)$$

such that $AB^* = BA^*$, $CD^* = DC^*$, and $AD^* - BC^* = I$. With the notation introduced thus far, the Stone-von Neumann theorem implies that given a symplectic $S \in \text{Sp}(G)$, there exists a metaplectic unitary U_S on $\mathbb{C}^{|G|}$ such that

$$U_S \pi(\lambda) U_S^{-1} = \psi(\lambda) \pi(S\lambda), \quad \lambda \in G^2, \quad (3.48)$$

where $\psi : G^2 \rightarrow \mathbb{T}$ is a second-degree character on G^2 . Our goal is to find U_S and ψ .

We pause to make a remark on the use of the second-degree character in 3.48. There are a variety of conventions in the literature regarding the use of second-degree characters, for

example, Equations 3.33 and 3.48. This usually corresponds to the phase factor of w and/or the use of the polarized versus non-polarized Heisenberg group. We have chosen to follow the convention of [25] in this section, which is why we use the non-involutive operators π that require a second-degree character.

Let $A \in \text{Aut}(G)$ and $C \in \text{End}(G)$ have $C = C^*$. We will decompose any metaplectic unitary into a product of matrices from three classes: the Fourier transform F , the dilation L_A , and the multiplication operator R_C . Define, for $v \in L^2(G)$,

- $Fv(k) = \frac{1}{\sqrt{|G|}} \sum_{m \in G} \overline{\langle k, m \rangle} v(m), \quad k \in G$
- $L_A v(k) = v(A^{-1}k), \quad k \in G$
- $R_C v(k) = \psi_C(k)v(k), \quad k \in G$

where we define the function $\psi_C : G \rightarrow \mathbb{C}$ by

$$\psi_C(k) = \exp(\pi i k^\top (I + N^{-1})C(I + N)k), \quad k \in G.$$

We note that ψ_C is a second-degree character for C .

Each of the above are metaplectic in their own right; we now give their corresponding symplectics and characters. Indeed, letting A and C be as above, and $l, m \in G$, we have

- $F\pi(l, m)F^{-1} = \langle m, l \rangle \pi(m, -l)$
- $L_A \pi(l, m)L_A^{-1} = \pi(Al, (A^*)^{-1}m)$
- $R_C \pi(l, m)R_C^{-1} = \overline{\psi_C(l)} \pi(l, Cl + m),$

so that $\psi^F(l, m) = \langle m, l \rangle$, $\psi_A^L(l, m) = 1$, and $\psi_C^R(l, m) = \overline{\psi_C(l)}$. We note that the range of any character is contained in the set of $|G|$ th roots of unity. One can deduce the corresponding symplectics from these equations:

$$S^F = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}, \quad S_A^L = \begin{pmatrix} A & 0 \\ 0 & (A^*)^{-1} \end{pmatrix}, \quad S_C^R = \begin{pmatrix} I & 0 \\ C & I \end{pmatrix} \quad (3.49)$$

One other important metaplectic is the inverse Fourier transform F^{-1} ; it is not difficult to see that

$$\psi^{F^{-1}}(\lambda) = \overline{\psi^F(\lambda)} \quad \text{and} \quad S^{F^{-1}} = (S^F)^{-1}.$$

Now, the process reduces to decomposing the symplectic into products of these three, then composing the corresponding component metaplectics together.

We will not attempt to reproduce and prove the entire decomposition in detail; we refer readers interested in learning more directly to [25]. Instead, we present our implementation of the algorithm in pseudocode. We mention the following composition law:

Lemma 3.3.3 (Metaplectic Composition Law). *If U_1 and U_2 satisfy 3.48 for S_1, ψ_1 and S_2, ψ_2 , respectively, then $U_1 U_2$ satisfies 3.48 for $S_1 S_2$ and $\psi(\lambda) = \psi_1(S_2 \lambda) \psi_2(\lambda)$.*

Proof. We see that

$$U_1 U_2 \pi(\lambda) U_2^{-1} U_1^{-1} = \psi_2(\lambda) U_1 \pi(S_2 \lambda) U_1^{-1} = \psi_1(S_2 \lambda) \psi_2(\lambda) \pi(S_1 S_2 \lambda) = \psi(\lambda) \pi(S_1 S_2 \lambda). \quad (3.50)$$

□

We now present the algorithm for computing a metaplectic and character corresponding to a given symplectic matrix. This algorithm has been implemented in `qWeyl`. Before we do so, we note that for any p dividing $|G|$, the matrix $A \bmod p$ created by reducing the components of A modulo p will necessarily be upper block triangular by the divisibility condition on representatives $[A]$ of A :

$$A \bmod p = \begin{pmatrix} A_1 & & & * \\ & A_2 & & \\ & & \ddots & \\ 0 & & & A_u \end{pmatrix} \quad (3.51)$$

We will use this notation in the algorithm. Recall also that we denote symplectic matrices in block form as

$$S = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \mathrm{Sp}(G),$$

for $A, B, C, D \in \mathrm{End}(G)$.

Algorithm 1 Given $G = \mathbb{Z}_{n_1} \times \cdots \times \mathbb{Z}_{n_d}$ with $n_1 | \dots | n_d$ and $S \in \mathrm{Sp}(G)$. We want to compute U_S and ψ that satisfy 3.48

Let **primes** be the set of primes that divide $|G|$, and ν be the product of these primes.

for p in **primes** **do**

 Compute $A \bmod p$

for each block A_j **do**

 Row reduce A_j , and denote by σ_j the set of indices of pivot columns

 Construct Θ_j , diagonal matrix with 0's in the columns indexed by σ_j , and 1's otherwise

end for

 Construct $\Theta^{(p)} = \mathrm{diag}(\theta_1, \dots, \theta_u)$

end for

Construct $\Theta = \sum_{p \text{ in primes}} \frac{\nu}{p} \Theta^{(p)}$

Set $A_0 = A + B\Theta$, $C_0 = C + D\Theta$

return Metaplectic $U_S = R_{C_0 A_0^{-1}} L_{A_0} F^{-1} R_{-A_0^{-1} B} F R_{-\theta}$ and character ψ determined using the composition law 3.3.3

One can show [25, Lemma 5] that the matrix A_0 in Algorithm 1 is necessarily invertible. We write the character explicitly as follows:

$$\begin{aligned} \psi(\lambda) &= \psi_{-\Theta}^R(\lambda) \psi^F(S_{-\Theta}^R \lambda) \psi_{-A_0^{-1} B}^R(S^F S_{-\Theta}^R \lambda) \psi^{F^{-1}}(S_{-A_0^{-1} B}^R S^F S_{-\Theta}^R \lambda) \\ &\quad \psi_{A_0}^L(S^{F^{-1}} S_{-A_0^{-1} B}^R S^F S_{-\Theta}^R \lambda) \psi_{C_0 A_0^{-1}}^R(S_{A_0}^L S^{F^{-1}} S_{-A_0^{-1} B}^R S^F S_{-\Theta}^R \lambda) \end{aligned} \quad (3.52)$$

So we now have a method to compute the metaplectic operators necessary for our finite dimensional Gaussian channels. We note that, in fact, this process easily generalizes to the case where $S \in \text{Sp}(\mathbb{R}^d)$ using analogous operators and the exact same algorithm. We refer the reader to [25, Section 4] for more details.

We end this section with an example of the Mathematica implementation. Initialize a two-particle system with dimensions 3 and 12:

```
In[10]:= qInit[2,{3,12}]
```

```
Out[10]= {q1,q2}
```

Here is a symplectic matrix. Let us verify that it is in fact symplectic:

```
In[11]:= S = {{0,2,1,2},{4,5,8,10},{0,1,1,0},{4,2,0,1}};
S // MatrixForm
symplecticQ[S]
```

```
Out[11]= 
$$\begin{pmatrix} 0 & 2 & 1 & 2 \\ 4 & 5 & 8 & 10 \\ 0 & 1 & 1 & 0 \\ 4 & 2 & 0 & 1 \end{pmatrix}$$

```

```
Out[12]= True
```

The functions `metaplecticU` and `phiMetaplectic` give us the metaplectic unitary and its character, respectively:

```
In[13]:= metaU = metaplecticU[S];
phi[z_] := phiMetaplectic[S,z]
```

We generate a random coordinate in phase space and verify equation 3.48:

```
In[14]:= metaU ** piSparse[lam = randomLambda[]] ** hc[metaU]
== phi[lam] piSparse[S.lam]
```

```
Out[14]= True
```

Up to this point we have been considering the Weyl representation on the *polarized* Heisenberg group $H^{pol}(G)$. It is possible to do the same for the $H(G)$. Let us define the Weyl operator

$$w(l, m)v(k) = \langle m, k \rangle_w v(k - l), \quad (3.53)$$

where the inner product is now

$$\langle m, k \rangle_w = \exp(-\pi i m N^{-1} k). \quad (3.54)$$

These operators are involutive.

The benefit of using these operators is that characters are no longer necessary in the metaplectic law 3.48. Indeed, turning back to Mathematica, where this definition of Weyl operators is implemented as `WSparse`:

```
In[15]:= metaU ** WSparse[lam = randomLambda[]] ** hc[metaU]
        == WSparse[S.lam]
```

```
Out[15]= True
```

3.4 Continuous Gaussian States

Like in the case of distributions, Gaussian states are typically defined in terms of their *characteristic function*.

Recall the *Weyl operators* on $L^2(\mathbb{R}^n)$,

$$w(z) = \exp\left(\frac{i}{2}uv\right)\hat{Z}_v\hat{X}_u, \quad z = (u, v) \in \mathbb{R}^n \times \mathbb{R}^n. \quad (3.55)$$

Given a state S on $L^2(\mathbb{R}^n)$, the characteristic function φ_S of S is given by

$$\varphi_S(z) = \text{Tr}(Sw(z)). \quad (3.56)$$

Definition 3.4.1 (Continuous Gaussian State). A state S on $L^2(\mathbb{R}^n)$ is called *Gaussian* if its characteristic function is of the form

$$\phi_S(z) = \exp\left(im(z) - \frac{1}{2}\alpha(z, z)\right), \quad (3.57)$$

where $m(z) = m^\top z$ is a linear form (corresponding to the *mean*) and $\alpha(z, z) = z^\top \alpha z$ is a quadratic form (corresponding to the *variance*).

Comparing these to the characteristic function of a Gaussian distribution 2.9 one can see the motivation to define the set of Gaussian states to be those states that have characteristic functions of this form. As mentioned in the introduction, the brute force generalization of Definition 3.4.1 to finite abelian groups is too restrictive. In the following section, we outline two equivalent characterizations that serve as motivations for our new notions discussed in Chapter 4.

3.4.1 Coherent States

We present the Heisenberg uncertainty relation and the states that saturate it, known as the minimal uncertainty states. We will also see the number states, an orthonormal system given by the familiar Hermite-Gaussian functions.

Given a state vector $\psi \in \mathcal{S}(\mathbb{R}) \subseteq L^2(\mathbb{R})$, we denote by

$$\begin{aligned} x &= \langle \psi | \hat{q} | \psi \rangle, & y &= \langle \psi | \hat{p} | \psi \rangle, \\ D_\psi(\hat{q}) &= \|(\hat{q} - x)\psi\|^2, & D_\psi(\hat{p}) &= \|(\hat{p} - y)\psi\|^2 \end{aligned}$$

the mean values and variances of observables \hat{q}, \hat{p} on the pure state $|\psi\rangle\langle\psi|$. Note that above and anywhere else it may appear, the difference $\hat{q} - x$ between an operator and a number

should be understood as $\hat{q} - xI$, where I is the identity on the space that \hat{q} is defined on. Fixing any real number ω , we see that

$$\begin{aligned} 0 &\leq \|[\omega(\hat{q} - x) + i(\hat{p} - y)]\psi\|^2 \\ &= \omega^2 D_\psi(\hat{q}) + i\omega \langle \psi | [\hat{q}, \hat{p}] | \psi \rangle + D_\psi(\hat{p}) \\ &= \omega^2 D_\psi(\hat{q}) - \omega\hbar + D_\psi(\hat{p}). \end{aligned}$$

This is a quadratic equation in the variable ω , which has discriminant $\hbar^2 - 4D_\psi(\hat{q})D_\psi(\hat{p})$. Since ω is real, we have

$$D_\psi(\hat{q})D_\psi(\hat{p}) \geq \frac{\hbar^2}{4}. \quad (3.58)$$

This inequality is saturated if and only if there exists ω such that

$$[\omega(\hat{q} - x) + i(\hat{p} - y)]\psi = 0.$$

This is in fact a differential equation:

$$[\omega(\xi - x) + (\hbar \frac{d}{d\xi} - iy)]\psi(\xi) = 0,$$

which has the normalized solution

$$\psi(\xi) = \sqrt[4]{\frac{\omega}{\pi\hbar}} \exp\left(\frac{iy}{\hbar}(\xi - \frac{x}{2}) - \frac{\omega(\xi - x)^2}{2\hbar}\right), \quad (3.59)$$

up to a constant factor with modulus one.

Inequality 3.58 is the *Heisenberg Uncertainty Relation*, and 3.59 are the *minimal uncertainty* state vectors. Some of these are plotted in Figure 3.1. We can recast the equation using

$$a = \frac{1}{\sqrt{2\hbar\omega}}(\omega\hat{q} + i\hat{p}), \quad \zeta = \frac{1}{\sqrt{2\hbar\omega}}(\omega x + iy) \quad (3.60)$$

to get

$$a|\zeta\rangle = \zeta|\zeta\rangle, \quad \zeta \in \mathbb{C}.$$

Here $|\zeta\rangle$ denotes the minimal uncertainty state vector with parameters x, y, ω . In particular,

$$a|0\rangle = 0,$$

where $|0\rangle$ is the vector parametrized by

$$\psi(\xi) = \sqrt[4]{\frac{\omega}{\pi\hbar}} \exp\left(-\frac{\omega\xi^2}{2\hbar}\right).$$

The operator a has adjoint $a^* = \frac{1}{\sqrt{2\hbar\omega}}(\omega\hat{q} - i\hat{p})$. Computing the commutator, we see

$$[a, a^*] = \frac{1}{2\hbar\omega}((\omega\hat{q} + i\hat{p})(\omega\hat{q} - i\hat{p}) - (\omega\hat{q} - i\hat{p})(\omega\hat{q} + i\hat{p})) = \frac{2i\omega}{2\hbar\omega}(\hat{p}\hat{q} - \hat{q}\hat{p}) = \frac{i}{\hbar}[\hat{p}, \hat{q}] = iI,$$

which is an equivalent reformulation of the commutation relation.

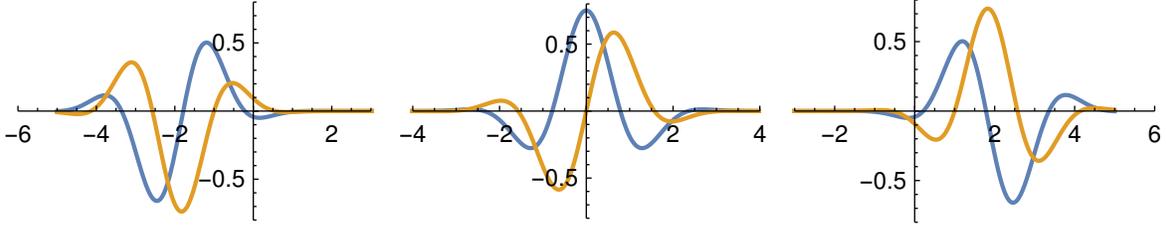


Figure 3.1: Plot of a minimal uncertainty state vector with momentum $y = 2$ and position $x = -2, 0, 2$, respectively. The real parts are blue, while the imaginary parts are yellow.

The essentially self-adjoint *number operator* $\mathcal{N} := a^*a = aa^* - I$ satisfies

$$\mathcal{N}|0\rangle = 0.$$

For any $n \in \mathbb{N}$, the associated *number states*

$$|n\rangle = \frac{(a^*)^n}{\sqrt{n!}}|0\rangle,$$

form an orthonormal system of eigenvectors for the number operator \mathcal{N} , where

$$\mathcal{N}|n\rangle = n|n\rangle, \quad n \in \mathbb{N}. \quad (3.61)$$

Indeed, orthonormality follows from the commutation relation, and one can see the eigenrelation with

$$\mathcal{N}|n\rangle = \frac{1}{\sqrt{n!}}a^*a(a^*)^n|0\rangle = \frac{1}{\sqrt{n!}}a^*(a^*a + I)(a^*)^{n-1}|0\rangle = \frac{1}{\sqrt{n!}}((a^*)^2a(a^*)^{n-1} + (a^*)^n)|0\rangle.$$

That this is a complete system in $L^2(\mathbb{R})$ will be justified presently. Some examples of number states are plotted in Figure 3.2.

One can parametrize the number states using the defining relation

$$\psi_n(\xi) = \frac{1}{\sqrt{2^n n!}} \sqrt{\frac{\omega^3}{\pi \hbar^3}} \exp\left(-\frac{\omega \xi^2}{2\hbar}\right) H_n\left(\sqrt{\frac{\omega}{\hbar}}\xi\right),$$

where H_n is the n th *Hermite polynomial*, given by

$$H_n(\xi) = (-1)^n e^{\xi^2} \frac{d^n}{d\xi^n} e^{-\xi^2}. \quad (3.62)$$

States of the form above are often called the *Hermite-Gaussians*, as they are a Hermite polynomial times a Gaussian factor. The Hermite polynomials form a complete orthogonal system in $L^2(\mathbb{R})$, which implies that the number states form an orthonormal basis in $L^2(\mathbb{R})$. This gives us the spectral decomposition

$$\mathcal{N} = \sum_{n=0}^{\infty} n|n\rangle\langle n|. \quad (3.63)$$

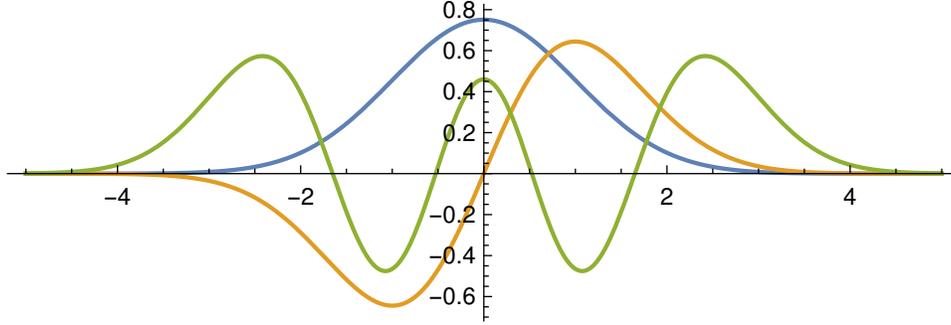


Figure 3.2: Plot of the number states: $n = 0$ is given in blue, $n = 1$ in yellow, and $n = 4$ in green.

One easily verifies the relations

$$a|n\rangle = \sqrt{n}|n-1\rangle, \quad a^*|n\rangle = \sqrt{n+1}|n+1\rangle.$$

We connect the number operator to the position and momentum operators via

$$\hbar\omega(\mathcal{N} + \frac{I}{2}) = \frac{1}{2}(\omega^2\hat{q}^2 + \hat{p}^2) =: \hbar\hat{H}. \quad (3.64)$$

We can now comment on the physical interpretation of the operators we have introduced. \hat{H} is the quantum analogue of the harmonic oscillator 3.1. Much like the classical harmonic oscillator, the quantum oscillator is an oft-used model for more complicated quantum dynamical systems. Since \mathcal{N} and \hat{H} commute, the Hermite-Gaussians form an orthonormal basis of eigenvectors for \hat{H} .

For any minimal uncertainty state vector corresponding to $\zeta \in \mathbb{C}$, we call $|\zeta\rangle\langle\zeta|$ a *coherent state*. We can calculate the inner product of two minimal uncertainty vectors to get

$$\langle\zeta_1|\zeta_2\rangle = \exp\left(-\frac{1}{2}(|\zeta_1|^2 + |\zeta_2|^2 - 2\bar{\zeta}_1\zeta_2)\right).$$

Indeed, letting $|\zeta_1\rangle = \psi_1$, $|\zeta_2\rangle = \psi_2$, and using another well-known formula for a Gaussian integral A.1, we see

$$\begin{aligned} \langle\zeta_1|\zeta_2\rangle &= \int_{-\infty}^{\infty} \overline{\psi_1(\xi)}\psi_2(\xi)d\xi \\ &= \sqrt{\frac{\omega}{\pi\hbar}} \int_{-\infty}^{\infty} \exp\left(-\left(\frac{\omega}{\hbar}\right)\xi^2 + \frac{1}{\hbar}(\omega x_1 + \omega x_2 - iy_1 + iy_2)\xi - \frac{1}{2\hbar}(\omega x_1^2 + \omega x_2^2 - iy_1x_1 + ix_2y_2)\right) \\ &= \sqrt{\frac{\omega}{\pi\hbar}} \sqrt{\frac{\pi\hbar}{\omega}} \exp\left(\frac{\left(\frac{\omega x_1 + \omega x_2 - iy_1 + iy_2}{\hbar}\right)^2}{4\left(\frac{\omega}{\hbar}\right)} - \frac{1}{2\hbar}(\omega x_1^2 + \omega x_2^2 - iy_1x_1 + ix_2y_2)\right) \\ &= \exp\left(-\frac{1}{4\hbar\omega}(\omega^2x_1^2 + y_1^2 + \omega^2x_2^2 + y_2^2 - 2(\omega^2x_1y_2 + y_1y_2 + i(\omega y_2x_1 - \omega y_1x_2)))\right) \\ &= \exp\left(-\frac{1}{2}(|\zeta_1|^2 + |\zeta_2|^2 - 2\bar{\zeta}_1\zeta_2)\right) \end{aligned}$$

We can use this relation to show the following special case:

$$\langle n|\xi\rangle = \frac{1}{\sqrt{n!}}\langle 0|a^n\xi\rangle = \frac{\xi^n}{\sqrt{n!}}\langle 0|\xi\rangle = \frac{\xi^n}{\sqrt{n!}}\exp\left(-\frac{|\xi|^2}{2}\right). \quad (3.65)$$

This will be useful shortly.

We can parametrize the minimal uncertainty states by their expected values for position x and momentum y , so that $|x, y\rangle := |\zeta\rangle$. Recalling the Weyl operators 3.21, a straightforward calculation shows that these operators shift the expected values of the minimal uncertainty states in the following way:

$$\begin{aligned} |x, y\rangle &= w(y/\hbar, -x/\hbar)|0, 0\rangle \\ w(y_2/\hbar, -x_2/\hbar)|x_1, y_1\rangle &= \exp\left(\frac{i}{2\hbar}[(x_1, y_1), (x_2, y_2)]\right)|x_1 + x_2, y_1 + y_2\rangle. \end{aligned}$$

Some call the Weyl operators *displacement operators* for this reason.

The set of minimal uncertainty state vectors $\{|\zeta\rangle \mid \zeta \in \mathbb{C}\}$ is *overcomplete* in the following sense:

$$\frac{1}{\pi} \int_{\mathbb{C}} |\zeta\rangle\langle\zeta| d^2\zeta = I,$$

where $d^2\zeta = \frac{1}{2\hbar} dx dy$. Indeed, note that the operator on the left-hand side is invariant under conjugation by $w(u, v)$ for all $u, v \in \mathbb{R}$. By irreducibility of the Weyl representation, it must lie in $\mathbb{C}I$. One can fix the factor of $1/\pi$ by taking the expected value with any state, say the vacuum state $|0\rangle$, and using 3.65.

Now we have a way of assigning, for every $\zeta \in \mathbb{C}$, a pure state $|\zeta\rangle\langle\zeta|$. Given a probability density $p(\zeta)$ on \mathbb{C} , we can define

$$S_p = \int_{\mathbb{C}} |\zeta\rangle\langle\zeta| p(\zeta) d^2\zeta,$$

which is a density operator. A state of this form is called *classical*, because it is a way of encoding a classical signal (that is, a probability distribution) as a quantum state. Choosing $p = \gamma$ to be a Gaussian density with zero mean and variance N , we obtain the density operator

$$S_0 = \frac{1}{\pi N} \int_{\mathbb{C}} |\zeta\rangle\langle\zeta| \exp\left(-\frac{|\zeta|^2}{N}\right) d^2\zeta. \quad (3.66)$$

We call this state *centered* because the distribution it is based on is centered. One can also consider the *shifted* state, that is, with variance N and mean μ ,

$$S_\mu = \frac{1}{\pi N} \int_{\mathbb{C}} |\zeta\rangle\langle\zeta| \exp\left(-\frac{|\zeta - \mu|^2}{N}\right) d^2\zeta. \quad (3.67)$$

Here μ is a complex number of the form $\mu = \frac{1}{\sqrt{2\hbar\omega}}(\omega m_q + im_p)$. Computing the characteristic function of this state, we get

$$\phi_{S_\mu}(z) = \exp\left(i(m_q x + m_p y) - \frac{\alpha}{2}(x^2 + y^2)\right), \quad (3.68)$$

so that S_μ is a Gaussian state. A characteristic function of this form is plotted in Figure 3.3.

One should note that the state S_μ is pure if and only if the variance N is zero; in this case, $S_\mu = |\mu\rangle\langle\mu|$.

We compute the matrix elements of the centred state S_0 using 3.65 and A.2:

$$\langle n|S_0|m\rangle = \frac{1}{\pi N} \int \frac{\zeta^n}{\sqrt{n!}} \frac{\bar{\zeta}^m}{\sqrt{m!}} \exp\left(-\frac{(N+1)|\zeta|^2}{N}\right) d^2\zeta = \delta_{nm} \frac{1}{N+1} \left(\frac{N}{N+1}\right)^n. \quad (3.69)$$

From this it follows that S_0 has the spectral decomposition

$$S_0 = \frac{1}{N+1} \sum_{n=1}^{\infty} \left(\frac{N}{N+1}\right)^n |n\rangle\langle n|. \quad (3.70)$$

Using the spectral decomposition of the number operator 3.63 and the definition of the Hamiltonian of the quantum harmonic oscillator 3.64, we can write the centered state S_0 as

$$S_{\hat{H},\theta} = \frac{\exp(-\theta\hat{H})}{\text{Tr}(\exp(-\theta\hat{H}))}. \quad (3.71)$$

This is a *Gibbs equilibrium state*, also known as a *thermal state*. Here the parameter θ gives the inverse temperature $\frac{\theta}{\hbar} = \frac{1}{\hbar\omega} \ln \frac{N+1}{N}$. We have therefore written the Gaussian state S_0 as a Gibbs state of a Hamiltonian which is quadratic in the observables \hat{q}, \hat{p} . It turns out that such Gibbs states exhaust all continuous Gaussian states, as we will see in the next section.

3.4.2 Gibbs Formalism

It turns out that the Gibbs state 3.71 is a powerful general framework for thinking about Gaussianity. If we let the Hamiltonian \hat{H} be a more general degree-2 polynomial in the canonical observables \hat{q}, \hat{p} , we can construct an entire class of states that behave well with respect to the Weyl operators and metaplectic unitaries that we have already discussed. We follow the process laid out in [30].

Recall the general form of the Gibbs state:

$$\frac{\exp(-\beta\hat{H})}{\text{Tr} \exp(-\beta\hat{H})}, \quad (3.72)$$

where \hat{H} is the *Hamiltonian* of the quantum harmonic oscillator, and β is a constant that is inversely related to *temperature*. As we let $\beta \rightarrow \infty$, the system cools and we obtain the vacuum state, $|0\rangle\langle 0|$.

We look at this formalism in a more general way; we fix a “second-order” Hamiltonian, that is, one that can be expressed as a polynomial of order at most 2 in the canonical variables \hat{q}, \hat{p} . It suffices to consider such “quadratic” systems in many cases, as often the higher order terms are negligible (or assumed to be). Notice that the Hamiltonian of the quantum harmonic oscillator 3.64 is quadratic. We consider a system with n degrees of freedom; this could model either a single particle moving in n spatial dimensions, or n particles moving on a line.

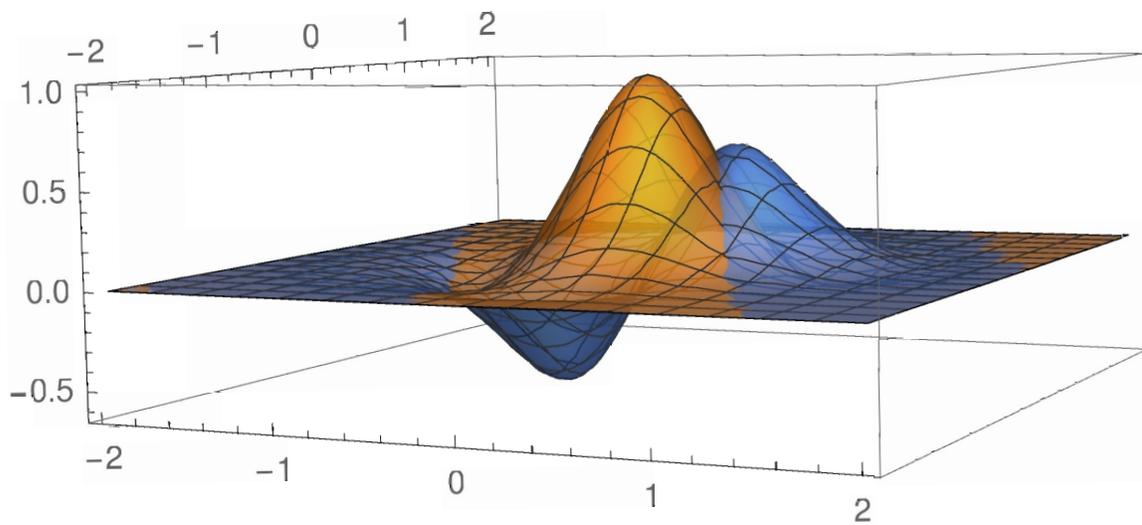


Figure 3.3: Plot of the single mode Gaussian characteristic function ϕ_{S_μ} with $m_p = -2$, $m_q = 1$, and $\alpha = 2$. Orange is the real part, while blue is imaginary. Compare this to the characteristic function for a Gaussian distribution given in Figure 2.2.

Each degree of freedom has a pair of canonical variables, \hat{q}_j, \hat{p}_j for $j = 1, \dots, n$. We recall the Heisenberg Canonical Commutation Relations (CCR),

$$[\hat{q}_i, \hat{p}_j] = \delta_{ij}, \quad [\hat{q}_i, \hat{q}_j] = 0 \quad [\hat{p}_i, \hat{p}_j] = 0 \quad (3.73)$$

The Hilbert space of this system is $L^2(\mathbb{R}^n)$. We denote, in full generality, a second-order Hamiltonian by

$$\hat{H} = \frac{1}{2} \hat{r}^\top H \hat{r} + \hat{r}^\top r. \quad (3.74)$$

Here \hat{r} is a column vector of the canonical operators, $\hat{r}^\top = (\hat{q}_1, \hat{p}_1, \dots, \hat{q}_n, \hat{p}_n)$, $r \in \mathbb{R}^{2n}$, and $H \in M_{2n}(\mathbb{R})$ is a symmetric matrix, called the *Hamiltonian matrix*. Note that we have ordered the canonical observables in a different way than in the previous section. We may take H symmetric because any antisymmetric component will become a multiple of the identity, by the CCR. This corresponds to adding a constant to the Hamiltonian, and can be ignored. The factor of $\frac{1}{2}$ is for convenience and this will become clear shortly. We note that we can view the CCR in terms of \hat{r} :

$$[\hat{r}, \hat{r}^\top] = i\Omega, \quad (3.75)$$

where the left-hand side is the *commutator matrix*, given by $M_{ij} = [\hat{r}_i, \hat{r}_j]$. Here Ω is again the matrix of the symplectic form, now given by $\Omega = \bigoplus_{j=1}^n \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

The benefit to using symmetric matrices H is that they are necessarily diagonalizable; this will be a powerful tool for “decoupling” composite systems and writing a given Hamiltonian in an easy to understand diagonal form. In addition, we will ask that the Hamiltonian matrix be positive definite. Such matrices give rise to Hamiltonians which are bounded below, and this will ensure that the system is “stable”, in the sense that the energy is bounded below.

It follows from [30] that any continuous Gaussian state S may be written as

$$S = \frac{\exp(-\beta \hat{H})}{\text{Tr} \exp(-\beta \hat{H})}, \quad (3.76)$$

with positive β and \hat{H} defined by r and H as in equation 3.74. The limiting case

$$S = \lim_{\beta \rightarrow \infty} \frac{\exp(-\beta \hat{H})}{\text{Tr} \exp(-\beta \hat{H})}. \quad (3.77)$$

recovers all coherent states.

We now outline Serafini’s argument which shows that an arbitrary Gaussian state as in 3.76 can be written in a canonical diagonal form. First, we note that since H is positive definite, it is invertible, so that setting $\bar{r} = -H^{-1}r$ (and ignoring the constant this introduces), we get

$$\hat{H}' = \frac{1}{2} (\hat{r} - \bar{r})^\top H (\hat{r} - \bar{r}), \quad (3.78)$$

which is equivalent to the general form for a second-order Hamiltonian given in equation 3.74. We can view the action of \bar{r} as a *displacement*; one recalls that this is exactly the action of the *Weyl operators*, Equation 3.55.

Proposition 3.4.2 (Action of Weyl operators). *We have*

$$w(\Omega^{-1}\bar{r})^{-1}\hat{r}w(\Omega^{-1}\bar{r}) = \hat{r} - \bar{r}. \quad (3.79)$$

One should understand the above formula as one Weyl operator acting on each operator in the vector \hat{r} . The proof follows from the CCR, and can be found in [30, Page 40]. This result further reinforces the justification for calling the Weyl operators “displacement operators.”

Combining equation 3.79 with 3.78, we get

$$\hat{H}' = \frac{1}{2}(\hat{r} - \bar{r})^\top H(\hat{r} - \bar{r}) = \frac{1}{2}w(\Omega^{-1}\bar{r})^{-1}\hat{r}^\top H\hat{r}w(\Omega^{-1}\bar{r}). \quad (3.80)$$

Note that the matrix H commutes with the Weyl operators as the former simply represents a set of coefficients and is not a true operator. Thus, this equation tells us that, up to the action of a Weyl operator, one may set the vector \bar{r} in the quadratic Hamiltonian \hat{H}' to zero without loss of generality.

We have removed the linear part of the general quadratic Hamiltonian by the previous discussion. We now aim to show that the quadratic part may be diagonalized via metaplectic unitaries, and to do so we consider dynamics in the Heisenberg picture 2.4 where states stay the same and the observables evolve with time. Recall that the vector of canonical observables \hat{r} evolves under the centered Hamiltonian $\hat{H} = \frac{1}{2}\hat{r}^\top H\hat{r}$, by definition, according to

$$\frac{d}{dt}\hat{r} = \frac{i}{2}[\hat{H}, \hat{r}].$$

We compute, first componentwise:

$$\begin{aligned} \frac{d}{dt}\hat{r}_j &= \frac{i}{2}[\hat{H}, \hat{r}_j] = \frac{i}{2}\sum_{kl}[\hat{r}_k H_{kl}\hat{r}_l, \hat{r}_j] \\ &= \frac{i}{2}\sum_{kl}H_{kl}(\hat{r}_k[\hat{r}_l, \hat{r}_j] + [\hat{r}_k, \hat{r}_j]\hat{r}_l) = \sum_{kl}\Omega_{jk}H_{kl}\hat{r}_l. \end{aligned}$$

One can see the last equality by considering that r_j is either q_k or p_k for some k , and noticing that its commutator with \hat{r}_k can only be $\pm i$ when $k \pm 1 = j$, depending on the parity of j . We can recast these equations in vector form,

$$\frac{d}{dt}\hat{r} = \Omega H\hat{r},$$

a differential equation that is easily solved: $\hat{r}(t) = \exp(\Omega Ht)\hat{r}(0)$. Thus, by 2.4, we may write

$$\exp\left(\frac{it}{2}\hat{r}^\top H\hat{r}\right)\hat{r}\exp\left(-\frac{it}{2}\hat{r}^\top H\hat{r}\right) = \exp(\Omega Ht)\hat{r}, \quad (3.81)$$

and since unitary conjugation preserves the CCR, it follows that $\exp(\Omega Ht) \in \text{Sp}(\mathbb{R}^n)$. By what we have seen in Section 3.3, we expect $\exp\left(\frac{it}{2}\hat{r}^\top H\hat{r}\right)$ to be *metaplectic* (compare with Equation 3.46). We denote

$$\hat{U}_H = \exp\left(\frac{i}{2}\hat{r}^\top H\hat{r}\right). \quad (3.82)$$

The *normal mode decomposition* is but one of many results related to reducing an arbitrary matrix to a canonical form, often diagonal. This particular version will prove very useful to us, as the transforming matrix is symplectic, and the elements on the diagonal are pairs of positive numbers. We quote it from [30, page 42]:

Theorem 3.4.3 (Normal Mode Decomposition). *Given a $2n \times 2n$ positive definite real matrix H , there exists a symplectic transformation $S \in Sp(\mathbb{R}^{2n})$ such that*

$$SHS^\top = D, \text{ with } D = \text{diag}(d_1, d_1, \dots, d_n, d_n), \quad (3.83)$$

where d_j is a positive real number, for all j .

The proof is based mostly on linear algebra and is not instructive to our narrative, so we refer the reader to the source given above. But how does one obtain the matrix S and the d_j 's (henceforth known as *symplectic eigenvalues*)? These come from the matrix $i\Omega H$. To see this, first note that if $\Omega_1 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ is the 2×2 symplectic form, then the matrix

$$\bar{v} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix} \quad (3.84)$$

diagonalizes Ω_1 : that is, $\bar{v}\Omega_1\bar{v}^* = -i\sigma_z$, where $\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ is the Pauli z matrix. Thus,

$\Omega = \bigoplus^n \Omega_1$ is diagonalized by $\bar{V} = \bigoplus^n \bar{v}$.

Now, rearrange 3.83 to get $H = S^{-1}D(S^\top)^{-1}$, and then

$$i\Omega H = i\Omega S^{-1}D(S^\top)^{-1} = iS^\top \Omega D(S^\top)^{-1} = S^\top \bar{V}^* \left(\bigoplus_{j=1}^n d_j \sigma_z \right) \bar{V} (S^\top)^{-1}.$$

Then, indeed, $i\Omega H$ is diagonalizable by $L := \bar{V}(S^\top)^{-1}$ with symplectic eigenvalues d_1, \dots, d_n . From this, we can conclude that the symplectic matrix that diagonalizes H into normal modes is $S = \bar{V}^* L^{*-1}$.

Every pair of observables that satisfy the CCR are equally valid choices for the canonical observables. Since a symplectic matrix preserves the symplectic form and thus the CCR, the effect of S is to define a new pair of canonical observables that are linearly related to the original ones. We have shown that in these variables, the Hamiltonian is diagonal, and we will presently show that under such a Hamiltonian, the system behaves like a set of n harmonic oscillators. Thinking geometrically, this means that, if a system evolves under a second-order Hamiltonian, then there is some linear transformation we can perform such that the motion is simple harmonic oscillation in n degrees of freedom. We will also show that in this case the n symplectic eigenvalues correspond to the frequencies of oscillation in each dimension.

Now we show that second-order Hamiltonian with zero displacement $\hat{H}' = \hat{r}^\top H \hat{r}$ is in fact a sum of harmonic oscillators. Indeed, by the normal mode decomposition, there is a symplectic matrix S^H (which is the inverse transpose of the one obtained through the normal mode decomposition) such that

$$H = S^{H^\top} \left(\bigoplus_{j=1}^n \omega_j I_2 \right) S^H, \quad (3.85)$$

where I_2 is the 2×2 identity matrix. Thus, we can write

$$\hat{H}' = \hat{r}^\top S^{H\top} \left(\bigoplus_{j=1}^n \omega_j I_2 \right) S^H \hat{r}. \quad (3.86)$$

But there exists a metaplectic operator \hat{U} such that $S^H \hat{r} = \hat{U} \hat{r} \hat{U}^*$, so that

$$\begin{aligned} \hat{r}^\top H \hat{r} &= \hat{U} \hat{r}^\top \hat{U}^* \left(\bigoplus_j \omega_j I_2 \right) \hat{U} \hat{r} \hat{U}^* \\ &= \hat{U} \hat{r}^\top \left(\bigoplus_j \omega_j I_2 \right) \hat{r} \hat{U}^* \\ &= \hat{U} \sum_{j=1}^n \omega_j (\hat{q}_j^2 + \hat{p}_j^2) \hat{U}^*. \end{aligned}$$

One clearly recognizes the Hamiltonian of the harmonic oscillator in the summation. Thus, every second-order Hamiltonian is unitarily equivalent to the Hamiltonian of a set of free, non-interacting harmonic oscillators.

Now that we have decoupled a general second-order Hamiltonian into n harmonic oscillators, we need only figure out what the Gibbs state 3.72 looks like for such a simple Hamiltonian. Denote, for a frequency ω_j ,

$$\hat{H}_{\omega_j} = \frac{\omega_j}{2} (\hat{q}_j^2 + \hat{p}_j^2).$$

However, recall that we have shown in equation 3.64 that this can be equivalently expressed in terms of the annihilation and creation operators \hat{a}_j, \hat{a}_j^* as

$$\hat{H}_{\omega_j} = \omega_j \left(\hat{a}_j^* \hat{a}_j + \frac{I}{2} \right).$$

We have already discussed that this implies that the eigenstates of \hat{H}_{ω_j} are exactly the number states $\{|n\rangle_j \mid n \in \mathbb{N}\}$, and the state $|m\rangle_j$ has eigenvalue $\omega_j(m + \frac{1}{2})$. We use the subscript j as a reminder that we have several subsystems. Then the exponential of \hat{H}_{ω_j} can be written

$$\exp(-\beta \hat{H}_{\omega_j}) = \sum_{m=0}^{\infty} \exp(-\beta \omega_j (m + 1/2)) |m\rangle \langle m|_j. \quad (3.87)$$

The normalization factor can be calculated as

$$\begin{aligned} \frac{1}{\text{Tr} \exp(-\beta \hat{H}_{\omega_j})} &= \frac{1}{\exp(-\beta \omega_j \frac{1}{2}) \sum_{m=0}^{\infty} \exp(-\beta \omega_j m)} \\ &= \frac{1 - \exp(-\frac{\beta \omega_j}{2})}{\exp(-\frac{\beta \omega_j}{2})} \\ &= \exp\left(\frac{\beta \omega_j}{2}\right) - \exp\left(-\frac{\beta \omega_j}{2}\right). \end{aligned}$$

We can now combine all of the above results to get the most general reduced Hamiltonian. Given a $2n \times 2n$ positive definite matrix H , a vector of displacements $r \in \mathbb{R}^{2n}$, and a positive parameter β , define $\bar{r} = -H^{-1}r$. We can find a Weyl operator $w(\Omega^{-1}\bar{r})$, a metaplectic operator \hat{U} , and a list of frequencies $\{\omega_j\}_{j=1}^n$ such that

$$\hat{H} = (\hat{r} - \bar{r})^\top H (\hat{r} - \bar{r}) = w(\Omega^{-1}\bar{r})^{-1} \hat{U} \left(\sum_{j=1}^n \hat{H}_{\omega_j} \right) \hat{U}^* w(\Omega^{-1}\bar{r}), \quad (3.88)$$

and its corresponding Gibbs state

$$S = \frac{\exp(-\beta\hat{H})}{\text{Tr} \exp(-\beta\hat{H})} \quad (3.89)$$

$$= w(\Omega^{-1}\bar{r})^{-1} \hat{U} \left(\frac{\bigotimes_{j=1}^n \exp(-\beta\hat{H}_{\omega_j})}{\prod_{j=1}^n \text{Tr} \exp(-\beta\hat{H}_{\omega_j})} \right) \hat{U}^* w(\Omega^{-1}\bar{r}) \quad (3.90)$$

$$= \left(\prod_{j=1}^n (1 - \exp(-\beta\omega_j)) \right) w(\Omega^{-1}\bar{r})^{-1} \hat{U} \left(\bigotimes_{j=1}^n \left(\sum_{m=0}^{\infty} \exp(-\beta\omega_j m) |m\rangle \langle m|_j \right) \right) \hat{U}^* w(\Omega^{-1}\bar{r}). \quad (3.91)$$

It follows that the characteristic function of S is of the form 3.4.1. The pure state limit as $\beta \rightarrow \infty$ can be given simply as

$$S = w(\Omega^{-1}\bar{r})^{-1} \hat{U} \left(\bigotimes_{j=1}^n |0\rangle \langle 0|_j \right) \hat{U}^* w(\Omega^{-1}\bar{r}). \quad (3.92)$$

These expressions give the most general Gaussian states.

We summarize the above process in Algorithm 2:

Algorithm 2 Given $2n \times 2n$ positive definite matrix H , positive β , displacement vector $r \in \mathbb{R}^{2n}$, calculate S , the Gaussian state corresponding to \hat{H}

Compute $\bar{r} = -H^{-1}r$

Compute $w(\Omega^{-1}\bar{r})$

Diagonalize $i\Omega H$ to obtain S^H and symplectic eigenvalues $\{\omega_j\}_{j=1}^n$

Use the continuous version of Algorithm 3.3.3 with input S^H to obtain metaplectic operator \hat{U}

return S using equation 3.89

Thus ends our treatment of the construction of a general Gibbs state. We will see that this process is much easier to generalize to the discrete case than the classical method of Holevo, largely because it focuses less on the canonical operators \hat{q}, \hat{p} which, as we will see, are problematic in finite dimensions.

Now we will show the implementation of the symplectic diagonalization procedure for a positive semidefinite matrix. Define such a matrix with integer coefficients of size 4×4 :

```
In[16]:= n = 2; psd = randomIntegerPSD[2n,3];
         psd // MatrixForm
```

```
Out[16]=

$$\begin{pmatrix} 10 & 5 & 7 & 7 \\ 5 & 6 & 3 & 5 \\ 7 & 3 & 6 & 3 \\ 7 & 5 & 3 & 9 \end{pmatrix}$$

```

The following computes L and the symplectic S using the algorithm above:

```
In[17]:= L = getL[psd];
S = psdSymplecticDiagonalize[psd];
L // MatrixForm
S // MatrixForm
```

```
Out[17]=

$$\begin{pmatrix} 0.12089 - 0.341012 i & 0.223181 - 0.116587 i & 0. - 0.273485 i & 0.273248 - 0.17123 i \\ 0.12089 + 0.341012 i & 0.223181 + 0.116587 i & 0. + 0.273485 i & 0.273248 + 0.17123 i \\ -0.12089 + 0.0832516 i & -0.223181 - 0.601461 i & 0. - 0.334094 i & 0.226752 + 0.576283 i \\ -0.12089 - 0.0832516 i & -0.223181 + 0.601461 i & 0. + 0.334094 i & 0.226752 - 0.576283 i \end{pmatrix}$$

```

```
Out[18]=

$$\begin{pmatrix} -0.315283 & 0.922186 & -0.46305 & 0.739574 \\ -0.603539 & 0.326917 & -0.738933 & 0. \\ -1.47159 & -0.203692 & 1.40999 & 0.817428 \\ 0.546057 & -0.295781 & -0.554794 & 0. \end{pmatrix}$$

```

Note that the symplectic obtained does not necessarily have integer coefficients. The following verifies that the appropriate relations hold:

```
In[19]:= I L.omega[n].psd.Inverse[L] // MatrixForm // Chop
S.omega[n].S^T // MatrixForm // Chop
S.psd.S^T // MatrixForm // Chop
symplecticEigenvalues[psd]
```

```
Out[19]=

$$\begin{pmatrix} 10.3811 & 0 & 0 & 0 \\ 0 & -10.3811 & 0 & 0 \\ 0 & 0 & 0.481643 & 0 \\ 0 & 0 & 0 & -0.481643 \end{pmatrix}$$

```

```
Out[20]=

$$\begin{pmatrix} 0 & 1. & 0 & 0 \\ -1. & 0 & 0 & 0 \\ 0 & 0 & 0 & 1. \\ 0 & 0 & -1. & 0 \end{pmatrix}$$

```

```
Out[21]=

$$\begin{pmatrix} 10.3811 & 0 & 0 & 0 \\ 0 & 10.3811 & 0 & 0 \\ 0 & 0 & 0.481643 & 0 \\ 0 & 0 & 0 & 0.481643 \end{pmatrix}$$

```

```
Out[22]=

$$\left\{ -\sqrt{54 + 7\sqrt{59}}, \sqrt{54 + 7\sqrt{59}}, -\frac{5}{\sqrt{54 + 7\sqrt{59}}}, \frac{5}{\sqrt{54 + 7\sqrt{59}}} \right\}$$

```

3.4.3 The Wigner Function

Given a state vector $\psi \in L^2(\mathbb{R})$, one can compute the probability distribution $\rho(q)$ of the position of the particle with the norm squared, $\rho(q) = |\psi(q)|^2$. Thus, it is relatively easy to picture the location distribution from ψ . One can use the Fourier transform to do the same for momentum: that is, if $\varphi = F\psi$, then $\sigma(p) = |\varphi(p)|^2$ is the distribution of momentum. The Wigner function arises from a desire to visualize *both* the position and momentum distributions at the same time, with one function.

Another desirable property of the Wigner function is its connection to the phase space from classical mechanics. Indeed, these two notions combine to give an *equivalent formulation of quantum mechanics*: that is, one can understand everything in quantum mechanics using Wigner functions on phase space instead of using operators on Hilbert space. We will not cover the entirety of this formalism; we ignore Wigner function dynamics, and only mention a few of the desirable qualities as they pertain to our applications. However, one should note the power of these functions, and the reader is referred to [11] for a readable introduction and the references within for more details.

Suppose we have a pure state ψ . The *Wigner Function* W of ψ is a real-valued function on the phase space V given by

$$W(p, q) = \frac{1}{2\pi\hbar} \int e^{-ipx/\hbar} \psi(q + x/2) \psi^*(q - x/2) dx. \quad (3.93)$$

It is not immediately obvious that W is real valued: if one takes the conjugate and does a change of variables $x \mapsto -x$, one can verify this. We can also now see that W does in fact encode both the position and momentum distributions: indeed, using

$$\int e^{\frac{ipx}{\hbar}} dp = 2\pi\hbar \delta_0(x), \quad (3.94)$$

which is to be understood in a measure theoretical sense, one can compute the *marginal distribution*, that is, the distribution along just one of the two variables, as

$$\int W(q, p) dp = \psi^*(q) \psi(q) \quad \text{and} \quad \int W(q, p) dq = \varphi^*(p) \varphi(p). \quad (3.95)$$

Geometrically, this amounts to viewing the Wigner distribution parallel to the p or q axes, respectively, making it effectively a projection.

One can also use equation 3.94 to show that W is normalized:

$$\begin{aligned} \int \int W(q, p) dp dq &= \int \int \frac{1}{2\pi\hbar} \int e^{-ipx/\hbar} \psi(q + x/2) \psi^*(q - x/2) dx dp dq \\ &= \frac{1}{2\pi\hbar} \int \int \left(\int e^{-ipx/\hbar} dp \right) \psi(q + x/2) \psi^*(q - x/2) dx dq \\ &= \int \int \delta_0(x) \psi(q + x/2) \psi^*(q - x/2) dx dq \\ &= \int \psi(q) \psi^*(q) dq \\ &= 1. \end{aligned}$$

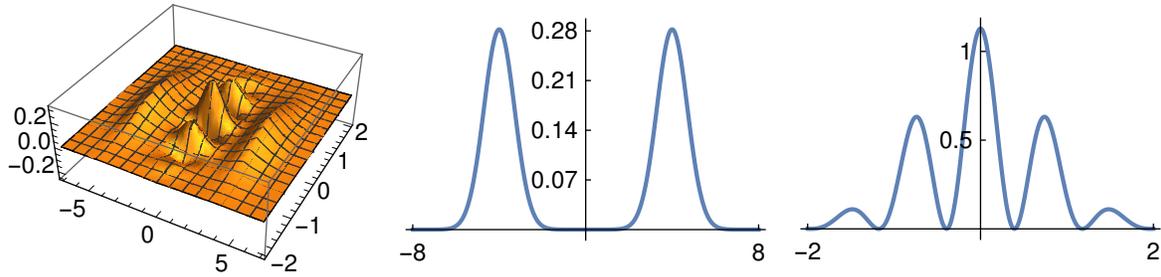


Figure 3.4: On the left, a plot of the Wigner function of a pure state $\rho = |\psi\rangle\langle\psi|$ that is a sum of two minimal uncertainty states: that is, $|\psi\rangle = 1/\sqrt{2}(|\zeta_1\rangle + |\zeta_2\rangle)$, where $|\zeta_j\rangle$ has position $x = 1$ and momentum $y = \pm 4$. Note the negative sections. Note the uneven dimensions on the position and momentum axes. In the middle, the marginal distribution of position, and on the right, that of momentum. Notice that these are both bona fide probability distributions.

Thus the Wigner function seems like a good candidate for a probability distribution that represents the probability distributions of both position and momentum. There is a problem, however, in that the Wigner function may take negative values: we give an example of this in Figure 3.4. The negativity of the Wigner function has been linked to non-classical features of the state, see, for example, [31]. This begs the question: can we classify states with non-negative Wigner functions?

The well-known answer to this question, at least for pure states, was given for single particle systems in 1974 by Hudson [24] and later extended to multi-particle systems by Soto and Claverie [32]:

Theorem 3.4.4 (Hudson's Theorem, generalized by Soto and Claverie). *Let $\psi \in L^2(\mathbb{R}^n)$ be a state vector. The Wigner function of ψ is non-negative if and only if ψ is a Gaussian state.*

Hudson's proof is short and direct, but indeed it only holds for pure states ψ . By the additivity of the expectation, we can see that any mixture of Gaussian states will also have a non-negative Wigner function. Some work [10] has been done toward the understanding of mixed states with positive Wigner function, but no constructive characterization has been found.

In Figure 3.4, we have an example of a Wigner function that has negative sections. The state in question is a pure state of the form $\rho = |\psi\rangle\langle\psi| = \frac{1}{2}(|\zeta_1\rangle\langle\zeta_1| + |\zeta_1\rangle\langle\zeta_2| + |\zeta_2\rangle\langle\zeta_1| + |\zeta_2\rangle\langle\zeta_2|)$. Note that while ρ is pure, $|\psi\rangle$ is not a minimal uncertainty state (rather, it's a combination of minimal uncertainty states), so that this example does not contradict Hudson's Theorem. In Figure 3.5, $\rho = 1/2(|\zeta_1\rangle\langle\zeta_1| + |\zeta_2\rangle\langle\zeta_2|)$ is a mixture of two pure minimal uncertainty states, which should each have nonnegative Wigner, and by additivity of the Wigner function so should their mixture. Both of these figures have been adapted from [11].

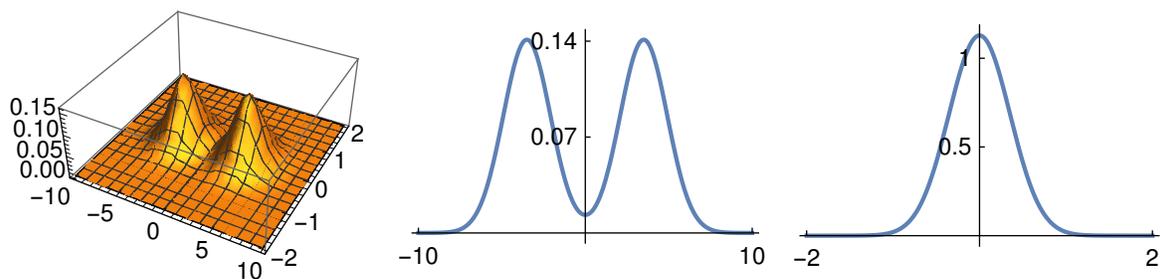


Figure 3.5: On the left, a plot of the Wigner function of a mixture ρ of two minimal uncertainty (ie Gaussian) states, with momentum $y = 2$ and positions $x = \pm 3$. That is, $\rho = 1/2(|\zeta_1\rangle\langle\zeta_1| + |\zeta_2\rangle\langle\zeta_2|)$. In the middle, the marginal distribution of position, and on the right, that of momentum. Again, note the uneven domains of these two plots.

Chapter 4

Approaches to Discrete Gaussian States

In the continuous case, the theory revolves around the self-adjoint operators of position and momentum, denoted \hat{q}, \hat{p} , respectively, and their canonical commutation relation,

$$[\hat{q}, \hat{p}] = iI. \tag{4.1}$$

It is easy to see that for any two matrices $A, B \in M_d(\mathbb{C})$, $\text{Tr}([A, B]) = 0$. Thus, there are no matrices $p, q \in M_d(\mathbb{C})$ that satisfy

$$[q, p] = iI.$$

Stones's theorem together with Weyl's formalism motivates a paradigm shift, where instead of considering self-adjoint observables we work with unitary operators.

We begin by reviewing Gross' version of Hudson's theorem which shows that pure states have nonnegative Wigner function if and only if they are *stabilizer states*, a popular and well-studied class of states in the literature, already considered by many to be the discrete analogue of Gaussian states. Next, we outline the recently introduced discrete analogues of the minimal uncertainty and number states, which, as we have seen, form the building blocks for two different equivalent notions of Gaussianity in the continuous case.

We then introduce a new notion of Gaussian state, based on a discretization of the Gibbs formalism seen in section 3.4.2. As this all takes place in the finite-dimensional setting, we give `Mathematica` implementations of almost all of the topics seen in this chapter.

4.1 Hudson states

Much work has been done to extend the idea of the Wigner function to finite-dimensional systems: we point in particular to [18]. In the discrete case it continues to be only a *quasi* probability distribution, and indeed it shares other characteristics of the continuous distribution, such as translational covariance when multiplying by discrete Weyl operators and the ability to compute marginal probabilities that behave as classical probability distributions. See equation 4.7 for the definition. These similarities suggest the possibility of another: if there is a class of pure states whose Wigner functions are non-negative, then it might make sense to label these "Gaussian" in analogy to the continuous case.

In fact, David Gross [20] has characterized, at least for systems with odd dimension, the set of such states, often called *stabilizer states* in the quantum information literature, in the following theorem:

Theorem 4.1.1 (Gross' Theorem). *Let d be odd, $n \in \mathbb{N}$ and $\psi \in L^2(\mathbb{Z}_d^n)$ be a state vector. The Wigner function of ψ is non-negative if and only if ψ is a stabilizer state.*

If ψ has full support, that is, $\psi(q) \neq 0$ for all $q \in \mathbb{Z}_d^n$, then ψ is a stabilizer state if and only if it has the form

$$\psi(q) \propto \exp\left(\frac{2\pi i}{d}(q\theta q + xq)\right),$$

where $q, x \in \mathbb{Z}_d^n$ and θ is a $n \times n$ symmetric matrix with entries in \mathbb{Z}_d .

Note first of all the similarities between the state vector given here and a continuous Gaussian state vector (that is, a minimal uncertainty state, 3.59). Indeed, Gross notes that this theorem adds further evidence to the idea that stabilizer states are in fact the natural finite-dimensional analogue of Gaussian states. We will review the argument given by Gross; along the way, we will see the discretization of several notions from the continuous case, and in particular we will see the discrete phase space.

We consider a d -dimensional quantum system for an odd number d . Let $L^2(\mathbb{Z}_d)$ be its Hilbert space, and denote by $\{|0\rangle, \dots, |d-1\rangle\}$ the standard basis, labeled by elements of \mathbb{Z}_d . We will refer to \mathbb{Z}_d as the *configuration space* and denote it by Q .

We now construct the discrete *Weyl operators*. Letting $\chi(q) = \exp(\frac{2\pi i}{d}q)$, define the *shift* and *boost* operators by the relations

$$X(p)|x\rangle = |x+p\rangle, \quad Z(q)|x\rangle = \chi(qx)|x\rangle.$$

Note that these are nearly identical to those given in the continuous case. Then for $p, q \in \mathbb{Z}_d$, we define the Weyl operators by

$$w(p, q) = \chi(-2^{-1}pq)Z(q)X(p).$$

As 2^{-1} does not exist in a ring of the form \mathbb{Z}_{2n} , we see the necessity of odd d .

Note the phase factors on the operator $w(z)$ for $z \in Q \times Q$. The factor χ is introduced to make $w(z)$ *involutive*, so that $w(z)^* = w(z)^{-1} = w(-z)$.

For $z, z' \in Q \times Q$, where $z = (p, q)$ and $z' = (p', q')$, one can directly calculate the composition law

$$w(z)w(z') = \chi(2^{-1}[z, z'])w(z+z'), \tag{4.2}$$

where $[z, z']$ is the now-familiar symplectic inner product. As we will often work in $Q \times Q$, we henceforth call it the *phase space*, and denote $V = \mathbb{Z}_d \times \mathbb{Z}_d$. One should pause and compare this to the notion of phase space discussed above.

We can generalize this notion naturally to the case of n subsystems. Consider now a configuration space $Q = \mathbb{Z}_d \times \dots \times \mathbb{Z}_d$. We can write the multiplication of two elements $p, q \in Q$ by $pq = \sum_i^n p_i q_i \in Q$. The Hilbert space corresponding to this system is $\mathcal{H} = L^2(\mathbb{Z}_d) \otimes \dots \otimes L^2(\mathbb{Z}_d)$ and is spanned by $\{|q\rangle\}_{q \in Q}$, and denoting w_j the Weyl operator defined on the j th system as above, we can write

$$w(p, q) = w(p_1, \dots, p_n, q_1, \dots, q_n) = w_1(p_1, q_1) \otimes \dots \otimes w_n(p_n, q_n).$$

The associated phase space is again given by $V = Q \times Q$.

Denote by $\delta_{p,q}$ the Kronecker delta, that is, the function that is 1 if $p = q$ and 0 otherwise. One can compute

$$\text{Tr}w(p, q) = |Q|\delta_{p,0}\delta_{q,0}. \quad (4.3)$$

This, in combination with the composition law 4.2, shows that the Weyl operators $\{w(z)\}_{z \in V}$ form an orthonormal basis of the space of operators on \mathcal{H} with respect to the Hilbert-Schmidt inner product $\frac{1}{|Q|}\text{Tr}(\cdot^* \cdot)$. Thus, for a density operator ρ , we may define the *characteristic function* ϕ_ρ by its expansion coefficients in the Weyl basis:

$$\phi_\rho(z) = |Q|^{-1}\text{Tr}(w(z)^*\rho) \quad (4.4)$$

Note that our definition of the Weyl operators ensures that ϕ_ρ is *Hermitian*: that is, $\phi_\rho(-z) = \overline{\phi_\rho(z)}$. Indeed, using the fact that $\overline{\text{Tr}(A)} = \text{Tr}(A^*)$ for a matrix A ,

$$\phi_\rho(-z) = |Q|^{-1}\text{Tr}(w(-z)^*\rho) = |Q|^{-1}\overline{\text{Tr}(\rho w(-z))} = |Q|^{-1}\overline{\text{Tr}(w(z)^*\rho)} = \overline{\phi_\rho(z)}. \quad (4.5)$$

Hermiticity is an important feature of characteristic functions in probability theory and so it is worth adding a phase factor to preserve this.

As in the continuous case, the Wigner function is the symplectic Fourier transform of the characteristic function. We define the symplectic Fourier transform $F_S f$ for a function $f : V \rightarrow \mathbb{C}$ by

$$(F_S f)(a) = |Q|^{-1/2} \sum_{b \in V} \bar{\chi}([a, b]) f(b). \quad (4.6)$$

Here, we define the character χ for an element $q \in Q$ by $\chi(q) = \chi(q_1) \dots \chi(q_n)$. Now, for a characteristic function ϕ_ρ , we can compute

$$\begin{aligned} (F_S \phi_\rho)(a) &= |Q|^{-2} \sum_{b \in V} \bar{\chi}([a, b]) \text{Tr}(w(b)^*\rho) \\ &= |Q|^{-1} \text{Tr} \left(|Q|^{-1} \sum_b \bar{\chi}([a, b]) w(b)^* \right) \rho. \end{aligned}$$

Defining the *point space operators* $A(a) = |Q|^{-1} \sum_b \bar{\chi}([a, b]) w(b)^*$, we can define the *Wigner function associated to ρ* by

$$W_\rho(a) = |Q|^{-1} \text{Tr}(A(a)\rho). \quad (4.7)$$

The discrete Wigner function satisfies most of the relations of the continuous one. We will show a few here for completeness. The following adaptation of the sum of roots of unity will be required throughout:

$$\sum_{x \in Q} \chi(xy) = |Q|\delta_{0,y}. \quad (4.8)$$

The first thing to note is that, for a pure state $\rho = |\psi\rangle\langle\psi|$, the Wigner function is computed very similarly to the continuous one:

$$W_\rho(p, q) = \frac{1}{d^2} \text{Tr} \left(\sum_{x,y} \bar{\chi}([(p, q), (x, y)]) w(x, y)^* \left(\sum_{i,j} \rho_{i,j} |i\rangle\langle j| \right) \right)$$

$$\begin{aligned}
&= \frac{1}{d^2} \text{Tr} \left(\sum_{x,y} \chi(qx - py + \frac{1}{2}xy) \sum_{i,j} \chi(-xi) \rho_{i,j} |i-y\rangle \langle j| \right) \\
&= \frac{1}{d^2} \sum_k \sum_{x,y} \chi(qx - py + \frac{1}{2}xy) \sum_{i,j} \chi(-xi) \rho_{i,j} \langle k|i-y\rangle \langle j|k\rangle \\
&= \frac{1}{d^2} \sum_{x,y} \chi(qx - py + \frac{1}{2}xy) \sum_k \chi(-x(k+y)) \rho_{k+y,k} \\
&= \frac{1}{d^2} \sum_y \chi(-py) \sum_k \rho_{k+y,k} \sum_x \chi(x(q - \frac{1}{2}y - k)) \\
(\text{so } k = q - \frac{1}{2}y) \quad &= \frac{1}{d} \sum_y \chi(-py) \rho_{q+1/2y, q-1/2y} \\
&= \frac{1}{d} \sum_y \bar{\chi}(py) \psi(q + \frac{1}{2}y) \bar{\psi}(q - \frac{1}{2}y),
\end{aligned}$$

where we have identified $\rho_{i,j}$ with $\psi(i)\bar{\psi}(j)$. From this we can see that the Wigner function behaves well with respect to mixed states.

The Wigner function again encodes the probability distribution of both position and momentum in its marginals:

$$\begin{aligned}
\sum_p W_\rho(p, q) &= \frac{1}{d} \sum_p \left(\sum_y \bar{\chi}(py) \psi(q + \frac{1}{2}y) \bar{\psi}(q - \frac{1}{2}y) \right) \\
&= \frac{1}{d} \sum_y \psi(q + \frac{1}{2}y) \bar{\psi}(q - \frac{1}{2}y) \sum_p \chi(-py) \\
&= \frac{1}{d} d \psi(q) \bar{\psi}(q) \\
&= |\psi(q)|^2.
\end{aligned}$$

Using the Fourier transform, it is possible to use a similar calculation to get the momentum distribution. Finally, the Wigner function is once again normalized: from the last equation, summing along the momentum variable p gives $|\psi|^2$ which is equal to 1.

We have already quoted Gross' main result, (Theorem 4.1.1), which states that a pure state has non-negative Wigner function if and only if it is a *stabilizer state*. We have not discussed stabilizer states yet, so we will give a very brief exposition now. First introduced by Gottesman [19], stabilizer states are the mutual eigenvectors of sets of commuting Weyl operators. Let us unpack this statement.

Recall that, for a group G that acts on a set X , the *stabilizer* G_x of a point $x \in X$ is the set of group elements that fix x :

$$G_x = \{g \in G \mid gx = x\}.$$

One can also consider the stabilizer of a subset $Y \subseteq X$, denoted G_Y . We can view the Weyl operators as a group that acts on the set of quantum state vectors.

Note that, by the composition law 4.2, two Weyl operators $w(z_1)$ and $w(z_2)$ commute if and only if $[z_1, z_2] = 0$. The image of a set of points $S \subseteq V$ under the Weyl representation

consists of mutually commuting operators if and only if the symplectic form vanishes on the entire set: that is, $w(S) = \{w(z) \mid z \in S\}$ is mutually commuting if and only if $[z_1, z_2] = 0$ for all $z_1, z_2 \in S$. In this case, $w(S)$ has a set of common eigenvectors. One can show that if, in addition, the set S is large enough, the eigenspaces corresponding to these eigenvectors become non-degenerate [20, Lemma 8], and thus we can single out vectors based on the set of Weyl operators that stabilize them.

Given a subset S of the phase space V , we can define its *symplectic complement* S^Δ by

$$S^\Delta = \{v \in V \mid [z, v] = 0 \text{ for all } z \in S\}.$$

This set is always a subgroup under addition. We say that a set S is *isotropic* if $S \subseteq S^\Delta$, *co-isotropic* if $S^\Delta \subseteq S$, and *Lagrangian* if $S = S^\Delta$, in which case S is a subgroup.

If d is prime, then a simple example of a Lagrangian subgroup of $V = \mathbb{Z}_d \times \mathbb{Z}_d$ can be given as the points on a *line* L through the origin. Such a line is a set of the form $L_m = \{(mq, q) \mid q \in \mathbb{Z}_d\}$. It is easy to see that any set of this form is Lagrangian: indeed, if $(mq_1, q_1), (mq_2, q_2) \in L_m$, then $[(mq_1, q_1), (mq_2, q_2)] = mq_1q_2 - mq_2q_1 = 0$ for all $q_1, q_2 \in \mathbb{Z}_d$.

Suppose S is a Lagrangian subgroup of V , and let $v \in V$. Up to a phase, the stabilizer state $|S, v\rangle$ corresponding to S and v is the unique state vector that satisfies the eigenvalue equations

$$\chi([v, z])w(z)|S, v\rangle = |S, v\rangle$$

for all $z \in S$. One can verify that the one-dimensional projection operator corresponding to this state is given by

$$|S, v\rangle\langle S, v| = |S|^{-1} \sum_{z \in S} \chi([v, z])w(z).$$

The uniqueness follows from a dimension-counting argument, and comparing that to the possible number of characters in S (see [20, Corollary 26]).

It is well-known (see [13], for example) that the coefficients of stabilizer state vectors are described by quadratic forms. We quote the following to make this notion precise [20, Lemma 18]:

Lemma 4.1.2. *Let ψ be a state vector on a d^n dimensional system. Then ψ is a stabilizer state with $\psi(q) \neq 0$ for all $q \in Q$ if and only if there exists a symmetric $n \times n$ matrix θ and an $x \in Q$ such that*

$$\psi(q) = \omega^{q\theta q + xq},$$

where ω is a root of unity.

The vector ψ in the form above is what one might consider the analogue of a *pure Gaussian state* in finite dimensions. These are exactly the pure states with nonnegative Wigner functions; due to the parallels to Hudson's theorem, we call states of this form *Hudson states*.

Analogously to the continuous case, there are limited results classifying mixed states with non-negative Wigner function. It is easy to see that a convex combination of Hudson states will have non-negative Wigner function, and one might conjecture that all non-negative Wigner states are mixtures of Hudson states. However, a simple counterexample is given by Gross [20, Section 5]. In the next section we will give an entire class of mixed states with non-negative Wigner function whose component states have negative Wigner function.

We soon begin with the `Mathematica` implementation. First, we need to revisit the Weyl operators. While these operators are unitary and involutive, it turns out that they are not *periodic*: that is, due to the factor of $1/2$ in the phase, it is possible that

$$w(p, q) \neq w(p + d, q), \quad \text{or} \quad w(p, q) \neq w(p, q + d). \quad (4.9)$$

While this is not a big problem, we find that this aperiodicity complicates plots of the characteristic function. Because of this, we find it appropriate to introduce an additional phase factor:

$$w(p, q) = (-1)^{pq} \chi\left(-\frac{1}{2}pq\right) Z(q) X(p). \quad (4.10)$$

This additional factor guarantees periodicity, makes the characteristic function more symmetric, does not affect involutivity, and does not complicate formulae unduly. For example, the commutation relation can be rewritten as

$$w(p, q)w(p', q') = (-1)^{-(pq' + qp')} \chi(2^{-1}[(p, q), (p', q')]) w(p + p', q + q'). \quad (4.11)$$

Now, initialize a single 13-dimensional system:

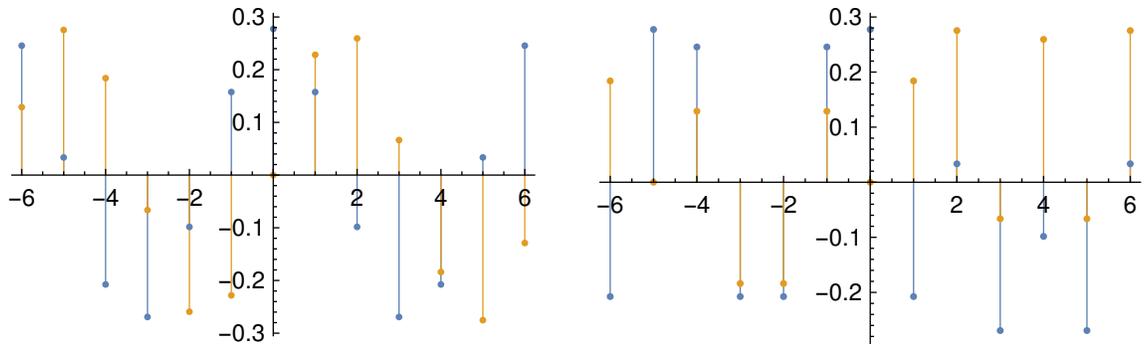
```
In[23]:= qInit[1, {13}]
```

```
Out[23]= {q1}
```

The function `complexKetPlot` takes a ket and plots it; here are two Hudson states.

```
In[24]:= GraphicsRow[{complexKetPlot[ket1=gaussianKet[1, 0, 2], False],
  complexKetPlot[ket2=gaussianKet[1, 3, 2], False]}, ImageSize->Full]
```

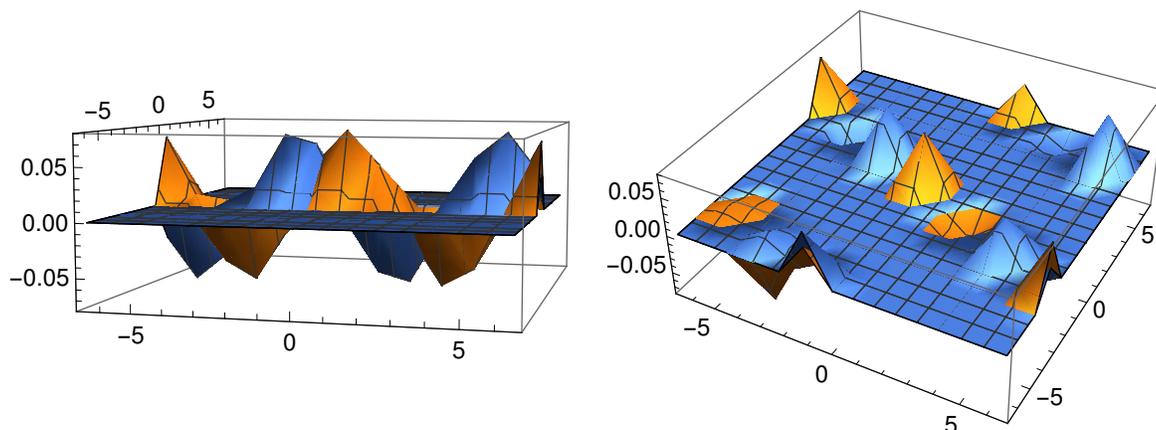
```
Out[24]=
```



The Weyl operators are implemented as `WSparse`. It is named this way because it is implemented as a *sparse array*: that is, a matrix where most of the entries are zero. Many numerical analysis programs have specialized functions to deal with sparse arrays quickly and efficiently. For example, the `SparseArray` object in `Mathematica` only stores the nonzero values, and multiplication between two such objects ignores any rows that are all or mostly zero, thus saving a lot of time.

```
In[25]:= WSparse[3, 2]
```

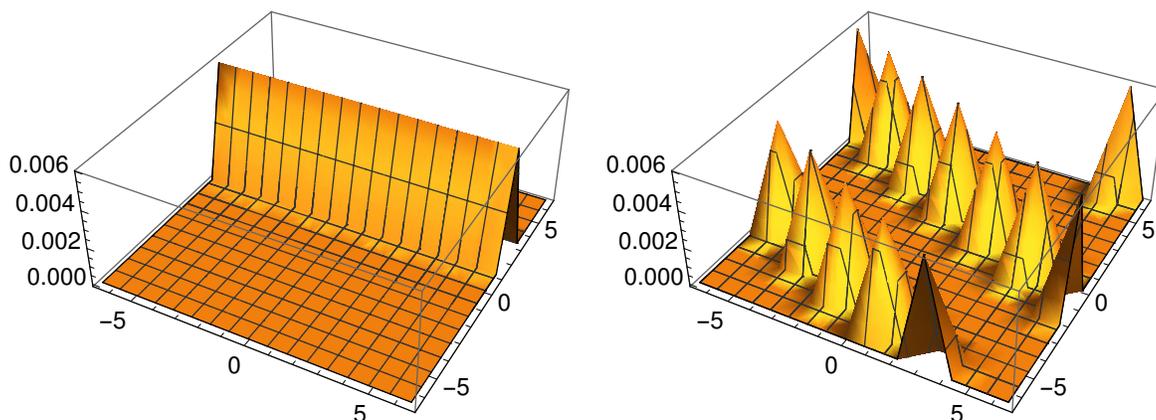

Out[28]=



And here is the Wigner function of both states. As they are Hudson states, we expect the Wigner functions to be nonnegative.

```
In[29]:= GraphicsRow[{charfcnPlot[wigner[ket1**hc[ket1]]],
charfcnPlot[wigner[ket2**hc[ket2]]]}, ImageSize→Full]
```

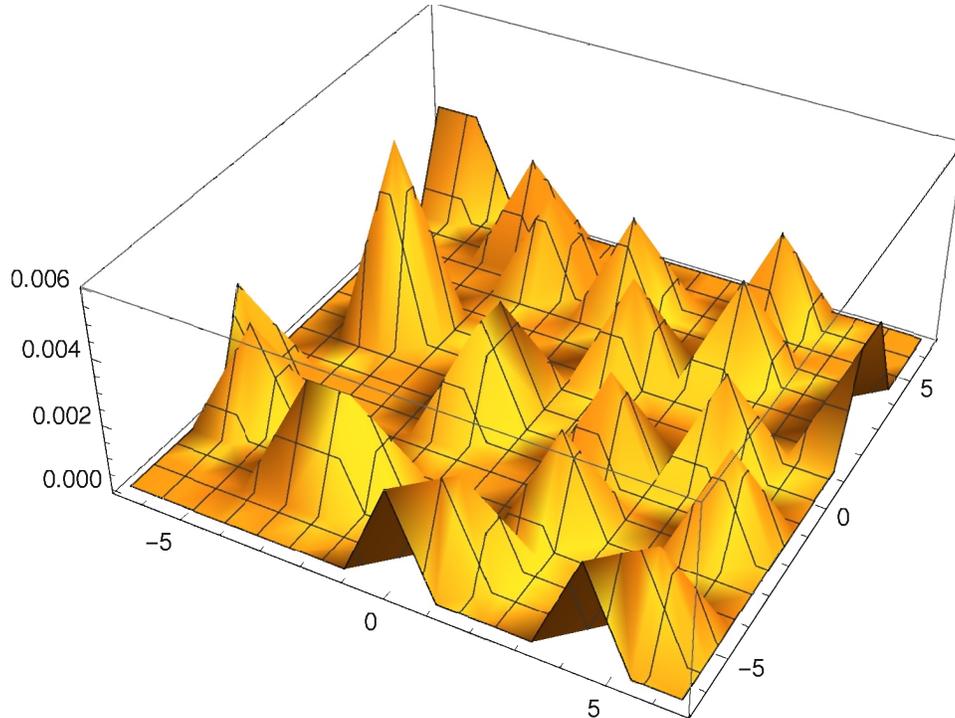
Out[29]=



The Wigner function of a mixture of Hudson states; as expected, this Wigner function is also nonnegative.

```
In[30]:= ket3=gaussianKet[1,2,4]//N;ket4=gaussianKet[1,4,7]//N;
charfcnPlot[wigner[ $\frac{1}{2}$ (ket3**hc[ket3] + ket4**hc[ket4])]]
```

Out[30]=



4.2 Bernstein States

Minimal uncertainty states for finite-dimensional systems have recently appeared in [21]. Motivated by the representation

$$S = \frac{1}{\pi} \int_{\mathbb{C}} |\zeta\rangle \langle \zeta| p(\zeta) d^2\zeta \quad (4.12)$$

of continuous Gaussian states as Gaussian combinations of coherent states, a natural discrete analogue would be “Gaussian” combinations of minimal uncertainty states from [21]. As Gaussian distributions over finite abelian groups are quite restrictive, we use the notion of Bernstein Gaussian distributions as our convex coefficients. We begin with an outline of the main results in [21].

We consider a d -dimensional system with Hilbert space $\mathcal{H} = L^2(\mathbb{Z}_d)$. Fix the standard basis $\{|q\rangle\}_{q=0}^{d-1}$ which we will interpret as the *position* coordinate q . We fix also another basis related by the (inverse) Fourier transform,

$$|\tilde{p}\rangle = \frac{1}{\sqrt{d}} \sum_{q=0}^{d-1} \chi(pq) |q\rangle, \quad p = 0, \dots, d-1,$$

where again $\chi(q) = e^{2\pi i q/d}$. We interpret the index p as the *momentum* coordinate. These two bases are *unbiased*, meaning that the overlap $|\langle q|\tilde{p}\rangle| = \frac{1}{\sqrt{d}}$ for all indices q, p .

The discrete Fourier transform has its own version of an uncertainty principle: indeed, if $x = \{x_1, \dots, x_d\}$ is a signal of length d , and $X = \{X_1, \dots, X_d\}$ is the Fourier transform of x ,

let N_x be the number of nonzero elements of x and respectively for N_X . Then one can show [14] that we have the following lower bound:

$$N_x \cdot N_X \geq d.$$

So if the signal x is sharply concentrated on a single point, then the signal X must be smeared out and nonzero everywhere. We seek an analogous notion for quantum states.

In order to do this we recall the unitary *shift* X and *boost* Z operators we saw in the previous section. One can see that they can be written in terms of our two bases as

$$X = \sum_{p=0}^{d-1} \chi(-p) |\widetilde{p}\rangle \langle \widetilde{p}|, \quad Z = \sum_{q=0}^{d-1} \chi(q) |q\rangle \langle q|.$$

Note that the shift and boost operators act as translation operators for the position and momentum bases, respectively:

$$X|q\rangle = |q+1\rangle, \quad Z|\widetilde{p}\rangle = |\widetilde{p+1}\rangle$$

Note that the minus in the exponent in the definition of the shift operator X is there because we related our bases by the *inverse* Fourier transform. We accept this notational inconvenience to simplify later formulae. Recall that the coordinates obey addition modulo the dimension d :

$$|q+d\rangle = |q\rangle, \quad |\widetilde{p+d}\rangle = |\widetilde{p}\rangle.$$

We can write a general state $|\phi\rangle$ in terms of either basis:

$$|\phi\rangle = \sum_{q=0}^{d-1} c_q |q\rangle = \sum_{p=0}^{d-1} \widetilde{c}_p |\widetilde{p}\rangle, \quad (4.13)$$

where c_q and \widetilde{c}_p are expansion coefficients in the position and momentum bases, respectively. Then we can compute the expected values for X and Z for the state $|\phi\rangle$ in two different ways:

$$\begin{aligned} \langle \phi | X | \phi \rangle &= \sum_{q=0}^{d-1} c_{q+1}^* c_q = \sum_{p=0}^{d-1} \chi(-p) |\widetilde{c}_p|^2 \\ \langle \phi | Z | \phi \rangle &= \sum_{q=0}^{d-1} \chi(q) |c_q|^2 = \sum_{p=0}^{d-1} \widetilde{c}_{p+1}^* \widetilde{c}_p \end{aligned}$$

Let us examine the expectation $\langle \phi | Z | \phi \rangle$ in terms of the position expansion c_q . This is a convex combination of the d roots of unity $\chi(q)$ with coefficients $|c_q|^2$. Thus, the value $\langle \phi | Z | \phi \rangle$ lies somewhere in the regular d -gon with vertices given by the points $\{\chi(q)\}$. It is clear that $|\langle \phi | Z | \phi \rangle|$ lies between 0 and 1. The only way for it to be 1 is if the position coordinate has a sharp value, say $c_{q_j} = 1$ for some j . In this case, $\langle \phi | Z | \phi \rangle = \chi(q_j)$. Otherwise, we have $|\langle \phi | Z | \phi \rangle| < 1$. On the other hand, say the position coordinate is equally weighted, so that $|c_q|^2 = 1/d$ for all q . In this case, $\langle \phi | Z | \phi \rangle$ is located at the origin, so that $|\langle \phi | Z | \phi \rangle| = 0$. Thus,

we may consider $|\langle\phi|Z|\phi\rangle|$ as a measure of how sharply the position coordinate is distributed. By a similar discussion, we can view $|\langle\phi|X|\phi\rangle|$ as a measure of how sharply the momentum coordinate is distributed.

Note that the quantities $|\langle\phi|Z|\phi\rangle|$ and $|\langle\phi|X|\phi\rangle|$ can not simultaneously be 1. Indeed, if $|\langle\phi|Z|\phi\rangle| = 1$, then its coefficients $\{c_q\}$ must be 0 except for one of them. In this case, $|\langle\phi|X|\phi\rangle| = 0$, as can be seen by its expansion in terms of c_q .

We define the *certainty* C of a state $|\phi\rangle$ to be

$$C(|\phi\rangle) = |\langle\phi|X|\phi\rangle\langle\phi|Z|\phi\rangle|,$$

which can be thought of as a simultaneous measure of the uncertainty with respect to the two sets of coordinates. If $C = 0$, then the state $|\phi\rangle$ is sharply localized in one of the two bases, so we have maximal certainty about the position or momentum of the state. A larger C indicates we are less certain about both bases, as the name ‘‘certainty’’ indicates.

Note that the certainty C is trivially bounded above by 1. A natural question is whether this bound is optimal, and for which states do we see the optimal value realized. The answer to these questions will be given by the next theorem in relation to the *Harper operator*,

$$H = \frac{X + X^* + Z + Z^*}{4}.$$

This operator has physical significance, which we will explore in section 5. The following is from [21].

Theorem 4.2.1 (Discrete Minimal Uncertainty States). *For any state vector $|\phi\rangle$, the certainty C is bounded by the inequality*

$$C(|\phi\rangle) = |\langle\phi|X|\phi\rangle\langle\phi|Z|\phi\rangle| \leq h^2, \quad (4.14)$$

where h is the greatest eigenvalue of the Harper operator H . Equality holds if and only if

$$|\phi\rangle = X^\alpha Z^\beta |\Gamma\rangle, \quad (4.15)$$

where $|\Gamma\rangle$ is the eigenvector of H with the greatest eigenvalue, and $\alpha, \beta \in \{0, \dots, d-1\}$.

Since they maximize the certainty we call $|\alpha, \beta\rangle := X^\alpha Z^\beta |\Gamma\rangle$ the *minimal uncertainty states*. One can easily see that they are not mutually orthogonal; it is harder to see that they comprise an overcomplete set in \mathbb{C}^d if only if d is odd [21, Section V]. A minimal uncertainty state is plotted in Figure 4.1.

As in the continuous case, we now have a way of assigning, to every point $\zeta = (\alpha, \beta)$ in phase space V , a pure state $|\zeta\rangle\langle\zeta|$. Following the recipe of 3.66, we define, for a probability distribution $p(\zeta)$ on V , the state

$$S_p = \sum_{\zeta} p(\zeta) |\zeta\rangle\langle\zeta|.$$

Specializing p to be Gaussian in the sense of Bernstein, we arrive at a new class of states:

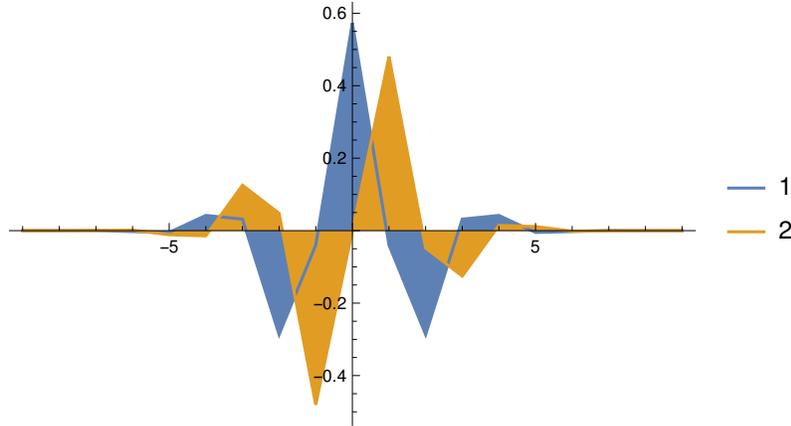


Figure 4.1: Plot of the minimal uncertainty state $|0, 5\rangle$, in dimension 19, with real part in blue and imaginary in yellow. Compare this to the continuous minimal uncertainty state in Figure 3.1.

Definition 4.2.2 (Bernstein States). A density operator ρ on $L^2(\mathbb{Z}_d)$ is called *Bernstein* if there exists a Bernstein Gaussian distribution γ such that $\rho = S_\gamma$.

Let us give an example of such a state. Let L be a Lagrangian subgroup of V . Taking $\gamma(\zeta) = |L|^{-1}\chi_L(\zeta)$ to be the uniform distribution over the elements of L , we get

$$S_\gamma = \frac{1}{|L|} \sum_{\zeta \in L} |\zeta\rangle\langle\zeta|. \quad (4.16)$$

We note that the set $w(L)$ stabilizes the state S_L , in that $w(\zeta)S_Lw(\zeta)^* = S_L$ for all $\zeta \in L$. One can see this by recalling that L is a subgroup under addition and by using the commutation relation.

These Bernstein states are a discrete analogue to the states given in 3.66, where we choose to use a Bernstein notion of Gaussianity. It is also reasonable to choose a Gaussian probability distribution, which is known to simply be a point mass. So for a point $\zeta \in V$ we could write

$$S_\zeta = |\zeta\rangle\langle\zeta| \quad (4.17)$$

and this is the analogue of the zero-variance degenerate Gaussian distribution.

These states have at least one interesting feature: as we will see in the **Mathematica** implementation, though all of the minimal uncertainty states have negative sections in their Wigner function, many Bernstein states have nonnegative Wigner function. As we have discussed in section 4.1, a mixture of states with nonnegative Wigner function has nonnegative Wigner function, but a mixture of states with negative Wigner function having nonnegative Wigner function is interesting. We have strong numerical evidence to suggest that every Bernstein state has nonnegative Wigner function.

Initialize a system with dimension 13:

```
In[31]:= qInit[1, {13}]
```

```
Out[31]= {q1}
```

We verify the certainty relation for the minimal uncertainty states $|\Gamma\rangle$ and $|1,2\rangle$:

```
In[32]:= gamma=minUncerVec[];
minUncerVal[]^2
Abs[(hc[gamma]**Z[1,1]**gamma **hc[gamma]**X[1,1]**gamma)[[1]]//Chop]
gamma=ABstate[1,2];
Abs[(hc[gamma]**Z[1,1]**gamma **hc[gamma]**X[1,1]**gamma)[[1]]//Chop]
```

```
Out[32]= 0.785604
```

```
Out[33]= 0.785604
```

```
Out[34]= 0.785604
```

These values are equal, as expected. We can generate a line in phase space, and check if it is a Lagrangian subset:

```
In[35]:= line=lineInPhaseSpace[1,3]
lagrangianQ[line]
```

```
Out[35]= {{0,0},{3,1},{6,2},{9,3},{12,4},{2,5},{5,6},{8,7},
{11,8},{1,9},{4,10},{7,11},{10,12}}
```

```
Out[36]= True
```

We can also check if subsets are isotropic, meaning they are contained in their symplectic complement, or coisotropic, meaning they contain their symplectic complement.

```
In[37]:= points = line[{{2,3,4,8}}]
isotropicQ[points]
coisotropicQ[Union[sympComplement[points],{{10,11}}]]
```

```
Out[37]= {{3,1},{6,2},{9,3},{8,7}}
```

```
Out[38]= True
```

```
Out[39]= True
```

The Wigner function of a pure minimal uncertainty state is negative:

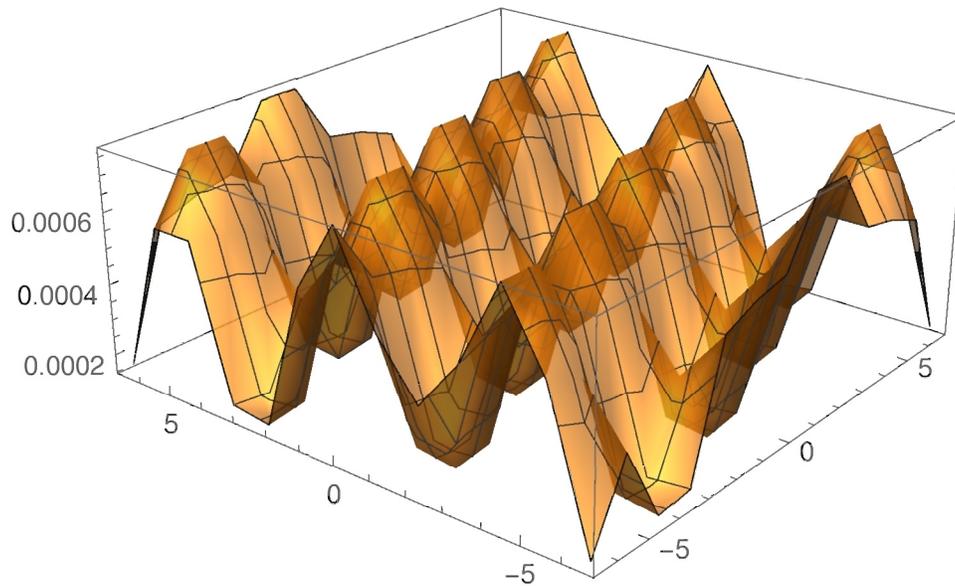
```
In[40]:= AllTrue[NonNegative[wigner[(ab=ABstate[1,3])**hc[ab]]//Chop],TrueQ]
```

```
Out[40]= False
```

But if we take a uniform combination over a line in phase space, we get a nonnegative Wigner function:

```
In[41]:= charfcnPlot[wigner[subsetState[line]//Chop]//Chop]
```

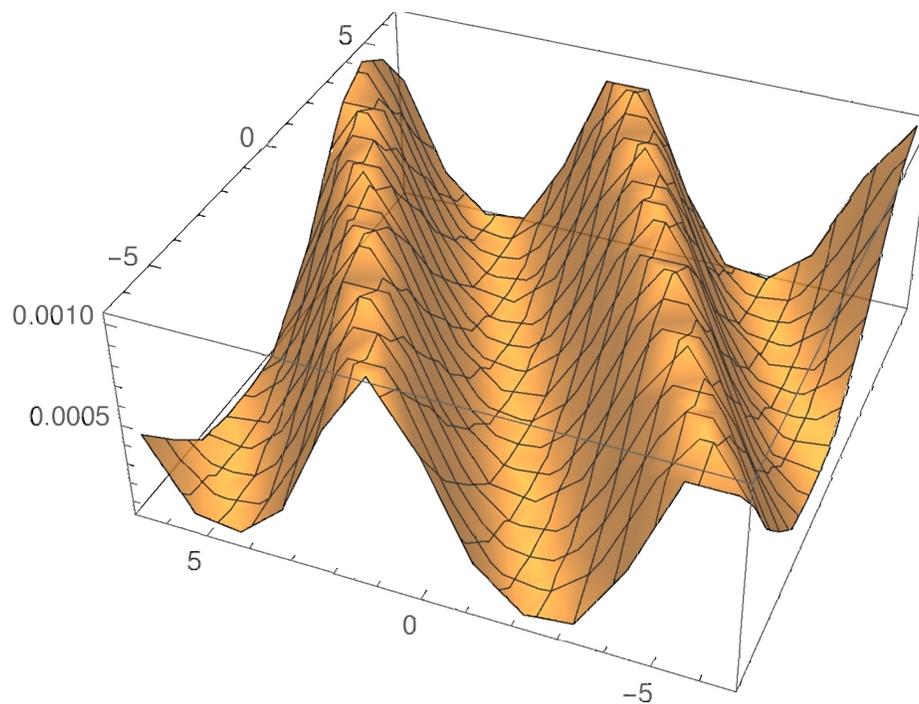
```
Out[41]=
```



We can even choose a line that has been shifted:

```
In[42]:= charfcnPlot[wigner[subsetState[lineInPhaseSpace[1,2,{0,4}]]//Chop] //Chop]
```

```
Out[42]=
```



4.3 The Harper Functions

We would now like to better understand the Harper operator, $H = 1/4(X + X^* + Z + Z^*)$. We will justify in section 5.2 that this operator is a choice for a discrete version of the Hamiltonian of the harmonic oscillator. Since H is self-adjoint, it has real eigenvalues, and one can show using the Perron-Frobenius theorem (see [21, Section B.2]) that the eigenvector associated to its eigenvalue with largest modulus is real and nondegenerate. We denoted this vector $|\Gamma\rangle$ and called it a *minimal uncertainty state*. We will see in this section another way to interpret this state and the rest of the eigenstates of H .

If we accept H as our harmonic oscillator Hamiltonian, then in analogy to the continuous case, the eigenvectors of H should be discrete versions of the *number states* $|n\rangle$. Recall that the number states were given by the *Hermite-Gaussians* ψ_n in equation 3.61, and these were named for their eigenvalue relation with the number operator \mathcal{N} given by $\mathcal{N}|n\rangle = n|n\rangle$.

Barker in [6] describes the following method for classifying the eigenvectors of H in a space of dimension d . First, note that H commutes with the square of the Fourier transform, also called the *parity operator* since $F^2|x\rangle = |-x\rangle$. This operator has two eigenspaces, corresponding to the eigenvalues ± 1 . Since H commutes with F^2 , they share a common set of eigenvectors. Separate these eigenvectors based on whether they belong to the 1 or -1 eigenspace, then define $|0\rangle, |2\rangle, \dots$ by ordering the former by decreasing eigenvalue h_0, h_2, \dots , and define $|1\rangle, |3\rangle, \dots$ by doing the same to the latter. Thus, we have a set of eigenvectors $\{|n\rangle\}_{n=0}^{d-1}$ and their corresponding eigenvalues $\{h_n\}_{n=0}^{d-1}$. We emphasize that this is *not* the standard basis. We can write

$$H = \sum_{n=0}^{d-1} h_n |n\rangle \langle n| \quad (4.18)$$

in analogy to the continuous case, and this will be handy in the next section.

We call the vectors $|n\rangle$ the *Harper functions* and consider them to be discrete analogues of the number states. In fact, in section 5.1 we will show that the Harper functions *converge* to the Hermite-Gaussians, in a certain sense.

We note that the Harper operator H is a special case of the more general operator

$$H_\theta = \cos \theta \frac{X + X^*}{2} + \sin \theta \frac{Z + Z^*}{2} \quad (4.19)$$

where $\theta = \pi/4$. This operator is also self-adjoint and commutes with the parity operator, so we may play the same game of defining number states, denoted now $|n\rangle_\theta$ with eigenvalues $h_{\theta,n}$. These are analogues of the *squeezed states*, which is clear when comparing the plots of these states to those with different values of θ , as we show in the `Mathematica` implementation below. We can again write the spectral decomposition of H_θ as

$$H_\theta = \sum_{n=0}^{d-1} h_{\theta,n} |n\rangle_\theta \langle n|_\theta. \quad (4.20)$$

Here we drop one of the θ 's for conciseness. Also, in the future we will drop the θ from $h_{\theta,n}$ when the squeezing parameter is clear.

There is no known general formula for the Harper functions $|n\rangle$, other than the vacuum state up to dimension 6, which is given in [26]. Indeed, even the eigenvalues h_j have not been

derived analytically and can only be found numerically. Having a closed form for either of these would be a big step in the direction of understanding better the Harper operator and its implications to quantum information.

The Harper functions are implemented in `qWeyl` as follows. First, initialize a 21 dimensional system:

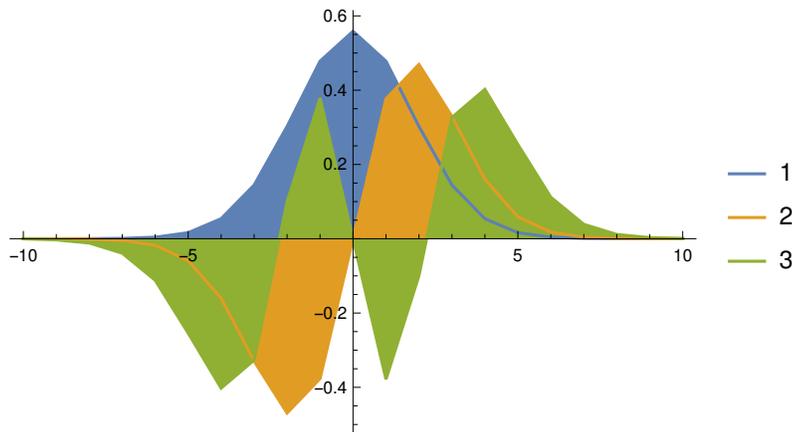
```
In[43]:= qInit[1, {21}]
```

```
Out[43]= {q1}
```

The following generates the Harper functions for $\theta = \pi/4$ and plots $|0\rangle$, $|1\rangle$, and $|3\rangle$:

```
In[44]:= harp2 = generateHarperFunctions [Pi/4];
ketLinePlot[{harp2[[1]], harp2[[2]], harp2[[4]]}]
```

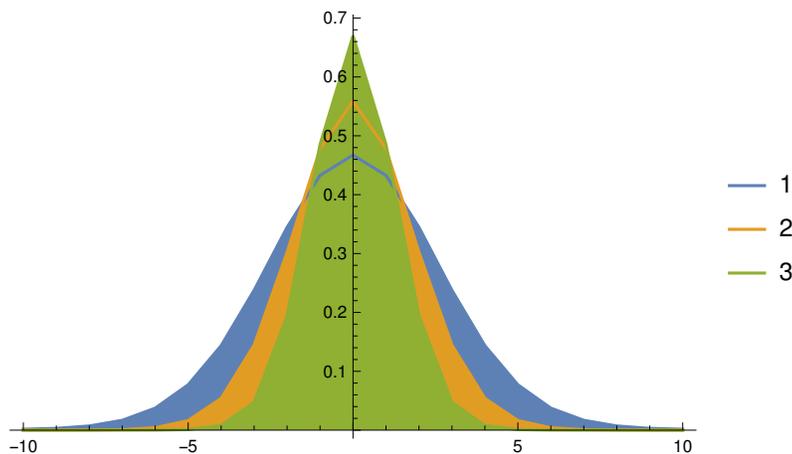
```
Out[44]=
```



The following compares $|0\rangle_\theta$ for $\theta = \pi/12, \pi/4, 5\pi/12$:

```
In[45]:= harp1 = generateHarperFunctions [Pi/12];
harp3 = generateHarperFunctions [5Pi/12];
ketLinePlot[{harp1[[1]], harp2[[1]], harp3[[1]]}]
```

```
Out[45]=
```



4.4 Discrete Gibbs Formalism

In the continuous setting, Gaussian quantum states can be characterized as thermal states of quadratic Hamiltonians in the canonical variables. As shown in [30, Chapter 3], this is equivalent to being in the orbit of Gibbs states of the harmonic oscillator Hamiltonian under conjugation by Weyl operators and metaplectic unitaries. Viewing Harper's equation as a discrete analogue of the harmonic oscillator Hamiltonian, we obtain a discrete analogue of Gaussian states via Gibbs states for Harper's Hamiltonian.

There is good reason to consider the orbit of a state under Weyl operators and metaplectic unitaries. This is a well-studied class of unitaries, typically called the *Clifford unitaries* in the literature. See [19] for more on this group.

We start by considering the (possibly squeezed) Harper operator

$$H_\theta = \cos \theta \left(\frac{X + X^*}{2} \right) + \sin \theta \left(\frac{Z + Z^*}{2} \right). \quad (4.21)$$

This operator is self-adjoint. We note that this expression behaves well with respect to conjugations by Weyl operators, since by the commutation relation this essentially amounts to multiplication by a phase. Also, this form is preserved by conjugation by a metaplectic operator, as we shall see. For these reasons, we try to make this the form of our general Hamiltonian.

For any unitary U and phase $\psi \in \mathbb{T}$, we define the *cosine operator*

$$C_U^\psi = \frac{1}{2}(\psi U + \psi^{-1} U^*). \quad (4.22)$$

Then it is clear that C_U^ψ is self adjoint. Compare this to taking the real part of a complex number, $\Re(z) = 1/2(z + \bar{z})$. Recalling the commutation relation $ZX = \omega XZ$, where $\omega = e^{\frac{2\pi i}{d}}$, we note the following relations for when U is a boost or shift operator:

$$\begin{aligned} ZC_X^\psi Z^* &= \frac{1}{2}(\psi ZXZ^* + \psi^{-1} ZX^*Z^*) = \frac{1}{2}(\psi\omega X + \psi^{-1}\omega^{-1}X^*) = C_X^{\omega\psi} \\ XC_Z^\psi X^* &= \frac{1}{2}(\psi XZX^* + \psi^{-1} XZ^*X^*) = \frac{1}{2}(\psi\omega^{-1}Z + \psi^{-1}\omega Z^*) = C_Z^{\omega^{-1}\psi} \end{aligned}$$

Also, if U_S is the metaplectic associated to some symplectic S so that $U_S w(z) U_S^* = w(Sz)$, then

$$U_S C_{w(z)}^\psi U_S = \frac{1}{2} U_S (\psi w(z) + \psi^{-1} w(-z)) U_S^* = \frac{1}{2} (\psi w(Sz) + \psi^{-1} w(-Sz)) = C_{w(Sz)}^\psi \quad (4.23)$$

So we can use Weyl operators to modulate the phase, and metaplectics to change the operator. This motivates the following definition:

Definition 4.4.1. We say that a state ρ on $L^2(\mathbb{Z}_d)$ is *Gibbs Gaussian* or simply *Gaussian* if it is of the form

$$\rho = \frac{\exp(-\beta H)}{\text{Tr} \exp(-\beta H)}, \quad (4.24)$$

for some positive parameter β and a Hamiltonian H of the form

$$H = \cos \theta C_{w(z_1)}^{\psi_1} + \sin \theta C_{w(z_2)}^{\psi_2}, \quad (4.25)$$

where z_1, z_2 are points in the phase space V such that $[z_1, z_2] = 1$, and ψ_1, ψ_2 are d th roots of unity.

The requirement that $[z_1, z_2] = 1$ will become clear shortly. We now describe how one can use Weyl operators and metaplectics to transform 4.25 into 4.21

If $[z_1, z_2] = 1$ then the matrix with columns z_1 and z_2 is symplectic. Set S to be the inverse of this matrix. Then

$$Sz_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ and } Sz_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (4.26)$$

Thus, if U_S is the corresponding metaplectic, we have

$$U_S H U_S^* = \cos \theta C_Z^{\psi_1} + \sin \theta C_X^{\psi_2} \quad (4.27)$$

Now, there exists k and l such that

$$\psi_1 = \exp\left(\frac{2\pi i}{d}k\right) \text{ and } \psi_2 = \exp\left(\frac{2\pi i}{d}l\right), \quad (4.28)$$

so if we conjugate by $w(l, -k)$ we get

$$w(l, -k)U_S H U_S^* w(l, -k)^* = \cos \theta C_Z + \sin \theta C_X = H_\theta \quad (4.29)$$

where the lack of superscript indicates that the phase is equal to 1. Thus, we have shown that our general Hamiltonian indeed reduces to the squeezed Harper operator after conjugation by Clifford unitaries, like in the continuous case. This reduction is part of what motivates us to choose this framework as our de facto Gaussian state.

Let us now examine the form of the Gibbs state with a Hamiltonian $H = U_S^* W(z)^* H_\theta W(z) U_S$. Exponentiating, we have

$$\exp(-\beta H) = \exp(-\beta U_S^* W(z)^* H_\theta W(z) U_S) = U_S^* W(z)^* \exp(-\beta H_\theta) W(z) U_S. \quad (4.30)$$

Recalling that we can write

$$H_\theta = \sum_{j=0}^{d-1} h_j |j\rangle \langle j|_\theta, \quad (4.31)$$

we can expand the exponential as

$$\exp(-\beta H_\theta) = \sum_{j=0}^{d-1} e^{-\beta h_j} |j\rangle \langle j|_\theta \quad (4.32)$$

By linearity of the trace, we have

$$\text{Tr} \exp(-\beta H) = \text{Tr} \left(\sum_{j=0}^{d-1} e^{-\beta h_j} |j\rangle \langle j|_\theta \right) = \sum_{j=0}^{d-1} e^{-\beta h_j}. \quad (4.33)$$

Combining these all, we have the most general form of a Gibbs state:

$$\rho = \left(\sum_{j=0}^{d-1} e^{-\beta h_j} \right)^{-1} U_S^* W(z)^* \left(\sum_{j=0}^{d-1} e^{-\beta h_j} |j\rangle\langle j|_\theta \right) W(z) U_S. \quad (4.34)$$

If one had an analytic expression for the state $|j\rangle\langle j|$, they could compute $U_S^* W(z)^* |j\rangle\langle j| W(z) U_S$, which would simplify this expression immensely.

At this time we are unaware of any transformation that will affect θ ; it seems like this is the natural analogue for the frequencies ω_j in the continuous case.

We have limited evidence to suggest that H given in 4.25 is in fact the most general form of Hamiltonian that one could consider, but it seems a good place to start. One way we could convince ourselves would be to figure out in what way H could be viewed as *quadratic*, as in the continuous case. If we could phrase $H = \cos \theta C_Z + \sin \theta C_X$ through some quadratic form, or find auxiliary terms to add that would make it quadratic, then that would make this definition more convincing. One approach could be through second-order difference equations; another could be through the use of a *sine operator*, given by $S_U = 1/2i(U - U^*)$. For more on sine and cosine operators see [27].

Just like in the continuous case, it makes sense to consider the limiting case as $\beta \rightarrow \infty$, which corresponds to temperature approaching absolute zero. In this case, it can again be shown that the Gibbs state of the squeezed Harper operator H_θ from equation 4.21 converges to the squeezed vacuum state $|0\rangle\langle 0|_\theta$: indeed, since

$$H_\theta = \sum_{j=0}^{d-1} h_j |j\rangle\langle j|_\theta \quad (4.35)$$

we have

$$\rho = \frac{\exp(-\beta H_\theta)}{\text{Tr} \exp(-\beta H_\theta)} = \sum_{j=0}^{d-1} \frac{\exp(-\beta h_j)}{\sum_{k=0}^{d-1} \exp(-\beta h_k)} |j\rangle\langle j|_\theta. \quad (4.36)$$

Let us consider the inverse of the coefficients, starting with $j = 0$:

$$\left(\frac{\exp(-\beta h_0)}{\sum_{k=0}^{d-1} \exp(-\beta h_k)} \right)^{-1} = 1 + \sum_{k \neq j} \exp(-\beta(h_k - h_0)) \xrightarrow{\beta \rightarrow \infty} 1, \quad (4.37)$$

since $h_k - h_0 > 0$ for all k . If $j > 0$, we have

$$\left(\frac{\exp(-\beta h_0)}{\sum_{k=0}^{d-1} \exp(-\beta h_k)} \right)^{-1} = \sum_{l=0}^{j-1} \exp(-\beta(h_l - h_j)) + 1 + \sum_{k=j+1}^{d-1} \exp(-\beta(h_k - h_j)) \xrightarrow{\beta \rightarrow \infty} \infty, \quad (4.38)$$

since $l < j$ implies $h_l - h_j < 0$, so that the first sum diverges to infinity. Thus, the only term in the sum 4.36 that survives is the one with $j = 0$, so indeed $\rho \rightarrow |0\rangle\langle 0|_\theta$ as β approaches infinity.

We note that the minimal uncertainty state $|\Gamma\rangle$ from 4.2.1 is the same as the zero state $|0\rangle$; thus, through our Gibbs formalism we recover the minimal uncertainty states by taking $\beta \rightarrow \infty$ as in the continuous case.

We now proceed with the *Mathematica* implementation of these ideas. Initialize a system of dimension 7:

```
In[46]:= qInit[1,{7}]
```

```
Out[46]= {q1}
```

We pick two points in phase space with symplectic product equal to 1:

```
In[47]:= z1 = {-3,5}; z2 = {-2,3};
          symp[z1,z2]
```

```
Out[47]= 1
```

We verify that they can be sent to the basis vectors using the symplectic T :

```
In[48]:= T = Inverse[Transpose[{z1,z2}]]
          T.z1
          T.z2
          psi1 =  $\chi$ [1, 3];
          psi2 =  $\chi$ [1, 5];
```

```
Out[48]= {{3,2},{-5,-3}}
```

```
Out[49]= {1,0}
```

```
Out[50]= {0,1}
```

Construct our Hamiltonian:

```
In[51]:= ham = 1/2(cosine[WSparse[z1],psi1] + cosine[WSparse[z2],psi2])
```

Construct the metaplectic corresponding to T :

```
In[52]:= metaT1 = metaplecticU[T];
```

How does metaplectic conjugation affect our Hamiltonian?

```
In[53]:= metaT1 ** ham ** hc[metaT1] ==
          1/2(cosine[WSparse[T.z1]] + cosine[WSparse[T.z2]])
```

```
Out[53]= True
```

Finally, check that we can reduce to the Harper operator:

```
In[54]:= WSparse[{5, -3}] ** metaT1 ** ham ** hc[metaT1]
          ** hc[WSparse[{5, -3}]]
          == 1/2 (cosine[X[1, 1]] + cosine[Z[1, 1]])
```

```
Out[54]= True
```

Let us do some visualization for a sample Gibbs state: initialize a 19-dimensional system and generate some points z_1, z_2 .

```
In[55]:= qInit[1, {19}]
```

```
Out[55]:= {q1}
```

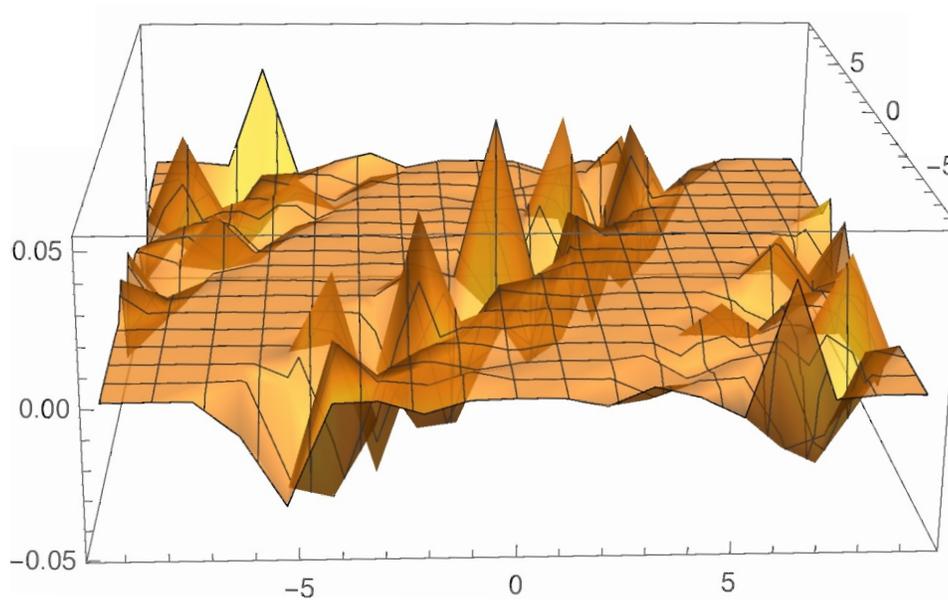
```
In[56]:= z1 = {2, 1}; z2 = {5, 3};
symp[z1, z2]
```

```
Out[56]:= 1
```

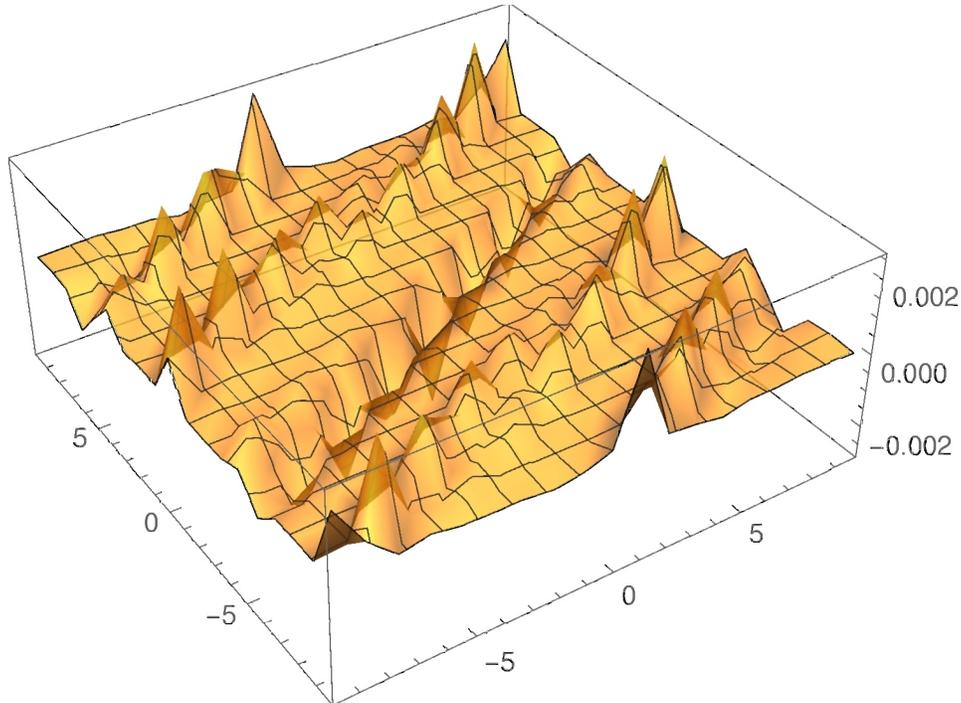
Here is the characteristic function and Wigner function of this Gibbs state:

```
In[57]:= ham = 1/2 (cosine[WSparse[z1]] + cosine[WSparse[z2]]) // Chop;
charfcnPlot[gibbs[ham, 75] // Chop, False]
charfcnPlot[wigner[gibbs[ham, 75]] // Chop]
```

```
Out[57]=
```



Out[58]=



Note that the Wigner function takes negative values.

4.5 Comparisons

We now have three definitions for Gaussian states in discrete phase space. How do these three definitions relate? At the very least, we know that there is no combination of Weyl and metaplectic operators that will transform a Hudson state into a Gibbs state, which we show presently. At this time, we are not sure how Gibbs states and Bernstein states relate, however numerical evidence suggests they are distinct.

4.5.1 Hudson States are not in the Gibbs Framework

In this section, we prove that Hudson states can not be obtained through a Clifford operation acting on a Gibbs state, at least for some dimensions d . This fact is in parallel with their corresponding notions from probability theory; despite their equivalence in the continuous case, they are distinct in finite dimensions.

Suppose, in a single particle system of dimension d , we are given a Hudson state with variance θ and mean x , which is given by

$$\psi_{\theta,x}(q) = \frac{\lambda}{\sqrt{d}} \exp\left(\frac{2\pi i}{d}(xq + q^2\theta)\right). \quad (4.39)$$

for some $|\lambda| = 1$. We can set both the variance and mean to zero using a chirp and a boost:

$$R_{-2\theta}V_{-x}\psi_{\theta,x}(q) = \frac{\lambda}{\sqrt{d}}(1, 1, \dots, 1). \quad (4.40)$$

Applying a Fourier transform, we get

$$\frac{\lambda}{\sqrt{d}}F(1, \dots, 1) = \lambda(1, 0, \dots, 0) = \lambda\delta_0 \quad (4.41)$$

where we will use δ_i to denote the i th basis vector. So we can use Clifford operations to transform any Hudson state into a standard basis vector. Since any Gibbs state can be acquired from the vacuum state $|\Gamma\rangle$, if we can show that there is no way to transform δ_0 into $|\Gamma\rangle$ by means of Weyl operators and metaplectics, we will have shown that these classes are distinct, up to Clifford unitaries.

Suppose there was a Clifford unitary U that transformed δ_0 into $|\Gamma\rangle$. By inspection it obviously can not be Weyl (these operations only translate or modulate by complex numbers, and $|\Gamma\rangle$ is real), so there must be some symplectic T such that $U = U_T$. Then

$$\delta_0 = U_T|\Gamma\rangle \quad (4.42)$$

Using only Weyl operators and δ_0 , we can construct an orthonormal basis for \mathbb{C}^d : indeed, $\{\delta_s = w(0, s)\delta_0 \mid s \in \mathbb{Z}_d\}$ is such a basis. Then by assumption, and the defining relation of the metaplectic $U_T w(z)U_T^* = w(Tz)$,

$$\delta_s = w(0, s)\delta_0 = w(0, s)U_T|\Gamma\rangle = U_T w(T(0, s))|\Gamma\rangle. \quad (4.43)$$

Since $\{\delta_s\}_{s \in \mathbb{Z}_d}$ is an orthonormal basis and U_T is unitary, $\{w(T(0, s))|\Gamma\rangle\}_{s \in \mathbb{Z}_d}$ should also be an orthonormal basis. So our problem reduces to the existence of such a T .

Finding this symplectic T would be very difficult, but fortunately thanks to the power of computers we can suggest a brute-force alternative: since there are only d^2 Weyl operators on \mathbb{C}^d , we can simply check them all! We proceed as follows: list the elements $z_j \in V$ in some way, and construct the *Gram matrix* $M = (M_{i,j})_{i,j}$ given by

$$M_{i,j} = \langle \Gamma | W(z_i)^* W(z_j) | \Gamma \rangle.$$

If there is some subset U of V such that $\{W(z)|\Gamma\rangle\}_{z \in U}$ is an orthonormal basis of \mathbb{C}^d , then this will correspond to rows and columns of zeros in the Gram matrix M . So we must construct this matrix, and check if there are enough zeros in the correct places, and we will see if there is such a basis. Existence of a basis will not be sufficient, but it is certainly necessary, by our discussion above.

We now turn to Mathematica to test this. Initialize a single system of dimension 7, and define a Hudson state:

```
In[59]:= qInit[1, {7}]
Out[59]:= {q1}
In[60]:= gauss = gaussianKet[1, theta=2, x=4]
Out[60]:= 
$$\begin{pmatrix} 0.377964 \\ 0.235657 - 0.295505 i \\ -0.084105 + 0.368488 i \\ -0.084105 + 0.368488 i \\ 0.235657 - 0.295505 i \\ 0.377964 \\ -0.084105 - 0.368488 i \end{pmatrix}$$

{ket[q1]}
```

We verify our decomposition into the first standard basis ket;

```
In[61]:= vec=FourierU[]**ChirpU[{-2theta}]**WSparse[-x,0]**gauss
```

```
Out[61]= 
$$\begin{pmatrix} 1. \\ 0. \\ 0. \\ 0. \\ 0. \\ 0. \\ 0. \end{pmatrix}$$

{ket[q1]}
```

The following helper functions allow us to compute the Gram matrix; since we need to use same Weyl operators several times, we first generate them as a list Wmat.

```
In[62]:= vecToMat[y_]:= {Mod[Floor[y/nlist[[1]]],nlist[[1]]],
  Mod[y,nlist[[1]]]}
(Wmat[#]=WSparse[vecToMat[#]])&/@
  Range[0,nlist[[1]]^2-1];
helper[x_,y_]:= (hc[minUncerVec[]]**hc[Wmat[x]**Wmat[y]
  **minUncerVec[]][[1]]
```

The following line takes several minutes to run;

```
In[63]:= Table[helper[i,j],{i,0,nlist[[1]]^2-1},{j,0,nlist[[1]]^2-1}]
// MatrixForm // Chop
```

This prints out a 49×49 complex matrix; we show just the first 10×4 block;

```
Out[63]= 
$$\begin{pmatrix} 1. & 0.800826 & 0.438339 & 0.215539 \\ 0.800826 & 1. & 0.800826 & 0.438339 \\ 0.438339 & 0.800826 & 1. & 0.800826 \\ 0.215539 & 0.438339 & 0.800826 & 1. \\ 0.215539 & 0.215539 & 0.438339 & 0.800826 \\ 0.438339 & 0.215539 & 0.215539 & 0.438339 \\ 0.800826 & 0.438339 & 0.215539 & 0.215539 \\ 0.800826 & 0.563475 - 0.271355 i & 0.193883 - 0.243122 i & 0.0182799 - 0.0800896 i \\ -0.62541 & -0.721519 + 0.347465 i & -0.389937 + 0.488965 i & -0.0691961 + 0.303168 i \\ 0.310965 & 0.563475 - 0.271355 i & 0.499307 - 0.626111 i & 0.139167 - 0.609729 i \end{pmatrix}$$

```

There appears to be no zeros, and in fact, the following command prints the indices of all the zeros in the matrix;

```
In[64]:= Position[%,0]
```

```
Out[64]= {}
```

So in dimension 7, it seems that Hudson states and Gibbs states are distinct, up to Clifford operations. Let us try again in dimension 9;

```

In[65]:= qInit[1, {9}]
          (Wmat[#]=WSparse[vecToMat[#]])&/@Range[0,nlist[[1]]^2 -1];

Out[65]= {q1}

In[66]:= Position[Table[helper[i, j], {i, 0, nlist[[1]]^2-1},
                    {j, 0, nlist[[1]]^2-1}] // Chop, 0]

Out[66]= {}

```

The two sets of states are also distinct in dimension 9.

There is no reason we must restrict to odd dimensions, and there are cases in the literature where small dimensions, such as dimension 6, behave differently than odd or prime dimensions; see [9]. Let us try in dimension 6;

```

In[67]:= qInit[1, {6}]

Out[67]= {q1}

```

We do the same computation as above,

```

In[68]:= (Wmat[#]=WSparse[vecToMat[#]])&/@Range[0,nlist[[1]]^2 -1];
          helper[x_, y_]
            :=(hc[minUncerVec[]]**hc[Wmat[x]**Wmat[y]**minUncerVec[]])[[1]]
          vertexList = Position[Table[helper[i, j], {i, 0, nlist[[1]]^2-1},
                                    {j, 0, nlist[[1]]^2-1}] //Chop, 0]

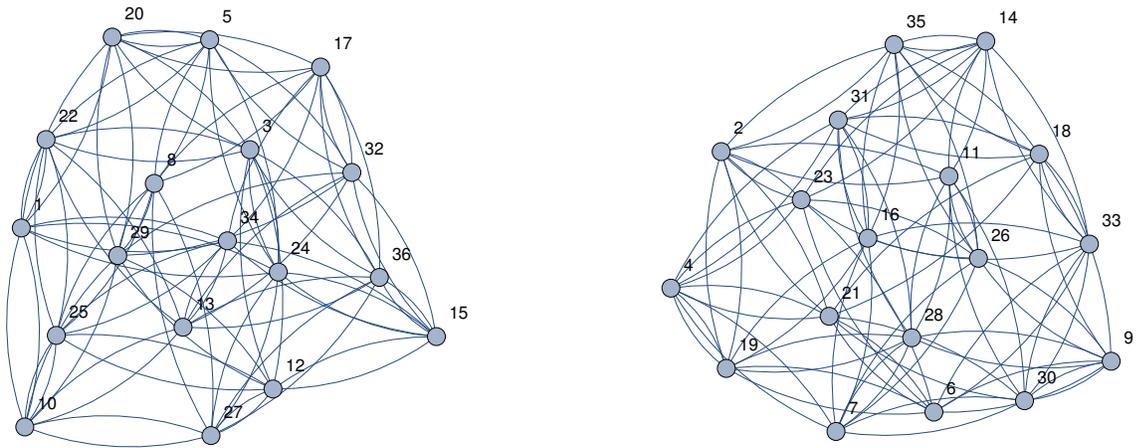
Out[68]= {{1, 10}, {1, 20}, {1, 22}, {1, 24}, {1, 34}, {2, 11}, {2, 19}, {2, 21}, {2, 23},
          {2, 35}, {3, 12}, {3, 20}, {3, 22}, {3, 24}, {3, 36}, {4, 7}, {4, 19}, {4, 21}, {4, 23},
          {4, 31}, {5, 8}, {5, 20}, {5, 22}, {5, 24}, {5, 32}, {6, 9}, {6, 19}, {6, 21}, {6, 23},
          {6, 33}, {7, 4}, {7, 16}, {7, 26}, {7, 28}, {7, 30}, {8, 5}, {8, 17}, {8, 25}, {8, 27},
          {8, 29}, {9, 6}, {9, 18}, {9, 26}, {9, 28}, {9, 30}, {10, 1}, {10, 13}, {10, 25}, {10, 27},
          {10, 29}, {11, 2}, {11, 14}, {11, 26}, {11, 28}, {11, 30}, {12, 3}, {12, 15}, {12, 25},
          {12, 27}, {12, 29}, {13, 10}, {13, 22}, {13, 32}, {13, 34}, {13, 36}, {14, 11}, {14, 23},
          {14, 31}, {14, 33}, {14, 35}, {15, 12}, {15, 24}, {15, 32}, {15, 34}, {15, 36}, {16, 7},
          {16, 19}, {16, 31}, {16, 33}, {16, 35}, {17, 8}, {17, 20}, {17, 32}, {17, 34}, {17, 36},
          {18, 9}, {18, 21}, {18, 31}, {18, 33}, {18, 35}, {19, 2}, {19, 4}, {19, 6}, {19, 16},
          {19, 28}, {20, 1}, {20, 3}, {20, 5}, {20, 17}, {20, 29}, {21, 2}, {21, 4}, {21, 6},
          {21, 18}, {21, 30}, {22, 1}, {22, 3}, {22, 5}, {22, 13}, {22, 25}, {23, 2}, {23, 4},
          {23, 6}, {23, 14}, {23, 26}, {24, 1}, {24, 3}, {24, 5}, {24, 15}, {24, 27}, {25, 8},
          {25, 10}, {25, 12}, {25, 22}, {25, 34}, {26, 7}, {26, 9}, {26, 11}, {26, 23}, {26, 35},
          {27, 8}, {27, 10}, {27, 12}, {27, 24}, {27, 36}, {28, 7}, {28, 9}, {28, 11}, {28, 19},
          {28, 31}, {29, 8}, {29, 10}, {29, 12}, {29, 20}, {29, 32}, {30, 7}, {30, 9}, {30, 11},
          {30, 21}, {30, 33}, {31, 4}, {31, 14}, {31, 16}, {31, 18}, {31, 28}, {32, 5}, {32, 13},
          {32, 15}, {32, 17}, {32, 29}, {33, 6}, {33, 14}, {33, 16}, {33, 18}, {33, 30}, {34, 1},
          {34, 13}, {34, 15}, {34, 17}, {34, 25}, {35, 2}, {35, 14}, {35, 16}, {35, 18}, {35, 26},
          {36, 3}, {36, 13}, {36, 15}, {36, 17}, {36, 27}}

```

and find that there are many indices with 0's! We need to find out if there is an orthonormal basis among these vectors. To do this, we appeal to graph theory; a *graph* $G = (V, E)$ is a set of *vertices* V and a set of *edges* E between them. Let us construct a graph G with vertices $V = \{1, \dots, 36\}$ corresponding to the 36 vectors $\{v_j\}_{j=1}^{36}$, with an edge (i, j) if and only if $\langle v_i | v_j \rangle = 0$. Mathematica has a suite of functions devoted to graph theory, such as `Graph`;

```
In[69]:= makeEdge[{x_, y_}] := UndirectedEdge[x, y]
G = Graph[Range[1, 36], makeEdge[#] & /@ %, VertexLabels -> "Name"]
```

```
Out[69]=
```



If there is a subset $V' \subset V$ such that $\{v_j\}_{j \in V'}$ is an orthonormal basis, then $E' = \{(i, j) \mid i, j \in V'\}$ should be a subset of E . That is, there should be an edge between every pair of elements in V' ; this is called a *clique*. So finding an orthonormal basis is equivalent to finding a clique with d elements. The Mathematica function `FindClique` will find all cliques in G . Any pair of vertices with an edge between them forms a clique with two elements. We claim that these are the only cliques that exist in G . Indeed,

```
In[70]:= Sort[FindClique[G, 2, All]] ==
Sort[Select[vertexList, #[[1]] < #[[2]] &]]
```

```
Out[70]= True
```

In words, the list of all cliques is the same as the list of all vertices, which proves that there are only two-element cliques in G . This proves that there is no orthonormal basis among the vectors in dimension 6.

What we have shown above is by no means a proof that Gibbs states are always distinct from Hudson states in the discrete case, but it does provide strong evidence. This process is not optimized to run in high dimensions, but by computing small chunks of the Gram matrix we can see further evidence that this is true in general. At the very least, we can state the following theorem:

Theorem 4.5.1. *There exists a number $d \in \mathbb{N}$ such that the set of Gibbs states (defined in Definition 4.4.1) is distinct from the set of Hudson states (defined in Lemma 4.1.2) in dimension d .*

Chapter 5

Convergence between Discrete and Continuous Quantum Systems

It is common practice in mathematics and physics to approximate continuous objects or systems by discrete ones. In the context of quantum systems, a particularly relevant approximation scheme has been developed by Barker in a series of papers [6, 1, 2, 3, 4, 5]. In this chapter we establish a new equivalent formulation of this approximation scheme.

5.1 Inductive Resolutions

We begin with an outline of Barker's construction.

Let $L^2(\mathbb{R})$ be the space of square-integrable functions on the real line. Let \mathcal{N} be an infinite subset of the natural numbers \mathbb{N} , and for each $d \in \mathcal{N}$, let $L^2(\mathbb{Z}_d)$ be a d -dimensional Hilbert space. Consider vectors $\psi \in L^2(\mathbb{R})$ and $\psi_d \in L^2(\mathbb{Z}_d)$ for each $d \in \mathcal{N}$.

The approximation involves rescaling and sampling the function ψ so that it becomes a state vector in $L^2(\mathbb{Z}_d)$. To do this, we use the scaling factor

$$v(d) = \sqrt[4]{\frac{d}{2\pi}}.$$

Recall that the Schwartz space $\mathcal{S}(\mathbb{R})$ is dense in $L^2(\mathbb{R})$. For each $d \in \mathcal{N}$, define the *restriction map* $\text{res}_d : \mathcal{S}(\mathbb{R}) \rightarrow L^2(\mathbb{Z}_d)$ by

$$\text{res}_d(\psi)(x) = v(d)^{-1}\psi(v(d)^{-2}x), \quad x \in \mathbb{Z}_d. \tag{5.1}$$

It is easy to see that res_d is linear.

Given $d \in \mathcal{N}$, the restriction map samples ψ at the points $v(d)^{-2}x$ for $x \in \mathbb{Z}_d$. We note that we use *minimal* representatives for elements of \mathbb{Z}_d : that is, $\mathbb{Z}_d = \{[d/2], \dots, [d/2] - 1\}$. This way, we embed \mathbb{Z}_d into the real line symmetrically about zero. As d gets large, the set of sample points $v(d)^{-2}x$ sweeps a larger range and the points get closer together: indeed, the distance between two points is $v(d)^{-2} = \sqrt{2\pi/d}$. Defining

$$\mathbb{R}(\mathcal{N}) = \{v(d)^{-2}x | d \in \mathcal{N}, x \in \mathbb{Z}_d\} = \bigcup_{d \in \mathcal{N}} \{v(d)^{-2}x | x \in \mathbb{Z}_d\},$$

we see that $\mathbb{R}(\mathcal{N})$ is a dense set in \mathbb{R} .

Given a vector $\psi \in L^2(\mathbb{R})$ and a sequence of vectors (ψ_d) where $\psi_d \in L^2(\mathbb{Z}_d)$ for each $d \in \mathcal{N}$, we say that ψ_d induces ψ and write $\psi = \lim_{d \in \mathcal{N}} \psi_d$ if, for all $\phi \in \mathcal{S}(\mathbb{R})$,

$$\langle \phi | \psi \rangle = \lim_{d \in \mathcal{N}} \langle \text{res}_d(\phi) | \psi_d \rangle.$$

By the denseness of $\mathcal{S}(\mathbb{R})$ in $L^2(\mathbb{R})$, the sequence (ψ_d) can converge to at most one vector ψ .

We have our first example of induction: the restrictions of Schwartz functions converge to themselves under the inner product.

Lemma 5.1.1. *For any $\psi, \varphi \in \mathcal{S}(\mathbb{R})$, we have*

$$\langle \varphi | \psi \rangle = \lim_{d \rightarrow \infty} \langle \text{res}_d(\varphi) | \text{res}_d(\psi) \rangle.$$

Proof. Without loss of generality, we may assume $\varphi, \psi \geq 0$. Set $\psi_n = \text{res}_d(\psi)$, $\phi_n = \text{res}_d(\phi)$. Since Schwartz functions are Riemann integrable, taking the natural partitions of $[-q, q]$ through our embedding of \mathbb{Z}_d into \mathbb{R} , it follows that

$$\begin{aligned} \int_{-q}^q \varphi(t)^* \psi(t) dt &= \lim_{d \in \mathcal{N}} \sum_{\substack{|x| \leq qv(d)^2 \\ x \in \mathbb{Z}_d}} v(d)^{-2} \varphi(v(d)^{-2}x)^* \psi(v(d)^{-2}x) \\ &= \lim_{d \in \mathcal{N}} \sum_{\substack{|x| \leq qv(d)^2 \\ x \in \mathbb{Z}_d}} \varphi_d(x)^* \psi_d(x), \end{aligned}$$

for any $q > 0$. Also, by Cauchy-Schwartz,

$$\left(\lim_{d \in \mathcal{N}} \sum_{\substack{|x| > qv(d)^2 \\ x \in \mathbb{Z}_d}} \varphi_d(x)^* \psi_d(x) \right)^2 = \left(\int_{\{|t| > q\}} \varphi(t)^* \psi(t) dt \right)^2 \leq \int_{\{|t| > q\}} |\varphi(t)|^2 dt \int_{\{|t| > q\}} |\psi(t)|^2 dt. \quad (5.2)$$

The right hand side goes to zero as q goes to infinity, so we have our result. \square

Using the previous result, it immediately follows that $\psi = \lim_{d \in \mathcal{N}} \text{res}_d(\psi)$, or, the restrictions of a vector induce itself.

We quote the following result of Barker [1, Theorem 2.5], which applies the ideas above to show that the Harper functions induce the Hermite-Gaussians:

Theorem 5.1.2. *Let \mathcal{N} be an infinite set of natural numbers with the condition that, if $d, d' \in \mathcal{N}$ with $d \leq d'$, then d divides d' and the integer d'/d is a square. Let ψ_n denote the n th Hermite-Gaussian and $|n\rangle_d$ be the n th Harper function in $L^2(\mathbb{Z}_d)$. Then for each $j \in \mathbb{N}$, $\psi_n = \lim_{d \in \mathcal{N}} |n\rangle_d$.*

That is, the j th Harper functions induce the j th Hermite-Gaussian. We illustrate this theorem in Figure 5.1.

We now reformulate Barker's induction into a form which is easier to work with, and which further justifies the terminology. For each $d \in \mathcal{N}$, define the map $i_d : L^2(\mathbb{Z}_d) \rightarrow L^2(\mathbb{R})$ by

$$i_d(\psi_d)(t) = v(d) \sum_{x \in \mathbb{Z}_d} \psi_d(x) \chi_{B_d(x)}(t),$$

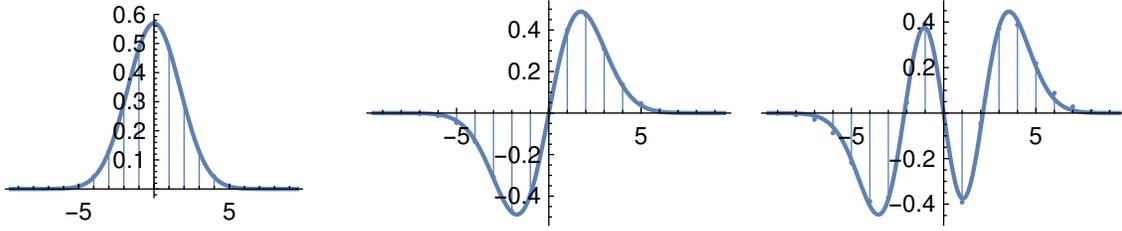


Figure 5.1: Comparison of the Harper functions to the Hermite-Gaussians. Here, the dimension of the discrete system is 19, and we show comparisons of the zeroth, first and third functions.

where $B_d(x) = B_{v(d)^{-2}/2}(xv(d)^{-2}) = \left(v(d)^{-2}(x - \frac{1}{2}), v(d)^{-2}(x + \frac{1}{2})\right)$. In this way, we *induce* a step function out of a discrete function ψ_d . Now we will say that ψ_d *converges to* ψ if the sequence $(i_d(\psi_d))$ converges to ψ in $L^2(\mathbb{R})$, that is, if $\|\psi - i_d(\psi_d)\|_2 \rightarrow 0$.

A few things to note; first, i_d is an isometry: indeed,

$$\|i_d(\psi_d)\|_{L^2(\mathbb{R})}^2 = v(d)^2 \sum_x |\varphi_d(x)|^2 \|\chi_{B_d(x)}\|_{L^2(\mathbb{R})}^2 = \|\varphi_d\|_{L^2(\mathbb{Z}_d)}^2,$$

since each $B_d(x)$ has measure $v(d)^{-2}$. Second, i_d is a right inverse to a natural extension of Barker's restriction map res_d to functions that are continuous on a neighborhood of $xv(d)^{-2}$ for $x \in \mathbb{Z}_d$. Indeed,

$$\text{res}_d(i_d(\psi_d))(x) = v(d)^{-1}(i_d(\psi_d))(xv(d)^{-2}) = \sum_{y \in \mathbb{Z}_d} \psi_d(y) \chi_{B_d(y)}(xv(d)^{-2}) = \psi_d(x),$$

since $xv(d)^{-2} \in B_d(x)$. Third, i_d is the natural way to ascribe a step function to the partitions defined by the scaling parameter $v(d)$. The following lemma will be useful in the proof of the main result in this section:

Lemma 5.1.3. *If $\varphi \in \mathcal{S}(\mathbb{R})$, define*

$$\varphi_d(t) := i_d(\text{res}_d(\varphi))(t) = \sum_{x \in \mathbb{Z}_n} \varphi(v(d)^{-2}x) \chi_{B_d(x)}(t) \in L^2(\mathbb{R}).$$

Then $\|\varphi - \varphi_d\|_2 \rightarrow 0$.

Proof. Given $\varepsilon > 0$, pick q such that

$$\int_{|t|>q} |\varphi(t)|^2 dt < \frac{\varepsilon^2}{64}.$$

Since φ is Schwartz, it is uniformly continuous, so there exists a $\delta > 0$ such that

$$|s - t| < \delta \text{ implies } |\varphi(s) - \varphi(t)| < \frac{\varepsilon}{2\sqrt{q}}.$$

Denoting by $l(I)$ the length of the interval I , we have

$$\|\varphi_d \chi_{(q,\infty)}\|_2^2 = \sum_{x \in \mathbb{Z}_d} |\varphi(xv(d)^{-2})|^2 l((q, \infty) \cap B_d(x)) \xrightarrow{n \rightarrow \infty} \int_q^\infty |\varphi(t)|^2 dt,$$

and similarly for $\|\varphi_d \chi_{(-\infty, -q)}\|_2^2$. With these two, we may find N large enough that for any $d \geq N$,

$$|\|\varphi_d \chi_{(q,\infty)}\|_2 - \|\varphi \chi_{(q,\infty)}\|_2| < \frac{\varepsilon}{4}, \quad \text{and} \quad |\|\varphi_d \chi_{(-\infty, -q)}\|_2 - \|\varphi \chi_{(-\infty, -q)}\|_2| < \frac{\varepsilon}{4}. \quad (5.3)$$

Without loss of generality we may assume this N is large enough that $v(d)^{-2} = \sqrt{2\pi/d} < \delta$ and $\sqrt{d\pi/2} > q$, for all $d \geq N$. Then for $d \geq N$,

$$\begin{aligned} \|(\varphi - \varphi_d) \chi_{(q,\infty)}\|_2 &\leq \|\varphi \chi_{(q,\infty)}\|_2 + \|\varphi_d \chi_{(q,\infty)}\|_2 \\ &< \frac{\varepsilon}{8} + \frac{\varepsilon}{4} + \frac{\varepsilon}{8} = \frac{\varepsilon}{2}. \end{aligned}$$

Similarly, $\|(\varphi - \varphi_d) \chi_{(-\infty, -q)}\|_2 < \varepsilon/2$. Also,

$$\begin{aligned} \int_{-q}^q |\varphi(t) - \varphi_d(t)|^2 dt &= \sum_{x \in \mathbb{Z}_d} \int_{(-q,q) \cap B_d(x)} |\varphi(t) - \varphi(xv(d)^{-2})|^2 dt \\ &< \sum_{x \in \mathbb{Z}_d} \int_{(-q,q) \cap B_d(x)} \frac{\varepsilon^2}{4q} dt = \frac{\varepsilon^2}{2}. \end{aligned}$$

Thus, for $d \geq N$,

$$\begin{aligned} \|\varphi - \varphi_d\|_2^2 &= \|(\varphi - \varphi_d) \chi_{(-\infty, -q)}\|_2^2 + \int_{-q}^q |\varphi(t) - \varphi_d(t)|^2 dt + \|(\varphi - \varphi_d) \chi_{(q,\infty)}\|_2^2 \\ &< \frac{\varepsilon^2}{4} + \frac{\varepsilon^2}{2} + \frac{\varepsilon^2}{4} = \varepsilon^2. \end{aligned}$$

□

Theorem 5.1.4. *Let $\psi \in L^2(\mathbb{R})$ and $\psi_d \in L^2(\mathbb{Z}_d)$ be states for all d in a subset $\mathcal{N} \subseteq \mathbb{N}$ satisfying the conditions in Theorem 5.1.2. Then*

$$(\psi_d) \text{ induces } \psi \text{ if and only if } \lim_{d \in \mathcal{N}} \|\psi - i_d(\psi_d)\|_2 = 0.$$

Proof. Suppose (ψ_d) induces ψ . Then for all $\varphi \in \mathcal{S}(\mathbb{R})$,

$$\langle \varphi | \psi \rangle = \lim_{d \in \mathcal{N}} \langle \text{res}_d(\varphi) | \psi_d \rangle = \lim_{d \in \mathcal{N}} \langle i_d(\text{res}_d(\varphi)) | i_d(\psi_d) \rangle = \lim_{d \in \mathcal{N}} \langle \varphi_d | i_d(\psi_d) \rangle.$$

By Lemma 5.1.3, we have

$$|\langle \varphi_d | i_d(\psi_d) \rangle - \langle \varphi | i_d(\psi_d) \rangle| = |\langle \varphi_d - \varphi | i_d(\psi_d) \rangle| \leq \|\varphi_d - \varphi\|_2 \|i_d(\psi_d)\|_2 = \|\varphi_d - \varphi\|_2,$$

which converges to zero, so that

$$\langle \varphi | \psi \rangle = \lim_{d \in \mathcal{N}} \langle \varphi | i_d(\psi_d) \rangle.$$

By the boundedness of $\|i_d(\psi_d)\|_2$, it follows that $i_d(\psi_d) \rightarrow \psi$ weakly. Indeed, for any $\varphi \in L^2(\mathbb{R})$ and $\varepsilon > 0$, by denseness there is a $\varphi_\varepsilon \in \mathcal{S}(\mathbb{R})$ such that $\|\varphi - \varphi_\varepsilon\|_2 < \varepsilon$. Then,

$$\begin{aligned} |\langle \varphi | \psi \rangle - \langle \varphi | i_d(\psi_d) \rangle| &\leq |\langle \varphi | \psi \rangle - \langle \varphi_\varepsilon | \psi \rangle| + |\langle \varphi_\varepsilon | \psi \rangle - \langle \varphi_\varepsilon | i_d(\psi_d) \rangle| + |\langle \varphi_\varepsilon | i_d(\psi_d) \rangle - \langle \varphi | i_d(\psi_d) \rangle| \\ &= |\langle \varphi - \varphi_\varepsilon | \psi \rangle| + |\langle \varphi_\varepsilon | \psi - i_d(\psi_d) \rangle| + |\langle \varphi - \varphi_\varepsilon | i_d(\psi_d) \rangle| \\ &\leq \|\varphi - \varphi_\varepsilon\|_2 \|\psi\|_2 + |\langle \varphi_\varepsilon | \psi - i_d(\psi_d) \rangle| + \|\varphi - \varphi_\varepsilon\|_2 \|i_d(\psi_d)\|_2 \\ &< \varepsilon. \end{aligned}$$

Finally, we have $\|i_d(\psi_d)\|_2 \rightarrow \|\psi\|_2 = 1$, so the convergence is in norm.

Conversely, suppose $i_d(\psi_d) \rightarrow \psi$ in $L^2(\mathbb{R})$. Then we also have weak convergence, so for any $\varphi \in \mathcal{S}(\mathbb{R})$ we have

$$\langle \varphi | \psi \rangle = \lim_{d \in \mathcal{N}} \langle \varphi | i_d(\psi_d) \rangle.$$

By Lemma 5.1.3,

$$|\langle \varphi_d | i_d(\psi_d) \rangle - \langle \varphi | i_d(\psi_d) \rangle| \leq \|\varphi_d - \varphi\|_2 \rightarrow 0,$$

for any $\varphi \in \mathcal{S}(\mathbb{R})$. Thus, $\lim_{d \in \mathcal{N}} \langle \varphi_d | i_d(\psi_d) \rangle$ exists and

$$\langle \varphi | \psi \rangle = \lim_{d \in \mathcal{N}} \langle \varphi_d | i_d(\psi_d) \rangle = \lim_{d \in \mathcal{N}} \langle i_d(\text{res}_d(\varphi)) | i_d(\psi_d) \rangle = \lim_{d \in \mathcal{N}} \langle \text{res}_d(\varphi) | \psi_d \rangle.$$

□

We immediately have the following:

Corollary 5.1.5. *The Harper functions converge to the Hermite-Gaussians.*

5.2 The Harper Operator

We have mentioned that the Harper operator is related to the quantum harmonic oscillator, and we used it to derive the Harper functions, which we have claimed are analogous to the number states. We have not provided any justification for these facts yet. In this section, we review the numerical argument of [21] to justify this choice of operator.

5.2.1 Harper Convergence

Hashimoto and others [21] give a numerical argument to justify the choice of the Harper operator $H = \frac{1}{4}(X + X^* + Z + Z^*)$ as the Hamiltonian of the discrete harmonic oscillator, in the sense that as the dimension d becomes large, an eigenstate of H approaches an eigenstate of the Hamiltonian of a continuous harmonic oscillator \hat{H} . Indeed, their argument shows that $|\Gamma\rangle$ can approach a coherent state with any variance σ^2 . We will review their argument here, providing extra details where necessary.

First, we recall what might be the most fundamental result in numerical analysis, Taylor's Theorem:

Theorem 5.2.1 (Taylor's Theorem). *If $f : \mathbb{R} \rightarrow \mathbb{R}$ is at least twice differentiable, then if ε is small,*

$$f(x + \varepsilon) = f(x) + \varepsilon f'(x) + \frac{1}{2}\varepsilon^2 f''(x) + O(\varepsilon^3).$$

This presentation of Taylor's theorem has been modified for our purposes and might be unfamiliar at first. The requirement that ε is small is required to replace the usual $O(\varepsilon^3)$ with $O(\varepsilon^2)$: this follows since if $\varepsilon < 1$, $\varepsilon^2 > \varepsilon^3$. We can use this version to immediately prove the following approximation of the second derivative:

$$\varepsilon^2 f''(x) = f(x + \varepsilon) + f(x - \varepsilon) - 2f(x) + O(\varepsilon^2) \quad (5.4)$$

Indeed, this follows from adding the expressions for $f(x - \varepsilon)$ and $f(x + \varepsilon)$ obtained from 5.2.1. We use δ^2 to denote the central second difference, so that if $g : \mathbb{Z} \rightarrow \mathbb{C}$, we have

$$\delta^2 g(a) = g(a - 1) - 2g(a) + g(a + 1).$$

In the theory of difference equations, this plays a role analogous to the second derivative. Notice the similar forms of this with the right hand side of 5.4.

Now, consider a system with dimension d and an eigenstate $\phi : \mathbb{Z}_d \rightarrow \mathbb{C}$ of H with eigenvalue λ . Then $H\phi(a) = \lambda\phi(a)$ for $a \in \mathbb{Z}_d$. Notice that

$$\begin{aligned} H\phi(a) &= \frac{1}{4}(X + X^* + Z + Z^*)\phi(a) \\ &= \frac{1}{4}(\phi(a + 1) + \phi(a - 1) + (e^{\frac{2\pi ia}{d}} + e^{-\frac{2\pi ia}{d}})\phi(a)) \\ &= \frac{1}{4}(\phi(a + 1) + \phi(a - 1) + 2\cos\left(\frac{2\pi a}{d}\right)\phi(a)) = \lambda\phi(a) \end{aligned} \quad (5.5)$$

We show that, as $d \rightarrow \infty$, $|\Gamma\rangle$ approaches the centered coherent state with variance σ^2 (for any choice of σ^2). Indeed, for any d , find ε such that

$$v(d) = \sqrt{\frac{\sigma}{\varepsilon}} = \sqrt[4]{\frac{d}{2\pi}},$$

so that $\varepsilon \rightarrow 0$ as $d \rightarrow \infty$. Recalling our choice of minimal modular representatives for \mathbb{Z}_d , we define a position variable $x = a\varepsilon$, so that as d gets large, x acts as a continuous variable in \mathbb{R} . We note that $O(\varepsilon^2) = O(\frac{1}{d})$.

Using the identity $\cos(2\theta) = 1 - 2\sin^2(\theta)$, rearranging, and dividing by ε^2 , we can rewrite 5.5 as

$$-\frac{1}{2}\frac{\delta^2\phi(a)}{\varepsilon^2} + \frac{2}{\varepsilon^2}\sin^2\left(\frac{\pi}{d}a\right)\phi(a) = \frac{2}{\varepsilon^2}(1 - \lambda)\phi(a). \quad (5.6)$$

By interpolating on the points $(\sigma av(d)^{-2}, \sqrt{\sigma}\phi(a)v(d)^{-1}) = (x, \sqrt{\varepsilon}\phi(a))$, we obtain a smooth function $\psi(x)$ such that $\psi(a\varepsilon) = \sqrt{\varepsilon}\phi(a)$, for all $a \in \mathbb{Z}_d$. We can use 5.4 to write

$$\frac{\delta^2\phi(a)}{\varepsilon^2}\sqrt{\varepsilon} = \psi''(a\varepsilon) + O\left(\frac{1}{d}\right), \quad a \in \mathbb{Z}_d \quad (5.7)$$

and the Taylor series $\sin^2(\theta) = \theta^2 + O(\theta^4)$ to write

$$\frac{2}{\varepsilon^2} \sin^2\left(\frac{\pi}{d}a\right) = \frac{2}{\varepsilon^2} \left(\frac{\pi^2}{d^2} a^2 + O\left(\left(\frac{\pi}{d}a\right)^4\right) \right) = \frac{(a\varepsilon)^2}{2\sigma^4} + O\left(\frac{1}{d}\right), \quad a \in \mathbb{Z}_d. \quad (5.8)$$

Doing this for d very large, we can think of this as an equation that holds over \mathbb{R} . Then, dropping lower order terms, we get

$$-\frac{1}{2}\psi''(x) + \frac{x^2}{2\sigma^4}\psi(x) = \frac{2}{\varepsilon^2}(1 - \lambda)\psi(x), \quad (5.9)$$

which is the equation of the harmonic oscillator with angular frequency $\frac{1}{\sigma^2}$.

The corresponding ground state wave function is given by the Gaussian

$$\psi(x) \propto \exp\left(-\frac{x^2}{2\sigma^2}\right),$$

which is a Gaussian with variance σ^2 . Thus, the state $|\Gamma\rangle$ approaches this state ψ as we take the dimension d to infinity. This numerical argument justifies our choice of the Harper operator as the discrete Hamiltonian of a harmonic oscillator.

Chapter 6

Conclusion

In this work, we have provided a general framework for Gaussian states in discrete phase space. We introduced two new classes of Gaussian states via Bernstein distributions and a generalized Gibbs formalism. We explored examples and connections with the existing notion of stabilizer states. In particular, we showed that our Gaussian states are distinct from those in [20]. We proved that one can recover continuous Gaussian states as suitable limits of discrete ones. The behavior of different classes of Gaussian states mimics that of Gaussian probability distributions; they are equivalent in the continuous case, but distinct in the discrete.

We created an entirely new package `qWeyl` in `Mathematica` to aid in numerical computations. Our hope is that other researchers and graduate students will find this work and use the framework provided to tackle interesting questions related to states and channels. The associated `Mathematica` package should be a good way to get one's hands dirty experimenting with these objects as well as assisting in proofs.

One of the main motivations for this project was to investigate discrete analogues of Gaussian quantum channels. We believe this work has laid the foundations for such investigations and that the `qWeyl` package will be instrumental for future exploration of this topic.

This work raises questions which deserve further study. For example, examining the graphs in Figures 4.1 and 3.1, one might expect a convergence type result such as the ones exhibited in Section 5. Also, the set of Bernstein states and the set of Gibbs states appear to be distinct, but we have only limited numerical evidence at this time. All numerical examples of Bernstein states have nonnegative Wigner functions; could this be true in general? In what sense are our discrete Hamiltonians 4.25 *quadratic*? Lastly, Barker gives a notion of convergence of operators in [4]; could one extend this to density operators, and use this to show that our discrete Gibbs states approximate the continuous Gibbs states?

Appendix A

Gaussian Integrals

Here are two well known integral formulae; these are commonly known as *Gaussian Integrals*, due to the similarity of the integrand to a Gaussian probability distribution.

$$\int_{-\infty}^{\infty} \exp(-ax^2 + bx + c) dx = \sqrt{\frac{\pi}{a}} \exp\left(\frac{b^2}{4a} + c\right). \quad (\text{A.1})$$

$$\int_{-\infty}^{\infty} x^{2n} \exp(-ax^2) dx = \frac{(2n)!}{2^{2n} n! a^n} \sqrt{\frac{\pi}{a}} \quad (\text{A.2})$$

Appendix B

Code to Generate Figures

All figures within this document were generated using `Mathematica` and the package `qWeyl`. Many of these figures are better understood if one is able to rotate them, which is obviously not possible on paper. Also, it can be helpful when beginning to use a software package to have a series of worked examples to play with. For these reasons, we include in this appendix the code used to generate the figures shown in the main document. One should be able to run any of these and obtain the corresponding figure.

In all cases, one must first load in the package `qWeyl`;

```
In[71]= <<qWeyl'
```

Figure 2.1: Plot of three Gaussian probability distributions.

```
In[72]= plot = Plot[{PDF[NormalDistribution[0, 1], x],  
  PDF[NormalDistribution[3, 0.75], x],  
  PDF[NormalDistribution[-2, 2], x]}, {x, -7, 7},  
  AspectRatio -> 1/4, Filling -> Axis]
```

Figure 2.2: Plot of the single-variate Gaussian characteristic function.

```
In[73]= characteristic[mu_, sigmasquared_, z_] :=  
  Exp[I z mu - 1/2 z^2 sigmasquared]  
char1[z_] := characteristic[-2, 2, z]  
plot = Plot[{Re[char1[z]], Im[char1[z]]}, {z, -3, 3},  
  PlotRange -> Full, AspectRatio -> 1/2]
```

Figure 3.1: One of the continuous minimal uncertainty states. Change the values in `minUncerState` to plot others.

```
In[74]= minUncerState[mom_, pos_, zeta_, omega_: 1, hbar_: 1] :=  
  Surd[omega/(Pi hbar), 4]  
  Exp[(I mom)/hbar (zeta - pos/2) - (omega (zeta - pos)^2)/(2 hbar)]  
state[z_] := minUncerState[2, -2, z]  
plot1 = Plot[{Re[state[z]], Im[state[z]]}, {z, -5, 5},  
  PlotRange -> {{-6, 3}, {-.8, .8}}]
```

Figure 3.2: Continuous number states.

```
ln[75]:= plot = Plot[{hermiteGaussian[0, x], hermiteGaussian[1, x],
                    hermiteGaussian[4, x]}, {x, -5, 5}, AspectRatio -> 1/3]
```

Figure 3.3: Characteristic function of a continuous Gaussian state.

```
ln[76]:= cntsCharFcn[p_, q_, alpha_, x_, y_] :=
          Exp[I (q x + p y) - alpha/2 (x^2 + y^2)] // N
chrF[x_, y_] := cntsCharFcn[-2, 1, 2, x, y]
Plot3D[{Re[chrF[x, y]], Im[chrF[x, y]]}, {x, -2, 2}, {y, -2, 2},
        PlotStyle -> Opacity[0.7], PlotRange -> Full]
```

Figure 3.4: An example of a Wigner function and its marginals.

```
ln[77]:= coherentWigner[q1_, p_, a_, b_] :=
          1/(2 Pi (1 + Exp[-b^2/a^2]))
          Exp[-a^2 p^2] (Exp[-((q1 - b)^2/a^2)] + Exp[-((q1 + b)^2/a^2)]
          + 2 Exp[-q1^2/a^2] Cos[2 b p])
wig[q1_, p1_] := coherentWigner[q1, p1, 1, 4]
plot1 = Plot3D[wig[q1, p1], {q1, -7, 7}, {p1, -2, 2},
               PlotRange -> Full]
posMarginal[x_] := Integrate[wig[x, p], {p, -10, 10}]
momMarginal[x_] := Integrate[wig[q1, x], {q1, -10, 10}]
plot2 = Plot[posMarginal[x], {x, -8, 8},
             Ticks -> {{-8, 8}, {0.07, .14, .21, .28}}]
plot3 = Plot[momMarginal[x], {x, -2, 2},
             Ticks -> {{-2, 2}, {0.5, 1}}]
```

Figure 3.5: An example of a nonnegative Wigner function and its marginals. The plotting functions are identical to the previous example.

```
ln[78]:= mixedCoherentWigner[q1_, p_, a_, b_] :=
          1/(2 Pi) Exp[-a^2 p^2] (Exp[-((q1 - b)^2/a^2)] +
          Exp[-((q1 + b)^2/a^2)])
wig[q1_, p1_] := mixedCoherentWigner[q1, p1, 1, 4]
```

Figure 4.1: A discrete minimal uncertainty state.

```
ln[79]:= qInit[1, {19}]
alpha=0; beta=5;
ketLinePlot[{re[ab=ABstate[alpha, beta]], im[ab]}]
```

Figure 5.1: Shows how the Harper functions induce the Hermite-Gaussians. This gives the zeroth functions; change k1 to see the others.

```
In[80]:= qInit[1, {n=19}]
          k1=0;
          harp=generateHarperFunctions[];
          plot1=ketPlot[harp[[k1+1]]];
          plot2=Plot[v[n]-1hermiteGaussian[k1, v[n]-2x],
                    {x, -n/2, n/2}, PlotRange→Full];
          p1=Show[plot2, plot1]
```

Bibliography

- [1] L. Barker. The discrete fractional Fourier transform and Harper's Equation. *Mathematika*, 47:281–297, 2000.
- [2] L. Barker. Continuum quantum systems as limits of discrete quantum systems, I: State vectors. *Journal of Functional Analysis*, 186(1):153–166, 2001.
- [3] L. Barker. Continuum quantum systems as limits of discrete quantum systems, II: State Functions. *Journal of Functional Analysis*, 186(1):153–166, 2001.
- [4] L. Barker. Continuum quantum systems as limits of discrete quantum systems. III. Operators. *Journal of Mathematical Physics*, 42(10):4653–4668, 2001.
- [5] L. Barker. Continuum quantum systems as limits of discrete quantum systems. IV. Affine canonical transforms. *Journal of Mathematical Physics*, 44(4):1535–1553, 2003.
- [6] L. Barker, Ç. Candan, T. Hakioglu, M. A. Kutay, and H. M. Ozaktas. The discrete harmonic oscillator, Harper's equation, and the discrete fractional Fourier transform. *Journal of Physics A: Mathematical and General*, 33(11):2209–2222, 2000.
- [7] S. Bernstein. On a Characteristic Property of the Normal Law (in Russian). *Trudy Leningrad Polytechnic Institute*, 3:21–22, 1941.
- [8] L. Bing-Ren. *Introduction to Operator Algebras*. World Scientific, 1992.
- [9] S. Brierley. *Introduction to Operator Algebras*. PhD thesis, University of York, 2009.
- [10] T. Bröcker and R. F. Werner. Mixed states with positive Wigner functions. *Journal of Mathematical Physics*, 36(1):62–75, 1995.
- [11] W. B. Case. Wigner functions and Weyl transforms for pedestrians. *American Journal of Physics*, 76(10):937–946, 2008.
- [12] M. de Gosson. *Symplectic Geometry and Quantum Mechanics*, volume 66. Birkhäuser Basel, Basel, 2006.
- [13] J. Dehaene and B. De Moor. Clifford group, stabilizer states, and linear and quadratic operations over GF(2). *Physical Review A - Atomic, Molecular, and Optical Physics*, 68(4), 2003.

- [14] D. L. Donoho and P. B. Stark. Uncertainty Principles and Signal Recovery. *SIAM Journal on Applied Mathematics*, 49(3):906–931, 1989.
- [15] G. Feldman. *Functional Equations and Characterization Problems on Locally Compact Abelian Groups*. European Mathematical Society Publishing House, 2008.
- [16] G. B. Folland. *Harmonic Analysis in Phase Space*. Princeton University Press, 1989.
- [17] G. B. Folland. *A Course in Abstract Harmonic Analysis*. Chapman and Hall/CRC, 2016.
- [18] K. S. Gibbons, M. J. Huffman, and W. K. Wootters. Discrete phase space based on finite fields. *Physical Review A - Atomic, Molecular, and Optical Physics*, 70(6), 2004.
- [19] D. Gottesman. *Stabilizer Codes and Quantum Error Correction*. PhD thesis, Caltech, 1997.
- [20] D. Gross. Hudson’s theorem for finite-dimensional quantum systems. *Journal of Mathematical Physics*, 47(12):1–17, 2006.
- [21] T. Hashimoto, A. Hayashi, and M. Horibe. Minimum-uncertainty states and completeness of non-negative quasiprobability of finite-dimensional quantum systems. *Physical Review A*, 99(2), 2019.
- [22] C. J. Hillar and D. L. Rhea. Automorphisms of finite abelian groups. *American Mathematical Monthly*, 114(10):917–923, 2007.
- [23] A. S. Holevo. *Quantum Systems, Channels, Information*. De Gruyter, 2012.
- [24] R. L. Hudson. When is the Wigner quasi-probability density non-negative? *Reports on Mathematical Physics*, 6(2):249–252, 1974.
- [25] N. Kaiblinger and M. Neuhauser. Metaplectic operators for finite abelian groups and \mathbb{R}^d . *Indagationes Mathematicae*, 20(2):233–246, 2009.
- [26] M. A. Marchioli and P. E. Mendonça. Theoretical formulation of finite-dimensional discrete phase spaces: II. On the uncertainty principle for Schwinger unitary operators. *Annals of Physics*, 336:76–97, 2013.
- [27] M. A. Marchioli and M. Ruzzi. Theoretical formulation of finite-dimensional discrete phase spaces: I. Algebraic structures and uncertainty principles. *Annals of Physics*, 327(6):1538–1561, 2012.
- [28] A. Prasad. An easy proof of the Stone-von Neumann-Mackey Theorem. *Expositiones Mathematicae*, 29(1):110–118, 2011.
- [29] M. P. Quine. On Three Characterizations of the Normal Distribution. *Probability and Mathematical Statistics*, 14(2):257–263, 1993.
- [30] A. Serafini. *Quantum Continuous Variables*. Chapman and Hall, 2017.

- [31] F. Siyouri, M. El Baz, and Y. Hassouni. The negativity of Wigner function as a measure of quantum correlations. *Quantum Information Processing*, 15(10):4237–4252, 2016.
- [32] F. Soto and P. Claverie. When is the Wigner function of multidimensional systems nonnegative? *Journal of Mathematical Physics*, 24(1):97–100, 1983.
- [33] J. Stillwell. *Naive Lie Theory*, volume 46 of *Undergraduate Texts in Mathematics*. Springer New York, 2008.
- [34] C. Weedbrook, S. Pirandola, R. García-Patrón, N. J. Cerf, T. C. Ralph, J. H. Shapiro, and S. Lloyd. Gaussian quantum information. *Reviews of Modern Physics*, 84(2):621–669, 2012.
- [35] A. Weil. Sur certains groupes d’opérateurs unitaires. *Acta Mathematica*, 111:143–211, 1964.