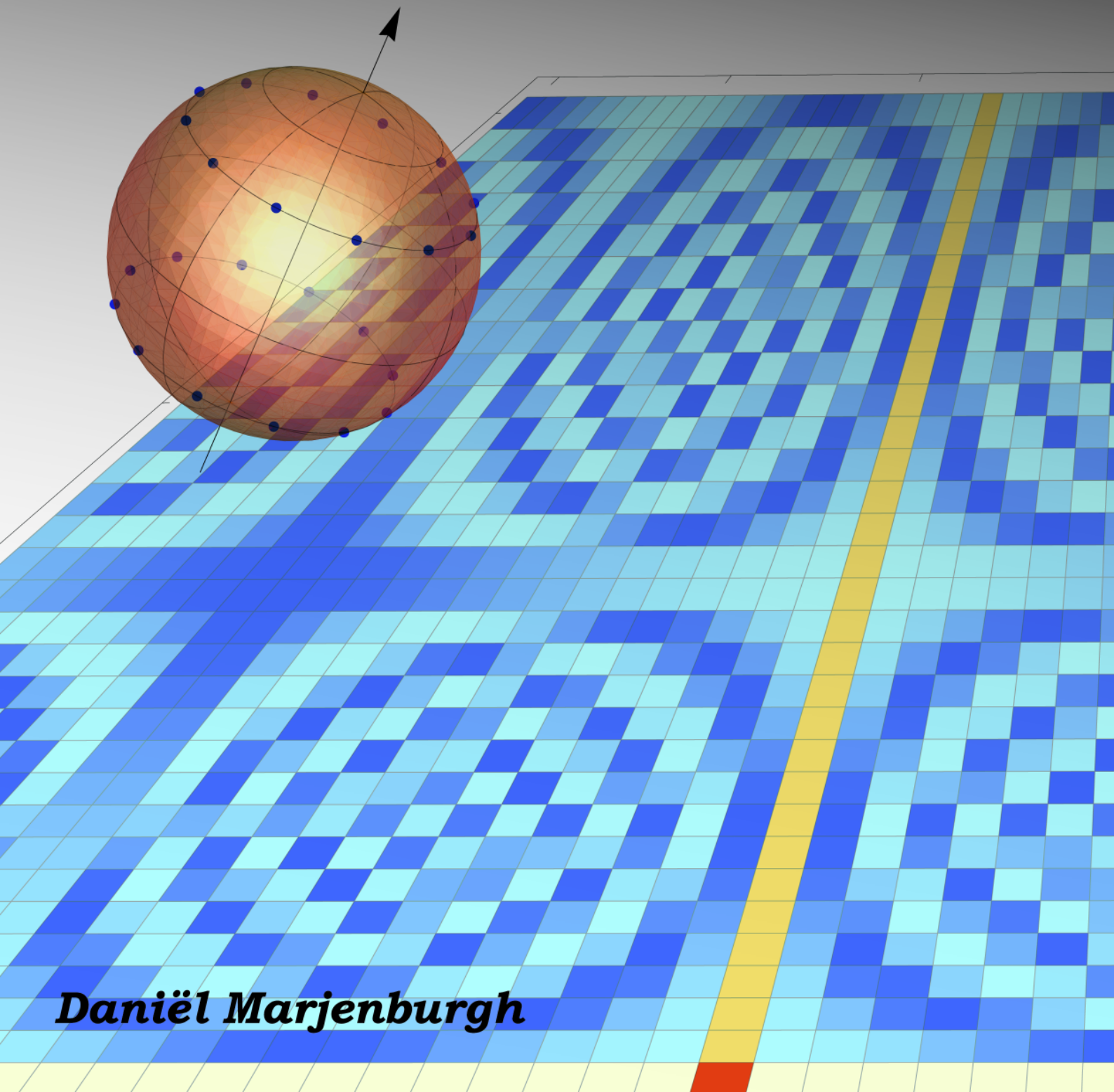
Geometrical Representations of Finite Dimensional Quantum Systems



Daniël Marjenburgh

Geometrical representations of finite dimensional quantum systems

Visualizing quantum states with phase spaces and spheres

Daniël Marjenburgh

March, 2010

Universiteit Leiden

Masterscriptie Theoretische Natuurkunde
Begeleider: Prof. G. Nienhuis
Leiden Institute of Physics, Quantum Optics and Quantum Information
Huygens Laboratorium, Universiteit Leiden

Contents

Preface

In the last few decades of the 21st century, an exciting new discipline has slowly been born out of a marriage between *Quantum Physics* and *Information Theory*. It is the field of *Quantum Information* which has, in turn, given rise to new areas of research such as *Quantum Computation*.

What *is* quantum computation? Before we answer that question, let's consider computation in the classical (non-quantum) sense, as a subject matter of computer Science. When we consider the field of Computer Science we see that it is a very theoretical and mathematical discipline dealing with the foundations of information and computation. Even the most complex computer algorithms in a PC, mobile phone or supercomputer can be conceptually brought down to the storage, manipulation and readout of basic units of information called *bits*. A bit is simply a system that can exist in either one of two states. For simplicity, these states are called 0 and 1, allowing us to speak of information in a device-independent way. This abstraction of reducing complex computational processes to simple manipulation of bits (through logic-gates) is very powerful. In reality of course, since an actual computer is physical, a physical realization of all computer components is needed. In practice, bits are encoded in the form of electric currents in circuit boards for example, or electromagnetic waves in wireless communication. Storage of bits can be realized through utilization of magnetic media (Hard Disk Drives, Flash memory) and optical discs (CD's, DVD's, Blu-ray) to name a few. For efficient manipulation of bits a modern processor has hundreds of millions of transistors. It is clear that a deep understanding of physics and materials science is needed to construct an actual computer.

Due to tremendous advances in the development of electronic circuitry over the last 30 years, the number of transistors that can be placed on an integrated circuit against constant costs is expected to double every two years. This trend in growing computer power, termed *Moore's Law*, has held since around 1970 and will likely continue to do so for another decade[1]. Because

[1]Originally Moore stated a doubling of transistors every 12 months, or a doubling in

of the constant miniaturization of computer circuitry this growth is expected to fail in the near future; with the current fabrication techniques, unwanted quantum size effects are steadily playing a greater role, interfering with the performance. Some revolutionary new breakthroughs in information processing will be needed to continue the growth.

Although the continuation of Moore's Law may be an interesting commercial challenge, one possible solution has exciting prospects from a scientific point of view. Since computers are physical, every computation is in essence a physical process. Rather than trying to continue to improve circuits that are used for the *classical* way of information processing (with bits), we should consider the laws of Nature that govern these processes and investigate how to apply these laws to our cause of computation. The laws are those of Quantum Mechanics and the basic unit serving as an analogue to the classical bit is the *Qubit* (Quantum bit), realized by a two-level quantum system.

A qubit differs in a number of important ways from a classical bit and the interesting ways in which multiple qubits can couple to each other through the phenomenon of entanglement allows us to use this as a resource to devise algorithms (called *quantum algorithms* naturally) that are able to perform computations which are believed to be impossible to do in any reasonable amount of time on a classical computer. A few such algorithms have been constructed which serve as a basis for all others. Unfortunately, the workings of even the simplest quantum algorithm, named the *Deutsch algorithm* which serves as a basic example of a quantum algorithm which a classical algorithm can not match, is conceptually not easy to 'see'. Although the Deutsch problem is simple to understand and its solution can be understood with basic quantum mechanics, actually devising the algorithm to solve the Deutsch problem is not very straightforward due to the nonintuitive nature of quantum behavior. This seems to be the norm in all quantum algorithms, which stands in contrast with classical algorithms which are generally more intuitive.

The goal of this thesis is to consider ways to represent finite dimensional quantum systems in a more intuitive way, through geometrical representations and phase-space functions. Aside from using these representations as a potentially helpful tool for devising quantum algorithms, it is a powerful way to gain physical insight into the often abstract description of a quantum system in terms of a complex Hilbert space. To be fair, gaining insight into the workings of nature, certainly that part that is beyond our senses, is a much more compelling stimulus to do research for a theoretical physicist.

---

performance every 18 months. This was later adjusted to a doubling every two years.

The setup of this thesis is as follows: In the introductory chapter I will briefly discuss the basic mathematical description of qubits (and quNits), entanglement, some basic quantum algorithms and establish the notation used throughout the thesis. Chapter two deals with the description of phase space, both classical and quantum mechanical. Chapter three is dedicated to the exploration of a phase space description of finite dimensional quantum systems. In chapter four we consider a geometric representation which is well suited to systems subject to rotational and/or permutational symmetries, such as systems describing spin angular momentum.

Chapter 1

Introduction

This chapter will cover a very brief introduction into the aspects of finite dimensional quantum systems that are the topic of this thesis. The first section describes the basic Hilbert space formulation of quantum mechanics, but is mainly intended to introduce the notation and terms we will use. The notation is highly standardized and widely used; however, there are some cases where different conventions are in use. I will comment about them in the cases I am aware of.

At the end of the chapter I will discuss some simple quantum algorithms and highlight the difficulties associated with it.

1.1 The Quantum Mechanical State Space

The mathematical description of a quantum mechanical system is that of complex Hilbert space $\mathcal{H}$. In the case where the dimension of $\mathcal{H}$ is finite this is simply a vector space over the field of complex numbers $\mathbb{C}$, supplied with an inner product. State vectors are denoted by kets $|\cdot\rangle$, where the dot $\cdot$ is replaced by a symbol labeling the state. The inner product between $|\phi\rangle$ and $|\psi\rangle$ is $\langle\phi|\psi\rangle$. The inner product is conjugate-symmetric: $\langle\phi|\psi\rangle = \langle\psi|\phi\rangle^*$ and linear in the second argument: $\langle\phi|\left(a|\psi_1\rangle + b|\psi_2\rangle\right) = a\langle\phi|\psi_1\rangle + b\langle\phi|\psi_2\rangle$. State vectors are given by a ket of unit length: $\langle\psi|\psi\rangle = 1$.
The state $|\psi\rangle$ is can not be distinguished from $e^{i\gamma}|\psi\rangle$; the phase γ has no observable effect. Sometimes, then, it is more convenient to consider these states identical and consider the space of pure density operators: $|\psi\rangle\langle\psi|$, where $|\psi\rangle$ is a state vector. Mixtures of states are just as easily handled in the density operator formalism. A general density operator is given by a convex linear combination of pure state ones:

$$\hat{\rho} = \sum_{i=1}^{n} p_i |\psi_i\rangle\langle\psi_i|, \qquad \sum_{i=1}^{n} p_i = 1$$

Operators are functions $\mathcal{H} \to \mathcal{H}$ and are always denoted with a hat, e.g: $\hat{A}$. Physical observables are associated with a Hermitian operator: $\hat{A}^\dagger = \hat{A}$, where the dagger denotes Hermitian conjugation. Upon measurement of $\hat{A}$ we obtain one of its eigenvalues λ_i. The probability of measuring λ_i is $\mathrm{tr}\left(\hat{\rho}\hat{P}_{\lambda_i}\right)$, where $\hat{P}_{\lambda_i}$ is the projection upon the eigenspace of $\hat{A}$ belonging to the eigenvalue λ_i. The average value of $\hat{A}$ is $\langle\hat{A}\rangle = \mathrm{tr}\left(\hat{\rho}\hat{A}\right)$.

For calculations we often choose an orthonormal basis; an eigenbasis associated with an observable for example. A neutral way to write this basis in quantum information, analogous to bit-representation, is $\{|n\rangle\}_{n=0}^{N-1}$. This is called the *computational basis*, although the term refers to the notation used to label the basis and not actually the basis itself.
An exception used often in this thesis is when we consider spin systems. States are in this case usually represented in the simultaneous eigenbasis of the spin operators $\hat{S}^2$ and $\hat{S}_z$ and written $|j,m\rangle$, where j is (half)-integer and m runs from $-j$ to j in integer steps. This basis is called the *angular momentum basis*.
The convention used to identify these bases in this $N = 2j+1$ dimensional Hilbert space is: $|j,m\rangle_{AM} \leftrightarrow |j-m\rangle_C$. For example, for a spin $j = 2$ system, we can make calculations in the computational basis $\{|0\rangle, |1\rangle, |2\rangle, |3\rangle, |4\rangle\}$ or the angular momentum basis $\{|2,2\rangle, |2,1\rangle, |2,0\rangle, |2,-1\rangle, |2,-2\rangle\}$. We identify $|1\rangle$ in the computational basis with $|2,1\rangle$ in the angular momentum basis etc. For spin 1/2, $|0\rangle$ is identified with spin-up and $|1\rangle$ with spin-down. This may at first sight seem contrary to natural expectations, but the reason will become clear in chapter 4.

1.2 Composite systems and entanglement

The composition of two quantum systems A and B is described by taking the tensor product $\mathcal{H}_A \otimes \mathcal{H}_B$ of the Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$ of the two subsystems. A system of particular interest in quantum computing is the N-qubit system: $\mathcal{H}_2^{\otimes N}$, where $\mathcal{H}_2$ is the Hilbert space for a single qubit. A basis for the N-qubit space is given by the tensor products of the single qubit basis states: $\{|x_1\rangle \otimes |x_2\rangle \otimes \cdots \otimes |x_N\rangle\}$ where each x_i is either 0 or 1. We will abbreviate the ket $|x_1\rangle \otimes |x_2\rangle \otimes \cdots \otimes |x_N\rangle$ by $|x_1\rangle|x_2\rangle \cdots |x_N\rangle$ or simply $|x_1 x_2 \cdots x_n\rangle$. So for a 2-qubit system the computational basis is given by $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$. These basis states are examples of *product states*. These are states of the form $|\psi\rangle_A|\phi\rangle_B$, where A, B denote the two subsystems. For product states, measurements and operations performed locally on one system bear no influence on the probability distribution for a measurement on the other system. States which are not product states are called *entangled states*. There are various degrees of entanglement and for systems consisting of a large number of subsystems there are many com-

plicated ways in which entanglement can occur. For a two-qubit system, a basis of maximally entangled states is given by the Bell states:

$$|\psi^-\rangle = \frac{1}{\sqrt{2}}\left(|01\rangle - |10\rangle\right)$$

$$|\psi^+\rangle = \frac{1}{\sqrt{2}}\left(|01\rangle + |10\rangle\right)$$

$$|\phi^-\rangle = \frac{1}{\sqrt{2}}\left(|00\rangle - |11\rangle\right)$$

$$|\phi^+\rangle = \frac{1}{\sqrt{2}}\left(|00\rangle + |11\rangle\right)$$

1.3 Quantum Algorithms

The purpose of a quantum computer is, eventually, to execute algorithms on N-qubit systems. Manipulation of bits is described by application of unitary operators, called *quantum gates* in this context. A prime example of a quantum gate operating on a single qubit is the *Hadamard* gate:

$$H = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

We now have the bare essentials needed to understand the simplest quantum algorithm: The Deutsch algorithm [DJ92, CEMM98].

1.3.1 The Deutsch Problem

The Deutsch problem is stated as follows. Suppose we have a black box that reads a single bit input x and gives an output bit $f(x)$, i.e. it acts as a function $f : \{0,1\} \to \{0,1\}$. Such a black box is called an *oracle*. The problem is to determine what this black box does; what function is f? It is clear that there are four possibilities for f, as shown in table 1.1. f can have a constant output of 0 or 1 on both inputs, or a balanced output of one 0 and one 1. How many *queries* are minimally necessary to ensure complete knowledge of f? Since each query generates at most 1 bit of information and knowledge of f (full knowledge of which of the 4 possibilities) equals 2 bits of information, we need at least 2 queries. This is not where a quantum computer outperforms a classical one, since each readout (measurement) of a qubit also gives no more than 1 bit of information. However, suppose we are not interested in the complete knowledge of f, but that we are only interested in knowing whether f is constant ($f(0) = f(1)$) or balanced ($f(0) \neq f(1)$), which is one bit of

f(0)	f(1)
0	0
0	1
1	0
1	1

Table 1.1: The possible black box functions f

information. Classically, this would still require two queries, but a quantum computer *can* do this in **one** query!
First, let's consider how to implement the function f in a black box. Since quantum evolutions are unitary, the box should enact a reversible operation. This is done, conveniently, with two registers. The input state is $|x\rangle|y\rangle$. The first ket is the data register, the second the target. The output is then $\hat{U}_f|x\rangle|y\rangle = |x\rangle|y \oplus f(x)\rangle$, which is clearly reversible. If the second qubit starts out as $|0\rangle$, it will simply become $|f(x)\rangle$.
The algorithm goes as follows:

1. Start with the state $|01\rangle$, apply a Hadamard gate to both qubits to obtain the state:

$$|\psi_1\rangle = \frac{1}{2}(|0\rangle + |1\rangle)(|0\rangle - |1\rangle) = \frac{1}{2}(|00\rangle - |01\rangle + |10\rangle - |11\rangle)$$

2. After the black box, the state has become:

$$\begin{aligned}\hat{U}_f|\psi_1\rangle &= \frac{1}{2}(|0\rangle|f(0)\rangle - |0\rangle|1 \oplus f(0)\rangle + |1\rangle|0 \oplus f(1)\rangle - |1\rangle|1 \oplus f(1)\rangle) \\ &= \frac{1}{2}(|0\rangle(|f(0)\rangle - |1 \oplus f(0)\rangle) + |1\rangle(|f(1)\rangle - |1 \oplus f(1)\rangle)) \\ &= \frac{1}{2}((-1)^{f(0)}|0\rangle(|0\rangle - |1\rangle) + (-1)^{f(1)}|1\rangle(|0\rangle - |1\rangle))\end{aligned}$$

   This is a product state, the first register is in the state:

$$\frac{1}{\sqrt{2}}((-1)^{f(0)}|0\rangle + (-1)^{f(1)}|1\rangle$$

3. Finally, we apply a Hadamard gate to the first register.
   If f is constant, then, apart from a global phase, the state becomes $|0\rangle$.
   If f is balanced, then the first register is $|1\rangle$.

The possibility of quantum computers to perform computations that can't be matched by classical computers is an exciting idea. However, it is not at all obvious that these particular steps in the algorithm lead to the desired result[1]. What are we to make of this? Should we accept the inherent difficulties associated with quantum programming? Or is there an alternative? Is the fact that such a relatively simple problem as the Deutsch problem requires a lot of trial and error to solve a sign that we are approaching it from an awkward angle, and that the notation used is clumsy and ill-suited

[1] In fact, the algorithm presented here differs from the original one proposed by Deutsch in 1985, which was not deterministic. It was improved to a deterministic one and generalized to N-qubits by Deutsch and Jozsa in 1992, which required 2 evaluations. In 1997 Cleve et al. improved it to require one query. As this is the simplest quantum algorithm, it is clear from this that the task of *quantum programming* is difficult and nonintuitive.

for this purpose?

In this thesis we consider various ways to visualize finite dimensional quantum systems. The hope is that with a clear picture of what's going on, we can gain increased insight into the physics of these systems and make the workings of these algorithms more transparent.

Bibliography

[CEMM98] R. Cleve, A. Ekert, C. Macchiavello, and M. Mosca. Quantum algorithms revisited. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 454(1969):339–354, 1998.

[DJ92] David Deutsch and Richard Jozsa. Rapid solution of problems by quantum computation. *Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences*, 439(1907):553–558, 1992.

Chapter 2

Phase Space

In this chapter we will go over the notion of phase space and phase space distributions. The first two sections are dedicated to a conceptual and largely non-technical overview of the phase space description of a classical system. The third section covers the phase space approach to quantum mechanical systems. Section four explains the concept of tomography and how it is related to our purposes. We will then, in section five, derive the celebrated Wigner function from a simple first principles approach that will guide us in finding a generalization for finite dimensional quantum systems, which is the subject of chapter 2.

Most of the material in the first sections is well known to nearly all physicists. Various notations are introduced in this chapter, where we have strived for uniformity and consistency throughout this thesis. Therefore, at least a quick read-through is recommended.
The goal and purpose here is simplicity, not thoroughness. Therefore, we will focus on a one-particle system with a single position degree of freedom.

2.1 Classical Phase Space

One of the simplest and useful kind of systems to analyze in classical mechanics are those of a particle with a single degree of freedom (e.g. position x) subject to a potential $V(x)$. The evolution of the particle is completely determined when we have specified the position and momentum of the particle at a certain moment in time. Each pair (x, p) thus corresponds to a *state* of the system. The space holding all possible states of the system is called the *phase space* [LL82, Gol01]. In the present example, it takes the form of a two-dimensional plane with the position variable x on one axis and the momentum variable p on the other. Each point in phase space thus uniquely corresponds to a state of the system under consideration. As the state of

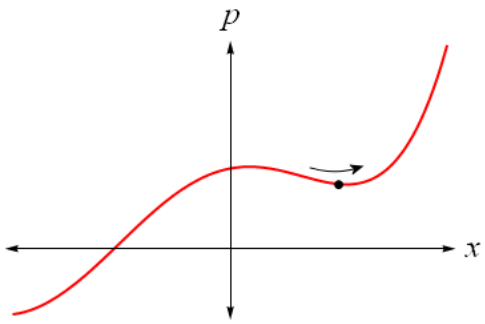


Figure 2.1: The classical phase space description for a particle with a single position degree of freedom. The system point follows a trajectory in the phase plane.

the system evolves, the system point follows a trajectory in phase space, see figure 2.1. The evolution of the system point is governed by Hamilton's equations of motion:

$$\dot{x} = \frac{\partial H}{\partial p}, \qquad \dot{p} = -\frac{\partial H}{\partial x}$$

Where $H(x,p) = \frac{p^2}{2m} + V(x)$ is the *Hamiltonian*. Because of conservation of energy, the point is constrained to move on a surface of constant energy: $H(x,p) = E$.

2.1.1 The Harmonic Oscillator

As an example, we will consider the classical harmonic oscillator with Hamiltonian $H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2$. As we see, the 'surface' of constant energy is an ellipse. As the system evolves, the system point moves clockwise along the ellipse.

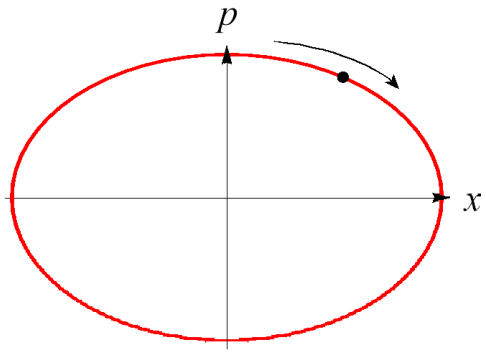


Figure 2.2: Phase portrait of a classical harmonic oscillator.

2.2 Classical Phase Space Distributions

Even though classical physics is deterministic, there are ways in which probability distributions can arise. For example, when we do not have enough information regarding the state of the system, we might wish to reflect this in our description by assigning a probability distribution $\rho(x,p)$ on the phase space. The $\rho(x,p)dxdp$ is then the probability that the system point is inside a phase space box of volume $dxdp$ centered around (x,p). In classical physics, this probability is a reflection of our ignorance about the *true* state of the system. Physicists often adhere to a frequency interpretation of probability in which $\rho(x,p)$ describes huge collection of identical systems, called an *ensemble*. The probability $\rho(x,p)dxdp$ is then the fraction of ensemble members inside the phase space box. Philosophical issues aside, phase space distributions are of great use in the analysis of classical Hamiltonian systems and are still used today (2010). For example, in the analysis of systems exhibiting chaos [Str00].

2.3 Quantum Mechanics in Phase Space

Now that we've (briefly) handled classical mechanics in phase space, let's examine a quantum system of a particle with one position degree of freedom. As quantum mechanics is more fundamental than classical mechanics, we cannot in honesty *derive* the path from a classical to a quantum description. There are some rules, however, which allow us to consider the quantum system corresponding to a classical system through a process called 'quantization'. Here are the top three ways to quantize a classical system:

1. **Canonical Quantization**: This is the standard operator approach developed by Heisenberg, Schrödinger, Dirac and others in the 1920's. Here the dynamical (conjugate) variables x and p become operators $\hat{x}$ and $\hat{p}$ acting in a Hilbert space which satisfy the canonical commutation relation $[\hat{x},\hat{p}] = i\hbar$.

2. **Path integral approach**: This approach to quantization was already conceived by Dirac in 1933 [Dir33] and worked out to a complete method by Feynman in 1948 [FH65, Fey48].

3. **Weyl Quantization**: Also known as "Phase Space Quantization" and "Moyal Quantization". It arose out of the work of Wigner (1932) [Wig32] and Weyl (1927) [Wey27]. The complete formulation was pulled together by the works of Groenewold (1946) [Gro46] and Moyal (1949) [Moy49].

These quantization schemes are all logically self-contained and equivalent. Weyl Quantization shows that quantum mechanics can be formulated fully

and autonomously in phase space, without recourse to Hilbert Space or operators. Our interest will not lie in this particular quantization scheme, we shall work quite often in the well known Hilbert Space formulation. Rather, we will use phase space as an aid to more fully seeing and understanding the structure of quantum systems. This all begins with the Wigner function.

2.3.1 The Wigner Function

In 1932 Eugene Wigner published a paper on quantum corrections to thermodynamic equilibrium where he introduced what is now known as the Wigner function [Wig32]. It is a phase space probability density function representing a quantum state. Its definition for a state given by the density operator $\hat{\rho}$ is:

$$W(x,k) \equiv \frac{1}{2\pi}\int_{-\infty}^{+\infty} dy \langle x - \frac{y}{2}|\hat{\rho}|x + \frac{y}{2}\rangle e^{iky} \tag{2.1}$$

In contrast to the classical case, we have favored the **wavenumber** k over the momentum $p = \hbar k$ to attain a cleaner notation without Planck's constants.

Although Wigner made no mention on how he arrived at this formula, it obeys some very desirable properties which make it such a useful tool. Instead of listing or deriving these properties here, we will adopt a first principles approach to defining a useful phase space distribution. In doing so, we will *derive* the Wigner function and some of its properties. There are a number of reasons for this approach:

1. Illuminating structure. This approach elucidates what characterizes the uniqueness of the Wigner function, and shows what properties can be derived.

2. Generality. When generalizing Wigner functions to different phase spaces (i.e. not just planes), we can carry over the defining properties of a Wigner function to guide us to a definition in a more general setting.

3. Elegance. All said and done, this path to the Wigner function turns out to be quite simple and elegant.

2.4 Phase Space Isometries

In this section we will discuss some useful operations that represent the two *orientation preserving isometries*[1] of the phase space. Since our phase space is a plane, any orientation preserving isometry can be built by combining:

[1] An isometry is a transformation of the plane that leaves all distances unaltered. Orientation preserving means that the 'handedness' of the plane is unchanged (i.e. no mirror image).

1. **Translations**

   and

2. **Rotations about the origin**

These isometries can be represented by unitary operators. Our purpose here is to introduce these operators and discuss their properties.

Because we would like to focus on the structure of the dynamics, we introduce the following dimensionless variables which take the place of the position and momentum operators $\hat{x}, \hat{p}$:

$$\hat{X} = \kappa\hat{x}, \qquad \hat{K} = \frac{\hat{p}}{\hbar\kappa}$$

Here, κ is a free to choose constant which can be adjusted to the particular problem at hand. For a harmonic oscillator, for example, a conventional choice is: $\kappa = \sqrt{m\omega\hbar}$.
We will refer to $\hat{X}$ and $\hat{K}$ simply as the position and momentum operators from now on. The canonical commutation relation takes the form:

$$[\hat{X}, \hat{K}] = i$$

And we also define the ladder operators $\hat{a}, \hat{a}^\dagger$ by:

$$\hat{a} = \frac{\hat{X} + i\hat{K}}{\sqrt{2}}, \qquad \hat{a}^\dagger = \frac{\hat{X} - i\hat{K}}{\sqrt{2}}$$
$$\hat{X} = \frac{\hat{a} + \hat{a}^\dagger}{\sqrt{2}}, \qquad \hat{K} = \frac{\hat{a} - \hat{a}^\dagger}{\sqrt{2}i}$$

Such that $[\hat{a}, \hat{a}^\dagger] = 1$. Lastly, we have the number operator $\hat{n} = \hat{a}^\dagger\hat{a}$ whose eigenstates are the *number states* $\{|n\rangle\}_{n=0}^{\infty}$.

2.4.1 Translations

It is well known that the momentum operator $\hat{K}$ generates translations in position space:

$$\hat{U}_a|x\rangle = \exp\left(-ia\hat{K}\right)|x\rangle = |x + a\rangle,$$

and the position operator $\hat{X}$ generates translations in momentum space:

$$\hat{V}_b|\widetilde{k}\rangle = \exp\left(ib\hat{X}\right)|k\rangle = |\widetilde{k + b}\rangle$$

We will use the tilde$\widetilde{}$over kets to accentuate that it is a momentum basis vector. We can combine the two operators in a single *displacement operator* $\hat{D}(b, a)$ that represents arbitrary translations in phase space:

$$\hat{D}(b, a) = \exp\left(i(a\hat{X} - b\hat{K}\right) = \hat{V}_a\hat{U}_b \exp\left(-\frac{iab}{2}\right) = \hat{U}_b\hat{V}_a \exp\left(\frac{iab}{2}\right)$$

Several useful properties of the displacement operator are shown in table 2.1.

2.4.2 Rotations

To find the generator of rotations, we take a cue from the classical simple harmonic oscillator. We know that here a system point rotates clockwise with angular frequency ω. An analogous result holds for the quantum harmonic oscillator. We consider the quantum mechanical states that are as 'classical as possible': the *coherent states* [Gla63b, Gla63a]. These are eigenstates of the annihilation operator $\hat{a}$ which spectrum is the complex plane:

$$\hat{a}|\alpha\rangle = \alpha|\alpha\rangle$$

The real and imaginary parts of α are related to the position and momentum, as is obvious by the relations:

$$\hat{X} = \frac{\hat{a} + \hat{a}^\dagger}{\sqrt{2}}, \qquad \hat{K} = \frac{\hat{a} - \hat{a}^\dagger}{\sqrt{2i}}$$

Therefore, a coherent state $|\alpha\rangle$ can be represented by a point in the phase plane. This point will rotate clockwise around the origin with angular frequency ω. So, thinking of the evolution operator $\hat{U}(t) = \exp\left(-i\hat{H}t/\hbar\right)$, where $\hat{H} = \hbar\omega(\hat{n} + 1/2)$, we can guess that the operator $\hat{R}(\theta) = \exp(i\theta\hat{n})$ performs a counterclockwise rotation in phase space. This is quickly verified for coherent states:

$$\hat{R}(\theta)|\alpha\rangle = e^{i\theta\hat{n}}e^{-|\alpha|^2/2}\sum_{n=0}^{\infty}\frac{\alpha^n}{\sqrt{n!}}|n\rangle = e^{-|\alpha|^2/2}\sum_{n=0}^{\infty}\frac{e^{i\theta n}\alpha^n}{\sqrt{n!}}|n\rangle = |\alpha e^{i\theta}\rangle$$

Since the coherent states form an overcomplete basis, $\hat{R}(\theta)$ represents rotations in phase space in general.
In two dimensions, a reflection about the origin sends the point (x, k) to $(-x, -k)$ and is equivalent to a rotation of 180° about the origin. If we consider the parity operator $\hat{P}$, defined by $\hat{P}|x\rangle = |-x\rangle$, then it is easily shown that $\hat{P} = \hat{R}(\pi)$.

Eigenstates of the Isometry Operators

It is clear from the expression of $\hat{R}(\theta) = \exp(i\theta\hat{n})$ that the number states $|n\rangle$ are eigenstates of the rotation operator. Therefore, their phase space representation must be rotationally symmetric about the origin. To find a set of eigenstates belonging to $\hat{D}(b, a)$, we use the knowledge that the position eigenstates $\{|x\rangle\}$ are eigenstates of $\hat{D}(0, 1) = \exp\left(i\hat{X}\right)$. Therefore, we define a rotated version of the position operator, $\hat{X}_\theta$, called a *rotated quadrature operator*:

$$\hat{X}_\theta = \hat{R}(\theta)\hat{X}\hat{R}^\dagger(\theta) = \hat{X}\cos\theta + \hat{K}\sin\theta$$

$\hat{D}(b_1,a_1)\hat{D}(b_2,a_2) = \hat{D}(b_2,a_2)\hat{D}(b_1,a_1)\exp\left(i(a_1b_2 - b_1a_2)\right)$	
$\hat{D}^\dagger(b,a) = \hat{D}(-b,-a)$	$\hat{D}(b,a) = \hat{U}_b\hat{V}_a e^{iab/2} = \hat{V}_a\hat{V}_b e^{-iab/2}$
$\hat{D}(b,a)\lvert x\rangle = e^{ia(x+b/2)}\lvert x+b\rangle$	$\hat{D}^r(b,a) = \hat{D}(rb,ra)$
$\hat{D}(b,a)\lvert\widetilde{k}\rangle = e^{-ib(k+a/2)}\lvert\widetilde{k+a}\rangle$	$\hat{D}(b,a)\hat{P}\hat{D}(b,a) = \hat{P}$
$\hat{P}\lvert\alpha\rangle = \lvert -\alpha\rangle$	$\hat{R}(\theta)\hat{a}^\dagger R^\dagger(\theta) = \hat{a}^\dagger e^{i\theta}$
$\hat{R}(\theta)\hat{a}\hat{R}^\dagger(\theta) = \hat{a}e^{-i\theta}$	$\hat{D}(re^{i\theta})\lvert\psi_t^\theta\rangle = \lvert\psi_{t+r}^\theta\rangle$
$\hat{R}(\theta)\hat{X}_\phi\hat{R}^\dagger(\theta) = \hat{X}_{\phi+\theta}$	$\hat{R}(\theta)\hat{D}(re^{i\phi})\hat{R}^\dagger(\theta) = \hat{D}(re^{i(\phi+\theta)})$
$\hat{R}(\theta)e^{ir\hat{X}_\phi}\hat{R}^\dagger(\theta) = e^{ir\hat{X}_{\phi+\theta}}$	$\hat{D}^\dagger(re^{i\theta})X_\phi\hat{D}(re^{i\theta}) = \hat{X}_\phi + r\cos(\theta-\phi)$

Table 2.1: Useful identities for the translation, rotation and quadrature operators

Consequently, the *rotated quadrature eigenstates* $\{|\psi_t^\theta\rangle\}$ are:

$$|\psi_t^\theta\rangle = \hat{R}(\theta)|t\rangle$$

Such that

$$\hat{X}_\theta|\psi_t^\theta\rangle = t|\psi_t^\theta\rangle$$

Table 2.1 contains a selection of useful properties concerning $\hat{D}, \hat{R}$ and their eigenstates.

2.5 Tomography

The word 'tomos' is ancient Greek, meaning 'cut','slice' or 'section'. Tomography, then, is imaging by sections. What exactly is meant by a 'section' or 'slice' in this context and why it is relevant to us will be explained below.

First, imagine that we are given an ensemble of particles in an identical quantum state, unknown to us. We would like to find out what this state is, i.e: find the density matrix $\hat{\rho}$. This is clearly an important problem for quantum state determination and quantum state preparation. Without any prior information, we cannot possibly determine $\hat{\rho}$ uniquely from a single measurement. Furthermore, we can't determine $\hat{\rho}$ from an arbitrary number of measurements of a single observable, like position. We could reconstruct $\mathrm{tr}\left(\hat{\rho}|x\rangle\langle x|\right) = \langle x|\hat{\rho}|x\rangle$ to arbitrary precision[2] with position measurements

[2]Theoretically speaking. Inevitable experimental uncertainties should be taken into account.

only, but this won't allow us to uniquely write down $\hat{\rho}$. To see this, consider the case where $\hat{\rho}$ is a pure state $|\psi\rangle\langle\psi|$. The wavefunction is $\langle x|\psi\rangle = \psi(x)$. With position measurements, we are able to reconstruct $|\psi(x)|^2$ to within any desired precision. But this does not allow us to distinguish between $\psi(x)$ and $\psi(x)e^{i\alpha(x)}$, where $\alpha(x)$ some function of x. As a concrete example, note that the entire set of plane waves $\psi(x) = \frac{1}{\sqrt{2\pi}} \exp(ikx)$ have a uniform position distribution and correspond to different momentum eigenstates of momentum $\hbar k$.

From the above, it seems we need to 'probe' different aspects/properties of the quantum state in order to 'reconstruct' it. In general, such a reconstruction procedure must depend heavily on the physical context, such as the accessible physical observables.
Let us be guided by our phase space setting to see what a position measurement entails. Since a phase space distribution $W(x,k)$ should, at least classically, be thought of as a *joint* probability distribution for the position and momentum, knowledge of the state's position distribution equals knowledge of the *marginal* distribution for the position. This is obtained by summing/integrating over the momentum variable:

$$\mathrm{tr}\left(\hat{\rho}|x\rangle\langle x|\right) = \int_{-\infty}^{+\infty} W(x,k)dk$$

The obtained function is called the *projection* of the phase space distribution along the k-direction. Similarly, by making measurements of the momentum (or wavenumber), we find the *projection* of W along the x-direction. These marginal distributions are still insufficient to determine $\hat{\rho}$ by far, as illustrated in figure 2.3.
The interpretation of these measurements as projections suggests that we should consider projections along arbitrary directions, parameterized by an angle θ. It turns out, remarkably, that from these projections the entire phase space distribution can be reconstructed. What we are measuring, the set of projections, is called the *Radon transform*[3] of W. To reconstruct W, we need to apply the inverse-Radon transform.

This is in itself a beautiful result, although some justifications regarding the physical meaning or feasibility of 'measuring projections along an arbitrary angle in phase space' is necessary. Integration along a parallel set of lines whose unit normal makes an angle θ with the x-axis amounts to measuring the observable $\hat{X}_\theta = \hat{X}\cos\theta + \hat{K}\sin\theta$. There are various homodyne detection schemes in optics which allows us to measure $\hat{X}_\theta$ [SBRF93][ZVB04]. They are widely used experimentally for tomographic reconstructions of quantum

[3]See appendix B for a mathematical description.

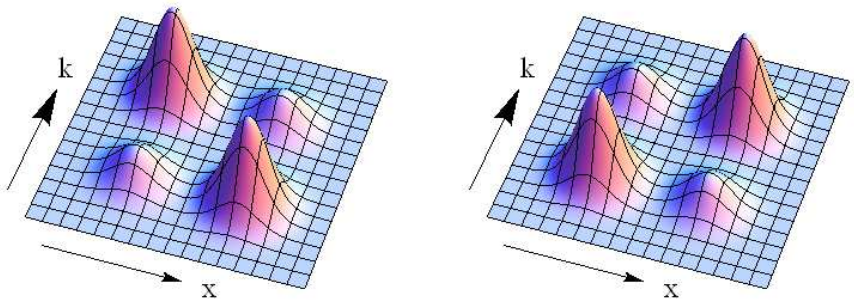


Figure 2.3: Two different Wigner functions having the same marginal position and momentum distributions. They are related by a quarter rotation.

states. For a more in-depth overview I will refer the reader to reference [LR09].

2.6 The Wigner Function from First Principles

2.6.1 Defining properties

In this section we will derive a phase space distribution $W(x,k)$ for a given density operator $\hat{\rho}$ by imposing essentially only two properties:

1. Galileo covariance
2. Tomographic property

The first property is a natural (nonrelativistic) physical requirement, as the choice of the origin of the phase space is essentially arbitrary by the equivalence principle. Since a Galileo boost of the system is identical to a translation in phase space, we will from now on refer to this property as *translation covariance*. The second property can be regarded as a *defining* property of the Wigner function. We want the marginal probability distributions along any direction to match with the probability distribution of the corresponding rotated quadrature observable.

2.6.2 Phase Point Operators

To cast the above requirements in a more mathematical form, we will introduce the concept of 'phase point operators'. By the tomographic property, the correspondence between $\hat{\rho}$ and $W(x,k)$ has to be linear, i.e. a convex linear combination of two density matrices $\lambda\hat{\rho}_1+(1-\lambda)\hat{\rho}_2$ with $0 \leq \lambda \leq 1$, leads to the same combination $\lambda W_1(x,k) + (1-\lambda)W_2(x,k)$ of the corresponding

Wigner functions $W_1(x,k)$ and $W_2(x,k)$. This is intuitively what we expect when we regard $\hat{\rho}$ as a statistical mixture of states and $W(x,k)$ as a probability density distribution. It follows that the correspondence between $\hat{\rho}$ and $W(x,k)$ can be written as:

$$W(x,k) = \frac{1}{2\pi}\mathrm{tr}\left(\hat{\rho}\hat{A}(x,k)\right) \tag{2.2}$$

The set of operators $\hat{A}(x,k)$ associated with each point in phase space are called *phase point operators*. We can use equation (2.2) to translate the requirements on $W(x,k)$ to requirements on $\hat{A}(x,k)$.

2.6.3 Translation Covariance

Galilean boosts of the system are represented by the *displacement operator* $\hat{D}(b,a)$:

$$\hat{D}(b,a) = \exp\left(i(a\hat{X} - b\hat{K})\right)$$

which generates translations in position and momentum. A boost of the system by (b,a) means the graph of W will be shifted a corresponding distance in phase space. So we require:

$$\hat{\rho} \to \hat{D}(b,a)\hat{\rho}\hat{D}^\dagger(b,a) \Rightarrow W(x,k) \to W(x-b,k-a)$$

As mentioned before, we can translate this to a constraint on the phase point operators $\hat{A}(x,k)$. We must have:

$$\mathrm{tr}\left(\hat{D}(b,a)\hat{\rho}\hat{D}^\dagger(b,a)\hat{A}(x,k)\right) = W(x-b,k-a) = \mathrm{tr}\left(\hat{\rho}\hat{A}(x-b,k-a)\right)$$

Since this relation has to hold for every density operator $\hat{\rho}$, this means $\hat{D}^\dagger(b,a)\hat{A}(x,k)\hat{D}(b,a) = \hat{A}(x-b,k-a)$ by result (3) in Appendix A. This means we only have *one* independent phase point operator, say $\hat{A}(0,0)$. The others are related through:

$$\hat{A}(x,k) = \hat{D}(x,k)\hat{A}(0,0)\hat{D}^\dagger(x,k) \tag{2.3}$$

2.6.4 The Tomographic Property

Recall the definition of the rotated quadrature operator:

$$\hat{X}_\theta = \hat{X}\cos\theta + \hat{K}\sin\theta$$

Its eigenstates are the rotated quadrature states:

$$\hat{X}_\theta|\psi_t^\theta\rangle = t|\psi_t^\theta\rangle$$

Considering the line through the origin, the tomographic property therefore requires:

$$\mathrm{tr}\left(\hat{\rho}|\psi_0^\theta\rangle\langle\psi_0^\theta|\right) = \int_{-\infty}^{+\infty} W(x,k)\delta(x\cos\theta + k\sin\theta)dxdk$$

Utilizing the correspondence (2.2) and proposition (3) in Appendix A as before, we find the following requirement for the phase point operators:

$$\frac{1}{2\pi}\iint_{\Phi} \hat{A}(x,k)\delta(x\cos\theta + k\sin\theta)dxdk = |\psi_0^\theta\rangle\langle\psi_0^\theta| \tag{2.4}$$

The above integral is over the entire phase space, which we will denote by Φ.

2.6.5 Radon Inversion

Equations (2.3) and (2.4) are the mathematical forms of the two requirements we ask of $\hat{A}(x,k)$. These properties can be combined into a single equation by using: $\hat{D}(b,a)\hat{D}(d,c) = \hat{D}(d,c)\hat{D}(b,a)e^{i(ad-bc)}$ (table 2.1). This gives:

$$\iint_{\Phi} \hat{A}(x,k)\delta(t - x\cos\theta - k\sin\theta)dxdk = 2\pi|\psi_t^\theta\rangle\langle\psi_t^\theta| \tag{2.5}$$

We have therefore fully specified the Radon transform of $\hat{A}$. We will now invert this expression to find $\hat{A}(x,k)$ by following the inversion process described in Appendix B.2.

First we will take the Fourier transform with respect to t, yielding the intermediate function $\hat{G}(r,\theta)$:

$$\hat{G}(r,\theta) = \sqrt{2\pi}\int_{-\infty}^{+\infty} e^{-irt}|\psi_t^\theta\rangle\langle\psi_t^\theta|dt = \sqrt{2\pi}\int_{-\infty}^{+\infty} e^{-ir\hat{X}_\theta}|\psi_t^\theta\rangle\langle\psi_t^\theta|dt = \sqrt{2\pi}e^{-ir\hat{X}_\theta}$$

Taking $a = r\cos\theta$ and $b = r\sin\theta$, we can write $\hat{\tilde{G}}(a,b)$ in terms of the displacement operator:

$$\hat{\tilde{G}}(a,b) = \sqrt{2\pi}e^{-i(a\hat{X}+b\hat{K})} = \sqrt{2\pi}\hat{D}(b,-a)$$

The last step is to take the inverse 2D Fourier Transform of $\hat{\tilde{G}}(a,b)$. We thus find:

$$\hat{A}(x,k) = \frac{1}{2\pi}\iint_{\Phi} e^{i(ax+bk)}\hat{D}(b,-a)dadb$$

In particular:

$$\hat{A}(0,0) = \frac{1}{2\pi} \iint_{\Phi} \hat{D}(b,-a) da db$$

To see what the action of this operator amounts to, we let it act on a position basis ket $|x\rangle$. We get:

$$\hat{A}(0,0)|x\rangle = \frac{1}{2\pi} \iint_{\Phi} \hat{D}(b,-a)|x\rangle da db = \frac{1}{2\pi} \iint_{\Phi} e^{-ia(x+\frac{b}{2})}|x+b\rangle da db$$

$$= \int_{-\infty}^{+\infty} \delta(x+\frac{b}{2})|x+b\rangle db = 2\int_{-\infty}^{+\infty} \delta(b+2x)|x+b\rangle db = 2|-x\rangle$$

Therefore $\hat{A}(0,0) = 2\hat{P}$, where $\hat{P}$ is the parity operator we mentioned in 2.4.2. By the translation property (2.3) we easily find the other phase point operators:

$$\hat{A}(x,k) = 2\hat{D}(x,k)\hat{P}\hat{D}^{\dagger}(x,k) \tag{2.6}$$

Using the definition of the phase point operators in equation (2.2) we can work out the following explicit expressions for the Wigner function:

$$W(x,k) = \frac{1}{\pi} \int_{-\infty}^{+\infty} \langle x-y|\hat{\rho}|x+y\rangle e^{2iky} dy \tag{2.7}$$

$$= \frac{1}{2\pi} \int_{-\infty}^{+\infty} \langle x-\frac{y}{2}|\hat{\rho}|x+\frac{y}{2}\rangle e^{iky} dy \tag{2.8}$$

in agreement with (2.1).

2.6.6 A Final Check

There is one possible caveat in the previous derivation. We know that Radon transforming a function and then inverse Radon transforming it yields back the original function. In our case however, we started with an expression $2\pi|\psi_\theta^t\rangle\langle\psi_\theta^t|$, *assumed* it to be the Radon transform of some function $\hat{A}(x,k)$ and then inverse Radon transformed to find what the $\hat{A}(x,k)$ should be. It is not necessary, however, that the Radon transform of $\hat{A}(x,k)$ yields back $2\pi|\psi_\theta^t\rangle\langle\psi_\theta^t|$, it might be that $2\pi|\psi_\theta^t\rangle\langle\psi_\theta^t|$ lies outside the range of the Radon transform. To check that this is not the case, we'll plug (2.6) into the Radon transform and verify that we indeed get the right projection operator. We parameterize the line $x\cos\theta + y\sin\theta = t$ by $(-r\sin\theta + t\cos\theta, r\cos\theta + t\sin\theta)$, where the r is the parameter. So we evaluate:

$$\frac{1}{2\pi} \int_{-\infty}^{+\infty} \hat{A}(-r\sin\theta + t\cos\theta, r\cos\theta + t\sin\theta) dr \tag{2.9}$$

We can rewrite the integrand with a combination of translations and rotations:

$$\hat{A}(-r\sin\theta + t\cos\theta, r\cos\theta + t\sin\theta) = \hat{D}(te^{i\theta})\hat{R}(\theta)\hat{A}(0,r)\hat{R}^{\dagger}(\theta)\hat{D}^{\dagger}(te^{i\theta})$$

$\hat{A}(0,r)$ is simply $2\hat{V}_r\hat{P}\hat{V}_r^{\dagger}$. Inserting the closure relation $\int |x\rangle\langle x| = 1$ between $\hat{V}_r$ and $\hat{P}$, we get:

$$\hat{A}(0,r) = 2\hat{V}_r \cdot 1 \cdot \hat{P}\hat{V}_r^{\dagger} = 2\int_{-\infty}^{+\infty} e^{irx}|x\rangle\langle -x|e^{irx}dx$$

Therefore, integrating this with respect to r:

$$2\int_{-\infty}^{\infty}\int_{-\infty}^{+\infty} e^{i2rx}|x\rangle\langle -x|dxdr = 4\pi\int_{-\infty}^{+\infty}\delta(2x)|x\rangle\langle -x|dx = 2\pi|0\rangle\langle 0| = 2\pi|\psi_0^0\rangle\langle\psi_0^0|$$

Putting it all together, we find that expression (2.9) equals:

$$\frac{1}{2\pi}\hat{D}(te^{i\theta})\hat{R}(\theta)\left(2\pi|\psi_0^0\rangle\langle\psi_0^0|\right)R^{\dagger}(\theta)\hat{D}^{\dagger}(te^{i\theta}) = |\psi_t^{\theta}\rangle\langle\psi_t^{\theta}|$$

So we verified that we indeed get the original function back.

2.7 Summary and Discussion

In this chapter we have given a quick overview of the concept of phase space. We have introduced the Wigner function and emphasized that the Wigner function is set apart from other phase space distributions by what we have called the *tomographic property*. The tomographic projection corresponds physically to a marginal distribution and this is what makes use of the Wigner function practical. Finally, we have shown that this tomographic property together with Galilean covariance completely and uniquely determines the Wigner function as introduced by Wigner himself in [Wig32].
Although the Wigner function has the flavor of a joint probability distribution of the position and momentum of a particle, it is not reconcilable with a true probability distribution. Arguing from Heisenberg's uncertainty principle or Bohr's complementarity principle, it is not meaningful to assign simultaneously both a definite position and a definite momentum to the particle. Thus the quantity $W(x,k)dxdk$ can generally not be interpreted as the probability that the particle has position in the range $[x, x+dx]$ and momentum in the range $[\hbar k, \hbar(k+dk)]$ as it does in the classical case. Mathematically, the failure of interpreting the Wigner function as a probability density is reflected by the possibility that $W(x,k)$ can become negative in certain regions. The Wigner function is, however, real, normalized and gives

the correct marginal probability distributions for all observables $\hat{X}_\theta$.

Given the phase space, there are several distributions on could define. The Wigner function is the most popular, given it is the unique distribution with the properties we described. We should mention two other alternatives: the Q- and P-distributions, which are related to the Wigner function. We will not go into these other distributions in any depth, but for completeness, appendix C discusses them. The definition of the Q-distribution requires coherent states, which will be useful for us in chapter 4.

Bibliography

[Dir33] P.A.M. Dirac. The Lagrangian in quantum mechanics. *Physikalische Zeitschrift der Sowjetunion*, Band 3(Heft 1):64Ű72, 1933.

[Fey48] R. P. Feynman. Space-time approach to non-relativistic quantum mechanics. *Rev. Mod. Phys.*, 20(2):367–387, Apr 1948.

[FH65] R.P. Feynman and A.R. Hibbs. *Quantum Mechanics and Path Integrals*. McGraw-Hill, 1965.

[Gla63a] Roy J. Glauber. Coherent and incoherent states of the radiation field. *Phys. Rev.*, 131(6):2766–2788, Sep 1963.

[Gla63b] Roy J. Glauber. Photon correlations. *Phys. Rev. Lett.*, 10(3):84–86, Feb 1963.

[Gol01] H. Goldstein. *Classical Mechanics*. Number ISBN 0201657023. Addison Wesley, 3rd edition, 2001.

[Gro46] H.J. Groenewold. On the principles of elementary quantum mechanics. *Physica*, 12(7):405 – 460, 1946.

[LL82] L.D. Landau and E.M. Lifshitz. *Mechanics*, volume 1 of *Course of Theoretical Physics*. Butterworth-Heinemann, 3rd edition, 1982.

[LR09] A. I. Lvovsky and M. G. Raymer. Continuous-variable optical quantum-state tomography. *Rev. Mod. Phys.*, 81(1):299–332, Mar 2009.

[Moy49] J. E. Moyal. Quantum mechanics as a statistical theory. *Mathematical Proceedings of the Cambridge Philosophical Society*, 45(01):99–124, 1949.

[SBRF93] D. T. Smithey, M. Beck, M. G. Raymer, and A. Faridani. Measurement of the Wigner distribution and the density matrix of a light mode using optical homodyne tomography: Application to

squeezed states and the vacuum. *Phys. Rev. Lett.*, 70(9):1244–1247, Mar 1993.

[Str00] Steven Strogatz. *Nonlinear Dynamics and Chaos*. Number ISBN 0-7382-0453-6. Perseus Publishing, 2000.

[Wey27] H. Weyl. Quantenmechanik und gruppentheorie. *Z. Phys.*, 46:1–46, 1927.

[Wig32] E. P. Wigner. On the quantum correction for thermodynamic equilibrium. *Phys. Rev.*, 40, 1932.

[ZVB04] Alessandro Zavatta, Silvia Viciani, and Marco Bellini. Tomographic reconstruction of the single-photon fock state by high-frequency homodyne detection. *Phys. Rev. A*, 70(5):053821, Nov 2004.

Chapter 3

Discrete Wigner Function

In this chapter we will investigate possible generalizations of the continuous Wigner function, discussed in the previous chapter, to finite dimensional quantum systems. We will see that this adapting to finite dimensions is not without difficulties. The dimension of the Hilbert space in question turns out to play a crucial role.

In the first section we will discuss some very useful properties of the continuous Wigner function. We would like to transfer as many of these properties to the finite dimensional case. In section two we start our descent into finite dimensional Hilbert space, starting by describing the phase space we will be using. Section three then covers a discussion of a Finite Radon Transform and its inverse in a form which is ideally suited for our purposes. After a thorough investigation we can then apply it to find an expression for the discrete[1] Wigner function from first principles, analogous to our approach in the first chapter, this is done in section four. In section five we investigate some properties of the Wigner function we have arrived at. We end with a summary and discussion of our results.

3.1 Some properties of the Continuous Wigner Function

In this section we will describe a few desirable properties of the continuous Wigner function and their relevance. These derivations are most easily done by investigating the properties of the phase point operators $\hat{A}(x,k)$. To

[1] By discrete, we explicitly mean *finite*.

recap, the definition of the Wigner function of a density matrix is:

$$\begin{aligned} W(x,k) &= \frac{1}{2\pi}\mathrm{tr}\left(\hat{\rho}\hat{A}(x,k)\right) \\ \hat{A}(x,k) &= 2\hat{D}(x,k)\hat{P}\hat{D}^{\dagger}(x,k) \end{aligned} \tag{3.1}$$

The first two properties are of course translation covariance and the tomographic property which allows us to reconstruct the Wigner function from a set of measurements. These are the properties we imposed and so this is not surprising.

The Wigner function is real and normalized. The first one follows from the Hermiticity of the phase point operators, the second from the tomographic property, since each quadrature basis $\{|\psi_t^{\theta}\rangle\}_t$ is complete.

Although we have not imposed any requirements on rotational aspects, the Wigner function is also well behaved under rotations. That is, if we rotate our system, the Wigner function rotates covariantly:

$$\hat{\rho} \to \hat{R}(\theta)\hat{\rho}\hat{R}^{\dagger}(\theta) \Rightarrow W(z) \to W(ze^{i\theta}),$$

where we have used a complex number representation $z = x + iy \leftrightarrow (x, y)$ for clarity.

The last important property is the *overlap property*:

$$\mathrm{tr}\left(\hat{\rho}\hat{\rho}'\right) = 2\pi \int_{\Phi} W_{\hat{\rho}}(x,k)W_{\hat{\rho}'}(x,k)dxdk$$

Thus the amount of overlap between two Wigner functions is directly related to the 'amount of overlap' between the corresponding quantum states.
Also, if we transform an observable $\hat{O}$ to a phase space function $O(x,k)$ in a similar, though not identical, manner as a density operator:

$$O(x,k) = \mathrm{tr}\left(\hat{O}\hat{A}(x,k)\right), \tag{3.2}$$

then the overlap property says that we can evaluate the average value of $\hat{O}$ as:

$$\langle\hat{O}\rangle = \int_{\Phi} O(x,k)W(x,k)dxdk,$$

which is what we would have classically. Note the factor of 2π that is present when transforming a density operator to a phase space function (equation 3.1), but absent when transforming an observable to a phase space function (equation 3.2).
The overlap property is a natural consequence of the tomographic property and translation covariance and will continue to hold in the finite dimensional case. We will show and prove this in section 3.5.4.

3.2 Finite Dimensional Phase Space

We now turn to the case where our quantum system is described by a finite, N-dimensional Hilbert space $\mathcal{H}$. The first task is to define the arena in which our Wigner function will live. What should our phase space look like?
In the continuous case, the phase space axes correspond to two canonically conjugate variables, $\hat{X}$ and $\hat{K}$, such that $[\hat{X}, \hat{K}] = i$. In finite dimensional Hilbert space this commutation relation can no longer hold. An easy way to see this is to take the trace of the commutator on both sides, together with $\mathrm{tr}\left(\hat{X}\hat{K}\right) = \mathrm{tr}\left(\hat{K}\hat{X}\right)$:

$$\mathrm{tr}\left([\hat{X}, \hat{K}]\right) = \mathrm{tr}\,(i) = iN$$

while[2]

$$\mathrm{tr}\left([\hat{X}, \hat{K}]\right) = \mathrm{tr}\left(\hat{X}\hat{K} - \hat{K}\hat{X}\right) = \mathrm{tr}\left(\hat{X}\hat{K}\right) - \mathrm{tr}\left(\hat{K}\hat{X}\right) = 0$$

First we need to assign a basis $\{|n\rangle\}_{n=0}^{N-1}$ in $\mathcal{H}$ which serves as the analogue of the position basis in the continuous case. We will call this the *computational basis*. By the tomographic property, we require that the sum of the Wigner function down any vertical line give an outcome probability for a 'position' measurement. Repeating the same argument as in the previous chapter, this means that the sum over the phase point operators $\hat{A}(n,k)$ down a vertical must equal something proportional to a projection operator $|n\rangle\langle n|$. Since there are N distinct operators, there should be N distinct vertical lines to sum over. The same argument for horizontal lines shows that a phase space of $N \times N$ points, laid out like a grid is a natural choice. There is another reason why this is a preferable choice. The number of independent real parameters to uniquely characterize an N-dimensional density matrix is N^2-1. We need $2N^2$ real parameters for an arbitrary complex matrix, N^2 if it is Hermitian and $N^2 - 1$ if it has unit trace. A density operator also has the property that it is positive. This is, however, a nonholonomic constraint[3], so we can disregard it in our counting of the degrees of freedom[4]. Specification of a normalized, real valued function defined on an $N \times N$ grid also requires $N^2 - 1$ real parameters. This means that there will be no redundancy in our definition with this choice.

[2] This argument fails for an infinite-dimensional Hilbert space, because the trace is then no longer a well defined operation on all operators. In particular, the trace of I is not defined.

[3] A nonholonomic constraint is a constraint which does not reduce the number of degrees of freedom. In this case it takes the form: $\lambda_i \geq 0$, where the λ_i are the eigenvalues of $\hat{\rho}$.

[4] By degrees of freedom, we simply mean the number of independent real parameters required to fully specify, without redundancy, our object of interest.

It should be noted in advance that this choice is not without difficulties, which will become clear later in this chapter. There exist formulations which utilize a $2N \times 2N$ phase space to combat these difficulties [HB80, Leo95, Leo96] and these have been used in descriptions of quantum algorithms [MPS02a, MPS02b]. The only real issue with this choice is the arising redundancy; the phase space is four times as large as it needs to be. In this thesis we will only investigate the possibility of defining a discrete Wigner function on an $N \times N$ phase space.

3.2.1 The Conjugate Basis and the Displacement Operator

We have already chosen a computational basis $\{|n\rangle\}_{n=0}^{N-1}$, functioning like the 'position' variable, as it does in the continuous case. What bases should we choose for the 'momentum' basis, labeled $\{|\widetilde{k}\rangle\}_{k=0}^{N-1}$? As they are conjugate, per Heisenberg's relation, precise knowledge of the one variable should imply complete uncertainty about the other. That is, we require:

$$|\langle n|\widetilde{k}\rangle|^2 = \frac{1}{N} \tag{3.3}$$

The value of $1/N$ is fixed by normalization of the basis vectors. A pair of bases satisfying equation (3.3) are called *mutually unbiased*. In the continuous case actually, the uncertainty relation gives the exact basis change between the position and momentum eigenstates:

$$\begin{aligned} \langle x|\widetilde{k}\rangle &= \frac{1}{\sqrt{2\pi}}\, e^{ikx} \\ |\widetilde{k}\rangle &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} dx\, e^{ikx} |x\rangle \end{aligned} \tag{3.4}$$

The N-dimensional analogues of these are:

$$\begin{aligned} \langle n|\widetilde{k}\rangle &= \frac{1}{\sqrt{N}}\, e^{\frac{2\pi ikn}{N}} \\ |\widetilde{k}\rangle &= \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} e^{\frac{2\pi ikn}{N}} |n\rangle \end{aligned} \tag{3.5}$$

Just as the position and momentum bases are related by a Fourier transform, so are the computational and conjugate bases related by a *discrete Fourier transform*:

$$|\widetilde{k}\rangle = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \exp\left(\frac{2\pi ikn}{N}\right) |n\rangle \tag{3.6}$$

$$|n\rangle = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \exp\left(-\frac{2\pi ikn}{N}\right) |\widetilde{k}\rangle \tag{3.7}$$

We introduce the discrete translation operators $\hat{U}$ and $\hat{V}$ by their action on these bases:

$$\hat{U}|n\rangle = |n+1\rangle$$
$$\hat{V}|\widetilde{k}\rangle = |\widetilde{k+1}\rangle$$

In the continuous case, the momentum and position operators act as generators of space and momentum respectively and we have:

$$\hat{U}_x|\widetilde{k}\rangle = \exp(-ikx)\,|\widetilde{k}\rangle$$
$$\hat{V}_y|n\rangle = \exp(iny)\,|n\rangle$$

It's easy to check that the discrete analogues of these equations hold:

$$\hat{U}^x|\widetilde{k}\rangle = \exp\left(-\frac{2\pi ikx}{N}\right)|\widetilde{k}\rangle$$
$$\hat{V}^y|n\rangle = \exp\left(\frac{2\pi iny}{N}\right)|n\rangle$$

From the above, it should be clear that we need to consider the labels in our basis kets $|n\rangle$ periodically, i.e. modulo N. Since translations should be unitary, finiteness of our Hilbert space together with $\hat{U}|n\rangle = |n+1\rangle$ implies that $\hat{U}|N-1\rangle = |N\rangle = |0\rangle$. Therefore, our phase space is not actually an $N \times N$ grid, but an $N \times N$ discrete torus (Fig. 3.1): If we walk along the horizontal or vertical directions, we end up at the same spot after N steps.

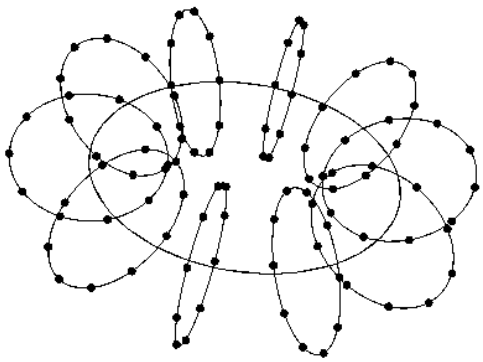
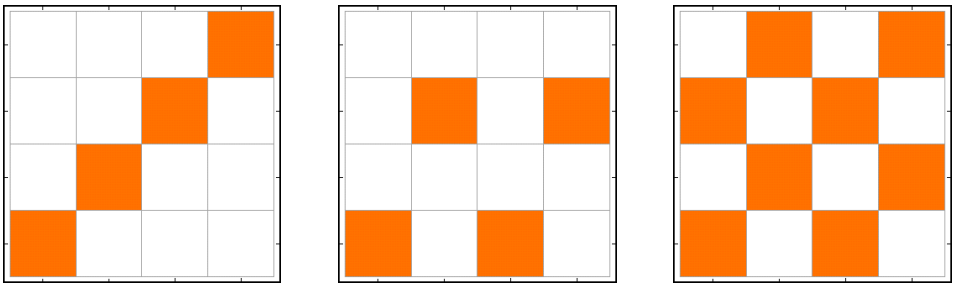


Figure 3.1: The topological structure of our finite phase space for $N = 10$. A discrete torus with ten circles, each having ten points. A point is specified by a pair (n, k) denoting the circle number and point on that circle.

3.2.2 Lines in Phase Space

If we wish to generalize the tomographic property to our discrete phase space, we need to understand what we mean by a 'line' in this case. In the



(a) $n - k = 0$ (b) $2n - k = 0$ (c) $2n - 2k = 0$

Figure 3.2: Three lines in phase space for $N = 4$. (a) is satisfactory, (b) is less well behaved, but will be necessary, (c) should not be allowed.

Euclidean plane, a line extends indefinitely in both directions. Since our discrete phase space is periodic, a line would be 'wrapped around' the plane.

Intuitively, we want lines to be 'straight', meaning the ratio between horizontal and vertical displacement of two points on the line should be the same. We can also go the algebraic route and say we define a line by the set of points (n, k) satisfying the equation $bn - ak = c$, where all parameters take values in $\mathbb{Z}_N$ (the set of integers modulo N) and $(a, b) \neq (0, 0)$. Although these definitions are nearly identical and may seem like a good choice, they contain some nasty cases which we would not like to include.
To illustrate this with an example, let's take $N = 4$. In $\mathbb{Z}_4$ the equation $2n - 2k = 0$ is satisfied when $n - k = 0$ or $n - k = 2$. This 'line' fills out *half* of the entire phase space (8 points). The line $n - k = 0$ has only 4 points and is *contained* in the previous line (Fig. 3.2).

A different kind of problem may arise for lines like $2n - k = 0$. This line intersects two of the horizontal lines twice. When we repeatedly translate this line one step horizontally we create only one other line and their union does not fill out the entire phase space.

At this point, a little excursion into some elementary number theory is necessary. The mathematical reason for multiple intersects with a horizontal axis is that the slope, 2 in the example, is not invertible modulo 4. That is, there is no element $x \in \mathbb{Z}_4$ such that $2x = 1$. An element that is invertible is called a *unit*. When every element of $\mathbb{Z}_N$ except 0 is a unit, $\mathbb{Z}_N$ is called a *field*[5]. Thus in a field, division through any nonzero number is possible. A basic result in number theory is that an element $x \in \mathbb{Z}_N$ is invertible if and only if the greatest common divisor $\gcd(x, N) = 1$. In this case x and N are

[5]The mathematical algebraic meaning of *field* is thus completely different from the same word in physics, e.g. in the *electric field*.

called *coprime*. By definition, the number of positive integers less than or equal to N that are coprime to N is equal to Euler's totient function $\varphi(N)$. If all positive integers less than N are coprime to N, then N has to be a prime number. Therefore, $\mathbb{Z}_N$ is a field if and only if N is a prime number.

Considering the important role lines will play we would like to impose to following requirements on them:

- Each line must contain N points
- Each line intersects each row and column exactly once

These restrictions ensure that each line has $N-1$ other parallel[6] lines which together fill out the phase space. such a set of N parallel lines is called a *parallel class*. Mathematically this means that we require a and b in the equation $bn - ak = c$ to be units or 0 (but not both zero). The number of 'slanted' directions (i.e. not horizontal or vertical) in our phase space by this rule is then $\varphi(N)$, since we can always write the equation $bn - ak = c$ in the form $k = mn + t$ where the coefficient of k is 1. So each unit gives rise to a certain *direction* in phase space. Together with the horizontal and vertical directions, there are $\varphi(N) + 2$ directions. Therefore there are also $\varphi(N) + 2$ parallel classes, each containing N lines.

Unfortunately, although we want to take all lines of the form $k = mn + t$, with m a unit, into account, these are not sufficient in general when N is not a prime number. We can see this if we make a count of the number of degrees of freedom. How many independent real parameters can we distill if we measure in a basis belonging to a parallel class? There are N distinct outcomes whose probability add to 1. So for each measurement corresponding to a direction we can learn $N-1$ real parameters. Thus we need $N+1$ directions to learn the $(N+1)(N-1) = N^2 - 1$ real parameters specifying the density matrix, therefore we need $\varphi(N) = N - 1$ which is only true if N is prime.
There is another way to see that the $\varphi(N) + 2$ directions give insufficient information to reconstruct the Wigner function when N is not prime. Suppose we wish to learn the value of $W(0,0)$ when we are given the sum over lines P_{λ_i} for each of the 6 lines λ_i in the figure below. Clearly only the lines through the origin contribute to information about the value $W(0,0)$. The solution is to take the total sum of all the given projections. Each point is counted once except for the origin, which is contributed 6 times. Since the Wigner function is normalized, we find that the value of $W(0,0)$ is:

$$W(0,0) = \frac{1}{5}\left(\sum_{i=1}^{6} P_{\lambda_i} - 1\right)$$

[6] Two lines are parallel when one can be made to coincide with the other by translating each point by the same vector modulo N.

Contrast this with the case where $N = 4$. Since now we cannot *reach* each point of the phase space from lines through the origin, our information is no longer sufficient. We need to include lines $k = mn + t$ where m is not a unit. This will cause redundant *oversampling* of some points in the Radon transform[7]. The direction (b, a) along the line $an - bk = t$ must satisfy $\gcd(b, a, N) = 1$ to ensure it has N points, but we don't need *all* such directions. How many and which directions are necessary to ensure that the finite Radon transform arising from this is invertible has been worked out by A. Kingston and I. Svalbe [KS07]. Although the finite Radon transform developed in their paper has likely been intended mainly for image processing applications, the periodicity of the $N \times N$ array is perfectly suited for our purposes of investigating discrete Wigner functions.

3.3 The Finite Radon Transform (FRT)

The continuous Radon transform (or rather its *inverse*) (Appendix B) has found an incredible number of uses, for example in medical imaging, seismology, signal analysis, digital image processing and quantum state tomography. We have found that we can invert the Radon transform given the continuous set of projections for each angle θ. In reality, we can only obtain a finite set of projections. The question of what set of projection angles to take and how to reduce consequential artifacts in the reconstructed image is an important and active area of research. For digital image processing purposes, it may not always be advantageous to consider an $N \times N$ image as a periodic array, or to require an exact inverse transform as information loss is acceptable in image compression. For this reason, there have been different constructions generalizing the continuous Radon transform to the finite case. For our purposes, it is clear that we must have a definition which considers periodic arrays and has an exact inverse. The first definition for periodic arrays was developed by Matúš and Flusser (1993) [MF93]. This definition only applied to arrays of size $p \times p$, where p is a prime. Hsung, Lun and Siu (1996) developed a periodic finite Radon transform which also applied to dyadic arrays of size $2^n \times 2^n$ [HS96]. Kingston extended this to a definition on prime-adic arrays of size $p^n \times p^n$ where p is prime. The generalized finite Radon transform for any square $N \times N$ array was completed by Kingston and Svalbe (2007) [KS07]. In this section we will briefly summarize the results of the paper before we apply them to our purposes.

Each parallel class is given by a direction vector (b, a) such that the lines in that class are of the form $an - bk = t$ for $t = 0, \ldots, N - 1$. This way,

[7] A more convincing proof of why we need other lines to create oversampling is given in 3.3.1, after the definition of the finite Radon transform

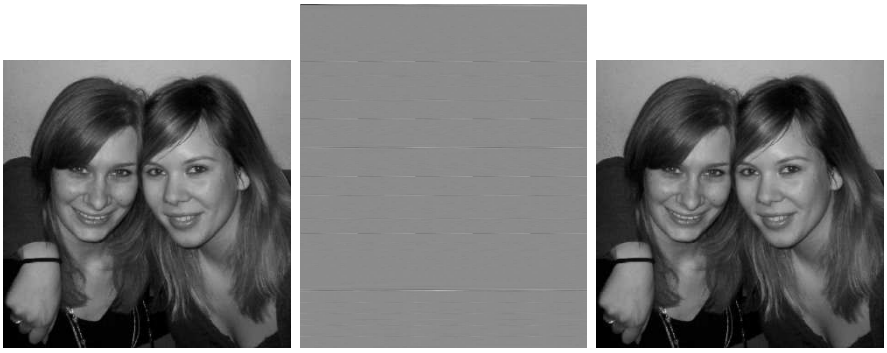


Figure 3.3: (left) A $5^4 \times 5^4$ image of Astrid and Eline, (middle) the 625×750 Radon transformed image, rescaled to the interval $[0, 255]$ and (right) the exactly reconstructed image using the IFRT of section 3.3.1.

the line through the origin and direction (b, a) is given by the set of points (sb, sa) $(s = 0, \ldots, N-1)$. The set of direction vectors is denoted by Θ_N. This set and the way it is constructed depends on whether N is prime (the simplest case), N is a prime power or neither (the most complex case). The construction of Θ_N, though described in [KS07], is actually not important to our purposes. We will therefore assume Θ_N given with the understanding that, for N not prime, oversampling occurs.
The number of lines passing through the origin going and the point (x, y) is given by the oversampling function $v_N(x, y)$. When we choose a direction $(b, a) \in \Theta_N$ and consider the sequence of points (sb, sa), for $s = 0, \ldots, N-1$, this gives rise to a sequence $v_N(sb, sa)$. A remarkable result is that this sequence is *independent* of the direction (b, a). Therefore we can consider the oversampling function as a function of one variable only: $v_N(s)$.

With these definitions in place we can state what the FRT (Finite Radon Transform) of a function $W(n, k)$ defined on $\mathbb{Z}_N \times \mathbb{Z}_N$ is:

$$R_{b,a}(t) = \sum_{n,k=0}^{N-1} W(n, k)\delta_N(an - bk - t), \qquad \text{for } (b, a) \in \Theta_N, \tag{3.8}$$

where δ_N is the periodic Kronecker delta: $\delta_N(x) = 1$ if $x \equiv 0 \mod N$ and $\delta_N(x) = 0$ otherwise.

As an example, we consider $N = 5^4 = 625$. Figure 3.3 shows a gray-scale 625×625 image, which we can regard as a discrete phase space function. The pixel values are bytes, ranging from 0 (black) to 255 (white). The number of directions can be shown to be $5^4(1 + 1/5) = 750$, so the Radon transformed image will have dimensions 625×750.

3.3.1 The inverse Finite Radon Transform (IFRT)

There are several ways to invert the FRT. Before stating the method we will use, we will take the (inverse) Fourier transform of $R_{b,a}(t)$ with respect to t, just like in the continuous case. As a result we will see why oversampling is necessary. We get:

$$\begin{aligned}\sum_{t=0}^{N-1} R_{b,a}(t)e^{-2\pi iut/N} &= \sum_{x,y,t=0}^{N-1} W(x,k)\delta_N(ax-by-t)e^{-2\pi iut/N} \\ &= \sum_{x,y=0}^{N-1} W(x,k)e^{2\pi iu(by-ax)/N} = \widehat{W}(-ua,ub),\end{aligned}$$

where $\widehat{W}$ is the 2D discrete Fourier transform of W. We see that, unless the set $(-ua, ub)$ with $(b,a) \in \Theta_N$ and $u = 0,\ldots,N-1$, cover the entire grid, $\widehat{W}$ is not completely specified. Since the function $\widehat{W}$ holds the same information as W and $R_{b,a}(t)$ (as they are related by Fourier transforms), we cannot reconstruct W if the set of lines through the origin, generated by $(b,a) \in \Theta_N$ do not cover the entire grid.

Inversion of equation (3.8) is done by the following steps, as explained in [KS07].

1. From $R_{b,a}(t)$, we find the *filtered back-projection* $\tilde{R}_{b,a}(\eta)$, defined by the circular convolution:

$$\tilde{R}_{b,a}(\eta) = \sum_{\tau=0}^{N-1} R_{b,a}(\eta-\tau)H(\tau), \tag{3.9}$$

   where the Kernel H is given by:

$$H(\tau) = \sum_{w=0}^{N-1} \frac{1}{v_N(w)} \exp\left(\frac{2\pi i\tau w}{N}\right) \tag{3.10}$$

2. With $\tilde{R}_{b,a}(\eta)$ in hand, the original function is given by:

$$W(n,k) = \frac{1}{N^2} \sum_{(b,a)\in\Theta_N} \tilde{R}_{b,a}(an-bk) \tag{3.11}$$

To prove that this indeed inverses the FRT, we can plug equation (3.8) into equation (3.11), using equations (3.9) and (3.10). We start by evaluating

$\tilde{R}_{b,a}(\eta)$:

$$\begin{aligned}\tilde{R}_{b,a}(\eta) &= \sum_{x,y,\tau,w=0}^{N-1} \frac{W(x,y)}{v_N(w)}\,\delta_N(ax-by-\eta+\tau)e^{2\pi i\tau w/N} \\ &= \sum_{x,y,w=0}^{N-1} \frac{W(x,y)}{v_N(w)} e^{2\pi i(ax-by-\eta)w/N}\end{aligned}$$

Now we use the transform given in equation (3.11) to get:

$$\begin{aligned}&\frac{1}{N^2}\sum_{(b,a)\in\Theta_N}\sum_{x,y,w=0}^{N-1} \frac{W(x,y)}{v_N(w)}\, e^{2\pi i(a(x-n)-b(y-k))w/N} \\ =&\frac{1}{N^2}\sum_{(b,a)\in\Theta_N}\sum_{x,y,w=0}^{N-1} \frac{W(x,y)}{v_N(wa,-wb)}\, e^{2\pi i(wa(x-n)-wb(y-k))/N}\end{aligned} \tag{3.12}$$

In the last step we have used that $v_N(w) = v_N(wa,-wb)$, which holds for all $(b,a)\in\Theta_N$[8]. Now the sum over the (b,a) and w together can be viewed as a sum over the entire array where the point (x_0,y_0) is visited $v_N(x_0,y_0)$ times, by definition of the oversampling function. Therefore, we can perform the substitution:

$$\sum_{(b,a)\in\Theta_N}\sum_{w=0}^{N-1}\frac{f(wa,-wb)}{v_N(w)} = \sum_{x_0,y_0} f(x_0,y_0)$$

Using this, expression (3.12) can be written as:

$$\begin{aligned}&\frac{1}{N^2}\sum_{x,y,x_0,y_0=0}^{N-1} W(x,y)e^{2\pi i(x_0(x-n)+y_0(y-k))/N} \\ =&\sum_{x,y=0}^{N-1} W(x,y)\delta_N(x-n)\delta_N(y-k) = W(n,k)\end{aligned}$$

Thus we have indeed recovered the original function.

3.4 The Discrete Wigner Function from the IFRT

The definition of the discrete Wigner function goes naturally in terms of phase point operators $A(n,k)$ on our $N\times N$ grid. We thus define:

$$W(n,k) = \frac{1}{N}\mathrm{tr}\left(\hat{\rho}\hat{A}(n,k)\right) \tag{3.13}$$

[8]This doesn't mean that $(b,a)\in\Theta_N$ implies $(a,-b)\in\Theta_N$ too, just that the oversampling sequence is identical: $v_N(w) = v_N(wb,wa) = v_N(wa,-wb)$

Now that we have the FRT and its inverse, we can apply it to the phase space operators $\hat{A}(n,k)$. Analogous to the continuous case, we consider lines given by $an - bk = t$. The sum of $\hat{A}(n,k)$ over this line should be proportional to the projection operator corresponding to that line. Because $\hat{D}(b,a)$ is unitary and $\hat{D}(b,a)^N = \hat{D}(Nb,Na) = (-1)^{abN}I$, its eigenvalues are $(-1)^{abN}\omega^t$, for $t = 0,\ldots,N-1$ and with $\omega = \exp(2\pi i/N)$[9]. We let $|\psi_t^{b,a}\rangle$ denote the eigenvector of $\hat{D}(b,a)$ with eigenvalue $(-1)^{baN}\omega^t$. Therefore:

$$\hat{D}(b,a) = (-1)^{abN}\sum_{s=0}^{N-1}\omega^s|\psi_s^{b,a}\rangle\langle\psi_s^{b,a}|$$

and

$$\hat{D}(wb,wa) = (-1)^{abNw}\sum_{s=0}^{N-1}\omega^{ws}|\psi_s^{b,a}\rangle\langle\psi_s^{b,a}|$$

The above expression has the form of a DFT. This allows us to invert it and write the projection operators $|\psi_t^{b,a}\rangle\langle\psi_t^{b,a}|$ in terms of the displacement operator:

$$|\psi_t^{b,a}\rangle\langle\psi_t^{b,a}| = \frac{1}{N}\sum_{w=0}^{N-1}(-1)^{abNw}\omega^{-wt}\hat{D}(wb,wa) \tag{3.14}$$

We characterize our discrete Wigner function by the tomographic property and translation covariance. Mathematically these are the finite dimensional counterparts to equations (2.3) and (2.4) in chapter 2:

1. Tomographic property:

$$\hat{R}_{b,a}(0) = \frac{1}{N}\sum_{x,y=0}^{N-1}\hat{A}(x,y)\delta(ax-by) = |\psi_0^{b,a}\rangle\langle\psi_0^{b,a}|$$

2. Translation covariance:

$$\hat{A}(x,y) = \hat{D}(x,y)\hat{A}(0,0)\hat{D}^\dagger(x,y)$$

These can be combined into the single equation:

$$\hat{R}_{b,a}(t) = \frac{1}{N}\sum_{x,y=0}^{N-1}\hat{A}(x,y)\delta(ax-by-t) = |\psi_t^{b,a}\rangle\langle\psi_t^{b,a}| \tag{3.15}$$

[9]This is only true if $\gcd(b,a,N) = 1$, which holds for all the directions we will consider.

We only have to concentrate on $\hat{A}(0,0)$ because of the translation property. Applying (3.11):

$$\begin{aligned}
\hat{A}(0,0) &= \frac{1}{N^2} \sum_{(b,a)\in\Theta_N} \sum_{\tau=0}^{N-1} \hat{R}_{b,a}(-\tau) H(\tau) \\
&= \frac{1}{N} \sum_{(b,a)\in\Theta_N} \sum_{\tau,w=0}^{N-1} \frac{1}{v_N(w)} |\psi_{-\tau}^{b,a}\rangle\langle\psi_{-\tau}^{b,a}| \omega^{\tau w} && \text{(equation (3.11))} \\
&= \frac{1}{N^2} \sum_{(b,a)\in\Theta_N} \sum_{\tau,w,s=0}^{N-1} \frac{(-1)^{abNs}}{v_N(w)} \hat{D}(sb,sa) \omega^{\tau(w+s)} && \text{(equation (3.14))} \\
&= \frac{1}{N} \sum_{(b,a)\in\Theta_N} \sum_{w,s=0}^{N-1} \frac{(-1)^{abNs}}{v_N(w)} \hat{D}(sb,sa) \delta_N(w+s) && \text{(Sum over } \tau\text{)} \\
&= \frac{1}{N} \sum_{(b,a)\in\Theta_N} \sum_{s=0}^{N-1} \frac{(-1)^{abNs}}{v_N(-s)} \hat{D}(sb,sa) && \text{(Sum over } w\text{)}
\end{aligned}$$

Lastly, if use $v_N(-s) = v_N(s) = v_N(sb,sa)$ and $(-1)^{abNs} = (-1)^{(sb)(sa)N}$, we get the expression:

$$\hat{A}(0,0) = \frac{1}{N} \sum_{(b,a)\in\Theta_N} \sum_{s=0}^{N-1} \frac{(-1)^{(sb)(sa)N}}{v_N(sb,sa)} \hat{D}(sb,sa) \tag{3.16}$$

Now we would like, again, to change the sum over $(b,a) \in \Theta_N$ and s in the above equation to a sum over the entire grid where the point (x,y) is visited $v_N(x,y)$ times. If we do so, it would become:

$$\hat{A}(0,0) = \frac{1}{N} \sum_{x,y=0}^{N-1} (-1)^{xyN} \hat{D}(x,y) \tag{3.17}$$

This is quite a simple result. From the translation property we can find the expression for a general phase point operator:

$$\hat{A}(n,k) = \frac{1}{N} \sum_{x,y=0}^{N-1} (-1)^{xyN} \hat{D}(x,y) e^{2\pi i(kx-ny)/N}$$

It turns out that this definition works beautifully when N is odd. Unfortunately, when N is even this derivation is subtly flawed. Changing the sum over (b,a) and s in equation (3.16) into a sum over the grid is justified, as is done in subsection 3.3.1, when the function that is summed over is doubly periodic, with periods N, i.e. $f(x,y) = f(x+N,y) = f(x,y+N)$ for all x,y. Unfortunately, the displacement operator has period $2N$, since:

$$\hat{D}(x+\lambda N, y+\mu N) = (-1)^{y\lambda + x\mu + \lambda\mu N} \hat{D}(x,y)$$

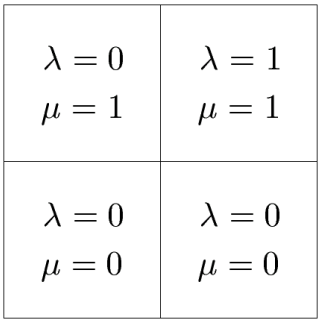


Figure 3.4: The $N \times N$ phase space extended to $2N \times 2N$. The grid is divided into four blocks of size $N \times N$. The summand in equation (3.16) generally depends on λ and μ.

As an example, we consider $N = 4$. We can take for our direction set:

$$\Theta_4 = \{(1,0),(1,1),(1,2),(1,3),(0,1),(2,1)$$

The point $(3,2)$ is visited once, by $(1,2) \in \Theta_N$ for $s = 3$, since $3(1,2) = (3,6)$, which equals the point $(3,2)$ on the phase space. However, we may not replace $\hat{D}(3,6)$ with $\hat{D}(3,2)$, because $\hat{D}(3,6) = -\hat{D}(3,2)$.
It is not clear how to deal effectively with this problem to obtain a closed form for $\hat{A}(0,0)$ for even N and further investigation is needed. For odd N we can show that equation 3.17 works beautifully. This we will do in the next section.

3.5 Investigating the discrete Wigner function

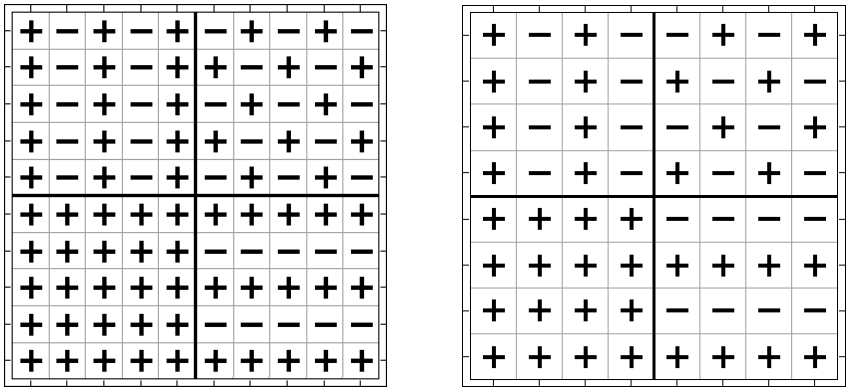
In this section we will investigate some properties of the Wigner function yielded from equation (3.16). We will have to take cases between even and odd dimensionality of the Hilbert space and these are treated in their relevant subsections.

We have seen that the sum over points in equation (3.16) extends beyond the $N \times N$ grid for certain directions $(b,a) \in \Theta_N$ and $s = 0, 1, \ldots N-1$ and is problematic since the displacement operator does not have period N. To investigate this issue, we rewrite equation (3.16) over a function where the arguments sb and sa are reduced modulo N. Therefore we define:

$$\alpha = sa - \lfloor \frac{sa}{N} \rfloor N, \qquad \mu = \lfloor \frac{sa}{N} \rfloor$$

$$\beta = sb - \lfloor \frac{sb}{N} \rfloor N, \qquad \lambda = \lfloor \frac{sb}{N} \rfloor$$

α and β are respectively sb and sa reduced modulo N. For μ and λ it is only important to know whether they are even or odd, so we can consider them modulo 2. They are 0 or 1 depending on which quadrant of the periodic $2N \times 2N$ grid you are in, see the schematic figure 3.4. With this we can



(a) N odd (b) N even

Figure 3.5: The structure of the signs for odd and even N. The points show the sign of $\hat{D}(b,a)$ relative to $\hat{D}(\beta,\alpha)$

rewrite equation (3.16) in the form:

$$\hat{A}(0,0) = \frac{1}{N} \sum_{(b,a)\in\Theta_N} \sum_{s=0}^{N-1} (-1)^{(\alpha\lambda+\beta\mu)(N-1)} (-1)^{\alpha\beta N} \frac{\hat{D}(\alpha,\beta)}{v_N(\alpha,\beta)} \tag{3.18}$$

We see that, when N is odd, the summand is independent of λ and μ and the step of changing to a sum over the grid leading to equation (3.17) is justified. For even N, however, we have to take into account the points at which a minus sign occurs.

3.5.1 When N is odd

For odd N, we can write $(-1)^{xyN} = (-1)^{xy}$. To see what the action of $\hat{A}(0,0)$ in equation (3.17) amounts to, we let it act on a computational basis ket $|n\rangle$:

$$\hat{A}(0,0)|n\rangle = \frac{1}{N} \sum_{x,y=0}^{N-1} (-1)^{xy} \hat{D}(x,y)|n\rangle = \frac{1}{N} \sum_{x,y=0}^{N-1} e^{\frac{2\pi i}{N}(n+x(N+1)/2)y} |n+x\rangle$$

When N is odd, $(N+1)/2$ is an integer and the above equation is a DFT:

$$\hat{A}(0,0)|n\rangle = \sum_{x=0}^{N-1} \delta_N \left(n + x\frac{N+1}{2} \right) |x+n\rangle$$

The solution to $n + \frac{1}{2}x(N+1) = 0$ exists and is unique when N is odd. The inverse of $\frac{1}{2}(N+1)$ modulo N is 2, since $2 \cdot \frac{1}{2}(N+1) = N+1 = 1 \mod N$. Therefore $n + \frac{1}{2}x(N+1) = 0$ is satisfied only for $x = -2n$. We find:

$$\hat{A}(0,0)|n\rangle = |-2n+n\rangle = |-n\rangle$$

So we can identify $\hat{A}(0,0)$ with the finite parity operator $\hat{P}$, similar to the continuous case. The general form of the phase point operators and the Wigner function thus becomes:

$$\hat{A}(n,k) = \hat{D}(n,k)\hat{P}\hat{D}^{\dagger}(n,k) = \hat{D}(2n,2k)\hat{P} \tag{3.19}$$

$$W(n,k) = \frac{1}{N}\mathrm{tr}\left(\hat{\rho}\hat{A}(n,k)\right) = \frac{1}{N}\sum_{y=0}^{N-1}\langle n-y|\hat{\rho}|n+y\rangle e^{4\pi iky/N} \tag{3.20}$$

As a check that the tomographic property is satisfied, we consider a direction vector $(b,a) \in \Theta_N$ and sum over the corresponding line through the origin. In the second step we insert the closure relation $\sum_t |\psi_t^{b,a}\rangle\langle\psi_t^{b,a}| = 1$.

$$\frac{1}{N}\sum_{r=0}^{N-1}\hat{A}(rb,ra) = \frac{1}{N}\sum_{r=0}^{N-1}\hat{D}(2rb,2ra)\cdot 1\cdot\hat{P}$$

$$= \frac{1}{N}\sum_{r,t=0}^{N-1}\omega^{2rt}|\psi_t^{b,a}\rangle\langle\psi_{-t}^{b,a}| = \sum_{t=0}^{N-1}\delta_N(2t)|\psi_t^{b,a}\rangle\langle\psi_{-t}^{b,a}|$$

which yields $|\psi_0^{b,a}\rangle\langle\psi_0^{b,a}|$ as desired.

A direct consequence of the tomographic property is that the Wigner function is normalized. This follows from:

$$\frac{1}{N}\sum_{n,k=0}^{N-1}\hat{A}(n,k) = \sum_{n=0}^{N-1}|n\rangle\langle n| = 1 \tag{3.21}$$

$$\sum_{n,k=0}^{N-1}W(n,k) = \sum_{n,k=0}^{N-1}\frac{1}{N}\mathrm{tr}\left(\hat{\rho}\hat{A}(n,k)\right) = \mathrm{tr}\left(\hat{\rho}\right) = 1$$

Since $\hat{A}(0,0)$ is Hermitian, the Wigner function is real.

Another important property of the continuous Wigner function that we have not imposed explicitly, but is nevertheless satisfied, is the overlap property, which follows from the orthogonality of the phase point operators. In part 3.5.4 we will show that this follows generally from the tomographic and translational properties, but it is easy to verify directly for odd N from equation (3.19).

The Wigner function obtained here is exactly the same as posed before by Wootters [Woo87], intended for (odd) prime powers, which was noted to also work satisfactorily for all odd N. It was, however, unclear what requirements were necessary or sufficient for tomographic purposes. With the IFRT approach presented here, we hope to have elucidated in a geometrical way that equation (3.19) is the *unique* form of the Wigner function obeying translation covariance and the tomographic property.

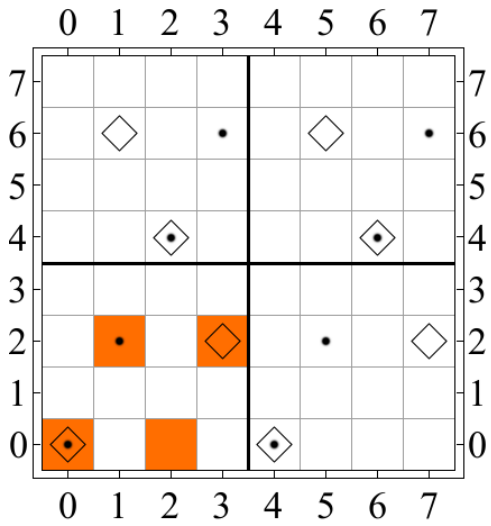


Figure 3.6: The extended phase space for $N = 4$. There are two inequivalent directions yielding the same line $2x - y = 0 \mod 4$, but different lines modulo 8. The line with the dots is generated by $(1, 2)$ and $(-1, -2)$, the line with the diamonds by $(1, 3)$ and $(1, -2)$.

3.5.2 When N is even

For even N we also would like to reduce equation (3.16), but we have seen from equation (3.18) that the summand depends on how (sb, sa) extends into the region outside the $N \times N$ grid. In fact, $\hat{A}(0, 0)$ depends on the set Θ_N as we can generally choose several direction vectors which generate the same line in phase space, but extend differently when we consider the extended $2N \times 2N$ phase space. In other words, they generate the same lines modulo N, but different lines modulo $2N$.
To illustrate with the case $N = 4$, the line with direction $(1, 2)$ can also be generated by any of the direction vectors $(-1, 2), (1, -2)$ and $(-1, -2)$, see figure 3.6.
The choice of $(1, 2)$ or $(-1, -2)$ yields an extra minus sign for the point $(3, 2)$, but choosing $(-1, 2)$ or $(1, -2)$ yields a minus sign for the point $(1, 2)$. A similar choice occurs for the line with direction $(2, 1)$. Therefore, there are four inequivalent choices which lead to four (slightly) different discrete Wigner functions.
It's hard to miss the observation that, for $N = 4$, the places where extra minuses can occur are precisely the places that are missed when we only consider lines through the origin where both b and a are units. Whether this is generally true has not been investigated yet.

To verify that the tomographic property is obeyed, we should be able to

work directly from equation (3.16), but we did not get a definite result from that yet. Numerical calculations have shown, however, that equation (3.16) gives a discrete Wigner function that obeys the tomographic property with orthogonal phase point operators for $N = 2^2, 2^3$ and 2^4. These properties both fail, however, for $N = 6, 10$ and 12. This suggests that the tomographic property, as we have formulated, can not be fulfilled when N is an even number that is not a power of 2. Furthermore, it does seem to hold for powers of 2, which includes the important case of a system of N qubits.

As an example of a discrete Wigner function for $N = 4$, we take the direction set $\Theta_4 = \{(1,0),(1,1),(1,2),(1,-1),(0,1),(2,1)\}$. The resulting phase operator $\hat{A}(0,0)$ is then:

$$\hat{A}(0,0) = \begin{pmatrix} 1 & \frac{1}{2\sqrt{2}}(\gamma^*+1) & -\frac{i}{2} & \frac{1}{2\sqrt{2}}(\gamma^*+1) \\ \frac{1}{2\sqrt{2}}(\gamma+1) & 0 & \frac{1}{2\sqrt{2}}(\gamma-1) & \frac{1}{2} \\ \frac{i}{2} & \frac{1}{2\sqrt{2}}(\gamma^*-1) & 0 & \frac{1}{2\sqrt{2}}(\gamma^*-1) \\ \frac{1}{2\sqrt{2}}(\gamma+1) & \frac{1}{2} & \frac{1}{2\sqrt{2}}(\gamma-1) & 0 \end{pmatrix},$$

where $\gamma = \frac{1}{\sqrt{2}}(1+i) = \exp(\pi i/4)$.
Together with $\hat{A}(n,k) = \hat{D}(n,k)\hat{A}(0,0)\hat{D}^\dagger(n,k)$ this defines the whole Wigner function $W(n,k) = \frac{1}{2}\mathrm{tr}\left(\hat{\rho}\hat{A}(n,k)\right)$. The distribution of the 'discrete quadrature states' $|\psi_t^{b,a}\rangle$ are as we expected; The Wigner function is zero, except on the line $ax - by = t$, where it takes on the constant value $1/N = 1/4$. Note, however, that for the state $|\psi_0^{3,2}\rangle$ we get the distribution we would expect for the line $2x - y = 2$ and *not* the line $2x - 3y = 0$. The reason for this is that we did not choose the direction $(3,2)$ as part of Θ_4. If we choose $(3,2) \in \Theta_4$ instead of $(1,2)$, we would get the expected distribution for $|\psi_0^{3,2}\rangle$ and our definition for $\hat{A}(0,0)$ would be different. Since the density matrix can be reconstructed from the Wigner function, they must contain the same information, and it is straightforward to check that $|\psi_0^{3,2}\rangle\langle\psi_0^{3,2}| = |\psi_2^{1,2}\rangle\langle\psi_2^{1,2}|$. We have thus gained important geometrical insights in the reason why the Wigner function is not unique, how much choice we have and how they are related.

3.5.3 A link with mutually unbiased bases

In this part, we go back to a discussion concerning general N. In subsection 3.2.1 we mentioned the concept of unbiased bases. It has been noted there is a natural link between the lines in discrete phase space and mutually unbiased bases [B+07, GHW04]. Here we show some insight in the overlap property and how it is linked with the unbiasedness of two bases.

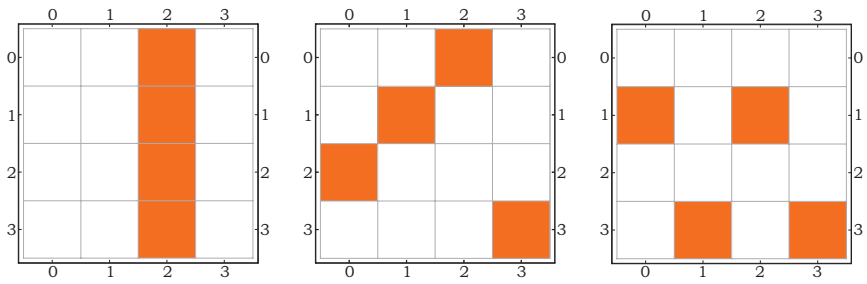


Figure 3.7: Discrete Wigner function of the quadrature states $|\psi_2^{0,1}\rangle, |\psi_3^{1,1}\rangle$ and $|\psi_2^{1,2}\rangle$.

We know that every direction $(b, a) \in \Theta_N$ is associated with a basis $\{|\psi_t^{b,a}\rangle\}$. Given two directions, how 'different' are the two? This is an important question in quantum state reconstruction [WF89, AB01]. The overlap property (equation (3.23)) shows that the more the two Wigner functions of two states $|\psi\rangle, |\phi\rangle$ overlap, the greater their fidelity $|\langle\phi|\psi\rangle|^2$. We expect, although we have not proven so yet, that the 'direction-' or 'finite quadrature' states $|\psi_t^{b,a}\rangle$ have a Wigner function that is zero everywhere, except on the line $ax - by = t$, where they take on the value $1/N$. From equation (3.23), we see that this means that the fidelity between two quadrature states is equal to $1/N$ times the number of intersection points. When every line with one direction intersects every line in another direction precisely once, the corresponding bases are unbiased. This is true for every direction in our phase space only when N is a prime number, which gives a direct way to create $N + 1$ mutually unbiased bases. We simply need the eigenvectors of $\hat{D}(1, a)$ for $a = 0, \ldots, N - 1$ and $\hat{D}(0, 1)$, the last giving the computational basis.

We now show that the Wigner function of the finite quadrature states $|\psi_t^{b,a}\rangle$ have the expected Wigner function from the tomographic property and translation covariance. Consider the relation:

$$\begin{aligned}\langle\psi_t^{b,a}|\hat{A}(n,k)|\psi_t^{b,a}\rangle &= \langle\psi_t^{b,a}|\hat{D}(rb,ra)\hat{A}(x,k)\hat{D}^\dagger(rb,ra)|\psi_t^{b,a}\rangle \\ &= \langle\psi_t^{b,a}|\hat{A}(n+rb,k+ra)|\psi_t^{b,a}\rangle\end{aligned}$$

Since $W_{|\psi_t^{b,a}\rangle}(n,k) = \frac{1}{N}\langle\psi_t^{b,a}|\hat{A}(n,k)|\psi_t^{b,a}\rangle$ is the Wigner function for a pure quadrature state, the above is equivalent to:

$$W_{|\psi_t^{b,a}\rangle}(n,k) = W_{|\psi_t^{b,a}\rangle}(n+rb,k+ra), \qquad r = 0, \ldots, N-1$$

So the Wigner function has the same value along every line with direction (b, a). The tomographic property, with normalization, then immediately tells us that:

$$W_{|\psi_t^{b,a}\rangle}(n,k) = \frac{1}{N}\,\delta_N(ax - by - t)$$

As a consequence of the above, we can calculate how 'unbiased' two quadrature bases are.

$$\begin{aligned}|\langle\psi_t^{b,a}|\psi_s^{d,c}\rangle|^2 &= \frac{1}{N}\sum_{x,y=0}^{N-1} \mathrm{tr}\left(\hat{A}(x,y)|\psi_s^{d,c}\rangle\right)\delta_N(ax-by-t)\\ &= \sum_{x,y=0}^{N-1} W_{|\psi_s^{d,c}\rangle}\delta_N(ax-by-t)\\ &= \frac{1}{N}\sum_{x,y=0}^{N-1}\delta_N(cx-dy-s)\delta_N(ax-by-t)\end{aligned}$$

Only the points (x,y) that satisfy both $ax-by=t$ and $cx-dy=s$ contribute in the above sum. Therefore, the fidelity of $|\psi_t^{b,a}\rangle$ and $|\psi_s^{d,c}\rangle$ is equal to $1/N$ times the number of intersection points of their associated lines in phase space. When N is a prime number, any two lines from different parallel classes intersect exactly one. Therefore each of the $N+1$ quadrature bases are unbiased in this case.

3.5.4 The Overlap property

In the previous part we showed that $\langle\psi_t^{b,a}|\hat{A}(n,k)|\psi_t^{b,a}\rangle = \delta_N(ax-by-t)$. We know show that the tomographic property with translation covariance imply that the phase point operators form an orthogonal set. This is quite easy from the linearity of the Radon transform. The main observation is that $\hat{A}(n,k)$ is the inverse Radon transform of $N|\psi_t^{b,a}\rangle\langle\psi_t^{b,a}|$ and the trace $\mathrm{tr}\left(\hat{A}(0,0)\hat{A}(n,k)\right)$, which depends linearly on $\hat{A}(n,k)$, must therefore be the inverse Radon transform of

$$N\mathrm{tr}\left(\hat{A}(0,0)|\psi_t^{b,a}\rangle\langle\psi_t^{b,a}|\right) = N^2 W_{|\psi_t^{b,a}\rangle}(0,0) = N\delta_N(t)$$

The proof is given in more detail below.

First we note that the trace $\mathrm{tr}\left(\hat{A}(n_1,k_1)\hat{A}(n_2,k_2)\right)$ depends only on the relative coordinates (x_2-x_1, k_2-k_1), since:

$$\begin{aligned}\mathrm{tr}\left(\hat{A}(n_1,k_1)\hat{A}(n_2,k_2)\right) &= \mathrm{tr}\left(\hat{D}(n_1,k_1)\hat{A}(0,0)\hat{D}^\dagger(n_1,k_1)\hat{A}(n_2,k_2)\right)\\ &= \mathrm{tr}\left(\hat{A}(0,0)\hat{A}(n_2-n_1,k_2-k_1)\right)\end{aligned}$$

Therefore we consider the function:

$$f(n,k) = \mathrm{tr}\left(\hat{A}(0,0)\hat{A}(n,k)\right)$$

and our goal is to find this function. Note that $f(n,k)$ is periodic on our $N \times N$ phase space, because $\hat{A}(n,k)$ is. The main step is to simply take its finite Radon transform:

$$\begin{aligned}\mathcal{R}^f_{b,a}(t) &= \sum_{x,y=0}^{N-1} f(n,k)\delta_N(ax-by-t) \\ &= \mathrm{tr}\left(\hat{A}(0,0)\left[\sum_{x,y=0}^{N-1}\hat{A}(x,y)\delta_N(ax-by-t)\right]\right) \\ &= N\mathrm{tr}\left(\hat{A}(0,0)|\psi^{b,a}_t\rangle\langle\psi^{b,a}_t|\right)\end{aligned}$$

The last expression is just equal to $N\delta_N(t)$, as we showed in subsection 3.5.3. It is not hard to imagine what function has a Radon transform which is zero everywhere, except for lines through the origin ($t = 0$), where it equals the constant N. We can simply take $f(n,k) = N\delta_N(n)\delta_N(k)$ and indeed, if we invert $\mathcal{R}^f_{b,a}(t) = N\delta_N(t)$ with the IFRT, we find:

$$\mathrm{tr}\left(\hat{A}(0,0)\hat{A}(n,k)\right) = N\delta_N(n)\delta_N(k)$$

and by translation we get the general result:

$$\mathrm{tr}\left(\hat{A}(n,k)\hat{A}(m,l)\right) = N\delta_N(n-m)\delta_N(k-l) \tag{3.22}$$

We conclude that the tomographic property and translation covariance imply the orthogonality of the phase point operators.

Because the phase operators form an orthogonal basis, from the definition (3.13) it follows that the Wigner function values are the coefficients of the expansion of the density matrix:

$$\hat{\rho} = \frac{1}{N}\sum_{n,k=0}^{N-1}\mathrm{tr}\left(\hat{\rho}\hat{A}(n,k)\right)\hat{A}(n,k) = \sum_{n,k=0}^{N-1} W(n,k)\hat{A}(n,k)$$

Multiplying by $\hat{\rho}'$ and taking the trace in the above expression readily yields the overlap property:

$$\mathrm{tr}\left(\hat{\rho}\hat{\rho}'\right) = N\sum_{n,k=0}^{N-1} W(n,k)W'(n,k) \tag{3.23}$$

Thus we can transform any observable $\hat{O}$ to a phase space function:

$$O(n,k) = \mathrm{tr}\left(\hat{O}\hat{A}(n,k)\right)$$

and the expectation value of $\hat{O}$ can be calculated from the classical relation:

$$\langle \hat{O} \rangle = \sum_{n,k=0}^{N-1} W(n,k) O(n,k),$$

The overlap property allows a straightforward phase space method of calculating the *fidelity* of two quantum states.

3.6 Summary and discussion

There is a large amount of literature available on finite quasi-probability distributions. It has long ago been recognized by Wootters [Woo87] that, if N is a prime number, there is a Wigner function which reflects naturally all the properties of the continuous case. For odd N, a discrete Wigner function was given by Cohendet *et al* [CCSSC88]. For even dimensions (greater than two), Leonhardt [Leo95, Leo96] considered enlarging the phase space to $2N \times 2N$, which has the drawback of introducing redundancy and 'ghost variables'. In this chapter we have approached the problem of finding a finite discrete Wigner from a novel first principles approach using the inverse finite Radon transform developed by Kingston and Svalbe [KS07]. From requiring two fundamental properties (the tomographic one and translation covariance) we have arrived at an expression (from eqn 3.16) which the phase point operators must satisfy. There were consequently two main cases we needed to consider separately, depending on the dimensionality N of the Hilbert space:

1. **When N is odd**
   We have found the *unique* expression for the phase point operators such that the Wigner function satisfies the tomographic property and translation covariance is $\hat{A}(n,k) = \hat{D}(n,k)\hat{P}\hat{D}^{\dagger}(n,k)$. Even though the tomographic was imposed by requiring that the sum of $W(n,k)$ over the line $an - bk = t \mod N$ should equal $\langle \psi_t^{b,a} | \hat{\rho} | \psi_t^{b,a} \rangle$ for all $(b,a) \in \Theta_N$, it actually holds for all directions (b,a) for which $\gcd(a,b,N) = 1$. We can therefore do without the explicit mention of Θ_N.

2. **When N is even**
   For even N we consider again two cases: N is a power of two, or N is even, but not a power of two.

   (a) **When $N = 2^n$**
   We have found a definition for a discrete Wigner function obeying the tomographic property and translation covariance and shown an expression for $\hat{A}(0,0)$ explicitly for $N = 4$. The definition

for $N = 2^n$ is not unique and depends on the set Θ_N, which thus needs explicit mention. We suspect that the existence (and non-uniqueness) of this discrete Wigner function is guaranteed, although we have not proved so. Numerical calculations have shown the tomographic property and translation covariance are obeyed for $N = 2, 2^2, 2^3, 2^4$ for certain Θ_N.

(b) **When N is even and not a prime power**
For this last case we suspect no discrete Wigner function exists satisfying the tomographic property and translation covariance. Numerical calculations have indicated this for $N = 6, 10$ and 12.

There are still some open questions concerning the discrete Wigner function for even N. Natural topics for further research would be a (possible) proof that the described Wigner function exists for $N = 2^n$, but not for other even N. In the case where $N = 2^n$, where non-uniqueness is an issue, is there a natural choice for Θ_N, gotten by imposing more structure? For example, if we consider the change of basis from $\{|\widetilde{k}\rangle\}$ to $\{|\psi_t^{1,1}\rangle\}$, this can be seen as a 'rotation about $\pi/4$' in the continuous case. Are similar results available in the discrete case? How does a Fractional fourier transform act and how are these related to the transforms between the quadrate bases?
Another question, coming back to the main point of this thesis, is whether we can really use this description to gain insight into quantum processes. What do various quantum algorithms look like in discrete phase space?[10]

Although it seems that, unfortunately, there are fundamental difficulties in defining a Wigner function for arbitrary N. We have found one which (presumably) works for $N = 2^n$, which includes the important case of a system of N qubits. We therefore hope that the Wigner function presented here is of use in illuminating the workings of N-qubit systems and an aid in visualizing quantum algorithms.

Bibliography

[AB01] Roberth Asplund and Gunnar Björk. Reconstructing the discrete Wigner function and some properties of the measurement bases. *Phys. Rev. A*, 64(1):012106, Jun 2001.

[B+07] G. Björk et al. Mutually unbiased bases and discrete Wigner functions. *J. Opt. Soc. Am. B*, 24(2):371–378, February 2007.

[10] A nice point to mention: The front cover of this thesis shows the discrete Wigner function of a state after a few iterations of Grover's search algorithm [Gro96] for $N = 31$. The Fourier-like interference pattern is unmistakable.

[CCSSC88] O Cohendet, P Combe, M Sirugue, and M Sirugue-Collin. A stochastic treatment of the dynamics of an integer spin. *Journal of Physics A: Mathematical and General*, 21(13):2875, 1988.

[GHW04] K. S. Gibbons, M.J. Hoffman, and W.K. Wootters. Discrete phase space based on finite fields. *Phys. Rev. A*, 70(062101), 2004.

[Gro96] Lov K. Grover. A fast quantum mechanical algorithm for database search. In *Proceedings of the twenty-eighth annual ACM symposium on Theory of computing*, STOC '96, pages 212–219, New York, NY, USA, 1996. ACM.

[HB80] J. H. Hannay and M. V. Berry. Quantization of linear maps on a torus-fresnel diffraction by a periodic grating. *Physica D: Nonlinear Phenomena*, 1(3):267 – 290, 1980.

[HS96] D. Hsung, T. Lun and W. Siu. The discrete periodic radon transform. *IEEE Transactions on Signal Processing*, 44(10):2651–2657, 1996.

[KS07] Andrew Kingston and Imants Svalbe. Generalised finite radon transform for n x n images. *Image Vision Comput.*, 25(10):1620–1630, 2007.

[Leo95] Ulf Leonhardt. Quantum-state tomography and discrete Wigner function. *Phys. Rev. Lett.*, 74(21):4101–4105, May 1995.

[Leo96] Ulf Leonhardt. Discrete Wigner function and quantum-state tomography. *Phys. Rev. A*, 53(5):2998–3013, May 1996.

[MF93] F. Matús and J. Flusser. Image representation via a finite radon transform. *IEEE Trans. Pattern Anal. Mach. Intell.*, 15(10):996–1006, 1993.

[MPS02a] César Miquel, Juan Pablo Paz, and Marcos Saraceno. Quantum computers in phase space. *Phys. Rev. A*, 65(6):062309, Jun 2002.

[MPS02b] César Miquel, Juan Pablo Paz, and Marcos Saraceno. Quantum computers in phase space. *Phys. Rev. A*, 65(6):062309, Jun 2002.

[WF89] William K Wootters and Brian D Fields. Optimal state-determination by mutually unbiased measurements. *Annals of Physics*, 191(2):363 – 381, 1989.

[Woo87] W.K. Wootters. A Wigner-function formulation of finite-state quantum mechanics. *Ann. Phys. (N.Y.)*, 176:1–21, 1987.

Chapter 4

Rotationally Covariant Representations

In the previous chapter we introduced a visualization of a finite dimensional quantum state by a generalization of the Wigner function to the discrete case. By translational covariance and unitarity of the translation operators, it was natural to impose the periodicity of our phase space. The topological structure was thus that of a discrete torus. When we consider what finite-dimensional systems naturally occur in nature, the *spin* of a particle quickly comes to mind. Although we could use the Wigner function to describe spins, it may be somewhat awkward to do so, because it is hard to imagine a physical meaning to translating a Wigner function when it describes a spin j particle. We can only mathematically identify the $N = 2j+1$ states $|j, m\rangle$ with $|n\rangle$. We would therefore like a (visual) representation of a spin state that is not translation covariant, but rotationally covariant. In other words: the representation should be well behaved under the rotation group $SO(3)$. This is the topic of this chapter.

A well-known way to visualize a spin $1/2$ particle state is by a point on a unit sphere: the Bloch sphere. This representation is rotationally covariant and is used in almost every introductory presentation on qubits. A summary of this representation is given in section 4.1. We then generalize this to a representation for arbitrary spin j particles and arrive at the little-known Majorana representation[1]. Some properties of this representation are then investigated and a link is made with the Schwinger boson representation. Lastly, we try to generalize the Majorana representation even further to include N spin j particles. The difficulty that enters is that the value of the

[1] This is unrelated to the Majorana representation of Gamma matrices in quantum field theory.

spin is no longer fixed and we try to separate the subspaces belonging to different spin values with the Schur transform.

4.1 The Bloch sphere

The conventional way to describe a pure quantum state is with a normalized state vector $|\psi\rangle$ in a Hilbert space $\mathcal{H}$. Complete specification of a state for calculational purposes is ordinarily done by choosing an orthonormal basis for $\mathcal{H}$ (a convenient one for the context at hand) and specifying the components of $|\psi\rangle$ with respect to this basis. This is a rather abstract approach and important properties may be hard to read off from this representation. Also, the representation by components is not unique due to the requirement that the state be normalized and that the global phase has no physical significance. Furthermore, there may be additional structure in the particular context of the problem (i.e. rotational covariance) which the component representation fails to point out or take advantage of. We strive to find a more intuitive way to think about spin.

In representing spin angular momentum, we often use the eigenstates of the square of the total spin S^2 and the z-component of the spin S_z as a basis. For a system of given spin j, where j can only take on a nonnegative integer or half-integer value, the eigenvalues of S_z (denoted by m) can take on the $2j+1$ values from $-j$ to j in integer steps. The basis is then denoted by $\{|j,m\rangle\}_{m=-j}^{j}$. For a spin 1/2 particle, an arbitrary state can be written:

$$|\psi\rangle = \alpha|\frac{1}{2},\frac{1}{2}\rangle + \beta|\frac{1}{2},-\frac{1}{2}\rangle$$

Where $|\alpha|^2 + |\beta|^2 = 1$ and only the phase difference between α and β is of physical significance.
The normalization and insignificance of the global phase means that this representation has the drawback that it contains respectively redundant and useless information and it breaks the natural rotational symmetry by singling out the z-direction as special.
There is a very convenient representation for a spin 1/2 particle which takes advantage of the rotational covariance for spin systems. Perhaps a simpler way to put rotational covariance into words is that nature doesn't care which direction in space we call the z-axis. To derive this representation, let's analyze the spin vector operator $\hat{\vec{S}} = \hat{S}_x\vec{e}_x + \hat{S}_y\vec{e}_y + \hat{S}_z\vec{e}_z$. We can only measure one component of the spin at once, so let's choose a general direction given by the unit vector $\vec{r} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$, where (θ,ϕ) are the polar and azimuthal angles of the point on the unit sphere given by $\vec{r}$. A straightforward calculation yields:

$$\vec{r}\cdot\vec{S} = \frac{1}{2}\begin{pmatrix} \cos\theta & \sin\theta e^{-i\phi} \\ \sin\theta e^{i\phi} & -\cos\theta \end{pmatrix}$$

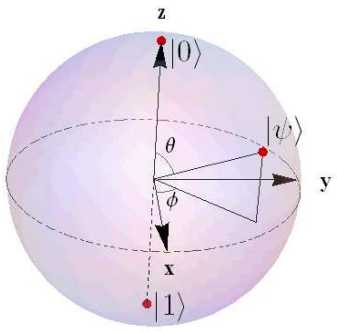


Figure 4.1: The qubit state $|\psi\rangle = \cos\frac{\theta}{2}|0\rangle + \sin\frac{\theta}{2}e^{i\phi}|1\rangle$ represented on the Bloch sphere. The basis states $|0\rangle$ and $|1\rangle$ are situated at the north and south pole respectively.

The eigenvalues are $\pm 1/2$ and the corresponding eigenvectors are:

$$|\vec{r}, +\frac{1}{2}\rangle = \begin{pmatrix} \cos\frac{\theta}{2} \\ \sin\frac{\theta}{2}e^{i\phi} \end{pmatrix} \qquad |\vec{r}, -\frac{1}{2}\rangle = \begin{pmatrix} \sin\frac{\theta}{2} \\ -\cos\frac{\theta}{2}e^{i\phi} \end{pmatrix} = \begin{pmatrix} \cos(\frac{\pi-\theta}{2}) \\ \sin(\frac{\pi-\theta}{2})e^{i(\phi+\pi)} \end{pmatrix}$$

As can be seen from the components of the above expression, the form of $|\vec{n}, \frac{1}{2}\rangle$ is that of the most general spin $1/2$ state. Therefore, every spin $1/2$ state can be represented uniquely by a point (θ, ϕ) on a unit sphere, called the *Bloch sphere*[2] This state is then the spin-up state for the direction it is pointing and the antipodal point represents the spin-down state for that direction.

To keep in line with the notation used in information theory, we will identify $|\frac{1}{2}, \frac{1}{2}\rangle$ with $|0\rangle$ and $|\frac{1}{2}, -\frac{1}{2}\rangle$ with $|1\rangle$.
The Bloch Sphere representation is widely used as a way to visualize a qubit state. Although any two-level system can be treated as a qubit and thus be represented on the Bloch sphere, the rotational covariant aspects in this Bloch Sphere for systems other than spin $1/2$ system may not have much meaning. Nevertheless, treating a non-spin system as a system with fictitious spin, allowing us to use the angular momentum formalism and representations is sometimes insightful [HN04, Mal04, GK06].

[2] In (quantum) optics, the sphere is used to specify the polarization state of a plane wave and is called the Poincaré Sphere. The north- and south poles represent right- and left circular polarization respectively and the equator forms the set of linearly polarized states.

4.2 Generalizing to higher spin systems: The Majorana Representation

In this section we will look for a generalization of the Bloch sphere representation which is valid for spin values other than 1/2. The end result will be that we can represent a pure spin j state uniquely by $2j$ points on the unit sphere. Rotation the system about an axis rotates the entire sphere with its points correspondingly. This representation is called the Majorana representation [Maj32]. We will construct the representation and motivate our steps as we go along.

If we wish to generalize the Bloch sphere to higher spin values, it helps to first consider the degrees of freedom of our system. In a Hilbert space of dimension N, a componentwise representation of an arbitrary state vector is given by N complex numbers, or $2N$ real numbers. Normalization and arbitrary global phase give two independent constraints, bringing the required number of real parameters required to uniquely specify a state to $2N-2$. For a system with given spin j the Hilbert space dimension is $2j+1$. Thus there are $4j$ degrees of freedom in this case. If we are to use points on a sphere to represent our state it is clear that we need $2j$ points since each point requires 2 degrees of freedom for its specification. To do this, we will first try to find a way to construct the qubit state $|\psi\rangle$ from a point on the Bloch sphere in a way that respects rotational covariance. Then we will generalize the procedure to multiple points on the sphere.

In space, rotations are given by elements of the rotation group $SO(3)$. The rotation of a spin j state about an axis given by the unit vector $\vec{n}$ by angle α, which we combine into the vector $\alpha\vec{n} = \vec{\alpha}$, is represented by the action of the unitary operator $U(\vec{\alpha}) = \exp(-i\vec{\alpha}\cdot\vec{J})$. Thus the correspondence between a physical rotation in 3-space and the corresponding rotation in the state space is:

$$R(\vec{\alpha}) \in SO(3) \leftrightarrow \hat{U}(\vec{\alpha}) = \exp(-i\vec{\alpha}\cdot\vec{J}) \in \mathcal{H} \tag{4.1}$$

The condition for rotational covariance is that it shouldn't matter whether we rotate the sphere and all its points by $R(\vec{\alpha})$ or act on the state with $\hat{U}(\vec{\alpha})$, in the end they should both represent the same state (although a global phase may be introduced). So if we represent the state given by $2j$ points $\vec{r}_1,\ldots,\vec{r}_{2j}$ as $|\vec{r}_1,\ldots,\vec{r}_{2j}\rangle$, then we must have:

$$|R\vec{r}_1,\ldots,R\vec{r}_{2j}\rangle = \hat{U}|\vec{r}_1,\ldots,\vec{r}_{2j}\rangle \tag{4.2}$$

To ensure this is true, we should try and find $|\psi\rangle$ from inner product expressions like $\langle\phi|\vec{r}_1,\ldots,\vec{r}_{2j}\rangle$ which are unchanged by unitary transformations. Taking example from the Bloch Sphere, we can find the point on the Bloch Sphere corresponding to the state $|\psi\rangle$ by solving $\langle\vec{s}|\psi\rangle = 0$ for $\vec{s}$. The point

on the Bloch sphere is exactly the one antipodal to $\vec{s}$: $|\psi\rangle = |-\vec{s}\rangle$. We can generalize this procedure we can find a representation of a spin j state that is represented by a *single* point on the sphere and which transforms covariantly under rotations. The obvious candidates are the spin coherent states (see Appendix D). Since the set of spin coherent states (SCS) allow a straightforward generalization of the Bloch sphere, we generalize our procedure by solving $\langle\vec{s}|\psi\rangle = 0$ for $\vec{s}$. We will see this yields $2j$ solutions $\vec{s}_i$ in general. The antipodal points of these solutions will be the points on the sphere representing $|\psi\rangle$.

For calculational purposes, it is often useful to represent the points on the sphere in the complex plane by means of stereographic projection to the extended complex plane $\mathbb{C} \cup \{\infty\}$ (Appendix D[3]). A point $\vec{r}$ on the sphere with coordinates (θ, ϕ) will be projected to the complex number $w = \tan\frac{\theta}{2}e^{i\phi}$ and its antipodal point is projected to $-1/w^*$. The expansion of a spin coherent state $|z\rangle$ in the angular momentum basis is:

$$|z\rangle = \frac{1}{(1+|z|^2)^j} \sum_{m=-j}^{j} \binom{2j}{j+m}^{1/2} z^{j-m}|j,m\rangle \tag{4.3}$$

Forming the inner product of $|\psi\rangle = \sum a_m|j,m\rangle$ with $|z\rangle$, we find:

$$\langle z|\psi\rangle = \frac{1}{(1+|z|^2)^j} \sum_{m=-j}^{j} \binom{2j}{j+m}^{1/2} a_m z^{*j-m}$$

This is a normalization factor times a polynomial of maximum degree $2j$ and has $2j$ zero's as expected. It can happen that the degree of the polynomial is lower because the highest indexed coefficients a_m can be zero. In that case, the zero's lie at infinity and these are mapped to the south pole under inverse stereographic projection. This translates to points on the north pole after taking the antipodal points of the zero's.

We are now practically done with finding a generalization of the Bloch sphere. We are left to show this representation is rotationally covariant and this is quite simple. If $\vec{r}$ is a point on the sphere for the state $|\psi\rangle$, then $\langle-\vec{r}|\psi\rangle = 0$, by construction. If $\hat{U}|\psi\rangle$ is the state $|\psi\rangle$ after rotation, then $R\vec{r}$ is a point on the sphere corresponding to this rotated state, since $\langle-R\vec{r}|\hat{U}|\psi\rangle = \langle-\vec{r}|\hat{U}^\dagger\hat{U}|\psi\rangle = \langle-\vec{r}|\psi\rangle = 0$. Therefore the points on the sphere move covariantly under a rotation.

To summarize, we can uniquely represent a spin j state by $2j$ points on a sphere in a rotationally covariant way, this is called the Majorana Representation [Maj32]. The correspondence between the component representation and the $2j$ points on the sphere is:

[3] As explained in the appendix, we use projection from the south pole.

1. Form $\psi(z) = \langle z|\psi\rangle$, this will give the following polynomial with $2j$ zero's in the extended complex plane[4]:

$$\psi(z) = \frac{1}{(1+|z|^2)^j} \sum_{m=-j}^{j} \binom{2j}{j+m}^{1/2} (z^*)^{j-m} a_m$$

2. Use (inverse) stereographic projection from the South Pole on the zero's to make $2j$ points on the sphere.

3. The $2j$ antipodal points give you the $2j$ points in the Majorana representation.

To find the component representation from the Majorana representation we reverse the process. We present a slightly different, but equivalent way, which is easier in practice:

1. Project the $2j$ points (*not* the antipodal ones) from the sphere to the complex plane. Name the projected points, say, $w_1, \ldots, w_{2j}$.

2. Form the polynomial:

$$P(z) = \prod_{i=1}^{2j}(1+zw_i) = \sum_{m=-j}^{j} c_m z^{j-m}$$

   to find the coefficients $c_{-j}, \ldots, c_j$.

3. Replace z^{j-m} with $\binom{2j}{j+m}^{-\frac{1}{2}}|j,m\rangle$. The state is:

$$|\psi\rangle = C \sum_{m=-j}^{j} \binom{2j}{j+m}^{-1/2} c_m |j,m\rangle,$$

   where C is a normalization constant.

4.2.1 Working with the Majorana representation

At first sight, the correspondence between the Majorana and component representations appears pretty complex. To get a feel for it we will work out some examples of useful and interesting states.

[4] In practice it is usually more convenient to find the zero's immediately form the *Majorana Polynomial*: $P_\psi = \sum_{m=-j}^{j} \binom{2j}{j+m}^{1/2} a_m z^{j-m}$

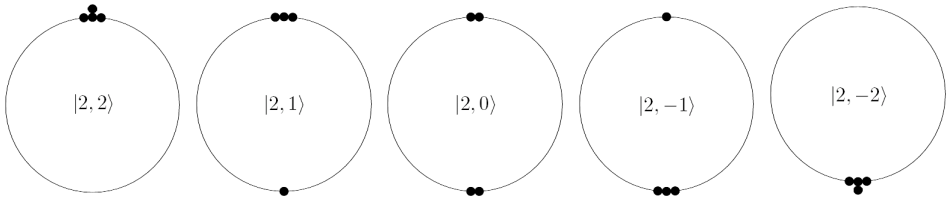


Figure 4.2: Schematic of the angular momentum basis states $|j,m\rangle$ for spin value $j=2$ in the Majorana representation. Degenerate points have been displaced for to make their multiplicity apparent.

- Spin coherent states
  For starters, let's see what the Majorana representation for the spin coherent state $|\vec{r}\rangle$ is. From equation D.4 in the appendix we have:

$$|\langle\vec{n}|\vec{r}\rangle| = \left|\frac{1+\vec{n}\cdot\vec{r}}{2}\right|^{2j}$$

  This is zero only when $\vec{n} = -\vec{r}$ and this zero has multiplicity $2j$. Therefore the Majorana representation of $|\vec{r}\rangle$ consist of all $2j$ points being situated at position $\vec{r}$. Note that for $j=\frac{1}{2}$ this is indeed the Bloch sphere, as each qubit is a spin coherent state.

- Next we look at the standard angular momentum basis states $|j,m\rangle$. From the expression for a SCS (4.3) we see that the Majorana polynomial is:

$$P_{|j,m\rangle} = \binom{2j}{j+m} z^{j-m}$$

  This equation has $j-m$ zero's at $z=0$. The remaining $j+m$ zero's are located at infinity. $z=0$ gets projected to the north pole and the point at infinity to the south pole. After taking antipodal points we see that $|j,m\rangle$ is represented by $j+m$ points on the north pole and $j-m$ points on the south pole, see figure 4.2.

Of potential interest is the application of the Majorana representation to systems of qubits. There's no clear immediate way to generalize to the space of multiple qubits in which the total spin j is not fixed. However, if we limit ourselves to the symmetric subspace of N qubits, then the spin j attains its maximal value of $N/2$ and we can use the Majorana representation.

- The GHZ-state $|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle^{\otimes N} + |1\rangle^{\otimes N})$ for N-qubits has Majorana polynomial:

$$P_{GHZ} = \frac{1}{\sqrt{2}}(1+z^N) = 0$$

  The $N=2j$ zero's lie on the unit circle and form the vertices of a regular N-sided polygon in the xy-plane. This remains true on the

sphere since the unit circle is mapped to itself under inverse stereographic projection.
As we can see from figure 4.3, a rotation of $2\pi/N$ about the z-axis returns the state to itself.

- The Majorana representation has been helpful as an inspiration to look for states which differ as much from the classical states (spin coherent states) as possible and this has let to the definition of anticoherent spin states [Zim06], which should be as 'nonclassical' as possible. Generally, the points for anticoherent states are spread out over the sphere. When an arbitrary rotation is performed (for example, by an unknown magnetic field), a class of states which are most sensitive against such a change are anticoherent states whose vertices form a regular polyhedron. This give rise to five classes of states, for $N = 4, 6, 8, 12$ and 20 corresponding to the five Platonic solids, see figure 4.4. It has been shown these are the optimal states for keeping reference frames aligned [KDD08].

4.3 The Schwinger boson representation

There is another, more widely known, formalism to work with angular momentum, introduced by Julian Schwinger [Sch65]. In this representation the angular momentum operators are expressed in terms of the creation and annihilation operators of two independent harmonic oscillators or field modes.

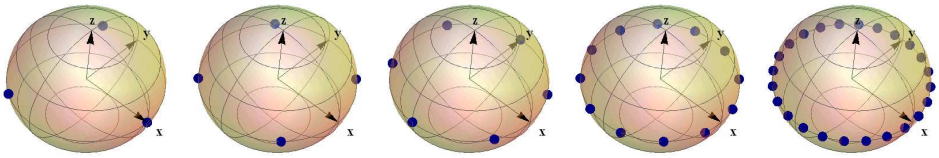


Figure 4.3: Majorana representation for the GHZ-states for $3, 4, 6, 12$ and 24 particles respectively.

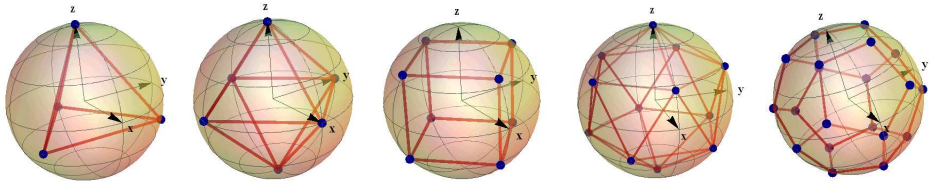


Figure 4.4: Majorana representation of states corresponding to the five Platonic solids.

The relation is expressed by the Schwinger-boson transform:

$$\vec{J} \leftrightarrow \frac{1}{2}(\hat{a}^\dagger \ \hat{b}^\dagger)\vec{\sigma}\begin{pmatrix}\hat{a}\\ \hat{b}\end{pmatrix}$$

where $\vec{\sigma}$ is the Pauli vector:

$$\vec{\sigma} = \hat{\sigma}_x\vec{e}_x + \hat{\sigma}_y\vec{e}_y + \hat{\sigma}_z\vec{e}_z = \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\vec{e}_x + \begin{pmatrix}0 & -i\\ i & 0\end{pmatrix}\vec{e}_y + \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\vec{e}_z$$

A calculation of the components gives:

$$\hat{J}_x = \frac{1}{2}(\hat{a}^\dagger b + \hat{a}\hat{b}^\dagger), \hat{J}_y = \frac{1}{2i}(\hat{a}^\dagger\hat{b} - a\hat{b}^\dagger), \hat{J}_z = \frac{1}{2}(\hat{a}^\dagger\hat{a} - \hat{b}^\dagger\hat{b})$$

and the spin raising- and lowering operators are:

$$\hat{J}_+ = \hat{a}^\dagger\hat{b}, \qquad \hat{J}_- = \hat{a}\hat{b}^\dagger$$

It can be verified that the commutation relations which generate the angular momentum algebra are conserved by using the commutation relations of the creation and annihilation operators:

$$[\hat{a}, \hat{a}^\dagger] = [\hat{b}, \hat{b}^\dagger] = 1, \qquad [\hat{a}, \hat{b}] = [\hat{a}, \hat{b}^\dagger] = 0$$

$$[\hat{J}_i, \hat{J}_j] = \frac{1}{4}\left[(\hat{a}^\dagger \ \hat{b}^\dagger)\sigma_i\begin{pmatrix}\hat{a}\\ \hat{b}\end{pmatrix}, (\hat{a}^\dagger \ \hat{b}^\dagger)\sigma_j\begin{pmatrix}\hat{a}\\ \hat{b}\end{pmatrix}\right] = i\epsilon^{ijk}\hat{J}_k$$

$$[\hat{J}_z, \hat{J}^\pm] = \pm\hat{J}^\pm$$

The basis states, which are products of number states $|n_A \quad n_B\rangle$ is identified with $|j, m\rangle$ by $j = (n_A + n_B)/2, m = (n_A - n_B)/2$. Note that a comma is used to distinguish the angular momentum basis from the Schwinger-boson basis. So

$$|j + m \quad j - m\rangle = |j, m\rangle$$

Therefore the state $|j, m\rangle$ can be created from the *vacuum state* $|vac\rangle = |0 \quad 0\rangle$ with $n_A = n_B = 0$ by applying the creation operators:

$$|j, m\rangle = \frac{(\hat{a}^\dagger)^{j+m}}{\sqrt{(j+m)!}}\frac{(\hat{b}^\dagger)^{j-m}}{\sqrt{(j-m)!}}|vac\rangle$$

We shortly draw attention to the fact that $\hat{a}^\dagger$ is applied $j + m$ times, $\hat{b}^\dagger$ is applied $j - m$ times and that in the Majorana representation this state has $j + m$ points on the north pole and $j - m$ points on the south pole. This connection will be explored in the next sections.

4.3.1 Spin coherent states in the Schwinger boson representation

If we consider the simple case of a single spin 1/2 particle state, or a single photon in a superposition of two modes, then this SCS can be written as:

$$|\theta,\phi\rangle = \cos\frac{\theta}{2}|1 \quad 0\rangle + \sin\frac{\theta}{2}e^{i\phi}|0 \quad 1\rangle = \left(\cos\frac{\theta}{2}\hat{a}^\dagger + \sin\frac{\theta}{2}e^{i\phi}\hat{b}^\dagger\right)|vac\rangle$$

If we define $\hat{c}^\dagger(\theta,\phi) = \cos\frac{\theta}{2}\hat{a}^\dagger + \sin\frac{\theta}{2}e^{i\phi}\hat{b}^\dagger$, then we can consider it to be the operator which, when applied to the vacuum state, creates a point on the Bloch-sphere at position (θ,ϕ). Taking this interpretation we can ask whether repeated application of $c^\dagger(\theta,\phi)$ creates a SCS with appropriate j. Writing out the definition of a SCS in the Schwinger-boson representation gives:

$$\begin{aligned}|\theta,\phi\rangle = & \\ &\sum_{m=-j}^{j}\binom{2j}{j+m}^{1/2}\left[\cos\frac{\theta}{2}\right]^{j+m}\left[e^{i\phi}\sin\frac{\theta}{2}\right]^{j-m}\frac{(\hat{a}^\dagger)^{j+m}(\hat{b}^\dagger)^{j-m}}{\sqrt{(j+m)!}\sqrt{(j-m)!}}|vac\rangle \\ &= \frac{1}{\sqrt{(2j)!}}\left(\sum_{m=-j}^{j}\binom{2j}{j+m}\left[\hat{a}^\dagger\cos\frac{\theta}{2}\right]^{j+m}\left[\hat{b}^\dagger e^{i\phi}\sin\frac{\theta}{2}\right]^{j-m}\right)|vac\rangle\end{aligned}$$

The last expression in parenthesis is recognized to be a binomial expansion. We therefore have our result:

$$|\theta,\phi\rangle = \frac{1}{\sqrt{(2j)!}}\left(\hat{a}^\dagger\cos\frac{\theta}{2} + \hat{b}^\dagger\sin\frac{\theta}{2}e^{i\phi}\right)^{2j}|vac\rangle = \frac{[\hat{c}^\dagger(\theta,\phi)]^{2j}}{\sqrt{(2j)!}}|vac\rangle$$

Alternatively, we can consider a rotated version of the operator $\hat{a}^\dagger$:

$$\hat{a}^\dagger \to \hat{R}(\theta,\phi)\hat{a}^\dagger\hat{R}(\theta,\phi)^\dagger$$

Giving:

$$|\theta,\phi\rangle = \hat{R}(\theta,\phi)\frac{(\hat{a}^\dagger)^{2j}}{\sqrt{(2j)!}}\hat{R}(\theta,\phi)^\dagger|vac\rangle$$

for $j = 1/2$. Note that $R(\theta,\phi)^\dagger|vac\rangle = |vac\rangle$. In our case the expression for the rotation operator is:

$$\hat{R}(\theta,\phi) = \exp(\alpha\hat{J}_- - \alpha^*\hat{J}_+) = \exp(\alpha\hat{a}\hat{b}^\dagger - \alpha^*\hat{a}^\dagger\hat{b})$$

where $\alpha = \frac{\theta}{2}e^{i\phi}$. By applying a Baker-Campbell-Hausdorff relation to this rotated creation operator it can be shown that:

$$\hat{R}(\theta,\phi)\hat{a}^\dagger\hat{R}^\dagger(\theta,\phi) = \hat{c}^\dagger(\theta,\phi)$$

An obvious corollary of this is that:

$$\hat{R}(\theta,\phi)\frac{\left(\hat{a}^\dagger\right)^{2j}}{\sqrt{(2j)!}}\hat{R}^\dagger(\theta,\phi)=\frac{[\hat{c}^\dagger(\theta,\phi)]^{2j}}{\sqrt{(2j)!}}$$

Therefore we have extended the previous result that $\hat{c}^\dagger$ is an operator which creates points on the Majorana sphere to the case of SCS's. In the next section we will show that this is true in general, completing the connection between the Majorana and Schwinger representations.

4.3.2 The connection between the Schwinger-boson and Majorana representations

In the Schwinger-boson picture we were able to interpret the rotated creation operator $\hat{c}^\dagger$ as an operator creating points on the Majorana sphere, at least for SCS's. In the Majorana picture however, an arbitrary state $|\psi\rangle$ of spin j is represented by a non-ordered collection of $2j$ points on the sphere. If we label these $2j$ points corresponding to the state $|\psi\rangle$ by (θ_i,ϕ_i) for $i=1,2,...,2j$, we pose the question whether it is true that

$$|\psi\rangle=\frac{C}{\sqrt{(2j)!}}\left(\prod_{i=1}^{2j}c^\dagger(\theta_i,\phi_i)\right)|vac\rangle$$

In that case, the equivalence and exact relation of the Schwinger and Majorana representations is demonstrated and we could have started from the Schwinger representation to introduce the Majorana picture. The relation is, as could be expected, true. The normalization constant C, however, depends on the angles between the points $\vec{r}_i$ in a nontrivial way.
For the expansion of the product in parenthesis, we will need to invoke expressions involving the elementary symmetric polynomials s_n in the variables $w_i=\tan\frac{\theta_i}{2}e^{i\phi_i}$, see table 4.1. Explicitly, the expansion gives:

$$\prod_{i=1}^{2j}c^\dagger(\theta_i,\phi_i)=\left(\prod_{i=1}^{2j}\cos\frac{\theta_i}{2}\right)\sum_{m=-j}^{j}s_{j-m}(a^\dagger)^{j+m}(b^\dagger)^{j-m}$$

This yields as expression for the state:

$$|\psi\rangle=C\left(\prod_{i=1}^{2j}\cos\frac{\theta_i}{2}\right)\sum_{m=-j}^{j}\binom{2j}{j+m}^{-1/2}s_{j-m}|j,m\rangle \tag{4.4}$$

To find the corresponding points on the sphere we form the inner product of

Definition of the k-th elementary symmetric polynomial in $(w_1, \ldots, w_N)$
$s_k = \sum_{1 \le i_1 < \ldots < i_k \le N} w_{i_1} \times \ldots \times w_{i_k}$
Example, for $N = 4$:
$s_0 = 1$
$s_1 = w_1 + w_2 + w_3 + w_4$
$s_2 = w_1w_2 + w_1w_3 + w_1w_4 + w_2w_3 + w_2w_4 + w_3w_4$
$s_3 = w_1w_2w_3 + w_1w_2w_4 + w_1w_3w_4 + w_2w_3w_4$
$s_4 = w_1w_2w_3w_4$

Table 4.1: Definition of the elementary symmetric polynomials.

$|\psi\rangle$ with a SCS $|z\rangle$:

$$\begin{aligned}\langle z|\psi\rangle &= N(z)C\left(\prod_{i=1}^{2j}\cos\frac{\theta_i}{2}\right)\sum_{m=-j}^{j}(z^*)^{j-m}s_{j-m} \\ &= N(z)C\left(\prod_{i=1}^{2j}\cos\frac{\theta_i}{2}\right)\prod_{i=1}^{2j}(1+z^*w_i)\end{aligned}$$

$N(z)$ is the normalization constant in front of the SCS $|z\rangle$. From this expression we can see that the $2j$ zeros p_i of this polynomial are $p_i = -1/w_i^*$, thus after stereographic projection and taking antipodal points we find that the $2j$ points on the Majorana sphere corresponding to $|\psi\rangle$ are indeed the (θ_i, ϕ_i) created by $c^\dagger(\theta_i, \phi_i)$. This concludes the proof of the very geometric connection between the Schwinger and Majorana representations.

4.3.3 Calculations in the Majorana picture

As we can see from equation (4.4), the relation between the points on the Majorana sphere and the coefficients in the component representation is rather complex and difficult to work with. Ideally, to use the Majorana picture in conjunction with the Schwinger formalism effectively we would like to express all quantities of interest in terms of the $2j$ points (θ_i, ϕ_i) (or unit vectors $\vec{r}_i$, or complex numbers w_i when it proves more convenient).
In equation (4.4) we have shown that:

$$|\vec{r}_1, \ldots, \vec{r}_{2j}\rangle = C\left(\prod_{i=1}^{2j}\cos\frac{\theta_i}{2}\right)\sum_{m=-j}^{j}\binom{2j}{j+m}^{-1/2}s_{j-m}|j,m\rangle$$

The normalization constant C is a function of the of the $2j$ points w_i and we'll write $C = C(\vec{w})$ where $\vec{w} = (w_1, \ldots, w_{2j})$ as shorthand notation. If we make the convention that we fix the global phase by making the coefficient of $|j, j\rangle$ real and positive as we do with SCS's, then $C(\vec{w})$ is real and positive and can be found by taking the inner product with $|j, j\rangle = |\vec{z}^{\otimes 2j}\rangle$. Explicitly:

$$C(\vec{w}) = \left[\left(\prod_{i=1}^{2j} \cos^2 \frac{\theta_i}{2} \right) \left(\sum_{m=-j}^{j} \binom{2j}{j+m}^{-1} |s_{j-m}(\vec{w})|^2 \right) \right]^{-\frac{1}{2}}$$

In the special case of a SCS, all w_i's are identical and $s_{j-m} = \binom{2j}{j-m} w^{j-m}$, so then:

$$\begin{aligned} C(w) &= \left[\left(\cos^2 \frac{\theta}{2} \right)^{2j} \left(\sum_{m=-j}^{j} \binom{2j}{j+m} |w|^{2(j-m)} \right) \right]^{-\frac{1}{2}} \\ &= \left[\left(\frac{1}{1+|w|^2} \right)^{2j} \left(1+|w|^2\right)^{2j} \right]^{-\frac{1}{2}} = 1 \end{aligned}$$

It is clear from the fact that the $2j$ points are unordered that only symmetric expressions involving w_i should enter all our calculations. In fact, the relation between a product state $|\vec{r}_1\rangle|\vec{r}_2\rangle \cdots |r_{2j}\rangle$ of qubits, where each $|\vec{r}_i\rangle$ can be represented by a point on the Bloch sphere, and the state $|\vec{r}_1, \cdots, \vec{r}_{2j}\rangle$ for the symmetrized state with spin j is direct and intuitive: The symmetrized expression of $|\vec{r}_1\rangle|\vec{r}_2\rangle \cdots |r_{2j}\rangle$ is proportional to $|\vec{r}_1, \cdots, \vec{r}_{2j}\rangle$.
To see this, consider the unitary symmetrization operator:

$$\hat{S} = \frac{1}{N!} \sum_{\sigma \in S_N} \hat{P}_\sigma$$

In the present case, $N = 2j$ is the number of particles of the system. A straightforward calculation reveals that:

$$|\vec{r}_1, \cdots, \vec{r}_{2j}\rangle = C\hat{S}|\vec{r}_1\rangle|\vec{r}_2\rangle \cdots |\vec{r}_{2j}\rangle \tag{4.5}$$

Taking the inner product of both sides with a coherent state $|\vec{r}\rangle^{\otimes 2j}$ we get:

$$C = \frac{\langle \vec{r}|^{\otimes 2j} |\vec{r}_1, \cdots, \vec{r}_{2j}\rangle}{\prod_{i=1}^{2j} \langle \vec{r}|\vec{r}_i\rangle}$$

From this we see that C is a geometrical quantity depending only on the angles (or inner products) of the vectors $\vec{r}_i$. The expression is quite complicated for even moderately large values of N. Values of C for the first few numbers integers are shown in table 4.2 in terms of the vectors $\vec{r}_i$.

$N = 2j$	C^{-2}
1	1
2	$\frac{1}{4}(3 + \vec{r}_1 \cdot \vec{r}_2)$
3	$\frac{1}{6}(3 + \vec{r}_1 \cdot \vec{r}_2 + \vec{r}_1 \cdot \vec{r}_3 + \vec{r}_2 \cdot \vec{r}_3)$
4	$\frac{1}{48}(15 + 5\sum_{i<j} \vec{r}_i \cdot \vec{r}_j + ((\vec{r}_1 \cdot \vec{r}_2)(\vec{r}_3 \cdot \vec{r}_4) + (\vec{r}_1 \cdot \vec{r}_3)(\vec{r}_2 \cdot \vec{r}_4) + (\vec{r}_1 \cdot \vec{r}_4)(\vec{r}_2 \cdot \vec{r}_3)))$

Table 4.2: The normalization constant C for various values of j

For arbitrary states, the expression for C has been calculated in [Lee88], in terms of $\sigma_{ij} = \sin^2 \frac{\angle(\vec{r}_i, \vec{r}_j)}{2} = \frac{1}{2}(1 - \vec{r}_i \cdot \vec{r}_j)$, which is equal to the square of half the length of the chord between $\vec{r}_1$ and $\vec{r}_2$. The expression for C is:

$$C^{-2} = \sum_{m=0}^{\lfloor N/2 \rfloor} (-1)^m \frac{(N-m)!}{N!} D_m^N$$

Where D_m^N is a sum of terms, each term involving m products of σ_{ij} where no repetitions of indices are allowed in each term. The indices take values from 1 to N. For example:

$$D_1^3 = \sigma_{12} + \sigma_{13} + \sigma_{23}$$

$$\begin{aligned} D_2^5 = {} & \sigma_{12}\sigma_{34} + \sigma_{12}\sigma_{35} + \sigma_{12}\sigma_{45} + \sigma_{13}\sigma_{24} + \sigma_{13}\sigma_{25} + \sigma_{13}\sigma_{45} + \sigma_{14}\sigma_{23} + \sigma_{14}\sigma_{25} \\ & + \sigma_{14}\sigma_{35} + \sigma_{15}\sigma_{23} + \sigma_{15}\sigma_{24} + \sigma_{15}\sigma_{34} + \sigma_{23}\sigma_{45} + \sigma_{24}\sigma_{35} + \sigma_{25}\sigma_{34} \end{aligned}$$

$$D_3^6 = \sigma_{12}\sigma_{34}\sigma_{56}$$

4.4 Multipartite systems and entanglement

As explained in the appendix (D), the symmetric state space of N qubits is mathematically identical to that of a *single* spin system with spin $j = N/2$. In this case, the state $|\vec{r}_1, \cdots, \vec{r}_{2j}\rangle$ is proportional to the symmetrized product state $|\vec{r}_1\rangle|\vec{r}_2\rangle \cdots |\vec{r}_{2j}\rangle$. The spin coherent states are exactly the product states in the symmetric subspace, i.e. of the form $|\vec{r}\rangle^{\otimes N}$. These are represented on the Majorana sphere by a single point at $\vec{r}$ with multiplicity $2j$. On the other hand, the maximally entangled Bell states:

$$|\Psi^+\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle)$$

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$$

$$|\Phi^-\rangle = \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle)$$

for $N = 2$ have their points as far apart as possible, on opposite points on the sphere.

This leads one to naturally wonder whether the amount of 'spread' in the points on the sphere is in some fashion a measure of the degree of entanglement. To investigate this question, we will look at a specific measure of entanglement: the geometric entanglement [VPRK97][WG03], defined for a pure state as:

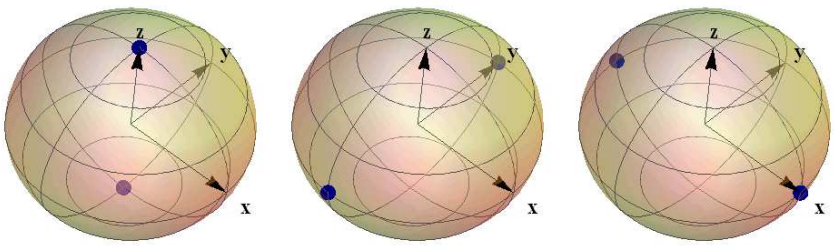
$$E_g(|\psi\rangle) := \min_{|\phi\rangle \in \mathcal{H}_{prod}} -\log |\langle\phi|\psi\rangle|^2$$

Here, $\mathcal{H}_{prod}$ is the set of product states. So it is geometrically a measure of the distance to a closest product state. We will limit our analysis to the symmetric state space, so we can use the Majorana representation.

Since for a system of qubits, every symmetric product state is a coherent state, we can parameterize $\mathcal{H}_{prod}$ by the angles (θ, ϕ) of the SCS $|\vec{r}\rangle$. Using equation (4.5) we easily find for the inner product we want to maximize:

$$|\langle\vec{r}|\vec{r}_1, \cdots, \vec{r}_{2j}\rangle|^2 = |C|^2 \prod_{i=1}^{2j} |\langle\vec{r}|\vec{r}_i\rangle|^2$$

This is zero whenever $\vec{r}$ is antipodal to one of the points $\vec{r}_i$. So intuitively, we would like $\vec{r}$ to be as close as possible to the points $\vec{r}_i$, but more importantly, as far away as possible from the antipodes. When the points are very much spread out over the sphere, $\vec{r}$ will nowhere be very far from an antipodal point and so the inner product will be relatively small and the degree of



(a) $|\Psi^+\rangle$ (b) $|\Phi^+\rangle$ (c) $|\Phi^-\rangle$

Figure 4.5: Majorana representation of the three 'spin 1' Bell states. From this picture it is immediately obvious that they are rotated versions of each other and eigenstates of respectively $\hat{J}_z, \hat{J}_y$ and $\hat{J}_x$ with eigenvalue 0. $|\Phi^+\rangle$ and $|\Phi^-\rangle$ are related by a 90° rotation about the z-axis. A 180° rotation about their corresponding axes leave the state unchanged.

entanglement high.

For some specific calculations, let us first look at the basis states $|j,m\rangle$. These are represented by $j+m$ points on the north pole and $j-m$ points on the south pole. The states $|j,j\rangle$ and $|j,-j\rangle$ are the only coherent states and thus the only product states. A straightforward calculation gives:

$$|\langle\vec{r}|j,m\rangle| = |\langle\vec{r}|\vec{z}\rangle|^{j+m}|\langle\vec{r}|-\vec{z}\rangle|^{j-m} = \binom{2j}{j+m}^{1/2}\left(\cos\frac{\theta}{2}\right)^{j+m}\left(\sin\frac{\theta}{2}\right)^{j-m}$$

The maximum is found by simple differentiation which leads to the result that a closest coherent state is given when $\cos\theta = \frac{m}{j}$. The azimuthal angle ϕ is arbitrary, which is immediately obvious from the symmetry of the problem. The value of the inner product is then:

$$|\langle\theta,\phi|j,m\rangle|^2 = \binom{2j}{j+m}\left(\frac{j+m}{2j}\right)^{j+m}\left(\frac{j-m}{2j}\right)^{j-m}$$

Next, let's consider the GHZ states: $(|0\rangle^{\otimes n}+|1\rangle^{\otimes n})/\sqrt{2}$. From figure 4.3 it is immediately clear that most distant state $|\bar{\psi}\rangle$ is either at the north- or south pole. Both immediately give:

$$|\langle\bar{\psi}|GHZ\rangle|^2 = 1/2$$

Lastly, we will give a simple geometric way to find the closest product state for the case of two qubits $|\vec{r}_1\rangle$ and $|\vec{r}_2\rangle$. We can exploit the rotational invariance of E_g. If the angle between $\vec{r}_1$ and $\vec{r}_2$ is θ', we can perform a rotation to bring $\vec{r}_1$ to the north pole. Then we perform a rotation about the z-axis to bring $\vec{r}_2$ in the xz-plane. So we have $\vec{r}_1 = (0,0,1)$ and $\vec{r}_2 = (\sin\theta', 0, \cos\theta')$ without loss of generality. We want to maximize:

$$|\langle\theta,\phi|0,0\rangle\langle\theta,\phi|\theta',0\rangle|^2 = \cos^2\frac{\theta}{2}\cos^2\left(\frac{\theta'-\theta}{2}\right)$$

The maximum is found by taking the derivative with respect to θ, which leads to:

$$\tan\frac{\theta}{2} = \tan\left(\frac{\theta'-\theta}{2}\right) \Rightarrow \theta = \frac{\theta'}{2} \tag{4.6}$$

In general, then, for two qubits $|\vec{r}_1\rangle$ and $|\vec{r}_2\rangle$, the closest symmetric product state is given by the vector which is the arithmetic mean $\vec{r} = (\vec{r}_1+\vec{r}_2)/|\vec{r}_1+\vec{r}_2|$. When the mean is zero, the points are antipodal and the state is a rotation of the basis state $|j=1,m=0\rangle$ for which we already determined that any vector $\vec{r}$ perpendicular to the axis joining $\vec{r}_1$ and $\vec{r}_2$ will determine a closest product state.

4.5 Generalization to arbitrary N spin j states

The Majorana representation is a representation of the state of a particle with a *given* value of j. In general, though, when combining two systems with angular momenta j_1 and j_2, the eigenvalues of the total angular momentum squared $J^2 = (\vec{J}_1 + \vec{J}_2)^2$ can take values $j(j+1)$ with j between $|j_1 - j_2|$ and $j_1 + j + 2$ in integer steps. Fore example, as is well-known, the space of two spin 1/2 particles is spanned by the singlet ($j = 0$) and triplet ($j = 1$) states. As we combine more particles, the possible values of j become highly degenerate. This is because the possible values of j only increase linearly, while the dimension $(2j+1)^N$ of the system increases exponentially. We will denote the degeneracy of the spin value j by combining N spin 1/2 particles by $W(N, j)$.

4.5.1 Combining N spin 1/2 particles

There are many ways to calculate $W(N, j)$ and the answer was already given by Dicke [Dic54] in his paper on superradiance. When we add spin 1/2 to a system of spin j, the spin of the product space can take the values $j - \frac{1}{2}$ and $j + \frac{1}{2}$, provided the former is not negative. Thus, by combining N spin 1/2 we can arrive at a particular spin value by different paths as shown schematically below. The value of $W(N, j)$ is thus equal to the number of paths to a certain point in a random walk with a hard reflecting boundary. By this analogy, we can use the result in [Orl03] which tackles this problem. The result, translated to our purposes is:

$$W(N,j) = \binom{N}{\frac{N}{2}+j} - \binom{N}{\frac{N}{2}+j+1} = \binom{N}{\frac{N}{2}+j}\frac{(2j+1)}{(\frac{N}{2}+j+1)}$$

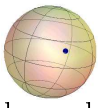
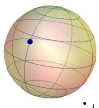
A general expression for combining N spin j particles, for arbitrary j is more complicated and was calculated by [Mih77, Ras77]. For simplicity, we will focus only on a collection of N spin 1/2 particles.

4.5.2 The general idea

In this subsection, the idea behind the generalization of the Majorana representation will be explained. A more rigorous description will be given in the following sections.

If we think about generalizing the Majorana representation to an N qubit system, it is clear that we need to introduce multiple spheres because of degeneracy. For three qubits, we can have $j = \frac{1}{2}$ with multiplicity two and $j = \frac{3}{2}$ with multiplicity one. A quick counting of the degrees of freedom shows that the number of points is $2 \cdot 1 + 1 \cdot 3 = 5$, which equals 10 real parameters, while we need $2 \cdot 2^3 - 2 = 14$ real parameters to specify a 3-qubit state. The reason is that, in the Majorana representation, we have

already imposed the normalization and phase-constraints on each sphere. The parts of the state vector in the subspaces belonging to different j may have a relative phase difference, or a different magnitude. Schematically, the description of a state is thus:

$|\psi\rangle = C_{1/2,1} \times$  $+ C_{1/2,2} \times$  $+ C_{3/2,1} \times$ 

This introduces three phases and three magnitudes, given by the complex coefficients $C_{\frac{1}{2},1}, C_{\frac{1}{2},2}, C_{\frac{3}{2},1}$. The normalization and phase constraints hold for these coefficients, so these account for the 'missing' 4 parameters.

In general, we will have $\lfloor \frac{N}{2} + 1 \rfloor$ subspaces, one for each value of j. The idea is to represent a state vector in the basis $|j, p; m\rangle$, where p labels the degeneracy of j:

$$|\psi\rangle = \sum_{j,p,m} c_{j,p,m} |j, p; m\rangle$$

In the above sum, j runs from 0 ($\frac{1}{2}$) if N is even (odd) to $\frac{N}{2}$. The basis states with given values of (j, p) span a $2j + 1$ dimensional subspace. If we introduce the projection operator $\hat{P}_{j,p}$ which projects on this subspace, we can write:

$$\hat{P}_{j,p} = \sum_{m=-j}^{j} |j, p; m\rangle\langle j, p; m|, \qquad \sum_{j} \sum_{p=1}^{W(N,j)} \hat{P}_{j,p} = 1$$

$$|\psi\rangle = \sum_{j} \sum_{p=1}^{W(N,j)} \hat{P}_{j,p} |\psi\rangle$$

The vector $\hat{P}_{j,p}|\psi\rangle$ holds all information on the points of the sphere labeled (j, p). We introduce the normalized version as $|\psi_{j,p}\rangle$:

$$|\psi_{j,p}\rangle = \exp(-i \mathrm{arg} c_{j,p;j}) \frac{P_{j,p}|\psi\rangle}{\sqrt{\langle\psi|P_{j,p}|\psi\rangle}}$$

We have chosen the phase of the coefficient corresponding to the highest spin to be zero, as was the practice in the previous sections. We can now write:

$$|\psi\rangle = \sum_{j,p} C_{j,p} |\psi_{j,p}\rangle$$

$$C_{j,p} = \sqrt{\langle\psi|P_{j,p}|\psi\rangle} \exp(i \mathrm{arg} c_{j,p;j})$$

And we can fix the phase of $C_{\frac{N}{2},1}$ to be zero. Each of the states $|\psi_{j,p}\rangle$ are exactly of the form used in the previous sections with given j. The coefficients $C_{j,p}$ can be captured in a normalized complex vector $\vec{C}$ holding information

on the relative phases and weights of the spheres.

A great advantage of this representation is that global rotations are easily visualized. Each sphere will rotate independently from each other and the relative phases and magnitudes will remain unchanged. An issue we have not touched yet, though, is: how do we decide on the degenerate subspaces for a given j? For this, we look for another natural group acting on our system of N-qubits besides the rotation group: the symmetric group S_N consisting of permutations of N objects. The group actions $SU(2)$ and S_N work together nicely: Under a permutation of N qubits,

4.5.3 A little representation theory

Everything in this section is most easily captured in the language of representation theory. We will therefore quickly provide some definitions and basic results that are useful to us. For a more complete treatment that is accessible to physicists, see for example [Sag01].

For a group G and a vector space V, we speak of a *representation* when we associate a linear transformation on V for each element $g \in G$, such that $g_1(g_2 v) = (g_1 g_2 v)$ and the identity element of G acts as the identity on V. The vector space V is called the representation space, although it is often just called the representation. Examples of representations are the group of translations of the plane $\mathbf{R}^2$, or the group $SO(3)$ acting on the Hilbert space of a spin j particle through the identification given in equation (4.1).

Sometimes the vector space V falls apart into pieces under the action of G. That means that there exists a non-trivial subspace $W \subset V$, such that $gW \in W$ for all $g \in G$. W is called an invariant subspace. So V can be 'broken up', *reduced* into W and its complement $W^\perp$, which are subspaces of lower dimension than V. Suppose we continue this process by breaking up W and $W^\perp$ when possible, just like decomposing a natural number into a product of primes. In the end we will have written V as a direct sum of subspaces W_i which are atomic in the sense that they cannot be broken up further, i.e. they have no nontrivial invariant subspaces and are called *irreducible*. These subspaces W_i by themselves form irreducible representations of G. In this way, every representation can be written as a sum of irreducible representations and, like prime factorization, this decomposition is unique.

The state space $\mathcal{H}_s$ of a spin s particle is a $2s+1$ dimensional, irreducible representation of the rotation group $SO(3)$. When combining N such particles, our space is $\mathcal{H}_s^{\otimes N}$, which can be written as a sum of irreducible representations:

$$\mathcal{H}_s^{\otimes N} = \sum_j \mathcal{H}_j^{\oplus W_s(N,j)},$$

where $W_s(N,j)$ is the number of times the irreducible representation $\mathcal{H}_j$ occurs in the expansion. The problem of finding the multiplicities $W_s(N,j)$ is exactly the same as finding the degeneracy of spin j when combining N spin s particles, described in the previous subsection. For $s = \frac{1}{2}$, the multiplicity of $\mathcal{H}_j$ in $\mathcal{H}_{\frac{1}{2}}^{\otimes N}$ is just the degeneracy $W(N,j)$ introduced earlier.

4.5.4 Irreducible representations of the symmetric group

A basic result in the representation theory of the symmetric group[5] is that the number of irreducible representations of S_N is equal to the number of conjugacy classes of S_N. The elements of a certain conjugacy class in S_N all have the same *cycle*-type, which follows from the relation:

$$\sigma(a_1, a_2, \ldots, a_N)\sigma^{-1} = (\sigma(a_1), \sigma(a_2), \ldots, \sigma(a_N))$$

The reverse, that permutations with the same cycle-type are conjugate is also true.
If we order the cycle-decomposition of a permutation from largest cycle-length to shortest, we can denote them by a partition of N. I.e. the permutation $(1374)(25)(6)$ in S_7 has cycle-type $(4,2,1)$. This means that the conjugacy classes, and thus the irreducible representations can be labeled by partitions of N.
A way to visualize a partition is by drawing a *Young diagram*. This is a stack of left-justified rows, where the rows are non increasing from top to bottom. Figure 4.6 shows a Young diagram for the partition $(5,4,1)$.
The number of Young Diagrams $p(N)$ of size N grows rapidly. Its generating function is given by:

$$\sum_{n=0}^{\infty} p(n)x^n = \prod_{n=1}^{\infty} \frac{1}{1-x^n}$$

There is no closed formula for $p(N)$, but its asymptotic behavior is $p(N) \sim \exp(\sqrt{2N/3}\pi)/4N\sqrt{3}$ [HR18]. It will delve us too deep into the representation theory of the symmetric groups too explain how to build the irreducible representation corresponding to a Young diagrams of a given shape, but

[5] The symmetric group S_N is the permutation group consisting of all permutations of N objects.

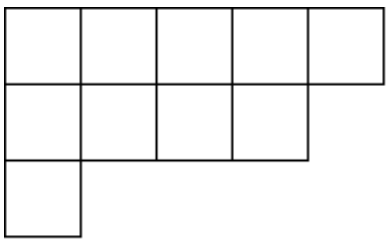
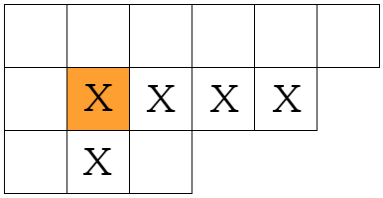


Figure 4.6: Young diagram corresponding to the partition $(5,4,1)$, which is one of the 42 partitions of the number 10.



(a) The hook length corresponding to the colored cell is the number of **X**'s.

9	7	6	3	2	1
5	3	2			
4	2	1			
1					

(b) A Young Diagram with the hook lengths of each cell.

Figure 4.7

we will state some important results. For a simple introductory treatment, see [Sag01]. For a treatment with applications in chemistry, see [Pau95] and for the slightly more abstract minded we recommend [Ful96].

As stated before, with every partition λ of N we can associate an irreducible representation S^λ of S_N which can be represented by a Young diagram of size N. The dimension of S^λ is given by the *hook length formula*:

$$\dim S^\lambda = \frac{N!}{\text{product of hook lengths in } \lambda}$$

The hook-length (see figure 4.7) of a box (i, j) in λ is the number of boxes to the right of (i, j), plus the number below (i, j), plus one for the box (i, j) itself. To find the dimension of the irreducible representation S^λ we take $N!$ and divide by the hook-lengths of all the boxes in λ.

- **Example 1.**
  As an example, we take the (trivial) partition N. Its Young diagram consists of just one row of length N. The hook-lengths are immediately seen to be $1, 2, \ldots, N$. So it corresponds to a $N!/N! = 1$-dimensional representation, which turns out to be the trivial representation: all permutations act as the identity. We will see later that this relates to the completely symmetrized subspace of our multipartite state space.

- The second obvious partition is $\lambda = 1 + 1 + \cdots + 1$. The Young diagram is a single column of length N. Again the dimension is seen to be one and it corresponds to the *alternating representation*, where each partition acts as ± 1 according to its sign.

- One more example that we will use the case where $\lambda = L + (N - L)$. The Young diagram has two rows of lengths L and $N - L$ respectively. We require, of course, that $L \geq N$. In this case, we can group the

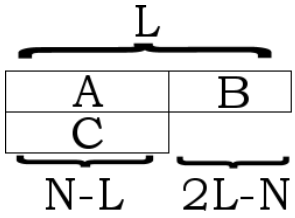


Figure 4.8: A Young diagram with two rows, split into three parts. The products of the hook lengths of each of the parts A, B and C can be written down explicitly.

boxes in three parts, A, B and C as in figure 4.8. The products of the hook-lengths for each of the parts are:

$$\begin{aligned} A: &\quad \frac{(L+1)!}{(2L-N+1)!} \\ B: &\quad (2L-N)! \\ C: &\quad (N-L)! \end{aligned}$$

So the dimension of the corresponding irreducible representation is:

$$\dim S^{(L,N-1)} = \frac{N!(2L-N+1)!}{(N-L)!(2L-N)!} = \binom{N}{L}\frac{2L-N+1}{L+1}$$

4.5.5 Schur-Weyl duality

In our generalization of the Majorana representation, we have separated our Hilbert space $\mathcal{H}_s^{\otimes N}$ for N spin s particles into subspaces with constant s, labeled by degeneracy p. In other words, we have written $\mathcal{H}_j^{\otimes N}$ as a direct sum of irreducible representations of the rotation group. The action of a rotation for a single spin s particle was given in equation (4.1). Since the a global rotation simply factors into a product of separate rotations for each particle:

$$e^{-i\alpha\vec{n}\cdot\vec{S}_T} = e^{-i\alpha\vec{n}\cdot\vec{S}_1}e^{-i\alpha\vec{n}\cdot\vec{S}_2}\dots e^{-i\alpha\vec{n}\cdot\vec{S}_N}$$

The action of a rotation $\hat{U} = e^{-i\alpha\vec{n}\cdot\vec{S}_T}$ on a product state is simply:

$$\hat{U}|x_1\rangle|x_2\rangle\dots|x_N\rangle = \hat{U}|x_1\rangle\hat{U}|x_2\rangle\dots\hat{U}|x_N\rangle \tag{4.7}$$

As mentioned before, we have another group which naturally acts on $\mathcal{H}_s^{\otimes N}$: the symmetric group S_N. If σ a permutation of N objects, then the corresponding permutation operator $\hat{P}_\sigma$ acts on a basis state as:

$$\hat{P}_\sigma|x_1\rangle|x_2\rangle\cdots|x_N\rangle = |x_{\sigma^{-1}(1)}\rangle|x_{\sigma^{-1}(2)}\rangle\cdots|x_{\sigma^{-1}(N)}\rangle \tag{4.8}$$

The inverse permutation σ^{-1} occurs, because we wish to permute the *particles*, or *kets* and not the labels. For example, when we have three particles and the permutation (123), we get:

$$\hat{P}_{(123)}|x_1\rangle|x_2\rangle|x_3\rangle = |x_3\rangle|x_1\rangle|x_2\rangle$$

The first ket ($|x_1\rangle$) went to the second ($|x_2\rangle$) etc. So we see the labels are transformed with the inverse permutation.

We would like to combine the representations of $SO(3)$ and S_N together nicely. We can combine the representations by considering the product $SO(3) \times S_N$, whereby the actions are given by the unitary operators $\hat{U}\hat{P}$, where $\hat{U}$ is a rotation and (or a general unitary operator) $\hat{P}$ a permutation. By choosing the identity for $\hat{U}$ or $\hat{P}$ the action reduces to that of P and $SO(3)$ respectively. It's clear from equations (4.7) and (4.8) that the two actions commute: $\hat{U}(\hat{n},\alpha)\hat{P}_\sigma = \hat{P}_\sigma\hat{U}(\hat{n},\alpha)$ for all $\hat{n},\alpha$ and σ. The full consequence of the above is called *Schur-Weyl duality*, a theorem which states that:

$$\mathcal{H}_s^{\otimes N} = \bigoplus_{\lambda \in P(N,2s+1)} \mathcal{H}_{j(\lambda)} \otimes S^\lambda$$

Where $P(N,d)$ is the set of partitions of N into $\leq d$ parts. $\mathcal{H}_j$ and S^λ denote, as before, irreducible representations of $SO(3)$ and S_N respectively. So Schur-Weyl duality asserts that $\mathcal{H}_s^{\otimes N}$ falls apart into a set of mutually orthogonal subspaces, labeled by $\lambda \in P(N, 2j+1)$, on which both $SO(3)$ and S_N act irreducibly. Thus instead of labeling our spheres by the pair (j,p), we choose to label them by a partition λ.

For spin $\frac{1}{2}$, we only consider partitions consisting of 1 or 2 parts (i.e. one or two rows in the Young diagram of λ). It can be shown that the spin value corresponding to a partition λ is half the difference of the row lengths. A way to verify this statement is to consider the dimension of the S^λ, which must equal the degeneracy $W(N,j)$. The top row has length $\frac{N}{2}+j$ and the bottom one $\frac{N}{2}-j$ so that there are N boxes in total and the difference in row lengths is $2j$. Thus we can write the partition as $\lambda = (\frac{N}{2}+j)+(\frac{N}{2}-j)$. Using the earlier result in equation (4.5.4) with $L = \frac{N}{2}+j$, we find:

$$W(N,j) = \binom{N}{\frac{N}{2}+j} \frac{2j+1}{\frac{N}{2}+j+1},$$

in agreement with our earlier results (equation (4.5.1)).

For the partition $\lambda = N$ we have already seen that corresponds to the trivial representation of the symmetric group. The corresponding spin value achieves the maximum value of $N/2$ and is thus $N+1$-dimensional. Therefore, the subspace of maximum angular momentum is completely symmetric

with respect to any permutation of particles, which is also shown in appendix D.
As a last example, for two qubits, the two partitions of two are $\lambda_1 = 1 + 1$ and $\lambda = 2$. Latter corresponds to the maximum spin (spin 1) subspace, which consist of the triplet states. The former corresponds to the singlet state and is antisymmetric under the exchange of particles.

When we use the Schur basis as our basis for the generalized Majorana representation, the action of a (global) rotation simply rotates each sphere independently and a permutation of the particles permutes the spheres (or the coefficients $C_{j,p}$) in the way corresponding to the irreducible representation.

4.6 Summary and discussion

We have investigated the Majorana representation, which represents a spin j particle by $2j$ points on a sphere. A link was made with the, better known, Schwinger boson representation and we found the explicit relation linking the two. For the purpose of gaining insight in N-qubit systems, we formulated a generalization of the Majorana representation with the aid of Schur-Weyl duality. In this representation, the action of certain quantum gates may be visualized. We note, however, that the irreducible representations of the symmetric group for even moderately high values of N act in a nonintuitive way. Furthermore, the representation acts only covariantly under global rotations (a uniform rotation of all qubits). An action like the spin flip of a single qubit generally acts very nontrivially in this picture. Further investigation into the structure of this representation is needed[6].

Although the Majorana representation has proven to be useful in providing insight into various areas of physics [Zim06, KDD08, MS07], the structure of the generalized representation is a lot more intricate and complex. A light on the horizon is the usefulness of a *Quantum Schur Transform*: a physical realization of the unitary transformation from the computational basis to the Schur basis. Current quantum algorithms, such as Shor's and Grover's are based on the Quantum Fourier Transform. It has been shown that the quantum Schur transform can be accomplished efficiently on a quantum computer and opens the way for a whole new class of quantum algorithms [BCH06]. The generalized Majorana representation may form a valuable aid in understanding the physics and structure of these algorithms.

During this research, an unexpected duality has shown itself. The two transforms (Fourier and Schur), which both turn out to be important in quantum

[6] During the writing of this thesis, I have come upon the following paper (arXiv:0910.3075v1) [Kol09] published on arXiv.org which discusses essentially the same idea as presented here.

computation, are both fundamental in the two representations (Wigner and Majorana respectively) discussed in this thesis.

Bibliography

[BCH06] Dave Bacon, Isaac L. Chuang, and Aram W. Harrow. Efficient quantum circuits for schur and clebsch-gordan transforms. *Phys. Rev. Lett.*, 97(17):170502, Oct 2006.

[Dic54] R. H. Dicke. Coherence in spontaneous radiation processes. *Phys. Rev.*, 93(1):99, Jan 1954.

[Ful96] William Fulton. *Young Tableaux: With Applications to Representation Theory and Geometry.* Number ISBN 0521567246. Cambridge University Pres, 1996.

[GK06] T. Gopinath and Anil Kumar. Geometric quantum computation using fictitious spin- 12 subspaces of strongly dipolar coupled nuclear spins. *Phys. Rev. A*, 73(2):022326, Feb 2006.

[HN04] H. L. Haroutyunyan and G. Nienhuis. Resonances for coupled bose-einstein condensates. *Phys. Rev. A*, 70(6):063603, Dec 2004.

[HR18] G. H. Hardy and S. Ramanujan. Asymptotic formulae in combinatory analysis. *Proc. London Math. Soc.*, 2(19):75–115, 1918.

[KDD08] Piotr Kolenderski and Rafal Demkowicz-Dobrzanski. Optimal state for keeping reference frames aligned and the platonic solids. *Phys. Rev. A*, 78(5):052333, Nov 2008.

[Kol09] Piotr Kolenderski. Geometry of pure states of n spin-j system. arXiv:0707.3168, 2009.

[Lee88] C. T. Lee. Wehrl's entropy of spin states and lieb's conjecture. *Journal of Physics A: Mathematical and General*, 21(19):3749–3761, 1988.

[Maj32] E. Majorana. Atomi orientati in campo magnetico variabile. *Nuovo Cimento*, 9(2):43–50, 1932.

[Mal04] G. B. Malykin. Use of the poincare sphere in polarization optics and classical and quantum mechanics. review. *Radiophysics and Quantum Electronics*, 40(3):175–195, March 2004.

[Mih77] V V Mihailov. Addition or arbitrary number of identical angular momenta. *Journal of Physics A: Mathematical and General*, 10(2):147, 1977.

[MS07] H. Mäkelä and K.-A. Suominen. Inert states of spin-s systems. *Phys. Rev. Lett.*, 99(19):190408, Nov 2007.

[Orl03] M. Orlowski. 1-d random walk in presence of a totally reflecting barrier. *Physica Status Solidi (B)*, 239(1):158–167, 2003.

[Pau95] Ruben Pauncz. *The Symmetric Group in Quantum Chemistry*. Number ISBN 0849382912. CRC-Press, 1st edition, 1995.

[Ras77] M A Rashid. Addition of arbitrary number of identical angular momenta. *Journal of Physics A: Mathematical and General*, 10(8):L135, 1977.

[Sag01] B. E. Sagan. *The Symmetric Group: Representations, Combinatorial Algorithms, and Symmetric Functions*. Number ISBN 0387950672. Springer, 2nd edition, 2001.

[Sch65] J. Schwinger. *Quantum Theory of Angular Momentum*. Academic Press, New York, l. biedenharn and h. van dam edition, 1965.

[VPRK97] V. Vedral, M. B. Plenio, M. A. Rippin, and P. L. Knight. Quantifying entanglement. *Phys. Rev. Lett.*, 78(12):2275–2279, Mar 1997.

[WG03] Tzu-Chieh Wei and Paul M. Goldbart. Geometric measure of entanglement and applications to bipartite and multipartite quantum states. *Phys. Rev. A*, 68(4):042307, Oct 2003.

[Zim06] Jason Zimba. "anticoherent" spin states via the Majorana representation. *Electronic Journal of Theoretical Physics*, (10):143–156, 2006.

Appendices

Appendix A

Mathematical Results

What follows are a few derivations, some of the results are referred to from the thesis.

Proposition 1. *Let A be a diagonalizable operator. If $\mathrm{tr}\,(\rho A) = 0$ for any positive*[1] *operator ρ, then $A = 0$.*

Proof. Since A is diagonalizable, we can write $A = \sum_i a_i |i\rangle\langle i|$. Since the result has to hold for all positive ρ, we pick $\rho = |j\rangle\langle j|$ to find:

$$\mathrm{tr}\,(\rho A) = \sum_i a_i |\langle j|i\rangle|^2 = a_j = 0$$

So each a_i is zero and therefore $A = 0$. □

Proposition 2. *If $\mathrm{tr}\,(\rho A)$ is real for any positive operator ρ, then A is Hermitian.*

Proof. We split A into its Hermitian and anti-Hermitian parts: $A = B + iC$ with B, C both Hermitian. Then, since $\mathrm{tr}\,(\rho A) = \mathrm{tr}\,(\rho B) + i\mathrm{tr}\,(\rho C)$ is real, we have $\mathrm{tr}\,(\rho C) = 0$. Now apply proposition 1. □

Proposition 3. *If A and B are Hermitian matrices and $\mathrm{tr}\,(\rho A) = \mathrm{tr}\,(\rho B)$ for any positive matrix ρ, then $A = B$.*

Proof. Write the equation as $\mathrm{tr}\,(\rho(A - B)) = 0$ and apply proposition 1. □

Proposition 4. *If $\omega \neq 1$ is a complex number and $\omega^N = 1$, then:*

$$\frac{1}{N}\sum_{n=0}^{N-1} \omega^{nx} = \delta_N(x) := \begin{cases} 1 & \text{if } x = 0 \pmod N \\ 0 & \text{otherwise} \end{cases}$$

[1] A positive operator is one for which $\langle\psi|\rho|\psi\rangle \geq 0$ for all $|\psi\rangle$. If $\langle\psi|\rho|\psi\rangle > 0$ for all $|\psi\rangle$ we call it positive definite. This appears to be standard terminology in physics- but not in mathematics literature.

Proof. The sum is simply a geometric series.
Since $(1+\omega+\omega^2+\cdots+\omega^{N-1})(1-\omega)=1-\omega^N$, we have:

$$\sum_{n=0}^{N-1}\omega^n=\frac{1-\omega^N}{1-\omega}=0$$

Replacing ω with ω^x, the result follows immediately. □

Appendix B

The Radon Transform

B.1 Definition of the Radon transform

The Radon Transform was introduced in section 2.5 on page 13. For a function of two variables $u(x, k)$, the Radon transform $\mathcal{R}[u]$ of u takes a line λ in the plane and returns the integral of u along that line: $\mathcal{R}[u](\lambda) = \int_\lambda u d\lambda$. This is similar to what is done in, for example, medical X-ray tomography, where X-ray photos are shot from different directions through an object, see figure B.1. As the rays pass through the object, the beam is attenuated at a rate which depends on the density u of the tissue at each particular point. The intensity of the emerging beam is then recorded. From this, we can

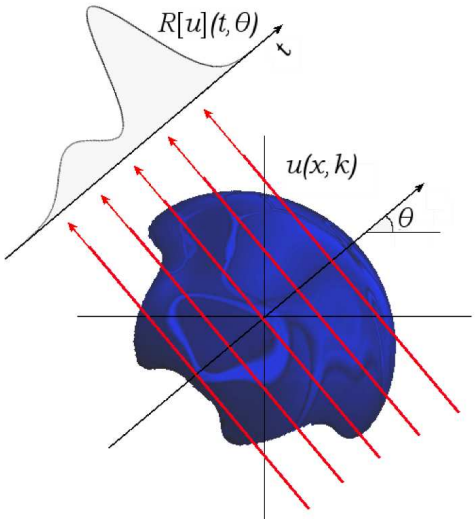


Figure B.1: The Radon transform gives the integral of your initial function along any line. The lines are parameterized by $t \in \mathbb{R}$ and $\theta \in [0, \pi)$.

learn the value of $\int_\lambda u d\lambda$ for each line λ. The function which takes a line and produces the integral of a function $u(x,k)$ along that line is called the Radon transform $\mathcal{R}[u]$ of u, i.e. $\mathcal{R}[u](\lambda) = \int_\lambda u d\lambda$. The problem is then to invert the transform and recover the function u from its projections which is, remarkably, possible!

Our first task is to parameterize the lines in the plane. It is clear that we need two parameters for a line. The choice $ax - by = t$ may seem natural, but is actually a poor choice. We would like to have a set of parallel lines to be parameterized by one parameter and the angle the lines in this parallel class makes with the x-axis by another. Parallel lines all share the same unit normal $\hat{n} = (\cos\theta, \sin\theta)$, so we use the angle θ the unit normal makes with the x-axis to specify the orientation of the line, as illustrated in figure B.2. We get a unique correspondence between a line through the origin and θ

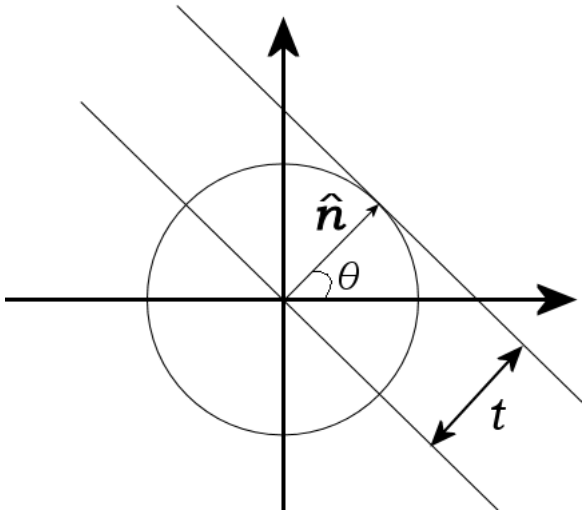


Figure B.2: Coordinates used in parameterizing the lines in the plane

if we restrict θ to the interval $[0, \pi)$, so that $\hat{n}$ never points into the lower half-plane. Any line parallel to the one through the origin with unit normal $\hat{n}$ can then be achieved by adding the displacement vector $t\hat{n}$ to each of its points, where $-\infty < t < \infty$.
So we identify to a line a pair (t, θ), where t is the signed distance of the line from the origin and θ is the angle the normal makes with the horizontal. The equation for the line given by (t, θ) is given by $\vec{r} \cdot \hat{n} = t$, or $x\cos\theta + k\sin\theta = t$. The Radon transform is then defined by:

$$\mathcal{R}[u](t,\theta) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} u(x,k)\delta(t - x\cos\theta - k\sin\theta)dxdk \qquad \text{(B.1)}$$

B.2 Inverting the Radon transform

It is most remarkable that equation (B.1) can be inverted to yield back $u(x,k)$. What is perhaps even more amazing is that the inverse solution becomes obvious by taking the *Fourier* transform of $\mathcal{R}[u]$ with respect to t (!). This will give rise to an intermediate function which we shall call $G(r,\theta)$. Thus:

$$G(r,\theta) = \mathcal{F}\left(\mathcal{R}[u]\right)(r,\theta)$$

The evaluation is simple:

$$G(r,\theta) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} u(x,k)e^{-irt}\delta(t - x\cos\theta - k\sin\theta)dxdkdt$$

We integrate this with respect to t:

$$G(r,\theta) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} u(x,k)e^{-i(xr\cos\theta + kr\sin\theta)}dxdk$$

If we let $a = r\cos\theta$ and $b = r\sin\theta$, then it is clear that G is nothing but the two dimensional Fourier transform of u at the point (a,b)(!), apart from a normalization factor. Therefore, if we let $\tilde{G}(a,b) = G(r,\theta)$ we can reconstruct u by taking the inverse (2D) Fourier transform of $\tilde{G}$.

$$u(x,k) = \frac{1}{(2\pi)^{3/2}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{i(ax+bk)}\tilde{G}(a,b)dadb$$

Appendix C

Phase Space Functions

In this part we will discuss two more phase space functions and their relations to each other and the Wigner function. Most of this discussion follows the treatment by Schleich [Sch01].
A classical distribution $P(x,p)$ in phase allows us to calculate the average value of any function $O(x,p)$ defined on the phase space by:

$$\langle O(x,p)\rangle = \int O(x,p)P(x,p)dx\,dp$$

In quantum mechanics, the observable becomes an operator $\hat{O}(\hat{x},\hat{p})$ and we wish to calculate the average value $\langle\hat{O}(\hat{x},\hat{p})\rangle$ by using a phase space distribution in analogy with the classical case. However, since $\hat{x}$ and $\hat{p}$ do not commute, it is not obvious how to obtain a phase space function representing $\hat{O}$. There are various ways to make this transition from $\hat{O}$ to a function $O(x,p)$ on the phase space, depending on different orderings of the operators.

C.1 The Q-Distribution

Even through the Wigner function is the most widely used distribution used to visualize a quantum state, it is not the only one. Here we describe the Q-distribution[1] with the help of coherent states. the Q-distribution has the nice property that it is everywhere positive and the distribution is easy to obtain from the density matrix. The downside is that the way observables transform to phase space is rather complex. The definition of the Q-distribution of a density matrix $\hat{\rho}$ is:

$$Q(z) = \frac{1}{\pi}\langle z|\hat{\rho}|z\rangle \tag{C.1}$$

Although we write $Q(z)$, we mean $Q(x,p)$ since Q is defined on the phase space. The identification is simply $z = x + ip$. So the Q-function at (x,p) is

[1] Also known as the Q-function or Husimi-Kano distribution.

proportional to the expectation value of the density operator in the coherent state $|z\rangle = |x + ip\rangle$. Because ρ is a positive operator, we see that Q is nowhere negative.

We have already mentioned that ordering of operators comes into play. If we consider an operator $\hat{O}$, we can write this as a function of the creation- and annihilation operators $\hat{a}^\dagger$ and a. With the help of the commutation relation $[a, a^\dagger] = 1$ we can always expand $\hat{O}$ in terms where products of creation- and annihilation operators are ordered like ${a^\dagger}^n a^m$, where the creation operators are always placed on the left of the annihilation operators. This is *normal ordering*. The ordering where the creation operators are placed on the right of the annihilation operators is called *antinormal ordering*. If we expand the density operator ρ normally:

$$\hat{\rho} = \sum_{n,m} \rho_{nm}^{(N)} {a^\dagger}^n a^m$$

Then we get for the average of $\hat{O}$:

$$\langle \hat{O} \rangle = \mathrm{tr}\left(\hat{\rho}\hat{O}\right) = \sum_{n,m} \rho_{nm}^{(N)} \mathrm{tr}\left({a^\dagger}^n a^m \hat{O}\right)$$

Because $\mathrm{tr}\left({a^\dagger}^n a^m \hat{O}\right) = \mathrm{tr}\left(a^m \hat{O} {a^\dagger}^n\right)$ We see that if we use the antinormal ordering expansion of $\hat{O}$, then the powers of $a^\dagger$ and a add, and we get:

$$\hat{O} = \sum_{k,l} O_{kl}^{(A)} a^k {a^\dagger}^l \tag{C.2}$$

$$\langle \hat{O} \rangle = \sum_{n,m,k,l} \rho_{nm}^{(N)} O_{kl}^{(A)} \mathrm{tr}\left({a^\dagger}^{n+l} a^{m+k}\right)$$

The trace is evaluated with the closure relation of coherent states:

$$\mathrm{tr}\left({a^\dagger}^{n+l} a^{m+k}\right) = \frac{1}{\pi} \int \langle z | {a^\dagger}^{n+l} a^{m+k} | z \rangle d^2 z = \frac{1}{\pi} \int {z^*}^{n+l} z^{m+k} d^2 z$$

Then the expression for $\langle \hat{O} \rangle$ can be written as:

$$\langle \hat{O} \rangle = \frac{1}{\pi} \int \left(\sum_{n,m} \rho_{n,m}^{(N)} {z^*}^n z^m \right) \left(\sum_{k,l} O_{kl}^{(A)} z^k {z^*}^l \right) d^2 z$$

If we recognize that:

$$Q(z) = \frac{1}{\pi} \langle z | \hat{\rho} | z \rangle = \frac{1}{\pi} \sum_{n,m} \rho_{nm}^{(N)} {z^*}^n z^m$$

And we define the phase space function:

$$O(z) = \sum_{k,l} O_{kl}^{(A)} z^k z^{*l},$$

found by substituting $a \to z$ and $a^\dagger \to z^*$ in the antinormal expansion of $\hat{O}$, then $\langle \hat{O} \rangle$ can be written as:

$$\langle \hat{O} \rangle = \int Q(z) O(z) d^2 z,$$

as desired.

Although every operator can be expanded in the coherent states by using the closure relation twice:

$$\hat{O} = \frac{1}{\pi^2} \iint d^2 z d^2 w \langle w|\hat{O}|z\rangle |w\rangle\langle z|$$

it is actually possible to have a *diagonal* representation, owing to the overcompleteness of the coherent states. This is actually what the antinormal ordering accomplishes. If we insert the closure relation between a^k and $a^{\dagger l}$ in equation (C.2), we get:

$$\hat{O} = \sum_{k,l} O_{kl}^{(A)} a^k a^{\dagger l} = \frac{1}{\pi} \int \sum_{k,l} O_{kl}^{(A)} z^k z^{*l} |z\rangle\langle z| = \frac{1}{\pi} \int O^{(A)} |z\rangle\langle z|$$

C.2 The P-distribution

It is clear from the previous section that another phase space distribution is easily found by taking the antinormal ordering expansion of $\hat{\rho}$ and the normal ordering expansion of $\hat{O}$:

$$\hat{\rho} = \sum_{n,m} \rho_{nm}^{(A)} a^n a^{\dagger m}$$

$$\hat{O} = \sum_{k,l} \rho_{kl}^{(N)} a^{\dagger k} a^l$$

A similar calculation as before gives:

$$\langle \hat{O} \rangle = \frac{1}{\pi} \int \left(\sum_{n,m} \rho_{n,m}^{(A)} z^n z^{*m} \right) \left(\sum_{k,l} O_{kl}^{(N)} z^{*k} z^l \right) d^2 z$$

This time the phase space function of $\hat{O}$ is defined as:

$$O(z) = \langle z|\hat{O}|z\rangle,$$

and we define the phase space distribution of $\hat{\rho}$:

$$P(z) = \frac{1}{\pi} \sum_{n,m} \rho^{(A)}_{n,m} z^n z^{*m}$$

which can be obtained from the antinormal ordering expansion of $\hat{\rho}$ by replacing $a \to z$ and $a^\dagger \to z^*$. This phase space distribution is called the P-distribution[2]. We see that the transformation roles of the density operator and the observables is interchanged with respect to the Q-distribution. In this case, it is the phase space function of $\hat{O}$ that is the expectation in a coherent state, and the distribution $\hat{\rho}$ are the coefficients when expanded diagonally in the coherent states:

$$\hat{\rho} = \int P(z)|z\rangle\langle z| \tag{C.3}$$

Although this looks very nice formally, the P-function is generally highly singular and not a true probability distribution; it has to be treated as a *generalized function*, or *distribution* in the functional analysis sense. For example, it is immediate from equation (C.3) that for a coherent state $\hat{\rho} = |\alpha\rangle\langle\alpha|$, we have $P(z) = \delta(z - \alpha)$. For other often used states, like number states or squeezed states, P can become even more singular, involving arbitrary high derivatives of delta functions.

C.3 The Q-distribution on a sphere

Since by definition, the Q-distribution gives the expectation value of a density matrix $\hat{\rho}$ between coherent states up to normalization (see equation C.1), the Q-distribution can be straightforwardly generalized to the sphere for spin systems by taking the expectation value between spin coherent states (Appendix D):

$$Q(\theta, \phi) = \frac{2j+1}{4\pi} \langle \theta, \phi | \hat{\rho} | \theta, \phi \rangle$$

Where the normalization constant is taken into account.

More generally, a Q-distribution can in principle be defined for generalized coherent states [Per72], but the corresponding P-representation for which the operators should transform can be ill-behaved under non-pathological conditions [MACS03].

Bibliography

[MACS03] N. Mukunda, Arvind, S. Chaturvedi, and R. Simon. Generalized coherent states and the diagonal representation for operators. *Journal of Mathematical Physics*, 44(6):2479–2506, 2003.

[2] Also called the Glauber-Sudarshan P-distribution.

[Per72] A. M. Perelomov. Coherent states for arbitrary lie group. *Commun. Math. Phys.*, 26(3):222–236, 1972.

[Sch01] W. P. Schleich. *Quantum Optics in Phase Space*. Number ISBN: 352729435X. Wiley-VCH, 1st edition, February 2001.

Appendix D

Spin Coherent States

Because of the usefulness and clear physical meaning of coherent states introduced by Glauber [Gla63b, Gla63a] and Sudarshan [Sud63] for EM-field modes in 1963, different ways to generalize the concept of coherent to arbitrary quantum systems have been introduced. In 1972 Perelomov [Per72] and Gilmore [Gil72] introduced a generalization, allowing the construction of coherent states for an arbitrary Lie Group[1] which share many properties analogues to the Glauber coherent states.

D.1 Generalized coherent states

Here we will give a rough and intuitive overview of the construction of generalized coherent states[2]. The basic setting is the Hilbert space of the quantum system together with a dynamical (Lie) group acting on the space. What this Lie group is will depend on the system in question and different groups will generally result in a different set of generalized coherent states. For example, in the original setting of a single mode EM-field (or a one-dimensional harmonic oscillator), we have the displacement operators: $\hat{D}(\alpha) = \exp\left(\alpha a^\dagger - \alpha^* a\right)$. The set of all displacement operators forms a Lie group acting on the space. Then all we need is some 'reference' state, which may be particular to the problem. Different choices of the reference state will generally create a different set of coherent states. After this choice, coherent states are simply the *displaced reference states*. In other words: the set of coherent states is the *orbit* of the reference state under the action of the Lie group. Again, in the case of Glauber coherent states, we choose the

[1] A Lie group is a group in which the group elements can be smoothly parameterized. The group of translations and the group rotations of a space are prime examples of Lie groups.

[2] This discussion is highly simplified and intended to give a general idea of generalized coherent states. For a more rigorous and much more detailed treatment of coherent states in general, see [ZFG90]

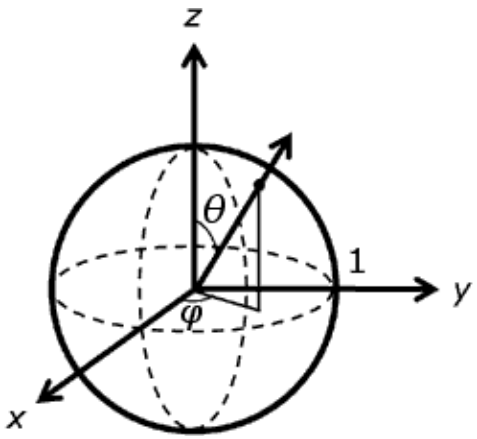


Figure D.1: Definition of the spherical coordinate angles.

vacuum state $|0\rangle$ as our reference state and the set of coherent states are the states $|\alpha\rangle = \hat{D}(\alpha)|0\rangle$ for all $\alpha \in \mathcal{C}$. If we had chosen a number state $|n\rangle$ as our reference state, yielding displaced number states as our set of 'coherent states'.
It can be shown that generalized coherent states generated this way obey two familiar properties:

1. They are non-orthogonal (except perhaps on a set of measure zero)

2. They are overcomplete and satisfy a closure relation:

$$\int dx|x\rangle\langle x| \propto I$$

D.2 Coherent states of the rotation group

The most used form of coherent states, after the Glauber coherent states, are obtained from considering a spin j system. The Lie group is the group of rotations and the reference state is a state of *maximum angular momentum* along some direction, say $|j, j\rangle$. These states are called spin coherent states (SCS's)[3] and were originally introduced by Radcliffe [Rad71] and Arecchi et al [ACGT72] at around same time the construction of generalized coherent states was introduced.

SCS's are rotations of the state $|j, j\rangle$, which is an eigenstate of $\vec{e}_z \cdot \vec{\hat{J}} = \hat{J}_z$ corresponding to the maximum eigenvalue $m = j$. Rotated versions will naturally be eigenstates of the angular momentum along some direction $(\vec{n} \cdot \vec{\hat{J}})$

[3] These are also often referred to as atomic coherent states, and sometimes as angular- or Bloch coherent states.

corresponding to the maximum eigenvalue j. We can therefore represent each SCS as a point on a sphere: the point signifying the direction along which the angular momentum is maximal. There are three representations that we will often use to specify such a point:

1. By the spherical coordinate angles (θ, ϕ), as in figure D.1. The SCS is written $|\theta, \phi\rangle$.

2. By a unit vector $\vec{r} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$. The SCS is written $|\vec{r}\rangle$.

3. By a complex number w obtained by stereographic projection from the south pole, see figure D.2. The triangle formed by the south pole, the point (θ, ϕ) and the origin is equilateral with base angle $\theta/2$. By elementary trigonometry, considering the right triangle formed by the origin, the south pole and the projected point w, we have $|w| = \tan\frac{\theta}{2}$. It is also not hard to see that $\arg(w) = \phi$. So we have the correspondence:
$$w = \frac{x + iy}{1 + z} = \tan\frac{\theta}{2}\, e^{i\phi}$$
The SCS is then written as $|w\rangle$, where w can take values in the extended complex plane $\mathcal{C} \cup \{\infty\}$. The point at infinity corresponds to the south

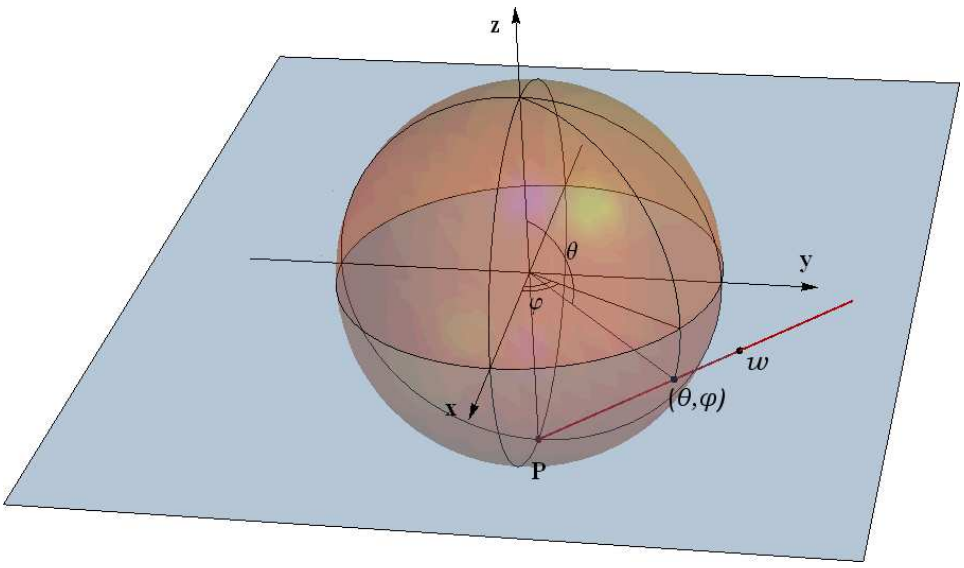


Figure D.2: Stereographic projection from the south pole. The point (θ, ϕ) on the sphere is mapped to the point $w = \tan\frac{\theta}{2}e^{i\phi}$ on the plane. Geometrically, w is the intersection of the plane with the line through the south pole P and the point (θ, ϕ)

pole on the sphere. Projection from the south pole is taken as to generalize known form $\cos\frac{\theta}{2}|0\rangle + \sin\frac{\theta}{2}e^{i\phi}|1\rangle$ for the Bloch sphere.

D.3 Representing rotations of the sphere

As mentioned before, all SCS's are rotated states of $|j,j\rangle$ ($|\vec{e}_z\rangle$ or $|0\rangle$ in the vector and complex notations). An arbitrary rotation can be given by a composing 3 rotations. First a rotation of γ about the z-axis, followed by a rotation of β about the y-axis and finally again a rotation about the z-axis of angle α. The three angles are the Euler angles. Alternatively we can specify an axis $\vec{n}$ and an angle ψ over which to rotate. For calculational purposes it is useful to know how the representations $\vec{r}$, and w act under a rotation.

We start by finding an expression of where the vector $\vec{r}$ goes after a rotation about $\vec{n}$ and angle ψ. We call the rotated vector $\vec{r}'$. First we note that, unless $\vec{r}$ and $\vec{n}$ are parallel, the three vectors $\vec{n}$ and $\vec{n}\times\vec{r}$ and $\vec{r}-(\vec{n}\cdot\vec{r})\vec{n}$ form an orthogonal basis. The expansion of $\vec{r}'$ in this basis is simply[4]:

$$\begin{aligned}\vec{r}' &= \cos\psi(\vec{r}-(\vec{r}\cdot\vec{n})\vec{n}) + \sin\psi(\vec{n}\times\vec{r}) + (\vec{r}\cdot\vec{n})\vec{n} \\ &= \cos\psi\vec{r} + (1-\cos\psi)(\vec{r}\cdot\vec{n})\vec{n} + \sin\psi(\vec{n}\times\vec{r})\end{aligned}$$

From the theory of spins and representations groups of rotations, we know that a rotation can be represented by an $SU(2)$ matrix:

$$U = \begin{pmatrix} \alpha & \beta \\ -\beta^* & \alpha^* \end{pmatrix}, \qquad |\alpha|^2 + |\beta|^2 = 1 \tag{D.1}$$

The rotation about axis $\vec{n}$ and angle ψ is represented by:

$$U = \exp\left(-i\psi\vec{n}\cdot\hat{\vec{\sigma}}/2\right) = \cos\frac{\psi}{2}I - i\sin\frac{\psi}{2}(\hat{n}\cdot\vec{\sigma})$$

Rotation of the unit sphere leads to a transformation of the complex plane through stereographic projection. The corresponding transformation is a *unitary Möbius transformation*. The Möbius transformation corresponding to equation D.1 is:

$$f_U(w) = w' = \frac{\alpha^* w - \beta^*}{\beta w + \alpha}, \qquad |\alpha|^2 + |\beta|^2 = 1$$

[4] It is easy to see when you compare with a rotation about the z-axis where the image is $\vec{r}' = \cos\psi\vec{e}_x + \sin\psi\vec{e}_y + (\vec{r}\cdot\vec{e}_z)\vec{e}_z$. The vector $\vec{n}$ plays the role of the z-axis and $\vec{n}\times\vec{r}$ and $\vec{r}-(\vec{n}\cdot\vec{r})\vec{n}$ span the plane perpendicular to the rotation axis.

D.4 Expressions of spin coherent states

As mentioned before, all SCS's are rotated states of $|j,j\rangle$ ($|\vec{e}_z\rangle$ or $|0\rangle$ in the vector and complex notations). An arbitrary rotation can be given in terms of the three Euler angles:

$$\hat{R}(\alpha,\beta,\gamma) = \exp(-i\alpha\hat{J}_z)\exp(-i\beta\hat{J}_y)\exp(-i\gamma\hat{J}_z)$$

We can rotate the state $|j,j\rangle$ on the north pole to $|\theta,\phi\rangle$ by the rotation $\hat{R}(\phi,\theta,-\phi)$, so:

$$|\theta,\phi\rangle = \hat{R}(\phi,\theta,-\phi)|j,j\rangle$$

To find the coefficients $\langle j,m|\theta,\phi\rangle$ we note that:

$$\begin{aligned}\langle j,m|\theta,\phi\rangle &= \langle j,m|\exp(-i\phi J_z)\exp(-i\theta J_y)\exp(i\phi J_z)|j,j\rangle \\ &= e^{i\phi(j-m)}d^j_{m,j}(\theta)\end{aligned} \tag{D.2}$$

where $d^j_{m,m'}(\beta) = \langle j,m|\exp(-i\beta J_y)|j,m'\rangle$ is the small Wigner d-matrix. The general expression is [Ros67]:

$$\begin{aligned}d^j_{m'm}(\beta) &= [(j+m')!(j-m')!(j+m)!(j-m)!]^{1/2}(\cos\frac{\beta}{2})^{2j} \\ &\times\sum_s \frac{(-1)^{m'-m+s}}{(j+m-s)!s!(m'-m+s)!(j-m'-s)!}\left(\cos\frac{\beta}{2}\right)^{m-m'-2s}\left(\sin\frac{\beta}{2}\right)^{m'-m+2s}\end{aligned}$$

where the sum over s taken such that no factorials of negative numbers occur. Fortunately, we don't need the entire expression as we are only interested in the case where $m' = j$. We get:

$$d^j_{j,m}(\theta) = \binom{2j}{j+m}^{1/2}\left(\cos\frac{\theta}{2}\right)^{j+m}\left(\sin\frac{\theta}{2}\right)^{j-m} \tag{D.3}$$

With the coefficients in hand from equations D.2 and D.3. We can write down the expansion of $|\theta,\phi\rangle$ in the basis $\{|j,m\rangle\}$:

$$|\theta,\phi\rangle = \sum_{m=-j}^{j}\binom{2j}{j+m}^{1/2}\left[\cos\frac{\theta}{2}\right]^{j+m}\left[\sin\frac{\theta}{2}\right]^{j-m}e^{i(j-m)\phi}|j,m\rangle$$

Or, in the complex number representation:

$$\begin{aligned}|w\rangle &= \frac{1}{(1+|w|^2)^j}\sum_{m=-j}^{j}\binom{2j}{j+m}^{1/2}w^{j-m}|j,m\rangle \\ &= \frac{1}{(1+|w|^2)^j}\sum_{m=-j}^{j}\frac{(wJ^-)^{j+m}}{(j+m)!}|j,j\rangle \\ &= \frac{1}{(1+|w|^2)^j}\exp(wJ^-)|j,j\rangle\end{aligned}$$

From this, the analogy with the Glauber coherent states becomes apparent. To further the analogy, let's find the SCS counterpart to the displacement operator in terms of the 'creation' and 'annihilation' operators J_- and J_+:

$$\begin{aligned}\hat{R}(\phi,\theta,-\phi) &= \exp(-i\phi\hat{J}_z)\exp(-i\theta\hat{J}_y)\exp(i\phi\hat{J}_z) \\ &= \exp\left[-i\theta(\hat{J}_y\cos\phi - \hat{J}_x\sin\phi)\right] = \exp(\alpha\hat{J}^- - \alpha^*\hat{J}^+)\end{aligned}$$

Where in the last expression $\alpha = \frac{1}{2}\theta e^{i\phi}$. Another way to write the SCS is therefore:

$$|\theta,\phi\rangle = \exp(\alpha\hat{J}^- - \alpha^*\hat{J}^+)|j,j\rangle,$$

which resembles the expression for the Glauber coherent states:

$$|\alpha\rangle = \exp(\alpha\hat{a}^\dagger - \alpha^*\hat{a})|0\rangle$$

D.5 Properties of spin coherent states

As mentioned before, the SCS's share some properties common to the familiar Glauber coherent states. One is that they form an overcomplete set. The overlap of two coherent states $|z\rangle = |\theta,\phi\rangle$ and $|w\rangle = |\theta',\phi'\rangle$ is given by:

$$\begin{aligned}\langle w|z\rangle &= \left[\frac{(1+w^*z)^2}{(1+|z|^2)(1+|w|^2)}\right]^j \\ &= e^{i(\phi-\phi')j}\left[\cos\frac{1}{2}(\theta-\theta')\cos\frac{1}{2}(\phi-\phi') - i\cos\frac{1}{2}(\theta+\theta')\sin\frac{1}{2}(\phi-\phi')\right]^{2j}\end{aligned}$$

And their modulus square can be written in a geometrical and coordinate-independent form:

$$|\langle\vec{r}\,'|\vec{r}\rangle|^2 = \left(\frac{1+\vec{r}\cdot\vec{r}\,'}{2}\right)^{2j} = \left(\cos\frac{1}{2}\Theta\right)^{4j}, \tag{D.4}$$

where Θ is the angle between $\vec{r}'$ and $\vec{r}$. Geometrically, $\frac{1+\vec{r}\cdot\vec{r}'}{2}$ is half the distance from the center of the sphere to the chord joining the two points. The SCS's satisfy a (non-unique) closure relation:

$$\frac{2j+1}{4\pi}\int_{S_2} d\Omega|\theta,\phi\rangle\langle\theta,\phi| = 1$$

For the corresponding relation on the complex plane we compute the Jacobian of the transformation:

$$d\Omega = \sin\theta d\theta d\phi = \frac{4d^2w}{(1+|w|^2)^2}$$

$$\frac{2j+1}{\pi}\int \frac{|w\rangle\langle w|}{(1+|w|^2)^2}d^2w = 1$$

The physical relation between the vector $\vec{r}$ of the SCS $|\vec{r}\rangle$ and its angular momentum is neatly expressed by the expectation value of the angular momentum vector operator $\hat{\vec{J}} = (\hat{J}_x, \hat{J}_y, \hat{J}_z)$:

$$\langle\vec{r}|\hat{\vec{J}}|\vec{r}\rangle = j\vec{r}$$

D.6 Spin coherent states in N-qubit systems

For a spin 1/2 system (or more generally, a qubit), every state is a spin coherent state. For higher values of spin, this is obviously no longer true. When we combine N spin 1/2 particles, the total angular momentum j can take values from 0 or $\frac{1}{2}$ (depending on whether N is even or odd respectively) to $N/2$ in integer steps. The subspace of $j_{max} = N/2$ is non-degenerate and is precisely the symmetric subspace[5] $\mathcal{H}^{(s)}$ of the total Hilbert space. This follows immediately by realizing that $|j_{max}, j_{max}\rangle$ is the product state where are spin are 'up' in the z-direction. This state is symmetric and the other $|j_{max}, m\rangle$ states are generated by application of the (symmetric) lowering operator $\hat{J}_- = \sum \hat{J}_-^{(i)}$, where $\hat{J}_-^{(i)}$ acts only in the subspace on the i-th particle. A calculation quickly reveals that $|j_{max}, m\rangle$ is the symmetrized state where $j+m$ spins are 'up' and $j-m$ are 'down'. For example, when $N = 3$:

$$|\frac{3}{2}, \frac{3}{2}\rangle = |000\rangle$$

$$|\frac{3}{2}, \frac{1}{2}\rangle = \frac{1}{\sqrt{3}}\left(|001\rangle + |010\rangle + |100\rangle\right)$$

$$|\frac{3}{2}, \frac{-1}{2}\rangle = \frac{1}{\sqrt{3}}\left(|011\rangle + |101\rangle + |110\rangle\right)$$

$$|\frac{3}{2}, -\frac{3}{2}\rangle = |111\rangle$$

These are exactly the projections of the computational basis states $\{|x_1 x_2 \ldots x_n\rangle\}$ (see section 1.2) onto $\mathcal{H}^{(s)}$. Therefore the subspace of maximum j is precisely the symmetric subspace[6].

Every coherent state is obtained by a rotation of $|j_{max}, j_{max}\rangle = |0\rangle^{\otimes N}$. If the unitary representation of the rotation on a single qubit is given by $\hat{U}$,

[5] A symmetric state in this context means unchanged under any permutation of the particle states.

[6] This observation allows one to treat N-boson systems as an equivalent spin system with $j = N/2$. The fictitious spin states $|j, m\rangle$ are called Dicke states and were introduced in Dicke's paper on superradiance [Dic54].

then in the N-qubit system it takes the product form $\hat{U}^{\otimes N}$. If the rotation maps the North Pole to the general point $\vec{r}$, then:

$$|j_{max}, \vec{r}\rangle = \hat{U}^{\otimes N}|0\rangle^{\otimes N} = |\vec{r}\rangle^{\otimes N}$$

Therefore, every coherent state is a symmetric product state[7]. Conversely, a symmetric product state must be of the form $|\phi\rangle^{\otimes N}$. Since every qubit state is a coherent state $|\vec{r}\rangle$ for some $\vec{r}$, every symmetric product state is also a coherent state[8]. This is a feature of the 'classicality' of coherent states. The Glauber coherent states can be characterized as 'states of minimum' uncertainty: they obtain the bound of the Heisenberg uncertainty relation for the quadrature operators $\hat{x}$ and $\hat{p}$. Coherent states are in a sense the most *classical* quantum states. It is interesting that the SCS's are classical in the sense that, for composite systems, they are all non-entangled states. Furthermore, every state that is orthogonal to all symmetric states is entangled. To see why this is so, consider that for a product state $|\psi\rangle = |\vec{s}_1\rangle \otimes \cdots \otimes |\vec{s}_N\rangle$ and a coherent state $|R\rangle = |\vec{r}\rangle^{\otimes N}$ we have:

$$\langle R|\psi\rangle = \prod_{i=1}^{N} \langle \vec{r}|\vec{s}_i\rangle$$

It is always possible to find an $\vec{r}$ different from $-\vec{s}_i$ for $i = 1, \ldots, N$ to make the inner product nonzero. This means that $|\psi\rangle$ does not lie in the orthogonal complement of $\mathcal{H}^{(s)}$[9].

Bibliography

[ACGT72] F. T. Arecchi, Eric Courtens, Robert Gilmore, and Harry Thomas. Atomic coherent states in quantum optics. *Phys. Rev. A*, 6(6):2211–2237, Dec 1972.

[DB09a] P. Dömötör and M. G. Benedict. Global entanglement and coherent states in an -partite system. *The European Physical Journal D - Atomic, Molecular, Optical and Plasma Physics*, 53(2):237–242, Jun 2009.

[7] This trivial result can just as easily be generalized to a composite system of arbitrary spins, since clearly, the spin is maximal precisely when all the subsystem-spins are coherent and pointing in the same direction.

[8] This is no longer true for quNits with $N > 2$. For example, the symmetric product state $|1,0\rangle|1,0\rangle$ of two spin 1 particles is written in the total angular momentum basis as $\sqrt{\frac{2}{3}}|2,0\rangle - \frac{1}{\sqrt{3}}|0,0\rangle$ which has no well defined value of j.

[9] These results were also recently published in [DB09a] using a more elaborate method. Even more recently, two publications were published [DB09b] with claim the result was generalized to quNits, in contrast to the counterexample $|1,0\rangle|1,0\rangle$. I've not yet investigated these papers in depth, although the discrepancy is suspected to come from a difference in the definition.

[DB09b] Piroska Dömötör and Mihaly G Benedict. Entanglement and coherent states. *Physica Scripta*, T135:014030 (3pp), 2009.

[Dic54] R. H. Dicke. Coherence in spontaneous radiation processes. *Phys. Rev.*, 93(1):99, Jan 1954.

[Gil72] Robert Gilmore. Geometry of symmetrized states. *Annals of Physics*, 74(2):391 – 463, 1972.

[Gla63a] Roy J. Glauber. Coherent and incoherent states of the radiation field. *Phys. Rev.*, 131(6):2766–2788, Sep 1963.

[Gla63b] Roy J. Glauber. Photon correlations. *Phys. Rev. Lett.*, 10(3):84–86, Feb 1963.

[Per72] A. M. Perelomov. Coherent states for arbitrary lie group. *Commun. Math. Phys.*, 26(3):222–236, 1972.

[Rad71] J M Radcliffe. Some properties of coherent spin states. *Journal of Physics A: General Physics*, 4(3):313–323, 1971.

[Ros67] M. E. Rose. *Elementary theory of angular momentum*. Number ISBN 0486684806. John Wiley & Sons, 1967.

[Sud63] E. C. G. Sudarshan. Equivalence of semiclassical and quantum mechanical descriptions of statistical light beams. *Phys. Rev. Lett.*, 10(7):277–279, Apr 1963.

[ZFG90] Wei-Min Zhang, Da Hsuan Feng, and Robert Gilmore. Coherent states: Theory and some applications. *Rev. Mod. Phys.*, 62(4):867–927, Oct 1990.