Implementing Grover's search algorithm using the one-way quantum computing model and photonic orbital angular momentum

by

Humairah Bassa

Submitted in fulfilment of the academic requirements for the degree of
Master of Science in the School of Physics,
University of KwaZulu-Natal, Durban

October, 2011

As the candidate's supervisor I have approved this dissertation for submission.

Signed: _____ Name: _____ Date: _____

UNIVERSITY OF
KWAZULU-NATAL

Abstract

Standard quantum computation proceeds via the unitary evolution of physical qubits (two-level systems) that carry the information. A remarkably different model is one-way quantum computing where a quantum algorithm is implemented by a set of irreversible measurements on a large array of entangled qubits, known as the cluster state. The order and sequence of these measurements allow for different algorithms to be implemented. With a large enough cluster state and a method in which to perform single-qubit measurements the desired computation can be realised.

We propose a potential implementation of one-way quantum computing using qubits encoded in the orbital angular momentum degree of freedom of single photons. Photons are good carriers of quantum information because of their weak interaction with the environment and the orbital angular momentum of single photons offers access to an infinite-dimensional Hilbert space for encoding information. Spontaneous parametric down-conversion is combined with a series of optical elements to generate a four-photon orbital angular momentum entangled cluster state and single-qubit measurements are carried out by means of digital holography. The proposed set-up, which is based on an experiment that utilised polarised photons, can be used to realise Grover's search algorithm which performs a search through an unstructured database of four elements. Our application is restricted to a two-dimensional subspace of a multi-dimensional system, but this research facilitates the use of orbital angular momentum qubits for quantum information processing and points towards the usage of photonic qudits (multi-level systems).

We also review the application of Dirac notation to paraxial light beams on a classical and quantum level. This formalism is generally employed in quantum mechanics but the analogy with paraxial optics allows us to represent the classical states of light by means of Dirac kets. An analysis of the analogy between the classical and quantum states of light using this formalism, is presented.

Preface

The work described in this dissertation was carried out in the School of Physics, University of KwaZulu-Natal, Durban, from February 2010 to October 2011, under the supervision of Prof. Thomas Konrad.

The studies represent original work by the author and have not otherwise been submitted in any form for any degree or diploma to any tertiary institution. Where use has been made of the work of others it is duly acknowledged in the text.

Declaration 1 – Plagiarism

I, **Humairah Bassa** declare that

1. The research reported in this dissertation, except where otherwise indicated, is my original research.

2. This dissertation has not been submitted for any degree or examination at any other university.

3. This dissertation does not contain other persons' data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons.

4. This dissertation does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:

    a. Their words have been rewritten but the general information attributed to them has been referenced.
    b. Where their exact words have been used, then their writing has been placed in italics and inside quotation marks, and referenced.

5. This dissertation does not contain text, graphics or tables copied and pasted from the Internet, unless specifically acknowledged, and the source being detailed in the dissertation and in the Bibliography.

Signed: _____

Declarations 2 – Publications

Publication 1

H. Bassa and T. Konrad, "Representation of the quantum and classical states of light carrying orbital angular momentum" in Proceedings of SAIP2011, the 56th Annual Conference of the South African Institute of Physics, edited by I. Basson and A.E. Botha (University of South Africa, Pretoria, 2011), pp. 752 - 757 ISBN: 978-1-86888-688-3.

Signed: _____

Acknowledgements

My first thanks must go to Prof. Thomas Konrad who has persevered with me through the eventful years of both my Honours and Masters projects. His enthusiasm and passion for science has stimulated my interest in the field. His advice, patience and encouragement, particularly in the difficult times, has been greatly appreciated.

I am incredibly grateful to Dr. Stef Roux from the National Laser Centre, who has always been available to answer my many (and sometimes very obvious) questions either in person or via email. I would also like to thank him for his helpful comments on my progress reports during the last two years and on one or two of the chapters in this dissertation.

I must express my gratitude to the National Institute of Theoretical Physics for the financial support that made this research possible.

Thanks to my co-supervisor, Prof. Andrew Forbes for his continuing assistance and for allowing me to interact with his diverse group at the National Laser Centre from whom I have gained indispensable theoretical and experimental knowledge. I am particularly grateful to Ms. Angela Dudley for taking me under her wing during my visit to the National Laser Centre and for providing me with valuable advice and many helpful documents.

A heartfelt thank you to Dr. Sujit Choudhary for taking on the daunting task of proofreading this dissertation during his busy schedule and for many interesting conversations.

Finally, I could not have reached this point without the continual support and love of my family and friends. Thanks to my mother, Rashida, and father, Farouk, – your example has given me the courage to do what I love, the strength to survive the difficult times, and the confidence to believe in myself. To my siblings and friends, thanks for the welcome distractions of movie nights, picnics in the park and especially for dealing with my mood swings and unavailability at times.

Contents

List of Figures

Chapter 1

Introduction

Quantum mechanics is a set of laws of nature that govern the behaviour of small objects like atoms (the building blocks of matter) and photons (the building blocks of light). These laws are strikingly different from those that determine the behaviour of large objects that we see in everyday life. There is experimental evidence which demonstrates that the physics of a quantum particle cannot be accurately described by a theory which is local and realistic (cf. for example, [1]). In the quantum regime we experience fascinating features of nature which include quantum superposition and quantum entanglement. Quantum superposition refers to the ability of a quantum particle to exist in a superposition of a number of characteristic states simultaneously; and, when a measurement is performed, the superposition collapses into one of these states with a certain probability. Quantum entanglement allows for the characteristic states of two particles to be linked in such a way that measuring the state of one instantaneously defines the state of the other, even if they are separated by arbitrarily large distances. Although this feature was termed "*spooky action at a distance*" by Albert Einstein [2], it does not allow for superluminal communication.

Over the last 20 years, physicists have realised that the astonishing properties of quantum mechanics could be harnessed in many physical applications. Of special interest to us is the development of incredibly powerful computers. Why should we be thinking about developing these technologies at this point in time? The reason is two-fold. Nowadays, silicon chips made up of transistors are the essential components of any computer. Each of the transistors acts as a switch and is used to represent a binary digit (or *bit*) of information: so the "off" position represents a "0" and the "on" position represents a "1". The number of transistors on a chip doubles approximately every two years, according to an empirically based claim known as *Moore's law* [3]. This means that transistors have to continually shrink in size and we can estimate that by the year 2020 one bit of information will be carried by a single atom or particle, which leads to the quantum regime. Therefore, the computing industry is already heading in the direction of quantum computation even though some do not acknowledge this trend at the moment. The second reason is that quantum computers will be extremely powerful, and increase the speed of the computers we have today, not linearly, but exponentially. One of the first algorithms demonstrating the power of quantum computers was devised by Peter Shor in 1994 [4]. Shor's algorithm allows one to find the prime factors of an arbitrarily large integer in polynomial time and this problem is believed to have no efficient solution on a classical computer [5]. The computational power of quantum

computers might help with the prediction of weather and natural disasters, tracing the motion of stars and planets and the simulations of events like earthquakes [5]. With interest from the military, medicine and science fields these computers will possibly make a bigger impact in this information age than the computers that exist now did.

Why will quantum computers be more powerful than their classical counterparts? Information in a quantum computer is encoded as quantum bits (or *qubits*) which, unlike the bit, can exist in a superposition of the states "0" or "1". This superposition of qubits gives quantum computers their inherent parallelism. While an N-bit register can be in one of 2^N states only, an N-qubit register can be in a superposition of all 2^N possible states. An operation on the N-bit register would allow us to evaluate a function at only one of the 2^N possible values while a unitary operation on an N-qubit register would enable the evaluation of the function at each of the 2^N values simultaneously. Consequently, a quantum computer will succeed in solving certain problems qualitatively faster than is possible on a classical computer. An example of such a task is a search through an unstructured database which is done on a quantum computer using Grover's search algorithm. This algorithm utilises quantum superposition to search through different parts of the database at the same time. To solve the search problem, a quantum register is set in a superposition of all the elements in the database and by simple quantum operations on the register, the correct solution can be obtained with a high probability. Classically, to find the correct solution one would need to search through at least half of the items in the database on average whereas Grover's algorithm would provide a quadratic speed-up by performing the search in only a small number of steps (of the order $O(\sqrt{N})$ if there are N elements in the database). Grover's algorithm is very useful in computation and cryptography where fast searching is essential.

The basic principle of the standard model of quantum computation is that we begin with a quantum register (made up of several qubits) which contains all the input information. By performing specific unitary operations (or quantum logic gates) on this input register, information can be processed and at the end of the computation the result is read out by measurements on the qubits in the output register. This model can, in principle, be simulated by a classical computer, albeit inefficiently.

In 2001, Hans Briegel and his student Robert Raussendorf developed a remarkably different model for quantum computation, called one-way quantum computing [6]. The main resource for one-way quantum computing is a network of entangled qubits, usually called a *cluster state*. Information is processed by measuring each of the qubits in the cluster state individually since a measurement of one of the entangled qubits immediately projects the other qubits into different states. Depending on the order and type of measurements that we perform, different quantum algorithms can be implemented. The output of the computation is left on the qubits that remain after the measurement sequence and reading out this state will give the user the result of the computation. By increasing the size of the initial cluster state we can perform more measurements and thus implement more complex and more powerful computational tasks. Quantum measurements are irreversible since they collapse a superposition into a definite state and that is reason for this model being referred to as *one-way* quantum computing. Although the measurements are irreversible it is still able to simulate the standard model of quantum computation.

In 2005, Philip Walther and his colleagues [7] were able to demonstrate the working principles of one-way quantum computing by employing the polarisation degree of freedom of photons to carry information. A vertically (horizontally) polarised photon represents the logical state "0" ("1") while a photon polarised along $45°$ denotes the superposition "0" + "1". Walther *et al.* used a process called spontaneous parametric down-conversion in order to generate polarisation-entangled photon pairs and with a series of optical elements they were able to combine two entangled photon pairs and create a four-photon cluster state. This cluster state allowed them to realise single- and two-qubit gates as well as to demonstrate Grover's search algorithm with a database of four elements. Interestingly, in this case only a single step is needed to obtain the correct solution!

In this dissertation we propose an experiment utilising the orbital angular momentum of photons which is based on the aforementioned implementation by Walther's group. But what is orbital angular momentum, how can we encode information using this degree of freedom and why would we want to do so? While spin angular momentum can be associated with the polarisation of light, orbital angular momentum arises from light beams possessing an azimuthal phase structure in the plane perpendicular to the direction of propagation, which is independent of its polarisation. This causes the beam to spiral in a corkscrew-like manner around the beam's central axis. For normal laser light, if we connect all the points in the wave that have the same phase we could map out a surface that looks like a flat plane. However, for light carrying orbital angular momentum (also referred to as *twisted light*), the wavefronts form a corkscrew shape. The number of twists per wavelength denotes the orbital angular momentum carried by the beam. Therefore, a single helix will carry one unit of orbital angular momentum, a double helix (which resembles DNA) would carry two units of orbital angular momentum, etc. Although these beams may seem exotic, Les Allen [8] demonstrated, in 1992, that a special set of twisted light beams, called Laguerre-Gaussian beams, could be easily generated in the laboratory. Nowadays, an even simpler method involving digital holography can be used to create beams with absolutely any finite number of twists.

Until Allen's seminal paper, physicists believed that orbital angular momentum was a classical feature of light since it is quite difficult to see what twisted light means on the single-photon level. However, Allen *et al.* proved that Laguerre-Gaussian beams possess a well-defined orbital angular momentum (in units of $\hbar$) per photon — so the orbital angular momentum is quantised and a property of the individual photons, as well. A photon is an elementary excitation of a certain mode of an electromagnetic field. A photon that excites a pure Laguerre-Gaussian mode is in a quantum state with a well-defined orbital angular momentum in units of $\hbar$. It is possible to prepare a superposition of a single excitation of modes with different orbital angular momenta which corresponds to the preparation of a superposition of orbital angular momentum states of a photon.

By creating twisted light we actually twist each of the individual photons which means that each photon in the beam gains the same definite orbital angular momentum. While polarisation is limited to only two states (horizontal and vertical, or right- and left-circular), orbital angular momentum can take on an unlimited number of possible states. This feature is, therefore, very useful since there is no theoretical limit to the amount of information that we can encode using photons which carry orbital angular momentum. This is advantageous because it means that it may be possible to use fewer information carriers for the same calculations than using the two dimensions

of polarisation. It has been proven that higher-dimensional states are more robust against environmental noise, which leads to the decoherence (i.e. the collapse of a quantum superposition due to the interaction with the environment) [9, 10]. Another important experiment that confirms the quantisation of orbital angular momentum was done by Alois Mair together with Anton Zeilinger's group [11], which demonstrates that two photons can be entangled in their orbital angular momentum degree of freedom using spontaneous parametric down-conversion and that the orbital angular momentum is conserved in this process. Furthermore, this process was used in a recent ground-breaking experiment done by A. C. Dada *et al.* [12] to demonstrate the preparation of a genuine 11-dimensional orbital angular momentum entangled two-photon state.

The research described in this dissertation leads to a theoretical basis for a proof-of-principle experiment of one-way quantum computing using the orbital angular momentum of single photons. The hope is that this will facilitate the use of the multi-dimensional system provided by photonic orbital angular momentum for quantum information processing protocols.

1.1 Overview of the Dissertation

This dissertation is comprised of two main parts. The first part (Chapters 2-5) deals with orbital angular momentum which is used to encode information in single photons while the second part (Chapters 6-8) introduces concepts related to quantum computing and one-way quantum computing, in particular. Chapter 9 is the main research chapter which combines knowledge from the first two parts to devise a proposal for an experiment utilising photonic orbital angular momentum for one-way quantum computing.

In Chapter 2 we introduce the concept of orbital angular momentum for classical paraxial light beams. We find solutions to the paraxial wave equation which form a basis for the description of all paraxial light beams. We also discuss the curious trajectory of the Poynting vector (which indicates the flow of energy in an electromagnetic field) for beams carrying orbital angular momentum.

In Chapter 3 we give a recipe for the quantisation of the electromagnetic field and then discuss the orbital angular momentum of light on the single-photon level. We review the use of Dirac formalism for classical light beams and discuss the confusion that may arise when utilising this formalism since one may actually believe that one has transitioned to the quantum regime. The analogy and difference between the quantum and classical states of light is also addressed.

Chapter 4 describes, in sufficient detail, a non-linear process known as spontaneous parametric down-conversion which is used to generate orbital angular momentum entangled photon pairs. These entangled photon pairs are then combined with the optical elements described in Chapter 5 in order to generate a four-photon orbital angular momentum entangled cluster state. Detail is given with regard to the transformation of twisted light beams on both a classical and quantum level.

A selection of quantum computational concepts are described in Chapter 6 including the idea of a

universal set of gates, the standard model of quantum computation and Grover's search algorithm. Chapter 7 provides a detailed description of one-way quantum computing wherein we concentrate on a few unique features of this model and discuss how we can inscribe a circuit from the standard model onto a cluster state by means of a particular series of single-qubit measurements. Walther's experiment is examined in detail in Chapter 8 by carefully working through each of the steps needed to prepare the cluster state and carry out single-photon polarisation measurements. Chapter 9 describes how we can transcribe this experiment to the use of orbital angular momentum qubits. We employ spontaneous parametric down-conversion to produce two entangled photon pairs and explain how these pairs can be combined using certain optical components to produce a four-photon cluster state. We then explain how single-qubit measurements can be performed in order to implement a two-bit version of Grover's algorithm.

1.2 How to read this dissertation

Because of the somewhat diverse nature of this dissertation there are many different selections of chapters which a particular reader may be interested in. Here we present a guide for those who wish to read this dissertation more selectively.

For readers who are new to the field of orbital angular momentum on both a quantum and classical level, Chapters 2, 3 and 5 give a good introduction.

Those interested in digital holography can find a fairly in-depth analysis in Section 4.1. A demonstration of this technology for projective measurements onto different orbital angular momentum states is provided in Chapter 9.

For readers interested in spontaneous parametric down-conversion we would suggest Chapter 4 and Sections 8.2 and 9.2 for an application of this knowledge in an experiment.

Readers from a classical physics background can read Chapters 3 and 6 to obtain a basic idea about how single photons can be used to encode information and the basic features of quantum computation.

Lastly, those who wish to explore the one-way quantum computation model would be advised to read Chapter 7 for the working principles and Chapters 8 and 9 for an implementation of a selection of these concepts.

Chapter 2

Orbital Angular Momentum of Classical Light

The laws which govern the behaviour of electromagnetic radiation, and more specifically light, were first compiled into a self-consistent theory by James Maxwell [13] in 1864 and since then the four fundamental equations of electromagnetism have been known as Maxwell's equations. It is well known from Maxwell's theory that light carries both energy and an associated momentum, which consists of linear and angular contributions. In 1909, Poynting proved that circularly polarised light carries spin angular momentum, which is related to the polarisation of the light beam [14]. This was experimentally verified by Beth, 20 years later, using a quarter-wave plate suspended from a fine quartz fibre [15]. The discovery of light carrying orbital angular momentum, however, was only made much later by Allen *et al.* [8] who were able to prove that a light beam with a specific azimuthal phase dependence possesses a well-defined orbital angular momentum.

In this chapter we will give a short introduction to the concept of orbital angular momentum for paraxial light beams from a classical perspective. In Sections 2.1 and 2.2 we present Maxwell's equations and a derivation of the paraxial wave equation. This will assist us in a short discussion about the orbital angular momentum content of paraxial light beams with an azimuthal phase structure of $\exp(-il\phi)$, which is given in Section 2.3. In Section 2.4 we look at two complete and orthogonal sets of solutions to the paraxial wave equation: namely the Hermite-Gaussian modes for rectangular geometry, and Laguerre-Gaussian modes for cylindrical geometry. These modes are particularly important for our purposes because Hermite-Gaussian modes are the usual outputs from lasers since most laser cavities contain elements with rectangular symmetry, and Laguerre-Gaussian modes contain an azimuthal phase factor which allows them to possess orbital angular momentum. Finally, in Section 2.5 the trajectory of the Poynting vector of Laguerre-Gaussian modes, which shows the flow of energy in the beam, is examined.

2.1 Maxwell's Equations

The four Maxwell's equations which describe the behaviour of light in free space are given by:

$$\nabla \cdot \mathbf{E} \;=\; 0 \tag{2.1}$$

$$\nabla \cdot \mathbf{B} \;=\; 0 \tag{2.2}$$

$$\nabla \times \mathbf{E} \;=\; -\frac{\partial \mathbf{B}}{\partial t} \tag{2.3}$$

$$\nabla \times \mathbf{B} \;=\; \varepsilon_0 \mu_0 \frac{\partial \mathbf{E}}{\partial t} \tag{2.4}$$

where ε_0 and μ_0 are the permittivity and permeability of free space, respectively. We can obtain the wave equation for the electric field by taking the curl of Eq. (2.3) and substituting Eq. (2.4) in the resulting equation

$$\nabla^2 \mathbf{E} - \frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} = 0, \tag{2.5}$$

where $c = 1/\sqrt{\mu_0 \varepsilon_0}$ is the speed of the light in vacuum.

2.2 The Paraxial Wave Equation

The coherent radiation emitted from lasers is usually a collimated monochromatic light beam whose transverse extent, or beam width, is large compared to its wavelength. Although they share many properties with plane waves, the intensity of these beams are concentrated near the axis rather than uniform, and their wavefronts (which are the surfaces consisting of a set of points with equal phase at a fixed instant in time) are slightly spherical. The paraxial approximation is valid for this type of light because the beam divergence is small, therefore we can use it to simplify the wave equation as we will now demonstrate.

The wave equation holds for each component of the electric field vector, i.e. each component V of this vector must satisfy the scalar wave equation

$$\nabla^2 V - \frac{1}{c^2} \frac{\partial^2 V}{\partial t^2} = 0. \tag{2.6}$$

We can use separation of variables by assuming that $V(x, y, z, t) = V(x, y, z) \exp(i\omega t)$ is a solution to Eq. (2.6) to arrive at the time-independent scalar wave equation, known as the Helmholtz equation

$$\left(\nabla^2 + k^2\right) V(x, y, z) = 0, \tag{2.7}$$

where $k = \omega/c$ is the wave number. If we take the z-axis as the direction of propagation, we can write

$$V = u(x, y, z) \exp(-ikz), \tag{2.8}$$

where $u(x, y, z)$ is the complex scalar function (also referred to as the *mode function*) describing the distribution of the field amplitude. This mode function represents the spatial modulation of a

plane wave travelling in the z-direction. Substituting into Eq. (2.7) yields

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) u(x, y, z) - 2ik \frac{\partial u}{\partial z} = 0. \tag{2.9}$$

We can now invoke the paraxial approximation since laser light is collimated and the mode function $u(x, y, z)$ varies slowly in the longitudinal direction z compared to the transverse directions x and y. This is stated mathematically as

$$\left| \frac{\partial^2 u}{\partial z^2} \right| \ll \left| \frac{\partial^2 u}{\partial x^2} \right|, \left| \frac{\partial^2 u}{\partial y^2} \right|, \left| 2k \frac{\partial u}{\partial z} \right|. \tag{2.10}$$

Eq. (2.9) thus simplifies to the paraxial wave equation

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) u(x, y, z) - 2ik \frac{\partial u}{\partial z} = 0. \tag{2.11}$$

2.3 Orbital angular momentum of a paraxial light beam

Much of the research in this dissertation involves utilising the orbital angular momentum (OAM) of light so it would serve us well to know which type of light beams may carry OAM. A plane wave has a linear momentum, $\mathbf{p} = \mathbf{E} \times \mathbf{B}$, in the direction of propagation so this type of wave cannot have an angular momentum, $\mathbf{L} = \mathbf{r} \times (\mathbf{E} \times \mathbf{B})$, in the same direction [8]. However, if our electric and magnetic fields had small components in the direction of propagation then we would obtain an angular momentum in that direction (i.e. $\mathbf{L}$ would have a z component). Moreover, it has been shown in [8] that any light beam with an amplitude distribution of the form $u(r, \phi, z) = u_0(r, z) \exp(-il\phi)$, which satisfies the paraxial wave equation, will possess a well-defined OAM. We will now try to sketch the proof behind this deduction using [16] as a guide.

For this discussion, and the work done in this dissertation, we will consider only linearly polarised light since we are interested in the orbital component of the angular momentum and we do not want to have complex expressions involving the spin angular momentum. We will use the Lorentz gauge for convenience and take the z direction as the direction of propagation of the light.

Let us begin with a complex-valued vector potential $\mathbf{A}$ that is linearly polarised along the x direction

$$\mathbf{A} = u(x, y, z) \exp[i(\omega t - kz)]\hat{\mathbf{x}}. \tag{2.12}$$

Using this vector potential we can obtain expressions for the electric and magnetic fields in terms of the mode function u in the paraxial approximation. In this approximation the second derivatives of the $\mathbf{B}$ and $\mathbf{E}$ fields and the products of first derivatives are ignored and $\partial u / \partial z$ is taken to be

small compared with ku [8]. The $\mathbf{B}$ and $\mathbf{E}$ fields are thus given by

$$\mathbf{B} \;=\; \nabla \times \mathbf{A} \;=\; -ik \exp[i(\omega t - kz)] \left(u\hat{\mathbf{y}} - \frac{i}{k}\frac{\partial u}{\partial y}\hat{\mathbf{z}} \right), \tag{2.13}$$

$$\mathbf{E} \;=\; -\frac{\partial \mathbf{A}}{\partial t} - \nabla \Psi \;=\; -i\omega t\mathbf{A} - \frac{ic}{k}\nabla(\nabla \cdot \mathbf{A})$$

$$\;=\; -i\omega \exp[i(\omega t - kz)] \left(u\hat{\mathbf{x}} - \frac{i}{k}\frac{\partial u}{\partial x}\hat{\mathbf{z}} \right), \tag{2.14}$$

respectively, where Ψ is a scalar potential and we have used the gauge condition $\Psi = (ic/k)\,\nabla \cdot \mathbf{A}$ in Eq. (2.14). Now both the $\mathbf{B}$ and $\mathbf{E}$ fields have z components which we have specified as a prerequisite for the beam to carry angular momentum in the direction of propagation. These fields can be used to evaluate the time-averaged linear momentum density

$$\mathbf{p} = \varepsilon_0 \langle \mathbf{E}_{\mathrm{real}} \times \mathbf{B}_{\mathrm{real}} \rangle. \tag{2.15}$$

We can extract the real parts of the $\mathbf{B}$ and $\mathbf{E}$ fields from Eqs. (2.13) and (2.14), respectively, as follows:

$$\mathbf{B}_{\mathrm{real}} = \tfrac{1}{2}(\mathbf{B} + \mathbf{B}^*) \tag{2.16}$$

$$\mathbf{E}_{\mathrm{real}} = \tfrac{1}{2}(\mathbf{E} + \mathbf{E}^*). \tag{2.17}$$

Substituting Eqs. (2.13), (2.14), (2.16) and (2.17) in Eq. (2.15) and after removing all oscillating terms since they average to zero over time, we obtain

$$\begin{aligned}
\mathbf{p} &= \varepsilon_0 \langle \mathbf{E}_{\mathrm{real}} \times \mathbf{B}_{\mathrm{real}} \rangle \\
&= \frac{\varepsilon_0}{4}\Big[(\mathbf{E} \times \mathbf{B}^*) + (\mathbf{E}^* \times \mathbf{B}) \Big] \\
&= i\omega\frac{\varepsilon_0}{4}\left[u^*\left(\frac{\partial u}{\partial x}\hat{\mathbf{x}} + \frac{\partial u}{\partial y}\hat{\mathbf{y}} \right) - u\left(\frac{\partial u^*}{\partial x}\hat{\mathbf{x}} + \frac{\partial u^*}{\partial y}\hat{\mathbf{y}} \right) + 2k|u|^2\hat{\mathbf{z}} \right] \\
&= i\omega\frac{\varepsilon_0}{4}\Big[u^*\nabla u - u\nabla u^* + 2k|u|^2\hat{\mathbf{z}} \Big], \tag{2.18}
\end{aligned}$$

where we have included the term of order zero (i.e. $\partial u/\partial z$) so that the grad terms could be written concisely. This form of the linear momentum density is useful because it allows for a simple conversion between Cartesian and cylindrical coordinates.

We are considering amplitude distributions of the form $u(r, \phi, z) = u_0(r, z)\exp(-il\phi)$. The azimuthal component of the linear momentum density when applied to such an amplitude distribution is

$$\begin{aligned}
p_\phi = \varepsilon_0 \langle \mathbf{E} \times \mathbf{B} \rangle_\phi &= \frac{i\omega\varepsilon_0}{4}\left[\frac{1}{r}u^*(r, \phi, z)\frac{\partial u(r, \phi, z)}{\partial \phi} - \frac{1}{r}u(r, \phi, z)\frac{\partial u^*(r, \phi, z)}{\partial \phi} \right] \\
&= \frac{\omega\varepsilon_0 l}{2r}|u(r, \phi, z)|^2, \tag{2.19}
\end{aligned}$$

where we have used the gradient operator in cylindrical coordinates (i.e. $\nabla = \frac{\partial}{\partial r}\hat{\mathbf{r}} + \frac{1}{r}\frac{\partial}{\partial \phi}\hat{\phi} + \frac{\partial}{\partial z}\hat{\mathbf{z}}$) and selected only the ϕ component. The magnitude of the angular momentum density along the beam axis j_z is then found to be

$$j_z = (\mathbf{r} \times \mathbf{p})_z = (\mathbf{r} \times \langle \mathbf{E} \times \mathbf{B} \rangle)_z = rp_\phi = \frac{\varepsilon_0 \omega l}{2} \left| u(r, \phi, z) \right|^2,\qquad(2.20)$$

and the energy density of such a beam is

$$\mathbf{w} = p_z c = c\varepsilon_0 \langle \mathbf{E} \times \mathbf{B} \rangle_z = \frac{\varepsilon_0 \omega^2}{2} \left| u(r, \phi, z) \right|^2.\qquad(2.21)$$

If we integrate the angular momentum and energy densities over the xy plane, the ratio of angular momentum to energy per unit length of the beam is obtained

$$\frac{J_z}{\mathbf{W}} = \frac{\int\int(\mathbf{r} \times \langle \mathbf{E} \times \mathbf{B} \rangle)_z r dr d\phi}{c\int\int\langle \mathbf{E} \times \mathbf{B} \rangle_z r dr d\phi} = \frac{l}{\omega},\qquad(2.22)$$

while the ratio of the angular momentum to linear momentum is

$$\frac{J_z}{P} = \frac{\int\int(\mathbf{r} \times \langle \mathbf{E} \times \mathbf{B} \rangle)_z r dr d\phi}{\int\int\langle \mathbf{E} \times \mathbf{B} \rangle_\phi r dr d\phi} = \frac{l}{k} = l\frac{\lambda}{2\pi}.\qquad(2.23)$$

Since we have specified that the beam we are considering is linearly polarised, the only type of angular momentum that it can possess is OAM. This type of angular momentum arises as a result of the azimuthal phase dependence $\exp(-il\phi)$ of the amplitude distribution. From Eqs. (2.22) and (2.23) we can see that the OAM of the beam is well-defined and proportional to l.

2.4 Beam-like solutions to the paraxial wave equation

The simplest non-trivial solution to the Helmholtz equation (Eq. (2.7)) is the well-known spherical wave

$$u(x, y, z) = \frac{\exp(-ik\rho)}{\rho},\qquad(2.24)$$

where $\rho = \sqrt{x^2 + y^2 + z^2}$. If we consider the situation where we have small angles between the distance vector $\boldsymbol{\rho}$ and the z-axis (which is the case for paraxial light beams) then we can invoke the Fresnel approximation by expanding the distance ρ as a power series and neglecting all terms higher than quadratic

$$\rho = z + \frac{x^2 + y^2}{2z}.\qquad(2.25)$$

This relation will be used when rewriting the phase shift factor $\exp(-ik\rho)$ in Eq. (2.24). In the denominator of Eq. (2.24) however, we shall drop even the quadratic term and replace ρ simply by z [17]. The spherical wave can now be written as

$$u(x, y, z) \approx \frac{1}{z} \exp\left[-ikz - \frac{ik\left(x^2 + y^2\right)}{2z} \right],\qquad(2.26)$$

which is the paraxial approximation of the spherical wave. One can verify that this wave is now an analytical solution to the paraxial wave equation rather than the Helmholtz equation. Our aim is to model laser light therefore we use this solution to help us to find beam-like solutions to the paraxial wave equation in the sections that follow, using [18] as a guide. A complete derivation of each of the solutions that we will now look at can be found in Appendix A but we shall highlight the key points that are required for each case.

2.4.1 The Gaussian Beam and its Properties

A coherent light beam with a Gaussian intensity profile is one of the most important solutions to the paraxial wave equation. It is often referred to as the *fundamental mode* since it is the lowest order mode emitted by a laser. In order to form the Gaussian beam we use a trial solution which has the same form as the paraxial approximation to the spherical wave

$$u_{\mathrm{G}}(r,z) = C \exp\left(-if(z) - \frac{ikr^2}{2q(z)}\right),$$ (2.27)

where $r^2 = x^2 + y^2$, C is a constant and $f(z)$ and $q(z)$ are analytical functions of z. The parameter $f(z)$ represents a complex phase shift which is associated with the propagation of the light beam, and $q(z)$ is a complex parameter which describes the Gaussian variation in beam intensity with distance r from the propagation axis, as well as the curvature of the wavefront which is spherical near the axis. If we substitute this ansatz into the paraxial wave equation and simplify we can obtain equations for $f(z)$ and $q(z)$ and after further calculations we arrive at the commonly used expression for the Gaussian beam

$$u_{\mathrm{G}}(r,z) = C\frac{w_0}{w(z)}\exp\left(-\frac{r^2}{w^2(z)}\right)\exp\left(-ik\frac{r^2}{2R(z)} - i\Phi^{\mathrm{G}}(z)\right),$$ (2.28)

with

$$w(z) = w_0\sqrt{1 + \left(\frac{z}{z_{\mathrm{R}}}\right)^2}$$ (2.29)

$$R(z) = z\left[1 + \left(\frac{z_{\mathrm{R}}}{z}\right)^2\right]$$ (2.30)

$$\Phi^{\mathrm{G}}(z) = \arctan\left(\frac{z}{z_{\mathrm{R}}}\right)$$ (2.31)

$$z_{\mathrm{R}} = \frac{\pi w_0^2}{\lambda},$$ (2.32)

where the factors on the left hand side of Eq. (2.28) will now be explained.

The beam radius or spot size $w(z)$ is a measure of how the amplitude of the electric field decreases when moving away from the axis of propagation in the transverse direction. This falloff is Gaussian, as seen in Fig. 2.1, and $w(z)$ is the radial distance from the axis of propagation where the amplitude is $1/e$ of its value on that axis. At $z = 0$ the beam radius will be at its minimum w_0,

and this is called the beam waist. The wavefront at the point $z = 0$ is a plane. The radius of curvature $R(z)$ describes the curvature of the wavefront intersecting the beam axis at z. It represents the spherical distortion of the wavefront compared to a plane wave. The Rayleigh range z_R gives us information about the spread of the beam. If we choose $z = z_R$, then $w(z) = \sqrt{2}w_0$ which means that at this z, the cross section of the beam is twice as large as the cross section at $z = 0$. A short z_R means that the beam diverges rapidly. The parameter $\Phi^G(z)$ is known as the Gouy phase and it represents the phase shift between the Gaussian beam and an ideal plane wave. This phase shift results in a slightly larger distance between wavefronts of a Gaussian beam compared to an ideal plane wave which means that the wavefronts have to propagate somewhat faster, leading to an increased local phase velocity [17]. Finally, the parameter $w_0/w(z)$ gives the expected decrease in the amplitude on the beam axis due to the expansion of the beam.

The time-averaged intensity of a Gaussian beam is given by

$$I_G(r, z) = |u(r, z)|^2 = I_0 \left(\frac{w_0}{w(z)} \right)^2 \exp \left(-\frac{2r^2}{w(z)^2} \right), \tag{2.33}$$

where $I_0 = |C|^2$. The intensity of this beam has a Gaussian distribution, hence the name. The beam has a circular cross section (Fig. 2.1) given by a Gaussian function, but the width of the distribution changes as the beam propagates. As mentioned before, the wavefronts (Fig. 2.2) are slightly curved compared to plane waves whose wavefronts are perpendicular to the z-axis.
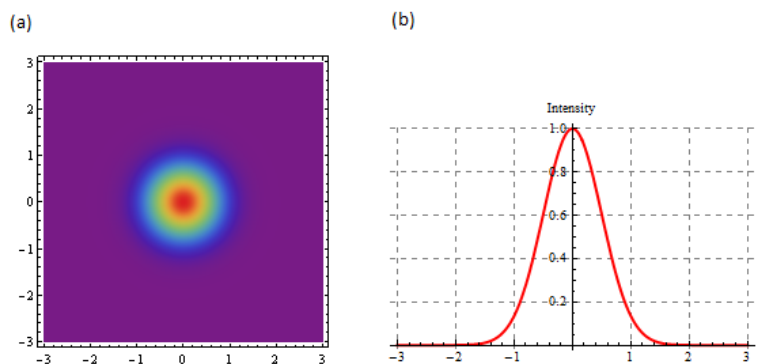
Figure 2.1: Gaussian Beam: (a) Cross-sectional beam structure in the transverse plane. (b) One-dimensional cross section of the beam intensity.

2.4.2 Higher Order Modes

The Gaussian beam that we have discussed in the previous section is only the lowest-order solution of an infinite family of higher-order solutions to the paraxial wave equation. If we chose a complete and orthogonal basis set, which are called the *modes of propagation*, then any real laser beam can be described as a superposition of solutions from this set. When light propagates in free

space, its intensity profile will generally change during propagation. For these modes of propagation, however, the shape of the amplitude profile remains constant, even though there may be a rescaling of the profile, an overall change in optical phase, and possibly also a change in the total optical power [19]. The design of the laser determines which basis set of solutions is most useful. Since most laser cavities have components with rectangular symmetry, Hermite-Gaussian modes are commonly used for the description of the laser beams that are emitted. Although Laguerre-Gaussian modes have been produced directly in laser systems with cylindrical symmetry, they are more easily produced by the external conversion of Hermite-Gaussian modes. This is because if there is even a slight departure form cylindrical symmetry, the emitted beam would be better described as a superposition of Hermite-Gaussian modes [20].
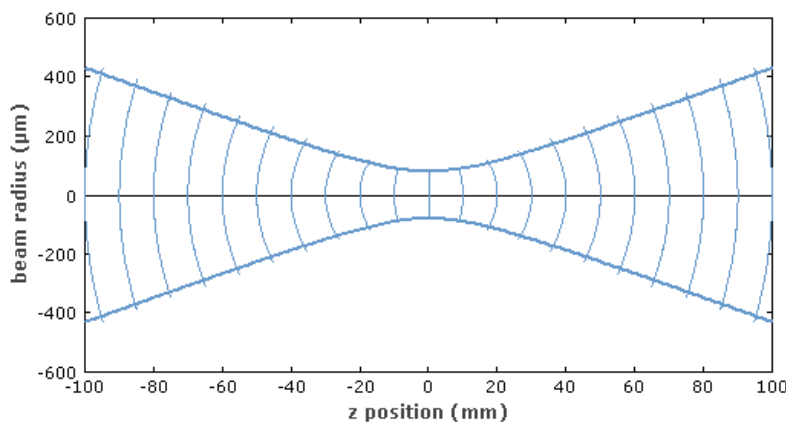
Figure 2.2: The cross section of the beam profile in the xz plane illustrates the spherical wavefronts of a Gaussian beam [19].

We will now show how the same trial solution approach of the previous section can be extended to find higher-order Hermite-Gaussian (HG) or Laguerre-Gaussian (LG) eigensolutions to the paraxial wave equation.

Hermite-Gaussian modes

Hermite-Gaussian (HG) modes can be used to describe the output of lasers whose cavity design is rectangular. In order to arrive at a complete expression for the HG modes we can use the ansatz

$$u(x,y,z) = g\left(\frac{x}{w(z)}\right) h\left(\frac{y}{w(z)}\right) \exp[-iP(z)]u_{\mathrm{G}}(r,z), \qquad (2.34)$$

where g is a function x and z, h is a function of y and z and $w(z)$ is the distance-dependent scaling factor of the beam (given by Eq. (2.29)). We can substitute this trial solution into the paraxial wave equation and on simplifying we find that the functions g and h can be written as Hermite

polynomials

$$g\left(\frac{x}{w(z)}\right) = H_n\left(\frac{\sqrt{2}x}{w(z)}\right), \tag{2.35}$$

$$h\left(\frac{y}{w(z)}\right) = H_m\left(\frac{\sqrt{2}y}{w(z)}\right), \tag{2.36}$$

since the differential equations for these functions are of the same form as the differential equation for a Hermite polynomial $H_m(x)$ of order n

$$\frac{d^2 H_n}{dx^2} - 2x\frac{dH_n}{dx} + 2nH_n. \tag{2.37}$$

The complex phase shift $P(z)$ is given by

$$P(z) = -(n+m)\arctan\left(\frac{z}{z_R}\right), \tag{2.38}$$

therefore the Gouy phase of an HG mode $\Phi^{HG}(z)$ is

$$\Phi^{HG}(m,n;z) = (n+m+1)\arctan\left(\frac{z}{z_R}\right) = (n+m+1)\Phi^{G}(z), \tag{2.39}$$

and the final expression for the amplitude is

$$u_{nm}^{HG}(x,y,z) = C_{nm}^{HG}\frac{w_0}{w(z)}H_n\left(\frac{\sqrt{2}x}{w(z)}\right)H_m\left(\frac{\sqrt{2}y}{w(z)}\right)$$
$$\times \exp\left(-\frac{r^2}{w^2(z)}\right)\exp\left(-ik\frac{r^2}{2R(z)} + i\Phi^{HG}(z)\right). \tag{2.40}$$

The beam parameters in Eq. (2.40) are defined as for the Gaussian beam. Apart from a phase term, an HG beam can be described as the product of Hermite polynomials and a Gaussian function, hence the name. For $n = m = 0$ we obtain the fundamental, or Gaussian, mode. Thereafter, the order of the mode for each of these higher-order modes is given by the sum $n+m$ where n and m determine the shape of the intensity profile in the x and y directions, respectively. The size of the profile of any HG mode depends on z while its shape remains constant as the beam propagates along the z-axis. The intensity distribution (Fig. 2.3) has n nodes in the horizontal direction and m nodes in the vertical direction. We can see that the size of the intensity distribution increases with the mode index while the peak intensity decreases.

The radius of curvature of the beam $R(z)$ depends on z but is the same for all HG modes implying that the wavefront curvature changes in the same way for all modes. The wavefronts are a series of discs separated by the wavelength of the light [20]. The Gouy phase shift for HG modes is a function of the mode numbers which means that the phase velocity and spacing between the wavefronts increases with mode number. This phase shift also results in modes of different orders having different resonant frequencies [17].
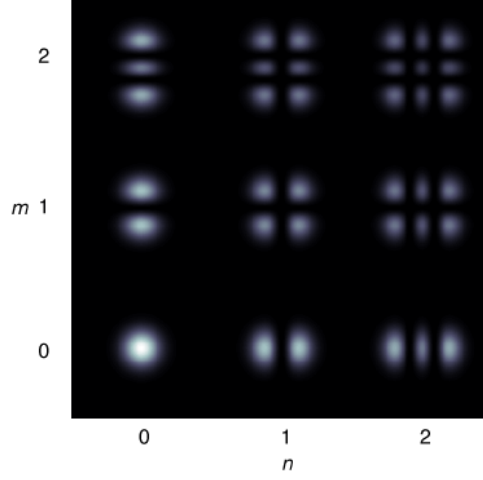
Figure 2.3: The intensity distribution of the first few HG modes. We can see the width of the intensity distribution increases with mode index, while the peak intensity decreases [21].

Laguerre-Gaussian modes

We now look for beam-like solutions to the paraxial wave equation which may possess OAM. These solutions must have an azimuthal phase dependence of $\exp(-il\phi)$ since we know, from Section 2.3, that such a beam will carry OAM. We shall once again use u_G as a starting point and expand it to form our ansatz using cylindrical coordinates

$$u(r, \phi, z) = F(r, z) \exp(-il\phi) u_G(r, z). \tag{2.41}$$

We can insert this ansatz into the paraxial equation and after some manipulation we find that

$$F(r, w) = \left(\frac{\sqrt{2}r}{w(z)} \right)^l L_p^l \left(\frac{2r^2}{w(z)^2} \right) \exp \left[i(2p + l) \arctan \left(\frac{z}{z_R} \right) \right], \tag{2.42}$$

where $p = \min(n, m)$ is the radial mode index, $l = m - n$ is the azimuthal mode index and L_p^l is are the associated Laguerre polynomials

$$L_p^l(x) = \frac{1}{p!} \sum_{j=1}^{p} \frac{p!}{j!} \binom{l+p}{p-j} (-x^j). \tag{2.43}$$

The Gouy phase of an LG is thus given by

$$\Phi^{LG}(p, l; z) = (2p + |l| + 1) \arctan \left(\frac{z}{z_R} \right) = (2p + |l| + 1) \Phi^G(z). \tag{2.44}$$

The final expression for the amplitude distribution of an LG mode is

$$u_{pl}^{\mathrm{LG}}(r, \phi, z) = C_{pl}^{\mathrm{LG}} \frac{w_0}{w(z)} \left(\frac{\sqrt{2}r}{w(z)} \right)^{|l|} L_p^{|l|} \left(\frac{2r^2}{w(z)^2} \right) \exp \left(-\frac{r^2}{w^2(z)} \right)$$

$$\times \exp \left(-il\phi - \frac{ikr^2}{2R(z)} + i\Phi^{\mathrm{LG}}(z) \right). \tag{2.45}$$

The parameters $w(z)$ and $R(z)$ are defined as for the Gaussian beam. The LG mode is a product of a Laguerre polynomial and a Gaussian function, hence the name. The order of an LG mode is given by $2p + |l|$. If we set $p = |l| = 0$ we obtain the same mode as the HG mode of order zero which is, of course, the Gaussian mode. This equation had to be adapted because both positive and negative values of l are important to us. Therefore, we observe that in some instances the absolute value of l is used instead of l: firstly in the Laguerre polynomials because they are only defined for $l \geq 0$; secondly in the amplitude factor $r^{|l|}$ because r^{-l} would result in an infinite amplitude as $r \to 0$, which is unphysical; and lastly in the Gouy phase because LG modes of the same order $(2p + |l|)$ must have the same Gouy phase [22]. One can see that the only difference between an LG mode u_{pl}^{LG} and $u_{p(-l)}^{\mathrm{LG}}$ is in the azimuthal phase factors $\exp(-il\phi)$ and $\exp(il\phi)$, respectively

The intensity profile of LG beams (Fig. 2.4) depend on the values of p and l. For $l \neq 0$ the beam intensity will be an annular ring whose radius increases with $|l|$ and the number of annular rings (or radial nodes) is given by $p + 1$.
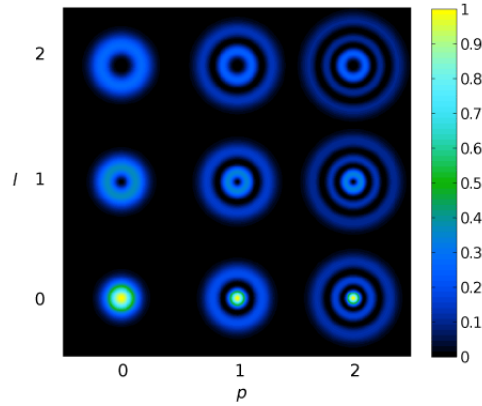
Figure 2.4: The intensity distribution of the first few LG modes. We can see the width of the intensity distribution increases with the azimuthal mode index, while the peak intensity decreases. The number of concentric rings is given by $p + 1$ [21].

All LG beams, with $l \neq 0$ have helical wavefronts (Fig. 2.5), as a result of the azimuthal phase factor $\exp(-il\phi)$ where l determines the number of spirals of the wavefronts which are completed

in a single wavelength. The handedness of the helical wavefronts is linked to the sign of the az-imuthal index l and can be chosen by convention. The helical form of the wavefronts results in an on-axis phase discontinuity and thus a zero on-axis intensity. This result can also be seen from the azimuthal phase structure of the beam which can be represented by the surface plot of a spiral (Fig. 2.6). The phase varies l times from 0 to 2π around the phase singularity as we encircle the beam. At the phase singularity the phase is undefined giving rise to zero intensity at this point.
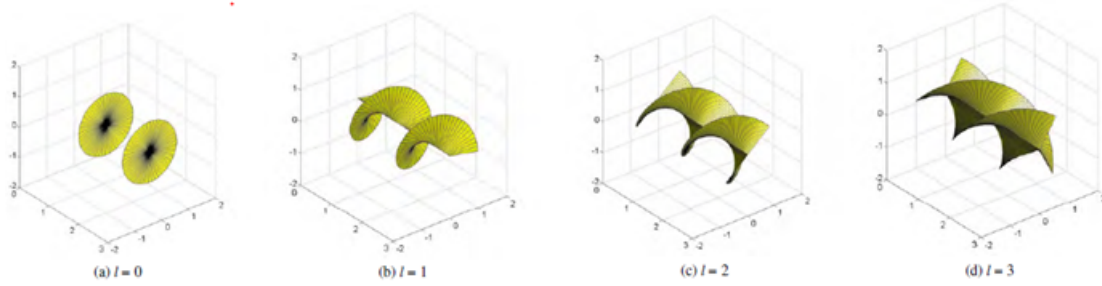
Figure 2.5: The helical wavefronts of the first few LG modes. The number of spirals of the wavefront depends on the value of l [22].
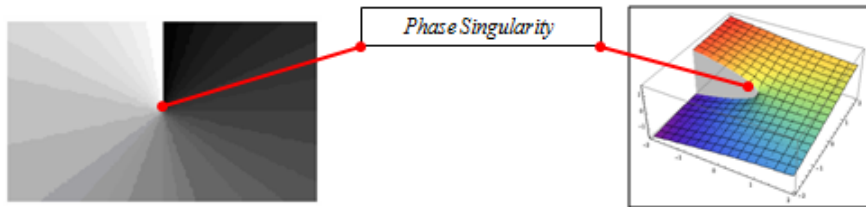
Figure 2.6: The phase can be represented by the following surface plot of a spiral. Increasing the number of times the phase varies from 0 (black) to 2π (white) around the phase singularity increases the order of the beam or OAM carried by the beam. At the phase singularity the phase is undefined giving rise to zero intensity at this point [23].

Relationship between HG and LG modes

In some cases it might be necessary to change our basis from HG modes, the usual modes of the laser, to LG modes, the modes that contain OAM. The decomposition of an LG mode in terms of HG modes, which connects a combination of Hermite polynomials to a Laguerre polynomial, was first given in [24]. The LG modes can be expressed as a sum of HG modes of lower or equal order

$$u_{pl}^{\text{LG}}(x,y,z) = \sum_{k=0}^{N} i^k b(n,m,k) u_{N-k,k}^{\text{HG}}(x,y,z),\qquad(2.46)$$

with real coefficients

$$b(n,m,k) = \sqrt{\frac{(N-k)!k!}{2^N n!m!}} \frac{1}{k!} \frac{d^k}{dt^k}[(1-t)^n(1+t)^m]|_{t=0}, \tag{2.47}$$

where $N = n + m$. Eq. (2.46) is very similar to the decomposition of an HG mode that is rotated at $45°$ about the principal axis, which is given by

$$u_{nm}^{HG@45}(x,y,z) = \sum_{k=0}^{N} b(n,m,k) u_{\mathrm{HG}}^{N-k,k}(x,y,z). \tag{2.48}$$

The only difference between the two decompositions is the i^k factor which gives a relative phase shift of $\pi/2$ between successive components in the u_{pl}^{LG} decomposition while all the components in the $u_{nm}^{\mathrm{HG@45}}$ decomposition are in phase. This shows that we can convert from an HG mode to an LG mode provided that the correct $\pi/2$ phase shift is obtained. Fig. 2.7 shows the decomposition of two first-order LG mode and two first-order HG mode rotated by $45°$. The conversion presented here is essential for a particular method of transforming HG modes into LG modes which will be discussed in Chapter 5.
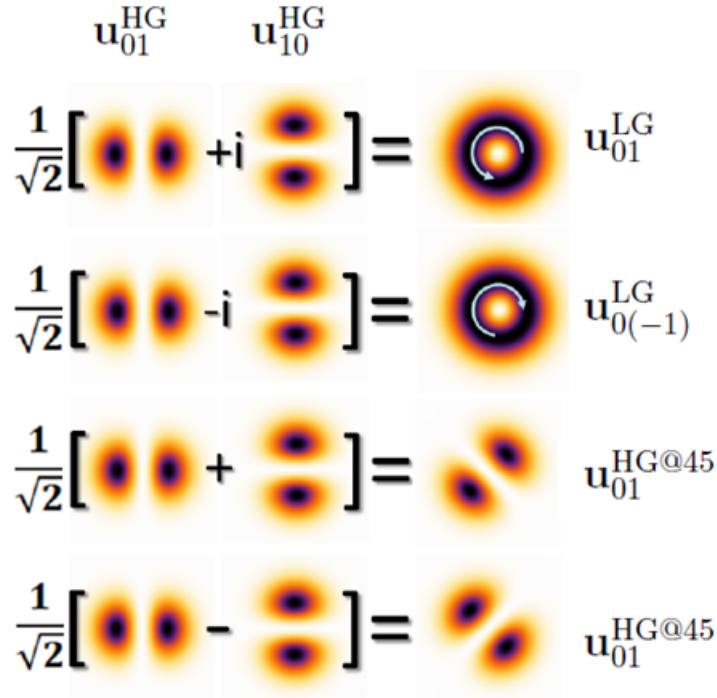
Figure 2.7: Decompositions of the u_{01}^{LG}, $u_{0(-1)}^{\mathrm{LG}}$, $u_{01}^{\mathrm{HG@45}}$ and $u_{01}^{\mathrm{HG@-45}}$ modes in the HG basis [25].

2.5 The Poynting vector and OAM of an LG mode

It is known that energy is propagated by waves, in general, and electromagnetic waves, in particular. We need to quantify the power flow associated with this energy and this is done by means of the Poynting vector. The Poynting vector of light $\mathbf{S}$ shows us the direction in which the energy travels in an electromagnetic wave and has the dimensions of energy per unit time per unit area. In free space the linear momentum density is $(\mathbf{E} \times \mathbf{B})/c^2$ and can be given in terms of the Poynting vector as $\mathbf{S}/c^2$. We are concerned only with the direction of the Poynting vector for LG modes which we will now derive by following the method used in [16].

Let us, once again, consider a linearly polarised beam with an LG amplitude distribution. Since the wavefronts of LG beams are helical, rather than perpendicular to the direction of propagation, the Poynting vector will not be parallel to the direction of propagation. We know from above that, apart from a constant factor, the Poynting vector is equal to the linear momentum density and since we are concerned only with its direction, we may use the linear momentum density to determine the trajectory of the Poynting vector.

If we evaluate the linear momentum density (Eq. (2.18)) for the LG mode function (Eq. (2.45)), the real part of $u^*\nabla u - u\nabla u^*$ does not contribute and if we now ignore the $\partial/\partial z$ terms (since they of order zero) we find that

$$\mathbf{p} = \frac{\varepsilon_0}{2} \left[\frac{\omega k r}{R(z)}\hat{\mathbf{r}} + \frac{\omega l}{r}\hat{\boldsymbol{\phi}} + \omega k \hat{\mathbf{z}} \right] \left| u_{pl}^{\mathrm{LG}} \right|^2 . \tag{2.49}$$

Here the radial component $\hat{\mathbf{r}}$ of the linear momentum corresponds to the spread of the beam; the azimuthal component $\hat{\boldsymbol{\phi}}$ is proportional to l and gives rise to OAM in the z direction and the $\hat{\mathbf{z}}$ component is related to the linear momentum in the direction of propagation. The relative size of these components determines the rotation of the Poynting vector [20]. From this equation we can see that the Poynting vector will have an azimuthal component as it propagates away from the beam waist.

The rate of change of the angle of rotation θ of the Poynting vector (Fig. (2.8)) with distance z away from the beam waist (at $z = 0$) is given by

$$\frac{\partial \theta}{\partial z} = \frac{1}{r}\frac{p_\phi}{p_z} = \frac{l}{kr^2(z)}. \tag{2.50}$$

In [8] the authors selected a constant radius r away from the beam axis for Eq. (2.50) and, therefore, depicted the trajectory of the Poynting vector as a *multi-turn* spiral with a pitch (i.e. the distance in z taken to complete a single helix) given by

$$z_p = \frac{2\pi k r^2}{l}. \tag{2.51}$$

This interpretation is, however, misleading since it neglects the radial component of the Poynting vector which is given in Eq. (2.49). To obtain the correct representation for the trajectory of the

Poynting vector we must take into account the divergence of the beam, which means that we cannot use a constant radial value r since such a point will not lie within the relative intensity profile of the beam as the beam propagates. We must select a radius that corresponds to a fixed point in the relative intensity profile which means we require the radius that scales with the beam width as the beam propagates, i.e. $r(z)/w(z)$, to be constant.
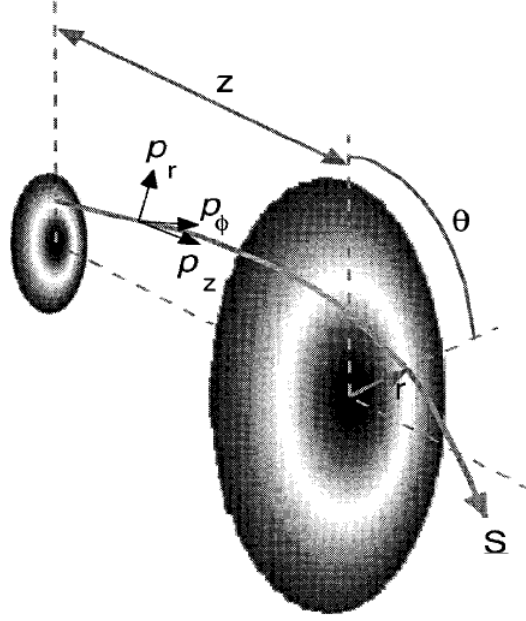
Figure 2.8: The trajectory of the Poynting vector and the components of linear momentum density. Here θ is the angle of rotation of the Poynting vector **S** [26].

The rate of rotation of the angle of the Poynting vector is then

$$\frac{\partial \theta}{\partial z} = \frac{l}{2} \left(\frac{w(z)}{r(z)} \right)^2 \frac{z_{\mathrm{R}}}{z_{\mathrm{R}}^2 + z^2}, \tag{2.52}$$

where we have used Eqs. (2.29) and (2.32). The angle of rotation of the Poynting vector from the beam waist at $z = 0$ is obtained by integrating Eq. (2.52) with respect to z

$$\theta = \frac{l}{2} \left(\frac{w(z)}{r(z)} \right)^2 \arctan \left(\frac{z}{z_{\mathrm{R}}} \right). \tag{2.53}$$

The angle of rotation of the Poynting vector is thus proportional to the change in the Gouy phase from that at the beam waist [16]. For an LG mode with $p = 0$ and $l \neq 0$ mode (which is subset of modes that we will be working with), the intensity distribution is given by a single ring. The radius corresponding to the maximum amplitude in the mode r_{max} is given by

$$r_{\mathrm{max}}(z) = \sqrt{\frac{w(z)l}{2}}. \tag{2.54}$$

The angle of rotation of the Poynting vector associated with a field of this radius is

$$\theta = \arctan\left(\frac{z}{z_{\mathrm{R}}}\right),$$ (2.55)

which is, surprisingly, independent of l.

It is, therefore, clear that the trajectory of the Poynting vector does not follow the helical structure of the wavefront and cannot be described by a *multi-turn* spiral with the number of turns depending on the value of l. Rather, in this case ($p = 0, l \neq 0$), it rotates by exactly $\pi/2$ either side of the beam waist as the light propagates to the far field [16]. The trajectory of the Poynting vector for a simple LG mode is shown in Fig. 2.9. The total rotation is usually of the order of one complete revolution or less for low l values [20] and the pitch z_p of this spiral depends on $r(z)$. At peak intensity, the pitch is such that it becomes a straight line at an angle to the beam axis (Fig. 2.10) [26].
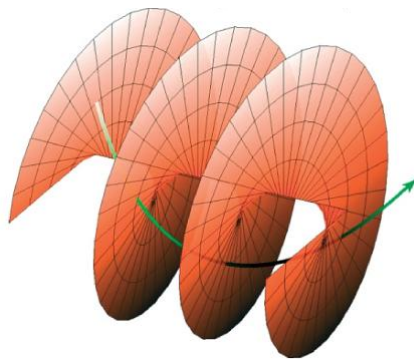
Figure 2.9: The helical wavefronts of an LG mode with $p = 0, l = 1$ and the trajectory of the associated Poynting vector [27].
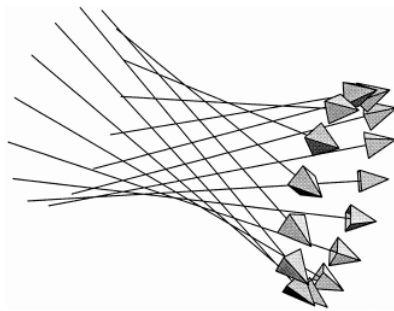
Figure 2.10: The straight line trajectories mapped out by the Poynting vectors at the ring of maximum intensity of a single-ringed LG mode [28].

Chapter 3

Quantum Orbital Angular Momentum

In the year 1900, Max Planck [29] postulated that light could only be absorbed and emitted in discrete bundles of energy. Five years later, Albert Einstein [30] explained the photoelectric effect by suggesting that light consists of discrete energy-packets, or *light quanta* as he called it. Thus began the development of the quantum theory of light as we now know it, which deviates from the 19th century classical theory of Maxwell and Einstein's light quanta are now referred to as photons — a term coined by Gilbert Lewis [31] in 1936.

Photons are one of the main carriers of information that are used in quantum computation. Information can be encoded into the different degrees of freedom of this quantum object including its polarisation, energy and linear momentum. Of interest to us is employing the OAM of photons which allows access to an infinite number of quantum states and is thus optimal for efficient quantum information processing.

This chapter serves as an introduction to the concept of a photon and the representation of its quantum state. We will concentrate specifically on photons that make up an LG laser beam since they are known to possess OAM. In Section 3.1 we show that the OAM of a beam which has an azimuthal phase of the form $\exp(-il\phi)$ is quantised, therefore it carries a discrete OAM of $l\hbar$ per photon along its direction of propagation [8]. Section 3.2 presents the quantisation of classical fields as a means to describe the general states of photons with emphasis on the state of a photon in an LG mode. Dirac notation and operator algebra is applied to LG beams [32], as an alternative to using the states of single photons within the beam, in Section 3.3. Finally, in Section 3.4 we discuss the analogy between the representation of the classical and quantum states of light when using Dirac formalism.

3.1   The OAM of a photon in an LG beam

It is well known from the quantum theory of light that each photon carries a well-defined energy, $E = \hbar\omega$, and a linear momentum of $p = \hbar k$. In 1936, Beth [15] was able to calculate the spin angular momentum of a photon, following Poynting's classical thought experiment. Using both the wave and quantum theories of light, he found that the ratio of spin angular momentum to en-

ergy for a circularly polarised light beam was $\pm 1/\omega$. Quantum mechanically, he had assumed that the beam contained N photons with a total energy of $E = N\hbar\omega$ and, therefore, possessed a spin angular momentum of $\pm\hbar$ per photon.

In Chapter 2, and according to Allen *et al.* [8], the ratio of angular momentum to energy per unit length for a linearly polarised LG beam was found to be l/ω while the ratio of angular to linear momentum density was l/k. Analogously to Beth's method, we assume that the beam is made up of N photons with a total energy of $E = N\hbar\omega$ and a total linear momentum of $p = N\hbar k$, therefore the beam must possess an OAM of $l\hbar$ per photon.

3.2 Quantisation of EM fields and the states of a photon

Maxwell's equations and the classical description of light were given in Chapter 2. We now wish to move to the quantum regime and find a description of light in terms of photons which are seen as elementary excitations of the electromagnetic field. In order to accomplish this we must quantise the electromagnetic field by first finding a representation for the free classical field in the plane-wave basis and then replacing the mode amplitudes by quantum operators, as we will now show.

3.2.1 The free classical field

It is often easier to solve Maxwell's equations for potentials rather than the electric field **E** and the magnetic field **B**. Using Maxwell's equations and some mathematical identities we can write the **B** and **E** fields in terms of a vector potential **A** and a scalar potential Ψ as

$$\mathbf{B} = \nabla \times \mathbf{A}, \tag{3.1}$$

$$\mathbf{E} = -\frac{\partial \mathbf{A}}{\partial t} - \nabla\Psi. \tag{3.2}$$

We are dealing with an electromagnetic field in free space (i.e. in the absence of charge or current sources) so it is convenient to adopt the Coulomb gauge ($\nabla \cdot \mathbf{A} = 0$) and eliminate the scalar potential ($\Psi = 0$). This leads to reduced equations for **B** and **E** which still satisfy Maxwell's equations

$$\mathbf{B} = \nabla \times \mathbf{A}, \tag{3.3}$$

$$\mathbf{E} = -\frac{\partial \mathbf{A}}{\partial t}, \tag{3.4}$$

and the final Maxwell equation (Eq. (2.4)) transforms into a wave equation for **A**

$$\nabla^2\mathbf{A} - \frac{1}{c^2}\frac{\partial^2\mathbf{A}}{\partial t^2} = 0. \tag{3.5}$$

If we treat the electromagnetic waves as if they were in a cubic region of space of finite volume $V = L^3$ and subject to periodic boundary conditions then the solution to Eq. (3.5) is a superposition of moving plane waves and can be written as

$$\mathbf{A}(\mathbf{r}, t) = \frac{1}{\sqrt{V}} \sum_{\mathbf{k},s} \boldsymbol{\epsilon}_{\mathbf{k},s} \left(c_{\mathbf{k},s} e^{i(\mathbf{k}\cdot\mathbf{r} - \omega_{\mathbf{k}}t)} + c_{\mathbf{k},s}^* e^{-i(\mathbf{k}\cdot\mathbf{r} - \omega_{\mathbf{k}}t)} \right), \tag{3.6}$$

where $s = \{1, 2\}$, $k_i = 2n_i\pi/L$ for $i = 1, 2, 3$ and $|\mathbf{k}| = k = \omega/c$. The vectors $\boldsymbol{\epsilon}_{\mathbf{k},s}$, which describe the polarisation of the field, are orthonormal ($\boldsymbol{\epsilon}_{\mathbf{k},s} \cdot \boldsymbol{\epsilon}_{\mathbf{k},s'} = \delta_{ss'}$) and lie in the transverse plane perpendicular to the longitudinal wave vector $\mathbf{k}$ ($\boldsymbol{\epsilon}_{\mathbf{k},s'} \cdot \mathbf{k} = 0$). A mode of an electromagnetic field is specified by a particular wave vector and a corresponding polarisation vector.

We can substitute Eq. (3.6) in Eqs. (3.3) and (3.4) to obtain the representations for the $\mathbf{E}$ and $\mathbf{B}$ fields:

$$\mathbf{E} = \frac{i}{\sqrt{V}} \sum_{\mathbf{k},s} \omega_{\mathbf{k}} \boldsymbol{\epsilon}_{\mathbf{k},s} \left[c_{\mathbf{k},s} e^{i(\mathbf{k}\cdot\mathbf{r}-\omega_{\mathbf{k}}t)} - c_{\mathbf{k},s}^* e^{-i(\mathbf{k}\cdot\mathbf{r}-\omega_{\mathbf{k}}t)} \right], \tag{3.7}$$

$$\mathbf{B} = \frac{i}{c\sqrt{V}} \sum_{\mathbf{k},s} \omega_{\mathbf{k}} \left(\frac{\mathbf{k}}{|\mathbf{k}|} \times \boldsymbol{\epsilon}_{\mathbf{k},s} \right) \left[c_{\mathbf{k},s} e^{i(\mathbf{k}\cdot\mathbf{r}-\omega_{\mathbf{k}}t)} - c_{\mathbf{k},s}^* e^{-i(\mathbf{k}\cdot\mathbf{r}-\omega_{\mathbf{k}}t)} \right]. \tag{3.8}$$

The Hamiltonian, which represents the total energy of the electromagnetic field in the cube, is then

$$\mathcal{H} = \frac{1}{2} \int_V \left(\varepsilon_0 \mathbf{E} \cdot \mathbf{E} + \frac{1}{\mu_0} \mathbf{B} \cdot \mathbf{B} \right) d^3\mathbf{r} = 2\varepsilon_0 \sum_{\mathbf{k},s} \omega_{\mathbf{k}}^2 c_{\mathbf{k},s} c_{\mathbf{k},s}^*, \tag{3.9}$$

when substituting for $\mathbf{E}$ and $\mathbf{B}$ from Eq. (3.7) and Eq. (3.8), respectively. It is noteworthy that the Hamiltonian of the electromagnetic field can be related to that of uncoupled harmonic oscillators. In order to see this we introduce the generalised position and momentum coordinates $q_{\mathbf{k},s}$ and $p_{\mathbf{k},s}$, respectively, which are canonical conjugate coordinates, so that

$$c_{k,s} = \frac{1}{2\omega_{\mathbf{k}}\sqrt{\varepsilon_0}} \left(\omega_{\mathbf{k}} q_{\mathbf{k},s} + i p_{\mathbf{k},s} \right),$$

$$c_{k,s}^* = \frac{1}{2\omega_{\mathbf{k}}\sqrt{\varepsilon_0}} \left(\omega_{\mathbf{k}} q_{\mathbf{k},s} - i p_{\mathbf{k},s} \right), \tag{3.10}$$

resulting in

$$\mathcal{H} = \frac{1}{2} \sum_{\mathbf{k},s} \left(p_{\mathbf{k},s}^2 + \omega_{\mathbf{k}}^2 q_{\mathbf{k},s}^2 \right). \tag{3.11}$$

This resembles the Hamiltonian of many uncoupled one-dimensional harmonic oscillators of unit mass with momenta $p_{\mathbf{k},s}$ and positions $q_{\mathbf{k},s}$.

3.2.2 Quantisation of the free field

We can now quantise the electromagnetic field by transforming the coordinates, $p_{\mathbf{k},s}$ and $q_{\mathbf{k},s}$, which are real numbers, into quantum mechanical operators, $\hat{\mathbf{p}}_{\mathbf{k},s}$ and $\hat{\mathbf{q}}_{\mathbf{k},s}$, which satisfy the commutation relations

$$\left[\hat{\mathbf{q}}_{\mathbf{k},s}, \hat{\mathbf{q}}_{\mathbf{k}',s'} \right] = 0 = \left[\hat{\mathbf{p}}_{\mathbf{k},s}, \hat{\mathbf{p}}_{\mathbf{k}',s'} \right], \tag{3.12}$$

$$\left[\hat{\mathbf{q}}_{\mathbf{k},s}, \hat{\mathbf{p}}_{\mathbf{k}',s'} \right] = i\hbar \delta^{(3)} \left(\mathbf{k} - \mathbf{k}' \right) \delta_{s,s'}. \tag{3.13}$$

The quantum mechanical Hamiltonian is, therefore,

$$\hat{\mathbf{H}} = \frac{1}{2} \sum_{\mathbf{k},s} \left(\hat{\mathbf{p}}_{\mathbf{k},s}^2 + \omega_{\mathbf{k}}^2 \hat{\mathbf{q}}_{\mathbf{k},s}^2 \right). \tag{3.14}$$

The Hamiltonian is the operator associated with the energy of a system and a first step towards calculating the evolution of a system consists of finding the eigenvalues of the Hamiltonian and the associated eigenstates. Hence, it would be desirable to have a simpler form of the Hamiltonian given in Eq. (3.14) which can be achieved by diagonalising the Hamiltonian. To do this we introduce the non-Hermitian (and therefore non-observable) operators $\hat{\mathbf{a}}_{\mathbf{k},s}$ and $\hat{\mathbf{a}}_{\mathbf{k},s}^{\dagger}$ to replace $\hat{\mathbf{q}}_{\mathbf{k},s}$ and $\hat{\mathbf{p}}_{\mathbf{k},s}$

$$\hat{\mathbf{a}}_{\mathbf{k},s} = \frac{1}{\sqrt{2\hbar\omega_{\mathbf{k}}}} \left(\omega_{\mathbf{k}}\hat{\mathbf{q}}_{\mathbf{k},s} + i\hat{\mathbf{p}}_{\mathbf{k},s}\right),$$

$$\hat{\mathbf{a}}_{\mathbf{k},s}^{\dagger} = \frac{1}{\sqrt{2\hbar\omega_{\mathbf{k}}}} \left(\omega_{\mathbf{k}}\hat{\mathbf{q}}_{\mathbf{k},s} - i\hat{\mathbf{p}}_{\mathbf{k},s}\right).$$

(3.15)

These operators satisfy the commutation relations

$$\left[\hat{\mathbf{a}}_{\mathbf{k},s}, \hat{\mathbf{a}}_{\mathbf{k}',s'}\right] = 0 = \left[\hat{\mathbf{a}}_{\mathbf{k},s}^{\dagger}, \hat{\mathbf{a}}_{\mathbf{k}',s'}^{\dagger}\right], \tag{3.16}$$

$$\left[\hat{\mathbf{a}}_{\mathbf{k},s}, \hat{\mathbf{a}}_{\mathbf{k}',s'}^{\dagger}\right] = \delta^{(3)}\left(\mathbf{k} - \mathbf{k}'\right)\delta_{s,s'}, \tag{3.17}$$

and are extremely useful on account of their special properties, which we will review shortly. We can now redefine the Hamiltonian as

$$\hat{\mathbf{H}} = \sum_{\mathbf{k},s} \hbar\omega_k \left(\hat{\mathbf{a}}_{\mathbf{k},s}^{\dagger}\hat{\mathbf{a}}_{\mathbf{k},s} + \frac{1}{2}\right). \tag{3.18}$$

It is clear from this equation that a quantum mechanical harmonic oscillator is associated with each mode $\mathbf{k}$ of the radiation field. Since there are infinitely many modes $\mathbf{k}$, the ground state energies $\hbar\omega_{\mathbf{k}}/2$ of these oscillators sum up to infinity. This problem was solved by setting the ground state energy (or vacuum energy) to zero, which amounts to omitting the factor $1/2$ in the Hamiltonian $\hat{\mathbf{H}}$.

The operators $\hat{\mathbf{a}}_{\mathbf{k},s}^{\dagger}$ and $\hat{\mathbf{a}}_{\mathbf{k},s}$ are known as the creation and annihilation operators respectively, since they create and destroy a quantum of energy $\hbar\omega_{\mathbf{k}}$, or a photon, in the mode of the electromagnetic field defined by the wave vector $\mathbf{k}$ and the polarisation vector $\epsilon_{\mathbf{k},s}$. The number of photons, or elementary excitations, in this mode is given by the non-negative eigenvalue $n_{\mathbf{k},s}$ of the number operator $\hat{\mathbf{n}}_{\mathbf{k},s} = \hat{\mathbf{a}}_{\mathbf{k},s}^{\dagger}\hat{\mathbf{a}}_{\mathbf{k},s}$. The excitation level of the mode is determined by its eigenstate $|n_{\mathbf{k},s}\rangle$ and the creation and annihilation operators are characterised by their effect on this state:

$$\hat{\mathbf{a}}_{\mathbf{k},s}^{\dagger}|n_{\mathbf{k},s}\rangle = \sqrt{n_{\mathbf{k},s} + 1}|n_{\mathbf{k},s} + 1\rangle, \tag{3.19}$$

$$\hat{\mathbf{a}}_{\mathbf{k},s}|n_{\mathbf{k},s}\rangle = \sqrt{n_{\mathbf{k},s}}|n_{\mathbf{k},s} - 1\rangle. \tag{3.20}$$

The vacuum state of the field is defined by

$$\hat{\mathbf{a}}_{\mathbf{k},s}|0\rangle = 0, \tag{3.21}$$

and the eigenstate $|n_{\mathbf{k},s}\rangle$ can be written in terms of the creation operator and the vacuum state as

$$|n_{\mathbf{k},s}\rangle = \frac{1}{\sqrt{n_{\mathbf{k},s}!}} \left(\hat{\mathbf{a}}_{\mathbf{k},s}^{\dagger}\right)^{n_{\mathbf{k},s}} |0\rangle. \tag{3.22}$$

These eigenstates are referred to as *number states* or *Fock states* and form a complete orthogonal set.

In order to find the quantum mechanical representations of the vector potential $\mathbf{A}$ and the $\mathbf{E}$ and $\mathbf{B}$ fields we need to convert the complex numbers $c_{\mathbf{k},s}$ and $c_{\mathbf{k},s}^{*}$ into the operators $\hat{a}_{\mathbf{k},s}$ and $\hat{a}_{\mathbf{k},s}^{\dagger}$. The relationship between these quantities is found by comparing Eqs. (3.10) and (3.15):

$$\hat{c}_{\mathbf{k},s} = \frac{\hbar}{\sqrt{2\omega_{\mathbf{k}}\varepsilon_0}}\hat{a}_{\mathbf{k},s}, \tag{3.23}$$

$$\hat{c}_{\mathbf{k},s}^{\dagger} = \frac{\hbar}{\sqrt{2\omega_{\mathbf{k}}\varepsilon_0}}\hat{a}_{\mathbf{k},s}^{*}, \tag{3.24}$$

where we have transformed the complex numbers $c_{\mathbf{k},s}$ of the classical regime into quantum operators $\hat{c}_{\mathbf{k},s}$ so that the comparison in this equation was possible. Hence, the representations for the quantised $\mathbf{A}$ and the $\mathbf{E}$ and $\mathbf{B}$ fields are given by

$$\hat{\mathbf{A}} = \sum_{\mathbf{k},s} \left(\frac{\hbar}{2\omega_{\mathbf{k}}\varepsilon_0 V}\right)^{\frac{1}{2}} \boldsymbol{\epsilon}_{\mathbf{k},s} \left[\hat{a}_{\mathbf{k},s} e^{i(\mathbf{k}\cdot\mathbf{r}-\omega_{\mathbf{k}}t)} + \hat{a}_{\mathbf{k},s}^{\dagger} e^{-i(\mathbf{k}\cdot\mathbf{r}-\omega_{\mathbf{k}}t)}\right], \tag{3.25}$$

$$\hat{\mathbf{E}} = -\frac{\partial \mathbf{A}}{\partial t} = i \sum_{\mathbf{k},s} \left(\frac{\hbar\omega_{\mathbf{k}}}{2\varepsilon_0 V}\right)^{\frac{1}{2}} \boldsymbol{\epsilon}_{\mathbf{k},s} \left[\hat{a}_{\mathbf{k},s} e^{i(\mathbf{k}\cdot\mathbf{r}-\omega_{\mathbf{k}}t)} - \hat{a}_{\mathbf{k},s}^{\dagger} e^{-i(\mathbf{k}\cdot\mathbf{r}-\omega_{\mathbf{k}}t)}\right], \tag{3.26}$$

$$\hat{\mathbf{B}} = \nabla \times \mathbf{A} = \frac{i}{c} \sum_{\mathbf{k},s} \left(\frac{\hbar\omega_{\mathbf{k}}}{2\varepsilon_0 V}\right)^{\frac{1}{2}} \left(\frac{\mathbf{k}}{|\mathbf{k}|} \times \boldsymbol{\epsilon}_{\mathbf{k},s}\right) \left[\hat{a}_{\mathbf{k},s} e^{i(\mathbf{k}\cdot\mathbf{r}-\omega_{\mathbf{k}}t)} - \hat{a}_{\mathbf{k},s}^{\dagger} e^{-i(\mathbf{k}\cdot\mathbf{r}-\omega_{\mathbf{k}}t)}\right]. \tag{3.27}$$

All three of these operators above are self-adjoint (i.e. $\hat{\mathbf{A}} = \hat{\mathbf{A}}^{\dagger}$) but only $\hat{\mathbf{E}}$ and $\hat{\mathbf{B}}$ are directly measurable and therefore represent observables. These fields operators, unlike the Hamiltonian, are time-dependent and correspond to the Heisenberg picture of quantum mechanics which describes the evolution of the system by means of time-dependent operators rather than time-dependent states. If we want to obtain their representation in the Schrödinger picture (where states are time-dependent and operators are not), then we simply neglect the time-dependent factors, $\exp(i\omega_{\mathbf{k}}t)$.

As mentioned before, the creation operator in Eq. (3.23) creates a photon in mode $\mathbf{k}$ of the electromagnetic field which describes its energy, its polarisation and its spatial mode. Information can be encoded in any of these degrees of freedom of the electromagnetic field, either in the number of photons in the mode (as we have seen) or in the field itself, through its time-frequency distribution, its polarisation or its OAM [33].

3.2.3 Representation of a single photon in an OAM mode

In the previous section we have discussed the decomposition of an electromagnetic field in the plane-wave basis. This is, however, not the only basis in which the field can be described. For example, paraxial light that is emitted from a laser cavity with rectangular symmetry can be decomposed as a sum of the orthogonal HG modes. We are concerned with paraxial beams that carry OAM and a convenient basis for their description is the complete and orthogonal basis of LG

modes since these modes are eigenmodes of the OAM operator [34]. We will now demonstrate the decomposition of the electromagnetic field in this basis according to the method used by Calvo *et al.* [35] and show how to represent the state of a single photon in an LG mode. This section serves to introduce the main idea behind the decomposition of light in different bases and provides a summary of some key points but a complete derivation can be found in both [35] and [36].

Let us consider the continuous version of the vector potential $\mathbf{A}$:

$$\hat{\mathbf{A}}(\mathbf{r}, t) = \sum_s \int d^3\mathbf{k} \left(\frac{\hbar}{16\pi^3 \varepsilon_0 \omega_\mathbf{k}}\right)^{\frac{1}{2}} \boldsymbol{\epsilon}_{\mathbf{k},s} \left[\hat{\mathbf{a}}_{\mathbf{k},s} e^{i(\mathbf{k}\cdot\mathbf{r} - \omega_\mathbf{k} t)} + \text{h.c.}\right], \quad (3.28)$$

where h.c. represents the Hermitian conjugate of the first term. This equation is true for a general electromagnetic light beam, but we are interested in paraxial light beams so we need find a representation of Eq. (3.28) in the paraxial limit. Calvo *et al.* derived an equivalent expression for $\mathbf{A}$ which clearly distinguishes the paraxial and non-paraxial contributions of the beam.

In order to do this the wave vector $\mathbf{k}$ is first separated into its longitudinal (k_z) and transverse (q) parts

$$\mathbf{k} = q\mathbf{q} + k_z\hat{\mathbf{z}}, \quad (3.29)$$

and a generalised dispersion relation is introduced

$$\frac{k_z}{k_0} = 1 - \vartheta^2, \quad (3.30)$$

where $\vartheta := q/\sqrt{2k_0^2}$ is defined as the parameter that governs the degree of paraxiality. It is worthwhile to observe that $\vartheta\sqrt{2} = q/k_0$ is approximately equal to the divergence angle of a Gaussian beam [36]. This allows for the vector potential to be written as

$$\hat{\mathbf{A}}(\mathbf{r}, t) = \sum_s \int_0^\infty dk_0 \int d^2\mathbf{q} \left[\frac{\hbar(1+\vartheta^2)^2}{16\pi^3 \varepsilon_0 c k_0 \sqrt{1+\vartheta^4}}\right]^{\frac{1}{2}} \boldsymbol{\epsilon}_{\mathbf{k},s} \Bigg\{\hat{\mathbf{a}}_{\mathbf{k},s} e^{ik_0(z-ct)}$$
$$\times \exp\left[i\mathbf{q}\cdot\mathbf{r}_\perp - ik_0\vartheta^2 z - ick_0\left(\sqrt{1+\vartheta^4} - 1\right)t\right] + \text{h.c.}\Bigg\}, \quad (3.31)$$

where $\omega_\mathbf{k} = ck_0$ and $\mathbf{r}_\perp = (x, y)$ is the two-dimensional transverse position vector, so that the paraxial and non-paraxial contributions are clearly displayed. By taking the paraxial limit, $\vartheta \ll 1$, the vector potential could then be expressed as

$$\hat{\mathbf{A}}(\mathbf{r}, t) = \sum_s \int_0^\infty dk_0 \left(\frac{\hbar}{16\pi^3 \varepsilon_0 c k_0}\right)^{\frac{1}{2}} \boldsymbol{\epsilon}_s \left[\hat{\mathbf{a}}_s(\mathbf{q}, k_0) e^{ik_0(z-ct)} e^{i(\mathbf{q}\cdot\mathbf{r}_\perp - k_0\vartheta^2 z)} + \text{h.c.}\right], \quad (3.32)$$

where the ϑ^2 term in the second phase factor is retained because it is the only relevant paraxial contribution and the polarisation vectors are now independent of the wave vectors.

The aim was to express the vector potential $\hat{\mathbf{A}}$ in terms of the transverse basis of LG modes.

To do this one should first observe that the field amplitude of an LG mode propagating in the z direction can be represented as [37]

$$LG_{pl}(\mathbf{r}_\perp, z) = \frac{1}{(2\pi)^2} \int d^2\mathbf{q} \mathcal{L}\mathcal{G}_{pl}(\mathbf{q}, z)e^{i\mathbf{q}\cdot\mathbf{r}_\perp}, \tag{3.33}$$

where, for simplicity, we have used the notation LG_{pl} instead of u_{pl}^{LG} (given in Eq. (2.45)) to refer to an LG mode with mode indices p and l. $\mathcal{L}\mathcal{G}_{pl}$ denotes the field amplitude of the Fourier-transformed LG modes and can be written as

$$\mathcal{L}\mathcal{G}_{pl}(\rho, \varphi) = \sqrt{\frac{\omega_0^2 p!}{2\pi(|l| + p)!}} \left(\frac{\omega_0 \rho}{\sqrt{2}}\right)^{|l|} L_p^{|l|}\left(\frac{\omega_0^2 \rho^2}{2}\right) \exp\left(-\frac{\omega_0^2 \rho^2}{4}\right)$$
$$\times \exp\left(il\varphi - i\frac{\pi}{2}(2p + |l|)\right), \tag{3.34}$$

where ρ and φ denote the frequency-space cylindrical coordinates. The closure relation for these modes is given by

$$\sum_{p,l} \mathcal{L}\mathcal{G}_{pl}^*(\mathbf{q})\mathcal{L}\mathcal{G}_{pl}(\mathbf{q}') = \delta^{(2)}(\mathbf{q} - \mathbf{q}'). \tag{3.35}$$

These modes are also known as the angular spectrum of the LG modes and can be represented by the inverse Fourier transform of the LG mode amplitudes

$$\mathcal{L}\mathcal{G}_{pl}(\mathbf{q}, z) = \int d^2\mathbf{r}_\perp LG(\mathbf{r}_\perp, z) e^{-i\mathbf{q}\cdot\mathbf{r}_\perp}. \tag{3.36}$$

Using the paraxial wave equation (Eq. (2.11)) and Eq. (3.33), it can easily be shown that if the angular spectrum of the field is known at $z = 0$ then its propagation in the z direction is

$$\mathcal{L}\mathcal{G}_{pl}(\mathbf{q}, z) = \mathcal{L}\mathcal{G}(\mathbf{q}, 0)e^{-ik_z z}. \tag{3.37}$$

Using this preliminary information we can see why Calvo *el al.* identified the $\mathbf{q}$-integrand on the right hand side of Eq. (3.32) as the paraxial angular spectrum and thus simplified the exponential factor in Eq. (3.32) as follows

$$e^{i\left(\mathbf{q}\cdot\mathbf{r}_\perp - k_0 \vartheta^2 z\right)} = \int d^2\mathbf{q}' e^{i\left(\mathbf{q}'\cdot\mathbf{r}_\perp - k_0 \vartheta^2 z\right)} \delta^{(2)}\left(\mathbf{q} - \mathbf{q}'\right)$$

$$= \int d^2\mathbf{q}' e^{i\left(\mathbf{q}'\cdot\mathbf{r}_\perp - k_0 \vartheta^2 z\right)} \left[\sum_{p,l} \mathcal{L}\mathcal{G}_{pl}^*(\mathbf{q})\,\mathcal{L}\mathcal{G}_{pl}(\mathbf{q}')\right]$$

$$= \sum_{p,l} \mathcal{L}\mathcal{G}_{pl}^*(\mathbf{q}) \int d^2\mathbf{q}' e^{i\left(\mathbf{q}'\cdot\mathbf{r}_\perp - k_0 \vartheta^2 z\right)} \mathcal{L}\mathcal{G}_{pl}(\mathbf{q}')$$

$$= \sum_{p,l} \mathcal{L}\mathcal{G}_{pl}^*(\mathbf{q}) LG_{pl}(\mathbf{r}_\perp, z). \tag{3.38}$$

They could then introduce the creation and annihilation operators for the LG modes:

$$\hat{\mathbf{a}}^{\dagger}_{s,p,l}(k_0) = \int d^2\mathbf{q} \mathcal{LG}_{pl}(\mathbf{q})\hat{\mathbf{a}}^{\dagger}_s(\mathbf{q}, k_0), \tag{3.39}$$

$$\hat{\mathbf{a}}_{s,p,l}(k_0) = \int d^2\mathbf{q} \mathcal{LG}^*_{pl}(\mathbf{q}) \hat{\mathbf{a}}_s(\mathbf{q}, k_0). \tag{3.40}$$

By using the commutation relation $\left[\hat{\mathbf{a}}_s(k_0), \hat{\mathbf{a}}^{\dagger}_{s'}(k'_0)\right] = \delta_{ss'}\delta(k_0 - k'_0)$ and the orthogonality condition $\int d^2\mathbf{q} \mathcal{LG}^*_{pl}(\mathbf{q})\mathcal{LG}_{p'l'}(\mathbf{q}) = \delta_{pp'}\delta_{ll'}$ they found that the new creation and annihilation operators satisfy the commutation relation

$$\left[\hat{\mathbf{a}}_{s,p,l}(k_0), \hat{\mathbf{a}}^{\dagger}_{s',p',l'}(k'_0)\right] = \delta_{ss'}\delta_{ll'}\delta_{pp'}\delta(k_0 - k'_0). \tag{3.41}$$

Eq. (3.32) thus becomes

$$\hat{\mathbf{A}}(\mathbf{r},t) = \sum_s \int_0^{\infty} dk_0 \left(\frac{\hbar}{16\pi^3\varepsilon_0 ck_0}\right)^{\frac{1}{2}} \boldsymbol{\epsilon}_s \left[\hat{\mathbf{a}}_{s,p,l}(k_0) e^{ik_0(z-ct)} LG_{pl}(\mathbf{r}_{\perp}, z) + \text{h.c.}\right]. \tag{3.42}$$

The creation operator in the Hermitian conjugate part of Eq. (3.42) creates a photon in the mode (s, p, l) of the field which specifies its polarisation and, more specifically, its spatial distribution which is that of a particular LG mode. The most general paraxial single-photon state in the LG basis is then given by

$$|\psi\rangle = \sum_{s,p,l} \int_0^{\infty} dk_0 C_{s,p,l}(k_0) \hat{\mathbf{a}}^{\dagger}_{s,p,l}(k_0) |0\rangle, \tag{3.43}$$

where the complex coefficients $C_{s,p,l}$ satisfy the normalisation condition $\sum_{s,p,l} \int dk_0 |C_{s,p,l}|^2 = 1$. This is a Fock state and the complex coefficients represent the probability amplitude of finding the photons in the state $|s, p, l, k_0\rangle$ with polarisation s, wave vector k_0 along the z-axis and corresponding to an excitation of an LG mode with radial mode index p and azimuthal mode index l. The Fock state of a photon in a linearly polarised, monochromatic beam in the LG basis is given by

$$|\psi_s\rangle = \sum_{p,l} f_{p,l} \hat{\mathbf{a}}^{\dagger}_{s,p,l}(k_0) |0\rangle, \tag{3.44}$$

where the complex coefficients $f_{p,l}$ once again satisfy the normalisation condition $\sum_{p,l} |f_{p,l}|^2 = 1$ and represent the probability of finding the photon in the state $|pl\rangle$.

For our proof-of-principle experiment we use photons in a single linearly polarised monochromatic LG mode (with specific p and l values). The state of a single photon in this type of mode is given by

$$|pl\rangle = \hat{\mathbf{a}}^{\dagger}_{s,p,l}(k_0) |0\rangle = \int d^2\mathbf{q} \mathcal{LG}_{pl}(\mathbf{q})\hat{\mathbf{a}}^{\dagger}_s(\mathbf{q}, k_0) |0\rangle. \tag{3.45}$$

Furthermore, we are concerned only with the OAM of the LG mode and, therefore, we use a subset of the LG modes with all values of l but with $p = 0$. The state of such a photon (which possesses

an OAM of $l\hbar$) can thus be represented by

$$|l\rangle = \int d^2 \mathbf{q} \mathcal{L}\mathcal{G}_{0l}(\mathbf{q})\hat{\mathbf{a}}_s^\dagger(\mathbf{q}, k_0)|0\rangle. \tag{3.46}$$

3.3 Dirac Notation for Laser beams

The well-known analogy between paraxial optics and quantum mechanics has motivated many authors [32, 38, 39] to apply Dirac notation, a technique generally used to represent states in quantum mechanics, to represent the field amplitude of paraxial light beams. Operator algebra is then employed to describe the propagation of these light beams through free space and lenses. We shall now briefly review the use of Dirac notation in quantum mechanics. This will assist us in demonstrating the application of this formalism to paraxial light beams in a fairly transparent manner. We shall then mention a few cases in which the application of Dirac notation and operator algebra has simplified the analysis of paraxial light beam propagation.

3.3.1 Preliminary information: Dirac notation

In quantum mechanics the state of a physical system can be described by a wave function $\Psi(\mathbf{r}, t)$. The change of the state of the system in time is governed by the Schrödinger equation. In the case of the free particle in a two-dimensional box, for example, the time evolution can be given by

$$\begin{aligned} i\hbar \frac{\partial}{\partial t}\Psi(\mathbf{r}, t) &= \hat{\mathbf{H}}\Psi(\mathbf{r}, t) \\ &= \frac{\hat{\mathbf{p}}^2}{2m}\Psi(\mathbf{r}, t) \\ &= -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)\Psi(\mathbf{r}, t), \end{aligned} \tag{3.47}$$

where $\hat{\mathbf{H}} = \hat{\mathbf{p}}^2/(2m)$ is the Hamiltonian operator and the momentum operator is defined as $\hat{\mathbf{p}} = -i\hbar\nabla$. The dynamics of the system is encapsulated in the expression

$$\Psi(\mathbf{r}, t) = \hat{\mathbf{U}}\Psi(\mathbf{r}, t_0), \tag{3.48}$$

where the operator $\hat{\mathbf{U}}$ is aptly named the *time-evolution operator* and is given by

$$\hat{\mathbf{U}} = \exp\left[-\frac{i}{\hbar}\hat{\mathbf{H}}(t - t_0)\right]. \tag{3.49}$$

Dirac notation, named for Paul Dirac, is an abstract mathematic formalism which allows us to represent the wave function $\Psi(\mathbf{r}, t)$ by a ket $|\Psi(t)\rangle$. This is a complex vector, called a *state vector*, in a complex vector space with inner product, called a Hilbert space. The notation $\langle\Psi(t)|$ is used to denote the dual vector to $|\Psi(t)\rangle$, and is known as a *bra*. One possible basis for this space is

given by the position kets $|\mathbf{r}\rangle$ and in this notation if we take the inner product with the state vector then we can recover the wave function,

$$\langle \mathbf{r}|\Psi(t)\rangle = \Psi(\mathbf{r}, t). \tag{3.50}$$

The normalisation condition (and inner product) for wave functions in this space is given by

$$\begin{aligned} 1 &= \langle \Psi(t)|\Psi(t)\rangle \\ &= \int d\mathbf{r}\langle \Psi(t)|\mathbf{r}\rangle\langle \mathbf{r}|\Psi(t)\rangle \\ &= \int d\mathbf{r}\Psi(\mathbf{r}, t)\Psi^*(\mathbf{r}, t). \end{aligned} \tag{3.51}$$

Finally, the time-evolution of the state vector can be calculated using the expression

$$|\Psi(t)\rangle = \hat{\mathbf{U}}|\Psi(0)\rangle. \tag{3.52}$$

3.3.2 Application to paraxial light beams

The paraxial wave equation

$$2ik\frac{\partial}{\partial z}u(x, y, z) = -\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)u(x, y, z), \tag{3.53}$$

has two distinct properties which suggest that Dirac notation and operator algebra could be applied to the longitudinal propagation of paraxial light beams. Firstly, the form of the paraxial wave equation is analogous to the Schrödinger equation given in Eq. (3.47) with the z coordinate replacing the time variable t and the wavelength replacing Planck's constant. Secondly, the paraxial wave equation has the explicit solution

$$u(x, y, z) = \exp\left[\frac{i(z - z_0)}{2k}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)\right]u(x, y, z_0), \tag{3.54}$$

which tells us that if we know the transverse profile of the beam at a particular point z_0 we can use the exponential factor to find the transverse profile of the beam at a different position z. This means that this exponential factor is analogous to the time-evolution operator. We can, therefore, transform Eq. (3.54) into an operator relation in Dirac notation as

$$|u(z)\rangle = \hat{\mathbf{U}}|u(z_0)\rangle. \tag{3.55}$$

The longitudinal propagation operator $\hat{\mathbf{U}}$ is given by

$$\hat{\mathbf{U}} = \exp\left[-\frac{i(z - z_0)}{2k}\hat{\mathbf{p}}^2\right], \tag{3.56}$$

where $\hat{\mathbf{p}}^2 = \hat{\mathbf{p}}_x^2 + \hat{\mathbf{p}}_y^2$ is the abstract operator representative of $-\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)$.

The commutation relations between the position and momentum operators define a Heisenberg

algebra and resemble the ones in quantum mechanics, with $\hbar$ set to 1. The state vector $|u(z)\rangle$ lies in the Hilbert space of $L^2(\mathbb{C}^2)$ and is normalised

$$\langle u(z)|u(z)\rangle \equiv \int d^2\mathbf{r}_\perp u^*(\mathbf{r}_\perp, z)u(\mathbf{r}_\perp, z) = 1, \tag{3.57}$$

where $\mathbf{r}_\perp = (x, y)^T$ is the position vector in the transverse plane. As well as showing the normalisation of the state vectors, this equation shows how the inner product is defined on this space. When $|u(z)\rangle$ is normalised for one position z, it is automatically normalised for all other z values [32]. One possible set of basis vectors in this space are the position kets $|\mathbf{r}_\perp\rangle$, which are eigenvectors of the transverse position operator $\hat{\mathbf{r}}_\perp = (\hat{\mathbf{x}}, \hat{\mathbf{y}})$ that acts on state vectors in this space. If we take the inner product of the state vector $|u(z)\rangle$ with a basis vector we recover the beam profile

$$\langle \mathbf{r}_\perp|u(z)\rangle = u(x, y, z). \tag{3.58}$$

One should observe that even though we are using the Dirac notation, which is usually used in quantum mechanics, we are still in the purely classical regime. We are not dealing with single photons, but rather a beam which consists of a stream of photons.

The expectation value of an operator $\hat{\mathbf{O}}$ is defined in the standard way as

$$\langle O \rangle = \langle u(z)|\hat{\mathbf{O}}|u(z)\rangle = \int d^2\mathbf{r}_\perp u^*(\mathbf{r}_\perp, z)\hat{\mathbf{O}}u(\mathbf{r}_\perp, z). \tag{3.59}$$

In this equation the states are z-dependent and the operators "stationary", which is the Schrödinger picture. In the Heisenberg picture the operators become z-dependent, so that the evolution of the expectation value of an operator can be expressed as

$$\langle O \rangle = \langle u(z_0)|\hat{\mathbf{U}}\hat{\mathbf{O}}\hat{\mathbf{U}}|u(z_0)\rangle. \tag{3.60}$$

Another complete orthogonal basis set for the state vectors $|u(z)\rangle$ is given by the transverse LG modes, $|u_{pl}(z)\rangle$. This is because the LG modes LG_{pl} are a set of solutions to the paraxial wave equation in cylindrical coordinates. Also, the $|u_{pl}(z)\rangle$ functions are stationary in the z coordinate (which is analogous to stationary states in time in quantum mechanics), except for a z-dependent scale change and a radial phase factor. So some properties such as the shape and the azimuthal phase dependence remain invariant [40]. A state vector in this space can, therefore, be expanded as a superposition of LG modes

$$|u(z)\rangle = \sum_{p,l} \alpha_{p,l}|u_{pl}(z)\rangle, \tag{3.61}$$

where $\alpha_{p,l} = \langle u_{pl}(z)|u(z)\rangle$ since the LG modes are orthonormal and $\sum_{p,l}|\alpha_{p,l}|^2 = 1$ since the state vectors are normalised.

It is well known that the analytic form of the LG modes resemble the wave functions of the stationary states of a two-dimensional quantum mechanical harmonic oscillator. Therefore, Nienhuis

and Allen [32] were able to use operator algebra to obtain ladder operators of the harmonic oscillator which allow one to generate higher-order LG modes from the fundamental Gaussian mode $|u_{00}(z)\rangle$.

We can see how the operator formalism simplifies the calculations of some quantities associated with the electric field by considering an example given in [40]. The full scalar version of the electric field, which includes the time dependence, can be written as

$$\langle \mathbf{r}_\perp | E \rangle = E_0 e^{i(kz - wt)} \langle \mathbf{r}_\perp | u \rangle \tag{3.62}$$

where E_0 is the amplitude of the field and $u \equiv u(z)$, for convenience. If $u(r, \phi, z) = u_0(r, z) e^{il\phi}$, which is the form of an LG mode, then the expectation value of the angular momentum operator is

$$\langle L_z \rangle = \left\langle E \left| i \frac{\partial}{\partial \phi} \right| E \right\rangle = |E_0|^2 l \langle u | u \rangle. \tag{3.63}$$

The linear momentum of the state in the direction of propagation z is given by

$$\langle p_z \rangle = \left\langle E \left| i \frac{\partial}{\partial z} \right| E \right\rangle = |E_0|^2 k \langle u | u \rangle, \tag{3.64}$$

while the energy of the state amounts to

$$\langle W \rangle = \int \int dx dy E^*(\mathbf{r}_\perp, t) \left(i \frac{\partial}{\partial t} \right) E(\mathbf{r}_\perp, t) = |E_0|^2 w \langle u | u \rangle. \tag{3.65}$$

Therefore, the ratio of the OAM carried by the wave to its energy is (cp. Eq. (2.22))

$$\frac{\langle L_z \rangle}{\langle W \rangle} = \frac{l}{\omega}, \tag{3.66}$$

while the ratio of OAM to linear momentum is (cp. Eq. (2.23))

$$\frac{\langle L_z \rangle}{\langle p_z \rangle} = \frac{l}{k}. \tag{3.67}$$

These relations are valid classically as well as quantum mechanically, since $\hbar$ does not appear in them. Every photon has an energy of $\hbar\omega$ and a linear momentum of $\hbar k$ so for the ratios to be valid on a single-photon level, each photon must possess a well-defined OAM of $l\hbar$. Thus we have, in a simple way, arrived at the conclusion reached by Allen *et al.* [8], that every photon in an LG beam carries a quantised OAM of $l\hbar$ per photon. Also note that due to Eq. (3.60) the quantities in Eqs. (3.63)–(3.65) are all conserved during free space propagation since the operators in these equations commute with $\hat{\mathbf{p}}^2$ of the propagation operator $\hat{\mathbf{U}}$ [39].

This operator formalism that we have developed is not limited to free-space propagation. Indeed many authors [32, 38, 39] have applied it to the propagation of beams through thin lenses. We can

implement the effect of such a lens with a focal length f by means of an operator $\hat{\mathbf{T}}_{\text{real}}$ that is a function of the transverse coordinate operators $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ [38]:

$$\hat{\mathbf{T}}_{\text{real}} = \exp\left[-\frac{ik}{2f}\left(\hat{\mathbf{x}}^2 + \hat{\mathbf{y}}^2\right)\right].\tag{3.68}$$

Let us consider a system which consists of two identical cylindrical lenses orientated along the x-axis with focal length f and separated from each other by a distance $2d$ with free propagation between the lenses. The propagation operator for this system is given by

$$\hat{\mathbf{U}}_{\text{cyl}} = \hat{\mathbf{T}}_{\text{real}}\hat{\mathbf{U}}\hat{\mathbf{T}}_{\text{real}} = \exp\left[-i\frac{k\hat{\mathbf{x}}^2}{2f}\right]\exp\left[-i\frac{d\hat{\mathbf{p}}^2}{k}\right]\exp\left[-i\frac{k\hat{\mathbf{x}}^2}{2f}\right].\tag{3.69}$$

Using this propagation operator van Enk and Nienhuis [39] were able to prove, in a simple way, that if $d = f/\sqrt{2}$ then this set-up converts an LG mode which contains OAM, into an HG beam that contains no OAM. This set-up is known as an astigmatic laser mode converter (cf. Section 5.2), and in [41], Allen *et al.* arrived at this conclusion in a more complex way.

3.4   Analogy

Consider the state in Eq. (3.44) which is a single-mode Fock state in the LG basis. This state represents the creation of a photon or a single excitation with a probability amplitude $f_{p,l}$ in the LG mode LG_{pl} and vacuum in all other modes (i.e. modes with different p and l values). In [42] the authors showed that this state could be written as $|\psi\rangle = \int\langle\mathbf{r}_\perp|\psi\rangle|\mathbf{r}_\perp\rangle d^2\mathbf{r}_\perp$, where $\psi(\mathbf{r}_\perp) = \langle\mathbf{r}_\perp|\psi\rangle$ is the transverse wave function of the photon in position representation. They then adopted operator algebra by denoting the wave function as the ket $|\psi(\mathbf{r}_\perp, z)\rangle$ and used the operators we have discussed in the previous section to propagate this wave function in free space and through thin lenses. Their research served as confirmation that optical elements function analogously on a single-photon level.

We have seen that a general paraxial light beam in classical optics with state vector $|u(z)\rangle$ can be expressed as a superposition of the LG state vectors $|u_{pl}(z)\rangle$ since they form a basis for the description of such beams (cf. Eq. (3.61)). With

$$|\langle\mathbf{r}_\perp|u(z)\rangle|^2 = \sum_{p,l}|\alpha_{p,l}|^2|\langle\mathbf{r}_\perp|u_{pl}(z)\rangle|^2,\tag{3.70}$$

we obtain the intensity of the electric field, where $|\alpha_{p,l}|^2$ represents the intensity of the LG mode LG_{pl}. Note that the state $|pl\rangle$ represents a single excitation of the LG mode LG_{pl} and the $|u_{pl}(z)\rangle$ state vectors are normalised. Consequently, the probability of obtaining an excitation of the LG mode LG_{pl}, given by $|f_{p,l}|^2$, is identical to the intensity of that mode, $|\alpha_{p,l}|^2$.

Although there is an analogy between the state of a classical light beam and that of a single photon we can observe the difference if we perform a measurement of each of these entities. $|\alpha_{p,l}|^2$ is a measurable and predetermined quantity while $|f_{p,l}|^2$ represents only the probability of obtaining an excitation of the LG mode LG_{pl}, or the probability of getting a "click" in a detector when

projecting onto the state $|pl\rangle$. Moreover, the single-photon state is fundamentally different from the classical beam since it is possible to entangle single photons so that their state before a measurement is not clearly defined. Another example of the difference would be the Hong-Ou-Mandel effect [43] which states that if two single photons arrive at a $50/50$ beam splitter simultaneously then both photons must exit from the same output port. These phenomenon do not occur in the case of classical beams.

Chapter 4

Spontaneous Parametric Down-Conversion

In quantum information processing, the physical system used to encode information is governed by the laws of quantum mechanics, which means that it exhibits properties that cannot be explained classically. One of these fascinating features of quantum mechanics is entanglement. Entanglement allows for the quantum states of two particles to be connected in such a way that measuring the state of one particle instantaneously defines the state of the other even though neither particle is in its own well-defined state before the measurement [11]. This non-classical correlation between entangled particles lies at the heart of quantum mechanics. It plays a crucial role in quantum information processing and quantum computing since it allows certain tasks to be performed at speeds much faster and in a more secure way than would be possible classically. This quantum mechanical phenomenon will be described in more detail in Chapter 6.

Since entanglement is one of the key ingredients for our proof-of-principle theoretical experiment we need a source of entangled photons. Spontaneous parametric down-conversion provides a reliable, albeit probabilistic, source of such photons. When an intense pump laser beam is incident on a non-linear crystal there is a possibility that two lower-frequency photons will be generated from a single pump photon and this non-linear optical phenomenon is known as spontaneous parametric down-conversion (SPDC). The generated two-photon state can be entangled in polarisation, but it also has many other accessible degrees of freedom that can be used to carry entanglement. These include linear momentum [44], energy-time [45], and time-bins [46]. Of special interest to us is the entanglement of the OAM states of a generated biphoton, which was demonstrated in [11]. An immediate advantage of this multi-dimensional entangled state is that it offers us a large Hilbert space for quantum information processing.

This chapter provides a brief outline of the process of SPDC because a detailed analysis is out of the scope of this research. Section 4.1 explains the non-linearity that exists in some optical materials while Section 4.2 shows how the second-order non-linearity can be exploited to produce three-wave mixing processes such as SPDC. We then give short description of the SPDC process in Section 4.3 including a discussion about suitable crystals for inducing this process and phase-

matching conditions that must be satisfied in order for entangled photon pairs to be produced. For a more in-depth look at the SPDC process we calculate the effective Hamiltonian and the resulting two-photon state (in Section 4.4) which allows us to show how the phase-matching conditions arise. In the final two sections we discuss the entanglement of the down-converted photons in two degrees of freedom, namely polarisation (in Section 4.5) and OAM (in Section 4.6).

4.1 Non-linear optical materials

When an electromagnetic field propagates through a dielectric medium, the electric field acts on each particle (electrons, atoms or molecules) in the medium causing a redistribution of the charges. There is a displacement of the positive charges in the direction of the field and the negative charges in the opposite direction resulting in the polarisation of the material. For small electric field amplitudes, as is the case in linear optics, there is an approximately linear relationship between the induced dielectric polarisation $\mathbf{P}$ and the incident electric field

$$\mathbf{P} \approx \varepsilon_0 \chi \mathbf{E}, \tag{4.1}$$

where ε_0 is the electric permittivity of free space and χ is the linear susceptibility which is a measure of how easily the dielectric material polarises in response to an electric field. We are, however, interested in non-linear optical phenomena where lasers, which usually output large amplitude electric fields, are used. Therefore, higher-order contributions of the electric field become significant and a non-linear response of the material to the incident field appears

$$\mathbf{P}(\mathbf{E}) = \varepsilon_0 \left[\chi^{(1)}\mathbf{E} + \chi^{(2)}\mathbf{E}^2 + \chi^{(3)}\mathbf{E}^3 + \dots \right], \tag{4.2}$$

where $\chi^{(n)}$ are the electric susceptibility tensors of order n (with $\chi^{(1)} \equiv \chi$). The non-linear susceptibilities usually have small magnitudes (i.e. $\chi^{(1)} \approx 1$, $\chi^{(2)} \approx 10^{-10}$ and $\chi^{(3)} \approx 10^{-17}$).

In optics the dielectric material that we use is a crystal. A crystal is called non-linear if it has a non-linear response to an incident electric field (as in Eq. (4.2)). Although all crystalline materials give a non-zero $\chi^{(3)}$ coefficient only an anisotropic or non-centrosymmetric crystal (which is a crystal with no symmetry centre) gives a non-zero square term in the polarisation expansion ($\chi^{(2)} \neq 0$ in Eq. (4.2)). As a result, a crystal with a $\chi^{(2)}$-type non-linearity is also birefringent. Normally crystals are only weakly non-linear, i.e. $\chi^{(1)} \gg \chi^{(2)}\mathbf{E} \gg \chi^{(3)}\mathbf{E}^2$.

The linear polarisability $\chi^{(1)}$ determines the index of refraction and describes the linear reaction of crystals, like refraction and dispersion; the second-order non-linear susceptibility $\chi^{(2)}$ leads to non-linear effects like three-wave mixing processes and the third-order non-linear susceptibility $\chi^{(3)}$ describes the tripling of the light frequency and hyper-parametric scattering [47].

4.2 Second-order Non-linear Processes

We will now study the second term in Eq. (4.2) which is called the second-order non-linear response (or non-linear polarisation) of a dielectric material and can be written as

$$\mathbf{P}^{(2)} = \varepsilon_0 \chi^{(2)}\mathbf{E} \cdot \mathbf{E}, \tag{4.3}$$

where $\chi^{(2)}$ is the second-order non-linear susceptibility. Note that there are three waves present, two *pump* waves and a third *response* wave characterised by $\mathbf{P}^{(2)}$. For this reason, second-order non-linear optical processes are sometimes referred to as *three-wave mixing processes*. Four possible three-wave mixing processes that can be obtained from this non-linearity. The main reference for this explanation is [48].

For simplicity in this explanation we neglect the tensor nature of the polarisation and assume that the electric fields and polarisation are scalar quantities. Let us consider an electric field made up of two pump fields oscillating at angular frequencies ω_1 and ω_2 such that

$$E(\mathbf{r}, t) = \tfrac{1}{2}\Big(E_1 \exp(i(k_1 r - \omega_1 t)) + E_2 \exp(i(k_2 r - \omega_2 t)) + \text{c.c.}\Big), \qquad (4.4)$$

where E_1 and E_2 are electric field amplitudes and c.c. represents the complex conjugate, to be incident on a birefringent medium. Substituting into the non-linear polarisation gives

$$P^{(2)}(t) = \varepsilon_0 \chi^{(2)} E^2 \qquad (4.5)$$

$$= \tfrac{1}{2}\varepsilon_0 \chi^{(2)}\left[|E_1|^2 + |E_2|^2 \right] +$$

$$\tfrac{1}{4}\varepsilon_0 \chi^{(2)}\left[E_1^2 \exp(i(2k_1 r - 2\omega_1 t)) + E_2^2 \exp(i(2k_2 r - 2\omega_2 t)) + \right.$$

$$2E_1 E_2 \exp(i((k_1 + k_2)r - (\omega_1 + \omega_2)t)) +$$

$$\left. 2E_1 E_2^* \exp(i((k_1 - k_2)r - (\omega_1 - \omega_2)t)) + \text{c.c.}\right]. \qquad (4.6)$$

The first two terms in Eq. (4.6) correspond to the non-linear process known as *optical rectification* which is the generation of a direct current when an intense optical field passes through a non-linear medium. The second two terms represents *second harmonic generation*. Each of the incident pump fields produce a second harmonic field that is twice the frequency, or half the wavelength, of the original field. This is only possible if the second harmonic field is generated in phase with the pump field, and this condition is known as phase-matching. The fifth term shows that the original pump beams can generate a field oscillating at a frequency that is the sum of the original two frequencies, hence this process in known as *sum-frequency generation*. The final term represents a process called *difference-frequency generation* which is the generation of a field with a frequency that is the difference of the original pump beams. Each of these processes can be cascaded. This means a second harmonic, for example, would be able to generate its own second harmonic, which would be the fourth harmonic of the original beam. However, only one of these processes will be possible when using the second-order medium and this is determined by the phase-matching conditions. The processes that are not properly phase-matched will be suppressed. The phase-matching condition will be explained in more detail in Section 4.3.

We are interested in reverse process to sum-frequency generation, which is also possible. This process, known as *spontaneous parametric down-conversion* (SPDC), occurs when a pump field oscillating at an angular frequency of ω_p generates two new optical fields such that $\omega_1 + \omega_2 = \omega_p$. On a single-photon level this process allows a pump photon which is incident on a non-linear anisotropic birefringent crystal to produce an entangled photon pair.

4.3 Fundamentals of SPDC

In SPDC an intense pump laser beam is incident on a non-linear birefringent crystal. Most of the photons from the pump beam continue straight through the crystal. However, occasionally the interaction between a single pump photon (with wave vector $\mathbf{k}_p$ and frequency ω_p) and a molecule of the crystal results in the pump photon being converted into two photons called the signal and idler with wave vectors $\mathbf{k}_s$, $\mathbf{k}_i$ and frequencies ω_s, ω_i respectively. The frequency of the pump is typically in the ultraviolet or violet range, while the down-converted fields have frequencies in the visible or near infrared region of the spectrum [37].

The conversion of a single photon is spontaneous since SPDC is stimulated by vacuum fluctuations. It is a parametric process because the incident energy is totally transferred to the generated photons and not to the molecule, i.e. the down-conversion medium is left unchanged by the process. It is a down-conversion process since each of the generated photons has a lower energy than the incident photon [49].

Under certain conditions (which will be discussed in the sections that follow) the state of the two down-converted photons will be entangled, meaning that it cannot be written as a product of two one-photon states and must be regarded as a single entity. The entanglement arises from the various constraints and conservation laws present in this non-linear process of SPDC. Since the pump beam is annihilated to produce the signal and idler photons, conservation of energy requires that the frequencies of the down-converted photons add up to the frequency of the pump

$$\omega_p = \omega_s + \omega_i. \tag{4.7}$$

This is also known as the frequency-matching condition. Conservation of momentum requires that

$$\mathbf{k}_p = \mathbf{k}_s + \mathbf{k}_i. \tag{4.8}$$

This is referred to as the phase-matching condition. By calculating the effective Hamiltonian for the SPDC process, which will be done in Section 4.4, we can verify the validity of the frequency- and phase-matching conditions.

4.3.1 Selecting the correct crystal for SPDC

The frequency- and phase-matching conditions that we have just mentioned allow us to select the correct crystal to induce SPDC. We can see why this is so if we follow the argument given in [50]. For this purpose we first note that the wave vectors in Eq. (4.8) can be expressed as

$$\mathbf{k}_j = \frac{n_j(\omega_j)\omega_j}{c}\hat{\mathbf{s}}_j, \tag{4.9}$$

where $n_j(\omega_j)$ (for $j = p, s, i$) is the dispersive refractive index of a material that depends on ω_j, $\hat{\mathbf{s}}_j$ is the unit vector pointing in the direction of $\mathbf{k}_j$ and c is the speed of light. Using this relationship we can rewrite Eq. (4.8) as

$$n_p(\omega_p)\omega_p\hat{\mathbf{s}}_p = n_s(\omega_s)\omega_s\hat{\mathbf{s}}_s + n_i(\omega_i)\omega_i\hat{\mathbf{s}}_i. \tag{4.10}$$

For degenerate down-conversion we have $\omega_s = \omega_i = \omega_p/2$ which is a choice of frequencies that satisfies the frequency-matching condition. In this case the minimal magnitude of $|\mathbf{k}_s| + |\mathbf{k}_i|$ that can still satisfy the phase-matching condition occurs when the down-converted fields and the pump beam are all collinear (so that the unit vectors $\hat{\mathbf{s}}_j$ point in the same direction). Eq. (4.10) thus becomes

$$n_p(\omega_p)\omega_p = n_s\left(\frac{\omega_p}{2}\right)\frac{\omega_p}{2} + n_i\left(\frac{\omega_p}{2}\right)\frac{\omega_p}{2}. \tag{4.11}$$

We can see that in order for this equation to be satisfied the refractive indices for the pump and target modes must be the same, i.e. $n_p = n_s = n_i$. This condition is true in an isotropic crystal where we have $n_p = n_s = n_i = n$, and in this instance Eq. (4.11) simplifies to

$$n(\omega_p) = n\left(\frac{\omega_p}{2}\right), \tag{4.12}$$

However, since the refractive index n for most dielectric materials decreases with increasing frequency, both frequency- and phase-matching conditions cannot be simultaneously satisfied in an isotropic medium [50].

With the help of an anisotropic birefringent crystal, where photons with different polarisations have different refractive indices, this task can be easily solved. Such a crystal allows the pump beam and at least one of the down-converted light beams to have opposite polarisations because it has two different refractive indices n_o and n_e for the ordinarily (*o*) and extraordinarily (*e*) polarised light, respectively. The simplest birefringent crystal is a uniaxial crystal which is one that has a single axis of symmetry called the optic axis (the z-axis, say). As a result, it responds differently to light polarised parallel to the optic axis (the extraordinary ray, or *e-ray*) and light polarised perpendicular to the optic axis (the ordinary ray, or *o-ray*). If $n_e < n_o$ then the crystal is a *negative* uniaxial crystal. In this case the e-ray experiences a refractive index that depends on the direction of propagation while the o-ray undergoes a constant larger refractive index. It is then possible to mix the waves in the anisotropic material so that the phase-matching condition is satisfied. For example, the extraordinary refractive index n_e for a pump beam with a wavelength of 405 nm is equal the ordinary refractive index n_o for the degenerate down-converted beam at a wavelength of 810 nm, i.e.

$$n_e(\omega_p) = n_o\left(\frac{\omega_p}{2}\right). \tag{4.13}$$

This means that if a pump beam that is polarised parallel to the optic axis (i.e. extraordinarily polarised), parametric down-conversion into a mode co-propagating with the pump beam but with orthogonal polarisation to it would be allowed. From this discussion we can see that an ideal crystal for inducing SPDC could be a negative uniaxial, anisotropic birefringent crystal. Therefore, for our proof-of-principle we use a β-Barium-Borate (BBO) crystal which is an example of such a crystal.

4.3.2 Types of phase-matching

In a uniaxial crystal there are two types of phase-matching that are commonly used, known as type-I and type-II phase-matching. In type-I phase-matching the two coaxial down-converted

beams have the same polarisation and this polarisation is orthogonal to that of the pump beam. For example, an ordinarily (*o*) polarised pump photon can create two extraordinarily (*e*) polarised photons. So type-I can be summarised as follows:

$$e \rightarrow oo \quad \text{or} \quad o \rightarrow ee \qquad \text{(Type-I)}. \qquad (4.14)$$

The down-converted photons emerge from the crystal along different directions and the possible directions of these conjugate photons form a set of concentric cones (Fig. 4.1 (a)). The opening angles of the cones depend on the frequency of the light, so different cones have different colours and the opening angles decrease with an increase in frequency [51]. Changing the angle of the pump beam with respect to the crystal allows us to change the size of the emission cone [52]. The phase-matching condition implies that the conjugate photons must emerge on opposite sides of the concentric cones (Fig. 4.1 (b)). We can see that there are infinite number of ways of choosing the signal and idler photons. By selecting a specific frequency for one of the photons we immediately constrain the frequency of the conjugate photon, so only two of all possible cones will be considered [49]. In the degenerate case the two down-converted cones overlap and collapse into one and the signal and idler photons can be found on opposite sides of the same beam.
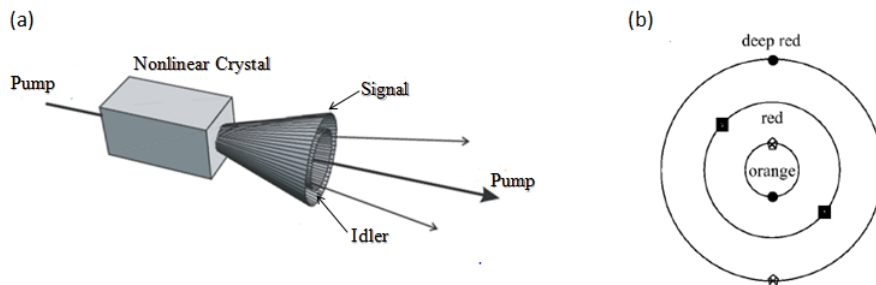
Figure 4.1: (a) A pump beam strikes a non-linear crystal with type-I phase-matching which emits coaxial SPDC cones. (b) A cross section of the emitted light cones. Like symbols (which lie on opposite sides of the cones) represent conjugate photons that satisfy the phase-matching condition. Note that the squares on the middle circle represent the degenerate case where the photons have exactly half the energy of the pump [53].

In type-II phase-matching the two down-converted beams have orthogonal polarisations. Because of birefringence effects the generated photons are emitted along two cones, one for the ordinary (*o*) wave and another for the extraordinary (*e*) wave as indicated in Fig. 4.2. So type-II can be summarised as:

$$e \rightarrow eo \quad \text{or} \quad o \rightarrow eo \qquad \text{(Type-II)}. \qquad (4.15)$$

As a result of the birefringence of our nonlinear crystal, tilting it allows us to increase the opening angle of the two cones so that they eventually intersect [54]. Fig. 4.2 shows this for the degenerate case (i.e. $\omega_s = \omega_i$). In the collinear case the two cones intersect at one line which is parallel to the pump beam. After more tilting of the crystal we obtain the non-collinear case where the two

cones intersect at two separate points. These intersection points are symmetric about the pump, thus a signal photon in one of the intersections will always have its corresponding idler in the other intersection and vice versa [54]. This means that the signal and the idler are indistinguishable in their spatial modes and thus we can easily obtain a polarisation entangled photon pair, which will be explained later.
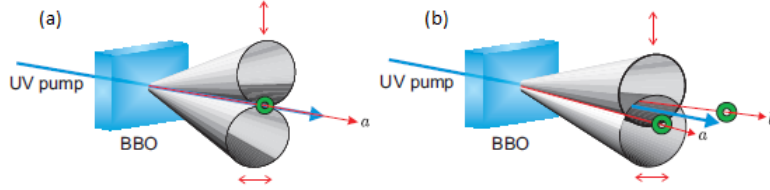
Figure 4.2: Degenerate Type-II SPDC. (a) In the collinear case the two emission cones intersect at a single point so photons of orthogonal polarisation are emitted in pairs into one spatial mode. (b) In the non-collinear case entangled photon pairs can be obtained in the spatial modes at the points where the two cones intersect [54].

4.4 The effective Hamiltonian

The usual way to treat SPDC is to start with the Hamiltonian of the process. This is because the Hamiltonian allows us to describe the energy transfer between different modes (i.e. from the pump beam to the signal and idler beams) due to the non-linear susceptibility of the crystal used. In order to describe this Hamiltonian we first calculate the energy density of the classical electric field, which includes the effects induced by the non-linear susceptibility, and then quantise the electric field, as we will now show. The interaction Hamiltonian can be used to calculate the state that is generated by the SPDC process. The calculations provided here are along the lines of the ones given in [37] and [49].

The non-linear crystal which is illuminated by the pump beam in SPDC is a dielectric medium of volume V. The electromagnetic energy density in the crystal is thus given by

$$\mathcal{H}(t) = \frac{1}{2} \int_V d^3\mathbf{r} \Big[\mathbf{D}(\mathbf{r},t) \cdot \mathbf{E}(\mathbf{r},t) + \mathbf{B}(\mathbf{r},t) \cdot \mathbf{H}(\mathbf{r},t) \Big], \qquad (4.16)$$

where $\mathbf{D}$ is the displacement vector, $\mathbf{B}$ is the magnetic induction and $\mathbf{H}$ is the magnetic field. The definition $\mathbf{D} = \varepsilon_0 \mathbf{E}(\mathbf{r},t) + \mathbf{P}(\mathbf{r},t)$ can then be used along with Eq. (4.2) to rewrite the Hamiltonian

$$\mathcal{H}(t) = \mathcal{H}_0(t) + \mathcal{H}_I^{(2)}(t), \qquad (4.17)$$

where $\mathcal{H}_0(t)$ contains the usual energy density of the electric field with decoupled harmonic oscillator modes while the *perturbation* $\mathcal{H}_I^{(2)}(t)$ contains the effect of the second-order non-linear

susceptibility of the material given by

$$\mathcal{H}_I^{(2)}(t) \equiv \mathcal{H}_I(t) = \frac{\varepsilon_0}{2} \int_V d^3\mathbf{r} \, \mathbf{E}(\mathbf{r}, t) \cdot \mathbf{P}^{(2)}(\mathbf{r}, t)$$

$$= \frac{\varepsilon_0}{2} \int_V d^3\mathbf{r} \, \chi_{jkm}^{(2)} E_{pj} E_{sk} E_{im}, \tag{4.18}$$

where $\mathbf{P}^{(2)}$ is the second-order non-linear component of the electric polarisation, $\chi_{jkm}^{(2)}$ is the second-order non-linear susceptibility tensor, V is the interaction volume covered by the pump laser beam, and repeated indices imply a sum. Each of the indices j, k and m correspond to the Cartesian coordinate axes x, y and z. The subscripts p, s and i refer to the pump, signal and idler beams, respectively. Note that if higher-order susceptibilities have to be considered, the interaction Hamiltonian density also takes the form $\mathbf{E} \cdot \mathbf{P}$.

From this point on we concern ourselves only with the non-linear portion of the Hamiltonian because this is the part that will allow us to calculate the result of the SPDC process. When the field is quantised, $E(\mathbf{r}, t)$ becomes a Hilbert space operator $\hat{\mathbf{E}}(\mathbf{r}, t)$ which can be decomposed as

$$\hat{\mathbf{E}}(\mathbf{r}, t) = \hat{\mathbf{E}}^{(+)}(\mathbf{r}, t) + \hat{\mathbf{E}}^{(-)}(\mathbf{r}, t), \tag{4.19}$$

where $\hat{\mathbf{E}}^{(+)}(\mathbf{r}, t)$ and $\hat{\mathbf{E}}^{(-)}(\mathbf{r}, t)$ are the positive- and negative-frequency parts of the electric field operator, respectively. These parts can be expressed in the plane-wave basis as (cf. Eq. (3.26))

$$\hat{\mathbf{E}}^{(+)}(\mathbf{r}, t) = \int_V d^3\mathbf{k} \left(\frac{\hbar \omega_{\mathbf{k}}}{2\epsilon_0 V} \right)^{\frac{1}{2}} \hat{\mathbf{a}}_{\mathbf{k}} e^{i(\mathbf{k} \cdot \mathbf{r} - \omega_{\mathbf{k}} t)} = \left[\hat{\mathbf{E}}^{(-)}(\mathbf{r}, t) \right]^\dagger. \tag{4.20}$$

If we substitute this equation into the classical Hamiltonian and retain only those terms that are important for the case when the signal and idler modes are initially in the vacuum state we obtain a Hermitian operator given by [53]

$$\hat{\mathbf{H}}_I = \frac{\varepsilon_0}{2} \int_V d^3\mathbf{r} \, \chi^{(2)} \hat{\mathbf{E}}_p^{(+)}(\mathbf{r}, t) \hat{\mathbf{E}}_s^{(-)}(\mathbf{r}, t) \hat{\mathbf{E}}_i^{(-)}(\mathbf{r}, t) + \text{h.c.}, \tag{4.21}$$

where $\chi^{(2)}$ is a single component of the susceptibility tensor.

4.4.1 The two-photon state

Let us look at the state that is produced by this Hamiltonian. We assume that the non-linear interaction is turned on at time $t_0 = 0$ when the system is in the initial state

$$|\Psi(0)\rangle \equiv |\Psi_0\rangle = |1\rangle_p |0\rangle_s |0\rangle_i, \tag{4.22}$$

which is a state of the field when we consider a single pump photon and the signal and the idler photons in the vacuum state. The state at time t is then given by

$$|\Psi(t)\rangle = \hat{\mathbf{U}}(t)|\Psi_0\rangle, \tag{4.23}$$

where

$$\hat{\mathbf{U}}(t) = \exp\left(\frac{1}{i\hbar}\int_0^t d\tau \hat{\mathbf{H}}_I(\tau)\right), \tag{4.24}$$

is the time-evolution operator. If the pump field is sufficiently weak, such that the interaction time t is small compared to the average time between the down-conversions, then we can expand the operator in Eq. (4.24) in power series as

$$\hat{\mathbf{U}}(t)|\Psi_0\rangle = |\Psi_0\rangle + \left(\frac{1}{i\hbar}\int d\tau \hat{\mathbf{H}}_I(\tau)\right)|\Psi_0\rangle$$
$$- \left(\frac{1}{2\hbar^2}\int\int d\tau d\tau' \hat{\mathbf{H}}_I(\tau)\hat{\mathbf{H}}_I(\tau')\right)|\Psi_0\rangle + \dots \tag{4.25}$$

The first term in this equation represents the initial state, the second term is the two-photon state $|\Psi_2\rangle$, the third term is the four-photon state $|\Psi_4\rangle$, etc. Let us consider the two-photon state

$$|\Psi_2\rangle = \left(\frac{1}{i\hbar}\int d\tau \hat{\mathbf{H}}_I(\tau)\right)|1\rangle_p|0\rangle_s|0\rangle_i$$
$$= \frac{\varepsilon_0\chi^{(2)}}{2i\hbar}\int_0^t d\tau \int_V d^3\mathbf{r}\,\hat{\mathbf{E}}_p^{(+)}(\mathbf{r},t)\hat{\mathbf{E}}_s^{(-)}(\mathbf{r},t)\hat{\mathbf{E}}_i^{(-)}(\mathbf{r},t)|1\rangle_p|0\rangle_s|0\rangle_i, \tag{4.26}$$

where we have assumed that $\chi^{(2)}$ is constant and the Hermitian conjugate portion of the Hamiltonian disappears because it contains annihilation operators acting on the vacuum. To get an idea of the action of this Hamiltonian on the initial state, we can consider the Hamiltonian in terms of the creation and annihilation operators only (for a specific wave vector $\mathbf{k}$)

$$|\Psi_2\rangle \propto \hat{\mathbf{a}}_p\hat{\mathbf{a}}_s^\dagger\hat{\mathbf{a}}_i^\dagger|1\rangle_p|0\rangle_s|0\rangle_i = |0\rangle_p|1\rangle_s|1\rangle_i. \tag{4.27}$$

This means that, in the simplest case, with the signal and idler beams initially in vacuum states, a single pump beam photon is converted into two optical photons: one in the signal beam and the other in the idler.

We want to find the conditions when this conversion will occur, so it is necessary to calculate the action of the Hamiltonian more explicitly. We will now use [55] as the main reference for this part of the discussion. Since the SPDC process depends on the second-order non-linear susceptibilities it has an efficiency in the range of 10^{-7} to 10^{-11}, which means there is a very low probability of a pump photon being down-converted. In order to obtain a substantial amount of photons being emitted into the signal and idler beams it is necessary to pump the medium with a very strong coherent field which can be modelled as a classical field. This is acceptable if we are interested in interactions over a short enough time such that the depletion of pump photons can be ignored [53]. This is known as the parametric approximation. Therefore, the electric field operator for the pump beam $\hat{\mathbf{E}}_p(\mathbf{r},t)$ may be replaced by the classical field in the form of a plane wave

$$\hat{\mathbf{E}}_p^{(+)}(\mathbf{r},t) = E_p e^{i(\mathbf{k}_p\cdot\mathbf{r}-\omega_p t)}, \tag{4.28}$$

where E_p is the classical amplitude of the pump beam. We can then reduce Eq. (4.26) as

$$|\Psi_2\rangle = \int d^3\mathbf{k}_s \int d^3\mathbf{k}_i \Phi\left(\mathbf{k}_s,\omega_s,\mathbf{k}_i,\omega_i\right)\hat{\mathbf{a}}_{k_s}^\dagger\hat{\mathbf{a}}_{k_i}^\dagger|0\rangle_s|0\rangle_i, \tag{4.29}$$

where Φ is known as the mode function and is defined as

$$\Phi(\mathbf{k}_s, \omega_s, \mathbf{k}_i, \omega_i) \propto \int_0^t d\tau \int_V d^3\mathbf{r} E_p e^{i(\mathbf{k}_p - \mathbf{k}_s - \mathbf{k}_i)\cdot\mathbf{r}} e^{i(\omega_s + \omega_i - \omega_p)t}. \tag{4.30}$$

The mode function contains all the information about the generated two-photon system in space and frequency, not only about their individual state but about their correlations as well. If we split the volume integration into its transverse and longitudinal parts the mode function becomes

$$\Phi = C_1 \int_0^t d\tau e^{i(\omega_s + \omega_i - \omega_p)t} \int_0^L dz e^{i(k_{p_z} - k_{s_z} - k_{i_z})z} \int_A d^2\mathbf{r}_\perp e^{-i(\mathbf{q}_s + \mathbf{q}_i - \mathbf{q}_p)\cdot r_\perp}, \tag{4.31}$$

where A and L are the area and length of the crystal, respectively, and all constant factors have been collected into the constant C_1. Note that we have also separated the wave vector into its longitudinal and transverse parts (i.e. $\mathbf{k} = k_z\hat{\mathbf{z}} + \mathbf{q}$) which is advantageous since we are considering paraxial beams. If we assume that the integrated volume over the transverse components, x and y, is infinite then we obtain a delta function for the corresponding wave vectors of the fields:

$$\int_A d^2\mathbf{r}_\perp e^{-i(\mathbf{q}_s + \mathbf{q}_i - \mathbf{q}_p)\cdot\mathbf{r}_\perp} = (2\pi)^2\delta(\mathbf{q}_s + \mathbf{q}_i - \mathbf{q}_p). \tag{4.32}$$

We could have a similar integral for the longitudinal part $2\pi\delta(k_{p_z} - k_{s_z} - k_{i_z})$ which leads to a delta function for the wave vectors,

$$(2\pi)^3\delta(\mathbf{k}_s + \mathbf{k}_i - \mathbf{k}_p). \tag{4.33}$$

However, if finite length of the crystal has to be taken into account, the delta function in the z-direction has to be replaced by a sinc function:

$$\int_0^z dz e^{i\Delta k_z} = \frac{2}{\Delta k_z L}\sin\left(\frac{\Delta k_z L}{2}\right) = \text{sinc}\left(\frac{\Delta k_z L}{2\pi}\right), \tag{4.34}$$

where L is the length of the crystal and $\Delta k_z = k_{p_z} - k_{s_z} - k_{i_z}$. The interaction time t is long enough that the integration over time can be approximated as a delta function as well

$$2\pi\delta(\omega_s + \omega_i - \omega_p). \tag{4.35}$$

Thus for a reasonably steady state assumption the two-photon state is

$$|\Psi_2\rangle = A_0 \int d^3\mathbf{k}_s \int d^3\mathbf{k}_i \delta(\omega_s + \omega_i - \omega_p)\delta(\mathbf{q}_s + \mathbf{q}_i - \mathbf{q}_p)\text{sinc}\left(\frac{\Delta k_z L}{2\pi}\right)$$
$$\times \hat{\mathbf{a}}_{k_s}^\dagger \hat{\mathbf{a}}_{k_i}^\dagger |0\rangle_s|0\rangle_i, \tag{4.36}$$

where A_0 is a new constant. We can observe that this result shows that the signal and idler photons are only produced in pairs. Also, the delta and sinc functions mean that we can only have pairs if the frequency- and phase-matching conditions are met. The frequency-matching condition implies that the energies of the two created photons equal the energy of one pump photon, therefore it is also called the energy conservation condition. The phase-matching condition is related to the fact that the pump beam must stay in phase with the signal and idler beams in order for the signal power to interfere constructively throughout the crystal [56].

4.5 Polarisation Entanglement

Since we have two different types of phase-matching which allow the signal and idler to be generated with either identical polarisation (Type-I) or orthogonal polarisations (Type-II), we have two corresponding types of polarisation-entangled states that can be formed from SPDC. This discussion will be limited to degenerate down-conversion of a monochromatic pump beam so $\omega_s = \omega_i = \omega_p/2$.

4.5.1 Type-II

We will consider non-collinear type-II SPDC first because it is the simpler of the two cases. The first *direct* source of polarisation-entangled photons (which did not require post-selection of a subspace of the produced photon states) was suggested and experimentally demonstrated by Kwiat *et al.* in 1995 [57]. As we have mentioned before, the down-converted photons are emitted into two cones — one ordinarily and one extraordinarily polarised. At the intersection of the cones we will obtain polarisation-entangled states. We use the notation $|V\rangle$ and $|H\rangle$ to represent vertically (ordinarily) and horizontally (extraordinarily) polarised single-photon states. In front of the intersection points where the cones overlap, two screens with pinholes labelled a and b are used to select the signal-idler pair according to the directions of their wave vectors, with all other pairs being neglected. For the selected conjugate pair of photons there will be an ambiguity as to whether the signal or idler photons are vertically or horizontally polarised, as indicated in Fig. 4.3.

Eq. (4.36) gives the general two-photon state that is produced in the SPDC process. For this case, however, the two-photon state can be written as

$$|\Psi_2\rangle = A_0 \left(\hat{\mathbf{a}}^\dagger_{Vs} \hat{\mathbf{a}}^\dagger_{Hi} + \hat{\mathbf{a}}^\dagger_{Hs} \hat{\mathbf{a}}^\dagger_{Vi} \right) |0\rangle_s |0\rangle_i, \tag{4.37}$$

where the operators $\hat{\mathbf{a}}^\dagger_{Vs}$, $\hat{\mathbf{a}}^\dagger_{Hs}$, $\hat{\mathbf{a}}^\dagger_{Vi}$ and $\hat{\mathbf{a}}^\dagger_{Hi}$ are the creation operators for photons with vertical and horizontal polarisation for the signal and idler beams, respectively. The sum appears because the down-converted photons are orthogonally polarised and we cannot distinguish whether the signal photon that is vertically polarised and the idler horizontally polarised or vice versa. We have assumed perfect phase-matching and degenerate down-conversion of a monochromatic pump beam so the delta and sinc functions equal unity. Also, we have dropped the integration over all possible wave vectors because the pinholes at the points where the signal and idler beams overlap post-select a specific value for each of the wave vectors. The normalised two-photon state that is produced is

$$|\Psi_2\rangle = \tfrac{1}{\sqrt{2}} \left(|H\rangle_a |V\rangle_b + e^{i\alpha} |V\rangle_a |H\rangle_b \right), \tag{4.38}$$

where the relative phase α is determined by the phase-matching conditions and depends on various factors such as the type and length of the crystal, and the frequencies and wave vectors of the pump and down-converted photons. This phase can be set by using an additional birefringent phase shifter in one of the optical parts after the crystal, or even by slightly rotating the crystal.
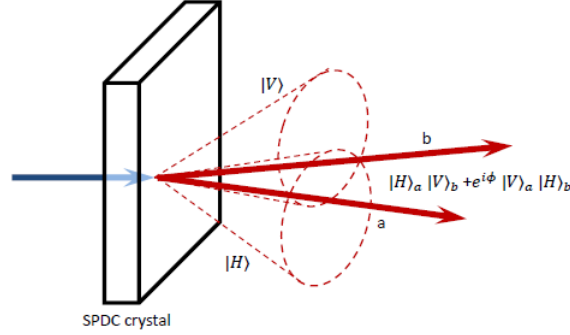
Figure 4.3: Type-II SPDC using a single non-linear crystal. The signal and idler photons are each emitted along two cones. At the intersection of the two cones (a and b) one can collect the polarisation-entangled photon pairs [56].

The advantage of this set-up is that the entangled photons are produced in two different spatial modes, so that they can be easily collected into single-mode fibres for measurement purposes. The disadvantage, however, is that we lose the majority of photons which are produced in SPDC because we use only those photons from the intersection of the down-conversion cones. Also note that because of the birefringence of the non-linear crystal we have different refractive indices for the signal and idler photons which causes a separation of the signal and idler beams (transversal walk-off) and different group velocities of idler and signal photons (longitudinal walk-off) [52]. Hence we would be able to distinguish between the two photons. This walk-off can be compensated for by placing a non-linear crystal, half the length of the down-conversion crystal but identical otherwise, in each of the output modes [56].

4.5.2 Type-I

In type-I SPDC we cannot obtain a polarisation-entangled photon pair with a single non-linear crystal. Consider a crystal that is cut for type-I SPDC, with its optic axis aligned in the vertical plane to be pumped with a vertically polarised beam. The two-photon state that is produced is

$$|\Psi_2\rangle = A_0 \left(\hat{\mathbf{a}}^\dagger_{Hs} \hat{\mathbf{a}}^\dagger_{Hi} \right) |0\rangle_s |0\rangle_i = A_0 |H\rangle_s |H\rangle_s. \tag{4.39}$$

Alternatively if the optic axis is aligned in the horizontal plane and the crystal is pumped by a horizontally polarised beam then we have

$$|\Psi_2\rangle = A_0 |V\rangle_s |V\rangle_i, \tag{4.40}$$

where we have made the same assumptions as in the type-II case. So we see that it is not possible to produce polarisation-entanglement in this case.

In 1999, however, Kwiat *et al.* [58] demonstrated polarisation entanglement using type-I SPDC with two non-linear crystals oriented in perpendicular planes to each other (see Fig. 4.4) and pumped by a diagonally polarised beam (i.e. with a polarisation of $|H\rangle_p + |V\rangle_p$). In this way, entangled photon pairs can be produced either in the first crystal with its optic axis oriented along

the horizontal plane ($|V\rangle_p \to |H\rangle_s |H\rangle_i$) or in the second crystal oriented along the vertical plane ($|H\rangle_p \to |V\rangle_s |V\rangle_i$). As long as there is no information left which would allow one to infer from which crystal a photon originated, the resulting two-photon state is entangled

$$|\Psi_2\rangle = \tfrac{1}{\sqrt{2}}(|H\rangle_a |H\rangle_b + e^{i\alpha}|V\rangle_a |V\rangle_b). \tag{4.41}$$

The phase α can be adjusted by tilting the down-conversion crystals themselves; by introducing a birefringent phase shift to one of the output beams; or by controlling the relative phase between the horizontal and vertical components of the pump beam [56]. The entangled state in Eq. (4.41) could also be produced using type-II SPDC by inserting a half-wave plate (which carries out the operation: $|H\rangle \to |V\rangle$ and vice versa) into one spatial mode after the crystal.
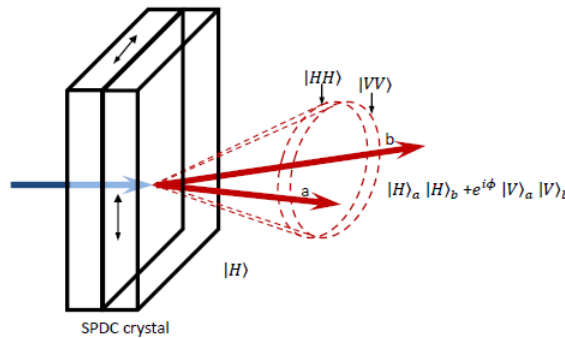
Figure 4.4: Type-I SPDC set-up. Two type-I non-linear crystals are "sandwiched" together and rotated by $90°$ with respect to each other. The crystal are pumped by a diagonally polarised beam so that the first crystal produces $|HH\rangle$ pairs and the second generates $|VV\rangle$ pairs. If there is no way of distinguishing from which crystal the photons come, then the produced state is entangled in polarisation [56].

The advantage of this method is that a very large proportion of the produced photons are entangled. However, in order to have indistinguishability in this configuration, the crystals need to be thin, thus limiting the brightness of this kind of source [56]. Similarly to the type-II case, we must compensate for the transverse and longitudinal walk-off. If a pulsed pump is used it is also necessary to compensate for the longitudinal walk-off between the horizontal and vertical components of the pump beam [37].

4.5.3 Double Pair emission

We see from Eq. (4.25) there exists the possibility of obtaining a four-photon event. The four-photon state given by

$$|\Psi_4\rangle = \frac{1}{2}\left(\frac{H_I t}{\hbar}\right)^2 |0\rangle_s |0\rangle_i. \tag{4.42}$$

In type-II down-conversion and in terms of creation and annihilation operators this leads to

$$|\Psi_4\rangle = \frac{\eta^2}{2} \left(\hat{\mathbf{a}}_{Vs}^\dagger \hat{\mathbf{a}}_{Hi}^\dagger + \hat{\mathbf{a}}_{Hs}^\dagger \hat{\mathbf{a}}_{Vi}^\dagger \right)^2 |0\rangle_s |0\rangle_i \tag{4.43}$$

$$= \frac{\eta^2}{2} \left(\hat{\mathbf{a}}_{Vs}^{\dagger 2} \hat{\mathbf{a}}_{Hi}^{\dagger 2} + \hat{\mathbf{a}}_{Hs}^{\dagger 2} \hat{\mathbf{a}}_{Vi}^{\dagger 2} + 2\hat{\mathbf{a}}_{Vs}^\dagger \hat{\mathbf{a}}_{Hi}^\dagger \hat{\mathbf{a}}_{Hs}^\dagger \hat{\mathbf{a}}_{Vi}^\dagger \right) |0\rangle_s |0\rangle_i, \tag{4.44}$$

where $\eta = A_0 t / \hbar$. The final four-photon entangled state is therefore given by

$$|\Psi_4\rangle = |2H\rangle_a |2V\rangle_b + |2V\rangle_a |2H\rangle_b + |1H, 1V\rangle_a |1H, 1V\rangle_b. \tag{4.45}$$

where, for example, $2H_a$ denotes two horizontally polarised photons in beam a while $2V_b$ refers to two vertically polarised photons in beam b. Note that the probability of obtaining a two-photon event is $p_2 \simeq A_0^2$ while the probability of obtaining a four-photon event is $p_2 \simeq A_0^4$.

4.6 Orbital angular momentum entanglement

In 2001, Mair *et al.* [11] showed that OAM is conserved in individual down-conversion events at the single-photon level and that the states of the signal and idler could be entangled in OAM. In order to measure the OAM correlations, and thus prove the first half of the previous statement, they performed coincidence measurements of the OAM of the down-converted photons. They found that a pump laser beam with $l_p = -1, 0, 1$ only produced photons with $l_p = l_s + l_i$, which means that the OAM was conserved. However, this confirms only a classical correlation between the OAM of the down-converted photons. To prove that the photons are indeed entangled in OAM, Mair *et al.* first projected one of the down-converted photons into a superposition of OAM states. They then measured the state of the second photon and found that it was not a mixture as it would be if there was only a classical correlation between the photons; rather, the state of the second photon was in a coherent superposition of OAM states as well. This means that the two photons were initially entangled in their OAM degree of freedom.

4.6.1 The OAM entangled two-photon state

We would now like to describe the two-photon state that is entangled in the OAM. For this purpose we shall follow the method used in [59]. We consider a thin BBO crystal cut for type-I SPDC to be illuminated by a pump beam propagating in the z-direction and assume perfect phase-matching. We use a single type-I crystal because we are interested in generating only OAM entangled photons and do not need the polarisation entanglement. The pump, signal and idler beams are paraxial which means that the wave vector $\mathbf{k}$ is almost parallel to its direction of propagation with a minimal spread in the transverse direction. We can, therefore, consider the longitudinal (k_z) and transverse components ($\mathbf{q}$) of the wave vector in this equation separately,

$$\mathbf{k} = k_z \hat{\mathbf{z}} + \mathbf{q}. \tag{4.46}$$

Eq. (4.29) can then be written as

$$|\Psi_2\rangle = \int d\omega_s d\omega_i d^2\mathbf{q}_s d^2\mathbf{q}_i \Phi(\mathbf{q}_s, \omega_s, \mathbf{q}_i, \omega_i) \hat{\mathbf{a}}_s^\dagger(\mathbf{q}_s, \omega_s) \hat{\mathbf{a}}_i^\dagger(\mathbf{q}_i, \omega_i) |0\rangle_s |0\rangle_i. \tag{4.47}$$

We would like to consider degenerate down-conversion of a monochromatic pump beam which means we select a specific value for the signal and idler beams ($\omega_s = \omega_i = \omega_p/2$). Eq. (4.47) therefore reduces to

$$|\Psi_2\rangle = \int d^2\mathbf{q}_s d^2\mathbf{q}_i \Phi(\mathbf{q}_s, \mathbf{q}_i)\hat{\mathbf{a}}_s^\dagger(\mathbf{q}_s)\hat{\mathbf{a}}_i^\dagger(\mathbf{q}_i)|0\rangle_s|0\rangle_i. \tag{4.48}$$

Normalisation of the state requires that $\int d^2\mathbf{q}_s d^2\mathbf{q}_i |\Phi(\mathbf{q}_s, \mathbf{q}_i)|^2 = 1$. Since a photon state in an LG mode can be written as

$$|pl\rangle = \int d^2\mathbf{q}\, \mathcal{LG}_{pl}(\mathbf{q})\hat{\mathbf{a}}^\dagger(\mathbf{q})|0\rangle, \tag{4.49}$$

where $\mathcal{LG}_{pl}(\mathbf{q})$ is given in cylindrical coordinates by Eq. (3.34), we can use $I = \sum_{p,l}|pl\rangle\langle pl|$ to rewrite Eq. (4.48) in terms of LG modes which are the eigenstates of the OAM operator

$$|\Psi_2\rangle = \sum_{p_s, l_s}\sum_{p_i, l_i} C_{p_s, l_s}^{p_i, l_i}|p_s l_s\rangle|p_i l_i\rangle, \tag{4.50}$$

where

$$C_{p_s, l_s}^{p_i, l_i} = \int d^2\mathbf{q}_s d^2\mathbf{q}_i \Phi(\mathbf{q}_s, \mathbf{q}_i)[\mathcal{LG}_{p_s l_s}]^*[\mathcal{LG}_{p_i l_i}]^*. \tag{4.51}$$

The probability of finding the signal photon in the state $|p_s l_s\rangle$ and the idler in the state $|p_i l_i\rangle$ is given by $P_{p_s, l_s}^{p_i, l_i} = |C_{p_s, l_s}^{p_i, l_i}|^2$. The pump beam carries a well-defined OAM of l_p so the mode function Φ can be described by an LG transverse spatial distribution given by Eq. (2.45). Assuming that the crystal length L is much smaller than the Rayleigh range z_R of the pump beam, Eq. (4.51) can be written as [37, 60]

$$C_{p_s, l_s}^{p_i, l_i} = \int_0^{2\pi} d\phi \int_0^\infty r dr\, LG_{p_p l_p}(r, \phi)[LG_{p_s l_s}(r, \phi)]^*[LG_{p_i l_i}(r, \phi)]^*, \tag{4.52}$$

where we have used the Fourier transform (cf. Section 3.2.3). After substituting Eq. (2.45) into Eq. (4.52) we find that the integral over the azimuthal phase ϕ is

$$\int_0^\infty d\phi \exp[i(l_p - l_s - l_p)\phi] = 2\pi\delta_{l_p, l_s + l_i}. \tag{4.53}$$

This gives us the conservation of OAM law: $l_p = l_s + l_i$ (see Fig. 4.5) which results in entanglement between the signal and idler photons. We will restrict ourselves to states with $p_s = p_i = 0$ so that we can write

$$|l_s\rangle|l_i\rangle \equiv |l_s, p_s = 0\rangle|l_i, p_i = 0\rangle, \tag{4.54}$$

and our final two-photon OAM entangled state is

$$|\Psi_2\rangle = \sum_{l_s}\sum_{l_i} C_{l_s, l_i}|l_s\rangle|l_i\rangle. \tag{4.55}$$

This is a multi-dimensional OAM entangled two-photon state, which in general will also contain terms with the radial mode index $p \neq 0$. This state shows that neither photon possesses a well-defined OAM after SPDC but the measurement of one photon defines its OAM state and projects the other into the corresponding OAM state.
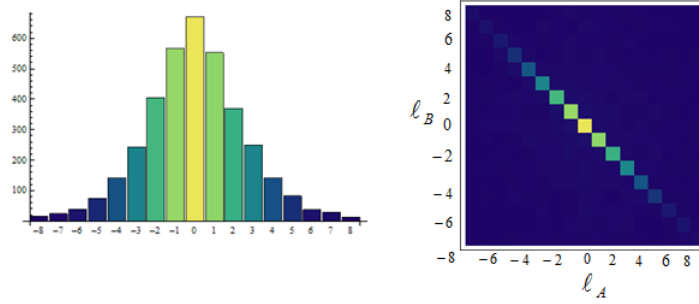
Figure 4.5: The conservation of OAM in SPDC for a pump beam of with $l_p = 0$, i.e. $l_A + l_B = 0$ [61].

Chapter 5

Optical toolkit for Orbital Angular Momentum

The OAM states of photons allow us access to an infinite-dimensional Hilbert space for encoding quantum information. Quantum information processing in this multi-dimensional system requires a toolkit of optical devices which enable us to generate single photons, and to manipulate and measure their quantum states. There has been significant research focused in this area since the discovery by Allen *et al.* [8] that photons in an LG beam have a quantised OAM. This has resulted in a set of optical elements that are analogous to those used to manipulate the spin angular momentum (or polarisation) states of photons. In this chapter we will give a short description of the components that allow us to alter the OAM states of photons. The selection of components discussed here are only those which are necessary for our theoretical proof-of-principle experiment. Each component will be described by their action on the classical and quantum states of light in order to gain a complete understanding of their diverse applications.

Section 5.1 gives a short introduction to conventional and digital holography wherein we describe how a computer-generated hologram together with a spatial light modulator can be used to produce the desired light beams, which in our case are the LG modes. We then classify the different types of holograms that are available to us and explain our choices of holograms for the proof-of-principle experiment. Manipulation and detection of LG beams and the OAM states of photons is also possible with computer-generated holograms, as we will show. Astigmatic mode-converters are introduced in Section 5.2, where we give a detailed explanation of the process of mode conversion between LG and HG modes using a pair of cylindrical lenses. Section 5.3 demonstrates how Dove prisms can be used to manipulate the OAM states of photons while Section 5.4 shows that we can combine mode converters and Dove prisms in order to obtain a rotation matrix for first-order LG modes. Finally, two interferometric methods for sorting the OAM states of single photons are described in Section 5.5.

5.1 Digital holography and Spatial Light Modulators

Holography is a technique that makes it possible to reproduce the image of an object that is no longer present. It involves recording both the phase and amplitude of the light waves that are scattered from an object. Since all recording materials respond only to the changes of intensity in the image, it is necessary to convert the phase information into variations in intensity [62]. In the conventional method of holography this could be done by first illuminating the object one wished to reproduce with a laser beam. The interference pattern between the light scattered from the object (called the object beam) and a reference beam was then recorded on photographic film, and this recording was known as the hologram. The intensity at any point in this interference pattern also depends on the phase of the object beam so the hologram will contain information on both the phase and amplitude of the object beam, as desired. A three-dimensional optical image of the original object could then be created by illuminating the hologram with the reference beam under the same angle of incidence as was used when recording it. The result is actually two images — a real image (often called the conjugate image) and a virtual image (often called the primary image, which is a twin of the real image with the same amplitude but opposite phase). Part of the incident beam also passes through the hologram unmodulated. Depending on the technique used to develop the hologram (i.e. on-axis or off-axis holography), the resulting beams can either be superimposed or separated, as we will see later. My main reference sources for this section will be [62] (which gives a simple but valuable introduction to holography) and [33] (which has a section explaining how information can be encoded into the transverse spatial modes of light).

5.1.1 A Computer-Generated Hologram

Unlike conventional holography, digital holography allows us to create the image of an object that does not physically exist. All we need is the mathematical description of the interference pattern between the object we wish to produce and a reference beam. This interference pattern can be generated by a computer and recorded onto a suitable device which will act as the hologram, hence it is called a *computer-generated hologram*. The recording device we will use is a spatial light modulator (SLM).

An SLM uses a liquid crystal display (LCD) to modulate the phase (or amplitude) of an incident beam and create a variety of beam shapes. A grey-scale computer-generated hologram is programmed onto the SLM which can be addressed by a circuit board via a cable (see Fig. 5.1). The SLM uses the hologram to shape the correct output beam (or image) by altering the phase of the incident beam according to the shade of grey present at each pixel of the hologram. SLMs are commonly used for digital holography because they are now commercially available and they can be dynamically updated without any need to realign the system (i.e. they can be reconfigured easily and efficiently [63]) which makes them very useful for optical configurations compared to fixed holograms.

Although an SLM is very versatile device, there are efficiency problems associated with the structure of the LCD of the SLM. The efficiency problems are, specifically, reduced diffraction and mode-conversion efficiencies. Diffraction efficiency is the fraction of the total power that is trans-

mitted into the desired diffraction order while mode-conversion efficiency describes the proportion of the diffracted power which is converted into the desired target mode.
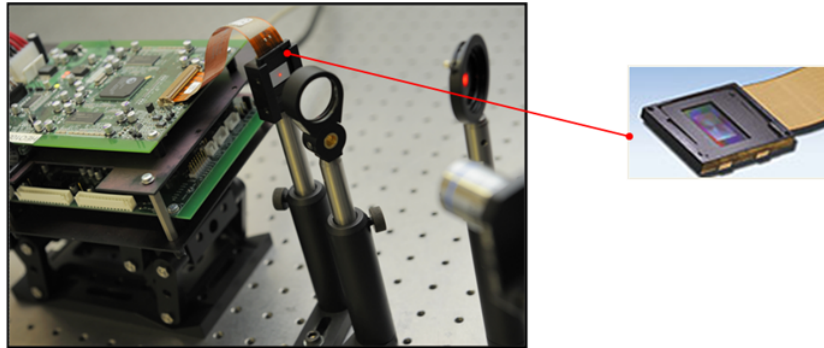
Figure 5.1: An SLM consists of a liquid crystal display which is programmed and addressed by a circuit board via a cable [23].

The first problem is that not all of the incident light is diffracted to produce the desired beam, which results in a reduced diffraction efficiency. Inefficient diffraction occurs for two reasons. Firstly, the LCD of the SLM consists of a grid of pixels and the spacing between the pixels causes the LCD to act as a two-dimensional grating. This results in light being diffracted into many orders. Also, when light is incident on a non-pixel (i.e. a space between pixels) the light is lost, and not reflected or transmitted. The fraction of the area of the LCD that is covered by pixels is referred to as the fill factor. A fill factor of 90% means that 90% of the of the area contains pixels while the other 10% is made up of non-pixels. A decrease in the fill factor decreases the diffraction efficiency.

The second problem is that not all of the light that is diffracted into the correct diffraction order is actually transformed into the target mode. Since the reflected light from the surface of the LCD and the diffracted light from the gaps between the pixels do not undergo phase modulation, the unmodulated incident light overlaps with the phase modulated light (which is our target mode) in the diffraction order of interest (see Fig. 5.5 (a)). This results in a reduced mode-conversion efficiency.

5.1.2 Types of holograms

On- or Off-axis Holograms

There are many different ways to classify a hologram. This category differentiates between holograms on the basis of the angle between the reference and the object beam that was used when constructing the interference pattern. For on-axis holography these two beams and the reconstructed beam are all collinear. This makes it easy to align the system to reconstruct the desired beam. As we have previously mentioned, the output of the hologram is the desired beam, the con-

jugate beam(s) (of equal amplitude and opposite phase as the desired beam) and the unmodulated incident beam. In the case of on-axis holography these beams are all superimposed so it can be difficult to separate out the desired beam.

For an off-axis hologram the reference and object beams are aligned at an angle with respect to each other. We know that the interference of a plane reference and object beam aligned in this manner results in a fringe pattern of light (constructive interference) and dark (destructive interference) regions with the intensity varying sinusoidally across the interference pattern (Fig. 5.2). The spacing between the fringes is determined by the angle between the two waves and by the wavelength of the light. When the hologram containing this fringe pattern is programmed onto the SLM it acts as a diffraction grating for when the object beam is reconstructed. So in off-axis holography we obtain a fringe pattern which is superimposed over the broad features of the interference pattern. The fringe pattern produces beams which are diffracted at different angles θ_n so the desired beam can be easily isolated from the conjugate beam(s). Note that by superimposing a grating on an on-axis hologram we can produce an off-axis hologram.

Reflection or Transmission Holograms

The next category is transmission or reflection. A transmission hologram is one where we consider the object and reference beam to be propagating in the same direction when forming the interference pattern while for a reflection hologram the two beams are incident from opposite sides. The reconstructed beam for the transmission (reflection) hologram is then in the transmitted (reflected) spatial mode [33]. To simplify the explanations that follow we will consider transmission holograms.

Thin or Thick Holograms

One of the most important categories used to differentiate between holograms is based on the dimensions in which the information in a hologram is stored. A thick hologram stores information in three dimensions (i.e. it has a z component) so the interference pattern varies with depth as well as in the transverse direction. It is for this reason that this hologram is also referred to as a *volume hologram*. The analysis of the field emitted after an incident beam has traversed the hologram is very complicated so, for simplicity, we will be using a thin hologram. A thin hologram stores the information in the transverse (xy) plane. This hologram is characterised by a transmission function, which is a complex function describing the interference pattern. Therefore, the hologram modifies the incident beam at each point in its transverse profile by means of a point-wise product with the transmission function

$$u_{\text{out}}(\boldsymbol{\rho}, z) = T(\boldsymbol{\rho}, z) u_{\text{in}}(\boldsymbol{\rho}, z), \tag{5.1}$$

where $\boldsymbol{\rho}$ is a two-dimensional vector in the transverse plane. Note that the transmission function is also referred to as a transfer function or as the transmittance of the hologram. The spatially varying complex amplitude transmission function is given by

$$T(\boldsymbol{\rho}) = |T(\boldsymbol{\rho})| \exp[i\Phi(\boldsymbol{\rho})]. \tag{5.2}$$

Amplitude or Phase Holograms

The interference pattern of a hologram can be stored in the transmission function as a phase or amplitude modulation and the hologram is named accordingly. We will now look for expressions for the transmission functions that characterise each of these holograms.

A thin amplitude hologram does not impose a phase shift on the incident beam ($\Phi(\boldsymbol{\rho}) = 0$) so the associated transmission function depends only on the amplitude $|T(\boldsymbol{\rho})|$ which varies across the hologram. The transmission function is, therefore, described in terms of its amplitude transmittance in the transverse plane which is the change in the amplitude of the incident field just before ($|u_{\text{in}}|$) and just after ($|u_{\text{out}}|$) traversing the hologram,

$$T(\boldsymbol{\rho}) = \frac{|u_{\text{out}}(\boldsymbol{\rho})|}{|u_{\text{in}}(\boldsymbol{\rho})|}. \tag{5.3}$$

We can see from this equation that the range of values for T is $0 \leq T \leq 1$. The amplitude hologram modulates the incident beam according to the intensity I of the interference pattern so we assume a linear relationship between the transmission function and the intensity

$$T_{\text{amp}}(\boldsymbol{\rho}) = 1 - \beta I(\boldsymbol{\rho}) = t_0 + \beta \cos[\alpha], \tag{5.4}$$

where t_0 is the average amplitude transmittance, β is the amplitude of the spatial variation of T and α is the grating pattern which results when the interference pattern between the object and reference beams is recorded onto the hologram. If both of these waves are plane waves then $\alpha = 2\pi x/\Lambda$ where Λ is the spacing between the fringes. The amplitude of the desired beam, which is diffracted into the first order, is at a maximum for $t_0 = \beta = 1/2$ so we have

$$T_{\text{amp}}(\boldsymbol{\rho}) = \tfrac{1}{2}(1 + \cos(\alpha)) \tag{5.5}$$

$$= \tfrac{1}{2} + \tfrac{1}{4}\exp[\alpha] + \tfrac{1}{4}\exp[-\alpha], \tag{5.6}$$

We can see that the output field of an incident beam would consist of an unmodulated component corresponding to half the incident amplitude (and a quarter of the incident energy) and two beams diffracted into the first order each of which corresponds to a quarter of the incident amplitude (and $1/16$ of the incident energy). This means that at least half the incident energy is absorbed by the hologram and it has a very low diffraction efficiency of 6.25%.

A phase hologram modulates the phase of the incident beam, and the phase shift produced is linearly proportional to the intensity in the interference pattern,

$$\Phi(\rho) = 1 - \delta I(\boldsymbol{\rho}) = \Phi_0 + \delta \cos[\alpha], \tag{5.7}$$

where Φ_0 is a constant phase, δ is the amplitude of the phase modulation and α is responsible for the fringe pattern. For the phase grating there is no amplitude modulation so $|T(\boldsymbol{\rho})| = T_0$ and this represents the absorption coefficient of the hologram. The transmission function can be expressed as

$$T_{\text{phase}}(\boldsymbol{\rho}) = T_0 \exp\left(i\delta \cos[\alpha]\right), \tag{5.8}$$

where we have neglected a global phase shift. Note that this transmission function is the Fourier transform of the interference pattern [64]. Eq. (5.8) can be expanded as a Fourier series

$$T_{\text{phase}}(\boldsymbol{\rho}) = T_0 \sum_{n=-\infty}^{\infty} i^n J_n(\delta) \exp(-in\alpha), \tag{5.9}$$

where J_n is a Bessel function of the first kind, of order n. This equation shows that an incident beam is diffracted into various orders and the amplitude of the beam which is diffracted into the nth diffraction order is proportional to the value of the Bessel function $J_n(\delta)$. Only the beam in the first order contributes to the desired image so the diffraction efficiency is given by $J_1^2(\delta)$. This results in a maximum of 60% of the amplitude of the incident beam being diffracted into the first order which is a diffraction efficiency of 30% compared to 6.25% allowed by the amplitude hologram. The low energy dissipation and high diffraction efficiency of a phase hologram compared to an amplitude hologram makes it a more suitable choice for quantum optical experiments.

Because we are using computer-generated holograms, we are not limited to these simple sinusoidal holograms [65]. In practice we can transform the transmission function into any periodic function (say g) of the fringe pattern α. The transmission functions for the amplitude and phase holograms can be given by

$$T_{\text{amp}}(\boldsymbol{\rho}) = \tfrac{1}{2}(1 + g(\alpha)), \tag{5.10}$$

$$T_{\text{phase}}(\boldsymbol{\rho}) = T_0 \exp\left(i\delta g(\alpha)\right), \tag{5.11}$$

respectively, where $g(\alpha)$ is a function that varies between -1 and $+1$ with a period of 2π, so that the range of the transmission function remains as $0 \leq T \leq 1$. To increase the diffraction efficiency of the hologram to a theoretical value of 100% we can blaze the hologram. In practice this value depends on the efficiency of the device used but the effect of blazing is maximising the diffraction efficiency as allowed by the device. The periodic function $g(\alpha)$ for a blazed hologram is given by

$$g(\alpha) = \tfrac{1}{2\pi} \operatorname{mod}(\alpha, 2\pi). \tag{5.12}$$

A blazed grating consists of a grid of parallel prisms instead of alternating opaque and transparent strips so it has the shape of a sawtooth wave as opposed to the sinusoidal grating given by $\cos(\alpha)$ [66]. We are only interested in phase holograms for which the corresponding transmission function is then given by

$$T(\boldsymbol{\rho}) = T_0 \exp\left[\frac{i\delta}{2\pi} \operatorname{mod}(\alpha, 2\pi)\right], \tag{5.13}$$

where $\operatorname{mod}(a, b) = a - b \operatorname{int}(a/b)$. This transmission function can be expanded in a Fourier series [33],

$$T(\boldsymbol{\rho}) = \sum_{n=-\infty}^{\infty} \operatorname{sinc}(n\pi + \delta) \exp(in\alpha). \tag{5.14}$$

Once again, this hologram diffracts the incident paraxial beam into a series of modes in different diffraction orders with efficiencies $|\operatorname{sinc}(n\pi + \delta)|^2$. The nth order is diffracted at a particular angle θ_n and experiences a specific transverse phase shift imposed by the hologram. The light can be

selectively directed into the desired diffracted order by setting $\delta = n\pi$ so that we have maximum diffraction efficiency to this n^{th} order. The asymmetrical fringe pattern is required to diffract light to only one side of the zeroth order (or to one side of the order that would usually contain the unmodulated light) [33].

5.1.3 Producing LG modes with Digital Holography

Our aim is to produce LG modes using the technology we have just described. For this purpose we first need to calculate the interference pattern between the desired LG mode and a reference wave, such as a Gaussian beam. The interference is then programmed onto the SLM and acts as a hologram for reconstructing the original LG mode (and its complex conjugate) when illuminated by a Gaussian beam. We will consider two methods of producing LG modes: on-axis holography by means of grey-level patterns of spiral surfaces and off-axis holography by means of *fork* holograms.

Off-axis Holography and Fork Holograms

For off-axis holography we need to calculate the interference pattern between the desired LG mode and a reference beam propagating at a small angle with respect to each other. Since we are using a thin hologram we consider the interference pattern in a particular plane, for example, the transverse plane. For simplicity, we will consider the interference pattern between a reference plane wave and a *vortex* beam which has the same azimuthal phase structure as the desired LG beam. This is acceptable because we will mainly be using this interference pattern to produce phase holograms which retain the phase information and ignore the amplitude variations [65].

The reference plane wave is given by

$$u_{\text{ref}} = E_0 e^{i(k_x x + k_z z)}, \tag{5.15}$$

where E_0 is the amplitude of the wave and k_x and k_z are the components of the wave vector k in the x and z directions respectively. The angle of incidence for this beam is $\psi = \arcsin(k_x/k)$. To simplify the calculation, we assume that the two beams have equal amplitudes of E_0, so the vortex beam can be written as

$$u_{\text{obj}} = E_0 e^{il\phi}. \tag{5.16}$$

The interference pattern produced by the desired beam and the plane wave at $z = 0$ is

$$I = |u_{\text{ref}} + u_{\text{obj}}|^2 \tag{5.17}$$

$$= 2|E_0|^2 \left(1 + \cos(k_x x - l\phi)\right) \tag{5.18}$$

$$= 2|E_0|^2 \left(1 + \cos\left(\frac{2\pi x}{\Lambda} - l\phi\right)\right), \tag{5.19}$$

where I is the irradiance on a screen at $z = 0$ and we have used $k = 2\pi/\lambda$ and $\Lambda = \lambda \sin\psi$, which is the spatial period of the fringes away from the central dislocation. An example of such an interference pattern between a plane wave and an vortex beam with azimuthal mode index $l = 1$ appears in Fig. 5.2 (b). As we can see, this interference pattern differs slightly from the rectilinear

interference pattern between two plane waves (Fig. 5.2 (a)) because the vortex beams (just like the LG modes) have l intertwined helical wavefronts as a result of their azimuthal phase distribution. The interference pattern, therefore, has l additional lines forming a multi-pronged fork on the beam axis. This hologram is aptly termed a *fork* hologram and the number of fork prongs at the centre depends on the azimuthal index l of the vortex beam used to create it.
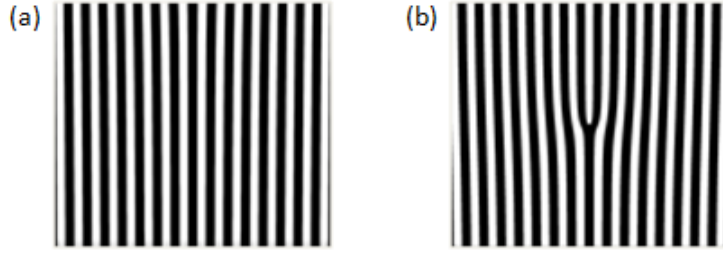
Figure 5.2: The interference pattern produced by: (a) two plane waves. (b) a plane wave an a vortex beam with azimuthal $l = 1$. Each of the waves propagate with at angle with respect to the other.

This interference pattern can now be used to form an amplitude hologram with a transmission function given by Eq. (5.4) so we have

$$T_{\text{amp}}(\boldsymbol{\rho}) = \frac{1}{2}\left(1 + \cos\left[\frac{2\pi x}{\Lambda} - l\phi\right]\right). \tag{5.20}$$

The fringe pattern for this hologram is given by $2\pi x/\Lambda - l\phi$ and if $l = 0$ we obtain the rectangular strip grating which is the interference pattern between two plane waves. Let us consider a plane wave with an amplitude A_0 traveling along the z-axis, to be incident on this hologram. If we use Eq. (5.1) we find that

$$\begin{aligned}
u_{\text{out}}(x, y) &= \frac{A_0}{2}(1 + \cos(k_x x - l\phi)) \\
&= \frac{A_0}{2} + \frac{A_0}{4}\exp(-il\phi)e^{ik_x x} + \frac{A_0}{4}\exp(il\phi)e^{-ik_x x}.
\end{aligned} \tag{5.21}$$

So the output beam consists of the unmodified plane wave propagating along the z-axis and two conjugate beams diffracted into the first orders ($n = \pm 1$) at an angle $\theta_n = \pm\psi$ with their phase profiles being modified by the phase shift $\exp(\pm il\phi)$. The two diffracted beams are vortex beams with azimuthal mode indices $\pm l$.

For our purposes we produce a phase hologram with a transmission function given by Eq. (5.8),

$$T_{\text{phase}}(x, y) = T_0 \exp\left(i\delta\cos\left[\frac{2\pi x}{\Lambda} - l\phi\right]\right), \tag{5.22}$$

where T_0 is the absorption coefficient of the hologram and δ is the amplitude of the phase modulation. Using Eq. (5.9) the transmission function of this hologram can be described by a Fourier series in polar coordinates (ρ, ϕ) as

$$T_{\text{phase}}(\rho, \phi) = \sum_{n=-\infty}^{\infty} T_n \exp\left[-in\left(\Delta m\phi - \frac{2\pi}{\Lambda}\rho\cos\phi\right)\right], \tag{5.23}$$

where $T_n = i^n T_0 J_n(\delta)$ (with $|T_n|^2$ representing the diffraction efficiency of the nth order) and we have replaced l by Δm for convenience. According to Eq. (5.23) a set of diffracted light beams or diffraction orders is formed behind the grating propagating in directions specified by the condition

$$\sin\theta_n = \frac{2\pi n}{k\Lambda}. \tag{5.24}$$

We can see that the nth term in Eq. (5.23) produces the nth diffraction order. If a Gaussian beam is incident on the hologram, the desired LG mode is reconstructed in the first positive diffraction order ($n = +1$) with an azimuthal index $l = \Delta m$ while its complex conjugate is produced in the first negative diffraction order ($n = -1$) with azimuthal index of $l = -\Delta m$. More generally, a beam formed at the nth diffraction order is an LG mode with an azimuthal index $l = n\Delta m$.

A problem arises from the fact that the hologram only introduces the correct phase term to the incident beam while leaving the intensity cross section unchanged — the intensity cross section does not match that of a perfect LG mode even though it has a perfect helical phase structure [67]. So when illuminated with a Gaussian beam, a fork hologram with Δm fork prongs actually produces a first-order diffracted beam that corresponds to a superposition of LG modes, all having the same l index but with a range of values for the radial index, p. The exact decomposition depends on the waist size of the incident beam, but in most cases the $p = 0$ mode is dominant. This is a mode with a single annular ring. For our purposes we are interested in the OAM of the beam and hence only the azimuthal index l is of interest to us. Therefore, we will consider LG modes with any azimuthal index but with radial index $p = 0$.

Blazing this hologram will allow us to improve the diffraction efficiency, as mentioned before. Using Eq. (5.13) we find that the transmission function for this blazed hologram can be expressed as

$$T(\rho, \phi) = T_0 \exp\left[\frac{i\delta}{2\pi}\bmod\left(\Delta m\phi - \frac{2\pi}{\Lambda}\rho\cos\phi, 2\pi\right)\right], \tag{5.25}$$

where Δm is the number of phase dislocations in the hologram and $\bmod(a, b) = a - b\,\text{int}\,(a/b)$. Fig. 5.3 shows a blazed hologram that is used produce an LG mode with $l = 1$. Blazing allows us to reduce the fraction of intensity diffracted to higher orders, with azimuthal indices that differ from that of the LG mode that we desire to produce. However, it does not affect the radial indices, so a superposition of LG modes with the same l and different p are still produced, with $p = 0$ being dominant.
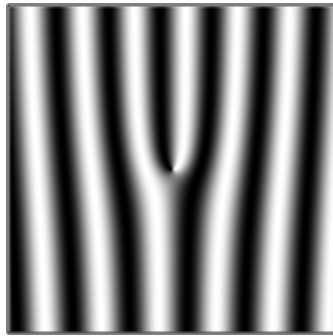
Figure 5.3: A computer-generated blazed fork hologram that can be used to produce an LG mode with $l = 1$ and $p = 0$.

On-axis Holography and Spiral Phase Patterns

We now consider a simple type of on-axis holography whereby we impose the desired phase structure on an incident Gaussian beam in order to produce the desired LG mode. Since the phase of an on-axis LG beam can be represented by the surface plot of a spiral, the phase hologram is a grey-level spiral phase pattern (Fig. 5.4). Each shade of grey on the contour plot indicates a phase shift in the range 0 (black) to 2π (white). By increasing the number of times the phase varies from 0 to 2π we can increase the phase shift imposed by the hologram on the incident beam. This means that we can increase the azimuthal index l of the LG beam that will be produced by an incident Gaussian beam. For example, the phase pattern in Fig. 5.4 (a) varies from zero to 2π once, therefore an LG mode with $l = 1$ will be produced while for Fig. 5.4 (b) the phase varies three times from zero to 2π producing an LG beam with $l = 3$. The transmission function for such a hologram is given by

$$T_{\text{on}}(\boldsymbol{\rho}) = \text{mod}[l\phi, 2\pi]. \tag{5.26}$$
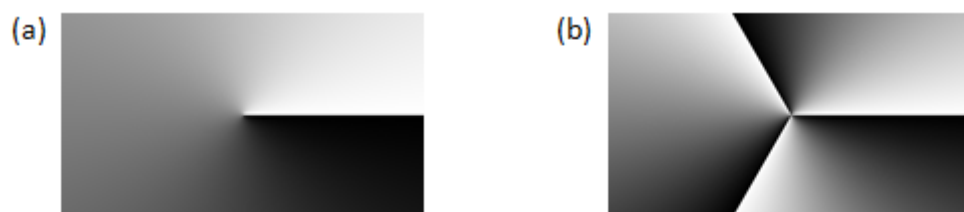
Figure 5.4: Computer-generated grey-level spiral phase patterns which allow us to produce an LG mode with azimuthal index: (a) $l = 1$ (b) $l = 3$ when a Gaussian beam is incident on them.

These spiral holograms do not diffract all the light into the desired mode so the beam that is produced is a mixture of the desired LG beam and the undiffracted incident Gaussian beam. This means that for an LG mode with azimuthal index l and radial index $p = 0$ we will have l phase

singularities in the beam instead of just one, as seen in Fig. 5.5 (a). The addition of a blazed grating to the hologram spatially separates the desired beam and the unmodulated incident beam, as seen in Fig. 5.5 (b).
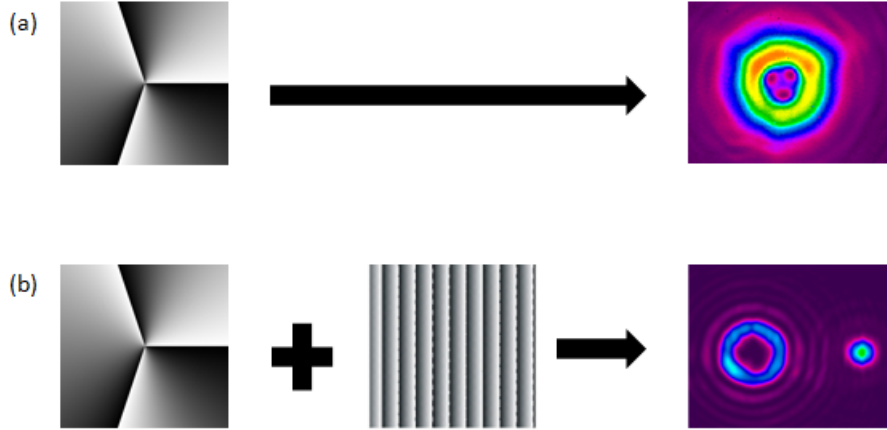
Figure 5.5: (a) A spiral phase pattern (left) used to produce an LG beam with $l = 3$ results in three phase singularities in the beam's cross section (right). (b) With the addition of a blazed grating the desired LG mode can be spatially separated from the unmodulated incident Gaussian beam.

The transmission function with the addition of the blazed grating is then given by

$$T_{\mathrm{on}}(\rho) = \mathrm{mod}\left[l\phi - \frac{2\pi}{\Lambda}\rho\cos\phi, 2\pi\right], \tag{5.27}$$

which is identical to the transmission function for the fork hologram given in Eq. (5.25). We can see from this equation and Fig. 5.6 that the addition of a grating to an on-axis hologram results in an off-axis hologram, as we have mentioned before. If the SLM device was perfect then higher orders of diffraction would not be produced. However, the pixelation of the SLM device means that a completely smooth phase transition from 0 to 2π in each line of the grating is not possible. This gives rise to the higher diffraction orders.

5.1.4 Diverse Uses of Fork Holograms

Fork holograms can be used for the generation, manipulation and detection of LG beams with azimuthal mode index l and the OAM states of photons $|l\rangle$ which is specified by Eq. (3.46). We will now describe how each of these actions may be performed by noting their effect on the index l of LG beams and on the OAM states $|l\rangle$ of a photon.

Figure 5.6: When a grating is superimposed on an on-axis hologram (our spiral phase pattern) we obtain an off-axis hologram (our fork hologram).

For the generation of an LG mode, a Gaussian beam (which has $l = 0$) must be incident on a fork hologram with Δm dislocations so that an LG beam with $l = \Delta m$ can be produced in the first diffraction order. The nth diffraction order contains an LG mode with $l = n\Delta m$. Quantum mechanically this means that if a photon in the state $|l = 0\rangle$ (with no OAM) is incident on the hologram with Δm dislocations then a photon in the first diffraction order will be in the state $|\Delta m\rangle$ (with an OAM of $\Delta m\hbar$) and a photon in the nth diffraction order will be in the state $|n\Delta m\rangle$ (with an OAM of $l = n\Delta m\hbar$). For example, if we want to produce an LG mode with azimuthal index $l = 3$ (or a photon with an OAM of $3\hbar$) in the first diffraction order then we allow a Gaussian beam (or a photon with zero OAM) to be incident on a hologram with $\Delta m = 3$ dislocations.
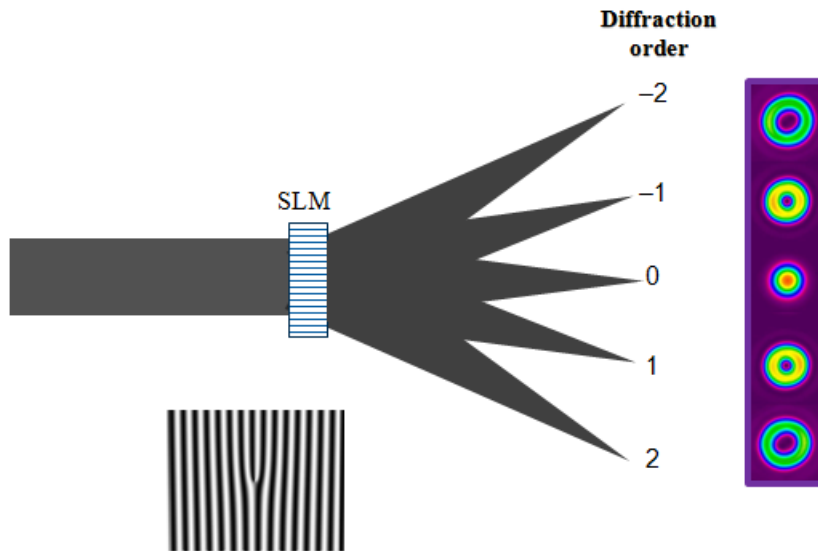
Figure 5.7: Diffraction that occurs when a Gaussian beam (l=0) is incident on a fork hologram with $\Delta m = 1$ dislocation.

Fork holograms can be used for the manipulation of LG modes and the OAM of photons as follows. If an LG mode with an azimuthal index l_{in} is incident on a hologram with Δm dislocations then the azimuthal index of the LG mode in the nth diffraction order after the hologram is given by $l_{\text{out}} = l_{\text{in}} + n\Delta m$. Quantum mechanically this means that if a photon in the state $|l_{\text{in}}\rangle$ is incident of a fork hologram with Δm dislocations then a photon in the nth diffraction order will be in the OAM state $|l_{\text{out}}\rangle = |l_{\text{in}} + n\Delta m\rangle$ with an OAM of $(l_{\text{in}} + n\Delta m)\hbar$. For example, if we want to increase the azimuthal index of an LG mode by two (or the OAM of a photon by $2\hbar$), then we use a hologram with $\Delta m = 2$ dislocations and the desired beam (or photon) will be in the first diffraction order.

For the detection of LG modes or the OAM states of photons we reverse the generation process. We note that the Gaussian beam is the only mode that can be coupled into a single-mode fibre for detection. So for the detection of a particular LG mode in the first diffraction order we aim to produce a Gaussian in that order, which means we want to "flatten the wavefront". In order to do this the azimuthal index of the incoming beam must be equal to the number of dislocations of the hologram, which results in a Gaussian beam in the $n = -1$ diffraction order. For example, if we have an input LG beam with $l = 2$ and $\Delta m = 2$ dislocations in the hologram a Gaussian beam will be formed at the diffraction order $n = -1$ and using a pin hole (so as to only allow in the Gaussian mode) a detection will be made (Fig. 5.8). Quantum mechanically, if we want to detect a photon of an OAM of $3\hbar$ in the state $|l = 3\rangle$, then using a hologram with $\Delta m = 3$ dislocations (i.e. $l = \Delta m$) we can produce a photon with zero OAM in the first diffraction order, which can be coupled into a single-mode fibre. This method allows us to carry out projective measurements in our theoretical proof-of-principle experiment.
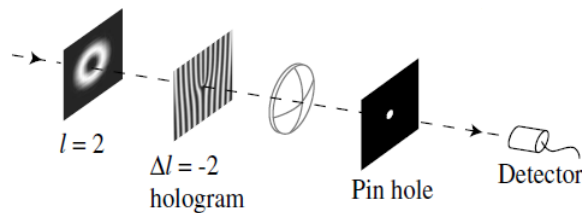
Figure 5.8: A hologram can be used to flatten the wavefronts of light modes with specific values of l, which makes it possible to focus these modes (but no others) through a pinhole, behind which they can be detected. This set-up also works for single photons [68].

5.1.5   Measuring superpositions of OAM light

In the preceding sections we have shown that it is possible to produce a single LG mode using a phase hologram. But what if we wanted to generate a superposition of LG modes? A method in which to do this was developed by Padgett *et al.* in [69]. We will now try to briefly sketch the idea behind this technique which can be used to generate an equal superposition of LG modes with azimuthal mode indices $\pm l$. We shall consider on-axis holography which was discussed in Section 5.1.3 where the azimuthal phase distribution of the desired LG mode (in the form of a

grey-level spiral phase pattern) was imposed on an incident Gaussian beam.

We wish to form a hologram which consists of an equal superposition of LG modes (with azimuthal mode indices $\pm l$) with a relative phase 2α between them so we need to impose the azimuthal phase

$$\frac{1}{\sqrt{2}}\Big[\exp(i\alpha)\exp(il\phi) + \exp(-i\alpha)\exp(-il\phi)\Big]. \tag{5.28}$$

It is clear that, apart from some constants, the phase we wish to impose is $\cos(l\phi + \alpha)$. However, because we are using a phase hologram we will not be able to reproduce smooth sinusoidal variation in the azimuthal phase — instead we produce a square-wave variation which alternates between the values -1 and $+1$, as shown in Fig. 5.9.
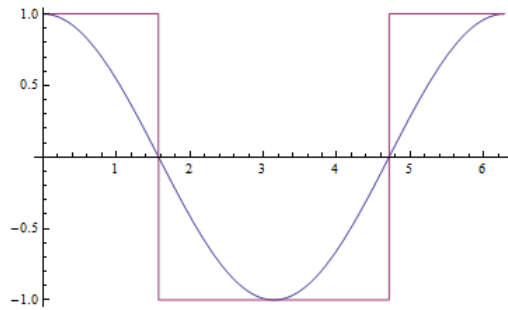
Figure 5.9: The sinusoidal phase variation $cos(l\phi + \alpha)$ that we wish to produce is represented by the blue curve while the square-wave curve s that we do produce appears in purple.

Nevertheless, it is well known that the Fourier series of the square-wave can be given in terms of cosines as

$$s(l\phi + \alpha) = \frac{2}{\pi}\cos(l\phi + \alpha) + \frac{2}{3\pi}\cos(3l\phi + \alpha) + \frac{2}{5\pi}\cos(5l\phi + \alpha) + \dots \tag{5.29}$$

Observe that the first term is proportional to the phase that we want to impose and that the amplitudes of the other factors tail off quickly as the $|l|$ increases. Therefore, we can regard the higher orders as an efficiency loss and the correct hologram is effectively generated to an excellent approximation. An example of a phase hologram that we can generate using this technique is given in Fig. 5.10. If a Gaussian beam is incident on this hologram the desired superpositions of LG modes is generated.

The phase hologram in Fig. 5.10 can also be used to project onto state

$$B(\alpha) = \frac{1}{\sqrt{2}}\left(|l = -1\rangle + e^{i2\alpha}|l = 1\rangle\right). \tag{5.30}$$

If a photon in the state $|l = 0\rangle$ is incident on the hologram then a photon in an OAM state which is an equally weighted superpositon of $|\pm l\rangle$ is generated. This means that in reverse, this hologram

will result in a detector "click" for $|l\rangle$ modes 50% of the time and $|-l\rangle$ modes 50% of the time, which is equivalent to projecting onto the superposition state.

Figure 5.10: Holograms used to generate an equal superposition of LG modes with azimuthal indices $l = \pm 1$ and different relative phases: (a) $\alpha = 0$ and (b) $2\alpha = \pi$.

5.1.6 Modulating the intensity of a phase hologram

We have seen that phase-only holograms allow us to form beams with the desired azimuthal phase distribution but with a uniform intensity distribution. However, it is possible to obtain precise control over both the phase and the intensity that is required, albeit at a reduction in the overall efficiency [70]. For our phase holograms the efficiency with which light is diffracted to the first order depends on the depth of the blazing function, for which the maximum efficiency is 2π. Varying this depth over the cross section of the hologram allows the intensity at various positions to be reduced, with any unwanted light being directed into the zeroth order. This technique allows us to accurately manipulate and detect the spatial modes we desire and to project onto the pure superposition given by Eq. (5.30). Examples of phase holograms with the appropriate intensity modulation (or blazing) appear in Fig. 5.11.
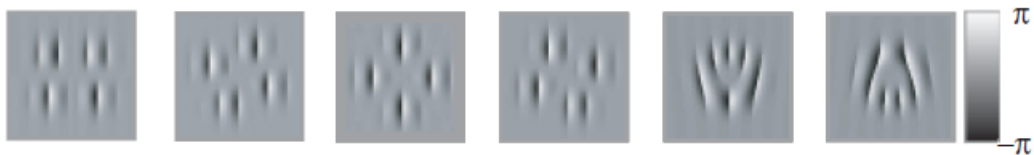
Figure 5.11: Phase holograms with the appropriate intensity blazing that can be used to generate equal superpositions of LG modes with azimuthal indices $l = \pm 2$ and different relative phases [70].

5.2 Astigmatic Laser Mode Converters

Although holograms are commonly used today to produce light which carries OAM, this was not the method that was first used by Allen *et al.* in their seminal work [8] on the OAM of light beams. Instead, they used a pair of cylindrical lenses which were able to convert HG modes into

LG modes of the same order with 100% efficiency. The production of LG modes using these cylindrical lenses is based on the fact that an HG mode rotated at $45°$ (or a diagonal HG mode) and an LG mode of the same order can be decomposed into the same set of HG modes where, in the case of the LG modes, there is a phase shift between the individual components of this decomposition (cf. Section 2.4.2). This idea was first used by Tamm and Weiss [71] who were able to convert an $HG_{n=1,m=0}$ mode into a $LG_{p=0,1=1}$ mode. This happens to be the simplest case to understand so we will now explain it. For simplicity, we will now refer to the HG modes as HG_{nm}, the diagonal HG modes as DGH_{nm} and the LG modes as LG_{pl} instead of the mode function convention that was used in Chapter 2.

5.2.1 Producing a first-order LG mode

If we want to convert a first-order HG mode (HG_{10}) into a first-order LG mode (LG_{01}) then we need to follow two steps. Firstly, we need to rotate the HG mode by $45°$ so that it becomes a first-order diagonal mode ($DGH_{n=1,m=0}$) which is a superposition of two first-order HG modes

$$DHG_{10} = \tfrac{1}{\sqrt{2}} \left(HG_{10} + HG_{01} \right). \tag{5.31}$$

The first-order LG mode is a superposition of these same modes with a relative phase shift of $\pi/2$

$$LG_{01} = \tfrac{1}{\sqrt{2}} \left(HG_{10} + iHG_{10} \right), \tag{5.32}$$

so the second step would be to introduce this relative phase shift between the two HG modes. It is important to note that the LG mode, the diagonal HG modes and the component HG modes are all of the same order, i.e. $N = n + m = 2p + l = 1$. Since we know (from Chapter 2) that an LG mode and a diagonal HG mode can both be described in terms of the same superposition of HG modes, we can apply this method of conversion to any HG mode with indices m and n to form an LG with indices $l = m - n$ and $p = \min(m, n)$ which will be of the same order as the HG mode. In order to carry out the conversion it is necessary to find a method of introducing a $\pi/2$ phase shift between successive mode components. This can be done by means of cylindrical lenses which allow us to alter the Gouy phase of the constituent modes according to their mode indices and hence introduce a relative phase between them. We will now see how this is possible.

5.2.2 Exploiting the Gouy phase with cylindrical lenses

The Gouy phase of an HG mode (Eq. (2.39)) represents the change in phase in the beam when going through the beam waist as compared to a plane wave. So in order to introduce a relative phase between the HG modes we need to alter the phase at the beam waist. Since the Gouy phase depends on the Rayleigh range, a first step would be to change the Rayleigh range without altering the position of the beam waist. (Note that the Rayleigh ranges z_{Rx} and z_{Ry} govern the divergence (or the diameter) of the beam in the x and y planes, respectively.) This can be done by inserting a cylindrical lens with the correct focal length into the Gaussian beam at $z = -d$ (see Fig. 5.12). Depending on which axis the cylindrical lens is aligned, it allows us to change the Rayleigh range z_R of an incident HG mode in the xz or yz plane. So if we align the lens parallel to the x-axis, then only the Rayleigh range in the yz plane is affected and the Rayleigh ranges z_{Rx} and z_{Ry} are no

longer equal. The beam becomes elliptical, or astigmatic, with a new Gouy that has a contribution from each transverse direction

$$\exp\left[-i\left(\left(n + \tfrac{1}{2}\right)\Phi_x^{HG}(z) + \left(m + \tfrac{1}{2}\right)\Phi_y^{HG}(z)\right)\right], \tag{5.33}$$

with

$$\Phi_x^{HG}(z) = \arctan\left(\frac{z}{z_{\mathrm{R}x}}\right), \tag{5.34}$$

$$\Phi_y^{HG}(z) = \arctan\left(\frac{z}{z_{\mathrm{R}y}}\right), \tag{5.35}$$

where we have taken the beam waist to be at $z = 0$.
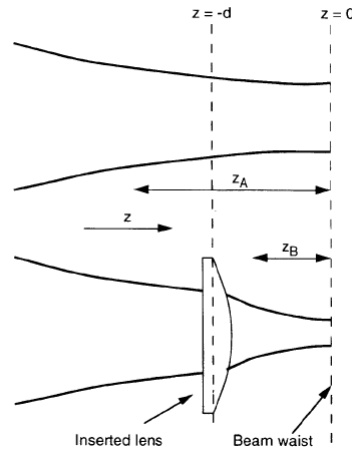
Figure 5.12: By inserting a cylindrical lens into a Gaussian beam we can change the Rayleigh range of the beam from z_A to z_B without altering the position of the beam waist [72].

The next step is to introduce a second identical cylindrical lens at $z = d$ (Fig. 5.13) aligned along the same axis, so that the Rayleigh range changes back to its original value (i.e. $z_{\mathrm{R}y}$ is once again equal to $z_{\mathrm{R}x}$). The beam is, therefore, astigmatic only in the confined region between the lenses (which are a distance $2d$ apart) and isotropic elsewhere. Because of the differing Rayleigh ranges in the xz and yz planes, when HG modes of the same order $N = n + m$ pass through the lenses they will undergo different phase shifts. The change in the Gouy phase of an HG mode oriented along the axis of the lenses, when passed through the region between them is

$$\Delta\Phi^{HG} = \left(n + \tfrac{1}{2}\right)\Delta\Phi_x^{HG} + \left(m + \tfrac{1}{2}\right)\Delta\Phi_y^{HG} \tag{5.36}$$

$$= \frac{(n + m + 1)}{2}\left[\Delta\Phi_x^{HG} + \Delta\Phi_y^{HG}\right] + \frac{(n - m)}{2}\left[\Delta\Phi_x^{HG} - \Delta\Phi_y^{HG}\right], \tag{5.37}$$

with

$$\Delta\Phi_x^{HG} = \Phi_x^{HG}(d) - \Phi_x^{HG}(-d) = 2\arctan\left(\frac{d}{z_{\mathrm{R}x}}\right),\tag{5.38}$$

$$\Delta\Phi_y^{HG} = \Phi_y^{HG}(d) - \Phi_y^{HG}(-d) = 2\arctan\left(\frac{d}{z_{\mathrm{R}y}}\right),\tag{5.39}$$

where we have written Eq. (5.37) in this form for convenience.

We now consider a diagonal HG_{nm} mode, which is aligned at $45°$ to the principal axis of the lenses, so that it can be expanded into HG modes of the same order $n+m$ oriented along the lens axes. The relative phase between the successive terms in the expansion is

$$\Delta\theta = 2\left[\arctan\left(\frac{d}{z_{\mathrm{R}x}}\right) - \arctan\left(\frac{d}{z_{\mathrm{R}y}}\right)\right].\tag{5.40}$$

This is because the successive terms in the expansion are all of the same order ($N = n+m$) so the first term in Eq. (5.37) cancels. For the second term, the difference in the value $(n-m)$ for successive terms is 2 hence we arrive at the relative phase given by Eq. (5.40). The range of values this phase can take on is from 0 to π. If this phase difference is set to $\pi/2$ the pair of cylindrical lenses introduces a factor of i^k in front of each term in the diagonal HG mode (Eq. (2.48)) so that it is converted into an LG mode with mode indices $l = n - m$ and $p = \min(n, m)$.
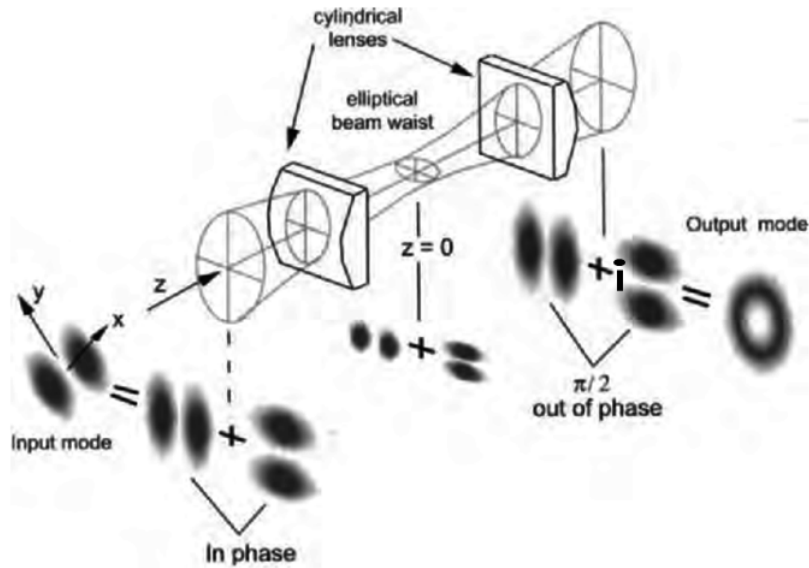
Figure 5.13: The cylindrical lens mode converter for the conversion of an HG_{10} mode into the corresponding LG_{01} mode [72].

5.2.3   Constructing a $\pi/2$ Mode Converter

Now we need to know what parameters allow us to select the relative phase that we require. Since we know that the beam is isotropic outside of the cylindrical lenses, the radius of the beam waist in the x and y directions must be equal at the positions of the cylindrical lenses ($z = \pm d$). Imposing this condition and the mode-matching condition we can find that the Rayleigh range, and hence the phase shift (Eq. (5.40)), depend on the focal length f of the cylindrical lenses as well as the distance $2d$ between them. Carrying out a simple calculation we can find that the phase difference can be set to $\pi/2$ by selecting the distance $d = f/\sqrt{2}$ and using an input beam which has a Rayleigh range $z_{\mathrm{R}y} = f + d = \left(1 + 1/\sqrt{2}\right) f$. This Rayleigh range can be achieved by using a spherical lens (called mode-matching lenses) before the cylindrical lenses and a second one after them which is used to collimate the beam. Therefore, this set of spherical and cylindrical lenses can create a phase difference of $\pi/2$ between consecutive HG modes which allows the conversion of a diagonal HG mode into an LG mode. An LG mode can also be converted to an HG mode by using the same combination of lenses to remove the relative phases, hence this set-up is aptly referred to as a *mode converter*, and more specifically in this case, a $\pi/2$ mode converter. This converter is the OAM equivalent of a quarter wave plate (QWP) since a QWP converts linearly polarised light into circularly polarised light which possesses spin angular momentum while a $\pi/2$ converter changes a beam with spherical wavefronts into a beam with helical wavefront which possesses OAM.

5.2.4   Constructing a π Mode Converter

It is possible to design a π converter which transforms any mode into its mirror image, i.e. it reverses the sense of the helical wavefronts (or the azimuthal index) for an LG mode while a diagonal HG mode is converted into a anti-diagonal HG mode. In order to obtain this converter we need to increase the separation between the lenses to $2f$ so that the phase difference between the constituent HG modes is increased to π. This allows us to introduce a relative phase of $(-1)^k$ between the components of the decompositions given in Eqs. (2.46) and (2.48). This converter is the OAM equivalent of a half wave plate for polarisation which reverses the sense of transmitted circularly polarised light. It is, however, often easier to use the optically equivalent Dove prism (cf. Section 5.3) instead of constructing this converter.

5.3   Dove prisms

The Dove Prism is a well-known optical element which is constructed from a truncated right-angled prism. When light is incident on one sloped face of this prism, total internal reflection takes place within the prism from the longest (bottom) face and the light exits from the other sloped face. Since only one reflection takes place, images that traverse the Dove prism are inverted about the x-axis, depending on the orientation of the coordinate axis with respect to the optic axis of the Dove prism. If the longest face of the Dove prism lies in the xz plane (Fig. 5.15) then the coordinate axis of the image will coincide with the optic axis of the Dove prism and the image will be flipped about the x-axis (i.e. the y-axis of the image is inverted). Alternatively, if the Dove prism is rotated along its longitudinal axis, it has a fascinating property which allows

the transmitted image and associated phase structure to be rotated at twice the rate (or twice the angular velocity) of the prism. This means that the Dove prism is able to rotate a beam of light at an arbitrary angle, therefore it can act as a beam rotator so it is very useful for applications in interferometry and astronomy [59]. We are interested in the action of a Dove prism on the OAM of a light beam or photon and we will now discuss this with our main reference being [59].
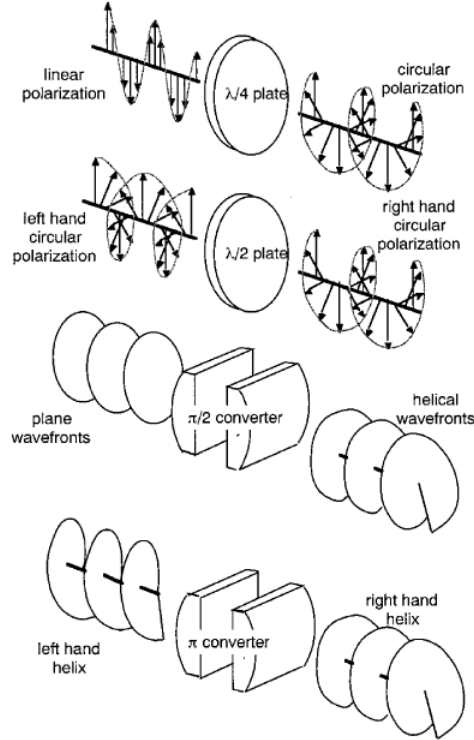
Figure 5.14: Quarter wave plates play the same role for polarisation as $\pi/2$ converters play for polarisation while half wave plates and π converters play equivalent roles for polarisation and OAM, respectively [73].

5.3.1   Manipulation of OAM using the Dove Prism

The light beams we are interested in are LG modes which carry an OAM of $l\hbar$ per photon. These modes have a spatial profile given by

$$u_{pl}^{\mathrm{LG}}(r,\phi) \simeq r^{|l|}e^{il\phi}. \tag{5.41}$$

When such a light beam traverses a Dove prism the output has a well-defined OAM of $-l\hbar$ per photon, with a spatial profile of

$$u_{pl}^{\mathrm{LG}}(r,\phi) \simeq r^{|l|}e^{-il\phi}\exp(-il\alpha), \tag{5.42}$$

where $\alpha/2$ is the angle of rotation of the Dove prism. So, as well as changing the sign of the OAM, the Dove prism introduces a phase shift of the form $l\alpha$. As an aside we note that if the Dove prism rotates continuously, the angle of rotation changes with time, and therefore the phase shift $l\alpha(t)$ depends on time. This effect allows us to observe the rotational frequency shift of light beams [74].
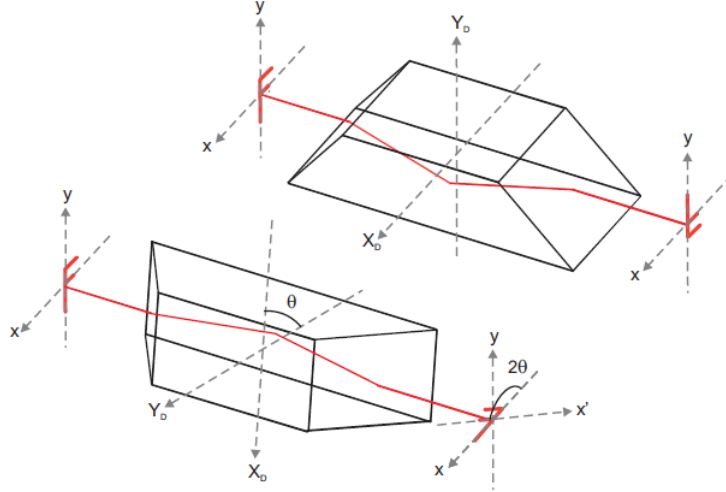
Figure 5.15: Example of how a Dove prism works. If the coordinate axis of the image coincides with the optical axis of the Dove prism flips the image (top figure). If the Dove prism is rotated an angle α about its optical axis, images passing through the prism are flipped and rotated through an angle 2α (bottom image). The phase is rotated by 2α, as well [59].

Calculation: A Rotated Dove Prism

Let us derive the expression for the sign change of the OAM ($l\hbar \rightarrow -l\hbar$) and the introduction of the phase $-l\alpha$ that occurs when an LG beam traverses a Dove prism. We can rewrite the spatial profile of an LG beam in Cartesian coordinates as

$$u_{pl}^{\text{LG}}(x,y) \simeq (x+iy)^l, \tag{5.43}$$

where $x = r\cos\phi$ and $y = r\sin\phi$. We will now consider the action of a prism rotated through an angle $\alpha/2$ on this LG mode. We know that the Dove prism performs the mirror transformation $x \rightarrow x, y \rightarrow -y$ when the coordinate axis of the beam and the optic axis of the Dove prism coincide. However, when the Dove prism is rotated, these two axes no longer coincide and the mirror transformation is now performed in a different plane. Consequently, we have to rotate the coordinate system of the beam through the same angle, $\alpha/2$, as the Dove prism is rotated. If the coordinates of this rotated coordinate system are (x_R, y_R) then the transformation of the original

coordinates (x, y) is given by

$$x = x_R \cos\left(\frac{\alpha}{2}\right) + y_R \sin\left(\frac{\alpha}{2}\right), \tag{5.44}$$

$$y = -x_R \sin\left(\frac{\alpha}{2}\right) + y_R \cos\left(\frac{\alpha}{2}\right), \tag{5.45}$$

and the incident beam can be written as

$$u_{pl}^{\mathrm{LG}}(x, y) \simeq \left[\left(\cos\left(\frac{\alpha}{2}\right) - i \sin\left(\frac{\alpha}{2}\right)\right)(x_R + i y_R)\right]^l. \tag{5.46}$$

The mirror transformation is performed in this plane so that

$$u_{pl}^{\mathrm{LG}}(x, y) \simeq \left[\left(\cos\left(\frac{\alpha}{2}\right) - i \sin\left(\frac{\alpha}{2}\right)\right)(x_R - i y_R)\right]^l. \tag{5.47}$$

The coordinate system is then rotated back to its original position so the rotated coordinates (x_R, y_R) can now be written as

$$x_R = x \cos\left(\frac{\alpha}{2}\right) - y \sin\left(\frac{\alpha}{2}\right), \tag{5.48}$$

$$y_R = x \sin\left(\frac{\alpha}{2}\right) + y \cos\left(\frac{\alpha}{2}\right), \tag{5.49}$$

and we finally obtain the complete transformation effected by the rotated Dove prism on an incident LG beam

$$u_{pl}^{\mathrm{LG}}(x, y) \simeq \left[(\cos\alpha - i \sin\alpha)(x - iy)\right]^l \tag{5.50}$$

$$= (x - iy)^l \exp(-il\alpha), \tag{5.51}$$

which is the expression that we wanted to achieve. This illustrates a reversal in the sense of the helical wavefronts as well as a rotation of the wavefront of the incident beam through an angle α i.e. $r^{|l|} \exp(il\phi) \to r^{|l|} \exp(-il(\phi + \alpha))$. Quantum mechanically we note that when a photon in the OAM state $|l\rangle$ traverses the rotated Dove prism it undergoes the transformation $|l\rangle \to \exp(-il\alpha)|-l\rangle$.

5.3.2 Combination of two Dove prisms

If we want to rotate the LG beam or the OAM state of a photon without affecting the sign of its OAM then we need to use two Dove prisms. The first prism is rotated by $\alpha/2$ and performs the transformation we have just described while the second prism is aligned along the xz plane so it merely performs the mirror transformation. Hence the total transformation of a the spatial profile of a light beam is $r^{|l|} \exp(il\phi) \to r^{|l|} \exp(il(\phi - \alpha))$ and the transformation of a photon is $|l\rangle \to \exp(-il\alpha)|l\rangle$.

5.4   Rotation Matrix for LG modes

In the HG basis of first-order modes, HG_{10} and HG_{01}, we can obtain the rotation matrix by using two Dove prisms aligned with an angle $\pi/2$ between them is given by [75]

$$U_{\text{HG}} = \left[\begin{array}{cc} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{array} \right]. \tag{5.52}$$

In order to obtain a rotation matrix for first-order LG modes we sandwich the Dove prisms between two $\pi/2$ converters. The first converter changes the LG basis to the HG basis which is a superposition of LG modes, the pair of Dove prisms carry out the rotation matrix in this basis and the second $\pi/2$ converter changes the basis back to the LG modes. In this manner, we are able to obtain the rotation matrix for LG modes.

5.5   Interferometer/Sorters

One method to sort photons according to their OAM is the interferometric technique introduced Leach *et al.* [68] where they use a Mach-Zehnder interferometer with a Dove prism inserted in each arm (Fig. 5.16). The Dove prisms are rotated at an angle $\alpha/2$ with respect to each other and, therefore, introduce an l-dependent phase shift of $\Delta\Phi_\alpha = l\alpha$ between the two optical paths in the interferometer. The rotation angle can be chosen such that photons with a specific value of l interfere constructively at, and exit through, the first output port while those with other l values leave through the second output port. For example, in Fig. 5.16, if the relative angle between the two prisms is $\alpha/2 = \pi/2$ then the phase difference between the two arms in the interferometer is $\Delta\Phi_\alpha = l\pi$, and this allows us to distinguish between photons with odd and even l values. This is because photons which l values that satisfy $(l \bmod 2) = 0$, which is all the even numbers, will exit through Port A1 whilst those which satisfy $(l \bmod 2) = 1$, which is the set of all odd numbers, will emerge through Port B1. The sorter is the OAM equivalent of a polarising beam splitter since it selects an optical path based on OAM — one path for each of the distinguishable states.
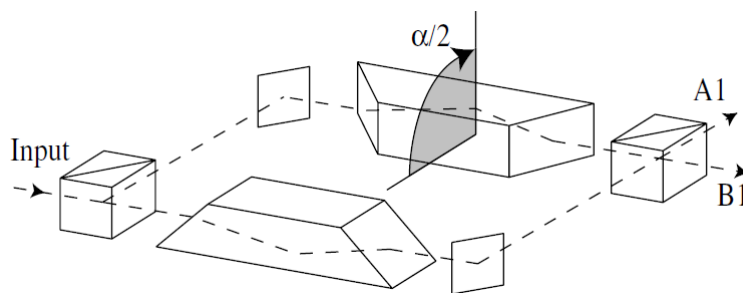
Figure 5.16: A Mach-Zehnder interferometer with a Dove prism placed in each arm. The beams in the two arms are rotated with respect to each other through an angle α, where $\alpha/2$ is the relative angle between the Dove prisms. If $\alpha/2 = \pi/2$, this device sorts photons with even values of l into Port A1 and those with odd values of l into Port B1 [68].

This interferometer is a *binary branching device* and can be extended to sort through arbitrarily large number of OAM states with cascading Mach-Zehnder interferometers having different rotation angles (see Fig. 5.17). By using a rotation angle of $\alpha/2 = \pi/2^n$, where n is a non-negative integer, all the even l values can be further sorted so that photons with $(l \bmod 2^{n+1}) = 0$ exit through one output port while those with $(l \bmod 2^{n+1}) = 2^n$ leave through the other. To sort through odd values of l, an appropriate hologram of $\Delta m = 1$ is used to convert the odd values of l to even values. The branching method that we have just mentioned can then be used to further distinguish between different l values. The problem with introducing holograms however, is that they result in optical losses and increase the complexity of the optical set-up, so it is not ideal for quantum optical experiments involving single photons.
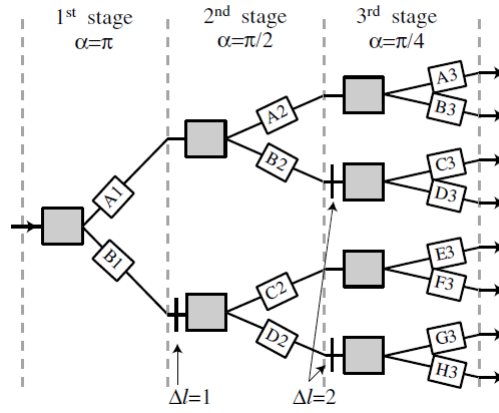
Figure 5.17: First three stages of a general sorting scheme. The gray boxes each represent an interferometer of the form shown in Fig. 5.16 with different angles between the Dove prisms. The first stage introduces a phase shift of a $\alpha = \pi$ and so sorts multiples of 2: even values of l into Port A1 and odd values of l into Port B1. The odd-l photons then pass though an $\Delta m = 1$ hologram so that they become even-l photons. The second stage introduces a phase shift of a $\alpha = \pi/2$, so it sorts even-l photons into even and odd multiples of 2. The $\Delta m = 2$ hologram is required before the photons are sorted further in the third stage [68].

The optical losses that occurs as a result of the holograms led Wei *et al.* to develop a simplified version of Leach's idea [76]. The sorter they developed now included a phase shifter in the form of a thin glass which introduces a phase offset of $\Delta\Phi_c = -k\alpha$ where k can take on values $0, 1, 2, \ldots, 2^n - 1$. This interferometer can be used to sort photons with $(l \bmod 2^n) = k$ into two groups: $(l \bmod 2^{n+1}) = k$ and $(l \bmod 2^{n+1}) = 2^n + k$. In the Leach's interferometer we would have to resort to using holograms to sort between the values $l = \pm 1, \pm 3, \ldots$ i.e. $l \bmod 2 = 1$. Using this interferometer however, with $k = 1$ and $n = 1$, we merely need to select the angle of rotation $\alpha/2 = \pi/4$ and phase offset $\Delta\Phi_c = -\alpha = -\pi/2$ to sort values $(l \bmod 4) = 1$, that is $l = \ldots, -3, 1, 5, 9, \ldots$ leave through port 1 while values of $(l \bmod 4) = 3$, that is $l = \ldots, -1, 3, 7, \ldots$ exit through port 2. This setting will be essential for our proposed experiment.
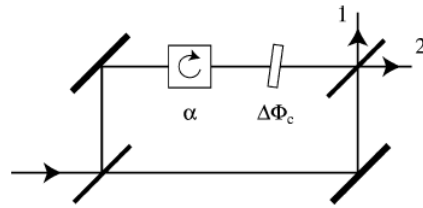
Figure 5.18: Mach-Zehnder interferometer with a beam rotator (rotation angle α) and a delay plate (phase shift $\Delta\Phi_c$) in one arm. The phase difference introduced by the beam rotator is $l\alpha$. The phase offset introduced by the phase shifter is independent of l. This device can sort light beams and even individual photons according to their l values, which exit the interferometer either through output port 1 or 2 [76].

Chapter 6

Quantum Computation Concepts

With the development of semiconductor transistors in 1948 came the advent of a new technological age — the age of electronic information processing. In a computer circuit a transistor acts as a switch and is the basic building block from which all microchips are built. It is used to represent a classical binary digit (or *bit*) since it switches between two distinct binary states: "0" if no current passes through and "1" if current does pass through. According to an empirical observation known as *Moore's law*, the number of transistors on a silicon microchip doubles roughly every two years. This means that computer power and information processing speeds increase exponentially while microchips continue to shrink in size. The miniaturisation of electronics is, however, reaching fundamental physical limitations. If Moore's law holds, then by the year 2020 the transistors will be about the size of an atom. At this point the classical notions which have been used to develop the electronics industry so far will become inadequate because quantum effects will start to dominate. Rather than being a problem, however, these quantum effects present us with the possibility to instigate the next technological revolution: the quantum information age. Many new ideas have formed out of the new quantum information paradigm, including quantum teleportation and quantum key distribution, but of interest to us is the development of quantum computation.

Quantum computing exploits the unique properties of quantum mechanics, instead of classical physics, in order to produce more efficient algorithms for computation. The advantage of quantum computing is that it provides inherent parallelism due to its quantum mechanical nature. Although it is possible for a quantum algorithm to be simulated by a classical computer, it is impossible to implement the simulation efficiently since the amount of resources required increases exponentially with the complexity of the computational problem. Therefore, quantum computers have an essential speed advantage over classical computers. This speed-up is so great that researchers believe no possible amount of progress in classical computation will be able to compete with the power of a quantum computer. Many algorithms, such as Grover's search algorithm and Shor's factorisation algorithm, which demonstrate the abilities of quantum computers have already been experimentally implemented, albeit on a very small scale using few qubits.

In this chapter a few basic concepts of quantum computation will be explained. In Section 6.1 we offer a short description of the basic unit for computation — the *qubit*, which is the quantum mechanical analogue of the classical bit. In Section 6.2, we move onto the quantum mechanical

feature of entanglement which gives quantum computation the essential speed-up over the classical method. A set of universal gates, which are the basic elements that allow us to compute a quantum algorithm, is described in Section 6.3. Section 6.4 provides a basic introduction to the standard model of quantum computation while Section 6.5 explains how Grover's search algorithm can be demonstrated using a quantum computer. Finally, in Section 6.6 one can find that criteria that must be satisfied in order to implement a quantum computer.

6.1 Qubit

In classical computation the bit is the fundamental unit for storing information. The bit can take on one of two discrete values: "0" or "1", and with a sufficiently long string of bits, any information can be encoded. The analogous concept in quantum computation is the qubit which, in contrast to its classical counterpart, can be in a coherent superposition of two states

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle. \tag{6.1}$$

where $|\psi\rangle$, $|0\rangle$ and $|1\rangle$ are ket-vectors in a two-dimensional Hilbert space $\mathcal{H}_2$ and α and β are complex numbers. The states $|0\rangle$ and $|1\rangle$ correspond to the logical states 0 and 1, respectively, of the classical bit and are thus mutually exclusive, i.e. their inner product is equal to zero. These states are known as the computational basis and form an orthonormal basis for the Hilbert space $\mathcal{H}_2$.

Classically, by examining the bit we can determine its state with certainty. Quantum mechanically, however, we can only measure the probability for the qubit to be in a certain state: $|\alpha|^2$ gives the probability of obtaining the result "0" while $|\beta|^2$ is the probability to obtain "1". Since the probabilities must equal unity we have the normalisation condition: $|\alpha|^2 + |\beta|^2 = 1$. This means that, in general, the state of a qubit is a unit vector in $\mathcal{H}_2$. Note that even though the qubit can exist as a superposition, in a measurement it can, at maximum, carry only one bit of classical information.

Any state in $\mathcal{H}_2$ can be written in a different orthonormal basis which spans the space. If the state can be represented by a sum of basis vectors, as in Eq. (6.1), then it is called a pure state. If such a description is not possible and the state is a statistical mixture of pure states, then it is called a mixed state. Both pure and mixed states can be written in terms of a density matrix, ρ. For example, Eq. (6.1) can be rewritten as

$$\rho = |\psi\rangle\langle\psi| = |\alpha|^2|0\rangle\langle0| + \alpha\beta^*|0\rangle\langle1| + \alpha^*\beta|1\rangle\langle0| + |\beta|^2|1\rangle\langle1|$$
$$= \begin{bmatrix} |\alpha|^2 & \alpha\beta^* \\ \alpha^*\beta & |\beta|^2 \end{bmatrix}. \tag{6.2}$$

The diagonal elements represent the probability amplitudes of the basis states while the off-diagonal elements contain information about the coherence between the two basis states, i.e. about the purity of the superposition. In order for the state to be physical, the density matrix must satisfy certain conditions:

1. it must be Hermitian ($\rho = \rho^{\dagger}$);

2. it must be non-negative (real, non-negative eigenvalues — $\lambda_j \geqslant 0$);

3. its trace (the sum of the elements on the main diagonal) must equal one, which is the normalisation condition.

The qubit gives quantum computation an advantage over classical computing because an N-bit register can store only one of the N possible binary numbers, while an N-qubit register can store all N numbers simultaneously, as a superposition. Therefore, as the number of qubits in the register increases, the capacity for storage of information increases exponentially. Qubits also allow for a speed-up by parallel processing. This can be seen from the fact that in classical computing, in order to evaluate a function at N values we would need to do each computation individually, whereas in quantum computing unitary operations on the N-qubit register can evaluate the function at all N values simultaneously.

In principle, any quantum mechanical system composed of two distinct (or orthogonal) states or levels can act as a qubit. Thus an example of a qubit could be a single atom with two different energy levels (*ground* and *excited* state), the spin of an electron (*spin-up* and *spin-down* state) or the polarisation of a photon (*horizontal* and *vertical* states). For our purposes we will be considering two OAM states, namely $|l = -1\rangle$ and $|l = +1\rangle$, which is a two-dimensional subspace of an infinite-dimensional space.

Geometrically we can visualise the Hilbert space of our OAM qubit as a sphere, called the Bloch sphere. This sphere is analogous to the Poincàre sphere for the spin angular momentum states of light, as was shown by Padgett and Courtial in [77]. Each point on the Bloch sphere represents a different pure state and its antipode represents the orthogonal state. Mixed states can also be represented as vectors that lie in the interior of the sphere. In general the state $|\mathbf{a}\rangle$, can be written as

$$|\mathbf{a}\rangle = \cos\left(\frac{\theta_a}{2}\right)|l = 1\rangle + e^{i\phi_a}\sin\left(\frac{\theta_a}{2}\right)|l = -1\rangle, \tag{6.3}$$

and corresponds to the vector $\mathbf{a} = (\sin\theta_a\cos\phi_a, \sin\theta_a\sin\phi_a, \cos\theta_a)$ on the Bloch sphere (Fig. 6.1) with latitude $0 \leqslant \theta \leqslant \pi$ and longitude $0 \leqslant \varphi < 2\pi$. The states $|l = +1\rangle$ and $|l = -1\rangle$ correspond to the poles of the sphere ($\theta_a = 0, \pi$) respectively, while equally weighted superpositions of $|l = +1\rangle$ and $|l = -1\rangle$ correspond to points around the equator ($\theta_a = \pi/2$), with the equatorial position being determined by the phase term ϕ_a between the two states.

In addition to qubits, we can also store information using quantum systems with larger state spaces. Such a d-level quantum system is called a *qudit*. A natural example of a qudit in our case would be the OAM states of a photon. However, their state space is infinite-dimensional and, as we have mentioned, we had to artificially restrict the system to two levels in order to obtain an OAM qubit. In the same way we can restrict the system to d levels. One advantage of using qudits over qubits for quantum computation is that more complicated calculations can be performed with the same amount of resources. Since the OAM of a photon can be easily transformed into a qudit it would make sense to utilise the extra degrees of freedom it provides. Just as for a qubit, a qudit can be

represented as a superposition of the d computational basis states,

$$|\psi\rangle = \sum_{i=0}^{d-1} \alpha_i |i\rangle. \tag{6.4}$$
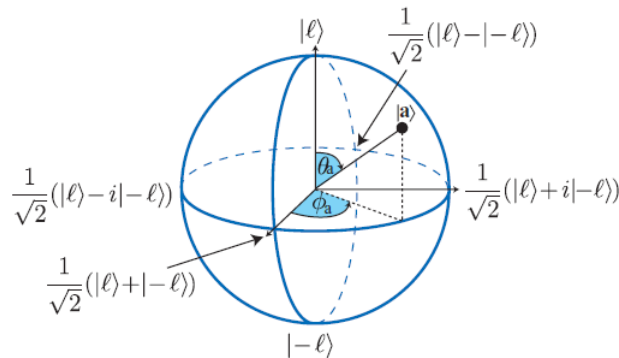
Figure 6.1: Bloch sphere equivalent for OAM states $|l = \pm 1\rangle$ [70].

6.2 Entanglement

One of the most significant consequences of using qubits is that the combined states of the qubits can exhibit correlations which cannot be observed when using classical systems. The quantum mechanical feature associated with these non-classical correlations is called entanglement. In 1935, Erwin Schrödinger stated that entanglement is *"not one but rather the characteristic trait of quantum mechanics, the one that enforces its entire departure from classical lines of thought"* [78]. It is the fundamental resource for quantum computation which allows for a speed-up over classical computation and it is the reason that quantum computation cannot be simulated efficiently on a classical computer. More importantly, a highly entangled multi-partite state is the essential ingredient for one-way quantum computation.

Two qubits are entangled if their combined pure state is not separable, i.e. it cannot be written as the product of the states of each of the individual qubits. Measuring the state of one qubit instantaneously defines the state of the other even though each qubit does not have a definite state before the measurement. Let us illustrate this concept with an example.

Consider a bipartite system in the state

$$|\Phi^+\rangle = \tfrac{1}{2}\Big[|0\rangle_1|0\rangle_2 + |1\rangle_1|1\rangle_2\Big], \tag{6.5}$$

where the subscripts 1 and 2 refer to qubits 1 and 2, respectively. This state cannot be factorised — it cannot be written as $|\Phi^+\rangle = |\psi\rangle_1 \otimes |\psi\rangle_2$, where each qubit has a definite state. We say that the

two qubits are entangled since they can only be described by a single quantum state, even though they may be separated by large distances. Until we perform a measurement, no information is available about each individual qubit, only about both jointly. Entanglement implies that if a measurement is made the results are always perfectly correlated. For example, if qubit 1 is measured and projected onto the state $|0\rangle$ then qubit 2 is projected onto the state $|0\rangle$ as well, and vice versa. So by measuring the state of one qubit we can obtain the state of the other instantaneously. Interestingly, entanglement can occur between particles that are arbitrarily separated, both spatially and temporarily.

If we wish to entangle two qubits ($|\psi\rangle_1$ and $|\psi\rangle_2$) then we have to allow for some kind of interaction between them. The use of only local operations (specified by the operator $\hat{\mathbf{A}}$) and classical communication would result in a series of operators acting on one qubit and a separate series of operations acting on the other. The final state will always remain separable,

$$(\hat{\mathbf{A}}_1 \otimes \hat{\mathbf{A}}_2)(|\psi\rangle_1 \otimes |\psi\rangle_2) = \hat{\mathbf{A}}_1|\psi\rangle_1 \otimes \hat{\mathbf{A}}_2|\psi\rangle_2. \tag{6.6}$$

We, therefore, use the non-linear process of spontaneous parametric down-conversion (cf. Chapter 4) in order to create entangled photon pairs for our proposed experiment.

6.3 Universal Set of Gates

From computational theory we know that a minimal set of all possible logic gates is necessary to perform any computation. This set is known as a *universal* set and in the classical case it is given by the NAND gate. All logic gates can be constructed from an appropriate number of NAND gates.
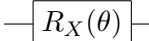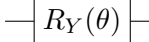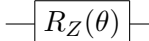
In analogy to classical logic gates, unitary operations on qubits can be thought of as quantum logic gates since they are the building blocks for a quantum computer. Unlike classical gates however, quantum gates are reversible. It is well known that any unitary operation that acts on a discrete number of qubits can be approximated with arbitrary accuracy by a series of single- and two-qubit gates [5]. This set of single-qubit gates (which carry out arbitrary single-qubit operations) together with an entangling two-qubit gate is sufficient to implement any quantum algorithm, so they form a universal set of gates for quantum computation.

One possible universal set of gates is given by the two-qubit controlled-NOT (CNOT) gate and the single-qubit Hadamard (H) and $\pi/8$-phase (T) gates. Another possible set is the two-qubit controlled phase (CZ) gate and rotations operators ($R_X(\theta)$, $R_Y(\theta)$, $R_Z(\theta)$). Let us now explain the effect of these quantum gates when applied to the computational basis states $|0\rangle$ and $|1\rangle$.

The CNOT gate takes in and outputs two qubits, the control bit and the target bit. The control bit is unaffected throughout the computation. If the control bit is "0" then the target bit remains unchanged, but if the control bit is "1" then the target bit is flipped. The operation of the CZ gate works in a similar way but instead of flipping the target bit when the control bit is "1", the target bit obtains a phase shift of π. The Hadamard gate allows for the equal superposition of the two orthogonal basis states to be formed out of one of these states. It also allows for the reverse

operation, reducing a superposition to a single state. The $\pi/8$-phase gate introduces a relative phase shift of $\pi/4$ between the orthogonal basis states of the qubit. Finally, the rotation operators $R_X(\theta)$, $R_Y(\theta)$ and $R_Z(\theta)$ perform rotations of the qubit on the Bloch about the axes X, Y and Z, respectively, by an angle θ. Table 6.1 contains the unitary matrices for each of these gates.

Table 6.1: Unitary matrices for a few quantum gates.

Name	Gate	Operation
Hadamard	$-\boxed{H}-$	$\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$
$\frac{\pi}{8}$	$-\boxed{T}-$	$\begin{bmatrix} 1 & 0 \\ 0 & e^{i\frac{\pi}{4}} \end{bmatrix}$
$R_X(\theta)$	$-\boxed{R_X(\theta)}-$	$\begin{bmatrix} \cos\frac{\theta}{2} & -i\sin\frac{\theta}{2} \\ -i\sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{bmatrix}$
$R_Y(\theta)$	$-\boxed{R_Y(\theta)}-$	$\begin{bmatrix} \cos\frac{\theta}{2} & -\sin\frac{\theta}{2} \\ \sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{bmatrix}$
$R_Z(\theta)$	$-\boxed{R_Z(\theta)}-$	$\begin{bmatrix} e^{-i\frac{\theta}{2}} & 0 \\ 0 & e^{i\frac{\theta}{2}} \end{bmatrix}$
Controlled NOT (CNOT)		$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$
Controlled Phase (CZ)		$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$

6.4 The standard model of quantum computation

The most widely used model for quantum computation at present is the quantum circuit model, or the standard model of quantum computation. This model is a generalisation of the classical circuit model which uses Boolean logic operators like AND and NOT. Information is stored in a register of n qubits which is initialised in the state $|0\rangle^{\otimes n}$. Such a state can be reliably reproduced with low variability and is thus referred to as a *fiducial state*. The circuit is then made of wires which carry the information and quantum logic gates which perform arbitrary unitary operations on the qubits in order to carry out a computational task (Fig. 6.2). The gates which are used in the circuit are part of a universal set of single- and two-qubit gates, as mentioned in the previous section. A large entangled state is developed before final measurements are made to give the result of the computation. The entangled state is projectively measured in some basis to convert the quantum state into a string of logical bits. Because measurement results are probabilistic, classical post-processing of these results is necessary in order to obtain the solution to the computational problem.

The size of a quantum circuit which is needed to implement a quantum algorithm is given by the number of gates in the circuit. In order for this circuit to be efficient, the number of gates must not increase exponentially as the size of the input register increases. The circuit must be scalable such that the algorithm can be implemented with a larger input register but with only a modest amount of additional computational resources, i.e. the circuit must be polynomially bounded in the number of input bits.
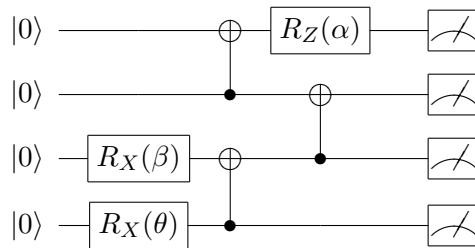
Figure 6.2: The standard model of quantum computing begins with a register, which contains four qubits in this case. Each qubit is initialised to a fiducial state, $|0\rangle$ for example. Single-qubit rotations and the controlled-NOT gate, which form a universal set of quantum gates, process the information in the register. A highly entangled state is obtained, which is then measured to give the output of the computation.

6.5   A Quantum Algorithm: Grover's Search Algorithm

Grover's search algorithm is one of the fundamental algorithms in quantum computing that stimulated interest in the field. This quantum algorithm was invented by Lov Grover in 1996 [79] and performs an efficient search through an unstructured database. It is an important algorithm since fast searches are essential in solving difficult computational problems [7]. A common example of an unstructured search is looking for a person's name in a phone directory when given a specific telephone number. If there are N entries in the database then classically the number of steps needed to locate the correct element would be $O(N)$ and $O(N/2)$ on average. Grover's algorithm however, utilises quantum parallelism to enhance the probability of finding the desired outcome in only $O(\sqrt{N})$ steps. Although this presents only a polynomial advantage over the classical case (compared to the usual exponential speed-up of other quantum algorithms) this speed-up is considerable when the database is large. As for all quantum algorithms, Grover's algorithm outputs the correct solution with only a high probability and by performing the algorithm repeatedly we have a better chance of obtaining the desired output.

Let us briefly outline the basic idea of this algorithm. Consider a database of N elements where one of these is the solution to our search problem. We have a device called an *oracle* which recognises and marks the correct solution. At this point we imagine that the database is a black box where the solution is marked. Since we cannot look into the box and identify the marked element our only option is to perform some operations on the elements in the black box so as to increase the probability of finding the correct solution, and decrease the probability of obtaining any of the

other elements, when a measurement is carried out. This operation is called the *inversion-about-the-mean* operation. By performing this operation together with the oracle several times we can obtain the solution with a very high probability.

6.5.1 Mathematical description

We will now give a mathematical description of the algorithm. For simplicity we will consider a database of N elements with only one solution to the search problem and $N = 2^n$ where n is the number of qubits required for the computation. This explanation is along the lines of the one given in [5]. As an aside note that since we are using qubits, the number of resources needed for this algorithm is given by $\log_2 N$ instead of N bits as in the classical case.

Let us say that we want to search through a database of N elements y_x, where x is the index of the element. So for example, y_0 is the first element in the database with index $x = 0$. Rather than search through the elements directly we concentrate on the indices of these elements, which is just a number in the range from 0 to $N - 1$. The algorithm begins by creating a superposition of the indices in the database

$$|\psi\rangle = \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} |x\rangle. \tag{6.7}$$

For the search problem we can define a function f such that $f(x) = 1$ when x indexes the solution to the problem and $f(x) = 0$ otherwise. We can construct a device called the oracle which marks the solution to the problem. The oracle takes an instance of the index register $|x\rangle$ and a single oracle qubit $|q\rangle$, as input. The unitary operation of the oracle O is defined as

$$|x\rangle|q\rangle \xrightarrow{O} |x\rangle|q \oplus f(x)\rangle, \tag{6.8}$$

where $\oplus$ denotes addition modulo 2 and the oracle qubit $|q\rangle$ is flipped if $f(x) = 1$ and remains unchanged otherwise. This operation can be simplified as

$$|x\rangle|q\rangle \xrightarrow{O} (-1)^{f(x)}|x\rangle|q\rangle, \tag{6.9}$$

and the operation of the oracle on the index register is, therefore,

$$|x\rangle \xrightarrow{O} (-1)^{f(x)}|x\rangle. \tag{6.10}$$

This means that the oracle marks the index of the solution with a π phase shift (or a minus sign). If we performed a measurement on the index register at this point we would not be able to tell which element is the solution since phase shifts are irrelevant in a measurement and each of the elements are equally weighted in the superposition. Therefore, there is an equal probability of obtaining any of the elements as a solution. We need to increase the probability of finding the marked element so we apply the operation

$$2|\psi\rangle\langle\psi| - I, \tag{6.11}$$

to the index register. This operation is the inversion-about-the-mean operation that we have previously mentioned, and if it is used together with the oracle we obtain the Grover iteration

$$G = (2|\psi\rangle\langle\psi| - I)O. \tag{6.12}$$

By applying the Grover iteration to the index register several times we can obtain the correct solution with a high probability. Remarkably, for a database of four elements (where we use just two qubits) we can obtain the solution with 100% probability with a single application of the Grover iteration.

6.5.2 Geometrical interpretation

A geometrical interpretation of the algorithm will allow us to better understand it and to determine the number of operations that will have to be applied in order to obtain the correct result. Let the superposition of all elements which are not a solution to the problem be denoted by

$$|\alpha\rangle = \frac{1}{\sqrt{N-1}} \sum_{x=0}^{N-1} |x\rangle, \tag{6.13}$$

and $|\beta\rangle$ be the solution to the problem. We can rewrite the initial superposition (Eq. (6.7)) as

$$|\psi\rangle = \sqrt{\frac{N-1}{N}}|\alpha\rangle + \sqrt{\frac{1}{N}}|\beta\rangle. \tag{6.14}$$

Let us see what is the result of applying the Grover iteration to this state. Since the oracle merely marks the solution in the superposition with a minus sign it can be represented geometrically as a reflection about the vector $|\alpha\rangle$ in the plane defined by $|\alpha\rangle$ and $|\beta\rangle$ (Fig. 6.3), i.e.

$$O(a|\alpha\rangle + b|\beta\rangle) = a|\alpha\rangle - b|\beta\rangle. \tag{6.15}$$

In the same way, the inversion-about-the-mean operation performs a reflection in the same plane about the vector $|\psi\rangle$. Therefore, the complete Grover iteration is actually a rotation in the plane. Several applications of the Grover iteration rotates the superposition $|\psi\rangle$ closer and closer to the solution $|\beta\rangle$ which means that with every iteration the amplitude of the solution in the superposition increases. As a result, when a measurement is made after several iterations, the solution is obtained with a high probability.

The angle of rotation θ (in Fig. 6.3) is a real number in the range 0 to $\pi/2$, chosen such that

$$\sin\theta = \frac{2\sqrt{N-1}}{N}. \tag{6.16}$$

The number of times that the Grover iterations must be applied is given by

$$R \approx \frac{\pi\sqrt{N}}{4}. \tag{6.17}$$

So $R = O(\sqrt{N})$ Grover iterations are necessary to obtain the correct solution with high probability, which is a quadratic advantage over the $O(N)$ oracle calls that are required classically [5].
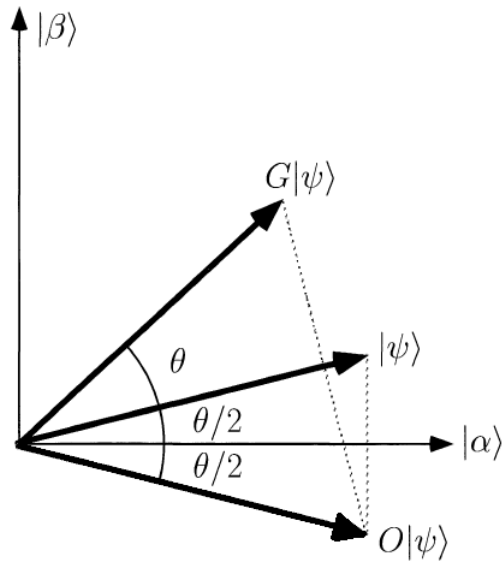
Figure 6.3: A geometric representation of a single Grover iteration. The oracle O reflects the superposition of all database elements $|\psi\rangle$ about $|\alpha\rangle$ (the superposition of all elements that are not the solution to the search problem). The $G = (2|\psi\rangle\langle\psi| - I)O$ operation then reflects the new state about $|\psi\rangle$. After repeated applications of the Grover iteration, the superposition $|\psi\rangle$ is rotated closer to the solution $|\beta\rangle$ so that there is a very high probability of obtaining the solution when a measurement is made in the computational basis [5].

6.6 The DiVincenzo Criteria

In 1996, David DiVincenzo formalised the criteria that the physical system supporting a quantum computer must fulfil. These five criteria are now known as the DiVincenzo criteria and have inspired new theoretical and experimental research in quantum information processing. They apply explicitly to the standard model and can be stated as [80]:

1. A scalable system with well-characterised qubits;
2. The ability to initialise the state of the qubits to a simple fiducial state, such as $|000\ldots\rangle$;
3. Long relevant decoherence times, much longer than the gate operation time;
4. A universal set of quantum gates;
5. A qubit-specific measurement capability.

In terms of applications, these criteria can be explained as follows. The physical system supporting the qubits should be scalable, which means that it should be possible generate many qubits without incurring an exponential cost of resources (such as time, space or energy). Once this is achieved it should be possible to initialise the qubits in a well-defined state. It must then be possible to apply a universal set of quantum gates on the qubits and the amount of resources necessary to implement these gates should not grow exponentially. Also, to allow for efficient quantum computation the time scales on which the gates act on the qubits should be sufficiently small compared to the time it takes for the qubit system to experience any decoherence. Finally, in order to carry out a

computation it should be possible to measure specific qubits individually, in a reliable and efficient manner.

Photons, which are the physical systems used in our proof-of-principle experiment, do meet almost all of these criteria but it is quite difficult to realise two-qubit gates due to the lack of photon-photon interaction. Also, both state initialisation and detection is currently a probabilistic process [81]. They do, however, have many advantages such as being robust to environmental noise and being the fastest information carries since they travel at the speed of light. It has been shown that, in principle, linear-optical quantum computing is always possible [82].

Chapter 7

One-way quantum computing

In 2000, Hans Briegel and his student Robert Raussendorf [6] created a completely different approach to quantum computing from the standard model, called *one-way quantum computing*. Here, all the entanglement that is ever needed for the computation is generated at the very beginning of the computation by creating a highly entangled multi-partite state called the *cluster state*. Information is processed and read out via a series of single-qubit projective measurements on each of the qubits in the cluster state. Depending on the choice and order of these measurements different algorithms can be executed. Remarkably, although this type of computation uses only quantum measurements which are not reversible or unitary, we are able to simulate the circuit model and implement a universal set of gates. Since measurements are inherently probabilistic we can make the computation deterministic by means of a process called classical feed-forward.

The word one-way is a direct translation of the German word *Einweg*, which could also be translated as disposable, and is meant to reflect the nature of the cluster state since this entanglement resource is destroyed by the measurements and can thus be only be used once. In some sense the cluster state serves as the hardware for the one-way quantum computer while the measurement sequence is the software or program. This model provides a scalable means of quantum computation and could lead to dramatic simplifications in the experimental implementations [81].

In this chapter we will explain the main features of this measurement-based model for universal quantum computation. Section 7.1 introduces the concept of a cluster state as the main resource for one-way quantum computing and illustrates the graphical representation of this family of quantum states. Section 7.2 gives us a detailed description of the implementation of single-qubit gates by measurements only. Section 7.3 demonstrates the idea of classical feed-forward of measurement outcomes as a means to neutralise the randomness of measurements. Section 7.4 discusses how two-qubit entangling gates are inherent in the model, and together with the single-qubit gates this makes up a universal set. Section 7.5 describes a unique set of gates called the Clifford group gates which require only Pauli measurements while Section 7.6 provides a method for etching out the cluster states that you require from a square cluster state by means of Pauli-Z measurements. Finally, Section 7.7 ends off with a summary of the one-way quantum computing model. The main resources for this chapter are [81, 83, 84].

7.1 Cluster states

Cluster states are the essential entanglement resource for the one-way quantum computer. Single-qubit measurements on this state can be used to imprint a circuit on it and process information. This state is universal for quantum computation which means that with an appropriate and large enough cluster state, any quantum logical operation can be carried out. Because of the importance of cluster states for one-way quantum computing this model is also referred to as *cluster state quantum computation*.

The preparation of the cluster state is a two-step process. Firstly, all qubits are prepared in the superposition state

$$|+\rangle = \tfrac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \tag{7.1}$$

where $|0\rangle$ and $|1\rangle$ are the computational basis states. Secondly, the controlled phase (CZ) operation

$$CZ = |0\rangle\langle 0| \otimes \mathbb{I} + |1\rangle\langle 1| \otimes Z, \tag{7.2}$$

is applied between neighbouring qubits in order to entangle them. If the control qubit is $|0\rangle$ then the target qubit does not change but if the control qubit is $|1\rangle$ then the Pauli-Z gate is applied to the target qubit. This gate is symmetric with respect to the two qubits, which means that it does not matter which qubit acts as the control and which is the target — the same entangled state is formed after application of the CZ gate. Since CZ gates are represented by diagonal matrices they commute with each other and can be applied to all the qubits in the cluster state simultaneously, or in an arbitrary order. Therefore, the entanglement process requires only a single step.

We note that in the circuit model all qubits are initialised in the $|0\rangle$ state. Preparing the qubits in this state would not be useful for creating a cluster state since no entanglement can be formed between the qubits — the CZ gate has no effect when the control qubit is in the $|0\rangle$ state. If we do initialise the qubits in the $|0\rangle$ state then we must apply the Hadamard (H) to all qubits before applying the CZ gates. This operation can be done simultaneously since Hadamard transformations commute with each other.

7.1.1 Examples of cluster states

At this point it would serve us well to illustrate how a few simple cluster states can be formed by using the two steps mentioned above. In each case the control qubit will be expanded as in Eq. (7.1) while the target qubit will be written simply as $|+\rangle$. Remember that we have freedom in choosing which qubit should be the control and which will be the target. We shall exclude normalisation factors for the moment since these are only important when we are concerned with measurements, i.e. when we need to the know the probability of obtaining a certain measurement outcome.

Two-qubit cluster state

1. Initialise the qubits: $|+\rangle_1 \otimes |+\rangle_2 = (|0\rangle_1 + |1\rangle_1) \otimes |+\rangle_2$.

2. Apply the CZ operation: $|0\rangle_1|+\rangle_2 + |1\rangle_1|-\rangle_2$.

In the second step we have used the relation $Z|+\rangle = |-\rangle$ where $|-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$. Note that two states are local unitarily equivalent if we can map from one to the other by the application of local unitaries (which are single-qubit gates, not entangling two-qubit gates). The application of the Hadamard to the second qubit results in a state that is unitarily equivalent to the two-qubit cluster state

$$|\Psi^+\rangle = |00\rangle + |11\rangle. \tag{7.3}$$

This state is one of the famous maximally-entangled two-qubit states called the *Bell states*.

Three-qubit cluster state

1. Initialise the qubits: $|+\rangle_1(|0\rangle_2 + |1\rangle_2)|+\rangle_3$.

2. Apply the CZ operation: $|+\rangle_1|0\rangle_2|+\rangle_3 + |-\rangle_1|1\rangle_2|-\rangle_3$.

Now if we apply the Hadamard to qubits 1 and 3 we find that the cluster state is unitarily equivalent to

$$|\Psi_{\text{GHZ}}\rangle = |000\rangle + |111\rangle, \tag{7.4}$$

which is the well-known Greenberger-Horne-Zeilinger (GHZ) state.

Four-qubit cluster state

1. Initialise the qubits: $(|0\rangle_1 + |1\rangle_1)|+\rangle_2(|0\rangle_3 + |1\rangle_3)|+\rangle_4$.

2. Apply CZ: $|0\rangle_1|+\rangle_2|0\rangle_3|+\rangle_4 + |0\rangle_1|-\rangle_2|1\rangle_3|-\rangle_4 + |1\rangle_1|-\rangle_2|0\rangle_3|+\rangle_4 + |0\rangle_1|+\rangle_2|1\rangle_3|-\rangle_4$.

If we apply the Hadamard to qubits 1 and 4 and simplify, we find that this cluster state is unitarily equivalent to

$$|\Psi_{\text{C}_4}\rangle = |0000\rangle + |0011\rangle + |1100\rangle - |1111\rangle, \tag{7.5}$$

which is the state that Walther *et al.* [7] produced for their proof-of-principle experiment on one-way quantum computing, as we will see in the Chapter 8.

7.1.2 Graphical Representation

Clearly, using the state vector representation becomes increasing complex as more qubits are added to the cluster state. Although it is easy to calculate the state vector it is often a tedious, inefficient process and quite useless since it tells you nothing about the computations that can be performed using this cluster state. A more effective description of the cluster state is obtained if each cluster state is represented by a graph. A graph, in mathematics, is a particular kind of diagram composed of vertices (represented by points) which are connected by edges (represented by lines between two vertices). In quantum information, graph states are a particular class of quantum states that can be represented by a graph. Note that it is not possible to represent all states by a graph, only certain families.

Cluster states are a subset of graph states whose underlying graph is a square grid (Fig. 7.1). Each vertex corresponds to a qubit initialised in the $|+\rangle$ state and an edge between two vertices denotes the application of a CZ gate between the corresponding adjacent qubits. A general graph state (which does not have the grid structure of a cluster state) is more flexible than the cluster state since it often requires fewer qubits to implement the same algorithm, but cluster states are generally generated more efficiently [84].

Figure 7.1: Graphical representation of a particular cluster state. The vertices symbolise qubits while the edges between two vertices indicate the application of a CZ gate between the corresponding qubits.

7.2 Single-qubit gates

In the standard model of quantum computation a circuit that contains a universal set of gates (which could consist of single-qubit rotations and an entangling two-qubit gate) is the necessary requirement for executing a quantum algorithm. In one-way quantum computation all we need is a series of single-qubit measurements on a sufficiently large cluster state in order to simulate any quantum circuit. The number of measurements required is polynomial, and sometimes even linear, in the number of gates required for the original circuit.

We shall now show how single-qubit gates can be accomplished using the cluster state model. For this purpose we introduce the concept of one-bit teleportation [85], which is represented by a simple quantum circuit, and demonstrate how we can map these circuits to measurements on the cluster state.

Consider the following circuit:

$$
\begin{array}{c}
|\psi\rangle \quad \bullet \quad \boxed{H} \quad \boxed{\measuredangle} \quad \text{``}m\text{''} \\
|+\rangle \quad \bullet \quad \text{——————} \quad X^m H |\psi\rangle
\end{array}
\tag{7.6}
$$

where X is the Pauli-X operator whose action can be defined as: $X|+\rangle = |+\rangle$ and $X|-\rangle = -|-\rangle$. Here, the metre represents a measurement in the computational basis where the measurement outcome "m" can take on the value 0 or 1. We can easily see why the result is true. Let the first
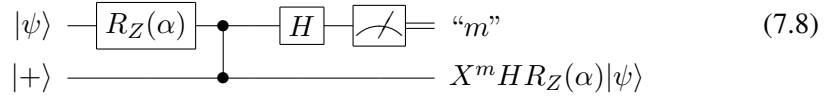
physical qubit be in the state $|\psi\rangle = a|0\rangle + b|1\rangle$ while the second physical qubit is in the $|+\rangle$ state. The CZ gate acts on both input qubits resulting in the entangled state $a|0\rangle|+\rangle + b|1\rangle|-\rangle$. The Hadamard then acts on the first qubit so that the entangled state becomes $a|+\rangle|+\rangle + b|-\rangle|-\rangle$. In this form we note that the state is fully symmetric — we cannot identify which physical qubit possesses the logical input state $|\psi\rangle$. We say that the logical state is encoded in both qubits, i.e. it is shared between both physical qubits. This process is referred to as redundant encoding.

We can express the final state differently as:

$$\begin{aligned}
|\Psi_{\text{out}}\rangle &= \tfrac{1}{\sqrt{2}}(|0\rangle \otimes (a|+\rangle + b|-\rangle) + |1\rangle \otimes (a|+\rangle - b|-\rangle)) \\
&= \tfrac{1}{\sqrt{2}}(|0\rangle \otimes H|\psi\rangle + |1\rangle \otimes XH|\psi\rangle).
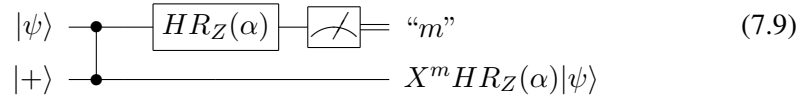\end{aligned} \tag{7.7}$$

Now if we, for example, project the first qubit onto the state $|0\rangle$ then the measurement outcome is $m = 0$ and the second qubit collapses onto the state $H|\psi\rangle$. Hence, we can see that the state left on the second qubit after a measurement is $X^m H|\psi\rangle$ which means that the state that was initially on the first qubit is mapped onto the second qubit after undergoing some unitary operations.

Now let us add a single-qubit rotation around the z-axis, $R_Z(\alpha)$, to the state of first qubit:

$$
\begin{array}{c}
|\psi\rangle \!-\!\boxed{R_Z(\alpha)}\!-\!\bullet\!-\!\boxed{H}\!-\!\measuredangle\!\!=\! \text{``}m\text{''} \\
|+\rangle \!-\!\!-\!\!-\!\!-\!\bullet\!-\!\!-\!\!-\!\!-\! X^m H R_Z(\alpha)|\psi\rangle
\end{array}
\tag{7.8}
$$

At the moment this circuit represents a teleportation of the already rotated state $R_Z(\alpha)|\psi\rangle$. However, the $R_Z(\alpha)$ rotation commutes with the CZ gate so we can interchange the order of these two gates:

$$
\begin{array}{c}
|\psi\rangle \!-\!\bullet\!-\!\boxed{HR_Z(\alpha)}\!-\!\measuredangle\!\!=\! \text{``}m\text{''} \\
|+\rangle \!-\!\bullet\!-\!\!-\!\!-\!\!-\!\!-\! X^m H R_Z(\alpha)|\psi\rangle
\end{array}
\tag{7.9}
$$

We note here that the measurement in the computational basis $\{|0\rangle, |1\rangle\}$ is preceded by the unitaries $R_Z(\alpha)$ and H. Let us write $U = R_Z(\alpha)H$, then this process can be represented as

$$U|m\rangle\langle m|U^\dagger, \tag{7.10}$$

where $m = \{0, 1\}$. This is equivalent to projecting onto the state $U^\dagger|m\rangle$. We can, therefore, absorb the unitaries into the measurement and redefine the measurement basis as

$$\begin{aligned}
B(\alpha) &= R_Z^\dagger(\alpha)H\{|0\rangle, |1\rangle\} \\
&= R_Z(-\alpha)\{|+\rangle, |-\rangle\} \\
&= \{|+\alpha\rangle, |-\alpha\rangle\} \quad \text{where} \quad |\pm\alpha\rangle = \tfrac{1}{\sqrt{2}}\left(|0\rangle \pm e^{i\alpha}|1\rangle\right).
\end{aligned} \tag{7.11}$$

This basis corresponds to a measurement along an axis defined by α on the equator of the Bloch sphere.

If we take $a = b = \frac{1}{\sqrt{2}}$ then $|\psi\rangle \equiv |+\rangle$ and our circuit is:

$$
\begin{array}{c}
|+\rangle \;\bullet\; B(-\alpha) \\
|+\rangle \;\bullet\; \qquad X^m H R_Z(\alpha)|+\rangle
\end{array}
\tag{7.12}
$$

This circuit demonstrates precisely how the one-way quantum computing model works! The highlighted box shows that the two qubits are prepared in the $|+\rangle$ state and then entangled by the CZ gate which is, of course, our two-qubit cluster state. Measuring the first qubit in the basis $B_1(-\alpha)$ implements the single-qubit gate $HR_Z(\alpha)$ on the input state $|+\rangle$, with the result being transferred to the second qubit. Although we measure the first qubit, no quantum information is lost. Instead, the input state of the first qubit is teleported to the second qubit after undergoing the unitary operations $R_Z(\alpha)$ and H. The Pauli-X operator is called a byproduct operator because it occurs as a result of the randomness which is intrinsic to measurement-based protocols.

We can represent one-bit teleportation circuit (Eq. (7.12)) in graphical notation as

$$
\underset{B(-\alpha)}{\overset{1}{\bigcirc}} \text{———} \overset{2}{\bigcirc}
\tag{7.13}
$$

where we have conveniently specified the measurement basis for qubit 1 within the corresponding vertex. The circuit for the logical qubit is

$$
|+\rangle \;\text{———}\; \boxed{HR_Z(\alpha)} \;\text{———}
\tag{7.14}
$$

It is essential to observe that there is a clear distinction between the *physical qubits* which make up the cluster state and on which physical measurements are performed and the *logical qubits* on which the actual computation is carried out. The physical qubits of the cluster state are the entanglement resource for the computation while the logical qubits represent the quantum information that is processed during the computations. The entire information of the logical input state is initially stored in the multi-partite correlations of the cluster, with the individual physical qubits being completely undefined and, therefore, carrying no information about the input state at all [81]. This can be explained as follows. The state of the logical qubit can be written onto the first physical qubit before the entanglement process. For generic cluster states the physical qubits are prepared in the $|+\rangle$ state and then entangled, so the state of the logical qubit is $|+\rangle_L$, where the subscript L is used for clarity. After entanglement of neighbouring physical qubits, the logical state becomes encoded into the correlations between the physical qubits — we can say that the logical state is actually distributed throughout the cluster. Note that each one-dimensional linear cluster can process only a single logical qubit.

Arbitrary single-qubit rotations

According to Euler's theorem an arbitrary single-qubit rotation can be decomposed as a product of three rotations — the first and third around one axis and the second around an axis orthogonal to the other. If we concatenate four copies of the $HR_Z(\alpha)$ gates we can obtain this rotation

$$R(\theta) = HR_Z(\gamma)HR_Z(\beta)HR_Z(\alpha)HR_Z(0) \tag{7.15}$$
$$= R_X(\gamma)R_Z(\beta)R_X(\alpha), \tag{7.16}$$

where we have used $HR_ZH = R_X$. An arbitrary rotation of a logical qubit $|+\rangle_L$ can, therefore, be represented in circuit notation as:

$$|+\rangle_L \quad \boxed{H} \boxed{HR_Z(\alpha)} \boxed{HR_Z(\beta)} \boxed{HR_Z(\gamma)} \tag{7.17}$$

This circuit can be implemented by four single-qubit measurements on a sufficiently large cluster state,

$$\tag{7.18}$$
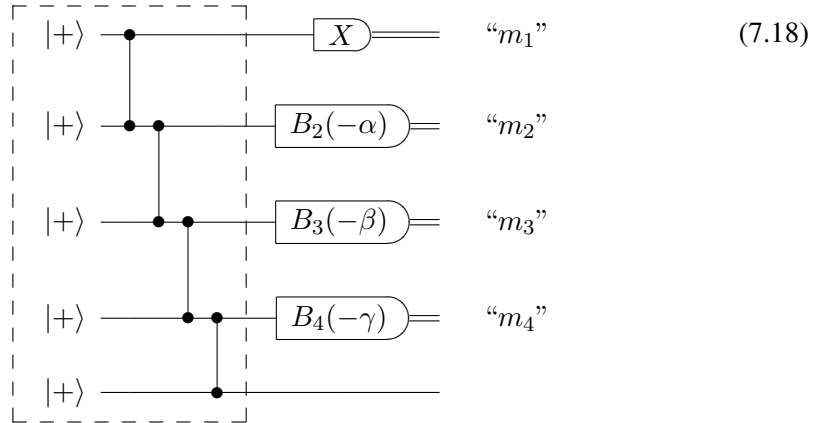
where, for brevity, we denote a measurement in the basis $B(0)$ with an X because it is actually a measurement of the Pauli-X observable; and the highlighted box represents a five-qubit cluster state. This measurement sequence results in the output state

$$|\Psi_{\text{out}}\rangle = (X^{m_4}HR_Z(\gamma))(X^{m_3}HR_Z(\beta))(X^{m_2}HR_Z(\alpha))(X^{m_1}H)|+\rangle_L. \tag{7.19}$$

If we consider the case $m_1 = m_2 = m_3 = m_4 = 0$, which is the case when we project onto the desired state and not to its antipodal state on the Bloch sphere, then we obtain the desired arbitrary single-qubit rotation (Eq. (7.15)). However, in practice we have to consider the cases when the measurement outcomes m_j are other than zero since measurements are inherently probabilistic. In these cases we have to find a way to discard the byproduct operators X. The way that we deal with the byproduct operators is to permute them to the end of the computation, before the measurement in the computational basis, where these errors can be corrected for by classical post-processing. This process will be discussed in detail in the next section.

7.3 Feed-forward and adaptive measurements

Let us first note that, mathematically, single qubit transformations are described by the group of $SU(2)$ matrices. The generators of $SU(2)$ are the Pauli matrices. We have seen that an arbitrary $SU(2)$ rotation of a logical state $|+\rangle_L$ can be preformed using a five-qubit cluster state. Specific single-qubit measurements on the first four qubits transforms the logical state with the final result being left on the fifth qubit where it is read out in the computational basis. The byproduct operators that occur as a result of the randomness of each measurement can be dealt with by propagating these operators to the end of the computation, before read-out of the final results. Here they never need to be physically applied, but can be accounted for by interpreting the final measurement outcomes. Therefore, these known errors are harmless and can be corrected by classical computation.

A Pauli-Z operator adds only a phase to the computational basis elements. When we carry out a measurement we cannot identify a phase so we can disregard the effect of this operator. In quantum information terminology, Pauli-Z operations commute with measurements and thus have no effect on their outcome. Any Pauli-X operation simply flips the measurement result, and thus can be corrected via classical post-processing. For example, if a Pauli-X operation occurs before a measurement in the computational basis then the measurement outcome $m = 0$ will now correspond to the state $X|0\rangle = |1\rangle$. The main point to note is that we know which Paulis will appear at the end of the computation so we do not have to correct for them, instead we reinterpret the measurement results. Therefore, our aim now is to transfer all the Paulis to the end of the circuit where they do not matter any more. In general interchanging the order of unitary operations is not trivial, but with Paulis it is easy especially when propagating these operations past Clifford group unitaries and single-qubit rotations, as we will now see.

Let us first show how we can commute the Pauli byproduct operators through the rotation gates. The Pauli group P consists of the operators X, Y, Z and $\mathbb{I}$. Pauli operators either commute ($P_1 P_2 = P_2 P_1$) or anti-commute ($P_1 P_2 = -P_2 P_1$). The rotation operator can be represented as $R_p = \exp\left[-i\frac{\theta}{2}p\right]$ where p is an element of the Pauli group. Therefore, if p and P anti-commute,

$$
\begin{aligned}
R_p(\alpha)P_i = \exp\left[-i\frac{\theta}{2}p\right]P_i &= \left[\cos\left(\frac{\alpha}{2}\right)\mathbb{I} - i\sin\left(\frac{\alpha}{2}\right)p\right]P_i \\
&= P_i\left[\cos\left(\frac{\alpha}{2}\right)\mathbb{I} + i\sin\left(\frac{\alpha}{2}\right)p\right] \\
&= P_i R_p(-\alpha).
\end{aligned} \tag{7.20}
$$

Clearly, after the order has been swapped, P_i remains unchanged while the angle of the rotation gate is now reversed. If p and P_i commute, the two can be interchanged without altering either one, i.e. $R_p(\alpha)P_i = P_i R_p(\alpha)$.

As well as rotation gates, we need to consider gates like the Hadamard (H), controlled-NOT (CNOT) and CZ. These belong to a group known as the *Clifford group*. Although it bears his name, this group was not discovered by William Clifford and is not related to the field of Clifford algebra. The Clifford group, in quantum information, is the set of unitaries which map Pauli operations to other Pauli operations under conjugation. In graph theory this group would be referred

to as the normaliser of the Pauli group. For further information about the Clifford group and its properties, we would refer the reader to [86, 87]. For a Clifford operator C and Paulis P_1 and P_2 we have the relation

$$CP_1C^\dagger = P_2. \tag{7.21}$$

Multiplying this equation from the right by C gives us a permutation relation

$$CP_1 = P_2C. \tag{7.22}$$

Thus, on interchanging the order of a Clifford gate and a Pauli byproduct operator, the Clifford gate remains invariant while the Pauli operator remains a (possibly different) Pauli [84]. It is always possible to permute a Pauli operation past a Clifford gate, provided that we conjugate it

$$CP|\psi\rangle \equiv C(CPC^\dagger)|\psi\rangle. \tag{7.23}$$

A specific example of this process would be

$$ZH|\psi\rangle = H(HZH^\dagger)|\psi\rangle$$
$$= HX|\psi\rangle. \tag{7.24}$$

We can summarise the two rules for permutating Pauli byproduct operators as follows:

1. The rotation gate changes to a rotation about the inverse angle while the Pauli byproduct operator remains unchanged on permutation past the rotation gate (if the Pauli and the rotation operator anti-commute). For example, we could have $R_Z(\alpha)X = XR_Z(-\alpha)$ and $R_Z(\alpha)Z = ZR_Z(\alpha)$.
2. The Clifford gate remains the same while the Pauli byproduct operator is conjugated on permutation past the Clifford gate. For example, $HX = ZH$ and $HZ = XH$.

If we now apply these rules to Eq. (7.19) we find that it can be written as

$$|\Psi_{\text{out}}\rangle = X^{m_2+m_4}Z^{m_1+m_3}R_X((-1)^{m_1+m_3}\gamma)R_Z((-1)^{m_2}\beta)R_X((-1)^{m_1}\alpha)|+\rangle_L, \tag{7.25}$$

where we, have once again, used the relation $HR_ZH = R_X$. In circuit notation this is

$$|+\rangle_L \;-\; \boxed{R_X((-1)^{m_1}\alpha)} \;-\; \boxed{R_Z((-1)^{m_2}\beta)} \;-\; \boxed{R_X((-1)^{m_1+m_3}\gamma)} \;-\; \boxed{Z^{m_1+m_3}} \;-\; \boxed{X^{m_2+m_4}} \;-$$

Here, the measurement bases depend on the preceding measurement results, so if we carry out each one sequentially and choose $\phi_1 = (-1)^{m_1}\alpha$, $\phi_2 = (-1)^{m_2}\beta$ and $\phi_3 = (-1)^{m_1+m_3}\gamma$, we can deterministically obtain the rotation given in Eq. (7.15). In graphical notation this measurement process is represented as
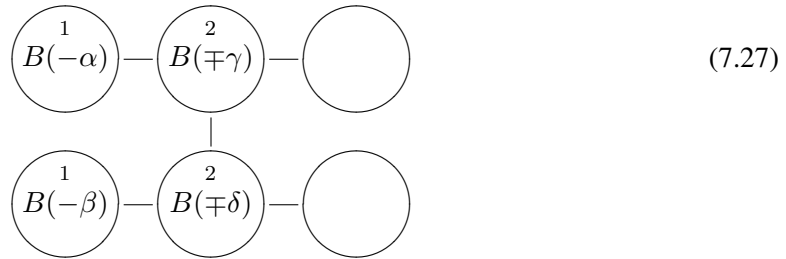

$$\tag{7.26}$$

Adapting measurement bases based on the outcome of previous measurements is called feed-forward. This is a generic feature of one-way quantum computation, occurring for all but the Clifford group of gates, as we will see in Section 7.5. This dependency means that there is a minimum number of time-steps in which any one-way quantum computation can be implemented. Deterministic computation is possible as long as measurements are correctly ordered and adaptive.
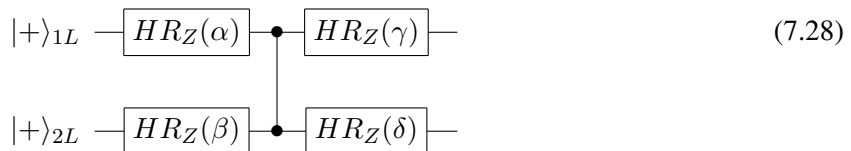
7.4   Two-qubit gates

So far we have seen how single-qubit gates can be implemented in one-way quantum computation using a specific sequence of measurements on a linear cluster state. However, for universal quantum computation, entangling two-qubit gates are also necessary. One such gate is the CZ gate. Implementing the CZ gate is particularly simple since this two-qubit gate is built into the one-way quantum computing model. Adding a CZ gate to the circuit for the logical qubits is equivalent to adding a link or bond between horizontal linear cluster states.

For example, consider the following cluster state [83],

$$
\begin{array}{ccc}
\overset{1}{B(-\alpha)} - \overset{2}{B(\mp\gamma)} - \bigcirc \\[2mm]
\overset{1}{B(-\beta)} - \overset{2}{B(\mp\delta)} - \bigcirc
\end{array}
\tag{7.27}
$$

The labelling indicates the time-ordering of measurements. This means that the qubits labelled 1 must be measured before qubits labelled 2 because the measurement bases of the second set of qubits have to be adapted based on the earlier measurements, as we have seen in the previous section. The qubits with the same label can be measured either sequentially or simultaneously. The circuit that this cluster implements is

$$
\begin{array}{c}
|+\rangle_{1L} - \boxed{HR_Z(\alpha)} - \bullet - \boxed{HR_Z(\gamma)} - \\[2mm]
|+\rangle_{2L} - \boxed{HR_Z(\beta)} - \bullet - \boxed{HR_Z(\delta)} -
\end{array}
\tag{7.28}
$$

The CZ gate that appears in the circuit is implemented by the bond between the two horizontal linear clusters, i.e. the bond between the qubits labelled 2. Remember that each horizontal cluster state represents a single logical qubit. We apply the CZ gate as if we are using the circuit model. We can explain the implementation of the gate as follows. First measure qubits labelled 1 in the basis $B(-\alpha)$ and $B(-\beta)$. This implements the gates $HR_Z(\alpha)$ and $HR_Z(\beta)$ (modulo Pauli errors) on logical qubits 1 and 2, respectively. The result is then teleported to the physical qubits labelled 2. At this point we apply a CZ gate to these physical qubits which is equivalent to applying it to the logical qubits. However, because the measurements and the CZ operations in the sequence

commute, we could have done this CZ operation at the beginning of the computation. This means that if we want to implement a circuit with a CZ gate between logical qubits all we need to do is to create a cluster state with an extra bond between the horizontal clusters.

The CZ gate is part of the Clifford group gate so we can propagate errors the byproduct operators past these as described in the previous section. Consider the output of the circuit (Eq. (7.28))

$$|\Psi_{\text{out}}\rangle = [H_1 R_{Z_1}(\gamma) \otimes H_2 R_{Z_2}(\delta)]CZ[H_1 R_{Z_1}(\alpha) \otimes H_2 R_{Z_2}(\beta)]|+\rangle_{1L}|+\rangle_{2L} \qquad (7.29)$$

The permutation rules in this case are:

$$CZ[\mathbb{I}_1 \otimes X_2] = [Z_1 \otimes X_2]CZ \qquad (7.30)$$
$$CZ[\mathbb{I}_1 \otimes Z_2] = [\mathbb{I}_1 \otimes Z_2]CZ \qquad (7.31)$$

7.5 Clifford Group Gates

Let us study the implementation of three Clifford gates — the Hadamard, the $\pi/2$-phase and the CNOT gates, in the one-way quantum computation model. The Hadamard- and $\pi/2$-phase gates are special with regard to the measurements which are performed on the cluster to realise these gates, compared to the arbitrary single-qubit rotations [88]. If we look at Eq. (7.15) we note that there are specific Euler angles with which we can execute any $SU(2)$ transformation. Depending on the byproduct operator on the input side, the Euler angles that implement a Hadamard-gate are given by $\alpha = \pm\pi/2$, $\beta = \pm\pi/2$ and $\gamma = \pm\pi/2$ while achieving a $\pi/2$-phase gate requires $\alpha = 0$, $\beta = \pm\pi/2$ and $\gamma = 0$. The measurement bases $B(\alpha)$ and $B(-\alpha)$ coincide for the angles 0 and $\pm\pi/2$. The measurement basis $B(0)$ corresponds to a measurement of the Pauli-X observable and the measurement basis $B(\pm\pi/2)$ corresponds to a measurement of Pauli-Y. A change of the measurement angle from $\pi/2$ to $-\pi/2$ has only the effect of interchanging the two states of the measurement basis in Eq. (7.11), but it does not change the basis itself. These measurements can, therefore, be realised simultaneously since there is no need for adaptive measurements. The error operators are propagated to the end as usual so that we can implement the Hadamard gate with byproduct operators

$$U_{\text{byprod}} = X^{m_1+m_3+m_4} Z^{m_2+m_3}, \qquad (7.32)$$

and for the $\pi/2$-phase gate with byproduct operators

$$U_{\text{byprod}} = X^{m_2+m_4} Z^{m_1+m_2+m_3+1}. \qquad (7.33)$$

The analysis for the CNOT gate is more complicated but we notice from Fig. 7.2 that all Clifford group gates require only Pauli measurements. This is related to the fact the Clifford group maps Pauli operators to other Pauli operators.

7.6 Pauli-Z measurements

Generally, when employing the one-way quantum computation model, we construct a square lattice and we need to be able to etch the required cluster state out of it. For example, we might have

the cluster state

$$(7.34)$$

and we want to obtain the state

$$(7.35)$$

We can do this by measuring the central qubit in the computational basis, i.e. by a Pauli-Z measurement. This serves to remove the edges attached to the central vertex if $m = 0$, which means that this type of measurement effectively disentangles this qubit from the cluster state. If $m = 1$ the Pauli-Z correction is applied to all the qubits that were initially entangled to this qubit. As usual, we can propagate these Pauli errors through to the end of the computation according to the rules we have mentioned in Section 7.3.
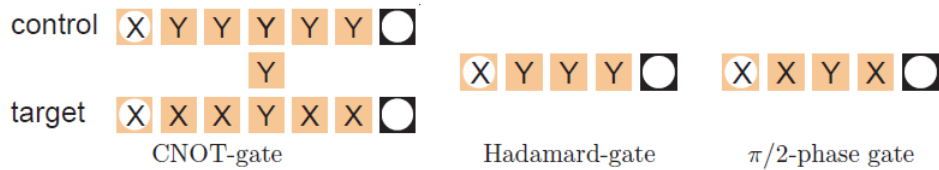
Figure 7.2: Realisation of Clifford group gates in the one-way model: the CNOT, Hadamard gate and $\pi/2$-phase gates. In this graphical notation each square denotes a physical qubit, where the input and output qubits are represented by a white circle within the square. The measurement bases are specified within the square. Note that Clifford group gates require only Pauli measurements [89].

The reason why the Z measurements merely delete the qubits from the cluster is as follows. The Z measurements commute with the CZ operations since they are diagonal in the same basis, therefore we can interchange the order of these operations. A Z measurement is represented by the projector $|0\rangle\langle 0|$. Applying the Z measurement before the CZ means that when forming the cluster state, we would prepare the specific qubit in the $|0\rangle$ state instead of $|+\rangle$ before entangling all the qubits. Of course, if a CZ gate is applied to a $|0\rangle$ state (as the target or control qubit) there is no change to the state, so it is equivalent to not entangling it at all. If we prepare the $|1\rangle$ state instead all the other qubits get a Z flip. Thus, we see that the Z measurement simply deletes the qubit.

Measuring certain qubits in a square lattice in the Z basis leaves a set of horizontal wires which can be used to implement single qubit gates, and connected regions which give us our entangling two-qubit gates. In this kind of initial cluster state the single-qubit wires turn out to be away from each other and not immediately adjacent (cf. Fig. 7.3). They are connected by a bridge between them with at least one extra qubit. Pauli measurements on the qubits that make up the bridge (i.e.

the vertical line of qubits connecting two horizontal wires) allows us to implement a two-qubit gate, as seen for the CNOT gate (Fig. 7.2).

7.7 Summary of one-way quantum computing

One-way quantum computing can be summarised as follows:

1. Create a sufficiently large cluster state by preparing qubits in the $|+\rangle$ state and entangling the qubits with the CZ operation.
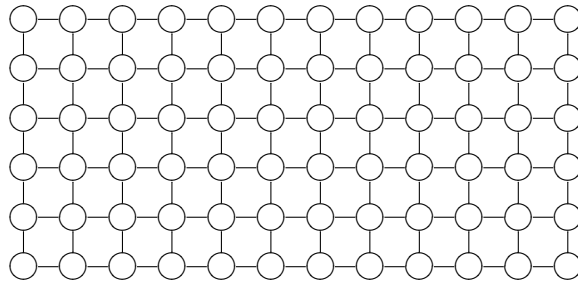
2. Remove unwanted qubits by performing Pauli-Z measurements, thereby shaping the quantum circuit. This etching process can be done simultaneously since Z measurements commute with each other.
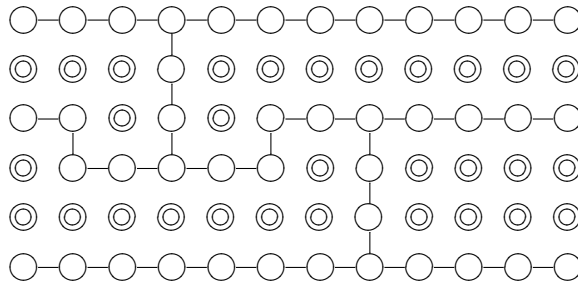
3. Each horizontal line represents a logical qubit. A projective measurement in the $B(-\alpha)$ basis on a physical qubit implement the $HR_Z(\alpha)$ gate on the corresponding logical qubit. Feed-forward of each measurement result is done to adapt future measurement settings and Pauli byproduct operators are commuted to the end of the computation. Remember that if Pauli measurements on a set of qubits are required for the computation, then these can be done simultaneously since they are not adaptive. Two-qubit gates are implemented by vertical bonds connecting the one-dimensional (horizontal) cluster chains and are thus inherent in the cluster (see Fig. 7.3).

4. Finally, the output is stored on the remaining qubits of the cluster state and the result of the computation is read out in the computational basis.

Figure 7.3: Measurements on single qubits in the cluster state allow for the processing of information. Vertical arrows indicate measurements of the Pauli-X observable while tilted arrows refer to measurements in the equatorial (X-Y) plane of the Bloch sphere [6].

Note that with a large enough cluster state and a means in which to carry out single-qubit measurements, any algorithm can be computed. Since the entanglement (which is usually the most difficult part of the quantum computation) can be done beforehand and single-qubit me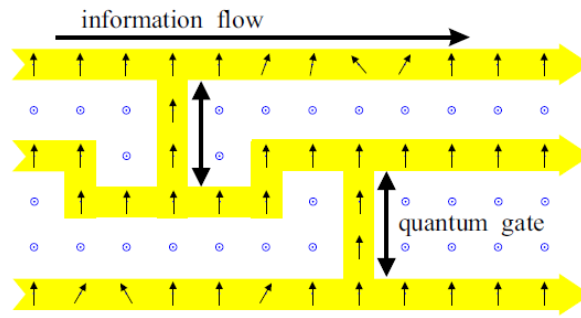asurements are usually very simple to implement this model for quantum computation is efficient and scalable. A remarkable and surprising feature of this model is that the etching with Z measurement and all other Pauli measurements (which are required for Clifford group gates), including the measurement of the output qubits in the computational basis, can be performed simultaneously at the beginning of the computation, after which all other measurements can be executed.

Cluster states that implement the single- or two-qubit gates can be considered to be the building blocks of the one-way quantum computer and these gates can be concatenated. Suppose we want to apply the $SU(2)$ unitaries U_1 and U_2 to the input state $|+\rangle_L$. To implement U_1 we apply single-qubit measurements on four of the five qubits in a linear cluster state (as discussed in Section 6.2). The output is then left on the fifth qubit and this becomes the input qubit (qubit 1 in Eq. (7.26)) for the next measurement sequence, when it is entangled with four other qubits to make up another five-qubit linear cluster. Since all measurements in first sequence commute with the entangling CZ operation we can instead apply all entangling operations (for both clusters) first, to obtain a longer initial cluster state before measuring any of the qubits.

One-way quantum computation differs substantially from the conventional unitary model of quantum computing and these differences have led to new insights into requirement for quantum computation [7]. It also challenges our conventional understanding of quantum measurement as a process that produces irreducible randomness and that inherently destroys quantum coherence [81]. There are many other interesting features of one-way quantum computing, but a detailed discussion is out of the scope of this dissertation. We, therefore, refer the reader to [83, 84, 89, 88, 90] where one will no doubt be fascinated by the intricate workings and diverse applications of this unique model.

Chapter 8

One-way quantum computing with polarisation

In 2005, the first proof-of-principle experiment to demonstrate the concept of one-way quantum computing was implemented by Walther *et al.* [7]. The non-linear process of spontaneous parametric down-conversion (SPDC) and a specific combination of linear optical elements was employed in order to generate a four-qubit spatial-mode and polarisation-entangled cluster state. Simple projective measurements on the cluster state were accomplished through the use of an appropriate arrangement of polarisers and birefringent wave plates. The experimental set-up was effective for the implementation of a set of universal quantum gates as well as the a simple quantum algorithm — the two-qubit Grover search algorithm. The flexibility and simplicity of the one-way model was thus demonstrated since a four-partite cluster state was sufficient to display these features of quantum computation.

In this chapter we will explain the basic results of the proof-of-principle experiment that was carried out by Walther *et al.*. Section 8.1 provides some preliminary information about the optical elements used in the experiment to manipulate the polarisation states of photons. Section 8.2 presents a relatively detailed explanation of the construction of the optical cluster state which is the entanglement resource for one-way quantum computing. Section 8.3 shows how the produced cluster state can be transformed into different one- and two-dimensional four-qubit cluster states. Section 8.4 explains how single-qubit measurements (which are an essential component of one-way quantum computing) are preformed on the polarisation qubits. The implementation of a universal set of quantum gates and the demonstration of the two-qubit Grover algorithm are discussed in Sections 8.5 and 8.6, respectively.

8.1 Preliminaries: Optical elements for polarisation

This proof-of-principle experiment uses photonic qubits which carry quantum information in their polarisation degree of freedom. Therefore, it would serve us well to have a brief outline of the optical elements that are used to manipulate the polarisation degree of single photons in the experiment.

8.1.1 Wave Plates

A wave plate is an optical retardation component which alters the polarisation state of a photon (or light, in general) by creating a phase shift between its two perpendicular components, $|H\rangle$ and $|V\rangle$. The action of a half-wave plate (HWP) and a quarter-wave plate (QWP) can be written in the form of unitary matrices as:

$$U_{\text{HWP}}\left(\tfrac{\theta}{2}\right) = e^{i\frac{\pi}{2}}\left(\begin{array}{cc} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{array}\right), \tag{8.1}$$

$$\text{and} \quad U_{\text{QWP}}\left(\theta\right) = \tfrac{1}{\sqrt{2}}\left(\begin{array}{cc} 1 + i\cos\theta & i\sin\theta \\ i\sin\theta & 1 - i\cos\theta \end{array}\right). \tag{8.2}$$

8.1.2 Polarisers

A polariser transmits photons polarised along its transmission axis and absorbs photons perpendicular to it. The polarisation state of the photon after traversing the polariser is aligned with the axis of the polariser. Thus the polariser projects the state of a photon onto the state aligned with its transmission axis. The action of the polariser can be represented by the projection operator of quantum mechanics. For example, if the polariser is aligned horizontally then

$$\hat{\mathbf{P}}_H = |H\rangle\langle H| = \left(\begin{array}{cc} 1 & 0 \\ 0 & 0 \end{array}\right), \tag{8.3}$$

and if it was aligned along the vertical axis

$$\hat{\mathbf{P}}_V = |V\rangle\langle V| = \left(\begin{array}{cc} 0 & 0 \\ 0 & 1 \end{array}\right). \tag{8.4}$$

8.1.3 Polarising Beam Splitters

The polarising beam splitter (PBS) transmits horizontally polarised photons $|H\rangle$ and reflects vertically polarised photons $|V\rangle$. In order to obtain a coincidence click (or a simultaneous detection of a *single* photon) in both detectors at the output ports we would either have to have: (a) two orthogonally polarised photons entering from the same input port (as in Fig. 8.1) or (b) two photons which have the same polarisation entering from the different input ports. The PBS is used in the experimental set-up to post-select the cases where coincidence clicks are obtained, as will be explained in Section 8.2.
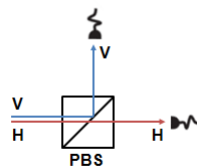

Figure 8.1: The PBS reflects vertically polarised light while transmitting horizontally polarised light.

8.1.4 Spontaneous Parametric down-conversion

In the experiment non-collinear type-II SPDC (cf. Chapter 4), which emits entangled photons into spatial modes a and b, is used to generate entangled photons. The Hamiltonian for the process is

$$\hat{\mathbf{H}} = A_0 \left(\hat{\mathbf{a}}_H^\dagger \hat{\mathbf{b}}_V^\dagger + e^{i\alpha} \hat{\mathbf{a}}_V^\dagger \hat{\mathbf{b}}_H^\dagger \right) + \text{h.c.}, \tag{8.5}$$

where the relative phase α can be selected by tilting the compensation crystals, or the non-linear crystal itself. Using the Taylor expansion, the quantum mechanical state that is emitted by the non-linear crystal becomes

$$\begin{aligned}
|\psi\rangle &= Z e^{-i\hat{\mathbf{H}}} |0\rangle \\
&= Z \exp\left(-iA_0 \left(\hat{\mathbf{a}}_H^\dagger \hat{\mathbf{b}}_V^\dagger + \hat{\mathbf{a}}_V^\dagger \hat{\mathbf{b}}_H^\dagger \right) \right) |0\rangle \\
&= Z \left(1 - \underbrace{iA_0 \left(\hat{\mathbf{a}}_H^\dagger \hat{\mathbf{b}}_V^\dagger + \hat{\mathbf{a}}_V^\dagger \hat{\mathbf{b}}_H^\dagger \right)}_{\text{2 photon state}} - \underbrace{\frac{A_0^2}{2} \left(\hat{\mathbf{a}}_H^\dagger \hat{\mathbf{b}}_V^\dagger + \hat{\mathbf{a}}_V^\dagger \hat{\mathbf{b}}_H^\dagger \right)^2}_{\text{4 photon state}} + \dots \right) |0\rangle, \tag{8.6}
\end{aligned}$$

where Z is a normalisation constant and $|0\rangle$ is the vacuum state. The four highly entangled two-photon states that can be produced in SPDC are known as the *Bell states* and are given by:

$$|\Phi^\pm\rangle = \tfrac{1}{\sqrt{2}} \left(|H\rangle|H\rangle \pm |V\rangle|V\rangle \right), \tag{8.7}$$

$$|\Psi^\pm\rangle = \tfrac{1}{\sqrt{2}} \left(|H\rangle|V\rangle \pm |V\rangle|H\rangle \right). \tag{8.8}$$

8.2 Production of the Cluster State

One of the main difficulties that arises in one-way quantum computation is the preparation of an appropriate cluster state. Projective measurements on the cluster state can be implemented with relative ease, as will be shown later. The target state for this proof-of-principle experiment is given by

$$\begin{aligned}
|\Phi_{\text{cluster}}\rangle = \tfrac{1}{2} \Big(& |H\rangle_1 |H\rangle_2 |H\rangle_3 |H\rangle_4 + |H\rangle_1 |H\rangle_2 |V\rangle_3 |V\rangle_4 \\
& + |V\rangle_1 |V\rangle_2 |H\rangle_3 |H\rangle_4 - |V\rangle_1 |V\rangle_2 |V\rangle_3 |V\rangle_4 \Big), \tag{8.9}
\end{aligned}$$

The set-up for the experiment is show in Fig. 8.2. The principle source of entangled photons in this experiment is non-collinear type-II SPDC. An ultraviolet (UV) pulse passes through a β-barium borate (BBO) crystal which probabilistically emits highly entangled photon pairs into the forward-oriented spatial modes a and b. This pulse is then reflected at the pump mirror and traverses the crystal for a second time, producing entangled pairs in the backward-oriented modes c and d. These photon pairs (which are initially in a separable state) are transformed by a series of optical elements in the experimental set-up and coherently combined in order to create the required highly entangled four-photon cluster state. We will now explain this process in detail.
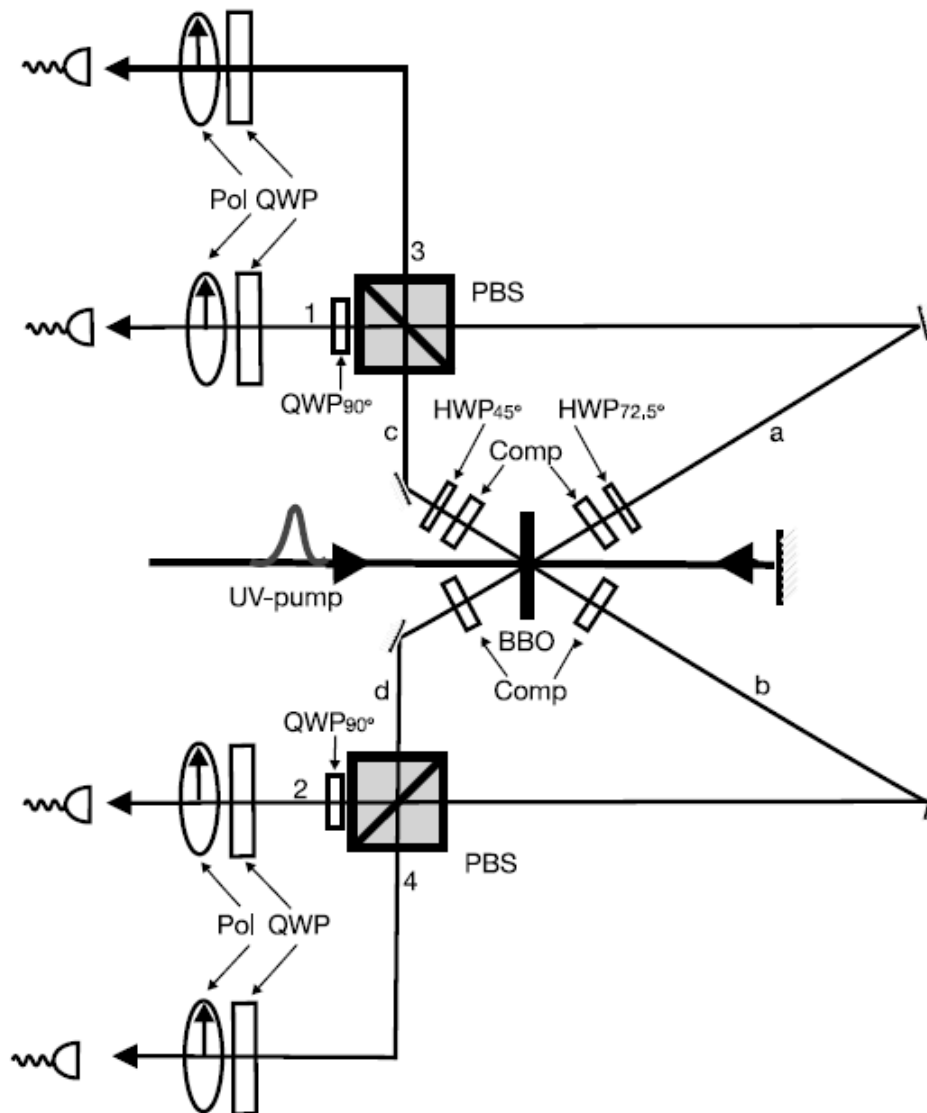
Figure 8.2: Schematic drawing of the set-up used to generate the polarisation-entangled four-photon linear cluster state [7].

Looking at Eq. (8.6) we see that SPDC allows for the probabilistic production of a single photon pair (which is a two-photon state), a double photon pair (which is a four-photon state), and higher orders as well. The possible emission of a single pair in either the forward or backward direction is approximately described by the Hamiltonian [91]

$$\hat{\mathbf{H}} \propto p_{\mathrm{f}} \left(\hat{\mathbf{a}}_H^\dagger \hat{\mathbf{b}}_V^\dagger + e^{i\alpha_{ab}} \hat{\mathbf{a}}_V^\dagger \hat{\mathbf{b}}_H^\dagger \right) + p_{\mathrm{b}} e^{i\Delta\phi} \left(\hat{\mathbf{c}}_H^\dagger \hat{\mathbf{d}}_V^\dagger + e^{i\alpha_{cd}} \hat{\mathbf{c}}_V^\dagger \hat{\mathbf{d}}_H^\dagger \right) + \mathrm{h.c.}, \tag{8.10}$$

where, for example, $\hat{\mathbf{a}}_H^\dagger$ represents the creation operator which creates a horizontally polarised photon in the spatial mode a; p_{f} (p_{b}) denotes the probability for the emission of photons into the modes a and b (c and d); and $\Delta\phi$ is the relative phase between the two emission probabilities which can be set by the pump mirror position. The compensation crystals are then used to counter transverse and longitudinal walk-off which arise as a result of the birefringence of the BBO crystal (cf. Section 4.5.1). The tilt of these crystals allow one to select the relative phases, which are chosen to be $\alpha_{ab} = \pi$ and $\alpha_{cd} = 0$ in this experiment.

The next optical element encountered is the half-wave plates (HWPs) at $45°$ in modes a and c, which is a single HWP for each photon pair. This element rotates the polarisation of a photon by $90°$ ($|H\rangle \leftrightarrow |V\rangle$) altering Eq. (8.10) to give the final Hamiltonian of the process

$$\hat{\mathbf{H}} \propto \left(\hat{\mathbf{a}}_H^\dagger \hat{\mathbf{b}}_H^\dagger - \hat{\mathbf{a}}_V^\dagger \hat{\mathbf{b}}_V^\dagger \right) + p_r e^{i\Delta\phi} \left(\hat{\mathbf{c}}_H^\dagger \hat{\mathbf{d}}_H^\dagger + \hat{\mathbf{c}}_V^\dagger \hat{\mathbf{d}}_V^\dagger \right) + \mathrm{h.c.}, \tag{8.11}$$

where $p_r = p_{\mathrm{b}}/p_{\mathrm{f}}$ denotes the relative probability of emission into the backward direction (c-d) compared to the forward direction (a-b). If we apply this Hamiltonian to the vacuum, we can obtain the two-photon state

$$\hat{\mathbf{H}}|0\rangle = (|H\rangle_a |H\rangle_b - |V\rangle_a |V\rangle_b) + p_r e^{i\Delta\phi}(|H\rangle_c |H\rangle_d + |V\rangle_c |V\rangle_d). \tag{8.12}$$

This state is a superposition of the *Bell states* $|\Phi^-\rangle$ in the forward direction and $|\Phi^+\rangle$ in the backward direction. This two-photon state does not contribute to the formation of cluster state but it is important for alignment purposes [81].

The production of the four-qubit cluster state is, however, based on simultaneous four-photon events, so we will now have to focus on the third term in Eq. (8.6). By applying the Hamiltonian twice to the vacuum we can obtain the four-photon state

$$\hat{\mathbf{H}}^2|0\rangle = \left[\left(\hat{\mathbf{a}}_H^\dagger \hat{\mathbf{b}}_H^\dagger - \hat{\mathbf{a}}_V^\dagger \hat{\mathbf{b}}_V^\dagger \right)^2 + p_r^2 e^{i2\Delta\phi} \left(\hat{\mathbf{c}}_H^\dagger \hat{\mathbf{d}}_H^\dagger + \hat{\mathbf{c}}_V^\dagger \hat{\mathbf{d}}_V^\dagger \right)^2 \right.$$
$$\left. + 2 p_r e^{i\Delta\phi} \left(\hat{\mathbf{a}}_H^\dagger \hat{\mathbf{b}}_H^\dagger - \hat{\mathbf{a}}_V^\dagger \hat{\mathbf{b}}_V^\dagger \right) \left(\hat{\mathbf{c}}_H^\dagger \hat{\mathbf{d}}_H^\dagger + \hat{\mathbf{c}}_V^\dagger \hat{\mathbf{d}}_V^\dagger \right) \right] |0\rangle. \tag{8.13}$$

The first two terms in this equation refer to the case of double-pair emission into the forward and backward modes, respectively, i.e. where two photon pairs are emitted into each of the modes, a-b and c-d. The third term encapsulates the situation where we have the simultaneous emission of a single photon pair into each of the modes a-b and c-d, which occurs with the same probability as the double-pair emission into either set of modes.

Note that, for convenience, we shall neglect the probability of emission p_r and the relative phase $e^{i\Delta\phi}$ for the moment, and include it later in the final equation for the cluster state. We begin by considering the case where we have the simultaneous emission of a single photon pair into each of the modes a-b and c-d (which is the third term in Eq. (8.13)),

$$|\phi_{\text{single}}\rangle = (|H\rangle_a|H\rangle_b - |V\rangle_a|V\rangle_b) \otimes (|H\rangle_c|H\rangle_d + |V\rangle_c|V\rangle_d)$$
$$= |\Phi^-\rangle_{ab}|\Phi^+\rangle_{cd}. \tag{8.14}$$

At the moment there are two individual entangled photon pairs — one in the modes a and b and another in modes c and d. In order to form the required four-photon spatial-mode and polarisation-entangled cluster state, there has to be a mechanism to connect spatial modes a and c and modes b and d. This is done by means of PBSs, which is the next optical element that the state encounters. Note that the cluster state is prepared by means of post-selection: specifically, by considering only the cases where we have the simultaneous emission of four photons (as we are) and then selecting only the cases where we obtain the simultaneous detection of a *single* photon in each of the output ports 1, 2, 3 and 4 of the PBSs (which is called *four-fold coincidences*). Table 8.1 summarises the action of the PBS which channels the polarisation states from the input mode a, b, c, and d into specific output modes according to the rules mentioned in Section 8.1.3. The state after the PBSs is therefore given by

$$|\phi_{\text{single}}\rangle = \tfrac{1}{2}\left(|H\rangle_1|H\rangle_2|H\rangle_3|H\rangle_4 - |V\rangle_1|V\rangle_2|V\rangle_3|V\rangle_4\right), \tag{8.15}$$

where we have neglected the terms that do not contribute to four-fold coincidences. This state gives only two of the four terms that are required for the cluster state.

Table 8.1: The horizontal and vertical components of the photon pairs are channeled from input modes a, b, c and d to output modes 1, 2, 3 and 4 by the PBSs.

Horizontal Components	Vertical Components
$|H\rangle_a \rightarrow |H\rangle_1$	$|V\rangle_a \rightarrow |V\rangle_3$
$|H\rangle_b \rightarrow |H\rangle_2$	$|V\rangle_b \rightarrow |V\rangle_4$
$|H\rangle_c \rightarrow |H\rangle_3$	$|V\rangle_c \rightarrow |V\rangle_1$
$|H\rangle_d \rightarrow |H\rangle_4$	$|V\rangle_d \rightarrow |V\rangle_2$

We now have to consider the case of double-pair emission in each direction, which occurs with equal probability as the single pair emission. The first term in Eq. (8.13) represents the case when we have a double-pair being emitted into the forward direction a-b

$$|\phi_{\text{double}}^{ab}\rangle = |2H\rangle_a|2H\rangle_b + |2V\rangle_a|2V\rangle_b - |1H, 1V\rangle_a|1H, 1V\rangle_b. \tag{8.16}$$

where, for example, $|2H\rangle_a$ refers to two horizontally polarised photons in mode a. Using Table 8.1, we see that the state after the PBSs is given by

$$|\phi_{\text{double}}^{ab}\rangle = -|H\rangle_1|H\rangle_2|V\rangle_3|V\rangle_4, \tag{8.17}$$

where only the third term in Eq. (8.16) contributes to four-fold coincidences. Similarly, the second term in Eq. (8.13) results in the state

$$|\phi_{\text{double}}^{cd}\rangle = |V\rangle_1|V\rangle_2|H\rangle_3|H\rangle_4, \tag{8.18}$$

after the PBSs. The two double-pair emission terms will be in a coherent superposition if we fix the position of the pump mirror so that they arrive at the PBSs simultaneously [81],

$$|\phi_{\text{double}}\rangle = \tfrac{1}{2}\left(-|H\rangle_1|H\rangle_2|V\rangle_3|V\rangle_4 + |V\rangle_1|V\rangle_2|H\rangle_3|H\rangle_4\right). \tag{8.19}$$

This state gives the two terms that are required for the cluster state in addition to the two that occur in Eq. (8.15),

$$\begin{aligned}
|\Phi_{\text{produced}}\rangle = \tfrac{1}{2}\,(&|H\rangle_1|H\rangle_2|H\rangle_3|H\rangle_4 - |H\rangle_1|H\rangle_2|V\rangle_3|V\rangle_4 \\
+\ &|V\rangle_1|V\rangle_2|H\rangle_3|H\rangle_4 - |V\rangle_1|V\rangle_2|V\rangle_3|V\rangle_4).
\end{aligned} \tag{8.20}$$

However, we notice that we have obtained a separable state (i.e. it can be factorised). In order for this state to be entangled the phase of second term in has to be shifted by π. This can be done by using a HWP in one mode (a in the experimental set-up).

The addition of the HWP at an angle $\theta/2$ in mode a transforms the third of term of Eq. (8.16) of as follows

$$\begin{aligned}
|\phi_{\text{double}}^{ab}\rangle = (&-\cos\theta\sin\theta|2H\rangle_a + \cos\theta\sin\theta|2V\rangle_a \\
+\ &\cos^2\theta|1H,1V\rangle_a - \sin^2\theta|1H,1V\rangle_a) \otimes |1H,1V\rangle_b.
\end{aligned} \tag{8.21}$$

The post-selection of four-fold coincidences causes this state to evolve to [7]

$$|\phi_{\text{double}}^{ab}\rangle = -\cos 2\theta|H\rangle_1|H\rangle_2|V\rangle_3|V\rangle_4, \tag{8.22}$$

after the PBSs. So for $\theta > 45°$ we obtain a sign change. However, the HWP in mode a also affects the single photon pair that is emitted into mode a-b, and thus it changes the contribution of the single-pair state (Eq. (8.14)) to the cluster state,

$$|\Phi^-\rangle_{ab} \otimes |\Phi^+\rangle_{cd} \xrightarrow{HWP} (\cos\theta|\Phi^-\rangle_{ab} + \sin\theta|\Psi^+\rangle_{ab}) \otimes |\Phi^+\rangle_{cd}. \tag{8.23}$$

According to the rules of the PBS, only the $|\Phi^-\rangle_{ab}$ part of the photon pair in mode a-b leads to four-fold coincidences and thus contributes to the formation of the cluster state. The $|\Psi^-\rangle_{ab}$ term results in two photons being directed to the same output port of the PBS so it does not affect the cluster state. Therefore, the HWP only reduces the amplitudes of the single-pair contribution, without adding additional terms to the cluster state,

$$|\phi_{\text{single}}\rangle = \tfrac{1}{2}\cos\theta\left(|H\rangle_1|H\rangle_2|H\rangle_3|H\rangle_4 - |V\rangle_1|V\rangle_2|V\rangle_3|V\rangle_4\right). \tag{8.24}$$

Putting all the terms together, including the rates of emission and phases, we obtain the state

$$\begin{aligned}
|\Phi_{\text{cluster}}\rangle = \tfrac{1}{2}\,\Big(&p_r\cos\theta e^{i\Delta\phi}|H\rangle_1|H\rangle_2|H\rangle_3|H\rangle_4 - \cos 2\theta|H\rangle_1|H\rangle_2|V\rangle_3|V\rangle_4 \\
&+p_r^2 e^{i2\Delta\phi}|V\rangle_1|V\rangle_2|H\rangle_3|H\rangle_4 - p_r\cos\theta e^{i\Delta\phi}|V\rangle_1|V\rangle_2|V\rangle_3|V\rangle_4\Big).
\end{aligned} \tag{8.25}$$

In order to obtain the target cluster state (Eq. (8.9)) we need to match the amplitudes and phases of all terms in the above equation. We, therefore, have to tune the parameters θ and p_r. The phase $\Delta\phi$ has to be adjusted to a multiple of 2π and kept fixed during the experiment [81]. A simple calculation then shows that the required values of the parameters are: $p_r = 1/\sqrt{3}$ and $\theta = 54.7°$ for which we will get perfect phase-matching of the prefactors and obtain the target state. The value of p_r is acceptable since typical emission rates of the forward- and backwards-emitted pairs in experiments are $30000 s^{-1}$ and $10000 s^{-1}$ two-photon coincidences, respectively [81]. Note that the extra HWP at $\theta/2 = 27.5°$ can be combined with the HWP at $45°$ in mode a, therefore we just have a single HWP rotated at $72.5°$. The QWPs at $90°$ after the PBSs are used to counter the birefringence effects of these elements.

8.3 Equivalence of cluster states

The experimentally produced cluster state is quite versatile as it can easily be transformed into different four-qubit cluster states. At this point we can define the states $|H\rangle \equiv |0\rangle$, $|V\rangle \equiv |1\rangle$ and $|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$. The produced cluster state is then equivalent to the one-dimensional (horizontal) linear four-qubit cluster

$$|\Phi_{\text{lin}^4}\rangle = \tfrac{1}{2}\Big(|0\rangle_1|+\rangle_2|0\rangle_3|+\rangle_4 + |0\rangle_1|-\rangle_2|1\rangle_3|-\rangle_4$$
$$+ |1\rangle_1|-\rangle_2|0\rangle_3|+\rangle_4 + |1\rangle_1|+\rangle_2|1\rangle_3|+\rangle_4\Big), \qquad (8.26)$$

as well as the two-dimensional horseshoe cluster state, $|\Phi_{\subset 4}\rangle$, after undergoing the unitary operation $H_1 \otimes I_2 \otimes I_3 \otimes H_4$, where $H_i(I_i)$ denotes the Hadamard (Identity) gate on the ith qubit. By measuring the first qubit of $|\Phi_{\text{lin}^4}\rangle$ in the computational basis $\{|0\rangle, |1\rangle\}$, we disentangle it from the rest of the qubits and thus obtain a three-qubit linear cluster state

$$|\Phi_{\text{lin}^3}\rangle = \tfrac{1}{\sqrt{2}}|0\rangle_1 \otimes \Big(|+\rangle_2|0\rangle_3|+\rangle_4 + |-\rangle_2|1\rangle_3|-\rangle_4\Big). \qquad (8.27)$$

The produced cluster state is also the unitary equivalent of the two-dimensional box cluster state which is used for the implementation of the Grover algorithm. To obtain the box cluster state

$$|\Phi_{\text{box}}\rangle = \tfrac{1}{2}\Big(|0\rangle_1|+\rangle_2|0\rangle_3|+\rangle_4 + |0\rangle_1|-\rangle_2|1\rangle_3|-\rangle_4$$
$$+ |1\rangle_1|-\rangle_2|0\rangle_3|-\rangle_4 + |1\rangle_1|+\rangle_2|1\rangle_3|+\rangle_4\Big), \qquad (8.28)$$

we need to perform the Hadamard transformation on each of the four qubits of the produced cluster state, as well as swap (or relabel) qubits 2 and 3. Practically, the Hadamard operation can be implemented using a HWP at $22.5°$.

8.4 Projective Measurements

In order to process information in the one-way quantum computing model we need to be able to perform projective measurements on the cluster state in the basis

$$B_j(\alpha) = \{|+\alpha\rangle_j, |-\alpha\rangle_j\} \text{ with } |\pm\alpha\rangle_j = \tfrac{1}{\sqrt{2}}\left(|0\rangle_j \pm e^{i\alpha}|1\rangle_j\right), \qquad (8.29)$$

where $\alpha \in [0, 2\pi]$ and j denotes the qubit. The outcome s_j of a measurement of physical qubit j is defined as 0 if the measurement outcome is $| + \alpha\rangle_j$, and as 1 if $| - \alpha\rangle_j$ is the outcome. (Note that we use s_j to refer to measurement outcomes, instead of m_j as done in the previous chapter, because this was the notation used by Walther *et al.* in [7].)

These projective measurements on each of the four polarisation qubits in cluster state are easily performed through the use of a QWP and a linear polariser (Pol) followed by a single-photon counting detector in each of the four output ports of the PBSs. The linear polariser allows us to project onto any of the linear polarisation states: horizontal $(|H\rangle)$, vertical $(|V\rangle)$, diagonal $\left(|D\rangle = \frac{1}{\sqrt{2}}(|H\rangle + |V\rangle)\right)$ and anti-diagonal $\left(|A\rangle = \frac{1}{\sqrt{2}}(|H\rangle - |V\rangle)\right)$. For example, if we want to project onto the state $|V\rangle$ we align the polariser along the vertical axis so that it transmits only the vertical polarisation component of the incoming photon. If the detector registers a "click" then it means that only the $|V\rangle$ component of the state before the PBS plays a role — effectively this state can be viewed as if it were projected onto the state $|V\rangle$. For the detection of linear polarised states the linear polariser and QWP are aligned along the same axis.

A QWP (at $45°$) is used to change linearly polarised light to circularly polarised light and vice versa. So if we want to project onto the right-circularly polarised state we use a linear polariser orientated at $45°$ and a QWP at $(45° + 45° = 90°)$ before the polariser. To project onto the circular and elliptically polarised states which are part of the $B(\alpha)$ basis the angle between the fast axis of the QWP and the horizontal must be equal to twice the angle α of the measurement basis while the polariser can be orientated either horizontally or vertically. Table 8.2 summarises the alignment of the optical elements for different projective measurements.

Table 8.2: The orientation angles of the quarter-wave plate (QWP) and the linear polariser (Pol) with the horizontal for different polarisation projective measurements.

Projection State	QWP($°$)	Pol ($°$)
$\lvert H\rangle$	0	0
$\lvert V\rangle$	+90	+90
$\lvert D\rangle = \frac{1}{\sqrt{2}}(\lvert H\rangle + \lvert V\rangle)$	+45	+45
$\lvert A\rangle = \frac{1}{\sqrt{2}}(\lvert H\rangle - \lvert V\rangle)$	−45	−45
$\lvert R\rangle = \frac{1}{\sqrt{2}}(\lvert H\rangle + i\lvert V\rangle)$	+45	+90
$\lvert L\rangle = \frac{1}{\sqrt{2}}(\lvert H\rangle - i\lvert V\rangle)$	−45	+90
$\lvert \pm\alpha\rangle$	2α	0 or 90

8.5 Quantum Gates

Walther *et al.* [7] were able to use their experimentally-produced cluster state to show that a universal set of quantum gates could be implemented using the method of one-way quantum computation. The universal set that they chose consisted of the single-qubit rotation gates $(R(\alpha))$ and

the entangling two-qubit controlled phase (CZ) gate since this set was easiest to realize with their set-up and it was useful for the Grover algorithm, as we will see in Section 8.6. An appropriate series of single-qubit measurements in the basis $B(\alpha)$ on the physical qubits in the cluster allow for the application of different quantum gates on the encoded input state, as shown in Fig. 8.3.
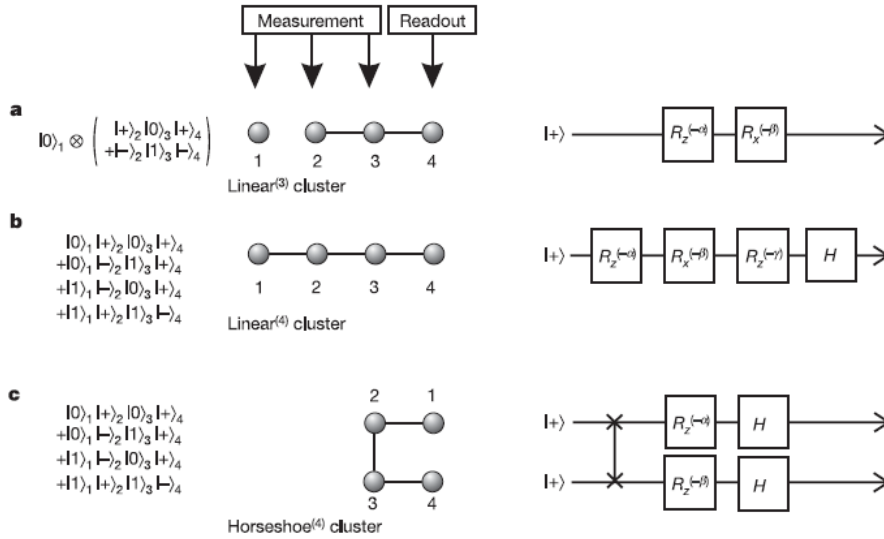

Figure 8.3: Examples of small cluster states and the circuits they implement. For (a) the three-qubit linear cluster and (b) the four-qubit linear cluster, consecutive measurements on qubits 1, 2, and 3 implement a series of single-qubit rotation gates for the encoded input state with the output being available on qubit 4. (c) For the two-dimensional horseshoe cluster state consecutive measurements on physical qubits 2 and 3, implement an entangling CZ gates as well as single-qubit rotation gates on the encoded input state with the output being available on qubits 1 and 4 [7].

Note that one of the principle features of one-way quantum computing is the distinction between the *physical* qubits (in this case the polarisation states of the four photons) which make up the cluster state and on which physical measurements are performed and the *logical* qubits on which the actual computation is carried out. The physical qubits of the cluster state are the entanglement resource for the computation while the logical qubits represent the quantum information that is processed during the computations. It is important to note that the entire information of the logical input state is initially stored in the multi-partite correlations of the cluster, with the individual physical qubits being completely undefined and, therefore, not carrying any information about the input at all [81]. This can be explained as follows. The state of the logical qubit can be written onto the first physical qubit before the entanglement process. In most cases the physical qubit is prepared in the $|+\rangle$ and then entangled, so the state of the logical qubit is $|+\rangle_E$. After entanglement of neighbouring physical qubits, the logical state becomes encoded into the correlations between the physical qubits — we can say that the logical state is actually distributed throughout the cluster. Each one-dimensional linear cluster can process only a single logical qubit.

Every projective measurements on the physical qubits in the cluster transforms the encoded input and then transfers the logical state from one physical qubit to another. Following a precise sequence of measurements, the final output is available on the remaining physical qubits of the cluster where it can be read out. Therefore, it is said that the information is transferred non-locally; information is carried by correlations between the physical qubits [7]. It is interesting to note that while the entanglement between the physical qubits in general decreases as a result of the measurements, the entanglement between logical qubits may increase [7]. This will be made clear with the example of the horseshoe cluster state (Fig. 8.3 (c)). One-way quantum computing consists of only single-qubit projective measurements and no entangling two-qubit gates between physical qubits. Therefore, if all the information was stored on the physical qubits themselves then performing single-qubit measurements on these qubits would destroy the coherence which is essential for quantum computation to proceed [92].

A measurements of a physical qubit in the basis $B(\alpha)$ allows one to implement the rotation gate $R_Z(-\alpha)$ gate followed by the Hadamard operation on the logical qubit — if the encoded input state is $|\Psi_{\text{in}}\rangle = |+\rangle_{1E}$ then after a measurement of the physical qubits $|\Psi_{\text{out}}\rangle = HR_Z(-\alpha)|\Psi_{\text{in}}\rangle$. With this preliminary information we can now understand the implementation of an arbitrary single-qubit rotation with the four-qubit linear cluster state (Fig. 8.3 (b)). Consecutive measurements on the first three physical qubits in the bases $B_1(\alpha)$, $B_2(\beta)$ and $B_3(\gamma)$ results in the output for the encoded (or logical) qubit being

$$
\begin{aligned}
|\Psi_{\text{out}}\rangle &= HR_Z(-\gamma)HR_Z(-\beta)HR_Z(-\alpha)|\Psi_{\text{in}}\rangle \\
&= HR_Z(-\gamma)R_X(-\beta)R_Z(-\alpha)|\Psi_{\text{in}}\rangle,
\end{aligned}
\tag{8.30}
$$

where we have used the relation $HR_ZH = R_X$. This output state is then available on physical qubit 4 where it can be read out. Fig. 8.3 (a) represents a three-qubit cluster state which is formed from the four-qubit linear cluster by measuring the first qubit in the computational basis $\{|0\rangle, |1\rangle\}$. After consecutive measurements of qubits 2 and 3 in the basis $B_2(-\alpha)$ and $B_3(-\alpha)$, respectively, the logical state $|\Psi_{\text{in}}\rangle = |+\rangle_{1E}$ is transformed to

$$
\begin{aligned}
|\Psi_{\text{out}}\rangle &= HR_Z(-\beta)HR_Z(-\alpha)|\Psi_{\text{in}}\rangle \\
&= R_X(-\beta)R_Z(-\alpha)|\Psi_{\text{in}}\rangle.
\end{aligned}
\tag{8.31}
$$

In one-way quantum computing two-qubit gates are implemented via the vertical bonds connecting the one-dimensional (horizontal) cluster and are thus inherent in the cluster. As an example, let us consider the horseshoe cluster state (Fig. 8.3 (c)). If we look at the shape of this cluster we see that is a two-dimensional cluster, i.e. it is made up of two linear horizontal clusters that are connected by a vertical bond. However, we know from Section 8.3 that this cluster state has the same mathematical structure as the linear cluster state. So what makes it two-dimensional? It is the order in which the measurements are performed, of course! Measurements of qubits connected by a vertical line can be done simultaneously, while measurements of qubits connected by a horizontal line must be done sequentially. As mentioned before, each horizontal line represented a different logical qubit. If we choose to measure qubits 2 and 3 first instead of qubit 1 as in the

linear case then the logical input state for this cluster is $|\Psi_{\text{in}}\rangle = |+\rangle_{1E}|+\rangle_{2E}$, which is "stored" on these qubits. By measuring qubits 2 and 3 simultaneously in the basis $B_2(-\alpha)$ and $B_3(-\beta)$ the output state for the logical qubit becomes

$$|\Psi_{\text{out}}\rangle = [H_1 \otimes H_2] \left[R_{Z_1}(-\alpha) \otimes R_{Z_2}(-\beta) \right] CZ |\Psi_{\text{in}}\rangle. \tag{8.32}$$

The output state is available for read-out on qubits 1 and 4. From this example, we see that simultaneous measurements of two physical qubits that are entangled (i.e. connected by a vertical bond) can generate entanglement between the logical qubits, while obviously destroying the entanglement between the physical qubits. Note that one of the strengths of the one-way quantum computing model is that by merely changing the order and sequence of the single-qubit measurements on a resource state different computations can be computed.

8.6   The Grover Algorithm

In order to show that one-way quantum computing is a viable option for quantum computing Walther *et al.* had to also show that a quantum algorithm could be performed using their cluster state. They selected the Grover's search algorithm because it is provably more efficient than the best classical algorithm and because fast searching is crucial for solving difficult problems [7]. Although they could implement only a two-qubit version of the Grover algorithm (which searches through an unstructured database of four elements) this was already sufficient to demonstrate the simplicity and power of the one-way quantum computing model. This is because it required only four simple projective measurements on the polarisation qubits in the cluster state. The two-qubit Grover algorithm, in general, also shows the efficiency of quantum computing compared to classical computing since the two-qubit case requires only one calculation of the Grover iteration while classically we would need three evaluations in the worst case and on average about 2.25 evaluations.

8.6.1   The circuit two-qubit Grover algorithm

The circuit for the single Grover iteration that is required for the two-qubit case appears in Fig. 8.4 and we will now briefly explain how the circuit accomplishes its purpose. The usual method for the Grover algorithm is to begin with a superposition of all the elements in the database, which is the two-qubit input state $|\Psi_{\text{in}}\rangle = |+\rangle|+\rangle$

$$|\Psi_{\text{in}}\rangle = |0\rangle|0\rangle + |0\rangle|1\rangle + |1\rangle|0\rangle + |1\rangle|1\rangle. \tag{8.33}$$

The Grover iteration can be broken down into four steps [5]:

1. Apply the oracle;
2. Apply the Hadamard transform on each of the n qubits;
3. Perform a conditional phase where every element except $|0\rangle$ receives a π phase-shift (or minus sign);
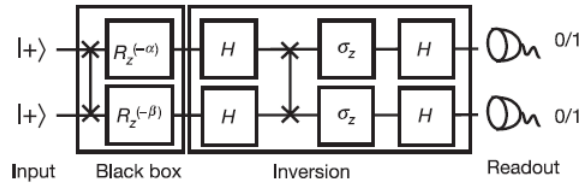4. Once again apply the Hadamard transform on each the n qubits.

Figure 8.4: The quantum circuit implementing Grover's search algorithm for two qubits [7].

The oracle or black box is made up of a CZ gate followed by single-qubit rotation gates $R_Z(-\alpha)$ and $R_Z(-\beta)$ on each of the input qubits. The choice of α and β allows us to mark any of the four database elements as the solution for the search problem, as shown in Table 8.3. Marking the element just means that the solution to the problem gets a minus sign or a π phase shift. Since we cannot identify a phase in a measurement we note that these settings are, in principle, hidden. As explained in Chapter 6, the oracle merely identifies the solution to the search problem and marks it. The aim of the algorithm is to find the marked element.

Table 8.3: The black box settings. The choice of rotation angles allows us to mark different elements as the solution to the search problem.

Marked Element	α	β
$|0\rangle|0\rangle$	π	π
$|0\rangle|1\rangle$	π	0
$|1\rangle|0\rangle$	0	π
$|1\rangle|1\rangle$	0	0

The Hadamard transform is then applied to each of the two qubits as required by Step 2. The σ_Z operation is defined as: $|0\rangle \rightarrow |0\rangle$ and $|1\rangle \rightarrow -|1\rangle$. The CZ together with σ_Z operation allows us to mark all elements except for $|0\rangle$ with a minus sign, thus fulfilling Step 3 of the algorithm. We can test this by applying these operations on the input state $|\Psi_{\text{in}}\rangle$,

$$
\begin{aligned}
|\Psi_{\text{in}}\rangle &\rightarrow [\sigma_{Z_1} \otimes \sigma_{Z_2}] \, CZ \, [|0\rangle|0\rangle + |0\rangle|1\rangle + |1\rangle|0\rangle + |1\rangle|1\rangle] \\
&= [\sigma_{Z_1} \otimes \sigma_{Z_2}] \, [|0\rangle|0\rangle + |0\rangle|1\rangle + |1\rangle|0\rangle - |1\rangle|1\rangle] \\
&= |0\rangle|0\rangle - |0\rangle|1\rangle - |1\rangle|0\rangle - |1\rangle|1\rangle.
\end{aligned}
\tag{8.34}
$$

Finally, the algorithm requires the Hadamard to be once again applied to the two qubits, as done in the circuit. The result is then read out in the computational basis $\{|0\rangle, |1\rangle\}$. We can now summarise the transformations performed by the circuit on the input state as follows

$$
|\Psi_{\text{out}}\rangle = [H_1 \otimes H_2] \, [\sigma_{Z_1} \otimes \sigma_{Z_2}] \, CZ \, [H_1 \otimes H_2] \, [R_{Z_1}(-\alpha) \otimes R_{Z_2}(-\beta)] \, CZ |\Psi_{\text{in}}\rangle.
\tag{8.35}
$$

8.6.2 The one-way two-qubit Grover algorithm

This two-qubit Grover algorithm can be implemented in the one-way quantum computing model using the box cluster state (Fig. 8.5). Let us see why this is possible. Simultaneous measurements on the physical qubits 1 and 4 in the basis $B_1(\alpha)$ and $B_4(\beta)$ entangle the two encoded input qubits with the CZ gate as well as perform the single-qubit rotations, $R_Z(-\alpha)$ and $R_Z(-\beta)$, and the Hadamard gate on each of the encoded qubits. As mentioned, the choice of α and β allows us to mark any of the four elements. The transformation of the encoded input state $|\Psi_{\text{in}}\rangle = |+\rangle|+\rangle$ can be written as follows

$$|\Psi_{\text{out}}\rangle \rightarrow [H_1 \otimes H_2] [R_{Z_1}(-\alpha) \otimes R_{Z_2}(-\beta)] \, CZ |\Psi_{\text{in}}\rangle. \tag{8.36}$$

This state is transferred to qubits 2 and 3. The two measurements performed so far fulfil the first two steps of the Grover algorithm.
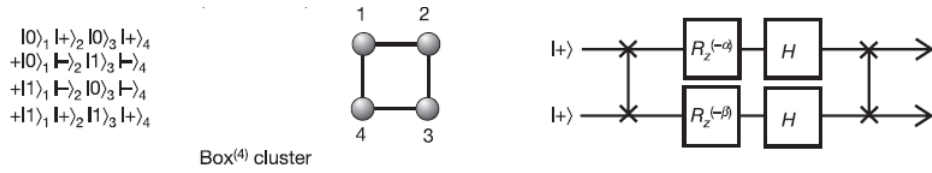

Figure 8.5: The box cluster state that is used to realise Grover's search algorithm in one-way quantum computing [7].

The next two steps are done by measuring qubits 2 and 3. If these measurements are done simultaneously in the computational basis $\{|0\rangle, |1\rangle\}$ then the CZ gate will be implemented. However, if we look at the circuit in Fig. 8.4 we see that we still need to account for the σ_Z and Hadamard gates. We can implement these gates by measuring qubit 2 and 3 in the basis $B(\pi) = \{|+\pi\rangle, |-\pi\rangle\}$ instead since this is equivalent to first implementing the σ_Z followed by a Hadamard gate and read-out in the computational basis,

$$
\begin{aligned}
\{\sigma_Z H|0\rangle \, , \, \sigma_Z H|1\rangle\} &= \left\{ \sigma_Z \tfrac{1}{\sqrt{2}} \big[|0\rangle + |1\rangle \big] \, , \, \sigma_Z \tfrac{1}{\sqrt{2}} \big[|0\rangle - |1\rangle \big] \right\} \\
&= \left\{ \tfrac{1}{\sqrt{2}} \big[|0\rangle - |1\rangle \big] \, , \, \tfrac{1}{\sqrt{2}} \big[|0\rangle + |1\rangle \big] \right\} \\
&= \{|+\pi\rangle \, , \, |-\pi\rangle\} \\
&\equiv \{|-\rangle, \, |+\rangle\} \, . \tag{8.37}
\end{aligned}
$$

The measurement of a qubit j in the basis $B_j(\alpha)$ leads to a measurement outcome s_j which is 0 if the measurement outcome is $|+\alpha\rangle_j$, and as 1 if the outcome is $|-\alpha\rangle_j$. If we post-select the cases where the measurement on qubits 1 and 4 gives $s_1 = s_4 = 0$ then there is a one-to-one correspondence between the measurement outcomes $\{s_2, s_3\}$ and the states that were marked by the black box, so we can easily identify the solution to the search problem. For example, if our measurement outcome was $\{s_2 = 0, s_3 = 1\}$ then the state that was originally marked by the black box was $|0\rangle|1\rangle$ and it is, therefore, the solution that we were looking for. Table 8.4 contains

all the possible combinations of results for the case $s_1 = s_4 = 0$. However, post-selection of the cases $s_1 = s_4 = 0$ reduces the probability of success to $1/4$.

Table 8.4: Identifying the solution to the search problem where $s_1 = s_4 = 0$. Read-out measurements on qubits 2 and 3 are made in the bases $B_{2,3}(\pi)$.

Qubit 2 before measurement	Qubit 3 before measurement	s_2	s_3	Solution
$|-\rangle$	$|-\rangle$	0	0	$|0\rangle|0\rangle$
$|-\rangle$	$|+\rangle$	0	1	$|0\rangle|1\rangle$
$|+\rangle$	$|-\rangle$	1	0	$|1\rangle|0\rangle$
$|+\rangle$	$|+\rangle$	1	1	$|1\rangle|1\rangle$

If we want the computation to be deterministic we need to include the possibilities where the measurement outcomes are other than $s_1 = 0$ and $s_4 = 0$. In these cases we have to take into account feed-forward, which means we have to use these measurement results to determine the meaning of the final read-out. Feed-forward is a key feature of one-way quantum computing and in these cases it is performed by carrying out *bitwise addition*: $\{s_2 \oplus s_4, s_3 \oplus s_1\}$ and interpreting the meaning of the result. Obviously this addition can be used in the cases where $s_1 = s_4 = 0$ as well. Table 8.5 shows how to reinterpret all possible sets of measurement results to obtain the solution.

Table 8.5: Identifying the solution to the search problem for all possible measurement results.

	$\{\mathbf{s}_1 = 0, \mathbf{s}_4 = 0\}$	$\{\mathbf{s}_1 = 0, \mathbf{s}_4 = 1\}$	$\{\mathbf{s}_1 = 1, \mathbf{s}_4 = 0\}$	$\{\mathbf{s}_1 = 1, \mathbf{s}_4 = 1\}$
$\{\mathbf{s}_2 = 0, \mathbf{s}_3 = 0\}$	$|0\rangle|0\rangle$	$|1\rangle|0\rangle$	$|0\rangle|1\rangle$	$|1\rangle|1\rangle$
$\{\mathbf{s}_2 = 0, \mathbf{s}_3 = 1\}$	$|0\rangle|1\rangle$	$|1\rangle|1\rangle$	$|0\rangle|0\rangle$	$|1\rangle|0\rangle$
$\{\mathbf{s}_2 = 1, \mathbf{s}_3 = 0\}$	$|1\rangle|0\rangle$	$|0\rangle|0\rangle$	$|1\rangle|1\rangle$	$|0\rangle|1\rangle$
$\{\mathbf{s}_2 = 1, \mathbf{s}_3 = 1\}$	$|1\rangle|1\rangle$	$|0\rangle|1\rangle$	$|1\rangle|0\rangle$	$|0\rangle|0\rangle$

We know that the box cluster state is required to implement the Grover algorithm and this cluster state is formed by performing the Hadamard operation on each of the physical qubits in the produced cluster state. Although the Hadamard can be implemented by means of a HWP at $22.5°$, we could equivalently perform the Hadamard on the measurement basis, i.e. we *redefined* the output basis. So instead of measuring the qubit in $B_j(\alpha)$ we measure $HB_j(\alpha)$. Also remember that since the box cluster is produced from the cluster state by swapping qubits 2 and 3 in addition to the Hadamard operations, we have to relabel our measurement bases accordingly. The measurement bases used for the algorithm were $B(0)$ and $B(\pi)$. Applying the Hadamard to these two measurement bases gives: $B(0) \rightarrow |0\rangle$ and $B(\pi) \rightarrow |1\rangle$. In terms of polarisation these measurement bases represent a projection onto the $|H\rangle$ or $|V\rangle$ states, respectively. In order to physically implement the projective measurements the linear polariser and QWP were used, as described in Section 8.4.

Chapter 9

One-way quantum computing with orbital angular momentum

The OAM of photons offers access to an infinite-dimensional Hilbert space for encoding quantum information and thereby opens up many exciting opportunities in the realm of quantum information processing. In terms of quantum computation, OAM qudits could provide the possibility to compute more complex tasks with fewer resources compared to polarisation qubits. The intriguing possibilities has prompted extensive research into the generation and manipulation of the OAM states of photons resulting in a well-developed set of optical tools for practical quantum information processing protocols with OAM. This toolkit has allowed us to devise a theoretical proof-of-principle experiment for the implementation of one-way quantum computing with OAM qubits as a first step towards the realisation of a generalised form of one-way quantum computing by means of qudits.

For the sake of clarity and a self-contained presentation we present a brief description of each of the optical elements that are required for the proposed experimental set-up, in Section 9.1. In Section 9.2 we explain the procedure for developing an OAM entangled cluster state which is analogous to the one produced by Walther *et al.* in [7]. Projective measurements on the OAM qubits in the cluster can be performed by means of computer-generated holograms and single-mode fibres, as shown in Section 9.3. Section 9.4 describes how the set-up can be used to realise a two-qubit version of Grover's search algorithm. Finally, a concise summary of the results of this chapter can be found in Section 9.5

9.1   Preliminaries: Optical elements for OAM

The set of optical components necessary to understand the experimental set-up (Fig. 9.5) are listed in this section in order of occurrence. A more comprehensive description of each of these devices can be found the preceding chapters — i.e. spontaneous parametric down-conversion in Chapter 4, Dove prisms in Section 5.3, the astigmatic laser mode converter in Section 5.2, the sorters in Section 5.5 and lastly, the computer-generated holograms in Section 5.1.

9.1.1 Spontaneous parametric down-conversion

If an intense laser pump beam is incident on a non-linear crystal cut for type-I phase-matching then entangled photon pairs are probabilistically generated from a single pump photon. The Hamiltonian for the process in the case of an incident photon in the OAM state $|l = 0\rangle$ is given by

$$\hat{\mathbf{H}} = A_0 \sum_{l=-\infty}^{\infty} C_{l,-l}\hat{\mathbf{a}}_l^\dagger \hat{\mathbf{b}}_{-l}^\dagger + \text{h.c.}, \tag{9.1}$$

where $C_{l,-l} = C_{-l,l}$. Using the Taylor expansion, the quantum mechanical state that is emitted by the non-linear crystal is

$$|\psi\rangle = Ze^{-i\hat{\mathbf{H}}}|0\rangle$$

$$= Z \exp\left(-iA_0 \sum_{l=-\infty}^{\infty} C_{l,-l}\hat{\mathbf{a}}_l^\dagger \hat{\mathbf{b}}_{-l}^\dagger\right)|0\rangle$$

$$= Z\left(1 - iA_0 \underbrace{\sum_{l=-\infty}^{\infty} C_{l,-l}\hat{\mathbf{a}}_l^\dagger \hat{\mathbf{b}}_{-l}^\dagger}_{\text{2 photon state}} - \frac{A_0^2}{2}\underbrace{\left(\sum_{l=-\infty}^{\infty} C_{l,-l}\hat{\mathbf{a}}_l^\dagger \hat{\mathbf{b}}_{-l}^\dagger\right)^2}_{\text{4 photon state}} + \dots\right)|0\rangle, \tag{9.2}$$

where Z is a normalisation constant and $|0\rangle$ is the vacuum state.

9.1.2 Dove Prism

In the experimental set-up we utilise the mirror transformation property of the Dove prism, i.e. $|l\rangle \rightarrow |-l\rangle$. In our case the Dove prism will be used to transform $|l = 1\rangle$ to $|l = -1\rangle$ and vice versa. In the two-dimensional subspace defined by the OAM states $|l = \pm 1\rangle$ and superpositions thereof, this corresponds to a bit flip which is the action of the Pauli-X operator.

9.1.3 Astigmatic Mode Converter

The OAM equivalent of a HWP is the astigmatic laser mode converter (AMC). If it is rotated by an angle $\theta/2$, the action of the AMC on the OAM subspace $|\pm l\rangle$ can be written in the form of a unitary matrix as

$$U_{\text{AMC}}\left(\tfrac{\theta}{2}\right) = \frac{1}{\sqrt{2}}\begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}. \tag{9.3}$$

9.1.4 Sorters

The sorter is a Mach Zehnder interferometer with a beam rotator (rotation angle α) and a phase plate (phase shift $\Delta\Phi_c$) in one arm. Depending on the values of α and $\Delta\Phi_c$ we can sort between individual photons according to their l values, which will then exit the interferometer either through output port 1 or 2.

For our purposes we are required to sort between the OAM states $|l = +1\rangle$ and $|l = -1\rangle$. This is possible if we choose the angles $\alpha = \Delta\Phi_c = \pi/2$ which results in photons in the state $|l = 1\rangle$ exiting from port 1 while those with $|l = -1\rangle$ leave through port 2. Furthermore, we will be interested in the cases where we have two inputs into the sorter and we obtain a coincidence click (or a simultaneous detection of a *single* photon) in both detectors at the output ports. This will occur when we have: (a) two photons in orthogonal OAM states entering from the same input port or (b) the two photons in the same OAM state entering from different input ports. Please refer to Appendix B for a detailed calculation.
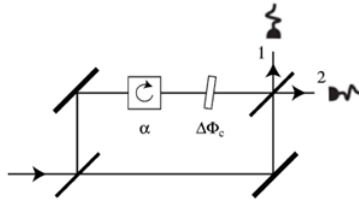

Figure 9.1: The sorter allows photons with different OAM values to exit through different output ports [76].

9.1.5 Holograms and the SLMs

Computer-generated holograms can be used to perform projective measurements onto different OAM states. The hologram is formed by calculating the interferogram between a Gaussian beam ($l = 0$) and an LG beam and the result is then programmed onto a spatial light modulator (SLM). Depending on the azimuthal mode index l of the LG beam this will result in the hologram having $\Delta m = l$ dislocations or fork prongs (refer to Fig. 9.2). An incoming photon in the state $|l = 0\rangle$ passing through the dislocation of the hologram is diffracted and at the nth diffraction order a photon in the state $|l = n\Delta m\rangle$ is formed. Reversing the process, a photon in the OAM state $|m\rangle$ before the grating will be transformed into the state $|l = 0\rangle$ in the first diffraction order and using a single-mode fibre (so as to allow only the $|l = 0\rangle$ state), a detection will be made. This procedure serves as a projection onto the state $|m\rangle$.
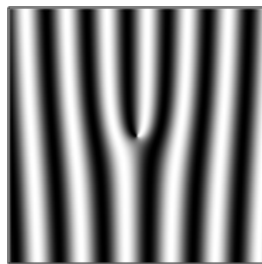

Figure 9.2: A fork hologram with a single dislocation or fork prong at the centre.

We will be interested in making projective measurements onto an equal superposition of two OAM states $|\pm l\rangle$. These measurements require a hologram which is formed using a superposition of the LG modes with azimuthal mode indices $\pm l$ instead of a single LG mode. An example of such a hologram is given in Fig. 9.3. If a photon in the state $|l = 0\rangle$ is incident on the hologram then a photon in an equally weighted superpositon of the OAM states $|\pm l\rangle$ is generated. This means that in reverse, this hologram will result in a detector "click" for $|l\rangle$ states 50% of the time and $|-l\rangle$ states 50% of the time, which is equivalent to projecting onto the superposition state.
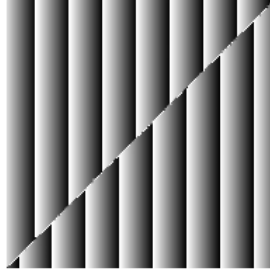

Figure 9.3: A hologram which is used to measure an equal superposition of OAM states $|l = 1\rangle$ and $|l = -1\rangle$.

9.2 Production of the Cluster State

The target state for our proof-of-principle is analogous to the cluster state that was generated by Walther in [7] and is given by

$$|\Psi_\mathrm{C}\rangle = \tfrac{1}{2}\Big[|l = -1\rangle_1|l = -1\rangle_2|l = -1\rangle_3|l = -1\rangle_4 + |l = -1\rangle_1|l = -1\rangle_2|l = 1\rangle_3|l = 1\rangle_4$$
$$+ |l = 1\rangle_1|l = 1\rangle_2|l = -1\rangle_3|l = -1\rangle_4 - |l = 1\rangle_1|l = 1\rangle_2|l = 1\rangle_3|l = 1\rangle_4\Big]. \quad (9.4)$$

To generate this cluster state, which serves as an entanglement substrate for one-way quantum computing, we employ the *rail-cross* scheme that was used in [7]. The principle resource for the OAM entangled photons is SPDC using a β-Barium Borate (BBO) crystal cut for type-I phase-matching so that both photons carry the same linear polarisation. A Gaussian ($l = 0$) ultraviolet laser pulse passes through a single BBO crystal which probabilistically releases highly entangled photons in the forward direction into spatial modes *a-b*. The pulse is then reflected by a pump mirror so that it traverses the BBO crystal once again, producing highly entangled photon pairs in the backward direction into modes *c-d*. Fig. 9.4 depicts this first step of the preparation process for the cluster state.
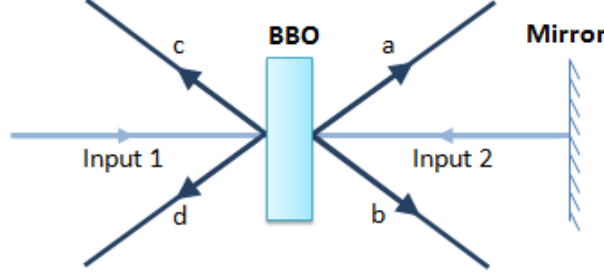
Figure 9.4: Spontaneous Parametric Down Conversion. A UV pulse passes through a BBO crystal probabilistically releasing highly entangled photon pairs in the forward direction a-b. This pulse is then reflected by a mirror and passes through the crystal for a second time creating entangled photon pairs in the backward direction c-d.

The possible emission of an entangled photon pair in the forward-directional modes a-b and backward-directional modes c-d can be specified by the Hamiltonian [91]

$$\hat{\mathbf{H}} = \sum_{l=-\infty}^{\infty} C_{l,-l}\hat{\mathbf{a}}_l^\dagger\hat{\mathbf{b}}_{-l}^\dagger + p_r e^{i\Delta\phi} \sum_{l=-\infty}^{\infty} C_{l,-l}\hat{\mathbf{c}}_l^\dagger\hat{\mathbf{d}}_{-l}^\dagger, \tag{9.5}$$

where, for example, $\hat{\mathbf{a}}_l^\dagger$ represents the creation of a photon with an OAM of $l\hbar$ in the spatial mode a; p_r is relative probability of photon pairs being emitted into the backward direction c-d compared to the forward direction a-b; and $\Delta\phi$ is the relative phase between emissions of photon pairs in either direction which can be set by the pump mirror. If we apply this Hamiltonian to the vacuum we can obtain a single-pair production (or a two-photon state) in either direction

$$|\Psi_2\rangle \propto \hat{\mathbf{H}}|0\rangle = \sum_{l=-\infty}^{\infty} C_{l,-l}|l\rangle_a|-l\rangle_b + p_r e^{i\Delta\phi} \sum_{l=-\infty}^{\infty} C_{l,-l}|l\rangle_c|-l\rangle_d. \tag{9.6}$$

These photon pairs are entangled in their OAM states as well as in their spatial modes. The Hamiltonian generates two multi-dimensional OAM entangled photon pairs, so each photon pair represents two entangled *qudits*. For our proof-of-principle we want to restrict ourselves to working with *qubits*, therefore we need to select a two-dimensional subspace of the multi-dimensional space. In practice this is possible if we make measurements that are sensitive only to the photons spontaneously prepared in this two-dimensional subspace [70]. In essence, we perform a post-selection of certain l values. Because of the eventual projection onto the qubit space we can simplify our calculations by taking into account only those terms that are part of our subspace, as we will now show.

Selecting a suitable subspace is one of the critical steps in the preparation of the target state — a poor choice could lead to unnecessary complications in the proposed experimental set-up. A sensible choice seems to be $|l = \pm 1\rangle$ since we have seen that there is an analogy between these

OAM states and the polarisation states $|H\rangle$ and $|V\rangle$, specifically the fact that these states can be represented on a Bloch sphere equivalent to the one used for polarisation (cf. Section 6.1). We can thus redefine our Hamiltonian as

$$\hat{\mathbf{H}}_{\text{2D}} \propto \left(\hat{\mathbf{a}}_{-1}^{\dagger}\hat{\mathbf{b}}_{1}^{\dagger} + \hat{\mathbf{a}}_{1}^{\dagger}\hat{\mathbf{b}}_{-1}^{\dagger}\right) + p_r e^{i\Delta\phi}\left(\hat{\mathbf{c}}_{-1}^{\dagger}\hat{\mathbf{d}}_{1}^{\dagger} + \hat{\mathbf{c}}_{1}^{\dagger}\hat{\mathbf{d}}_{-1}^{\dagger}\right). \tag{9.7}$$

When applying this Hamiltonian to the vacuum we obtain a photon pair in the forward and backward directions, each of which is entangled in two OAM states,

$$\begin{aligned}|\Psi_{\text{2D}}\rangle = \tfrac{1}{\sqrt{2}}\Big[&\left(|\boldsymbol{l} = -1\rangle_a|\boldsymbol{l} = 1\rangle_b + |\boldsymbol{l} = 1\rangle_a|\boldsymbol{l} = -1\rangle_b\right) \\ &+ p_r e^{i\Delta\phi}\left(|\boldsymbol{l} = -1\rangle_c|\boldsymbol{l} = 1\rangle_d + |\boldsymbol{l} = 1\rangle_c|\boldsymbol{l} = -1\rangle_d\right)\Big].\end{aligned} \tag{9.8}$$

This two-photon state does not contribute to the formation of the cluster state but it could be useful for alignment purposes [81].

Since the production of our cluster state is based on the simultaneous emission of four photons, just as in [7], we need to apply the Hamiltonian twice to the vacuum

$$\begin{aligned}\hat{\mathbf{H}}_{\text{2D}}^2|0\rangle = \Big[&\left(\hat{\mathbf{a}}_{-1}^{\dagger}\hat{\mathbf{b}}_{1}^{\dagger} + \hat{\mathbf{a}}_{1}^{\dagger}\hat{\mathbf{b}}_{-1}^{\dagger}\right)^2 + p_r^2 e^{i2\Delta\phi}\left(\hat{\mathbf{c}}_{-1}^{\dagger}\hat{\mathbf{d}}_{1}^{\dagger} + \hat{\mathbf{c}}_{1}^{\dagger}\hat{\mathbf{d}}_{-1}^{\dagger}\right)^2 \\ &+ 2p_r e^{i\Delta\phi}\left(\hat{\mathbf{a}}_{-1}^{\dagger}\hat{\mathbf{b}}_{1}^{\dagger} + \hat{\mathbf{a}}_{1}^{\dagger}\hat{\mathbf{b}}_{-1}^{\dagger}\right)\left(\hat{\mathbf{c}}_{-1}^{\dagger}\hat{\mathbf{d}}_{1}^{\dagger} + \hat{\mathbf{c}}_{1}^{\dagger}\hat{\mathbf{d}}_{-1}^{\dagger}\right)\Big]|0\rangle.\end{aligned} \tag{9.9}$$

The first and second terms in this equation represent the emission of two photon pairs (called *double-pair* emission) in the forward direction into spatial modes *a-b* and in the backward direction into spatial modes *c-d*, respectively. The third term denotes the simultaneous production of a single photon pair in either direction. Fig. 9.5 illustrates the sequence of optical elements that will allow this four-photon state to be transformed into the target cluster state (Eq. (9.4)). We will now be evaluating each of these terms separately. For this reason we shall disregard the prefactors that occur in Eq. (9.9), and concentrate solely on the creation operators, until we present the final result at the end of this section.

We will first focus on the simultaneous production of the single photon pairs,

$$\begin{aligned}|\Psi_{\text{single}}^{(0)}\rangle &= \left[\left(\hat{\mathbf{a}}_{-1}^{\dagger}\hat{\mathbf{b}}_{1}^{\dagger} + \hat{\mathbf{a}}_{1}^{\dagger}\hat{\mathbf{b}}_{-1}^{\dagger}\right)\left(\hat{\mathbf{c}}_{-1}^{\dagger}\hat{\mathbf{d}}_{1}^{\dagger} + \hat{\mathbf{c}}_1{}^{\dagger}\hat{\mathbf{d}}_{-1}^{\dagger}\right)\right]|0\rangle \\ &= \tfrac{1}{\sqrt{2}}\left(|\boldsymbol{l} = -1\rangle_a|\boldsymbol{l} = 1\rangle_b + |\boldsymbol{l} = 1\rangle_a|\boldsymbol{l} = -1\rangle_b\right) \\ &\otimes \tfrac{1}{\sqrt{2}}\left(|\boldsymbol{l} = -1\rangle_c|\boldsymbol{l} = 1\rangle_d + |\boldsymbol{l} = 1\rangle_c|\boldsymbol{l} = -1\rangle_d\right).\end{aligned} \tag{9.10}$$

The first optical elements in the set-up that transform this state are the Dove prisms in modes *b* and *d*, which results in the state

$$\begin{aligned}|\Psi_{\text{single}}^{(1)}\rangle &= \tfrac{1}{\sqrt{2}}\left(|\boldsymbol{l} = -1\rangle_a|\boldsymbol{l} = -1\rangle_b + |\boldsymbol{l} = 1\rangle_a|\boldsymbol{l} = 1\rangle_b\right) \\ &\otimes \tfrac{1}{\sqrt{2}}\left(|\boldsymbol{l} = -1\rangle_c|\boldsymbol{l} = -1\rangle_d + |\boldsymbol{l} = 1\rangle_c|\boldsymbol{l} = 1\rangle_d\right).\end{aligned} \tag{9.11}$$
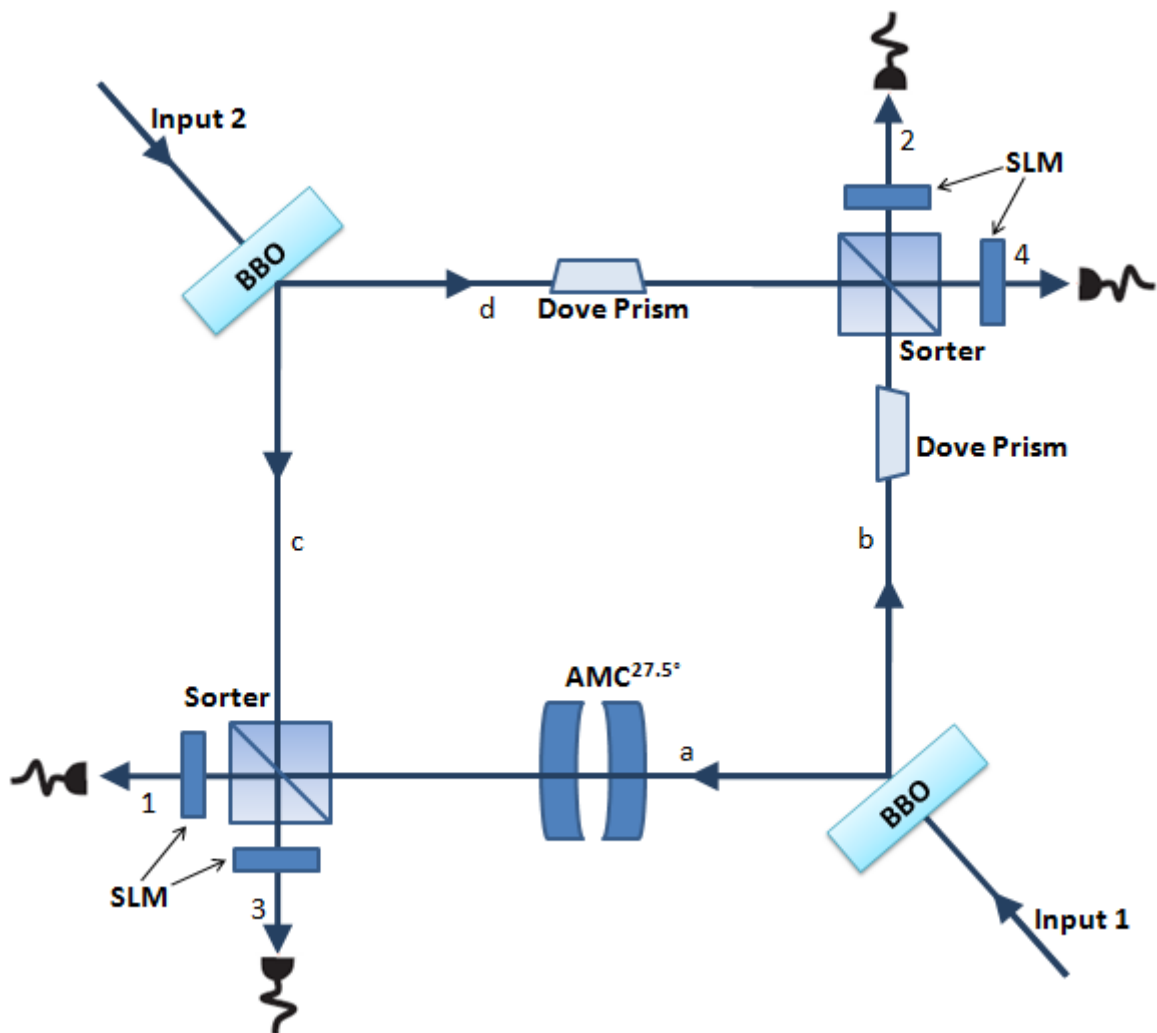
Figure 9.5: A schematic diagram of the proposed set-up for the construction of the OAM entangled cluster state. This state is based on the simultaneous emission of four photons and the post-selection of four-fold coincidence counts in the output ports 1–4. The Dove prisms and the astigmatic mode converter (AMC) at $27.5°$ allows us to obtain the correct cluster state. OAM projective measurements on each qubit are made using computer-generated holograms programmed onto a Spatial Light Modulator (SLM) and a single-mode fibre followed by a detector in each output port.

This state contains two distinct entangled photon pairs, one in spatial modes a and b and one in modes c and d. We have to entangle all four photons in order to obtain the cluster state — in particular, we must entangle spatial modes a and c and b and d of the photon pairs. This is done by post-selection of certain components, as will be discussed below.

Let us consider the effect of the sorters. The two sorters channel the OAM states of the four photons from the input modes a, b, c and d to output modes 1, 2, 3 and 4 as shown in Table 9.1. At this point it is important to note that our cluster state arises from the post-selection of only four-photon events from the SPDC process as well as post-selecting only the cases where we obtain the simultaneous detection of a single photon in each of the four output ports 1, 2, 3 and 4 (called four-fold coincidences) just as was done in [7]. Photons coming from the same input port must be in different OAM states while photons coming from different input ports must have the same OAM state in order to obtain a coincidence click in each of the detectors after a sorter.

Table 9.1: The OAM states of the photon pairs are channeled from input modes a, b, c and d to output modes 1, 2, 3 and 4 by the sorters.

$\lvert l = -1 \rangle_a \to \quad \lvert l = -1 \rangle_1$	$\lvert l = 1 \rangle_a \to \ -\lvert l = 1 \rangle_3$
$\lvert l = -1 \rangle_b \to \quad \lvert l = -1 \rangle_2$	$\lvert l = 1 \rangle_b \to \ -\lvert l = 1 \rangle_4$
$\lvert l = -1 \rangle_c \to \ -\lvert l = -1 \rangle_3$	$\lvert l = 1 \rangle_c \to \quad \lvert l = 1 \rangle_1$
$\lvert l = -1 \rangle_d \to \ -\lvert l = -1 \rangle_4$	$\lvert l = 1 \rangle_d \to \quad \lvert l = 1 \rangle_2$

The post-selection of four-fold coincidences results in Eq. (9.11) being transformed to

$$\lvert \Psi^{(2)}_{\text{single}} \rangle = \tfrac{1}{2}\Big[\lvert l = -1 \rangle_1 \lvert l = -1 \rangle_2 \lvert l = -1 \rangle_3 \lvert l = -1 \rangle_4$$
$$+ \ \lvert l = 1 \rangle_1 \lvert l = 1 \rangle_2 \lvert l = 1 \rangle_3 \lvert l = 1 \rangle_4 \Big]. \tag{9.12}$$

after the sorters. This gives us only two of the four terms that are required for the target cluster state (Eq. (9.4)). However, we see from Eq. (9.9) that with equal probability as the single-pair production there exists the case of double-pair emission of entangled photons in the forward and backward directions. Considering each of these cases will give us the two remaining terms for the cluster state.

Double-pair emission into the forward direction a-b results in the four-photon state

$$\lvert \Psi^{(0)}_{\text{double-}ab} \rangle = \left(\hat{\mathbf{a}}^{\dagger}_{-1} \hat{\mathbf{b}}^{\dagger}_{1} + \hat{\mathbf{a}}^{\dagger}_{1} \hat{\mathbf{b}}^{\dagger}_{-1} \right)^2 \lvert 0 \rangle$$
$$= \lvert l = -1, l = -1 \rangle_a \lvert l = 1, l = 1 \rangle_b + \lvert l = 1, l = 1 \rangle_a \lvert l = -1, l = -1 \rangle_b$$
$$+ \lvert l = -1, l = 1 \rangle_a \lvert l = -1, l = 1 \rangle_b, \tag{9.13}$$

where, for example, $\lvert l = -1, l = -1 \rangle_a$ denotes two photons with an OAM of $-\hbar$ in the spatial

mode a. The Dove prisms cause this state to evolve to

$$|\Psi^{(1)}_{\text{double-}ab}\rangle = |l = -1, l = -1\rangle_a|l = -1, l = -1\rangle_b + |l = 1, l = 1\rangle_a|l = 1, l = 1\rangle_b$$
$$+ |l = -1, l = 1\rangle_a|l = -1, l = 1\rangle_b, \tag{9.14}$$

and after the traversing the sorters

$$|\Psi^{(2)}_{\text{double-}ab}\rangle = |l = -1\rangle_1|l = -1\rangle_2|l = 1\rangle_3|l = 1\rangle_4. \tag{9.15}$$

We obtain this state because only the third term in Eq. (9.14) results in four-fold coincidence clicks and contributes to the formation of the cluster state. The two other components represent the cases where we have two photons exiting through a single output port and are thus discarded. Similarly, double-pair emission in the backward direction c-d results in the state

$$|\Psi^{(2)}_{\text{double-}cd}\rangle = |l = 1\rangle_1|l = 1\rangle_2|l = -1\rangle_3|l = -1\rangle_4, \tag{9.16}$$

in the output ports of the sorters. The two double-pair emission states will be in a coherent super-position if we adjust the position of the pump mirror (shown in Fig. 9.4) so that the amplitudes of both of these emissions arrive at the sorters simultaneously

$$|\Psi^{(2)}_{\text{double}}\rangle = \tfrac{1}{2}\Big(|l = -1\rangle_1|l = -1\rangle_2|l = 1\rangle_3|l = 1\rangle_4$$
$$+ |l = 1\rangle_1|l = 1\rangle_2|l = -1\rangle_3|l = -1\rangle_4 \Big). \tag{9.17}$$

By combining the single and double-pair contributions that we have just calculated we can identify the adjustments that need to be made to the state we have produced so far, in order for the desired cluster state to be formed,

$$|\Psi_{\text{C}}\rangle = |\Psi^{(2)}_{\text{single}}\rangle + |\Psi^{(2)}_{\text{double}}\rangle$$
$$= \tfrac{1}{2}\Big[|l = -1\rangle_1|l = -1\rangle_2|l = -1\rangle_3|l = -1\rangle_4$$
$$+ |l = -1\rangle_1|l = -1\rangle_2|l = 1\rangle_3|l = 1\rangle_4 + |l = 1\rangle_1|l = 1\rangle_2|l = -1\rangle_3|l = -1\rangle_4$$
$$+ |l = 1\rangle_1|l = 1\rangle_2|l = 1\rangle_3|l = 1\rangle_4 \Big], \tag{9.18}$$

where we have excluded the probability of emission and the phase factors for the moment, as we have previously mentioned. Comparing this equation to Eq. (9.4) we find that in order to obtain the required cluster state we need to shift the phase of the last term by π. This can be done by introducing the astigmatic mode converter (AMC) at an angle $\theta/2$ in mode a. The AMC enables the following transformations

$$|l = -1\rangle \xrightarrow{AMC} \cos\theta|l = -1\rangle + \sin\theta|l = 1\rangle, \tag{9.19}$$

$$|l = 1\rangle \xrightarrow{AMC} \sin\theta|l = -1\rangle - \cos\theta|l = 1\rangle. \tag{9.20}$$

With the addition of the AMC, the single-pair component of the produced cluster state becomes

$$
\begin{aligned}
|\Psi_{\text{single}}\rangle = \tfrac{1}{\sqrt{2}} \Big[& (\cos\theta|l=-1\rangle_a + \sin\theta|l=1\rangle_a)|l=-1\rangle_b \\
+ & (\sin\theta|l=-1\rangle_a - \cos\theta|l=1\rangle_a)|l=1\rangle_b \Big] \\
\otimes \tfrac{1}{\sqrt{2}} \Big[& |l=-1\rangle_c|l=-1\rangle_d + |l=1\rangle_c|l=1\rangle_d \Big].
\end{aligned} \tag{9.21}
$$

which results in the state

$$
\begin{aligned}
|\Psi_{\text{single}}\rangle = \tfrac{1}{2}\cos\theta \Big[& |l=-1\rangle_1|l=-1\rangle_2|l=-1\rangle_3|l=-1\rangle_4 \\
- & |l=1\rangle_1|l=1\rangle_2|l=1\rangle_3|l=1\rangle_4 \Big],
\end{aligned} \tag{9.22}
$$

after the interferometers, since all other terms do not contribute to four-fold coincidences. The AMC, however, also affects the double-pair in mode a. Only the third term in Eq. (9.14) contributes to the production of the cluster state and the AMC transforms this state to

$$
\begin{aligned}
|\Psi_{\text{double}}^{ab}\rangle = \Big[& \sin\theta\cos\theta|l=-1,l=-1\rangle_a - \sin\theta\cos\theta|l=1,l=1\rangle_a \\
- & \cos 2\theta|l=-1,l=1\rangle_a \Big] \otimes |l=-1,l=1\rangle_b,
\end{aligned} \tag{9.23}
$$

which results in the state

$$
|\Psi_{\text{double}}^{ab}\rangle = -\cos 2\theta|l=-1\rangle_1|l=-1\rangle_2|l=1\rangle_3|l=1\rangle_4, \tag{9.24}
$$

after the interferometers.

Putting all these terms together once again, and including the probability of emission and the phases that occur in Eq. (9.9), we arrive at

$$
\begin{aligned}
|\Psi_{\text{C}}\rangle = \tfrac{1}{2} \Big[& p_r\cos\theta e^{i\Delta\phi}|l=-1\rangle_1|l=-1\rangle_2|l=-1\rangle_3|l=-1\rangle_4 \\
- & \cos 2\theta|l=-1\rangle_1|l=-1\rangle_2|l=1\rangle_3|l=1\rangle_4 \\
+ & p_r^2 e^{i2\Delta\phi}|l=1\rangle_1|l=1\rangle_2|l=-1\rangle_3|l=-1\rangle_4 \\
+ & p_r\cos\theta e^{i\Delta\phi}|l=1\rangle_1|l=1\rangle_2|l=1\rangle_3|l=1\rangle_4 \Big].
\end{aligned} \tag{9.25}
$$

Note that the factor 2 which appears in Eq. (9.9) does not appear here since each of the components of the four-photon state have the same probability amplitude otherwise the state would be factorisable instead of entangled (cf. Eq. (4) in [93]). Our aim was to arrive at Eq. (9.4) and for $\theta > 45°$ we obtain the correct signs for the cluster state. Simple simultaneous equations to equalise all the prefactors result in the values $p_r = 1/\sqrt{3}$ and $\theta = 54.7°$. The value for p_r is acceptable since usual emission rates in the forward and backward direction are $30000s^{-1}$ and $10000s^{-1}$ two-photon coincidences, respectively (i.e $p_r^2 = 10000/30000 = 1/3$) [81]. The AMC has to be rotated by $27.5°$ and the pump mirror must be used to set the phase of the back-reflected pair such that $\Delta\phi = 2\Delta\phi = 2n\pi$ where n is an integer. This value must be fixed throughout the experiment.

9.3   Projective Measurements

In order to process information in the one-way quantum computing model we need to be able to perform projective measurements on the cluster state in the basis

$$B_j(\alpha) = \{|+\alpha\rangle_j, |-\alpha\rangle_j\} \ \text{ with } \ |\pm\alpha\rangle_j = \tfrac{1}{\sqrt{2}}(|0\rangle_j \pm e^{i\alpha}|1\rangle_j) \tag{9.26}$$

where $\alpha \in [0, 2\pi]$ and j denotes the qubit. The outcome m_j of a measurement of physical qubit j is defined as 0 if the measurement outcome is $|+\alpha\rangle_j$, and as 1 if $|-\alpha\rangle_j$ is the outcome. This measurement basis corresponds to a set of projections onto the states that lie on the equator of the Bloch sphere. This Bloch sphere is the geometrical representation of a qubit with basis states $\{|0\rangle, |1\rangle\}$ which lie at the poles. When dealing with OAM qubits we have an analogous Bloch sphere but the basis states in our case are $\{|l = -1\rangle, |l = 1\rangle\}$. We, therefore, redefine the measurement basis as

$$B_j(\alpha) = \{|+\alpha\rangle_j, |-\alpha\rangle_j\} \ \text{ with } \ |\pm\alpha\rangle_j = \tfrac{1}{\sqrt{2}}\left(|l = -1\rangle_j \pm e^{i\alpha}|l = 1\rangle_j\right). \tag{9.27}$$

Projective measurements in this basis on each of the four OAM qubits in cluster state can be easily performed through the use of a computer-generated hologram and a single-mode fibre followed by a single-photon counting detector in each of the four output ports of the sorters. A fork hologram allows us to project onto the linear OAM states $|l = \pm 1\rangle$. For example, if we want to project onto the state $|l = -1\rangle$ we use a hologram with a single dislocation. If a photon in the state $|l = -1\rangle$ is incident on such a hologram then this photon will be transformed into the state $|l = 0\rangle$ in the first diffraction order and using the single-mode fibre (so as to only allow in the $|l = 0\rangle$ state) a detection will be made. If the detector registers a "click" then it means that only the $|l = -1\rangle$ component of the state before the sorter plays a role — effectively this state can be viewed as if it were projected onto the state $|l = -1\rangle$. We can project onto $|l = 1\rangle$ in a similar fashion with the same hologram rotated by $180°$.

In order to project onto superpositions of the states we use a hologram of the form given in Fig. 9.3. This hologram will result in a detector "click" for $|l = -1\rangle$ modes 50% of the time and $|l = 1\rangle$ modes 50% of the time, which is equivalent to projecting onto the superposition state. By changing the relative phase between the between the LG modes used to form the hologram we can project onto any of the $|\pm\alpha\rangle$ states.

9.4   Grover's search algorithm with an OAM cluster state

We wish to demonstrate a two-qubit Grover algorithm using our OAM entangled cluster state, based on the implementation in [7] using the a cluster state with polarisation qubits. The Grover algorithm was introduced in Chapter 6 where we explained the workings of this algorithm in detail. The circuit model as well as the one-way quantum computing method for executing a two-qubit version of this algorithm was described in Section 8.6. In this section we illustrate how this algorithm can be realised using OAM qubits.

First note the Hadamard operation in terms of the basis states given in Eq. (9.27) is defined as

$$H|l = -1\rangle = \tfrac{1}{\sqrt{2}}\left(|l = -1\rangle + |l = 1\rangle\right),$$
$$H|l = 1\rangle = \tfrac{1}{\sqrt{2}}\left(|l = -1\rangle - |l = 1\rangle\right). \tag{9.28}$$

Now, Grover's search algorithm is implemented using a box cluster state (refer to Eq. (8.28) and Fig. 8.5). This two-dimensional state can be obtained from the produced cluster state by performing the unitary operation $H_1 \otimes H_2 \otimes H_3 \otimes H_4$ (where H_1 represents the Hadamard operation on the first physical qubit) and swapping or relabeling qubits 2 and 3 [7]. The Hadamard operation can be physically implemented with an AMC rotated at 22.5°. However, as mentioned in Section 8.6.2, it is equivalent and often easier to incorporate the Hadamard operation, which is meant to be performed on the individual qubits, into the corresponding measurement bases. This means that we basically redefined the basis in which a particular qubit is to be measured by applying the Hadamard to it instead of physically implementing the Hadamard on the associated qubit. For example, if the Hadamard is to be carried out on qubit 1 and the measurement basis for qubit 1 is

$$B_1(0) = \{|+\rangle_1 , |-\rangle_1\} \tag{9.29}$$
$$= \left\{ \tfrac{1}{\sqrt{2}}\big(|l = -1\rangle_1 + |l = 1\rangle_1\big) , \tfrac{1}{\sqrt{2}}\big(|l = -1\rangle_1 - |l = 1\rangle_1\big)\right\}, \tag{9.30}$$

then we perform the measurement on qubit 1 in the basis

$$HB_1(0) = \{|l = -1\rangle_1, |l = 1\rangle_1\}, \tag{9.31}$$

instead, which is a simple projection onto the basis states.

In order to implement the Grover algorithm we need only perform measurement in the bases $B(0)$ or $B(\pi)$ on each of the four qubits in our cluster state, as explained in Section 8.6.2. Since the Hadamard has to be performed on all four qubits we will actually only have to perform measurements in the bases $HB(0) = \{|l = -1\rangle, |l = 1\rangle\}$ and $HB(\pi) = \{|l = 1\rangle, |l = -1\rangle\}$. Each of these bases represent projections onto the basis states. These projections can be easily realised using fork holograms and single-mode fibres, as explained in the previous section.

9.5 Summary

We have developed an experimental set-up that allows us to realise a four-qubit OAM entangled cluster state

$$|\Psi_C\rangle = \tfrac{1}{2}\Big[|l = -1\rangle_1|l = -1\rangle_2|l = -1\rangle_3|l = -1\rangle_4 + |l = -1\rangle_1|l = -1\rangle_2|l = 1\rangle_3|l = 1\rangle_4$$

$$+ |l = 1\rangle_1|l = 1\rangle_2|l = -1\rangle_3|l = -1\rangle_4 - |l = 1\rangle_1|l = 1\rangle_2|l = 1\rangle_3|l = 1\rangle_4\Big]. \tag{9.32}$$

With this cluster state and a means to do single-qubit projective measurements (cf. Section 9.4) we have the necessary ingredients to realise a simple one-way quantum computing scheme. Our set-up enables us to implement a two-qubit version of the Grover search algorithm using the protocol described in Section 8.6.2. Measurements on the equivalence class of cluster states described Section 8.3 allow one two implement single- and two-qubit quantum logic gates which are necessary for universal quantum computation.

Chapter 10

Conclusion and Outlook

The one-way quantum computing model allows for a simple method in which to implement quantum algorithms by means of single-qubit measurements on a sufficiently large entanglement resource, called the cluster state. We designed a practical scheme to create a four-partite cluster state using qubits encoded in the OAM degree of freedom of single photons. This set-up, which was transcribed from the proof-of-principle experiment of Walther *et al.* [7], involves the non-linear process of spontaneous parametric down-conversion followed by a set of linear optical elements and post-selection of certain detection events. Necessary single-qubit projective measurements in the equatorial plane of the Bloch sphere can be easily performed using phase holograms, by employing the method developed by Padgett *et al.* in [69]. As shown by Walther *et al.* [7], this cluster state has a few unitarily equivalent states and different measurement sequences on the individual qubits allow one to execute a universal set of single- and two-qubit quantum logic gates. The fact that these gates have been realised with the same experimental set-up and with the same four-qubit linear cluster state underlines in a striking way the flexibility, power and simplicity of the one-way approach.

A two-bit form of Grover's search algorithm, which performs a search through an unstructured database of four elements, can be realised through a specific measurement pattern on the produced cluster state. This algorithm is important because it demonstrates how quantum parallelism can be utilised to give a speed-up (in this case, a quadratic speed-up) over classical algorithms. The implementation using the framework of one-way quantum computing gives an additional speed-up since only four simple single-qubit measurements are necessary to provide the correct solution with 100% accuracy instead of executing the ten quantum logic gates that are required for the circuit model. This is because gates which belong to the Clifford group (for example, the Hadamard and CZ), which are sequential in the circuit model, are parallelisable in the one-way model. Grover's algorithm has a useful application in the field of cryptography since it is theoretically possibly to use this algorithm to crack the Data Encryption Standard (DES), a standard which is used to protect, amongst other things, financial transactions between banks [94].

While the polarisation of photons provides a natural qubit, being discrete and two-dimensional, it is also limited by this feature. OAM, on the other hand, is a natural qudit (d-level system). For our proposed proof-of-principle experiment we selected only a two-dimensional subspace of all

possible OAM states, thereby choosing to work with pseudo qubits. However, our result lays the foundation for the use of photonic OAM for applications in quantum information processing. Although qudits do not provide an exponential speed-up over qubit-based quantum computing, they may provide improvements in computing power in practical situations where systems are subject to physical constraints. There are two existing generalisations of the qubit one-way quantum computing model to higher-dimensional Hilbert spaces [95, 96], so the next step would be to create an OAM entangled qudit cluster state for one-way quantum computing. Note that there are a few limitations for using qudits encoded in the OAM of photons with current technology. For example, it is quite difficult to carry out unitary transformations from one arbitrary OAM mode into another. The mode converter (cf. Section 5.2), which allows the perfect conversion between LG and HG modes, is only able to perform certain transformations and is, therefore, quite inflexible. Although arbitrary spatial transformations are possible using holograms they are probabilistic and non-reversible.

One important improvement for the experiment that we have proposed is incorporating the feed-forward of measurement results and adaptive measurements which allows the computation to be deterministic. This has been done for the polarisation case using PBSs, Pockel cells (which can act as variable wave plates) and single-mode fibres [97, 98]. The PBSs allow for the recording of both measurement results: "0" (when we project onto the desired state) or "1" (when we project onto its antipodal state in the equatorial plane of the Bloch sphere). The measurement bases can then be adapted by employing fast-switching and Pockel cells to change the basis of the analyser (which is made up of a QWP, a linear polariser and a PBS) [81]. Our impression is that this experiment can easily be adapted to the OAM case using a combination of elements from the toolkit that we have described in Chapter 5 since many of these act analogously to the polarisation case. However, a more thorough investigation is necessary.

The only one of DiVincenzo's five criteria (cf. Section 6.6) for building a quantum computer that is not satisfied by this experiment is that it is not scalable to an arbitrary number of qubits. This is because the cluster state is constructed by spontaneous parametric down-conversion which is a highly probabilistic process since only one in 10^6 photons from the pump is actually down-converted. The formation of the four-qubit cluster state also depends on the selection of specific cases where we obtain a detection in each of the four output ports. The problem with this is that the probability of n-photon coincidences decreases exponentially with n. Our method suffices for a proof-of-principle, but new techniques are necessary for the efficient generation of larger cluster states for the development of a practical quantum computer.

Another important component of this research comprised a review of the Dirac notation for paraxial light beams. This formalism, which is usually used to represent quantum states, can also be employed to represent the classical states of paraxial light, which are given by the electric field vector. The use of Dirac kets and operator algebra for the propagation of paraxial light beams through optical elements simplifies the analysis, as demonstrated in [32, 38, 39, 40]. In [42] the authors showed that it was possible to apply operator algebra to the wave function of a paraxial photon and propagate it through the optical elements, thus demonstrating that these components act analogously on a single-photon level. Although this analogy exists, we also note that there is

a significant difference between the classical and quantum states of paraxial light beams. The LG modes $u_{pl}(z)$ form an orthogonal basis for the classical states of the electromagnetic field $u(z)$ so a general state of a paraxial light beam can be expressed as a linear combination of these LG modes in terms of Dirac notation (i.e $|u(z)\rangle = \sum_{p,l} c_{pl}|u_{pl}(z)\rangle$). The state vectors $|u(z)\rangle$ are normalised, therefore $|c_{pl}|^2 = 1$. A single-photon Fock state in the LG basis can be expressed as $|\psi\rangle = \sum_{p,l} c_{pl}|pl\rangle$. In a measurement we see the difference between a single-photon state and a paraxial beam of light. With $|\langle \mathbf{r}|u(z)\rangle|^2$ we obtain the intensity of the electric field, which is the sum of the intensities $|c_{p,l}|^2$ of each of the LG modes, u_{pl}. This is a measurable and predetermined quantity. However, in quantum optics with $|\psi|^2$ the square of the coefficients $|c_{p,l}|^2$ give only the probability of finding a photon in the mode u_{pl}.

Quantum information processing is one of the most exciting and diverse fields of science today. There are many different designs for quantum computers and several of these have been demonstrated on a small scale. Single photons that carry quantum information in their OAM degree of freedom are promising candidates for the realisation of quantum computers. Photons are easily available, they tend to not interact with their environment and, therefore, suffer from negligible decoherence. Their excellent manipulability allows for the precise execution of single-qubit operations. However, the absence of photon-photon interaction makes two-qubit operations, required for universal quantum computing, very difficult to realise. The one-way quantum computing model allows us to carry out all the two-qubit operations right at the beginning of the computation by creating a suitable entanglement resource, which is much easier to carry out. Although the experimental set-up presented in this dissertation corresponds to a proof-of-principle, it is of significant progress once realised and will hopefully pave the way for ever more exciting inventions and experimental demonstrations in the future. It is our belief that a true quantum computer and quantum communication network will be constructed through a combination of optical techniques and quantum memory using atoms.

Appendix A

Solutions to the paraxial wave equation

In this Appendix we derive the fundamental solution as well as two sets of higher-order solutions to the paraxial wave equation

$$\nabla_\perp^2 u(x, y, z) - 2ik\frac{\partial u}{\partial z} = 0, \tag{A.1}$$

where $\nabla_\perp^2 = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)$.

A.1 Gaussian Beam

In order to find the fundamental solution to the paraxial wave equation we use an ansatz which is an extension of the paraxial spherical wave of Eq. (2.26)

$$u_{\mathrm{G}}(r, z) = C \exp\left[-i\left(f(z) + \frac{ikr^2}{2q(z)}\right)\right], \tag{A.2}$$

where $r^2 = x^2 + y^2$ and C is a constant which depends on the amplitude of the Gaussian beam and is determined by the boundary conditions. By substituting this ansatz into the paraxial wave equation we find

$$\exp\left(-if(z)\right)\nabla_\perp^2\left[\exp\left(-\frac{ikr^2}{2q(z)}\right)\right] - 2ik\frac{\partial}{\partial z}\left[\exp\left(-if(z)\right)\exp\left(-\frac{ikr^2}{2q(z)}\right)\right] = 0. \tag{A.3}$$

If we carry out the differential operations in the previous equation we obtain

$$\exp\left(-if(z)\right)\exp\left(-\frac{ikr^2}{2q(z)}\right)\left[-\frac{2ik}{q(z)} - \frac{k^2r^2}{q^2(z)} - 2ik\left(-i\frac{\partial}{\partial z}f(z) + \frac{ikr^2}{2q^2(z)}\frac{\partial}{\partial z}q(z)\right)\right] = 0, \tag{A.4}$$

which can be simplified to give

$$2k\frac{\partial}{\partial z}f(z) + \frac{2ik}{q(z)} + \frac{k^2r^2}{q^2(z)}\left[1 - \frac{\partial}{\partial q(z)}\right] = 0. \tag{A.5}$$

By comparing equal powers in r we obtain the relations

$$\frac{k^2 r^2}{q^2(z)} \left[1 - \frac{\partial}{\partial q(z)} \right] = 0,$$

$$\frac{\partial}{\partial z} q(z) = 1, \tag{A.6}$$

and

$$2k \frac{\partial}{\partial z} f(z) + \frac{2ik}{q(z)} = 0,$$

$$\frac{\partial}{\partial z} f(z) = -\frac{i}{q(z)}. \tag{A.7}$$

Integration of Eq. (A.6) yields,

$$q(z) = z + q_0. \tag{A.8}$$

For convenience we now introduce two real parameters $R(z)$ and $w(z)$ which are related to the complex parameter $q(z)$ as follows

$$\frac{1}{q(z)} = \frac{1}{R(z)} - i \frac{\lambda}{\pi w^2(z)}. \tag{A.9}$$

To standardise the constants of integration we assume a plane wavefront at an arbitrary reference point $z = 0$, i.e. we take $R(0) \equiv \infty$. Thus

$$\frac{1}{R(0)} = 0 \quad \text{and} \quad -i \frac{\lambda}{\pi w_0^2} = \frac{1}{q_0}, \tag{A.10}$$

where $w_0 \equiv w(0)$. This leads to

$$q_0 = \frac{i \pi w_0^2}{\lambda} = i z_{\mathrm{R}} \tag{A.11}$$

where $z_{\mathrm{R}} = \pi w_0^2 / \lambda$ is the Rayleigh range, so that $q(z)$ is now given by

$$q(z) = z + i z_{\mathrm{R}}. \tag{A.12}$$

Eq. (A.9) can be written in terms of the Rayleigh range as

$$\frac{1}{q(z)} = \frac{1}{R(z)} - i \frac{\lambda}{\pi w^2(z)} = \frac{1}{z + i z_{\mathrm{R}}} = \frac{z - i z_{\mathrm{R}}}{z^2 + z_{\mathrm{R}}^2}. \tag{A.13}$$

Equating the real and imaginary parts of $q(z)$ we obtain

$$\frac{1}{R(z)} = \frac{z}{z^2 + z_{\mathrm{R}}^2} \quad \text{and} \quad -\frac{i\lambda}{\pi w^2(z)} = -\frac{i z_{\mathrm{R}}}{z^2 + z_{\mathrm{R}}^2}, \tag{A.14}$$

or in the standard form

$$R(z) = z \left[1 + \left(\frac{z_{\mathrm{R}}}{z} \right)^2 \right], \tag{A.15}$$

$$w^2(z) = w_0^2 \left[1 + \left(\frac{z}{z_{\mathrm{R}}} \right)^2 \right]. \tag{A.16}$$

To calculate the complex phase shift at a distance z away from the waist we can insert Eq. (A.12) into Eq. (A.7),

$$\frac{\partial}{\partial z} f(z) = -\frac{i}{q(z)} = -\frac{i}{z + iz_{\mathrm{R}}}. \tag{A.17}$$

Integration of this equation yields

$$f(z) = -i \ln\left[1 - i\frac{z}{z_{\mathrm{R}}}\right] = -i\left[\ln\sqrt{1 + \left(\frac{z}{z_{\mathrm{R}}}\right)^2} - i\arctan\left(\frac{z}{z_{\mathrm{R}}}\right)\right]. \tag{A.18}$$

which allows us to obtain an expression for the first phase factor in Eq. (A.2),

$$\exp[-if(z)] = \frac{\exp\left[i\arctan\left(\frac{z}{z_{\mathrm{R}}}\right)\right]}{\sqrt{1 + \left(\frac{z}{z_{\mathrm{R}}}\right)^2}} = \frac{w_0}{w(z)}\exp\left[i\Phi^{\mathrm{G}}(z)\right], \tag{A.19}$$

where we have used Eq. (A.16) and $\Phi^{\mathrm{G}}(z) = \arctan\left(\frac{z}{z_{\mathrm{R}}}\right)$ is the Gouy phase. The final form of the Gaussian beam is, therefore,

$$\begin{aligned}
u_{\mathrm{G}}(r, z) &= C_{\mathrm{G}}\exp\left[-i\left(f(z) + \frac{kr^2}{2q(z)}\right)\right] \\
&= C_{\mathrm{G}}\frac{w_0}{w(z)}\exp\left[-\frac{r^2}{w^2(z)}\right]\exp\left[-i\left(\frac{kr^2}{2R(z)} - \Phi^{\mathrm{G}}(z)\right)\right].
\end{aligned} \tag{A.20}$$

A.2 Hermite-Gaussian modes

In order to obtain higher-order solutions (in rectangular coordinates) to the paraxial wave equation we can extend the Gaussian solution as

$$u(r, z) = F(x, y, z)u_{\mathrm{G}}(r, z) \tag{A.21}$$

$$= g\left(\frac{x}{w(z)}\right) h\left(\frac{x}{w(z)}\right) \exp[-iP(z)]u_{\mathrm{G}}(r, z). \tag{A.22}$$

If we substitute this ansatz into Eq. (A.1) we obtain

$$F\nabla_\perp^2 u_{\mathrm{G}} + 2\left[\nabla_\perp F \cdot \nabla_\perp u_{\mathrm{G}}\right] + u_{\mathrm{G}}\nabla_\perp^2 F - 2iku_{\mathrm{G}}\frac{\partial}{\partial z}F - 2ikF\frac{\partial}{\partial z}u_{\mathrm{G}} = 0, \tag{A.23}$$

where we have excluded the arguments of the functions F and u_{G} for simplicity. The sum of the first and fifth terms satisfies the paraxial wave equation so Eq. (A.23) reduces to

$$\frac{g''}{g} + 2ik\left[\frac{dw}{dz} - \frac{w}{q}\right]x\frac{g'}{g} + \frac{h''}{h} + 2ik\left[\frac{dw}{dz} - \frac{w}{q}\right]y\frac{h'}{h} - 2kw^2\frac{dP}{dz} = 0, \tag{A.24}$$

where the primes indicate derivatives with respect to the argument of the function and $q \equiv q(z)$ is the complex beam parameter. We can use Eq. (A.9), Eq. (A.15) and Eq. (A.16) to simplify the following factor:

$$\frac{dw}{dz} - \frac{w}{q} = \frac{w}{R} - w\left(\frac{1}{R} - \frac{i\lambda}{\pi w^2}\right) = \frac{i\lambda}{\pi w}. \tag{A.25}$$

This allows us to reduce Eq. (A.24)

$$\frac{g''}{g} - 4\xi\frac{g'}{g} + \frac{h''}{h} - 4\eta\frac{h'}{h} - 2kw^2\frac{dP}{dz} = 0, \tag{A.26}$$

where $\xi = x/w$ and $\eta = y/w$.

Note that we are trying to find a solution to the paraxial wave equation that describes HG modes propagating in space. We must therefore be able to identify the functions g and h with the Hermite polynomials. A Hermite polynomial of order n has the following differential equation

$$\frac{d^2}{dt^2}H_n(t) - 2t\frac{d}{dt}H_n(t) + 2nH_n(t) = 0. \tag{A.27}$$

We can see that Eq. (A.26) contains terms that resemble those found in the differential equation Eq. (A.27). With the simple change of variables $t = \sqrt{2}\xi = \sqrt{2}x/w$ and $\tau = \sqrt{2}\eta = \sqrt{2}y/w$, Eq. (A.26) is transformed into

$$\frac{1}{g}\left[\frac{d^2g}{dt^2} - 2t\frac{dg}{dt}\right] + \frac{1}{h}\left[\frac{d^2h}{d\tau^2} - 2\tau\frac{dh}{d\tau}\right] - kw^2\frac{dP}{dz} = 0. \tag{A.28}$$

Thus it is clear that we can write the functions g and h as Hermite polynomial

$$g\left(\frac{x}{w}\right) = H_n(t) = H_n\left(\sqrt{2}\frac{x}{w}\right), \tag{A.29}$$

$$h\left(\frac{y}{w}\right) = H_m(t) = H_m\left(\sqrt{2}\frac{y}{w}\right), \tag{A.30}$$

if we require that

$$kw^2 \frac{dP}{dz} = -2(n+m).$$
(A.31)

This requirement allows us to determine the phase factor $P(z)$ if we set $P(0) = 0$,

$$\frac{dP}{dz} = -\frac{2}{kw^2}(n+m) = -(n+m)\frac{z_R}{z_R^2 + z^2},$$
(A.32)

and integration yields

$$P(z) = -(m+n)\arctan\left(\frac{z}{z_R}\right).$$
(A.33)

Combining all these results we can write the HG mode as

$$u_{nm}^{HG} = C_{nm}^{HG}\frac{w_0}{w(z)}H_n\left(\frac{\sqrt{2}x}{w}\right)H_m\left(\frac{\sqrt{2}y}{w}\right)\exp\left(-\frac{r^2}{w^2(z)}\right)$$

$$\times \exp\left[-i\left(k\frac{r^2}{2R(z)} - \Phi^{HG}(z)\right)\right],$$
(A.34)

where $\Phi^{HG} = (n+m+1)\arctan\left(\frac{z}{z_R}\right)$ is the Gouy phase of the HG mode.

A.3   Laguerre-Gaussian mode

In cylindrical coordinates, the solution to the paraxial wave equation is called the LG beam since it can be expressed as the product of a Laguerre polynomial and the Gaussian function, as we will now show. Although it is possible to find this solution using the method of the previous section, we present a slightly different technique here which involves the separation of variables [99].

The paraxial wave equation in cylindrical coordinates is given by

$$\left[\frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial \phi^2} + 2ik\frac{\partial}{\partial r} \right] u(r, \phi, z) = 0. \tag{A.35}$$

We assume a solution of the form

$$u = h(r, z)\exp[-il\phi]u_{\mathrm{G}}(r, z), \tag{A.36}$$

where l is an integer and we use the form of u_{G} given in Eq. (A.2) for the substitutions that follow since the r and z components are clearly separated here, which simplifies the calculations. Substituting Eq. (A.36) into Eq. (A.35) gives

$$\frac{\partial^2 h}{\partial r^2} + \frac{1}{r}\frac{\partial h}{\partial r} - \frac{2ikr}{q}\frac{\partial h}{\partial r} - \frac{l^2}{r^2}h - 2ik\frac{\partial h}{\partial z} = 0, \tag{A.37}$$

where we have used Eq. (A.7). Since q is a function of z we cannot use separation of variables to solve the equation above. In order to use this method we must make the following change of variables

$$r' = \frac{\sqrt{2}}{w}r \quad \text{and} \quad z' = z. \tag{A.38}$$

With this new set of variables we can write the fifth term in Eq. (A.37) as

$$\begin{aligned}
-2ik\frac{\partial h}{\partial z} &= -2ik\left(\frac{\partial h}{\partial r'}\frac{\partial r'}{\partial z'} + \frac{\partial h}{\partial z} \right) \\
&= -2ik\left[-\frac{z}{z^2 + z_{\mathrm{R}}^2}\frac{\partial h}{\partial r'} + \frac{\partial h}{\partial z'} \right]
\end{aligned} \tag{A.39}$$

where we have used Eq. (A.16). After some simplifications and the use of Eq. (A.9) we arrive at

$$\frac{\partial^2 h}{\partial r'^2} + \frac{1}{r'}\frac{\partial h}{\partial r'} - 2r'\frac{\partial h}{\partial r'} - \frac{l^2}{r'^2}h - ik\omega^2\frac{\partial h}{\partial z'} = 0. \tag{A.40}$$

If $l \neq 0$ the value of the function h on the axis must be zero, so we can define this function as

$$h = r'^l G(r', z'). \tag{A.41}$$

Substituting this form into Eq. (A.40) we obtain

$$\frac{\partial^2 G}{\partial r'^2} + \frac{2l + 1 - 2r'^2}{r'}\frac{\partial G}{\partial r'} - 2lG - ik\omega^2\frac{\partial G}{\partial z'} = 0. \tag{A.42}$$

Note that we are trying to find a solution to the paraxial wave equation that describes LG modes propagating in space. We must, therefore, be able to identify the function h with a Laguerre polynomial. A Laguerre polynomial of order p has the following differential equation

$$\xi\frac{d^2}{d\xi^2}L_p^l(\xi) + (l + 1 - \xi)\frac{d}{d\xi}L_p^l(\xi) + pL_p^l(\xi) = 0. \tag{A.43}$$

Eq. (A.42) contains terms that resemble those in Eq. (A.43) but we still require a another change of variable to make this resemblance more visible. We use the simple transformation

$$\xi = r'^2, \tag{A.44}$$

and substitute this into Eq. (A.42) to obtain

$$\xi\frac{\partial^2 G}{\partial\xi^2} + (l + 1 - \xi)\frac{\partial G}{\partial\xi} - \frac{l}{2}G - \frac{ik\omega^2}{4}\frac{\partial G}{\partial z'} = 0. \tag{A.45}$$

This equation can now be solved by the separation of variables

$$G = M(\xi)Z(z'), \tag{A.46}$$

and substituting into Eq. (A.45) gives

$$\xi\frac{d^2 M}{d\xi^2} + (l + 1 - \xi)\frac{dM}{d\xi} + pM = 0, \tag{A.47}$$

$$\frac{dZ}{dz'} - \frac{2i(2p + l)}{kw^2}Z = 0, \tag{A.48}$$

where p is an integer. These equations now have the solutions

$$M(\xi) = L_p^l(\xi), \tag{A.49}$$

$$Z(z') = \exp\left[i(2p + l)\arctan\left(\frac{z}{z_\mathrm{R}}\right)\right], \tag{A.50}$$

where the Laguerre polynomial can be expressed as

$$L_p^l(\xi) = \sum_{k=0}^{p}\frac{(p + l)!(-\xi)^k}{(l + k)!k!(p - k)!}. \tag{A.51}$$

The amplitude distribution of the LG beam is thus

$$u_{pl}^\mathrm{LG}(r, \phi, z) = C_{pl}^\mathrm{LG}\frac{w_0}{w(z)}\left(\frac{\sqrt{2}r}{w(z)}\right)^l L_p^l\left(\frac{2r^2}{w^2(z)}\right)\exp\left(-\frac{r^2}{w^2(z)}\right)$$
$$\times \exp\left[-i\left(\frac{kr^2}{2R(z)} + l\phi - \Phi^\mathrm{LG}(z)\right)\right], \tag{A.52}$$

where $\Phi^\mathrm{LG} = (2p + l + 1)\arctan\left(\frac{z}{z_\mathrm{R}}\right)$ is the Gouy phase of the LG mode.

Appendix B

Sorters

The sorter is a Mach-Zehnder interferometer with a beam rotator (rotation angle α) and a phase plate (phase shift $\Delta\Phi_c$) in one arm. Depending on the values of α and $\Delta\Phi_c$ we can sort between individual photons according to their l values, which will then exit the interferometer either through output port 1 or 2.
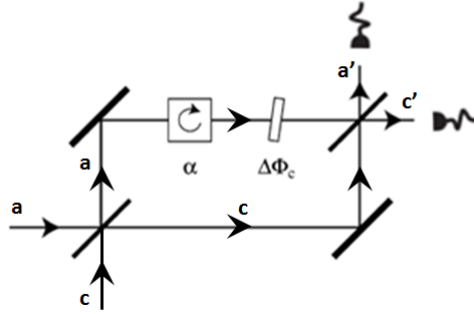

Figure B.1: A sorter to distinguish between different OAM states. Adapted from Fig. 2 in [76].

Let us first consider the input state coming from path a

$$|\Psi\rangle_a = \sum_l n_l |l\rangle_a. \tag{B.1}$$

The first optical element that transforms this state is the beam splitter. The action of the beam splitter is given by

$$|\Psi\rangle_a = \tfrac{1}{\sqrt{2}}\Big[|\Psi\rangle_a + |\Psi\rangle_c\Big],$$
$$|\Psi\rangle_c = \tfrac{1}{\sqrt{2}}\Big[-|\Psi\rangle_a + |\Psi\rangle_c\Big], \tag{B.2}$$

After the first beam splitter this becomes

$$|\Psi\rangle_a \xrightarrow{BS1} \tfrac{1}{\sqrt{2}}\Big[|\Psi\rangle_a + |\Psi\rangle_c\Big]$$
$$= \tfrac{1}{\sqrt{2}}\sum_l n_l\Big[|l\rangle_a + |l\rangle_c\Big], \tag{B.3}$$

The beam rotator (rotation angle α) and phase shifter (phase shift $\Delta\Phi_c$) in arm a causes the state to evolve to

$$\tfrac{1}{\sqrt{2}}\sum_l n_l\left[e^{i(l\alpha+\Delta\Phi_c)}|l\rangle_a + |l\rangle_c\right]. \tag{B.4}$$

The final optical element that this state encounter is the second beam splitter whose action is defined as

$$|l\rangle_a = \tfrac{1}{\sqrt{2}}\left(|l\rangle_{a'} + |l\rangle_{c'}\right),$$
$$|l\rangle_c = \tfrac{1}{\sqrt{2}}\left(-|l\rangle_{a'} + |l\rangle_{c'}\right), \tag{B.5}$$

so the state now becomes

$$\tfrac{1}{2}\sum_l n_l\left[e^{i(l\alpha+\Delta\Phi_c)}\left(|l\rangle_{a'} + |l\rangle_{c'}\right) + \left(-|l\rangle_{a'} + |l\rangle_{c'}\right)\right]$$

$$= \tfrac{1}{2}\sum_l n_l\left[\left(e^{i(l\alpha+\Delta\Phi_c)} - 1\right)|l\rangle_{a'} + \left(e^{i(l\alpha+\Delta\Phi_c)} + 1\right)|l\rangle_{c'}\right]$$

$$= \tfrac{1}{2}\sum_l n_l\left[e^{i(l\alpha+\Delta\Phi_c)/2}\left(e^{i(l\alpha+\Delta\Phi_c)/2} - e^{-i(l\alpha+\Delta\Phi_c)/2}\right)|l\rangle_{a'}\right.$$

$$\left. + e^{i(l\alpha+\Delta\Phi_c)/2}\left(e^{i(l\alpha+\Delta\Phi_c)/2} + e^{-i(l\alpha+\Delta\Phi_c)/2}\right)|l\rangle_{c'}\right]$$

$$= \sum_l n_l\left[ie^{i(l\alpha+\Delta\Phi_c)/2}\sin\left(\frac{l\alpha+\Delta\Phi_c}{2}\right)|l\rangle_{a'} + e^{i(l\alpha+\Delta\Phi_c)/2}\cos\left(\frac{l\alpha+\Delta\Phi_c}{2}\right)|l\rangle_{c'}\right]. \tag{B.6}$$

Therefore, by choosing different values for α and $\Delta\Phi_c$ we can sort between different OAM states $|l\rangle$. This will become clearer if we look at a case that is useful for us. Consider the input state

$$|\Psi\rangle_a = \tfrac{1}{\sqrt{2}}\Big(|l=-1\rangle_a + |l=1\rangle_a\Big). \tag{B.7}$$

If we take $\alpha = \Delta\Phi_c = 90°$ then we obtain

$$|\Psi\rangle_a \rightarrow -\tfrac{1}{\sqrt{2}}|l=1\rangle_{a'} + \tfrac{1}{\sqrt{2}}|l=-1\rangle_{c'}. \tag{B.8}$$

To understand Table 9.1 we have to rename $a' \rightarrow 3$ and $c' \rightarrow 1$

$$|\Psi\rangle_a \rightarrow \tfrac{1}{\sqrt{2}}|l=-1\rangle_1 - \tfrac{1}{\sqrt{2}}|l=1\rangle_3 \tag{B.9}$$

Now, let us consider the case where we have an input state coming from path c

$$|\Psi\rangle_c = \sum_l n_l |l\rangle_c. \tag{B.10}$$

After the first beam splitter this becomes

$$|\Psi\rangle_c \xrightarrow{BS1} \frac{1}{\sqrt{2}} \sum_l n_l \left[-|l\rangle_a + |l\rangle_c \right]. \tag{B.11}$$

The beam rotator (rotation angle α) and phase shifter (phase shift $\Delta\Phi_c$) in arm a causes this state to evolve to

$$\frac{1}{\sqrt{2}} \sum_l n_l \left[-e^{i(l\alpha+\Delta\Phi_c)} |l\rangle_a + |l\rangle_c \right]. \tag{B.12}$$

After the second beam splitter this state is transformed to

$$\begin{aligned}
&\frac{1}{2} \sum_l n_l \left[-e^{i(l\alpha+\Delta\Phi_c)} \left(|l\rangle_{a'} + |l\rangle_{c'} \right) + \left(-|l\rangle_{a'} + |l\rangle_{c'} \right) \right] \\
&= \frac{1}{2} \sum_l n_l \left[-\left(e^{i(l\alpha+\Delta\Phi_c)} + 1 \right) |l\rangle_{a'} - \left(e^{i(l\alpha+\Delta\Phi_c)} - 1 \right) |l\rangle_{c'} \right] \\
&= -\sum_l n_l e^{i(l\alpha+\Delta\Phi_c)/2} \left[\cos\left(\frac{l\alpha + \Delta\Phi_c}{2} \right) |l\rangle_{a'} + i\sin\left(\frac{l\alpha + \Delta\Phi_c}{2} \right) |l\rangle_{c'} \right]. \tag{B.13}
\end{aligned}$$

Now, for our input state

$$|\Psi\rangle_c = \frac{1}{\sqrt{2}} \left(|l = -1\rangle_c + |l = 1\rangle_c \right) \tag{B.14}$$

and $\alpha = \Delta\Phi_c = 90°$, we obtain

$$|\Psi\rangle_c \rightarrow -\frac{1}{\sqrt{2}} |l = -1\rangle_{a'} + \frac{1}{\sqrt{2}} |l = 1\rangle_{c'}. \tag{B.15}$$

Once again, we have to rename $a' \rightarrow 3$ and $c' \rightarrow 1$,

$$|\Psi\rangle_c \rightarrow \frac{1}{\sqrt{2}} |l = 1\rangle_1 - \frac{1}{\sqrt{2}} |l = -1\rangle_3. \tag{B.16}$$

If we relabel $a \rightarrow b$ and $c \rightarrow d$ and rename $a' \rightarrow 4$ and $c' \rightarrow 2$ we will obtain the remaining components in Table 9.1.

Bibliography

[1] A. Aspect, J. Dalibard, and G. Roger, "Experimental test of Bell's inequalities using time-varying analyzers," Phys. Rev. Lett. **49**, 1804–1807 (1982).

[2] A. Einstein, *The Born-Einstein Letters; Correspondence between Albert Einstein and Max and Hedwig Born from 1916 to 1955* (Walker, 1971).

[3] G. E. Moore, "Cramming more components onto integrated circuits," Electronics **8**, 114–117 (1965).

[4] P. W. Shor, "Algorithms for quantum computation: Discrete logarithms and factoring," in "Proc. 35th Annual Symposium on Foundations of Computer Science," , S. Goldwasser, ed. (IEEE Computer Society Press, USA, 1994), pp. 124–134.

[5] M. A. Nielsen and I. L. Chuang, *Quantum Information and Quantum Computation* (Cambridge University Press, United Kingdom, 2000).

[6] R. Raussendorf and H. J. Briegel, "A one-way quantum computer," Phys. Rev. Lett. **86**, 5188–5191 (2001).

[7] P. Walther, K. J. Resch, T. Rudolph, E. Schenck, H. Weinfurter, V. Vedral, M. Aspelmeyer, and A. Zeilinger, "Experimental one-way quantum computing," Nature **434**, 169–176 (2005).

[8] L. Allen, M. W. Beijersbergen, R. J. C. Spreeuw, and J. P. Woerdman, "Orbital angular momentum of light and the transformation of Laguerre-Gaussian laser modes," Phys. Rev. A **45**, 8185–8189 (1992).

[9] D. Kaszlikowski, P. Gnaciński, M. Żukowski, W. Miklaszewski, and A. Zeilinger, "Violations of local realism by two entangled N-dimensional systems are stronger than for two qubits," Phys. Rev. Lett. **85**, 4418–4421 (2000).

[10] D. Collins, N. Gisin, N. Linden, S. Massar, and S. Popescu, "Bell inequalities for arbitrarily high-dimensional systems," Phys. Rev. Lett. **88**, 040404 (2002).

[11] A. Mair, A. Vaziri, G. Weihs, and A. Zeilinger, "Entanglement of the orbital angular momentum states of photons," Nature **412**, 313–316 (2001).

[12] A. C. Dada, J. Leach, G. S. Buller, M. J. Padgett, and E. Andersson, "Experimental high-dimensional two-photon entanglement and violations of generalised bell inequalities," Nature Physics **7**, 677–680 (2011).

[13] J. C. Maxwell, "A dynamical theory of the electromagnetic field," Phil. Trans. R. Soc. Lond. **155**, 459–512 (1865).

[14] J. H. Poynting, "The wave motion of a revolving shaft, and a suggestion as to the angular momentum in a beam of circularly polarized light," Proc. Roy. Soc. London Ser. A **82**, 560–567 (1909).

[15] R. A. Beth, "Mechanical detection and measurement of the angular momentum of light," Phys. Rev. **50**, 115–125 (1936).

[16] L. Allen, M. J. Padgett, and M. Babiker, "The orbital angular momentum of light," (Elsevier, 1999), pp. 291–372.

[17] A. E. Siegman, *Lasers* (University Science Books, Mill Valley, Calif., 1986).

[18] H. Kogelnik and T. Li, "Laser beams and resonators," Proc. IEEE **54**, 1314–1317 (1966).

[19] R. P. C. GmbH., "Encyclopedia of laser physics and technology," `http://www.rp-photonics.com/encyclopedia.html` (2011).

[20] M. J. Padgett and L. Allen, "The poynting vector in Laguerre-Gaussian laser modes," Opt. Comm. **121**, 36–40 (1995).

[21] A. Freise and K. Strain, "Interferometer techniques for gravitational-wave detection," Living Reviews in Relativity **13** (2010).

[22] K. Palmer, "Angular momentum of electromagnetic radiation," Master's thesis, Uppsala University (2007).

[23] A. Dudley, National Laser Centre, CSIR, PO Box 395, Pretoria, South Africa, 0001 (personal communication, 2010).

[24] E. Abramochkin and V. Volostnikov, "Beam transformations and nontransformed beams," Opt. Comm. **83**, 123–135 (1991).

[25] E. Karimi, "Generation and manipulation of laser beams carrying orbital angular momentum for classical and quantum information applications," Ph.D. thesis, Universita degli Studi di Napoli Federico II (2009).

[26] L. Allen and M. J. Padgett, "The poynting vector in Laguerre-Gaussian beams and the interpretation of their angular momentum density," Opt. Comm. **184**, 67–71 (2000).

[27] M. J. Padgett and L. Allen, *Twisted Photons: Applications of Light with Orbital Angular Momentum* (Wiley-VCH, Germany, 2011), chap. the Orbital Angular Momentum of Light: An Introduction, pp. 1–12.

[28] J. Courtial and M. J. Padgett, "Limit to the orbital angular momentum per unit energy in a light beam that can be focussed onto a small particle," Opt. Comm. **173**, 269–274 (2000).

[29] M. Planck, "Entropy and temperature of radiant heat," Ann. Phys. **1**, 719–737 (1901).

[30] A. Einstein, "Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt," Ann. Phys. **17**, 132–148 (1905).

[31] G. N. Lewis, "The conservation of photons," Nature **118**, 874–875 (1936).

[32] G. Nienhuis and L. Allen, "Paraxial wave optics and harmonic oscillators," Phys. Rev. A **48**, 656–665 (1993).

[33] N. K. Langford, "Encoding, manipulating and measuring quantum information in optics," Ph.D. thesis, University of Queensland (2007).

[34] G. Nienhuis, *Structured Light and Its Applications: An introduction to Phase-Structured Beams and Nanoscale Optical Forces* (Elsevier, United Kingdom, 2008), chap. Angular Momentum and Vortices in Optics, pp. 19–62.

[35] F. G. Calvo, A. Picon, and E. Bagan, "Quantum field theory of photons with orbital angular momentum," Phys. Rev. A **73**, 013805 (2006).

[36] A. Aiello and J. P. Woerdman, "Exact quantization of a paraxial electromagnetic field," Phys. Rev. A **72**, 060101 (2005).

[37] S. P. Walborn, C. H. Monken, S. Pádua, and P. H. S. Ribeiro, "Spatial correlations in parametric down-conversion," Phys. Rep. **495**, 87–139 (2010).

[38] D. Stoler, "Operator methods in physical optics," J. Opt. Soc. Am. **71**, 334–341 (1981).

[39] S. J. van Enk and G. Nienhuis, "Eigenfunction description of laser beams and orbital angular momentum of light," Opt. Comm. **94**, 147–158 (1992).

[40] Y. Y. Schechner and J. Shamir, "Parameterization and orbital angular momentum of anisotropic dislocations," J. Opt. Soc. Am. A **13**, 967–973 (1996).

[41] M. W. Beijersbergen, L. Allen, H. van der Veen, and J. P. Woerdman, "Astigmatic laser mode converters and transfer of orbital angular momentum," Opt. Comm. **96**, 123–132 (1993).

[42] D. S. Tasca, R. M. Gomes, F. Toscano, P. H. S. Ribeiro, and S. P. Walborn, "Continuous-variable quantum computation with spatial degrees of freedom of photons," Phys. Rev. A **83**, 052325 (2011).

[43] C. K. Hong, Z. Y. Ou, and L. Mandel, "Measurement of subpicosecond time intervals between two photons by interference," Phys. Rev. Lett. **59**, 2044–2046 (1987).

[44] J. G. Rarity and P. R. Tapster, "Experimental violation of bells inequality based on phase and momentum," Phys. Rev. Lett. **64**, 24952498 (1990).

[45] D. V. Strekalov, T. B. Pitman, A. V. Sergienko, Y. H. Shih, and P. G. Kwiat, "Postselection-free energy-time entanglement," Phys. Rev. A. **54**, R1–R4 (1996).

[46] J. Brendel, N. Gisin, W. Tittel, and H. Zbinden, "Pulsed energy-time entangled twin-photon source for quantum communication," Phys. Rev. Lett. **82**, 25942597 (1999).

[47] D. N. Klyshko, *Photons and nonlinear optics* (Gordon and Breach Science Publishers, Great Britain, 1988).

[48] C. Staus, *A New Configuration for Efficient Far-Infrared Difference Frequency Generation: Nested Waveguides* (BiblioBazaar LLC, Charleston, USA, 2011).

[49] C. I. O. Tamayo, "Spatial characterization of two-photon states," Ph.D. thesis, Universitat Politècnica de Catalunya, Barcelona (2009).

[50] P. H. Shun, "Towards a high quality polarization-entangled multi-photon source," Master's thesis, National University of Singapore (2009).

[51] P. G. Kwiat, "Nonclassical effects from spontaneous parametric down conversion: Adventures in quantum wonderland," Ph.D. thesis, University of California at Berkeley (1993).

[52] Y. Nazirizadeh, "Compact source for polarization entangled photon pairs," Master's thesis, Max-Planck-Institute for Quantum Optics (2005).

[53] C. C. Gerry and P. L. Knight, *Introductory Quantum Optics* (Cambridge University Press, 2005).

[54] M. Rådmark, "Photonic quantum information and experimental tests of foundations of quantum mechanics," Ph.D. thesis, Stockholm University, Sweden (2010).

[55] Y. Shih, "Entangled biphoton source - property and preparation," Rep. Prog. Phys. **66**, 1009 (2003).

[56] D. R. Hamel, "Realization of novel entangled photon sources using periodically poled materials," Master's thesis, University of Waterloo, Canada (2010).

[57] P. G. Kwiat, K. Mattle, H. Weinfurter, A. Zeilinger, A. V. Sergienko, and Y. Shih, "New high-intensity source of polarization-entangled photon pairs," Phys. Rev. Lett. **75**, 4337–4341 (1995).

[58] P. G. Kwiat, E. Waks, A. G. White, I. Appelbaum, and P. H. Eberhard, "Ultrabright source of polarization-entangled photons," Phys. Rev. A **60**, R773–R776 (1999).

[59] N. G. Rodriguez, "Measurement of the spatial shape of photons," Ph.D. thesis, Universitat Politècnica de Catalunya (2010).

[60] A. M. Yao, "Angular momentum decomposition of entangled photons with an arbitrary pump," New J. Phys. **13**, 053048 (2011).

[61] A. Forbes, National Laser Centre, CSIR, PO Box 395, Pretoria, South Africa, 0001 (personal communication, 2011).

[62] P. Hariharan, *Basics of Holography* (Cambridge University Press, United Kingdom, 2002).

[63] E. Yao, S. Franke-Arnold, J. Courtial, M. J. Padgett, and S. M. Barnett, "Observation of quantum entanglement using spatial light modulators," Opt. Express **14**, 13089–13094 (2006).

[64] C. Guo, X. Liu, X. Ren, and H. Wang, "Optimal annular computer-generated holograms for the generation of optical vortices," J. Opt. Soc. Am. A **22**, 385–390 (2005).

[65] N. R. Heckenberg, R. McDuff, C. P. Smith, H. Rubinsztein-Dunlop, and M. J. Wegener, "Laser beams with phase singularities," Opt. Quant. Elect. **24**, S951–S962 (1992). 10.1007/BF01588597.

[66] J. F. James, *A Student's Guide to Fourier Transforms: With Applications in Physics and Engineering* (Cambridge University Press, United Kingdom, 2011).

[67] M. R. Dennis, K. O'Holleran, and M. J. Padgett, "Chapter 5 singular optics: Optical vortices and polarization singularities," (Elsevier, 2009), pp. 293–363.

[68] J. Leach, M. J. Padgett, S. M. Barnett, S. Franke-Arnold, and J. Courtial, "Measuring the orbital angular momentum of a single photon," Phys. Rev. Lett. **88**, 257901 (2002).

[69] B. Jack, J. Leach, H. Ritsch, S. M. Barnett, M. J. Padgett, and S. Franke-Arnold, "Precise quantum tomography of photon pairs with entangled orbital angular momentum," New J. Phys. **11**, 103024 (2009).

[70] B. Jack, A. M. Yao, J. Leach, J. Romero, S. Franke-Arnold, D. G. Ireland, S. M. Barnett, and M. J. Padgett, "Entanglement of arbitrary superpositions of modes within two-dimensional orbital angular momentum state spaces," Phys. Rev. A **81**, 043844 (2010).

[71] C. Tamm and C. O. Weiss, "Bistability and optical switching of spatial patterns in a laser," J. Opt. Soc. Am. B **7**, 1034–1038 (1990).

[72] M. J. Padgett, J. Arlt, and N. Simpson, "An experiment to observe the intensity and phase structure of Laguerre-Gaussian laser modes," (1996).

[73] L. Allen, J. Courtial, and M. J. Padgett, "Matrix formulation for the propagation of light beams with orbital and spin angular momenta," (1999).

[74] J. Courtial, D. A. Robertson, K. Dholakia, L. Allen, and M. J. Padgett, "Rotational frequency shift of a light beam," Phys. Rev. Lett. **81**, 4828 (1998).

[75] B. R. Gadway, J. W. Noe, and M. G. Cohen, "Universal quantum gates for order n=1 orbital angular momentum states," (Optical Society of America, 2008), p. JSuA1.

[76] H. Wei, X. Xue, J. Leach, M. J. Padgett, S. M. Barnett, S. Franke-Arnold, E. Yao, and J. Courtial, "Simplified measurement of the orbital angular momentum of single photons," Opt. Comm. **223**, 117–122 (2003).

[77] M. J. Padgett and J. Courtial, "Poincaré-sphere equivalent for light beams containing orbital angular momentum," Opt. Lett. **24**, 430–432 (1999).

[78] E. Schrödinger, "Die gegenwärtige situation in der quantenmechanik," Naturwissenschaften **23**, 807 (1935).

[79] L. K. Grover, "A fast quantum mechanical algorithm for database search," in "Proc. of the 28th ACM STOC," (1996), pp. 212–219.

[80] D. P. DiVincenzo, "The physical implementation of quantum computation," Fortschritte der Physik **48**, 771–784 (2000).

[81] R. Prevedel, "Experimental all-optical one-way quantum computing," Ph.D. thesis, Universität Wien (2008).

[82] E. Knill, R. Laflamme, and G. Milburn, "A scheme for efficient quantum computation with linear optics," Nature **409**, 46–52 (2001).

[83] M. Nielsen, "Cluster-state quantum computation," Rep. Math. Phys. **57**, 147–161 (2006).

[84] D. E. Browne and H. J. Briegel, *Lectures on quantum information* (Wiley-VCH, 2007), chap. One-way quantum computation – a tutorial introduction, pp. 359–380.

[85] X. Zhou, D. W. Leung, and I. L. Chuang, "Quantum logic gate constructions with one-bit teleportation," Tech. rep. (2000).

[86] M. Hein, W. Dür, J. Eisert, R. Raussendorf, M. V. der Nest, and H. J. Briegel, "Entanglement in graph states and its applications," in "Proceedings of International School of Physics Enrico Fermi," , vol. 162, G. Casati, D. Shepelyansky, P. Zoller, and G. Benenti, eds., Società Italiana di Fisica (IOS Press, Italy, 2006), vol. 162, pp. 115–212.

[87] D. Gottesman, "An introduction to quantum error correction and fault-tolerant quantum computation," in "Proceedings of symposia in applied mathematics," , vol. 68, S. J. Lomonaco, ed. (American Mathematical Society, Washington, DC, 2009), vol. 68, pp. 13–61.

[88] R. Raussendorf, D. E. Browne, and H. J. Briegel, "The one-way quantum computer – a non-network model of quantum computation," J. Mod. Opt. **49**, 1299 (2002).

[89] R. Raussendorf, D. E. Browne, and H. J. Briegel, "Measurement-based quantum computation on cluster states," Phys. Rev. A **68**, 022312 (2003).

[90] E. T. Campbell and J. Fitzsimons, "An introduction to one-way quantum computing in distributed architectures," arXiv:quant-ph/0906.2725 (2009).

[91] C. Simon and J. Pan, "Polarization entanglement purification using spatial entanglement," Phys. Rev. Lett. **89**, 257901 (2002).

[92] A. M. Childs, D. W. Leung, and M. A. Nielsen, "Unified derivations of measurement-based schemes for quantum computation," Phys Rev. A **71** (2005).

[93] H. Weinfurter and M. Żukowski, "Four-photon entanglement from down-conversion," Phys. Rev. A **64**, 010102 (2001).

[94] S. Bone and M. Castro, "A brief history of quantum computing," `http://www.doc.ic.ac.uk/~nd/surprise_97/journal/vol4/spb3/` (1997).

[95] W. Hall, "Cluster state quantum computation with for many level systems," arXiv:quant-ph/0512130 (2006).

[96] D. L. Zhou, B. Zeng, Z. Xu, and C. P. Sun, "Quantum computation based on *d*-level cluster state," Phys. Rev. A **68**, 062303 (2003).

[97] R. Prevedel, P. Walther, F. Tiefenbacher, P. Böhi, R. Kaltenbaek, T. Jennewein, and A. Zeilinger, "High-speed linear optics quantum computing using active feed-forward," Nature **445**, 65–69 (2007).

[98] P. Böhi, R. Prevedel, T. Jennewein, A. Stefanov, F. Tiefenbacher, and A. Zeilinger, "Implementation and characterization of active feed-forward for deterministic linear optics quantum computing," Appl. Phys. B **89**, 499–505 (2007).

[99] K. Zhang and D. Li, *Electromagnetic Theory for Microwaves and Optoelectronics* (Springer, Germany, 1999).