Integral and Fractional Orbital Angular Momentum of Light

By

Jörg Bernhard Götte

A thesis submitted for the
degree of Doctor of Philosophy
at the University of Strathclyde

Department of Physics
December 2006

Abstract

Orbital angular momentum of light is a new field of research which is concerned with the mechanical and optical effects of light with a helical phase structure. In this thesis we ask fundamental questions on the properties of light carrying orbital angular momentum.

We discuss the uncertainty relation for angle and angular momentum on the example of orbital angular momentum of light. The lower bound in the angular uncertainty relation is state dependent, which requires a distinction between states satisfying the equality in the uncertainty relation and states giving a minimum in the uncertainty product. We examine these special states and their uncertainty product. We show that for both kinds of states, the uncertainty product can be surprisingly large.

We propose an experimentally testable criterion for an EPR paradox for orbital angular momentum and azimuthal angle. The criterion is designed for an experimental demonstration using orbital angular momentum of light. For the interpretation of future experimental results from the proposed setup, we include a model for the indeterminacies inherent to the angular position measurement. We show how angular apertures can be used to determine the angle, and we discuss the effects of this measurement on the proposed criterion. We show that for a class of aperture functions a demonstration of an angular EPR paradox, according to our criterion, is to be expected.

The quantum theory of rotation angles is generalised to non-integer values of the orbital angular momentum. This requires the introduction of an additional parameter, the orientation of a phase discontinuity associated with fractional values of the orbital angular momentum. We apply our formalism to the propagation of light modes with fractional orbital angular momentum in the paraxial and non-paraxial regime.

Acknowledgements

I wish to express my gratitude to:

My supervisors, Professor Stephen M. Barnett and Dr. Sonja Franke-Arnold, for their patience, support and guidance.

Dr. Roberta Zambrini of the Universitat de las Illes Baleares, Palma de Mallorca, for her enthusiasm and encouragement

Professor Miles Padgett, Dr. Eric Yao and Dr. Jonathan Leach from the University of Glasgow for helpful discussions and their insight on the experimental aspects of orbital angular momentum of light

Dr. Paul M Radmore from the University College London for his expertise on the subject of finding approximate differential equations with exact solutions

Professor Sir Michael Berry from the University of Bristol for his insight in light emerging from fractional phase steps

Thomas Brougham, Sarah Croke, David Grant, Dr. Kieran Hunter, Dr. Graeme McCartney, Douglas Murray, Dr. Eng-Kiang Tan, Dr. Nicholas Whitlock, who all are or were fellow PhD students, and with whom I had the pleasure of sharing an office, for discussions on, and more importantly, off physics

Dr. Richard Martin and Dr. Andrew Scroggie, for administering the CNQO computer suite and for their help with computational worries

My fiancée, for enduring the stressful times and the longest days in the desperate try to finish this thesis before our wedding day

My family, for their never ending support and encouragement since my earliest days

Contents

List of Figures

I Chapter

Introduction

When a scientific work tries to define orbital angular momentum of light this is often done by setting it against spin angular momentum. This is a rather natural way to proceed because it has long been known that light posseses spin angular momentum, which is observable in the polarisation of the light. But this alone would not necessarily explain why for light, spin angular momentum is the more familiar concept. After all, the total angular momentum of light within the field theory of electrodynamics can be derived directly from the Maxwell equations (Maxwell, 1873). But it has become commonplace that the polarisation of light is widely used in experimental tests of the foundation of quantum mechanics (Aspect et al., 1982*b*) and also for optical implementations of quantum information processes (Bouwmeester et al., 2001). It is possible, for example, to create light entangled in the orthogonal polarisation states. This property has been used in the most accurate experiments to measure the violation of Bell inequalities (Aspect et al., 1982*a*). Entanglement also plays a crucial role in quantum information. The experimental implementations of protocols for quantum key distribution often make use of the orthogonal polarisation states (Gisin et al., 2002). In other words spin angular momentum has become a piece of technology; it is well understood how to generate light with spin angular momentum, even for entangled pairs of photons, and also the measurement of spin or polarisation of light is well studied. It is this degree of control which makes spin angular momentum so widely used in quantum information applications.

For orbital angular momentum of light this control became possible only after it was discovered by Allen et al. (1992) that light beams carry a well defined orbital angular momentum. Allen et al. (1992) showed that Laguerre-Gaussian light beams familiar from paraxial optics have a total angular momentum which separates into a spin angular momentum part, solely determined by the polarisation of the light, and an orbital angular momentum part, connected to the azimuthal phase structure. This discovery made orbital angular momentum of light subject to an intense theoretical and experimental study (Allen et al., 2003). The work on the theoretical part was focussed on understanding the separation of the total angular momentum into spin and orbital angular momentum, as the well-defined separation of Allen et al. (1992) had been derived only in the paraxial approximation. But with the introduction of the orbital angular momentum flux as defining quantity it became possible to extend this separation beyond paraxial optics (Barnett, 2002). On the experimental side Laguerre-Gaussian beams have been used in the context of optical tweezers, a tightly focussed laser beam used to trap and manipulate micron sized particles in three dimensions. At first only the annular intensity profile typical for Laguerre-Gaussian beams has

been used to trap particles more efficiently than with conventional laser beams, but it became soon possible to transfer the angular momentum of the light beam to the particle and to set it in a spinning motion around its axis (He et al., 1995). It is also possible to set particles in an orbiting motion as demonstrated by O'Neil et al. (2002). In this respect orbital angular momentum of light has become a piece of technology which finds wider use in other areas of science from atomic physics to microbiology.

Entanglement is a very puzzling property of quantum mechanics and a powerful tool for quantum information. It has been shown that photon pairs can be entangled in their orbital angular momentum states (Mair et al., 2001). As these states span a discrete Hilbert space of infinite dimensions the entanglement would be multi-dimensional rather than two-dimensional as for photons entangled in the two orthogonal polarisation states. Orbital angular momentum states of photons can thus be used as an optical implementation of quNits (Molina-Terizza et al., 2001), that is states carrying quantum information in a multi-dimensional Hilbert space. Applications in quantum information and fundamental questions on the foundations of quantum mechanics often require experiments on a single photon level. Whereas the experimental techniques are well known for spin angular momentum, an efficient, interferometric measurement of the angular momentum content at a single photon level became possible only recently (Leach et al., 2002, 2004*a*). A number of requirements to use orbital angular momentum in quantum information have thus been met. Indeed a free-space communication system which utilises the higher dimensions of orbital angular momentum has been experimentally implemented (Gibson et al., 2004).

In principle orbital angular momentum has the potential to become as widely used as spin angular momentum. There are, however, some difficulties with transmitting the information. So far, no way has been found to transmit different modes of light carrying orbital angular momentum in the same optical fibre. Also, the mechanical stress in optical fibres due to bends can change the orbital angular momentum content. Atmospheric turbulence affects the phase structure of light beams and limits the transmission range for free-space communication (Paterson, 2005). Measurements on orbital angular momentum will be more involved due to the higher dimensions of the underlying Hilbert space. On the other hand, orbital angular momentum of light is a new field of research compared to spin angular momentum and there are a number of fundamental questions still to be asked. Some of these questions will be answered in this thesis. Although the theoretical work in this thesis is always closely connected to an experimental realisation, the thesis is not concerned with a particular application of orbital angular momentum of light. It is nevertheless the hope of the author that some of the discoveries from this thesis lead to future applications which exploit the powerful properties of orbital angular momentum of light. Only when orbital angular momentum has become as familiar and widely used as spin angular momentum, it will be possible to define it on its own terms rather than setting it against spin angular momentum.

In CHAPTER II starting on page 5 we discuss the foundations for the orbital angular momentum within classical electrodynamics. We review the concept of a continuity equation for the total angular momentum of the electromagnetic field. From the continuity equation we derive an expression for the angular momentum flux. We

introduce Laguerre-Gaussian beams and other beams carrying orbital angular momentum, and we show that the total angular momentum flux separates into a spin and an orbital part.

Whereas chapter II thus covers the classical aspects of orbital angular momentum of light, CHAPTER III is concerned from page 39 on with the description of angle and angular momentum in quantum mechanics. The angle operator and its eigenstates require a limiting procedure to be correctly represented within the concept of Hilbert spaces. We therefore present the quantum theory of rotation angles and summarise some of the results which are relevant for the following chapters.

In CHAPTER IV starting on page 49 we look into the uncertainty relation for angular momentum and angle. In contrast to the uncertainty relation for linear position and momentum the lower bound in the angular uncertainty relation is state-dependant. States satisfying the equality therefore do not have to give a minimum in the uncertainty product. Moreover, these special states come in two varieties, with small and large angle uncertainty, depending on whether the angle uncertainty is smaller or larger than for a flat wavefunction. We discuss the four different classes of states by deriving the defining equations and we present their solutions.

Some of the analytic expressions of the special states in chapter IV find their use also in CHAPTER V beginning on page 92. In this chapter we look into the possibility to demonstrate an EPR paradox for angular momentum and angle. For an experimentally testable criterion we propose the use of angular apertures. The particular form of the aperture is determined by an angle probability density, which ties in with the special states from chapter IV.

In CHAPTER VI starting on page 124 we turn our attention to fractional orbital angular momentum. Light with fractional orbital angular momentum is created with help of an edge dislocation which induces a phase discontinuity. This edge dislocation destroys the rotational symmetry and light beams with fractional orbital angular momentum are thus characterised by a dark line in the direction of the dislocation. We present a quantum formulation of fractional orbital angular momentum based on the fundamental theory from chapter III on page 39, and we apply our formulation to the propagation of light with fractional orbital angular momentum.

The thesis concludes with CHAPTER VII in which we summarise the previous chapters and put them into the perspective of future work.

Bibliography

Allen, L., Barnett, S. M. & Padgett, M. J. (2003), *Optical Angular Momentum*, Institute of Physics Publishing, Ltd., Bristol. ISBN: 0-7503-0901-6.

Allen, L., Beijersbergen, M. W., Spreeuw, R. J. C. & Woerdman, J. P. (1992), 'Orbital angular momentum of light and the transformation of Laguerre-Gaussian modes', *Physical Review A* **45**(11), 8185–8190. Reprinted in (Allen et al., 2003, Paper 2.1). DOI: 0.1103/PhysRevA.45.8185.

Aspect, A., Dalibard, J. & Roger, G. (1982*a*), 'Experimental Test of Bell's Inequalities

Using Time-Varying Analyzers', *Physical Review Letters* **49**(25), 1804 - 1806. DOI: 10.1103/PhysRevLett.49.1804.

Aspect, A., Grangier, P. & Roger, G. (1982*b*), 'Experimental Realization of Einstein-Podolsky-Rosen-Bohm Gedankenexperiment: A New Violation of Bell's Inequalities', *Physical Review Letters* **49**(2), 91 – 94. DOI: 10.1103/PhysRevLett.49.91.

Barnett, S. M. (2002), 'Optical angular-momentum flux', *Journal of Optics B* **4**(2), S7–S16. DOI: 10.1088/1464-4266/4/2/361.

Bouwmeester, D., Ekert, A. & Zeilinger, A., eds (2001), *The Physics of Quantum Information*, Springer Verlag, Berlin.

Gibson, G., Courtial, J., Padgett, M. J., Vastnetsov, V., Pas'ko, V., Barnett, S. M. & Franke-Arnold, S. (2004), 'Free-space information transfer using light beams carrying orbital angular momentum', *Optics Express* **12**(22), 5448–5456.

Gisin, N., Ribordy, G., Tittel, W. & Zbinden, H. (2002), 'Quantum cryptography', *Reviews of Modern Physics* **74**(1), 145–195. DOI: 10.1103/RevModPhys.74.145.

He, H., Friese, M. E. J., Heckenberg, N. R. & Rubinsztein-Dunlop, H. (1995), 'Direct Observation of Transfer of Angular Momentum to Absorptive Particles from a Laser Beam with a Phase Singularity', *Physics* **75**(5), 826–829. DOI: 10.1103/PhysRevLett.75.826.

Leach, J., Courtial, J., Skeldon, K., Barnett, S. M., Franke-Arnold, S. & Padgett, M. J. (2004), 'Interferometric Methods to Measure Orbital and Spin, or the Total Angular Momentum of a Single Photon', *Physical Review Letters* **92**, 013601. DOI: 10.1103/PhysRevLett.92.013601.

Leach, J., Padgett, M. J., Barnett, S. M., Franke-Arnold, S. & Courtial, J. (2002), 'Measuring the Orbital angular momentum of a single Photon', *Physical Review Letters* **88**, 257901. DOI: 10.1103/PhysRevLett.88.257901.

Mair, A. E., Vaziri, A., Weihs, G. & Zeilinger, A. (2001), 'Entanglement of the orbital angular momentum states of photons', *Nature* **412**, 313–316. Reprinted in (Allen et al., 2003, Paper 8.1).

Maxwell, J. C. (1873), *A Treatise on Electricity and Magnetism*, Clarendon Press Series, Oxford at the Clarendon Press, Oxford. Reprinted in Maxwell (1998).

Molina-Terizza, G., Torres, J. P. & Torner, L. (2001), 'Management of the Angular Momentum of Light: Preparation of Photons in Multidimensional Vector States of Angular Momentum', *Physical Review Letters* **88**(1), 013601. DOI: 10.1103/PhysRevLett.88.013601.

O'Neil, A. T., MacVicar, I., Allen, L. & Padgett, M. J. (2002), 'Intrinsic and Extrinsic Nature of the Orbital Angular Momentum of a Light Beam', *Physical Review Letters* **88**(5), 053601. DOI: 10.1103/PhysRevLett.88.053601.

Paterson, C. (2005), 'Atmospheric Turbulence and Orbital Angular Momentum of Single Photons for Optical Communication', *Physical Review Letters* **94**, 153901. DOI: 10.1103/PhysRevLett.94.153901.

II Chapter

Light beams carrying orbital angular momentum

For optical fields the notion of a total angular momentum has long been known. But the concept of a light beam carrying orbital angular momentum had been unfamiliar until it was discovered that Laguerre-Gaussian beams, within the paraxial approximation, carry a well defined orbital angular momentum. This discovery started the modern interest in orbital angular momentum of light. In this chapter we discuss the theoretical framework of orbital angular momentum of light in terms of fields and light beams and how to generate them.

1 Introduction

For Aristotle comets[1] were hairy objects not travelling the heavenly spheres but earthly exhalations reaching the upper atmosphere. This opinion prevailed through the Middle Ages, though from the 16th century on astronomers observed that comets circle the sun on elliptical orbits. At the end of the 17th century Newton applied his law of gravitation to the motion of comets which provided an explanation for the observed orbits. By then at the latest it was accepted that the Aristotelian hairs always point away from the sun and that they are not hairs but rather tails of evaporating material reflecting the sunlight. Even if this does not entirely conform with the modern explanation of comet tails due to solar wind and radiation pressure, it shows that the concept of light having mechanical properties was not unthinkable for the early physicists.

A quantitative treatment of the mechanical effects became possible only after light had been integrated into Maxwell's dynamical theory of electromagnetic waves. With this theory Poynting derived a continuity equation for the energy in the electromagnetic field (Poynting, 1884). After Heaviside introduced the vectorial notation for the Maxwell equations this continuity equation could be written in its modern form using the Poynting vector. Interestingly, the linear momentum density in the electromagnetic field is also given by the Poynting vector apart from constant factors depending on the chosen system of units. Poynting also derived an expression for the angular momentum of circular polarised light by means of a mechanical analogue in the form of a rotating shaft (Poynting, 1909). Later, Poynting's expression has been verified by measuring the torque on a quarter wave-plate due to circular polarised light (Beth, 1936). By this time quantum mechanics had been firmly established and Beth

[1] from Greek ἡ κόμη, the hair or κομήτης, long-haired

showed that the same quantitative result is obtained whether the torque is calculated classically or from the assumption that each photon carries an angular momentum of $\hbar$.

While spin angular momentum in light beams was thus established in terms of circular polarised light, OAM of light beams remained an unfamiliar concept. But this does not mean that orbital angular momentum of light in general has not been studied. In particular the concept of orbital angular momentum of optical fields has a substantial literature. In fact, the continuity equation for the angular momentum of an electromagnetic field seems to suggest that there is only orbital angular momentum in the electromagnetic field and no spin angular momentum (Römer & Forger, 1993). The total angular momentum density is given by the cross product of the radial vector and the linear momentum density, which is a form typical for orbital angular momentum. This surprising result shows the difficulties in splitting the total angular momentum in a spin and orbital part for any non-specific form of the electromagnetic field. But Allen et al. (1992) discovered that light beams, familiar from paraxial optics or laser theory and realisable in a laboratory, can carry a well-defined quantity of orbital angular momentum. This discovery started the current scientific activity on optical angular momentum, which has remained a lively field of research ever since.

In this introductory chapter we review the concept of orbital angular momentum of light beams. Starting from the Maxwell equations we look at the continuity equation for linear and angular momentum of the electromagnetic field. We then derive the exact Helmholtz equation and the paraxial approximation to study light beams with orbital angular momentum in the paraxial and non-paraxial regime. This allows us to establish that light beams with a particular azimuthal phase structure have orbital angular momentum. At the end of this chapter we will describe some methods to generate these beams in a laboratory.

2 Mechanical properties of optical fields

The set of Maxwell equations describes the electromagnetic field including its mechanical properties. The conservation of the mechanical properties such as energy, linear and angular momentum are expressed by continuity equations. These equations describe the change of quantities in a continuum and are therefore stated in terms of densities. The rate of change for a density ρ is given by the divergence of the respective flux density $\mathbf{j}$ and the source density q:

$$\frac{\partial \rho^{\mathrm{a}}}{\partial t} + \nabla \cdot \mathbf{j}^{\mathrm{a}} = q^{\mathrm{a}}, \tag{II.1}$$

where the superscript 'a' indicates a locally conserved quantity. The charge density ρ and the charge flux density $\mathbf{j}$ will be written without superscript. The source density for the electric charge is zero as the electric charge is globally conserved. The relevant continuity equation for the study of orbital angular momentum governs the local conversation of the total angular momentum. However, for the electromagnetic field the continuity equations for the mechanical quantities are intricately linked; the flux density of the energy is also the density for the linear momentum and the angular momentum densities are directly related to the linear momentum densities. In the

following we therefore review the continuity equations for energy, linear and angular momentum of the electromagnetic field.

2.1 Energy of the electromagnetic field

Electromagnetic energy can be converted into mechanical or thermal energy. This conversion of energy will be mediated through work exerted on a charge through the Lorentz force

$$\mathbf{F} = q\mathbf{E} + \eta \mathbf{v} \times \mathbf{B}, \tag{II.2}$$

which is stated here for a point charge q with velocity $\mathbf{v}$. The constant η is determined by the choice of the unit system. In this chapter we will use vectorial notation along with notation in components. The Cartesian component of a vector $\mathbf{a}$ will be denoted by a_i, where $i = 1, 2, 3$ represents the x, y or z component respectively. The Cartesian components of the Lorentz force may thus be we written as

$$F_i = qE_i + \eta \epsilon_{ijk} v_j B_k, \tag{II.3}$$

where we have used the Levi-Civita symbol ϵ_{ijk} to write the cross product $\mathbf{v} \times \mathbf{B}$ in components. The Levi-Civita symbol is equal to 1 if the combination of the indices ijk is an even permutation of 123 and it is equal to -1 for an odd permutation. Otherwise the Levi-Civita symbol is equal to zero. We also use the convention of implicitly summing over doubly occurring indices. The work exerted on a point charge between two arbitrary times t_0 and t_1 is thus given by

$$W(t_1, t_2) = \int_{t_1}^{t_2} \mathrm{d}t \, v_i F_i = q \int_{t_1}^{t_2} v_i E_i. \tag{II.4}$$

The magnetic field does not contribute to the exerted work as the cross product $\mathbf{v} \times \mathbf{B}$ is perpendicular to $\mathbf{v}$. For a continuous distribution of charges ρ in a velocity field $\mathbf{v}$ the energy of the electromagnetic fields is expressed as

$$q^{\mathrm{E}} = -j_i E_i = -\mathbf{j} \cdot \mathbf{E}, \tag{II.5}$$

where the charge flux density $\mathbf{j}$ is given by $\rho \mathbf{v}$. The exerted work reduces the energy of the electric field which explains the minus sign in the expression above. The expression in Eq. (II.5) has the dimension of power which is lost due to exerted work. The loss in power is balanced by the rate of change of the energy of the electromagnetic field and the energy flux. But in the absence of any charged particles the continuity equation has no source density term and the rate of change of the energy density is balanced by the divergence of the energy density flux only.

To find the expressions for the energy density and the energy flux density we can use the Maxwell equations for external charge and flux densities (Maxwell, 1873;

Jackson, 1998)

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}, \tag{II.6a}$$

$$\nabla \times \mathbf{E} = -\eta \frac{\partial \mathbf{B}}{\partial t}, \tag{II.6b}$$

$$\nabla \cdot \mathbf{B} = 0, \tag{II.6c}$$

$$\nabla \times \mathbf{B} = \eta \mu_0 \left(\mathbf{j} + \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right). \tag{II.6d}$$

The constant η has been introduced to make the set of Maxwell equations readily transferable from one system of units to the other (Römer & Forger, 1993). The values of the constants η, μ_0 and ϵ_0 are connected by the following relation in every system:

$$\eta^2 \mu_0 \epsilon_0 = \frac{1}{c^2}, \tag{II.7}$$

where c is the speed of light. On substituting Eq. (II.6d) into Eq. (II.5) on the preceding page and using Eq. (II.6b) to substitute $\nabla \times \mathbf{E}$ we find that

$$-\mathbf{j} \cdot \mathbf{E} = \frac{\partial}{\partial t} \left(\frac{\epsilon_0}{2} \mathbf{E}^2 + \frac{1}{2\mu_0} \mathbf{B}^2 \right) + \nabla \cdot \left(\frac{1}{\eta \mu_0} \mathbf{E} \times \mathbf{B} \right). \tag{II.8}$$

This has the required form of a continuity equation for the energy of the electromagnetic field if we identify

$$\rho^{\mathrm{E}} \equiv u = \frac{\epsilon_0}{2} \mathbf{E}^2 + \frac{1}{2\mu_0} \mathbf{B}^2 \tag{II.9}$$

with the energy density of the electromagnetic field and

$$\mathbf{j}^{\mathrm{E}} \equiv \mathbf{S} = \frac{1}{\eta \mu_0} \mathbf{E} \times \mathbf{B} \tag{II.10}$$

with the energy flux density. The quantity $\mathbf{S}$ is often called the Poynting vector as it was Poynting who discovered the law for the transfer of energy in the electromagnetic field (Poynting, 1884). The continuity equation is therefore also known as Poynting theorem. A continuity equation does not determine the density and flux density completely; the continuity equation Eq. (II.1) on page 6 is also fulfilled if the density and the flux density are augmented by two arbitrary vector fields $\mathbf{C}$ and $\mathbf{F}$ in the following way

$$\tilde{\rho}^{\mathrm{E}} = \rho^{\mathrm{E}} + \nabla \cdot \mathbf{C}, \tag{II.11a}$$

$$\tilde{\mathbf{j}}^{\mathrm{E}} = \mathbf{j}^{\mathrm{E}} - \frac{\partial}{\partial t} \mathbf{C} + \nabla \times \mathbf{F}. \tag{II.11b}$$

If $\mathbf{C}$ falls off sufficiently quickly, both expressions for the energy density give rise to the same total energy. The additional term $\nabla \times \mathbf{F}$ appears only as a boundary term for integrals of the flux density over open surfaces. However, it is possible to determine the energy density uniquely within a relativistic consideration (Jackson, 1998; Landau & Lifshitz, 1975) (see App. II.A on page 31). Finally, it should be noted that the Poynting vector in Eq. (II.10) is not a generally valid expression for the energy flux density, but rather in the special case of external sources in vacuum, where the

electric displacement $\mathbf{D}$ is related to the electric field $\mathbf{E}$ by $\mathbf{D} = \epsilon_0 \mathbf{E}$ and ϵ_0 is the free space electric permittivity. Also, the magnetic induction $\mathbf{B}$ has to be related to the magnetic field $\mathbf{H}$ by $\mathbf{B} = \mu_0 \mathbf{H}$ and μ_0 is the free space magnetic permittivity. For the propagation of light beams we will treat the electromagnetic wave as a closed system without any sources. In this case the expression for the energy flux density is rigorous (Rohrlich, 1970).

2.2 Linear momentum of the electromagnetic field

The linear momentum is a vector and we will use the notation in components to express the continuity equation for a vector quantity:

$$\frac{\partial \rho_k^{\mathrm{a}}}{\partial t} + \frac{\partial j_{ik}^{\mathrm{a}}}{\partial x_i} = q_k^{\mathrm{a}}, \tag{II.12}$$

where the superscript 'a' denotes a locally preserved quantity. For the linear momentum the rate of change of the momentum density and the divergence of the momentum flux density are balanced by the momentum source density. It is not surprising that the source density is given by the Lorentz force:

$$q_k^p \equiv f_k = -\rho E_k - \eta \epsilon_{klm} j_l B_m. \tag{II.13}$$

On using all Maxwell equations in Eq. (II.6) on the previous page to manipulate the expression for the source density the continuity equation for the linear momentum can be formulated if we identify

$$\rho_k^p \equiv \frac{1}{c^2} \mu_0 S_k = \eta \epsilon_0 \epsilon_{klm} E_l B_m \tag{II.14}$$

with the linear momentum density. This is identical to the energy flux density apart from the factor μ_0 / c^2. The momentum flux density is given by

$$j_{ik}^p \equiv -T_{ik} = \epsilon_0 \left(\frac{1}{2} \mathbf{E}^2 \delta_{ik} - E_i E_k \right) + \frac{1}{\mu_0} \left(\frac{1}{2} \mathbf{B}^2 \delta_{ik} - B_i B_k \right). \tag{II.15}$$

Whenever the momentum flux density is symmetric, the angular momentum continuity equation may be fulfilled with a density, flux density and source density which are given by a vectorial product of the radial position and the respective linear density. This is similar to a rigid body where the angular momentum is solely due to orbital angular momentum which is given by the cross product $\mathbf{L} = \mathbf{r} \times \mathbf{P}$ and where the torque $\mathbf{T} = \mathbf{r} \times \mathbf{F}$ is completely determined by the force $\mathbf{F}$. In a continuum theory this indicates that there is no intrinsic angular momentum and no intrinsic torque, and that for sufficiently small volumes all angular quantities are already given by the corresponding linear quantity. For classical fluids this a well justified assumption, but circular polarised light carries spin angular momentum. This shows the difficulties in assigning the terms spin and orbital angular momentum to specific parts of the total angular momentum. For light beams this difficulty has been resolved by using the angular momentum flux rather than the angular momentum density itself (Barnett, 2002).

2.3 Angular momentum of the electromagnetic field

The density, flux density and source density for the angular momentum are directly related to the respective linear quantities:

$$\rho_k^L = \epsilon_{klm} x_l \rho_m^P, \tag{II.16a}$$

$$j_{ik}^L = \epsilon_{ijl} x_j j_{lk}^P, \tag{II.16b}$$

$$q_k^L = \epsilon_{klm} x_l q_m^P. \tag{II.16c}$$

The given angular quantities fulfil the angular momentum continuity equation owing to the symmetry of the linear momentum flux density j_{lk}^P. The additional term in the divergence of the flux density

$$\frac{\partial}{\partial x_k} j_{ik}^L = \epsilon_{ijl} j_{lj}^P + \epsilon_{ijl} x_j \frac{\partial}{\partial x_k} j_{lk}^P \tag{II.17}$$

vanishes, because the terms in the sum $\epsilon_{ijl} j_{lj}^P$ are anti-symmetric in the last two indices and terms with $j = l$ are equal to zero because of the Levi-Civita symbol. The whole continuity equation for the angular momentum can thus be written as

$$\epsilon_{ijk} x_j \left(\frac{\partial}{\partial t} \rho_k^P + \frac{\partial}{\partial x_l} j_{kl}^P - q_k^P \right) = 0, \tag{II.18}$$

which is always fulfilled because of the validity of the continuity equation for linear momentum.

For the purpose of this thesis we will be concerned with the Maxwell equations in free space, so that all the source densities are equal to zero. To simplify the notation for the calculation of the angular momentum of light beams we introduce the symbols $\mathbf{l}$ for the angular momentum density

$$\mathbf{l} = \epsilon_0 \mathbf{r} \times (\mathbf{E} \times \mathbf{B}) \quad \text{with} \quad l_i \equiv \rho_i^L = \eta \epsilon_0 \left(E_i x_j B_j - B_i x_j E_j \right), \tag{II.19}$$

and $\mathbf{M}$ for the angular momentum flux density which has the components

$$M_{ik} \equiv j_{ik}^L = \epsilon_{ilk} x_l u - \epsilon_{ilm} x_l \left(\epsilon_0 E_m E_k + \frac{1}{\mu_0} B_m B_k \right), \tag{II.20}$$

where u is the energy density from Eq. (II.9) on page 8. For a light beam with an optical axis along the z direction the angular momentum flux through a plane of constant z is given by the integral

$$\mathcal{M}_{zz} \equiv \mathcal{M}_{33} = \iint \mathrm{d}x \, \mathrm{d}y \, M_{33}. \tag{II.21}$$

This quantity has been shown to have a well defined spin and orbital angular momentum part even beyond the paraxial approximation (Barnett, 2002). Earlier work showed that the angular momentum density can be separated into a well defined spin and orbital angular momentum part only within the paraxial approximation (Allen et al., 1992; Barnett & Allen, 1994).

With the help of these expression we are in the position to calculate the orbital

angular momentum of light beams. Light beams are solutions to either the exact wave equation or its paraxial approximation. In the next section we derive the exact wave equation, or Helmholtz equation, and we discuss how this wave equation can be approximated for well collimated beams near the axis.

3 Wave equations

In this section we derive the Helmholtz equation and its paraxial approximation for electromagnetic waves in free space. For a consistent application of the paraxial approximation the wave equation has to be derived for the vector potential rather than the electric field or the magnetic induction.

3.1 Helmholtz equation

We are mostly concerned with the propagation of waves carrying orbital angular momentum in free space and we thus do not consider the presence of charge densities ρ or flux densities $\mathbf{j}$ in the Maxwell equations. In this case the Maxwell equations are a set of homogenous differential equations for the electric field $\mathbf{E}$ and the magnetic induction $\mathbf{B}$:

$$\nabla \cdot \mathbf{E} = 0, \tag{II.22a}$$

$$\nabla \times \mathbf{E} = -\eta \frac{\partial \mathbf{B}}{\partial t}, \tag{II.22b}$$

$$\nabla \cdot \mathbf{B} = 0, \tag{II.22c}$$

$$\nabla \times \mathbf{B} = \eta \epsilon_0 \mu_0 \frac{\partial \mathbf{E}}{\partial t}. \tag{II.22d}$$

On taking the curl of either Eq. (II.22b) or Eq. (II.22d) and substituting the other equation it is possible to derive two wave equations for $\mathbf{E}$ and $\mathbf{B}$. For the electric field the double curl $\nabla \times (\nabla \times \mathbf{E})$ may be replaced by $\nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E}$ and analogously for the magnetic induction $\mathbf{B}$. In both cases the first term is equal to zero because of the Maxwell equations (II.22a) and (II.22c). We are thus left with the wave equations

$$\nabla^2 \mathbf{E} - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \mathbf{E} = 0, \tag{II.23a}$$

$$\nabla^2 \mathbf{B} - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \mathbf{B} = 0. \tag{II.23b}$$

In the following we introduce complex electric and magnetic fields and we limit our attention to monochromatic beams with angular frequency ω. This allows us to separate the electric and magnetic field into a part with the spatial dependence and a time dependent factor $\exp(\mathrm{i}\omega t)$:

$$\mathbf{E} = \mathrm{Re}[\boldsymbol{\mathcal{E}} \exp(-\mathrm{i}\omega t)] \quad \text{and} \quad \mathbf{B} = \mathrm{Re}[\boldsymbol{\mathcal{B}} \exp(-\mathrm{i}\omega t)]. \tag{II.24}$$

By substituting this separated ansatz into the wave equations (II.23) on the previous page we obtain the Helmholtz equations for $\mathcal{E}$ and $\mathcal{B}$:

$$\nabla^2\mathcal{E} + k^2\mathcal{E} = 0, \tag{II.25a}$$

$$\nabla^2\mathcal{B} + k^2\mathcal{B} = 0, \tag{II.25b}$$

where $k = \omega/c$ is the overall wavenumber.

As detailed later, the paraxial approximation to the Helmholtz equation (II.25a) is inconsistent with the Maxwell equation (II.22a) on the previous page (Lax et al., 1975). To avoid these inconsistencies it is possible to use a Helmholtz equation for the vector potential $\mathbf{A}$ which is not required to have a vanishing divergence in the Lorentz gauge. The vector potential is defined over the magnetic induction by

$$\mathbf{B} = \nabla \times \mathbf{A}. \tag{II.26}$$

On substituting the definition of the vector potential into Faraday's law in Eq. (II.22b) we find that the curl of the combined vector field $\mathbf{E} + \eta(\partial\mathbf{A}/\partial t)$ is equal to zero:

$$\nabla \times \left(\mathbf{E} + \eta\frac{\partial}{\partial t}\mathbf{A}\right) = 0. \tag{II.27}$$

This is an indication for the existence of a scalar potential Φ defined as

$$\nabla\Phi = -\left(\mathbf{E} + \eta\frac{\partial}{\partial t}\mathbf{A}\right). \tag{II.28}$$

The vector potential and the scalar potential are not completely determined by the electric field and the electric induction. The derivative of a scalar function ζ may be added to the vector potential if the time-derivative of ζ is subtracted from the scalar potential:

$$\mathbf{A} \longrightarrow \mathbf{A}' = \mathbf{A} + \nabla\zeta \qquad \Phi \longrightarrow \Phi' = \Phi - \eta\frac{\partial}{\partial t}\zeta, \tag{II.29}$$

where ζ is an arbitrary function of $\mathbf{r}$ and t. The relevant gauge for our purposes is the Lorentz gauge, as the Coulomb gauge requires that the divergence of the vector potential is equal to zero, which would be inconsistent with the paraxial approximation (Davis, 1979). The Lorentz gauge is given by the condition

$$\nabla \cdot \mathbf{A} + \frac{1}{\eta c^2}\frac{\partial}{\partial t}\Phi = 0. \tag{II.30}$$

This also requires that the scalar function ζ obeys a wave equation:

$$\nabla^2\zeta - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\zeta = 0. \tag{II.31}$$

With the help of the Eqs. (II.26) and (II.28) it is possible to substitute the electric field and the magnetic induction in the Maxwell equations with the scalar and vector potential. The substitution satisfies the Maxwell equations (II.22b) and (II.22c) on the previous page identically. The dynamic behaviour of $\mathbf{A}$ and Φ is determined by the two remaining Maxwell equations in Eq. (II.22) on the preceding page. Together with

the Lorentz condition in Eq. (II.30) on the previous page the two remaining Maxwell equations yield two wave equations for the scalar and the vector potential:

$$\nabla^2\Phi - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\Phi = 0, \tag{II.32a}$$

$$\nabla^2\mathbf{A} - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\mathbf{A} = 0. \tag{II.32b}$$

The Helmholtz equation for the vector potential is obtained with the ansatz $\mathbf{A} = \mathrm{Re}[\boldsymbol{\mathcal{A}}\exp(-\mathrm{i}\omega t)]$ for a monochromatic wave. With $k = \omega/c$ this yields the following equation

$$\nabla^2\boldsymbol{\mathcal{A}} + k^2\boldsymbol{\mathcal{A}} = 0. \tag{II.33}$$

Once a solution for the vector potential has been found, the electric field can be calculated from the vector potential via

$$\boldsymbol{\mathcal{E}} = \mathrm{i}\eta\omega\left(k^{-2}\nabla(\nabla\cdot\boldsymbol{\mathcal{A}}) + \boldsymbol{\mathcal{A}}\right). \tag{II.34}$$

For this relation we have assumed the monochromatic form for the scalar and vector potential. The given equation follows then from the definition of the scalar potential in Eq. (II.28) on the preceding page and the Lorentz condition in Eq. (II.30) on the previous page.

The vectorial form of the Helmholtz equation can be turned into a scalar wave equation by choosing a constant polarisation vector of unit modulus. We are considering the case, where the vector fields have a polarisation transverse to the optical axis, which is taken along the z direction. We thus write

$$\boldsymbol{\mathcal{V}} = \mathbf{t}\xi(\mathbf{r}), \tag{II.35}$$

where $\boldsymbol{\mathcal{V}}$ stands for the electric field, the magnetic induction or the vector potential and $\mathbf{t}$ is the polarisation vector of unit modulus. With this ansatz the vector form of the Helmholtz equation (II.25) on the preceding page and (II.33) can be turned into a scalar form:

$$\nabla^2\xi + k^2\xi = 0. \tag{II.36}$$

This equation is the starting point for the analysis of the exact form of light beams with orbital angular momentum. In the next section we find an approximation of the Helmholtz equation for beams forming a small angle with the optical axis.

3.2 Paraxial approximation

If the propagation vector $\mathbf{k}$ is inclined by a small angle with respect to the z axis, this vector is paraxial. In the far field this inclination of the propagation vector corresponds to the angle between light beam and optical axis. The dominant component of a paraxial wavevector $\mathbf{k}$ is the z component with

$$k_z = \sqrt{k^2 - \kappa^2} \approx k - \frac{\kappa^2}{2k}, \tag{II.37}$$

where $\kappa = \sqrt{k_x^2 + k_y^2}$ is the transverse wave number. As an ansatz for a solution to the scalar Helmholtz equation (II.36) on the preceding page we consider a function which explicitly depends on the phase factor $\exp(ikz)$:

$$\zeta(\mathbf{r}) = u(\mathbf{r}) \exp(ikz). \tag{II.38}$$

Here, $u(\mathbf{r})$ is an amplitude distribution and not to be confused with the energy density of the electric field in Eq. (II.9) on page 8. In our case the amplitude distribution describes the transverse profile of the light beam. This profile may change with the distance z owing to diffraction or propagation effects. Compared with the primary variation $\exp(ikz)$ these changes will be small for a well collimated beam. This is the reason why the primary factor $\exp(ikz)$ has been written explicitly. On substituting the ansatz of Eq. (II.38) into the scalar Helmholtz equation (II.36) on the previous page we find a partial differential equation for the amplitude distribution u:

$$\nabla_t^2 u + \frac{\partial^2}{\partial z^2} u + 2ik \frac{\partial}{\partial z} u = 0. \tag{II.39}$$

This differential equation can be approximated by neglecting the second derivative of u with respect to z in comparison with the other two terms. If the profile changes only slowly with z the transverse variation of the profile can become dominant so that the following relation holds:

$$\left| \frac{\partial^2 u}{\partial z^2} \right| \ll \left| \nabla_t^2 u \right|. \tag{II.40}$$

If the paraxial approximation is applied to the Helmholtz equation for the electric field the relation above becomes untenable. For a free electric field the Maxwell equation (II.22a) on page 11 requires that the divergence of the electric field vanishes identically. For the transversely polarised electric field of Eq. (II.35) on the previous page this requires in turn that the directional derivative in the transverse direction $\nabla_t E$ is equal to zero everywhere:

$$\nabla \cdot \mathbf{E} = \nabla \cdot \mathrm{t}E = \nabla_t E = 0. \tag{II.41}$$

But this also requires that the transverse Laplacian $\nabla_t^2 E$ is equal to zero, which contradicts the relation in Eq. (II.40). This problem has first been noted by Lax et al. (1975) who, in order to analyse this shortcoming of the paraxial approximation, expanded the electrical field in terms of a parameter $f = \omega_0 / l$, the ratio of the beam-waist and the diffraction length. Using the Helmholtz equation for the vector potential in the Lorentz gauge avoids these inconsistencies and Davis (1979) showed that the results are equivalent in the lowest orders. The second estimate for the paraxial approximation is less problematic. If the profile is slowly varying in z the second derivative is much smaller than the first derivative multiplied by the overall wavenumber:

$$\left| \frac{\partial^2 u}{\partial z^2} \right| \ll k \left| \frac{\partial u}{\partial z} \right|. \tag{II.42}$$

Neglecting the second order derivative with respect to z constitutes the paraxial approximation and results in the paraxial wave equation:

$$\nabla_t^2 u + 2ik\frac{\partial}{\partial z}u = 0. \tag{II.43}$$

Other than in the initial motivation the derivation of the paraxial approximation makes no use of the fact that the ratio of transverse wavenumber to overall wavenumber is small for paraxial beams. It is therefore interesting to derive a systematic expansion in terms of the parameter κ/k.

In order to do so, the whole Helmholtz equation has to be transformed with a Fourier transform to obtain an algebraic equation for the different wavenumbers. When the full Helmholtz equation (II.36) on page 13 is transformed, the result confirms the exact relation

$$-\kappa^2 - k_z^2 + k^2 = 0 \tag{II.44}$$

between the wavenumbers in Eq. (II.37) on page 13. But if we substitute the paraxial ansatz $\xi = u(\mathbf{r})\exp(ikz)$ before transforming the full Helmholtz equation we can use an expansion in terms of κ/k to obtain the paraxial approximation. The Fourier transform of the differential equation

$$\left(\nabla^2 + k^2\right)\xi = \left(\nabla^2 + k^2\right)u(\mathbf{r})\exp(ikz) = 0 \tag{II.45}$$

results in

$$\left[-\kappa^2 - (k-k_z)^2 + 2k(k-k_z)\right]\tilde{\xi}(\mathbf{k}) = 0. \tag{II.46}$$

where $\tilde{\xi}(\mathbf{k}) = \mathcal{F}\{u(\mathbf{r})\exp(ikz)\}$ is the Fourier transform of ξ. This is equivalent to the result $(-\kappa^2 - k_z^2 + k^2)\tilde{\xi}(\mathbf{k}) = 0$ from the Fourier transformation of the Helmholtz equation (II.36) on page 13 without substituting ξ. But the difference $k - k_z$ can be written in terms of the parameter κ/k. Starting from the expression for k_z in Eq. (II.37) on page 13 we can expand k_z in terms of $\varepsilon = (\kappa/k)^2$:

$$k_z = k\left(1 - \frac{\varepsilon}{2} - \frac{\varepsilon^2}{8} + \ldots\right). \tag{II.47}$$

Substituting this expansion into the transformed Helmholtz equation (II.46) shows that the lowest order in $(k-k_z)^2$ is ε^2, whereas $2k(k-k_z)$ is of first order in ε. Neglecting the higher order term is equivalent to neglecting the second derivative of u with respect to z in Eq. (II.39) on the preceding page. This is not surprising as the term $(k-k_z)^2$ in the transformed Helmholtz equation (II.46) comes from the term $(\partial^2 u/\partial z)^2$ in the Helmholtz equation (II.39) on the preceding page. The approximated form of the transformed Helmholtz equation is given by

$$\left[-\kappa^2 + 2k(k-k_z)\right]\tilde{\xi}(\mathbf{k}) = 0. \tag{II.48}$$

On applying the back transform we recover the paraxial wave equation in the form of Eq. (II.43). This is not straight forward to see from Eq. (II.48), but by considering a

Fourier transform of the paraxial wave equation in the form of

$$\left[\nabla_t^2 u + 2ik\frac{\partial}{\partial z}u\right]\exp(ikz) = 0, \tag{II.49}$$

one obtains the approximated form in Eq. (II.48) on the previous page. The paraxial approximation can thus truly be seen as a small angle approximation, where the parameter $\varepsilon = (\kappa/k)^2$ is small. The equivalence will become most apparent in the next section where we consider light beams carrying orbital angular momentum. In particular Bessel beams as solutions to the Helmholtz equation or the paraxial equation are transformed from one into the other by the same small angle approximation that transforms the full Helmholtz equation into the approximated, paraxial form.

4 Paraxial and non paraxial optics

In this section we present light beams as solution to the paraxial and non-paraxial wave equation, which is the full Helmholtz equation derived earlier in this chapter. The orbital angular momentum (OAM) of light has first been studied in the paraxial approximation, as it seemed to be a well-defined quantity only within this approximation. But with the discovery of the angular momentum flux as the defining quantity for angular momenta in electromagnetic fields it became possible to extend the analysis into the non-paraxial regime. For this thesis the non-paraxial regime will become important for the propagation of light with fractional mean OAM. This is why we present first the well established Laguerre-Gaussian beams in the paraxial approximation, and later we seek an extension to the non-paraxial regime.

4.1 Paraxial light beams

Solutions to the paraxial wave equation are well known from laser optics (Siegman, 1986). In particular Hermite-Gaussian beams are common for physical systems with a rectangular symmetry. Hermite-Gaussian beams do not carry OAM but they form a complete set of solutions. It is therefore possible to transform between these beams and Laguerre-Gaussian beams, another set of complete solutions. Laguerre-Gaussian beams are solutions to the paraxial wave equation in polar coordinates. In many respects the most simple solution are Bessel beams for which transverse parts and longitudinal parts are completely factorised. These beams also form a complete set of solutions, but it is the fact that Bessel beams are solutions to both the paraxial and the full Helmholtz equations which makes them useful for an analysis of the propagation of beams with OAM in the non-paraxial regime.

4.1.1 Hermite-Gaussian beams

In Cartesian coordinates it is possible to separate the amplitude function $u(x,y,z)$ into two functions, either of which depends on one transverse variable, x or y, and the longitudinal variable z in the following way:

$$u_{nl}^{\mathrm{HG}}(x,y,z) = u_n^{\mathrm{HG}}(x,z)u_l^{\mathrm{HG}}(y,z), \tag{II.50}$$

where n and l give the order of the Hermite polynomial in u_n or u_l. Each of these functions obeys a paraxial wave equation in one transverse dimension. In the following we consider only one function and we choose x as the single transverse variable. The paraxial equation for $u_n(x,z)$ has the form:

$$\left(\frac{\partial^2}{\partial x^2} + 2ik\frac{\partial}{\partial z}\right) u_n^{\mathrm{HG}}(x,z) = 0. \tag{II.51}$$

A normalised solution to this equation is given by the product of a Gaussian with a Hermite polynomial in the following form:

$$\begin{aligned} u_n^{\mathrm{HG}}(x,z) = \frac{C_n^{\mathrm{HG}}}{\sqrt{w(z)}} \exp\left(ik\frac{x^2 z}{2(z_R^2+z^2)}\right) \exp\left(-\frac{x^2}{w^2(z)}\right) \\ \times \exp(-i(n+\frac{1}{2})\chi(z)) H_n\left(\frac{\sqrt{2}x}{w(z)}\right). \end{aligned} \tag{II.52}$$

Here, C_n^{HG} is a normalisation constant and H_n is the Hermite polynomial of order n. Owing to the orthogonality of the Hermite polynomials (Stephenson & Radmore, 1993) the Hermite-Gaussian beams are orthonormal when integrated over the transverse variable x:

$$\int_{-\infty}^{\infty} u_n^{\mathrm{HG}}(x,z)\left(u_m^{\mathrm{HG}}(x,z)\right)^* \mathrm{d}x = \delta_{nm}. \tag{II.53}$$

The width of the beam changes on propagation and is determined by the Gaussian spot size $w(z)$, which gives the radial distance at which the Gaussian term falls off to $1/e$ of its value on axis:

$$w(z)^2 = \frac{2(z_R^2+z^2)}{kz_R} = w_0^2\left[1+\left(\frac{z}{z_R}\right)^2\right]. \tag{II.54}$$

The Rayleigh range z_R is a measure for the focal region of the beam and is given by

$$z_R = \frac{\pi w_0^2}{\lambda}, \tag{II.55}$$

where λ is the wavelength and w_0 is the beam waist at the focal point, which is also the origin of the z axis. The term $(n+1/2)\chi(z)$ is the Gouy phase which describes the phase jump of π that occurs over the focal region of any spherical converging wave (Gouy, 1890; Siegman, 1986). This behaviour is apparent from the form of the function $\chi(z)$:

$$\tan\chi = \frac{z}{z_R}. \tag{II.56}$$

In figure (II.1) on the following page the intensity profiles for a number of different Hermite-Gaussian beams are shown. The profiles are plotted for the combined Hermite-Gaussian mode $u_{nl}^{\mathrm{HG}}(x,y,z)$. The indices n and l indicate the number of nodal lines in the respective x and y direction.

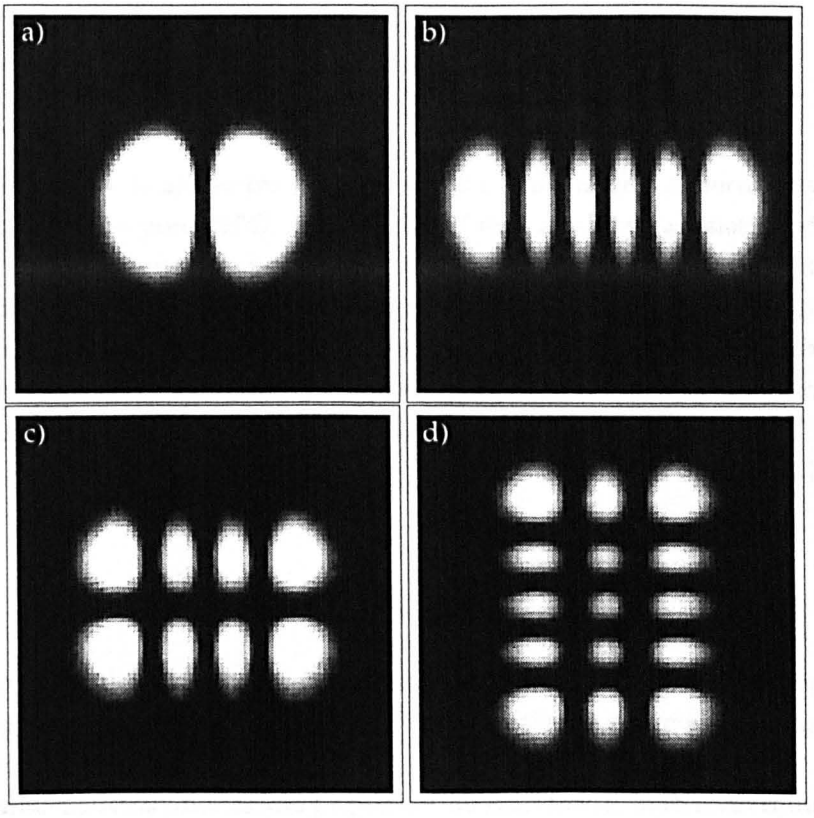

Figure II.1 — Intensity profiles of Hermite-Gaussian beams. a) $n = 1, l = 0$. b) $n = 5, l = 0$. c) $n = 3, l = 1$. d) $n = 2, l = 4$.

4.1.2 Laguerre-Gaussian beams

Another complete set of beams are the Laguerre-Gaussian modes, but instead of the rectangular symmetry of the Hermite-Gaussian modes these beams shows a rotational symmetry. Consequently, the Laguerre-Gaussian beams are solutions to the paraxial wave equation in cylindrical coordinates:

$$\left(\frac{1}{\rho}\frac{\partial}{\partial\rho} + \frac{\partial^2}{\partial\rho^2} + \frac{1}{\rho^2}\frac{\partial^2}{\partial\varphi^2} + 2ik\frac{\partial}{\partial z}\right) u_{mp}^{\mathrm{LG}} = 0. \tag{II.57}$$

Here, ρ is the radial coordinate and φ is the azimuthal or polar angle. The normalised form of the Laguerre-Gaussian beams is given by

$$\begin{aligned} u_{mp}^{\mathrm{LG}}(\rho,\varphi,z) = {} & \frac{C_{mp}^{\mathrm{LG}}}{\sqrt{w(z)}}\left(\frac{\rho\sqrt{2}}{w(z)}\right)^{|m|} \exp\left(-\frac{\rho^2}{w^2(z)}\right) L_p^{|m|}\left(\frac{2\rho^2}{w^2(z)}\right) \\ & \times \exp\left(ik\frac{\rho^2 z}{2(z_R^2+z^2)}\right) \exp(im\varphi)\exp(-i(2p+|m|+1)\chi(z)) \end{aligned} \tag{II.58}$$

Here, $L_p^{|m|}$ are the generalised Laguerre polynomials and C_{mp}^{LG} is the appropriate normalisation constant. The Laguerre-Gaussian beams form an orthonormal set in the

mode index p when integrated over the radial coordinate:

$$\int_0^\infty \rho u_{np}^{\mathrm{LG}}(\rho,\varphi,z)\left(u_{nq}^{\mathrm{LG}}(\rho,\varphi,z)\right)^* \mathrm{d}\rho = \delta_{pq}. \tag{II.59}$$

The orthogonality is due to the properties of the generalised Laguerre polynomials (Abramowitz & Stegun, 1974). The profiles of the Laguerre-Gaussian modes show concentric rings, whose number is determined by the mode index p (see figure (II.2)). The other mode index is contained in the azimuthal phase term $\exp(im\varphi)$, which gives rise to $|m|$ intertwined helical wave-fronts, that is surfaces of equal phase. The handedness of these helixes is determined by the sign of m. This azimuthal phase structure is the cause for the OAM of Laguerre-Gaussian modes. The Laguerre-Gaussian modes are also orthonormal in the azimuthal mode index m when integrated over the azimuthal variable φ:

$$\int_0^{2\pi} u_{np}^{\mathrm{LG}}(\rho,\varphi,z)\left(u_{mq}^{\mathrm{LG}}(\rho,\varphi,z)\right)^* \mathrm{d}\rho = \delta_{nm}. \tag{II.60}$$

In all generality the choice of the 2π radians interval for the angle φ is arbitrary and it is more general to integrate from θ_0 to $\theta_0 + 2\pi$. This will be discussed in more detail in the next chapter.

Hermite-Gaussian modes and Laguerre-Gaussian modes both form complete sets and one family of beams can therefore be represented as a superposition of the other. Beijersbergen et al. (1993) published a very clear analysis of the mode conversion, but instead of using the indices m and p to describe the Laguerre-Gaussian modes, they considered indices n and l analogously to the Hermite-Gaussian beams. The two sets of indices are related by

$$m = n - l \quad \text{and} \quad p = \min(n,l). \tag{II.61}$$

Any Laguerre-Gaussian mode can now be expressed in terms of Hermite-Gaussian modes as follows:

$$u_{nl}^{\mathrm{LG}}(x,y,z) = \sum_{t=0}^{n+l} \mathrm{i}^t b(n,l,t) u_{n+l-t,t}^{\mathrm{HG}}(x,y,z). \tag{II.62}$$

The coefficients are given by:

$$b(n,l,t) = \sqrt{\frac{(n+l)!t!}{2^{n+l}n!l!}} \times \frac{1}{k!}\frac{\mathrm{d}^t}{\mathrm{d}s^t}\left((1-s)^n(1-s)^m\right)|_{s=0}. \tag{II.63}$$

This transformation is based on earlier results by Abramochkin & Volostnikov (1991) (see App. II.B on page 33). Hermite-Gaussian and Laguerre-Gaussian beams can also be generated from a Gaussian seed function by acting on the fundamental Gaussian function with differential operators (Enderlein & Pampaloni, 2004). This provides another way of transforming one family of modes into the other.

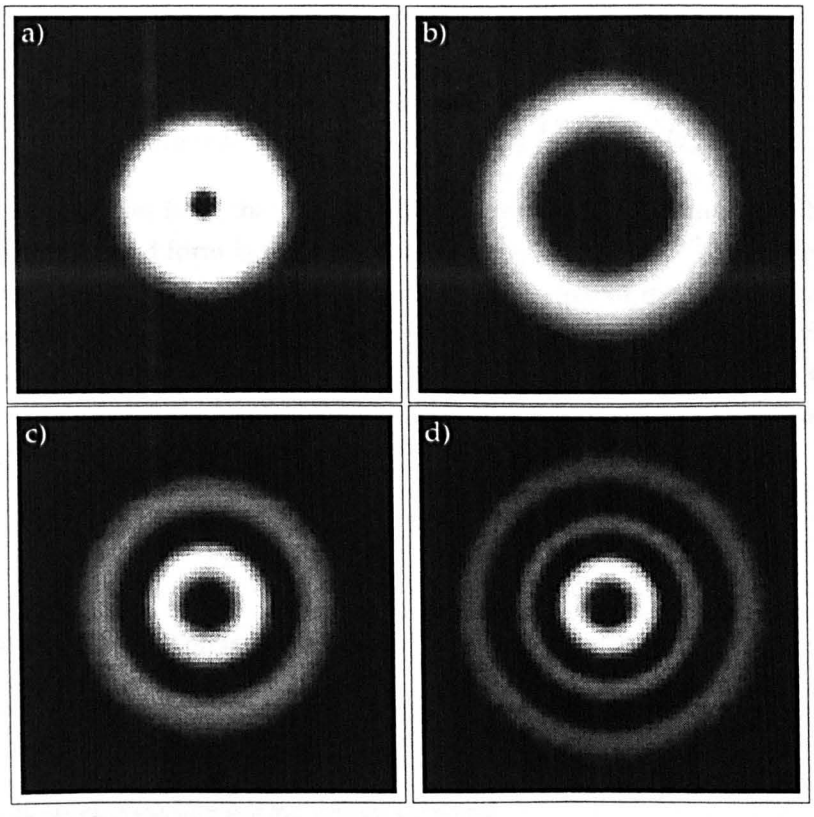

Figure II.2 — Intensity profiles of Laguerre-Gaussian beams. a) $m = 1, p = 0$. b) $m = 5, p = 0$. c) $m = 2, p = 1$. d) $m = 2, p = 2$.

4.1.3 Complex argument form

Both, the Hermite-Gaussian modes and the Laguerre-Gaussian modes can also be written with complex arguments. This is due to a freedom in the scaling parameter for the Hermite or Laguerre polynomials in the respective rectangular or cylindrical paraxial equations. In the standard form, where the polynomials have real arguments this scaling parameter is chosen differently from the scaling parameter of the Gaussian term. This results in the known form of the Hermite-Gaussian and Laguerre-Gaussian modes in Eqs. (II.52) on page 17 and (II.58) on page 18, where the Gaussian term is split into a real and complex part. The real part contains the Gaussian spot size $w(z)$ and the complex part contains the radius of curvature $R(z) = (z^2 + z_R^2)/z$.

This 'inelegant' difference in the scaling parameter is mended in the complex argument form, which is therefore also called the 'elegant' form. In this form the scaling parameter for the polynomial is chosen to be identical to the scaling parameter of the Gaussian term and is basically determined by a function $\tilde{q}$ which is given in terms of the radius of curvature $R(z)$ and and Gaussian spot size $w(z)$:

$$\frac{1}{\tilde{q}(z)} \equiv \frac{1}{R(z)} + \mathrm{i}\frac{\lambda}{\pi w^2(z)} = \frac{\mathrm{i}}{z_R + \mathrm{i}z}. \tag{II.64}$$

If we use this scaling parameter the expression for the Hermite-Gaussian mode changes

to

$$\breve{u}_n^{\mathrm{HG}}(x,z) = \breve{C}_n^{\mathrm{HG}} \left(\frac{z_R}{z_R + \mathrm{i}z}\right)^{\frac{n+1}{2}} H_n\left(\sqrt{\frac{kx^2}{2(z_R + \mathrm{i}z)}}\right) \exp\left(-\frac{kx^2}{2(z_R + \mathrm{i}z)}\right), \quad \text{(II.65)}$$

which is of a simpler form than the standard Hermite-Gaussian modes. The Gouy phase of the standard form is related to the term $z_R/(z_R = \mathrm{i}z)^{(n+1)/2}$ in the elegant form. The form of this term is also altered by the change of the scaling parameter. The elegant variant represents an equally valid and complete set of solutions to the paraxial wave equations. But, whereas the Hermite-Gaussian modes in the standard form are orthogonal to each other, the modes in the elegant form are biorthogonal to a set of adjoint functions $\breve{v}_n^{HG}$ given by:

$$\breve{v}_n^{HG}(x,z) = H_n\left(\sqrt{\frac{kx^2}{2(z_R + \mathrm{i}z)}}\right). \quad \text{(II.66)}$$

The orthogonality relation is now given by

$$\int_{-\infty}^{\infty} \breve{u}_n^{\mathrm{HG}}(x,z) \left(\breve{v}_m^{HG}(x,z)\right)^* \mathrm{d}x = c_n \delta_{nm}, \quad \text{(II.67)}$$

where c_n is an appropriate normalisation constant.

The Laguerre-Gaussian beams can also be given in an elegant form. By replacing the standard argument $\sqrt{2}\rho/w(z)$ with the complex argument $\sqrt{-\mathrm{i}k\rho^2/(2\tilde{q}(z))}$ the Laguerre polynomial has the same argument as the Gaussian term. However, this also affects the Gouy phase and the term $(\sqrt{2}\rho/w(z))^{|m|}$. The final result is given by

$$\begin{aligned}\breve{u}_{mp}^{\mathrm{LG}}(\rho,\varphi,z) = \breve{C}_{mp}^{\mathrm{LG}} \left(\frac{k}{2}\right)^{|m|+p+1} \rho^m \left(\frac{1}{z_R + \mathrm{i}z}\right)^{|m|+p+1} \exp(\mathrm{i}m\varphi) \\ \times L_p^{|m|}\left(\frac{k\rho^2}{2(z_R + \mathrm{i}z)}\right) \exp\left(-\frac{k\rho^2}{2(z_R + \mathrm{i}z)}\right),\end{aligned} \quad \text{(II.68)}$$

which is identical to the paraxial limit of a non-paraxial extension to Laguerre-Gaussian beams discussed by Barnett & Allen (1994).

4.1.4 Bessel beams

The full Helmholtz equation supports a class of solutions, whose transverse profile is invariant on propagation. Consequently, these beams have been called 'diffraction-free beams' by Durnin et al. (1987). They are exact solutions to the full Helmholtz equation, but they have an infinitely extended transverse profile. Under laboratory conditions the realisation of such beams will be limited by the physical requirement of a finite aperture (Durnin, 1987). The radial, azimuthal and axial parts are all separated in the Bessel beams, which have an amplitude distribution of the form:

$$\zeta_m^{\mathrm{B}}(\rho,\varphi,z) = C_m^e J_m(\kappa\rho) \exp(\mathrm{i}m\varphi) \exp(\mathrm{i}k_z z), \quad \text{(II.69)}$$

where C_m^e is a normalisation constant and no summation over m is implied. Bessel beams have the same azimuthal phase factor $\exp(\mathrm{i}m\varphi)$ as the Laguerre-Gaussian

modes, and hence carry also OAM. In the transition from the full Helmholtz equation to the paraxial wave equation the transverse part of the differential equation is left unchanged. But the relation between the different wavenumbers k, κ and k_z changes. The ansatz for the paraxial wave equation Eq. (II.43) on page 15 is given by

$$u_m^B(\rho, \varphi, z) = C_m^p J_m(\kappa\rho) \exp(im\varphi) \exp(i\tilde{k}_z z), \tag{II.70}$$

where C_m^p is the normalisation constant in the paraxial case and $\tilde{k}_z$ is a reduced longitudinal wave number with $\tilde{k}_z = k_z - k$. This is due to the separated factor $\exp(ikz)$ in the paraxial ansatz (see Eq. (II.39) on page 14). On comparing the paraxial wave equation (II.43) on page 15 with the defining differential equation for Bessel functions (see (Stephenson & Radmore, 1993))

$$(\kappa\rho)^2 J_m''(\kappa\rho) + \kappa\rho J_m'(\kappa\rho) + \left((\kappa\rho)^2 - m^2\right) J_m(\kappa\rho) = 0, \tag{II.71}$$

one can see that u_m^B will be a solution of the paraxial wave equation if

$$\kappa^2 = -2k\tilde{k}_z \quad \longrightarrow \quad k_z = k - 2\frac{\kappa^2}{2k}. \tag{II.72}$$

This form of the longitudinal wave number looks identical to approximated form for small κ/k. This is not surprising as u_m^B is the solution to the paraxial wave equation. In the next section we will see that the Bessel beams are also solutions to the full Helmholtz equation but with a different expression for k_z.

The transverse profiles of the Bessel beams are shown in figure (II.3) on the following page. Bessel beams do not change on propagation and there is no diffraction or dilation of the profile. The difference between paraxial and non-paraxial Bessel beams is only apparent in the longitudinal factor. The transverse profiles are the same for both cases.

4.2 Non-paraxial light beams

Bessel beams are also solutions to the full Helmholtz equation which can be seen if we substitute the ansatz ζ_m^B from Eq. (II.69) on the previous page into the full Helmholtz equation in cylindrical coordinates (see Eq. (II.25) on page 12):

$$\left(\frac{1}{\rho}\frac{\partial}{\partial\rho} + \frac{\partial^2}{\partial\rho^2} + \frac{1}{\rho^2}\frac{\partial}{\partial\varphi^2} + \frac{\partial^2}{\partial z^2} + k^2\right)\zeta_m^B = 0, \tag{II.73}$$

To satisfy Bessel's equation (II.71) the expression for k_z has to be different from the paraxial case and is given by exact expression in Eq. (II.37) on page 13:

$$\kappa^2 = k^2 - k_z^2 \quad \longrightarrow \quad k_z = \sqrt{k^2 - \kappa^2}. \tag{II.74}$$

Bessel modes can thus be transformed from solving the Helmholtz equation to solve the paraxial equation, by the same change in the expression for k_z that characterises the paraxial approximation. Together with the completeness of the Bessel functions this allows us to develop a non-paraxial extension of the Laguerre-Gaussian modes.

Using the completeness of the Bessel beams it is possible to expand the Laguerre-

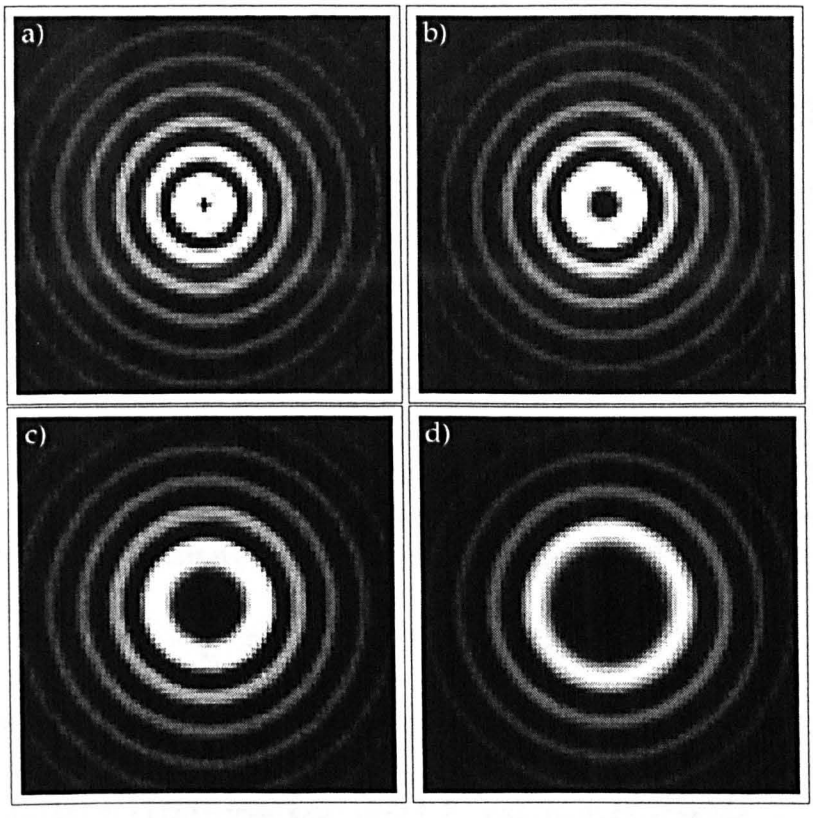

Figure II.3 — Intensity profiles of paraxial Bessel beams. a) $m = 1$. b) $m = 2$. c) $m = 4$. d) $m = 6$.

Gaussian modes in the following way

$$u_{mp}^{\mathrm{LG}}(\rho, \varphi, z) = \int_0^\infty d_{mp}(\kappa) J_m(\kappa\rho) \exp(\mathrm{i}m\varphi) \exp(-\mathrm{i}\kappa^2 z/(2k)) \mathrm{d}\kappa. \tag{II.75}$$

The exact form of the function $d_{mp}(\kappa)$ can be determined from the expression for the paraxial Laguerre-Gaussian beam. To transform the superposition into a solution of the full Helmholtz equation we have to change the longitudinal phase factor. The radial and the azimuthal part are left unchanged. In this way we can construct a non-paraxial form of the Laguerre-Gaussian beam:

$$\zeta_{mp}^{\mathrm{LG}}(\rho, \varphi, z) = \int_0^k d_{mp}(\kappa) J_m(\kappa\rho) \exp(\mathrm{i}m\varphi) \exp(\mathrm{i}\sqrt{k^2 - \kappa^2} z) \mathrm{d}\kappa. \tag{II.76}$$

The two expansions differ not only in the longitudinal part, but also in the boundaries of the integration. While the paraxial expansion includes all transverse wavevectors $0 \leq \kappa < \infty$, the non-paraxial expansion is limited to the range $0 \leq \kappa \leq k$. For values of $\kappa > k$ the contributions to the expansion become evanescent. The general scheme for obtaining a non-paraxial extension of the Laguerre-Gaussian modes is shown in

the following diagram:

$$\begin{array}{ccc} \nabla^2\xi + k^2\xi = 0 & \longrightarrow & \xi^{\mathrm{LG}}_{mp} = \displaystyle\int_0^k d_{mp}(\kappa) J_m(\kappa\rho)\exp(\mathrm{i}m\phi)\exp(\mathrm{i}\sqrt{k^2-\kappa^2}z)\mathrm{d}\kappa, \\ \Big\downarrow \frac{\kappa^2}{k^2}\ll 1 \ \xi = u\exp(\mathrm{i}kz) & & \Big\uparrow \frac{\kappa^2}{k^2}\ll 1 \ \xi = u\exp(\mathrm{i}kz) \\ \nabla_t^2 u + 2\mathrm{i}k\partial_z u = 0 & \longrightarrow & u^{\mathrm{LG}}_{mp} = \displaystyle\int_0^\infty d_{mp}(\kappa) J_m(\kappa\rho)\exp(\mathrm{i}m\phi)\exp\left(-\mathrm{i}\frac{\kappa^2}{2k}z\right)\mathrm{d}\kappa. \end{array}$$

To determine the function $d_{mp}(\kappa)$ we can multiply both sides of Eq. (II.75) on the previous page by $\rho J_m(\kappa'\rho)$ and integrate both sides over ρ. On applying the Fourier-Bessel theorem in the form (Gray & Mathews, 1895) (see App. VI.D on page 160):

$$\int_0^\infty \rho J_m(\kappa\rho) J_m(\kappa'\rho)\mathrm{d}\rho = \frac{1}{\kappa'}\delta(\kappa-\kappa'). \tag{II.77}$$

we can write the function $d_{mp}(\kappa)$ in integral form:

$$d_{mp}(\kappa) = \kappa\exp(-\mathrm{i}m\varphi)\exp\left(\mathrm{i}\frac{\kappa^2}{2k}z\right)\int_0^\infty \rho u^{\mathrm{LG}}_{mp}(\rho,\varphi,z) J_m(\kappa\rho)\mathrm{d}\rho. \tag{II.78}$$

It is interesting to relate this integral to a different approach of obtaining a non-paraxial extension to the Laguerre-Gaussian modes by Barnett & Allen (1994) who looked into a general expression for an electric field, that is a solution to the Helmholtz equation and satisfies the transversality condition. In the transverse plane this field is also given as an expansion in Bessel functions similar to Eq. (II.76) on the preceding page. The beam has been constructed as a non-paraxial extension of the Laguerre-Gaussian beam, but in the paraxial limit the elegant form of the Laguerre-Gaussian modes is recovered. This step involves the following integral (Gradshteyn & Ryzhik, 2000; Barnett & Allen, 1994):

$$\begin{aligned} &\int_0^\infty \mathrm{d}\kappa \exp\left(-\frac{\kappa^2}{2k}(z_R+\mathrm{i}z)\right)\left(\frac{\kappa}{k}\right)^{2p+m+1} J_m(\kappa\rho) \\ &\quad = \frac{p!}{2}k\left(\frac{k\rho}{2}\right)^m\left(\frac{2}{k(z_R+\mathrm{i}z)}\right)\exp\left(-\frac{\kappa\rho^2}{2(z_R+\mathrm{i}z)}\right) L_p^{|m|}\left(\frac{\kappa\rho^2}{2(z_R+\mathrm{i}z)}\right). \end{aligned} \tag{II.79}$$

Apart from the azimuthal phase factor $\exp(\mathrm{i}m\varphi)$ and the normalisation constant this is essentially the Laguerre-Gaussian mode in its elegant form (cf. Eq. (II.68)). This fact allows us to substitute the Laguerre-Gaussian mode $u^{\mathrm{LG}}_{nl}(x,y,z)$ in the integral (II.78) with the elegant form $\breve{u}^{\mathrm{LG}}_{mp}$ given by

$$\breve{u}^{\mathrm{LG}}_{mp} = U_{mp}\exp(\mathrm{i}m\varphi)\int_0^\infty \mathrm{d}\kappa \exp\left(-\frac{\kappa^2}{2k}(z_R+\mathrm{i}z)\right)\left(\frac{\kappa}{k}\right)^{2p+m+1} J_m(\kappa\rho), \tag{II.80}$$

so that it is possible to apply the Fourier-Bessel theorem in Eq. (II.78). The final result

has the rather simple form

$$d_{mp}(\kappa) = \exp\left(-\frac{\kappa^2}{2k} z_R\right) \left(\frac{\kappa}{k}\right)^{2p+|m|+1}. \tag{II.81}$$

On substituting this result into the expansion in Eq. (II.76) we obtain a non-paraxial extension of the Laguerre-Gaussian modes. Non-paraxial solutions will become important later in this thesis, when we calculate the propagation of modes with fractional orbital angular momentum.

5 Spin and orbital angular momentum

In this section we discuss the concepts of spin angular momentum (SAM) and orbital angular momentum (OAM) in light beams. Previously it has been thought impossible to separate the total angular momentum into a spin and an angular part outside the paraxial approximation. But with the introduction of the angular momentum flux as the defining property for the optical angular momentum it became possible to extend earlier results from the paraxial approximation into the non-paraxial regime. We will thus present an analysis of the SAM and OAM in terms of the angular momentum flux. To allow for circular polarisation associated with SAM of light we consider a vectorial electromagnetic field.

5.1 Angular momentum flux

In section 2.3 we have defined the angular momentum flux density and the total angular momentum flux. Here, we are discussing why this quantity provides a meaningful separation into SAM and OAM parts. It is interesting to note that it is not possible to distinguish the spin part from the orbital part by their behaviour under a change of axis. As long as the total linear momentum flux through a surface is perpendicular to this surface, the total angular momentum flux is the same for every axis parallel to the beam (Barnett, 2002). A similar result has also been derived for the angular momentum by Berry (1998).

Although the expression in Eq. (II.20) on page 10 is a valid expression for the angular momentum flux density, it is helpful to substitute the real electric field and the real magnetic induction with the complex field amplitudes of a monochromatic beam with frequency ω (see Eq. (II.24) on page 11). For the complex field amplitudes the Maxwell equations (II.22) on page 11 are given by

$$\mathcal{E}_j = \mathrm{i}\frac{c^2}{\eta\omega}\varepsilon_{jkl}\frac{\partial}{\partial x_k}\mathcal{B}_l, \tag{II.82a}$$

$$\mathcal{B}_j = -\mathrm{i}\frac{1}{\eta\omega}\varepsilon_{jkl}\frac{\partial}{\partial x_k}\mathcal{B}_l. \tag{II.82b}$$

With help of the complex field amplitudes it is possible to calculate the cycle-averaged angular momentum flux density $\mathring{M}_{zz}$. The real field components in Eq. (II.20) on page 10 are expressed as sums of the complex field amplitudes and their conjugates. On integration over the cycle time $T = 2\pi/\omega$ only terms without the time-varying factor $\exp(-\mathrm{i}\omega t)$ are left (see App. II.C). This allows us to write the cycle average of

the angular momentum flux as (Barnett, 2002):

$$\mathring{M}_{zz} = \frac{1}{2}\,\mathrm{Re}\left[y\left(\epsilon_0\mathcal{E}_x\mathcal{E}_y^* + \frac{1}{\mu_0}\mathcal{B}_x^*\mathcal{B}_z\right) - x\left(\epsilon_0\mathcal{E}_y\mathcal{E}_z^* + \frac{1}{\mu_0}\mathcal{B}_y^*\mathcal{B}_z\right)\right]. \tag{II.83}$$

The z components of the electric field and the magnetic induction can be eliminated by use of the Maxwell equations (II.82) on the previous page. The total, cycle-averaged angular momentum flux is obtained by integrating this quantity over the whole transverse plane. By partial integration and subsequent manipulations this integral can be brought into the following form:

$$\begin{aligned}\mathring{\mathcal{M}}_{zz} = \frac{\epsilon_0 c^2}{2\omega}\,\mathrm{Re}\Bigg\{-\mathrm{i}\iint \rho\,\mathrm{d}\rho\,\mathrm{d}\varphi\,\Big[&(\mathcal{E}_x\mathcal{B}_x^* + \mathcal{E}_y\mathcal{B}_y^*)\\ &+\frac{1}{2}\left(-\mathcal{B}_x^*\frac{\partial}{\partial\varphi}\mathcal{E}_y + \mathcal{E}_y\frac{\partial}{\partial\varphi}\mathcal{B}_x^* - \mathcal{E}_x\frac{\partial}{\partial\varphi}\mathcal{B}_y^* + \mathcal{B}_y^*\frac{\partial}{\partial\varphi}\mathcal{E}_x\right)\Big]\Bigg\},\end{aligned} \tag{II.84}$$

which is given in cylindrical polar coordinates. In this form the separation of the total angular momentum flux into a spin and orbital part becomes most apparent. The spin part is given by

$$\mathring{\mathcal{M}}_{zz}^{\mathrm{spin}} = \frac{\epsilon_0 c^2}{2\omega}\,\mathrm{Re}\left\{-\mathrm{i}\iint \rho\,\mathrm{d}\rho\,\mathrm{d}\varphi(\mathcal{E}_x\mathcal{B}_x^* + \mathcal{E}_y\mathcal{B}_y^*)\right\} \tag{II.85}$$

and depends on the polarisation of the light beam, whereas the orbital part is connected to changes in the azimuthal variable:

$$\mathring{\mathcal{M}}_{zz}^{\mathrm{orbital}} = \frac{\epsilon_0 c^2}{4\omega}\,\mathrm{Re}\left\{-\mathrm{i}\iint \rho\,\mathrm{d}\rho\,\mathrm{d}\varphi\left(-\mathcal{B}_x^*\frac{\partial}{\partial\varphi}\mathcal{E}_y + \mathcal{E}_y\frac{\partial}{\partial\varphi}\mathcal{B}_x^* - \mathcal{E}_x\frac{\partial}{\partial\varphi}\mathcal{B}_y^* + \mathcal{B}_y^*\frac{\partial}{\partial\varphi}\mathcal{E}_x\right)\right\}. \tag{II.86}$$

The physical significance of this separation follows on considering the effects of birefringence on the beam. The imparted phase shift will be different for the x and y components but it will not incur an azimuthal phase shift. The orbital part is therefore left unchanged. On the other hand, if an optical element imprints an azimuthal phase shift all the components will change in the same way, which leaves the spin part unchanged.

5.2 Angular momenta of light beams

We are now able to calculate the spin and orbital parts of the total OAM flux of a light beam beyond the paraxial approximation. To this end we consider a class of general non-paraxial beams introduced by Barnett & Allen (1994). The components

of the electric field for this beam are given by

$$\mathcal{E}_x = \alpha \int_0^k \mathrm{d}\kappa\, d_{mp}(\kappa) \exp(\mathrm{i}m\varphi) \exp(\mathrm{i}\sqrt{k^2-\kappa^2}) J_m(\kappa\rho), \tag{II.87a}$$

$$\mathcal{E}_y = \beta \int_0^k \mathrm{d}\kappa\, d_{mp}(\kappa) \exp(\mathrm{i}m\varphi) \exp(\mathrm{i}\sqrt{k^2-\kappa^2}) J_m(\kappa\rho), \tag{II.87b}$$

$$\mathcal{E}_z = \int_0^k \mathrm{d}\kappa\, d_{mp}(\kappa) \exp(\mathrm{i}m\varphi) \exp(\mathrm{i}\sqrt{k^2-\kappa^2}) \frac{\kappa}{2\sqrt{k^2-\kappa^2}} \tag{II.87c}$$
$$\times \left[(\mathrm{i}\alpha - \beta)\exp(-\mathrm{i}\varphi) J_{m-1}(\kappa\rho) - (\mathrm{i}\alpha + \beta)\exp(-\mathrm{i}\varphi) J_{m+1}(\kappa\rho)\right].$$

This beam is a solution to the vector form of the Helmholtz equation (II.25a) on page 12 and satisfies also the transversality condition in $\nabla \cdot \mathbf{E} = 0$. The complex numbers α and β are chosen such that their squared moduli sum to unity: $|\alpha|^2 + |\beta|^2 = 1$. The components of the magnetic induction are given by Eq. (II.82) on page 25 (see (Barnett, 2002)):

$$\mathcal{B}_x = \frac{1}{\omega\eta} \int_0^k \mathrm{d}\kappa\, \frac{d_{mp}(\kappa) \exp(\mathrm{i}\sqrt{k^2-\kappa^2}z)}{\sqrt{k^2-\kappa^2}} \left[-\beta \frac{2k^2-\kappa^2}{2} \mathrm{e}^{\mathrm{i}m\varphi} J_m(\kappa\rho)\right. \tag{II.88a}$$
$$\left. + \frac{\kappa^2}{4}\left(\mathrm{e}^{\mathrm{i}(m-2)\varphi}(\mathrm{i}\alpha - \beta) J_{m-2}(\kappa\rho) - \mathrm{e}^{\mathrm{i}(m+2)\varphi}(\mathrm{i}\alpha + \beta) J_{m+2}(\kappa\rho)\right)\right],$$

$$\mathcal{B}_y = \frac{1}{\omega\eta} \int_0^k \mathrm{d}\kappa\, \frac{d_{mp}(\kappa) \exp(\mathrm{i}\sqrt{k^2-\kappa^2}z)}{\sqrt{k^2-\kappa^2}} \left[\alpha \frac{2k^2-\kappa^2}{2} \mathrm{e}^{\mathrm{i}m\varphi} J_m(\kappa\rho)\right. \tag{II.88b}$$
$$\left. + \frac{\kappa^2}{4}\left(\mathrm{e}^{\mathrm{i}(m-2)\varphi}(-\alpha - \mathrm{i}\beta) J_{m-2}(\kappa\rho) - \mathrm{e}^{\mathrm{i}(m+2)\varphi}(\alpha - \mathrm{i}\beta) J_{m+2}(\kappa\rho)\right)\right],$$

$$\mathcal{B}_z = \frac{1}{\omega\eta} \int_0^k \mathrm{d}\kappa\, \frac{\kappa d_{mp}(\kappa) \exp(\mathrm{i}\sqrt{k^2-\kappa^2}z)}{2} \tag{II.88c}$$
$$\times \left[(-\alpha + \mathrm{i}\beta)\mathrm{e}^{\mathrm{i}(m+1)\varphi} J_{m+1}(\kappa\rho) - (\alpha + \mathrm{i}\beta)\mathrm{e}^{\mathrm{i}(m-1)\varphi} J_{m-1}(\kappa\rho)\right].$$

For this particular beam the total linear momentum flux has no component in the transverse direction. The total angular momentum flux is therefore the same for every axis parallel to the z direction.

Originally, Barnett & Allen (1994) considered a specific form of the function $d_{mp}(\kappa)$ which is different from the result in Eq. (II.81) on page 25. But there is some degree of freedom in choosing this function, so that it is equally possible to use the function $d_{mp}(\kappa)$ given in Eq. (II.81). This particular function falls off sufficiently quickly as κ tends to zero, which ensures a finite energy, momentum and angular momentum per unit length. If we calculate now the spin and orbital part of the total angular momentum according to Eqs. (II.85) and (II.86) on the previous page we find

$$\mathring{\mathcal{M}}_{zz}^{\mathrm{spin}} = \mathrm{i}(\alpha\beta^* - \alpha^*\beta) \frac{\pi\epsilon_0 c^2}{2\eta\omega^2} \int_0^k \mathrm{d}\kappa\, |d_{mp}(\kappa)|^2 \frac{2k^2-\kappa^2}{\kappa\sqrt{k^2-\kappa^2}}, \tag{II.89}$$

$$\mathring{\mathcal{M}}_{zz}^{\mathrm{orbital}} = m \frac{\pi\epsilon_0 c^2}{2\eta\omega^2} \int_0^k \mathrm{d}\kappa\, |d_{mp}(\kappa)|^2 \frac{2k^2-\kappa^2}{\kappa\sqrt{k^2-\kappa^2}}. \tag{II.90}$$

Only terms for which the azimuthal factor $\exp(\mathrm{i}m\varphi)$ is multiplied by its conjugate contribute to the total flux, this also ensures that only products of Bessel functions with equal order have to be considered. These terms can be calculated using the

Fourier-Bessel theorem (see Eq. (II.77) on page 24). In the original paper on angular momentum of light Allen et al. (1992) examined the ratio between angular momentum density and energy flux both per unit length. This established the fact that Laguerre-Gaussian beams, or in fact any beam with the particular azimuthal phase factor $\exp(\mathrm{i}m\varphi)$, has well defined OAM. On considering the ratio between either the spin or orbital part of the total angular momentum flux through the transverse plane and the total energy flux though the same plane we can recover the original results beyond the paraxial approximation. The cycle-averaged energy flux through the transverse plane is given by the integral (see Eq. (II.10) on page 8):

$$\begin{aligned} \overset{\circ}{j}{}_z^{\mathrm{E}} &= \frac{1}{2\eta\mu_0\omega}\,\mathrm{Re}\iint \rho\mathrm{d}\rho\mathrm{d}\phi\left(\mathcal{E}_x\mathcal{B}_y^* - \mathcal{E}_y\mathcal{B}_x^*\right), \\ &= \frac{\pi}{2\eta\mu_0\omega}\int_0^k \mathrm{d}\kappa|d_{mp}(\kappa)|^2\frac{2k^2-\kappa^2}{\kappa\sqrt{k^2-\kappa^2}}. \end{aligned} \tag{II.91}$$

This expression includes the same integral as the spin and orbital parts of the total angular momentum flux. The ratio of the respective quantities is thus given by:

$$\frac{\overset{\circ}{\mathcal{M}}{}_{zz}^{\mathrm{spin}}}{\overset{\circ}{j}{}_z^{\mathrm{E}}} = \frac{\sigma_z}{\omega}, \qquad \frac{\overset{\circ}{\mathcal{M}}{}_{zz}^{\mathrm{orbital}}}{\overset{\circ}{j}{}_z^{\mathrm{E}}} = \frac{m}{\omega}, \qquad \frac{\overset{\circ}{\mathcal{M}}{}_{zz}}{\overset{\circ}{j}{}_z^{\mathrm{E}}} = \frac{\sigma_z + m}{\omega}. \tag{II.92}$$

Here, we have set $\sigma_z = \mathrm{i}(\alpha\beta^* - \alpha^*\beta)$, which determines the polarisation. For circular polarised light $\beta = \alpha\exp(\pm\mathrm{i}\pi/2)$, so that σ_z takes the values ± 1.

This analysis shows that the azimuthal phase factor $\exp(\mathrm{i}m\varphi)$ is responsible for the well-defined OAM. Laguerre-Gaussian beams and Bessel beams share this phase factor and in the next section we are looking in more detail into the azimuthal phase structure of these beams.

5.3 Azimuthal phase structure

The previous analysis showed that beams with a phase factor $\exp(\mathrm{i}m\varphi)$ have a well defined orbital angular momentum. As the ratio of the orbital angular momentum flux to the energy flux is given by m/ω and the energy of a single photon is given by $\hbar\omega$, this seems to suggest that each photon in a beam with the phase factor $\exp(\mathrm{i}m\varphi)$ carries an OAM of $\hbar m$. But the analysis of the light beams with OAM has been entirely classical so far, so that this tempting conclusion remains to be supported by a quantum theory which will be introduced in the next chapter.

However, it is possible to link the OAM and the associated phase factor to an optical vortex (Nye & Berry, 1974). Optical vortices are phase singularities of the optical field and hence have zero amplitude. Light beams with a phase factor of $\exp(\mathrm{i}m\varphi)$ have an optical vortex on axis with a topological charge of m. This can be seen more clearly from the definition of the vortex charge, which is calculated by integrating along a closed path around the singularity (Nye, 1999):

$$Q = \frac{1}{2\pi}\oint \mathrm{d}\chi = \frac{1}{2\pi}\oint \mathbf{ds}\cdot\nabla\chi, \tag{II.93}$$

where $\mathbf{ds}$ is the line element. For a azimuthal phase of $\chi = m\varphi$ this integral results in

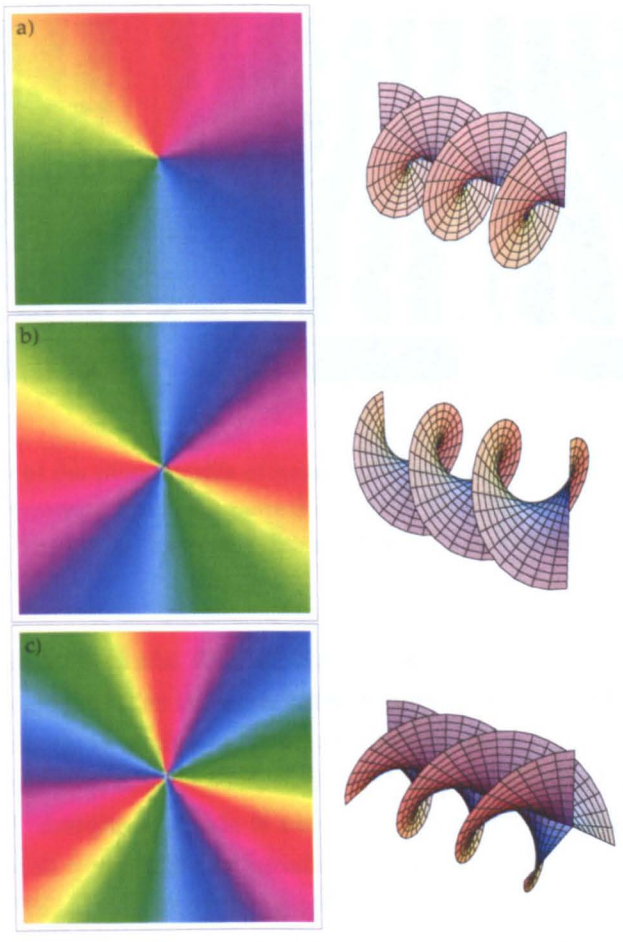

Figure II.4 — The left panel shows plots of the phase structure associated with a phase factor $\exp(im\varphi)$ for different values of m. When propagated light beams with the depicted phase structures have surfaces of equal phase in form of intertwined helices. a) $m = 1$. b) $m = -2$. c) $m = 3$.

a vortex charge of m. The connection between the vortex charge, the azimuthal phase structure and the OAM content of a beam will be revisited later in this thesis, when we examine light beams with fractional OAM. To complete this section we show the phase structures for a number of different values of m in figure (II.4). One can see that for a phase factor of $\exp(im\varphi)$ there are $|m|$ phase ramps in a complete circle. Whether the phase ramps clockwise or anti-clockwise depends on the sign of m. This can be seen from the sequence of the colours in figure (II.4). On propagation we can identify $|m|$ intertwined helices as surfaces of equal phase, which are shown on the right panel in figure (II.4). The handedness of these helices is determined by the sign of m.

6 Generating light beams with orbital angular momentum

In most experimental setups the transverse field distribution of a laser beam is best described by a superposition of Hermite-Gaussian modes. If the laser cavity has rectangular symmetry this follows naturally from symmetry of the Hermite-Gaussian modes, but even for a cavity with circular symmetry the output mode is hardly ever a

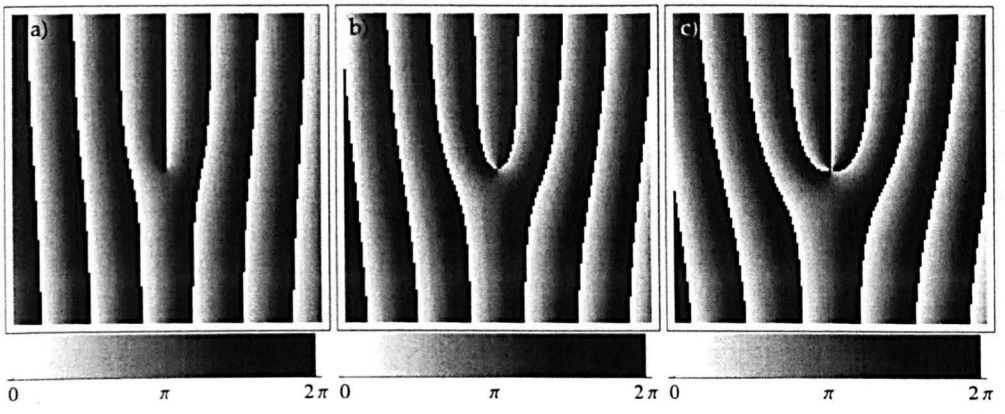

Figure II.5 — Plot of the phase profile of holographic patterns for generating light carrying orbital angular momentum. The patterns are designed to give the emerging light an azimuthal phase factor of $\exp(im\varphi)$ in the first order diffracted beam. a) $m = 1$. b) $m = 2$. c) $m = 3$.

pure Laguerre-Gaussian mode. As discussed earlier in this chapter (see section 4.1.2) Laguerre-Gaussian beams can be represented as superpositions of Hermite-Gaussian beams. In a circular cavity astigmatism breaks the frequency degeneracy between the constituent Hermite-Gaussian modes, so that instead of the pure Laguerre-Gaussian mode a so-called 'doughnut'-mode is observed. One way to generate a Laguerre-Gaussian mode is to remove the astigmatism, either by specially designing the laser cavity (Tamm & Weis, 1990) or by using cylindrical lenses (Beijersbergen et al., 1993).

In this thesis we are more concerned with two other methods of generating Laguerre-Gaussian modes. Rather than transforming Hermite-Gaussian modes into Laguerre-Gaussian modes these two methods change the orbital angular momentum content within the family of Laguerre-Gaussian modes. In this way a Laguerre-Gaussian mode with nonzero orbital angular momentum can be generated from a pure Gaussian mode. One way of manipulating light beams in this way is to use spiral phase plates (Beijersbergen et al., 1994). These are optical elements with an optical thickness that increases with the azimuthal angle, until it drops sharply after completing the circle. As a result an incident Gaussian beam emerges with a helical phase front. Spiral phase plates require a very small mechanical tolerance to ensure that the step height corresponds exactly to an integer multiple of the optical wavelength. Spiral phase plates can also be deliberately designed such that the step height does not correspond to an integer multiple. This gives rise to the phenomenon of fractional OAM which will be covered in more detail in chapter VI.

Holograms give results which are very similar to spiral phase plates. They resemble diffraction gratings, but modified such that the first order diffracted beam has a specific phase and amplitude structure. The required diffraction grating can be produced experimentally by interfering the desired light beam with a reference beam, say a pure Gaussian beam. The interference pattern can be put on a holographic film which, once photographically developed, generates the desired beam when placed in the reference beam. It is also possible to calculate the required diffraction grating. This has the advantage that the resulting pattern can be created on a spatial light modulator, a commercially available optical element which manipulates the intensity

or phase of the light beam. With the help of spatial light modulators the pattern of the diffraction grating can be modified in real-time. This makes computer-generated holograms a very versatile experimental tool, which is used widely to manipulate light beams carrying OAM (Allen et al., 2003). In figure (II.5) on the previous page we show the holographic patterns designed to create light beams with an azimuthal phase factor $\exp(im\varphi)$ for $m = 1, 2, 3$. The patterns show a characteristic fork dislocation, where the number of prongs corresponds to $|m|$.

7 Summary

In this chapter we have given an overview of the classical aspects of light carrying orbital angular momentum. Starting from the Maxwell equations we have derived analytic expressions for the angular momentum density of the electromagnetic field and the corresponding flux density and source density. These expressions were needed to identify the angular momentum flux density as the defining quantity for the orbital angular momentum (OAM) of light. This is because the flux density shows a well defined separation into a spin and an orbital part beyond the paraxial approximation. We have calculated the spin and orbital angular momentum content for a class of non-paraxial light beams. Contained in this class is a non-paraxial extension of the Laguerre-Gaussian beams, which has been derived in this chapter. This particular light beam and also other beams of light carrying OAM share the same phase factor $\exp(im\varphi)$. Such a phase factor is directly related to an optical vortex on the axis and is also responsible for the OAM in the beam.

In the next chapter we will review the quantum mechanical description of the azimuthal angle φ, and its conjugate variable, the OAM. The classical expressions, and in particular the non-paraxial expansion into Bessel beams will be needed later in this thesis, when we examine light beams with fractional OAM.

II.A Relativistic considerations

In section 2 we have derived the expressions for the energy, the linear momentum and the angular momentum of the electromagnetic field. By using the general form of a continuity equation and the Maxwell equations (II.6) on page 8 it is not possible to determine the quantities uniquely. In the relativistic formulation this ambiguity is resolved. We are using Greek indices to denote the time, and the three spatial components of vectors and tensors. We distinguish covariant components from contravariant components by writing the former with a lowered index and the latter with a raised index. A Greek index can thus take on the values 0, 1, 2 and 3 and if the same index is used for two components, one covariant and one contravariant, a summation over this index is implied. Roman indices will be used to denote the spatial components only.

The starting point of the derivation is the rank two, anti-symmetric field strength tensor (Jackson, 1998)

$$F^{\alpha\beta} = \frac{\partial}{\partial x^\alpha} A^\beta - \frac{\partial}{\partial x^\beta} A^\alpha, \tag{II.94}$$

which is defined over derivatives of the relativistic four potential. This potential has

components $A^0 = \Phi$, and the A^i are identical to the components of the vector potential. The field strength tensor contains components of the electrical field and the magnetic induction (Römer & Forger, 1993)

$$F^{0i} = -\frac{E_i}{\eta c}, \qquad \text{and} \qquad F^{ij} = -\epsilon_{ijk} B_k. \tag{II.95}$$

For the three-dimensional quantities E_i and B_i we do not have to distinguish between covariant and contravariant components and we can continue to write them with a lower, Roman index. The quantities of energy and linear momentum of the field are combined in the symmetric energy-momentum tensor:

$$T^{\alpha\beta} = \frac{1}{\mu_0}\left(g^{\alpha\gamma}F_{\gamma\delta}F^{\delta\beta} + \frac{1}{4}g^{\alpha\beta}F_{\gamma\delta}F^{\gamma\delta}\right) = \frac{1}{\mu_0}\left(F^{\alpha}{}_{\gamma}F^{\delta\beta} + \frac{1}{4}g^{\alpha\beta}F_{\gamma\delta}F^{\gamma\delta}\right). \tag{II.96}$$

Here, $g^{\alpha\beta} = g_{\alpha\beta}$ is the diagonal metric tensor of special relativity, with elements $g_{00} = 1$ and $g_{11} = g_{22} = g_{33} = -1$. This tensor contains the energy and momentum density in the components

$$T^{00} \equiv \rho^{\mathrm{E}} = \frac{1}{2}\left(\epsilon_0 E_i E_i + \frac{1}{\mu_0} B_i B_i\right), \tag{II.97}$$

$$T^{0i} = T^{10} \equiv \frac{j_i^{\mathrm{E}}}{c} = c\rho_i^{\mathrm{P}} = \frac{1}{\eta\mu_0 c}\epsilon_{ijk} E_j B_k. \tag{II.98}$$

The remaining components of the energy-momentum are identical to the linear momentum flux density, which is denoted by the same symbol:

$$T^{ij} = T^{ji} \equiv T_{ij} = \epsilon_0\left(\frac{1}{2}E_k E_k \delta_{ij} - E_i E_j\right) + \frac{1}{\mu_0}\left(\frac{1}{2}B_k B_k \delta_{ij} - B_i B_j\right). \tag{II.99}$$

The continuity equations for energy and linear momentum in the absence of any sources can be combined to

$$\frac{\partial}{\partial x^{\alpha}} T^{\alpha\beta} = 0. \tag{II.100}$$

In the relativistic derivation the energy and momentum densities are uniquely determined.

The angular momentum content of the field is described by the rank three tensor

$$M^{\alpha\beta\gamma} = x^{\beta} T^{\gamma\alpha} - x^{\gamma} T^{\beta\alpha}, \tag{II.101}$$

which is anti-symmetric in the last two indices. This tensor is also a locally conserved quantity and obeys the continuity equation

$$\frac{\partial}{\partial x^{\alpha}} M^{\alpha\beta\gamma} = 0. \tag{II.102}$$

The angular momentum density can be derived from the rank three tensor in the following way:

$$\frac{1}{2}\epsilon_{ijk} M^{0jk} \equiv \rho_i^{\mathrm{L}} = \eta\epsilon_0\left(E_i x_j B_j - B_i x_j E_j\right). \tag{II.103}$$

In the continuity equation (II.102) the angular momentum density is differentiated

with respect to time; the spatial derivatives thus act on components of $M^{\alpha\beta\gamma}$ which form the angular momentum flux density:

$$\begin{aligned}\frac{1}{2}\varepsilon_{ijk}M^{ljk} &= M^{l}{}_{i} \equiv j^{\mathrm{L}}_{li}, \\ &= \varepsilon_{ijk}x_j\left[\epsilon_0\left(\frac{1}{2}E_nE_n\delta_{kl} - E_kE_l\right) + \frac{1}{\mu_0}\left(\frac{1}{2}B_nB_n\delta_{kl} - B_kB_l\right)\right].\end{aligned} \tag{II.104}$$

This is equivalent to the angular momentum flux density in Eq. (II.20) on page 10.

II.B Beam transformations

According to Abramochkin & Volostnikov (1991) a linear combination of products of Hermite polynomials for the two transverse coordinates may be turned into a Laguerre polynomial. The connection between the two different polynomials is given by

$$\begin{aligned}&\sum_{k=0}^{n+m}(2\mathrm{i})^k P_k^{(n-k,m-k)}(0)H_{n+m-k}(x)H_k(y) = \\ &\qquad = 2^{n+m}\begin{cases}(-1)^m m!(x+\mathrm{i}y)^{(n-m)}L_m^{n-m}(x^2+y^2) & n \geq m, \\ (-1)^n n!(x+\mathrm{i}y)^{(m-n)}L_n^{m-n}(x^2+y^2) & m \geq n.\end{cases}\end{aligned} \tag{II.105}$$

where $P_k^{(n-k,m-k)}(0)$ is a Jacobi polynomial defined by the equation

$$P_k^{(n-k,m-k)}(0) = \frac{(-1)^k}{2^k k!}\frac{\mathrm{d}^k}{\mathrm{d}t^k}\left[(1-t)^n(1+t)^m\right]|_{t=0}. \tag{II.106}$$

To prove the given relation we consider the following two equivalent expressions:

$$\int_{\mathbb{R}^2}\exp\left[-\xi^2-\eta^2+2\mathrm{i}(x\xi+y\eta)\right]\mathrm{d}\xi\,\mathrm{d}\eta = \pi\exp\left(-x^2-y^2\right), \tag{II.107a}$$

$$\int_{\mathbb{R}^2}\exp\left[-\xi^2-\eta^2+\mathrm{i}z(\xi-\mathrm{i}\eta)+\mathrm{i}z^*(\xi+\mathrm{i}\eta)\right]\mathrm{d}\xi\,\mathrm{d}\eta = \pi\exp\left(-zz^*\right), \tag{II.107b}$$

where $z = x+\mathrm{i}y$ and $z^* = x-\mathrm{i}y$. In order to make use of Rodrigues' formulae for Hermite and Laguerre polynomials (Abramowitz & Stegun, 1974) both lines are differentiated by the operators $\partial_x^m\partial_y^n$ and $\partial_z^m\partial_{z^*}^n$ respectively. For the first line in Eq. (II.107a) this operation yields

$$\begin{aligned}(2\mathrm{i})^{m+n}&\iint_{\mathbb{R}^2}\exp\left[-\xi^2-\eta^2+2\mathrm{i}(x\xi+y\eta)\right]\xi^m\eta^n\mathrm{d}\xi\,\mathrm{d}\eta = \\ &= \pi\frac{\partial^m\partial^n}{\partial x^m\partial y^n}\exp\left(-x^2-y^2\right), \\ &= \pi(-1)^{m+n}\exp\left(-x^2-y^2\right)H_m(x)H_n(y).\end{aligned} \tag{II.108}$$

The Laguerre polynomial in (II.105) contains both z and its complex conjugate z^* in its argument. For two variables Rodrigues' formula for Laguerre polynomials changes

into:

$$L_n^m(zz^*) = \frac{\mathrm{d}^n}{\mathrm{d}(zz^*)^n}\exp(-zz^*)(zz^*)^m(zz^*)^n, \tag{II.109a}$$

$$= \frac{1}{n!\exp(-zz^*)(zz^*)^m}\frac{z^{m+n}}{z^n}\frac{\mathrm{d}^n}{\mathrm{d}(z^*)^n}\exp(-zz^*)(z^*)^m(z^*)^n, \tag{II.109b}$$

$$= \frac{1}{n!\exp(-zz^*)(zz^*)^m}\frac{(z^*)^{m+n}}{(z^*)^n}\frac{\mathrm{d}^n}{\mathrm{d}z^n}\exp(-zz^*)z^m z^n. \tag{II.109c}$$

Therefore, the differentiation in (II.107b) on the previous page may be expressed in terms of Laguerre polynomials as follows:

$$\frac{\partial^m\partial^n}{\partial z^m\partial(z^*)^n}\exp(-zz^*) = \frac{\partial^m}{\partial z^m}(-z)^n\exp(-zz^*),$$

$$= (-1)^n m!\exp(-zz^*)z^{n-m}L_m^{n-m}(zz^*), \quad \text{for } n \geq m, \tag{II.110a}$$

$$= \frac{\partial^n}{\partial(z^*)^m}(-z^*)^n\exp(-zz^*),$$

$$= (-1)^m n!\exp(-zz^*)(z^*)^{m-n}L_m^{m-n}(zz^*), \quad \text{for } m \geq n. \tag{II.110b}$$

By operating with $\partial_z^m\partial_{z^*}^n$ on both sides of Eq. (II.107b) on the preceding page, the left hand side will be transformed, such that Eq. (II.108) on the previous page can be used. The right hand side can be written in terms of Laguerre functions (see Eq. (II.110b)), so that eventually a relation between Hermite functions and Laguerre functions is established. The differentiation of the left hand side of (II.107b) on the previous page yields

$$\partial_z^m\partial_{z^*}^n\int_{\mathbb{R}^2}\mathrm{e}^{\left[-\xi^2-\eta^2+\mathrm{i}z(\xi-\mathrm{i}\eta)+\mathrm{i}z^*(\xi+\mathrm{i}\eta)\right]}\mathrm{d}\xi\,\mathrm{d}\eta = \mathrm{i}^{m+n}\int_{\mathbb{R}^2}\mathrm{e}^{\left[-\xi^2-\eta^2+\mathrm{i}z(\xi-\mathrm{i}\eta)+\mathrm{i}z^*(\xi+\mathrm{i}\eta)\right]}(\xi-\mathrm{i}\eta)^m(\xi+\mathrm{i}\eta)^n\,\mathrm{d}\xi\,\mathrm{d}\eta\,. \tag{II.111}$$

The factor $(\xi-\mathrm{i}\eta)^m(\xi+\mathrm{i}\eta)^n$ may be expressed using a Taylor series in the following way:

$$(\xi-\mathrm{i}\eta)^m(\xi+\mathrm{i}\eta)^n = \left(1-\mathrm{i}\frac{\eta}{\xi}\right)^m\left(1+\mathrm{i}\frac{\eta}{\xi}\right)^n\xi^{m+n},$$

$$= \sum_{k=0}^{m+n}\frac{1}{k!}\frac{\partial^k}{\partial t^k}\left[(1-t)^n(1+t)^m\right]|_{t=0}t^k \quad \text{for } t = -\mathrm{i}\frac{\eta}{\xi}. \tag{II.112}$$

On substitutung the Taylor expansion into Eq. (II.111), this equation may be rewritten such that Eq. (II.108) on the preceding page can be used to write the integral in terms

of Hermite functions:

$$\begin{aligned}
&\sum_{k=0}^{m+n} \mathrm{i}^{m+n} \frac{1}{k!} \frac{\partial^k}{\partial t^k} \left[(1-t)^n (1+t)^m\right]|_{t=0} (-\mathrm{i})^k \\
&\quad \times \int_{\mathbb{R}^2} \exp\left[-\xi^2 - \eta^2 + \mathrm{i}z(\xi - \mathrm{i}\eta) + \mathrm{i}z^*(\xi + \mathrm{i}\eta)\right] \xi^{m+n-k} \eta^k \mathrm{d}\xi \, \mathrm{d}\eta \,, \\
&\quad = \sum_{k=0}^{m+n} (2\mathrm{i})^k \frac{(-1)^{m+n}}{2^{m+n}} \frac{(-1)^k}{2^k k!} \frac{\partial^k}{\partial t^k} \left[(1-t)^n (1+t)^m\right]|_{t=0} \pi \\
&\quad \times \exp\left(-x^2 - y^2\right) H_{m+n-k}(x) H_n(y).
\end{aligned} \tag{II.113}$$

Using the equivalence of $x^2 + y^2$ and zz^*, the expression for the Jacobi polynomials in (II.106) on page 33 and comparing the terms with the differentiated right hand side of Eq. (II.111), yields finally the relation between Hermite and Laguerre functions in Eq. (II.105) on page 33.

II.C Cycle average

For a monochromatic wave with angular frequency ω we can write a vector quantity, such as the electric field or the magnetic induction, in the form:

$$\mathbf{V}(\mathbf{r}, t) = \mathrm{Re}[\boldsymbol{\mathcal{V}}(\mathbf{r} \exp(-\mathrm{i}\omega t)] = \frac{1}{2} \left[\boldsymbol{\mathcal{V}} \exp(-\mathrm{i}\omega t) - \boldsymbol{\mathcal{V}}^* \exp(\mathrm{i}\omega t)\right]. \tag{II.114}$$

In calculating derived quantities, such as the energy density or the momentum density of the electromagnetic field it may be advantageous to concentrate only on the spatial dependence and not on the periodic time variation. We therefore take the average value of a quantity during a whole cycle of period $T = 2\pi/\omega$. For example the energy density of the electromagnetic field is given by (see Eq. (II.9) on page 8) (Born & Wolf, 1999):

$$\rho^{\mathrm{E}} = u = \frac{\epsilon_0}{2} \mathbf{E}^2 + \frac{1}{2\mu_0} \mathbf{B}^2 = u_e + u_m, \tag{II.115}$$

which separates into an electric and a magnetic part. The time average of the electric energy density is calculated according to:

$$\begin{aligned}
\mathring{u}_e = \frac{\epsilon_0}{2T} \int_0^T \mathbf{E}^2 \mathrm{d}t &= \frac{\epsilon_0}{8T} \int_0^T \left[\mathcal{E}^2 \exp(-2\mathrm{i}\omega t) + 2\mathcal{E}\mathcal{E}^* + (\mathcal{E}^*)^2 \exp(2\mathrm{i}\omega t)\right] \mathrm{d}t, \\
&= \frac{\epsilon_0}{4} \mathcal{E}\mathcal{E}^*
\end{aligned} \tag{II.116}$$

and analogously the cycle averaged magnetic energy density is $\mathring{u}^b = \mathcal{B}\mathcal{B}^*/(4\mu_0)$. As an example of the cycle average of a quantity involving two different fields we calculate the cycle average of the Poynting vector (see Eq. (II.10) on page 8):

$$\begin{aligned}
\mathring{\mathbf{S}} &= \frac{1}{\eta\mu_0} \frac{1}{T} \int_0^T \mathbf{E} \times \mathbf{B} \, \mathrm{d}t, \\
&= \frac{1}{\eta\mu_0} \frac{1}{4T} \int_0^T \left[\mathcal{E} \times \mathcal{B} \mathrm{e}^{-2\mathrm{i}\omega t} + 2\,\mathrm{Re}[\mathcal{E} \times \mathcal{B}^*] + \mathcal{E}^* \times \mathcal{B}^* \mathrm{e}^{2\mathrm{i}\omega t}\right] \mathrm{d}t, \\
&= \frac{1}{2\eta\mu_0} \mathrm{Re}[\mathcal{E} \times \mathcal{B}^*].
\end{aligned} \tag{II.117}$$

All cycle averages of the quantities in section II 5.1 are calculated similarly.

Bibliography

Abramochkin, E. & Volostnikov, V. (1991), 'Beam transformation and nontransformed beams', *Optics Communications* **83**(1,2), 123–135.

Abramowitz, M. & Stegun, I. S. (1974), *Handbook of Mathematical Functions*, Dover Publications, Inc., Mineola, New York. Reprint. Originally published: National Bureau of Standards, corrected edition, 1964. ISBN: 0-486-61272-4.

Allen, L., Barnett, S. M. & Padgett, M. J. (2003), *Optical Angular Momentum*, Institute of Physics Publishing, Ltd., Bristol. ISBN: 0-7503-0901-6.

Allen, L., Beijersbergen, M. W., Spreeuw, R. J. C. & Woerdman, J. P. (1992), 'Orbital angular momentum of light and the transformation of Laguerre-Gaussian modes', *Physical Review A* **45**(11), 8185–8190. Reprinted in (Allen et al., 2003, Paper 2.1). DOI: 0.1103/PhysRevA.45.8185.

Barnett, S. M. (2002), 'Optical angular-momentum flux', *Journal of Optics B* **4**(2), S7–S16. DOI: 10.1088/1464-4266/4/2/361.

Barnett, S. M. & Allen, L. (1994), 'Orbital angular momentum and nonparaxial light beams', *Optics Communication* **110**(5-6), 670–678. DOI: doi:10.1016/0030-4018(94)90269-0.

Beijersbergen, M. W., Allen, L., van der Veen, H. E. L. O. & Woerdman, J. P. (1993), 'Astigmatic laser mode converters and transfer of orbital angular momentum', *Optics Communication* **96**(1-3), 123–132. DOI: doi:10.1016/0030-4018(93)90535-D.

Beijersbergen, M. W., Coerwinkel, R. P. C., Kristensen, M. & Woerdman, J. P. (1994), 'Helical-wavefrontnext term laser beams produced with a spiral phaseplate', *Optics Communication* **112**(5-6), 321–327. DOI: 10.1016/0030-4018(94)90638-6.

Berry, M. V. (1998), Paraxial beams of spinning light, *in* M. S. Soskin & M. V. Vasnetsov, eds, 'Singular optics', Vol. 3487, International Conference on Singular Optics, SPIE, pp. 6–11.

Beth, R. A. (1936), 'Mechanical detection and measurement of the angular momentum of light', *Physical Review* **50**, 115–125.

Born, M. & Wolf, E. (1999), *Principles of Optics*, 7th expanded edn, Cambrige University Press, Cambridge. reprinted 2003. ISBN: 0-521-64222-1.

Davis, L. W. (1979), 'Theory of electromagnetic beams', *Physical Review A* **19**(3), 1177–1179. DOI: 10.1103/PhysRevA.19.1177.

Durnin, J. (1987), 'Exact solutions for nondiffracting beams. I. The scalar theory', *Journal of the Optical Society of America A* **4**(4), 651.

Durnin, J., Micelli, J. J. & Eberly, J. H. (1987), 'Diffraction-Free Beams', *Physical Review Letters* **58**(15), 1499–1501. DOI: 10.1103/PhysRevLett.58.1499.

Enderlein, J. & Pampaloni, F. (2004), 'Unified operator approach for deriving Hermite-Gaussian and Laguerre-Gaussian laser modes', *Journal of the Optical Society of America A* **21**(8), 1553–1558.

Gouy, L. G. (1890), 'Sur une propriete nouvelle des ondes lumineuses', *Compted Rendus de l'Academie des Sciences* **110**, 1251.

Gradshteyn, I. S. & Ryzhik, I. M. (2000), *Tables of Integrals, Series, and Products*, 6th edn, Academic Press, Inc., San Diego. ISBN: 0-12-294757-6.

Gray, A. G. & Mathews, G. B. (1895), *A Treatise on Bessel functions and their Application to Physics*, Macmillan and Co., London. The authors do not use the notation of the Dirac delta distribution, which had been not yet been introduced at the time the book was published.

Jackson, J. D. (1998), *Classical Electrodynamics*, 3rd edn, John Wiley & Sons, Inc., New York. ISBN: 0-471-30932-X.

Landau, L. D. & Lifshitz, E. M. (1975), *The Classical Theory of Fields*, Vol. 2 of *Course of Theoretical Physics*, 4th revised english edition edn, Butterworth-Heinemann, Burlington.

Lax, M., Louisell, W. H. & McKnight, B. (1975), 'From Maxwell to paraxial wave optics', *Physical Review A* **11**(4), 1365–1370. DOI: 10.1103/PhysRevA.11.1365.

Maxwell, J. C. (1873), *A Treatise on Electricity and Magnetism*, Clarendon Press Series, Oxford at the Clarendon Press, Oxford. Reprinted in Maxwell (1998).

Nye, J. F. (1999), *Natural Focussing and the Fine Structure of Light*, Institute of Physics Publishing, Bristol.

Nye, J. F. & Berry, M. V. (1974), 'Dislocations in Wave Trains', *Proceedings of the Royal Society of London, Series A* **336**(1605), 165–190.

Poynting, J. H. (1884), 'On the transfer of energy in the electromagnetic field', *Philosophical Transactions* **175**, 343–361. Reprinted in Poynting (1920*a*).

Poynting, J. H. (1909), 'The wave motion of a revolving shaft, and a suggestion as to the angular momentum in a beam of circularly polarised light', *Proceedings of the Royal Society of London, Series A* **82**, 560–567. Reprinted in Poynting (1920*b*) and (Allen et al., 2003, Paper 1.1).

Rohrlich, F. (1970), 'Electromagnetic Momentum, Energy, and Mass', *American Journal of Physics* **38**(11), 1310–1316. DOI: 10.1119/1.1976082.

Römer, H. & Forger, M. (1993), *Klassische Feldtheorie*, VCH Verlagsgesellschaft, Weinheim.

Siegman, A. E. (1986), *Lasers*, University Science Books, Sausalito. ISBN: 0-935702-11-3.

Stephenson, G. & Radmore, P. M. (1993), *Advanced Mathematical Methods for Engineerging and Science Students*, Cambrige University Press, Cambridge. ISBN: 0-521-36860-X.

Tamm, C. & Weis, C. O. (1990), 'Bistability and optical switching of spatial patterns in a laser', *Journal of the Optical Society of America B* 7(6), 1034–1038.

III Chapter

Quantum formulation of angle and orbital angular momentum

The quantum mechanical description of the azimuthal rotation angle and the orbital angular momentum around the polar axis differs greatly from the description of position and momentum. This has led to a controversy over the existence of a Hermitian angle operator in the literature. It is possible to circumvent some problems associated with Hermiticity of an angle operator by using trigonometric function of an angle operator as a basis for the quantum mechanical description. This approach, however, does not allow us to study the properties of the angle operator itself. By using a state space of an arbitrarily large yet finite number of dimensions, it is possible to introduce angle and orbital angular momentum as a conjugate pair of variables both represented by Hermitian operators. After physical and measurable quantities have been calculated in this state space the number of dimensions is allowed to tend to infinity in a limiting procedure.

In this thesis we have adopted the latter approach as it allows for a rigorous examination of the properties of the angle and orbital angular momentum operators. This chapter contains a review of the formal description of angle and orbital angular momentum in quantum mechanics.

1 Introduction

The correct description of a periodic variable such as a rotation angle or the optical phase has been a long standing problem in quantum mechanics. At the base of the problem is the question of whether the variable itself is restricted in a 2π radian range or whether it evolves continuously without bound. In the latter case the periodicity of the system would be contained in derived physical quantities, which would be periodic in the variable. For the purpose of this thesis we are mostly concerned with the azimuthal rotation angle as conjugate variable to the orbital angular momentum (OAM) around a beam axis. The difficulties in the quantum mechanical description, however, are directly connected to the periodicity of the variable, and similar problems arise for the conjugate pair of photon number and optical phase. Naturally both settings are related to the correct description of the phase of a harmonic oscillator. This is why most of the literature we are referring to in this chapter covers the question of the phase of the harmonic oscillator or equivalently the phase of a single mode optical field.

If the phase of the harmonic oscillator χ and the rotation angle φ are treated as

continuously evolving variables, the classical Poisson brackets with the Hamiltonian H, or the OAM along the rotation axis L_z, have a standard form for a conjugate pair of variables (Pegg & Barnett, 1989; Barnett & Pegg, 1990). But a problem arises if these Poisson brackets are translated into quantum mechanical commutators according to the correspondence principle. Louisell (1963) noted that the commutator $[\hat{\chi}, \hat{N}] = \mathrm{i}$ shows inconsistencies. This commutator corresponds to the Poisson bracket $\{\chi, H\}$ because the Hamiltonian can be written in terms of the number operator $\hat{N}$ (Loudon, 2000). In the number state basis $\{|n\rangle\}_{n\in\mathbb{N}_0}$, where n can take on every positive integer value and zero, the matrix elements of the commutator are given by

$$\langle n|[\hat{\chi}, \hat{N}]|n'\rangle = \mathrm{i}\delta_{nn'}\,. \tag{III.1}$$

But this requires that all diagonal matrix element of the phase operator are undefined, as the number states are the eigenstates of the number operator with $\hat{N}|n\rangle = n|n\rangle$ and the matrix elements of the phase operator $\hat{\chi}$ are therefore given by

$$(n' - n)\langle n|\hat{\chi}|n'\rangle = \delta_{nn'}\,. \tag{III.2}$$

A similar inconsistency can be derived for OAM and angle (Barnett & Pegg, 1990). The commutator $[\hat{\varphi}, \hat{L}_z] = \mathrm{i}\hbar$ corresponding to the classical Poisson bracket $\{\varphi, L_z\} = 1$ leads to undetermined diagonal matrix elements of the angle operator:

$$(m' - m)\langle m|\hat{\varphi}|m'\rangle = \mathrm{i}\hbar\delta_{mm'}\,, \tag{III.3}$$

where $|m\rangle, m \in \mathbb{Z}$ are eigenstates of $\hat{L}_z$ with eigenvalues $\hbar m$. It is important to note that despite all similarities in the problem of finding the correct quantum mechanical description for either the phase or the rotation angle, there is a difference in the eigenvalue spectrum of the respective conjugate variable. Whereas the angular momentum eigenstates $|m\rangle$ can take on all integer values, positive, negative and zero, the number states $|n\rangle$ are only defined for positive values of n and zero.

A different commutator is obtained if one assumes that the periodic variable is strictly bound within a 2π radian interval from θ_0 to $\theta_0 + 2\pi$. The angle or phase operator acts then not as simple multiplication, but it is modified by a series of step functions (Judge & Lewis, 1963):

$$Y(\varphi) = \varphi - 2\pi\sum_{p=1}^{\infty}\Theta(\varphi - \theta_0 - 2p\pi) + 2\pi\sum_{p=1}^{\infty}\Theta(-\varphi - \theta_0 - 2p\pi), \tag{III.4}$$

where $\Theta(x)$ is the unit step function with $\Theta(x) = 0$ for $x < 0$ and $\Theta(x) = 1$ for $x > 0$. The commutator $[\hat{\varphi}_{\theta_0}, \hat{L}_z]$ exhibits a series of Dirac-δ functions:

$$[\hat{\varphi}_{\theta_0}, \hat{L}_z] = \mathrm{i}\hbar[1 - 2\pi\delta_{2\pi}(\varphi - \theta_0)], \tag{III.5}$$

where the index θ_0 on the angle operator $\hat{\varphi}$ indicates the starting angle of the 2π radian range. The 2π-periodic δ-function is defined in App. V.B on page 118. Similar results can be obtained for the periodic phase variable (Pegg & Barnett, 1997). Interestingly, the correspondence between Poisson bracket and commutator is restored, if

the classical phase or angle variable is treated similarly (Barnett & Pegg, 1989); instead of a smoothly evolving variable, the angle or the phase show discontinuities whenever the end of the 2π radian is reached.

While this approach yields correct results for the matrix elements of the angle or phase operator in the basis states of the conjugate variable (Pegg et al., 1990), it fails to give well-defined matrix elements for the OAM operator $\hat{L}_z$ or number operator $\hat{N}$ in the respective basis states of the angle or phase operator (Barnett & Pegg, 1989, 1990). For the matrix elements of the number operator $\hat{N}$ in the phase state basis this is due to the missing negative numbers in the spectrum of the number operator (Pegg & Barnett, 1997). But even in the presence of negative quantum numbers, as in the case of the OAM operator $\hat{L}_z$, a careful analysis shows that the matrix elements in the angle state basis are undefined (Barnett & Pegg, 1990). This is because phase and angle states, as eigenstates of a Hermitian phase or angle operator, cannot be represented in an infinite dimensional Hilbert space. This shortcoming of the Hilbert space to accommodate angle or phase states has been seen as an indication that a Hermitian angle or phase operator does not exist. However, eigenstates of the linear momentum and position operator also are not elements of the Hilbert space $L^2(-\infty, \infty)$, on which their wavefunctions in position or momentum representation is defined. The usual identification as generalised functions in the form of Dirac δ-distributions involves no difficulties in most cases.

The search for a quantum mechanical description of the angle or phase operator has attracted some controversy (see (Pegg & Barnett, 1997) and the extensive bibliography therein). A very different approach exists which does not consider a Hermitian phase operator, but trigonometric operators (Susskind & Glogower, 1964):

$$\widehat{\exp}(\mathrm{i}\chi) = \sum_{n=0}^{\infty} |n\rangle\langle n+1|, \tag{III.6a}$$

$$\widehat{\exp}(-\mathrm{i}\chi) = \sum_{n=0}^{\infty} |n+1\rangle\langle n|. \tag{III.6b}$$

This approach has been developed for the phase variable. For a unique determination in the form given above Susskind & Glogower had to add the postulate that the projector onto the vacuum state $|\psi\rangle\langle 0|$ is absent in the expressions in Eq. (III.6). These operators are not unitary and do not commute. It is therefore not possible to deduce the existence of a unique phase operator in this approach, as the operators in Eq. (III.6) cannot be functions of a common phase operator. The approach by Susskind & Glogower can also be applied to angle and OAM (Carruthers & Nieto, 1968). In this case the operators are unitary, because of the inclusion of negative quantum numbers for the OAM. However, the angle operator deduced from the exponential operator acts on the infinite dimensional space and the pair of angle and OAM operator in this approach suffers from the same difficulties with the indefiniteness of matrix elements as mentioned above.

The problem with a phase or angle operator is that their eigenstates cannot be accommodated in an infinite dimensional space. It is, however, possible to define a Hermitian phase and angle operator on a finite dimensional space. To this end, Barnett & Pegg (1989, 1990) introduced a state space Ψ of $2L+1$ dimensions, where L

is arbitrary but finite. Within this state space the eigenstates of the operators form an orthonormal basis and all matrix elements are well defined. The operators are Hermitian and owing to the finite dimension also self-adjoint. After physical results have been calculated, such as expectation values of matrix elements, L is allowed to tend to infinity, restoring an infinite, but countable set of OAM or number states and a respective dense set of angle or phase states. The limiting procedure of starting in a space with a finite number of dimensions, which is allowed to tend to infinity only after measurable results have been calculated, is therefore crucial to the approach by Barnett & Pegg (1989). The mathematical reasoning behind this approach is that the convergence behaviour for operators is different from that of matrix elements or expectation values. The former are defined by their actions on elements of a Hilbert space whereas the latter are elements of the space of complex numbers. The question of convergence of the phase operator has been addressed by Vaccaro (1995) who introduced an infinite Hilbert space on which a unitary and Hermitian phase operator can be defined. The phase operator introduced by Barnett & Pegg (1989) converges strongly on this particular Hilbert space.

In the following we present the quantum mechanical description of rotation angles in the approach by Barnett & Pegg (1990) in more detail. This is a very elegant formulation which allows us to investigate properties of the angle and OAM operators directly. In the remainder of the thesis we will refer to this chapter repeatedly.

2 Quantum theory of rotation angles

The elegance in the finite dimensional approach lies in the fact that all desired properties of a pair of conjugate operators occur naturally within this method. Angle and OAM are complementary variables, they act as mutual generators and they also give the correct commutator analogous to Eq. (III.5) on page 40. Obviously an elegant formulation alone is not a proof for the validity of a theory. A consistent formulation though is a necessary requirement and in this section we show that the finite-dimensional approach by Barnett & Pegg (1990) describes angle and OAM consistently.

In this section we look first at the definition of angle and OAM states in the finite dimensional state space. With these states we can then proceed to calculate the commutator for angle and OAM. Angle states cannot be prepared in an experiment, and we therefore review the concept of physical states at the end of this section.

2.1 Angle and orbital angular momentum states

The finite-dimensional space state Ψ is spanned by $2L+1$ eigenstates $|m\rangle$ of the OAM operator $\hat{L}_z$. The action of $\hat{L}_z$ on one of the eigenstates is defined by $\hat{L}_z|m\rangle = m\hbar|m\rangle$. The positive integer L is arbitrary and the quantum number m takes on the values $-L, -L+1, \ldots, L$. Angle states can be unambiguously decomposed in the basis of the OAM states by writing:

$$|\varphi\rangle = \frac{1}{\sqrt{2L+1}} \sum_{m=-L}^{L} \exp(-im\varphi)|m\rangle. \tag{III.7}$$

In this form angle states have a periodic structure; the state $|\varphi\rangle$ is the same as $|\varphi + 2\pi\rangle$. This is only natural, as rotation angles differing by 2π radians are indistinguishable from each other. All distinct angle states are therefore contained in an interval ranging from an arbitrary starting angle θ_0 up to, but not including, $\theta_0 + 2\pi$, that is $\varphi \in [\theta_0, \theta_0 + 2\pi)$. This interval includes an infinite number of angle states which are not all mutually orthogonal:

$$\langle\varphi|\varphi'\rangle = \frac{1}{2L+1}\frac{\sin[(2L+1)(\varphi'-\varphi)/2]}{\sin[(\varphi'-\varphi)/2]}. \tag{III.8}$$

The states $|\varphi\rangle$ and $|\varphi'\rangle$ are orthogonal if the difference $\varphi' - \varphi$ is a nonzero, integer multiple of $2\pi/(2L+1)$. It is thus possible to form an orthonormal basis of $2L+1$ angle states $|\theta_n\rangle$ by selecting values $|\theta_n\rangle$ as

$$\theta_n = \theta_0 + \frac{2\pi n}{2L+1}, \quad n = 0, 1, \ldots, 2L. \tag{III.9}$$

With this choice for the allowed angles all states in the set $\{|\theta_n\rangle\}_{n=0,\ldots,2L}$ are orthonormal to each other:

$$\langle\theta_n|\theta_{n'}\rangle = \delta_{nn'}. \tag{III.10}$$

It is very important to note that the choice of θ_0 is arbitrary and determines the particular basis set. Angle states from two different basis sets, with two different values for θ_0, will not be orthogonal in general. The choice of θ_0 determines a particular angle operator which will not commute with other angle operators for a different starting angle θ_0. The angle operator $\hat{\varphi}_{\theta_0}$ thus has to be labelled with the starting angle θ_0 indicating that the eigenvalues are given by Eq. (III.9):

$$\hat{\varphi}_{\theta_0} = \sum_{n=0}^{2L} \theta_n |\theta_n\rangle\langle\theta_n| = \theta_0 + \sum_{n=0}^{2L} \frac{2\pi n}{2L+1} |\theta_n\rangle\langle\theta_n|. \tag{III.11}$$

Angle and OAM are both defined on a space of $2L+1$ dimensions. In this formulation both basis sets can be used to represent any element of the state space Ψ equally well. In particular the angle basis can be represented in the basis states of the OAM operator and vice versa. It is perhaps surprising that in the finite-dimensional approach the angle and OAM bases span the same state space, while the two state spaces for OAM and angle are different in character when working directly in infinite dimensions. There, the OAM is characterised as a discrete and unbounded variable, while the angle is bounded and continuous. However, in mathematics a number of projections are known which map an unbounded domain on to a compact range. For example the whole complex plane can be projected onto a Riemann sphere (Silverman, 1984). With an increasing L the interval for $m \in [-L, L]$ is expanded, while the interval for the angle $\theta_n \in [\theta_0, \theta_0 + 2\pi]$ remains the same. The spacing of two consecutive angles, however, decreases with an increasing L. In the limit of $L \to \infty$ the OAM eigenvalues take on all numbers $m \in \mathbb{Z}$, while the angle eigenvalues form a dense set on the circle $[\theta_0, \theta_0 + 2\pi]$. The starting angle θ_0 retains its importance in the limit of $L \to \infty$. As mentioned above, θ_0 determines a particular angle operator and the uncertainty and variance of this operator will be different from other angle operators for different

values of θ_0.

2.2 Commutator $[\hat{\varphi}_{\theta_0}, \hat{L}_z]$

The expectation value of the commutator determines the uncertainty relation, which will be studied in more detail in the next chapter. We therefore derive the commutator for the angle and OAM operator in both representations. In order to express the angle operator in the OAM basis we have to calculate the matrix elements in the expansion

$$\hat{\varphi}_{\theta_0} = \sum_{m,m'=-L}^{L} |m'\rangle\langle m'|\hat{\varphi}_{\theta_0}|m\rangle\langle m|. \tag{III.12}$$

On using the completeness of the angle states we can calculate the matrix elements $\langle m'|\hat{\varphi}_{\theta_0}|m\rangle$ as a sum over exponentials:

$$\langle m'|\hat{\varphi}_{\theta_0}|m\rangle = \sum_{n=0}^{2L} \theta_n \langle m'|\theta_n\rangle\langle\theta_n|m\rangle = \frac{1}{2L+1}\sum_{n=0}^{2L} \theta_n \exp[\mathrm{i}(m-m')\theta_n]. \tag{III.13}$$

The diagonal matrix elements involve a summation of the positive integer n from zero to $2L$. This sum is given by $(2L+1)L$, which allows us to evaluate the diagonal matrix elements to:

$$\langle m|\hat{\varphi}_{\theta_0}|m\rangle = \theta_0 + \frac{2\pi L}{2L+1}. \tag{III.14}$$

For the off-diagonal elements geometric progression (see App. VI.C on page 159) can be used to calculate the sum:

$$\langle m'|\hat{\varphi}_{\theta_0}|m\rangle = \frac{2\pi}{2L+1}\frac{\exp[\mathrm{i}(m-m')\theta_0]}{\exp[\mathrm{i}(m-m')2\pi/(2L+1)]-1}. \tag{III.15}$$

With the help of these matrix elements we can express the angle operator in the OAM basis as

$$\hat{\varphi}_{\theta_0} = \theta_0 + \frac{2\pi}{2L+1}\left[L + \sum_{\substack{m,m'=-L \\ m\neq m'}}^{L} \frac{\exp[\mathrm{i}(m-m')\theta_0]|m'\rangle\langle m|}{\exp[\mathrm{i}(m-m')2\pi/(2L+1)]-1}\right]. \tag{III.16}$$

Using the expression of the angle operator in the OAM basis in Eq. (III.16) we can write the commutator in the same basis as

$$[\hat{\varphi}_{\theta_0}, \hat{L}_z] = \frac{2\pi\hbar}{2L+1}\sum_{\substack{m,m'=-L \\ m\neq m'}}^{L} \frac{(m-m')\exp[\mathrm{i}(m-m')\theta_0]|m'\rangle\langle m|}{\exp[\mathrm{i}(m-m')2\pi/(2L+1)]-1}. \tag{III.17}$$

To express the commutator in the angle state basis we can use the representation of $|m\rangle$ in terms of angle states:

$$|m\rangle = \sum_{n=0}^{2L} |\theta_n\rangle\langle\theta_n|m\rangle = \frac{1}{\sqrt{2L+1}}\sum_{n=0}^{2L} \exp(\mathrm{i}m\theta_n)|\theta_n\rangle, \tag{III.18}$$

to write the OAM operator in the angle state basis:

$$\hat{L}_z = -\frac{\hbar}{2} \sum_{\substack{n,n'=0 \\ n \neq n'}}^{2L} \frac{(-1)^{n-n'} |\theta_{n'}\rangle\langle\theta_n|}{\sin[(n-n')\pi/(2L+1)]}. \tag{III.19}$$

The commutator can then be written in the angle state basis as:

$$[\hat{\varphi}_{\theta_0}, \hat{L}_z] = \mathrm{i}\frac{\hbar\pi}{2L+1} \sum_{\substack{n,n'=0 \\ n \neq n'}}^{2L} \frac{(n-n')(-1)^{n-n'} |\theta_{n'}\rangle\langle\theta_n|}{\sin[(n-n')\pi/(2L+1)]}. \tag{III.20}$$

The commutator has well defined matrix elements and does not suffer from the difficulties mentioned in the introduction to this chapter.

2.3 Physical states

Angle states are impossible to realise experimentally. For all practical purposes it is more interesting to consider physical states $|p\rangle$ for which all the moments of the OAM operator $\hat{L}_z$ are finite. Within the finite state space Ψ this condition seems to be fulfilled trivially. As long as the sum of the coefficients $\langle m|p\rangle$ in the decomposition of $|p\rangle$ into the angular momentum basis

$$|p\rangle = \sum_{m=-L}^{L} \langle m|p\rangle |m\rangle \tag{III.21}$$

is finite, any finite moment $\langle p|\hat{L}_z^n|p\rangle, n \in \mathbb{N}$ will be finite. But a moment of an operator is a physical quantity for which we can allow L to tend to infinity. If we require all moments to be finite even in the limit of $L \to \infty$, the coefficients have to fall off sufficiently quickly. Such states may be approximated to any degree by an expansion in the form of Eq. (III.21) where the coefficients $\langle m|p\rangle$ are zero for $|m| > M$. The bound M is as large as required for the desired accuracy in the approximation but less than L. If we restrict the domain of $\hat{\varphi}_{\theta_0}$ to physical states, we can employ a physical angle operator in the limit of large L by expanding the exponential in the denominator:

$$(\hat{\varphi}_{\theta_0})_p = \theta_0 + \pi - \mathrm{i} \sum_{\substack{m,m \\ m \neq m'}} \frac{\exp[\mathrm{i}(m-m')\theta_0]}{m-m'} |m'\rangle\langle m|. \tag{III.22}$$

We should stress that this is a deviation from the rule that the limit of $L \to \infty$ is only to be taken after physical results have been calculated. This is why the approximated form in Eq. (III.22) can only replace $\hat{\varphi}_{\theta_0}$ when operating on physical states, and not on the angle states themselves, which cannot be approximated in the same way as physical states. Using the physical form of the angle operator we can also give a specialised expression for the commutator when acting on physical states:

$$[\hat{\varphi}_{\theta_0}, \hat{L}_z]_p = -\mathrm{i}\hbar \sum_{\substack{m,m' \\ m \neq m'}} \exp[\mathrm{i}(m-m')\theta_0] |m'\rangle\langle m|. \tag{III.23}$$

It is possible to calculate matrix elements for this physical commutator without another limiting procedure. Again, this is a deviation from the normal procedure, where the limit is taken only after expectation values and matrix elements have been calculated. But for the matrix elements in the OAM basis the relevant limit has been applied already by forming the physical angle operator. The physical commutator has matrix elements given by

$$\langle m'|[\hat{\varphi}_{\theta_0}, \hat{L}_z]_p]|m\rangle = -\mathrm{i}\hbar(1-\delta_{mm'})\exp[\mathrm{i}(m-m')\theta_0]. \tag{III.24}$$

Orbital angular momentum eigenstates are physical states and this is why we can use the physical form of the commutator to calculate the matrix elements in Eq. (III.24). However, the coefficients in the decomposition of an OAM eigenstate into the angle basis do not fall off. Therefore the commutator in Eq. (III.20) on the previous page needs to contain values of $n-n'$ up to $2L+1$. It is possible to substitute the definition of angle states in terms of OAM states (see Eq. (III.7) on page 42)

$$|\theta_n\rangle = \frac{1}{\sqrt{2L+1}} \sum_{m=-L}^{L} \exp(-\mathrm{i}m\theta_n)|m\rangle \tag{III.25}$$

into the expression for the physical commutator in the OAM basis:

$$[\hat{\varphi}_{\theta_0}, \hat{L}_z]_p = \mathrm{i}\hbar[1-(2L+1)|\theta_0\rangle\langle\theta_0|]. \tag{III.26}$$

For any physical state $|p\rangle$ the expectation value of the commutator will be

$$\langle p|[\hat{\varphi}_{\theta_0}, \hat{L}_z]_p|p\rangle = \mathrm{i}\hbar[1-(2L+1)|\langle p|\theta_0\rangle|^2]. \tag{III.27}$$

If we introduce an angle probability distribution $P(\theta_n)$ within the finite state space Ψ, the probability that the angle is found within $\delta\theta$ from θ_n is given by $P(\theta_n)\delta\theta$. The distance between two consecutive angles $\delta\theta$ is given by $\delta\theta = 2\pi/(2L+1)$. With this probability density the expectation value of the commutator can be written as

$$\langle p|[\hat{\varphi}_{\theta_0}, \hat{L}_z]_p|p\rangle = \mathrm{i}\hbar[1-2\pi P(\theta_0)]. \tag{III.28}$$

In the limit of $L \to \infty$ the spacing between two consecutive angles tends to zero and the probability distribution turns into a probablity density of the continuous variable φ, which is normalised according to:

$$\int_{\theta_0}^{\theta_0+2\pi} P(\varphi)\mathrm{d}\varphi = 1. \tag{III.29}$$

In later chapters we will be using the continuous probability density when we deal with physical states. The expectation value of the commutator also dictates the lower bound for the uncertainty product. For physical states we can write

$$(\Delta L_z)_p(\Delta\varphi)_p \geq \frac{\hbar}{2}|1-2\pi P(\theta_0)|, \tag{III.30}$$

but it is important to keep in mind that this is an approximated form, which is derived from the exact form in Eq. (III.20) on page 45. The angle uncertainty $(\Delta\varphi)_p$ is thus given by the square root of the angle variance for a physical state $(\Delta\varphi)_p^2$:

$$(\Delta\varphi)_p^2 = \int_{\theta_0}^{\theta_0+2\pi} \varphi^2 P(\varphi)\mathrm{d}\varphi - \left(\int_{\theta_0}^{\theta_0+2\pi} \varphi P(\varphi)\mathrm{d}\varphi\right)^2. \tag{III.31}$$

For the OAM uncertainty the formulation in terms of a continuous angle representation requires more care. A direct representation of $\hat{L}_z$ as derivative with respect to φ can lead to problems even for physical states. A physical state was defined above as a state for which all moments of the OAM operator $\hat{L}_z$ are finite. This includes the second moment $\langle\hat{L}_z^2\rangle_p$ which forms part of the OAM variance. The coefficients $\langle m|p\rangle$ in the decomposition in Eq. (III.21) on page 45 thus have to fall off as m^{-2} for large m. As pointed out by Pegg et al. (2005), states with this fall-off characteristic, show a discontinuity in the gradient of the angle representation $p(\varphi) = \langle\varphi|p\rangle$. Acting with $\hat{L}_z^2$ on such a state can lead to singularities which render the states unnormalisable. By using the self-adjointness of the OAM operator we can represent $\langle\hat{L}_z^2\rangle_p$ by

$$\langle\hat{L}_z^2\rangle_p = \int_{\theta_0}^{\theta_0+2\pi} \left|-\mathrm{i}\hbar\frac{\mathrm{d}p(\varphi)}{\mathrm{d}\varphi}\right|^2 \mathrm{d}\varphi. \tag{III.32}$$

The OAM uncertainty is thus given by the square root of

$$(\Delta L_z)_p^2 = \hbar^2 \int_{\theta_0}^{\theta_0+2\pi} \left|\frac{\mathrm{d}p}{\mathrm{d}\varphi}\right|^2 \mathrm{d}\varphi + \hbar^2 \left(\int_{\theta_0}^{\theta_0+2\pi} \frac{\mathrm{d}p}{\mathrm{d}\varphi}\mathrm{d}\varphi\right)^2. \tag{III.33}$$

In this class of physical states fall the intelligent states, which satisfy the equality in the uncertainty relation. Another class of physical states are the minimum uncertainty product states, which give a constrained or global minimum in the uncertainty product. The next chapter examines the properties of both kind of states.

3 Summary

In this chapter we have presented a consistent description of angle and orbital angular momentum (OAM) as a pair of conjugate variables in quantum mechanics. In this approach angle and OAM are defined on a finite-dimensional state space of arbitrary dimension. Starting in a finite dimensional space and using a limiting procedure to allow the number of dimensions to tend to infinity after physical results have been calculated, avoids inconsistencies which we have highlighted in the first part of this chapter. The second part introduces the quantum theory for rotation angles in some detail. In later chapters we will derive aspects of the OAM of light with the help of this theory.

Bibliography

Barnett, S. M. & Pegg, D. T. (1989), 'On the Hermitian optical phase operator', *Journal*

of Modern Optics **36**(1), 7–19.

Barnett, S. M. & Pegg, D. T. (1990), 'Quantum theory of rotation angles', *Physical Review A* **41**(7), 3427–3435. DOI: 10.1103/PhysRevA.41.3427.

Carruthers, P. & Nieto, M. M. (1968), 'Phase and Angle Variables in Quantum Mechanics', *Reviews of Modern Physics* **40**(2), 411–440. DOI: 0.1103/RevModPhys.40.411.

Judge, D. & Lewis, J. T. (1963), 'On the commutator $[L_z, \varphi]$', *Physics Letters* **5**(3), 190.

Loudon, R. (2000), *The Quantum Theory of Light*, 3rd edn, Oxford University Press, Oxford. ISBN: 0-19-850176-5.

Louisell, W. H. (1963), 'Amplitude and phase uncertainty relations', *Physics Letters* **7**(1), 60–61. DOI: http://dx.doi.org/10.1016/0031-9163(63)90442-6.

Pegg, D. T. & Barnett, S. M. (1989), 'Phase properties of the quantized singemode eletromagnetic field', *Physical Review A* **39**(4), 1665–1675. DOI: 10.1103/PhysRevA.39.1665.

Pegg, D. T. & Barnett, S. M. (1997), 'Tutorial review: Quantum optical phase', *Journal of Modern Optics* **44**(2), 225–264.

Pegg, D. T., Barnett, S. M., Zambrini, R., Franke-Arnold, S. & Padgett, M. (2005), 'Minimum uncertainty states of angular momentum and angular position', *New Journal of Physics* **7**, 62. DOI: 10.1088/1376-2630/7/1/062.

Pegg, D. T., Vaccaro, J. A. & Barnett, S. M. (1990), 'Quantum-optical phase and canonical conjugation', *Journal of Modern Optics* **37**(11), 1703–1710.

Silverman, R. A. (1984), *Complex Analysis with Applications*, Dover Publications, Inc., Mineola, New York. Reprint. Originally published: Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1973, c1974. ISBN: 0-486-64762-5.

Susskind, L. & Glogower, J. (1964), 'Quantum mechanical phase and time operator', *Physics* **1**, 49–61.

Vaccaro, J. A. (1995), 'Phase operators on Hilbert space', *Physical Review A* **51**(4), 3309–3317. DOI: 10.1103/PhysRevA.51.3309.

IV Chapter

Special states for the angular uncertainty relation

The uncertainty product for two non-commuting observables has a lower bound given by the expectation value of the commutator for the two observables. For angular momentum and angle this lower bound depends on the form of the state for which the uncertainties and the expectation values are calculated. This requires a distinction between states satisfying the equality in the uncertainty relation, that is intelligent states, and states giving a minimum in the uncertainty product. In this chapter we explain the difference between those two kinds of states and derive their defining equations.

1 Introduction

The uncertainty principle limits the precise knowledge of all physical quantities of a system. It finds its most familiar expression in the Heisenberg uncertainty principle (Heisenberg, 1927), where the product of the uncertainties for linear momentum Δp_x and position Δx has a constant lower bound $\hbar/2$:

$$\Delta p_x \Delta x \geq \frac{\hbar}{2}, \tag{IV.1}$$

and equivalently for the y and z direction. For such a constant lower bound states satisfying the equality in the uncertainty relation also minimise the uncertainty product. In general however, the lower bound depends on the form of the state for which the uncertainties are calculated. This is because for two arbitrary Hermitian operators $\hat{A}$ and $\hat{B}$ the lower bound in the uncertainty relation is given by the expectation value of the commutator $[\hat{A}, \hat{B}]$ (Robertson, 1929):

$$\Delta A \Delta B \geq \frac{1}{2}|\langle[\hat{A}, \hat{B}]\rangle|. \tag{IV.2}$$

In this case states which give a minimum in the uncertainty product $\Delta A \Delta B$ do not necessarily fulfil the equality (IV.2), but will always obey the inequality. On the other hand states satisfying the equality in the uncertainty relation do not need to have the smallest possible uncertainty product, as there might be states which have a smaller uncertainty product, without obeying the equality in the uncertainty relation. It is therefore required to distinguish between states satisfying the equality in Eq. (IV.2), that is intelligent states (Aragone et al., 1974, 1976), and minimum uncertainty product states. Mathematically the distinction arises from two different eigenvalue equa-

tions for intelligent states and minimum uncertainty product states. For the Heisenberg uncertainty relation the solutions to the eigenvalue equations are identical and intelligent states and minimum uncertainty product states have the same Gaussian wavefunction.

The uncertainty relation for orbital angular momentum (OAM) and angle has a state dependent lower bound. In its approximate form (Barnett & Pegg, 1990) (cf. Eq. (III.30) on page 46)

$$\Delta L_z \Delta\varphi \geq \frac{\hbar}{2} |1 - 2\pi P(\theta_0)| \tag{IV.3}$$

the lower bound depends on the probability density $P(\theta_0)$ at the edge of the chosen 2π radian interval for the angle φ. The probability density for OAM eigenstates is a flat function with $P(\varphi) = 1/(2\pi)$ for all φ in the interval $[\theta_0, \theta_0 + 2\pi)$. The lower bound in Eq. (IV.3) may thus become equal to zero, and the global minimum in the uncertainty product is obtained for the eigenstates of the OAM operator $\hat{L}_z$. But it is also possible to identify states which minimise the uncertainty product $\Delta L_z \Delta\varphi$ for a given uncertainty in the OAM or in the angle. These states are called constrained minimum uncertainty product (CMUP) states (Pegg et al., 2005). For the Heisenberg uncertainty relation (IV.1) on the previous page, that is for linear momentum and position, these CMUP states are identical to the global minimum uncertainty product states and the intelligent states.

The angular uncertainty relation (IV.3) differs from the linear case also in the interval on which the respective observables are defined. Whereas in the Heisenberg uncertainty relation the position is defined on an infinite interval $x, y, z \in (-\infty, \infty)$, the angle is bounded in the 2π radian interval $[\theta_0, \theta_0 + 2\pi)$. This allows solutions to be normalised in the angular case which are disregarded in the linear case on the grounds that they do not represent physical, normalisable states. The wavefunctions of these solutions are peaked at the edges of the 2π radian interval for the angle and consequently the angle uncertainty tends to be larger than for states with a wavefunction peaked in the centre of the interval. The intelligent and CMUP states thus appear in two variants with small and large angle uncertainties. The distinction is most apparent for the uncertainty product $\Delta L_z \Delta\varphi$ which is bounded by $\hbar/2$ for states with small angle uncertainties but unbounded for the large-uncertainty case.

The family of special states for the angular uncertainty relation has been studied both experimentally and theoretically in a series of papers. Intelligent states with small angle uncertainties have been first introduced by Franke-Arnold et al. (2004) to verify the form of the angular uncertainty relation in an optical experiment. The distinction between intelligent states and CMUP states in the small-uncertainty regime has been presented by Pegg et al. (2005). The work on the large-uncertainty regime is an original contribution to the literature and published in (Götte et al., 2005) for the intelligent states and in (Götte et al., 2006*b*) for the CMUP states. The analysis for the large-uncertainty states would be incomplete without a comparison to the small uncertainty case, and we therefore present the theory for the complete family of special states in this chapter.

2 Intelligent states

Intelligent states satisfy the equality in the uncertainty relation (IV.2) on page 49. We first derive a necessary and sufficient condition for the equality to hold in form of an eigenvalue equation, which will then be used to determine the intelligent states in both variants.

2.1 Equality condition

For two Hermitian operators $\hat{A}$ and $\hat{B}$ the Schrödinger form of the uncertainty relation is given by (Schrödinger, 1930)

$$(\Delta A)^2(\Delta B)^2 \geq \frac{1}{4}\left(\langle[\hat{A},\hat{B}]_+\rangle - 2\langle\hat{A}\rangle\langle\hat{B}\rangle\right)^2 + \frac{1}{4}|\langle[\hat{A},\hat{B}]\rangle|^2, \tag{IV.4}$$

where $[\hat{A},\hat{B}]_+ = \hat{A}\hat{B} + \hat{B}\hat{A}$ denotes the anti-commutator of the two operators $\hat{A}$ and $\hat{B}$. In the derivation of the uncertainty relation (Schwabl, 1995) positive terms related to the anti-commutator are omitted in order to recover the compact form given in Eq. (IV.2) [see App. IV.A]. For the equality in Eq. (IV.4) these terms become important as they strictly have to be equal to zero. We introduce two associated Hermitian operators $\hat{a} = \hat{A} - \langle\hat{A}\rangle$ and $\hat{b} = \hat{B} - \langle\hat{B}\rangle$ for the two Hermitian operators in Eq. (IV.4). Using these associated operators the uncertainty relation may be written as

$$(\Delta A)^2(\Delta B)^2 \geq \frac{1}{4}\langle[\hat{a},\hat{b}]_+\rangle^2 + \frac{1}{4}|\langle[\hat{A},\hat{B}]\rangle|^2. \tag{IV.5}$$

If we take the square on both sides in the Robertson form of the uncertainty relation (IV.2) on page 49 and compare this to Eq. (IV.5) we find that the term with the anti-commutator is additionally present in the Schrödinger form. This does not affect the validity of the inequality (IV.2) as the expectation value of the Hermitian operator $[\hat{a},\hat{b}]_+$ is always positive and can thus be omitted. However, for the equality to hold in Eq. (IV.2) this additional term has to be taken into account, and $\langle[\hat{a},\hat{b}]_+\rangle$ is required to be equal to zero:

$$\langle[\hat{a},\hat{b}]_+\rangle = \langle\hat{a}\hat{b} + \hat{b}\hat{a}\rangle = \langle\hat{a}\hat{b}\rangle + \langle\hat{b}\hat{a}\rangle = 0. \tag{IV.6}$$

It is worth mentioning that $\langle\hat{a}\hat{b} + \hat{b}\hat{a}\rangle$ is the quantum correlation function $C_{1,1}$ or the covariance for the original observables A and B (Bohm, 1951). By requiring $\langle\hat{a}\hat{b} + \hat{b}\hat{a}\rangle$ to be equal to zero we demand that A and B are uncorrelated with respect to $C_{1,1}$. In the derivation of the uncertainty relation [see App. IV.A] the Schwartz inequality is used to give an upper bound for the modulus square of the scalar product $\langle\hat{a}\psi|\hat{b}\psi\rangle$ in terms of the product $\langle\hat{a}^2\rangle\langle\hat{b}^2\rangle$. The first term turns into the lower bound in Eqs. (IV.2) and (IV.5), while the latter term is equal to the product of variances $(\Delta A)^2(\Delta B)^2$. The equality in the Schwartz inequality holds if and only if the two state vectors in the scalar product are linearly dependent or parallel. Applied to Eq. (IV.5) this means that there exists a complex number $z \in \mathbf{C}$ such that

$$|\hat{a}\psi\rangle = z|\hat{b}\psi\rangle \quad \text{or} \quad \langle\hat{a}\psi| = z^*\langle\hat{b}\psi|, \tag{IV.7}$$

due to the Hermiticty of $\hat{a}$ and $\hat{b}$. On substituting this parallelity condition into Eq. (IV.6) on the previous page we can ascertain that z has to be purely imaginary:

$$0 = \langle \hat{a}\psi|\hat{b}\psi\rangle + \langle \hat{b}\psi|\hat{a}\psi\rangle = (z^* + z)\langle\psi|\hat{b}^2|\psi\rangle = (z^* + z)(\Delta b)^2. \tag{IV.8}$$

The last step follows from the fact that $\langle\hat{b}\rangle = 0$ and therefore $(\Delta b)^2 = \langle\hat{b}^2\rangle$. The Hermitian operators $\hat{b}$ and $\hat{B}$ have the same variance $(\Delta b)^2 = (\Delta B)^2$, so that $(\Delta b)^2 = 0$ only for an the eigenstate of $\hat{B}$. This is a rather specific case, but as mentioned above we reach the global minimum of the uncertainty product $\Delta L_z \Delta\varphi$ for an eigenstate of the OAM operator, and this state also satisfies the equality in the uncertainty relation. More generally, however, for the equality (IV.8) to hold $z^* + z$ has to be equal to zero, which requires z to be purely imaginary. We can therefore write Eq. (IV.7) on the previous page as an eigenvalue equation with the real parameter $\lambda \in \mathbb{R}$:

$$|\hat{a}\psi\rangle = \mathrm{i}\lambda|\hat{b}\psi\rangle. \tag{IV.9}$$

If we now apply this equality condition to the angular uncertainty relation and identify $\hat{a} = \hat{L}_z - \langle\hat{L}_z\rangle$ and $\hat{b} = \hat{\varphi} - \langle\hat{\varphi}\rangle$ we arrive at the eigenvalue equation determining the intelligent states for OAM and angle (Franke-Arnold et al., 2004):

$$\left[\hat{L}_z - \langle\hat{L}_z\rangle\right]|g\rangle = \mathrm{i}\hbar\lambda\left[\hat{\varphi}_\theta - \langle\hat{\varphi}_\theta\rangle\right]|g\rangle, \tag{IV.10}$$

where $|g\rangle$ denotes an intelligent state.

Projecting Eq. (IV.10) onto $\langle\varphi|$ yields a differential equation if we identify in the angle representation the OAM operator $\hat{L}_z$ as derivative with respect to φ and $\hat{\varphi}_{\theta_0}$ as multiplictive operator $Y(\varphi) = \varphi + 2\pi n$:

$$\left[\mathrm{i}\hbar\frac{\partial}{\partial\varphi} + \langle\hat{L}_z\rangle + \mathrm{i}\hbar\lambda(Y(\varphi) - \langle\hat{\varphi}_\theta\rangle)\right] g(\varphi) = 0. \tag{IV.11}$$

Only for a continuous wavefunction is the representation of $\hat{L}_z$ as derivative with respect to φ well defined. Owing to the periodic character this continuous behaviour is also required for the end points of the interval $[\theta_0, \theta_0 + 2\pi]$. It is more convenient to work in units where $\hbar = 1$, but to the same effect we can introduce a dimensionless OAM operator $\hat{m} = \hat{L}_z/\hbar$. In terms of this operator the differential equation may be written without $\hbar$:

$$\left[\mathrm{i}\frac{\partial}{\partial\varphi} + \langle\hat{m}\rangle + \mathrm{i}\lambda(Y(\varphi) - \langle\hat{\varphi}_\theta\rangle)\right] g(\varphi) = 0. \tag{IV.12}$$

Solutions to this equation are the intelligent states for OAM and angle. The sign of λ determines the curvature of the wavefunction at $\theta_0 + \pi$, the centre of the allowed range. If we restrict this analysis to cases where the mean values are equal to zero, we find that the second derivative of $g(\varphi)$ is given by

$$\frac{\partial^2 g}{\partial\varphi^2} = \lambda^2 Y^2(\varphi)g(\varphi) - \lambda g(\varphi). \tag{IV.13}$$

For a positive λ this leads to a peak in the centre of allowed range and turning points

at $\varphi = \pm 1/\sqrt{\lambda}$. For negative λ the wavefunction is peaked at the edges and has no turning points.

2.2 Wavefunction of intelligent states

A family of solutions to the differential equation (IV.12) on the previous page is given by truncated exponentials, which are characterised by the parameter λ (Galindo & Pascual, 1990; Franke-Arnold et al., 2004):

$$g(\varphi) = N_g \exp(\mathrm{i}\langle\hat{m}\rangle\varphi) \exp\left(-\frac{\lambda}{2}(Y(\varphi) - \langle\hat{\varphi}\rangle)^2\right). \tag{IV.14}$$

Here, N_g is the normalisation constant for the intelligent states. The normalisation integral is calculated on the interval $[\theta_0, \theta_0 + 2\pi]$ and $Y(\varphi)$ can thus be replaced with φ. The resulting normalisation constant can be given in terms of the error function erf (Abramowitz & Stegun, 1974):

$$\frac{1}{N_g^2} = \int_{\theta_0}^{\theta_0+2\pi} \exp\left(-\lambda(\varphi - \langle\hat{\varphi}\rangle)^2\right) = \frac{\mathrm{erf}(\pi\sqrt{\lambda})}{\sqrt{\lambda/\pi}}. \tag{IV.15}$$

Discontinuities in the wavefunction lead to infinite uncertainties for the OAM, we therefore require the wavefunction $g(\varphi)$ to be continuous at the ends of the 2π radian interval such that $g(\theta_0) = g(\theta_0 + 2\pi)$. This ensures that the mean angle lies in the middle of its definition range so that $\langle\hat{\varphi}\rangle = \theta_0 + \pi$ and that the OAM mean is an integer. The mean $\langle\hat{m}\rangle$ contributes only to the phase of the wavefunction. We can therefore restrict the analysis of intelligent states to the case of zero OAM mean. This is because the unitary OAM shift operator $\exp(\mathrm{i}\hat{\varphi}_{\theta_0}k)$ shifts the angular spectrum uniformly by an integer number $k \in \mathbb{Z}$, but this does not change ΔL_z (Barnett & Pegg, 1990) [see App. IV.B]. Also, owing to the unitary character of $\exp(\mathrm{i}\hat{\varphi}_{\theta_0}k)$, the angle probability density $P(\varphi) = |g(\varphi)_{\theta_0}|^2$ is left invariant and so is $\Delta\varphi$. Therefore, $\exp(\mathrm{i}\hat{\varphi}_{\theta_0}k)|g\rangle$ is also an intelligent state with a shifted OAM mean. Until now we have not specified the starting point of the 2π radian interval θ_0 and the wavefunction g is given in complete general terms. But in the following we will specifically set $\theta_0 = -\pi$, such that the intelligent states have $\langle\hat{\varphi}\rangle = 0$. Jointly with choosing $\langle\hat{m}\rangle = 0$ this simplifies the notation in subsequent calculations and graphs without loss of generality.

The parameter λ determines the intelligent state completely. For $\lambda > 0$ we have wavefunctions with a peak in the middle of the allowed range of angles. The wavefunction for these states has the form of a truncated Gaussian and the intelligent states represented by such wavefunctions will have a smaller angle uncertainty than states with $\lambda < 0$ and a positively curved wavefunction with peaks at the edges of the chosen 2π radian interval. For $\lambda = 0$ the angle wavefunction is flat which corresponds to the OAM eigenstates. This particular state divides the small angle uncertainty regime with $\lambda > 0$ from the large angle uncertainty regime with $\lambda < 0$. For $\lambda < 0$ the argument of the error function in Eq. (IV.15) turns imaginary. It follows from the definition of the error function (Abramowitz & Stegun, 1974) that an error function

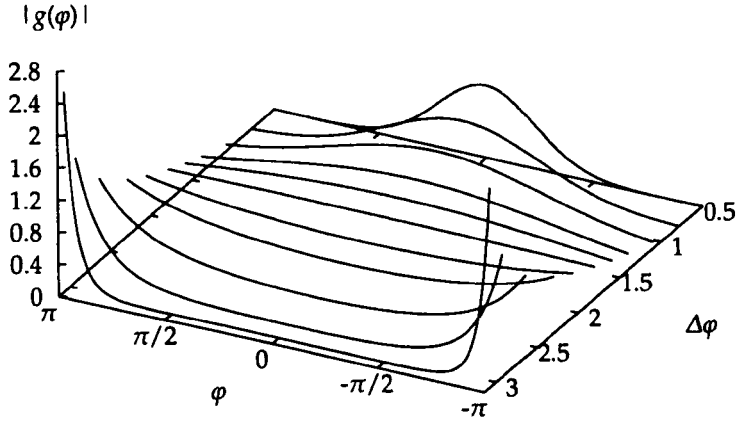

Figure IV.1 — The angle wavefunction of the intelligent states plotted for different values of $\Delta\varphi$. For $\Delta\varphi = \pi/\sqrt{3}$ [$\lambda = 0$] the wavefunction is flat and represents an OAM eigenstate. For smaller values of $\Delta\varphi$ [$\lambda > 0$] the wavefunction has the form of a truncated Gaussian and is peaked around $\varphi = 0$. Intelligent states which are peaked at the edges have an angle uncertainty $\Delta\varphi > \pi/\sqrt{3}$ [$\lambda < 0$].

with an imaginary argument is itself imaginary:

$$\operatorname{erf}(\mathrm{i}y) = \mathrm{i}\frac{2}{\sqrt{\pi}}\int_0^y \exp(\tau^2)\mathrm{d}\tau, \quad y \in \mathbb{R}. \tag{IV.16}$$

To plot the complex error function we use an approximation in terms of an infinite series (Abramowitz & Stegun, 1974):

$$\begin{aligned}\operatorname{erf}(x+\mathrm{i}y) = \operatorname{erf}(x) + \frac{\exp(-x^2)}{2\pi x}\left[(1-\cos(2xy)) + \mathrm{i}\sin(2xy)\right] + \\ + \frac{2}{\pi}\exp(-x^2)\sum_{n=1}^{\infty}\frac{\exp(-\frac{1}{4}n^2)}{n^2+4x^2}\left[\xi_n(x,y) + \mathrm{i}\zeta_n(x,y)\right] + \epsilon(x,y),\end{aligned} \tag{IV.17a}$$

where the functions ξ_n and ζ_n are given by

$$\xi_n(x,y) = 2x - 2x\cosh(ny)\cos(2xy) + n\sinh(ny)\sin(2xy), \tag{IV.17b}$$

$$\zeta_n(x,y) = 2x\cosh(ny)\sin(2xy) + n\sinh(ny)\cos(2xy). \tag{IV.17c}$$

The error in this expansion can be estimated to be $|\epsilon(x,y)| \approx 10^{-16}|\operatorname{erf}(x+\mathrm{i}y)|$. Within this error the approximation also gives an imaginary value for an imaginary argument. The graphs in the analysis of intelligent states by Galindo & Pascual (1990) and Götte et al. (2005) show the physical quantities of the intelligent states as a function of λ. In this thesis, however, we use the angle uncertainty $\Delta\varphi$ to characterise intelligent states in order to compare them with the constrained minimum uncertainty

product states (CMUP) states later in this chapter.

2.3 Angle and angular momentum uncertainty

In the following we present mathematical expressions and graphs for the uncertainties of angle and angular momentum and their product. In contrast to the treatment in the literature (Franke-Arnold et al., 2004; Götte et al., 2005), we give expression valid for intelligent states in both variants; for the small and large angle uncertainty case.

2.3.1 Angle uncertainty

The angle variance is given by $(\Delta\varphi)^2 = \langle\hat{\varphi}^2\rangle - \langle\hat{\varphi}\rangle^2$. As our solutions are symmetric about $\varphi = 0$, the mean value of $\hat{\varphi}$ is zero and the angle variance is simply the expectation value of $\hat{\varphi}^2$

$$(\Delta\varphi)^2 = \int_{-\pi}^{\pi} \varphi^2 P(\varphi)\mathrm{d}\varphi = \frac{1}{N_g^2}\int_{-\pi}^{\pi} \varphi^2 \exp(-\lambda\varphi^2)\mathrm{d}\varphi. \tag{IV.18}$$

The functional form of the wavefunction allows us to express the angle variance as the derivative of the normalisation constant N_g with respect to λ:

$$(\Delta\varphi)^2 = \frac{1}{N_g^2}\int_{-\pi}^{\pi} \frac{\mathrm{d}}{\mathrm{d}\lambda}\exp(-\lambda\varphi^2)\mathrm{d}\varphi = -\frac{1}{N_g^2}\frac{\mathrm{d}N_g^2}{\mathrm{d}\lambda}. \tag{IV.19}$$

Using the Leibniz formula to evaluate the derivative of the error function, an expression for the angle variance can be obtained on substituting the normalisation constant from Eq. (IV.15) on page 53 into Eq. (IV.19):

$$(\Delta\varphi)^2 = \left(\frac{1}{2\lambda} - \frac{\sqrt{\pi}\exp(-\pi^2\lambda)}{\sqrt{\lambda}\,\mathrm{erf}(\pi\sqrt{\lambda})}\right). \tag{IV.20}$$

The square of a self-adjoint operator is self-adjoint and hence we expect the variance of $\hat{\varphi}$ to be real. The angle uncertainty is thus given by the positive square root of Eq. (IV.20) and the expression holds for positive and negative λ. A plot of $\Delta\varphi$ as a function of λ is given in figure (IV.2) on the following page. In the limit of $\lambda \to \infty$ the angle uncertainty $\Delta\varphi$ tends to zero. But it is worth mentioning that the behaviour of $(\Delta\varphi)^2$ in Eq. (IV.20) is dominated by the $1/(2\lambda)$ term. For $\lambda \to -\infty$ the angle uncertainty $\Delta\varphi$ tends to its supremum π. That π is indeed the supremum of $\Delta\varphi$ for any kind of normalised state can be seen from an estimate of the integral in Eq. (IV.18):

$$(\Delta\varphi)^2 = 2\int_0^{\pi} \varphi^2 P(\varphi)\mathrm{d}\varphi \leq 2\pi^2\int_0^{\pi} P(\varphi)\mathrm{d}\varphi = \pi^2. \tag{IV.21}$$

For $\lambda = 0$ the wavefunction becomes flat and represents an OAM eigenstate. Although this is the global minimum uncertainty product state, it is also an intelligent state. The angle uncertainty for this states is given by

$$(\Delta\varphi)^2 = 2\int_0^{\pi} \varphi^2 \frac{1}{2\pi}\mathrm{d}\varphi = \frac{\pi^2}{3}. \tag{IV.22}$$

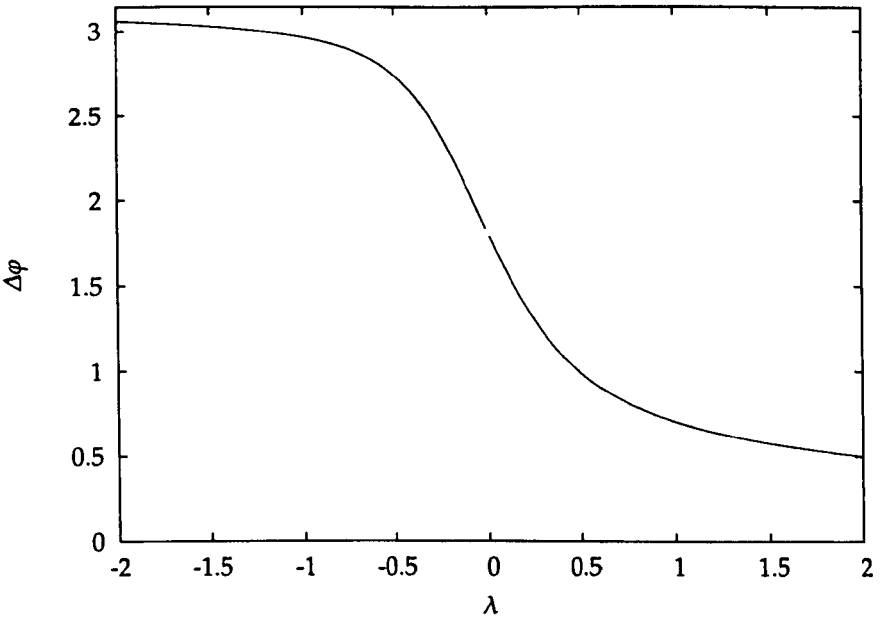

Figure IV.2 — The angle uncertainty $\Delta\varphi$ for the intelligent states. The parameter value $\lambda = 0$ distinguishes between the truncated Gaussians with $\lambda > 0$ (Franke-Arnold et al., 2004) and the large-uncertainty intelligent states with $\lambda < 0$ (Götte et al., 2005). The angle uncertainty is bounded by $\Delta\varphi \leq \pi$.

This particular state is the dividing point between small uncertainty intelligent states with $\Delta\varphi < \pi/\sqrt{3}$ and large-uncertainty intelligent states with $\Delta\varphi > \pi/\sqrt{3}$.

2.3.2 Angular momentum uncertainty

Using the continuous wavefunction in Eq. (IV.14) and the representation of $\hat{L}_z$ as a derivative requires some care. In section 2.1 we have argued that for the intelligent states the first derivative of $g(\varphi)$ with respect to φ is well behaved, while the second derivative is not. This is the reason why we can use $-i\hbar(d/d\varphi)$ as a valid representation of $\hat{L}_z$ in the eigenvalue equation for the intelligent states.

To calculate the OAM uncertainty ΔL_z from its variance $(\Delta L_z)^2 = \langle \hat{L}_z^2 \rangle - \langle \hat{L}_z \rangle^2$ we have to operate with $\hat{L}_z^2$ on an intelligent state. Identifying $\hat{L}_z^2$ with $-\hbar^2(d^2/d\varphi^2)$ does not lead to a well behaved result because of the discontinuity of the wavefunction at the edges of the 2π radian range. It is possible, however, to express $\langle \hat{L}_z^2 \rangle$ in terms of the first derivative with respect to φ. To prove this relation we have to represent the states in the arbitrarily large state space Ψ of $2L + 1$ dimensions (see Ch. III), in which $\hat{L}_z$ is Hermitian and self-adjoint. But as the relation is derived for expectation values, and therefore a physical quantity, it is admissible to take the limit of $L \to \infty$ and to revert to the notation of continuous wavefunctions afterwards. Moreover, for expectation values it is not relevant in which representation the operator is given or in which basis the intelligent state is decomposed. If we write $|g\rangle$ in its OAM decomposition in the state space Ψ

$$|g\rangle = \sum_{m=-L}^{L} g_m |m\rangle \tag{IV.23}$$

then we have for the expectation value $\langle \hat{L}_z^2 \rangle$:

$$\langle \hat{L}_z^2 \rangle = \sum_{m=-L}^{L} |g_m|^2 \langle m|\hat{L}_z^2|m\rangle = \sum_{m=-L}^{L} |g_m|^2 \langle \hat{L}_z m|\hat{L}_z m\rangle = \sum_{m=-L}^{L} |g_m|^2 \, |\hat{L}_z|m\rangle \, |^2. \quad \text{(IV.24)}$$

Here, $|\hat{L}_z|m\rangle|^2$ denotes the norm of the the state vector $\hat{L}_z|m\rangle$. As $|m\rangle$ is an eigenstate of $\hat{L}_z$ and obeys the orthogonality condition $\langle m|m'\rangle = \delta_{mm'}$, we can exchange the order in which we take the norm and evaluate the sum over m. The equality in Eq. (IV.24) may thus be continued by

$$\langle \hat{L}_z^2 \rangle = \sum_{m=-L}^{L} |g_m|^2 \, |\hat{L}_z|m\rangle \, |^2 = \left| \sum_{m=-L}^{L} \hat{L}_z g_m |m\rangle \right|^2 = |\hat{L}_z|g\rangle \, |^2. \quad \text{(IV.25)}$$

If we now revert to the infinite dimensional space we can represent $\hat{L}_z$ by $-i\hbar(d/d\varphi)$ in the angle representation. Taking the norm for the continuous wavefunction corresponds to the integral of the modulus square of the wavefunction over the full 2π radian interval:

$$\langle \hat{L}_z^2 \rangle = \int_{-\pi}^{\pi} \left| \frac{\hbar}{i} \frac{d}{d\varphi} g(\varphi) \right|^2 d\varphi. \quad \text{(IV.26)}$$

On substituting the wavefunction in Eq. (IV.14) on page 53 into this expression we find that the dimensionless OAM uncertainty Δm can be expressed in terms of the angle uncertainty $\Delta\varphi$:

$$(\Delta m)^2 = \frac{1}{N_g^2} \int_{-\pi}^{\pi} |\lambda|^2 \varphi^2 \exp(|\lambda|\varphi^2) d\varphi = |\lambda|^2 (\Delta\varphi)^2. \quad \text{(IV.27)}$$

To obtain the equality in the uncertainty relation for linear position and momentum, the uncertainties have to be inversely proportional to each other. In the angular case the uncertainties are directly proportional to each other; if we have a large angle uncertainty the OAM uncertainty is also large and yet the uncertainty product satisfies the equality in the uncertainty relation. The reason for this behaviour lies in the state dependent right hand side of Eq. (IV.3) on page 50.

A plot of the OAM uncertainty Δm as a function of the angle uncertainty $\Delta\varphi$ is shown in figure (IV.3) on the following page. As $\Delta\varphi$ tends to π for $\lambda \to -\infty$, the OAM uncertainty Δm tends to infinity. We have discussed the behaviour of $(\Delta\varphi)^2$ for $\lambda \to \infty$ earlier [see section IV 2.3.2] and noted that $1/(2\lambda)$ is the dominant term. The behaviour of Δm for $\lambda \to \infty$ follows from Eq. (IV.27) and we find that Δm varies as $\sqrt{|\lambda|/2}$ for $\lambda \to \infty$. This behaviour can be seen in the inset in figure (IV.3) on the following page.

2.3.3 Uncertainty product

The left hand side of the uncertainty relation (IV.3) on page 50 is given by the uncertainty product $\Delta L_z \Delta\varphi$. If we use the result for the dimensionless OAM uncertainty Δm from Eq. (IV.27) and the expression for $\Delta\varphi$ in Eq. (IV.18) on page 55 the

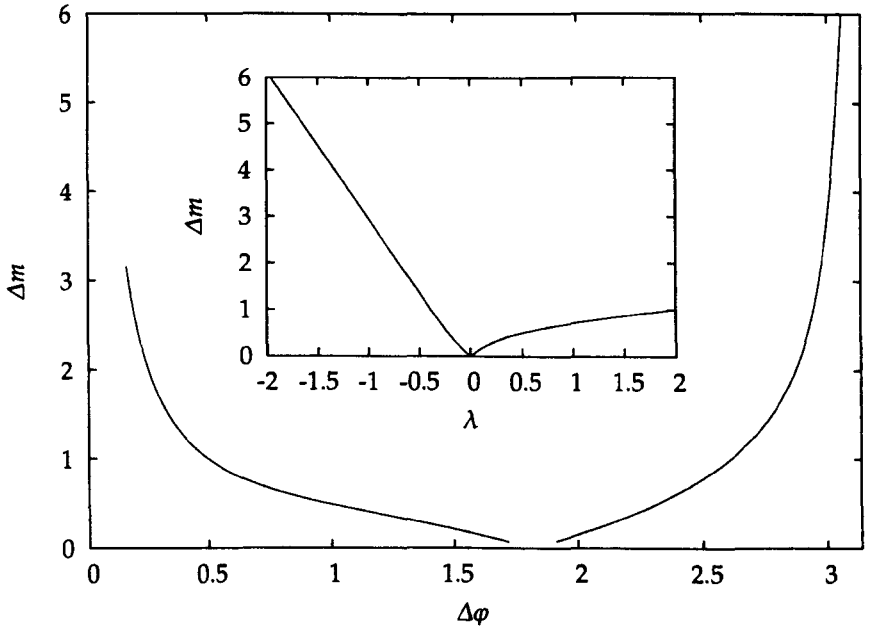

Figure IV.3 — The OAM uncertainty for intelligent states. The plot shows Δm as a function of $\Delta\varphi$. The plot is parameterised over λ and owing to the asymptotic behaviour of Δm as a function of λ a large parameter range has to be chosen in order to continue the plot for $\Delta\varphi \to 0$. The lack of plot points around $\Delta\varphi = \pi/\sqrt{3}$ is also a consequence of the parametrisation, as Δm and $\Delta\varphi$ are evaluated for equidistant values of λ. The behaviour of Δm as a function of λ is shown in the inset.

uncertainty product can be expressed as

$$\Delta m \Delta\varphi = |\lambda|(\Delta\varphi)^2 = \left|\lambda\left(\frac{1}{2\lambda} - \frac{\sqrt{\pi}\exp(-\pi^2\lambda)}{\sqrt{\lambda}\,\mathrm{erf}(\pi\sqrt{\lambda})}\right)\right|. \tag{IV.28}$$

As $(\Delta\varphi)^2$ is positive we can include it in the modulus. For our choice of θ_0 the left hand side of the uncertainty relation for $\Delta m\Delta\varphi$ is given by $|1 - 2\pi P(\pi)|/2$. If we substitute the expression for the wavefunction with the normalisation constant from Eq. (IV.15) on page 53 to calculate $P(\pi)$ we obtain

$$\frac{1}{2}|1 - 2\pi P(\pi)| = \frac{1}{2}\left|1 - 2\sqrt{\lambda}\pi\frac{\exp(-\pi^2\lambda)}{\mathrm{erf}(\pi\sqrt{\lambda})}\right| \tag{IV.29}$$

for the lower bound in the uncertainty relation. If we compare this expression with Eq. (IV.28), we can see that the two sides of the uncertainty relation are equal, as they should be, for intelligent states. In the limit of $\lambda \to -\infty$ the uncertainty product tends to infinity as $(\Delta\varphi)^2$ approaches π. The surprising result is that for the large-uncertainty intelligent states with $\lambda < 0$ the uncertainty product can become arbitrarily large and yet the equality in Eq. (IV.3) on page 50 is still satisfied. For the small-uncertainty intelligent states with $\lambda < 0$ the uncertainty product tends to $1/2$ because the $(\Delta\varphi)^2$ behaves dominantly as $1/(2\lambda)$ for $\lambda \to \infty$. The graph of the uncertainty product as a function of $\Delta\varphi$ is given in figure (IV.4) on the next page.

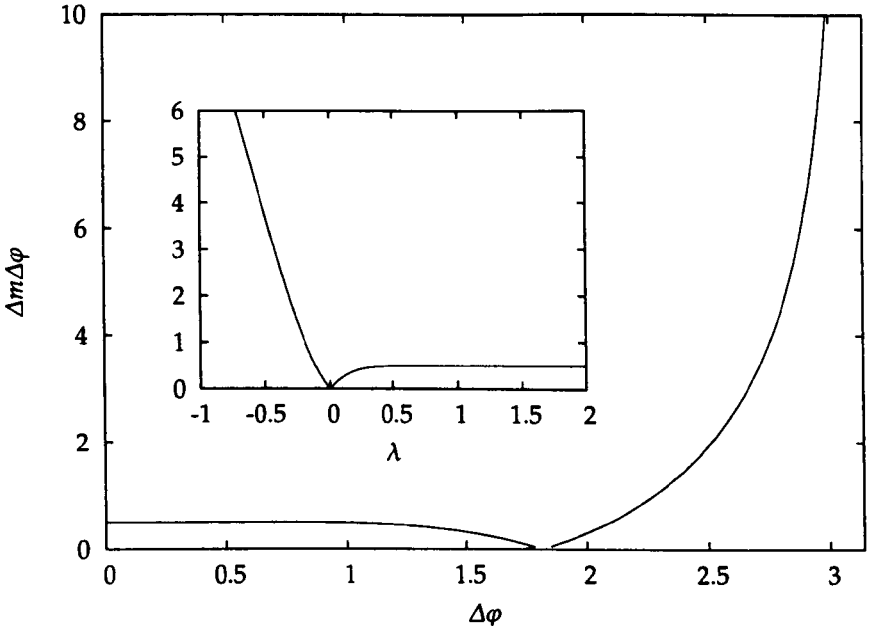

Figure IV.4 — The uncertainty product $\Delta m\Delta\varphi$ as a function of $\Delta\varphi$. For the large-uncertainty intelligent states ($\Delta\varphi > \pi/\sqrt{3}$ or $\lambda < 0$) the uncertainty product has no upper bound. The inset shows a plot of the uncertainty product against λ for comparison with the plots of $\Delta\varphi$ and Δm (see figures (IV.2) to (IV.3) on pages 56-58; see also a similar plot in (Galindo & Pascual, 1990)).

2.4 Orbital angular momentum distribution

The probability amplitudes g_m of the OAM are calculated from the wavefunction in Eq. (IV.14) on page 53 by means of a Fourier coefficient transform:

$$g_m = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} g(\varphi) \exp(-im\varphi) \mathrm{d}\varphi. \tag{IV.30}$$

By virtue of the symmetry of the wave function the transformation may be specialised to the Fourier cosine transform

$$g_m = \frac{1}{\sqrt{2\pi}} \int_{0}^{\pi} g(\varphi) \exp(im\varphi) \mathrm{d}\varphi + \text{c.c.}. \tag{IV.31}$$

Here, c.c. denotes the complex conjugate of the previous expression, and we have written $\cos(m\varphi)$ as $(\exp(im\varphi) + \text{c.c.})/2$. On substituting the wavefunction $g(\varphi)$ from Eq. (IV.14) into the Fourier transform, we can write the integral as an error function if we add and subsequently subtract a term to write the exponent as a complete square. In particular we can rewrite the integrand according to

$$\exp\left(-\frac{\lambda}{2}\varphi^2\right) \exp(im\varphi) = \exp\left(-\frac{m^2}{2\lambda}\right) \exp\left(\frac{m}{\sqrt{2\lambda}} + i\frac{\sqrt{\lambda}\varphi}{\sqrt{2}}\right)^2. \tag{IV.32}$$

The first exponential can be taken out of the integral and we can transform the whole exponent of the second exponential by setting

$$i\tau = \frac{m}{\sqrt{2\lambda}} + i\frac{\sqrt{\lambda}\varphi}{\sqrt{2}}, \tag{IV.33}$$

so that we can write the integral as

$$g_m = \frac{1}{\sqrt{2\pi}} N_g \exp\left(-\frac{m^2}{2\lambda}\right) \frac{\sqrt{2}}{\sqrt{\lambda}} \int_{\tau(0)}^{\tau(\pi)} \exp(-\tau^2) d\tau + \text{c.c.}. \tag{IV.34}$$

The upper and lower boundaries in the integral are given by $\tau(0) = -im/(\sqrt{2\lambda})$ and $\tau(\pi) = -im/(\sqrt{2\lambda}) + \sqrt{\lambda}\pi/\sqrt{2}$. The integral in the boundaries from $\tau(0)$ to $\tau(\pi)$ can be split up as a difference between two integrals, each starting at zero and ending at $\tau(\pi)$ and $\tau(0)$ respectively. These single integrals may be written as complex error functions

$$g_m = \frac{1}{\sqrt{2\pi}} N_g \exp\left(-\frac{m^2}{2\lambda}\right) \frac{\sqrt{\pi}}{\sqrt{2\lambda}} (\operatorname{erf}(\tau(\pi)) - \operatorname{erf}(\tau(0))) + \text{c.c.}. \tag{IV.35}$$

The term containing $\operatorname{erf}(\tau(0))$ will always cancel with its complex conjugate, regardless of the sign of λ. Either $\lambda > 0$ and $\operatorname{erf}(\tau(0))$ is imaginary, or $\lambda < 0$ and $\operatorname{erf}(\tau(0))$ is real and the factor $\sqrt{\pi}/\sqrt{2\lambda}$ is imaginary. In both cases only the $\operatorname{erf}(\tau(\pi))$ term is left in the calculation of the g_m. Adding the complex conjugate yields twice the real part of the entire expression:

$$g_m = 2\,\mathrm{Re}\left[\frac{1}{\sqrt{2\pi}} N_g \exp\left(-\frac{m^2}{2\lambda}\right) \frac{\sqrt{\pi}}{\sqrt{2\lambda}} \operatorname{erf}(\tau(\pi))\right]. \tag{IV.36}$$

Depending on whether λ is positive or negative we can write the g_m differently (see App. IV.C). For the truncated Gaussians with $\lambda > 0$ only the error function is a complex number. Consequently, the g_m may be written as

$$g_m = \frac{1}{\sqrt{\lambda}} N_g \exp\left(-\frac{m^2}{2\lambda}\right) \operatorname{Re}\operatorname{erf}\left(-i\frac{m}{\sqrt{2\lambda}} + \frac{\sqrt{\lambda}\pi}{\sqrt{2}}\right) \quad \lambda > 0. \tag{IV.37}$$

For the large-uncertainty intelligent states the factor $\sqrt{\pi}/\sqrt{2\lambda}$ is imaginary and the additional imaginary unit interchanges real and imaginary part of the complex error function. In this case we can set $\lambda = -|\lambda|$ an we have for the g_m

$$g_m = \frac{1}{\sqrt{|\lambda|}} N_g \exp\left(\frac{m^2}{2|\lambda|}\right) \operatorname{Im}\operatorname{erf}\left(-\frac{m}{\sqrt{2|\lambda|}} + i\frac{\sqrt{|\lambda|}\pi}{\sqrt{2}}\right) \quad \lambda < 0. \tag{IV.38}$$

The divergent exponential $\exp(m^2/(2|\lambda|))$ is counterbalanced by the imaginary part of the complex error function which decays to zero in this case. This can be seen if the approximate expression for the complex error function in Eq. (IV.17a) on page 54 is substituted into Eq. (IV.38).

The distribution of the probability amplitudes g_m for four different intelligent states is shown in figure (IV.5) on the next page. For the small-uncertainty intelligent

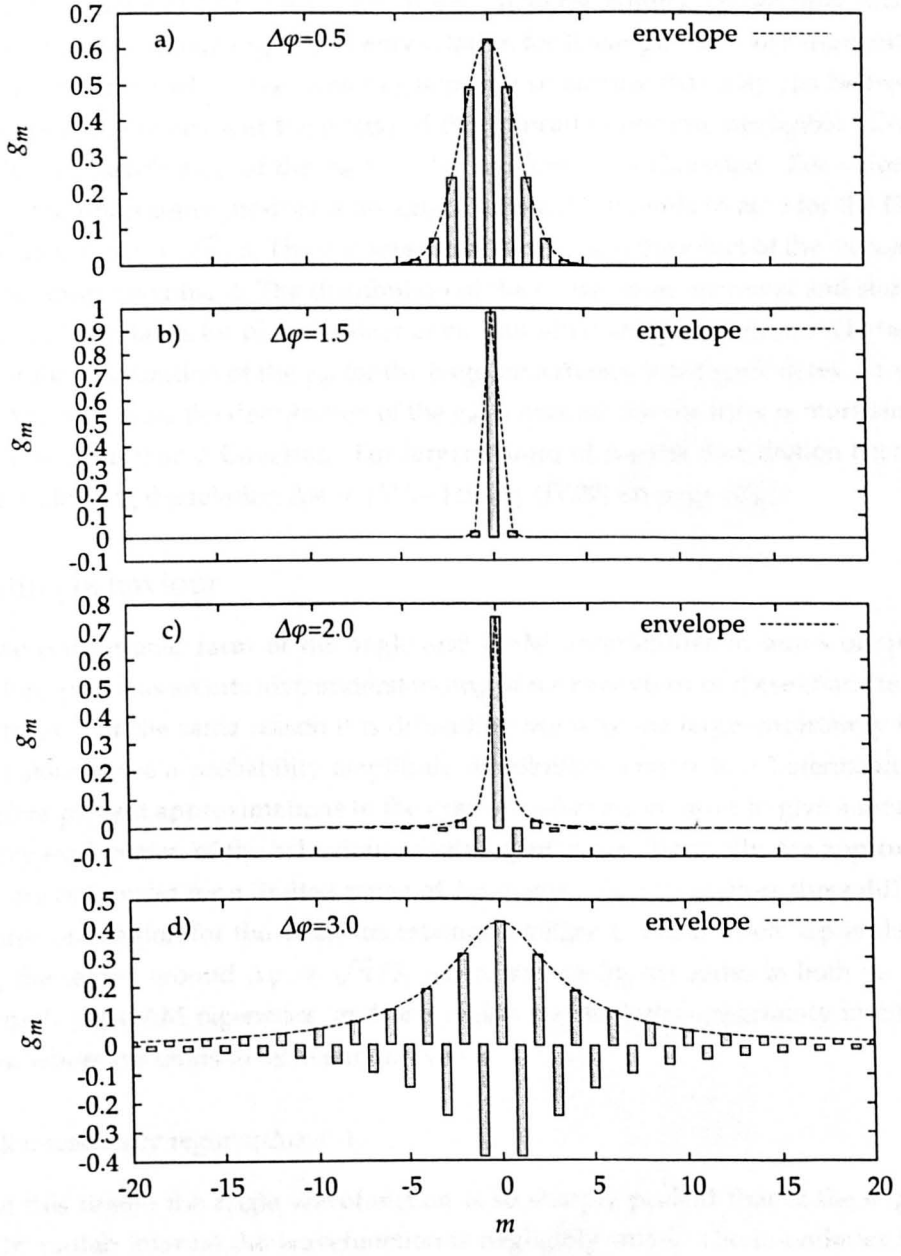

Figure IV.5 — The distribution of probability amplitudes g_m for the OAM decomposition of intelligent states. The series of plots shows the distribution of probability amplitudes g_m for four different intelligent states specified by their angle uncertainty $\Delta\varphi$. a) $\Delta\varphi = 0.5$ corresponds to a range in figure (IV.4) on page 59 where the uncertainty product is constant. The wavefunction is a narrow truncated Gaussian and the distribution of the g_m has also Gaussian form. b) for $\Delta\varphi = 1.5$ the uncertainty product differs from the linear case, and the effect of the truncation of the Gaussian wavefunction become more visible. The state is close to the OAM eigenstate and the distribution of the g_m is narrow. c) $\Delta\varphi = 2.0$ corresponds to a flat large-uncertainty intelligent state. The distribution is still narrow but alternating and has a form similar to a Lorentzian. d) $\Delta\varphi = 3.0$ is at the upper end of possible values for $\Delta\varphi$. $\Delta\varphi$ and Δm are proportional to each other, so a large angle uncertainty, for a wavefunction sharply peaked at $-\pi$ and π, corresponds to a wide distribution of the g_m.

states with $\Delta\varphi < \pi/\sqrt{3}$ one can distinguish two regions for the uncertainty product [see figure (IV.4) on page 59]. For $\Delta\varphi < 1$ the uncertainty product has roughly the constant value of $1/2$. This is the range where the angular uncertainty relation behaves like the Heisenberg uncertainty relation for linear position and momentum. It is also the range where the wavefunctions are so narrow that they can be treated as extended Gaussians and the effects of the truncation become negligible. Consequently, the distribution of the g_m has also the form of a Gaussian. For values of $\Delta\varphi > 1$ the uncertainty product is no longer constant but tends to zero for the OAM eigenstates at $\Delta\varphi = \sqrt{\pi}/3$. The wavefunction flattens and the effect of the truncation becomes more prominent. The distribution of the g_m becomes narrower and starts to alternate at the flanks for higher values of m. This alternating behaviour is characteristic for the distribution of the g_m for the large-uncertainty intelligent states. Close to the OAM eigenstate the distribution of the g_m is narrow, but the form is more similar to a Lorentzian than a Gaussian. For larger values of $\Delta\varphi$ the distribution becomes wider, following the relation $\Delta m = |\lambda|\Delta\varphi$ [cf. Eq. (IV.27) on page 57].

2.5 Limiting behaviour

The complicated form of the angle and OAM uncertainties in terms of special functions prevents an intuitive understanding of the behaviour of these characteristic quantities. For the same reason it is difficult to see why the large-uncertainty intelligent states have a probability amplitude distribution similar to a Lorentzian. We therefore present approximations to the exact expressions in order to give a more accessible explanation of the behaviour of intelligent states. Naturally, the approximations are only valid for a limited range of $\Delta\varphi$ values. We will explore three different regions: one region for the small-uncertainty intelligent states, where $\Delta\varphi$ is close to zero, the region around $\Delta\varphi = \sqrt{\pi}/3$, where the intelligent states in both varieties approach the OAM eigenstate, and one region for the large-uncertainty intelligent states, where $\Delta\varphi$ tends to its maximum value.

2.5.1 Small uncertainty regime: $\Delta\varphi < 1$

In this region the angle wavefunction is so sharply peaked that at the edges of the 2π radian interval the wavefunction is negligibly small. The boundaries in the integrals to calculate the angle uncertainty or the OAM probability amplitudes may thus be extended to infinity. The wavefunction in this approximation is an extended Gaussian and is given here in its normalised form:

$$g_{\mathrm{sml}}(\varphi) = \left(\frac{\lambda}{\pi}\right)^{\frac{1}{4}} \exp\left(-\frac{\lambda}{2}\varphi^2\right) \quad \lambda > 0. \tag{IV.39}$$

If one compares this result with the exact wavefunction from Eq. (IV.14) on page 53 it becomes evident that the error function in the normalisation constant in Eq. (IV.15) results from the restriction of the definition range for the angle to a 2π radian interval. For large positive λ the error function tends to one, which gives the Gaussian wavefunction in Eq. (IV.39). The variance of the angle within this approximation is

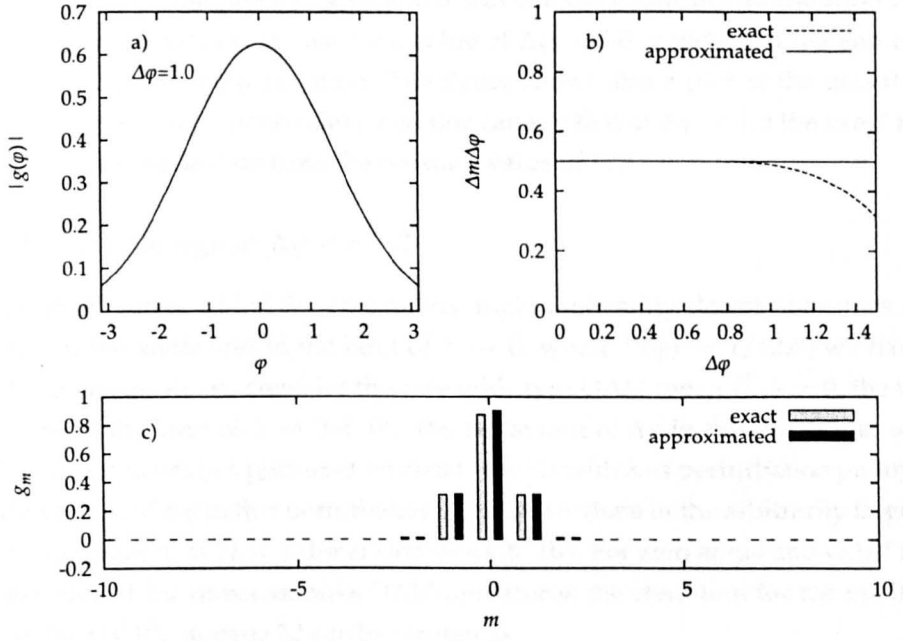

Figure IV.6 — Characteristic quantities for intelligent states for small angle uncertainties. a) shows the exact and approximated wavefunction for $\Delta\varphi = 1.0$. b) shows the uncertainty product as a function of $\Delta\varphi$. In this approximation the uncertainty product has a constant value as for the Heisenberg uncertainty relation. For $\Delta\varphi > 1$ the exact result deviates from the constant value. c) shows the distribution of probability amplitudes. Although the values for the exact and approximated results are very similar, the exact results have alternating positive and negative values away from the central peak.

calculated from Eq. (IV.18) on page 55:

$$(\Delta\varphi)^2_{\mathrm{sml}} = \frac{1}{2\lambda} \quad \lambda > 0, \tag{IV.40}$$

which corresponds to the dominant term in Eq. (IV.20) on page 55 for large λ. The OAM variance can be calculated from Eq. (IV.27) on page 57 which also holds in this approximation. As λ is always positive for $\Delta\varphi < 1$ we have

$$(\Delta m)^2_{\mathrm{sml}} = \frac{\lambda}{2} \quad \lambda > 0. \tag{IV.41}$$

The uncertainty product in this approximation has the constant value of $1/2$. This is a form which corresponds to the intelligent states for the Heisenberg uncertainty relation. It satisfies the equality (IV.3) on page 50 because $P_{\mathrm{sml}}(\pi)$ is negligible compared to one. Moreover, as the lower bound in the uncertainty relation in this approximation is constant, the intelligent states are also the minimum uncertainty product states. The distribution of probability amplitudes is calculated from the wavefunction using a Fourier transform. In this approximation the boundaries in the Fourier integral (IV.30) on page 59 are extended to infinity and the g_m follow a Gaussian distribution

$$(g_m)_{\mathrm{sml}} = \left(\frac{1}{\lambda\pi}\right)^{\frac{1}{4}} \exp\left(-\frac{1}{2\lambda}m^2\right) \quad \lambda > 0. \tag{IV.42}$$

In figure (IV.6) on the preceding page the wavefunction and the distribution of the probability amplitudes is shown for a value of $\Delta\varphi = 1.0$, which is at the end of the valid range for this approximation. The figure shows also a plot of the uncertainty product over the angle uncertainty, and one can see that at $\Delta\varphi = 1.0$ the exact result for $\Delta m \Delta\varphi$ starts to deviate from the constant value of $1/2$.

2.5.2 Flat wavefunction regime: $\Delta\varphi \approx \pi/\sqrt{3}$:

For small values of $|\lambda|$ the continuous angle probability density becomes a flat function of the angle and in the limit of $\lambda \to 0$, where $P(\varphi) = 1/(2\pi)$ we have an OAM eigenstate. As we consider the case with zero OAM mean $\langle \hat{L}_z \rangle = 0$, the OAM eigenstate in the limit of $\lambda \to 0$ is $|0\rangle$. The behaviour of $\Delta\varphi$ in this parameter region can be explained using a perturbation ansatz for $|0\rangle$ with λ as perturbation parameter. The derivation of $\Delta\varphi$ in this perturbative approach is done in the arbitrarily large, but finite state space Ψ of $2L+1$ dimension (see Ch. III). For zero angle and OAM mean and in terms of the dimensionless OAM operator $\hat{m}$ the condition for the intelligent state in Eq. (IV.10) on page 52 can be written as

$$\hat{m}|g\rangle - \mathrm{i}\lambda\hat{\varphi}|g\rangle = 0. \tag{IV.43}$$

For small λ we can use the perturbation ansatz $|g_{\mathrm{per}}\rangle = |0\rangle + \lambda|g^{(1)}\rangle$. Substituting this ansatz in the condition (IV.43) yields at first order in λ

$$\lambda\left(\hat{m}|g^{(1)}\rangle - \mathrm{i}\hat{\varphi}|0\rangle\right) = 0. \tag{IV.44}$$

Without loss of generality we can write the perturbative state $|g^{(1)}\rangle$ as a superposition of OAM eigenstates $|m\rangle, m = -L, -L+1, \ldots, -1, 1, \ldots, L$ without a contribution from $|0\rangle$:

$$|g^{(1)}\rangle = \sum_{\substack{m=-L \\ m\neq 0}}^{L} g_m^{(1)}|m\rangle. \tag{IV.45}$$

The OAM eigenstate $|0\rangle$ is a physical state, that is a state which may be approximated to any desired accuracy by the expansion $\sum_m b_m|m\rangle$, where the coefficients $b_m = 0$ for $|m| > M$ (see section III 2.3). Here, the sum includes integer values of m and the bound M is sufficiently large to guarantee the desired accuracy but always less than L. Restricting the domain of the angle operator to these physical states simplifies the expression for the angle operator and changes the summation to include an infinite number of OAM states $|m\rangle$. Using the definitions for the OAM operator and for the angle operator for physical states in (Barnett & Pegg, 1990) (see also Ch. III) we can calculate the resulting states $\hat{m}|g^{(1)}\rangle$ and $\hat{\varphi}|0\rangle$:

$$\hat{m}|g^{(1)}\rangle = \sum_{\substack{m=-L \\ m\neq 0}}^{L} m g_m^{(1)}|m\rangle = \sum_{m=-L}^{L} m g_m^{(1)}|m\rangle, \tag{IV.46}$$

$$\hat{\varphi}|0\rangle = -\mathrm{i}\sum_{m\neq 0} \frac{\exp(\mathrm{i}m\pi)}{m}|m\rangle. \tag{IV.47}$$

On substituting these results in the condition for the intelligent states at first order in λ from Eq. (IV.44) on the previous page we find an equation to determine the $g_m^{(1)}$:

$$\lambda \sum_{m\neq 0} \left(m g_m^{(1)} - \frac{\exp(\mathrm{i}m\pi)}{m} \right) |m\rangle = 0. \tag{IV.48}$$

For the linearly independent basis states $|m\rangle$ the coefficients in Eq. (IV.48) have to vanish for the sum to result in zero. This requires $g_m^{(1)} = -(-1)^m/m^2$. For small λ the intelligent state $|g_{\mathrm{per}}\rangle$ may thus be written as

$$|g_{\mathrm{per}}\rangle = N_{\mathrm{per}} \left(|0\rangle - \lambda \sum_{m\neq 0} \frac{(-1)^m}{m^2} |m\rangle \right), \tag{IV.49}$$

where N_{per} is the normalisation constant in the perturbative approach. The OAM amplitudes of the state $|g_{\mathrm{per}}\rangle$ are given by $(g_0)_{\mathrm{per}} = N_{\mathrm{per}}$ and $(g_m)_{\mathrm{per}} = \lambda g_m^{(1)} N_{\mathrm{per}}$. The angle representation of this state is obtained by projection onto an angle eigenstate. We hereby calculate a physical quantity, the angle probability amplitude, and we can now allow L to tend to infinity, which yields the continuous wavefunction of the intelligent state for small λ:

$$\begin{aligned} \langle \varphi | g_{\mathrm{per}} \rangle = g(\varphi)_{\mathrm{per}} &= \frac{N_{\mathrm{per}}}{\sqrt{2\pi}} \left(1 - \lambda \sum_{m\neq 0} \frac{(-1)^m}{m^2} \exp(\mathrm{i}m\varphi) \right), \\ &= \frac{N_{\mathrm{per}}}{\sqrt{2\pi}} \left[1 + \lambda \left(\frac{\pi^2}{6} - \frac{\varphi^2}{2} \right) \right]. \end{aligned} \tag{IV.50}$$

The sum in the last equation has been evaluated using the contour integration method (Stephenson & Radmore, 1993) (see App. A on page 166), which can also be used to calculate the normalisation constant from the requirement that the OAM probabilities $|(g_m)_{\mathrm{per}}|^2$ sum to unity:

$$\sum_m |(g_m)_{\mathrm{per}}|^2 = 1 = \left(1 + \lambda^2 \sum_{m\neq 0} \frac{1}{m^4} \right) N_{\mathrm{per}}^2 = \left(1 + \lambda^2 \frac{\pi^4}{45} \right) N_{\mathrm{per}}^2. \tag{IV.51}$$

The angle variance for the perturbative state can be calculated at first order from the continuous wavefunction in Eq. (IV.50):

$$(\Delta\varphi)_{\mathrm{per}}^2 = \langle \hat{\varphi}^2 \rangle = N_{\mathrm{per}}^2 \left(\pi^2 \frac{1}{3} - \lambda \pi^4 \frac{4}{45} + \mathrm{O}(\lambda^2) \right), \tag{IV.52}$$

where $\mathrm{O}(\lambda^2)$ indicates terms of order λ^2 which are negligible for $\lambda \to 0$. The angle uncertainty is given by the square root of the variance to first order in λ and for small λ we can approximate N_{per} and the square root. This yields at first order in λ the angle uncertainty

$$(\Delta\varphi)_{\mathrm{per}} \approx \frac{\pi}{\sqrt{3}} \left(1 - \frac{2}{15} \lambda \pi^2 \right) \quad \text{for } \lambda \to 0. \tag{IV.53}$$

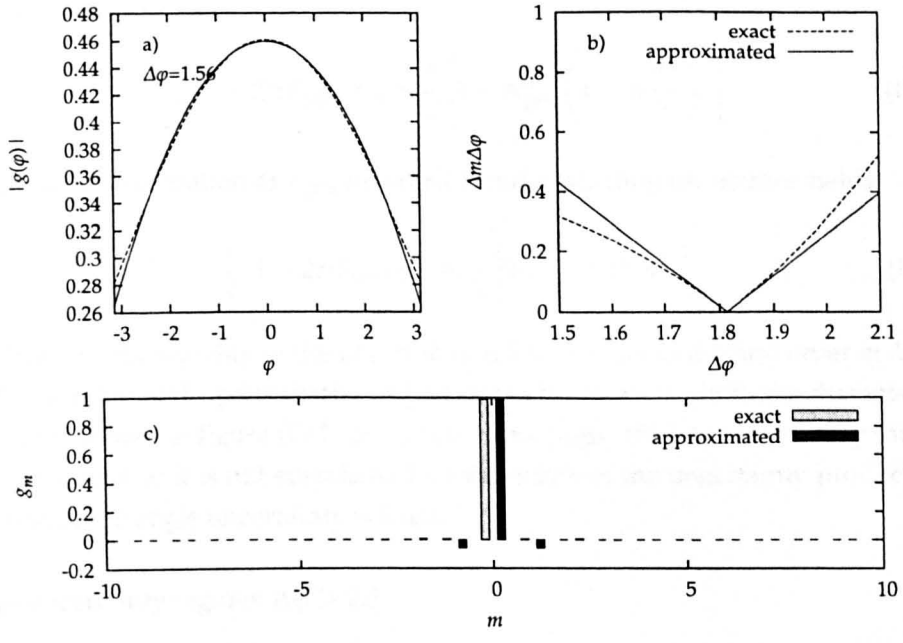

Figure IV.7 — Characteristic quantities for intelligent states in the perturbation approach. a) shows the exact and approximated wavefunction for $\Delta\varphi = 1.56$. b) shows the uncertainty product as a function of $\Delta\varphi$. The approximated graphs are linear as all quantities have been calculated to the first order in λ. c) shows the distribution of probability amplitudes.

For $\lambda \to 0$ this perturbative result is consistent with $\Delta\varphi = \pi/\sqrt{3}$ for the OAM eigenstate.

The OAM uncertainty cannot be calculated from the relation (IV.27) for wavefunctions in form of Gaussians or truncated Gaussians. Instead, we can calculate $(\Delta m)_{\text{per}}$ directly from the OAM decomposition in Eq. (IV.49) on the preceding page. With N_{per} determined by Eq. (IV.51) on the previous page the $(g_m)_{\text{per}}$ are appropriately normalised. The OAM variance is thus given by

$$(\Delta m)^2_{\text{per}} = \sum_{m\neq 0} m^2 |(g_m)_{\text{per}}|^2 = N^2_{\text{per}}\lambda^2 \sum_{m\neq 0} \frac{1}{m^2} = N^2_{\text{per}}\lambda^2 \frac{\pi^2}{3}. \tag{IV.54}$$

Approximating N_{per} for small λ yields the angular momentum uncertainty in the perturbation approach

$$(\Delta m)_{\text{per}} \approx |\lambda| \frac{\pi}{\sqrt{3}} \quad \text{for } \lambda \to 0. \tag{IV.55}$$

Here, we have to take the modulus of λ to ascertain a positive $(\Delta m)_{\text{per}}$. To the first order in λ the uncertainty product in the perturbation approach is approximately given by

$$(\Delta m)_{\text{per}}(\Delta\varphi)_{\text{per}} \approx |\lambda| \frac{\pi^2}{3} \quad \text{for } \lambda \to 0. \tag{IV.56}$$

On substituting the expression for the perturbative wavefunction in Eq. (IV.50) on the preceding page for $P_{\text{per}}(\pi) = |g(\pi)_{\text{per}}|^2$, the right hand side of the uncertainty

relation becomes

$$\frac{1}{2}\left|1-2\pi P_{\mathrm{per}}(\pi)\right| = \frac{1}{2}\left|1-N_{\mathrm{per}}^2\left(1-\lambda\frac{\pi^2}{3}\right)^2\right|. \tag{IV.57}$$

Using the approximation of N_{per} for small λ and evaluating the square yields

$$\frac{1}{2}\left|1-2\pi P_{\mathrm{per}}(\pi)\right| \approx \frac{1}{2}\left|2\lambda\frac{\pi^2}{3}+\mathrm{O}(\lambda^3)\right| \tag{IV.58}$$

which shows the equality in the uncertainty relation holds to the first order in λ.

The validity of the perturbative approach can be estimated from the characteristic quantities shown in figure (IV.7) on the previous page. We have taken only the first order in λ and so it is not surprising that the graph of the uncertainty product as a function of the angle uncertainty is linear.

2.5.3 Large uncertainty regime: $\Delta\varphi > 2.5$

The angle and angular momentum uncertainty are determined by the normalisation constant [see Eq. (IV.18) on page 55 and (IV.27) on page 57]. In the negative φ range and within the integral in Eq. (IV.15) on page 53, the wavefunction can be approximated with an exponentially decreasing function. In the limit of large negative λ this approach provides simple approximate expressions, which explain the behaviour of the uncertainties for large angle uncertainties. However, it is worth stressing that we do not derive an approximate expression for the wavefunction in this approximation directly; we only approximate the wavefunction within an integral.

The normalisation integral (IV.15) on page 53 can be rewritten as

$$\frac{1}{N_g^2} = 2\exp(|\lambda|\pi^2)\int_{-\pi}^{0}\exp[|\lambda|(\varphi+\pi)(\varphi-\pi)]\mathrm{d}\varphi \quad \lambda<0. \tag{IV.59}$$

For large $|\lambda|$ only a small region around $-\pi$ will contribute significantly to the integral, and we can therefore approximate the factor $(\varphi-\pi)$ in the exponential with -2π and extend the upper integration boundary to infinity. This results in the normalisation constant N_{exp}:

$$\frac{1}{N_{\mathrm{exp}}^2} = 2\exp(|\lambda|\pi^2)\int_{-\pi}^{\infty}\exp(-2\pi|\lambda|(\varphi+\pi))\mathrm{d}\varphi = \frac{1}{|\lambda|\pi}\exp(|\lambda|\pi^2) \quad \lambda<0. \tag{IV.60}$$

Using this normalisation constant to calculate the probability density in this approximation yields

$$P_{\mathrm{per}}(\pi) = |\lambda|\pi \quad \lambda<0. \tag{IV.61}$$

We can thus give an approximation of the right hand side of the uncertainty relation (IV.3) in the limit of large $|\lambda|$. This approximation is valid for $\Delta\varphi > 2.5$ which corresponds to $\lambda < -1/3$. In this range $P(\pi) > 1$ and the lower bound in the uncertainty relation can be written as

$$\frac{1}{2}\left|1-2\pi P_{\mathrm{exp}}(\pi)\right| = |\lambda|\pi^2-\frac{1}{2} \quad \lambda<0. \tag{IV.62}$$

To calculate the angle uncertainty we can use the relation (IV.19) on page 55, although in this regime we do not have an approximate expression for the wavefunction, for which the equality in Eq. (IV.19) would hold. The argument is that our approach allows us to approximate the normalisation constant and its functional dependence on λ for large $\Delta\varphi$. In this region we can thus use the derivative of the normalisation constant N_{exp} to calculate the square of the angle uncertainty:

$$(\Delta\varphi)^2_{\text{exp}} = N^2_{\text{exp}} \frac{\mathrm{d}}{\mathrm{d}|\lambda|} \frac{\exp(|\lambda|\pi^2)}{|\lambda|\pi^2} = \pi^2 - \frac{1}{|\lambda|}. \tag{IV.63}$$

For the angular momentum uncertainty we can use the relation in Eq. (IV.27) on page 57

$$(\Delta m)^2_{\text{exp}} = |\lambda|^2(\Delta\varphi)^2_{\text{exp}} = |\lambda|^2\pi^2 - |\lambda|. \tag{IV.64}$$

As a consequence of Eq. (IV.27) the uncertainty product in this approximation is given by $|\lambda|(\Delta\varphi)^2_{\text{exp}}$. On substituting Eq. (IV.63) into Eq. (??) on page ?? we find for the uncertainty product"

$$(\Delta m)^2_{\text{exp}}(\Delta\varphi)^2_{\text{exp}} = |\lambda|\pi^2 - 1. \tag{IV.65}$$

If we compare this uncertainty product with the lower bound in the uncertainty relation (IV.62) on the preceding page, we find that the two expressions are only approximately equal

$$|\lambda|\pi^2 - 1 \simeq |\lambda|\pi^2 - \frac{1}{2} \quad \lambda < 0. \tag{IV.66}$$

This approximation is valid for large $\Delta\varphi$ which corresponds to $\lambda \longrightarrow -\infty$. Clearly, for sufficiently large values of $|\lambda|$ the discrepancy becomes negligible.

A similar approximation as for the normalisation integral can be used in the Fourier integral to calculate the probability amplitudes. This leads to an integrand in form of an exponential and we expect a distribution of probability amplitudes in form of a Lorentzian. Owing to the symmetry of the wavefunction we can specialise the Fourier transform to the cosine form

$$g_m = \frac{2N_{\text{int}}}{\sqrt{2\pi}} \int_{-\pi}^{0} \exp(\frac{|\lambda|}{2}\varphi^2)\cos(m\varphi)\mathrm{d}\varphi. \tag{IV.67}$$

In essence the integrand can be approximated in the same way as for the wavefunction approximation. The Fourier integrand can be written as $\exp[|\lambda|(\varphi+\pi)(\varphi-\pi)/2]\exp(|\lambda|\pi^2/2)$. Within the integration interval $-\pi \leq \varphi < 0$ and for large $|\lambda|$ this integrand is only significantly different from zero around the lower boundary $-\pi$. This allows us to approximate the integrand with $\exp[|\lambda|(\varphi+\pi)(-\pi)]\exp(|\lambda|\pi^2/2)$ and to extend the integration interval in Eq. (IV.67) to infinity, which results in an angular momentum distribution in the shape of a Lorentzian:

$$\begin{aligned}(g_m)_{\text{Lor}} &= \frac{N_{\text{Lor}}}{\sqrt{2\pi}} \exp(\frac{|\lambda|}{2}\pi^2) \int_{-\pi}^{\infty} \exp(-\pi|\lambda|(\varphi+\pi))\cos(m\varphi)\mathrm{d}\varphi, \\ &= \frac{N_{\text{Lor}}}{\sqrt{2\pi}}(-1)^m \exp(\frac{|\lambda|}{2}\pi^2)\left(\frac{2|\lambda|\pi}{|\lambda|^2\pi^2 + m^2}\right).\end{aligned} \tag{IV.68}$$

Here, wave are using the index Lor to indicate the Lorentzian distribution of probability amplitudes. The distinction between the approximation for the normalisation

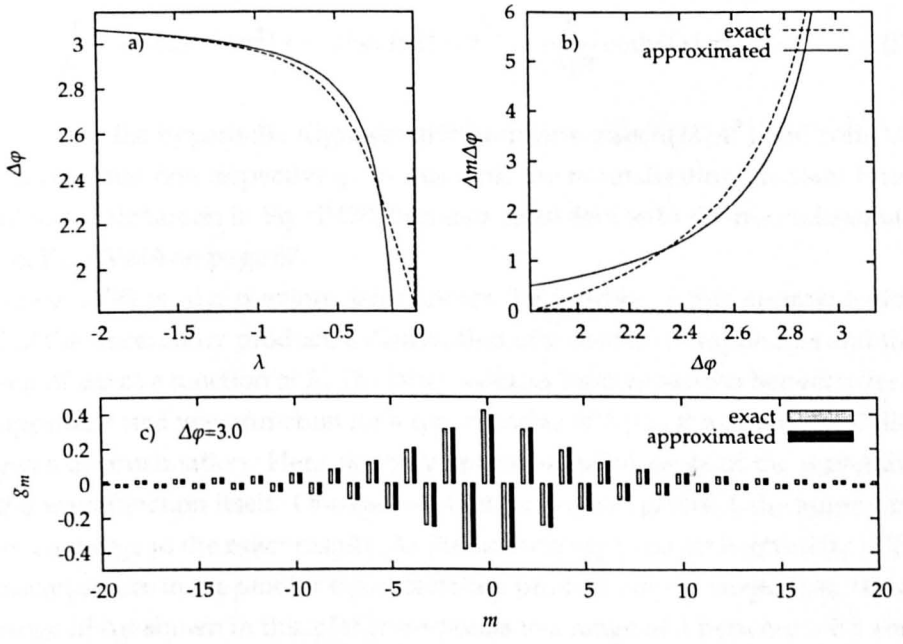

Figure IV.8 — For large $\Delta\varphi$ characteristic quantities may be approximated by treating the wavefunction as an exponential in the integrand of the normalisation integral and in the Fourier integral to calculate the OAM probability amplitudes. a) shows the approximation of $\Delta\varphi$ as a function of λ. One can see that the approximate value approaches the exact results for large negative λ. b) shows the uncertainty product as a function of $\Delta\varphi$. c) If applied to the Fourier integral this analysis yields an explanation for the distribution of probability amplitudes g_m which is approximately Lorentzian for large negative λ. The graph compares the exact results with a Lorentzian distribution calculated in this approximation for $\Delta\varphi = 3.0$.

integral in Eq. (IV.60) on page 67 and the Fourier integral in Eq. (IV.68) on page 68 becomes important for the normalisation constants, which differ from each other in the two approximations. The condition on the OAM probabilities $|(g_m)_{\mathrm{Lor}}|^2$ to sum to one can be used to determine the normalisation constant N_{Lor} consistently in this approximation. From Eq. (IV.68) on page 68 we have

$$\sum_{m=-\infty}^{\infty} |c_m|^2 = 1 = 2N_{\mathrm{Lor}}^2 |\lambda|^2 \pi \exp(|\lambda|\pi^2) \sum_{m=-\infty}^{\infty} \left(|\lambda|^2\pi^2 + m^2\right)^{-2}. \tag{IV.69}$$

The summation can be executed using the contour integration method (Stephenson & Radmore, 1993) (see App. A), which results in the following expression for N_{Lor}:

$$\frac{1}{N_{\mathrm{Lor}}^2} = \exp(|\lambda|\pi^2)\left(\pi\,\mathrm{cosech}^2(|\lambda|\pi^2) + \frac{1}{|\lambda|\pi}\coth(|\lambda|\pi^2)\right). \tag{IV.70}$$

For $\lambda \to -\infty$ the hyperbolic trigonometric functions $\mathrm{cosech}(|\lambda|\pi^2)$ and $\coth(|\lambda|\pi^2)$ tend to zero and one respectively. In this limit the normalisation constant from the Lorentzian distribution in Eq. (IV.70) becomes consistent with the normalisation constant in Eq. (IV.60) on page 67.

Figure (IV.8) on the previous page shows the validity of this approximation on hand of the uncertainty product, a distribution of probability amplitudes and the behaviour of $\Delta\varphi$ as a function of λ. The latter replaces the comparison between the exact and approximated wavefunction for a typical value of $\Delta\varphi$ in the range of validity for the given approximation. Here, we only approximate integrals of the wavefunction not the wavefunction itself. One can see that for large negative λ the approximated values converge to the exact results. As the uncertainty product is given by $|\lambda|(\Delta\varphi)^2$ the discrepancies in the plot for the uncertainty product may be surprising. However, the range of $\Delta\varphi$ shown in this plot corresponds to a range of λ between -0.5 and -1. In this range the approximated values differ greatly. Also, every discrepancy in $\Delta\varphi$ is multiplied in the plot of the uncertainty product, as both axes depend on $\Delta\varphi$: the abscissa linearly and the ordinate quadratically.

3 CMUP states

The constrained minimum uncertainty product (CMUP) states minimise the uncertainty product for the additional constraint of a given uncertainty in the angle or in the orbital angular momentum (OAM). The form of the CMUP states is the same, whether the additional constraint is a given uncertainty in the angle or in the OAM. This is because the defining equation is derived using a variation of the uncertainty product $\Delta L_z \Delta\varphi$ and the given uncertainty in either the angle or the OAM is taken into account by Lagrange multipliers. Whether an uncertainty in the angle is given or an uncertainty in the OAM, we have the same linear combination of variances with undetermined multipliers in both cases. In this section we first present the variational approach leading to the eigenvalue equation which defines the CMUP states. We then present the solutions for small and large angle uncertainties.

3.1 Variation method

Using variational methods to determine minimum uncertainty product states has first been introduced by Jackiw (1968). Here, this method is applied for finding states which minimise the uncertainty product $\Delta A \Delta B$. Seeking states which give a minimum in the uncertainty product for a given uncertainty in the angle or for given uncertainty in the OAM is equivalent to finding a minimum for the uncertainty that is not given as a constraint. In the derivation presented in this section we are minimising the variance and not the uncertainty. But as the uncertainty is defined as the square root of the variance, uncertainty and variance are related to each other by a monotonic function and therefore minimising the variance is equivalent to minimising the uncertainty. In the following we are going to use $(\Delta A)^2$ for the variance which is to be minimised and $(\Delta B)^2$ for the variance given as a constrained in the variation. We thus require that a CMUP state $|f\rangle$ minimises the free variance $(\Delta A)^2$, but with the constraint of keeping the given variance $(\Delta B)^2$ constant and obeying the normalisation condition $\langle f|f\rangle = 1$.

The variables in the variation are $\langle f|$ and $|f\rangle$ and they are treated as independent in the variation of $(\Delta A)^2$. The additional requirements can be expressed formally as holonomic constraints, also depending on $\langle f|$ and $|f\rangle$. An extremum for $(\Delta A)^2$ under these constraints will be obtained if the variation of $(\Delta A)^2$ and the holonomic constraints is stationary. Introducing Lagrange multipliers μ and ν this condition can be expressed as

$$\delta(\Delta A)^2 = \mu\delta(\Delta B)^2 + \nu\delta\langle f|f\rangle. \tag{IV.71}$$

On multiplying the whole equation with a third undetermined multiplier the stationary condition can be written as a linear combination of variations:

$$\lambda'\delta(\Delta A)^2 + \mu'\delta(\Delta B)^2 + \nu'\delta\langle f|f\rangle = 0. \tag{IV.72}$$

The same linear combination of variations with undetermined Lagrange multipliers would be obtained, if we interchanged the roles of $(\Delta A)^2$ and $(\Delta B)^2$. This explains why the defining equation for the CMUP state is the same whether the constraint is a given variance in the angle or in the OAM. As it is not necessary to distinguish between the variance given as a constraint and the variance which is to be minimised we replace the variances $(\Delta A)^2$ and $(\Delta B)^2$ with the variances for angle and OAM in no particular order.

We have seen for the intelligent states that the operator $\exp(ik\hat{\varphi}_{\theta_0})$ shifts the angular spectrum uniformly by k without changing $(\Delta L_z)^2$ or $(\Delta\varphi^2)$. We can therefore choose the OAM mean to be zero. This holds, whether $(\Delta L_z)^2$ is given as constraint, or whether it is to be minimised, and we can replace $(\Delta L_z)^2$ with $\langle \hat{L}_z^2\rangle$ in the variation. In App. IV.D we show that for the angle variance minimising $(\Delta\varphi)^2$ is equivalent to minimising $\langle\hat{\varphi}^2\rangle$. We can therefore rewrite Eq. (IV.72) in terms of $\langle\hat{L}_z^2\rangle$ and $\langle\hat{\varphi}^2\rangle$:

$$\lambda'\delta\langle f|\hat{L}_z^2|f\rangle + \mu'\delta\langle f|\hat{\varphi}^2|f\rangle + \nu'\delta\langle f|\hat{I}|f\rangle = 0. \tag{IV.73}$$

It has been shown by Pegg et al. (2005) that for the CMUP states the coefficients f_m in

the OAM decomposition

$$|f\rangle = \sum_m f_m |m\rangle \quad f(\varphi) = \frac{1}{\sqrt{2\pi}} \sum_m f_m \exp(im\varphi) \tag{IV.74}$$

can be assumed to be real, regardless whether the given constraint is an uncertainty in the angle or in the OAM. The variation for the variance of the OAM is therefore given by (Summy & Pegg, 1990):

$$\begin{aligned} \delta\langle f|\hat{L}_z^2|f\rangle &= (\delta\langle f|)\hat{L}_z^2|f\rangle + \langle f|\hat{L}_z^2(\delta|f\rangle), \\ &= \sum_m \delta f_m \left(\langle m|\hat{L}_z^2|f\rangle + \langle f|\hat{L}_z^2|m\rangle\right) = 2(\delta\langle f|)\hat{L}_z^2|f\rangle \end{aligned} \tag{IV.75}$$

The same reasoning can be applied to the variation $\delta\langle f|f\rangle$ and, by using the wavefunction in Eq. (IV.74), to the variation of the angle variance $\delta\langle f|\hat{\varphi}^2|f\rangle$. The linear combination of variations in Eq. (IV.73) on the preceding page may thus be turned into

$$2(\delta\langle f|)(\lambda'\hat{L}_z^2 + \mu'\hat{\varphi}^2 + \nu'\hat{I})|f\rangle = 0. \tag{IV.76}$$

For $|f\rangle$ to minimise the variance in the angle or OAM under the constraints of normalisation and keeping the other variance constant, this equation has to vanish for all variations $(\delta\langle f|)$. This gives a condition on the linear combination of operators applied to $|f\rangle$ and we find as eigenvalue equation for the CMUP states

$$\hat{L}_z^2|f\rangle = \left(-\nu\hat{\varphi}^2 + \mu\hat{I}\right)|f\rangle \tag{IV.77}$$

Here, we have divided the whole expression by λ' and we have chosen the sign of the undetermined multipliers μ, ν according to the expressions in (Pegg et al., 2005; Götte et al., 2006*b*). The identification of the OAM operator $\hat{L}_z$ as derivative with respect to φ sets additional requirements on the wavefunction representing CMUP states (Pegg et al., 2005; Götte et al., 2005). The wavefunction in the angle representation $f(\varphi) = \langle\varphi|f\rangle$ has to be an element of C^1, which is the set of continuously differentiable functions. The question of differentiability is of particular importance at the boundaries of the 2π radian interval, on which the angle wavefunction is defined. Whereas intelligent states are continuous, they do not have a continuous first derivative at $\varphi = \pm\pi$. CMUP states, however, need to have a continuous first derivative at the boundaries, because representing $\hat{L}_z^2$ by the differential operator $-\hbar(\partial^2/\partial\varphi^2)$ has to be well defined. Together with the periodicity of the wavefunction this requires $\partial f/\partial\varphi = 0$ at $\varphi = \pm 0$. In the following we are using the dimensionless OAM operator $\hat{m} = \hat{L}_z/\hbar$ such that the eigenvalue equation (IV.77) may be turned into a differential equation for the CMUP wavefunction $f(\varphi)$:

$$\frac{\partial^2}{\partial\varphi^2} f(\varphi) = \left(\nu\varphi^2 - \mu\right) f(\varphi). \tag{IV.78}$$

From Eq. (IV.78) follows that the curvature at $\varphi = 0$ is determined by μ. Owing to the undetermined multipliers, however, the curvature is not constant over the whole 2π radian interval. To satisfy the continuous boundary condition $\partial f/\partial\varphi = 0$ at $\varphi =$

$\pm\pi$ the wavefunction has to have a turning point with $\partial^2 f/\partial\varphi^2 = 0$ and thus the Lagrange multipliers are either both positive or both negative. Whereas the case of both multipliers being positive leads to the small-uncertainty CMUP states with a peak in the middle of 2π radian interval, the case of both multipliers being negative gives rise to the large-uncertainty CMUP states with peaks on either side of the angle range.

3.2 Wavefunction of CMUP states

A number of approximate solutions of Eq. (IV.78) on the previous page for special cases and the exact solution in the small uncertainty regime has been discussed by Pegg et al. (2005). Here, we are presenting the exact solution for the CMUP states for small and large angle uncertainties. The solution is given in terms of a Gaussian modulated by a confluent hypergeometric function. For the large uncertainty case the arguments of the Gaussian and the hypergeometric function become complex. We therefore present a power series solution in the large-uncertainty case to approximate the wavefunction in the large uncertainty regime.

3.2.1 Direct solution

The differential equation (IV.78) on the preceding page can be solved by introducing substitutions which fix the ratio of the Lagrange multiplier μ and ν:

$$\eta = \sqrt{2}\nu^{\frac{1}{4}}\varphi \quad \text{and} \quad a = -\frac{\mu}{2\sqrt{\nu}}. \tag{IV.79}$$

The parameters ν and μ take on positive and negative values and their sign distinguishes between small-uncertainty and large-uncertainty CMUP states. For the moment, we treat CMUP states in their generality and leave the signs of ν and μ undetermined. The new variables η and a may thus be imaginary. In these new variables the differential equation (IV.78) on the previous page can be rewritten as

$$\left[\frac{\partial}{\partial\eta^2} - \left(\frac{\eta^2}{4} + a\right)\right] f(\eta) = 0. \tag{IV.80}$$

By making an ansatz for the solution of this differential equation in terms of a modulated Gaussian wavefunction $f(\eta) = \exp(-\eta^2/4)M(\eta)$ we find that $M(\eta)$ has to satisfy the differential equation

$$\left[\frac{\partial^2}{\partial\eta^2} - \eta\frac{\partial}{\partial\eta} + \left(a + \frac{1}{2}\right)\right] M(\eta) = 0. \tag{IV.81}$$

On substituting $\chi = \eta^2/2$ this differential equation can be turned into Kummer's equation (Abramowitz & Stegun, 1974):

$$\chi\frac{\partial^2 M}{\partial\chi^2} + \left(\frac{1}{2} - \chi\right)\frac{\partial M}{\partial\chi} + \left(\frac{1}{2}a + \frac{1}{4}\right) M(\chi) = 0. \tag{IV.82}$$

Solutions to this equation are given by the confluent hypergeometric function $M(a/2 + 1/4, 1/2, \chi) = M(a/2 + 1/4, 1/2, \eta^2/2)$. For the small-uncertainty case, with $\nu, \mu > 0$

the wavefunction for the CMUP states in the original variables is given by (Pegg et al., 2005):

$$f(\varphi) \propto \exp\left(-\frac{\sqrt{\nu}}{2}\varphi^2\right) M\left(-\frac{1}{4}\frac{\mu}{\sqrt{\nu}} + \frac{1}{4}, \frac{1}{2}, \sqrt{\nu}\varphi^2\right), \quad \mu, \nu > 0. \tag{IV.83}$$

For the large-uncertainty case with $\nu, \mu < 0$, we write explicitly $\eta = -\sqrt{2}(-1)^{\frac{1}{4}}|\nu|^{\frac{1}{4}}\varphi$ and $a = \mathrm{i}\mu/(2\sqrt{|\nu|})$. This leads to a wavefunction with complex arguments (Götte et al., 2006*b*):

$$f(\varphi) \propto \exp\left(-\mathrm{i}\frac{\sqrt{|\nu|}}{2}\varphi^2\right) M\left(\frac{\mathrm{i}}{4}\frac{\mu}{\sqrt{|\nu|}} + \frac{1}{4}, \frac{1}{2}, \mathrm{i}\sqrt{|\nu|}\varphi^2\right), \quad \mu, \nu < 0. \tag{IV.84}$$

To determine the two multipliers μ and ν we solve the differential equation (IV.80) on the preceding page in the substituted variables for a given a. The parameter ν is then determined from the condition that the position of first extremum η_0 of $f(\eta)$ has to correspond to $\varphi = \pm\pi$, such that

$$\varphi = \pm\frac{\eta}{\eta_0}\pi \quad \text{and} \quad \nu = \begin{cases} \eta_0^4/(4\pi^4) & \nu > 0, \\ -\eta_0^4/(4\pi^4) & \nu < 0. \end{cases} \tag{IV.85}$$

From Eq. (IV.79) on the previous page follows that μ is given by $\mu = 2\sqrt{\nu}a$ for $\mu, \nu > 0$ and $\mu = -2\mathrm{i}\sqrt{|\nu|}a$ for $\mu, \nu < 0$. Series expansions of Kummer's function for numerical calculation can be found in the literature (Abramowitz & Stegun, 1974). However, by using different substitutions a modified differential equation with real parameters can be obtained. This differential equation can be solved using a series expansion.

3.2.2 Series expansion

By modifying the substitutions for the large-uncertainty case and making it independent of the sign of ν

$$\tilde{\eta} = \sqrt{2}|\nu|^{\frac{1}{4}}\varphi \quad \text{and} \quad \tilde{a} = -\frac{\mu}{2\sqrt{|\nu|}}, \tag{IV.86}$$

we change the sign of the η^2 term in the differential equation (IV.80) on the preceding page

$$\left[\frac{\partial}{\partial\tilde{\eta}^2} + \left(\frac{\tilde{\eta}^2}{4} - \tilde{a}\right)\right] f(\tilde{\eta}) = 0. \tag{IV.87}$$

The two differential equations are equivalent, and by making the ansatz $f(\tilde{\eta}) = \exp(-\mathrm{i}\tilde{\eta}^2/4)M(\tilde{\eta})$, we recover Kummer's equation on substituting $\tilde{\chi} = \mathrm{i}\tilde{\eta}^2/2$. After substituting the original variables according to Eq. (IV.86) the solution $f(\tilde{\eta})$ is identical to the wavefunction in Eq. (IV.84) obtained by the direct solution. However, in the variables given by Eq. (IV.86) the parameter a is real and we can solve the differential equation (IV.87) with a series expansion in the form

$$f(\tilde{\eta}) = \sum_{j=0}^{\infty} b_j \tilde{\eta}^j \tag{IV.88}$$

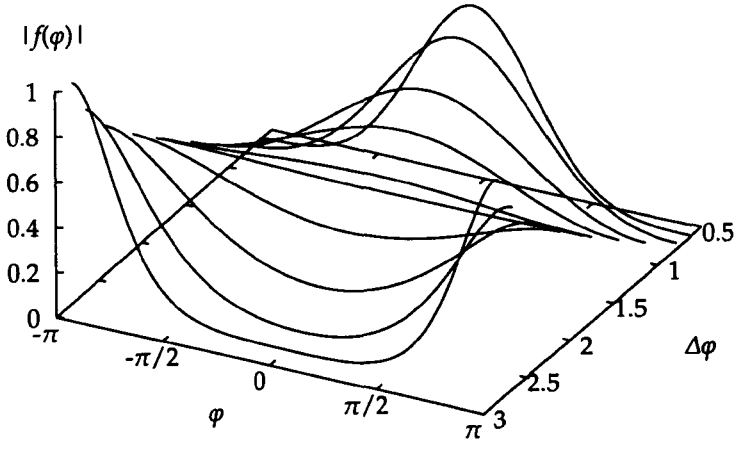

Figure IV.9 — The angle wavefunction of the CMUP states plotted for different values of $\Delta\varphi$. For $\Delta\varphi = \pi/\sqrt{3}$ the wavefunction is flat and represents an OAM eigenstate. In the limit of $\Delta\varphi \to 0$ the small-uncertainty CMUP states become sharply peaked and very similar to the intelligent states in the same limit. For larger angle uncertainties the CMUP states are different from the intelligent states [see figure (IV.1) on page 54].

On substituting the power series into the differential equation (IV.87) on the previous page we find the recurrence relation for the coefficients b_j

$$b_{j+2} = \frac{\tilde{a}b_j - \frac{1}{4}b_{j-2}}{(j+2)(j+1)} \quad \text{and} \quad b_2 = \frac{\tilde{a}}{2}b_0. \tag{IV.89}$$

The coefficient b_0 is determined by the normalisation of the wavefunction. It is difficult to anticipate the convergence of the power series over a large range of values for $\tilde{a}$. But within the numerical implementation of the power series the algorithm tests if the contribution of an additional order falls below a set value relative to the power series up to this order. Should that be the case, the power series is terminated at this point, otherwise the next order contribution is calculated. With the help of the power series we are able to determine the position of the first maximum numerically, and therefore the values of μ and ν.

A plot of the wavefunction for CMUP states for the small and the large uncertainty case is shown in figure (IV.9).

3.3 Angle and angular momentum uncertainty

Using Eq. (IV.79) on page 73 the angle variance $\langle\hat{\varphi}^2\rangle$ is given by $\langle\hat{\varphi}^2\rangle = (\pi^2/\eta_0^2)\langle\hat{\eta}^2\rangle$ for values of η in $[-\eta_0, \eta_0)$. The variance of the angular momentum operator is given by Eq. (IV.78) on page 72 in terms of μ, ν and $\langle\hat{\varphi}^2\rangle$, which results in the following

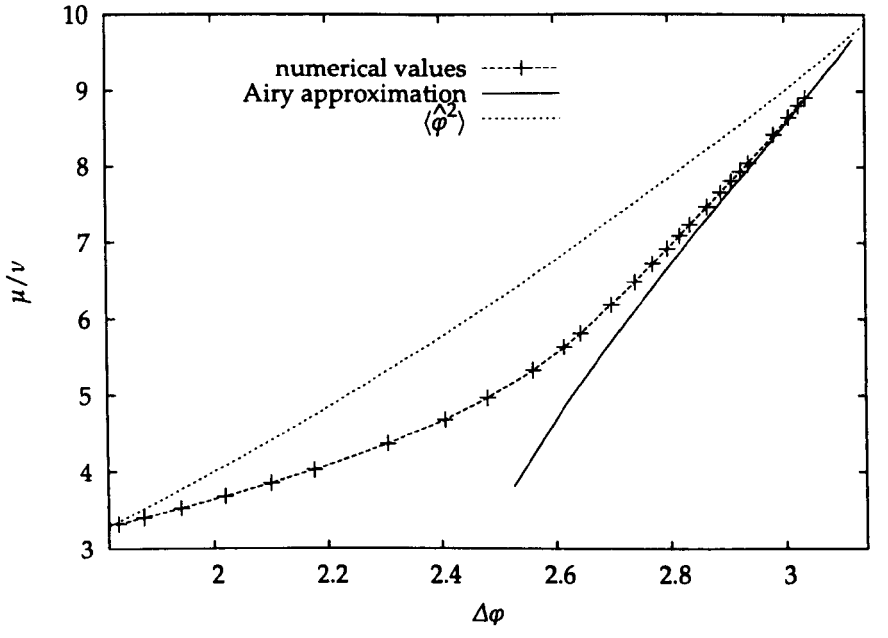

Figure IV.10 — The ratio of the two Lagrange multipliers μ and ν determines the limiting behaviour of the uncertainty product (see Eq. (IV.90a)). For $\Delta\varphi \to \pi$ the ratio μ/ν tends to π^2, but more slowly than $\langle\hat{\varphi}^2\rangle$. The uncertainty product thus tends to infinity. The plot of μ/ν in the Airy approximation shows the region of validity for this approximation.

expression for the product of the variances $\langle\hat{\varphi}^2\rangle$ and $\langle\hat{m}^2\rangle$:

$$\langle\hat{\varphi}^2\rangle\langle\hat{m}^2\rangle = \langle\hat{\varphi}^2\rangle\left(\mu - \nu\langle\hat{\varphi}^2\rangle\right), \tag{IV.90a}$$

$$= \langle\hat{\eta}^2\rangle\left(-a + \frac{1}{4}\langle\hat{\eta}^2\rangle\right). \tag{IV.90b}$$

The limiting behaviour of the uncertainty product is directly connected to the behaviour of the ratio μ/ν. For $a \to 0$, μ and ν tend to zero individually but their ratio μ/ν approaches $\pi^2/3$ (see figure (IV.10)). The variance $\langle\hat{\varphi}^2\rangle$ takes on the value of $\pi^2/3$ and the overall product of variances vanishes. For $a \to \infty$, both μ and $\nu \to -\infty$, but the ratio μ/ν approaches π^2. The variance $\langle\hat{\varphi}^2\rangle$, however, tends to its maximum π^2 faster than μ/ν such that the uncertainty product in Eq. (IV.90a) tends to infinity. The resulting behaviour of the uncertainty product as a function of $\Delta\varphi$ is given in figure (IV.11) on the next page. As in the small-uncertainty case for $\Delta\varphi < \pi/\sqrt{3}$ the uncertainty product is smaller for the CMUP states than for the intelligent states while still obeying the uncertainty relation (IV.3) on page 50. This is possible because of the smaller probability density $P(\pi)$ at the edge of the chosen 2π radian interval. But in comparison to the small uncertainty CMUP states, the difference in the uncertainty product between intelligent states and CMUP states in the large-uncertainty regime is enhanced over the small-uncertainty regime. This goes along with a significant difference in the wavefunction for intelligent states and CMUP states for the same $\Delta\varphi$ in this region (see figure (IV.12) on page 78). In connection with figure (IV.11) on the following page it is appropriate to clarify the meaning of minimising the uncertainty product under a constraint. For CMUP states with small and large angle uncertain-

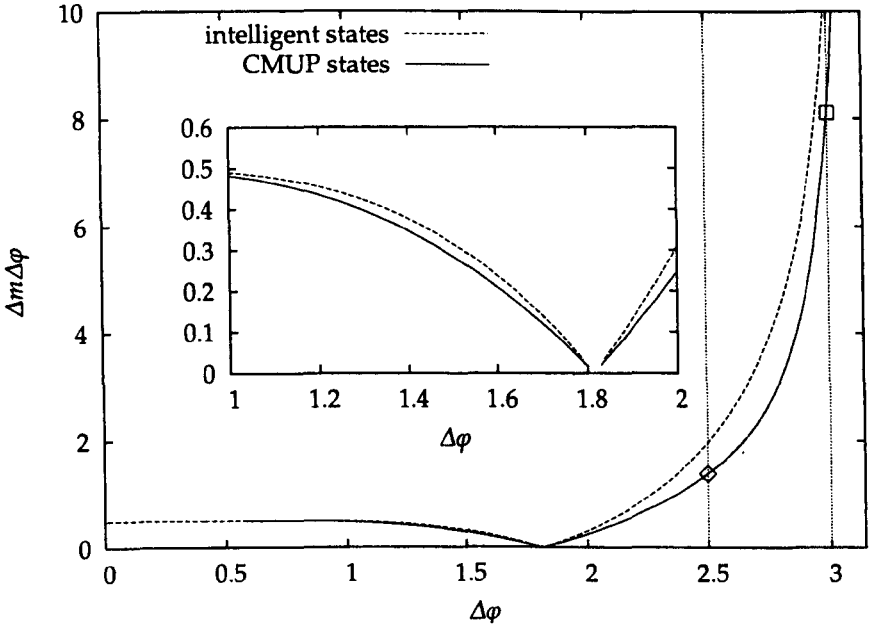

Figure IV.11 — Plot of the uncertainty product as a function of $\Delta\varphi$. The graphs of the intelligent states (Franke-Arnold et al., 2004; Götte et al., 2005) and small-uncertainty CMUP states (Pegg et al., 2005) are shown for comparison (see also plot of intelligent states in (Galindo & Pascual, 1990)). The difference in the uncertainty product between intelligent states and CMUP states is significantly enhanced in the large-uncertainty regime for $\Delta\varphi > \pi/\sqrt{3}$. For two values of $\Delta\varphi$ (marked by the dotted lines with the symbols $\Diamond$ and $\square$) the difference in the wavefunction is shown in figure (IV.12) on the next page. The inset shows an enlargement around the global minimum in the uncertainty product for $\Delta\varphi = \pi/\sqrt{3}$.

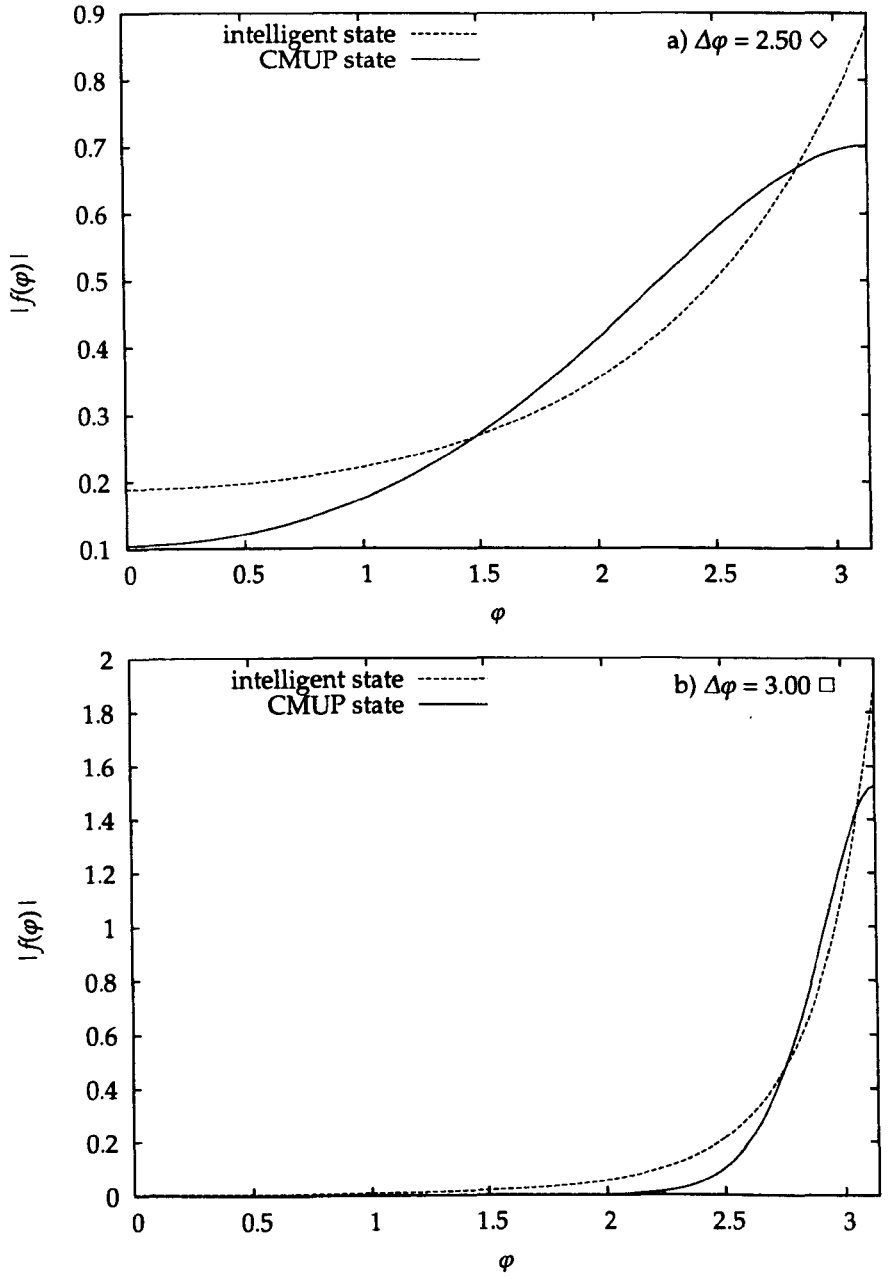

Figure IV.12 — Comparison of the wavefunction for intelligent states and CMUP states for the same $\Delta\varphi$ in the large-uncertainty regime. The difference in the uncertainty product for the two values a) $\Delta\varphi = 2.5\,\Diamond$ and b) $\Delta\varphi = 3.0\,\square$ can be seen in figure (IV.11) on the preceding page. The position of the two values for a) and b) is marked in figure (IV.11) on the previous page by dotted lines with the respective symbols.

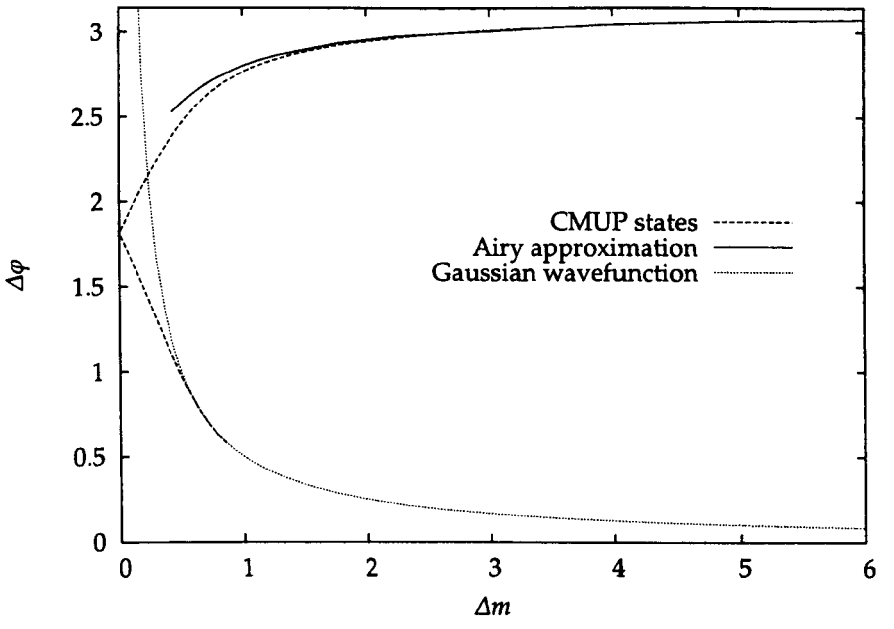

Figure IV.13 — If the uncertainty product is minimised for a given uncertainty in the OAM, two minima can be obtained. The first and smaller minimum is obtained for the small-uncertainty CMUP states while a secondary minimum is found for the large-uncertainty CMUP states. For comparison the limiting cases for these two kinds of states are shown. For $\Delta\varphi \to 0$ the small uncertainty states become Gaussians (Pegg et al., 2005), whereas the large-uncertainty states are approximatively given by Airy functions for $\Delta\varphi \to \pi$ (Götte et al., 2006*b*).

ties the OAM uncertainty can take on all positive real values. Δm is zero for the OAM eigenstates at $\Delta\varphi = \pi/\sqrt{3}$ and it approaches infinity for $\Delta\varphi \to 0$ and $\Delta\varphi \to \pi$. Minimising the uncertainty product for a given Δm yields two constrained minima. The smaller constrained minimum is obtained for the small-uncertainty CMUP states and corresponds to an angle uncertainty $\Delta\varphi < \pi/\sqrt{3}$. A secondary minimum, however, is obtained for the large-uncertainty CMUP states corresponding to $\Delta\varphi > \pi/\sqrt{3}$ (see figure (IV.13)). On the other hand minimising the uncertainty product for a given $\Delta\varphi$ results in a unique minimum. Whether this minimum is obtained for small-uncertainty or large-uncertainty CMUP states depends on the given $\Delta\varphi$. Owing to the complexity of the CMUP states we are not able to give an analytical explanation of the limiting behaviour in simple terms. Also, our method to determine the first maximum of the wavefunction numerically fails for very sharply peaked wavefunctions. In the following we therefore present an approximate expression for the wavefunction in terms of Airy functions, which allows us to calculate the variance $\Delta\varphi$ analytically.

3.4 Orbital angular momentum distribution

For completeness we present the OAM distribution for CMUP states in this section. The OAM probabilities have been calculated numerically with help of the power series discussed in section IV 3.2.2. The graph of the distributions is shown in fig-

ure (IV.14) on the following page. In contrast to the intelligent states, the OAM distribution of CMUP states has a Gaussian form in the small and large uncertainty regime.

3.5 Limiting behaviour

In this section we investigate the CMUP states for the same regions as for the intelligent states. For the intelligent states the two extreme cases for $\Delta\varphi \to 0$ and $\Delta\varphi \to \pi$ are very distinct, but for the CMUP states one might expect that the large uncertainty CMUP states are a shifted version of the small uncertainty states owing to the smooth boundary conditions. As we will see in this section this is not the case, and the CMUP states in the small and large uncertainty regime have different functional forms.

3.5.1 Small uncertainty regime: $\Delta\varphi < 1$

The uncertainty product in this region is approximately constant, and therefore intelligent states and CMUP states are identical and we can refer to the analysis in section 2.5.1. It is, however, interesting to point out that it is possible to have Gaussians as a solution to the defining differential equation (IV.78) on page 72. If we make the ansatz $f(\varphi) \propto \exp(-\sigma\varphi^2)$, we find that that the constant $\sigma \in \mathbb{R}$ has to fulfill the relation

$$-2\sigma + 4\sigma^2\varphi^2 = \nu\varphi^2 - \mu. \tag{IV.91}$$

This is the case for $\nu = 4\sigma^2$ and $\mu = 2\sigma$ or equivalently $\mu^2/\nu = 1$, and this is always possible as ν and μ are positive for the small uncertainty case. In terms of the scaled variables in Eq. (IV.79) on page 73 this gives $a = 1/2$ which corresponds to the limit of $\Delta\varphi = 0$. In this extreme case the wavefunction is therefore indeed Gaussian. As pointed out in section 2.5.1 the wavefunctions in the whole region for $\Delta\varphi < 1$ are approximatively Gaussians.

3.5.2 Large uncertainty regime: $\Delta\varphi \to \pi$

Judging from the plot of the wavefunction for the CMUP states in figure (IV.9) on page 75 the wavefunction for large uncertainty CMUP with $\Delta\varphi \to \pi$ states looks similar to the Gaussian wavefunction for $\Delta\varphi \to 0$ shifted by π. But if we make the same simple analysis as in section 3.5.1 we find that the wavefunction cannot be Gaussian. By making the ansatz $f(\varphi) \propto \exp(-\sigma(\varphi - \pi)^2), \sigma > 0$ as solution for the differential equation (IV.78) on page 72 for $\varphi > 0$ we find that σ has to fulfil the equation

$$-2\sigma + 4\sigma^2(\varphi - \pi)^2 = \nu\varphi^2 - \mu. \tag{IV.92}$$

For the large-uncertainty CMUP states $\nu, \mu < 0$ and therefore σ cannot obey the relation $4\sigma^2\varphi^2 = \nu\varphi^2$ for all φ. It is, however, possible to find an approximate expression for the large- uncertainty CMUP states in the limit $\Delta\varphi \to \pi$. To be more precise, rather than finding an approximate solution to the exact differential equation, we find an exact solution to an approximate version of the defining differential equation (IV.78) on page 72.

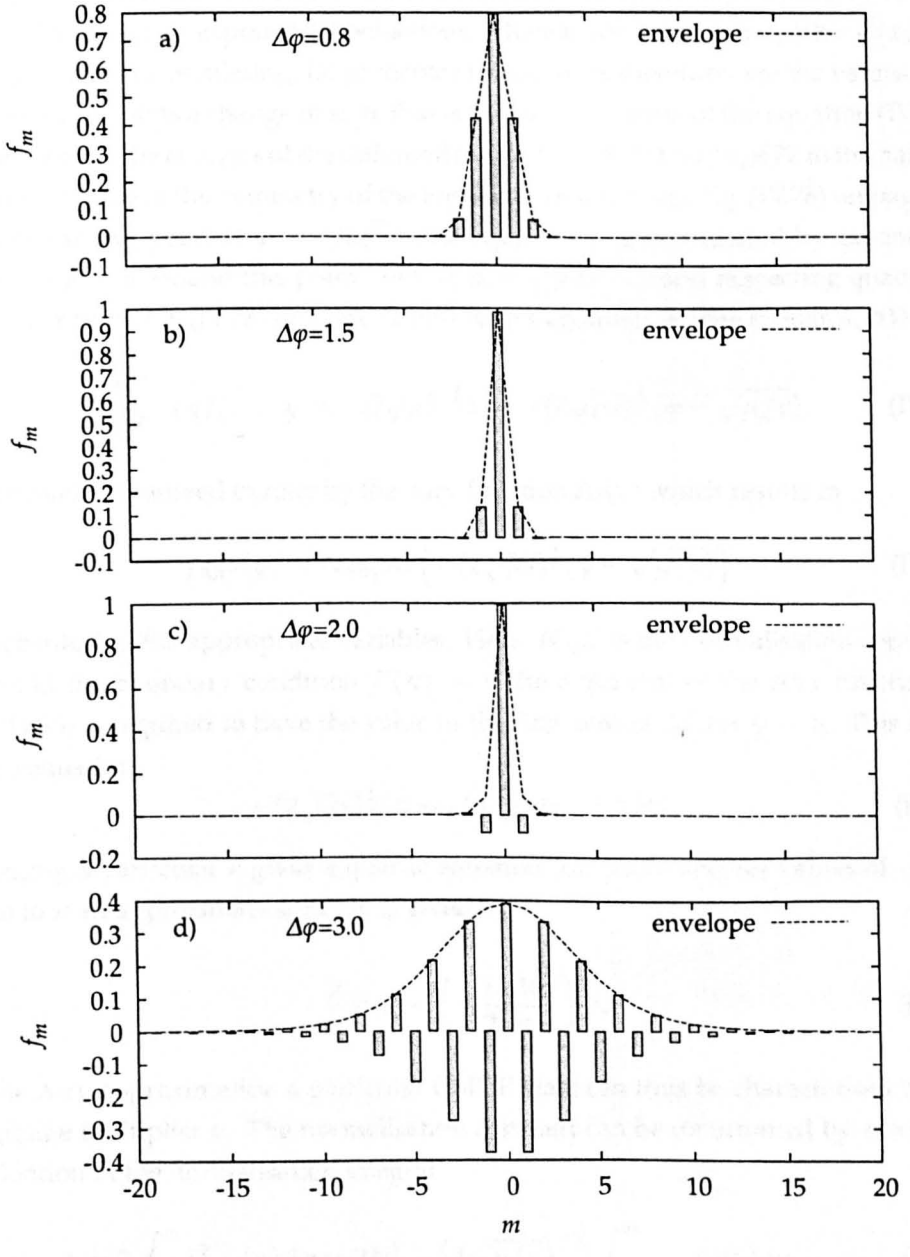

Figure IV.14 — The distribution of the OAM probabilities for the CMUP states is approximatively Gaussian for both small-uncertainty and large-uncertainty states. a) for $\Delta\varphi < 1$ the uncertainty product is constant and intelligent states and CMUP have the same Gaussian probability distribution (see figure (IV.5) on page 61). b) for $\Delta\varphi = 1.5$ the probability distribution is different to the distribution for intelligent states of the same angle uncertainty. d) for $\Delta\varphi = 3.0$ the wavefunction of the CMUP states is approximately given by the decaying tail of an Airy function.

The behaviour of the solution for a general differential equation of the form

$$\frac{\partial^2 f}{\partial x^2} = C(x)f(x) \tag{IV.93}$$

is partly determined by the sign of the function $C(x)$. Should $C(x)$ be purely positive we would expect an exponential behaviour, whereas for a purely negative $C(x)$ the solution would be oscillating. Of particular importance, therefore, are the values of x where $C(x)$ exhibits a change of sign, that is the turning points of the equation (IV.93). We can restrict the analysis of the differential equation (IV.78) on page 72 to the half interval $[0, \pi)$ due to the symmetry of the equation. In this range Eq. (IV.78) on page 72 has one turning point at $\varphi = \sqrt{\mu/\nu}$. The equation is approximated by expanding $C(x) = \nu\varphi^2 - \mu$ around this point. Setting $\varphi = \sqrt{\mu/\nu} + x$ and neglecting quadratic terms in x turns Eq. (IV.78) on page 72 into Airy's equation (Vallée & Soares, 2004):

$$\frac{\partial^2 f}{\partial y^2} = yf, \qquad y = -(2\sqrt{\mu\nu})^{\frac{1}{3}}x = -(2\sqrt{\mu\nu})^{\frac{1}{3}}(\varphi - \sqrt{\mu/\nu}). \tag{IV.94}$$

This equation is solved exactly by the Airy function $\mathrm{Ai}(y)$ which results in

$$f_{\mathrm{Airy}}(\varphi) = N_{\mathrm{Airy}}\mathrm{Ai}\left(-(2\sqrt{\mu\nu})^{\frac{1}{3}}(\varphi - \sqrt{\mu/\nu})\right) \tag{IV.95}$$

on substituting the appropriate variables. Here, N_{Airy} is the normalisation constant. To fulfill the boundary condition $f'(\pi) = 0$ the argument of the Airy function in Eq. (IV.95) is required to have the value of the first zero of Ai' for $\varphi = \pi$. This leads to the equation

$$-(2\sqrt{\mu\nu})^{\frac{1}{3}}(\pi - \sqrt{\mu/\nu}) \cong -1.0188. \tag{IV.96}$$

Choosing a particular ν gives a quartic equation for $\sqrt{\mu/\nu}$ and for values of $\sqrt{\mu/\nu}$ close to π an approximate solution is given by

$$\frac{\mu}{\nu} \approx \pi - \left(\frac{-1.0188}{2\nu\pi}\right)^{\frac{1}{3}}. \tag{IV.97}$$

In the Airy approximation a particular CMUP state can thus be characterised by the Lagrange multiplier ν. The normalisation constant can be determined by analytical evaluation of the normalisation integral

$$1 = 2\int_0^{\pi} f_{\mathrm{Airy}}^2(\varphi)\mathrm{d}\varphi \approx 2N_{\mathrm{Airy}}^2\left(2\sqrt{\mu/\nu}\right)^{-\frac{1}{3}}\int_{-1.0188}^{\infty}\mathrm{Ai}(y)\mathrm{d}y. \tag{IV.98}$$

In the last step we have extended the range of integration from $y(\varphi = 0)$ to infinity. In the region where the Airy approximation is applicable ($\Delta\varphi \to \pi$), the wavefunction decays to zero sufficiently quickly for small angles so that extending the upper bound in the integral does not significantly change the normalisation integral. Primitives of products of Airy functions can be calculated using the method of Albright (1977). This results in

$$N_{\mathrm{Airy}} = (\mu\nu)^{\frac{1}{12}}\left((1.0188)^{\frac{1}{2}}(0.5357)2^{\frac{1}{3}}\right)^{-1}, \tag{IV.99}$$

where $\mathrm{Ai}(y = -1.0188) = 0.5357$. In figure (IV.15) on the following page a compar-

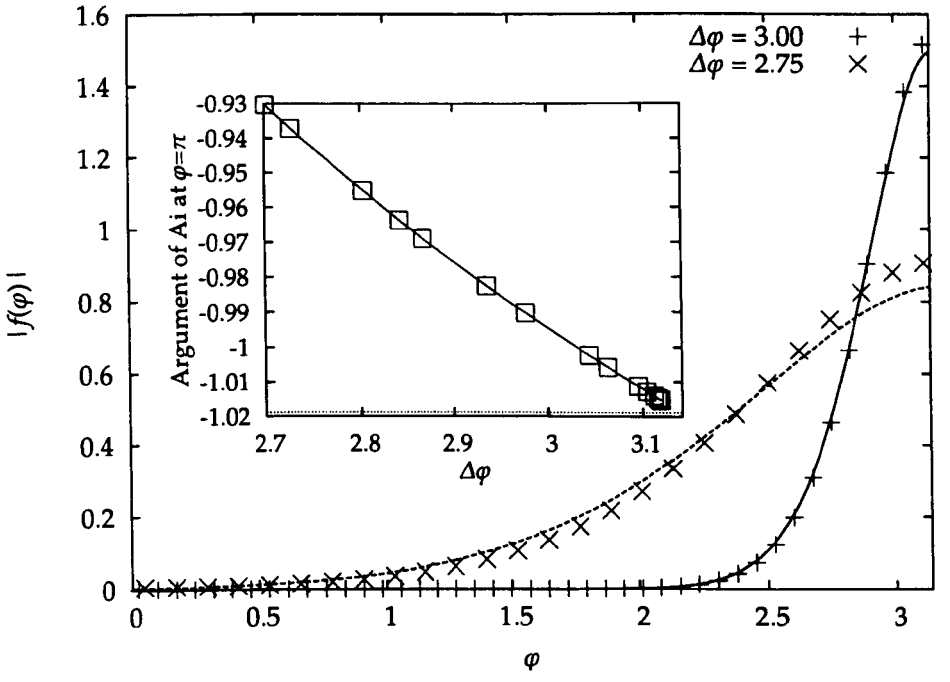

Figure IV.15 — Plot showing a comparison of the Airy approximation (continuous lines) with the numerically calculated wavefunction (individual points). For $\Delta\varphi = 3$ (+) the Airy approximation shows a good agreement with the numerical results. The inset shows the deviation of the argument of the Airy function Ai at $\varphi = \pi$ from -1.0188 (marked by the horizontal dotted line), the position of the first maximum of the Airy function Ai.

ison of the numerically calculated wavefunction and the wavefunction in the Airy approximation is shown. The approximation becomes better for values of $\Delta\varphi$ closer to π. An inset in figure (IV.15) gives the deviation of the argument of the Airy function in Eq. (IV.95) on the previous page from the exact value of $y = -1.0188$ due to the approximation in Eq. (IV.97) on the preceding page of the quartic equation (IV.96) determining μ/ν. Within the Airy approximation the integral for the angle variance can be calculated analytically using the method of Albright (1977):

$$(\Delta\varphi)^2 = \frac{\mu}{\nu} + \frac{2}{3}(1.0188)\,(2\sqrt{\mu\nu})^{-\frac{1}{3}}\,\sqrt{\mu/\nu} + \frac{1}{5}(1.0188^{-1} + 1.0188^{2})\,(2\sqrt{\mu\nu})^{-\frac{2}{3}}\,. \tag{IV.100}$$

As in the calculation of the normalisation constant in Eq. (IV.98) on the previous page the upper boundary in the integration has been extended to infinity. On multiplying Eq. (IV.100) by ν one can see in Eq. (IV.90a) on page 76 that $\nu\langle\hat{\varphi}^2\rangle$ will always be smaller than μ resulting in an unbounded uncertainty product. Within the Airy approximation the uncertainty product can be calculated for values of $\Delta\varphi$ much closer to π than in the numerical calculation. This is due to the fact that our numerical determination of the first maximum fails for large values of a. In the Airy approximation a numerical search for the first maximum is not necessary. The uncertainty product calculated in the Airy approximation is compared with the numerical results in figure (IV.16) on the following page.

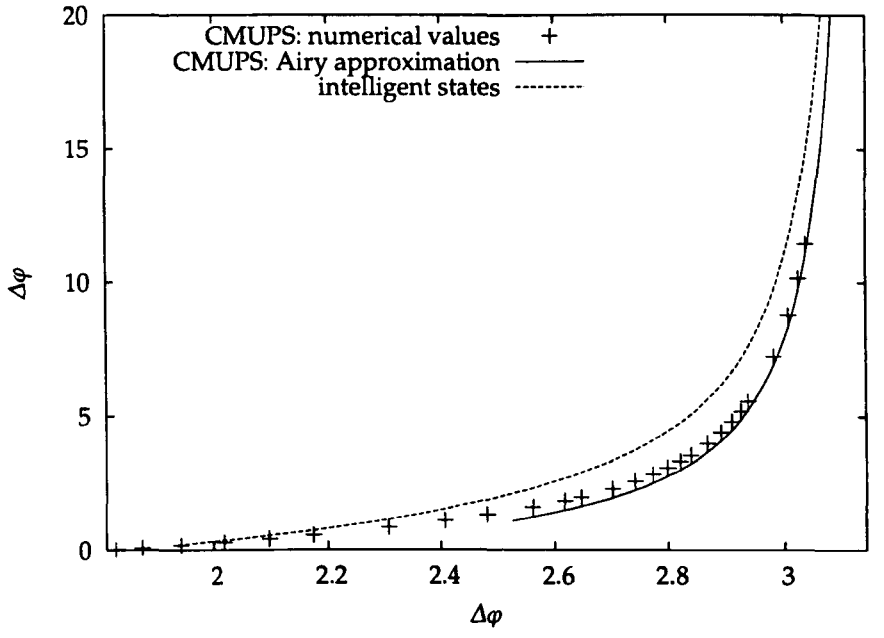

Figure IV.16 — Comparison of the uncertainty product calculated in the Airy approximation and for numerical results. In difference to figure (IV.11) on page 77, the graph shows only the large-uncertainty region ($\Delta\varphi > \pi/\sqrt{3}$) and the ordinate is extended to larger values of $\Delta\varphi\Delta m$. The Airy approximation explains the behaviour of $\Delta\varphi\Delta m$ in a region where our numerical calculation fails.

4 Conclusions

The angular uncertainty relation gives rise to a family of special states: intelligent states satisfying the equality in the uncertainty relation and states giving a minimum in the uncertainty product. The distinction arises from the state dependent lower bound for the uncertainty relation (IV.3) on page 50, which reaches the global minimum for orbital angular momentum (OAM) eigenstates. But one can also look for minimising the uncertainty product for a given uncertainty in the angle or OAM. This leads to the constrained minimum uncertainty product (CMUP) states, which have the same form whether the given uncertainty is for the angle or for the OAM. The fact that the angle is defined on a 2π radian range adds further to the complexity of the special states for the angular uncertainty relation. On a finite interval states can be normalised which could not be normalised on an infinite interval. These states are peaked at the edge of the 2π radian interval and have a larger angle uncertainty than states peaked at the centre of the allowed range of angles. CMUP states and intelligent states therefore come in two varieties, with small and large angle uncertainties. The OAM eigenstate is the dividing point and, at the same time, intelligent state and minimum uncertainty product state.

In this chapter we have derived the defining equations for the whole family of states. The solution to the differential equations (IV.12) on page 52 and (IV.78) on page 72 gives the wavefunction for intelligent states and CMUP states, and we have used these wavefunctions to calculate the uncertainties in angle or OAM and their product. In the case of the intelligent states is was possible to find analytical expressions, but for the CMUP states all the characteristics had to be calculated numerically.

In both cases, however, we have presented approximations for a particular range of values for $\Delta\varphi$ to give a better understanding and an explanation of the observed phenomena.

In the theoretical framework by Barnett & Pegg (1990) choosing a specific 2π radian range determines the angular uncertainty, the uncertainty relation and so the intelligent states and CMUP states. Our analysis of the CMUP states in the two extreme cases of $\Delta\varphi \to 0$ and $\Delta\varphi \to \pi$ shows that these wavefunctions cannot be transformed into each other simply by shifting the range of allowed angles. In the case of $\Delta\varphi \to 0$ the wavefunction is approximatively Gaussian, whereas for $\Delta\varphi \to \pi$ the wavefunction can be excellently approximated by an Airy function. In the article by Pegg et al. (2005) the possibility is discussed to distinguish between intelligent states and optical states in an experiment. Our comparison of the uncertainty product for CMUP states and intelligent states shows that the difference in the uncertainty product is enhanced in the large-uncertainty regime. This goes in hand with a significant difference in the wavefunction. Judging from these results it could be advantageous to perform the experiment with large-uncertainty intelligent and CMUP states.

Truncated Gaussians, the functional form of the wavefunction for intelligent states, will also be used in the next chapter to define angular apertures used in the derivation of an angular EPR paradox.

IV.A Uncertainty relation

The derivation of the uncertainty relation can be found in standard quantum mechanic textbooks. It is given here, to explain the derivation of Eq. (IV.4) on page 51, the starting point for the analysis of angular intelligent states. We follow the treatment by Louisell (1973).

In quantum mechanics physical observables are represented by Hermitian operators. The measurement of an observable A shows, in general, fluctuations about the average value $\langle\hat{A}\rangle$. The mean- square deviation from the average value, also called variance, is given by

$$(\Delta A)^2 = \langle\hat{A}^2\rangle - \langle\hat{A}\rangle^2. \tag{IV.101}$$

If we disregard fluctuations in the measurement originating from the measuring instruments then the variance is only zero if and only if the mean value $\langle\hat{A}\rangle$ is evaluated for an eigenstate of $\hat{A}$.

If we now have two observables A and B which do not commute, we can define two new Hermitian operators $\hat{a}$ and $\hat{b}$ by setting

$$\hat{a} = \hat{A} - \langle\hat{A}\rangle \quad \text{and} \quad \hat{b} = \hat{B} - \langle\hat{B}\rangle. \tag{IV.102}$$

As $\langle\hat{A}\rangle$ and $\langle\hat{B}\rangle$ are numbers and not operators, $\hat{a}$ and $\hat{b}$ share the same commutation relation as $\hat{A}$ and $\hat{B}$:

$$[\hat{a},\hat{b}] = [\hat{A},\hat{B}]. \tag{IV.103}$$

From Eq. (IV.102) it follows that $\langle\hat{a}\rangle = \langle\hat{b}\rangle = 0$ and so $\hat{a}$ and $\hat{b}$ have the same variances as $\hat{A}$ and $\hat{B}$

$$(\Delta a)^2 = (\Delta A)^2 = \langle\hat{a}^2\rangle \quad \text{and} \quad (\Delta b)^2 = (\Delta B)^2 = \langle\hat{b}^2\rangle. \tag{IV.104}$$

Owing to the Hermiticity of the operators $\hat{a}$ and $\hat{b}$ the mean values $\langle\hat{a}^2\rangle$ and $\langle\hat{b}^2\rangle$ may be written as $\langle\hat{a}\psi|\hat{a}\psi\rangle$ and $\langle\hat{b}\psi|\hat{b}\psi\rangle$ respectively. The product of the variances $(\Delta A)^2(\Delta B)^2$ may thus be written as $\langle\hat{a}\psi|\hat{a}\psi\rangle\langle\hat{b}\psi|\hat{b}\psi\rangle$. We may now use the Schwartz inequality which asserts that for any two state vectors $|\varphi\rangle$ and $|\chi\rangle$ the modulus square of the scalar product $|\langle\varphi|\chi\rangle|^2$ is smaller than or equal to the product $\langle\varphi|\varphi\rangle\langle\chi|\chi\rangle$. And the equality holds if and only if the two state vectors are linearly dependent. If we apply the Schwartz inequality to the product $\langle\hat{a}\psi|\hat{a}\psi\rangle\langle\hat{b}\psi|\hat{b}\psi\rangle$ and use Eq. (IV.104) on the preceding page we arrive at the inequality

$$(\Delta a)^2(\Delta b)^2 \geq |\langle\psi|\hat{a}\hat{b}|\psi\rangle|^2. \tag{IV.105}$$

The product $\hat{a}\hat{b}$ may be written in terms of the anti-commutator and commutator of these two operators:

$$\begin{aligned}\hat{a}\hat{b} &= \frac{1}{2}\left(\hat{a}\hat{b}+\hat{b}\hat{a}\right)+\frac{1}{2}\left(\hat{a}\hat{b}-\hat{b}\hat{a}\right),\\ &= \frac{1}{2}[\hat{a},\hat{b}]_+ + \frac{1}{2}[\hat{a},\hat{b}].\end{aligned} \tag{IV.106}$$

The anti-commutator of two Hermitian operators is again Hermitian, but the Hermitian conjugate of the commutator of two Hermitian operator is its own negative. Therefore multiplying the commutator $[\hat{a},\hat{b}]$ by the imaginary unit gives an overall Hermitian operator $\mathrm{i}[\hat{a},\hat{b}]$ (Bohm, 1951). The product $\hat{a}\hat{b}$ may thus be rewritten as the sum of two Hermitian operators

$$\hat{a}\hat{b} = \frac{1}{2}[\hat{a},\hat{b}]_+ - \mathrm{i}\frac{1}{2}\mathrm{i}[\hat{a},\hat{b}]. \tag{IV.107}$$

On substituting this form into Eq. (IV.105) we note that the mean value of a Hermitian operator is a real number. In Eq. (IV.105) we therefore have the modulus square of a complex number with $\langle[\hat{a},\hat{b}]_+/2\rangle$ as the real part and $\langle-\mathrm{i}[\hat{a},\hat{b}]/2\rangle$ as the imaginary part. The inequality (IV.105) may thus be rewritten in terms of the anti- commutator and commutator

$$\begin{aligned}(\Delta a)^2(\Delta b)^2 &\geq \frac{1}{4}\langle[\hat{a},\hat{b}]_+\rangle^2+\frac{1}{4}\langle-\mathrm{i}[\hat{a},\hat{b}]\rangle^2,\\ &= \frac{1}{4}\langle[\hat{a},\hat{b}]_+\rangle^2+\frac{1}{4}|\langle[\hat{a},\hat{b}]\rangle|^2.\end{aligned} \tag{IV.108}$$

Here, we have used the fact that $\langle-\mathrm{i}[\hat{a},\hat{b}]\rangle$ is a real number and we can thus identify $\langle-\mathrm{i}[\hat{a},\hat{b}]\rangle^2 = |\langle-\mathrm{i}[\hat{a},\hat{b}]\rangle|^2$. Taking the imaginary unit out of the mean value then leads to the form given in Eq. (IV.108). Using the identities in Eqs. (IV.103) to (IV.104), we may substitute the product of variances $(\Delta a)^2(\Delta b)^2$ and the commutator $[\hat{a},\hat{b}]$ in Eq. (IV.108) with $(\Delta A)^2(\Delta B)^2$ and $[\hat{A},\hat{B}]$ to arrive at Eq. (IV.5) on page 51.

IV.B Orbital angular momentum shift operator

In the derivation of the wavefunction for intelligent states and CMUP states we consider the case of zero OAM mean without loss of generality. This is because the class of unitary OAM shift operators $\exp(\mathrm{i}k\hat{\varphi}_{\theta_0})$ with $k \in \mathbb{Z}$ changes the OAM of a

state by k (Barnett & Pegg, 1990). This class of operators forms a one a parameter group with k as the parameter of OAM shift (Holevo, 2001).

In quantum mechanics OAM and angle are two incompatible observables. As such the angle operator acts as the generator of OAM (Galindo & Pascual, 1990). Within the state space Ψ of $2L+1$ dimensions [see Ch. III] the unitary OAM shift operator is constructed as

$$\hat{U}_k = \exp(\mathrm{i}k\hat{\varphi}_{\theta_0}) = \exp\left(\mathrm{i}k\sum_{n=0}^{2L}\theta_n|\theta_n\rangle\langle\theta_n|\right), \tag{IV.109}$$

where $k = -L, -L+1, \ldots, L$. The dyadic product $|\theta_n\rangle\langle\theta_n| = \hat{P}(\theta_n)$ is a projector onto the angle states $|\theta_n\rangle$ and as such idempotent, that is $\hat{P}(\theta_n)^\nu = \hat{P}(\theta_n)$ for $\nu \in \mathbb{N}$. Owing to the orthogonal angle states the product of two projectors onto different angle states results in zero: $\hat{P}(\theta_n)\hat{P}(\theta_{n'}) = \delta_{nn'}$. If we use these properties of the projector in a series expansion of the exponential we find

$$\begin{aligned}\exp(\mathrm{i}k\hat{\varphi}_{\theta_0}) = \sum_{\nu=0}^{\infty}\frac{1}{\nu!}\left(\mathrm{i}k\sum_{n=0}^{2L}\hat{P}(\theta_n)\right)^\nu &= \sum_{n=0}^{2L}\sum_{\nu=0}^{\infty}\frac{1}{\nu!}(\mathrm{i}k\theta_n)^\nu\hat{P}(\theta_n),\\ &= \sum_{n=0}^{2L}\exp(\mathrm{i}k\theta_n)\hat{P}(\theta_n).\end{aligned} \tag{IV.110}$$

Applied to an OAM eigenstate $|m\rangle$ the OAM shift operator $\exp(\mathrm{i}k\hat{\varphi}_{\theta_0})$ changes the OAM quantum number by k

$$\begin{aligned}\exp(\mathrm{i}k\hat{\varphi}_{\theta_0})|m\rangle &= \sum_{n=0}^{2L}\exp(\mathrm{i}k\theta_n)\hat{P}(\theta_n)\sum_{n'=0}^{2L}\frac{\exp(\mathrm{i}m\theta_n)}{\sqrt{2L+1}}|\theta_n'\rangle,\\ &= \sum_{n=0}^{2L}\frac{\exp(\mathrm{i}(m+k)\theta_n)}{\sqrt{2L+1}}|\theta_n\rangle = |m\oplus k\rangle.\end{aligned} \tag{IV.111}$$

Here, we have used the decomposition of an OAM eigenstate in terms of angle eigenstates as discussed in chapter III. The addition is to be understood modulo $2L+1$ such that

$$|m\oplus k\rangle = \begin{cases}\exp[\mathrm{i}(2L+1)\theta_0]|m+k-(2L+1)\rangle & m+k > L\\ \exp[-\mathrm{i}(2L+1)\theta_0]|m+k+(2L+1)\rangle, & m+k < -L.\end{cases} \tag{IV.112}$$

The symbol for the modulo addition $\oplus$ will be used elswhere in this thesis for a different modulo addition. The particular meaning will be clear from the context. For a general state $|\psi\rangle$ the shift operator $\exp(\mathrm{i}k\hat{\varphi}_{\theta_0})$ changes the mean value of the orbital angular momentum by k, where the change is again understood modulo $2L+1$. If we consider a shifted state $|\psi'\rangle = \exp(\mathrm{i}k\hat{\varphi}_{\theta_0})|\psi\rangle$ we find that the mean value $\langle\hat{m}\rangle_{\psi'}$ is

given by:

$$\begin{aligned}\langle\hat{m}\rangle_{\psi'} &= \lim_{L\to\infty}\sum_{m,m'=-L}^{L} g^*_{m'}g_m\langle m'|\exp(-ik\hat{\varphi}_{\theta_0})\hat{m}\exp(-ik\hat{\varphi}_{\theta_0})|m\rangle,\\ &= \lim_{L\to\infty}\sum_{m,m'=-L}^{L} g^*_{m'}g_m\langle m'\oplus k|\hat{m}|m\oplus k\rangle,\\ &= \sum_{m=-\infty}^{\infty}|g_m|^2(m+k) = \langle\hat{m}\rangle_\psi + k,\end{aligned} \tag{IV.113}$$

where we have used the fact that the OAM probabilities g_m sum to unity. The modulo addition is no longer required as we have taken the limit of $L\to\infty$ in the last step. To show that the OAM variance is left invariant we calculate Δm^2 for the states $|\psi\rangle$ and $|\psi'\rangle = \exp(ik\hat{\varphi}_{\theta_0})|\psi\rangle$:

$$(\Delta m)^2_\psi = \langle\hat{m}^2\rangle_\psi - (\langle\hat{m}\rangle_\psi)^2. \tag{IV.114}$$

The effect of the shift operator on the mean value has been calculated in Eq. (IV.113). The mean square for the shifted state is given by:

$$\begin{aligned}\langle\hat{m}^2\rangle_{\psi'} &= \lim_{L\to\infty}\sum_{m,m'=-L}^{L} g^*_{m'}g_m\langle m'|\exp(-ik\hat{\varphi}_{\theta_0})\hat{m}^2\exp(-ik\hat{\varphi}_{\theta_0})|m\rangle,\\ &= \lim_{L\to\infty}\sum_{m,m'=-L}^{L} g^*_{m'}g_m\langle m'\oplus k|\hat{m}^2|m\oplus k\rangle,\\ &= \sum_{m=-\infty}^{\infty}|g_m|^2(m+k) = \langle\hat{m}^2\rangle_\psi + 2\langle\hat{m}\rangle k + k^2.\end{aligned} \tag{IV.115}$$

This completes the proof hat the OAM shift operator leaves the variance unchanged.

To show hat the unitary shift operators form a one parameter group we have to prove that $\hat{U}_k\hat{U}_{k'} = \hat{U}_{k+k'}$:

$$\begin{aligned}\hat{U}_k\hat{U}_{k'} &= \sum_{n,n'=0}^{2L}\exp(ik\theta_n)\exp(ik'\theta_{n'})\hat{P}(\theta_n)\hat{P}(\theta'_n),\\ &= \sum_{n=0}^{2L}\exp(i(k+k')\theta_n)\hat{P}(\theta_n) = \hat{U}_{k\oplus k'},\end{aligned} \tag{IV.116}$$

where the addition is understood modulo $2L+1$ such that $k\oplus k'$ lies between $-L$ ans L. The unitarity of the OAM shift operator is a consequence of the Hermiticity of the angle operator $\hat{\varphi}_{\theta_0}$, but we prove here that the Hermitian conjugate of the shift operator $\hat{U}_k$ is its inverse $\hat{U}_{-k}$ shifting the OAM by $-k$

$$\hat{U}^\dagger_k = \sum_{n=0}^{2L}\exp(-ik\theta_n)\hat{P}^\dagger(\theta_n) = \sum_{n=0}^{2L}\exp(-ik\theta_n)\hat{P}(\theta_n) = \hat{U}_{-k}. \tag{IV.117}$$

Therefore $\hat{U}^\dagger_k\hat{U}_k = \hat{U}_{-k}\hat{U}_k = \hat{U}_0 = \hat{I}$, where $\hat{I}$ is the identity operator and $\hat{U}_0$ is identified as the neutral element of the one-parameter group.

IV.C Infinite series for the complex error function

Owing to the complex error function it is difficult to gain a qualitative insight of some of the physical properties of the intelligent states. In particular the behaviour of the OAM probability amplitudes is hard to understand from the Eqs. (IV.37) to (IV.38). Using the approximate expression of the complex error function in terms of an infinite series in Eq. (IV.17a) on page 54 can explain some aspects of the distribution of the probability amplitudes.

For the small-uncertainty case we find that the expression for the g_m takes on the following form after substituting the complex error function with the infinite series for $\lambda > 0$:

$$g_m = \frac{N_{\text{int}}}{\sqrt{\lambda}} \exp\left(-\frac{m^2}{2\lambda}\right)\left[\operatorname{erf}\left(\frac{\sqrt{\lambda}\pi}{\sqrt{2}}\right) + \exp\left(-\frac{\lambda\pi^2}{2}\right)\left\{\frac{1}{\sqrt{2\lambda}\pi^2} \times (1-(-1)^m) + 2\sqrt{2\lambda}\sum_{n=1}^{\infty}\frac{\exp(-\frac{1}{4}n^2)}{n^2+2\lambda\pi^2}\left[1-(-1)^m\cosh\left(-n\frac{m}{\sqrt{2\lambda}}\right)\right]\right\}\right]. \tag{IV.118}$$

One can see that the alternating factor contained in the sum becomes important for $|m| \to \infty$, as the hyperbolic cosine will become much larger than one. In the central region around $m = 0$ and in particular for large λ the distribution follows dominantly the Gaussian given by $\exp(-m^2/(2\lambda))$. This behaviour can be seen in the graphs in figure (IV.5) on page 61.

In the large uncertainty case the expression shows the alternating behaviour over the whole range of m. This can be seen on substituting the infinite series for the complex error function in Eq. (IV.38) for $\lambda < 0$:

$$g_m = (-1)^m \frac{N_{\text{int}}}{\sqrt{|\lambda|}} \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{\exp(-\frac{1}{4}n^2)}{n^2+2m^2/|\lambda|} n \sinh\left(n\frac{\sqrt{|\lambda|}\pi}{\sqrt{2}}\right). \tag{IV.119}$$

The fraction inside the sum has the form of a Lorentzian distribution with a weight factor given by the exponential in n^2 and the hyperbolic sine function. In the limit of large $|\lambda|$ the contribution from a single dominant Lorentzian distribution with the largest weight becomes more prominent. This explains the validity of our statement that the distribution of the g_m in the limit of $|\lambda| \to \infty$ is approximately Lorentzian.

IV.D Zero angular mean

We have seen for the intelligent states that the unitary operator $\exp(ik\hat{\varphi}_{\theta_0})$ shifts the OAM mean by k without changing the OAM variance $(\Delta L_z)^2$ or the angle variance $(\Delta\varphi)^2$. Whether the OAM is given as a constraint or whether it is the variance to be minimised, in both cases we are free to choose the OAM mean. For simplicity we choose $\langle \hat{L}_z \rangle = 0$. Similarly, a shift in the angular coordinate does not change $(\Delta L_z)^2$. However, the angle variance for a shifted angular coordinate is

$$(\Delta\varphi)^2 = \int_{\tau_0}^{\tau_0+2\pi} P(\varphi-\Phi)Y^2(\varphi)\mathrm{d}\varphi - \left(\int_{\tau_0}^{\tau_0+2\pi} P(\varphi-\Phi)Y(\varphi)\mathrm{d}\varphi\right)^2, \tag{IV.120}$$

where $Y(\varphi) = Y(\varphi + 2\pi n), n \in \mathbb{Z}$ is the angle representation of $\hat{\varphi}_{\theta_0}$ and Φ is the shift in the angle coordinate. On differentiating Eq. (IV.120) on the preceding page with respect to Φ we find a condition on Φ to have a minimum in the angle variance. By using the identity $\mathrm{d}/(\mathrm{d}\Phi)P(\varphi - \Phi) = -\mathrm{d}/(\mathrm{d}\varphi)P(\varphi - \Phi)$ and partial integration we obtain for the derivative of $(\Delta\varphi)^2$:

$$\frac{\mathrm{d}}{\mathrm{d}\Phi}(\Delta\varphi)^2 = 2\int_{\tau_0}^{\tau_0+2\pi} P(\varphi - \Phi)Y(\varphi)\mathrm{d}\varphi. \qquad \text{(IV.121)}$$

In order to have a minimum, or more generally an extremum, this has to be zero, which corresponds to choosing Φ such that the mean value $\langle\hat{\varphi}_{\theta_0}\rangle = 0$. Hence, minimising $(\Delta\varphi)^2$ is consistent with minimising $\langle\hat{\varphi}^2_{\theta_0}\rangle$.

Bibliography

Abramowitz, M. & Stegun, I. S. (1974), *Handbook of Mathematical Functions*, Dover Publications, Inc., Mineola, New York. Reprint. Originally published: National Bureau of Standards, corrected edition, 1964. ISBN: 0-486-61272-4.

Albright, J. R. (1977), 'Integrals of products of Airy functions', *Journal of Physics A* **10**(4), 485–490.

Aragone, C., Chalbaud, E. & Salamó, S. (1976), 'On intelligent spin states', *Journal of Mathematical Physics* **17**(11), 1963–1971.

Aragone, C., Guerri, G., Salamó, S. & Tani, J. L. (1974), 'Intelligent spin states', *Journal of Physics A* **7**(15), L149–L151.

Barnett, S. M. & Pegg, D. T. (1990), 'Quantum theory of rotation angles', *Physical Review A* **41**(7), 3427–3435. DOI: 10.1103/PhysRevA.41.3427.

Bohm, D. (1951), *Quantum Theory*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey. Reprinted in 1989, Dover Publications, Inc., New York, ISBN: 0-486-65969-0.

Franke-Arnold, S., Barnett, S. M., Yao, E., Leach, J., Courtial, J. & Padgett, M. (2004), 'Uncertainty principle for angular position and angular momentum', *New Journal of Physics* **6**, 103. DOI: 10.1088/1367-2630/6/1/103.

Galindo, A. & Pascual, P. (1990), *Quantum Mechanics*, Vol. I, Springer Verlag, Berlin. ISBN: 3-540-51406-6.

Götte, J. B., Radmore, P. M., Zambrini, R. & Barnett, S. M. (2006), 'Angular minimum uncertainty states with large uncertainties', *Journal of Physics B* **39**, 2791–2801. DOI: doi:10.1088/0953-4075/39/12/013.

Götte, J. B., Zambrini, R., Franke-Arnold, S. & Barnett, S. M. (2005), 'Large-uncertainty intelligent states for angular momentum and angle', *Journal of Optics B* **7**, S563–S571. DOI: 10.1088/1464-4266/7/12/019.

Heisenberg, W. (1927), 'Über den anschaulichen Inhalt der quantentheoretischen Kinematik and Mechanik', *Zeitschrift für Physik* **43**, 172.

Holevo, A. S. (2001), *Statistical Structure of Quantum Theory*, no Monographs 67 *in* 'Lecture Notes in Physics', Springer Verlag, Berlin. ISBN: 3-540-42082-7.

Jackiw, R. (1968), 'Minimum uncertianty product, number-phase uncertainty product and Coherent States', *Journal of Mathematical Physics* **9**(3), 339–346.

Louisell, W. H. (1973), *Quantum Statistical Properties of Radiation*, Pure and Applied Optics, John Wiley & Sons, Inc., New York. ISBN: 0-471-54785-9.

Pegg, D. T., Barnett, S. M., Zambrini, R., Franke-Arnold, S. & Padgett, M. (2005), 'Minimum uncertainty states of angular momentum and angular position', *New Journal of Physics* **7**, 62. DOI: 10.1088/1376-2630/7/1/062.

Robertson, H. P. (1929), 'The Uncertainty Principle', *Physical Review* **34**, 163–164.

Schrödinger, E. (1930), 'Zum Heisenbergschen Unschärfeprinzip', *Sitzungsberichte der Preussischen Akademie der Wissenschaften Physikalisch-mathematische Klasse* **XIX**, 296–303.

Schwabl, F. (1995), *Quantum Mechanics*, 2nd revised edn, Springer Verlag, Berlin. ISBN: 3-540-59187-7.

Stephenson, G. & Radmore, P. M. (1993), *Advanced Mathematical Methods for Engineering and Science Students*, Cambrige University Press, Cambridge. ISBN: 0-521-36860-X.

Summy, G. S. & Pegg, D. T. (1990), 'Phase optimized quantum states of light', *Optics Communications* **77**(1), 75–79. DOI: doi:10.1016/0030-4018(90)90464-5.

Vallée, O. & Soares, M. (2004), *Airy functions and applications to Physics*, Imperial College Press, London. ISBN: 1-86094-478-7.

V Chapter

Angular EPR paradox

The Einstein-Poldosky-Rosen (EPR) *Gedankenexperiment* reveals an apparent paradox if the premises in the EPR argument of local realism and a complete description of the physical reality by quantum mechanics are held to be true. In juxtaposition to EPR's solution to the paradox of concluding that quantum mechanics is incomplete, J. S. Bell showed that any attempt to complete quantum theory by the introduction of 'hidden variables' leads to an observable contradiction with quantum mechanics. This opened the possibility to test experimentally the validity of local 'hidden variable' theories. Subsequent experiments gave results in favour of traditional quantum mechanics, which requires a non-local description of the entangled quantum system considered by EPR. In this chapter we reformulate the EPR *Gedankenexperiment* for angle and orbital angular momentum. By considering indeterminacies in the measurement of the conjugate variables we derive an experimentally testable criterion for the angular version of the EPR paradox.

1 Introduction

With their *Gedankenexperiment* Einstein, Podolsky and Rosen (EPR) asked questions about the completeness of quantum mechanics (Einstein et al., 1935). The critical point in the original form of the argument is the apparent violation of the Heisenberg uncertainty principle (Heisenberg, 1927) for an entangled quantum state if one adheres to the idea of local realism. In short this term refers to the notion that measurements on one subsystem do not influence the other subsystem instantaneously. Inferring with certainty the value of an observable in one subsystem from a measurement on the other subsystem and without disturbing the system according the locality argument constitutes for EPR an element of reality just as a direct measurement on the first subsystem would do.

The premise of local realism in the argument seemed so naturally compelling for EPR that they concluded from the apparent violation of the uncertainty principle that their other premise, the completeness of quantum theory, has to be wrong. The theories which would complete quantum theory and which would explain why the outcome of the measurement can be inferred are usually called 'hidden variable theories'. According to these theories the statistical character of quantum mechanics originates in the same way as in classical statistical mechanics, from averages of states which are more completely determined by supplementary parameters. These parameters or variables are called 'hidden' because they are not accessible for experimental preparation. If states with hidden variables could be prepared experimentally quantum

mechanics could be proven to be wrong by direct observation. In general, however, hidden variable states are required to give the measurable quantum mechanical results only when averaged over the additional, hidden variables. This restores the statistical character of quantum mechanics and the applicability of the Heisenberg uncertainty principle.

Various mathematical considerations exists which attempt to prove that 'hidden variables' cannot exists (von Neumann, 1932; Jauch & Piron, 1963), but they all have been found to make assumptions on the hidden variables which only have to be fulfilled by quantum mechanical states (Bell, 1966). Bell proved that hidden variable theories cannot exist in accordance with the postulates of quantum mechanics, but it did not prove that hidden variables theories cannot exist outside quantum mechanics, or indeed that they cannot complete quantum mechanics. These proofs are often stated in terms of a contradiction: assumptions are made on the hidden variable states and the conclusions drawn from the assumptions lead to a contradiction which shows that the assumptions have been wrong. In hindsight, after reading the analysis of Bell (1966), it is not surprising to reach a contradiction in these arguments if the hidden variable states are required to fulfil some rules of quantum mechanical states, but it shows the conceptual difficulty in finding a suitable criterion which does not make use of quantum mechanics. Bell (1964) showed that hidden variable theories which would explain the violation of the uncertainty principle while retaining the assumption of locality would give predictions for observable quantities which differ from the predictions of quantum theory. Bell bases his argument on a version of the EPR paradox according to Bohm & Aharonov (1957) who considered two spin 1/2 particles in a singlet state. Using discrete variables such as the orientation of spin components has a considerable advantage for experiments over the continuous variables of linear position and momentum considered originally by EPR. It thus became possible to design conclusive experiments which would answer the question about the completeness of quantum theory. The subsequent experiments by Aspect et al. (1982*a*,*b*) on an optical version of the discrete EPR paradox using the polarisation of photons were in agreement with the quantum mechanical predictions.

EPR paradoxes can be construed for any two complimentary observables, but the experimental demonstration requires a source generating the correlations in both variables. Optical parametric down conversion, a non-linear process, in which a birefringent crystal is used to convert one incident photon into two photons of lower energy, is a well studied source of entangled photon pairs (Louisell et al., 1961; Rubin et al., 1994; Kwiat et al., 1995). This is why the implications of the EPR paradox have been mainly tested on optical systems, for example on the polarisation of photons (Aspect et al., 1982*a*; Weihs et al., 1998), quadrature phase components (Reid, 1989; Ou et al., 1982) or directly on the optical version of EPR's original example, the linear momentum and linear position of photons (Howell et al., 2004). In this chapter we study the EPR paradox for different variables, orbital angular momentum (OAM) and angle of light. The relation between OAM and its conjugate variable, the angular position, is fundamentally different from other systems, because OAM is a discrete observable of infinite dimension and the angular position is continuous and bounded. The entanglement for OAM of photon pairs generated in parametric down

conversion has been confirmed both experimentally (Mair et al., 2001) and theoretically (Franke-Arnold et al., 2003). It is therefore interesting to examine the possibility to demonstrate an angular EPR paradox for this pair of observables, in particular as the necessary experimental techniques have already been employed in recent work in this field (Franke-Arnold et al., 2004; Leach et al., 2002).

The fundamental interest in studying EPR paradoxes is the startling revision of our understanding of correlated systems. The correlations for physical systems which exhibit EPR paradoxes cannot be explained by local interactions. This is the result from Bell's consideration and Aspect's experiments. But it is not obvious that this can be demonstrated for any two complimentary observables. This point is of particular importance if the measurement of the variables cannot be precise, not due to the statistical character of quantum mechanics, but because the variable is continuous and can only be determined within a range of values. The inherent indeterminacy in measuring the angle will be a major aspect in studying the nature of the EPR correlations for angle and orbital angular momentum (OAM). Apart from the fundamental interest in EPR correlations, the criterion for an EPR paradox for OAM and angular position provides a tool to characterise entanglement for these observables. Using variances in violations of local uncertainty relations as entanglement criteria has received renewed interest. Hofmann & Takeuchi (2003) have shown that variances of special observables can be used to detect entanglement in finite-dimensional systems. This approach has been generalised to arbitrary observables by Gühne (2004).

We first review the paradox arising from the EPR *Gedankenexperiment* in more detail. We then devise a criterion for an EPR paradox for OAM and angle, which is suitable for experimental tests. The criterion takes indeterminacies in the angle measurement into account which are inherent to the measurement of a continuous variable. We propose an experimental scheme and discuss the necessary measurements and finally we are modeling the measurement process theoretically. For that we have to make assumptions on the correlations of a photon pair generated by parametric down conversion. We show that under these assumptions we can expect a violation of the angular uncertainty relation and hence a demonstration of the angular EPR paradox. This work is an original contribution to the literature and has been published in part in (Götte et al., 2006*a*).

2 Formulation of the paradox

In this section the criterion for a demonstration of the angular EPR paradox is derived and discussed. We first reexamine the original EPR *Gedankenexperiment* and explain why the inclusion of statistical and methodical errors requires a reformulation in terms of conditional variances. The violation of the Heisenberg uncertainty relation in the original EPR argument is based on inference, we therefore have to distinguish clearly between inferred and measured quantities. With this distinction we are able to give an experimentally testable criterion for an angular EPR paradox in terms of measured and inferred conditional variances for angle and OAM.

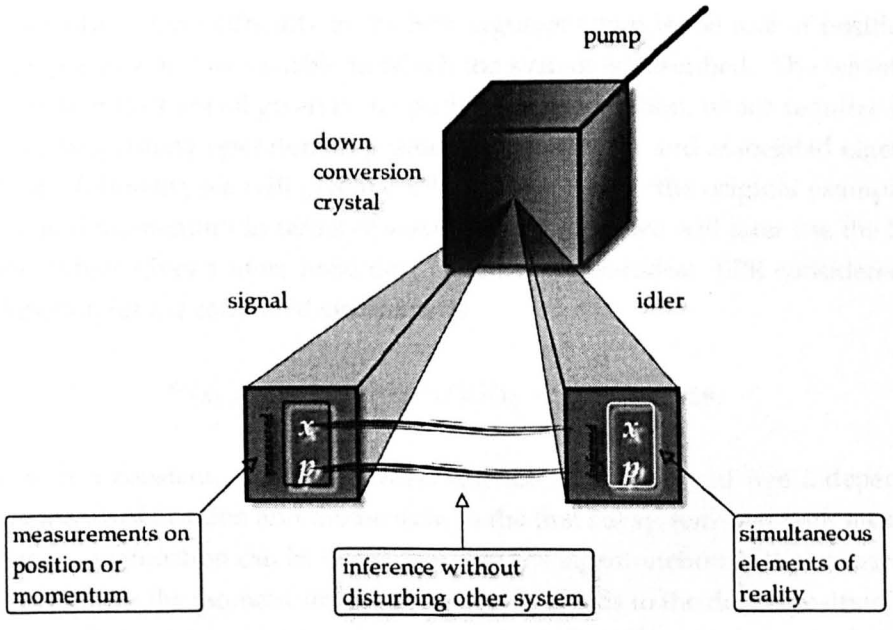

Figure V.1 — Schematic picture of the EPR *Gedankenexperiment*. Two independent measurements on position and momentum on one subsystem are used to infer the corresponding physical quantity in the other subsystem. The chosen example shows a photon pair generated by a parametric down conversion process. After the parametric down conversion the two subsystems do not interact and according to EPR a measurement on one subsystem cannot have an instantaneous effect on the other subsystem. The inference with certainty constitutes an element of reality for the predicted values of position and momentum in the second subsystem. With the definition of reality by EPR the inferred values have a simultaneous element of reality which contradicts the Heisenberg uncertainty relation (see Howell et al. (2004)).

2.1 EPR paradox

The EPR paradox describes the apparent violation of the uncertainty principle resulting from measurements on correlated, spatially separated systems. The original EPR argument considers correlations that are strong enough to predict or infer with certainty the values of observables in one subsystem from measurements on the other, separated, subsystem without disturbing in any way the first subsystem. The ability to predict with certainty the value of an observable defines, according to EPR, an element of reality. However, non-commuting observables, cannot have a simultaneous reality, an expression of which is the uncertainty principle (Robertson, 1929). The tension between local elements of reality and quantum complementarity leads to the paradox.

The mathematical formulation of the original argument considers two separated systems which are permitted to interact for a certain time, but after this time the interaction stops and the two subsystems are separated according to the locality argument. This setup is shown in figure (V.1) on the example of a photon pair generated by parametric down conversion. The state of the combined system is given by the wavefunction $\Psi(x_1, x_2)$, where x_1 and x_2 are variables to describe the two subsystems, here signal and idler. Although EPR state their argument in all generality they consider a specific example for the conjugate physical quantities of linear position

and momentum. One difficulty in the EPR argumentation is the role of position as physical quantity and as variable in which the system is described. The wavefunctions and operators are all given in the position representation, which requires some care in distinguishing operators in position representation and associated eigenvalues. In the following we will give the EPR argument as in the original example for position and momentum in terms of wavefunctions, and we will later use the Dirac notation, which gives a more lucid description of the paradox. EPR considered the wavefunction for the combined system to be

$$\Psi(x_1, x_2) = \int_{-\infty}^{\infty} \exp[(\mathrm{i}/\hbar)(x_1 - x_2 + x_0)p)\mathrm{d}p, \tag{V.1}$$

where x_0 is a constant. Using this wavefunction EPR examined two independent measurements of position and momentum in the first subsystem. For both measurements the wavefunction can be decomposed in the eigenfunction of the corresponding operator. For the momentum measurement this leads to the decomposition

$$\Psi(x_1, x_2) = \int_{-\infty}^{\infty} \psi_p(x_2) u_p(x_1)\mathrm{d}p, \tag{V.2}$$

where $u_p(x_1)$ is the eigenfunction of the momentum operator in position representation:

$$u_p(x_1) = \exp[(\mathrm{i}/\hbar)px_1]. \tag{V.3}$$

This requires that $\psi_p(x_2)$ has the form

$$\psi(x_2) = \exp[-(\mathrm{i}/\hbar)(x_2 - x_0)p], \tag{V.4}$$

which is also an eigenfunction of the momentum operator, but now in the second subsystem with the eigenvalue $-p$. If a measurement of the position in the first subsystem gives a result of p, the first subsystem will be left in the state given by the eigenfunction $u_p(x_1)$, while the other subsystem will be in a state described by the wavefunction $\psi_p(x_2)$. As this is an eigenfunction of the momentum operator a subsequent measurement on the momentum in the second subsystem would give the corresponding eigenvalue $-p$. After measuring p in the first subsystem the momentum in the other subsystem is known with certainty and the uncertainty of the momentum measurement in the second subsystem in zero. In a similar way a position measurement can be considered. Here, the decomposition is in terms of the eigenfunctions of the position operator $v_x(x_1)$:

$$\Psi(x_1, x_2) = \int_{-\infty}^{\infty} \phi_x(x_2) v_x(x_1)\mathrm{d}x. \tag{V.5}$$

The eigenfunctions $v_x(x_1)$ in the position representation are given by the Dirac δ-function $v_x(x_1) = \delta(x_1 - x)$. The weight function $\phi_x(x_2)$ is then given by the integral representation

$$\phi_x(x_2) = \int_{-\infty}^{\infty} \exp[(\mathrm{i}/\hbar)(x - x_2 + x_0)p\mathrm{d}p = 2\pi\hbar\delta(x - x_2 + x_0). \tag{V.6}$$

This Dirac δ-function is the eigenfunction of the multiplicative position operator in the second subsystem for the eigenvalue $x + x_0$. By following the same reasoning as above for the momentum a measurement of the position in the first subsystem with the result x determines the outcome of a position measurement in the second subsystem with certainty. Position and momentum in the second subsystem can thus be inferred with certainty from measurements in the first subsystem. The two measurements in the first subsystem are independent, not simultaneous and leave the first subsystem in two different states. But this is of no relevance for the EPR argument. For EPR the possibility to predict position or momentum in the second subsystem from a measurement in the first subsystem which does not disturb the second subsystem gives the corresponding physical quantity in the second subsystem an element of reality. In the definition of reality according to EPR a measurement in one subsystem which does not disturb the other subsystem cannot have an influence on the reality of physical quantities in the other, undisturbed subsystem. Therefore, the inferred position and the inferred momentum in the second subsystem have simultaneous elements of reality, which contradicts the Heisenberg uncertainty principle.

The original EPR *Gedankenexperiment* considers an idealised situation. The quantum state given by EPR on the example of the position and momentum is – in the modern language of entanglement – a maximally entangled state (Nielsen & Chuang, 2000). This becomes obvious if the state given by the wavefunction in Eq. (V.2) on the previous page is written in the Dirac notation. From the reasoning for the momentum measurement follows that the state $|\Psi\rangle$ is given by

$$\Psi(x_1, x_2) = \langle x_1, x_2|\Psi\rangle \longrightarrow |\Psi\rangle = \int \mathrm{d}p|-p\rangle_2|p\rangle_1, \tag{V.7}$$

where $|p\rangle_i$ denotes the momentum eigenstate in subsystem $i = 1, 2$ with eigenvalue p. This is a continuous Schmidt decomposition where all Schmidt coefficients are equal and the state is thus maximally entangled. Also, the measurement of the observables is assumed to be infinitely precise. For OAM and angular position this idealised setting would require a parametric down conversion process which creates an entangled photon pair, perfectly correlated in OAM and angular position. An errorless measurement of the OAM on the signal photon could then be used to infer the OAM of the idler photon, and an errorless measurement of the azimuthal angle on the signal photon would allow us to predict precisely the angle of the idler photon. As these measurements on the signal photon 'do not disturb the idler photon in any way', the predictions would constitute simultaneous elements of reality for the OAM and the azimuthal angle of the idler photon. We stress that the possibility to predict observables of the idler photon with certainty depends on the ability to measure the observables on the signal photon without error.

In particular for a continuous observable a measurement with infinite precision cannot be realised experimentally. A typical experimental setup would allow us to determine whether a continuous variable falls into a previously specified range. To analyse the possibility of demonstrating an EPR paradox experimentally a more realistic situation has to be studied. This requires the consideration of non-maximal correlations and of measurements with finite precision leading to an error in infer-

ring one observable from a measurement on the other subsystem. The size of this error determines whether the EPR paradox can be demonstrated in the considered experimental setup (Reid, 1997).

2.2 Indeterminacies in Preparation and Measurement

The inclusion of indeterminacies in preparation and measurement requires a reformulation of the EPR paradox. If the correlation between the two subsystem is not perfect it will not be possible to infer with certainty the value of a physical quantity in the other subsystem. Furthermore, the measurement on the first subsystem may have an experimental error, this again would render a precise prediction impossible. The question of experimental indeterminacies in the measurement will be of great importance for the angular version of the EPR paradox. For the moment we restrict the reformulation to the case, where the physical quantities may be measured precisely but the imperfect correlations forbid a inference with certainty. In this scenario we can consider conditional variances $\mathrm{var}[x_2|x_1]$, that is variances for the position in the second subsystem, say the idler, given that a measurement in the first subsystem, the signal, gives a value of x_1. For two perfectly correlated subsystems a measurement of a single value of x_1 in the signal would correspond to a single value in the idler and the conditional variance would thus be equal to zero. But in general the imperfect correlations will introduce a spread in the measurement of x_2. In a similar way a conditional variance for the momentum $\mathrm{var}[p_2|p_1]$ can be constructed. By using a local version of the Heisenberg uncertainty relation for the second subsystem $\Delta x_2 \Delta p_2 > \hbar/2$, one can infer the minimum variance $\min \mathrm{var}[x_2|p_1]$, which is still in accordance with the uncertainty relation. To distinguish measured and inferred quantities we are going to label them with the indices 'm' and 'i' respectively. Within the EPR argument the measurement on the signal, which sets the condition in the variance $\mathrm{var}[x_2|x_1]$, does not have an instantaneous influence on the other subsystem. In this sense the measured variance $\mathrm{var}[x_2|x_1]_\mathrm{m}$ and the inferred variance $\mathrm{var}[x_2|p_1]_\mathrm{i}$ can be compared and if

$$\mathrm{var}[x_2|x_1]_\mathrm{m} < \min \mathrm{var}[x_2|p_1]_\mathrm{i} \tag{V.8}$$

the EPR paradox would be demonstrated for the particular set of conditions. In this formulation the two positions have a different character, x_2 is a position variable, and the variance $\mathrm{var}[x_2|p_1]_\mathrm{i}$ gives a measure for the spread in this variable, but x_1 and also p_1 designate a specific position and a specific momentum. If the inequality (V.8) can be demonstrated for one set of conditions x_1 and p_1 it does not necessarily constitute a demonstration of the EPR paradox in general. This is particularly valid as we consider imperfect correlations, and we cannot assume that the correlations in the position are the same for all values of x_1 nor that the momenta are identically correlated for all values or p_1. For this reason we are taking the average on both sides in Eq. (V.8) over the respective conditions x_1 and p_1

$$\langle \mathrm{var}[x_2|x_1]_\mathrm{m} \rangle_{x_1} < \langle \min \mathrm{var}[x_2|p_1]_\mathrm{i} \rangle_{p_1} . \tag{V.9}$$

The averaging involves the probability densities $\mathcal{P}(x_1)$ and $\mathcal{P}(p_1)$. These are the local probability densities for the variables x_1 and p_1, and give the probability for a local measurement of the position or momentum in the first subsystem. A criterion which compares a measured variance in the momentum $\mathrm{var}[p_2|p_1]_m$ with an inferred variance in the position $\mathrm{var}[x_2|x_1]_i$ can be derived analogously.

Another conceptual difficulty arises, when we take into account that the measurement will not be precise. This has implications not only for measuring the conditional variance in an experiment but it also affects the condition in the conditional variances. Until now the condition was set by the measurement of a single value of the particular position or the momentum. For the angular version of the EPR paradox we consider OAM and angle. Orbital angular momentum is a discrete variable and the condition can be set to the measurement of a single value m_1. But for the angle a measurement can only determine if the angle falls within a range of angles. Similarly, a condition can only be set on a range of angles. For the OAM we denote the conditional variance with $\mathrm{var}[m_2|m_1]$, i.e. the variance of m_2 in the idler under the condition that a measurement on the signal photon yields a value of m_1. The condition for the azimuthal position will be set by passing an angular aperture. This gives a conclusive yes or no answer just as measuring a particular value of m_1. The angular aperture will be denoted by $A_1(\varphi_1)$, where φ_1 gives the orientation of the mask. A convenient choice of φ_1 for a symmetric aperture would be the central angle. A simple example of these angular aperture is shown in figure (V.2a) on page 101; a hard-edged aperture in form of a circle segment. But it is possible to realise smooth apertures which gradually become more opaque (see figure (V.2b)). The conditional variance for the angle can thus be written as $\mathrm{var}[\varphi_2|A_1(\varphi_1)]$, which should be read as the variance of the angle φ_2 in the idler under the condition that in the signal the photon passes the mask A_1 positioned at φ_1.

2.3 Criterion for an angular EPR paradox

In the previous sections 2.1 and 2.2 we have emphasizes the importance of inference in EPR type arguments and the use of conditional variance to consider indeterminacies. To set a condition for the angle, angular apertures are used which have an effect on the range of angles that are detected behind such a mask. In the simplest case of a 'cake-slice' mask as in figure (V.2) on page 101 the angular position of the photon has to fall in a range of angles set by the mask.

To demonstrate an angular version of the EPR paradox the measured conditional variance $\mathrm{var}[m_2|m_1]_m$ can be compared to the inferred quantity $\mathrm{var}[m_2|A_1(\varphi_1)]_i$. The reasoning is identical to the previous section 2.2: from a measured conditional variance $\mathrm{var}[\varphi_2|A_1(\varphi_1)]_m$ a minimum variance can be inferred which is still in accordance with the local angular uncertainty relation (III.30) on page 46. If the measured variance for m_2 is smaller than the minimum inferred variance, that is if

$$\mathrm{var}[m_2|m_1]_m < \min \mathrm{var}[m_2|A_1(\varphi_1)]_i, \tag{V.10}$$

the angular EPR paradox would be demonstrated for this particular choice of conditions. To arrive at a general statement, however, the average has to be taken on both

sides over the respective conditions. For the OAM this is done by summing over all possible values of m_1 weighted with the probabilities $|c_{m_1}|^2$. For the inferred variance, however, the average is taken over the orientation of the angular aperture φ_1:

$$\langle \mathrm{var}[m_2|m_1]_\mathrm{m}\rangle_{m_1} < \langle \min \mathrm{var}[m_2|A_1(\varphi_1)]_\mathrm{i}\rangle_{\varphi_1} . \tag{V.11}$$

The weight function in this average is the probability that the photon passes the angular aperture A_1 at the orientation φ_1. This is analogous to the case of position and momentum in section 2.2. There, the condition was set to a single value of the position or momentum in the signal. The average contained therefore the probability for this particular value of x_1 or p_1. Here, the condition is set by detecting the photon behind the aperture and consequently the average contains the probability for passing the mask at the given orientation. The criterion for a demonstration reads thus as

$$\sum_{m_1} |c_{m_1}|^2 \mathrm{var}[m_2|m_1]_\mathrm{m} < \int_{-\pi}^{\pi} \mathrm{d}\varphi_1 \mathcal{P}[A_1(\varphi_1)] \min \mathrm{var}[m_2|A_1(\varphi_1)]_\mathrm{i} . \tag{V.12}$$

This criterion depends on the form of the aperture A_1, but the dependence on the particular form of the aperture is very much different from the dependence on the orientation. The latter is connected to the correlations in the angle for the entangled photon pair. We cannot assume a priori a particular form of the correlations and we therefore have to average over all orientation angles φ_1 to achieve generality for the criterion. The influence of the aperture in the criterion is connected to the way in which the angle is measured. This will become more obvious when the apertures are translated into probability densities for the angle.

In the variance $\mathrm{var}[m_2|A_1(\varphi_1)]_\mathrm{i}$ the condition is set by passing an angular aperture, which restricts the range of angles. The most simple example of such an aperture is a 'cake-slice' mask, which only lets through photons within a limited range of angles and blocks all others. If a photon is detected behind such a mask it must have been in the range of angles allowed by the mask. We can thus assign a probability density for the angular position of the passing photon. For a 'cake-slice' mask all allowed angles are equally likely and the probability density is given by a rectangular function. But we are also considering apertures which are represented by smooth functions such as shown in figure (V.2) on the following page.

The assignment of a probability density to a detected photon is at the heart of the EPR argument. Prior to the detection the angular position is completely undetermined. Setting the condition by the angular aperture and detecting a photon behind this aperture is a rough measurement of the angular position. The angular probability density of the photon that has passed the mask is thus given by the corresponding probability density of the angular aperture. It is, however, different from the angular probability density before passing the aperture. We write $P_1(\tau_1;\varphi_1)$ for the probability density representing the angular aperture $A_1(\varphi_1)$. The measured conditional variance is thus given by $\mathrm{var}[\varphi_2|P_1(\tau_1;\varphi_1)]_\mathrm{m}$, which should be read as the variance of the angle φ_2 in the idler under the condition that the paired photon in the signal has passed the angular aperture described by the probability density $P_1(\tau_1;\varphi_1)$. The probability $\mathcal{P}[A_1(\varphi_1)]$ in Eq. (V.12) translates thus into $\mathcal{P}[P_1(\tau_1;\varphi_1)]$ which is the

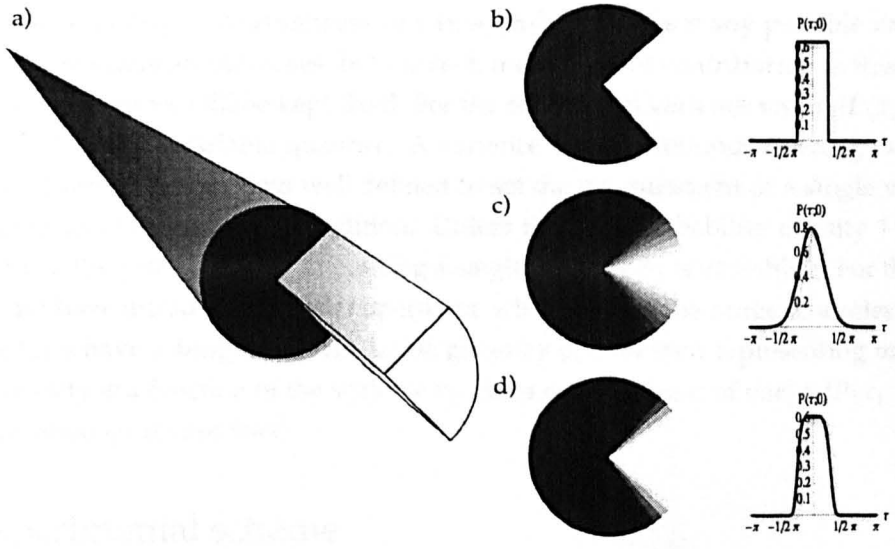

Figure V.2 — The graph shows different angular apertures which are used to set a condition on the angular position of the photon in the signal. b) shows a hard-edged aperture in form of a circle segment. The corresponding probability density is given by a rectangular function. c) shows a smooth aperture represented by a truncated Gaussian. d) shows a smooth aperture represented by truncated super-Gaussians. These functions have a higher power in the exponential than Gaussians, and resemble the rectangular function while retaining the smoothness.

probability that a photon passes the angular aperture represented by $P_1(\tau_1;\varphi_1)$. This probability is given by the integral over the probability density $\mathcal{P}(\tau_1)$ before passing the angular aperture and $P_1(\tau_1;\varphi_1)$ representing the angular aperture:

$$\mathcal{P}[P_1(\tau_1;\varphi_1)] = \int_{-\pi}^{\pi} \mathrm{d}\tau_1 \mathcal{P}(\tau_1) P_1(\tau_1;\varphi_1). \tag{V.13}$$

We will write as short hand notation for this integral $[\mathcal{P} \circledast P_1](\varphi_1)$. For symmetric apertures the integral is identical to the convolution $\mathcal{P} * P_1$, but not in general (see App. V.C). The graphical meaning of this operation, however, is very similar and the notation $[\mathcal{P} \circledast P_1](\varphi_1)$ makes it clear that the construct does not depend on the variable τ_1 but only on the orientation of the aperture φ_1. In the special case, where the aperture has the form of a Dirac δ-function centred at φ_1 the integral in Eq. (V.13) reduces to the probability density $\mathcal{P}(\varphi_1)$. This is a very reasonable result, as the case of a precise condition discussed in section 2.2 can be seen as the limiting case of a condition set by an aperture with infinitely small opening. The criterion can now be written in terms of probability densities alone:

$$\sum_{m_1} |c_{m_1}|^2 \mathrm{var}[m_2|m_1]_{\mathrm{m}} < \int_{-\pi}^{\pi} \mathrm{d}\varphi_1 [\mathcal{P} \circledast P_1](\varphi_1) \min \mathrm{var}[m_2|P_1(\tau_1;\varphi_1)]_{\mathrm{i}}\,. \tag{V.14}$$

A similar criterion can also be derived when a measured $\mathrm{var}[\varphi_2|P(\tau_1;\varphi_1)]_{\mathrm{m}}$ is compared to an inferred $\mathrm{var}[\varphi_2|m_1]_{\mathrm{i}}$.

At the end of this section it is worthwhile to reflect on the meaning of the different variables in this criterion. In the conditional variance for the OAM m_2 is a variable

physical quantity; a measurement of $\mathrm{var}[m_2|m_1]_\mathrm{m}$ involves many possible values of m_2 as measurement outcomes, but for each measurement contributing to this conditional variance m_1 will be kept fixed. For the conditional variance $\mathrm{var}[\varphi_2|P(\tau_1;\varphi_1)]_\mathrm{m}$ the angle φ_2 is a variable quantity. A variance for a continuous quantity is a well defined object, but it is less well defined to set the measurement of a single value of a continuous variable as a condition. Unless the local probability density $\mathcal{P}(\varphi_1)$ is singular, the probability for measuring a single value of φ_1 is vanishing. For that reason we have introduced angular apertures, which restrict the range of angles. These apertures have a designated orientation given by φ_1, and their representing probability density is a function of the variable τ_1. For a measurement of $\mathrm{var}[\varphi_2|P(\tau_1;\varphi_1)$ the orientation φ_1 is kept fixed.

3 Experimental scheme

In this section the requirements for an experimental implementation of the criterion (V.14) on the previous page are discussed. We first examine conditional probabilities and how they are used to calculate the measured and inferred conditional variances. We then propose experimental schemes for measuring the conditional probabilities of the angle and the OAM. In the last section we discuss experimentally relevant phenomena that are not taken into account in our criterion.

3.1 Conditional variances from measurement

In the previous section we repeatedly distinguished between measured and inferred conditional variances. The term 'measured' in this distinction indicates that the conditional variance, say $\mathrm{var}[m_2|m_1]_\mathrm{m}$, is calculated directly from measured conditional probabilities. For $\mathrm{var}[m_2|m_1]_\mathrm{m}$ these probabilities are the conditional probabilities $|c[m_2|m_1]_\mathrm{m}|^2$. An experimental scheme to measure $|c[m_2|m_1]_\mathrm{m}|^2$ will be presented in the next section. But once the conditional probabilities are known from measurements, the conditional variance can be calculated by summation over m_2:

$$\mathrm{var}[m_2|m_1]_\mathrm{m} = \sum_{m_2} |c[m_2|m_1]_\mathrm{m}|^2 m_2^2 - \left(\sum_{m_2} |c[m_2|m_1]_\mathrm{m}|^2 m_2\right)^2 . \qquad \text{(V.15)}$$

In practice, only a limited number of m_2 values will be taken into account for the summation. Judging from the experimental and theoretical studies on the entanglement of OAM (Mair et al., 2001; Franke-Arnold et al., 2003), the distribution of conditional probabilities $|c[m_2|m_1]_\mathrm{m}|^2$ will be sharply peaked for $m_2 = -m_1$. For values of m_2 much different from $-m_1$ the conditional probability will be very small and these terms will not contribute significantly to the variance.

The inferred conditional variance, say $\mathrm{var}[m_2|P_1(\tau_1;\varphi_1)]_\mathrm{i}$, on the other hand is inferred from conditional probabilities of the conjugate variable, here $P[\varphi_2|P_1(\tau_1;\varphi_1)]_\mathrm{m}$. This is precisely the point, where the inference enters the line of argument in our criterion. One way to infer $\mathrm{var}[m_2|P_1(\tau_1;\varphi_1)]_\mathrm{i}$ from $P[\varphi_2|P_1(\tau_1;\varphi_1)]_\mathrm{m}$ in practice is by using the Fourier relation between OAM and angle. From the measured conditional

probability density $P[\varphi_2|P_1(\tau_1;\varphi_1)]_m$ a conditional wavefunction can be derived:

$$\psi[\varphi_2|P_1(\tau_1;\varphi_1)] = (P[\varphi_2|P_1(\tau_1;\varphi_1)]_m)^{1/2}\exp[i\alpha(\varphi_2)]. \tag{V.16}$$

Here, the phase $\alpha(\varphi_2)$ is undetermined, as the measured probability densities only give the modulus square of the wavefunction. The wavefunction is then transformed into a conditional probability amplitude for the OAM via a Fourier transform:

$$c[m_2|P_1(\tau_1;\varphi_1)]_i = \frac{1}{\sqrt{2\pi}}\int_{-\pi}^{\pi} d\varphi_2 \exp(im_2\varphi_2)\,(P[\varphi_2|P_1(\tau_1;\varphi_1)]_m)^{1/2}\exp(i\alpha(\varphi_2)). \tag{V.17}$$

From the conditional probability amplitudes we can calculate the conditional variance $\mathrm{var}[m_2|P(\tau_1;\varphi_1)]_i$ by taking the sum over all m_2 values

$$\mathrm{var}[m_2|P(\tau_1;\varphi_1)]_i = \sum_{m_2}|c[m_2|P_1(\tau_1;\varphi_1)]_i|^2 m_2^2 - \left(\sum_{m_2}|c[m_2|P_1(\tau_1;\varphi_1)]_i|^2 m_2\right)^2. \tag{V.18}$$

This is the inferred variance which can be minimised and then compared to the measured quantity $\mathrm{var}[m_2|m_1]_m$. The phase $\alpha(\varphi_2)$ will be determined by the minimisation of the conditional variance $\mathrm{var}[m_2|P_1(\tau_1;\varphi_1)]_i$ as detailed in App. V.A.

Before discussing the experimental schemes to measure the conditional probabilities we present our criterion written in terms of conditional probabilities alone:

$$\begin{aligned}\sum_{m_1}|c_{m_1}|^2\left[\sum_{m_2}|c[m_2|m_1]_m|^2 m_2^2 - \left(\sum_{m_2}|c[m_2|m_1]_m|^2 m_2\right)^2\right] \\ < \int_{-\pi}^{\pi} d\varphi_1 [\mathcal{P}\circledast P_1](\varphi_1)\min\left[\sum_{m_2}|c[m_2|P_1(\tau_1;\varphi_1)]_i|^2 m_2^2\right. \\ \left. - \left(\sum_{m_2}|c[m_2|P_1(\tau_1;\varphi_1)]_i|^2 m_2\right)^2\right].\end{aligned} \tag{V.19}$$

With this formulation and the presented method of inferring $c[m_2|P_1(\tau_1;\varphi_1)]_i$ from the measured quantity $P[\varphi_2|P_1(\tau_1;\varphi_1)]_m$ the criterion can be used directly on experimental data.

3.2 Measuring conditional probabilities

An experimental demonstration of the angular EPR paradox requires the measurement of conditional probabilities. From these probabilities the measured and inferred conditional variances can be calculated. In the following we propose experimental schemes to measure the conditional probabilities for the angle and OAM.

3.2.1 Orbital angular momentum

In order to measure $P[m_2|m_1]_m$ a condition in the signal has to be set such that only photons with an OAM of m_1 are detected, and in coincidence with the detection in the signal the OAM m_2 has to be measured in the idler. The condition in the sig-

nal can be set by an optical component which changes the OAM by $-m_1$ such that the photon has an OAM of zero after the optical element (Vaziri et al., 2002; Franke-Arnold et al., 2004). In this setup it is convenient to use a spatial light modulator to set the condition. A spatial light modulator is a phase-only diffractive component which can be readily programmed for different phase patterns changing the OAM by a different value of $-m_1$. In this way any condition m_1 can be realised without replacing optical elements. Behind the phase component the light is passed through a pinhole. Only light beams with zero OAM have on-axis intensity. The pinhole thus filters out photons whose OAM has not been changed to zero by the phase component and only photons with an OAM of m_1 are detected behind the pinhole.

The measurement of m_2 has to be in coincidence with the detection of a photon behind the pinhole in the signal. Ideally the measurement should distinguish several values for m_2. By repeating the measurement for the same condition m_1 several times the probability distribution $P[m_2|m_1]$ can be measured. An experimental scheme to sort OAM states according to the symmetry of their phase structure has been experimentally implemented by Leach et al. (2002) and a simplification has been proposed by (Wei et al., 2003). This sorting scheme works in stages, the first stage is able to distinguish even and odd values of m_2, which measures m_2 mod 2. Subsequent stages would be able to measure m_2 mod 4, m_2 mod 8 and so on. In general n stages are able to measure the OAM m_2 mod 2^n. The experimental implementation of this sorting scheme gets increasingly difficult with the number of stages, but for a strong correlation between m_1 and m_2 only a limited number of m_2 values around $-m_1$ will have to be distinguished in the experiment. Such strong correlations are to be expected and the knowledge of the possible outcomes of a measurement of m_2 can be used to associate the measurement of m_2 mod 2^n with the most likely value for m_2. But for a small n it is possible that different values of m_2 with a similar probability are sorted into the same outcome. If, for example $n = 3$, the OAM m_2 mod 8 would be measured in the idler and the values of $m_2 = -3$ and $m_2 = 5$ would be sorted into the same outcome. If the condition in the signal is set to $m_1 = 0$, one would expect fluctuations of m_2 around zero. The conditional probabilities $P[m_2 = -3|m_1 = 0]$ and $P[m_2 = 5|m_1 = 1]$ might thus have a similar value. Choosing a suitable number of stages is therefore important in measuring $P[m_2|m_1]$. A schematic representation of the experimental setup to measure this quantity is given in figure (V.3) on the following page.

3.2.2 Angular position

To calculate the conditional variance $\mathrm{var}[\varphi_2|P_1(\tau_1;\varphi_1)]_m$, the conditional probability density $P[\varphi_2|P_1(\tau_1;\varphi_1)]_m$ has to be measured. We have discussed how the condition can be set with help of an angular aperture in an earlier section (see section 2.3 and figure (V.4) on page 106). To measure $P[\varphi_2|P_1(\tau_1;\varphi_1)]_m$ the angle φ_2 in the idler has to be measured in coincidence with a photon being detected behind the angular aperture in the signal. The same difficulty in setting the condition in the conditional variance appears here for measuring the angle φ_2. Experimentally it will not be possible to measure the conditional probability density for a single angle, instead a suitable aperture can be used to test for a range of angles. Analogously to the conditioning

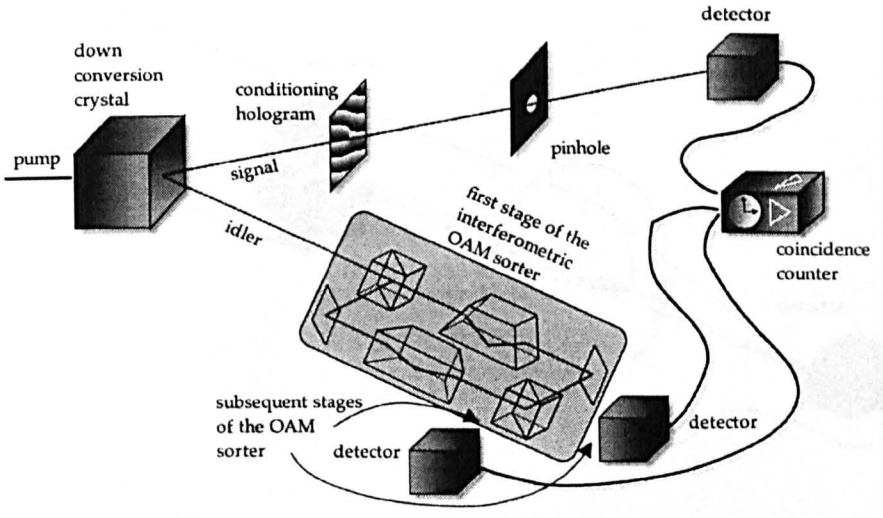

Figure V.3 — Experimental scheme for measuring $\mathrm{var}[m_2|m_1]_\mathrm{m}$. In the signal a hologram is used to single out a particular value of m_1 as a condition. The OAM distribution of the idler is determined with help of an interferometric OAM sorter (Leach et al., 2002). Only the first stage of the sorter is shown here, additional stages are added where indicated by arrows. Each stage doubles the possible outcomes and therefore the number of detectors. Eventually the signals from all detectors are transmitted to a coincidence counter.

aperture, the measured quantity is $P[A_2(\varphi_2)|P_1(\tau_1;\varphi_1)]_\mathrm{m}$, where the aperture A_2 is oriented at a particular angle φ_2. The aperture can be described by a probability density $P_2(\tau_2;\varphi_2)$, where the variable τ_2 is used to describe the function and φ_2 indicates the orientation.

By using a probability density we can make statements about the likelihood of the angular position φ_2. We cannot speak of the actual angle φ_2, as detecting the photon behind an angular aperture only allows us to assign the corresponding probability density to the state. But with this probability density we are able to calculate the likelihood of the angular position φ_2. For an aperture in form of a narrowly peaked Gaussian it is much more likely that the actual angular position is within an interval around the peak than in an interval of the same size at the flanks of the Gaussian. In this sense the measured probability can be written as $P[P_2(\tau_2;\varphi_2)|P_1(\tau_1;\varphi_1)]_\mathrm{m}$, and for a very narrow aperture A_2 oriented at φ_2 this measurement will give a good estimate of $P[\varphi_2|P_1(\tau_1;\varphi_1)]_\mathrm{m}$:

$$P[\varphi_2|P_1(\tau_1;\varphi_1)]_\mathrm{m} \approx P[P_2(\tau_2;\varphi_2)|P_1(\tau_1;\varphi_1)]_\mathrm{m} \quad \text{for suitable } P_2(\tau_2;\varphi_2). \tag{V.20}$$

The error made in this approximation is then given by the variance of τ_2 for the probability density P_2.

The scheme to measure $P[P_2(\tau_2;\varphi_2)|P_1(\tau_1;\varphi_1)]_\mathrm{m}$ is basically shown in figure (V.5). Spatial light modulators can also be used to create intensity masks. With help of such an intensity mask any chosen $P_1(\tau_1;\varphi_1)$ can be set as a condition for the signal. For the idler the aim is to determine φ_2 as exactly as possible. To achieve this a spatial light modulator can be programmed for a narrow angular aperture $P_2(\tau_2;\varphi_2)$ centred at φ_2, which would then be varied over φ_2 in a 2π radian interval. Eventually

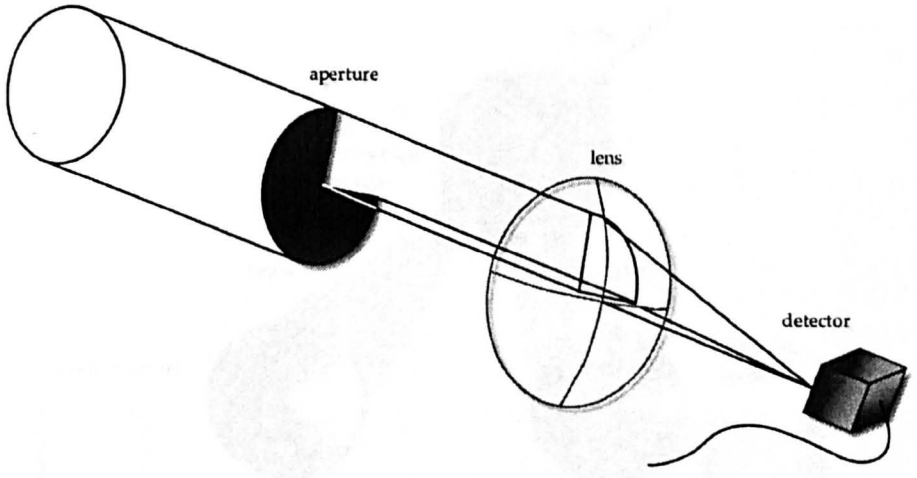

Figure V.4 — Measurement scheme for the azimuthal position. A photon is said to have a particular probability density for the azimuthal angle if it is detected after passing an aperture corresponding to this probability density. Experimentally these apertures may be shaped using a spatial light modulator.

this would lead to $P[P_2(\tau_2;\varphi_2)|P_1(\tau_1;\varphi_1)]_m$ which not only depends on the condition $P_1(\tau_1;\varphi_1)$ but also on the chosen analysing aperture $P_2(\tau_2;\varphi_2)$. This is the experimentally measurable estimate of the quantity $P[\varphi_2|P_1(\tau_1;\varphi_1)]_m$ used in the formulation of the angular EPR paradox (see section 2.3). Obviously there are experimental limitations: a narrow aperture would transmit only little intensity, which would reduce the detection rate in the conditional measurement. Also, spatial light modulators have a finite size and resolution, which limits the ability to distinguish between similar apertures.

3.3 Experimental limitations

Our presented experimental scheme does not take into account a number of phenomena which could possibly have an influence on the measurements. For one misalignment has an affect on the probability distribution, either by introducing a spread in the distribution, or by shifting the distribution by a constant, or by a combination of both (Vaziri et al., 2002; Vastnetsov et al., 2005). These misalignment effects can be characterised by a dimensionless variance introduced by Zambrini & Barnett (2006). But also the exact effect of the angular apertures on the angular position is not known. In the theoretical derivation in section 2.3 we assumed that an angular aperture described by the probability density $P(\tau_1;\varphi_1)$ imprints this probability density on the passing photon exactly. In the next section, where we model the measurements theoretically, we will present a result which questions this assumption. Rectangular probability densities give a conditional wavefunction which has singular derivatives for particular angles. The inferred variance thus tends to infinity and the angular EPR criterion would be fulfilled for any finite measured conditional variance. We therefore study a class of continuously differentiable aperture functions which asymptotically approximate rectangular functions. These functions have a similar form to truncated Gaussians but with a higher power in the exponential and are therefore called trun-

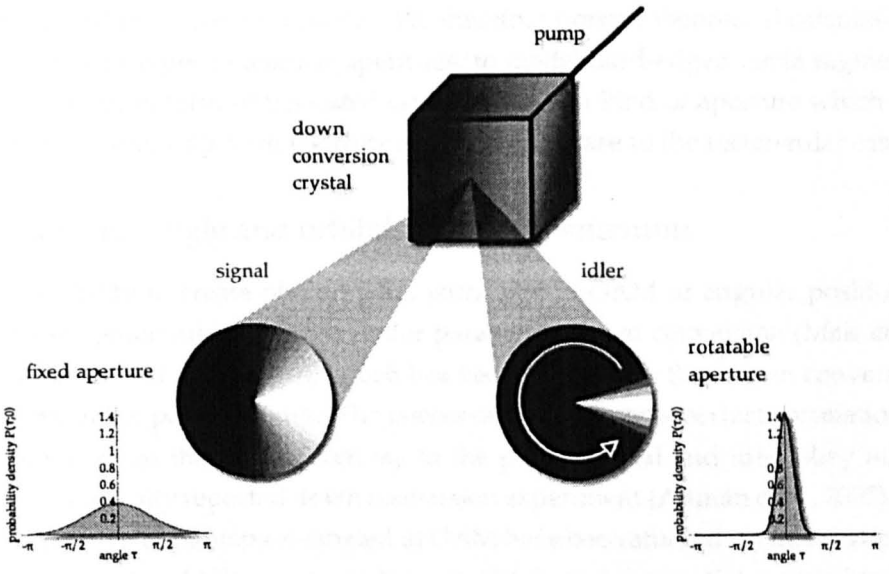

Figure V.5 — Schematic representation of the conditional measurement. A fixed angular aperture in the signal beam sets the condition and a rotatable aperture in the idler beam can be used to measure a conditional probability density for the azimuthal position.

cated super-Gaussians. These functions do not show the same divergent behaviour for the inferred conditional variance although for a very large power they differ only slightly from rectangular functions. It is therefore interesting to ask if diffraction on the hard edges of a circle segment could have an influence on the angular probability density. A related problem lies in the limited resolution of the spatial light modulators. For two apertures represented by very similar probability densities the resolution of the spatial light modulator might not be sufficient to distinguish between them. This is particularly relevant for the case where the power in the truncated super-Gaussian is chosen to be so large that the difference to a rectangular function falls below the resolution of the spatial light modulator. As will be shown in the next section the theoretical prediction for the two cases are very different, whereas the experiments would probably give similar results.

Some of these questions will be discussed in connection with the theoretical results presented in the next section. A comprehensive analysis of further experimental effects will only be possible with a particular experiment and experimental data.

4 Theoretical results

Our criterion (V.19) on page 103 has been devised to be tested on experimental data. For that it does not make any assumption on the correlation for angle and OAM in the two subsystems. In lack of experimental data we have to make an assumption on these correlations, and in the first part of this section we derive the form of the correlations from a simple phase matching argument. With this assumption we are able to model the measurement process and to predict the measurement results. This is more than a mathematical exercise as the theoretical results can be compared to future experimental results, which would provide more insight in both, the experi-

mental data and the theoretical model. We therefore present theoretical calculations for three different types of angular apertures, to model hard-edged circle segments, smooth apertures in form of truncated Gaussians, and a kind of aperture which can be changed continuously from the truncated Gaussian case to the rectangular case.

4.1 Corellations for angle and orbital angular momentum

The possibility to create photon pairs entangled in OAM or angular position is based on the conservation of OAM under parametric down conversion (Mair et al., 2001; Franke-Arnold et al., 2003), which has been studied for thin down conversion crystals and in the paraxial limit. The conservation leads to a perfect correlation in the OAM values so that for a given m_p in the pump, signal and idler obey $m_p = m_1 + m_2$. In a recently reported down conversion experiment (Altman et al., 2005), the spatial correlation of photons entangled in OAM have been studied. In this particular experiment signal and idler cone overlap completely and the spatial correlations are such that for a fixed detector position in the signal, the coincidence pattern in the idler shows two distinct spots equally separated in angle from the position exactly opposite the signal detector on the phase matching ring. The vertex of the separation angle is on the pump axis. In our work we are concerned with a different angle: the azimuthal position of a photon in a beam is measured from the beam axis, i.e. in a down conversion experiment from the signal and idler axis respectively. For non degenerate down conversion crystals these two angles are not identical. The question if they are compatible for the degenerate case is a very interesting one, but it will not be examined in the scope of this thesis. Therefore, for our theoretical modelling, we make use of the phase matching condition for parametric down conversion. The entanglement in OAM and angle is a consequence of this phase matching (Allen et al., 2003).

For a plane wave pump the conservation of transverse momentum requires, that the two-photon wavefunction for the signal and idler has to be of the form $\delta(k_{1,x} + k_{2_x})\delta(k_{1,y} + k_{2,y})$ (Allen et al., 2003). Using a simplified approach one can argue that the transverse spatial correlations in the far field originate from the momentum conservation under parametric down conversion. Identifying transverse momentum components in the near field with spatial coordinates in the far field allows us to write the spatial dependence of the wavefunction in position representation as

$$\delta(x_1 + x_2)\delta(y_1 + y_2) = \frac{1}{\rho_1}\delta(\rho_1 - \rho_2)\delta_{2\pi}(\tau_1 - \tau_2 - \pi), \tag{V.21}$$

where ρ and τ are the radial and azimuthal coordinates and $\delta_{2\pi}$ is the 2π periodic delta function (see App. V.B). From this result we can expect that the azimuthal angles of the photons in the far field obey $\tau_1 = \tau_2 + \pi$, so that the signal and idler photons appear on opposite sides of their respective cones. The correlation in angular momentum follows on writing Eq. (V.21) in terms of its angular Fourier components:

$$\delta(x_1 + x_2)\delta(y_1 + y_2) = \frac{1}{\rho_1}\delta(\rho_1 - \rho_2)\frac{1}{2\pi}\sum_{m=-\infty}^{\infty}(-1)^m \exp(\mathrm{i}m\tau_1)\exp(-\mathrm{i}m\tau_2), \tag{V.22}$$

which is an entangled superposition of states with zero total OAM. A more detailed analysis which considers a specific parametric down conversion process shows that more complicated dependencies of the wavefunction on the azimuthal angles are also possible (Barbosa & Arnaut, 2002).

4.2 Theoretical modelling

To give a quantitative result we model the measurement of $P[P_2(\tau_2;\varphi_2)|P_1(\tau_1;\varphi_1)]$ under the two assumptions that the photon pair is perfectly correlated in angular position and that the angle probability density behind the aperture is exactly given by the function describing the aperture. Under these assumptions the probability to detect a photon behind the aperture $P_2(\tau_2;\varphi_2)$ in the idler given that the paired photon is detected behind the aperture $P_1(\tau_1;\varphi_1)$ in the signal is given by the probability densities in signal and idler after passing the apertures and the two-photon probability density describing the correlated photon pair:

$$P_{\mathrm{dtc}}(\varphi_2) = \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \mathrm{d}\tau_1\mathrm{d}\tau_2 |\psi(\tau_1,\tau_2)|^2 P_1(\tau_1;\varphi_1)P_2(\tau_2;\varphi_2). \tag{V.23}$$

Using the assumption of perfect correlations the probability density will be sharply peaked for $\tau_1 - \tau_2 = \pi$ and in that sense we may use a 2π-periodic δ-function (see App. V.B) to approximate $|\psi(\tau_1,\tau_2)|^2$

$$|\psi(\tau_1,\tau_2)|^2 \approx \delta_{2\pi}(\tau_1 - \tau_2 - \pi). \tag{V.24}$$

Using an implicit periodicity the probability densities P_1 and P_2 can be written as $P_1(\tau_1 - \varphi_1)$ and $P_2(\tau_2 - \varphi_2)$. This notation implies that there exists a fundamental form $P_i(\tau_i)$ which can be displaced by $\varphi_i, i = 1,2$. The displacement of the probability density corresponds to a rotation of the angular aperture. Using the fact that for periodic functions the integration boundaries can be uniformly displaced without altering the integral, we can write for the detection probability

$$P_{\mathrm{dtc}}(\varphi_2) = \int_{-\pi}^{\pi} \mathrm{d}\tau_1 P_1(\tau_1 - \varphi_1)P_2(\tau_1 + \pi - \varphi_2) = [P_1 \circledast P_2](\varphi_1 - \varphi_2 + \pi), \tag{V.25}$$

where we have used the short hand notation from Eq. (V.13) on page 101. For the calculation of this integral it is important to use the periodicity of the probability densities if the argument lies outside the 2π radian interval. The detection probability $P_{\mathrm{dtc}}(\varphi_2)$ is the measured conditional probability $P[P_2(\tau_2;\varphi_2)|P_1(\tau_1;\varphi_1)]_{\mathrm{m}}$.

4.3 Aperture functions

We model the conditional angle measurement for three different apertures. The probability densities or aperture functions are given by rectangular functions, truncated Gaussians and truncated super-Gaussians respectively. Rectangular aperture functions represent hard-edged circle segments. Using these apertures to determine the angle in the range set by the aperture is a simple and clear concept. But the calculated inferred conditional variance tends to infinity and the criterion in Eq. (V.14)

would be fulfilled trivially. Apertures represented by truncated Gaussians have been used in the experimental confirmation of the angular uncertainty relation (Franke-Arnold et al., 2004) and truncated super-Gaussian are able to interpolate between rectangular and truncated Gaussian apertures.

All aperture functions discussed in this section are symmetric, i.e. $P_i(\tau_i - \varphi_i) = P_i(\varphi_i - \tau_i)$ for $i = 1,2$. The overlap integral in Eq. (V.25) on the preceding page can therefore be turned into the convolution of the two probability densities

$$P_{\mathrm{dtc}}(\varphi_2) = \int_{-\pi}^{\pi} P_1(\tau_1 - \varphi_1)P_2(\varphi_2 + \pi - \tau_1) = [P_1 * P_2](\varphi_2 - \varphi_1 + \pi), \qquad \text{(V.26)}$$

where the asterisk denotes the convolution. The theoretically motivated correlations from section 4.1 are uniform and the detection probability depends only on the difference between the orientations of the apertures. It is therefore possible to set $\varphi_1 = \pi$ and to vary φ_2. The calculated quantity $P_{\mathrm{dtc}}(\varphi_2) = P[P_2(\tau_2;\varphi_2)|P_1(\tau_1;\varphi_1)]_{\mathrm{m}}$ is therefore given as a function of φ_2 alone.

Aperture functions have to be periodic by nature, as they represent rotatable angular apertures. But writing the periodic behaviour explicitly in the functional form complicates the equations unnecessarily. We therefore introduce a new variable ξ, which is periodic in τ_i and also contains the orientation of the aperture φ_i for $i = 1,2$:

$$\xi_i(\tau_i;\varphi_i) = [(\tau_i - \varphi_i + \pi) \bmod 2\pi] - \pi \quad i = 1,2. \qquad \text{(V.27)}$$

This variable ranges from $-\pi$ to π, regardless of our choice of the original 2π radian interval $[\theta_0;\theta_0 + 2\pi)$. In the following we will present the aperture functions as a function of ξ_i. Additionally the aperture function will be characterised by a parameter σ_i associated with the width of the respective aperture function.

4.3.1 Rectangular functions

The probability functions describing rectangular apertures can be given in terms of Heaviside step functions $H(\xi)$:

$$P_i(\xi_i) = \frac{1}{2\sigma_i} H(\xi_i + \sigma_i)H(-\xi_i + \sigma_i). \qquad \text{(V.28)}$$

The conditional probability density is given by the convolution of the probability densities for signal and idler according to Eq. (V.26) and is a function of the orientation φ_2 as detailed in the previous section:

$$[P_1 * P_2](\varphi_2) = \frac{1}{\sigma_1\sigma_2} \begin{cases} \sigma_2 & |\varphi_2| < (\sigma_1 - \sigma_2)/2, \\ (\sigma_1 + \sigma_2)/2 - \varphi_2 & (\sigma_1 - \sigma_2)/2 \le |\varphi_2| < (\sigma_1 + \sigma_2)/2, \\ 0 & (\sigma_1 + \sigma_2)/2 \le |\varphi_2|. \end{cases} \qquad \text{(V.29)}$$

Here, we have chosen $\sigma_1 > \sigma_2$, but the convolution is commutative (see App. V.C) and we can interchange σ_1 and σ_2 for the case $\sigma_2 < \sigma_1$. The conditional wavefunction is given by the square root of $[P_1 * P_2](\varphi_2)$ according to Eq. (V.16) on page 103. The phase is determined by minimising the angular momentum variance and in App. V.A it is shown that choosing a constant phase is one possibility to minimise the angular

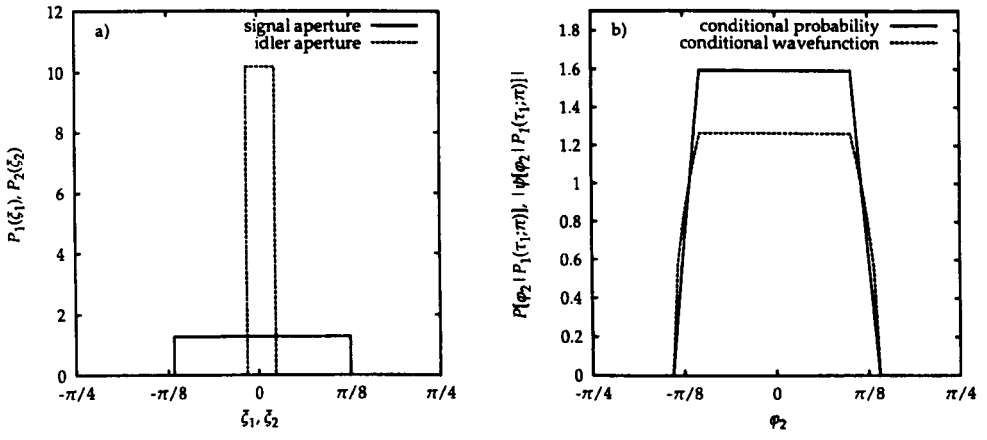

Figure V.6 — In a) the probability densities representing the apertures in the signal and idler beam are plotted for a width in the idler and signal of $\sigma_1 = 1/8\pi$ and $\sigma_2 = 1/128\pi$ respectively. The resulting conditional probability density and wavefunction are shown in b).

momentum variance. Graphs of the rectangular apertures and the resulting conditional probability density and wavefunction are shown in Fig (V.6).

The Fourier integral in Eq. (V.17) on page 103 can be calculated analytically by using the Fresnel sine and cosine integrals $\mathcal{S}_2$ and $\mathcal{C}_2$ (Gradshteyn & Ryzhik, 2000):

$$c[m_2|P_1(\tau_1;\pi)] = (\sigma_2)^{-1/2}\left(\frac{\sin(|m_2|(\sigma_1+\sigma_2)/2)}{|m_2|\sqrt{|m_2|}}\mathcal{C}_2(|m_2|(\sigma_2)) - \frac{\cos(|m_2|(\sigma_1-\sigma_2)/2)}{|m_2|\sqrt{|m_2|}}\mathcal{S}_2(|m_2|(\sigma_2))\right). \tag{V.30}$$

The distribution of conditional probability amplitudes is given in figure (V.7) on the following page for the analytical calculation and a numerical integration. The conditional variance is shown in figure (V.7) over the value of m_2 at which the sum in Eq. (V.18) on page 103 is truncated. The Fresnel integrals tend to $1/2$ for large arguments and the conditional amplitudes $c[m_2|P_1(\tau_1;\pi)]$ thus vary with $|m_2|^{-3}$ for large m_2. This explains the logarithmic dependence on the maximum number of terms in the sum. As a consequence the conditional variance tends to infinity. The reason for this behaviour lies the singular derivative of the conditional wavefunction at $\varphi_2 = (\sigma_1 - \sigma_2)/2$. As pointed out in by Pegg et al. (2005) wavefunctions with discontinuities in the first derivative generally have infinite variance.

4.3.2 Truncated Gaussians

Aperture functions in form of truncated Gaussians have been used in the experiment studying the angular uncertainty principle (Franke-Arnold et al., 2004). In this case the probability densities P_i are given by:

$$P_i(\xi_i) = \frac{\sqrt{\sigma_i}}{\sqrt{\pi}\mathrm{erf}(\pi\sqrt{\sigma_i})}\exp\left(-\sigma_i\xi_i^2\right), \tag{V.31}$$

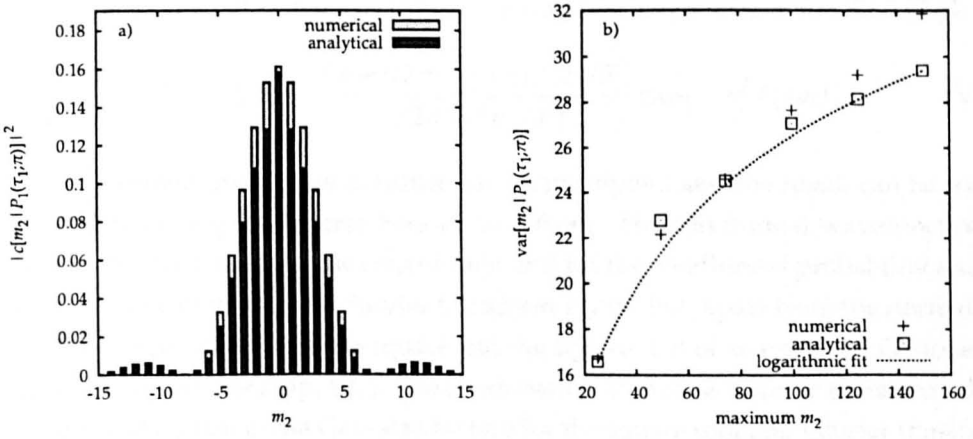

Figure V.7 — a) Numerically and analytically calculated angular momentum distribution for rectangular apertures. In b) the conditional variance is plotted. The logarithmic increase in the variance can be seen.

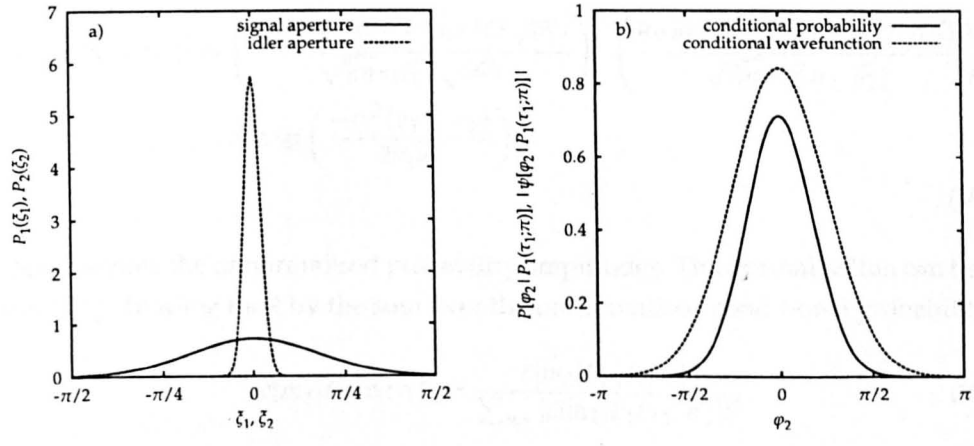

Figure V.8 — In a) the probability densities for truncated Gaussian apertures in the signal and idler beam are plotted for a width in the idler and signal of $\sigma_1 = 16/\pi^2$ and $\sigma_2 = 4096/\pi^2$ respectively. These values for $\sigma_i, i = 1,2$ give a comparable width with the rectangular case. The resulting conditional probability density and wavefunction are shown in b).

where erf denotes the error function. The graph for truncated Gaussian apertures is shown in figure (V.8) for comparable widths. The resulting conditional probability density and the wavefunction are also plotted in figure (V.8).

The conditional probability density $P[\varphi_2|P(\tau_1;\pi)]$ can be calculated analytically using the convolution theorem for the Fourier transform (see App. V.C):

$$\mathcal{F}^{-1}\{\mathcal{F}\{P_1(\xi_i)\}\mathcal{F}\{P_2(\xi_2)\}\} = P_1 * P_2. \tag{V.32}$$

This way avoids the complication of taking the periodic behaviour of the probability densities into account when calculating the convolution integral. The Fourier transform of truncated Gaussians is known from section IV 2.4 on page 59 of the previous

chapter on intelligent states. For the chosen notation here we have for $\mathcal{F}\{P_i(\xi_i)\}$:

$$\mathcal{F}\{P_i(\xi_i)\} = \frac{\mathrm{Re}\,\mathrm{erf}(2\pi\sigma_i + \mathrm{i}m)/(2\sqrt{\sigma_i})}{\sqrt{2\pi}\mathrm{erf}(\pi\sqrt{\sigma_i})} \exp(-m^2/(4\sigma_i)). \tag{V.33}$$

The transformed probability densities can be multiplied and the result can be transformed back using the inverse Fourier transform. The conditional wavefunction is taken as the square root of the convolution and for the conditional probability amplitudes we have to invoke the Fourier transform again. But, apart from the normalisation, the Fourier transform, the square and the square root of an extended Gaussian is again a Gaussian (see App. V.C). The combination of inverse Fourier transform, division of the exponent in the Gaussian by two for the square root and Fourier transform is therefore equivalent to multiplying the exponent in the original Gaussian by two. By using this fact we can give an approximate result for the conditional probability amplitudes, but we have to renormalise the probability distribution for calculating the conditional variance:

$$\begin{aligned}\tilde{c}[m_2|P_1(\tau_1;\pi)] \approx &\left(\frac{\mathrm{Re}\,\mathrm{erf}(2\pi\sigma_1 + \mathrm{i}m)/(2\sqrt{\sigma_1})}{\sqrt{2\pi}\mathrm{erf}(\pi\sqrt{\sigma_1})}\right)^2 \left(\frac{\mathrm{Re}\,\mathrm{erf}(2\pi\sigma_2 + \mathrm{i}m)/(2\sqrt{\sigma_2})}{\sqrt{2\pi}\mathrm{erf}(\pi\sqrt{\sigma_2})}\right)^2 \\ &\times \exp\left(\frac{-m^2(\sigma_1+\sigma_2)}{2\sigma_1\sigma_2}\right).\end{aligned} \tag{V.34}$$

Here $\tilde{c}$ denote the unnormalised probability amplitudes. The normalisation can be restored by dividing the $\tilde{c}$ by the sum over the unnormalised conditional probabilities:

$$c[m_2|P_1(\tau_1;\pi)] = \frac{\tilde{c}[m_2|P_1(\tau_1;\pi)]}{\sum_{m_2}|\tilde{c}[m_2|P_1(\tau_1;\pi)]|^2}. \tag{V.35}$$

For the chosen widths σ_1 and σ_2 the agreement with the numerical solution is excellent as can be seen in Fig. (V.9) on the next page. This is because the single probability densities P_1 and P_2 are sufficiently narrow and decay quickly enough towards the boundaries, so that extending the integral boundaries does not change the integrals significantly. The resulting conditional variance is shown for numerical and analytical calculations in figure (V.9) on the following page. For the truncated Gaussians the conditional variance $\mathrm{var}[m_2|P_1(\tau_1;\pi)]$ converges and does not change with the number of m_2 values taken into account in the summation in Eq. (V.18) on page 103.

4.3.3 Truncated super Gaussians

With truncated super-Gaussian apertures we can gradually go from the rectangular case to the Gaussian. This is achieved by an additional parameter γ in the exponential :

$$P_i(\xi_i) = \frac{\gamma}{\sqrt{\sigma_i}(\Gamma(\frac{1}{2\gamma}) - \Gamma(\frac{1}{2\gamma};(\sigma_i\pi^2)^\gamma)} \exp\left[-(\sigma_i\xi^2)^\gamma\right], \tag{V.36}$$

where $\Gamma(\cdot)$ is the complete Gamma function and $\Gamma(\cdot;\cdot)$ the incomplete Gamma function (Abramowitz & Stegun, 1974). The effect of this parameter on the probability

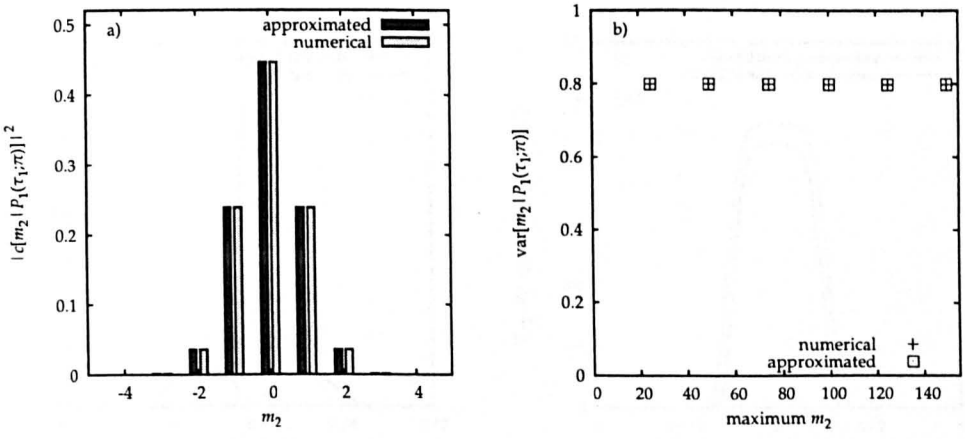

Figure V.9 — a) Numerically and analytically calculated OAM distribution for truncated Gaussians apertures. The analytical solution is approximated by neglecting the effects of the finite interval and treating the truncated Gaussians as extended. This allows us to calculate the probability amplitudes using the convolution theorem and the fact that square roots of Gaussians are again Gaussians. The resulting conditional variance can be seen in b). For the truncated Gaussian apertures the conditional variance converges.

density can be seen in figure (V.10) on the following page.

For the truncated super-Gaussians all calculations have been done numerically and one can see in figure (V.11) on the next page that, even for high values of γ, the probability distribution, $|c[m_2|P_1(\tau 1;\pi)]|^2$, differs substantially from the rectangular case. The dependence of the conditional variance on the maximum value of m_2 and γ is also shown in figure (V.11). For small values of γ the variance converges quickly and does not depend much on the number of m_2 values in the sum. For higher values the variance converges more slowly. Even for the highest value of $\gamma = 80$ a logarithmic increase, as for the rectangular case, cannot be seen.

4.4 Discussion of theoretical results

The theoretical modelling of a conditional angle measurement leads to a conditional variance significantly different from zero. This is the result from all angular apertures. However, for rectangular apertures the inferred conditional variance tends to infinity, which would allow a demonstration of the angular EPR paradox according to our criterion in Eq. (V.14) on page 101 for all measured conditional variances $var[m_2|m_1]_m$ with finite value. This is certainly not a sensible result and the reason for this divergent behaviour lies in the discontinuous derivative of the conditional wavefunction. But of course the inferred conditional variance will always depend on the chosen apertures, not only because the condition is set by passing an aperture, but also because an aperture is used to measure the angle. In the theoretical analysis we assumed that the measuring aperture P_2 is very narrow, and in that case it is the aperture in the signal P_1 which determines mostly the broadness of the conditional probability density $P[P_2(\tau_2;\varphi_2)|P_1(\tau_1;\varphi_1)]$. Choosing a broad aperture P_1 will in all generality give a broad conditional probability density and the angular position in the conditional wavefunction will not be very localised. This in turn causes the OAM

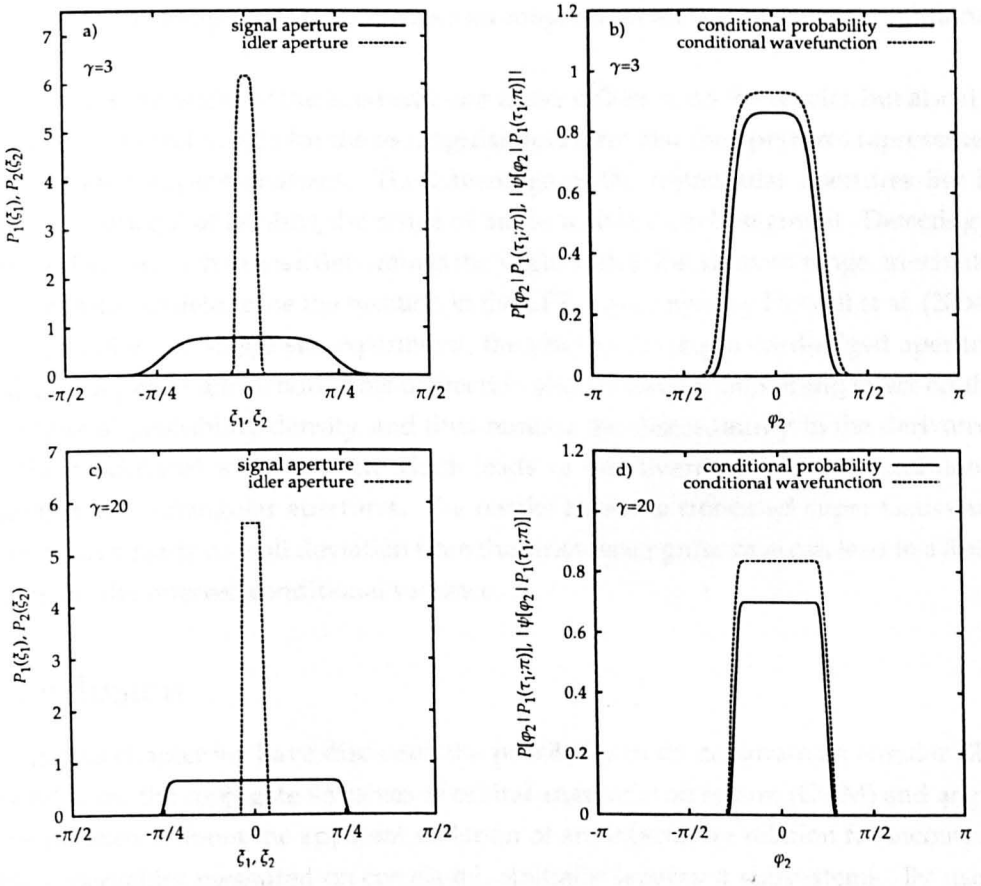

Figure V.10 — a) Probability densities for apertures in form of truncated super-Gaussians in the signal and idler beam ($\sigma_1 = 16/\pi^2$, $\sigma_2 = 4096/\pi^2$, $\gamma = 3$) b) Resulting conditional probability density and wavefunction. c) Probability densities representing the apertures in the signal and idler beam [σ_1 and σ_2 as in a), $\gamma = 20$]. *d*) Resulting conditional probability density and wavefunction.

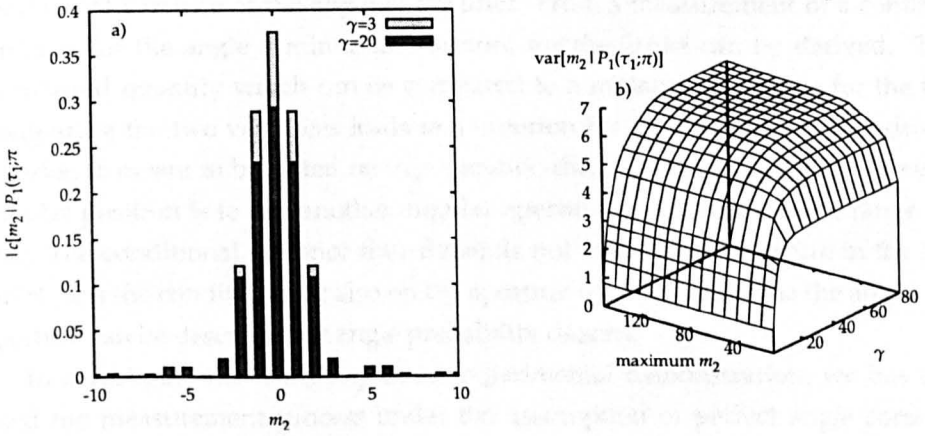

Figure V.11 — a) Numerically calculated conditional angular momentum distribution for truncated super-Gaussian apertures. The distribution is plotted for two different values of γ. In b) the dependence of the conditional variance on maximum value of m_2 and γ is shown.

distribution to be very narrow and thus the inferred conditional variance to be small. Choosing a broad aperture sets therefore a stronger limit for the measured conditional variance.

Not only the width of the aperture has a direct influence on the results, but also the form. This is most visible for the rectangular apertures and the apertures represented by truncated super-Gaussians. The advantage of the rectangular apertures lies in the clear concept of limiting the range of angle within a circle segment. Detecting a photon behind such a mask determines the angle within the allowed range, much like slits are used to determine the position in the EPR experiment by Howell et al. (2004). And just like in a single slit experiment, the photon passing a hard-edged aperture will be subject to diffraction. This diffraction should have a smoothing effect on the conditional probability density, and thus remove the discontinuity in the derivative of the conditional wavefunction which leads to the diverging inferred conditional variance for rectangular apertures. The results from the truncated super-Gaussians show that already a small deviation from the strict rectangular case can lead to a finite value for the inferred conditional variance.

5 Conclusion

In this chapter we have discussed the possibility to demonstrate an angular EPR paradox for the conjugate variables of orbital angular momentum (OAM) and angle. The paradox is about the apparent violation of an uncertainty relation for incompatible observables measured on correlated, spatially separated subsystems. By using photon pairs entangled in OAM and in angle these subsystems can be realised in an optical experiment. We have found a testable criterion for an angular EPR paradox, which takes experimental indeterminacies into account. For that we have reformulated the EPR paradox using conditional variances, that is variances of observables from one subsystem given a preset outcome on the other subsystem. To devise a conditional variance for the angular position we make use of angular apertures. The detection of a photon behind such an aperture in the signal sets a condition in the conditional variance of the angle in the idler. From a measurement of a conditional variance for the angle a minimum variance for the OAM can be derived. This is an inferred quantity which can be compared to a measured variance for the OAM. Comparing the two variances leads to a criterion for an angular EPR paradox. The criterion is meant to be tested on experimental data. One possibility to measure the angular position is to use another angular aperture which restricts the range of angles. The conditional variance thus depends not only on the aperture in the signal, which sets the condition, but also on the aperture used to determine the angle. These aperture can be described by angle probability density.

To investigate the feasibility of an experimental demonstration, we have modelled the measurement process under the assumption of perfect angle correlation. Also, angular apertures to set the condition or to determine the angle were assumed to impose their probability characteristics exactly on transmitted photons. Under these assumptions we have studied different classes of aperture functions for the final conditional OAM variance. Rectangular functions lead to a divergent conditional

OAM variance, which does not set any lower bound for the OAM correlations. Truncated Gaussians result in a quickly converging variance significantly different from zero. Truncated super-Gaussians, which can be varied from the rectangular case to the truncated Gaussians, provide an aperture which leads to a convergent variance. For the case of a truncated super-Gaussian which is very similar to a rectangular function, one can combine the clear concept of measuring the angle with an angular aperture with a finite conditional variance for the OAM. The conditional variances obtained from the theoretical modeling show that an angular EPR paradox can be demonstrated. Given the current state of experiments we expect an implementation of our criterion to be able to demonstrate the EPR paradox for OAM and azimuthal position.

V.A Minimisation

Calculating the conditional wavefunction from the conditional probability leaves the phase $\alpha(\varphi_2)$ undetermined. We find that if the variance is calculated for a constant phase α then the minimum variance $\min \mathrm{var}[m_2|P_1(\tau_1;\varphi_1)]_i$ is obtained. To show this we assume the conditional wavefunction to be of the form

$$\psi[\varphi_2|P_1(\tau_1;\varphi_1)] = A\exp[\mathrm{i}\alpha(\varphi_2)], \tag{V.37}$$

where $A = \sqrt[4]{|\psi[\varphi_2|P_1(\tau_1;\varphi_1)]|^2}$ is a positive, real function, which is periodic in φ_2. Applying the OAM operator to the wavefunction yields

$$\begin{aligned}\hat{L}_z\psi[\varphi_2|P_1(\tau_1;\varphi_1)] &= -\mathrm{i}\hbar\psi'[\varphi_2|P_1(\varphi_1)],\\ &= -\mathrm{i}\hbar A'(\varphi_2)\exp[\mathrm{i}\alpha(\varphi_2)] + \hbar A(\varphi_2)\alpha'(\varphi_2)\exp[\mathrm{i}\alpha(\varphi_2)],\end{aligned} \tag{V.38}$$

where the primes denote derivatives with respect to φ_2. The periodicity of the conditional probability density and the conditional wavefunction requires that both A and α are periodic in φ_2. The expectation value of $\hat{L}_z$ is thus given by

$$\langle\hat{L}_z\rangle = \hbar\int_{-\pi}^{\pi}\mathrm{d}\varphi_2\alpha'(\varphi_2)P[\varphi_2|P_1(\tau_1;\varphi_1)], \tag{V.39}$$

where we used the fact that $A^2(\varphi_2) = P[\varphi_2|P_1(\tau_1;\varphi_1)]$. The first term in the second line of the expression for the expectation value vanishes on integration as the integrand AA' can be written as the derivative of $A^2/2$ with respect to φ_2. Owing to the periodicity of A the integral evaluates to zero. The OAM variance requires the expectation value of $\hat{L}_z^2$:

$$\begin{aligned}\hat{L}_z^2\psi[\varphi_2|P_1(\tau_1;\varphi_1)] = &-\hbar^2A''(\varphi_2)\exp[\mathrm{i}\alpha(\varphi_2)] - 2\mathrm{i}\hbar^2A'(\varphi_2)\alpha(\varphi_2)\exp[\mathrm{i}\alpha(\varphi_2)]\\ &-\mathrm{i}\hbar^2A(\varphi_2)\alpha(\varphi_2)\exp[\mathrm{i}\alpha(\varphi_2)] + \hbar^2A(\varphi_2)[\alpha'(\varphi_2)]^2\exp[\mathrm{i}\alpha(\varphi_2)].\end{aligned} \tag{V.40}$$

On integration the middle two terms vanishes as the combined integrand can be written as derivative of $A^2a'/2$ with respect to φ_2. The expectation value of L_z^2 is thus

given by:

$$\langle \hat{L}_z^2 \rangle = -\hbar^2 \int_{-\pi}^{\pi} \mathrm{d}\varphi_2 A(\varphi_2) A''(\varphi_2) + \hbar^2 \int_{-\pi}^{\pi} \mathrm{d}\varphi_2 [\alpha'(\varphi_2)]^2 P[\varphi_2 | P_1(\tau_1; \varphi_1)]. \quad \text{(V.41)}$$

The variance of L_z^2 is given by $\langle \hat{L}_z^2 \rangle - \langle \hat{L}_z \rangle^2$. The second term in Eq. (V.41) together with the square of the expectation value of L_z from Eq. (V.39) on the preceding page is the variance of α'. The variance of L_z is thus given by the variance of L_z with a constant α and the variance of α':

$$\mathrm{var}\, L_z = [\mathrm{var}\, L_z]_{\alpha=\mathrm{const}} + \hbar^2 \mathrm{var}\, \alpha'. \quad \text{(V.42)}$$

The variance of α' is either positive or zero, if $\alpha' = \mathrm{const}$. The minimum of $\mathrm{var} L_z$ is therefore given by the minimum of $[\mathrm{var}\, L_z]_{\alpha=\mathrm{const}}$. Variances are calculated for a given state, and by specifying the probability density, the wavefunction for this state is determined apart from the phase. From Eq. (V.42) follows that a constant phase gives the minimum variance for the OAM.

V.B Periodic δ-function

The periodic δ function can be written as a sum of conventional δ functions (Peřinová et al., 1998):

$$\delta_{2\pi}(\theta) = \sum_{k=-\infty}^{\infty} \delta(\theta - 2\pi k). \quad \text{(V.43)}$$

From this definition follows the property for an infinitely extended integral

$$\int_{-\infty}^{\infty} \mathrm{d}\theta \delta_{2\pi}(\theta) f(\theta) = \sum_{k=-\infty}^{\infty} f(2\pi k). \quad \text{(V.44)}$$

For a finite integral only a finite number of terms appear in the sum. An interesting case occurs for a integration interval with a length of an integer multiple of 2π. Depending on the exact position of the integral boundaries, one δ function from the sum in Eq. (V.43) is fully contained or two δ functions are at the boundaries. This is the case for the expectation value of the commutator between number and phase operator (Pegg & Barnett, 1997) or OAM and angle operator (Barnett & Pegg, 1990) and causes the expectation value to depend on the exact choice of the integration interval (see Ch. III).

V.C Periodic convolution

The convolution of two periodic functions v, w on a 2π radian interval is defined as

$$[v * w](\theta) = \int_{-\pi}^{\pi} \mathrm{d}\chi v(\chi) w(\theta - \chi). \quad \text{(V.45)}$$

For the evaluation of this integral it is important to make use of the periodicity of the functions. This is of particular importance if the function does not decay quickly towards the boundaries. We can study the effect on the example of truncated Gaussians.

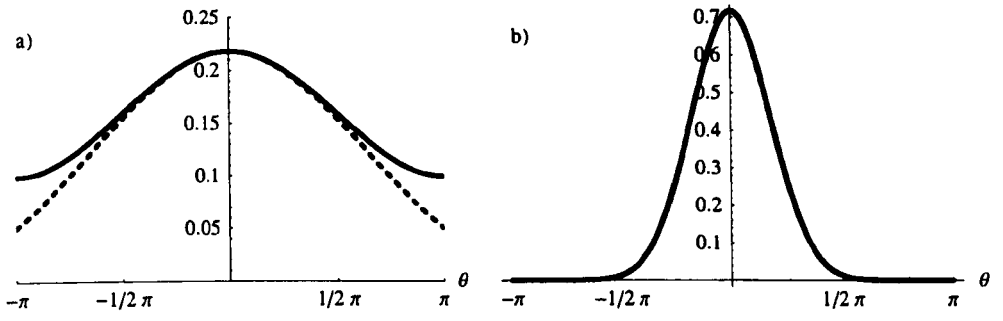

Figure V.12 — Graphs of the convolution of truncated Gaussians. a) shows the difference between the evaluation of the convolution according to Eq. (V.46) (solid lines) and an integration which disregards the periodic behaviour (dashed lines) for the values of $\sigma_1 = 2/\pi^2$ and $\sigma_2 = 4/\pi^2$. b) shows the difference for $\sigma_1 = 16/\pi^2$ and $\sigma_2 = 4096/\pi^2$. Solid and dashed lines are inseparable.

With the definition in Eq. (V.31) on page 111 the periodicity is given implicitly by using the variable ξ_i. Using the periodicity explicitly the integration in Eq. (V.45) on the preceding page has to be evaluated in the following way:

$$[v * w](\theta) = \int_{-\pi}^{\theta-\pi} \mathrm{d}\chi v(\chi) w(\theta - \chi - 2\pi) + \int_{\theta-\pi}^{\pi} \mathrm{d}\chi v(\chi) w(\theta - \chi) \quad \theta > 0, \quad \text{(V.46a)}$$

$$[v * w](\theta) = \int_{-\pi}^{\pi+\theta} \mathrm{d}\chi v(\chi) w(\theta - \chi) + \int_{\pi+\theta}^{\pi} \mathrm{d}\chi v(\chi) w(\theta - \chi + 2\pi) \quad \theta < 0. \quad \text{(V.46b)}$$

The effect can be seen in figure (V.12). The solid line is the evaluation of the integral according to Eq. (V.46), while the dashed line is an integration in which the truncated Gaussians are treated as extended, and no amendments have been made for the case where the argument of w is outside $[-\pi, \pi)$. The graph shows that for narrowly peaked functions which decay sufficiently quickly towards to the boundaries the correct convolution is excellently approximated by evaluating the integral without taking the periodic behaviour into account.

In the derivation of the criterion for the angular EPR paradox we have introduced the operation $v \circledast w$ in Eq. (V.13) on page 101 and Eq. (V.25) on page 109:

$$[v \circledast w](\theta) = \int_{-\pi}^{\pi} \mathrm{d}\chi v(\chi) w(\chi - \theta), \quad \text{(V.47)}$$

which is an overlap integral for v and w, where w is shifted by θ. From the definition in Eq. (V.45) on the previous page it is clear that for a symmetric w the shifted integral is equivalent to a convolution.

The Fourier transform and the convolution for periodic function obey the convolution theorem (Stephenson & Radmore, 1993):

$$\mathcal{F}\{v * w\} = A^{-1} \bar{v}(s) \bar{w}(s), \quad \text{(V.48)}$$

where $\bar{v}(s)$ and $\bar{w}(s)$ are the Fourier transforms of $v(\theta)$ and $w(\theta)$. The constant factor

A depends on the choice of multiplicative factors in the Fourier transform:

$$\mathcal{F}\{v(\theta)\} = \bar{v}(s) = A\int_{-\pi}^{\pi} \mathrm{d}\theta v(\theta)\exp(-\mathrm{i}s\theta), \tag{V.49a}$$

$$\mathcal{F}^{-1}\{\bar{v}(s)\} = v(\theta) = \frac{1}{2\pi A}\sum_{s}\bar{v}(s)\exp(\mathrm{i}s\theta). \tag{V.49b}$$

For the Fourier relation between angle and OAM we have chosen $A = 1/\sqrt{2\pi}$ (cf. Eq. (V.17) on page 103). With this notation it is straightforward to proof the relation in Eq. (V.48) on the previous page:

$$\begin{aligned}
\mathcal{F}\{v * w\} &= A\int_{-\pi}^{\pi}\mathrm{d}\theta\int_{-\pi}^{\pi}\mathrm{d}\zeta v(\zeta)w(\theta-\zeta)\exp(-\mathrm{i}s\theta),\\
&= A\int_{-\pi}^{\pi}\mathrm{d}\theta\int_{-\pi}^{\pi}\mathrm{d}\zeta v(\zeta)\frac{1}{2\pi A}\sum_{t}\bar{w}(t)\exp[\mathrm{i}t(\theta-\zeta)]\exp(-\mathrm{i}s\theta),\\
&= \frac{1}{2\pi A}\sum_{t}\bar{v}(t)\bar{w}(t)\int_{-\pi}^{\pi}\mathrm{d}\theta\exp[\mathrm{i}(t-s)\theta],\\
&= \frac{1}{A}\bar{v}(s)\bar{w}(s).
\end{aligned} \tag{V.50}$$

Here, we have used an integral representation for $\delta(t-s)$ in form of

$$\int_{-\pi}^{\pi}\mathrm{d}\theta\exp[\mathrm{i}(t-s)\theta] = 2\pi\delta(t-s). \tag{V.51}$$

This shows the convolution theorem for the Fourier transform on a 2π radian interval.

The convolution theorem finds its use in calculating the conditional probability amplitudes $c[m_2|P_1(\tau_1;\pi)]$ for truncated Gaussian apertures. From Eqs. (V.17) on page 103 and (V.26) on page 110 it follows that the conditional probability amplitudes are given by the Fourier transform of the square root of $P_1 * P_2$:

$$c[m_2|P_1(\tau_1;\pi)] = \mathcal{F}\left\{\sqrt{P_1 * P_2}\right\}. \tag{V.52}$$

The Fourier transform of $P_1 * P_2$ would simply be given by the product of the Fourier transforms $\bar{P}_1$ and $\bar{P}_2$. For extended Gaussians of the form $\exp(-\sigma m^2)$ the inverse Fourier transform yields another Gaussian with a reciprocal width $\exp[-\xi^2/(4\sigma)]$. The factor of four depends on our particular choice of the form of the Gaussian. Taking the square root would result in $\exp[-\xi^2/(8\sigma)]$ and the Fourier transform of that would give $\exp(-2\sigma m^2)$. For an extended Gaussian the process of inverse Fourier transform, square root and Fourier transform results in multiplying the original exponent by two:

$$\begin{array}{ccc}
\mathcal{F}\{P_1 * P_2\} = \bar{P}_1\bar{P}_2 & \xrightarrow{\mathcal{F}^{-1}} & P[\varphi_2|P_1(\tau_1;\pi)],\\
(\cdot)^2\big\downarrow & & \big\downarrow\sqrt{\cdot}\\
c[m_2|P_1(\tau_1;\pi)] & \xleftarrow{\mathcal{F}} & \psi[\varphi_2|P_1(\tau_1;\pi)].
\end{array} \tag{V.53}$$

The resulting distribution of OAM probabilities will in general not be normalised. For truncated Gaussians this approach works only approximately. For narrow truncated Gaussians the effects introduced by the truncation at π and $-\pi$ are very small, this

is why this approach gives results in excellent agreement with a direct numerical calculation.

Bibliography

Abramowitz, M. & Stegun, I. S. (1974), *Handbook of Mathematical Functions*, Dover Publications, Inc., Mineola, New York. Reprint. Originally published: National Bureau of Standards, corrected edition, 1964. ISBN: 0-486-61272-4.

Allen, L., Barnett, S. M. & Padgett, M. J. (2003), *Optical Angular Momentum*, Institute of Physics Publishing, Ltd., Bristol. ISBN: 0-7503-0901-6.

Altman, A. R., Köprülü, K. G., Corndorf, E., Kumar, P. & Barbosa, G. A. (2005), 'Quantum imaging of nonlocal spatial correlations induced by orbital angular momentum', *Physical Review Letters* **94**(12), 123601. DOI: 10.1103/PhysRevLett.94.123601.

Aspect, A., Dalibard, J. & Roger, G. (1982*a*), 'Experimental Test of Bell's Inequalities Using Time-Varying Analyzers', *Physical Review Letters* **49**(25), 1804 – 1806. DOI: 10.1103/PhysRevLett.49.1804.

Aspect, A., Grangier, P. & Roger, G. (1982*b*), 'Experimental Realization of Einstein-Podolsky-Rosen-Bohm Gedankenexperiment: A New Violation of Bell's Inequalities', *Physical Review Letters* **49**(2), 91 – 94. DOI: 10.1103/PhysRevLett.49.91.

Barbosa, G. A. & Arnaut, H. H. (2002), 'Twin photons with angular-momentum entanglement: Phase matching', *Physical Review A* **65**, 053801. DOI: 10.1103/PhysRevA.65.053801.

Barnett, S. M. & Pegg, D. T. (1990), 'Quantum theory of rotation angles', *Physical Review A* **41**(7), 3427–3435. DOI: 10.1103/PhysRevA.41.3427.

Bell, J. S. (1964), 'On the Einstein-Podolsky-Rosen paradox', *Physics* **1**, 195–200. Reprinted in (Bell, 2004, Paper 2).

Bell, J. S. (1966), 'On the problem of hidden variables in quantum mechanics', *Reviews of Modern Physics* **38**(3), 447 – 452. Reprinted in (Bell, 2004, Paper 1). DOI: 10.1103/RevModPhys.38.447.

Bohm, D. & Aharonov, Y. (1957), 'Discussion of Experimental Proof for the Paradox of Einstein, Rosen, and Podolsky', *Physical Review* **108**(4), 1070 – 1076. DOI: 10.1103/PhysRev.108.1070.

Einstein, A., Podolsky, B. & Rosen, N. (1935), 'Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?', *Physical Review* **47**, 777–780. DOI: http://link.aps.org/abstract/PR/v47/p777 DOI: 10.1103/PhysRev.47.777.

Franke-Arnold, S., Barnett, S. M., Padgett, M. J. & Allen, L. (2003), 'Two-photon entanglement of orbital angular momentum states', *Physical Review A* **65**, 033823. DOI: 10.1103/PhysRevA.65.033823.

Franke-Arnold, S., Barnett, S. M., Yao, E., Leach, J., Courtial, J. & Padgett, M. (2004), 'Uncertainty principle for angular position and angular momentum', *New Journal of Physics* **6**, 103. DOI: 10.1088/1367-2630/6/1/103.

Götte, J. B., Franke-Arnold, S. & Barnett, S. M. (2006), 'Angular EPR paradox', *Journal of Modern Optics* **53**(5 - 6), 627 – 645.

Gradshteyn, I. S. & Ryzhik, I. M. (2000), *Tables of Integrals, Series, and Products*, 6th edn, Academic Press, Inc., San Diego. ISBN: 0-12-294757-6.

Gühne, O. (2004), 'Charcterizing Entanglement via Uncertainty Relations', *Physical Review Letters* **92**(11), 117903. DOI: 10.1103/PhysRevLett.92.117903.

Heisenberg, W. (1927), 'Über den anschaulichen Inhalt der quantentheoretischen Kinematik and Mechanik', *Zeitschrift für Physik* **43**, 172.

Hofmann, H. F. & Takeuchi, S. (2003), 'Violation of local uncertainty relations as a signature of entanglement', *Physical Review A* **68**, 032103. DOI: 10.1103/PhysRevA.68.032103.

Howell, J. C., Bennink, R. S., Bentley, S. J. & Boyd, R. W. (2004), 'Realization of the Einstein-Podolsky-Rosen Paradox Using Momentum- and Position-Entangles Photons from Spontaneous Parametric Down Conversion', *Physical Review Letters* **92**(21), 210403. DOI: 10.1103/PhysRevLett.92.210403.

Jauch, J. M. & Piron, C. (1963), 'Can Hidden Variables be excluded in Quantum Mechanics?', *Helvetica Acta Physics* **36**, 827–837.

Kwiat, P. G., Mattle, K., Weinfurter, H. & Zeilinger, A. (1995), 'New High-Intensity Source of Polarization-Entangled Photon Pairs', *Physical Review Letters* **75**(24), 4337–4341. DOI: 10.1103/PhysRevLett.75.4337.

Leach, J., Padgett, M. J., Barnett, S. M., Franke-Arnold, S. & Courtial, J. (2002), 'Measuring the Orbital angular momentum of a single Photon', *Physical Review Letters* **88**, 257901. DOI: 10.1103/PhysRevLett.88.257901.

Louisell, W. H., Yariv, A. & Siegman, A. E. (1961), 'Quantum Fluctuations and Noise in Parametric Processes. I.', *Physical Review* **124**(6), 1646–1654. DOI: 10.1103/PhysRev.124.1646.

Mair, A. E., Vaziri, A., Weihs, G. & Zeilinger, A. (2001), 'Entanglement of the orbital angular momentum states of photons', *Nature* **412**, 313–316. Reprinted in (Allen et al., 2003, Paper 8.1).

Nielsen, M. A. & Chuang, I. L. (2000), *Quantum Computation and Quantum Information*, Cambridge University Press, Cambridge.

Ou, Z. Y., Pereira, S. F. & Kimble, H. J. (1982), 'Realization of the Einstein-Podolsky-Rosen paradox for continuous variables', *Physical Review Letters* **68**(25), 3663 – 3666. DOI: 10.1103/PhysRevLett.68.3663.

Pegg, D. T. & Barnett, S. M. (1997), 'Tutorial review: Quantum optical phase', *Journal of Modern Optics* **44**(2), 225–264.

Pegg, D. T., Barnett, S. M., Zambrini, R., Franke-Arnold, S. & Padgett, M. (2005), 'Minimum uncertainty states of angular momentum and angular position', *New Journal of Physics* **7**, 62. DOI: 10.1088/1376-2630/7/1/062.

Peřinová, V., Lukš, A. & Peřina, J. (1998), *Phase in Optics*, Vol. 15 of *World Scientific Series in Contemporary Chemical Physics*, World Scientific, Singapore.

Reid, M. (1997), 'Macroscopic elements of reality and the Einstein– Podolsky–Rosen paradox', *Quantum and Semiclassical Optics* **9**, 489–499.

Reid, M. D. (1989), 'Demonstration of the Einstein–Podolsky–Rosen paradox using nondegenerate parametric amplification', *Physical Review A* **40**(2), 913–923. DOI: 10.1103/PhysRevA.40.913.

Robertson, H. P. (1929), 'The Uncertainty Principle', *Physical Review* **34**, 163–164.

Rubin, M. H., Klyshko, D. N., Shih, Y. H. & Sergienko, A. V. (1994), 'Theory of two-photon entanglement in type-II optical parametric down-conversion', *Physical Review A* **50**(6), 5122–5133. DOI: 10.1103/PhysRevA.50.5122.

Stephenson, G. & Radmore, P. M. (1993), *Advanced Mathematical Methods for Engineering and Science Students*, Cambrige University Press, Cambridge. ISBN: 0-521-36860-X.

Vastnetsov, M. V., Pas'ko, V. A. & Soskin, M. S. (2005), 'Analysis of orbital angular momentum of a misaligned optical beam', *New Journal of Physics*. DOI: 10.1088/1367-2630/7/1/046.

Vaziri, A., Weihs, G. & Zeilinger, A. (2002), 'Superpositions of the orbital angular momentum for applications in quantum experiments', *Journal of Optics B* **4**, S47–S51. DOI: 10.1088/1464-4266/4/2/367.

von Neumann, J. (1932), *Mathematische Grundlagen der Quantenmechanik*, Springer Verlag, Berlin. English translation: (von Neumann, 1955).

Wei, H., Xue, X., Leach, J., Padgett, M. J., Barnett, S. M., Franke-Arnold, S., Yao, E. & Courtial, J. (2003), 'Simplified measurement of the orbital angular momentum of single photons', *Optics Communication* **223**(1-3), 117–122. DOI: doi:10.1016/S0030-4018(03)01619-5.

Weihs, G., Jennwein, T. & Simon, C. (1998), 'Violation of Bell's Inequality under Strict Einstein Locality Conditions', *Physical Review Letters* **81**(23), 5039 – 5043. DOI: 10.1103/PhysRevLett.81.5039.

Zambrini, R. & Barnett, S. M. (2006), 'Quasi-Intrinsic Angular Momentum and the Measurement of Its Spectrum', *Physical Review Letters* **96**, 113901. DOI: 10.1103/PhysRevLett.96.113901.

VI Chapter

Fractional Orbital Angular Momentum

Light beams carrying integer values of orbital angular momentum have a phase structure which is stable under propagation. The phase structure contains a number of intertwined helices corresponding to the charge of the optical vortex imprinted by an optical device. The optical devices which generate light beams with orbital angular momentum can also be designed for fractional phase steps. In place of a vortex with integer charge these phase steps imprint a vortex with fractional charge which cannot propagate. Instead, the fractional vortex splits up in a number of vortices with charge ± 1. Light beams emerging from a fractional phase step are not stable under propagation and show a line of low light intensity. For odd half-integer phase steps additional vortices are created in this region. The theoretical description of these fractional modes is facilitated by a generalisation of the quantum mechanical description of orbital angular momentum to non-integer values. The generalised theory is applied to simulate the propagation of fractional modes in the paraxial and nonparaxial regime.

1 Introduction

By investigating special states connected to the angular uncertainty relation and by formulating an angular EPR paradox orbital angular momentum (OAM) and angle have been established as conjugate variables. The connection between angle and OAM is visible in the wavefunction of angular momentum eigenstates which is given by $\exp(-im\varphi)$ in the unnormalised angle representation. A wavefunction of this form has a phase singularity at the origin. Beams with OAM can be generated by phase sensitive optical components such as spiral phase plates. On propagation the phase singularity turns into an optical vortex of charge m. The azimuthal phase structure of beams with OAM consists of $|m|$ intertwined helices and the handedness of the helices is given by the sign of m. The phase structure of light beams emerging from integer phase steps is stable under propagation apart from dilation. In particular the mean value of the OAM in the beam remains unchanged and there is a direct correspondence between the charge of the optical vortex and the OAM in the light beam. It is possible, however, to design these phase sensitive optical components for fractional phase steps (Oemrawsingh et al., 2004*b*). In this case the generating optical device imprints a phase structure of $\exp(iM\phi)$, where M is not restricted to an integer value. Light beams emerging from a fractional phase step are not stable under propagation and have a more complex phase structure and intensity profile. The intensity profile shows a line of darkness along the direction of the phase discontinuity of the optical

device. In the phase profile the fractional charge vortex on the axis splits up in a number of charge ± 1 vortices which are situated in the central area of the beam but not directly on the axis. For half-integer phase steps a series of additional optical vortices with alternating charge is created in the line of low intensity (Berry, 2004; Leach et al., 2004*b*). Moreover, only for integer and half-integer values of M the mean value of the OAM in the beam is equal to step height but not in general.

The dependence of the fractional modes on the orientation of the phase discontinuity finds interesting applications in the study of entangled photon pairs carrying OAM (Oemrawsingh et al., 2004*a*). Half-integer spiral phase plates imprint a π radians phase shift which results in a line of zero intensity along the orientation of the step. Additionally, the quantum states which describe the light emerging from two half-integer spiral phase plates are orthogonal if their relative orientation is π radians. This opens the possibility to use the orientation of the phase plates as measurement settings in a Bell-type inequality (Aiello et al., 2005; Oemrawsingh et al., 2005). Until now, however, the theoretical description of such fractional OAM modes has been based on modelling the generating optical device. In this chapter we present a fundamental theory of fractional OAM by generalising the quantum theory of rotation angles (Barnett & Pegg, 1990) to fractional values of M. In light of recent applications of integer OAM in quantum key distribution (Spedalieri, 2006) and the conversion of spin angular momentum to OAM in an optical medium (Marrucci et al., 2006) a rigid theory of fractional OAM is important for possible applications of fractional OAM in quantum communication. We compare our generalised theory with recent experimental and theoretical work. With our approach we are able to calculate the propagation of light modes with fractional OAM in the paraxial and non-paraxial regime. The work in this chapter is an original contribution to the literature and has been accepted for publication.

2 Generating fractional orbital angular momentum

In this section we present two methods of generating light with fractional OAM. Both methods have been experimentally implemented and theoretically studied. We review the corresponding theory in both cases and reproduce important results to which we will refer later in this chapter. These two methods have been used for different aspects of fractional OAM. While the half-integer spiral phase plates have been used to examine two-photon entanglement (Aiello et al., 2005; Oemrawsingh et al., 2005), holograms with fractional step height have been used to study the vortex structure of non-integer vortex beams (Berry, 2004; Leach et al., 2004*b*).

2.1 Spiral phase plates

The production of high-quality half-integer spiral phase plates has first been reported by Oemrawsingh et al. (2004*b*). The term half-integer spiral phase plates refers to odd half-integer values of the imprinted vortex charge which is given by a closed

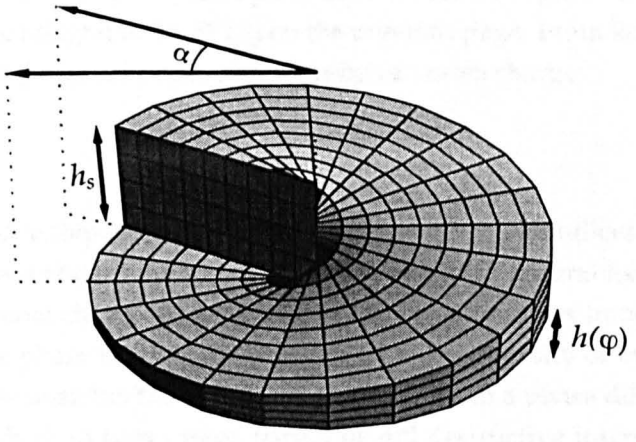

Figure VI.1 — Sketch of a spiral phase plate without a base ($h_0 = 0$). The step is oriented at the azimuthal angle $\varphi = \alpha$.

path integral around the phase singularity (Nye, 1999):

$$Q = \frac{1}{2\pi} \oint \mathrm{d}\chi = \frac{1}{2\pi} \oint \mathrm{d}\mathbf{s} \cdot \nabla\chi, \tag{VI.1}$$

where $\mathbf{s}$ is the line element and χ is the phase of the field. A spiral phase plate is a transparent plate of refractive index n with a base and a height on top of the base, which is proportional to the azimuthal angle φ (see figure (VI.1)):

$$h(\varphi) = h_s \frac{\varphi}{2\pi} + h_0. \tag{VI.2}$$

Here, h_s is the step height and h_0 the base height of the device. The phase change is determined by the ratio of the optical path length l and the wavelength of the light λ:

$$\Delta\chi = 2\pi \frac{l}{\lambda}. \tag{VI.3}$$

The optical path length for a path c is given by the integral over the refractive index along that path (Born & Wolf, 1999):

$$l = \int_c n(s)\mathrm{d}s. \tag{VI.4}$$

Owing to the spiral structure of the phase plate the optical path length depends on the azimuthal angle φ:

$$l(\varphi) = \int_{h_0}^{h(\varphi)} n\mathrm{d}z + \int_{h(\varphi)}^{h_0+h_s} n_0\mathrm{d}z = (n - n_0)h_s \frac{\varphi}{2\pi}. \tag{VI.5}$$

If the spiral phase plate is illuminated by a plane wave or an area of a beam with constant phase the change in the phase is solely due to the difference in the optical path length. The base of a spiral phase plate causes a uniform phase shift and does not contribute to the integral in Eq. (VI.1) on the previous page. From Eqs. (VI.1) and (VI.3) it follows that the spiral phase plate imprints a vortex charge

$$Q = h_s \frac{(n - n_0)}{\lambda}. \tag{VI.6}$$

By choosing a suitable step height or difference in the refractive indices it is possible to design spiral phase plates with a fractional step height. In general for spiral phase plates with a fractional charge the orientation of the step becomes important, as the discontinuity in the phase leads to a visible line of lower intensity or even complete darkness. This is because the fractional step height leads to a phase difference other than 2π radians which in turn causes partial or full destructive interference along the discontinuity. Oemrawsingh et al. (2004*b*) showed that the spiral phase plate is a unitary optical element by illuminating a spiral phase plate and its optical conjugate phase plate with a Gaussian beam. The conjugate spiral phase plate is a second spiral phase plate which is identical to the first, but has opposite helicity.

The unitarity of the spiral phase plate is also reflected in the spiral phase plate operator introduced by Oemrawsingh et al. (2004*a*). The operator is defined by its action on an eigenstate of the OAM operator $\hat{L}_z$:

$$\hat{S}(\alpha, M')|m\rangle = |(M' + m)(\alpha)\rangle, \tag{VI.7}$$

where $M' = m' + \mu'$ and $m' \in \mathbb{Z}$ with $\mu' \in [0, 1)$. The corresponding wave function in the angle representation shows a discontinuity at the angle α:

$$\langle\varphi|\hat{S}(\alpha, M')|m\rangle = \frac{\exp(im\varphi)}{\sqrt{2\pi}} \exp[iM'(\varphi - \alpha)] \times \begin{cases} \exp(iM'2\pi) & 0 < \varphi < \alpha, \\ 1 & \alpha < \varphi < 2\pi. \end{cases} \tag{VI.8}$$

On using the expression for the wavefunction in Eq. (VI.8), one can calculate the overlap for two states with equal fractional OAM M and with a relative orientation β. The overlap probability is given by:

$$|\langle M(0)|M(\beta)\rangle|^2 = \left(1 - \frac{\beta}{\pi}\right)^2 \sin^2(\mu\pi) + \cos^2(\mu\pi), \tag{VI.9}$$

where we have split $M = m + \mu$ into an integer part m and a fractional part μ. For $\mu = 1/2$ and $\beta = \pi$ the overlap is zero and the states are orthogonal. It is this property that renders half-integer spiral phase plates interesting for studying two photon entanglement. Choosing two orthogonal orientations provides an orthonormal, two-dimensional basis which describes a dichotomic subspace of the OAM space (Aiello et al., 2005). The integer part of the OAM does not enter the overlap probability and by using half integer spiral phase plates with two orthogonal orientations the infinite number of dimensions in the OAM space is projected onto two dimensions. This opens the possibility to use the orientations of the spiral phase plates in a Clauser-Horne-type inequality as proposed by Aiello et al. (2005). This is, however, a differ-

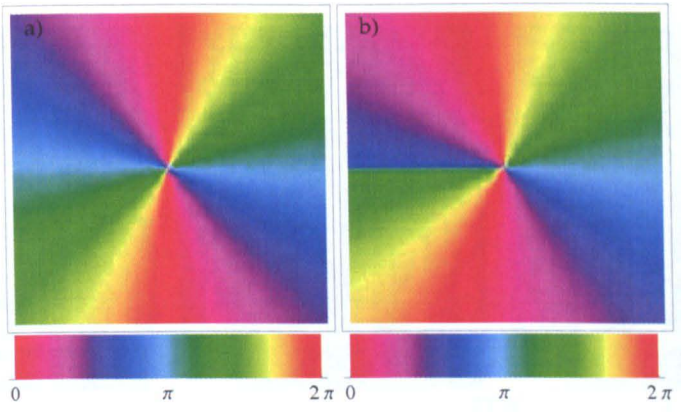

Figure VI.2 — Phase profile for holograms with integer and fractional phase steps. a) shows the phase profile of $\exp(im\varphi)$ with $m = -2$. b) shows the phase profile of $\exp(iM\varphi)$ with $M = -1.7$. An additional discontinuity is visible in the latter, where the phase jumps by less than 2π. The orientation of this discontinuity corresponds to the orientation of the step in the spiral phase plates.

ent setup from a continuous variation of the relative orientation β which has been studied by Oemrawsingh et al. (2005).

2.2 Fractional phase steps with holograms

The use of spatial light modulators to produce holograms with fractional phase steps has first been reported by Leach et al. (2004*b*). When compared to an integer phase step hologram, a hologram with a fractional phase step shows an additional radial discontinuity, where the phase jumps by an amount smaller than 2π (see figure (VI.2)). The orientation of this discontinuity corresponds to the orientation of the step in the spiral phase plates. To displace different orders of diffracted beams the phase profile of a blazed grating is added to the phase ramp via a modulo 2π radian addition. This creates the characteristic fork dislocation (see Ch. II). For a fractional phase step the number of prongs is given by the modulus of the nearest integer. The additional radial phase discontinuity is clearly visible in the fork dislocation pattern (see figure (VI.3) on the following page). Changing the orientation of the phase step in the spiral phase plates simply corresponds to rotating the whole spiral phase plate. For holograms, the orientation of the phase discontinuity and the direction of the grating can be changed independently. Rotating the whole hologram including the grating would not only change the orientation of the phase discontinuity but also the direction of the angular displacement for the different order of diffraction. Leaving the alignment of the grating vertically while changing the orientation of the radial phase discontinuity results in a hologram different from those shown in figure (VI.3) on the next page with a different angle between radial discontinuity and the primary direction of the grating. With help of these holograms with fractional phase steps Leach et al. (2004*b*) confirmed theoretical predictions of Berry (2004) about the vortex structure emerging from fractional phase steps. The analysis of the optical vortices is twofold; first the field amplitude behind a phase component with an integer phase step is written as a superposition of plane waves, then

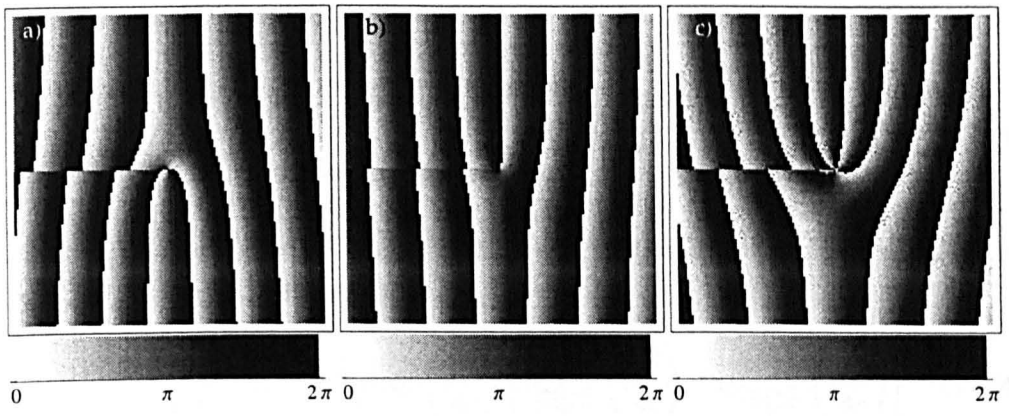

Figure VI.3 — Phase profiles for holograms with fractional phase steps. The phase profile includes a blazed grating to displace different order of diffraction. a) shows the phase profile for $M = -1.7$. The fork dislocation has two prongs as -2 is the nearest integer. b) for $M = 1.1$ there is only one prong and the phase profile is only slightly different from $m = 1$ (see figure (II.5) on page 30). c) for $M = 3.5$ three prongs are clearly visible and a fourth one is forming. Here, the radial discontinuity corresponds to a π phase jump.

the wave emerging from a fractional phase step is expressed as a superposition of waves emerging from integer phase steps. In the following we are reproducing some important results from Berry (2004), to which we will refer later in this section when we apply our fundamental theory to the propagation of fractional modes.

The superposition of the propagating wave into plane waves with a transverse wavevector $\mathbf{K}$ can be written as

$$\Psi_m(\mathbf{r}) = \iint_{\mathbf{K}} \mathrm{d}\mathbf{K}\, a_m(\mathbf{K}) \exp\left[\mathrm{i}\left(\mathbf{R}\cdot\mathbf{K} + z\sqrt{1-\kappa^2}\right)\right], \tag{VI.10}$$

where $\kappa = |\mathbf{K}|$ and the overall wavenumber k has been set to $k = 1$. Waves with $\kappa > 1$ are evanescent, but they are included in the superposition, because they originate from the on-axis singularity and they are needed to describe the singularity correctly (Roux, 2003). The Fourier coefficients $a_m(\mathbf{K})$ are given by

$$a_m(\mathbf{k}) = \frac{1}{4\pi^2} \iint_{\mathbf{R}} \mathrm{d}\mathbf{R} \exp[\mathrm{i}(\mathbf{K}\cdot\mathbf{R} + m\varphi) = \frac{|m|(-1)^{|m|}}{2\pi\kappa^2} \exp(\mathrm{i}m\varphi_{\mathbf{k}}), \tag{VI.11}$$

where $\varphi_{\mathbf{k}}$ the azimuthal angle in a polar representation of $\mathbf{K}$. The propagating wave can be given in the exact form and in the paraxial approximation by choosing the respective propagation factor (see section II.4 on page 16). For the exact wave this leads to

$$\Psi_m(\mathbf{r}) = \exp(\mathrm{i}m\varphi)|m| \int_0^\infty \frac{\mathrm{d}\kappa}{\kappa} J_{|m|}(\kappa\rho) \exp(\mathrm{i}z\sqrt{1-K^2}), \tag{VI.12}$$

where J_n denotes the Bessel function of the first kind (Abramowitz & Stegun, 1974). In the paraxial approximation $\sqrt{1-\kappa^2}$ is replaced by $1-\kappa^2/2$ and the form of the

wave changes to

$$\begin{aligned}\psi_m(\mathbf{r}) &= \exp[\mathrm{i}(m\varphi + z)]|m| \int_0^\infty \frac{\mathrm{d}\kappa}{\kappa} J_{|m|}(\kappa\rho)\exp(-\mathrm{i}z\kappa^2/2), \\ &= \exp[\mathrm{i}(m\varphi + z)P_m(\frac{\rho}{\sqrt{z}}).\end{aligned} \tag{VI.13}$$

On manipulating the integral in Eq. (VI.13) we can give an expression for the P_m in closed form in terms of Bessel functions (see App. VI.D):

$$P_m(\frac{\rho}{\sqrt{z}}) = \sqrt{\frac{\pi}{8}}(-\mathrm{i})^{|m|/2}\frac{\rho}{\sqrt{z}}\exp\left(\mathrm{i}\frac{\rho^2}{4z}\right)\left[J_{\frac{|m|-1}{2}}\left(\mathrm{i}\frac{\rho^2}{4z}\right) - \mathrm{i}J_{\frac{|m|+1}{2}}\left(\mathrm{i}\frac{\rho^2}{4z}\right)\right]. \tag{VI.14}$$

For even values of m the expression for the P_m contains Bessel functions of half-integer order, which can be expressed in terms of finite trigonometric sums (Abramowitz & Stegun, 1974). It is worth noting that the wave in the paraxial integration (see Eq. (VI.13)) does not contain evanescent waves.

Waves emerging from fractional phase steps are treated as a superposition of waves from integer phase steps. The superposition is determined by the Fourier series:

$$\exp(\mathrm{i}M\varphi) = \frac{\exp(\mathrm{i}M\pi)\sin(M\pi)}{\pi}\sum_{m=-\infty}^{\infty}\frac{\exp(\mathrm{i}m\varphi)}{M-m}, \tag{VI.15}$$

where $M = m + \mu$, with $m \in \mathbb{Z}$ and $\mu \in [0,1)$. The exact wave emerging from a fractional phase step is thus given by:

$$\Psi_M(\mathbf{r}) = \frac{\exp(\mathrm{i}M\pi)\sin(M\pi)}{\pi}\sum_{m=-\infty}^{\infty}\frac{\Psi_m(\mathbf{r})}{M-m}. \tag{VI.16}$$

The corresponding wave in the paraxial approximations is:

$$\psi_M(\mathbf{r}) = \frac{\exp[\mathrm{i}(z + M\pi)\sin(M\pi)}{\pi}\sum_{m=-\infty}^{\infty}\frac{\exp(\mathrm{i}m\varphi)P_m(\rho/\sqrt{z})}{M-m}. \tag{VI.17}$$

One quantity of interest is the OAM mean $\bar{M}$ of light emerging from a fractional phase step. For the initial wave $\exp(\mathrm{i}M\varphi)$ in Eq. (VI.15) the OAM mean can be calculated using the contour integration method to evaluate sums (see App. A). We present here the final result which has been reported by Berry (see (Leach et al., 2004*b*)):

$$\bar{M} = M - \frac{\sin(2M\pi)}{2\pi}. \tag{VI.18}$$

We will derive an identical result in the next section where we present a fundamental theory for fractional OAM based on the generalisation of the quantum theory of rotation angles by Barnett & Pegg (1990). The OAM mean of the propagated wave will be examined in a later section to compare the measured and simulated results from Leach et al. (2004*b*) with our numerical calculation of the propagated wave emerging from a fractional phase step.

3 Generalised theory for fractional OAM states

In this section we present the quantum theory of fractional OAM. From a generalised expression for OAM states we are able to calculate the overlap of two such states. This expression can be specialised to the case, where the fractional value M is identical for both states and the relative orientation β can be varied over a 2π radian interval. These results will be compared with results from Oemrawsingh et al. (2005). From the general expression for the overlap amplitudes we derive the decomposition of fractional OAM states into the OAM basis. It is shown that states with fractional OAM form an alternative basis for the OAM state space, if the fractional part and the orientation are kept fixed.

3.1 Construction of fractional OAM states

The properties of the angle operator are rigorously derived in the arbitrarily large, yet finite state space of $2L+1$ dimensions (Barnett & Pegg, 1990) which has been introduced in chapter III. This space is spanned by the OAM states $|m\rangle$ with m ranging from $-L, -L+1, \ldots, L$. Accordingly, the 2π radian interval $[\theta_0, \theta_0 + 2\pi)$ is spanned by $2L+1$ orthonormal angle states $|\theta_n\rangle$ with $\theta_n = \theta_0 + 2\pi n/(2L+1)$. Here, θ_0 determines the starting point of the interval and with it a particular angle operator $\hat{\phi}_\theta$. Only after physical results have been calculated within this state space L is allowed to tend to infinity, which recovers the result of an infinite but countable number of basis states for the OAM and a dense set of angle states within a 2π radian interval. The analysis of fractional OAM is conducted in this finite dimensional space. After we have calculated expectation values or probabilities we give final results in the limit of $L \to \infty$.

A quantum state with fractional OAM is denoted by $|M\rangle$, where $M = m + \mu$ with m as the integer part and $\mu \in [0, 1)$ as the fractional part. Choosing 'fractional' over 'non-integer' follows the terminology in earlier work, but μ can take on every real number between zero and one and not just rational numbers. This distinction, however, is not relevant for all practical purposes. Such a state can be created using holograms (Leach et al., 2004*b*) or spiral-phase plates (Oemrawsingh et al., 2004*b*). The angle decomposition of $|M\rangle$ is given by:

$$\begin{aligned} |M\rangle &= (2L+1)^{1/2} \sum_{n=0}^{2L} \exp(\mathrm{i}M\theta_n)|\theta_n\rangle, \\ &= (2L+1)^{1/2} \sum_{n=0}^{2L} \exp(\mathrm{i}m\theta_n)\exp(\mathrm{i}\mu\theta_n)|\theta_n\rangle. \end{aligned} \tag{VI.19}$$

In general $\exp(\mathrm{i}\mu\theta_n)$ will be a multivalued function, but a common way to render a multi-valued function single-valued is the introduction of a branch cut (Conway, 1978), which restricts the range of the function. Usually the branch cut is chosen to be along the negative real axis in the complex plane, such that the restricted function has no discontinuities (see App. VI.A). In our case we would like to choose the position of the branch cut α, that is the position of the discontinuity in the function $\exp(\mathrm{i}\mu\theta_n)$, independently of θ_0, the starting angle of the 2π radian interval. For that we intro-

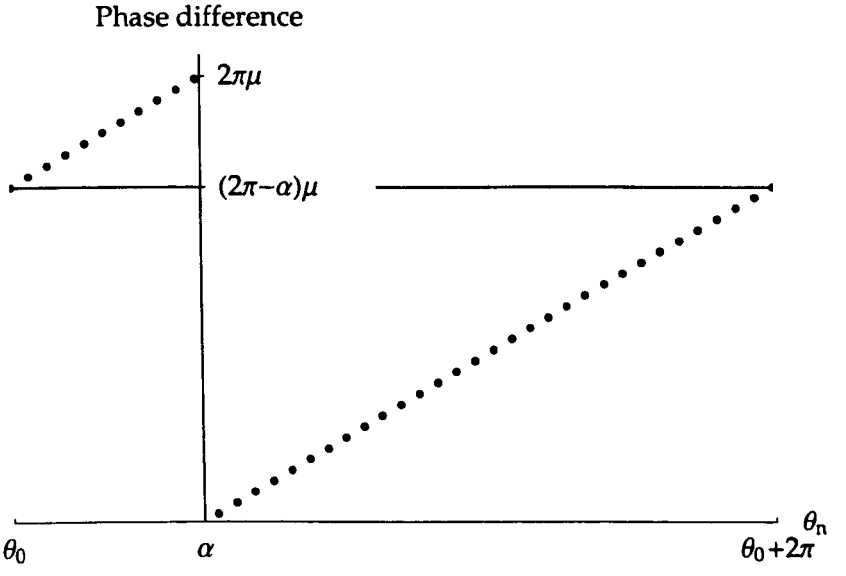

Figure VI.4 — Graph of the phase discontinuity in fractional orbital angular momentum states. The position of the discontinuity α is chosen independently from the starting angle of the 2π radian interval, on which the angle is defined. In the state space Ψ the angle is a discrete variable with $2L+1$ angle eigenstates $|\theta_n\rangle, n = 0, \ldots, 2L$ (see Ch. III).

duce an integer-valued function $f_\alpha(\theta_n)$ which takes on the values 1 and 0, depending whether θ_n is smaller or greater than α:

$$f_\alpha(\theta_n) : \{\theta_n\}_{n=0,\ldots,2L} \longrightarrow \{0,1\}, \quad \theta_n \longmapsto \begin{cases} 1 & \theta_0 < \theta_n < \theta_0 + \alpha, \\ 0 & \theta_0 + \alpha < \theta_n < \theta_0 + 2\pi, \end{cases} \tag{VI.20}$$

where $\alpha \in [0, 2\pi)$. The effect can be seen in figure (VI.4). We should note that according to the mathematical definition we do not create a branch of the multivalued function, as we explicitly include the discontinuity. With this construction the fractional state $|M\rangle$ can be written as

$$\begin{aligned} |M(\alpha)\rangle &= \frac{\exp(-\mathrm{i}\mu\alpha)}{\sqrt{2L+1}} \sum_{n=0}^{2L} \exp(\mathrm{i}m\theta_n) \exp[\mathrm{i}\mu(\theta_n + 2\pi f_\alpha(\theta_n))] |\theta_n\rangle, \\ &= \frac{\exp(-\mathrm{i}\mu\alpha)}{\sqrt{2L+1}} \sum_{n=0}^{2L} \exp(\mathrm{i}M\theta_n) \exp[\mathrm{i}2\pi\mu f_\alpha(\theta_n)] |\theta_n\rangle. \end{aligned} \tag{VI.21}$$

The factor $\exp(-\mathrm{i}\mu\alpha)$ ensures that the fractional OAM state remains in the same 'branch' independent of the value of α. The fractional OAM state $|M(\alpha)\rangle$ depends now on the orientation of the discontinuity α.

3.2 Overlap of fractional OAM states

The overlap of two states with fractional values of OAM and two different orientations is given by $\langle M'(\alpha')|M(\alpha)\rangle$. It is possible to introduce a unitary operator

$\hat{U}_m(\beta), \beta \in [0, 2\pi)$ which formally rotates the position of the discontinuity by writing

$$\hat{U}_m(\beta)|M'(\alpha)\rangle = \exp[\mathrm{i}(m - m')\beta]|M'(\alpha \oplus \beta)\rangle. \tag{VI.22}$$

The addition $\alpha \oplus \beta$ is required to yield a result in the range $[0, 2\pi)$. This can be achieved by a modulo addition in the form $\alpha \oplus \beta = \alpha + \beta \mod 2\pi$. The set of operators $\{\hat{U}_m(\beta)\}_{\beta \in [0,2\pi)}$ forms a one-parameter group under multiplication with the continuous parameter β (see App. VI.B). On applying the operator $\hat{U}_{m'}(-\alpha')$ to the state $|M(\alpha)\rangle$ and $\hat{U}^\dagger_{m'}(-\alpha')$ to $\langle M'(\alpha')|$ the orientation of the states in the overlap is altered such that one orientation is zero and the other is equal to the difference of the original orientations. This yields

$$\begin{aligned} \langle M'(\alpha')|M(\alpha)\rangle &= \langle M'(\alpha')|\hat{U}^\dagger_{m'}(-\alpha')\hat{U}_{m'}(-\alpha')|M(\alpha)\rangle, \\ &= \exp[(\mathrm{i}(m - m')\alpha']\langle M'(0)|M(\beta)\rangle, \end{aligned} \tag{VI.23}$$

where we have set $\beta = \alpha \oplus (-\alpha')$. The moduli of the rotated and unrotated overlaps are the same so we proceed in calculating $\langle M'(0)|M(\beta)\rangle$ only. With the definition of the fractional OAM state in Eq. (VI.21) on the previous page the overlap is given by

$$\langle M'(0)|M(\beta)\rangle = \frac{\exp(-\mathrm{i}\mu\beta)}{2L+1} \sum_{n=0}^{2L} \exp[\mathrm{i}(M - M')\theta_n] \exp[\mathrm{i}\mu 2\pi f_\beta(\theta_n)], \tag{VI.24}$$

where we have used the fact that $\{|\theta_n\rangle\}_{n=0,\ldots,2L}$ is an orthonormal set (see Eq. (III.10)). The sum in the above equation (VI.24) can be split in two parts by introducing an index N with $\theta_N < \beta \leq \theta_{N+1}$ corresponding to the different cases in the definition of f_β in Eq. (VI.20) on the previous page:

$$\begin{aligned} \langle M'(0)|M(\beta)\rangle = \frac{\exp(-\mathrm{i}\mu\beta)}{2L+1} \Bigg[&\exp(\mathrm{i}\mu 2\pi) \sum_{n=0}^{N} \exp[\mathrm{i}(M - M')\theta_n] \\ &+ \sum_{n=N+1}^{2L} \exp[\mathrm{i}(M - M')\theta_n] \Bigg]. \end{aligned} \tag{VI.25}$$

On substituting the expression for the angle θ_n in Eq. (III.9) on page 43 it is possible to evaluate the sums using geometric progression (see App. VI.C):

$$\begin{aligned} \langle M'(0)|M(\beta)\rangle = \exp(-\mathrm{i}\mu\beta) &\frac{\exp[\mathrm{i}(M - M')\theta_0]}{2L+1} \\ &\times \Bigg[\exp(\mathrm{i}\mu 2\pi) \frac{1 - \exp[\mathrm{i}(M - M')\frac{2\pi(N+1)}{2L+1}]}{1 - \exp[\mathrm{i}(M - M')\frac{2\pi}{2L+1}]} \\ &+ \frac{\exp[\mathrm{i}(M - M')\frac{2\pi(N+1)}{2L+1}] - \exp[\mathrm{i}(M - M')2\pi]}{1 - \exp[\mathrm{i}(M - M')\frac{2\pi}{2L+1}]} \Bigg]. \end{aligned} \tag{VI.26}$$

The overlap is a physical result and we can take the limit of $L \to \infty$ by expanding the exponentials in the denominators. In this limit the set of angles $\{\theta_n\}_{n=0,\ldots,2L}$ becomes

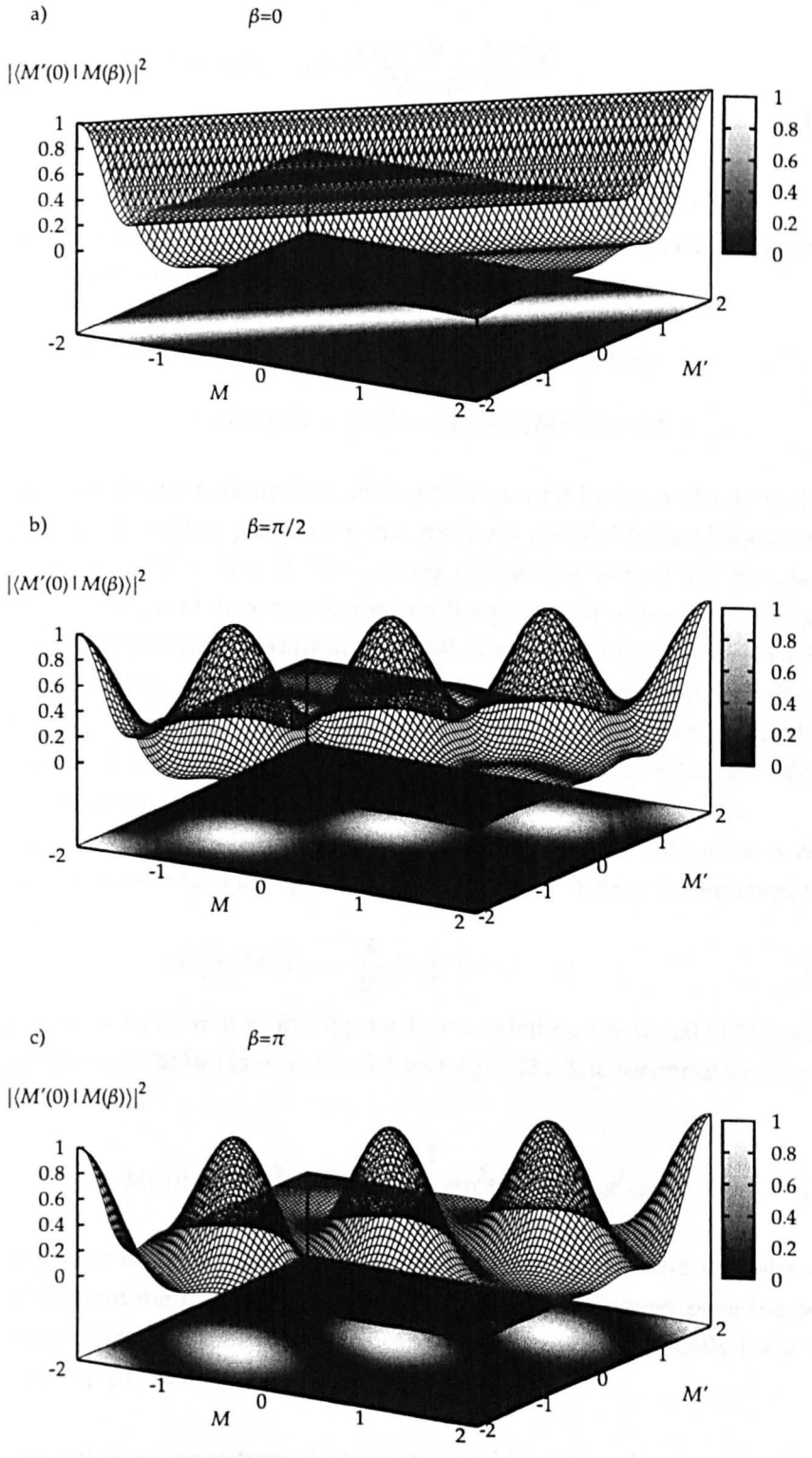

Figure VI.5 — Plot of the overlap probability for two general OAM states $|M\rangle$ and $|M'\rangle$ for three different values of the difference in the orientation β. a) $\beta = 0$: For $M - M' = K \in \mathbb{Z}$ the overlap probability is independent of the particular values M and M'. b) $\beta = \pi/2$: The overlap probability generally depends on the fractional values, but does not reach zero for $M = M'$. c) $\beta = \pi$: For $M = M'$ and $\mu = \mu' = 1/2$, the overlap probability reaches zero.

dense and one can write for $\lim_{L\to\infty} \frac{2\pi(N+1)}{2L+1} = \lim_{L\to\infty} \theta_{N+1} = \beta$:

$$\begin{aligned}\langle M'(0)|M(\beta)\rangle = \exp(-i\mu\beta)\frac{i\exp[i(M-M')\theta_0]}{(M-M')2\pi} \\ \times \{\exp(i\mu 2\pi)\left[1-\exp[i(M-M')\beta]\right] \\ + \left[\exp[i(M-M')\beta] - \exp[i(M-M')2\pi]\right]\}.\end{aligned} \tag{VI.27}$$

From this expression it is possible to calculate the modulus square of the overlap $\langle M'(0)|M(\beta)\rangle$ with help of several trigonometric identities:

$$\begin{aligned}|\langle M'(0)|M(\beta)\rangle|^2 = \frac{1}{(M-M')^2\pi^2}\Big[\sin^2(M\pi) + \sin^2(M'\pi) \\ - 2\cos[(M-M')(\pi-\beta)]\sin(M\pi)\sin(M'\pi)\Big].\end{aligned} \tag{VI.28}$$

The overlap probability is plotted for different values of β in figure (VI.5) on the previous page. For $\beta = 0$ the plot shows that there are parallel lines of equal overlap probability for $M - M' = K \in \mathbb{R}$. The overlap probability along these lines is given by $\sin^2(\pi K)/(\pi^2 K^2)$ and does not depend on the particular values of μ and μ'. For $\beta = \pi/2$ there are primary maxima at $M = M' = m \in \mathbb{Z}$ and secondary maxima at $M - M' = \pm 2$, with $\mu, \mu' = 1/2$. There are also further local maxima, at half-integer pairs of M, M'. Along the diagonal $M = M$ the overlap probability is significantly lower at $\mu, \mu' = 1/2$ but not equal to zero. This is in contrast to the case of $\beta = \pi$ where the overlap probability reaches zero for $M = M'$ with $\mu, \mu' = 1/2$.

This can be seem more clearly by considering the particular case of $M = M'$. On expanding the exponentials in Eq. (VI.27) for $M - M' \to 0$, the overlap amplitude is given by

$$\langle M(0)|M(\beta)\rangle = \frac{\beta}{2\pi}(\exp(i\mu 2\pi) - 1) + 1. \tag{VI.29}$$

This result can be compared to the approach of modelling fractional OAM states by Oemrawsingh et al. (2004*a*) (see section 2.1 on page 125). The resulting overlap probabilities are identical:

$$|\langle M(0)|M(\beta)\rangle|^2 = \left(1 - \frac{\beta}{\pi}\right)^2 \sin^2(\mu\pi) + \cos^2(\mu\pi). \tag{VI.30}$$

The overlap probability only depends on the fractional part μ and the relative orientation β and not the integer part m. In figure (VI.6) on the next page the overlap probability is plotted as a function of β for different values of μ. Only for $\mu = 1/2$ does the overlap probability reach zero for $\beta = \pi$.

3.3 Orbital angular momentum distribution of fractional states

The expression for the general overlap of two fractional angular momentum states can be specialised to yield the coefficients in an integer OAM decomposition of fractional OAM. By choosing $M' = m' \in \mathbb{Z}$ we can evaluate the probability amplitudes

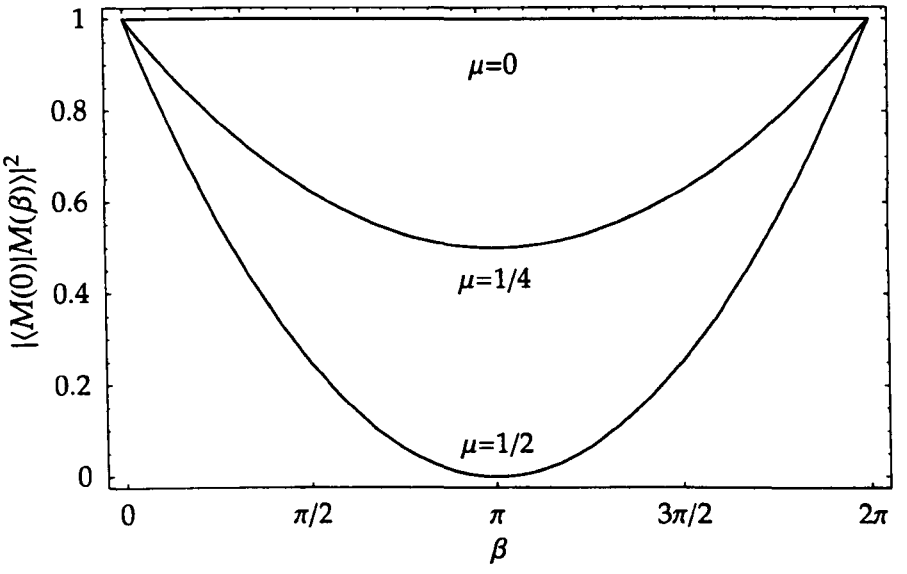

Figure VI.6 — Plot of the overlap probability for two fractional orbital angular momentum states with $M = M'$ as a function of the relative orientation β. The overlap probability reaches zero for $\mu = 1/2$ and $\beta = \pi$ (cf. Oemrawsingh et al. (2004*a*) and Eq. (VI.9) on page 127).

$c_{m'}[M(\beta)] = \langle m'(0)|M(\beta)\rangle$ from Eq. (VI.27) on the preceding page:

$$c_{m'}[M(\beta)] = \langle m'|M(\beta)\rangle = \exp(-\mathrm{i}\mu\beta)\frac{\mathrm{i}\exp[\mathrm{i}(M-m')\theta_0]}{2\pi(M-m')} \times \left[\exp[\mathrm{i}(M-m')\beta]\left(1-\exp(\mathrm{i}\mu 2\pi)\right)\right]. \tag{VI.31}$$

Only the complex argument of the probability amplitudes $c_{m'}[M(\beta)]$ depends on the relative orientation; the probabilities $P_{m'}(M)$, given by the modulus square of the probability amplitudes, are independent of β:

$$P_{m'}(M) = |c_{m'}[M(\beta)]|^2 = \frac{1-\cos(2\mu\pi)}{(M-m')^2 2\pi^2} = \frac{\sin^2(\mu\pi)}{(M-m')^2\pi^2}. \tag{VI.32}$$

Integer OAM states form an orthonormal basis, that is they obey $\langle m|m'\rangle = \delta_{mm'}$ for $m, m' \in \mathbb{Z}$. This particular result is recovered for the OAM probabilities $P_{m'}(M)$ from Eq. (VI.31). For $M = m \in \mathbb{Z}$ and $m \neq m'$ the fractional part μ is zero resulting in a vanishing OAM probability. For $M = m'$, we can determine the value of $P_{m'}(m')$ by the limiting procedure:

$$P_{m'}(m) = \lim_{\mu\to 0} P_{m'}(m+\mu) = \frac{1}{2\pi^2}\lim_{\mu\to 0}\frac{1-\cos(2\mu\pi)}{(m+\mu-m')^2} = \delta_{mm'}. \tag{VI.33}$$

For integer values of M the OAM distribution is thus singular, consisting of a single non-vanishing probability at $M = m'$. For fractional values of M, however, the probabilities are peaked around the nearest integer to M, as can be seen in figure (VI.7) on the following page.

Owing to the completeness of the OAM basis states $|m'\rangle$, the sum of the the prob-

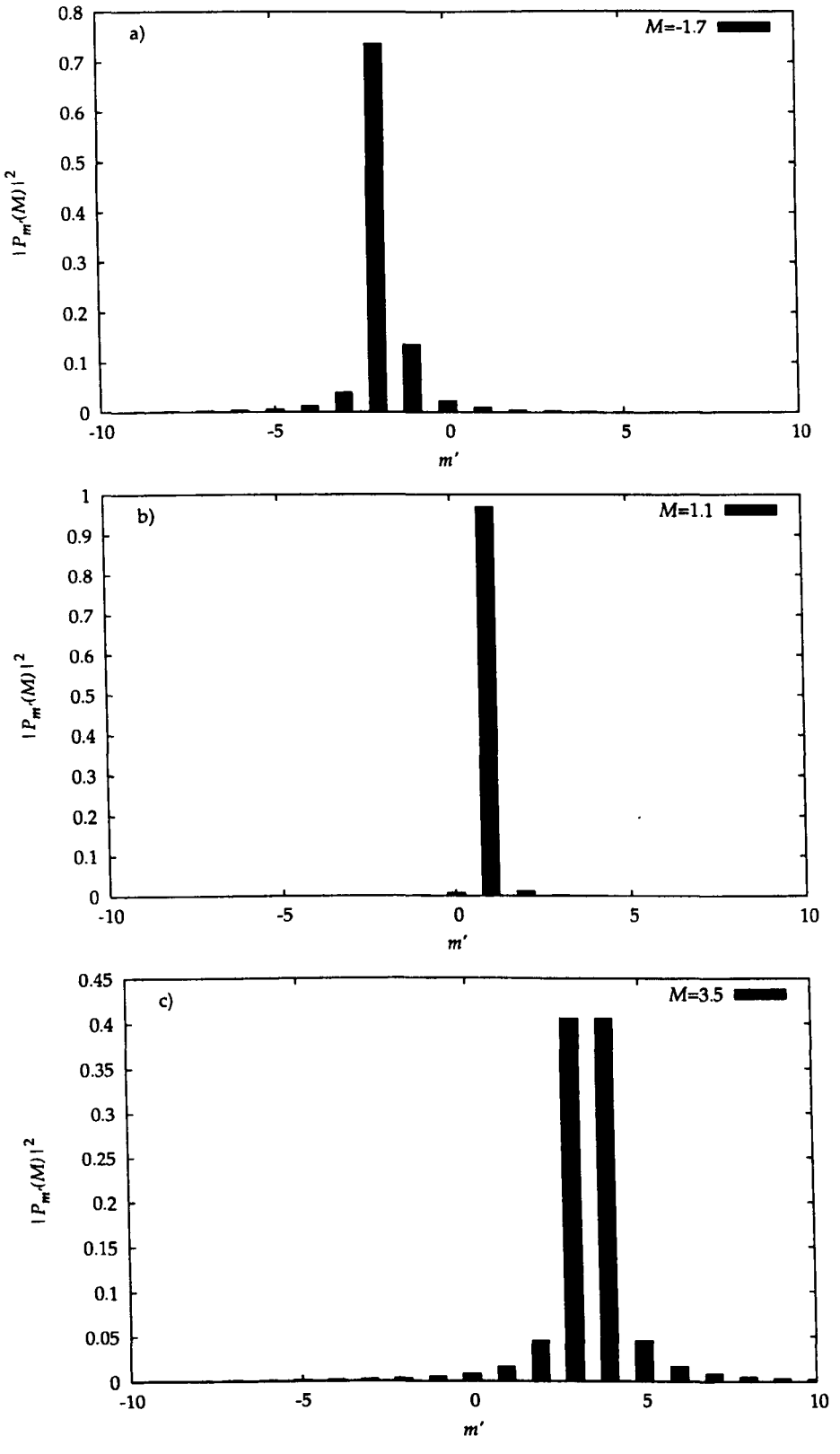

Figure VI.7 — Plot of the angular momentum distribution for fractional states. The distribution has a peak at the nearest integer to M. For $\mu = 0.5$ this results in two peaks of equal height at the two neighbouring integers. Also, the spread in the distribution is determined by the fractional value μ. In b) for $\mu = 0.1$ the distribution is concentrated at $m' = 1$ and only the nearest integers $m' = 0, 2$ have a probability visibly different from zero. In c) for $\mu = 0.5$, the distribution has the widest spread.

abilities $P_{m'}(M)$ adds to unity:

$$\sum_{m'=-\infty}^{\infty} P_{m'}(M) = \sum_{m'=-\infty}^{\infty} \frac{1-\cos(2M\pi)}{(M-m')^2 2\pi^2} = 1, \tag{VI.34}$$

where we have used $\cos(2\mu\pi) = \cos(2M\pi)$. The relevant summation can be executed using the contour integration method (see App. A on page 166), which yields

$$\sum_{m'=-\infty}^{\infty} \frac{1}{(M-m')^2} = -\frac{\pi^2}{\sin^2(\pi M)}. \tag{VI.35}$$

On using trigonometric identities this shows that the probabilities in Eq. (VI.34) sum indeed to unity. In a similar way the mean value of the OAM can be calculated:

$$\bar{M} = \sum_{m'=-\infty}^{\infty} m' P_{m'}(M) = \frac{1-\cos(2M\pi)}{2\pi^2} \sum_{m'=-\infty}^{\infty} \frac{m'}{(M-m')^2}. \tag{VI.36}$$

In this form, however, the summation cannot be evaluated using the contour integration method directly, because the overall expression does not fall off as m'^{-2} for large $|m'|$ (Stephenson & Radmore, 1993) (see App. A). By splitting the sum at $m' = 0$ and substituting $m' \to -m'$ in the sum from $m' = -\infty$ to $m' = 0$ we can transform the summation into a form which falls off as m'^{-2}:

$$\begin{aligned} \sum_{m'=-\infty}^{\infty} \frac{m'}{(M-m')^2} &= 4M \sum_{m'=0}^{\infty} \frac{m'^2}{(m'^2-M^2)^2} = 2M \sum_{m'=-\infty}^{\infty} \frac{m'^2}{(m'^2-M^2)^2}, \\ &= 2M \left(\frac{\pi^2}{2\sin^2(M\pi)} - \frac{\pi}{2} \frac{\cot(M\pi)}{M} \right). \end{aligned} \tag{VI.37}$$

This results in an OAM mean which is equal to M only at integer and half-integer values of M (cf. Eq. (VI.18) on page 130):

$$\bar{M} = M - \frac{\sin(2M\pi)}{2\pi}. \tag{VI.38}$$

A plot of the OAM mean as a function of M is given in figure (VI.8) on the next page. The OAM variance, however, tends to infinity

$$(\Delta M)^2 = \sum_{m'=-\infty}^{\infty} m'^2 P_{m'}(M) - \left(\sum_{m'=-\infty}^{\infty} m' P_{m'}(M) \right)^2 \to \infty. \tag{VI.39}$$

While the second term is the square of the mean value and therefore finite, the summation in the first term is divergent:

$$\sum_{m'=-\infty}^{\infty} \frac{m'^2}{(M-m')^2} = \sum_{m'=-\infty}^{\infty} 1 + \sum_{m'=-\infty}^{\infty} \frac{2Mm'}{(M-m')^2} - \sum_{m'=-\infty}^{\infty} \frac{M^2}{(M-m')^2}. \tag{VI.40}$$

The first sum causes the divergence whereas the second and third sum have been calculated in the sum of the probabilities (see Eq. (VI.34)) and the OAM mean (see Eq. (VI.36)) and are finite. From a more general point of view it is not surprising

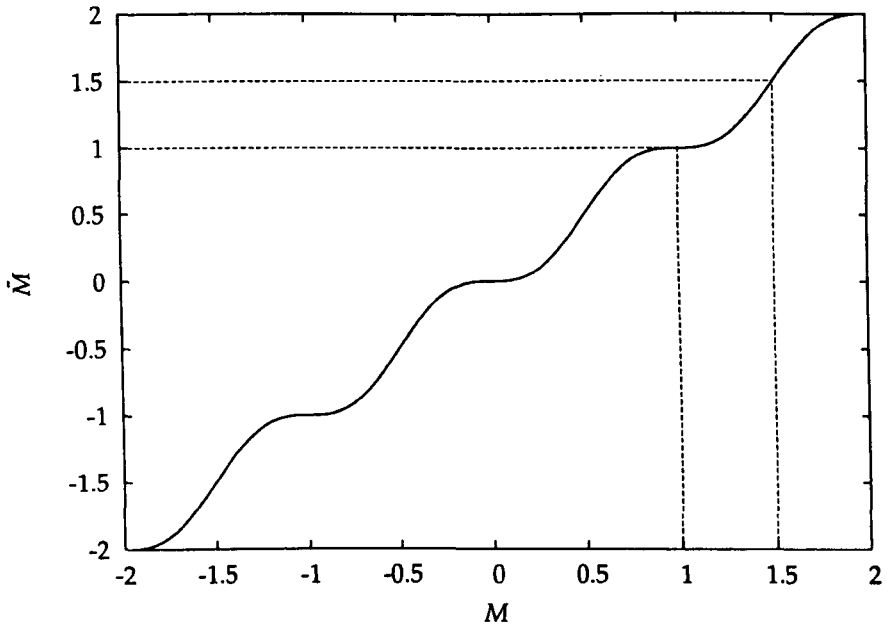

Figure VI.8 — Plot of the angular momentum mean value for fractional angular momentum states. Only at integer and half integer values of M is the angular momentum mean $\bar{M}$ equal to M. In general the relation between $\bar{M}$ and M is given in Eq. (VI.36).

that the OAM variance is divergent. The discontinuity in the phase leads also to a discontinuity in the angle wavefunction. For such a discontinuous wavefunction the mean square OAM will be divergent as demonstrated by Pegg et al. (2005). This is due to the singularity in the derivative, which also causes the rectangular apertures from the previous chapter to have an infinite inferred conditional variance (cf. section (4.3.1) on page 110).

The question of the OAM mean value for fractional states will be discussed also in the next section, where we calculate the propagation of fractional OAM beams.

3.4 Transformed basis

The spiral phase plate operator $\hat{S}(\alpha, M')$ introduced by Oemrawsingh et al. (2004*a*) is unitary and symmetric (Aiello et al., 2005). Therefore, the basis of OAM states $\{|m\rangle\}_{m\in\mathbf{Z}}$ can be transformed to a different basis $\{|(m+M')(\alpha)\rangle\}_{m\in\mathbf{Z}}$ with help of the unitary operator $\hat{S}(\alpha, M')$. This transformation is in part a rotation of the basis, not unlike rotating a canonical basis in a Euclidean space, but it is worth noting that the rotation only becomes apparent for fractional OAM. The basis of integer OAM states $\{|m\rangle\}_{m\in\mathbf{Z}}$ is left invariant by a rotation around α. Owing to the unitarity of $\hat{S}(\alpha, M')$ the transformed basis $\{|(m+M')(\alpha)\rangle\}_{m\in\mathbf{Z}}$ is also orthonormal and complete.

In the construction of states with fractional OAM in section VI 3.1 we have not introduced an operator to generate states with fractional OAM from states with integer OAM. We thus cannot show that a corresponding operator in our approach is unitary and symmetric. But we can show that the set of fractional OAM states $\{|(m+M')(\alpha)\rangle\}_{m\in\mathbf{Z}}$, which is obtained by applying the spiral phase plate operator $\hat{S}(\alpha, M')$ to the canonical basis of the OAM space $\{|m\rangle\}_{m\in\mathbf{Z}}$, is orthonormal and

complete. We prove these results in the arbitrarily large yet finite state space Ψ introduced in chapter III. The finite number of dimensions requires a bit more care in the operation of the spiral phase plate operator $\hat{S}(\alpha, M')$, where $M' = m' + \mu'$. This operator not only adds a fractional part μ' to the OAM, but also changes the integer part by m'. To ensure that the transformed basis is again fully contained in the state space Ψ, the addition has to be calculated such that values of $m + m' > L$ are mapped to $m + m' - (2L + 1)$. Correspondingly values of $m + m' < -L$ are mapped to $m + m' + (2L + 1)$. The fractional value $\mu' \in [0, 1)$ is unaffected by this restriction. With the particular form of adding to the integer part of the OAM every operator $\hat{S}(\alpha, M')$ can be treated as an operator with $M' = \mu \in [0, 1)$, that is with an integer part equal to zero. In the state space Ψ the transformed basis is thus given by $\{|m + \mu(\alpha)\rangle\}_{m=-L,L+1,\ldots,L}$. To change the absolute orientation of the fractional states we make use of the unitary operator $\hat{U}_m(\alpha)$ which in contrast to the spiral phase plate operator $\hat{S}(\alpha, M')$ only rotates the fractional states and does not change the OAM (see App. VI.B). We first show that the transformed basis states are orthonormal. In the transformed basis the orientation is always the same which is why we can change the relative orientation to zero with help of the unitary operator $\hat{U}_\alpha$:

$$\begin{aligned}\langle (m' + \mu)(\alpha) | (m + \mu)(\alpha) \rangle &= \langle (m' + \mu)(\alpha) | \hat{U}^\dagger_{m'}(-\alpha) \hat{U}_{m'}(-\alpha) | (m + \mu)(\alpha) \rangle, \\ &= \exp[-\mathrm{i}(m' - m)\alpha] \langle (m' + \mu)(0) | (m + \mu)(0) \rangle = \delta_{mm'}.\end{aligned} \tag{VI.41}$$

The orthonormality follows from Eq. (VI.27) on page 135 for $m \neq m'$ and from Eq. (VI.29) on page 135 for $m = m'$. For the completeness of the set $\{|m + \mu(\alpha)\rangle\}_{m \in \mathbf{Z}}$ we have to show that the sum of the projectors $|(m + \mu)(\alpha)\rangle\langle(m + \mu)(\alpha)|$ is equal to the identity operator:

$$\sum_{m=-L}^{L} |(m + \mu)(\alpha)\rangle\langle(m + \mu)(\alpha)| = \mathbb{1}. \tag{VI.42}$$

By acting on the sum with the operator $\hat{U}_m(-\alpha)$ from the left and with $\hat{U}^\dagger_m(-\alpha)$ from the right we can change the orientation in the fractional states to zero:

$$\begin{aligned}\mathbb{1} = \hat{U}^\dagger_m(-\alpha)\mathbb{1}\hat{U}^\dagger_(- \alpha) = \hat{U}^\dagger_m(-\alpha) \sum_{m=-L}^{L} |(m + \mu)(\alpha)\rangle\langle(m + \mu)(\alpha)| \hat{U}^\dagger_m(-\alpha), \\ = \sum_{m=-L}^{L} |(m + \mu)(0)\rangle\langle(m + \mu)(0)|.\end{aligned} \tag{VI.43}$$

To prove this result we make use of the decomposition of fractional OAM states in

Eq. (VI.21) on page 132:

$$\begin{aligned}\mathbb{1} &= \sum_{m=-L}^{L} \frac{1}{2L+1} \sum_{n,n'=0}^{2L} \exp[iM(\theta_n - \theta_{n'})]|\theta_n\rangle\langle\theta_{n'}|, \\ &= \frac{1}{2L+1} \sum_{n,n'=0}^{2L} \exp[i\mu(\theta_n - \theta_{n'})]|\theta_n\rangle\langle\theta_{n'}| \sum_{m=-L}^{L} \exp\left[im\frac{2\pi(n-n')}{2L+1}\right], \qquad \text{(VI.44)} \\ &= \frac{1}{2L+1} \sum_{n,n'=0}^{2L} \exp[i\mu(\theta_n - \theta_{n'})]|\theta_n\rangle\langle\theta_{n'}|(2L+1)\delta_{nn'} = \sum_{n=0}^{2L} |\theta_n\rangle\langle\theta_n|.\end{aligned}$$

In the state space Ψ the set of angle states $\{|\theta_n\rangle\}_{n\in\mathbf{N}}$ is complete which concludes the proof of the completeness of the transformed basis $\{|(m+\mu)(\alpha)\rangle\}_{m\in\mathbf{Z}}$. This establishes formally the equivalence between the spiral phase plate operator $\hat{S}(\alpha, M')$ and the fractional orbital angular momentum states $|(m+\mu)(\alpha)\rangle$. In the next section we will see that our approach to fractional orbital angular momentum is also applicable to the propagation of modes emerging from a fractional phase step.

4 Propagation of fractional modes

In this section we apply the theory of fractional OAM to calculate the propagation of a light beam after passing an optical component which changes the OAM by a fractional value M. The input is a Gaussian with zero OAM and we will use a decomposition in terms of Bessel beams to calculate the propagation of the emerging beam in the paraxial and non-paraxial regime.

4.1 Bessel decomposition of fractional modes

In chapter II we have presented Bessel beams as one form of light beams carrying OAM. One interesting property of Bessel beams is that they are solutions to the paraxial wave equation as well as the Helmholtz equation. By changing the expression for the longitudinal wavenumber we can switch between the different solutions in the same way as Berry (2004) in his analysis of waves evolving from fractional and integer phase steps (see section 2.2). But in contrast to this analysis we consider a Gaussian input beam with a finite beam width w_0. The beam emerging from spiral phase plate can thus be written as

$$\Psi(\rho, \varphi, z=0) = \exp\left(-\frac{\rho^2}{2w_0^2}\right)\exp(iM\varphi). \qquad \text{(VI.45)}$$

Another advantage of Bessel functions as solutions to the exact and paraxial wave equation is that the dependence on the radial and longitudinal coordinates factorises. We thus can express the Gaussian beam directly behind the phase step in a superposition of Bessel beams with different transverse wavenumbers κ. The propagated wave is obtained by multiplying each term in the superposition with the adequate propagation factor for the exact or paraxial form. The superposition of Bessel beams is specific for every integer value of the OAM and the decomposition of the wave behind a frac-

tional phase step requires a summation over the integer values (see Eq. (VI.15) on page 130):

$$\Psi(\rho,\varphi,z=0)=\sum_{m'=-\infty}^{\infty}c_{m'}[M(\beta)]\exp(im'\varphi)\int \mathrm{d}\kappa\, d_{m'}(\kappa)J_{m'}(\kappa\rho). \tag{VI.46}$$

The coefficients $c_{m'}[M(\beta)]$ are given by the overlap between the integer and fractional OAM states $\langle m'(0)|M(\beta)\rangle$ (see Eq. (VI.31) on page 136). This follows on multiplying both expressions for $\Psi(\rho,\varphi,z=0)$ in Eqs. (VI.45) and (VI.46) by $\exp(-im''\varphi)$ and integration over φ. The decomposition of the radial part is given by:

$$\exp\left(-\frac{\rho^2}{2w_0^2}\right)=\int \mathrm{d}\kappa\, d_{m'}(\kappa)J_{m'}(\kappa\rho), \tag{VI.47}$$

which has to be fulfilled for every m'. The coefficients $d_{m'}(\kappa)$ can be calculated using the Fourier-Bessel theorem (see App. VI.D):

$$\int_0^\infty \mathrm{d}\rho\,\rho J_m(\kappa\rho)J_m(\kappa'\rho)=\frac{1}{\kappa}\delta(\kappa'-\kappa). \tag{VI.48}$$

By multiplying both sides in Eq. (VI.47) by $J_{m'}(\kappa'\rho)$ and integration over ρ we find that the coefficients $d_{m'}(\kappa)$ are given by the integral:

$$d_{m'}(\kappa)=\kappa\int_0^\infty \mathrm{d}\rho\,\rho J_{m'}(\kappa\rho)\exp\left(-\frac{\rho^2}{2w_0^2}\right). \tag{VI.49}$$

The solution of this integral can be written in terms of modified Bessel functions I_ν with a half-integer order ν (Gradshteyn & Ryzhik, 2000). For all positive half-integer orders the modified Bessel functions increase exponentially. This behaviour changes for $\nu=-\frac{1}{2}$ and $\nu=-\frac{3}{2}$ which renders the integral expression for the coefficients in terms of modified Bessel functions invalid for $m'\leq -2$. We therefore have to evaluate the integral in Eq. (VI.49) for negative values of m' separately by using $J_{-m'}(\kappa\rho)=(-1)^{m'}J_{m'}(\kappa\rho)$ (Stephenson & Radmore, 1993). The coefficients in Eq. (VI.49) may thus be written as

$$\begin{aligned} d_{m'}(\kappa)=\kappa^2\frac{\sqrt{\pi}}{8}\left(2w_0^2\right)^{\frac{3}{2}}\exp\left(-\frac{\kappa^2w_0^2}{4}\right)\times \\ \left[I_{\frac{|m'|-1}{2}}\left(-\frac{\kappa^2w_0^2}{4}\right)-I_{\frac{|m'|+1}{2}}\left(-\frac{\kappa^2w_0^2}{4}\right)\right]\times\begin{cases}(-1)^{|m'|} & m'<0,\\ 1 & m'\geq 0.\end{cases} \end{aligned} \tag{VI.50}$$

The coefficients $d_{m'}$ show the same behaviour on inversion of the index m' as the Bessel functions $J_{m'}$; the coefficients for negative m' are related to the coefficients with positive m' by $d_{-m'}=(-1)^{m'}d_{m'}$. This is why in figures (VI.9) and (VI.10) on the next page the coefficients $d_{m'}$ are plotted only for positive m'. From Eq. (VI.50) one can see that the coefficients scale linearly with the beamwaist w_0. For large values of κ an asymptotic expansion of the modified Bessel functions is advantageous for the numerical evaluation (see App. VI.E).

On substituting the coefficients in Eq. (VI.46) the wave emerging from a fractional

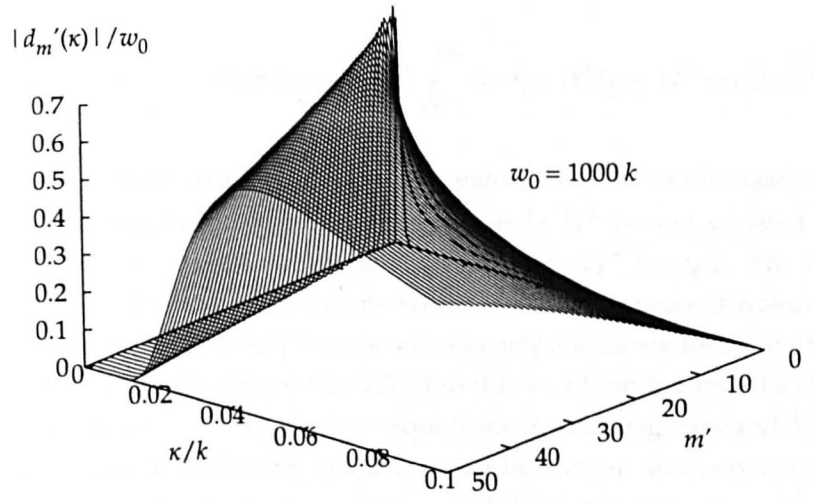

Figure VI.9 — Plot of the coefficients $d_{m'}(\kappa)$ in the Bessel decomposition in Eq. (VI.47) on the preceding page for $w_0 = 1000\lambda/(2\pi)$. The coefficients scale linearly with w_0. The coefficients for negative values of m' are given by $d_{-m'} = (-1)^{m'} d_{m'}$.

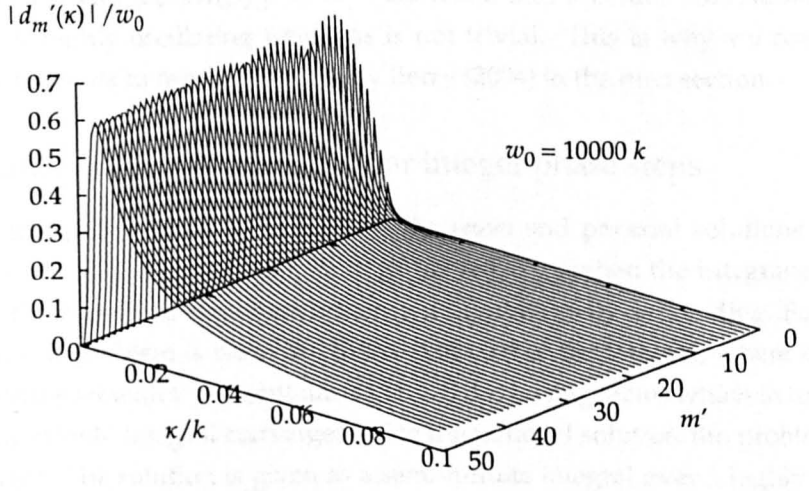

Figure VI.10 — Plot of the coefficients $d_{m'}(\kappa)$ in the Bessel decompostion in Eq. (VI.47) on the preceding page for $w_0 = 10000\lambda/(2\pi)$.

phase step is determined at $z = 0$. To calculate the propagated wave we have to add the appropriate propagation factor in the integral in Eq. (VI.46) on page 142. For the exact form the propagation factor is given by $\exp(i\sqrt{k^2 - \kappa^2}z)$ and the propagated wave reads as

$$\Psi(\rho, \varphi, z) = \sum_{m'=-\infty}^{\infty} c_{m'}[M(\beta)] \exp(im'\varphi) \int_0^{\infty} d\kappa\, d_{|m'|}(\kappa) J_{|m'|}(\kappa\rho) \exp(i\sqrt{k^2 - \kappa^2}z). \tag{VI.51}$$

A distinction between negative and positive values of m' is not necessary in this formulation. For negative m' the alternating factor $(-1)^m$ is compensated by the alternating sign of the Bessel functions $J_{-m'}(\kappa\rho) = (-1)^{m'} J_{m'}(\kappa\rho)$. The integration boundaries are set from zero to infinity which includes evanescent waves for $\kappa > k$. As pointed out by Roux (2003) the evanescent components are necessary to describe the singularity correctly (see section 2.2). But it is worth noting that the character of the integral changes at $\kappa = k$, where the square root turns imaginary and the exponential changes from an oscillating to a decreasing behaviour. In the paraxial solution, where the propagation factor is $\exp[ik(1 - \kappa^2/(2k^2)z]$ this distinction does not exist. The z dependent term $\exp(ikz)$ in the propagation factor can be written in front of the integral, which gives for the paraxial solution

$$\begin{aligned} \psi(\rho, \varphi, z) = &\sum_{m'=-\infty}^{\infty} c_{m'}[M(\beta)] \exp(im'\varphi) \exp(ikz) \\ &\times \int_0^{\infty} d\kappa\, d_{|m'|}(\kappa) J_{|m'|}(\kappa\rho) \exp\left(-i\frac{\kappa^2}{2k}z\right). \end{aligned} \tag{VI.52}$$

The integrals for the exact and paraxial solution are solved numerically to give the intensity and the phase profile of the propagated wave. The number of contributing integer orbital angular momentum modes is restricted in the numerical calculation. Only modes with $|c_{m'}[M(\beta)]| > 10^{-4}$ are taken into account. The numerical integration of highly oscillating functions is not trivial. This is why we compare our numerical results to results obtained by Berry (2004) in the next section.

4.2 Comparison with Berry's result for integer phase steps

The integrals in the expressions for the exact and paraxial solutions are calculated numerically. In particular for large values of z , when the integrand oscillates rapidly, the numerical algorithm becomes computationally demanding. For the exact solution this problem is restricted to the first part of the integral, where $\kappa < k$. The evanescent part with $\kappa > k$ contains a quickly decreasing factor which in turn renders the semi-infinite integral convergent. For the paraxial solution the problem is more prominent. The solution is given as a semi-infinite integral over a highly oscillating function. The only converging factors are the coefficients $d_{m'}(\kappa)$, but for larger values of m' the coefficients tend to zero more slowly. For these reasons it is not guaranteed that the adaptive numerical algorithms used in the calculation of the integrals give reliable results. We therefore compare our numerical calculations to analytical and numerical results from Berry (2004) for integer phase steps. In his article Berry presents the intensity profile of a propagated wave emerging from a phase step with

$m = 1$ for different values of z. As Berry considers a plane wave incident on the spiral phase plate, we approximate this setup with a very large beamwaist for the incident Gaussian beam in our approach. We choose a beamwaist of $w_0 = 1000000\ \lambda/(2\pi)$. For visible light with a wavelength of roughly one micron this corresponds to a beamwaist of about 1 decimetre.

The profiles are compared in figure (VI.11) on the next page. The values for Berry's results are manually determined from the graphs in the article. To plot the graphs for different propagation distances on the same scale the radial position ρ is divided by the square root of z. One can see from figure (VI.11) on the following page that the agreement is excellent for $z = 1, 5, 200$ and that there is a small difference between our results and Berry's results in the graph for $z = 50$. The reason for this discrepancy is not known, but the values follow a similar behaviour with a relative difference of under 3%. In Berry's approach the paraxial solution is given in closed form in terms of modified Bessel functions (see Eq. (VI.13) on page 130). The fact that the numerical integration of Eq. (VI.52) on the preceding page reproduces analytical results shows the accuracy of our integration method.

It is interesting to examine, why the two different expressions for the exact and paraxial wave give similar results. If we compare Eqs. (VI.12) and (VI.51) it becomes obvious that the coefficients $d_{m'}(\kappa)$ in our approach carry the dependence on the beamwaist w_0. In Berry's expression the coefficients are replaced by $1/\kappa$, but for large w_0 and $m' = 1$ the behaviour of $d_1(\kappa)$ is similar to $1/\kappa$ (see figure (VI.10) on page 143). The integrands in Eq. (VI.12) on page 129 and Eq. (VI.51) on the previous page are therefore similar. For other beamwaists w_0 and in particular for other values of m' the behaviour of $d_{m'}(\kappa)$ is much different from $1/\kappa$. For the paraxial solutions in Eqs. (VI.13) on page 130 and (VI.52) on the previous page the reasoning is similar. The fact that the general form of the coefficients $d_{m'}(\kappa)$ in Eq. (VI.50) on page 142 can be written in terms of Bessel functions similar to Berry's paraxial solution in Eq. (VI.14) on page 130 is not directly related to the agreement seen in figure (VI.11) on the following page. It is rather a consequence of the recurrent integral over quadratic exponentials combined with Bessel functions. The two expressions describe two different entities; while Berry's expression in Eq. (VI.14) on page 130 describes the modulus of the paraxial wave from an integer phase step completely, the coefficient in Eq. (VI.50) on page 142 only determines the contribution from one transverse wavevector to the Bessel decomposition in Eq. (VI.47) on page 142.

The series of graphs in figure (VI.11) on the following page also compares the exact and paraxial solution. The paraxial solution gives an accurate result near the vortex, but the oscillations far from the vortex are only correctly described if the $\rho/z \ll 1$ and $z \gg 1$. Within Berry's analysis the intensity profile in the paraxial solution does not change under propagation in the scaled coordinate $\rho/\sqrt{z}$ (see Eq. (VI.14) on page 130). The excellent agreement with results obtained by different means shows that our approach and the numerical integration of the exact and paraxial solution give reliable results.

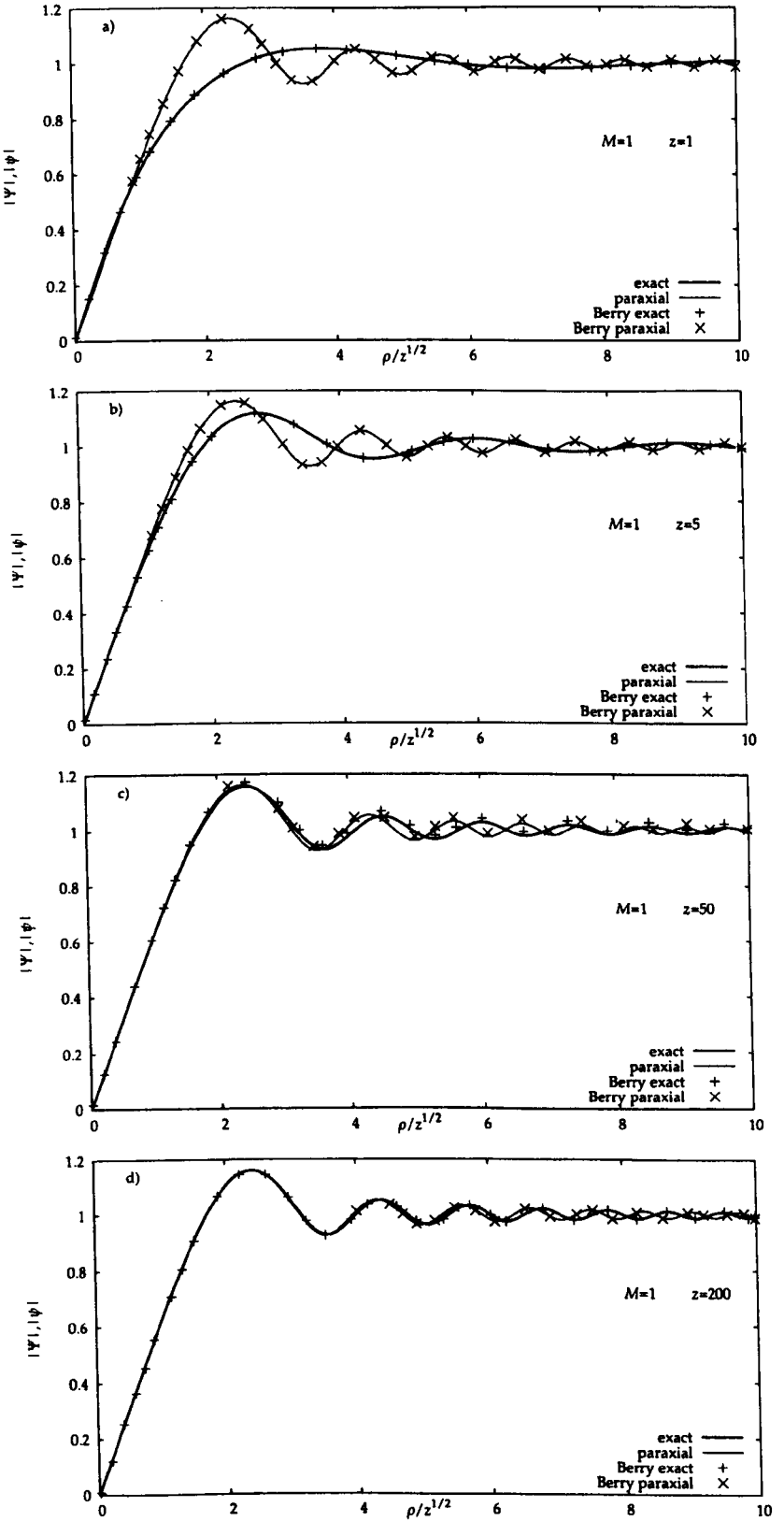

Figure VI.11 — Plot of the intensity profile of a wave emerging from an integer phase step for different propagation distances. The graphs compare results reported in (Berry, 2004) to the Bessel decomposition in section (4.1) on page 141.

4.3 Propagated phase and intensity profiles

The phase and intensity profiles are calculated for three different fractional phase steps with $M = -1.7, 1.1$ and $M = 3.5$. All three cases are given for the exact and paraxial solution. For the paraxial solution and a beam width $w_0 = 10000\ \lambda/(2\pi)$ the intensity profile remains unchanged on propagation and the phase profile simply rotates around the propagation axis. This is why we present four profiles for the exact solution at the propagation distances $kz = 1, 5, 50, 200$ and two for the paraxial solution at $kz = 1, 200$. The graphs are presented in the series of figures (VI.12) to (VI.17) from page 148 on. In the intensity profiles the most obvious feature is the radial line of low or zero intensity. In particular for $M = -1.7$ (see figure (VI.12) on the following page) and $M = 3.5$ (see figure (VI.17) on page 153) the line is clearly visible. In the case of $M = 3.5$ the intensity along this line drops to zero. For $M = 1.7$ the line has a lower intensity than the dark fringes but it is not equal to zero. The line of low intensity is framed by two lines of higher intensity, and this feature hints at the existence of a faint line of slightly lower intensity in the graphs for $M = 1.1$ (see figure (VI.14) on page 150). The other important feature in the intensity profiles is the number and position of the spots of zero intensity in the centre of the beam. The graphs for $M = -1.7, 1.1, 3.5$ show 2, 1 and 3 such spots respectively. These spots correspond to optical vortices and the number of vortices is given by the modulus of the nearest integer to M. While a beam with integer OAM $m \in \mathbb{Z}$ propagates with an optical vortex of charge m on the axis, beams with fractional OAM only show vortices with charge ± 1. None of these vortices is on the axis, but the whole central region has low intensity. Another prominent feature are the diffraction fringes surrounding the axis, due to diffraction on the initial singularity at the centre, and the radial fringes in the direction of α due to diffraction on the discontinuity in the spiral phase plate or hologram. For a fractional part $\mu = 1/2$ the contrast in the fringes is highest (cf. figure (VI.16) with figures (VI.12) and (VI.14)).

The phases profiles show clearly that the fractional strength vortex splits up in vortices with strength ± 1. The phase is represented by a colour circle, such that the colour for the phase of 0 and 2π are identical. A vortex corresponds to a point where all colours meet and for a vortex of strength ± 1 a circle surrounding the vortex goes through the colour circle once. The sign determines if the colours run through the circle clockwise or anti-clockwise. The difference becomes obvious if one compares the vortices in figure (VI.13) on page 149 with the vortices in the figures (VI.15) on page 151 and (VI.17) on page 153. Although vortices are the prominent feature in the phase profiles it is worth noting that the profiles are not determined by the position and charge of the vortices alone. In the phase profile for $M = -1.7$ in figure (VI.13) the corresponding colours are not opposite to each other. If the field would solely be given by a superposition of two charge -1 vortices the phase would show a symmetry with respect to the connecting line. The reason for this asymmetric behaviour is the presence of the initial radial phase discontinuity. On propagation the discontinuity smoothens, which can be seen from the sequence of phase profiles for the exact solution in figure (VI.13). In figure (VI.13a) for $kz = 1$ the line is smaller than in the subsequent profiles in figures (VI.13b) to (VI.13d) for $kz = 5, 50$ and $kz = 200$. From this series of phase profiles one can also see that the position of the vortices is

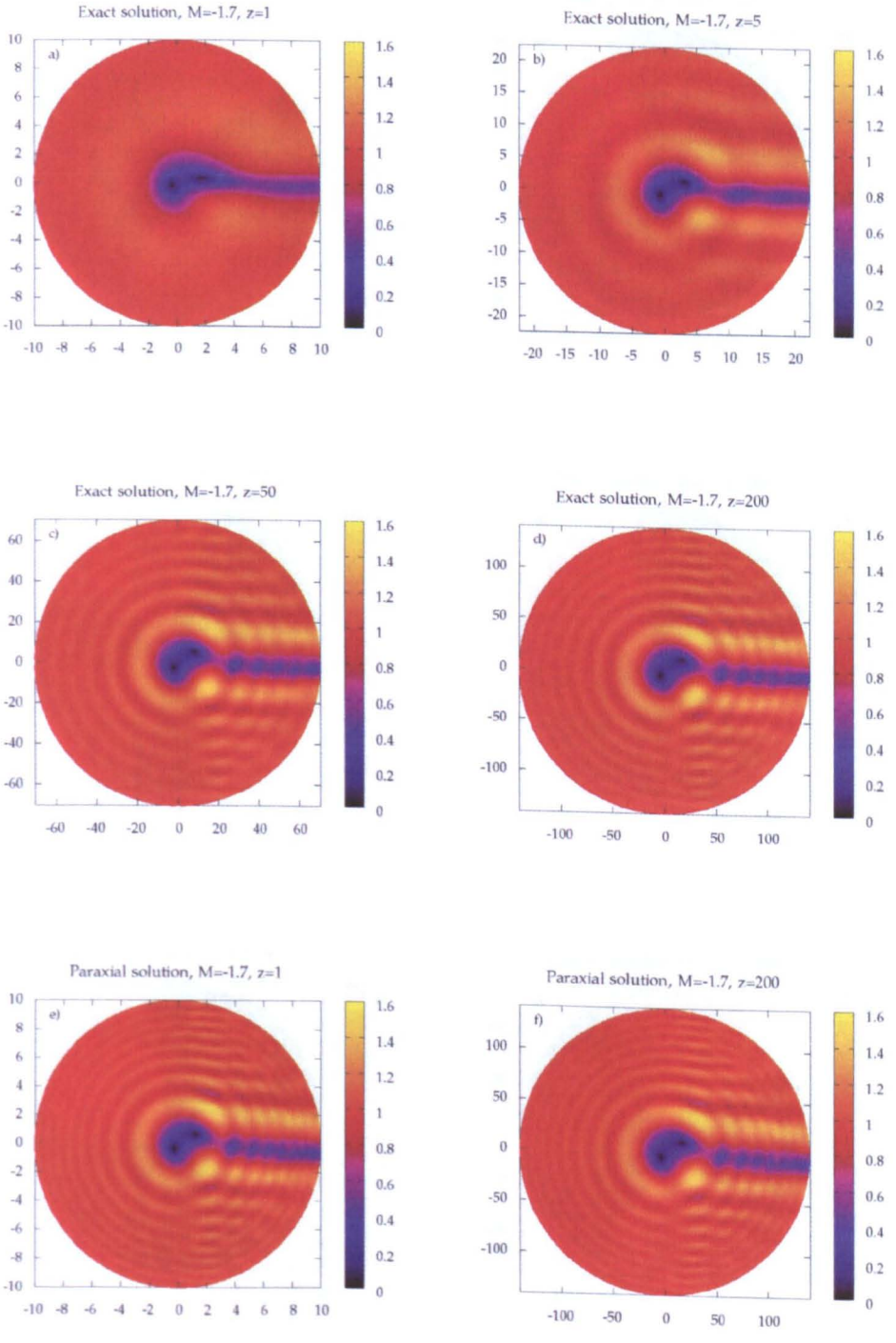

Figure VI.12 — Intensity profiles for $M = -1.7$. The graphs a-d) show the exact solution for the propagation distance $kz = 1, 5, 50, 200$ respectively. The graphs e) and f) show the paraxial solution for the distances $kz = 1$ and $kz = 200$. The radial line of low intensity is clearly visible.

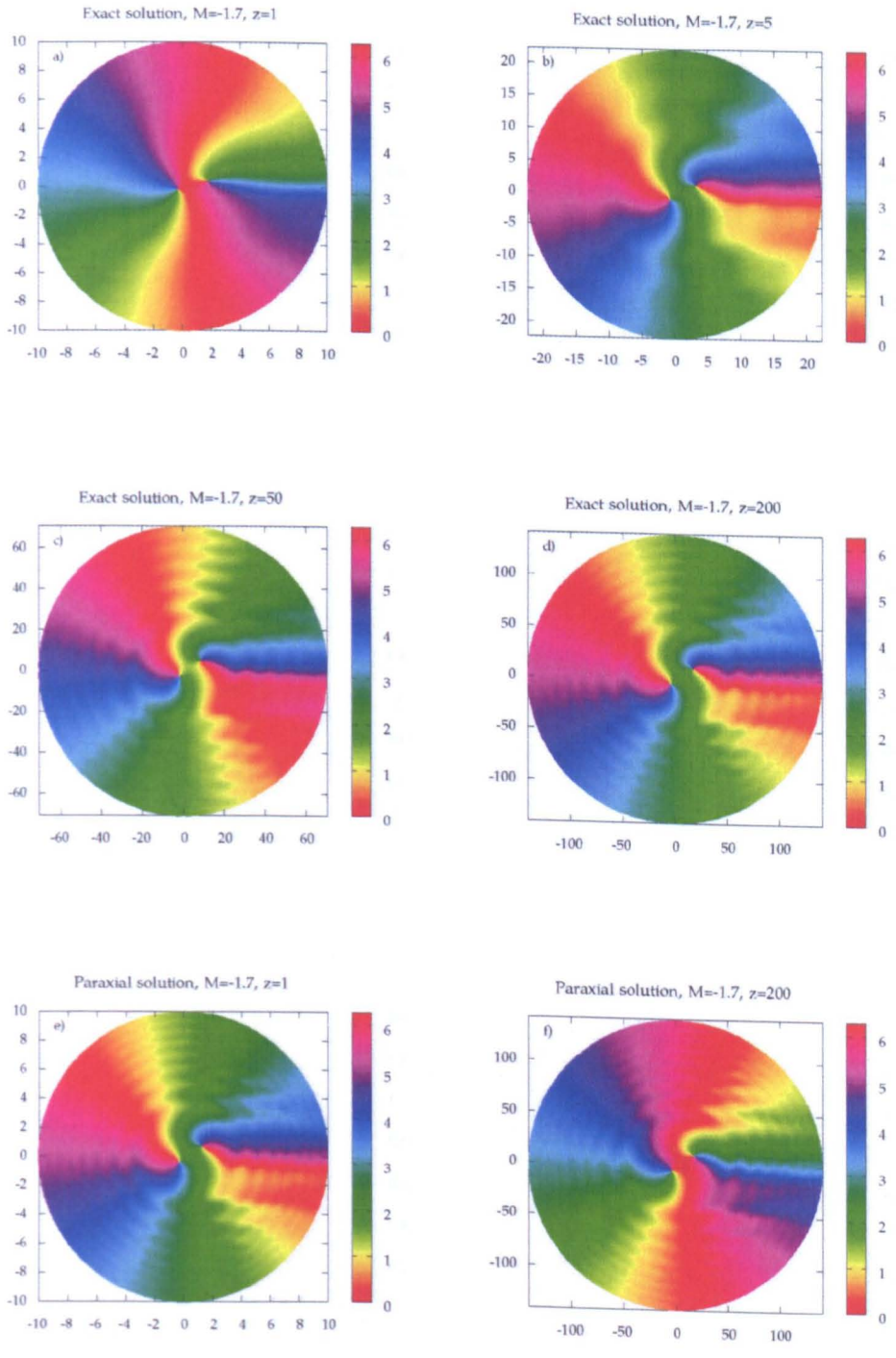

Figure VI.13 — Phase profiles for $M = -1.7$. The graphs a-d) show the exact solution for the propagation distance $kz = 1, 5, 50, 200$ respectively. The graphs e) and f) show the paraxial solution for the distances $kz = 1$ and $kz = 200$. The vortex with a fractional charge of $M = -1.7$ splits up in two vortices with charge -1.

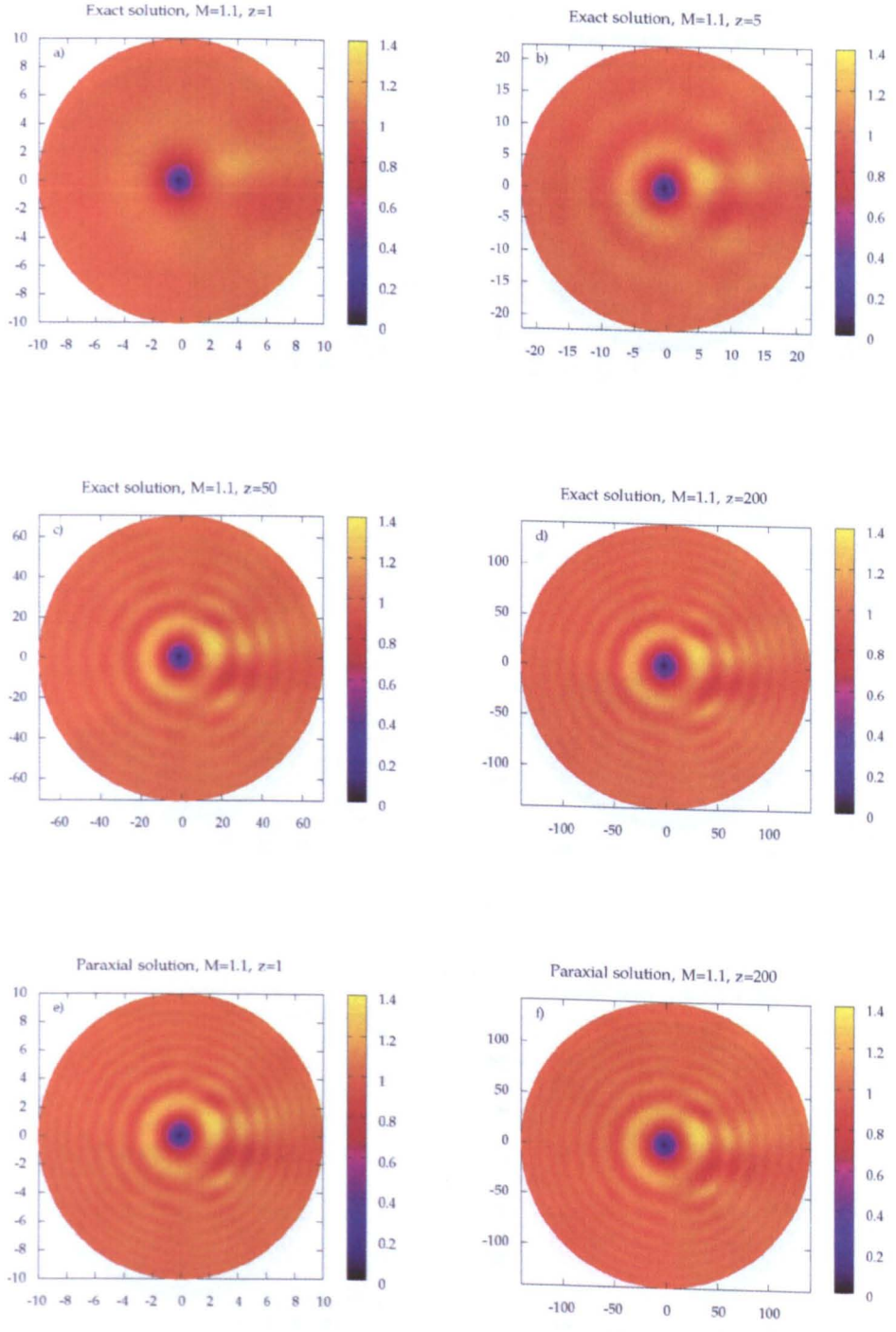

Figure VI.14 — Intensity profiles for $M = 1.1$. The graphs a-d) show the exact solution for the propagation distance $kz = 1, 5, 50, 200$ respectively. The graphs e) and f) show the paraxial solution for the distances $kz = 1$ and $kz = 200$. For a fractional phase step close to an integer phase step the radial line of lower intensity is only very faintly visible.

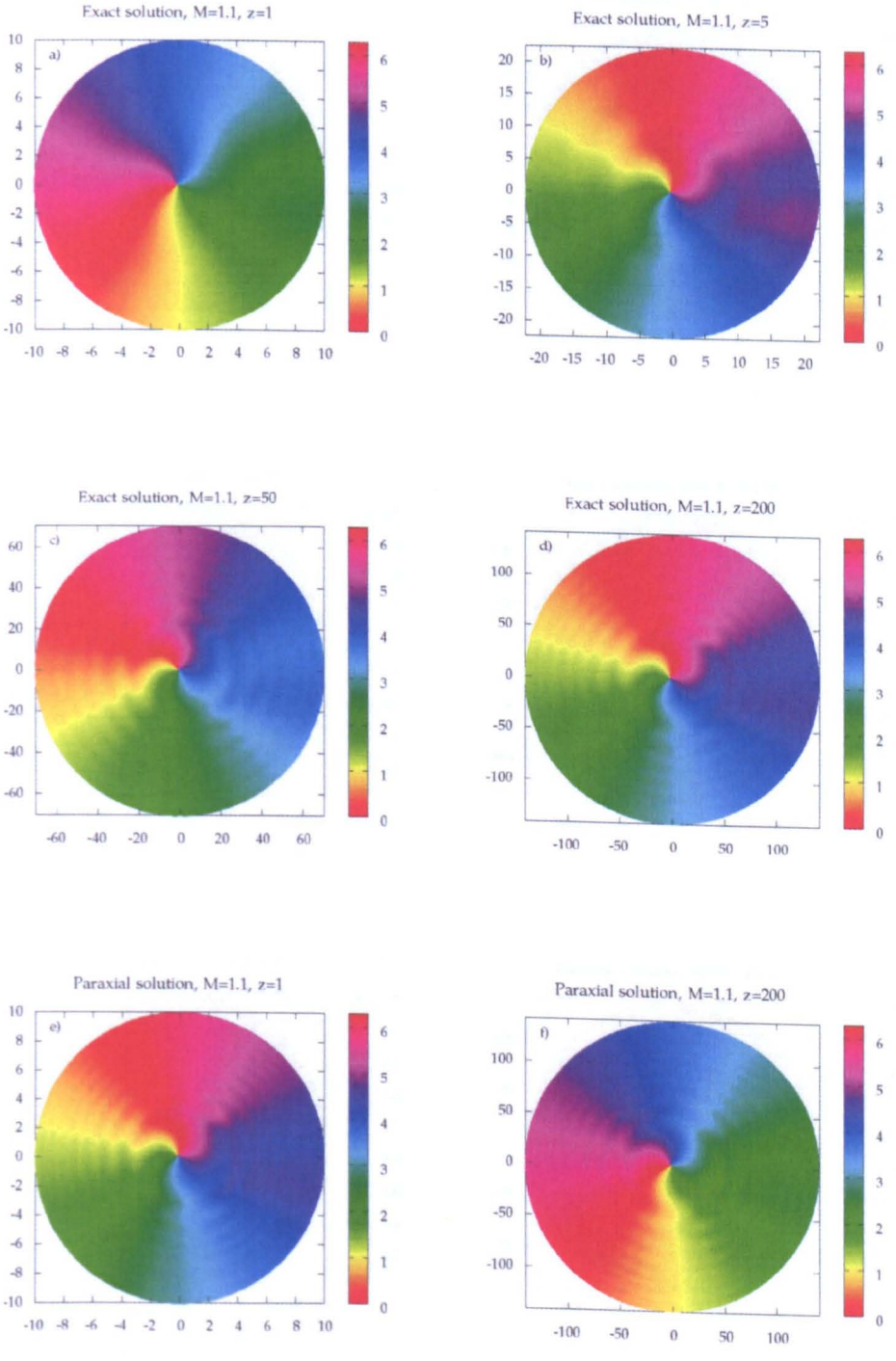

Figure VI.15 — Phase profiles for $M = 1.1$. The graphs a-d) show the exact solution for the propagation distance $kz = 1, 5, 50, 200$ respectively. The graphs e) and f) show the paraxial solution for the distances $kz = 1$ and $kz = 200$.

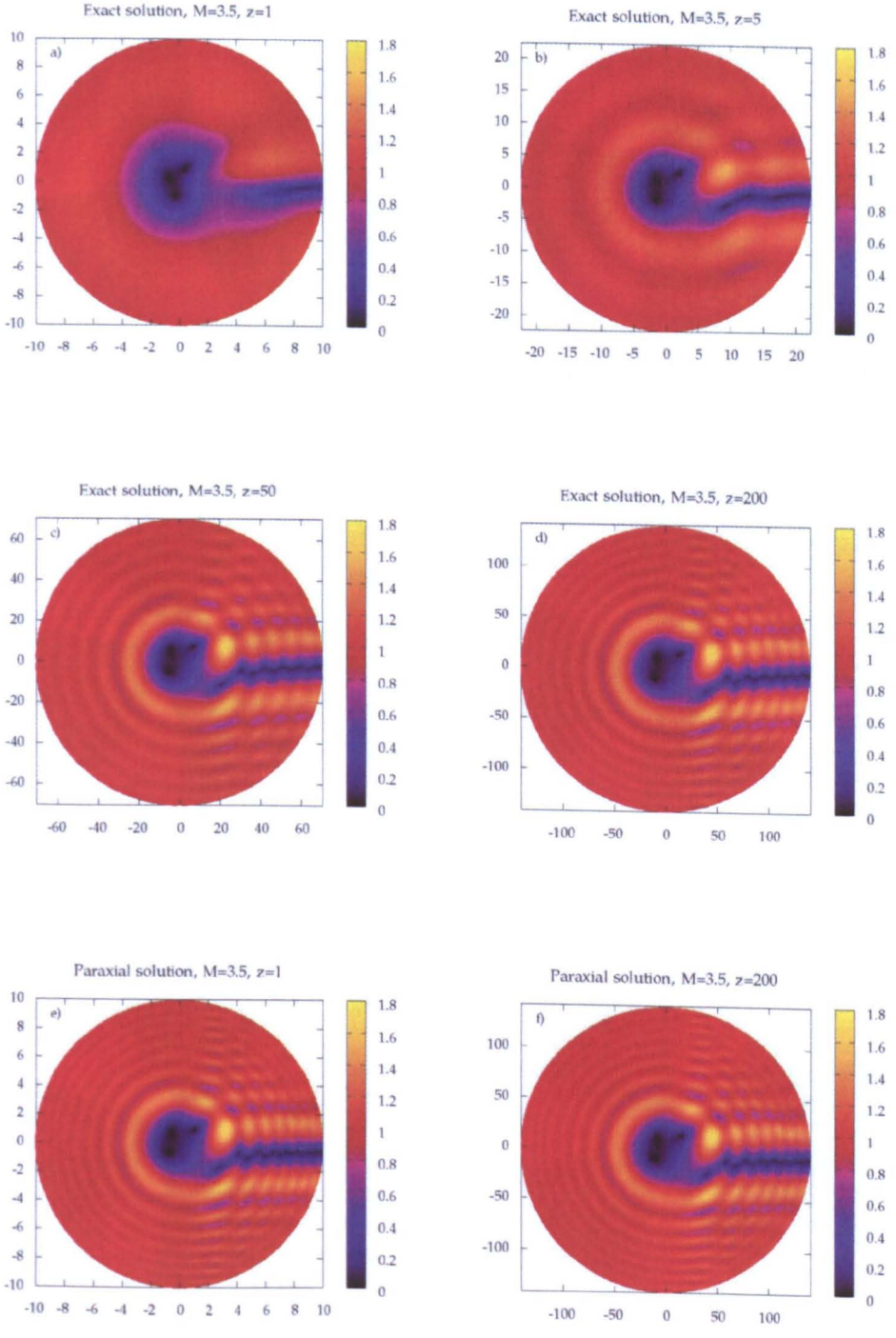

Figure VI.16 — Intensity profiles for $M = 3.5$. The graphs a-d) show the exact solution for the propagation distance $kz = 1, 5, 50, 200$ respectively. The graphs e) and f) show the paraxial solution for the distances $kz = 1$ and $kz = 200$. For fractional phase steps with $\mu = 1/2$ the intensity along the radial line drops to zero.

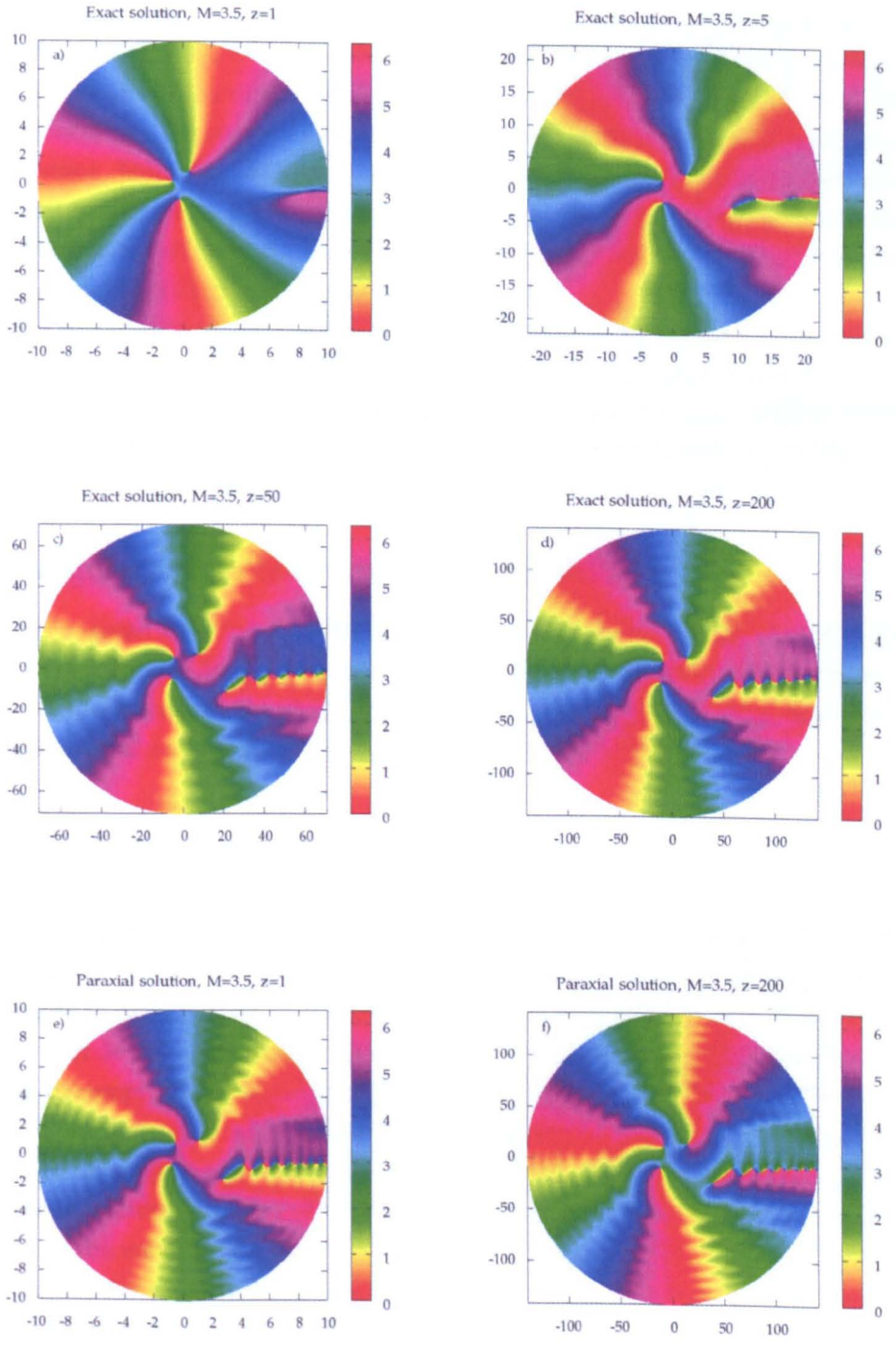

Figure VI.17 — Phase profiles for $M = 3.5$. The graphs a-d) show the exact solution for the propagation distance $kz = 1, 5, 50, 200$ respectively. The graphs e) and f) show the paraxial solution for the distances $kz = 1$ and $kz = 200$. The phase profiles show the creation of additional vortices on propagation.

changing slightly. In (VI.13a) the two vortices are positioned near the x axes and the connecting line is only slightly tilted with respect to the axes. On propagation the tilting increases. The most prominent feature in the phase profiles, however, is the creation of additional vortices in the region of zero intensity for half-integer phase steps. In figure (VI.17a) for $kz = 1$ one can see that there are no additional vortices, while in figure (VI.17b) one pair of vortices has formed and another pair is in the process of forming. At greater propagating distances the series of vortex pairs moves closer together. From the different rotation sense for the colours it can be seen that the vortices in this series have alternating charge. Berry (2004) points out that the vortices annihilate each other differently depending if the phase step has a fractional part μ of just under $1/2$ or just over $1/2$.

On propagation fractional modes develop a number of features which can only be seen in the exact solution. Although the paraxial solution approaches the exact solution for $z \to \infty$ the paraxial solution does not show the changing position of the integer vortices or the formation and movement in the chain of vortices. In the next section we discuss how the propagation affects the mean of the OAM.

4.4 Orbital angular momentum mean of propagated fields

The propagated field for the exact and paraxial solution are integrated numerically and we calculate the OAM mean from the propagated waves. This requires a suitable normalisation, as the numerical values are only calculated on a finite disc. The OAM mean at a propagation distance z is thus given by the integral

$$\bar{L}_z(z) = \int_{\mathcal{D}(\rho)} \mathrm{d}\rho\, \mathrm{d}\varphi\, \rho \Psi^* \hat{L}_z \Psi \times \left(\int_{\mathcal{D}(\rho)} \mathrm{d}\rho\, \mathrm{d}\varphi\, \rho \Psi^* \Psi \right)^{-1}, \qquad \text{(VI.53)}$$

where $\mathcal{D}(\rho)$ is a disc of radius ρ around the z axes. The graphs in the figures (VI.12), (VI.14) and (VI.16) show a smooth intensity profile without discontinuities. It is therefore possible to represent the OAM operator simply as derivative with respect to φ. The only part in the exact and paraxial wave which depends on the azimuthal angle is the decomposition into OAM eigenfunctions (see Eqs. (VI.51) and (VI.52) on page 144). The derivative with respect to φ thus leads to a factor of im' for every term in the OAM decomposition. The terms together with the respective integral from the Bessel decomposition are summed over m' which yields $\hat{L}_z \Psi(\rho, \varphi, z)$. The complex conjugate of the wave $\Psi^*(\rho, \varphi, z)$ is calculated in a similar manner from Eq. (VI.51) or Eq. (VI.52). In the numerical calculation of the OAM mean the product of these two terms is then summed over ρ and φ in accordance with the discretisation of the disc $\mathcal{D}(\rho)$. The integral over the modulus square of the wave is calculated analogously. In figure (VI.18) on the next page it can be seen that the numerical values deviate from the analytical formula in Eq. (VI.36) on page 138. The plot shows the difference in the OAM mean for the exact solution at the propagation distances $kz = 1$ and $kz = 5$. The difference in the plots for $kz = 1$ and $kz = 5$ can be explained if we assume that at a given propagation distance not all OAM eigenstates contribute to the OAM mean. This is possible because the corresponding eigenmodes to the OAM eigenstates have different spatial dimensions. For $kz = 50, 200$ and the paraxial solutions the devia-

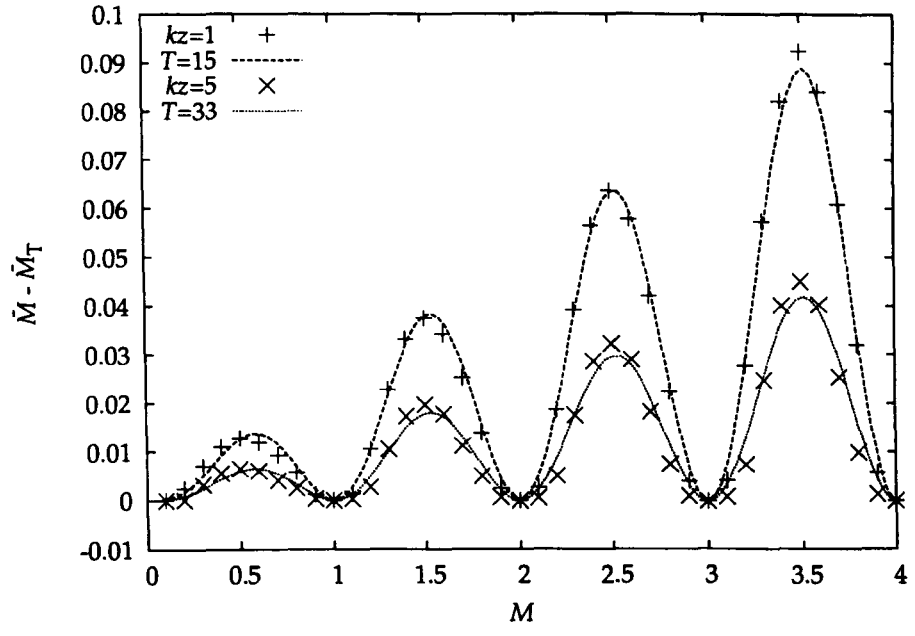

Figure VI.18 — Plot of the difference between the analytical expression for the angular momentum mean from Eq. (VI.36) on page 138 and numerically calculated values. The continuous curves are based on a model to explain the differences by assuming that the numerical values for the angular momentum mean are calculated on a limited subset of the OAM eigenstates $\{|m\rangle\}_{m=-T,-T+1\ldots,T}$.

tions from the analytical expression are so small that other effects are more dominant (see figure (VI.19) on the next page).

By introducing an index T at which the sum over m' in Eq. (VI.36) on page 138 is truncated, we calculate the OAM mean on a restricted subset of the OAM eigenstates:

$$\bar{M}_T = \sum_{m'=-T}^{T} m' P_{m'}(M) = \frac{1-\cos(2M\pi)}{2\pi^2} \sum_{m'=-T}^{T} \frac{m'}{(M-m')^2}. \tag{VI.54}$$

An analytical expression for this sum can be found by subtracting it from the untruncated sum. This leaves two infinite sums which are divergent if evaluated on their own. We thus have to introduce a limiting procedure to calculate the overall sum:

$$\bar{M}_T = \frac{1-\cos(2M\pi)}{2\pi^2} \left[\sum_{m'=-\infty}^{\infty} \frac{m'}{(M-m')^2} - \lim_{L\to\infty} \left(\sum_{m'=-L}^{-(T+1)} \frac{m'}{(M-m')^2} + \sum_{m'=(T+1)}^{L} \frac{m'}{(M-m')^2} \right) \right]. \tag{VI.55}$$

The first sum is the OAM mean given in Eq. (VI.36) on page 138. The sums in the limiting procedure can be combined in the same way as in the original calculation of

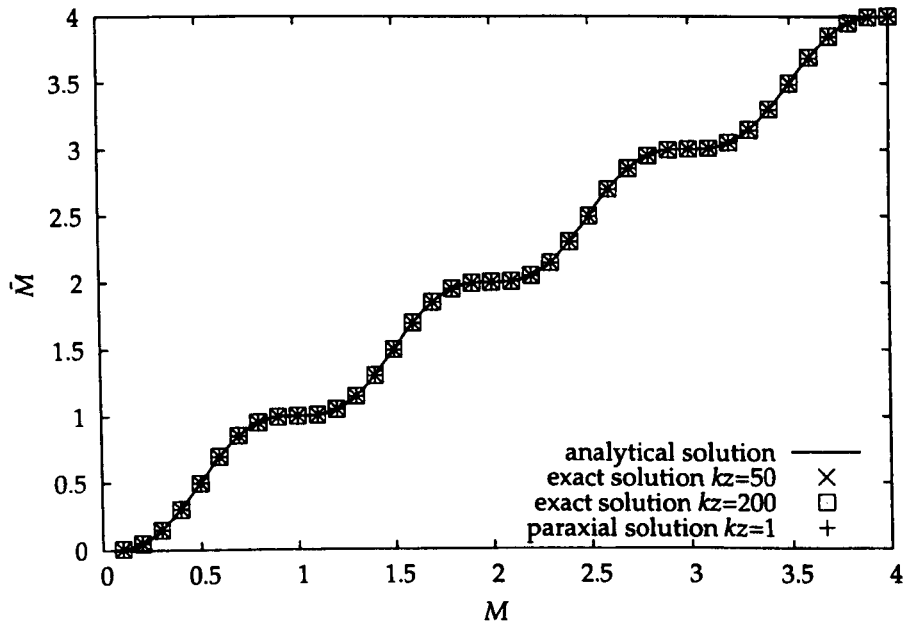

Figure VI.19 — Plot of the numerically calculated mean value of the OAM for different propagation distances. For smaller propagation distances not all OAM eigenmodes are contained within the calculated area. But for $kz = 50$ and $kz = 200$ the small differences to the analytical expression in Eq. (VI.36) on page 138 cannot be explained by a truncated set of modes. The paraxial solution does not show different values for the OAM mean at different propagation distances.

the OAM mean. The resulting sum is convergent and the limit $L \to \infty$ can be taken:

$$\lim_{L\to\infty} \sum_{m'=T+1}^{L} \frac{m'^2}{(m'^2 - M^2)^2} \approx 4M \int_{N+1}^{\infty} \frac{m'^2}{(m'^2 - M^2)^2} \mathrm{d}m' . \tag{VI.56}$$

Considering a truncated OAM mean aims to explain the differences between the analytical and numerical values for the OAM mean qualitatively. For a qualitative explanation we can approximate the integrand for $T \gg M$ in Eq. (VI.56) by $(m')^{-2}$. This allows us to calculate the fundamental integral in Eq. (VI.56) and hence to give an approximate expression for the truncated OAM mean:

$$\bar{M}_T \approx M\left(1 - \frac{2(1-\cos(2\pi M)}{T\pi^2}\right) - \frac{\sin(2\pi M)}{2\pi}. \tag{VI.57}$$

From this formula and the expression for the OAM mean in Eq. (VI.36) it is possible to calculate the difference $\bar{M} - \bar{M}_T$. We have fitted the analytical expression on the numerical data to estimate the value of T for the different propagation distances. The numerical fit yields real numbers for T, but in figure (VI.18) on the preceding page we have plotted the analytical expression for $\bar{M} - \bar{M}_T$ for the nearest integer. This is consistent with our assumption that the deviation from the analytical value for the OAM mean is caused by a limitation in the number of integer OAM eigenstates contributing to the sum for the mean value.

5 Conclusion

In this chapter we have derived a quantum formulation of fractional orbital angular momentum (OAM). This formulation is a generalisation of the quantum theory of rotation angles by Barnett & Pegg (1990) to fractional values of the OAM. Fractional values of OAM require the introduction of a branch cut in the angle representation of the angular momentum eigenfunctions. The orientation of the branch cut is an additional parameter in the description of fractional OAM. We have calculated the overlap of two general fractional OAM states. We have found that for odd half-integer values of the orbital angular momentum and a relative orientation of π radian the overlap is zero and the corresponding states are orthogonal. This confirms earlier work on fractional OAM by Oemrawsingh et al. (2004*a*) who consider half-integer spiral phase plates. On establishing an analogy between the description of these spiral phase plates in terms of a unitary operator and our generalised theory of fractional OAM we are able to identify the orientation of the branch cut with the orientation of the step in spiral phase plates.

We have applied our theory to the propagation of light beams with fractional OAM. For light carrying integer OAM a stable optical vortex of corresponding integer strength forms on the propagation axes. For fractional OAM no fractional strength vortices are stable on propagation, and instead a number of strength ± 1 vortices are formed in the central region of the beam. Additionally a line of low intensity is visible corresponding to the orientation of the branch cut or the phase step discontinuity in the generating optical device. For odd half-integer phase steps a chain of additional vortices with alternating charge is formed in this region of low intensity. We have confirmed theoretical and experimental results regarding the evolution of optical vortices in light beams with fractional OAM. Further to existing theories we have calculated the propagated fields in the paraxial and non-paraxial regime for an incident beam of finite width. The gradual formation of additional vortices in the region of low intensity is a phenomenon which is only visible in the non-paraxial solution. We have calculated the mean value of the OAM for different propagation distances. For small propagation distances the numerical values differ from the analytically calculated values. To explain these differences we assume that for small propagation distances not all OAM eigenmodes are contained in the numerical calculation. On calculating the OAM mean on a truncated set of OAM states and comparing the analytical results to the numerical values we find that our assumptions explain the differences accurately.

VI.A Branch cut

A branch of a many-valued function is a simply connected region in the complex plane on which the original function is continuous and single valued. The branch is obtained by cutting the complex plane along a simple arc starting at the origin and reaching to infinity. Branches of arbitrary functions can be rigorously defined over the branches of the logarithm (Conway, 1978). But it is possible to illustrate the idea behind a branch cut without going into the mathematical definition (Hahn & Epstein, 1996).

The function $f(z) = \sqrt[3]{z} = z^{1/3}$ has three different values for each $z \neq 0$. However, it is possible to construct a simply connected region not containing the origin in which the value of the function is unambiguously determined once a value is given at any point in the region. If we cut the complex plane along the negative real axis and include the origin in the cut, we end up with a simply connected region. The value of the function $f(z)$ in this region is completely determined once we choose $\sqrt[3]{1} = 1, \exp(i2\pi/3)$ or $\exp(i4\pi/3)$. Each of these branches is single-valued if z is restricted to this region. Although for every z in the complex plane there is now only one value of $f(z)$ in one branch, this does not mean, that one branch covers the whole complex plane. By introducing branches of the multi-valued functions we have restricted the range of the function $f(z)$. The polar angle θ in the representation $z = |z|\exp(i\theta)$ ranges either from 0 to 2π or from $-\pi$ to π. Depending on the range of θ the first branch with $\sqrt[3]{1} = 1$ contains numbers with their complex argument ranging from 0 to $2\pi/3$ or from $-\pi/3$ to $\pi/3$. The second branch starts where the first ended and so on, until the largest complex argument in the third branch connects to the lowest argument of the first branch with a difference of 2π.

VI.B Rotation operator for fractional angular momentum states

Quantum states with fractional OAM depend on α, the orientation of the phase discontinuity. To rotate the orientation we introduce a class of operators $\hat{U}_m(\beta), m \in \mathbb{R}, \beta \in [0, 2\pi)$ defined by their action on a state with fractional OAM:

$$\hat{U}_m(\beta)|M'(\alpha)\rangle = \exp[i(m - m')\beta]|M(\alpha \oplus \beta)\rangle, \tag{VI.58}$$

where β is the angle through which the discontinuity is rotated. The addition $\alpha \oplus \beta = (\alpha + \beta) \mod 2\pi$ yields a result in the range $[0, 2\pi)$. For $m = m'$, that is if m is equal to the integer part of M, this operator is a pure rotation of the orientation α, but for $m \neq m'$ the rotated state acquires also a phase shift unless $\beta \mod 2\pi = 0$. Consequently, the operator acts as identity operator if β is an integer multiple of 2π. The eigenstates of this operator are the integer OAM states. Acting with $\hat{U}_m(\beta)$ on an integer state $|m\rangle, m \in \mathbb{Z}$, results only in a phase shift:

$$\hat{U}_m(\beta)|m'\rangle = \exp[i(m - m')\beta]|m'\rangle. \tag{VI.59}$$

All the eigenvalues have unit modulus and the operator $\hat{U}_m(\beta)$ is therefore unitary with $\hat{U}_m(\beta) = \hat{U}_m(-\beta) = \hat{U}_m^\dagger(\beta)$.

From the definition of the operator in Eq. (VI.22) it is obvious that the set of operators cannot form a group under multiplication for arbitrary combinations of parameters β and m. But for a fixed m the set of operators $\{\hat{U}_m(\beta)\}_{\beta\in[0,2\pi)}$ does form a one-parameter group under multiplication with the group parameter β (Galindo & Pascual, 1990). As proof we show that $\hat{U}_m(\beta')\hat{U}_m(\beta)$ can always be written as $\hat{U}_m(\gamma)$

with $\gamma \in [0, 2\pi)$:

$$\begin{aligned}\hat{U}_m(\beta')\hat{U}_m(\beta)|M'(\alpha)\rangle &= \hat{U}_m(\beta')\exp[\mathrm{i}(m-m')\beta]|M'(\alpha\oplus\beta)\rangle,\\ &= \exp[\mathrm{i}(m-m')(\beta+\beta')]|M'(\alpha\oplus\beta\oplus\beta')\rangle,\\ &= \hat{U}_m(\beta\oplus\beta')|M'(\alpha)\rangle.\end{aligned} \tag{VI.60}$$

Therefore $\hat{U}_m(\beta')\hat{U}_m(\beta) = \hat{U}_m(\beta\oplus\beta')$ as the modulo 2π addition $\beta\oplus\beta'$ is always in the interval $[0, 2\pi)$. There is also a neutral element in form of the identity operator for $\beta = 0$. Moreover, to every operator $\hat{U}_m(\beta)$ exists an inverse element in the form $\hat{U}_m^\dagger(\beta) = \hat{U}_m(-\beta)$.

The product of two operators with the same orientation but different values for the fractional OAM can be combined to give

$$\hat{U}_m(\beta)\hat{U}_{m'}(\beta) = \exp[-\mathrm{i}(m+m')\beta]U_{m+m'}(\beta\oplus\beta). \tag{VI.61}$$

The set of operators $\{\hat{U}_m(\beta)\}_{m\in\mathbb{Z}}$ does not form a group under multiplication for $\beta \neq 0$. For $\beta = 0$ these operators are all identity operators.

We have introduced the operator $\hat{U}_m(\beta)$ to facilitate the calculation of the general overlap $\langle M'(\alpha')|M(\alpha)\rangle$ in section VI 3.2. The mathematical properties suggest, that the action of this operator on a state $|M'(\alpha)\rangle$ is independent of the fractional part μ'. A physical implementation of this operator would be an optical device with two edge dislocations; one to compensate the original dislocation at an angle α, the other to imprint the new phase discontinuity at an angle $\alpha\oplus\beta$. This implementation, however, would be specific for given fractional part μ'. As the effect of the edge dislocation is completely specified by the fractional part μ' and the orientation α it would be possible to change the orientation of the state $|M'(\alpha)\rangle$ with an implementation of the operator $\hat{U}_m(\beta)$, where the integer part of M' is different from m, which causes the overall phase factor in definition of the rotation operator in Eq. (VI.22) on page 133.

VI.C Geometric progression

A geometric progression is a sequence of numbers such that any two consecutive members in the sequence have a common ratio. A geometric progression can be written as

$$a, ar, ar^2, ar^3, \ldots, \tag{VI.62}$$

where $r \neq 0$ is the common ratio and a is a scaling factor. A sum of the numbers in a geometric progression is called a geometric series. A finite geometric series can be evaluated by the following consideration. The sum over $N+1$ members of in the geometric progression is given by

$$\sum_{n=0}^{N} ar^n = a + ar + ar^2 + ar^3 + \cdots + ar^n. \tag{VI.63}$$

By multiplying both sides of the equation with $(1-r)$ any two consecutive terms other than the first or last term cancel. The sum on the right hand side is thus reduced

to

$$(1-r)\sum_{n=0}^{N} ar^n = a - ar^N. \tag{VI.64}$$

The geometric series starting at $n = 0$ can be evaluated to

$$\sum_{n=0}^{N} ar^n = \frac{a(1-r^N)}{1-r}. \tag{VI.65}$$

If the sum does not start at $n = 0$ but at different order $M \leq N$ the geometric series can be evaluated by subtraction from the full series:

$$\sum_{n=M}^{N} ar^n = \sum_{n=0}^{N} ar^n - \sum_{n=0}^{N} ar^n = \frac{a(1-r^N)}{1-r} - \frac{a(1-r^M)}{1-r} = \frac{a(r^M-r^N)}{1-r}. \tag{VI.66}$$

The exponentials in the calculation of the general overlap of two states with fractional OAM in Eq. (VI.25) on page 133 are members of a geometric progression with the common ration $\exp[\mathrm{i}(M-M')2\pi/(2L+1)]$.

VI.D Fourier-Bessel theorem

The Fourier-Bessel theorem is closely connected to the Hankel transform (Arfken, 1985). The Hankel transform of order k is defined as:

$$F_k(\alpha) = \mathcal{H}_k\{f(t)\} = \int_0^\infty f(t)J_k(\alpha t)\mathrm{d}t, \tag{VI.67a}$$

$$f(t) = \mathcal{H}_k^{-1}\{F_k(\alpha)\} = \int_0^\infty f_k(\alpha)J_k(\alpha t)\mathrm{d}\alpha. \tag{VI.67b}$$

On substituting $F_k(\alpha)$ into the inverse Hankel transform we obtain:

$$f(t) = \int_0^\infty \mathrm{d}\alpha \int_0^\infty \mathrm{d}t' f(t')J_k(\alpha t')J_k(\alpha t), \tag{VI.68}$$

from which we can deduce the Fourier-Bessel theorem:

$$\int_0^\infty \mathrm{d}\alpha J_n(\alpha t)J_n(\alpha t') = \frac{1}{t'}\delta(t-t'). \tag{VI.69}$$

The connection between Hankel transforms and Fourier transforms is discussed in (Goodman, 2005).

VI.E Asymptotic Expansion for Bessel functions

The modified Bessel functions in the expression for the coefficients in the Bessel decomposition in Eq. (VI.50) on page 142 show a exponential increase for larger arguments. Overall the coefficients $d_m(\kappa)$ tend to zero for $\kappa \to \infty$ but the numerical calculation may cause overflow errors for large values of κ owing to the divergent behaviour of the modified Bessel functions. By using a asymptotic expansion for large values of κ the difficulties in the numerical evaluation of the $d_m(\kappa)$ can be avoided.

For large arguments z with $|\arg(z)| < \pi/2$ modified Bessel functions can be approximated by the expansion

$$I_\nu(z) \approx \frac{\exp(z)}{\sqrt{2\pi z}} \left(P_\nu(\mathrm{i}z) - \mathrm{i}Q_\nu(\mathrm{i}z)\right), \tag{VI.70}$$

where P_ν and Q_ν are given by:

$$P_\nu(\xi) = 1 + \sum_{n=1}^{\infty} (-1)^n \frac{\prod_{k=0}^{2n-1}[\mu - (2k+1)^2]}{(2n)!(8\xi)^{2n}}, \tag{VI.71a}$$

$$Q_\nu(\xi) = \sum_{n=0}^{\infty} (-1)^n \frac{\prod_{k=0}^{2n}[\mu - (2k+1)^2]}{(2n+1)!(8\xi)^{2n+1}}, \tag{VI.71b}$$

with $\mu = 4\nu^2$. With help of this asymptotic expansion the coefficients $d_m(\kappa)$ may be approximated for large κ to:

$$\begin{aligned} d_m(\kappa) = \frac{\kappa w_0^2}{2} \Big[&\left(P_{(m-1)/2}(\mathrm{i}\kappa^2 w_0^2/4) - P_{(m+1)/2}(\mathrm{i}\kappa^2 w_0^2/4)\right) \\ &-\mathrm{i}\left(Q_{(m-1)/2}(\mathrm{i}\kappa^2 w_0^2/4) - Q_{(m+1)/2}(\mathrm{i}\kappa^2 w_0^2/4)\right)\Big]. \end{aligned} \tag{VI.72}$$

In the plots in figures (VI.9) and (VI.10) on page 143 the coefficients have been calculated with the asymptotic expansion for values of $\kappa w_0 > 50$. The plots shows a smooth transition to the region where the asymptotic expansion has been used to calculate the values of the coefficients $d_m(\kappa)$.

Bibliography

Abramowitz, M. & Stegun, I. S. (1974), *Handbook of Mathematical Functions*, Dover Publications, Inc., Mineola, New York. Reprint. Originally published: National Bureau of Standards, corrected edition, 1964. ISBN: 0-486-61272-4.

Aiello, A., Oemrawsingh, S. S. R., Eliel, E. R. & Woerdman, J. P. (2005), 'Nonlocality of high-dimensional two-photon orbital angular momentum states', *Physical Review A* **72**, 052114. DOI: 10.1103/PhysRevA.72.052114.

Arfken, G. (1985), *Mathematical Methods for Physicists*, 3rd edn, Academic Press, Inc., Orlando. ISBN: 0-12-059820-5.

Barnett, S. M. & Pegg, D. T. (1990), 'Quantum theory of rotation angles', *Physical Review A* **41**(7), 3427–3435. DOI: 10.1103/PhysRevA.41.3427.

Berry, M. V. (2004), 'Optical vortices evolving from helicoidal integer and fractional phase steps', *Journal of Optics A* **6**(2), 259–269. DOI: 10.1088/1464-4285/6/2/018.

Born, M. & Wolf, E. (1999), *Principles of Optics*, 7th expanded edn, Cambrige University Press, Cambridge. reprinted 2003. ISBN: 0-521-64222-1.

Conway, J. B. (1978), *Functions of One Complex Variable*, no 11 *in* 'Graduate Text in Mathematics', 2nd edn, Springer Verlag.

Galindo, A. & Pascual, P. (1990), *Quantum Mechanics*, Vol. I, Springer Verlag, Berlin. ISBN: 3-540-51406-6.

Goodman, J. W. (2005), *Introduction to Fourier Optics*, 3rd edn, Roberts and Company Publishers, Englewood, Colorado.

Gradshteyn, I. S. & Ryzhik, I. M. (2000), *Tables of Integrals, Series, and Products*, 6th edn, Academic Press, Inc., San Diego. ISBN: 0-12-294757-6.

Hahn, L. & Epstein, B. (1996), *Classical Complex Analysis*, Jones and Bartlett Publishers, Boston.

Leach, J., Yao, E. & Padgett, M. J. (2004), 'Observation of the vortex structure of a non-integer vortex beam', *New Journal of Physics* **6**, 71. DOI: 10.1088/1367-2630/6/1/071.

Marrucci, L., Manzo, C. & Paparo, D. (2006), 'Optical Spin-to-Orbital Angular Momentum Conversion in Inhomogeneous Anisotropic Media', *Physical Review Letters* **96**(16), 163905. DOI: 10.1103/PhysRevLett.96.163905.

Nye, J. F. (1999), *Natural Focussing and the Fine Structure of Light*, Institute of Physics Publishing, Bristol.

Oemrawsingh, S. S. R., Aiello, A., Eliel, E. R., Nienhaus, G. & Woerdman, J. P. (2004*a*), 'How to observe High-Dimensional Two-Photon Entanglement with Only Two Detectors', *Physical Review Letters* **92**, 217901. DOI: 10.1103/PhysRevLett.92.217901.

Oemrawsingh, S. S. R., Eliel, E. R., Woerdman, J. P., Verstegen, E. J. K., Kloosterboer, J. G. & t' Hooft, G. W. (2004*b*), 'Half-integral spiral phase plates for optical wavelengths', *Journal of Optics A* **6**, S288–290.

Oemrawsingh, S. S. R., Ma, X., Voigt, D., Aiello, A., Eliel, E. R., 't Hooft, G. W. & Woerdman, J. P. (2005), 'Experimental Demonstration of Fractional Orbital Angular Momentum Entanglement of Two Photons', *Physical Review Letters* **95**(24), 240501. DOI: 10.1103/PhysRevLett.95.240501.

Pegg, D. T., Barnett, S. M., Zambrini, R., Franke-Arnold, S. & Padgett, M. (2005), 'Minimum uncertainty states of angular momentum and angular position', *New Journal of Physics* **7**, 62. DOI: 10.1088/1376-2630/7/1/062.

Roux, F. S. (2003), 'Optical vortex density limitations', *Optics Communication* **223**(1-3), 31–37. DOI: 10.1016/S0030-4018(03)01626-2.

Spedalieri, F. M. (2006), 'Quantum key distribution without reference frame alignment: Exploiting photon orbital angular momentum', *Optics Communications* **260**, 340–346. DOI: 10.1016/j.optcom.2005.10.001.

Stephenson, G. & Radmore, P. M. (1993), *Advanced Mathematical Methods for Engineering and Science Students*, Cambrige University Press, Cambridge. ISBN: 0-521-36860-X.

VII Chapter

Conclusions

Orbital angular momentum (OAM) of light is still a relatively young area of research and although its 'foundations have been firmly established'[1] there are still a number of open questions on a fundamental level. In this thesis we have sought to investigate some of these questions. These question are not directly connected to each other and we will draw conclusions for each of them separately. In the end of this chapter we will put the results of this thesis in context which earlier and future work in this field.

In CHAPTER IV we have examined four classes of special states connected to the angular uncertainty relation. The lower bound in the uncertainty relation for angle and OAM is state dependent which requires a distinction between intelligent states, that is states which satisfy the equality in the uncertainty relation, and minimum uncertainty product states. In this thesis we have treated the angle as a variable which is strictly defined on a 2π radian interval as outlined in chapter III. Wavefunctions in the angle representation which have peaks on either edge of the allowed range can therefore be normalised. This gives rise to two varieties of states; states having an angle uncertainty smaller than $\pi/\sqrt{3}$ and states having an angle uncertainty larger than $\pi/\sqrt{3}$. The OAM eigenstates with a flat wavefunction and an angle uncertainty of $\pi/\sqrt{3}$ separate the small uncertainty states from the large uncertainty states. These states give the global minimum for the uncertainty product, but one can also identify constrained minimum uncertainty product (CMUP) states, which minimise the uncertainty product for a given uncertainty in the angle or in the OAM. In this chapter we have discussed all four classes of states, by deriving the defining equations and giving their solution. Of particular interest are the extreme cases for angle uncertainties either close to zero or close to the maximum of π. For the intelligent states the wavefunctions of these two extreme situations are very different, because intelligent states only have to be continuous at the edges of the 2π interval, but do not need to be continuously differentiable. But for the CMUP states, the first derivative has to be continuous and one might be tempted to regard the large uncertainty CMUP states as a shifted version of the small uncertainty CMUP states, but we have shown that this is not the case. The wavefunction of CMUP states with very small angle uncertainties approaches a Gaussian form, whereas the wavefunction of CMUP states with a very large uncertainty is approximatively given by the decaying part of an Airy function. This shows the importance of the choice for the particular 2π radian interval.

In CHAPTER V we have derived a criterion for the demonstration of an EPR paradox for OAM and angle. Within the derivation we have placed a particular emphasis on the fact that it is not possible to have infinitely precise measurements. For

[1] Allen et al. (2003)

the OAM these difficulties in the measurement are technical problems in a potential experiment, but for the angle it is inherently impossible to measure a single value. However, EPR paradoxes are stated in terms of apparent violations of inequalities and it is possible to demonstrate such a violation even for measurements with realistic uncertainties. This is why in the formulation of the criterion we have used angular apertures which test for a range of angles. With help of these angular apertures we were able to conceive a measurement scheme for conditional variance for the angle. The measurement is designed for a down conversion experiment, in which one angular aperture is used to set a condition on the angular position of the signal, while the angular position of the idler is measured with help of another angular aperture. By using the locality and realism arguments, akin to those occurring in EPR paradoxes, we can infer from the conditional angle variance a conditional variance for the OAM. The inferred conditional variance can be compared to a measured conditional variance for the OAM. For a criterion of an EPR paradox we have to take into account that the influence of the condition on the conditional variance is not uniform. The criterion derived in this chapter therefore also compares averaged conditional variances, where the average is taken over the local probability.

In CHAPTER VI we have found a way to describe the phenomenon of fractional OAM within the framework of the quantum theory of rotation angles reviewed in chapter III. We have generalised this theory by allowing non-integer values of the OAM. In a decomposition into angle states fractional OAM leads to a multi-valued complex factor, which we are requires the introduction of a branch cut. This leads to additional parameter in the description of the fractional states: the orientation of the complex phase discontinuity. We have shown that this phase discontinuity corresponds to an edge dislocation necessary in the creation of light beams with fractional OAM. Finally we have applied the generalised theory to numerical simulations of the propagation of light carrying fractional OAM.

At the time of this thesis the scientific field of OAM of light is almost 15 years old. In the area of optical tweezers, where light is used to trap and manipulate small particles, it has become a part of technology, not only used in physics laboratories but also in biology and other sciences. The motivation for further research in this area comes from desire or necessity to manipulate the particles in more and more sophisticated ways. Although this does not necessarily add to the fundamental knowledge on OAM, it is certainly a great success for this young area of research. This thesis, however, is concerned with the foundations of OAM of light and it is interesting to ask how the results of this thesis have enhanced our knowledge. In the opinion of the author it is very important to keep in mind that the OAM of light is a physical realisation where the variables of angle and OAM can be studied within the framework of quantum mechanics. This applies in particular to the results from chapter IV. The theoretical work on the special states connected to the angular uncertainty principle is not restricted to implementations with light carrying OAM. The form of the wavefunction for intelligent states and CMUP states will be the same regardless of the particular physical realisation. However, an experiment designed to distinguish between intelligent states and CMUP states would possibly use OAM of light. For such an experiment large uncertainty states with their enhanced difference between

intelligent states and CMUP states would be a natural choice.

Although our criterion for an angular EPR paradox in chapter V has been designed for an optical experiment, the underlying principle would still hold for a different physical system. But our motivation to pursue the idea of an angular EPR paradox is firmly rooted in the possibility to demonstrate the paradox in an optical experiment. EPR paradoxes can be constructed theoretically for all complimentary variables, but an experimental demonstration will not always be possible. It is the hope of the author that this experiment will be realised in the near future, because only then will it be possible to assess how feasible this criterion is. For future applications of this criterion as a possible signature of entanglement, it is important to keep the experimental demonstration as simple as possible. With the experience gained from performing the experiment it might be possible to construct a different signature of entanglement for OAM which is better suited for applications in quantum information.

The motivation for our work on fractional OAM in chapter VI does not come from a possible implementation with light. Some optical experiments on fractional OAM have have already been performed, but the theoretical explanation lacked a connection to the quantum theory of rotation angles in chapter III. With this connection in place fractional OAM is placed on a sound theoretical foundation. The possibilities for further research in this area are many. It already connects two very different areas: optical vortices and the study of two-photon entanglement. It would also be very interesting to use light beams with fractional OAM in connection with Bose-Einstein condensates.

Finally, we hope that this fascinating field of research remains an active one. Not only in the search of new possible applications but also on a fundamental level. Only then OAM of light will be named alongside spin angular momentum and not set against it.

A Appendix

Contour integration method to evaluate infinite sums

Contour integration can be used to sum infinite series if the member of the series fall off sufficiently quickly. Suppose, $f(z)$ is analytic at the integers $z \in \mathbb{Z}$ and approaches zero as $|z|^{-2}$ or faster for $|z| \to \infty$. Consider now the function

$$F(z) = \pi \cot(\pi z) f(z), \tag{A.1}$$

which is devised to have simple poles at the integers $n \in \mathbb{Z}$ with residues at these poles given by $f(n)$ (Stephenson & Radmore, 1993). In addition to the poles at the integers from the $\cot(\pi z)$ part, this function could also have poles from the function $f(z)$. By construction there are no poles of $f(z)$ at $z \in \mathbb{Z}$. If we integrate $F(z)$ around a square $\mathcal{S}_N$ with corners at the points $z = (N + 1/2)(\pm 1 \pm i)$ in the complex plane, then this square contains $2N + 1$ integers on the real axis. Using the Cauchy residue theorem allows us to evaluate the integral as the sum over the residues of $F(z)$ at the poles of $F(z)$ within the square $\mathcal{S}_N$:

$$\oint_{\mathcal{S}_N} F(z)dz = 2\pi i \times \begin{bmatrix} \text{sum of the residues at the poles of} \\ \pi \cot(\pi z) f(z) \text{ inside the square } \mathcal{S}_N \end{bmatrix}. \tag{A.2}$$

The poles of $\pi \cot(\pi z) f(z)$ are the integers $n \in \mathbb{Z}$ and the poles of $f(z)$. We can thus split the sum of the residues in Eq. (A.2) according to:

$$\oint_{\mathcal{S}_N} F(z)dz = 2\pi i \times \left\{ \sum_{n=-N}^{N} f(n) + \begin{bmatrix} \text{sum of the residues of } \pi \cot(\pi z) f(z) \text{ at} \\ \text{the poles of } f(z) \text{ inside the square } \mathcal{S}_N \end{bmatrix} \right\}. \tag{A.3}$$

If now the contour integration of $F(z) = \pi \cot(\pi z) f(z)$ along the square $\mathcal{S}_N$ vanishes for $N \to \infty$, the sum of the $f(n)$ and the sum over the residues in Eq. (A.3) have to add to zero. Because $|\cot(\pi z)|$ is bounded on $\mathcal{S}_N$ as $N \to \infty$ the contour integral tends to zero, as $f(z)$ falls off sufficiently quickly for $|z| \to \infty$. The infinite sum of the $f(n)$ is thus given by:

$$\sum_{n=-\infty}^{\infty} f(n) = - \begin{bmatrix} \text{sum of the residues of } \pi \cot(\pi z) f(z) \text{ at} \\ \text{the poles of } f(z) \end{bmatrix}. \tag{A.4}$$

If we want to sum a series with alternating signs we have to use a function $F(z) = \pi \operatorname{cosec}(\pi z)$ with residues at the poles $z = n \in \mathbb{N}$ given by $(-1)^n f(n)$. The contour integral tends to zero for $N \to \infty$ because $|\operatorname{cosec}(\pi z)|$ is bounded for $|z| \to \infty$. The

corresponding result to Eq. (A.4) on the previous page is

$$\sum_{n=-\infty}^{\infty} (-1)^n f(n) = -\left[\begin{array}{l}\text{sum of the residues of } \pi \operatorname{cosec}(\pi z) f(z) \\ \text{at the poles of } f(z)\end{array}\right]. \qquad \text{(A.5)}$$

The contour integration method to sum infinite series has been used widely in this thesis, most often to calculate orbital angular momentum uncertainties.

Bibliography

Abramochkin, E. & Volostnikov, V. (1991), 'Beam transformation and nontransformed beams', *Optics Communications* **83**(1,2), 123–135.

Abramowitz, M. & Stegun, I. S. (1974), *Handbook of Mathematical Functions*, Dover Publications, Inc., Mineola, New York. Reprint. Originally published: National Bureau of Standards, corrected edition, 1964. ISBN: 0-486-61272-4.

Aiello, A., Oemrawsingh, S. S. R., Eliel, E. R. & Woerdman, J. P. (2005), 'Nonlocality of high-dimensional two-photon orbital angular momentum states', *Physical Review A* **72**, 052114. DOI: 10.1103/PhysRevA.72.052114.

Albright, J. R. (1977), 'Integrals of products of Airy functions', *Journal of Physics A* **10**(4), 485–490.

Allen, L., Barnett, S. M. & Padgett, M. J. (2003), *Optical Angular Momentum*, Institute of Physics Publishing, Ltd., Bristol. ISBN: 0-7503-0901-6.

Allen, L., Beijersbergen, M. W., Spreeuw, R. J. C. & Woerdman, J. P. (1992), 'Orbital angular momentum of light and the transformation of Laguerre-Gaussian modes', *Physical Review A* **45**(11), 8185–8190. Reprinted in (Allen et al., 2003, Paper 2.1). DOI: 0.1103/PhysRevA.45.8185.

Altman, A. R., Köprülü, K. G., Corndorf, E., Kumar, P. & Barbosa, G. A. (2005), 'Quantum imaging of nonlocal spatial correlations induced by orbital angular momentum', *Physical Review Letters* **94**(12), 123601. DOI: 10.1103/PhysRevLett.94.123601.

Aragone, C., Chalbaud, E. & Salamó, S. (1976), 'On intelligent spin states', *Journal of Mathematical Physics* **17**(11), 1963–1971.

Aragone, C., Guerri, G., Salamó, S. & Tani, J. L. (1974), 'Intelligent spin states', *Journal of Physics A* **7**(15), L149–L151.

Arfken, G. (1985), *Mathematical Methods for Physicists*, 3rd edn, Academic Press, Inc., Orlando. ISBN: 0-12-059820-5.

Aspect, A., Dalibard, J. & Roger, G. (1982*a*), 'Experimental Test of Bell's Inequalities Using Time-Varying Analyzers', *Physical Review Letters* **49**(25), 1804 – 1806. DOI: 10.1103/PhysRevLett.49.1804.

Aspect, A., Grangier, P. & Roger, G. (1982*b*), 'Experimental Realization of Einstein-Podolsky-Rosen-Bohm Gedankenexperiment: A New Violation of Bell's Inequalities', *Physical Review Letters* **49**(2), 91 – 94. DOI: 10.1103/PhysRevLett.49.91.

Barbosa, G. A. & Arnaut, H. H. (2002), 'Twin photons with angular-momentum entanglement: Phase matching', *Physical Review A* **65**, 053801. DOI: 10.1103/PhysRevA.65.053801.

Barnett, S. M. (2002), 'Optical angular-momentum flux', *Journal of Optics B* **4**(2), S7–S16. DOI: 10.1088/1464-4266/4/2/361.

Barnett, S. M. & Allen, L. (1994), 'Orbital angular momentum and nonparaxial light beams', *Optics Communication* **110**(5-6), 670–678. DOI: doi:10.1016/0030-4018(94)90269-0.

Barnett, S. M. & Pegg, D. T. (1989), 'On the Hermitian optical phase operator', *Journal of Modern Optics* **36**(1), 7–19.

Barnett, S. M. & Pegg, D. T. (1990), 'Quantum theory of rotation angles', *Physical Review A* **41**(7), 3427–3435. DOI: 10.1103/PhysRevA.41.3427.

Beijersbergen, M. W., Allen, L., van der Veen, H. E. L. O. & Woerdman, J. P. (1993), 'Astigmatic laser mode converters and transfer of orbital angular momentum', *Optics Communication* **96**(1-3), 123–132. DOI: doi:10.1016/0030-4018(93)90535-D.

Beijersbergen, M. W., Coerwinkel, R. P. C., Kristensen, M. & Woerdman, J. P. (1994), 'Helical-wavefrontnext term laser beams produced with a spiral phaseplate', *Optics Communication* **112**(5-6), 321–327. DOI: 10.1016/0030-4018(94)90638-6.

Bell, J. S. (1964), 'On the Einstein-Podolsky-Rosen paradox', *Physics* **1**, 195–200. Reprinted in (Bell, 2004, Paper 2).

Bell, J. S. (1966), 'On the problem of hidden variables in quantum mechanics', *Reviews of Modern Physics* **38**(3), 447 – 452. Reprinted in (Bell, 2004, Paper 1). DOI: 10.1103/RevModPhys.38.447.

Bell, J. S. (2004), *Speakable and Unspeakable in Quantum Mechanics*, 2nd edn, Cambrige University Press, Cambridge. ISBN: 0-521-52338-9.

Berry, M. V. (1998), Paraxial beams of spinning light, *in* M. S. Soskin & M. V. Vasnetsov, eds, 'Singular optics', Vol. 3487, International Conference on Singular Optics, SPIE, pp. 6–11.

Berry, M. V. (2004), 'Optical vortices evolving from helicoidal integer and fractional phase steps', *Journal of Optics A* **6**(2), 259–269. DOI: 10.1088/1464-4285/6/2/018.

Beth, R. A. (1936), 'Mechanical detection and measurement of the angular momentum of light', *Physical Review* **50**, 115–125.

Bohm, D. (1951), *Quantum Theory*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey. Reprinted in 1989, Dover Publications, Inc., New York, ISBN: 0-486-65969-0.

Bohm, D. & Aharonov, Y. (1957), 'Discussion of Experimental Proof for the Paradox of Einstein, Rosen, and Podolsky', *Physical Review* **108**(4), 1070 – 1076. DOI: 10.1103/PhysRev.108.1070.

Born, M. & Wolf, E. (1999), *Principles of Optics*, 7th expanded edn, Cambrige University Press, Cambridge. reprinted 2003. ISBN: 0-521-64222-1.

Bouwmeester, D., Ekert, A. & Zeilinger, A., eds (2001), *The Physics of Quantum Information*, Springer Verlag, Berlin.

Carruthers, P. & Nieto, M. M. (1968), 'Phase and Angle Variables in Quantum Mechanics', *Reviews of Modern Physics* **40**(2), 411–440. DOI: 0.1103/RevModPhys.40.411.

Conway, J. B. (1978), *Functions of One Complex Variable*, no 11 *in* 'Graduate Text in Mathematics', 2nd edn, Springer Verlag.

Davis, L. W. (1979), 'Theory of electromagnetic beams', *Physical Review A* **19**(3), 1177–1179. DOI: 10.1103/PhysRevA.19.1177.

Durnin, J. (1987), 'Exact solutions for nondiffracting beams. I. The scalar theory', *Journal of the Optical Society of America A* **4**(4), 651.

Durnin, J., Micelli, J. J. & Eberly, J. H. (1987), 'Diffraction-Free Beams', *Physical Review Letters* **58**(15), 1499–1501. DOI: 10.1103/PhysRevLett.58.1499.

Einstein, A., Podolsky, B. & Rosen, N. (1935), 'Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?', *Physical Review* **47**, 777–780. DOI: http://link.aps.org/abstract/PR/v47/p777 DOI: 10.1103/PhysRev.47.777.

Enderlein, J. & Pampaloni, F. (2004), 'Unified operator approach for deriving Hermite-Gaussian and Laguerre-Gaussian laser modes', *Journal of the Optical Society of America A* **21**(8), 1553–1558.

Franke-Arnold, S., Barnett, S. M., Padgett, M. J. & Allen, L. (2003), 'Two-photon entanglement of orbital angular momentum states', *Physical Review A* **65**, 033823. DOI: 10.1103/PhysRevA.65.033823.

Franke-Arnold, S., Barnett, S. M., Yao, E., Leach, J., Courtial, J. & Padgett, M. (2004), 'Uncertainty principle for angular position and angular momentum', *New Journal of Physics* **6**, 103. DOI: 10.1088/1367-2630/6/1/103.

Galindo, A. & Pascual, P. (1990), *Quantum Mechanics*, Vol. I, Springer Verlag, Berlin. ISBN: 3-540-51406-6.

Gibson, G., Courtial, J., Padgett, M. J., Vastnetsov, V., Pas'ko, V., Barnett, S. M. & Franke-Arnold, S. (2004), 'Free-space information transfer using light beams carrying orbital angular momentum', *Optics Express* **12**(22), 5448–5456.

Gisin, N., Ribordy, G., Tittel, W. & Zbinden, H. (2002), 'Quantum cryptography', *Reviews of Modern Physics* **74**(1), 145–195. DOI: 10.1103/RevModPhys.74.145.

Goodman, J. W. (2005), *Introduction to Fourier Optics*, 3rd edn, Roberts and Company Publishers, Englewood, Colorado.

Götte, J. B., Franke-Arnold, S. & Barnett, S. M. (2006*a*), 'Angular EPR paradox', *Journal of Modern Optics* **53**(5 - 6), 627 – 645.

Götte, J. B., Radmore, P. M., Zambrini, R. & Barnett, S. M. (2006*b*), 'Angular minimum uncertainty states with large uncertainties', *Journal of Physics B* **39**, 2791–2801. DOI: doi:10.1088/0953-4075/39/12/013.

Götte, J. B., Zambrini, R., Franke-Arnold, S. & Barnett, S. M. (2005), 'Large-uncertainty intelligent states for angular momentum and angle', *Journal of Optics B* **7**, S563–S571. DOI: 10.1088/1464-4266/7/12/019.

Gouy, L. G. (1890), 'Sur une propriete nouvelle des ondes lumineuses', *Compted Rendus de l'Academie des Sciences* **110**, 1251.

Gradshteyn, I. S. & Ryzhik, I. M. (2000), *Tables of Integrals, Series, and Products*, 6th edn, Academic Press, Inc., San Diego. ISBN: 0-12-294757-6.

Gray, A. G. & Mathews, G. B. (1895), *A Treatise on Bessel functions and their Application to Physics*, Macmillan and Co., London. The authors do not use the notation of the Dirac delta distribution, which had been not yet been introduced at the time the book was published.

Gühne, O. (2004), 'Charcterizing Entanglement via Uncertainty Relations', *Physical Review Letters* **92**(11), 117903. DOI: 10.1103/PhysRevLett.92.117903.

Hahn, L. & Epstein, B. (1996), *Classical Complex Analysis*, Jones and Bartlett Publishers, Boston.

He, H., Friese, M. E. J., Heckenberg, N. R. & Rubinsztein-Dunlop, H. (1995), 'Direct Observation of Transfer of Angular Momentum to Absorptive Particles from a Laser Beam with a Phase Singularity', *Physics* **75**(5), 826–829. DOI: 10.1103/PhysRevLett.75.826.

Heisenberg, W. (1927), 'Über den anschaulichen Inhalt der quantentheoretischen Kinematik and Mechanik', *Zeitschrift für Physik* **43**, 172.

Hofmann, H. F. & Takeuchi, S. (2003), 'Violation of local uncertainty relations as a signature of entanglement', *Physical Review A* **68**, 032103. DOI: 10.1103/PhysRevA.68.032103.

Holevo, A. S. (2001), *Statistical Structure of Quantum Theory*, no Monographs 67 *in* 'Lecture Notes in Physics', Springer Verlag, Berlin. ISBN: 3-540-42082-7.

Howell, J. C., Bennink, R. S., Bentley, S. J. & Boyd, R. W. (2004), 'Realization of the Einstein-Podolsky-Rosen Paradox Using Momentum- and Position-Entangles Photons from Spontaneous Parametric Down Conversion', *Physical Review Letters* **92**(21), 210403. DOI: 10.1103/PhysRevLett.92.210403.

Jackiw, R. (1968), 'Minimum uncertianty product, number-phase uncertainty product and Coherent States', *Journal of Mathematical Physics* **9**(3), 339–346.

Jackson, J. D. (1998), *Classical Electrodynamics*, 3rd edn, John Wiley & Sons, Inc., New York. ISBN: 0-471-30932-X.

Jauch, J. M. & Piron, C. (1963), 'Can Hidden Variables be excluded in Quantum Mechanics?', *Helvetica Acta Physics* **36**, 827–837.

Judge, D. & Lewis, J. T. (1963), 'On the commutator $[L_z, \varphi]$', *Physics Letters* **5**(3), 190.

Kwiat, P. G., Mattle, K., Weinfurter, H. & Zeilinger, A. (1995), 'New High-Intensity Source of Polarization-Entangled Photon Pairs', *Physical Review Letters* **75**(24), 4337–4341. DOI: 10.1103/PhysRevLett.75.4337.

Landau, L. D. & Lifshitz, E. M. (1975), *The Classical Theory of Fields*, Vol. 2 of *Course of Theoretical Physics*, 4th revised english edition edn, Butterworth-Heinemann, Burlington.

Lax, M., Louisell, W. H. & McKnight, B. (1975), 'From Maxwell to paraxial wave optics', *Physical Review A* **11**(4), 1365–1370. DOI: 10.1103/PhysRevA.11.1365.

Leach, J., Courtial, J., Skeldon, K., Barnett, S. M., Franke-Arnold, S. & Padgett, M. J. (2004*a*), 'Interferometric Methods to Measure Orbital and Spin, or the Total Angular Momentum of a Single Photon', *Physical Review Letters* **92**, 013601. DOI: 10.1103/PhysRevLett.92.013601.

Leach, J., Padgett, M. J., Barnett, S. M., Franke-Arnold, S. & Courtial, J. (2002), 'Measuring the Orbital angular momentum of a single Photon', *Physical Review Letters* **88**, 257901. DOI: 10.1103/PhysRevLett.88.257901.

Leach, J., Yao, E. & Padgett, M. J. (2004*b*), 'Observation of the vortex structure of a non-integer vortex beam', *New Journal of Physics* **6**, 71. DOI: 10.1088/1367-2630/6/1/071.

Loudon, R. (2000), *The Quantum Theory of Light*, 3rd edn, Oxford University Press, Oxford. ISBN: 0-19-850176-5.

Louisell, W. H. (1963), 'Amplitude and phase uncertainty relations', *Physics Letters* **7**(1), 60–61. DOI: http://dx.doi.org/10.1016/0031-9163(63)90442-6.

Louisell, W. H. (1973), *Quantum Statistical Properties of Radiation*, Pure and Applied Optics, John Wiley & Sons, Inc., New York. ISBN: 0-471-54785-9.

Louisell, W. H., Yariv, A. & Siegman, A. E. (1961), 'Quantum Fluctuations and Noise in Parametric Processes. I.', *Physical Review* **124**(6), 1646–1654. DOI: 10.1103/PhysRev.124.1646.

Mair, A. E., Vaziri, A., Weihs, G. & Zeilinger, A. (2001), 'Entanglement of the orbital angular momentum states of photons', *Nature* **412**, 313–316. Reprinted in (Allen et al., 2003, Paper 8.1).

Marrucci, L., Manzo, C. & Paparo, D. (2006), 'Optical Spin-to-Orbital Angular Momentum Conversion in Inhomogeneous Anisotropic Media', *Physical Review Letters* **96**(16), 163905. DOI: 10.1103/PhysRevLett.96.163905.

Maxwell, J. C. (1873), *A Treatise on Electricity and Magnetism*, Clarendon Press Series, Oxford at the Clarendon Press, Oxford. Reprinted in Maxwell (1998).

Maxwell, J. C. (1998), *A Treatise on Electricity and Magnetism*, Oxford Classical Texts in the Physical Sciences, Oxford University Press, Oxford.

Molina-Terizza, G., Torres, J. P. & Torner, L. (2001), 'Management of the Angular Momentum of Light: Preparation of Photons in Multidimensional Vector States of Angular Momentum', *Physical Review Letters* **88**(1), 013601. DOI: 10.1103/PhysRevLett.88.013601.

Nielsen, M. A. & Chuang, I. L. (2000), *Quantum Computation and Quantum Information*, Cambridge University Press, Cambridge.

Nye, J. F. (1999), *Natural Focussing and the Fine Structure of Light*, Institute of Physics Publishing, Bristol.

Nye, J. F. & Berry, M. V. (1974), 'Dislocations in Wave Trains', *Proceedings of the Royal Society of London, Series A* **336**(1605), 165–190.

Oemrawsingh, S. S. R., Aiello, A., Eliel, E. R., Nienhaus, G. & Woerdman, J. P. (2004*a*), 'How to observe High-Dimensional Two-Photon Entanglement with Only Two Detectors', *Physical Review Letters* **92**, 217901. DOI: 10.1103/PhysRevLett.92.217901.

Oemrawsingh, S. S. R., Eliel, E. R., Woerdman, J. P., Verstegen, E. J. K., Kloosterboer, J. G. & t' Hooft, G. W. (2004*b*), 'Half-integral spiral phase plates for optical wavelengths', *Journal of Optics A* **6**, S288–290.

Oemrawsingh, S. S. R., Ma, X., Voigt, D., Aiello, A., Eliel, E. R., 't Hooft, G. W. & Woerdman, J. P. (2005), 'Experimental Demonstration of Fractional Orbital Angular Momentum Entanglement of Two Photons', *Physical Review Letters* **95**(24), 240501. DOI: 10.1103/PhysRevLett.95.240501.

O'Neil, A. T., MacVicar, I., Allen, L. & Padgett, M. J. (2002), 'Intrinsic and Extrinsic Nature of the Orbital Angular Momentum of a Light Beam', *Physical Review Letters* **88**(5), 053601. DOI: 10.1103/PhysRevLett.88.053601.

Ou, Z. Y., Pereira, S. F. & Kimble, H. J. (1982), 'Realization of the Einstein-Podolsky-Rosen paradox for continuous variables', *Physical Review Letters* **68**(25), 3663 – 3666. DOI: 10.1103/PhysRevLett.68.3663.

Paterson, C. (2005), 'Atmospheric Turbulence and Orbital Angular Momentum of Single Photons for Optical Communication', *Physical Review Letters* **94**, 153901. DOI: 10.1103/PhysRevLett.94.153901.

Pegg, D. T. & Barnett, S. M. (1989), 'Phase properties of the quantized singemode eletromagnetic field', *Physical Review A* **39**(4), 1665–1675. DOI: 10.1103/PhysRevA.39.1665.

Pegg, D. T. & Barnett, S. M. (1997), 'Tutorial review: Quantum optical phase', *Journal of Modern Optics* **44**(2), 225–264.

Pegg, D. T., Barnett, S. M., Zambrini, R., Franke-Arnold, S. & Padgett, M. (2005), 'Minimum uncertainty states of angular momentum and angular position', *New Journal of Physics* **7**, 62. DOI: 10.1088/1376-2630/7/1/062.

Pegg, D. T., Vaccaro, J. A. & Barnett, S. M. (1990), 'Quantum-optical phase and canonical conjugation', *Journal of Modern Optics* **37**(11), 1703–1710.

Peřinová, V., Lukš, A. & Peřina, J. (1998), *Phase in Optics*, Vol. 15 of *World Scientific Series in Contemporary Chemical Physics*, World Scientific, Singapore.

Poynting, J. H. (1884), 'On the transfer of energy in the electromagnetic field', *Philosophical Transactions* **175**, 343–361. Reprinted in Poynting (1920*a*).

Poynting, J. H. (1909), 'The wave motion of a revolving shaft, and a suggestion as to the angular momentum in a beam of circularly polarised light', *Proceedings of the Royal Society of London, Series A* **82**, 560–567. Reprinted in Poynting (1920*b*) and (Allen et al., 2003, Paper 1.1).

Poynting, J. H. (1920*a*), On the transfer of energy in the electromagnetic field, *in* G. A. Shakespear & G. Barlow, eds, 'Collected Scientific Papers by John Henry Poynting', Cambridge at the University Press, Cambridge.

Poynting, J. H. (1920*b*), The wave motion of a revolving shaft, and a suggestion as to the angular momentum in a beam of circularly polarised light, *in* G. A. Shakespear & G. Barlow, eds, 'Collected Scientific Papers by John Henry Poynting', Cambridge at the University Press, Cambridge.

Reid, M. (1997), 'Macroscopic elements of reality and the Einstein– Podolsky–Rosen paradox', *Quantum and Semiclassical Optics* **9**, 489–499.

Reid, M. D. (1989), 'Demonstration of the Einstein–Podolsky–Rosen paradox using nondegenerate parametric amplification', *Physical Review A* **40**(2), 913–923. DOI: 10.1103/PhysRevA.40.913.

Robertson, H. P. (1929), 'The Uncertainty Principle', *Physical Review* **34**, 163–164.

Rohrlich, F. (1970), 'Electromagnetic Momentum, Energy, and Mass', *American Journal of Physics* **38**(11), 1310–1316. DOI: 10.1119/1.1976082.

Römer, H. & Forger, M. (1993), *Klassische Feldtheorie*, VCH Verlagsgesellschaft, Weinheim.

Roux, F. S. (2003), 'Optical vortex density limitations', *Optics Communication* **223**(1-3), 31–37. DOI: 10.1016/S0030-4018(03)01626-2.

Rubin, M. H., Klyshko, D. N., Shih, Y. H. & Sergienko, A. V. (1994), 'Theory of two-photon entanglement in type-II optical parametric down-conversion', *Physical Review A* **50**(6), 5122–5133. DOI: 10.1103/PhysRevA.50.5122.

Schrödinger, E. (1930), 'Zum Heisenbergschen Unschärfeprinzip', *Sitzungsberichte der Preussischen Akademie der Wissenschaften Physikalisch-mathematische Klasse* **XIX**, 296–303.

Schwabl, F. (1995), *Quantum Mechanics*, 2nd revised edn, Springer Verlag, Berlin. ISBN: 3-540-59187-7.

Siegman, A. E. (1986), *Lasers*, University Science Books, Sausalito. ISBN: 0-935702-11-3.

Silverman, R. A. (1984), *Complex Analysis with Applications*, Dover Publications, Inc., Mineola, New York. Reprint. Originally published: Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1973, c1974. ISBN: 0-486-64762-5.

Spedalieri, F. M. (2006), 'Quantum key distribution without reference frame alignment: Exploiting photon orbital angular momentum', *Optics Communications* **260**, 340–346. DOI: 10.1016/j.optcom.2005.10.001.

Stephenson, G. & Radmore, P. M. (1993), *Advanced Mathematical Methods for Engineering and Science Students*, Cambrige University Press, Cambridge. ISBN: 0-521-36860-X.

Summy, G. S. & Pegg, D. T. (1990), 'Phase optimized quantum states of light', *Optics Communications* **77**(1), 75–79. DOI: doi:10.1016/0030-4018(90)90464-5.

Susskind, L. & Glogower, J. (1964), 'Quantum mechanical phase and time operator', *Physics* **1**, 49–61.

Tamm, C. & Weis, C. O. (1990), 'Bistability and optical switching of spatial patterns in a laser', *Journal of the Optical Society of America B* **7**(6), 1034–1038.

Vaccaro, J. A. (1995), 'Phase operators on Hilbert space', *Physical Review A* **51**(4), 3309–3317. DOI: 10.1103/PhysRevA.51.3309.

Vallée, O. & Soares, M. (2004), *Airy functions and applications to Physics*, Imperial College Press, London. ISBN: 1-86094-478-7.

Vastnetsov, M. V., Pas'ko, V. A. & Soskin, M. S. (2005), 'Analysis of orbital angular momentum of a misaligned optical beam', *New Journal of Physics*. DOI: 10.1088/1367-2630/7/1/046.

Vaziri, A., Weihs, G. & Zeilinger, A. (2002), 'Superpositions of the orbital angular momentum for applications in quantum experiments', *Journal of Optics B* **4**, S47–S51. DOI: 10.1088/1464-4266/4/2/367.

von Neumann, J. (1932), *Mathematische Grundlagen der Quantenmechanik*, Springer Verlag, Berlin. English translation: (von Neumann, 1955).

von Neumann, J. (1955), *Mathematical Foundations of Quantum Mechanics*, Princeton University Press.

Wei, H., Xue, X., Leach, J., Padgett, M. J., Barnett, S. M., Franke-Arnold, S., Yao, E. & Courtial, J. (2003), 'Simplified measurement of the orbital angular momentum of single photons', *Optics Communication* **223**(1-3), 117–122. DOI: doi:10.1016/S0030-4018(03)01619-5.

Weihs, G., Jennwein, T. & Simon, C. (1998), 'Violation of Bell's Inequality under Strict Einstein Locality Conditions', *Physical Review Letters* **81**(23), 5039 – 5043. DOI: 10.1103/PhysRevLett.81.5039.

Zambrini, R. & Barnett, S. M. (2006), 'Quasi-Intrinsic Angular Momentum and the Measurement of Its Spectrum', *Physical Review Letters* **96**, 113901. DOI: 10.1103/PhysRevLett.96.113901.