New Journal of Physics
The open access journal at the forefront of physics

Deutsche Physikalische Gesellschaft DPG IOP Institute of Physics

PAPER • **OPEN ACCESS**

Goos–Hänchen and Imbert–Fedorov shifts from a quantum-mechanical perspective

To cite this article: Falk Töppel *et al* 2013 *New J. Phys.* **15** 113059

New Journal of Physics
The open access journal for physics

Goos–Hänchen and Imbert–Fedorov shifts from a quantum-mechanical perspective

Falk Töppel[1,2,3,4], **Marco Ornigotti**[1,2] **and Andrea Aiello**[1,2]
[1] Max Planck Institute for the Science of Light, Günther-Scharowsky-Straße 1/Building 24, D-91058 Erlangen, Germany
[2] Institute for Optics, Information and Photonics, Universität Erlangen-Nürnberg, Staudtstraße 7/B2, D-91058 Erlangen, Germany
[3] Erlangen Graduate School in Advanced Optical Technologies (SAOT), Paul-Gordan-Straße 6, D-91052 Erlangen, Germany
E-mail: falk.toeppel@mpl.mpg.de

New Journal of Physics **15** (2013) 113059 (15pp)
Received 23 July 2013
Published 27 November 2013
Online at http://www.njp.org/
doi:10.1088/1367-2630/15/11/113059

Abstract. We study the classical optics effects known as Goos–Hänchen and Imbert–Fedorov shifts, occurring when reflecting a bounded light beam from a planar surface, by using a quantum-mechanical formalism. This new approach allows us to naturally separate the spatial shift into two parts, one independent on orbital angular momentum (OAM) and the other one showing OAM-induced spatial-versus-angular shift mixing. In addition, within this quantum-mechanical-like formalism, it becomes apparent that the angular shift is proportional to the beams angular spread, namely to the variance of the transverse components of the wave vector. Moreover, we extend our treatment to the enhancement of beam shifts via weak measurements and relate our results to the recent experiments.

[4] Author to whom any correspondence should be addressed.

Contents

1. Introduction

Theoretical physics is all about describing nature in terms of mathematics. However, often there is more than one way to do so. Therefore, different mathematical formalisms are frequently developed in physics to describe one and the same physical phenomenon. In fact, this is even useful since each description offers its own viewpoint onto the underlying physics and some viewpoints are more suited to observe certain details than others. Famous examples of this fact include Newtonian and Hamiltonian formulation of classical mechanics [1] or the Heisenberg and Schrödinger picture in quantum mechanics (QM) [2]. While the Newtonian mechanics relies on the system's forces, Hamiltonian mechanics rests on the system's energy. Moreover, while the Heisenberg picture resembles classical dynamics, the Schrödinger picture stresses the wave-like properties of quantum particles, e.g. electrons.

Following the idea of changing the perspective, in this work we propose a different treatment of beam shifts, by describing this purely classical phenomenon with the mathematical formalism of proper QM. Already in 1987 Player [3] used a QM-like formalism to calculate transverse beam shifts as expectation values of Hermitian operators. Earlier derivations of beam shifts using a classical treatment can be found in [4–6]. Quite later, in 2004 Onoda *et al* [7] used the concept of Berry phase and QM conservation laws to predict the existence of a Hall effect of light. A similar but distinct treatment of this phenomenon was also furnished by Bliokh and Bliokh [8, 9]. From the experimental point of view, the connection between classical beam shifts and QM was probably first exploited in 2008 by Hosten and Kwiat [10] who used a well-known quantum-weak-measurement technique to measure the spin Hall effect of light occurring in optical refraction. However, shortly afterwards Aiello and Woerdman [11] showed that such a connection has a pure formal character and that the Hosten and Kwiat experiment also admits a fully classical optics description. In 2009, Aiello *et al* [12] used again a QM formalism to illustrate the 'duality' existing between spatial and angular beam shifts and in 2010 Merano *et al* [13] had written spatial and angular real-valued physical shifts as weighted sums of real and imaginary parts of complex-valued shifts, respectively. This formalism was further elaborated in early 2012 by Aiello [14]. Finally, later in 2012 Dennis and Götte in two excellent papers [15, 16] provided for a unified view of all polarization-dependent beam shift phenomena by exploiting classical/quantum analogies.

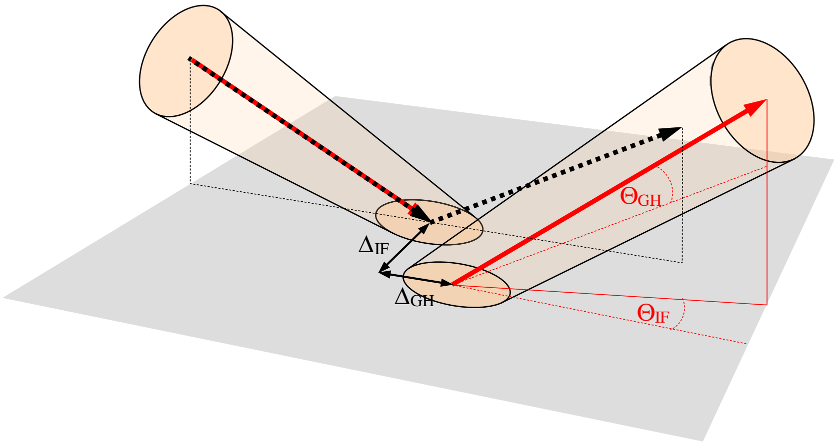

Figure 1. Visualization of the beam shifts occurring upon reflection of a bounded beam from a planar surface.

The aim of this work is to move a step forward in the 'quantum' direction by adopting an *ab initio* quantum formalism for a unified description of *all* beam shift phenomena. By exploiting the formal analogy between the paraxial wave equation and the two-dimensional Schrödinger equation [17], we can represent beam propagation as a 'time' evolution generated by the displacement operator quadratic in the 'momentum' operator and, therefore, calculate both spatial and angular shifts by using a common formalism. As will be shown later, in this manner all the beam shift phenomena will manifest a natural connection.

Beam shifts are deviations from geometric optics (ray optics) predictions that a beam with finite transverse extent experiences on reflection and/or refraction from a planar surface (see figure 1). The Goos–Hänchen (GH) [18–21] and Imbert–Fedorov (IF) shifts [22–25] are the most celebrated examples thereof (see figure 1). The spatial and angular shifts do occur in the plane of incidence (GH shifts) as well as orthogonal to the plane of incidence (IF shifts). Beam shifts have been the subject of extensive studies in the past decades both theoretically and experimentally for different kind of surfaces [26–29], different beam shapes [13, 30–32] and they have also been recently studied for the non-monochromatic case [33]. For a detailed review and further information on this topic, we refer the reader to [14, 34] and references therein.

One of the main advantages given by a quantum mechanical approach to the description of beam shift phenomena is that it furnishes in a natural manner the reason why the angular GH and IF shifts are proportional to the angular aperture of the beam, i.e. proportional to the variance of the transverse component of the wave vector. Furthermore, in this context it also clearly appears that the spatial shifts naturally separate into two terms, one independent of and the other dependent on the beams orbital angular momentum (OAM).

We proceed with the following agenda: section 2 fixes the notation for the solution of the paraxial wave equation. Hereafter, these results are translated into a quantum-mechanical notation (section 3). The main part of this work is contained in section 4. There, the problem of the beam shifts is studied in terms of the quantum-mechanical formalism derived earlier, including a description of the enhancement of the beam shifts through weak measurements. It follows section 5 where we discuss our results and relate them to the experiments on weak measurements. We conclude our paper with some final remarks in section 6.

2. Paraxial wave equation: classical optics notation

The scalar wave equation of a monochromatic electromagnetic field[5] $E(\boldsymbol{r}, t) = u(x, y, z)\, \mathrm{e}^{\mathrm{i}k(z-ct)}$ propagating in free space mainly along the z direction is well approximated by the paraxial wave equation [36]

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + 2\mathrm{i}k \frac{\partial}{\partial z} \right) u(x, y, z) = 0 \tag{1}$$

with $k = |\boldsymbol{k}| > 0$ being the modulus of the wave vector $\boldsymbol{k}$. It is well known that the normalized fundamental solution $f(x, y, z)$ of this equation is a Gaussian beam

$$f(x, y, z) = \sqrt{\frac{kL}{\pi}} \frac{1}{z - \mathrm{i}L} \exp\left(\frac{\mathrm{i}k}{2} \frac{x^2 + y^2}{z - \mathrm{i}L} \right). \tag{2}$$

Here $L > 0$ is an arbitrary length (commonly known in optics as the Rayleigh range) that fixes the width of the intensity distribution $|f(x, y, z)|^2$

$$\langle x^2 \rangle \equiv \iint x^2 \, |f(x, y, z)|^2 \, \mathrm{d}x \, \mathrm{d}y = \frac{z^2 + L^2}{2kL} = \langle y^2 \rangle, \tag{3}$$

evaluated at $z = 0$:

$$\langle x^2 \rangle \big|_{z=0} = \langle y^2 \rangle \big|_{z=0} = \frac{L}{2k} \equiv \frac{w_0^2}{4}, \tag{4}$$

where we have introduced the so-called *waist* of the beam $w_0 > 0$.

It is also useful to introduce the two-dimensional Fourier representation, i.e. the angular spectrum, of the fundamental solution $f(x, y, z)$ as follows:

$$\begin{aligned} \widetilde{f}(k_x, k_y, z) = \widetilde{f}(\boldsymbol{K}, z) &= \frac{1}{2\pi} \iint f(x, y, z) \exp\left[-\mathrm{i}(x k_x + y k_y) \right] \mathrm{d}x \, \mathrm{d}y \\ &= \mathrm{i} \sqrt{\frac{L}{\pi k}} \exp\left[-\frac{\mathrm{i}}{2k} \left(k_x^2 + k_y^2 \right) (z - \mathrm{i}L) \right], \end{aligned} \tag{5}$$

which is normalized as well.

3. Paraxial wave equation: quantum mechanics notation

It is well known that paraxial classical optics is formally equivalent to two-dimensional QM [17, 37–39] and thus, we shall rewrite all the aforementioned results in quantum-mechanical notation. However, although we are about to use the mathematical formalism of QM to describe the propagation of a paraxial beam, it is worth stressing once more that the physics underneath is purely classical, and no quantum character of the electromagnetic field is taken into account at this level of description.

Given a generic function $\psi(\boldsymbol{r}) = \psi(\boldsymbol{R}, z)$ and its two-dimensional Fourier transform $\widetilde{\psi}(\boldsymbol{K}, z)$ we define the position and momentum eigenkets $|\boldsymbol{R}\rangle$ and $|\boldsymbol{K}\rangle$, respectively, via the

[5] Please note that here we follow the notation introduced in [35], i.e. we represent three-dimensional vectors by lowercase boldface symbols, e.g. $\boldsymbol{r} = \{x, y, z\} = \{x_1, x_2, x_3\}$, and by capital boldface symbols its transverse components, e.g. $\boldsymbol{r} = \{\boldsymbol{R}, z\}$ with $\boldsymbol{R} = \{x, y\} = \{x_1, x_2\}$.

relations $\psi(\boldsymbol{R}, z) = \langle \boldsymbol{R}|\psi(z)\rangle$ and $\widetilde{\psi}(\boldsymbol{K}, z) = \langle \boldsymbol{K}|\psi(z)\rangle$. We assume that both $|\boldsymbol{R}\rangle$ and $|\boldsymbol{K}\rangle$ form a complete and orthonormal basis. Explicitly, for $|\boldsymbol{R}\rangle$ this means

$$\int |\boldsymbol{R}\rangle\langle \boldsymbol{R}| \mathrm{d}^2 R = \hat{\mathbf{1}} \quad \text{and} \quad \langle \boldsymbol{R}|\boldsymbol{R}'\rangle = \delta\left(\boldsymbol{R} - \boldsymbol{R}'\right), \tag{6}$$

where with $\hat{\mathbf{1}}$ we have denoted the identity operator. The two-dimensional Fourier transform fixes the value of $\langle \boldsymbol{R}|\boldsymbol{K}\rangle$:

$$\begin{aligned}\widetilde{\psi}(\boldsymbol{K}, z) = \langle \boldsymbol{K}|\psi(z)\rangle &= \int \langle \boldsymbol{K}|\boldsymbol{R}\rangle\langle \boldsymbol{R}|\psi(z)\rangle\, \mathrm{d}^2 R \\ &= \frac{1}{2\pi}\int \mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{R}}\, \psi(\boldsymbol{R}, z)\, \mathrm{d}x\, \mathrm{d}y,\end{aligned} \tag{7}$$

which implies $\langle \boldsymbol{K}|\boldsymbol{R}\rangle = \mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{R}}/(2\pi)$.

The position operator $\hat{\boldsymbol{R}} = \{\hat{x}, \hat{y}\} = \{\hat{x}_1, \hat{x}_2\}$ and momentum operator $\hat{\boldsymbol{K}} = \{\hat{k}_x, \hat{k}_y\} = \{\hat{k}_1, \hat{k}_2\}$ are defined via the eigenvalue equations $\hat{\boldsymbol{R}}|\boldsymbol{R}'\rangle = \boldsymbol{R}'|\boldsymbol{R}'\rangle$ and $\hat{\boldsymbol{K}}|\boldsymbol{K}'\rangle = \boldsymbol{K}'|\boldsymbol{K}'\rangle$. They fulfill the canonical commutation relations, i.e. $[\hat{x}_\alpha, \hat{k}_\beta] = \mathrm{i}\delta_{\alpha\beta}$ and $[\hat{x}_\alpha, \hat{x}_\beta] = [\hat{k}_\alpha, \hat{k}_\beta] = 0$ for $\alpha, \beta \in \{1, 2\}$.

Finally, the position operator in the momentum basis and the momentum operator in the position basis are represented by

$$\langle \boldsymbol{K}''|\hat{\boldsymbol{R}}|\boldsymbol{K}'\rangle = \mathrm{i}\frac{\partial}{\partial \boldsymbol{K}''}\delta(\boldsymbol{K}' - \boldsymbol{K}''), \tag{8a}$$

$$\langle \boldsymbol{R}''|\hat{\boldsymbol{K}}|\boldsymbol{R}'\rangle = \frac{1}{\mathrm{i}}\frac{\partial}{\partial \boldsymbol{R}''}\delta(\boldsymbol{R}' - \boldsymbol{R}''), \tag{8b}$$

respectively, where $\frac{\partial}{\partial \boldsymbol{V}}$ is a shorthand notation for the vector $\{\frac{\partial}{\partial v_1}, \frac{\partial}{\partial v_2}\}$.

In this quantum formalism, the paraxial wave equation (1) can be rewritten as

$$0 = \left[\frac{1}{2k}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) + \mathrm{i}\frac{\partial}{\partial z}\right] f(x, y, z) = \left\langle \boldsymbol{R}\left|\left(\mathrm{i}\frac{\partial}{\partial z} - \frac{1}{2k}\hat{\boldsymbol{K}}^2\right)\right| f(z)\right\rangle, \tag{9}$$

which is equivalent to the Schrödinger equation for the free propagation of a quantum particle of 'mass' k:

$$\mathrm{i}\frac{\partial}{\partial z}|f(z)\rangle = \frac{1}{2k}\hat{\boldsymbol{K}}^2|f(z)\rangle. \tag{10}$$

The formal solution of this equation can thus be written in an operator form as

$$|f(z)\rangle = \exp\left[-\frac{\mathrm{i}}{2k}\hat{\boldsymbol{K}}^2(z - z_0)\right]|f(z_0)\rangle, \tag{11}$$

where z_0 is an arbitrary real constant [17].

4. Beam shifts: from classical to quantum formalism

4.1. Quantum-mechanical representation of a reflection process

Owing to the one-to-one correspondence between the paraxial wave equation and the Schrödinger equation, the electric field $\boldsymbol{E}(\boldsymbol{r}, t)$ of a paraxial beam can be described in a way formally equivalent to the wave function of a nonrelativistic quantum-mechanical particle with spin 1/2. Let $\boldsymbol{e}_1 = \{1, 0\}$ and $\boldsymbol{e}_2 = \{0, 1\}$ be two unit vectors that span the transverse plane perpendicular to the beam propagation axis z. Then, we can write[6]

$$\boldsymbol{E}(\boldsymbol{r}, t) \propto \mathrm{e}^{\mathrm{i}(z-ct)} \sum_{\alpha=1}^{2} \boldsymbol{e}_\alpha \langle\alpha|\langle\boldsymbol{R}|\Psi(z)\rangle, \tag{12}$$

where $|\Psi(z)\rangle = |\psi(z)\rangle|A\rangle$, with $|A\rangle$ being a two-component spinor representing the polarization of the beam, i.e. $|A\rangle \equiv a_1\boldsymbol{e}_1 + a_2\boldsymbol{e}_2 = \{a_1, a_2\}$ and $|a_1|^2 + |a_2|^2 = 1$. Therefore,

$$\langle\alpha|\langle\boldsymbol{R}|\Psi(z)\rangle = \langle\boldsymbol{R}|\psi(z)\rangle\langle\alpha|A\rangle = \psi(\boldsymbol{R}, z)a_\alpha \tag{13}$$

with $\psi(\boldsymbol{R}, z)$ denoting a solution of the paraxial wave equation (1) and $|\alpha = 1\rangle \equiv \boldsymbol{e}_1 = \{1, 0\}$ as well as $|\alpha = 2\rangle \equiv \boldsymbol{e}_2 = \{0, 1\}$.

The reflection process may be described by means of the scattering (entangling) operator

$$\hat{S} = \sum_{\alpha=1}^{2} \hat{M}_{(\alpha)} \otimes \hat{P}_{(\alpha)} \tag{14}$$

where $\hat{P}_{(\alpha)}$ and $\hat{M}_{(\alpha)}$ are the polarization and the mode scattering operators associated with the considered reflection process, respectively. These operators are defined as

$$\hat{P}_{(\alpha)} = r_\alpha(\theta)|\alpha\rangle\langle\alpha| \tag{15}$$

with $r_1(\theta)$ and $r_2(\theta)$ being the Fresnel reflection coefficients evaluated at the incident angle θ (note the correspondence $1 \Leftrightarrow p$-polarization and $2 \Leftrightarrow s$-polarization) and

$$\langle\boldsymbol{R}|\hat{M}_{(\alpha)}|\psi(z)\rangle = \psi\left(-x + X_\alpha, y - Y_\alpha; z\right), \tag{16}$$

where the minus sign in the x-dependence of the shifted distribution in (16) is due to the parity inversion caused by the reflection as seen from the reflected-beam reference frame. We dedicate to this operation the 'bar' symbol: if $\mathbf{V} = \{v_x, v_y\}$ then $\bar{\boldsymbol{V}} = \{-v_x, v_y\}$. The vector state representing the electric field after reflection can be thus written as

$$\hat{S}|\Psi(z)\rangle = \sum_{\alpha=1}^{2} \hat{M}_{(\alpha)}|\psi(z)\rangle\,\hat{P}_{(\alpha)}|A\rangle = \sum_{\alpha=1}^{2} a_\alpha r_\alpha(\theta)\hat{M}_{(\alpha)}|\psi(z)\rangle|\alpha\rangle. \tag{17}$$

The four dimensionless quantities X_α and Y_α with $\alpha \in \{1, 2\}$ that appear in (16) are the *complex* shifts, whose explicit forms are given by [14]

$$X_1 = -\mathrm{i}\frac{\partial \ln r_1}{\partial\theta}, \qquad Y_1 = \mathrm{i}\frac{a_2}{a_1}\left(1 + \frac{r_2}{r_1}\right)\cot\theta, \tag{18a}$$

$$X_2 = -\mathrm{i}\frac{\partial \ln r_2}{\partial\theta}, \qquad Y_2 = -\mathrm{i}\frac{a_1}{a_2}\left(1 + \frac{r_1}{r_2}\right)\cot\theta. \tag{18b}$$

Finally, the vector wave function $\boldsymbol{\Psi}(\boldsymbol{R}, z) = \sum_{\alpha=1}^{2} \boldsymbol{e}_\alpha\langle\alpha|\langle\boldsymbol{R}|\hat{S}|\Psi(z)\rangle$ of a beam reflected by a plane surface may be written, with respect to a Cartesian reference frame attached to the reflected beam itself, as

$$\boldsymbol{\Psi}(\boldsymbol{R}, z) = \sum_{\alpha=1}^{2} \boldsymbol{e}_\alpha a_\alpha r_\alpha(\theta)\,\psi(\bar{\boldsymbol{R}} - \bar{\boldsymbol{R}}_\alpha, z), \tag{19}$$

[6] Please note that henceforth we are working in 'natural' units, where $k = 2\pi/\lambda = 1$ and [length] = [wavenumber] = 1.

where $\boldsymbol{R}_\alpha = \{X_\alpha, Y_\alpha\}$. This result coincides with the according expressions obtained in [13, 14]. The shifted function $\psi(\bar{\boldsymbol{R}} - \bar{\boldsymbol{R}}_\alpha, z)$ can be expanded in a Taylor series as follows:

$$\psi(\bar{\boldsymbol{R}} - \bar{\boldsymbol{R}}_\alpha, z) \cong \psi(\bar{\boldsymbol{R}}, z) - \boldsymbol{R}_\alpha \cdot \frac{\partial}{\partial \boldsymbol{R}}\psi(\bar{\boldsymbol{R}}, z) + \cdots. \tag{20}$$

Using this result in (19) yields

$$\boldsymbol{\Psi}(\boldsymbol{R}, z) \cong \sum_{\alpha=1}^{2} \boldsymbol{e}_\alpha a_\alpha r_\alpha(\theta)\left[1 - X_\alpha \frac{\partial}{\partial x} - Y_\alpha \frac{\partial}{\partial y}\right]\psi(\bar{\boldsymbol{R}}, z). \tag{21}$$

Here, the first term gives the geometric optics contribution to the reflected beam (it may be simply named the 'Fresnel term'), while the second and the third terms are responsible for the GH and the IF shifts, respectively.

To put (21) in a fully quantum-mechanical form, we need to introduce the three 2×2 matrices

$$\mathsf{F} \equiv \begin{bmatrix} r_1(\theta) & 0 \\ 0 & r_2(\theta) \end{bmatrix}, \tag{22a}$$

$$\mathsf{X} \equiv \begin{bmatrix} -\mathrm{i}\dfrac{\partial \ln r_1}{\partial \theta} & 0 \\ 0 & -\mathrm{i}\dfrac{\partial \ln r_2}{\partial \theta} \end{bmatrix}, \tag{22b}$$

$$\mathsf{Y} \equiv \begin{bmatrix} 0 & \mathrm{i}\left(1 + \dfrac{r_1}{r_2}\right)\cot\theta \\ -\mathrm{i}\left(1 + \dfrac{r_2}{r_1}\right)\cot\theta & 0 \end{bmatrix}, \tag{22c}$$

that represent the Fresnel reflection (F), the GH (X) and the IF (Y) shifts, respectively. Then it is not difficult to see by means of a straightforward calculation that

$$\begin{aligned} \boldsymbol{\Psi}(\boldsymbol{R}, z) &\cong \left[\mathsf{I} - \frac{\partial}{\partial x}\mathsf{X} - \frac{\partial}{\partial y}\mathsf{Y}\right] \cdot \mathsf{F} \cdot \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \psi(\bar{\boldsymbol{R}}, z) \\ &\cong \exp\left(-\hat{\boldsymbol{\mathcal{A}}} \cdot \frac{\partial}{\partial \boldsymbol{R}}\right)\psi(\bar{\boldsymbol{R}}, z) \cdot \begin{bmatrix} r_1 a_1 \\ r_2 a_2 \end{bmatrix}, \end{aligned} \tag{23}$$

where I denotes the 2×2 identity matrix and we have defined the matrix-valued 'spin operator' vector $\hat{\boldsymbol{\mathcal{A}}} = \{\mathsf{X}, \mathsf{Y}\} \equiv \{\mathsf{A}_1, \mathsf{A}_2\}$. Let $\hat{M}$ be the Hermitian and symmetric operator representing the mirror-symmetry reflection with respect to the x-axis. By definition, it acts upon the position eigenket $|\boldsymbol{R}\rangle = |x, y\rangle$ as follows: $\hat{M}|x, y\rangle = |-x, y\rangle \equiv |\bar{\boldsymbol{R}}\rangle$ and $\hat{M}^2|x, y\rangle = \hat{M}|-x, y\rangle = |x, y\rangle$, namely $\hat{M}^2 = \hat{\mathbf{1}}$. Then we can write

$$\psi(\bar{\boldsymbol{R}}, z) = \langle \bar{\boldsymbol{R}}|\psi(z)\rangle = \langle \boldsymbol{R}|\hat{M}|\psi(z)\rangle \equiv \langle \boldsymbol{R}|\bar{\psi}(z)\rangle. \tag{24}$$

With the help of (8a) and (13) it is straightforward to write (23) as

$$\boldsymbol{\Psi}(\boldsymbol{R}, z) = \langle \boldsymbol{R}|\Psi(z)\rangle = \langle \boldsymbol{R}|\exp(-\mathrm{i}\,\hat{\boldsymbol{\mathcal{A}}} \cdot \hat{\boldsymbol{K}})|\bar{\psi}(z)\rangle|A_{\mathsf{F}}\rangle, \tag{25}$$

where $|A_{\mathsf{F}}\rangle \equiv \mathsf{F}|A\rangle = \{r_1 a_1, r_2 a_2\}$. The operator $\hat{H}_{\mathrm{I}} = \hat{\boldsymbol{\mathcal{A}}} \cdot \hat{\boldsymbol{K}}$ is exactly what Hosten and Kwiat [10] call the 'interaction Hamiltonian' that couples the momentum of the meter to the 'spin observable' $\hat{\boldsymbol{\mathcal{A}}}$. However, it should be noticed that this 'Hamiltonian' is not Hermitian because $\hat{\boldsymbol{\mathcal{A}}} \neq \hat{\boldsymbol{\mathcal{A}}}^\dagger$. Therefore, the operator $\hat{\boldsymbol{\mathcal{A}}}$ does not really correspond to an observable. As a consequence of this, the operator $\exp(-\mathrm{i}\ \hat{\boldsymbol{\mathcal{A}}} \cdot \hat{\boldsymbol{K}})$ is not unitary. We shall find later that only the Hermitian combinations $\hat{\boldsymbol{\mathcal{A}}} + \hat{\boldsymbol{\mathcal{A}}}^\dagger$ and $-\mathrm{i}(\hat{\boldsymbol{\mathcal{A}}} - \hat{\boldsymbol{\mathcal{A}}}^\dagger)$ are indeed observables and yield to the measurable spatial and angular shifts, respectively.

Finally, by using (11), (25) can be recast in a fully quantum-mechanical form as

$$\begin{aligned} |\Psi(z)\rangle &= \exp(-\mathrm{i}\, \hat{\boldsymbol{\mathcal{A}}} \cdot \hat{\boldsymbol{K}}) |\bar{\psi}(z)\rangle |A_{\mathsf{F}}\rangle \\ &= \exp(-\mathrm{i}\, \hat{\boldsymbol{\mathcal{A}}} \cdot \hat{\boldsymbol{K}}) \exp\left(-\frac{\mathrm{i}}{2} \hat{\boldsymbol{K}}^2 z\right) |\bar{\psi}(0)\rangle |A_{\mathsf{F}}\rangle. \end{aligned} \tag{26}$$

It should be noticed that the free propagator $\exp(-\mathrm{i}\, \hat{\boldsymbol{K}}^2 z/2)$ and the interaction operator $\exp(-\mathrm{i}\, \hat{\boldsymbol{\mathcal{A}}} \cdot \hat{\boldsymbol{K}})$ do commute.

4.2. Ordinary (not weak) beam shifts

Beam shifts are quantified by the displacement of the centroid of the beam distribution after reflection with respect to the centroid of the reflected beam according to geometric optics. Hence, we calculate the expectation value of the position operator $\hat{\boldsymbol{R}}$ in the reference frame attached to the reflected beam, namely

$$\langle \boldsymbol{R} \rangle (z) = \frac{\langle \Psi(z) | \hat{\boldsymbol{R}} | \Psi(z) \rangle}{\langle \Psi(z) | \Psi(z) \rangle}, \tag{27}$$

where $|\Psi(z)\rangle$ has been defined for the reflected beam in (26). on using (26), the denominator of the expression above gives

$$\begin{aligned} \langle \Psi(z) | \Psi(z) \rangle &= \langle A_{\mathsf{F}} | \langle \bar{\psi}(0) | \mathrm{e}^{\frac{\mathrm{i}}{2} \hat{\boldsymbol{K}}^2 z} \mathrm{e}^{\mathrm{i} \hat{\boldsymbol{\mathcal{A}}}^\dagger \cdot \hat{\boldsymbol{K}}} \mathrm{e}^{-\mathrm{i} \hat{\boldsymbol{\mathcal{A}}} \cdot \hat{\boldsymbol{K}}} \mathrm{e}^{-\frac{\mathrm{i}}{2} \hat{\boldsymbol{K}}^2 z} | \bar{\psi}(0) \rangle | A_{\mathsf{F}} \rangle \\ &\cong \langle A_{\mathsf{F}} | \langle \bar{\psi}(0) | \left[1 - \mathrm{i} \left(\hat{\boldsymbol{\mathcal{A}}} - \hat{\boldsymbol{\mathcal{A}}}^\dagger \right) \cdot \hat{\boldsymbol{K}} + \cdots \right] | \bar{\psi}(0) \rangle | A_{\mathsf{F}} \rangle \\ &= \langle A_{\mathsf{F}} | A_{\mathsf{F}} \rangle \langle \bar{\psi}(0) | \bar{\psi}(0) \rangle + \text{quadratic terms in } \hat{\boldsymbol{\mathcal{A}}}, \end{aligned} \tag{28}$$

since $\langle \bar{\psi}(0) | \hat{\boldsymbol{K}} | \bar{\psi}(0) \rangle = 0$ because it amounts to the angular shift of the input beam which is, by definition, zero. If we assume that the input wave function is normalized, then $\langle \bar{\psi}(0) | \bar{\psi}(0) \rangle = \langle \psi(0) | \hat{M}^2 | \psi(0) \rangle = \langle \psi(0) | \psi(0) \rangle = 1$ and we can rewrite

$$\langle \Psi(z) | \Psi(z) \rangle \cong \langle A_{\mathsf{F}} | A_{\mathsf{F}} \rangle = |r_1 a_1|^2 + |r_2 a_2|^2 . \tag{29}$$

Equation (28) clearly illustrates a result that we already know from the conventional calculations, namely that the perturbative corrections to the denominator start at the second order [40].

The numerator of (27) reads as

$$\langle \Psi(z) | \hat{\boldsymbol{R}} | \Psi(z) \rangle = \langle A_{\mathsf{F}} | \langle \bar{\psi}(z) | \mathrm{e}^{\mathrm{i} \hat{\boldsymbol{\mathcal{A}}}^\dagger \cdot \hat{\boldsymbol{K}}} \hat{\boldsymbol{R}} \mathrm{e}^{-\mathrm{i} \hat{\boldsymbol{\mathcal{A}}} \cdot \hat{\boldsymbol{K}}} | \bar{\psi}(z) \rangle | A_{\mathsf{F}} \rangle, \tag{30}$$

and we can evaluate it first by noticing that

$$\mathrm{e}^{\mathrm{i} \hat{\boldsymbol{\mathcal{A}}}^\dagger \cdot \hat{\boldsymbol{K}}} \hat{\boldsymbol{R}} \mathrm{e}^{-\mathrm{i} \hat{\boldsymbol{\mathcal{A}}} \cdot \hat{\boldsymbol{K}}} = \mathrm{e}^{\mathrm{i} \hat{\boldsymbol{\mathcal{A}}}^\dagger \cdot \hat{\boldsymbol{K}}} \mathrm{e}^{-\mathrm{i} \hat{\boldsymbol{\mathcal{A}}} \cdot \hat{\boldsymbol{K}}} \left(\mathrm{e}^{\mathrm{i} \hat{\boldsymbol{\mathcal{A}}} \cdot \hat{\boldsymbol{K}}} \hat{\boldsymbol{R}} \mathrm{e}^{-\mathrm{i} \hat{\boldsymbol{\mathcal{A}}} \cdot \hat{\boldsymbol{K}}} \right) \tag{31a}$$

$$= \left(\mathrm{e}^{\mathrm{i} \hat{\boldsymbol{\mathcal{A}}}^\dagger \cdot \hat{\boldsymbol{K}}} \hat{\boldsymbol{R}} \mathrm{e}^{-\mathrm{i} \hat{\boldsymbol{\mathcal{A}}}^\dagger \cdot \hat{\boldsymbol{K}}} \right) \mathrm{e}^{\mathrm{i} \hat{\boldsymbol{\mathcal{A}}}^\dagger \cdot \hat{\boldsymbol{K}}} \mathrm{e}^{-\mathrm{i} \hat{\boldsymbol{\mathcal{A}}} \cdot \hat{\boldsymbol{K}}}. \tag{31b}$$

By using the Baker–Campbell–Hausdorff lemma [41] and the canonical commutation relation of $\hat{\boldsymbol{R}}$ and $\hat{\boldsymbol{K}}$, it is straightforward to calculate

$$e^{i\hat{\mathcal{A}}\cdot\hat{\boldsymbol{K}}}\hat{\boldsymbol{R}}e^{-i\hat{\mathcal{A}}\cdot\hat{\boldsymbol{K}}} = \hat{\boldsymbol{R}}+\hat{\mathcal{A}} \quad \text{and} \quad e^{i\hat{\mathcal{A}}^{\dagger}\cdot\hat{\boldsymbol{K}}}\hat{\boldsymbol{R}}e^{-i\hat{\mathcal{A}}^{\dagger}\cdot\hat{\boldsymbol{K}}} = \hat{\boldsymbol{R}}+\hat{\mathcal{A}}^{\dagger}. \tag{32}$$

Moreover, from the calculation of the denominator we already know that

$$e^{i\hat{\mathcal{A}}^{\dagger}\cdot\hat{\boldsymbol{K}}}e^{-i\hat{\mathcal{A}}\cdot\hat{\boldsymbol{K}}} \cong 1-i\left(\hat{\mathcal{A}}-\hat{\mathcal{A}}^{\dagger}\right)\cdot\hat{\boldsymbol{K}}+\cdots. \tag{33}$$

By inserting (32) and (33) into (31a) we obtain

$$\begin{aligned} e^{i\hat{\mathcal{A}}^{\dagger}\cdot\hat{\boldsymbol{K}}}\hat{\boldsymbol{R}}\,e^{-i\hat{\mathcal{A}}\cdot\hat{\boldsymbol{K}}} &\cong \left[1-i\left(\hat{\mathcal{A}}-\hat{\mathcal{A}}^{\dagger}\right)\cdot\hat{\boldsymbol{K}}+\cdots\right]\left(\hat{\boldsymbol{R}}+\hat{\mathcal{A}}\right) \\ &\cong \hat{\boldsymbol{R}}+\hat{\mathcal{A}}-i\left[\left(\hat{\mathcal{A}}-\hat{\mathcal{A}}^{\dagger}\right)\cdot\hat{\boldsymbol{K}}\right]\hat{\boldsymbol{R}}+\cdots \end{aligned} \tag{34a}$$

and, similarly, from (31b) we attain

$$\begin{aligned} e^{i\hat{\mathcal{A}}^{\dagger}\cdot\hat{\boldsymbol{K}}}\hat{\boldsymbol{R}}\,e^{-i\hat{\mathcal{A}}\cdot\hat{\boldsymbol{K}}} &\cong \left(\hat{\boldsymbol{R}}+\hat{\mathcal{A}}^{\dagger}\right)\left[1-i\left(\hat{\mathcal{A}}-\hat{\mathcal{A}}^{\dagger}\right)\cdot\hat{\boldsymbol{K}}+\cdots\right] \\ &\cong \hat{\boldsymbol{R}}+\hat{\mathcal{A}}^{\dagger}-i\,\hat{\boldsymbol{R}}\left[\left(\hat{\mathcal{A}}-\hat{\mathcal{A}}^{\dagger}\right)\cdot\hat{\boldsymbol{K}}\right]+\cdots, \end{aligned} \tag{34b}$$

where the quadratic and higher order terms in $\hat{\mathcal{A}}$ and $\hat{\mathcal{A}}^{\dagger}$ have been discarded in the last lines since the shifts X_α and Y_α with $\alpha \in \{1, 2\}$ are supposed to be small. The approximations (34a) and (34b) are not Hermitian, although the left sides of these equations are. To obtain a Hermitian quantity for the approximation of the left-hand side of (31a), we take a symmetric combination of (34a) and (34b). By substituting this into the numerator of (30) we find

$$\begin{aligned} \langle\Psi(z)|\hat{\boldsymbol{R}}|\Psi(z)\rangle &\cong \langle\bar{\psi}(z)|\hat{\boldsymbol{R}}|\bar{\psi}(z)\rangle\langle A_{\mathsf{F}}|A_{\mathsf{F}}\rangle+\frac{1}{2}\langle\bar{\psi}(z)|\bar{\psi}(z)\rangle\langle A_{\mathsf{F}}|\left(\hat{\mathcal{A}}+\hat{\mathcal{A}}^{\dagger}\right)|A_{\mathsf{F}}\rangle \\ &\quad -\frac{i}{2}\langle\bar{\psi}(z)|\left(\hat{\boldsymbol{K}}\hat{\boldsymbol{R}}+\hat{\boldsymbol{R}}\hat{\boldsymbol{K}}\right)|\bar{\psi}(z)\rangle\cdot\langle A_{\mathsf{F}}|\left(\hat{\mathcal{A}}-\hat{\mathcal{A}}^{\dagger}\right)|A_{\mathsf{F}}\rangle \\ &= \mathrm{Re}\langle A_{\mathsf{F}}|\hat{\mathcal{A}}|A_{\mathsf{F}}\rangle+\langle\bar{\psi}(z)|\left(\hat{\boldsymbol{K}}\hat{\boldsymbol{R}}+\hat{\boldsymbol{R}}\hat{\boldsymbol{K}}\right)|\bar{\psi}(z)\rangle\cdot\mathrm{Im}\langle A_{\mathsf{F}}|\hat{\mathcal{A}}|A_{\mathsf{F}}\rangle, \end{aligned} \tag{35}$$

where $\langle\bar{\psi}(z)|\bar{\psi}(z)\rangle = \langle\psi(z)|\psi(z)\rangle = 1$ and $\langle\bar{\psi}(z)|\hat{\boldsymbol{R}}|\bar{\psi}(z)\rangle = \langle\bar{\psi}(0)|\hat{\boldsymbol{R}}+z\hat{\boldsymbol{K}}|\bar{\psi}(0)\rangle = 0$ has been used. The latter equality is due to the result

$$e^{\frac{i}{2}\hat{\boldsymbol{K}}^2 z}\hat{\boldsymbol{R}}e^{-\frac{i}{2}\hat{\boldsymbol{K}}^2 z} = \hat{\boldsymbol{R}}+z\hat{\boldsymbol{K}}, \tag{36}$$

and the fact that the spatial and the angular shifts of the input beam are zero by definition. Note that in (35), the dyadic operator $\hat{\boldsymbol{K}}\hat{\boldsymbol{R}}+\hat{\boldsymbol{R}}\hat{\boldsymbol{K}}$ can be represented by a 2×2 matrix by recalling that $[\hat{\boldsymbol{K}}\hat{\boldsymbol{R}}+\hat{\boldsymbol{R}}\hat{\boldsymbol{K}}]_{\alpha\beta} = \hat{k}_\alpha\hat{x}_\beta+\hat{x}_\alpha\hat{k}_\beta$:

$$\hat{\boldsymbol{K}}\hat{\boldsymbol{R}}+\hat{\boldsymbol{R}}\hat{\boldsymbol{K}} = \begin{bmatrix} \hat{k}_x\hat{x}+\hat{x}\hat{k}_x & \hat{x}\hat{k}_y+\hat{y}\hat{k}_x \\ \hat{x}\hat{k}_y+\hat{y}\hat{k}_x & \hat{k}_y\hat{y}+\hat{y}\hat{k}_y \end{bmatrix}. \tag{37}$$

This matrix formulation is equivalent to equations (19) and (20) in [14] when calculating the shift of an OAM beam. This can be shown explicitly by calculating the expectation value

$$\begin{aligned} \langle\bar{\psi}(z)|\hat{\boldsymbol{R}}\hat{\boldsymbol{K}}|\bar{\psi}(z)\rangle &= \int d^2R\int d^2R'\langle\bar{\psi}(z)|\hat{\boldsymbol{R}}|\boldsymbol{R}'\rangle\langle\boldsymbol{R}'|\hat{\boldsymbol{K}}|\boldsymbol{R}\rangle\langle\boldsymbol{R}|\bar{\psi}(z)\rangle \\ &= -i\int d^2R\,\psi^*(\bar{\boldsymbol{R}},z)\boldsymbol{R}\frac{\partial}{\partial\boldsymbol{R}}\psi(\bar{\boldsymbol{R}},z) \end{aligned} \tag{38}$$

and therefore

$$\langle\bar{\psi}(z)|\left(\hat{\boldsymbol{K}}\hat{\boldsymbol{R}}+\hat{\boldsymbol{R}}\hat{\boldsymbol{K}}\right)|\bar{\psi}(z)\rangle = 2\mathrm{Re}\langle\bar{\psi}(z)|\hat{\boldsymbol{R}}\hat{\boldsymbol{K}}|\bar{\psi}(z)\rangle$$
$$= 2\mathrm{Im}\int \mathrm{d}^2R\,\psi^*(\bar{\boldsymbol{R}},z)\boldsymbol{R}\frac{\partial}{\partial\boldsymbol{R}}\psi(\bar{\boldsymbol{R}},z). \tag{39}$$

Thus, it is not by chance that in [14] the off diagonal matrix entries calculated for an OAM beam are proportional to the angular momentum carried by the beam itself. It is clear from (39) that the off-diagonal elements of $\hat{\boldsymbol{K}}\hat{\boldsymbol{R}}+\hat{\boldsymbol{R}}\hat{\boldsymbol{K}}$, namely $\hat{x}\hat{k}_y+\hat{y}\hat{k}_x$, are proportional to the z component of the angular momentum operator.

The z-dependence in the term $\langle\bar{\psi}(z)|\hat{\boldsymbol{K}}\hat{\boldsymbol{R}}+\hat{\boldsymbol{R}}\hat{\boldsymbol{K}}|\bar{\psi}(z)\rangle$ may be explicitly calculated:

$$\langle\bar{\psi}(z)|\hat{k}_\alpha\hat{x}_\beta|\bar{\psi}(z)\rangle = \langle\bar{\psi}(0)|\mathrm{e}^{\frac{\mathrm{i}}{2}\hat{\boldsymbol{K}}^2 z}\hat{k}_\alpha\hat{x}_\beta\mathrm{e}^{-\frac{\mathrm{i}}{2}\hat{\boldsymbol{K}}^2 z}|\bar{\psi}(0)\rangle$$
$$= \langle\bar{\psi}(0)|\hat{k}_\alpha(\hat{x}_\beta+z\,\hat{k}_\beta)|\bar{\psi}(0)\rangle$$
$$= \langle\bar{\psi}(0)|\hat{k}_\alpha\hat{x}_\beta|\bar{\psi}(0)\rangle + z\,\langle\bar{\psi}(0)|\hat{k}_\alpha\hat{k}_\beta|\bar{\psi}(0)\rangle. \tag{40}$$

Similarly

$$\langle\bar{\psi}(z)|\hat{x}_\alpha\hat{k}_\beta|\bar{\psi}(z)\rangle = \langle\bar{\psi}(0)|\hat{x}_\alpha\hat{k}_\beta|\bar{\psi}(0)\rangle + z\,\langle\bar{\psi}(0)|\hat{k}_\alpha\hat{k}_\beta|\bar{\psi}(0)\rangle, \tag{41}$$

where the Baker–Campbell–Hausdorff lemma has been used once again. Thus, (35) can be rewritten as

$$\langle\Psi(z)|\hat{\boldsymbol{R}}|\Psi(z)\rangle \cong \mathrm{Re}\langle A_{\mathsf{F}}|\hat{\boldsymbol{\mathcal{A}}}|A_{\mathsf{F}}\rangle + \langle\bar{\psi}(0)|\hat{\boldsymbol{K}}\hat{\boldsymbol{R}}+\hat{\boldsymbol{R}}\hat{\boldsymbol{K}}|\bar{\psi}(0)\rangle\cdot\mathrm{Im}\langle A_{\mathsf{F}}|\hat{\boldsymbol{\mathcal{A}}}|A_{\mathsf{F}}\rangle$$
$$+z\,\langle\bar{\psi}(0)|2\hat{\boldsymbol{K}}\hat{\boldsymbol{K}}|\bar{\psi}(0)\rangle\mathrm{Im}\langle A_{\mathsf{F}}|\hat{\boldsymbol{\mathcal{A}}}|A_{\mathsf{F}}\rangle. \tag{42}$$

The angular shift Θ contribution comes from the term proportional to z, whereas the spatial shift Δ is independent of z (see figure 1). Therefore, the first line of (42) gives the spatial shift. However, there are two contributions to the spatial shift. The first term of the first line of (42), proportional to the real part of $\hat{\boldsymbol{\mathcal{A}}}$, gives a spatial shift that only depends on the polarization properties of the beam and (through the matrix F) on the properties of the surface. Conversely, the second term on the same line, proportional to the imaginary part of $\hat{\boldsymbol{\mathcal{A}}}$, depends also on the spatial properties of the beam and yields the spatial-versus-angular shift mixing occurring, for example, for OAM beams. Such OAM-induced beam shifts were predicted theoretically by Fedoseyev [42, 43] and by Bliokh *et al* [44] and in the case of the IF shift observed experimentally by Dasgupta and Gupta [45]. Finally, the second line of the equation above gives the angular shift (z-dependent part of the total shift). We find that it is always proportional to the angular spread of the beam because it amounts to the momentum self-correlation matrix $\hat{\boldsymbol{K}}\hat{\boldsymbol{K}}$ whose diagonal elements give indeed the angular spread of the incident beam. Thus, the beams angular aperture is proportional to the variance of the transverse component of the k-vector.

To illustrate (42) in greater detail let us calculate it for the specific case of an input fundamental Gaussian beam of the form (2), namely for

$$\langle\boldsymbol{R}|\bar{\psi}(0)\rangle = \frac{\mathrm{i}}{\sqrt{\pi L}}\exp\left(-\frac{x^2+y^2}{2L}\right) \tag{43}$$

and

$$\langle\boldsymbol{K}|\bar{\psi}(0)\rangle = \mathrm{i}\sqrt{\frac{L}{\pi}}\exp\left(-\frac{k_x^2+k_y^2}{2/L}\right). \tag{44}$$

A straightforward calculation furnishes

$$\langle\bar{\psi}(0)|\hat{\boldsymbol{K}}\hat{\boldsymbol{R}}+\hat{\boldsymbol{R}}\hat{\boldsymbol{K}}|\bar{\psi}(0)\rangle=0 \quad \text{and} \quad \langle\bar{\psi}(0)|\hat{\boldsymbol{K}}\hat{\boldsymbol{K}}|\bar{\psi}(0)\rangle=\frac{1}{2L}\mathbf{1}. \tag{45a}$$

Finally, by using these results (42) reduces to

$$\langle\Psi(z)|\hat{\boldsymbol{R}}|\Psi(z)\rangle=\mathrm{Re}\langle A_{\mathsf{F}}|\hat{\boldsymbol{\mathcal{A}}}|A_{\mathsf{F}}\rangle+\frac{z}{L}\mathrm{Im}\langle A_{\mathsf{F}}|\hat{\boldsymbol{\mathcal{A}}}|A_{\mathsf{F}}\rangle. \tag{46}$$

This clear result beautifully illustrates how the real and the imaginary part of the interaction operator $\hat{\boldsymbol{\mathcal{A}}}$ yield to the spatial and the angular shifts, respectively. At the end of the day, by gathering all the results we can write

$$\frac{\langle\Psi(z)|\hat{\boldsymbol{R}}|\Psi(z)\rangle}{\langle\Psi(z)|\Psi(z)\rangle}=\mathrm{Re}\frac{\langle A_{\mathsf{F}}|\hat{\boldsymbol{\mathcal{A}}}|A_{\mathsf{F}}\rangle}{\langle A_{\mathsf{F}}|A_{\mathsf{F}}\rangle}+\frac{z}{L}\,\mathrm{Im}\frac{\langle A_{\mathsf{F}}|\hat{\boldsymbol{\mathcal{A}}}|A_{\mathsf{F}}\rangle}{\langle A_{\mathsf{F}}|A_{\mathsf{F}}\rangle}, \tag{47}$$

which, in terms of the matrices X and Y given by equations (22*b*) and (22*c*), respectively, may be rewritten as

$$\Delta_{\mathrm{GH}}=\mathrm{Re}\frac{\langle A_{\mathsf{F}}|\mathsf{X}|A_{\mathsf{F}}\rangle}{\langle A_{\mathsf{F}}|A_{\mathsf{F}}\rangle}, \qquad \Theta_{\mathrm{GH}}=\frac{1}{L}\,\mathrm{Im}\frac{\langle A_{\mathsf{F}}|\mathsf{X}|A_{\mathsf{F}}\rangle}{\langle A_{\mathsf{F}}|A_{\mathsf{F}}\rangle}, \tag{48a}$$

$$\Delta_{\mathrm{IF}}=\mathrm{Re}\frac{\langle A_{\mathsf{F}}|\mathsf{Y}|A_{\mathsf{F}}\rangle}{\langle A_{\mathsf{F}}|A_{\mathsf{F}}\rangle}, \qquad \Theta_{\mathrm{IF}}=\frac{1}{L}\,\mathrm{Im}\frac{\langle A_{\mathsf{F}}|\mathsf{Y}|A_{\mathsf{F}}\rangle}{\langle A_{\mathsf{F}}|A_{\mathsf{F}}\rangle}. \tag{48b}$$

These expressions are fully coincidental with the ones obtained by means of ordinary (classical) calculations for an input fundamental Gaussian beam [34].

4.3. *Weak measurements*

Theory [11, 15, 16, 46] points out a close connection between the beam shifts and the weak measurements [47, 48]. In the experiments [10, 49] weak measurements are frequently applied to enhance and thus observe beam shift effects. For this reason, we will extend our formalism now to weak measurements. To begin with, let us rewrite (26) as

$$|\Psi(z)\rangle=\exp(-\mathrm{i}\,\hat{\boldsymbol{\mathcal{A}}}\cdot\hat{\boldsymbol{K}})|\bar{\psi}(z)\rangle|A_{\mathsf{F}}\rangle, \tag{49}$$

which describes the beam up to the detector surface. Now, imagine putting in front of the detector a polarizer oriented along the direction $|B\rangle=b_1\boldsymbol{e}_1+b_2\boldsymbol{e}_2$, with $|b_1|^2+|b_2|^2=1$. As a consequence, the polarization of the beam will be projected along this direction and the resulting state will be

$$|\Psi(z)\rangle\rightarrow|B\rangle\langle B|\Psi(z)\rangle\equiv|\psi_B(z)\rangle|B\rangle, \tag{50}$$

where $|B\rangle\equiv\{b_1,b_2\}$ and

$$|\psi_B(z)\rangle=\langle B|\exp(-\mathrm{i}\,\hat{\boldsymbol{\mathcal{A}}}\cdot\hat{\boldsymbol{K}})|\bar{\psi}(z)\rangle|A_{\mathsf{F}}\rangle. \tag{51}$$

As usual, in the hypothesis of weak perturbation we can expand the exponential to obtain

$$\begin{aligned}|\psi_B(z)\rangle&\cong\langle B|(1-\mathrm{i}\,\hat{\boldsymbol{\mathcal{A}}}\cdot\hat{\boldsymbol{K}}+\cdots)|\bar{\psi}(z)\rangle|A_{\mathsf{F}}\rangle\\&=\langle B|A_{\mathsf{F}}\rangle\left[|\bar{\psi}(z)\rangle-\mathrm{i}\frac{\langle B|\hat{\boldsymbol{\mathcal{A}}}|A_{\mathsf{F}}\rangle}{\langle B|A_{\mathsf{F}}\rangle}\cdot\hat{\boldsymbol{K}}|\bar{\psi}(z)\rangle+\cdots\right]\\&\cong\langle B|A_{\mathsf{F}}\rangle\exp(-\mathrm{i}\,\boldsymbol{\mathcal{A}}_w\cdot\hat{\boldsymbol{K}})|\bar{\psi}(z)\rangle,\end{aligned} \tag{52}$$

where we have defined the vector-valued weak value

$$\boldsymbol{\mathcal{A}}_w = \frac{\langle B|\hat{\boldsymbol{\mathcal{A}}}|A_{\mathsf{F}}\rangle}{\langle B|A_{\mathsf{F}}\rangle} \tag{53}$$

with a post-selection probability $P_{\mathrm{ps}} = |\langle B|A_{\mathsf{F}}\rangle|^2/\langle A_{\mathsf{F}}|A_{\mathsf{F}}\rangle$. Since $\boldsymbol{\mathcal{A}}_w \in \mathbb{C}$ this quantity is not directly observable. However, it is simply related to the spatial and angular beam shifts as it can be seen by calculating explicitly the post-selected wave function in the position representation:

$$\begin{aligned}
\langle \boldsymbol{R}|\psi_B(z)\rangle &= \langle B|A_{\mathsf{F}}\rangle\langle \boldsymbol{R}|\exp(-\mathrm{i}\,\boldsymbol{\mathcal{A}}_w\cdot\hat{\boldsymbol{K}})|\bar{\psi}(z)\rangle \\
&= \langle B|A_{\mathsf{F}}\rangle \int \langle \boldsymbol{R}|\boldsymbol{K}\rangle\langle \boldsymbol{K}|\exp(-\mathrm{i}\,\boldsymbol{\mathcal{A}}_w\cdot\hat{\boldsymbol{K}})|\bar{\psi}(z)\rangle\,\mathrm{d}^2K \\
&= \frac{\langle B|A_{\mathsf{F}}\rangle}{2\pi}\int \exp(\mathrm{i}\boldsymbol{R}\cdot\boldsymbol{K})\exp(-\mathrm{i}\,\boldsymbol{\mathcal{A}}_w\cdot\boldsymbol{K})\langle \boldsymbol{K}|\bar{\psi}(z)\rangle\,\mathrm{d}^2K \\
&= \frac{\langle B|A_{\mathsf{F}}\rangle}{2\pi}\int \exp\big[\mathrm{i}\boldsymbol{K}\cdot(\boldsymbol{R}-\boldsymbol{\mathcal{A}}_w)\big]\widetilde{\psi}(\bar{\boldsymbol{K}},z)\,\mathrm{d}^2K \\
&= \langle B|A_{\mathsf{F}}\rangle\,\psi(\bar{\boldsymbol{R}}-\bar{\boldsymbol{\mathcal{A}}}_w,z).
\end{aligned} \tag{54}$$

At this point one may proceed in the usual manner by calculating the centroid of the shifted distribution $|\psi(\bar{\boldsymbol{R}}-\bar{\boldsymbol{\mathcal{A}}}_w,z)|^2$. Alternatively, we can start from (52) and calculate directly $\langle \boldsymbol{R}\rangle(z)$ from (27). By proceeding in exactly the same manner as before, first we find $\langle\psi_B(z)|\psi_B(z)\rangle = |\langle B|A_{\mathsf{F}}\rangle|^2 = |\langle B|\mathsf{F}|A\rangle|^2$. Then, we calculate the equivalent of (42) which yields to our final result

$$\begin{aligned}
\frac{\langle\psi_B(z)|\hat{\boldsymbol{R}}|\psi_B(z)\rangle}{\langle\psi_B(z)|\psi_B(z)\rangle} &= \frac{\boldsymbol{\mathcal{A}}_w+\boldsymbol{\mathcal{A}}_w^{\dagger}}{2} + \langle\bar{\psi}(0)|\hat{\boldsymbol{K}}\hat{\boldsymbol{R}}+\hat{\boldsymbol{R}}\hat{\boldsymbol{K}}|\bar{\psi}(0)\rangle\cdot\frac{\boldsymbol{\mathcal{A}}_w-\boldsymbol{\mathcal{A}}_w^{\dagger}}{2\mathrm{i}} \\
&+z\,\langle\bar{\psi}(0)|2\hat{\boldsymbol{K}}\hat{\boldsymbol{K}}|\bar{\psi}(0)\rangle\cdot\frac{\boldsymbol{\mathcal{A}}_w-\boldsymbol{\mathcal{A}}_w^{\dagger}}{2\mathrm{i}}.
\end{aligned} \tag{55}$$

Explicitly, for an input fundamental Gaussian beam, the post-selected GH and IF shifts are thus

$$\Delta_{\mathrm{GH}}^{\mathrm{ps}} = \mathrm{Re}\frac{\langle B|\mathsf{X}|A_{\mathsf{F}}\rangle}{\langle B|A_{\mathsf{F}}\rangle}, \qquad \Theta_{\mathrm{GH}}^{\mathrm{ps}} = \frac{1}{L}\,\mathrm{Im}\frac{\langle B|\mathsf{X}|A_{\mathsf{F}}\rangle}{\langle B|A_{\mathsf{F}}\rangle}, \tag{56a}$$

$$\Delta_{\mathrm{IF}}^{\mathrm{ps}} = \mathrm{Re}\frac{\langle B|\mathsf{Y}|A_{\mathsf{F}}\rangle}{\langle B|A_{\mathsf{F}}\rangle}, \qquad \Theta_{\mathrm{IF}}^{\mathrm{ps}} = \frac{1}{L}\,\mathrm{Im}\frac{\langle B|\mathsf{Y}|A_{\mathsf{F}}\rangle}{\langle B|A_{\mathsf{F}}\rangle}. \tag{56b}$$

5. Discussions

In the previous section, we have derived the following expressions for the (complex-valued) post-selection enhanced GH and IF shifts:

$$X \equiv \frac{\langle B|\mathsf{X}|A_{\mathsf{F}}\rangle}{\langle B|A_{\mathsf{F}}\rangle} = -\mathrm{i}\frac{b_1^*\frac{\partial r_1}{\partial\theta}a_1 + b_2^*\frac{\partial r_2}{\partial\theta}a_2}{b_1^* r_1 a_1 + b_2^* r_2 a_2}, \tag{57a}$$

$$Y \equiv \frac{\langle B|\mathsf{Y}|A_{\mathsf{F}}\rangle}{\langle B|A_{\mathsf{F}}\rangle} = \mathrm{i}(r_1+r_2)\cot\theta\,\frac{b_1^* a_2 - b_2^* a_1}{b_1^* r_1 a_1 + b_2^* r_2 a_2}. \tag{57b}$$

First of all, we notice that when either $a_1 = 0$ or $a_2 = 0$ one has

$$X|_{a_1=0} = -\frac{\mathrm{i}}{r_2}\frac{\partial r_2}{\partial \theta}, \qquad X|_{a_2=0} = -\frac{\mathrm{i}}{r_1}\frac{\partial r_1}{\partial \theta}, \tag{58a}$$

$$Y|_{a_1=0} = \mathrm{i}\frac{b_1^*}{b_2^*}\left(1+\frac{r_1}{r_2}\right)\cot\theta, \quad Y|_{a_2=0} = -\mathrm{i}\frac{b_2^*}{b_1^*}\left(1+\frac{r_2}{r_1}\right)\cot\theta. \tag{58b}$$

From (58*a*) it follows that the 'ordinary' GH shift, either spatial or angular, cannot be enhanced if one has either s $(a_1 = 0)$ or p $(a_2 = 0)$ input polarization. In contrast, for the same kind of input states the IF shift can be enhanced. For example, consider the experiment by Hosten and Kwiat [10], where the input is polarized horizontally ($|A\rangle = |H\rangle \Leftrightarrow a_2 = 0$) and post-selected with linear polarization: $|B\rangle = -\sin\Delta|H\rangle + \cos\Delta|V\rangle = b_1|\alpha = 1\rangle + b_2|\alpha = 2\rangle$. In this case, we obtain from the second equation of (58*b*):

$$Y|_{a_2=0} = \mathrm{i}\cot\Delta\left(1+\frac{r_2}{r_1}\right)\cot\theta, \tag{59}$$

which grows indefinitely when $\Delta \to 0$. In the case of real r_1 and r_2, Y becomes purely imaginary and the angular shift is therefore magnified, in agreement with Hosten and Kwiats experiment. Equations (57*a*) and (57*b*) are strictly valid until the denominators do not vanish. However, this is precisely what happens at Brewster's and null-reflection angles, as described with great detail in [50]. In those conditions, (57*a*) and (57*b*) cease to be valid near the singularity and the higher terms disregarded in (28), (34*a*) and (35) contribute [51].

Next, we discuss the question what is in general the best choice for the post-selection state in these weak measurements. The expressions in (57*a*) and (57*b*) are clearly singular when $\langle B|A_\mathsf{F}\rangle = 0$ which occurs for

$$\begin{aligned} |B\rangle : &= \{-r_2^* a_2^*, r_1^* a_1^*\} \\ &\equiv |A_\mathsf{F}^\perp\rangle. \end{aligned} \tag{60}$$

Hence, in principle, one could choose $|B\rangle \simeq |A_\mathsf{F}^\perp\rangle$ in order to increase the magnitude of X, but at the same time keep it finite. However, this is only part of the story because the choice of $|B\rangle$ also affects the value of the numerator in (57*a*) and (57*b*). Thus, one could choose $|B\rangle = \mathsf{U}|A_\mathsf{F}^\perp\rangle$, where U is an arbitrary 2×2 matrix (with $\mathsf{U} \neq \mathsf{I}$ in order to avoid the singularity) and use the three real parameters on which U may depend, in order to maximize (numerically) X and Y. This is far too complicated and hence we use a single real parameter, say $\Delta \in [0, 2\pi)$, and choose

$$|B\rangle = \cos\Delta|A_\mathsf{F}^\perp\rangle - \sin\Delta|A_\mathsf{F}\rangle, \tag{61}$$

in order to have a huge enhancement for $\Delta \simeq 0$. With this choice the post-selection probability

$$P_{\mathrm{ps}} = \frac{|\langle B|A_\mathsf{F}\rangle|^2}{\langle A_\mathsf{F}|A_\mathsf{F}\rangle} = \sin^2\Delta \tag{62}$$

is independent from the angle of incidence of the beam and can be kept constant during the experiment. The intensity of the beam behind the post-selecting polarizer is simply equal to $I_{\mathrm{ps}} = P_{\mathrm{ps}}\langle A_\mathsf{F}|A_\mathsf{F}\rangle = \sin^2\Delta(|r_1 a_1|^2 + |r_2 a_2|^2)$.

From (58*a*), (58*b*) and (61) it follows that, in general, the enhanced shift can always be written as the sum of the ordinary shift plus an enhancement term:

$$S = \frac{\langle B|\mathsf{S}|A_\mathsf{F}\rangle}{\langle B|A_\mathsf{F}\rangle} = \frac{\langle A_\mathsf{F}|\mathsf{S}|A_\mathsf{F}\rangle}{\langle A_\mathsf{F}|A_\mathsf{F}\rangle} + \frac{\langle B|A_\mathsf{F}^\perp\rangle}{\langle B|A_\mathsf{F}\rangle}\frac{\langle A_\mathsf{F}^\perp|\mathsf{S}|A_\mathsf{F}\rangle}{\langle A_\mathsf{F}|A_\mathsf{F}\rangle}$$
$$= \frac{\langle A_\mathsf{F}|\mathsf{S}|A_\mathsf{F}\rangle}{\langle A_\mathsf{F}|A_\mathsf{F}\rangle} - \cot\Delta\,\frac{\langle A_\mathsf{F}^\perp|\mathsf{S}|A_\mathsf{F}\rangle}{\langle A_\mathsf{F}|A_\mathsf{F}\rangle}, \tag{63}$$

where $S \in \{X, Y\}$, $\mathsf{S} \in \{\mathsf{X}, \mathsf{Y}\}$ and the completeness relation $|A_\mathsf{F}\rangle\langle A_\mathsf{F}| + |A_\mathsf{F}^\perp\rangle\langle A_\mathsf{F}^\perp| = \langle A_\mathsf{F}|A_\mathsf{F}\rangle\hat{\mathbf{1}}$ has been used. Note that since $\langle A_\mathsf{F}|A_\mathsf{F}\rangle = |r_1 a_1|^2 + |r_2 a_2|^2 = \langle A_\mathsf{F}^\perp|A_\mathsf{F}^\perp\rangle$, the second term in (63) can be written in the following more symmetric form:

$$\frac{\langle A_\mathsf{F}^\perp|\mathsf{S}|A_\mathsf{F}\rangle}{\langle A_\mathsf{F}|A_\mathsf{F}\rangle} = \frac{\langle A_\mathsf{F}^\perp|}{\sqrt{\langle A_\mathsf{F}^\perp|A_\mathsf{F}^\perp\rangle}}\,\mathsf{S}\,\frac{|A_\mathsf{F}\rangle}{\sqrt{\langle A_\mathsf{F}|A_\mathsf{F}\rangle}} \tag{64}$$

Furthermore, an explicit calculation gives

$$X = -\mathrm{i}\left[\frac{|a_1|^2\, r_1^* \frac{\partial r_1}{\partial\theta} + |a_2|^2\, r_2^* \frac{\partial r_2}{\partial\theta}}{|a_1 r_1|^2 + |a_2 r_2|^2} - \cot\Delta\,\frac{a_1 a_2\left(r_1\frac{\partial r_2}{\partial\theta} - r_2\frac{\partial r_1}{\partial\theta}\right)}{|a_1 r_1|^2 + |a_2 r_2|^2}\right], \tag{65a}$$

$$Y = \mathrm{i}(r_1 + r_2)\cot\theta\left[\frac{a_2 a_1^* r_1^* - a_1 a_2^* r_2^*}{|a_1 r_1|^2 + |a_2 r_2|^2} + \cot\Delta\,\frac{a_1^2 r_1 + a_2^2 r_2}{|a_1 r_1|^2 + |a_2 r_2|^2}\right]. \tag{65b}$$

We find that the enhancement term depends, in general, on the angle of incidence (θ), on the polarization of the state (a_1 and a_2) as well as on the properties of the reflecting surface (r_1 and r_2).

6. Conclusions

In this work, we have derived the spatial and angular GH and IF shifts of a beam with finite transversal extent after reflection using a quantum-mechanical notation. Studying these classical effects through the glasses of QM gave some new insights. Our main result is equation (42). It proves that the spatial shift consist of two parts, one showing spatial-versus-angular shift mixing occurring, for example, for OAM beams. Furthermore, it becomes apparent that the angular shift is proportional to the beams angular spread (variance of the transverse component of the wave vector). Moreover, we studied the enhancement of beam shifts due to weak measurements and related our results to the seminal experiment of Hosten and Kwiat. The results presented here are in full agreement with the ones presented by Dennis and Götte [15] and Götte and Dennis [16].

References

[1] Goldstein H, Poole C and Safko L 1980 *Classical Mechanics* 3rd edn (San Francisco, CA: Addison-Wesley)
[2] Sakurai J J and Napolitano J 2007 *Modern Quantum Mechanics* 2nd edn (San Francisco, CA: Pearson Education)
[3] Player M A 1987 *J. Phys. A: Math. Gen.* **20** 3667–78
[4] Schilling H 1965 *Ann. Phys.* **16** 122–34
[5] Fedoseev V G 1985 *Opt. Spectrosc. (USSR)* **58** 296–8

[6] Liberman V S and Zel'dovich B Ya 1992 *Phys. Rev.* A **46** 5199–207
[7] Onoda M, Murakami S and Nagaosa N 2004 *Phys. Rev. Lett.* **93** 083901
[8] Bliokh K Y and Bliokh Y P 2006 *Phys. Rev. Lett.* **96** 073903
[9] Bliokh K Y and Bliokh Y P 2007 *Phys. Rev.* E **75** 066609
[10] Hosten O and Kwiat P 2008 *Science* **319** 787–90
[11] Aiello A and Woerdman J P 2008 *Opt. Lett.* **33** 1437–9
[12] Aiello A, Merano M and Woerdman J P 2009 *Phys. Rev.* A **80** 061801
[13] Merano M, Hermosa N, Woerdman J P and Aiello A 2010 *Phys. Rev.* A **82** 023817
[14] Aiello A 2012 *New J. Phys.* **14** 013058
[15] Dennis M R and Götte J B 2012 *New J. Phys.* **14** 073013
[16] Götte J B and Dennis M R 2012 *New J. Phys.* **14** 073016
[17] Stoler D 1981 *J. Opt. Soc. Am.* **71** 334–41
[18] Goos F and Hänchen H 1947 *Ann. Phys.* **1** 333–46
[19] Artmann K 1948 *Ann. Phys.* **2** 87–102
[20] Chiu K W and Quinn J J 1972 *Am. J. Phys.* **40** 1847–51
[21] McGuirk M and Carniglia C K 1977 *J. Opt. Soc. Am.* **67** 103–7
[22] Fedorov F I 1955 *Dokl. Akad. Nauk SSSR* **105** 465–8
[23] Imbert C 1972 *Phys. Rev.* D **5** 787–96
[24] Costa de, Beauregard O and Imbert C 1972 *Phys. Rev. Lett.* **28** 1211–3
[25] Pillon F, Gilles H and Girard S 2004 *Appl. Opt.* **43** 1863–9
[26] Leung P T, Chen C W and Chiang H P 2007 *Opt. Commun.* **276** 206–8
[27] Merano M, Aiello A, 't Hooft G W, von Exter M P, Elier E R and Woerdman J P 2007 *Opt. Express* **15** 15928–34
[28] Tamir T 1986 *J. Opt. Soc. Am.* A **3** 558–65
[29] Landry G D and Maldonado T A 1996 *Appl. Opt.* **35** 5870–9
[30] Kozaki S and Sakurai H 1978 *J. Opt. Soc. Am.* **68** 508–14
[31] Golla D and Dutta Gupta S 2011 Goos–Hänchen shift for higher order Hermite–Gaussian beams arXiv:1011.3968v1
[32] Aiello A and Woerdman J P 2010 *Opt. Lett.* **36** 543–5
[33] Ornigotti M and Aiello A 2013 *J. Opt.* **15** 014004
[34] Bliokh K Y and Aiello A 2013 *J. Opt.* **15** 014001
[35] Berry M V 2009 *J. Opt. A: Pure Appl. Opt.* **11** 094001
[36] Siegman A E 1986 *Lasers* (Mill Valley: University Science Books)
[37] Meron M, Viccaro P J and Lin B 1999 *Phys. Rev.* E **59** 7152–5
[38] Kogelnik H and Li T 1966 *Appl. Opt.* **5** 1550–67
[39] Nienhuis G and Allen L 1993 *Phys. Rev.* A **48** 656–65
[40] Aiello A and Woerdman J P 2009 Theory of angular Goos–Hänchen shift near Brewster incidence arXiv:0903.3730
[41] Merzbacher E 1998 *Quantum Mechanics* 3rd edn (Hoboken: Wiley)
[42] Fedoseyev V G 2001 *Opt. Commun.* **193** 9–18
[43] Fedoseyev V G 2008 *J. Phys. A: Math. Theor.* **41** 505202
[44] Bliokh K Y, Shadrivov I V and Kivshar Y S 2009 *Opt. Lett.* **34** 389–91
[45] Dasgupta R and Gupta P K 2006 *Opt. Commun.* **257** 91–6
[46] Jozsa R 2007 *Phys. Rev.* A **76** 044103
[47] Aharonov Y, Albert D Z and Vaidman L 1988 *Phys. Rev. Lett.* **60** 1351–4
[48] Ritchie N W M, Story J G and Hulet R G 1991 *Phys. Rev. Lett.* **66** 1107–10
[49] Qin Y, Li Y, Feng X, Xiao Y-F, Yang H and Gong Q 2011 *Opt. Express* **19** 9636–45
[50] Götte J B and Dennis M R 2013 *Opt. Lett.* **38** 2295–7
[51] Töppel F, Ornigotti M and Aiello A 2014 in preparation