Bi-orthonormal sets of Gaussian-type modes

2005 J. Phys. A: Math. Gen. 38 9931

(http://iopscience.iop.org/0305-4470/38/46/003)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND GENERAL

J. Phys. A: Math. Gen. **38** (2005) 9931–9939 doi:10.1088/0305-4470/38/46/003

Bi-orthonormal sets of Gaussian-type modes

Martin J Bastiaans[1] **and Tatiana Alieva**[2]

[1] Faculteit Elektrotechniek, Technische Universiteit Eindhoven, Postbus 513, 5600 MB Eindhoven, Netherlands
[2] Facultad de Ciencias Físicas, Universidad Complutense de Madrid, Ciudad Universitaria s/n, Madrid 28040, Spain

E-mail: m.j.bastiaans@tue.nl and talieva@fis.ucm.es

Received 31 May 2005, in final form 23 September 2005
Published 2 November 2005
Online at stacks.iop.org/JPhysA/38/9931

Abstract
Based on the recently introduced orthonormal Hermite–Gaussian-type modes, a general class of sets of non-orthonormal Gaussian-type modes is introduced, along with their associated bi-orthonormal partner sets. The conditions between these two bi-orthonormal sets of modes have been derived, expressed in terms of their generating functions, and the relations with Wünsche's Hermite two-dimensional functions and the two-variable Hermite polynomials have been established. A closed-form expression for Gaussian-type modes is derived from their derivative and recurrence relations, which result from the generating function. It is shown that the evolution of non-orthonormal Gaussian-type modes under linear canonical transformations can be described by the same mechanism as used for the evolution of orthonormal Hermite–Gaussian-type modes, when, simultaneously, the associated bio-orthonormal modes are taken into account.

PACS numbers: 02.30.Nw, 02.30.Uu, 42.30.Kq, 42.60.Jf

1. Introduction

Hermite and Laguerre polynomials and their associated Hermite–Gaussian and Laguerre–Gaussian functions—or modes—are widely used in physics and information processing. Schemes to convert Hermite–Gaussian into Laguerre–Gaussian modes by means of appropriate linear canonical transformations, are well known; see [1, 2], for instance, in which mode converters were presented in the field of optics. In some recent papers, generalizations of these polynomials and their associated Gaussian functions were proposed; we mention Wünsche's Hermite and Laguerre two-dimensional polynomials and functions [3] and Abramochkin's Hermite–Laguerre–Gaussian modes [4]. For all these cases we have a generating function that has a Gaussian form.

In this paper, we propose a unified approach for the description of all polynomials and functions that are characterized by a Gaussian-type generating function, leading to a general class of sets of Gaussian-type modes; the polynomials and functions mentioned above, then appear as special cases. The general class contains not only the sets of orthonormal Hermite–Gaussian-type modes [5, 6]—with Hermite–Gaussian, Laguerre–Gaussian and Hermite–Laguerre–Gaussian modes as examples—but also includes mode sets that are not orthonormal. It will be shown that, in the non-orthonormal case, any set of modes has an associated bi-orthonormal partner set from the same class; in the orthonormal case, this bi-orthonormal partner set is then simply identical to the original set.

From the generating function, we will construct derivative relations and recurrence relations between Gaussian-type modes, and from these we will derive a closed-form expression for them. Furthermore, the evolution of Gaussian-type modes under linear canonical transformation is studied in phase space.

2. Orthonormal Hermite–Gaussian-type modes

The recently introduced general class of sets of orthonormal modes [5, 6] $\mathcal{H}_{n,m}(\boldsymbol{r};\boldsymbol{K},\boldsymbol{L})$, which we called Hermite–Gaussian-type modes because they can be generated by an appropriate linear canonical transformation of the Hermite–Gaussian modes, was defined by the generating function

$$2^{1/2}(\det\boldsymbol{K})^{1/2}\exp(-\boldsymbol{s}^{\mathrm{t}}\boldsymbol{M}\boldsymbol{s}+2\sqrt{2\pi}\,\boldsymbol{s}^{\mathrm{t}}\boldsymbol{K}\boldsymbol{r}-\pi\boldsymbol{r}^{\mathrm{t}}\boldsymbol{L}\boldsymbol{r}) = \sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\mathcal{H}_{n,m}(\boldsymbol{r};\boldsymbol{K},\boldsymbol{L})\left(\frac{2^{n+m}}{n!m!}\right)^{1/2}s_x^n s_y^m, \tag{1}$$

where we have introduced the column vector $\boldsymbol{s}=(s_x,s_y)^{\mathrm{t}}$ and three (possibly complex) 2×2 matrices $\boldsymbol{K}$, $\boldsymbol{L}=\boldsymbol{L}^{\mathrm{t}}$ and $\boldsymbol{M}=\boldsymbol{M}^{\mathrm{t}}$, and where the superscript $^{\mathrm{t}}$ denotes transposition. For the common Hermite–Gaussian modes, separable in x and y, we thus have

$$\boldsymbol{K}=\begin{pmatrix} w_x^{-1} & 0\\ 0 & w_y^{-1}\end{pmatrix},\qquad \boldsymbol{L}=\begin{pmatrix} w_x^{-2} & 0\\ 0 & w_y^{-2}\end{pmatrix},\qquad \boldsymbol{M}=\mathbf{I}. \tag{2}$$

As another example, we mention the common Laguerre–Gaussian modes, see, for instance, [2], for which we have

$$\boldsymbol{K}=\frac{1}{w\sqrt{2}}\begin{pmatrix} 1 & \mathrm{i}\\ \mathrm{i} & 1\end{pmatrix},\qquad \boldsymbol{L}=\frac{1}{w^2}\mathbf{I},\qquad \boldsymbol{M}=\begin{pmatrix} 0 & \mathrm{i}\\ \mathrm{i} & 0\end{pmatrix}. \tag{3}$$

The Hermite–Laguerre–Gaussian modes [4] arise for

$$\boldsymbol{K}=\frac{\mathrm{i}}{w\sqrt{2}}(\boldsymbol{M}-\mathbf{I}),\qquad \boldsymbol{L}=\frac{1}{w^2}\mathbf{I},\qquad \boldsymbol{M}=-\mathrm{i}\begin{pmatrix} \cos 2\alpha & \sin 2\alpha\\ \sin 2\alpha & -\cos 2\alpha\end{pmatrix}, \tag{4}$$

and reduce to Laguerre–Gaussian modes again for $\alpha=\pi/4$ and to Hermite–Gaussian modes for $\alpha=0$, although in a slightly modified form. Since we restricted ourselves in [5] and [6] to orthonormal modes, the symmetric matrix $\boldsymbol{M}$ was completely defined by the non-singular matrix $\boldsymbol{K}$,

$$\boldsymbol{M}=\boldsymbol{K}\boldsymbol{K}^{*-1}, \tag{5}$$

with the superscript * denoting complex conjugation, and so was the (positive-definite) real part $\mathrm{Re}\,\boldsymbol{L}$ of the symmetric matrix $\boldsymbol{L}$,

$$\mathrm{Re}\,\boldsymbol{L}=\frac{\boldsymbol{L}+\boldsymbol{L}^*}{2}=\boldsymbol{K}^{\mathrm{t}}\boldsymbol{K}^*. \tag{6}$$

For this reason we did not need to include $\boldsymbol{M}$ as a parameter in $\mathcal{H}_{n,m}(\boldsymbol{r};\boldsymbol{K},\boldsymbol{L})$: the set is completely determined by the matrices $\boldsymbol{K}$ (non-singular) and $\boldsymbol{L}$ (symmetric, and with its positive-definite real part related to $\boldsymbol{K}$ through equation (6)).

3. Bi-orthonormal Gaussian-type modes

To extend the class of sets of orthonormal modes $\mathcal{H}_{n,m}(\boldsymbol{r};\boldsymbol{K},\boldsymbol{L})$ to a more general class of sets $\mathcal{H}_{n,m}(\boldsymbol{r};\boldsymbol{K},\boldsymbol{L},\tilde{\boldsymbol{L}})$ that are no longer orthonormal, we will consider simultaneously a second set of non-orthonormal modes from the same class, which is bi-orthonormal to the first one. So we have the bi-orthonormality condition

$$\int\int_{-\infty}^{\infty}\mathcal{H}_{n,m}(\boldsymbol{r};\boldsymbol{K}_1,\boldsymbol{L}_1,\tilde{\boldsymbol{L}}_1)\mathcal{H}^*_{l,k}(\boldsymbol{r};\boldsymbol{K}_2,\boldsymbol{L}_2,\tilde{\boldsymbol{L}}_2)\,\mathrm{d}\boldsymbol{r}=\delta_{nl}\delta_{mk} \tag{7}$$

for the set $\mathcal{H}_{n,m}(\boldsymbol{r};\boldsymbol{K}_1,\boldsymbol{L}_1,\tilde{\boldsymbol{L}}_1)$ and its partner set $\mathcal{H}_{l,k}(\boldsymbol{r};\boldsymbol{K}_2,\boldsymbol{L}_2,\tilde{\boldsymbol{L}}_2)$, which by themselves are still defined through the generating function (1). Note that—instead of including the symmetric matrix $\boldsymbol{M}$ as a parameter—we have added a symmetric matrix $\tilde{\boldsymbol{L}}$ as an additional parameter in $\mathcal{H}_{n,m}(\boldsymbol{r};\boldsymbol{K},\boldsymbol{L},\tilde{\boldsymbol{L}})$, where $\tilde{\boldsymbol{L}}$ is related to $\boldsymbol{K}$, $\boldsymbol{L}$ and $\boldsymbol{M}$ as

$$\boldsymbol{L}+\tilde{\boldsymbol{L}}^*=2\boldsymbol{K}^{\mathrm{t}}\boldsymbol{M}^{-1}\boldsymbol{K}. \tag{8}$$

The reason for using $\tilde{\boldsymbol{L}}$ instead of $\boldsymbol{M}$ will become clear later; in particular, we will see that the real part of $\boldsymbol{L}+\tilde{\boldsymbol{L}}$ should be positive definite, the condition which is more explicit than the implicit condition for $\boldsymbol{M}$ that $\boldsymbol{K}^{\mathrm{t}}\boldsymbol{M}^{-1}\boldsymbol{K}$ should have a positive-definite real part. As these modes can no longer be generated by a linear canonical transformation of the common Hermite–Gaussian modes, we drop the word Hermitian and we will call them Gaussian-type modes, since all of them can be described by a generating function that has a Gaussian form. We recall that knowledge of the associated bi-orthonormal partner set $\mathcal{H}_{n,m}(\boldsymbol{r};\boldsymbol{K}_2,\boldsymbol{L}_2,\tilde{\boldsymbol{L}}_2)$ is valuable when we want to expand an arbitrary function $f(\boldsymbol{r})$ in terms of the modes $\mathcal{H}_{n,m}(\boldsymbol{r};\boldsymbol{K}_1,\boldsymbol{L}_1,\tilde{\boldsymbol{L}}_1)$ and we need to find the expansion coefficients c_{nm}:

$$f(\boldsymbol{r})=\sum_{n=0}^{\infty}\sum_{m=0}^{\infty}c_{nm}\mathcal{H}_{n,m}(\boldsymbol{r};\boldsymbol{K}_1,\boldsymbol{L}_1,\tilde{\boldsymbol{L}}_1), \tag{9}$$

$$c_{nm}=\int\int_{-\infty}^{\infty}f(\boldsymbol{r})\mathcal{H}^*_{n,m}(\boldsymbol{r};\boldsymbol{K}_2,\boldsymbol{L}_2,\tilde{\boldsymbol{L}}_2)\,\mathrm{d}\boldsymbol{r}. \tag{10}$$

To find the relations between $\boldsymbol{K}_1,\boldsymbol{L}_1,\boldsymbol{M}_1$ and $\boldsymbol{K}_2,\boldsymbol{L}_2,\boldsymbol{M}_2$, required by the bi-orthonormality condition (7), we consider the expression

$$\begin{aligned}J&=\sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\sum_{l=0}^{\infty}\sum_{k=0}^{\infty}\left(\frac{2^{n+m}}{n!m!}\right)^{1/2}s_x^ns_y^m\left(\frac{2^{l+k}}{l!k!}\right)^{1/2}t_x^lt_y^k\\&\quad\times\int\int_{-\infty}^{\infty}\mathcal{H}_{n,m}(\boldsymbol{r};\boldsymbol{K}_1,\boldsymbol{L}_1,\tilde{\boldsymbol{L}}_1)\mathcal{H}^*_{l,k}(\boldsymbol{r};\boldsymbol{K}_2,\boldsymbol{L}_2,\tilde{\boldsymbol{L}}_2)\,\mathrm{d}\boldsymbol{r}\\&=2(\det\boldsymbol{K}_1)^{1/2}(\det\boldsymbol{K}_2^*)^{1/2}\exp(-\boldsymbol{s}^{\mathrm{t}}\boldsymbol{M}_1\boldsymbol{s}-\boldsymbol{t}^{\mathrm{t}}\boldsymbol{M}_2^*\boldsymbol{t})\\&\quad\times\int\int_{-\infty}^{\infty}\exp[2\sqrt{2\pi}(\boldsymbol{s}^{\mathrm{t}}\boldsymbol{K}_1+\boldsymbol{t}^{\mathrm{t}}\boldsymbol{K}_2^*)\boldsymbol{r}-\pi\boldsymbol{r}^{\mathrm{t}}(\boldsymbol{L}_1+\boldsymbol{L}_2^*)\boldsymbol{r}]\,\mathrm{d}\boldsymbol{r},\end{aligned} \tag{11}$$

where $\boldsymbol{t}=(t_x,t_y)^{\mathrm{t}}$ in analogy with $\boldsymbol{s}=(s_x,s_y)^{\mathrm{t}}$, and where we have substituted from the generating function (1). The integral in this expression is of the form

$$\int\int_{-\infty}^{\infty}\exp(-\pi\boldsymbol{r}^{\mathrm{t}}\boldsymbol{P}\boldsymbol{r}-\mathrm{i}2\pi\boldsymbol{r}^{\mathrm{t}}\boldsymbol{q})\,\mathrm{d}\boldsymbol{r}=\frac{1}{\sqrt{\det\boldsymbol{P}}}\exp(-\pi\boldsymbol{q}^{\mathrm{t}}\boldsymbol{P}^{-1}\boldsymbol{q}), \tag{12}$$

which is a straightforward extension to more dimensions of a similar relation in the one-dimensional case; see, for instance, [7, equation (2.3.15.11)]. In equation (12) $\boldsymbol{P} = \boldsymbol{L}_1 + \boldsymbol{L}_2^*$ is a symmetric matrix and $\boldsymbol{q} = \mathrm{i}\sqrt{2/\pi}\left(\boldsymbol{K}_1^{\mathrm{t}}\boldsymbol{s} + \boldsymbol{K}_2^{*\mathrm{t}}\boldsymbol{t}\right)$ is a column vector; for the integral to converge, we have to require that the real part of $\boldsymbol{P}$ be positive definite, which at the same time makes $\boldsymbol{P}$ non-singular, see appendix A. The integral in equation (11) then leads directly to

$$[\det(\boldsymbol{L}_1 + \boldsymbol{L}_2^*)]^{-1/2} \exp\left\{\boldsymbol{s}^{\mathrm{t}}\left[\boldsymbol{K}_1\left(\frac{\boldsymbol{L}_1+\boldsymbol{L}_2^*}{2}\right)^{-1}\boldsymbol{K}_1^{\mathrm{t}}\right]\boldsymbol{s}\right\}$$
$$\times \exp\left\{\boldsymbol{t}^{\mathrm{t}}\left[\boldsymbol{K}_2\left(\frac{\boldsymbol{L}_2+\boldsymbol{L}_1^*}{2}\right)^{-1}\boldsymbol{K}_2^{\mathrm{t}}\right]^*\boldsymbol{t}\right\}$$
$$\times \exp\left\{2\boldsymbol{s}^{\mathrm{t}}\left[\boldsymbol{K}_1\left(\frac{\boldsymbol{L}_1+\boldsymbol{L}_2^*}{2}\right)^{-1}\boldsymbol{K}_2^{*\mathrm{t}}\right]\boldsymbol{t}\right\}$$

and we get

$$J = 2(\det \boldsymbol{K}_1)^{1/2}(\det \boldsymbol{K}_2^*)^{1/2}[\det(\boldsymbol{L}+\boldsymbol{L}^*)]^{-1/2}$$
$$\times \exp\left\{-\boldsymbol{s}^{\mathrm{t}}\left[\boldsymbol{M}_1 - \boldsymbol{K}_1\left(\frac{\boldsymbol{L}_1+\boldsymbol{L}_2^*}{2}\right)^{-1}\boldsymbol{K}_1^{\mathrm{t}}\right]\boldsymbol{s}\right\}$$
$$\times \exp\left\{-\boldsymbol{t}^{\mathrm{t}}\left[\boldsymbol{M}_2 - \boldsymbol{K}_2\left(\frac{\boldsymbol{L}_2+\boldsymbol{L}_1^*}{2}\right)^{-1}\boldsymbol{K}_2^{\mathrm{t}}\right]^*\boldsymbol{t}\right\}$$
$$\times \exp\left\{2\boldsymbol{s}^{\mathrm{t}}\left[\boldsymbol{K}_1\left(\frac{\boldsymbol{L}_1+\boldsymbol{L}_2^*}{2}\right)^{-1}\boldsymbol{K}_2^{*\mathrm{t}}\right]\boldsymbol{t}\right\}. \tag{13}$$

To get to the bi-orthonormality condition (7), we have to require that

$$\boldsymbol{M}_1 - \boldsymbol{K}_1\left(\frac{\boldsymbol{L}_1+\boldsymbol{L}_2^*}{2}\right)^{-1}\boldsymbol{K}_1^{\mathrm{t}} = \boldsymbol{0}, \tag{14}$$

$$\boldsymbol{M}_2 - \boldsymbol{K}_2\left(\frac{\boldsymbol{L}_2+\boldsymbol{L}_1^*}{2}\right)^{-1}\boldsymbol{K}_2^{\mathrm{t}} = \boldsymbol{0}, \tag{15}$$

$$\boldsymbol{K}_1\left(\frac{\boldsymbol{L}_1+\boldsymbol{L}_2^*}{2}\right)^{-1}\boldsymbol{K}_2^{*\mathrm{t}} = \mathbf{I}. \tag{16}$$

Note that bi-orthonormality imposes the additional conditions that $\boldsymbol{K}$ and $\boldsymbol{M}$ are non-singular and that the symmetric matrix $\boldsymbol{K}^{\mathrm{t}}\boldsymbol{M}^{-1}\boldsymbol{K}$ has a positive-definite real part. From equations (14)–(16) we easily verify the explicit relations between $\boldsymbol{K}_2, \boldsymbol{L}_2, \boldsymbol{M}_2$ and $\boldsymbol{K}_1, \boldsymbol{L}_1, \boldsymbol{M}_1$:

$$\boldsymbol{K}_2^* = \boldsymbol{M}_1^{-1}\boldsymbol{K}_1, \tag{17}$$

$$\boldsymbol{L}_2^* = 2\boldsymbol{K}_1^{\mathrm{t}}\boldsymbol{M}_1^{-1}\boldsymbol{K}_1 - \boldsymbol{L}_1, \tag{18}$$

$$\boldsymbol{M}_2^{-1} = \boldsymbol{M}_1^*. \tag{19}$$

We are now prepared to identify the associated matrices $\tilde{\boldsymbol{L}}_1$ and $\tilde{\boldsymbol{L}}_2$ in $\mathcal{H}_{n,m}(\boldsymbol{r}; \boldsymbol{K}_1, \boldsymbol{L}_1, \tilde{\boldsymbol{L}}_1)$ and $\mathcal{H}_{l,k}(\boldsymbol{r}; \boldsymbol{K}_2, \boldsymbol{L}_2, \tilde{\boldsymbol{L}}_2)$ with $\boldsymbol{L}_2$ and $\boldsymbol{L}_1$, respectively, with the additional condition that the real part of $\boldsymbol{L} + \tilde{\boldsymbol{L}}$ be positive definite. The two sets of modes $\mathcal{H}_{n,m}(\boldsymbol{r}; \boldsymbol{K}_1, \boldsymbol{L}_1, \boldsymbol{L}_2)$ and

Table 1. Bi-orthonormal sets of Gaussian-type modes $\mathcal{H}_{n,m}(\boldsymbol{r};\boldsymbol{K}_1,\boldsymbol{L}_1,\boldsymbol{L}_2)$ and $\mathcal{H}_{l,k}(\boldsymbol{r};\boldsymbol{K}_2,\boldsymbol{L}_2,\boldsymbol{L}_1)$: (a) orthonormal Hermite–Gaussian-type modes, with $2\boldsymbol{K}^{\mathrm{t}}\boldsymbol{K}^* = \boldsymbol{L}+\boldsymbol{L}^*$; (b) common Hermite–Gaussian modes (separable in x and y); (c) bi-orthonormal Gaussian-type modes, with $2\tilde{\boldsymbol{K}}^{\mathrm{t}}\boldsymbol{K}^* = \boldsymbol{L}+\tilde{\boldsymbol{L}}^*$; (d) Wünsche's Hermite two-dimensional functions, with $\tilde{\boldsymbol{K}}^{\mathrm{t}}\boldsymbol{K}^* = \mathbf{I}$; (e) two-variable Hermite polynomials, with $\boldsymbol{K}=\boldsymbol{K}^{\mathrm{t}}$; (f) common Hermite polynomials (separable in x and y).

	$\mathcal{H}_{n,m}(\boldsymbol{r};\boldsymbol{K}_1,\boldsymbol{L}_1,\boldsymbol{L}_2)$	$\mathcal{H}_{l,k}(\boldsymbol{r};\boldsymbol{K}_2,\boldsymbol{L}_2,\boldsymbol{L}_1)$	$\boldsymbol{M}_1$	$\boldsymbol{M}_2$
(a)	$\mathcal{H}_{n,m}(\boldsymbol{r};\boldsymbol{K},\boldsymbol{L},\boldsymbol{L})$	$\mathcal{H}_{l,k}(\boldsymbol{r};\boldsymbol{K},\boldsymbol{L},\boldsymbol{L})$	$\boldsymbol{K}\boldsymbol{K}^{*-1}$	$\boldsymbol{K}\boldsymbol{K}^{*-1}$
(b)	$\mathcal{H}_{n,m}(\boldsymbol{r};\mathbf{I},\mathbf{I},\mathbf{I})$	$\mathcal{H}_{l,k}(\boldsymbol{r};\mathbf{I},\mathbf{I},\mathbf{I})$	$\mathbf{I}$	$\mathbf{I}$
(c)	$\mathcal{H}_{n,m}(\boldsymbol{r};\boldsymbol{K},\boldsymbol{L},\tilde{\boldsymbol{L}})$	$\mathcal{H}_{l,k}(\boldsymbol{r};\tilde{\boldsymbol{K}},\tilde{\boldsymbol{L}},\boldsymbol{L})$	$\boldsymbol{K}(\tilde{\boldsymbol{K}}^{\mathrm{t}}\boldsymbol{K}^*)^{-1}\boldsymbol{K}^{\mathrm{t}}$	$\tilde{\boldsymbol{K}}(\boldsymbol{K}^{\mathrm{t}}\tilde{\boldsymbol{K}}^*)^{-1}\tilde{\boldsymbol{K}}^{\mathrm{t}}$
(d)	$\mathcal{H}_{n,m}(\boldsymbol{r};\boldsymbol{K},\mathbf{I},\mathbf{I})$	$\mathcal{H}_{l,k}(\boldsymbol{r};\tilde{\boldsymbol{K}},\mathbf{I},\mathbf{I})$	$\boldsymbol{K}\boldsymbol{K}^{\mathrm{t}}$	$\tilde{\boldsymbol{K}}\tilde{\boldsymbol{K}}^{\mathrm{t}}$
(e)	$\mathcal{H}_{n,m}(\boldsymbol{r};\boldsymbol{K},\mathbf{0},2\boldsymbol{K}^*)$	$\mathcal{H}_{l,k}(\boldsymbol{r};\mathbf{I},2\boldsymbol{K}^*,\mathbf{0})$	$\boldsymbol{K}$	$\boldsymbol{K}^{*-1}$
(f)	$\mathcal{H}_{n,m}(\boldsymbol{r};\mathbf{I},\mathbf{0},2\mathbf{I})$	$\mathcal{H}_{l,k}(\boldsymbol{r};\mathbf{I},2\mathbf{I},\mathbf{0})$	$\mathbf{I}$	$\mathbf{I}$

$\mathcal{H}_{l,k}(\boldsymbol{r};\boldsymbol{K}_2,\boldsymbol{L}_2,\boldsymbol{L}_1)$ then form a bi-orthonormal set when equation (16) is satisfied. The matrices $\boldsymbol{M}$ follow directly from equation (8) (or (14) and (15)); see also table 1 (row (c)). We easily verify that for any non-orthonormal set of Gaussian-type modes $\mathcal{H}_{n,m}(\boldsymbol{r};\boldsymbol{K},\boldsymbol{L},\tilde{\boldsymbol{L}})$ with arbitrarily given matrices $\boldsymbol{K}$ (non-singular), $\boldsymbol{L}$ (symmetric) and $\tilde{\boldsymbol{L}}$ (also symmetric, and satisfying the condition that the real part of $\boldsymbol{L}+\tilde{\boldsymbol{L}}$ be positive definite), we can always find its associated bi-orthonormal partner set $\mathcal{H}_{l,k}(\boldsymbol{r};\tilde{\boldsymbol{K}},\tilde{\boldsymbol{L}},\boldsymbol{L})$. As an easy example, we mention that the set of common Hermite polynomials $\mathcal{H}_{n,m}(\boldsymbol{r};\mathbf{I},\mathbf{0},2\mathbf{I})$, which arises for $\boldsymbol{K}=\mathbf{I}$, $\boldsymbol{L}=\mathbf{0}$, and $\boldsymbol{M}=\mathbf{I}$ (and thus $\tilde{\boldsymbol{L}}=2\mathbf{I}$), has as its bi-orthonormal partner set the modes $\mathcal{H}_{l,k}(\boldsymbol{r};\mathbf{I},2\mathbf{I},\mathbf{0})$ (and thus $\tilde{\boldsymbol{M}}=\mathbf{I}$); see also table 1 (row (f)). The strictly orthonormal case that we considered before in [5] and [6], arises when $\boldsymbol{M}=\boldsymbol{K}\boldsymbol{K}^{*-1}$, see equation (5), and $\operatorname{Re}\boldsymbol{L}=\boldsymbol{K}^{\mathrm{t}}\boldsymbol{K}^*$, see equation (6) (and thus $\tilde{\boldsymbol{L}}=\boldsymbol{L}$); see also table 1 (row (a)) and, for the common Hermite–Gaussian modes in particular, table 1 (row (b)).

Our results are in agreement with those of Wünsche [3] for the special choice $\boldsymbol{L}=\tilde{\boldsymbol{L}}=\mathbf{I}$. From the bi-orthonormality condition (16) we easily get $\tilde{\boldsymbol{K}}^*=\boldsymbol{K}^{\mathrm{t}-1}$, and we also have $\boldsymbol{M}=\boldsymbol{K}\boldsymbol{K}^{\mathrm{t}}$ and $\tilde{\boldsymbol{M}}^*=(\boldsymbol{K}\boldsymbol{K}^{\mathrm{t}})^{-1}$; see also table 1 (row (d)). With these choices of $\boldsymbol{K}$, $\boldsymbol{L}$ and $\tilde{\boldsymbol{L}}$, Wünsche's relation [3, equation (5.9)] between the two bi-orthonormal sets of Hermite two-dimensional functions $\mathcal{H}_{n,m}(\boldsymbol{r};\boldsymbol{K},\mathbf{I},\mathbf{I})$ and $\mathcal{H}_{l,k}(\boldsymbol{r};\tilde{\boldsymbol{K}},\mathbf{I},\mathbf{I})$, with $\boldsymbol{K}^{\mathrm{t}}\tilde{\boldsymbol{K}}^*=\mathbf{I}$, and our bi-orthonormality relation (7) are identical. The results are also in agreement with the set of two-variable Hermite polynomials [8, section 12.8] $\mathcal{H}_{n,m}(\boldsymbol{r};\boldsymbol{K},\mathbf{0},2\boldsymbol{K}^*)$ with $\boldsymbol{K}=\boldsymbol{K}^{\mathrm{t}}$ and its bi-orthonormal partner set $\mathcal{H}_{l,k}(\boldsymbol{r};\mathbf{I},2\boldsymbol{K}^*,\mathbf{0})$, for which the matrices $\boldsymbol{M}$ and $\tilde{\boldsymbol{M}}$ read $\boldsymbol{K}$ and $\boldsymbol{K}^{*-1}$, respectively; see also table 1 (row (e)).

4. A closed-form expression for Gaussian-type modes

From the generating function (1) we derive the derivative relations for $\mathcal{H}_{n,m}(\boldsymbol{r}) = \mathcal{H}_{n,m}(\boldsymbol{r};\boldsymbol{K},\boldsymbol{L},\tilde{\boldsymbol{L}})$

$$\begin{pmatrix} \frac{\partial\mathcal{H}_{n,m}(\boldsymbol{r})}{\partial x} \\ \frac{\partial\mathcal{H}_{n,m}(\boldsymbol{r})}{\partial y} \end{pmatrix} = -2\pi\mathcal{H}_{n,m}(\boldsymbol{r})\boldsymbol{L}\begin{pmatrix} x \\ y \end{pmatrix} + 2\sqrt{\pi}\boldsymbol{K}^{\mathrm{t}}\begin{pmatrix} \sqrt{n}\mathcal{H}_{n-1,m}(\boldsymbol{r}) \\ \sqrt{m}\mathcal{H}_{n,m-1}(\boldsymbol{r}) \end{pmatrix} \tag{20}$$

by differentiating it with respect to $\boldsymbol{r}$, and the recurrence relations

$$2\sqrt{\pi}\mathcal{H}_{n,m}(\boldsymbol{r})\boldsymbol{K}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \sqrt{n+1}\mathcal{H}_{n+1,m}(\boldsymbol{r}) \\ \sqrt{m+1}\mathcal{H}_{n,m+1}(\boldsymbol{r}) \end{pmatrix} + \boldsymbol{M}\begin{pmatrix} \sqrt{n}\mathcal{H}_{n-1,m}(\boldsymbol{r}) \\ \sqrt{m}\mathcal{H}_{n,m-1}(\boldsymbol{r}) \end{pmatrix} \tag{21}$$

by differentiating it with respect to s. The derivative and recurrence relations can be combined to yield

$$\begin{pmatrix} \sqrt{n+1}\mathcal{H}_{n+1,m}(\boldsymbol{r}) \\ \sqrt{m+1}\mathcal{H}_{n,m+1}(\boldsymbol{r}) \end{pmatrix} = \left[\sqrt{\pi}\tilde{\boldsymbol{K}}^{*^{\mathrm{t}-1}}\tilde{\boldsymbol{L}}^{*}\begin{pmatrix} x \\ y \end{pmatrix} - \frac{1}{2\sqrt{\pi}}\tilde{\boldsymbol{K}}^{*^{\mathrm{t}-1}}\begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \end{pmatrix}\right]\mathcal{H}_{n,m}(\boldsymbol{r}), \tag{22}$$

where we have substituted $\boldsymbol{M}\boldsymbol{K}^{\mathrm{t}^{-1}} = \tilde{\boldsymbol{K}}^{*^{\mathrm{t}-1}}$ and $2\boldsymbol{K} - \boldsymbol{M}\boldsymbol{K}^{\mathrm{t}^{-1}}\boldsymbol{L} = \tilde{\boldsymbol{K}}^{*^{\mathrm{t}-1}}\tilde{\boldsymbol{L}}^{*}$, which follow directly from the bi-orthonormality conditions (14)–(16). Equation (22) can be written in an operator notation as [9]

$$\begin{aligned} 2\sqrt{\pi(n+1)}\mathcal{H}_{n+1,m}(\boldsymbol{r}) &= \mathcal{P}_x\mathcal{H}_{n,m}(\boldsymbol{r}) \\ 2\sqrt{\pi(m+1)}\mathcal{H}_{n,m+1}(\boldsymbol{r}) &= \mathcal{P}_y\mathcal{H}_{n,m}(\boldsymbol{r}) \end{aligned} \tag{23}$$

with the operators

$$\begin{aligned} \mathcal{P}_x &= 2\pi(U_{11}x + U_{12}y) - Z_{11}\frac{\partial}{\partial x} - Z_{12}\frac{\partial}{\partial y} \\ \mathcal{P}_y &= 2\pi(U_{21}x + U_{22}y) - Z_{21}\frac{\partial}{\partial x} - Z_{22}\frac{\partial}{\partial y} \end{aligned} \tag{24}$$

and the matrices

$$\boldsymbol{U} = \tilde{\boldsymbol{K}}^{*^{\mathrm{t}-1}}\tilde{\boldsymbol{L}}^{*} \qquad \text{and} \qquad \boldsymbol{Z} = \tilde{\boldsymbol{K}}^{*^{\mathrm{t}-1}}. \tag{25}$$

Note that the operators $\mathcal{P}_x$ and $\mathcal{P}_y$ commute, since $\boldsymbol{Z}\boldsymbol{U}^{\mathrm{t}} = \boldsymbol{U}\boldsymbol{Z}^{\mathrm{t}}$, and that we are thus led to an alternative, closed-form expression for Gaussian-type modes:

$$\mathcal{H}_{n,m}(\boldsymbol{r};\boldsymbol{K},\boldsymbol{L},\tilde{\boldsymbol{L}}) = \frac{\mathcal{P}_x^n\mathcal{P}_y^m\mathcal{H}_{0,0}(\boldsymbol{r};\boldsymbol{K},\boldsymbol{L})}{2^{n+m}\sqrt{\pi^{n+m}n!m!}}. \tag{26}$$

We remark that the set of modes is defined by using $\mathcal{H}_{0,0}(\boldsymbol{r};\boldsymbol{K},\boldsymbol{L}) = 2^{1/2}(\det\boldsymbol{K})^{1/2}\exp(-\pi\boldsymbol{r}^{\mathrm{t}}\boldsymbol{L}\boldsymbol{r})$ as its root, whereas the operators $\mathcal{P}_x$ and $\mathcal{P}_y$ depend on the parameters $\tilde{\boldsymbol{K}}$ and $\tilde{\boldsymbol{L}}$ of the bi-orthonormal partner set. The case of strictly orthonormal modes, with $\tilde{\boldsymbol{K}} = \boldsymbol{K} = (\boldsymbol{a}+\mathrm{i}\boldsymbol{b})^{-1}$ and $\tilde{\boldsymbol{L}} = \boldsymbol{L} = (\boldsymbol{d}-\mathrm{i}\boldsymbol{c})(\boldsymbol{a}+\mathrm{i}\boldsymbol{b})^{-1}$, was reported in [9].

5. Evolution under linear canonical transformations

We now let non-orthonormal Gaussian-type modes $\mathcal{H}_{n,m}(\boldsymbol{r};\boldsymbol{K},\boldsymbol{L},\tilde{\boldsymbol{L}})$ propagate through a linear system $f_{\mathrm{o}}(\boldsymbol{r}) = \mathcal{L}\{f_{\mathrm{i}}(\boldsymbol{r})\}$ whose input–output relationship is described by a linear canonical transformation [10], and determine the generating function of the set of modes that appear at the output of this system. In optical terms we are thus considering a lossless first-order optical system—also called an $\boldsymbol{ABCD}$ system—described by its ray transformation matrix $\boldsymbol{T}$ [11], which relates the position $\boldsymbol{r}_{\mathrm{i}}$ and direction $\boldsymbol{q}_{\mathrm{i}}$ of an incoming ray to the position $\boldsymbol{r}_{\mathrm{o}}$ and direction $\boldsymbol{q}_{\mathrm{o}}$ of the outgoing ray:

$$\begin{pmatrix} \boldsymbol{r}_{\mathrm{o}} \\ \boldsymbol{q}_{\mathrm{o}} \end{pmatrix} = \begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{pmatrix}\begin{pmatrix} \boldsymbol{r}_{\mathrm{i}} \\ \boldsymbol{q}_{\mathrm{i}} \end{pmatrix} \equiv \boldsymbol{T}\begin{pmatrix} \boldsymbol{r}_{\mathrm{i}} \\ \boldsymbol{q}_{\mathrm{i}} \end{pmatrix}. \tag{27}$$

The ray transformation matrix of such a system is real and symplectic, yielding the relations

$$\begin{aligned} &\boldsymbol{A}\boldsymbol{B}^{\mathrm{t}} = \boldsymbol{B}\boldsymbol{A}^{\mathrm{t}}, \quad &&\boldsymbol{C}\boldsymbol{D}^{\mathrm{t}} = \boldsymbol{D}\boldsymbol{C}^{\mathrm{t}}, \quad &&\boldsymbol{A}\boldsymbol{D}^{\mathrm{t}} - \boldsymbol{B}\boldsymbol{C}^{\mathrm{t}} = \mathbf{I}, \\ &\boldsymbol{A}^{\mathrm{t}}\boldsymbol{C} = \boldsymbol{C}^{\mathrm{t}}\boldsymbol{A}, \quad &&\boldsymbol{B}^{\mathrm{t}}\boldsymbol{D} = \boldsymbol{D}^{\mathrm{t}}\boldsymbol{B}, \quad &&\boldsymbol{A}^{\mathrm{t}}\boldsymbol{D} - \boldsymbol{C}^{\mathrm{t}}\boldsymbol{B} = \mathbf{I}, \end{aligned} \tag{28}$$

or in a short-hand matrix notation,

$$\boldsymbol{T}^{-1} = \mathbf{J}\boldsymbol{T}^{\mathrm{t}}\mathbf{J} \qquad \text{with} \quad \mathbf{J} = \mathbf{J}^{-1} = \mathbf{J}^{*\mathrm{t}} = \mathrm{i}\begin{pmatrix} \mathbf{0} & -\mathbf{I} \\ \mathbf{I} & \mathbf{0} \end{pmatrix}. \tag{29}$$

Using the matrices $\boldsymbol{A}, \boldsymbol{B}$ and $\boldsymbol{D}$, and assuming that $\boldsymbol{B}$ is a non-singular matrix, we can represent the canonical transformation by the Collins integral [12]

$$f_{\mathrm{o}}(\boldsymbol{r}_{\mathrm{o}}) = \frac{1}{\sqrt{\det \mathrm{i}\boldsymbol{B}}} \int\int_{-\infty}^{\infty} f_{\mathrm{i}}(\boldsymbol{r}_{\mathrm{i}}) \exp\left[\mathrm{i}\pi\left(\boldsymbol{r}_{\mathrm{i}}^{\mathrm{t}}\boldsymbol{B}^{-1}\boldsymbol{A}\boldsymbol{r}_{\mathrm{i}} - 2\boldsymbol{r}_{\mathrm{i}}^{\mathrm{t}}\boldsymbol{B}^{-1}\boldsymbol{r}_{\mathrm{o}} + \boldsymbol{r}_{\mathrm{o}}^{\mathrm{t}}\boldsymbol{D}\boldsymbol{B}^{-1}\boldsymbol{r}_{\mathrm{o}}\right)\right] \mathrm{d}\boldsymbol{r}_{\mathrm{i}}, \quad (30)$$

where the output amplitude $f_{\mathrm{o}}(\boldsymbol{r})$ is expressed in terms of the input amplitude $f_{\mathrm{i}}(\boldsymbol{r})$. We remark that a linear canonical transformation $f_{\mathrm{o}}(\boldsymbol{r}) = \mathcal{L}\{f_{\mathrm{i}}(\boldsymbol{r})\}$ is unitary, i.e.,

$$\int\int_{-\infty}^{\infty} f_{\mathrm{i},1}(\boldsymbol{r}) f_{\mathrm{i},2}^{*}(\boldsymbol{r})\, \mathrm{d}\boldsymbol{r} = \int\int_{-\infty}^{\infty} f_{\mathrm{o},1}(\boldsymbol{r}) f_{\mathrm{o},2}^{*}(\boldsymbol{r})\, \mathrm{d}\boldsymbol{r}. \quad (31)$$

This implies that bi-orthonormality properties for input signals also hold for the corresponding output signals. Moreover, we easily see that the generating function undergoes the same transformation $\mathcal{L}$ as the signal $f(\boldsymbol{r})$ does.

With a Gaussian-type mode $\mathcal{H}_{n,m}(\boldsymbol{r}; \boldsymbol{K}_{\mathrm{i}}, \boldsymbol{L}_{\mathrm{i}}, \tilde{\boldsymbol{L}}_{\mathrm{i}})$ at the input of an $\boldsymbol{ABCD}$ system, we denote the output mode by $\mathcal{H}_{n,m}(\boldsymbol{r}; \boldsymbol{K}_{\mathrm{o}}, \boldsymbol{L}_{\mathrm{o}}, \tilde{\boldsymbol{L}}_{\mathrm{o}})$. Reasoning along the lines presented in [6], we get the input–output relationships

$$\boldsymbol{K}_{\mathrm{o}} = \boldsymbol{K}_{\mathrm{i}}(\boldsymbol{A} + \boldsymbol{B}\mathrm{i}\boldsymbol{L}_{\mathrm{i}})^{-1}, \quad (32)$$

$$\mathrm{i}\boldsymbol{L}_{\mathrm{o}} = (\boldsymbol{C} + \boldsymbol{D}\mathrm{i}\boldsymbol{L}_{\mathrm{i}})(\boldsymbol{A} + \boldsymbol{B}\mathrm{i}\boldsymbol{L}_{\mathrm{i}})^{-1}, \quad (33)$$

$$\boldsymbol{M}_{\mathrm{o}} = \boldsymbol{M}_{\mathrm{i}} - 2\mathrm{i}\boldsymbol{K}_{\mathrm{i}}\,(\boldsymbol{A} + \boldsymbol{B}\mathrm{i}\boldsymbol{L}_{\mathrm{i}})^{-1}\,\boldsymbol{B}\boldsymbol{K}_{\mathrm{i}}^{\mathrm{t}}. \quad (34)$$

Note that equation (33) is in fact the well-known bilinear $\boldsymbol{ABCD}$ law, and that equations (32) and (33) can be combined into

$$\begin{pmatrix} \mathbf{I} \\ \mathrm{i}\boldsymbol{L}_{\mathrm{o}} \end{pmatrix} \boldsymbol{K}_{\mathrm{o}}^{-1} = \begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{pmatrix} \begin{pmatrix} \mathbf{I} \\ \mathrm{i}\boldsymbol{L}_{\mathrm{i}} \end{pmatrix} \boldsymbol{K}_{\mathrm{i}}^{-1}. \quad (35)$$

It is not difficult to see that for $\boldsymbol{B} = \mathbf{0}$, in which case Collins integral (30) reduces to $f_{\mathrm{o}}(\boldsymbol{r}) = f_{\mathrm{i}}(\boldsymbol{A}^{-1}\boldsymbol{r}) \exp(\mathrm{i}\pi \boldsymbol{r}^{\mathrm{t}}\boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{r})/\sqrt{|\det \boldsymbol{A}|}$, these input–output relations remain valid. They also remain valid if $\boldsymbol{B}$ is a singular matrix while at the same time $\boldsymbol{B} \neq \mathbf{0}$, see, for instance, [10], where it was shown that any system with a singular matrix $\boldsymbol{B}$ can be represented as a cascade of two subsystems whose matrices $\boldsymbol{B}$ are non-singular; in that case we may simply use Collins integral (30) twice, for each subsystem separately. We thus conclude that while the generating function (1) keeps its form when Gaussian-type modes undergo a linear canonical transformation, we only have to replace the input matrices $\boldsymbol{K}_{\mathrm{i}}, \boldsymbol{L}_{\mathrm{i}}$ and $\boldsymbol{M}_{\mathrm{i}}$ by the output matrices $\boldsymbol{K}_{\mathrm{o}}, \boldsymbol{L}_{\mathrm{o}}$ and $\boldsymbol{M}_{\mathrm{o}}$, respectively, in accordance with the input–output relationships (32)–(34).

Equation (34), with which the propagation of the matrix $\boldsymbol{M}$ is described, can be neglected if we work simultaneously with the non-orthonormal set of modes and its associated bi-orthonormal partner set, and when we recall that bi-orthonormality properties are invariant under a unitary transformation, see equation (31). For both sets of modes, the matrices $\boldsymbol{K}$ and $\boldsymbol{L}$ in the output plane follow from equation (35), and the matrices $\boldsymbol{M}$ in the output plane then follow easily from equation (8) (or (14) and (15)). We can act even more efficiently if we are not interested in all the parameters of the associated bi-orthonormal set: we simply use equation (35)—or equations (32) and (33)—for the propagation of $\boldsymbol{K}$ and $\boldsymbol{L}$, combined with the bilinear relation

$$\mathrm{i}\tilde{\boldsymbol{L}}_{\mathrm{o}} = (\boldsymbol{C} + \boldsymbol{D}\mathrm{i}\tilde{\boldsymbol{L}}_{\mathrm{i}})(\boldsymbol{A} + \boldsymbol{B}\mathrm{i}\tilde{\boldsymbol{L}}_{\mathrm{i}})^{-1}, \quad (36)$$

cf equation (33), to find the propagation of the associated matrix $\tilde{\boldsymbol{L}}$.

If we express $\boldsymbol{K}_{\mathrm{i,o}}^{-1}$ in their real and imaginary parts, $\boldsymbol{K}_{\mathrm{i,o}}^{-1} = \boldsymbol{a}_{\mathrm{i,o}} + \mathrm{i}\boldsymbol{b}_{\mathrm{i,o}}$, and subsequently express $\mathrm{i}\boldsymbol{L}_{i,\mathrm{o}}$ as $\mathrm{i}\boldsymbol{L}_{\mathrm{i,o}} = (\boldsymbol{c}_{\mathrm{i,o}} + \mathrm{i}\boldsymbol{d}_{\mathrm{i,o}})\boldsymbol{K}_{\mathrm{i,o}} = (\boldsymbol{c}_{\mathrm{i,o}} + \mathrm{i}\boldsymbol{d}_{\mathrm{i,o}})(\boldsymbol{a}_{\mathrm{i,o}} + \mathrm{i}\boldsymbol{b}_{\mathrm{i,o}})^{-1}$, equation (35) can as well be expressed in the form [6]

$$\begin{pmatrix} \boldsymbol{a}_\mathrm{o} & \boldsymbol{b}_\mathrm{o} \\ \boldsymbol{c}_\mathrm{o} & \boldsymbol{d}_\mathrm{o} \end{pmatrix} = \begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{pmatrix} \begin{pmatrix} \boldsymbol{a}_\mathrm{i} & \boldsymbol{b}_\mathrm{i} \\ \boldsymbol{c}_\mathrm{i} & \boldsymbol{d}_\mathrm{i} \end{pmatrix}. \tag{37}$$

This propagation law resembles equations (12) and (29) in [13], where $\mathrm{i}(\boldsymbol{a} + \mathrm{i}\boldsymbol{b})$, $-\mathrm{i}(\boldsymbol{c} + \mathrm{i}\boldsymbol{d})$ correspond to the 'matricial rays' $\boldsymbol{Q}$, $\boldsymbol{P}$ [13, equation (11)], $\boldsymbol{Q}\sqrt{\pi} = \mathrm{i}\boldsymbol{K}^{-1} = \mathrm{i}(\boldsymbol{a} + \mathrm{i}\boldsymbol{b})$ and $\boldsymbol{P}\sqrt{\pi} = \lambda\boldsymbol{L}\boldsymbol{K}^{-1} = -\mathrm{i}\lambda(\boldsymbol{c} + \mathrm{i}\boldsymbol{d})$, with λ the wavelength of the light. From [6] we learn that in the case of strictly orthonormal modes, the 4×4 $\boldsymbol{abcd}$ matrices that arise in equation (37) are symplectic, and the four sub-matrices $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ and $\boldsymbol{d}$ thus satisfy relations of the form (28). Although such a nice property does not hold for non-orthonormal modes, we do have similar relations for bi-orthonormal partner sets,

$$(\tilde{\boldsymbol{a}}^\mathrm{t}\boldsymbol{c} - \tilde{\boldsymbol{c}}^\mathrm{t}\boldsymbol{a}) + (\tilde{\boldsymbol{b}}^\mathrm{t}\boldsymbol{d} - \tilde{\boldsymbol{d}}^\mathrm{t}\boldsymbol{b}) = \boldsymbol{0} \qquad \text{and} \qquad (\tilde{\boldsymbol{a}}^\mathrm{t}\boldsymbol{d} - \tilde{\boldsymbol{c}}^\mathrm{t}\boldsymbol{b}) + (\tilde{\boldsymbol{d}}^\mathrm{t}\boldsymbol{a} - \tilde{\boldsymbol{b}}^\mathrm{t}\boldsymbol{c}) = 2\mathbf{I}, \tag{38}$$

immediately resulting from the bi-orthonormality equation (16). In the special case that $\boldsymbol{L} = \tilde{\boldsymbol{L}}$ we even have

$$\tilde{\boldsymbol{a}}^\mathrm{t}\boldsymbol{c} = \tilde{\boldsymbol{c}}^\mathrm{t}\boldsymbol{a}, \qquad \tilde{\boldsymbol{b}}^\mathrm{t}\boldsymbol{d} = \tilde{\boldsymbol{d}}^\mathrm{t}\boldsymbol{b}, \qquad \tilde{\boldsymbol{a}}^\mathrm{t}\boldsymbol{d} - \tilde{\boldsymbol{c}}^\mathrm{t}\boldsymbol{b} = \tilde{\boldsymbol{d}}^\mathrm{t}\boldsymbol{a} - \tilde{\boldsymbol{b}}^\mathrm{t}\boldsymbol{c} = \mathbf{I}, \tag{39}$$

or, in short-hand matrix notation,

$$\tilde{\boldsymbol{m}}^\mathrm{t}\mathbf{J}\boldsymbol{m} = \mathbf{J} \qquad \text{with} \quad \boldsymbol{m} = \begin{pmatrix} \boldsymbol{a} & \boldsymbol{b} \\ \boldsymbol{c} & \boldsymbol{d} \end{pmatrix} \quad \text{and} \quad \tilde{\boldsymbol{m}} = \begin{pmatrix} \tilde{\boldsymbol{a}} & \tilde{\boldsymbol{b}} \\ \tilde{\boldsymbol{c}} & \tilde{\boldsymbol{d}} \end{pmatrix}, \tag{40}$$

cf equations (28) and (29), and the two bi-orthonormal partner sets may then be called 'bi-symplectic.' We easily verify that bi-symplecticity is preserved under linear canonical transformations: with $\boldsymbol{m}_\mathrm{i}$ and $\tilde{\boldsymbol{m}}_\mathrm{i}$ a pair of bi-symplectic matrices, $\tilde{\boldsymbol{m}}_\mathrm{i}^\mathrm{t}\mathbf{J}\boldsymbol{m}_\mathrm{i} = \mathbf{J}$, and with $\boldsymbol{T}$ a symplectic matrix, $\boldsymbol{T}^\mathrm{t}\mathbf{J}\boldsymbol{T} = \mathbf{J}$, describing the input–output relationship $\boldsymbol{m}_\mathrm{o} = \boldsymbol{T}\boldsymbol{m}_\mathrm{i}$ between the two matrices $\boldsymbol{m}_\mathrm{i}$ and $\boldsymbol{m}_\mathrm{o}$, we immediately derive $\tilde{\boldsymbol{m}}_\mathrm{o}^\mathrm{t}\mathbf{J}\boldsymbol{m}_\mathrm{o} = \tilde{\boldsymbol{m}}_\mathrm{i}^\mathrm{t}\boldsymbol{T}^\mathrm{t}\mathbf{J}\boldsymbol{T}\boldsymbol{m}_\mathrm{i} = \tilde{\boldsymbol{m}}_\mathrm{i}^\mathrm{t}\mathbf{J}\boldsymbol{m}_\mathrm{i} = \mathbf{J}$.

6. Conclusion

We have introduced a general class of sets of non-orthonormal Gaussian-type modes, along with their associated bi-orthonormal sets. The conditions between these two bi-orthonormal sets of modes have been derived, based on their generating functions, and the relation with other functions (the Hermite two-dimensional functions and the two-variable Hermite polynomials) has been established. In addition to their rather implicit definition in terms of a generating function, we have also derived a closed-form expression for these Gaussian-type modes. We have shown that the evolution of non-orthonormal Gaussian-type modes under a linear canonical transformation, described by a symplectic matrix, can be described by the same mechanism as is used for the evolution of orthonormal Hermite–Gaussian-type modes, when, simultaneously, the associated bio-orthonormal partners are taken into account. Finally, a subclass has been observed for which the two bi-orthogonal partners show a property that might be called bi-symplecticity.

It will be clear that the procedure of defining non-orthonormal Gaussian-type modes, based on a generalization of the generating function of the common Hermite–Gaussian modes as presented here, is not restricted to the two-dimensional case, but can easily be extended to more dimensions.

Acknowledgments

The Spanish Ministry of Education and Science is acknowledged for financial support: 'Ramon y Cajal' grant and project TIC 2002-01846 (TA) and SAB2004-0018 (MJB).

Appendix. Non-singularity of a symmetric matrix with a positive-definite real part

We consider a symmetric matrix $\boldsymbol{P} = \boldsymbol{P}^{\mathrm{t}} = \boldsymbol{X} + \mathrm{i}\boldsymbol{Y}$ whose real part $\boldsymbol{X}$ is positive definite: $\boldsymbol{r}^{\mathrm{t}}\boldsymbol{X}\boldsymbol{r} > 0$ for any real vector $\boldsymbol{r}$. In order for the matrix $\boldsymbol{P}$ to be singular, there should be at least one vanishing eigenvalue $\lambda_{\mathrm{o}} = 0$. Let us denote the corresponding eigenvector by $\boldsymbol{r}_{\mathrm{o}} = \boldsymbol{x}_{\mathrm{o}} + \mathrm{i}\boldsymbol{y}_{\mathrm{o}}$, and we have the relation $\boldsymbol{P}\boldsymbol{r}_{\mathrm{o}} = \lambda_{\mathrm{o}}\boldsymbol{r}_{\mathrm{o}} = \boldsymbol{0}$, leading to the following two equations for the real and imaginary parts separately: $\boldsymbol{X}\boldsymbol{x}_{\mathrm{o}} = \boldsymbol{Y}\boldsymbol{y}_{\mathrm{o}}$ and $\boldsymbol{X}\boldsymbol{y}_{\mathrm{o}} = -\boldsymbol{Y}\boldsymbol{x}_{\mathrm{o}}$. We now consider the quadratic form $\boldsymbol{x}_{\mathrm{o}}^{\mathrm{t}}\boldsymbol{X}\boldsymbol{x}_{\mathrm{o}}$ and substitute from these two equations: $\boldsymbol{x}_{\mathrm{o}}^{\mathrm{t}}\boldsymbol{X}\boldsymbol{x}_{\mathrm{o}} = \boldsymbol{x}_{\mathrm{o}}^{\mathrm{t}}\boldsymbol{Y}\boldsymbol{y}_{\mathrm{o}} = \boldsymbol{y}_{\mathrm{o}}^{\mathrm{t}}\boldsymbol{Y}\boldsymbol{x}_{\mathrm{o}} = -\boldsymbol{y}_{\mathrm{o}}^{\mathrm{t}}\boldsymbol{X}\boldsymbol{y}_{\mathrm{o}}$. In view of the positive definiteness of the matrix $\boldsymbol{X}$, the relation $\boldsymbol{x}_{\mathrm{o}}^{\mathrm{t}}\boldsymbol{X}\boldsymbol{x}_{\mathrm{o}} = -\boldsymbol{y}_{\mathrm{o}}^{\mathrm{t}}\boldsymbol{X}\boldsymbol{y}_{\mathrm{o}}$ cannot be true, and we conclude that a vanishing eigenvalue does not exist. In other words, a symmetric matrix whose real part is positive definite, is non-singular.

References

[1] Beijersbergen M W, Allen L, van der Veen H E L O and Woerdman J P 1993 Astigmatic laser mode converters and transfer of orbital angular momentum *Opt. Commun.* **96** 123
[2] Simon S and Agarwal G S 2000 Wigner representation of Laguerre–Gaussian beams *Opt. Lett.* **25** 1313
[3] Wünsche A 2000 General Hermite and Laguerre two-dimensional polynomials *J. Phys. A: Math Gen.* **33** 1603
Wünsche A 2000 *J. Phys. A: Math Gen.* **33** 3531 (erratum)
[4] Abramochkin E G and Volostnikov V G 2004 Generalized Gaussian beams *J. Opt. A: Pure Appl. Opt.* **6** S157
[5] Bastiaans M J and Alieva T 2005 Generating function for Hermite–Gaussian modes propagating through first-order optical systems *J. Phys. A: Math. Gen.* **38** L73
[6] Bastiaans M J and Alieva T 2005 Propagation law for the generating function of Hermite–Gaussian-type modes in first-order optical systems *Opt. Express* **13** 1107
[7] Prudnikov A P, Brychkov Yu A and Marichev O I (ed) 1986 *Integrals and Series, Volume 1, Elementary Functions* (New York: Gordon and Breach)
[8] Erdélyi A (ed) 1953 *Higher Transcendental Functions* vol 2 (New York: McGraw-Hill)
[9] Alieva T and Bastiaans M J 2005 Mode mapping in paraxial lossless optics *Opt. Lett.* **30** 1461
[10] Moshinsky M and Quesne C 1971 Linear canonical transformations and their unitary representations *J. Math. Phys.* **12** 1772
[11] Luneburg R K 1966 *Mathematical Theory of Optics* (Berkeley, CA: University of California Press)
[12] Collins S A 1970 Lens-system diffraction integral written in terms of matrix optics *J. Opt. Soc. Am.* **60** 1168
[13] Arnaud J A 1971 Mode coupling in first-order optics *J. Opt. Soc. Am.* **61** 751