From Kurt Bernardo Wolf, Optical Models and Symmetries. In: Taco D. Visser, editor, *Progress in Optics, Vol. 62*, Amsterdam: Elsevier, 2017, pp. 225-291.
ISBN: 978-0-12-811999-0

Elsevier

CHAPTER FOUR

Optical Models and Symmetries

Kurt Bernardo Wolf
Instituto de Ciencias Físicas, Universidad Nacional Autónoma de México, Cuernavaca, Morelos, Mexico

Contents

Progress in Optics, Volume 62
ISSN 0079-6638
http://dx.doi.org/10.1016/bs.po.2016.12.002

1. INTRODUCTION

It is not difficult to argue that optics offers a richer field for the recognition and use of symmetries than mechanics—classical or quantum. The harvest includes the geometric, wave, and finite (pixelated) models of optics; in turn, the first two encompass the global (4π), paraxial and metaxial (aberration) regimes, while the finite model, when the number and density of pixels increases without bound, limits to the continuous cases.

The *mother symmetry* of all these models is the Euclidean group and algebra of translations and rotations. This statement may appear disappointing at first sight because that is the symmetry of a homogeneous and isotropic vacuum. But quite on the contrary, as we shall show, this symmetry serves as a basis for the construction of phase space in geometric optics, as well as the Hilbert space formulation of wave optics. A *deformation* of this group to that of rotations produces the finite model of pixelated optical systems. Using various techniques of deformation and contraction we succinctly shift between them and their distinct regimes.

The theory of Lie algebras and groups was developed by mathematicians during the second half of the 19th century and applied to find and classify all possible crystallographic lattices in three dimensions—which are but discrete subgroups of the Euclidean group. Early researchers in quantum mechanics found that rotational bands in the spectra of atomic systems were naturally characterized by the underlying geometric symmetry, while systematic level degeneracies were due to hidden, higher symmetries. The second half of the 20th century became the heyday of Lie group theory as nuclear and elementary particle physics presented quantum-number patterns and conservation laws whose origin was understood to be due to symmetries of Hamiltonians that were themselves unknown; yet, they provided conservation laws and sum rules for the observed reaction rates.

The main reason for the statement in our first sentence is that mechanical Hamiltonians basically read $H = p^2/2m + V(q)$, separated into a fixed kinetic term of squared momentum p, and an in-principle arbitrary potential term, normally of position q alone. In optics on the other hand, we come to write and use evolution-generating Hamiltonians of various other forms, according to whether the regime is global, paraxial or metaxial, geometric, wave or finite; and a phase space with symplectic metric follows. Most generators are of *co*-variance, rather than *in*-variance transformations, and belong to finite-dimensional Lie algebras whose representation theory is well established. We shall basically work in $D = 3$ space dimensions, occasionally with the $D = 2$ case for clarity in some figures, but most developments are in principle valid for generic D dimensions.

Since the main preoccupation of early optical research was the faithful formation of images on a plane screen, the angles of incoming rays—their *momenta*—did not really matter, so they were not placed on the same level of interest as ray *positions*. The recent surge of literature on linear canonical transforms for the paraxial regime of geometric and wave optics has highlighted the necessity of treating both position and momentum as coordinates of a phase space where one-dimensional wavefields could be visually displayed on a plane through their Wigner function as a music sheet with time and frequency axes. As we shall show, the phase space and Wigner function constructs fit also some of the other optical models and regimes, albeit with different topologies, but based on purely group-theoretical premises and applicable to other Lie groups beside the Euclidean.

First of all, in Section 2 we introduce the Euclidean group and its structure. The symmetries of the basic objects of the geometric and wave models of light—a line and a plane in space—are presented in Section 3. The reader will appreciate that these are two among a number of other possibilities based on the symmetry of chosen fundamental "objects" within the same Euclidean mother group. The geometric model of light with a dynamical postulate builds a Hamiltonian system whose canonical and optical transformations are the subject of Section 4, while Section 5 builds a Hilbert space analogue for Helmholtz monochromatic wavefields.

The contraction of the Euclidean group along the evolution axis yields in Section 6 the paraxial models of geometric and wave optical models under the Heisenberg–Weyl group of translations in position and paraxial momentum. Linear transformations, obtained through a quadratic extension of this algebra in Section 7, lead to the symplectic group of canonical geometric and wave transformations, the former serving to introduce the unitary Fourier subgroup, and the latter realized by integral transforms.

Higher-order extensions, addressed in Section 8, take us to the metaxial regime where the classification and action of aberrations are set forth.

The Euclidean group of continuous optical models is actually the result of a contraction of a higher compact group: rotations in 4-space. There, the operators of position and momentum have finite spectra of equally spaced eigenvalues; in Section 9 we thus present discrete optical models on linear or rectangular pixelated screens and the unitary transformations under which no information is lost. Finally, when the number and density of pixels grows without bound, one recovers the continuous models based on the Euclidean group and canonical transformations. In Appendix A we review a Wigner function defined on the rotation group that allows us to plot discrete and finite signals on phase space, in particular of aberrated signals; this too contracts to the generally known Wigner function on paraxial phase space.

2. THE EUCLIDEAN GROUP

Consider a 3D (three-dimensional) space whose points are labeled by column vectors[1] $\vec{r} = (x, y, z)^\top \in \mathsf{R}^3$. The rigid transformations of this space are translations and rotations that we indicate, respectively, by 3-vectors $\vec{\tau} = (\tau_x, \tau_y, \tau_z)^\top \in \mathsf{T}_3 = \mathsf{R}^3$, and 3×3 real, special[2] orthogonal matrices $\mathbf{R}(\phi, \theta, \psi) \in \mathsf{SO}(3)$, $\mathbf{R}\,\mathbf{R}^\top = \mathbf{R}^\top\mathbf{R} = \mathbf{1}$, where ϕ, θ, and ψ are the Euler angles of rotation around the z-, x-, and z-axes.[3] The action of these transformations can be written as a 4×4 matrix in $3 + 1$ block-diagonal form,[4]

$$\mathbf{E}(\vec{\tau}, \mathbf{R})\begin{pmatrix} \vec{r} \\ 1 \end{pmatrix} := \begin{pmatrix} \mathbf{R} & \vec{\tau} \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \vec{r} \\ 1 \end{pmatrix} = \begin{pmatrix} \mathbf{R}\vec{r} + \vec{\tau} \\ 1 \end{pmatrix}. \tag{1}$$

It is immediate to verify through this 4×4 matrix realization that the set of transformations (1) form a *group*, that is, they satisfy the four axioms:

$$\text{composition}: \quad \mathbf{E}(\vec{\tau}_1, \mathbf{R}_1)\,\mathbf{E}(\vec{\tau}_2, \mathbf{R}_2) = \mathbf{E}(\vec{\tau}_1 + \mathbf{R}_1\vec{\tau}_2, \mathbf{R}_1\mathbf{R}_2), \tag{2}$$

$$\text{identity}: \quad \mathbf{E}(\vec{0}, \mathbf{1}) = \mathbf{1} \quad (4 \times 4 \text{ unit}), \tag{3}$$

[1] We indicate by $^\top$ the transpose of an array.

[2] That is, of unit determinant; thus, reflections across one coordinate, or 3D inversions, are not included in this group.

[3] A more common (if older) parametrization rotates around the z-, y-, and z-axes; ours has the advantage of generalizing easily to D dimensions by rotating successively in the 1-2, 2-3, 3-4, …, $(D-1)$-D planes.

[4] We use the notation $A := B$ when the symbol A is defined by the expression B.

$$\text{inverse}: \quad \mathbf{E}(\vec{\tau},\mathbf{R})^{-1} = \mathbf{E}(-\mathbf{R}^{-1}\vec{\tau},\mathbf{R}^{-1}), \tag{4}$$

$$\begin{aligned}\text{associativity}: \quad & \mathbf{E}(\vec{\tau}_1,\mathbf{R}_1)(\mathbf{E}(\vec{\tau}_2,\mathbf{R}_2)\mathbf{E}(\vec{\tau}_3,\mathbf{R}_3)) \\ & = (\mathbf{E}(\vec{\tau}_1,\mathbf{R}_1)\mathbf{E}(\vec{\tau}_2,\mathbf{R}_2))\mathbf{E}(\vec{\tau}_3,\mathbf{R}_3).\end{aligned} \tag{5}$$

This is the group of *inhomogeneous special orthogonal* transformations, denoted $\mathsf{ISO}(3)$ or, in common parlance, the *Euclidean* group E_3 in three dimensions. Its elements can be factored into translations and rotations, as

$$\mathbf{E}(\vec{\tau},\mathbf{R}) = \mathbf{E}(\vec{\tau},\mathbf{1})\mathbf{E}(\vec{0},\mathbf{R}). \tag{6}$$

The *manifold* of the Euclidean group has six coordinates: three for translations $\vec{\tau} \in \mathsf{R}^3$ and three for rotations through Euler angles $(\psi, \theta, \phi) \in \mathsf{S}^3$, where S^3 is the 3D sphere (in a 4D ambient space). Eqs. (2) and (6) also indicate that, while the group contains the two *subgroups* of translations and of rotations, they are *not* on the same footing since rotations act on translations but not vice versa. The structure of the Euclidean group is that of a *semidirect* product (Gilmore, 1978; Sudarshan & Mukunda, 1974; Wybourne, 1974) indicated

$$\mathsf{E}_3 := \mathsf{ISO}(3) = \mathsf{T}_3 \,\circledS\, \mathsf{SO}(3). \tag{7}$$

In such a composition the left factor (translations) is called the *invariant* subgroup, while the right factor (rotations) is the *factor* subgroup.

In (2) we see that the composition functions for the group parameters of a product of elements are *analytic* functions of the parameters of the factors. Thus E_3 is a *Lie* group, whose (local) structure is determined by the infinitesimal neighborhood of the identity element. When we abbreviate all coordinates by $\xi := (\vec{\tau},\mathbf{R})$, the action of a Euclidean group element $E(\xi')$ on functions $f(\xi)$ of the group manifold is $E(\xi') : f(\xi) = f(E(\xi')^{-1}\xi)$.[5] In the Taylor series around the identity element (3), the first derivatives provide the *generators* of the one-parameter subgroup lines and form the Lie *algebra*[6] e_3 of the Euclidean group. They yield the familiar operators of translation and rotation in skew-Hermitian form and on the six coordinates of the E_3 manifold,

[5] The *inverse* of $E(\xi')$ in the argument of $f(\xi)$ ensures that the action of two or more group transformations of the manifold $\{\xi\}$ maintain their order of application.

[6] A Lie algebra is the real vector space spanned by the generators of the group, with one extra operation: the *Lie bracket* $\{\!|\hat{A}, \hat{B}|\!\}$. This operation is bilinear $\{\!|\hat{A}, b\hat{B}+c\hat{C}|\!\} = b\{\!|\hat{A}, \hat{B}|\!\} + c\{\!|\hat{A}, \hat{C}|\!\}$, skew-symmetric $\{\!|\hat{A}, \hat{B}|\!\} = -\{\!|\hat{B}, \hat{A}|\!\}$, and satisfies the Jacobi's identity $\{\!|\hat{A}, \{\!|\hat{B}, \hat{C}|\!\}|\!\} + \{\!|\hat{B}, \{\!|\hat{C}, \hat{A}|\!\}|\!\} + \{\!|\hat{C}, \{\!|\hat{A}, \hat{B}|\!\}|\!\} = 0.$

$$\hat{T}_x^{\rm E} = -\partial_{\tau_x}, \quad \hat{T}_y^{\rm E} = -\partial_{\tau_y}, \quad \hat{T}_z^{\rm E} = -\partial_{\tau_z}, \tag{8}$$

$$\hat{J}_x^{\rm E} = -\tau_y\partial_{\tau_z} + \tau_z\partial_{\tau_y} + \cot\theta\cos\phi\,\partial_\phi + \sin\phi\,\partial_\theta - \frac{\cos\phi}{\sin\theta}\,\partial_\psi, \tag{9}$$

$$\hat{J}_y^{\rm E} = -\tau_z\partial_{\tau_x} + \tau_x\partial_{\tau_z} + \cot\theta\sin\phi\,\partial_\phi - \cos\phi\,\partial_\theta - \frac{\sin\phi}{\sin\theta}\,\partial_\psi, \tag{10}$$

$$\hat{J}_z^{\rm E} = -\tau_x\partial_{\tau_y} + \tau_y\partial_{\tau_x} - \partial_\phi. \tag{11}$$

The translation generators (8) perform as $\exp(\alpha_i\hat{T}_i^{\rm E})f(\tau_i, R) = f(\tau_i - \alpha_i, R)$ and do not affect the rotation parameters R, while the rotation generators (9)–(11) consist of two summands, which act on the translation and on the rotation parameters. Their *commutators*[7] reflect this:

$$[\hat{T}_i, \hat{T}_j] = 0, \quad [\hat{J}_i, \hat{T}_j] = \hat{T}_k, \quad [\hat{J}_i, \hat{J}_j] = \hat{J}_k, \tag{12}$$

where i, j, k are cyclic permutations of 1, 2, 3, respectively. This structure is common to Euclidean generators in all realizations below and characterizes the Euclidean algebra e_3 as a *semidirect sum* of the translation and rotation Lie algebras, $\mathsf{e}_3 := \mathsf{iso}(3) = \mathsf{t}_3 \oplus\!\!\!\!\leftarrow \mathsf{so}(3)$, following from (7) and written with lowercase letters.

3. THE FUNDAMENTAL OBJECT OF A MODEL

In the geometric model of optics in a 3D vacuum, light rays are idealized as straight directed lines in space, while in the wave model, wavefields are integrated out of directed plane waves, which in turn can be built out of Dirac-δ 2D planes in space (Luneburg, 1964; Wolf, 1989). Through translations T_3 and rotations $\mathsf{SO}(3)$, both can be brought to the following *fundamental objects*:

$$\mathcal{O}_{\rm G} := \text{the } z\text{-axis line}, \tag{13}$$

$$\mathcal{O}_{\rm W} := \text{the } x\text{–}y \text{ plane}. \tag{14}$$

Fundamental objects are determined by their *symmetry* groups: $\mathcal{O}_{\rm G}$ is invariant under translations along and rotations around the z-axis (but not inversions $z \leftrightarrow -z$; the line is *directed*), while $\mathcal{O}_{\rm W}$ is invariant under translations in the x–y plane and rotations around the z-axis (but not inversions).

[7] The commutators $[A, B] := AB - BA$ are a realization of Lie brackets.

Indicating by calligraphic font the abstract group elements realized by the 4×4 (boldface) matrices in (1), their respective invariance subgroups are

$$\begin{aligned} &\mathcal{H}_{\mathrm{G}}(s;\psi) : \mathcal{O}_{\mathrm{G}} = \mathcal{O}_{\mathrm{G}}, \\ &\mathcal{H}_{\mathrm{G}}(s,\psi) := \mathcal{E}\Big((0,0,s)^{\top}, \mathbf{R}_z(\psi)\Big) \in \mathsf{T}_z \otimes \mathsf{SO}(2)_z \subset \mathsf{E}_3, \end{aligned} \tag{15}$$

$$\begin{aligned} &\mathcal{H}_{\mathrm{W}}(t_x,t_y;\psi) : \mathcal{O}_{\mathrm{W}} = \mathcal{O}_{\mathrm{W}}, \\ &\mathcal{H}_{\mathrm{W}}(t_x,t_y;\psi) := \mathcal{E}\Big((t_x,t_y,0)^{\top}, \mathbf{R}_z(\psi)\Big) \in \mathsf{ISO}(2)_{x,y} \subset \mathsf{E}_3, \end{aligned} \tag{16}$$

where $\otimes$ indicates the direct product of groups. Indeed, *any* subgroup of E_3 can be used as a symmetry group to define a "fundamental object" for some model (useful or not) in crystallography, quantum mechanics, or optics.

Now consider factoring the generic E_3 group element in the following two forms, corresponding with the geometric (13)–(15) and wave (14)–(16) models,

$$\mathcal{E}(\vec{\tau}, \mathbf{R}(\phi,\theta,\psi)) = \mathcal{E}\Big((q_x,q_y,0)^{\top}, \mathbf{R}_z(\phi)\,\mathbf{R}_x(\theta)\Big)\mathcal{H}_{\mathrm{G}}(s;\psi), \tag{17}$$

$$\mathcal{E}(\vec{\tau}, \mathbf{R}(\phi,\theta,\psi)) = \mathcal{E}\Big((0,0,u)^{\top}, \mathbf{R}_z(\phi)\,\mathbf{R}_x(\theta)\Big)\mathcal{H}_{\mathrm{W}}(t_x,t_y;\psi). \tag{18}$$

While the right factors are elements of the symmetry groups of the geometric and wave fundamental objects $\mathcal{O}_{\mathrm{G}}$ and $\mathcal{O}_{\mathrm{W}}$, the left factors are not. The structure of (17) and (18) is that of a decomposition of E_3 into *cosets* by the *set* of elements in the subgroups $\mathcal{H}_{\mathrm{G}}$ and $\mathcal{H}_{\mathrm{W}}$, respectively.[8] Cosets are subsets of the group; their main properties are that they are *disjoint* (no overlap between any two), and that their union *covers* the group (it provides all elements of the group). The parameters of the *left* factors of (17) and (18) are the coordinates of the *manifold* of cosets in each model, i.e., of all straight lines or all planes in 3-space. When we multiply (17) or (18) on the left by a generic Euclidean transformation $\mathcal{E}(\vec{\tau}', \mathbf{R}') \in \mathsf{E}_3$ and again factor out the symmetry subgroup to the right, we have maps of the space of cosets, i.e., the Euclidean transformations of the manifold of all lines, or of all planes, among each other.

The set of all straight lines in the geometric model of optics is thus a 4D manifold parametrized by $\{q_x, q_y; \theta, \phi\} \in \mathsf{R}^2 \otimes \mathsf{S}^2$, while that of all planes in the wave model is a 3D manifold with coordinates by $\{u; \theta, \phi\} \in \mathsf{R} \otimes \mathsf{S}^2$. For both cases it will be useful to define the unit 3-vector on the sphere

[8] These are *right* cosets; if the symmetry group elements were on the left, they would be *left* cosets.

$$\vec{p}(\theta,\phi)=\begin{pmatrix}p_x(\theta,\phi)\\ p_y(\theta,\phi)\\ p_z(\theta)\end{pmatrix}:=\mathbf{R}_z(\phi)\,\mathbf{R}_x(\theta)\begin{pmatrix}0\\0\\1\end{pmatrix}$$
$$=\begin{pmatrix}\cos\phi & \sin\phi & 0\\ -\sin\phi & \cos\phi & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0 & 0\\ 0 & \cos\theta & \sin\theta\\ 0 & -\sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}0\\0\\1\end{pmatrix}=\begin{pmatrix}\sin\theta\sin\phi\\ \sin\theta\cos\phi\\ \cos\theta\end{pmatrix}, \tag{19}$$

that will be shown below to take the role of a *momentum* vector.

4. THE GEOMETRIC MODEL

In this section we shall introduce the Hamiltonian structure of the geometric model of light, based on the fundamental object $\mathcal{O}_{\mathrm{G}}$ in (13) and the Euclidean transformations that give rise to the manifold of straight directed lines in space. On these we shall then impress a dynamic postulate to describe their change of direction (momentum) due to the inhomogeneity of the medium determined by the gradient of a scalar *refractive index* function over the space of positions (Goodman, 1968).

4.1 Euclidean Group and Coset Parameters

From the composition of the Euclidean 4×4 matrix realization involving (1)–(2)–(6)–(15) we can relate the Euclidean group and coset parameters in (17) through

$$\begin{aligned}\tau_x &= q_x + s\sin\theta\sin\phi, & q_x &= \tau_x - sp_x,\\ \tau_y &= q_y + s\sin\theta\cos\phi, & q_y &= \tau_y - sp_y,\\ \tau_z &= \phantom{q_y +{}} s\cos\theta, & s &= \tau_z/p_z = \tau_z\sec\theta,\end{aligned} \tag{20}$$

where the components of the unit 3-vector $\vec{p}(\theta,\phi)$ are given in (19). The geometric meaning of these coordinates is shown (projected on 2D) in Fig. 1. Each geometric light ray is a coset by $\mathcal{H}_{\mathrm{G}}(s;\psi)$, where $s\in\mathsf{R}$ draws out the line and $\psi\in\mathsf{S}^1$ rotates around it preventing the attachment of any "flag" or "polarization" to this line. The manifold of straight lines is the 4D manifold of cosets

$$\wp := \{q_x,\, q_y; p_x,\, p_y; \sigma\} \in \mathsf{R}^2 \otimes \mathsf{D}^2 \otimes \mathsf{Z}_2, \tag{21}$$

where the 2-vector $\mathbf{q} := (q_x,\, q_y)^\top \in \mathsf{R}^2$ is the intersection of the line with the $z = 0$ plane. The first two components of the ray direction unit

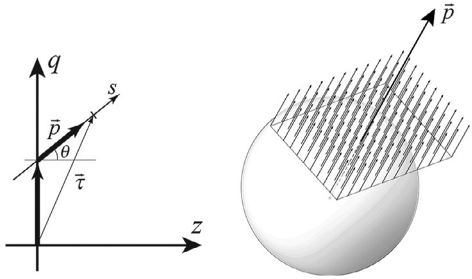

Fig. 1 *Left*: Relation between the 2D Euclidean group parameters $\{\tau_x, \tau_z; \theta\}$, $\vec{p} = (p, p_z)$, and the coset-separated parameters $\{q, p\}$ and $\{s\}$, with respect to the $\{q, z\}$ plane. *Right*: Rendering of the set of 3D geometric rays in the direction of $\vec{p}(\theta,\phi)$ parametrized by translations $\vec{\tau} \perp \vec{p}$.

3-vector $\vec{p}(\theta,\phi)$ form the 2-vector $\mathbf{p} := (p_x, p_y)^\top$; since $|\mathbf{p}| \le 1$, we must use $\sigma := \text{sign}\, p_z \in \{+,-\} =: \mathsf{Z}_2$ to distinguish between the two 2D unit disks, D^2_+ for "forward" rays where s grows in the z direction, and D^2_- for "backward" rays, where s grows along $-z$. There is a $p_z = 0$ ($\sigma = 0$) circle along the common boundary that stitches together the two disks, but being a 1D submanifold it can and will be ignored.

Beams of geometric light can be described by functions of the $\wp$ manifold of directed lines, $\rho(\mathbf{q}, \mathbf{p}, \sigma)$. The total amount of "geometric light" is then given by $\sum_\sigma \int_\wp \mathrm{d}\mu(w)\, \rho(w,\sigma)$, where $w := (\mathbf{q}, \mathbf{p})$; to determine the measure $\mathrm{d}\mu(w)$, we start with the invariant measure over the full E_3 group in terms of its six parameters, $\{\vec{\tau}, \mathbf{R}\}$ in (1), putting them in terms of the coordinates used in the coset decomposition $\{\mathbf{q}, \theta, \phi; s, \psi\}$ in (17). The measure can be found then through (20) because it *separates* into two differential forms[9]

$$\mathrm{d}^6 E(\vec{\tau},\mathbf{R}) = \mathrm{d}^3\vec{\tau}\,\mathrm{d}^3\mathbf{R}(\phi,\theta,\psi) = \mathrm{d}^4 w_{\mathrm{G}}(\mathbf{q},\mathbf{p})\,\mathrm{d}^2 h_{\mathrm{G}}(s,\psi), \tag{22}$$

$$\mathrm{d}^3\vec{\tau} = \mathrm{d}\tau_x\,\mathrm{d}\tau_y\,\mathrm{d}\tau_z, \qquad \mathrm{d}^4 w_{\mathrm{G}}(\mathbf{q},\mathbf{p}) = \mathrm{d}q_x\,\mathrm{d}q_y\,\mathrm{d}p_x\,\mathrm{d}p_y, \tag{23}$$

$$\mathrm{d}^3\mathbf{R}(\phi,\theta,\psi) = \mathrm{d}\phi\,\mathrm{d}\cos\theta\,\mathrm{d}\psi, \qquad \mathrm{d}^2 h_{\mathrm{G}}(s,\psi) = \mathrm{d}s\,\mathrm{d}\psi. \tag{24}$$

This measure $\mathrm{d}^6 E(\vec{\tau},\mathbf{R})$ is the unique (up to a numerical factor) *Haar* measure for all Euclidean transformations; it is *invariant* because, for all (fixed) $E' \in \mathsf{E}_3$, $\mathrm{d}^6(E'E) = \mathrm{d}^6 E = \mathrm{d}^6(E\,E')$. The measure on the space of rays on the z-screen (cosets) is *also* invariant: $\mathrm{d}^4(E' w_{\mathrm{G}}(\mathbf{q},\mathbf{p})) = \mathrm{d}^4 w_{\mathrm{G}}(\mathbf{q},\mathbf{p})$.

9 Differentials in coordinates that follow from coset decompositions always separate.

This states a Liouville-type conservation law: no light is created nor destroyed under rotations or translations of space.

Written in terms of the parameters of the manifold $\wp$ of directed lines (21) that constitute the geometric model, the generators of the Euclidean algebra of translations (8) and rotations (9)–(11) assume the following form

$$\hat{T}_x^{\mathrm{G}} = -\frac{\partial}{\partial q_x}, \quad \hat{T}_y^{\mathrm{G}} = -\frac{\partial}{\partial q_y}, \quad \hat{T}_z^{\mathrm{G}} = \frac{\sigma}{\sqrt{1-|\mathbf{p}|^2}}\mathbf{p}\cdot\frac{\partial}{\partial \mathbf{q}}, \tag{25}$$

$$\hat{J}_x^{\mathrm{G}} = \sigma\sqrt{1-|\mathbf{p}|^2}\frac{\partial}{\partial p_y} + \frac{\sigma\, q_y}{\sqrt{1-|\mathbf{p}|^2}}\mathbf{p}\cdot\frac{\partial}{\partial \mathbf{q}}, \tag{26}$$

$$\hat{J}_y^{\mathrm{G}} = -\sigma\sqrt{1-|\mathbf{p}|^2}\frac{\partial}{\partial p_x} - \frac{\sigma\, q_x}{\sqrt{1-|\mathbf{p}|^2}}\mathbf{p}\cdot\frac{\partial}{\partial \mathbf{q}}, \tag{27}$$

$$\hat{J}_z^{\mathrm{G}} = -\mathbf{q}\times\frac{\partial}{\partial \mathbf{q}} - \mathbf{p}\times\frac{\partial}{\partial \mathbf{p}}. \tag{28}$$

These operators satisfy the same commutation relations (12) that characterize any realization of the Euclidean algebra e_3.

4.2 Geometric and Dynamic Postulates

We have not yet said that $\wp$ is a phase space because from geometry we only showed that $(\mathbf{q}, \mathbf{p})$ is a 4D manifold with a Euclidean-invariant measure. To formulate useful geometric optics we need an extra postulate on its *dynamics*, namely the behavior of the light rays or beams, idealized as cosets $w_{\mathrm{G}}(\mathbf{q}, \mathbf{p}, \sigma)$ or distributions $\rho(\mathbf{q}, \mathbf{p}, \sigma)$, $\sigma = \operatorname{sign} p_z$, as they *evolve* when the reference $z = 0$ screen is translated along $z \in \mathsf{R}$, in a medium that is no longer the homogeneous vacuum assumed above. The rays will then generally not be straight, so we will have a *deformation* of the space of cosets $w_{\mathrm{G}}(\mathbf{q}, \mathbf{p}, \sigma) \in \wp$ that must nevertheless respect the invariance of its measure, with the conservation of its points and integrals (lest light be created or destroyed!).

To describe lines in 3-space that are generally not straight, we use the parametric *ray* 3-vector $\vec{q}(s) = (q_x(s), q_y(s), q_z(s))^\top$ for $s \in \mathsf{R}$, that will be subsequently projected as $\mathbf{q}(s)$ on the standard $z = 0$ screen. The rays $\vec{q}(s)$ are subject to the following two postulates:

- **Geometric postulate.** *Rays are continuous and piecewise differentiable.* This means that, except at points where they *break*, they have a tangent vector (indicated $\vec{p}(s)$), and they never disconnect. Since $|\mathrm{d}\vec{q}(s)| = \mathrm{d}s$, we can write

$$\frac{\mathrm{d}\vec{q}(s)}{\mathrm{d}s} = \frac{\vec{p}(s)}{|\vec{p}(s)|} =: \nabla_{\vec{p}}\, H(\vec{q},\vec{p}), \tag{29}$$

where we introduce $H(\vec{q},\vec{p}) = |\vec{p}|$ +arbitrary function of $\vec{q}$.
- **Dynamic postulate.** *The ray direction vector $\vec{p}(s)$ responds linearly to the local 3-space gradient of a real, region-wise differentiable scalar function $n(\vec{q})$ (the refractive index).* This is written as

$$\frac{\mathrm{d}\vec{p}(s)}{\mathrm{d}s} = \nabla_{\vec{q}}\, n(\vec{q}) =: -\nabla_{\vec{q}} H(\vec{q},\, \vec{p}). \tag{30}$$

From the two postulate equations we obtain $H(\vec{q},\, \vec{p}) = |\vec{p}| - n(\vec{q})$ +constant and, using the chain rule, $\mathrm{d}H(\vec{q}(s),\, \vec{p}(s))/\mathrm{d}s = 0$. Incorporating the constant into $n(\vec{q})$ so that $H = 0$, we find the tangent vector $\vec{p}$ to be of length

$$|\vec{p}| = n(\vec{q}) \quad \text{for all } \vec{q} \in \mathsf{R}^3. \tag{31}$$

Thus, to every point of the medium corresponds a sphere of radius $n(\vec{q}) > 0$, —the *Descartes* sphere—that "guides" the ray trajectory, and which proceeds obeying the two above postulates. The geometric and dynamic postulates imply two *conservation* laws: at a point $\bar{s}$ of the trajectory $\vec{q}(s)$, between neighboring points $s_\pm = \bar{s} \pm \varepsilon$ as $\varepsilon \to 0$, and with $\Delta\vec{p}(\bar{s}) = \vec{p}(s_+) - \vec{p}(s_-)$ being parallel to $\nabla_{\vec{q}}\, n(\vec{q}(\bar{s}))$, the two conservation laws are stated as

$$\text{ray continuity}: \quad \vec{q}(s_-) = \vec{q}(s_+), \tag{32}$$

$$\text{refraction law}: \quad \nabla_{\vec{q}}\, n(\vec{q}(\bar{s})) \times \vec{p}(s_+) = \nabla_{\vec{q}}\, n(\vec{q}(\bar{s})) \times \vec{p}(s_-). \tag{33}$$

These, plus piecewise differentiability of $\vec{q}(s)$, imply the two original postulates. In particular (33) yields the well-known equation

$$n_+ \sin\alpha_+ = n_- \sin\alpha_-, \qquad \begin{array}{l} n_\pm = |\vec{p}(s_\pm)| \\ \alpha_\pm := \angle\{\vec{p}(s_\pm), \nabla_{\vec{q}}\, n(\vec{q}(s_\pm))\}, \end{array} \tag{34}$$

which holds when the refractive index has a *finite* discontinuity at $\bar{s}$ and $\varepsilon \to 0$. This is of course known as the Ibn Sahl (Rashed, 1990, 1993), Snell, Descartes, and/or sine law of refraction.

4.3 Hamiltonian Structure and Phase Space

Two of the six coordinates $\{\vec{q}(s), \vec{p}(s)\}$ are redundant: the origin $s = 0$, and one of the three components of $\vec{p}(s)$ that lies on a Descartes sphere where we choose to discount the z-component. Noting that the triangle $\triangle(\mathrm{d}s, \mathrm{d}z, \mathrm{d}\mathbf{q})$ is similar and equally oriented with $\triangle(n, p_z, \mathbf{p})$, we can divide the two postulated equations, (29) and (30), by $\mathrm{d}z/\mathrm{d}s = p_z/n$, to obtain a new pair of *Hamilton* equations in the essential x, y components of $\vec{q}$ and $\vec{p}$,

$$\frac{\mathrm{d}\mathbf{q}}{\mathrm{d}z} = \frac{\mathbf{p}}{p_z} =: \frac{\partial h(\mathbf{q}, z; \mathbf{p}, \sigma)}{\partial \mathbf{p}}, \tag{35}$$

$$\frac{\mathrm{d}\mathbf{p}}{\mathrm{d}z} = \frac{n(\mathbf{q}, z)}{p_z}\frac{\partial n(\mathbf{q}, z)}{\partial \mathbf{q}} =: -\frac{\partial h(\mathbf{q}, z; \mathbf{p}, \sigma)}{\partial \mathbf{p}}, \tag{36}$$

where the *Hamiltonian function* is here

$$h(\mathbf{q}, z; \mathbf{p}, \sigma) := -p_z = -\sigma\sqrt{n(\mathbf{q}, z)^2 - |\mathbf{p}|^2} = -n(\mathbf{q}, z)\cos\theta, \tag{37}$$

and where θ is the angle between the ray direction 3-vector $\vec{p}$ and the z-axis. The 3-vector $\vec{q} = (\mathbf{q}, z)^\top$ thus includes now the evolution parameter z, while $\vec{p} = (\mathbf{p}, -h)^\top$ includes the (minus) Hamiltonian that guides its evolution. At the $z = 0$ screen and on the Descartes sphere of $|\vec{p}|$, the range of coordinates $(\mathbf{q}, \mathbf{p}, \sigma) \in \wp$ form the *phase space* manifold of the geometric model. This is a restricted definition of symplectic phase spaces, but is sufficient for our purposes. In particular, in a homogeneous medium $n(\mathbf{q}, z) = n$, $\partial n/\partial\vec{q} = \vec{0}$, free flight preserves the ray direction and its chart index σ, but shears the position coordinate of $\wp$

$$\begin{aligned} \mathbf{q}(z) &= \mathbf{q}(0) + z\,\mathbf{p}(0)/p_z(0) & \mathbf{p}(z) &= \mathbf{p}(0), \\ &= \mathbf{q}(0) + z\tan\theta, & h(z) &= h(0). \end{aligned} \tag{38}$$

It is time to introduce, for conceptual and computational ease, the *Poisson operator* of a scalar function $f(\mathbf{q}, \mathbf{p})$,

$$\{f, \circ\}_{(\mathbf{q},\mathbf{p})} := \frac{\partial f(\mathbf{q}, \mathbf{p})}{\partial \mathbf{q}} \cdot \frac{\partial}{\partial \mathbf{p}} - \frac{\partial f(\mathbf{q}, \mathbf{p})}{\partial \mathbf{p}} \cdot \frac{\partial}{\partial \mathbf{q}}. \tag{39}$$

This allows us to write the Hamilton equations (35) and (36) as a $2 \times 2 =$ 4-vector equation

$$\frac{\mathrm{d}}{\mathrm{d}z}\begin{pmatrix}\mathbf{q}\\ \mathbf{p}\end{pmatrix}=\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}\begin{pmatrix}\partial/\partial\mathbf{q}\\ \partial/\partial\mathbf{p}\end{pmatrix}h(\mathbf{q},\mathbf{p},z)=-\{h,\circ\}\begin{pmatrix}\mathbf{q}\\ \mathbf{p}\end{pmatrix}, \qquad (40)$$

where we omit the chart index σ assuming the rays not to "bend over" in the interval of interest. The form (40) is attractive because it shows this system in *evolution* form by identifying $\mathrm{d}/\mathrm{d}z \leftrightarrow \left(\begin{smallmatrix}0 & 1\\ -1 & 0\end{smallmatrix}\right) \leftrightarrow -\{h,\circ\}$. Systems governed by a Hamiltonian h in this form are *Hamiltonian systems*, with coordinates $\mathbf{p}$ of ray *momentum* that are *canonically conjugate* to coordinates of ray position $\mathbf{q}$.

4.4 Canonical and Optical Transformations

The invariance we demand of a Hamiltonian system is that if Eq. (40) is valid for the ray coordinates $(\mathbf{q},\,\mathbf{p})$ with $h(\mathbf{q},\,\mathbf{p},\,z)$ on a plane screen at z, they should continue to be valid on any other screen at z', where they are registered as $(\mathbf{Q}(\mathbf{q},\,\mathbf{p}),\,\mathbf{P}(\mathbf{q},\,\mathbf{p}))$ with $h(\mathbf{Q},\mathbf{P},z')$. From (39) we introduce the *Poisson bracket* of two differentiable functions f, g of $(\mathbf{q},\,\mathbf{p})$ defined by $\{f,\,g\} := \{f,\,\circ\}g = -\{g,\,\circ\}f$, and if necessary indicate by a subindex $(\mathbf{q},\,\mathbf{p})$ the variables with respect to which the derivatives are taken.[10] Then, replacing the differentials and partial derivatives of the new coordinates into (40), we find the conditions of invariance given succinctly by

$$\{Q_i,\,Q_j\}_{(\mathbf{q},\,\mathbf{p})}=0,\quad \{Q_i,\,P_j\}_{(\mathbf{q},\,\mathbf{p})}=\delta_{i,j},\quad \{P_i,\,P_j\}_{(\mathbf{q},\,\mathbf{p})}=0, \qquad (41)$$

for $i,\,j\in\{x,\,y\}$. Poisson brackets are a skew-symmetric bilinear form satisfying the conditions of a Lie bracket plus the Leibniz identity,[11] and to whose useful properties we shall return shortly.

The set of all transformations $(\mathbf{q},\mathbf{p})\leftrightarrow(\mathbf{Q},\mathbf{P})$ that leave (41) invariant form a group because the definition is transitive. Thus we have the functional group of all (linear and nonlinear) *canonical* transformations.[12] Moreover, when it maps the geometric-optical phase space $\wp$ bijectively onto itself (respecting the projections of the two Descartes spheres) we can call it an *optical* transformation. Under optical transformations all admissible rays are conserved. In the group of all canonical transformations some

[10] Poisson brackets are also a realization of Lie brackets, where the two partners commute under ordinary multiplication.

[11] The Leibniz identity is $\{fg,\,h\} = f\{g,\,h\} + \{f,\,h\}g$ for functions f, g, h.

[12] It is called *functional*, because its elements are defined by functions $\mathbf{Q}(\mathbf{q},\,\mathbf{p})$, $\mathbf{P}(\mathbf{q},\,\mathbf{p})$ that may carry an infinite number of parameters.

order can be established noting that, for any differentiable function $f(\mathbf{q}, \mathbf{p})$, the transformation

$$\exp\left(\tau\,\{f, \circ\}\right)\begin{pmatrix}\mathbf{q}\\ \mathbf{p}\end{pmatrix} \mapsto \begin{pmatrix}\mathbf{q}'(\mathbf{q}, \mathbf{p}; \tau)\\ \mathbf{p}'(\mathbf{q}, \mathbf{p}; \tau)\end{pmatrix} \quad \text{is canonical,} \tag{42}$$

where the Taylor series expansion of the exponential operator is

$$\exp\left(\tau\,\{f, \circ\}\right) = \sum_{n=0}^{\infty} \frac{(\tau\{f, \circ\})^n}{n!} = 1 + \sum_{n=1}^{\infty} \frac{\tau^n}{n!} \underbrace{\{\,\{\,\ldots\{\,\{}_{n \text{ brackets}} f, \circ\}, \circ\}\ldots, \circ\}, \circ\}. \tag{43}$$

These exponential operators can "jump into" the arguments of any infinitely differentiable function $F(\mathbf{q}, \mathbf{p})$, as (Steinberg, 1986)

$$e^{\tau\,\{f, \circ\}} F(\mathbf{q}, \mathbf{p}) = F(e^{\tau\,\{f, \circ\}}\mathbf{q},\ e^{\tau\,\{f, \circ\}}\mathbf{p}). \tag{44}$$

Canonical transformations also preserve the volume element $\mathrm{d}^2\mathbf{q} \wedge \mathrm{d}^2\mathbf{p}$ of the space of rays/cosets (23), and form one-parameter groups $e^{\tau_1\,\{f, \circ\}}\, e^{\tau_2\,\{f, \circ\}} = e^{(\tau_1 + \tau_2)\,\{f, \circ\}}$.

In homogeneous regions, where the refractive index n is constant, the generator functions of the Euclidean algebra in Poisson bracket form, obtained from (25)–(28), adopt the readily recognized form

$$\begin{array}{lll} \hat{T}_x^{\mathrm{G}} = \{p_x, \circ\}, & \hat{T}_y^{\mathrm{G}} = \{p_y, \circ\}, & \hat{T}_z^{\mathrm{G}} = \{\sigma\sqrt{n^2 - |\mathbf{p}|^2}, \circ\}, \\ \hat{J}_x^{\mathrm{G}} = \{q_y\,\sigma\sqrt{n^2 - |\mathbf{p}|^2}, \circ\}, & \hat{J}_y^{\mathrm{G}} = -\{q_x\,\sigma\sqrt{n^2 - |\mathbf{p}|^2}, \circ\}, & \hat{J}_z^{\mathrm{G}} = \{\mathbf{q} \times \mathbf{p}, \circ\}, \end{array} \tag{45}$$

where the Hamiltonian h of (37) appears repeatedly. These operators also satisfy the Euclidean algebra e_3 commutation relations (12). The generator functions inside the Poisson bracket yield $\vec{T}^{\mathrm{G}\,2} = n^2$, $\vec{T}^{\mathrm{G}} \cdot \vec{J}^{\mathrm{G}} = 0$, and $(\vec{J}^{\mathrm{G}})^2 = n^2|\mathbf{q}|^2 - (\mathbf{p} \cdot \mathbf{q})^2 = |\vec{q} \times \vec{p}|^2\big|_{\wp}$; the last is the square of the Petzval projected on the $z = 0$ screen with $\vec{p}$ on the Descartes sphere.

It is not guaranteed that series (42) and (43) will be easy to calculate or compute, nor that it will preserve $\wp$ globally. In general, subgroups of canonical transformations with a *finite* number of parameters are more amenable to ordered discussion.

4.5 Refracting Surfaces and the Root Transformation

The transformation of the manifold of rays due to refraction by a smooth but arbitrary surface $S(x, y, z) = 0$ between a medium with refracting index n and another n' cannot be put cogently in the evolution form (42) since the transformation is "sudden" and discontinuous. Yet it is clearly a most relevant transformation in optics. Note that this is distinct from the "thin lens" approximation in the paraxial regime (to be seen in Section 6) or the "potential jolts" used in quantum mechanics, because the surface S is generally not flat, so it does not act at one given z or time.

We shall assume that the surface S can be described, at least region-wise, by $z = \zeta(\mathbf{q})$, with a well-defined 3D normal vector

$$\vec{\Sigma}(\mathbf{q}) = \begin{pmatrix} \Sigma(\mathbf{q}) \\ -1 \end{pmatrix}, \quad \Sigma(\mathbf{q}) := \begin{pmatrix} \partial\zeta(\mathbf{q})/\partial q_x \\ \partial\zeta(\mathbf{q})/\partial q_y \end{pmatrix}, \tag{46}$$

that takes the place of the gradient of the refractive index, $\nabla_{\vec{q}}\, n(\vec{q})$ in (30) whose magnitude is now infinite, but whose direction is parallel to $\vec{\Sigma}(\mathbf{q})$. We may then resort to the two conservation laws (32) and (33) for position and momentum. As shown in Fig. 2, the coordinates of a ray before and after refraction, $(\mathbf{q}, \mathbf{p})$ and $(\mathbf{q}', \mathbf{p}')$ referred to the *same* screen $z = 0$, meet at point of impact $\bar{\mathbf{q}}$ after free flight (38) by $z = \zeta(\bar{\mathbf{q}})$. There, the component of momentum tangential to the surface, $\vec{\Sigma}(\mathbf{q}) \times \vec{p}$, is conserved. We thus have two 2-vector equations whose members we separate as

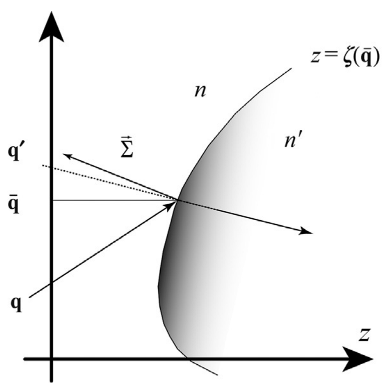

Fig. 2 An interface $z = \zeta(\mathbf{q})$ between homogeneous media n and n'. Referred to the same $z = 0$ screen, the incoming ray crosses it at $\mathbf{q}$, impacts the surface at $\bar{\mathbf{q}}$, refracts, and exits as having crossed the screen at $\mathbf{q}'$.

$$\mathbf{q} + \zeta(\bar{\mathbf{q}})\,\mathbf{p}/p_z = \bar{\mathbf{q}} = \mathbf{q}' + \zeta(\bar{\mathbf{q}})\,\mathbf{p}'/p'_z, \tag{47}$$

$$\mathbf{p} + \Sigma(\bar{\mathbf{q}})\,p_z =: \bar{\mathbf{p}} = \mathbf{p}' + \Sigma(\bar{\mathbf{q}})\,p'_z, \tag{48}$$

where we have defined $\bar{\mathbf{p}}$. Also, we note the resulting conservation of $\Sigma(\bar{\mathbf{q}}) \times (\mathbf{p} - \mathbf{p}') = 0$, and hence the coplanarity of $\vec{p}$, $\vec{p}'$ and $\vec{\Sigma}(\mathbf{q})$.

Solving the pair of simultaneous implicit equations (47) and (48) to find explicitly the transformation $\mathcal{S}_{n,\,n';\,\zeta}$ produced by the refracting surface $z = \zeta(\mathbf{q})$ between the media n and n',

$$\mathcal{S}_{n,\,n';\,\zeta} : \begin{pmatrix} \mathbf{q} \\ \mathbf{p} \end{pmatrix} = \begin{pmatrix} \mathbf{q}'(\mathbf{q},\mathbf{p};\zeta) \\ \mathbf{p}'(\mathbf{q},\mathbf{p};\zeta) \end{pmatrix}, \tag{49}$$

may seem (and is) daunting. But if we perform it via the intermediate step of using the barred coordinates $(\bar{\mathbf{q}},\bar{\mathbf{p}})$ in (47) and (48) and define the *root* transformation $\mathcal{R}_{n;\,\zeta}$ through (Navarro Saad & Wolf, 1986)

$$\bar{\mathbf{q}}(\mathbf{q},\mathbf{p}) = \mathcal{R}_{n;\,\zeta} : \mathbf{q} = \mathbf{q} + \zeta(\bar{\mathbf{q}})\,\mathbf{p}\Big/\sqrt{n^2 - |\mathbf{p}|^2}, \tag{50}$$

$$\bar{\mathbf{p}}(\mathbf{q},\mathbf{p}) = \mathcal{R}_{n;\,\zeta} : \mathbf{p} = \mathbf{p} + \zeta(\bar{\mathbf{q}})\,\sqrt{n^2 - |\mathbf{p}|^2}. \tag{51}$$

and its inverse,

$$\mathbf{q}'(\bar{\mathbf{q}},\bar{\mathbf{p}}) = \mathcal{R}^{-1}_{n';\,\zeta} : \bar{\mathbf{q}} = \bar{\mathbf{q}} - \zeta(\bar{\mathbf{q}})\,\mathbf{p}'\Big/\sqrt{n'^2 - |\mathbf{p}'|^2}, \tag{52}$$

$$\mathbf{p}'(\bar{\mathbf{q}},\bar{\mathbf{p}}) = \mathcal{R}^{-1}_{n';\,\zeta} : \bar{\mathbf{p}} = \bar{\mathbf{p}} - \zeta(\bar{\mathbf{q}})\,\sqrt{n'^2 - |\mathbf{p}'|^2}, \tag{53}$$

we will have factorized the refracting surface transformation (49) as

$$\mathcal{S}_{n,\,n';\,\zeta} = \mathcal{R}_{n;\,\zeta}\,\mathcal{R}^{-1}_{n';\,\zeta} \tag{54}$$

where each factor depends on the surface ζ and on one medium *only*.[13]

Instead of simultaneous implicit equations in two variables, Eq. (50) is implicit in $\bar{\mathbf{q}}$ only, and of the simpler form $\bar{\mathbf{q}} = \mathbf{q} + f(\bar{\mathbf{q}};\mathbf{p};n)$. When solved by repeated replacement (if at all possible), or Taylor expansions of $\zeta(\bar{\mathbf{q}})$ and $\bar{\mathbf{q}}(\mathbf{q},\mathbf{p})$ by powers in an aberration series, this can be now replaced into (51) to find explicitly $\bar{\mathbf{p}}(\mathbf{q},\mathbf{p})$. The inverse transformation (53) is now implicit in $\mathbf{p}'(\bar{\mathbf{q}},\bar{\mathbf{p}})$ and can be similarly solved, and explicitly replaced into (52).

[13] It may appear as if the two factors in (54) should be in the opposite order. As can be verified, the property of these operators to "jump into" the argument of functions as in (44), leads to the correct composition $\mathcal{S}_{n,\,n';\,\zeta} : \rho(\mathbf{q},\mathbf{p}) = \rho(\mathcal{S}_{n,\,n';\,\zeta} : \mathbf{q}, \mathcal{S}_{n,\,n';\,\zeta} : \mathbf{p})$.

Composition of both results then provides the refracting surface transformation (49). With the aid of symbolic computation (Wolf & Krötzsch, 1995) we have worked through aberration expansions up to total order seven in the four components of $\mathbf{q}$, $\mathbf{p}$, for axially symmetric refracting surfaces $\zeta(|\mathbf{q}|)$.

It may seem surprising that the root transformation (50) and (51) is *canonical* (although generally not *optical*); in fact it is only *locally* canonical, because its regions of validity in 4D phase space $\wp$ are bounded by submanifolds of rays that are *tangent* to the surface $z = \zeta(\mathbf{q})$. Beyond, they may be in another region or simply miss the surface. The proof of canonicity is quite simple (Wolf, 2004, chap. 4) when we use one of the four Hamilton *characteristic* functions, $F(\vec{q}', \vec{p})$, of a final position $\vec{q}'$ and initial momentum $\vec{p}$ to determine the remaining coordinates,

$$q_k = \frac{\partial F(\vec{q}', \vec{p})}{\partial p_k}, \quad p'_k = \frac{\partial F(\vec{q}', \vec{p})}{\partial q'_k}, \qquad k \in \{x, y, z\}. \tag{55}$$

From here it is easy to see that the basic Poisson brackets (41) are preserved. Now consider the *unit* transformation, whose characteristic function in 6D is $F_{\mathrm{id}}(\vec{q}', \vec{p}) = \vec{q}' \cdot \vec{p}$, and then restrict position $\vec{q}'$ to the surface $q_z = \zeta(\mathbf{q})$ and momentum $\vec{p}$ to the Descartes sphere $|\vec{p}| = n$ in (37). We are left with a 4D transformation whose Hamilton characteristic function is

$$F_{\mathrm{root}}(\bar{\mathbf{q}}, \mathbf{p}) := F_{\mathrm{id}}(\vec{q}', \vec{p})|_{\zeta, n} = \bar{\mathbf{q}} \cdot \mathbf{p} + \sqrt{n^2 - |\mathbf{p}|^2}. \tag{56}$$

When introduced in (55), with $\bar{\mathbf{q}}$ in place of $\vec{q}'$, and $\mathbf{p}$ in place of $\vec{p}$ this becomes the root transformation (50) and (51), proving its canonicity. The root operator $\mathcal{R}_{n;\zeta}$ canonically transforms (regions of) the phase space of rays at the standard screen $z = 0$, to regions of another phase space referred to the warped surface ζ in the medium n (Atzema, Krötzsch, & Wolf, 1997).

5. THE WAVE AND HELMHOLTZ MODELS

We return to the decomposition of the Euclidean group manifold into cosets, but now of the symmetry group $\mathcal{H}_{\mathrm{W}}(t_x, t_y; \psi) = \mathsf{E}_2$ of $\mathcal{O}_{\mathrm{W}}$ in (14) and (16), which determines the "wave" model of 2-planes in 3-space. The coordinates of the *manifold* of cosets are $\{u; \theta, \phi\}$, obtained from (18). The angular parameters $\{\theta, \phi\} \in \mathsf{S}^2$ mark the direction normal to the planes, $\vec{p}(\theta, \phi)$

in (19), which is the same as the Descartes sphere of the geometric model. Thereafter, we shall reduce the wave model to its monochromatic components that are solutions to the Helmholtz equation. There we shall have a map between functions on a sphere and fields (value *and* normal derivative) on the $z = 0$ screen. This map is unitary between the Hilbert space $\mathcal{L}^2(\mathsf{S}^2)$ of square-integrable functions on the sphere, and an interesting Hilbert space $\mathcal{H}_k$ on the screen that is characterized by a *nonlocal* inner product.

5.1 Coset Parameters for the Wave Model

For each direction $\vec{p}(\theta, \phi)$, the planes form a stack characterized by the coset parameter $u \in \mathsf{R}$ as shown in Fig. 3; it will be more convenient to instead use the normal distance from the origin, $s \in \mathsf{R}$. In place of (20) for the geometric model the change of variables is now

$$\begin{aligned} \tau_x &= t_x \cos\phi + t_y \cos\theta \sin\phi, \\ \tau_y &= -t_x \sin\phi + t_y \cos\theta \cos\phi, \qquad s := u\cos\theta, \\ \tau_z &= -t_x \sin\theta \ + u, \end{aligned} \tag{57}$$

As was the case for the geometric model (24), the invariant Haar measure of the mother group E_3 separates into the E_2-invariant measure on each plane (coset), and an invariant measure on the manifold of cosets (planes),

$$\mathrm{d}^6 g(\vec{\tau}, \mathbf{R}(\phi, \theta, \psi)) = \mathrm{d}^3 w_{\mathrm{W}}(s; \theta, \phi)\, \mathrm{d}^3 h_{\mathrm{W}}(t_x, t_y; \psi), \tag{58}$$

$$\mathrm{d}^3 w_{\mathrm{W}}(s; \theta, \phi) = \mathrm{d}s\, \sin\theta\, \mathrm{d}\theta\, \mathrm{d}\phi, \qquad \mathrm{d}^3 h_{\mathrm{W}}(t_x, t_y; \psi) = \mathrm{d}t_x\, \mathrm{d}t_y\, \mathrm{d}\psi. \tag{59}$$

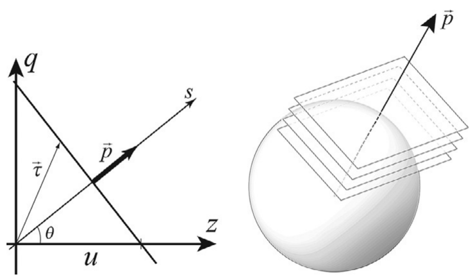

Fig. 3 *Left*: Relation between the 2D Euclidean group parameters $\{\tau_x, \tau_y; \theta\}$, $\vec{p} = (p, p_z)$, and the coset-separated parameters of the wave model, $\{u, \theta\}$ and $\{s\}$, with respect to the ambient $\{q, z\}$ space; *cf.* Fig. 1. *Right*: Rendering of the set of planes in a 3D space, parametrized by their distance $\{s\}$ to the origin and normal to the direction of $\vec{p}(\theta, \phi)$.

In this model, we have "beams" $\rho(s;\vec{p}(\theta,\phi))$ that stand for plane wavetrains $\rho_{(\theta,\phi)}(s)$ in every direction $\vec{p}(\theta,\phi)$ of the sphere, which integrate into wavefields in 3-space.

5.2 Euclidean Generators and Casimir Invariants

The Euclidean group E_3 presented in (2), is generated by a Lie *algebra* e_3, with three generators of translation $\hat{T}_i = \partial_{\tau_i}$, and three generators of rotations $\hat{J}_i$ as in quantum angular momentum theory, involving derivatives ∂_θ, ∂_ϕ, and ∂_ψ.

As before, when we change variables to coset parameters (57) and eliminate those of the symmetry subgroup, we are left with the generators of transformations in the manifold of this *wavetrain* model,

$$\begin{aligned} \hat{T}_x^{\mathrm{W}} &= -\sin\theta\sin\phi\,\partial_s, & \hat{J}_x^{\mathrm{W}} &= \cot\theta\sin\phi\,\partial_\phi - \cos\phi\,\partial_\theta, \\ \hat{T}_y^{\mathrm{W}} &= -\sin\theta\cos\phi\,\partial_s, & \hat{J}_y^{\mathrm{W}} &= \cot\theta\cos\phi\,\partial_\phi + \sin\phi\,\partial_\theta, \\ \hat{T}_z^{\mathrm{W}} &= -\cos\theta\,\partial_s, & \hat{J}_z^{\mathrm{W}} &= -\partial_\phi. \end{aligned} \tag{60}$$

These operators close under commutation into the Lie algebra e_3 with the relations (12), as all its other realizations do. The generator of translations along the s line of the wavetrain along $\vec{p}(\theta,\phi)$ is thus $\vec{p}(\theta,\phi)\,\partial_s$ and, since in vacuum $|\vec{p}|=1$, the invariant quadratic Casimir operators are

$$\hat{T}^{\mathrm{W}\,2} := \sum_{i=x,y,z} \hat{T}_i^{\mathrm{W}\,2} = \frac{\partial^2}{\partial s^2}, \qquad \sum_{i=x,y,z} \hat{T}_i^{\mathrm{W}} \hat{J}_i^{\mathrm{W}} = 0. \tag{61}$$

Functions in the eigenspace of $\hat{T}^{\mathrm{W}\,2}$ have the form

$$\rho_{(\theta,\phi)}(s) = f_k(\theta,\phi)\,\exp(\mathrm{i}ks), \quad k \in \mathsf{R} - \{0\}, \tag{62}$$

and span subspaces of eigenvalue $-k^2$ that will not mix under Euclidean transformations. These are *monochromatic* beam functions of *wavenumber* k. A monochromatic *wavefield* on $\vec{q} = (q_x, q_y, q_z)^\top \in \mathsf{R}^3$ is the integral of $f_k(\vec{p})$ over $\vec{p}$ in the Descartes sphere,

$$F_k(\vec{q}) = \frac{k}{2\pi}\int_{\mathsf{S}^2} \mathrm{d}^2 S(\vec{p})\, f_k(\vec{p})\,\exp(\mathrm{i}k\vec{p}\cdot\vec{q}), \tag{63}$$

without the now-redundant parameter s, which can be set to any constant (or integrated to a Dirac δ subsequently factored out), its phase attached to $f_k(\vec{p})$, and where we will henceforth omit the subindex with the constant k. The monochromatic wavefields (63) are solutions of the *Helmholtz* equation,

$$\left(\frac{\partial^2}{\partial q_x^2}+\frac{\partial^2}{\partial q_y^2}+\frac{\partial^2}{\partial q_z^2}\right)F(\vec{q})=-k^2\,F(\vec{q}). \tag{64}$$

5.3 Hilbert Space for Helmholtz Wavefields

As in the geometric model, we privilege the description of the objects on the $z=0$ plane screen, particularly because the 3D Helmholtz wavefields (63) actually depend on beam functions on the 2D surface of the sphere. The integration over the Descartes sphere using the two coordinates $\mathbf{p}=(p_x,p_y)^\top$ of $\vec{p}(\theta,\phi)$ requires $\sigma:=\operatorname{sign}p_z$ to distinguish between *two* unit 2D disks D^2 where $|\mathbf{p}|\le 1$, both having the measure $\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\phi=\mathrm{d}p_x\,\mathrm{d}p_y/p_z$, and $p_z=\sigma\sqrt{1-|\mathbf{p}|^2}$. Denoting by $f_\sigma(\mathbf{p})$ the beam functions in the forward or backward hemispheres, the Helmholtz wavefield on the plane of the screen and its normal derivative on the screen are given by the *wave* transform:

$$F(\mathbf{q}):=F(\vec{q})|_{q_z=0}\quad=\frac{k}{2\pi}\int_{\mathsf{D}^2}\frac{\mathrm{d}^2\mathbf{p}}{\sqrt{1-|\mathbf{p}|^2}}(f_+(\mathbf{p})+f_-(\mathbf{p}))\,e^{ik\mathbf{p}\cdot\mathbf{q}}, \tag{65}$$

$$F_z(\mathbf{q}):=\left.\frac{\partial F(\vec{q})}{\partial q_z}\right|_{q_z=0}\quad=\frac{\mathrm{i}\,k^2}{2\pi}\int_{\mathsf{D}^2}\mathrm{d}^2\mathbf{p}\,(f_+(\mathbf{p})-f_-(\mathbf{p}))\,e^{ik\mathbf{p}\cdot\mathbf{q}}. \tag{66}$$

With 2D Fourier analysis, the inverse wave transform between $f_\pm(\mathbf{p})$ on the two disks and the pair $\{F(\mathbf{q}),\,F_z(\mathbf{q})\}$ on the $z=0$ screen, is found to be

$$f_\pm(\mathbf{p})=\frac{k}{4\pi}\int_{\mathsf{R}^2}\mathrm{d}^2\mathbf{q}\left(\sqrt{1-|\mathbf{p}|^2}\,F(\mathbf{q})\pm\frac{1}{\mathrm{i}k}F_z(\mathbf{q})\right)e^{-\mathrm{i}k\mathbf{p}\cdot\mathbf{q}}. \tag{67}$$

Among wave phenomena, the energy in a wavefield is (proportional) to the integral of the absolute square of the wave function. For the beam functions on the sphere this is given naturally by the usual $\mathcal{L}^2(\mathsf{S}^2)$ inner product, $(f,g)_{\mathsf{S}^2}:=\int_{\mathsf{S}^2}\mathrm{d}^2\omega\,f(\omega)^*g(\omega)$, $\omega:=\{\theta,\phi\}$. Since the wave transform (65)–(67) is closely related with the Fourier transform, which is *unitary*, we can find the inner product for Helmholtz fields $\mathbf{F}(\mathbf{q}):=\{F(\mathbf{q}),\,F_z(\mathbf{q})\}$ over $\mathbf{q}\in\mathsf{R}^2$ on the screen through replacing (67) in $(f,g)_{\mathsf{S}^2}$ and performing one of the three resulting integrals. Thus we obtain the *nonlocal* inner product (Steinberg & Wolf, 1981; Wolf, 1989),

$$(\mathbf{F},\mathbf{G})_{\mathcal{H}_k} := \frac{k^2}{4\pi^2}\int_{\mathsf{R}^2} \mathrm{d}\mathbf{q}\int_{\mathsf{R}^2}\mathrm{d}\mathbf{q}' \times \Big(F(\mathbf{q})^*\mu'(|\mathbf{q}-\mathbf{q}'|)\,G(\mathbf{q}') + F_z(\mathbf{q})^*\mu(|\mathbf{q}-\mathbf{q}'|)\,G_z(\mathbf{q}')\Big), \tag{68}$$

which characterizes the *Helmholtz* Hilbert space $\mathcal{H}_k$, with the nonlocality measures obtained from the integration,

$$\mu'(v) = \pi\frac{j_1(v)}{v} = \pi\frac{\sin v - v\cos v}{v^3}, \quad \mu(v) = \frac{\pi}{k^2}j_0(v) = \frac{\pi}{k^2}\frac{\sin v}{v}, \tag{69}$$

with $v := k|\mathbf{q}-\mathbf{q}'|$, where $j_1(v)$ and $j_0(v)$ are the spherical Bessel functions, and where we note that $\mu'(v) = (k^2/v)\,\partial_v\mu(v)$. For models on N-dimensional screens one has the nonlocality given by Bessel functions of integer or half-integer index for N odd or even. The wave transform (65)–(67) is unitary and the Hilbert spaces $\mathcal{H}_k$ are *unique* for Euclidean-invariant systems, as shown in Steinberg and Wolf (1981). The Helmholtz wavefield energy is $\sim (\mathbf{F},\mathbf{F})_{\mathcal{H}_k}$; in this context, in González-Casanova and Wolf (1995) we have an algorithm to fit the minimum-energy Helmholtz wavefield or normal derivative to the values on a discrete and finite number of sensor points.

6. PARAXIAL MODELS

In the geometric and wave models mothered by the Euclidean group (2)–(5), we have the explicit realizations of the generating Lie algebra e_3 in Eqs. (25)–(28) and (60), respectively. The *paraxial limit* of these models is the regime where their ray directions or plane normals are infinitesimally close to the z-axis. Seen in the group E_3, we concentrate on vanishingly small rotations around the x- and y-axes, while rotations around the z-axis remain as such. We shall follow the structure of the mother Euclidean algebra as it contracts to the Heisenberg–Weyl algebra, whose group will provide the standard objects for classical and wave paraxial phase spaces, and then study their canonical transformations. Essentially, geometric paraxial phase space will be a plane $(\mathbf{q}, \mathbf{p}) \in \mathsf{R}^{2D}$ for D-dimensional screens ($D = 2$ as we have considered before), while for the wave model $\mathbf{q} \in \mathsf{R}^D$ and a phase will provide the argument for square-integrable wavefields $f(\mathbf{q}) \in \mathcal{L}^2(\mathsf{R}^D)$, compatible with the usual formalism of quantum mechanics. The transformations will correspond, in optics, to free flights, "thin" lenses, as well as harmonic waveguides of quadratic refractive index profile.

6.1 Contraction of the Euclidean to the Heisenberg–Weyl Algebra and Group

For clarity (and without much claim to rigor) we perform the *contraction* of the Euclidean algebra e_3 to the Heisenberg–Weyl algebra w_2 of 2D quantum mechanics, by rescaling the generators T_i, J_j, with the invariant $\sum_i T_i^2 = 1$, and defining

$$\begin{array}{lll} T_x^{(\varepsilon)} := \frac{1}{\varepsilon} T_x, & T_y^{(\varepsilon)} := \frac{1}{\varepsilon} T_y, & T_z^{(\varepsilon)} := T_z = \sqrt{1-\varepsilon^2(T_x^{(\varepsilon)2} + T_y^{(\varepsilon)2})}, \\ J_x^{(\varepsilon)} := \varepsilon J_x, & J_y^{(\varepsilon)} := \varepsilon J_y, & J_z^{(\varepsilon)} := J_z. \end{array} \tag{70}$$

Now let $\{\!|\circ,\circ|\!\}$ again stand for *Lie brackets*, which are Poisson brackets $\{\circ,\circ\}$ of functions of phase space in the geometric models, or commutators $[\circ,\circ]$ of operators acting on wavefunctions. The Lie bracket relations (12) for the rescaled generators (70) become

$$\begin{array}{ll} \{\!|T_i^{(\varepsilon)}, T_j^{(\varepsilon)}|\!\} = 0, & \{\!|J_z^{(\varepsilon)}, J_x^{(\varepsilon)}|\!\} = J_y^{(\varepsilon)} \\ \{\!|J_i^{(\varepsilon)}, T_j^{(\varepsilon)}|\!\} = T_k^{(\varepsilon)}, & \{\!|J_z^{(\varepsilon)}, J_y^{(\varepsilon)}|\!\} = -J_x^{(\varepsilon)} \\ & \{\!|J_x^{(\varepsilon)}, J_y^{(\varepsilon)}|\!\} = \varepsilon^2 J_z^{(\varepsilon)}, \end{array} \tag{71}$$

with i, j, k cyclic. When $\varepsilon \to 0$ reaches the limit, the structure of the Lie algebra *changes*: the z-translation generator in (70) becomes the unit $T_z^{(0)} = \hat{1}$ that has null Lie brackets with all others. The Lie bracket relations (71), regrouped as convenient, in the limit become

$$\begin{array}{lll} \{\!|T_i^{(0)}, T_j^{(0)}|\!\} = 0, & & \{\!|J_z^{(0)}, J_x^{(0)}|\!\} = J_y^{(0)}, \\ \{\!|J_i^{(0)}, T_i^{(0)}|\!\} = 0, & \{\!|J_x^{(0)}, T_y^{(0)}|\!\} = \hat{1}, & \{\!|J_z^{(0)}, J_y^{(0)}|\!\} = -J_x^{(0)}, \\ \{\!|J_z^{(0)}, T_x^{(0)}|\!\} = T_y^{(0)}, & \{\!|J_y^{(0)}, T_x^{(0)}|\!\} = -\hat{1}, & \{\!|J_x^{(0)}, J_y^{(0)}|\!\} = 0. \\ \{\!|J_z^{(0)}, T_y^{(0)}|\!\} = -T_x^{(0)}, & & \end{array} \tag{72}$$

Let us now finally change notation to

$$\begin{array}{lll} Q_x := J_x^{(0)}, & P_x := -\bar{\imath}\, T_y^{(0)}, & R := \bar{\imath} J_z^{(0)}, \\ Q_y := J_y^{(0)}, & P_y := \ \bar{\imath}\, T_x^{(0)}, & \hat{1} = \bar{\imath}\, T_z^{(0)}, \end{array} \tag{73}$$

where the factor $\bar{\imath}$ is the unit 1 in the *geometric* model of phase space coordinates and Poisson brackets; in the *wave* (or quantum mechanical) realization of the algebra by self-adjoint operators, $\bar{\imath}$ is the imaginary unit i. In these terms, their common Lie brackets are

$$\begin{array}{lll} \{\!\!\{Q_i, Q_j\}\!\!\} = 0, & & \\ \{\!\!\{P_i, P_j\}\!\!\} = 0, & \{\!\!\{R, Q_x\}\!\!\} = \bar{\imath}\, Q_y, & \{\!\!\{R, P_x\}\!\!\} = \bar{\imath}\, P_y, \\ & \{\!\!\{R, Q_y\}\!\!\} = -\bar{\imath}\, Q_x, & \{\!\!\{R, P_y\}\!\!\} = -\bar{\imath}\, P_x. \\ \{\!\!\{Q_i, P_j\}\!\!\} = \bar{\imath}\, \delta_{i,j} \hat{1}, & & \end{array} \tag{74}$$

This contraction leaves $\{Q_i, P_j, \hat{1}\} \in \mathsf{w}_2$, $i, j, \in \{x, y\}$, in semidirect sum with R, the generator of $\mathsf{so}(2)$ rotations in the x–y plane, i.e., $\mathsf{e}_3 \xrightarrow{\varepsilon \to 0} \mathsf{w}_2 \oplus\!\!\!\!\!\leftarrow \mathsf{so}(2)$, which continues to be a Lie algebra with six generators.

The Heisenberg–Weyl algebra w_2 has five generators, and its exponential is the 5-dimensional group W_2, whose elements can be parametrized as

$$\begin{aligned} w(\boldsymbol{\tau}, \boldsymbol{\rho}, \upsilon) &:= \exp\left(-\mathrm{i}(\boldsymbol{\tau}\cdot\mathbf{Q} + \boldsymbol{\rho}\cdot\mathbf{P} + \upsilon\hat{1})\right) \\ &= \exp(-\mathrm{i}\boldsymbol{\tau}\cdot\mathbf{Q})\exp(-\mathrm{i}\boldsymbol{\rho}\cdot\mathbf{P})\exp\left(-\mathrm{i}\left(\upsilon - \tfrac{1}{2}\boldsymbol{\tau}\cdot\boldsymbol{\rho}\right)\hat{1}\right) \\ &= \exp(-\mathrm{i}\boldsymbol{\rho}\cdot\mathbf{P})\exp(-\mathrm{i}\boldsymbol{\tau}\cdot\mathbf{Q})\exp\left(-\mathrm{i}\left(\upsilon + \tfrac{1}{2}\boldsymbol{\tau}\cdot\boldsymbol{\rho}\right)\hat{1}\right), \end{aligned} \tag{75}$$

with $\{\boldsymbol{\tau}, \boldsymbol{\rho}, v\} \in \mathsf{R}^4 \otimes \mathsf{S}^1$.[14] The product of two group elements is then

$$w(\boldsymbol{\tau}_1, \boldsymbol{\rho}_1, \upsilon_1)\, w(\boldsymbol{\tau}_2, \boldsymbol{\rho}_2, \upsilon_2) = w(\boldsymbol{\tau}_1 + \boldsymbol{\tau}_2,\ \boldsymbol{\rho}_1 + \boldsymbol{\rho}_2,\ \upsilon_1 + \upsilon_2 + \tfrac{1}{2}\boldsymbol{\tau}_1\cdot\boldsymbol{\rho}_2), \tag{76}$$

the unit element is $w\,(\mathbf{0}, \mathbf{0}, 0)$, the inverse of (75) is $w\,(-\boldsymbol{\tau}, -\boldsymbol{\rho}, -v)$, and associativity holds.

6.2 The Heisenberg–Weyl Algebra and Group

In the phase space coordinates familiar from mechanics, the paraxial models have momentum generators P_i that stem from the e_3 space translation generators T_i, position generators Q_i that stem from the J_i rotation generators that now generate translations of momentum, and T_z that has become $\bar{\imath}\,\hat{1} \mapsto 1$, commuting with all.

6.2.1 Geometric Model

In the same way that we selected the symmetry group of a fundamental object in E_3 to parametrize its manifold of cosets, in the W_2 Heisenberg–Weyl group for geometric-optical (and classic mechanical) models we select the fundamental object given by the 1-parameter subgroup $\{e^{\upsilon 1}\}$, so that the space of its cosets is the manifold of phase space points $\{\mathbf{q}, \mathbf{p}\} \in \mathsf{R}^4$—without

[14] We may instead decree that $\upsilon \in \mathsf{R}$ to have the *covering* group of the usually understood Heisenberg–Weyl group.

phases. The action of the group on this manifold is then through the exponentiated Poisson operators,

$$e^{\boldsymbol{\tau}\cdot\{\mathbf{p},\circ\}}\, f(\mathbf{q},\mathbf{p}) = f(\mathbf{q}-\boldsymbol{\tau},\mathbf{p}), \quad e^{\boldsymbol{\rho}\cdot\{\mathbf{q},\circ\}}\, f(\mathbf{q},\mathbf{p}) = f(\mathbf{q},\mathbf{p}+\boldsymbol{\rho}), \tag{77}$$

and $e^{\upsilon\{1,\,\circ\}} = 1$. Hence from a fundamental $(\mathbf{0},\,\mathbf{0})$ point we reach all other points in this phase space by translations. The generator R in (73) rotates jointly $\mathbf{q}$ and $\mathbf{p}$ in their x–y planes.

6.2.2 Wave Model

For the wave/quantum model, the fundamental object has the symmetry subgroup $\{e^{\mathrm{i}\boldsymbol{\tau}\cdot\mathbf{Q}}\}$, whose manifold is that of positions $\{\mathbf{q}\}\in \mathsf{R}^2$ and a phase. From here we obtain the realization of the group W_2 on functions $\psi(\mathbf{q})$ of the position manifold,[15]

$$e^{-\mathrm{i}\boldsymbol{\tau}\cdot\mathbf{P}}\psi(\mathbf{q}) = \psi(\mathbf{q}-\boldsymbol{\tau}),\; e^{-\mathrm{i}\boldsymbol{\rho}\cdot\mathbf{Q}}\psi(\mathbf{q}) = e^{-\mathrm{i}\boldsymbol{\rho}\cdot\mathbf{q}}\,\psi(\mathbf{q}),\; e^{-\mathrm{i}\upsilon\hat{1}}\psi(\mathbf{q}) = e^{-\mathrm{i}\upsilon}\,\psi(\mathbf{q}), \tag{78}$$

where the generators have the well-known form

$$P_i\psi(\mathbf{q}) = -\mathrm{i}\frac{\partial}{\partial q_i}\psi(\mathbf{q}), \qquad Q_i\psi(\mathbf{q}) = q_i\psi(\mathbf{q}), \quad i\in\{x,y\}, \tag{79}$$

while $\bar{\imath}\hat{1} = \mathrm{i}\hat{1}$.

This model thus realizes W_2 by unitary transformations (78) on $\mathcal{L}^2(\mathrm{R}^2)$, and of w_2 by operators (79) that are essentially self-adjoint. Different choices in the coset decomposition of W_2 yield other realizations of the Heisenberg–Weyl algebra and group (Wolf, 1975). Admittedly, the standard realizations (77) and (78) do not need the "fundamental object" approach for their construction, which we added only for completeness to show that they stem from contraction of the Euclidean models.

[15] If we require the physical *units* of the generators: in optics positions q_i have units of distance while momenta p_j have no units; hence $\bar{\imath}$ also has units of distance and one should introduce the reduced wavelength $\lambda\!\!\!^{-} := \lambda/2\pi = 1/k$, to have $\lambda\!\!\!^{-}1$ or $\lambda\!\!\!^{-}\,\mathrm{i}\hat{1}$ in the geometric or wave optical models. In quantum mechanics on the other hand, momentum has units of mass $\times$ distance/time, and $\bar{\imath}$ has units of action, so one has $\hbar\,\mathrm{i}\hat{1}$ with a *fixed* $\hbar := h/2\pi$.

7. LINEAR TRANSFORMATIONS OF PHASE SPACE

In both the classical geometric and wave/quantum models, the Heisenberg–Weyl algebra is special in having a *center*, i.e., the generator $\hat{1}$ that has null Lie brackets with all others, $\{\!|\hat{1}, \hat{A}|\!\} = 0$, $\hat{A} \in \mathsf{w}_2$. When we introduce the additional operation of *multiplication* between elements of an algebra, we generate its *covering* algebra, whose elements are $\hat{A}\hat{B}$, $\hat{A}\hat{B}\hat{C}$, etc.; this operation is commutative in classical models, and non-commutative, $\hat{A}\hat{B} = \hat{B}\hat{A} + \{\!|\hat{A}, \hat{B}|\!\}$ in wave/quantum models. The Lie bracket of the algebra can be extended to its covering through the Leibniz identity: $\{\!|\hat{A}\hat{B}, \hat{C}|\!\} = \hat{A}\{\!|\hat{B}, \hat{C}|\!\} + \{\!|\hat{A}, \hat{C}|\!\}\hat{B}$. Unit central elements with the property $\hat{1}\hat{A} = \hat{A}\hat{1} = \hat{A}$ allow the *quadratic* extension in the covering algebra to be a Lie algebra in its own right. In this way, out of w_2 we produce the 4D *real symplectic* algebra $\mathsf{sp}(4,\mathsf{R})$ that will generate the group of linear canonical transformations of interest in optics (Goodman, 1968; Kauderer, 1994).

7.1 Geometric Model

In the classical model, whose four w_2 generator functions commute, its quadratic extension contains the following 10 quadratic generator functions

$$\begin{array}{ccccc} q_x^2, & q_x q_y, & q_y^2, & q_x p_x, & q_x p_y, \\ p_x^2, & p_x p_y, & p_y^2, & q_y p_x, & q_y p_y. \end{array} \tag{80}$$

Linear combinations of these functions belong to a 10D linear vector space where Poisson brackets between any pair give back functions within that set (for example, $\{q_x q_y, q_x p_y\} = q_x^2$); hence this vector space is a Lie *algebra* $\mathsf{sp}(4,\mathsf{R})$, whose name will be justified below. Moreover, Poisson brackets of $\mathsf{sp}(4,\mathsf{R})$ elements with the original w_2 elements return *linear* combinations of elements of the latter (for example, $\{q_x q_y, p_x\} = q_y$). The exponential Taylor series (43) of $\exp(\tau\{a, \circ\})$, $\{a, \circ\} \in \mathrm{sp}(4,\mathsf{R})$ will thus produce finite *linear* transformations of the 4D manifold $(\mathbf{q}, \mathbf{p})$ that will form the Lie *group* $\mathsf{Sp}(4,\mathsf{R})$. These transformations will be *canonical* since they are generated through Poisson bracket operators.

The 4×4 $\mathsf{Sp}(4,\mathsf{R})$ matrices do not exhaust all the 16 independent linear transformations of the 4D manifold $\mathbf{w} := (q_x, q_y, p_x, p_y)^\top \in \mathsf{R}^4$. Canonicity demands that the fundamental Poisson brackets of the w_2 generators written in (41) be respected. These we can write in block matrix form as

$$\{\mathbf{w}^\top, \mathbf{w}\} := \left\{ (\mathbf{q}, \mathbf{p}), \begin{pmatrix} \mathbf{q} \\ \mathbf{p} \end{pmatrix} \right\} := \begin{pmatrix} \{q_i, q_j\} & \{p_i, q_j\} \\ \{q_i, p_j\} & \{p_i, p_j\} \end{pmatrix} = \begin{pmatrix} 0 & -\delta_{i,j} \\ \delta_{i,j} & 0 \end{pmatrix}, \quad (81)$$

and demand that when $\mathbf{w} \mapsto \mathbf{M}\,\mathbf{w}$, with a matrix $\mathbf{M}$, we satisfy

$$\boldsymbol{\Omega} := \{\mathbf{w}^\top, \mathbf{w}\} = \{(\mathbf{M}\mathbf{w})^\top, \mathbf{M}\mathbf{w}\} = \mathbf{M}\boldsymbol{\Omega}\mathbf{M}^\top. \quad (82)$$

In 2×2 block form, with $\mathbf{M} = \begin{pmatrix} \mathbf{a} & \mathbf{b} \\ \mathbf{c} & \mathbf{d} \end{pmatrix}$ and $\boldsymbol{\Omega} = \begin{pmatrix} \mathbf{0} & -\mathbf{1} \\ \mathbf{1} & \mathbf{0} \end{pmatrix} = -\boldsymbol{\Omega}^{-1}$, this reads

$$\begin{pmatrix} \mathbf{0} & -\mathbf{1} \\ \mathbf{1} & \mathbf{0} \end{pmatrix} = \begin{pmatrix} -\mathbf{a}\mathbf{b}^\top + \mathbf{b}\mathbf{a}^\top & -\mathbf{a}\mathbf{d}^\top + \mathbf{b}\mathbf{c}^\top \\ -\mathbf{c}\mathbf{b}^\top + \mathbf{d}\mathbf{a}^\top & -\mathbf{c}\mathbf{d}^\top + \mathbf{d}\mathbf{c}^\top \end{pmatrix}. \quad (83)$$

From here we conclude six independent conditions:

$$\mathbf{a}\mathbf{b}^\top, \quad \mathbf{c}\mathbf{d}^\top \quad \text{are symmetric (and also } \mathbf{a}^\top\mathbf{c}, \quad \mathbf{b}^\top\mathbf{d}), \quad \mathbf{a}\mathbf{d}^\top - \mathbf{b}\mathbf{c}^\top = \mathbf{1}, \quad (84)$$

$$\Rightarrow \quad \mathbf{M}^{-1} = \boldsymbol{\Omega}\mathbf{M}^\top\boldsymbol{\Omega}^{-1} = \begin{pmatrix} \mathbf{d}^\top & -\mathbf{b}^\top \\ -\mathbf{c}^\top & \mathbf{a}^\top \end{pmatrix} \in \mathsf{Sp}(4, \mathsf{R}). \quad (85)$$

Matrices $\mathbf{M}$ that satisfy (82) with the nondiagonal *metric* matrix $\boldsymbol{\Omega}$ are called *symplectic* (Guillemin & Sternberg, 1984; Kauderer, 1994). The product of two symplectic matrices is symplectic, the unit and inverses are symplectic. They form the Lie group $\mathsf{Sp}(4,\mathsf{R})$ of linear canonical transformations of phase space.

There are several schemes to sensibly organize the 10 generators of $\mathsf{sp}(4,\mathsf{R})$ in (80) and their corresponding $\mathsf{Sp}(4,\mathsf{R})$ group parameters. One scheme, which may be called "optical" favors separating the phase space transformations due to three thin anamorphic lens parameters, three free anisotropic flight parameters, and four ideal magnifiers,

$$\exp\left\{\sum_{i \le j} c_{i,j} q_i q_j, \circ\right\} \begin{pmatrix} \mathbf{q} \\ \mathbf{p} \end{pmatrix} = \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{c} & \mathbf{1} \end{pmatrix} \begin{pmatrix} \mathbf{q} \\ \mathbf{p} \end{pmatrix}, \quad \mathbf{c} = \begin{pmatrix} c_{x,x} & c_{x,y} \\ c_{x,y} & c_{y,y} \end{pmatrix}, \quad (86)$$

$$\exp\left\{\sum_{i \le j} b_{i,j} p_i p_j, \circ\right\} \begin{pmatrix} \mathbf{q} \\ \mathbf{p} \end{pmatrix} = \begin{pmatrix} \mathbf{1} & -\mathbf{b} \\ \mathbf{0} & \mathbf{1} \end{pmatrix} \begin{pmatrix} \mathbf{q} \\ \mathbf{p} \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} b_{x,x} & b_{x,y} \\ b_{x,y} & b_{y,y} \end{pmatrix}, \quad (87)$$

$$\exp\left\{\sum_{i,j} a_{i,j} q_i p_j, \circ\right\} \begin{pmatrix} \mathbf{q} \\ \mathbf{p} \end{pmatrix} = \begin{pmatrix} e^{-\mathbf{a}} & \mathbf{0} \\ \mathbf{0} & e^{\mathbf{a}^\top} \end{pmatrix} \begin{pmatrix} \mathbf{q} \\ \mathbf{p} \end{pmatrix}, \quad \mathbf{a} = \begin{pmatrix} a_{x,x} & a_{x,y} \\ a_{y,x} & a_{y,y} \end{pmatrix}. \quad (88)$$

Products between elements of the first two subgroups represent all paraxial optical setups of lenses and empty spaces.[16]

Another scheme to organize the generators (80) highlights the structure of the real symplectic algebras and groups by identifying those quadratic functions, linear combinations of (80), which generate *rotations* of phase space. To this end we profit from the 4×4 representation of $\mathbf{M} \in \mathsf{Sp}(4, \mathsf{R})$, to find the representation of the generating *algebra*, $\mathbf{m} \in \mathsf{sp}(4, \mathsf{R})$ for $\mathbf{M} = \exp(\varepsilon \mathbf{m})$ as $\varepsilon \to 0$, so that $\mathbf{M} \approx \mathbf{1} + \varepsilon \mathbf{m}$. Then, from (82),

$$\boldsymbol{\Omega} = \mathbf{M}\boldsymbol{\Omega}\mathbf{M}^\top \quad \Rightarrow \quad \mathbf{m}\boldsymbol{\Omega} = -\boldsymbol{\Omega}\mathbf{m}^\top. \tag{89}$$

Matrices satisfying the last equality represent $\mathsf{sp}(4,\mathsf{R})$ and are called infinitesimal symplectic, or better *Hamiltonian* matrices, since some end up in their own right generating the dynamics of mechanical systems with quadratic Hamiltonians.

Acting on the 4D vector space $\mathbf{w} = (q_x, q_y, p_x, p_y)^\top$, we see that *skew-symmetric* matrices $\mathbf{m} = -\mathbf{m}^\top$ which satisfy (89) generate rotations because $\exp(\theta \mathbf{m}) = \exp(-\theta \mathbf{m}^\top)$ is orthogonal. There are then four linearly independent such matrices, which include two *fractional Fourier transforms* (FrFT) (Mendlovic & Ozaktas, 1993; Ozaktas & Mendlovic, 1993a, 1993b; Ozaktas, Zalevsky, & Kutay, 2001; Sudarshan, Mukunda, & Simon, 1985) that rotate in q_i–p_i planes, cross-rotation (gyrations) in the q_x–p_y and q_y–p_x planes, and rotation in the x–y planes. Their generator functions and representing skew-symmetric matrices are

$$\text{isotropic FrFT}: \;\; \ell_0 := \tfrac{1}{4}\left(p_x^2 + p_y^2 + q_x^2 + q_y^2\right) \;\leftrightarrow\; \frac{1}{2}\begin{pmatrix} \mathbf{0} & \begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix} \\ \begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix} & \mathbf{0} \end{pmatrix}, \tag{90}$$

$$\text{anisotropic FrFT}: \;\; \ell_1 := \tfrac{1}{4}\left(p_x^2 - p_y^2 + q_x^2 - q_y^2\right) \;\leftrightarrow\; \frac{1}{2}\begin{pmatrix} \mathbf{0} & \begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix} \\ \begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix} & \mathbf{0} \end{pmatrix}, \tag{91}$$

$$\text{gyrations}: \;\; \ell_2 := \tfrac{1}{2}\left(p_x p_y + q_x q_y\right) \;\leftrightarrow\; \frac{1}{2}\begin{pmatrix} \mathbf{0} & \begin{smallmatrix} 0 & -1 \\ -1 & 0 \end{smallmatrix} \\ \begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix} & \mathbf{0} \end{pmatrix}, \tag{92}$$

[16] One would still have to show that *any* $\mathsf{Sp}(4,\mathsf{R})$ can be reached through products of elements in those two subgroups—and if so, with how many elements? In the 1D case of $\mathsf{Sp}(2,\mathsf{R})$ the answer is with up to three lenses and three empty spaces (Wolf, 2004, sect. 10.5), but in D dimensions I believe the question is still open.

$$\text{rotations}: \quad \ell_3 := \tfrac{1}{2}\left(q_x p_y - q_y p_x\right) \quad \leftrightarrow \quad \frac{1}{2}\begin{pmatrix} \begin{matrix} 0 & -1 \\ 1 & 0 \end{matrix} & \mathbf{0} \\ \mathbf{0} & \begin{matrix} 0 & -1 \\ 1 & 0 \end{matrix} \end{pmatrix}. \tag{93}$$

Under Poisson brackets, these functions close into a subalgebra of $\mathsf{sp}(4,\mathsf{R})$:

$$\{\ell_i, \ell_j\} = \ell_k, \quad i, j, k \text{ cyclic, and } \{\ell_0, \ell_i\} = 0, \tag{94}$$

that we identify as $\mathrm{u}(2) = \mathsf{u}(1) \oplus \mathsf{su}(2)$, the algebra of 2×2 skew-adjoint matrices, whose center $\mathsf{u}(1)$ is ℓ_0 in (90), the isotropic harmonic oscillator Hamiltonian that generates the fractional Fourier transform, while $\mathsf{su}(2)$ is the well-known angular momentum algebra that generates 3D rotations. The representing matrices satisfy the same algebra (94) under commutation. In Fig. 4 we show the generated $\mathsf{su}(2)$ rotations that include anisotropic Fourier transforms, gyrations, and rotations. This $\mathsf{u}(2)$ is called the *Fourier algebra* $\mathsf{u}_{\mathrm{F}}(2) \subset \mathsf{sp}(4,\mathsf{R})$ of *ortho-symplectic* matrices (Simon & Wolf, 2000).

The finite $\mathsf{Sp}(4,\mathsf{R})$ matrices generated by (90)–(93) are obtained by exponentiation. Since $\exp \frac{1}{2}\theta\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} \cos\frac{1}{2}\theta & -\sin\frac{1}{2}\theta \\ \sin\frac{1}{2}\theta & \cos\frac{1}{2}\theta \end{pmatrix}$, writing $c := \cos\frac{1}{2}\theta$ and $s := \sin\frac{1}{2}\theta$ for brevity, these are respectively

$$\begin{pmatrix} c\mathbf{1} & -s\mathbf{1} \\ s\mathbf{1} & c\mathbf{1} \end{pmatrix}; \quad \begin{pmatrix} c & 0 & -s & 0 \\ 0 & c & 0 & s \\ s & 0 & c & 0 \\ 0 & -s & 0 & c \end{pmatrix}, \quad \begin{pmatrix} c & 0 & 0 & -s \\ 0 & c & -s & 0 \\ 0 & s & c & 0 \\ s & 0 & 0 & c \end{pmatrix}, \quad \begin{pmatrix} \begin{matrix} c & -s \\ s & c \end{matrix} & \mathbf{0} \\ \mathbf{0} & \begin{matrix} c & -s \\ s & c \end{matrix} \end{pmatrix}. \tag{95}$$

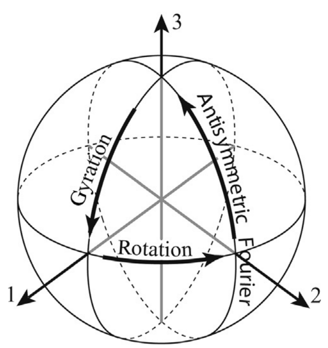

Fig. 4 The $\mathrm{SU}(2)_{\mathrm{F}}$ Fourier (sub)-group of anisotropic fractional Fourier transforms generated by (91), gyrations (92), and rotations (93). The isotropic $\mathrm{U}(1)_{\mathrm{F}}$ Fourier transforms generated by (90) commute with these and can be visualized as the product circle that closes the S^2 sphere of the figure into a S^3 sphere in a 4D ambient space.

Since the parameter ranges are all finite, $\theta \equiv \theta + 4\pi$, the total *volume* of the group is finite, i.e., it is *compact*; in fact, this $\mathsf{U}_{\mathrm{F}}(2)$ is the *maximal* compact subgroup of $\mathsf{Sp}(4,\mathsf{R})$. Note that the $\mathsf{SU}_{\mathrm{F}}(2)$ factor *covers twice* the normal rotation group $\mathsf{SO}(3)$ because of the argument $\frac{1}{2}\theta$ in the trigonometric functions, while the central factor $\mathsf{U}_{\mathrm{F}}(1)$ can be infinitely covered by R.

The remaining six independent generators of $\mathsf{sp}(4,\mathsf{R})$ are represented by symmetric matrices $\mathbf{m}$ in (89), whose exponentials follow $\exp \frac{1}{2}\zeta \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} \cosh \frac{1}{2}\zeta & \sinh \frac{1}{2}\zeta \\ \sinh \frac{1}{2}\zeta & \cosh \frac{1}{2}\zeta \end{pmatrix}$, and whose elements are unbounded. These and the previous considerations lead us to understand the $\mathsf{Sp}(4,\mathsf{R})$ 10D manifold of parameters as a higher-dimensional one-sheeted hyperboloid whose *waist* $\mathsf{U}_{\mathrm{F}}(1)$ allows for multiple covers. Its *double* cover is $\mathsf{Mp}(4,\mathsf{R})$, the *metaplectic* group.

To study the structure of $\mathsf{Sp}(4,\mathsf{R})$ it is helpful to use the accidental homomorphism of the Lie algebra $\mathsf{sp}(4,\mathsf{R})$ and the Lie algebra $\mathsf{so}(3,2)$ of infinitesimal 5×5 pseudo-orthogonal matrices under the metric $(+,+,+,-,-)$.[17] Using this feature, one can simplify the problem of finding all *inequivalent* optical systems between 2D screens, i.e., the independent matrix conjugation classes in $\mathsf{sp}(4,\mathsf{R})$, called *orbits*, obtained from $\alpha\mathbf{M}\mathbf{m}\mathbf{M}^{-1}$ by letting $\alpha \in \mathsf{R}$ and $\mathbf{M}$ roam over $\mathsf{Sp}(4,\mathsf{R})$. For $\mathsf{Sp}(2,\mathsf{R})$ this is easy: there are three orbit representatives corresponding to the 1D harmonic oscillator, the repulsive oscillator, and free flight. In $\mathsf{sp}(4,\mathsf{R})$ we have 4 continua of orbits, plus 12 points of isolated systems that are inequivalent to each other (Wolf, 2004, chap. 12; Khan & Wolf, 2002).

7.2 Wave Model: Canonical Integral Transforms

One of the most significant extensions in the theory of Fourier analysis during the last few decades is that of linear canonical transforms (Healy, Alper Kutay, Ozaktas, & Sheridan, 2015). This was born as the solution to a rather (now) evident problem in paraxial optics: the transfer function between input and output scalar wavefields, as formulated by Collins in 1970 and, almost

[17] This homomorphism is similar to the well-known infinitesimal unitary spin and rotation matrices, $\mathsf{su}(2)$ and $\mathsf{so}(3)$, as well as between infinitesimal pseudo-unitary, symplectic, and 3D Lorentz Lie algebras: $\mathsf{su}(1,1) = \mathsf{sp}(2,\mathsf{R}) = \mathsf{so}(2,1)$. Note carefully that while the Lie algebras are the *same*, their exponentiation to the corresponding Lie *groups* can lead to different *coverings*. Thus $\mathsf{SU}(2)$ covers $\mathsf{SO}(3)$ *twice*; similarly $\mathsf{SO}(2,1)$ is covered twice by $\mathsf{SU}(1,1) = \mathsf{Sp}(2,\mathsf{R})$, while $\mathsf{Sp}(2,\mathsf{R})$ is doubly covered by the group $\mathsf{Mp}(2,\mathsf{R})$ of integral transforms (to be seen in the next section), and also has an infinite cover $\overline{\mathsf{Sp}}(2,\mathsf{R})$.

simultaneously, as the solution to a question in mathematical physics, investigated by Marcos Moshinsky and Christiane Quesne, on the representation of the group of linear canonical transformations in quantum mechanics (Moshinsky & Quesne, 1971; Quesne & Moshinsky, 1971). The initially separate development of both lines of research and their intertwining provides in itself an interesting foray into the ways scientific knowledge propagates in pure and applied fields that was analyzed in Liberman and Wolf (2015).

7.2.1 Linear Maps of the Heisenberg–Weyl Algebra

Within the context of the optical models based on the quadratic extension of the Heisenberg–Weyl algebra and group, most of the exploratory work on the model of $\mathsf{Sp}(4,\mathsf{R})$ paraxial optical systems has been done above in the geometric realization. In the wave (or quantum) model we have the operators Q_i, P_j in (79) and $\mathrm{i}\hat{1}$, which also provide a quadratic extension of their wave W_2 realization.

Following Moshinsky and Quesne, for general dimension D we search for operators $\mathcal{C}_{\mathbf{M}}$, $\mathbf{M} \in \mathsf{Sp}(2D, \mathrm{R})$ such that, with the matrix inverse to $\mathbf{M} = \begin{pmatrix} \mathbf{a} & \mathbf{b} \\ \mathbf{c} & \mathbf{d} \end{pmatrix}$, namely $\mathbf{M}^{-1}$ given in (85),

$$\mathcal{C}_{\mathbf{M}} \begin{pmatrix} \mathbf{Q} \\ \mathbf{P} \end{pmatrix} \mathcal{C}_{\mathbf{M}}^{-1} = \mathbf{M}^{-1} \begin{pmatrix} \mathbf{Q} \\ \mathbf{P} \end{pmatrix} = \begin{pmatrix} \mathbf{d}^\top \mathbf{Q} - \mathbf{b}^\top \mathbf{P} \\ -\mathbf{c}^\top \mathbf{Q} + \mathbf{a}^\top \mathbf{P} \end{pmatrix}. \tag{96}$$

Now the action of this $\mathcal{C}_{\mathbf{M}}$ on functions $f(\mathbf{q}) \in \mathcal{L}^2(\mathrm{R}^D)$ can be found letting (96) act on some $f(\mathbf{q})$ to obtain 2D simultaneous equations of the form

$$\mathcal{C}_{\mathbf{M}}(\mathbf{Q} f(\mathbf{q})) = \mathbf{d}^\top \mathbf{Q}\, \mathcal{C}_{\mathbf{M}} f(\mathbf{q}) - \mathbf{b}^\top \mathbf{P}\, \mathcal{C}_{\mathbf{M}} f(\mathbf{q}), \tag{97}$$

$$\mathcal{C}_{\mathbf{M}}(\mathbf{P} f(\mathbf{q})) = -\mathbf{c}^\top \mathbf{Q}\, \mathcal{C}_{\mathbf{M}} f(\mathbf{q}) + \mathbf{a}^\top \mathbf{P}\, \mathcal{C}_{\mathbf{M}} f(\mathbf{q}). \tag{98}$$

Let now $f_{\mathbf{M}}(\mathbf{q}) := (\mathcal{C}_{\mathbf{M}} f)(\mathbf{q})$, writing $Q_j f(\mathbf{q}) = q_j f(\mathbf{q})$, $P_j f(\mathbf{q}) = -\mathrm{i}\partial_j f(\mathbf{q})$ and $\partial_j f(\mathbf{q}) := \partial f(\mathbf{q})/\partial q_j$. Then we have

$$\mathcal{C}_{\mathbf{M}}(q_i f(\mathbf{q})) = \sum_j (d_{j,i} q_j + \mathrm{i} b_{j,i} \partial_j) f_{\mathbf{M}}(\mathbf{q}), \tag{99}$$

$$-\mathrm{i}\, \mathcal{C}_{\mathbf{M}}(\partial_i f(\mathbf{q})) = \sum_j (-c_{j,i} q_j - \mathrm{i} a_{j,i} \partial_j) f_{\mathbf{M}}(\mathbf{q}). \tag{100}$$

7.2.2 Integral Transform Realization

We expect $\mathcal{C}_{\mathbf{M}} f(\mathbf{q})$ to be an *integral transform* of $f(\mathbf{q})$ because such are the Fourier transforms that belong to $\mathsf{Sp}(4D,\mathsf{R})$ in (90) and (91), and the Fresnel transform for (87), that is, with an integral kernel $C_{\mathbf{M}}(\mathbf{q}, \mathbf{q}')$, and of the form

$$\mathcal{C}_{\mathbf{M}} f(\mathbf{q}) = \int_{\mathrm{R}^D} \mathrm{d}^D \mathbf{q}'\, C_{\mathbf{M}}(\mathbf{q},\mathbf{q}') f(\mathbf{q}'). \tag{101}$$

Introducing this into (99) and (100) and integrating by parts the terms on the right-hand sides where the derivatives act on $f(\mathbf{q}')$ so that they act on the kernel, we obtain a set of differential equations that the kernel must satisfy,

$$q'_i C_{\mathbf{M}}(\mathbf{q},\mathbf{q}') = \sum_j (d_{j,i} q_j + \mathrm{i} b_{j,i} \partial_j) C_{\mathbf{M}}(\mathbf{q},\mathbf{q}'), \tag{102}$$

$$\partial'_i C_{\mathbf{M}}(\mathbf{q},\mathbf{q}') = \sum_j (\mathrm{i} c_{j,i} q_j - a_{j,i} \partial_j) C_{\mathbf{M}}(\mathbf{q},\mathbf{q}'). \tag{103}$$

Up to a constant factor, the solution is a complex Gaussian,

$$C_{\mathbf{M}}(\mathbf{q},\mathbf{q}') = K_{\mathbf{M}} \exp \mathrm{i}\left(\tfrac{1}{2}\mathbf{q}^\top \mathbf{b}^{-1}\mathbf{d}\mathbf{q} - \mathbf{q}^\top \mathbf{b}^{-1}\mathbf{q}' + \tfrac{1}{2}\mathbf{q}'^\top \mathbf{a}\mathbf{b}^{-1}\mathbf{q}'\right), \tag{104}$$

where the normalization constant K_{M} can be evaluated through asking for the Fresnel transform that corresponds to $C_{\mathrm{M(b)}}(\mathbf{q},\mathbf{q}') \to \delta^D(\mathbf{q}-\mathbf{q}')$ for $\mathbf{M}(\mathbf{b}) = \begin{pmatrix} \mathbf{1} & \mathbf{b} \\ \mathbf{0} & \mathbf{1} \end{pmatrix}$ as $\mathbf{b} \to \mathbf{0}$. Since $\mathbf{b}$ is then a symmetric matrix due to (84) it can be diagonalized to $(\beta_1, \beta_2, \ldots, \beta_D)$, and the exponent expanded to a sum over the coordinates where for each we use the Dirac-δ convergent limit for oscillating but decreasing Gaussians in the second and fourth complex-β quadrants,

$$\lim_{\beta\to 0} \frac{1}{\sqrt{2\pi\beta}} \exp\left(\mathrm{i}\frac{(q-q')^2}{2\beta}\right) = \sigma_\beta\, e^{\mathrm{i}\pi/4}\, \delta(q-q'), \quad \sigma_\beta := \begin{cases} +1, & \arg\beta \in [-\tfrac{1}{2}\pi, 0], \\ -1, & \arg\beta \in [\tfrac{1}{2}\pi, \pi]. \end{cases} \tag{105}$$

The normalization constant in (104) is then obtained as a product of limits for each coordinate,

$$K_{\mathbf{M}} = \frac{1}{\sqrt{(2\pi i)^D \det \mathbf{b}}} = \frac{e^{-\mathrm{i}\pi D/4} \exp\left(-\mathrm{i}\tfrac{1}{2}\arg\det\mathbf{b}\right)}{\sqrt{(2\pi)^D |\det\mathbf{b}|}}. \tag{106}$$

Having reached $\mathbf{b} = \mathbf{0}$, the decomposition of a generic $\mathbf{M}_o = \begin{pmatrix} \mathbf{a} & \mathbf{0} \\ \mathbf{c} & \mathbf{a}^{\top -1} \end{pmatrix} \in \mathsf{Sp}(2D, \mathsf{R})$ reads

$$(\mathcal{C}_{\mathrm{M}_o} f)(\mathbf{q}) = \frac{\exp\left(\mathrm{i}\tfrac{1}{2}\mathbf{q}^\top \mathbf{c}\mathbf{a}^{-1}\mathbf{q}\right)}{\sqrt{\det\mathbf{a}}} f(\mathbf{a}^{-1}\mathbf{q}). \tag{107}$$

Finally, when $\det\mathbf{b} = 0$ but $\mathbf{b} \neq \mathbf{0}$, we perform a similarity transformation to bring $\mathbf{b}$ to diagonal form and use (105) for its null eigenvalues.

7.2.3 Fractional Fourier and Canonical Transforms

The set of canonical integral transform kernels $C_{\mathbf{M}}(\mathbf{q},\mathbf{q}')$ in (102) are a *representation* of the group of 4×4 symplectic matrices $\mathbf{M} \in \mathsf{Sp}(2D, \mathsf{R})$, which we regard as a matrix of continuous, infinite rows and columns $(\mathbf{q},\mathbf{q}')$. Their basic group composition property was addressed by Moshinsky and Quesne (1971), who found that, for $\mathbf{M}_1\mathbf{M}_2 = \mathbf{M}_3$, their kernels compose as

$$\int_{\mathrm{R}^D} \mathrm{d}^D\mathbf{q}\; C_{\mathrm{M}_1}(\mathbf{q},\mathbf{q}')\, C_{\mathrm{M}_2}(\mathbf{q}',\mathbf{q}'') = \sigma_{1,\,2;\,3}\, C_{\mathrm{M}_2}(\mathbf{q},\mathbf{q}''), \tag{108}$$

$$\sigma_{1,\,2;\,3} := \mathrm{sign}\,(\det\mathbf{b}_3/\det\mathbf{b}_1\;\det\mathbf{b}_2), \tag{109}$$

where $\mathbf{b}_1$, $\mathbf{b}_2$, and $\mathbf{b}_3$ are the upper-right submatrices of the $\mathbf{M}$'s.

The "ambiguity sign" $\sigma_{1,\,2;\,3}$ in the group composition in (108) stems from the sign σ_β in (105) and would not go away through any redefinition of phases. Only somewhat later it was recognized that this sign is due to the multiple cover of the symplectic groups afforded by this set of kernels, whose topological features were studied by Valentin Bargmann for 1D in Bargmann (1947) and for ND in Bargmann (1970). His analysis follows the polar decomposition of complex numbers into a phase and a positive magnitude; for matrices, this is a decomposition into a unitary matrix (of the Fourier subgroup $\mathsf{U}(N)_\mathrm{F}$) and a positive definite matrix, with an appropriate choice of parameters. Unitary matrices in turn can be decomposed into the subgroup $\mathsf{SU}(N)_\mathrm{F}$ with unit determinant, and the subgroup of matrices which are the circle of isotropic fractional Fourier transforms $\mathsf{U}(1)_\mathrm{F}$ in the geometric model (90); the latter bear the onus of multivaluation because the manifolds of the other two factors are simply connected.

Fractional Fourier transforms $\mathcal{F}^\alpha$ of power α were defined in 1937 by Condon (1937) essentially on the path (103)–(105); they were rediscovered by Namias (1980), who found the kernel through the bilinear generating function of the harmonic oscillator wavefunctions, with a phase $e^{+\mathrm{i}\alpha/2}$ which guarantees that $\mathcal{F}^{\alpha_1}\mathcal{F}^{\alpha_2} = \mathcal{F}^{\alpha_1+\alpha_2}$ and $\mathcal{F}^4 = 1$ in 1D. Comparison with the canonical transform kernel[18] in (102) for the D-dimensional version shows that for angles and powers $\phi = \frac{1}{2}\pi\alpha$,

$$\mathcal{C}_{\mathbf{F}(\phi)} = \exp(\mathrm{i}D\alpha/\pi)\,\mathcal{F}^\alpha_{(D)} \quad \text{for} \quad \mathbf{F}(\phi) := \begin{pmatrix} \cos\phi\mathbf{1} & \sin\phi\mathbf{1} \\ -\sin\phi\mathbf{1} & \cos\phi\mathbf{1} \end{pmatrix}. \tag{110}$$

[18] The kernel of the Fourier transform we take as $\sim e^{-\mathrm{i}qq'}$; Namias's work uses $\sim e^{+\mathrm{i}qq'}$ instead since he follows the harmonic oscillator evolution.

We see that while the fractional Fourier transform in Condon (1937), Namias (1980), and McBride and Kerr (1987) is a fourth root of the unit transformation, $\mathcal{F}^4 = 1$, in D dimensions we have $\mathcal{C}_{\mathrm{F}}^4 = e^{-\mathrm{i}\pi D}1$, which is -1 when D is odd (including the basic 1D case), and 1 in even dimensions (including the $D = 2$ case of our concerns). The subgroup $\{\mathcal{C}_{\mathrm{F}(\phi)}\}_{\phi\in\mathsf{S}^1}$ of canonical transforms represents the quantum harmonic evolution cycle, thus distinct by a phase from the fractional Fourier transform defined in those references.

7.2.4 Canonical Transforms—Remarks and Extensions

Another special property of the wave/quantum canonical transforms is that they are generated by *second*-order differential operators, in complete algebraic correspondence with the classical counterpart of first-order Poisson operators given in (86)–(88) (Wolf, 1974). It is worth noting that the work of Lie (1888) and practically all subsequent work on Lie algebras (Gilmore, 1978) had dealt with first-order differential operators only, although the 1D time evolution generated by oscillator or waveguide Hamiltonians was used on occasion.

The triple connection between 2D×2D symplectic matrices, integral transforms, and exponentials of up-to-second-order differential operators, provides several computationally easy ways to find special function identities and Baker–Campbell–Hausdorff relations to factorize exponentials of noncommuting exponents (García-Bullé, Lassner, & Wolf, 1986).

An important special case arises when the optical setups are assumed to be axially symmetric, described by matrices $\mathbf{M} = \begin{pmatrix} a\mathbf{1} & b\mathbf{1} \\ c\mathbf{1} & d\mathbf{1} \end{pmatrix} \in \mathsf{Sp}(2,\mathsf{R})$. The integral kernel then involves Bessel functions in the radial coordinate, and phases $e^{\mathrm{i}m\theta}$ that are used to project $\sim J_m(rr'/b)$ kernels times oscillating exponentials. These *radial* canonical transform unitary kernels belong to the Bargmann *discrete* series (Bargmann, 1947) of irreducible representations of $\mathsf{Sp}(2,\mathsf{R})$, where Laguerre–Gauss beams are prominent. Another special (but less-known) case is that of *hyperbolic* canonical transforms where the $\mathbf{1}$'s in $\mathbf{M}$ are replaced by $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$'s. The kernel involves Hankel functions and lies in the *continuous* irreducible representation series of this group (Healy et al., 2015, chap. 1).

Note that the parameters of Sp (2D, R) can be complexified to build the Lie algebra $\mathsf{Sp}(2D, \mathsf{C})$, with the same properties (84) and (85), but its integral transform realization is only possible for $\mathcal{L}^2(\mathsf{R}^D)$ functions when the Gaussian factor to be integrated is in the lower complex half-plane, i.e., $\mathrm{Im}\ (\mathbf{ab}^{-1})_{i,j} \le 0$. The resulting integral transforms form a complex *semi*-group (i.e., with no guaranteed inverse), called HSp (2D,C). This allows one to treat diffusion phenomena with the same tools as for wave/quantum evolution. Thus the diffusion of harmonic, repulsive, and Airy heat distributions follows convertical ellipses, hyperbolas, and parabolas (Wolf, 1979, chap. 12), while oscillating Gaussians follow diverging lines. Moreover, with an extension of the *measure* $\mathrm{d}^D\mathbf{q}$ to an appropriate measure $\mu(\mathbf{q}, \mathbf{q}^*)$ $\mathrm{d}^D\mathrm{Re}\,\mathbf{q}\,\mathrm{d}^D\mathrm{Im}\,\mathbf{q}$ on the complex plane, these diffusive transforms can made *unitary* (Bargmann, 1970; Wolf, 1974).

8. THE METAXIAL REGIME

Beyond the quadratic extension of the Heisenberg–Weyl algebra to the real symplectic algebra, we can propose a nested family of extensions generated by polynomials in the phase space variables or operators of degree higher than those in (80), which will generate *nonlinear* (but *canonical*) transformation of phase space. First we shall build this *covering algebra* structure for 1D images, and then consider the aberration of 2D images by aligned axis-symmetric optical setups. The resulting classification of aberrations to third order follows roughly that of Seidel (1853), who was interested in image formation rather than phase space transformations, but departs from their subsequent treatment by Buchdahl (1970) for higher-order aberrations.

8.1 Aberrations in 2D Systems

In classical 2D paraxial systems, where screens are one-dimensional lines and phase space $(q, p) \in \mathsf{R}^2$ is two-dimensional, consider the monomials

$$M_{k,m}(p,q) := p^{k+m}q^{k-m}, \quad \text{of} \begin{cases} \text{rank} & k \in \{0, \frac{1}{2}, 1, \frac{3}{2}, 2, \ldots\}, \\ \text{weight} & m \in \{k, k-1, \ldots, -k\}. \end{cases} \tag{111}$$

For $k = 0$, $M_{0,0} = 1$, while for $k = \frac{1}{2}$, $M_{\frac{1}{2},\frac{1}{2}} = p$ and $M_{\frac{1}{2},-\frac{1}{2}} = q$ are the phase space variables. Then, for $k = 1$ we have the quadratic monomials $M_{1,m} = \{p^2, pq, q^2\} \in \mathsf{sp}(2,\mathsf{R})$, $m \in \{1, 0, -1\}$, that we saw in the last section. For general k, k' their Poisson brackets are

$$\{M_{k,m}, M_{k',m'}\} = 2(km' - k'm)M_{k+k'-1,m+m'}, \tag{112}$$

so they generate a countably infinite covering algebra of the classical Heisenberg–Weyl algebra which is graded by rank k.

Under linear transformations of phase space, (86)–(88) with $ad - bc = 1$, each set of $2k + 1$ monomials $M_{k,m}$ in each rank k will transform as a *multiplet*, i.e., only among themselves as,[19]

$$\mathcal{C}\begin{pmatrix} a & b \\ c & d \end{pmatrix} : \begin{pmatrix} q \\ p \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} \begin{pmatrix} q \\ p \end{pmatrix} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \begin{pmatrix} q \\ p \end{pmatrix} \Rightarrow \tag{113}$$

$$\mathcal{C}\begin{pmatrix} a & b \\ c & d \end{pmatrix} : M_{k,m}(q,p) = \sum_{m'=-k}^{k} D^k_{m,m'} \begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} M_{k,m'}(q,p), \tag{114}$$

$$D^k_{m,m'} \begin{pmatrix} a & b \\ c & d \end{pmatrix} := \sum_{m''} \binom{k+m}{m''-m'} \binom{k-m}{k-m''} \times a^{k-m''} b^{m''-m} c^{m''-m'} d^{k+m+m'-m''}, \tag{115}$$

where the matrices $D^k_{m,m'}(\mathbf{M})$ *represent* the linear canonical transformation $\mathcal{C}(\mathbf{M}) \equiv \mathcal{C}_{\mathbf{M}} \in \mathsf{Sp}(2, \mathsf{R})$.

Using (112) repeatedly on $\binom{q}{p}$, we find the action of the higher monomials $M_{k,m}$, $k > 1$ as generators of one-parameter groups of *nonlinear* transformations of phase space,

$$\exp\left(\alpha\{M_{k,m}, \circ\}\right) \begin{pmatrix} q \\ p \end{pmatrix} = \begin{pmatrix} q\left(1 + \sum_{n=1}^{\infty} \dfrac{(-\alpha)^n}{n!} c^{-}_{k,m;n} M_{n(k-1),nm}\right) \\ p\left(1 + \sum_{n=1}^{\infty} \dfrac{(+\alpha)^n}{n!} c^{+}_{k,m;n} M_{n(k-1),nm}\right) \end{pmatrix}, \tag{116}$$

with $c^{\sigma}_{k,m;n} := \prod_{s=0}^{n-1} (k + \sigma(2s-1)m)$.[20] In the series (116), the term linear in α ($n = 1$) is of degree $2k - 1$ in the phase space variables; the $\{M_{k,m}\}_{m=-k}^{k}$ are thus the generators of *aberrations* of *order* $A := 2k - 1$.[21]

Consider the following five aberrations of order 3, $\{M_{2,m}\}_{m=-2}^{2}$ where we now cut the series of $\exp \alpha\{M_{2,m}, \circ\}$ in (116) after the term linear in α, cubic in (q, p), and identify them by names that will be borne out through the spot diagrams of the case of 3D optics on 2D screens in the following section,[22]

[19] In comparing with Wolf (2004, eq. (13.5)) we note that instead of $\binom{p}{q}$ there we have here $\binom{q}{p}$. The expressions match when we exchange $a \leftrightarrow d$ and $b \leftrightarrow c$.

[20] Note that for $k = \frac{1}{2}, 1$ these coefficients become null after the first n-term.

[21] We shall here exclude half-integer ranks $\frac{1}{2}, \frac{3}{2}, \ldots$ for simplicity: they generate aberrations of even order such as due to misaligned 2D optical systems. Nevertheless they can be treated on equal footing with the axis-symmetric aberrations (Wolf, 2004, chap. 13).

[22] Perhaps I should apologize for the name of $M_{2,-2} = q^4$. It does not seem to have had any name before; its p-unfocusing property suggested the irreverent name of *pocus*.

$$\begin{array}{llll}
\text{MONOMIAL} & \text{MAP LINEAR IN } \alpha & & \text{NAME} \\
M_{2,2} = p^4, & : \begin{pmatrix} q \\ p \end{pmatrix} \mapsto \begin{pmatrix} q - 4\alpha p^3 \\ p \end{pmatrix}, & & \text{spherical aberration,} \\
M_{2,1} = p^3 q, & : \begin{pmatrix} q \\ p \end{pmatrix} \mapsto \begin{pmatrix} q - 3\alpha p^2 q \\ p + \alpha p^3 \end{pmatrix}, & & \text{coma,} \\
M_{2,0} = p^2 q^2, & : \begin{pmatrix} q \\ p \end{pmatrix} \mapsto \begin{pmatrix} q - 2\alpha p q^2 \\ p + 2\alpha p^2 q \end{pmatrix}, & & \text{astigmatism/ curvature of field,} \\
M_{2,-1} = p q^3, & : \begin{pmatrix} q \\ p \end{pmatrix} \mapsto \begin{pmatrix} q - \alpha q^3 \\ p + 3\alpha p q^2 \end{pmatrix}, & & \text{distortion,} \\
M_{2,-2} = q^4, & : \begin{pmatrix} q \\ p \end{pmatrix} \mapsto \begin{pmatrix} q \\ p + 4\alpha p^3 \end{pmatrix}, & & \textit{pocus.}
\end{array} \tag{117}$$

Yet except for $M_{k,\pm k}$, (117) are *not* canonical transformations of (q, p) because of terms in α^2 and higher; they will be canonical only if the full series (116) is kept. In order to cogently approximate the phase space transformations with (presumably *small*) aberrations, let us define *canonicity up to rank* K through *cutting* Poisson brackets between monomials (112), by defining

$$\{M_{k,m}, M_{k',m'}\}_{(K)} := \begin{cases} 2(km' - k'm) M_{k+k'-1,\, m+m'}, & k + k' - 1 \le K, \\ 0, & \text{otherwise.} \end{cases} \tag{118}$$

With the cut brackets, the transformation of (q,p) in (117) can be thus declared to be canonical—up to rank $K = 2$, or aberration order $A = 3$. Let us denote by $\mathcal{A}_k$ the linear vector space spanned by the monomials $M_{k,m}$, of elements $A_k = \sum_{m=-k}^{k} \alpha_{k,m} M_{k,m} \in \mathcal{A}_k$ with coefficients $\alpha_{k,m} \in \mathsf{R}$, which maps on itself irreducibly under linear $\mathsf{Sp}(2,\mathsf{R})$ transformations.

The next term in the series (116) is $\sim \alpha^2$; this brings in cut Poisson operator monomials $\{A_2, \{A_2, q\}\}_{(3)}$ and $\{A_2, \{A_2, p\}\}_{(3)}$, which are generally of degree 5 in the phase space variables. These terms are also brought in by the action $\{A_3, \circ\}_{(3)}$ of monomials of rank 3 (aberration order 5), which belong to the multiplet of seven generators in $\mathcal{A}_3$; the transformations will be canonical up to rank $K = 3$. With the cut Poisson brackets (118) we construct thus a Lie *group* that contains linear transformations $\mathcal{C}_{\mathbf{M}}$ (generated by the quadratic polynomials in $\mathcal{A}_1$), and aberrations of orders 3 and 5, generated by $\mathcal{A}_2$ and $\mathcal{A}_3$. The structure of this group is similar to that of the Euclidean group in Section 2; it is a semidirect product whose invariant subgroup—there translations, here aberrations of ranks 2 and 3—is now *not* abelian: $\{A_2, B_2\}_{(3)} \in \mathcal{A}_3$; the factor group was there $\mathsf{SO}(3)$ and is here $\mathbf{M} \in \mathsf{Sp}(2,\mathsf{R})$ acting through (114) on the aberration generators that we collectively indicate by $\mathbf{A} = \{A_3, A_2\}$. We thus build the *fifth-order aberration group* with elements written in the *factored-product* parametrization (Dragt, 1982a, 1982b, 1987), as

$$\mathcal{G}_3(\mathbf{A}, \mathbf{M}) = \exp(\{A_3, \circ\}_{(3)}) \exp(\{A_2, \circ\}_{(3)}) \mathcal{C}_{\mathbf{M}} \in \mathsf{A}_3 \circledast \mathsf{Sp}(2,\mathsf{R}). \quad (119)$$

The product of two $\mathsf{A}_3\mathsf{Sp}(2,\mathsf{R})$ elements, $\mathcal{G}_3(\mathbf{A},\mathbf{M})$ and $\mathcal{G}_3(\mathbf{B},\mathbf{N})$ is then

$$\begin{aligned}\mathcal{G}_3(\mathbf{A},\mathbf{M})\,\mathcal{G}_3(\mathbf{B},\mathbf{N}) &= \mathcal{G}_3(\mathbf{A},\mathbf{1})\,\mathcal{C}_{\mathbf{M}}\,\mathcal{G}_3(\mathbf{B},\mathbf{1})\,\mathcal{C}_{\mathbf{N}} \\ &= \mathcal{G}_3(\mathbf{A},\mathbf{1})\,\mathcal{G}_3(\mathcal{C}_{\mathbf{M}} : \mathbf{B},\mathbf{1})\,\mathcal{C}_{\mathbf{MN}} \\ &= \mathcal{G}_3(\mathbf{A} \sharp (\mathcal{C}_{\mathbf{M}} : \mathbf{B}), \mathbf{MN}),\end{aligned} \quad (120)$$

where $\mathcal{C}_{\mathbf{M}} : \mathbf{B} = \mathcal{C}_{\mathbf{M}}\,\mathbf{B}\,\mathcal{C}_{\mathbf{M}}^{-1}$ is the linear transformation in $\mathcal{A}_3 \cup \mathcal{A}_2$, and $\mathbf{A} \sharp \mathbf{B}$ is the compounding of aberrations through the cut Poisson bracket $\{\circ, \circ\}_{(3)}$ in (112), that we called the *gato* multiplication (Wolf, 2004). The $7 + 5$ aberration polynomials indicated by $\mathbf{A}$ contain

$$A_3(p,q) = \sum_{m=-3}^{3} \alpha_{3,m} M_{3,m}(p,q), \quad A_2(p,q) = \sum_{m=-2}^{2} \alpha_{2,m} M_{2,m}(p,q), \quad (121)$$

and their *gato* product $\mathbf{C} = \mathbf{A} \sharp \mathbf{B}$ has generating polynomials

$$C_2(p,q) = A_2(p,q) + B_2(p,q), \quad (122)$$

$$C_3(p,q) = A_3(p,q) + B_3(p,q) + \tfrac{1}{2}\{A_2, B_2\}(p,q), \quad (123)$$

while $\{A_2, B_3\}_{(3)} = 0$ and $\{A_3, B_3\}_{(3)} = 0$. The group unit is $\mathcal{G}_3(\mathbf{0},\mathbf{1})$, the inverse is $\mathcal{G}_3(\mathbf{A}, \mathbf{M})^{-1} = \mathcal{G}_3(-\mathcal{C}_{\mathbf{M}}^{-1} : \mathbf{A}, \mathbf{M}^{-1})$, and associativity holds. In Fig. 5 we show the maps produced by the monomial aberration functions in (116) on 2D phase space.

Application of $\mathcal{G}_3(\mathbf{A},\mathbf{M})$ to the phase space coordinates (q,p) first performs the linear canonical transformation $\mathcal{C}_{\mathbf{M}}$, then acts with $\exp\{A_2,\circ\}_{(3)} = 1 + \{A_2,\circ\} + \frac{1}{2}\{A_2,\{A_2,\circ\}\}$ producing a polynomial of degrees up to 5 in the phase space variables, and lastly acts with $\exp\{A_3,\circ\}_{(3)} = 1 + \{A_3,\circ\}_{(3)}$, yielding the result as a polynomial with terms of degrees 1, 3, and 5. The transformation will be canonical up to rank $K = 3$ and the number of parameters $\alpha_{k,m}$, $k = 1, 2, 3$ and $m|_{-k}^{k}$, is thus $3 + 5 + 7 = 15$. The explicit

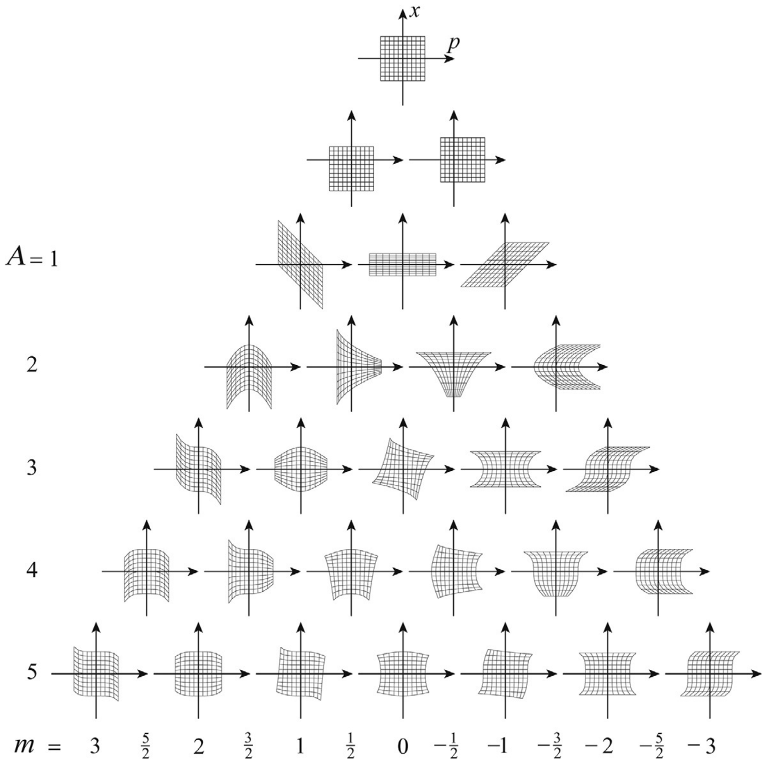

Fig. 5 Linear transformations and aberrations of classical phase space q, p, generated by the monomials (116), classified by aberration order $A = 2k - 1$ and weight m. The unit map is at the top; in the second row, Heisenberg–Weyl translations along q and p (zero aberration order). The three linear transformations ($A = 1$) are free propagation, magnification, and thin lens; higher-order aberrations of orders $A = 2, 3, 4, 5\ldots$ follow.

multiplication tables in terms of these aberration parameters for orders up to 7 are given in Wolf (2004, Part IV), which were calculated using Wolf and Krötzsch (1995).

8.2 Axis-Symmetric Aberrations in 3D Systems

Axis-symmetric linear transformations form a subgroup of the symplectic $\mathsf{Sp}(4,\mathsf{R})$ group of the paraxial model of Section 6, whose elements are block-diagonal, $\mathbf{M} = \begin{pmatrix} a\mathbf{1} & b\mathbf{1} \\ c\mathbf{1} & d\mathbf{1} \end{pmatrix}$, and whose generating Lie algebra is effectively $\mathsf{sp}(2,\mathsf{R})$. They represent 3D systems which are invariant under rotations around the z optical axis, and under reflections across planes that contain it. A basis for this algebra are the Poisson operators of the three functions[23]

$$M_{1,0,0} := |\mathbf{p}|^2, \quad M_{0,1,0} := \mathbf{p}\cdot\mathbf{q}, \quad M_{0,0,1} := |\mathbf{q}|^2. \tag{124}$$

The 3D counterpart of the monomials (111) are now

$$M_{k_+,k_0,k_-}(\mathbf{p},\mathbf{q}) := (|\mathbf{p}|^2)^{k_+}(\mathbf{p}\cdot\mathbf{q})^{k_0}(|\mathbf{q}|^2)^{k_-},$$
$$\text{of}\begin{cases}\text{rank} & k := k_+ + k_0 + k_- \in \{0,1,2,\ldots\},\\ \text{weight} & m := k_+ - k_- \in \{k, k-1, \ldots, -k\}.\end{cases} \tag{125}$$

We can organize the monomials M_{k_+,k_0,k_-} along the axes k_σ, $\sigma\in\{+,0,-\}$, where at a glance we see that this is the same diagram as that of the eigenstates of a 3D quantum harmonic oscillator, with k_σ energy quanta on the σ-axis. There is more than a passing analogy with boson $\mathsf{SU}(3)$ multiplets with (124) as quarks, of dimensions 1, 3, 6, 10, …; it directs us to reduce each rank-k multiplet with respect to a "symplectc spin-j" into sub-multiplets that will not mix among each other under the $\mathsf{Sp}(2,\mathsf{R})$ linear canonical transformations generated by (124), which are the *same* as (115) with $D^j_{m,m'}\begin{pmatrix} a & b\\ c & d\end{pmatrix}$, where

$$\begin{aligned}\text{rank } k \text{ even} &\;\Rightarrow\; j\in\{0,2,4,\ldots,k\},\\ k \text{ odd} &\;\Rightarrow\; j\in\{1,3,5,\ldots,k\}, \qquad m,m'|^j_{-j}.\end{aligned} \tag{126}$$

This classification of aberrations is shown in Fig. 6, with the 'symplectic harmonics" defined as (Wolf, 2004, sect. 14.2)

$$Y_{k,j,m}(\mathbf{p},\mathbf{q}) = (\mathbf{p}\times\mathbf{q})^{k-j}\; Y_{j,j,m}(\mathbf{p},\mathbf{q}),$$
$$Y_{j,j,m}(\mathbf{p},\mathbf{q}) = \frac{(j+m)!\,(j-m)!}{2^j(2j-1)!!}\sum_{\nu\in N(j,m)} 2^\nu \frac{|\mathbf{p}|^{j+m-\nu}}{\left[\frac{1}{2}(j+m-\nu)\right]!}\frac{(\mathbf{p}\cdot\mathbf{q})^\nu}{\nu!}\frac{|\mathbf{q}|^{j-m-\nu}}{\left[\frac{1}{2}(j-m-\nu)\right]!}, \tag{127}$$

where $N(j, m) := \{j-|m|, j-|m|-2, \ldots, 0 \text{ or } 1\}$ and $n!! := n\cdot(n-2)\ldots 2$ or 1. We note the presence of $\mathbf{p}\times\mathbf{q}$ as factor in (127), which by itself is not symmetric under reflections, but appears with even powers and is invariant under $\mathsf{Sp}(2,\mathsf{R})$. There are *six* third-order axis-symmetric 3D aberrations; in the Cartesian basis (125), $M_{2,0,0}$, $M_{1,1,0}$, …, $M_{0,0,2}$,

[23] We exclude the "angular momentum" function $\mathbf{p}\times\mathbf{q}$, which would be necessary for magnetic optics (Dragt, 2004).

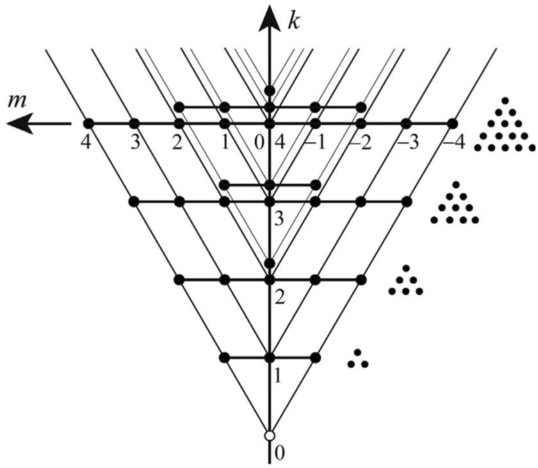

Fig. 6 The "symplectic harmonic" aberrations $Y_{k,j,m}(\mathbf{p},\mathbf{q})$ classified into multiplets j of the SO(3) rotation group (represented as *dots* joined by *horizontal lines*). For each k, on the *right*, the triangular multiplets of monomial aberrations M_{k_+,k_0,k_-} in (125) for the corresponding rank k. This is in exact analogy with the 3D quantum harmonic oscillator states.

where $M_{0,2,0}$ and $M_{1,0,1}$ have the same rank and weight k, m. In the symplectic basis (127) the last two are separated into $Y_{2,2,0}=\frac{1}{3}|\mathbf{p}|^2|\mathbf{q}|^2+\frac{2}{3}(\mathbf{p}\cdot\mathbf{q})^2$ and $Y_{2,0,0}=(\mathbf{p}\times\mathbf{q})^2$, where the former is part of the "spin" quintuplet $Y_{2,2,m}$, and the latter is an invariant singlet. In Fig. 7 we show the spot diagrams[24] of the Cartesian and "spin" multiplets of rank $k=2$.

The construction of the 3D axis-symmetric aberration group follows that of the 2D case in (119) and (120), except that $\mathcal{C}_{\mathbf{M}}:\mathbf{B}$ will now entail a 6×6 matrix in the Cartesian basis; in the symplectic spin basis this matrix is block-diagonal, with 5×5 and 1×1 submatrices. There is thus some computational advantage in using the spin basis for aberrations of higher order, but also in the geometric interpretation of the spot diagrams of these aberrations.

The use of aberration expansions may presently be obviated by fast and reliable ray-tracing computer algorithms. Still, this classification of phase space nonlinear maps clearly profits from the quantum harmonic oscillator state pattern and seems to be applicable to mechanical and other models, and extendable to a quantum/wave version. The monomials in (111) will straightforwardly "quantize" to essentially self-adjoint operators on

[24] A spot diagram in the optical context is the image of a nested set of cones (for various values of $|\mathbf{p}|$) that issue from a fixed point $\mathbf{q}$ away from the origin.

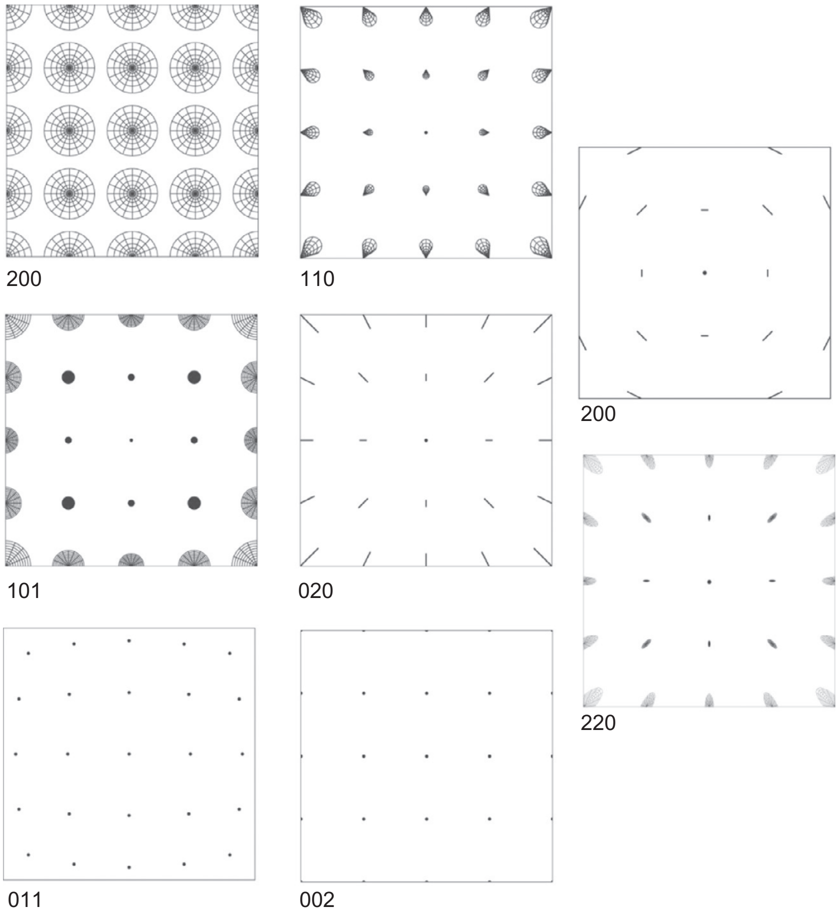

Fig. 7 *Left*: Spot diagrams of the monomial generator functions M_{k_+,k_0,k_-} in (125), indicated by the Cartesian indices $[k_+, k_0, k_-]$. *Right*: Spot diagrams of the two "symplectic harmonic" aberrations generated by $Y_{2,0,0}$ and $Y_{2,2,0}$, indicated by (k, j, m) that belong to the singlet and quintuplet irreducible representations, respectively, and are linear combinations of [1,0,1] and [0,2,0]. In [0, 1, 1] (distortion) and [0, 0, 2] (pocus) the spots are points; [0, 2, 0] and (2, 0, 0) are lines.

$\mathcal{L}^2(\mathrm{R}^2)$, only taking care to use the *Weyl symmetrization* for the noncommuting factors, which preserves their transformation properties under the linear symplectic subgroup. In Wolf (2004, chap. 15) we apply these techniques toward the correction of three fractional Fourier transform setups: a single axis-symmetric lens of polynomial surface, such a lens with a reflecting back surface, and a waveguide of polynomial refractive index profile.

9. DISCRETE OPTICAL MODELS

Many people would regard images on a finite pixelated screen as only distantly related to "continuous" geometric or wave optics; if at all, it would appear as obtained by sampling continuous images. But quite on the contrary, here we show that continuous optics is a *contraction* of the finite, pixelated, discrete model. Or conversely, we can say that the discrete model is a *deformation* of the continuous model (Boyer & Wolf, 1973; Wolf & Boyer, 1974) with a further Weyl noncompact-to-compact replacement. It is important to consider pixelated versions of optics because that is the nature of the data that is obtained from CCD sensor arrays, transformed and analyzed by computer algorithms (Pei & Ding, 2000; Pei & Yeh, 1997; Pei, Yeh, & Tseng, 1999).

With the same method we used to contract the Euclidean to the Heisenberg–Weyl Lie algebras for the paraxial model in Section 6, we shall contract the generators of rotation algebras to those of the Euclidean generators of Section 2. Then we shall show how pixelated screens contain a union of irreducible representations of the rotation group, and how they transform under the corresponding Fourier algebra (90)–(93) and group, including unitary rotations, gyrations, and preliminarily aberrations. Here, the requirement of canonicity is replaced by unitarity, i.e., reversibility and no loss of information.

9.1 The Contraction of so(4) to iso(3)

Consider the generators of rotations in a 4D space of Cartesian coordinates x_1, x_2, x_3, x_4, which span the special orthogonal Lie algebra $\mathsf{so}(4)$ of rotations in this space, realized as

$$\Lambda_{i,j} := \mathrm{i}\,(x_i\partial_j - x_j\partial_i), \qquad i,j|_1^4. \tag{128}$$

Now separate this set into two subsets: the generators $\{\Lambda_{1,2}, \Lambda_{1,3}, \Lambda_{2,3}\}$ of a subalgebra $\mathsf{so}(3)$, and the subset $\boldsymbol{\Lambda}_4 := \{\Lambda_{1,4}, \Lambda_{2,4}, \Lambda_{3,4}\}$ that transforms as a 3-vector under commutation with the former,

$$[\Lambda_{i,j}, \Lambda_{k,4}] = \mathrm{i}\,(\delta_{j,k}\Lambda_{i,4} - \delta_{i,k}\Lambda_{j,4}). \tag{129}$$

As in (70)–(74) we introduce a change of scale on the 3-vector, defining $\boldsymbol{\Lambda}_4{}^{(\varepsilon)} := \varepsilon\boldsymbol{\Lambda}_4$, with a parameter ε destined to vanish. The $\mathsf{so}(3)$ generators $J_k := \Lambda_{i,\,j}$ (i, j, k cyclic) are left unscathed, as well as the transformation property (129) of $\boldsymbol{\Lambda}_4{}^{(\varepsilon)}$ under rotation. But the commutator between any two of its components vanishes,

$$[\Lambda^{(\varepsilon)}_{i,4}, \Lambda^{(\varepsilon)}_{j,4}] = \varepsilon^2[\Lambda_{i,4}, \Lambda_{j,4}] = \mathrm{i}\,\varepsilon^2\Lambda_{j,i} \xrightarrow{\varepsilon\to 0} 0. \tag{130}$$

In the limit, J_i remain as rotation generators while $T_i = \Lambda^{(0)}_{i,4}$ become translation generators of the inhomogeneous special orthogonal group $\mathsf{ISO}(3)$. Since this E_3 was referred to as the mother group of optical models, $\mathsf{SO}(4)$ could be called the *grand-mother* group. Of course, this contraction process also applies in N dimensions.

9.2 The Plane Pixelated Screen

We are interested in the *four*-dimensional Lie algebra of rotations $\mathsf{so}(4)$ because this dimension is special: it applies to the "real" case of pixelated 2D displays, and the algebra $\mathsf{so}(4)$ is—unique among all orthogonal algebras—a *direct sum*,[25]

$$\mathsf{so}(4) = \mathsf{su}(2)_x \oplus \mathsf{su}(2)_y, \tag{131}$$

This can be seen through writing the generators (128) as

$$\begin{aligned} J^x_1 &= \tfrac{1}{2}(\Lambda_{2,3} + \Lambda_{1,4}), \quad J^x_2 = -\tfrac{1}{2}(\Lambda_{1,3} - \Lambda_{2,4}), \quad J^x_3 = \tfrac{1}{2}(\Lambda_{1,2} + \Lambda_{3,4}), \\ J^y_1 &= \tfrac{1}{2}(\Lambda_{2,3} - \Lambda_{1,4}), \quad J^y_2 = -\tfrac{1}{2}(\Lambda_{1,3} + \Lambda_{2,4}), \quad J^y_3 = \tfrac{1}{2}(\Lambda_{1,2} - \Lambda_{3,4}), \end{aligned} \tag{132}$$

and verifying that for i, j, k cyclic,

$$[J^x_i, J^x_j] = \mathrm{i}J^x_k, \quad [J^y_i, J^y_j] = \mathrm{i}J^y_k, \quad [J^x_i, J^y_j] = 0. \tag{133}$$

The rotation algebra $\mathsf{so}(4)$ has two Casimir invariant operators, $(\vec{J}^x)^2$ and $(\vec{J}^y)^2$; their eigenvalues $j_x(j_x{+}1)$ and $j_y(j_y{+}1)$ determine that the generators will have spectra composed of equidistant points $m_x|^{j_x}_{-j_x}$ and $m_y|^{j_y}_{-j_y}$, and eigenfunction multiplets of $N_x = 2j_x{+}1$ and $N_y = 2j_y{+}1$ functions.

The gist of defining a discrete model is to associate the generators of $\mathsf{su}(2)_x \oplus \mathsf{su}(2)_y$ to *operators* of *position* Q_x, Q_y with eigenvalues $q_x|^{j_x}_{-j_x}$, $q_y|^{j_y}_{-j_y}$, *momentum* P_x, P_y, and *mode* $H_x := J^x_3 + j_x 1$ and $H_y := J^y_3 + j_y 1$ with eigenvalues $n_x|^{2j_x}_0$, $n_y|^{2j_y}_0$. Then we can identify the pixels of an in general $N_x \times N_y$ rectangular screen with the elements of a complex matrix with rows and columns labeled by (q_x, q_y). On these we shall determine the *mode* functions $\Psi^{j_x,j_y}_{n_x,n_y}(q_x, q_y)$ of a finite model of the harmonic oscillator. This should not

[25] Although as Lie algebras $\mathsf{so}(3) = \mathsf{su}(2)$, when one examines the group *manifold* one finds that indeed $\mathsf{SO}(4) = \mathsf{SU}(2)_x \otimes \mathsf{SU}(2)_y$.

be surprising, since a point on a rotating sphere projects as harmonic motion on a screen.

9.3 The Kravchuk Oscillator States

Let us work first with an $N \times 1$ pixelated screen line. The identification of phase space operators suggested above sets

$$\begin{array}{llll} \text{position} & Q := J_1, & \text{momentum} & P := -J_2, \\ & K := J_3, & \text{mode} & H := K + j1. \end{array} \tag{134}$$

The $\mathsf{su}(2)$ commutation relations (133) then become

$$[K, Q] = -\mathrm{i}P, \quad [K, P] = \mathrm{i}Q, \quad [Q, P] = -\mathrm{i}K. \tag{135}$$

The first two commutators correspond to the two Hamilton equations for quantum position and momentum operators under a harmonic oscillator Hamiltonian $H_{\text{osc}} = \frac{1}{2}(P^2 + Q^2)$. However, the third $\mathsf{su}(2)$ commutator differs from the usual quantum commutator $[Q, P] = \mathrm{i}\lambda\!\!\!^{-} 1$, and this marks the difference between the finite and the "continuous" models of the harmonic oscillator.

The finite oscillator wavefunctions are the overlaps between the eigenfunctions of the position generator Q and the mode generator H. Since J_1 and J_3 are related by a $\frac{1}{2}\pi$ rotation around the J_2 axis, quantum angular momentum theory identifies the overlap as a Wigner *little-d* function (Biedenharn & Louck, 1981) for that angle, and given by

$$\Psi_n^j(q) := d^j_{n-j,q}\left(\tfrac{1}{2}\pi\right) \qquad n|_0^{2j}, \quad q|_{-j}^{j} \tag{136}$$

$$\begin{aligned} &= \frac{(-1)^n}{2^j}\sqrt{\binom{2j}{n}\binom{2j}{j+q}}\, K_n\left(j+q; \tfrac{1}{2}, 2j\right), \\ K_n\left(s; \tfrac{1}{2}, 2j\right) &= {}_2F_1(-n, -s; -2j; 2) = K_s\left(n; \tfrac{1}{2}, 2j\right), \end{aligned} \tag{137}$$

for $s|_0^{2j}$, where $K_n(s; \frac{1}{2}, 2j)$ is a symmetric Kravchuk polynomial (Krawtchouk, 1928), $\binom{m}{n}$ are the binomial coefficients, and ${}_2F_1(a, b; c; z)$ is the Gauss hypergeometric function. These we call the *Kravchuk functions* on the discrete positions of the finite oscillator model.

The Kravchuk functions belong to $\mathsf{su}(2)$ multiplets and have been detailed in several papers (Atakishiyev, Pogosyan, & Wolf, 2005; Atakishiyev & Wolf, 1997). They are shown in Fig. 8, where it can be seen that the lowest n-modes $\Psi_n^j(q)$ closely resemble the continuous

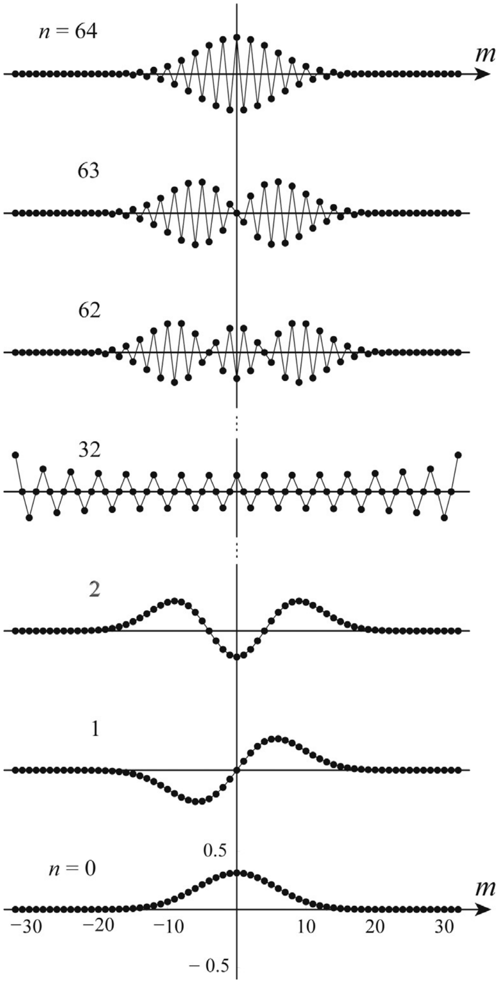

Fig. 8 The 1D Kravchuk states $\Psi_n^j(q)$ in (136), for $j = 32$, on the 65 points $m|_{-32}^{32}$ (joined by *straight lines* for visibility), and modes $n|_0^{64}$ from *bottom* to *top*. The $n = 0$ ground state of the finite oscillator is the square root of the binomial distribution, approximating the Gaussian harmonic oscillator wavefunction. The highest $n = 64$ state of the finite oscillator reproduces the ground state with a change of sign between neighboring points.

Hermite–Gauss states in continuous wave optics; they are real, have definite parity and, for higher $n > j$, they alternate in sign between each pair of neighboring points,

$$\Psi_n^j(-q) = (-1)^n \Psi_n^j(q), \qquad \Psi_{2j-n}^j(q) = (-1)^q \Psi_n^j(q). \tag{138}$$

The finite oscillator Kravchuk functions are orthonormal and complete in the N-dimensional vector space of images on the 1D pixelated screen,

$$\sum_{q=-j}^{j} \Psi_n^j(q)\, \Psi_{n'}^j(q) = \delta_{n,n'}, \quad \sum_{n=0}^{2j} \Psi_n^j(q)\, \Psi_n^j(q') = \delta_{q,q'}. \tag{139}$$

Functions $f(q)$ of $2j+1$ points $q|_{-j}^{j}$, are expanded in the Kravchuk basis as,

$$f(q) = \sum_{n=0}^{2j} f_n\, \Psi_n^j(q), \quad f_n := \sum_{q=-j}^{j} f(q)\, \Psi_n^j(q). \tag{140}$$

When the number and density of points grows without bound, $j \to \infty$, so that q becomes the real line, it can be shown that the Kravchuk functions (136) limit to the usual Hermite–Gaussian wavefunctions (Atakishiyev, Pogosyan, & Wolf, 2003). This is a contraction of $\mathrm{u}(2) = \mathsf{u}(1) \oplus \mathsf{su}(2)$, where $\mathsf{u}(1)$ provides the representation label j for $\mathsf{su}(2)$, to the oscillator algebra generated by $\{1, Q, P, H\}$.

The Kravchuk eigenfunctions satisfy $Q\Psi_n^j(q) = q\Psi_n^j(q)$ and $H\Psi_n^j(q) = n\Psi_n^j(q)$; hence the finite rotation generated by $\mathcal{F}^\alpha := e^{-\mathrm{i}\frac{1}{2}\pi\alpha H}$ (α mod 4), only multiplies them by phase,

$$\mathcal{F}^\alpha \Psi_n^j(q) = \exp\left(-\mathrm{i}\tfrac{1}{2}\pi n\alpha\right) \Psi_n^j(q), \tag{141}$$

and qualifies to be called the fractional *Fourier–Kravchuk* discrete transform.[26] This transform is realized by a matrix kernel that acts on the vector of discrete "image" values (Wolf & Krötzsch, 2007),

$$f(q) \mapsto \tilde{f}_\alpha(q) = (\mathcal{F}^\alpha f)(q) = \sum_{q'=-j}^{j} F_{q,q'}^j(\alpha) f(q'), \tag{142}$$

$$F_{q,q'}^j(\alpha) = \sum_{n=0}^{2j} \Psi_n^j(q)\, e^{-\mathrm{i}\frac{1}{2}\pi n\alpha}\, \Psi_n^j(q') = e^{-\mathrm{i}\frac{1}{2}\pi(q-q')}\, d_{q,q'}^j(\alpha). \tag{143}$$

Thus we have an $\mathsf{SO}(2)$ subgroup of rotations around the K axis with a phase built as a finite counterpart of the Namias expression (Namias, 1980) for α-fractional Fourier transforms. A rotation (141) by $\frac{1}{2}\pi$

[26] When the number of pixels $N = 2j + 1$ is odd, j is integer and we have a pixel at the center of the array. When N is even we are in the half-integer spin representations; the Fourier–Kravchuk transform (141) "corrects" the double spin range $\alpha \in [0, 4\pi)$ of $e^{-\mathrm{i}\frac{1}{2}\pi K\alpha}$ with the extra phase $e^{-\mathrm{i}\frac{1}{2}\pi j\alpha}$, as was the case in (110).

brings the Q axis of position onto the P axis of momentum, so the Fourier–Kravchuk transforms of $\Psi_n^j(q)$ are $\widetilde{\Psi}_n^j(p) = (-\mathrm{i})^n\, \Psi_n^j(p)$, as is familiar from quantum mechanics.[27]

9.4 2D Screens and $\mathsf{U}(2)_\mathrm{F}$ Transformations

Two-dimensional pixelated screens can be described basically as a Cartesian product of two one-dimensional ones. We now have two sets of generators (134) for $\mathsf{su}(2)_x \oplus \mathsf{su}(2)_y$, to name those in (132),

$$\begin{array}{lll} Q_x = J_1^x, & P_x = -J_2^x, & K_x = J_3^x, \\ Q_y = J_1^y, & P_y = -J_2^y, & K_y = J_3^y, \end{array} \tag{144}$$

satisfying (135) and mutually commuting—and $H_x = K_x + j_x\mathit{1}$, $H_y = K_y + j_y\mathit{1}$ that provide the pair of mode numbers n_x and n_y. The 2D Cartesian Kravchuk functions are (Atakishiyev, Pogosyan, Vicent, & Wolf, 2001b)

$$\begin{array}{l} \Psi_{n_x,n_y}^{(j_x,j_y)}(q_x,q_y) := \Psi_{n_x}^{(j_x)}(q_x)\Psi_{n_y}^{(j_y)}(q_y), \\ q_x|_{-j_x}^{j_x}, \quad n_x|_0^{2j_x}, \qquad q_y|_{-j_y}^{j_y}, \quad n_y|_0^{2j_y}. \end{array} \tag{145}$$

These N_xN_y Kravchuk functions can be arranged along axes of total mode $n := n_x + n_y$ and mode difference $m := n_x - n_y$ into the rhomboid pattern shown in Fig. 9. These modes are orthonormal and complete under the natural sesquilinear inner product on $\mathsf{C}^{N_xN_y}$. We shall consider first the general case of rectangular screens, with $j_x > j_y$; the special case $j_x = j = j_y$ will deserve some extra attention in the next subsection.

9.4.1 *Domestic Fourier–Kravchuk Transformations*

In two dimensions we have the Fourier group $\mathsf{U}(2)_\mathrm{F}$ generated by the Poisson operators of the classical functions $\ell_i|_{i=0}^3$ in (90)–(93). We evidently associate the isotropic ℓ_0 in (90), and anisotropic ℓ_1 in (91), to the fractional Fourier–Kravchuk transform seen in the last subsection. Note the factor $\frac{1}{4}$ in their expressions, and the factor $\frac{1}{2}$ in front of "physical" angular momentum in (93), which imply that we must take the *double* of the angle $\frac{1}{2}\pi\alpha$ in (141). Let $F_0 := \frac{1}{2}(H_x + H_y)$ and $F_1 := \frac{1}{2}(H_x - H_y)$; the isotropic and anisotropic fractional Fourier–Kravchuk transforms $\mathcal{F}_I(\chi) := \exp(-2\mathrm{i}\chi F_0)$ and $\mathcal{F}_A(\beta) := \exp(-2\mathrm{i}\beta F_1)$ are *domestic* to the discrete model and act on the Cartesian functions (145) as

[27] Note that the Fourier–Kravchuk transform is *not* "exactly" the discrete Fourier transform, but the form (143) indicates that, in the limit $j \to \infty$ referenced below, both converge to the Fourier integral transform kernel.

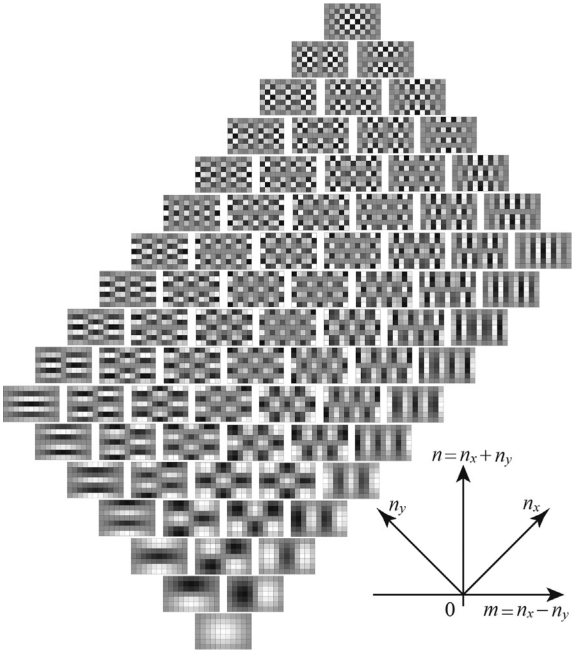

Fig. 9 The 11 × 7 rhomboid of Cartesian modes $\Psi^{(5,3)}_{n_x,n_y}(q_x,q_y)$ in (145), referred to axes $n_x|_0^{10}$ and $n_y|_0^6$ and also to axes n, m. In each screen, the pixels are numbered from the lower-left corner by $q_x|_{-5}^{5}$ and $q_y|_{-3}^{3}$. Gray-level densities are black and white for values from −1 to 1.

$$\mathcal{F}_I(\chi):\Psi^{(j_x,j_y)}_{n_x,n_y}(q_x,q_y)=\exp\left[-\mathrm{i}\chi(n_x+n_y)\right]\Psi^{(j_x,j_y)}_{n_x,n_y}(q_x,q_y), \tag{146}$$

$$\mathcal{F}_A(\beta):\Psi^{(j_x,j_y)}_{n_x,n_y}(q_x,q_y)=\exp\left[-\mathrm{i}\beta(n_x-n_y)\right]\Psi^{(j_x,j_y)}_{n_x,n_y}(q_x,q_y). \tag{147}$$

9.4.2 Imported Rotations

Next consider *rotations* of the pixelated image in a rectangular screen. We conjecture that if we use the well-known rotation coefficients—the Wigner $d^j_{\mu,\mu'}(2\theta)$'s with double angle—on the Cartesian Hermite–Gauss oscillator functions (Frank & van Isacker, 1994), and *import* (Barker, Çandan, Hakioğlu, Kutay, & Ozaktas, 2000) them to the discrete model, we should obtain a recognizable rotation of the pixelated image, which is real, and will be unitary (orthogonal) and thus reversible. Recall that the continuous quantum harmonic oscillator states (n_x, n_y) form an infinite pyramid with rungs $n_x+n_y=n|_0^\infty$ that are angular momentum multiplets of spin

$\lambda(n)=\frac{1}{2}n$ and z-projection $\frac{1}{2}(n_x-n_y)=\mu|_{-\lambda}^{\lambda}$. In the discrete model, however, we see in Fig. 9 that we have only the lowest part of that pyramid, in a rhombus where the spins $\lambda(n)$ and "z-projectons" μ,μ' are now constrained to the following ranges (for $j_x \geq j_y$):[28]

$$\begin{array}{lll} \text{lower triangle:} & & \\ \quad 0\leq n\leq 2j_y, & \lambda(n)=\frac{1}{2}n, & \mu=\frac{1}{2}(n_x-n_y); \\ \text{mid}-\text{rhomboid:} & & \\ \quad 2j_y<n<2j_x, & \lambda(n)=j_y, & \mu=j_y-n_y; \\ \text{upper triangle:} & & \\ \quad 2j_x\leq n\leq 2(j_x+j_y), & \lambda(n)=j_x+j_y-\frac{1}{2}n, & \mu=\frac{1}{2}(n_x-n_y)-j_x+n_y. \end{array} \tag{148}$$

Thus we posit that rotations $\mathcal{R}(\theta)$ of the discrete modes (145) are

$$\mathcal{R}(\theta):\Psi_{n_x,n_y}^{(j_x,j_y)}(q_x,q_y):=\sum_{n_x'+n_y'=n} d_{\mu,\mu'}^{\lambda(n)}(2\theta)\,\Psi_{n_x',n_y'}^{(j_x,j_y)}(q_x,q_y), \tag{149}$$

where μ,μ' are given in terms of n_x,n_y (unprimed and primed) by (148).

In Fig. 10 we rotate a 41×25 pixelated image of the letter "$\mathcal{B}$", white on black (1's on 0's), through six successive rotations by $\theta=\frac{1}{6}\pi$. We note that oscillations with small negative values appear in all intermediate positions; this type of Gibbs oscillation is common to all signal reconstruction algorithms with Fourier series when the signals have sharp edges. For $\theta=\frac{1}{2}\pi$ the image suffers expansion in x and compression in y with the concomitant oscillations, but for $\theta=\pi$ we recover the same, exact (inverted) image, as through a permutation of pixels. On square screens (Vicent & Wolf,2008, 2011), rotations by $\theta=\frac{1}{2}\pi$ are also permutations. This would be impossible

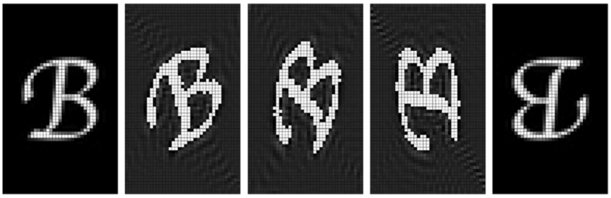

Fig. 10 Image of the letter "$\mathcal{B}$" on a 41×25 pixelated screen, $(j_x, j_y)=(20, 12)$, under successive rotations by $\frac{1}{6}\pi$ of the *left*, to $\frac{1}{3}\pi$, $\frac{1}{2}\pi$, and (extreme *right*) π. In the rotated images, the gray-level scale is rescaled so that the pixel values lie between 0 and 1.

[28] The ranges of n do overlap at $2j_y$ and at $2j_x$; we ascribe these to the triangles.

with any successively applied interpolation algorithm, since they inevitably loose information. Although the rotations in the present discrete model of rectangular screens are unitary, they also embody the *longest* computation algorithm, because each pixel on the transformed screen is a linear combination of *all* pixels in the original image.

9.4.3 Gyrations

Finally, consider the Fourier subgroup of *gyrations* $\mathcal{G}(\gamma)$ generated by the classical quadratic function ℓ_2 in (92) that generates rotations around the 2-axis of a mathematical "Fourier sphere." But since we already have domestic $\mathcal{F}_A(\beta)$ transformations (147) around the 1-axis, and the imported rotations $\mathcal{R}(\theta)$ around its 3-axis, remembering the double angle issue, we can write

$$\mathcal{G}(\gamma) := \mathcal{F}_A\left(\tfrac{1}{4}\pi\right) \mathcal{R}(\gamma)\, \mathcal{F}_A\left(-\tfrac{1}{4}\pi\right). \tag{150}$$

The action of gyrations on the basis of Cartesian Kravchuk modes is then

$$\begin{aligned} \mathcal{G}(\gamma) : \Psi^{(j_x, j_y)}_{n_x, n_y}(q_x, q_y) \\ := e^{-i\pi(n_x - n_y)/4} \sum_{n'_x + n'_y = n} d^{\lambda(n)}_{\mu,\mu'}(2\gamma)\, e^{+i\pi(n'_x - n'_y)/4}\, \Psi^{(j_x, j_y)}_{n'_x, n'_y}(q_x, q_y). \end{aligned} \tag{151}$$

where $\lambda(n)$, μ and μ' are again determined by j_x, n_x, j_y, n_y as in (148).

9.4.4 Laguerre–Kravchuk Modes

In wave models, fractional gyrations transform continuously the Hermite–Gauss beams from $\gamma = 0$, into Laguerre–Gauss beams for $\gamma = \frac{1}{4}\pi$ (Alieva, Bastiaans, & Calvo, 2005; Rodrigo, Alieva, & Bastiaans, 2011; Rodrigo, Alieva, & Calvo, 2007). In the present discrete model we show in Fig. 11 gyrations of $\Psi^{(j_x, j_y)}_{n_x, n_y}(q_x, q_y)$ for the quintuplet of $\lambda = 2$ states $n = 4$, noting that after a $\frac{1}{4}\pi$ we indeed obtain a credible discrete analogue of Laguerre–Gauss beams which, for lack of another name, we may call (*rectangular*) *Laguerre–Kravchuk* states,

$$\Lambda^{(j_x, j_y)}_{n,m}(q_x, q_y) := e^{-i\pi(n_x - n_y)/4} \sum_{n'_x + n'_y = n} d^{\lambda(n)}_{\mu,\mu'}\left(\tfrac{1}{2}\pi\right) e^{+i\pi(n'_x - n'_y)/4}\, \Psi^{(j_x, j_y)}_{n'_x, n'_y}(q_x, q_y). \tag{152}$$

Since these are complex functions, in Fig. 11 for $\gamma = \frac{1}{4}\pi$ we show separately the absolute values and phases. The chosen multiplet lies in the lower triangle of (148); the upper triangle yields the same absolute values with a

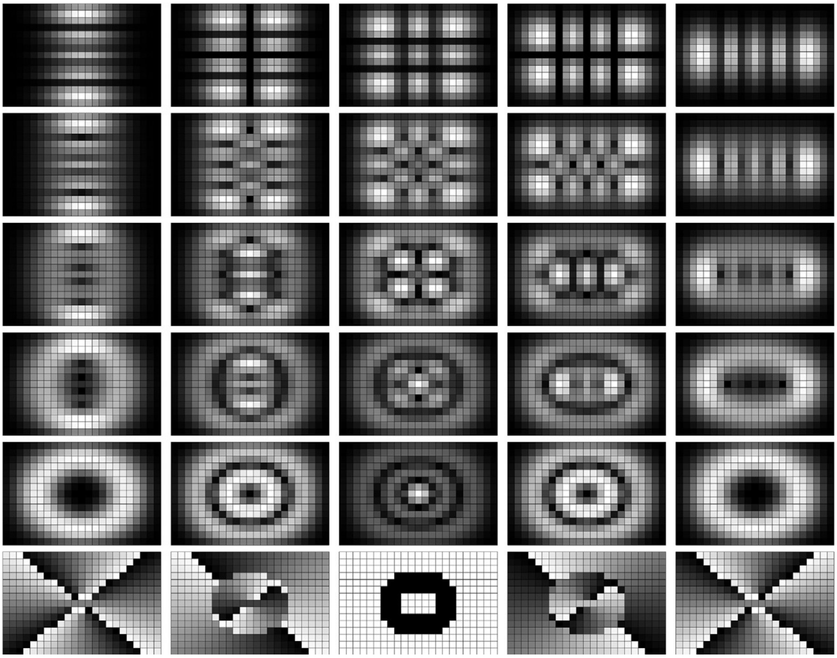

Fig. 11 Gyrations of the quintuplet $\lambda = 2$ ($m = -2, -1, 0, 1, 2$) of Cartesian Kravchuk modes $n = 4$ on 11×7 pixelated screens ($j_x = 5, j_y = 3$), through angles of $\gamma = 0, \frac{1}{16}\pi, \frac{1}{8}\pi, \frac{3}{16}\pi$, and $\frac{1}{4}\pi$ in the last two *lines*, where we show the absolute values (since the image values are complex), and the phase of the $\frac{1}{4}\pi$ gyration. The latter are the Laguerre–Kravchuk states of "rectangular angular momentum."

checkerboard of $e^{i\pi}$ differences in phase between neighbor pixels; multiplets in the mid-rhomboid follow a similar pattern. The functions (152) are also orthogonal and complete in $\mathsf{C}^{N_x N_y}$ so they can serve as an alternate basis for images; they transform under rotations by phases, with "angular momentum" number $m = 2\mu = n_x - n_y$, $|\mu| \leq \lambda(n)$, constrained by $n := n_x + n_y$ through (148).

9.5 Square and Circular Pixelated Screens

We can profit from a subalgebra chain of $\mathsf{so}(4)$ distinct from (131), namely the natural Gel'fand-Zetlin-type chain (Wong, 1967),

$$\mathsf{so}(4) \supset \mathsf{so}(3) \supset \mathsf{so}(2), \tag{153}$$

particularly when the screen is an $N \times N$ *square*, with $N = 2j+1$. The generators $\{\Lambda_{i,j}\}_{1=i<j}^{4}$, of $\mathsf{so}(4)$, with their commutation relations (129), will be

now renamed with *new* position and momentum operators identified by a circle superscript,

$$\begin{aligned} Q_x^\circ &:= \Lambda_{2,3} = Q_x + Q_y, & P_x^\circ &:= -\Lambda_{1,3} = P_x + P_y, \\ Q_y^\circ &:= \Lambda_{2,4} = P_x - P_y, & P_y^\circ &:= -\Lambda_{2,3} = Q_x - Q_y, \\ M^\circ &:= \Lambda_{3,4} = K_x - K_y, & K^\circ &:= \Lambda_{1,2} = K_x + K_y, \end{aligned} \tag{154}$$

where we note that M° and K° commute and thus can be both diagonal. From (129) we can see that M° generates rotations between the x- and y-components of Q° and of P° as angular momentum operators do, while K° generates rotations between Q_i° and the corresponding P_i° as the fractional Fourier transform does. On the other hand, note that the two "components of position," Q_x° and Q_y° do *not* commute, and neither do the two P°'s.

The subalgebra $\mathsf{so}(3)$ in (153) that we choose, which will yield the position of pixels on a screen, contains the generators Q_x°, Q_y°, and M°; call this subalgebra $\mathsf{so}(3)_\mathrm{Q}$. Its Casimir invariant operator will have the usual distribution of eigenvalues

$$R_\circ^2 := (Q_x^\circ)^2 + (Q_y^\circ)^2 + (M^\circ)^2 = \rho(\rho+1)\mathit{1}, \quad \rho|_0^{2j}, \tag{155}$$

and the basis elements are classified by the eigenvalues of $M^\circ = m\ \mathit{1}$, $m|_{-\rho}^{\rho}$. The operator $H^\circ := K^\circ + 2j\mathit{1}$ commutes with the generators of $\mathsf{so}(3)_\mathrm{Q}$ and is also diagonal with eigenvalues $n_x + n_y = n|_0^{4j}$ in the rhombus (148) with the mid-rhomboid now absent, as seen in Fig. 12.

Using the generators $\vec{J}^x$ and $\vec{J}^y$ in (132), the Cartesian modes $\Psi^{(j,j)}_{n_x,n_y} \equiv \Psi^j_{n,m}$ written now with indices n, m,[29] are

$$\begin{array}{lcccc} \text{eigenbasis of} & (\vec{J}^x)^2, & (\vec{J}^y)^2, & J_3^x, & J_3^y, \\ \text{with eigenvalues} & j(j+1), & j(j+1), & \tfrac12(n+m)-j, & \tfrac12(n-m)-j. \end{array} \tag{156}$$

We define states $Q^j_{\rho,m}$ related to the "position" $\mathsf{so}(3)_\mathrm{Q}$ subalgebra in (155), as

$$\begin{array}{lcccc} \text{eigenbasis of} & (\vec{J}^x+\vec{J}^y)^2, & \vec{J}^x\cdot\vec{J}^y, & R_\circ^2, & M^\circ, \\ \text{with eigenvalues} & 2j(j+1), & 0, & \rho(\rho+1), & m. \end{array} \tag{157}$$

Both $\{\Psi^j_{n,m}\}$ and $\{Q^j_{\rho,m}\}$ are orthogonal bases of $N^2 = (2j+1)^2$ states. Their *overlap* is clearly a *coupling* of two $\mathsf{so}(3)$ representations j to a third ρ, and thus given by Clebsch–Gordan coefficients $C^{j_1,\ j_2,\ j_3}_{m_1,m_2,m_3}$. Wigner's definition of

[29] Recall that $n = n_x + n_y$ and $m = n_x - n_y$, so $n_x = \frac12(n+m)$ and $n_y = \frac12(n-m)$.

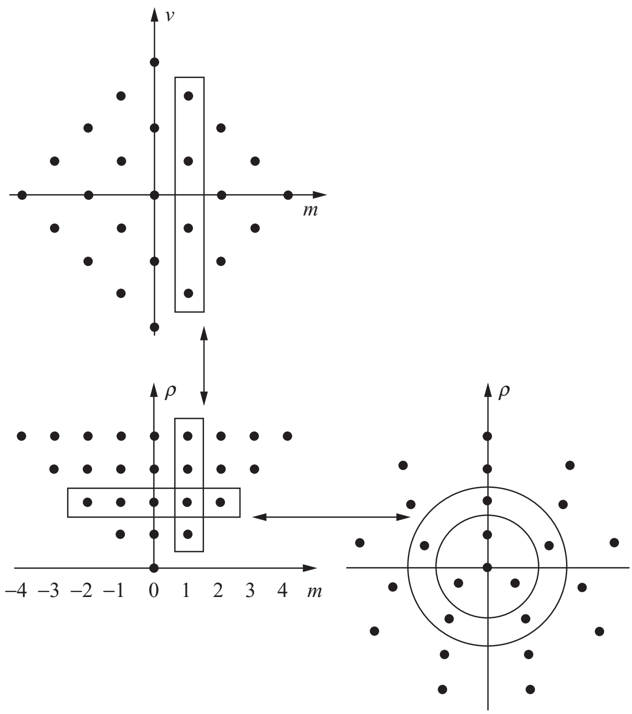

Fig. 12 *Top*: Cartesian states $\Psi^{(j,j)}_{n_x,n_y} \equiv \Psi^j_{n,m}$ (indicated by *dots*) in a symmetric so(4) multiplet on a square screen. *Left*: Cartesian states with the same $m = n_x - n_y$ are linearly combined (indicated by the *thin rectangles*) into "polar" states $\Phi^j_{\rho,m}$ with the same m, and belonging to so(3) multiplets $\rho|^{2j}_{|m|}$. *Right*: The finite Fourier transform maps the polar states $\Phi^j_{\rho,m}$, $m|^{\rho}_{-\rho}$ on a screen where the pixels (represented by *dots*) are on circles of radii $\rho|^{2j}_0$ and, on each *circle*, distributed by $2\rho+1$ angles θ_k, $k|^{\rho}_{-\rho}$.

these coefficients (Biedenharn & Louck, 1981), however, involves the subalgebra $\{\Lambda_{i,j}\}^3_{1\le i<j}$, while our $\mathrm{so}(3)_{\mathrm{Q}}$ is generated by $\{\Lambda_{i,j}\}^4_{2\le i<j}$. A rotation is necessary, which introduces a phase and sign reversal of the J^{γ}_3 eigenvalue, yielding the overlap (Atakishiyev, Pogosyan, Vicent, & Wolf, 2001a; Vicent & Wolf, 2008),

$$R^j_{n,m}(\rho) := (Q^j_{\rho,m}, \Psi^j_{n,m}) = \varphi^{j,n}_{\rho,m}\, C^{j,\qquad\quad j,\qquad\quad \rho}_{\frac{1}{2}(m+n)-j,\ \frac{1}{2}(m-n)+j,\ m} \tag{158}$$

$$\varphi^{j,n}_{\rho,m} := (-1)^{j+\rho+\frac{1}{2}(|m|-m)} e^{\mathrm{i}\frac{\pi}{2}n}, \tag{159}$$

which include the restrictions $0 \le \rho \le 2j$ and $|m| \le \rho$. We can regard $R^j_{n,m}(\rho)$ as a function of a *radius* ρ, on whose circle we have $2\rho+1$ pixels at equidistant angles ϕ_k, $k|^{\rho}_{-\rho}$, as shown in Fig. 12 (right).

Finally, we build the discrete basis of wavefields[30]

$$\Phi^j_{n,m}(\rho,\phi_k) := R^j_{n,m}(\rho)\,\frac{\exp\left(im\phi_k\right)}{\sqrt{2\rho+1}}, \quad \phi_k = \frac{2\pi k}{2\rho+1}. \tag{160}$$

These are shown in Fig. 13. By construction, they are orthonormal and complete under inner products over modes and angular momenta n, m and over *positions* of radius and angle ρ, ϕ_k on a polar-pixelated screen.[31] This pattern of pixels comes closest to contain pixels of equal size, except a bit near the origin $\rho = 0$. Another orthonormal and complete basis for modes and angular momenta n, m, are the Laguerre–Kravchuk states (152), $\Lambda^{(j,j)}_{n,m}(q_x, q_y)$ on the square screen Cartesian coordinates q_x,q_y.

We can thus transform between images f and f° on the Cartesian and polar screens through

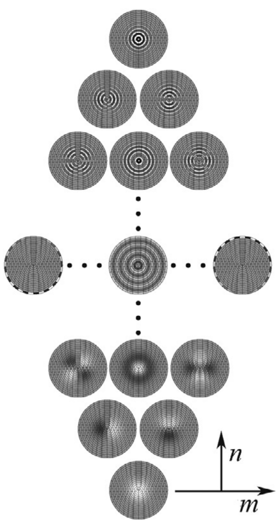

Fig. 13 The rhombus n, m of Laguerre–Kravchuk states $\Phi^j_{n,m}(\rho, \phi_k)$ in (160) for $j = 32$, on the circular pixelated screen for radii $\rho|_0^{64}$ and $2\rho+1$ angles $\phi_k = 2\pi k/(2\rho+1)$. The modes are complex, $\Phi^j_{n,-m}(\rho, \phi_k) = \Phi^j_{n,m}(\rho, \phi_k)^*$, so their real parts are shown on the right-hand side $m > 0$, and their imaginary part on the left-hand side $m < 0$; the $m = 0$ modes are real.

[30] Please note that there is an error in Eq. (34) of Vicent and Wolf (2008).

[31] One can add fixed angles θ_ρ to the ϕ_k's on each ρ-circle in the definition (160). This will only shift the starting angle on each circle. We have found it not inconvenient to let $\theta_\rho = 0$, even if this results in one radial line of $2j$ aligned pixels.

$$f(q_x,q_y) = \sum_{n,m} U^j(\rho,\phi_k\,;q_x,q_y)^* f^\circ(\rho,\phi_k),$$
$$f^\circ(\rho,\phi_k) = \sum_{q_x,q_y} U^j(\rho,\phi_k;q_x,q_y) f(q_x,q_y), \tag{161}$$

with a kernel that is the sum over the modes and angular momenta in the two bases,

$$U^j(\rho,\phi_k;q_x,q_y) := \sum_{n,m} \Phi^j_{n,m}(\rho,\phi_k)\, \Lambda^{(j,j)}_{n,m}(q_x,q_y)^*. \tag{162}$$

Since $\Phi^j_{n,m}(\rho,\phi_k) = \Phi^j_{n,-m}(\rho,\phi_k)^*$ and $\Lambda^{(j,j)}_{n,m}(q_x,q_y)^* = \Lambda^{(j_x,j_y)}_{n,-m}(q_x,q_y)$, this kernel is real. In Fig. 14 we show an image (a letter "R") mapped between a Cartesian and a polar screen. This map can be seen as the discrete analogue of *separation of variables* between continuous Cartesian and polar coordinates.

At this point the reader may rightfully suspect that images on *rectangular* pixelated screens can also be mapped faithfully on some other screen geometry. This has been tried by simply noting that the Clebsch–Gordan coefficients in (158) would now couple $j_x > j_y$ to radii $\rho|^{j_x+j_y}_{j_x-j_y}$, forming an *annular* screen, with the same angles $\phi_k|^{\rho}_{-\rho}$ as in (160). This could be perhaps useful in Newtonian telescopes. The result, however, (Urzúa & Wolf, 2016) shows the images on the annular screen to be very distorted even for small $j_x - j_y$ and hardly recognizable for larger values. We also considered possible *elliptic* screens, but the problem of distributing equally sized pixels along elliptic coordinates given by operator eigenvalues is challenging.

9.6 Aberrations of 1D Finite Discrete Signals

Transformations can be applied to finite data sets that will correspond to the geometric optical aberrations introduced in Section 8. Their study has been

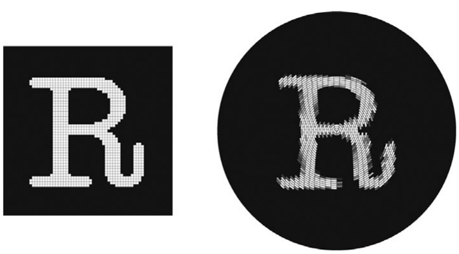

Fig. 14 The image "R" on a 32 × 32 Cartesian pixelated screen (valued 0 and 1), unitarily transformed onto a circular screen of 32^2 pixels.

concentrated on 1D discrete and finite signals on their *phase space* by means of a Wigner function on sets of $N = 2j+1$ points. This Wigner function is based on the algebra $\mathsf{su}(2) = \mathsf{so}(3)$ that will be given in Appendix A. Since there are N points, the acting matrices have to be $N \times N$, and thus there cannot be more than N^2 aberrations, which may be embedded in the Lie *unitary* group $\mathsf{U}(\mathrm{N})$ that will contain in particular all $N!$ pixel permutations.

To build the N^2 generators of 1D aberrations in an orderly fashion, we use again the monomials of classical phase space variables $M_{k,m}(p, q) := p^{k+m} q^{k-m}$ in (111), of rank k and weight m, and replace position q and momentum p with the $N \times N$ Hermitian and traceless matrices of the $\mathsf{su}(2)$ spin j representations (Biedenharn & Louck, 1981) according to (134), with diagonal

$$\begin{aligned} &\text{position}: q \mapsto \mathbf{Q} = \|Q_{q,q'}\|, \\ &\qquad Q_{q,q'} = q\,\delta_{q,q'}, \quad q|_{-j}^{j}, \end{aligned} \tag{163}$$

$$\begin{aligned} &\text{momentum}: p \mapsto \mathbf{P} = \|P_{q,q'}\|, \\ &P_{q,q'} = -\mathrm{i}\tfrac{1}{2}\sqrt{(j-q)(j+q+1)}\,\delta_{q+1,q'} + \mathrm{i}\tfrac{1}{2}\sqrt{(j+q)(j-q+1)}\,\delta_{q-1,q'}, \end{aligned} \tag{164}$$

$$\begin{aligned} &\text{mode} - j: \quad \mathbf{K} = \|K_{q,q'}\|, \\ &\quad K_{q,q'} = \tfrac{1}{2}\sqrt{(j-q)(j+q+1)}\,\delta_{q+1,q'} + \tfrac{1}{2}\sqrt{(j+q)(j-q+1)}\,\delta_{q-1,q'}, \end{aligned} \tag{165}$$

that satisfy

$$\mathbf{Q}^2 + \mathbf{P}^2 + \mathbf{K}^2 = j(j+1)\,\mathbf{1}. \tag{166}$$

Then we build a Hermitian product matrix out of the *three* matrices $\mathbf{Q}$, $\mathbf{P}$, $\mathbf{K}$ through their *Weyl-order product*. When there are n of these matrices, we sum over all $n!$ permutations of the factor operators $\mathbf{Q}$, $\mathbf{P}$, and $\mathbf{K}$, and divide by $n!$,

$$\{\mathbf{Q}^a, \mathbf{P}^b, \mathbf{K}^c\}_{\text{Weyl}} := \frac{1}{(a+b+c)!} \sum_{\text{permutations}} \overbrace{\mathbf{Q}\cdots\mathbf{Q}}^{a \text{ factors}}\ \overbrace{\mathbf{P}\cdots\mathbf{P}}^{b \text{ factors}}\ \overbrace{\mathbf{K}\cdots\mathbf{K}}^{c \text{ factors}}. \tag{167}$$

These will exponentiate to $N \times N$ unitary matrices, ensuring their reversibility and conservation of information. Eq. (166) is a restriction that we can choose to limit the powers of $\mathbf{K}$ to 0 or 1. In comparison with their classification in geometric optics, we thus have *two* pyramids of finite aberrations for each rank k

$$\mathbf{M}^0_{k,m} := \{\mathbf{P}^{k+m}, \mathbf{Q}^{k-m}\}_{\text{Weyl}}, \qquad m|_{-k}^{k}, \tag{168}$$

$$\mathbf{M}^1_{k,m} := \{\mathbf{P}^{k-\frac{1}{2}+m}, \mathbf{Q}^{k-\frac{1}{2}-m}, \mathbf{K}\}_{\text{Weyl}}, \quad m|_{-k+\frac{1}{2}}^{k-\frac{1}{2}} \tag{169}$$

for integer $0 \le 2k \le 2j$, i.e., aberration orders $0 \le A = 2k-1 \le N-1$.

Through symbolic computation and numerical evaluation, these matrices have been exponentiated and applied on 1D signals, shown in Figs. 15 and 16. There we render the action single aberrations (168) and (169) on the signal and on the phase space of finite systems, determined by the Wigner function given in Appendix A, to be compared with the classical deformations of phase space in Fig. 5. The signal exposed to aberrations is a rectangle function (top of Fig. 15); overall phases are generated by $\mathbf{M}^0_{0,0} = \mathbf{1}$, followed by the $\mathsf{SU}(2)$-linear transformations generated by $\mathbf{M}^0_{1/2,-1/2} = \mathbf{Q}$ and $\mathbf{M}^0_{1/2,1/2} = \mathbf{P}$ in the first pyramid. The second pyramid in Fig. 16 has on top $\mathbf{M}^1_{1/2,0} = \mathbf{K}$, corresponding to an oscillator Hamiltonian generating rotations of phase space. In the next rung, $\mathbf{M}^0_{1,m}$ are the finite counterparts of the linear canonical transformations of geometric optics, allowing for the deformation inherent in mapping the surface of the spheres on rectangles. The exponentiated aberration matrices (168) and (169) can be composed as in the geometric factored-product parametrization (119). In Rueda-Paz and Wolf (2011) we used this decomposition to simulate the aberrations of a 1D signal in a quasi-harmonic planar waveguide whose refractive index profile is

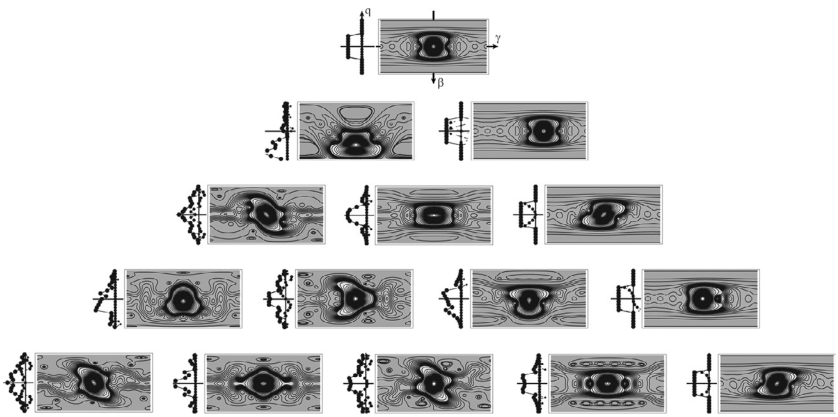

Fig. 15 Pyramid of aberrations (168) of order up to $A = 3$ on a 1D "rectangle" signal of $N = 21$ points. The signal is shown to the *left* of each subfigure, and its phase space Wigner function to the *right*. At the top is the original rectangle signal and its Wigner function on the flattened polar coordinates $0 \le \beta \le \pi$ and $-\pi < \gamma \le \pi$ of the sphere. The second row has the two SO(3)-"translations" in position and momentum (rank $\frac{1}{2}$, $m = \frac{1}{2}, -\frac{1}{2}$, aberration order 0). The following row of three aberrations corresponds to the *linear* transformations in continuous signals (free flight, squeezing, and lens; rank $k = 1$, $m = 1, 0, -1$, aberration order $A = 1$). There follow aberrations of orders 2 and 3. To highlight the behavior of the Wigner function near zero, the contour lines are chosen at $\{0, \pm 0.0001, \pm 0.001, \pm 0.01, 0.02, 0.03, \ldots, 0.15, 0.2, 0.3, \ldots, 3.0, 3.1\}$.

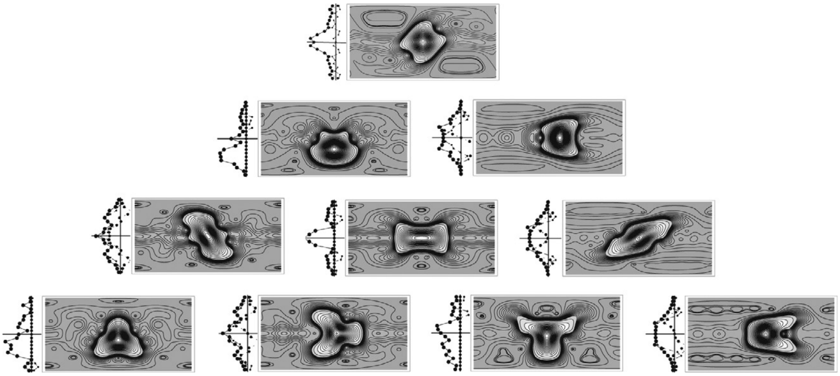

Fig. 16 Second pyramid of aberrations (169) of orders up to $A = 3$ on the same $N = 21$-point rectangle signal of the previous figure, with the same contour lines, axes, and values. At top, a 45 degree SO(3)-linear rotation generated by **K** (rank $k = \frac{1}{2}$, $m = 0$, aberration order 0), and its Wigner function. In the following rows, ranks $k = 1, \frac{3}{2}, 2$ (aberrations orders 1, 2, and 3); these are "**K**-repeaters" of aberrations of orders 0, 1, and 2 in the previous figure.

$n(q) \sim n_o q^2 + n_1 q^4$, along the z-axis of evolution. In this treatment of aberrations, we can pass directly from the geometric to the discrete model of optics.

ACKNOWLEDGMENTS

I thank the many colleagues who participated with me in the quest to find the symmetry that underlies optical models and ask leniency from those I have not mentioned. For the elaboration of the figures I acknowledge the indispensable help of Guillermo Krötzsch and of several students who worked out cases and examples. Support for the Óptica Matemática endeavors was provided by PAPIIT–DGAPA (Universidad Nacional Autónoma de México) funds along several years, the latest being project number 101115.

APPENDIX A. THE SU(2) WIGNER FUNCTION

We understand the Wigner function as the matrix elements of a Wigner *operator* that is an element of a group *ring*.[32] The Wigner operator may be defined broadly as the Fourier transform of a corresponding group. The Wigner function does not contain *more* information than the signal itself

[32] The group *ring* is the group G with the extra operation of *linear combination*. Its elements are of the form $\mathcal{A} = \sum_{g_i \in G} a_i g_i$ for $g_i \in G$ a discrete group and $a_i \in \mathsf{C}$ or, if the group is a continuous Lie group of elements $g(\vec{x})$, then $\mathcal{A} = \int_{\vec{x} \in G} d\mu(\vec{x})\, A(\vec{x})\, g(\vec{x})$, with the invariant Haar measure $d\mu(\vec{x})$ on the group manifold and $A(\vec{x}) \in \mathcal{L}^2(G)$.

(up to a total phase), but displays it in *phase space* (q, p) in a form that is friendly to the educated eye, much as a musical score is more informative than the pressure-wave register of a recorded tune (Forbes, Mańko, Ozaktas, Simon, & Wolf, 2000); while the latter is meaningless to visual inspection, the former can be recognized when played on an instrument, hummed by a human voice, or simply recalled from memory.

In 1932 Wigner proposed the original quasiprobability distribution function that we can write in 1D, between two states $\phi(q)$ and $\psi(q)$ in $\mathcal{L}^2(\mathsf{R})$, with a constant $\lambda\!\!\!^{-}$ (or $\hbar$), as

$$W(\phi, \psi|\, q, p, \lambda\!\!\!^{-}) = \frac{1}{2\pi\lambda\!\!\!^{-}} \int_{-\infty}^{\infty} \mathrm{d}x\, \phi\big(q - \tfrac{1}{2}x\big)^{*}\, e^{-ixp/\lambda\!\!\!^{-}}\, \psi\,\big(q - \tfrac{1}{2}x\big). \qquad \text{(A.1)}$$

This function (Hillery, O'Connell, Scully, & Wigner, 1984; Lee, 1995) is sesquilinear, $W(\phi, \psi|\, q, p) = W(\psi, \phi|\, q, p)^*$; for $\phi = \psi$ it is real (although not quite strictly positive); it is covariant under the Heisenberg–Weyl translations and, uniquely, under linear canonical transformations (García-Calderón & Moshinsky, 1980). It also has marginals and overlaps that allow the formulation of a quantum theory of measurement. It was introduced to optical models by Adolf Lohmann in the groundbreaking articles (Bartelt, Brenner, & Lohmann, 1980; Brenner & Lohmann, 1982; Lohmann, 1980). For specific systems and phase space manifolds, several distinct "Wigner functions" have been built with these properties that will be commented on below.

A.1 The Wigner Operator

In Wolf (1996), Atakishiyev, Chumakov, and Wolf (1998), and Ali, Atakishiyev, Chumakov, and Wolf (2000) we proposed an operator belonging to the *ring* of a D-dimensional Lie group G, whose generators X_n $(n|_1^D)$ form its Lie algebra. We use the *polar* parametrization of the group,[33] which we indicate using square brackets as $g[\vec{x}] = \exp i(\sum_{n=1}^{D} x_n X_n)$ so that the group identity is $g[\vec{0}]$ and the inverse is $g[\vec{x}]^{-1} = g[-\vec{x}]$. These coordinates $\{x_n\}_{n=1}^{D}$ can be treated as a "vector" but only extend over the manifold G of the group, $\vec{x} \in G \subset \mathsf{R}^D$. Let $\vec{\xi}$ be a vector in the full real manifold R^D and write, with some generality,

[33] We assume that the group G is of *exponential* type, i.e., that all its elements can be reached with the polar parametrization. This holds for the Heisenberg–Weyl, rotation, euclidean and all compact groups, but *not* for the $\mathsf{Sp}(2D, \mathsf{R})$ groups (Wolf, 2004, sect. 12.2).

$$\mathcal{W}(\vec{\xi}) := \int_G \mathrm{d}g[\vec{x}]\, \exp(-\mathrm{i}\vec{x}\cdot\vec{\xi})\, g[\vec{x}] = \int_G \mathrm{d}g[\vec{x}]\, \exp\Big(\mathrm{i}\,\vec{x}\cdot(\vec{X}-\vec{\xi})\Big), \quad \text{(A.2)}$$

where $\mathrm{d}g[\vec{x}]$ is the invariant Haar measure.[34] If the generators X_n were *numbers*, the function (A.2) would be simply $(2\pi)^D \prod_{n=1}^{D} \delta(\xi_n - X_n)$; the fact that the X_n are *operators* that will be applied to functions of the group, of coset spaces, or finite representation multiplets of the group, is what makes this Wigner operator interesting. We shall understand the manifold $\vec{\xi} \in \mathsf{R}^D$ to be the *meta-phase space* associated to the group G, but with the ordinary Euclidean measure $\mathrm{d}^D\vec{\xi}$.

Assume we have a Hilbert space of complex functions $\phi(h), \psi(h) \in \mathcal{H}$, where $h \in H$ may be the group itself, a space of cosets, a representation multiplet, or any arena for *unitary* action $g : \phi(h)$ by $g[\vec{x}] \in \mathsf{G}$, so that $g^\dagger = g^{-1}$, and with invariant measure $\mathrm{d}h$. The matrix elements of $\mathcal{W}(\vec{\xi})$ between two such functions is their Wigner *function*,

$$W(\phi,\psi|\vec{\xi}) := \int_H \mathrm{d}h\, \phi^*(h)\, \mathcal{W}(\vec{\xi}) : \psi(h) \quad \text{(A.3)}$$

$$= \int_H \mathrm{d}h \int_G \mathrm{d}g\, \phi^*(h)\, e^{-\mathrm{i}\vec{x}\cdot\vec{\xi}}\, (g : \psi)(h) \quad \text{(A.4)}$$

$$= \int_G \mathrm{d}g \int_H \mathrm{d}h\, (g^{-1/2} : \phi^*)(h)\, e^{-\mathrm{i}\vec{x}\cdot\vec{\xi}}\, (g^{1/2} : \psi)(h) \quad \text{(A.5)}$$

This is the structure which, for the Heisenberg–Weyl group, yields (A.1) with its left- and right-half translations (Wolf, 1996). Note that only in the polar parametrization are the *square roots* of group elements well defined: $(g[\vec{x}])^{1/2} = g[\frac{1}{2}\vec{x}]$. When we are given *density matrices*[35] ρ instead of pure states ϕ, ψ, the Wigner function is defined through $W(\rho|\vec{\xi}) = \mathrm{trace}\,(\mathcal{W}(\vec{\xi})\,\rho)$.

The Wigner operator (A.2) presents the following properties in $\mathcal{H}$, corresponding to those of the original Wigner function (A.1). The operator is *self-adjoint*: $\mathcal{W}(\vec{\xi})^\dagger = \mathcal{W}(\vec{\xi})$. It is *covariant* under similarity transformations by $g \in \mathsf{G}$, namely $g^{-1}\, \mathcal{W}(\vec{\xi})\, g = \mathcal{W}(D^{\mathrm{ad}}(g)\vec{\xi})$, where $D^{\mathrm{ad}}(g)$ is the $D \times D$ adjoint matrix representation of $g \in \mathsf{G}$; this also holds even when the transformation is an *outer* automorphism of the algebra, as the Heisenberg–Weyl generators under linear canonical transformations. The product integral

[34] The group G is assumed to be *unimodular*, i.e., that its *right*- and *left*-invariant measures are the same; this holds for a wide class of groups, but is *not* the case for the two-parameter *affine* group of translations and dilatations relevant for wavelets. In Ali et al. (2000) the expression (A.2) is generalized for such groups.

[35] A sum of ket-bra's, or *ideal projectors* in the ring.

$\int_{\mathfrak{R}^D} d\vec{\xi}\,|\mathcal{W}(\vec{\xi})|^2 \propto 1$ is proportional to the unit operator due to the Dirac δ's produced in one of the integrations, $\int_{\mathrm{R}^D} \mathrm{d}^D\vec{x}\exp\left(\mathrm{i}(\vec{x}-\vec{x}')\right)$. And finally, the Wigner operator *commutes* with all Casimir invariants of the algebra, so we may use the unitary irreducible matrix representations of the operators.

A.2 The SU(2) Wigner Matrix

Consider now the $N\times N$ representation of $\mathsf{SU}(2)$ of spin j ($N=2j+1$), with generators $\{J_i\}_{i=1}^3$ as given in (134). The polar parametrization uses the unit axis of rotation coordinates on the sphere, $\vec{v}(\theta,\phi)=\vec{x}/|\vec{x}|$, and the length $\eta := |\vec{x}|$, which is the rotation angle. The group elements are then

$$g[\vec{x}] = \exp\left(\mathrm{i}\vec{x}\cdot\vec{J}\right) = \exp\left[\mathrm{i}\eta(v_1J_1+v_2J_2+v_3J_3)\right], \tag{A.6}$$

and the Haar measure for continuous $\eta|_{-2\pi}^{2\pi}$,[36] $\theta|_0^{\pi}$ and $\phi|_{-\pi}^{\pi}$, is

$$\mathrm{d}g[\vec{x}] = \tfrac{1}{2}\sin^2\tfrac{1}{2}\eta\,\mathrm{d}\eta\,\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\phi. \tag{A.7}$$

Now let the Wigner operator (A.2) act on column vectors $\mathbf{f}=\{f_m\}_{m=-j}^{j}$ whose components are the N values of a finite signal on a 1D array of points. This will define a "Wigner *matrix*" $\mathbf{W}^j(\vec{\xi}) = \|W^j_{m,m'}(\vec{\xi})\|$ that *represents* the Wigner operator $\mathcal{W}(\vec{\xi})$, $\vec{\xi}\in\mathsf{R}^3$, for spin j,

$$\mathcal{W}(\vec{\xi}):\mathbf{f} = \int_G \mathrm{d}g[\vec{x}]\,\exp\left(-\mathrm{i}\vec{x}\cdot\vec{\xi}\right)\mathbf{D}^j(g[\vec{x}])\,\mathbf{f} =: \mathbf{W}^j(\vec{\xi})\,\mathbf{f}, \tag{A.8}$$

where $\mathbf{D}^j(g[\vec{x}]) = \|D^j_{m,m'}(g[\eta,\theta,\phi])\|$ are the $\mathsf{SU}(2)$ rotation matrices (called Wigner Big-D matrices (Biedenharn & Louck, 1981)) in polar parameters. These Wigner matrices qualify to be the "Fourier transform" of the irreducible representation matrices of the group $\mathsf{SU}(2)$, because

$$\mathbf{W}^j(\vec{\xi}) = \int_{\mathsf{SU}(2)} \nu(\vec{x})\,\mathrm{d}^3\vec{x}\exp\left(-\mathrm{i}\vec{x}\cdot\vec{\xi}\right)\mathbf{D}^j(g[\vec{x}]), \tag{A.9}$$

$$\mathbf{D}^j(g[\vec{x}]) = \frac{1}{(2\pi)^3\nu(\vec{x})}\int_{\mathrm{R}^3} \mathrm{d}^3\vec{\xi}\exp\left(\mathrm{i}\vec{x}\cdot\vec{\xi}\right)\mathbf{W}^j(\vec{\xi}), \tag{A.10}$$

with the weight $\nu(\vec{x})$ in the Haar measure $\mathrm{d}g[\vec{x}]=\nu(\vec{x})\,\mathrm{d}^3\vec{x}$ that we find in (A.7) as $\nu(\vec{x})=\nu(\eta)=\frac{1}{8}\,\mathrm{sinc}\,\frac{1}{2}\eta$.

[36] This range ensures that the group integration will contain both an element $g(\eta,\theta,\phi)$ and its inverse with $-\eta$.

The $\mathsf{SU}(2)$ Wigner *function* of the N-point finite signal vectors $\mathbf{f}_1$ and $\mathbf{f}_2$, is then a vector-and-matrix product,

$$W^j(\mathbf{f}_1,\mathbf{f}_2|\vec{\xi}) := \mathbf{f}_1^\dagger \mathbf{W}^j(\vec{\xi})\mathbf{f}_2, \qquad W^j(\mathbf{f}|\vec{\xi}) := W^j(\mathbf{f},\mathbf{f}|\vec{\xi}). \tag{A.11}$$

To complete the task of finding the Wigner matrix elements $W^j_{m,m'}(\vec{\xi})$ from (A.9) we have the Big-D matrix elements, normally written in Euler angles for spin j, $D^j_{m,m'}(\alpha,\beta,\gamma) = e^{\mathrm{i}m\alpha}\, d^j_{m,m'}(\beta)\, e^{\mathrm{i}m'\gamma}$. Here we have *polar* angles, so we can use the Wigner little-d functions to write (Biedenharn & Louck, 1981)

$$D^j_{m,m'}[\eta\, v(\theta,\phi)] = e^{\mathrm{i}(m'-m)\phi} \sum_{m''=-j}^{j} d^j_{m,m''}(\theta)\, e^{-\mathrm{i}m''\eta}\, d^j_{m',m''}(\theta). \tag{A.12}$$

The integration over $\mathsf{SU}(2)$ in (A.9) can be rotated so that $\vec{v}(\theta,\phi)$ is the 3-axis unit vector $\mathbf{k}$, and where the Wigner matrix is diagonal, $W^j_{m,m'}(\eta\mathbf{k}) = \delta_{m,m'}\, W^j_m(\eta)$; for $W^j_m(\eta)$ we can similarly write

$$W^j_{m,m'}[\eta\, \vec{v}(\theta,\phi)] = e^{\mathrm{i}(m'-m)\phi} \sum_{m''=-j}^{j} d^j_{m,m''}(\theta)\, W^j_{m''}(\eta)\, d^j_{m',m''}(\theta). \tag{A.13}$$

This allows us to separate the sphere manifold θ,ϕ of known functions, from the η-dependent diagonal elements of the Wigner function $W^j_m(\eta)$, $m|_{-j}^{j}$, which are the *eigenvalues* of the Wigner matrix. Calculated with some care,[37] these are

$$\begin{aligned} W^j_m(\eta) = (-1)^{2j+1}\frac{\pi}{4}\sum_{n=-j}^{j}\int_{-1}^{1} \mathrm{d}s\, \left(d^j_{m,n}(\arccos s)\right)^2 \\ \times \sin(2\pi\eta s)\left[\frac{1}{\eta s - n + 1} - \frac{2}{\eta s - n} + \frac{1}{\eta s - n - 1}\right]. \end{aligned} \tag{A.14}$$

This expression, replaced in (A.13), gives the matrix elements of the Wigner matrix for spin j; its dependence on the radial coordinate η is shown to be strongly peaked between j and $j+1$ (Atakishiyev et al., 1998), so we may display the Wigner function of a signal $\mathbf{f}$ on the *surface* of a *sphere* in the $\vec{\xi}$-space R^3 at the radius $\eta = |\vec{\xi}| = j + \frac{1}{2}$.

According to (134) we may identify the three continuous coordinates $\{\xi_i\}_{i=1}^3 \in \mathsf{R}$ of the Wigner function as position $\xi_1 = q$, momentum $\xi_2 = -p$, and $\xi_3 = \mu = n - j$ for mode $n|_0^{2j}$, each in its real line. Low modes $n \approx 0$ (see Fig. 8) register around the bottom pole of the sphere $\xi_3 \approx -j$,

[37] The poles from the brackets are canceled by the zeros of the sine function.

while the highest modes $n \approx 2j$ at the top pole. Plotting functions on spheres θ, ϕ is awkward in flat figures, so in Figs. 15 and 16 we resorted to show the angles $\theta|_{\frac{1}{2}\pi}^{\frac{3}{2}\pi}$ and $\phi|_{-\pi}^{\pi}$ as if they were Cartesian coordinates, with the bottom pole of the sphere at the center.[38]

A.3 Closing Remarks

This appendix mostly pertained the $\mathsf{SU}(2)$ Wigner function of the form (A.2), with phase space being a sphere—classically a symplectic manifold. $\mathsf{SU}(2)$ was also used by Agarwal et al. (Agarwal, 1981; Agarwal, Puri, & Singh, 1997; Dowling, Agarwal, & Schleich, 1994) to define a Wigner function using tensorial notation which was *prima facie* quite distinct from our presentation. Actually, the two are equivalent, as shown with some labor in Chumakov, Klimov, and Wolf (2000).

The structure of (A.2) has been used for other group rings: on the compact circle of $\mathsf{SO}(2)$ phase space is the set of integer points, and the Wigner function is the "sinc" interpolation between the absolute squares of the analyzed function (Nieto, Atakishiyev, Chumakov, & Wolf, 1998). The 2D Euclidean group $\mathsf{ISO}(2)$ phase space can be also reduced from the 3D manifold as in the present case and is a cylinder (Nieto et al., 1998). As we said before, the Heisenberg–Weyl group leads to standard Wigner function (A.1), and the 1D affine group was studied in Ali et al. (2000) to place wavelets. Phase space representations are useful when they are two-dimensional; although marginals—projections on lower-dimensional spaces—hold for all models.

Distinct "Wigner-type" functions can be obtained from (A.2) if we introduce to the integrand a function $K(\chi[\vec{x}])$ over the manifold of *equivalence classes* of the group G, so $\chi(g_c) = \chi(g\, g_c\, g^{-1})$. This plays the role of the *Cohen function* (Cohen, 1966; Lee, 1995) that defines Q-, P-, or Husimi functions, among others.

Further models that also use the basic structure of the Wigner function, with interesting properties of their own, include solutions of the Helmholtz equation (Wolf, Alonso, & Forbes, 1999), which led to in-depth studies by Gregory Forbes and Miguel Angel Alonso on electromagnetic fields (Alonso, 2009, 2011, 2015), and which settled some controversies in radiometry. Still other works have dealt where the position coordinate lies on a sphere or hyperboloid (Alonso, Pogosyan, & Wolf,2002, 2003).

[38] Other maps could perhaps be better even if not as immediate.

REFERENCES

Agarwal, G. S. (1981). Relation between atomic coherent-state representation, state multipoles, and generalized phase-space distributions. *Physical Review A*, *24*, 2889–2896.

Agarwal, G. S., Puri, R. R., & Singh, R. P. (1997). Atomic schrödinger cat states. *Physical Review A*, *56*, 2249–2254.

Ali, S. T., Atakishiyev, N. M., Chumakov, S. M., & Wolf, K. B. (2000). The Wigner function for general Lie groups and the wavelet transform. *Annales Henri Poincaré*, *1*, 685–714.

Alieva, T., Bastiaans, M. J., & Calvo, M. L. (2005). Fractional transforms in optical information processing. *EURASIP Journal in Signal Processing*, *2005*, 1498–1519.

Alonso, M. A. (2009). Diffraction of paraxial partially coherent fields by planar obstacles in the Wigner representation. *Journal of the Optical Society of America A*, *26*, 1588–1597.

Alonso, M. A. (2011). Wigner functions in optics: Describing beams as ray bundles and pulses as particle ensembles. *Advances in Optics and Photonics*, *3*, 272–365.

Alonso, M. A. (2015). In *Mathematical optics group (Rochester) tutorial*. http://www.optics.rochester.edu/workgroups/alonso/Home.html

Alonso, M. A., Pogosyan, G. S., & Wolf, K. B. (2002). Wigner functions for curved spaces I: On hyperboloids. *Journal of Mathematical Physics*, *43*, 5857–5871.

Alonso, M. A., Pogosyan, G. S., & Wolf, K. B. (2003). Wigner functions for curved spaces II: On spheres. *Journal of Mathematical Physics*, *44*, 1472–1489.

Atakishiyev, N. M., Chumakov, S. M., & Wolf, K. B. (1998). Wigner distribution function for finite systems. *Journal of Mathematical Physics*, *39*, 6247–6261.

Atakishiyev, N. M., Pogosyan, G. S., Vicent, L. E., & Wolf, K. B. (2001a). Finite two-dimensional oscillator II: The radial model. *Journal of Physics A*, *34*, 9399–9415.

Atakishiyev, N. M., Pogosyan, G. S., Vicent, L. E., & Wolf, K. B. (2001b). Finite two-dimensional oscillator I: The Cartesian model. *Journal of Physics A*, *34*, 9381–9398.

Atakishiyev, N. M., Pogosyan, G. S., & Wolf, K. B. (2003). Contraction of the finite one-dimensional oscillator. *International Journal of Modern Physics A*, *18*, 317–327.

Atakishiyev, N. M., Pogosyan, G. S., & Wolf, K. B. (2005). Finite models of the oscillator. *Physics of Particles and Nuclei*, *36*(Suppl. 3), 521–555.

Atakishiyev, N. M., & Wolf, K. B. (1997). Fractional Fourier-Kravchuk transform. *Journal of the Optical Society of America A*, *14*, 1467–1477.

Atzema, E. J., Krötzsch, G., & Wolf, K. B. (1997). Canonical transformations to warped surfaces: Correction of aberrated optical images. *Journal of Physics A*, *30*, 5793–5803.

Bargmann, V. (1947). Irreducible unitary representations of the Lorentz group. *Annals of Mathematics*, *48*, 568–642.

Bargmann, V. (1970). Group representation in Hilbert spaces of analytic functions. In P. Gilbert & R. G. Newton (Eds.), *Analytical methods in mathematical physics* (pp. 27–63). New York: Gordon & Breach.

Barker, L., Çandan, Ç., Hakioğlu, T., Kutay, M. A., & Ozaktas, H. M. (2000). The discrete harmonic oscillator, Harper's equation, and the discrete fractional Fourier transform. *Journal of Physics A*, *33*, 2209–2222.

Bartelt, H. O., Brenner, K.-H., & Lohmann, A. W. (1980). The Wigner distribution function and its optical production. *Optics Communication*, *32*, 32–38.

Biedenharn, L. C., & Louck, J. D. (1981). Angular momentum in quantum physics. Theory and application. In G.-C. Rota (Ed.), *Encyclopædia of mathematics and its applications*. Reading, MA: Addison-Wesley.

Boyer, C. P., & Wolf, K. B. (1973). Deformation of inhomogeneous classical Lie algebras to the algebras of the linear groups. *Journal of Mathematical Physics*, *14*, 1853–1859.

Brenner, K.-H., & Lohmann, A. H. (1982). Wigner distribution function display of complex 1D signals. *Optics Communication*, *42*, 310–314.

Buchdahl, H. (1970). *An introduction to Hamiltonian optics*. Cambridge: Cambridge University Press.

Chumakov, S. M., Klimov, A. B., & Wolf, K. B. (2000). On the connection of two Wigner functions for spin systems. *Physical Review A*, *61*(3), 034101.

Cohen, L. (1966). Generalized phase-space distribution functions. *Journal of Mathematical Physics*, *7*, 781–786.

Collins, S. A., Jr. (1970). Lens-system diffraction integral written in terms of matrix optics. *Journal of the Optical Society of America*, *60*, 1168–1177.

Condon, E. U. (1937). Immersion of the Fourier transform in a continuous group of functional transformations. *Proceedings of the National Academy of Sciences*, *23*, 158–163.

Dowling, J. P., Agarwal, G. S., & Schleich, W. P. (1994). Wigner distribution of a general angular-momentum state: Applications to a collection of two-level atoms. *Physical Review A*, *49*, 4101–4109.

Dragt, A. J. (1982a). Lectures on nonlinear orbit dynamics. In R. A. Carrigan, F. R. Huson, & M. Month (Eds.), *AIP conference proceedings* (Vol. 87, No. 1, pp. 147–313). AIP.

Dragt, A. J. (1982b). Lie algebraic theory of geometrical optics and optical aberrations. *Journal of the Optical Society of America*, *72*, 372–379.

Dragt, A. J. (1987). Elementary and advanced Lie algebraic methods with applications to accelerator design, electron microscopes, and light optics. *Nuclear Instruments and Methods in Physics Research A*, *258*, 339–354.

Dragt, A. J. (2004). *Lie methods for nonlinear dynamics with applications to accelerator physics*. Maryland: University of Maryland. http://www.physics.umd.edu/dsat/dsatliemethods.html.

Forbes, G. W., Manko, V. I., Ozaktas, H. M., Simon, R., & Wolf, K. B. (Eds.), (2000). Feature issue on Wigner distributions and phase space in optics. *Journal of the Optical Society of America A*, *17*, 2274.

Frank, A., & van Isacker, P. (1994). *Algebraic methods in molecular and nuclear structure physics*. New York: Wiley.

García-Bullé, M., Lassner, W., & Wolf, K. B. (1986). The metaplectic group within the Heisenberg–Weyl ring. *Journal of Mathematical Physics*, *27*, 29–36.

García-Calderón, G., & Moshinsky, M. (1980). Wigner distribution functions and the representation of canonical transformations. *Journal of Physics A*, *13*, L185–L189.

Gilmore, R. (1978). *Lie groups, Lie algebras, and some of their applications*. New York: Wiley Interscience.

González-Casanova, P., & Wolf, K. B. (1995). Interpolation of solutions to the Helmholtz equation. *Numerical Methods of Partial Differential Equations*, *11*, 77–91.

Goodman, J. W. (1968). *Introduction to Fourier optics*. New York: McGraw-Hill.

Guillemin, V., & Sternberg, S. (1984). *Symplectic techniques in physics*. Cambridge: Cambridge University Press.

Healy, J. J., Alper Kutay, M., Ozaktas, H. M., & Sheridan, J. T. (Eds.). (2015). *Linear canonical transforms, theory and applications: Vol. 198. Springer Series in Optical Sciences*. New York: Springer.

Hillery, M., O'Connell, R. F., Scully, M. O., & Wigner, E. P. (1984). Distribution functions in physics: Fundamentals. *Physics Reports*, *106*, 121–167.

Kauderer, M. (1994). *Symplectic matrices. First order systems and special relativity*. Singapore: World Scientific.

Khan, S. A., & Wolf, K. B. (2002). Hamiltonian orbit structure of the set of paraxial optical systems. *Journal of the Optical Society of America A*, *19*, 2436–2444.

Krawtchouk, M. (1928). Sur une généralisation des pôlynomes d'hermite. *Comptes Rendus de l'Académie des Sciences Series I - Mathématique*, *189*, 620–622.

Lee, H.-W. (1995). Theory and application of the quantum phase-space distribution functions. *Physics Reports*, *259*, 147–211.
Liberman, S., & Wolf, K. B. (2015). Independent simultaneous discoveries visualized through network analysis: The case of linear canonical transforms. *Scientometrics*, *104*, 715–735.
Lie, S. (1888). *Theorie der transformationsgruppen* (Vol. 1). Leipzig: B. G. Teubner.
Lohmann, A. (1980). The Wigner function and its optical production. *Optics Communication*, *42*, 32–37.
Luneburg, R. K. (1964). *Mathematical theory of optics*. Berkeley: University of California.
McBride, A. C., & Kerr, F. H. (1987). On Namias's fractional Fourier transforms. *IMA Journal of Applied Mathematics*, *39*, 159–175.
Mendlovic, D., & Ozaktas, H. M. (1993). Fractional Fourier transforms and their optical implementation: I. *Journal of the Optical Society of America A*, *10*, 1875–1881.
Moshinsky, M., & Quesne, C. (1971). Linear canonical transformations and their unitary representation. *Journal of Mathematical Physics*, *12*, 1772–1780.
Namias, V. (1980). The fractional order Fourier transform and its applications in quantum mechanics. *IMA Journal of Applied Mathematics*, *25*, 241–265.
Navarro Saad, M., & Wolf, K. B. (1986). Factorization of the phase–space transformation produced by an arbitrary refracting surface. *Journal of the Optical Society of America A*, *3*, 340–346.
Nieto, L. M., Atakishiyev, N. M., Chumakov, S. M., & Wolf, K. B. (1998). Wigner distribution function for finite systems. *Journal of Physics A*, *31*, 3875–3895.
Ozaktas, H. M., & Mendlovic, D. (1993a). Fourier transforms of fractional order and their optical interpretation. *Optics Communication*, *101*, 163–169.
Ozaktas, H. M., & Mendlovic, D. (1993b). Fractional Fourier transforms and their optical implementation: II. *Journal of the Optical Society of America A*, *10*, 2522–2531.
Ozaktas, H. M., Zalevsky, Z., & Kutay, M. A. (2001). *The fractional Fourier transform with applications in optics and signal processing*. Chichester: Wiley.
Pei, S.-C., & Ding, J.-J. (2000). Closed-form discrete fractional and affine transforms. *IEEE Transactions on Signal Processing*, *48*, 1338–1353.
Pei, S.-C., & Yeh, M.-H. (1997). Improved discrete fractional transform. *Optics Letters*, *22*, 1047–1049.
Pei, S.-C., Yeh, M.-H., & Tseng, C.-C. (1999). Discrete fractional Fourier transform based on orthogonal projections. *IEEE Transactions on Signal Processing*, *47*, 1335–1348.
Quesne, C., & Moshinsky, M. (1971). Linear canonical transformations and matrix elements. *Journal of Mathematical Physics*, *12*, 1780–1783.
Rashed, R. (1990). A pioneer in anaclastics—Ibn Sahl on burning mirrors and lenses. *ISIS*, *81*, 464–491.
Rashed, R. (1993). *Géométrie et dioptrique au x^e siècle: Ibn Sahl, al-qūhī, et Ibn al-Hayatham, Collection Sciences et Philosophie Arabes, Textes et Études*. Paris: Les Belles Lettres.
Rodrigo, J. A., Alieva, T., & Bastiaans, M. J. (2011). Phase space rotators and their applications in optics. In G. Cristóbal, P. Schelkens, & H. Thienpont (Eds.), *Optical and digital image processing: Fundamentals and applications* (pp. 251–271). Weinheim: Wiley-VCH Verlag.
Rodrigo, J. A., Alieva, T., & Calvo, M. L. (2007). Gyrator transform: Properties and applications. *Optics Express*, *15*, 2190–2203.
Rueda-Paz, J., & Wolf, K. B. (2011). Finite signals in planar waveguides. *Journal of the Optical Society of America A*, *28*, 641–650.
Seidel, L. (1853). Zur dioptrik. *Astronomische Nachrichten*, *871*, 105–120.
Simon, R., & Wolf, K. B. (2000). Fractional Fourier transforms in two dimensions. *Journal of the Optical Society of America A*, *17*, 2368–2381.

Steinberg, S. (1986). Lie series, Lie transformations, and their applications. In J. Sánchez-Mondragón & K. B. Wolf (Eds.), *Lie methods in optics: Vol. 250. Lecture notes in physics*. Heidelberg: Springer-Verlag

Steinberg, S., & Wolf, K. B. (1981). Invariant inner products on spaces of solutions of the Klein-Gordon and Helmholtz equations. *Journal of Mathematical Physics*, *22*, 1660–1663.

Sudarshan, E. C. G., & Mukunda, N. (1974). *Classical dynamics: a modern perspective*. New York: Wiley.

Sudarshan, E. C. G., Mukunda, N., & Simon, R. (1985). Realization of first order optical systems using thin lenses. *Optica Acta*, *32*, 855–872.

Urzúa, A. R., & Wolf, K. B. (2016). Unitary rotation and gyration of pixellated images on rectangular screens. *Journal of the Optical Society of America A*, *33*, 642–647.

Vicent, L. E., & Wolf, K. B. (2008). Unitary transformation between Cartesian- and polar-pixellated screens. *Journal of the Optical Society of America A*, *25*, 1875–1884.

Vicent, L. E., & Wolf, K. B. (2011). Analysis of digital images into energy-angular momentum modes. *Journal of the Optical Society of America A*, *28*, 808–814.

Wigner, E. (1932). On the quantum correction for thermodynamic equilibrium. *Physics Review*, *40*, 749–759.

Wolf, K. B. (1974). Canonical transforms I. Complex linear transforms. *Journal of Mathematical Physics*, *15*, 1295–1301.

Wolf, K. B. (1975). The Heisenberg-Weyl ring in quantum mechanics. In E. M. Loebl (Ed.), *Group theory and its applications* (Vol. 3, pp. 190–247). New York: Academic Press.

Wolf, K. B. (1979). *Integral transforms in science and engineering*. New York: Plenum Publishing Corporation.

Wolf, K. B. (1989). Elements of euclidean optics. In K. B. Wolf (Ed.), *Lie methods in optics, II workshop: Vol. 352. Lecture notes in physics* (pp. 115–162). Heidelberg: Springer-Verlag.

Wolf, K. B. (1996). Wigner distribution function for paraxial polychromatic optics. *Optics Communication*, *132*, 343–352.

Wolf, K. B. (2004). *Geometric optics on phase space*. Heidelberg: Springer-Verlag.

Wolf, K. B., Alonso, M. A., & Forbes, G. W. (1999). Wigner functions for Helmholtz wavefields. *Journal of the Optical Society of America A*, *16*, 2476–2487.

Wolf, K. B., & Boyer, C. P. (1974). The algebra and group deformations $I^m[SO(n)\otimes SO(m)] \Rightarrow SO(n,m)$, $I^m[U(n)\otimes U(m)] \Rightarrow U(n,m)$, and $I^m[Sp(n)\otimes Sp(m)] \Rightarrow Sp(n,m)$, for $1 \leq m \leq n$. *Journal of Mathematical Physics*, *15*, 2096–2101.

Wolf, K. B., & Krötzsch, G. (1995). `mexLIE 2`, *a set of symbolic computation functions for geometric aberration optics*. IIMAS–UNAM Manual No. 10. Mexico City.

Wolf, K. B., & Krötzsch, G. (2007). Geometry and dynamics in the fractional discrete Fourier transform. *Journal of the Optical Society of America A*, *24*, 651–658.

Wong, M. K. F. (1967). Representations of the orthogonal group. I. Lowering and raising operators of the orthogonal group and matrix elements of the generator. *Journal of Mathematical Physics*, *8*, 1899–1911.

Wybourne, B. G. (1974). *Classical groups for physicists*. New York: Wiley.