Relativistic aberration of optical phase space

Kurt Bernardo Wolf

Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas—Cuernavaca, Universidad Nacional Autónoma de México, Apartado Postal 139-B, 62191 Cuernavaca, Morelos, México

Received November 16, 1992; revised manuscript received March 5, 1993; accepted March 8, 1993

When an optical imaging system is boosted to relativistic velocities, the directions of the incoming light rays converge toward the pole of motion and the images are amplified in that direction and suffer a global comatic aberration. Our analysis uses succinctly the Hamilton–Lie formulation of optics to find the invariants in optical phase space of all Lorentz transformations.

1. INTRODUCTION

One favorite *Gedankenexperiment* in special relativity mounts an observer on a spaceship. Among other effects at higher velocities, photons will rain on him with increasingly forward-slanted trajectories so that the field of visible stars will shrink toward his direction of motion; the Doppler effect turns the light blue at the center and red at the periphery. Terrell[1] has studied the apparent images in a snapshot camera: differences in the time of photon flight will make three-dimensional objects near the observer's path appear to be rotated. Stereoscopic cameras with continuously open shutters, on the other hand, will produce shear in the images; these have been analyzed numerically and within approximations.[2] The controversy over the two interpretations has revived recently.[3]

The purpose of this paper is to discuss the relativistic aberrations of optical images from the point of view of geometric optics in phase space. Lorentz transformations of the sphere of ray directions entail a canonically conjugate transformation of the images on the screen of an imaging device. We find that when the direction of motion coincides with the optical axis, there is magnification and circular comatic aberration; when the axes do not coincide, the resulting spot diagrams present other interesting geometric properties.

In 1947 Bargmann[4] studied the representations of the two-dimensional relativistic group $SO(2,1)$ as ray transformations of the circle (with no reference to optics). Along the same lines, Boyer and the present author studied groups of deformations of N-dimensional spheres,[5] real, complex, and quaternionic.[6] Later work on the group-theoretical foundations of geometric and Helmholtz optics[7] found the natural action of the Lorentz group on spaces of optical rays and wave fronts. In Refs. 8 and 9 the Bargmann deformation algorithm was applied to geometric and wave-optical coset spaces of the Euclidean group of rigid translations and rotations. Here the basic premises are stated in phase-space language, and the conclusions follow succinctly from the complete set of invariants that we find.

In geometric optics, light rays in a homogeneous medium are idealized as lines in three space $\vec{q}(s)$, with a coordinate $s \in \mathcal{R}$ (the real line) along the ray. This is the path of a classical photon in time $t = sn/c$, where n is the refractive index and c the velocity of light in vacuum. The manifold of all such lines is optical phase space; it has two position and two momentum coordinates that satisfy Hamilton's equations of evolution referred to the standard $z = 0$ plane screen.[10] In Section 2 we recall these facts, simplifying the proofs of Ref. 11. Our basic premise is that all optical transformations are canonical[12]: a deformation of the sphere of ray directions necessarily entails a conjugate transformation of ray positions.

The relativistic addition of velocities deforms the sphere of ray directions toward the axis of motion. In Section 3 we present the well-known ray-direction transformation and in Section 4 the new canonically conjugate position transformation. If an imaging device is set to focus on a nearby point light source at rest, we deduce its image when it is in motion. It is important to distinguish between a near source and a star. To form the image of a near source, a system will collect rays from a pencil of rays that range over a finite angle; from stars, a telescopic imaging system will receive only a pencil of parallel rays, where the comatic phenomenon is not apparent.

Section 5 typifies telescopic, identity, and general linear systems. In Section 6 we analyze the relativistic boost transformation on screens perpendicular to the direction of motion. We show that the image will magnify and aberrate, suffering a series of circular comatic aberrations of increasing order that we have termed relativistic coma.[8] The global method of invariants is valid for all angles (4π sr) and is used in Section 7 to find the transformations for any screen orientation. The concluding Section 8 adds some remarks on the scope of this analysis.

2. OPTICAL PHASE SPACE

Consider the lines $\vec{q}(s)$, $s \in \mathcal{R}$, $\vec{q} \in \mathcal{R}^3$, their tangents $\vec{p}(s) \in \mathcal{R}^3$, and vanishing intervals $\mathrm{d}s$, in continuous, differentiable media. On purely geometric grounds we write

$$\frac{\mathrm{d}\vec{q}}{\mathrm{d}s} = \frac{\vec{p}}{|\vec{p}|} = \frac{\partial \mathcal{H}}{\partial \vec{p}}. \tag{2.1}$$

The first term is a unit vector because $\mathrm{d}s = |\mathrm{d}\vec{q}|$, and the last term proposes a function $\mathcal{H}(\vec{q},\vec{p})$ that will generate evolution along s.[12] Performing partial integration with respect to $\vec{p}$, up to an undetermined function n of $\vec{q}$ and

0740-3232/93/091925-10$06.00

additive constant, we obtain

$$\mathcal{H} = |\vec{p}| - n(\vec{q}). \tag{2.2}$$

The dynamic characterization of light in inhomogeneous optical media is the local form of Snell's law: the differential change of the line direction $d\vec{p}$ is ds times the space gradient of the refractive index:

$$\frac{d\vec{p}}{ds} = \frac{\partial n(\vec{q})}{\partial \vec{q}} = -\frac{\partial \mathcal{H}}{\partial \vec{q}}. \tag{2.3}$$

The generating function $\mathcal{H}$ is the same as in Eq. (2.2) when the arbitrary function $n(\vec{q})$ is indentified with the refractive index of physical optics, and it is the evolution Hamiltonian in the parameter s.

The Hamiltonian flow in s preserves the surfaces $\mathcal{H}(\vec{q},\vec{p}) = 0$; i.e.,

$$|\vec{p}| = n(\vec{q}) \tag{2.4}$$

holds, for all points of three space and any ray direction. In this way, the ray direction three vector $\vec{p}$ is constrained to range over a sphere $\mathcal{S}_2$, whose radius is the local refractive index. For image-forming devices we use Cartesian coordinates, defining the $z = 0$ plane as the stardard screen and the $+z$ direction as the optical axis. Since the Hamiltonian flow is constrained to $\mathcal{H}$ = constant surfaces, six coordinates $(\vec{q},\vec{p})$ are redundant for the system. We will parameterize (almost all) rays by four coordinates $(\mathbf{q},\mathbf{p})$, $\mathbf{q} = (q_x, q_y)$, $\mathbf{p} = (p_x, p_y)$, and a chart index σ = sign p_z, so $p_z = \sigma(n^2 - p_x{}^2 - p_y{}^2)^{1/2}$. This index distinguishes between rays in the hemisphere along the optical axis ($\sigma = +$) from those counter to it ($\sigma = -$). Only for $p_z = 0$ rays, i.e., those parallel to the screen, does this parameterization of the ray-vector bundle fail[13]; being a lower-dimensional manifold, this fact need not concern us overly. In the azimuth and collatitude coordinates (θ,ϕ) of the ray direction sphere,

$$|\mathbf{p}| = n \sin\theta, \tag{2.5a}$$

$$p_z = n\cos\theta. \tag{2.5b}$$

We have used the length s along the ray as evolution parameter in the Hamilton equations (2.1)–(2.3); now we reduce these to equations in the evolution parameter $q_z = z$, the distance along the optical axis.[14,15] The differentials ds, $d\mathbf{q}$, and dz form a right triangle in 3-space[11] that is similar and has the same orientation as the triangle of sides n, $\mathbf{p}$, and $p_z = \sigma(n^2 - p^2)^{1/2}$, where we indicate that $p^2 = |\mathbf{p}|^2$. Hence $ds = dz n/p_z$ ($p_z \neq 0$). Replacing the differentials reduces the Hamilton equations in 3-space to Hamilton equations on the two-dimensional plane of the screen $[\mathbf{q}(z),\mathbf{p}(z),\sigma]$,

$$\frac{d\mathbf{q}}{dz} = \frac{\mathbf{p}}{p_z} = \frac{\partial h}{\partial \mathbf{p}} = \{-h,\circ\}\mathbf{q}, \tag{2.6a}$$

$$\frac{d\mathbf{p}}{dz} = \frac{n}{p_z}\frac{\partial n}{\partial \mathbf{q}} = \frac{\partial h}{\partial \mathbf{q}} = \{-h,\circ\}\mathbf{p}, \tag{2.6b}$$

where the screen Hamiltonian function that generates z evolution is

$$h(\mathbf{q},\sigma,\mathbf{p};z) = -\sigma\sqrt{n(\mathbf{q},z)^2 - p^2} \tag{2.7}$$

and we have used the Lie–Poisson operator notation

$$\{f,\circ\}g = \{f,g\} = \frac{\partial f}{\partial \mathbf{q}}\cdot\frac{\partial g}{\partial \mathbf{p}} - \frac{\partial f}{\partial \mathbf{p}}\cdot\frac{\partial g}{\partial \mathbf{q}}, \tag{2.8}$$

applicable also for 2-vector functions.

Because they appear in a pair of Hamilton equations, $\mathbf{q}$ and $\mathbf{p}$ in each σ chart are canonically conjugate coordinates of the ray.[12] Position (or configuration) space $\mathbf{q}$ is the $\mathcal{R}^2$ plane of the screen. The conjugate coordinates $\mathbf{p}$ are the optical momentum-vector components on the screen; it is the projection (p_x, p_y) of the ray direction 3-vector $\vec{p}$ and has bounded length $|n \sin\theta| \leq n$. Notice carefully the momentum domains: the projection of the $\mathcal{S}_2$ surface onto its equatorial plane is composed of two open disks $\mathcal{D}_\sigma = \pm$ of radius n and of the connecting circle of $\sigma = 0$ rays that are parallel to the screen. Optical phase space is then a symplectic manifold $\mathcal{P} = \mathcal{R}^2 \times \mathcal{S}_2 \doteq \mathcal{R}^2 \times (\mathcal{D}_+ + \mathcal{D}_- + \mathcal{D}_0)$.

For example, in a homogeneous optical medium, with n = constant, the Hamilton equations may be readily integrated to

$$\mathbf{q}(z) = \mathbf{q}(0) + z\frac{\mathbf{p}}{\sigma(n^2 - p^2)^{1/2}}, \qquad \mathbf{p}(z) = \mathbf{p}(0),$$

$$\sigma(z) = \sigma. \tag{2.9}$$

The momentum of the ray $\mathbf{p}$ remains constant, while $\mathbf{q}$ increases along $\mathbf{p}$ by $z\tan\theta$ [cf. Eq. (2.5)]; $\mathbf{q}(z)$ is the intersection of the straight ray line with a screen that is translated along the optical axis by z. In this case the σ charts do not mix.

We recall that a system governed by a pair of integrable Hamilton equations evolves in such a way that the flow of its phase-space points $\mathbf{q} \mapsto \mathbf{q}' = \mathbf{q}'(\mathbf{q},\mathbf{p},\sigma;z)$, $\mathbf{p} \mapsto \mathbf{p}' = \mathbf{p}'(\mathbf{q},\mathbf{p},\sigma;z)$ preserves the basic Poisson brackets: $\{q_i,q_j\} = 0$, $\{q_i,p_j\} = \delta_{i,j}$, and $\{p_i,p_j\} = 0$, and the value of $h'(\mathbf{q},\mathbf{p},\sigma) = h(\mathbf{q}'(\mathbf{q},\mathbf{p},\sigma),\mathbf{p}'(\mathbf{q},\mathbf{p},\sigma),\sigma') = h(\mathbf{q},\mathbf{p},\sigma)$ for all z.[12] Liouville's theorem states that the flow of the phase space under Hamiltonian evolution is divergenceless, like that of an incompressible fluid. The flow is along the lines h = constant, so the Hamiltonian function is an invariant of the flow. Free flight (or displacement of the standard screen in the z direction) given by Eq. (2.9) is shown in Fig. 1 for two-dimensional phase space

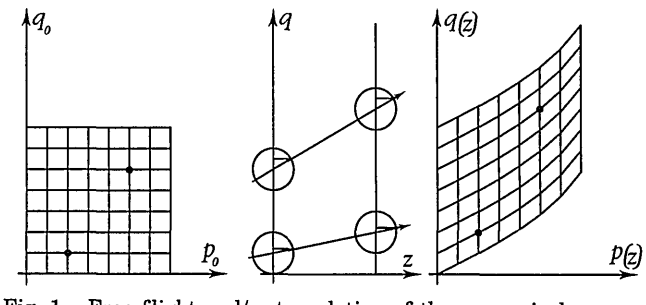

Fig. 1. Free flight and/or translation of the screen in homogeneous media. Points in a patch of (two-dimensional) optical phase space (left-hand screen) represent rays crossing the (one-dimensional) screen (middle panel). Free flight by generic z to a second (right-hand) screen deforms this patch of phase space (left-hand side of screen) carrying the rays. This transformation changes the position coordinate of the ray but maintains the ray directions invariant. Area elements are preserved, because the transformation is canonical.

(i.e., two-dimensional optics on one-dimensional screens). In this case, the incompressibility of the flow is equivalent to the preservation of Poisson brackets, because $dq' \wedge dp' = \{q', p'\} dq \wedge dp$.[13] In four or more dimensions, the conservation of Poisson brackets is a stronger statement than the incompressibility of the flow. Phase-space maps that preserve Poisson brackets are called canonical transformations.[12]

The basic assumption in Hamilton–Lie optics is canonicity: all optical transformations are canonical. (This principle is local; on the other hand, Fermat's minimum-action principle,[10] leading naturally to the Euler–Lagrange formulation of optics, is global.) Translations along the chosen optical z axis in an inhomogeneous medium are canonical, and so are x and y translations and screen rotations.[7] That optical transformations besides Euclidean ones are also canonical is what makes the canonical principle interesting. Refraction or reflection by a smooth arbitrary surface separating two media with indices $n \neq n'$ is not a Hamiltonian evolution process (in any obvious way, yet see Ref. 10); it is a finite, sudden transformation that is, nevertheless, canonical.[16] Our hypothesis is that velocity boosts of optical systems are also canonical transformations of optical phase space.

3. RELATIVISTIC DEFORMATION OF MOMENTUM SPACE

Consider now 4-vectors in the common notation of special relativity, $\mathsf{I} = (\vec{\mathsf{I}}, \mathsf{I}_0) = (\mathsf{I}_x, \mathsf{I}_y, \mathsf{I}_z, \mathsf{I}_0)$, and boost $\mathcal{B}_\beta$ in the z direction to a velocity $v = (c/n) \tanh \beta$, $\beta \in \mathcal{R}$, in a homogeneous medium of refractive index n, which we set henceforth equal to unity. This boost acts on the components of I through

$$\mathcal{B}_\beta: \mathsf{I}_x \mapsto \mathsf{I}_x' = \mathsf{I}_x, \qquad \mathcal{B}_\beta: \mathsf{I}_y \mapsto \mathsf{I}_y' = \mathsf{I}_y, \tag{3.1a}$$

$$\mathcal{B}_\beta: \mathsf{I}_z' \mapsto \mathsf{I}_z' = \mathsf{I}_z \cosh \beta + \mathsf{I}_0 \sinh \beta, \tag{3.1b}$$

$$\mathcal{B}_\beta: \mathsf{I}_0 \mapsto \mathsf{I}_0' = \mathsf{I}_z \sinh \beta + \mathsf{I}_0 \cosh \beta. \tag{3.1c}$$

Now let I be a lightlike vector, so that $|\vec{\mathsf{I}}| = |\mathsf{I}_0|$. The sphere $\mathcal{S}_2$ of ray directions $\vec{p}$ is a homogeneous coordinate set of the $\mathcal{R}^3$ space of 3-vectors $\vec{\mathsf{I}}$:

$$\vec{p} = (p_x, p_y, p_z) = \left(\frac{\mathsf{I}_x}{\mathsf{I}_0}, \frac{\mathsf{I}_y}{\mathsf{I}_0}, \frac{\mathsf{I}_z}{\mathsf{I}_0}\right), \tag{3.2}$$

Transformation (3.1) thus becomes the following transformation of optical momentum space and Hamiltonian:

$$\mathcal{B}_\beta: \mathbf{p} \mapsto \mathbf{p}' = \frac{\mathbf{p}}{\cosh \beta + p_z \sinh \beta}, \tag{3.3a}$$

$$\mathcal{B}_\beta: \mathbf{p}_z \mapsto \mathbf{p}_z' = \frac{p_z \cosh \beta + \sinh \beta}{\cosh \beta + p_z \sinh \beta}, \tag{3.3b}$$

where $\mathbf{p} = (p_x, p_y)$ as in Section 2. We see that $\mathcal{B}_\beta$ is a nonlinear, 1:1 deformation of phase space $\mathfrak{p}$. It mixes the σ charts; they are tracked by the sign of p_z and p_z' in Eq. (3.3b). We note the vector form $\mathbf{p}'(\mathrm{p}, \sigma; \beta) = P(|\mathbf{p}|^2, \sigma; \beta)\mathbf{p}$ of this momentum transformation.

In terms of the polar angles (θ, ϕ) of the ray-direction sphere, the z boost leaves the azimuth ϕ invariant, and the colatitude θ transforms by

$$\mathcal{B}_\beta: \frac{p}{p_z + 1} = \mathcal{B}_\beta: \tan \tfrac{1}{2}\theta \mapsto \tan \tfrac{1}{2}\theta' = \exp(-\beta) \tan \tfrac{1}{2}\theta. \tag{3.4}$$

This realization of the map $\mathcal{B}_\beta$ is known to group theorists as the Bargmann deformation of the circle[4] and is shown in Fig. 2. The projection of this transformation onto the screen is Eq. (3.3a). In ca. 1725 Bradley[17] observed the linear part of this action on star images for $\theta' - \theta \approx 20''$, recognized that it originated from the orbital motion of the Earth, and named it stellar aberration—a misnomer, since images remain pointlike, even though the star field is distorted.

The Bargmann Lorentz map $\mathcal{B}_\beta$ does not conserve length elements on the circle of Fig. 2; instead $d^2\mathbf{p}'/p'^2 = d^2\mathbf{p}/p^2$. The conserved quantity is, by Liouville's theorem, the four-dimensional volume element of $\mathfrak{p}$, $d\mathbf{q} \wedge d\mathbf{p}$. Thus transformations of momentum $\mathbf{p}$ are bound to canonically conjugate transformations of the position coordinates $\mathbf{q}$. The boost transformation of ray directions thus must entail a corresponding transformation of optical images.

4. RELATIVISTIC ABERRATION OF RAY POSITIONS

Now let us find the transformation of the image positions that is canonically conjugate to Eq. (3.3a) by building the invariants of the boost. We look at transformations in a neighborhood of β around zero. There we can expand Eq. (3.3a) to first order in $d\beta$ and, for some function $B(\mathbf{q}, \mathbf{p}, \sigma)$ to be found, write

$$\mathbf{p}'(\mathbf{p}, \sigma; d\beta) = \frac{\mathbf{p}}{1 + d\beta p_z} \approx \mathbf{p} - d\beta p_z \mathbf{p} = (\mathcal{I} + d\beta\{B, \circ\})\mathbf{p}. \tag{4.1}$$

This has the form of a second Hamilton evolution equation for $d\mathbf{p} = \mathbf{p}'(d\beta) - \mathbf{p}(0)$,

$$\frac{d\mathbf{p}}{d\beta} = -p_z \mathbf{p} = -\frac{\partial B}{\partial \mathbf{q}} = \{B, \circ\}\mathbf{p}. \tag{4.2}$$

We should now find the boost generator function $B(\mathbf{q}, \mathbf{p}, \sigma)$; its action on position space $\mathbf{q}'(\mathbf{q}, \mathbf{p}, \sigma; \beta)$ will be conjugate to that of $\mathbf{p}$ above.

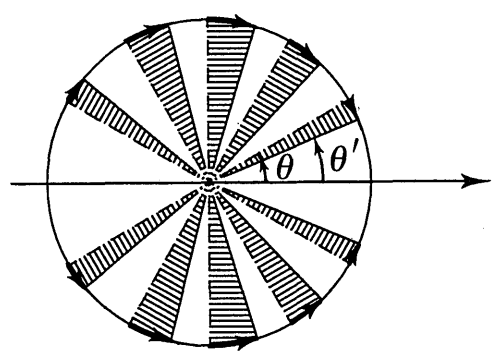

Fig. 2. Bargmann Lorentz transformation of the circle. This is the deformation of the colatitude circle of light-ray directions under a boost toward the right. Its projection on a vertical line is the action of boosts on optical momentum.

Since Poisson brackets involve derivatives, the degree of homogeneity of B in the components of $\mathbf{q}$ must be 1 or 0. Since the momentum transformation has a vector form, B must be a scalar function of phase space, i.e., invariant under rotations around the optical axis (in the x–y planes of $\mathbf{p}$ and $\mathbf{q}$) and under reflections across planes that contain that axis (say, $x \leftrightarrow -x$). The latter requirement excludes additive terms with a factor $\mathbf{q} \times \mathbf{p}$ from being present in B. The first two then allow only the forms $\mathbf{q} \cdot \mathbf{p} b(|\mathbf{p}|^2, \sigma) + a(|\mathbf{p}|^2, \sigma)$, where b can be found from Eq. (4.2). (Since $\partial B/\partial \mathbf{q} = b\mathbf{p}$, it follows that $b = p_z$) and a is arbitrary. The effect of an $a(|\mathbf{p}|^2)$ on $\mathbf{q}'(\mathbf{q}, \mathbf{p})$ is to introduce an additional spherical aberrationlike term $\beta a' \mathbf{p}$ (such as z translation when $a = p_z$); on momentum space, the presence of an a has no effect. Spherical aberration always can be factored off from the boost transformation (spherical aberration and coma generate a solvable subgroup of aberrations[18]), and for this reason we set $a = 0$. We can thus write the boost generator function

$$B(\mathbf{q}, \mathbf{p}, \sigma) = -\mathbf{q} \cdot \mathbf{p} p_z = -\mathbf{q} \cdot \mathbf{p} \sigma \sqrt{1 - p^2}. \tag{4.3}$$

Therefore the first Hamilton-type equation generated by the boost in the parameter β will be

$$\frac{d\mathbf{q}}{d\beta} = -p_z \mathbf{q} + \mathbf{q} \cdot \mathbf{p} \frac{\mathbf{p}}{p_z} = \frac{\partial B}{\partial \mathbf{p}} = \{B, \circ\}\mathbf{q}. \tag{4.4}$$

Since $B(\mathbf{q}, \mathbf{p}, \sigma)$ is independent of β, the infinitesimal transformation (4.1) may be integrated to a finite one as an exponential Lie–Poisson operator $\mathcal{B}_\beta = \exp(\beta\{B, \circ\})$. The boost $\mathcal{B}_\beta$ thus defined has a natural canonical action on optical phase space $\mathfrak{p}$ and in particular on the screen position coordinates of optical rays on the screen: $\mathbf{q}'(\mathbf{q}, \mathbf{p}; \beta) = \exp(\beta\{B, \circ\})\mathbf{q}$. We now shall find $\mathbf{q}'(\mathbf{q}, \mathbf{p}; \beta)$ explicitly without need of integration.

Since evidently $\{B, B\} = 0$, it follows that $B(\mathbf{q}, \mathbf{p}, \sigma)$ is an invariant under the transformation $\mathcal{B}_\beta$ that it generates. Moreover, in three-dimensional optics, because $\mathbf{q}'(d\beta)$ and $\mathbf{p}'(d\beta)$ are vector functions of $\mathbf{q}$ and $\mathbf{p}$, we can find a second quantity that is invariant under z boosts: the cross product (the square root of the Petzval[19]):

$$L = \mathbf{q} \times \mathbf{p} = q_x p_y - q_y p_x, \tag{4.5}$$

because $\{B, L\} = 0$, as we may easily verify. [In fact, it is a skew invariant because it changes sign, $L \mapsto -L$, under reflections $(q_x, q_y, p_x, p_y) \mapsto (-q_x, q_y, -p_x, p_y)$, as we already have seen.]

Finally, note that the repeated Poisson bracket of a scalar quantity (such as B) of the form $\mathbf{q} \cdot \mathbf{p} b(|\mathbf{p}|^2) + a(|\mathbf{p}|^2)$ with a vector quantity of the form $c(p^2)\mathbf{q} + \mathbf{q} \cdot \mathbf{p} d(p^2)\mathbf{p}$ is again of the latter form (for different functions c and d). Therefore the exponential map that integrates the evolution equations (4.2) and (4.4) leads to the generic form

$$\mathbf{q}'(\mathbf{q}, \mathbf{p}; \beta) = R(p^2; \beta)\mathbf{q} + \mathbf{q} \cdot \mathbf{p} S(p^2; \beta)\mathbf{p}. \tag{4.6}$$

The invariance of $\mathbf{q} \times \mathbf{p} = L = L' = R\mathbf{q} \times \mathbf{p}'$ and the known momentum transformation (3.3a) then imply that

$$R(p^2; \beta)^{-1} = \cosh \beta + \sqrt{1 - p^2} \sinh \beta. \tag{4.7}$$

We use this information in the boost invariant $B' = B$ to write

$$\begin{aligned} \mathbf{q} \cdot \mathbf{p} p_z &= \mathbf{q}' \cdot \mathbf{p}' p_z' \\ &= [R(p^2; \beta)\mathbf{q} + \mathbf{q} \cdot \mathbf{p} S(p^2; \beta)\mathbf{p}] \cdot [R(p^2; \beta)^{-1}\mathbf{p}] p_z' \\ &= \mathbf{q} \cdot \mathbf{p}[1 + p^2 S(p^2; \beta) R(p^2; \beta)^{-1}] p_z' \end{aligned} \tag{4.8}$$

and solve algebraically for the function S. It is

$$\begin{aligned} S(p^2; \beta) &= \frac{R(p^2; \beta)}{p^2}\left(\frac{p_z}{p_z'} - 1\right) \\ &= -\sinh \beta \frac{\cosh \beta + p_z \sinh \beta}{p_z \cosh \beta + \sinh \beta}. \end{aligned} \tag{4.9}$$

We have thus extended the boost transformation from ray directions to the full optical phase space $\mathfrak{p}$. The complete form of the boost transformation along the optical z axis is therefore

$$\mathcal{B}_\beta: \mathbf{q} \mapsto \mathbf{q}' = (\cosh \beta + p_z \sinh \beta) \times \left(\mathbf{q} - \frac{\mathbf{q} \cdot \mathbf{p} \sinh \beta}{p_z \cosh \beta + \sinh \beta}\mathbf{p}\right), \tag{4.10a}$$

$$\mathcal{B}_\beta: \mathbf{p} \mapsto \mathbf{p}' = \frac{\mathbf{p}}{\cosh \beta + p_z \sinh \beta}. \tag{4.10b}$$

To picture this transformation we draw the results for two-dimensional optics, where screens are one-dimensional (position $q \in \mathcal{R}$). There the boost transformation may be put in terms of the position and the azimuth angle as

$$q'(q, \theta; \beta) = q \frac{(\cosh \beta + \cos \theta \sinh \beta)^2}{\cosh \beta + \sec \theta \sinh \beta}, \tag{4.11a}$$

$$\sin \theta'(q, \theta; \beta) = \frac{\sin \theta}{\cosh \beta + \cos \theta \sinh \beta}. \tag{4.11b}$$

This is shown in Fig. 3 for $p \in (-1, 1)$. The (single) quantity invariant under boosts is $-B = q \sin \theta \cos \theta = (1/2)q \sin 2\theta$; the lines of flow of phase space in the figure are B = constant.

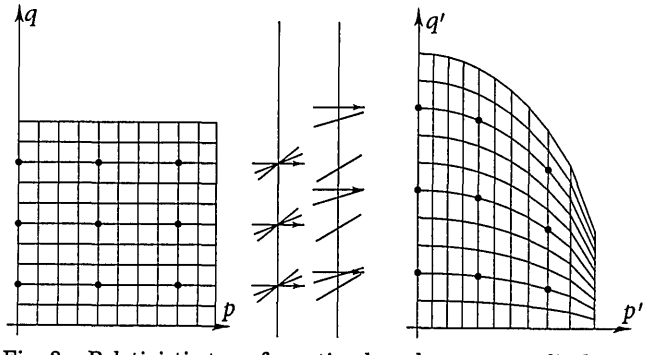

Fig. 3. Relativistic transformation by a boost perpendicular to the screen. As in Fig. 1, a patch of phase space (left-hand screen) is followed under the boost (right-hand screen). Three ray pencils that at rest converge to image points on the screen magnify and aberrate after the boost, no longer forming point images. The rays along the optical axis remain as such and define the tip of the comatic image; those that form a nonzero angle map 2:1 and fall below. A pencil of rays represented by a region of phase will yield its image on the screen when projected on the vertical q axis.

5. TELESCOPIC, IDENTITY, AND LINEAR IMAGING SYSTEMS

We now analyze the performance of optical systems moving at relativistic velocities, their objects, and their corresponding images, to highlight the central role of the identity system.

A system focused at infinity is telescopic[20]: it forms images of the incoming ray directions on its screen, thus measuring the momentum $\mathbf{p}$ of the rays. The points of phase space of Figs. 1 and 3 are projected on the p abscissas while integrated over the q ordinates of all rays that enter its front lens. If this system is aligned with its $+z$ axis along the direction of motion, the colatitude angle [Eq. (3.4)] will change as shown in Fig. 2. With increasing speed, the star images (ray directions) will crowd the optical center ($\mathbf{q} = \mathbf{0}$) of the screen and remain point images.

The identity system is a simple screen onto which rays fall. This screen senses the position $\mathbf{q}$ of the rays but not their direction of arrival. In Figs. 1 and 3 (for two-dimensional optics), the identity system registers the ordinate of the points of phase space, projecting them on the vertical q axis, integrating over the abscissa of ray angles.

Paraxial, axis-symmetric imaging optical systems are well known and often used as a class of ideal devices. They transform object rays $(\mathbf{p}^o, \mathbf{q}^o)$ into image rays $(\mathbf{p}^i, \mathbf{q}^i)$ linearly:

$$\mathcal{M}: \begin{bmatrix} \mathbf{q}^o \\ \mathbf{p}^o \end{bmatrix} \mapsto \begin{bmatrix} \mathbf{q}^i \\ \mathbf{p}^i \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} \mathbf{q}^o \\ \mathbf{p}^o \end{bmatrix}. \tag{5.1}$$

The canonicity of the transformation requires that the matrix $\mathbf{M}$ be unimodular, i.e., $\det \mathbf{M} = ad - bc = 1$. The use of linear systems explicitly assumes that the rays are paraxial: $|\mathbf{q}| \ll 1$ and $\theta \approx p/n \ll 1$. Classically, the generators of such systems are free flights and thin lenses. From these elements one produces compound systems by multiplying their matrices. When the system matrix is lower triangular, it is an imaging system,[20] since the image position will depend only on the object position:

$$\begin{bmatrix} \mathbf{q}^i \\ \mathbf{p}^i \end{bmatrix} = \begin{bmatrix} \mu & 0 \\ 1/f & 1/\mu \end{bmatrix} \begin{bmatrix} \mathbf{q}^o \\ \mathbf{p}^o \end{bmatrix} = \begin{bmatrix} \mu \mathbf{q}^o \\ 1/\mu \mathbf{p}^o + 1/f \mathbf{q}^o \end{bmatrix}, \tag{5.2}$$

where f is the focal length of the system and μ its magnification.

If a projector produces a pencil of rays $(\mathbf{q}^i, \mathbf{p})_{\mathbf{p}\in D}$, where D is a range of ray directions, that focuses at rest on the screen point $\mathbf{q}^i$, then, when boosted to relativistic motion, the pencil will become $[\mathbf{q}'(\mathbf{q}^i, \mathbf{p}; \beta), \mathbf{p}'(\mathbf{p}; \beta)]_{\mathbf{p}\in D}$, producing an image $\mathbf{q}'(\mathbf{q}^i, \mathbf{p}; \beta)_{\mathbf{p}\in D}$ that is given by Eqs. (4.10). For later analysis, let us Taylor expand this formula by powers of the phase-space variables as high as degree 5:

$$\mathbf{q}' = \exp(\beta)\mathbf{q} - n^{-2} \sinh \beta[\tfrac{1}{2}p^2\mathbf{q} + \mathbf{q} \cdot \mathbf{pp}] - n^{-4} \sinh \beta[\tfrac{1}{8}p^4\mathbf{q} + \tfrac{1}{2} \exp(-2\beta)p^2\mathbf{q} \cdot \mathbf{pp}] - \ldots, \tag{5.3}$$

$$\mathbf{p}' = \exp(-\beta)\mathbf{p} + \tfrac{1}{2}n^{-2} \sinh \beta \exp(-2\beta)p^2\mathbf{p} + \tfrac{1}{4}n^{-4} \sinh \beta \exp(-2\beta) \times (1 - \tfrac{1}{2} \exp(-2\beta))p^4\mathbf{p} + \ldots. \tag{5.4}$$

The first term in each of these series is linear, so the first-order approximation of the boost is thus that of a pure magnifying system:

$$\mathcal{B}_\beta^{[1]}: \begin{bmatrix} \mathbf{q} \\ \mathbf{p} \end{bmatrix} \mapsto \begin{bmatrix} e^\beta & 0 \\ 0 & e^{-\beta} \end{bmatrix} \begin{bmatrix} \mathbf{q} \\ \mathbf{p} \end{bmatrix}, \qquad q, p \ll 1. \tag{5.5}$$

Within the paraxial regime thus, a camera with the observer or a projector at the roadside will, after boosting in the optical axis direction, produce the same, expanded image: $(\mathbf{q}')^i = \mathbf{q}^i[\mathbf{q}'(\mathbf{q}^o, \mathbf{p}^o)] = \mu\mathbf{q}'(\mathbf{q}^o, \mathbf{p}^o) = \mu \exp(\beta)\mathbf{q}^o$ in the first case and $(\mathbf{q}^i)' = \mathbf{q}'(\mathbf{q}^i, \mathbf{p}^i) = \mathbf{q}'[\mu\mathbf{q}^o, \mathbf{p}^i(\mathbf{q}^o, \mathbf{p}^o)] = \exp(\beta)\mu\mathbf{q}^o$ in the second case. This is still a focused point image, since it is independent of the issuing ray directions $\mathbf{p}^o$. For general nonimaging systems [Eq. (5.1)], the boost matrix [Eq. (5.5)] multiplies the system matrix by the right-hand side in the first case (the relativistic boost affects the object rays) and by the left-hand side in the second (the boost acts on image rays).

Linear optical systems are the first-order approximation to a true optical transformation by refracting surfaces. The latter may be expanded in powers of the components of $\mathbf{q}$ and $\mathbf{p}$, as in Eqs. (5.3)–(5.4), and calculated up to some degree, the aberration order. The composition of systems and boosts of the last paragraph may be extended (with the aid of symbolic computation) for third, fifth, etc., aberration order. This process highlights the role of the identity system to separate the relativistic aberration phenomenon from the rest of the transformations of the system. Our paradigmal setup is thus a simple screen carried by the observer.

6. SPOT DIAGRAM OF RELATIVISTIC COMA

The spot diagram at $\mathbf{q}$ of an optical transformation $(\mathbf{q}, \mathbf{p}) \mapsto (\mathbf{q}', \mathbf{p}')$ is the map from a region $\mathbf{p} \in D$ of the sphere of object-ray directions onto a region $\mathbf{q}' \in D' \in \mathcal{R}^2$ on the screen for fixed $\mathbf{q}$. Quite universally, the region D is chosen to be a polar coordinate cap of the sphere of ray directions, with the colatitude θ ranging in steps from 0 to some maximal angle $\theta_{\max}$ that, in the present case, is bounded only by π. As we shall see, relativistic boosts give rise to spot diagrams that yield to simple geometrical analysis and deserve the name comatic.

We consider a fixed object point $\mathbf{q} = (q_x, q_y) = (q \cos \kappa, q \sin \kappa)$, and we let the ray $\mathbf{p} = (p_x, p_y) = (p \cos \phi, p \sin \phi)$, $p = n \sin \theta$, trace out a cone in ϕ, at a fixed colatitude angle θ. Then the vector $\mathbf{q}'(q, \kappa, p, \phi, ; \beta)$ in Eq. (5.3) will trace out a closed curve on the boosted screen. The curve may be computed easily in Cartesian components and is of the form

$$\mathbf{q}'[\mathbf{q}(q, \kappa), \mathbf{p}(p, \phi)] = \{[R + \tfrac{1}{2}p^2S]\mathbf{q} + (\tfrac{1}{2})qp^2S\mathbf{r}(\phi)\}, \tag{6.1a}$$

where

$$\mathbf{r}(\phi) = \begin{bmatrix} \cos(2\phi - \kappa) \\ \sin(2\phi - \kappa) \end{bmatrix}, \qquad -\pi < \phi \le \pi, \tag{6.1b}$$

with $R = R(p^2; \beta)$ and $S = S(p^2; \beta)$ as in Eqs. (4.6) *et seq.*

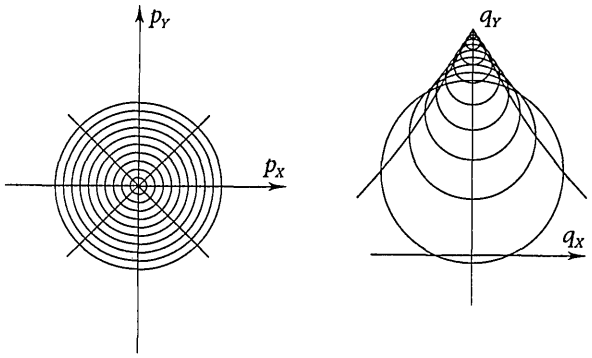

Fig. 4. The spot diagram of the relativistic boost transformation is the map $\mathbf{p} \mapsto \mathbf{q}'(\mathbf{p})$, where $\mathbf{p}$ ranges over a polar coordinate grid (left-hand diagram) and a corresponding region on the screen $\mathbf{q}'$ (right-hand diagram). The optical center is down, and, since the spot grows linearly in the direction of $\mathbf{q}$, this spot diagram is generic for all $\mathbf{q} \in \mathcal{R}^2$. (Compare with Fig. 3 for two-dimensional optics.)

To fifth-order aberration, we may write

$$\mathbf{q}' = \{\exp(\beta) - p^2 \sinh\beta - \tfrac{1}{8}p^4 \times [1 + \exp(-2\beta)]\sinh\beta - \ldots\}\mathbf{q} - [\tfrac{1}{2}p^2 q \sinh\beta + \tfrac{1}{4}p^4 q \exp(-2\beta)\sinh\beta + \ldots]\mathbf{r}(\phi). \quad (6.2)$$

In Fig. 4 we draw the spot diagram of the boost transformation at point $\mathbf{q} = (q, 0)$ for a range of directions up to $\theta_{\max} = 45°$. Since the spot grows linearly with q and rotates with the orientation of $\mathbf{q}$, this shape is generic.

The sets $\{\mathbf{q}'[\mathbf{q}(q,\kappa), \mathbf{p}(p,\phi)]\}_{-\pi<\phi\le\pi}$ are circles on the screen, drawn out twice by the vector $\mathbf{r}(\phi)$, with

center at $\mathbf{q}_{\text{center}}$

$$= [R + \tfrac{1}{2}Sp^2]\mathbf{q} = \exp(\beta)\mathbf{q} - p^2 \sinh\beta\mathbf{q} - \ldots, \quad (6.3a)$$

of radius ρ

$$= \tfrac{1}{2}qp^2 S = -\tfrac{1}{2}qp^2 \sinh\beta - \ldots. \quad (6.3b)$$

In this way we represent

$$\mathbf{q}'(\mathbf{q}, \mathbf{p}; \beta) = \mathbf{q}_{\text{center}}(\mathbf{q}, p^2; \beta) + \rho(q, p^2; \beta)\mathbf{r}(\kappa, \phi). \quad (6.3c)$$

For fixed $\mathbf{q}$ and letting the ray $\mathbf{p}$ roam on a polar coordinate grid of the forward-momentum disk ($\sigma = +$), the image on the boosted screen (Fig. 4) has the shape of a comet (coma). The apex of this coma is the map of the ray along the optical axis ($p = 0$), which is affected only by the scale factor $\exp(\beta)$ of paraxial transformation (5.5). The circles p = constant have their centers on the line between this apex and the optical center of the screen. As p grows, their centers move from the apex by $\delta = |\mathbf{q}_{\text{center}} - \exp(\beta)\mathbf{q}| = qp^2 \sinh\beta$, while their radii increase as $\rho = \tfrac{1}{2}qp^2 \sinh\beta$. The opening angle of the comet at the apex is 60° as shown in Fig. 4, because $\delta/\rho = \tfrac{1}{2} = \sin 30°$ is the half-angle. For growing p in Eq. (6.2), the p^4 and higher terms open the coma in the manner of a funnel.[8]

For $\beta > 0$, all rays with $\sigma = +$ fall onto the boosted screen with the same $\sigma' = -$. For $0 \le \theta < 1/2\pi$ ($0 \le p < 1$, $1 \ge p_z > 0$), the centers of the ϕ circles move from the apex toward the optical center, while their radii increase from $\rho(\theta = 0) = 0$ to $\rho[\theta = \tfrac{1}{2}\pi] = \tfrac{1}{2}q \cosh\beta$. When $\beta < 0$ the apex of the coma is at $\mathbf{q}'(\mathbf{p} = 0) = \exp(-|\beta|)\mathbf{q}$, nearer than $\mathbf{q}$ from the optical center, so the comet opens outward. Neither the denominator in Eq. (4.10b) nor the factor in Eq. (4.10a) vanishes, because $\cosh\beta > |\cos\theta \sinh\beta|$ for all $|\beta| < \infty$. However, the denominator in Eq. (4.10a) does have a simple zero at a finite critical angle θ_c such that $\cos\theta_c = \tanh|\beta|$, or $\tan\tfrac{1}{2}\theta_c = \exp(-|\beta|)$, whose image is $\theta' = \tfrac{1}{2}\pi$ in Eq. (4.11b), i.e., rays that after boosting fall on directions parallel to the screen. It is a mathematical singularity of the screen and chart coordinates. Beyond this critical colatitude value, $\sigma = +$ rays focused on the receding screen will appear to come from the backward hemisphere: in Eq. (3.3b), $\sigma' = \text{sign}\, p_z' = -$. Since our screen also records rays from the backward hemisphere, the coma figure generated by colatitude circles continues uneventfully in the interval $\theta_c < \theta \le \pi$ through the value $\theta = \tfrac{1}{2}\pi$. The circles shrink with further-increasing θ, and for the ray $\theta = \pi$ (from the backward pole) one obtains $\mathbf{q}' = \exp(|\beta|)\mathbf{q}$, i.e., the image of the forward pole when $\beta > 0$.

Figure 5 shows, as a function of θ,

The image of meridional rays ($\mathbf{q} \| \mathbf{p}$)

$$q_\| ' = (R + Sp^2)q, \quad (6.4a)$$

The center of the circles

$$q_{\text{center}} = [R + \tfrac{1}{2}Sp^2)q, \quad (6.4b)$$

The image of sagittal rays ($\mathbf{q} \perp \mathbf{p}$)

$$q_\perp ' = Rp^2 q. \quad (6.4c)$$

From Eqs. (4.10) we see that inverting the observer's velocity $\beta \leftrightarrow -\beta$ is equivalent to $p_z \leftrightarrow -p_z$; i.e., $\sigma \leftrightarrow -\sigma$. The spot around the backward pole is shown on the right-hand side of Fig. 5. When the region D is the full sphere of ray directions, we have a global coma spot diagram. It is a map of the sphere on the plane; it is 2:1, the two poles are mapped on the distinct apices $\mathbf{q}_{\pm\text{pole}} = \exp(\pm\beta)\mathbf{q}$ that open with 60° angles; and it is singular at the critical colatitude value θ_c. We thus characterize the relativistic boost beyond aberration-order expansions as a global aberration, i.e., valid for the full ray-direction sphere.[7]

In geometric optics, caustics appear when a spot diagram folds over itself, so rays accumulate at the fold

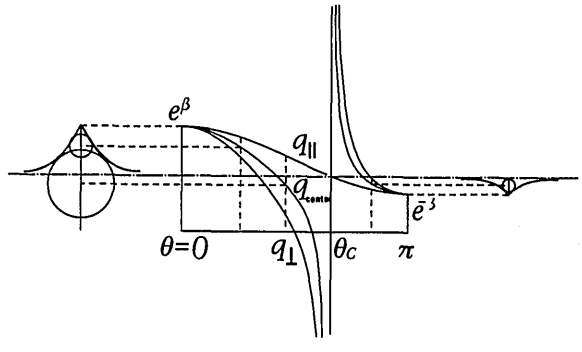

Fig. 5. Circles and caustics of relativistic coma. The middle portion shows the behavior of the center, upper, and lower edges of the circles in the spot diagram as functions of colatitude angle. To the left, the spot and the caustic around the forward ($+z$) ray; to the right, the same around the backward ($-z$) ray.

edge. For fixed $\mathbf{q}$, an area element of momentum space $dp_x \wedge dp_y = \frac{1}{2}dp^2 \wedge d\phi$ vanishes with the Hessian of the two components of $\mathbf{q}'$ with respect to the two components of $\mathbf{p}$. From Eq. (6.2) we obtain[13,21]

$$dq_x' \wedge dq_y' = 2\rho(\dot{q}_{\text{center}} \cos 2\phi + \dot{\rho})dp^2 \wedge d\phi, \qquad (6.5)$$

where the overdots indicate derivatives with respect to p^2. The Hessian vanishes when $\rho = 0$ (i.e., at the coma's apices) and also when the factor in parentheses vanishes: there are four lines $\pm\phi_o(p^2), \pm[\pi - \phi_o(p^2)]$ that map on the caustic edges. The caustics of the global relativistic coma spot diagram form a diamond shape, with two 60°-angle tips at $\mathbf{q}_{\pm\text{pole}} = \exp(\pm\beta)\mathbf{q}$, and two pairs of edges asymptotically tangent to a line perpendicular to $\mathbf{q}$ through $\mathbf{q}_{\text{asymp}} = \operatorname{sech}\beta\mathbf{q}$. This is read in Fig. 5.

7. LORENTZ MAPS OF OPTICAL PHASE SPACE

Imaging systems boosted along the optical axis are but a special situation of Lorentz (rotation and boost) aberration phenomena. The use that we have made of optical invariants to analyze z boosts extends easily to all Lorentz transformations, supported by the structure of Lie algebras and groups on optical phase space $\mathfrak{p}$.

Under rotations by α around the y axis, the transformation of the components of the lightlike vector $\mathfrak{l}$ of Section 3 is

$$\mathcal{J}_{y,\alpha}: \mathfrak{l}_0 \mapsto \mathfrak{l}_0' = \mathfrak{l}_0, \qquad \mathcal{J}_{y,\alpha}: \mathfrak{l}_y \mapsto \mathfrak{l}_y' = \mathfrak{l}_y, \qquad (7.1a)$$

$$\mathcal{J}_{y,\alpha}: \mathfrak{l}_x \mapsto \mathfrak{l}_x' = \mathfrak{l}_x \cos\alpha + \mathfrak{l}_z \sin\alpha, \qquad (7.1b)$$

$$\mathcal{J}_{y,\alpha}: \mathfrak{l}_z \mapsto \mathfrak{l}_z' = -\mathfrak{l}_x \sin\alpha + \mathfrak{l}_z \cos\alpha. \qquad (7.1c)$$

The transformation of optical momentum $\mathbf{p}$ is therefore

$$p_x' = p_x \cos\alpha + \sigma(1 - p_x^2 - p_y^2)^{1/2}\sin\alpha$$
$$= \sigma \sin\alpha + p_x \cos\alpha - \sigma\tfrac{1}{2}(p_x^2 + p_y^2)\sin\alpha - \ldots, \qquad (7.2a)$$

$$p_y' = p_y. \qquad (7.2b)$$

$$p_z' = -p_x \sin\alpha + p_z \cos\alpha, \qquad (7.2c)$$

where $\sigma = \pm$ is the chart index. Hemispheres mix under rotations; the sign σ' of p_z' is found from Eq. (7.2c). Series expansion of this rotation shows that a thin ray pencil $|p_x| \ll 1$ is translated to $\sigma \sin\alpha$ in the x direction, magnified by a factor $\cos\alpha$ in that direction, and suffers from a series of rotationally symmetric aberration terms $\sim(px^2 + p_y^2)^k$, $k = 1, 2, \ldots$.

In the same way as for the z boost generator in Section 4, we find the Lie–Poisson generator of infinitesimal y rotations, $\mathcal{J}_{y,d\alpha} = 1 + d\alpha\{J_y, \circ\}$, on optical phase space $\mathfrak{p}$. The generating function $J_y(\mathbf{q}, \mathbf{p})$ is found by integration of $\{J_y, p_x\} = \sigma(1 - p_x^2 - p_y^2)^{1/2}$ and $\{J_y, p_y\} = 0$; it is

$$J_y = q_x\sigma(1 - p_x^2 - p_y^2)^{1/2} = q_x p_z \qquad (7.3)$$

up to additive functions of $\mathbf{p}$ alone. Similarly, we find the Lie generator functions of rotations around the x and z axes for each σ chart.[22] They are

$$J_x = -q_y\sigma(1 - p_x^2 - p_y^2)^{1/2} = -q_y p_z, \qquad (7.4)$$

$$J_z = q_y p_x - q_x p_y = -\mathbf{q} \times \mathbf{p}. \qquad (7.5)$$

The last function appeared above as the Petzval invariant in Eq. (4.5) with the name $-L$ (it is not an invariant under x or y rotations, of course).

The generator functions of rotations J_x, J_y, and J_z in Eqs. (7.3)–(7.5) form the basis of a three-dimensional vector space of functions on $\mathfrak{p}$ that transform among themselves under rotations and close under Poisson brackets:

$$\{J_x, J_y\} = -J_z, \qquad \text{and cyclically.} \qquad (7.6)$$

The formalism of angular momentum in quantum mechanics and the geometric image aberrations that are due to screen rotation have the same Lie-algebraic structure. Hence

$$J^2 = J_x^2 + J_y^2 + J_z^2 = (q_x^2 + q_y^2)(1 - p^2) - (\mathbf{q} \times \mathbf{p})^2$$
$$= q^2 - (\mathbf{q} \cdot \mathbf{p})^2 \qquad (7.7)$$

is an invariant under rotations, because $\{\vec{J}, J^2\} = 0$. When we substitute Eqs. (7.2) into Eq. (7.7), from $J^2(\mathbf{q}', \mathbf{p}') = J^2(\mathbf{q}, \mathbf{p})$ it follows that the spot diagrams of rotation [$\mathbf{q}'(\mathbf{q}, \mathbf{p})$ for $\mathbf{q}$ fixed and rays ranging over colatitude circles] are conic sections (ellipses, parabolas, and hyperbolas). This is only the classical construction that cuts the circular cone of rays with the inclined-plane screen.

There are three independent boosts along the x, y, and z axes; their generators transform as the components of a 3-vector under rotations,

$$\{J_x, B_y\} = -B_z, \qquad \text{and cyclically,} \qquad (7.8)$$

and close with the rotation generators under Poisson brackets:

$$\{B_x, B_y\} = +J_z, \qquad \text{and cyclically.} \qquad (7.9)$$

We note the plus in Eq. (7.9) versus the minus in Eqs. (7.6) and (7.8); this distinguishes the Lorentz group $SO(3,1)$ from the four-dimensional rotation group $SO(4)$. From here we find that

$$B_x = q_x p_z^2 + p_y \mathbf{q} \times \mathbf{p} = q_x - p_x \mathbf{q} \cdot \mathbf{p}, \qquad (7.10)$$

$$B_y = q_y p_z^2 - p_x \mathbf{q} \times \mathbf{p} = q_y - p_y \mathbf{q} \cdot \mathbf{p}, \qquad (7.11)$$

$$B_z = -\mathbf{q} \cdot \mathbf{p}\sigma\sqrt{1 - p^2} = -p_z \mathbf{q} \cdot \mathbf{p}. \qquad (7.12)$$

We note that B_x and B_y are independent of the chart index $\sigma = \pm$; their action therefore will not mix the two hemispheres of rays.

The action of boosts along the y axis may be found as in Eqs. (3.1) and (7.1) on ray direction and on optical momentum. It is given by the first two of the following

equations:

$$p_x' = \frac{p_x}{\cosh\beta + p_y \sinh\beta}$$
$$= p_x \operatorname{sech}\beta - p_x p_y \operatorname{sech}\beta \tanh\beta + p_x p_y^2 \operatorname{sech}\beta \tanh^2\beta \pm \dots, \quad (7.13a)$$

$$p_y' = \frac{p_y \cosh\beta + n \sinh\beta}{\cosh\beta + p_y \sinh\beta}$$
$$= n \tanh\beta + p_y \operatorname{sech}^2\beta - p_y^2 \tanh\beta \operatorname{sech}^2\beta + p_y^3 \tanh^2\beta \operatorname{sech}^3\beta \pm \dots, \quad (7.13b)$$

$$p_z' = \frac{p_z}{\cosh\beta + p_y \sinh\beta}$$
$$= p_z \operatorname{sech}\beta - p_z p_y \operatorname{sech}\beta \tanh\beta + p_z p_y^2 \operatorname{sech}\beta \tanh^2\beta \pm \dots, \quad (7.13c)$$

where we write the first- (linear), second-, and third-order aberrations [cf. Eqs. (3.3) and (6.2) for z boosts and Eqs. (7.2) for y rotations]. A telescopic system focused on stars $|p_y| \ll 1$ and boosted in the y direction continues to draw sharp images, but from the series developments in Eqs. (7.13) it follows that the images are displaced to $p_y' = \tanh\beta$, and the star field is magnified by $\operatorname{sech}\beta \le 1$ in the x direction and by $\operatorname{sech}^2\beta$ in the y direction and with further asymmetric distorting terms in the aberration series.

Following our program, we now find the image transformations that accompany rotations and boosts of the sphere of ray directions. Again we study the invariants: under rotations the generating function J_y is invariant, while under boosts B_y is invariant. Furthermore, the Poisson bracket of the two functions is zero. Thus they are simultaneously invariant under their commuting finite Lorentz transformations. This is sufficient to permit us to find that $\mathbf{q}' = (q_x', q_y')$ as a function of $\mathbf{q}$ and $\mathbf{p}$, when we replace $\mathbf{p}'(\mathbf{p})$ correspondingly. For y rotations, $q_x p_z = J_y = J_y' = q_x' p_z'$ and Eq. (7.2c) lead to

$$q_x' = \frac{q_x}{\cos\alpha - (p_x/p_z)\sin\alpha}$$
$$= q_x \sec\alpha + q_x p_x \sec\alpha \tan\alpha + q_x p_x^2 \sec\alpha \tan^2\alpha + \dots. \quad (7.14a)$$

For the y component, $q_y(p_x^2 + p_z^2) - q_x p_x p_y = B_y = B_y' = q_y'(p_x'^2 + p_z'^2) - q_x' p_x' p_y'$, Eq. (7.2b), and Eq. (7.14a) yield

$$q_y' = q_y + \frac{q_x p_y}{p_z} \frac{\sin\alpha}{\cos\alpha - (p_x/p_z)\sin\alpha}$$
$$= q_y + q_x p_y \tan\alpha + q_x p_x p_y \tan^2\alpha + \dots. \quad (7.14b)$$

We see that q_x is magnified by $\sec\alpha \ge 1$ (the factor is inverse to that of p_x) and q_y has unit magnification. These facts may be found by ordinary vector analysis but here serve to illustrate the Lie method of invariants.

For y boosts, from $J_y = J_y'$ and Eq. (7.13a) we find the transformation for the coordinate across the direction of boost that is linear plus quadratic:

$$q_x' = q_x \cosh\beta - q_x p_y \sinh\beta. \quad (7.15a)$$

For the y component, $B_y = B_y'$ and Eq. (7.15a) yield, after rearrangement,

$$q_y' = (\cosh\beta + p_y \sinh\beta)(q_y \cosh\beta + \mathbf{q}\cdot\mathbf{p} \sinh\beta). \quad (7.15b)$$

This is an up-to-cubic transformation. To first order, q_x is magnified by $\cosh\beta \ge 1$ and q_y by $\cosh^2\beta$; these factors are inverse to those of their conjugate quantities in Eqs. (7.13). Paraxial images thus expand and distort. (The last two transformations, reported in Ref. 8, were calculated by back-and-forth guesswork with the aid of computer symbolic manipulation; here the use of invariants permits their direct algebraic derivation.)

To higher orders, phase space aberrates: the quadratic terms $\sim -q_x p_y \sinh\beta$ in q_x' and $\sim(q_x p_x + 2q_y p_y)\sinh\beta\cosh\beta$ in q_y' are the second-order aberrations. To draw the spot diagram of y boosts[7] and to understand the geometry in a simple way, we write

$$q_x' = q_x^o + v_x, \qquad q_x^o = q_x \cosh\beta, \qquad v_x = A\sin\phi,$$
$$q_y' = q_y^o + v_y, \qquad q_y^o = q_y \cosh^2\beta,$$
$$v_y = B\cos\phi + C\sin\phi, \quad (7.16)$$

where ϕ is the azimuth of the ray (counted from the x to the y axis). The colatitude θ appears in the coefficients [cf. Eqs. (7.15)]

$$A = q_x \sin\theta \sinh\beta, \qquad B = q_x \sin\theta \sinh\beta \cosh\beta,$$
$$C = 2q_y \sin\theta \sinh\beta \cosh\beta. \quad (7.17)$$

When $\mathbf{p}$ sweeps out in ϕ a circle of radius $p = \sin\theta$, the second-order aberration vector $\mathbf{v} = (v_x v_y)$ draws out a closed curve on the screen that conserves

$$\frac{B^2 + C^2}{A^2 B^2} v_x^2 - \frac{2C}{AB^2} v_x v_y + \frac{1}{B^2} v_y^2 = 1, \quad (7.18)$$

i.e., the equation of a conic. This may be compared to the equation for a standard ellipse in the x–y plane, of half-axes a and b, rotated by an angle τ (in the common sense of ϕ and κ on the screen), which is

$$\left(\frac{c^2}{a^2} + \frac{s^2}{b^2}\right)x^2 + 2cs\left(\frac{1}{a^2} - \frac{1}{b^2}\right)xy + \left(\frac{s^2}{a^2} + \frac{c^2}{b^2}\right)y^2 = 1, \quad (7.19)$$

where $s = \sin\tau$ and $c = \cos\tau$. Comparison of the three coefficients allows us to compute a, b, and $\tan 2\tau$ of the ellipses as functions of θ, q, κ, and the boost parameter β. The formulas simplify for $|\beta| \ll 1$; then Eqs. (7.17) become $A = B = \beta q_x p$ and $C = 2\beta q_y p$. In this case the ellipses are tilted by

$$\tau = \tfrac{1}{2}\kappa + \tfrac{1}{4}\pi \quad (7.20)$$

and have half-axes

$$a = \beta q(1 - \sin\kappa)\sin\theta, \qquad b = \beta q(1 + \sin\kappa)\sin\theta. \quad (7.21)$$

As we show in Fig. 6, for $\kappa = 0$ (positions on the x axis) the spot is a circle, while for $\kappa = \pm\frac{1}{2}\pi$ (the y axis) the spot degenerates to a line along the direction of the boost. In

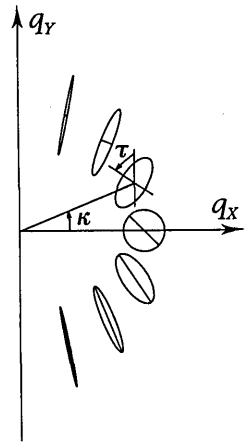

Fig. 6. Spot diagrams (to second order) of boosts along the y axis for various values of κ, the polar angle on the screen. The spot shapes are generic because their sizes increase linearly with q, the distance to the optical center.

general it is a rotated ellipse. The spot of second-order aberrations grows linearly with distance q to the optical center; its shape depends only on the polar angle κ on the screen. The shape shown in the figure is thus generic.

We found the rotation invariant J^2 in Eq. (7.7); from there and from Eqs. (7.10)–(7.12) we find that $J^2 - B^2$ and $\vec{J} \cdot \vec{B}$, the two Lorentz (Casimir) invariants,[23] are identically zero in the geometric-optics representation of the group. Therefore there are no further functionally independent invariants under boosts in any direction $\vec{w}$ and rotations around the $\vec{w}$ axis beyond the corresponding J_w and B_w.

8. CONCLUDING REMARKS

We presented with some care the foundations of Hamiltonian geometric optics on the phase space $\mathfrak{p}$ so that the basic hypothesis of canonicity of Lorentz transformations is evident and its conclusions are well grounded. Geometric optics and special relativity may seem to be unweddable theories, because the latter involves time in an essential way,[1,2] while the former divides the proportional s out of the evolution equations (the same remarks hold between Maxwell and Helmholtz optics[9]).

The Hamilton–Lie formulation of optics determines that the global action of Lorentz transformations results in aberrations of optical images. These aberrations involve magnification and circular comatic aberration when the direction of the boost coincides with the optical z axis. This spot diagram is a peculiar comatic 2:1 map of the sphere on the plane. (The spot diagram of y boosts in 4π is also 2:1, because a ray and its reflection $p_z \mapsto -p_z$ fall upon the same point of the screen, although on different sides of it.) These two spot diagrams are but two faces, related by screen rotation, of the relativistic aberration phenomenon. For every direction in which we point our system, we can perform a similar analysis. The spot diagrams are representations by pictures of the Lorentz-group elements.

There are other physical considerations to be taken into account before a *Gedankenexperiment* can shed the qualifier, since light is actually a quantum-wave phenomenon. In our context, however, the *Gedanken* part of this experiment illuminates the aberrations of one- and two-dimensional signals with bounded spectrum, as general as the concept of phase space itself. The Lorentz group $SO(3,1)$ is a finite-parameter Lie subgroup of all continuous, nonlinear transformations $\mathcal{T}(\mathcal{S}_2)$ of the sphere $\mathcal{S}_2$ of ray directions on itself. The extension of the action of $\mathcal{T}(\mathcal{S}_2)$ to the symplectic space $\mathfrak{p}$ is, as here, generated by functions that are linear in the components of position: of the form $f_x(\mathbf{p})q_x + f_y(\mathbf{p})q_y$. When this generating function is a z axis rotation invariant, the conjugate image aberration will always be a series of circular comas.

ACKNOWLEDGMENTS

I acknowledge help and encouragement on the relativistic coma aberration by N. M. Atakishiyev (Baku) and W. Lassner (Leipzig) as coauthors, J. O. Castañeda (Instituto Nacional de Astrofísica, Optica y Electrónica, Tonantzintla) as critical listener, and G. Krötzsch (Instituto de Investigaciones en Mathemáticas Aplicadas y en Sistemas, Universidad Nacional Autónoma de México, Cuernavaca) for the friendly graphical display.

REFERENCES AND NOTES

1. J. Terrell, "Invisibility of the Lorentz contraction," Phys. Rev. **116,** 1041–1045 (1959).
2. G. D. Scott and M. R. Viner, "The geometrical appearance of large objects moving at relativistic speeds," Am. J. Phys. **33,** 534–537 (1965); G. D. Scott and H. J. van Driel, "Geometrical appearances at relativistic speeds," Am. J. Phys. **38,** 971–977 (1970); A. Peres, "Relativistic telemetry," Am. J. Phys. **55,** 516–519 (1987); E. Sheldon, "The twists and turns of the Terrell effect," Am. J. Phys. **56,** 199–200 (1988).
3. J. Terrell, "The Terrell effect," Am. J. Phys. **57,** 9–10 (L) (1992); E. Sheldon, "The Terrell effect: eppure si contorce," Am. J. Phys. **57,** 487 (L) (1992); T. E. Phillips, "Relativity of aberration," Am. J. Phys. **57,** 553–555 (1992).
4. V. Bargmann, "Irreducible unitary representations of the Lorentz group," Ann. Math. **48,** 568–640 (1947).
5. K. B. Wolf, "A recursive method for the calculation of the SO_n, $SO_{n,1}$, and $ISO(n)$ representation matrices," J. Math. Phys. **12,** 197–206 (1971).
6. C. P. Boyer and K. B. Wolf, "Deformations of inhomogeneous classical Lie algebras to the algebras of the linear groups," J. Math. Phys. **14,** 1853–1859 (1973); K. B. Wolf and C. P. Boyer, "The algebra and group deformations $I^m[\mathrm{SO}(n) \otimes \mathrm{SO}(m)] \Rightarrow \mathrm{SO}(n,m)$, $I^m[U(n) \otimes U(m)] \Rightarrow U(n,m)$, and $I^m[\mathrm{Sp}(n) \otimes Sp(m)] \Rightarrow \mathrm{Sp}(n,m)$," J. Math. Phys. **15,** 2096–2101 (1974).
7. K. B. Wolf, "Elements of Euclidean optics," in *Lie Methods in Optics, Second Workshop,* K. B. Wolf, ed., Vol. 352 of Lecture Notes in Physics (Springer-Verlag, Heidelberg, 1989), pp. 116–162.
8. N. M. Atakishiyev, W. Lassner, and K. B. Wolf, "The relativistic coma aberration. I. Geometric optics," J. Math. Phys. **30,** 2457–2462 (1989).
9. N. M. Atakishiyev, W. Lassner, and K. B. Wolf," The relativistic coma aberration. II. Helmholtz wave optics," J. Math. Phys. **30,** 2463–2468 (1989).
10. A. J. Dragt, E. Forest, and K. B. Wolf, "Foundations of a Lie algebraic theory of geometrical optics," in *Lie Methods in Optics,* J. Sánchez Mondragón and K. B. Wolf, eds., Vol. 250 of Springer Lecture Notes in Physics (Springer-Verlag, Heidelberg, 1986), pp. 105–157.
11. T. Sekiguchi and K. B. Wolf, "The Hamiltonian formulation of optics," Am. J. Phys. **55,** 830–835 (1987).
12. H. Goldstein, *Classical Mechanics,* 2nd ed. (Addison-Wesley, Reading, Mass., 1980).
13. V. Guillemin and S. Sternberg, *Symplectic Techniques in Physics* (Cambridge U. Press, Cambridge, 1984).

14. A. J. Dragt, "Lie-algebraic theory of geometrical optics and optical aberrations," J. Opt. Soc. Am. **72,** 372–379 (1982).
15. E. López Moreno and K. B. Wolf, "De la ley de Snell–Descartes a las ecuaciones de Hamilton en el espacio fase de la óptica geométrica," Rev. Mex. Fís. **35,** 291–300 (1989); G. Krötzsch and K. B. Wolf, "Las tres caras de Hamilton en la óptica geométrica y en la mecánica," Rev. Mex. Fís. **36,** 724–735 (1990).
16. M. Navarro-Saad and K. B. Wolf, "Factorization of the phase-space transformation produced by an arbitrary refracting surface," J. Opt. Soc. Am. A **3,** 340–346 (1986).
17. J. Bradley, "Aberration," Philos. Trans. R. Soc. London **35,** 637 (1729).
18. K. B. Wolf, "Symmetry-adapted classification of aberrations," J. Opt. Soc. Am. A **5,** 1226–1232 (1988).
19. H. Buchdahl, *An Introduction to Hamiltonian Optics* (Cambridge U. Press, Cambridge, 1970), Sec. 122. Actually, traditional optics calls L^2 the Petzval sum: it is a third-order aberration, a linear combination of curvature of field and astigmatism, that is invariant under general paraxial transformations (Subsec. 5.3). For the relation between the two, see, e.g., Ref. 18.
20. O. Stavroudis, *The Optics of Rays, Wavefronts, and Caustics* (Academic, New York, 1972).
21. H. Flanders, *Differential Forms with Applications to the Physical Sciences* (Academic, New York, 1963).
22. K. B. Wolf, "Symmetry in Lie optics," Ann. Phys. **172,** 1–25 (1986).
23. R. Gilmore, *Lie Groups, Lie Algebras, and Some of Their Applications* (Wiley, New York, 1974).