Physica D 51 (1991) 189–199
North-Holland

Lie–Poisson description of Hamiltonian ray optics

Darryl D. Holm[a] and K. Bernardo Wolf[b]

[a]*Center for Nonlinear Studies and Theoretical Division, Los Alamos National Laboratory, MS B284, Los Alamos, NM 87545, USA*

[b]*Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas, Universidad Nacional Autonoma de Mexico-Cuernavaca, Apdo. Postal 20-726, 01000 Mexico D.F., Mexico*

We express classical Hamiltonian ray optics for light rays in axisymmetric fibers as a Lie–Poisson dynamical system defined in $\mathbb{R}^3$, regarded as the dual of the Lie algebra sp(2, $\mathbb{R}$). The ray-tracing dynamics is interpreted geometrically as motion in $\mathbb{R}^3$ along the intersections of two-dimensional level surfaces of the conserved optical Hamiltonian and the skewness invariant (the analog of angular momentum, conserved because of the axisymmetry of the medium). In this geometrical picture, a Hamiltonian level surface is a vertically oriented cylinder whose cross section describes the radial profile of the refractive index, and a level surface of the skewness function is a hyperboloid of revolution around a horizontal axis. Points of tangency of these surfaces are equilibria, which are stable when the Gaussian curvature of the Hamiltonian level surface (constrained by the skewness function) is negative definite at the equilibrium point. Examples are discussed for various radial profiles of the refractive index. This discussion places optical ray tracing in fibers into the geometrical setting of Lie–Poisson Hamiltonian dynamics and provides an example of optical ray trapping within separatrices (homoclinic orbits).

1. Optical phase space

The phase space of geometrical optics is four-dimensional. Referred to a standard planar screen, a ray is determined by two *position* coordinates $\boldsymbol{q} = (q_x, q_y)$ defining its intersection with the screen, and two *momentum* coordinates $\boldsymbol{p} = (p_x, p_y)$ that cue the projection onto the screen of a three-vector $\vec{n}$ tangent to the ray, whose length $n(\boldsymbol{q})$ is the refractive index of the medium at that point. See fig. 1. Only rays perpendicular to the *optical axis* cannot be parametrized in this manner.

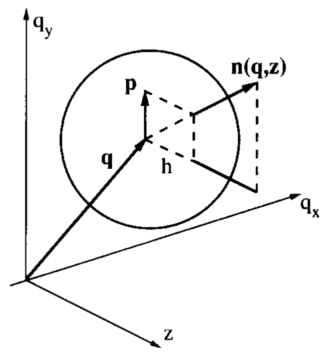

Fig. 1. Image screen phase space for geometrical optics.

The coordinate normal to the screen, z, extends along the optical axis. The projection along the optical axis of the vector $\vec{n}(q, z)$ – which in general is allowed to depend on z – is given by

$$h = \sqrt{n^2(\boldsymbol{q}, z) - p_x^2 - p_y^2} = -H^{\mathrm{opt}}.$$

The manifold $(\boldsymbol{q}, \boldsymbol{p})$ is symplectic, with Poisson bracket

$$\{F, G\} = \frac{\partial F}{\partial \boldsymbol{q}} \cdot \frac{\partial G}{\partial \boldsymbol{p}} - \frac{\partial F}{\partial \boldsymbol{p}} \cdot \frac{\partial G}{\partial \boldsymbol{q}},$$

and the evolution of the system is governed by Hamilton's equations,

$$\frac{\mathrm{d}\boldsymbol{q}}{\mathrm{d}z} = \{\boldsymbol{q}, H^{\mathrm{opt}}\} = \frac{\partial H^{\mathrm{opt}}}{\partial \boldsymbol{p}},$$

$$\frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}z} = \{\boldsymbol{p}, H^{\mathrm{opt}}\} = -\frac{\partial H^{\mathrm{opt}}}{\partial \boldsymbol{q}}.$$

2. Generators of flow

To any differentiable function $F(\boldsymbol{q}, \boldsymbol{p})$ we associate its *Lie operator* $\hat{F}(\boldsymbol{q}, \boldsymbol{p}; \partial/\partial \boldsymbol{q}, \partial/\partial \boldsymbol{p})$ [6],

$$\hat{F} = \{F, \circ\} = \frac{\partial F}{\partial \boldsymbol{q}} \cdot \frac{\partial}{\partial \boldsymbol{p}} - \frac{\partial F}{\partial \boldsymbol{p}} \cdot \frac{\partial}{\partial \boldsymbol{q}}.$$

A Lie operator $\hat{F}$ generates the *flow* of phase space $(\boldsymbol{q}, \boldsymbol{p})$ given by the vector field $(-\partial F/\partial \boldsymbol{p}, \partial F/\partial \boldsymbol{q})$, which is orthogonal to the gradient of F, $(\partial F/\partial \boldsymbol{q}, \partial F/\partial \boldsymbol{p})$, and, hence, is contained within the submanifolds $F(\boldsymbol{q}, \boldsymbol{p}) =$ constant, i.e. within its "level surfaces". The flow expressed by Hamilton's equations is the particular case for $F = -H^{\text{opt}} = h$. The evolution of an observable $g(\boldsymbol{q}, \boldsymbol{p})$ under the flow generated by an $\hat{F}$ will be governed by the analogous equation $\mathrm{d}g/\mathrm{d}\tau = \hat{F}g$, for flow parameter τ, and will be *canonical*, i.e. it will preserve Poisson brackets. If F is independent of τ we may integrate the flow equation into a *Lie transformation*,

$$g(\boldsymbol{q}, \boldsymbol{p}; \tau) = \exp(\tau \hat{F})\, g(\boldsymbol{q}, \boldsymbol{p}).$$

In particular, the skewness function

$$S = \boldsymbol{q} \times \boldsymbol{p} = q_x p_y - q_y p_x$$

generates rotations of phase space, of $\boldsymbol{q}$ and $\boldsymbol{p}$ jointly, each in its plane, around the optical axis. Its square, S^2, is called the *Petzval* invariant, and is conserved for ray optics in axisymmetric media.

3. Invariants in axisymmetric systems

When two functions, F and G, Poisson commute, $\{F, G\} = 0$, the flow generated by F takes place on the $G =$ constant submanifolds, and vice versa. This leads to the natural introduction of symmetry-adapted coordinates in phase space. Specifically, we take advantage of the fact that axisymmetric, translation-invariant fibers are described by Hamiltonians,

$$H = -\sqrt{n(q^2)^2 - p^2},$$

which Poisson commute with the skewness invariant (i.e. exhibit $\{H, S^2\} = 0$). We define the axisymmetric coordinates

$$X = q^2 \geq 0, \quad Y = p^2 \geq 0, \quad Z = \boldsymbol{q} \cdot \boldsymbol{p},$$

for which $\{S^2, X\} = \{S^2, Y\} = \{S^2, Z\} = 0$. In fact, any translation-invariant Hamiltonian describing a system that is symmetric under rotations around the optical axis (and under reflections across planes containing this axis) may be expressed as a function of only these axisymmetric variables, $H = H(X, Y, Z)$. Actually, the general *fiber* Hamiltonian is a function of only *two* of these variables, $H = H(X, Y)$, since no terms appear in H that depend on the relative angle φ between $\boldsymbol{q}$ and $\boldsymbol{p}$ through $\boldsymbol{q} \cdot \boldsymbol{p} = qp \cos \varphi = Z$.

The Hamiltonian flow will take place on level surfaces $H =$ constant, or

$$Y = n^2(X) - H^2.$$

These are surfaces generated by a line parallel to the Z-axis whose intercept describes the refractive index profile in the $Z = 0$ plane, and which are restricted to the first quadrant $X \geq 0$, $Y \geq 0$. See fig. 2.

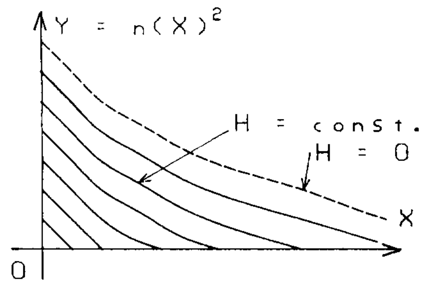

Fig. 2. Level surfaces of the Hamiltonian are cylinders whose intersection with the X–Y plane describes the radial refractive-index profile.

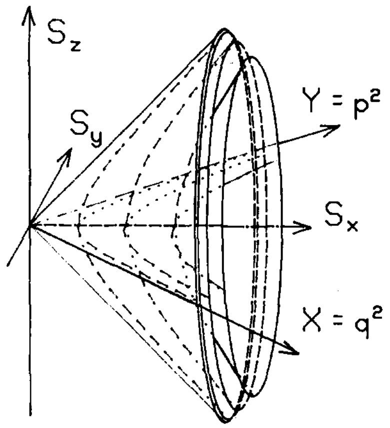

Fig. 3. Level surfaces of the skewness function are hyperboloids of revolution around the $X = Y$ axis at $Z = 0$.

Concurrently, the flow will lie on level surfaces $S^2 = \text{constant}$, determined by $(\boldsymbol{q} \times \boldsymbol{p})^2 = \boldsymbol{q}^2\boldsymbol{p}^2 - (\boldsymbol{q} \cdot \boldsymbol{p})^2 = \text{constant}$, or

$$XY - Z^2 = S^2 \geq 0.$$

The $S = \text{constant}$ submanifolds are seen to be hyperboloids of revolution around the $X = Y$ axis, extending up through the interior of the $S = 0$ cone, and lying between the X- and Y-axes. See fig. 3. The intercepts between the hyperboloids and the axis of revolution of the cone occur at $X = Y = S$. Through the discussion and examples in the following sections, the reader should come to appreciate the geometry introduced by the (X, Y, Z) coordinates and the level surfaces of H and S^2.

4. $\mathbb{R}^3$ vector formulation of ray optics: equilibria and stability conditions

Hamilton's equations for ray optics may be rewritten in terms of the $\mathbb{R}^3$ coordinates $\boldsymbol{x} = (X, Y, Z)$ in vector form as

$$\dot{\boldsymbol{x}} = \nabla 2S^2 \times \nabla H.$$

This vector-cross-product form of the optics equations ensures that both of the quantities H and S^2 are conserved, and that the motion of the system takes place in $\mathbb{R}^3$ along the intersections of the level surfaces of these two conserved quantities.

The level surfaces of H are cylinders extending along the Z-axis with cross-sections determined by the refractive index profile, while level surfaces of S^2 are hyperboloids of revolution around the line $X = Y$ in the positive quadrant of the X–Y plane. At the points of tangency of these level surfaces, the gradients of H and S^2 become collinear. These tangency points are the equilibria, i.e., the fixed points of the flow in $\mathbb{R}^3$.

At an equilibrium point, the constrained surface composed of the sum

$$H_S(\boldsymbol{x}) = H + \lambda S^2$$

has a critical point, for a multiplier λ that may depend upon the value of the equilibrium point, $\boldsymbol{x}_e$. At such a point, we have

$$\delta H_S = 0 = \mathrm{D}H_S(\boldsymbol{x}_e) \cdot \delta \boldsymbol{x}.$$

Written out, this becomes

$$\delta H_S = \left(\frac{1}{2H}\frac{\mathrm{d}n^2}{\mathrm{d}X} + \lambda Y\right)\delta X + \left(-\frac{1}{2H} + \lambda X\right)\delta Y - 2\lambda Z \delta Z.$$

At criticality each coefficient vanishes, so

$$Z_e = 0, \quad \lambda = \frac{1}{2H_e X_e} < 0, \quad X_e \frac{\mathrm{d}n^2}{\mathrm{d}X_e} + Y_e = 0.$$

Hence, critical points occur in the $Z = 0$ plane only, for values of H and S^2 such that the gradients $\nabla H = (1/2H)(\mathrm{d}n^2/\mathrm{d}X, -1)$ and $\nabla S^2 = (Y, X)$ are collinear. These are the equilibria of the ray optics system.

The second variation,

$$\delta^2 H_S = \delta \boldsymbol{x}^t \cdot \mathrm{D}^2 H_S(\boldsymbol{x}_e) \cdot \delta \boldsymbol{x},$$

is conserved by the linearized equations. (This is because $\delta^2 H_S$ is the Hamiltonian for the linearized dynamics around $\boldsymbol{x}_e$. See ref. [3], appendix A, and ref. [7] for discussions of linearized Hamiltonian dynamics.) When the equilibrium point $\boldsymbol{x}_e$ satisfies the conditions for $D^2 H_S(\boldsymbol{x}_e)$ to be definite in sign, the second variation $\delta^2 H_S$ may be taken as a norm for measuring perturbations from equilibrium, and its conservation in that case implies stability of the equilibrium. (Phase points that start near the equilibrium stay near it in the linearly conserved norm $\delta^2 H_S$ when $D^2 H_S(\boldsymbol{x}_e)$ is definite; so in that case the equilibrium is stable.) Consequently, $\boldsymbol{x}_e$ will be a stable equilibrium when the following symmetric matrix is definite:

$$D^2 H_S = \begin{pmatrix} A & B & 0 \\ B & C & 0 \\ 0 & 0 & -2\lambda \end{pmatrix},$$

where

$$A = -\frac{1}{2H}\left[\frac{1}{2H^2}\left(\frac{dn^2}{dX}\right)^2 - \frac{d^2 n^2}{dX^2}\right],$$

$$B = \frac{1}{4H^3}\frac{dn^2}{dX} + \lambda, \quad C = -\frac{1}{4H^3}.$$

Since -2λ and $-H$ are positive already, the remaining requirement for positive definiteness of $D^2 H_S$ is that its determinant $\det(D^2 H_S)$ be positive. Now, in $\mathbb{R}^3$ the Gaussian curvature of the constrained surface $H_S(\boldsymbol{x}) = H + \lambda S^2$ is given by [1]

$$K = -\frac{\det(D^2 H_S)}{\left(1 + |DH_S|^2\right)^{5/2}}.$$

Therefore, definiteness of $D^2 H_S$ implies definiteness of the Gaussian curvature of the constrained Hamiltonian surface $H_S(\boldsymbol{x})$ at its critical points. In our particular case, DH_S vanishes at equilibrium and

$$-K = \det(D^2 H_S) = \frac{2\lambda}{8H^4 X^2}\left(X^2 \frac{d^2 n^2}{dX^2} + 2X\frac{dn^2}{dX} + 2H^2\right).$$

The determinant is positive (and the Gaussian curvature of $H_S(\boldsymbol{x}_e)$ is negative), provided

$$\frac{d^2(X^2 n^2)}{dX^2} = X^2\frac{d^2 n^2}{dX^2} + 4X\frac{dn^2}{dX} + 2n^2 < 0,$$

which means that the profile of $X^2 n^2(X)$ is concave downward at a stable equilibrium.

This vector formulation provides a geometrical picture of Hamiltonian ray optics. It characterizes the motion as taking place along intersections in $\mathbb{R}^3$ of two-dimensional level surfaces of the Hamiltonian and the skewness invariant. Furthermore, it characterizes the equilibria both as tangent points of these level surfaces, and as critical points of the Hamiltonian constrained by the skewness invariant. Finally, this formulation determines the stability of the equilibria in terms of the Gaussian curvature of the constrained Hamiltonian at its critical points. In particular, points for which the H and S^2 level surfaces are tangent and possess curvatures of the opposite sign are stable equilibria. (When the two level surfaces have curvatures of opposite sign at the tangent point, nearby motions along their intersections describe nested ellipses on each surface, thereby implying stability.) On the other hand, equilibria at tangencies of H and S^2 having same-sign curvatures may be either stable or unstable. Such an equilibrium is stable (unstable) when the Gaussian curvature of the constrained Hamiltonian at the critical point is negative (positive). In terms of the refractive-index profile, such an equilibrium is stable provided the profile of $X^2 n^2(X)$ is concave downward when evaluated at the equilibrium point in the domain $X \geq 0$, $Y \geq 0$.

5. $\mathbb{R}^3$ Poisson brackets and Lie–Poisson brackets

The Poisson brackets among the axisymmetric variables X, Y and Z close among themselves,

$$\{X,Y\} = 4Z, \quad \{Y,Z\} = -2Y, \quad \{Z,X\} = -2X.$$

Consequently, we may re-express the equations of Hamiltonian ray optics in axisymmetric media with $H = H(X,Y)$ as

$$\dot{X} = \{X,H\} = \{X,Y\}\frac{\partial H}{\partial Y} = 4Z\frac{\partial H}{\partial Y},$$

$$\dot{Y} = \{Y,H\} = \{Y,X\}\frac{\partial H}{\partial X} = -4Z\frac{\partial H}{\partial X},$$

$$\dot{Z} = \{Z,H\} = \{Z,X\}\frac{\partial H}{\partial X} + \{Z,Y\}\frac{\partial H}{\partial Y} = -2X\frac{\partial H}{\partial X} + 2Y\frac{\partial H}{\partial Y}.$$

For functions F and G of the axisymmetric variables $\boldsymbol{x} = (X,Y,Z)$ in $\mathbb{R}^3$ we may introduce the following Poisson bracket in triple-scalar-product form:

$$\{F,G\}(\boldsymbol{x}) = \nabla 2S^2 \cdot \nabla F \times \nabla G.$$

This bracket satisfies the properties required of a Poisson bracket: it is bilinear, antisymmetric, and satisfies the Jacobi identity. In fact, the Jacobi identity is satisfied by this Poisson bracket for *any* continuous choice of the function S^2. We call such a bracket an $\mathbb{R}^3$ Poisson bracket. Note that $\{S^2, G\} = 0$ for any function G in $\mathbb{R}^3$. Such functions that Poisson commute with every function are called Casimirs. Clearly, the Casimirs are constants of the motion, since they Poisson commute with every Hamiltonian. Thus, the preservation of the level surfaces of the skewness invariant S^2 is built into the Poisson bracket description for Hamiltonian optics in axisymmetric media. The equations of ray optics in vector form now reappear as Hamilton's equations,

$$\dot{\boldsymbol{x}} = \{\boldsymbol{x}, H\} = \nabla 2S^2 \times \nabla H,$$

and the $\mathbb{R}^3$ geometry of the previous section is recovered as a Hamiltonian system.

In the present case, the function S^2 is quadratic. In such cases, the $\mathbb{R}^3$ Poisson bracket becomes a *Lie–Poisson* bracket, defined on the dual $\mathfrak{g}^*$ of the Lie algebra $\mathfrak{g}$ by the general formula, for $\mu \in \mathfrak{g}^*$,

$$\{F,G\}(\mu) = \left\langle \mu, \left[\frac{\partial F}{\partial \mu}, \frac{\partial G}{\partial \mu}\right]\right\rangle,$$

where $\langle\ ,\ \rangle$: $\mathfrak{g}^* \times \mathfrak{g} \to \mathbb{R}$ denotes the pairing between the Lie algebra and its dual, and $[\ ,\ ]$ is the Lie algebra product. (The Lie–Poisson bracket was introduced in ref. [4]. See refs. [5, 8] for modern discussions of Lie–Poisson brackets.) If we choose a basis $\xi_1, \ldots, \xi_n$ of $\mathfrak{g}$, the structure constants C_{ij}^k are defined by

$$[\xi_i, \xi_j] = C_{ij}^k \xi_k,$$

where we sum over repeated indices. Let $x^1, \ldots, x^n$ be the corresponding dual basis, with pairing given by $\langle x^i, \xi_j\rangle = \delta_j^i$. Then the Lie–Poisson bracket is expressible as

$$\{F,G\} = -C_{ij}^k \mu_k \frac{\partial F}{\partial \mu_i}\frac{\partial G}{\partial \mu_j},$$

upon identifying $\mu = \mu_i x^i$.

Referring to the $\mathbb{R}^3$ Poisson brackets among the variables X, Y and Z quickly determines the structure constants C_{ij}^k for the Lie–Poisson description of Hamiltonian ray optics. Setting $\mu_i = x_i$, $i = 1,2,3$, with $x_1 = X$, $x_2 = Y$, $x_3 = Z$, gives $C_{12}^3 = -4$, $C_{23}^2 = 2$, $C_{31}^1 = 2$, and the rest either vanish, or are obtained from antisymmetry under exchange of indices. These are the structure constants of any of the Lie algebras sp(2, $\mathbb{R}$), sl(2, $\mathbb{R}$), su(1, 1), or so(2, 1). For definiteness, we refer to the reduced description of Hamiltonian ray optics in terms of axisymmetric $\mathbb{R}^3$ variables as being "Lie–Poisson on sp(2, $\mathbb{R}$)*".

The Lie operator $\hat{F}$ defined by the sp(2, $\mathbb{R}$)* Lie–Poisson bracket is given by

$$\hat{F} = \{F, \cdot\} = (\nabla 2S^2 \times \nabla F) \cdot \nabla.$$

In particular, we have the Lie operators of the coordinate basis X, Y, Z,

$$\hat{X} = -4Z\frac{\partial}{\partial Y} - 2X\frac{\partial}{\partial Z}, \quad \hat{Y} = 4Z\frac{\partial}{\partial X} + 2Y\frac{\partial}{\partial Z},$$

$$\hat{Z} = 2X\frac{\partial}{\partial X} - 2Y\frac{\partial}{\partial Y}.$$

These operators satisfy the commutation relations

$$[\hat{X}, \hat{Y}] = -4\hat{Z}, \quad [\hat{Y}, \hat{Z}] = 2\hat{Y}, \quad [\hat{Z}, \hat{X}] = 2\hat{X},$$

which is in keeping with the general relation satisfied by Lie operators,

$$[\hat{F}, \hat{G}] = -\{F, G\}^{\wedge}.$$

A Lie operator $\hat{F}$ generates an SP(2, $\mathbb{R}$) group transformation that preserves the level surfaces of S^2. (These are the SP(2, $\mathbb{R}$) "co-adjoint orbits".) In particular, the SP(2, $\mathbb{R}$) Hamiltonian flow generated by $\hat{H}$ preserves the level surfaces of S^2 (since the Lie operator $\hat{S}^2$ vanishes).

6. Reconstruction

Reconstruction of the ray optics solution in terms of the symplectic image-screen variables $(\boldsymbol{q}, \boldsymbol{p})$ from the reduced, axisymmetric sp(2, $\mathbb{R}$)* variables (X, Y, Z) requires determination of the axial angle, $\theta(z)$. This may be done by a quadrature from the reduced solution in several ways, since the motion is symplectic when restricted to level surfaces of H, to those of S^2, or to those of any SL(2, $\mathbb{R}$) linear combination of H and S^2. For example, defining polar coordinates (r, θ) with $\boldsymbol{q} = r(\cos\theta, \sin\theta)$ leads, as usual, to canonically conjugate pairs of variables (r, p_r) and (θ, p_θ). These variables are related to the $\mathbb{R}^3$ variables by

$$X = q^2 = r^2, \quad Y = p^2 = p_r^2 + \frac{p_\theta^2}{r^2},$$

$$Z = \boldsymbol{q} \cdot \boldsymbol{p} = r p_r, \quad S = p_\theta,$$

and they appear in a symplectic Poisson bracket,

$$\{F, G\} = \frac{\partial F}{\partial r}\frac{\partial G}{\partial p_r} - \frac{\partial F}{\partial p_r}\frac{\partial G}{\partial r} + \frac{\partial F}{\partial \theta}\frac{\partial G}{\partial p_\theta} - \frac{\partial F}{\partial p_\theta}\frac{\partial G}{\partial \theta}.$$

The optical Hamiltonian for axisymmetric, translation-invariant media is given in terms of these canonically conjugate variables by

$$H = -\sqrt{n^2(r) - p_r^2 - \frac{p_\theta^2}{r^2}},$$

which, as expected, is independent of θ. Consequently, one may solve for $\theta(z)$ by integrating its Hamilton equation,

$$\frac{\mathrm{d}\theta}{\mathrm{d}z} = \frac{\partial H}{\partial p_\theta} = -\frac{p_\theta}{H}\frac{1}{r^2} = -\frac{S}{H}\frac{1}{X(z)}.$$

In this way, the solution for $X(z)$ in the reduced variables determines the axial angle $\theta(z)$ by a quadrature, since S and H are constants.

7. Free propagation in a homogeneous medium

As the first specific example in the $\mathbb{R}^3$ axisymmetric variables, consider ray propagation along the optical axis z in a homogeneous medium, n = constant, generated by the optical Hamiltonian as a Lie transformation. This is readily integrable and yields the well known result

$$\boldsymbol{q} \to \boldsymbol{q} + z\frac{\boldsymbol{p}}{\sqrt{n^2 - p^2}}, \quad \boldsymbol{p} \to \boldsymbol{p},$$

where $p/\sqrt{n^2 - p^2}$ is the tangent of the angle between the ray vector $\vec{\boldsymbol{n}}$ and the optical axis. The singularity for $H = -\sqrt{n^2 - p^2} = 0$ above is due to the inadequacy of the standard screen

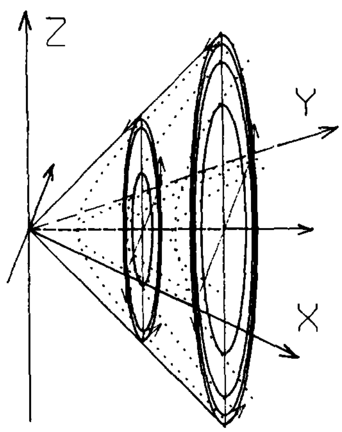

Fig. 4. Free motion in a homogeneous optical medium translates into the parabolas (heavy lines) that lie at the intersection of H = constant planes with the level surfaces of the skewness hyperboloid.

chart to describe grazing rays. Henceforth, we exclude the boundary, $H = 0$.

In the three-dimensional space (X, Y, Z), propagation through a homogeneous medium may be written as

$$X \to X + 2z\frac{Z}{H} + z^2\frac{Y}{H^2},$$
$$Y \to Y,$$
$$Z \to Z + z\frac{Y}{H}.$$

This is the family of parabolas shown in fig. 4 arising from the intersections of the planar level surfaces of H given by $Y = n^2 - H^2 > 0$, with the $S^2 > 0$ hyperboloids.

8. Elliptic-profile fibers

As a second example, consider the (exact, not "paraxial") propagation in an elliptic-profile graded-index fiber,

$$n^2(q^2) = n_0^2 - \nu^2 q^2, \quad \text{i.e. } n^2(X) = n_0^2 - \nu^2 X,$$
$$\nu^2 > 0.$$

(Again, this is not the "paraxial" parabolic profile $n(q^2) = n_0 - \nu q^2$.) The Hamiltonian is now

$$H = -\sqrt{n_0^2 - \nu^2 X - Y}$$

and the flow will take place on the *planes*

$$Y = n_0^2 - \nu X - H^2$$

intersecting the cone and hyperboloids of fig. 3. To compare, let us recall the exact solution obtained as a Lie transformation on $(\boldsymbol{q}, \boldsymbol{p})$ phase space,

$$\boldsymbol{q} \to \boldsymbol{q}\cos(\omega z) + \frac{1}{\nu}\boldsymbol{p}\sin(\omega z),$$
$$\boldsymbol{p} \to -\nu\boldsymbol{q}\sin(\omega z) + \boldsymbol{p}\cos(\omega z),$$
$$\omega = \frac{\nu}{H}.$$

Correspondingly we find in the (X, Y, Z) coordinates

$$X \to X\cos^2(\omega z) + \frac{1}{\nu}Z\sin(2\omega z) + \frac{1}{\nu^2}Y\sin^2(\omega z),$$
$$Y \to \nu^2 X\sin^2(\omega z) - \nu Z\sin(2\omega z) + Y\cos^2(\omega z),$$
$$Z \to -\frac{\nu}{z}X\sin(2\omega z) + Z\cos(2\omega z) + \frac{1}{2\nu}Y\sin(2\omega z),$$

and again H and S^2 are preserved. These orbits are ellipses, see fig. 5.

Let us analyze the information in fig. 5. The surface of the cone $S^2 = 0$ corresponds to meridional rays, i.e. rays in a plane with the optical axis where $\boldsymbol{q}$ and $\boldsymbol{p}$ are collinear on the screen, so that $\boldsymbol{q} \times \boldsymbol{p} = 0$. In the fiber, the sinusoidal motion in (X, Y, Z) remains in a plane with the optical axis, alternating between maximal elongation and zero momentum (on the X-axis), and zero elongation with maximal optical momentum (on the Y-axis) every quarter-cycle $(\boldsymbol{q} \to \boldsymbol{p})$, $(\boldsymbol{p} \to -\boldsymbol{q})$. As

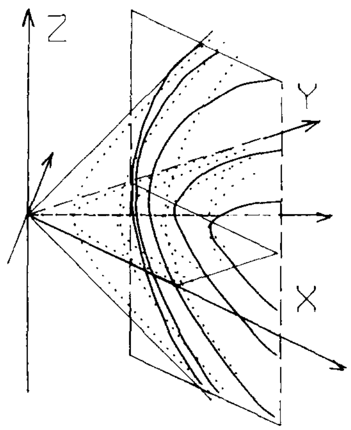

Fig. 5. Elliptical periodic motion in the optical phase space (X, Y, Z) takes place along intersections of a Hamiltonian level surface for an elliptic-profile graded-index fiber, with a level surface of a skewness hyperboloid.

discussed earlier, half-cycles $(\boldsymbol{q} \to -\boldsymbol{q})$, $(\boldsymbol{p} \to -\boldsymbol{p})$ correspond to the same point on the hyperboloid. The size of the elongation is given (and bounded) by $H^2 = n_0^2 - \nu^2 X_{\max} \geq 0$.

Sagittal rays are those rays for which $\boldsymbol{q}$ and $\boldsymbol{p}$ are orthogonal on the screen, $\boldsymbol{p} \cdot \boldsymbol{q} = Z = 0$. They are represented by *fixed* points on the $X = Y$ axis, i.e. as the points of contact between a $H =$ constant plane that is tangent at the tip of the hyperboloid at $X = Y = S$, where $H^2 = n_0^2 - (1 + \nu^2)S$. They appear as fixed points because the ray path in $(\boldsymbol{q}, \boldsymbol{p})$-space is a *circular helix*, and, by construction, the X–Y–Z coordinates identify rays related by a pure rotation around the optical axis. The axial angle $\theta(z)$ advances along these orbits at a constant rate.

Periodic motion on one of the interior hyperboloids cutting a finite circle or ellipse at the $H =$ constant plane, depicts the motion of a skew ray $XY > S^2 > 0$ along an elliptical helical path along the fiber. As $\boldsymbol{p}$ and $\boldsymbol{q}$ grow and wane, and change their *relative* orientation, the point in (X, Y, Z)-space progresses along its periodic orbit. Again the complete solution is reconstructed from a quadrature for $\theta(z)$.

9. General-profile fibers

Consider now a generic fiber in which the refractive index n is a function of $X = q^2$ that is only assumed to be positive and physically meaningful for $n \geq 1$. As we remarked before, the flow will lie on $H =$ constant vertically ruled surfaces with profiles $Y = n(X)^2 - H^2$. On these surfaces we may superpose the lines $Y = (S^2 + Z^2)/X$ stemming from the conservation of skewness, projected onto the $Z = 0$ plane. This is shown in fig. 6a for one S-level surface intersecting one H-level surface within the allowed region of the phase space quadrant, $X > 0, Y > 0, H^2 > 0$. Between the intersection points, shown in fig. 6c, periodic motion will occur.

When periodic motion occurs for some value of S, then also rays of neighboring skewness will perform similar motion. To increase skewness we may take rays with a larger optical momentum (larger angle with the optical axis) moving up fig. 6a, larger elongation (moving to the right), or increasing the angle between $\boldsymbol{p}$ and $\boldsymbol{q}$

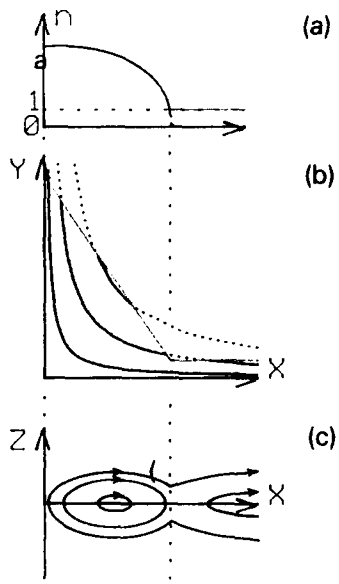

Fig. 6. (a) The refractive index profile of an elliptic index fiber. (b) The intersection of the level curves of S and $H(X, Y) = 0$ on the X–Y plane. (c) The orbit diagram projected on the X–Y plane.

($S = pq \sin\theta$) up to its maximum ($\theta = \frac{1}{2}\pi$, $|S| = \sqrt{XY}$). In the case of an elliptic-profile fiber, the orbit will shrink if the energy is maintained the same because only circular helical paths will be possible: i.e. fixed points.

We have thus far seen elliptic profile fibers, depicted in the X–Y plane as straight lines with negative slope, limiting to a horizontal line when the medium becomes homogeneous. Let us now present an example of a fiber that exhibits *hyperbolic* fixed points.

10. Hyperbolic fixed points and ray trapping

Fixed points of the axisymmetric ray tracing system occur in the $Z = 0$ plane only, at X–Y points for which the gradients $\nabla H^2 = (\mathrm{d}n^2/\mathrm{d}X, -1)$ and $\nabla S^2 = (Y, X)$ are collinear, giving $\nabla S^2 \times \nabla H^2 = 0$, and satisfying the tangency condition

$$Y = -X\frac{\mathrm{d}n^2}{\mathrm{d}X}.$$

These fixed points will be stable (elliptic), provided

$$\frac{\mathrm{d}^2(X^2 n^2)}{\mathrm{d}X^2} < 0;$$

otherwise, they will be unstable (hyperbolic).

As an example of a refractive-index profile whose ray path dynamics has hyperbolic fixed points, consider the parabolic profile

$$n^2(X) = n_0^2 + (n_1 - \nu X)^2,$$

where n_0^2, n_1 and ν are positive constants. In this case, the level surfaces of the Hamiltonian form a family of upright parabolas

$$Y = n_0^2 + (n_1 - \nu X)^2 - H^2,$$

whose minima occur along their common symmetry axis $X = n_1/\nu$. These parabolas intercept the positive Y-axis for $(n_0^2 + n_1^2) \geq H^2 \geq 0$.

The tangency condition for this choice of refractive index is the inverted parabola, independent of n_0^2,

$$Y = 2\nu X(n_1 - \nu X),$$

with maximum at the (X, Y) point $(n_1/2\nu, \frac{1}{2}n_1^2)$. The level surfaces of H^2 satisfy the tangency condition when

$$n_0^2 + (n_1 - \nu X)^2 - H^2 = Y = 2\nu X(n_1 - \nu X).$$

Thus, the tangent points occur in pairs at the roots of this quadratic equation in X,

$$X = \frac{1}{3\nu}\left(2n_1 \pm \sqrt{n_1^2 - 3n_0^2 + H^2}\right),$$

provided $n_0^2 < \frac{1}{3}(n_1^2 + H^2)$. Conversely, for $n_0^2 > \frac{1}{3}(n_1^2 + H^2)$, no tangency occurs in the positive X–Y quadrant. (The value $n_0^2 = \frac{1}{3}(n_1^2 + H^2)$ is a cusp (or "tangent") bifurcation point for the ray path equilibria.)

The stability condition at an equilibrium value of X is

$$\frac{\mathrm{d}^2(X^2 n^2)}{\mathrm{d}X^2} = 12\nu X(\nu X - n_1) + 2\left(n_1^2 + n_0^2\right) < 0.$$

In each pair of roots of the tangency condition, the larger one is unstable, while the smaller one is stable. The unstable equilibrium point at the larger root is connected to itself by a homoclinic orbit lying on one of the S^2-level surfaces and encircling the stable equilibrium at the smaller root. Initial values lying within the homoclinic loop are trapped forever within it on periodic orbits, while those lying initially outside the loop escape to infinity along the S^2 hyperboloid.

For the case of meridional rays ($S^2 = 0$), for example, tangencies occur with values $n_0^2 = 0 = H^2$ at the hyperbolic point $(n_1/\nu, 0)$, and at the elliptic point $(n_1/3\nu, \frac{4}{9}n_1^2)$. The hyperbolic point is connected to itself on the $S^2 = 0$ cone by a homoclinic orbit. The motion along the homoclinic orbit is given by

$$X = r_0^2 \tanh^2(n_1 z/r_0), \quad Y = n_1^2 \operatorname{sech}^4(n_1 z/r_0),$$

$$Z = r_0 n_1 \tanh(n_1 z/r_0) \operatorname{sech}^2(n_1 z/r_0),$$

where $r_0 = \sqrt{n_1/\nu}$.

Two closing remarks on the homoclinic orbits

The discussion we have given for the behavior of light in an optical fiber places ray optics within the modern geometric setting of the Lie–Poisson Hamiltonian systems and provides an example of ray trapping by homoclinic orbits in this setting. Certain perturbations may disrupt the homoclinic trapping. In particular, periodic perturbations in the refractive index along the optical axis may induce sensitivity of the motion to its initial conditions by causing a homoclinic tangle to form, across which phase points sufficiently near the original unperturbed homoclinic orbit may be transported. This phase space transport causes some loss of homoclinic trapping in the region of optical phase space near the original unperturbed homoclinic orbit. This loss of homoclinic trapping appears as a practically unpredictable (chaotic) wandering of the rays, which are caught in the homoclinic tangle produced by the periodic perturbations near the unperturbed homoclinic orbit. Such wandering motions due to periodic perturbations have been recently discussed in a ray-optics context by Holm and Kovačič [2].

In polar coordinates, Hamiltonian ray path dynamics reduces to phase-plane analysis in (r, p_r), but loses the global picture afforded by the $\mathrm{sp}(2,\mathbb{R})^*$ variables in $\mathbb{R}^3$. In particular, the $\mathbb{R}^3$ picture "isolates" the singularity in the phase plane at $r = 0$, by placing it at the vertex of the $p_\theta^2 = 0 = S^2$ cone, whose (X, Y, Z) coordinates are $(0,0,0)$. This conical vertex is a *fixed point* for any medium whose refractive index has finite radial derivative at the optical axis, $r = 0$. The stability of this fixed point determines whether optical ray trapping by separatrices occurs for a given radial refractive index profile. For example, the separatrix orbit will be *homoclinic* to the fixed point at the conical vertex, if one of the level surfaces of the Hamiltonian passes through the vertex of the cone. The nature of the separatrix orbit through the vertex of the $S^2 = 0$ cone is quite interesting. The periodic orbits which do not encircle the vertex will pass through the $Y = 0 = p_r^2$ axis of the cone, but not through the $X = 0 = r^2$ axis. Along these orbits, the canonical momentum p_r changes sign by going smoothly through zero at a finite value of the radius, r. This change of sign is detected in the $\mathbb{R}^3$ coordinates as a change in the sign of Z. Thus, periodic orbits not encircling the vertex are singly sheeted on the $S^2 = 0$ cone. However, along the periodic orbits on the cone which *do* encircle the vertex, the canonical momentum p_r reflects at a *nonzero value* back to $-p_r$ as the orbit crosses the $X = 0$ axis. In order to get back to the original state in the (r, p_r) phase plane, this reflection must occur twice; so these orbits in $\mathbb{R}^3$ are doubly sheeted on the $S^2 = 0$ cone. Thus, a separatrix orbit through the vertex of the $S^2 = 0$ cone separates the singly sheeted periodic orbits which do not pass through the $X = 0$ axis, from the doubly sheeted periodic orbits which do pass through that axis. (The periodic orbits on the hyperboloids within the cone are all singly sheeted; since these orbits never cross the $X = 0$ axis, no ambiguity arises in the sign of p_r for them.)

References

[1] H. Flanders, Differential Forms with Applications in the Physical Sciences (Academic Press, New York, 1963).

[2] D.D. Holm and G. Kovačič, Homoclinic chaos for ray optics in a fiber, Physica D 51 (1991) 177–188, these Proceedings.

[3] D.D. Holm, J.E. Marsden, T. Ratiu and A. Weinstein, Nonlinear stability of fluid and plasma equilibria, Phys. Rep. 123 (1985) 1–116.

[4] S. Lie, Collected Works, Vol. II (Teubner, Leipzig, 1890) p. 237.

[5] J.E. Marsden and T. Ratiu, Mechanics and Symmetry (1990), in preparation.
[6] A.J. Dragt, E. Forest and K.B. Wolf, Foundations of a Lie algebraic theory of geometrical optics, in: Lie Methods in Optics, eds. J.S. Mondragón and K.B. Wolf, Lecture Notes in Physics, Vol. 250 (Springer, New York, 1986).
[7] J.E. Marsden, T. Ratiu and G. Raugel, Symplectic connections and the linearization of Hamiltonian systems, Proc. R. Soc. Edinburgh A (1990), to appear.
[8] A. Weinstein, The local structure of Poisson manifolds, J. Diff. Geom. 18 (1983) 523–557.