1.7.5 Non-canonical $\mathbb{R}^3$ Poisson bracket for ray optics

The canonical Poisson bracket relations in (1.7.5) may be used to transform to another Poisson bracket expressed solely in terms of the variables $\mathbf{X} = (X_1, X_2, X_3) \in \mathbb{R}^3$ by using the chain rule again,

$$\frac{dF}{dt} = \{F, H\} = \frac{\partial F}{\partial X_i}\{X_i,\, X_j\}\frac{\partial H}{\partial X_j}. \tag{1.7.7}$$

Here, the quantities $\{X_i,\, X_j\}$, with $i, j = 1, 2, 3$, are obtained from Poisson commutator table in (1.7.5).

This chain rule calculation reveals that the Poisson bracket in the $\mathbb{R}^3$ variables (X_1, X_2, X_3) repeats the commutator table $[m_i,\, m_j] = c_{ij}^k m_k$ for the Lie algebra $sp(2, \mathbb{R})$ of Hamiltonian matrices in (1.7.4). Consequently, we may write this Poisson bracket equivalently as

$$\{F, H\} = X_k\, c_{ij}^k \frac{\partial F}{\partial X_i}\frac{\partial H}{\partial X_j}\,. \tag{1.7.8}$$

In particular, the Poisson bracket between two of these quadratic-monomial invariants is a linear function of them

$$\{X_i,\, X_j\} = c_{ij}^k X_k\,, \tag{1.7.9}$$

and we also have

$$\{X_l,\, \{X_i,\, X_j\}\} = c_{ij}^k\{X_l,\, X_k\} = c_{ij}^k c_{lk}^m X_m\,. \tag{1.7.10}$$

Hence, the Jacobi identity is satisfied for the Poisson bracket (1.7.7) as a consequence of

$$\begin{aligned}
&\{X_l,\, \{X_i,\, X_j\}\} + \{X_i,\, \{X_j,\, X_l\}\} + \{X_j,\, \{X_l,\, X_i\}\} \\
&= c_{ij}^k\{X_l,\, X_k\} + c_{jl}^k\{X_i,\, X_k\} + c_{li}^k\{X_j,\, X_k\} \\
&= \left(c_{ij}^k c_{lk}^m + c_{jl}^k c_{ik}^m + c_{li}^k c_{jk}^m\right) X_m = 0\,,
\end{aligned}$$

followed by comparison with equation (1.7.3) for the Jacobi identity in terms of the structure constants.

Remark 1.7.6 *This calculation for the Poisson bracket (1.7.8) provides an independent proof that it satisfies the Jacobi identity.*

The chain rule calculation (1.7.7) also reveals the following.

Theorem 1.7.7 *Under the map*

$$T^*\mathbb{R}^2 \to \mathbb{R}^3 : (\mathbf{q},\mathbf{p}) \to \mathbf{X} = (X_1, X_2, X_3)\,, \tag{1.7.11}$$

the Poisson bracket among the axisymmetric optical variables (1.4.5)

$$X_1 = |\mathbf{q}|^2 \geq 0\,, \quad X_2 = |\mathbf{p}|^2 \geq 0\,, \quad X_3 = \mathbf{p}\cdot\mathbf{q}\,,$$

may be expressed for $S^2 = X_1X_2 - X_3^2$ as

$$\begin{aligned} \frac{dF}{dt} = \{F, H\} &= \nabla F \,.\, \nabla S^2 \times \nabla H \\ &= -\frac{\partial S^2}{\partial X_l}\,\epsilon_{ljk}\frac{\partial F}{\partial X_j}\,\frac{\partial H}{\partial X_k}\,. \end{aligned} \tag{1.7.12}$$

Proof. This is a direct verification using formula (1.7.7). For example,

$$2\epsilon_{123}\frac{\partial S^2}{\partial X_3} = -4X_3, \quad 2\epsilon_{132}\frac{\partial S^2}{\partial X_2} = 2X_1, \quad 2\epsilon_{231}\frac{\partial S^2}{\partial X_1} = -2X_2\,.$$

(The inessential factors of 2 may be absorbed into the definition of the independent variable, which here is the time, t.) ■

The standard symbol ϵ_{klm} used in the last relation in (1.7.12) to write the triple scalar product of vectors in index form is defined as follows.

Definition 1.7.8 (Antisymmetric symbol ϵ_{klm})
The symbol ϵ_{klm} with $\epsilon_{123} = 1$ is the totally antisymmetric tensor in three dimensions: it vanishes if any of its indices are repeated and it equals the parity of the permutations of the set $\{1,2,3\}$ when $\{k,l,m\}$ are all different. That is,

$$\epsilon_{kkm} = 0 \text{ (no sum)}$$

and

$$\epsilon_{klm} = +1 \text{ (resp. } -1\text{) for even (resp. odd) permutations of } \{1,2,3\}\,.$$

Remark 1.7.9 *For three-dimensional vectors* $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, *one has*

$$(\mathbf{B}\times\mathbf{C})_l = \epsilon_{klm}B_mC_n \quad \text{and} \quad \big(\mathbf{A}\times(\mathbf{B}\times\mathbf{C})\big)_i = \epsilon_{ijk}A_j\epsilon_{klm}B_lC_m$$

Hence, the relation

$$\epsilon_{ijk}\epsilon_{klm} = \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}$$

verifies the familiar BAC minus CAB rule for the triple vector product. That is,

$$\mathbf{A}\times(\mathbf{B}\times\mathbf{C}) = \mathbf{B}(\mathbf{A}\cdot\mathbf{C}) - \mathbf{C}(\mathbf{A}\cdot\mathbf{B}).$$

Corollary 1.7.10 *The equations of Hamiltonian ray optics in axisymmetric translation-invariant media may be expressed with* $H = H(X_1, X_2)$ *as*

$$\dot{\mathbf{X}} = \nabla S^2 \times \nabla H\,, \quad \text{with} \quad S^2 = X_1X_2 - X_3^2 \geq 0\,. \qquad (1.7.13)$$

Thus, the flow preserves volume (that is, it satisfies $\operatorname{div}\dot{\mathbf{X}} = 0$*) and the evolution along the curve* $\mathbf{X}(z) \in \mathbb{R}^3$ *takes place on intersections of level surfaces of the axisymmetric media invariants* S^2 *and* $H(X_1, X_2)$ *in* $\mathbb{R}^3$.

Remark 1.7.11 *The Petzval invariant* S^2 *satisfies* $\{S^2, H\} = 0$ *with the bracket (1.7.12) for every Hamiltonian* $H(X_1, X_2, X_3)$ *expressed in these variables.*

Definition 1.7.12 (Casimir, or distinguished function) *A function that Poisson-commutes with all other functions on a certain space is the Poisson bracket's* ***Casimir****, or* ***distinguished function****.*

1.8 Equilibrium solutions

1.8.1 Energy-Casimir stability

Remark 1.8.1 (Critical energy plus Casimir equilibria) *A point of tangency of the level sets of Hamiltonian H and Casimir*

S^2 is an equilibrium solution of equation (1.7.13). This is because, at such a point, the gradients of the Hamiltonian H and Casimir S^2 are collinear; so the right-hand side of (1.7.13) vanishes. At such points of tangency, the variation of the sum $H_\Phi = H + \Phi(S^2)$ vanishes, for some smooth function Φ. That is,

$$\begin{aligned}\delta H_\Phi(\mathbf{X}_e) &= DH_\Phi(\mathbf{X}_e)\cdot\delta\mathbf{X} \\ &= \Big[\nabla H + \Phi'(S^2)\nabla S^2\Big]_{\mathbf{X}_e}\cdot\delta\mathbf{X} = 0\,,\end{aligned}$$

when evaluated at equilibrium points $\mathbf{X}_e$ where the level sets of H and S^2 are tangent.

Exercise. Show that a point $\mathbf{X}_e$ at which H_Φ has a critical point (i.e., $\delta H_\Phi = 0$) must be an equilibrium solution of equation (1.7.13). ★

Energy-Casimir stability of equilibria

The second variation of the sum $H_\Phi = H + \Phi(S^2)$ is a quadratic form in $\mathbb{R}^3$ given by

$$\delta^2 H_\Phi(\mathbf{X}_e) = \delta\mathbf{X}\cdot D^2 H_\Phi(\mathbf{X}_e)\cdot\delta\mathbf{X}\,.$$

Thus we have, by Taylor's theorem,

$$H_\Phi(\mathbf{X}_e + \delta\mathbf{X}) - H_\Phi(\mathbf{X}_e) = \frac{1}{2}\delta^2 H_\Phi(\mathbf{X}_e) + o(|\delta\mathbf{X}|^2)\,,$$

when evaluated at the critical point $\mathbf{X}_e$. Remarkably, the quadratic form $\delta^2 H_\Phi(\mathbf{X}_e)$ is the Hamiltonian for the dynamics linearised around the critical point. Consequently, the second variation $\delta^2 H_\Phi$ is preserved by the linearised dynamics in a neighbourhood of the equilibrium point.

Exercise. Linearise the dynamical equation (1.7.13) about an equilibrium $\mathbf{X}_e$ for which the quantity H_Φ has a critical point and show that the linearised dynamics conserves the quadratic form arising from the second variation. Show that the quadratic form is the Hamiltonian for the

linearised dynamics.
What is the corresponding Poisson bracket?
Does this process provides a proper bracket for the linearised dynamics? Prove that it does. ★

The ***signature*** of the second variation provides a method for determining the stability of the critical point. This is the ***energy-Casimir stability method***. This method is based on the following.

Theorem 1.8.2 *A critical point $\mathbf{X}_e$ of $H_\Phi = H + \Phi(S^2)$ whose second variation is definite in sign is a stable equilibrium solution of equation (1.7.13).*

Proof. A critical point $\mathbf{X}_e$ of $H_\Phi = H + \Phi(S^2)$ is an equilibrium solution of equation (1.7.13). Sign definiteness of the second variation provides a norm $\|\delta\mathbf{X}\|^2 = |\delta^2 H_\Phi(\mathbf{X}_e)|$ for the perturbations around the equilibrium $\mathbf{X}_e$ that is conserved by the linearised dynamics. Being conserved by the dynamics linearised around the equilibrium, this sign-definite distance from $\mathbf{X}_e$ must remain constant. Therefore, in this case, the absolute value of sign-definite second variation $|\delta^2 H_\Phi(\mathbf{X}_e)|$ provides a distance from the equilibrium $\|\delta\mathbf{X}\|^2$ which is bounded in time under the linearised dynamics. Hence, the equilibrium solution is stable.

■

Remark 1.8.3 *Even when the second variation is indefinite, it is still linearly conserved. However, an indefinite second variation does not provide a norm for the perturbations. Consequently, an indefinite second variation does not limit the growth of a perturbation away from its equilibrium.*

Definition 1.8.4 (Geometrical nature of equilibria)
An equilibrium whose second variation is sign-definite are called ***elliptic****, because the level sets of the second variation in this case*

make closed, nearly elliptical contours in its Euclidean neighbourhood. Hence, the orbits on these closed level sets remain near the equilibria in the sense of the Euclidean norm on $\mathbb{R}^3$. (In $\mathbb{R}^3$ all norms are equivalent to the Euclidean norm.)

An equilibrium with sign-indefinite second variation is called ***hyperbolic****, because the level sets of the second variation do not close locally in its Euclidean neighbourhood. Hence, in this case, an initial perturbation following a hyperbolic level set of the second variation may move out of the Euclidean neighbourhood of the equilibrium.*

1.9 Momentum maps

1.9.1 The action of $Sp(2,\mathbb{R})$ on $T^*\mathbb{R}^2 \simeq \mathbb{R}^2 \times \mathbb{R}^2$

The Lie group $Sp(2,\mathbb{R})$ of symplectic real matrices $M(s)$ acts diagonally on $\mathbf{z} = (\mathbf{q},\, \mathbf{p})^T \in T^*\mathbb{R}^2$ by matrix multiplication as

$$\mathbf{z}(s) = M(s)\mathbf{z}(0) = \exp(s\xi)\mathbf{z}(0)\,,$$

in which $M(s)JM^T(s) = J$ is a symplectic 2×2 matrix. The 2×2 matrix tangent to the symplectic matrix $M(s)$ at the identity $s = 0$ is given by

$$\xi = \Big[M'(s)M^{-1}(s)\Big]_{s=0}.$$

This is a 2×2 Hamiltonian matrix in $sp(2,\mathbb{R})$, satisfying

$$\xi J + J\xi = 0 \quad \text{so that} \quad J\xi J = \xi\,. \tag{1.9.1}$$

Exercise. Verify (1.9.1), cf. (1.6.8). ★

The vector field $\xi_M(\mathbf{z}) \in T\mathbb{R}^2$ may be expressed as a derivative,

$$\xi_M(\mathbf{z}) = \frac{d}{ds}\big[\exp(s\xi)\mathbf{z}\big]\Big|_{s=0} = \xi\mathbf{z}\,,$$

in which the diagonal action $(\xi\mathbf{z})$ of the Hamiltonian matrix (ξ) and the 2-component real multi-vector $\mathbf{z} = (\mathbf{q},\, \mathbf{p})^T$ has components $(\xi_{kl}q_l, \xi_{kl}p_l)^T$, with $k, l = 1, 2$. The matrix ξ is any linear combination of the traceless constant Hamiltonian matrices (1.6.5).

Definition 1.9.1 (Map $\mathcal{J} : T^*\mathbb{R}^2 \simeq \mathbb{R}^2 \times \mathbb{R}^2 \to sp(2,\mathbb{R})^*$) *The map, $\mathcal{J} : T^*\mathbb{R}^2 \simeq \mathbb{R}^2 \times \mathbb{R}^2 \to sp(2,\mathbb{R})^*$ is defined by*

$$
\begin{aligned}
\mathcal{J}^\xi(\mathbf{z}) &:= \Big\langle \mathcal{J}(\mathbf{z}),\, \xi \Big\rangle_{sp(2,\mathbb{R})^* \times sp(2,\mathbb{R})} \\
&= \Big(\mathbf{z},\, J\xi\mathbf{z} \Big)_{\mathbb{R}^2 \times \mathbb{R}^2} \\
&:= z_k (J\xi)_{kl} z_l \\
&= \mathbf{z}^T \cdot J\xi\mathbf{z} \\
&= \operatorname{tr}\Big((\mathbf{z} \otimes \mathbf{z}^T J)\xi \Big) ,
\end{aligned}
\tag{1.9.2}
$$

where $\mathbf{z} = (\mathbf{q},\, \mathbf{p})^T \in \mathbb{R}^2 \times \mathbb{R}^2$.

Remark 1.9.2 *The map $\mathcal{J}(\mathbf{z})$ given in (1.9.2) by*

$$
\mathcal{J}(\mathbf{z}) = (\mathbf{z} \otimes \mathbf{z}^T J) \in sp(2,\mathbb{R})^* \tag{1.9.3}
$$

sends $\mathbf{z} = (\mathbf{q},\mathbf{p})^T \in \mathbb{R}^2 \times \mathbb{R}^2$ to $\mathcal{J}(\mathbf{z}) = (\mathbf{z} \otimes \mathbf{z}^T J)$, which is an element of $sp(2,\mathbb{R})^*$, the dual space to $sp(2,\mathbb{R})$. Under the pairing $\langle\,\cdot\,,\,\cdot\,\rangle : sp(2,\mathbb{R})^* \times sp(2,\mathbb{R}) \to \mathbb{R}$ given by the trace of the matrix product, one finds the Hamiltonian, or phase space function,

$$
\Big\langle \mathcal{J}(\mathbf{z}),\, \xi \Big\rangle = \operatorname{tr}\left(\mathcal{J}(\mathbf{z})\, \xi \right) , \tag{1.9.4}
$$

for $\mathcal{J}(\mathbf{z}) = (\mathbf{z} \otimes \mathbf{z}^T J) \in sp(2,\mathbb{R})^*$ and $\xi \in sp(2,\mathbb{R})$.

Remark 1.9.3 (Map to axisymmetric invariant variables) *The map, $\mathcal{J} : T^*\mathbb{R}^2 \simeq \mathbb{R}^2 \times \mathbb{R}^2 \to sp(2,\mathbb{R})^*$ in (1.9.2) for $Sp(2,\mathbb{R})$ acting diagonally on $\mathbb{R}^2 \times \mathbb{R}^2$ in equation (1.9.3) may be expressed in matrix form as*

$$
\begin{aligned}
\mathcal{J} &= (\mathbf{z} \otimes \mathbf{z}^T J) \\
&= 2 \begin{pmatrix} \mathbf{p}\cdot\mathbf{q} & -|\mathbf{q}|^2 \\ |\mathbf{p}|^2 & -\mathbf{p}\cdot\mathbf{q} \end{pmatrix} \\
&= 2 \begin{pmatrix} X_3 & -X_1 \\ X_2 & -X_3 \end{pmatrix} .
\end{aligned}
\tag{1.9.5}
$$

This is none other than matrix form of the map (1.7.11) to axisymmetric invariant variables.

$$
T^*\mathbb{R}^2 \to \mathbb{R}^3 : (\mathbf{q},\mathbf{p})^T \to \mathbf{X} = (X_1, X_2, X_3) ,
$$

defined as

$$X_1 = |\mathbf{q}|^2 \geq 0\,, \quad X_2 = |\mathbf{p}|^2 \geq 0\,, \quad X_3 = \mathbf{p}\cdot\mathbf{q}\,. \tag{1.9.6}$$

Applying the momentum map $\mathcal{J}$ to the vector of Hamiltonian matrices $\mathbf{m} = (m_1, m_2, m_3)$ in equation (1.6.5) yields the individual components,

$$\mathcal{J}\cdot\mathbf{m} = 2\mathbf{X} \quad\Longleftrightarrow\quad \mathbf{X} = \frac{1}{2}\, z_k (J\mathbf{m})_{kl} z_l\,. \tag{1.9.7}$$

Thus, the map, $\mathcal{J} : T^*\mathbb{R}^2 \simeq \mathbb{R}^2 \times \mathbb{R}^2 \to sp(2,\mathbb{R})^*$ recovers the components of the vector $\mathbf{X} = (X_1, X_2, X_3)$ that are related to the components of the Petzval invariant by $S^2 = X_1X_2 - X_3^2$.

Exercise. Verify equation (1.9.7) explicitly by computing, for example,

$$\begin{aligned} X_1 &= \frac{1}{2}(\mathbf{q},\mathbf{p})\cdot(Jm_1)\begin{pmatrix}\mathbf{q}\\ \mathbf{p}\end{pmatrix} \\ &= \frac{1}{2}(\mathbf{q},\mathbf{p})\cdot\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}\begin{pmatrix}0 & 0\\ -2 & 0\end{pmatrix}\begin{pmatrix}\mathbf{q}\\ \mathbf{p}\end{pmatrix} \\ &= |\mathbf{q}|^2\,. \end{aligned}$$

★

Remark 1.9.4 (Momentum maps for ray optics)
Our previous discussions have revealed that the axisymmetric variables (X_1, X_2, X_3) in (1.9.6) generate the Lie group of symplectic transformations (1.6.3) as flows of Hamiltonian vector fields. It turns out that this result is connected to the theory of ***momentum maps****. Momentum maps take phase space coordinates $(\mathbf{q},\mathbf{p})$ to the space of Hamiltonians whose flows are canonical transformations of phase space. An example of a momentum map already appeared in Definition 1.3.18.*

The Hamiltonian functions for the one-parameter subgroups of the symplectic group $Sp(2,\mathbb{R})$ in the KAN decomposition (1.6.13) are given by

$$H_K = \frac{1}{2}(X_1 + X_2)\,, \quad H_A = X_3 \quad \text{and} \quad H_N = -X_1\,. \tag{1.9.8}$$

The three phase space functions,

$$H_K = \frac{1}{2}(|\mathbf{q}|^2 + |\mathbf{p}|^2)\,, \quad H_A = \mathbf{q}\cdot\mathbf{p}\,, \quad H_N = -|\mathbf{q}|^2\,, \tag{1.9.9}$$

map the phase space $(\mathbf{q},\mathbf{p})$ to these Hamiltonians whose corresponding Poisson brackets are the Hamiltonian vector fields for the corresponding one-parameter subgroups. These three Hamiltonians and, equally well, any other linear combinations of (X_1, X_2, X_3), arise from a single ***momentum map****, as we shall explain in Section 1.9.2.*

Remark 1.9.5 *Momentum maps are* ***Poisson maps****. That is, they map Poisson brackets on phase space into Poisson brackets on the target space.*

The corresponding Lie algebra product in $sp(2,\mathbb{R})$ was identified using Theorem 1.7.7 with the vector cross product in the space $\mathbb{R}^3$ by using the $\mathbb{R}^3$-bracket. The $\mathbb{R}^3$-brackets among the (X_1, X_2, X_3) closed among themselves. Therefore, as expected, the momentum map was found to be Poisson. In general, when the Poisson bracket relations are all linear, they will be Lie-Poisson brackets, defined below in Section 1.10.1.

1.9.2 Summary: Properties of momentum maps

A momentum map takes phase space coordinates $(\mathbf{q},\mathbf{p})$ to the space of Hamiltonians, whose flows are canonical transformations of phase space. The ingredients of the momentum map are: (i) a representation of the infinitesimal action of the Lie algebra of the transformation group on the coordinate space; and (ii) an appropriate pairing with the conjugate momentum space. For example, one may construct a momentum map by using the familiar pairing $\langle\langle \cdot\,,\,\cdot \rangle\rangle$ between momentum in phase space and the velocity in the tangent space of the configuration manifold that also appears in the Legendre transformation. For this pairing, the momentum map is derived from the cotangent lift of the infinitesimal action $\xi_M(\mathbf{q})$ of the Lie algebra of the transformation group on the configuration manifold to its action on the canonical momentum. In this case, the formula for the momentum map $\mathcal{J}(\mathbf{q},\, \mathbf{p})$ is

$$\mathcal{J}^{\xi}(\mathbf{q},\, \mathbf{p}) = \Big\langle \mathcal{J}(\mathbf{q},\, \mathbf{p}),\, \xi \Big\rangle = \Big\langle\Big\langle \mathbf{p},\, \xi_M(\mathbf{q}) \Big\rangle\Big\rangle\,, \tag{1.9.10}$$

in which the other pairing $\langle \cdot\,,\,\cdot \rangle$ is between the Lie algebra and its dual. This means the momentum map $\mathcal{J}$ for the Hamiltonian $\mathcal{J}^{\xi}$ lives in the dual space of the Lie algebra belonging to the Lie symmetry. The flow of its vector field $X_{\mathcal{J}^{\xi}} = \{\,\cdot\,,\, \mathcal{J}^{\xi}\}$ is the transformation of phase space by the cotangent lift of a Lie group symmetry infinitesimally generated for configuration space by $\xi_M(\mathbf{q})$. The computation of the Lagrange invariant S in (1.3.23) is an example of this type of momentum map.

Not all momentum maps arise as cotangent lifts. Momentum maps may also arise from the infinitesimal action of the Lie algebra on the phase space manifold $\xi_{T^*M}(\mathbf{z})$ with $\mathbf{z} = (\mathbf{q},\, \mathbf{p})$ by using the pairing with the symplectic form. The formula for the momentum map is then

$$\mathcal{J}^{\xi}(\mathbf{z}) = \Big\langle \mathcal{J}(\mathbf{z}),\, \xi \Big\rangle = \Big(\mathbf{z},\, J\xi_{T^*M}(\mathbf{z})\Big)\,, \tag{1.9.11}$$

where J is the symplectic form and $(\cdot\,,\,\cdot)$ is the inner product on phase space $T^*\mathbb{R} \simeq \mathbb{R} \times \mathbb{R}$ for n degrees of freedom. The transformation to axisymmetric variables in (1.9.5) is an example of a momentum map obtained from the symplectic pairing. Both of these approaches are useful and we have seen that both types of momentum maps are summoned when reduction by S^1 axisymmetry is applied in ray optics. The present chapter explores the consequences of S^1 symmetry and the reductions of phase space associated with the momentum maps for this symmetry.

The level sets of the momentum maps provide the geometrical setting for dynamics with symmetry. The components of the momentum map live on the dual of the Lie symmetry algebra, which is a linear space. The level sets of the components of the momentum map provide the natural coordinates for the reduced dynamics. Thus, the motion takes place in a ***reduced space*** whose coordinates are invariant under the original S^1 symmetry. The motion in the reduced space lies on a level set of the momentum map for the S^1 symmetry. It also lies on a level set of the Hamiltonian. Hence, the dynamics in the reduced space of coordinates that are invariant under the S^1 symmetry is confined to the intersections in the reduced space of the level sets of the Hamiltonian and the momentum map associated with that symmetry. Moreover, in most cases, restriction to either level set results in symplectic (canonical) dynamics.

After the solution for this S^1-reduced motion is determined, one must reconstruct the phase associated with the S^1 symmetry, which decouples from the dynamics of the rest of system through the process of reduction. Thus, each point on the manifolds defined by the level sets of the Hamiltonian and the momentum map in the reduced space is associated with an orbit of the phase on S^1. This S^1 phase must be reconstructed from the solution on the reduced space of S^1-invariant functions. The reconstruction of the phase is of interest in its own right, because it contains both geometric and dynamic components, as discussed in Section 1.12.2.

One advantage of this geometric setting is that it readily reveals how ***bifurcations*** arise under changes of parameters in the system, for example, under changes in parameters in the Hamiltonian. In this setting, bifurcations are topological transitions in the intersections of level surfaces of ***orbit manifolds*** of the Hamiltonian and momentum map. The motion proceeds along these intersections in the reduced space whose points are defined by S^1-invariant coordinates. These topological changes in the intersections of the orbit manifolds accompany qualitative changes in the solution behaviour, such as the change of stability of an equilibrium, or the creation or destruction of equilibria. The display of these changes of topology in the reduced space of S^1-invariant functions also allows a visual classification of potential bifurcations. That is, it affords an opportunity to organise the ***choreography of bifurcations*** that are available to the system as its parameters are varied. For an example of this type of geometric bifurcation analysis, see Section 4.5.5.

Remark 1.9.6 *The two results:*
(1) that the action of a Lie group G with Lie algebra $\mathfrak{g}$ on a symplectic manifold P should be accompanied by a momentum map $J : P \to \mathfrak{g}^*$; and
(2) that the orbits of this action are themselves symplectic manifolds, both occur already in [Lie1890]. See [We1983] for an interesting discussion of Lie's contributions to the theory of momentum maps.

The reader should consult [MaRa1994, OrRa2004] for more discussions and additional examples of momentum maps.

1.10 Lie-Poisson brackets

1.10.1 The $\mathbb{R}^3$-bracket for ray optics is Lie-Poisson

The Casimir invariant $S^2 = X_1X_2 - X_3^2$ for the $\mathbb{R}^3$-bracket (1.7.12) is quadratic. In such cases, one may write the Poisson bracket on $\mathbb{R}^3$ in the suggestive form with a pairing $\langle \cdot \, , \cdot \rangle$,

$$\{F, H\} = -\, X_k \, c_{ij}^k \frac{\partial F}{\partial X_i} \frac{\partial H}{\partial X_j} =: -\left\langle \mathbf{X} \, , \left[\frac{\partial F}{\partial \mathbf{X}} \, , \frac{\partial H}{\partial \mathbf{X}} \right] \right\rangle , \qquad (1.10.1)$$

where c_{ij}^k with $i, j, k = 1, 2, 3$ are the structure constants of a three-dimensional Lie algebra operation denoted as $[\,\cdot\, , \cdot\,]$. In the particular case of ray optics, $c_{12}^3 = 4$, $c_{23}^2 = 2$, $c_{31}^1 = 2$ and the rest of the structure constants either vanish, or are obtained from antisymmetry of c_{ij}^k under exchange of any pair of its indices. These values are the structure constants of the 2×2 Hamiltonian matrices (1.6.5), which represent any of the Lie algebras $sp(2, \mathbb{R})$, $so(2, 1)$, $su(1, 1)$, or $sl(2, \mathbb{R})$. Thus, the reduced description of Hamiltonian ray optics in terms of axisymmetric $\mathbb{R}^3$ variables may be defined on the dual space of any of these Lie algebras, say, $sp(2, \mathbb{R})^*$ for definiteness, where duality is defined by pairing $\langle \cdot \, , \cdot \rangle$ in $\mathbb{R}^3$ (contraction of indices). Since $\mathbb{R}^3$ is dual to itself under this pairing, upper and lower indices are equivalent.

Definition 1.10.1 (Lie-Poisson bracket)
*A **Lie-Poisson bracket** is a bracket operation defined as a linear functional of a Lie algebra bracket by a real-valued pairing between a Lie algebra and its dual space.*

Remark 1.10.2 *Equation (1.10.1) defines a Lie-Poisson bracket. Being a linear functional of an operation (the Lie bracket $[\cdot,\cdot]$) which satisfies the Jacobi identity, any Lie-Poisson bracket also satisfies the Jacobi identity.*

1.10.2 Lie-Poisson brackets with quadratic Casimirs

An interesting class of Lie-Poisson brackets emerges from the $\mathbb{R}^3$ Poisson bracket,

$$\{F, H\}_C := -\nabla C \cdot \nabla F \times \nabla H\,, \tag{1.10.2}$$

when its Casimir is the quadratic form on $\mathbb{R}^3$ given by $C = \frac{1}{2}\mathbf{X}^T \cdot \mathsf{K}\mathbf{X}$ associated with the 3×3 symmetric matrix $\mathsf{K}^T = \mathsf{K}$. This bracket may be written equivalently in various notations, including index form, $\mathbb{R}^3$ vector form, and Lie-Poisson form, as

$$\begin{aligned} \{F,\, H\}_{\mathsf{K}} &= -\nabla C \cdot \nabla F \times \nabla H \\ &= -X_l \mathsf{K}^{li} \epsilon_{ijk} \frac{\partial F}{\partial X_j} \frac{\partial H}{\partial X_k} \\ &= -\mathbf{X}\ \mathsf{K}\left(\frac{\partial F}{\partial \mathbf{X}} \times \frac{\partial H}{\partial \mathbf{X}}\right) \\ &=: -\left\langle \mathbf{X},\, \left[\frac{\partial F}{\partial \mathbf{X}},\, \frac{\partial H}{\partial \mathbf{X}}\right]_{\mathsf{K}} \right\rangle . \end{aligned} \tag{1.10.3}$$

Remark 1.10.3 *The triple scalar product of gradients in the $\mathbb{R}^3$-bracket (1.10.2) is the determinant of the Jacobian matrix for the transformation $(X_1, X_2, X_3) \to (C, F, H)$, which is known to satisfy the Jacobi identity. Being a special case, the Poisson bracket $\{F,\, H\}_{\mathsf{K}}$ also satisfies the Jacobi identity.*

In terms of the $\mathbb{R}^3$ components, the Poisson bracket (1.10.3) yields

$$\{X_j,\, X_k\}_{\mathsf{K}} = -X_l \mathsf{K}^{li} \epsilon_{ijk}\,. \tag{1.10.4}$$

The Lie-Poisson form in (1.10.3) associates the $\mathbb{R}^3$ bracket to a Lie algebra with structure constants given in the dual vector basis by

$$[\mathbf{e}_j,\, \mathbf{e}_k]_{\mathsf{K}} = \mathbf{e}_l \mathsf{K}^{li} \epsilon_{ijk} =: \mathbf{e}_l c^{\,l}_{jk}\,. \tag{1.10.5}$$

The Lie group belonging to this Lie algebra is the invariance group of the quadratic Casimir. Namely, it is the transformation group G_{K} with elements $O(s) \in G_{\mathsf{K}}$ with $O(t)|_{t=0} = Id$ whose action from the left on $\mathbb{R}^3$ is given by $\mathbf{X} \to O\mathbf{X}$, such that

$$O^T(t)\mathsf{K}O(t) = \mathsf{K} \tag{1.10.6}$$

or, equivalently,

$$\mathsf{K}^{-1}O^T(t)\mathsf{K} = O^{-1}(t)\,, \tag{1.10.7}$$

for the 3×3 symmetric matrix $\mathsf{K}^T = \mathsf{K}$. A matrix $O(t)$ satisfying (1.10.6) is called a ***quasi-orthogonal matrix*** with respect to K. That is, $O(t)$ is the similarity transformation of an orthogonal matrix by the symmetric matrix K.

These transformations $\mathbf{X} \to O\mathbf{X}$ are *not* orthogonal, unless $\mathsf{K} =$ Id. However, they do form a Lie group under matrix multiplication, since for any two of them O_1 and O_2, we have

$$(O_1O_2)^T\mathsf{K}(O_1O_2) = O_2^T(O_1^T\mathsf{K}O_1)O_2 = O_2^T\mathsf{K}O_2 = \mathsf{K}\,. \tag{1.10.8}$$

The corresponding Lie algebra $\mathfrak{g}_{\mathsf{K}}$ is the derivative of the defining condition of the Lie group (1.10.6), evaluated at the identity. This yields,

$$0 = \big[\dot{O}^T O^{-T}\big]_{t=0}\mathsf{K} + \mathsf{K}\big[O^{-1}\dot{O}\big]_{t=0}\,.$$

Consequently, if $\widehat{X} = [O^{-1}\dot{O}]_{t=0} \in \mathfrak{g}_{\mathsf{K}}$, the quantity $\mathsf{K}\widehat{X}$ is skew. That is,

$$(\mathsf{K}\widehat{X})^T = -\,\mathsf{K}\widehat{X}\,.$$

A vector representation of this skew matrix is provided by the following ***hat map*** from the Lie algebra $\mathfrak{g}_{\mathsf{K}}$ to vectors in $\mathbb{R}^3$,

$$\widehat{\ }:\ \mathfrak{g}_{\mathsf{K}} \to \mathbb{R}^3 \quad \text{defined by} \quad (\mathsf{K}\widehat{X})_{jk} = -\,X_l\mathsf{K}^{li}\epsilon_{ijk}\,. \tag{1.10.9}$$

When K is invertible, the hat map $(\widehat{\cdot})$ in (1.10.9) is a linear isomorphism. For any vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$ with components u^j, v^k, where $j, k = 1, 2, 3$, one computes

$$
\begin{aligned}
u^j (\mathsf{K}\widehat{X})_{jk} v^k &= -\mathbf{X} \cdot \mathsf{K}(\mathbf{u} \times \mathbf{v}) \\
&=: -\mathbf{X} \cdot [\mathbf{u}\,,\,\mathbf{v}]_{\mathsf{K}}\,.
\end{aligned}
$$

This is the Lie-Poisson bracket for the Lie algebra structure represented on $\mathbb{R}^3$ by the vector product,

$$
[\mathbf{u}, \mathbf{v}]_{\mathsf{K}} = \mathsf{K}(\mathbf{u} \times \mathbf{v})\,. \tag{1.10.10}
$$

Thus, the Lie algebra of the Lie group of transformations of $\mathbb{R}^3$ leaving invariant the quadratic form $C = \frac{1}{2}\mathbf{X}^T \cdot \mathsf{K}\mathbf{X}$ may be identified with the cross product of vectors in $\mathbb{R}^3$ by using the K-pairing instead of the usual dot product. For example, in the case of the Petzval invariant we have

$$
S^2 = X_1X_2 - X_3^2 = \mathbf{X} \cdot \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \mathbf{X}\,.
$$

Consequently,

$$
\mathsf{K} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix},
$$

for ray optics, with $\mathbf{X} = (X_1,\, X_2,\, X_3)^T$.

> **Exercise.** Verify that inserting this formula for K into formula (1.10.4) recovers the Lie-Poisson bracket relations (1.7.5) for ray optics (up to an inessential constant). ★

Hence, we have proved the following theorem.

Theorem 1.10.4 *Consider the* $\mathbb{R}^3$ *bracket in equation (1.10.3)*

$$
\{F, H\}_{\mathsf{K}} := -\nabla C_{\mathsf{K}} \cdot \nabla F \times \nabla H \quad \textit{with} \quad C_{\mathsf{K}} = \frac{1}{2}\mathbf{X} \cdot \mathsf{K}\mathbf{X}\,, \tag{1.10.11}
$$

in which $\mathsf{K}^T = \mathsf{K}$ *is a* 3×3 *real symmetric matrix and* $\mathbf{X}\in\mathbb{R}^3$. *The quadratic form* C_K *is the Casimir function for the Lie-Poisson bracket given by*

$$\{F,\,H\}_\mathsf{K} \;=\; -\mathbf{X}\cdot\mathsf{K}\left(\frac{\partial F}{\partial\mathbf{X}}\times\frac{\partial H}{\partial\mathbf{X}}\right), \qquad (1.10.12)$$

$$=: \; -\left\langle \mathbf{X},\,\left[\frac{\partial F}{\partial\mathbf{X}},\,\frac{\partial H}{\partial\mathbf{X}}\right]_\mathsf{K}\right\rangle, \qquad (1.10.13)$$

defined on the dual of the three-dimensional Lie algebra $\mathfrak{g}_\mathsf{K}$, *whose bracket has the following vector product representation for* $\mathbf{u},\,\mathbf{v}\in\mathbb{R}^3$,

$$[\mathbf{u},\mathbf{v}]_\mathsf{K} = \mathsf{K}(\mathbf{u}\times\mathbf{v}). \qquad (1.10.14)$$

This is the Lie algebra bracket for the Lie group G_K *of transformations of* $\mathbb{R}^3$ *given by action from the left* $\mathbf{X}\to O\mathbf{X}$, *such that* $O^T\mathsf{K}O = \mathsf{K}$, *thereby leaving the quadratic form* C_K *invariant.*

Definition 1.10.5 (The ad **and** ad* **operations)**
The adjoint (ad) *and coadjoint* (ad*) *operations are defined for the Lie-Poisson bracket (1.10.13) with quadratic Casimir,* $C_\mathsf{K} = \frac12\mathbf{X}\cdot\mathsf{K}\mathbf{X}$, *as follows.*

$$\langle\mathbf{X},\,[\mathbf{u},\mathbf{v}]_\mathsf{K}\rangle \;=\; \langle\mathbf{X},\,\mathrm{ad}_\mathbf{u}\mathbf{v}\rangle = \langle\mathrm{ad}^*_\mathbf{u}\mathbf{X},\,\mathbf{v}\rangle \qquad (1.10.15)$$

$$=\; \mathsf{K}\mathbf{X}\cdot(\mathbf{u}\times\mathbf{v}) = (\mathsf{K}\mathbf{X}\times\mathbf{u})\cdot\mathbf{v}.$$

Thus, we have explicitly,

$$\mathrm{ad}_\mathbf{u}\mathbf{v} = \mathsf{K}(\mathbf{u}\times\mathbf{v}) \quad \textit{and} \quad \mathrm{ad}^*_\mathbf{u}\mathbf{X} = -\mathbf{u}\times\mathsf{K}\mathbf{X}. \qquad (1.10.16)$$

These definitions of the ad and ad* operations yield the following theorem for the dynamics.

Theorem 1.10.6 (Lie-Poisson dynamics)
The Lie-Poisson dynamics (1.10.12) - (1.10.13) is expressed in terms of the ad *and* ad* *operations by*

$$\frac{dF}{dt} = \{F,\,H\}_\mathsf{K} \;=\; \left\langle \mathbf{X},\,\mathrm{ad}_{\partial H/\partial\mathbf{X}}\,\frac{\partial F}{\partial\mathbf{X}}\right\rangle$$

$$=\; \left\langle \mathrm{ad}^*_{\partial H/\partial\mathbf{X}}\mathbf{X},\,\frac{\partial F}{\partial\mathbf{X}}\right\rangle, \qquad (1.10.17)$$

so that the Lie-Poisson dynamics expresses itself as coadjoint motion,

$$\frac{d\mathbf{X}}{dt} = \{\mathbf{X}, H\}_{\mathsf{K}} = \mathrm{ad}^*_{\partial H/\partial \mathbf{X}}\mathbf{X} = -\frac{\partial H}{\partial \mathbf{X}} \times \mathsf{K}\mathbf{X}\,. \quad (1.10.18)$$

By construction, this equation conserves the quadratic Casimir, $C_{\mathsf{K}} = \frac{1}{2}\mathbf{X}\cdot\mathsf{K}\mathbf{X}$.

Exercise. Write the equations of coadjoint motion (1.10.18) for $\mathsf{K} = \mathrm{diag}(1,1,1)$ and $H = X_1^2 - X_3^2/2$. ★

1.11 Divergenceless vector fields

1.11.1 Jacobi identity

One may verify directly that the $\mathbb{R}^3$ bracket in (1.7.12) and in the class of brackets (1.10.11) does indeed satisfy the defining properties of a Poisson bracket. Clearly, it is a bilinear, skew-symmetric form. To show that it is also a Leibnitz operator that satisfies the Jacobi identity, we identify the bracket in (1.7.12) with the following ***divergenceless vector field*** on $\mathbb{R}^3$ defined by

$$X_H = \{\,\cdot\,, H\} = \nabla S^2 \times \nabla H \cdot \nabla \in \mathfrak{X}\,. \quad (1.11.1)$$

This isomorphism identifies the bracket in (1.11.1) acting on functions on $\mathbb{R}^3$ with the action of the divergenceless vector fields $\mathfrak{X}$. It remains to verify the Jacobi identity explicitly, by using the properties of the ***commutator of divergenceless vector fields***.

Definition 1.11.1 (Jacobi-Lie bracket)
The ***commutator*** *of two divergenceless vector fields* $u, v \in \mathfrak{X}$ *is defined to be*

$$[v, w] = [\mathbf{v}\cdot\nabla,\, \mathbf{w}\cdot\nabla] = \Big((\mathbf{v}\cdot\nabla)\mathbf{w} - (\mathbf{w}\cdot\nabla)\mathbf{v}\Big)\cdot\nabla\,. \quad (1.11.2)$$

The coefficient of the commutator of vector fields is called the ***Jacobi-Lie bracket****. It may be written without risk of confusion in the same notation as*

$$[v,\, w] = (\mathbf{v}\cdot\nabla)\mathbf{w} - (\mathbf{w}\cdot\nabla)\mathbf{v}\,. \tag{1.11.3}$$

In Euclidean vector components, the Jacobi-Lie bracket (1.11.3) is expressed as

$$[v,\, w]_i = w_{i,j}v_j - v_{i,j}w_j\,. \tag{1.11.4}$$

Here, a ***subscript comma denotes partial derivative****, e.g.,* $v_{i,j} = \partial v_i/\partial x_j$ *and one* ***sums repeated indices*** *over their range; for example,* $i, j = 1, 2, 3$*, in three dimensions.*

> **Exercise.** Show that $[v, w]_{i,i} = 0$ for the expression in (1.11.4); so the commutator of two divergenceless vector fields yields another one. ★

Remark 1.11.2 (Interpreting commutators of vector fields) *We may interpret a smooth vector field in* $\mathbb{R}^3$ *as the tangent at the identity* ($\epsilon = 0$) *of a one-parameter flow* ϕ_ϵ *in* $\mathbb{R}^3$ *parameterised by* $\epsilon \in \mathbb{R}$ *and given by integrating*

$$\frac{d}{d\epsilon}\bigg|_{\epsilon=0} = v_i(\mathbf{x})\frac{\partial}{\partial x_i}\,. \tag{1.11.5}$$

The characteristic equations of this flow are

$$\frac{dx_i}{d\epsilon} = v_i(\mathbf{x}(\epsilon))\,, \quad \textit{so that} \quad \frac{dx_i}{d\epsilon}\bigg|_{\epsilon=0} = v_i(\mathbf{x})\,, \quad i = 1, 2, 3\,. \tag{1.11.6}$$

Integration of the characteristic equations (1.11.6) yields the solution for the ***flow*** $\mathbf{x}(\epsilon) = \phi_\epsilon \mathbf{x}$ *of the vector field defined by (1.11.5), whose initial condition starts from* $\mathbf{x} = \mathbf{x}(0)$*. Suppose the characteristic equations for two such flows parameterised by* $(\epsilon, \sigma) \in \mathbb{R}$ *are given respectively by,*

$$\frac{dx_i}{d\epsilon} = v_i(\mathbf{x}(\epsilon)) \quad \textit{and} \quad \frac{dx_i}{d\sigma} = w_i(\mathbf{x}(\sigma))\,.$$

The difference of their cross derivatives evaluated at the identity yields the Jacobi-Lie bracket,

$$
\begin{aligned}
\frac{d}{d\epsilon}\Big|_{\epsilon=0}\frac{dx_i}{d\sigma}\Big|_{\sigma=0} - \frac{d}{d\sigma}\Big|_{\sigma=0}\frac{dx_i}{d\epsilon}\Big|_{\epsilon=0} &= \frac{d}{d\epsilon}w_i(\mathbf{x}(\epsilon))\Big|_{\epsilon=0} - \frac{d}{d\sigma}v_i(\mathbf{x}(\sigma))\Big|_{\sigma=0} \\
&= \frac{\partial w_i}{\partial x_j}\frac{dx_i}{d\epsilon}\Big|_{\epsilon=0} - \frac{\partial v_i}{\partial x_j}\frac{dx_i}{d\sigma}\Big|_{\sigma=0} \\
&= w_{i,j}v_j - v_{i,j}w_j \\
&= [v,w]_i \,.
\end{aligned}
$$

Thus, the Jacobi-Lie bracket of vector fields is the difference between the cross-derivatives with respect to their corresponding characteristic equations, evaluated at the identity. Of course, this difference of cross derivatives would vanish if each derivative were not evaluated ***before*** *taking the next one.*

The composition of Jacobi-Lie brackets for three divergenceless vector fields $u, v, w \in \mathfrak{X}$ has components given by

$$
\begin{aligned}
\big[u,[v,w]\big]_i &= u_k v_j w_{i,kj} + u_k v_{j,k} w_{i,j} - u_k w_{j,k} v_{i,j} \\
& \quad u_k w_j v_{i,jk} - v_j w_{k,j} u_{i,k} + w_j v_{k,j} u_{i,k} \,. \qquad (1.11.7)
\end{aligned}
$$

Equivalently, in vector form,

$$
\begin{aligned}
[u,[v,w]] &= \mathbf{uv} : \nabla\nabla\mathbf{w} + \mathbf{u}\cdot\nabla\mathbf{v}^T\cdot\nabla\mathbf{w}^T - \mathbf{u}\cdot\nabla\mathbf{w}^T\cdot\nabla\mathbf{v}^T \\
& \quad - \mathbf{uw} : \nabla\nabla\mathbf{v} - \mathbf{v}\cdot\nabla\mathbf{w}^T\cdot\nabla\mathbf{u}^T + \mathbf{w}\cdot\nabla\mathbf{v}^T\cdot\nabla\mathbf{u}^T \,.
\end{aligned}
$$

Theorem 1.11.3 *The Jacobi-Lie bracket of divergenceless vector fields satisfies the* ***Jacobi identity****,*

$$
[u,[v,w]] + [v,[w,u]] + [w,[u,v]] = 0 \,. \qquad (1.11.8)
$$

Proof. Direct verification using (1.11.7) and summing over cyclic permutations. ■

Exercise. Prove Theorem 1.11.3 in streamlined notation obtained by writing

$$
[v,\, w] = v(w) - w(v)\,,
$$

and using bilinearity of the Jacobi-Lie bracket. ★

Lemma 1.11.4 *The $\mathbb{R}^3$-bracket (1.7.12) may be identified with the divergenceless vector fields in (1.11.1) by*

$$[X_G, X_H] = -X_{\{G,H\}}\,, \tag{1.11.9}$$

where $[X_G, X_H]$ is the Jacobi-Lie bracket of vector fields X_G and X_H.

Proof. Equation (1.11.9) may be verified by a direct calculation,

$$\begin{aligned} [X_G, X_H] &= X_G X_H - X_H X_G \\ &= \{G,\cdot\}\{H,\cdot\} - \{H,\cdot\}\{G,\cdot\} \\ &= \{G,\{H,\cdot\}\} - \{H,\{G,\cdot\}\} \\ &= \{\{G,H\},\cdot\} = -X_{\{G,H\}}\,. \end{aligned}$$

■

Remark 1.11.5 *The last step in the proof of Lemma 1.11.4 uses the Jacobi identity for the class of $\mathbb{R}^3$-brackets in equation (1.10.2).*

1.11.2 Geometric forms of Poisson brackets

Determinant & wedge-product forms of the canonical bracket

For one degree of freedom, the canonical Poisson bracket $\{F, H\}$ is the same as the determinant for a change of variables $(q, p) \to (F(q,p), H(q,p))$,

$$\{F, H\} = \frac{\partial F}{\partial q}\frac{\partial H}{\partial p} - \frac{\partial H}{\partial q}\frac{\partial F}{\partial p} = \det\frac{\partial(F,H)}{\partial(q,p)}\,. \tag{1.11.10}$$

This may be written in terms of the differentials of the functions $(F(q,p), H(q,p))$ defined as

$$dF = \frac{\partial F}{\partial q}\,dq + \frac{\partial F}{\partial p}\,dp \quad\text{and}\quad dH = \frac{\partial H}{\partial q}\,dq + \frac{\partial H}{\partial p}\,dp\,, \tag{1.11.11}$$

by writing the canonical Poisson bracket $\{F, H\}$ as a phase space density

$$dF \wedge dH = \det\frac{\partial(F,H)}{\partial(q,p)}dq \wedge dp = \{F, H\}dq \wedge dp\,. \tag{1.11.12}$$

Here the wedge product $\wedge$ in $dF \wedge dH = -dH \wedge dF$ is introduced to impose the antisymmetry of the Jacobian determinant under interchange of its columns.

Definition 1.11.6 (Wedge product of differentials)
The wedge product of differentials (dF, dG, dH) of any smooth functions (F, G, H) is defined by its following three properties.

(i) $\wedge$ is anticommutative: $dF \wedge dG = -dG \wedge dF$;

(ii) $\wedge$ is bilinear: $(adF + bdG) \wedge dH = a(dF \wedge dH) + b(dG \wedge dH)$;

(iii) $\wedge$ is associative: $dF \wedge (dG \wedge dH) = (dF \wedge dG) \wedge dH$.

Remark 1.11.7 *These are the usual properties of area elements and volume elements in integral calculus. These properties also apply in computing changes of variables.*

> **Exercise.** Verify formula (1.11.12) from equation (1.11.11) and the linearity and antisymmetry of the wedge product, so that, e.g., $dq \wedge dp = -dp \wedge dq$ and $dq \wedge dq = 0$.
>
> ★

Determinant & wedge-product forms of the $\mathbb{R}^3$-bracket

The $\mathbb{R}^3$-bracket in equation (1.7.12) may also be rewritten equivalently as a Jacobian determinant, namely,

$$\{F,\, H\} = -\nabla S^2 \cdot \nabla F \times \nabla H = -\,\frac{\partial(S^2, F, H)}{\partial(X_1, X_2, X_3)}\,, \qquad (1.11.13)$$

where

$$\frac{\partial(F_1, F_2, F_3)}{\partial(X_1, X_2, X_3)} = \det\left(\frac{\partial \mathbf{F}}{\partial \mathbf{X}}\right). \qquad (1.11.14)$$

The determinant in three dimensions may be defined using the antisymmetric tensor symbol ϵ_{ijk} as

$$\epsilon_{ijk} \det\left(\frac{\partial \mathbf{F}}{\partial \mathbf{X}}\right) = \epsilon_{abc}\, \frac{\partial F_a}{\partial X_i} \frac{\partial F_b}{\partial X_j} \frac{\partial F_c}{\partial X_k}, \qquad (1.11.15)$$

where, as mentioned earlier, we sum on repeated indices over their range. We shall keep track of the antisymmetry of the determinant in three dimensions by using the ***wedge product*** ($\wedge$)

$$\det\left(\frac{\partial \mathbf{F}}{\partial \mathbf{X}}\right) dX_1 \wedge dX_2 \wedge dX_3 = dF_1 \wedge dF_2 \wedge dF_3 \,. \qquad (1.11.16)$$

Thus, the $\mathbb{R}^3$-bracket in equation (1.7.12) may be rewritten equivalently in wedge-product form as

$$\begin{aligned} \{F,\, H\}\, dX_1 \wedge dX_2 \wedge dX_3 &= -(\nabla S^2 \cdot \nabla F \times \nabla H)\, dX_1 \wedge dX_2 \wedge dX_3 \\ &= -dS^2 \wedge dF \wedge dH . \end{aligned}$$

Poisson brackets of this type are called ***Nambu brackets***, since [Na1973] introduced them in three dimensions. They can be generalised to any dimension, but this requires additional compatibility conditions [Ta1994].

1.11.3 Nambu brackets

Theorem 1.11.8 (Nambu brackets [Na1973])
*For any smooth functions $F\,H \in \mathcal{F}(\mathbb{R}^3)$ of coordinates $\mathbf{X} \in \mathbb{R}^3$ with volume element d^3X, the **Nambu bracket***

$$\{F,\, H\} : \mathcal{F}(\mathbb{R}^3) \times \mathcal{F}(\mathbb{R}^3) \to \mathcal{F}(\mathbb{R}^3)$$

defined by

$$\begin{aligned} \{F,\, H\}\, d^3X &= -\nabla C \cdot \nabla F \times \nabla H d^3X \\ &= -dC \wedge dF \wedge dH \,, \end{aligned} \qquad (1.11.17)$$

possesses the properties (1.3.4) required of a Poisson bracket for any choice of distinguished smooth function $C : \mathbb{R}^3 \to \mathbb{R}$.

Proof. The bilinear skew-symmetric Nambu $\mathbb{R}^3$ bracket yields the divergenceless vector field

$$X_H = \{\,\cdot\,,\, h\} = \nabla C \times \nabla H \cdot \nabla \,,$$

in which

$$\operatorname{div}(\nabla C \times \nabla H) = 0 \,.$$

Divergenceless vector fields are derivative operators that satisfy the Leibnitz product rule and the Jacobi identity. These properties hold in this case for any choice of smooth functions $C, H \in \mathcal{F}(\mathbb{R}^3)$. The other two properties – bilinearity and skew symmetry – hold as properties of the wedge product. Hence, the Nambu $\mathbb{R}^3$ bracket in (1.11.17) satisfies all the properties required of a Poisson bracket specified in Definition 1.3.4. ∎

1.12 Geometry of solution behaviour

1.12.1 Restricting axisymmetric ray optics to level sets

Having realised that the $\mathbb{R}^3$-bracket in equation (1.7.12) is associated to Jacobian determinants for changes of variables, it is natural to transform the dynamics of the axisymmetric optical variables (1.4.5) from three dimensions $(X_1, X_2, X_3) \in \mathbb{R}^3$ to one of its level sets $S^2 > 0$. For convenience, we first make a linear change of Cartesian coordinates in $\mathbb{R}^3$ that explicitly displays the axisymmetry of the level sets of S^2 under rotations, namely,

$$S^2 = X_1X_2 - X_3^2 = Y_1^2 - Y_2^2 - Y_3^2 \,, \tag{1.12.1}$$

with

$$Y_1 = \frac{1}{2}(X_1 + X_2) \,, \quad Y_2 = \frac{1}{2}(X_2 - X_1) \,, \quad Y_3 = X_3 \,.$$

In these new Cartesian coordinates $(Y_1, Y_2, Y_3) \in \mathbb{R}^3$, the level sets of S^2 are manifestly invariant under rotations about the Y_1-axis.

> **Exercise.** Show that this linear change of Cartesian coordinates preserves the orientation of volume elements, but scales them by a constant factor of one-half. That is, show
>
> $$\{F,\, H\}\, dY_1 \wedge dY_2 \wedge dY_3 = \frac{1}{2}\, \{F,\, H\}\, dX_1 \wedge dX_2 \wedge dX_3 \,.$$
>
> The overall constant factor of one-half here is unimportant, because it may be simply absorbed into the units of axial distance in the dynamics induced by the $\mathbb{R}^3$-bracket for axisymmetric ray optics in the Y-variables. ★

Each of the family of hyperboloids of revolution in (1.12.1) comprises a layer in the "hyperbolic onion" preserved by axisymmetric ray optics. We use hyperbolic polar coordinates on these layers in analogy to spherical coordinates,

$$Y_1 = S\cosh u\,, \quad Y_2 = S\sinh u\cos\psi\,, \quad Y_3 = S\sinh u\sin\psi\,. \tag{1.12.2}$$

The $\mathbb{R}^3$-bracket (1.7.12) thereby transforms into hyperbolic coordinates (1.12.2) as

$$\{F,\,H\}\,dY_1\wedge dY_2\wedge dY_3 = -\{F,\,H\}_{hyperb}\,S^2\,dS\wedge d\psi\wedge d\cosh u\,. \tag{1.12.3}$$

Note that the oriented quantity

$$S^2 d\cosh u\wedge d\psi = -S^2 d\psi\wedge d\cosh u\,,$$

is the ***area element on the hyperboloid*** corresponding to the constant S^2.

On a constant level surface of S^2 the function $\{F,H\}_{hyperb}$ only depends on $(\cosh u,\psi)$ so the Poisson bracket for optical motion on any *particular* hyperboloid is then

$$\begin{aligned}\{F,\,H\}\,d^3Y &= -\,S^2 dS\wedge dF\wedge dH \qquad (1.12.4)\\ &= -\,S^2 dS\wedge\{F,\,H\}_{hyperb}\,d\psi\wedge d\cosh u\,,\end{aligned}$$

with

$$\{F,\,H\}_{hyperb} = \left(\frac{\partial F}{\partial\psi}\frac{\partial H}{\partial\cosh u} - \frac{\partial H}{\partial\cosh u}\frac{\partial F}{\partial\psi}\right).$$

Being a constant of the motion, the value of S^2 may be absorbed by a choice of units for any given initial condition and the Poisson bracket for the optical motion thereby becomes ***canonical on each hyperboloid***,

$$\begin{aligned}\frac{d\psi}{dz} &= \{\psi,\,H\}_{hyperb} = \frac{\partial H}{\partial\cosh u}\,, \qquad (1.12.5)\\ \frac{d\cosh u}{dz} &= \{\cosh u,\,H\}_{hyperb} = -\,\frac{\partial H}{\partial\psi}\,. \qquad (1.12.6)\end{aligned}$$

In the Cartesian variables $(Y_1, Y_2, Y_3) \in \mathbb{R}^3$, one has $\cosh u = Y_1/S$ and $\psi = \tan^{-1}(Y_3/Y_2)$. In the original variables,

$$\cosh u = \frac{X_1 + X_2}{2S} \quad \text{and} \quad \psi = \tan^{-1} \frac{2X_3}{X_2 - X_1}.$$

Example 1.12.1 *For a paraxial harmonic guide, whose Hamiltonian is,*

$$H = \frac{1}{2}(|\mathbf{p}|^2 + |\mathbf{q}|^2) = \frac{1}{2}(X_1 + X_2) = Y_1\,, \tag{1.12.7}$$

the ray paths consist of circles cut by the intersections of level sets of the planes $Y_1 = const$ with the hyperboloids of revolution about the Y_1-axis, given by $S^2 = const$.

The dynamics for $\mathbf{Y} \in \mathbb{R}^3$ is given by

$$\dot{\mathbf{Y}} = \{\mathbf{Y}, H\} = \nabla_{\mathbf{Y}} S^2 \times \hat{\mathbf{Y}}_1 = 2\hat{\mathbf{Y}}_1 \times \mathbf{Y}\,, \tag{1.12.8}$$

on using the (1.12.1) to transform the $\mathbb{R}^3$ bracket in (1.7.12). Thus, for the paraxial harmonic guide, the rays spiral down the optical axis following circular helices whose radius is determined by their initial conditions.

Exercise. Verify that equation (1.12.3) transforms the $\mathbb{R}^3$-bracket from Cartesian to hyperboloidal coordinates, by using the definitions in equations (1.12.2). ★

Exercise. Reduce $\{F, H\}_{hyperb}$ to the conical level set $S = 0$. ★

Exercise. Reduce the $\mathbb{R}^3$ dynamics of (1.7.12) to level sets of the Hamiltonian

$$H = aX_1 + bX_2 + cX_3\,,$$

for constants (a, b, c). Explain how this reduction simplifies the equations of motion. ★

1.12.2 Geometric phase on level sets of $S^2 = p_\phi^2$

In polar coordinates, the axisymmetric invariants are

$$\begin{aligned} Y_1 &= \frac{1}{2}\Big(p_r^2 + p_\phi^2/r^2 + r^2\Big), \\ Y_2 &= \frac{1}{2}\Big(p_r^2 + p_\phi^2/r^2 - r^2\Big), \\ Y_3 &= rp_r\,. \end{aligned}$$

Hence, the corresponding volume elements are found to be

$$\begin{aligned} d^3Y &=: dY_1 \wedge dY_2 \wedge dY_3 \\ &= d\frac{S^3}{3} \wedge d\cosh u \wedge d\psi \\ &= dp_\phi^2 \wedge dp_r \wedge dr \end{aligned} \tag{1.12.9}$$

On a level set of $S = p_\phi$ this implies

$$S\,d\cosh u \wedge d\psi = 2\,dp_r \wedge dr\,, \tag{1.12.10}$$

so the transformation of variables $(\cosh u, \psi) \to (p_r, r)$ is ***canonical*** on level sets of $S = p_\phi$.

One recalls Stokes Theorem on phase space

$$\iint_A dp_j \wedge dq_j = \oint_{\partial A} p_j dq_j\,, \tag{1.12.11}$$

where the boundary of the phase space area ∂A is taken around a loop on a closed orbit. Either in polar coordinates or on an invariant hyperboloid $S = p_\phi$ this loop integral becomes

$$\begin{aligned} \oint \mathbf{p}\cdot d\mathbf{q} := \oint p_j dq_j &= \oint \Big(p_\phi d\phi + p_r dr\Big) \\ &= \oint \Big(\frac{S^3}{3}d\phi + \cosh u\, d\psi\Big)\,. \end{aligned}$$

Thus we may compute the total phase change around a closed periodic orbit on the level set of hyperboloid S from

$$\begin{aligned}\oint \frac{S^3}{3} d\phi &= \frac{S^3}{3}\Delta\phi \\ &= \underbrace{-\oint \cosh u \, d\psi}_{\text{Geometric } \Delta\phi} + \underbrace{\oint \mathbf{p}\cdot d\mathbf{q}}_{\text{Dynamic } \Delta\phi} . \end{aligned} \qquad (1.12.12)$$

Evidently, one may denote the total change in phase as the sum

$$\Delta\phi = \Delta\phi_{geom} + \Delta\phi_{dyn} \, ,$$

by identifying the corresponding terms in the previous formula. By the Stokes theorem (1.12.11), one sees that the geometric phase associated with a periodic motion on a particular hyperboloid is given by the hyperbolic solid angle enclosed by the orbit, times a constant factor depending on the level set value $S = p_\phi$. Thus, the name: ***geometric phase***.

1.13 Singular ray optics in anisotropic media

> Every ray of light has therefore two opposite sides. . . . And since the crystal by this disposition or virtue does not act upon the rays except when one of their sides of unusual refraction looks toward that coast, this argues a virtue or disposition in those sides of the rays which answers to and sympathises with that virtue or disposition of the crystal, as the poles of two magnets answer to one another. . . .
>
> —Newton, *Optiks* 1704

Some media have directional properties that are exhibited by differences in the transmission of light in different directions. This effect is seen, for example, in certain crystals. Fermat's principle for such media still conceives light rays as lines in space (i.e., no polarisation vectors, yet), but the refractive index along the paths of the rays in the medium is allowed to depend on both position and *direction.* In this case, Theorem 1.1.1 adapts easily to yield the expected 3D eikonal equation (1.1.9). However, in general, the Lagrangian in such a description is singular, as we shall explain. The Euler-Lagrange equation,

$$\frac{d}{ds}\left(\frac{\partial L}{\partial \dot{\mathbf{r}}(s)}\right) = \frac{\partial L}{\partial \mathbf{r}(s)}, \tag{1.13.1}$$

follows from the variational principle,

$$0 = \delta S = \delta \int_A^B L(\mathbf{r}(s), \dot{\mathbf{r}}(s))ds\,. \tag{1.13.2}$$

The Lagrangian in the case of an anisotropic optical medium is given by

$$L(\mathbf{r}(s), \dot{\mathbf{r}}(s)) = n(\mathbf{r}(s), \dot{\mathbf{r}}(s))\,|\dot{\mathbf{r}}(s)|\,. \tag{1.13.3}$$

Here the refractive index is modelled as a function both of position along the ray $\mathbf{r}(s)$ and the ray direction $\dot{\mathbf{r}}(s)$, which is a unit vector. The latter is defined when s is the arclength as

$$\widehat{\mathbf{s}} = \dot{\mathbf{r}}(s)/|\dot{\mathbf{r}}(s)| \quad \text{with} \quad |\dot{\mathbf{r}}| = 1\,. \tag{1.13.4}$$

Exercise. Show that the variation of the ray direction in (1.13.4) is related to the variation of the path $\delta\mathbf{r}(s)$ by

$$\delta\,\widehat{\mathbf{s}} = \frac{-1}{|\dot{\mathbf{r}}|}\,\widehat{\mathbf{s}} \times \Big(\widehat{\mathbf{s}} \times \delta\dot{\mathbf{r}}(s)\Big)\,.$$

★

The variational principle (1.13.2) with optical Lagrangian (1.13.3) implies the following ***3D eikonal equation*** for the vector $\mathbf{r}(s)$ defining the ray path,

$$\frac{d}{ds}\bigg(n(\mathbf{r},\widehat{\mathbf{s}})\widehat{\mathbf{s}} + \mathbf{A}(\mathbf{r},\widehat{\mathbf{s}})\bigg) = \frac{\partial n(\mathbf{r},\widehat{\mathbf{s}})}{\partial \mathbf{r}}\,. \tag{1.13.5}$$

Here, the ***anisotropy vector*** $\mathbf{A}(\mathbf{r},\widehat{\mathbf{s}})$ is defined as

$$\mathbf{A} := \frac{\partial n}{\partial \dot{\mathbf{r}}}\bigg|_{|\dot{\mathbf{r}}|=1} = -\,\widehat{\mathbf{s}} \times \left(\widehat{\mathbf{s}} \times \frac{\partial n}{\partial \widehat{\mathbf{s}}}\right). \tag{1.13.6}$$

The anisotropy vector $\mathbf{A}$ is the projection of the vector $\partial n/\partial\dot{\mathbf{r}}$ onto the plane that is normal to $\dot{\mathbf{r}}$ and tangent to the direction sphere $|\dot{\mathbf{r}}| = 1$.

In the $\dot{\mathbf{r}}$ notation, the ***3D optical momentum*** is defined as

$$\mathbf{p}_3 := \frac{\partial L}{\partial\dot{\mathbf{r}}(s)} = n(\mathbf{r},\dot{\mathbf{r}})\dot{\mathbf{r}} + \mathbf{A}(\mathbf{r},\dot{\mathbf{r}})\,. \tag{1.13.7}$$

Thus, the 3D optical momentum $\mathbf{p}_3$ lies in the plane spanned by the vectors $\dot{\mathbf{r}}$ and $\partial n/\partial\dot{\mathbf{r}}$, and these two vectors are orthogonal because of the constraint $|\dot{\mathbf{r}}| = 1$. The optical momentum is related to the tangent vector $\dot{\mathbf{r}}(s)$ along the ray path $\mathbf{r}(s)$ by

$$n(\mathbf{r},\dot{\mathbf{r}})\dot{\mathbf{r}} = \mathbf{p}_3 - \mathbf{A}(\mathbf{r},\dot{\mathbf{r}})\,, \tag{1.13.8}$$

whose norm is

$$|\mathbf{p}_3 - \mathbf{A}(\mathbf{r},\dot{\mathbf{r}})| = n(\mathbf{r},\dot{\mathbf{r}}) \quad \text{since} \quad \dot{\mathbf{r}}\cdot\mathbf{A} = 0\,. \tag{1.13.9}$$

Remark 1.13.1 *The anisotropy vector is orthogonal to the desired ray direction and is a prescribed function of it and the position along the ray path.*

Unfortunately, it is not possible to solve for the ray direction $\dot{\mathbf{r}}$, given the 3D optical momentum $\mathbf{p}_3$ and position $\mathbf{r}$. The 3D optical momentum decomposes conveniently into components which are parallel and perpendicular to $\dot{\mathbf{r}}$, as

$$\mathbf{p}_3 = n(\mathbf{r},\dot{\mathbf{r}})\dot{\mathbf{r}} + \mathbf{A}(\mathbf{r},\dot{\mathbf{r}}) =: \mathbf{p}_{3\parallel} + \mathbf{p}_{3\perp}\,.$$

However, media for which these functional relations are nontrivial do not in general admit solutions for the tangent vector $\dot{\mathbf{r}}(s)$ as a function of $(\mathbf{r}(s), \mathbf{p}_3(s))$. Thus, the ray direction is not solvable in general from the optical momentum and ray path.[5] However, the 3D eikonal equation (1.13.5) still holds and so does its associated ***anisotropic Huygens wave front description***,

$$\frac{\partial S(\mathbf{r},\dot{\mathbf{r}})}{\partial\mathbf{r}} = n(\mathbf{r},\dot{\mathbf{r}})\dot{\mathbf{r}} + \mathbf{A}(\mathbf{r},\dot{\mathbf{r}})\,,$$

[5] This is an example of a singular, or degenerate, Lagrangian.

whose norm yields the ***scalar Huygens equation for anisotropic media***,

$$\left|\frac{\partial S}{\partial \mathbf{r}}\right|^2 = n^2(\mathbf{r},\dot{\mathbf{r}}\,) + \left|\mathbf{A}(\mathbf{r},\dot{\mathbf{r}})\right|^2 .$$

Remark 1.13.2 (Ibn Sahl-Snell law for anisotropic media) *The statement of the Ibn Sahl-Snell law relation at discontinuities of the refractive index in anisotropic media is rather more involved than for isotropic media. A break in the direction* $\widehat{\mathbf{s}}$ *of the ray vector is still expected at any finite discontinuity in the refractive index* $n = |\mathbf{n}|$ *encountered along the ray path* $\mathbf{r}(s)$. *According to the eikonal equation for anisotropic media (1.13.5) the jump (denoted by* Δ*) in 3D optical momentum across the discontinuity must satisfy the relation*

$$\Delta\mathbf{p}_3 \times \frac{\partial n}{\partial \mathbf{r}} = \Delta\bigg(n(\mathbf{r},\widehat{\mathbf{s}})\widehat{\mathbf{s}} + \mathbf{A}(\mathbf{r},\widehat{\mathbf{s}})\bigg) \times \frac{\partial n}{\partial \mathbf{r}} = 0\,. \qquad (1.13.10)$$

This means the 2D projections of the 3D optical momenta $\mathbf{p}_3$ *and* $\mathbf{p}_3'$ *onto the plane of the discontinuity in refractive index will be invariant across the interface.*

Thus, preservation of the components of 3D optical momentum tangential to the discontinuity still holds, but a difficulty occurs because the ray direction and optical momentum are no longer colinear. Instead, they differ by the anisotropy vector, which is orthogonal to the desired ray direction and also depends as a prescribed function of ray direction on either side of the discontinuity.

The geometry for determining the refracted ray direction in an anisotropic medium thus becomes considerably more involved than the simple Ibn Sahl-Snell law of ray projection for isotropic media. There does exist a graphical construction (see, e.g., [Wo2004]), but its application in the Ibn Sahl-Snell law for construction of the break in ray direction at a discontinuity in refractive index in an anisotropic medium is problematic, unless the prescribed dependence of the anisotropy vector on the ray direction is rather simple.

An alternative argument

The loop integral argument in equations (1.1.16) - (1.1.18) reaches the same conclusion about the difficulty in determining the ray direc-

tions in general at an interface where the refractive index is discontinuous in an anisotopic medium. This argument proceeds by evaluating the loop integral of the Huygens phase,

$$\oint_P \nabla S(\mathbf{r}) \cdot d\mathbf{r} = \oint_P \Big(\mathbf{n}(\mathbf{r}) + \mathbf{A}(\mathbf{r}, \dot{\mathbf{r}}) \Big) \cdot d\mathbf{r} = 0 \,, \qquad (1.13.11)$$

taken around a closed path P that surrounds a boundary separating two different media. Letting the sides of the loop perpendicular to the interface shrink to zero implies that the tangential components of the momentum vectors must be preserved, in agreement with the previous argument. Consequently,

$$\Big((\mathbf{n}(\mathbf{r}) + \mathbf{A}(\mathbf{r}, \dot{\mathbf{r}})) - (\mathbf{n}'(\mathbf{r}) + \mathbf{A}'(\mathbf{r}, \dot{\mathbf{r}})) \Big) \times \hat{\mathbf{z}} = 0 \,, \qquad (1.13.12)$$

in agreement with equation (1.13.10). If ψ and ψ' are the angles of incident and transmitted ***momentum directions****, measured from the normal $\hat{\mathbf{z}}$ through the boundary, then preservation of the tangential components of the momentum vector means that the momentum vectors must lie in the same plane and the angles ψ and ψ' must satisfy*

$$\sqrt{n^2 + A^2} \sin \psi = \sqrt{(n')^2 + (A')^2} \sin \psi' \qquad (1.13.13)$$

Relation (1.13.13) determines the angles of incidence and transmission of the momentum directions. However, in general, it does not determine the ray directions. The ray directions are not invertible from the co-planar momentum directions, because the anisotropy vector in equation (1.13.7) shifts the ray vectors into different planes by an amount depending on the ray direction itself, not the momentum direction.

1.14 Ten geometrical features of ray optics

1. The design of axisymmetric planar optical systems reduces to multiplication of ***symplectic matrices*** corresponding to each element of the system, see Theorem 1.6.4.

2. Hamiltonian evolution occurs by ***canonical transformations***. Such transformations may be obtained by integrating the characteristic equations of Hamiltonian vector fields, which are defined by Poisson-bracket operations with smooth functions on phase space, as in the proof of Theorem 1.5.3.

3. The Poisson bracket is associated geometrically with the Jacobian for canonical transformations in Section 1.11.2. Canonical transformations are generated by Poisson-bracket operations and these transformations preserve the Jacobian.

4. A one-parameter symmetry, that is, an invariance under canonical transformations generated by a Hamiltonian vector field $X_{p_\phi} = \{\,\cdot\,, p_\phi\}$, separates out an angle, ϕ, whose canonically conjugate momentum p_ϕ is conserved. As discussed in Section 1.3.2, the conserved quantity p_ϕ may be an important bifurcation parameter for the remaining reduced system. The dynamics of the angle ϕ decouples from the reduced system and can be determined as a quadrature after solving the reduced system.

5. Given a symmetry of the Hamiltonian, it may be wise to transform from phase space coordinates to invariant variables as in (1.4.5). This transformation defines the quotient map, which quotients out the angle(s) conjugate to the symmetry generator. The image of the quotient map produces the orbit manifold, a reduced manifold whose points are orbits under the symmetry transformation. The corresponding transformation of the Poisson bracket is done using the chain rule as in (1.7.6). Closure of the Poisson brackets of the invariant variables amongst themselves is a necessary condition for the quotient map to be a momentum map, as discussed in Section 1.9.2.

6. Closure of the Poisson brackets among an odd number of invariant variables is no cause for regret. It only means that this Poisson bracket among the invariant variables is not canonical (symplectic). For example, the Nambu $\mathbb{R}^3$ bracket (1.11.17) arises this way.

7. The bracket resulting from transforming to invariant variables could also be Lie-Poisson. This will happen when the new invariant variables are quadratic in the phase space variables, as occurs for the Poisson brackets among the axisymmetric variables X_1, X_2 and X_3 in (1.4.5). Then the Poisson brackets among them are *linear* in the new variables with constant coefficients. Those constant coefficients are dual to the structure constants of a Lie algebra. In that case, the brackets will take the Lie-Poisson form (1.10.1) and the transformation to invariant variables will be the momentum map associated as in Remark 1.9.4 with the action of the symmetry group on the phase space.

8. The orbits of the solutions in the space of axisymmetric invariant variables in ray optics lie on the intersections of level sets of the Hamiltonian and the Casimir for the noncanonical bracket. The Petzval invariant S^2 in (1.7.13) is the Casimir for the Nambu bracket in $\mathbf{R}^3$, which for axisymmetric, translation-invariant ray optics is also a Lie-Poisson bracket. In this case, the ray paths are revealed when the Hamiltonian knife slices through the level sets of the Petzval invariant. These level sets are the layers of the *hyperbolic onion* shown in Figure 1.8. When restricted to a level set of the Petzval invariant, the dynamics becomes symplectic.

9. The *phases* associated with reconstructing the solution from the reduced space of invariant variables by going back to the original space of canonical coordinates and momenta naturally divide into their geometric and dynamic parts as in equation (1.12.12). In ray optics as governed by Fermat's principle, the geometric phase is related to the area enclosed by a periodic solution on a symplectic level set of the Petzval invariant S^2. This is no surprise, because the Poisson bracket on the level set is determined from the Jacobian using that area element.

10. A Lagrangian may be singular; that is, it may be degenerate, as occurs in the example of Fermat's principle in anisotropic media discussed in Section 1.13. This means the velocity cannot be

solved from the momentum and its conjugate coordinate. Even so, the dynamics resulting from the Lagrangian formulation of the problem may still be well-defined, in the sense that the solutions may still exist for the resulting ordinary differential equations.