

PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.

<http://hdl.handle.net/2066/170406>

Please be advised that this information was generated on 2019-11-22 and may be subject to change.

Chapter 3

Classical physics on a general phase space

Passing from finite phase spaces X to infinite ones yields many fascinating new phenomena, some of which even seem genuinely “emergent” in not having any finite-dimensional shadow, approximate or otherwise. Nonetheless, practically all results in the previous chapter remain valid, typically after the inclusion of some technical condition(s) that restrict the almost unlimited freedom allowed by infinite sets.

One of these restrictions is that in classical physics we assume that our phase space X is *locally compact Hausdorff*, where we recall that a space is:

- **compact** if every open cover has a finite subcover;
- **locally compact** if every point has a compact neighbourhood;
- **Hausdorff** (or T_2) if every pair of distinct points x, y can be separated by open sets (i.e., there are disjoint open sets U_x, U_y that contain x and y , respectively).

This combination of topological properties turns out to be very convenient; it incorporates spaces like \mathbb{R}^k (and more generally all non-pathological manifolds), or lattices like \mathbb{Z}^n (the price is that we exclude systems with an infinite number of degrees of freedom, such as classical field theories). A locally compact Hausdorff space X is **regular** in that each $x \in X$ and each closed set $F \subset X$ not containing x can be separated by open sets (i.e., there are disjoint open sets $U_x \ni x$ and $U_F \supset F$).

From the perspective of C^* -algebras, the main advantage of using this particular class of spaces is that they are naturally singled out by **Gelfand’s Theorem**:

Theorem 3.1. *Every commutative C^* -algebra A is isomorphic to $C_0(X)$ for some locally compact Hausdorff space X , which is unique up to homeomorphism.*

A proof may be found in Appendix C; here we just explain the notation and the main idea behind the proof (cf. Definition C.1, which we do not repeat).

First, $C_0(X)$ is the set of all continuous functions $f : X \rightarrow \mathbb{C}$ that **vanish at infinity**, i.e., for any $\varepsilon > 0$ the set $\{x \in X \mid |f(x)| \geq \varepsilon\}$ is compact, or, equivalently, for any $\varepsilon > 0$ there is a compact set $K \subset X$ such that $|f(x)| < \varepsilon$ for all $x \notin K$. For example, if $X = \mathbb{R}$, then $f(x) = \exp(-x^2)$ lies in $C_0(\mathbb{R})$. If X is compact, then $C_0(X) = C(X)$.

Second, $C_0(X)$ is a vector space under pointwise operations (including pointwise complex conjugation as the involution), and is a Banach space in the **sup-norm**

$$\|f\|_\infty = \sup_{x \in X} \{|f(x)|\}. \quad (3.1)$$

The space X making A isomorphic to $C_0(X)$, then, is the **Gelfand spectrum** $\Sigma(A)$ of A , which we already encountered (cf. Definition 1.4) as the set of nonzero algebra homomorphisms from A to \mathbb{C} . This set turns out to be a locally compact Hausdorff space in the topology of pointwise convergence, and the isomorphism $A \rightarrow C_0(X)$ is the **Gelfand transform** $a \mapsto \hat{a}$, where $\hat{a}(\omega) = \omega(a)$. Conversely, if X is given, then we associate the commutative C^* -algebra $C_0(X)$ to it, as in Chapter 1.

Generalizing Definition 1.14, as a special case of the notion of a state we have:

Definition 3.2. A state on $C_0(X)$ is a positive (and hence bounded) linear functional $\omega : C_0(X) \rightarrow \mathbb{C}$ with $\|\omega\| = 1$.

If X is compact, given positivity one has $\|\omega\| = 1$ iff $\omega(1_X) = 1$, cf. Lemma C.4.

The appropriate generalization of Theorem 1.15 then reads (cf. Corollary B.21):

Theorem 3.3. Let X be a locally compact Hausdorff space. There is a bijective correspondence between states on $C_0(X)$ and probability measures on X , namely

$$\varphi(f) = \int_X d\mu f, \quad f \in C_0(X). \quad (3.2)$$

Moreover, pure states correspond to Dirac measures and hence to points of X .

In particular, a nonzero linear functional $\omega : C_0(X) \rightarrow \mathbb{C}$ is multiplicative iff it is a pure state. This recovery of probability measures on phase space as states of the associated algebra of observables $C_0(X)$, and of points in phase space as the associated pure states, already familiar from the finite case, remains of great importance.

As in quantum mechanics, many interesting observables in classical mechanics fail to be bounded, let alone C_0 ; coordinate functions (on non-compact phase spaces) and the usual kinetic energy are a case in point. This is not a serious problem, especially not if, as we shall assume from now on, X is a (smooth) manifold (those unfamiliar with this notion may always have $X = \mathbb{R}^k$ in mind). In that case, there is a very natural class of (typically unbounded) functions on X , viz. $C^\infty(X) \equiv C^\infty(X, \mathbb{R})$, which form a commutative algebra just like $C_0(X) \equiv C_0(X, \mathbb{C})$, and provide the (algebraic) basis for the theory of symmetry and dynamics in classical physics, as we shall now show (the fact that functions in $C^\infty(X)$ may be freely added and multiplied provides a major simplification compared to unbounded operators in quantum mechanics, even self-adjoint ones, which are most easily treated by transforming them into bounded ones, as discussed in §B.21). In fact, the most natural mathematical setting of classical physics is not operator theory, or even symplectic geometry (as even mathematically minded people used to think until the 1980s), but rather the more general and flexible framework of **Poisson geometry**, to which we now turn.

3.1 Vector fields and their flows

We do not assume familiarity with differential geometry and analysis on manifolds, so in what follows one may assume that $M = \mathbb{R}^k$ for some k . However, whenever possible we will phrase definitions and results in such a way that their more general meaning should be clear to those who *are* familiar with differential geometry etc.

An *old-fashioned vector field* on $X = \mathbb{R}^k$ is a map

$$\xi : \mathbb{R}^k \rightarrow \mathbb{R}^k; \quad (3.3)$$

$$\xi(x) = (\xi^1(x), \dots, \xi^k(x)), \quad (3.4)$$

which describes something like a hyper-arrow at x . However, this is a coordinate-dependent object, which is hard to generalize to arbitrary manifolds. Therefore, in a modern approach a vector field is seen as the corresponding first-order differential operator $\xi : C^\infty(X) \rightarrow C^\infty(X)$ defined by

$$\xi f(x) = \sum_{j=1}^k \xi^j(x) \frac{\partial f(x)}{\partial x^j}. \quad (3.5)$$

To make the idea precise that a vector field on X is essentially the same as a first-order differential operator on $C^\infty(X)$, we note that it easily follows from (3.5) that

$$\xi(fg) = \xi(f)g + f\xi(g), \quad (3.6)$$

for any $f, g \in C^\infty(X)$, where the product fg is defined pointwise, i.e.,

$$(fg)(x) = f(x)g(x). \quad (3.7)$$

Similarly, we have pointwise addition and scalar multiplication, i.e., for $s, t \in \mathbb{R}$,

$$(sf + tg)(x) = sf(x) + tg(x). \quad (3.8)$$

This turns $C^\infty(X)$ into a commutative algebra (over \mathbb{R} , as $C^\infty(X) \equiv C^\infty(X, \mathbb{R})$).

A *derivation* of an algebra A (over \mathbb{R}) is a linear map $\delta : A \rightarrow A$ satisfying

$$\delta(ab) = \delta(a)b + a\delta(b). \quad (3.9)$$

Thus any vector field on X defines a derivation of the algebra $C^\infty(X)$ by (3.5). Conversely, a deep theorem of differential geometry states that for any manifold X , each derivation of $C^\infty(X)$ takes the form (3.5), at least locally (and for $X = \mathbb{R}^k$ also globally). Therefore, either as a definition or as a theorem, we often simply identify vector fields on X with derivations of $C^\infty(X)$. Derivations have a rich structure:

Definition 3.4. A (real) **Lie algebra** is a (real) vector space equipped with a bilinear map $[\cdot, \cdot] : A \times A \rightarrow A$ that satisfies $[a, b] = -[b, a]$ (and hence $[a, a] = 0$) as well as

$$[a, [b, c]] + [c, [a, b]] + [b, [c, a]] = 0 \quad (\text{Jacobi identity}). \quad (3.10)$$

It is easy to see that the set $\text{Vec}(X)$ of all old-fashioned vector fields ξ on X (i.e. in the sense (3.5)) forms a real Lie algebra under pointwise vector space operations (i.e., $(s\xi + t\eta)(f) = s\xi f + t\eta f$) and the natural bracket

$$[\xi, \eta] = \xi\eta - \eta\xi. \quad (3.11)$$

Similarly, the set $\text{Der}(A)$ of all derivations on some algebra is a Lie algebra under pointwise vector space operations and Lie bracket

$$[\delta_1, \delta_2] = \delta_1 \circ \delta_2 - \delta_2 \circ \delta_1. \quad (3.12)$$

Of course, the identification of $\text{Vec}(X)$ with $\text{Der}(C^\infty(X))$ identifies (3.11) and (3.12).

Vector fields (or, equivalently, derivations) may be “integrated”, at least *locally*, in the following sense. First, a **curve** through $x_0 \in X$ is a smooth map $c : I \rightarrow X$, where $I \subset \mathbb{R}$ is open and $c(t_0) = x_0$ for some $t_0 \in I$. We usually assume that $0 \in I$ with $t_0 = 0$ and hence $c(0) = x_0$. We then say that c **integrates** ξ near x_0 if

$$\dot{c}(t) = \xi(c(t)), \quad (3.13)$$

a somewhat symbolic equality that can be interpreted in two equivalent ways:

- Describing $c : I \rightarrow \mathbb{R}^k$ by k functions $c^j : I \rightarrow \mathbb{R}$ ($j = 1, \dots, k$), eq. (3.13) denotes

$$\frac{dc^j(t)}{dt} = \xi^j(c^1(t), \dots, c^k(t)), \quad j = 1, \dots, k. \quad (3.14)$$

- More abstractly, eq. (3.13) means that for any $f \in C^\infty(X)$ we have

$$\xi f(c(t)) = \frac{d}{dt} f(c(t)). \quad (3.15)$$

To pass from (3.15) to (3.14), we just have to recall (3.5), and note that

$$\frac{d}{dt} f(c(t)) = \frac{d}{dt} f(c^1(t), \dots, c^k(t)) = \sum_{j=1}^k \frac{dc^j(t)}{dt} \frac{\partial f(c(t))}{\partial x^j}. \quad (3.16)$$

The theory of ordinary differential equations shows that such local integral curves exist near any point $x_0 \in X$, and that they are unique in the following sense: if two curves $c_1 : I_1 \rightarrow X$ and $c_2 : I_2 \rightarrow X$ both satisfy (3.13) with $c_1(0) = c_2(0) = x_0$, then $c_1 = c_2$ on $I_1 \cap I_2$. However, curves that integrate ξ near some point may not be defined for all t , i.e., for $I = \mathbb{R}$. This makes the concept of a **flow** of a vector field ξ , which is meant to encapsulate all integral curves of ξ , a bit complicated. We start with the simplest case. We say that a vector field ξ is **complete** if for any $x_0 \in X$ there is a curve $c : \mathbb{R} \rightarrow X$ satisfying (3.13) with $c(0) = x_0$. The simplest example of a complete vector field is $X = \mathbb{R}$ and $\xi = d/dx$, so that $\varphi_t(x) = x + t$. For an incomplete example, take $X = \mathbb{R}$ and $\xi(x) = x^2 d/dx$. It can be shown that a vector field ξ with compact support (in the sense that the set $\{x \in X \mid \xi(x) \neq 0\}$ is bounded) is complete. In particular, any vector field on a compact manifold is complete.

Definition 3.5. Let X be a manifold and let $\xi \in \text{Vec}(X)$ be a complete vector field. A flow of ξ is a smooth map $\varphi : \mathbb{R} \times X \rightarrow X$, written

$$\varphi_t(x) \equiv \varphi(t, x), \quad (3.17)$$

that satisfies

$$\varphi_0(x) = x; \quad (3.18)$$

$$\varphi_s \circ \varphi_t = \varphi_{s+t}, \quad (3.19)$$

and that integrates ξ in the sense that for each $t \in \mathbb{R}$ and $x \in X$,

$$\xi(\varphi_t(x)) = \frac{d}{dt} \varphi_t(x). \quad (3.20)$$

As before, eq. (3.20) by definition means that for each $f \in C^\infty(X)$ we have

$$\xi f(\varphi_t(x)) = \frac{d}{dt} f(\varphi_t(x)), \quad (3.21)$$

or, equivalently, that in local coordinates, where

$$\varphi_t(x) = (\varphi_t^1(x), \dots, \varphi_t^k(x)), \quad (3.22)$$

we have

$$\frac{d\varphi_t^j(x)}{dt} = \xi^j(\varphi_t(x)), \quad j = 1, \dots, k. \quad (3.23)$$

Indeed, the flow φ of ξ gives the integral curve c of ξ through x_0 by

$$c(t) = \varphi_t(x_0). \quad (3.24)$$

According to the Picard–Lindelöf Theorem in the theory of ordinary differential equations, any complete vector field has a unique flow. In fact, the uniqueness part of this theorem implies that (3.19) is a consequence of (3.20) with (3.18), but it is convenient to state (3.19) separately, so as to make the point that the flow of a complete vector field ξ on X is a smooth \mathbb{R} -action on X , as defined by conditions (3.18) - (3.19), whose orbits integrate ξ . In particular, each $\varphi_t : X \rightarrow X$ is invertible, with inverse $\varphi_t^{-1} = \varphi_{-t}$. In particular, X is a disjoint union of the integral curves of ξ , which can never cross each other because of the uniqueness of the solution of the initial-value problem (3.13) with $c(0) = x_0$.

If ξ is not complete, we do the best we can by defining the set

$$D_\xi = \{(t, x) \in \mathbb{R} \times X \mid \exists c : I \rightarrow X, c(0) = x, t \in I\} \subset \mathbb{R} \times X, \quad (3.25)$$

where it is understood that c satisfies (3.13). Obviously $\{0\} \times X \subset D_\xi$, and (less trivially) it turns out that D_ξ is open. Then a flow of ξ is a map $\varphi : D_\xi \rightarrow X$ that satisfies (3.18) for all x , eq. (3.21) for $(t, x) \in D_\xi$, as well as (3.19) whenever defined.

3.2 Poisson brackets and Hamiltonian vector fields

To obtain flows, classical mechanics requires more than a manifold structure:

Definition 3.6. A **Poisson bracket** on a manifold X is a Lie bracket $\{-, -\}$ on (the real vector space) $C^\infty(X)$, such that for each $h \in C^\infty(X)$ the map

$$\xi_h : f \mapsto \{h, f\} \quad (3.26)$$

is a vector field on X (or, equivalently, a derivation of $C^\infty(X, \mathbb{R})$ with respect to its structure of a commutative algebra under pointwise multiplication). A manifold X equipped with a Poisson bracket is called a **Poisson manifold**, $(C^\infty(X), \{-, -\})$ is called a **Poisson algebra**, and ξ_h is called the **Hamiltonian vector field** of h .

Unfolding, we have a bilinear map $\{-, -\} : C^\infty(X) \times C^\infty(X) \rightarrow C^\infty(X)$ that satisfies

$$\{g, f\} = -\{f, g\}; \quad (3.27)$$

$$\{f, \{g, h\}\} + \{h, \{f, g\}\} + \{g, \{h, f\}\} = 0; \quad (3.28)$$

$$\{f, gh\} = \{f, g\}h + g\{f, h\}. \quad (3.29)$$

Bilinearity and the abstract properties (3.27) - (3.29) imply:

Proposition 3.7. Each Poisson bracket on X defines a Lie algebra homomorphism

$$C^\infty(X) \rightarrow \text{Der}(C^\infty(X)); \quad (3.30)$$

$$h \mapsto \delta_h, \quad (3.31)$$

or, equivalently, a Lie algebra homomorphism

$$C^\infty(X) \rightarrow \text{Vec}(X); \quad (3.32)$$

$$h \mapsto \xi_h. \quad (3.33)$$

The time-honored example is $X = \mathbb{R}^{2n}$, with coordinates $x = (p, q)$ and bracket

$$\{f, g\} = \sum_{j=1}^n \left(\frac{\partial f}{\partial p_j} \frac{\partial g}{\partial q^j} - \frac{\partial f}{\partial q^j} \frac{\partial g}{\partial p_j} \right). \quad (3.34)$$

In that case, the Hamiltonian vector field of h is obviously given by

$$\xi_h = \sum_{j=1}^n \left(\frac{\partial h}{\partial p_j} \frac{\partial}{\partial q^j} - \frac{\partial h}{\partial q^j} \frac{\partial}{\partial p_j} \right). \quad (3.35)$$

The flow of ξ_h gives the motion of a system with Hamiltonian h . Writing

$$\varphi_t(p, q) = (p(t), q(t)),$$

we see from (3.23) that this flow is given by **Hamilton's equations**

$$\frac{dp_j(t)}{dt} = -\frac{\partial h(p(t), q(t))}{\partial q^j}; \quad (3.36)$$

$$\frac{dq^j(t)}{dt} = \frac{\partial h(p(t), q(t))}{\partial p_j}. \quad (3.37)$$

Hamiltonians of the special form

$$h(p, q) = \frac{p^2}{2m} + V(q), \quad (3.38)$$

where $p^2 = \sum_j p_j^2$, give *Newton's equation* “ $F = ma$ ”, where $F_j = -\partial V / \partial q^j$, viz.

$$F_j(q(t)) = m \frac{d^2 q^j(t)}{dt^2}. \quad (3.39)$$

Proposition 3.8. *For any vector field ξ on a manifold X , we say that a function $f \in C^\infty(X)$ is **conserved** if f is constant along the flow of ξ . If X is a Poisson manifold and $\xi = \xi_h$ is Hamiltonian, then f is conserved iff $\{h, f\} = 0$.*

The proof is trivial. A Poisson bracket on X may also be defined in terms of a **Poisson tensor**. In coordinates, this is just an anti-symmetric matrix $B^{ij}(x)$ that satisfies

$$\sum_l \left(B^{li} \frac{\partial B^{jk}}{\partial x_l} + B^{lj} \frac{\partial B^{ki}}{\partial x_l} + B^{lk} \frac{\partial B^{ij}}{\partial x_l} \right) = 0, \quad (3.40)$$

for each (i, j, k) . In terms of B , the Poisson bracket is then defined abstractly by

$$\{f, g\} = B(df, dg), \quad (3.41)$$

using standard notation of differential geometry, or, in coordinates, by

$$\{f, g\}(x) = \sum_{i,j} B^{ij}(x) \frac{\partial f(x)}{\partial x^i} \frac{\partial g(x)}{\partial x^j}. \quad (3.42)$$

Conversely, a Poisson bracket must come from a Poisson tensor: for any derivation δ on $C^\infty(X)$, the function $\delta(g)$ depends linearly on dg , so if $\delta_f(g) = \{f, g\}$, then $\delta_f(g) = -\delta_g(f)$, so that $\{f, g\}$ depends linearly on both df and dg . This enforces (3.42), upon which (3.41) implies (3.40). A nice example is $X = \mathbb{R}^3$, with

$$\begin{aligned} \{f, g\}(\mathbf{x}) &= x \left(\frac{\partial f}{\partial y} \frac{\partial g}{\partial z} - \frac{\partial f}{\partial z} \frac{\partial g}{\partial y} \right) + y \left(\frac{\partial f}{\partial z} \frac{\partial g}{\partial x} - \frac{\partial f}{\partial x} \frac{\partial g}{\partial z} \right) + z \left(\frac{\partial f}{\partial x} \frac{\partial g}{\partial y} - \frac{\partial f}{\partial y} \frac{\partial g}{\partial x} \right); \\ B^{ij}(\mathbf{x}) &= \sum_k \varepsilon_{kij} x^k. \end{aligned} \quad (3.43)$$

Finally, we say that a Poisson manifold is **symplectic** if the corresponding Poisson tensor $B(x)$ is given by an *invertible* matrix, for each $x \in X$. This requires X to be *even-dimensional*. For example, \mathbb{R}^{2n} with Poisson bracket (3.34) is symplectic.

3.3 Symmetries of Poisson manifolds

Two equivalent notions of symmetries of classical physics suggest themselves: one is based on the idea of a Poisson *manifold* (X, B) , the other comes from the equivalent notion of a Poisson *algebra* $(C^\infty(X), \{, \})$.

Definition 3.9. 1. A *symmetry of a Poisson manifold* (X, B) is a diffeomorphism $\varphi : X \rightarrow X$ (that is, an invertible smooth map with smooth inverse) satisfying

$$\varphi_* B = B. \quad (3.44)$$

2. A *symmetry of a Poisson algebra* $(C^\infty(X), \{, \})$ is an invertible linear map $\alpha : C^\infty(X) \rightarrow C^\infty(X)$ that satisfies (for each $f, g \in C^\infty(X)$):

$$\alpha(fg) = \alpha(f)\alpha(g); \quad (3.45)$$

$$\alpha(\{f, g\}) = \{\alpha(f), \alpha(g)\}. \quad (3.46)$$

Let us define the push-forward φ_* in (3.44). We do this in terms of the *pullback* φ^* of a smooth (i.e., infinitely often differentiable) map $\varphi : X \rightarrow X$, defined as

$$\varphi^* : C^\infty(X) \rightarrow C^\infty(X); \quad (3.47)$$

$$\varphi^* f = f \circ \varphi. \quad (3.48)$$

If φ is a diffeomorphism, the *push-forward* φ_* of φ , which acts on derivations, is

$$\varphi_* : \text{Der}(C^\infty(X)) \rightarrow \text{Der}(C^\infty(X)); \quad (3.49)$$

$$(\varphi_* \delta)(f) = \delta(\varphi^* f) \circ \varphi^{-1}; \quad (3.50)$$

this may be checked to define a derivation, as follows:

$$\begin{aligned} (\varphi_* \delta)(f \cdot g) &= (\varphi^{-1})^* \delta(\varphi^*(f \cdot g)) \\ &= (\varphi^{-1})^* \delta(\varphi^*(f) \varphi^*(g)) \\ &= (\varphi^{-1})^* (\delta(\varphi^*(f)) \varphi^*(g) + \varphi^*(f) \delta(\varphi^*(g))) \\ &= (\varphi_* \delta)(f) \cdot g + f \cdot (\varphi_* \delta)(g). \end{aligned}$$

If, given coordinates $x = (x^1, \dots, x^k)$ on X , we now (without loss of generality) take our derivation δ to be a vector field $\xi = \sum_j \xi^j \partial / \partial x^j$, and write $\varphi(x) = (\varphi^1(x), \dots, \varphi^l(x))$, for the image $\varphi_*(\xi)$ we obtain

$$\begin{aligned} (\varphi_* \xi)(f)(x) &= (\xi(\varphi^* f))(\varphi^{-1}(x)) \\ &= \sum_j \xi^j(\varphi^{-1}(x)) \left(\frac{\partial}{\partial x^j} f \circ \varphi \right) (\varphi^{-1}(x)) \\ &= \sum_{j,k} \xi^k(\varphi^{-1}(x)) \frac{\partial f(x)}{\partial x^j} \frac{\partial \varphi^j}{\partial x^k} (\varphi^{-1}(x)), \end{aligned}$$

so that

$$\varphi_* \xi^j(x) = \sum_k \frac{\partial \varphi^j}{\partial x^k}(\varphi^{-1}(x)) \xi^k(\varphi^{-1}(x)), \quad (3.51)$$

or, equivalently,

$$\varphi_* \xi^j(\varphi(x)) = \sum_k \frac{\partial \varphi^j}{\partial x^k}(x) \xi^k(x), \quad (3.52)$$

which only depends on $\xi(x)$, so that for each $x \in X$, φ_* may be localized to a linear map $\varphi_*(x) : T_x X \rightarrow T_{\varphi(x)} X$. This may be done even if φ is not invertible. Physicists often write this as $\varphi(x) \equiv y = y(x^1, \dots, x^k)$, $\xi = v$, $\varphi_* \xi = v'$, so that we have a “covariant” transformation rule $(v')^i(y) = \sum_{j=1}^k \frac{\partial y^i(x)}{\partial x^j} v^j(x)$.

Taking tensor products, one obtains similar rules for higher-order tensors. For example, if $N = X$, the transformation rule for the Poisson tensor B reads

$$\varphi_* B^{ij}(\varphi(x)) = \sum_{m,n=1}^k \frac{\partial \varphi^i(x)}{\partial x^m} \frac{\partial \varphi^j(x)}{\partial x^n} B^{mn}(x), \quad (3.53)$$

so that, in coordinates, the invariance requirement (3.44) reads

$$\sum_{m,n=1}^k \frac{\partial \varphi^i(x)}{\partial x^m} \frac{\partial \varphi^j(x)}{\partial x^n} B^{mn}(x) = B^{ij}(\varphi(x)). \quad (3.54)$$

Theorem 3.10. *The two parts of Definition 3.9 are equivalent, in that:*

1. *Given a diffeomorphism $\varphi : X \rightarrow X$ satisfying (3.44), the map*

$$\alpha = \varphi^*, \quad (3.55)$$

i.e., $\alpha(f) = f \circ \varphi$, is linear, invertible, and satisfies (3.45) - (3.46).

2. *Given an invertible linear map $\alpha : C^\infty(X) \rightarrow C^\infty(X)$ that satisfies (3.45) - (3.46), there is a unique diffeomorphism $\varphi : X \rightarrow X$ inducing α as in (3.55).*

3. *This correspondence defines an anti-isomorphism between the group $\text{Diff}(X, B)$ of diffeomorphisms of X satisfying (3.44) and the group $\text{Aut}(C^\infty(X), \{, \})$ of invertible linear maps $\alpha : C^\infty(X) \rightarrow C^\infty(X)$ that satisfy (3.45) - (3.46).*

Here an **anti-isomorphism** of groups is just an isomorphism that inverts the order of multiplication. This complication may be removed by writing φ^{-1} instead of φ in (3.55), but that change would make the next proposition a bit less natural.

Proof. The first claim is true by construction. The hard part is the second claim, which follows from a more general result about manifolds (note that in our terminology, manifolds are by definition assumed to be Hausdorff):

Proposition 3.11. *Let X and Y be a smooth manifolds. Then (3.55) establishes a bijective correspondence between linear maps $\alpha : C^\infty(X) \rightarrow C^\infty(Y)$ satisfying (3.45) and smooth maps $\varphi : Y \rightarrow X$.*

The proof is quite similar to a central part of the proof of Gelfand duality for commutative C^* -algebras, in which (3.55) establishes a bijective correspondence between C^* -homomorphisms $\alpha : C(X) \rightarrow C(Y)$ and continuous maps $\varphi : Y \rightarrow X$, where X and Y are compact Hausdorff spaces; see §C.3 and especially Proposition C.22.

For any commutative real algebra A , let $\Sigma(A)$ be the space of non-zero algebra homomorphisms $\omega : A \rightarrow \mathbb{R}$ (these are just the non-zero multiplicative linear maps), equipped with the weakest topology that makes each function $\hat{a} : \Sigma(A) \rightarrow \mathbb{R}$ continuous, where $\hat{a}(\omega) = \omega(a)$. Furthermore, if B is another commutative real algebra, then any homomorphism $\alpha : A \rightarrow B$ induces a continuous map $\alpha^* : \Sigma(B) \rightarrow \Sigma(A)$ in the obvious way, that is, by $\alpha^* \omega = \omega \circ \alpha$. In the special case $A = C^\infty(X)$ (and similarly if $A = C(X)$), one has a canonical map $\text{ev}^X : X \rightarrow \Sigma(C(X))$, given by $\text{ev}_x^X(f) = f(x)$. The whole point (in which the entire difficulty of the proof lies) is that this map is a bijection (see Proposition C.21), which simultaneously equips X with a smooth structure that makes ev^X a diffeomorphism (by definition of the smooth structure on $\Sigma(C(X))$). In view of all this, given a multiplicative linear map $\alpha : C^\infty(X) \rightarrow C^\infty(Y)$, we obtain a continuous map $\varphi : Y \rightarrow X$ by

$$\varphi = (\text{ev}^Y)^{-1} \circ \alpha^* \circ \text{ev}^X. \quad (3.56)$$

Eq. (3.55) then holds by construction. Smoothness of φ , then, is a consequence of the fact that $\alpha(f) = f \circ \varphi$ must be a smooth function on Y for any $f \in C^\infty(X)$.

Applying this to the setting of Theorem 3.10 easily yields all claims. \square

In what follows, we look at smooth actions of Lie groups on (Poisson) manifolds X , in other words, at homomorphisms $\varphi : G \rightarrow \text{Diff}(X)$ or $\varphi : G \rightarrow \text{Diff}(X, B)$, where G is a Lie group, $\text{Diff}(X)$ is the group of all diffeomorphisms of a manifold, and $\text{Diff}(X, B)$ is the group of all diffeomorphisms of a Poisson manifold preserving the Poisson structure. Foregoing the underlying differential geometry, we take a pragmatic attitude and only study **linear Lie groups**, defined as closed subgroups G of $GL_n(\mathbb{R})$ or $GL_n(\mathbb{C})$, with group multiplication given by matrix multiplication and hence group inverse being matrix inverse. Here one may think of $SU(2) \subset GL_2(\mathbb{C})$ or $SO(3) \subset GL_3(\mathbb{R})$, but also abelian Lie groups like the additive groups \mathbb{R}^n fall under this scope, since one may identify $a \in \mathbb{R}^n$ with the $2n \times 2n$ -matrix

$$a \equiv \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}, \quad (3.57)$$

in which case matrix multiplication indeed reproduces addition. Similarly, the $2n + 1$ -dimensional **Heisenberg group** H_n is the group of real $(n + 2) \times (n + 2)$ -matrices

$$(a, b, c) = \begin{pmatrix} 1 & a^T & c + \frac{1}{2}a^T b \\ 0 & 1_n & b \\ 0 & 0 & 1 \end{pmatrix}, \quad (3.58)$$

where $a, b \in \mathbb{R}^n$, $c \in \mathbb{R}$, and $a^T b = \langle a, b \rangle$; this gives the multiplication rule

$$(a, b, c) \cdot (a', b', c') = (a + a', b + b', c + c' - \frac{1}{2}(\langle a, b' \rangle - \langle a', b \rangle)). \quad (3.59)$$

If G is a linear Lie group, its **Lie algebra** \mathfrak{g} may be defined as the vector space

$$\mathfrak{g} = \{A \in M_n(\mathbb{K}) \mid e^{tA} \in G \forall t \in \mathbb{R}\}, \quad (3.60)$$

where $\mathbb{K} = \mathbb{R}$ or \mathbb{C} , as determined by the embedding $G \subset GL_n(\mathbb{R})$ or $G \subset GL_n(\mathbb{C})$. Either way, \mathfrak{g} is seen as a *real* vector space, equipped with the **Lie bracket**

$$[A, B] = AB - BA. \quad (3.61)$$

This is trivially a bilinear antisymmetric map $\mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g}$ satisfying the Jacobi identity

$$[A, [B, C]] + [C, [A, B]] + [B, [C, A]] = 0, \quad (3.62)$$

which in turn expresses the fact that for fixed $A \in \mathfrak{g}$ the map $\delta_A : \mathfrak{g} \rightarrow \mathfrak{g}$ defined by

$$\delta_A(B) = [A, B] \quad (3.63)$$

is a derivation of \mathfrak{g} with respect to its Lie bracket, i.e.,

$$\delta_A([B, C]) = [\delta_A(B), C] + [B, \delta_A(C)]. \quad (3.64)$$

The **exponential map** $\exp : \mathfrak{g} \rightarrow G$ is then just given by its usual power series, which for matrices is norm-convergent. Conversely, one may pass from G to \mathfrak{g} through

$$A = \left. \frac{d}{dt}(e^{tA}) \right|_{t=0}. \quad (3.65)$$

If $G = \mathbb{R}^n$, we also have $\mathfrak{g} = \mathbb{R}^n$, and eq. (3.57) implies that \exp is the identity map.

For example, since $SO(3)$ is the subgroup of $GL_3(\mathbb{R})$ consisting of matrices R that satisfy $R^T R = 1_3$, its Lie algebra $\mathfrak{so}(3)$ consists of all matrices a that satisfy $a^T = -a$. As a vector space have $\mathfrak{so}(3) \cong \mathbb{R}^3$, which follows by choosing a basis

$$J_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, J_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, J_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \quad (3.66)$$

of the 3×3 real antisymmetric matrices. The commutators of these elements are

$$[J_1, J_2] = J_3; [J_3, J_1] = J_2; [J_2, J_3] = J_1. \quad (3.67)$$

For the Lie algebra of the Heisenberg group we obtain $\mathfrak{h}_n = \mathbb{R}^{2n+1}$, with basis

$$P_i = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -\mathbf{e}_i \\ 0 & 0 & 0 \end{pmatrix}, Q_j = \begin{pmatrix} 0 & \mathbf{e}_j^T & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, Z = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad (3.68)$$

where $(\mathbf{e}_1, \dots, \mathbf{e}_n)$ is the usual basis of \mathbb{R}^n , satisfying commutation relations

$$[P_i, Q_j] = \delta_{ij}Z; [P_i, P_j] = [Q_i, Q_j] = [P_i, Z] = [Q_j, Z] = 0. \quad (3.69)$$

3.4 The momentum map

Leaving out the Poisson structure for the moment, let X be a manifold, let G be a Lie group, and let $\varphi : G \rightarrow \text{Diff}(X)$ be a homomorphism; as already mentioned, this corresponds to a smooth action $\tilde{\varphi} : G \times X \rightarrow X$, which we simply write as

$$\gamma \cdot x \equiv \varphi_\gamma(x) \equiv \tilde{\varphi}(\gamma, x).$$

In terms of the pullback $\varphi_\gamma^*(f) = f \circ \varphi_\gamma$, we then automatically have

$$\varphi_\gamma^*(fg) = \varphi_\gamma^*(f)\varphi_\gamma^*(g). \quad (3.70)$$

For each $A \in \mathfrak{g}$ we then define a map $\delta_A : C^\infty(X) \rightarrow C^\infty(X)$ by

$$\delta_A f(x) = \frac{d}{dt} f(e^{-tA} \cdot x)|_{t=0}. \quad (3.71)$$

This map is obviously linear. Moreover, it can be shown that δ is well behaved:

Proposition 3.12. *The map $\delta : \mathfrak{g} \rightarrow \text{Der}(C^\infty(X))$, $A \mapsto \delta_A$ is a homomorphism of Lie algebra, i.e., each δ_A is a derivation, δ is linear in A , and, for each $A, B \in \mathfrak{g}$,*

$$[\delta_A, \delta_B] = \delta_{[A, B]}. \quad (3.72)$$

The proof relies on **Hadamard's Lemma**, which we only need for complete vector fields, or, equivalently, for derivations with complete flow (i.e., defined for all t).

Lemma 3.13. *If δ is a derivation of $C^\infty(X)$ with complete flow φ , and $f \in C^\infty(X)$, then there is a function $g(t, x) \equiv g_t(x)$ such that for all x and t ,*

$$g_0(x) = \delta f(x); \quad (3.73)$$

$$f(\varphi_t(x)) = f(x) + t g_t(x). \quad (3.74)$$

Indeed, if the flow is complete one may take

$$g_t(x) = \int_0^1 ds \dot{F}(st, x), \quad (3.75)$$

where $F(t, x) = f(\varphi_t(x))$ and (in Newton's notation) \dot{F} is the time derivative of F .

Proof. To prove that δ_A is linear in A , let φ be the flow of δ_A , i.e., $\varphi_t(x) = e^{-tA}x$. For $B \in \mathfrak{g}$, Hadamard's Lemma with $\delta \rightsquigarrow \delta_A$ and $x \rightsquigarrow e^{-tB}x$ then gives us

$$\begin{aligned} f(e^{-tA}e^{-tB}x) &= f(\varphi_t(e^{-tB}x)) = f(e^{-tB}x) + t g_t(e^{-tB}x); \\ \Rightarrow \frac{d}{dt} f(e^{-tA}e^{-tB}x)|_{t=0} &= \delta_B f(x) + g_0(x) = \delta_B f(x) + \delta_A f(x). \end{aligned} \quad (3.76)$$

On the other hand, since A and B are matrices, we may use the CBH-formula

$$e^{-tA}e^{-tB} = e^{-t(A+B) + \frac{1}{2}t^2[A,B] + O(t^3)}, \quad (3.77)$$

which gives $e^{-tA}e^{-tB} = e^{-t(A+B)}(1 + O(t^2))$, and hence

$$\frac{d}{dt}f(e^{-tA}e^{-tB}x)|_{t=0} = \frac{d}{dt}f(e^{-t(A+B)}x)|_{t=0} = \delta_{A+B}f(x). \quad (3.78)$$

Comparing (3.76) with (3.78) gives $\delta_{A+B} = \delta_A + \delta_B$. The property $\delta_{sA} = s\delta_A$ is trivial. We now prove (3.72). Within the (matrix) Lie algebra \mathfrak{g} we have

$$[A, B] = -\frac{d}{dt}(e^{-tA}Be^{tA})|_{t=0} = -\lim_{t \rightarrow 0} \frac{e^{-tA}Be^{tA} - B}{t}. \quad (3.79)$$

Furthermore, for any $g \in G$ one has $e^{gBg^{-1}} = ge^Bg^{-1}$, so linearity of δ gives

$$\begin{aligned} \delta_{[A,B]}f(x) &= -\lim_{t \rightarrow 0} \frac{1}{t} (\delta_{e^{-tA}Be^{tA}}f(x) - \delta_Bf(x)) \\ &= \lim_{t \rightarrow 0} \frac{1}{t} \left(\frac{d}{ds}f(e^{-tA}e^{sB}e^{tA}x) - \frac{d}{ds}f(e^{sB}x) \right) \\ &= \lim_{s,t \rightarrow 0} \frac{1}{st} (f(e^{-tA}e^{sB}e^{tA}x) - f(e^{-tA}e^{tA}e^{sB}x)) \\ &= \lim_{s,t \rightarrow 0} \frac{1}{st} (f \circ \varphi_t(e^{sB}e^{tA}x) - f \circ \varphi_t(e^{tA}e^{sB}x)) \\ &= \lim_{s,t \rightarrow 0} \left(\frac{1}{st} (f(e^{sB}e^{tA}x) - f(e^{tA}e^{sB}x)) + \frac{1}{s} (g_t(e^{sB}e^{tA}x) - g_t(e^{tA}e^{sB}x)) \right) \\ &= [\delta_A, \delta_B]f(x), \end{aligned}$$

since in the limit $t \rightarrow 0$ the third term in the penultimate line cancels the fourth. \square

Now suppose that, in addition, X is a Poisson manifold, and that each φ_γ acts on X as a Poisson symmetry, in that

$$\varphi_\gamma^*B = B, \quad (3.80)$$

cf. (3.44), or, equivalently, cf. (3.46),

$$\varphi_\gamma^* (\{f, g\}) = \{\varphi_\gamma^*(f), \varphi_\gamma^*(g)\}. \quad (3.81)$$

This implies, for each $A \in \mathfrak{g}$, and each $f, g \in C^\infty(X)$,

$$\delta_A(\{f, g\}) = \{\delta_A(f), g\} + \{f, \delta_A(g)\}. \quad (3.82)$$

Compare this with the following property δ_A already has since it is a derivation:

$$\delta_A(fg) = \delta_A(f)g + f\delta_A(g). \quad (3.83)$$

We may call a derivation $\delta : C^\infty(X) \rightarrow C^\infty(X)$ satisfying the like of (3.82), i.e.,

$$\delta(\{f, g\}) = \{\delta(f), g\} + \{f, \delta(g)\}, \quad (3.84)$$

a **Poisson derivation**. We are already familiar with a large class of Poisson derivations: for each $h \in C^\infty(X)$, the corresponding map δ_h defined by (3.26) is a Poisson derivation (this follows from the Jacobi identity). Let us call a Poisson derivation of the kind δ_h **inner**. This raises the question if our derivations δ_A are inner.

Definition 3.14. A **momentum map** for a Lie group G acting on a Poisson manifold X is a map

$$J : X \rightarrow \mathfrak{g}^* \quad (3.85)$$

such that for each $A \in \mathfrak{g}$,

$$\delta_A = \delta_{J_A}, \quad (3.86)$$

where the function $J_A \in C^\infty(X)$ is defined by

$$J_A(x) = \langle J(x), A \rangle \equiv J(x)(A). \quad (3.87)$$

In other words, for each $A \in \mathfrak{g}$ and $f \in C^\infty(X)$ we must have

$$\delta_A(f) = \{J_A, f\}. \quad (3.88)$$

A Lie group action admitting a momentum map is called **Hamiltonian**.

Equivalently, a momentum map is a linear map

$$J^* : \mathfrak{g} \rightarrow C^\infty(X) \quad (3.89)$$

such that $\delta_A = \delta_{J^*(A)}$; the connection between the two definitions is given by

$$J_A = J^*(A). \quad (3.90)$$

The pullback notation J^* would suggest that it is a map $C^\infty(\mathfrak{g}^*) \rightarrow C^\infty(X)$, which is not quite the case, but it is a near miss: we embed $\mathfrak{g} \hookrightarrow C^\infty(\mathfrak{g}^*)$ by $A \mapsto \hat{A}$, where $\hat{A}(\theta) = \theta(A)$, so $J^* : \mathfrak{g} \rightarrow C^\infty(X)$ is the restriction of the pullback J^* to \mathfrak{g} . Another near miss would be to read J^* as the adjoint to J , which maps $\mathfrak{g}^{**} \cong \mathfrak{g}$ to the ‘dual’ X^* , but since X may not be a vector space, this dual cannot be defined as in linear algebra, so instead of all linear maps from X to \mathbb{R} we might as well say that it consists of all smooth functions on X . Either way, the symbol J^* seems justified.

Proposition 3.15. Let G be a connected Lie group that acts on a Poisson manifold X . If this action is Hamiltonian (i.e., if it has a momentum map), then G acts on (X, B) by Poisson symmetries (in the sense that (3.81) holds).

Proof. An easy computation shows that (3.82) holds. We omit the proof of the fact that for *connected* Lie groups this “infinitesimal” property is equivalent to (3.81); this relies on the fact that G is generated by the image of the exponential map. \square

The converse is not true: if G acts by Poisson symmetries, the action is not necessarily Hamiltonian. For example, take $X = \mathbb{R}^2$, with the unusual Poisson bracket

$$\{f, g\}(p, q) = p \left(\frac{\partial f}{\partial p} \frac{\partial g}{\partial q} - \frac{\partial f}{\partial q} \frac{\partial g}{\partial p} \right), \quad (3.91)$$

and let $G = \mathbb{R}$ act on \mathbb{R}^2 by $b \cdot (p, q) = (p, q + b)$. This action satisfies (3.81), and has a single generator $\delta = -\partial/\partial q$. But there clearly is no function $J \in C^\infty(\mathbb{R}^2)$ such that $\{J, f\} = -\partial f/\partial q$ (it should be $J(p, q) = -\log(p)$, which is singular at $p = 0$).

However, in most “everyday situations” momentum maps exist:

1. Take $X = \mathbb{R}^6 = \mathbb{R}^3 \times \mathbb{R}^3$, with coordinates $x = (\mathbf{p}, \mathbf{q})$, where $\mathbf{p} = (p_1, p_2, p_3)$ and $\mathbf{q} = (q^1, q^2, q^3)$, equipped with the canonical Poisson bracket (3.34).

- a. Let $G = \mathbb{R}^6$ act on X by

$$(\mathbf{a}, \mathbf{b}) \cdot (\mathbf{p}, \mathbf{q}) = (\mathbf{p} + \mathbf{a}, \mathbf{q} + \mathbf{b}). \quad (3.92)$$

This action is Hamiltonian, with momentum map

$$J(\mathbf{p}, \mathbf{q}) = (\mathbf{q}, -\mathbf{p}). \quad (3.93)$$

- b. Let $G = SO(3)$ act on the same space X by

$$R \cdot (\mathbf{p}, \mathbf{q}) = (R\mathbf{p}, R\mathbf{q}). \quad (3.94)$$

Also this action is Hamiltonian, with momentum map

$$J(\mathbf{p}, \mathbf{q}) = \mathbf{p} \times \mathbf{q}. \quad (3.95)$$

2. Let $G = SO(3)$ act on $X = \mathbb{R}^3$, equipped with the Poisson bracket (3.43), through its defining representation. This action has a momentum map

$$J(\mathbf{x}) = \mathbf{x}, \quad (3.96)$$

where we have identified \mathfrak{g} with \mathbb{R}^3 by choosing the basis (3.66) of \mathfrak{g} , and have identified \mathfrak{g}^* with \mathfrak{g} (and hence with \mathbb{R}^3 also) by the usual inner product on \mathbb{R}^3 .

3. The previous example is a special case of the **Lie–Poisson structure**. Let G be a Lie group with Lie algebra \mathfrak{g} . Choose a basis (T_a) of \mathfrak{g} , with associated **structure constants** C_{ab}^c defined by the Lie bracket on \mathfrak{g} as

$$[T_a, T_b] = \sum_c C_{ab}^c T_c. \quad (3.97)$$

We write θ in the dual vector space \mathfrak{g}^* as $\theta = \sum_a \theta_a \omega^a$, where (ω_a) is the dual basis to a chosen basis (T_a) of \mathfrak{g} , i.e., $\omega_a(T_b) = \delta_{ab}$. In terms of these coordinates, the **Lie–Poisson bracket** on $C^\infty(\mathfrak{g}^*)$ is defined by

$$\{f, g\}(\theta) = C_{ab}^c \theta_c \frac{\partial f(\theta)}{\partial \theta_a} \frac{\partial g(\theta)}{\partial \theta_b}. \quad (3.98)$$

Equivalently, the Poisson bracket (3.98) may be defined by the condition

$$\{\widehat{A}, \widehat{B}\} = \widehat{[A, B]}, \quad (3.99)$$

where $A, B \in \mathfrak{g}$ and $\widehat{A} \in C^\infty(\mathfrak{g}^*)$ is the evaluation map $\widehat{A}(\theta) = \theta(A)$.

Now G canonically acts on \mathfrak{g}^* through the **coadjoint representation**, defined by

$$(x \cdot \theta)(A) = \theta(x^{-1}Ax). \quad (3.100)$$

This action is Hamiltonian with respect to the Lie–Poisson bracket (3.98), the associated momentum map simply being the identity map $\mathfrak{g}^* \rightarrow \mathfrak{g}^*$, as in (3.96). In other words, we have

$$J_A = \widehat{A}, \quad (3.101)$$

whose correctness may be verified from the computation

$$\begin{aligned} \delta_A \tilde{B}(\theta) &= \frac{d}{dt} \tilde{B}(e^{-tA} \cdot \theta)|_{t=0} = \frac{d}{dt} \theta(e^{tA} B e^{-tA})|_{t=0} \\ &= \theta([A, B]) = \widehat{[A, B]}(\theta) = \{\widehat{A}, \widehat{B}\}(\theta) \\ &= \{J_A, \widehat{B}\}(\theta). \end{aligned}$$

4. Let $X = T^*Q$ for some manifold Q . e.g. $Q = \mathbb{R}^n$ and hence $X = \mathbb{R}^{2n}$. We take

$$G = \text{Diff}(Q), \quad (3.102)$$

i.e., the diffeomorphism group of Q . This is an infinite-dimensional Lie group (if described in the right way). The defining action of $\varphi \in G$ on Q induces an action called φ^* on T^*Q , given (in coordinates) by

$$\varphi^*(p, q) = (p', q'); \quad (3.103)$$

$$(q^j)' = \varphi^j(q); \quad (3.104)$$

$$p'_i = \sum_{j=1}^n \frac{\partial(\varphi^{-1})^j(q)}{\partial q^i} p_j. \quad (3.105)$$

This may be taken as a definition, but in the language of differential geometry this comes down to the neater prescription that if $\theta = \sum_j p_j dq^j \in T_q^*Q$, then $\varphi^*\theta \in T_{\varphi(q)}^*Q$ is the one-form that maps a vector $X \in T_{\varphi(q)}Q$ to $\theta(\varphi_*^{-1}(X))$, i.e.,

$$(\varphi^*\theta)(X) = \theta(\varphi_*^{-1}(X)), \quad (3.106)$$

where $\varphi_*^{-1}(X) = \sum_j \varphi_*^{-1}(X)^j \partial/\partial q^j$ is given componentwise by, cf. (3.52),

$$\varphi_*^{-1}X^j = \sum_k \frac{\partial(\varphi^{-1})^j(q)}{\partial q^k} X^k. \quad (3.107)$$

If $Q = \mathbb{R}^3$ and $\varphi = R \in SO(3)$, then, using $R^{-1} = R^T$, we find that (3.104) - (3.105) simply become $R^*(\mathbf{p}, \mathbf{q}) = (R\mathbf{p}, R\mathbf{q})$, as in (3.94).

Furthermore, if $\varphi(\mathbf{q}) = \mathbf{q} + \mathbf{b}$, then the partial derivatives in (3.105) form the identity matrix, so that $\varphi^*(\mathbf{p}, \mathbf{q}) = (\mathbf{p}, \mathbf{q} + \mathbf{b})$. To show that the action of $\text{Diff}(Q)$ on T^*Q is Hamiltonian and compute its momentum map, we need to know that the Lie algebra of $\text{Diff}(Q)$ is the space $\text{Vec}(X)$ of all vector fields on Q , with its canonical Lie bracket (3.61)! We will not prove this, but the exponential map $\exp : \mathfrak{g} \rightarrow G$ is given through the flow φ of the vector field ξ on Q by (cf. (3.20))

$$e^{t\xi} = \varphi_t. \quad (3.108)$$

Theorem 3.16. *The action of $\text{Diff}(Q)$ on T^*Q has momentum map*

$$J_X(p, q) = -\sum_j p_j X^j(q), \quad (3.109)$$

and hence is Hamiltonian. Moreover, this momentum map satisfies

$$\{J_\xi, J_\eta\}_\xi = -J_{[\xi, \eta]}. \quad (3.110)$$

Proof. First note that $\varphi_t^{-1} = \varphi_{-t}$, so from (3.71), (3.108), and (3.104) - (3.105),

$$\begin{aligned} \delta_\xi f(p, q) &= \frac{d}{dt} f(\varphi_{-t}^*(p, q))|_{t=0} \\ &= \sum_{i,j} \frac{\partial f}{\partial p_i}(p, q) \frac{d}{dt} \left(\frac{\partial \varphi_t^j(q)}{\partial q^i} \right) \Big|_{t=0} p_j + \sum_i \frac{\partial f}{\partial q^i}(p, q) \frac{d}{dt} \varphi_{-t}^i(q) \Big|_{t=0} \\ &= \sum_{i,j} p_j \frac{\partial X^j(q)}{\partial q^i} \frac{\partial f}{\partial p_i}(p, q) - \sum_j X^j(q) \frac{\partial f}{\partial q^j}(p, q). \end{aligned}$$

From this and (3.109), using the canonical Poisson bracket (3.34) we find

$$\{J_\xi, f\} = \delta_\xi f.$$

Finally, verifying (3.110) is a simple exercise. \square

Thus the momentum map is a generalization of (minus) the momentum, whence its name; the quantity in (3.95) is (minus) the angular momentum. These annoying minus signs could be removed by putting a minus sign in (3.86), but that would have other negative (*sic*) consequences. For example, with our sign choice one often has

$$\{J_A, J_B\} = J_{[A, B]}, \quad (3.111)$$

in which case the accompanying map (3.89) is a homomorphism of Lie algebras, or, equivalently, J is a morphism with respect to the given Poisson bracket on X and the Lie–Poisson bracket on \mathfrak{g}^* . Such a momentum map is called *infinitesimally equivariant*, for if G is connected, (3.111) is equivalent to the equivariance property

$$J(g \cdot x) = g \cdot J(x). \quad (3.112)$$

Here the G -action on \mathfrak{g}^* on the right-hand side is the coadjoint representation.

All of this is true for our examples (3.95), (3.96), (3.101), and (3.109); in the latter case we note that the Lie bracket in the Lie algebra of $\text{Diff}(Q)$ is *minus* the commutator of vector fields. However, (3.111) does not always hold (in which case *a fortiori* also (3.112) fails). For example, it fails for (3.93): if we take the usual basis $(\mathbf{e}, \mathbf{f}) \equiv (e_1, e_2, e_3, f_1, f_2, f_3)$ of $\mathfrak{g} = \mathbb{R}^6$ and relabel $e_j \equiv Q_j$ and $f_i \equiv -P_i$, then

$$J_{P_i}(\mathbf{p}, \mathbf{q}) = p_i; \quad (3.113)$$

$$J_{Q_j}(\mathbf{p}, \mathbf{q}) = q_j, \quad (3.114)$$

cf. (3.93), and hence, although $[P_i, P_j] = [Q_i, Q_j] = [P_i, Q_j] = 0$, we obtain

$$\{J_{P_i}, J_{P_j}\} = \{J_{Q_i}, J_{Q_j}\} = 0; \quad (3.115)$$

$$\{J_{P_i}, J_{Q_j}\} = \delta_{ij} 1_{\mathbb{R}^6}. \quad (3.116)$$

Fortunately, in cases like that one can often find a central extension G_φ of G (see §5.10 below for notation) that acts on X through its quotient group G and does have an infinitesimally equivariant momentum map. In the case at hand, the Heisenberg group H_3 does the job, whose central elements $(0, 0, c)$ then act trivially on \mathbb{R}^6 . In terms of the generators (3.68) we take J_{P_i} and J_{Q_j} as in (3.113) - (3.114), and add $J_Z = 1_{\mathbb{R}^6}$; according to (3.69) and (3.115) - (3.116) we then have (3.111), as desired.

Finally, the above formalism leads to a clean formulation of **Noether's Theorem**, providing the well-known link between symmetries and conserved quantities:

Theorem 3.17. *Let X be a Poisson action equipped with a Hamiltonian action of some Lie group G (so that there is a momentum map $J : X \rightarrow \mathfrak{g}^*$). Suppose $h \in C^\infty(X)$ is G -invariant, in that $h(\gamma \cdot x) = h(x)$ for each $\gamma \in G$ and $x \in X$. Then for each $A \in \mathfrak{g}$, the function J_A is constant along the flow of the vector field X_h . In other words,*

$$J_A(\varphi_t(x)) = J_A(x) \quad (3.117)$$

for any $x \in X$ and any $t \in \mathbb{R}$ for which the flow $\varphi_t(x)$ of X_h is defined.

Proof. Using all assumptions as well as the definition of a flow, we compute:

$$\begin{aligned} \frac{d}{dt} J_A(\varphi_t(x)) &= X_h(J_A)(\varphi_t(x)) = \delta_h(J_A)(\varphi_t(x)) \\ &= \{h, J_A\}(\varphi_t(x)) = -\{J_A, h\}(\varphi_t(x)) \\ &= -\delta_A(h)(\varphi_t(x)) = \frac{d}{ds} h(e^{sA} \varphi_t(x))|_{s=0} \\ &= \frac{d}{ds} h(\varphi_t(x))|_{s=0} = 0. \quad \square \end{aligned}$$

For example, a Hamiltonian (3.38) has conserved (angular) momentum if the potential V is translation (rotation) invariant, reflecting (3.93) and (3.95), respectively.

Notes

The traditional symplectic approach to classical mechanics, culminating in the momentum map, is exhaustively covered in Guillemin & Sternberg (1984) and Abraham & Marsden (1985). A founding paper for Poisson geometry is Weinstein (1983). The modern Poisson approach to mechanics may be found in Marsden & Ratiu (1994), from which most of the material in this chapter originates.

Our proof of Proposition 3.11 is based on Navarro González & Sancho de Salas (2003), §2.1. Burtscher (2009) is a nice survey of many similar results.