I

Preliminaries

In this chapter, we collect some key notions and facts from the theory of Lie groups and Hamiltonian systems, as well as set up the notations.

1 Lie Groups and Lie Algebras

This section introduces the notions of a Lie group and the corresponding Lie algebra. Many of the basic facts known for finite-dimensional Lie groups are no longer true for infinite-dimensional ones, and below we illustrate some of the pathologies one can encounter in the infinite-dimensional setting.

1.1 Lie Groups and an Infinite-Dimensional Setting

The most basic definition for us will be that of a (transformation) group.

Definition 1.1 A nonempty collection G of transformations of some set is called a (transformation) *group* if along with every two transformations $g, h \in G$ belonging to the collection, the composition $g \circ h$ and the inverse transformation g^{-1} belong to the same collection G.

It follows directly from this definition that every group contains the identity transformation e. Also, the composition of transformations is an associative operation. These properties, associativity and the existence of the unit and an inverse of each element, are often taken as the definition of an abstract group.[1]

The groups we are concerned with in this book are so-called Lie groups. In addition to being a group, they carry the structure of a smooth manifold such that both the multiplication and inversion respect this structure.

[1] Here we employ the point of view of V.I. Arnold, that every group should be viewed as the group of transformations of some set, and the "usual" axiomatic definition of a group only obscures its true meaning (cf. [19], p. 58).

Definition 1.2 A *Lie group* is a smooth manifold G with a group structure such that the multiplication $G \times G \to G$ and the inversion $G \to G$ are smooth maps.

The Lie groups considered throughout this book will usually be infinite-dimensional. So what do we mean by an infinite-dimensional manifold? Roughly speaking, an infinite-dimensional manifold is a manifold modeled on an infinite-dimensional locally convex vector space just as a finite-dimensional manifold is modeled on $\mathbb{R}^n$.

Definition 1.3 Let V, W be *Fréchet spaces*, i.e., complete locally convex Hausdorff metrizable vector spaces, and let U be an open subset of V. A map $f : U \subset V \to W$ is said to be *differentiable* at a point $u \in U$ in a direction $v \in V$ if the limit

$$Df(u;v) = \lim_{t \to 0} \frac{f(u+tv) - f(u)}{t} \tag{1.1}$$

exists. The function is said to be continuously differentiable on U if the limit exists for all $u \in U$ and all $v \in V$, and if the function $Df : U \times V \to W$ is continuous as a function on $U \times V$. In the same way, we can build the second derivative $D^2 f$, which (if it exists) will be a function $D^2 f : U \times V \times V \to W$, and so on. A function $f : U \to W$ is called *smooth* or C^∞ if all its derivatives exist and are continuous.

Definition 1.4 A *Fréchet manifold* is a Hausdorff space with a coordinate atlas taking values in a Fréchet space such that all transition functions are smooth maps.

Remark 1.5 Now one can start defining vector fields, tangent spaces, differential forms, principal bundles, and the like on a Fréchet manifold exactly in the same way as for finite-dimensional manifolds.

For example, for a manifold M, a *tangent vector* at some point $m \in M$ is defined as an equivalence class of smooth parametrized curves $f : \mathbb{R} \to M$ such that $f(0) = m$. The set of all such equivalence classes is the *tangent space* $T_m M$ at m. The union of the tangent spaces $T_m M$ for all $m \in M$ can be given the structure of a Fréchet manifold TM, the tangent bundle of M. Now a smooth vector field on the manifold M is a smooth map $v : M \to TM$, and one defines in a similar vein the directional derivative of a function and the Lie bracket of two vector fields.

Since the dual of a Fréchet space need not be Fréchet, we define differential 1-forms in the Fréchet setting directly, as smooth maps $\alpha : TM \to \mathbb{R}$ such that for any $m \in M$, the restriction $\alpha|_{T_m M} : T_m M \to \mathbb{R}$ is a linear map. Differential forms of higher degree are defined analogously: say, a 2-form on a Fréchet manifold M is a smooth map $\beta : T^{\otimes 2} M \to \mathbb{R}$ whose restriction $\beta|_{T_m^{\otimes 2} M} : T_m^{\otimes 2} M \to \mathbb{R}$ for any $m \in M$ is bilinear and antisymmetric. The differential df of a smooth function $f : M \to \mathbb{R}$ is defined via the directional derivative, and this construction generalizes to smooth n-forms on a Fréchet manifold M to give the *exterior derivative* operator d, which maps n-forms to $(n+1)$-forms on M; see, for example, [231].

Remark 1.6 More facts on infinite-dimensional manifolds can be found in, e.g., [265, 157]. From now on, whenever we speak of an infinite-dimensional manifold, we implicitly mean a Fréchet manifold (unless we say explicitly otherwise). In particular, our infinite-dimensional Lie groups are *Fréchet Lie groups.*

Instead of Fréchet manifolds, one could consider manifolds modeled on Banach spaces. This would lead to the category of Banach manifolds. The main advantage of Banach manifolds is that strong theorems from finite-dimensional analysis, such as the inverse function theorem, hold in Banach spaces but not necessarily in Fréchet spaces. However, some of the Lie groups we will be considering, such as the diffeomorphism groups, are not Banach manifolds. For this reason we stay within the more general framework of Fréchet manifolds. In fact, for most purposes, it is enough to consider groups modeled on locally convex vector spaces. This is the setting considered by Milnor [265].

1.2 The Lie Algebra of a Lie Group

Definition 1.7 Let G be a Lie group with the identity element $e \in G$. The tangent space to the group G at its identity element is (the vector space of) the *Lie algebra* $\mathfrak{g}$ of this group G. The group multiplication on a Lie group G endows its Lie algebra $\mathfrak{g}$ with the following bilinear operation $[\ ,\ ] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$, called the *Lie bracket* on $\mathfrak{g}$.

First note that the Lie algebra $\mathfrak{g}$ can be identified with the set of left-invariant vector fields on the group G. Namely, to a given vector $X \in \mathfrak{g}$ one can associate a vector field $\widetilde{X}$ on G by left translation: $\widetilde{X}(g) = l_{g*}X$, where $l_g : G \to G$ denotes the multiplication by a group element g from the left, $h \in G \mapsto gh$. Obviously, such a vector field $\widetilde{X}$ is invariant under left translations by elements of G. That is, $l_{g*}\widetilde{X} = \widetilde{X}$ for all $g \in G$. On the other hand, any left-invariant vector field $\widetilde{X}$ on the group G uniquely defines an element $\widetilde{X}(e) \in \mathfrak{g}$.

The usual Lie bracket (or commutator) $[\widetilde{X}, \widetilde{Y}]$ of two left-invariant vector fields $\widetilde{X}$ and $\widetilde{Y}$ on the group is again a left-invariant vector field on G. Hence we can write $[\widetilde{X}, \widetilde{Y}] = \widetilde{Z}$ for some $Z \in \mathfrak{g}$. We define the *Lie bracket* $[X, Y]$ of two elements X, Y of the Lie algebra $\mathfrak{g}$ of the group G via $[X, Y] := Z$. The Lie bracket gives the space $\mathfrak{g}$ the *structure of a Lie algebra.*

Examples 1.8 Here are several finite-dimensional Lie groups and their Lie algebras:

- $\mathrm{GL}(n, \mathbb{R})$, the set of nondegenerate $n \times n$ matrices, is a Lie group with respect to the matrix product: multiplication and taking the inverse are smooth operations. Its Lie algebra is $\mathfrak{gl}(n, \mathbb{R}) = \mathrm{Mat}(n, \mathbb{R})$, the set of all $n \times n$ matrices.

- $\mathrm{SL}(n, \mathbb{R}) = \{A \in \mathrm{GL}(n, \mathbb{R}) \mid \det A = 1\}$ is a Lie group and a closed subgroup of $\mathrm{GL}(n, \mathbb{R})$. Its Lie algebra is the space of traceless matrices $\mathfrak{sl}(n, \mathbb{R}) = \{A \in \mathfrak{gl}(n, \mathbb{R}) \mid \operatorname{tr} A = 0\}$. This follows from the relation

$$\det(I + \epsilon A) = 1 + \epsilon \operatorname{tr} A + \mathcal{O}(\epsilon^2)\,, \qquad \text{as} \qquad \epsilon \to 0\,,$$

  where I is the identity matrix.
- $\mathrm{SO}(n, \mathbb{R})$ is a Lie group of transformations $\{A : \mathbb{R}^n \to \mathbb{R}^n\}$ preserving the Euclidean inner product of vectors (and orientation) in $\mathbb{R}^n$, i.e. $(Au,\ Av) = (u,\ v)$ for all vectors $u, v \in \mathbb{R}^n$. Equivalently, one can define

$$\mathrm{SO}(n, \mathbb{R}) = \{A \in \mathrm{GL}(n, \mathbb{R}) \mid AA^t = I,\ \det A > 0\}.$$

  The Lie algebra of $\mathrm{SO}(n)$ is the space of skew-symmetric matrices

$$\mathfrak{so}(n, \mathbb{R}) = \{A \in \mathfrak{gl}(n, \mathbb{R}) \mid A + A^t = 0\}\,,$$

  as the relation

$$(I + \epsilon A)(I + \epsilon A^t) = I + \epsilon(A + A^t) + \mathcal{O}(\epsilon^2)$$

  shows.
- $\mathrm{Sp}(2n, \mathbb{R})$ is the group of transformations of $\mathbb{R}^{2n}$ preserving the nondegenerate skew-product of vectors.

Exercise 1.9 Give an alternative definition of $\mathrm{Sp}(2n, \mathbb{R})$ with the help of the equation satisfied by the corresponding matrices for the following skew-product of vectors $\langle u, v\rangle := \sum_{j=1}^n (u_j v_{j+n} - v_j u_{j+n})$. Find the corresponding Lie algebra.

Exercise 1.10 Show that in all of Examples 1.8, the Lie bracket is given by the usual commutator of matrices: $[A, B] = AB - BA$.

The following examples are the first infinite-dimensional Lie groups we shall encounter.

Example 1.11 Let M be a compact n-dimensional manifold. Consider the set $\mathrm{Diff}(M)$ of diffeomorphisms of M. It is an open subspace of (the Fréchet manifold of) all smooth maps from M to M. One can check that the composition and inversion are smooth maps, so that the set $\mathrm{Diff}(M)$ is a Fréchet

Lie group; see [157].[2] Its Lie algebra is given by $\mathrm{Vect}(M)$, the Lie algebra of smooth vector fields on M.

Given a volume form μ on M, one can define the group of volume-preserving diffeomorphisms

$$S\mathrm{Diff}(M) := \{\phi \in \mathrm{Diff}(M) \mid \phi^*\mu = \mu\}\,.$$

It is a Lie group, since $S\mathrm{Diff}(M)$ is a closed subgroup of $\mathrm{Diff}(M)$. Its Lie algebra $S\mathrm{Vect}(M) := \{v \in \mathrm{Vect}(M) \mid \mathrm{div}(v) = 0\}$ consists of vector fields on M that are divergence-free with respect to the volume form μ.

Example 1.12 Let M be a finite-dimensional compact manifold and let G be a finite-dimensional Lie group. Set the *group of currents* on M to be $G^M = C^\infty(M, G)$, the group of G-valued functions on M. We can define a multiplication on G^M pointwise, i.e., we set $(\varphi \cdot \psi)(g) = \varphi(g)\psi(g)$ for all φ, $\psi \in G^M$. This multiplication gives G^M the structure of a (Fréchet) Lie group, as we discuss below.

Example 1.13 A slight, but important, generalization of the example above is the following: Let G be a finite-dimensional Lie group, and P a principal G-bundle over a manifold M. Denote by $\pi : P \to M$ the natural projection to the base. Define the Lie *group* $\mathrm{Gau}(P)$ *of gauge transformations* (or, simply, the *gauge group*) of P as the group of bundle (i.e., fiberwise) automorphisms: $\mathrm{Gau}(P) = \{\varphi \in \mathrm{Aut}(P) \mid \pi \circ \varphi = \pi\}$. The group multiplication is the natural composition of the bundle automorphisms. (Automorphisms of each fiber of P form a copy of the group G, and all together they define the associated bundle over M with the structure group G. The identity bundle automorphism gives the trivial section of this associated G-bundle, and the gauge transformation group consists of all smooth sections of it; see details in [265].) One can show that this is a Lie group (cf. [157]), and we denote the corresponding Lie algebra by $\mathfrak{gau}(P)$. For a topologically trivial G-bundle P, the group $\mathrm{Gau}(P)$ coincides with the current group G^M.

Exercise 1.14 Describe the Lie brackets for the Lie algebras in the last three examples.

Remark 1.15 For a Lie group G, the Lie bracket on the corresponding Lie algebra $\mathfrak{g}$, which we defined via the usual Lie bracket of left-invariant vector fields on the group, satisfies the following properties:

---

[2] In many analysis questions it is convenient to work with the larger space of diffeomorphisms $\mathrm{Diff}^s(M)$ of Sobolev class H^s. For $s > n/2 + 1$ these spaces are smooth Hilbert manifolds. On the other hand, the spaces $\mathrm{Diff}^s(M)$ are only topological (but not smooth) groups, since the composition of such diffeomorphisms is not smooth. Indeed, while the right multiplication $r_\phi : \psi \mapsto \psi \circ \phi$ is smooth, the left multiplication $l_\psi : \phi \mapsto \psi \circ \phi$ is only continuous, but not even Lipschitz continuous; see [95].

(i) it is antisymmetric in X and Y, i.e., $[X,Y] = -[Y,X]$, and
(ii) it satisfies the Jacobi identity:

$$[[X,Y],Z] + [[Z,X],Y] + [[Y,Z],X] = 0\,.$$

The Jacobi identity can be thought of as an infinitesimal analogue of the associativity of the group multiplication.

1.3 The Exponential Map

Definition 1.16 The *exponential map* from a Lie algebra to the corresponding Lie group $\exp : \mathfrak{g} \to G$ is defined as follows: Let us fix some $X \in \mathfrak{g}$ and let $\widetilde{X}$ denote the corresponding left-invariant vector field. The flow of the field $\widetilde{X}$ is a map $\phi_X : G \times \mathbb{R} \to G$ such that $\frac{d}{dt}\phi_X(g,t) = \widetilde{X}(\phi_X(g,t))$ for all t and $\phi_X(g,0) = g$. The flow ϕ_X is the solution of an ordinary differential equation, which, if it exists, is unique. In the case that the flow subgroup $\phi_X(e,.)$ exists for all $X \in \mathfrak{g}$, we define the exponential map $\exp : \mathfrak{g} \to G$ via the time-one map $X \mapsto \phi_X(e,1)$; see Figure 1.1.

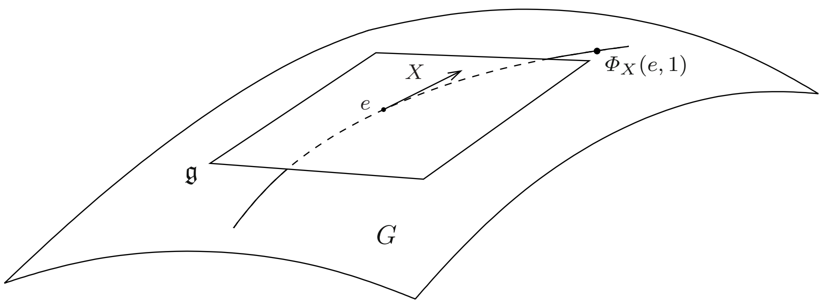


Fig. 1.1. The exponential map on the group G associates to a vector X the time-one map for the trajectory of a left-invariant vector field defined by X at $e \in G$.

Example 1.17 For each of the finite-dimensional Lie groups considered in Example 1.8, the exponential map is given by the usual exponential map for matrices:

$$\exp : A \mapsto \exp(A) = \sum_{n=0}^{\infty} \frac{1}{n!} A^n\,.$$

Remark 1.18 The definition of the exponential map relies on the existence and uniqueness of solutions of certain first-order differential equations. In general, solutions of differential equations in Fréchet spaces might not be unique.[3] However, the differential equation in the definition of the exponential map is of special type, which secures the solution's uniqueness upon fixing its initial condition. Namely, let $\phi : \mathbb{R} \to G$ be a smooth path in the Lie group G. Its derivative $\phi'(t) := \frac{d}{dt}\phi(t)$ is a tangent vector to the group G at the point $\phi(t)$. Translate this vector back to the identity via left multiplication by $\phi^{-1}(t)$. The corresponding element of the Lie algebra $\mathfrak{g}$ is denoted by $\phi^{-1}(t)\phi'(t)$ and is called the *left logarithmic derivative* of the path ϕ.

Now consider a Lie algebra element $X \in \mathfrak{g}$. By definition of the exponential map, the curve $\phi(t) = \exp(tX)$ satisfies the differential equation $\phi'(t) = \phi(t)X$ with the initial condition $\phi(0) = e$. So for all solutions of this differential equation, the left logarithmic derivative is given by the constant curve $X \in \mathfrak{g}$. Now the uniqueness of the exponential map is implied by the following Exercise.

Exercise 1.19 Show that two smooth paths ϕ, $\psi : \mathbb{R} \to G$ have the same left logarithmic derivative for all $t \in \mathbb{R}$ if and only if they are translations of each other by some constant element $g \in G$: $\phi(t) = g\,\psi(t)$ for all $t \in \mathbb{R}$. (Hint: see, e.g., [265].)

Remark 1.20 As far as the existence is concerned, the exponential map exists for all finite-dimensional Lie groups and more generally for Lie groups modeled on Banach spaces, as follows from the general theory of differential equations. However, there may exist infinite-dimensional Lie groups that do not admit an exponential map. Moreover, even in the cases in which the exponential map of an infinite-dimensional group exists, it can exhibit rather peculiar properties; see the examples below.

Example 1.21 For the diffeomorphism group $\mathrm{Diff}(M)$ the exponential map $\exp : \mathrm{Vect}(M) \to \mathrm{Diff}(M)$ has to assign to each vector field on M the time-one map for its flow. However, for a noncompact M this map may not exist: the corresponding vector field may not be complete. Indeed, for example, for the vector field $\xi = x^2\partial/\partial x$ on the real line $M = \mathbb{R}$, the time-one map of the flow is not defined on the whole of $\mathbb{R}$: the corresponding flow sends some points to infinity for the time less than 1! Fortunately, for compact manifolds M and smooth vector fields, the time-one maps of the corresponding flows, and hence the exponential maps, are well defined.

Note that the group of diffeomorphisms of a noncompact manifold is not complete, and hence it is not a Lie group in our sense. It is an important open problem to find a Lie group that is modeled on a complete space and does not admit an exponential map.

[3] For instance, the initial value problem $u(x, 0) = f(x)$ for the equation $u_t(x, t) = u_x(x, t)$ with $x \in [0, 1]$ has wave-type solutions $u(x, t) = f(x + t)$. For nonzero t such a solution $u(x, t)$ for $x \in [0, 1]$ depends on the extension of $f(x)$ to the segment $[-t, 1 - t]$. Due to arbitrariness in the choice of a smooth extension of f from $[0, 1]$ to $\mathbb{R}$, the solution to this initial value problem is not unique.

Let us return to the current group G^M, where the exponential map exists and can be used to give this group the structure of a Fréchet Lie group. Namely, the space $\mathfrak{g}^M = C^\infty(M, \mathfrak{g})$ endowed with the topology of uniform convergence is a Fréchet space. Moreover, the map $\exp : \mathfrak{g} \to G$ can be used to define a map $\widetilde{\exp} : \mathfrak{g}^M \to G^M$ pointwise. In a sufficiently small neighborhood of $0 \in \mathfrak{g}^M$, the map $\widetilde{\exp}$ is bijective. Thus it can be used to define a local system of open neighborhoods of the identity in G^M. We can use left translation to transfer this system to any point in G^M and thus define a topology on the group G^M. Again using the exponential map, we can define coordinate charts on G^M. This definition implies that multiplication and inversion in G^M are smooth maps. So G^M is an infinite-dimensional Lie groups (see, e.g., [157] for more details).

From the construction of the Lie group structure on G^M, it is clear that its Lie algebra is the current algebra $\mathfrak{g}^M$, and that the exponential map $\mathfrak{g}^M \to G^M$ is the map $\widetilde{\exp}$ described above. Note, however, that $\widetilde{\exp}$ is not, in general, surjective, even if $\exp : \mathfrak{g} \to G$ is surjective. As an example, take the manifold M to be the circle S^1 and G to be the group $\mathrm{SU}(2)$.

Exercise 1.22 Show that the map

$$\theta \mapsto \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix}$$

for $\theta \in S^1 = \mathbb{R}/2\pi\mathbb{Z}$ defines an element in G^{S^1} that does not belong to the image of the exponential map $\widetilde{\exp} : \mathfrak{g}^{S^1} \to G^{S^1}$.

In contrast to the exponential map in the case of the current group G^M, the exponential map $\exp : \mathrm{Vect}(M) \to \mathrm{Diff}(M)$ for the diffeomorphism group of a compact M is not, in general, even locally surjective already for the case of a circle.

Proposition 1.23 (see, e.g., [265, 301, 322]) *The exponential map* $\exp : \mathrm{Vect}(S^1) \to \mathrm{Diff}(S^1)$ *is not locally surjective.*

Proof. First observe that any nowhere-vanishing vector field on S^1 is conjugate under $\mathrm{Diff}(S^1)$ to a constant vector field. Indeed, if $\xi(\theta) = v(\theta)\frac{\partial}{\partial\theta}$ is such a vector field, we can define a diffeomorphism $\psi : S^1 \to S^1$ via $\psi(\theta) = a\int_0^\theta \frac{dt}{v(t)}$. Here, $a \in \mathbb{R}$ is chosen such that $\psi(2\pi) = 2\pi$. Then $\psi_*(\xi \circ \psi^{-1})$ is a constant vector field on S^1.

From this observation, one can conclude that any diffeomorphism of S^1 that lies in the image of the exponential map and that does not have any fixed points is conjugate to a rigid rotation of S^1. Hence in order to see that the exponential map is not locally surjective, it is enough to construct diffeomorphisms arbitrarily close to the identity that do not have any fixed points and that are not conjugate to a rigid rotation. For this, one can take diffeomorphisms without fixed points, but which have isolated periodic points, i.e., fixed points for a certain nth iteration of this diffeomorphism. Indeed, if such a diffeomorphism ψ belonged to the image of the exponential map, so would its nth power ψ^n. Then the corresponding vector field defining the ψ^n as the time-one map would either have zeros or be nonvanishing everywhere. In the former case, the n-periodic points of ψ must actually be its fixed points, while in the latter case, the diffeomorphism ψ^n, as well as ψ, would be conjugate to a rigid rotation and hence *all* points of ψ would be n-periodic. Both cases give us a contradiction.

Explicitly, a family of such diffeomorphisms can be constructed as follows: Let us identify S^1 with $\mathbb{R}/2\pi\mathbb{Z}$. Then consider the map $\psi_{n,\epsilon} : x \mapsto x + \frac{2\pi}{n} + \epsilon \sin(nx)$. For ϵ small enough, this is indeed a diffeomorphism of S^1. Furthermore, by choosing n large and ϵ small, the diffeomorphisms $\psi_{n,\epsilon}$ can be made arbitrarily close to the identity while having no fixed points. Finally, for $\epsilon \neq 0$, $\psi_{n,\epsilon}$ cannot be conjugate to a rigid rotation. If it were conjugate to a rotation, it would have to be the rotation $\psi_{n,0}$, since $\psi^n_{n,\epsilon}(0) = 0$. But in this case, we would have $\psi^n_{n,\epsilon} = \mathrm{id}$, which is not true for $\epsilon \neq 0$. □

1.4 Abstract Lie Algebras

As we have seen in the last section, the Lie bracket of two left-invariant vector fields $\widetilde{X}$ and $\widetilde{Y}$ on a Lie group G defines a bilinear map $[\,.\,,.\,] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ of the Lie algebra of G that is antisymmetric in X and Y and satisfies the Jacobi identity (1.2). These properties can be taken as the definition of an abstract Lie algebra:

Definition 1.24 An *(abstract) Lie algebra* is a real or complex vector space $\mathfrak{g}$ together with a bilinear map $[\,.\,,.\,] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ (the Lie bracket) that is antisymmetric in X and Y and that satisfies the Jacobi identity

$$[[X,Y],Z] + [[Z,X],Y] + [[Y,Z],X] = 0\,. \tag{1.2}$$

All the Lie algebras we have encountered so far as accompanying the corresponding Lie groups can also be regarded by themselves, i.e., as abstract Lie algebras. A famous theorem of Sophus Lie states that every finite-dimensional (abstract) Lie algebra $\mathfrak{g}$ is the Lie algebra of some Lie group G. In infinite dimensions this is no longer true in general.

Example 1.25 ([205, 207]) To illustrate the failure of Lie's theorem in an infinite-dimensional context, consider the Lie algebra of complex vector fields on the circle $\mathrm{Vect}^{\mathbb{C}}(S^1) = \mathrm{Vect}(S^1) \otimes \mathbb{C}$. Let us show that this Lie algebra cannot be the Lie algebra of any Lie group. First note that $\mathrm{Vect}^{\mathbb{C}}(S^1)$ contains as a subalgebra the Lie algebra $\mathrm{Vect}(S^1)$ of real vector fields on the circle, which is the Lie algebra of the group $\mathrm{Diff}(S^1)$.

Let G_1 denote the group $\mathrm{PSL}(2,\mathbb{R})$ and let G_k denote the k-fold covering of G_1. The group G_2 is isomorphic to $\mathrm{SL}(2,\mathbb{R})$, while for $k > 2$ it is known that the groups G_k have no matrix realization. The group $\mathrm{Diff}(S^1)$ contains each G_k as a subgroup. Namely, G_k is the subgroup corresponding to the Lie subalgebra $\mathfrak{g}_k$ spanned by the vector fields

$$\frac{\partial}{\partial\theta},\ \sin(k\theta)\frac{\partial}{\partial\theta},\ \cos(k\theta)\frac{\partial}{\partial\theta}\,.$$

(Note that each $\mathfrak{g}_k$ is isomorphic to $\mathfrak{sl}(2,\mathbb{R})$.)

Now suppose that there exists a complexification of the group $\mathrm{Diff}(S^1)$, i.e., a Lie group G corresponding to the complex Lie algebra $\mathrm{Vect}^{\mathbb{C}}(S^1)$. Such a group G would have to contain the complexifications of all the groups G_k. However, for $k > 2$ the groups G_k do not admit complexifications: the only complex groups corresponding to the Lie algebra $\mathfrak{sl}(2,\mathbb{C})$ are $\mathrm{SL}(2,\mathbb{C})$ and $\mathrm{PSL}(2,\mathbb{C})$.

More precisely, if the complex Lie group G existed, the real subgroups G_k would belong to the complex subgroups of G corresponding to complex subalgebras $\mathfrak{g}_k^{\mathbb{C}} \simeq \mathfrak{sl}(2,\mathbb{C})$. But these complex subgroups have to be isomorphic either to $\mathrm{SL}(2,\mathbb{C})$, which contains only $\mathrm{SL}(2,\mathbb{R}) = G_2$, or to $\mathrm{PSL}(2,\mathbb{C})$, which contains only $\mathrm{PSL}(2,\mathbb{R}) = G_1$. Thus the complex group G containing all G_k cannot exist, and hence there is no Lie group for the Lie algebra $\mathrm{Vect}^{\mathbb{C}}(S^1)$.

Lie algebra homomorphisms are defined in the usual way: A map $\rho : \mathfrak{g} \to \mathfrak{h}$ between two Lie algebras is a Lie algebra homomorphism if it satisfies $\rho([X,Y]) = [\rho(X),\rho(Y)]$ for all X, $Y \in \mathfrak{g}$. We will also need another important class of maps between Lie algebras called derivations:

Definition 1.26 A linear map $\delta : \mathfrak{g} \to \mathfrak{g}$ of a Lie algebra $\mathfrak{g}$ to itself is called a *derivation* if it satisfies

$$\delta([X,Y]) = [\delta(X),Y] + [X,\delta(Y)]$$

for all X, $Y \in \mathfrak{g}$.

Exercise 1.27 Define the map $\mathrm{ad}_X : \mathfrak{g} \to \mathfrak{g}$ associated to a fixed vector $X \in \mathfrak{g}$ via

$$\mathrm{ad}_X(Y) = [X,Y]\,.$$

Show that this is a derivation for any choice of X. (Hint: use the Jacobi identity.)

If a derivation of a Lie algebra $\mathfrak{g}$ can be expressed in the form ad_X for some $X \in \mathfrak{g}$, it is called an *inner derivation*; otherwise, it is called an *outer derivation* of $\mathfrak{g}$.

Exercise 1.28 Let δ be a derivation of a Lie algebra $\mathfrak{g}$, and suppose that $\exp(\delta) = \sum_{i=0}^{\infty} \frac{1}{i!}\delta^i$ makes sense (for example, suppose, the map δ is nilpotent). Show that the map $\exp(\delta)$ is an automorphism of the Lie algebra $\mathfrak{g}$.

Definition 1.29 A *subalgebra* of a Lie algebra $\mathfrak{g}$ is a subspace $\mathfrak{h} \subset \mathfrak{g}$ invariant under the Lie bracket in $\mathfrak{g}$. An *ideal* of a Lie algebra $\mathfrak{g}$ is a subalgebra $\mathfrak{h} \subset \mathfrak{g}$ such that $[X, \mathfrak{h}] \subset \mathfrak{h}$ for all $X \in \mathfrak{g}$.

The importance of ideals comes from the fact that if $\mathfrak{h} \subset \mathfrak{g}$ is an ideal, then the quotient space $\mathfrak{g}/\mathfrak{h}$ is again a Lie algebra.

Exercise 1.30 (i) Show that for an ideal $\mathfrak{h} \subset \mathfrak{g}$ the Lie bracket on $\mathfrak{g}$ descends to a Lie bracket on the quotient space $\mathfrak{g}/\mathfrak{h}$.

(ii) Show that if $\rho : \mathfrak{g} \to \widetilde{\mathfrak{g}}$ is a homomorphism of two Lie algebras, then the kernel $\ker \rho$ of ρ is an ideal in $\mathfrak{g}$.

Definition 1.31 A Lie algebra is *simple* (respectively, *semisimple*) if it does not contain nontrivial ideals (respectively, nontrivial abelian ideals).

Any finite-dimensional semisimple Lie algebra is a direct sum of nonabelian simple Lie algebras.

A group analogue of an ideal is the notion of a normal subgroup. A subgroup $H \subset G$ of a group G is called *normal* if $gHg^{-1} \subset H$ for all $g \in G$. Exercise 1.30 translates directly to normal subgroups.

2 Adjoint and Coadjoint Orbits

Writing out a linear operator in a different basis or a vector field in a different coordinate system has a far-reaching generalization as the adjoint representation for any Lie group. In this section we define the adjoint and coadjoint representations and the corresponding orbits for an arbitrary Lie group.

2.1 The Adjoint Representation

A *representation* of a Lie group G on a vector space V is a linear action φ of the group G on V that is smooth in the sense that the map $G \times V \to V$, $(g, v) \mapsto gv$, is smooth. If V is a real vector space, (V, φ) is called a real representation, and if V is complex, it is a complex representation. (Here V is assumed to be a Fréchet space, and, often, a Hilbert space. In the latter case, the representation is said to be unitary if the inner product on V is invariant under the action of G.)

Every Lie group has two distinguished representations: the adjoint and the coadjoint representations. Since they will play a special role in this book, we describe them in more detail.

Any element $g \in G$ defines an automorphism c_g of the group G by conjugation:

$$c_g : h \in G \mapsto ghg^{-1}.$$

The differential of c_g at the identity $e \in G$ maps the Lie algebra of G to itself and thus defines an element $\mathrm{Ad}_g \in \mathrm{Aut}(\mathfrak{g})$, the group of all automorphisms of the Lie algebra $\mathfrak{g}$.

Definition 2.1 The map $\mathrm{Ad} : G \to \mathrm{Aut}(\mathfrak{g})$, $g \mapsto \mathrm{Ad}_g$ defines a representation of the group G on the space $\mathfrak{g}$ and is called the *group adjoint representation*; see Figure 2.1. The orbits of the group G in its Lie algebra $\mathfrak{g}$ are called the *adjoint orbits* of G.

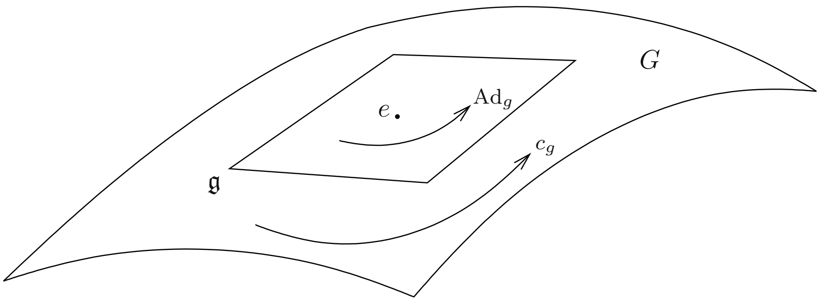


Fig. 2.1. Conjugation c_g on the group G generates the adjoint representation Ad_g on the Lie algebra $\mathfrak{g}$.

The differential of $\mathrm{Ad} : G \to \mathrm{Aut}(\mathfrak{g})$ at the group identity $g = e$ defines a map $\mathrm{ad} : \mathfrak{g} \to \mathrm{End}(\mathfrak{g})$, the *adjoint representation of the Lie algebra* $\mathfrak{g}$.

One can show that the bracket $[\ ,\ ]$ on the space $\mathfrak{g}$ defined via

$$[X, Y] := \mathrm{ad}_X(Y)$$

coincides with the bracket (or commutator) of the corresponding two left-invariant vector fields on the group G and hence with the Lie bracket on $\mathfrak{g}$ defined in Section 1.2.

Example 2.2

- Let $g \in \mathrm{GL}(n, \mathbb{R})$ and $A \in \mathfrak{gl}(n, \mathbb{R})$. Then $\mathrm{Ad}_g A = gAg^{-1}$. Hence the adjoint orbits are given by sets of similar (i.e., conjugate) matrices in $\mathfrak{gl}(n, \mathbb{R})$. The adjoint representation of $\mathfrak{gl}(n, \mathbb{R})$ is given by $\mathrm{ad}_A(B) = [A, B] = AB - BA$.

- The adjoint orbits of $\mathrm{SO}(3, \mathbb{R})$ are spheres centered at the origin of $\mathbb{R}^3 \simeq \mathfrak{so}(3, \mathbb{R})$ and the origin itself.
- The adjoint orbits of $\mathrm{SL}(2, \mathbb{R})$ are contained in the sets of similar matrices. By writing $A = \left(\begin{smallmatrix} a & b \\ c & -a \end{smallmatrix}\right) \in \mathfrak{sl}(2, \mathbb{R})$, one sees that the adjoint orbits lie in the level sets of $\Delta = -(a^2 + bc) = \text{const}$: matrices that are conjugate to each other have the same determinant. Note, however, that not all matrices in $\mathfrak{sl}(2, \mathbb{R})$ that have the same determinant are conjugate. For instance, the matrices with determinant $\Delta = 0$ constitute three different orbits: the origin and two other orbits, cones, passing through the matrices $\left(\begin{smallmatrix} 0 & \pm 1 \\ 0 & 0 \end{smallmatrix}\right)$, respectively. For $\Delta \neq 0$ the $\mathrm{SL}(2, \mathbb{R})$-orbits are either one-sheet hyperboloids or connected components of the two-sheet hyperboloids $a^2 + bc = \text{const}$, since the group $\mathrm{SL}(2, \mathbb{R})$ is connected.
- Let G be the set of orientation-preserving affine transformations of the real line. That is, $G = \{(a, b) \mid a, b \in \mathbb{R}\,,\ a > 0\}$, and $(a, b) \in G$ acts on $x \in \mathbb{R}$ via $x \mapsto ax + b$. The Lie algebra of G is $\mathbb{R}^2$, and its adjoint orbits are the affine lines
$$\{(\alpha, \beta) \in \mathbb{R}^2 \mid \alpha = \text{const} \neq 0, \quad \beta \text{ arbitrary}\}\,,$$
the two rays
$$\{(\alpha, \beta) \in \mathbb{R}^2, \alpha = 0, \beta < 0\} \quad \text{and} \quad \{(\alpha, \beta) \in \mathbb{R}^2, \alpha = 0, \beta > 0\}\,,$$
and the origin $\{(0, 0)\}$; see Figure 2.2.
- Let M be a compact manifold. The adjoint orbits of the current group $\mathrm{GL}(n, \mathbb{C})^M$ in its Lie algebra $\mathfrak{gl}(n, \mathbb{C})^M$ are given by fixing the (smoothly dependent) Jordan normal form of the current at each point of the manifold M.
- Let M be a compact manifold. The adjoint representation of $\mathrm{Diff}(M)$ on $\mathrm{Vect}(M)$ is given by coordinate changes of the vector field: for a $\phi \in \mathrm{Diff}(M)$ one has $\mathrm{Ad}_\phi : v \mapsto \phi_* v \circ \phi^{-1}$. The adjoint representation of $\mathrm{Vect}(M)$ on itself is given by the negative of the usual Lie bracket of vector fields: $\mathrm{ad}_v\, w = \frac{\partial v}{\partial x} w(x) - \frac{\partial w}{\partial x} v(x)$ in any local coordinate x.

Exercise 2.3 Verify the latter formula for the action of $\mathrm{Diff}(M)$ on $\mathrm{Vect}(M)$ from the definition of the group adjoint action. (Hint: express the diffeomorphisms corresponding to the vector fields $v(x)$ and $w(x)$ in the form
$$g(t): \quad x \mapsto x + tv(x) + o(t), \qquad h(s): \quad x \mapsto x + sw(x) + o(s), \qquad t, s \to 0,$$
and find the first several terms of $g(t)h(s)g^{-1}(t)$.)

2.2 The Coadjoint Representation

The dual object to the adjoint representation of a Lie group G on its Lie algebra $\mathfrak{g}$ is called the coadjoint representation of G on $\mathfrak{g}^*$, the dual space to $\mathfrak{g}$.

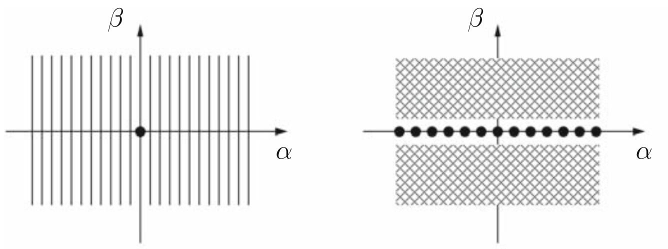


Fig. 2.2. Adjoint and coadjoint orbits of the group of affine transformations on the line.

Definition 2.4 The *coadjoint representation* Ad^* of the group G on the space $\mathfrak{g}^*$ is the dual of the adjoint representation. Let $\langle\ ,\ \rangle$ denote the pairing between $\mathfrak{g}$ and its dual $\mathfrak{g}^*$. Then the *coadjoint action of the group* G on the dual space $\mathfrak{g}^*$ is given by the operators $\mathrm{Ad}_g^* : \mathfrak{g}^* \to \mathfrak{g}^*$ for any $g \in G$ that are defined by the relation

$$\langle \mathrm{Ad}_g^*(\xi), X\rangle := \langle \xi, \mathrm{Ad}_{g^{-1}}(X)\rangle \tag{2.3}$$

for all ξ in $\mathfrak{g}^*$ and $X \in \mathfrak{g}$. The orbits of the group G under this action on $\mathfrak{g}^*$ are called the *coadjoint orbits* of G.

The differential $\mathrm{ad}^* : \mathfrak{g} \to \mathrm{End}(\mathfrak{g}^*)$ of the group representation $\mathrm{Ad}^* : G \to \mathrm{Aut}(\mathfrak{g}^*)$ at the group identity $e \in G$ is called the *coadjoint representation of the Lie algebra* $\mathfrak{g}$. Explicitly, at a given vector $Z \in \mathfrak{g}$ it is defined by the relation

$$\langle \mathrm{ad}_Z^*(\xi), X\rangle = -\langle \xi, \mathrm{ad}_Z(X)\rangle.$$

Remark 2.5 The dual space of a Fréchet space is not necessarily again a Fréchet space. In this case, instead of considering the full dual space to an infinite-dimensional Lie algebra $\mathfrak{g}$, we will usually confine ourselves to considering only appropriate "smooth duals," the functionals from a certain G-invariant Fréchet subspace $\mathfrak{g}_s^* \subset \mathfrak{g}^*$. Natural smooth duals will be different according to the type of the infinite-dimensional groups considered, but they all have a (weak) nondegenerate pairing with the corresponding Lie algebra $\mathfrak{g}$ in the following sense: for every nonzero element $X \in \mathfrak{g}$, there exists some element $\xi \in \mathfrak{g}_s^*$ such that $\langle \xi, X\rangle \neq 0$, and the other way around. This ensures that the coadjoint action is uniquely fixed by equation (2.3). The pair $(\mathfrak{g}_s^*, \mathrm{Ad}^*|_{\mathfrak{g}_s^*})$ is called the regular (or smooth) part of the coadjoint representation of G, and, abusing notations, we will usually skip the index s.

Example 2.6

- In the first three cases of Example 2.2, there exists a G-invariant inner product on $\mathfrak{g}$ that induces an isomorphism between $\mathfrak{g}$ and $\mathfrak{g}^*$ respecting the group actions. Hence the adjoint and coadjoint representations of the groups G are isomorphic, and the coadjoint orbits coincide with the adjoint ones.

- The group of affine transformations of the real line in Example 2.2 has two 2-dimensional coadjoint orbits, the upper and lower half-planes in $\mathbb{R}^2$, and a set of zero-dimensional orbits, namely, the points $(\alpha, 0)$ for each $\alpha \in \mathbb{R}$ (see Figure 2.2).
- For a compact manifold M with some fixed volume form $d\mathrm{Vol}$, we can define a nondegenerate G^M-invariant product on the current Lie algebra $\mathfrak{gl}(n,\mathbb{R})^M$ by setting

$$\langle X, Y \rangle = \int_M \operatorname{tr}(X(x) \cdot Y(x))\, d\mathrm{Vol}(x)$$

for $X, Y \in \mathfrak{gl}(n,\mathbb{R})^M$. This inner product can be used to identify the current algebra $\mathfrak{g}^M$ with a subspace in its dual $\mathfrak{g}^*_s \subset \mathfrak{g}^*$. The space $\mathfrak{g}^*_s$ is called the smooth (or regular) part of $\mathfrak{g}^*$. Thanks to the nondegenerate pairing, the smooth part of the coadjoint representation of G^M is isomorphic to the adjoint representation.

Note that each of the finite-dimensional coadjoint orbits above is even-dimensional. This is a consequence of the general fact that coadjoint orbits are symplectic manifolds, which we discuss later.

Remark 2.7 In what follows we pay particular attention to the structure and description of coadjoint orbits of infinite-dimensional Lie groups. We are interested in coadjoint orbits mostly because they appear as natural phase spaces of dynamical systems. Another reason to study coadjoint orbits comes from the orbit method. This is a general principle due to A. Kirillov, which asserts that the information on the set of unitary representations of a Lie group G is contained in the group coadjoint orbits. This method has become a powerful tool in the study of Lie groups and it has been worked out in detail for large classes of finite-dimensional Lie groups such as nilpotent and compact Lie groups (see [206]). In infinite dimensions, the correspondence between coadjoint orbits and unitary representations has been fully understood only for certain types of groups, e.g., for affine Lie groups (cf. [132, 322, 385]), although there are some indications that it works for other classes as well.

3 Central Extensions

In this section we collect several basic facts about central extensions of Lie groups and Lie algebras. One can think of a central extension of a Lie group G as a new bigger Lie group $\widetilde{G}$ fibered over the initial group G in such a way that the fiber over the identity $e \in G$ lies in the center of $\widetilde{G}$.

Central extensions of Lie groups appear naturally in representation theory and quantum mechanics when one lifts a group projective representation to an ordinary one: one often needs to pass to a central extension of the group to be able to do this. For us the main advantage of these extensions is that for many infinite-dimensional groups their central extensions have simpler and "more regular" structure of the coadjoint orbits, as well as more interesting dynamical systems related to them.

3.1 Lie Algebra Central Extensions

Definition 3.1 A *central extension of a Lie algebra* $\mathfrak{g}$ by a vector space $\mathfrak{n}$ is a Lie algebra $\widetilde{\mathfrak{g}}$ whose underlying vector space $\widetilde{\mathfrak{g}} = \mathfrak{g} \oplus \mathfrak{n}$ is equipped with the following Lie bracket:

$$[(X,u),(Y,v)]^{\sim} = ([X,Y], \omega(X,Y))$$

for some continuous bilinear map $\omega : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{n}$. (Note that ω depends only on X and Y, but not on u and v, which means that the extension is *central*: the space $\mathfrak{n}$ belongs to the center of the new Lie algebra, i.e., it commutes with all of $\widetilde{\mathfrak{g}}$: $[(0,u),(Y,v)] = 0$ for all $Y \in \mathfrak{g}$ and $u, v \in \mathfrak{n}$.) The skew symmetry and the Jacobi identity for the new Lie bracket on $\widetilde{\mathfrak{g}}$ are equivalent to the following conditions on the map ω. Such a map $\omega : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{n}$ has to be a 2-*cocycle on the Lie algebra* $\mathfrak{g}$, i.e., ω has to be bilinear and antisymmetric, and it has to satisfy the *cocycle identity*

$$\omega([X,Y],Z) + \omega([Z,X],Y) + \omega([Y,Z],X) = 0$$

for any triple of elements $X, Y, Z \in \mathfrak{g}$. (Here and below we always require Lie algebra cocycles to be continuous maps.)

A 2-cocycle ω on $\mathfrak{g}$ with values in $\mathfrak{n}$ is called a 2-*coboundary* if there exists a linear map $\alpha : \mathfrak{g} \to \mathfrak{n}$ such that $\omega(X,Y) = \alpha([X,Y])$ for all $X, Y \in \mathfrak{g}$. One can easily see that the central extension defined by such a 2-coboundary becomes the trivial extension by the zero cocycle after the change of coordinates $(X,u) \mapsto (X, u - \alpha(X))$.

Hence in describing different central extensions we are interested only in the 2-cocycles modulo 2-coboundaries, i.e., in the *second cohomology* $H^2(\mathfrak{g};\mathfrak{n})$ of the Lie algebra $\mathfrak{g}$ with values in $\mathfrak{n}$: $H^2(\mathfrak{g};\mathfrak{n}) = \mathcal{Z}(\mathfrak{g};\mathfrak{n})/\mathcal{B}(\mathfrak{g};\mathfrak{n})$, where $\mathcal{Z}(\mathfrak{g};\mathfrak{n})$ is the vector space of all 2-cocycles on $\mathfrak{g}$ with values in $\mathfrak{n}$, and $\mathcal{B}(\mathfrak{g};\mathfrak{n})$ is the subspace of 2-coboundaries.

Remark 3.2 A central extension of a Lie algebra $\mathfrak{g}$ by an abelian Lie algebra $\mathfrak{n}$ can be defined by the exact sequence

$$\{0\} \longrightarrow \mathfrak{n} \longrightarrow \widetilde{\mathfrak{g}} \longrightarrow \mathfrak{g} \longrightarrow \{0\}$$

of Lie algebras such that $\mathfrak{n}$ lies in the center of $\widetilde{\mathfrak{g}}$. A morphism of two central extensions is a pair (ν, μ) of Lie algebra homomorphisms $\nu : \mathfrak{n} \to \mathfrak{n}'$ and $\mu : \widetilde{\mathfrak{g}} \to \widetilde{\mathfrak{g}}'$ such that the following diagram is commutative:

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \mathfrak{n} & \longrightarrow & \widetilde{\mathfrak{g}} & \xrightarrow{\pi} & \mathfrak{g} & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle \nu} & & \downarrow{\scriptstyle \mu} & & \downarrow{\scriptstyle \mathrm{id}} & & \\ 0 & \longrightarrow & \mathfrak{n}' & \longrightarrow & \widetilde{\mathfrak{g}}' & \xrightarrow{\pi'} & \mathfrak{g} & \longrightarrow & 0. \end{array} \tag{3.4}$$

Two extensions are said to be equivalent if the map μ is an isomorphism and $\nu = \mathrm{id}$.

Exercise 3.3 Prove the following equivalence:

Proposition 3.4 *There is a one-to-one correspondence between the equivalence classes of central extensions of* $\mathfrak{g}$ *by* $\mathfrak{n}$ *and the elements of* $H^2(\mathfrak{g};\mathfrak{n})$.

Example 3.5 Consider the abelian Lie algebra $\mathfrak{g} = \mathbb{R}^2$, and let $\omega \in \Lambda^2(\mathbb{R}^2)$ be an arbitrary skew-symmetric bilinear form on $\mathbb{R}^2$. Then ω defines a 2-cocycle on $\mathbb{R}^2$ with values in $\mathbb{R}$ (in this case, the cocycle condition is trivial, since $\mathfrak{g}$ is abelian). The resulting central extension is $\widetilde{\mathfrak{g}} = \mathbb{R}^2 \oplus \mathbb{R}$ with Lie bracket $[(v_1,h_1),(v_2,h_2)] = (0,\omega(v_1,v_2))$. Moreover, since $\mathfrak{g}$ is abelian, $\mathcal{B}(\mathfrak{g};\mathbb{R}) = \{0\}$ whence $H^2(\mathfrak{g};\mathbb{R}) = \Lambda^2(\mathbb{R}^2) \cong \mathbb{R}$. Note that all $\omega \neq 0 \in \Lambda^2(\mathbb{R}^2)$ lead to isomorphic Lie algebras. The algebra $\widetilde{\mathfrak{g}}$ with a nonzero ω, i.e., a representative of this isomorphism class, is called the three-dimensional *Heisenberg algebra*.

By taking a nondegenerate skew-symmetric form ω in $\mathbb{R}^{2n}$, we can define in the same way the $(2n+1)$-dimensional Heisenberg algebra.

An infinite-dimensional analogue of the Heisenberg algebra is as follows. Consider the space $\mathfrak{g} = \{f \in C^\infty(S^1) \mid \int_{S^1} f\, d\theta = 0\}$ of smooth functions on the circle with zero mean and regard it as an abelian Lie algebra. Define the 2-cocycle by $\omega(f,g) = \int_{S^1} f'g\, d\theta$. (One can view this algebra and the corresponding cocycle as the "limit" $n \to \infty$ of the example above by considering the functions in Fourier components.)

Exercise 3.6 Check the skew-symmetry and the cocycle identity for $\omega(f,g)$.

Definition 3.7 A central extension $\widetilde{\mathfrak{g}}$ of $\mathfrak{g}$ is called *universal* if for any other central extension $\widetilde{\mathfrak{g}}'$, there is a unique morphism $\widetilde{\mathfrak{g}} \to \widetilde{\mathfrak{g}}'$ of the central extensions. If it exists, the universal central extension of a Lie algebra $\mathfrak{g}$ is unique up to isomorphism.

Remark 3.8 A sufficient condition for a Lie algebra $\mathfrak{g}$ to have a universal central extension is that $\mathfrak{g}$ be perfect, i.e., that it coincide with its own derived algebra: $\mathfrak{g} = [\mathfrak{g},\mathfrak{g}]$ (see, e.g., [276]). Any finite-dimensional semisimple Lie algebra is perfect. The universal central extension of a semisimple Lie algebra $\mathfrak{g}$ coincides with $\mathfrak{g}$ itself: such algebras do not admit nontrivial central extensions.

No abelian Lie algebra is perfect. Nevertheless, abelian Lie algebras can still have universal central extensions: for instance, the three-dimensional Heisenberg algebra is the universal central extension of the abelian algebra $\mathbb{R}^2$.

Example 3.9 Let M be a finite-dimensional manifold. One can show that the Lie algebra $\mathrm{Vect}(M)$ of vector fields on M is perfect. The universal central extension of $\mathrm{Vect}(M)$ for the case $M = S^1$ is called the Virasoro algebra, and we describe it in detail in Section 2 of Chapter II.

Example 3.10 For a simple Lie algebra $\mathfrak{g}$ and any n-dimensional compact manifold M, the current Lie algebra $\mathfrak{g}^M$ is perfect. (More generally, for any perfect finite-dimensional Lie algebra $\mathfrak{g}$ the Lie algebra $\mathfrak{g}^M$ is perfect.) Its universal central extension $\widetilde{\mathfrak{g}}^M$ can be constructed as follows. Let $\langle\ ,\ \rangle$ be a nondegenerate symmetric invariant bilinear form on $\mathfrak{g}$, where the invariance means that $\langle [A,B],C\rangle = \langle A,[B,C]\rangle$ for all A, B, $C \in \mathfrak{g}$. Denote by $\Omega^1(M)$ the set of 1-forms on M and let $d\Omega^0(M)$ be the subset of exact 1-forms. Now we define the 2-cocycle ω on $\mathfrak{g}^M$ with values in $\Omega^1(M)/d\Omega^0(M)$ via

$$\omega(X,Y) := \langle X, dY\rangle\,,$$

where X, $Y \in \mathfrak{g}^M$. The antisymmetry of ω is immediate, while the cocycle identity follows from the Jacobi identity in $\mathfrak{g}^M$ and the invariance of the bilinear form. So ω defines a central extension of $\mathfrak{g}^M$. For a proof of universality of this central extension see, e.g., [322, 247].

In the case of $M = S^1$ the corresponding space $\Omega^1(S^1)/d\Omega^0(S^1)$ is one-dimensional. The current algebra on S^1 is called the loop algebra associated to $\mathfrak{g}$, and it has the universal central extension by the $\mathbb{R}$- (or $\mathbb{C}$)-valued 2-cocycle

$$\omega(X,Y) := \int_{S^1} \langle X, dY\rangle\,.$$

We discuss loop algebras and their generalizations in detail in Sections 1 and 5 of Chapter II.

3.2 Central Extensions of Lie Groups

Central extensions of Lie groups can be defined similarly to those of Lie algebras. However, unlike the case of Lie algebras, not all group extensions can be described explicitly by cocycles. This is why we start with the alternative definition of the extensions via exact sequences.

Definition 3.11 A *central extension* $\widetilde{G}$ of a Lie group G by an abelian Lie group H is an exact sequence of Lie groups

$$\{e\} \to H \to \widetilde{G} \to G \to \{e\}$$

such that the image of H lies in the center of $\widetilde{G}$. (Here $\{e\}$ is the trivial group containing only the identity element.) Morphisms and equivalence of two central extensions are defined analogously to the case of Lie algebras.

If the central extension $\widetilde{G}$ is topologically a direct product of G and H, $\widetilde{G} = G \times H$ (or, equivalently, if there is a smooth section in the principal H-bundle $\widetilde{G} \to G$),[4] one can define the multiplication in $\widetilde{G}$ as follows:

$$(g_1, h_1) \cdot (g_2, h_2) = (g_1 g_2, \gamma(g_1, g_2) h_1 h_2)$$

for a smooth map $\gamma : G \times G \to H$, which is similar to the case of Lie algebra central extensions. The associativity of this multiplication corresponds to the so-called *group cocycle identity* on the map γ.

Definition 3.12 Let G and H be Lie groups and suppose H is abelian. A smooth map $\gamma : G \times G \to H$ that satisfies

$$\gamma(g_1 g_2, g_3)\gamma(g_1, g_2) = \gamma(g_1, g_2 g_3)\gamma(g_2, g_3)$$

is called a smooth *group* 2*-cocycle* on G with values in H.

A smooth 2-cocycle on G with values in H is called a 2*-coboundary* if there exists a smooth map $\lambda : G \to H$ such that $\gamma(g_1, g_2) = \lambda(g_1)\lambda(g_2)\lambda(g_1 g_2)^{-1}$. As before, the group 2-coboundaries correspond to the trivial group extensions, after a possible change of coordinates (more precisely, of the trivializing section for $\widetilde{G} \to G$). Similarly, two group 2-cocycles define isomorphic extensions if they differ by a 2-coboundary. This explains the following fact.

Proposition 3.13 *There is a one-to-one correspondence between the set of central extensions of G by H that admit a smooth section and the elements in the second cohomology group $H^2(G, H) := Z(G, H)/B(G, H)$. Here $Z(G, H)$ and $B(G, H)$ denote respectively the sets of smooth 2-cocycles and 2-coboundaries on G, with the natural abelian group structure.*

However, in contrast to the case of Lie algebras, there exist central extensions of Lie groups that do not admit a smooth section, and hence cannot be defined by smooth 2-cocycles. We will encounter examples for such groups in Chapter II.

A central extension of a Lie group G always defines a central extension of the corresponding Lie algebra. The converse need not be true: the existence of a Lie group for a given Lie algebra is not guaranteed in infinite dimensions. Instead, one says that a central extension $\widetilde{\mathfrak{g}}$ of a Lie algebra $\mathfrak{g}$ lifts to the group level if there exists a central extension $\widetilde{G}$ of the group G whose Lie algebra is given by $\widetilde{\mathfrak{g}}$. If the group central extension $\widetilde{G}$ by H is defined by a group 2-cocycle γ, one can recover the Lie algebra 2-cocycle defining the corresponding central extension $\widetilde{\mathfrak{g}}$ of the Lie algebra $\mathfrak{g}$ directly from the group cocycle γ by appropriate differentiation.

[4] We always require central extensions of Lie groups to have smooth local sections, in order to secure the existence of a continuous linear section for the corresponding Lie algebra extensions.

Proposition 3.14 *Let H be an abelian Lie group with a Lie algebra $\mathfrak{h}$, and let γ be an H-valued 2-cocycle on G defining a central extension $\widetilde{G}$. Then the $\mathfrak{h}$-valued 2-cocycle ω defining the corresponding central extension $\widetilde{\mathfrak{g}}$ of the Lie algebra $\mathfrak{g}$ is given by*

$$\omega(X,Y) = \frac{d^2}{dt\,ds}\Big|_{t=0,s=0}\gamma(g_t, h_s) - \frac{d^2}{dt\,ds}\Big|_{t=0,s=0}\gamma(h_s, g_t)\,,$$

where g_t is a smooth curve in G such that $\frac{d}{dt}|_{t=0}g_t = X$, and h_s is a smooth curve in G such that $\frac{d}{ds}|_{s=0}h_s = Y$.

Exercise 3.15 Prove the above proposition.

Example 3.16 Let G be $\mathbb{R}^2 = \{(a,b)\}$ with the natural abelian group structure. The three-dimensional *Heisenberg group* $\widetilde{G}$ can be defined as the following matrix group:

$$\widetilde{G} = \left\{ \begin{pmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix} \mid a, b, c \in \mathbb{R} \right\},$$

and it is a central extension of the group G. One verifies directly that the central extension is defined via the $\mathbb{R}$-valued group 2-cocycle γ given by $\gamma((a,b),(a',b')) = ab'$. Using Proposition 3.14, we see that the infinitesimal form of the cocycle γ is given by

$$\omega((A,B),(A',B')) = AB' - A'B,$$

so that the Lie algebra of $\widetilde{G}$ is the three-dimensional Heisenberg algebra discussed in Example 3.5.

4 The Euler Equations for Lie Groups

The Euler equations form a class of dynamical systems closely related to Lie groups and to the geometry of their coadjoint orbits. To describe them we start with generalities on Poisson structures and Hamiltonian systems, before bridging them to Lie groups. Although the manifolds considered in this section are finite-dimensional, we will see later in the book that most of the notions and formulas discussed here are applicable in the infinite-dimensional context (where the dual $\mathfrak{g}^*$ of a Lie algebra $\mathfrak{g}$ stands for its smooth dual).

4.1 Poisson Structures on Manifolds

Definition 4.1 A *Poisson structure* on a manifold M is a bilinear operation on functions

$$\{\ ,\ \}:\ \ C^\infty(M) \times C^\infty(M) \to C^\infty(M)$$

satisfying the following properties:

(i) antisymmetry:

$$\{f, g\} = -\{g, f\},$$

(ii) the Jacobi identity:

$$\{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = 0, \quad \text{and}$$

(iii) the Leibniz identity:

$$\{f, gh\} = \{f, g\}h + \{f, h\}g$$

for any functions $f, g, h \in C^\infty(M)$.

The first two properties mean that a Poisson structure defines a Lie algebra structure $\{\ ,\ \}$ on the space $C^\infty(M)$ of smooth functions on M, while the third property implies that $\{f, \cdot\} : C^\infty(M) \to C^\infty(M)$ is a derivation for any function $f \in C^\infty(M)$. Since each derivation on the space of functions is the Lie derivative along an appropriate vector field, the Poisson structure can be thought of as a map from functions to the corresponding vector fields on the manifold:

Definition 4.2 Let $H : M \to \mathbb{R}$ be any smooth function on a Poisson manifold M. Such a function H defines a vector field ξ_H on M by $L_{\xi_H} g = \{H, g\}$ for any test function $g \in C^\infty(M)$. The vector field ξ_H is called the *Hamiltonian field* corresponding to the *Hamiltonian function* H with respect to the Poisson bracket $\{\ ,\ \}$.

We call a function $F : M \to \mathbb{R}$ a *Casimir function* on a Poisson manifold M if it generates the zero Hamiltonian field, i.e., if the Poisson bracket of the function F with any other function vanishes everywhere on M.

Remark 4.3 Let M be a manifold with a Poisson structure $\{\ ,\ \}$, and we fix some point $m \in M$. All Hamiltonian vector fields on M evaluated at the point $m \in M$ span a subspace of the tangent space $T_m M$. Thus, the Poisson structure defines a distribution of such subspaces on the manifold M (i.e., a subbundle of the tangent bundle TM) by varying the point m.[5] Note that the dimension of this distribution can differ from one point to another. This distribution is integrable, according to the Frobenius theorem, since the commutator of two Hamiltonian vector fields is again Hamiltonian. Therefore, it gives rise to a (possibly singular) *foliation* of the Poisson manifold M [384].

---

[5] Alternatively, a Poisson structure can be defined by specifying a *bivector field* $\varPi$ on the manifold M, i.e., a section of $TM^{\wedge 2}$:

$$\{f, g\} = \varPi(df, dg).$$

Such a bivector field defines a distribution on M as the images of the map $\varPi : T^*M \to TM$.

The leaves of this foliation are called *symplectic leaves*. (In short, two points belong to the same symplectic leaf if they can be joined by a path whose velocity at any point is a Hamiltonian vector.)

Definition 4.4 A pair (N, ω) consisting of a manifold N and a 2-form ω on N is called a *symplectic manifold* if ω is closed ($d\omega = 0$) and nondegenerate. (In the case of infinite dimensions, the form ω is required to be nondegenerate in the sense that for each point $p \in N$ and any nonzero vector $X \in T_pN$, there exists another vector $Y \in T_pM$ such that $\omega_p(X, Y) \neq 0$.) The 2-form ω is called the *symplectic form* on the manifold N.

The reason for the name "symplectic leaves" in Remark 4.3 is that one can define a symplectic 2-form ω on each leaf. It suffices to fix its values on Hamiltonian vector fields ξ_f and ξ_g at any point:

$$\omega(\xi_f, \xi_g) := \{f, g\},$$

since the tangent space of each leaf is generated by Hamiltonian fields.

Exercise 4.5 Show that the 2-form ω defined on the leaf through a point $m \in M$ is closed and nondegenerate.

Note that Casimir functions, by definition, are constant on the leaves of the above foliation, and the codimension of generic symplectic leaves on a Poisson manifold M is equal to the number of (locally functionally) independent Casimir functions on M.

Example 4.6 For the Poisson structure

$$\{f, g\} := \frac{\partial f}{\partial x}\frac{\partial g}{\partial y} - \frac{\partial f}{\partial y}\frac{\partial g}{\partial x}$$

in $\mathbb{R}^3 = \{(x, y, z)\}$, its symplectic leaves are the planes $z = \text{const}$. The coordinate function z, or any function $F = F(z)$, is a Casimir function for this Poisson manifold.

Remark 4.7 Locally, a Poisson manifold near any point p splits into the product of a symplectic space and a Poisson manifold whose rank at p is zero [384]. The symplectic space is a neighborhood of the symplectic leaf passing through p, while the Poisson manifold of zero rank represents the transverse Poisson structure at the point p.

Below we will see that this splitting works in many (but not all!) infinite-dimensional examples: Poisson structures can have infinite-dimensional symplectic leaves and finite-dimensional Poisson transversals.

4.2 Hamiltonian Equations on the Dual of a Lie Algebra

Let G be a Lie group (finite- or infinite-dimensional) with Lie algebra $\mathfrak{g}$, and let $\mathfrak{g}^*$ denote (the smooth part of) its dual.

Definition 4.8 The natural *Lie–Poisson* (or *Kirillov–Kostant* Poisson) structure $\{\ ,\ \}_{LP}$ on the dual Lie algebra $\mathfrak{g}^*$,

$$\{\ ,\ \}_{LP} : C^\infty(\mathfrak{g}^*) \times C^\infty(\mathfrak{g}^*) \to C^\infty(\mathfrak{g}^*)\,,$$

is defined via

$$\{f, g\}_{LP}(m) := \langle [df_m,\ dg_m],\ m\rangle$$

for any $m \in \mathfrak{g}^*$ and any two smooth functions f, g on $\mathfrak{g}^*$; see Figure 4.1. (Here df_m is the differential of the smooth function f taken at the point m, understood as an element of the space $\mathfrak{g}$ itself, and $\langle\ ,\ \rangle$ is the natural pairing between the dual spaces $\mathfrak{g}$ and $\mathfrak{g}^*$.)

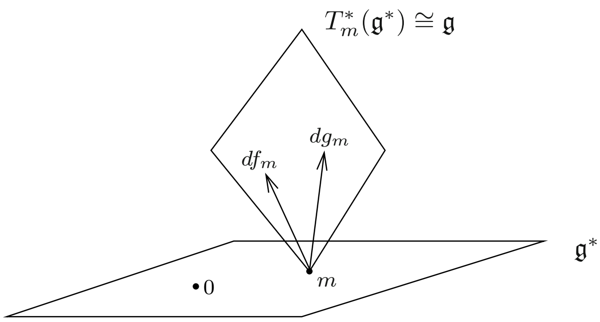


Fig. 4.1. Defining the Lie–Poisson structure: $df_m,\ dg_m \in \mathfrak{g}$, while $m \in \mathfrak{g}^*$.

Proposition 4.9 *The Hamiltonian equation corresponding to a function H and the natural Lie–Poisson structure $\{\ ,\ \}_{LP}$ on $\mathfrak{g}^*$ is given by*

$$\frac{d}{dt} m(t) = -\operatorname{ad}^*_{dH_{m(t)}} m(t).$$

This equation is called the Euler–Poisson equation *on $\mathfrak{g}^*$.*

Proof. Let $f \in C^\infty(\mathfrak{g}^*)$ be an arbitrary function. Then

$$\begin{aligned} L_{\xi_H} f(m) =& \{H, f\}(m) = \langle [dH_m, df_m], m\rangle \\ =& \langle \operatorname{ad}_{dH_m}(df_m), m\rangle = -\langle df_m, \operatorname{ad}^*_{dH_m}(m)\rangle. \end{aligned}$$

Since the Lie derivative of a function f along a vector field is the evaluation of the function's differential df on this vector field, this implies that $\xi_H(m) = -\operatorname{ad}^*_{dH_m}(m)$, which is the assertion. □

Corollary 4.10 *The symplectic leaves of $\{\ ,\ \}_{LP}$ on $\mathfrak{g}^*$ are the coadjoint orbits of G. In particular, all (finite-dimensional) coadjoint orbits have even dimension.*

PROOF. Denote by $\mathcal{O}_m$ the coadjoint orbit through a point $m \in \mathfrak{g}^*$ in the dual space. Let $H \in C^\infty(\mathfrak{g}^*)$ be a function on the dual. For any vector $v \in \mathfrak{g}$ of the Lie algebra, one can represent it as $v = dH_m$ by taking an appropriate function H. Therefore, one can obtain as Hamiltonian vectors $\operatorname{ad}^*_{dH_m}(m)$ at the point m all vectors in the image of $\operatorname{ad}^*_{\mathfrak{g}}(m)$, i.e., all vectors in the tangent space to the orbit $T_m\mathcal{O}_m := T_m(\operatorname{Ad}^*_G(m))$. By definition, all Hamiltonian vectors span the tangent space to any symplectic leaf at each point m, which proves that coadjoint orbits are exactly the symplectic leaves of the Lie–Poisson bracket. □

Corollary 4.11 *Let $A : \mathfrak{g} \to \mathfrak{g}^*$ be an invertible self-adjoint operator.*[6] *For the quadratic Hamiltonian function $H : \mathfrak{g}^* \to \mathbb{R}$ defined by $H(m) := \frac{1}{2}\langle m, A^{-1}m\rangle$ the corresponding Hamiltonian equation is*

$$\frac{d}{dt}m(t) = -\operatorname{ad}^*_{A^{-1}m(t)} m(t)\,. \tag{4.5}$$

Indeed, $dH_m(m) = A^{-1}m$ for any $m \in \mathfrak{g}^*$.

Definition 4.12 An invertible self-adjoint operator $A : \mathfrak{g} \to \mathfrak{g}^*$ defining the quadratic Hamiltonian H is called an *inertia operator on* $\mathfrak{g}$.

4.3 A Riemannian Approach to the Euler Equations

It turns out that the Euler–Poisson equations with quadratic Hamiltonians have a beautiful Riemannian reformulation.

V. Arnold suggested in [12] the following general setup for the Euler equation describing a geodesic flow on an arbitrary Lie group. Consider a (possibly infinite-dimensional) Lie group G, which can be thought of as the configuration space of some physical system. (Examples from [12, 18]: SO(3) for a rigid

[6] Note that one can define the "self-adjointness property" for an operator from a space to its dual, similarly to a self-adjoint operator acting on a given space with respect to a fixed pairing.

body or the group $S\mathrm{Diff}(M)$ of volume-preserving diffeomorphisms for an ideal fluid filling a domain M.) The tangent space at the identity of the Lie group G is the corresponding Lie algebra $\mathfrak{g}$. Fix some (positive definite) quadratic form, the energy, on $\mathfrak{g}$. We consider left (or right) translations of this quadratic form to the tangent space at any point of the group (the "translational symmetry" of the energy). In this way, the energy defines a left- (respectively, right-) invariant Riemannian metric on the group G. The geodesic flow on G with respect to this energy metric represents extremals of the least-action principle, i.e., possible motions of our physical system.[7] To describe a geodesic on the Lie group G with an initial velocity $v(0)$, we transport its velocity vector at any moment t to the identity of the group using the left (respectively, right) translation. This way we obtain the evolution law for $v(t)$ on the Lie algebra $\mathfrak{g}$.

To fix the notation, let $(\ ,\ )$ be some left-invariant metric on the group G. The geodesic flow with respect to this metric is a dynamical system on the tangent bundle TG of the group G. We can pull back this system to the Lie algebra $\mathfrak{g}$ of the group G by left translation. That is, if $g(t)$ is a geodesic in the group G with tangent vector $g'(t)$, then the pullback $v(t) = l^*_{g(t)^{-1}} g'(t)$ is an element of the Lie algebra $\mathfrak{g}$. (In the case of a right-invariant metric, we set $v(t) = r^*_{g(t)^{-1}} g'(t)$.) Hence, the geodesic equations for $g(t)$ give us a dynamical system

$$\frac{d}{dt} v(t) = B(v(t)) \tag{4.6}$$

on the Lie algebra $\mathfrak{g}$ of the group G, where $B : \mathfrak{g} \to \mathfrak{g}$ is a (nonlinear) operator.

Definition 4.13 The dynamical system (4.6) on the Lie algebra $\mathfrak{g}$ describing the evolution of the velocity vector of a geodesic in a left-invariant metric on the Lie group G is called the *Euler (or Euler–Arnold) equation* corresponding to this metric on G.

It turns out that the Euler equation for a Lie group G can be viewed as a Hamiltonian equation on the dual of the Lie algebra $\mathfrak{g}$ in the following way. Observe that the metric $(\ ,\ )_e$ at the identity $e \in G$ defines a nondegenerate bilinear form on the Lie algebra $\mathfrak{g}$, and therefore, it also determines an inertia operator $A : \mathfrak{g} \to \mathfrak{g}^*$ such that $(v, w)_e = \langle A(v), w \rangle$ for all $v, w \in \mathfrak{g}$. This identification $A : \mathfrak{g} \to \mathfrak{g}^*$ allows one to rewrite the Euler equation on the dual space $\mathfrak{g}^*$; see Figure 4.2. Now, setting $m = A(v)$, one can relate the geodesic equation (4.6) on the Lie algebra $\mathfrak{g}$ to the Hamiltonian equation on the dual $\mathfrak{g}^*$ with respect to the Hamiltonian function $H(m) = \frac{1}{2}\langle m, A^{-1} m \rangle$:

[7] Usually, the finite-dimensional examples below are related to the left invariance, while the infinite-dimensional ones with the right invariance of the metric. In particular, for a rigid body one has to consider left translations on SO(3), while for fluids, one must consider the right ones on $S\mathrm{Diff}(M)$.

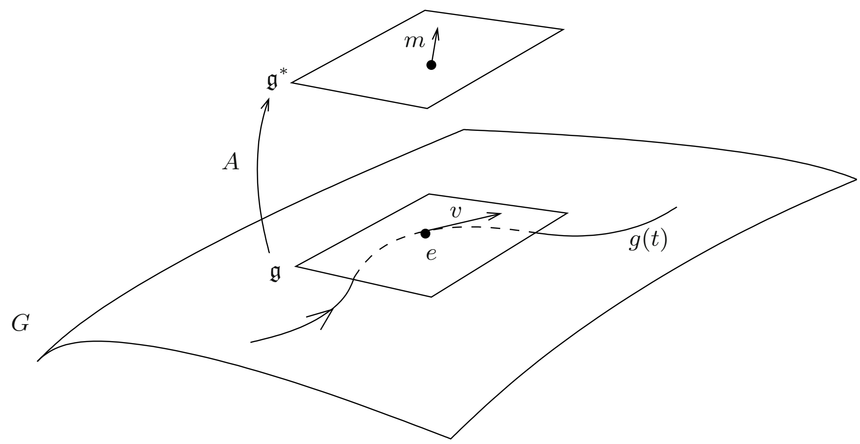


Fig. 4.2. The vector v in the Lie algebra $\mathfrak{g}$ traces the evolution of the velocity vector of a geodesic $g(t)$ on the group. The inertia operator A sends v to a vector m in the dual space $\mathfrak{g}^*$.

Theorem 4.14 (Arnold [12, 13, 18]) *For the left-invariant metric on a group generated by an inertia operator $A : \mathfrak{g} \to \mathfrak{g}^*$, the Euler (or the geodesic) equation (4.6) assumes the form*

$$\frac{d}{dt} m(t) = -\operatorname{ad}^*_{A^{-1}m(t)} m(t)$$

on the dual space $\mathfrak{g}^*$.[8]

We postpone the proof of this theorem until the end of this section.

Remark 4.15 The underlying reason for the Riemannian reformulation is the fact that any geodesic problem in *Riemannian* geometry can be described in terms of *symplectic* geometry. Geodesics on M are extremals of a quadratic Lagrangian on TM (coming from the metric on M). They can also be described by the Hamiltonian flow on T^*M for the quadratic Hamiltonian function obtained from the Lagrangian via the Legendre transform.

If the manifold is a group G with a left-invariant metric, then there exists the group action on the tangent bundle TG, as well as on the cotangent bundle T^*G. The left translations on the group trivialize the cotangent bundle $T^*G \cong G \times \mathfrak{g}^*$ and identify any cotangent space of G with $\mathfrak{g}^*$. By taking the quotient with respect to the group action, from the (symplectic) cotangent bundle T^*G we obtain the dual Lie algebra $\mathfrak{g}^* = T^*G|_e$ equipped with the negative

[8] Note that these signs are different from the conventions in the book [24], since we have used a different definition of Ad^* here (see equation (2.3)) in order to have the group coadjoint representation, rather than the antirepresentation.

of the Lie–Poisson structure (cf. Section 5 below, on symplectic reduction). The Hamiltonian function on T^*G is dual to the Riemannian metric (viewed as a form on TG), and its restriction to $\mathfrak{g}^*$ is the quadratic form $H(m) = \frac{1}{2}\langle m, A^{-1}m\rangle$, where $m \in \mathfrak{g}^*$.

The geodesics of a left-invariant metric on G correspond to the Hamiltonian function $H(m)$ with respect to the standard Lie–Poisson structure.

Remark 4.16 Instead of using a left-invariant metric on G, we could have used a right-invariant one. This changes the signs in the Euler equation, so that one obtains

$$\frac{d}{dt}m = ad^*_{A^{-1}m}(m)\,.$$

Now the geodesics in a right-invariant metric correspond to the Hamiltonian $-H(m)$.

Example 4.17 Let us consider the group SO(3). The Lie algebra $\mathfrak{so}(3)$ of SO(3) can be identified with $\mathbb{R}^3$ such that the Lie bracket on $\mathfrak{g}$ is the cross product on $\mathbb{R}^3$: $[u, v] = v \times u$. Let A be a symmetric nondegenerate 3×3 matrix, which we view as an inertia operator for a left-invariant metric on SO(3). Then by Arnold's theorem, the Euler equation on $\mathfrak{so}(3)^*$ is given by

$$\frac{d}{dt}m = m \times A^{-1}m.$$

For $A = \mathrm{diag}(I_1, I_2, I_3)$ one obtains the classical Euler equations for a rigid body in $\mathbb{R}^3$:

$$\frac{d}{dt}m_i = (I_k^{-1} - I_j^{-1})m_j m_k$$

for (i, j, k) being a cyclic permutation of $(1, 2, 3)$. Similarly, for $G = \mathrm{SO}(n)$, one obtains the Euler equation for a higher-dimensional rigid body (see Remark 4.28 below).

Example 4.18 Many other conservative dynamical systems in mathematical physics also describe geodesic flows on appropriate Lie groups. In Table 4.1 we list several examples of such systems to emphasize the range of applications of this approach. The choice of a group G (column 1) and an energy metric E (column 2) defines the corresponding Euler equations (column 3).

We discuss many of these examples later in the book. There are plenty of other interesting systems that fit into this framework, such as, e.g., the super-KdV equation or gas dynamics. This list is by no means complete, and we refer to [24, 252] for more details.

Group	*Metric*	*Equation*
SO(3)	$\langle\omega, A\omega\rangle$	Euler top
$\mathrm{SO}(3) \ltimes \mathbb{R}^3$	quadratic forms	Kirchhoff equation for a body in a fluid
SO(n)	Manakov′s metrics	n-dimensional top
Diff(S^1)	L^2	Hopf (or, inviscid Burgers) equation
Virasoro	L^2	KdV equation
Virasoro	H^1	Camassa–Holm equation
Virasoro	$\dot{H}^1$	Hunter–Saxton (or Dym) equation
SDiff(M)	L^2	Euler ideal fluid
SDiff(M)	H^1	averaged Euler flow
$\mathrm{SDiff}(M) \ltimes \mathrm{SVect}(M)$	$L^2 + L^2$	Magnetohydrodynamics
Maps(S^1, SO(3))	H^{-1}	Heisenberg magnetic chain

Table 4.1: Euler equations related to various Lie groups.

Now we return to the proof of Arnold's theorem.

Proof of Theorem 4.14. Consider the energy function (or Lagrangian) $L : TG \to \mathbb{R}$ defined by the left-invariant metric $(\ , \ )$ on the group G:

$$L(g, v) = \frac{1}{2}(v, v)_g \, ,$$

where $(\ , \ )_g$ is the metric at the point $g \in G$. Then by definition, a geodesic path $g(t)$ on G satisfies the variational principle

$$\delta \int L(g(t), g'(t))dt = 0 \tag{4.7}$$

with fixed endpoints. (Here and later, δ denotes the variational derivative, and the prime $'$ stands for the time derivative d/dt.)

To simplify the notation, we write $g^{-1}(t)g'(t)$ for $l^*_{g^{-1}(t)}g'(t)$. (If G is a matrix group, this notation agrees with the usual meaning of the expression $g^{-1}(t)g'(t)$ as a matrix product.) Since the metric $(\ , \ )$ on the group G is left-invariant, we can write

$$(g'(t), g'(t))_{g(t)} = (g^{-1}(t)g'(t), g^{-1}(t)g'(t))_e \, .$$

Then we can calculate

$$\delta \int \frac{1}{2}(g^{-1}g', g^{-1}g')_e dt = \int (\delta(g^{-1}g'), g^{-1}g')_e dt \, . \tag{4.8}$$

Note that we have

$$\delta(g^{-1}g') = g^{-1}\delta g' - g^{-1}\delta g g^{-1}g' = (g^{-1}\delta g)' + [g^{-1}g', g^{-1}\delta g] \, ,$$

since

$$(g^{-1}\delta g)' = g^{-1}\delta g' - g^{-1}g'g^{-1}\delta g\,.$$

Thus, the left-hand-side in equation (4.8) becomes

$$\begin{aligned}\int(\delta(g^{-1}g'), g^{-1}g')_e dt &= \int((g^{-1}\delta g)', g^{-1}g')_e dt + \int([g^{-1}g', g^{-1}\delta g], g^{-1}g')_e dt\\ &= -\int(g^{-1}\delta g, (g^{-1}g')')_e dt + \int([g^{-1}g', g^{-1}\delta g], g^{-1}g')_e dt\,,\end{aligned}$$

where we have used integration by parts in the last step. (Since we confined ourselves to variations of the path g with fixed endpoints, we do not pick up any boundary terms in the integration by parts.)

Now set $v(t) := g^{-1}(t)g'(t)$, and let $A : \mathfrak{g} \to \mathfrak{g}^*$ be the inertia operator defined by the metric $(\ ,\ )_e$: $(u, w)_e = \langle u, Aw\rangle$. Then the right-hand side in the latter equation becomes

$$\begin{aligned}-\int(g^{-1}\delta g, (g^{-1}g')')_e dt &+ \int([g^{-1}g', g^{-1}\delta g], g^{-1}g')_e dt\\ &= -\int\langle g^{-1}\delta g, (Av)'\rangle dt + \int\langle \mathrm{ad}_v(g^{-1}\delta g), Av\rangle dt\\ &= -\int\langle g^{-1}\delta g, (Av)'\rangle dt - \int\langle g^{-1}\delta g, \mathrm{ad}_v^*(Av)\rangle dt = 0\,.\end{aligned}$$

This implies

$$(Av)' = -\,\mathrm{ad}_v^*(Av)\,.$$

Rewriting this equation in terms of $m = Av$ finishes the proof of Theorem 4.14. □

4.4 Poisson Pairs and Bi-Hamiltonian Structures

A *first integral* (or a conservation law) for a vector field ξ on a manifold M is a function on M invariant under the flow of this field. In this section we will show that if the vector field ξ is a Hamiltonian vector field with respect to two different Poisson structures on the manifold M that are compatible in a certain sense, there is a way of constructing first integrals for such a field.

Definition 4.19 Two Poisson structures $\{\ ,\ \}_0$ and $\{\ ,\ \}_1$ on a manifold M are said to be *compatible* (or *form a Poisson pair*) if for every $\lambda \in \mathbb{R}$ the linear combination $\{\ ,\ \}_0 + \lambda\{\ ,\ \}_1$ is again a Poisson bracket on M.

A dynamical system $\frac{d}{dt}m = \xi(m)$ on M is called *bi-Hamiltonian* if the vector field ξ is Hamiltonian with respect to both structures $\{\ ,\ \}_0$ and $\{\ ,\ \}_1$.

Our main example of a manifold that admits a Poisson pair is the dual space $\mathfrak{g}^*$ of a Lie algebra $\mathfrak{g}$. One Poisson structure on the space $\mathfrak{g}^*$ is given by the usual Lie–Poisson bracket $\{\ ,\ \}_{LP}$. We can define a second Poisson structure on $\mathfrak{g}^*$ by "freezing" the Lie–Poisson bracket at any point $m_0 \in \mathfrak{g}^*$:

Definition 4.20 The constant Poisson bracket on $\mathfrak{g}^*$ associated to a point $m_0 \in \mathfrak{g}^*$ is the bracket $\{\ ,\ \}_0$ defined on two smooth functions f, g on $\mathfrak{g}^*$ by

$$\{f, g\}_0(m) := \langle [df_m, dg_m], m_0 \rangle \,.$$

The Poisson bracket $\{\ ,\ \}_0$ depends on the *freezing point* $m_0 \in \mathfrak{g}^*$. Note that at the point m_0 itself the two Poisson brackets $\{\ ,\ \}_{LP}$ and $\{\ ,\ \}_0$ coincide. While the symplectic leaves of the Lie–Poisson bracket $\{\ ,\ \}_{LP}$ are the coadjoint orbits $\mathcal{O}_m$ of the Lie group G, the symplectic leaves of the constant bracket $\{\ ,\ \}_0$ are given by all translations of the tangent space $T_{m_0}\mathcal{O}_{m_0}$ to the coadjoint orbit $\mathcal{O}_{m_0}$ through the point m_0 (see Figure 4.3).

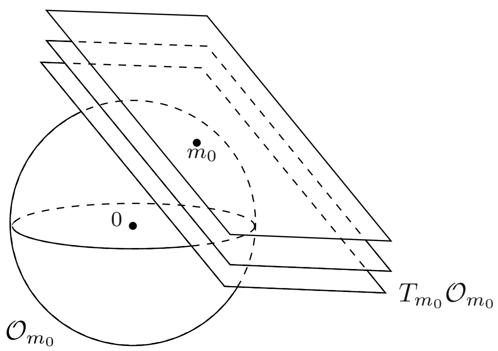


Fig. 4.3. Coadjoint orbit $\mathcal{O}_{m_0}$ through m_0 and leaves of the Poisson bracket frozen at m_0.

Lemma 4.21 *The Poisson brackets $\{\ ,\ \}_{LP}$ and $\{\ ,\ \}_0$ are compatible for every "freezing point" $m_0 \in \mathfrak{g}^*$.*

Proof. We have to check that $\{\ ,\ \}_\lambda = \{\ ,\ \}_{LP} + \lambda\{\ ,\ \}_0$ is a Poisson bracket on $\mathfrak{g}^*$ for all $\lambda \in \mathbb{R}$. The latter is true since $\{\ ,\ \}_\lambda$ is simply the bracket $\{\ ,\ \}_{LP}$ shifted by $-\lambda m_0$. □

Remark 4.22 Explicitly, the Hamiltonian equation on $\mathfrak{g}^*$ with the Hamiltonian function F and computed with respect to the constant Poisson structure frozen at a point $m_0 \in \mathfrak{g}^*$ has the following form:

$$\frac{dm}{dt} = -\mathrm{ad}^*_{dF_m} m_0 \,, \tag{4.9}$$

as a modification of Proposition 4.9 shows.

Going back to the general situation, let $\{\ ,\ \}_0$ and $\{\ ,\ \}_1$ be a Poisson pair on a manifold M. In this case one can generate a bi-Hamiltonian dynamical system by producing a sequence of Hamiltonians in involution, according to the following *Lenard–Magri scheme* [246, 321]. Consider the Poisson bracket $\{\ ,\ \}_\lambda = \{\ ,\ \}_0 + \lambda\{\ ,\ \}_1$ for any λ. Let h_λ be a *Casimir function* on M for this bracket, i.e., a function on the manifold M that is parametrized by λ and satisfies $\{h_\lambda, f\}_\lambda = 0$ for all smooth functions $f \in C^\infty(M)$ and $\lambda \in \mathbb{R}$. Furthermore, suppose that the function h_λ can be expanded into a power series in λ, i.e., that we can write

$$h_\lambda = \sum_{i=0}^{\infty} \lambda^i\, h_i\,, \tag{4.10}$$

where each coefficient h_i is a smooth function on M. Any function h_i defines a Hamiltonian vector field ξ_i on M with respect to the Poisson bracket $\{\ ,\ \}_1$ by setting $\{h_i, f\}_1 = L_{\xi_i} f$ for all $f \in C^\infty(M)$.

Theorem 4.23 *The functions $h_i,\ i = 0, 1, \ldots$ are Hamiltonians of a hierarchy of bi-Hamiltonian systems. In other words, each function h_i generates the Hamiltonian vector field ξ_i on M with respect to the Poisson bracket $\{\ ,\ \}_1$, which is also Hamiltonian for the other bracket $\{\ ,\ \}_0$ with the Hamiltonian function $-h_{i+1}$:*

$$\{h_i, f\}_1 = L_{\xi_i} f = -\{h_{i+1}, f\}_0$$

for any f. Other functions $h_j,\ j \neq i$, are first integrals of the corresponding dynamical systems ξ_i.

In other words, the functions $h_i,\ i = 0, 1, \ldots$ are in involution with respect to each of the two Poisson brackets $\{\ ,\ \}_0$ and $\{\ ,\ \}_1$:

$$\{h_i, h_j\}_k = 0$$

for all $i \neq j$ and for $k = 0, 1$.

Proof. Since h_λ is a Casimir function for the Poisson bracket $\{\ ,\ \}_\lambda$, we have $\{h_\lambda, f\}_\lambda = 0$ for all smooth functions f on M. Substituting for h_λ its power series expansion (4.10), we get

$$0 = \{h_\lambda, f\}_\lambda = \left\{\sum_{i=0}^{\infty} \lambda^i h_i\,, f\right\}_\lambda = \left\{\sum_{i=0}^{\infty} \lambda^i h_i\,, f\right\}_0 + \lambda \left\{\sum_{i=0}^{\infty} \lambda^i h_i\,, f\right\}_1 .$$

Collecting the coefficients at the powers of λ we find that $\{h_0, f\}_0 = 0$ and

$$\{h_i, f\}_0 = -\{h_{i-1}, f\}_1\,.$$

The first identity expresses the fact that h_0 is a Casimir function for the bracket $\{\ ,\ \}_0$. The next one says that the Hamiltonian field for h_1 with respect to $\{\ ,\ \}_0$ coincides with the Hamiltonian field for $-h_0$ and the bracket $\{\ ,\ \}_1$, and so on.

To see that every function h_i is a first integral for the equation generated by h_j with respect to each bracket, we have to show that $\{h_i, h_j\}_k = 0$ for $i \neq j$ and $k = 0, 1$. Indeed, for instance, for $i < j$ and $k = 1$ we have

$$\{h_i, h_j\}_1 = -\{h_i, h_{j+1}\}_0 = \{h_{i-1}, h_{j+1}\}_1 = \cdots = -\{h_0, h_{i+j+1}\}_0 = 0,$$

since h_0 is a Casimir for the bracket $\{.,.\}_0$, i.e., in involution with any function, and in particular, with h_{i+j+1}. □

Remark 4.24 The fact that $\{h_i, h_j\}_k = 0$ for $k = 0, 1$ means that the functions h_j are first integrals of the Hamiltonian vector fields ξ_i. So if the functions h_j are independent, Theorem 4.23 provides us with an infinite list of first integrals for each of the fields ξ_i. In this case one says that the h_i are the Hamiltonians of a hierarchy of bi-Hamiltonian systems. We will treat the KdV equation as a bi-Hamiltonian system from this viewpoint in Section II.2.4.

Exercise 4.25 Suppose that a manifold M admits two compatible Poisson structures $\{\ ,\ \}_0$ and $\{\ ,\ \}_1$. Show that if symplectic leaves of $\{\ ,\ \}_\lambda = \{\ ,\ \}_0 + \lambda\{\ ,\ \}_1$ are of codimension greater than 1, and if there are several independent Casimirs $h_\lambda^{(1)}, h_\lambda^{(2)}, \ldots$, then all the coefficients of their expansions in λ are in mutual involution with respect to both brackets, e.g., $\{h_i^{(1)}, h_j^{(2)}\}_k = 0$.

4.5 Integrable Systems and the Liouville–Arnold Theorem

The more first integrals a dynamical system has, the less chaotically it behaves. For a Hamiltonian system the notion of complete integrability corresponds to the "least chaotic" and "most ordered" structure of its trajectories.

Definition 4.26 A Hamiltonian system on a symplectic $2n$-dimensional manifold M is called *(completely) integrable* if it has n integrals in involution that are functionally independent almost everywhere on M. The Hamiltonian function is one of the above first integrals. (Alternatively, one can avoid specifying which of them is a Hamiltonian and describe an *integrable system* as a set of n functions $f_1, \ldots, f_n$ that are functionally independent almost everywhere and commute pairwise,

$$\{f_i, f_j\} = 0 \quad \text{for all } 1 \leq i, j \leq n\,,$$

with respect to the natural Poisson bracket defined by the symplectic structure on M.)

Example 4.27 Every Hamiltonian system with one degree of freedom is completely integrable, since it always possesses one first integral, the Hamiltonian function itself. This purely dimensional argument implies, for example, that the Euler equation of a three-dimensional rigid body is a completely integrable Hamiltonian system on the coadjoint orbits of SO(3).

Indeed, the configuration space of the Euler top, a three-dimensional rigid body with a fixed point, is the set of all rotations of the Euclidean space, i.e., the Lie group SO(3). The motion of the body is described by the Euler equation on the body angular momentum m in the corresponding phase space, $\mathfrak{so}(3)^*$; see Example 4.17. Note that the conservation of the total momentum $|m|^2$ corresponds to the restriction of the angular momentum evolution to a particular coadjoint orbit, a two-dimensional sphere centered at the origin of $\mathfrak{so}(3)^* \cong \mathbb{R}^3$. Hence the Euler equation for the rigid body is a Hamiltonian system on a two-dimensional symplectic sphere, while the Hamiltonian function is given by the kinetic energy of the body.

Remark 4.28 A more complicated example is a rotation of an n-dimensional rigid body, where the dimensional consideration is not sufficient. Free motions of a body with a fixed point at its mass center are described by the geodesic flow on the group $\mathrm{SO}(n)$ of all rotations of Euclidean space $\mathbb{R}^n$. The group $\mathrm{SO}(n)$ can be regarded as the configuration space of this system. The left-invariant metric on $\mathrm{SO}(n)$ is defined by the quadratic form $-\operatorname{tr}(\omega D\omega)$, where $\omega \in \mathfrak{so}(n)$ is the body's angular velocity and $D = \operatorname{diag}(d_1, \ldots, d_n)$ defines the inertia ellipsoid. The corresponding inertia operator $A : \mathfrak{so}(n) \to \mathfrak{so}(n)^*$ has a very special form: $A(\omega) = D\omega + \omega D$.

Now the evolution of the angular momentum is in the space $\mathfrak{so}(n)^*$, the phase space of the n-dimensional top. The dimension of generic coadjoint orbits in $\mathfrak{so}(n)^*$ is equal to the integer part of $(n-1)^2/2$. Therefore the energy invariance alone is insufficient to guarantee the integrability of the Euler equation for an n-dimensional rigid body. The existence of sufficiently many first integrals and complete integrability in the general n-dimensional case were established by Manakov in [249]. In this paper the argument translation (or freezing) method was discovered and applied to find first integrals in this problem.

We note that the above inertia operators (or equivalently, the corresponding left-invariant metrics) form a variety of dimension n in the $n(n-1)/2$-dimensional space of equivalence classes of symmetric matrices on the Lie algebra $\mathfrak{so}(n)$. For $n > 3$ such quadratic forms are indeed very special in the space of all quadratic forms on this space. The geodesic flow on the group $\mathrm{SO}(n)$ equipped with an arbitrary left-invariant Riemannian metric is, in general, nonintegrable.

Other examples of integrable systems include, for instance, the geodesics on an ellipsoid [281] and the Calogero–Moser systems [280, 66, 67].

The following Liouville–Arnold theorem explains how the sufficient number of first integrals simplifies the Hamiltonian system. Consider a common level set of the first integrals

$$M_c = \{m \in M \mid f_i(m) = c_i,\ i = 1, \ldots, n\}\,.$$

Theorem 4.29 (Liouville–Arnold [11, 18]) *For a compact manifold M, connected components of noncritical common level sets M_c of the n first integrals are n-dimensional tori, while the Hamiltonian system defines a (quasi-) periodic motion on each of them. In a neighborhood of such a component in M there are coordinates $(\varphi_1, \ldots, \varphi_n, I_1, \ldots, I_n)$, where φ_i are angular coordinates along the tori and I_i are first integrals, such that the dynamical system assumes the form $\dot\varphi_i = \Omega_i(I_1, \ldots, I_n)$ and the symplectic form is $\omega = \sum_{i=1}^n dI_i \wedge d\varphi_i$.*

The coordinates φ_i and I_i are called the *angle* and *action coordinates*, respectively. For the case of a noncompact M, one has a natural $\mathbb{R}^n$-action on the levels M_c, coming from the commuting Hamiltonian vector fields corresponding to the Hamiltonian functions $f_i, i = 1, \ldots, n$.

Note that the symplectic form ω vanishes identically on any level set M_c, so that each regular level set is a Lagrangian submanifold of the symplectic manifold M. (By definition, a *Lagrangian submanifold* $L \subset M$ of a symplectic manifold M is an isotropic submanifold of maximal dimension: for a $2n$-dimensional M, a Lagrangian submanifold L is n-dimensional and satisfies $\omega|_L \equiv 0$.)

Remark 4.30 While in finite dimensions there are many definitions of complete integrability of a Hamiltonian system and they are all more or less equivalent, this question is more subtle in infinite dimensions. One can start defining such systems based on the existence of action-angle coordinates, or on bi-Hamiltonian structures, or on the existence of an infinite number of "sufficiently independent" first integrals, or, even by requiring an explicit solvability. These definitions lead, generally speaking, to inequivalent notions, and precise relations between these definitions in infinite dimensions are yet to be better understood.

There are, however, examples of infinite-dimensional systems in which most, if not all, of these definitions work. This is the case, for example, for the celebrated Korteweg–de Vries equation. Other systems for which several approaches are also known are the Kadomtsev–Petviashvili equation, the Camassa–Holm equation, and many others, some of which we will encounter later in the book.

5 Symplectic Reduction

The Noether theorem in classical mechanics states that a Lagrangian system with extra symmetries has an invariant of motion. Hence in describing such a system one can reduce the dimensionality of the problem by "sacrificing this invariance." The notion of symplectic reduction can be thought of as a Hamiltonian analogue of the latter: If a symplectic manifold admits an appropriate group action, then this action can be "factored out." The quotient is a new symplectic manifold of lower dimension.

This construction can be used in both ways. On the one hand, one can reduce the dimensionality of certain systems that admit extra symmetries. On the other hand, certain complicated physical systems can be better understood by realizing them as the result of symplectic reduction from much simpler systems in higher dimensions.

5.1 Hamiltonian Group Actions

Consider a finite-dimensional symplectic manifold (M, ω), i.e., a manifold M equipped with a nondegenerate closed 2-form ω. Let G be a connected Lie group with Lie algebra $\mathfrak{g}$ and suppose that the exponential map exists. If the group G acts smoothly on M, each element X of the Lie algebra $\mathfrak{g}$ defines a vector field ξ_X on the manifold M as an infinitesimal action of the group:

$$\xi_X(m) := \frac{d}{dt}|_{t=0} \exp(tX)m \,.$$

The *action* of the group G on the manifold M is called *symplectic* if it leaves the symplectic form ω invariant, i.e., if $g^*\omega = \omega$ for all $g \in G$.

Exercise 5.1 Show that for the symplectic group action, the vector field ξ_X for any $X \in \mathfrak{g}$ is symplectic, i.e., the 1-form $\iota_{\xi_X}\omega$ is closed. (Hint: use the *Cartan homotopy formula* on differential forms, $L_\xi = \iota_\xi\, d + d\,\iota_\xi$, where L_ξ means the Lie derivative along ξ and the operators ι_ξ and d stand for the inner and outer derivatives of forms.)

The closedness of the 1-form means that it is locally exact, and hence the field ξ_X is locally Hamiltonian: in a neighborhood of each point of the manifold M, there exists a function H_X such that $\iota_{\xi_X}\omega = dH_X$. In general, this field is not necessarily defined by a univalued Hamiltonian function on the whole of M. Even if we suppose that such a Hamiltonian function exists, it is defined only up to an additive constant.

Definition 5.2 The *action* of a Lie group G on M is called *Hamiltonian* if for every $X \in \mathfrak{g}$ there exists a globally defined Hamiltonian function H_X that can be chosen in such a way that the map $\mathfrak{g} \to \mathbb{C}^\infty(M)$, associating to X the corresponding Hamiltonian H_X, is a Lie algebra homomorphism of the Lie algebra $\mathfrak{g}$ to the Poisson algebra of functions on M:

$$H_{[X,Y]} = \{H_X, H_Y\} \,.$$

Exercise 5.3 Prove that for a Hamiltonian G-action on M the Lie algebra isomorphism is *equivariant*, i.e.,

$$H_{\mathrm{Ad}_g X}(m) = H_X(g(m))$$

for all $g \in G$, $X \in \mathfrak{g}$, and $m \in M$.

Definition 5.4 Assume that the action of a group G on M is Hamiltonian. Then the *moment map* is the map $\Phi : M \to \mathfrak{g}^*$ defined by

$$H_X(m) = \langle \Phi(m), X \rangle ,$$

where $\langle\ ,\ \rangle$ denotes the pairing between $\mathfrak{g}$ and $\mathfrak{g}^*$.

In other words, given a vector X from the Lie algebra $\mathfrak{g}$, the moment map sends points of the manifold to the values of the Hamiltonian function H_X at those points.

Summarizing the above definitions, a symplectic G-action on a symplectic manifold M is called *Hamiltonian* if there exists a G-equivariant smooth map $\Phi : M \to \mathfrak{g}^*$ (the *moment map*) such that for all $X \in \mathfrak{g}$, we have $d\langle \Phi, X\rangle = \iota_{\xi_X}\omega$. Any vector field ξ_X on M that comes from an element $X \in \mathfrak{g}$ for such a group action has the Hamiltonian function $H_X = \langle \Phi, X\rangle$.

Exercise 5.5 Consider $M = \mathbb{R}^2$ with the standard symplectic form $\omega = dp \wedge dq$ and the group $U(1)$ acting on $\mathbb{R}^2$ by rotations. Show that this action is Hamiltonian with the moment map $\Phi(p, q) = \frac{1}{2}(p^2 + q^2)$.

Exercise 5.6 Consider the coadjoint action of a Lie group G on the dual of its Lie algebra. Show that this action restricted to any coadjoint orbit $\mathcal{O} \subset \mathfrak{g}^*$ is Hamiltonian with the moment map being the inclusion $\iota : \mathcal{O} \hookrightarrow \mathfrak{g}^*$.

Exercise 5.7 Generalize the definition of Hamiltonian group actions to Poisson manifolds and show that the coadjoint action of a Lie group G on the dual $\mathfrak{g}^*$ of its Lie algebra is Hamiltonian with the moment map given by the identity map $\mathrm{id} : \mathfrak{g}^* \to \mathfrak{g}^*$.

5.2 Symplectic Quotients

Let (M, ω) be a symplectic manifold with a Hamiltonian action of the group G. The equivariance of the moment map $\Phi : M \to \mathfrak{g}^*$ implies that the inverse image $\Phi^{-1}(\lambda)$ of a point $\lambda \in \mathfrak{g}^*$ is a union of G_λ-orbits, where $G_\lambda := \{g \in G \mid \mathrm{Ad}_g^*(\lambda) = \lambda\}$ is the stabilizer of λ. The symplectic reduction theorem below states that if λ is a regular value of the moment map, and if the set $\Phi^{-1}(\lambda)/G_\lambda$ of G_λ-orbits in $\Phi^{-1}(\lambda)$ is a manifold, then it acquires a natural symplectic structure from the one on M.

Exercise 5.8 Suppose that G is a finite-dimensional Lie group acting on the symplectic manifold M in a Hamiltonian way with a moment map Φ. For $m \in M$, let $G.m$ denote the G-orbit through m and let G_m denote the stabilizer of m with the Lie algebra $\mathfrak{g}_m$. Show that the kernel of the differential $d\Phi_m$ of the moment map Φ at any point $m \in M$ is given by

$$\ker(d\Phi_m) = (T_m(G.m))^\omega := \{\xi \in T_m(M) \mid \omega(\xi, \chi) = 0 \text{ for all } \chi \in T_m(G.m)\} .$$

(Here $(T_m(G.m))^\omega$ is the symplectic orthogonal complement to the tangent space $T_m(G.m)$ in $T_m(M)$.)

Show that the image of the differential $d\Phi_m$ is

$$\operatorname{im}(d\Phi_m) = \operatorname{ann}(\mathfrak{g}_m) := \{\lambda \in \mathfrak{g}^* \mid \lambda(X) = 0 \text{ for all } X \in \mathfrak{g}_m\} .$$

Conclude that an element $\lambda \in \mathfrak{g}^*$ is a regular value of the moment map, i.e., $d\Phi_m$ is surjective for all $m \in \Phi^{-1}(\lambda)$, if and only if for all $m \in \Phi^{-1}(\lambda)$ the stabilizer G_m is discrete.

The restriction of the symplectic form ω to the level set $\Phi^{-1}(\lambda)$ of the moment map is not necessarily symplectic, since it might acquire a kernel.

Exercise 5.9 Show that the foliation of $\Phi^{-1}(\lambda)$ by the kernels of ω is the foliation into (connected components of) G_λ-orbits. (Hint: for a regular value λ the preimage $\Phi^{-1}(\lambda)$ is a smooth submanifold of M, and the exercise above gives

$$\begin{aligned}\ker \omega|_{\Phi^{-1}(\lambda)} &= T_m\Phi^{-1}(\lambda) \cap (T_m\Phi^{-1}(\lambda))^\omega \\ &= T_m\Phi^{-1}(\lambda) \cap (\ker d\Phi_m)^\omega \\ &= T_m\Phi^{-1}(\lambda) \cap T_m(G.m) = T_m(G_\lambda.m)\,.)\end{aligned}$$

Hence, if the quotient space of the level $\Phi^{-1}(\lambda)$ over the G_λ-action is reasonably nice, the 2-form ω descends to a symplectic form on this quotient; see Figure 5.1. This is made precise in the following reduction theorem.

Theorem 5.10 (Marsden–Weinstein [254], Meyer [261]) *Suppose that λ is a regular value of the moment map and suppose that $\Phi^{-1}(\lambda)/G_\lambda$ is a manifold (this condition is satisfied if, for example, G_λ is compact and acts freely on $\Phi^{-1}(\lambda)$). Then there exists a unique symplectic structure ω_λ on the reduced space $\Phi^{-1}(\lambda)/G_\lambda$ such that*

$$\iota^*\omega = \pi^*\omega_\lambda.$$

(Here, ι denotes the embedding $\Phi^{-1}(\lambda) \hookrightarrow M$ and π stands for the projection $\Phi^{-1}(\lambda) \to \Phi^{-1}(\lambda)/G_\lambda$.)

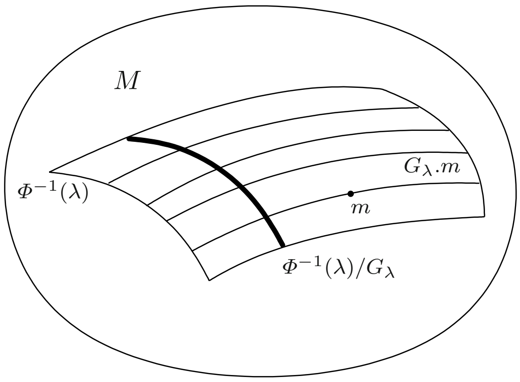


Fig. 5.1. The orbits of the stabilizer group G_λ in the preimage $\Phi^{-1}(\lambda)$.

The resulting manifold $\Phi^{-1}(\lambda)/G_\lambda$ of the above *symplectic reduction* (also known as *Hamiltonian* or *Marsden–Weinstein reduction*) is called the *symplectic quotient*.

Finally, if $H : M \to \mathbb{R}$ is a Hamiltonian function invariant under the G-action, it descends to a function H_λ on the quotient space $\Phi^{-1}(\lambda)/G_\lambda$. Furthermore, if two G-invariant functions F and H on M Poisson commute with respect to the Poisson structure on M defined by the symplectic form ω, the corresponding functions on the quotient $\Phi^{-1}(\lambda)/G_\lambda$ still Poisson commute with respect to the quotient Poisson structure.

Example 5.11 Consider the manifold $M = \mathbb{C}^{n+1}$ with its standard symplectic structure $\omega = \frac{i}{2}\sum dz_i \wedge d\bar{z}_i$. The group $U(1) = \mathbb{R}/\mathbb{Z}$ acts on $\mathbb{C}^{n+1}$ by rotation: $z \mapsto e^{2\pi it}z$. The moment map for this action is given by $\Phi(z) = \pi\|z\|^2$. The reduced space $\Phi^{-1}(1)/U(1)$ is the complex projective space $\mathbb{CP}^n$ with the symplectic form being (a multiple of) the Fubini–Study form.

Let $f_i : \mathbb{C}^{n+1} \to \mathbb{R}$ denote the function $f_i(z) = \|z_i\|^2$. The functions f_i are invariant under the $U(1)$-action on $\mathbb{C}^{n+1}$, and the corresponding Hamiltonian functions on the symplectic quotient generate the rotations in the $\mathbb{C}$-hyperplanes $\{z_i = \text{const}\}$ in $\mathbb{CP}^n$.

6 Bibliographical Notes

There are many books on Lie groups and Lie algebras covering the material of this chapter ([345, 383, 93, 60] to name a few). The treatment of various aspects of infinite-dimensional Lie groups can be found in [12, 157, 265], as well as in the thorough monographs [301, 322]. The calculus on infinite-dimensional manifolds is developed in [157, 222]. For more details on central extensions (Section 3) we recommend [322, 276], while for the Euler equations on groups (Section 4) see [18, 24, 252].

The Lie–Poisson structure on the dual space of a Lie algebra was already known to Sophus Lie. The symplectic structure on the coadjoint orbits of a Lie group goes back to Kirillov, Kostant, and Souriau. The local description of a Poisson manifold, including its symplectic realizations and the transversal Poisson structure, was given in [384]. A generalization of Poisson manifolds are Jacobi manifolds, functions on which form a Lie algebra (but not necessarily a Poisson algebra). Jacobi manifolds are also called local Lie algebras, and their local structure was described in [199].

There is vast literature on integrable systems; see e.g., [18, 31, 90, 165, 280, 299, 330], and in particular, on integrability of the Euler equations on dual Lie algebras [47, 48, 114, 300, 311]. The Euler equations for an n-dimensional rigid body were considered back in the nineteenth century by Cayley, Frahm, and others; cf., e.g., [128, 266]. Schottky proved the integrability of these equations in the four-dimensional case [339], by giving an explicit theta-function solution for SO(4). The idea of freezing the Lie–Poisson bracket, which resolved the integrability issue in the general n-dimensional case, was proposed by Manakov in [249].

Among other mechanisms of integrability, not discussed in this book, we would like to mention the Adler–Kostant–Symes theorem [3, 213, 362, 328], the R-matrix method, see, e.g., [348, 330]; integrability of geodesic flows on quadrics and related systems, see [281, 23]; as well as discrete analogues of the Euler equation on Lie groups [282, 373, 45].

For the description of the symplectic reduction we followed [154] and [259]. The construction of symplectic reduction does not generalize directly to the case of infinite-dimensional manifolds and infinite-dimensional Lie groups. In spite of the lack of a sufficiently broad general theory, there are many concrete constructions in various infinite-dimensional situations. In particular, Hamiltonian actions of loop groups and, more generally, Hamiltonian actions of gauge transformation groups on Riemann surfaces, which we are going to deal with further in this book, have been considered, for example, in the papers [28, 260].

2 Diffeomorphisms of the Circle and the Virasoro–Bott Group

This section deals with the Lie group $\mathrm{Diff}(S^1)$ of orientation-preserving diffeomorphisms of the circle and its Lie algebra $\mathrm{Vect}(S^1)$ of smooth vector fields on the circle, as well as with their central extensions. We start by showing that the Lie algebra of vector fields on the circle admits a unique nontrivial central extension, the so-called Virasoro algebra. This central extension gives rise to a central extension of the Lie group of circle diffeomorphisms, which is called the Virasoro–Bott group. Similar to the case of loop groups, the Virasoro–Bott group has a "nicer" coadjoint representation than the nonextended group of circle diffeomorphisms. Its coadjoint orbits can be classified in terms of conjugacy classes of the finite-dimensional group $\mathrm{SL}(2,\mathbb{R})$. Finally, we describe the Euler equations corresponding to right-invariant metrics on the Virasoro–Bott group and encounter the KdV and related partial differential equations among them.

2.1 Central Extensions

Let us consider the Lie algebra $\mathrm{Vect}(S^1)$ of smooth vector fields on the circle. After fixing a coordinate θ on the circle, any vector field can be written as $f(\theta)\partial_\theta$, where f is a smooth function on S^1 and ∂_θ stands for $\frac{\partial}{\partial\theta}$. Under this identification, the commutator of two elements in $\mathrm{Vect}(S^1)$ is given by

$$[f(\theta)\partial_\theta, g(\theta)\partial_\theta] = (f'(\theta)g(\theta) - g'(\theta)f(\theta))\partial_\theta\,,$$

where f' denotes the derivative in θ of the function f.[11]

Definition / Proposition 2.1 *The map* $\omega : \mathrm{Vect}(S^1) \times \mathrm{Vect}(S^1) \to \mathbb{R}$ *given by*

$$\omega(f(\theta)\partial_\theta, g(\theta)\partial_\theta) = \int_{S^1} f'(\theta)g''(\theta)d\theta \tag{2.5}$$

is a nontrivial 2*-cocycle on* $\mathrm{Vect}(S^1)$*, called the* Gelfand–Fuchs cocycle*. The corresponding central extension of* $\mathrm{Vect}(S^1)$ *is called the* Virasoro algebra *and is denoted by* $\mathfrak{vir}$*.*

Exercise 2.2 Prove the cocycle identity for ω.

The following proposition shows that the Virasoro algebra is the unique (up to isomorphism) nontrivial central extension of the Lie algebra $\mathrm{Vect}(S^1)$.

Proposition 2.3 *The second continuous cohomology group* $H^2(\mathrm{Vect}(S^1), \mathbb{R})$ *is one-dimensional and is generated by the Gelfand–Fuchs cocycle* ω*.*

[11] Note that this Lie bracket is the negative of the commonly assumed commutator of vector fields, as the calculations in Exercise 2.3 of Chapter I shows; see [24].

PROOF. The proof of this proposition (see [322]) is similar to that of Proposition 1.6. Our goal is to show that, up to a coboundary, any continuous 2-cocycle ω on the Lie algebra $\mathrm{Vect}(S^1)$ is a multiple of the Gelfand–Fuchs cocycle. First, let us extend the cocycle ω from $\mathrm{Vect}(S^1)$ to a complex bilinear form on the complexification $\mathrm{Vect}(S^1)_{\mathbb{C}} = \mathrm{Vect}(S^1) \otimes \mathbb{C}$ of the Lie algebra $\mathrm{Vect}(S^1)$. An element $f(\theta)\partial_\theta \in \mathrm{Vect}(S^1)_{\mathbb{C}}$ can be expanded into a Fourier series

$$f(\theta) = \sum f_n e^{in\theta}\,.$$

By continuity, the cocycle ω is completely determined by its values on the basis fields $L_n = ie^{in\theta}\partial_\theta$. Note that the commutator of the fields L_n and L_m is given by

$$[L_n, L_m] = (m-n)L_{n+m}\,.$$

The cocycle identity for ω and the triple L_0, L_m, L_n gives

$$\omega([L_0, L_m], L_n) + \omega(L_m, [L_0, L_n]) = \omega(L_0, [L_m, L_n])\,,$$

which implies that the cohomology class of the cocycle ω is unchanged under rotations of S^1 that are generated by the vector field L_0. Indeed, the right-hand side of the equation above is an *exact cocycle* (i.e., coboundary) $d\alpha$, where α is the linear functional on $\mathrm{Vect}(S^1)$ defined by $\alpha(L_m) := \omega(L_0, L_m)$. (Here by definition $d\alpha(L_n, L_m) := \alpha([L_n, L_m])$.) In particular, the cocycle obtained from ω by averaging over S^1 belongs to the same cohomology class as ω. Therefore, we can assume ω to be *rotation invariant*, i.e.,

$$\omega([L_0, L_m], L_n) + \omega(L_m, [L_0, L_n]) = 0\,. \tag{2.6}$$

Set $\omega_{n,m} := \omega(L_n, L_m)$. Then the commutator relation of the fields L_n and L_m and equation (2.6) imply

$$m\omega_{m,n} + n\omega_{m,n} = 0\,.$$

This implies that $\omega_{m,n} = 0$ for $m + n \neq 0$. Antisymmetry of the cocycle ω implies $\omega_{n,-n} = \omega_{-n,n}$, so that it is enough to determine $\omega_{n,-n}$ for $n \in \mathbb{N}$.

The cocycle identity for ω evaluated on the triple L_m, L_n, L_{-m-n} implies

$$(m-n)\omega_{m+n,-n-m} + (2m+n)\omega_{n,-n} - (2n+m)\omega_{m,-m} = 0\,.$$

In particular, for $m = 1$ the equation above reads as follows:

$$(-n+1)\omega_{n+1,-n-1} + (n+2)\omega_{n,-n} - (2n+1)\omega_{1,-1} = 0\,.$$

Hence $\omega_{n,-n}$ is defined recursively once $\omega_{1,-1}$ and $\omega_{2,-2}$ are fixed. This shows that the space of the bilinear forms ω that satisfy the 2-cocycle condition is at most two-dimensional. Two linear independent elements of this space are given by $\omega_{n,-n} = n^3$ and $\omega_{n,-n} = n$. But the 2-cocycle defined by $\omega_{n,-n} = n$

is exact, since it coincides with $d\widetilde{\alpha}$, where $\widetilde{\alpha}$ is the linear functional defined by $\widetilde{\alpha}(L_n) = -\frac{1}{2}\delta_{n,0}$.

So up to a 2-coboundary, any 2-cocycle ω has the "cubic" form

$$\omega(L_n, L_m) = c\delta_{n,-m}n^3$$

for some $c \in \mathbb{C}$.

It remains to show that the "cubic" cocycle ω is nontrivial. Suppose that $\omega = d\beta$ for some 1-cocycle β. This means that β is a linear map and $\omega(L_n, L_m) = \beta([L_n, L_m])$. In particular, we have $\beta([L_n, L_{-n}]) = 2ni\beta(L_0)$, which shows that in this case $\omega(L_n, L_{-n})$ would have to depend linearly on n. This contradiction completes the proof. $\square$

Our next goal is to show that the central extension of the Lie algebra of vector fields $\mathrm{Vect}(S^1)$ defined by the Gelfand–Fuchs cocycle ω can be lifted to a central extension of the group of circle diffeomorphisms $\mathrm{Diff}(S^1)$. It turns out that the situation here is much simpler than that in the case of the loop groups. The central extension of the group $\mathrm{Diff}(S^1)$ corresponding to the Lie algebra $\mathfrak{vir}$ is topologically trivial and hence can be defined by a continuous group 2-cocycle.

Let $\varphi : \theta \mapsto \varphi(\theta)$ be a diffeomorphism of the circle, and φ' stands for its derivative in θ.

Definition / Proposition 2.4 *The map $B : \mathrm{Diff}(S^1) \times \mathrm{Diff}(S^1) \to S^1$ given by*

$$(\varphi, \psi) \mapsto \frac{1}{2}\int_{S^1} \log(\varphi \circ \psi)' d \log \psi'$$

is a continuous 2-cocycle on the group $\mathrm{Diff}(S^1)$. The Lie algebra of the corresponding central extension $\widehat{\mathrm{Diff}}(S^1)$ is the Virasoro algebra $\mathfrak{vir}$. The 2-cocycle B is called the Bott cocycle, *and the corresponding central extension of the group $\mathrm{Diff}(S^1)$ is called the* Virasoro–Bott group.

Proof. To show that the map B defines a group 2-cocycle, we have to check the identity

$$B(\varphi \circ \psi, \eta) + B(\varphi, \psi) = B(\varphi, \psi \circ \eta) + B(\psi, \eta)\,.$$

It is provided by the chain rule, which immediately gives

$$B(\varphi\circ\psi, \eta) = \frac{1}{2}\int_{S^1} \log(\varphi\circ\psi\circ\eta)'\, d\log\eta' = \frac{1}{2}\int_{S^1} \log(\varphi'\circ\psi\circ\eta)\, d\log\eta' + B(\psi, \eta)$$

and

$$B(\varphi, \psi \circ \eta) = \frac{1}{2} \int_{S^1} \log(\varphi \circ \psi \circ \eta)' \, d \log(\psi \circ \eta)'$$
$$= B(\varphi, \psi) + \frac{1}{2} \int_{S^1} \log(\varphi' \circ \psi \circ \eta) \, d \log \eta' \, .$$

Now we verify that the infinitesimal version of the Bott group cocycle B coincides with the Gelfand–Fuchs Lie algebra 2-cocycle ω. Let $f\partial_\theta$ and $g\partial_\theta$ be two smooth vector fields on S^1 and consider the corresponding flows φ_s and ψ_t on S^1, starting at the identity diffeomorphism: $\varphi_0 = \psi_0 = \mathrm{id}$.

We have to check that

$$\omega(f\partial_\theta, g\partial_\theta) = \frac{d^2}{dt\, ds}\Big|_{t=0,s=0} B(\varphi_t, \psi_s) - \frac{d^2}{dt\, ds}\Big|_{t=0,s=0} B(\psi_s, \varphi_t)$$

(see Proposition 3.14 of Chapter I). The latter holds, since

$$\frac{d}{dt}\Big|_{t=0} B(\varphi_t, \psi_s) = \frac{1}{2} \int_{S^1} \big(\log'(\varphi_0 \circ \psi_s)'\big) \, (f \circ \psi_s)' d \log \psi_s'$$
$$= \frac{1}{2} \int_{S^1} (f' \circ \psi_s) \, d \log \psi_s'$$

and

$$\frac{d}{ds}\Big|_{s=0} \frac{1}{2} \int_{S^1} (f' \circ \psi_s) \, d \log \psi_s' = \frac{1}{2} \int_{S^1} f' dg' \, .$$

Similarly, we obtain

$$\frac{d^2}{dt\, ds}\Big|_{t=0,s=0} B(\psi_s, \varphi_t) = \frac{1}{2} \int_{S^1} g' df' = -\frac{1}{2} \int_{S^1} f' dg' \, ,$$

which, combined with the equation above, yields the assertion. □

2.2 Coadjoint Orbits of the Group of Circle Diffeomorphisms

Before we start classifying the coadjoint orbits of the Virasoro group, let us take a look at the coadjoint representation of the nonextended group of orientation-preserving diffeomorphisms of the circle. Observe that the dual spaces to the infinite-dimensional Lie algebras considered below are always understood as smooth duals, i.e., they are identified with appropriate spaces of smooth functions.

Let $\mathrm{Diff}(S^1)$ be the group of all orientation-preserving diffeomorphisms of S^1 and let $\mathrm{Vect}(S^1)$ be its Lie algebra.

Proposition 2.5 ([202]) *The (smooth) dual space* $\mathrm{Vect}(S^1)^*$ *is naturally identified with the space of quadratic differentials* $\Omega^{\otimes 2}(S^1) = \{u(\theta)(d\theta)^2\}$ *on the circle. The pairing is given by the formula*

$$\langle u(\theta)(d\theta)^2,\ v(\theta)\partial_\theta\rangle = \int_{S^1} u(\theta)v(\theta)\ d\theta$$

for any vector field $v(\theta)\partial_\theta \in \mathrm{Vect}(S^1)$. *The coadjoint action coincides with the action of a diffeomorphism on the quadratic differential: for a diffeomorphism* $\varphi \in \mathrm{Diff}(S^1)$ *the action is*

$$\mathrm{Ad}^*_{\varphi^{-1}} : \quad u\,(d\theta)^2 \mapsto u(\varphi)\cdot(\varphi')^2\,(d\theta)^2 = u(\varphi)\cdot(d\varphi)^2.$$

It follows from this proposition that the square root $\sqrt{u(\theta)(d\theta)^2}$ (when it makes sense) transforms under a diffeomorphism as a differential 1-form. In particular, if the function $u(\theta)$ does not have any zeros on the circle (say, $u(\theta) > 0$), then $\Phi(u(\theta)(d\theta)^2) := \int_{S^1}\sqrt{u(\theta)}\,d\theta$ is a Casimir function, i.e., an invariant of the coadjoint action. One can see that there is only one Casimir function in this case, since the corresponding orbit has codimension 1 in the dual space $\mathrm{Vect}(S^1)^*$. Indeed, there exists a diffeomorphism that sends the quadratic differential $u(\theta)(d\theta)^2$ without zeros to the constant quadratic differential $u_0(d\theta)^2$, where the constant u_0 is such that $\sqrt{u_0}$ is the average value of the 1-form $\sqrt{u(\theta)}\,d\theta$ on the circle:

$$2\pi\sqrt{u_0} = \int_{S^1}\sqrt{u(\theta)}\,d\theta\,.$$

The value u_0 parametrizes the orbits close to $u(\theta)(d\theta)^2$, and hence all these orbits have codimension 1 in $\Omega^{\otimes 2}(S^1)$. The stabilizer of a constant quadratic differential is the group S^1 of rigid rotations, so that the orbit through $u(\theta)(d\theta)^2$ is diffeomorphic to $\mathrm{Diff}(S^1)/S^1$.

On the other hand, if a differential $u(\theta)(d\theta)^2$ changes sign on the circle, then the integrals

$$\int_a^b \sqrt{|u(\theta)|}\,d\theta\,,$$

evaluated between any two consecutive zeros a and b of the function $u(\theta)$, are invariant. In particular, since $u(\theta)$ has at least two zeros, the coadjoint orbit of such a differential $u(\theta)(d\theta)^2$ necessarily has codimension higher than 1, and there exist coadjoint orbits of the group $\mathrm{Diff}(S^1)$ of arbitrarily high codimension. The classification of orbits in $\mathrm{Vect}(S^1)^*$ was described in [201, 203].

Remark 2.6 One can show that if the function $u(\theta)$ has two simple zeros, changing sign exactly twice on the circle, then the corresponding coadjoint orbit of the group $\mathrm{Diff}(S^1)$ has codimension 2; see [203]. (The corresponding two Casimirs are the integrals of $\sqrt{|u(\theta)|}\,d\theta$ over two different parts of the circle between these two zeros, while there are no extra local invariants at zeros themselves: quadratic differentials with simple zeros are all locally diffeomorphic to $\pm\theta\,(d\theta)^2$.)

In other words, in a family of quadratic differentials $\bar{u}^\epsilon := u^\epsilon(\theta)\,(d\theta)^2$, where the function u^ϵ is everywhere positive for $\epsilon > 0$, has a double zero for $\epsilon = 0$, and has two simple zeros for $\epsilon < 0$ (e.g., $u^\epsilon = \cos\theta + 1 + \epsilon$) the codimension of the coadjoint orbit of $\bar{u}^\epsilon = u^\epsilon(d\theta)^2$ changes from 1 for $\epsilon > 0$ to 2 for $\epsilon \leq 0$, since the number of Casimirs jumps from 1 to 2. (Note that for $\epsilon = 0$ the orbit codimension of $\bar{u}^0$ is also 2, since the existence of a double zero imposes an extra constraint on a quadratic differential.)

This change of "codimension parity" of the (infinite-dimensional) coadjoint orbits is rather surprising, since in finite dimensions the existence of a symplectic structure on each coadjoint orbit forces all of them to be even-dimensional, and hence codimensions of coadjoint orbits for a given (finite-dimensional) group are always of the same parity: either all even or all odd. However, for $\mathrm{Vect}(S^1)^* = \Omega^{\otimes 2}(S^1)$ we observe that there exist orbits of both codimensions 1 and 2!

In particular, this shows that the Weinstein theorem [384] on the existence of the transverse Poisson structure to symplectic leaves does not hold for infinite-dimensional Poisson manifolds; cf. Remark I.4.7. Indeed, on a two-dimensional transversal to $\bar{u}^0$ in $\mathrm{Vect}(S^1)^*$, neighboring coadjoint orbits of $\bar{u}^\epsilon$ have traces of both codimensions 1 and 2. One can consider the following example, clarifying how the change of parity can occur for infinite-dimensional symplectic leaves. Define the Poisson structure in an infinite-dimensional vector space $\{(x_0, x_1, x_2, \dots)\}$ by the bivector field

$$\Pi = x_0\,\partial_{x_1} \wedge \partial_{x_2} + \partial_{x_2} \wedge \partial_{x_3} + \partial_{x_3} \wedge \partial_{x_4} + \dots .$$

Its symplectic leaves are hyperplanes $\{x_0 = \mathrm{const} \neq 0\}$, while for $x_0 = 0$ the symplectic leaves are planes $\{x_0 = 0,\, x_1 = \mathrm{const}\}$ of codimension 2.

In the next section, however, we shall see that coadjoint orbits of the Virasoro group, the central extension of the diffeomorphism group $\mathrm{Diff}(S^1)$, do respect the codimension parity and behave much more like finite-dimensional coadjoint orbits.

2.3 The Virasoro Coadjoint Action and Hill's Operators

Let $\mathfrak{vir}$ be the Virasoro algebra, whose elements are pairs $(f(\theta)\partial_\theta, c)$, where $f(\theta)\partial_\theta$ is a vector field and c is a real number. We can think of its (smooth) dual space as the space of pairs $\mathfrak{vir}^* = \{(u(\theta)(d\theta)^2, a)\}$ consisting of a quadratic differential and a real number (the cocentral term). The pairing between $\mathfrak{vir}$ and $\mathfrak{vir}^*$ is given by

$$\langle (f(\theta)\partial_\theta, c), (u(\theta)(d\theta)^2, a)\rangle = \int_{S^1} f(\theta)u(\theta)d\theta + c \cdot a\,.$$

Our goal in this section is to derive a classification of the coadjoint orbits of the Virasoro–Bott group $\widehat{\mathrm{Diff}}(S^1)$. This classification turns out to be similar to that of the coadjoint orbits of the centrally extended loop groups in Section 1.2.

We begin by noticing that the center of the group $\widehat{\mathrm{Diff}}(S^1)$ acts trivially on the dual space $\mathfrak{vir}^*$. This is why to describe the coadjoint representation of the Virasoro–Bott group we need the action of (nonextended) diffeomorphisms only.

Definition / Proposition 2.7 *The coadjoint action of a diffeomorphism $\varphi \in \mathrm{Diff}(S^1)$ on the dual $\mathfrak{vir}^*$ of the Virasoro algebra is given by the following formula:*

$$\mathrm{Ad}^*_{\varphi^{-1}} : \quad (u\,(d\theta)^2,\, a) \mapsto \big(u(\varphi)\cdot(\varphi')^2\,(d\theta)^2 + aS(\varphi)\,(d\theta)^2,\, a\big)\,, \tag{2.7}$$

where

$$S(\varphi) = \frac{\varphi'\varphi''' - \frac{3}{2}(\varphi'')^2}{(\varphi')^2}$$

is the Schwarzian derivative *of the diffeomorphism φ.*

The same formula can be used to define the Schwarzian derivative $S(\phi)$ for a smooth map $\phi : \mathbb{R} \to \mathbb{R}$ or $\mathbb{R} \to \mathbb{R}P^1 \simeq S^1$.

Proof. The coadjoint action of the Virasoro algebra is defined by the identity

$$\langle \mathrm{ad}^*_{(v\partial_\theta, b)}(u(d\theta)^2,\, a), (w\partial_\theta, c)\rangle = -\langle (u(d\theta)^2,\, a), [(v\partial_\theta, b), (w\partial_\theta, c)]\rangle.$$

Using the definition of the Virasoro commutator and integrating by parts we obtain that the right-hand side is equal to

$$\int_{S^1} -w(2uv' + u'v + av''')\; d\theta\,.$$

Thus the coadjoint operator is

$$\mathrm{ad}^*_{(v\partial_\theta, b)}(u(d\theta)^2,\, a) = -((2uv' + u'v + av''')(d\theta)^2,\; 0). \tag{2.8}$$

It remains to check that equation (2.7) indeed defines a representation of the group $\mathrm{Diff}(S^1)$ on the space $\mathfrak{vir}^*$ and that the infinitesimal version of this action is given by equation (2.8). Both assertions can be checked by direct calculations, which we leave to the reader. □

Exercise 2.8 Prove the following transformation law for the Schwarzian derivative:

$$S(\varphi \circ \psi) = (S(\varphi) \circ \psi) \cdot (\psi')^2 + S(\psi)\,. \tag{2.9}$$

Check that the formula (2.7) defines a group representation by using this law.

It turns out to be more convenient to regard the dual Virasoro space $\mathfrak{vir}^*$ not as the space of pairs $\{(u(\theta)(d\theta)^2, a)\}$, but as the space of *Hill's operators*, i.e., differential operators $a\partial_\theta^2+u(\theta)$, where ∂_θ^2 stands for the second derivative $d^2/d\theta^2$. Indeed, the group action on Hill's operators

$$\mathrm{Ad}^*_{\varphi^{-1}}: \quad a\partial_\theta^2 + u(\theta) \mapsto a\partial_\theta^2 + u(\varphi)\cdot(\varphi')^2 + aS(\varphi) \tag{2.10}$$

has the following nice geometric interpretation (see, e.g., [342, 202, 205, 304]).

Look at a hyperplane $a=$ const corresponding to nonzero a in the dual space $\mathfrak{vir}^*$. For instance, we fix $a=1$ and consider Hill's operators of the form $\partial_\theta^2 + u(\theta)$, where θ is a coordinate on S^1. Let f and g be two independent solutions of the corresponding Hill differential equation

$$(\partial_\theta^2 + u(\theta))y = 0 \tag{2.11}$$

for an unknown function y. Although this equation has periodic coefficients, the solutions need not necessarily be periodic, but instead are defined over $\mathbb{R}$. Consider the ratio $\eta := f/g : \mathbb{R} \to \mathbb{R}P^1$. (Below we use the same notation θ for the coordinate on the circle $S^1 = \mathbb{R}/2\pi\mathbb{Z}$ and on its cover $\mathbb{R}$.)

Proposition 2.9 *The potential u is (one-half) the Schwarzian derivative of the ratio η:*

$$u = \frac{S(\eta)}{2}.$$

Proof. First we note that the *Wronskian*

$$W(f,g) := \det\begin{pmatrix} f & g \\ f' & g'\end{pmatrix} = fg' - f'g$$

is constant, since it satisfies $W' = 0$. For two independent solutions the Wronskian does not vanish, and we normalize W by setting $W=-1$.

This additional condition allows one to find the potential u from the ratio η. Indeed, first one reconstructs the solutions f and g from the ratio η by differentiating:

$$\eta' = \frac{f'g - fg'}{g^2} = \frac{-W}{g^2} = \frac{1}{g^2}\,.$$

Therefore,

$$g = \frac{1}{\sqrt{\eta'}}, \qquad f = g\cdot\eta = \frac{\eta}{\sqrt{\eta'}}.$$

Given two solutions f and g, one immediately finds the corresponding differential equation they satisfy by writing out the following 3×3 determinant:

$$\det\begin{bmatrix} y & f & g \\ y' & f' & g' \\ y'' & f'' & g''\end{bmatrix} = 0\,.$$

Since f and g satisfy the equation $y'' + u \cdot y = 0$, one obtains from the determinant above that

$$u = -\det \begin{bmatrix} f' & g' \\ f'' & g'' \end{bmatrix} .$$

The explicit formula for u expressed in terms of η turns out to be one-half the Schwarzian derivative of η. □

Corollary 2.10 *The Schwarzian derivative $S(\eta)$ is invariant with respect to a Möbius transformation $\eta \mapsto (a\eta+b)/(c\eta+d)$, where a, b, c, d are real numbers such that $ad - bc = 1$.*

In particular, if η itself is a Möbius transformation $\eta : \theta \mapsto (a\theta+b)/(c\theta+d)$, then $S(\eta) = S(\mathrm{id}) = 0$, where $\mathrm{id} : \theta \mapsto \theta$.

Proof. Indeed, for a given potential u the solutions f and g of the corresponding differential equation are not defined uniquely, but up to a transformation of the pair (f, g) by a matrix from $\mathrm{SL}(2, \mathbb{R})$. Then the ratio η changes by a Möbius transformation. Thus Möbius equivalent ratios η correspond to the same potential $u = S(\eta)/2$. For the identity diffeomorphism $\mathrm{id} : \theta \mapsto \theta$, the explicit formula for the Schwarzian derivative gives $S(\mathrm{id}) = 0$. □

Proposition 2.11 *The Virasoro coadjoint action of a diffeomorphism φ on the potential $u(\theta)$ gives rise to a diffeomorphism change of coordinate in the ratio η:*

$$\varphi : \quad \eta(\theta) \to \eta(\varphi(\theta)) .$$

Proof. We look at the corresponding infinitesimal action on the solutions of the differential equation $(\partial_\theta^2 + u(\theta))y = 0$. For a diffeomorphism $\varphi^{-1}(\theta) = \theta + \epsilon v(\theta)$ close to the identity, consider the infinitesimal Virasoro action of φ^{-1} on the potential $u(\theta)$:

$$u \mapsto u + \epsilon \cdot \delta u, \quad \text{where} \quad \delta u = 2uv' + u'v + \frac{1}{2}v'''$$

(cf. formula (2.8) for $a = \frac{1}{2}$ and note that we are considering the action of φ^{-1}). It is consistent with the following action on a solution y of the above differential equation:

$$y \mapsto y + \epsilon \cdot \delta y, \quad \text{where} \quad \delta y = -\frac{1}{2}yv' + y'v.$$

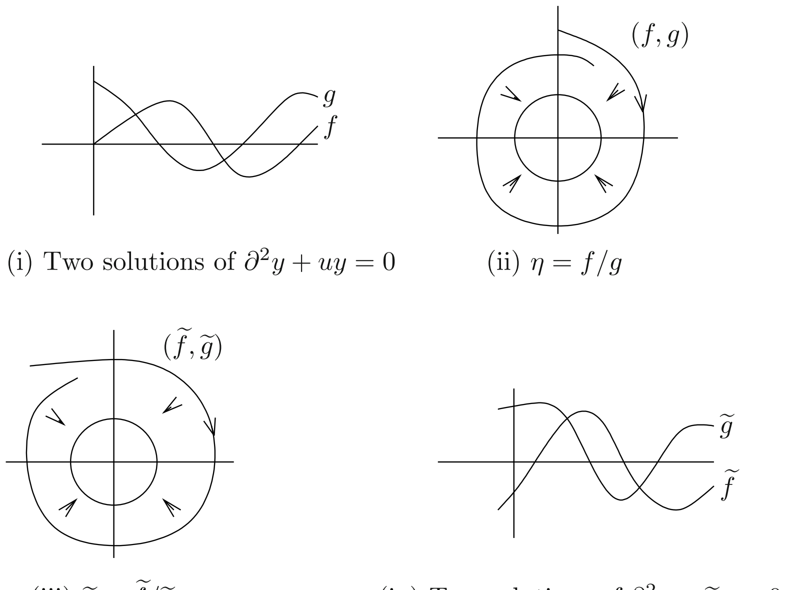
The consistency means that $(\partial_\theta^2 + u + \epsilon \cdot \delta u)(y + \epsilon \cdot \delta y) = 0 + \mathcal{O}(\epsilon^2)$.

Note that the action $\epsilon \cdot \delta y = \epsilon \cdot (-\frac{1}{2} y v' + y' v)$ is an infinitesimal version of the following action of the diffeomorphism $\varphi^{-1}(\theta) = \theta + \epsilon v(\theta)$ on y:

$$\varphi^{-1} : \quad y(\theta) \mapsto y(\varphi(\theta))(\varphi'(\theta))^{-1/2}.$$

Thus solutions to Hill's equation transform as densities of degree $-1/2$. Therefore the ratio η of two solutions transforms as a function under a diffeomorphism action. $\square$

In short, to calculate the coadjoint action on the potential u one can first pass from this potential to the ratio of two solutions of the corresponding Hill equation, then change the variable in the ratio, and finally take the Schwarzian derivative of the new ratio to reconstruct the new potential $\mathrm{Ad}^*_{\varphi^{-1}} u$ (see Figure 2.1).



(i) Two solutions of $\partial^2 y + uy = 0$ (ii) $\eta = f/g$

(iii) $\widetilde{\eta} = \widetilde{f}/\widetilde{g}$ (iv) Two solutions of $\partial^2 y + \widetilde{u} y = 0$

Fig. 2.1. Schematic picture of the action of a diffeomorphism of S^1 on Hill's operators and their solutions: To two solutions of the equation $(\partial^2 + u)y = 0$, one associates their ratio $\eta : \mathbb{R} \to \mathbb{R}P^1$. A diffeomorphism φ acts on the ratio η by reparametrization, and one reconstructs the corresponding solutions and Hill's operator $\partial^2 + \widetilde{u}$ from the new ratio $\widetilde{\eta}$.

Now we return to our goal, the classification of the Virasoro coadjoint orbits. Any (a = const)-hyperplane in the Virasoro dual $\mathfrak{vir}^*$ is invariant under the coadjoint action (see equation (2.10)), and identified with Hill's operators for $a \neq 0$. While all of the above considerations of Hill's operators were of local nature (local in θ), now we will make use of the fact that Hill's operators are periodic: $u(\theta)$ is defined on a circle.

Consider the universal covering $\widetilde{\mathrm{SL}}(2,\mathbb{R})$ of the group $\mathrm{SL}(2,\mathbb{R})$. The group $\widetilde{\mathrm{SL}}(2,\mathbb{R})$ admits an outer automorphism of order 2, taking the inverse of a matrix. One can see that the identity is its only fixed point on the universal covering group. Consider the set $(\widetilde{\mathrm{SL}}(2,\mathbb{R}) \setminus \{\mathrm{id}\})/\mathbb{Z}_2$, where we first dropped the identity element from $\widetilde{\mathrm{SL}}(2,\mathbb{R})$ before taking the quotient. The main result of this section is the following theorem.

Theorem 2.12 ([342, 202]) *Given $a \neq 0$, there is a one-to-one correspondence between the set of coadjoint orbits of the Virasoro–Bott group in the hyperplane $\{a\partial_\theta^2 + u(\theta)\} \subset \mathfrak{vir}^*$ and the set of conjugacy classes in the quotient $(\widetilde{\mathrm{SL}}(2,\mathbb{R}) \setminus \{\mathrm{id}\})/\mathbb{Z}_2$*

Proof. Consider a pair (f,g) of linearly independent solutions of the Hill equation $(a\partial_\theta^2 + u(\theta))y = 0$. For a periodic potential $u(\theta)$ these solutions are quasiperiodic, i.e., the values $(f(\theta), g(\theta))$ and $(f(\theta+2\pi), g(\theta+2\pi))$ are related by a monodromy matrix $M \in \mathrm{SL}(2,\mathbb{R})$:

$$(f(\theta + 2\pi), g(\theta + 2\pi)) = (f(\theta), g(\theta))\, M\,. \tag{2.12}$$

Recall that we use the same notation θ for the coordinate on the circle $S^1 = \mathbb{R}/2\pi\mathbb{Z}$ and on its cover $\mathbb{R}$. Note that the monodromy matrix M can be viewed as an element in the universal cover $\widetilde{\mathrm{SL}}(2,\mathbb{R})$, where the lift to the cover is provided by the fundamental solution $(f(\theta), g(\theta))$, starting at the identity: $\left(\begin{smallmatrix} f & g \\ f' & g' \end{smallmatrix}\right)|_{\theta=0} = \mathrm{id}$.

Similarly, the values of the "projective solution," the ratios $\eta(\theta) := f(\theta)/g(\theta)$ and $\eta(\theta + 2\pi) := f(\theta + 2\pi)/g(\theta + 2\pi)$, are related by a Möbius transformation $\mathcal{M} \in \mathrm{PSL}(2,\mathbb{R}) = \mathrm{SL}(2,\mathbb{R})/\{\pm\,\mathrm{id}\}$. The monodromy matrix M (respectively $\mathcal{M}$) changes to a conjugate matrix if we pick a different pair of solutions (f,g) for the same differential equation.

Now regard the ratio $\eta = f/g$ for $\theta \in [0, 2\pi]$ as a map $\eta:\ [0,2\pi] \to \mathbb{R}P^1$ describing a motion ("rotation") along the circle $\mathbb{R}P^1 \simeq S^1$. One can see that the condition $W \neq 0$ on the Wronskian is equivalent to the condition $\eta' = -W/g^2 \neq 0$, i.e., that the rotation "does not stop." Choosing the positive sign of the Wronskian, $W > 0$, we can assume that the rotation always goes in the negative direction: $\eta' < 0$.

Recall that the Virasoro action on η is, in fact, a circle reparametrization for the coordinate θ. By a diffeomorphism change of the coordinate $\theta \mapsto \varphi(\theta)$, one can always turn the map $\eta:\ [0,2\pi] \to \mathbb{R}P^1$ into a *uniform* rotation along $\mathbb{R}P^1$, while keeping the boundary values of $\eta(\theta)$ on the segment $[0,2\pi]$ satisfying the monodromy relation $\eta(\theta + 2\pi) = \eta(\theta)\mathcal{M}$. Furthermore, the number of rotations (the "winding number") for the map $\eta: [0, 2\pi] \to \mathbb{R}P^1$ does not change under a reparametrization by a circle diffeomorphism φ. In other words, the orbits of the maps η (or, equivalently, of the potentials $\{u(x)\}$) are described by one continuous parameter (the conjugacy class of M) and one discrete parameter (the winding number). One can see that these two parameters together encode nothing else but the conjugacy class of the monodromy matrices M in the universal covering of $\mathrm{SL}(2, \mathbb{R})$.

Note that the choice in the sign of the Wronskian reflects the $\mathbb{Z}_2$-action on the universal covering $\widetilde{\mathrm{SL}}(2, \mathbb{R})$. Indeed, with this choice ($W > 0$) the path η always goes in the negative direction ($\eta' < 0$), so one can reach only the "negative half" of the conjugacy classes in the universal cover $\widetilde{\mathrm{SL}}(2, \mathbb{R})$.

Finally, note that the identity matrix in the universal covering $\widetilde{\mathrm{SL}}(2, \mathbb{R})$ (or in its projectivization $\widetilde{\mathrm{SL}}(2, \mathbb{R})/\mathbb{Z}_2$) cannot be obtained as a monodromy matrix for the maps $\eta: [0, 2\pi] \to \mathbb{R}P^1$. Indeed, any map η starting at the identity has to move out from it, since $\eta'(0) \neq 0$. □

Corollary 2.13 *The Virasoro orbits in the hyperplane* $\{a\partial_\theta^2 + u(\theta) \mid a = a_0\} \subset \mathfrak{vir}^*$ *with fixed* $a_0 \neq 0$ *are classified by the Jordan normal form of matrices in* $\mathrm{SL}(2, \mathbb{R})$ *and by a positive integer parameter, the winding number. In this hyperplane* $\{a = a_0\}$ *of the dual* $\mathfrak{vir}^*$ *the orbit containing Hill's operator* $a\partial_\theta^2 + u(\theta)$ *has codimension equal to the codimension in* $\mathrm{SL}(2, \mathbb{R})$ *of the conjugacy class of the monodromy matrix* M *corresponding to this Hill's operator.*

Matrices in the group $\mathrm{SL}(2, \mathbb{R})$ split into three classes, whose normal forms are the exponentials of the following three classes in the Lie algebra $\mathfrak{sl}(2, \mathbb{C})$, the complexification of $\mathfrak{sl}(2, \mathbb{R})$:

$$(i) \begin{pmatrix} \mu & 0 \\ 0 & -\mu \end{pmatrix}, \quad (ii) \begin{pmatrix} 0 & \pm 1 \\ 0 & 0 \end{pmatrix}, \text{ and } \quad (iii) \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}; \tag{2.13}$$

see Figure 2.2. The codimensions of the corresponding conjugacy classes in $\mathrm{SL}(2, \mathbb{R})$ are 1 in cases (i) and (ii), and 3 in case (iii). Note that the set of real matrices that are exponentials of (i)-type matrices consists of the elliptic and hyperbolic parts: rotation matrices (for $\mu \in i\mathbb{R}$) and hyperbolic rotations (for $\mu \in \mathbb{R}$). Furthermore, hyperbolic rotations correspond to one-sheeted hyperboloids. Rotations in the clockwise and counterclockwise directions correspond to different sheets of two-sheeted hyperboloids, and they belong to different conjugacy classes in $\mathrm{SL}(2, \mathbb{R})$. (The rotation by 180° has a three-dimensional stabilizer and corresponds to a one-point conjugacy class.) The group $\mathrm{SL}(2, \mathbb{R})$ is topologically a solid torus, and the adjacency of conjugacy classes described in Figure 2.2 is observed near both id and $-$ id in this group.

The equality of the codimensions of the Virasoro coadjoint orbits in $\mathfrak{vir}^*$ and the codimensions of (the conjugacy classes of) the corresponding

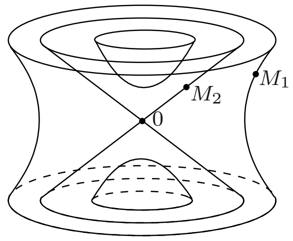


Fig. 2.2. The points M_1, M_2, and 0 in $\mathfrak{sl}(2,\mathbb{R})$ (which is a local picture of the group $\mathrm{SL}(2,\mathbb{R})$) correspond to Virasoro orbits of types (*i*), (*ii*), and (*iii*) respectively.

monodromy matrices in $\mathrm{SL}(2,\mathbb{R})$ follows from the smooth dependence on a parameter in the above classification. (The versal deformations of the Virasoro orbits can be defined in terms of the Jordan–Arnold normal forms of the monodromy matrices depending on a parameter; cf. [15, 233, 306].) Alternatively, one can describe the dimensions of the corresponding stabilizers; see [202, 342]. To visualize (the three-dimensional transversal to) the set of the Virasoro orbits, one can imagine the universal covering of $\mathrm{SL}(2,\mathbb{R})$ as a cylinder filled with an infinite number of copies of Figure 2.2, stacked one on top of another, while the $\mathbb{Z}_2$-quotient keeps only "half" of this infinite cylinder.

Remark 2.14 Regarded as homogeneous spaces, the orbits of type (*i*) are often denoted by $\mathrm{Diff}(S^1)/S^1$, the notation $\mathrm{Diff}(S^1)/\mathbb{R}^1$ stands for (*ii*) (and sometimes for the case $\mu \in \mathbb{R}$ in (*i*)), and $\mathrm{Diff}(S^1)/\mathrm{SL}(2,\mathbb{R})$ corresponds to (*iii*).

To see the reasoning for this, we describe the stabilizers for coadjoint orbits containing constant elements, i.e., Hill's operators $\partial_\theta^2 + u(\theta)$ with constant potentials $u(\theta) \equiv p = \mathrm{const}$. For such an operator, the corresponding monodromy matrix $M_p \in \mathrm{SL}(2,\mathbb{R})$ is given explicitly:

$$M_p = \begin{pmatrix} \cos(2\pi\sqrt{p}) & \frac{1}{\sqrt{p}}\sin(2\pi\sqrt{p}) \\ -\sqrt{p}\sin(2\pi\sqrt{p}) & \cos(2\pi\sqrt{p})\,, \end{pmatrix} \qquad \text{for } \ p > 0\,,$$

$$M_p = \begin{pmatrix} \cosh(2\pi\sqrt{-p}) & \frac{1}{\sqrt{-p}}\sinh(2\pi\sqrt{-p}) \\ \sqrt{-p}\sinh(2\pi\sqrt{-p}) & \cosh(2\pi\sqrt{-p}) \end{pmatrix} \qquad \text{for } \ p < 0\,,$$

and

$$M_0 = \begin{pmatrix} 1 & 2\pi \\ 0 & 1 \end{pmatrix} \qquad \text{for } \ p = 0\,.$$

One can see that for Hill's differential operators with potentials $p < 0$ or $p = 0$ the stabilizer is $\mathbb{R}$. In the case $p > 0$ it is the group S^1 of rigid rotations, provided that $M_p \neq \pm\,\mathrm{id} \in \mathrm{SL}(2,\mathbb{R})$. Finally, the stabilizer is three-dimensional, once the monodromy M_p is plus or minus the identity matrix, i.e., the exponential of type (*iii*) in the list (2.13). This can be the case only if $p = m^2/4$

for some $m \in \mathbb{N}$. In the latter case, the stabilizer of the corresponding Hill's equation is the m-fold covering of $\mathrm{PSL}(2,\mathbb{R})$ in $\mathrm{Diff}(S^1)$.

Finally, we note that the trace function $\mathrm{tr} : \mathrm{SL}(2,\mathbb{R}) \to \mathbb{R}$ is invariant on conjugacy classes in $\mathrm{SL}(2,\mathbb{R})$ and it gives rise to a Casimir function on the hyperplane $\{a = a_0\} \subset \mathfrak{vir}^*$ for the Lie–Poisson bracket: $\mathrm{tr}\{a_0\partial_\theta^2 + u(\theta)\} := \mathrm{tr}(M)$, where M is a monodromy matrix of the given Hill operator. Along with the value of a, this is the only Casimir for generic Virasoro orbits, since their codimension in the hyperplane $\{a = a_0\}$ is equal to 1.

2.4 The Virasoro–Bott Group and the Korteweg–de Vries Equation

The *Korteweg–de Vries* (or KdV) equation is the nonlinear evolution equation

$$u_t = -3uu' - au''' \,, \tag{2.14}$$

which describes traveling waves in a shallow canal. Here, u is a function of the time variable t and one space variable θ, u_t and u' denote the corresponding partial derivatives in t and θ, and a is a nonzero constant. A brief history of this equation can be found in [294].

In this section we show how the Korteweg–de Vries equation appears as the Euler equation with respect to a certain right-invariant metric on the Virasoro–Bott group. Recall that the Euler equation with respect to a right-invariant metric on a Lie group G is a dynamical system on the corresponding Lie algebra $\mathfrak{g}$ describing the evolution of the tangent vector along a geodesic on G, where this vector is pulled back to the Lie algebra of G by right translation; see Section I.4.

Consider the L^2-inner product on the Virasoro algebra $\mathfrak{vir} = \mathrm{Vect}(S^1) \oplus \mathbb{R}$ defined by

$$\langle (v(\theta)\partial_\theta, a), (w(\theta)\partial_\theta, c)\rangle = \int_{S^1} v(\theta)w(\theta)d\theta + a \cdot c \,. \tag{2.15}$$

Extend this quadratic form to every tangent space on the Virasoro–Bott group by right translations to define a (weak) right-invariant L^2-metric on the group.

Theorem 2.15 ([305]) *The Euler equation for the right-invariant L^2-metric on the Virasoro group is (the family of) the KdV equation:*

$$u_t = -3uu' - au''' \,, \tag{2.16}$$

$$a_t = 0 \,. \tag{2.17}$$

PROOF. According to Arnold's Theorem I.4.14, the Euler equation on $\mathfrak{g}^*$ for the *right-invariant* metric on the group G has the form

$$\frac{d}{dt}m(t) = \mathrm{ad}^*_{A^{-1}m(t)}\, m(t)\,, \tag{2.18}$$

where $m(t)$ is a point in $\mathfrak{g}^*$, and $A : \mathfrak{g} \to \mathfrak{g}^*$ is the inertia operator defined by the metric (using the left-invariant metric would give the minus sign in the equation; cf. Remark I.4.16).

In the Virasoro coadjoint action, an element $(v\partial_\theta, c) \in \mathfrak{vir}$ of the Virasoro algebra acts on an element $(u(d\theta)^2, a) \in \mathfrak{vir}^*$ of the dual space via

$$\mathrm{ad}^*_{(v\partial_\theta,c)}(u(d\theta)^2, a) = ((-2v'u - vu' - av''')(d\theta)^2, 0)\,;$$

see equation (2.8). Furthermore, the L^2-inner product gives rise to the "identity" inertia operator $A : \mathfrak{vir} \to \mathfrak{vir}^*$:

$$(u\partial_\theta, a) \mapsto (u(d\theta)^2, a)\,,$$

mapping a vector field $u(\theta)\partial_\theta$ to the quadratic differential $u(\theta)(d\theta)^2$ with the same function $u(\theta)$.

Then, substituting $(u(d\theta)^2, a)$ for m, we see that the Euler equation (2.18) becomes

$$\frac{d}{dt}(u(d\theta)^2, a) = ((-3uu' - au''')(d\theta)^2, 0)\,,$$

from which one immediately reads off the KdV equation (2.16), (2.17). □

The component a does not change with time (see equation (2.17))and plays the role of a constant parameter in the KdV equation. It has the physical meaning of the characteristic thickness of the shallow-water approximation (see, e.g., [228], p. 169).

Remark 2.16 One can study more general metrics on the Virasoro algebra, some of which are of particular interest in mathematical physics. Consider, for instance, the following two-parameter family of weighted $H^1_{\alpha,\beta}$-inner products on $\mathfrak{vir}$:

$$\langle (v\partial_\theta, b), (w\partial_\theta, c)\rangle_{H^1_{\alpha,\beta}} = \int_{S^1} (\alpha\, vw + \beta\, v'w')\, d\theta + bc\,.$$

The case $\alpha = 1, \quad \beta = 0$ corresponds to the L^2 inner product above, while $\alpha = \beta = 1$ corresponds to the H^1 product.

Theorem 2.17 ([192]) *The Euler equations for the right-invariant $H^1_{\alpha,\beta}$-metric (with $\alpha \neq 0$) on the Virasoro group are given by the following system:*

$$\alpha(u_t + 3uu') - \beta((u'')_t + 2u'u'' + uu''') + au''' = 0\,, \tag{2.19}$$

$$a_t = 0\,. \tag{2.20}$$

Exercise 2.18 Give a proof of the latter theorem along the lines of the proof of the L^2-case above. (Hint: The inertia operator for the weighted H^1 metric is

$$A : (v\partial_\theta, a) \mapsto ((\Lambda v)(d\theta)^2, \, a),$$

where $\Lambda := \alpha - \beta\partial_\theta^2$ is a second-order differential operator. Verify that in terms of $v = \Lambda^{-1}u$ the Euler equation has the form

$$\frac{d}{dt}(\Lambda v) = -2(\Lambda v)v' - (\Lambda v')v + av''' \,,$$

which is equivalent to equation (2.19).)

Remark 2.19 For $\alpha = 1$, $\beta = 0$, equation (2.19) is the KdV equation (2.16). For $\alpha = \beta = 1$, one recovers the *Camassa–Holm equation* (see [268]). For $\alpha = 0$, $\beta = 1$, equation (2.19) becomes the *Hunter–Saxton equation*. We note that in the case of $\alpha = 0$, the $H^1_{\alpha,\beta}$-metric becomes the homogeneous $\dot{H}^1$-metric, which is degenerate. Therefore, to define the Euler equations one has to pass to the homogeneous space $\widehat{\mathrm{Diff}}(S^1)/S^1$ (or $\mathrm{Diff}(S^1)/S^1$) and define the geodesic flow on it; see details in [192]. It turns out that the space $\mathrm{Diff}(S^1)/S^1$ equipped with the $\dot{H}^1$-metric is isometric to an open subset of an L^2-sphere; see [237, 238]. This isometry, in particular, allows one to extend solutions of the Hunter–Saxton equation beyond breaking time and interpret them after wave-breaking in an appropriate weak sense.

We also note that the case $a = 0$ corresponds to the nonextended Lie algebra $\mathrm{Vect}(S^1)$ of vector fields on the circle, rather than to the Virasoro algebra $\mathfrak{vir}$. In the nonextended case, depending on the values of α and β, one obtains the Hopf (or inviscid Burgers) equation $u_t + 3uu' = 0$ or the nonextended Camassa–Holm equation [69, 127]

$$u_t + 3uu' + 2u'u'' + uu''' + (u'')_t = 0 \,.$$

2.5 The Bi-Hamiltonian Structure of the KdV Equation

The KdV equation is not only a Hamiltonian system; it also exhibits strong integrability properties. As we discussed before, there are various definitions of what an integrable infinite-dimensional system is: one can require from the system either the existence of action-angle coordinates, or the existence of "sufficiently many" independent integrals of motion, or some other properties, which may differ substantially in infinite dimensions. In this section we show that the KdV equation is not only Hamiltonian, but in fact bi-Hamiltonian, thus exhibiting one of the "strongest forms" of integrability. More precisely, in addition to being Hamiltonian with respect to the Lie–Poisson bracket on the dual space $\mathfrak{vir}^*$, this equation turns out to be Hamiltonian with respect to another compatible Poisson structure on the same space.

Recall that for any Lie algebra $\mathfrak{g}$, every point $m_0 \in \mathfrak{g}^*$ gives rise to a "constant" Poisson bracket $\{\ ,\ \}_0$ on $\mathfrak{g}^*$ by "freezing" the usual Lie–Poisson bracket $\{\ ,\ \}_{LP}$ at the point m_0. The constant Poisson bracket is defined for smooth functions f, g on $\mathfrak{g}^*$ by

$$\{f, g\}_0(m) := \langle [df_m, dg_m], m_0 \rangle \,;$$

see Section I.4.4. Furthermore, the Poisson brackets $\{\ ,\ \}_{LP}$ and $\{\ ,\ \}_0$ are compatible for all choices of the point m_0 (see Lemma I.4.21). The main goal of this section is to show that for a certain choice of the "freezing point" $m_0 \in \mathfrak{vir}^*$, the KdV equation is Hamiltonian with respect to the constant Poisson structure $\{\ ,\ \}_0$. Note that other equations discussed above (Camassa–Holm, Hunter–Saxton) have a similar bi-Hamiltonian structure, but related to different choices of the "freezing point"; see [192].

Theorem 2.20 *The KdV equation (2.14) is Hamiltonian with respect to the constant Poisson bracket on* $\mathfrak{vir}^*$ *with the "freezing point"* $m_0 = (\frac{1}{2}(d\theta)^2, 0) \in \mathfrak{vir}^*$.

PROOF. Let $F(u, a)$ be a function on $\mathfrak{vir}^*$ and let $(v\partial_\theta, b) := (\delta F/\delta u, \delta F/\delta a) \in \mathfrak{vir}$ be the (variational) derivative $dF_{(u(d\theta)^2, a)}$ of F at $(u(d\theta)^2, a)$. Then the Hamiltonian equation with the Hamiltonian function F, computed with respect to the constant Poisson structure "frozen" at $(u_0(dx)^2, a_0)$, has the form

$$\frac{d}{dt}(u(d\theta)^2, a) = \mathrm{ad}^*_{(v\partial_\theta, b)}(u_0(d\theta)^2, a_0) = -\big((2u_0v' + (u_0)'v + a_0v''')(d\theta)^2, 0\big)\,.$$

(Here we use Remark I.4.22 and the explicit form (2.8) of the coadjoint action ad^* for the Virasoro algebra.)

Now specifying the "freezing" point to $(u_0(d\theta)^2, a_0) = (\frac{1}{2}(d\theta)^2, 0) \in \mathfrak{vir}^*$, we come to

$$\frac{d}{dt}(u(d\theta)^2, a) = -(v'(d\theta)^2, 0)\,, \tag{2.21}$$

where v is defined as the "partial derivative" of the functional F, i.e., $v\partial_\theta = \delta F/\delta u$.

Next, consider the functional F of the form

$$F(u, a) = \int_{S^1} \left(\frac{1}{2}u^3 - \frac{a}{2}(u')^2\right) d\theta\,.$$

By definition, the *variational derivative* $(\delta F/\delta u, \delta F/\delta a) \in \mathfrak{vir}$ of the functional F is determined by the following identity satisfied for any $(\xi(d\theta)^2, c) \in \mathfrak{vir}^*$:

$$\left\langle (\xi(d\theta)^2, c)\,, \left(\frac{\delta F}{\delta u}, \frac{\delta F}{\delta a}\right)\right\rangle = \frac{d}{d\epsilon}\Big|_{\epsilon=0} F(u + \epsilon\xi, a + \epsilon c).$$

For equation (2.21) we need only the partial derivative $\delta F/\delta u$, and we find it as follows:

$$
\begin{aligned}
\frac{d}{d\epsilon}\Big|_{\epsilon=0} F(u+\epsilon\xi, a) &= \frac{d}{d\epsilon}\Big|_{\epsilon=0} \int_{S^1} \left(\frac{1}{2}(u+\epsilon\xi)^3 - \frac{a}{2}(u'+\epsilon\xi')^2\right) d\theta \\
&= \int_{S^1} \left(\frac{3}{2}u^2\xi - au'\xi'\right) d\theta = \int_{S^1} \left(\frac{3}{2}u^2\xi + au''\xi\right) d\theta \\
&= \left\langle \left(\left(\frac{3}{2}u^2 + au''\right)\partial_\theta, a\right), (\xi(d\theta)^2, 0)\right\rangle .
\end{aligned}
$$

Hence we obtain the derivative $\delta F/\delta u = (\frac{3}{2}u^2 + au'')\partial_\theta$. Since $v\partial_\theta = \delta F/\delta u$, now we substitute $v = \frac{3}{2}u^2 + au''$ into the equation (2.21), which gives

$$
u_t = -3uu' - au''' ,
$$

the KdV equation. □

Corollary 2.21 *The KdV equation is bi-Hamiltonian with respect to the compatible Poisson structures* $\{\ ,\ \}_{LP}$ *and* $\{\ ,\ \}_0$ *on* $\mathfrak{vir}^*$*, where* $\{\ ,\ \}_0$ *denotes the constant Poisson structure, frozen in the point* $(\frac{1}{2}(d\theta)^2, 0) \in \mathfrak{vir}^*$.

The constant bracket for the KdV is usually called the *first KdV structure*, or the Gardner–Faddeev–Zakharov bracket [140, 392], while the linear Lie–Poisson structure is called the *second KdV*, or the Magri *bracket* [246]. The analogues of these structures for higher-order differential operators are called the first and second Adler–Gelfand–Dickey structures; see Section 4.

Remark 2.22 Similarly, one can show that both the Camassa–Holm and Hunter–Saxton equations are bi-Hamiltonian with respect to the Lie–Poisson structure and a constant structure on $\mathfrak{vir}^*$. The respective "freezing" points of the constant Poisson bracket on $\mathfrak{vir}^*$ are $m_1 = (\frac{1}{2}(d\theta)^2, 1)$ for the Camassa–Holm equation and $m_2 = (0, 1)$ for the Hunter–Saxton equation (see Figure 2.3).

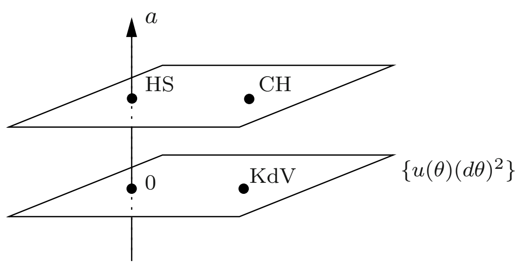


Fig. 2.3. The freezing points of the constant Poisson bracket on $\mathfrak{vir}^*$ that give rise to the bi-Hamiltonian structures for the KdV, the Camassa–Holm, and the Hunter–Saxton equations respectively.

Remark 2.23 The bi-Hamiltonian nature of the KdV equation allows one to obtain the whole hierarchy of the KdV Hamiltonians via the Lenard–Magri scheme discussed in the introduction.

Namely, consider the linear combination of the Poisson brackets for the KdV: the Virasoro Lie–Poisson and the constant ones:

$$\{\ ,\ \}_\lambda := \{\ ,\ \}_{LP} + \lambda^2\{\ ,\ \}_0\,.$$

Since the corresponding brackets are compatible, $\{\ ,\ \}_\lambda$ defines a Poisson bracket for all $\lambda \in \mathbb{R}$, which is the usual Lie–Poisson bracket on $\mathfrak{vir}^*$ shifted in the direction of $m_0 = (\frac{1}{2}(d\theta)^2, 0) \in \mathfrak{vir}^*$. (Note that one uses here the parameter λ^2 rather than λ for the expansion below to have a simpler form.)

Recall that we think of the dual space $\mathfrak{vir}^* = \{a\partial_\theta^2 + u(\theta)\}$ as the space of Hill's operators. The monodromy of a differential operator $\partial_\theta^2 + u(\theta)$, which is a matrix in $\mathrm{SL}(2,\mathbb{R})$, changes to a conjugate matrix under the Virasoro coadjoint action (see Theorem 2.12). Therefore, the trace of the monodromy is a Casimir function for the Poisson bracket $\{\ ,\ \}_{LP}$ on (the hyperplane $a = 1$ in) the dual space $\mathfrak{vir}^*$. The same reasoning allows one to obtain the following result.

Lemma 2.24 *Let M_λ denote the monodromy of the differential operator $\frac{d^2}{d\theta^2} + u(\theta) - \lambda^2$. Then the function*

$$h_\lambda := \log(\mathrm{tr}(M_\lambda))$$

is a Casimir function for the Poisson bracket $\{\ ,\ \}_\lambda$ on the space $\mathfrak{vir}^*$.

Indeed, $\{\ ,\ \}_\lambda$ is the usual Lie–Poisson bracket on $\mathfrak{vir}^*$ shifted in the direction of $m_0 = (0\cdot\partial_\theta^2 + 1/2) \in \mathfrak{vir}^*$, and so, instead of the Lie–Poisson Casimir $\mathrm{tr}(M)$, we can use the shifted Casimir $\mathrm{tr}(M_\lambda)$ or any function of it, in particular, $\log(\mathrm{tr}(M_\lambda))$.

Finally, by expanding h_λ into a power series in λ^{-1} one produces first integrals of the KdV equation:

$$h_\lambda = 2\pi\lambda - \sum_{n=1}^{\infty} h_{2n-1}\lambda^{1-2n}\,,$$

where

$$h_1 = \frac{1}{2}\int_{S^1} u\, d\theta, \quad h_3 = \frac{1}{8}\int_{S^1} u^2\, d\theta, \quad h_5 = \frac{1}{16}\int_{S^1} (u^3 - \frac{1}{2}(u')^2)\, d\theta, \quad \dots;$$

see, e.g., [31, 37].

2.6 Bibliographical Notes

The cohomology of the Lie algebra of vector fields on the circle was computed by Gelfand and Fuchs in [143], where the term Virasoro algebra was coined after the paper [376] (see also [118, 138]). The Bott cocycle on the group $\mathrm{Diff}(S^1)$ first appeared in [53].

The group $\mathrm{Diff}(S^1)$ (and hence the Virasoro group) does not admit a natural complexification, i.e., a group corresponding to the complexified Lie algebra $\mathrm{Vect}(S^1)_{\mathbb{C}}$; see [205, 322] and Example I.1.25. However, for a cone in this complex Lie algebra consisting of those vector fields that "point outside of the circle," there exists a semigroup. This *annulus semigroup* appeared in the papers of Neretin [291] and Segal [343]. For more details on representations and applications of the Virasoro group and algebra see [292, 322, 153].

The classification of the coadjoint orbits of the Virasoro group can be found in the literature under different guises: as a classification of projective structures on the circle by Kuiper [223], as a classification of Hill's operators by Lazutkin and Pankratova [233], and in the present form, as Virasoro orbits, in the papers by Kirillov [202, 205] and Segal [342]; see also [164, 304, 388, 32]. The adjoint orbits of the diffeomorphism group of the circle were described in [164].

The Virasoro coadjoint orbit $\mathrm{Diff}(S^1)/S^1$ can also be understood as the universal Teichmüller space [326]; cf. [284, 285]. Curvatures of a Kähler metric on this orbit were described in [208], while its complex Hilbert manifold structure is discussed in [338, 363]. The Virasoro group itself admits a complex structure and can be viewed as a holomorphic $\mathbb{C}^*$-bundle over its orbit $\mathrm{Diff}(S^1)/S^1$ [236].

There is a vast literature related to the geometry and Hamiltonian properties of the KdV equation, which is one of the key examples in any book on soliton theory. The description of the KdV equation, as well as its super-analogue, as an Euler–Arnold equation on the Virasoro group with respect to the L^2-metric can be found in [305, 344, 346]. More general, H^1-type, metrics on this group were considered in [268, 189, 192], and we followed the latter paper in our exposition. Regularity properties of the Riemannian exponential maps for these and other Sobolev metrics on the Virasoro and the diffeomorphism groups are described in [73, 74].

The Adler–Gelfand–Dickey structures [3, 141, 142] are the generalizations of the KdV Poisson structures from Hill's operators to linear differential operators of higher order, and we discuss them in Section 4.

For the algebro-geometric approach to constructing solutions of the KdV equation we address the reader to [216]; the description of the corresponding infinite-dimensional Grassmann manifolds can be found in [347] and the references therein. Various analytical aspects of the KdV theory, its spectral theory, and the angle-action variables are discussed in the book [182]. The geometry related to the KAM theory for near-integrable Hamiltonian systems, applied in the infinite-dimensional context, e.g., to the KdV-type equations, is discussed in [224, 182].

3 Around the Chern–Simons Functional

In this section we move from surfaces to threefolds, along with connections and bundles on them. By studying the topological Chern–Simons action functional on the space of connections on a three-dimensional manifold with boundary, we recover the definition of the symplectic structure on the moduli space of flat connections on a compact Riemann surface. Similarly, the holomorphic Chern–Simons functional on $\bar{\partial}$-connections over three-dimensional Fano manifolds is related to the holomorphic symplectic structure on the moduli spaces of stable bundles over $K3$ or abelian surfaces.

Furthermore, the corresponding path integrals for these Chern–Simons functionals in the abelian case can be used to define the Gauss linking number of oriented curves in three-dimensional space and its holomorphic analogue, the polar linking number of holomorphic curves.

3.1 A Reminder on the Lagrangian Formalism

A motion of a particle on a manifold can be described by the least action principle. Consider an action functional

$$S[q] = \int_{t_0}^{t_1} L(q(t), \dot{q}(t), t)\, dt$$

defined on the space $\mathcal{C}[t_0, t_1]$ of smooth maps $q : [t_0, t_1] \to M$ of the interval $[t_0, t_1]$ to the manifold M. Here L is a (time-dependent) Lagrangian function, $L : TM \times \mathbb{R} \to \mathbb{R}$, which we assume to depend only on t, q, and its first derivative $\dot{q} := dq/dt$.

For a path variation δq one can find the corresponding *variation of the action functional*, i.e., the linear-in-δq term of the difference $S[q + \delta q] - S[q]$:

$$\delta S[q] = \int_{t_0}^{t_1} E\, \delta q\, dt + p\, \delta q|_{t_0}^{t_1} ,$$

where

$$E := \frac{\partial L(q, \dot{q}, t)}{\partial q} - \frac{d}{dt} \frac{\partial L(q, \dot{q}, t)}{\partial \dot{q}}$$

and

$$p := \frac{\partial L(q, \dot{q}, t)}{\partial \dot{q}} .$$

(Here and below we assume the summation over the coordinates $q = (q^1, \dots, q^d)$: $p\delta q := \sum_j p_j \delta q^j , \quad p_j := \partial L(q, \dot{q}, t)/\partial \dot{q}^j$, etc.)

Exercise 3.1 Prove the variation formula. (Hint: use integration by parts.)

This way the variation δS can be regarded as a 1-form on the infinite-dimensional space $\mathcal{C}[t_0, t_1]$ of "virtual trajectories" of the particle.

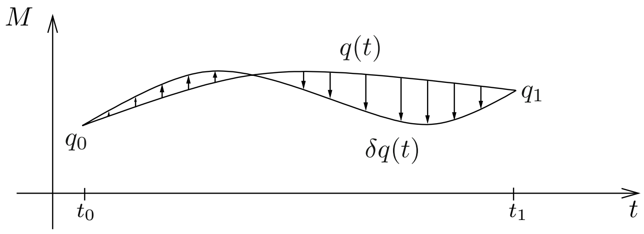


Fig. 3.1. A small variation of the path $q(t)$ with fixed endpoints.

Definition 3.2 The *least action principle* states that the actual trajectories of the particle are the critical points of this action functional: $\delta S[q] = 0$.

By confining ourselves to variations with fixed ends, $\delta q(t_0) = \delta q(t_1) = 0$, we come to a necessary condition on the extremals. Namely, actual particle trajectories satisfy the *Euler–Lagrange equation* $E = 0$, i.e.,

$$\frac{\partial L(q, \dot{q}, t)}{\partial q} - \frac{d}{dt}\frac{\partial L(q, \dot{q}, t)}{\partial \dot{q}} = 0\,.$$

Denote by $\mathcal{E}[t_0, t_1]$ the space of all solutions to the Euler–Lagrange equation, i.e., the space of such trajectories.

Exercise 3.3 A free particle of mass m moving in the space $\mathbb{R}^d$ with a potential energy $V : \mathbb{R}^d \to \mathbb{R}$ has the Lagrangian $L(q, \dot{q}, t) = m|\dot{q}|^2/2 - V(q)$, the difference of its kinetic and potential energies. Prove that the Euler–Lagrange equation for this L gives the *Newton equation* of motion:

$$m\ddot{q} = -\operatorname{grad} V(q)\,.$$

Now we restrict the variation 1-form δS to the space of extremals $\mathcal{E}[t_0, t_1]$, which is singled out by the Euler–Lagrange equation. On this space of "trajectories with free ends" we obtain

$$\delta S = p\,\delta q|_{t_0}^{t_1} = \sigma_1 - \sigma_0\,, \tag{3.5}$$

where $\sigma_i := p\,\delta q|_{t_i}$, $i = 0, 1$ are the corresponding 1-forms on $\mathcal{C}[t_0, t_1]$. One can regard the above as a relation between these three 1-forms: σ_0, σ_1, and δS, which holds for their restrictions to the space of extremals $\mathcal{E}[t_0, t_1]$.

Now, by applying the exterior differential δ (on the infinite-dimensional manifold $\mathcal{C}[t_0, t_1]$) to both sides of the relation (3.5) above and using $\delta^2 = 0$, we obtain $\delta\sigma_0 = \delta\sigma_1$, which holds on $\mathcal{E}[t_0, t_1]$. This means that the space $\mathcal{E}[t_0, t_1]$ turns out to be naturally equipped with a closed 2-form ω defined by

$$\omega := \delta\sigma_0 = \delta\sigma_1\,.$$

Definition 3.4 A manifold N equipped with a closed 2-form ω (not necessarily nondegenerate) is called *presymplectic*.

Consider the distribution of null-spaces of this 2-form in N.

Exercise 3.5 (i) Assuming that this distribution has constant rank, prove that it is integrable, i.e., it is tangent to a foliation in N.

(ii) Assuming that this null-foliation is a fibration $\pi : N \to N'$, prove that the base of this fibration carries a natural symplectic structure, i.e., (N', ω') is a symplectic manifold such that $\pi^* \omega' = \omega$.

The above discussion shows that whenever the space of extremals $\mathcal{E}[t_0, t_1]$ is a manifold, it is in fact a *presymplectic* manifold. However, the 2-form ω is often degenerate. The *phase space* $\mathcal{P}$ of the particle can be described as the corresponding *symplectic* manifold. (Here we implicitly assume that various regularity conditions are satisfied to guarantee that both $\mathcal{E}[t_0, t_1]$ and the phase space are smooth manifolds.)

Exercise 3.6 Check that for the above example of a particle motion in $\mathbb{R}^d$ this definition of the phase space $\mathcal{P}$ coincides with $T^*\mathbb{R}^d$ equipped with the natural symplectic structure.

Remark 3.7 [393, 79, 341] The discussed Lagrangian formalism can be generalized to infinite-dimensional target manifolds M or to higher-dimensional domains instead of the interval $[t_0, t_1]$. These are the objects that a field theory deals with. Consider, for example, a local action functional

$$S[\varphi] = \int_N L(\varphi(x), \partial\varphi(x))\, d^n x$$

describing a field theory on an n-dimensional manifold N with boundary ∂N. Here $x = (x_1, \ldots, x_n)$ are local coordinates on N, φ is a map from N to a target manifold M or a section of some bundle on N, $\partial\varphi$ are the first derivatives of φ, while the Lagrangian L can depend on additional structures on N. As in the one-dimensional situation described above, one can pose a variational problem $\delta S[\varphi] = 0$, which leads to the Euler–Lagrange equations.

Suppose first that $N = I \times \Sigma$, where I is an interval and a manifold Σ has dimension $n-1$. One can consider $t \in I$ as the time variable and identify the field theory with an infinite-dimensional classical mechanics, where the space of maps $\varphi : \Sigma \to M$ plays the role of the target. In particular, one has a presymplectic manifold of extremals $\mathcal{E}_N$ and the symplectic phase space $\mathcal{P}$ associated to N (or, rather, to Σ).

Alternatively, one can associate the phase spaces $\mathcal{P}_0$ and $\mathcal{P}_1$ to the corresponding boundary components $\partial N = \Sigma_1 - \Sigma_0$ of N, and equip the total phase space $\mathcal{P}_0 \times \mathcal{P}_1$ with the product symplectic structure. There is a natural projection α_N of the space $\mathcal{E}_N$ of extremals into the product $\mathcal{P}_0 \times \mathcal{P}_1$, since it "tautologically" projects to each factor: one describes the extremals via different boundary components, taking the orientation of the latter into account. Then the relation $0 = \delta^2 S = \delta\sigma_1 - \delta\sigma_0$ that held on $\mathcal{E}_N$ now reads that the image $\alpha_N(\mathcal{E}_N)$ of $\mathcal{E}_N$ is an *isotropic submanifold* in the symplectic manifold $\mathcal{P}_0 \times \mathcal{P}_1$.

Definition 3.8 A submanifold of a symplectic manifold is *isotropic* if the restriction of the symplectic form to this submanifold is zero.

Exercise 3.9 Let $f : (N_1, \omega_1) \to (N_2, \omega_2)$ be a diffeomorphism between two symplectic manifolds. Prove that f is a symplectic map, i.e., $\omega_1 = f^*\omega_2$, if and only if the graph of f is an isotropic submanifold in the symplectic manifold $(N_1 \times N_2,\, \omega_1 \ominus \omega_2)$.

(An isotropic submanifold of maximal possible dimension, which is equal to half the dimension of the symplectic manifold, is called a *Lagrangian submanifold*; cf. Section I.4.5. This is the case for the graph of f.)

One can see that the image $\alpha_N(\mathcal{E}_N)$ is indeed isotropic in $\mathcal{P}_0 \times \mathcal{P}_1$, since the 2-form $\delta\sigma_1 - \delta\sigma_0$ is exactly the restriction of the product symplectic structure of $\mathcal{P}_0 \times \mathcal{P}_1$ (with different orientations of the boundary components) to this image.

The latter formulation of the presymplectic/isotropic properties of the space of extremals $\mathcal{E}_N$ extends naturally to the general case of a manifold N with boundary consisting of several components $\Sigma_1, \ldots, \Sigma_k$. Associate the phase space $\mathcal{P}_j$ to each component Σ_j, thinking of a neighborhood of Σ_j in N as a product $I \times \Sigma$. One has the relations $\delta S = \sigma_1 + \cdots + \sigma_k$ and $\delta\sigma_1 + \cdots + \delta\sigma_k = 0$ on the space of extremals $\mathcal{E}_N$, where σ_j stands for the contribution of the corresponding boundary component. The latter shows that the image $\alpha_N(\mathcal{E}_N)$ under the natural map $\alpha_N : \mathcal{E}_N \to \mathcal{P}_1 \times \cdots \times \mathcal{P}_k$ is *isotropic* with respect to the product symplectic structure on the phase space $\mathcal{P}_1 \times \cdots \times \mathcal{P}_k$. We refer to [341, 79] for more details.

Remark 3.10 The philosophy of holomorphic orientation (see Sections 2.2 and 2.3) can be applied to field-theoretic notions in the following way. Suppose we have an action functional

$$\mathcal{S}[\varphi] = \int_M L(\varphi, \partial\varphi)\, d^n x$$

on *smooth* fields φ (e.g., functions, connections, etc.) on a *real* (oriented) manifold M, and this functional is defined by an n-form $L\, d^n x$, which depends on the fields and their derivatives.

Then one can suggest the following complex analogue $\mathcal{S}_{\mathbb{C}}$ of the action functional $\mathcal{S}$ for a *complex* n-dimensional manifold X equipped with a "polar orientation," i.e., with a holomorphic or meromorphic n-form μ:

$$\mathcal{S}_{\mathbb{C}}[\varphi] := \int_X \mu \wedge L(\varphi, \bar{\partial}\varphi)\, d^n \bar{x}\,.$$

Here φ stands for *smooth* fields on a complex manifold X. Now the $(0,n)$-form $L\,d^n\bar{x}$ is integrated against the holomorphic orientation μ over X.

Furthermore, the interrelation between the extremals of the real functional $\mathcal{S}[\varphi]$ (on smooth fields) on the real manifold M and the boundary values of those fields on ∂M is replaced by the analogous interrelation for the complex functional $\mathcal{S}_{\mathbb{C}}[\varphi]$ (still on smooth fields) on a complex manifold X (equipped with an n-form μ) and on the polar divisor $Y := \mathrm{div}_\infty \mu \subset X$ (equipped with the residue $(n-1)$-form $\nu := \mathrm{res}\ \mu$).

The above discussion will allow us to see in the next two sections how the symplectic structures on the moduli of flat connections and holomorphic bundles on surfaces arise naturally from the Lagrangian formalism related to the topological and holomorphic Chern–Simons functionals.

3.2 The Topological Chern–Simons Action Functional

Let N be a real compact oriented three-dimensional manifold with boundary $\partial N = \Sigma$. As usual in the "real case," we take G to be a compact simply

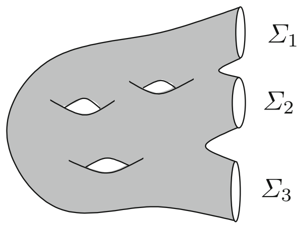


Fig. 3.2. Three-dimensional manifold N with boundary $\partial N = \Sigma_1 \cup \Sigma_2 \cup \Sigma_3$.

connected simple Lie group with the corresponding Lie algebra $\mathfrak{g}$. Denote the nondegenerate invariant (Killing) bilinear form on $\mathfrak{g}$ by $\mathrm{tr}(XY) := \langle X, Y\rangle$. Fix a trivial G-bundle E over N and let $\mathcal{A}$ denote the space of connections in the bundle E. Upon fixing a reference flat connection, we think of $\mathcal{A}$ as the space $\Omega^1(N,\mathfrak{g})$.

Definition 3.11 The topological *Chern–Simons action functional* is the following real-valued function on the space of connections $\mathcal{A}$:

$$\mathrm{CS}(A) := \int_N \mathrm{tr}(A \wedge dA) + \frac{2}{3}\int_N \mathrm{tr}(A \wedge A \wedge A)\,,$$

where a connection $A \in \mathcal{A}$ is understood as a $\mathfrak{g}$-valued 1-form on N.

Proposition 3.12 *The set of extremals, i.e., solutions of the Euler–Lagrange equation, for the Chern–Simons functional* CS *is the space of flat connections in the G-bundle E over the manifold N.*

PROOF. For a small variation δA of a connection $A \in \mathcal{A}$ the corresponding variation of the functional is

$$\begin{aligned}\delta\,\mathrm{CS} &= \int_N \mathrm{tr}(\delta A \wedge dA) + \int_N \mathrm{tr}(A \wedge d\delta A) + 2\int_N \mathrm{tr}(\delta A \wedge A \wedge A) \\ &= \int_N d\ \mathrm{tr}(A \wedge \delta A) + 2\int_N \mathrm{tr}\left(\delta A \wedge (dA + A \wedge A)\right) \\ &= \int_{\partial N} \mathrm{tr}(A \wedge \delta A) + 2\int_N \mathrm{tr}\left(\delta A \wedge (dA + A \wedge A)\right)\,,\end{aligned}$$

where at the last step we used the Stokes formula.

By imposing the boundary condition $\delta A|_{\partial N} = 0$ on variations δA, we obtain the Euler–Lagrange equation

$$dA + A \wedge A = 0\,,$$

i.e., the equation of vanishing curvature $F(A) = 0$ on N. Hence the space of solutions of this equation is exactly the space of flat connections on the real threefold N. □

The first term in the above calculation of $\delta\,\mathrm{CS}$ gives the boundary contribution, the 1-form $\sum \sigma_j$ on the extremals, where the summation is taken over the boundary components of ∂N. Take $N = I \times \Sigma$ to be a finite cylinder over a closed two-dimensional surface Σ. Then the presymplectic structure on the space of flat connections on N, i.e., on the extremals for our action functional, is $\omega = \delta\sigma$ for

$$\sigma := \int_\Sigma \mathrm{tr}(a \wedge \delta a)\,,$$

where $a := A|_\Sigma$ denotes the restriction of a flat connection A from the manifold N to either of its boundary components Σ. (Here we omit the index $j = 0, 1$ for σ_j, since $\omega = \delta\sigma_0 = \delta\sigma_1$.)

Exercise 3.13 Verify that the 2-form $\omega = \delta\sigma$ is degenerate on the space of flat connections on the surface Σ exactly along the gauge equivalence classes of the connections $\{a\}$. (Hint: the 2-form $\delta\sigma = \int_\Sigma \mathrm{tr}(\delta a \wedge \delta a)$ is the restriction of the canonical 2-form ω from the set of all connections to the subset of flat connections on Σ; cf. Definition 2.1.)

Thus the moduli space of flat connections $\mathcal{M}^\Sigma$ on the surface Σ appears as the natural symplectic (or phase) space for this presymplectic space of flat connections on Σ, and we obtain yet another definition of the symplectic structure on $\mathcal{M}^\Sigma$ from Section 2.1.

Corollary 3.14 *The moduli space $\mathcal{M}^\Sigma$ of flat connections on a surface Σ is naturally symplectic as the phase space for extremals of the Chern–Simons action functional for connections on the threefold $N = I \times \Sigma$.*

Remark 3.15 To see why this action functional is called *topological* we now check the invariance property of the Chern–Simons action with respect to gauge transformations of the connections. Let M be a compact three-dimensional manifold *without boundary* and suppose that A and $\widetilde{A}$ are connections in a G-bundle over M that are sent to each other by a gauge transformation g:

$$\widetilde{A} = gAg^{-1} - dgg^{-1}\,.$$

Then the Chern–Simons actions for them are related as follows:

$$\mathrm{CS}(\widetilde{A}) = \mathrm{CS}(A) + \frac{1}{3}\int_M \mathrm{tr}\left(g^{-1}dg \wedge g^{-1}dg \wedge g^{-1}dg\right)\,.$$

Recall that the 3-form $\frac{1}{24\pi^2}\,\mathrm{tr}(g^{-1}dg)^{\wedge 3}$ is the pullback under the map $g : M \to G$ of an integral closed 3-form η on the compact simply connected simple Lie group G (see Proposition 2.16 in Appendix A.2; cf. Section II.1.3). Thus the integral of this form depends only on topological properties of the map g and can be expressed as

$$\frac{1}{24\pi^2}\int_M \mathrm{tr}(g^{-1}dg)^{\wedge 3} = \int_M g^*\eta\,,$$

which is an integer, since the 3-form η generates $H^3(G, \mathbb{Z})$. The latter implies that the exponential $\exp\left(\frac{i}{4\pi}\,\mathrm{CS}(A)\right)$ is gauge invariant:

$$\frac{i}{4\pi}\,\mathrm{CS}(\widetilde{A}) - \frac{i}{4\pi}\,\mathrm{CS}(A) = 2\pi i \cdot \frac{1}{24\pi^2}\int_M \mathrm{tr}(g^{-1}dg)^{\wedge 3} \in 2\pi i \cdot \mathbb{Z}\,.$$

Remark 3.16 An interesting integer-valued invariant for a homology 3-sphere M was introduced by Casson and is closely related to the gauge-theoretic constructions above [70]. Roughly speaking, the *Casson invariant* $\mathrm{Cas}(M)$ is defined as the algebraic number of the conjugacy classes of irreducible SU(2)-representations of the fundamental group $\pi_1(M)$. In other words, it counts the number of irreducible flat SU(2)-connections on M modulo conjugation. The homology restriction on the threefold M is related to the fact that if $H_1(M) \neq 0$, then the moduli space of flat connections on M might not be zero-dimensional, and in particular, it would not consist of a finite number of points. The reason for restricting to SU(2) is clarified in the following exercise.

Exercise 3.17 Show that the only reducible representation $\rho : \pi_1(M) \to \mathrm{SU}(2)$ is the trivial one. (Hint: Reducible representations of $\pi_1(M)$ in SU(2) are necessarily abelian and hence factor through the homology $H_1(M)$. This homology group is trivial for a homology 3-sphere.)

Now consider a Heegaard splitting of M into two handlebodies $M = M_1 \cup_\Sigma M_2$ glued together along their common boundary, an embedded surface $\Sigma \subset M$. Consider the moduli space $\mathcal{M}^\Sigma$ of flat connections in the trivial

SU(2)-bundle on the surface Σ. Define two submanifolds L_1 and L_2 of the symplectic manifold $\mathcal{M}^\Sigma$ as those (equivalence classes of) flat connections on the surface Σ that extend to M_1 and M_2 respectively. One can show that these submanifolds are Lagrangian. Their intersection points $L_1 \cap L_2$ correspond to flat connections extendable to the whole of M. Thus the Casson invariant is defined as the intersection number of these submanifolds,

$$\mathrm{Cas}(M) = \#(L_1, L_2)\,,$$

where we assume that the submanifolds intersect transversally, and exclude the intersection corresponding to the trivial representation; see details, for example, in [364].

3.3 The Holomorphic Chern–Simons Action Functional

A complex three-dimensional manifold X equipped with a nowhere vanishing meromorphic 3-form μ can be regarded as a complex analogue of a real oriented manifold with boundary, following the general philosophy that we adopted in Sections 2.2 and 2.3. Accordingly, one can complexify the Lagrangian formalism to this situation. Here we define a holomorphic analogue of the Chern–Simons action functional for (X, μ) and relate it to Mukai's holomorphic symplectic structures on moduli of holomorphic bundles over complex surfaces, following [195, 85].

Let $G_{\mathbb{C}}$ be a complex simple and simply connected Lie group and $E_{\mathbb{C}}$ a complex $G_{\mathbb{C}}$-bundle over the manifold X. As before, let us denote by $\mathcal{A}^X_{\mathbb{C}}$ the space of $(0,1)$-connections in the bundle $E_{\mathbb{C}}$.

Definition 3.18 The *holomorphic Chern–Simons action functional* $\mathrm{CS}_{\mathbb{C}} : \mathcal{A}^X_{\mathbb{C}} \to \mathbb{C}$ is defined via

$$\mathrm{CS}_{\mathbb{C}}(A) := \int_X \mu \wedge \left(\langle A \wedge \bar{\partial} A \rangle + \frac{2}{3} \langle A \wedge A \wedge A \rangle \right)$$

for any $(0,1)$-connection $A \in \mathcal{A}^X_{\mathbb{C}}$ thought of as a $\mathfrak{g}_{\mathbb{C}}$-valued $(0,1)$-form on X. As usual, we assume that the 3-form μ has only first-order poles, and hence the integral above is well defined.

Proposition 3.19 *The extremals of the holomorphic Chern–Simons functional are holomorphic structures in the complex bundle $E_{\mathbb{C}}$.*

Proof. Indeed, in the same way as in the real case and by using the Cauchy–Stokes formula we come to the Euler–Lagrange equation

$$\bar{\partial} A + A \wedge A = 0$$

in the holomorphic setting. Its solutions are $(0,1)$-connections A with vanishing $(0,2)$-curvature, $F^{0,2}(A) = 0$, and each such connection defines the corresponding holomorphic structure in the complex bundle $E_{\mathbb{C}}$. □

Consider now the "boundary term" of the variation $\delta\,\mathrm{CS}_{\mathbb{C}}$, which now descends to the polar divisor of the meromorphic 3-form μ. Denote this polar divisor by $Y := \mathrm{div}_{\infty}\mu \subset X$. Note that the residue $\nu := \mathrm{res}_Y\,\mu$ is a nonvanishing 2-form on the divisor Y, since μ itself is nonvanishing (see Exercise 2.19). In particular, the canonical bundle of Y has to be trivial, so that Y is either a $K3$ surface or a complex torus.

To define the presymplectic structure in the real case we considered a cylinder $M = I \times \Sigma$ over a Riemann surface Σ. Here we look at the complex analogue of such a cylinder. Namely, let $X = \mathbb{CP}^1 \times Y$ be the product of $\mathbb{CP}^1$ and a $K3$ surface or abelian surface Y. Suppose that Y is endowed with a holomorphic (necessarily nonvanishing) 2-form ν, and consider the meromorphic 3-form $\mu = (dz/z) \wedge \nu$ on X, where dz/z is a 1-form on the complex line $\mathbb{CP}^1$. One can see that $\nu = \mathrm{res}_{z=0}\,\mu = -\,\mathrm{res}_{z=\infty}\,\mu$.

Now the variation of the holomorphic Chern–Simons functional satisfies the relation $\delta\,\mathrm{CS}_{\mathbb{C}} = \sigma_{0,\mathbb{C}} + \sigma_{\infty,\mathbb{C}}$ on the space of extremals, which are the integrable $(0,1)$-connections on X, i.e., the connections with vanishing $(0,2)$-curvature. Here $\sigma_{0,\mathbb{C}}$ and $\sigma_{\infty,\mathbb{C}}$ stand for the contributions of the corresponding components $z = 0$ and $z = \infty$ of the polar divisor of μ.

This allows us to introduce the *holomorphic presymplectic* structure $\omega_{\mathbb{C}} = \delta\sigma_{\mathbb{C}}$ on the "boundary values" of the extremals, i.e., on the space of integrable connections on the surface Y. Explicitly, the holomorphic 1-form $\sigma_{\mathbb{C}}$ is

$$\sigma_{\mathbb{C}} := \int_Y \nu \wedge \mathrm{tr}(a \wedge \delta a)\,,$$

where $a := A|_{z=0}$ is the restriction of a $(0,1)$-connection A in $E_{\mathbb{C}}$ from the threefold X to the surface Y (understood as one component $\{z = 0\} \times Y \subset X$ of the polar divisor of μ), δa is the corresponding variation of a, and $\nu = \mathrm{res}_{z=0}\,\mu$ is a holomorphic 2-form on Y.

One can show that, similarly to the real case, the presymplectic structure $\omega_{\mathbb{C}}$ is degenerate along the orbits of the action of the complex group of gauge transformations $G^Y_{\mathbb{C}}$ on integrable $(0,1)$-connections (i.e., holomorphic structures) in the bundle $E_{\mathbb{C}}$ over Y. After taking the quotient with respect to the group action, we obtain a nondegenerate holomorphic symplectic structure on the moduli space of (stable) holomorphic bundles on the $K3$ or abelian surface Y. (Here, as usual, we are concerned with the moduli space only locally around a smooth point.) Thus the holomorphic Lagrangian formalism gives an alternative approach to Mukai's result discussed before:

Theorem 3.20 ([283]) *There exists a holomorphic symplectic structure $\omega_{\mathbb{C}}$ on the moduli space $\mathcal{M}_Y$ of stable holomorphic $G_{\mathbb{C}}$-bundles over a $K3$ or abelian surface Y.*

Remark 3.21 It turns out that there exists a holomorphic analogue of the Casson invariant for a Calabi–Yau manifold X; see [85, 366]. Instead of a Heegaard splitting of a real manifold, one considers a degeneration of this CY manifold to an intersection of two Fano manifolds. The divisor of intersection is a $K3$ or abelian surface, and one counts in a special way the holomorphic bundles over Y extendable to both of these two Fano manifolds.

We also note that the holomorphic Chern–Simons action functional has more complicated transformation properties with respect to gauge transformations. After a "large" gauge transformation, the value of the functional differs by a multiple of the integrals $\int_X \mu \wedge g^*\eta$. The latter can be viewed as the integrals of the meromorphic 3-form μ over the three-cycles in X that are Poincaré dual to the 3-form $g^*\eta$ for various maps $g : X \to G_{\mathbb{C}}$. The values of these integrals can form a lattice or even a dense set in $\mathbb{C}$; hence considering the exponential similar to $\exp\left(\frac{i}{4\pi} CS(A)\right)$ does not allow one to extract a gauge-invariant quantity in the holomorphic setting.

3.4 A Reminder on Linking Numbers

Let M be a simply connected oriented manifold and let γ_1 and γ_2 be two nonintersecting oriented closed curves in M. Pick an oriented surface $D_1 \subset M$ (a Seifert surface for the curve γ_1) such that the curve γ_1 is the oriented boundary of the surface D_1 and such that D_1 and γ_2 intersect transversally.

Definition 3.22 The *linking number* $\mathrm{lk}(\gamma_1, \gamma_2)$ of the curves γ_1 and γ_2 is the intersection number of the surface D_1 and the curve γ_2, i.e., the number of intersections of the curve γ_2 with the surface D_1 counted with orientation (see Figure 3.3):

$$\mathrm{lk}(\gamma_1, \gamma_2) = \#(D_1, \gamma_2).$$

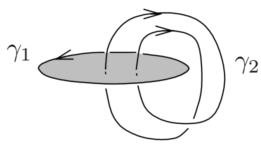


Fig. 3.3. Linking of two oriented curves.

The sign at each intersection point is obtained by forming there a frame from the orientation frames for D_1 and γ_2, and comparing it with the orientation of the ambient manifold M.

Proposition 3.23 *The linking number* $\mathrm{lk}(\gamma_1, \gamma_2)$ *is*

(i) independent of the choice of a Seifert surface D_1*,*

(ii) symmetric in γ_1 *and* γ_2*,*

(iii) invariant with respect to isotopy of the curves, provided they do not intersect each other,

(iv) well defined in any (not necessarily simply connected) oriented three-dimensional manifold M*, provided that both curves* γ_1 *and* γ_2 *are homologous to* 0 *in* M*.*

Note that if the manifold M is not simply connected and only one of the curves is homologous to 0 in M, but the other is not, the linking number might not be well defined. For instance, take $M = \mathbb{T}^3$ and two curves, one of which is homologous to 0, while the other is a generator in $H_1(\mathbb{T}^3, \mathbb{Z})$. Then by taking different Seifert surfaces for the first curve one obtains either 0 or 1 for their linking number; see Figure 3.4.

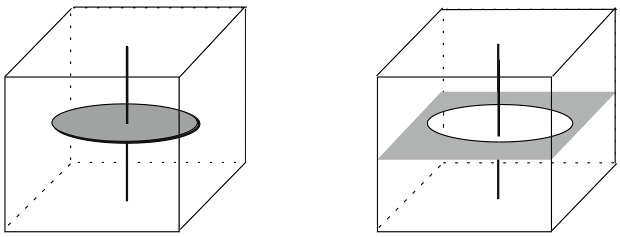


Fig. 3.4. Two Seifert surfaces for the horizontal circle in the cube-torus $\mathbb{T}^3 = \mathbb{R}^3/\mathbb{Z}^3$, one "inside" and one "outside," give linking numbers ± 1 and 0, respectively, for the intersection with the "vertical" cycle.

Exercise 3.24 Prove the above proposition. Furthermore, show also that the linking number is actually invariant when the curve γ_1 changes to a curve (or a collection of curves) $\widetilde{\gamma}_1$ homologous to γ_1 in the complement $M \setminus \gamma_2$ (see Figure 3.5).

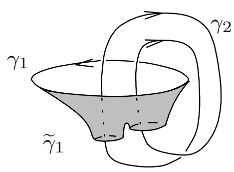


Fig. 3.5. The homologous curves γ_1 and $\widetilde{\gamma}_1$ have the same linking number with the curve γ_2.

Needless to say, the linking number easily generalizes to manifolds of any dimension n, provided that the linking submanifolds are homologous to zero and have "linking dimensions": the sum of their dimensions equals $n-1$.

Remark 3.25 There exists the Gauss integral formula for the linking number of two curves γ_1 and γ_2 in $\mathbb{R}^3$. Recall it here in a somewhat "symbolic form," which we need further.

Let $\Delta \subset M \times M$ denote the diagonal in $M \times M$ and let δ stand for its Poincaré dual current, a closed 3-form supported on the diagonal, $[\delta] \in H^3(M \times M, \mathbb{R})$. Then we can write

$$\mathrm{lk}(\gamma_1, \gamma_2) = \#(\Delta\,, D_1 \times \gamma_2) = \int_{x\in D_1} \int_{y\in\gamma_2} \delta(x,y)\,, \tag{3.6}$$

where $\#(\Delta\,, D_1 \times \gamma_2)$ is the intersection number of Δ and $D_1 \times \gamma_2 \subset M \times M$. One can split the 3-form δ into the homogeneous components

$$\delta = \delta^{3,0} + \delta^{2,1} + \delta^{1,2} + \delta^{0,3}\ ,$$

where $\delta^{i,j}$ denotes the component that is an i-form on the first factor of $M \times M$, and a j-form on the second factor. Note that in equation (3.6) we had to integrate only over the component $\delta^{2,1}$, since all the other integrals vanish. This component is an exact 2-form in x on D_1, which allows us to apply the Stokes formula:

$$\mathrm{lk}(\gamma_1, \gamma_2) = \int_{x\in D_1} \int_{y\in\gamma_2} \delta^{2,1}(x,y) = \int_{x\in\gamma_1} \int_{y\in\gamma_2} d_x^{-1}\delta^{2,1}(x,y)\ .$$

For $\mathbb{R}^3$ the $(1,1)$-form $d_x^{-1}\delta^{2,1}(x,y)$ on the torus $\gamma_1 \times \gamma_2$ assumes the standard *Gauss form*

$$\frac{1}{4\pi} \cdot \frac{(\overrightarrow{x-y}, \overrightarrow{dx}, \overrightarrow{dy})}{\|\overrightarrow{x-y}\|^3}\,,$$

where $(\cdot,\cdot,\cdot)$ is the mixed product of three vectors in $\mathbb{R}^3$.

Remark 3.26 In what follows we need a bit of calculus of such δ-type forms. Let δ_γ be the Dirac δ-type 2-form supported on a closed oriented curve γ in a simply connected threefold M. (Alternatively, the curve γ can be regarded as a de Rham current, a linear functional on 1-forms on M, whose value is the integral of the 1-form over γ.) The integral of this 2-form δ_γ over a two-dimensional surface counts the intersection number of this surface with the curve γ. Then by using the decomposition of the diagonal 3-form δ into the homogeneous components, we can express

$$\delta_\gamma(x) = \int_{y\in\gamma} \delta^{2,1}(x,y)\,,$$

where we denote the coordinates on the first and the second factors of $M \times M$ by x and y respectively. Choose a surface $D \subset M$ whose boundary is $\gamma = \partial D$. Similarly, we can define the δ-type 1-form supported on the surface D by

$$\delta_D(x) = \int_{y \in D} \delta^{1,2}(x, y) \,.$$

The relation $\partial D = \gamma$ is equivalent to the relation between the corresponding δ-forms: $d_x \delta_D(x) = \delta_\gamma(x)$, due to the Stokes theorem, or more explicitly,

$$\delta_\gamma(x) = \int_{y \in D} d_x(\delta^{1,2})(x, y) \ ,$$

where d_x denotes the exterior derivative applied to the x-coordinates only. Finally, if γ_1 and γ_2 are two nonintersecting curves, we have

$$\mathrm{lk}(\gamma_1, \gamma_2) = \int\limits_{x \in D_1} \delta_{\gamma_2}(x) = \int\limits_M \delta_{D_1}(x) \wedge \delta_{\gamma_2}(x) = \int\limits_M \delta_{D_1} \wedge d\delta_{D_2} \,, \tag{3.7}$$

where $\partial D_2 = \gamma_2$. The latter form suggests a common nature of the linking number and the $A \wedge dA$-part of the Chern–Simons functional, which we are going to study below.

3.5 The Abelian Chern–Simons Path Integral and Linking Numbers

We start with a reminder on finite-dimensional Gaussian integrals. Let (x, Qx) be a symmetric negative-definite form in the Euclidean $\mathbb{R}^n$. The classical Gauss integral

$$\int_{\mathbb{R}} \exp(-qx^2/2) dx = \sqrt{2\pi/q}$$

has the multidimensional analogue

$$\int_{\mathbb{R}^n} e^{\frac{1}{2}(x,Qx)} \, d^n x = \left(\frac{(2\pi)^n}{\det(-Q)} \right)^{\frac{1}{2}} .$$

Now fix a vector $J \in \mathbb{R}^n$ and consider the integral

$$Z_Q(J) := \int_{\mathbb{R}^n} e^{\frac{1}{2}(x,Qx)+(x,J)} \, d^n x = \int_{\mathbb{R}^n} e^{S_J(x)} \, d^n x$$

corresponding to the shift $S_J(x) := \frac{1}{2}(x, Qx) + (x, J)$ of the quadratic form by a linear term. (The initial integral is $Z_Q(0)$.) This integral can easily be solved by completing the square. Indeed, let x_0 be a solution of the equation $Qx_0 + J = 0$, i.e., $x_0 = -Q^{-1}J$. Then by introducing a shifted variable $\widetilde{x} = x - x_0$ and using the translation invariance of the measure $d^n x$, we obtain

$$
\begin{aligned}
Z_Q(J) &= \int_{\mathbb{R}^n} e^{S_J(\widetilde{x}+x_0)}\, d^n x \\
&= \int_{\mathbb{R}^n} \exp\left\{\frac{1}{2}(\widetilde{x}+x_0, Q(\widetilde{x}+x_0)) + (\widetilde{x}+x_0, J)\right\} d^n x \\
&= \int_{\mathbb{R}^n} \exp\left\{\frac{1}{2}(\widetilde{x}, Q\widetilde{x}) + \frac{1}{2}(x_0, Qx_0) + (x_0, J)\right\} d^n \widetilde{x} \\
&= e^{S_J(x_0)} \int_{\mathbb{R}^n} e^{\frac{1}{2}(\widetilde{x}, Q\widetilde{x})}\, d^n\widetilde{x} = e^{\frac{1}{2}(x_0,J)}\, Z_Q(0)\,.
\end{aligned}
$$

Thus, we have

$$
\frac{Z_Q(J)}{Z_Q(0)} = e^{S_J(x_0)} = e^{\frac{1}{2}(x_0,J)} = e^{-\frac{1}{2}(Q^{-1}J,J)}\,. \tag{3.8}
$$

Remark 3.27 When the space $\mathbb{R}^n$ is replaced by some infinite-dimensional vector space, the integrals defining $Z_Q(0)$ and $Z_Q(J)$ usually do not make sense. However, one can "calculate" their ratio, which often turns out to be well defined. Note that the second of the equivalent expressions for the ratio $Z_Q(J)/Z_Q(0)$ in formula (3.8) has the form $\exp(\frac{1}{2}(x_0,J)) = \exp(-\frac{1}{2}(x_0,Qx_0))$, which allows us to avoid looking for the inverse Q^{-1} of the corresponding operator in the infinite-dimensional space.

Consider an application of this idea to the abelian Chern–Simons path integral. Let $\mathcal{A}$ be the space of connections in a U(1)-bundle over a real three-dimensional simply connected manifold M without boundary. We can think of such connections as real-valued 1-forms on M. Denote by $\mathrm{CS} : \mathcal{A} \to \mathbb{R}$ the Chern–Simons action functional on $\mathcal{A} = \Omega^1(M,\mathbb{R})$, which now becomes a quadratic form

$$
\mathrm{CS}(A) = \int_M A \wedge dA\,,
$$

since the group U(1) is abelian and the cubic term $A \wedge A \wedge A$ vanishes. Note that the kernel of this quadratic form is the space of exact 1-forms $d\Omega^0 \subset \Omega^1(M,\mathbb{R})$.

Fix some linear functional J on $\Omega^1(M,\mathbb{R})$, i.e., a de Rham current on this space, and define

$$
S_J(A) := \frac{1}{2}\int_M A \wedge dA + \int_M A \wedge J\ .
$$

for $A \in \Omega^1(M,\mathbb{R})$. We also impose the condition $dJ = 0$, so that the linear term $\int_M A \wedge J$ is well defined on the quotient $\Omega^1(M)/d\Omega^0(M)$. Now make the following "formal" definition.

Definition 3.28 The *abelian Chern–Simons path integral* is the expression

$$Z_{\mathrm{CS}}(J) := \int_{\Omega^1/d\Omega^0} e^{S_J(A)}\, DA\,,$$

where DA stands for a translation-invariant measure on the infinite-dimensional space $\Omega^1(M)/d\Omega^0(M)$.

Rather than trying to define the measure and the path integral precisely, we are going to see what the above formal manipulations with Gaussian integrals give us in this situation, where, in a sense, the operator Q is replaced by the outer derivative d. By formula (3.8) for the ratio $Z_Q(J)/Z_Q(0)$ we obtain

$$\frac{Z_{\mathrm{CS}}(J)}{Z_{\mathrm{CS}}(0)} = e^{S_J(A_0)} = e^{\frac{1}{2}\int_M A_0\wedge J}\,,$$

where A_0 is a solution of the equation $dA_0 + J = 0$. (Recall that J is a closed current on a simply connected M, and hence it is exact, i.e., this equation formally has a solution.)

Now we would like to specify the functional J on 1-forms $A \in \Omega^1(M, \mathbb{R})$ to be the integral of the form over a collection of curves in the simply connected manifold M. Let γ_i, $i = 1, \dots, k$, be closed oriented nonintersecting curves in the manifold M. We set $J = \sum_i q_i\, \delta_{\gamma_i}$, where δ_{γ_i} is the δ-type 2-form on M supported on the curve γ_i, while q_i are real parameters. By applying the calculus of δ-forms (see Remark 3.26) we obtain that the ratio $Z_{\mathrm{CS}}(J)/Z_{\mathrm{CS}}(0)$ assumes the following explicit form:

$$\begin{aligned}
\frac{Z_{\mathrm{CS}}(J)}{Z_{\mathrm{CS}}(0)} &= \exp\left\{\frac{1}{2}\int_M A_0\wedge J\right\} = \exp\left\{\frac{1}{2}\int_M A_0\wedge \sum_i q_i \int_{y\in\gamma_i}\delta^{2,1}(x,y)\right\}\\
&= \exp\left\{\frac{1}{2}\int_M A_0\wedge \sum_i q_i \int_{y\in D_i} d_x\delta^{1,2}(x,y)\right\}\\
&= \exp\left\{\frac{1}{2}\int_M -dA_0\wedge \sum_i q_i \int_{y\in D_i} \delta^{1,2}(x,y)\right\}\\
&= \exp\left\{\frac{1}{2}\int_M \left(\sum_j q_j \int_{z\in\gamma_j}\delta^{2,1}(x,z)\right)\wedge\left(\sum_i q_i\int_{y\in D_i}\delta^{1,2}(x,y)\right)\right\}\\
&= \exp\left\{\frac{1}{2}\sum_{i,j} q_iq_j\int_M \delta_{\gamma_j}(x)\wedge\delta_{D_i}(x)\right\} = \exp\left\{\frac{1}{2}\sum_{i,j} q_iq_j\; lk(\gamma_j,\gamma_i)\right\}
\end{aligned}$$

Here we have used the Stokes theorem, as well as the definition of A_0 as a solution of the equation $dA_0 + J = 0$.

Corollary 3.29 ([340, 318]) *For the functional J defined as the integral of 1-forms over a collection of curves in a threefold, the ratio $Z_{\mathrm{CS}}(J)/Z_{\mathrm{CS}}(0)$ counts the pairwise linking numbers of these curves.*

Note also that above, in the latter sum, we had to assume that $i \neq j$, so that the linking number was defined. The case of self-linking is much more subtle. It leads to divergences of the path integral and requires some additional specifications, such as framing, for its normalization; see [54]. The value $Z_{\mathrm{CS}}(0)$ in the case without any curve corresponds to the Ray–Singer torsion of the manifold M [340].

The topological Chern–Simons path integral has a holomorphic analogue.

Definition 3.30 (cf. [390]) For a three-dimensional Calabi–Yau manifold X with a holomorphic 3-form μ the *holomorphic abelian Chern–Simons path integral* is the expression

$$Z_{\mathbb{C}S}(J) := \int_{\Omega^{0,1}/\bar\partial\Omega^{0,0}} e^{S_{\mathbb{C}J}(A)}\, DA\,,$$

where

$$S_{\mathbb{C}J}(A) := \frac{1}{2}\int_X \mu \wedge A \wedge \bar\partial A + \langle \mathbb{C}J, A\rangle$$

is the quadratic form shifted by the linear functional $\mathbb{C}J$ on the space of $(0,1)$-connections $A \in \Omega^{0,1}(X,\mathbb{C})$.

Remark 3.31 For a complex curve $C \subset X$ equipped with a holomorphic 1-form α define the linear functional on $(0,1)$-connections A by assigning $\langle \mathbb{C}J_C, A\rangle := \int_C \alpha \wedge A$. Similarly to the topological case, if such a functional $\mathbb{C}J$ corresponds to a collection of complex curves, the holomorphic abelian Chern–Simons path integral can be described in terms of the *polar linking number*, a holomorphic analogue of the Gauss linking number, which we define in Section 4.3. The relation of this functional with the holomorphic analogue of linking was established in [134, 195, 366].

The abelian theory is a particular case of the general Chern–Simons path integral. In the topological case we consider a link $L = \cup_i \gamma_i$ in a compact real threefold M. Let $\mathcal{A}$ be the affine space of all connections in the (trivial) G-bundle over M for a compact simply connected simple Lie group G. We identify $\mathcal{A}$ with the space $\Omega^1(M,\mathfrak{g})$ of 1-forms on M with values in the Lie algebra $\mathfrak{g}$ of G. Finally, let $G^M = C^\infty(M,G)$ be the group of gauge transformations in the bundle.

Definition 3.32 The *nonabelian Chern–Simons path integral* for a link $L \subset M$ is the following function of a parameter k:

$$Z_{\mathrm{CS}}(L;k) = \int\limits_{\mathcal{A}/G^M} \Big\{ \exp\Big\{ ik \int_M \mathrm{tr}\Big(A \wedge dA + \frac{2}{3} A \wedge A \wedge A \Big) \Big\} \times \prod_{\gamma_i \subset L} \mathrm{tr}\Big(P \exp \int_{\gamma_i} A \Big) \Big\} DA,$$

where $P\exp$ is the path-ordered exponential integral of a nonabelian connection A over γ_i, and DA is an appropriate measure on the moduli space of the connections $\mathcal{A}/G^M$.

Remark 3.33 Witten showed in [389] that for $M = S^3$ and $G = \mathrm{SU}(2)$ this path integral leads to the Jones polynomial for the link L. Other link or knot invariants can be obtained by changing the group. Note that they are always Vassiliev-type invariants of finite order [35, 36]. There are various ways to give $Z_{\mathrm{CS}}(L; k)$ and the corresponding link invariants rigorous definitions (see, e.g., the combinatorial [327] or probabilistic [5] approaches).

The extension of these results to a holomorphic version of the nonabelian Chern–Simons path integral is an intriguing open problem. The more complicated gauge transformation property of the holomorphic Chern–Simons action functional already makes the first step, writing out the corresponding path integral for an arbitrary collection of complex curves in a Calabi–Yau threefold, a serious problem; see some discussion in [391, 134, 366].

3.6 Bibliographical Notes

The Chern–Simons functional was introduced in [72]. For the relation of the abelian Chern–Simons functional to linking numbers we refer to [340, 318]. The appearance of the Jones polynomial and other knot invariants from the Chern–Simons functional was discovered by Witten [389]; see more details in [35, 210]. An excellent account of the relation between this functional to knot theory is contained in the book by Atiyah [27]. The relation between the Chern–Simons functional and the Vassiliev knot invariants is described in [36, 210].

The holomorphic Chern–Simons functional was introduced in [390] and studied in a number of papers [85, 134, 195, 196, 367]. For a higher-dimensional version of the Chern–Simons functional and its relation to linking numbers of several submanifolds see [124].

The classical Lagrangian formalism can be found, for example, in [18]. The formalism of the Lagrangian field theory was described in [393]; see also the presentations in [79, 341] for more details and examples. For preliminaries on linking numbers one can look at any book on differential topology, e.g., [162]. The question of when the space of extremals (more precisely, geodesics on a manifold) is a smooth manifold by itself is addressed, for example, in [38, 39].

http://www.springer.com/978-3-540-77262-0

The Geometry of Infinite-Dimensional Groups
Khesin, B.; Wendt, R.
2009, XII, 304 p., Hardcover
ISBN: 978-3-540-77262-0