Hamiltonian Systems and their Integrability

Michèle Audin

Institut de Recherche Mathématique Avancée, Université Louis Pasteur et CNRS, 7 rue René Descartes, 67084 Strasbourg cedex, France

E-mail address: `Michele.Audin@math.u-strasbg.fr`

URL: `http://www-irma.u-strasbg.fr/~maudin`

Studious readers are reading in the libraries. Teachers are giving their lessons. Students are taking notes. Accountants are lining up columns of figures. Apprentice pastry cooks are stuffing cream into rows of cream puffs. Pianists are playing their scales. Sitting deep in thought at their tables, writers are forming lines of words.

Georges Perec [**68**]

I dedicate this book to the memory of Nicole Desolneux-Moulis.

Translated by: ANNA PIERREHUMBERT

Contents

Appendices

Introduction

This book is an introduction to some topological, analytical, and algebro-geometric aspects of the theory of completely integrable Hamiltonian systems.

Hamiltonian systems are differential equations that describe the movements of an object whose energy is conserved (conservative mechanics). Certain of these systems can be integrated "by quadratures", these are the completely integrable Hamiltonian systems. One example of such a system, which all children know well, is that of the motion of a top. The motion of a free particle on a surface of revolution or an ellipsoid are other simple examples of completely integrable Hamiltonian systems.

What Is in This Book. The theory of integrable systems finds methods, followers, and applications in most of the fields of mathematics. It has as a result been the subject of an immense amount of work. No book—even one ten times thicker than this booklet—could address all these aspects.

Integrable systems belong to the field of (conservative) Hamiltonian mechanics, which today means the field of symplectic geometry. They have many conserved quantities traditionally called "first integrals". It is this definition which Chapter I addresses.

The heart of the material presented here is the Arnold–Liouville theorem, which describes the properties of solutions for which the values of these conserved quantities are generic. There are certain coordinates in which these systems take a very simple form. We have said that these systems are "integrable by quadratures". Geometrically, this means that their trajectories lie on tori and, furthermore, are *linear* on these tori (Figure 1). This is the Arnold–Liouville theorem and Chapter II is devoted to its proof and to some examples.

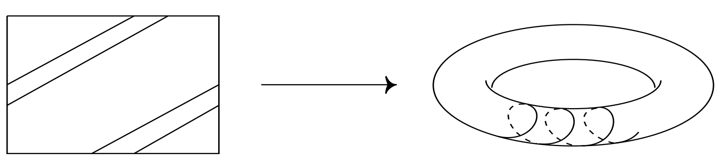


Figure 1

Next the material branches into two independent directions:

- In Chapter III, we turn to the algebraic theory of differential equations. The Arnold–Liouville theorem asserts that the solutions near a solution that is general enough have the same nature as their neighbor (again, see Figure 1). It is then tempting to look at what happens in general for the solutions that are (infinitesimally) near a trajectory of a Hamiltonian vector field. The word "infinitesimally" indicates that the differential equation has been linearized. With suitable hypotheses of analyticity, we can then make use of the Picard–Vessiot theory (differential Galois theory). The most interesting result in this area is a theorem of Morales and Ramis [**62**], which asserts that the differential Galois group of the linearized equation cannot be too complicated when the system is integrable (its identity component must be an Abelian group). This theorem can be used to show that certain Hamiltonian systems are *not* completely integrable.

- In Chapter IV, we turn to algebraic geometry, which in some cases is able to give us a more precise, even effective, version of the Arnold–Liouville theorem. When the differential system can be written as a "Lax equation", there is associated with it an algebraic curve (or family of algebraic curves) which will lead to a (family of) Jacobian(s), which are tori ready to play the role of Liouville tori. In particular, we will see how this machinery gives us tools to the answer to concrete, naive questions such as determining whether a value of the first integrals is generic.

Finally, two appendices summarize some useful definitions and properties from differential Galois theory and the theory of algebraic curves, respectively.

While the methods of algebraic geometry have been the object of a great many works, including many books, the same cannot be said for the "Galois-theoretic" methods. In writing this book, one of my goals was to understand and popularize the beautiful Morales–Ramis theorem by presenting it to a less specialized audience than the original texts were written for. Although the criterion for non-integrability it provides belongs to the world of differential Galois theory, it certainly still is about objects belonging to the world of geometry!

This is why I tried

- to use the most classical and especially the most elementary examples possible. The main example used here to illustrate the Galois methods (a special case of the Hénon–Heiles system leading to the Airy equation) is, probably, the simplest possible one although it does not appear to be in any of the original articles;
- to clarify the situation a bit by separating
  - what is drawn from the general theory of differential equations (the variational equation, initial forms, first integrals, Galois groups)
  - from what is actually drawn from symplectic geometry.

  Similarly, I have tried to give the essential ingredients of the proofs in order to get straight to the desired goal;
- to maintain an acceptable level; in fact, these notes were developed from a graduate course (cours de DEA) given to real students (who contributed much to the simplicity of this book).

I am a firm believer in the virtue of examples, and so there are numerous examples (classical ones, as I have said, but they are the best) in Chapters I and II. While the mathematics used in Chapter III are no more difficult than in the first two chapters, the *applications* of the Morales–Ramis theorem are more elaborate. Therefore, unfortunately, Chapter III contains mainly academic examples. Chapter IV, on the other hand, is almost solely composed of examples—the main one being that of geodesics on quadrics.

The Exercises. Each chapter ends with a section of exercises. Most of these are direct applications of the material presented in that chapter. Against tradition, none of the important proofs have been left as an exercise.

The Index. I made a particular effort to make the index as complete and redundant as possible, both essential properties for its usefulness.

Notation. The notation is as standard as possible. There is no reason to let $\mathbb{P}^n(\mathbb{C})$ be the stabilizer of an element x under the action of a group G, or to let G_x be a symplectic form, or to let ω to be complex projective space, so I have simply conformed to tradition. I will use *slanted* characters to designate a notion that I am in the process of defining. I will use *italics* to emphasize something.

What Is Not in This Book. As I have said, there are many things, including

- the use of algebro-geometric machinery for counting Liouville tori, which would only be a repetition of [**8**];
- examples of infinite-dimensional integrable systems: no Korteweg–de Vries equation, no inverse scattering, all of which one can be found in Moser's little book [**66**] and Segal's article in [**42**];
- questions of perturbation theory, also found in [**66**].

Bibliographic Note. I will utilize the basics of differential calculus on manifolds (vector fields, differential forms, Lie derivative) which the readers can find, for example, in [**54**] or in the first volume of [**75**]. I will also use the definition of Lie groups and Lie algebras (and not much more)[1] and the basic concepts of group actions on manifolds (which the readers can find in [**21**] and also [**15**]). A part from this, the book is "self-contained" in the sense that all the definitions of objects utilized in the text and all of the proofs, except for

- the uniqueness of Picard–Vessiot extensions and the fact that the Galois group is algebraic (see [**58**]);
- the triviality of holomorphic vector bundles on open Riemann surfaces, which I use very marginally and which can be found in [**29**];
- some results for Riemann surfaces (completion, Riemann–Roch,...), used in Chapter IV and listed explicitly in Appendix B, which can be found in [**72**].

In 1999, I gave three courses on integrable systems, where I presented some material contained in this book:

[1] Again, see [**54**].

- in Lisbon in June, where I particularly addressed algebro-geometric methods (Chapter IV);
- at the Centre International de Recherches en Mathématiques (Marseilles) in July, where I briefly discussed Galois-theoretic methods (Chapter III in this book), but not the algebro-geometric part, at the summer school "Symétries et application moment", organized by Patrick Iglesias and Elisa Prato;
- in Strasbourg in the fall, for a graduate course (Cours de DEA) where I presented all the material found in this book, except Chapter IV and, in particular, the material on Galois methods in a fair amount of detail.

This book is a development of the notes I wrote for the "students" of all these courses.

Acknowledgements. I would like to thank all the participants of the various courses I gave, and especially the warm audience of the Lisbon course. The tenacity and fidelity of those subjected to me twice a day, every day for two weeks, touched me greatly.

A special thanks goes to Ana Cannas da Silva, who not only organized the course in Lisbon[2] but also actively followed the course in Marseilles, providing remarks, questions, and numerous corrections and improvements.

The material on Galois methods was put together during the working group on the Morales–Ramis theorem that Claudine Mitschi and myself organized in Strasbourg during the spring of 1999. Most of what I understand about differential Galois theory comes from her presentations and explanations. She is also owed thanks for reading a draft of part of this book and suggestion several improvements.

I would also like to thank Leonor Godinho, Christelle Holtzmann, Marie-Laure Kostyra, and Julien Stiker for reading all or parts of various preliminary versions of this text and for the comments and corrections they provided.

A good-natured anonymous referee expressed a few wishes that I was happy to fulfill. I would like to thank him for the resulting improvements.

Finally, I thank Raymond Seroul, who designed the most beautiful figures in this book, Donald Babbitt, who carefully edited the English tranlation, and Claude Sabbah, who patiently edited twice the LaTeX files—It goes without saying that I, the author, am the only responsible for the remaining errors.

[2]Um mês de junho muito agradável em Lisboa... É um grande prazer agradecer o convite.

CHAPTER I

Introduction to Integrable Systems

Integrable systems are special examples of differential systems defined by "Hamiltonian" vector fields[1] on a symplectic manifold.

In this chapter, I briefly review some useful concepts from symplectic geometry and give some examples of integrable systems. I refer readers to their preferred books on symplectic geometry (e.g., [**56, 15**] and others) and especially [**4**].

I.1. Symplectic Manifolds

I.1.a. Symplectic Vector Spaces. A symplectic vector space is a real vector space equipped with a *non-degenerate* skew-symmetric bilinear form. The prototypical (and, as we will see, essentially unique) example of a symplectic vector space is $\mathbb{R}^n \times \mathbb{R}^n$ equipped with the form

$$\omega((q,p),(q',p')) = p \cdot q' - p' \cdot q,$$

where, as usual, the dot $\cdot$ denotes the Euclidean dot product in $\mathbb{R}^n$.

In fact, as the following proposition asserts, this is the only example of a symplectic vector space.

Proposition I.1.1. *Let ω be a nondegenerate, skew-symmetric bilinear form on a finite-dimensional vector space E. Then there is a basis $(e_1, \dots, e_n, f_1, \dots, f_n)$ (called a symplectic basis) such that $\omega(e_i, f_j) = \delta_{i,j}$ and $\omega(e_i, e_j) = \omega(f_i, f_j) = 0$.*

Proof. Since ω is nondegenerate, it is nonzero, and we can find two vectors e_1 and f_1 such that $\omega(e_1, f_1) = 1$. One verifies that ω, restricted to the (symplectic) orthogonal complement (with respect to ω) of the plane $\langle e_1, f_1 \rangle$, is a *non-degenerate* skew-symmetric form. Using the fact that a skew-symmetric bilinear form on a one-dimensional space is zero everywhere, induction on the dimension yields the desired result. □

In particular, the dimension of E is even and is the only isomorphism invariant of (E, ω). In a symplectic basis, the matrix of the bilinear form ω is

$$J = \begin{pmatrix} 0 & \mathrm{Id} \\ -\mathrm{Id} & 0 \end{pmatrix}.$$

[1]A vector field on a manifold *is* a differential equation, the differential equation whose solutions are its integral curves. See [**3**] and Footnote 1 in Appendix A.

EXAMPLE I.1.2. If $(e_1, \dots, e_n)$ is an orthonormal basis for $\mathbb{R}^n$, then

$$((0, e_1), \dots, (0, e_n), (e_1, 0), \dots, (e_n, 0))$$

is a symplectic basis for $\mathbb{R}^n \times \mathbb{R}^n$.

Exercises I.1, I.2.

The Symplectic Group. This is the group of isometries of ω i.e., the linear isomorphisms g of E such that

$$\omega(gX, gY) = \omega(X, Y) \quad \text{for all } X, Y \in E.$$

In a symplectic basis, where $\omega(X,Y)$ is written as ${}^tXJY$, the symplectic group consists of matrices g such that

$${}^tgJg = J.$$

This is an algebraic subgroup of $\mathrm{GL}(2n;\mathbb{R})$, since the equations defining it are polynomial equations in the entries of g; we denote this group by $\mathrm{Sp}(2n;\mathbb{R})$. Its Lie algebra, the tangent space at the identity, consists of matrices satisfying

$${}^tAJ + JA = 0,$$

a linear equation obtained by differentiating ${}^tgJg - J = 0$ at the identity.
Exercises I.5, I.25.

Isotropic and Lagrangian Subspaces. We say that a subspace F of a symplectic space E is *isotropic* if it is contained in its symplectic orthogonal complement ($F \subset F^\circ$) of F. We say that F° is *co-isotropic*.

Since the form ω is nondegenerate, the equality

$$\dim F + \dim F^\circ = \dim E$$

holds for any subspace F. If we let $\dim E = 2n$, then $\dim F \leqslant n$ for any isotropic subspace F. There are in fact isotropic subspaces of dimension exactly n, for example the subspace generated by the "half" $(e_1, \dots, e_n)$ of the symplectic basis. These subspaces are called *Lagrangian subspaces*.
Exercises I.3, I.4, I.6.

I.1.b. Definition of Symplectic Manifold. We now try to understand what the definition a symplectic *manifold* W ought to be. First of all, the tangent space T_xW at each x should be equipped with a skew-symmetric bilinear form ω_x such that all of these forms together define a a differential 2-form ω . We also want each ω_x to be non-degenerate. If the (necessarily even) dimension of W is $2n$, we can express this as[2]

$$\wedge^n \omega_x \neq 0 \in \bigwedge^{2n} T_x^\star W$$

for all x. In other words, the $2n$-form $\omega^{\wedge n}$ must be a volume form on W. In particular, there must be a volume form defined on W which implies that W must be an orientable manifold.

But this definition does not yet suffice. We also want the calculus arising from the form ω on W to be locally the same as the usual one on $\mathbb{R}^n \times \mathbb{R}^n$ arising from

[2]See Exercise I.2 if necessary.

the "constant" form (as described at the beginning of Section I.1.a). We write this form as a differential form as follows: let $X = (q, p)$ and $X' = (q', p')$ be vectors in $\mathbb{R}^n \times \mathbb{R}^n$, then we have

$$\omega(X, X') = p \cdot q' - p' \cdot q.$$

In other words,

$$\omega = \sum_{i=1}^{n} dp_i \wedge dq_i = d\Big(\sum_{i=1}^{n} p_i dq_i\Big),$$

and, in particular, ω is an *exact* form. We could require a symplectic form to be exact. Unfortunately, however, this would rule out compact manifolds as symplectic manifolds. This is because of the following result.

PROPOSITION I.1.3. *No 2-form defined on a compact manifold W is both nondegenerate and exact.*

PROOF. We have noted that a form ω is nondegenerate if and only if $\omega^{\wedge n}$ is a volume form. Since W is compact, it follows from de Rham cohomology that no volume form on W is exact[3]. Now, if ω were exact, then we would have $\omega = d\alpha$, but then the volume form $\omega^{\wedge n}$ would be exact as well: $\omega^{\wedge n} = d(\alpha \wedge \omega^{\wedge(n-1)})$. $\square$

In fact, if we want to compute *locally* as in $\mathbb{R}^n \times \mathbb{R}^n$, then what we need is a condition of local exactness, that is: ω is *closed*. This leads us to the good definition[4].

DEFINITION I.1.4. A *symplectic manifold* is a pair (W, ω) where W is a manifold and ω is a closed, nondegenerate 2-form. Such a form ω is said to be *symplectic*.

Of course, we will often abuse notation by simply calling W a symplectic manifold, without mentioning the form ω when there is no cause for confusion.

I.1.c. Examples of Symplectic Manifolds. The space $\mathbb{R}^n \times \mathbb{R}^n$ with the form $\sum dp_i \wedge dq_i$ is clearly a symplectic manifold.

Cotangent Bundles. Let V be a manifold, and consider the total space of its cotangent bundle with the projection

$$\pi : T^\star V \longrightarrow V.$$

On $T^\star V$, there is a *canonical* differential 1-form α, called the *Liouville form* (see Exercise I.7), defined by

$$\alpha_{(x,\varphi)}(X) = \varphi(T_{(x,\varphi)}\pi(X)),$$

where $x \in V$, $\varphi \in T^\star_x V$, and $X \in T_{(x,\varphi)}(T^\star V)$. It is easy to check that $\omega = d\alpha$ is nondegenerate (and even easier to check that it is closed!). If $(q_1, \dots, q_n)$ are local coordinates on V and $(p_1, \dots, p_n)$ are the "dual" coordinates, then $(p_1, \dots, p_n, q_1, \dots, q_n)$ is a system of local coordinates such that $\alpha = \sum p_i dq_i$ and $\omega = \sum dp_i \wedge dq_i$. We often say that ω is the "canonical symplectic form" on the

[3] For a basic knowledge of de Rham cohomology, see, for example, [**54**] or Volume I, Chapter VII or [**75**].

[4] See also Footnote 20.

cotangent bundle[5]. The example above of $\mathbb{R}^n \times \mathbb{R}^n$ can be considered as the special case where $V = \mathbb{R}^n$.

The Space of Lines. The space $\mathcal{D}_n$ of oriented (affine) lines in $\mathbb{R}^{n+1}$ is a symplectic manifold of dimension $2n$. It can be identified with

$$\left\{(p,u) \in \mathbb{R}^{n+1} \times \mathbb{R}^{n+1} \mid \|p\|^2 = 1 \text{ and } p \cdot u = 0\right\}$$

(see Figure 1), which is symplectic as a submanifold of $\mathbb{R}^{n+1} \times \mathbb{R}^{n+1}$ or as the (co)tangent bundle of the sphere S^n.

Exercises I.7, I.9, I.13.

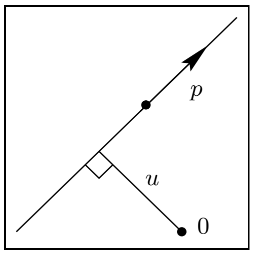


FIGURE 1. The space of lines

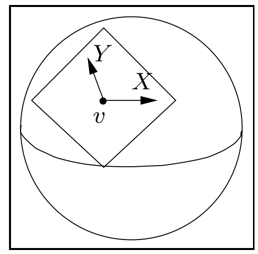


FIGURE 2. The symplectic form on the sphere

Surfaces. Any differential 2-form on a surface W is closed. Moreover, in dimension 2, saying that a 2-form is nondegenerate is equivalent to saying that it is nowhere zero. On a surface, the notion of a symplectic form thus coincides with that of a volume form. Consequently all orientable surfaces can be considered to be symplectic manifolds.

EXAMPLE I.1.5. We consider the unit sphere S^2 in $\mathbb{R}^3$. Its tangent space at a point v is the plane orthogonal to the unit vector v. Set

$$\omega_v(X,Y) = v \cdot (X \times Y)$$

(see Figure 2). This is a nondegenerate (straightforward check), skew-symmetric, bilinear form, and thus a symplectic form.

Exercise I.8.

It can be shown (this is the famous Theorem of Darboux; see Section I.5) that all symplectic manifolds are locally isomorphic to $\mathbb{R}^n \times \mathbb{R}^n$ with the form $\sum dp_i \wedge dq_i$.

Complex Symplectic Manifolds. In Chapter III, we will need *complex* symplectic manifolds. A *complex symplectic* manifold is:

- a complex analytic manifold W (the change of coordinate maps are analytic)

[5]This is another example of an exact symplectic form—on a noncompact manifold, of course.

- equipped with a nondegenerate, closed, complex 2-form ω (on each tangent space, ω is a nondegenerate complex bilinear form) that is *holomorphic*, i.e., we have in complex analytic local coordinates that
$$\omega = \sum f_{i,j} dz_i \wedge dz_j,$$
where the $f_{i,j}$ are holomorphic functions.

I.1.d. Hamiltonian Vector Fields. If $H : W \to \mathbb{R}$ is a function, the symplectic form allows us to associate to H a vector field X_H, which is a sort of gradient called the *Hamiltonian vector field* (or sometimes called the "symplectic gradient"). This is the vector field determined by the relation

$$\omega_x(Y, X_H(x)) = (dH)_x(Y) \quad \text{for all } Y \in T_x W$$

(or, equivalently, by $i_{X_H}\omega = -dH$).

Note that the vector field X_H vanishes at x if and only if x is a critical point of the function H

$$X_H(x) = 0 \Longleftrightarrow (dH)_x = 0.$$

In particular, the singularities (or zeros) of a Hamiltonian vector field are the critical points of a function.

Note also that the function H is constant on the trajectories, or integral curves, of the field X_H: since ω_x is skew-symmetric, we have $(dH)(X_H) = 0$ or $X_H \cdot H = 0$. Exercise I.11.

I.1.e. The Poisson Bracket. Let f and g be two differentiable functions on W. We define their "Poisson bracket" by the formula

$$\{f, g\} = X_f \cdot g = dg(X_f).$$

It is easy to check that

$$\{f, g\} = dg(X_f) = \omega(X_f, X_g) = -\omega(X_g, X_f) = -df(X_g) = -X_g \cdot f = -\{g, f\}.$$

This sequence of equalities shows that the Poisson bracket is skew-symmetric in f and g. Also by definition, it is a derivation (in each of its variables), that is, it satisfies the Leibniz identity

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

The Poisson Bracket and Hamiltonian Vector Fields. From the general formula

$$\mathcal{L}_X i_Y - i_Y \mathcal{L}_X = i_{[X,Y]}$$

and the Cartan formula

$$\mathcal{L}_X = d i_X + i_X d,$$

we see that

$$\begin{aligned} i_{[X_f, X_g]}\omega &= \mathcal{L}_{X_f} i_{X_g}\omega - i_{X_g}\mathcal{L}_{X_f}\omega \\ &= d i_{X_f} i_{X_g}\omega + i_{X_f} d i_{X_g}\omega - i_{X_g} d i_{X_f}\omega - i_{X_g} i_{X_f} d\omega \\ &= d i_{X_f} i_{X_g}\omega = d(\omega(X_g, X_f)) = d\{f, g\}, \end{aligned}$$

that is,

$$[X_f, X_g] = X_{\{f,g\}}.$$

Thus, we also have that

$$[X_f, X_g] \cdot h = \{\{f, g\}, h\}.$$

We will deduce from this that the Poisson bracket also satisfies the Jacobi identity

$$\{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = 0$$

and thus it defines a Lie algebra structure on $\mathcal{C}^\infty(W)$. Moreover the map

$$\begin{aligned} \mathcal{C}^\infty(W) &\longrightarrow \mathcal{X}(W) \\ f &\longmapsto X_f \end{aligned}$$

is a morphism from the Lie algebras $\mathcal{C}^\infty(W)$ (equipped with the Poisson bracket) into the Lie algebra of vector fields (equipped with the Lie bracket of vector fields).

Proof of the Jacobi Identity. We apply the definition of the bracket of vector fields,

$$[X_f, X_g] \cdot h = X_f \cdot (X_g \cdot h) - X_g \cdot (X_f \cdot h),$$

and that of the Poisson bracket together with the equality shown above to obtain

$$\begin{aligned} \{\{f, g\}, h\} &= [X_f, X_g] \cdot h \\ &= X_f \cdot (X_g \cdot h) - X_g \cdot (X_f \cdot h) \\ &= X_f \cdot \{g, h\} - X_g \cdot \{f, h\} \\ &= \{f, \{g, h\}\} - \{g, \{f, h\}\}. \end{aligned}$$

Since the Poisson bracket is skewsymmetric, this is equivalent to the Jacobi identity. □

Remark I.1.6. More generally, a *Poisson bracket* on a manifold V is a Lie algebra structure, denoted by $\{\,,\,\}$, on the vector space $\mathcal{C}^\infty(V)$ of functions on V that satisfies the Leibniz identity. There will be some non-symplectic examples in Section IV.4.a.

Exercises I.10, I.15, I.16, I.17.

I.2. Hamiltonian Systems, Examples

A Hamiltonian system is simply a differential system associated to a Hamiltonian vector field. It is a differential equation

$$\dot{x}(t) = X_H(x(t))$$

for some function H on a symplectic manifold W. For example, if H is a function on $\mathbb{R}^n \times \mathbb{R}^n$, then the associated Hamiltonian system is the differential system[6]

$$\begin{cases} \dot{q} = \dfrac{\partial H}{\partial p} \\ \dot{p} = -\dfrac{\partial H}{\partial q}, \end{cases}$$

where, in abbreviated notation, $q = (q_1, \dots, q_n)$ and $p = (p_1, \dots, p_n)$.

[6] See Exercise I.10 for an expression in terms of coordinates for the field X_H.

Examples of Hamiltonian systems that I present in this book include pendulums, solids (tops), geodesic flows (which get special treatment in Section I.4), and some other academic examples.

I.2.a. The Harmonic Oscillator. A harmonic oscillator is a differential system describing, for example, the motion of a spring (without damping) or an RLC electrical circuit without resistor ($R = 0$). The corresponding differential equation is $\ddot{q} = -aq$, for some real $a > 0$.

In this case, the symplectic manifold is $\mathbb{R} \times \mathbb{R}$, and we assume a to be 1 (see Exercise I.12 for the general case). The Hamiltonian is then

$$H = \frac{1}{2}p^2 + \frac{1}{2}q^2,$$

and the differential system is simply

$$\dot{q} = p, \quad \dot{p} = -q.$$

This is of course equivalent to the differential equation $\ddot{q} = -q$, and its solutions are given by $q = A\cos(t - t_0)$, $p = -A\sin(t - t_0)$. The supports of the trajectories are the level sets of H, i.e., the circles centered at the origin.

We can construct Hamiltonian systems of higher dimension by coupling several harmonic oscillators, that is considering Hamiltonians on $\mathbb{R}^{2n}$ of the form

$$H(q_1, \ldots, q_n, p_1, \ldots, p_n) = \frac{1}{2}\sum p_i^2 + \sum a_i q_i^2.$$

Exercise I.12.

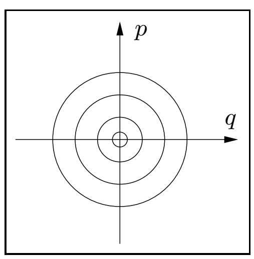


FIGURE 3. The harmonic oscillator

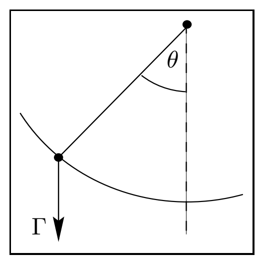


FIGURE 4. The simple pendulum

I.2.b. Rotations on the Sphere. As we saw in Example I.1.5, the unit sphere S^2 in the Euclidean space $\mathbb{R}^3$ can be equipped with a form ω defined by

$$\omega_v(X, Y) = v \cdot (X \times Y).$$

We consider the Hamiltonian $H(x, y, z) = z$. We call e_3 the "third" basis vector, so that $H(v) = v \cdot e_3$. The Hamiltonian vector field is the vector field X tangent to S^2 and satisfying

$$v \cdot (X \times Y) = -(dH)_v(Y) = -Y \cdot e_3$$

for all $Y \in T_v S^2 = v^\perp$. In other words, we have

$$Y \cdot (v \times X - e_3) = 0$$

for all $Y \in v^\perp$. This condition is thus equivalent to the fact that $v \times X - e_3$ is collinear to v, which implies $X_H(v) = v \times e_3$. The trajectories of X_H are the circular intersections of the sphere with horizontal planes.

I.2.c. The Simple Pendulum. The simple pendulum is the mechanical system consisting of a point mass attached to the end of a massless inelastic cord for which the other end is fixed, moving in a constant gravitational field (see Figure 4).

We assume here that the mass moves in a vertical plane. We denote by θ the angle between the cord and the vertical axis (direction of gravity). The differential equation describing the motion of the point mass is then $\ddot{\theta} = -\sin\theta$.

The phase space is $S^1 \times \mathbb{R}$. Using a real variable q whose reduction modulo 2π is θ and the dual variable $p = \dot{q}$, we see that this differential equation is equivalent to the system

$$\dot{q} = p, \qquad \dot{p} = -\sin q.$$

This is the Hamiltonian system associated to the function $H = \frac{1}{2}p^2 - \cos q = \frac{1}{2}p^2 - \cos\theta$, the total energy of the simple pendulum.
Exercise I.18.

We now consider some examples "with two degrees of freedom", i.e., on four-dimensional symplectic manifolds.

I.2.d. The Hénon–Heiles System. The symplectic manifold is $\mathbb{R}^2 \times \mathbb{R}^2$, and the Hamiltonian is the function

$$H = \frac{1}{2}(p_1^2 + p_2^2) - q_2^2(A + q_1) - \frac{\lambda}{3}q_1^3.$$

The vector field X_H is

$$X_H(q_1, q_2, p_1, p_2) = (p_1, p_2, q_2^2 + \lambda q_1^2, 2q_2(A + q_1)).$$

Let E be the (symplectic) subspace defined by $q_2 = p_2 = 0$. If $x = (q_1, 0, p_1, 0)$ is a point in E, we have

$$X_H(x) = (p_1, 0, \lambda q_1^2, 0).$$

Thus the Hamiltonian vector field X_H is *tangent* to the subspace E, and thus there are solutions to the Hamiltonian system which are contained in E. When $\lambda = 0$, these are the lines $(p_1 t + q_1^0, 0, p_1, 0)$. This "academic" example and its solutions will be utilized in Chapter III.
Exercise I.19.

I.2.e. An Anharmonic Oscillator. The symplectic manifold is still $\mathbb{R}^4$ and the Hamiltonian is

$$H = \frac{1}{2}(p_1^2 + p_2^2) + \lambda q_2^2 + (q_1^2 + q_2^2)^2,$$

with λ a nonzero constant. The Hamiltonian vector field is

$$X_H(q_1, q_2, p_1, p_2) = (p_1, p_2, -4(q_1^2 + q_2^2)q_1, -2\lambda q_2 - 4(q_1^2 + q_2^2)q_2).$$

As in Section I.2.d, the vector field X_H is tangent to the subspace E defined by $q_2 = p_2 = 0$. The solutions contained in this plane are supported by the curves

$$\frac{1}{2}p_1^2 + q_1^4 = h, \quad q_2 = p_2 = 0.$$

I.2.f. The Spherical Pendulum. In the system called the "simple pendulum", the mass moves in a plane. Now we allow it to move in space, namely on the surface of a sphere centered at the fixed end of the cord. By a "spherical pendulum" we mean the system defined on the unit sphere in $\mathbb{R}^3$ by the Hamiltonian

$$H = \frac{1}{2}\|p\|^2 - \Gamma \cdot q,$$

where p and q are vectors in $\mathbb{R}^3$ and Γ is the vertical gravitational field.

The phase space is

$$TS^2 = \left\{(q,p) \in \mathbb{R}^3 \times \mathbb{R}^3 \mid \|q\|^2 = 1 \text{ and } q \cdot p = 0\right\}$$

(which, of course, is the tangent bundle of the sphere S^2), equipped with the restriction of the symplectic form on $\mathbb{R}^3 \times \mathbb{R}^3$, which continues to be a symplectic form (see Exercise I.13 if necessary).

The Hamiltonian vector field is the unique vector field (Y, X) tangent to TS^2 satisfying

$$(dH)_{(q,p)}(\eta, \xi) = \omega((\eta, \xi), (Y, X)) = \xi \cdot Y - \eta \cdot X$$

for all vectors (η, ξ) tangent to TS^2. Now,

$$(dH)_{(q,p)}(\eta, \xi) = p \cdot \xi - \Gamma \cdot \eta,$$

and the fact that (Y, X) is tangent to TS^2 at (q, p) is expressed by the conditions $Y \cdot q = 0$ and $Y \cdot p + q \cdot X = 0$. The unique solution is $Y = p$ and $X = \Gamma - (q \cdot \Gamma + \|p\|^2)q$, and the Hamiltonian system is then

$$\begin{cases} \dot{q} = p \\ \dot{p} = \Gamma - (q \cdot \Gamma + \|p\|^2)q. \end{cases}$$

Exercises I.13, I.14.

I.2.g. The Rigid Body With A Fixed Point. Here we consider a rigid body of mass 1 (for mathematical convenience) with center of gravity G, fixed point O and in a constant gravitational field.

In writing the equations describing the motion, it is convenient to utilize a frame attached to the rigid body. The constant gravitational field is a vector Γ depending on time (in the moving frame). Also, we let M be the angular momentum It is related to the instantaneous rotation Ω of the solid by the relation $M = \mathcal{J}(\Omega)$, where $\mathcal{J}$ is a certain constant symmetric endomorphism called *the matrix of inertia* (see, for example, [**8**]).

Then the total energy of the solid is

$$H(\Gamma, M) = \underbrace{\tfrac{1}{2} M \cdot \Omega}_{\text{kinetic}} + \underbrace{\Gamma \cdot L}_{\text{potential}},$$

where L denotes the vector $\overrightarrow{GO}$.

By applying the laws of mechanics, one can show that the differential system describing the motion of the solid is

$$\begin{cases} \dot{\Gamma} = \Gamma \times \Omega \\ \dot{M} = M \times \Omega + \Gamma \times L. \end{cases}$$

Of course, the vector Γ is no longer constant; however, it remains a unit vector[7]. We also know that the quantity $\Gamma \cdot M$, the moment of the solid with respect to the vertical axis[8], is conserved. Thus the vectors Γ and M are constrained to remain on the submanifold

$$W_a = \left\{ (\Gamma, M) \in \mathbb{R}^3 \times \mathbb{R}^3 \mid \|\Gamma\|^2 = 1 \text{ and } \Gamma \cdot M = a \right\}.$$

This submanifold is diffeomorphic to the tangent bundle of the sphere S^2 via the map

$$(\Gamma, M) \longmapsto (\Gamma, M - a\Gamma).$$

We can equip it with a symplectic form ω for which the differential system is the Hamiltonian system associated to the energy H. Be careful, ω is *not* the canonical symplectic structure for the tangent bundle of S^2. Instead, it is the form[9]

$$\omega_{(\Gamma,M)}((\xi,\eta),(\xi',\eta')) = (\xi \times M + \Gamma \times \eta) \cdot \xi' + (\Gamma \times \xi) \cdot \eta'.$$

Exercise I.20.

I.3. Completely Integrable Systems

I.3.a. Definition of a Completely Integrable System. The Hamiltonian H is constant along integral curves of the Hamiltonian system it defines, which we can write as

$$X_H \cdot H = 0 \text{ or } dH(X_H) = 0.$$

We say that H is a *first integral* for the system. More generally, a function $f : W \to \mathbb{R}$ that is constant along the trajectories of a vector field X is called a *first integral*. For a Hamiltonian vector field X_H, the equality $X_H \cdot f = 0$ is equivalent to $\{f, H\} = 0$.

Informally, we say that a Hamiltonian system is completely integrable if there are "as many first integrals in involution as possible." We now explain what this means formally.

- Let $f_1, \ldots, f_k$ be first integrals of the system X_H. We say that they are "in involution" if $\{f_i, f_j\} = 0$ for all i and j. Each f_i is constant along the trajectories of the Hamiltonian vector field defined by the others.
- We could also talk of a subalgebra[10] $\mathcal{A}$ of $\mathcal{C}^\infty(W)$ containing H that is Abelian in the sense that

$$f, g \in \mathcal{A} \Longrightarrow \{f, g\} = 0.$$

[7] If the units are suitably chosen.

[8] By definition, "vertical axis" is the direction of the gravitational field.

[9] The manifold W_a is an "adjoint orbit", which yields the form in question. See, for example, [**8**].

[10] See [**77**].

- Now, we turn to what is meant by "as many as possible." At each point x in W, the subspace of T_xW generated by the Hamiltonian vector fields of the functions in $\mathcal{A}$ is isotropic
$$\omega(X_f, X_g) = \pm\{f, g\} = 0,$$
and so its dimension is at most $n = \frac{1}{2}\dim W$. Thus for each x in an dense open subset of W, we require that this subspace have maximal dimension n.
- Finally, note that the vectors X_{f_i} are independent at x if and only if the linear forms $(df_i)_x$ are independent.

DEFINITION I.3.1. The Hamiltonian vector field X_H on a symplectic manifold W of dimension $2n$ is *completely integrable* if it has n independent first integrals in involution.

EXAMPLE I.3.2. On $\mathbb{R}^n \times \mathbb{R}^n$ with the standard symplectic form, any Hamiltonian that depends only on the coordinates p_i,
$$H = H(p_1, \dots, p_n),$$
is completely integrable: the functions p_i are independent first integrals in involution.

Any Hamiltonian system on a surface is completely integrable. Similarly, a Hamiltonian system on a symplectic manifold of dimension 4 is integrable if and only if it has a "second first integral" that is functionally independent of the energy first integral H.

I.3.b. The Harmonic Oscillator and the Simple Pendulum. See Sections I.2.a and I.2.c for a description of these systems, which are Hamiltonian systems on manifolds of dimension 2 and thus (automatically) integrable.

I.3.c. The Hénon–Heiles System. There exists a second first integral for this system (introduced in Section I.2.d) for certain values of the parameters A and λ. In particular, this is the case for $\lambda = 6$.

Exercise I.21.

I.3.d. The Anharmonic Oscillator. It is easy to check that the function
$$K = \frac{1}{2}p_1^2 - \frac{1}{\lambda}(q_1p_2 - q_2p_1)^2 + q_1^2(q_1^2 + q_2^2)$$
is a first integral for the Hamiltonian system considered in Section I.2.e and that is independent of H.

I.3.e. The Spherical Pendulum. We retain the notation introduced in Section I.2.f. The system is invariant under rotations about the vertical axis, which allows us to find easily a second first integral, i.e., the moment about this axis
$$K(q, p) = (q \times p) \cdot \Gamma.$$

One can show directly that $\{H, K\} = 0$ (see Exercise I.22), but it is more enlightening to give a geometric argument. The Hamiltonian vector field associated to K is

$$X_K(q,p) = (-q \times \Gamma, -p \times \Gamma),$$

whose flow is that of rotations about the vertical axis,

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

(this is what it means for K to be the moment about this axis). This flow preserves the submanifold TS^2, since

$$\|R_t q\|^2 = \|q\|^2 = 1 \text{ and } R_t p \cdot R_t p = q \cdot p = 0.$$

Similarly, it fixes H because it leaves the vertical vector Γ invariant. Then, by differentiating the relation $H(\varphi_t(x)) = H(x)$ at $t = 0$, we see that the Poisson bracket of H and K is zero. Thus the spherical pendulum is a completely integrable system.

Exercise I.22.

I.3.f. A Rigid Body: the Euler–Poinsot Case. This is the rigid body described in Section I.2.g when the fixed point is the center of gravity ($G = 0, L = 0$). It is easy to check that $K(M) = \|M\|^2$ is a first integral for this system and that is independent of the energy integral.

Remark I.3.3. In this case, the Hamiltonian system

$$\begin{cases} \dot{\Gamma} = \Gamma \times \Omega \\ \dot{M} = M \times \Omega \end{cases}$$

is reduced to the system $\dot{M} = M \times \Omega$ (with $M = \mathcal{J}(\Omega)$), which describes the motion of a "free" body (without gravity).

I.3.g. The Top. We continue to consider a rigid body, but this time one with an axis of rotation OG. This means that there is an orthonormal frame where the last vector is $L = \overrightarrow{GO}$ and the matrix of inertia is the diagonal matrix (ℓ, ℓ, m). We choose the units so that $\ell = 1$. Then the condition is that $M - \Omega$ is collinear to L.

This case, which Lagrange studied in 1788, is that of a "top". The tops one plays with do not really have a fixed point, but instead are constrained to having one end O of their axis of rotation in a given horizontal plane. For more about this system, see the classic work by Klein and Sommerfeld [**49**]. In this book, however, we consider "Lagrange's top" to be a reasonable approximation to the toy one usually plays with.

In this case, it is clear that the moment about the axis of rotation is a first integral.

Exercise I.20.

I.3.h. The Kowalevski Top. We continue to consider a rigid body with a fixed point but this time when the matrix of inertia $\mathcal{J}$ is the diagonal matrix $\mathcal{J} = (2, 2, 1)$ in a frame where the *first* vector is collinear to L. As in the case of Lagrange's top, there is an equatorial plane, but this time it contains the axis $\overrightarrow{OG}$. In the same basis, we write

$$M = \begin{pmatrix} u \\ v \\ w \end{pmatrix}, \quad \Omega = \begin{pmatrix} p \\ q \\ r \end{pmatrix}, \quad \Gamma = \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{pmatrix}, \quad L = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}.$$

Then the function

$$K = |(p + iq)^2 + (\gamma_1 + i\gamma_2)|^2$$

is a first integral called the "Kowalevski integral" (see [**53**]).

Remark I.3.4. These three special cases are the only cases of rigid bodies with a fixed point that are integrable systems. More precisely, in the other cases, there are no additional independent meromorphic integrals. (The polynomial analog is due to Husson [**43**] and the meromorphic case is due to Ziglin [**84, 85**]. This can also be established with the help of the Morales–Ramis theorem [**62**], as in Chapter III, see [**57**].)

I.3.i. Other Examples. There are many other interesting examples of integrable systems, notably among geodesic flows, which I will address in the next section. See, for example, the list on page 13 of [**8**]. Two of the most remarkable systems that will not appear in this book are the systems studied by Goldman [**35**][11] and Hitchin [**41**].

Exercise I.26.

I.4. Geodesic Flows

In this section we consider a Riemannian manifold V. Its metric defines a function $L : TV \to \mathbb{R}$, the "Lagrangian"

$$L(q, \xi) = \frac{1}{2}\langle \xi, \xi \rangle_q,$$

where q is a point in V and ξ is a vector in T_qV.

The geodesics of V are the extrema of the functional

$$c \longmapsto \int \|\dot{c}(t)\|^2 \, dt$$

and are given by the Euler–Lagrange equation. We thus let $p = \partial L/\partial \xi$ (which is the linear form $\langle \xi, \cdot \rangle_q$). The Euler–Lagrange equation can now be written as: $\dot{p} = \partial L/\partial q$. The Hamiltonian[12] is then the function on $T^\star V$ defined by

$$H(q, p) = p \cdot \xi - L(q, \xi) = \langle \xi, \xi \rangle_q - \frac{1}{2}\langle \xi, \xi \rangle_q = \frac{1}{2}\langle \xi, \xi \rangle_q$$

(to be written in terms of p).

[11] An overview of some integrable systems on moduli spaces can be found in [**9**].
[12] See [**2**] for its construction.

Geodesics on a Hypersurface and the Space of Lines. Let $Q \subset \mathbb{R}^{n+1}$ be a hypersurface equipped with the metric induced by the Euclidean metric. We can study the geodesic flow on Q by using a system defined on the space of lines obtained by associating to each geodesic (a curve on Q) the curve of its tangents.

The essential tool for doing this is the "Melrose" hexagonal diagram

$$\begin{tikzcd}
 & & T^\star\mathbb{R}^{n+1} & & \\
j^\star T^\star\mathbb{R}^{n+1} \arrow[urr, hook] \arrow[dd] & & & & S(T^\star\mathbb{R}^{n+1}) \arrow[ull, hook'] \arrow[dd] \\
 & & X \arrow[ull, hook'] \arrow[urr, hook] & & \\
T^\star Q & & & & \mathcal{D}_n \\
 & & \Sigma \arrow[ull] \arrow[urr] \arrow[uu, hook] & &
\end{tikzcd}$$

where

- the top of the diagram is the ambient symplectic manifold, here $T^\star\mathbb{R}^{n+1}$ (as usual, (p,q) denotes a point in this space);
- the second line contains the two hypersurfaces $j^\star T^\star\mathbb{R}^{n+1}$ (defined by $q \in Q$ where $j : Q \subset \mathbb{R}^{n+1}$ is the inclusion map) and $S(T^\star\mathbb{R}^{n+1})$ (defined by $\|p\|^2 = 1$);
- the third line contains their intersection X;
- the next line contains the two spaces of characteristics (see Exercise I.9) of the two hypersurfaces, i.e., the cotangent bundle to Q and the space $\mathcal{D}_n$ of (geodesic!) lines of $\mathbb{R}^{n+1}$;
- finally at the bottom of the diagram, Σ is the unitary cotangent bundle of Q, a hypersurface in $T^\star Q$ whose space of characteristics is the space of geodesics on Q, viewed here as
$$\Sigma = \left\{(p,q) \mid \|p\|^2 = 1 \text{ and } p \in T_q^\star Q\right\}.$$
The map $\Sigma \to \mathcal{D}_n$ sends (p,q) to the line in the direction p and passing through q.

Thus we have described Σ as a hypersurface in the two symplectic manifolds $T^\star Q$ and $\mathcal{D}_n$ as well as a submanifold of codimension 3 of the ambient space $T^\star\mathbb{R}^{n+1}$, from which come all the symplectic structures that we utilize here. Hence the characteristics of Σ are the same whether we consider Σ

- as a submanifold of $\mathcal{D}_n$, the space of lines tangent to Q,
- or as the unitary bundle in $T^\star Q$—the space of characteristics is that of the geodesics on Q.

In other words, any characteristic of the space of lines tangent to Q in the space of lines consists of lines tangent to a geodesic in Q. Or, in the case of a hypersurface of

$\mathbb{R}^{n+1}$, we can consider the geodesic flow as a Hamiltonian system on the symplectic manifold $\mathcal{D}_n$ of lines.

I.4.a. Geodesics on Surfaces of Revolution. We consider a surface of revolution in the Euclidean space $\mathbb{R}^3$. We parametrize it by a cylinder $(a,b)\times\mathbb{R}/2\pi\mathbb{Z}$

$$(s,\theta)\longmapsto\big(f(s)\cos\theta, f(s)\sin\theta, g(s)\big),$$

where the meridian curve is parametrized by the arc length s. A tangent vector ξ written as

$$\xi=\sigma\frac{\partial}{\partial s}+\tau\frac{\partial}{\partial\theta}$$

has norm

$$\|\xi\|^2_{(s,\theta)}=\frac{1}{2}(\sigma^2+f(s)^2\tau^2)=L\big((s,\theta),(\sigma,\tau)\big).$$

The corresponding cotangent variables p_s and p_θ are

$$p_s=\frac{\partial L}{\partial\sigma}=\sigma \text{ and } p_\theta=\frac{\partial L}{\partial\tau}=f(s)^2\tau.$$

Thus the Hamiltonian is

$$H((s,\theta),(p_s,p_\theta))=\frac{1}{2}(\sigma^2+f(s)^2\tau^2)=\frac{1}{2}\Big(p_s^2+\frac{1}{f(s)^2}p_\theta^2\Big),$$

and the Hamiltonian system is

$$\begin{cases}\dot{s}=p_s\\ \dot{\theta}=\dfrac{1}{f(s)^2}p_\theta\\ \dot{p}_s=\dfrac{f'(s)}{f(s)^3}p_\theta^2\\ \dot{p}_\theta=0.\end{cases}$$

Since H does not depend on θ, it is clear that p_θ is a second independent first integral. This is the simplest way to describe the Clairaut integral, whose classical description[13] appears in Exercise I.23.

Exercise I.23.

Remark I.4.1. As in the cases of the spherical pendulum and the top (see Sections I.3.e and I.3.g), the second first integral is a moment about the axis of revolution.

I.4.b. Geodesics on Quadrics. We now consider an ellipsoid in $\mathbb{R}^{n+1}$, that is, a hypersurface defined by

$$f(x)=\langle A^{-1}x,x\rangle-1=0.$$

where A is a positive definite symmetric matrix A of order $n+1$.

[13]See, for example, volume III of [**75**].

Let $t \mapsto x(t)$ be a geodesic on this hypersurface and $y(t) = \dot{x}(t)$ be its tangent vector. Saying that x is a geodesic is equivalent to saying that its acceleration $\dot{y}(t)$ has no tangent component or, in other words,

$$\dot{y}(t) = \lambda \operatorname{grad}_{x(t)} f.$$

Here $\operatorname{grad}_x f = 2A^{-1}x$ and we can easily compute λ by differentiating the relation $\langle A^{-1}x, y\rangle = 0$ (which says that y is tangent to the ellipsoid)

$$0 = \langle A^{-1}\dot{x}, y\rangle + \langle A^{-1}x, \dot{y}\rangle = \langle A^{-1}y, y\rangle + \lambda\langle A^{-1}x, A^{-1}x\rangle.$$

The differential system that yields the geodesics on the ellipsoid is thus

$$\begin{cases} \dot{x} = y \\ \dot{y} = -\dfrac{\langle A^{-1}y, y\rangle}{\|A^{-1}x\|^2} A^{-1}x. \end{cases}$$

Integrability (the Classical Approach). Classical theorems due to Jacobi and Chasles allow us to show that this system is completely integrable[14].

I especially like the geometry of this example and will devote several pages to it. Readers pressed for time can skip it on first reading (but that would be unfortunate!).

The family of "confocal" quadrics to our ellipsoid plays a very important role. These are the quadrics Q defined by an equation

$$f_z(x) = \big\langle (A - z\operatorname{Id})^{-1}x, x\big\rangle - 1 = 0$$

(with the original ellipsoid Q_0 corresponding to the value 0 for the real parameter z). For the case of conics ($n = 1$), this is the family of conics having the same foci as the given ellipse, hence the term "confocal quadric." In higher dimensions, the remarkable characteristic of this family is that the dual family is a pencil, that is, it is *linear*.

The first result utilized here is a theorem due to Jacobi that states a generic point of the space is contained in $n+1$ quadrics of the family which are orthogonal at this point of intersection.

The second one, due to Chasles, states that a generic line of the space is tangent to n quadrics in the family and that the tangent hyperplanes to these n quadrics (at the points where the line is tangent to them) are pairwise orthogonal.

Because this is very beautiful, I give a sketch of the proofs of these results below. Already, notice that the line $\{u + tp | t \in \mathbb{R}\}$ is tangent to the quadric Q_z exactly when the quadratic equation in t

$$f_z(u + tp) = 0$$

has a double root. Concretely, here the space Σ from the Melrose diagram is the hypersurface in $\mathcal{D}_n$ defined by $\Psi_z(p, u) = 0$, where

$$\Psi_z(p,u) = \big\langle (A - z\operatorname{Id})^{-1}u, p\big\rangle^2 - \big(\langle (A - z\operatorname{Id})^{-1}p, p\rangle\big)\big(\langle (A - z\operatorname{Id})^{-1}u, u\rangle - 1\big).$$

[14]The proof is also classical and can be found, for example, in [**76**]. It was known to Jacobi [**45**]. There are more modern proofs as well, using techniques of algebraic geometry. For more about geodesics on quadrics, see, for example, [**79, 50, 65, 6**].

This is a function of z of the form

$$\Psi_z(p,u) = \frac{Q_{p,u}(z)}{\prod(\alpha_i - z)},$$

where Q is a polynomial of degree n and where we supposed that A is a diagonal with eigenvalues α_i.

The condition that the line (p,u) is tangent to the quadric Q_z is thus equivalent to the fact that z is a root of a certain polynomial $Q_{(p,u)}(z)$ of degree n whose roots, Chasles tells us, are real. Locally, we have functions

$$z_1(p,u), \ldots, z_n(p,u)$$

describing the n quadrics to which the line $\{u + tp | t \in \mathbb{R}\}$ is tangent.

If we consider the Hamiltonian system describing the geodesics on the hypersurface Q as a system on the space $\mathcal{D}_n$ of lines, the functions $z_1, \ldots, z_n$ (locally) define functions on this space, and we now show (following Jacobi and Chasles) that these functions are in involution.

Let X_i be the Hamiltonian field associated to the function z_i. It is tangent to the characteristics of the level hypersurfaces of z_i, that is, the space of lines tangent to the quadric Q_{z_i} (for a fixed value of z_i). Thanks to the Melrose diagram discussed at the beginning of Section I.4, we know that this hypersurface consists of lines tangent to a geodesic of Q_{z_i}. We also know that these lines are tangent to Q_{z_j} for $j \neq i$. Thus z_j is constant along the flow of X_i, or in other words, $\{z_i, z_j\} = 0$.

Two Theorems on Confocal Families. Let us now prove the Jacobi and Chasles theorems that we have just used.

Proof of the Jacobi Theorem. We begin by translating the statement into one in terms of the dual family. We will show that every affine hyperplane of $(\mathbb{R}^{n+1})^\star$ not passing through the origin[15] is tangent to n quadrics of the pencil and that the vectors corresponding to the points of tangency are pairwise orthogonal. The equation of quadrics of the pencil is

$$\langle (A - z\,\mathrm{Id})x, x\rangle = 1.$$

Consider the quadratic form $B(x) = \langle Ax, x\rangle - \langle x, y\rangle$ and a line ℓ passing through the eigenvector corresponding to the eigenvalue z. Thus the quadratic form $B - z\,\mathrm{Id}$ and its polar form are zero on ℓ. Let y be a point in the space $\mathbb{R}^{n+1}$. The dual hyperplane is determined by $\langle x, y\rangle = 1$; it intersects the line ℓ at a point x_0 on the quadric $\langle (A - z\,\mathrm{Id})x, x\rangle = 1$, since $(B - z\,\mathrm{Id})x_0 = 0$. Moreover, since the polar form on $B - z\,\mathrm{Id}$ is also zero on ℓ, the quadric defined by $A - z\,\mathrm{Id}$ is tangent to the hyperplane at x_0.

The points of tangency of the hyperplane dual to y with the quadrics in the pencil are the eigenvectors of the quadratic form B. Since B is a real quadratic form, it has n independent orthogonal eigenvectors. □

[15] This is the condition of being generic in the original statement.

Proof of the Chasles Theorem. We choose a line ℓ and project (orthogonally) $\mathbb{R}^{n+1}$ onto the hyperplane $\ell^\perp$. We consider the apparent contours[16] of the quadrics of the family onto $\ell^\perp$.

This family of hypersurfaces of $\ell^\perp$ is still a confocal family. We utilize the duality that transforms the apparent contour of Q on $\ell^\perp$ into the intersection of the dual quadric with $\ell^\perp$. Since the intersection of a linear family of quadrics with a hyperplane is certainly a linear family of quadrics, our family of apparent contours is indeed a confocal family.

We now apply Jacobi's theorem to the point of intersection of ℓ with $\ell^\perp$. It lies on n quadrics of the family of apparent contours which is equivalent to saying the line ℓ is tangent to n quadrics of the original confocal family. $\square$

I will come back to this system and its integrability at length in Chapter IV.

Exercise I.24.

I.5. Appendix: The Theorem of Darboux

This is a theorem that asserts that, locally, all symplectic forms are isomorphic.

Theorem I.5.1 ([**64, 80**]). *Let ω_0 and ω_1 be two symplectic forms on W that are equal at x. Then there is a neighborhood U_0 of x in W and a map*

$$\psi : (U_0, x) \longrightarrow (W, x)$$

such that $\psi^\star\omega_1 = \omega_0$.

Remark I.5.2. The map ψ is then a local diffeomorphism since $\omega_0^{\wedge n} = \psi^\star\omega_1^{\wedge n}$ and these two $2n$-forms are volume forms.

Corollary I.5.3 (Theorem of Darboux). *Let x be a point on a symplectic manifold W with symplectic form ω. Then there exist local coordinates $(p_1, \dots, p_n, q_1, \dots, q_n)$ centered at x such that $\omega = \sum dp_i \wedge dq_i$.*

Proof of the Corollary. Let U be a neighborhood of 0 in $\mathbb{R}^{2n}$ and $f : U \to W$ a coordinate chart centered at x. We consider two symplectic forms on U:

- the form $f^\star\omega$
- the constant form $(f^\star\omega)_0$.

By definition, these two forms are equal at 0, so the Theorem asserts that they are diffeomorphic. The coordinates $(q_1, \dots, q_n, p_1, \dots, p_n)$ in a symplectic basis for $(f^\star\omega)_0$, transported by ψ, have the desired property. $\square$

Proof of the Theorem. We apply the path method of Moser (see [**64**]). Namely, for each $t \in [0,1]$, we consider the form $\omega_t = \omega_0 + t(\omega_1 - \omega_0)$. It is clearly closed. Since ω_0 and ω_1 are equal at x, it is nondegenerate at x and thus on some neighborhood of x since the property of being nondegenerate is an open property. Then, since the interval $[0,1]$ is compact, we can find a neighborhood of x on which all the ω_t are symplectic forms.

[16] The apparent contour of one of the quadrics is the set of critical values of the projection. It consists of the images of the points where ℓ is tangent to the quadric.

Since ω_0 and ω_1 are closed, the same is true for $\omega_0 - \omega_1$, and thus we can find a 1-form α such that $d\alpha = \omega_0 - \omega_1$ in a neighborhood of x. Let f be a function defined on a neighborhood of x such that $(df)_x = \alpha_x$. Then $(\alpha - df)_x = 0$ and $d(\alpha - df) = d\alpha = \omega_0 - \omega_1$. By replacing α by $\alpha - df$, we can thus assume that $\alpha_x = 0$.

Since the 2-form ω_t is symplectic, it defines a duality between the tangent and cotangent bundles, yielding for each t, a vector field X_t dual to α by ω_t (i.e., such that $i_{X_t}\omega_t = \alpha$). We now consider X_t as a vector field (depending on the time t) that is zero at x for every t. The flow φ_t of X_t fixes x, and thus we can find a neighborhood U of x on which φ_t is defined and satisfies $\varphi_t(U) \subset U$.

Hence we have

$$\frac{d}{dt}\left[\varphi_t^\star \omega_t\right] = \varphi_t^\star \left[\frac{d\omega_t}{dt} + \mathcal{L}_{X_t}\omega_t\right] = \varphi_t^\star \left[\omega_1 - \omega_0 + \omega_0 - \omega_1\right] = 0,$$

since $\mathcal{L}_{X_t}\omega_t = di_{X_t}\omega_t + i_{X_t}d\omega_t = d\alpha$ (from Cartan's formula and the definition of α). Thus the form $\varphi_t^\star\omega_t$ does not depend on t and equals ω_0 at $t = 0$. The desired result follows by taking $\psi = \varphi_1$. □

Remark I.5.4. The statement analogous to Theorem I.5.1 for Riemannian metrics is definitely false. In the case of a Riemannian metric, there is a *local* invariant distinguishing a neighborhood of a point from its tangent space, namely, the curvature. We have just shown that there is *no* local invariant in symplectic geometry.

A more general theorem can be obtained by essentially the same method [**80**]. It describes the tubular neighborhoods of isotropic submanifolds of W (those on which ω is identically zero). In particular, it states that if L is an isotropic submanifold of maximal dimension (so $\dim L = \frac{1}{2}\dim W$, and L is called *Lagrangian*), a tubular neighborhood of L in W is (symplecticly) isomorphic to a neighborhood of the zero section of $T^\star L$ (with the canonical symplectic form).

Exercises

Exercise I.1. We consider $\mathbb{C}^n$ as a *real* vector space of dimension $2n$. Show that the (real) bilinear form

$$\omega(X, Y) = \operatorname{Im}\langle X, Y\rangle$$

where the bracket denotes the standard Hermitian form on $\mathbb{C}^n$, is alternating and nondegenerate. Compare this with the form defined on $\mathbb{R}^n \times \mathbb{R}^n$ in Section I.1.a. Construct a symplectic basis from the unitary complex basis of $\mathbb{C}^n$.

Exercise I.2. Let ω be a skew-symmetric bilinear form on a vector space E. Show that there exists a basis $e_1, \ldots, e_r, f_1, \ldots, f_r, g_1, \ldots, g_r$ of E such that

$$\omega(e_i, f_j) = \delta_{i,j}, \quad \omega(e_i, e_j) = \omega(e_i, g_\ell) = \omega(f_i, g_\ell) = \omega(g_\ell, g_m) = 0$$

for all indices i, j, ℓ, and m.

Let $\dim E = 2n$. Show that $\omega^{\wedge n} \neq 0$ if and only if ω is nondegenerate.

Exercise I.3. Let F be a k-dimensional isotropic subspace of a symplectic vector space E. Show that there exists a symplectic basis $(e_1, \dots, e_n, f_1, \dots, f_n)$ of E such that $(e_1, \dots, e_k)$ is a basis of F.

Exercise I.4 (Symplectic Reduction). Let F be an isotropic subspace of a symplectic vector space E. Show that the symplectic form on E defines a symplectic form on the quotient space F°/F.

Let L be a Lagrangian subspace of E such that $L + F^\circ = E$. Show that $L \cap F^\circ$ *embeds*[17] as a Lagrangian subspace of F°/F.

Exercise I.5 (The Symplectic Group). Show that $\mathrm{Sp}(2; \mathbb{R}) = \mathrm{SL}(2; \mathbb{R})$.

Compute the dimension of the kernel of the linear map $A \mapsto {}^tAJ + JA$, and deduce that the dimension of $\mathrm{Sp}(2n; \mathbb{R})$ is $2n^2 + n$.

Given a symplectic basis of $\mathbb{R}^{2n}$, we write a $2n \times 2n$ matrix as an $n \times n$ block matrix

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}.$$

What conditions on A, B, C, and D make this matrix an element of the Lie algebra $\mathfrak{sp}(2n)$? Of the Lie group $\mathrm{Sp}(2n)$?[18]

Exercise I.6. Let L be a Lagrangian subspace of a symplectic vector space E. Show that the group of symplectic automorphisms of E that restrict to the identity on L is isomorphic to the group of matrices of the form

$$\begin{pmatrix} \mathrm{Id} & B \\ 0 & \mathrm{Id} \end{pmatrix},$$

where B is a symmetric matrix. Show that this group is Abelian.

Exercise I.7 (The Liouville Form is Canonical). A differential 1-form on a manifold V is a section of the cotangent bundle $T^\star V$

$$\eta : V \longrightarrow T^\star V \qquad \eta(x) \in T^\star_x V.$$

(Do you agree?) Letting α be the Liouville form, show that

$$\eta^\star \alpha = \eta.$$

Exercise I.8 (Symplectic Form on the Torus). Show that the standard symplectic form $\omega = \sum dp_i \wedge dq_i$ on $\mathbb{R}^{2n}$ defines a symplectic form on the torus $T^{2n} = \mathbb{R}^{2n}/\mathbb{Z}^{2n}$ when we pass to the quotient.

Exercise I.9 (Symplectic Reduction—Continuation). Let (W, ω) be a symplectic manifold. If $V \subset W$ is a hypersurface, what can be said about the rank of ω on V? Suppose that the integral curves of the kernel of $\omega_{|V}$ (the "characteristics") form a manifold Z. Show that ω defines a symplectic structure[19] on Z.

[17] A minor miracle of symplectic geometry.

[18] This is another way of computing the dimension of the symplectic group.

[19] The symplectic manifold obtained is called "the space of characteristics".

- Determine the symplectic manifold obtained when $W = \mathbb{R}^{n+1} \times \mathbb{R}^{n+1}$ and $V = S^n \times \mathbb{R}^{n+1}$.
- Consider the set of geodesics (oriented but non-parametrized curves) on a Riemannian manifold, which we suppose is a manifold. Show that it can be equipped with a symplectic structure.

EXERCISE I.10 (Hamiltonian Vector Field and the Poisson Bracket in Coordinates). Show that on $W = \mathbb{R}^n \times \mathbb{R}^n$ with the standard symplectic form, we have

$$X_H(q_1, \dots, q_n, p_1, \dots, p_n) = \left(\frac{\partial H}{\partial p_1}, \dots, \frac{\partial H}{\partial p_n}, -\frac{\partial H}{\partial q_1}, \dots, -\frac{\partial H}{\partial q_n}\right)$$

for any function H and

$$\{f, g\} = \sum_{i=1}^{n} \left(\frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i} - \frac{\partial g}{\partial p_i}\frac{\partial f}{\partial q_i}\right)$$

for all functions f and g.

What are the values of the Poisson brackets $\{p_i, p_j\}$, $\{p_i, q_j\}$, $\{q_i, q_j\}$?

EXERCISE I.11. Show that any Hamiltonian vector field on a compact symplectic manifold has a zero.

EXERCISE I.12. Consider the phase space $\mathbb{R} \times \mathbb{R}$, and fix a real number $a > 0$. Write out and solve the Hamiltonian system associated to the function $H = \frac{1}{2}p^2 + \frac{1}{2}aq^2$, and draw its trajectories.

EXERCISE I.13. Show that the standard symplectic form on $\mathbb{R}^{n+1} \times \mathbb{R}^{n+1}$ restricts to a symplectic form on the submanifold

$$TS^n = \left\{(q, p) \mid \|q\|^2 = 1 \text{ and } q \cdot p = 0\right\}.$$

EXERCISE I.14. Let W be a symplectic manifold. Suppose that V is a symplectic submanifold of W, that is, that the symplectic form on W defines a *symplectic* form on V.

- Show that, for any $v \in V$,
$$T_vV \oplus (T_vV)^\circ = T_vW.$$
- Let $H : W \to \mathbb{R}$ be a function. Show that the Hamiltonian vector field of the restriction of H to V is the projection of the Hamiltonian vector field X_H to the tangent bundle of V (for the decomposition obtained above).
- Use this method to compute the Hamiltonian vector field for the spherical pendulum.

EXERCISE I.15. Let X and Y be two vector fields that are *locally* Hamiltonian, i.e., such that $i_X\omega$ and $i_Y\omega$ are closed forms. Show that their bracket $[X, Y]$ is (globally) Hamiltonian, i.e. that the form $i_{[X,Y]}\omega$ is exact.

EXERCISE I.16. In defining Hamiltonian vector fields and the Poisson bracket, the *only* property of the symplectic form that we have used is that it is nondegenerate. Thus let ω be a nondegenerate 2-form. Determine $(d\omega)_x(X, Y, Z)$ when X, Y, and Z are vectors in T_xW that are the values at x of the Hamiltonian vector fields associated to three function f, g, and h. Show that ω is closed if and only if the Poisson bracket it defines satisfies the Jacobi identity[20].

EXERCISE I.17. Let $\mathcal{Q}$ be the vector space of quadratic forms on $\mathbb{R}^{2n}$. What is its dimension?

Consider the standard symplectic form $\omega = \sum dp_i \wedge dq_i$ on $\mathbb{R}^{2n}$. Show that $\mathcal{Q}$ is invariant under the Poisson bracket and that the Lie algebra $\left(\mathcal{Q}, \{\ ,\ \}\right)$ is isomorphic to the symplectic Lie algebra $\mathfrak{sp}(2n;\mathbb{R})$[21].

EXERCISE I.18. In the (q,p) plane, draw the level sets of the Hamiltonian $H = \frac{1}{2}p^2 - \cos q$ for the simple pendulum and the trajectories of the vector field (see also Figure 4 in Chapter II).

EXERCISE I.19. Consider the Hénon–Heiles Hamiltonian

$$H = \frac{1}{2}(p_1^2 + p_2^2) - q_2^2(A + q_1) - \frac{\lambda}{3}q_1^3$$

(see Section I.2.d) with $\lambda \neq 0$.

- Draw the trajectories of X_H contained in the plane $q_2 = p_2 = 0$.
- Determine explicitly the solution $(q_1(t), p_1(t))$ such that

$$\frac{1}{2}p_1^2(t) - \frac{\lambda}{3}q_1^3(t) = 0.$$

EXERCISE I.20 (The Rigid Body). Show that the bilinear forms $\omega_{(\Gamma,M)}$ defined by

$$\omega_{(\Gamma,M)}\big((\xi,\eta),(\xi',\eta')\big) = (\xi \times M + \Gamma \times \eta)\cdot\xi' + (\Gamma\times\xi)\cdot\eta'$$

define a symplectic form ω on the 4-dimensional manifold

$$W_a = \left\{(\Gamma, M) \in \mathbb{R}^3 \times \mathbb{R}^3 \mid \|\Gamma\|^2 = 1 \text{ and } \Gamma\cdot M = a\right\}.$$

Using the notation from Section I.2.g, check that the Hamiltonian vector field associated to $H = \frac{1}{2}M\cdot\Omega + \Gamma\cdot L$ is

$$X_H(\Gamma, M) = (\Gamma\times\Omega, M\times\Omega + \Gamma\times L).$$

Let $L = 0$ (this is the Euler–Poinsot case). Show that the function K defined by

$$K(M) = \|M\|^2$$

commutes with H.

Finally, suppose that $M - \Omega$ is collinear with L (this is the case of the top). Show that the function K defined by

$$K(M) = M\cdot\Gamma$$

commutes with H.

[20]This is the best possible justification for requiring the closure condition in the definition of a symplectic form: the Poisson bracket should define a Lie algebra structure.

[21]This is a third way to compute the dimension of the symplectic group.

EXERCISE I.21. Consider the function K defined on $\mathbb{R}^2 \times \mathbb{R}^2$ by

$$K = q_2^4 + 4q_1^2 q_2^2 + 4p_2(p_2 q_1 - p_1 q_2) + 8A(q_2^2 q_1 - p_2^2 + 2Aq_2^2).$$

Show that the Hénon–Heiles Hamiltonian (see Section I.2.d) commutes[22] with K if and only if $\lambda = 6$.

EXERCISE I.22. Consider the Hamiltonian H of the spherical pendulum and the moment

$$K(q,p) = q_1 p_2 - q_2 p_1.$$

By calculating the Poisson bracket $\{H, K\}$, show that K is indeed a first integral.

EXERCISE I.23 (The Clairaut Integral). Consider a surface of revolution. For a geodesic $c(t)$ on the surface, let $\alpha(t)$ be the angle the vector $\dot{c}(t)$ makes with the meridian passing through $c(t)$, and let $r(t)$ be the distance from $c(t)$ to the axis of rotation. Show that $r(t)\sin\alpha(t)$ does not depend on t (by showing that, in the notation of Section I.4.a, we have $p_\theta(t) = r(t)\sin\alpha(t)$).

EXERCISE I.24. Using the notation of Section I.4.b, show that the Hamiltonian system associated to the function $\frac{1}{2}\Psi_z$ on the space $\mathcal{D}_n$ of lines is

$$\begin{cases} \dot{p} = \Gamma p \\ \dot{u} = \Gamma u - (A - z\,\mathrm{Id})^{-1} p, \end{cases}$$

where Γ is the matrix with entries $\Gamma_{i,j} = \dfrac{p_i u_j - p_j u_i}{(\alpha_i - z)(\alpha_j - z)}$.

EXERCISE I.25. The complex[23] vector space $\mathbb{C}^{2n}$ is equipped with the symplectic form $\sum dp_i \wedge dq_i$. Consider a matrix $A \in \mathfrak{sp}(2n;\mathbb{C})$.

Show that the characteristic polynomial of A has the form

$$P(\lambda) = \lambda^{2r} Q(\lambda) Q(-\lambda)$$

for some integer r and some polynomial Q.

Show that there exist symplectic subspaces $E_1, \dots, E_k$ of $\mathbb{C}^{2n}$, pairwise orthogonal and invariant under A, such that $\mathbb{C}^{2n} = E_1 \oplus \cdots \oplus E_k$ and such that each E_i has a symplectic basis in which the matrix $A_{|E_i}$ has one the following forms

$$\left[\begin{array}{cccc|cccc} \alpha & & & & & & & \\ -1 & \ddots & & & & & & \\ & \ddots & \ddots & & & 0 & & \\ & & -1 & \alpha & & & & \\ \hline & & & & -\alpha & 1 & & \\ & & & & & \ddots & \ddots & \\ & 0 & & & & & \ddots & 1 \\ & & & & & & & -\alpha \end{array}\right], \quad \left[\begin{array}{cccc|cccc} 0 & & & & & & & \\ -1 & \ddots & & & & & & \\ & \ddots & \ddots & & & 0 & & \\ & & -1 & 0 & & & & \\ \hline 0 & & & & 0 & 1 & & \\ & \ddots & & & & \ddots & \ddots & \\ & & \ddots & & & & \ddots & 1 \\ & & & 1 & & & & 0 \end{array}\right],$$

[22] The second first integral comes from [**20**].

[23] One can obtain analogous results in the real case. As always in linear algebra, this case is somewhat more delicate but not much more difficult.

where α is complex number and where all the entries not written are zero.

Now consider the vector space $\mathbb{C}^{2r}$ with a matrix B of one of the above forms. Show that the subspace

$$Z_B = \{M \in \mathfrak{sp}(2r;\mathbb{C}) \mid [M,B] = 0\}$$

is a Lie subalgebra of $\mathfrak{sp}(2r;\mathbb{C})$ containing an Abelian Lie subalgebra of dimension r.

Exercise I.26 (Normal Form for Quadratic Forms). This exercise is a simple reformulation of the one above. Let Q be a quadratic form on the complex vector space $\mathbb{C}^{2n}$ equipped with the complex symplectic form $\sum dp_i \wedge dq_i$. Show that there is a decomposition of $\mathbb{C}^{2n}$ into a sum of symplectic subspaces $E_1, \ldots, E_k$, pairwise orthogonal for the symplectic form *and* for the quadratic form Q, such that

- $\mathbb{C}^{2n} = E_1 \oplus \cdots \oplus E_k$ and
- each E_i has a symplectic basis in which the quadratic form $Q_{|E_i}$ is either of the form

$$\alpha(p_1q_1 + \cdots + p_rq_r) - (p_2q_1 + \cdots + p_rq_{r-1})$$

or

$$p_2q_1 + \cdots + p_rq_{r-1} + \frac{1}{2}q_r^2.$$

Show that a quadratic Hamiltonian on $\mathbb{C}^{2n}$ must have n independent first integrals in involution. In particular, it is integrable[24].

[24] These results (including the real case) are due to Williamson, in [**82**] for the normal forms and in [**83**] for the integrability.

CHAPTER II

Action–Angle Variables

In this chapter, I prove the Arnold–Liouville theorem[1], which describes a symplectic manifold and completely integrable Hamiltonian system in a neighborhood of a common regular level set of first integrals for this system. With a compactness hypothesis, these regular level sets are Lagrangian tori on which the Hamiltonian system is *linear*.

The "semi-local" model is that of a Hamiltonian action of a torus—an important concept in symplectic geometry. We thus open this chapter with a discussion of Hamiltonian torus actions.

For the proof of the theorem itself, I have closely followed Duistermaat's beautiful article [**24**]. The general commentaries about Lagrangian fibrations can also be found in [**4**].

Terminology. The notion of "action-angle *variables*" is exactly the same as that of "action angle *coordinates*". We will utilize the former more classical language in this book.

II.1. Hamiltonian Torus Actions

II.1.a. Symplectic Actions. Let G be a Lie group acting on a symplectic manifold (W, ω). We say that the action is *symplectic* if G leaves ω invariant, that is, if we have

$$g^\star \omega = \omega$$

for all g in G. This can be stated in terms of fundamental vector fields that describe infinitesimally the action of G. For any vector X in the Lie algebra $\mathfrak{g}$ of G, we define a vector field $\widetilde{X}$ on W by

$$\widetilde{X}(x) = \frac{d}{dt}(\exp(tX) \cdot x)_{|t=0}.$$

This is called the *fundamental vector field* associated to X. By differentiating the relation $g^\star \omega = \omega$, we see that

$$\forall\, X \in \mathfrak{g} \qquad \mathcal{L}_{\widetilde{X}} \omega = 0.$$

Since the form ω is closed, Cartan's formula implies that $d i_{\widetilde{X}} \omega = 0$, and thus the forms $i_{\widetilde{X}} \omega$ are *closed*.

[1]This theorem is stated (and proved) in Arnold's book [**2**] under the name of "théorème de Liouville". As I learned in San Vũ-Ngọc's thesis [**78**], one can also find this theorem in an older but forgotten paper by Mineur [**60**].

II.1.b. Hamiltonian Circle Actions. Let G be the circle $S^1 = \mathbb{R}/\mathbb{Z}$. The Lie algebra is $\mathbb{R}$. Let X be the canonical basis vector of $\mathbb{R}$. This vector defines "the" fundamental vector field of the action of S^1. The action is symplectic if and only if the form $i_{\widetilde{X}}\omega$ is closed. If it is *exact* (that is, if there is a function H such that $i_{\widetilde{X}}\omega = -dH$), we say that the action is *Hamiltonian*.

Periodic Hamiltonians. Now going in the other direction we begin with a function H on a symplectic manifold. It defines a vector field X_H. We have seen in the 2-dimensional examples (see Sections I.2.a and I.2.b) that this vector field often has periodic orbits[2].

We consider the case of a Hamiltonian vector field that is *truly* periodic (that is, not only are all its orbits periodic, but all its orbits have the same period). We can then consider the corresponding flow of the vector field as a circle action. We thus have a function $H : W \to \mathbb{R}$ whose Hamiltonian vector field X_H is the fundamental vector field of an action of the circle S^1. This means that S^1 acts on W as

$$\begin{aligned} S^1 \times W &\longrightarrow W \\ (t, x) &\longmapsto t \cdot x, \end{aligned}$$

and that the vector tangent to the orbit at x is $X_H(x)$. As Figure 1 indicates, we can also write

$$\frac{d}{d\theta}(e^{i\theta} \cdot x)_{|\theta=0} = X_H(x).$$

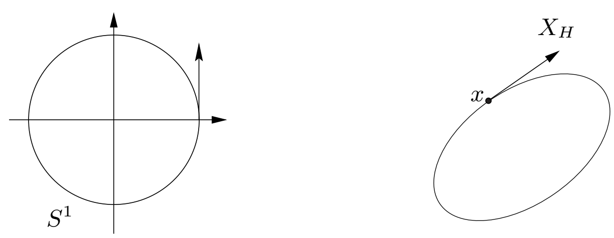


FIGURE 1. The fundamental vector field

We say that H is a *periodic Hamiltonian*, or, equivalently, that the S^1-action is Hamiltonian. We also say that H is a moment map for the action of the circle.

EXAMPLE II.1.1. The simplest example is that of the harmonic oscillator

$$H = \frac{1}{2}p^2 + \frac{1}{2}q^2$$

(see Section I.2.a), which is the moment map for the action of S^1 on $\mathbb{R}^2$ by rotations about the origin.

Exercises II.1, II.2.

To conclude this section, I include a result which will be useful in the proof of the Arnold–Liouville theorem.

[2]The Arnold–Liouville theorem that we discuss in this chapter is a very precise way of generalizing this remark.

PROPOSITION II.1.2. *Let (V, ω) be an* exact *symplectic manifold, and let λ be a 1-form on V satisfying $\omega = d\lambda$. Then any symplectic action of S^1 on V is Hamiltonian with Hamiltonian*

$$J(v) = \int_{S^1 \cdot v} \lambda.$$

PROOF. The proof is straightforward: we compute $(dJ)_v(Z)$ for Z in T_vV and show that

$$(dJ)_v(Z) = \omega(Z, Y(v)),$$

where Y is the fundamental vector field of the action of the circle.

Thus let v be a point in V and Z be a vector tangent to V at this point. We begin by extending Z to an invariant vector field along the orbit of v by setting

$$Z(t \cdot v) = T_v t(Z).$$

We then extend Z to a vector field (still called Z) on V and set

$$J(v) = \int_0^1 \lambda_{t \cdot v}(Y(t \cdot v))dt$$

(here, as the readres have understood, t denotes both an element of $S^1 = \mathbb{R}/\mathbb{Z}$ and the corresponding element in $[0, 1]$) and differentiate. We must calculate the derivative of the function

$$v \longmapsto \lambda_{t \cdot v}(Y(t \cdot v))$$

on Z (for a fixed t in S^1), which we see is equal to

$$(\mathcal{L}_Z \lambda)_{t \cdot v}(Y(t \cdot v)) + \lambda_{t \cdot v}((\mathcal{L}_Z Y)(t \cdot v)).$$

The second term of this sum is zero. Indeed, we have $\mathcal{L}_Z Y = -\mathcal{L}_Y Z$, and $(\mathcal{L}_Y Z)(t \cdot v)$ depends only on the field Z along the orbit of v (see Lemma III.1.3). Then, since Z is invariant along this orbit, we have $\mathcal{L}_Y Z(t \cdot v) = 0$.

We now compute the first term. We have

$$(d\lambda)(Z, Y) = Z \cdot \lambda(Y) - Y \cdot \lambda(Z) - \lambda([Z, Y]).$$

We have already said that the bracket $[Z, Y]$ is zero at the points $t \cdot v$; thus the last term is zero, and we have

$$(\mathcal{L}_Z \lambda)_{t \cdot v}(Y(t \cdot v)) = (d\lambda)_{t \cdot v}(Z(t \cdot v), Y(t \cdot v)) + (\mathcal{L}_Y \lambda)_{t \cdot v}(Z(t \cdot v)).$$

To calculate $(dJ)_v(Z)$, we have to integrate the remaining two terms on $[0, 1]$.

We start by the second term. By differentiating the function

$$t \longmapsto \lambda_{t \cdot v}(Z(t \cdot v)),$$

we get

$$\frac{d}{dt}\big(\lambda_{t \cdot v}(Z(t \cdot v))\big) = (\mathcal{L}_Y \lambda)_{t \cdot v}(Z(t \cdot v)) + \lambda_{t \cdot v}\big((\mathcal{L}_Y Z)(t \cdot v)\big) = (\mathcal{L}_Y \lambda)_{t \cdot v}(Z(t \cdot v)).$$

Then integrating gives us

$$\int_0^1 (\mathcal{L}_Y \lambda)_{t \cdot v}(Z(t \cdot v))dt = \int_0^1 \frac{d}{dt}\big(\lambda_{t \cdot v}(Z(t \cdot v))\big)dt = 0.$$

Up to now, we have only used the fact that ω is exact and that we can extend Z to an invariant vector field along the orbit of v. We now will use the fact that Y is the fundamental vector field of a *symplectic* action; that is, we have

$$\mathcal{L}_Y\omega = 0, \text{ or } di_Y\omega = 0.$$

Since Y and Z are invariant along the orbit of v, we see that the function

$$t \longmapsto \omega_{t\cdot v}\big(Z(t\cdot v), Y(t\cdot v)\big)$$

is constant. Indeed, its derivative with respect to t is

$$\Big(\frac{d}{dt}\omega_{t\cdot v}\Big)\big(Z(t\cdot v), Y(t\cdot v)\big) + \omega_{t\cdot v}\big((\mathcal{L}_Y Z)_{t\cdot v}, Y(t\cdot v)\big) + \omega_{t\cdot v}\big(Z(t\cdot v), (\mathcal{L}_Y Y)(t\cdot v)\big),$$

a sum of three terms all of which are equal to zero. Finally, we have that

$$\int_0^1 \omega_{t\cdot v}\big(Z(t\cdot v), Y(t\cdot v)\big)dt = \omega_v(Z, Y(v)),$$

that is,

$$(dJ)_v(Z) = \omega_v(Z, Y(v)),$$

which is what we wanted to prove. □

Remark II.1.3. We did not actually use the full hypothesis that the action is symplectic ($\mathcal{L}_Y\omega = 0$). Instead, we only needed the fact that

$$i_Y\mathcal{L}_Y\omega = 0.$$

Thus the proposition remains true under this weaker condition.

II.1.c. Hamiltonian Torus Actions. More generally, if a torus T (written as a product $(S^1)^k$) acts on a symplectic manifold W, we obtain k vector fields $X_1, \dots, X_k$, the fundamental vector fields of the k factors. We say that the action is *Hamiltonian* if each of these is a Hamiltonian vector field. By putting together all of these periodic Hamiltonians, we obtain the map

$$H = (H_1, \dots, H_k) : W \longrightarrow \mathbb{R}^k,$$

which is called a *moment map*[3].

The Lie brackets of the fundamental vector fields $X_1, \dots, X_k$ are 0 since the group T^k and its Lie algebra are Abelian[4]. This implies that $\{H_i, H_j\} = 0$ and the subspace of T_xW generated by the vectors $X_1(x), \dots, X_k(x)$ is isotropic (see Section I.3).

We deduce from this that H is constant on the orbits. We have

$$dH_j(X_i) = \{H_i, H_j\} = 0,$$

and thus $(dH)_x$ is zero on the subspace generated by the fundamental vector fields, that is, on the tangent space to the orbit of x.

[3] It is easy to find definitions that are more intrinsic (here I made use of a basis) and more general (by replacing the torus by a Lie group). See, for example, [**15**].

[4] Here we use the fact that the bracket of fundamental vector fields is the fundamental vector field of the bracket.

Orbits, Effective Actions. By using equivariant tubular neighborhoods of the orbits, one can show[5] that

- the stabilizers of the points in a neighborhood of x are subgroups of the stabilizer of x (since the torus T is Abelian).
- if W is connected, there is a dense open set in W whose points have the same stabilizer (since the group is compact and Abelian), and, in particular, this stabilizer is contained in all the stabilizers of the points of W.

We say that the action is *effective* if each element of the torus (except 1) effectively displaces a point on the manifold

$$\forall\, t \in T - \{1\}\,,\ \exists\, x \in W, \quad t \cdot x \neq x.$$

This is to say, equivalently, that the intersection of all the stabilizers of points of W is the trivial subgroup. Thus there is a dense open set of points of W with a trivial stabilizer. Consequently the orbits of these points are tori T.

We suppose that the torus T^k acts effectively on the manifold W. The vectors $X_1(x), \dots, X_k(x)$ are independent on the dense open subset U of W consisting of points x with trivial stabilizer.

As in Section I.3, we deduce that the maximum possible dimension of a torus acting effectively on a $2n$-dimensional symplectic manifold is n. Hence the moment map of this action is an integrable system.

Example II.1.4. In Example II.1.1, we saw that the harmonic oscillator $H = \frac{1}{2}(p^2 + q^2)$ is the moment map for the action of S^1 on $\mathbb{R}^2$ by rotation about the origin. More generally,

$$H = \frac{1}{2}(p_1^2 + q_1^2, \dots, p_n^2 + q_n^2)$$

is the moment map for the "product" action of the torus T^n on $\mathbb{R}^{2n}$. The most beautiful of all these Hamiltonian actions[6] is that of the torus T^n on the projective space $\mathbb{P}^n(\mathbb{C})$. See Exercise II.5.

Exercises II.3, II.4, II.5; see also IV.20.

II.2. The Arnold–Liouville Theorem

This theorem states that in a neighborhood of a compact connected component of a level set of an integrable system, we are in the situation of a Hamiltonian action of a torus.

Here we consider an integrable system, that is, a $2n$-dimensional symplectic manifold (W, ω) with a map $f : W \to \mathbb{R}^n$ whose components $f_1, \dots, f_n$ are independent and in involution.

[5] See, for example, [**21**].

[6] The generalization to the action of the torus T^n of diagonal unitary matrices by conjugation on the Hermitian matrices is not bad either! See [**15**] for an elementary discussion (including figures) of these examples.

II.2.a. The Local Action of $\mathbb{R}^n$. The manifold W is equipped with a local action of $\mathbb{R}^n$ that is locally free at regular points of $f_1, \dots, f_n$ as we show now. Let $X_1, \dots, X_n$ be the Hamiltonian vector fields associated to the f_i. These vector fields commute

$$[X_i, X_j] = X_{\{f_i, f_j\}} = 0.$$

We then obtain the action of $\mathbb{R}^n$ by integrating, i.e.,

$$t \cdot x = \varphi_n^{t_n} \circ \varphi_{n-1}^{t_{n-1}} \circ \cdots \circ \varphi_1^{t_1}(x),$$

where φ_i is the flow of X_i and $t = (t_1, \dots, t_n) \in \mathbb{R}^n$ is sufficiently close to 0 (so that $\varphi_i^{t_i}$ is defined).

This local action is locally free[7] on the open set of regular points in W, since the vector fields X_i are independent at these points.

When the Vector Fields are Complete. Let us make now the additional hypothesis that the vector fields X_i are complete (so the flows $\varphi_i^{t_i}$ are defined for all values of t_i). We then have a locally free action of $\mathbb{R}^n$ on the open set of regular points on W. Since the vector fields X_i are tangent to the common level sets of the f_i, this action preserves these level sets.

The connected components of the level sets of f are homogeneous spaces of $\mathbb{R}^n$ by discrete subgroups. The discrete subgroups of $\mathbb{R}^n$ are the $\mathbb{Z}^k$ lattices of the vector subspaces of dimension k. Then any connected component of the regular level sets is $\mathbb{R}^{n-k} \times T^k$ for some k where $0 \leqslant k \leqslant n$.

Liouville Tori. The compact connected components of the regular level sets are thus tori T^n, and these tori are Lagrangian.

Remark II.2.1. We do not need to require that the flows are complete to obtain this result since on a compact connected component, the flow *is* complete.

Remark II.2.2. There are examples of integrable systems whose integrals are not proper. This is the case for intsance for the Hénon-Heiles system (see Section I.3.c) and for another celebrated system, the Toda system. See, for example, [**31, 33**] for the Hénon-Heiles system and [**7**] for the Toda system.

Exercises II.7, II.8, II.9.

II.2.b. Construction of "Action" Variables. We can describe the results of Section II.2.a by saying that the level sets of the first integrals have a natural affine structure (see Exercise II.8).

Remark II.2.3. This affine structure is defined by integrating the Hamiltonian vector fields of the first integrals. In particular, the solutions of the Hamiltonian system are linear in this affine structure. I will return to this tautology in Chapter IV.

[7]Recall that a local action of a Lie group G on a manifold W is *locally free* if the fundamental vector fields are independent at some point on W. See also Exercise II.7.

There is also a transverse affine structure defined in terms of the action variables, which we will now construct.

Fix a regular value q of f and assume $f^{-1}(q)$ has a compact connected component which we denote by F_q. We want to study what happens near F_q. Since the regular values of f form an open set, there is a neighborhood U of q in $\mathbb{R}^n$ consisting entirely of regularly values.

A Neighborhood of a Liouville Torus. Since F_q is compact, there is a relatively compact neighborhood V of F_q in the manifold W such that $f(\overline{V}) \subset U$. Let g be the restriction of f to V. The map $g : V \to U$ is a submersion that I will assume is proper. By shrinking V and U as necessary, we can trivialize the fibration g (this is the Ehresmann fibration theorem [**25, 26**]) and obtain a diffeomorphism

$$\tau : V \longrightarrow U \times F_q$$

such that $\tau^{-1}(\{u\} \times F_q) = g^{-1}(u)$, the fiber of u.

As we have remarked, F_q and the other fibers are tori. Using τ, we can also construct sections of g. Fix a point p in F_q. In $U \times F_q$, we have the obvious section $u \mapsto (u, p)$. We look at its image σ under τ^{-1} (see Figure 2)

$$\sigma(u) = \tau^{-1}(u, p).$$

This is a section of g.

Deformation of the Lattice. We now use this section σ to show that the tori $g^{-1}(u)$ can be continuously deformed, in the sense that $g^{-1}(u)$ is the quotient of $\mathbb{R}^n$ by the stabilizer[8] $P(\sigma(u))$, which depends only on u and, as a function of u, is $\mathcal{C}^\infty$.

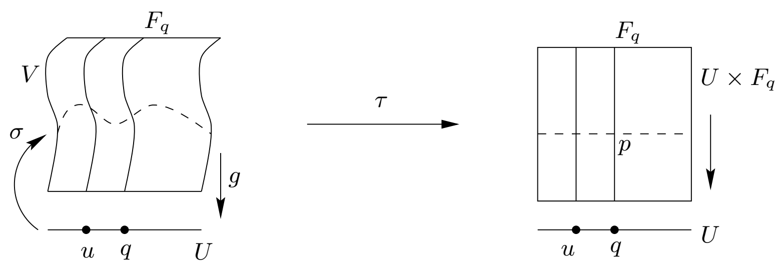


FIGURE 2

The idea is to apply the implicit function theorem to the equation

$$t \cdot \sigma(u) = \sigma(u).$$

We fix $t_0 \in P(p) \subset \mathbb{R}^n$. Let A be a neighborhood of t_0 in $\mathbb{R}^n$ and B a neighborhood of p in F_q, with A and B sufficiently small so that

$$t \in A \text{ and } u \in U \Longrightarrow \tau(t \cdot \sigma(u)) \in \{u\} \times B.$$

[8]In this proof, I use the notation $P(p)$ to denote the stabilizer of the point p, a lattice in $\mathbb{R}^n$, with the letter P standing for "periods." These lattices, which play an essential role in the theory and particularly in this chapter, were foreshadowed by the cream puffs of this book's opening quotation.

Since $F_q \cong \mathbb{R}^n/P(p)$, we can think of the open subset B of F_q as an open subset of $\mathbb{R}^n$.

We have thus defined a map

$$\begin{aligned}\Theta : A \times U &\longrightarrow B\\ (t,u) &\longmapsto \tau(t\cdot\sigma(u)) - p\end{aligned}$$

(with subtraction in the sense of the affine structure defined on the torus F_q). We then have

$$\begin{aligned}\Theta(t,u) = 0 &\Longleftrightarrow \tau(t\cdot\sigma(u)) = p\\ &\Longleftrightarrow t\cdot\sigma(u) = \sigma(u)\\ &\Longleftrightarrow t \in P(\sigma(u)).\end{aligned}$$

The partial differential $(d_t\Theta)_{(t,u)}$ of Θ with respect to the variable t is the composition of the differential $(d\tau)_{t\cdot\sigma(u)}$ of the diffeomorphism τ and the differential of $t \mapsto t\cdot\sigma(u)$. This has rank n since it sends the vectors $\dfrac{\partial}{\partial t_i}$ of the canonical basis to the independent vectors $X_i(\sigma(u))$.

Since the partial differential $(d_t\Theta)_{(t,u)}$ has rank n, the implicit function theorem gives us a function $T : U \to \mathbb{R}^n$ (by shrinking again the open set U if necessary) such that

$$\begin{cases}(t,u) \in A\times U \text{ and } t = T(u) \Leftrightarrow t\cdot\sigma(u) = \sigma(u)\\ T(p) = t_0.\end{cases}$$

Then, by letting t_0 run over a basis of the lattice $P(p)$, we obtain a basis of the lattice $P(\sigma(u))$ depending in a $\mathcal{C}^\infty$ fashion on u.

The Action of the Torus T^n. Let $(T_1(u),\dots,T_n(u))$ (for $u \in U$) be such a basis. This is a basis for the lattice $P(\sigma(u))$ with periods in $\mathbb{R}^n$. The corresponding vectors $Y_1(v),\dots,Y_n(v)$ in T_vV form a basis of the vectors tangent to the fiber (Figure 3).

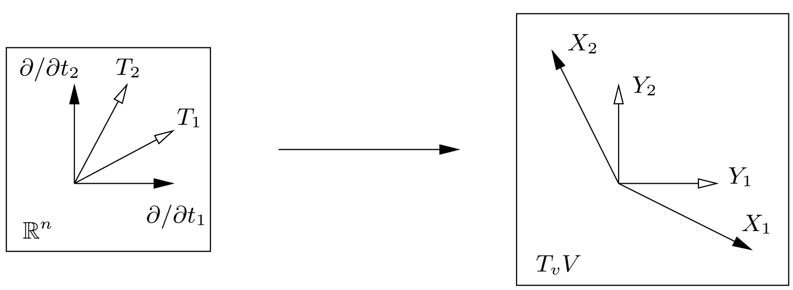


FIGURE 3

For those who like formulas: writing $T_i(u)$ in the canonical basis of $\mathbb{R}^n$ as

$$T_i(u) = \sum_{i=1}^{n} T_{i,j}(u)\frac{\partial}{\partial t_i},$$

we have

$$Y_i(v) = \sum_{j=1}^{n} T_{i,j}(f(v))X_j(v).$$

The vector fields $Y_1, \ldots, Y_n$ are, by definition, periodic with period 1. Moreover, if we let ψ_i^t be the flow of Y_i, we have that

$$\psi_i^t(v) = (tT_i(f(v))) \cdot v, .$$

This implies that

$$\begin{aligned} \psi_i^t \circ \psi_j^{t'}(v) &= tT_i(f(\psi_j^{t'}(v))) \cdot \psi_j^{t'}(v) \\ &= tT_i(f(v)) \cdot (t'T_j(f(v)) \cdot v) \\ &= tt'T_i(f(v))T_j(f(v)) \cdot v \end{aligned}$$

(by using the fact that $f(\psi_j^{t'}(v)) = f(v)$), and thus Y_i and Y_j commute.

We have thus defined an action of the torus $T^n = \mathbb{R}^n/\mathbb{Z}^n$ on V. We conclude that the Y_i are periodic with period 1, *independent of* u.

A Primitive of the Symplectic Form. We will construct the action coordinates by integrating a 1-form on the trajectories of Y_i. We now turn to constructing this 1-form. We can assume that U is a ball[9]; in which case V can be retracted to F_q. We have in particular, that $H^2(V;\mathbb{R}) \cong H^2(F_q;\mathbb{R})$.

Now, since F_q is isotropic, the class $[\omega]$ of the symplectic form ω in $H^2(F_q;\mathbb{R})$ is zero. Thus it is also zero in $H^2(V;\mathbb{R})$, i.e., the symplectic form is exact on V. Let λ be one of its primitives (i.e., λ satisfies $\omega = d\lambda$).

Definition of Action Variables. We set

$$J_i(v) = \int_0^1 \lambda(Y_i) \circ \psi_i^t(v)dt.$$

Here we are integrating the function of t obtained by evaluating the form λ at the point $\varphi_i^t(v)$ on the tangent vector Y_i. By setting $\gamma_i(t) = \psi_i^t(v)$, we define a loop parametrized by $[0,1]$ that is the orbit of v under the action of the circle corresponding to Y_i. Then the function $J_i(v)$ is none other than the integral of the form λ along the path γ_i

$$J_i(v) = \int_{\gamma_i} \lambda.$$

Proposition II.2.4. *The action of the torus T^n on V is Hamiltonian, and its moment map is*

$$J = (J_1, \ldots, J_n) : V \longrightarrow \mathbb{R}^n.$$

Proof. We will prove this by applying Proposition II.1.2 (or, more precisely, the variation of it given in Remark II.1.3) to the exact symplectic manifold V, for the action of each of the factors of T^n. The only thing we need to check is that

[9] See, however, Section II.2.d.

each of the fundamental vector fields Y_i satisfies $i_{Y_i}\mathcal{L}_{Y_i} = 0$. By Cartan's formula, we have

$$\begin{aligned} i_{Y_i}\mathcal{L}_{Y_i}\omega &= i_{Y_i} d i_{Y_i}\omega \\ &= i_{Y_i} d\left(\sum T_{i,j} df_j\right) \\ &= i_{Y_i}\left(\sum dT_{i,j} \wedge df_j\right). \end{aligned}$$

But Y_i is tangent to the fibers of f and $T_{i,j}$ is constant on these fibers. Thus Y_i lies in the kernel of $dT_{i,j}$ and in that of df_j. □

Some Consequences. Notice that of course $J(v)$ depends only on $f(v)$: it is a moment map, it is constant on the orbits. This could also be checked directly.

Notice also that the map J has rank n (and so the action of the torus is effective). Indeed, we have

$$(dJ)_v(Z) = \big(\omega_v(Z, Y_1), \ldots, \omega_v(Z, Y_n)\big),$$

and this linear map has rank n since the vectors Y_i are independent and the form ω is nondegenerate.

Exercise II.10.

II.2.c. Construction of "Angle" Variables. We summarize the previous episodes. We have a symplectic fiber bundle

$$J : V \longrightarrow U$$

whose fibers are Lagrangian tori and an action of T^n with moment map J.

As above, we can utilize a section σ of the fiber bundle and a diffeomorphism τ mapping it to $U \times F_q$. Since F_q is a torus $\mathbb{R}^n/\mathbb{Z}^n$, we have coordinates $(\theta_1, \ldots, \theta_n)$ on F_q and thus coordinates $(J_1, \ldots, J_n, \theta_1, \ldots, \theta_n)$ on V.

In these coordinates, the symplectic form can be written as

$$\omega = \sum dJ_i \wedge d\theta_i + \sum a_{i,j} dJ_i \wedge dJ_j$$

(where no terms including $d\theta_i \wedge d\theta_j$ appear since the fibers are Lagrangian).

We set $\eta = \sum a_{i,j} dJ_i \wedge dJ_j$. *A priori*, the $a_{i,j}$ could depend on the θ_k; however, they do not since the fact that ω is closed implies that η is as well, and thus η does not depend on the θ_k. More intrinsically, $\eta = \sigma^\star\omega$ is closed since ω is.

The form η is thus a closed form on U. We have already supposed (and utilized the fact) that U is contractible. Thus η is exact, so $\eta = dA$ for some form A on U of the form

$$A = \sum A_i dJ_i.$$

The A_i are functions on U. We use them to modify the section σ as follows: set

$$\varphi_i = \theta_i - A_i$$

(the new "angles") or

$$s(J_1, \ldots, J_n) = (J_1, \ldots, J_n, \theta_1 - A_1, \ldots, \theta_n - A_n),$$

or abbreviating,

$$s(J_1, \ldots, J_n) = \sigma(J_1, \ldots, J_n) - (A_1, \ldots, A_n).$$

Finally,

$$s^\star\omega = \sigma^\star\omega - dA = 0,$$

and thus

$$\omega = \sum dJ_i \wedge d\varphi_i.$$

REMARK II.2.5. We can modify J by adding a constant and modify φ by adding a function of J (only), and still obtain "action–angle" variables.

The Hamiltonian System in Action–Angle Variables. It has a particularly simple form. First of all, the Hamiltonian H is in the algebra generated by the first integrals, and thus is a function of $J_1, \ldots, J_n$ only. We write this as

$$H(x) = h(J_1, \ldots, J_n).$$

Next, since the coordinates $(J_1, \ldots, J_n, \varphi_1, \ldots, \varphi_n)$ are symplectic, the Hamiltonian system can be written in these coordinates as

$$\dot{J}_i = -\frac{\partial h}{\partial \varphi_i} = 0 \text{ and } \dot{\varphi}_i = \frac{\partial h}{\partial J_i} = a_i(J_1, \ldots, J_n).$$

Using this form, it is easy to check that the solutions are angles φ_i depending linearly on t on the level sets of J. It is clear that such a system is integrable by quadratures, hence the terminology "integrable system".

II.2.d. The Global Problem. One interesting question is to know "how far" we can extend these action–angle variables. To do this, we return to the proof and examine the various shrinking operations we applied to the open set U. For simplicity, we suppose that f is proper. We realize successively that it suffices for the open set U to satisfy the following properties:

- First of all, U is connected and contained in the set of regular values of f.
- Over U, the bundle $g : V \to U$ is trivial. This is certainly the case if U is contractible[10].
- The implicit function theorem can be applied globally: since the derivative of the map Θ is of rank n as soon as the X_i are independent and thus over the open set of regular values of f.
- We needed to use a primitive of the symplectic form on V. To do this, we used the isomorphism

  $$H^2(U \times F_q; \mathbb{R}) \cong H^2(F_q; \mathbb{R})$$

  which, according to the Künneth formula, holds whenever

  $$H^1(U; \mathbb{R}) = 0 \text{ and } H^2(U; \mathbb{R}) = 0,$$

  and in particular whenever U is contractible.

We have shown the following more global result.

[10] However, note that since the torus bundle $V \to U$ is not *a priori* principal, it is not necessarily trivial even when $H^2(U; \mathbb{Z})$ is zero. See the examples II.3.c and II.3.d.

PROPOSITION II.2.6. *Let U be a contractible open subset of $\mathbb{R}^n$ consisting of regular values of f and let V be a connected component of $f^{-1}(U)$ such that $f_{|V} : V \to U$ has a compact fiber. Then there are global action–angle coordinates on V.*

The foundational article by Duistermaat [**24**] includes a more general discussion where it is shown that the absence of monodromy, in particular, the simply-connectedness of U and the triviality of $H^2(U;\mathbb{Z})$, implies the existence of global action–angle variables.

REMARK II.2.7. If one of the components of f, for example the Hamiltonian, is proper, then the compactness hypothesis is automatically satisfied.

EXAMPLE II.2.8. In the case of geodesics on quadrics, one can show (see Section IV.3.c and [**6**]) that the connected components of the set of regular values are simplices, and in particular contractible, yielding action–angle coordinates over each of the components (see Section IV.3.d and [**10**]). However, note that the inverse images of these components are never connected[11].

EXAMPLE II.2.9 (Counterexamples). There are two classical counterexamples, the spherical pendulum and the symmetric top (see [**24**] or [**22**]). In both cases, f has an isolated singular point in the middle of the open set of its regular points (see Exercises II.13 and IV.12, as well as Figure 7). The open set U in question is not simply connected, thus failing to satisfy the hypotheses of Proposition II.2.6. One can show that the fibration $V \to U$ is not trivial, and thus there are *no* global action–angle coordinates for these two systems. See [**24**] and [**22**] (and also [**11**], for a different point of view).

II.2.e. The Real Analytic Case. The proof we used here does not work well in the real analytic case. Indeed, the proof used the Ehresmann fibration theorem, which does not apply to the real analytic case: there are real analytic families that are not (analytically) locally trivial. Nonetheless, the real analytic version of the Arnold–Liouville theorem does hold. Instead of using the fact that the fibration is locally trivial, followed by the section σ and the implicit function theorem, we can directly show that the level sets form an analytic family of tori, with the lattice depending analytically on u (see, for example, [**44**]). Then the rest of the proof follows in the same way as for the complex case.

II.3. Examples

We now return to the examples described in Chapter I to study what the Arnold–Liouville theorem says about these cases, mainly from a topological point of view.

[11] An assertion that is easy to check: geodesics are *oriented* curves, and this helps the regular level sets to have an even number of connected components. For a precise counting of Liouville tori, see [**6**].

II.3.a. The Harmonic Oscillator. All the values of $H = \frac{1}{2}p^2 + \frac{1}{2}q^2$ except 0 are regular. The Liouville tori are the circles centered at the origin. The angle variable is the angle θ of polar coordinates, and the action coordinate is $\frac{1}{2}r^2$ (that is, H).

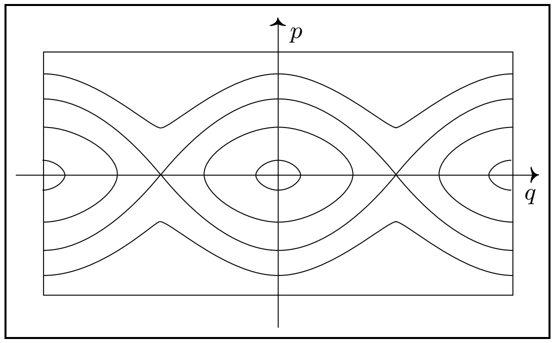


FIGURE 4. The Simple Pendulum

II.3.b. The simple pendulum. Figure 4 shows the level sets of the Hamiltonian H, which are the supports for the solutions to the differential system associated to

$$H = \frac{1}{2}p^2 - \cos q.$$

The ovals such as the ones shown in Figure 4 correspond to small (almost periodic) movements of the pendulum.

Although this 2-dimensional example is quite simple, it is still worthwhile to make sure one understands it completely (see Exercise II.12).

II.3.c. The Spherical Pendulum. The regular values and the Liouville tori for the spherical pendulum are derived in great detail and down-to-earth way in [**22**], where one can also find a long discussion of the existence, or rather the non-existence, of global action-angle variables. It is thus not unfair to propose these results as an exercise, which, in fact, appears twice in this book: first using a direct method (Exercise II.13) and then by a more elegant, calculation-free one using the methods of algebraic geometry presented in Chapter IV (Exercise IV.12).

II.3.d. The Top. We use notation from Sections I.2.g and I.3.g. The phase space is the manifold

$$W_a = \left\{ (\Gamma, M) \in \mathbb{R}^3 \times \mathbb{R}^3 \mid \|\Gamma\|^2 = 1 \text{ and } \Gamma \cdot M = a \right\},$$

and the Hamiltonian is $H = \frac{1}{2}M \cdot \Omega + \Gamma \cdot L$. Let

$$H'(\Gamma, M) = \frac{1}{2}\|M\|^2 + \Gamma \cdot L.$$

It is easy to show[12] that H' is a first integral for the motion of the top, and that the functions (H', K) and (H, K) have the same rank.

[12]The gaps in this proof are enumerated in Exercise II.15.

We determine now the regular values of the map (H', K). The rank of this map at the point (M, Γ) is that of the pair of linear forms

$$\begin{cases} (dH')_{(\Gamma,M)}(\gamma,\mu) = L\cdot\gamma + M\cdot\mu \\ (dK)_{(\Gamma,M)}(\gamma,\mu) = L\cdot\mu \end{cases}$$

on the vector subspace

$$\begin{cases} \Gamma\cdot\gamma = 0 \\ M\cdot\gamma + \Gamma\cdot\mu = 0. \end{cases}$$

Thus we need to determine the rank of the matrix $\begin{pmatrix} L & 0 & \Gamma & M \\ M & L & 0 & \Gamma \end{pmatrix}$. Saying that this rank is not maximal is equivalent to having

$$\begin{cases} xL & + & & + & z\Gamma & + & tM & = & 0 \\ xM & + & yL & + & & + & t\Gamma & = & 0 \end{cases}$$

for some x, y, z, and t not all zero.

- If these two equations are independent, then the vectors L, Γ, and M are collinear; thus $L = \varepsilon\Gamma$ (with $\varepsilon = \pm 1$ corresponding to the case where the axis of the top is vertical) and $M = a\Gamma$. At these points, (H', K) has rank zero and takes the values $H' = \frac{1}{2}a^2 + \varepsilon$ and $K = \varepsilon a$.
- If they are related with, say, $t \neq 0$, there exists $\chi \neq 0$ such that $\chi x = t$, $\chi y = x$, $\chi t = z$. Setting $t = 1$, we see that there exists a real number λ such that $\lambda^{-1}\Gamma + M + \lambda L = 0$.

The regular values of (H', K) are those for which all the roots of the polynomial

$$Q(\lambda) = \left\|\Gamma + M\lambda + L\lambda^2\right\|^2$$

are simple.

In the case where $a < 2$, the image of the map (H', K) is very similar to that of the map (H, K) for the spherical pendulum (see Figure 7). In particular, the set of regular values is not simply connected. See [**22**] for a proof that the fibration by the Liouville tori also has no global section. We now determine these Liouville tori.

Proposition II.3.1. *The regular level sets of* (H', K) *are connected.*

Proof. The results of Exercise II.15 imply that the set of regular values of (H', K) is connected (although not necessarily simply connected). It thus suffices to show that certain regular level sets are connected. I will show this is the case for those those for which K is submersive. Let k be a regular value of

$$K : W_a \longrightarrow \mathbb{R}.$$

The group of rotations about the axis L of the top preserves $K^{-1}(k)$ by definition: K is a moment map for the action of the circle on W_a. Thus we can consider the "reduced symplectic manifold" $W_{a,k} = W_a/S^1$ (see Exercise II.4). Since W_a has dimension 4, $W_{a,k}$ is a surface. Since H' commutes with K, we can pass to the quotient and define a Hamiltonian (which we will still denote by H')

$$H' : W_{a,k} \longrightarrow \mathbb{R}.$$

The level set corresponding to the value (h', k) of (H', K) is a circle bundle on the level set h' of H' in $W_{a,k}$. To establish the proposition, it suffices to show that this is connected.

Consider the orbit of a point (Γ, M) of $K^{-1}(k)$. In the orthonormal basis used to diagonalize the inertia matrix $\mathcal{J}$ (see Section I.3.g), set

$$\Gamma = \begin{pmatrix}\gamma_1\\ \gamma_2\\ \gamma_3\end{pmatrix} \text{ and } M = \begin{pmatrix}u\\ v\\ k\end{pmatrix}$$

or, better yet,

$$\Gamma = \begin{pmatrix}\gamma_1 + i\gamma_2\\ \gamma_3\end{pmatrix} \text{ and } M = \begin{pmatrix}u + iv\\ k\end{pmatrix}.$$

The circle acts as follows

$$\theta \cdot (\Gamma, M) = \left(\begin{pmatrix}e^{i\theta}(\gamma_1 + i\gamma_2)\\ \gamma_3\end{pmatrix}, \begin{pmatrix}e^{i\theta}(u + iv)\\ k\end{pmatrix}\right).$$

Note also that if $M \cdot \Gamma = a$, we have $k = \pm a$, which is a critical value of K as noted above. Since we assume k is a regular value, we have $\gamma_1 + i\gamma_2 \neq 0$, and each orbit has a unique representative with $\gamma_1 > 0$ and $\gamma_2 = 0$. We fix a γ_3 in $(-1, 1)$. Then $\gamma_1 = \sqrt{1 - \gamma_3^2}$ and $\Gamma \cdot M = a$ implies $u = (a - k\gamma_3)/\gamma_1$with no restriction on v. We have thus defined an embedding of $\mathcal{O}_{c,k}$ into $(-1, 1) \times \mathbb{R}$

$$(\Gamma, M) \longmapsto (\gamma_3, v).$$

Notice that, in $(-1, 1) \times \mathbb{R}$, the level sets of the reduced Hamiltonian H' are

$$H'(\gamma_3, v) = \frac{1}{2}\left[\left(\frac{a - k\gamma_3}{\gamma_1}\right)^2 + v^2 + k^2\right] + \gamma_3.$$

On the level set corresponding to h', we see that

$$v^2 = (1 - \gamma_3^2)(h' - k^2 - 2\gamma_3) - (a - k\gamma_3)^2,$$

whose intersection with $(-1, 1) \times \mathbb{R}$ is connected. □

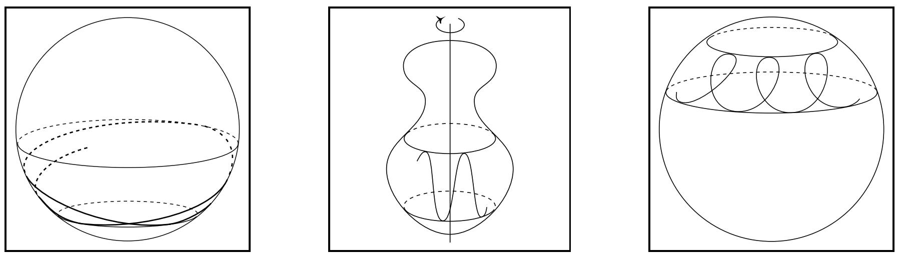
II.3.e. Geodesics on Surfaces of Revolution. I will present the example of the geodesic flow on surfaces of revolution in Exercise II.14. Recall that we showed this system to be completely integrable in Section I.4.a. This exercise aims to convince the readers that the picture illustrating the cover of [**30**] is actually a Liouville torus of the geodesic flow on a surface of revolution.

II.3.f. Geodesics on the Ellipsoid. I will describe the regular values and determine action variables for geodesics on ellipsoids in Chapter IV, by using methods of algebraic geometry. These methods can also be used to count the Liouville tori (see [**6**]).

Exercises II.11, II.12, II.13, II.14, II.15.

II.3.g. Annuli and Tori, a Remark. The figure representing the geodesics on a surface of revolution appearing on the cover of [**30**] closely resembles figures representing the motion of the free end of a top's axis (see [**2**] or [**8**], for example), the motion of the weight of a spherical pendulum, or the geodesics on an ellipsoid (see [**40**] or [**6**]).

In each of these cases, we see a solution of an integrable system in a configuration space: the geodesic on the surface or the movement of a point on a sphere. A few such trajectories are shown in Figure 5. Note that in each of these cases, we have an annulus (topologically, a cylinder) and that the solution in question oscillates between the two limiting "circles" of the annulus.



(a) the spherical pendulum (b) a surface of revolution (c) the top

FIGURE 5. Annuli...

What each of these figures represents is a Liouville torus (which lives in a 4-dimensional phase space) projected onto a 2-dimensional space. This is illustrated in Figure 6.

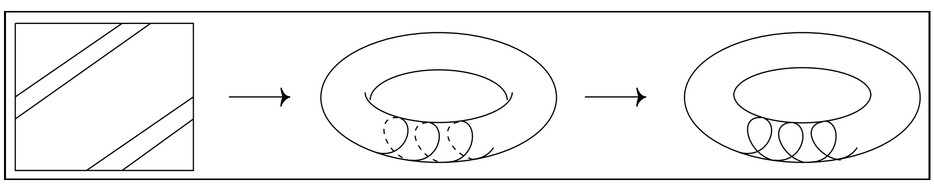


FIGURE 6. ... or tori?

Exercises

EXERCISE II.1. Show that the action of S^1 on the unit sphere in $\mathbb{R}^3$ by rotations about the z-axis is Hamiltonian with the moment map equal to the height function z. (See Section I.2.b.)

EXERCISE II.2. The torus $T^2 = \mathbb{R}^2/\mathbb{Z}^2$ inherits a symplectic form from $\mathbb{R}^2$ (see Exercise I.8). Check that the formula

$$t \cdot (x, y) = (x + t, y)$$

defines a symplectic action of $S^1 = \mathbb{R}/\mathbb{Z}$ on T^2. Is this action Hamiltonian?

EXERCISE II.3. The symplectic manifold (W, ω) is equipped with a symplectic action of a torus T. Let X and Y be two fundamental vector fields of this action. Show that the function

$$x \longmapsto \omega_x(X(x), Y(x))$$

is constant. Deduce that this action equips the Lie algebra $\mathfrak{t}$ of T with a skew-symmetric bilinear form. Under what conditions is the action Hamiltonian?

EXERCISE II.4 (Symplectic Reduction—Continued). Let F be a periodic Hamiltonian on a symplectic manifold W, and let f be a regular value of F. Check that S^1 acts on the submanifold $F^{-1}(f)$ and that this action is locally free.

Now suppose that the action is (globally) free. Show that the symplectic form on W defines a symplectic form on the quotient manifold $F^{-1}(f)/S^1$ (see Exercise I.4).

Let G be another function on W commuting with F. Show that G defines a function on the reduced symplectic manifold $F^{-1}(f)/S^1$.

EXERCISE II.5 (Complex Projective Space). Using the method developed in Exercise II.4, define a symplectic structure on the complex projective space

$$\mathbb{P}^n(\mathbb{C}) = \left(\mathbb{C}^{n+1} - \{0\}\right)/\mathbb{C}^\star = S^{2n+1}/S^1.$$

The torus T^{n+1} acts on $\mathbb{R}^{n+1} \times \mathbb{R}^{n+1} = \mathbb{C}^{n+1}$ as term-by-term multiplication

$$(t_1, \ldots, t_{n+1}) \cdot (z_1, \ldots, z_{n+1}) = (t_1 z_1, \ldots, t_{n+1} z_{n+1}).$$

Show that this action is Hamiltonian. What is its moment map?

Continuing to use the method of Exercise II.4, deduce from this a Hamiltonian action of the torus T^{n+1} on the projective space $\mathbb{P}^n(\mathbb{C})$. Is it effective? Show that the image of $\mathbb{P}^n(\mathbb{C})$ under the moment map of this action is an n-dimensional simplex[13].

EXERCISE II.6 (Periodic Hamiltonians are Integrable). Recall that[14], if f is a regular value of a function F on a compact manifold W, a neighborhood of the regular level $F^{-1}(f)$ in W is diffeomorphic to

$$F_1^{-1}(f) \times (f - \varepsilon, f + \varepsilon).$$

Assume now that F is a periodic Hamiltonian. Consider a trajectory Γ of X_F contained in a regular level $F^{-1}(f)$. Assume that that the circle S^1 acts freely on the level set $F^{-1}(f)$. By utilizing Exercise II.4 and $n-1$ independent functions in involution on a neighborhood of the image of Γ in $F^{-1}(f)/S^1$, show that F is integrable on a neighborhood of Γ in W.

EXERCISE II.7. Show that the action of the Lie group G on the manifold W is locally free if and only if there exist a point x in W and a neighborhood U of 1 in G such that $G_x \cap U = \varnothing$.

[13]This is a special case of a famous theorem due to Atiyah, Guillemin, and Sternberg [**5, 38**], which states that the image of the moment map of the action of a torus on a compact connected symplectic manifold is a convex polytope.

[14]This can be shown by integrating the gradient of F for a Riemannian metric. Notice that, in this decomposition, F can be written as $F(x, t) = t$.

Exercise II.8 (Affine Structure of Fibers). Show that the fibers of a Lagrangian fibration have a canonical affine structure.

Exercise II.9. Show that every compact connected component of a fiber of a Lagrangian fibration is isomorphic to a torus (as a manifold with an affine structure).

Exercise II.10 (Transverse Affine Structure). Let $g : V \to U$ be a Lagrangian fibration.

- Let q be a point in U. Show that the cotangent space $T_q^\star U$ acts on $g^{-1}(q)$.
- Suppose that the fibers of g are compact and connected. Show that these fibers can be identified with the quotient of $T_q^\star U$ by a lattice P_q and that there exist 1-forms $\alpha_1, \dots, \alpha_n$ in a neighborhood of q such that $(\alpha_1(q), \dots, \alpha_n(q))$ is a basis of P_q.
- Show that these forms are closed.
- Let $J_1, \dots, J_n$ be local primitives of the α_i on a neighborhood U' of q. Show that
$$(J_1, \dots, J_n) : U' \longrightarrow \mathbb{R}^n$$
is a local chart for the base U of the fibration.
- We have thus constructed an atlas on U. Show that the changes of coordinates between charts lie in the subgroup of the affine group consisting of all the affine transformations with linear component in $\mathrm{GL}(n; \mathbb{Z})$.

Exercise II.11. Determine the action–angle variables for the Hamiltonian system on $\mathbb{R}^2$ corresponding to $H = \frac{1}{2}p^2 + \frac{1}{2}aq^2$.

Exercise II.12. Determine the regular values of H, the Liouville tori, and an action variable for the spherical pendulum, and give a geometric description of this action variable.

Exercise II.13 (Regular Values for the Spherical Pendulum). We use the notation from Section I.2.f. Show that the horizontal motions of the spherical pendulum are critical (a result due to Huygens, see [**24**]). Write the vector q in the form
$$q = \frac{1}{\alpha^2}\Gamma + q'$$
for some constant α such that $\alpha^2 > 1$ and a vector q' orthogonal to Γ, and consider the vector
$$p = \alpha q' \times \Gamma,$$
which itself is horizontal as well. Show that
$$X_H(q, p) = -\alpha X_K(q, p),$$
and deduce that the values
$$\begin{cases} H = \dfrac{1}{2}\alpha^2 - \dfrac{3}{2\alpha^2} \\ K = -\alpha + \dfrac{1}{\alpha^3} \end{cases} \qquad \text{for } \alpha^2 > 1$$
are critical.

Next show that the motions where the pendulum remains vertical are critical. Consider $(q, p) = (\pm\Gamma, 0)$. Show that $X_K(q, p) = 0$. Deduce from this that the values $(H, K) = (\pm 1, 0)$ are critical.

Show that the set of regular values of the map $(H, K) : TS^2 \to \mathbb{R}^2$ looks like the non-shaded portion of Figure 7 (note that this open set is *not* simply connected[15]). Show that all the regular level sets of (H, K) are connected and that the level set $(H, K) = (1, 0)$ is a "pinched" torus.

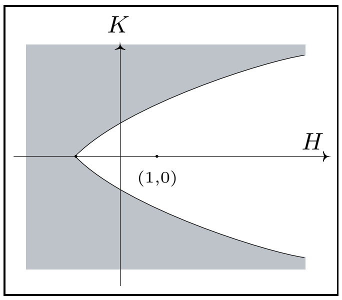


FIGURE 7

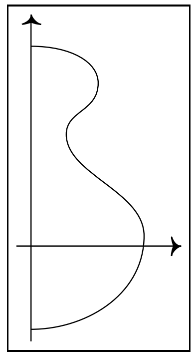


FIGURE 8

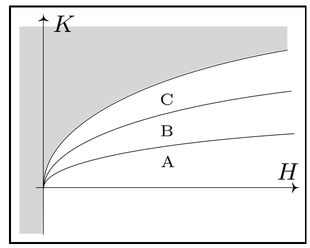


FIGURE 9

EXERCISE II.14 (Liouville Tori for Surfaces of Revolution). Consider the surface of revolution generated by the curve shown in Figure 8 that we assume is parametrized by arc length (using the notation of Section I.4.a). Determine the regular values and show that the image of the map (H, K) in $\mathbb{R}^2$ resembles the non-shaded portion of Figure 9. What do the parabolas in this figure correspond to? Show that the regular level sets consist of

- two[16] tori in parts A and C, and
- four tori in part B.

What can be said about the level sets corresponding to the points on the parabolas?

Show that the geodesics that are not meridians are condemned to oscillate between two parallels of the surface. What can be said about these parallels? Give an interpretation in the case of geodesics on the sphere seen as a surface of revolution about the z-axis.

EXERCISE II.15. Fill in the gaps in the study of the top carried out in Section II.3.d):

- Show that H' is a first integral for the motion of the top and that the maps (H', K) and (H, K) have the same rank.
- Let $Q(\lambda) = \left\|\Gamma + M\lambda + L\lambda^2\right\|^2$ (a polynomial of degree 4 in λ). Show that (h', k) is a regular value of (H', K) if and only if the polynomial Q has

[15]See Exercise IV.12 for a calculation-free proof.

[16]See Footnote 11.

no multiple roots. Determine the set of (regular and critical) values of (H', K) (see figures in [**8**]).

- Draw (the real part of) the elliptic curve

$$y^2 = (1 - x^2)(h' - k^2 - 2x) - (a - kx)^2$$

and check that it has two connected components, one of which is compact and projects into the open interval $x \in (-1, 1)$.

CHAPTER III

Integrability and Galois Groups

The Arnold–Liouville Theorem, Revisited Through Differential Galois Theory

The Arnold–Liouville theorem affirms that the solutions close to a sufficiently general solution of an integrable system closely "resemble" that solution. In this chapter, we will study (infinitesimally) close solutions by using the algebraic theory of linear differential equations.

Following a "variational" tradition going back to Poincaré [**69**], we can associate a linear differential equation, called the variational equation, to each trajectory of a vector field X on a manifold W. Ziglin [**84**] has shown and efficiently used the fact that any first integral of X allows us to define a first integral of the variational equation that is invariant under the Galois group.

In the specific case of a generic trajectory of an integrable system, it is easy to show that the Galois group is "small" (as an algebraic group!). It is connected, Abelian, and often trivial. A theorem of Morales and Ramis [**62, 61**] (see also [**12**]) states that along any trajectory, its identity component must be an Abelian group.

This theorem already has many applications: see [**62, 63, 46, 61**] and, since the French edition of this book appeared [**13, 14, 57, 19**]. It can serve as a criterion for non-integrability, helping to answer questions such as, "Given a Hamiltonian system, how can we show that it is (or, rather, that it is not) completely integrable?" or, more naively, "If I could not find any first integrals, is it because I am clumsy, or because there actually are *not* any?" Although it is much easier to prove, the Morales–Ramis theorem implies a classical theorem of Ziglin [**84**]. It is strictly stronger[1] in the sense that it allows us to show that certain systems, for which Ziglin's method fails, are not integrable .

In this chapter, I explain some of these properties of the Galois group while attempting to minimize the calculations and separating what belongs to the general theory of differential equations from what belongs specifically to the symplectic framework. At the end of the chapter I give a relatively simple example of application[2].

[1] We know this a *priori*, because the Galois group contains more non-commutativity sources than the monodromy group used by Ziglin. However, the hypotheses used by Ziglin are less restrictive, since they do not assume that the first integrals commute.

[2] This is probably the simplest example possible; surprisingly, it is not found in [**62**] or [**61**].

III.1. Variational Equations, Galois Groups, and First Integrals

I begin by reviewing how to associate a linear differential equation, the variational equation, to a trajectory of a vector field on a manifold. Then, I show how the first integrals of the vector field yield first integrals of the linear differential equation that are invariant under the Galois group of this equation.

III.1.a. The Trajectory Γ. Let X be a vector field on a complex analytic manifold W. The time t is also a complex variable[3].

Consider a trajectory Γ of X that is not reduced to a point (that is, a non-stationary trajectory). This is a complex curve, a Riemann surface, but it is not necessarily *embedded* in W. More precisely, for any point x_0 of W, there exists an open set U of $\mathbb{C}$ and a neighborhood V of x_0 such that the flow of X is defined on $U \times V$

$$\begin{aligned} U \times V &\longrightarrow W \\ (t, x) &\longmapsto \varphi_t(x). \end{aligned}$$

We define a topology on W using as a basis the sets $\{\varphi_t(x)\}_{t \in U}$. Then the trajectory passing through x_0 is the connected component of x_0 for this topology. This is the curve we denote by Γ. It has two important properties:

- it is equipped with a coordinate, the time t,
- it is equipped with an injective embedding $i : \Gamma \to W$.

REMARK III.1.1. The curve Γ is globally parametrized by t in an open subset of $\mathbb{C}$, but this parametrization is not injective in general. In fact, it is often an infinite covering: think of the case of a differential equation whose solutions are transcendental functions, for instance of the vector field $X(x) = x$ on $\mathbb{C}$.

III.1.b. The Variational Equation. Traditionally (see [**69**]), one considers the variational equation as the one describing the solutions that are "infinitesimally close" to a given solution. If $x(t)$ and $y(t)$ are solutions that are close, then for a suitable coordinate chart, we can write

$$y(t) = x(t) + Y(t).$$

Then, to first order, we have

$$\frac{dY}{dt} = \frac{dy}{dt} - \frac{dx}{dt} = X(y(t)) - X(x(t)) = (dX)_{x(t)}(Y(t)).$$

This *linear* differential equation in Y is the variational equation that I will now define in a more intrinsic way.

[3]For applications to actual differential systems, we will content ourselves with real analytic systems that we will hasten to complexify. This is a somewhat delicate operation, to be performed carefully, that we do not dwell on in these notes.

The Operator D. Consider the vector bundle $i^\star TW \to \Gamma$. A section of this bundle is a field of vectors that are tangent to W along Γ. For example, $X = \dfrac{d}{dt}$ is one such section.

LEMMA III.1.2. *The Lie derivative $\mathcal{L}_X$ defines an operator D on the local sections of $i^\star TW$ satisfying*

$$D(fY) = \dot{f}Y + fDY$$

for any function f on Γ.

PROOF. We define DY by

- extending Y to $\widetilde{Y}$ on a neighborhood of Γ (or of a point on Γ, since everything is local);
- calculating $\mathcal{L}_X\widetilde{Y}$ $(= [X, \widetilde{Y}])$;
- restricting this vector field to Γ.

We check that the result does not depend on the choice of the extension $\widetilde{Y}$ of Y by observing that if Z is identically zero on its restriction to Γ, we have that $\mathcal{L}_X Z$ is zero on its restriction to Γ (in other words, X differentiates only in the direction of Γ). This is what Lemma III.1.3 below asserts.

The "linearity" with respect to functions is an immediate consequence of the formula

$$\mathcal{L}_X(fY) = (X \cdot f)Y + f\mathcal{L}_X Y. \qquad \square$$

LEMMA III.1.3. *If Z is a vector field on W defined in a neighborhood of Γ and identically zero on Γ, then $\mathcal{L}_X Z$ is zero at every point of Γ.*

PROOF. To define $\mathcal{L}_X Z$, we use the flow φ_t on X

$$\mathcal{L}_X Z = \frac{d}{dt}\left((\varphi_t)_\star Z\right)_{|t=0}.$$

The vector $Z(x)$ is zero at every point x of Γ, and thus the same is true for $(\varphi_t)_\star Z(x)$. Hence $\mathcal{L}_X Z$ is zero at every point of Γ. $\square$

REMARK III.1.4. In more pedantic terms, $\nabla = D \otimes dt$ is a connection on the bundle $i^\star TW$. If we allow ourselves to add the zeros of X or points at infinity to Γ, the form dt becomes a singular form on Γ, making it useful to think in terms of a (singular) connection ∇ rather than the operator D. The "meromorphic" case will be particularly interesting. In this case, we will speak of an *extended* trajectory. We will need this for some later applications, see Section III.5.

The Variational Equation. The *variational equation* is the linear equation

$$DY = 0 \qquad \text{or} \qquad \nabla Y = 0.$$

Note that X itself is obviously a solution to the variational equation.

For those who like coordinates, we write this equation in matrix form. Let $(e_1(t), \ldots, e_m(t))$ be a basis for the local sections of the vector bundle $E \to \Gamma$ and write a section Y of E as $Y(t) = \sum y_i(t)e_i(t)$; then

$$\begin{aligned} DY &= \sum_{i=1}^{m} (\dot{y}_i e_i + y_i De_i) \\ &= \sum_{i=1}^{m} \Big(\dot{y}_i e_i + y_i \sum_{j=1}^{m} a_{i,j} e_j\Big) \\ &= \sum_{i=1}^{m} \Big(\dot{y}_i + \sum_{k=1}^{m} a_{k,i} y_k\Big) e_i. \end{aligned}$$

If y is the column vector of the y_i and $A(t)$ the matrix with entries $A(t)_{i,j} = a_{j,i}(t)$, we then have

$$\dot{y} + A(t)y = 0.$$

Even more locally, consider an open set W of $\mathbb{C}^m$. In this case we can treat the vector field X as a map

$$\begin{aligned} W &\longrightarrow \mathbb{C}^m \\ x &\longmapsto X(x). \end{aligned}$$

Restricting the canonical basis $\left(\dfrac{\partial}{\partial x_1}, \dots, \dfrac{\partial}{\partial x_m}\right)$ of $\mathbb{C}^m$ to the trajectory, using it as our basis $(e_1(t), \dots, e_m(t))$, we have

$$\left[X, \frac{\partial}{\partial x_i}\right] = -\frac{\partial X}{\partial x_i},$$

so that

$$A(t) = -(dX)_{x(t)}.$$

Then the variational equation is the *linearization* $\dot{y} = (dX)_{x(t)}y$ of the original differential equation. This is the variational equation "à la Poincaré" alluded to at the beginning of the chapter.

EXAMPLES III.1.5.

(1) The Case of a Linear System. As above, let W be an open subset of $\mathbb{C}^m$. In addition, suppose that the vector field X is linear i.e., it is a matrix $X \in M_m(\mathbb{C})$ with $X(x) = X \cdot x$. In this case, the variational equation along a solution is just the equation itself.

(2) The Case of the Hénon–Heiles Hamiltonian. We use the notation of Section I.2.d. The Hamiltonian system is

$$\begin{cases} \dot{q}_1 = p_1 \\ \dot{q}_2 = p_2 \\ \dot{p}_1 = q_2^2 + \lambda q_1^2 \\ \dot{p}_2 = 2(A + q_1)q_2. \end{cases}$$

We have already noted that this system has solutions in the plane $q_2 = p_2 = 0$. Along a solution $(q_1(t), 0, p_1(t), 0)$, the variational equation is the system

$$\begin{cases} \dot{Q}_1 = P_1 \\ \dot{Q}_2 = P_2 \\ \dot{P}_1 = 2\lambda q_1 Q_1 \\ \dot{P}_2 = 2AQ_2 + 2q_1 Q_2. \end{cases}$$

(3) The Case of a Geodesic Flow. Consider a Riemannian manifold V and a geodesic $c(t)$. This is the image in V of an integral curve of the geodesic flow considered as a flow on $T^\star V$ (see Section I.4). The solutions $Y(t)$ of the variational equation on $T^\star V$ can be projected in V as tangent fields along the geodesic. These are the famous "Jacobi fields", which are solutions to the linear differential equation

$$\ddot{J}(t) + R(\dot{c}(t), J(t))\dot{c}(t) = 0$$

(see, for example, [**75**] and [**30**]).

Exercises III.1, III.2.

III.1.c. Initial Forms. Consider now a function f on W. Let x be a point in W. Recall the following lemma.

LEMMA III.1.6. *If* $(df)_x = 0, \ldots, (d^{k-1}f)_x = 0$*, then* $(d^k f)_x$ *is well-defined and is a homogeneous polynomial of degree* k

$$(d^k f)_x : T_x W \longrightarrow \mathbb{C}.$$

PROOF. Let U be a neighborhood of x in W and $\varphi : U \to \mathbb{C}^m$ a coordinate chart with $\varphi(x) = 0$. For $Y \in T_x W$, we set

$$(d^k f)_x(Y) = d^k(f \circ \varphi^{-1})_0(T_x\varphi(Y)).$$

It is easy to check that the result does not depend on the coordinate chart chosen when the preceding derivatives of f with respect to x are zero (we leave the details as an exercise). $\square$

Exercise III.3.

If the first non-zero derivative is $(d^k f)_x$, we let $f_x^\circ = (d^k f)_x / k!$, so that $f_x^\circ(Y)$ is the homogeneous term of lowest degree in the Taylor series of f at x. In local coordinates $z \in \mathbb{C}^m$ centered at x as above, we have

$$f \circ \varphi^{-1}(z) = P_k(z) + \text{higher order terms},$$

where P_k is a homogeneous polynomial of degree k. We then have

$$f_x^\circ(Y) = P_k(y),$$

where y denotes the vector $T_x\varphi(Y) \in \mathbb{C}^m$.

First Integrals. We now suppose that the function f is a first integral of the differential equation X (and thus it is constant on the trajectories). This can be expressed by the fact that $f(\varphi_t(x))$ does not depend on t, or, infinitesimally, by the equation $X \cdot f = 0$.

We first assume that f is a holomorphic function. For all x in W, let $v_x(f)$ be the order of the zero of f at x.

LEMMA III.1.7 (Ziglin [**84**]). *If f is a first integral of X, then the function $x \mapsto v_x(f)$ is constant on the trajectory Γ of X.*

PROOF. By assumption, $f \circ \varphi_t(x)$ does not depend on t. Thus if

$$(df)_{\varphi_t(x)} = 0, \ldots, (df^{k-1})_{\varphi_t(x)} = 0 \text{ and } (d^k f)_{\varphi_t(x)} \neq 0$$

for some value of t, the same is true for all values of t. □

We now define a function

$$f^\circ : i^\star TW \longrightarrow \mathbb{C}$$

by setting

$$f^\circ(y) = \frac{1}{k!}(d^k f)_x(y) \quad \text{for } k = v_x(f)$$

for each $x \in \Gamma$ and $y \in T_x W$. This is called the *initial form* of f. (The alternative ugly terminology "junior term" comes from a brutal translation of [**84**] but, unfortunately, it seems to be well-established.)

EXAMPLE III.1.8. Once again, we consider the Hénon–Heiles system:

$$H = \frac{1}{2}(p_1^2 + p_2^2) - q_2^2(A + q_1) - \frac{\lambda}{3}q_1^3$$

along the trajectory $\Gamma : t \mapsto (q_1(t), 0, p_1(t), 0)$ contained in the plane $q_2 = p_2 = 0$. We already know that the Hamiltonian H is a first integral of the field X_H. Its initial form along Γ is

$$\begin{aligned} H^\circ(t, Q_1, Q_2, P_1, P_2) &= (dH)_{(q_1(t),0,p_1(t),0)}(Q_1, Q_2, P_1, P_2) \\ &= -\lambda q_1(t)^2 Q_1 + p_1(t) P_1. \end{aligned}$$

REMARK III.1.9. As we see in this example, the initial form may depend explicitly on the time t.

EXAMPLE III.1.10. The Hamiltonian vector field associated to the function

$$H = \frac{1}{2}(p_1^2 + p_2^2) + \lambda q_2^2 + (q_1^2 + q_2^2)^2, \qquad \lambda \neq 0$$

(the anharmonic oscillator of Section I.2.e) has a particular solution Γ whose support is the curve defined by

$$\frac{1}{2}p_1^2 + q_1^4 = 0, \quad q_2 = p_2 = 0$$

and parametrized by

$$q_2 = p_2 = 0, \qquad q_1 = \frac{i}{\sqrt{2}\,t}, \qquad p_1 = -\frac{i}{\sqrt{2}\,t^2}, \qquad t \in \mathbb{R}^\star.$$

The initial form of H along Γ is

$$H^\circ = p_1P_1 + 4q_1^3Q_1,$$

and that of the first integral K described in Section I.3.d is the quadratic form

$$K^\circ = \Big(q_1^2 - \frac{1}{\lambda}p_1^2\Big)Q_2^2 + \frac{2}{\lambda}q_1p_1Q_2P_2 - \frac{1}{\lambda}q_1^2P_2^2.$$

REMARK III.1.11. Suppose that the function f is only *meromorphic*. If we write it as the quotient of two holomorphic functions, we then see that the initial form is defined (and is well-defined) as the quotient of the initial forms of these functions. Note that f° is a rational function in each fiber of $i^\star TW$ and a meromorphic function on Γ.

Here is a third easy lemma, coming from [**69**] and [**84**].

LEMMA III.1.12. *If f is a first integral of the vector field X, then f° is a first integral of the variational equation.*

PROOF. Using coordinates (t, y) on W in the neighborhood of $x \in \Gamma$ in which $X = \partial/\partial t$ and Γ is the t-axis, we see that f is a function of y alone and that f° is its lowest order homogeneous term. Then the solutions of the variational equation are the constant vectors Y, on which f° is certainly constant. The case where f is only meromorphic is left as an exercise. $\square$

Exercises III.4, III.5.

Invariance Under the Galois Group. The following lemma, which follows easily from the fact that $t \mapsto f^\circ(Y(t))$ is constant when Y is a solution of the variational equation, completes this subsection.

LEMMA III.1.13 (Morales–Ramis [**62**]). *The Galois group of the variational equation leaves the initial form f° of the meromorphic integral f invariant.*

The readers who do not know enough of differential Galois theory can look at Appendix A and the references given there.

PROOF. If Y is a solution, then the function $t \mapsto f^\circ(Y(t))$ lies in the Picard–Vessiot extension of $(\mathcal{M}(\Gamma), d/dt)$ determined by the variational equation (where, as usual, $\mathcal{M}(\Gamma)$ denotes the field of meromorphic functions on Γ). Since it is constant, it lies in the invariant field under the Galois group. We fix a point x_0 of Γ, and consider the map

$$\begin{aligned} T_{x_0}W &\longrightarrow \mathbb{C} \\ Y &\longmapsto f^\circ_{x_0}(Y). \end{aligned}$$

For each solution $Y(t)$ of D with $Y(0) = Y$, we have

$$f^\circ_{x(t)}(Y(t)) = f^\circ_{x_0}(Y)$$

since f° is a first integral.

Now, let σ be an element of the Galois group of the variational equation. If $Y \in T_{x_0}W$, then $\sigma(Y)$ is the value at x_0 of the image of the solution whose value at x_0 is Y. Hence we have

$$\begin{aligned} f^\circ_{x_0}(Y) &= f^\circ_{x(t)}(Y(t)) \\ &= \sigma \cdot f^\circ_{x(t)}(Y(t)), \text{ since } \sigma \text{ fixes constants} \\ &= f^\circ_{x(t)}(\sigma(Y(t))), \text{ since } \sigma \text{ acts on the solutions while fixing } \mathcal{M}(\Gamma) \\ &= f^\circ_{x_0}(\sigma(Y)), \text{ since } f^\circ \text{ is constant on the solutions.} \end{aligned}$$

Thus the Galois group, viewed as a subgroup of $\mathrm{GL}(T_xW)$, fixes the rational function f°_x. □

III.1.d. Ziglin's "Lemma". This lemma says that if a vector field X has k independent first integrals, then the same is true for the corresponding variational equation.

We consider a neighborhood of $0 \in \mathbb{C}^m$, with coordinates $(x_1, \ldots, x_m)$.

THEOREM III.1.14 (Ziglin's lemma [**84**]). *Let $f_1, \ldots, f_k$ be (germs of) functionally independent meromorphic functions defined on a neighborhood of the origin in $\mathbb{C}^m$. Then there exist polynomials $P_1, \ldots, P_k$, with $P_i \in \mathbb{C}[z_1, \ldots, z_i]$, such that the initial forms at the origin of the k functions $g_i = P_i(f_1, \ldots, f_i)$ are rational functions that are algebraically independent in $\mathbb{C}(x_1, \ldots, x_m)$.*

In other words, even if the initial forms of the f_i are not independent at the point x, there are k first integrals in the algebra generated by the first integrals whose initial forms at x are independent. A proof following Ziglin's original article [**84**] (see also [**17**]) can be found in Section III.6, along with a discussion of the concepts of independence used in the statement of the theorem.

III.2. The Case of a Hamiltonian System

In this section, we will apply the above constructions to Hamiltonian systems on symplectic manifolds.

A Convention. We will assume that the symplectic manifold on which the Hamiltonian system is defined is a *real analytic*; in fact, in almost all of the examples considered in these notes it is either $\mathbb{R}^{2n}$ or a submanifold of $\mathbb{R}^{2n}$ determined by polynomial equations. Naturally, we will also assume the symplectic forms and the Hamiltonians in question to be real analytic[4].

In this way, everything can be complexified as necessary: differential Galois theory works on any *algebraically closed* field (see Appendix A) and for *analytic* or meromorphic functions (see Footnote 5).

From now on we assume that W is a symplectic manifold with symplectic form ω and that X is a Hamiltonian vector field.

[4]or possibly only meromorphic

III.2.a. The Poisson Bracket and Initial Forms.

LEMMA III.2.1. *If f and g are meromorphic functions on W that commute with respect to the Poisson bracket defined by ω, then the initial forms at x are functions on T_xW that commute with respect to the Poisson bracket defined by ω_x on this vector space.*

PROOF. I will consider only holomorphic functions, leaving the meromorphic case to brave readers. We begin by expanding f and g in power series in a neighborhood of x

$$f(x+h) = f(x) + f_x^\circ(h) + \text{higher order homogeneous terms}$$

and similarly for g, so that

$$\{f, g\}(x+h) = \{f_x^\circ, g_x^\circ\}_x(h) + \text{higher order terms},$$

where $\{\ ,\ \}_x$ denotes the Poisson bracket associated to ω_x on T_xW. If $\{f, g\} = 0$, then, in particular, $\{f_x^\circ, g_x^\circ\}_x = 0$, i.e., f_x° and g_x° commute. □

REMARK III.2.2. We carried out these calculations in terms of local coordinates. Recall (see the discussion around Lemma III.1.6) that the higher order terms depend on the choice of these local coordinates but that this is not the case for the fact that they are of higher order!

III.2.b. The Galois Group is Symplectic. We first note that the Galois group is symplectic in the Hamiltonian case.

LEMMA III.2.3. *The Galois group of the variational equation, treated as a subgroup of the linear group* $\mathrm{GL}(T_xW)$*, is a subgroup of the symplectic group* $\mathrm{Sp}(T_xW)$.

This result is a special case of a general theorem due to Kolchin (see [**51**]). I give a proof in Section A.4. The geometric property which is the "raison d'être" for this lemma is expressed by the following result.

LEMMA III.2.4. *The variational equation for a Hamiltonian system is of the form $Y' + AY = 0$ with $A(t) \in \mathfrak{sp}(2n; \mathbb{C})$ for all t.*

PROOF. Let X be a vector field preserving the symplectic form ω (in particular, this is the case when X is Hamiltonian), so that $\mathcal{L}_X\omega = 0$. If Y and Z are two vector fields defined along the trajectory Γ, we write

$$0 = (\mathcal{L}_X\omega)(Y, Z) = X \cdot \omega(Y, Z) - \omega(DY, Z) - \omega(Y, DZ).$$

In terms of matrices (having chosen a basis), this implies that

$$\begin{aligned}\frac{d}{dt}\omega\big(Y(t), Z(t)\big) &= \omega(Y' + AY, Z) + \omega(Y, Z' + AZ)\\ &= \omega(Y', Z) + \omega(Y, Z') + \omega(AY, Z) + \omega(Y, AZ).\end{aligned}$$

Thus the sum of the extra terms is zero, i.e., the matrix A satisfies

$$\omega(AY, Z) + \omega(Y, AZ) = 0$$

for all Y and Z, which is precisely the definition of the Lie algebra $\mathfrak{sp}(2n; \mathbb{C})$. □

III.2.c. The Normal Variational Equation. In the case of a Hamiltonian system, we can use symplectic reduction to decrease the order of the linearized equation thereby simplifying matters significantly.

We consider an isotropic vector subbundle $E \subset i^\star TW$ and its orthogonal complement (with respect to ω) E°. Since E is isotropic, we have $E \subset E^\circ$ (it is this inclusion that says E is isotropic).

LEMMA III.2.5. *The bundle E is invariant under D if and only if the bundle E° is.*

Thus D defines an operator, that we denote D_N, on the symplectic bundle E°/E.

PROOF. Let Y and Z be sections of E and E°, respectively. Then $\omega(Y, Z)$ is identically zero, so that we have

$$0 = X \cdot \omega(Y, Z) = (\mathcal{L}_X \omega)(Y, Z) + \omega(DY, Z) + \omega(Y, DZ).$$

Since $\mathcal{L}_X \omega = 0$, this implies that

$$\omega(DY, Z) + \omega(Y, DZ) = 0$$

for all Y and Z, giving us the desired result. □

If the trajectory Γ is not stationary, the vector bundle $i^\star TW$ must still contain an isotropic subbundle, the tangent bundle $T\Gamma$ generated by X. Let Q be the level set of the function H containing Γ. Then the orthogonal complement of $T\Gamma$ is $i^\star TQ$: since $\omega(X, Y) = -dH(Y)$, the orthogonal complement of X is just the kernel of dH.

It is clear that the subbundle $T\Gamma$ of $i^\star TW$ is invariant under D. Indeed, if X and Y are tangent to Γ, then $\mathcal{L}_X Y = [X, Y]$ is tangent to Γ. Hence the lemma can be applied to

$$T\Gamma \subset i^\star TQ \subset i^\star TW.$$

We denote the reduced bundle[5] $i^\star TQ/T\Gamma$ by N (for "normal," since it is the symplectic normal bundle of Γ). It is equipped with the operator D_N, and the equation $D_N Z = 0$ is called the *normal* variational equation. Just as for the usual variational equation, it is sometimes convenient to consider the connection $\nabla_N = D_N \otimes dt$.

EXAMPLE III.2.6 (Hénon–Heiles—Continuation). We return to the linear system obtained by linearizing the Hénon–Heiles Hamiltonian system (see Section I.2.d and Example III.1.5). Consider a solution Γ contained both in the plane $q_2 = p_2 = 0$ and the level set $H = 0$. For simplicity, we write $(q_1, p_1) = (x, y)$. Then the support of the solution Γ is the curve

$$\frac{1}{2}y^2 - \frac{\lambda}{3}x^3 = 0 \qquad (x, y) \neq (0, 0).$$

[5] See Exercise I.4.

The symplectic bundle normal to Γ can be identified with the trivial bundle with fiber the (Q_2, P_2) plane. Then the normal variational equation can be written as

$$\begin{cases} \dot{Q}_2 = P_2 \\ \dot{P}_2 = 2(A + q_1(t))Q_2 \end{cases}$$

or, by setting $q_1 = x$ and $Q_2 = \xi$, as the equation

$$\ddot{\xi} - 2(A + x(t))\xi = 0.$$

Example III.2.7 (Reduction for Jacobi Fields). In the case of a variational equation along a geodesic of a Riemannian manifold, which is a differential equation of order $2n$, one often only considers to Jacobi fields orthogonal to the geodesic. In the cotangent bundle of V, this corresponds exactly to the reduction operation given above. In this context, it should not come as a surprise that the space of Jacobi fields has dimension $2n - 2$.

III.3. The Case of an Integrable System

In this section, I study what symplectic geometry says about the Galois group of a variational equation in the case of a completely integrable system. As we have seen above (Lemma III.2.3), the Galois group is a subgroup of the symplectic group; now, in addition, it must preserve the n initial forms, which Ziglin's lemma (Theorem III.1.14) allows us to assume to be independent.

As a geometric introduction, I begin (Section III.3.a) with an explanation of what the Arnold–Liouville theorem (Chapter II) says about the Galois group along a generic trajectory. Then (Section III.3.b), I derive the restriction placed on the Galois group by the preservation of n independent linear forms. The moral of these two sections is that along a generic trajectory, the Galois group is "tiny," and certainly Abelian and connected. Finally, I prove the more refined non-integrability theorem of Morales and Ramis (Section III.3.c), which states that, for any trajectory, the Galois group is "almost" Abelian. We will need to wait for the reduction technique that will be introduced in Section III.4, before applying this result to non-integrability. Before doing that, I give an application to the calculation of some Galois groups.

III.3.a. Along a Generic Trajectory of an Integrable System. Consider an integrable system on a (real analytic) symplectic manifold, and suppose that one of the first integrals (the Hamiltonian itself, for example) is proper. Then we can invoke the Arnold–Liouville theorem[6] to write the system in terms of "action–angle" variables.

The Hamiltonian System in Action–Angle Variables. Consider now an open subset U of $\mathbb{R}^n$, a torus $T^n = \mathbb{R}^n/\mathbb{Z}^n$, and the manifold $W = U \times T^n$ with the symplectic form

$$\omega = dx_1 \wedge d\theta_1 + \cdots + dx_n \wedge d\theta_n,$$

[6] For a real analytic version, see, for example, [**44**].

where the x_i are the coordinates in U and the $d\theta_i$ are the 1-forms on T^n corresponding to the coordinates $\theta_1, \ldots, \theta_n$ of $\mathbb{R}^n$.

Given a function $h : U \to \mathbb{R}$, we are interested in the Hamiltonian system associated to the composition

$$H : W \longrightarrow U \xrightarrow{\ h\ } \mathbb{R},$$

that is, in the differential system

$$\dot{x}_i = 0 \qquad \dot{\theta}_i = \frac{\partial h}{\partial x_i}(x_1, \ldots, x_n)$$

(defined here on the covering $U \times \mathbb{R}^n \to W$). Since the functions x_i are n independent first integrals in involution, this Hamiltonian system is completely integrable.

The Solutions. Choose a solution

$$x_i = a_i, \qquad \theta_i = \frac{\partial h}{\partial x_i}(a_1, \ldots, a_n)t + \theta_i^0 \qquad 1 \leqslant i \leqslant n,$$

which is globally parametrized by $t \in \mathbb{R}$. The trajectory Γ is

(a) $\mathbb{R}$ if the partial derivatives $\dfrac{\partial h}{\partial x_i}(a_1, \ldots, a_n)$ are not all rational multiples of the same rational number $T > 0$,

(b) $\mathbb{R}/T\mathbb{Z}$ otherwise.

The Solutions of the Variational Equation. The variational equation along Γ is simply the linearized system

$$\dot{X}_i = 0 \qquad \dot{\Theta}_i = \sum_{j=1}^{n} \frac{\partial^2 h}{\partial x_i \partial x_j}(a_1, \ldots, a_n)X_j,$$

with solutions

$$X_i = A_i \qquad \Theta_i = \Big(\sum_{j=1}^{n} \frac{\partial^2 h}{\partial x_i \partial x_j}(a_1, \ldots, a_n)A_j\Big)t + \Theta_i^0.$$

Note that

(a) When $\Gamma = \mathbb{R}$, the solutions are well defined on Γ.

(b) When $\Gamma = \mathbb{R}/T\mathbb{Z}$, we have two possibilities:

(b′) There is an integer m such that $\Theta_i(t+mT) = \Theta_i(t)$ for all t, implying that the solutions are well defined on a finite covering on Γ. Such an integer exists if and only if the numbers

$$\sum_{j=1}^{n} \frac{\partial^2 h}{\partial x_i \partial x_j}(a_1, \ldots, a_n)A_j$$

are rational for any choice of the A_j, and this occurs exactly when the Hessian matrix of h at $(a_1, \ldots, a_n)$ is identically zero. In this case, the solutions of the variational equation are the constants. An interesting special case is when h is linear. This the case where we consider a Hamiltonian action of the torus T^n and where the fundamental vector field X_H is associated to a closed subgroup of this torus. See more generally Exercise III.6.

(b″) Otherwise, some of the solutions can be defined only on the universal covering $\mathbb{R}$ of Γ.

As we have already noted, it is preferable to complexify the situation so that the Picard–Vessiot method can be utilized efficiently. This is why we have only considered real analytic objects. We now consider the complexified trajectory $\Gamma_{\mathbb{C}}$, here the germ of $\mathbb{C} \supset \mathbb{R}$ or $\mathbb{C}/T\mathbb{Z} \supset \mathbb{R}/T\mathbb{Z}$. Concretely, the field $\mathcal{M}(\Gamma_{\mathbb{C}})$ of functions on $\Gamma_{\mathbb{C}}$ is the field of fractions of

- the ring of complex analytic functions defined on a neighborhood of $\mathbb{R} \subset \mathbb{C}$ if $\Gamma = \mathbb{R}$, or
- the ring of periodic complex analytic functions of period T defined in a neighborhood of $\mathbb{R} \subset \mathbb{C}$ if $\Gamma = \mathbb{R}/T\mathbb{Z}$. In this case, we set $z = \exp(it/T)$; then the elements of $\mathcal{M}(\Gamma_{\mathbb{C}})$ are the quotients of the analytic functions (of z) defined on a neighborhood of $S^1 \subset \mathbb{C}^\star$.

The Galois Group. Cases (a) and (b′) yield trivial Galois groups. In case (b″), the Picard–Vessiot extension is generated by t. The Galois group is a 1-dimensional additive group. It is represented in the symplectic group by

$$c \longmapsto \begin{pmatrix} \mathrm{Id} & 0 \\ cA & \mathrm{Id} \end{pmatrix},$$

where A is the Hessian matrix of h at $(a_1, \dots, a_n)$. We have thus proved:

PROPOSITION III.3.1. *Let H be an integrable Hamiltonian on a real analytic symplectic manifold. The Galois group of the variational equation along a trajectory contained in a compact component of a regular level set of the first integrals is a connected subgroup of dimension $\leqslant 1$ of the symplectic group.*

Exercises III.6, III.7, III.8.

III.3.b. Subgroups of the Symplectic Group Preserving Linear Forms. We have said that the Galois group is a symplectic group (Lemma III.2.3) and that is preserves initial forms. When these are independent linear forms, we have the following simple result.

LEMMA III.3.2. *The subgroup of the symplectic group* $\mathrm{Sp}(2n)$ *consisting of transformation preserving n independent linear forms in involution is a conjugate of the group of matrices of the form*

$$\begin{pmatrix} \mathrm{Id} & 0 \\ A & \mathrm{Id} \end{pmatrix},$$

where A is an $n \times n$ symmetric matrix.

PROOF. This can be proved by elementary methods of linear algebra (see Exercise I.6). First of all, the n linear forms in involution generate a Lagrangian subspace of the dual space. We next look for the symplectic transformations of the dual that restrict to the identity on this Lagrangian. This is done by extending the n forms to a symplectic basis. Finally in this dual basis, our transformations are the desired matrices. □

REMARKS III.3.3.

(1) The first integrals may be independent along Γ even if Γ is not contained in a regular level set (see, for example, Exercise III.17).
(2) This group has dimension $n(n+1)/2$ which gives a very crude bound when applied to the Galois group where we have already seen (at least in the case where one of the first integrals is proper) that it has dimension at most 1.
(3) The group obtained here is isomorphic to the *additive* group of the vector space of $n \times n$ symmetric matrices. In Example III.3.14, we will see that the Galois group may contain multiplicative factors when there are initial forms of higher degree.
(4) This group is connected as well as Abelian. Since it is an *additive* group, it is easy to check that all its algebraic subgroups are connected (see Exercise III.8).

COROLLARY III.3.4. *Let H be a completely integrable Hamiltonian and let Γ be a trajectory of X_H along which the first integrals are independent. Then the Galois group of the variational equation along Γ is a connected Abelian group.*

PROOF. Let x be a point on Γ. Since the first integrals at x are assumed to be independent, their initial forms are independent linear forms[7] on T_xW. We apply Lemma III.3.2. The Galois group is thus an algebraic subgroup of the additive group of symmetric matrices and thus it is Abelian and connected, as noted above. □

REMARK III.3.5. We do not require here that the level set containing Γ to be regular but only that the points of Γ are regular. In addition, we also do not require that one of the first integrals is proper. Compare this result to Proposition III.3.1.

Next is an example where Lemma III.3.2 is applicable but where the Galois group is of dimension 2 (and where we cannot apply Proposition III.3.1, as none of the integrals is proper and the level set is not regular).

EXAMPLE III.3.6 (An Integrable Hénon–Heiles Hamiltonian). We consider the Hénon–Heiles Hamiltonian:

$$H = \frac{1}{2}(p_1^2 + p_2^2) - q_2^2 q_1 - \frac{\lambda}{3} q_1^3$$

on $\mathbb{R}^4$. When $\lambda = 6$, we know a second first integral, i.e., the function

$$K = q_2^4 + 4q_1^2 q_2^2 + 4p_2(p_2 q_1 - p_1 q_2)$$

(see Exercise I.21).

We consider the trajectory Γ of X_H defined by

$$q_2 = p_2 = 0, \qquad p_1^2 = 4q_1^3 + 2h \quad \text{for } h \neq 0.$$

We use the Weierstrass $\wp$-function[8] associated with the smooth elliptic curve Γ to parametrize the trajectory by the time t

$$q_1 = \wp(t), \qquad p_1 = \wp'(t).$$

[7] See Remark III.3.12.

[8] See Appendix B for the differential equation satisfied by the Weierstrass $\wp$-function.

The variational equation along Γ is the system

$$\begin{cases} \dot{Q}_1 = P_1 & \dot{P}_1 = 12q_1Q_1 \\ \dot{Q}_2 = P_2 & \dot{P}_2 = 2q_1Q_2, \end{cases}$$

which is equivalent to the two (independent) equations

$$\ddot{Q}_1 = 12\wp(t)Q_1 \text{ and } \ddot{Q}_2 = 2\wp(t)Q_2.$$

Each of these is an example of a "Lamé equation". The first has $Q_1 = \wp(t)$ as an obvious solution, while the second is just as obviously satisfied by $Q_2 = \wp'(t)$. It is a classical and a fairly easy problem[9] to show that neither of these has another biperiodic solution, that the Picard–Vessiot extension has transcendence degree 1 in both cases and that the Galois group is additive. The Galois group of the system is thus the group of matrices of the form

$$\left(\begin{array}{c|c} \mathrm{Id} & 0 \\ \hline \begin{matrix} \alpha & 0 \\ 0 & \beta \end{matrix} & \mathrm{Id} \end{array}\right) \quad \text{with } \alpha, \beta \in \mathbb{C},$$

which is a connected group of dimension 2.

III.3.c. The Morales–Ramis Theorem. Results like Proposition III.3.1 and Lemma III.3.2 impose severe restrictions on the Galois groups along generic trajectories of a completely integrable Hamiltonian system. Because these restrictions can be summarized by the word "generic", these results cannot serve as a criterion for non-integrability. The Morales–Ramis theorem, on the other hand, provides such a criterion. It is based on a result, that we consider to be the key lemma of [**62**], a lemma analogous to III.3.2 but a little more general although not much harder to prove.

LEMMA III.3.7 (Key Lemma). *Let (E, ω) be a symplectic vector space of dimension $2n$. Let $F_1, \dots, F_n$ be functionally independent $\mathcal{C}^1$ functions in involution on an open subset of E. If $\mathfrak{g}$ is a Lie subalgebra of $\mathfrak{sp}(E)$ preserving the F_i, then $\mathfrak{g}$ is Abelian.*

In the statement of this lemma, the Lie algebra $\mathfrak{g}$ is a Lie sub-algebra of $\mathfrak{sp}(E)$, so any $Z \in \mathfrak{g}$ can be considered as an endomorphism of E. Thus we can consider the map

$$x \longmapsto Z \cdot x$$

as a vector field on E. Saying that Z fixes a function F defined on an open subset of E is then equivalent to saying that the vector field Z fixes F (that is, $Z \cdot F = 0$), or, equivalently, that $(dF) \circ Z = 0$.

PROOF OF THE KEY LEMMA. Let V be a common level set of the F_i, which we suppose is regular. Its tangent space, the intersection $\bigcap(\ker dF_j)$, is generated by the Hamiltonian vector fields $X_1, \dots, X_n$ of the functions $F_1, \dots, F_n$.

[9]See the chapter on Lamé equations in the second volume of [**39**].

Let $Z \in \mathfrak{g}$. Since $dF_i \circ Z = 0$, $Z(x) \in \ker(dF_i)_x$ for all i. Thus

$$Z(x) \in \bigcap_{i=1}^{n} \ker(dF_i)_x = \langle X_1(x), \dots, X_n(x)\rangle^\circ = \langle X_1(x), \dots, X_n(x)\rangle$$

for $x \in V$. Thus Z can be written as $Z = \sum \lambda_j X_j$ for some functions $\lambda_1, \dots, \lambda_n$.

Since $Z \in \mathfrak{sp}(E)$, we have $\mathcal{L}_Z\omega = 0$ (see Exercise III.10). This implies that $di_Z\omega = 0$, i.e., $i_Z\omega$ is a closed form. Thus the vector field Z is Hamiltonian, here it is globally Hamiltonian, since E is a vector space, and thus $i_Z\omega$ is (globally) exact. Hence $Z = X_g$ for some quadratic form g. Then

$$[X_i, Z] = [X_i, X_g] = X_{\{F_i, g\}} = 0,$$

since $\{F_i, g\} = Z \cdot F_i = 0$. But

$$\Big[X_i, \sum_{j=1}^{n} \lambda_j X_j\Big] = \sum_{j=1}^{n} \big((X_i \cdot \lambda_j) X_j + \lambda_j [X_i, X_j]\big) = \sum_{j=1}^{n} (X_i \cdot \lambda_j) X_j,$$

since the F_i commute. Eventually $X_i \cdot \lambda_j = 0$ for all i and j, so the functions $\lambda_1, \dots, \lambda_n$ are constant on V. Hence $Z = \sum \lambda_j(V) X_j$.

Let Z' be another element of $\mathfrak{g}$, which we write as $Z' = X_{g'}$ and $Z' = \sum \lambda'_j(V) X_j$ as above. We compute

$$[Z, Z'] = \big[X_g, \sum \lambda'_j(V) X_j\big] = \sum \big(X_g \cdot \lambda'_j(V)\big) X_j + \sum \lambda'_j(V)[Z, X_j].$$

Since both terms of the last sum are zero, the vector fields Z and Z' commute. Hence (see Exercise III.9), the endomorphisms Z and Z' commute in $\mathfrak{g}$ and $\mathfrak{g}$ is indeed Abelian. □

Exercises III.9, III.10, III.11.

Remark III.3.8. Note that we have utilized the power of infinitesimal calculus and, in particular, the fact that the group G is a Lie group; note that this applies to the Galois group, which is an algebraic group. Unfortunately, we lose all control over properties of connectedness.

Remark III.3.9. Under the hypotheses of Lemma III.3.2, the proof of the key lemma gives us for the Lie algebra the result corresponding to the conclusion of Lemma III.3.2, i.e., the matrices Z are linear combinations of the $F_i X_{F_j}$, symmetric in i and j.

This brings us to the Morales–Ramis theorem.

Theorem III.3.10 ([**62**]). *Let Γ be a trajectory of the Hamiltonian vector field X on the symplectic manifold W of dimension $2n$. If X has n independent first integrals in involution, then the Lie algebra of the Galois group of the variational equation along Γ is Abelian.*

PROOF. Fix a point x on Γ, and let E be the symplectic vector space T_xW and ω the (constant) symplectic form ω_x. Ziglin's lemma (Theorem III.1.14) provides us with n independent functions on this vector space, the initial forms at x of n independent first integrals of X. We denote them by $F_1, \dots, F_n$.

By virtue of Lemma III.2.1, these functions are in involution, by virtue of Lemma III.1.13, the elements of the Galois group preserve the F_i. By virtue of Lemma III.2.3, the Galois group is a subgroup of the symplectic group. The theorem is then a consequence of the key lemma. □

REMARK III.3.11. This theorem must be refined for use in applications. In particular, we will have to add singular points and/or points at infinity to Γ. However, I have preferred to present this "simple" version since it captures the main ideas of the result. The applications discussed in Section III.5 will use more refined versions.

REMARK III.3.12 (A Remark About Independence). Ziglin's lemma seems to address only the question of independence at a point. By letting $f_1, \dots, f_k$ be first integrals and φ_t be the flow of the Hamiltonian vector field X, we see that

$$T_{x_0}\varphi_t(X_{f_i}(x_0)) = X_{f_i}(\varphi_t(x_0)),$$

since $[X, X_{f_i}] = 0$. Thus if the vectors $X_{f_1}(x_0), \dots, X_{f_k}(x_0)$ are independent, their images under the isomorphism $(\varphi_t)_\star$ are independent as well. We conclude that if the first integrals $f_1, \dots, f_k$ are independent at $x_0 \in \Gamma$, then they are independent at every point on Γ.

III.3.d. Applications to Calculating Galois Groups. In addition to giving us the non-integrability criterion of Theorem III.3.10, the proof of the key lemma also has applications in determining the "size" of the Galois group.

Consider the case where k initial forms of the first integrals are linear and the remaining $n-k$ are quadratic. Then the key lemma implies that the elements of the Lie algebra of the group G are

$$Z(x) = \sum_{1\leqslant i\leqslant j\leqslant k} a_{i,j}\left(F_i(x)X_{F_j} + F_j(x)X_{F_i}\right) + \sum_{j=k+1}^{n} b_j X_{Q_j}(x),$$

which imply the following bound on the dimension of this Lie algebra

$$\frac{k(k+1)}{2} + n - k.$$

We also note that this bound improves for small k, i.e., the more the level set is critical, the better result the key lemma yields for the Galois group.

In the case of a system with two degrees of freedom and a trajectory contained in a regular level set of a Hamiltonian along which the second first integral is quadratic, the dimension of the Galois group is bounded by 2. We have encountered examples of this in Section III.3.b.

We consider the case of a "very" singular level set. If we write

$$Z(x) = \sum_{i=1}^{n} \lambda_i(x) X_{F_i}(x),$$

$Z(x)$ is a linear vector field. If F_i is a homogeneous polynomial of degree d_i, we see that $X_{F_i}(x)$ is a vector field whose components are homogeneous polynomials of degree $d_i - 1$ in x. When $d_i - 1 \geqslant 2$, this forces the coefficient $\lambda_i(x)$ to be identically zero. This gives us the following refined result.

PROPOSITION III.3.13. *Suppose that along Γ, the initial forms of the first integrals consist of*

- *k linear forms,*
- *ℓ quadratic forms,*
- *and $n-k-\ell$ polynomials of degree $\geqslant 3$.*

Then the Lie algebra of the group G is Abelian with dimension $\leqslant \dfrac{k(k+1)}{2} + \ell$.

EXAMPLE III.3.14. Here is an example where the key lemma helps us determine the Galois group. This is the exampe of the anharmonic oscillator that we already considered in Sections I.2.e and I.3.d and Example III.1.10. Recall that the Hamiltonian for the anharmonic oscillator is

$$H = \frac{1}{2}(p_1^2 + p_2^2) + \lambda q_2^2 + (q_1^2 + q_2^2)^2, \qquad \lambda \neq 0$$

that we consider along the solution Γ

$$q_2 = p_2 = 0, \qquad q_1 = \frac{i}{\sqrt{2}\,t}, \qquad p_1 = -\frac{i}{\sqrt{2}\,t^2}, \qquad t \in \mathbb{R}^\star.$$

The solution Γ has as support the curve $\frac{1}{2}p_1^2 + q_1^4 = 0$, $q_2 = p_2 = 0$. The initial forms of the first integrals H and K (see Section III.1.10) along Γ are

$$H^\circ = p_1 P_1 + 4q_1^3 Q_1$$

and

$$K^\circ = \Big(q_1^2 - \frac{1}{\lambda}p_1^2\Big)Q_2^2 + \frac{2}{\lambda}q_1 p_1 Q_2 P_2 - \frac{1}{\lambda}q_1^2 P_2^2.$$

The associated Hamiltonian vector fields are

- the constant field

$$X_{H^\circ} = \big(p_1, 0, -4q_1^3, 0\big)$$

- the linear field

$$X_{K^\circ} = \left(0, \frac{2}{\lambda}q_1 p_1 Q_2 - \frac{2}{\lambda}q_1^2 P_2, 0, -2\Big(q_1^2 - \frac{1}{\lambda}p_1^2\Big)Q_2 - \frac{2}{\lambda}q_1 p_1 P_2\right).$$

The elements Z of the Lie algebra of the Galois group are the linear combinations

$$Z = AH^\circ X_{H^\circ} + BX_{K^\circ}.$$

Here it is nicer to write these as block matrices by writing the coordinates in the order q_1, p_1, q_2, p_2. Then the matrix Z takes the form

$$Z = \begin{bmatrix} 4Aq_1^3p_1 & Ap_1^2 & 0 & 0 \\ -16Aq_1^6 & -4Aq_1^3p_1 & 0 & 0 \\ 0 & 0 & \dfrac{2B}{\lambda}q_1p_1 & -\dfrac{2B}{\lambda}q_1^2 \\ 0 & 0 & -2B(q_1^2 - \dfrac{1}{\lambda}p_1^2) & -\dfrac{2B}{\lambda}q_1p_1 \end{bmatrix}.$$

The matrix Z is nilpotent when $B = 0$ and diagonalizable when $A = 0$. The second term represents the normal Galois group and. The Lie algebra of the Galois group is the sum of

- a tangent component contained in the Lie algebra of matrices $\begin{pmatrix} 0 & 0 \\ A & 0 \end{pmatrix}$
- and a normal component contained in the Lie algebra of matrices $\begin{pmatrix} B & 0 \\ 0 & -B \end{pmatrix}$.

The identity component of the Galois group is thus contained in the group of matrices

$$\left(\begin{array}{cc|cc} 1 & 0 & & \\ a & 1 & \multicolumn{2}{c}{0} \\ \hline & & b & 0 \\ \multicolumn{2}{c|}{0} & 0 & b^{-1} \end{array}\right) \quad \text{with } (a,b) \in \mathbb{C} \times \mathbb{C}^\star.$$

The Variational Equation. In this case, the variational equation is

$$\begin{cases} \dot{Q}_1 = P_1 & \dot{P}_1 = -12q_1^2Q_1 = \dfrac{6}{t^2}Q_1 \\ \dot{Q}_2 = P_2 & \dot{P}_2 = -2(\lambda + 2q_1^2)Q_2 = -2\Big(\lambda - \dfrac{1}{t^2}\Big)Q_2. \end{cases}$$

As we could expect from what we have seen above, the system is composed of two independent subsystems equivalent to the two equations

$$\ddot{Q}_1 = \frac{6}{t^2}Q_1, \qquad \ddot{Q}_2 = -2\Big(\lambda - \frac{1}{t^2}\Big)Q_2.$$

The solutions of the first are $Ct^3 + Dt^{-2}$; its Galois group is trivial and thus (strictly!) contained in that of the matrices $\begin{pmatrix} 1 & 0 \\ a & 1 \end{pmatrix}$.

The second is a "Whittaker equation", and what we have seen above implies that its Galois group is a subgroup of the multiplicative group of diagonal matrices. It is a classical result (that is to say, one well known to specialists) that this group is in fact the entire multiplicative group.

III.4. Towards Applications: The Reduction

It might appear surprising that a statement like Theorem III.3.10 would have applications. If we want to use it to show that a Hamiltonian system on a manifold of dimension $2n$ is not integrable, we would need not only to find a particular[10] solution, the curve Γ, but, in addition, we would need to show that the Galois group of the corresponding variational equation is sufficiently non-abelian[11].

The variational equation has order $2n$. Although we can reduce the order using the fact that X is a solution, it still remains fairly impractical after doing so, even for a Hamiltonian system with only two degrees of freedom ($n = 2$).

As we have seen in Section III.2.c, the fact that the system is Hamiltonian allows us to utilize a standard process of symplectic geometry and Hamiltonian mechanics to *reduce* the order of the differential equation by constructing a linear differential equation of order $2n - 2$ whose Galois group is forced to be Abelian, analogous to the result of Theorem III.3.10.

For a system with two degrees of freedom, we thus obtain a second order equation, and there exists an algorithm that allows to know whether the identity component of the Galois group is not solvable (and, in particular, not Abelian). This is Kovacic's algorithm (see, for example, [**17**]), which has been successfully used in applications termed "non-academic" by Morales and Ramis (see [**62, 61**] and [**46**]).

Remark III.4.1. Lemma III.2.5 allows us to reduce the order of the variational equation. In fact the more independent first integrals in involution along Γ we know, the greater the reduction. If $f_1 = H, f_2, \ldots, f_k$ are independent and satisfy $\{f_i, f_j\} = 0$, then along Γ the corresponding Hamiltonian vector fields generate an isotropic vector subbundle of rank k whose orthogonal complement is the tangent bundle of the common level set of the f_i containing Γ. The reduction process allows us to construct a linear differential equation of order $2n - 2k$.

III.4.a. The Galois Group of the Reduced Equation.

As we did in Section III.3.b, we begin with a simple result from symplectic geometry.

Proposition III.4.2. *The subgroup of the symplectic group* $\mathrm{Sp}(2n)$ *consisting of transformations that preserve k independent linear forms in involution is a conjugate of an extension of a subgroup of* $\mathrm{Sp}(2n-2k)$ *by a subgroup of the group of matrices of the form*

$$\left[\begin{array}{cc|cc} \mathrm{Id} & 0 & 0 & 0 \\ \alpha & \mathrm{Id} & 0 & 0 \\ \hline A & \beta & \mathrm{Id} & \beta' \\ \alpha' & 0 & 0 & \mathrm{Id} \end{array}\right],$$

Proof. We adapt the proof of Lemma III.3.2. Let G be the group in question and F the isotropic subspace of the dual $E^\star$ of the vector space generated by

[10] and good enough...

[11] In short, the theorem is easy to prove because the Galois group is big enough to have a Lie algebra, while it seems hard to apply because this Galois group is too big to be calculated.

the k linear forms. The elements of G determine transformations of the reduced symplectic vector space F°/F. In other words, we have a morphism of groups

$$G \longrightarrow \mathrm{Sp}(2n-2k).$$

Let us determine its kernel. This is the subgroup of $\mathrm{Sp}(2n)$ consisting of the linear transformations of $E^\star$ that when restricted to the co-isotropic subspace F° are linear maps of the form

$$x \longmapsto x + f(x)$$

for some linear map $f : F^\circ \to F$ that is zero on F. In a symplectic basis of the form

$$F_1, \dots, F_k \text{ (our linear forms)}, F_{k+1}, \dots, F_n, g_1, \dots, g_n,$$

F° is the subspace generated by all the F_i and $g_{k+1}, \dots, g_n$. In this basis, the matrices of the transformations in question are the matrices

$$\left[\begin{array}{cc|cc} \mathrm{Id} & \alpha & A & \alpha' \\ 0 & \mathrm{Id} & \beta & 0 \\ \hline 0 & 0 & \mathrm{Id} & 0 \\ 0 & 0 & \beta' & \mathrm{Id} \end{array}\right].$$

This gives us the desired result by transposition. □

Remark III.4.3. The group of matrices of this form, as well as all of its algebraic subgroups, are *connected*. To check this, we order the basis vectors as follows

$$F_1, \dots, F_k, g_1, \dots, g_k, F_{k+1}, \dots, F_n, g_{k+1}, \dots, g_n$$

and write the matrices as block matrices

$$\left[\begin{array}{c|c|c} \mathrm{Id} & A & a \\ \hline 0 & \mathrm{Id} & 0 \\ \hline 0 & b & \mathrm{Id} \end{array}\right],$$

where A is a $k \times k$ matrix, a a matrix with k rows and $2n-2k$ columns, and b a matrix with $2n-2k$ rows and k columns. The group structure on the triples (A, a, b) is given by

$$(A, a, b) \cdot (A', a', b') = (A + A' + ab', a + a', b + b').$$

This group is an extension of the additive group of the vector space $M_{k,2n-2k} \times M_{2n-2k,k}$ by that of the $k \times k$ matrices. This implies the connectedness of all its algebraic subgroups.

The group of all the symplectic transformations preserving k linear independent forms has the same number of connected components as its image in $\mathrm{Sp}(2n-2k)$.

Proposition III.4.4. *Let H be a completely integrable Hamiltonian on a symplectic manifold of dimension $2n$, and let Γ be a trajectory along which k of its first integrals are independent. Then the Lie algebra of the Galois group of the normal variational equation is Abelian.*

PROOF. We fix a point x of Γ and consider the symplectic group of $T_xW = E$. Let G be the group of *all* the symplectic transformations of E preserving the initial forms of the first integrals (and, in particular, the linear forms $F_1, \ldots, F_k$, the initial forms of k independent integrals). From the key lemma (Lemma III.3.7), the Lie algebra of G is Abelian.

We have just seen that G is equipped with a projection onto $\mathrm{Sp}(2n-2k)$. Let G_N be its image in this group. Finally, consider the Galois group of the normal variational equation (reduced by the k integrals). This is a subgroup of the symplectic group of the reduced space F°/F, and it preserves the reductions of the linear forms; thus it is a subgroup of G_N. Since G_N is a quotient of G, its Lie algebra is Abelian, and since the reduced Galois group is a subgroup of G_N, it also is Abelian. □

It is in fact Proposition III.4.4 that has applications to non-integrability.

The Reduced Galois Group is a Quotient of the Galois Group. Before moving on to applications, let us make a parenthetical observation. Proposition III.4.2 concerns any subgroup of the symplectic group preserving linear forms; we will now give the corresponding result for Galois groups. Surprisingly (and this surprise is the reason why we give this result here), the proof yields exactly the same groups as Proposition III.4.2.

PROPOSITION III.4.5. *Let H be a Hamiltonian on a symplectic manifold of dimension $2n$. Suppose that Γ is a nonstationary trajectory of X_H, and that the system has k independent first integrals in involution along Γ. Then the Galois group of the variational equation along Γ is isomorphic to an extension of a subgroup of* $\mathrm{Sp}(2n-2k)$ *by a subgroup of matrices of the form*

$$\left[\begin{array}{cc|cc} \mathrm{Id} & 0 & 0 & 0 \\ \alpha & \mathrm{Id} & 0 & 0 \\ \hline A & \beta & \mathrm{Id} & \beta' \\ \alpha' & 0 & 0 & \mathrm{Id} \end{array}\right].$$

This is of course a direct corollary of Proposition III.4.2: if $H = f_1, \ldots, f_k$ are the first integrals and x is a point on Γ, the initial forms of the f_i are k independent linear forms in involution on T_xW, and the Galois group is a subgroup of the symplectic group preserving them. The following proof shows that we can not obtain a better result by making the group in question a Galois group, at least if we do not use any specific properties of the trajectory.

REMARK III.4.6. In the proof, we will construct the solutions of D by using those of D_N. To do this we need *bases*. We will use the fact that the holomorphic bundle $i^\star TW \to \Gamma$ on the noncompact Riemann surface Γ is analytically trivial, a non-trivial result that can be found in [**29**]. The key ingredients of the proof of this theorem are

- It is possible to trivialize the bundle among the $\mathcal{C}^\infty$ bundles on Γ.

- There exist many holomorphic functions on Γ, or, more precisely, the group $H^1(\mathcal{O}_\Gamma)$ is trivial, and thus a $\mathcal{C}^\infty$ trivialization will yield a holomorphic trivialization.

Similarly, the proof provides a *symplectic* analytic trivialization in the case of a symplectic bundle.

"Differential Equation" Proof of Proposition III.4.5. The Hamiltonian vector fields $X_1 = X_H$, $X_i = X_{f_i}$ (for $2 \leqslant i \leqslant k$) generate an isotropic subbundle F of $i^\star TW \to \Gamma$ of rank k. Choose vector fields $Y_1(t), \dots, Y_k(t)$ along Γ such that

$$\omega(X_i, Y_j) = \delta_{i,j} \text{ and } \omega(Y_i, Y_j) = 0.$$

Any section Z of $i^\star TW$ can be written as

$$Z = \sum_{i=1}^k \lambda_i X_i + \sum_{i=1}^k \mu_i Y_i + Z_N,$$

where Z_N is orthogonal (with respect to ω) to the symplectic subbundle of $i^\star TW$ generated by the X_i and Y_j: from the way we chose these vector fields, we can identify the reduced symplectic bundle F°/F to a subbundle F_N of $i^\star TW$, of which Z_N is a section.

Since the vector field X_H is a Hamiltonian vector field, we have $\mathcal{L}_X\omega = 0$, from which we deduce that the operator D is symplectic: $\omega(DU, V) + \omega(U, DV) = 0$. We also see that it defines an operator D_N on the reduced bundle (see [**62**]): by differentiating the equalities $\omega(Z_N, X_i) = 0$ we obtain that $\omega(DZ_N, X_i) = 0$ for all i, and thus that we can write

$$DZ_N = \sum \alpha_i(Z_N) X_i + (DZ_N)_N,$$

with the last term denoting the projection of DZ_N onto F_N and representing $D_N Z_N$ via the isomorphism of F_N with F°/F.

Now by differentiating the equalities $\omega(X_i, Y_j) = \text{constant}$, we have that

$$\omega(X_i, DY_j) = 0 \text{ for all } i \text{ and } j,$$

from which we deduce that DY_j is orthogonal to all the X_j, that is

$$DY_j = \sum_{i=1}^k a_{j,i} X_i + B_j,$$

where B_j is a section of F_N.

Finally, if Z is any section of $i^\star TW$ written as

$$Z = \sum_{i=1}^k \lambda_i X_i + \sum_{i=1}^k \mu_i Y_i + Z_N,$$

we have that

$$DZ = \sum_{i=1}^k \Big(\dot\lambda_i + \sum_{j=1}^k \mu_j a_{j,i}\Big) X_i + \sum_{i=1}^k \dot\mu_i Y_i + \Big(D_N Z_N + \sum_{i=1}^k \mu_i B_i\Big).$$

The differential equation $DZ = 0$ is thus equivalent to the system

$$\begin{cases} \dot\lambda_i + \sum_j a_{j,i}\mu_j + \alpha_i(Z_N) = 0 & \text{for } 1 \leqslant i \leqslant k \\ \dot\mu_i = 0 & \text{for } 1 \leqslant i \leqslant k \\ D_N Z_N + \sum_i \mu_i B_i = 0. & \end{cases}$$

In the vector space of solutions there is the subspace of codimension k of solutions to

$$\begin{cases} \dot\lambda_i + \alpha_i(Z_N) = 0 \quad \text{for } 1 \leqslant i \leqslant k \\ D_N Z_N = 0, \end{cases}$$

corresponding to taking all the μ_j to be zero. These solutions are obtained from the solutions of the normal variational equation $D_N Z_N = 0$ by the quadratures

$$\lambda_i = -\int \alpha_i(Z_N).$$

The solution $Z_N = 0$ yields the constants λ_i, i.e., the vectors X_i. Then if $Z_1, \ldots, Z_{2n-2k}$ is a basis for the solutions of $D_N Z_N = 0$, the expressions

$$-\sum_{i=1}^{k} \Big(\int \alpha_i(Z_\ell) \Big) X_i + Z_\ell$$

are $2n - 2k$ independent solutions (which are also independent of the X_i).

Hence a fundamental matrix of solutions is of the form

$$\left[\begin{array}{c|c|c} \mathrm{Id} & A & \alpha \\ \hline 0 & \mathrm{Id} & 0 \\ \hline 0 & \beta & Z_0 \end{array}\right],$$

where the blocks correspond to the blocks $\langle X_1, \ldots, X_k\rangle, \langle Y_1, \ldots, Y_k\rangle, \langle F_N\rangle$ and Z_0 is a fundamental solution matrix of the normal variational equation.

This shows that the Picard–Vessiot extension of the variational equation is an extension of that of the normal variational equation, and that the Galois group G_N of the normal variational equation is a quotient of the Galois group G of the variational equation.

Thus it is easy to evaluate the Galois group by "solving" the differential system, that is to say by being more precise about the form taken by the matrices A, x, β in the fundamental matrix.

We have seen that the matrix α consists of the primitives of the $-\alpha_i(Z_\ell)$ (for $1 \leqslant i \leqslant k$ and $1 \leqslant \ell \leqslant 2n-2k$). The matrices A and β are obtained by considering the solutions where one of the μ_i is a nonzero constant. We fix an i between 1 and k, choose

$$\mu_j = 0 \quad \text{for } j \neq i \text{ and } \mu_i = 1$$

and solve

$$\begin{cases} \dot\lambda_j^i + a_{j,i} + \alpha_i(Z_N^i) = 0 \qquad \text{for } 1 \leqslant j \leqslant k \\ D_N Z_N^i + B_i = 0. \end{cases}$$

The second equation is solved by utilizing the variation of constants technique which yields solutions of the form

$$Z_N^i = Z_0 \cdot v_0 - \int Z_0 \cdot B_i.$$

We can then solve the first by a new quadrature.

Consider now an element σ of the Galois group G of the variational equation and suppose that it lies in the kernel of the morphism $G \to G_N$. That is, we suppose that σ induces the identity on the Picard–Vessiot extension of the normal variational equation. Let us find out how σ transforms the fundamental matrix written above.

- First, note that σ fixes the X_i obtained for μ_j equal to zero and constant λ_j.
- On the other hand, σ fixes the Z_ℓ and transforms the matrix α of the primitives of the $-\alpha_i(Z_\ell)$ into $\alpha + a$, where a is a constant matrix (which could be identically zero, if the $\alpha_i(Z_\ell)$ have primitives in the field of functions on our curve Γ).
- To finish, consider the effect of σ on solutions of the form

  $$Y_i + \sum_{j=1}^{k} \lambda_j^i X_j + Z_N^i.$$

  We have that $\sigma(Z_N^i) = Z_N^i + Z_0 \cdot b_i$, where b_i is a constant vector in F_N and $\sigma(\lambda_j^i) = \lambda_j^i + C_{i,j} + \alpha_i(Z_0 \cdot b_i)$.

Finally all the elements of G that are sent to the identity in G_N form a subgroup of the group of matrices of the form

$$\left[\begin{array}{c|c|c} \mathrm{Id} & C & a \\ \hline C & \mathrm{Id} & b \\ \hline a & 0 & \mathrm{Id} \end{array}\right],$$

as stated (modulo the ordering of the basis vectors). □

Note that this proof contains the following result (which is classical but not logically necessary here).

Corollary III.4.7. *The Galois group of D_N is a quotient of the Galois group of D by an Abelian group.*

Exercises III.13, III.17.

III.5. The Example of the Hénon–Heiles System

Consider the Hénon–Heiles system, which we have already encountered several times in these notes. The Hamiltonian can be found in Section I.2.d, and, as we saw in Section I.3.c, it is integrable when λ equals 6. We will now show that this case is exceptional and that the system is not integrable in general.

To apply the techniques developed above, we first must choose a trajectory. We have seen (also in Section I.2.d) that the system has solutions in the plane $q_2 = p_2 = 0$. We split our discussion into two cases.

- When $\lambda \neq 0$, we will take
$$q_1(t) = \frac{6}{\lambda t^2}$$
(which is the answer to Exercise I.19).
- If $\lambda = 0$, we will take
$$q_1(t) = \frac{1}{2}t - A.$$

We will consider the *normal* variational equation along this trajectory, which here is the second order equation
$$\ddot{\xi}(t) - 2(A + x(t))\xi = 0$$
(see Example III.2.6), and, specifically:

- the equation $\ddot{\xi}(t) - 2\Big(A + \dfrac{6}{\lambda t^2}\Big)\xi = 0$ for $\lambda \neq 0$
- the equation $\ddot{\xi} - t\xi = 0$ for $\lambda = 0$.

These equations are well categorized by the botanists of differential equation

The $\lambda \neq 0$ Case. We begin by making the small change of variables $t = uz$ to obtain the equation
$$\frac{d^2\xi}{dz^2} - \Big(2Au^2 + \frac{12}{\lambda z^2}\Big)\xi = 0.$$

- If $A = 0$, there is not much to do, the equation keeps its original form and is a "Cauchy equation".
- If $A \neq 0$, we can choose u so that $2Au^2 = 1/4$, which puts the equation in the form
$$\frac{d^2\xi}{dz^2} - \Big(\frac{1}{4} + \frac{12}{\lambda z^2}\Big)\xi = 0.$$
This is a "Whittaker equation", which we have already encountered as a special case of the anharmonic oscillator in Example III.3.14.

The $\lambda = 0$ Case. In this case, the equation is an "Airy equation".

These equations have been categorized, studied, explained, and all together thoroughly explored. The most interesting thing here is to see how we can apply the Morales–Ramis theorem. Since this is the simplest example and since it isabsent from the literature, we will go through this case in detail. Necessary hints and references for the other cases will also be provided.

III.5.a. The Case of the Airy Equation.

Extending Γ. In this case, the curve Γ

$$t \longmapsto \Big(q_1(t) = \frac{1}{2}t - A, 0, p_1(t) = \frac{1}{2}, 0\Big)$$

is a complex line that is globally parametrized by $t \in \mathbb{C}$. The differential field to consider is thus the field of meromorphic functions on $\mathbb{C}$.

From existence theorems for solutions to linear differential equations, we know that the solutions of the equation $\ddot{\xi} - t\xi = 0$ are holomorphic functions on all of $\mathbb{C}$; in other words, they are *entire* functions. In this case, the Picard–Vessiot extension is trivial: all the solutions are already in the base differential field!

Nonetheless, this equation is famous and there are good reasons for this. Most notably, its solutions are entire but certainly not rational (this is an easy exercise; see Exercise III.14). Moreover, if we treat it as a differential equation with coefficients in the "small" field $\mathbb{C}(t)$, then its Galois group is exactly $\mathrm{SL}(2;\mathbb{C})$ (and its Lie algebra is certainly not Abelian; see Section A.5). The readers have probably noticed that $\mathrm{Sp}(2;\mathbb{C}) = \mathrm{SL}(2;\mathbb{C})$ (for the real case, see Exercise I.5).

The ideal would thus be to replace $\Gamma = \mathbb{C}$ by a curve whose function field is exactly the rational function field $\mathbb{C}(t)$. To do this, we just need to add a point at infinity (see Exercise A.4) and consider $\Gamma' = \Gamma \cup \{\infty\}$ instead.

Then we must complete and extend everything as follows:

- We complete $W = \mathbb{C}^4$ by adding a hypersurface at infinity. The simplest way to do this is to add a hyperplane and to consider $W' = \mathbb{P}^4(\mathbb{C})$. The new manifold W' then has homogeneous coordinates
$$[x_1, x_2, y_1, y_2, z] \quad \text{with } q_1 = \frac{x_1}{z}, \text{ etc.}$$
and W corresponds to $z \neq 0$.
- We complete $\Gamma = \big\{[\frac{1}{2}t - A, 0, \frac{1}{2}, 0, 1]\big\}_{t\in\mathbb{C}}$ by adding the point (at infinity) with homogeneous coordinates $[1, 0, 0, 0, 0]$.
- We extend the symplectic form ω to a form ω' on W' (see Exercise III.16). The form ω' is meromorphic and has poles on the hyperplane at infinity (which corresponds to $z = 0$).
- We extend the Hamiltonian H to a meromorphic function on W' which presents no difficulties since H is a polynomial.
- We then ask ourselves if there exists a function K on W that extends to a meromorphic function on W' such that H and K commute. Thus we restrict ourselves to rational first integrals in the coordinates q_i and p_i.

Now, it remains for us to convince ourselves that we can apply a variant of Theorem III.3.10 to this new situation and deduce from it the assertion of the following proposition:

PROPOSITION III.5.1. *The Hamiltonian* $H = \frac{1}{2}(p_1^2 + p_2^2) - q_2^2(A + q_1)$ *has no second first integral that is rational in* (q_1, q_2, p_1, p_2).

PROOF. The normal variational equation along Γ' is an Airy equation. The identity component of its Galois group on $\mathbb{C}(t) = \mathcal{M}(\Gamma')$ is exactly the group $\mathrm{SL}(2;\mathbb{C})$, and thus it is not Abelian (see Section A.5). $\square$

Remark III.5.2. The completed curve Γ is still simply connected, and the corresponding Galois group contains no monodromy (in other words, the solutions are entire). As academic as this example is, it cannot be solved by the methods of [**84**].

III.5.b. The Whittaker Equation. The solution Γ that we have chosen can be parametrized by $t \in \mathbb{C} - \{0\}$ as follows

$$t \longmapsto \left(q_1(t) = \frac{6}{\lambda t^2}, 0, p_1(t) = -\frac{12}{\lambda t^3}, 0\right).$$

The point $q_1 = 0$, $p_1 = 0$ is a singular point for the Hamiltonian and a zero for the corresponding vector field. Therefore, it is natural to consider the curve $\overline{\Gamma}$ obtained by adding this singular point to Γ.

This time, the extended curve $\overline{\Gamma}$ lies in $W = \mathbb{C}^4$, which we do not need to complete. It is parametrized by $t \in (\mathbb{C} - \{0\}) \cup \{\infty\}$.

By applying an *ad hoc* version of their theorem to the curve $\overline{\Gamma}$, Morales and Ramis have explained (see [**62**], and also [**61**]) that when the normal variational equations is a Whittaker equation (that is, when $A \neq 0$), the system is not integrable (with a meromorphic second first integral), except possibly when

$$\frac{6}{\lambda} = \frac{k(k+1)}{2} \quad \text{for some } k \in \mathbb{Z}.$$

Remarks III.5.3.

(1) The case $\lambda = 6$ corresponds to $k = 1$, and so the Morales–Ramis theorem does not forbid its integrability. Of course, this is a good thing, because we found a second first integral for this case in Exercise I.21. In [**61**, Chapter VI], Morales uses a refinement of these methods to show that only when λ equals 1, 2, 6, or 16 can the system be integrable.

(2) We have already encountered a Whittaker equation with an Abelian Galois group in Example III.3.14. In fact, we showed that its Galois group is a subgroup of the group $\mathbb{C}^\star$, by using the integrability of the system in question (the anharmonic oscillator).

(3) As we have already noted, the case $A = 0$ corresponds to a Cauchy equation. The Galois group, at least on $\mathbb{C}(t)$, is Abelian (see Example A.3.3), and so we are unable to conclude anything in this case.

Exercises III.14, III.15, III.16.

III.6. Appendix: Proof of Ziglin's Lemma

I begin with a few explanations on the statement of Theorem III.1.14.

It is about germs, which just means that we are looking at power series: the ring of germs of holomorphic functions at 0 is the ring $\mathbb{C}\{x_1, \ldots, x_m\}$ of convergent power series, while the ring of germs of meromorphic functions is its field of fractions.

The theorem also mentions two different concepts of independence:

- The f_i are functionally independent, i.e., the derivatives must be linearly independent (this is an open condition).

- The initial forms are algebraically independent in the field of rational functions. Now, in $\mathbb{C}(x_1, \dots, x_m)$, the properties of algebraic and functional independence coincide[12].

Another useful remark on independence is the following proposition whose proof will appear along the way to proving Ziglin's lemma.

PROPOSITION III.6.1. *If the initial forms $f_1^\circ, \dots, f_k^\circ$ at 0 are independent, then $f_1, \dots, f_k$ are functionally independent.*

Before giving the proof, here is a simple but enlightening example involving only two functions ($k = 2$) and three variables ($m = 3$).

EXAMPLE III.6.2. We consider the two functions

$$f_1 = x_1^2 \text{ and } f_2 = x_1 + x_1x_2x_3,$$

which are clearly and functionally independent. It is equally clear that their initial forms at 0,

$$f_1^\circ = x_1^2 \text{ and } f_2^\circ = x_1,$$

are dependent, since

$$f_1^\circ = (f_2^\circ)^2 .$$

Then by taking

$$g_1 = f_1 \text{ and } g_2 = f_2^2 - f_1 = x_1^2x_2^2x_3^2 + 2x_1^2x_2x_3,$$

we obtain the independent initial forms

$$g_1^\circ = x_1^2 \text{ and } g_2^\circ = 2x_1^2x_2x_3.$$

Note that g_2 was constructed by using a polynomial annihilating f_1° and f_2°. A generalization of this construction will be used in the proof of the theorem.

We will need some notation and a few lemmas for the proof. We let $d(f) = \deg f^\circ$ ($= \deg u^\circ - \deg v^\circ$ if $f = u/v$) and $d(0) = +\infty$.

LEMMA III.6.3. *Let $f : U \to \mathbb{C}$ be a meromorphic function. Then*

$$d\Big(\frac{\partial f}{\partial x_i}\Big) \geqslant d(f) - 1,$$

with equality if and only if $\dfrac{\partial f}{\partial x_i} \neq 0$.

PROOF. If f is holomorphic at 0, we can write $f = \sum a_\alpha x^\alpha$ and $f^\circ = a_\gamma x^\gamma$, where $|\gamma| = d(f)$ is minimal. Then

$$\frac{\partial f}{\partial x_i} = \sum \alpha_i a_\alpha \frac{x^\alpha}{x_i},$$

[12] This statement is left as an exercise, although not an entirely trivial one. To prove it, one needs to use the fact that any derivative of a field K of characteristic zero can be lifted to a derivative of an extension L/K.

so that

$$d\Big(\frac{\partial f}{\partial x_i}\Big) = \deg\Big(\frac{\partial f}{\partial x_i}\Big)^\circ \begin{cases} = \quad |\gamma|-1 & \text{if } \dfrac{\partial f^\circ}{\partial x_i} \neq 0 \\ > \quad |\gamma|-1 & \text{if } \dfrac{\partial f^\circ}{\partial x_i} = 0. \end{cases}$$

The case where f is only meromorphic is left to the readers... □

... as is the (analogous) proof of the following lemma.

Lemma III.6.4. *Let $f_1, \ldots, f_k$ be meromorphic functions and $i_1, \ldots, i_k$ be integers in $\{1, \ldots, m\}$. Then*

$$d\Big(\frac{D(f_1, \ldots, f_k)}{D(x_{i_1}, \ldots, x_{i_k})}\Big) \geqslant \sum d(f_i) - k,$$

with equality if and only if

$$\frac{D(f_1^\circ, \ldots, f_k^\circ)}{D(x_{i_1}, \ldots, x_{i_k})} \neq 0.$$

With this lemma, we can already prove Proposition III.6.1.

Proof of Proposition III.6.1. If the initial forms $f_1^\circ, \ldots, f_k^\circ$ are independent, there are indices $i_1, \ldots, i_k$ such that

$$\frac{D(f_1^\circ, \ldots, f_k^\circ)}{D(x_{i_1}, \ldots, x_{i_k})} \neq 0.$$

Then Lemma III.6.4 implies that

$$d\Big(\frac{D(f_1, \ldots, f_k)}{D(x_{i_1}, \ldots, x_{i_k})}\Big) = \sum d(f_i) - k < +\infty.$$

Hence the Jacobian cannot be identically zero, and the f_i are independent. □

Lemma III.6.5. *Let $f_1, \ldots, f_k$ be meromorphic functions on U. Define a gradation on the ring of polynomials $\mathbb{C}[z_1, \ldots, z_k]$ by assigning the weight $d(f_i)$ to the variable z_i. Let $P \in \mathbb{C}[z_1, \ldots, z_k]$ be a homogeneous polynomial of weight ν. Then*

$$d(P(f_1, \ldots, f_k)) \geqslant \nu,$$

with equality if and only if

$$P(f_1^\circ, \ldots, f_k^\circ) \neq 0,$$

in which case we have

$$(P(f_1, \ldots, f_k))^\circ = P(f_1^\circ, \ldots, f_k^\circ).$$

Proof. I will only consider the case where all the f_i are holomorphic, leaving the meromorphic case as an exercise. We have

$$f_i = f_i^\circ + \text{higher order terms},$$

and thus

$$P(f_1, \ldots, f_k) = P(f_1^\circ, \ldots, f_k^\circ) + \text{higher order terms}.$$

If the term $P(f_1^\circ, \ldots, f_k^\circ)$ is nonzero, it is the initial form, and we obtain the last equality given in the statement as well as the equality of ν with $d(P(f_1, \ldots, f_k))$. Otherwise, the initial form has higher degree. □

PROOF OF ZIGLIN'S LEMMA. We proceed by induction on k. It suffices thus to carry out the last step. We suppose that $f_1^\circ, \dots, f_{k-1}^\circ$ are algebraically independent. If $(f_1, \dots, f_k)$ are meromorphic functions, we set

$$\mu(f_1, \dots, f_k) = \inf_{i_1, \dots, i_k} d\Big(\frac{D(f_1, \dots, f_k)}{D(x_{i_1}, \dots, x_{i_k})}\Big) - \sum_{i=1}^{k} d(f_i) + k,$$

which is a nonnegative integer by Lemma III.6.4. We prove the "Lemma" by induction on this integer μ. If $\mu = 0$, we have equality in Lemma III.6.4 and thus the initial forms $f_1^\circ, \dots, f_k^\circ$ are independent. Now suppose that $\mu > 0$. We show that one can find a function g_k which is a polynomial in the f_i and is such that

$$\mu(f_1, \dots, f_{k-1}, g_k) < \mu(f_1, \dots, f_k).$$

From Lemma III.6.4, we have that

$$\frac{D(f_1, \dots, f_k)}{D(x_{i_1}, \dots, x_{i_k})} = 0 \quad \text{for any } i_1, \dots, i_k.$$

Hence the initial forms $f_1^\circ, \dots, f_k^\circ$ are functionally related, and thus they are algebraically dependent (a result we decided to accept above). Thus let P be an algebraic relation, i.e., a polynomial in $\mathbb{C}[z_1, \dots, z_k]$ such that $P(f_1^\circ, \dots, f_k^\circ) = 0$. We choose P of minimal degree.

We equip the ring $\mathbb{C}[z_1, \dots, z_k]$ with the gradation of Lemma III.6.5 where the weight of z_i is the degree of f_i°. By decomposing P into a sum of terms homogeneous in this graduation, we can suppose that this polynomial is homogeneous of weight ν.

We have made the hypothesis that $f_1^\circ, \dots, f_{k-1}^\circ$ are algebraically independent thus we know that the polynomial $\partial P/\partial z_k$ is nonzero and that its degree is less than that of P. Since P was chosen to have minimal degree, we have

$$\frac{\partial P}{\partial z_k}(f_1^\circ, \dots, f_k^\circ) \neq 0.$$

Set $g_k = P(f_1, \dots, f_k)$ (see Example III.6.2), and let $i_1, \dots, i_k$ be the indices such that the minimum is achieved in the definition of μ. Then we have

$$\begin{aligned}
&\mu(f_1, \dots, f_{k-1}, g_k) \\
&\quad\leqslant d\Big(\frac{D(f_1, \dots, f_{k-1}, g_k)}{D(x_{i_1}, \dots, x_{i_k})}\Big) - \sum_{i=1}^{k-1} d(f_i) - d(g_k) + k \\
&\quad\leqslant d\Big(\frac{D(f_1, \dots, f_k)}{D(x_{i_1}, \dots, x_{i_k})}\frac{\partial P}{\partial z_k}(f_1, \dots, f_k)\Big) - \sum_{i=1}^{k} d(f_i) + d(f_k) - d(g_k) + k \\
&\quad\leqslant d\Big(\frac{\partial P}{\partial z_k}(f_1, \dots, f_k)\Big) + \bigg(d\Big(\frac{D(f_1, \dots, f_k)}{D(x_{i_1}, \dots, x_{i_k})}\Big) - \sum_{i=1}^{k} d(f_i) + k\bigg) + d(f_k) - d(g_k).
\end{aligned}$$

We now evaluate these terms one by one:

- By the minimality of the degree of P, the polynomial $\partial P/\partial z_k$ is not zero on $f_1^\circ, \dots, f_k^\circ$. Therefore, Lemma III.6.5 implies that
$$\begin{aligned} d\Big(\frac{\partial P}{\partial z_k}(f_1, \dots, f_k)\Big) &= \text{ weight of } \frac{\partial P}{\partial z_k}(z_1, \dots, z_k) \\ &= \text{ weight of } P - \text{ weight of } z_k \\ &= \nu - d(f_k) \end{aligned}$$
by the definition of the gradation.
- From the choice of the indices $i_1, \dots, i_k$, we have
$$d\Big(\frac{D(f_1, \dots, f_k)}{D(x_{i_1}, \dots, x_{i_k})}\Big) - \sum_{i=1}^{k} d(f_i) + k = \mu(f_1, \dots, f_k).$$
- Finally, Lemma III.6.5 implies that $d(g_k) > \nu$ since $g_k = P(f_1, \dots, f_k)$, which is zero on the f_i.

Thus our sequence of inequalities gives
$$\mu(f_1, \dots, f_{k-1}, g_k) < (\nu - d(f_k)) + \mu(f_1, \dots, f_k) + d(f_k) - \nu < \mu(f_1, \dots, f_k),$$
which is exactly what we need to conclude the desired result via induction. □

Exercises

Exercise III.1. Give a proof of Lemma III.1.3 using (well-chosen!) local coordinates on W.

Exercise III.2. Let X be the fundamental vector field for the action of the circle S^1 on a manifold W. Describe the trajectories of X in terms of this action. Fixing a trajectory, describe the solutions of the variational equation.

Exercise III.3. Let U and V be neighborhoods of 0 in $\mathbb{C}^m$, and let $\psi : V \to U$ be a diffeomorphism such that $\psi(0) = 0$. Consider a function f on U. Evaluate $d^2(f \circ \psi)_0$ and show that, if 0 is not a critical point of f,
$$d^2(f \circ \psi)_0 \neq (d^2 f)_0 \circ T_0\psi.$$

Exercise III.4. Check that the initial form H° from Example III.1.8 is constant along the solutions of the linear system given in Example III.1.5.

Exercise III.5. Similarly, check that the initial forms H° and K° from Example III.1.10 (the anharmonic oscillator) are constant along the trajectory Γ.

Exercise III.6. Under the hypotheses of Exercise II.6 on periodic Hamiltonians, show that the Galois group along all the nonstationary trajectories is finite[13].

Exercise III.7 (Singular Action–Angle Variables). Consider a Hamiltonian system written in the (real analytic, symplectic) coordinates
$$x_1, \dots, x_k, \theta_1, \dots, \theta_k, \xi_{k+1}, \dots, \xi_n, \eta_{k+1}, \dots, \eta_n \qquad (k \geqslant 1),$$

[13] In a neighborhood of one these trajectories, a periodic Hamiltonian is always integrable (see Exercise II.6).

$$\begin{cases} \dot{x}_i \ = \ 0 & \dot{\theta}_i \ = \ \dfrac{\partial h}{\partial x_i}(x_1, \ldots, x_k, \tau_{k+1}, \ldots, \tau_n) \\ \dot{\xi}_j \ = \ \dfrac{\partial h}{\partial \tau_j}\eta_j & \dot{\eta}_j \ = \ -\dfrac{\partial h}{\partial \tau_j}\xi_j. \end{cases}$$

Here τ_j denotes the quadratic form

$$\tau_j = \frac{1}{2}(\xi_j^2 + \eta_j^2).$$

Check that $x_1, \ldots, x_k, \tau_{k+1}, \ldots, \tau_n$ are n first integrals in involution[14]. Determine explicitly the trajectories contained in a level set $\tau_{k+1} = \cdots = \tau_n = 0$, and check that they are $\mathbb{R}$ or $\mathbb{R}/\mathbb{Z}$.

Determine the variational equation along such a trajectory and its solutions.

In the case where the trajectory is periodic (although the arbitrarily close ones may not be), describe the Picard–Vessiot extension. Show that the Galois group contains a 1-dimensional additive component and multiplicative components, and that its dimension is $\leqslant 1 + n - k$.

Exercise III.8. Let G be an algebraic subgroup (that is, G is described by polynomial equations) of the additive group $\mathbb{C}^m$. Show that, if G contains a nonzero element z of $\mathbb{C}^m$, G contains the line generated by z. Deduce from this that G is a vector subspace of $\mathbb{C}^m$ and is, in particular, connected.

Exercise III.9. Let Y and Z be endomorphisms of E. Check that their bracket $[Y, Z]$ as vector fields on E is the vector field corresponding to the endomorphism $ZY - YZ$ (the opposite of the bracket of the endomorphisms).

Exercise III.10. Check that an endomorphism Z of E lies in the Lie algebra $\mathfrak{sp}(E)$ if and only if, as a vector field, it fixes the symplectic form (i.e., if and only if $\mathcal{L}_Z \omega = 0$).

Exercise III.11. Consider the vector space $\mathbb{C}^{2n}$ with the complex form $\sum dp_i \ \wedge \ dq_i$.

Letting F be a linear form, what can be said about the corresponding Hamiltonian field? Let F and G be two linear forms. The map

$$Y \longmapsto F(Y)X_G + G(Y)X_F$$

defines a vector field on $\mathbb{C}^{2n}$. Show that it is the Hamiltonian vector field associated to a quadratic Hamiltonian, and determine this Hamiltonian.

Now, suppose that $F_1, \ldots, F_n$ are independent linear forms in involution. Show that there is a symplectic basis of $\mathbb{C}^{2n}$ in which the linear vector field

$$Y \longmapsto F_j(Y)X_{F_i} + F_i(Y)X_{F_j}$$

has a matrix whose only nonzero entries are on the i-th row and the j-th column or the j-th row and the i-th column of the upper right block.

Finally, suppose that $f_1, \ldots, f_n$ are n independent analytic functions in involution and that Γ is a trajectory of X_{f_1} contained in a regular common level set

[14] This type of singular action–angle variables is considered by Eliasson in [**27**]; see also [**23**] and [**44**].

of the f_i. Show that the proof of the key lemma (Lemma III.3.7) implies that the Lie algebra of the Galois group of the variational equation is conjugate to a Lie subalgebra of the Abelian Lie algebra of matrices $\left[\begin{array}{c|c} 0 & 0 \\ \hline B & 0 \end{array}\right] \in \mathfrak{sp}(2n;\mathbb{C})$[15].

Exercise III.12. Consider the Hamiltonian[16]

$$H = \alpha_1 q_1 p_1 + \alpha_2 q_2 p_2$$

on $\mathbb{R}^4$ (where α_1 and α_2 are nonzero constants to be given later). Show that the Hamiltonian system has solutions contained in the plane $q_2 = p_2 = 0$ and choose such a solution Γ. Check that the Riemann surface $\Gamma_{\mathbb{C}}$ is (the germ of) $\mathbb{C}/(\frac{2i\pi}{\alpha_1})\mathbb{Z}$ and that the field of meromorphic functions is the field of fractions of the ring of germs of holomorphic functions in $z = e^{\alpha_1 t}$.

Suppose that α_1/α_2 is a nonzero, non-integral rational number. Determine the Galois group.

More generally, consider the quadratic Hamiltonian

$$H = \alpha_1 q_1 p_1 + \cdots + \alpha_n q_n p_n$$

on $\mathbb{R}^{2n}$ along a trajectory

$$q_1 = ae^{\alpha_1 t}, \qquad p_1 = be^{-\alpha_1 t}, \qquad q_2 = \cdots = q_n = p_2 = \cdots = p_n = 0.$$

and adjust α_i to obtain any product that one would wish of cyclic groups and multiplicative groups as the Galois group.

Exercise III.13. Let k be a differential field on $\mathbb{C}$ and let a be an element of k. Consider the linear differential system

$$\begin{cases} \mu' = 0 \\ z' = a\mu. \end{cases}$$

Show that the Picard–Vessiot extension K has transcendence degree at most 1. We suppose the degree is 1 and there is an element Z of K such that $K = k(Z)$ and $Z' = a$. What are the solutions of the system? Show that the Galois group is conjugated to the "additive group" of matrices $\begin{pmatrix} 1 & c \\ 0 & 1 \end{pmatrix}$.

Exercise III.14. Show that the Airy equation $y'' - ty = 0$ has no rational solution.

Exercise III.15. Let f be an entire function on $\mathbb{C}$ satisfying an algebraic equation with coefficients in $\mathbb{C}(t)$. Show that there exists constants A, B, and C such that

$$|f(t)| \leqslant A + B|t|^C$$

for large enough $|t|$. Deduce from this that f is a polynomial.

[15] Compare this result with Lemma III.3.2.

[16] Like all quadratic Hamiltonians, H is integrable (see [**83**] and Exercise I.26). Here $K = q_2 p_2$, for example, is a second first integral.

Exercise III.16. Show that the form $\omega = dp_1 \wedge dq_1 + dp_2 \wedge dq_2$ on $\mathbb{C}^4$ can be extended to a *meromorphic* form on $\mathbb{P}^4(\mathbb{C})$.

Exercise III.17. Consider an analytic function H on the complex symplectic vector space $\mathbb{C}^4$. We suppose that the plane

$$E : q_2 = p_2 = 0$$

is stable under H in the sense that

$$(\forall\, x \in E) \qquad X_H(x) \in E.$$

Show that H takes the form

$$H(q_1, q_2, p_1, p_2) = h(q_1, p_1) + \frac{1}{2}p_2^2 f + q_2 p_2 \varphi + \frac{1}{2} q_2^2 g$$

for some functions h on $\mathbb{C}^2$, f, φ, and g on $\mathbb{C}^4$.

Choose a trajectory Γ of X_H contained in E that we suppose does not reduce to a point. What is the degree of the initial form H° of H along Γ?

Determine the variational equation along Γ. Show that the reduced symplectic bundle along Γ can be identified with the trivial bundle $\Gamma \times (Q_2, P_2)$, and determine the normal variational equation satisfied by (Q_2, P_2).

Is the Hamiltonian

$$H = \frac{1}{2}(p_1^2 + p_2^2) + \frac{1}{2} q_2^2 (q_1 + 3 q_2 p_2^3)$$

integrable?

Now consider the Hamiltonian

$$H = \frac{1}{2}p_1^2 - A q_1^{2+2r} + q_1^r q_2^2,$$

where $r \geqslant 1$ is an integer. Show that the Hamiltonian system associated to H is integrable. (Try calculating the Poisson bracket of H with a function of q_2 alone.)

Check that there is a value of A such that X_H has a trajectory contained in E of the form

$$q_1 = t^\alpha, \quad p_1 = \alpha t^{\alpha - 1}$$

for some real number α (to be determined). Is it contained in a regular level set of H? For this value of A and trajectory Γ, determine the Galois group of the normal variational equation along Γ.

CHAPTER IV

An Introduction to Lax Equations

The Arnold–Liouville Theorem, Revisited Through Lax Equations

Solving certain completely integrable Hamiltonian systems involves algebraic curves and functions related to these curves: elliptic curves and Weierstrass functions, for example. An elegant way to understand the appearance of these curves is to write the differential system in the form of a Lax equation, namely in the form

$$\frac{dA_\lambda}{dt} = A_\lambda B_\lambda - B_\lambda A_\lambda = [A_\lambda, B_\lambda],$$

where A_λ and B_λ are matrices whose entries are the dynamical variables that interest us. The additional independent parameter λ plays an important role by bringing into play algebraic curves, those of equations $\det(A_\lambda - \mu \,\mathrm{Id}) = 0$ (the spectral curves). With the aid of the eigenvectors of the matrices A_λ, we can finally connect geometry (the family of common level sets of the first integrals) to an algebraic model (the family of Jacobians of the spectral curves). Recall that the Jacobian of a curve is a complex torus, and that its real part is a union of real tori, just like a regular level set of an integrable system.

It is hardly an exaggeration to say that this chapter can be thought of as an *effective* version of Chapter II. I will try to show how these algebro-geometric methods can be used to effectively solve some questions left hanging by the Arnold–Liouville theorem:

- *The affine structure of level sets* (linearization on the Liouville tori). Often the solutions of the differential system can be considered to be "linearized on the Jacobian of the curve"—which can be seen as an algebraic version[1] of the Arnold–Liouville theorem (see Theorem IV.2.5).
- *The determination of the regular values.* These techniques allow us to relate the regularity of a given level set to the smoothness of the corresponding spectral curve, thus easily yielding which values of the first integrals are regular.
- *The counting of the Liouville tori and the study of their bifurcations.* We can also use these methods to count the Liouville tori on a given regular level set, and even to study their bifurcations. This point of view, which was used successfully in [**16**] to determine the Liouville tori of the

[1] But this version is much more precise: the Hamiltonian vector fields are linear for an affine structure on the level sets they themselves define—a tautology that was the basis of Remark II.2.3—and here the Jacobians burst in fully armed... with their canonical affine structure!

Kowalevski top, and in [**6**] to count those of the geodesics on quadrics, is developed in [**8**].

- *Construction of action variables.* Instead of considering a single regular level set, we consider this time the whole family of these level sets (which is an open subset of the symplectic manifold). From the algebraic point of view, we can model this by the family of Jacobians sitting over the family of spectral curves.

In this chapter the readers will find an introduction to Lax equations and spectral curves (Section IV.1), the definition of the eigenvector map and a linearization theorem (Section IV.2), and in an appendix a survey of a method for constructing Lax equations (Section IV.4). The simple example of the spherical pendulum will be used throughout the chapter to illustrate the methods I will present. Section IV.3 contains a fairly complete study of the more elaborate example of geodesics on quadrics.

IV.1. Why Algebraic Curves?

IV.1.a. A Simple Example. We return to the example of the spherical pendulum (Sections I.2.f and I.3.e).

Choose an orthonormal basis whose third vector e_3 is vertical (that is, parallel to the gravitational field Γ): $e_3 = -\Gamma$. Let us determine the height q_3 of the weight as a function of time.

The differential equations are

$$\dot{q} = p \text{ and } \dot{p} = -e_3 - (-q_3 + \|p\|^2)q,$$

and so, in particular, $\dot{q_3} = p_3$. Using the two first integrals $H = \frac{1}{2}\|p\|^2 + q_3$ and $K = q_1p_2 - q_2p_1$ and the relations $\|q\|^2 = 1$ and $p \cdot q = 0$, we can easily eliminate the other unknowns and find that

$$\dot{q}_3^2 = p_3^2 = 2(H - q_3)(1 - q_3^2) - K^2.$$

This leads us to consider the curve with equation $y^2 = P(x)$ with $P(x) = 2(H - x)(1 - x^2) - K^2$.

This is a cubic curve in the plane $\mathbb{C}^2$. By adding a point at infinity we can treat it as an "elliptic curve". We can also easily put this equation in the standard form: $y^2 = 4x^3 - g_2x - g_3$ and parametrize it by a lattice in $\mathbb{C}$ and a Weierstrass $\wp$-function. This allows us to effectively solve the equations (Exercise IV.1).

Since the polynomial P is negative at -1 and 1, it has two real roots in the interval $(-1, 1)$ and one in $(1, +\infty)$. Now, since the weight is restricted to the unit sphere, its height must be between -1 and 1. It is thus limited by the parallels corresponding to the two roots of P contained in this interval (see Figure 5 of Chapter II). From this, we can draw two lessons:

- The elliptic curve $y^2 = P(x)$ arises naturally in the search for solutions.
- The *real* geometry of this algebraic curve gives us information about the dynamics and the geometry of the system.

We will now present a theoretical framework that explains why many completely integrable systems are connected to algebraic curves.

REMARK IV.1.1. There are very simple integrable systems with no hope of being approachable by algebro-geometric methods. Think, for example, of the geodesic flow on a surface of revolution (see Section I.4.a).

Exercises IV.1, IV.2.

IV.1.b. Lax Equations and Algebraic Curves. A Lax equation with spectral parameter is an equation of the form

$$\frac{d}{dt}A_\lambda = [A_\lambda, B_\lambda],$$

where

- A and B are matrices and $[\ ,\ ]$ is the matrix bracket $[A,B] = AB - BA$,
- these matrices are polynomials in λ, an independent variable called the *spectral parameter*,
- B may depend on A (so the equation is not linear).

REMARK IV.1.2. We could also consider Lax equations in which A and B are linear operators that are more general than matrices, for example differential operators. See [**66**] and [**42**] for an introduction and further references.

EXAMPLE IV.1.3. We identify the vectors in $\mathbb{R}^3$ with the skew-symmetric matrices and the vector product with the matrix bracket (as in Exercise IV.3). Then the differential equation

$$\frac{d}{dt}\left(q + \lambda(p \times q) - \lambda^2\Gamma\right) = [q + \lambda(p \times q) - \lambda^2\Gamma, p \times q - \lambda\Gamma]$$

for $(q,p) \in TS^2$, is equivalent to

$$\begin{cases} \dot q = q \times (p \times q) = p \\ \dot p \times q + p \times \dot q = \dot p \times q = -q \times \Gamma, \end{cases}$$

namely, to the system

$$\begin{cases} \dot q = p \\ \dot p = \Gamma + \alpha q = -\Gamma - (q \cdot \Gamma + \|p\|^2)q. \end{cases}$$

We recognize the Hamiltonian system describing the spherical pendulum[2] (Section I.2.f).

EXAMPLE IV.1.4. If Γ, M, L, and Ω are skew-symmetric matrices of order 3, the differential equation

$$\frac{d}{dt}\left(\Gamma + M\lambda + L\lambda^2\right) = \left[\Gamma + M\lambda + L\lambda^2, \Omega + L\lambda\right]$$

[2]This is Alexei Reyman who gave me this Lax equation when I was preparing my Lisbon course and wanted to illustrate it by an example that was not already in [**8**]. Thanks to him!

is a Lax equation with spectral parameter. One can check (Exercise IV.4) that when $M-\Omega$ is collinear to L, this equation is equivalent to the Hamiltonian system describing the motion of a symmetric top (Sections I.2.g and I.3.g).

EXAMPLE IV.1.5 (The Euler–Arnold Equation). In Section I.3.f, we have said that the motion of a free body can be described by the differential system $\dot{M} = M \times \Omega$, where M and Ω are related by $M = \mathcal{J}\Omega$. By treating a vector in $\mathbb{R}^3$ as a skew-symmetric matrix and the vector product as the matrix bracket (see Exercise IV.3 if necessary), we can write this as

$$\dot{M} = [M, \Omega], \quad M = J\Omega + \Omega J,$$

where J is a diagonal matrix (a function of $\mathcal{J}$!). This new way of writing the equation makes sense in dimension N, i.e., we can take M and Ω to be skew-symmetric matrices of order N (and thus elements of $\mathfrak{so}(N)$) and J to be a symmetric matrix.

This system was put in the form of a Lax equation by Manakov [**59**]

$$\frac{d}{dt}\left(M + \lambda J^2\right) = \left[M + \lambda J^2, \Omega + \lambda J\right].$$

It also describes the geodesic flow of an invariant metric on the group $\mathrm{SO}(N)$ (see [**2**, Appendix 2]), which is why this differential equation is called the Arnold–Euler equation.

Exercises IV.13, IV.4, IV.5.

IV.1.c. Isospectral Deformations and the Spectral Curve. The matrix bracket is the infinitesimal form of conjugation. Consequently, a Lax equation is the infinitesimal form of an equality

$$A_\lambda(t) = U_\lambda(t)^{-1} A_\lambda(t_0) U_\lambda(t).$$

Skeptical readers are invited to work out Exercise IV.6. Those who want to be "absolutely sure" should even work out Exercise IV.7. Now that everyone is convinced, the matrix A_λ remains its own conjugate, and, in particular, its eigenvalues are independent of the time t. Thus we say that the Lax equation determines "isospectral" deformations.

The Spectral Curve. A differential system written as a Lax equation always has many first integrals: all the eigenvalues of the matrix A_λ, or, if one prefers, all the coefficients of its characteristic polynomial.

Here is where the spectral parameter λ begins to play an important role: the characteristic polynomial $\det(A_\lambda - \mu \,\mathrm{Id})$ of the matrix is a polynomial in *two* variables λ and μ. And what is a polynomial in two variables if not the equation of an algebraic curve in the plane $\mathbb{C}^2$?

This curve describes the spectrum of A_λ, it is thus called the "spectral curve".

Remark that to each value of λ there are associated N values of μ, the N eigenvalues of the matrix A_λ (with multiplicities)[3], where N is the size of the matrix A_λ. In other words, the spectral curve is a ramified covering of $\mathbb{C}$ of degree N.

[3]From now on, it would be wise to take the polynomials in question as complex and thus as defining complex curves. We will remind ourselves of their real origins when we will need it.

REMARK IV.1.6. The equation $\det(A_\lambda - \mu\,\mathrm{Id}) = 0$ defines an affine curve

$$C_0 = \{(\lambda,\mu) \in \mathbb{C}\times\mathbb{C} \mid \det(A_\lambda - \mu\,\mathrm{Id}) = 0\},$$

but it can be completed and normalized by adding points over $\lambda = \infty$ (see Appendix B). We thus obtain a curve C (a Riemann surface) together with a covering $\lambda : C \to \mathbb{P}^1$. I will discuss this completion more concretely in Section IV.2.c.

REMARK IV.1.7. The spectral curve does not depend on the matrix B. What the latter describes is a differential equation, i.e., a flow on a set of isospectral matrices A.

EXAMPLE IV.1.8. We return to the Lax equation describing the spherical pendulum (Example IV.1.3). The characteristic polynomial of a skew-symmetric matrix A of order 3 is

$$\det(A - \mu\,\mathrm{Id}) = -\mu(\mu^2 + \|A\|^2).$$

Hence we will consider here the equation of the spectral curve[4]

$$\mu^2 + \left\|q + \lambda(p\times q) - \lambda^2\Gamma\right\|^2 = \mu^2 + \lambda^4 + 2K\lambda^3 + 2H\lambda^2 + 1,$$

where we find the first integrals H (energy) and K (momentum with respect to the vertical axis) as expected (see Section I.2.f). This curve, completed by adding two points above $\lambda = \infty$, has genus 1 (see Exercise B.1 if necessary). We remark that it contains no real points; however, this does not prevent it from being useful (see, in particular, Exercise IV.12).

Exercises IV.6, IV.7, IV.8, IV.9, IV.10.

IV.2. A Linearization Theorem

We obtained the spectral curve by considering the eigenvalues of the matrix A_λ. We will now consider and use, even exploit, its eigenvectors.

IV.2.a. Definition of the Eigenvector Map. Fix now the curve C, and consider the isospectral family $\mathcal{T}_C$ of matrices A_λ for which C is the spectral curve (which means fixing the values of the first integrals, i.e., the characteristic polynomial).

Suppose that the polynomials A_λ have a simple spectrum for almost all values of λ, so that the eigenspaces are lines, and that these lines span the space $\mathbb{C}^N$. Then $\lambda : C \to \mathbb{P}^1$ is a ramified covering of degree N.

PROPOSITION IV.2.1. *Let A_λ be a polynomial of matrices over $\mathcal{T}_C$. If the curve C is smooth, then there exists a unique complex line bundle on C defined as the subbundle of $C \times \mathbb{C}^N$ whose fiber at $(\lambda,\mu) \in C$ is, when μ is a simple eigenvalue of A_λ, the eigenspace of A_λ associated to μ.*

[4]We could eliminate the superfluous factor μ by replacing the 3×3 matrices by complex 2×2 matrices, namely via the isomorphism $\mathfrak{so}(3) \cong \mathfrak{su}(2)$. See [**32**] for applications of this remark to studying the top.

PROOF. Consider the submanifold V of $C \times \mathbb{P}^{N-1}(\mathbb{C})$ defined by

$$V = \{(\lambda, \mu, d) \mid d \subset \ker(A_\lambda - \mu \, \mathrm{Id})\}$$

(where d denotes a line in $\mathbb{C}^N$, considered as a point in $\mathbb{P}^{N-1}(\mathbb{C})$) and the submanifold of V

$$V_0 = \{(\lambda, \mu, d) \in V \mid \mu \text{ is a simple eigenvalue of } A_\lambda\}.$$

Let F be the set of points of C where two eigenvalues coincide. Since F is finite under our hypotheses, it must be the ramification locus of the covering $\lambda : C \to \mathbb{P}^1$. We want to show that the bijection

$$\begin{aligned} f : C - F &\longrightarrow V_0 \\ (\lambda, \mu) &\longmapsto (\lambda, \mu, d) \end{aligned}$$

has a holomorphic extension to the entire curve C, with values in the closure $\overline{V}_0$ of V_0.

Let a be a point in F. Since the curve C is smooth, it has a local coordinate z centered at a, and we have, for nonzero z,

$$f(z) = (z, [x_1(z), \ldots, x_N(z)])$$

(where the coordinate z represents the points (λ, μ) of the curve in a neighborhood of a and $[x_1, \ldots, x_N]$ the homogeneous coordinates for $\mathbb{P}^{N-1}(\mathbb{C})$). In this expression, the functions x_i are holomorphic except at 0, where they may have a pole. Thus, for m sufficiently large (but not too large!), we can extend f on a neighborhood of 0 by setting

$$\varphi(z) = (z, [z^m x_1(z), \ldots, z^m x_N(z)]). \qquad \square$$

REMARK IV.2.2. Note that this proof required not only that C is smooth but that it is a curve.

We thus have constructed a map taking values in the Picard group

$$\varphi_C : \mathcal{T}_C \longrightarrow \mathrm{Pic}^d(C)$$

(for a certain integer[5] d). To the polynomial A_λ, φ_C associates the dual L of bundle V of eigenvectors. (We have chosen to utilizer the dual because this is the one who has positive degree: the coordinates of the eigenvectors are holomorphic sections.) We call the map φ_C the *eigenvector map*.

EXAMPLE IV.2.3. It is often possible to produce a divisor which represents the image of a point in $\mathcal{T}_C$. We will see this in the example of geodesics on quadrics, which we will study in Section IV.3.b. In the case where the Lax matrix is a 3×3 skew-symmetric matrix, it is easy to calculate its eigenvectors and obtain a concrete description of the map φ_C. In particular, this is the case for the spherical pendulum (see Exercise IV.11) or for the top (this time, the exercise is worked out in [**8**]).

[5] The degree d is a function of the size of the matrix A_λ and its degree as a polynomial in λ, that it is not indispensable to make explicit here.

We want to understand what $\varphi_C(A_\lambda(t))$ describes in $\operatorname{Pic}^d(C)$ when $A_\lambda(t)$ is a solution to a Lax equation. Since $A_\lambda(t)$ is a solution of a differential equation, we are interested in the differential equation satisfied by the image. Thus we want is to calculate the tangent map associated to φ_C

$$T\varphi_C : T\mathcal{T}_C \longrightarrow T\operatorname{Pic}^d(C).$$

This calculation was performed in great generality by Griffiths [**36**] (see also [**8**, Appendix 3], which closely follows [**36**]). The special case of Lax equations obtained from R-matrices on loop algebras is due to Adler and van Moerbeke [**1**] and Reyman [**70**]. This is the case I explain now in a very down-to-earth way, following the presentation in [**42**].

IV.2.b. The Torus $\operatorname{Pic}^d(C)$ and Its Tangent Space. The tangent map of φ_C takes values in the tangent bundle of $\operatorname{Pic}^d(C)$, so I start by describing this bundle.

We think of a line bundle as being described by a cocycle (see, e.g., Appendix B) and consider the exact "exponential sequence"

$$H^1(C;\mathbb{Z}) \longrightarrow H^1(C;\mathcal{O}_C) \xrightarrow{\exp} H^1(C;\mathcal{O}_C^\star) \xrightarrow{\deg} H^2(C;\mathbb{Z})$$

which presents the vector space $H^1(C;\mathcal{O}_C)$ as a covering of $\operatorname{Pic}^0(C)$, and thus as its tangent space (and also as the tangent space of $\operatorname{Pic}^d(C)$).

If $\beta \in H^1(C;\mathcal{O}_C)$ is the class of a 1-cocycle, then $\exp\beta$ describes a bundle of degree 0. Similarly, if we fix a bundle of degree d described by a cocycle c, then $\exp(t\beta)c$ describes a bundle L_t of degree d that depends *linearly* on t.

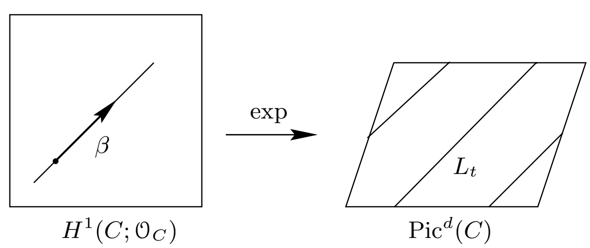


FIGURE 1

This is what I explain now:

- these linear families are the images of the solutions of a certain Lax equation,
- more precisely, to each β I will associate a matrix B_λ such that the bundle defined by $\exp(t\beta)c$ is the image of a solution $A_\lambda(t)$ of $\dot{A}_\lambda = [A_\lambda, B_\lambda]$,
- to do this, I will give a fairly concrete description of the elements of $H^1(C;\mathcal{O}_C)$ by using an embedding of C in a complex surface S.

IV.2.c. The Spectral Curve As Seen in $\mathcal{O}(n)$. The surface S in which we embed the spectral curve C is simply the total space of the bundle $\mathcal{O}(n)$ on $\mathbb{P}^1$, where n is the degree of A_λ as a polynomial in λ.

Let $U_0 = \mathbb{P}^1 - \{\infty\}$ and $U_\infty = \mathbb{P}^1 - \{0\}$. Recall (see Appendix B) that the bundle $\mathcal{O}(n)$ can be obtained by gluing $U_0 \times \mathbb{C}$ and $U_\infty \times \mathbb{C}$ via

$$(\lambda, \mu) \longmapsto \Big(\frac{1}{\lambda}, \frac{\mu}{\lambda^n}\Big).$$

And the notation is good: λ is the "affine" coordinate (centered at 0) and setting $\lambda' = 1/\lambda$ (the coordinate centered at ∞), we have

$$\begin{aligned}\det(A_\lambda - \mu\,\mathrm{Id}) &= \det\Big(\sum_{i=0}^{n} A_i\lambda^i - \mu\,\mathrm{Id}\Big)\\ &= \det\Big(\frac{1}{\lambda'^n}\Big(\sum_{i=0}^{n} A_i\lambda'^{n-i} - \lambda'^n\mu\,\mathrm{Id}\Big)\Big).\end{aligned}$$

Hence we can consider any point (λ, μ) in C as a point in the total space of $\mathcal{O}(n)$ allowing us to see the curve C as embedded in the surface S.

We can also interpret this embedding by saying that μ is a section of the bundle $\lambda^\star\mathcal{O}(n)$ on C.

REMARK IV.2.4. This is indeed an embedding of the *complete* curve C. The partial compactification of $\mathbb{C}^2$ in the total space of the bundle $\mathcal{O}(n)$ on $\mathbb{P}^1(\mathbb{C})$ is actually the most natural place to construct the completion of the affine curve.

IV.2.d. The Elements of $H^1(\mathcal{O}_C)$ as Polynomials in (λ, μ). The advantage of seeing C in S is that $H^1(S; \mathcal{O}_S)$ can be described simply, thereby giving us a very concrete description of the elements of $H^1(C; \mathcal{O}_C)$.

The surface S is the union of the two open sets $U_0 \times \mathbb{C}$ and $U_\infty \times \mathbb{C}$. We show that $H^1(U_0 \times \mathbb{C}, \mathcal{O}_S)$, $H^1(U_\infty \times \mathbb{C}, \mathcal{O}_S)$, and $H^1((U_0 \cap U_\infty) \times \mathbb{C}, \mathcal{O}_S)$ are zero, so we can calculate $H^1(S; \mathcal{O}_S)$ simply by using the 1-cocycles of the covering of S by these two open sets. In this way $H^1(S; \mathcal{O}_S)$ can be interpreted as the vector space of holomorphic functions on $\mathbb{C}^\star \times \mathbb{C}$, that is to say, the series

$$\sum_{i\in\mathbb{Z}}\sum_{j\in\mathbb{N}} a_{i,j}\lambda^i\mu^j$$

modulo those that can be extended to $U_0 \times \mathbb{C}$, i.e., the series $\sum_{i\in\mathbb{N}}\sum_{j\in\mathbb{N}} a_{i,j}\lambda^i\mu^j$, and those that can be extended to $U_\infty \times \mathbb{C}$, i.e., the series

$$\sum_{i\in\mathbb{N}}\sum_{j\in\mathbb{N}} a_{i,j}\lambda'^i\mu^j = \sum_{i\in\mathbb{N}}\sum_{j\in\mathbb{N}} a_{i,j}\lambda^{-nj-i}\mu^j.$$

We now consider the embedding of C in S. The polynomial $\det(A_\lambda - \mu\,\mathrm{Id})$ has degree nN, and the functions on S which are 0 on C are those in the ideal generated by this polynomial. We can write this as an exact sequence of sheaves

$$0 \longrightarrow \mathcal{O}_S(-nN) \xrightarrow{\det(A_\lambda - \mu\,\mathrm{Id})} \mathcal{O}_S \longrightarrow \mathcal{O}_C \longrightarrow 0.$$

Since S is covered by two open sets, the group $H^2(S;\mathcal{O}(-nN))$ is trivial. Thus the inclusion $C \to S$ induces a surjection

$$H^1(S;\mathcal{O}_S) \longrightarrow H^1(C;\mathcal{O}_C)$$

and we can say more precisely that the elements of $H^1(C;\mathcal{O}_C)$ can be described as elements of $H^1(S;\mathcal{O}_S)$ modulo the equation of C, that is $\det(A_\lambda - \mu\,\mathrm{Id}) = 0$.

In conclusion, any element of $H^1(C;\mathcal{O}_C)$ can be represented as a Laurent polynomial of the form

$$\beta(\lambda,\mu) = \frac{1}{\lambda^k}\sum_{i=0}^{N-1} b_i(\lambda)\mu^i$$

for an integer k and polynomials b_i.

IV.2.e. Cocycles and Lax Equations. As we have already noted, for a cocycle c determining a bundle of degree d on C, that $\exp(t\beta)c$ also determines a bundle L_t of degree d. As we announced it, we will identify L_t with the image of a matrix $A_\lambda(t)$.

THEOREM IV.2.5. *Let* $\beta(\lambda,\mu) = \dfrac{1}{\lambda^k}\sum_{i=0}^{N-1} b_i(\lambda)\mu^i$ *be an element of* $H^1(C;\mathcal{O}_C)$*, and let*

$$B_\lambda = \Big(\frac{1}{\lambda^k}\sum_{i=0}^{N-1} b_i(\lambda)A_\lambda^i\Big)_+$$

be the polynomial part of $\beta(\lambda, A_\lambda)$*. Then the image of a solution* $A_\lambda(t)$ *of the Lax equation* $\dot A_\lambda = [A_\lambda, B_\lambda]$ *under the eigenvector map* φ_C *is a bundle* L_t *determined by a cocycle* $\exp(t\beta)c$*.*

EXAMPLE IV.2.6. Continuing to consider the Lax equation for the spherical pendulum (see Example IV.1.3), we have

$$A_\lambda = q + \lambda(p\times q) - \lambda^2\Gamma \text{ and } B_\lambda = p\times q - \lambda\Gamma = \big(\lambda^{-1}A_\lambda\big)_+ .$$

Thus the eigenvector map linearizes the flow on the Jacobian of C. Here the cocycle β is $\lambda^{-1}\mu$.

REMARKS IV.2.7.

(1) The choice of c is related to the choice of the initial condition $\varphi_C(A_\lambda(0))$.
(2) As Griffiths explains in [**36**], we expect *a priori* that B_λ is not unique (we could add any polynomial in A_λ) and this foreshadows the cohomological nature of the objects in question.
(3) This result has some beautiful applications. For example, one can use it to determine the regular values of an integrable system. See Exercise IV.12 for the simple case of the spherical pendulum and Section IV.3 for the case of geodesics on quadrics.

PROOF. For this proof, it is simpler to work with the eigenvectors themselves. Thus let $v(t)$ be a section of $L_t^\star$. Abusing notation, I will also write v for the section over U_0 and v' the same section over U_∞. In this way we have

$$v = \exp(-t\beta(\lambda,\mu))\gamma v'$$

(with the $-$ sign since v is a section of $L_t^\star$, and with γ denoting a cocycle that determines a bundle of degree $-d$, for example $\gamma = c^{-1}$). Differentiating with respect to t yields

$$\dot{v} = -\beta v + \exp(-t\beta)\gamma\dot{v}'.$$

But since v is an eigenvector of A_λ with eigenvalue μ, i.e., $A_\lambda v = \mu v$, we have that

$$\beta(\lambda,\mu)v = \beta(\lambda, A_\lambda)v.$$

We decompose $\beta(\lambda,\mu)$ into a polynomial component and a "Laurent polynomial without constant term" component

$$\beta(\lambda,\mu) = \beta(\lambda,\mu)_+ - \beta(\lambda,\mu)_-$$

and thus we see that

$$\begin{aligned}\dot{v} + \beta(\lambda, A_\lambda)_+ v &= \beta(\lambda, A_\lambda)_- v + \exp(-t\beta(\lambda,\mu))\gamma\dot{v}' \\ &= \exp(-t\beta)\gamma\left(\dot{v}' + \beta(\lambda, A_\lambda)_- v'\right).\end{aligned}$$

In this equality, the left side is holomorphic in λ and the right side is holomorphic in λ^{-1}, so that the equality defines a section of $L_t^\star$. Finally, $\beta(\lambda, A_\lambda)_+$ is what we want to call B_λ so that this section can be written over U_0 as

$$\dot{v} + B_\lambda v.$$

We have assumed that the eigenvalues of A_λ are, generically, distinct, so that $\dot{v} + B_\lambda v$ is an eigenvector of A_λ with eigenvalue μ, and thus

$$\begin{cases} A_\lambda v = \mu v \\ A_\lambda(\dot{v} + B_\lambda v) = \mu(\dot{v} + B_\lambda v). \end{cases}$$

This implies that

$$A_\lambda \dot{v} + (A_\lambda B_\lambda v - B_\lambda A_\lambda v) - \mu\dot{v} = 0,$$

or, in other words,

$$A_\lambda \dot{v} + [A_\lambda, B_\lambda]v - \mu\dot{v} = 0.$$

By comparing this equality to the one we get by differentiating $A_\lambda v - \mu v = 0$ with respect to the time t, we get

$$\dot{A}_\lambda v = [A_\lambda, B_\lambda]v.$$

Then, since we can find a basis of eigenvectors for A_λ, we see that A_λ does in fact satisfy the differential equation

$$\dot{A}_\lambda = [A_\lambda, B_\lambda]$$

for the matrix B_λ we have constructed. □

Exercises IV.11, IV.12.

IV.3. The case of Geodesics on Quadrics

In this section, I utilize and illustrate the methods presented so far to determine the regular values and action variables for geodesics on quadrics. Notation is that of Section I.4.b. The matrix Γ is the matrix with entries $\Gamma_{i,j} = (p_i u_j - p_j u_i)/\alpha_i\alpha_j$.

IV.3.a. The Lax Equation and the Spectral Curve. The system is described (by Moser in [**65**], and) by Lax equations obtained by the method explained in Section IV.4. We define $p \otimes p$ and $p \times u$ to be the matrices with entries

$$(p \otimes p)_{i,j} = p_i p_j, \qquad (p \wedge u)_{i,j} = p_i u_j - p_j u_i.$$

It is easy to prove the following result (Exercise IV.14).

PROPOSITION IV.3.1. *The Hamiltonian system describing the geodesics of the quadric Q_0 on the space $\mathcal{D}_n$ of lines of $\mathbb{R}^{n+1}$ is equivalent to the Lax equation $\dot{A}_\lambda = [A_\lambda, B_\lambda]$, where*

$$A_\lambda = \alpha\lambda^2 + (p \wedge u)\lambda - p \otimes p \text{ and } B_\lambda = \alpha^{-1}\lambda + \Gamma.$$

Exercise IV.14.

We now determine the spectral curve. The matrix $p \otimes p$ is the matrix of the orthogonal projection onto the line generated by p, while the image of the matrix $p \wedge u$ is the subspace generated by p and u. Thus A_λ is a perturbation of the diagonal matrix α by terms of rank at most 2. It is easy to calculate its characteristic polynomial[6] (see [**65**]).

Considering the form of the matrix A_λ, it makes sense to call the eigenvalue $\lambda^2 z$. We then write

$$A_\lambda - \lambda^2 z \operatorname{Id} = \lambda^2(\alpha - z \operatorname{Id})\left[\operatorname{Id} + \lambda^{-1} \underbrace{(\alpha - z \operatorname{Id})^{-1}(p \wedge u - \lambda^{-1} p \otimes p)}_{C}\right].$$

We choose a basis $e_1, \dots, e_{n+1}$ to write

$$\begin{aligned}\det(\operatorname{Id} + \lambda^{-1}C) e_1 \wedge \cdots \wedge e_{n+1} &= (\operatorname{Id} + \lambda^{-1}C)^{\wedge n+1}(e_1 \wedge \cdots \wedge e_{n+1}) \\ &= e_1 \wedge \cdots \wedge e_{n+1} + \lambda^{-1}(\operatorname{Tr} C) e_1 \wedge \cdots \wedge e_{n+1} \\ &\quad + \lambda^{-2}(\operatorname{Tr} C^{\wedge 2}) e_1 \wedge \cdots \wedge e_{n+1}\end{aligned}$$

(rank $C \leqslant 2$, implying that $C^{\wedge 3} = 0$, thus there are no other terms). But $\alpha - z \operatorname{Id}$ is diagonal and $p \wedge u$ has no diagonal term, so the trace of the matrix $(\alpha - z \operatorname{Id})^{-1} p \wedge u$ must be zero. Hence

$$\operatorname{Tr} C = -\lambda^{-1} \operatorname{Tr}((\alpha - z \operatorname{Id})^{-1} p \otimes p) = -\lambda^{-1} q_z(p),$$

where q_z denotes the quadratic form $\sum x_i^2/(\alpha_i - z)$ (we use the notation φ_z for its polar form). Similarly, $p \otimes p$ has rank 1 and we compute

$$\operatorname{Tr} C^{\wedge 2} = \operatorname{Tr}\left(\left((\alpha - z \operatorname{Id})^{-1} p \wedge u\right)^2\right) = q_z(p) q_z(u) - \varphi_z(p, u)^2.$$

Finally,

$$\begin{aligned}\det(A_\lambda - \lambda z \operatorname{Id}) &= \lambda^{2n+2} \det(\alpha - z \operatorname{Id})\left(1 - \lambda^{-2}(q_z(p) - (q_z(p) q_z(u) - \varphi_z(p, u)^2))\right) \\ &= \lambda^{2n+2} \det(\alpha - z \operatorname{Id})\left\{1 - \lambda^{-2}(\underbrace{\varphi_z(p, u)^2 - q_z(p)(q_z(u) - 1)}_{\Psi_z(p,u)})\right\}.\end{aligned}$$

[6]In [**65**], Moser considers more generally Lax equations with a matrix A_λ that is a perturbation of rank at most 2 of an invertible diagonal matrix (which includes the Euler–Arnold equation, which we discussed in Example IV.1.5).

The rational function $\Psi_z(p,u)$ can be written as

$$\Psi_z(p,u)=\frac{Q_{p,u}(z)}{\prod(\alpha_i-z)},$$

where $Q_{p,u}$ is a polynomial of degree n in the variable z whose coefficients are functions in p and u.

Remark IV.3.2. The rational function $\Psi_z(p,u)$ is just the discriminant of the quadratic polynomial (in t)

$$q_z(u+tp)-1=f_z(u+tp),$$

which we discussed in Section I.4.b. Thus $\Psi_z(p,u)=0$ if and only if the line of the $\{u+tp\}$ is *tangent* to the quadric Q_z. Viewed as a function of the space of lines $\mathcal{D}_n$ in $\mathbb{R}$, Ψ_z is the Hamiltonian of the system describing the geodesics on the quadric Q_z.

We can also decompose

$$\frac{Q_{p,u}(z)}{\prod(\alpha_i-z)}=\sum_{i=1}^{n+1}\frac{F_i(p,u)}{\alpha_i-z}$$

with

$$F_i(p,u)=p_i^2+\sum_{j\neq i}\frac{(p_iu_j-p_ju_i)^2}{\alpha_i-\alpha_j}\qquad\text{for } 1\leqslant i\leqslant n+1.$$

Note that the $n+1$ functions F_i satisfy $\sum F_i=1$.

Remark IV.3.3. In Section I.4.b, we utilized the roots $z_1,\dots,z_n$ of the polynomial $Q_{p,u}$ as first integrals. The coefficients of the characteristic polynomial are also first integrals as are the functions F_i, the “Uhlenbeck integrals”. The z_i are in involution, similarly so are the F_i. This is also a consequence of Proposition IV.4.1 and of the construction of the Lax equation.

The Spectral Curve. The spectral curve is the completion C (normalized at infinity) of the affine curve defined by

$$\lambda^2=\Psi_z(p,u)=\frac{Q_{p,u}(z)}{\prod(\alpha_i-z)}.$$

The function $z:C\to\mathbb{P}^1$ is a double covering of $\mathbb{P}^1$ ramified at the α_i, at the roots of $Q_{p,u}$, and at ∞ (there is only one point over $z=\infty$ on C).

Thus the curve C is hyperelliptic with genus n. Let τ denote the hyperelliptic involution $\tau(\lambda,z)=(-\lambda,z)$.

I will always think of C as the completed curve, even when I use the affine equation. Note that C is a real curve (its equation is real). When there is a risk of confusion, I will call $C_{\mathbb{R}}$ its real part.

Remark IV.3.4. We have $Q(\alpha_i)=F_i$, thus if $F_i(p,u)=0$, α_i is a root of $Q_{p,u}$ and the corresponding curve is singular or, more precisely, it is no longer a curve of genus n.

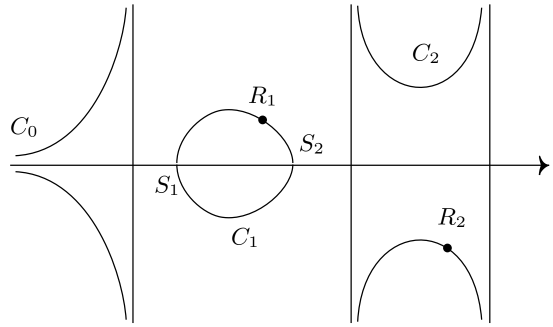


Figure 2. The real part of the spectral curve

The Real Part of the Spectral Curve. We mention some properties of the real part of the curve C, which will be useful for determining the action variables.

Proposition IV.3.5. *If the curve C is smooth with genus n, then its real part $C_{\mathbb{R}}$ has $n+1$ connected components $C_0, \ldots, C_n$ and $z(C_i) \subset (\alpha_i, \alpha_{i+1})$ (where $\alpha_0 = \infty$).*

Remark IV.3.6. As we noted in Section I.4.b, Chasles' theorem implies that the roots of the polynomial $Q_{p,u}$ are real. What we have here is actually a bit more precise, since we know where these roots are relative to the α_i. We can reformulate the proposition by saying that the roots $z_1, \ldots, z_n$ of p,u are real and that if we order $\{z_1, \ldots, z_n, \alpha_1, \ldots, \alpha_{n+1}\}$ as $-\infty < b_1 < \cdots < b_{2n+2}$, then $z_i = b_{2i-1}$ or b_{2i}. This is a classic result [**50, 67, 6**].

Recall (again from Section I.4.b) that the roots $z_1, \ldots, z_n$ of $Q_{p,u}$ determine n quadrics $Q_{z_1}, \ldots, Q_{z_n}$ of the confocal family defined by the α_j, i.e., the ones to which the line in question (an element of $\mathcal{D}_n$) is tangent. The position of the z_i with respect to the α_j determines the type of the real quadrics Q_{z_i}. For example, the configuration shown in Figure 2,

$$\alpha_1 < z_1 < z_2 < \alpha_2 < \alpha_3$$

corresponds to two hyperboloids of one sheet.

Proof. We begin by determining some real points on C. Instead of considering $\Psi_z(p,u)$ in the form $Q(z)/\prod(\alpha_i - z)$, we use the form given in its definition (see Section I.4.b)

$$\Psi_z(p,u) = \varphi_z(p,u)^2 - q_z(p)[q_z(u) - 1].$$

We immediately see that the equation of C is satisfied, for example, by $\lambda = \pm\varphi_z(p,u)$ with $q_z(p) = 0$ or with $q_z(u) = 1$. If $\zeta_1, \ldots, \zeta_n$ are the n solutions of $q_z(p) = 0$, we let R_i denote the point on C with coordinates $(\lambda, z) = (-\varphi_{\zeta_i}(p,u), \zeta_i)$. We write

$$\sum \frac{p_i^2}{\alpha_i - z} = \frac{\mathcal{R}(z)}{\prod(\alpha_i - z)} \text{ with } \mathcal{R}(z) = \sum_{i=1}^{n+1} p_i^2 \prod_{j \neq i} (\alpha_i - z)\,,$$

so that $\mathcal{R}(\alpha_i) = p_i^2 \prod_{j \neq i} (\alpha_i - \alpha_j)$. Thus $\mathcal{R}(\alpha_1) > 0$, the signs of $\mathcal{R}(\alpha_i)$ alternate, and $\mathcal{R}$ has a real root in each interval $[\alpha_i, \alpha_{i+1})$ (with $\zeta_i = \alpha_i$ only if $p_i = 0$).

We now consider the roots of Q. We simply look at the sign of $\Psi - z$:

- We have $\Psi_{-\infty} > 0$ and $\Psi_{+\infty} < 0$.
- The sign changes at each α_i and at each of the real roots of Q (and nowhere else).
- We know that $R_i \in C_{\mathbb{R}}$ and, by definition, that $\Psi_{\zeta_i}(p, u) = \varphi_{\zeta_i}(p, u)^2 \geqslant 0$.

Thus we observe that the sign of Ψ_z changes at least two times in each interval $(-\infty, \zeta_1]$, $[\zeta_1, \zeta_2)$, ..., $[\zeta_{n-1}, \zeta_n]$, hence that Q has at least one root in each of these intervals. Since its degree is n, this is enough to prove the proposition. □

Remark IV.3.7. The (real) points $R_1, \ldots, R_n$ depend not only on the level sets of the first integrals associated to the curve C but also on the point (p, u) used to define them. These will reappear below as a form of the eigenvector map.

Later, we will use the points $S_1, \ldots, S_n$ of C whose coordinates are $(0, z_1), \ldots, (0, z_n)$, respectively. Figure 2 represents one of the four $(= 2^n)$ possible configurations of the real part of the curve C, with the points R_i and S_i for the case when $n = 2$.

IV.3.b. The Divisor Associated to the Eigenvectors. We will now determine the image of the point (p, u) under the eigenvector map as an explicit form of the divisor of the zeros of a *holomorphic* section of the vector bundle dual to the eigenvector bundle. Equivalently, we will construct a *nowhere zero* meromorphic section of the eigenvector bundle and determine the divisor of its poles.

Concretely, we want to understand how an eigenvector of the matrix A_λ can become infinite. Thus we fix a point $(p, u) \in \mathcal{D}_n$. Any vector in $\mathbb{R}^{n+1}$ can be written uniquely as $\alpha p + \mu u + v$, where v is orthogonal to the subspace generated by p and u. To ensure that the eigenvector is nonzero, we normalize by taking $\alpha = 1$. Then we look for a vector of the form

$$w = p + \mu(\lambda, z) u + v(\lambda, z) \text{ in the kernel } \ker(A_\lambda - \lambda^2 z \operatorname{Id}).$$

More precisely, we look for the divisor of the poles of such a section of L.

Since $\|p\| = 1$, $p + \mu u + v$ can be infinite only if $\mu u + v$ is. So for a given (p, u), we look for the values of (λ, z) such that $\ker(A_\lambda - \lambda^2 z \operatorname{Id})$ contains a vector orthogonal to p.

We recall that $A_\lambda - \lambda^2 z \operatorname{Id}$ has the form $-p \otimes p + \lambda M_{\lambda, z}$, and that $p \otimes p$ is zero on the vectors orthogonal to p. Thus we look for the (λ, z) on C such that

$$\ker(\lambda M_{\lambda, z}) \cap p^\perp \neq 0,$$

where $M_{\lambda, z} = p \wedge u + \lambda(\alpha - z)$.

We first look at the case where $\lambda = 0$. Then $z = \infty$ or $Q_{p,u}(z) = 0$, and the eigenvectors of the matrix $A_\lambda - \lambda^2 z \operatorname{Id} = -p \otimes p$ are

- the vector p for the eigenvalue -1 and
- all the vectors in $p^\perp$ for the eigenvalue 0.

The eigenvalue is

$$\mu = \lambda^2 z = \frac{zQ(z)}{\prod(\alpha_i - z)} = \begin{cases} -1 & \text{if } z = \infty \\ 0 & \text{if } Q(z) = 0, \end{cases}$$

but in the first case, our normalization implies that the eigenvector is p and thus is bounded. On the other hand, if z_i denotes a root of Q (i.e., the point on the curve is one of the S_i), our eigenvector tends to infinity, since it is, as it must be, orthogonal to p and normalized to $p + \mu + v$, the limits of the eigen-directions are the n normals of the n quadrics defined by $z_1, \dots, z_n$. These n vectors are usually called the Chasles frame: another theorem of Chasles asserts that these n normals are pairwise orthogonal.

Finally, we suppose that $\lambda \neq 0$. Then we look for the kernel of the restriction of $M_{\lambda,z}$ to $p^\perp$. Explicitly, we have

$$\begin{aligned} M_{\lambda,z}(\mu u + v) &= \mu(p \wedge u)(u) + \lambda(\alpha - z)(\mu u + v) \\ &= \mu \|u\|^2 p + \lambda (\alpha - z) (\mu u + v) . \end{aligned}$$

If $\mu = 0$ or $u = 0$, then $M_{\lambda,z}(v) = \lambda(\alpha - z)v$ is zero only if $v = 0$. Thus we can suppose that $\mu = 1$ and $u \neq 0$. Our equation then becomes

$$\|u\|^2 p + \lambda (\alpha - z) (u + v) = 0,$$

that is

$$v = -\frac{\|u\|^2}{\lambda} (\alpha - z)^{-1} p - u$$

with $v \cdot p = 0$, that is, $\left[(\alpha - z)^{-1} p\right] \cdot p = 0$, so $q_z(p) = 0$, and with $v \cdot u = 0$, we finally see that $\lambda = -\varphi_z(p, u)$. We have thus retrieved the points R_i which we used to describe the real part of C and have proved the following proposition.

Proposition IV.3.8. *The image $\varphi_C(p, u)$ is the class of the divisor*

$$R_1 + \cdots + R_n + S_1 + \cdots + S_n$$

in $\operatorname{Pic}^{2n} C$.

The divisor $S_1 + \cdots + S_n$ does not depend on (p, u), so we can content ourselves with $R_1 + \cdots + R_n$ and its class in $\operatorname{Pic}^n(C)$. But we know that the points R_i are real and distributed among the components $C_1, \dots, C_n$ of $C_{\mathbb{R}}$.

Proposition IV.3.9. *The eigenvector map*

$$\varphi_C : \mathcal{T}_C \longrightarrow \operatorname{Pic}^n(C)$$

takes its values in the real torus $C_1 \times \cdots \times C_n \subset \operatorname{Pic}^n(C)$.

Proof. The only thing which is left to show that the real torus $C_1 \times \cdots \times C_n$ is actually embedded in $\operatorname{Pic}^n(C)$. We see it as a subspace of the Cartesian product $C_{\mathbb{R}}^n$, or even as a subspace of the symmetric product $C_{\mathbb{R}}^{(n)} = C_{\mathbb{R}}^n / \mathfrak{S}_n$ by utilizing the ordering on $\mathbb{R}$. This symmetric product is a subset of the symmetric product $C^{(n)}$

of the complex curve, which can be mapped into the Picard group via the linear equivalence relation. In short, we have

$$C_1 \times \cdots \times C_n \subset C_{\mathbb{R}}^{(n)} \subset C^{(n)} \xrightarrow{u} \operatorname{Pic}^n(C),$$

where only the last arrow is not injective. However, the following lemma implies that its restriction to $C_1 \times \cdots \times C_n$ is injective.

LEMMA IV.3.10. *Let V be the open set of $C^{(n)}$ defined by*

$$V = \{P_1 + \cdots + P_n \mid P_i \neq \tau(P_j) \text{ and } z(P_i) \neq \infty\}.$$

The restriction of the map u to V is injective.

This is a direct application of the Riemann–Roch theorem, its proof will be given below in Section IV.3.e. It is clear that $C_1 \times \cdots \times C_n$ is contained in V, since the C_i are stable under the hyperelliptic involution τ and C_0, which does not appear here, is the only component of $C_{\mathbb{R}}$ containing points with z infinite. □

Starting from Proposition IV.3.9 we will be able to treat the eigenvector map φ_C as one with values in $C_1 \times \cdots \times C_n$. In particular, we can forget the Jacobian and the complex curve.

IV.3.c. Regular Values. The family of first integrals $F_1, \ldots, F_{n+1}$ define a map

$$F : \mathcal{D}_n \longrightarrow E,$$

where E is the hyperplane with equation $\sum x_i = 1$ in $\mathbb{R}^{n+1}$. Here is a nice way of applying Theorem IV.2.5.

THEOREM IV.3.11. *A point in E is a regular value of F if and only if the corresponding spectral curve is smooth with genus n.*

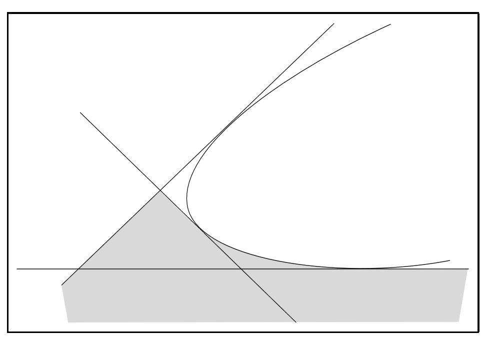


FIGURE 3. Regular values for $n = 2$

REMARK IV.3.12. Although it can be tedious to determine for which values of (p, u) the map F has rank n, it is easy to decide when the curve C is smooth with genus n: for this to be true it is necessary and sufficient that the ramification points of z are distinct, that is to say, if $Q_{p,u}$ has n distinct roots that are distinct from the α_i.

The critical values of the map F are thus on the discriminant of the polynomials $Q_{p,u}$ and on the hyperplanes $F_i = 0$, which are tangent to it. Figure 3 shows these critical points for $n = 2$, that is to say, in the case of quadrics in $\mathbb{R}^3$. The shaded region consists of the values that are effectively attained, i.e., those that correspond to the z_i that satisfy the conditions of Proposition IV.3.5 or Remark IV.3.6.

Remark IV.3.13. As already noted in Example II.2.8, the connected components of the set of regular values are contractible.

The proof of Theorem IV.3.11 is based on the following lemma, which describes the tangent map associated to the eigenvector map φ_C.

Lemma IV.3.14. *The images under the tangent map associated to φ_C of the Hamiltonian vector fields associated to the functions $F_1, \dots, F_n$ are independent vectors in the tangent space $H^1(C; \mathcal{O}_C)$ of $\operatorname{Pic}(C)$.*

I will prove this lemma at the end of this section (Section IV.3.e) by utilizing the linearization theorem (Theorem IV.2.5) via the following lemma:

Lemma IV.3.15. *The matrix B_λ satisfies $B_\lambda = \left(\lambda^3 A_\lambda^{-1}\right)_+$.*

Proof. First, we note that this takes place in the space of matrices of order $n+1$ and that A_λ^{-1} is indeed a polynomial in A_λ. The proof is then a simple calculation

$$\begin{aligned} A_\lambda &= \lambda^2(\alpha + (p \wedge u)\lambda^{-1} - (p \otimes p)\lambda^{-2}) \\ &= \lambda^2 \alpha \underbrace{(\operatorname{Id} + \alpha^{-1}(p \wedge u)\lambda^{-1} - \alpha^{-1}(p \otimes p)\lambda^{-2})}_{D}. \end{aligned}$$

Thus we have $A_\lambda^{-1} = \lambda^{-2} D^{-1} \alpha^{-1}$ with

$$D^{-1} = \operatorname{Id} - \alpha^{-1}(p \wedge u)\lambda^{-1} \mod \lambda^{-2},$$

and, finally,

$$\left(\lambda^3 A_\lambda^{-1}\right)_+ = \lambda\alpha^{-1} - \alpha^{-1}(p \wedge u)\alpha^{-1} = \lambda\alpha^{-1} + \Gamma. \qquad \square$$

Proof of Theorem IV.3.11. The map $\varphi_C : \mathcal{T}_C \to \operatorname{Pic}(C)$ is well defined as soon as C is smooth. If the n vectors have independent images, then they must be independent themselves, and the curve C corresponds to a regular value.

To establish the converse, we look at each component of the discriminant:

- The hyperplanes $z = \alpha_i$ or $F_i = 0$ are critical since the images of the (p, u), with $p_i = 0$ and $u_i = 0$ annihilate dF_i.
- Consider the points of the discriminant of Q, that is to say, the level sets $F_1, \dots, F_n$ such that Q has a double root z. This level set contains common tangents to two quadrics that coincide. For reasons of degree, these lines are contained in Q_z. Such a line is clearly a fixed point of the flow of Ψ_z, and thus the level set is critical. $\square$

Here is another application of Lemma IV.3.14, which we be useful in Section IV.3.d for determining the action variables.

COROLLARY IV.3.16. *The map $\varphi_C : \mathcal{T}_C \to C_1 \times \cdots \times C_n$ is a covering.*

PROOF. The tangent map is everywhere injective and the two manifolds are compact. $\square$

REMARK IV.3.17. Since $C_1 \times \cdots \times C_n$ is a torus, this corollary guarantees that $\mathcal{T}_C$ is union of tori... which the Arnold–Liouville theorem already told us long ago. This does not help in counting the tori (the covering has degree 2^{n+1} and $\mathcal{T}_C$ could have any number of components between 1 and 2^{n+1}). Nevertheless, we can still use algebraic geometry to determine the number of Liouville tori (see [**6**]). This number is a power of 2, with the exponent being the number of types of real quadrics appearing twice in the list $Q_{z_1}, \ldots, Q_{z_n}$. For example, there are two Liouville tori when $n = 2$ (see Figure 3), except for values lying in the small region with a boundary on the discriminant, in which case there are four Liouville tori (see Exercise IV.15).

Exercise IV.15.

IV.3.d. Action Variables. We now turn to the question of action variables. We no longer restrict ourselves to a level set $\mathcal{T}_C$, or to a curve C, or even to a Jacobian, but instead we will return to the global setting of the symplectic manifold $\mathcal{D}_n$ or at least to a connected neighborhood of a Liouville torus. Indeed, as promised in Example II.2.8, we will construct action variables that are "as global as possible", that is, global over each component of the set of regular values of the first integrals.

We thus fix one of these components Δ, that is to say, we fix the position of the z_i with respect to the α_j or the type of the real parts of the quadrics $Q_{z_1}, \ldots, Q_{z_n}$, or the real part of the spectral curves. We consider the family $\mathcal{C}$ of spectral curves that is a convenient completion of

$$\left\{(F, \lambda, z) \in \Delta \times \mathbb{R} \times \mathbb{R} \mid \lambda^2 = \sum \frac{F_j}{\alpha_j - z}\right\}$$

over Δ: the points at infinity are the same on all the curves, so there is no problem in completing all the curves as a family. We will also use the product bundle $\mathcal{C}_n$ viewed as the family[7] of the $C^n_{\mathbb{R}}$ over Δ.

As one might expect (see Section II.2.b), we construct the action variables by integrating a primitive of the symplectic form on loops forming a basis for the fundamental group of the Liouville tori. We already have what we need to construct such a basis: up to a covering, a Liouville torus is a product $C_1 \times \cdots \times C_n$. We also have a (global) primitive of the symplectic form on the symplectic manifold $\mathcal{D}_n$ containing the Liouville tori. This, in fact, is isomorphic to TS^n. The form

$$\alpha = \sum_{i=1}^{n+1} u_i dp_i$$

is a primitive of the symplectic form.

[7]This is a baby version of the relative Jacobian, which will suffice for our purposes because of Proposition IV.3.9.

There is still a bit of work to do since the basis of the fundamental group is from "the Jacobian way", while the 1-form comes from "the Liouville way".

Thus we begin by constructing a form on $\mathcal{C}_n$. We make use of the form λdz on each of the curves C, which then defines by the same formula a form on $\mathcal{C}$ with which we finally construct a form

$$\sigma = \sum_{i=1}^{n} \lambda_i dz_i$$

on $\mathcal{C}_n$.

Remark IV.3.18. The form λdz is a meromorphic form on the complex curve C. It has a double pole at the point $z = \infty$ of C, which lies on none of the other components $C_1, \ldots, C_n$ of $C_{\mathbb{R}}$ that interest us (this is an easy exercise; see Exercise B.12).

We send $F^{-1}(\Delta)$ into $\mathcal{C}_n$ by the eigenvector map

$$\varphi : (p, u) \longmapsto (F(p, u), R_1, \ldots, R_n),$$

and we have the following lemma (whose calculation-intensive proof is given in Section IV.3.e).

Lemma IV.3.19. *The form σ satisfies*

$$\varphi^\star \sigma = 2(-1)^{n-1}\alpha.$$

Finally, we denote by $C_1(F), \ldots, C_n(F)$ the n components $C_1, \ldots, C_n$ of the real part of the curve corresponding to F.

Theorem IV.3.20. *Let U be a connected component of the inverse image of a connected component of the set of regular values. There are nonzero integers $m_1, \ldots, m_n$ such that*

$$\begin{aligned} U &\longrightarrow \mathbb{R}^n \\ (p, u) &\longmapsto \Big(\frac{m_1}{2} \int_{C_1(F)} \sigma, \ldots, \frac{m_n}{2} \int_{C_n(F)} \sigma \Big) \end{aligned}$$

are action variables.

Proof. The m_i are the smallest integers such that there is a basis $\gamma_1, \ldots, \gamma_n$ of the fundamental group of a Liouville tori contained in U with the image of γ_i under φ being m_i times the loop C_i. Their existence is assured by Corollary IV.3.16. □

IV.3.e. The Proofs of Three Lemmas.

Proof of Lemma IV.3.10. We fix points $P_1, \ldots, P_n$ of C such that $P_1 + \cdots + P_n$ lies in V. The lemma asserts that the only meromorphic functions on the curve C having, at worst, simple poles at $P_1, \ldots, P_n$ are the constant functions. This is a direct application of the Riemann–Roch theorem, which tells us that the dimension of the space of these functions is

$$\deg(P_1 + \cdots + P_n) - n + 1 + i = 1 + i$$

where i is the dimension of the space of holomorphic functions on C having a zero at each P_k. Since the curve C is hyperelliptic with genus n, we know all the holomorphic forms on C explicitly

$$\frac{g(z)}{y}dz \text{ for } g \text{ a polynomial of degree } \leqslant n-1.$$

If $\ell = \deg g$, the divisor of this form is

$$\left(\frac{g(z)}{y}dz\right) = Z_1 + \tau Z_1 + \cdots + Z_\ell + \tau Z_\ell + (n-\ell-1)\infty,$$

where we have denoted by ∞ the divisor of degree 2 of the poles of z and Z_i is a point on C above a root of the polynomial g.

Now, in order to have

$$Z_1 + \tau Z_1 + \cdots + Z_\ell + \tau Z_\ell + (n-\ell-1)\infty \geqslant P_1 + \cdots + P_n,$$

it is necessary that $P_i = \tau(P_j)$ or that $z(P_i) = \infty$ at least once. Thus if $P_1 + \cdots + P_n$ lies in V, we must have $i = 0$. $\square$

Proof of Lemma IV.3.14. We show that the Hamiltonian vector fields corresponding to the functions $\Psi_{z_1}, \ldots, \Psi_{z_n}$ have independent images. By definition, each Ψ_{z_i} is a linear combination of $F_1, \ldots, F_n$ (replace F_{n+1} by $1 - \sum F_i$), so the images of the Hamiltonian vector fields of $F_1, \ldots, F_n$ also generate a subspace of dimension n.

Consider the image of the Hamiltonian vector field of Ψ_0 (that of Ψ_z is obtained by replacing everywhere α_i by $\alpha_i - z$). By Lemma IV.3.15 and Theorem, we see that the corresponding eigenvector bundle can be constructed from a cocycle $\exp(t\beta)c$ with

$$\beta(\lambda, \mu) = \lambda^3 \mu^{-1} = \lambda^3 \lambda^{-2} z^{-1}$$

seen as a cocycle of the covering of C by the two open sets U_0 (where $\lambda \neq \infty$) and U_∞ (where $\lambda \neq 0$) representing an element of $H^1(C; \mathcal{O}_C)$. Thus this cocycle represents the image of the Hamiltonian vector field of Ψ_0 under the tangent map associated to the eigenvector map.

By replacing everywhere α_i by $\alpha_i - z_j$, we see that the images of the Hamiltonian vector fields corresponding to $\Psi_{z_1}, \ldots, \Psi_{z_n}$ are the classes of the cocycles

$$\lambda (z - z_1)^{-1}, \ldots, \lambda (z - z_n)^{-1}$$

in $H^1(C; \mathcal{O}_C)$. To complete the proof of the lemma, we just need to verify that these are independent. This is a residue calculation. We will choose a basis $\eta_1, \ldots, \eta_n$ of the vector space $H^0(C; \Omega^1_C)$ of holomorphic forms such that the matrix with entries

$$\operatorname{res}_{\lambda=0}\left(\frac{\lambda \eta_i}{z - z_j}\right)$$

is (diagonal and) invertible. This will give us an injective linear map

$$H^1(C; \mathcal{O}_C) \longrightarrow H^0(C; \Omega^1_C)^\star$$

sending the classes of the cocycles $\lambda(z - z_j)^{-1}$ to a basis of $H^0(C; \Omega^1_C)^\star$, thus proving the independence of these cocycles.

By writing the equation of C in the form

$$\left[\lambda \prod (\alpha_s - z)\right]^2 = Q(z) \prod (\alpha_s - z)\,,$$

(that is, $y^2 = P(z)$ for a polynomial P of degree $2n+1$), we see that the holomorphic forms on C are all of the forms

$$\frac{f(z)dz}{\lambda \prod(\alpha_j - z)},$$

where f is a polynomial of degree at most $n-1$. We let P_i be the polynomial

$$P_i(z) = \prod_{j \neq i}(z - z_j) = \frac{Q(z)}{z - z_i}.$$

Since the z_i are distinct, the P_i are a basis of the space of polynomials of degree at most $n-1$, and the forms

$$\eta_i = \frac{P_i(z)dz}{\lambda \prod(\alpha_s - z)} \qquad 1 \leqslant i \leqslant n$$

are a basis of $H^0(C;\Omega^1_C)$. We now consider the meromorphic forms

$$\gamma_{i,j} = \frac{\lambda \eta_i}{z - z_j} \qquad 1 \leqslant i,j \leqslant n.$$

The form $\gamma_{i,j}$ has a pole at a point S_k only if $i = j = k$, and it has no pole at ∞. Then we have

$$\operatorname{res}_{\lambda=0}(\gamma_{i,j}) = \delta_{i,j} \operatorname{res}_{S_i} \frac{P_i(z)dz}{(z - z_i)\prod(\alpha_s - z)},$$

and we see that the matrix is diagonal. In a neighborhood of S_i, λ is a coordinate on C and $z = z_i + A_i\lambda^2 + o(\lambda^2)$, from which we deduce the value of the residue

$$\operatorname{res}_{S_i} \frac{P_i(z)dz}{(z - z_i)\prod(\alpha_s - z)} = 2\frac{\prod_{j\neq i}(z_i - z_j)}{\prod_s(\alpha_s - z_i)},$$

hence, the diagonal terms of the matrix are nonzero, which is what we wanted to prove. □

Proof of Lemma IV.3.19. It seems to me that this result ought to be tautological, this is why I am ashamed, for lack of a better method, to give a proof based on pure calculation. Let us evaluate

$$\varphi^\star \sigma = -\sum_{i=1}^{n} \varphi_{\zeta_i}(p,u)d\zeta_i.$$

We write the relation

$$\mathcal{Q}(z,p) = q_z(p) = \sum_{j=1}^{n+1} \frac{p_j^2}{\alpha_j - z} = \frac{\mathcal{R}(z)}{\prod(\alpha_j - z)},$$

differentiate it and then evaluate it at one of the roots ζ (one of the ζ_i) of $\mathcal{R}$ to obtain

$$\frac{\partial \mathcal{Q}}{\partial z}(\zeta,p)d\zeta + 2\varphi_\zeta(p,dp) = 0,$$

so that

$$d\zeta = -\frac{2\varphi_\zeta(p, dp)}{\dfrac{\partial \mathcal{Q}}{\partial z}(\zeta, p)}$$

and

$$-\varphi_\zeta(p, u)d\zeta = 2\frac{\varphi_\zeta(p, u)\varphi_\zeta(p, dp)}{\dfrac{\partial \mathcal{Q}}{\partial z}(\zeta, p)}.$$

On the other hand, $\mathcal{R}(z) = \prod(\zeta_k - z)$ and thus

$$\frac{\partial \mathcal{Q}}{\partial z}(\zeta_i, p) = \frac{\mathcal{R}'(\zeta_i)}{\prod(\alpha_j - \zeta_i)} = -\frac{\prod_{k\neq i}(\zeta_k - \zeta_i)}{\prod_j(\alpha_j - \zeta_i)}.$$

We get

$$\varphi^\star\sigma = 2\sum_{i=1}^{n}\frac{\varphi_{\zeta_i}(p, u)\varphi_{\zeta_i}(p, dp)}{\dfrac{\partial \mathcal{Q}}{\partial z}(\zeta_i, p)} = 2\sum_{j=1}^{n+1}A_j p_j dp_j$$

with

$$A_j = \sum_{i=1}^{n}\frac{\varphi_{\zeta_i}(p, u)}{(\alpha_j - \zeta_i)\dfrac{\partial \mathcal{Q}}{\partial z}(\zeta_i, p)}.$$

Let us evaluate A_j. Since all the j play exactly the same role and there are already so many indices to keep track of, I will limit myself to calculating A_1

$$\begin{aligned}
A_1 &= -\sum_{i=1}^{n}\frac{\varphi_{\zeta_i}(p, u)}{\prod_{k\neq i}(\zeta_k - \zeta_i)}\frac{\prod_{\ell\neq 1}(\alpha_\ell - \zeta_i)}{\alpha_1 - \zeta_i}\\
&= -\sum_{i=1}^{n}\frac{\prod_{\ell\neq 1}(\alpha_\ell - \zeta_i)}{\prod_{k\neq i}(\zeta_k - \zeta_i)}\sum_{j=1}^{n+1}\frac{p_j u_j}{\alpha_j - \zeta_i}\\
&= -p_1u_1\sum_{i=1}^{n}\frac{\prod_{\ell\neq 1}(\alpha_\ell - \zeta_i)}{(\alpha_1 - \zeta_i)\prod_{k\neq i}(\zeta_k - \zeta_i)} - \sum_{j=2}^{n+1}p_j u_j\sum_{i=1}^{n}\frac{\prod_{\ell\neq 1,j}(\alpha_\ell - \zeta_i)}{\prod_{k\neq i}(\zeta_k - \zeta_i)}\\
&= -Bp_1u_1 - \sum_{j=2}^{n+1}C_j p_j u_j.
\end{aligned}$$

...we first calculate the C_j. To do this, expand the rational fraction

$$D(Z) = \frac{\prod_{\ell\neq 1,j}(\alpha_\ell - Z)}{\prod_{k\neq 1}(\zeta_k - Z)}$$

into partial fractions (the term $i = 1$ in the sum defining C_j is $D(\zeta_1)$)

$$D(Z) = 1 - \sum_{i=2}^{n}\frac{\prod_{\ell\neq 1,j}(\alpha_\ell - \zeta_i)}{(Z - \zeta_i)\prod_{k\neq 1,i}(\zeta_k - \zeta_i)}.$$

By replacing Z by ζ_1 in this equality, we see that $C_j = 1$. Next, we calculate B

$$\begin{aligned}
B &= \sum_{i=1}^{n} \frac{\prod_{\ell\neq 1}(\alpha_\ell - \zeta_i)}{(\alpha_1 - \zeta_i)\prod_{k\neq i}(\zeta_k - \zeta_i)} \\
&= \frac{1}{\prod_{i=1}^{n}(\alpha_1 - \zeta_i)} \sum_{i=1}^{n} \frac{\prod_{\ell\neq i}(\alpha_1 - \zeta_\ell)\prod_{\ell\neq 1}(\alpha_\ell - \zeta_i)}{\prod_{k\neq i}(\zeta_k - \zeta_i)} \\
&= \frac{1}{\mathcal{R}(\alpha_1)} E(\alpha_1),
\end{aligned}$$

where $E(Z)$, a polynomial of degree at most $n-1$, equals $(-1)^{n-1}\prod_{\ell\neq 1}(\alpha_\ell - \zeta_i)$ at ζ_i, so that

$$E(Z) = (-1)^{n-1}\prod_{s\neq 1}(\alpha_s - Z) - \prod_{i=1}^{n}(\alpha_1 - \zeta_i).$$

We deduce from this the equality

$$B = \frac{(-1)^{n-1}\prod_{s\neq 1}(\alpha_s - \alpha_1)}{\mathcal{R}(\alpha_1)} - 1.$$

Since we also have that $\mathcal{R}(\alpha_1) = p_1^2\prod_{k\neq 1}(\alpha_k - \alpha_1)$, B is simply equal to $-1 + (-1)^{n-1}/p_1^2$, and, eventually,

$$\begin{aligned}
A_1 &= -p_1 u_1\Big(\frac{(-1)^{n-1}}{p_1^2} - 1\Big) - \sum_{j=2}^{n+1} p_j u_j \\
&= (-1)^{n-1}\frac{u_1}{p_1},
\end{aligned}$$

using the fact that the vectors p and u are orthogonal. Similarly, $A_j = (-1)^{n-1}u_j/p_j$, and, finally,

$$\varphi^\star\sigma = 2\sum A_j p_j dp_j = 2(-1)^{n-1}\sum_{j=1}^{n+1} u_j dp_j,$$

which is what we wanted to prove. □

IV.4. Appendix: How to Construct Lax Equations (an Overview)

A Lax equation is a differential equation in a space of matrices and we have seen that it has many first integrals. To understand it as a completely integrable Hamiltonian system, we also need a minimum amount of symplectic data. In fact, the space of matrices has exactly what we need to make Hamiltonian systems. This is because the (duals of the) Lie algebras are naturally equipped with Poisson brackets. I begin here with a review of this notion.

IV.4.a. Poisson Structure on the Dual of a Lie Algebra. I give here a quick description of the Kirillov–Kostant–Souriau structure (see [**48, 52, 74**]) on the dual of a Lie algebra. For more details, see Kirillov's book [**48**] and the symplectic geometry texts cited earlier, especially [**56**] and [**15**].

The Dual of a Lie Algebra. As we know, a (real or complex) vector space $\mathfrak{g}$ is a Lie algebra if it is equipped with a skew-symmetric bilinear bracket

$$[\ ,\ ] : \mathfrak{g} \times \mathfrak{g} \longrightarrow \mathfrak{g}$$

satisfying the Jacobi identity

$$[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0 \qquad \forall\, X,Y,Z \in \mathfrak{g}.$$

Consider now the dual vector space $\mathfrak{g}^\star$. Note that the Lie bracket of $\mathfrak{g}$ does not define a Lie algebra structure on $\mathfrak{g}^\star$ in any natural way. Instead, it defines a *Poisson bracket* (which was discovered and studied by Kirillov, Kostant, and Souriau). If $f \in \mathcal{C}^\infty(\mathfrak{g}^\star)$ is a function and $\xi \in \mathfrak{g}^\star$ is a point[8], then $df(\xi)$ is a linear form on $T_\xi \mathfrak{g}^\star = \mathfrak{g}^\star$, which we can consider as an element of $\mathfrak{g}$ via biduality. For $f, g \in \mathcal{C}^\infty(\mathfrak{g}^\star)$, define

$$\{f,g\}(\xi) = \langle \xi, [df(\xi), dg(\xi)]\rangle,$$

where $\langle\ ,\ \rangle$ is the duality. This gives us a Poisson bracket: $\{\ ,\ \}$ is obviously skew-symmetric, satisfies the Jacobi identity, and also satisfies the Leibniz identity because it only involves differentials of functions.

Although it has a Poisson bracket, there is no reason why $\mathfrak{g}^\star$ should be a symplectic manifold. Indeed, the example given in Exercise IV.3 has odd dimension. Nonetheless, the "Poisson manifold" $\mathfrak{g}^\star$ (as all Poisson manifolds) is foliated by symplectic manifolds (see [**81**] for the geometry of Poisson manifolds). Here the leaves of the foliation are the coadjoint orbits. See Exercise IV.19.

Just as on a symplectic manifold, we can define Hamiltonian vector fields for the functions on $\mathfrak{g}^\star$. The vector field X_f is the vector field defined by the way it derives functions

$$X_f \cdot g = \{f,g\}.$$

Adjoint and Coadjoint Representations. Finally, recall that ad (for the adjoint map) denotes the map

$$\begin{aligned} \mathfrak{g} &\longrightarrow \operatorname{End} \mathfrak{g} \\ X &\longmapsto (Y \mapsto [Y,X]) \end{aligned}$$

and $\mathrm{ad}^\star$ (the coadjoint map) denotes its transpose

$$\mathfrak{g} \longrightarrow \operatorname{End} \mathfrak{g}^\star$$

which is defined by

$$\langle \mathrm{ad}^\star_X \xi, Y\rangle = \langle \xi, \mathrm{ad}_{-X} Y\rangle = \langle \xi, [X,Y]\rangle.$$

Exercises IV.13, IV.16, IV.17, IV.18, IV.19, IV.20, IV.21.

IV.4.b. R-Matrices and Lax Equations.

---

[8] I use systematically capital Latin letters for elements in $\mathfrak{g}$ and Greek letters for elements in $\mathfrak{g}^\star$.

Classical R-Matrices. I follow closely the article [**73**] of Semenov-Tian-Shanski. In this article, we learn that a classical R-matrix is a linear map $R:\mathfrak{g}\to\mathfrak{g}$ (where $\mathfrak{g}$ is a Lie algebra) satisfying a "Yang–Baxter" equation.

Any linear map R can be used to define a new skew-symmetric bracket on $\mathfrak{g}$

$$[X,Y]_R = [RX,Y]+[X,RY].$$

The Yang–Baxter equation, which we do not need to write down here, simply expresses the condition that this new bracket satisfies the Jacobi identity, in other words, that $(\mathfrak{g},[\ ,\ ]_R)$ is a Lie algebra. In particular, $\mathfrak{g}^\star$ is then equipped with a second Poisson structure, naturally written as $\{\ ,\ \}_R$.

Also as we learn in reading [**73**], the first use of R-matrices is to construct integrable systems.

Proposition IV.4.1.

(a) *If φ is an invariant function on $\mathfrak{g}^\star$, then*

$$X^R_\varphi(\xi) = \operatorname{ad}^\star_{Rd\varphi(\xi)}\cdot\xi.$$

(b) *If φ and ψ are invariant, then*

$$\{\varphi,\psi\} = \{\varphi,\psi\}_R = 0.$$

In this statement, X^R_φ is the Hamiltonian vector field associated to φ via the Poisson bracket $\{\ ,\ \}_R$. Thus Assertion (a) defines a Hamiltonian system, while Assertion (b) provides a list of first integrals.

Proof. The invariant functions on $\mathfrak{g}^\star$ discussed here are the functions invariant under the coadjoint action (Exercise IV.16) or, equivalently, the functions satisfying

$$\langle\xi,[d\varphi(\xi),Y]\rangle = 0 \quad \text{for all } \xi\in\mathfrak{g}^\star, Y\in\mathfrak{g}.$$

But then we have

$$\begin{aligned}\{\varphi,\psi\}_R(\xi) &= \langle\xi,[Rd\varphi(\xi),d\psi(\xi)]\rangle + \langle\xi,[d\varphi(\xi),Rd\psi(\xi)]\rangle\\ &= \langle\xi,[Rd\varphi(\xi),d\psi(\xi)]\rangle\end{aligned}$$

for any function ψ, which proves the first assertion. Moreover, if ψ is also invariant, then the remaining term is zero, giving us the second assertion. □

If $\mathfrak{g}$ is equipped with an invariant non-degenerate symmetric bilinear form, we can use this bilinear form to identify $\mathfrak{g}$ with $\mathfrak{g}^\star$ and ad with $\operatorname{ad}^\star$. Thus the Hamiltonian system corresponding to the invariant function φ can be written as

$$\frac{d}{dt}A = [A, R\nabla_A\varphi],$$

where ∇ denotes the gradient. This is indeed a Lax equation.

Proposition IV.4.1 contains a list of first integrals in involution. To get an integrable system, we still need to be in a symplectic manifold of the right dimension, namely, of sufficiently small dimension (since the number of independent functions is limited by the rank of the Lie algebra $\mathfrak{g}$).

Infinite-Dimensional R-Matrices and Algebraic Curves. Although we have utilized duality and, worse, bi-duality in a significant way, an important class of maps involving R-matrices is in the framework of infinite-dimensional Lie algebras of the form $\mathfrak{g}[\lambda,\lambda^{-1}]$, where $\mathfrak{g}$ is a (honest) Lie algebra of matrices, λ is an independent variable, and R is the endomorphism associated to the decomposition

$$\mathfrak{g}[\lambda,\lambda^{-1}] = \mathfrak{g}[\lambda] + \lambda^{-1}\mathfrak{g}[\lambda^{-1}]$$

via $R(P) = P_+ - P_-$ (where P_+ and P_- are the polynomial and Laurent components of P, respectively).

In this framework, we obtain Lax equations "with spectral parameter": the coefficients of the matrices are polynomials in λ and λ^{-1}.

The algebras $\mathfrak{g}[\lambda,\lambda^{-1}]$ (and their various completions) are called "loop algebras". If $\mathfrak{g}$ is equipped with an involution σ, we can also construct a subalgebra of $\mathfrak{g}[\lambda,\lambda^{-1}]$, the "twisted loop algebra"

$$\mathcal{L}(\mathfrak{g},\sigma) = \left\{A \in \mathfrak{g}[\lambda,\lambda^{-1}] \mid A(-\lambda) = \sigma(A(\lambda))\right\}.$$

The algebra $\mathcal{L}(\mathfrak{g},\sigma)$ is also equipped with a natural R-matrix corresponding, in the same way as above, to the decomposition into negative and positive powers of λ.

Remark IV.4.2. It is convenient to use *Laurent* polynomials in λ to write the "R-matrix" in this case. To obtain true polynomials in λ, we can just multiply the matrix A_λ by a suitable power of λ (as we did in Section IV.1.b).

The invariant functions on $\mathfrak{g}$ are in involution and provide us with numerous first integrals, since they are polynomials in λ whose coefficients are all conserved quantities. And as we noted in Section IV.1.b, they give us something even better: if a matrix $A_\lambda(t)$ is a solution to a differential equation of the form

$$\frac{d}{dt}A_\lambda(t) = [A_\lambda(t), B_\lambda(t)],$$

then its eigenvalues do not depend on the time t and its characteristic polynomial defines an algebraic curve (the spectral curve)

$$\left\{(\lambda,\mu) \mid \det(A_\lambda - \mu\,\mathrm{Id}) = 0\right\}.$$

Remark IV.4.3. The matrices B_λ obtained by this method are precisely the type studied in Section IV.2.e, and so the results of that section apply: the solutions of a Hamiltonian system constructed in this way are linearized on Jacobians.

Exercise IV.22.

IV.4.c. Construction of Examples. As we have seen, the R-matrix method can be used to construct many integrable systems written as Lax equations. But conversely, it is not always obvious to put a given integrable system into the form of a Lax equation. It is very easy for the simple examples of the spherical pendulum and the symmetric top (the Lax equations given in Section IV.1.b were not hard to invent). It is less obvious, in the case of geodesics on quadrics (which is due to Moser [**65**]). The case of the Kowalevski top, for example, resisted for a long time. Here I show a general construction yielding the Lax equations for all these systems.

One of its most beautiful successes (and certainly its "raison d'être") is the case of Kowalevski (see [**18**]). I will just summarize here the *ad hoc part* of [**71**] where the readers can find more details and numerous examples. The argument will involve what Reyman and Semenov call "multidimensional tops".

I will go through the construction fairly quickly, and the readers can skip this section in the first reading: it is indeed very easy to check that the Lax equations obtained are in fact equivalent to the systems they are supposed to represent.

We consider a group G of matrices and a compact subgroup K, the subgroup of the fixed points of an involution σ. Our examples will simply be

- $G = \mathrm{SL}(n;\mathbb{R})$, $K = \mathrm{SO}(n)$
- $G = \mathrm{SO}(p,q)$, $K = \mathrm{SO}(p) \times \mathrm{SO}(q)$,

and I will not do any more general theory here. The derivative of σ at 1 is also a (linear) involution of the Lie algebra $\mathfrak{g}$, which we will again call σ. We decompose this Lie algebra as

$$\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p},$$

where $\mathfrak{k}$ is the Lie algebra of K (and the vector space of fixed points of σ) and $\mathfrak{p}$ is the eigenspace of σ for the eigenvalue -1.

Examples IV.4.4.

- If $G = \mathrm{GL}(n;\mathbb{R})$ and $\sigma(A) = ({}^tA)^{-1}$, we have that $K = \mathrm{SO}(n)$ so that $\mathfrak{g}$ is the Lie algebra of matrices, $\mathfrak{k} = \mathfrak{so}(n)$ is the Lie algebra of skew-symmetric matrices, and $\mathfrak{p}$ is the vector space of symmetric matrices.
- If $G = \mathrm{SO}(p,q)$, still with $\sigma(A) = ({}^tA)^{-1}$, we have that $K = \mathrm{SO}(p) \times \mathrm{SO}(q)$ and $\mathfrak{g} = \mathfrak{so}(p,q)$, the space of $(p+q)\times(p+q)$ matrices such that ${}^tX = -JXJ$ (where J is the matrix of the standard quadratic form with signature (p,q)). The Lie algebra $\mathfrak{k}$ is $\mathfrak{so}(p)\times\mathfrak{so}(q)$ and $\mathfrak{p}$ is the vector space of matrices $\begin{pmatrix} 0 & s \\ {}^ts & 0 \end{pmatrix}$.

In both of these cases we identify $\mathfrak{g}$ with $\mathfrak{g}^\star$ via the invariant nondegenerate bilinear form

$$(X,Y) \longmapsto -\operatorname{Tr}(XY),$$

This also identifies $\mathfrak{k}$ with $\mathfrak{k}^\star$ and $\mathfrak{p}$ with $\mathfrak{p}^\star$.

We now fix a matrix $a \in \mathfrak{p}^\star = \mathfrak{p}$, and consider the set of matrices of the form

$$a\lambda + \ell + s\lambda^{-1} \quad \text{with } \ell \in \mathfrak{k}^\star \text{ and } s \in \mathfrak{p}^\star.$$

This is an (affine) subspace of the dual $\mathcal{L}(\mathfrak{g},\sigma)^\star$ of the twisted loop algebra and is equipped with the Poisson structure of the dual of $\mathfrak{k}\oplus\mathfrak{p}$.

The group K acts on $\mathfrak{k}^\star$ and $\mathfrak{p}^\star$ (Exercise IV.18). Let V be a manifold equipped with an action of K and an equivariant map $j : V \to \mathfrak{p}^\star$. We extend the action of K on V to a Hamiltonian action on $T^\star V$ with moment map

$$\mu : T^\star V \longrightarrow \mathfrak{k}^\star$$

(Exercise IV.20). We deduce from this a map

$$\begin{array}{ccccc} T^\star V & \longrightarrow & \mathfrak{k}^\star \oplus \mathfrak{p}^\star & \longrightarrow & \mathcal{L}(\mathfrak{g},\sigma)^\star \\ (v,\varphi) & \longmapsto & (\mu(v,\varphi), j(v)) & \longmapsto & a\lambda + \mu(v,\varphi) + j(v)\lambda^{-1}, \end{array}$$

which preserves Poisson brackets (Exercise IV.21). Then from Proposition IV.4.1 we deduce the the following:

Proposition IV.4.5. *The invariant functions of A_λ are in involution with respect to the Poisson bracket of $T^\star K$ (respectively of $T^\star V$).*

We still need to choose a Hamiltonian. We fix an invariant function φ on $\mathfrak{g}$. To the function φ we associate a (completely integrable) Hamiltonian on the cotangent bundle $T^\star K$ via

$$H(k,\kappa) = -\frac{1}{2}\operatorname{Tr}(\mathcal{J}(\kappa)\kappa) - \operatorname{Tr}(bk^{-1}fk),$$

where

- the symmetric endomorphism $\mathcal{J}$ is $\varphi''(a)$,
- the "kinetic component" is quadratic and given by $\mathcal{J}(\kappa)$,
- the b that appears in the "potential component" is $b = d\varphi(a)$.

We consider the Lax equations $\dot{A}_\lambda = [A_\lambda, B_\lambda]$ with

$$A_\lambda = a\lambda + \kappa + (k^{-1}fk)\lambda^{-1} \text{ and } B_\lambda = b\lambda + \mathcal{J}(\kappa).$$

This construction gives Lax equations for the following systems (with some generalizations to higher dimensions).

Example IV.4.6 (The Spherical Pendulum). Here $G = \mathrm{SO}(3,1)$ and $K = \mathrm{SO}(3)$. Then $\mathfrak{p} = \mathbb{R}^3$ with the usual action of SO(3). The manifold V is the sphere S^2 and $j : S^2 \to \mathbb{R}^3$ is the standard embedding. The matrix A_λ is

$$-\left(\begin{array}{c|c} 0 & \Gamma \\ \hline {}^t\Gamma & 0 \end{array}\right)\lambda + \left(\begin{array}{c|c} p\wedge q & 0 \\ \hline 0 & 0 \end{array}\right) + \left(\begin{array}{c|c} 0 & q \\ \hline {}^tq & 0 \end{array}\right)\lambda^{-1}.$$

The simplest Hamiltonian is

$$H = \frac{1}{2}\sum p_i^2 + \sum a_i q_i$$

(in the notation of Sections I.2.f and IV.1.a, $a = \Gamma = -e_3$). This gives rise to the Lax equation of Exercise IV.5. Note, however, that this is *not* the Lax pair we encountered in Example IV.1.3. As one might expect, Lax pairs are not unique.

Example IV.4.7 (The Symmetric Top). The same method also gives us a Lax pair for the symmetric top. Here again $G = \mathrm{SO}(3,1)$, $K = \mathrm{SO}(3)$, but, this time, the Hamiltonian is

$$H = \frac{1}{2}M\cdot\mathcal{J}(M) + \Gamma\cdot e_3.$$

The Lax pair is (A_λ, B_λ), where

$$A_\lambda = \begin{pmatrix} 0 & \Gamma \\ {}^t\Gamma & 0 \end{pmatrix}\lambda^{-1} + \begin{pmatrix} -M & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & L \\ {}^tL & 0 \end{pmatrix}\lambda$$

and B_λ is the polynomial component $\left(A_\lambda^2\right)_+$ of A_λ^2.

The corresponding Lax equation is equivalent to the system describing the symmetric top (Exercise IV.10). Once again, this is *not* the Lax pair we encountered in Example IV.1.4.

Example IV.4.8 (The Kowalevski Top). This is the most beautiful success of this construction. We get it in the case where $G = \mathrm{SO}(3,2)$ and $K = \mathrm{SO}(3) \times \mathrm{SO}(2)$ (see [**18, 71**], as well as the applications in [**16**] and [**8**, Chapter III]).

Example IV.4.9 (Geodesics on Quadrics). Here $G = \mathrm{SL}(n+1;\mathbb{R})$ and $K = \mathrm{SO}(n+1)$, and the decomposition $\mathfrak{g} = \mathfrak{so}(n) \oplus \mathfrak{p}$ is the decomposition of a matrix into the sum of a skew-symmetric matrix and a symmetric matrix. We choose a to be the diagonal matrix $\alpha = (\alpha_1, \dots, \alpha_{n+1})$ of the quadratic form $\langle Ax, x\rangle$ in an orthonormal basis. The manifold V is the sphere S^n, and the equivariant map j, which sends it into the space of symmetric matrices, is $p \mapsto -p \otimes p$.

The moment map for the action of $\mathrm{SO}(n+1)$ on $T^\star S^n$ is given by $\mu(p,u) = p \wedge u$ (Exercise IV.20). We find that A_λ is the matrix

$$A_\lambda = \lambda\alpha + p \wedge u - \lambda^{-1} p \otimes p,$$

which is, within a factor of λ, the same as was used in Section IV.3.

Exercises

Exercise IV.1. Determine the solutions $q_3(t)$ of the differential equation describing the height q_3 of the spherical pendulum utilizing the Weierstrass $\wp$-function associated to the elliptic curve obtained in Section IV.1.a.

Exercise IV.2. For the symmetric top, let x_3 be the height of the end of the axis of rotation. Show that x_3 satisfies a differential equation of the form $\dot{x}_3^2 = P(x_3)$, where P is a cubic polynomial[9] with two roots in $(-1,1)$ and one in $(1,+\infty)$ (see [**2**]).

Deduce from this that the end of the top's axis oscillates between two parallel circles on the unit sphere (as experience would suggest).

Exercise IV.3. Show that the isomorphism of $\mathbb{R}^3$ onto the vector space of skew-symmetric matrices

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \longmapsto \begin{pmatrix} 0 & -z & y \\ z & 0 & -x \\ -y & x & 0 \end{pmatrix}$$

sends the vector product to the matrix bracket.

Exercise IV.4. Show that the Lax equation given in Example IV.1.4 is equivalent to the system

$$\dot{\Gamma} = [\Gamma, \Omega], \quad \dot{M} = [M, \Omega], \quad \dot{L} = [M - \Omega, L].$$

[9]This polynomial has already appeared at the end of the proof of Proposition II.3.1 and in Exercise II.15.

Deduce from this that it is also equivalent to the Hamiltonian system describing the motion of the solid in the case of a symmetric top (see Section I.3.g).

Exercise IV.5. Fix a vector $\Gamma \in \mathbb{R}^3$, and consider the 4×4 matrix

$$-\left(\begin{array}{c|c} 0 & \Gamma \\ \hline {}^t\Gamma & 0 \end{array}\right)\lambda^2 + \left(\begin{array}{c|c} p \times q & 0 \\ \hline 0 & 0 \end{array}\right)\lambda + \left(\begin{array}{c|c} 0 & q \\ \hline {}^tq & 0 \end{array}\right),$$

where p and q are vectors in $\mathbb{R}^3$ and where the vector product $p \times q$ is identified with the corresponding 3×3 skew-symmetric matrix (see Exercise IV.3 for this identification). To simplify notation, we write this matrix in the form

$$A_\lambda = -(\Gamma \otimes e_4)\lambda^2 + (p \wedge q)\lambda + (q \otimes e_4).$$

The differential equation

$$\frac{d}{dt}A_\lambda = [A_\lambda, -(\Gamma \otimes e_4)\lambda + p \wedge q]$$

is a Lax equation with spectral parameter. Show that it is equivalent to the differential system

$$\begin{cases} \dot{q} = q \times (p \times q) \\ \dot{\overbrace{p \times q}} = -q \times \Gamma. \end{cases}$$

Now, suppose that $\|q\|^2 = 1$ and $p \cdot q = 0$ (i.e., $(q,p) \in TS^2$). Show that the Lax equation is equivalent to

$$\begin{cases} \dot{q} = p \\ \dot{p} \times q = -q \times \Gamma, \end{cases}$$

that is to say, the Hamiltonian system for the spherical pendulum.

Exercise IV.6 (Isospectral Deformations). Let $U(t)$ be an invertible matrix solution of the differential equation $\dot{U}_\lambda = U_\lambda B_\lambda$. Show that A_λ satisfies

$$\frac{d}{dt}\left(U_\lambda A_\lambda U_\lambda^{-1}\right) = 0$$

if and only if it is a solution of $\dot{A}_\lambda = [A_\lambda, B_\lambda]$. Then show that the solutions of the Lax equation are the matrices $A_\lambda(t) = U_\lambda(t)^{-1}A_\lambda(t_0)U_\lambda(t)$.

Exercise IV.7 (A Pedestrian Proof). Show that if $A_\lambda(t)$ satisfies a Lax equation, then $\operatorname{Tr}(A_\lambda^m)$ is constant for all m and thus the coefficients of the powers of λ in the polynomial $\operatorname{Tr}(A_\lambda^m)$ are first integrals for the Lax equation.

Exercise IV.8 (An Exercise of Minor Interest). Calculate the characteristic polynomials of the matrices in Exercise IV.5 to show that

$$\det(A_\lambda - \mu\,\mathrm{Id}) = \mu^4 - \mu^2\|q\|^2 + \mu^2\lambda^2(\|p\|^2 + 2q_3) - \mu^2\lambda^4 - (p_1q_2 - p_2q_1)^2\lambda^6$$

and retrieve the first integrals H and K given in Sections I.2.f and I.3.e. The equation of the spectral curve is

$$\mu^4 - \mu^2 + 2\mu^2\lambda^2 H - \mu^2\lambda^4 - K^2\lambda^6 = 0.$$

EXERCISE IV.9. We return to the Lax equation of Example IV.1.4. Show that the characteristic polynomial of $\Gamma + M\lambda + L\lambda^2$ is

$$-\mu(\mu^2 + 1 + 2(\Gamma \cdot M)\lambda + (\|M\|^2 + 2\Gamma \cdot L)\lambda^2 + 2(M \cdot L)\lambda^3 + \lambda^4).$$

Retrieve the first integrals for the motion of the top given in Sections I.2.g and I.3.g.

EXERCISE IV.10. Consider the Lax equation

$$\frac{dA_\lambda}{dt} = [A_\lambda, B_\lambda],$$

where A_λ and B_λ are in $\mathfrak{so}(3,1)$ and are defined by

$$A_\lambda = \begin{pmatrix} 0 & \Gamma \\ {}^t\Gamma & 0 \end{pmatrix} \lambda^{-1} + \begin{pmatrix} -M & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & L \\ {}^tL & 0 \end{pmatrix} \lambda$$

and B_λ is the polynomial component $\left(A_\lambda^2\right)_+$ of A_λ^2 (see Example IV.4.7). Show that this is equivalent to the system describing the symmetric top.

EXERCISE IV.11. Let A be the skew-symmetric matrix

$$A = \begin{pmatrix} 0 & -z & y \\ z & 0 & -x \\ -y & x & 0 \end{pmatrix},$$

and let μ be an eigenvalue of A such that $\mu^2 + x^2 + y^2 + z^2 = 0$. Check that

$$V = \begin{pmatrix} \dfrac{-xz - \mu y}{x^2 + y^2} \\ \dfrac{-yz + \mu x}{x^2 + y^2} \\ 1 \end{pmatrix}$$

is an eigenvector of A for the eigenvalue μ. We now suppose that

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = q + \lambda(p \wedge q) - \lambda^2\Gamma,$$

in other words, we consider the example of the spherical pendulum (Section IV.1.3). Show[10] that the divisor of the poles in V has the form

$$R_+ + R_- + \infty$$

where ∞ denotes the divisor of degree 2 over $\lambda = \infty$ on C and R_+ and R_- are the points with coordinates

$$\lambda(R_\pm) = -\frac{q_1 \mp iq_2}{v_1 \mp iv_2} \text{ and } \mu(R_\pm) = \pm i\left(q_3 + \lambda(R_\pm)v_3 + \lambda(R_\pm)^2\right)$$

—where v_1, v_2, and v_3 are the coordinates of $p \times q$.

[10] This calculation is given explicitly (for the top) in [**8**].

EXERCISE IV.12 (Regular Values for the Spherical Pendulum (again)). Fix a point $(H,K)\in\mathbb{C}^2$ and consider the spectral curve C with equation

$$\mu^2+\lambda^4+2K\lambda^3+2H\lambda^2+1=0$$

as well as the level set $\mathcal{T}_C$ in the complexification[11] of TS^2.

Suppose that the curve C is smooth.

- Show that the eigenvector map φ_C is constant on the orbits of the vector field X_K (see Section I.3.e). Determine the points of $\mathcal{T}_C$ where X_K vanishes and the possible values of H and K.
- We said (in Example IV.2.6) that the image of the vector field X_H under the tangent map associated to φ_C is the cocycle class $\beta=\lambda^{-1}\mu$. Calculate $\operatorname{res}_{\lambda=0}(\beta d\lambda/\mu)$, and deduce that the image of X_H is never zero.

From this, deduce that if the complex curve C is smooth, then the level set $\mathcal{T}_C$ is regular.

Show that the curve C is singular exactly when H and K satisfy

$$\begin{cases} H=\dfrac{1}{2}u^2-\dfrac{3}{2u^2} \\[2ex] K=-u+\dfrac{1}{u^3} \end{cases}\qquad \text{for some } u\in\mathbb{C}^\star$$

(hint: write that the polynomial $\lambda^4+2K\lambda^3+2H\lambda^2+1$ has a double root u). Show that the level set $\mathcal{T}_C$ is critical when C is singular (use the horizontal motions, see Exercise II.13).

We return (at last!) to the real setting. Determine the image of the map $(H,K):TS^2\to\mathbb{R}^2$ and the set of its regular values (finding again Figure 7 of Chapter II)[12].

EXERCISE IV.13. Consider the Euclidean space $\mathbb{R}^3$. Show that the vector product defines a Lie algebra structure on $\mathbb{R}^3$. Using the Euclidean structure, identify it with its dual. Show that the coordinate functions x, y, and z satisfy

$$\{x,y\}(v)=\left\langle v,\left[\begin{pmatrix}a\\0\\0\end{pmatrix},\begin{pmatrix}0\\b\\0\end{pmatrix}\right]\right\rangle=abc$$

(with $v={}^t(a,b,c)$) and, similarly, that

$$\{x,y\}(v)=\{y,z\}(v)=\{z,x\}(v)=abc.$$

EXERCISE IV.14. Show that the Lax equation given in the statement of Proposition IV.3.1 is equivalent to the system

$$\begin{cases}\dot p\wedge u+p\wedge\dot u=[p\wedge u,\Gamma]-[p\otimes p,\alpha^{-1}]\\ p\otimes\dot p+\dot p\otimes p=[p\otimes p,\Gamma.]\end{cases}$$

---

[11]This is just $\left\{(q,p)\in\mathbb{C}^3\times\mathbb{C}^3 \mid \|q\|^2=1 \text{ and } q\cdot p=0\right\}$.

[12]See [**11**] for a study of the monodromy of this system utilizing methods of algebraic geometry.

Using the Hamiltonian system corresponding to the Hamiltonian Ψ_0 given in Exercise I.24, show that this system is equivalent to the Lax equation.

EXERCISE IV.15. Here we address the problem of geodesics on quadrics in $\mathbb{R}^3$ (in the notation of Section IV.3, this is the case $n = 2$). Determine which type of quadrics Q_{z_1}, Q_{z_2} correspond to the different connected components of the set of regular values of the map F (the shaded regions of Figure 3). Give the corresponding forms of the real part of the spectral curve.

Consider a critical value F on the discriminant of the polynomial Q (but not on one of the lines $F_i = 0$). The level set $\mathcal{T}_F$ corresponds in $\mathcal{D}_n$ to the limiting situation of the common tangents to two confocal quadrics which just coincided. What is the real type of these quadrics? What are the common tangents? Show that the critical level set $\mathcal{T}_F$ consists of four disjoint circles[13].

EXERCISE IV.16. The group G acts on itself by conjugation. Let g also denote the map

$$\begin{aligned} G &\xrightarrow{g} G \\ h &\longmapsto ghg^{-1} \end{aligned}$$

Show that the maps $T_1 g : \mathfrak{g} \to \mathfrak{g}$ define an action of G on $\mathfrak{g}$. This is called the *adjoint action*, and we write $T_1 g(X) = \operatorname{Ad}_g X$. What relation is there between Ad and ad? Determine the action and the orbits when $G = \mathrm{SO}(3)$.

The *coadjoint action* $\mathrm{Ad}^\star$ of G on $\mathfrak{g}^\star$ is defined by

$$\langle \operatorname{Ad}_g^\star \xi, X\rangle = \langle \xi, \operatorname{Ad}_{g^{-1}} X\rangle.$$

What relation is there between $\mathrm{Ad}^\star$ and $\mathrm{ad}^\star$? What is the fundamental field on $\mathfrak{g}^\star$ associated to X by the coadjoint action?

EXERCISE IV.17. Suppose that the Lie algebra $\mathfrak{g}$ is equipped with a nondegenerate bilinear form that is invariant under the adjoint action of the group G (see Exercise IV.16). Show that duality can be used to define a Poisson bracket on $\mathfrak{g}$. Show that this is the case when $G = \mathrm{SO}(n)$ (and, more generally, when G is compact) or when $G = \mathrm{SL}(n;\mathbb{C})$.

EXERCISE IV.18. Consider a Lie group G equipped with an involution σ. Let K be the subgroup of its fixed points. Decompose the Lie algebra $\mathfrak{g}$ of G

$$\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$$

where the components are the eigenspaces of σ for the eigenvalues 1 and -1, respectively. Check that the restriction to K of the coadjoint action of G fixes the subspaces $\mathfrak{k}^\star$ and $\mathfrak{p}^\star$.

Suppose that G is the group $\mathrm{SO}(3,1)$ of isometries of the diagonal form with signature $(3,1)$ and that $\sigma(A) = {}^tA^{-1}$ so that $K = \mathrm{SO}(3)$. Show that $\mathfrak{p}$ can be identified with $\mathbb{R}^3$ and that the operation of K on $\mathfrak{p}^\star$ can be identified with the usual action of $\mathrm{SO}(3)$ by isometries.

[13] This hints at the fact that close regular level sets consist of four Liouville tori [**6**].

Exercise IV.19 (The Symplectic Form on Coadjoint Orbits). For each ξ in $\mathfrak{g}^\star$, Define a skew-symmetric bilinear form ω_ξ on $\mathfrak{g}$ by setting

$$\omega_\xi(X,Y) = \langle \xi, [X,Y]\rangle = -\langle \operatorname{ad}^\star_X \xi, Y\rangle.$$

What is the kernel of ω_ξ? Show that the formula

$$\widetilde{\omega}_\xi(\operatorname{ad}^\star_X \xi, \operatorname{ad}^\star_Y \xi) = \omega_\xi(X,Y)$$

defines a symplectic form $\widetilde{\omega}$ on the coadjoint orbit of ξ and show that the Poisson bracket corresponding to $\widetilde{\omega}$ is the Poisson bracket of $\mathfrak{g}^\star$.

Exercise IV.20. Assume that G is a Lie group acting on a manifold V. Show that the formula

$$g\cdot(v,\varphi) = \big(g\cdot v, {}^tT_v g^{-1}(\varphi)\big)$$

defines an action of G on $T^\star V$. Show that this action preserves the Liouville form α.

For $X \in \mathfrak{g}$, let $\widetilde{X}$ and $\approx{X}$ be the corresponding fundamental fields on V and $T^\star V$, respectively. Define a map

$$\mu_X : T^\star V \longrightarrow \mathbb{R}$$

by $\mu_X(v,\varphi) = \langle \varphi, \widetilde{X}(v)\rangle$. Show that $\mu_X(v,\varphi) = i_{\approx{X}}\alpha$, and deduce that

$$d\mu_X = -i_{\approx{X}} d\alpha,$$

implying that $\approx{X}$ is the Hamiltonian vector field corresponding to μ_X. Finally, define the *moment map*

$$\mu : T^\star V \longrightarrow \mathfrak{g}^\star$$

by $\langle \mu(v,\varphi), X\rangle = \mu_X(v,\varphi)$. Show that μ is equivariant, that is

$$\operatorname{Ad}^\star_g \mu(v,\varphi) = \mu(g\cdot(v,\varphi)).$$

Consider the natural action of $\mathrm{SO}(n)$ by isometries on the unit sphere S^{n-1} in $\mathbb{R}^n$. Determine the moment map for the action on $T^\star S^{n-1}$.

Exercise IV.21. We use the notation from Exercise IV.18. Suppose that the group K also acts on a manifold V, and let $\mu : T^\star V \to \mathfrak{k}^\star$ be the moment map (see Exercise IV.20). Finally, suppose that $j : V \to \mathfrak{p}^\star$ is a K-equivariant map, and consider the map

$$\begin{aligned} \widetilde{\mu} : T^\star V &\longrightarrow \mathfrak{k}^\star \oplus \mathfrak{p}^\star \\ (v,\varphi) &\longmapsto (\mu(v,\varphi), j(v)). \end{aligned}$$

We want to show that this map preserves Poisson brackets; that is, if f and g are two functions on $\mathfrak{k}^\star \oplus \mathfrak{p}^\star$, then

$$\{f\circ\widetilde{\mu}, g\circ\widetilde{\mu}\}_{T^\star V} = \{f,g\}_{\mathfrak{k}^\star\oplus\mathfrak{p}^\star}\circ\widetilde{\mu}.$$

First, show that it suffices to check this when f and h are linear functions on $\mathfrak{k}^\star\oplus\mathfrak{p}^\star$, in which case the above relation can be written as

$$\{X\circ\widetilde{\mu}, Y\circ\widetilde{\mu}\}_{T^\star V} = [X,Y]_{\mathfrak{g}}\circ\widetilde{\mu}.$$

Show that this holds when $X, Y \in \mathfrak{p}$, then when $X, Y \in \mathfrak{k}$, and finally when $X \in \mathfrak{p}$ and $Y \in \mathfrak{k}$ to achieve the desired result.

EXERCISE IV.22. Let $\mathfrak{g}$ be a Lie algebra, and let $\mathfrak{a}$ and $\mathfrak{b}$ be Lie sub-algebras. Suppose that the vector space $\mathfrak{g}$ is a direct sum of the two subspaces $\mathfrak{a}$ and $\mathfrak{b}$, and let P_+ and P_- be the projections onto $\mathfrak{a}$ and $\mathfrak{b}$, respectively, corresponding to this decomposition. Show that the linear map

$$R = \frac{1}{2}(P_+ - P_-) : \mathfrak{g} \longrightarrow \mathfrak{g}$$

is an R-matrix.

Let $\mathfrak{g} = \mathfrak{sl}(n;\mathbb{C})$ be the Lie algebra of complex matrices with trace zero, $\mathfrak{a}$ the Lie sub-algebra of lower triangular matrices (with trace zero), and $\mathfrak{b}$ that of skew-symmetric matrices. Show that $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{b}$ (as vector spaces).

Consider the invariant bilinear form on $\mathfrak{g}$

$$\langle X, Y \rangle = \operatorname{Tr}(XY).$$

Show that it is non-degenerate, that $\mathfrak{a}^\perp$ is the subspace of strictly lower triangular matrices (with diagonal entries all zero), and that $\mathfrak{b}^\perp$ is that of symmetric matrices.

Consider the function φ defined by $\varphi(A) = \frac{1}{2}\operatorname{Tr}(A^2)$. Determine its gradient. Show that the Hamiltonian system associated to φ on $\mathfrak{a}^\star = \mathfrak{b}^\perp$ is

$$\frac{dA}{dt} = [A, B],$$

where B is the skew-symmetric matrix whose entries above the diagonal are those of the symmetric matrix A. Show that the functions $\operatorname{Tr}(A^k)$ are first integrals of this system[14].

[14] This construction is the origin of a Lax equation for a famous Hamiltonian system, the (non-periodic) Toda system. See [**1**].

APPENDIX A

What One Needs to Know About Differential Galois Theory

In this appendix, I give the definitions of the basic notions of differential Galois theory (differential field, Picard–Vessiot extension, Galois group). For more details, see [**47**] and [**58**].

A.1. Differential Fields

We consider here a *differential field*, that is, a field k equipped with a *derivation*, i.e., an additive operator $'$ that satisfies

$$(ab)' = a'b + ab'.$$

We suppose that the field of constants of k (the set of elements a such that $a' = 0$; see Exercise A.3) is isomorphic to $\mathbb{C}$ (it does not really matter that this is exactly $\mathbb{C}$, what is especially important is that this is an *algebraically closed* field of characteristic zero).

We will also use *differential rings*, with the obvious definition. A *differential ideal* of a differential ring is an ideal that is closed under derivation.

Example A.1.1. Let Γ be a connected open subset of $\mathbb{C}$, and let $k = \mathcal{M}(\Gamma)$ be the field of meromorphic functions on Γ. This is a differential field (with the usual derivation of functions) whose field of constants is indeed isomorphic to $\mathbb{C}$.

This is the most basic example one should keep in mind. More generally, one can consider the field of meromorphic functions on a Riemann surface, but we have to choose a derivation.

Example A.1.2. Let Γ be a connected (but not necessarily compact) Riemann surface equipped with a holomorphic vector field X. Let $k = \mathcal{M}(\Gamma)$ be the field of meromorphic functions on Γ. The vector field X defines a derivation[1] on k

$$X(fg) = X(f)g + fX(g).$$

If X is not identically zero, then the functions f such that $X(f) = 0$ are constant on the open set where X is nonzero, and thus are constant on Γ.

Example A.1.3. The field $\mathbb{C}(t)$ of rational fractions in the variable t with the derivation $\dfrac{d}{dt}$ is a differential field.

[1]Indeed, a vector field is also a derivation of functions.

The field $\mathbb{C}(t)$ is the field of meromorphic functions on the Riemann sphere, and thus it is a special case of the example of the field of meromorphic functions on a Riemann surface (Example A.1.2). See Exercise A.4.

Naturally, a differential field that is a field extension of the field k whose derivation is an extension of that of k is called a *differential extension* of k.

Of course, there are algebraic extensions in this theory. For example, $\mathbb{C}(t,u)/(u^2-t)$ is an extension of degree 2 of $\mathbb{C}(t)$ and we can extend the derivation by setting $u' = 1/2u$. But the real purpose of this theory is to study extensions (usually transcendental) obtained in solving (linear) differential equations. Thus we consider an operator L on k of the form

$$L \cdot y = \sum_{i=0}^{m} a_i y^{(i)}$$

where $a_i \in k$ and $y^{(i)}$ denotes the i-th derivative of y. More generally, we can consider an $m \times m$ matrix A with coefficients in k and study the extensions of k in which the equations $L \cdot y = 0$ or $Y' + AY = 0$ have solutions.

EXAMPLES A.1.4. We consider the example where $k = \mathbb{C}(t)$.

(1) All the solutions of the differential equation $y' = 0$ are in k.
(2) This is not the case for $y' = ay$, where we need to perform quadratures and to use exponentials.
(3) Let α be a complex number. A necessary and sufficient condition that all the solutions of the differential equation $y' = \alpha y/t$ are in k is that $\alpha \in \mathbb{Z}$.

Exercises A.1, A.2, A.3, A.4.

A.2. The Picard–Vessiot Extension

We consider a differential system $Y' + AY = 0$, where A is an $m \times m$ matrix with entries in k. It is not hard to check that its solutions form a vector space of dimension at most m over the field of constants $\mathbb{C}$. We will show that there always exists a differential extension K of k for which the solutions form a vector space of dimension exactly m. We will even construct a "minimal" extension having this property, in the sense that:

- its field of constants is $\mathbb{C}$ (so we do not need to add constants),
- there exists a "fundamental" solution with entries in K, that is, a square matrix Z whose columns form a basis for the solutions, or in other words, an invertible matrix Z such that $Z' + AZ = 0$,
- the field K is generated over k by the entries of Z.

Such an extension, the differential version of the decomposition field for an algebraic equation, is called a "Picard–Vessiot extension". A theorem due to Kolchin [**51**] asserts that such an extension exists and is unique up to isomorphism. We now give a simple construction of this extension (see [**58**] or [**55**]).

We begin by crudely constructing an extension that we know contains a fundamental solution by considering the ring

$$\mathfrak{A} = k[Z_{1,1}, \ldots, Z_{1,m}, Z_{2,1}, \ldots, Z_{m,m}, (\det Z)^{-1}].$$

Here we have added m^2 variables, considered as the entries of a matrix Z and we require that this matrix Z be invertible[2]. Then we define a derivation on the ring $\mathfrak{A}$ by extending that of k via the rule

$$Z'_{i,j} = (Z')_{i,j} \text{ and } Z' = -AZ.$$

By definition, Z is a fundamental solution and $\mathfrak{A}$ is generated over k by the entries of this solution. Unfortunately, $\mathfrak{A}$ is not a field.

We then choose a differential ideal $\mathfrak{q}$ in $\mathfrak{A}$ that is maximal among the differential ideals. Note that $\mathfrak{q}$ may not be a maximal ideal; however, as the following lemma shows, $\mathfrak{A}/\mathfrak{q}$ is still an integral domain[3].

LEMMA A.2.1. *If $\mathfrak{q}$ is a differential ideal of a differential ring $\mathfrak{A}$, then its radical $\sqrt{\mathfrak{q}}$ is a differential ideal.*

PROOF. Let $z \in \sqrt{\mathfrak{q}}$, so that $z^r \in \mathfrak{q}$ for some r. We differentiate z^r and show using induction that

$$(z^r)^{(r)} = r(z')^r + za$$

for some $a \in \mathfrak{A}$, or in other terms, we have

$$(z')^r = \frac{1}{r}\left((z^r)^{(r)} - za\right).$$

We know that $z^r \in \mathfrak{q}$, which is a differential ideal. Hence $(z')^r \in \sqrt{\mathfrak{q}}$ and thus $z' \in \sqrt{\mathfrak{q}}$. □

LEMMA A.2.2. *If $\mathfrak{q}$ is a differential ideal that is maximal among differential ideals of the differential ring $\mathfrak{A}$, then the quotient ring $\mathfrak{B} = \mathfrak{A}/\mathfrak{q}$ is an integral domain.*

PROOF. First, note that $\mathfrak{B}$ contains no nilpotent elements. Indeed, the previous lemma asserts that the nilpotent elements form a differential ideal of $\mathfrak{B}$ (as it is the radical of the zero differential ideal), and $\mathfrak{B}$ contains no proper differential ideal.

Let a be a nonzero element of $\mathfrak{B}$. We consider the ideal I of elements that annihilate a. We want to show that $I = \{0\}$. For b in I, we have

$$0 = (ab)' = a'b + ab',$$

so that

$$ab' = -a'b \text{ and } (ab')^2 = -a'(ba)b' = 0.$$

Thus ab' is nilpotent and thus it is zero. Thus $b' \in I$. Hence I is a differential ideal of $\mathfrak{A}$, and thus $I = \{0\}$ and $\mathfrak{B}$ is an integral domain. □

Now knowing that the quotient ring $\mathfrak{B} = \mathfrak{A}/\mathfrak{q}$ is an integral domain, we define K (which is the Picard–Vessiot extension we have been looking for) as its field of fractions. One can show (once again, see [**58**]) that:

[2]In the terminology of differential equations, $\det Z$ is the *Wronskian*.

[3]The proof given here is inspired—to say the very least—by the appendix of [**55**].

- The field K is a differential field. (This follows since $\mathfrak{B}$ is a differential ring because the ideal $\mathfrak{q}$ is differential, and the derivation on $\mathfrak{B}$ can be extended to one on K (Exercise A.2).)
- The field of constants of K is the same as that of k (Exercise A.7).
- The matrix z of the images of $Z_{i,j}$ is a fundamental matrix of solutions (it has been constructed for exactly that purpose).

To establish the uniqueness of K up to isomorphism, one uses again the fact that the field of constants is algebraically closed, but we will not discuss this here (see [**58**] instead). Consequently, the result of the construction does not depend on the choice of ideal $\mathfrak{q}$, a property we will exploit below.

Exercises A.5, A.6, A.7, A.8.

A.3. The Galois Group

As in the theory of algebraic equations, the Galois group of the equation $Y' + AY = 0$ is the group of (differential!) automorphisms of K fixing k. That is, these are the automorphisms $\sigma : K \to K$ such that

$$\sigma_{|_k} = \mathrm{Id}_k \text{ and } (\sigma(u))' = \sigma(u') \quad \text{for all } u \text{ in } K.$$

We write this group as $\mathrm{Gal}(Y' + AY)$.

A.3.1. Representation in the Linear Group. As the Galois group of an algebraic equation of degree m can be represented in the symmetric group $\mathfrak{S}_m$, the Galois group of a differential equation of order m can be represented in the linear group $\mathrm{GL}(m;\mathbb{C})$ (here m is the order of the differential equation or the size of matrix A). Here is how this is done. If z_i is a column of a fundamental matrix of solutions z, it is a solution to the differential equation. Thus we have

$$z_i' + Az_i = 0,$$

and thus,

$$0 = \sigma(z_i' + Az_i) = \sigma(z_i') + A\sigma(z_i),$$

since σ fixes k and A has entries in k. Finally, since σ is differential, we see that

$$(\sigma(z_i))' + A\sigma(z_i) = 0.$$

Hence $\sigma(z_i)$ is also a solution of the equation and we can express it in the basis $(z_1, \ldots, z_m)$,

$$\sigma(z_i) = \sum_{j=1}^{m} c_{j,i} z_j$$

for some constants $c_{j,i} \in \mathbb{C}$. Since σ is an automorphism, $(\sigma(z_1), \ldots, \sigma(z_m))$ forms a basis for the vector space of solutions and the matrix c with entries $(c_{j,i})$ is invertible. The morphism

$$\begin{aligned} \mathrm{Gal}(Y' + AY) &\longrightarrow \mathrm{GL}(m;\mathbb{C}) \\ \sigma &\longmapsto c \end{aligned}$$

is injective. We have shown that the Galois group is indeed represented in $\mathrm{GL}(m;\mathbb{C})$. We could also show that it is an *algebraic* subgroup of $\mathrm{GL}(m;\mathbb{C})$

(this is a theorem due to Kolchin). The proof consists of "transforming" the ideal $\mathfrak{q}$ used in the construction of the Picard–Vessiot extension into an ideal of $\mathbb{C}[Y_{1,1}\dots,Y_{m,m},\det(Y_{i,j})^{-1}]$ (the ring of the linear group) that vanishes exactly on the elements of the Galois group. It uses techniques of commutative algebra that I have not assumed the readers of this book are familiar with; see [**58**] for this proof.

REMARK A.3.1. The representation obtained depends on the choice of the basis $(z_1,\dots,z_m)$, but a different choice clearly yields a conjugate representation.

A.3.2. Examples. Here are two very classical examples of calculations of Galois groups.

EXAMPLE A.3.2. We return to the equation $y' = \dfrac{\alpha}{t}y$ over $k = \mathbb{C}(t)$ (see Example A.1.4 and Exercise A.8).

- The extension $k(t^\alpha)/k$ is trivial when $\alpha \in \mathbb{Z}$.
- It is algebraic if and only if $\alpha \in \mathbb{Q}$. If $\alpha = p/q$ is an irreducible fraction with $q > 0$, then the Picard–Vessiot extension is $\mathbb{C}(t)[u]/\langle u^q - t^p\rangle$, which is algebraic of degree q. Since $u^q = t^p$, we have
$$u' = \frac{p}{q}\frac{t^{p-1}}{u^{q-1}},$$
implying that the differential and algebraic Galois groups coincide. Here the algebraic Galois group is the group of σ defined by $\sigma(u) = \zeta u$ for ζ a q-th root of 1, and is thus isomorphic to $\mathbb{Z}/q\mathbb{Z}$.
- If $\alpha \notin \mathbb{Q}$, then the extension is transcendental (with transcendence degree 1). The Galois group is the linear group $\mathbb{C}^\star$ itself, which acts on the 1-dimension space of solutions by multiplication.

EXAMPLE A.3.3 (The Cauchy Equation). We consider the equation
$$y'' - \frac{a}{t^2}y = 0, \qquad a \in \mathbb{C}$$
over $k = \mathbb{C}(t)$. We consider first its relation with the first-order equations
$$y' - \frac{\alpha}{t}y = 0$$
for $\alpha \in \mathbb{C}$ to be chosen. If u is a solution of this equation, then $tu' = \alpha u$. By differentiating and multiplying by t, we obtain
$$t^2u'' + (\alpha - \alpha^2)u = 0.$$
Thus if α is chosen so that
$$\alpha^2 - \alpha - a = 0,$$
the solutions of the equation $ty' = \alpha y$ are also solutions of $t^2y'' = ay$.

Moreover, if α_1 and α_2 are the two roots of this quadratic equation and u_1 and u_2 are solutions of the corresponding differential equations, we have
$$\begin{vmatrix} u_1 & u_2 \\ u_1' & u_2' \end{vmatrix} = (\alpha_2 - \alpha_1)\frac{u_1u_2}{t}$$

so that if $\alpha_1 \neq \alpha_2$, the two solutions obtained are independent. Assume now that $a \neq -1/4$, so that u_1 and u_2 are independent. Then we obtain the Picard–Vessiot extension of the equation $y'' = ay/t^2$ by composing those of the $y' = \alpha_i y/t$. If $y' = \alpha_i y/t$, we have that

$$\sigma(y)' = \sigma(y') = \sigma\left(\frac{\alpha_i}{t}y\right) = \frac{\alpha_i}{t}\sigma(y)$$

for any σ in the Galois group, since σ fixes $\mathbb{C}(t)$. Hence the Galois group fixes the solution lines of the two first-order equations and thus is contained in the subgroup of $\mathrm{GL}(2;\mathbb{C})$ consisting of diagonal matrices. In particular, it is Abelian.

Exercises A.9, A.10, A.11.

A.3.3. The Galois Correspondence. As in classical Galois theory, there is a correspondence between the intermediate differential extensions of a differential extension K/k and the (algebraic) subgroups of the Galois group $\mathrm{Gal}(K/k)$. An analog of the "fundamental theorem" can be proved (see [**58**]). This states that if K/k is a Picard–Vessiot extension and L is a differential subfield of K and an extension of k, then the extension L/k is a Picard–Vessiot extension if and only if the Galois group of K/L is normal in that of K/k. In this case, we have

$$\mathrm{Gal}(L/k) \cong \mathrm{Gal}(K/k) \big/ \mathrm{Gal}(K/L).$$

A.4. The Galois Group of a Hamiltonian System is Symplectic

I show now that, in the case of a Hamiltonian system, the Galois group is in fact a subgroup of the symplectic group $\mathrm{Sp}(2n;\mathbb{C})$ (which is itself a subgroup of $\mathrm{GL}(2n;\mathbb{C})$), that is to say, I prove Lemma III.2.3.

Consider the complex symplectic vector space $\mathbb{C}^{2n}$. The symplectic form is the bilinear form whose matrix in the canonical basis is $J = \begin{pmatrix} 0 & \mathrm{Id} \\ -\mathrm{Id} & 0 \end{pmatrix}$. The symplectic group $\mathrm{Sp}(2n;\mathbb{C})$ consists of matrices M satisfying the algebraic equation ${}^tMJM = J$ (see Section I.1.a).

Proof of Lemma III.2.3. First, we consider the group $\mathrm{GL}(2n;\mathbb{C})$ from an algebraic point of view. This is the subspace of the set of matrices defined by the equation $\det Z \neq 0$. In other words, the ring of algebraic functions on the linear group $\mathrm{GL}(2n;\mathbb{C})$ is the ring $\mathbb{C}[Z_{i,j}, (\det Z)^{-1}]$. (Note the analogy with the ring $\mathfrak{A}$ used in the construction of the Picard–Vessiot extension, which is the ring of $\mathrm{GL}(m;k)$.)

We now look at its subgroup $\mathrm{Sp}(2n;\mathbb{C})$, which is defined by the ideal I generated by ${}^tZJZ - J$. We will make a clever choice for the ideal $\mathfrak{q}$ utilized in the construction of the Picard–Vessiot extension. This should be a (maximal differential) ideal of the ring $\mathfrak{A} = k[Z_{i,j}, (\det Z)^{-1}]$, the ring of the group $\mathrm{GL}(2n;k)$. It is natural to consider the ideal $(I) = I\mathfrak{A}$ of the group $\mathrm{Sp}(2n;k)$ and then to choose $\mathfrak{q}$ to be a maximal differential ideal containing (I). The key in making this choice possible is the following lemma.

Lemma A.4.1. *If $A \in \mathfrak{sp}(2n;\mathbb{C})$, then the ideal (I) is a differential ideal.*

Recall that A, the matrix of the system, is utilized to define the derivation of $\mathfrak{A}$ and that, in the case of a Hamiltonian system, it lies in the Lie algebra of the symplectic group (this is Lemma III.2.4). Let us assume Lemma A.4.1 for now and see how to use it to complete the proof.

Since $\mathfrak{q}$ is a maximal ideal containing (I), the matrix z that is the image of Z in $\mathfrak{B} = \mathfrak{A}/\mathfrak{q}$ lies in the symplectic group on $\mathfrak{B}$. Moreover, the Galois group of $Y' + AY$, the group of differentiable k-automorphisms of the field of fractions K of $\mathfrak{B}$, is also the group of differential k-automorphisms of the ring $\mathfrak{B}$.

Let σ be an element of this Galois group. Since $z \in \mathrm{Sp}(\mathfrak{B})$, we also have that $\sigma(z) \in \mathrm{Sp}(\mathfrak{B})$ (the equations defining the symplectic group have coefficients in the field of constants $\mathbb{C}$). As noted above, the matrix $\sigma(z)$ is a fundamental matrix of solutions, and thus it can be expressed in terms of a matrix $c \in \mathrm{GL}(2n;\mathbb{C})$

$$\sigma(z) = zc.$$

In this equation, z and $\sigma(z)$ are symplectic with entries in $\mathfrak{B}$, while c is *a priori* only in the linear group, but with *constant* entries. Thus, in fact, $c \in \mathrm{Sp}(\mathfrak{B}) \cap \mathrm{GL}(2n;\mathbb{C}) = \mathrm{Sp}(2n;\mathbb{C})$.

In conclusion, the representation $\sigma \mapsto c$ of the Galois group in the linear group has values in the symplectic group. $\square$

Proof of Lemma A.4.1. We want to show that the ideal (I) generated by the equations of the symplectic group is differential, i.e., that for any u in (I), u' is also in (I). The elements of (I) are of the form

$$u = \sum a_{i,j} \left({}^tZJZ - J\right)_{i,j} \quad \text{with } a_{i,j} \in \mathfrak{A}.$$

Thus

$$u' = \sum a'_{i,j} \left({}^tZJZ - J\right)_{i,j} + \sum a_{i,j} \left({}^tZJZ - J\right)'_{i,j}.$$

The first term is clearly in (I). Let us show that the second is zero. Indeed, we have

$$\begin{aligned}\left({}^tZJZ - J\right)' &= {}^tZ'JZ + {}^tZJZ' \quad \text{since } J \text{ is constant}\\ &= -{}^tZ{}^tAJZ - {}^tZJAZ \quad \text{since } Z' = -AZ\\ &= -{}^tZ({}^tAJ + JA)Z\\ &= 0 \quad \text{since } A \in \mathfrak{sp}(2n).\end{aligned}$$

$\square$

This proof can be easily adapted to more general situations than the symplectic group. The good statement is:

Proposition A.4.2. *Let G be an algebraic subgroup of the linear group $\mathrm{GL}(m;\mathbb{C})$, and let $\mathfrak{g}$ be its Lie algebra. If $A \in \mathfrak{g} \otimes k$, then $\mathrm{Gal}(Y' + AY)$ is a subgroup of G.*

Proof. Replace I by the ideal defining the subgroup G in the ring of $\mathrm{GL}(m;\mathbb{C})$. To check that the proof of Lemma A.4.1 applies to this situation, we just need to check that $f(Z)' = 0$ for $f \in I$ and $A \in \mathfrak{g}$. We calculate

$$f(Z)' = (df)_Z(Z') = -(df)_Z(AZ),$$

and it remains to check that $(df)_Z(AZ) = 0$ if $(df)_{\mathrm{Id}}(A) = 0$. This follows by setting Z to Id and using the fact that f is a relation defining a subgroup. □

A.5. An Example: the Airy Equation

I work through the example of the Airy equation in detail since I have utilized it in Section III.5.

A.5.1. Calculation of the Galois Group. Consider the differential equation $y'' - ty = 0$ over the field $\mathbb{C}(t)$ of rational fractions. We prove the following proposition.

PROPOSITION A.5.1. *The identity component of the Galois group of* $y'' - ty = 0$ *over* $\mathbb{C}(t)$ *is not Abelian.*

To do this, we prove two easy lemmas.

LEMMA A.5.2. *The Galois group* $\mathrm{Gal}(y'' - ty)$ *is a subgroup of* $\mathrm{SL}(2;\mathbb{C})$.

PROOF. Let (u, v) be a basis of the vector space of solutions in a Picard–Vessiot extension. We calculate the derivative of the Wronskian

$$w' = \begin{vmatrix} u & v \\ u' & v' \end{vmatrix}' = (uv' - u'v)' = uv'' - u''v = u(tv) - (tu)v = 0.$$

Thus w is a constant. By multiplying one of the solutions by a constant, we can suppose that $w = 1$. Next, for $\sigma \in \mathrm{Gal}(y'' - ty)$ we have

$$1 = \sigma(1) = \sigma(w) = \begin{vmatrix} \sigma(u) & \sigma(v) \\ \sigma(u') & \sigma(v') \end{vmatrix} = \det\sigma \begin{vmatrix} u & v \\ u' & v' \end{vmatrix} = \det\sigma.$$

Hence the Galois group is indeed contained in $\mathrm{SL}(2;\mathbb{C})$. □

LEMMA A.5.3. *The Picard–Vessiot extension is not finite.*

In particular, we note that the Galois group $\mathrm{Gal}(y'' - ty)$ is not finite.

PROOF. If it were finite, then the solutions, which we know are entire, would be solutions of polynomial equations with coefficients in $\mathbb{C}(t)$, and thus polynomials (Exercise III.15). But this is impossible (Exercise III.14). □

PROOF OF PROPOSITION A.4.2. Suppose that the identity component of the Galois group were Abelian. It is not difficult to show[4] that if an infinite algebraic subgroup G of $\mathrm{SL}(2;\mathbb{C})$ has an Abelian identity component, then

- either the elements of G are simultaneously triangular
- or G has a subgroup (that we call H) of index 2 whose elements are simultaneously diagonal.

[4]See Proposition A.5.5.

In the first case, there is a solution u of the Airy equation such that $\sigma(u)$ is collinear with u for all $\sigma \in \mathrm{Gal}(y'' - ty)$. Then we have

$$\sigma\left(\frac{u'}{u}\right) = \frac{\lambda u'}{\lambda u} = \frac{u'}{u},$$

implying that $u'/u \in \mathbb{C}(t)$. Let $z = u'/u$. Then, by differentiating $u' = zu$, we find that $(t - z^2)u = z'u$ and thus z is a solution of the "Riccati" differential equation $z' = t - z^2$.

In the second case, let (u, v) be a basis of common eigenvectors for the elements of H. Then, we see, for $\sigma \in G$,

$$\text{either } \sigma^2\left(\frac{u'}{u}\right) = \frac{u'}{u}, \qquad \text{or } \sigma\left(\frac{u'}{u}\right) = \frac{u'}{u}.$$

Hence u'/u lies in a quadratic extension L of $\mathbb{C}(t)$ (with Galois group G/H). Similarly, $z = u'/u$ is a solution of the differential equation $z' = t - z^2$.

Thus, if we suppose that G° is Abelian, this would imply the Riccati equation $z' = t - z^2$ has either a rational solution or a solution in a quadratic extension of $\mathbb{C}(t)$. The end of the proof consists in showing that this is impossible. This is elementary but rather calculation-intensive.

In the first case, we quickly arrive at a contradiction. We write z as a quotient of polynomials

$$z = \frac{f(t)}{g(t)}, \quad \deg f = m, \quad \deg g = n.$$

By differentiating, the Riccati equation implies that

$$f'g - fg' + f^2 = tg^2.$$

If $m > n$, then the degree of the left side is $2m$, which is strictly greater than the degree $2n+1$ of the right side. If $m \leqslant n$, then degree of the left side is $\leqslant 2n$, while that of the right is $2n+1$. This is the contradiction we were looking for.

Obtaining a contradiction in the second case is somewhat more involved, but the idea is the same, i.e., to show there would exist a rational function which satisfies a complicated differential equation that, in fact, has no rational solution. The function z satisfies

$$\begin{cases} z' = t - z^2 \\ z^2 + rz + s = 0, \quad r, s \in \mathbb{C}(t). \end{cases}$$

We will eliminate z and s to obtain a differential equation satisfied by the rational function r. We differentiate the above algebraic equation to obtain

$$2zz' + r'z + rz' + s' = 0,$$

and then utilize the Riccati equation to obtain

$$-2z^3 - rz^2 + (2t + r')z + rt + s' = 0.$$

Then we add $2z(z^2 + rz + s)$, which equals zero, to see that

$$rz^2 + (2t + 2s + r')z + rt + s' = 0,$$

which is another quadratic equation satisfied by z. We deduce that

$$\begin{cases} 2t + 2s + r' = r^2 \\ rt + s' = rs. \end{cases}$$

The first equation gives us s as a function of r

$$s = \frac{r^2 - r'}{2} - t.$$

By differentiating this same equation, we express s' in terms of r

$$s' = rr' - 1 - \frac{r''}{2},$$

and we put this into the second equation. After all this gymnastics, we have shown that r must be a solution of the differential equation

$$r'' - 3rr' + r^3 - 4rt + 2 = 0.$$

Finally, we show that there is no rational fraction $r = f/g$ satisfying this differential equation. Set $m = \deg f$ and $n = \deg g$. We substitute $r = f/g$ into the differential equation supposedly satisfied by r and multiply it by g^3 to eliminate denominators, the numerator becoming a linear combination of the polynomials

$$g^2 f'',\ fgg'',\ gf'g',\ f(g')^2,\ ff'g,\ f^2g',\ f^3,\ tfg^2,\ g^3.$$

If $m = n$, the term tfg^2 does not vanish. If $m > n$, we have the same problem with f^3, and if $n > m+1$, g^3 remains. Unfortunately, if $n = m+1$, the terms tfg^2 and g^3 have the same degree, so we need to look at their coefficients a little more closely.

We expand r as a Laurent series around one of its poles a

$$r = \frac{c_k}{(t-a)^k} + \cdots + \frac{c_1}{(t-a)} + \text{ a holomorphic function.}$$

To be a solution of the complicated differential equation satisfied by r, we must have equality between the two largest of the numbers $k+2$, $2k+1$, $3k$, so $k = 1$. In this same differential equation, a term $c/(t-a)$ gives

$$\frac{2c}{(t-a)^3} \text{ in } r'',\ \frac{3c^2}{(t-a)^3} \text{ in } -3rr', \text{ and } \frac{c^3}{(t-a)^3} \text{ in } r^3,$$

and thus c must satisfy $c(c^2+3c+2) = 0$, and either $c = -1$ or $c = -2$ if a is truly a pole of r.

Finally, since $r = f/g$ with $\deg g = \deg f + 1$, we have $f(t) = \alpha t^m + \cdots$, $g(t) = t^{m+1} + \cdots$, and

$$r(t) = \sum_{i=1}^{m+1} \frac{c_i}{t - a_i} \quad \text{with } c_i = -1 \text{ or } -2.$$

Hence $\alpha = c_1 + \cdots + c_{m+1} \in \mathbb{Z}$. But to eliminate the terms in tfg^2 (coming from $-4tr$) and g^3 (coming from $+2$), we must have $-4\alpha + 2 = 0$ which is impossible since α is an integer! □

Remark A.5.4. By refining the proof a bit, one can show that the Galois group is exactly $\mathrm{SL}(2;\mathbb{C})$. See [**47**], the origin of this "algebraic" proof.

A.5.2. Subgroups of $\mathrm{SL}(2;\mathbb{C})$**.** For completeness, I establish the properties of the subgroups of $\mathrm{SL}(2;\mathbb{C})$ that I utilized in the proof of Proposition A.4.2.

PROPOSITION A.5.5. *Let G be an algebraic subgroup of* $\mathrm{SL}(2;\mathbb{C})$ *whose identity component G° is an Abelian group. Then*

- *either G is finite,*
- *or G has a subgroup of index 2 consisting of elements diagonalized by the same basis,*
- *or all the elements of G are triangular in the same basis.*

PROOF. We begin by examining the identity component G°.

- Either $G^\circ = \{\mathrm{Id}\}$,
- or G° contains an element $h \neq \pm\,\mathrm{Id}$, in which case:
  - either h is diagonalizable in a basis (u, v): since $G \subset \mathrm{SL}(2;\mathbb{C})$, the eigenvectors of h are distinct and since G° is Abelian, all the elements of G° are diagonalizable in the basis (u, v),
  - or h is not diagonalizable, in which case, there is an eigenvector u (since h is a complex endomorphism) that is an eigenvector for all the elements of the Abelian group G°. Thus all elements are triangular in a basis (u, v).

We now look at the group G itself. When $G^\circ = \{\mathrm{Id}\}$, G is finite (using the fact that G has a finite number of connected components; this is the only place where we use, in a very weak way, the fact that G is algebraic. If $G^\circ \neq \{\mathrm{Id}\}$, let us show that G fixes the set of common eigenvectors of the elements of G°. Let u be such a vector and $g \in G$, $h \in G^\circ$, we have

$$(ghg^{-1})(g(u)) = gh(u) = g(\lambda u) = \lambda g(u),$$

thus $g(u)$ is an eigenvector of ghg^{-1}. We recall that the identity component G° is a normal subgroup of G and therefore ghg^{-1} is an element of G°.

- If the subgroup G° is not diagonalizable, we conclude that the unique (common) eigenline of its elements is fixed by G, and thus all the elements of H are triangular in a basis (u, v) where the first vector is an eigenvector.
- If (u, v) is a basis for which the subgroup G° is diagonal, then any element of G sends u to an eigenvector of G°, that is to a λu or λv.
  - If they all fix the line generated by u, they also fix the one generated by v, so all the elements of G are diagonal in the basis (u, v).
  - If an element g of G sends u to λv, then the matrix of g in the basis (u, v) is $\begin{pmatrix} 0 & -\lambda^{-1} \\ \lambda & 0 \end{pmatrix}$, so that $g^2 = -\,\mathrm{Id}$. Thus G has a subgroup of index 2 consisting of diagonal elements in the basis (u, v). $\square$

Easy Exercises on Differential Galois Theory

EXERCISE A.1. Is the ring of $\mathcal{C}^r$ functions on $\mathbb{R}$ a differential ring (for the usual derivative of functions)? What about the ring of $\mathcal{C}^\infty$ functions?[5]

EXERCISE A.2. Show that a derivation on an integral domain can be extended uniquely to a derivation on its field of fractions.

EXERCISE A.3 (Field of Constants). Let k be a differential field and let $C \subset k$ be the set of its *constants*, i.e., the elements a such that $a' = 0$. Show that C is a field, that $'$ is a C-linear operator, and that k is a C-algebra.

EXERCISE A.4. Let Γ be the complex projective line $\mathbb{P}^1(\mathbb{C})$ considered as $\Gamma = \mathbb{C} \cup \{\infty\}$. Check that the field of meromorphic functions on Γ is the field $\mathbb{C}(t)$ of rational functions and that the vector field $\dfrac{d}{dt}$ on $\mathbb{C}$ extends to a vector field X on Γ. Determine its zeros (see Figure 1).

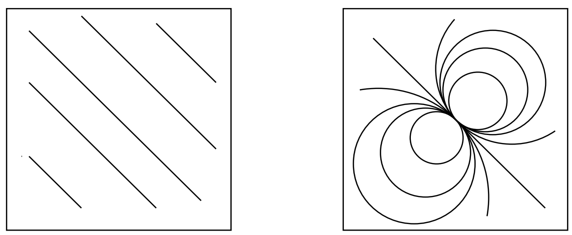


FIGURE 1. Trajectories of $\dfrac{d}{dt}$ near 0 and ∞

EXERCISE A.5. Write a system $Y' + AY = 0$ that is equivalent to the equation $L \cdot y = 0$.

EXERCISE A.6. Consider the differential field $k = \mathbb{C}(t)$, the ring $\mathbb{C}[[t]]$ of formal power series, and its field of fractions K, so that K is a differential extension of k.

Show that the equation $y' = y$ has no nontrivial solution in k but has solutions in K. Explain why K is not a Picard–Vessiot extension of this equation.

EXERCISE A.7. Let k be a differential field with field of constants C and let $\mathfrak{B}$ be a differential k-algebra. Suppose that $\mathfrak{B}$ has no proper differential ideal. Show that the constants of the field of fractions of $\mathfrak{B}$ are the constants of $\mathfrak{B}$. Let $C_{\mathfrak{B}}$ be this field of constants.

Now, suppose that $C = \mathbb{C}$ and that $\mathfrak{B}$ is a finite-type k-algebra. Let $\mathfrak{m}$ be a (ordinary) maximal ideal. The projection $\mathfrak{B} \to \mathfrak{B}/\mathfrak{m}$ identifies k with a subfield. Show that the extension $(\mathfrak{B}/\mathfrak{m})/k$ is algebraic and that any constant of $\mathfrak{B}$ is algebraic over $\mathbb{C}$. Finally, show that $C_{\mathfrak{B}} = \mathbb{C}$.

[5]The ring of $\mathcal{C}^\infty$ functions is never an integral domain, unlike the ring of holomorphic functions on a connected open subset of $\mathbb{C}$, which is because of the principle of isolated zeros. This is one of the reasons that differential Galois theory is restricted to analytic functions.

Exercise A.8. Consider the differential field $k = \mathbb{C}(t)$ and the differential equation

$$y' = \frac{\alpha}{t} y \qquad (\text{for } \alpha \in \mathbb{C}).$$

Show that this equation has a rational solution if and only if $\alpha \in \mathbb{Z}$, then that it has an algebraic solution over $\mathbb{C}(t)$ if and only if $\alpha \in \mathbb{Q}$.

Show that $K = k(t^\alpha)$ is a Picard–Vessiot extension of this equation.

Exercise A.9. Given two complex numbers a and b, determine the Galois group over $\mathbb{C}(t)$ of the linear differential equation with constant coefficients

$$y'' - 2by' + a^2 y = 0.$$

Exercise A.10. Let k be a differential field with field of constants $\mathbb{C}$ and derivation $a \mapsto a'$. Consider the differential equation

$$y'' - ay' + by = 0$$

and a Picard–Vessiot extension K of this equation. The Wronskian of two independent solutions u and v is denoted by

$$w = \begin{vmatrix} u & v \\ u' & v' \end{vmatrix},$$

which is an element of K. Show that $w \in k$ if and only if the Galois group is a subgroup of $\mathrm{SL}(2; \mathbb{C})$. Calculate w', and deduce that the Galois group is a subgroup of $\mathrm{SL}(2; \mathbb{C})$ if and only if $a = c'/c$ for some element c of k.

Exercise A.11. Determine the Galois group of the Cauchy equation

$$y'' - \frac{a}{t^2} y = 0,$$

depending on whether or not the roots of $\alpha^2 - \alpha - a = 0$ are in $\mathbb{Q}$.

APPENDIX B

What One Needs to Know About Algebraic Curves

In this Appendix, I describe the properties of algebraic curves and their Jacobians that are used in this book. For more about complex curves (or Riemann surfaces), see the books by Griffiths and Harris [**37**], Farkas and Kra [**28**], and Reyssat [**72**].

B.1. Complex Curves and Riemann Surfaces

B.1.1. Completion of Affine Curves. The algebraic curves appearing in Chapter IV are given by polynomial equations $P(x, y) = 0$; such curves are called "affine" curves (since they lie in the affine plane $\mathbb{C}^2$). As such they are not compact: they always have points at infinity. They can even be singular. Nonetheless, the singularities can be removed and the curves can be completed to become Riemann surfaces.

A *Riemann surface* is a complex analytic manifold of dimension 1, which we will assume here to also be compact and connected[1]. To a curve C_P defined by

$$C_P = \left\{(x, y) \in \mathbb{C}^2 \mid P(x, y) = 0\right\}$$

is associated a (well-defined) compact Riemann surface X_P such that there exist two finite subsets $E \subset C_P$ and $F \subset X_P$ with $C_P - E = X_P - F$. Thus the two surfaces coincide except for a finite number of points: X_P is compact while C_P never is, and X_P is smooth while C_P may be singular at some points of E (if C_P is smooth, then $E = \varnothing$, and X_P is a completion of C_P).

We say that the Riemann surface X_P is the "normalization" of C_P. We can also complete a singular affine curve by adding points "at infinity", namely, without changing its singular points. In fact, this cannot be avoided if we want to complete a family of curves in a uniform enough way, a situation we encounter in integrable systems, and thus in this book.

This is a very concrete operation, which we can perform by separating the branches at infinity without worrying about what happens at finite distance. A simple example is found in Exercise B.1.

Examples B.1.1.

- We complete $\mathbb{C}$ (which can be viewed as the planar affine curve with equation $y = 0$) by adding a point at infinity. The Riemann surface

[1] Although we did use non-compact Riemann surfaces in Chapter III.

thus obtained is the complex projective line $\mathbb{P}^1 = \mathbb{P}^1(\mathbb{C})$ or the "Riemann sphere".

- Consider the affine curve with equation $y^2 = Q(x)$, where $Q(x)$ is a cubic polynomial. If the roots of Q are distinct, then the curve is smooth, and we can complete it by adding a point at infinity. The Riemann surface obtained is an "elliptic curve".

B.1.2. The Genus of a Curve. The most fundamental invariant of a curve is its "genus", an integer g for which there are numerous equivalent definitions[2].

The genus fully determines the topology of the Riemann surface X, since it is the number of its handles or, equivalently, we have

$$H_1(X;\mathbb{Z}) \cong \mathbb{Z}^{2g}.$$

It also tells us a lot about the analytic structure, since it is also the dimension of the complex vector space of holomorphic forms on X, that we write (with notation that will be explained in Section B.3),

$$g = \dim H^0\left(\Omega_X^1\right),$$

where Ω_X^1 denotes the *sheaf* of germs of holomorphic forms and $H^0\left(\Omega_X^1\right)$ the space of its global sections, the space of holomorphic forms.

Examples B.1.2.

- The projective line $\mathbb{P}^1$ is a Riemann surface of genus 0 (it is a sphere).
- Elliptic curves can be obtained as quotients of $\mathbb{C}$ by lattices. Therefore, these are topologically tori and thus curves of genus 1. The Weierstrass $\wp$-function associated to the lattice Λ is the solution of a differential equation of the form
$$\wp'(z)^2 = 4\wp(z)^3 - g_2\wp(z) - g_3,$$
so that $z \mapsto (\wp(z), \wp'(z))$ parametrizes the curve with equation $y^2 = 4x^3 - g_2x - g_3$. This yields an isomorphism of $\mathbb{C}/\Lambda$ with the completion of this curve (with the poles of $\wp$ being sent to the point at infinity).

The most useful result about curves is certainly the Riemann–Roch theorem, a very precise theorem about the existence of meromorphic functions (see Section B.4.2). One of its consequences is the equivalence of the two definitions of genus I just mentioned.

An easier, but very useful, result is the Riemann–Hurwitz formula, which concerns a map $f : X \to X'$ between two Riemann surfaces. The formula in question is the result of a simple calculation of the Euler characteristic, allowing the genus of X and of X' to be compared.

For example, if the map f has degree 2, then the Riemann–Hurwitz formula asserts that the Euler characteristic $2 - 2g$ of X is twice the Euler characteristic $2 - 2g'$ of X' minus the number of ramification points.

[2]There are eighteen of them in [**72**].

Examples B.1.3.

- A double covering of $\mathbb{P}^1$ ramified at four points is a curve of genus 1.
- Curves with a map of degree 2 taking values in $\mathbb{P}^1$ are called "hyperelliptic" curves. These appear often throughout this book. They are the curves with an affine equation of the form $y^2 = Q(x)$ (the map of degree 2 is $(x, y) \mapsto x$, see Exercise B.1). There is an involution defined on these curves, called the "hyperelliptic involution", which is simply the map $(x, y) \mapsto (x, -y)$.

  By applying the Riemann–Hurwitz formula, we see that the genus of the curve is g if $\deg Q = 2g + 1$ or $2g + 2$.

B.2. Line Bundles on Curves

A (holomorphic) line bundle on a Riemann surface X is a complex surface L with a projection $\pi : L \to X$ whose fibers are complex 1-dimensional vector spaces and which is locally trivial in the sense that:

- The curve X is covered by open sets U such that there exists a homeomorphism φ_U making the diagrams

$$\begin{tikzcd} \pi^{-1}(U) \arrow[rr, "\varphi_U"] \arrow[dr, "\pi"'] & & U \times \mathbb{C} \arrow[dl] \\ & U & \end{tikzcd}$$

  commutative.
- If U and V are two open sets of the covering, then

$$\varphi_V \circ \varphi_U^{-1}(x, \alpha) = (x, f_{VU}(x) \cdot \alpha)$$

  over $U \cap V$, for some holomorphic function f_{VU} that does not vanish. The function f_{VU} is called the transition function between the two trivializations (U, φ_U) and (V, φ_V).

Examples B.2.1.

- The trivial bundle, $L = X \times \mathbb{C}$ with the natural projection (we can take $U = X$, $f_{VU} \equiv 1$).
- The tautological bundle $\mathcal{O}(-1) \to \mathbb{P}^1$ (I will explain the notation in Section B.4.1)

$$\mathcal{O}(-1) = \left\{(\ell, v) \in \mathbb{P}^1 \times \mathbb{C}^2 \mid v \in \ell\right\} \subset \mathbb{P}^1 \times \mathbb{C}^2$$

  (the points ℓ of $\mathbb{P}^1$ are the lines in $\mathbb{C}^2$ passing through 0) with the natural projection to $\mathbb{P}^1$. The bundle can be trivialized on the two open sets U_0 (where $\ell \neq \mathbb{C} \times 0$, i.e., $y \neq 0$) and U_∞ (($\ell \neq 0 \times \mathbb{C}$, $x \neq 0$).

  Write φ_0 and φ_∞ for φ_{U_0} and φ_{U_∞}. The local trivializations are

$$\begin{aligned} U_0 \times \mathbb{C} &\xrightarrow{\varphi_0^{-1}} \pi^{-1}(U_0) \\ (z, b) &\longmapsto ([z, 1], (bz, b)), \end{aligned}$$

and, similarly,

$$\begin{aligned} U_\infty \times \mathbb{C} &\xrightarrow{\varphi_\infty^{-1}} \pi^{-1}(U_\infty) \\ (z', a) &\longmapsto ([1, z'], (a, az')). \end{aligned}$$

Then by identifying $U_0 \cap U_\infty$ with $\mathbb{C}^\star$ via the coordinate z of U_0, we can write $\varphi_\infty \circ \varphi_0^{-1}$ as

$$\begin{array}{ccccccc} \mathbb{C}^\star \times \mathbb{C} & \longrightarrow & U_0 \times \mathbb{C} & \longrightarrow & U_\infty \times \mathbb{C} & \longrightarrow & \mathbb{C}^\star \times \mathbb{C} \\ (z, b) & \longmapsto & (z, b) & \longmapsto & \left(\dfrac{1}{z}, bz\right) & \longmapsto & \left(z, \dfrac{b}{z}\right). \end{array}$$

Thus the transition function is $f_{\infty 0}(z) = 1/z$.

- The tangent bundle. Let (U, g_U) be an atlas of of coordinate charts of X, i.e., $g_U : U \to \mathbb{C}$ is a homeomorphism and $g_V \circ g_U^{-1}$ is a holomorphic function on $g_U(U \cap V)$. The bundle defined by gluing together the $U \times \mathbb{C}$ by $f_{VU}(x) = (g_V \circ g_U^{-1})'(x)$ is the tangent bundle of X, which we write as TX.
- Its dual is the cotangent bundle, $T^\star X$ or K_X, which is also called the "canonical bundle"[3].

There is an obvious notion of isomorphism for bundles.

Group Structure. The tensor product of two holomorphic line bundles is again a holomorphic line bundle, and the same is true for the dual of a line bundle. The tensor product thus endows the set of isomorphism classes of line bundles on X with a group structure, the identity is the trivial bundle and the inverse of a bundle is its dual (see Exercises B.2, B.3 and B.4).

EXAMPLE B.2.2. On $\mathbb{P}^1$, $\mathcal{O}(1)$ denotes the dual of the tautological bundle $\mathcal{O}(-1)$, and $\mathcal{O}(n)$ is the $|n|$-th tensor power of $\mathcal{O}(1)$ (if n is positive) or of $\mathcal{O}(-1)$ (if n is negative). See Exercise B.5.

Sections. On a compact Riemann surface, the only holomorphic functions are the constants (a direct application of the maximum modulus principle). Some line bundles, though, have holomorphic *sections*, which are the closest analog to functions.

A holomorphic (resp., meromorphic) "section" of $\pi : L \to X$ is a map $s : X \to L$ such that $\pi \circ s = \mathrm{Id}_X$ such that for any local trivialization,

$$U \xrightarrow{s} \pi^{-1}(U) \xrightarrow{\varphi_U} U \times \mathbb{C},$$

the map $\varphi_U \circ s(z) = (z, s_U(z))$ being holomorphic, resp. meromorphic. In either case we also have that on $U \cap V$, $s_U = f_{UV} s_V$, where f_{UV} is the transition function mentioned above.

All line bundle have meromorphic sections. This is not trivial but follows from general results about the existence of functions (the Riemann–Roch theorem).

[3] If X is an analytic manifold of dimension n, then $T^\star X$ is a vector bundle of rank n and its maximal exterior product $\wedge^n T^\star X$ is is denoted by K_X and is called the canonical bundle. It is is also a line bundle. Note that these two objects coincide for curves ($n = 1$).

B.3. Sheaves and Cocycles

The set of germs of sections of a bundle form a sheaf[4]. I will not make a general theory of sheaves here (see [**34**]).

Examples B.3.1.

- The sheaf $\mathcal{O}_X$ of germs of holomorphic functions corresponds to the trivial bundle $X \times \mathbb{C}$.
- The sheaf Ω^1_X of germs of holomorphic 1-forms corresponds to the bundle $T^\star X$ (or K_X, since X is a curve).

Sheaves are designed to make cohomology. We will content ourselves to a description of the groups $H^0(X; \mathcal{F})$ and $H^1(X; \mathcal{F})$ when $\mathcal{F}$ is the sheaf of sections of a line bundle on X.

The vector space $H^0(X; \mathcal{F})$ is the space of global sections of the sheaf $\mathcal{F}$. For example, saying that the holomorphic functions on X are constant is equivalent to saying that $H^0(X; \mathcal{O}_X) = \mathbb{C}$, and saying that the holomorphic forms form a space of dimension g is equivalent to saying that $\dim H^0(X; \Omega^1_X) = g$.

Example B.3.2. One can check (Exercise B.6) that $H^0(\mathbb{P}^1; \mathcal{O}(n))$ is isomorphic to the vector space of polynomials of degree at most n if n is positive and is the trivial vector space otherwise.

The group $H^1(X; \mathcal{F})$ is the group of the 1-cocycles. We fix a covering of X. A 1-cocycle is a set of sections (f_{UV}) of the sheaf $\mathcal{F}$ over the intersection $U \cap V$ of any two sets of the covering, satisfying a compatibility condition (precisely called the "cocycle condition") on the intersections of three open sets of the covering. Like any cohomology group, $H^1(X; \mathcal{F})$ is a quotient which in our case is the quotient of the group of cocycles by the subgroup of coboundaries (which are the cocycles of the form $f_{UV} = f_U - f_V$, where f_U and f_V are defined on U and V).

The result of this operation *depends* on the choice of the covering. To define $H^1(X; \mathcal{F})$, one needs to take into account the choice of the covering. For a curve X covered by two affine open sets U_0 and U_∞ (the case we are interested in), $H^1(X; \mathcal{F})$ is indeed the quotient group of 1-cocycles for the covering $X = U_0 \cup U_\infty$ by the group of coboundaries.

Remark B.3.3. The same result holds for the total space of the bundle $\mathcal{O}(n)$ on $\mathbb{P}^1$ covered by the two open sets $U_0 \times \mathbb{C}$ and $U_\infty \times \mathbb{C}$; this is a result we used in Section IV.2.d.

Example B.3.4. Consider the bundle $\mathcal{O}^\star_X$ of germs of *non-vanishing* holomorphic functions on the curve X. A 1-cocycle is exactly the same thing as a family of transition functions defining a line bundle on X. Moreover, a cocycle determines the trivial bundle if and only if it is a coboundary.

The set of isomorphism classes of line bundles on X can be identified with $H^1(X; \mathcal{O}^\star_X)$. Notice that this is a group.

[4]However, there are many sheaves that do not come from line bundles, or even vector bundles.

We still have to mention the unavoidable "exponential exact sequence". At the level of sheaves, this is

$$0 \longrightarrow \mathbb{Z} \longrightarrow \mathcal{O}_X \xrightarrow{\exp} \mathcal{O}_X^\star \longrightarrow 0,$$

and at the level of cohomology groups this is

$$\longrightarrow H^1(X;\mathbb{Z}) \longrightarrow H^1(\mathcal{O}_X) \longrightarrow H^1(\mathcal{O}_X^\star) \xrightarrow{\deg} H^2(X;\mathbb{Z}) \longrightarrow,$$

with a "degree" homomorphism that will be discussed later.

Finiteness. When X is compact and $\mathcal{F}$ is the sheaf of sections of a line bundle, the spaces $H^\star(X;\mathcal{F})$ have finite dimension. This is not a trivial result; see, for example, [**72**] or [**37**].

B.4. The Picard Group, the Riemann–Roch Theorem, the Jacobian

B.4.1. Divisors and the Picard Group. To each meromorphic section of a line bundle is associated a divisor on the curve. A divisor is a formal linear combination

$$D = \sum_{i\in I} m_i P_i,$$

where I is a finite set, $m_i \in \mathbb{Z}$, and $P_i \in X$. For example, the meromorphic function f on X defines the divisor

$$(f) = \text{ zeros of } f - \text{ poles of } f$$

(counted with multiplicities). A meromorphic form similarly defines a divisor as well.

The divisors form a group $\mathrm{Div}(X)$, the free Abelian group generated by the points of X.

There is a natural homomorphism from the group $\mathrm{Div}(X)$ to $\mathbb{Z}$, the *degree*

$$\deg : \sum m_i P_i \longmapsto \sum m_i.$$

For example, functions have as many zeros as poles, so the divisors of functions have degree 0. Those of meromorphic forms have degree $2g-2$ (another consequence of the Riemann–Roch theorem).

The Picard Group. The group $\mathrm{Div}(X)$ is both huge and not very significant. To give it more meaning, we define an equivalence relation on it that takes into account the analytic structure of the surface, this is the notion of "linear equivalence"

$$D \sim D' \Leftrightarrow \text{ there exists a meromorphic function } f \text{ such that } D - D' = (f).$$

The quotient group is the *Picard group* $\mathrm{Pic}(X)$. Since the divisors of functions have degree zero, the degree homomorphism is well-defined on $\mathrm{Pic}(X)$. The component with degree d is denoted $\mathrm{Pic}^d(X)$. The component $\mathrm{Pic}^0(X)$ is a subgroup.

Divisors and Line Bundles. To a meromorphic section s of a line bundle $L \to X$, is associated a divisor (just as for a function)

$$(s) = \text{ zeros of } s - \text{ poles of } s.$$

For example, the line bundle $\mathcal{O}(n)$ on $\mathbb{P}^1$ has a section whose divisor is of degree n, justifying this notation.

If s and s' are two meromorphic sections of the same bundle L, s/s' is a meromorphic function on X. Indeed, on each open set U trivializing L, s_U/s'_U is a meromorphic function g_U and all the g_U can be glued together to form a global function since, on an intersection $U \cap V$,

$$s_U = f_{UV} s_V, \qquad s'_U = f_{UV} s'_V.$$

Hence

$$\frac{s_U}{s'_U} = \frac{s_V}{s'_V},$$

and $g_U = g_V$ on $U \cap V$.

We thus have a homomorphism from the group of isomorphism classes of line bundles to the Picard group. This homomorphism is an isomorphism. The key remark leading to this is that a divisor determines a line bundle, which we now discuss.

Line Bundles and Divisors. We begin with the case of a point P. Cover X by an open set U of a chart around P and $V = X - \{P\}$. If z is a local coordinate centered at P, then it is a holomorphic function that is non-zero on $U \cap V = U - \{P\}$. We glue $U \times \mathbb{C}$ and $V \times \mathbb{C}$ over $U \cap V$ via

$$(z, \alpha) \longmapsto (z, z\alpha).$$

It is clear that the line bundle obtained has a holomorphic section with divisor P: this is the section that is identically 1 on V and z on U. We let L_P be this line bundle.

More generally, if $D = \sum m_i P_i$ is a divisor, we define $L_D = \bigotimes L_{P_i}^{\otimes m_i}$. (Here, of course, $L^{\otimes m}$ means $(L^\star)^{\otimes(-m)}$ when $m < 0$.)

Degree. Finally, note that by the definition of the degree of a line bundle as the connecting homomorphism in the exponential exact sequence, the degree of the bundle L_P is 1, and the two notions of degree coincide. We deduce, for example, that the degree of $\mathcal{O}(n)$ is n, justifying again our notation.

B.4.2. The Riemann–Roch Theorem. If D is a divisor on the curve X with genus g, we denote by L_D the line bundle thus associated to D or the sheaf of its local sections. The Riemann-Roch theorem then asserts that

$$\dim H^0(X; L_D) - \dim H^1(X; L_D) = \deg D - g + 1.$$

REMARK B.4.1. The space $H^0(X; L_D)$ is that of global holomorphic sections of L_D. By the definition of L_D, it is the space of meromorphic sections of X with divisor satisfying

$$(f) + D \geqslant 0$$

(in the sense that all the coefficients are non-negative). Similarly, $H^1(X; L_D)$ is isomorphic to the dual of $H^0(X; \Omega^1_X \otimes L^\star_D)$ (Serre duality), so this space has the same dimension as the space of meromorphic forms η on X with divisor satisfying

$$(\eta) - D \geqslant 0$$

(see Exercise B.10). The Riemann–Roch theorem is thus a theorem about the existence of functions. See [**72**] or [**37**][5].

B.4.3. The Jacobian. To any curve of genus g is associated a complex torus of dimension g, its Jacobian, which can be described with the aid of integration of differential forms or by divisors and line bundles.

We begin with the integration of holomorphic forms. Let X be a complex curve. There is a natural map (integration)

$$\begin{aligned} H_1(X; \mathbb{Z}) &\longrightarrow H^0\left(\Omega^1_X\right)^\star \\ \gamma &\longmapsto \left(\omega \mapsto \int_\gamma \omega\right), \end{aligned}$$

which can be shown to be injective. Its image is the "lattice of periods" $\Lambda \cong \mathbb{Z}^{2g}$ in $H^0\left(\Omega^1_X\right)^\star \cong \mathbb{C}^g$ (by utilizing the Riemann–Roch theorem). The Jacobian is the quotient torus

$$\operatorname{Jac}(X) = H^0\left(\Omega^1_X\right)^\star / \Lambda,$$

which a complex torus of dimension g. The tangent space of $\operatorname{Jac}(X)$ at each point is canonically isomorphic to $H^0\left(\Omega^1_X\right)^\star$, which, in turn, is isomorphic to $H^1(\mathcal{O}_X)$, the first cohomology group of the sheaf $\mathcal{O}_X$ of germs of holomorphic functions on X. This isomorphism is a simple special case of Serre duality.

B.4.4. The Abel–Jacobi Map. The Abel–Jacobi map is simply the integration map

$$\begin{aligned} \operatorname{Div}^0(X) &\longrightarrow H^0\left(\Omega^1_X\right)^\star / \Lambda \\ \sum (P_j - Q_j) &\longmapsto \left(\omega \mapsto \int_{Q_j}^{P_j} \omega\right) \end{aligned}$$

(one can integrate on any path connecting Q_j to P_j, the result in the quotient does not depend on the choice of path). The Abel–Jacobi theorem asserts that this defines an isomorphism

$$\operatorname{Pic}^0(X) \longrightarrow \operatorname{Jac}(X).$$

If Q is a given point on X, the map

$$\begin{aligned} u : X &\longrightarrow \operatorname{Jac}(X) \\ P &\longmapsto \left(\omega \mapsto \int_Q^P \omega\right) \end{aligned}$$

[5]Some proofs seem short and elementary. This is the case for those remaining in an algebraic framework [**37**]. It is also the case for some proofs written in an analytic framework... however, one must at least use the nontrivial fact that the spaces $H^k(X; L)$ are finite dimensional.

is injective if $g \geqslant 1$, and, in particular, an isomorphism if $g = 1$. In other words, up to the choice of the point Q, the curves of genus 1 are isomorphic to their Jacobians, and thus are groups.

Note also that u induces an isomorphism

$$u_\star : H_1(X;\mathbb{Z}) \longrightarrow H_1(\mathrm{Jac}(X);\mathbb{Z})$$

by the definition of $\mathrm{Jac}(X)$.

The divisors that are linear combinations of points with positive coefficients are called "effective". Note that the effective divisors of degree d are exactly the points in the symmetric product

$$X^{(d)} = X^d/\mathfrak{s}_d.$$

The natural map

$$X^{(d)} \longrightarrow \mathrm{Pic}^d(X)$$

is surjective for $d \geqslant g$.

Exercises to Check Comprehension of the Definitions

Exercise B.1. Let P be a polynomial of degree $2n+2$ with only simple roots. Consider the affine curve C with equation $y^2 = P(x)$. Show that C is smooth.

Consider the projective completion of this curve (in $\mathbb{P}^2(\mathbb{C})$). Check that it has a unique point at infinity, which is always a singular point, and that in the neighborhood of this point, the curve consists of two smooth branches that are tangent at the point at infinity. Construct a (compact connected) Riemann surface X by adding two points to C.

Show that the map $x : C \to \mathbb{C}$ can be extended to a meromorphic function on X. Show that the genus of X is n.

Exercise B.2. If the bundles L and L' are given by the transition functions f_{UV} and f'_{UV} (for the same covering), what transition functions determine the line bundle $L \otimes L'$?

Exercise B.3. If L is determined by the transition functions f_{UV}, what are the transition functions of the dual $L^\star$?

Exercise B.4. Show that if L is a line bundle on a curve, then $L \otimes L^\star$ is canonically trivial.

Exercise B.5. Show that the transition function of the bundle $\mathcal{O}(n) \to \mathbb{P}^1$ on $U_0 \cap U_\infty$ is $f_{\infty,0}(z) = z^n$.

Exercise B.6. A section of $\mathcal{O}(n)$ is given by two holomorphic functions s_0 and s_∞ satisfying $s_0(z) = z^n s_\infty(z')$ on $\mathbb{C}^\star$.

For positive n, show that s_0 is a polynomial of degree at most n and, conversely, that such a polynomial defines a section on $\mathcal{O}(n)$.

What happens when n is negative?

EXERCISE B.7. Let P be a point in $\mathbb{P}^1$. Show that L_P is isomorphic to a bundle $\mathcal{O}(n)$ for some integer n and determine this integer. Show that the group $\operatorname{Pic}(\mathbb{P}^1)$ is isomorphic to $\mathbb{Z}$.

EXERCISE B.8. Let Λ be a lattice in $\mathbb{C}$. Show that the form dz (on $\mathbb{C}$) defines a holomorphic form on $X = \mathbb{C}/\Lambda$ with no zeros. What can be said about the cotangent bundle of X?

EXERCISE B.9 (The Cotangent Bundle of $\mathbb{P}^1$). Show that the form dz (on $\mathbb{C}$) defines a meromorphic form on $\mathbb{P}^1$ and that the cotangent bundle of $\mathbb{P}^1$ is isomorphic to the bundle[6] $\mathcal{O}(-2)$.

EXERCISE B.10. Let P be a point on the curve X so that the line bundle L_P has a section whose divisor is P. Show that $H^0(X; L_P)$ can be identified with the space of meromorphic functions on X having, at worst, a simple pole at P and no other pole. More generally, if D is a divisor, show that $H^0(X; L_D)$ is isomorphic to the vector space of meromorphic functions f satisfying

$$(f) + D \geqslant 0.$$

Fix a meromorphic form ω on X. Show that $H^0(X; \Omega^1_X \otimes L^\star_D)$ can be identified with the vector space of meromorphic functions f on X satisfying

$$(f\omega) - D \geqslant 0$$

and to the vector space of meromorphic forms η satisfying

$$(\eta) - D \geqslant 0.$$

EXERCISE B.11. Consider the Riemann surface constructed in Exercise B.1, with its meromorphic functions x and y. Determine the divisors of the functions x and y and of the form dx.

Show that the forms $x^m dx/y$ are holomorphic on X for $0 \leqslant m \leqslant n-1$. These n holomorphic forms are independent, and thus are a basis for $H^0(X; \Omega^1_X)$.

EXERCISE B.12. As in Section IV.3, consider the curve C determined by

$$\lambda^2 = \frac{Q(z)}{\prod(\alpha_j - z)}.$$

Let ∞ be its unique point above $z = \infty$, $S_1, \ldots, S_n$ the ramification points where $Q = 0$, and $A_1, \ldots, A_{n+1}$ the points (∞, α_j). Show that the divisors of λ and dz are

$$(\lambda) = S_1 + \cdots + S_n + \infty - A_1 - \cdots - A_{n+1}$$

and

$$(dz) = S_1 + \cdots + S_n + A_1 + \cdots + A_{n+1} - 3\infty.$$

Deduce from this that λdz has a double pole at ∞ and no other pole.

[6] The tangent bundle is thus $\mathcal{O}(2)$. Figure 1 in Appendix A shows its holomorphic section d/dz with its double root at infinity.

Bibliography

1. Mark Adler and Pierre van Moerbeke, *Completely integrable systems, Euclidean Lie algebras and curves,* and *Linearization of Hamiltonian systems, Jacobi varieties and representation theory*, Advances in Math. **38** (1980), 267–317 and 318–379.
2. Vladimir I. Arnold, *Mathematical methods in classical mechanics*, Springer, 1978.
3. ———, *Ordinary differential equations*, Universitext, Springer-Verlag, Berlin, 2006, Translated from the Russian by Roger Cooke, Second printing of the 1992 edition.
4. Vladimir I. Arnold and Alexander B. Givental, *Symplectic geometry*, Dynamical systems IV, Encyclopaedia of Math. Sci., Springer, 1985, pp. 1–138.
5. Michael F. Atiyah, *Convexity and commuting Hamiltonians*, Bull. London Math. Soc. **14** (1982), 1–15.
6. Michèle Audin, *Courbes algébriques et systèmes intégrables: géodésiques des quadriques*, Expositiones Math. **12** (1994), 193–226.
7. ———, *Vecteurs propres de matrices de Jacobi*, Ann. Inst. Fourier **44** (1994), 1505–1517.
8. ———, *Spinning tops, a course on integrable systems*, Cambridge University Press, 1996, Russian Transl., Regular and chaotic dynamics, Moscou, 1999, Japanese Transl., Kyoritsu, 2000.
9. ———, *Lectures on gauge theory and integrable systems*, Séminaire de Mathématiques Supérieures de Montréal, Kluwer, 1997.
10. ———, *Eigenvectors of Lax matrices, spaces of hyperelliptic curves and action coordinates for Moser systems*, Regular and chaotic dynamics **5** (2000), 67–88.
11. ———, *Hamiltonian monodromy* via *Picard-Lefschetz theory*, Commun. Math. Phys. **229** (2002), 459–489.
12. ———, *Intégrabilité et non-intégrabilité de systèmes hamiltoniens, [d'après S. Ziglin, J. Morales-Ruiz, J.-P. Ramis,...]*, Séminaire Bourbaki 2000/2001, Astérisque, vol. 282, Société Mathématique de France, Paris, 2002, pp. 113–135.
13. ———, *Exemples de hamiltoniens non intégrables en mécanique analytique réelle*, Ann. Fac. Sci. Toulouse **12** (2003), 1–23.
14. ———, *La réduction symplectique appliquée à la non-intégrabilité du problème du satellite*, Ann. Fac. Sci. Toulouse **12** (2003), 25–46.
15. ———, *Torus actions on symplectic manifolds*, Progress in Math., Birkhäuser, 2004, Revised and enlarged edition.
16. Michèle Audin and Robert Silhol, *Variétés abéliennes réelles et toupie de Kowalevski*, Compositio Math. **87** (1993), 153–229.
17. Alberto Baider, Richard C. Churchill, David L. Rod, and Michael F. Singer, *On the infinitesimal geometry of integrable systems*, Fields Institute communications, American Math. Soc. **7** (1996), 5–56.
18. Alexandre Ivanovitch Bobenko, Alexei Georgevitch Reiman, and Mikhail Arsenievitch Semenov-Tian-Shansky, *The Kowalevski top 99 years later*, Commun. Math. Phys. **122** (1989), 321–354.
19. Delphine Boucher, *Non-complete integrability of a satellite in circular orbit.*, Port. Math. **63** (2006), no. 1, 69–89.
20. Tassos Bountis, Harvey Segur, and Franco Vivaldi, *Integrable Hamiltonian systems and the Painlevé property*, Phys. Rev. A (3) **25** (1982), 1257–1264.
21. Glen E. Bredon, *Introduction to compact transformation groups*, Pure and Applied Mathematics, vol. 46, Academic Press, New York, 1972.
22. Richard Cushman and Larry Bates, *Global aspects of classical integrable systems*, Birkhäuser, 1997.

23. Nicole Desolneux-Moulis, *Dynamique des systèmes hamiltoniens complètement intégrables sur les variétés compactes*, Géométrie symplectique et mécanique (La Grande Motte, 1988), Springer, Berlin, 1990, pp. 75–83.
24. J. J. Duistermaat, *On global action-angle coordinates*, Comm. Pure Appl. Math. **33** (1980), 687–706.
25. Charles Ehresmann, *Sur les espaces fibrés différentiables*, C. R. Acad. Sc. Paris **224** (1947), 1611–1612.
26. ______, *Les connexions infinitésimales dans un espace fibré différentiable*, Colloque de topologie (espaces fibrés), Bruxelles, 1950, Georges Thone, Liège, 1951, pp. 29–55.
27. Hakan Eliasson, *Normal forms for Hamiltonian systems with Poisson commuting integrals—elliptic case*, Comment. Math. Helv. **65** (1990), no. 1, 4–35.
28. Hershel Farkas and Irwin Kra, *Riemann surfaces*, Graduate Texts in Math., vol. 71, Springer, 1980.
29. Otto Forster, *Lectures on Riemann surfaces*, Springer, New York, 1981, Translated from the German by Bruce Gilligan.
30. Sylvestre Gallot, Dominique Hulin, and Jacques Lafontaine, *Riemannian geometry*, Universitext, Springer, 1987.
31. Lubomir Gavrilov, *Bifurcations of invariant manifolds in the generalized Hénon-Heiles system*, Physica D **34** (1989), 223–239.
32. Lubomir Gavrilov and Angel Zhivkov, *The complex geometry of the Lagrange top*, Enseign. Math. (2) **44** (1998), 133–170.
33. Sophie Gérardy, *Le système de Hénon-Heiles*, Ann. Fac. Sci. Toulouse Math. **6** (1997), 633–652.
34. Roger Godement, *Topologie algébrique et théorie des faisceaux*, Hermann, Paris, 1973, Troisième édition revue et corrigée, Publications de l'Institut de Mathématique de l'Université de Strasbourg, XIII, Actualités Scientifiques et Industrielles, No. 1252.
35. William Goldman, *The symplectic nature of fundamental groups of surfaces*, Advances in Math. **54** (1984), 200–225.
36. Philip A. Griffiths, *Linearizing flows and a cohomological interpretation of Lax equations*, Amer. J. of Math. **107** (1985), 1445–1483.
37. Philip A. Griffiths and Joe Harris, *Principles of algebraic geometry*, Wiley, 1978.
38. Victor Guillemin and Shlomo Sternberg, *Convexity properties of the moment mapping*, Invent. Math. **67** (1982), 491–513.
39. Georges-Henri Halphen, *Traité des fonctions elliptiques et de leurs applications*, Gauthier-Villars, Paris, 1886.
40. David Hilbert and Stefan Cohn-Vossen, *Anschaulige geometrie*, Springer, 1932.
41. Nigel J. Hitchin, *Stable bundles and integrable systems*, Duke Math. J. **54** (1987), 91–114.
42. Nigel J. Hitchin, Graeme Segal, and Richard Ward, *Integrable systems, twistors, loop groups and Riemann surfaces*, Oxford Graduate texts in mathematics, Oxford Science Publications, 1999.
43. Édouard Husson, *Recherche des intégrales algébriques dans le mouvement d'un solide pesant autour d'un point fixe*, Ann. Fac. Sci. Toulouse **8** (1906), 73–152.
44. Hidekazu Ito, *Action-angle coordinates at singularities for analytic integrable systems*, Math. Z. **206** (1991), 363–407.
45. Carl Jacobi, *Vorlesungen über Dynamik*, Gesammelte Werke, Supplementband, Berlin, 1884.
46. Emmanuelle Julliard Tosel, *Non intégrabilité algébrique et méromorphe de problèmes de N corps*, Thèse, Université Paris 7, 1999.
47. Irving Kaplansky, *An introduction to differential algebra*, Hermann, Paris, 1976, Actualités Scientifiques et Industrielles, No. 1251, Publications de l'Institut de Mathématique de l'Université de Nancago, No. V.
48. A. A. Kirillov, *Elements of the theory of representations*, Springer-Verlag, Berlin, 1976, Translated from the Russian by Edwin Hewitt, Grundlehren der Mathematischen Wissenschaften, Band 220.
49. Felix Klein and Arnold Sommerfeld, *Theorie des Kreisels*, Teubner, Leipzig, 1897.
50. Horst Knörrer, *Geodesics on the ellipsoid*, Invent. Math. **39** (1980), 119–143.
51. Ellis Kolchin, *Differential algebra and algebraic groups*, Academic Press, New York, 1973.
52. Bertram Kostant, *Quantization and representation theory I: prequantization*, Lectures in modern analysis and applications III, Lecture Notes in Math., vol. 170, Springer, 1970.

53. Sophie Kowalevski, *Sur le problème de la rotation d'un corps solide autour d'un point fixe*, Acta Math. **12** (1889), 177–232.
54. Jacques Lafontaine, *Introduction aux variétés différentielles*, Presses universitaires de Grenoble, 1996.
55. A. H. M. Levelt, *Differential Galois theory and tensor products*, Indag. Math. (N.S.) **1** (1990), 439–449.
56. Paulette Libermann and Charles-Michel Marle, *Symplectic geometry and analytic mechanics*, Reidel, Boston, 1987.
57. Andrzej Maciejewski and Maria Przybylska, *Differential Galois approach to the Non-Integrability of the Heavy Top Problem*, Ann. Fac. Sci. Toulouse **14** (2005), 123–160.
58. Andrew Magid, *Lectures on differential Galois theory*, University lecture series, vol. 7, American Mathematical Society, 1994.
59. S. V. Manakov, *Note on the integration of Euler's equations of the dynamics of an n-dimensional rigid body*, Funct. Anal. Appl. **11** (1976), 328–329.
60. Henri Mineur, *Réduction des systèmes mécaniques à n degrés de liberté admettant n intégrales premières uniformes en involution aux systèmes à variables séparées*, Journ. de Math. **15** (1936), 385–389.
61. Juan Morales-Ruiz, *Differential Galois theory and non-integrability of Hamiltonian systems*, Progress in Math., Birkhäuser, 1999.
62. Juan Morales-Ruiz and Jean-Pierre Ramis, *Galoisian obstructions to integrability of Hamiltonian systems*, Methods Appl. Anal. **8** (2001), 33–95.
63. ———, *Galoisian obstructions to integrability of Hamiltonian systems, II*, Methods Appl. Anal. **8** (2001), 97–111.
64. Jürgen Moser, *On the volume elements on a manifold*, Trans. Amer. Math. Soc. **120** (1965), 286–294.
65. ———, *Geometry of quadrics and spectral theory*, The Chern Symposium, Springer (1980), 147–188.
66. ———, *Integrable Hamiltonian systems and spectral theory*, Accademia Nazionale dei Lincei, Rome, 1983.
67. David Mumford, *Tata lectures on theta II*, Progress in Math., Birkhäuser, 1984.
68. Georges Perec, *Espèces d'espaces*, Galilée, 1985, Edited and translated by John Sturrock, Penguin classics, 2008.
69. Henri Poincaré, *Les méthodes nouvelles de la mécanique céleste. Tome I*, Librairie Scientifique et Technique Albert Blanchard, Paris, 1987, Réimpression de l'édition originale de 1899.
70. Alexei Georgevitch Reiman, *Integrable Hamiltonian systems connected with graded Lie algebras*, J. Soviet Math. **19** (1982), 1507–1545.
71. Alexei Georgevitch Reyman and Mikhail Arsenievitch Semenov-Tian-Shanski, *Group theoretical methods in the theory of finite dimensional Integrable systems*, Dynamical systems VII, Encyclopaedia of Math. Sci., vol. 16, Springer, 1994, pp. 116–225.
72. Eric Reyssat, *Quelques aspects des surfaces de Riemann*, Progress in Math., vol. 77, Birkhäuser, 1990.
73. Mikhail Arsenievitch Semenov-Tian-Shanski, *What a classical r-matrix is*, Funct. Anal. Appl. **17** (1983), 259–272.
74. Jean-Marie Souriau, *Structure of dynamical systems*, Progress in Mathematics, vol. 149, Birkhäuser Boston Inc., Boston, MA, 1997, A symplectic view of physics, Translated from the French by C. H. Cushman-de Vries, Translation edited and with a preface by R. H. Cushman and G. M. Tuynman.
75. Michael Spivak, *Differential geometry*, Publish or perish, 1975.
76. Sergei Tabachnikov, *Billiards*, Panoramas et Synthèses, vol. 1, Société Mathématique de France, Paris, 1995.
77. Pol Vanhaecke, *Integrable systems in the realm of algebraic geometry*, second ed., Lecture Notes in Mathematics, vol. 1638, Springer-Verlag, Berlin, 2001.
78. San Vũ-Ngọc, *Sur le spectre des systèmes complètement intégrables semi-classiques avec singularités*, Thèse, Université de Grenoble, 1998.
79. Karl Weierstraß, *Über die geodätischen Linien auf dem dreiachsigen Ellipsoid*, Math. Werke I (1861), 257–266.
80. Alan Weinstein, *Lectures on symplectic manifolds*, CBMS Regional Conference Series in Mathematics, vol. 29, Amer. Math. Soc., 1977.

81. ———, *The local structure of Poisson manifolds*, Journal Diff. Geom. **18** (1983), 523–557.
82. John Williamson, *On the algebraic problem concerning the normal forms of linear dynamical systems*, Amer. J. Math. **58** (1936), 141–163.
83. ———, *An algebraic problem involving the involutory integrals of linear dynamical systems*, Amer. J. Math. **62** (1940), 881–911.
84. Sergei Lvovitch Ziglin, *Branching of solutions and non existence of first integrals in Hamiltonian mechanics I*, Funct. Anal. Appl. **16** (1982), 181–189.
85. ———, *Branching of solutions and non existence of first integrals in Hamiltonian mechanics II*, Funct. Anal. Appl. **17** (1983), 6–17.

Index