The University of Melbourne

Classical Integrable Systems and Linear Flow on Tori

Gus Schrader

Honours Thesis, November 2009

Supervisor: Dr. Omar Foda

Abstract

We study two approaches to classical integrable systems. The first considers such systems within the framework of Hamiltonian mechanics. We explain how results from symplectic and Poisson geometry can be used to obtain insight into the dynamics of Liouville integrable models. The second approach we discuss applies to systems of ordinary differential equations that can be written as Lax equations with a spectral parameter. Such equations have no *a priori* Hamiltonian content. Through the Adler-Kostant-Symes (AKS) construction, however, we can produce Hamiltonian systems on coadjoint orbits in the dual space to a Lie algebra whose equations of motion take the Lax form. We outline an algebraic-geometric interpretation of the flows of these systems, which are shown to describe linear motion on a complex torus.

Contents

Acknowledgments

I would like to extend my sincere thanks to my supervisor, Omar Foda, who devoted so much of his time to learning the material with me. I am also indebted to Michael Couch, Paul Norbury, Emma Carberry, Craig Hodgson and Peter Tingley for useful and stimulating discussions related to the subject matter of this thesis. Finally, I should like to thank my parents for all the support they provided throughout the year.

Chapter 0

Introduction

This thesis is devoted to topological and geometric aspects of the theory of classical integrable systems. As such, it is surely natural to begin by explaining the meaning of the term 'integrable system'. It turns out that this is not an entirely straightforward task: indeed, as Hitchin points out in [18], it is one to which entire books[1] have been devoted. Consequently, we merely aim to sketch a few of the various concepts of integrability, and outline some of their common features.

Historically, the definition of an integrable system can be traced back to the work of Liouville on Hamiltonian mechanics. In order to understand the definition of a Liouville integrable system, it is necessary to develop Hamiltonian mechanics on symplectic and Poisson manifolds. Chapter 1 consists of a brief overview of these ideas. Before launching into details, however, it is perhaps useful to recall some of the physical motivations behind the theory.

Consider a particle in n-dimensional Euclidean space $\mathbb{R}^n$, with position coordinates $(q_1, \ldots, q_n)$. The set of all possible positions and momenta of the particle is the **phase space** $\mathbb{R}^{2n}$. If we write p_i for the component of the particle's momentum in the ith direction, a point in phase space can be described by the coordinates $(q_1, \ldots, q_n, p_1, \ldots, p_n)$. A **Hamiltonian** $H : \mathbb{R}^{2n} \to \mathbb{R}$ is simply a smooth real-valued function on phase space. The dynamical system defined by the system of ordinary differential equations

$$\frac{dq_i}{dt} = \frac{\partial H}{\partial p_i} \qquad \frac{dp_i}{dt} = -\frac{\partial H}{\partial q^i} \tag{0.1}$$

is called the **Hamiltonian dynamical system** described by H, and the equations of motion (0.1) are known as **Hamilton's equations**.

Example 0.1 The 1-dimensional simple harmonic oscillator is an example of a Hamiltonian dynamical system. The phase space is $\mathbb{R}^2$ with coordinates (q, p), and the system's Hamiltonian is

$$H = \frac{1}{2}\left(p^2 + q^2\right)$$

The Hamiltonian dynamical system defined by H is the system of differential equations

$$\frac{dq}{dt} = p \qquad \frac{dp}{dt} = -q$$

[1]For instance, [33]

which is equivalent to the well known harmonic oscillator equation of motion

$$\frac{d^2q}{dt^2} + q = 0$$

Roughly speaking, a Liouville integrable system on a phase space $M = \mathbb{R}^{2n}$ is one which possesses n functionally independent conserved quantities: that is, n functions defined on the system's phase space whose level sets are preserved by the flow of the system. The Liouville-Arnold theorem states that under mild hypotheses, the common level set of these conserved quantities is diffeomorphic to an n-dimensional torus $T^n = \{(\varphi_1, \ldots, \varphi_n) \bmod 2\pi\}$. Additionally, the system's time evolution is linear on this torus, so that the angular coordinates φ_i satisfy

$$\frac{d\varphi_i}{dt} = \omega_i$$

for some constants $\omega_i \in \mathbb{R}$.

Example 0.2 The harmonic oscillator of Example 0.1 is a Liouville integrable Hamiltonian system: the Hamiltonian itself[2] provides the only neccesary conserved quantity. Introduce polar coordinates (r, θ) on phase space defined by

$$q = r\sin\theta, \quad p = r\cos\theta$$

In these coordinates, Hamilton's equations become

$$\frac{dr}{dt} = 0, \quad \frac{d\theta}{dt} = 1$$

Fixing a level set $H(q,p) = E$, we see that the phase space is fibered into the circles (*i.e.* 1-dimensional tori) $r = \sqrt{2E}$, and the system's time evolution is linear in the angular coordinate θ on these tori.

In Chapter 3, we study Hamiltonian systems defined on coadjoint orbits in the dual space to a Lie algebra. These coadjoint orbits turn out to be natural examples of symplectic manifolds. We outline a method, known as the Adler-Kostant-Symes (AKS) construction, which allows us to systematically construct Liouville-integrable Hamiltonian systems on coadjoint orbits. The equations of motion obtained *via* the AKS scheme take the form

$$\frac{dA}{dt} = [A, B] \tag{0.2}$$

where A and B are square matrices and $[\cdot, \cdot]$ denotes the matrix commutator. Equations of this form are known as **Lax equations**, and have been studied independently of the Hamiltonian formulation we have described[3]. Indeed, they form another possible definition of integrability: as system of ordinary differential equations is **Lax integrable** if it can be cast in the Lax form (0.2).

[2]Indeed, in Chapter 1, we shall we that the Hamiltonian function of any Hamiltonian system is a conserved quantity

[3]See [7], [31]

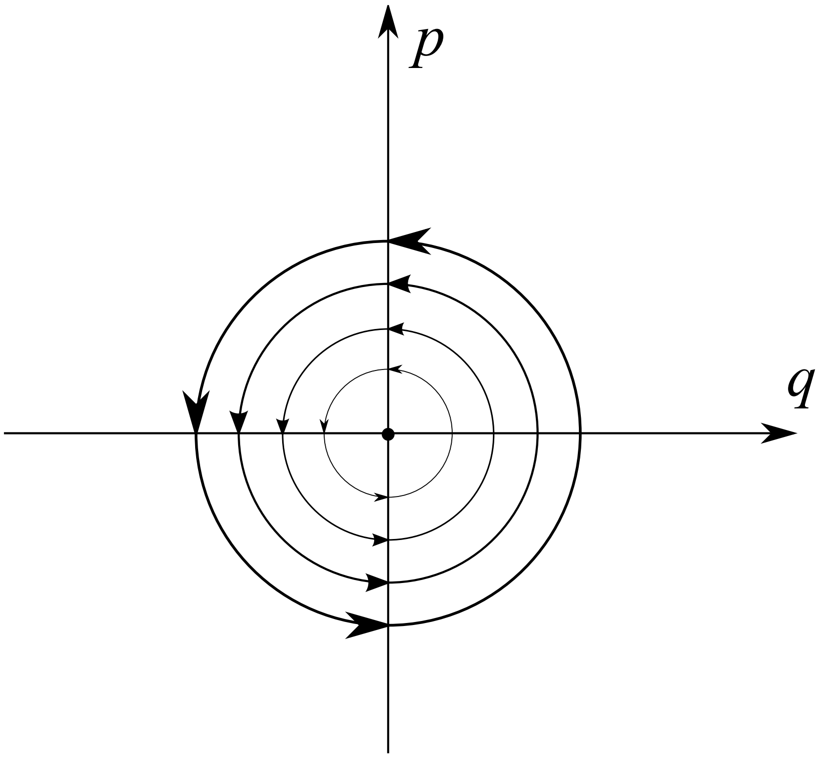

Figure 1: The foliation of the harmonic oscillator's phase space by Liouville tori

Example 0.3 The equations of motion for the simple harmonic oscillator can be written in Lax form. If we form the matrices

$$A = \begin{pmatrix} p & q \\ q & -p \end{pmatrix} \qquad B = \begin{pmatrix} 0 & -\frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix}$$

we see that the Lax equation (0.2) is equivalent to Hamilton's equations (0.1).

We should warn that this example is deceptively simple: in general, we allow the entries of the matrix B to be functions of the entries of A. Thus, Lax equations are *a priori* nonlinear.

As in the case of a Liouville integrable system, given any system of ordinary differential equations in Lax form we obtain a natural ring of functions invariant under time evolution.

Proposition 0.4 *Let $A(z;t)$ be the solution to the Lax equation (0.2). Then the functions*

$$F_p(A) = \mathrm{tr}\,(A^p)$$

are conserved quantities.

Proof. Using the differential equation (0.2) and the fact the trace is cyclic[4], we can directly compute

$$\begin{aligned}\frac{d}{dt}\mathrm{tr}\left(A^p\right) &= ptr\left([A,B]A^{p-1}\right)\\ &= \mathrm{ptr}\left(ABA^{p-1} - BA^p\right)\\ &= \mathrm{ptr}\left(BA^p - BA^p\right) = 0\end{aligned}$$

□

If $\lambda_1, \ldots, \lambda_r$ are the eigenvalues of A, then we recognise $F_p(A) = \sum_{i=1}^{p} \lambda_i^r$ as the p-th power symmetric function[5] in the λ_i. Since the power symmetric functions form a basis for the ring of symmetric functions in the λ_i, we see that any symmetric function in the eigenvalues of the matrix A is necessarily a conserved quantity of the system.

We can consider a generalization of the Lax equation (0.2) by replacing the matrices A and B by **matrix polynomials** $A(z), B(z)$ in a spectral parameter $z \in \mathbb{C}$. A matrix polynomial $A(z)$ is simply an expression of the form

$$A(z) = A_d z^d + A_{d-1}z^{d-1} + \ldots + A_0$$

where the A_i are square matrices. It is important to note that the spectral parameter z is simply a formal placeholder which we use to write the system of equations. In particular, z is not regarded as a function of time. By demanding that the equation

$$\frac{dA(z;t)}{dt} = [A(z;t), B(z;t)] \tag{0.3}$$

holds in the coefficient of each power of z, we obtain a system of matrix ordinary differential equations.

Example 0.5 Consider the following system[6] of differential equations in the three unknown functions $u_1(t), u_2(t), u_3(t)$:

$$\begin{aligned}\dot{u}_1 &= 2u_2u_3\\ \dot{u}_2 &= 2u_3u_1\\ \dot{u}_3 &= 2u_1u_2\end{aligned}$$

One checks that we can express these alternatively as the Lax equation $\frac{dA(z;t)}{dt} = [A(z;t), B(z;t)]$, where

$$A(z) = \begin{pmatrix} iu_1 & -u_2 \\ -u_2 & -iu_1 \end{pmatrix} + \begin{pmatrix} 0 & -2u_3 \\ 2u_3 & 0 \end{pmatrix} z + \begin{pmatrix} -iu_1 & -u_2 \\ -u_2 & +iu_1 \end{pmatrix} z^2$$

$$B(z) = \begin{pmatrix} 0 & 2iu_3 \\ -2iu_3 & 0 \end{pmatrix} + \begin{pmatrix} -u_1 & iu_2 \\ iu_2 & +u_1 \end{pmatrix} z$$

[4]That is, $\mathrm{tr}(AB) = \mathrm{tr}(BA)$

[5]See [23]

[6]This corresponds to a scaled version of the equations of motion for the Euler top, a well-known *Liouville* integrable system. See section 2.2.

The introduction of the spectral parameter z allows us to give Lax equations an attractive geometric interpretation, which we outline in Chapter 5. The fundamental observation in this geometric approach is that the zero set of the characteristic polynomial

$$\det(A(z) - w\mathbb{1}) = 0$$

defines an affine plane curve, which we can complete to obtain a compact Riemann surface C called the **spectral curve** of $A(z)$. The coefficients of the characteristic polynomial of $A(z)$ can be expressed as linear combinations of the traces of suitable powers of $A(z)$ and are thus conserved quantities of the system (0.3). Hence the spectral curve of $A(z;t)$ is invariant under time evolution. Next, given a matrix polynomial $A(z)$ with spectral curve C, one defines a holomorphic line bundle associated to $A(z)$ as a subbundle of $C \times \mathbb{C}^m$, whose fibre over a generic point (z, w) can be identified with the eigenspace $E_{z,w}$ of $A(z)$ corresponding to an eigenvalue w. The space of isomorphism classes of holomorphic line bundles (of a fixed degree) over C is called the Jacobian of C and has the structure of complex torus. We conclude by relating this algebraic-geometric picture to the Hamiltonian formalism of the AKS scheme: we will show that if $A(z)$ satisfies a Lax equation of the AKS form, then the image of $A(z)$ in this complex torus evolves linearly in time.

Chapter 1

Hamiltonian Mechanics

This chapter provides an introduction to Hamiltonian mechanics, which provides a natural framework to define and study classical integrable systems. Hamiltonian mechanics is a beautiful subject, with a long history and an extensive literature. As such, the treatment here is not intended to be exhaustive. An excellent starting point for further details on classical mechanics is [4], which gives a very clear and self-contained exposition of the fundamentals of symplectic geometry, emphasising the connections with physics. The lecture notes [10] are also extremely useful, and have the benefit of including a large number of interesting exercises. For technical details on Poisson manifolds, as well as a treatment of mechanics on infinite dimensional manifolds, see [24].

1.1 Symplectic Geometry

At the heart of the mathematical description of Hamiltonian mechanics is the notion of a symplectic manifold, which generalizes the concept of a phase space of a mechanical system. We shall outline the fundamentals of symplectic geometry, emphasising features of particular relevance to mechanics.

We begin by describing the concept of a symplectic vector space. Let V be a real vector space. Recall that a **bilinear form** on V is a map $\omega : V \times V \to \mathbb{R}$ satisfying

$$\omega(au + bv, w) = a\omega(u, w) + b\omega(v, w) \text{ for all } u, v, w \in V, a, b \in \mathbb{R}$$

$$\omega(u, av + bw) = a\omega(u, v) + b\omega(u, w) \text{ for all } u, v, w \in V, a, b \in \mathbb{R}$$

A bilinear form ω is **skew-symmetric** if for all $u, v \in V$, $\omega(u, v) = -\omega(v, u)$. We denote the set of all skew-symmetric bilinear forms on V by $\bigwedge^2 V^*$. A bilinear form $\omega : V \times V \to \mathbb{R}$ is said to be **non-degenerate** if $\omega(u, v) = 0$ for all $v \in V$ implies $u = 0$.

Definition 1.1 ***(Symplectic vector space)*** *A symplectic vector space is a pair (V, ω), where ω is a non-degenerate, skew-symmetric bilinear form on V.*

Example 1.2 Consider the vector space $\mathbb{R}^{2n} = \text{Span}\{e_1, \ldots, e_n, f_1, \ldots, f_n\}$ equipped with the non-degenerate, skew-symmetric bilinear form ω defined by

$$\omega(e_i, e_j) = 0 \quad \omega(f_i, f_j) = 0 \quad \omega(e_i, f_j) = \delta_{ij}$$

Then $(\mathbb{R}^{2n}, \omega)$ is a symplectic vector space.

Indeed, this example really is the *only* example of a symplectic vector space:

Proposition 1.3 ***(Symplectic basis)*** *Let V be a real vector space, and let ω be a non-degenerate, skew-symmetric bilinear form on V. Then there is a basis $\{e_1, \dots, e_n, f_1, \dots, f_n\}$ for V such that*

$$\omega(e_i, f_j) = \delta_{ij}, \quad \omega(e_i, e_j) = 0 = \omega(f_i, f_j) \tag{1.1}$$

Proof. The proof is essentially a skew-symmetric version of the Gram-Schmidt orthogonalisation procedure. We begin by choosing a non-zero vector $e_1 \in V$. By the non-degeneracy of ω, there exists f_1 such that $\omega(e_1, f_1) \neq 0$. Rescaling if necessary, we may assume that $\omega(e_1, f_1) = 1$. Now define the subspaces V_1, V_1^ω of V by

$$V_1 = \text{span}\,\{e_1, f_1\}$$

$$V_1^\omega = \{v \in V : \omega(v, w) = 0 \text{ for all } w \in V_1\}$$

Observe that $V = V_1 + V_1^\omega$. Indeed, if $v \in V$, set $a = \omega(v, e_1)$ and $b = \omega(v, f_1)$. Then $v + af_1 - be_1 \in V_1^\omega$, $be_1 - af_1 \in V_1$, and

$$v = (v + af_1 - be_1) + (be_1 - af_1)$$

Moreover, $V_1 \cap V_1^\omega = \{0\}$, since if $v = ce_1 + df_2 \in V_1 \cap V_1^\omega$, then $\omega(v, e_1) = 0$ implies $b = 0$, while $\omega(v, f_1) = 0$ implies $a = 0$. Thus $V = V_1 \oplus V_1^\omega$. Now if $V_1^\omega = \{0\}$ then the proof is complete. If not, we observe that the restriction of ω to V_1^ω is again non-degenerate. Thus we may apply the argument above to the vector space V_1^ω to obtain $V_1^\omega = V_2 \oplus V_2^\omega$. Repeating, we obtain a decomposition of V as

$$V = V_1 \oplus V_2 \oplus \ldots \oplus V_n$$

where $\dim V = 2n$. If $W_i = \text{span}\{e_i, f_i\}$, then $\{e_1, \dots, e_n, f_1, \dots f_n\}$ is a basis for V such that $\omega(e_i, f_j) = \delta_{ij}$, $\omega(e_i, e_j) = 0 = \omega(f_i, f_j)$. □

A basis of the form (1.1) is often referred to as a symplectic (or canonical) basis. Note that Proposition 1.3 also leads us to the following observation:

Corollary 1.4 *Any symplectic vector space is even dimensional*

Our next goal is to investigate the subspaces of a symplectic vector space.

Definition 1.5 ***(Subspaces of symplectic vector spaces)*** *Let (V, ω) be a symplectic vector space and let $F \subset V$ be a vector subspace. The symplectic orthogonal complement of F, denoted F^ω, is the subspace*

$$F^\omega := \{v \in V : \omega(v, f) = 0 \textit{ for all } f \in F\}$$

A subspace F is said to be isotropic if it is contained in its symplectic orthogonal complement, i.e. $F \subset F^\omega$.

Given a subspace $F \subset V$, let F^* denote the dual space of linear functionals on F. By the non-degeneracy of the form ω, the map

$$\phi_\omega : V \to F^*, \qquad v \mapsto \omega(v, -)$$

is surjective and has kernel F^ω. Hence, we find that

$$\dim F + \dim F^\omega = \dim V$$

Thus, if V is a $2n$-dimensional symplectic vector space and F is an isotropic subspace, then $\dim F \leq n$.

Definition 1.6 ***(Lagrangian subspace)*** *Let (V, ω) be a $2n$-dimensional symplectic vector space. A subspace $F \subset V$ is Lagrangian if it is isotropic and of maximal dimension n.*

The prototypical example of a Lagrangian subspace is the space spanned by the vectors $\{e_1, \dots e_n\}$ in a symplectic basis for a $2n$ dimensional symplectic vector space. We now describe an appropriate notion of isomorphism for symplectic vector spaces:

Definition 1.7 ***(Linear symplectomorphism)*** *Let (V, ω) be a symplectic vector space. A linear symplectomorphism of V is a vector space isomorphism $g : V \to V$ such that for all $u, v \in V$*

$$\omega(g(u), g(v)) = \omega(u, v)$$

Under the operation of composition, the symplectomorphisms of (V, ω) form a group. In order to describe this group, let us choose a symplectic basis $\{e_1, \dots, e_n, f_1, \dots, f_n\}$ for V. Then with respect to this basis, the form ω is represented by the matrix

$$J = \begin{pmatrix} 0 & \mathbb{1}_n \\ -\mathbb{1}_n & 0 \end{pmatrix}$$

where $\mathbb{1}_n$ is the $n \times n$ identity matrix. Then the group of symplectomorphisms of (V, ω) is the subgroup of $GL(2n, \mathbb{R})$ consisting of $2n \times 2n$ matrices g such that

$$g^T J g = J \tag{1.2}$$

This is precisely the definition of the classical Lie group $Sp(2n, \mathbb{R})$. Differentiating the relation (1.2) at the identity, we observe that the Lie algebra of $Sp(2n, \mathbb{R})$ consists of $2n \times 2n$ matrices X satisfying

$$JX + X^T J = 0$$

We now aim to generalize the definition of a symplectic vector space and decide what it might mean for a differentiable manifold to be ‘symplectic’. The basic idea is the following: if M is a differentiable manifold and ω is a smooth differential 2-form on M, then at each point x the tangent space T_xM is equipped with a skew-symmetric bilinear form ω_x. In the spirit of the definition of a symplectic vector space, we will demand that each form $\omega_x \in \bigwedge^2 T_x^*M$ be non-degenerate. As we have already seen, this requirement forces the manifold M to be even dimensional.

Definition 1.8 ***(Non-degenerate differential form)*** *A differential 2-form ω on a $2n$-dimensional smooth manifold M is non-degenerate if at each point $x \in M$, $\omega_x \in \bigwedge^2 T_x^*M$ is a nondegenerate form on the tangent space to M at x.*

Finally, we will also require that our differential 2-form be closed, i.e. $d\omega = 0$, where d denotes the exterior derivative. We summarise these considerations in the following definition:

Definition 1.9 ***(Symplectic manifold)*** *A symplectic form on a $2n$-dimensional manifold M is a closed, non-degenerate differential 2-form on M. A symplectic manifold is a pair (M, ω) where M is a smooth manifold and ω is a symplectic form*

Remark: Note that the non-degeneracy of a 2-form ω on M is equivalent to the statement that for all $x \in M$,

$$\wedge^n \omega_x \neq 0 \in \bigwedge^{2n} T_x^* M \cong \mathbb{R}$$

(i.e. the matrix representing ω has non-zero determinant). Hence if ω is non-degenerate, then the $2n$-form $\wedge^n \omega$ is a nowhere-vanishing form of top degree, or in other words, a volume form on M. In particular, we see that any manifold admitting a non-degenerate differential 2-form must be orientable.

In many of the symplectic manifolds arising from Hamiltonian mechanics, the symplectic form ω is actually exact: there exists a one form α such that $\omega = d\alpha$. If a symplectic manifold is closed, however, the situation is necessarily different:

Proposition 1.10 *If (M, ω) is a closed $2n$-dimensional symplectic manifold, then ω is not exact.*

Proof. Suppose for a contradiction that there exists a 1-form α on M such that $\omega = d\alpha$. By the non-degeneracy of ω, its top exterior power $\omega^{\wedge n}$ is a volume form on M. Now since ω is closed, we see that this volume form is also exact:

$$\omega^{\wedge n} = d(\alpha \wedge \omega^{\wedge n-1})$$

But by Stokes' theorem, the fact $\partial M = 0$ implies

$$\begin{aligned} \mathrm{vol}(M) &= \int_M \omega^{\wedge n} = \int_M d(\alpha \wedge \omega^{\wedge n-1}) \\ &= \int_{\partial M} \alpha \wedge \omega^{\wedge n-1} = 0 \end{aligned}$$

which is a contradiction. Hence ω cannot be exact. □

Note that this implies that S^{2n} cannot be a symplectic manifold if $n > 1$: for $n > 1$ the second de Rham cohomology of S^{2n} is trivial, so every closed 2-form must be exact.

Example 1.11 Consider the vector space $\mathbb{R}^{2n}$ with basis $\{e_1, \ldots, e_n, f_1, \ldots, f_n\}$, and let $\{p_1, \ldots, p_n, q_1, \ldots, q_n\} \in C^\infty(V)$ be the dual basis of linear coordinate functions on V. Then the 2-form

$$\omega = d\mathbf{q} \wedge d\mathbf{p} := \sum_{i=1}^n dq_i \wedge dp_i$$

is a symplectic form on V. In this case the symplectic form ω is not only closed, but exact: we have

$$\omega = -d\left(\sum_{i=1}^n p_i dq_i\right)$$

Example 1.12 Any 2-dimensional orientable smooth manifold is a symplectic manifold, since any volume form is symplectic: it must be closed since it is of top degree. For instance, the 2-sphere S^2 and the 2-torus T^2 both admit the structure of a symplectic manifold. In view of the Proposition 1.10, their symplectic forms cannot be exact.

The next example is of fundamental importance in Hamiltonian mechanics:

Example 1.13 (Cotangent bundles) Let X be a smooth manifold and let T^*X be its cotangent bundle, with the bundle projection map

$$\pi : T^*X \to X$$

The **tautological 1-form** α on the manifold T^*X is defined by

$$\alpha_{(x,\phi)}(v) = (\pi^*\phi)_{(x,\phi)}(v)$$

where $x \in X$ and $\phi \in T^*_x X$, so that (x, ϕ) is a point of the cotangent bundle T^*X, and $v \in T_{(x,\phi)}(T^*X)$ is a tangent vector to T^*X at (x, ϕ). It is important to remember that α is a 1-form on the manifold T^*X, not X. We can then define a 2-form ω on T^*X by $\omega = -d\alpha$. In order to understand these definitions, it is instructive to consider the expression for α in local coordinates. Let $\{U, (q_1, \ldots, q_n)\}$ be a coordinate chart on X, and let $\{T^*U, (q_1, \ldots, q_n, p_1, \ldots, p_n)\}$ be the corresponding cotangent coordinate chart: that is, the point $(q_1, \ldots, q_n, p_1, \ldots, p_n)$ corresponds to the point $\sum_{i=1}^n p_i \,(dq_i)_{\mathbf{q}}$ of the cotangent bundle. Expressing α in these local coordinates, we find

$$\alpha = \sum_{i=1}^{n} p_i dq_i$$

and hence

$$\omega = \sum_{i=1}^{n} dq_i \wedge dp_i$$

In particular, it is evident from this formula that the form ω is non-degenerate. Since ω is exact, it is certainly closed, and thus ω defines a symplectic form on T^*X. The form ω is often referred to as the **canonical symplectic form** on the cotangent bundle. Observe that the symplectic structure on the vector space $\mathbb{R}^{2n} = \mathbb{R}^n \times \mathbb{R}^n$ considered in Example 1.11 corresponds to the special case $X = \mathbb{R}^n$.

Next, we investigate some properties of smooth maps between symplectic manifolds:

Definition 1.14 ***(Symplectomorphism)*** *Let (M_1, ω_1) and (M_2, ω_2) be symplectic manifolds. A diffeomorphism $\phi : M_1 \to M_2$ is a symplectomorphism if $\phi^*\omega_2 = \omega_1$, i.e. for all $p \in M_1$ and $X, Y \in T_pM_1$,*

$$\omega_2(d\phi(X), d\phi(Y)) = \omega_1(X, Y)$$

where $d\phi : TM_1 \to TM_2$ is the derivative of ϕ.

We can also generalize the notions of isotropic and Lagrangian subspaces of symplectic vector spaces to the setting of symplectic manifolds:

Definition 1.15 ***(Lagrangian submanifold)*** *Let (M,ω) be a symplectic manifold of dimension $2n$, and let $i : X \hookrightarrow M$ be a submanifold of M. If $i^*\omega = 0$, then X is said to be an isotropic submanifold of M. If $\dim X = n$, then X is a Lagrangian submanifold of M.*

Example 1.16 Let X be an n-dimensional manifold, and let $M = T^*X$ be its cotangent bundle equipped with the canonical symplectic form. At each point $x \in X$, the cotangent space $T^*_x X$ is an n-dimensional submanifold of M. If $(q_1, \ldots, q_n, p_1, \ldots, p_n)$ is a set of cotangent coordinates on M centered at x, then

$$T^*_x X = \{m \in M : q_1(m) = q_2(m) = \cdots = q_n(m) = 0\}$$

Thus the tangent spaces to $T^*_x X$ are spanned by the tangent vectors $\left\{ \frac{\partial}{\partial p_i} \right\}_{i=1}^n$. If $i : T^*_x X \hookrightarrow M$ is inclusion,

$$(i^*\omega)\left(\frac{\partial}{\partial p_i}, \frac{\partial}{\partial p_j}\right) = \sum_{k=1}^n dq_k \wedge dp_k \left(\frac{\partial}{\partial p_i}, \frac{\partial}{\partial p_j}\right) = 0$$

and we conclude that $T^*_x X$ is isotropic. Since $\dim X = n$, $T^*_x X$ is in fact a Lagrangian submanifold of M.

Recall the very simple classification of symplectic vector spaces given in Proposition 1.3: any two symplectic vector spaces of the same dimension are symplectomorphic. Needless to say, such a simple statement cannot hold for symplectic manifolds: obstructions are given by global topological properties such as compactness and contractibility. It is a remarkable fact, however, that all symplectic manifolds are locally isomorphic to the 'standard' symplectic manifold $(\mathbb{R}^{2n}, \omega)$ described in Example 1.11. This fact is known as Darboux's theorem. The proof we present here is due to Moser (see [11]).

Theorem 1.17 ***(Moser's Trick)*** *Let M be a smooth manifold and let ω_0, ω_1 be two symplectic forms on M that are equal at the point $m \in M$. Then there exist neighbourhoods U_0, U_1 of m in M and a diffeomorphism $\psi : U_0 \to U_1$ such that $\psi(m) = m$ and $\psi^*\omega_1 = \omega_0$.*

Proof. The proof relies on an argument known as Moser's trick. For $t \in [0,1]$, consider the 2-form given by

$$\omega_t = \omega_0 + t(\omega_1 - \omega_0)$$

Since ω_1 and ω_0 are both closed, so is ω_t. Observe that since ω_0 and ω_1 agree at m, the form ω_t also agrees with both of these at m. In particular, ω_t is non-degenerate at m. Since non-degeneracy is an open condition[1], there exists some open neighbourhood of m on which ω_t is non-degenerate. Since the closed interval $[0,1]$, is compact, we can find an open neighbourhood of m on which all the ω_t are non-degenerate and thus symplectic. Now since the form $\omega_0 - \omega_1$ is closed, by the Poincare lemma there exists a 1-form α such that on a sufficiently small open neighbourhood of m, we have $d\alpha = \omega_0 - \omega_1$. Notice that by modifying α to $\alpha - df$, where f is a function defined on a neighbourhood of m such that $(df)_m = \alpha_m$, we may assume that $\alpha_m = 0$.

[1] In local coordinates it corresponds to a determinant being non-zero

The next step in the argument is the 'trick' due to Moser: by the non-degeneracy of ω_t, the equation

$$i_{X_t}\omega_t = \alpha$$

can be uniquely solved at each $t \in [0,1]$, yielding a time-dependent vector field X_t which vanishes at m for all t since $\alpha_m = 0$. This implies that the flow φ_t of X_t fixes m. Thus, if U_0 is a neighbourhood of m on which the flow φ_t is defined, $V = \varphi_t(U_0)$ also contains m. Let us now consider the Lie derivative of the form ω_t in the direction of the vector field X_t. By Cartan's magic formula and the fact that ω_t is closed, we find

$$\mathcal{L}_{X_t}\omega_t = di_{X_t}\omega_t + i_{X_t}d\omega_t = d\alpha = \omega_0 - \omega_1$$

With this in mind, we can compute

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

and conclude that the form $\varphi_t^*\omega_t$ is time-independent, so that $\varphi_t^*\omega_t = \omega_0$ for all t. Thus, taking $U_1 = \varphi_1(U_0)$ and $\psi = \varphi_1$ yields the result. □

Corollary 1.18 ***(Darboux's Theorem)*** *Let (M,ω) be a symplectic manifold and let m be a point of M. Then there exists a coordinate chart $U = (q_1,\ldots,q_n \ldots p_1 \ldots, p_n)$ centered at m such that*

$$\omega|_U = \sum_{i=1}^{n} dq_i \wedge dp_i$$

Proof. Let $\{U,(x_1,\ldots,x_n,y_1\ldots,y_n)\}$ be a coordinate chart containing m. By our classification of symplectic vector spaces, we may assume that our coordinates are chosen so that

$$\omega_m = \sum_{i=1}^{n}(dx_i)_m \wedge (dy_i)_m$$

Now consider the following two symplectic forms defined on U: the form $\omega_0 = \omega$, and the constant form $\omega_1 = \omega_m$. Since these forms agree at m, Theorem 1.17 implies that there exists a smooth map ψ between neighbourhoods of m such that $\psi^*\omega_1 = \omega_0$. Now by the naturality of the exterior derivative,

$$\omega_0 = \psi^*\omega_1 = \sum_{i=1}^{n} d(x_i \circ \phi) \wedge d(y_i \circ \phi)$$

and thus taking new coordinate functions $q_i = x_i \circ \phi, p_i = y_i \circ \phi$ we obtain the result. □

The link between symplectic manifolds and Hamiltonian dynamical systems is the concept of a Hamiltonian vector field. Let (M,ω) is a symplectic manifold and let $H : M \to \mathbb{R}$ be a smooth function on M. By the non-degeneracy of ω, the map

$$i : \mathrm{Vect}(M) \to \Omega^1(M)$$
$$X \mapsto i_X\omega$$

is a bijection: given a vector field X on M, there exists a unique 1-form α such that

$$\alpha = \omega(X, \cdot)$$

This allows us to make the following definition:

Definition 1.19 ***(Hamiltonian vector field)*** *Let (M, ω) be a symplectic manifold, and let $H : M \to \mathbb{R}$ be a smooth function on M. The Hamiltonian vector field of H, denoted X_H, is the unique vector field on M satisfying*

$$i_{X_H}\omega = dH$$

Given a smooth function H on a symplectic manifold (M, ω), a Hamiltonian vector field allows us to construct a dynamical system. We say that the Hamiltonian system on M with Hamiltonian function H is the smooth dynamical system defined by the flow φ_t of the Hamiltonian vector field X_H. The equation of motion

$$\frac{d}{dt}\varphi_t(x) = X_H(x)$$

is known as Hamilton's equation. By the basic existence, uniqueness and smoothness results from the theory of ordinary differential equations, the flow of a vector field is locally a diffeomorphism. If the manifold M is not compact, however, it is important to realise that the flow of a vector field may not exist for all time. We shall always assume that we have restricted to a sufficiently small time interval so that the flow of a particular vector field does indeed exist.

Example 1.20 Let $H : \mathbb{R}^{2n} \to \mathbb{R}$ be a smooth function on the symplectic manifold $\mathbb{R}^{2n} = \{q_1, \ldots, q_n, p_1 \ldots, p_n\}$ equipped with the canonical symplectic form $\omega = \sum_{i=1}^{n} dq_i \wedge dp_i$. Then the Hamiltonian vector field X_H of H is given by

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

Hence the equations of motion for the Hamiltonian system defined on $\mathbb{R}^{2n}$ by H take the familiar form

$$\frac{dq_i}{dt} = \frac{\partial H}{\partial p_i}, \ \frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i}$$

We now establish some basic properties of Hamiltonian vector fields and Hamiltonian dynamical systems.

Proposition 1.21 ***(Conservation of energy)*** *The Hamiltonian function H is constant on a level curve of the Hamiltonian vector field X_H.*

Proof. Let $c(t)$ be an integral curve of X_H, so that $c'(t) = X_H(c(t))$. Then by definition of X_H and the skew-symmetry of the symplectic form ω, we have

$$dH(c'(t)) = \omega\left(X_H(c(t)), c'(t)\right) = \omega\left(X_H(c(t)), X_H(c(t))\right) = 0$$

which proves the proposition. □

We shall say that the flow φ_t of a vector field X consists of symplectomorphisms if $\varphi_t^*\omega = \omega$ for all t such that φ_t is defined.

Definition 1.22 ***(Locally Hamiltonian vector field)*** *A vector field X on a symplectic manifold (M, ω) is locally Hamiltonian if the 1-form $i_X\omega$ is closed.*

The reason for this terminology is the following: if X is locally Hamiltonian, by the Poincare lemma there exists in the neighbourhood of any point a locally defined function H such that $i_X\omega = dH$.

Proposition 1.23 ***(Hamiltonian flows consist of symplectomorphisms)*** *The flow of a vector field X consists of symplectomorphisms if and only if X is locally Hamiltonian.*

Proof. The flow of X consists of symplectomorphisms if and only if $\frac{d}{dt}(\varphi_t^*\omega) = 0$. The condition that X is locally Hamiltonian is equivalent to $\mathcal{L}_X\omega = 0$, since by Cartan's magic formula and the fact that ω is closed,

$$\mathcal{L}_X(\omega) = i_X d\omega + d i_X\omega = d i_X\omega$$

But then since

$$\frac{d}{dt}(\varphi_t^*\omega) = \varphi_t^*\mathcal{L}_{X_H}\omega$$

the result follows. □

We can give a characterisation of symplectomorphism based on Hamiltonian vector fields as follows:

Proposition 1.24 *Let (M_1, ω_1), (M_2, ω_2) be symplectic manifolds, and let $\phi : M_1 \to M_2$ be a diffeomorphism. Then ϕ is a symplectomorphism if and only if for all $H \in C^\infty(M_2)$ and $m \in M_1$*

$$(d\phi)_m \cdot X_{H\circ\phi}(m) = X_H(\phi(m)) \tag{1.3}$$

Proof. Suppose firstly that ϕ is a symplectomorphism. Then for all $v \in T_mM_1$, we have

$$\omega_2(d\phi_m \cdot X_{H\circ\phi}(m), v) = \omega_1\left(X_{H\circ\phi}(m), (d\phi_m)^{-1}(v)\right)$$

However, by definition of a Hamiltonian vector field and the chain rule

$$\begin{aligned}\omega_1\left(X_{H\circ\phi}(m), (d\phi_m)^{-1}(v)\right) &= d(H\circ\phi)_m \cdot (d\phi_m)^{-1}(v) \\ &= dH_{\phi(m)} \cdot d\phi_m \cdot (d\phi_m)^{-1}(v) \\ &= \omega_2(X_H(\phi(m)), v))\end{aligned}$$

So for all $v \in T_m M_1$, $\omega_2\left(X_H\left(\phi(m)\right), v)\right) = \omega_2\left((d\varphi)_m \cdot X_{H\circ\phi}(m), v\right)$. Thus by non-degeneracy of ω_2 we conclude $(d\phi)_m \cdot X_{H\circ\phi}(m) = X_H\left(\phi(m)\right)$.

Conversely, suppose φ satisfies 1.4. We show that for all $v, w \in T_m M_1$,

$$\omega_1(v, w) = \omega_2(d\phi_m \cdot v, d\phi_m \cdot w)$$

Choosing a smooth function $H \in C^\infty(M_2)$ such that $X_{H\circ\varphi}(m) = v$, we have

$$\begin{aligned}\omega_1(v, w) &= \omega_1(X_{H\circ\varphi}(m), v)\\ &= \omega_2\left((X_H(\varphi(m)), d\varphi_m \cdot w\right)\\ &= \omega_2\left(d\varphi_m \cdot v, d\varphi_m \cdot w\right)\end{aligned}$$

which shows that φ is a symplectomorphism. □

We shall now define the Poisson bracket on a symplectic manifold, which will enable us to give a particularly elegant description of Hamiltonian mechanical systems.

Definition 1.25 ***(Poisson bracket on a symplectic manifold)*** *Let (M, ω) be a symplectic manifold, and let $F, G : M \to \mathbb{R}$ be two smooth functions on M. The Poisson bracket of F and G is the smooth function on M defined by*

$$\{F, G\}(m) := \omega\left(X_F(m), X_G(m)\right)$$

Example 1.26 If our symplectic manifold is $\mathbb{R}^{2n}$ equipped with the canonical symplectic form, we recover the familiar form of the Poisson bracket from classical mechanics:

$$\{F, G\} = \sum_{i=1}^{n} \frac{\partial F}{\partial q_i}\frac{\partial G}{\partial p_i} - \frac{\partial F}{\partial p_i}\frac{\partial G}{\partial q_i}$$

Poisson brackets enable us to give another useful characterisation of symplectomorphism:

Proposition 1.27 ***(Symplectomorphisms preserve Poisson brackets)*** *Let (M_1, ω_1), (M_2, ω_2) be symplectic manifolds. A diffeomorphism $\varphi : M_1 \to M_2$ is a symplectomorphism if and only if it preserves Poisson brackets, that is,*

$$\{F, G\} \circ \varphi = \{F \circ \varphi, G \circ \varphi\}$$

for all $F, G \in C^\infty(M_2)$.

Proof. This is essentially a restatement of Proposition 1.24. If $F, G \in C^\infty(M_2)$,

$$\{F, G\} \circ \varphi(m) = d(F \circ \varphi)_m \cdot (d\varphi^{-1})_{\varphi(m)} \cdot X_G(\varphi(m))$$

while

$$\{F \circ \varphi, G \circ \varphi\}(m) = d(F \circ \varphi)_m \cdot X_{G\circ\varphi}(m)$$

Applying Proposition 1.24, the result follows. □

We can also write down an important relation between the Poisson bracket of functions and the Lie bracket of their Hamiltonian vector fields:

Proposition 1.28 *For all F, G, we have*

$$X_{\{F,G\}} = -[X_F, X_G]$$

Proof. To avoid a proliferation of subscripts, we set $X = X_F, Y = X_G$. Now using the identity

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

together with Cartan's magic formula and the fact that ω is closed, we compute

$$\begin{aligned} i_{[X,Y]}\omega = \mathcal{L}_X i_Y \omega - i_Y \mathcal{L}_X \omega &= d i_X i_Y \omega + i_X d i_Y \omega - i_Y d i_X \omega - i_Y i_X d\omega \\ &= d i_X i_Y \omega + i_X d(dF) - i_Y d(dG) \\ &= d\left(\omega(Y, X)\right) \\ &= d\left(\{g, f\}\right) \end{aligned}$$

which yields the result. □

Together with the Jacobi identity for the Lie bracket of vector fields, Proposition 1.28 immediately yields

Corollary 1.29 *The Poisson bracket satisfies the Jacobi identity, that is,*

$$\{F, \{G, H\}\} + \{H, \{F, G\}\} + \{G, \{H, F\}\} = 0$$

Thus the Poisson bracket makes $C^\infty(M)$ into an (infinite dimensional) Lie algebra[2].

Finally, Poisson brackets enable us to elegantly describe the time evolution of an observable $F \in C^\infty(M)$:

Proposition 1.30 *Let H be a Hamiltonian function, X_H the corresponding Hamiltonian vector field, and φ_t the flow of X_H. Then*

$$\frac{d}{dt}(F \circ \varphi_t) = \{F \circ \varphi_t, H\} = \{F, H\} \circ \varphi_t$$

Proof. By the chain rule,

$$\frac{d}{dt}(F \circ \varphi_t)(m) = dF_{\varphi_t(m)} \cdot X_H\left(\varphi_t(m)\right) = \{F, H\}(\varphi_t(m))$$

Now since φ_t is a symplectomorphism it preserves Poisson brackets, and we have

$$\{F, H\}(\varphi_t(m)) = \{F \circ \varphi_t(m), H \circ \varphi_t(m)\} = \{F \circ \varphi_t, H\}(m)$$

since $H \circ \varphi_t = H$ by the conservation of energy lemma. □

[2]Moreover, Proposition 1.28 states that the mapping $C^\infty(M) \to \mathrm{Vect}(M), H \mapsto X_H$ is a Lie algebra anti-homomorphism.

1.2 Group Actions

The notion of a symmetry of a physical system plays an important role in most formulations of classical mechanics. In the context of Hamiltonian mechanics, this idea is formalised by considering the actions of Lie groups on symplectic manifolds.

Definition 1.31 ***(Lie group)*** *A Lie group is a group G which is also a smooth manifold on which inversion $G \to G$ and group multiplication $G \times G \to G$ are smooth maps.*

Perhaps the most basic example of example of a Lie group is the group $GL_n(\mathbb{R})$ of invertible $n \times n$ real matrices. Since $GL_n(\mathbb{R})$ is an open subset of the vector space of all $n \times n$ real matrices, it is evidently a smooth manifold of dimension n^2 and the group multiplication and inversion are readily seen to be smooth maps. One can show (see [1], page 253) that if G is a Lie group and H is a subgroup which is a closed subset of G, then H is a submanifold of G and hence a Lie group in its own right. Thus closed subgroups of $Gl_n(\mathbb{R})$ such as SL_n, O_n, SO_n, Sp_{2n} also define Lie groups. Such groups are referred to as **matrix Lie groups**. We now introduce the notion of the action of a Lie group on a smooth manifold.

Definition 1.32 ***(Group action)*** *Let G be a Lie group with identity element e and let M be a smooth manifold. A (left) action of G on M is a smooth map $\Phi : G \times M \to M$ such that*

$$\Phi(e, x) = x \text{ for all } x \in M; \text{ and}$$

$$\Phi(g, \Phi(h, x)) = \Phi(gh, x) \text{ for all } g, h \in G \text{ and } x \in M$$

When describing the action of a group, we shall interchangeably use the notations $\Phi(g, x) = \Phi_g(x) = g \cdot x$.

To illustrate the definition of a group action, we note that any Lie group G acts on itself by left-multiplication:

$$g \mapsto L_g : G \to G$$

$$L_g(h) = gh$$

A Lie group G can also act on itself by conjugation, or inner automorphisms:

$$g \mapsto I_g : G \to G$$

$$I_g(h) = ghg^{-1}$$

Before proceeding further, we must make some general remarks on group actions. An action of G on M is said to be

1. **Free** if $\Phi_g(x) = x$ implies $x = e$
2. **Transitive** if for any $m, n \in M$, there exists $g \in G$ such that $\Phi(g, m) = n$
3. **Proper** if the mapping $\tilde{\Phi} : G \times M \to M \times M$ defined by $\tilde{\Phi}(g, m) = (m, \Phi(g, m))$ is a proper map[3].

[3] Recall that a map is proper if the preimage of a compact set is compact

Let Φ_g be an action of a Lie group G on a manifold M. For $m \in M$, the **stabilizer** (or **isotropy subgroup**) of m is the subgroup $G_m \subset G$ defined by

$$G_m := \{g \in G | \Phi_g(m) = m\}$$

By the continuity of the map $\Phi_m : G \to M$, $g \mapsto \Phi(g, m)$, the stabilizer $G_m = \Phi_m^{-1}(m)$ is closed in G and is thus a submanifold of G. If $m \in M$, the **orbit** through m under the action of G is the subset $\Omega_m \subset M$ defined by

$$\Omega_m = \{\Phi(g, m) | g \in G\}$$

The relation of belonging to the same orbit defines an equivalence relation on M, and we call the set of equivalence classes the orbit space of M under G, denoted M/G. An important special case is when M is a Lie group K and G is a closed subgroup of K, with the action of G on K given by left-multiplication. This action is free and proper[4], and the orbit space K/G is the space of left cosets of G in K. It turns out that the orbit space of M under a free and proper Lie group action can be given a natural smooth manifold structure:

Proposition 1.33 *If $\Phi : G \times M \to M$ is a proper and free action, then the orbit space M/G can be given the structure of a smooth manifold of dimension* $\dim M - \dim G$ *such that the orbit projection $\pi : M \to M/G$ is a smooth submersion.*

For a proof, the reader is referred to page 153 of [22]

This result implies that orbits in M under the action of G can also be given a natural smooth manifold structure. This structure is defined by the requirement that the natural bijection of the coset space G/G_m with the orbit Ω_m be a diffeomorphism, where the smooth structure on G/G_m comes from Proposition 1.33. It can be shown (see [1], p.265) that an orbit $\Omega_m \hookrightarrow M$ is an injectively immersed submanifold of M. Thus, we shall always regard the orbits Ω_m as injectively immersed submanifolds of M diffeomorphic to G/G_m.

If the Lie group G is compact, then one can show (again, see [1]) that the orbits of its action are closed embedded submanifolds of M.

For $g \in G$, let $L_g : G \to G$ denote left multiplication by g. We say that a vector field $X \in \mathrm{Vect}(G)$ is **left-invariant** if for all $g, h \in G$ we have

$$d\left(L_g\right)_h \cdot X(h) = X(gh)$$

Observe that since the action of G on itself by left-multiplication is transitive, there exists a unique left-invariant vector field X_ξ on G such that $X_\xi(e) = \xi \in T_eG$. Thus, the space of left-invariant vector fields on G is finite-dimensional, and can be identified with T_eG, the tangent space to G at the identity.

It can be shown (see Lemma 15.15 of [22]) that any left-invariant vector fields X_ξ is complete, so that there exists a unique integral curve $\gamma_\xi(t) : \mathbb{R} \to G$ defined for all t such that $\gamma_\xi(0) = e$ and $\gamma_\xi'(t) = X_\xi(\gamma_\xi(t))$. Moreover, we have[5] that

$$\gamma_\xi(s + t) = \gamma_\xi(s)\gamma_\xi(t)$$

[4] See [22] p. 160, Theorem 7.15

[5] Indeed, considered as functions of t, the left invariance of X_ξ implies that both sides of the equation above are solutions of the ODE $\rho'(t) = X_{\rho(t)}$ which agree at $t = 0$, and thus they must agree for all time.

where the product on the right-hand-side is group multiplication in G). In view of this fact, we say that the integral curve $\gamma_\xi(t)$ defines a **1-parameter subgroup** of G.

Definition 1.34 ***(Exponential map)*** *The exponential map* $\exp : T_eG \to G$ *is defined by*

$$\exp(\xi) = \gamma_\xi(1)$$

One can check (see [22] p.385) that the exponential map is smooth, a local diffeomorphism, and satisfies $\exp(t\xi) = \gamma_\xi(t)$. If G is a matrix Lie group, so that its Lie algebra $\mathfrak{g}$ is identified with a subspace of the space of $n \times n$ matrices, one finds that the exponential map defined above agrees with the familiar notion of the matrix exponential. We also note that by definition of the exponential map, the flow of the left-invariant vector field X_ξ is given by $\varphi_t(g) = R_{\exp t\xi} g$, where $R_h : G \to G$ denotes right multiplication by h.

We now introduce the concept of a Lie algebra, which plays a central role in the theory of integrable systems.

Definition 1.35 ***(Lie algebra)*** *A Lie algebra is a vector space* $\mathfrak{g}$ *together with a bilinear, skew-symmetric bracket*

$$[\cdot, \cdot] : \mathfrak{g} \to \mathfrak{g}$$

(called the Lie bracket) which satisfies the Jacobi identity

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

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

Example 1.36 The vector space $\mathbb{R}^3$ with the Lie bracket given by the cross product of vectors is a Lie algebra.

Given a Lie group G, we can equip the vector space $\mathfrak{g} = T_eG$ with the structure of a Lie algebra in the following fashion. Recall that $\mathfrak{g}$ can be identified with the space of left-invariant vector fields on G via

$$X_\xi \longmapsto X_\xi(e)$$

Since the Lie-bracket of left-invariant vector fields is again left-invariant, and the Lie bracket of vector fields satisfies the Jacobi identity, it follows that the Lie bracket of vector fields gives $\mathfrak{g}$ the structure of a Lie algebra. If $\xi, \eta \in T_eG$, we have

$$[\xi, \eta] = [X_\xi, X_\eta](e)$$

Let us now return to the conjugation action of G on itself by the inner automorphisms I_g. Since the inner automorphisms fix the identity element $e \in G$, differentiating at e yields an action of G on the tangent space T_eG. Thus we obtain an action of G on its Lie algebra $\mathfrak{g}$ given by

$$g \mapsto Ad_g : \mathfrak{g} \to \mathfrak{g}$$

where $Ad_g(\xi) = d(I_g)_e(\xi)$. This action is called the **adjoint action** of G on $\mathfrak{g}$.

Definition 1.37 ***(Infinitesimal generator)****Let $\Phi : G \times M \to M$ be an action of a Lie group G on a smooth manifold M, and let $\xi \in \mathfrak{g}$. Then the infinitesimal generator of the action corresponding to ξ is the vector field ξ_M on M defined by*

$$\xi_m(x) = \left.\frac{d}{dt}\Phi\left(\exp(t\xi), x\right)\right|_{t=0}$$

To illustrate this concept, let us compute the infinitesimal generators for the adjoint action of G on $\mathfrak{g}$. The standard notation for the infinitesimal generator of the adjoint action corresponding to an element $\xi \in \mathfrak{g}$ is ad_ξ. Now by the definition above, we have

$$\begin{aligned}
\mathrm{ad}_\xi &= \left.\frac{d}{dt}Ad_{\exp t\xi}(\eta)\right|_{t=0} \\
&= \left.\frac{d}{dt}d\left(L_{\exp t\xi} \circ R_{\exp -t\xi}\right)_e(\eta)\right|_{t=0}
\end{aligned}$$

Since the flow of the left invariant vector field X_ξ is given by $\varphi_t(g) = R_{\exp t\xi}g$, by the dynamic definition of the Lie bracket of vector fields we can write

$$\begin{aligned}
\left.\frac{d}{dt}d\left(R_{\exp -t\xi} \circ L_{\exp t\xi}\right)_e(\eta)\right|_{t=0} &= \left.\frac{d}{dt}d(\varphi_t^{-1})_{\varphi_t(e)} \cdot X_\eta(\varphi_t(e))\right|_{t=0} \\
&= [X_\xi, X_\eta](e) \\
&= [\xi, \eta]
\end{aligned}$$

and thus we see that $\mathrm{ad}_\xi(\eta) = [\xi, \eta]$.

Let us unravel these definitions in the familiar context of matrix Lie groups. Let $G \subset GL_n(\mathbb{R})$ be a matrix Lie group, and let $\mathfrak{g}$ be its tangent space at the identity, which is identified with a subspace of the space of $n \times n$ matrices. In this case, the exponential map is given by

$$\exp(\xi) = \sum_{k=0}^{\infty} \frac{\xi^k}{k!}$$

We see that the adjoint action of G on $\mathfrak{g}$ is given by

$$\mathrm{Ad}_A(\xi) = \left.\frac{d}{dt}A\exp(t\xi)A^{-1}\right|_{t=0} = A\xi A^{-1}$$

The infinitesimal generator ad_η is also easily computed:

$$\begin{aligned}
\mathrm{ad}_\eta(\xi) &= \left.\frac{d}{dt}\mathrm{Ad}_{\exp t\eta}(\xi)\right|_{t=0} \\
&= \left.\frac{d}{dt}\exp(t\eta)\xi\exp(-t\eta)\right|_{t=0} \\
&= \eta\xi - \xi\eta
\end{aligned}$$

and we see that the Lie bracket on $\mathfrak{g}$ is given by the commutator of matrices.

Returning to the general setting, let $\mathfrak{g}$ be the Lie algebra of a Lie group G and let $\mathfrak{g}^*$ be the dual vector space of linear functionals on $\mathfrak{g}$. The adjoint action of G on $\mathfrak{g}$ naturally induces an action of G on $\mathfrak{g}^*$ called the **coadjoint action**, which is defined by

$$\langle Ad_g^* \alpha, X \rangle = \langle \alpha, Ad_{g^{-1}} X \rangle$$

where $\langle \cdot, \cdot \rangle$ denotes the natural pairing between vectors and linear functionals.

The infinitesimal generator of the coadjoint action corresponding to $\xi \in \mathfrak{g}$, denoted ad_ξ^*, satisfies

$$\langle ad_\xi^*(\alpha), \eta \rangle = \langle \alpha, \mathrm{ad}_{-\xi}(\eta) \rangle = -\langle \alpha, [\xi, \eta] \rangle$$

Suppose that there exists a symmetric, non-degenerate, bilinear form $\langle \ , \ \rangle$ on $\mathfrak{g}$ such that for all $X, Y \in \mathfrak{g}$ and for all $g \in G$,

$$\langle Ad_g X, Ad_g Y \rangle = \langle X, Y \rangle \tag{1.4}$$

A form satisfying this relation above is said to be **adjoint-invariant**[6]. Differentiating the condition (1.4) yields the following equality:

$$\langle X, \mathrm{ad_Y}(Z) \rangle + \langle Y, \mathrm{ad_Z}(X) \rangle = 0, \qquad X, Y, Z \in \mathfrak{g}$$

We will call a form satisfying this differentiated relation ad^*-invariant. An adjoint-invariant form establishes an isomorphism $\mathfrak{g} \cong \mathfrak{g}^*$ under which the adjoint action of G on $\mathfrak{g}$ is identified with its coadjoint action on $\mathfrak{g}^*$: if $X \in \mathfrak{g}$, this identification is given by

$$X \longmapsto \langle X, \cdot \ \rangle, \qquad \mathrm{Ad}_g(X) \longmapsto \mathrm{Ad}^*_{g^{-1}}(\langle X, \cdot \ \rangle)$$

Under this isomorphism, we identify a function H on $\mathfrak{g}^*$ with the function f on $\mathfrak{g}$ defined by

$$f(X) = H(\langle X, \cdot \ \rangle)$$

The form $\langle \cdot, \cdot \rangle$ also allows us to define gradients of functions on $\mathfrak{g}$, which will be used in Chapter 3 to write Hamilton's equations in Lax form.

Definition 1.38 ***(Gradient of a function)*** *Let V be a vector space with a symmetric bilinear form $\langle \cdot, \cdot \rangle$ and let f be a smooth function on V. Making the usual identification $T_v V \sim V$, the gradient of f is the unique vector field ∇f on V such that for all $X \in V$,*

$$df(X) = \langle \nabla f, X \rangle$$

Let us now focus on the coadjoint action of a Lie group G. It turns out that the orbits in $\mathfrak{g}^*$ under the coadjoint action of G are examples of symplectic manifolds. In Chapter 3, we will construct integrable Hamiltonian systems on these manifolds whose equations of motion take the Lax form.

Definition 1.39 ***(Coadjoint orbit)*** *Let G be a Lie group with Lie algebra $\mathfrak{g}$ and denote by $\mathfrak{g}^*$ the vector space dual to $\mathfrak{g}$. Write $\mathrm{Ad}_g^* : \mathfrak{g}^* \to \mathfrak{g}^*$ for the coadjoint action of G on $\mathfrak{g}^*$. Then if $\alpha \in \mathfrak{g}^*$, the orbit*

$$\Omega_\alpha = \left\{ Ad_g^* \cdot \alpha | g \in G \right\}$$

is said to be the coadjoint orbit through α.

[6] If $\mathfrak{g}$ is semisimple, the Killing form is an example of such a form

Example 1.40 We consider the example of the matrix Lie group $G = SO(3)$, the group of rotations of 3-dimensional Euclidean space:

$$SO(3) = \{g \in M_3(\mathbb{R}) | \det(g) = 1, g^T g = e\}$$

We let the reader verify that the $SO(3)$ is a submanifold of $M_3(\mathbb{R})$, and that its Lie algebra $T_e SO(3) = \mathfrak{so}(3)$ is the vector space of skew-symmetric 3×3 matrices

$$\mathfrak{so}(3) = \{X \in M_3(\mathbb{R}) | X + X^T = 0\}$$

with Lie bracket given by the commutator of matrices. Consider the vector space isomorphism

$$\phi : \mathbb{R}^3 \to \mathfrak{so}(3)$$

$$(x, y, z) \mapsto \begin{pmatrix} 0 & -z & y \\ z & 0 & -x \\ -y & x & 0 \end{pmatrix}$$

If make $\mathbb{R}^3$ into a Lie algebra with Lie bracket given by the cross product, a simple calculation shows that ϕ is a **Lie algebra homomorphism**, that is,

$$\phi\left([v, w]_{\mathbb{R}^3}\right) = \phi(v \times w) = \phi(v)\phi(w) - \phi(w)\phi(v) = [\phi(v), \phi(w)]_{\mathfrak{so}(3)}$$

and thus we may identify $\mathfrak{so}(3)$ with the Lie algebra $(\mathbb{R}^3, \times)$. Recall that if $\xi \in \mathfrak{so}(3)$, the adjoint action of $SO(3)$ is given by $\mathrm{Ad}_g(\xi) = g\xi g^{-1}$. One checks that for all matrices $A \in SO(3)$

$$\phi(A \cdot v) = A\phi(v)A^{-1}$$

and[7] so we see that the adjoint action of $SO(3)$ is identified with the action of $SO(3)$ on $\mathbb{R}^3$ by matrix multiplication. Now the Euclidean dot product on $\mathbb{R}^3$ defines a non-degenerate symmetric bilinear form on $\mathbb{R}^3 \sim \mathfrak{so}(3)$. By the condition $A^T A = e$ in the definition of the group $SO(3)$, this form is adjoint invariant. Thus, we may use the dot product to identify the coadjoint and adjoint actions of $SO(3)$. If $w \in \mathbb{R}^3 \cong (\mathbb{R}^3)^* \cong \mathfrak{so}(3)^*$, we have

$$\mathrm{Ad}^*_A(w) = A^{-1} \cdot w$$

and thus the coadjoint orbit through w is

$$\Omega_w = \{\mathrm{Ad}^*_A(w) | A \in SO(3)\} = \{Aw | A \in SO(3)\}$$

So if $w \neq 0$, the coadjoint orbit through w is the sphere of radius $\|w\|$, while the coadjoint orbit through 0 consists of a single point.

As we have mentioned, the reason for our interest in coadjoint orbits is that they are natural examples of symplectic manifolds. The tangent space to a coadjoint orbit Ω through $\alpha \in \mathfrak{g}^*$ is easily described. Recall that by Proposition 1.33, the natural projection map

$$\varphi_\alpha : G \to \Omega$$

$$g \mapsto \mathrm{Ad}^*_g \cdot \alpha$$

is a smooth submersion. Thus its derivative at the identity

$$d(\varphi_\alpha)_e : \mathfrak{g} \to T_\alpha \mathfrak{g}^*$$

[7] Where $A \cdot v$ denotes the matrix A multiplying the vector v

$$X \mapsto \mathrm{ad}_X^*(\alpha)$$

establishes an isomorphism

$$\mathfrak{g}/\ker\left(d(\varphi_\alpha)_e\right) \to T_\alpha\Omega$$

Now for any $\alpha \in \mathfrak{g}^*$, we can define a skew-symmetric bilinear form on the Lie algebra $\mathfrak{g}$ by

$$\omega_\alpha(X, Y) = \langle \alpha, [X, Y] \rangle = -\langle \mathrm{ad}_X^*(\alpha), Y \rangle$$

where once again the angle brackets denote the natural pairing between functionals and vectors. Observe that this form may be degenerate if there exists a non-zero vector $X \in \mathfrak{g}$ is such that $\mathrm{ad}_X^* \alpha = 0$. However ω_α does define a non-degenerate form on the quotient space $\mathfrak{g}/\ker\left(d(\varphi_\alpha)_e\right)$. Thus, we can define a non-degenerate 2-form ω on the coadjoint orbit Ω by

$$\omega(\mathrm{ad}_X^*(\alpha), \mathrm{ad}_Y^*(\alpha)) = \omega_\alpha(X, Y) = \langle \alpha, [X, Y] \rangle$$

Proposition 1.41 *The 2-form ω is closed, and thus defines a symplectic form on the coadjoint orbit Ω.*

Proof. We claim that for any $X, Y, Z \in \mathfrak{g}$ and for any $\alpha \in \mathfrak{g}^*$, we have

$$d\omega(\mathrm{ad}_X^*(\alpha), \mathrm{ad}_Y^*(\alpha), \mathrm{ad}_Z^*(\alpha)) = 0$$

For each $X \in \mathfrak{g}$, define a function H_X on Ω by $H_X(\alpha) = \langle \alpha, X \rangle$. Note that for all tangent vector fields ad_Y^* we have

$$\omega(\mathrm{ad}_Y^*(\alpha), \mathrm{ad}_X^*(\alpha)) = \langle \alpha, [Y, X] \rangle = -\langle \mathrm{ad}_Y^*(\alpha), X \rangle = H_X(\mathrm{ad}_Y^*(\alpha))$$

where as usual, we are making the identification $T_\alpha(\mathfrak{g}^*) \cong \mathfrak{g}^*$. Since the function H_X is linear, we recognise $H_X(\mathrm{ad}_Y^*(\alpha))$ as the derivative of H_X in the direction $\mathrm{ad}_Y^*(\alpha)$. Thus for all tangent vector fields ad_Y^*, we have

$$\omega(\mathrm{ad}_Y^*, \mathrm{ad}_X^*) = -dH_X(\mathrm{ad}_Y^*)$$

and thus $i_{\mathrm{ad}_X^*}\omega = dH_X$. Observe that by construction, ω is G-invariant: that is, for all $g \in G$,

$$\left(\mathrm{Ad}_g^*\right)^* \omega = \omega$$

Thus the Lie derivative of ω satisfies $\mathcal{L}_{\mathrm{ad}_Y^*}\omega = 0$ for all $Y \in \mathfrak{g}$. Now, by Cartan's magic formula, we have

$$0 = \mathcal{L}_{\mathrm{ad}_X^*}\omega = d(i_{\mathrm{ad}_X^*}\omega) + i_{\mathrm{ad}_X^*}d\omega = -d(dH_X) + i_{\mathrm{ad}_X^*}d\omega = i_{\mathrm{ad}_X^*}d\omega$$

Hence we conclude that $i_{\mathrm{ad}_X^*}d\omega = 0$ for all $X \in \mathfrak{g}$, and since the vector fields of the form ad_X^* span the tangent spaces to Ω, this implies $d\omega = 0$ as claimed.

□

Corollary 1.42 *All coadjoint orbits are even dimensional*

Example 1.43 Recall that the generic orbits of the coadjoint action of $SO(3)$ on $\mathfrak{so}(3)^* \cong \mathbb{R}^3$ are spheres centered at the origin. Let $\mathbf{\Pi}$ be such a sphere, and let $v \in \mathbf{\Pi}$. Then if $X, Y \in T_v\mathbf{\Pi}$, the coadjoint orbit symplectic form on $\mathbf{\Pi}$ is given by

$$\omega(X, Y) = \langle v, X \times Y \rangle$$

where $\times$ is the cross product of vectors, which we recognise as the standard Euclidean area form on the sphere.

1.3 Poisson Manifolds

In order to interpret Lax equations as Hamiltonian dynamical systems, it will be necessary to describe Hamiltonian mechanics on a broader class of spaces than symplectic manifolds. This is accomplished by generalising the structure of the symplectic Poisson bracket as follows.

Definition 1.44 ***(Poisson manifold)*** *A Poisson manifold* $(P, \{,\})$ *is a manifold* P *together with a bilinear Poisson bracket* $\{\cdot,\cdot\} : C^\infty(P) \times C^\infty(P) \to C^\infty(P)$ *satisfying*

P1: $\{\cdot,\cdot\}$ *is a derivation: for all* $f, g, h \in C^\infty(P)$,

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

P2: *Considered as a Lie bracket on* $C^\infty\ (P)$*,* $\{\cdot,\cdot\}$ *makes* $C^\infty\ (P)$ *into a Lie algebra; that is is* $\{\cdot,\cdot\}$ *is skew-symmetric and satisfies the Jacobi identity*

We see that the canonical Poisson bracket defined on a symplectic manifold indeed meets the requirements of this definition. In the symplectic case, the starting point for our formulation of Hamiltonian mechanics was the notion of Hamiltonian vector fields. We can easily define a similar notion on a Poisson manifold using the natural correspondence between derivations of $C^\infty(P)$ and vector fields on P; every derivation determines a unique vector field on P and vice versa. Since by the definition above $\{\cdot, h\} : C^\infty(P) \to C^\infty(P)$ is a derivation, we can associate a unique vector field to h in the following fashion:

Definition 1.45 ***(Hamiltonian vector field)*** *The Hamiltonian vector field associated to the function* $h \in C^\infty\ (P)$ *is the unique vector field* X_h *such that for all* $f \in C^\infty\ (P)$

$$X_h[f] = df(X_h) = \{f, h\}$$

Since $\{\cdot, G\} : C^\infty(P) \to C^\infty(P)$ is a derivation, we see that $\{F, G\}(p)$ depends on F only through the cotangent vector $dF_p \in T_p^*P$. Similarly, $\{F, G\}(p)$ depends on G only through the covector $dG_p \in T_p^*$P. Since the differentials of functions are sections of the cotangent bundle T^*P, it follows that the Poisson bracket is equivalent to a skew-symmetric, bilinear map

$$B : T^*P \times T^*P \to \mathbb{R}$$

such that

$$B_p(dF_p, dG_p) = \{F, G\}(p) \tag{1.5}$$

Hence we may think of B as a section of $\bigwedge^2 TP$, that is, a contravariant 2-tensor field on P. We call this tensor the **Poisson tensor** associated to the Poisson manifold P. In local coordinates $(x_1, \ldots, x_n)$ on M, we can expand the tensor B as

$$B = \sum_{i,j=1}^{n} B^{ij}(x_1 \ldots, x_n) \frac{\partial}{\partial x_i} \otimes \frac{\partial}{\partial x_j}$$

From the defining relation 1.5, it follows that the components B^{ij} are given by

$$B^{ij} = \{x_i, x_j\}$$

Thus, the Poisson bracket on a Poisson manifold is entirely determined by the Poisson bracket of coordinate functions on P. This observation proves particularly useful in computations.

Suppose that (M, ω) is a connected symplectic manifold. By the non-degeneracy of ω, any function $H \in C^\infty(M)$ such that $\{H, F\} \equiv 0$ for all $F \in C^\infty(M)$ must be a constant function. For a Poisson manifold, however, the situation may be very different:

Definition 1.46 ***(Casimir function)*** *Let P be a Poisson manifold. A function $H \in C^\infty(P)$ such that $\{H, F\} = 0$ for all $F \in C^\infty(M)$ is called a Casimir function on P.*

We now present an important example of a class of Poisson manifolds which are not in general symplectic.

The dual vector space $\mathfrak{g}^*$ to a Lie algebra $\mathfrak{g}$ is a natural example of a Poisson manifold. We define a Poisson bracket on $\mathfrak{g}^*$ as follows: if $f, g \in C^\infty(\mathfrak{g}^*)$ are two smooth functions on $\mathfrak{g}^*$, at each point $\alpha \in \mathfrak{g}^*$ their differentials df_α, dg_α are linear functionals on the tangent space $T_\alpha\mathfrak{g}^*$. Making the usual identification $T_\alpha\mathfrak{g}^* \cong \mathfrak{g}^*$, we can consider df_α and dg_α as elements of $(\mathfrak{g}^*)^* \cong \mathfrak{g}$. We shall make no notational distinction between the differential df_α and its representative in the Lie algebra $\mathfrak{g}$. If we let $\langle\ ,\ \rangle : \mathfrak{g}^* \times \mathfrak{g} \to \mathbb{C}$ denote the natural pairing between covectors in $\mathfrak{g}^*$ and vectors in $\mathfrak{g}$, for all $\xi \in T_\alpha(\mathfrak{g}^*) \cong \mathfrak{g}^*$, this identification reads

$$df_\alpha(\xi) = \langle \xi, df_\alpha \rangle$$

We then define the **Lie-Poisson bracket** of f and g to be the following function on $\mathfrak{g}^*$:

$$\{f, g\}(\alpha) := \langle \alpha, [df_\alpha, dg_\alpha] \rangle$$

This bracket is evidently bilinear and skew-symmetric. The fact that is satisfies the Leibniz rule follows easily from the Leibniz rule of exterior differentiation:

$$\begin{aligned} \{f, gh\}(\alpha) &= \langle \alpha, [df_\alpha, d(gh)_\alpha] \rangle \\ &= \langle \alpha, [df_\alpha, g(\alpha) dh_\alpha] \rangle + \langle \alpha, [df_\alpha, h(\alpha) dg_\alpha] \rangle \\ &= g(\alpha)\{f, h\}(\alpha) + h(\alpha)\{f, g\} \end{aligned}$$

In order to prove the Jacobi identity for the Lie-Poisson bracket, by our discussion of the Poisson tensor it suffices to verify it for a set of linear coordinate functions $\{z^i\}$ on $\mathfrak{g}^*$. If $\{X_i\}$ is a basis for $\mathfrak{g}$ such that $\langle z^i, X_j\rangle = \delta_{ij}$ we have

$$\{z^i, \{z^j, z^k\}\}(\alpha) = \langle \alpha, [dz^i, d\{z^j, z^k]\}_\alpha\rangle = \langle \alpha, [X_i, [X_j, X_k]\rangle$$

and thus the Jacobi identity for the Lie-Poisson bracket follows from the Jacobi identity in the Lie algebra $\mathfrak{g}$. Hence the Lie-Poisson bracket does indeed equip $\mathfrak{g}^*$ with the structure of a Poisson manifold. Note that if the dimension of $\mathfrak{g}^*$ is odd, then it is an example of Poisson manifold which is not symplectic.

The Hamiltonian vector field of a function with respect to the Lie-Poisson structure on $\mathfrak{g}^*$ is easily described:

Proposition 1.47 *The Hamiltonian vector field with respect to the Lie-Poisson structure associated to $H \in C^\infty(\mathfrak{g}^*)$ is*

$$X_H(\alpha) = \mathrm{ad}^*_{dH_\alpha}(\alpha)$$

Proof. By definition, the Hamiltonian vector field of H is the unique vector field $X_H \in \mathrm{Vect}(M)$ such that for all $f \in C^\infty(\mathfrak{g}^*)$

$$\{f, H\} = X_H[f]$$

Evaluated at a point $\alpha \in \mathfrak{g}^*$, this condition reads

$$\langle \alpha, [df_\alpha, dH_\alpha]\rangle = \langle X_H(\alpha), df_\alpha\rangle$$

But recalling the definition of the infinitesimal adjoint action, we see that

$$\langle \alpha, [df_\alpha, dH_\alpha]\rangle = \langle \alpha, -\mathrm{ad}_{dH_\alpha}(df_\alpha)\rangle = \langle \mathrm{ad}^*_{dH_\alpha}(\alpha), df_\alpha\rangle$$

and so we conclude that $X_H(\alpha) = \mathrm{ad}^*_{dH_\alpha}(\alpha)$ as claimed. □

This description of the Hamiltonian vector fields with respect to the Lie-Poisson bracket has some useful corollaries. Recall that $h \in C^\infty(\mathfrak{g}^*)$ is (infinitesimally) coadjoint invariant if for all $\alpha \in \mathfrak{g}^*, X \in \mathfrak{g}$ it satisfies

$$\langle \mathrm{ad}^*_X(\alpha)\ , dh_\alpha\rangle = 0 = \langle \alpha\ , [dh_\alpha, X]\rangle$$

Corollary 1.48 *The set of coadjoint invariant functions on $\mathfrak{g}^*$ is precisely the set of Casimir functions for the Lie-Poisson bracket, i.e. the set of functions which Poisson commute with all other functions with respect to the Lie-Poisson bracket.*

Proof. This follows from the fact that for $f, g \in C^\infty(\mathfrak{g}^*)$

$$\{f, g\}(\alpha) = \langle \alpha, [df_\alpha, dg_\alpha]\rangle = \langle \mathrm{ad}^*_{dg_\alpha}(\alpha), df_\alpha\rangle$$

□

Given a smooth function defined on our Poisson manifold P,(i.e. an observable), we want to understand how such a function varies along the integral curves of a Hamiltonian vector field. This is the content of the following observation:

Lemma 1.49 *Let $f, h \in C^\infty(P)$ and let φ_t be the flow of the Hamiltonian vector field X_h. Then*

$$\frac{d}{dt}(f \circ \varphi_t) = \{f, h\} \circ \varphi_t = \{f \circ \varphi_t, h\}$$

Proof. We simply compute

$$\begin{aligned}\frac{d}{dt}(f \circ \varphi_t)(p) &= df_{\varphi_{t(p)}}(X_h(\varphi_t(p))) \\ &= \{f, h\}(\varphi_t(p)) \\ &= df_{\varphi_{t(p)}} \circ (d\varphi_t)_z(X_h(p)) \\ &= d(f \circ \varphi_t)_z(X_h(p)) \\ &= \{f \circ \varphi_t, h\}(p)\end{aligned}$$

□

Next we consider smooth maps between Poisson manifolds:

Definition 1.50 ***(Poisson maps)*** *A smooth map $\varphi : (P_1, \{\cdot, \cdot\}_1) \to (P_2, \{\cdot, \cdot\}_2)$ is called a* Poisson map *if it preserves the Poisson bracket structure: that is*

$$\{f \circ \varphi, g \circ \varphi\}_1 = \{f, g\}_2 \circ \varphi$$

On a symplectic manifold, we know that the flow of a Hamiltonian vector field is a 1-parameter family of symplectomorphisms. By Proposition 1.27, this means that the flow consists of transformations which preserve the symplectic Poisson bracket. The analogue of this result is also true for Poisson manifolds:

Proposition 1.51

The flow φ_t of a Hamiltonian vector field X_h consists of Poisson maps: that is for all t for which φ_t is defined,

$$\{f \circ \varphi_t, g \circ \varphi_t\} = \{f, g\} \circ \varphi_t$$

Proof. Define a function K by

$$K = \{f \circ \varphi_t, g \circ \varphi_t\} - \{f, g\} \circ \varphi_t$$

Using the Poisson bracket's bilinearity, we compute the time derivative of K:

$$\frac{dK}{dt} = \{\frac{d}{dt}(f \circ \varphi_t), g \circ \varphi_t\} + \{f \circ \varphi_t, \frac{d}{dt}(g \circ \varphi_t)\} - \frac{d}{dt}(\{f, g\} \circ \varphi_t)$$

Now we rewrite this using the result of the previous lemma:

$$\frac{dK}{dt} = \{\{f \circ \varphi_t, h\}, g \circ \varphi_t\} + \{f \circ \varphi_t, \{g \circ \varphi_t, h\}\} - \{\{f, h\} \circ \varphi_t, h\}$$

Applying the Jacobi identity yields $\frac{dK}{dt} = \{K, h\}$, which has the unique solution $K_t = K_0 \circ \varphi_t$. But since $K_0(p) = 0$ for all $p \in P$, we conclude $K_t(p) = 0$ for all t, proving the proposition. □

The next proposition shows how Poisson maps allow us to 'pull back' Hamiltonian systems:

Proposition 1.52 *Suppose $\varphi : P_1 \to P_2$ is a Poisson map and let $H \in C^\infty(P_2)$ be a smooth function defined on P_2. Denote by ψ_t the flow of the Hamiltonian vector field X_H and let ρ_t be the flow of the Hamiltonian vector field $X_{H\circ\varphi}$. Then $\varphi \circ \rho_t = \psi_t \circ \varphi$, and $(d\varphi)_p \cdot (X_{H\circ\varphi}(p)) = X_H(\varphi(p))$ for all $p \in P_1$.*

Proof. Observe that for any point $p \in P_1$

$$\begin{aligned}\frac{d}{dt}G((\varphi \circ \rho_t)(p)) &= \frac{d}{dt}(G \circ \varphi)(\rho_t(p)) \\ &= d(G \circ \varphi)_{\rho_t(p)} \cdot X_{H\circ\varphi}(\rho_t(p)) \\ &= \{G \circ \varphi, H \circ \varphi\}(\rho_t(p)) \\ &= \{G, H\}(\varphi \circ \rho_t)(p)\end{aligned}$$

That is,

$$dG_{\varphi\circ\rho_t(z)} \cdot \left(\frac{d}{dt}(\varphi \circ \rho\ _t)(p)\right) = dG_{\varphi\circ\rho_t(p)} \cdot X_H(\varphi \circ \rho_t(p))$$

So we see that $(\varphi \circ \rho_t)(p)$ is an integral curve of X_H passing through the point $\varphi(p)$ on P_2 at $t = 0$. But of course, by definition of ψ_t, $(\psi_t \circ \varphi)(p)$ is another such curve. So the uniqueness of integral curves means that

$$(\varphi \circ \rho_t)(z) = (\psi_t \circ \varphi)(z) \tag{1.6}$$

Differentiating the identity (1.6) with respect to time yields

$$(d\varphi)_p \cdot (X_{H\circ\varphi}(p)) = X_H(\varphi(p)) \tag{1.7}$$

□

As in the symplectic case, we have:

Proposition 1.53 *If $f, h \in C^\infty(P)$, and X_f, X_h are their respective Hamiltonian vector fields, then the following identity holds:*

$$[X_f, X_h] = -X_{\{f,h\}}$$

where $[\cdot, \cdot]$ denotes the Lie bracket of vector fields on P.

Proof. This is a consequence of the Jacobi identity for the Poisson bracket. Considering the actions of the vector fields as derivations, for any function $K \in C^\infty(P)$ we have

$$\begin{aligned}[X_f, X_h][K] &= X_f[X_h[K]] - X_h[X_f[K]] \\ &= \{\{K, h\}, f\} - \{\{K, f\}, h\} \\ &= -\{K, \{f, h\}\} \\ &= -X_{\{f,h\}}[K]\end{aligned}$$

□

1.4 Submanifolds of Poisson Manifolds

Given a submanifold S of a Poisson manifold P, it is natural to ask if we can use the Poisson structure of P to S make into a Poisson manifold in its own right.

Definition 1.54 ***(Poisson immersion)*** *Let P be a Poisson manifold and let $i : S \to P$ be an injective immersion. Then i is a Poisson immersion if any Hamiltonian vector field defined on some open subset of P containing $i(S)$ is in the range of the derivative map $d_p i$ at all points $i(p)$ for $p \in S$.*

Proposition 1.55 *An immersion $i : S \to P$ is Poisson if and only if the following condition is satisfied: if $F, G : V \subset S \to \mathbb{R}$ are functions defined on an open subset $V \subset S$ and $\bar{F}, \bar{G} : U \to \mathbb{R}$ are extensions of $F \circ i^{-1}$, $G \circ i^{-1}$ to an open neighbourhood U of $i(V)$ in P, then $\{\bar{F}, \bar{G}\}$ is well-defined and independent of the extensions.*

For the proof, the reader is referred to page 301 of [24]. Proposition 1.55 implies that an injectively immersed submanifold of a Poisson manifold inherits a well-defined Poisson structure precisely when the inclusion $I : S \hookrightarrow P$ is a Poisson immersion. Note also that by definition of a Poisson immersion, this means that a submanifold S of a Poisson manifold is Poisson if and only if every locally defined Hamiltonian vector field is tangent to S.

Example 1.56 We consider the Poisson manifold $\mathfrak{g}^*$ equipped with the Lie-Poisson bracket. Recall from Proposition 1.55 that the Hamiltonian vector field associated to $H \in C^\infty(\mathfrak{g}^*)$ is $X_H(\alpha) = \mathrm{ad}^*_{dH_\alpha}(\alpha)$. Thus, Proposition 1.54 implies that the coadjoint orbits are injectively immersed Poisson submanifolds of $\mathfrak{g}^*$, and thus inherit a Poisson bracket from the Lie-Poisson bracket on $\mathfrak{g}^*$. We now show that this Poisson bracket agrees with the symplectic Poisson bracket arising from the coadjoint orbit symplectic structure we constructed previously. To see this, note that the Hamiltonian vector field of a function $F \in C^\infty(\Omega)$ with respect to the orbit symplectic structure is given by $X^\Omega_F(\alpha) = \mathrm{ad}^*_{dF}(\alpha)$, since for any tangent vector $\mathrm{ad}^*_Y(\alpha) \in T_\alpha\Omega$

we have

$$\begin{aligned} dF_\alpha \cdot \mathrm{ad}^*_Y(\alpha) &= \langle \mathrm{ad}^*_Y(\alpha), dF_\alpha \rangle \\ &= \langle \alpha, [dF_\alpha, Y] \rangle \\ &= \omega_\alpha(dF_\alpha, Y) \\ &= \omega(\mathrm{ad}^*_{dF}(\alpha), \mathrm{ad}^*_Y(\alpha)) \end{aligned}$$

So we see that the Hamiltonian vector field of F with respect to the orbit symplectic structure is the same as the Hamiltonian vector field with respect to the inherited Lie-Poisson structure, and thus the two Poisson structures coincide.
We will say that submanifold $M \subset \mathfrak{g}^*$ is ad^*-invariant if for all $X \in \mathfrak{g}$, $m \in M$

$$\mathrm{ad}^*_X(m) \in T_m M$$

Note that Proposition 1.55 implies that the Poisson submanifolds of $\mathfrak{g}^*$ are precisely the ad^*-invariant submanifolds.

Let us return to the description of the Poisson manifold $\mathfrak{g}^*$ given in the example above: $\mathfrak{g}^*$ is a disjoint union of its coadjoint orbits, which are injectively immersed symplectic submanifolds. Moreover, the restriction of the Lie-Poisson bracket to the coadjoint orbits is equivalent to the natural symplectic structure on the orbits. It turns out that a similar picture also holds for a general Poisson manifold:

Definition 1.57 ***(Symplectic leaves)*** *If P is a Poisson manifold, we say two points z_1, z_2 are on the same symplectic leaf if there is a piecewise-smooth curve joining z_1 and z_2, and each piece of this curve is an integral curve of a locally defined Hamiltonian vector field.*

We then have

Theorem 1.58 ***(Symplectic stratification theorem)*** *If P is a finite dimensional Poisson manifold, then P is the disjoint union of its symplectic leaves. Each symplectic leaf is an injectively immersed Poisson submanifold, and is also a symplectic manifold with symplectic form induced by the Poisson structure.*

For a proof, the reader is directed to [1], page 302.

1.5 Moment Maps

Let M be a Poisson manifold, G be a Lie group and let $\Phi : G \times M \to M$ be an action of G on M. As usual, we use the notations $\Phi(g, m) = \Phi_g(m) = g \cdot m$ interchangeably. We say that the action is **canonical** if for all $g \in G$, the maps $\Phi_g : M \to M$ are Poisson maps: that is, for all $g \in G$ and $F_1, F_2 \in C^\infty(M)$

$$\{F_1, F_2\} \circ \Phi_g = \{F_1 \circ \Phi_g, F_2 \circ \Phi_g\}$$

If we denote by ξ_M the infinitesimal generator of the action corresponding to $\xi \in \mathfrak{g}$, differentiating this condition yields

$$\xi_M \left[\{F_1, F_2\}\right] = \{\xi_M[F_1], F_2\} + \{F_1, \xi_M[F_2]\}$$

Definition 1.59 ***(Poisson action)*** *Let M be a Poisson manifold and let G be a Lie group acting canonically on M. We say that the action of G on M is a Poisson action if for all $\xi \in \mathfrak{g}$ there exists a smooth function $J^\xi \in C^\infty(M)$ such that ξ_M is the Hamiltonian vector field of J^ξ, that is,*

$$X_{J^\xi} = \xi_M$$

Let us note that the functions J^ξ are not unique, since any two functions whose difference is a Casimir will have the same Hamiltonian vector field. Moreover, we might as well assume that the map $J : \mathfrak{g} \to C^\infty(M), \xi \mapsto J^\xi$ is linear: if $\{e_i\}$ is a basis for $\mathfrak{g}$, then the new functions $\tilde{J}^\xi$ defined by

$$\tilde{J}^{\xi_1 e_1 + \ldots \xi_n e_n} = \sum_i \xi_i J^{e_i}$$

also satisfy the definition.

Thus a Poisson action of G on M gives rise to a map

$$\mathfrak{g} \times M \to \mathbb{R}, \quad (\xi, m) \mapsto J^\xi(m)$$

such that if we fix a point $m \in M$, we obtain a linear map

$$\mu_m : \mathfrak{g} \to \mathbb{R}$$

or, in other words, a linear functional $\mu_m \in \mathfrak{g}^*$. So the assignment

$$m \mapsto \mu_m$$

determines a map $\mu : M \to \mathfrak{g}^*$ such that

$$\langle \mu(m), \xi \rangle = J^\xi(m)$$

Definition 1.60 ***(Moment map)*** *Let G be a Lie group acting canonically on a Poisson manifold M, and suppose that there is a linear map $J : \mathfrak{g} \to C^\infty(M)$ such that*

$$X_{J^\xi} = \xi_M$$

for all $\xi \in \mathfrak{g}$. Then the map $\mu : M \to \mathfrak{g}^*$ satisfying

$$\langle \mu(m), \xi \rangle = J^\xi(m)$$

for all $\xi \in \mathfrak{g}, m \in M$ is called a moment map for the G-action.

One reason for the importance of moment maps in mechanics is that they make explicit the connection between symmetries of a Hamiltonian system and its conserved quantities:

Theorem 1.61 ***(Noether's theorem)*** *Let M be a Poisson manifold admitting a Poisson group action by a Lie group G, with corresponding moment map μ. Then if a function H is a G-invariant, that is,*

$$H \circ \Phi_g = H$$

for all $g \in G$, the moment map is a conserved quantity for the Hamiltonian dynamical system on M with Hamiltonian H. That is, if φ_t is the flow of the Hamiltonian vector field X_H, then

$$\mu \circ \varphi_t = \mu$$

Proof. Differentiating the invariance condition $H \circ \Phi_g = H$ along suitable curves in G implies that $\xi_M[H] = 0$. But by definition of J^ξ this implies $\{J^\xi, H\} = 0$, and we see that J^ξ, and thus μ, is a conserved quantity under the Hamiltonian flow of H. $\square$

The moment maps of interest to us will satisfy the following compatibility condition with the coadjoint action:

Definition 1.62 ***(Equivariant moment map)*** *Let M be a Poisson manifold admitting a Poisson group action by a Lie group G, with corresponding moment map μ. The moment map μ is said to be equivariant if for all $g \in G$,*

$$\mu \circ \Phi_g = \mathrm{Ad}^*_{g^{-1}} \circ \mu$$

Lemma 1.63 *If μ is an equivariant moment map, then for all $\xi, \eta \in \mathfrak{g}$ the corresponding linear map $J : \mathfrak{g} \to C^\infty(M)$ satisfies*

$$J^{[\xi,\eta]} = \{J^\xi, J^\eta\}$$

Proof. By the equivariance condition $\mu \circ \Phi_g = \mathrm{Ad}^*_{g^{-1}} \circ \mu$, for all $\xi \in \mathfrak{g}$, $g \in G$ and $m \in M$ we have

$$\begin{aligned} \langle \mathrm{Ad}^*_{g^{-1}} \mu(m), \xi \rangle &= \langle \mu(m), \mathrm{Ad}_g \xi \rangle \\ &= J^\xi(g \cdot m) \end{aligned}$$

Differentiating this equality along a smooth path $g = g(t)$ with $g(0) = e$ and $g'(0) = \eta$ at $t = 0$ yields

$$\langle \mu(m), [\xi, \eta] \rangle = J^{[\xi,\eta]} = dJ^\xi(\eta_M(m))$$

and noticing $dJ^\xi(\eta_M(m)) = \{J^\xi, J^\eta\}$ we obtain the result. $\square$

Corollary 1.64 *An equivariant moment map is a Poisson map with respect to the Lie-Poisson structure on $\mathfrak{g}^*$*

Proof. Given a point $\mu(m) = \alpha \in \mathfrak{g}^*$ and $F, G \in C^\infty(\mathfrak{g}^*)$, write $\xi = dF_\alpha, \eta = dG_\alpha$. Using the Lemma 1.63, we can compute

$$\{F, G\} \circ \mu(m) = \langle \alpha, [\xi, \eta] \rangle = J^{[\eta,\xi]}(m) = \{J^\xi, J^\eta\}$$

Now for all $v \in T_m M$, by the chain rule we have

$$\begin{aligned} d(F \circ \mu)_m \cdot v &= (dF)_\alpha \cdot (dJ)_m \cdot v \\ &= \langle \xi, (dJ)_m \cdot v \rangle \\ &= (dJ^\xi)_m \cdot v \end{aligned}$$

and so we see that the differentials of the functions $F \circ \mu$ and J^ξ on M are equal. Since by its derivation property the Poisson bracket depends only on these differentials, we conclude that

$$\{F, G\} \circ \mu(m) = \{J^\xi, J^\eta\}(m) = \{F \circ \mu, G \circ \mu\}(m)$$

and thus μ is a Poisson map $\square$

1.6 Symplectic Reduction

We conclude our overview of Hamiltonian mechanics by with a discussion of the reduction of a Hamiltonian system on a symplectic manifold with respect to a symmetry. Let (M, ω) be a symplectic manifold and let G be a Lie group G acting canonically M and admitting an equivariant moment map $\mu : M \to \mathfrak{g}^*$. From this data, we aim to construct a new symplectic manifold N which we interpret as the symplectic quotient of M by G. The manifold N is called the symplectic (or Marsden-Weinstein) reduction of M by G. If $m \in M$, let $\mathfrak{g}_p$ denote the G_p denote the stabilizer subgroup of p, and let $\mathfrak{g}_p$ be its Lie algebra. Let Ω_p be the G-orbit in M through the point p, which, as usual, we regard as injectively immersed submanifold of M diffeomorphic to G/G_p. Recall that the tangent space $T_p\Omega_p \subset T_pM$ is identified as

$$T_p\Omega_p = \{\xi_M(p) | \xi \in \mathfrak{g}\}$$

where ξ_M is the infinitesimal generator corresponding to $\xi \in \mathfrak{g}$.

Proposition 1.65 *If $d\mu_p$ is the derivative of the moment map at $p \in M$ and $\mathfrak{g}_p^0 \subset \mathfrak{g}^*$ is the annihilator of the stabilizer algebra $\mathfrak{g}_p$, then* $\text{im}\,(d\mu_p) = \mathfrak{g}_p^0$

Proof. By definition of the moment map, for all $v \in T_pM$ and $\xi \in \mathfrak{g}$ we have

$$\langle d\mu_p(v), \xi \rangle = dJ_p^\xi(v) = \omega_p(\xi_M(p), v)$$

Thus, since $\alpha = 0 \iff \langle \alpha, \xi \rangle = 0$ for all $\xi \in \mathfrak{g}$ we find

$$\ker\,(d\mu_p) = \{v \in T_pM | \omega_p(\xi_M(p), v) = 0 \text{ for all } \xi \in \mathfrak{g}\} = (T_p\Omega_p)^{\omega_p}$$

where $(T_p\Omega_p)^{\omega_p}$ is the orthogonal complement of the subspace $T_p\Omega_p \subset T_pM$ with respect to the form ω_p on T_pM. Moreover, observe that by definition of the stabilizer subgroup, the infinitesimal generator corresponding to $\xi \in \mathfrak{g}_p$ satisfies $\xi_M(p) = 0$. Hence, for all $v \in T_pM$, $\xi \in \mathfrak{g}_p$

$$\langle d\mu_p(v), \xi \rangle = \omega_p(\xi_M(p), v) = 0$$

and so we see that the image of T_pM under $d\mu_p$ is contained in $\mathfrak{g}_p^0 \subset \mathfrak{g}^*$, the annihilator of the subspace $\mathfrak{g}_p \subset \mathfrak{g}$. Counting dimensions, we see that indeed $\text{im}\,(d\mu_p) = \mathfrak{g}_p^0$. ☐

Let us now consider $\mu^{-1}(0) \subset M$, the preimage of $0 \in \mathfrak{g}^*$ under the moment map. Observe that by the equivariance of μ, $\mu^{-1}(0)$ is preserved by the action of G. Suppose that the action of G on $\mu^{-1}(0)$ is free, so that for each $p \in \mu^{-1}(0)$, the stabilizer subgroups G_p are all trivial. Then $\mathfrak{g}_p = \{0\}$ for all $p \in \mu^{-1}(0)$, which means $\mathfrak{g}_p^0 = \mathfrak{g}^*$ and we conclude that $d\mu_p$ is surjective. So under the assumption that G acts freely on $\mu^{-1}(0)$, 0 is a regular value of the moment map, and thus $\mu^{-1}(0)$ is a submanifold of M of dimension $\dim M - \dim G$. Moreover, the tangent space $T_p\mu^{-1}(0)$ is given by

$$T_p\mu^{-1}(0) = \ker\,(d\mu_p) = (T_p\Omega_p)^{\omega_p}$$

and so we find that $T_p\mu^{-1}(0)$ and $T_p\Omega_p$ are symplectic orthogonal complements. Since $T_P\Omega_p \subset T_p\mu^{-1}(0)$ this means that the G-orbits are isotropic.

The symplectic reduction procedure relies on the following result[8] in linear algebra:

[8] In [6], Audin refers to this fact as 'a minor miracle of symplectic geometry'

Lemma 1.66 *Let (V, ω) be a symplectic vector space and let $I \subset V$ be an isotropic subspace and let I^ω be its orthogonal complement. Then ω induces a canonical symplectic form $\tilde{\omega}$ on the quotient space I^ω / I*

Proof. If $[u], [v]$ denote equivalence classes in I^ω/I with representatives $u, v \in I^\omega$, the form $\tilde{\omega}$ is defined by

$$\tilde{\omega}([u], [v]) = \omega(u, v)$$

Since $\omega(i, j) = 0$ for all $i, j \in I$, the form $\tilde{\omega}$ is well-defined. To check that $\tilde{\omega}$ is non-degenerate, note that if $u \in I^\omega$ satisfies $\omega(u, v) = 0$ for all $v \in I^\omega$, then $u \in (I^\omega)^\omega = I$, and thus $[u] = 0$ □

With this result in mind, we can construct the Marsden-Weinstein reduction of M by G. If we assume that the action of G on $\mu^{-1}(0)$ is proper in addition to being free, then the orbit space $N := \mu^{-1}(0)/G$ is a smooth manifold of (even)dimension $\dim M - 2\dim G$, whose tangent space at $[p]$ is identified with the quotient space $T_p\mu^{-1}(0)/T_p\Omega_p$, where Ω_p is the G-orbit through p.

Now recall that at each point $p \in \mu^{-1}(0)$, $T_p\Omega_p$, is an isotropic subspace of the symplectic vector space T_pM, ω_p. So by the preceding lemma, ω_p defines a symplectic form $\tilde{\omega}_p$ on the quotient space $T_p\mu^{-1}(0)/T_p\Omega_p \sim T_{[p]} \sim T_{[p]}N$. And our definition is independent of the choice of representative $[p]$, since the action of G is canonical, so that $\Phi_g^*\omega = \omega$ for all $g \in G$. Thus letting $[p]$ vary, we obtain a non-degenerate 2-form $\tilde{\omega}$ on N.

It remains to show that $\tilde{\omega}$ is closed. Let $i : \mu^{-1}(0) \to M$ be inclusion and $\pi : \mu^{-1}(0) \to N$ be the natural projection. Then, by construction, the reduced form $\tilde{\omega}$ on N satisfies $i^*\omega = \pi^*\tilde{\omega}$. So by naturality of d,

$$\pi^* d\tilde{\omega} = d\pi^*\tilde{\omega} = di^*\omega = i^*d\omega = 0$$

Since the natural projection π is a submersion, the fact that $\pi^*d\tilde{\omega} = 0$ implies $d\tilde{\omega} = 0$. So we see that $\tilde{\omega}$ is indeed closed, and thus defines a symplectic form on N.

Example 1.67 (The Neumann model) We apply the technique of symplectic reduction to the Neumann model, which describes the motion of a point particle constrained to the surface of the unit sphere $S^{n-1} \subset \mathbb{R}^n$ and subjected to harmonic oscillator forces. The system's phase space M is the cotangent bundle of the sphere $S^{n-1} \subset \mathbb{R}^{n-1}$, which can be identified with the tangent bundle via the Euclidean metric:

$$M = T^*S^{n-1} \cong \left\{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^n \times \mathbb{R}^n : \mathbf{x}^T\mathbf{x} = 1, \mathbf{x}^T\mathbf{y} = 0\right\} \subset \mathbb{R}^n \times \mathbb{R}^n$$

Fix a diagonal $n \times n$ matrix A with distinct eigenvalues

$$A = \operatorname{diag}(\alpha_1, \ldots, \alpha_n) \text{ }, \alpha_i \neq \alpha_j \text{ if } i \neq j$$

Then the Neumann Hamiltonian corresponding to the oscillator strengths $\{\alpha_i\}_{i=1}^n$ is

$$H : M \to \mathbb{R}, \quad H(\mathbf{x}, \mathbf{y}) = \frac{1}{2}\left(\mathbf{x}^T A\mathbf{x} + \mathbf{y}^T\mathbf{y}\right) \tag{1.8}$$

It is in fact easier to understand the Hamiltonian formulation of the Neumann model by first passing to the larger phase space $T^*\mathbb{R}^n \cong \mathbb{R}^{2n}$ and then performing a symplectic reduction with to a symmetry. Because of the constraints $\mathbf{x}^T\mathbf{x} = 1$, $\mathbf{x}^T\mathbf{y} = 0$, the Hamiltonian may be equivalently chosen as

$$\phi(\mathbf{x}, \mathbf{y}) = \frac{1}{2}[(\mathbf{x}^T\mathbf{x})(\mathbf{y}^T\mathbf{y}) + \mathbf{x}^T A\mathbf{x} - (\mathbf{x}^T\mathbf{y})^2] \tag{1.9}$$

If we define an angular momentum matrix $J_{kl} = [x_k y_l - x_l y_k]_{k,l=1,\dots,n}$, observe that the Hamiltonian ϕ can also be expressed as

$$\phi(\mathbf{x}, \mathbf{y}) = \frac{1}{2}\sum_{i=1}^{n} \alpha_i x_i^2 + \frac{1}{4}\sum_{k\neq l} J_{kl}^2$$

Now the Hamiltonian ϕ is invariant under the 1-parameter group T of diffeomorphisms

$$(\mathbf{x}, \mathbf{y}) \mapsto (\mathbf{x}, \mathbf{y} + \lambda\mathbf{x}), \;\; \lambda \in \mathbb{R}$$

Moreover, this action is canonical with respect to the standard symplectic structure $\omega = d\mathbf{x} \wedge d\mathbf{y}$ on $\mathbb{R}^{2n}$. Indeed, it admits a moment map

$$J(\mathbf{x}, \mathbf{y}) = \frac{1}{2}\left(\mathbf{x}^T\mathbf{x} - 1\right)$$

Hence we can perform a Marsden-Weinstein reduction with respect to this group. Fixing the value of the moment to $J = 0$, we have

$$J^{-1}(0) = \left\{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^n \times \mathbb{R}^n : \mathbf{x}^T\mathbf{x} = 1\right\}$$

To complete the symplectic reduction, we must take the quotient of $J^{-1}(0)$ under the action of T. We observe that orbit space $J^{-1}(0)/T$ can be identified with the original Neumann phase space T^*S^{n-1} via the symplectomorphism

$$\varphi : J^{-1}(0)/T \mapsto T^*S^{n-1} \subset \mathbb{R}^{2n}, \;\; [(\mathbf{x}, \mathbf{y})] \mapsto (\mathbf{x}, \mathbf{y} - (\mathbf{y}^T\mathbf{x})\mathbf{x})$$

Hence the flows generated by the Neumann Hamiltonian (1.9) are obtained from the unconstrained flows generated by the Hamiltonian (1.10) by the 'orthogonal projection':

$$(\mathbf{x}(t), \mathbf{y}(t)) \longmapsto \left(\mathbf{x}(t), \mathbf{y}(t) - \mathbf{x}^T(t)\mathbf{y}(t)\mathbf{x}(t)\right)$$

Chapter 2

The Liouville-Arnold Theorem

In classical mechanics, it is often possible to use principles such as conservation of energy and conservation of momentum to obtain both quantitative and qualitative information about a physical problem. Intuitively, we expect that the more conserved quantities one can identify, the more one should be able to say about a system. In the context of Hamiltonian mechanics on symplectic manifolds, this intuition is supported by the Liouville-Arnold theorem. The statement and proof of the Liouville-Arnold theorem that we present are based on Arnold's classic book [4] There exist several refinements of the theorem that describe a locally defined set of action variables I_j in addition to the angle coordinates φ_j discussed here. Together, the complete set of action-angle variables I_j, φ_j forms a system of local Darboux coordinates. For a proof of the existence of these coordinates, the reader is directed to [6]. In general, obtaining an explicit parameterisation of a system's action variables is highly non-trivial. In Babelon's book [7], this problem is discussed at length. Duistermaat's article [11] gives a detailed discussion of the topological obstructions to extending the action-angle coordinates globally.

2.1 The statement

Theorem 2.1 ***(Liouville-Arnold Theorem)*** *Let $(M\,\omega)$ be a $2n$-dimensional symplectic manifold, and let $F_1, \dots, F_n$ be a set of n smooth functions on M whose Poisson brackets satisfy $\{F_i, F_j\} \equiv 0,\ i, j = 1, 2 \dots, n$. Suppose that on the level set*

$$M_c = \{x \in M | F_i(x) = c_i,\ i = 1, \dots, n\}$$

the F_i are functionally independent; that is at each point of M_c their differentials satisfy $dF_1 \wedge \dots \wedge dF_n \neq 0$. Then

1. *The level set M_c is a smooth Lagrangian submanifold of M which is invariant under the flow of the Hamiltonian system with Hamiltonian function[1] $H = F_1$.*
2. *Provided M_c is compact and connected, it is diffeomorphic to an n-dimensional torus*

$$T^n = \{(\varphi_1, \dots, \varphi_n) \bmod 2\pi\}$$

3. *The flow of the Hamiltonian system with Hamiltonian function $H = F_1$ is linear on the torus M_c, so that the angular coordinates $\varphi_1, \dots, \varphi_n$ on M_c satisfy*

$$\frac{d\varphi_i}{dt} = \omega_i$$

for some constants $\omega_i \in \mathbb{R}$.

Proof. 1. By the assumption that $dF_1 \wedge \dots \wedge dF_n \neq 0$ on M_c, it follows immediately from the inverse function theorem that M_c is an n-dimensional embedded submanifold of M. Observe that the n Hamiltonian vector fields X_{F_i} are all tangent to M_c, since for all $1 \leq i, j \neq n$,

$$\{F_i, F_j\} = dF_i \cdot X_{F_j} = 0$$

Moreover, since the map

$$T_x^* M \to T_x M, \quad dF_x \mapsto X_F(x)$$

is an isomorphism, and the n 1-forms dF_i are linearly independent at each point on M_c, we see that the n Hamiltonian vector fields $X_{F_1}, \dots, X_{F_n}$ span each tangent space $T_x M_c$. But then since $\omega(X_{F_i}, X_{F_j}) = \{F_i, F_j\} = 0$, we see that the restriction of ω to $T_x M_c$ is identically zero, showing that M_c is a Lagrangian submanifold of M. And since under the Hamiltonian flow of $H = F_1$ we have $\frac{dF_i}{dt} = \{F_i, F_1\} = 0$, we see that M_c is invariant under this flow.

2. Let us now assume that the level set M_c is compact and connected. By the compactness assumption, the flow of any vector field on M_c is complete. By Proposition 1.28 of Chapter 1, the Lie brackets of the Hamiltonian vector fields X_{F_i}, X_{F_j} satisfy $[X_{F_j}, X_{F_i}] = -X_{\{F_i, F_j\}} = 0$. Consequently, their flows $\rho_i^{t_i}$:

[1] Or indeed any Hamiltonian dynamical system whose Hamiltonian function H can be expressed as a function of the F_i

$M_c \times \mathbb{R} \to M_c$ all commute: that is, $\rho_i^{t_i} \circ \rho_j^{t_j} = \rho_j^{t_j} \circ \rho_i^{t_i}$. Since all these flows are complete, we can use them to define an action of the abelian group $\mathbb{R}^n$ on M_c by

$$(t_1, \ldots, t_n) \cdot x = \rho_1^{t1} \circ \ldots \circ \rho_n^{t_n}(x)$$

for $(t_1, \ldots, t_n) \in \mathbb{R}^n$ and $x \in M_c$. We claim that this action of $\mathbb{R}^n$ on M_c is transitive: that is, for each $x \in M_c$, the map $\Phi_x : \mathbb{R}^{2n} \to M_c, (t_1, \ldots, t_n) \mapsto (t_1, \ldots, t_n) \cdot x$, whose image is a priori the $\mathbb{R}^n$-orbit through x, is in fact surjective. Indeed by the independence of the vector fields X_{F_i}, the map ρ_x is a local diffeomorphism, and thus its image is an open subset of M_c. Hence the orbits are disjoint open subsets of M_c, and so if M_c is connected it must consist of a single orbit.

We now fix a point $x \in M_c$. Let $\Sigma = \{(t_1, \ldots, t_n) \in \mathbb{R}^n | \rho_1^{t1} \circ \ldots \circ \rho_n^{t_n}(x) = x\}$ be the stabilizer subgroup of x. Recall from our discussion of group actions in Chapter 1 that the orbit of x, which is all of M_c, is diffeomorphic to $\mathbb{R}^n/\Sigma$. Since the mapping Φ_x is a local diffeomorphism, we see that the subgroup Σ must be discrete. We now appeal to the following fact: any discrete subgroup $\Sigma \subset \mathbb{R}^n$ is a integral lattice $\mathbb{Z}^k$ generated by some k-tuple of vectors, *i.e.*

$$\Sigma = \{a_1\mathbf{e}_1 + \ldots + a_k\mathbf{e}_k | a_i \in \mathbb{Z}\} \cong \mathbb{Z}^k$$

This is can be easily proved inductively: in the case $n = 1$, it is clear that a discrete subgroup $\Sigma \subset \mathbb{R}$ is must consist of integer multiples of a vector $e_1 \subset \Sigma$ with smallest Euclidean length. Indeed, if there were a vector $v \in \Sigma$ and some $k \in \mathbb{Z}$ such that $v = re_1$ and $k < r < k + 1$, then the length of the vector $v - ke_1$ would be strictly less than that of e_1 itself, a contradiction. If $n = 2$, we again choose $e_1 \in \Sigma$ to be a vector with smallest Euclidean length. Let $V_1 = \text{span}(e_1)$. If $\Sigma \subset V_1$, then by analysis of the case $n = 1$ the proof is complete. If not, we choose a vector e_2 at (non-zero) minimal distance from the line V_1. We claim that Σ consists of integer linear combinations of e_1, e_2. For a contradiction, suppose that exists a vector v which is not contained in the integer lattice spanned by e_1, e_2. If we think of this lattice as partitioning the plane into parallelograms spanned by the vectors e_1, e_2 as shown in Figure 2.1, then the vector v cannot coincide with any of the vertices of these parallelograms. So as in the case $n = 1$, by translating v by a suitable integer linear combination of e_1 and e_2 we can obtain a vector v' which is closer to V_1 than e_2, a contradiction.

Continuing similarly, we eventually obtain a set of linearly independent vectors $e_1, \ldots, e_k$ such that $\Sigma = \{a_1\mathbf{e}_1 + \ldots + a_k\mathbf{e}_k | a_i \in \mathbb{Z}\}$.

Now the quotient space $\mathbb{R}^n/\mathbb{Z}^k$ is diffeomorphic to the cylinder $T^k \times \mathbb{R}^{n-k}$. Since M_c is compact, we must have $k = n$, and we see that M_c is indeed diffeomorphic to an n-torus.

3. Finally, the angular coordinates $\{(\varphi_1, \ldots, \varphi_n) \bmod 2\pi\}$ are defined as follows: if $y \in M_c$ satisfies $y = \Phi_x(v)$, where $v = v_1e_1 + \ldots + v_ne_n$, then $\varphi_i(y) = 2\pi v_i$. Observe that by construction, under the flow of X_{F_1} these coordinates satisfy

$$\frac{d\varphi_i}{dt} = \omega_i$$

for some constants ω_i, which are determined by the writing the first standard basis vector $(1, 0, \ldots, 0) \in \mathbb{R}^n$ as a linear combination of the vectors e_i.

□

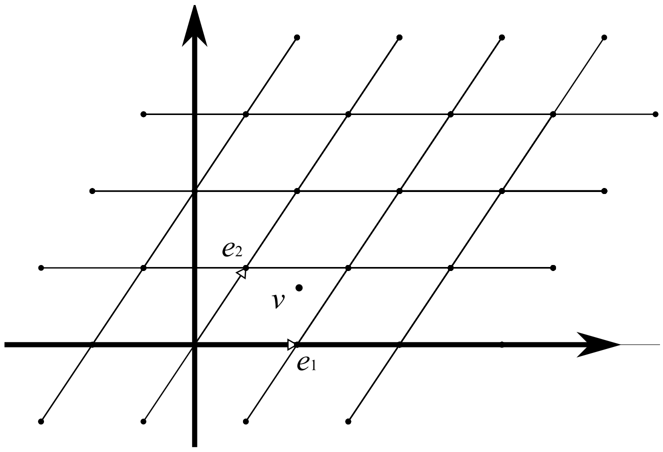

Figure 2.1: The partition of the plane into parallelograms whose vertices represent integral linear combinations of the vectors e_1 and e_2.

2.2 Examples of integrable systems

Example 2.2 (Two-dimensional systems) Observe that any Hamiltonian system defined on a two-dimensional symplectic manifold is trivially Liouville integrable: the Hamiltonian provides the required conserved quantity.

Example 2.3 (The Kepler Problem) The classical two-body (Kepler) problem is a Hamiltonian system defined on the symplectic manifold $(T^*\mathbb{R}^3 = \{(q_1, q_2, q_3, p_1, p_2, p_3)\}, \omega = \sum_{i=1}^n dq_i \wedge dp_i)$. The Hamiltonian is

$$H = \frac{1}{2}\sum_{i=1}^{3} p_i^2 + V(r)$$

where V is a smooth function of $r = \sqrt{q_1^2 + q_2^2 + q_3^2}$. Define the components[2] of angular momentum J_i by $J_1 = q_2p_3 - p_3q_2, J_2 = q_3p_1 - p_1q_3, J_3 = q_1p_2 - p_1q_2$. Then one verifies by direct computation that the functions H, J_3 and $J^2 := J_1^2 + J_2^2 + J_3^2$ pairwise Poisson commute and are generically functionally independent. Hence the Kepler problem is a Liouville integrable system.

Example 2.4 (The Euler Top) The Euler top is a Hamiltonian system that describes a rotating rigid body in $\mathbb{R}^3$ that is attached to a fixed point but not subjected to any external forces. The system is defined on the Poisson manifold $\mathfrak{so}(3)^*$ equipped with the Lie-Poisson bracket. Let J_1, J_2, J_3 be coordinate functions on $(\mathfrak{so}^3)^* \cong \mathbb{R}^3$. For fixed distinct constants[3] $I_1, I_2, I_3 \in \mathbb{R}$, the Hamiltonian for the Euler top is

$$H = \frac{1}{2}\sum_{i=1}^{3} \frac{J_i^2}{I_i}$$

[2] As a vector, $(J_1, J_2, J_3) = \mathbf{J} = \mathbf{q} \times \mathbf{p}$

[3] These are the *principal moments of inertia*. See [4] for a discussion of their physical meaning

Observe that the function $J^2 = J_1^2 + J_2^2 + J_3^2$ is a Casimir. Restricting to a symplectic leaf $\Sigma = J^{-1}(c) \subset \mathfrak{so}(3)^*$, we obtain a Hamiltonian system on a symplectic manifold of dimension 2, which is necessarily integrable.

Example 2.5 (The Neumann Model) The Neumann model is another example of a Liouville integrable system. Recall that the system's Hamiltonian can be taken to be

$$\phi(\mathbf{x}, \mathbf{y}) = \frac{1}{2}\sum_{i=1}^{n} \alpha_i x_i^2 + \frac{1}{4}\sum_{k \neq l} J_{kl}^2$$

where $J_{kl} = [x_k y_l - x_l y_k]_{k,l=1,\ldots,n}$ is the angular momentum matrix. Define a rational function in the parameter λ by

$$\Delta(\lambda) = -\frac{1}{8}\sum_{k=1}^{n}\sum_{l \neq k} \frac{J_{kl}^2}{(\lambda - \alpha_k)(\lambda - \alpha_l)} + \frac{1}{4}\sum_{k=1}^{n} \frac{x_k^2}{\lambda - \alpha_k}$$

Writing $a(\lambda) = \prod_{i=1}^{n}(\lambda - \alpha_i)$, observe that

$$\mathcal{P}(\lambda) = a(\lambda)\Delta(\lambda) =: \mathcal{P}_{n-1}\lambda^{n-1} + \mathcal{P}_{n-2}\lambda^{n-2} + \ldots + \mathcal{P}_0$$

is a polynomial in λ of degree $n-1$. Let us define n functions H_k by

$$\mathcal{P}(\lambda) = a(\lambda)\sum_{k=1}^{n} \frac{H_k}{\lambda - \alpha_k}$$

Using the fact that $\mathcal{P}(\alpha_k) = \prod_{i \neq k}(\alpha_i - \alpha_k)H_k$, we see that

$$H_k = \frac{1}{4}\sum_{l \neq k}^{n} \frac{J_{kl}^2}{\alpha_k - \alpha_l} + \frac{1}{4}x_k^2$$

Computing directly from the definition of $\mathcal{P}(\lambda)$, we find[4]

$$\mathcal{P}_{n-1} = \frac{1}{4}\sum_{k=1}^{n} x_k^2 = \frac{1}{4}$$

and

$$\mathcal{P}_{n-2} = \frac{1}{2}\phi(\mathbf{x}, \mathbf{y}) - \frac{1}{4}\sum_{k=1}^{n}\alpha_k$$

However from the definition of the H_k,

$$\mathcal{P}_{n-1} = \sum_{k=1}^{n} H_k$$

$$\mathcal{P}_{n-2} = \sum_{k=1}^{n}\alpha_k H_k - \sum_{k=1}^{n}\alpha_k \sum_{j=1}^{n} H_j = \sum_{k=1}^{n}\alpha_k H_k - \frac{1}{4}\sum_{k=1}^{n}\alpha_k$$

[4]In view of the constraint $\mathbf{x}^T\mathbf{x} = 1$

since $\sum_{j=1}^{n} H_j = \mathcal{P}_{n-1} = 1/4$.

Thus, we see that the functions H_k satisfy $\sum_{k=1}^{n} H_k = \frac{1}{4}$, and the Hamiltonian ϕ can be expressed in terms of the H_k as

$$\phi(\mathbf{x}, \mathbf{y}) = 2 \sum_{k=1}^{n} H_k$$

A lengthy but direct computation using the canonical commutation relations $\{x_i, x_j\} = 0 = \{y_i, y_j\}, \{x_i, y_j\} = \delta_{ij}$ shows that the H_k Poisson commute, and that any $n-1$ of them are generically functionally independent on $T^*S^{n-1} \subset \mathbb{R}^{2n}$. The functions H_k are known as the Devaney-Uhlenbeck conserved quantities. Thus, the Neumann model is a Liouville integrable system.

Chapter 3

Lax equations as Hamiltonian flows

This chapter is a bridge between the integrable Hamiltonian systems discussed in Chapter 2 and the Lax equations we shall consider in Chapter 5. We describe the AKS method of constructing integrable Hamiltonian systems on coadjoint orbits, whose equations of motion are of the Lax form. The treatment given here closely follows that of Audin in [6]. Other useful references include the encyclopaedic survey of [31] and the articles [2],[3]. Following the paper [26] of Adams, Harnad and Hurtubise, we shall apply the AKS construction to the case of a loop algebra to obtain Lax pairs with a spectral parameter. In particular, we shall see how the example of the Neumann model fits naturally into this framework.

3.1 Classical R-matrices

Let $(\mathfrak{g},\ [\cdot,\cdot])$ be a Lie algebra, and let $R:\mathfrak{g}\to\mathfrak{g}$ be a linear operator on $\mathfrak{g}$. Such an operator allows us to define a new bracket operation $[\cdot,\cdot]_R:\mathfrak{g}\times\mathfrak{g}\to\mathfrak{g}$ by

$$[X,Y]_R=[RX,Y]+[X,RY]$$

For any choice of operator R, the bracket $[\cdot,\cdot]_R$ is evidently bilinear and skew-symmetric. If it satisfies the Jacobi identity, we say that R is an R-**matrix** for the Lie algebra $\mathfrak{g}$. By definition, an R-matrix defines new Lie algebra structure $(\mathfrak{g},\ [\cdot,\cdot]_R)$ on the vector space $\mathfrak{g}$. We will denote this modified Lie algebra $\mathfrak{g}_R$.

Let us examine a situation in which such an R-matrix arises naturally. Suppose the Lie algebra $\mathfrak{g}$ has a decomposition as a vector space direct sum of two Lie subalgebras $\mathfrak{a}$ and $\mathfrak{b}$:

$$\mathfrak{g}=\mathfrak{a}\oplus\mathfrak{b}$$

Note that we do not require that $[a,b]=0$ for $a\in\mathfrak{a},b\in\mathfrak{b}$; this is a direct sum of vector spaces, not Lie algebras. Let P_+ denote the linear operator corresponding to orthogonal projection onto the subspace $\mathfrak{a}$ with respect to this decomposition, and let P_- denote the orthogonal projection onto $\mathfrak{b}$. For $X\in\mathfrak{g}$, we shall write $P_+X=X_+$, and $P_-X=X_-$.

Proposition 3.1 *The following operator R defines an R-matrix:*

$$R:\mathfrak{g}\to\mathfrak{g}\ ,\qquad X\mapsto\frac{1}{2}\left(P_+-P_-\right)$$

Proof. To see that the Jacobi identity is satisfied by the modified bracket $[\cdot,\cdot]_R$, we observe that

$$\begin{aligned}[X,Y]_R &= \frac{1}{2}\left([X_+ - X_-, Y] + [X, Y_+ - Y_-]\right)\\ &= \frac{1}{2}\left([X_+ - X_-, Y_+ + Y_-] + [X_+ + X_-, Y_+ - Y_-]\right)\\ &= [X_+, Y_+] - [X_-, Y_-]\end{aligned}$$

Then since

$$\begin{aligned}[[X,Y]_R, Z]_R &= [([X,Y]_R)_+, Z_+] - [([X,Y]_R)_-, Z_-]\\ &= [[X_+,Y_+],Z_+] - [[X_-,Y_-],Z_-]\end{aligned}$$

it is evident that the Jacobi identity for $[\cdot,\cdot]_R$ follows from the Jacobi identity in the Lie algebras $\mathfrak{a}$ and $\mathfrak{b}$. □

Using the Lie bracket $[\cdot,\cdot]_R$ on $\mathfrak{g}$ defined by an R-matrix, we obtain a new Lie-Poisson bracket $\{\cdot,\cdot\}_R$ on the vector space[1] $\mathfrak{g}^*$:

$$\{f,g\}_R := \langle \alpha, [df_\alpha, dg_\alpha]_R\rangle = \langle \alpha, [Rdf_\alpha, dg_\alpha]\rangle + \langle \alpha, [df_\alpha, Rdg_\alpha]\rangle$$

Given a function $h \in C^\infty(\mathfrak{g}^*)$ we can consider its Hamiltonian vector field X_h^R with respect to the new Poisson bracket $\{\cdot,\cdot\}_R$. If h is a Casimir function on $(\mathfrak{g}^*, \{\cdot,\cdot\})$, its Hamiltonian vector field with respect to the Poisson bracket $\{\cdot,\cdot\}_R$ has a particularly simple form:

Proposition 3.2 *If h is a Casimir function on the Poisson manifold $(\mathfrak{g}^*, \{\cdot,\cdot\})$ then its Hamiltonian vector field X_h^R with respect to the Poisson bracket $\{\cdot,\cdot\}_R$ is given by*

$$X_h^R(\alpha) = \mathrm{ad}^*_{Rdh_\alpha}(\alpha)$$

Proof. For all $f \in C^\infty(\mathfrak{g}^*)$, the Hamiltonian vector field X_h^R must satisfy

$$\begin{aligned}X_h^R[f] &= \{f,h\}_R\\ &= \langle \alpha, [df_\alpha, dh_\alpha]_R\rangle\\ &= \langle \alpha, [Rdf_\alpha, dh_\alpha]\rangle + \langle \alpha, [df_\alpha, Rdh_\alpha]\rangle\\ &= \langle X_h^R(\alpha), df_\alpha\rangle\end{aligned}$$

Since h is a Casimir function, for all $X \in \mathfrak{g}$ we have

$$\langle \mathrm{ad}^*_X(\alpha)\ , dh_\alpha\rangle = -\langle \alpha\ , [X, dh_\alpha]\rangle = 0$$

Hence $\langle \alpha, [Rdf_\alpha, dh_\alpha]\rangle = 0$, and thus

$$\langle \alpha, [df_\alpha, dh_\alpha]_R\rangle = \langle \alpha, [df_\alpha, Rdh_\alpha]\rangle = \langle \mathrm{ad}^*_{R\cdot dh_\alpha}(\alpha), df_\alpha\rangle$$

from which we observe that $X_h^R(\alpha) = \mathrm{ad}^*_{Rdh_\alpha}(\alpha)$ as claimed. □

[1](Recall that via the natural isomorphism $(\mathfrak{g}^*)^* \sim \mathfrak{g}$, we identify df_α with an element of $\mathfrak{g}$).

Let us also observe that this proposition implies that any two Casimirs f and g with respect to the unmodified Lie-Poisson bracket $\{\cdot,\cdot\}$ Poisson commute with respect to the modified bracket $\{\cdot,\cdot\}_R$.

What relevance does all this have to the construction of integrable systems? By definition, any Casimir function H on $\mathfrak{g}^*$ is in the centre of the Poisson algebra $(\mathfrak{g}^*\,\{\cdot,\cdot\})$. Hence the Hamiltonian dynamical system on $(\mathfrak{g}^*\,\{\cdot,\cdot\})$ generated by taking H as the Hamiltonian will be trivial: the equation of motion in $\mathfrak{g}^*$ takes the form

$$\dot{\alpha} = \mathrm{ad}^*_{dH}(\alpha) = 0$$

and the values of all observables will be constant in time. However, these Casimir functions will not in general lie in the centre of the Poisson algebra $(\mathfrak{g}^*\,\{\cdot,\cdot\}_R)$. Crucially, we have seen that all these functions still Poisson commute with respect to the bracket $\{\cdot,\cdot\}_R$. Choosing such a function H as a Hamiltonian for a dynamical system on the Poisson manifold $(\mathfrak{g}^*\ ,\{\cdot,\cdot\}_R)$ yields the equation of motion

$$\dot{\alpha} = X^R_h(\alpha) = \mathrm{ad}^*_{R\cdot dH}(\alpha)$$

Restricting to a symplectic leaf Σ of the Poisson manifold $(\mathfrak{g}^*, \{\cdot,\cdot\}_R)$, we obtain a symplectic manifold whose Poisson bracket is simply the restriction of the Lie-Poisson bracket $\{\cdot,\cdot\}_R$ on $\mathfrak{g}^*$. Hence any Casimir function on the Poisson manifold $(\mathfrak{g}^*, \{\cdot,\cdot\})$ restricts to a function on the symplectic manifold Σ which Poisson commutes[2] with the restriction of any other such Casimir function. Choosing such a function as the Hamiltonian for a dynamical system on Σ, we obtain a non-trivial Hamiltonian system equipped with a set of commuting functions: the restrictions of the ring of Casimir invariant functions on $(\mathfrak{g}^*, \{\cdot,\cdot\})$. This suggests the possibility of constructing integrable systems in the sense of the Liouville-Arnold theorem.

In order make the connection with Lax equations, suppose that there exists a non-degenerate, adjoint invariant form on $\mathfrak{g}$. As described in Chapter 1, such a form establishes an isomorphism of $\mathfrak{g}$ with $\mathfrak{g}^*$ which allows us to identify a function H on $\mathfrak{g}^*$ with a function f on $\mathfrak{g}$. Under this isomorphism, the coadjoint action on $\mathfrak{g}^*$ is transformed to the adjoint action on $\mathfrak{g}$. Then with these identifications, the Hamiltonian flow of $f \sim H$ in $\mathfrak{g} \sim \mathfrak{g}^*$ above takes the Lax form:

$$\frac{dA}{dt} = -\mathrm{ad}_{RdH}(A) = [A, RdH] = [A, R\nabla f]$$

where $A \in \mathfrak{g}$, and ∇f is the gradient of f with respect to the form $\langle\cdot,\cdot\rangle$ as defined in Chapter 1.

3.2 The AKS construction

We now specialize to the case in which our Lie algebra admits a vector space decomposition into two subalgebras, so that we can write $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{b}$. There is a method, known as the Adler-Kostant-Symes (AKS) construction, which allows us to build Hamiltonian systems on the dual $\mathfrak{a}^*$ of the subalgebra $\mathfrak{a}$ which possess many conserved quantities in involution. The main idea of the construction is to use an embedding $\mathfrak{a}^* \hookrightarrow \mathfrak{g}^*$ to consider Hamiltonian systems on $\mathfrak{a}^*$ as restrictions of Hamiltonian systems on $\mathfrak{g}^*$.

We begin by fixing some notation. As before, let P_+ and P_- be the projection operators onto $\mathfrak{a}$ and $\mathfrak{b}$ respectively with respect to the decomposition $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{b}$. We will

[2]With respect to the restriction of $\{\cdot,\cdot\}_R$

again assume that we have a bilinear, symmetric, non-degenerate and adjoint-invariant form $\langle \cdot, \cdot \rangle$ on $\mathfrak{g}$, which allows us to identify $\mathfrak{g}$ with $\mathfrak{g}^*$ in the fashion described previously. We let $[\cdot, \cdot]$ denote the natural Lie bracket on $\mathfrak{g}$, and let $[\cdot, \cdot]_R$ denote the bracket arising from the R-matrix $R = \frac{1}{2}(P_+ - P_-)$. The corresponding Lie-Poisson brackets on $\mathfrak{g}^*$ are denoted $\{\cdot, \cdot\}$ and $\{\cdot, \cdot\}_R$ respectively. Now let H be a Casimir function on the Poisson manifold $(\mathfrak{g}^*, \{\cdot, \cdot\})$. As we have already noted, the invariant bilinear form allows us to view H as an adjoint-invariant function f on the Lie algebra $\mathfrak{g}$.

Let $\mathfrak{b}^0$ be the annihilator in $\mathfrak{g}^*$ of the subspace $\mathfrak{b} \subset \mathfrak{g}$, and let $\mathfrak{a}^0$ be the corresponding annihilator of $\mathfrak{a}$. We may identify $\mathfrak{g}^*$ with $\mathfrak{a}^* \oplus \mathfrak{b}^*$ by identifying $\mathfrak{a}^*$ with $\mathfrak{b}^0$, and identifying $\mathfrak{b}^*$ with $\mathfrak{a}^0$. The invariant bilinear form establishes an isomorphism of $\mathfrak{g}^*$ with $\mathfrak{g}$ in the usual fashion. Under this isomorphism we see that $\mathfrak{a}^* \sim \mathfrak{b}^0$ is identified with $\mathfrak{b}^\perp$, the orthogonal complement of $\mathfrak{b}$ with respect to the form, while $\mathfrak{b}^* \sim \mathfrak{a}^0$ is identified with $\mathfrak{a}^\perp$. With these identifications in place, we can write the ad* action of $\mathfrak{a}$ on $\mathfrak{a}^* \sim \mathfrak{b}^\perp$ in terms of the Lie bracket in $\mathfrak{g}$. For $X \in \mathfrak{a}$ and $Y \in \mathfrak{b}^\perp \sim \mathfrak{a}^*$, we have

$$(\mathrm{ad}^*_{\mathfrak{a}})_X (Y) = P_{\mathfrak{b}^\perp}[X, Y]$$

where $P_{\mathfrak{b}^\perp}$ denotes the orthogonal projection onto $\mathfrak{b}^\perp$ with respect to the decomposition $\mathfrak{g} = \mathfrak{a}^\perp \oplus \mathfrak{b}^\perp$.

Now let $\Gamma \subset \mathfrak{b}^\perp \sim \mathfrak{a}^*$ be a Poisson submanifold of $\mathfrak{a}^*$, and let $f \in C^\infty(\mathfrak{g})$ be the representative of a Casimir function H on $\mathfrak{g}^*$. Notice that the restriction of the form $\langle \cdot, \cdot \rangle$ to the subspace $\mathfrak{b}^\perp \subset \mathfrak{g}$ remains non-degenerate. Therefore if we restrict f and consider it as a function on the $\mathfrak{b}^\perp$, it has a gradient $\nabla_{\mathfrak{b}^\perp} f$, defined as the unique vector field on $\mathfrak{b}^\perp$ such that for all c in $\mathfrak{b}^\perp$

$$df(c) = \langle \nabla_{\mathfrak{b}^\perp} f, c \rangle$$

If we consider f as a function defined on all of $\mathfrak{g}$, however, its $\mathfrak{g}$-gradient is defined by

$$df(X) = \langle \nabla f, X \rangle = \langle (P_+ + P_-)\nabla f, X \rangle$$

If $X \in \mathfrak{b}^\perp$ comparing these two formulas yields

$$\nabla_{\mathfrak{b}^\perp} f(X) = P_+ \nabla f(X)$$

The standard Lie-Poisson bracket on $\mathfrak{a}^* \cong \mathfrak{b}^\perp$ is

$$\{f, g\}(c) = \langle c, [\nabla_{\mathfrak{b}^\perp} f, \nabla_{\mathfrak{b}^\perp} g] \rangle = \langle c, [P_+ \nabla f, P_+ \nabla g] \rangle$$

However, the restriction of the R-matrix bracket $\{\cdot, \cdot\}_R$ to $\mathfrak{b}^\perp$ is given by

$$\begin{aligned} \{f, g\}(c) = \langle c, [\nabla f, \nabla g]_R \rangle &= \langle c, [P_+ \nabla f, P_+ \nabla g] \rangle - \langle c, [P_- \nabla f, P_- \nabla g] \rangle \\ &= \langle c, [P_+ \nabla f, P_+ \nabla g] \rangle \end{aligned}$$

since $[P_- \nabla f, P_- \nabla g] \in \mathfrak{b}$. Hence the Lie-Poisson bracket on $\mathfrak{a}^*$ agrees with the restriction of the R-modified Poisson bracket to $\mathfrak{a}^* \subset \mathfrak{g}^*$. This allows us to bring the results about R-matrices from the previous section to bear on the situation. Recall that the R-matrix is given by $R = 1/2(P_+ - P_-)$. Observe that if H is a Casimir function on $\mathfrak{g}^*$, then

$$0 = \mathrm{ad}^*_{dH_\alpha}(\alpha) = \mathrm{ad}^*_{(dH_\alpha)+}(\alpha) + \mathrm{ad}^*_{(dH_\alpha)-}(\alpha)$$

and thus

$$\mathrm{ad}^*_{RdH_\alpha}(\alpha) = -\mathrm{ad}^*_{(dH_\alpha)-}(\alpha)$$

So if $H \in C^\infty(\mathfrak{a}^*)$ is the restriction of a Casimir for $\mathfrak{g}^*$, and $f \in C^\infty(\mathfrak{b}^\perp)$ is the corresponding function on $\mathfrak{b}^\perp \sim \mathfrak{a}^*$, then the Hamiltonian system defined by H on Γ has the equation of motion

$$\frac{dX}{dt} = [X, RdH] = [X, P_-\nabla f]$$

Observe that since $P_-\nabla f \in \mathfrak{b}$, $X \in \mathfrak{b}^\perp$ and the form is adjoint-invariant, we have $[c, P_-\nabla f] \in \mathfrak{b}^\perp$, so that the Hamiltonian flows remain within $\mathfrak{b}^\perp$. Moreover, we know that all restrictions of Casimir functions Poisson commute with respect to the bracket $\{\cdot,\cdot\}_R$, and hence with respect to the natural Lie-Poisson bracket on $\mathfrak{a}^*$. In summary, we have proved:

Theorem 3.3 ***(AKS Theorem)*** *Let $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{b}$ be a Lie algebra which is the vector space direct sum of two Lie subalgebras. Assume that $\mathfrak{g}$ is endowed with an adjoint-invariant, non-degenerate symmetric bilinear form, so that $\mathfrak{g} = \mathfrak{a}^\perp \oplus \mathfrak{b}^\perp$, and $\mathfrak{b}^\perp$ is identified with $\mathfrak{a}^*$. Let $\Gamma \subset \mathfrak{b}^\perp$ be a Poisson subspace of $\mathfrak{a}^* \sim \mathfrak{b}^\perp$ with respect to the Lie-Poisson structure on $\mathfrak{a}^*$, and let H be the restriction to Γ of a Casimir function for the Lie-Poisson bracket on $\mathfrak{g}$. Identifying H with a function f on $\mathfrak{g}$, the differential equation for the Hamiltonian flow of H is*

$$\frac{dX}{dt} = [X, (\nabla f)_-]$$

where the subscript denotes the orthogonal projection onto $\mathfrak{b}$ with respect to the decomposition $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{b}$. Additionally, the restrictions of all Casimir functions on $\mathfrak{g}^*$ to Γ pairwise Poisson-commute with respect to the Lie-Poisson bracket on $\mathfrak{a}^*$.

We conclude our discussion of the AKS construction by mentioning a slight generalisation of the statement above. Let $\mathfrak{a}$ be a Lie algebra. An element $\alpha \in \mathfrak{a}^*$ is called an **infinitesimal character** of $\mathfrak{a}^*$ if $\mathrm{ad}^*_X(\alpha) = 0$ for all $X \in \mathfrak{a}$.

Now suppose as before that we have a Lie algebra admitting a decomposition $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{b}$. An element $Y \in \mathfrak{a}^* \sim \mathfrak{b}^\perp$ will be an infinitesimal character of $\mathfrak{a}$ if $P_{\mathfrak{b}^\perp}[X, Y] = 0$ for all $X \in \mathfrak{a}$. If Y is an infinitesimal character of $\mathfrak{a}$ and $H : \mathfrak{g}^* \to \mathbb{R}$ is a Casimir function on $\mathfrak{g}^*$, we define the **shifted Hamiltonian** H_Y by

$$H_Y(\alpha) = H(\alpha + Y)$$

Observe that $(dH_Y)_\alpha = (dH)_{\alpha+Y}$. So by the AKS theorem, we see that if f, g are two Casimir functions on $\mathfrak{g}^*$, then the restrictions to $\mathfrak{a}^*$ of the shifted Hamiltonians f_Y, g_Y satisfy $\{f_Y, g_Y\}_{\mathfrak{a}^*} = 0$, where $\{\cdot,\cdot\}_{\mathfrak{a}^*}$ is the Lie-Poisson bracket in $\mathfrak{a}^*$. Moreover, the differential equation for the Hamiltonian flow corresponding to the shifted Hamiltonian f_Y is

$$\frac{dX}{dt} = [X + Y, (d_{X+Y} f)_-]$$

where $X \in \mathfrak{b}^\perp \sim \mathfrak{a}^*$ and $(d_{X+Y} f)_- \in \mathfrak{b}$.

3.3 Construction of Lax equations with a spectral parameter

We will now show how the AKS formalism leads naturally to Hamiltonian systems whose dynamics are described by Lax equations with a spectral parameter. Let $\tilde{\mathfrak{gl}}_r =$

$\mathfrak{gl}_r(\mathbb{C}) \otimes \mathbb{C}[\lambda, \lambda^{-1}]]$ be the loop algebra of semi-infinite formal Laurent series in λ with coefficients in $\mathfrak{gl}_r(\mathbb{C})$. An element $X(\lambda) \in \tilde{\mathfrak{gl}}_r$ has the form

$$X(\lambda) = \sum_{i=-\infty}^{m} X_i \lambda^i$$

If $Y(\lambda) = \sum_{j=-\infty}^{l} Y_j \lambda^j$ the bracket $[X(\lambda), Y(\lambda)]$ is defined by

$$[X(\lambda), Y(\lambda)] := \sum_{k=-\infty}^{m+l} \sum_{i+j=k} [X_i, Y_j] \lambda^k$$

The inner sums are all finite, while the outer one is formal. The algebra $\tilde{\mathfrak{gl}}_r$ has a non-degenerate, ad-invariant bilinear form defined by

$$\langle X(\lambda), Y(\lambda) \rangle := \operatorname{tr} \left(X(\lambda) Y(\lambda) \right)_0$$

where $(X(\lambda))_0$ is the constant term in the series $X(\lambda)$. Observe that if $X(\lambda) = \sum_{i=-\infty}^{m} X_i \lambda^i$,

$$\langle X(\lambda), \lambda^k \mathbb{1} \rangle = a_{-k}$$

In order to apply the AKS theorem, we consider the vector space direct sum decomposition

$$\tilde{\mathfrak{gl}}_r = \tilde{\mathfrak{gl}}_r^+ \oplus \tilde{\mathfrak{gl}}_r^-$$

where $\tilde{\mathfrak{gl}}_r^+$ is the subalgebra consisting of polynomials in λ and $\tilde{\mathfrak{gl}}_r^-$ the subalgebra of series with strictly negative powers of λ. The bilinear form gives identifications

$$\left(\tilde{\mathfrak{gl}}_r^+ \right)^* \cong \left(\tilde{\mathfrak{gl}}_r \right)^\perp = \left(\tilde{\mathfrak{gl}}_r \right)_0$$

where $\left(\tilde{\mathfrak{gl}}_r \right)_0 = \lambda \tilde{\mathfrak{gl}}_r^-$. In order to find some candidates for Hamiltonians, we must identify some Casimir functions on $\left(\tilde{\mathfrak{gl}}_r \right)^*$.

Proposition 3.4 ***(AKS Hamiltonians)*** *The functions ϕ_{jk} on $\tilde{\mathfrak{gl}}_r^*$ defined by*

$$\phi_{jk}(X(\lambda)) = \frac{1}{k} \operatorname{tr} \left(\lambda^j X^k(\lambda) \right)_0$$

are Casimirs.

Proof. We begin by observing that

$$d\phi_{jk}(X(\lambda)) = \lambda^j X^{k-1}(\lambda)$$

since given a curve $X(\lambda)(t)$ with $X(\lambda)(0) = X(\lambda), X(\lambda)'(0) = Y(\lambda)$, in view of the fact that the trace is cyclic we have

$$\begin{aligned} \left. \frac{d}{dt} \phi_{jk}(X(\lambda)(t)) \right|_{t=0} &= \left. \frac{1}{k} \frac{d}{dt} \operatorname{tr} \left(\lambda^j X^k(\lambda)(t) \right)_0 \right|_{t=0} \\ &= \operatorname{tr} \left((\lambda^j X^{k-1}(\lambda) Y(\lambda) \right)_0 \\ &= \langle \lambda^j X^{k-1}(\lambda), Y(\lambda) \rangle \end{aligned}$$

Hence for all $X(\lambda) \in \left(\tilde{\mathfrak{gl}}_r^*\right)$ and all $Y(\lambda) \in \tilde{\mathfrak{gl}}_r$, we have

$$\langle X(\lambda), [d\phi_{jk}, Y(\lambda)]\rangle = \operatorname{tr}\left(\lambda^j (X^k(\lambda)Y(\lambda)\right)_0 - \operatorname{tr}\left(\lambda^j X(\lambda)Y(\lambda)X^{k-1}(\lambda)\right)_0 = 0$$

again since the trace is cyclic. We conclude that the ϕ_{jk} are coadjoint invariant and thus Casimirs. □

The final step in applying the AKS theorem is to identify some suitable finite-dimensional Poisson subspaces of $\left(\tilde{\mathfrak{gl}}_r^+\right)^*$. To this end, we fix n distinct complex numbers $\{\alpha_1, \dots \alpha_n\}$. Define a polynomial $a(\lambda) \in \mathbb{C}[\lambda]$ by

$$a(\lambda) = \prod_{i=1}^{n} (\lambda - \alpha_i)$$

Since $X(\lambda) \in \tilde{\mathfrak{gl}}_r^+$ is a matrix polynomial in λ, evaluation at $\lambda = \alpha_i$ gives an element $X(\alpha_i) \in \mathfrak{gl}_r(\mathbb{C})$. For $n \leq N$, write

$$G_r^n := GL_r(\mathbb{C}) \times \dots \times GL_r(\mathbb{C}) \qquad (n \text{ times})$$

Denote by

$$\mathfrak{g}_r^n = \mathfrak{gl}_r(\mathbb{C}) \times \dots \mathfrak{gl}^r(\mathbb{C}) \qquad (n \text{ times})$$

its Lie algebra. We define a surjective Lie algebra homomorphism

$$\mathcal{A} : \left(\tilde{\mathfrak{gl}}_r^+\right) \to \mathfrak{g}_r^n$$

$$X(\lambda) \mapsto (X(\alpha_1), \dots, X(\alpha_n))$$

Since $\mathcal{A}$ is surjective, the dual map $\mathcal{A}^* : (\mathfrak{g}_r^n)^* \to \left(\tilde{\mathfrak{gl}}_r^+\right)^*$ is injective. If we identify $(\mathfrak{g}_r^n)^* \cong \mathfrak{g}_r^n$ by taking traces componentwise, and identify $\left(\tilde{\mathfrak{gl}}_r^+\right)^* \cong \left(\tilde{\mathfrak{gl}}_r\right)_0$, we have

$$\mathcal{A}^*(Y_1, \dots Y_n) = \sum_{k=0}^{\infty} \left(\sum_{i=1}^{n} Y_i \alpha_i^k \right) \lambda^{-k} \tag{3.1}$$

Because of the geometric series identity

$$(\lambda - \alpha_i) \sum_{k=0}^{\infty} Y_i \alpha_i^k \lambda^{-k} = \lambda Y_i$$

we will use the notation

$$\sum_{k=0}^{\infty} \left(\sum_{i=1}^{n} Y_i \alpha_i^k \right) \lambda^{-k} = \lambda \sum_{i=1}^{n} \frac{Y_i}{\lambda - \alpha_i}$$

We now argue that the map $\mathcal{A}^*$ is in fact a Poisson map. This follows from the following more general proposition:

Lemma 3.5 *If $\phi : \mathfrak{g} \to \mathfrak{h}$ is a Lie algebra homomorphism, the dual map $\phi^* : \mathfrak{h}^* \to \mathfrak{g}^*$ is a Poisson map with respect to the standard Lie-Poisson structures on the Poisson manifolds $\mathfrak{g}^*$ and $\mathfrak{h}^*$.*

Proof. Given functions F, G defined on $\mathfrak{g}^*$, we have

$$\{F \circ \phi^*, G \circ \phi^*\}(\mu) = \langle \mu, [d(F \circ \phi^*)_\mu, d(G \circ \phi^*)_\mu] \rangle$$

However since ϕ is a linear map, we have

$$d(F \circ \phi^*)_\mu = \phi \cdot dF_{\phi(\mu)}$$

Thus,

$$\begin{aligned} \{F \circ \phi^*, G \circ \phi^*\}(\mu) &= \langle \mu, [\phi \cdot dF_{\phi(\mu)}, \phi \cdot dG_{\phi(\mu)}] \rangle \\ &= \langle \mu, \phi \cdot [dF_{\phi(\mu)}, dG_{\phi(\mu)}] = \langle \phi^* \mu, [dF_{\phi(\mu)}, dG_{\phi(\mu)}] \rangle \\ &= \{F, G\} \circ \phi^* \end{aligned}$$

and we conclude that ϕ^* is indeed Poisson. □

So in view of the preceding proposition, the fact that $\mathcal{A}$ is a homomorphism of Lie algebras implies that the dual map $\mathcal{A}^*$ is a Poisson map with respect to the standard Lie-Poisson structures on the Poisson manifolds $(\mathfrak{g}_r^n)^*$ and $\left(\tilde{\mathfrak{gl}}_r^+\right)^*$. Moreover, it is evident from the explicit formula for $\mathcal{A}^*$ that its image is $\mathrm{ad}_{\tilde{\mathfrak{gl}}_r^+}$-invariant, and thus a Poisson subspace of $\left(\tilde{\mathfrak{gl}}_r^+\right)^*$:

Proposition 3.6 *The Poisson map $\mathcal{A}^*$ is an isomorphism of $(\mathfrak{g}_r^n)^*$ with the finite-dimensional Poisson subspace*

$$\left(\mathfrak{gl}_r^+\right)_A^* = \left\{ \sum_{i=1}^n \frac{\lambda X_i}{\lambda - \alpha_i} : X_i \subset \mathfrak{gl}_r \right\}$$

It follows that the restriction of $\mathcal{A}^*$ to a symplectic leaf of $(\mathfrak{g}_r^n)^*$ is a symplectomorphism onto a symplectic leaf in $\left(\mathfrak{gl}_r^+\right)^*$. But the symplectic leaves of $(\mathfrak{g}_r^n)^*$ are simply the orbits under coadjoint action of G_r^n. We naturally identify such an orbit as the product of n $GL_r(\mathbb{C})$ coadjoint orbits. So we have realised our goal of describing some finite-dimensional symplectic leaves of $\left(\mathfrak{gl}_r^+\right)^*$. Explicitly, if

$$\mathcal{N}(\lambda) = \sum_{i=1}^n \frac{\lambda X_i}{\lambda - \alpha i}$$

then the symplectic leaf $\mathcal{O}_{\mathcal{N}(\lambda)}$ through $\mathcal{N}(\lambda)$ is given by

$$\mathcal{O}_{\mathcal{N}(\lambda)} = \left\{ \sum_{i=1}^n \frac{\lambda g_i X_i g_i^{-1}}{\lambda - \alpha_i} : g_i \in GL_r(\mathbb{C}) \right\}$$

Applying the AKS theorem, we obtain:

Theorem 3.7 *Let $\mathcal{F}$ denote the restriction of the ring of Casimir functions on $\tilde{\mathfrak{gl}}_r^*$ to the symplectic leaf $\mathcal{O}_{\mathcal{N}(\lambda)}$ through the point $\mathcal{N}(\lambda) \in \left(\mathfrak{gl}_r^+\right)^*$. Then the ring $\mathcal{F}$ Poisson commutes with respect to the Poisson structure on $\mathcal{O}_{\mathcal{N}(\lambda)}$, and if $\phi \in \mathcal{F}$, the Hamiltonian flow for ϕ on $\mathcal{O}_{\mathcal{N}(\lambda)}$ through the point $\mathcal{N}(\lambda)$ is given by*

$$\frac{d}{dt}\mathcal{N}(\lambda; t) = [(d\phi(\mathcal{N}(\lambda; t))_+, \mathcal{N}(\lambda; t)]$$

Note that although the intermediate stages of this construction occur in an infinite-dimensional loop algebra, the resulting Hamiltonian dynamical systems obtained by the restriction described above are defined on finite-dimensional symplectic manifolds isomorphic to coadjoint orbits.

The restriction to a finite-dimensional Poisson submanifold of $\left(\mathfrak{gl}_r^+\right)^*$ also allows us to recast the Hamiltonian flow above into a convenient form involving matrix polynomials. Specifically, if $\mathcal{N}(\lambda) \in \mathcal{O}_{\mathcal{M}(\lambda)}$, then we can equally well think of $\mathcal{N}(\lambda)$ as a matrix polynomial $A(\lambda)$ of degree $d = n - 1$, by writing

$$A(\lambda) = \lambda^{-1} a(\lambda) \mathcal{N}(\lambda)$$

Thus the Hamiltonian flow corresponding to an AKS Hamiltonian $\phi \in \mathcal{F}$ is equivalent to the following matrix polynomial Lax equation:

$$\frac{d}{dt} A(\lambda; t) = \left[(d\phi(\mathcal{N}(\lambda; t))_+ \, , A(\lambda; t) \right] \tag{3.2}$$

The following proposition gives an alternative characterisation of these equations of motion:

Proposition 3.8 *Let $A(\lambda; t)$ be the solution of a Lax equation of the form*

$$\frac{d}{dt} A(\lambda; t) = \left[\left(P(A(\lambda; t), \lambda^{-1}) \right)_+ \, , A(\lambda; t) \right] \tag{3.3}$$

where $P(x, \lambda^{-1})$ is a formal polynomial in the variables x and λ^{-1}. Write

$$\mathcal{N}(\lambda; t) = \frac{\lambda}{a(\lambda)} A(\lambda; t)$$

Then there exists a coadjoint invariant function $\phi \in \mathcal{F}$ such that $\mathcal{N}(\lambda; t)$ describes the Hamiltonian flow for ϕ on $\left(\mathfrak{gl}_r^+\right)^*$.

Proof. Observe from considerations of linearity that it suffices to prove the proposition when P is a monomial $P(A(\lambda), \lambda^{-1}) = \lambda^{-j} \left(A(\lambda)\right)^k$. In this case, equation (3.3) implies

$$\frac{d\mathcal{N}(\lambda; t)}{dt} = \left[\left(\lambda^{-j} \left(\lambda^{-1} a(\lambda) \mathcal{N}(\lambda) \right)^k \right)_+ , \mathcal{N}(\lambda; t) \right]$$

Therefore we must find a Casimir function $\phi \in \mathcal{F}$ such that

$$d\phi(X(\lambda)) = \lambda^{-j-k} a^k(\lambda) X^k(\lambda)$$

But recall that the differentials of the AKS Hamiltonians ϕ_{mn} defined in Proposition 3.4 satisfy

$$d\phi_{mn}(X(\lambda)) = \lambda^m X^{n-1}(\lambda)$$

Hence it follows that the desired Hamiltonian ϕ can be expressed as a linear combination of the functions ϕ_{mn}, which proves the proposition. □

In summary, applying the AKS construction in a loop algebra yields Hamiltonian systems defined on finite-dimensional coadjoint orbits which possess a non-trivial ring of Poisson-commuting functions. Moreover, Hamilton's equations for these systems can be cast in the characteristic Lax form (3.3).

Example 3.9 (The Neumann model) We show how the methods of this section can be used to interpret the Neumann model as describing Hamiltonian flow on the dual of the positive part of the loop algebra $\tilde{\mathfrak{sl}}_2(\mathbb{R})$. Recall that the unreduced Neumann phase space is the symplectic manifold $(\mathbb{R}^{2n}, \omega)$, where $\omega = d\mathbf{x} \wedge d\mathbf{y}$ is the canonical symplectic form on $T^*\mathbb{R}^n \cong \mathbb{R}^{2n}$ Following the approach in [26], we will use a moment map to embed this phase space in $\tilde{\mathfrak{sl}}_2(\mathbb{R})^*$. It will be useful identify points $(\mathbf{x}, \mathbf{y})$ of $\mathbb{R}^{2n}$ with n-tuples of row vectors, i.e.

$$(\mathbf{x}, \mathbf{y}) \sim (v_1, \ldots, v_n) = ((x_1\ y_1), \ldots (x_n\ y_n))$$

Define a matrix J by

$$J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

If $\xi = (\xi_1, \ldots, \xi_n)$ and $\eta = (\eta_1, \ldots, \eta_n)$ are tangent vectors, then we can express the symplectic form ω as

$$\omega(\xi, \eta) = \frac{1}{2} \sum_{i=1}^{n} \left[\xi_i J \eta_i^T + \eta_i J^T \xi_i^T\right]$$

Define a group G by

$$G := SL_2(\mathbb{R}) \times \ldots \times SL_2(\mathbb{R}) \qquad (n \text{ times})$$

Denote by $\mathfrak{sl}_2^n$ the Lie algebra of G, noting that

$$\mathfrak{sl}_2^n \cong \mathfrak{sl}_2(\mathbb{R}) \times \cdots \times \mathfrak{sl}_2(\mathbb{R}) \qquad (n \text{ times})$$

The group G acts on $\mathbb{R}^{2n}$ by right multiplication: if $\mathbf{g} = (g_1 \ldots, g_n) \in G$,

$$\mathbf{g} \cdot (v_1, \ldots, v_n) = \left(v_1 g_1^{-1}, \ldots, v_n g_1^{-1}\right)$$

This action is Poisson, with equivariant moment map

$$I : \mathbb{R}^{2n} \to (\mathfrak{sl}_r^n)^*$$

$$I(v_1, \ldots, v_n)(X_1, \ldots X_n) = -\sum_{i=1}^{n} \operatorname{tr}(v_i X_i J v_i^T)$$

for $(X_1, \ldots, X_n) \in \mathfrak{sl}_2^n$. If we identify $\mathfrak{sl}_2^n \sim (\mathfrak{sl}_2^n)^*$ by taking traces componentwise, we can write this as

$$I(v_1, \ldots, v_n) = \left(-J v_1^T v_1, \ldots, -J v_n^T v_n\right)$$

Now I is a Poisson map, since equivariant moment maps are Poisson with respect to the Lie-Poisson structure. Let

$$\mathcal{A}^* : (\mathfrak{sl}_2^n)^* \to \left(\tilde{\mathfrak{sl}}_2^+\right)^*$$

$$(Y_1, \ldots Y_n) \mapsto \lambda \sum_{i=1}^{n} \frac{Y_i}{\lambda - \alpha_i}$$

be the Poisson map defined in equation 3.1. Then the composite $\sigma = \mathcal{A}^* \circ I$ defines a Poisson map from $\mathbb{R}^{2n}$ to $\left(\tilde{\mathfrak{sl}}_2(\mathbb{R})\right)^*$. Explicitly, we have

$$\sigma(\mathbf{x}, \mathbf{y}) = \lambda \sum_{i=1}^{n} \begin{pmatrix} \frac{-x_i y_i}{\lambda - \alpha_i} & \frac{-y_i^2}{\lambda - \alpha_i} \\ \frac{x_i^2}{\lambda - \alpha_i} & \frac{x_i y_i}{\lambda - \alpha_i} \end{pmatrix}$$

Now observe that for any $Y \in \mathfrak{sl}_2$, the element $\lambda Y \in \left(\tilde{\mathfrak{sl}}_r\right)_0 \sim \left(\tilde{\mathfrak{sl}}_2\right)^*$ is an infinitesimal character of $\left(\tilde{\mathfrak{sl}}_2\right)^*$. So taking

$$Y = \begin{pmatrix} 0 & -\frac{1}{2} \\ 0 & 0 \end{pmatrix}$$

we consider the Hamiltonian flow on $\left(\tilde{\mathfrak{sl}}_2^+\right)^*$ generated by the shifted Hamiltonian

$$\phi(X(\lambda)) = \frac{1}{2}\mathrm{tr}\left((\lambda Y + X(\lambda))^2\right)_0$$

The pullback by σ of ϕ yields the following Hamiltonian H on $\mathbb{R}^{2n}$:

$$H(\mathbf{x}, \mathbf{y}) = \frac{1}{2}[(\mathbf{x}^T\mathbf{x})(\mathbf{y}^T\mathbf{y}) + \mathbf{x}^T A\mathbf{x} - (\mathbf{x}^T\mathbf{y})^2]$$

and we recognise this as the Neumann Hamiltonian of Section 2.2.

Although σ is not injective, its fibres are simply the orbits of a finite group of reflections generated by

$$(x_i, y_i) \mapsto (-x_i, -y_i)$$

Thus, up to a quotient by this group[3], the Neumann model is equivalent to the Hamiltonian flow on $\left(\tilde{\mathfrak{sl}}_2^+\right)^*$ generated by the Hamiltonian ϕ. The Hamiltonian flow on $\left(\tilde{\mathfrak{sl}}_2^+\right)^*$ is described by the Lax equation

$$\frac{d}{dt}\mathcal{N}(\lambda) = [\mathcal{N}(\lambda), B(\lambda)] \tag{3.4}$$

where

$$B(\lambda) = (d\phi(\mathcal{N}))_+ \tag{3.5}$$

Expressed in terms of the Neumann dynamical variables via the Poisson map σ, the matrices $\mathcal{N}(\lambda)$ and $B(\lambda)$ are given by

$$\mathcal{N}(\lambda) = \lambda \sum_{i=1}^{n} \begin{pmatrix} \frac{-x_i y_i}{\lambda - \alpha_i} & -\frac{1}{2} - \frac{-y_i^2}{\lambda - \alpha_i} \\ \frac{x_i^2}{\lambda - \alpha_i} & \frac{x_i y_i}{\lambda - \alpha_i} \end{pmatrix}$$

$$B(\lambda) = \begin{pmatrix} \mathbf{x}^T\mathbf{x} & \lambda + \mathbf{y}^T\mathbf{y} \\ -\mathbf{x}^T\mathbf{x} & -\mathbf{x}^T\mathbf{x} \end{pmatrix}$$

Writing $a(\lambda) = \prod_{i=1}^{n}(\lambda - \alpha_i)$, define a matrix polynomial $A(\lambda)$ by

$$A(\lambda) = \frac{-2a(\lambda)}{\lambda}\mathcal{N}(\lambda)$$

[3]Indeed, the ambiguity can be resolved along the flows by continuity

Then the Hamiltonian flow (3.6) is equivalent to the Lax equation

$$\frac{dA(\lambda)}{dt} = [A(\lambda), B(\lambda)]$$

Observe that the matrix $B(\lambda)$ can be expressed

$$B(\lambda) = P(x, \lambda^{-1})_+$$

for the polynomial $P(x, \lambda^{-1}) = x\lambda^{1-n}$.
Let us concluding by noting that we can now give an elegant proof of the fact that the Devaney-Uhlenbeck functions H_k Poisson commute. Indeed, observe that we can write

$$H_k = \mathrm{Res}_{\lambda=0}\left(\frac{a(\lambda)\mathrm{det}\mathcal{N}(\lambda)}{\lambda^{k+1}}\right)$$

Since we can express the determinant $\mathrm{det}\mathcal{N}(\lambda)$ as a sum of traces of various powers of $\mathrm{N}(\lambda)$, it follows that each H_k is a linear combination of the AKS conserved quantities ϕ_{jk} defined in Proposition 3.4. In the language of Section 2.1, the determinant $\det\mathcal{N}(\lambda)$ nothing but the rational function $\Delta(\lambda)$ used to define the Devaney-Uhlenbeck invariants.

Chapter 4

Riemann Surfaces

In this chapter, we present the elements of the theory of Riemann surfaces that we will use to study Lax equations with a spectral parameter. Just as in the case of Hamiltonian mechanics, this is an enormous subject that has been studied in the literature from almost every angle imaginable[1]. Our treatment is necessarily cursory, and we shall omit the proofs of most results. Of the many excellent references on Riemann surfaces, the author particularly recommends Gunning's three volumes [14],[16],[15] as well as Miranda's book [25].

4.1 Basic Definitions

Suppose that M is a 2-dimensional topological manifold with an atlas of coordinate charts $\mathcal{A} = \{U_\alpha, z_\alpha\}$, where each map z_α is a bijection of an (open) subset $U_\alpha \subset M$ with an open subset $z_\alpha(U_\alpha) \subset \mathbb{R}^2$. Making the identification of $\mathbb{R}^2$ with the complex line $\mathbb{C}$, the coordinate transition functions

$$f_{\alpha\beta} = z_\alpha \circ z_\beta^{-1} : z_\beta(U_\alpha \cap U_\beta) \subset \mathbb{R}^2 \sim \mathbb{C} \to z_\alpha(U_\alpha \cap U_\beta) \subset \mathbb{C}$$

can be thought of as continuous complex-valued functions defined on various open subsets of $\mathbb{C}$.

Definition 4.1 ***(Riemann surface)*** *Let M be a 2-dimensional smooth manifold with an atlas of coordinate charts $\mathcal{A} = \{U_\alpha, z_\alpha\}$. The atlas $\mathcal{A}$ is called a complex analytic atlas for M if the coordinate transition functions $z_\alpha \circ z_\beta^{-1}$ are all holomorphic functions. Two complex analytic atlases on M are equivalent if their union is also a complex analytic atlas, and an equivalence class of complex analytic atlas is called a complex analytic structure on M. The pair $(M, \mathcal{A})$ where $\mathcal{A}$ is a complex analytic structure is called a Riemann surface.*

Remark: In a similar spirit, we can identify the target spaces of the coordinate charts on a smooth manifold M of dimension $2n$ with the complex space $\mathbb{C}^n$, leading to an analogous definition of an n-dimensional complex manifold.

[1] Note the title of Jost's book [20]

Simple examples of Riemann surfaces are given by the complex line $\mathbb{C}$ and its open subsets. For us, however, the compact Riemann surfaces will be the main objects of interest. Notice that a Riemann surface is always orientable: an orientation is induced by the holomorphic transition functions in its complex analytic atlas. So by the classification of surfaces, it follows that any compact and connected Riemann surface is homeomorphic to a surface of genus g. The simplest example of a compact Riemann surface is the Riemann sphere $\mathbb{P}^1$, which is the 2-sphere endowed with a complex analytic structure. This complex analytic structure be described as follows: take the unit sphere $S^2 \subset \mathbb{R}^3$, and project stereographically from the north pole N onto the sphere's equatorial plane, which we identify with the complex line $\mathbb{C}$. This defines a bijection $z_0 : S^2 - \{N\} \to \mathbb{R}^2 \sim \mathbb{C}$. Similarly, projecting from the south pole onto the equatorial plane yields $z_1 : S^2 - \{S\} \to \mathbb{C}$. Examining the geometry of the situation, we check that the transition map $z_1 \circ z_0^{-1} :\to \mathbb{C} - \{0\} \to \mathbb{C} - \{0\}$ is given by $z \mapsto \frac{1}{z}$ and is thus holomorphic, making S^2 into a compact Riemann surface called the **Riemann sphere**. Since every point except for the north pole N is contained in the chart z_0, it is useful to think of the Riemann sphere as the complex line of the variable z_0 together with an extra point, ∞. We will use the notation $\mathbb{P}^1$ for the Riemann sphere. The reason for this is that one obtains the same Riemann surface by considering the complex projective space $\mathbb{CP}^1$ of lines in $\mathbb{C}^2$. As a set, $\mathbb{CP}^1$ can be described as the quotient space of $\mathbb{C}^2 - \{0\} = \{[v_1, v_2] \neq (0,0) \in \mathbb{C}^2\}$ under the relation

$$[v_1, v_2] \sim [\lambda v_1, \lambda v_2], \lambda \in \mathbb{C}^*$$

where $\mathbb{C}^* = \mathbb{C} - \{0\}$. Thus $\mathbb{CP}^1$ can be covered by the sets

$$U_0 = \{[v_1, v_2] : v_1 \neq 0\}, \qquad U_1 = \{[v_1, v_2] : v_2 \neq 0\}$$

and the maps

$$z_0 : U_0 \to \mathbb{C}, \quad [v_1, v_2] \to \frac{v_2}{v_1}$$
$$z_1 : U_1 \to \mathbb{C}, \quad [v_1, v_2] \to \frac{v_1}{v_2}$$

define bijections of U_1 and U_2 respectively with $\mathbb{C}$. Evidently, the transition map $z_1 \circ z_0^{-1} : \mathbb{C} - \{0\} \to \mathbb{C} - \{0\}$ is given by $z \mapsto \frac{1}{z}$, and so we see that $\mathbb{CP}^1$ and the Riemann sphere share the same complex analytic structure.

We now consider some further examples of Riemann surfaces:

Example 4.2 (Graphs of holomorphic functions) If $V \subset \mathbb{C}$ is an open connected subset of the complex line and f is a holomorphic function defined on V, then the graph

$$\Gamma = \{(z, f(z)) \subset \mathbb{C}^2 | z \in V\}$$

has the structure of a Riemann surface: it has a complex analytic atlas composed of a single chart, the whole set Γ, together with the projection map $\pi : \Gamma \to V, (z, f(z)) \mapsto z$.

Example 4.3 (Smooth affine plane curves) Suppose that $f \in \mathbb{C}[z, w]$ is a polynomial, and let

$$X = \{(z, w) \in \mathbb{C}^2 | f(z, w) = 0\}$$

be the zero set of f. The set $X \subset \mathbb{C}^2$ is called an *affine plane curve*. In order to give X the structure of a Riemann surface, it will be necessary to impose an extra condition on the polynomial f, which is related to the **holomorphic implicit function theorem**:

Theorem 4.4 ***(Holomorphic Implicit Function Theorem)*** *Let $f(z,w) \in \mathbb{C}^2$ be a polynomial, let $X = \{(z,w) \in \mathbb{C}^2 | f(z,w) = 0\}$ be its zero set, and let p be a point of X. Then if $\frac{\partial f}{\partial w}(p) \neq 0$ there exists a holomorphic function $g(z)$ defined in an open neighbourhood of z_0 such that, near p, X is equal to the graph $\{w = g(z)\} \subset \mathbb{C}^2$.*

For a proof of the holomorphic implicit function theorem, see [13]. In view of this result, we say that the algebraic plane curve $X = \{(z,w) \in \mathbb{C}^2 | f(z,w) = 0\}$ is **smooth** if at all points $p \in X$, at least one of the partial derivatives $\frac{\partial f}{\partial w}(p)$ and $\frac{\partial f}{\partial z}(p)$ is non-zero. By the holomorphic implicit function theorem, it follows that each point of X has an open neighbourhood which is the graph of a holomorphic function, and we can take the union of these neighbourhoods to obtain a complex analytic atlas on X. Thus smooth affine plane curves are examples of Riemann surfaces. We shall say that an affine plane curve is **irreducible** if its defining polynomial $f(z,w)$ is irreducible. It can be shown[2] that an irreducible affine plane curve is a connected subset of $\mathbb{C}^2$. Moreover, given irreducible affine plane curve X, there is a unique[3] compact Riemann surface M_X and an inclusion $X \hookrightarrow M_X$ such that X and M_X coincide except for a finite number of points: that is, there exist two finite subsets $E \subset X$ and $F \subset M_X$ with $X - E = M_X - F$. We say that the compact Riemann surface M_X is the **completion** of the affine plane curve X.

Example 4.5 (Complex tori) We now come to an example of complex analytic manifolds of dimension greater than 1. Any n-dimensional complex vector space $\mathbb{C}^n$ can also be considered as a $2n$-dimensional real vector space $\mathbb{R}^{2n}$. Recall from our discussion of the Liouville-Arnold theorem that a **lattice subgroup** of the real vector space $\mathbb{R}^{2n}$ is an additive subgroup $\mathbb{Z}^k \subset \mathbb{R}^{2n}$ generated by k vectors $e_1, \ldots, e_k$ which are linearly independent over $\mathbb{R}$. We define a lattice in a complex vector space $\mathbb{C}^n$ to be a lattice subgroup of the associated real vector space $\mathbb{R}^{2n}$. Now let $\Lambda \cong \mathbb{Z}^{2n}$ be a lattice in $\mathbb{C}^n$ of maximal dimension $2n$. Then as a smooth manifold, the quotient space $\mathbb{C}^n/\Lambda \cong \mathbb{R}^{2n}/\Lambda$ is diffeomorphic to a $2n$-dimensional torus T^{2n}. Moreover, we can use the orbit projection $\pi : \mathbb{R}^{2n} \sim \mathbb{C}^n \to T^{2n}$ to define a complex analytic structure on the quotient space T^{2n}: to obtain a complex analytic chart containing a point in T^{2n}, we simply restrict to sufficiently small open neighbourhood so that the restriction of the covering projection is a bijection.

Now if Λ and Λ' are two maximal lattices in $\mathbb{C}^n$, the tori $\mathbb{R}^{2n}/\Lambda$ and $\mathbb{R}^{2n}/\Lambda'$ are always diffeomorphic as smooth manifolds. In general, however, they are *not* isomorphic as complex manifolds. For an introduction to the surprisingly subtle geometry of complex tori, the reader is directed to [14].

[2] See [13]

[3] Up to holomorphic isomorphism

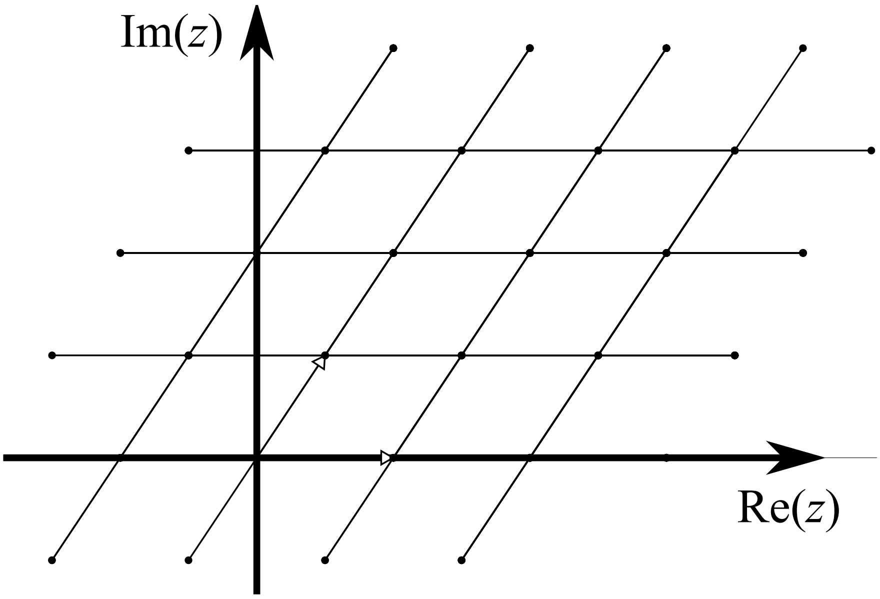

Figure 4.1: A maximal lattice in $\mathbb{C} \cong \mathbb{R}^2$.

A complex analytic structure on a 2-dimensional manifold M allows us to define the notion of a holomorphic function on M:

Definition 4.6 ***(Holomorphic functions)*** *Let M be a Riemann surface with complex analytic atlas $\{U_\alpha, z_\alpha\}$. If $V \subset M$ is an open subset of M, a map $f : M \to \mathbb{C}$ is called a holomorphic function on V if for each nonempty intersection $V \cap U_\alpha$ the function*

$$f \circ z_\alpha^{-1} : z_\alpha(V \cap U_\alpha) \subset \mathbb{C} \to \mathbb{C}$$

is a holomorphic. We say that a function $f : M \to \mathbb{C}$ is holomorphic at a point $p \in M$ if there exists an open set $V \subset M$ containing p such that f is holomorphic on V.

It turns out that the theory of holomorphic functions on compact Riemann surfaces is rather poor:

Proposition 4.7 *If M is a compact, connected Riemann surface, every holomorphic function defined on M is constant.*

Proof. Since M is compact, the continuous function $|f|$ attains its maximum at a point $p \in M$. So in a coordinate chart (U_α, z_α) containing p, the modulus of the holomorphic function $f \circ z_\alpha^{-1}$ has an interior maximum. By the maximum modulus principle, we see that f is locally constant on the neighborhood U_α. Then by the identity theorem together with the connectedness of M it follows that f must be constant on all of M. □

Thus the ring of holomorphic functions on a compact connected Riemann surface is an object of little interest. However since Riemann surfaces are locally isomorphic to the complex line $\mathbb{C}$, it makes sense to generalize the idea of holomorphic functions to holomorphic maps between Riemann surfaces:

Definition 4.8 ***(Holomorphic maps)*** *A map $f : \tilde{M} \to M$ between two Riemann surfaces $\tilde{M}$ and M is called holomorphic if for each point $p \in M$ there exist coordinate charts $\tilde{z}_\alpha : \tilde{U}_\alpha \to \mathbb{C}$ and $z_\beta : U_\beta \to \mathbb{C}$ such that $p \in \tilde{U}_\alpha$ and $f(p) \in U_\beta$, and the map $z_\beta \circ f \circ \tilde{z}_\alpha^{-1}$ is holomorphic in some open neighbourhood of $\tilde{z}_\alpha(p) \in \mathbb{C}$.*

Suppose that f is a meromorphic function defined on an open subset D of the complex line. At points where $f(z)$ has a pole, the function $\frac{1}{f(z)}$ is holomorphic. So thinking of the Riemann sphere as $\mathbb{P}^1 = \mathbb{C} \cup \infty$, we can define a holomorphic map $F : D \to \mathbb{P}^1$ by

$$F(z) = \begin{cases} f(z) & \text{if } z \text{ is not a pole of } f, \\ \infty & \text{if } z \text{ is a pole of } f. \end{cases}$$

Thus meromorphic functions on domains in $\mathbb{C}$ can be interpreted as holomorphic maps to $\mathbb{P}^1$. We can generalise this example to define meromorphic functions on arbitrary Riemann surfaces:

Definition 4.9 ***(Meromorphic functions)*** *A meromorphic function on a Riemann surface M is a holomorphic map $F : M \to \mathbb{P}^1$.*

We now present some basic properties of holomorphic maps between Riemann surfaces. Proofs can be found in [25]. The first of these is another illustration of the rigidity of holomorphic maps defined on compact Riemann surfaces:

Proposition 4.10 *Let M be a compact Riemann surface. If there exists a non-constant holomorphic map $F : M \to N$ from M to a Riemann surface N, then N must be compact, and F is surjective*

Together with the fact that the zeroes of a holomorphic function form a discrete set, this implies:

Proposition 4.11 *Let $F : M \to N$ be a non-constant holomorphic map between compact Riemann surfaces. Then for each point $q \in N$, the preimage $F^{-1}(q) \subset M$ is a non-empty finite set.*

Another useful result is the following:

Proposition 4.12 ***(Local normal form of a holomorphic map)*** *Let $F : M \to N$ be a non-constant holomorphic map between Riemann surfaces. Then for each point $p \in M$ there is a unique natural number $m \in \mathbb{N}$ such that the following property holds: for every chart $z_2 : U_2 \to \mathbb{C}$ on N centered at $F(p)$ there exists a chart $z_1 : U_1 \to \mathbb{C}$ on M centered at p such that*

$$z_2 \circ F \circ z_1(p) = z_1(p)^m$$

The proof, which is based on the holomorphic implicit function theorem, can be found on page 44 of [25]. The integer m_p that describes the normal form of $F : M \to N$ at $p \in M$ is called the **multiplicity** of F at p. A point $p \in M$ is called a **ramification point** of F if $m_p > 1$. The **ramification index** of F at p is the integer $e_p = m_p - 1$.

The image in N of a ramification point under F is called a **branch point** of F. The ramification points of F correspond to points where the the derivative of F vanishes in local coordinates. So by the results above, the ramification points and branch points of a holomorphic map between compact Riemann surfaces are finite sets. It is possible to show (see [25], p.47) that the sum of the multiplicities of a holomorphic map F over the point in a fibre of F is an invariant:

Proposition 4.13 *Let $F : M \to N$ be a nonconstant holomorphic map between compact Riemann surfaces. Then the integer*

$$d_q(F) = \sum_{p \in F^{-1}(q)} m_p(F)$$

is a constant independent of the point q, called the degree of the map F.

Example 4.14 (Affine plane curves) Consider a smooth affine plane curve $X = \{(z, w) \in \mathbb{C}^2 | f(z, w) = 0\}$ equipped with the holomorphic projection map $\pi : C \to \mathbb{C}, \pi(z, w) = z$. A point $p \in C$ is a ramification point of π if and only if $\frac{\partial f}{\partial w}(p) = 0$. Indeed, if $\frac{\partial f}{\partial w}(p) \neq 0$, then in a neighbourhood of p the map π defines a chart on C, and hence π is a bijection and has multiplicity one at p. Conversely, suppose that $\frac{\partial f}{\partial w}(p) = 0$. Since C is smooth, we must have $\frac{\partial f}{\partial z}(p) \neq 0$, so the coordinate w defines a chart on C in a neighbourhood of p. By the holomorphic implicit function theorem, in this neighbourhood C is the graph of a holomorphic function $g(w)$, and in this chart the map π is given by $w \mapsto g(w)$. Differentiating the relation $f(g(w), w) = 0$ with respect to w at p, we obtain

$$\frac{\partial f}{\partial z}(p) g'(w_0) + \frac{\partial f}{\partial w}(p) = 0$$

Since $\frac{\partial f}{\partial z}(p) \neq 0$ and $\frac{\partial f}{\partial w}(p) = 0$, it follows that $g'(w_0) = 0$, and hence π is ramified at p. Similarly, a point $p \in C$ is a ramification point of the map $(z, w) \mapsto w$ if and only if $\frac{\partial f}{\partial z}(p) = 0$.

Recall that a non-constant meromorphic function on a compact Riemann surface M is equivalent to a non-constant holomorphic map $F : M \to \mathbb{P}^1$. In view of Proposition 4.12, we say that such a map F of degree r exhibits M as an r-sheeted **branched covering** of $\mathbb{P}^1$. Knowledge of the ramification points of the map F allows us to compute the genus of M. Let Q denote the set of ramification points of F. Choose a triangulation[4] of $\mathbb{P}^1$ with v vertices, e edges and f faces, such that each branch point of F is a vertex. This is always possible since the branch points of F form a finite set. Since the Euler characteristic of the sphere is 2, we have $v - e + f = 2$. Now via the map F this triangulation lifts to a triangulation of M which also has rf faces and re edges. However since every ramification point F is a vertex in the lifted triangulation, the number of vertices in the lifted triangulation will be given by

$$v' = rv - \sum_{q \in Q} e_q$$

[4] A triangulation of M is a decomposition of M into disjoint closed subsets, each homeomorphic to a triangle, such that any two triangles are either disjoint, intersect only at single vertex, or intersect along a single edge. The existence of triangulations is a standard result; see [25]

where e_q is the ramification index of F at q. Now recall that the genus of a surface is related to its Euler characteristic χ by the formula

$$\chi = 2 - 2g$$

Since the genus of the Riemann sphere is 0, we conclude that the genus of M determined by the relation

$$2 - 2g = rf - re + rv - \sum_{q \in Q} e_q = 2r - \sum_{q \in Q} e_q$$

This is known as the **Riemann-Hurwitz** formula.

4.2 Complex Line Bundles

As we have emphasised, the requirement that a function on a compact Riemann surface be globally holomorphic is a rigid one. In order to obtain a greater variety of maps which still respect the complex analytic structure, we defined meromorphic functions and more generally holomorphic maps between Riemann surfaces. It turns out that we can make a slightly more subtle generalisation by allowing the target spaces of our functions to vary from point to point. This leads to the concept of a complex line bundle.

Definition 4.15 ***(Holomorphic line bundles)*** *A holomorphic line bundle over a Riemann surface M is a pair (L, π) where L is a complex manifold and π is a surjective holomorphic map $\pi : L \to M$ such that*

1. *For every point p on M, the fibre $\pi^{-1}(p)$ is a one-dimensional complex vector space*
2. *There exists an open covering $\{U_\alpha\}$ of M and homeomorphisms $\varphi_\alpha : \pi^{-1}(U_\alpha) \to U_\alpha \times \mathbb{C}$ such that $\text{proj}_{\mathbb{C}} \circ \varphi_\alpha = \pi$, where $\text{proj}_{\mathbb{C}}$ is the canonical projection $\text{proj}_{\mathbb{C}} : U_\alpha \times \mathbb{C} \to U_\alpha$, and,*
3. *The maps $\varphi_\beta \circ \varphi_\alpha^{-1} : U_\alpha \times \mathbb{C} \to U_\beta \times \mathbb{C}$ take the form $(p, v) \mapsto (p, f_{\beta\alpha}\ (p)v)$, where $f_{\beta\alpha}(p)$ is a non-vanishing holomorphic function defined on $U_\alpha \cap U_\beta$.*

We will also wish to consider bundles whose standard fibre is not the complex line $\mathbb{C}$, but rather some higher dimensional complex vector space:

Definition 4.16 ***(Holomorphic vector bundles)*** *A holomorphic vector bundle of rank n over a Riemann surface M is a pair (E, π) where complex manifold E and π is a surjective holomorphic map $\pi : V \to M$ such that*

1. *For every point p on M, the fibre $\pi^{-1}(p)$ has the structure of an n dimensional complex vector space, called the standard fibre.*
2. *There exists an open covering $\{U_\alpha\}$ of M and homeomorphisms $\varphi_\alpha : \pi^{-1}(U_\alpha) \to U_\alpha \times \mathbb{C}^n$ such that $pr \circ \varphi_\alpha = \pi$, where $\text{proj}_{\mathbb{C}}$ is the canonical projection $\text{proj}_{\mathbb{C}} : U_\alpha \times \mathbb{C}^n \to U_\alpha$, and,*
3. *The maps $\varphi_\beta \circ \varphi_\alpha^{-1} : U_\alpha \times \mathbb{C}^n \to U_\beta \times \mathbb{C}^n$ take the form $(p, v) \mapsto (p, f_{\beta\alpha}\ (p)v)$, where $f_{\beta\alpha} : U_\alpha \cap U_\beta \to GL_n(\mathbb{C})$ is a holomorphic map defined from $U_\alpha \cap U_\beta$ into the space of invertible $n \times n$ complex-valued matrices.*

The maps φ_α in the definitions above are called **local trivializations** of a bundle, since they guarantee that the bundle is locally isomorphic to a trivial bundle. We call the maps $f_{\alpha\beta}$ the **transition functions** of the bundle. Since $f_{\alpha\beta} = \varphi_\alpha \circ \varphi_\beta^{-1}$, we see that

$$f_{\beta\alpha} = f_{\beta\alpha}^{-1}, \qquad f_{\alpha\beta}.f_{\beta\gamma}.f_{\gamma\alpha} = 1$$

This second identity is known as the cocycle condition. Let us note that given an open cover $\{U_\alpha\}$ of M and holomorphic maps $f_{\alpha\beta} : U_\alpha \cap U_\beta \to GL_n(\mathbb{C})$ satisfying the two conditions above, we can define a rank n vector bundle by gluing together the trivial bundles $U_\alpha \times \mathbb{C}^n$ using the $f_{\alpha\beta}$. More precisely, we take the disjoint union of all the $U_\alpha \times \mathbb{C}^n$, and then quotient by the following equivalence relation: if $(m, v) \in U_\alpha \times \mathbb{C}^n$ and $(m', v') \in U_\beta \times \mathbb{C}^n$, $(m, v) \sim (m', v')$ if and only if $m = m'$ and $v' = f_{\beta\alpha}(p)v$. The cocycle condition guarantees that this is indeed an equivalence relation, and we denote an equivalence class by $[(p, v)]$. It is a straightforward matter (see [22]) to check that that the space obtained from this construction together with the projection map π defined by $\pi([p, v]) = p$, does indeed have the structure of a holomorphic line bundle over M.

For the moment, let us restrict our attention to holomorphic line bundles. The simplest example of a holomorphic line bundle is just the Cartesian product $M \times \mathbb{C}$ of a Riemann surface with the complex line. The projection map is just the canonical projection into the first factor of the product. This is known as the trivial line bundle over M. The following are some non-trivial examples of holomorphic line bundles:

Example 4.17 (Point bundles) Let p be a point on a compact Riemann surface M, and let (U_0, z) be a coordinate chart with local analytic parameter z containing p such that $z(p) = 0$. Define another open set U_1 on M by $U_1 = M - \{p\}$. Then since the function $z(p)$ does not vanish on $U_0 \cap U_1 = U_0 - \{p\}$, we define a line bundle over M with a single[5] transition function $f_{01} = z$. This is called the point bundle associated to the point $p \in M$, and is often denoted L_p.

Example 4.18 (The line bundle $\mathcal{O}(n)$) Let $\mathbb{P}^1$ be the Riemann sphere and let U_0 be the chart with analytic coordinate $z(p)$ obtained by stereographic projection from the north pole. Let U_1 be the chart obtained from projection from the south pole. Then on $U_0 \cap U_1 \cong \mathbb{C}^*$, the functions $z^n(p),\ n \in \mathbb{Z}$ are holomorphic and non-vanishing, so we can use them as a transition functions to define line bundles over the sphere. We shall denote these bundles by $\mathcal{O}(n)$. Note that $\mathcal{O}(1) = L_S$, where S is the south pole.

Example 4.19 (Canonical bundle) Let $\{U_\alpha, z_\alpha\}$ be a complex analytic atlas of a Riemann surface M. Then on each non-empty intersection $U_\alpha \cap U_\beta$, the coordinate z_β can be expressed as an analytic function of z_α. Since the transition maps in the complex analytic atlas are invertible, the functions $f_{\alpha\beta} = \frac{dz_\beta}{dz_\alpha}$ are non-vanishing and holomorphic. Moreover, the chain rule implies that the $f_{\alpha\beta}$ satisfy the cocycle condition $f_{\alpha\beta}.f_{\beta\gamma}.f_{\gamma\alpha} = 1$, and so they define a holomorphic line bundle over M. This line bundle is called the canonical bundle, or holomorphic cotangent bundle, of M.

[5]Note that the cocycle condition is automatically satisfied whenever an open cover consists of only two open sets

We now explain what is meant by a morphism in the category of holomorphic vector bundles over a Riemann surface M:

Definition 4.20 ***(Vector bundle morphisms)*** *A morphism between holomorphic vector bundles $\pi : E \to M$ and $\tilde{\pi} : \tilde{E} \to M$ over the same Riemann surface M is a holomorphic map $\Psi : E \to \tilde{E}$ such that $\tilde{\pi} \circ \Psi = \pi$, and the map from $\pi^{-1}(p)$ to $\tilde{\pi}^{-1}(p)$ given by the restriction of Ψ to a fibre is a linear map between complex vector spaces.*

Let us again return to the rank 1 case of holomorphic line bundles. If we think of line bundles of being built from their transition functions, it is important to know when two sets of transition functions define the isomorphic line bundles over M. Firstly, notice that if L is a line bundle with local trivialisations $(U_\alpha, \varphi_\alpha)$ and $\mathcal{V} = \{V_\alpha\}$ is any refinement of the open cover $\mathcal{U} = \{U_\alpha\}$, then the pairs $(V_\alpha, \varphi_\alpha|_{V_\alpha})$ also define local trivialisations of L.

Thus if we have two bundles over M defined via transition functions over different open coverings of M, by passing to a common refinement we may assume that the bundles are trivialised with respect to the same open cover of M.

Suppose now that the holomorphic line bundles $\pi : L \to M$ and $\tilde{\pi} : \tilde{L} \to M$ are isomorphic. This means that there is a bijective holomorphic map $I : L \to \tilde{L}$ satisfying $\tilde{\pi} \circ I = \pi$, and such that the restriction of I to each fibre is a linear isomorphism. If L and $\tilde{L}$ are have local trivialisations $\{U_\alpha, \varphi_\alpha\}$ and $\{U_\alpha, \psi_\alpha\}$ by refining again if necessary, we may assume that

$$I(\pi^{-1}(U_\alpha)) = (\tilde{\pi})^{-1}(U_\alpha)$$

Observe that we can define holomorphic maps $\lambda_\alpha : U_\alpha \to \mathbb{C}^*$ by

$$\lambda_\alpha(p) = \mathrm{proj}_{\mathbb{C}}(\psi_\alpha \circ I \circ \varphi_\alpha^{-1})(p)$$

Then if $\{f_{\alpha\beta}\}$ are the transition functions of L and $\{\tilde{f}_{\alpha\beta}\}$ are the transition functions of $\tilde{L}$, on the intersection $U_\alpha \cap U_\beta$ we find that

$$\tilde{f}_{\alpha\beta} = \lambda_\alpha f_{\alpha\beta} \lambda_\beta^{-1} \tag{4.1}$$

Now suppose conversely that the we begin with two line bundles L, $\tilde{L}$ on M defined by transition functions $f_{\alpha\beta}, \tilde{f}_{\alpha\beta}$ respectively with respect to an open cover $\{U_\alpha\}$ of M. Suppose further that for each U_α there exists a holomorphic map $\lambda_\alpha : U_\alpha \to \mathbb{C}^*$ such that on each intersection $U_\alpha \cap U_\beta$

$$\tilde{f}_{\alpha\beta} = \lambda_\alpha f_{\alpha\beta} \lambda_\beta^{-1}$$

Then if we think of points of L as being the equivalence classes $[p, v]$ in the quotient construction of the line bundle L, and points of $\tilde{L}$ as equivalence classes $\{p, v\}$, then the map $I : L \to \tilde{L}$ sending $[p, v]$ to $\{p, v\}$ defines an isomorphism of holomorphic line bundles. Thus, we see that two line bundles are isomorphic if and only if they admit local trivialisations satisfying the condition (4.1) for some collection of nonvanishing holomorphic maps λ_α defined on each open set U_α.

Given vector spaces V and W, we can form the direct sum $V \oplus W$, or their tensor product $V \otimes W$, or their consider their dual spaces V^* and W^*. We can adapt these constructions to vector bundles over a fixed Riemann surface in straightforward ways by performing suitable operations on transition functions. For instance, the tensor product of bundles E and F is defined as follows: if $e_{\alpha\beta} : U_\alpha \to GL_n(\mathbb{C})$ and $f_{\alpha\beta}$ are the transition functions for E and F respectively, then the transition functions for $E \otimes F$ are given by $t_{\alpha\beta} = e_{\alpha\beta} \otimes f_{\alpha\beta}$. Similarly, given a vector bundle E with standard fibre V and transition functions $t_{\alpha\beta}$, we can build the dual bundle E^*, whose standard fiber is V^* and whose transition functions are given by $(t^T_{\alpha\beta})^{-1}$, where $t^T_{\alpha\beta}$ the transpose of the matrix $t_{\alpha\beta}$. In the case of a line bundle L with non-vanishing holomorphic transition functions $f_{\alpha\beta}$, the dual bundle L^* has transition functions $\frac{1}{f_{\alpha\beta}}$. For this reason, it is sometimes called the inverse bundle to L.

Given a rank m vector bundle V over a Riemann surface M, we can build a line bundle $\det V$ over M called the **determinant bundle** of V. If the transition maps of V are given by $m \times m$ invertible matrices $g_{\alpha\beta}$ defined over $U_\alpha \cap U_\beta$, then the determinant bundle is the holomorphic line bundle with transition functions given by $\det g_{\alpha\beta}$. It follows from the basic properties of matrix determinants that these transition functions obey the cocycle conditions and hence define a holomorphic line bundle over M.

Finally, suppose that $f : M \to N$ is a holomorphic map between compact Riemann surfaces that exhibits M as an r-sheeted branched covering of N. If L is a line bundle over N, then the map f induces a line bundle f^*L over M called the pullback bundle of L by f, defined as follows: if U_α is an open cover of N and L is defined by transition functions $\{U_\alpha, g_{\alpha\beta}\}$, then f^*L is defined by the transition functions $g_{\alpha\beta} \circ f$ with respect to the open cover $\{f^{-1}(U_\alpha)\}$ of M. Categorically speaking, as a complex manifold f^*L is the *fibre product*

$$f^*L = \{(x, q) \in L \times M : \pi(x) = f(q)\}$$

4.3 Sections of Holomorphic Line Bundles

Definition 4.21 ***(Holomorphic sections)*** *A holomorphic section s of a holomorphic line bundle $\pi : L \to M$ is a holomorphic map $s : M \to L$ such that $\pi \circ s = \mathrm{id}_M$*

When a section is restricted to a local trivialization $\varphi_\alpha : \pi^{-1}(U_\alpha) \to U_\alpha \times \mathbb{C}$, it is represented by a holomorphic function s_α such that $\varphi_\alpha \circ s = (id_{U_\alpha}, s_\alpha)$. In the trivialization φ_β we also have $\varphi_\beta \circ s = (id_{U_\beta}, s_\beta)$, so on the overlap $U_\alpha \cap U_\beta$, we must have $s_\alpha = f_{\alpha\beta} s_\beta$. Hence we can also think of a section of a line bundle as a collection of locally defined holomorphic functions s_α which patch together via the bundle's transition functions. We define a zero of a holomorphic section to be any point $p \in M$ such that in a local trivialization φ_α around p, the local function s_α satisfies $s_\alpha(p) = 0$. Notice that since the bundle's transition functions are non-vanishing, the notion of a zero of a section is well-defined and independent of the local trivialization we chose to examine. Since we can add sections pointwise and multiply them by complex numbers, the set of all holomorphic sections of a line bundle forms forms a complex vector space. Any holomorphic line bundle has a holomorphic section called the zero section, which is identically zero in each local trivialization. We now look at some more interesting examples of holomorphic sections:

Example 4.22 (Sections of the trivial bundle) If $M \times \mathbb{C}$ is the trivial bundle over M, then a holomorphic section of M is simply a holomorphic function $f : M \to \mathbb{C}$. So by Proposition 4.7, we see that the sections of a trivial bundle must all be constant. Also, observe that if a line bundle L over a Riemann surface M has a non-vanishing, holomorphic section s, then L is isomorphic to the trivial bundle *via* the following map:

$$\Psi : M \times \mathbb{C} \to L, \qquad (p, v) \mapsto v.s(p)$$

Example 4.23 (Sections of point bundles) If L_p is the point bundle associated to the point p on a Riemann surface M, then we can define a non-trivial section as follows. Using the cover $M = U_0 \cup U_1$ from the definition of L_p, we define a section s by the pair of local functions $s_0 = z$, $s_1 = 1$, where z is the local analytic coordinate centered at p on U_0. Since $f_{01} = z$, we see that this indeed defines a section of L_p.

Example 4.24 (Sections of $\mathcal{O}(n)$) In the case of the bundle $\mathcal{O}(n)$ over $\mathbb{P}^1$, we can give a very concrete description of the vector space of holomorphic sections. In the chart U_0 with local coordinate z, a section is represented by a holomorphic function s_0 defined on U_0, with Taylor series $s_0 = \sum_{k=0}^{\infty} a_k z^k$. On U_1 with local parameter w, we have $s_1 = \sum_{k=0}^{\infty} b_k w^k$. On the intersection $U_0 \cap U_1$, we have $z = \frac{1}{w}$, and $s_0 = f_{01} s_1 = z^n s_1$. That is,

$$\sum_{k=0}^{\infty} a_k z^k = z^n \sum_{k=0}^{\infty} b_k w^k = \sum_{k=0}^{\infty} b_k z^{n-k}$$

Equating coefficients and requiring both sides to be holomorphic tells us that $a_k = 0$ for $k > n$, and $a_k = b_{n-k}$ for $0 \leq k \leq n$. So any section is described by the $n+1$ non-zero coefficients of its Taylor series, and hence the space of holomorphic sections of $\mathcal{O}(n)$ is of dimension $n + 1$.

Example 4.25 (Sections of the canonical bundle) Let K be the canonical bundle over a Riemann surface M. A global section of K is a collection of holomorphic functions s_α defined on a coordinate covering (U_α, z_α) of M which satisfy $s_\alpha = \frac{dz_\beta}{dz_\alpha} s_\beta$ on each non-empty intersection $U_\alpha \cap U_\beta$. Given the form of these transition functions, we can interpret such a section as a globally defined complex valued holomorphic 1-form on M. The transition functions of K are equivalent to the consistency requirement that

$$s_\alpha dz_\alpha = s_\beta dz_\beta$$

Suppose now that $\{U_\alpha\}$ is an open cover of a Riemann surface M and L is a holomorphic line bundle over M defined by transition functions $f_{\alpha\beta}$. If s_α is a collection of meromorphic functions defined on each open set U_α which satisfies $s_\alpha = f_{\alpha\beta} s_\beta$ on each non-empty intersection $U_\alpha \cap U_\beta$, then we say the $\{s_\alpha\}$ define a **meromorphic section** s of L. A holomorphic section $t = \{t_\alpha\}$ of a line bundle L gives rise to a meromorphic section t^{-1} of the line bundle L^*, where t^{-1} is defined by the collection of locally defined meromorphic functions $\frac{1}{t_\alpha}$. Similarly, given a holomorphic section r of a line bundle $\tilde{L}$

defined by local holomorphic functions r_α, we can define a the product of t and r to be a section $t \cdot r$ of $L \otimes \tilde{L}$ which is defined by the local functions $t_\alpha r_\alpha$. Note that the ratio of two sections r and s of a holomorphic line bundle L defines a meromorphic section of the trivial bundle, that is, a meromorphic function on M.

The **divisor** of a holomorphic section is its set of zeroes, counted with multiplicity. It is traditional to represent divisors as weighted sums of points. For instance, if a section s has a simple zero at p_1 and a zero of order 3 at p_2, then the divisor of s is written $\mathcal{D}(s) = p_1 + 3p_2$. The divisors of holomorphic sections of a line bundle actually reveal a great deal of information about the bundle itself. For instance, suppose that a line bundle L has a non-trivial holomorphic section s which has zeroes at points $p_1, ..., p_n$, with the zero at p_k having multiplicity m_k. Let L_p^{-m} denote the tensor product of m copies of the dual of the point bundle L_p associated to $p \in M$, and let s_p^{-m} be the product of m copies of the meromorphic section s_p^{-1}, where s_p is a section of L_p with a simple zero at p. Then the line bundle $L \otimes L_{p_1}^{-m_1} \otimes \cdots \otimes L_{p_n}^{-m_n}$ has a non-vanishing holomorphic section $s \cdot s_{p_1}^{-m_1} \cdots s_{p_k}^{-m_k}$. This implies that the bundle $L \otimes L_{p_1}^{-m_1} \otimes \cdots \otimes L_{p_n}^{-m_n}$ is trivial, and so we find that L is isomorphic to $L_{p_1}^{m_1} \otimes \cdots \otimes L_{p_n}^{m_n}$. In other words, we have shown that if a holomorphic line bundle admits a non-trivial holomorphic section, then it is isomorphic to a tensor product of point bundles.

4.4 Sheaves

Many fundamental objects in topology and geometry, such as manifolds and vector bundles, are required to be locally 'standard'. In the case of holomorphic line bundles, we were able to use the bundle's local structure to describe global sections as a collections of functions defined on open sets obeying certain transformation laws on their intersections. The technology of sheaves gives us a systematic way to understand the interaction between this sort of local data and the corresponding global objects.

Definition 4.26 ***(Presheaf)*** *Let M be a topological space and let $\mathcal{C}$ be a category. A presheaf $\mathcal{F}$ on M with values in the category $\mathcal{C}$ associates to each open set $U \subset M$ an object $\mathcal{F}(U) \in \mathrm{Obj}(\mathcal{C})$, and to each inclusion $U \subset V$ of open sets U and V a restriction morphism $r_{UV} : \mathcal{F}(V) \to \mathcal{F}(U) \in \mathrm{Mor}_{\mathcal{C}}(\mathcal{F}(V), \mathcal{F}(U)$ such that the following conditions hold:*

1. $r_{UU} = id$
2. *if $U \subset V \subset W$, then $r_{UV} \circ r_{VW} = r_{UW}$.*

Remark: Observe that if $\mathrm{Open_M}$ is the category whose objects are the open sets of M and whose morphisms are inclusions of open sets, then a presheaf is a contravariant functor from $\mathrm{Open_M}$ to the category C.

Notice that in the setting of presheaves, we do not yet have a means of 'gluing' together local data. This motivates the next definition, that of a sheaf:

Definition 4.27 ***(Sheaf)*** *A presheaf $\mathcal{F}$ over M is called a sheaf if the following condition (the sheaf axiom) is satisfied whenever we can write an open set $U \subset M$ as $U = \bigcup_\alpha U_\alpha$ for some collection of open sets $\{U_\alpha\}$:*

Given a collection of elements $s_\alpha \in \mathcal{F}(U_\alpha)$ such that for each pair of open sets U_α, U_β , $r_{U_\alpha \cap U_\beta U_\alpha}(s_\alpha) = r_{U_\alpha \cap U_\beta U_\beta}(s_\beta)$, there exists a unique $s \in \mathcal{F}(U)$ such that $r_{U_\alpha U}(s) = s_\alpha$ for all α.

Example 4.28 (Some examples of sheaves)

1. If M is a Riemann surface, we can define a presheaf of abelian groups $\mathcal{O}_M$ which associates to each open subset $U \subset M$ the abelian group of holomorphic functions defined on U, the group multiplication being pointwise addition of functions. The restriction homomorphisms are the natural ones obtained by restricting the domain of a function. The sheaf axiom is easily seen to hold[6], so that $\mathcal{O}_M$ defines a sheaf of abelian groups.
2. We can also define a sheaf $\mathcal{O}^*_M$ which associates to each U the abelian group of non-vanishing holomorphic functions on U, where we take the group operation to be pointwise multiplication of functions, equipped with the natural restriction homomorphisms.
3. Let G be an abelian group, and let p be a point of M. The **skyscraper sheaf** G_p is defined by

$$G_p(U) = \begin{cases} G & \text{if } p \in U, \\ 0 & \text{otherwise} \end{cases}$$

4. The constant sheaves $\mathbb{C}$ and $\mathbb{Z}$ associate to U the locally constant functions over U taking values in $\mathbb{C}$ or $\mathbb{Z}$, again with the natural restriction homomorphisms.
5. If $\pi : L \to M$ is a holomorphic line(or vector) bundle over a Riemann surface M, then we can form a sheaf $\mathcal{O}(L)$ which associates to an open set $U \in M$ the space of holomorphic sections of the restricted bundle $\pi^{-1}(U)$ over U. Note that if L is the trivial line bundle its sections are just holomorphic functions, and so $\mathcal{O}(L) = \mathcal{O}_M$. Also note that in the case of the canonical bundle, we obtain a sheaf Ω^1_M which associates to each open set U the holomorphic one-forms defined over U.

Although the sheaf axiom was easily seen to hold in the examples above, it is important to realise that there do exist presheaves which are not sheaves:

Example 4.29 (Failure of the sheaf axiom) Let M be a topological space and G be an abelian group. Observe that we can define a presheaf G^M by associating to each open set $U \subset M$ the constant functions from U to G, and using the natural restriction maps. However, considering an open set with multiple connected components, we observe that the sheaf axiom does not hold for the presheaf G^M.

Our next task is to define a morphism in the category of sheaves:

[6] As it will for any presheaf of functions satisfying a 'local' property such as continuity or smoothness, equipped with the natural restriction maps.

Definition 4.30 ***(Morphism of sheaves)*** *A morphism or sheaf map $\phi : \mathcal{S} \to \mathcal{T}$ between sheaves $\mathcal{S}$ and $\mathcal{T}$ over a Riemann surface M is a collection of homomorphisms $\phi_U : \mathcal{S} \to \mathcal{T}$ for each open set $U \subset M$, such that for all open subsets $V \subset U$, the ϕ_U commute with the restriction homomorphisms of $\mathcal{S}$ and $\mathcal{T}$:*

$$\begin{array}{ccc} \mathcal{S}(U) & \xrightarrow{\phi_U} & \mathcal{T}(U) \\ {\scriptstyle r_{VU}}\downarrow & & \downarrow{\scriptstyle r'_{VU}} \\ \mathcal{S}(V) & \xrightarrow{\phi_V} & \mathcal{T}(V) \end{array}$$

Example 4.31 The identity sheaf map $Id_{\mathcal{S}} : \mathcal{S} \to \mathcal{S}$ assigns to each abelian group $\mathcal{S}(U)$ the identity homomorphism $\phi_U = id_{\mathcal{S}(U)}$. Clearly the ϕ_U commute with the restriction homomorphisms of $\mathcal{S}$, so Id is indeed a sheaf map.

In order to extract very local information about a sheaf, it is useful to introduce the concept of a stalk.

Definition 4.32 ***(Stalk)*** *If $\mathcal{S}$ is a sheaf of abelian groups on a topological space M, then the stalk $\mathcal{S}_p$ of $\mathcal{S}$ at the point $p \in M$ is the abelian group defined as the direct limit*

$$\mathcal{S}_p = \varinjlim \{\mathcal{S}(U) : p \in U\}$$

The direct limit in the definition above is perhaps best spelled out in words: the stalk at p is the set of all pairs (U, s) where U is an open set containing p, modulo the equivalence relation $(U, s) \sim (V, t)$ if and only if there exits an open subset $W \subset U \cap V$ such that $r_{WV}t = r_{WU}s$. The stalk has the structure of an abelian group, with addition defined by

$$[(U, s)] + [(V, t)] = [(U \cap V, r_{U\cap VU}s + r_{U\cap VV})]$$

If U is an open set containing the point p, and $s \in \mathcal{S}(U)$, then we denote by s_p the equivalence class of s in the stalk $\mathcal{S}_p$. An element $s_p \in \mathcal{S}_p$ is called the **germ** of s at p.

Example 4.33 Let M be a Riemann surface and let $\mathcal{O}_M$ be the sheaf of holomorphic functions on M. At a point $p \in M$, let z a local analytic coordinate on M centered at p. Then the stalk $(\mathcal{O}_M)_p$ is given by the convergent power series in z. If $\mathcal{S}_p$ is the skyscraper sheaf considered previously, then the stalk of $\mathcal{S}_p$ is G at p and the trivial group everywhere else.

We observe that a sheaf map $\phi : \mathcal{S} \to \mathcal{T}$ naturally induces homomorphisms of stalks $\phi_p : S_p \to T_p$ for all $p \in M$ as follows:

$$\phi_p \cdot [(U, s)] = [(U, \phi_U(s))]$$

At this point we must mention a somewhat technical point concerning sheaves and presheaves. Recall from Example 4.29 that it is possible that a presheaf over a topological space may not satisfy the sheaf axiom. However, it can be shown (see Chapter 1 of [14]) that given any presheaf S over a space M one naturally associate a sheaf $\mathcal{S}$ to S called the **sheafification** of S whose stalks will be isomorphic to those of the presheaf S. If S is already a sheaf, then the sheafification of S is simply S itself. For the details of this construction, the reader is referred to [14].

The following Proposition, whose proof can be found on page 287 of [25], clarifies the notions of injectivity and surjectivity for sheaf maps.

Lemma 4.34 *Let $\phi : \mathcal{Q} \to \mathcal{R}$ be a sheaf map. Then the following are equivalent:*

1. *For each open subset $U \subset M$, the map $\phi_U : \mathcal{Q}(U) \to \mathcal{R}(U)$ is injective (respectively, bijective)*
2. *Each map of stalks induced by ϕ is injective (respectively, bijective).*

Crucially, this lemma is false if we replace the words 'injective' or 'bijective' with 'surjective'. Indeed, let $X = \mathbb{C} - \{0\}$, and let $\mathcal{O}_M$ be the sheaf of holomorphic functions on M. Then the sheaf map defined by $\phi_U(f) = \exp(2\pi i f)$ can be easily seen to be surjective on each stalk $(\mathcal{O}_M)_p$: one simply considers a open neighbourhood of p small enough that a single-valued complex logarithm may be locally defined. Since there is no continuous, let alone holomorphic, logarithm defined $\mathbb{C} - \{0\}$, however, we see that the map ϕ_M fails to be surjective.

With Lemma 4.35 in mind, we define an exact sequence of sheaves as follows:

Definition 4.35 ***(Exact sequence of sheaves)*** *The sequence*

$$\cdots \longrightarrow \mathcal{T} \longrightarrow \mathcal{S} \longrightarrow \mathcal{Q} \longrightarrow \cdots$$

is said to be exact at $\mathcal{S}$ if the sequence of stalks

$$\cdots \longrightarrow \mathcal{T}_p \longrightarrow \mathcal{S}_p \longrightarrow \mathcal{Q}_p \rightarrow \cdots$$

is exact for all $p \in M$.

The following are examples of short exact sequences of sheaves:

Example 4.36 (Exponential sheaf sequence) On a Riemann surface M, the exponential sheaf sequence is

$$0 \longrightarrow \mathbb{Z} \longrightarrow \mathcal{O}_M \xrightarrow{\exp} \mathcal{O}_M^* \longrightarrow 0$$

where the map $\mathbb{Z} \to \mathcal{O}_M$ is the natural inclusion, and the exponentiation map exp is defined by $\exp_U : \mathcal{O}_M(U) \to \mathcal{O}_M^*(U), \quad f \mapsto \exp(2\pi i f)$. The sequence is exact, since the kernel of the map $\exp_U$ is precisely the locally constant integer-valued functions over U. And the exponentiation map from $\mathcal{O}_M$ to $\mathcal{O}_M^*$ is surjective, since when restricted to a sufficiently small simply connected open set U, every non-vanishing holomorphic function has a well-defined complex logarithm.

Example 4.37 If we let K denote the canonical bundle of holomorphic 1-forms on a Riemann surface M, and $\mathcal{O}(K)$ denote its sheaf of holomorphic sections, we have the following exact sequence of sheaves:

$$0 \longrightarrow \mathbb{C} \longrightarrow \mathcal{O}_M \xrightarrow{d} \mathcal{O}(K) \longrightarrow 0$$

where d denotes the sheaf map $\mathcal{O}_M(U) \to \mathcal{O}(K)(U)$, $f(z_\alpha) \mapsto \frac{df}{dz_\alpha} dz_\alpha$ corresponding to taking the differential of a holomorphic function to obtain a holomorphic 1-form. That this sequence is exact follows from the Poincare lemma[7].

Example 4.38 Let L be a holomorphic line bundle over a Riemann surface M, and let L_p be the point bundle associated to a point $p \in M$. Let L_p^{-n} be the tensor product of n copies of the dual bundle L_p^*, and let s_p^n be a section of L_p which has a zero of order n at p and no other zeroes. Then there is a short exact sequence of sheaves

$$0 \longrightarrow \mathcal{O}(L \otimes L_p^{-n}) \xrightarrow{s_p^n} \mathcal{O}(L) \longrightarrow \mathcal{Q} \longrightarrow 0$$

where the sheaf map s_p is given by multiplication by the section s_p, and where $\mathcal{Q}$ is the skyscraper sheaf

$$\mathcal{Q}(U) = \begin{cases} \mathbb{C}^n & \text{if } p \in U, \\ 0 & \text{otherwise} \end{cases}$$

To see this, note that the stalk of $\mathcal{Q}$ at any point $q \neq p$ must be zero, since for a sufficiently small open set U containing q, multiplication by s_p^n establishes an isomorphism of $\mathcal{O}(L \otimes L_p^{-n})(U)$ with $\mathcal{O}(L)(U)$. The stalk of $\mathcal{Q}$ at p is the quotient of the space of germs of sections of L at p by the subspace consisting germs of sections of L vanishing to order n at p. Using the description of germs of holomorphic sections in terms of convergent power series, we see that a section of L vanishing to order n at p is represented by a series of the form $\sum_{k \geq 0} a_k z^{n-k}$. Hence the stalk $\mathcal{Q}_p$ is isomorphic to $\mathbb{C}^n$, the isomorphism being given by associating to any germ $\sum_{k \geq 0} a_k z^k$ the n-tuple $(a_0, \ldots, a_n)$

In order to study the sheaves of sections of holomorphic vector bundles in more detail, it will be necessary to introduce a further concept, that of locally free sheaves.

Definition 4.39 ***(Ringed space)*** *A ringed space is a topological space M together with a sheaf of commutative rings $\mathcal{O}_M$ on M.*

For us, the most important example of a ringed space is a compact Riemann surface M together with the sheaf of holomorphic functions on M, $\mathcal{O}_M$. On a locally ringed space, it is possible to define sheaves of modules:

[7] See [14], p.68.

Definition 4.40 ***(Sheaf of modules)*** *A sheaf of abelian groups $\mathcal{S}$ on a ringed space $(M, \mathcal{O}_M)$ is called a sheaf of $\mathcal{O}_M$-modules if for each open set $U \subset M$, the abelian group $\mathcal{S}(U)$ has the structure of an $\mathcal{O}_M(U)$-module, and the restriction maps of $\mathcal{S}$ are compatible with the module structures induced by the restriction maps of $\mathcal{O}_M$.*

Observe that the stalk $\mathcal{S}_p$ of a sheaf $\mathcal{S}$ of $\mathcal{O}_M$-modules naturally inherits the structure of an $(\mathcal{O}_M)_p$-module. A sheaf of $\mathcal{O}_M$-modules $\mathcal{S}$ is called **locally free** of rank m if for all $p \in M$, the stalk $\mathcal{S}_p$ is a free $(\mathcal{O}_M)_p$-module of rank m. The prototypical example of a locally free sheaf on a compact Riemann surface M is the sheaf of holomorphic sections of a vector bundle on M. Indeed, it can be shown (see [16], page 14) that the functor associating to a holomorphic vector bundle E its sheaf of holomorphic sections $\mathcal{O}(E)$ is an equivalence of categories, so that every locally free sheaf of $\mathcal{O}_M$-modules arises as the sheaf of sections of a holomorphic vector bundle. Because of this correspondence, the terms 'holomorphic vector bundle' and 'locally free sheaf of $\mathcal{O}_M$-modules' are often used interchangeably. Locally free sheaves of rank 1, which correspond to sheaves of sections of holomorphic line bundles, are also referred to as **invertible sheaves**[8].

Given sheaves of $\mathcal{O}_M$ modules, we can consider their direct sums and tensor products: for instance, if $\mathcal{S}_2, \mathcal{S}_2$ are two sheaves of $\mathcal{O}_M$-modules over M, then their tensor product is defined as the sheafification of the presheaf $\{\mathcal{S}_2(U) \otimes_{\mathcal{O}_M(U)} \mathcal{S}_2(U)\}$.

Finally, we want to understand how locally free sheaves behave under holomorphic maps bewteen Riemann surfaces. Suppose that M and N are two compact Riemann surfaces and $f : M \to N$ is a holomorphic map. Given a sheaf $\mathcal{F}$ of $\mathcal{O}_N$-modules over N, there exists an induced sheaf $f^*\mathcal{S}$ of $\mathcal{O}_M$-modules called the inverse image sheaf of $\mathcal{F}$ over M which is defined as follows.

Firstly, given any sheaf of abelian groups $\mathcal{S}$ over N, we define a sheaf $f^{-1}(\mathcal{S})$ of abelian groups over M as the sheafification of the presheaf of abelian groups on M which associates to an open set $U \subset M$ the abelian group $\mathcal{S}(f(U))$. Observe that in the case of our sheaf of $\mathcal{O}_N$-modules $\mathcal{F}$, the sheaf $f^{-1}(\mathcal{F})$ has the structure of an $f^{-1}(\mathcal{O}_N)$-module. Now since the sheaf of rings $f^{-1}(\mathcal{O}_N)$ can be naturally identified with a subsheaf of $\mathcal{O}_M$, we can view $\mathcal{O}_M$ as a sheaf of modules over $f^{-1}(\mathcal{O}_N)$. So we define the **inverse image sheaf** $f^*\mathcal{S}$ of $\mathcal{O}_M$ by

$$f^*\mathcal{S} := \mathcal{O}_M \otimes_{f^{-1}(\mathcal{O}_N)} f^{-1}(\mathcal{F})$$

The reader is invited to check that if $\mathcal{O}(L)$ is the sheaf of holomorphic sections of a line bundle L over N, then the inverse image $f^*(\mathcal{O}(L))$ can be identified with the sheaf of holomorphic sections of the pullback bundle f^*L over M.

Given a sheaf $\mathcal{S}$ abelian groups over M, we can also define an induced sheaf of abelian groups $f_*(\mathcal{S})$ over N, called the **direct image** sheaf, as the presheaf[9] that associates to an open set $U \subset N$ the group $\mathcal{S}(f^{-1}(U))$. Observe that the sheaf of rings $\mathcal{O}_N$ is naturally identified as a subsheaf[10] of $\mathcal{O}_M$. Therefore if $\mathcal{F}$ is a sheaf of $\mathcal{O}_M$-modules, the direct image sheaf $f_*(\mathcal{S})$ can be considered as a sheaf of $\mathcal{O}_N$-modules over N.

The following Proposition summarises some of the key properties of inverse image and direct image sheaves. The proof can be found in [16], page 54.

[8] Inversion corresponds to passing to the sheaf of sections of the dual bundle

[9] Note that this presheaf is indeed a sheaf, and thus requires no sheafification.

[10] via the embedding $h \mapsto h \circ f$ of $f_*(\mathcal{O}_M)$ for $h \in \mathcal{O}_N(U)$).

Proposition 4.41 ***(Induced mappings of locally free sheaves)*** *Let $f : M \to N$ be a complex analytic mapping between compact Riemann surfaces which exhibits M as an r-sheeted branched covering of N, let L be a line bundle over M, and let $\mathcal{O}(L)$ be its sheaf of sections, which is a rank 1 locally free sheaf of $\mathcal{O}_M$-modules over M. Then:*

1. *The direct image $f_*(\mathcal{O}(L))$ is a locally free sheaf of $\mathcal{O}_N$-modules of rank r.*
2. ***(Projection formula)*** *For any locally free sheaf B of rank 1 over N,*

$$f_* \left(f^*\mathcal{O}(B) \otimes_{\mathcal{O}_M} L\right) \cong \mathcal{O}(B) \otimes_{\mathcal{O}_N} f_*(\mathcal{O}(L)) \tag{4.2}$$

Proof. For the proof of the projection formula, which is an exercise in applying the definitions of tensor products and inverse images, we refer the reader to [16], page 54. The proof of item 1, however, is quite instructive. To show that $f_*(\mathcal{O}(L))$ is a locally free sheaf of $\mathcal{O}_N$-modules of rank r, we must show that each point $p \in N$ has a neighbourhood U on which

$$f_*(\mathcal{O}(L))(U) \cong \mathcal{O}_N(U) \oplus \cdots \oplus \mathcal{O}_N(U) \qquad r \text{ times}$$

If p is not a branch point of the covering map f, then the preimage $f^{-1}(U)$ of a sufficiently small neighbourhood U of p will consist of r disjoint open sets V_i such that the restriction of f to each V_i is a holomorphic bijection of V_i with U. So by definition of the direct image, we have

$$f_*(\mathcal{O}(L))(U) = \mathcal{O}(L)(f^{-1}(U)) = \bigoplus_{i=1}^{r} \mathcal{O}_N(V_i)$$

as required. If p is a branch point, so that the fibre $f^{-1}(p)$ contains a ramification point q, the situation is slightly different. In a local coordinate patch (V, z) on M centered at q, the Local Normal Form theorem 4.12 implies f takes the form $z \mapsto z^k$ for some integer $k > 1$, so that $w = z^k$ gives a local coordinate on $f(V)$. Assuming that V is sufficiently small, the holomorphic sections of L over V correspond to power series in the variable z convergent on V. Now observe that for suitable power series $h_0(w), \ldots, h_{k-1}(w)$ in the variable w, any convergent power series $H(z)$ in the coordinate z can be written

$$H(z) = h_0(w) + zh_1(w) + \cdots + z^{k-1}h_{k-1}(w)$$

Thus, $\mathcal{O}(L)(V)$ is isomorphic to the direct sum of k copies of the space of holomorphic functions on $f(V)$:

$$\mathcal{O}(L)(V) = \bigoplus_{i=0}^{k-1} z^i \cdot \mathcal{O}_N(f(V))$$

Since the map f has degree r, by applying this reasoning to each point in the fibre $f^{-1}(p)$ we can obtain an neighbourhood U of p on which $f_*(\mathcal{O}(L))$ is again isomorphic to the direct sum of r copies of $\mathcal{O}_N(U)$. We conclude that $f_*(\mathcal{O}(L)))$ is a locally free sheaf of $\mathcal{O}_N$-modules of rank r. □

In view of the fact proved above, given a degree r holomorphic map $F : M \to N$ we can consider the direct image construction as an operation which converts a line bundle L over M to rank r vector bundle f_*L over N.

4.5 Sheaf Cohomology

Although the language of sheaves makes many local statements easy to formulate, we would like to be able to draw global conclusions. We now introduce the fundamentals of the theory of sheaf cohomology, which allows us to define groups which reflect the global structure of a sheaf.
We begin by defining the cohomology of a sheaf with respect to an open cover. Let M be a topological space, and let $\mathcal{U} = \{U_\alpha\}$ be a locally finite open cover of M. We can use this data to define a simplicial complex $N(\mathcal{U})$, called the **nerve** of $\mathcal{U}$, as follows. For the vertices of $N(\mathcal{U})$, we simply take the entire collection of open sets U_α of the covering. If $(U_{\alpha_0}, \dots, U_{\alpha_p})$ is a collection of $p+1$ open sets in $\mathcal{U}$ with $U_{\alpha_0} \cap \dots \cap U_{\alpha_p} \neq \phi$, we say that $U_{\alpha_0}, \dots, U_{\alpha_p}$ span a **p-simplex** $\sigma = (U_{\alpha_0}, \dots, U_{\alpha_p})$. The intersection $|\sigma| = U_{\alpha_0} \cap \dots \cap U_{\alpha_p}$ is called the **support** of σ. Now suppose that $\mathcal{S}$ is a sheaf of abelian groups on M. A **p-cochain of the cover $\mathcal{U}$ with coefficients in the sheaf** $\mathcal{S}$ is a function f which assigns to each p-simplex $\sigma \in N(\mathcal{N})$ an element $f(\sigma) \in \mathcal{S}(|\sigma|)$. We denote the set of all p-cochains by $C^p(\mathcal{U}, \mathcal{S})$. Observe that $C^p(\mathcal{U}, \mathcal{S})$ has a natural abelian group structure: if $f, g \in C^p(\mathcal{U}, \mathcal{S})$, we define $(f+g)(\sigma) := f(\sigma) + g(\sigma)$. For brevity, if $\sigma = (U_{\alpha_0}, \dots, U_{\alpha_p})$, we shall often write $f_{\alpha_0,\dots,\alpha_p} := f(\sigma)$. Now we can define boundary homomorphisms $\delta^p : C^p \to C^{p+1}$ as follows:

$$\{\delta^p(f)\}_{\alpha_0.\dots\alpha_{p+1}} = \sum_{i=0}^{p+1} (-1)^i r_i(f_{\alpha_0\dots\hat{\alpha}_i\dots\alpha_{p+1}})$$

where r_i denotes the restriction homomorphism from $\mathcal{S}(U_{\alpha_0} \cap \dots \cap \hat{U}_{\alpha_i} \cap \dots \cap U_{\alpha_{p+1}})$ to $\mathcal{S}(U_{\alpha_0} \cap \dots \cap U_{\alpha_{p+1}})$. To illustrate the definition of these maps, if $f \in C^0$, $(\delta^0 \cdot f)_{\alpha\beta} = f_\beta - f_\alpha$. And if $h \in C^1$, $(\delta^1 \cdot h)_{\alpha\beta\gamma} = f_{\beta\gamma} - f_{\alpha\gamma} + f_{\alpha\beta}$. A simple calculation shows that $\delta^{p+1} \circ \delta^p = 0$, allowing us to define the pth cohomology group of $\mathcal{S}$ relative to the cover U_α by

$$H^p(\{U_\alpha\}, \mathcal{S}) = \frac{\ker(\delta^{p+1})}{\operatorname{im}(\delta^p)}$$

In order to remove the dependence of these groups on the choice of covering, one takes the direct limit of these groups over all locally finite open coverings of M, with respect to the partial order on coverings given by one covering being a refinement of another. As a result of this process we obtain sequence of groups $H^p(M, \mathcal{S})$. However for almost all practical purposes, this definition proves extremely cumbersome. Fortunately, we are able to appeal to the following useful result:

Theorem 4.42 ***(Leray's theorem)*** *Let M be a topological space and $\mathcal{S}$ a sheaf of abelian groups on M. Let $\mathcal{U} = \{U_\alpha\}$ be a locally finite open covering of M such that for every q-fold intersection $\sigma = U_{\alpha_0} \cap \dots \cap U_{\alpha_k}$, $H^q(\sigma, \mathcal{S}) = 0$ for $q \geq 1$. Then $H^p(\mathcal{U}, \mathcal{S}) = H^p(M, \mathcal{S})$ for all p.*

For a proof, see [14]. An open cover satisfying the hypotheses of Leray's theorem for a sheaf $\mathcal{S}$ is called Leray open cover with respect to $\mathcal{S}$.

Although their definitions may seem complicated, many sheaf cohomology groups have very natural geometric interpretations. If $\pi : L \to M$ is a holomorphic line bundle over a Riemann surface M, and $\mathcal{O}(L)$ is its sheaf of holomorphic sections, then we can think of an element of $H^0(M, \mathcal{O}(L))$ as a collection of local sections f_α defined over each open set U_α in a Leray cover of M, such that on an intersection $U_\alpha \cap U_\beta$, $\delta_0 \cdot f = f_\beta - f_\alpha = 0$.

But this is precisely the condition for our local sections f_α to patch together consistently to give us a well-defined global section of L. Hence the group $H^0(M, \mathcal{O}(L))$ is identified with the space of global holomorphic sections of the line bundle L. We shall sometimes abbreviate the notation for this group to simply $H^0(M, L)$. Similarly, the group $H^0(M, \mathcal{O}_M)$ is identified with the globally defined holomorphic functions on M.

Another important sheaf cohomology group, which we have actually already considered, is the group $H^1(M, \mathcal{O}^*_M)$, where M is a Riemann surface and $\mathcal{O}^*_M$ is the sheaf of non-vanishing holomorphic functions on M. To interpret this group, recall that an element f of $\ker\left(\delta_1 : C^1 \to C^2\right)$ assigns to each non-empty intersection of open sets $U_\alpha \cap U_\beta$ in a Leray cover of M a non-vanishing holomorphic function $f_{\alpha\beta}$ such that $(\delta_1 \cdot f)_{\alpha\beta\gamma} = f_{\alpha\beta}.f_{\beta\gamma}.f_{\gamma\alpha} = 1$. But recall that this is precisely the condition required for the collection of functions $f_{\alpha\beta}$ to define the transition functions for a holomorphic line bundle L over M. Moreover, we've already seen that two sets of transition functions $f_{\alpha\beta}$ and $g_{\alpha\beta}$ define the same holomorphic line bundle if and only if we can write $f_{\alpha\beta} = \frac{h_\alpha}{h_\beta} \cdot g_{\alpha\beta}$ for some nonvanishing holomorphic functions h_α and h_β defined on U_α and U_β respectively. We can think of the functions h_α as representing a 0-cochain $h \in C^0$. But by definition of the coboundary map, $\frac{h_\alpha}{h_\beta} = (\delta_0 \cdot h)_{\alpha\beta}$. So we see that we can identify the group

$$H^1(M, \mathcal{O}^*_M) = \frac{\ker\left(\delta_1 : C^1 \to C^2\right)}{\mathrm{im}\left(\delta_0 : C^0 \to C^1\right)}$$

with the space of isomorphism classes of holomorphic line bundles over the Riemann surface M.

We now present, without proof, some fundamental technical results from the theory of sheaf cohomology. The proofs can be found in Chapter 3 of [14]. We begin with some important finiteness theorems. For brevity, if $\mathcal{O}(L)$ is the sheaf of holomorphic sections of a holomorphic line bundle over M, we shall write $H^p(M, L)$ for the p-th sheaf cohomology group of $\mathcal{O}(L)$.

> **Theorem 4.43** *If M is a compact Riemann surface, and L is a holomorphic line bundle over M, then the cohomology groups $H^p(M, \mathcal{O}(L))$ vanish for $p \geq 2$. Moreover, the cohomology groups $H^p(M, \mathbb{Z})$ and $H^p(M, \mathbb{C})$ vanish for $p \geq 3$.*

This last assertion follows from the fact that the sheaf cohomology groups $H^p(M, \mathbb{Z})$ are isomorphic to the usual singular cohomology groups of M. We also have:

> **Theorem 4.44** *If M is a compact Riemann surface, and L is a holomorphic line bundle over M, then the cohomology groups $H^p(M, \mathcal{O}(L))$ are finite dimensional complex vector spaces for $p = 0, 1$.*

Following the general pattern of a cohomology theory, sheaf maps induce homomorphims of the cochain groups which commute with the coboundary maps, and a short exact sequence

$$0 \longrightarrow \mathcal{T} \longrightarrow \mathcal{S} \longrightarrow \mathcal{Q} \longrightarrow 0$$

of sheaves over M induces a long exact sequence of sheaf cohomology groups:

$$0 \to H^0(M, \mathcal{T}) \to H^0(M, \mathcal{S}) \to H^0(M, \mathcal{Q}) \to H^1(M, \mathcal{T}) \to$$

$$\to H^1(M,\mathcal{S}) \to H^1(M,\mathcal{Q}) \to H^2(M,\mathcal{T}) \to \cdots$$

The connecting homomorphism Δ from $H^p(M,\mathcal{Q})$ to $H^{p+1}(M,\mathcal{T})$ is obtained by the usual diagram chasing: if $f \in C^p(\mathcal{Q})$ represents a class in $H^p(M,\mathcal{Q})$, then by the short exact sequence of sheaves, there exists[11] an element $h \in C^p(\mathcal{S})$ such that $h \mapsto f$ under the induced mapping of cochain groups . Since these induced homomorphisms commute with coboundary maps, we see that the element $\delta_p \cdot h$ of $C^{p+1}(\mathcal{S})$ lies in the kernel of the induced map from $C^{p+1}(\mathcal{S}) \to C^{p+1}(\mathcal{Q})$, so again using the short exact sequence, we can identify h with a unique element $t \in C^{p+1}(\mathcal{Q})$. It is straightforward to check that $\delta_{p+1} \cdot t = 0$, and that t does not depend on our arbitrary choice of element $h \in C^p(\mathcal{S})$, so we can define $\Delta([f]) = [t]$, where the square brackets denote taking cohomology classes.

Let us look at some applications of the long exact sequence of sheaf cohomology groups. Recall the exponential exact sequence for a compact, connected Riemann surface M:

$$0 \longrightarrow \mathbb{Z} \longrightarrow \mathcal{O}_M \xrightarrow{\exp} \mathcal{O}_M^* \longrightarrow 0$$

The corresponding long exact sequence reads

$$0 \to H^0(M,\mathbb{Z}) \to H^0(M,\mathcal{O}_M) \to H^0(M,\mathcal{O}_M^*) \to H^1(M,\mathbb{Z}) \to \tag{4.3}$$
$$\to H^1(M,\mathcal{O}_M) \to H^1(M,\mathcal{O}_M^*) \to H^2(M,\mathbb{Z}) \to 0$$

Now since M is connected, $H^0(M,\mathbb{Z}) \cong \mathbb{Z}$, and since M is topologically a compact orientable surface of genus g, $H^1(M,\mathbb{Z}) \cong \mathbb{Z}^{2g}$ and $H^2(M,\mathbb{Z}) \cong \mathbb{Z}$. Further, since the only holomorphic functions on a compact Riemann surface are the constants, we have that $H^0(M,\mathcal{O}_M) \cong \mathbb{C}$, and $H^0(M,\mathcal{O}_M^*) \cong \mathbb{C}^*$. We also know that the exponentiation map from $\mathbb{C}$ to $\mathbb{C}^*$ is surjective. So by exactness, we see that $H^1(M,\mathbb{Z}) = \mathbb{Z}^{2g}$ injects into $H^1(M,\mathcal{O}_M)$. Hence we can write the interesting part of our long exact sequence as

$$0 \longrightarrow \frac{H^1(M,\mathcal{O}_M)}{\mathbb{Z}^{2g}} \longrightarrow H^1(M,\mathcal{O}_M^*) \longrightarrow \mathbb{Z} \longrightarrow 0 \tag{4.4}$$

Now recall that we can interpret the space $H^1(M,\mathcal{O}_M^*)$ as the space of holomorphic isomorphism classes of holomorphic line bundles over M. The coboundary homomorphism $c : H^1(M,\mathcal{O}_M^*) \to H^2(M,\mathbb{Z})$ in the sequence above is called the **characteristic homomorphism**, and if $L \in H^1(M,\mathcal{O}_M^*)$ represents an such an isomorphism class, the image $c(L)$ is called the **Chern class** of L. The class $c(L) \in H^2(M,\mathbb{Z})$ measures the topological properties of L. Indeed, it can be shown[12] that $c(L) = c(\tilde{L})$ if and only if there is a C^∞ line bundle isomorphism $L \cong \tilde{L}$. It is important to note that this isomorphism need not be holomorphic, however: on Riemann surfaces of genus $g > 0$ there exist holomorphic line bundles with equal Chern class which are not holomorphically isomorphic. Now since a compact Riemann surface M is a compact, orientable 2-manifold, it follows that $H^2(M,\mathbb{Z}) \cong \mathbb{Z}$. Thus, an isomorphism $H^2(M,\mathbb{Z}) \cong \mathbb{Z}$ allows us to consider $c(L)$ as an integer, which we call the degree of L, and denote $\deg(L)$. For more details regarding this identification, the reader is directed to [14], page 101.

We now present a few basic facts about the degree of a holomorphic line bundle. Firstly, since the trivial bundle is represented by identity element of $H^1(M,\mathcal{O}_M^*)$ and the characteristic map is a homomorphism, the degree of the trivial line bundle is zero. Further,

[11]Note that h may not be unique
[12]See[14] page 99

since the tensor product of line bundles corresponds to multiplication of transition functions, we have that $\deg(L\otimes\tilde{L}) = \deg L + \deg \tilde{L}$. One can also show[13] that if a line bundle L over M has a holomorphic section s with zeroes at points $p_1, ..., p_n$, with the zero at p_k having multiplicity m_k, then $\deg L = \sum_k m_k$.

Returning to the exact sequence 4.2, it can be shown [14] that the image of $H^1(M,\mathbb{Z}) \cong \mathbb{Z}^{2g}$ in $H^1(M,\mathcal{O}_M) \cong \mathbb{C}^g$ is a lattice subgroup. Hence it follows that the space of isomorphism classes of degree 0 line bundles on M has the structure of a g-dimensional complex torus $\mathbb{C}^g/\mathbb{Z}^{2g}$. This complex torus is called the **Jacobian**[15] of the Riemann surface Γ, which we shall denote $\mathrm{Jac}(M)$. Observe that via the bijective map given by tensoring with any fixed line bundle of degree d, we can also regard the space of isomorphism classes of line bundles of degree d over M as a complex torus isomorphic to $\mathrm{Jac}(M)$. Consequently, we shall write $\mathrm{Jac}^d(M)$ for this latter space. A complex torus that is isomorphic to the Jacobian of some compact Riemann surface is called an **abelian variety**[16].
In Chapter 5, we will consider a Riemann surface M of genus g and the space $\mathrm{Jac}^{g-1}(M)$ of isomorphism classes of degree $g-1$ holomorphic line bundles over M. The **theta divisor** $\Theta \subset \mathrm{Jac}^{g-1}(M)$ is the subset of $\mathrm{Jac}^{g-1}(M)$ defined by

$$\Theta = \left\{ L \in \mathrm{Jac}^{g-1}(M) | L \cong L_{p_1} \cdots L_{p_{g-1}} \quad \text{for some} \quad p_1, \ldots, p_{g-1} \in M \right\}$$

In words, a degree $g-1$ line bundle represents a point of the theta divisor if and only if it is isomorphic to a product of point bundles over M. Equivalently, a line bundle represents a point of the theta divisor if and only if it admits a non-trivial holomorphic section. The complement of the theta divisor in $\mathrm{Jac}^{g-1}(M)$ is called the **affine Jacobian**[17] of M.

We now state a useful analytic result, the Serre duality theorem, which enables us to efficiently count the dimensions of sheaf cohomology groups:

Theorem 4.45 ***(Serre Duality)*** *Let M be a Riemann surface, L be a line bundle over M, and K be the canonical bundle of M. Then there is an non-degenerate, bilinear pairing of finite dimensional complex vector spaces*

$$\Omega : H^1(M,L) \times H^0(M, K \otimes L^*) \longrightarrow \mathbb{C}$$

which induces an isomorphism of complex vector spaces

$$H^1(M,L) \cong H^0(M, K\otimes L^*)^*$$

.

For a self-contained proof of Serre's duality theorem, the reader is directed to chapter 6 of [14].

Let us apply Serre duality to the long exact sequence of cohomology groups arising from the following short exact sequence of sheaves over a compact connected Riemann surface M of genus g:

$$0 \longrightarrow \mathbb{C} \longrightarrow \mathcal{O}_M \xrightarrow{d} \mathcal{O}(K) \longrightarrow 0$$

[13] See [14], page 103

[14] see [14], p.136

[15] There are other possible definitions of the Jacobian of a compact Riemann surface, but all describe isomorphic complex tori

[16] For an explanation of the terminology, see [13]

[17] It can be identified with an affine algebraic variety; again see [13]

If we let K denote the canonical bundle on M, we obtain the following long exact sequence of complex vector spaces:

$$0 \to H^0(M,\mathbb{C}) \to H^0(M,\mathcal{O}_M) \to H^0(M,K) \to H^1(M,\mathbb{C}) \to$$

$$\to H^1(M,\mathcal{O}_M) \to H^1(M,K) \to H^2(M,\mathbb{C}) \to 0$$

Now as usual we know that $H^0(M,\mathcal{O}_M) \cong \mathbb{C}$, and the connectedness, compactness and orientability of M ensure that $H^0(M,\mathbb{C}) \cong \mathbb{C} \cong H^2(M,\mathbb{C})$. And since M has genus g, we have $H^1(M,\mathbb{C}) \cong \mathbb{C}^{2g}$. Moreover, by the Serre duality theorem, we have

$$\dim H^1(M,\mathcal{O}_M) = \dim H^1(M, M \times \mathbb{C}) = \dim H^0(M,K)^* = \dim H^0(M,K)$$

where $M \times \mathbb{C}$ is the trivial line bundle over M. Using Serre duality again yields

$$\dim H^1(M,K) = \dim H^0(M, K \otimes K^*) = \dim H^0(M,\mathcal{O}_M) = 1$$

We now appeal to the fact that in an exact sequence of vector spaces beginning and ending with the trivial space 0, the alternating sum of the dimensions of the spaces in the sequence vanishes[18]. Summing dimensions in the cohomology long exact sequence, we obtain

$$1 - 1 + \dim H^0(M,K) - 2g + \dim H^0(M,K) - 1 + 1 = 0$$

Hence $\dim H^0(M,K) = g$: that is, the dimension of the space of holomorphic 1-forms on M is equal to the topological genus of M. This is an example of the interplay between analytic and topological invariants in the theory of Riemann surfaces which foreshadows the nature of the powerful Riemann-Roch theorem.

Let us now make some brief remarks about the classification of holomorphic line bundles and vector bundles over the Riemann sphere $\mathbb{P}^1$. Returning to the exact sequence (4.4) and setting $M = \mathbb{P}^1$, we obtain

$$0 \to 0 \to H^1(\mathbb{P}^1, \mathcal{O}^*) \to \mathbb{Z} \to 0$$

Thus there is exactly one isomorphism class of holomorphic line bundles on $\mathbb{P}^1$ of each degree $d \in \mathbb{Z}$. So two line bundles over $\mathbb{P}^1$ are isomorphic if and only if they have the same degree. There is also a very simple classification of the rank n holomorphic vector bundles over $\mathbb{P}^1$ which we will use in Chapter 5.

Theorem 4.46 ***(Birkhoff-Grothendieck)*** *If E is a rank n holomorphic vector bundle over $\mathbb{P}^1$, then*

$$E \cong \mathcal{O}(a_1) \oplus \cdots \mathcal{O}(a_n)$$

for some integers $a_1, \ldots, a_n \in \mathbb{Z}$.

Although the Birkhoff-Grothendieck theorem can be proved using the technology developed in this chapter, for reasons of space we refer the reader to the proof on page 30 of [18]. We often encounter Riemann surfaces exhibited as n-sheeted branched covers of $\mathbb{P}^1$. Via the direct image construction, we can transform the problem of studying a line bundle over such a Riemann surface to the study of a rank n vector bundle on $\mathbb{P}^1$, and these bundles are classified by the Birkhoff-Grothendieck theorem.

[18] This is easy to prove by induction on the length of the sequence

4.6 The Riemann-Roch Theorem

We now state a fundamental theorem in the theory of Riemann surfaces, the Riemann-Roch theorem. Given a holomorphic line bundle L over a compact Riemann surface M, Riemann-Roch relates the bundle's analytic data (the dimensions of the cohomology groups of its sheaf of holomorphic sections) to its topological data (the Chern class of L and the genus of M).

Theorem 4.47 ***(Riemann-Roch)*** *Let L be a holomorphic line bundle over a compact, connected Riemann surface M of genus g. Then*

$$\dim H^0(M,L) - \dim H^1(M,L) = \deg L + 1 - g \qquad (4.5)$$

Proof. We begin by noting that for the trivial line bundle $L = M \times \mathbb{C}$, $\dim H^0(M,L) = \dim H^0(M,\mathcal{O}_M) = 1$, and by Serre duality, $\dim H^1(M,L) = \dim H^0(M,K) = g$. Since the trivial bundle has degree zero, we see that indeed

$$\dim H^0(M,L) - \dim H^1(M,L) = 1 - g = 0 + 1 \cdot (1-g)$$

So the Riemann-Roch formula holds for the trivial line bundle. We now proceed inductively: assuming that the formula holds for a bundle L, we show that it holds for the bundle $L \otimes L_p$, where L_p is the point bundle associated to $p \in M$. Consider the following short exact sequence of sheaves, discussed in Example 4.38:

$$0 \longrightarrow \mathcal{O}(L \otimes L_p^*) \longrightarrow \mathcal{O}(L) \longrightarrow \mathcal{Q} \longrightarrow 0$$

This induces the following long exact sequence of cohomology groups:

$$0 \to H^0(M, L \otimes L_p^*) \to H^0(M,L) \to H^0(M,\mathcal{Q}) \to H^1(M, L \otimes L_p^*) \to$$

$$\to H^1(M,L) \to H^1(M,\mathcal{Q}) \to 0$$

Since the sheaf $\mathcal{Q}$ is a skyscraper sheaf with stalk $\mathbb{C}$ at p, we have that $H^0(M,\mathcal{Q}) = \mathbb{C}$ and $H^1(M,\mathcal{Q}) = 0$. Now since in an exact sequence of vector spaces beginning and ending with $\{0\}$, the alternating sum of the dimensions of the spaces in the sequence vanishes, we have

$$\dim H^0(M, L \otimes L_p^*) - \dim H^1(M, L \otimes L_p^*) = \dim H^0(M,L) - \dim H^1(M,L) - 1$$

Now since the characteristic map is a homomorphism and $\deg L_p^* = -1$, we have that $\deg(L \otimes L_p^*) = \deg L - 1$. From this, and our assumption that

$$\dim H^0(M,L) - \dim H^1(M,L) = \deg L + (1-g) \qquad (4.6)$$

we see that

$$\dim H^0(M, L \otimes L_p^*) - \dim H^1(M, L \otimes L_p^*) = \deg(L \otimes L_p^*) + 1 - g$$

so if the formula holds for L, then it also holds for $L \otimes L_p^*$ for any $p \in M$. An analogous argument shows that if the formula holds for L, then it also holds for $L \otimes L_p$.

In order to prove the theorem for a general line bundle, we show that any holomorphic line bundle L is isomorphic to the tensor product of a finite number of point bundles and their inverses. To this end, let s_q be a holomorphic section of L_q with a single simple zero at q, and consider the short exact sequence of sheaves

$$0 \longrightarrow \mathcal{O}(L) \longrightarrow \mathcal{O}(L \otimes L_q^n) \longrightarrow \mathcal{R} \longrightarrow 0$$

where the map from $\mathcal{O}(L) \longrightarrow \mathcal{O}(L \otimes L_q^n)$ is given by multiplication by the section s_q^n of L_q^n. Recall that the sheaf $\mathcal{R}$ is a skyscraper sheaf whose only non-zero stalk is $\mathcal{R}_q = \mathbb{C}^n$. Hence, $H^0(M, \mathcal{R}) = \mathbb{C}^n$ and $H^p(M, \mathcal{R}) = 0$ for $p \geq 1$. Summing dimensions in the corresponding long exact sequence of cohomology spaces yields

$$\dim H^0(M, L \otimes L_q^n) = \dim H^1(M, L \otimes L_q^n) + \dim H^0(M, L) - \dim H^1(M, L) + n$$

$$\geq \dim H^0(M, L) - \dim H^1(M, L) + n$$

Now since all the vector spaces above are finite dimensional, we can choose n sufficiently large that the right hand side is strictly positive. Hence for such n, $\dim H^0(M, L \otimes L_q^n) \geq 0$, so $L \otimes L_q^n$ has a non-trivial holomorphic section σ, which has some finite number of zeroes at points $p_1, ..., p_k$, each having multiplicity $m_1, ..., m_k$. Consequently, the bundle $L \otimes L_q^n \otimes L_{p_1}^{-m_1} \otimes \cdots L_{p_k}^{-m_k}$ has a nowhere vanishing holomorphic section, and is thus trivial. Therefore $L = L_q^{-n} \otimes L_{p_1}^{m_1} \otimes \cdots L_{p_k}^{m_k}$, so we see that we can indeed write an arbitrary holomorphic line bundle as tensor product of finitely of point bundles and their inverses. But the Riemann-Roch theorem was shown to hold for the trivial bundle, and we have also shown that if it holds for L, it must also hold for L_p^* and L_p. It therefore follows that the Riemann-Roch theorem holds for any holomorphic line bundle L. □

Remark: The Riemann-Roch formula can also be shown to hold for a rank m vector bundle V over a Riemann surface M. We refer the reader to [18] for the proof, which proceeds by induction on the rank of the bundle.

We conclude our discussion of Riemann surfaces by mentioning a consequence of the Riemann-Roch theorem that will prove useful in Chapter 5.

Proposition 4.48 *Let M and N be Riemann surfaces, and let $\pi : M \to N$ be a holomorphic map that exhibits M as an r-fold branched covering of N. If V is a rank n-vector bundle M, g and $\tilde{g}$ are the genera of N and M respectively, and $d, \tilde{d}$ the degrees of V and $\pi_* V$ respectively, then following relationship holds*

$$\tilde{d} - rn(\tilde{g} - 1) = d - n(g - 1)$$

Chapter 5

Lax equations and Riemann surfaces

We shall now apply the machinery assembled in Chapter 4 to give a geometric interpretation of Lax equations with a spectral parameter. We shall see that for the Lax pairs $\dot{A} = [A, B]$ arising from the AKS construction, we can interpret the differential equation as describing the linear motion of a point on a complex torus which is the Jacobian of a compact Riemann surface. The treatment given here is based on the articles [26],[28] as well as the survey [31]. The proof of the linearization theorem that we present follows closely Hitchin's treatment in [18]. A proof can also be found in the Griffiths' paper [13]. Other useful references on the subject include Chapter 5 of Babelon's book [7], as well as the article [3] by Adler and van Moerbeke. The specific example of the Neumann model is treated extensively in Harnad's lecture notes [17]

5.1 Matrix Polynomials and Spectral Curves

Let $M(r,d)$ be the set of all degree d polynomials with coefficients in the space of $r \times r$ complex matrices:

$$M(r,d) = \left\{ A_d z^d + A_{d-1} z^{d-1} + \ldots + A_0 : A_i \in \mathfrak{gl}_r(\mathbb{C}) \right\}$$

If $A(z) \in M(r,d)$ is such a matrix polynomial, its characteristic equation defines an element $P(z,w)$ of the polynomial ring $\mathbb{C}[z,w]$:

$$P(z,w) = \det(A(z) - w\mathbb{1}) = w^r + a_1(z) w^{r-1} + \ldots + a_r(z)$$

Observe that each $a_i \in \mathbb{C}[z]$ satisfies $\deg(a_i) \leq id$.

Definition 5.1 *The affine spectral curve of a matrix polynomial $A(z) \in M(r,d)$ is the affine plane curve*

$$C_0 = \left\{ (z,w) \in \mathbb{C}^2 : P(z,w) = 0 \right\}$$

Consider the complex projective line $\mathbb{P}^1$ with its standard covering by two affine coordinate charts (U_0, z) and $(U_1, \tilde{z})$. At a point p in the intersection $U_0 \cap U_1 \cong \mathbb{C}^*$

these coordinates are related by $\tilde{z}(p) = 1/z(p)$. Recall that with respect to this covering of $\mathbb{P}^1$, the line bundle $\mathcal{O}(d)$ is defined by the transition function

$$t_{01} : U_0 \cap U_1 \to \mathbb{C}^* \ , p \mapsto z^d(p)$$

Let T be the total space of the line bundle $\mathcal{O}(d)$, and $\rho : T \to \mathbb{P}^1$ the bundle projection. Given a matrix polynomial $A(z)$, we will construct the completion of its affine spectral curve inside the complex manifold T. We can cover T by two open sets, $W_0 = \rho^{-1}(U_0)$ and $W_1 = \rho^{-1}(U_1)$. Since the restriction of $\mathcal{O}(d)$ to each U_i is trivial, both W_0 and W_1 are isomorphic to $\mathbb{C}^2$. Take[1] (z, w) to be analytic coordinates on W_0 and $(\tilde{z}, \tilde{w})$ to be analytic coordinates on W_1. On the intersection $W_0 \cap W_1$, these coordinates are related by

$$\tilde{z} = \frac{1}{z}$$
$$\tilde{w} = \frac{w}{z^d}$$

Let us now consider the line bundle $\rho^*\mathcal{O}(rd)$ over T. With respect to the cover $T = W_0 \cup W_1$, this line bundle is defined by the transition function

$$\tau_{01} : W_0 \cap W_1 \to \mathbb{C}^*, \ p \mapsto z^{rd}(p)$$

Let $A(z) \in M(r, d)$ be a matrix polynomial with characteristic polynomial $P(z, w)$. We define a section $\mathcal{P}(A)$ of $\rho^*\mathcal{O}(rd)$ by the following pair of locally defined functions

$$\mathcal{P}(A)_{W_0} : W_0 \to \mathbb{C}, \ (z, w) \mapsto P(z, w)$$

$$\mathcal{P}(A)_{W_1} : W_1 \to \mathbb{C}, \ (\tilde{z}, \tilde{w}) \mapsto \tilde{z}^{rd} P(\tilde{z}^{-1}, \tilde{z}^{-d}\tilde{w})$$

Definition 5.2 ***(Spectral curve of a matrix polynomial)*** *The spectral curve C of a matrix polynomial $A(z) \in M(r, d)$ is the zero set of the section $\mathcal{P}(A)$:*

$$C = \{p \in T : \mathcal{P}(A)(p) = 0\}$$

Suppose that $A(z)$ has characteristic polynomial

$$P(z, w) = \det(A(z) - w\mathbb{1}) = w^r + a_1(z)w^{r-1} + \ldots + a_r(z)$$

Then

$$\tilde{z}^{rd} P(\tilde{z}^{-1}, \tilde{z}^{-d}\tilde{w}) = \tilde{w}^r + \tilde{a}_1(\tilde{z})\tilde{w}^{r-1} + \ldots + \tilde{a}_r(\tilde{z})$$

where $\tilde{a}_i$ is a polynomial in $\tilde{z}$ of degree less than or equal to ir. Since for any fixed value of z or $\tilde{z}$ the coefficients of w^r in $P(z, w)$ and $\tilde{w}^r$ in $\tilde{z}^{rd}P(\tilde{z}, \tilde{w})$ never vanish, the map $\pi = \rho|_C$ exhibits C as an r-fold branched cover of $\mathbb{P}^1$. Hence the spectral curve is a (possibly singular) curve C embedded in the complex surface T, whose 'affine part' $C \cap W_0$ is exactly the affine spectral curve we defined previously. From now on we shall assume that all spectral curves are smooth and all characteristic polynomials are irreducible. Under these assumptions, the spectral curve C associated to a matrix polynomial $A(z) \in M(r, d)$ defines a compact, connected Riemann surface. Note that a branch point $z \in \mathbb{P}^1$ of the covering map $\pi : C \to \mathbb{P}^1$ corresponds to the matrix $A(z)$ having fewer than r distinct eigenvalues, while a branch point at $z = \infty$ corresponds to the leading coefficient A_d of $A(z)$ having fewer than r distinct eigenvalues.

[1]We think of z as the base coordinate and w as the fibre coordinate.

Example 5.3 An affine hyperelliptic curve C_0 is the locus of points

$$C_0 = \{(z, w) \in \mathbb{C}^2 : w^2 = P(z)\}$$

where

$$P(z) = (z - \alpha_1)(z - \alpha_2) \cdots (z - \alpha_n)$$

is a polynomial with n distinct roots $\alpha_1, ..., \alpha_n \in \mathbb{C}$. Following Mumford in [29], we consider matrices of the form

$$A(z) = \begin{pmatrix} V(z) & W(z) \\ U(z) & -V(z) \end{pmatrix}$$

where $U(z)$ and $W(z)$ are monic polynomials of degree $d-1$ and d respectively, and $V(z)$ is of degree $\leq d-2$ for some fixed positive integer d. The characteristic polynomial of $A(z)$ is given by

$$P(z, w) = w^2 - (V^2 + UW)(z) =: w^2 - F(z)$$

Assume that the polynomials U, W, V are generic in the sense that $F(z)$ has no repeated roots:

$$F(z) = \prod_{i=1}^{2d-1} (z - \alpha_i), \ , \alpha_i \neq \alpha_j \ \text{ if } \ i \neq j$$

In this case, the affine spectral curve of $A(z)$ is hyperelliptic. Following the method outlined above, we think of

$$P(z, w) = w^2 - \prod_{i=1}^{2d-1} (z - \alpha_i)$$

as being the local representative of a global section of $\rho^* \mathcal{O}(2d)$ in the local trivialization $(z, w) \in V_1 \times \mathbb{C}$. It follows that in the other trivialization $(\tilde{z}, \tilde{w}) \in V_2 \times \mathbb{C}$, this section is represented by the local function

$$\tilde{\mathcal{P}}(\tilde{z}, \tilde{w}) = \tilde{w}^2 - \tilde{z} \prod_{i=1}^{2d-1} (1 - \alpha_i \tilde{z})$$

The spectral curve C of $A(z)$ is the zero set in $\mathcal{O}(d)$ of this section. We observe that there is only one 'point at infinity', the point $(\tilde{z} = 0, \tilde{w} = 0)$. This is readily seen to be a smooth point. It is, however, an order 2 branch point of the covering $\pi : C \to \mathbb{P}^1$. We also note that $\pi : C \to \mathbb{P}^1$ has $2d-1$ order 2 branch points $\{(z = \alpha_i, w = 0)\}_{i=1}^{2d-1}$ at finite distance. Hence $\pi : C \to \mathbb{P}^1$ has $2d$ branch points of order 2, and thus by the Riemann-Hurwitz formula we compute its genus g to be

$$g = d - 1$$

We now study some general properties of spectral curves. Let $T = W_0 \cup W_1$ be the open cover of the total space of the line bundle $\mathcal{O}(n)$ we defined previously. We begin by making an important observation about this cover:

Lemma 5.4 *The cover $T = W_0 \cup W_1$ is a Leray open cover with respect to any coherent analytic sheaf $\mathcal{F}$ over T.*

Proof. We have seen that both W_0 and W_1 are isomorphic to $\mathbb{C} \times \mathbb{C}$. So W_0, W_1 is a cover of T by Stein manifolds. By Cartan's Theorem B, such a covering is a Leray cover with respect to any coherent analytic sheaf on T. □

We will exploit this fact to give explicit descriptions of the cohomology groups of sheaves on T. In particular, if $\mathcal{O}_T$ denotes the sheaf of holomorphic functions on T, we can compute $H^1(T, \mathcal{O}_T)$. By definition, an element of $H^1(T, \mathcal{O}_T)$ is represented by a holomorphic function defined on $U_1 \cap U_2 \cong \mathbb{C}^* \times \mathbb{C}$, given by a Laurent series of the form

$$\sum_{j=-\infty}^{\infty} \sum_{k=0}^{\infty} a_{ij} z^j w^k$$

modulo the functions which extend holomorphically to U_1, which have the form

$$\sum_{j=0}^{\infty} \sum_{k=0}^{\infty} b_{ij} z^j w^k$$

and modulo those which extend to U_2, which are of the form

$$\sum_{j=0}^{\infty} \sum_{k=0}^{\infty} c_{ij} z^{-dk-j} w^k$$

Hence we find that

Proposition 5.5 *$H^1(T, \mathcal{O}_T)$ is spanned by the monomials*

$$\left\{ w^i z^j \ , \ i > 0 \ , id < j < 0 \right\}$$

By definition of the spectral curve C associated to $A(z)$, multiplication by the defining section $\mathcal{P}(A)$ gives a short exact sequence of sheaves

$$0 \to \rho^* \mathcal{O}(-rd) \to \mathcal{O}_T \to \mathcal{O}_C \to 0$$

where we identify the quotient sheaf with $\mathcal{O}_C$, the sheaf of holomorphic functions on C. Consider the following portion of the corresponding long exact sequence in cohomology:

$$H^1(T, \mathcal{O}_T) \to H^1(C, \mathcal{O}_T) \to H^2(T, \rho^* \mathcal{O}(-rd))$$

Since there is Leray cover for T with respect to $\rho^* \mathcal{O}(-rd)$ which has only two open sets, it follows that $H^2(T, \rho^* \mathcal{O}(-rd)) = 0$. Hence the map from $H^1(T, \mathcal{O}_T)$ to $H^1(C, \mathcal{O}_C)$ is surjective. So any element of $H^1(C, \mathcal{O}_C)$ can be described as an element of $H^1(T, \mathcal{O}_T)$ modulo the relation $\det(A(z) - w\mathbb{1}) = 0$ defining the spectral curve. Since we can use this relation to eliminate any powers of w higher than w^{r-1}, we observe that the monomials $w^i z^j \ , \ 0 < i \leq r-1 \ , id < j < 0$ form a basis for $H^1(C, \mathcal{O}_C)$. Noting that

$$(d-1) + (2d-1) + \cdots + (r-1)(d-1) = \frac{1}{2}(r-1)(dr-2)$$

we have:

Proposition 5.6 *Let C be the spectral curve associated to the matrix polynomial $A(z) \in M(r,d)$. If C is smooth, its genus is given by $g = \frac{1}{2}(r-1)(dr-2)$*

We will be also be interested pullback bundle $\pi^*\mathcal{O}(d)$ over the spectral curve C. With respect to the covering of C by the two charts $V_1 = C \cap U_1$ and $V_2 = C \cap U_2$, then this bundle is defined by the transition function $t_{12} = z^n(p),\ p \in V_0 \cap V_1$. We can define a section w of this bundle by the two functions $w_{V_1} = w$, $w_{V_2} = \tilde{w}$. This section contains the data of the eigenvalues of the matrix polynomial $A(z)$. Of course, these eigenvalues cannot be expressed as single-valued functions of a parameter on $\mathbb{P}^1$: there are, after all, generically r eigenvalues corresponding to each value of z. However on the r-fold cover of $\mathbb{P}^1$ given by the spectral curve, we find that we can indeed represent the eigenvalues as a single-valued global section of the line bundle $\pi^*\mathcal{O}(n)$.
We conclude our discussion of spectral curves by noting a consequence of the smoothness assumption:

Definition 5.7 ***(Regular matrix)*** *A matrix $M \in GL_r(\mathbb{C})$ is regular if the following equivalent conditions hold:*

1. *The minimal polynomial of M is equal to its characteristic polynomial*
2. *The conjugacy class of M in $GL_r(\mathbb{C})$ is of maximal dimension $r^2 - r$*
3. *There exists a vector $v \in \mathbb{C}^r$ such that $v, Mv, M^2v, \dots$ span $\mathbb{C}^r$*
4. *All the eigenspaces of M are of dimension 1*

Proposition 5.8 *Let $A(z) \in M(r,d)$ be a matrix polynomial with smooth spectral curve X. Then for all fixed $a \in \mathbb{P}^1$, the matrix $A(a)$ is regular (for $z = \infty$, this means the leading matrix A_d is regular).*

Proof. As usual, we write

$$A(z) = A_d z^d + A_{d-1} z^{d-1} + \ldots + A_0$$

Now suppose that at $z_0 \in \mathbb{C}$ the matrix $A(z_0)$ is not regular. Then there exists w_0 such that $(z_0, w_0) \in C_0$ and $\operatorname{rank}(A(z_0) - w_0.\mathbb{1}) \leq r-2$. Let $\Delta_{ij}(z,w)$ be the (i,j)-th minor of the matrix $A(z) - w\mathbb{1}$. Since $A(z_0)$ is of rank $\leq r-2$, $\Delta_{ij} = 0$ for all i,j. Set $A_{ij}(z) := (A(z))_{ij}$. Then computing the partial derivatives

$$\partial_w P(z,w)|(z_0,w_0) = \sum_{i=1}^{r} \Delta_{ii}(z_0,w_0) = 0$$

$$\partial_z P(z,w)|(z_0,w_0) = \sum_{i,j=1}^{r} \frac{\partial A_{ij}}{\partial z}(z_0)\Delta_{ij}(z_0,w_0) = 0$$

shows that z_0, w_0 is a singular point of C. Applying similar considerations to

$$\tilde{z}^{rd} P(\tilde{z}^{-1}, \tilde{z}^{-d}\tilde{w}) = \det\left(\tilde{z}^d A\left(\frac{1}{\tilde{z}}\right) - \tilde{w}\right)$$

and noting that

$$\tilde{z}^d A\left(\frac{1}{\tilde{z}}\right)\Bigg|_{\tilde{z}=0} = A_d$$

we see that if A_d fails to be regular, then there exists $\tilde{w}_0$ such that $(\tilde{z}, \tilde{w}) = (0, \tilde{w}_0)$ is a singular point of C. This proves the proposition. □

5.2 The Eigenvector Bundle

Suppose that $A(z) \in M(r,d)$ is a matrix polynomial whose spectral curve C is smooth. By Proposition 5.8, we know that for all $z \in \mathbb{P}^1$, the eigenspaces of $A(z)$ are one-dimensional. We will now study these eigenspaces in more detail. Let $\Psi(p) = (\psi_1(p), \ldots, \psi_r(p))$ be an eigenvector associated to the point $p \in C$: if $p = (z,w)$, then $\Psi(p)$ is an eigenvector of the matrix $A(z)$ with eigenvalue w. Let $e_1, \ldots, e_r$ be the standard basis for $\mathbb{C}^r$ and $\langle , \rangle$ the inner product defined by declaring the e_i an orthonormal basis. Define r open subsets of C by

$$U_i = \{p \in C : \exists\ v \in \ker(A(z) - w\mathbb{1}) \text{ such that } \langle v, e_i \rangle \neq 0\}$$

Let us choose to normalize the eigenvector $\Psi(p)$ by setting $\psi_1 \equiv 1$. This defines r functions $\psi_1 \equiv 1, \psi_2, \ldots \psi_r$ on the open set U_1.

Proposition 5.9 *The components $(\psi_1, \psi_2, \ldots \psi_r)$ of the normalized eigenvector $\Psi(p)$ define meromorphic functions on the spectral curve C*

Proof. At each point $p \in U_1$ there is a unique eigenvector Ψ of A normalized by $\psi_1 = 1$. The unnormalized components of this eigenvector will be suitable minors of the matrix $A(z) - w\mathbb{1}$ and hence will depend polynomially on the two meromorphic functions z, w on C. After dividing by the first component of this vector, all other components depend rationally on z and w and are thus define meromorphic functions on C. $\square$

We will use this set of meromorphic functions to construct a line bundle over C which encodes the data of the eigenspaces of $A(z)$:

Proposition 5.10 *Suppose that $A(z) \in M(r,d)$ is a matrix polynomial whose spectral curve C is smooth. Then there exists a unique line bundle $\mathcal{L}$ over C, embedded as a subbundle of $C \times \mathbb{C}^m$, whose fibre over the point $p \in C$ is the one-dimensional eigenspace of the matrix $A(z(p))$ corresponding to the eigenvalue $w(p)$.*

Proof. When restricted to the set $U_1 \subset C$ the r functions $(1, \psi_2, \ldots \psi_r)$ are holomorphic. Hence they define a holomorphic map

$$\Psi : U_1 \to \mathbb{P}^{r-1} \quad p \mapsto [1 : \psi_2 : \cdots : \psi_r]$$

This map can be extended to a holomorphic map $\Psi : C \to \mathbb{P}^{r-1}$ as follows. Let p be a point where at least one of the ψ_i has a pole. Denote by k the maximum of the orders of the poles of the ψ_i at p; by assumption $k \geq 1$. Let (V, x) be a local analytic coordinate chart on C centered at p. Then we define a holomorphic map $V \to \mathbb{P}^1$

$$x \to [x^k : x^k\psi_2 : \ldots : x^k\psi_r]$$

By definition of the homogeneous coordinates on $\mathbb{P}^1$ this map agrees with the map Ψ on $V \cap U_1$. Applying this procedure at each point $p \in C - U_1$ defines a unique extension of Ψ to a holomorphic map $\Psi : C \to \mathbb{P}^{r-1}$. Now let $\mathcal{T}$ be the tautological bundle over $\mathbb{P}^{m-1}$; that is, the subbundle of $\mathbb{P}^{m-1} \times \mathbb{C}^m$ whose fiber over the point $p \in \mathbb{P}^{m-1}$ is the one-dimensional subspace of $\mathbb{C}^m$ represented by p. Our line bundle $\mathcal{L}$ is then simply $\Psi^*\mathcal{T}$, the pullback by Ψ of the tautological bundle. $\square$

It is in fact the dual bundle to $\mathcal{L}$ that will play a fundamental role in our discussion:

Definition 5.11 ***(Eigenvector bundle)*** *Let $A(z) \in M(r,d)$ be a matrix polynomial with smooth spectral curve C and let $\mathcal{L}$ be the line bundle constructed in Proposition 5.10. The eigenvector bundle associated to $A(z)$ is the line bundle $L = \mathcal{L}^*$.*

We will now give a description of the line bundles $\mathcal{L}$ and L in terms of local trivializations and transition functions. Since $\varrho : \mathcal{L} \to C$ is a subbundle of $C \times \mathbb{C}^r$, points $x \in \varrho^{-1}(p)$ determine vectors in $\mathbb{C}^m$. We define local trivializations $\varphi_i : \varrho^{-1}(U_i) \to U_i \times \mathbb{C}$ as follows:

$$\varphi_i(x) = (\varrho(x), \langle x, e_i \rangle)$$

The map φ is evidently analytic and fibrewise linear, and since $\langle x, e_i \rangle \neq 0$ for $\pi(x) \in U_i$, φ is a bijection and hence defines a local trivialization of $\mathcal{L}$ over U_i.
We obtain transition functions t_{ij} on each intersection $U_i \cap U_j$ by

$$t_{ij} = \varphi_i \circ \varphi_j^{-1}$$

Explicitly, if $p \in C$ we have

$$\begin{aligned} t_{ij}(p) &= \frac{\langle v, e_i \rangle}{\langle v, e_j \rangle}, \ v \in \varrho^{-1}(p) \\ &= \frac{\psi_i}{\psi j} \end{aligned}$$

Note that these functions are well-defined: they are independent of the choice of vector $v(x)$ in the eigenspace $\varrho^{-1}(x)$. By definition, the dual bundle L is described by the transition functions

$$f_{ij}(x) = \frac{1}{t_{ij}(x)}$$

The eigenvector bundle L has r natural holomorphic sections $s_1, \ldots, s_r$ which correspond to projection onto the linear coordinates of $\mathbb{C}^r$. With respect to the trivialization of L we have constructed, the section s_k is represented by the following collection of holomorphic functions defined on the U_i

$$s_k|_{U_i} = \frac{\psi_k}{\psi_i}$$

where the ψ_j are suitable restrictions of the meromorphic normalized eigenvector components we constructed previously. Since for all but finitely many $z \in \mathbb{P}^1$ the eigenvectors of $A(z)$ span $\mathbb{C}^m$, the sections s_i are linearly independent. By their construction in in terms of components of eigenvectors, the sections s_i have the property :

Lemma 5.12

The standard basis sections s_i and the meromorphic function w on C satisfy

$$ws_i = \sum_{j=1}^{r} A_{ij}(z) s_j$$

We now determine some basic properties of the line bundle L. We will use the abbreviation

$$L(k) := L \otimes \pi^* \mathcal{O}(k), \ k \in \mathbb{Z}$$

Proposition 5.13 *Suppose that $A(z) \in M(r,d)$ is a matrix polynomial whose spectral curve C is smooth and has r distinct points above $z = \infty$. Then the eigenvector bundle L corresponding to $A(z)$ satisfies* $\dim H^0(C, L(-1)) = 0$.

Proof. Let $P_\infty = (P_1, \ldots, P_r)$ be the r distinct points above $z = \infty$, and write $C_0 = C - \{P_1, \ldots P_r\}$ for the affine spectral curve. Recall that C_0 is the affine variety corresponding to the ideal in $\mathbb{C}[z,w]$ generated by the characteristic polynomial $P(z,w)$. Let $R_0 = \mathbb{C}[z,w]/\langle P(z,w)\rangle$ be the coordinate ring of the affine variety C_0, and denote by R the polynomial ring $\mathbb{C}[z]$. Let V be the subspace of $H^0(C,L)$ generated by the sections $s_1, \ldots, s_r$. We begin the proof with a simple observation: if $v \in V$ and $zv \in H^0(C,L)$, then $v = 0$. Indeed if $zv \in H^0(C,L)$, then v must vanish at each point over infinity. But since the eigenspaces at these points span $\mathbb{C}^r$, this implies $v = 0$.

Let $\mathcal{M}$ be the space of meromorphic sections of L whose restrictions to C_0 are holomorphic. Since polynomials in z are holomorphic on C_0, $\mathcal{M}$ has the structure of an R_0-module. We claim that the natural map $r : V \otimes R \to \mathcal{M}$ given by multiplication of sections by functions is surjective. Observe that by Lemma 5.12, $r(V \otimes R)$ is an R_0-module, since the collection of linear coordinates $(s_1, \ldots, s_r)$ satisfy $ws_i = \sum_{j=1}^r A_{ij}(z)s_j$. Now suppose for a contradiction that $r(V \otimes R)$ is a proper R_0 submodule of $\mathcal{M}$. Then there exists a maximal ideal $\mathfrak{m}$ in R_0 such that $r(V \otimes R) \subset \mathfrak{m}\mathcal{M}$. Associated to any maximal ideal in the coordinate ring R_0 is a point $p \in C_0$ such that $\mathfrak{m} = \{g \in R_0 : g(p) = 0\}$. Therefore every element of $r(V \otimes R)$ vanishes at p. This is absurd, since there must exist a linear coordinate which does not vanish on the eigenspace of $A(z)$ corresponding to the point $p = (z,w)$. Thus we see that r is indeed surjective.

We conclude the argument by showing that $H^0(C,L) = V$. By the surjectivity of r, any element of $\mathcal{M}$, and thus any element $t \in H^0(C,L)$, may be written $t = \sum_{i=0}^m v_i z^i$. Suppose that $v_m \geq 0$. Then $v_m z = sz^{1-m} - \sum_{i \leq 0} v_{i+m-1} z^i$. Note that the right hand side of this expression is manifestly holomorphic at $z = \infty$. Consequently $v_m z \in H^0(C,L)$, which shows $v_m = 0$. Thus $H^0(C,L) = V$ as claimed. It now follows readily that $H^0(C, L(-1)) = \{0\}$, since any non-trivial section of $L(-1)$ would give rise to a section of L vanishing at each point P_i over infinity.

□

Corollary 5.14 *Under the assumptions of Proposition 5.13, the dimension of the space of holomorphic sections of L is* $\dim H^0(C,L) = r$.

Proof. Suppose that $p \in \mathbb{P}^1$ is not a branch point of the r-fold covering $\pi : C \to \mathbb{P}^1$, and let s_p be a section of the bundle $\mathcal{O}(1)$ over $\mathbb{P}^1$ with a single simple zero at p. Then the composite $s_p \circ \pi$ defines a section of the pullback bundle $\pi^*\mathcal{O}(1)$ over C which has r distinct zeroes. Now consider the short exact sequence of sheaves

$$0 \to \mathcal{O}(L(-1)) \to \mathcal{O}(L) \to \mathcal{Q} \to 0$$

where the map from $\mathcal{O}(L(-1))$ to $\mathcal{O}(L)$ is multiplication by the section $s_p \circ \pi$. The corresponding long exact sequence in cohomology reads

$$0 \to H^0(C, L(-1)) \to H^0(C,L) \to \mathbb{C}^r \to H^1(C, L(-1)) \to H^1(C,L) \to 0 \qquad (5.1)$$

By Proposition 5.13, $\dim H^0(C, L(-1)) = 0$, so that $H^0(C, L)$ injects into $\mathbb{C}^r$. Thus $\dim H^0(C, L) \leq r$. But we have already explicitly constructed r linearly independent elements of $H^0(C, L)$, the sections $\{s_i\}$. Hence $\dim H^0(C, L) = r$, and the $\{s_i\}$ form a basis. □

We will refer to the basis s_i as the standard basis for $H^0(C, L)$. We can also compute the degree of such an eigenvector bundle:

Proposition 5.15 *Under the assumptions of Proposition 5.13, the degree of the eigenvector bundle over a spectral curve C of genus g is given by*

$$\deg L = r + g - 1$$

Proof. Counting dimensions in the long exact sequence (5.1), we observe that

$$\dim H^1(C, L(-1)) = \dim H^1(C, L)$$

. By Serre duality, we have

$$H^1(C, L(-1)) \cong H^0(C, K(1)L^{-1})^*, \qquad H^1(C, L) \cong H^0(C, KL^{-1})^*$$

We claim that $\dim H^1(C, L) = \dim H^0(C, KL^{-1}) = 0$. Consider the injective maps from $H^0(C, KL^{-1})$ to $H^0(C, K(1)L^{-1})$ given by multiplication by some non-trivial section of the pullback bundle $\pi^*\mathcal{O}(1)$. Observe that given any point $p \in C$, there exists a section of $\pi^*\mathcal{O}(1)$ which vanishes at p; this is a consequence of the fact that any two point bundles on $\mathbb{P}^1$ are isomorphic. Since

$$\dim H^0(C, KL^{-1}) = \dim H^0(C, K(1)L^{-1})$$

any injective map between these spaces is also surjective. But this means that every section of $K(1)L^{-1}$ vanishes at each point $p \in C$, which implies

$$\dim H^0(C, K(1)L^{-1}) = 0 = \dim H^0(C, KL^{-1}) = H^1(C, L)$$

So applying the Riemann-Roch theorem to L,

$$\dim H^0(C, L) - \dim H^1(C, L) = r - 0 = \deg L + 1 - g$$

We conclude that $\deg(L) = r + g - 1$ as claimed. □

5.3 The Jacobian of a Spectral Curve

Suppose that C is a spectral curve of genus g with r distinct points above $z = \infty$. Denote by M_P the subset of $M(r, d)$ consisting of matrix polynomials whose spectral curve is C. Let $\mathrm{Jac}^{g-1}(C)$ be the space of isomorphism classes of holomorphic line bundles on C of degree $g - 1$. Recall from Chapter 4 that $\mathrm{Jac}^{g-1}(C)$ has the structure of a g-dimensional complex torus which is isomorphic to the Jacobian of C. From now on, we will refer to $\mathrm{Jac}^{g-1}(C)$ simply as the Jacobian of C. We also recall that the theta divisor $\Theta \subset \mathrm{Jac}^{g-1}(C)$ is the subset of the Jacobian consisting of line bundles admitting a non-trivial holomorphic section:

$$\Theta = \left\{ B \in J^{g-1}(C) : H^0(C, B) \neq 0 \right\}$$

Given a matrix polynomial $A(z) \in M_P$, by Proposition 5.13 its eigenvector bundle L has degree $r + g - 1$ and satisfies $\dim H^0(C, L(-1)) = 0$. Since the pullback bundle $\pi^*\mathcal{O}(1)$ has a section with r distinct simple zeroes, it has degree r, which implies that the bundle $L(-1)$ must have degree $g - 1$. Therefore the construction of the eigenvector bundle L from a matrix $A(z) \in M_P$ defines a map

$$l : M_P \to J^{g-1}(C) - \Theta, \quad A(z) \mapsto L(-1)$$

Using the direct image construction, we can also obtain matrix polynomials from a line bundle $B \in \mathrm{Jac}^{g-1}(C) - \Theta$.

Proposition 5.16 *Let $A(z), \tilde{A}(z) \in M_P$, and let $L, \tilde{L}$ be the corresponding eigenvector bundles. Suppose there exists a constant matrix $P \in GL_r(\mathbb{C})$ such that $A(z) = P\tilde{A}(z)P^{-1}$. Then $L \cong \tilde{L}$.*

Proof. Let s_i be the standard basis for $H^0(C, L)$ and let t_i be the corresponding basis for $H^0(C, \tilde{L})$. Consider the section $\sigma = \sum_j P_{1j} t_j$ of $\tilde{L}$. We claim the divisor of σ is the same of that of the section s_1 of L. Let $p \in C$ be contained in a local trivialization U. Denote by $\mathbf{s}$ the vector $(s_1|_U, \ldots s_r|_U)$ where $s_i|_U$ is the local holomorphic function representing the section s_i in the trivialization U; define a vector $\mathbf{t}$ similarly. Since the eigenspaces of $A(z)$ are one-dimensional for each p there exists a function $k(p)$ such that $\mathbf{s} = k(p) \cdot P\mathbf{t}$, that is,

$$s_i|_U = k(p) \sum_j P_{ij} t_j|_U$$

By definition of the bases s_i, t_i in terms of the components of normalized eigenvectors, $k(p)$ must be a non-vanishing holomorphic function. But then since $s_1 = k(p)\sigma$, we see that the holomorphic sections σ and s_1 have the same divisor, from which it follows that $L \cong \tilde{L}$.

□

Proposition 5.17 *Let L be a holomorphic line bundle over C such that $L(-1) \in J^{g-1}(C) - \Theta$. Then the direct image $\pi_* L$ is isomorphic to the trivial rank r vector bundle E over $\mathbb{P}^1$.*

Proof. Since $L(-1) \in J^{g-1}(C) - \Theta$, we have $\deg(L(-1)) = g - 1$ and $H^0(C, L(-1)) = 0$. It follows that $\deg L = r + d - 1$. By Proposition 4.41, we can identify the direct image $\pi_* L$ with a rank r vector bundle E over $\mathbb{P}^1$. Recall that from Proposition 4.48 that if $\pi : M \to N$ is an r-fold branched covering and V is a rank n-vector bundle over M, then

$$\tilde{d} - rn(\tilde{g} - 1) = d - n(g - 1)$$

where g and $\tilde{g}$ are the genera of N and M respectively, and $d, \tilde{d}$ the degrees of V and $\pi_* V$ respectively. In our case this formula yields deg $\pi_* L(-1) = 0$. So by the Birkhoff-Grothendieck theorem,

$$\pi_* L(-1) = E \cong \mathcal{O}(a_1) \oplus \cdots \oplus \mathcal{O}(a_k)$$

with $\sum_i a_i = 0$. However, by the projection formula in Proposition 4.41, we have

$$\pi_* L \cong E \otimes \mathcal{O}(-1) \cong \mathcal{O}(a_1 - 1) \oplus \cdots \oplus \mathcal{O}(a_r - 1)$$

By definition of the direct image, $H^0(C, L) = H^0(\mathbb{P}^1, \pi_* L) = 0$. Hence each a_i must satisfy $a_i \leq 0$. Since $\sum_i a_i = 0$ this implies $a_i = 0$ for all i, and thus E is isomorphic to the trivial bundle over $\mathbb{P}^1$. □

Observe that by the projection formula, this proposition implies that $\pi_* L(d) \cong E(d)$, where $E(d) = E \otimes \mathcal{O}(d)$.

Recall from our discussion of spectral curves in Section 5.1 that with respect to the cover $C = V_1 \cup V_2$, the line bundle $\pi^*\mathcal{O}(d)$ can be described by the cocycle z^d. This bundle has a canonical section w, which is described by the pair of local holomorphic functions $(w, \tilde{w})$. Multiplication by w defines a map $W : H^0(C, L) \to H^0(C, L(d))$. So by definition of the direct image, W induces a map $\tilde{W} : H^0(\mathbb{P}^1, E) \to H^0(\mathbb{P}^1, E(d))$.

Now recall that $H^0(\mathbb{P}^1, E) \cong \mathbb{C}^m$. Let us fix a specific isomorphism by choosing a basis $(e_1, \ldots, e_m)$ for $H^0(\mathbb{P}^1, E)$. Recall that $H^0(\mathbb{P}^1, \mathcal{O}(d))$ can be identified with the space of polynomials of degree less than or equal to d. Therefore our choice of basis for $H^0(\mathbb{P}^1, E)$ determines a basis for $H^0(\mathbb{P}^1, E(d))$ of the form

$$\mathcal{B} = \{e_i \otimes z^k : 1 \leq i \leq r, 0 \leq k \leq d\}$$

Thus $\tilde{W}$ defines a matrix $A(z) \in M(r, d)$:

$$\tilde{W}(e_i) = \sum_{j=1}^{r} A_{ij}(z) \otimes e_j$$

Observe that passing to a different choice of basis for $H^0(\mathbb{P}^1, E)$ corresponds to conjugation of $A(z)$ by a matrix $P \in GL_r(\mathbb{C})$. Now by definition of the section w, the matrix $A(z)$ satisfies $P(z, A(z)) = 0$. Since the spectral polynomial $P(z, w)$ was assumed to be irreducible, the Cayley-Hamilton theorem implies that the characteristic polynomial of $A(z)$ is equal to $P(z, w)$, so that $A(z) \in M_P$.

Hence we have defined a map from $J^{g-1}(C) - \Theta$ to the set of conjugacy classes of matrix polynomials with spectral curve C. We claim that this map is an inverse to the eigenvector mapping:

Theorem 5.18 ***(Beauville)*** *Suppose that the spectral curve associated to P is smooth with r distinct points above $z = \infty$. Then the group $GL_r(\mathbb{C})$ acts on M_P by conjugation, and there is a bijection between the quotient space and the affine Jacobian $J^{g-1}(C) - \Theta$ of C.*

Proof. Let $A(z)$ represent a conjugacy class in M_P, and let L_A be the image of this conjugacy class under the eigenvector mapping. Let $\{s_i\}$ be the standard basis of sections of L_A. The bundle L_A can be described by the two local trivialisations $V_1 \times \mathbb{C}$ and $V_2 \times \mathbb{C}$ which are related by a transition function $f_{12} : V_1 \cap V_2 \to \mathbb{C}^*$. In these trivialisations, the basis s_i of $H^0(C, L)$ is represented by pairs $(b_i, \tilde{b}_i)$ where $b_i : V_1 \to \mathbb{C}, \tilde{b}_i : V_2 \to C$ and $b_i = f_{12}\tilde{b}_i$. By Lemma 5.12, we have that

$$wb_i = \sum_j A_{ij}(z) b_j$$

$$\tilde{w}\tilde{b}_i = \sum_{j=1}^{r} \tilde{z}^d A_{ij}(z) \tilde{b}_j$$

Hence we recover the matrix $A(z)$ as the matrix with respect to the basis s_i of the map $H^0(\mathbb{P}^1, E) \to H^0(\mathbb{P}^1, E(d))$ induced by multiplication by the section w.

Conversely, suppose we begin with a line bundle $B \in J^{g-1} - \Theta$. Denote $\mathcal{L} = B(1)$, and let us assume again that $\mathcal{L}$ is trivialized over the two sets V_1, V_2. A choice of basis of $H^0(C, \mathcal{L})$ is represented by r sets of pairs $(b_i, \tilde{b}_i)$ where $b_i : V_1 \to \mathbb{C}, \tilde{b}_i : V_2 \to C$ and $b_i = f_{12}\tilde{b}_i$. The matrix associated to $\mathcal{L}$ and this choice of basis is defined by $wb_i = \sum_j A_{ij}(z)b_j$. Observe that since $\pi_*(L)$ is trivial, at each point $p \in C$ at least one b_i satisfies $b_i(p) \neq 0$. Now let L_A be the eigenvector bundle corresponding to the matrix $A(z)$, and let s_i be the usual canonical basis of sections. By an identical argument to the one used in the proof of Proposition 5.17, we conclude that the divisors of b_i and s_i are equal. Hence $\mathcal{L} \cong L_A$, which proves that the maps are indeed mutual inverses.

□

5.4 Geometric Interpretation of Lax Equations

Let $A(z)$ and $B(z)$ be matrix polynomials, and consider the Lax equation

$$\frac{dA(z;t)}{dt} = [A(z;t), B(z;t)] \ , \ A(z;0) = A_0 \tag{5.2}$$

Recall from Proposition 0.1 in the Introduction that the coefficients of the characteristic polynomial of $A(z)$ are preserved by the time evolution of this system. In particular, this means that if we fix an initial condition $A(z;0)$ corresponding to a characteristic polynomial P and spectral curve C, the Lax equation (5.2) describes a dynamical system on M_P, the subset of $M(r,d)$ consisting of matrix polynomials whose spectral curve is C. Thus, we can use the machinery assembled in this chapter to study the system's time evolution. By Beauville's theorem, we can identify M_P with the affine Jacobian $J^{g-1}(C) - \Theta$ of C. Hence the Lax equation (5.2) describes a dynamical system on an open subset of the complex torus

$$\text{Jac}^{g-1}(C) \cong \frac{H^1(C, \mathcal{O}_C)}{H^1(C, \mathbb{Z})} \cong \frac{\mathbb{C}^g}{\mathbb{Z}^{2g}}$$

We say that a curve $c(t)$ in $\text{Jac}^{g-1}(C)$ represents linear motion on the Jacobian if $c(t)$ is the image of a line $t \mapsto \beta_0 + t\beta$ in $H^1(C, \mathcal{O}_C)$ under the covering map exp : $H^1(C, \mathcal{O}_C) \to \text{Jac}^{g-1}(C)$.

Remarkably, for a familiar form of the matrix $B(z)$, we shall see that the Lax equation (5.2) describes linear motion on $\text{Jac}^{g-1}(C)$.

As we saw in Section 3.3, applying the AKS construction to the loop algebra $\tilde{\mathfrak{gl}}_m(\mathbb{C})$ yields Hamiltonian dynamical systems whose equations of motion take the form

$$\frac{d}{dt}A(z;t) = \Big[\big(Q(A(z), z^{-1})\big)_+ , A(z;t)\Big] \tag{5.3}$$

where $A(z)$ is an $m \times m$ matrix polynomial, $Q(x, z)$ is a polynomial in x and z^{-1}, and the subscript $+$ denotes taking the holomorphic part of the matrix Laurent series $Q(A(z), z^{-1})$. Now let C be the invariant spectral curve of the matrix polynomial $A(z;t)$ corresponding to some fixed initial condition $A(z;0)$. Recall from our discussion of spectral curves that $H^1(C, \mathcal{O}_C)$ can be explicitly described as the complex vector space spanned by the monomials

$$\left\{w^i z^j \ , \ i > 0 \ , id < j < 0\right\}$$

As we will see, the matrix polynomial $B(z) = \big(Q(A(z), z^{-1})\big)_+$ is related to a cohomology class $\beta(z, w) \in H^1(C, \mathcal{O}(C))$, which determines the direction of a linear flow on $\text{Jac}^{g-1}(C) - \Theta$:

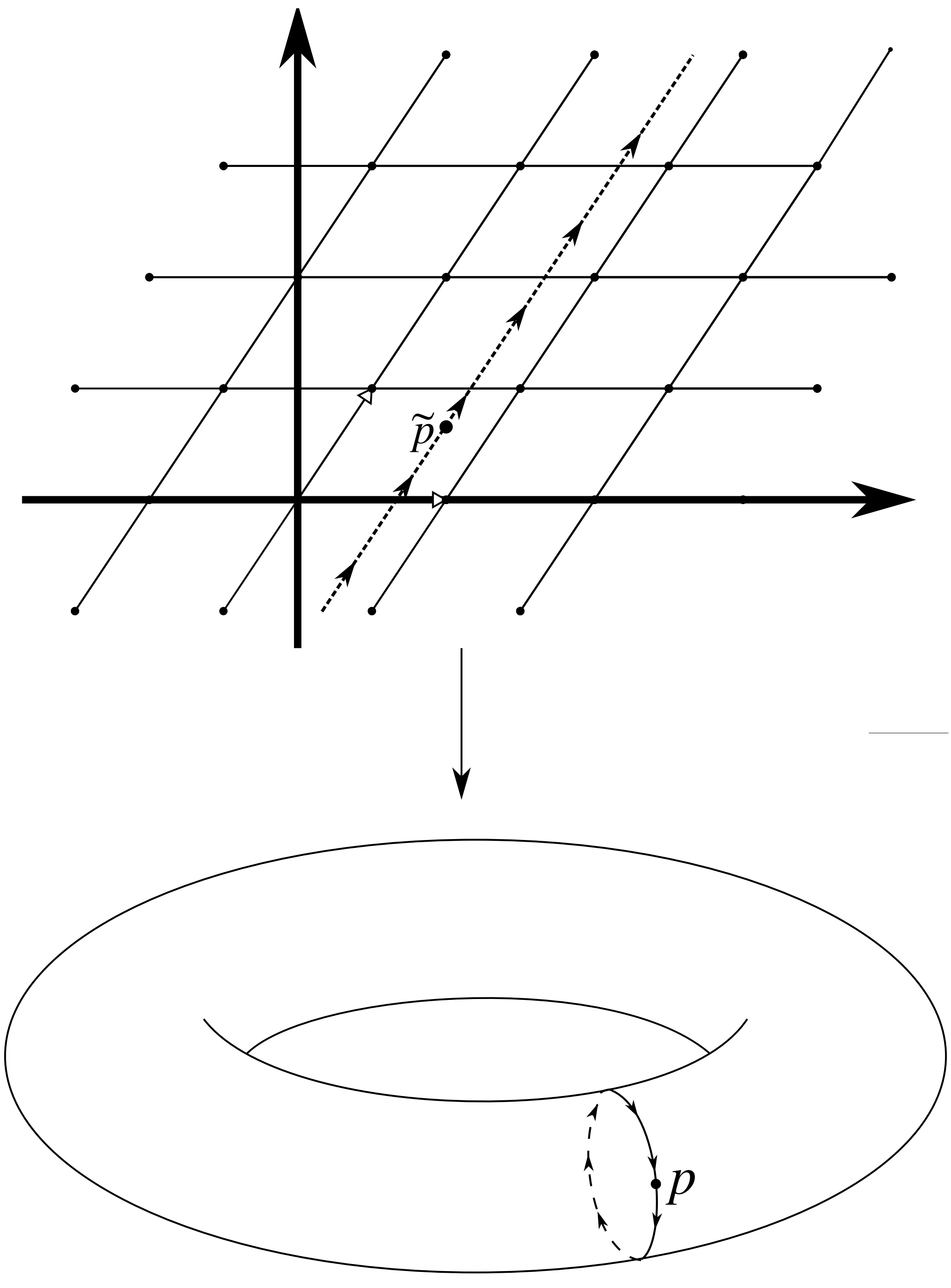

Figure 5.1: Linear motion on a 1 dimensional complex torus.

Theorem 5.19 ***(Linearization of flows)*** *Let $\beta(z,w)$ be an element of $H^1(C,\mathcal{O}_C)$, and let*

$$B(z) = (\beta(z, A(z)))_+$$

be the holomorphic part of the matrix Laurent series $\beta(z, A(z))$. Then the image of a solution $A(z;t)$ of the Lax equation

$$\frac{dA(z;t)}{dt} = [A(z;t), B(z;t)] \ , \ A(z;0) = A_0 \tag{5.4}$$

under the eigenvector map $l : M_P \to J^{g-1}(C) - \Theta$ is a line bundle $L(t)$ defined by the cocycle $c\exp(t\beta) \in H^1(C, \mathcal{O}_C^*)$, where $L(0) = l(A(z;0))$. Hence $L(t)$ evolves linearly in $\mathrm{Jac}^{g-1}(C)$.

Proof. In view of the Beauville correspondence, the strategy is to study the linear flow of $L(t)$ on $J^{g-1}(C) - \Theta$ in the direction $\beta(z,w)$ and show that we can find representatives $\tilde{A}(z;t)$ of the corresponding conjugacy classes of matrix polynomials such that equation (5.4) is satisfied. Consider the line bundle $L(t)$ with transition function $c\exp(t\beta(z,w))$. Let $s_k(t)$ denote the standard basis for the space of holomorphic sections of $L(t)$. In the $C \cap U$ trivialization, these sections are represented by locally defined holomorphic functions which we shall also denote $s_k(t)$. We denote their representatives in the $C \cap V_2$ trivialization by $\tilde{s}_k(t)$. On the intersection $V_1 \cap V_2$, we have

$$s_k(t) = c\exp(t\beta(z,w)) \cdot \ \tilde{s}_k(t)$$

Differentiating with respect to time yields

$$\frac{ds_k}{dt} = \beta(z,w) \cdot s_k(t) + c\exp(t\beta(z,w)) \cdot \frac{d\tilde{s}_k}{dt} \tag{5.5}$$

Using Lemma 5.12, we have that

$$ws_k(t) = \sum_{j=1}^{m} A_{kj}(z) s_j(t) \tag{5.6}$$

Hence we can write equation (5.5) as

$$\frac{ds_k}{dt} = \sum_j [\beta(z, A(z))]_{kj} \cdot s_j(t) + c\exp(t\beta(z,w)) \cdot \frac{d\tilde{s}_k}{dt}$$

Also note that since the spectrum of $A(z)$ is constant in time, differentiating (5.6) gives

$$w\frac{ds_k}{dt} = \sum_{j=1}^{m} \left(\frac{dA_{kj}}{dt} s_j(t) + A_{kj}(z) \frac{ds_j}{dt} \right) \tag{5.7}$$

Now consider the matrix-valued Laurent series $\beta(z, A(z))$. We can split it into two terms: one, β_+, which contains powers of z greater than or equal to 0, and another, β_-, containing strictly negative powers of z. We then have

$$\frac{ds_k}{dt} - \sum_j [\beta_+]_{kj} \cdot s_j(t) = \sum_j [\beta_-]_{kj} \cdot s_j(t) + c\exp(t\beta(z,w)) \cdot \frac{d\tilde{s}_k}{dt}$$

$$= c\exp(t\beta(z,w)) \left(\sum_j [\beta_-]_{kj} \tilde{s}_j(t) + \frac{d\tilde{s}_k}{dt} \right)$$

By our splitting of the powers of z in $\beta(z, A(z))$, we see that this defines a set of global holomorphic sections $\tau^{(k)}$ of $L(t)$. We can write these sections as a complex linear combination of the standard basis sections $s_i(t)$of $L(t)$:

$$\tau^{(k)} = \sum_{i=1}^{m} C_{ki} s_i(t)$$

In this fashion we obtain a new time-dependent matrix $C(t)$. In the $C \cap V_1$ trivialization, this statement reads

$$\frac{ds_k}{dt} - \sum_{i=1}^{m} [\beta_+]_{ki} \cdot s_i(t) = \sum_{i=1}^{m} C_{ki} s_i(t)$$

Now, substituting this expression into (5.7) yields:

$$w\frac{ds_k}{dt} = \sum_{j=1}^{m} \left(\frac{dA_{kj}}{dt} s_j + A_{kj} \frac{ds_j}{dt} \right) = \sum_{j=1}^{m} \left(\frac{dA_{kj}}{dt} s_j + \sum_{i=1}^{m} A_{kj} (C_{ji} + [\beta_+]_{ji}) \cdot s_i \right)$$

However, we can also write

$$\begin{aligned} w\frac{ds_k}{dt} &= w \sum_{i=1}^{m} \left([\beta_+]_{ki} \cdot s_i + C_{ki} s_i \right) \\ &= \sum_{i,j=1}^{m} A_{ji} \left([\beta_+]_{ki} + C_{kj} \right) s_j \end{aligned}$$

Comparing these expressions, we have

$$\sum_{i,j=1}^{m} A_{ji} \left([\beta_+]_{ik} + C_{kj} \right) s_j = \sum_{j=1}^{m} \left(\frac{dA_{jk}}{dt} s_j + \sum_{i=1}^{m} A_{jk} (C_{ji} + [\beta_+]_{ij}) \cdot s_i \right)$$

Equating coefficients of the basis section s_l yields

$$\frac{d}{dt} A_{kl} = \sum_i \left(C_{ki} + (\beta_+)_{ki} \right) A_{il} - \sum_j A_{kj} - \left(C_{jl} + (\beta_+)_{jl} \right)$$

Hence we find that

$$\frac{dA}{dt} = [A, \beta_+ + C(t)]$$

We want to eliminate the explicitly time-dependent term $C(t)$ from this equation. In order to accomplish this, we consider the conjugate $\tilde{A}(z) = PA(z)P^{-1}$ of $A(z)$, where P is a time-dependent(but not z-dependent) invertible matrix satisfying

$$\frac{dP}{dt} = P(t)C(t)$$

Then we find that we can express the time evolution of $\tilde{A}(z)$ as

$$\frac{d\tilde{A}}{dt} = [\tilde{A}, \beta_+(z, \tilde{A}(z))]$$

where $\beta_+(z, \tilde{A}(z)) = P\beta_+(z, A(z))P^{-1}$. □

Let us once again emphasise the meaning of this result: it tells us that any Lax pair with a spectral parameter obtained by applying the AKS construction in a loop algebra describes the linear motion of a point in a Jacobian.

Remark: Observe that if $\mathcal{Q}(z, A(z))$ is a polynomial in z and $A(z)$, making the transformation

$$B(z) \mapsto B(z) + \mathcal{Q}(z, A(z))$$

has no effect on the form of the Lax equation. This is a reflection of the cohomological nature of $B(z)$ we have just described.

Remark: As we have seen, we can interpret the Lax equation (5.2) as describing linear flow on $\text{Jac}^{g-1}(C) - \Theta$. The solution $A(z;t)$ of the equation must develop a singularity when this flow meets the theta divisor.

Example 5.20 (Nahm's equations) Following Hitchin in [18], we consider Nahm's equations, a set of 3 coupled nonlinear matrix ODE's which arise as a dimensional reduction of the self-dual Yang-Mills equations: for 2×2 matrices $T_1(t), T_2(t), T_3(t) \in \mathfrak{gl}_2(\mathbb{C})$, they read

$$\frac{dT_1}{dt} = [T_2, T_3]$$

$$\frac{dT_2}{dt} = [T_3, T_1]$$

$$\frac{dT_3}{dt} = [T_1, T_2]$$

Setting

$$A(z) = (T_1 + iT_2) + 2T_3 z - (T_1 - iT_2)z^2$$

we compute

$$\begin{aligned} \frac{dA}{dt} &= [T_2 - iT_1, T_3] + 2[T_1, T_2]z - [T_2 + iT_1, T_3]z^2 \\ &= [A, -iT_3 + i(T_1 - iT_2)z] \end{aligned}$$

and we see that the dynamics can be expressed in Lax form. From Proposition 5.6, we see that a generic (i.e. smooth) spectral curve has genus

$$g = \frac{1}{2}(2-1)(2 \cdot 2 - 2) = 1$$

Moreover, we note that the matrix $B(z)$ may be written

$$B(z) = -i\left(\frac{A}{z}\right)_+$$

Hence Nahm's equations can be interpreted as describing a linear flow on the Jacobian of C, in the direction described by the cocycle

$$\beta(z, w) = -i\frac{w}{z}$$

Let us also note an interesting special case of the Nahm equations: if

$$T_1 = u_1 \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$$

$$T_2 = u_2 \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$$

$$T_3 = u_3 \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

then we recover the Lax equation for the rigid body discussed in example 0.3.

Example 5.21 (The Neumann model revisited) The Neumann model also fits into this geometric framework as follows. Let $\mathcal{N}(z)$ be the matrix-valued rational function described in Section 3.3:

$$\mathcal{N}(z) = \mathcal{N}(z) = z \sum_{i=1}^{n} \begin{pmatrix} \frac{-x_i y_i}{z-\alpha_i} & -\frac{1}{2} - \frac{-y_i^2}{z-\alpha_i} \\ \frac{x_i^2}{z-\alpha_i} & \frac{x_i y_i}{z-\alpha_i} \end{pmatrix}$$

Writing $a(z) = \prod_{i=1}^{n}(z - \alpha_i)$, and setting

$$A(z) = \frac{-2a(z)}{z}\mathcal{N}(z) =: \begin{pmatrix} V(z) & U(z) \\ W(z) & -V(z) \end{pmatrix}$$

we recognise that the matrix $A(z)$ is of the 'Mumford' type considered in example 5.3: U is monic of degree n, while W, V are monic of degree $n-1$. Hence the spectral curve C is generically smooth and hyperelliptic of genus $g = n - 1$. Recall that the dynamics can be written

$$\frac{dA}{dt} = [A(z), B(z)]$$

where

$$B(z) = \sum_{i=1}^{n} \begin{pmatrix} x_i y_i & z + y_i^2 \\ -x_i^2 & -x_i y_i \end{pmatrix}$$

The matrix $B(z)$ can be expressed as

$$B(z) = (P(z, A(z)))_+$$

where $P(z, w) = (wz^{1-n})$. Hence the flow of the Neumann model is seen to linearize on the affine hyperelliptic Jacobian $\mathrm{Jac}^{g-1}(C) - \Theta$ and the direction of the linear flow is represented by the cohomology class $\beta(z, w) = wz^{1-n}$.

Chapter 6

Conclusion

6.1 Summary

Our discussion of classical integrable systems began within the framework of Hamiltonian mechanics. As we explained in Chapter 2, the Liouville-Arnold theorem guarantees that the flow of a Hamiltonian system with sufficiently many Poisson commuting conserved quantities is linearized on the common level sets of these quantitites. Moreover, provided these level sets are compact and connected, they are diffeomorphic to tori.
The Adler-Kostant-Symes method was used to give a systematic construction of Hamiltonian systems with non-trivial rings of Poisson commuting functions. We showed in Chapter 3 that the equations of motion for these systems can be cast in the form of Lax equations involving matrix polynomials. In Chapter 5, we studied these Lax equations in detail. From the data of a matrix polynomial $A(z)$, we explained how to construct a compact Riemann surface C and a holomorphic line bundle L over C. Crucially, since the solutions $A(z;t)$ of a Lax equation with spectral parameter have the same characteristic polynomial, the Riemann surface C was seen to be invariant under time evolution. As explained in Chapter 4, if C has genus g, the set of isomorphism classes of holomorphic line bundles of fixed degree over C has the structure of a g-dimensional complex torus called the Jacobian of C. For the Lax equations corresponding to the equations of motion for an AKS Hamiltonian system, it was shown that the image of the solution $A(z;t)$ evolves linearly in this Jacobian.

6.2 Further Developments

The algebraic-geometric machinery developed in Chapter 5 has a wide range of applications in the theory of integrable systems, as well as in broader areas of mathematics. In [6], Audin uses the map associating a matrix polynomial $A(z)$ to its eigenvector bundle to obtain further topological information about Liouville integrable Hamiltonian systems. In particular, Audin's methods allow her to elegantly determine which level sets of a system's conserved quantities are regular, as well as to study the behaviour of the Liouville tori as one passes through a critical value.

The connection between Liouville integrability and Lax equations is further developed in [27]. Adams, Harnad and Hurtubise study Liouville integrable Hamiltonian systems on a symplectic manifold M, and give an expression for the symplectic form on M in terms of 'spectral Darboux coordinates' associated to line bundles over a family of spectral curves embedded in a two-dimensional complex manifold. They show that their construction is equivalent to the classical separation of variables technique used to construct a system's action-angle coordinates. This is applied to a range of examples, including the Neumann model.

The Neumann model itself also has important applications. In particular, the quasi-periodic 'finite-gap' solutions of the famous KdV equation can be studied as solutions of the Neumann model. The reader interested in further details is referred to [12].

Finally, we mention a remarkable generalisation, due to Hitchin, of the algebraic-geometric ideas explained in Chapter 5. One begins by fixing a compact Riemann surface Σ of genus g. The space of isomorphism classes of stable[1] vector bundles of degree d and rank r over Σ has the structure of a complex manifold $\mathcal{M}$, and the holomorphic cotangent bundle of $\mathcal{M}$ is a complex symplectic manifold: that is, a complex manifold equipped with a closed holomorphic 2-form. The definition of Liouville integrability can be generalised to complex symplectic manifolds as follows: we say that a Hamiltonian system on a complex symplectic manifold N of complex dimension $2n$ is **algebraically completely integrable** if there exist n Poisson commuting holomorphic functions on N whose generic common level set is isomorphic to an open subset of an abelian variety on which the Hamiltonian vector fields of these functions are linear. In [19], Hitchin shows that the cotangent bundle of $\mathcal{M}$ supports a natural algebraically completely integrable system. In the case $\Sigma = \mathbb{P}^1$, we recover the Lax formalism discussed in Chapter 5. When Σ is of higher genus, it becomes more difficult to explicitly obtain the form of the equations of motion of Hitchin's integrable system. In [7], however, it is explained that the genus 1 case corresponds to a well-known classical integrable model, the Calogero-Moser system.

[1] For a definition, see [19]

Bibliography

[1] R. Abraham and J.E. Marsden. *Foundations of Mechanics.* Addison-Wesley, New York, 1978.

[2] M. Adler and P. van Moerbeke. Completely Integrable Systems, Euclidean Lie Algebras, and Curves. *Adv. Math.*, 38:267–317, 1980.

[3] M. Adler and P. van Moerbeke. Linearization of Hamiltonian Systems, Jacobi Varieties and Representation Theory. *Adv. Math.*, 38:318–379, 1980.

[4] V. I. Arnold. *Mathematical Methods of Classical Mechanics.* Springer-Verlag, New York, 1978.

[5] M. Audin. *Spinning Tops: A Course on Integrable Systems.* Cambridge University Press, Cambridge, 1996.

[6] M. Audin. *Hamiltonian Systems and Their Integrability.* American Mathematical Society, Rhode Island, 2008.

[7] O. Babelon, D. Bernard, and M. Talon. *Introduction to Classical Integrable Systems.* Cambridge University Press, Cambridge, 2003.

[8] A. Beauville. Jacobiennes des courbes spectrales et systemes hamiltoniens completement integrables. *Acta Math.*, 164:211–235, 1990.

[9] A.V. Bolsinov and A.T. Fomenko. *Integrable Hamiltonian systems : Geometry, Topology, Classification.* Chapman and Hall/CRC, Boca Raton, 2004.

[10] A. Cannas Da Silva. *Lectures on Symplectic Geometry.* Springer-Verlag, Berlin, 2001.

[11] J Duistermaat. On global action-angle coordinates. *Commun. Pure and Applied Math.*, 33:687–706, 1980.

[12] A.I. Bobenko E.D. Belokolos. *Algebro-Geometric Approach to Non-Linear Differential Equations.* Springer-Verlag, Berlin, 1984.

[13] P. Griffiths and J. Harris. *Principles of Algebraic Geometry.* Wiley, New York, 1978.

[14] R.C. Gunning. *Lectures on Riemann surfaces.* Princeton University Press, Princeton, New Jersey, 1966.

[15] R.C. Gunning. *Lectures on Riemann surfaces, Jacobi varieties.* Princeton University Press, Princeton, New Jersey, 1967.

[16] R.C. Gunning. *Lectures on vector bundles over Riemann surfaces.* Princeton University Press, Princeton, New Jersey, 1967.

[17] J Harnad. Loop groups, r-matrices and separation of variables. In *Integrable Systems: From Classical to Quantum*, volume 26 of *CRM Proceedings and Lecture Notes*, pages 21–52, Rhode Island, 2000. American Mathematical Society.

[18] N. J. Hitchin, G. Segal, N. Woodhouse, and R.S. Ward. *Twistors, Loop Groups and Riemann Surfaces.* Oxford University Press, Oxford, 1997.

[19] N.J. Hitchin. Stable bundles and integrable systems. *Duke Math. J.*, 54:91–114, 1987.

[20] J. Jost. *Compact Riemann Surfaces: An Introduction to Contemporary Mathematics.* Springer-Verlag, Berlin, 2000.

[21] A. A. Kirillov. *Lectures on the Orbit Method.* American Mathematical Society, Rhode Island, 2004.

[22] J.M. Lee. *Introduction to Smooth Manifolds.* Springer-Verlag, New York, 2003.

[23] I. Macdonald. *Hall Polynomials and Symmetric Functions.* Oxford University Press, Oxford, 1979.

[24] J. E. Marsden and T. S. Ratiu. *Introduction to Mechanics and Symmetry 2nd Edition.* Springer-Verlag, New York, 1999.

[25] R Miranda. *Algebraic curves and Riemann surfaces.* American Mathematical Society, Rhode Island, 1995.

[26] J. Harnad M.R. Adams and J. Hurtubise. Isospectral Hamiltonian Flows in Finite and Infinite Dimensions ii. Integration of Flows. *Commun. Math. Phys.*, 134:555–585, 1990.

[27] J. Harnad M.R. Adams and J. Hurtubise. Darboux Coordinates and Liouville-Arnold Integration in Loop Algebras. *Commun. Math. Phys.*, 155:385–413, 1993.

[28] J. Harnad M.R. Adams and E. Previato. Isospectral Hamiltonian Flows in Finite and Infinite Dimensions I. Generalised Moser Systems and Moment Maps into Loop Algebras. *Commun. Math. Phys.*, 117:451–500, 1988.

[29] D. Mumford. *Tata Lectures on Theta II.* Birkhauser, Basel, 1984.

[30] A.G. Reyman and M.A. Semenov-Tian-Shansky. Reduction of Hamiltonian systems, affine Lie algebras and Lax equations II. *Invent. Math.*, 63:423–432, 1981.

[31] A.G. Reyman and M.A. Semenov-Tian-Shansky. Group theoretical methods in the theory of finite dimensional integrable systems. In *Dynamical Systems VII*, volume VII of *Encyclopaedia of Math. Sci.* Springer-Verlag, New York, 1994.

[32] M.A. Semenov-Tian-Shansky. What a classical r-matrix is. *Funct. Anal. Appl.*, 17:259–272, 1983.

[33] V. Zakharov. *What is integrability?* Springer-Verlag, Berlin, 1991.