INTRODUCTION TO CLASSICAL INTEGRABLE SYSTEMS

O. Babelon, D. Bernard, M. Talon.

November 21, 2000

Contents

Introduction.

The aim of this book is to introduce the reader to classical integrable systems. Because the subject has been developed by several schools having different perspectives, it may appear fragmented at first sight. We develop here the thesis that it has a profound unity and that the various approaches are simply changes of point of view on the same underlying reality. The more one understands each approach, the more one sees their unity. At the end one get a very small set of interconnected methods.

This fundamental fact sets the tone of the book. We hope in this way to convey the reader to the extraordinary beauty of the structures emerging in this field, which have illuminated many other branches of theoretical physics.

The subject of integrable systems is born together with Classical Mechanics, with a quest of exact solutions to Newton's equations of motion. It turned out that apart from the Kepler problem which was solved by Newton himself, after two centuries of hard investigations, only a handful of other cases were found. In the nineteenth century, Liouville finally provided a general framework characterizing the cases where the equations of motion are "solvable by quadratures" only. All examples previously found indeed pertained to this setting. The subject stayed dormant until the second half of the twentieth century when Gardner, Greene, Kruskal and Miura invented the Classical Inverse Scattering Method for the Korteweg–de Vries equation, which had been introduced in fluid mechanics. Soon afterwards, the Lax formulation was discovered, and the connection with integrability was unveiled by Faddeev and Zakharov. This was the signal for a revival of the subject which led to an enormous amount of results, and truly general structures emerged, which organized the subject. More recently, the extension of these results to Quantum Mechanics already led to remarkable results and is still a very active field of research.

Let us give a general overview of the ideas we present in this book. They all find their roots in the notion of Lax pairs. It consists in presenting the equations of motion of the system in the form $\dot{L}(\lambda) = [L(\lambda), M(\lambda)]$, where the matrices $L(\lambda)$ and $M(\lambda)$ depend on the dynamical variables and on a parameter λ called the spectral parameter. Importance of Lax pairs stems from the following simple remark: the Lax equation is an isospectral evolution equation for the Lax matrix $L(\lambda)$. It follows that the curve defined by the equation $\det(L(\lambda) - \mu I) = 0$ is time–independent. This curve, called the spectral curve, can be seen as a Riemann surface. Its moduli contain the conserved quantities. This immediately introduces the two main structures into the theory: groups enter through the Lie algebra involved in the commutator $[L, M]$, while complex analysis enters through the spectral curve.

As integrable systems are rather rare one naturally expects strong constraints on the matrices $L(\lambda)$ and $M(\lambda)$. Constructing consistent Lax matrices may be achieved by introducing a factorization problem in appropriate groups. Taking into account the spectral parameter promotes this group to a loop group. The factorization problem may then be viewed as a Riemann–Hilbert problem, a central tool of this subject. Lax pairs are the new concept in the subject, which was not available to the mathematicians of the 19th century.

In the group theoretical setting, solving the equations of motion amounts to solving the factorization problem. In the analytical setting, solutions are obtained by considering the eigenvectors of the Lax matrix. At any point of the spectral curve there exists, by definition, an eigenvector of $L(\lambda)$ with eigenvalue μ. This defines an analytic line bundle $\mathcal{L}$ on the spectral curve with prescribed Chern class. The time evolution is described as follows: if $\mathcal{L}(t)$ is the line bundle at time t then $\mathcal{L}(t)\mathcal{L}^{-1}(0)$ is of Chern class 0, i.e. is a point on the Jacobian of the spectral curve. It is a beautiful result that this point evolves linearly on the Jacobian. As a result, one can express the dynamical variables in terms of θ functions defined on the Jacobian of the spectral curve. The two methods are related as follows: the factorization problem in loop group defines transition functions for the line bundle $\mathcal{L}$.

It is well–known that the Riemann–Hilbert problem is also at the heart of the so–called isomonodromic deformations problem, first considered by Riemann and further studied by Birkhoff. Here the main object of study is the equation $(\partial_\lambda + M_\lambda(\lambda))\psi = 0$, where $M(\lambda)$ depends rationnally on λ. The solution ψ acquires non trivial monodromy when λ describes a loop around a pole of M_λ. The isomonodromy problem consists in finding all M_λ with prescribed monodromy data. In general, the solution is not unique, and depends on a number of continuous parameters, the so–called isomonodromic deformation parameters. We show that the deformation equations with respect to these parameters are an integrable system. In our context this amounts to replacing the Lax matrix by a first orded linear differential operator $L(\lambda) \to \partial_\lambda + M_\lambda(\lambda)$. The θ functions of the isospectral approach are then extended to more general objects called the τ functions in the isomonodromic approach.

One can study the behaviour around each singularity of the differential operator quite independently. In the group theoretical version, the above extension of the framework corresponds to centrally extending the loop groups. Around a singularity the most general extended group is the group $GL(\infty)$ which corresponds to the KP hierarchy. It can be represented in a fermionic Fock space. Fermionic monomials acting on the vacuum yield decomposed vectors, which describe an infinite Grassmannian introduced by Sato. In this setting, the time flows are induced by the action of commuting one parameter subgroups, and the τ function is defined on the Grassmannian, i.e. the orbit of the vacuum, and characterizes it. Finally the equations of motion, written in the bilinear Hirota form identify with the Plücker equations of the Grassmannian.

The book is organized around these ideas. We have tried, as much as possible, to be self–contained, and to achieve that each Chapter can be studied quite independently. Generally we first explain methods and then show how they can be applied to particular examples, even though this does not correspond to the historical development of the subject.

In the first Chapter we introduce the classical definition of integrable systems through the Liouville theorem. We present the Lax pair formulation, and describe the symplectic structure which is encoded into the so–called r–matrix form. In the second Chapter we explain how to construct Lax pairs with spectral parameter, for finite and infinite dimensional systems. This introduces immediately a factorization problem in loop groups. We also introduce, at this early stage, the notion of τ functions. In the third Chapter we discuss abstractly the group theoretical aspects of the theory. We then describe the analytical aspects of the theory, in Chapter four, and illustrate the general constructions by the examples of the closed Toda chain in Chapter five, and the Calogero model in Chapter six.

The following two Chapters seven and eight describe respectively the isomonodromic deformation problem and the infinite Grassmannian. This puts the previous studies in proper perspective and allows to enter the analysis of field theoretical models in the following Chapters.

Chapters, nine and ten, are devoted to the classical study of the KP and KdV hierarchy. We develop and use the formalism of pseudo–differential operators which allows to give simple proofs of the main formal properties. Finite zone solutions of KdV allow to make contact with integrable systems of finite dimensionality and soliton solutions.

In the following three Chapters, eleven, twelve, thirteen, we study the class of Toda and Sine–Gordon field theories. We use this opportunity to exhibit the relations between their conformal and integrable properties. The Sine–Gordon model is presented in the framework of the Classical Inverse Scattering Method. This well–known and very ingenious method is usually exploited to solve the KdV equation. Here we felt it was rewarding to show how it applies to a slightly more involved situation.

The last three Chapters may be viewed as Mathematical appendices, provided to help the reader. First we present the basic facts of symplectic geometry, which is the natural language to speak about Classical Mechanics and integrable systems. Since mathematical tools from Riemann surfaces and Lie groups are used almost everywhere, we have written two Chapters presenting them in a concise way. We hope that they will be useful at least as an introduction and to fix notations. Of course we could not provide all proofs for these theories in such a small exposition, but we hope to give the general ideas and their relations.

Let us describe briefly how we have limited our discussion. First we have tried to remain consistently at a relatively elementary mathematical level, and have been obliged to exclude some important developments which require more advanced mathematics. We have not tried to make an exhaustive list of integrable systems. Another aspect of the theory we do not touch upon is the study of perturbation of integrable systems. All these subjects are very interesting by themselves, but the present book is big enough! A most active field of recent research is concerned with Quantum integrable systems or the closely related field of Exactly soluble models in Statistical mechanics, beautifully described in the book of Baxter. When writing this book we always had the quantum theory present in mind, and have introduced all classical objects which have a well–known quantum counterpart, or are semi classical limits of quantum objects. This explains our emphasis over Hamiltonians methods, Poisson brackets, classical r–matrices, and in particular the description of Poisson–Lie groups in the Chapter on symplectic spaces. So although nothing quantum is discussed in this book, a large part of the apparatus necessary to understand the literature on quantum integrable systems is in fact present.

The bibliography for integrable systems would fill a book by itself. We have made no attempt to provide one. Instead, we give, at the end of each Chapter, a short list of references, which complements and enhance the material presented in the Chapter, and we highly encourage the reader to consult them. Of course these references are far from complete, and we apologize to the numerous authors having contributed to the domain, and whose due credit is not acknowledged. Finally we want to thank our many colleagues from whom we learned so much and with whom we have discussed many parts of this book.

Chapter 1

Integrable dynamical systems

We introduce the classical definition of integrable systems through the Liouville theorem, namely systems for which n conserved quantities in involution are known on a phase space of dimension $2n$. The Liouville theorem asserts that the equation of motion can be solved by quadrature. The notion of Lax pair is then introduced. It is shown that, for integrable systems, the Poisson brackets are encoded into the so–called r–matrix form. Finally we present some historical examples of integrable systems.

1.1 Introduction.

In Classical Mechanics the state of the system is described by a point in phase space. This is generally a space of even dimension coordinatized by position q_i and momenta p_i. The Hamiltonian is a function on phase space, denoted $H(p_i, q_i)$. The equations of motion are a first order differential system taking the Hamiltonian form:

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial q_i} \tag{1.1}$$

Here and in the following, a dot will refer to a time derivative. For any function $F(p, q)$ on phase space, this implies that $F(p(t), q(t))$ obeys:

$$\dot{F} \equiv \frac{dF}{dt} = \{H, F\}, \quad \{F, G\} = \sum_i \frac{\partial F}{\partial p_i}\frac{\partial G}{\partial q_i} - \frac{\partial G}{\partial p_i}\frac{\partial F}{\partial q_i}$$

where $\{F, G\}$ is called the Poisson bracket. The quantity $H(p, q)$ is automatically conserved under time evolution, so that the motion takes place on the subvariety of phase space defined by $H = E$ constant.

Historically, it proved very difficult to find dynamical systems such that eqs.(1.1) could be solved exactly. However there is a general framework where explicit solution is guaranteed. This means that the solutions can be obtained by solving a finite number of algebraic equations and computing finite number of integrals. These dynamical systems are the Liouville integrable systems that we will consider in this book. A dynamical system on a phase space of dimension $2n$ is Liouville integrable if one knows n independent functions F_i on the phase space which Poisson *commute*, that is $\{F_i, F_j\} = 0$. The Hamiltonian is assumed to be a function of the F_i.

In order to understand the geometry of the situation, let us discuss a very simple example: the harmonic oscillator. The phase space is of dimension 2 and the Hamiltonian is $H = \frac{1}{2}(p^2 + \omega^2 q^2)$ with Poisson bracket $\{p, q\} = 1$. The phase space is fibered into ellipses $H = E$ except for the point $(0, 0)$ which is a stationnary point. An adapted coordinate system ρ, θ is given by: $p = \rho \cos(\theta), \quad q = \frac{\rho}{\omega}\sin(\theta)$ and the non vanishing Poisson bracket is $\{\rho, \theta\} = \omega/\rho$. In these coordinates the flow reads:

$$\rho = \sqrt{2E}, \quad \theta = \omega t + \theta_0$$

i.e. the flow takes place on the above ellipsis.

This can be straightforwardly generalized to a direct sum of n harmonic oscillators with

$$H = \sum_{i=1}^{n} \frac{1}{2}(p_i^2 + \omega_i^2 q_i^2)$$

We do have n conserved quantities in involution, $F_i = \frac{1}{2}(p_i^2 + \omega_i^2 q_i^2)$, and the level manifold M_f, i.e. the set of points of phase space such that $F_i = f_i$, is an n–dimensional real torus, which is explicitly a cartesian product of n circles. The motion takes place on these tori which foliate the phase space. We can intoduce n angles θ_i as above, conjugated to the F_i, which evolve linearly in time with frequency ω_i taking initial values θ_{i0}. A generic orbit of the dynamical flow is dense on the torus when the ω_i are rationnally independent. Let us also notice that these tori are Lagrangian subvarieties of phase space, that is the restriction of the symplectic form on them vanishes.

For Liouville integrable systems, we shall assume that the conserved quantities are well–behaved so that the n dimensional surfaces M_f defined by $F_i = f_i$ are generically regular, and foliate the phase space. This does not precludes the existence of singular points such as $p_i = q_i = 0$ in the above example of the harmonic oscillator. In this setting we are now going to prove the Liouville theorem and show that the geometry of the situation is analogous to that of the harmonic oscillator example.

1.2 The Liouville theorem.

We consider a dynamical Hamiltonian system with phase space M, *dim* $M = 2n$. Introduce canonical coordinates p_i, q_i such that the non degenerate Poisson bracket reads $\{p_i, q_j\} = \delta_{ij}$. As usual a non degenerate Poisson bracket on M is equivalent to the data of a non-degenerate closed 2-form ω, $d\omega = 0$, defined on M, see Chapter [14]. In the canonical coordinates $\omega = \sum_j dp_j \wedge dq_j$. Let H be the Hamiltonian of the system.

Definition. *The system is Liouville integrable if it possesses n independent conserved quantities F_i, $i = 1, \cdots, n$, $\{H, F_j\} = 0$, in involution*

$$\{F_i, F_j\} = 0$$

The independence means that at generic points (i.e. anywhere except on a set of measure 0), the dF_i are linearly independent, or that the tangent space of the surface $F_i = f_i$ exists everywhere and is of dimension n. There cannot be more than n independent quantities in involution otherwise the Poisson bracket would be degenerate. In particular, the Hamiltonian H is a function of the F_i's.

The Liouville theorem. *The solution of the equations of motion of an integrable system is obtained by quadrature.*

Proof. Let $\alpha = \sum_i p_i dq_i$ be the canonical 1-form and $\omega = d\alpha = \sum_i dp_i \wedge dq_i$ be the symplectic 2-form on the phase space M. We will construct a canonical transformation $(p_i, q_i) \to (F_i, \Psi_i)$ such that the F_i's are among the new coordinates. i.e., a transformation such that

$$\omega = \sum_i dp_i \wedge dq_i = \sum_i dF_i \wedge d\Psi_i$$

If we succeed to do that, the equations of motion become trivial:

$$\begin{aligned} \dot{F}_j = \{H, F_j\} &= 0 \\ \dot{\psi}_j = \{H, \psi_j\} &= \frac{\partial H}{\partial F_j} = \Omega_j(F) \end{aligned} \quad (1.2)$$

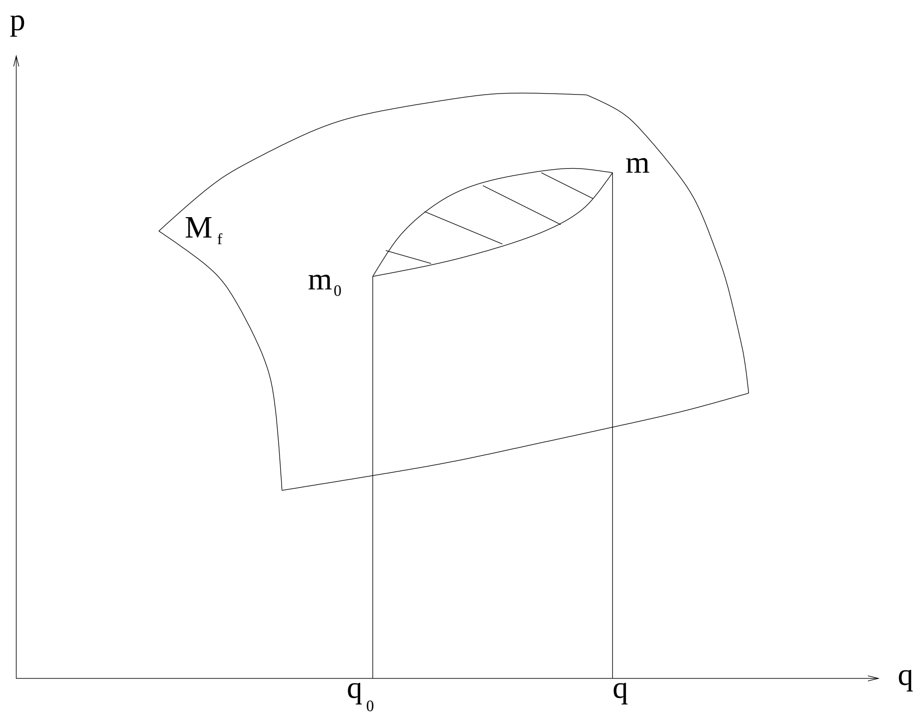
Their solutions are:

$$F_j(t) = F_j(0), \quad \psi_j(t) = \psi_j(0) + t\Omega_j$$

To construct this canonical transformation, we exhibit its generating function S. Let M_f be the level manifold $F_i(p,q) = f_i$. Suppose we can solve for p_i, $p_i = p_i(f,q)$, and consider the function

$$S(F,q) = \int_{m_0}^{m} \alpha = \int_{q_0}^{q} \sum_i p_i(f,q) dq_i,$$

where the integration path is drawn on M_f and goes from the point of coordinate $(p(f,q_0),q_0)$ to the point $(p(f,q),q)$, where q_0 is some reference value.



A leaf M on phase space

If this function exists, i.e. if it does not depend on the path from m_0 to m, it is the function we are looking for. Indeed, from the definition of S, $p_j = \frac{\partial S}{\partial q_j}$. Defining ψ_j by

$$\psi_j = \frac{\partial S}{\partial F_j},$$

we have

$$dS = \sum_j \psi_j dF_j + p_j dq_j,$$

Since $d^2S = 0$ we deduce that $\omega = \sum_j dp_j \wedge dq_j = \sum_j dF_j \wedge d\psi_j$. This shows that if S is a well defined function then the transformation is canonical.
To show that S exists, we must prove that it is independent of the integration path, ie. we have to prove that

$$d\alpha|_{M_f} = \omega|_{M_f} = 0$$

Let X_i be the Hamiltonian vector field associated to F_i, defined by $dF_i = \omega(X_i, \cdot)$,

$$X_i = \sum_k \frac{\partial F_i}{\partial q_k}\frac{\partial}{\partial p_k} - \frac{\partial F_i}{\partial p_k}\frac{\partial}{\partial q_k}.$$

These vector fields are tangent to the manifold M_f because the F_j are in involution,

$$X_i(F_j) = \{F_i, F_j\} = 0$$

Since the F_j are assumed to be independent functions, the tangent space to the submanifold M_f is generated at each point $m \in M$ by the vectors $X_i|_m$ $(i = 1, \ldots, n)$. But then $\omega(X_i, X_j) = dF_i(X_j) = 0$ and we have proved that $\omega|_{M_f} = 0$, and therefore S exists. ■

We have effectively obtained the solution of the equations of motion through one quadrature (to calculate the function S) and some "algebraic manipulation" (to express the p's as functions of q and F)

Remark 1. From the closedness of α on M_f, the function S is unchanged by *continuous* deformations of the path (m_0, m). However, if M_f has non trivial cycles, which is generically the case, S is a multivalued function defined in a neighborhood of M_f. The variation over a cycle

$$\Delta_{cycle} S = \int_{cycle} \alpha$$

is a function of F only. This induces a multivaluedness of the variables ψ_j: $\Delta_{cycle}\psi_j = \frac{\partial}{\partial F_j}\Delta_{cycle}S$. For instance, in the case of harmonic oscillators, we see that above each point $(q_1, \cdots, q_n)$ we have 2^n points on the M_f level surface, due to the independent choices of sign in $p_i = \pm\sqrt{2f_i - \omega_i^2 q_i^2}$. So we have many choices for the path of integration, reflecting the topology of the torus.

Remark 2. The definition we have given of a Liouville integrable system requires some care. Given any Hamiltonian H, the Darboux theorem, see Chapter [14], implies that we can always find *locally* a system of canonical coordinates on phase space $(P_1, \cdots, P_n; Q_1, \cdots, Q_n)$, with $H = P_1$, hence fulfilling the hypothesis of Liouville theorem. For integrable systems we require that the conserved quantities are globally defined on a sufficiently large open set, and that the surfaces $F_i = f_i$ are well–behaved and foliate the phase space. This is not generally the case for the P_i constructed by the Darboux theorem. Moreover in all known examples, the conserved quantities are even analytic functions of canonical coordinates on some open domain and the solutions of the equations of motion are similarly analytic.

Remark 3. Using the Poisson commuting functions F_i, one can solve *simultaneously* the n time evolution equations $dF/dt_i = \{F_i, F\}$, since:

$$\frac{\partial}{\partial t_i}\frac{\partial}{\partial t_j}F - \frac{\partial}{\partial t_j}\frac{\partial}{\partial t_i}F = \{F_i, \{F_j, F\}\} - \{F_j, \{F_i, F\}\} = \{\{F_i, F_j\}, F\} = 0$$

Since the Hamiltonian vector fields are well–defined and linearly independent everywhere, the flows define a locally free (no fixed points) and transitive (goes everywhere) action of a small open set in $\mathbb{R}^n$ on the surface M_f. Assuming that M_f is connected and compact, the flows extend to all values of the times t_i and fill the whole surface M_f, hence we have a surjective action of $\mathbb{R}^n$ on M_f. The stabilizer of a point is an abelian discrete subgroup of $\mathbb{R}^n$ since the action is locally free, so it is of the form $\mathbb{Z}^n$. Thus M_f appears as the quotient of $\mathbb{R}^n$ by $\mathbb{Z}^n$, i.e. a torus. This refinement, a theorem due to Arnold, of the Liouville theorem shows that, under suitable global hypothesis, the phase space is indeed foliated by n dimensional tori, called the Liouville tori.

1.3 Action-angle variables

As already noticed in the proof of the Liouville theorem, the level manifold M_f has non trivial cycles. Under suitable compacity and connexity hypothesis, the M_f are indeed tori T_n. This points to the introduction of angle variables to describe the motion along the cycles. The torus T_n is isomorphic to a product of n circles C_i hence we may choose special angular coordinates on M_f: the n angles φ_i $(i = 1, \ldots, n)$ associated to n fundamental cycles C_i $(i = 1, \ldots, n)$. The variables φ_i $(i = 1, \ldots, n)$ are the angle variables. By definition,

$$\frac{1}{2\pi} \int_{C_j} d\varphi_i = \delta_{ij}.$$

Let us define the action variables I_j as the integrals of the canonical 1–form over the cycles C_j,

$$I_j = \frac{1}{2\pi} \int_{C_j} \alpha.$$

The I_j are functions of the F's and we suppose they are independent, so that if the values of I_j $(j = 1, \ldots, n)$ are known, then M_f is determined. The I_j are the action variables and are constants of motion.

Proposition. *The variables I_j and φ_j are canonically conjugate.*

Proof. To prove the proposition, we describe the canonical transformation from the system of coordinates (p, q) to the system (I, φ) with its generating function. Let

$$S(I, q) = \int_{m_0}^{m} \alpha$$

as above, that is to say the same function S, but in terms of the variables I instead of F. Denoting by θ_j the variable conjugate to I_j, the canonical transformation generated by S is defined by

$$p_k = \frac{\partial S}{\partial q_k} \qquad\qquad \theta_k = \frac{\partial S}{\partial I_k}.$$

We must show that $\theta_k = \varphi_k$. For this it is sufficient to show that θ_k satisfies

$$\frac{1}{2\pi} \int_{C_j} d\theta_k = \delta_{jk}.$$

Since by definition of θ_k,

$$\int_{C_j} d\theta_k = \frac{\partial}{\partial I_k} \int_{C_j} dS,$$

and since on the manifold M_f, $dI_k = 0$, we have

$$\int_{C_j} d\theta_k = \frac{\partial}{\partial I_k} \int_{C_j} \frac{\partial S}{\partial q} dq = \frac{\partial}{\partial I_k} \int_{C_j} \alpha = 2\pi\delta_{jk}.$$

This proves that $\theta_k = \varphi_k$ and the proposition. ∎

1.4 Lax pairs

The new concept which emerged from the modern studies of integrable systems is the notion of Lax pairs. A Lax pair L, M consists of two matrices, functions on the phase space of the system, such that the Hamiltonian evolution equations may be written as

$$\frac{dL}{dt} \equiv \dot{L} = [L, M] \tag{1.3}$$

Here, $[L, M] = LM - ML$ denotes the commutator of the matrices L and M.

The immediate interest in the existence of such a pair lies in the fact that it allows for an easy construction of conserved quantities. Indeed the solution of eq(1.3) is of the form

$$L(t) = g^{-1}(t)L(0)g(t),$$

where the invertible matrix $g(t)$ is determined by the equation

$$M = g^{-1}\frac{dg}{dt}$$

It follows that if $I(L)$ is a function of L invariant by conjugation $L \to g^{-1}Lg$, then $I(L(t))$ is a constant of the motion. Such functions are functions of the eigenvalues of L. We say that the evolution equation (1.3) is isopsectral which means that the spectrum of L is preserved by the time evolution.

Integrability of the system in the sense of Liouville demands (i) that the number of independent conserved quantities equals the number of degree of freedom, and (ii) that these conserved quantities are in involution.

Remark. A Lax pair is by no means unique: even the size of the matrices may be changed. There is also a natural gauge transformation group acting on the Lax pair:

$$L \longrightarrow g^{-1}Lg, \quad M \longrightarrow g^{-1}Mg + g^{-1}\frac{dg}{dt}$$

where g is an invertible matrix, function on phase space.

Let us present some simple examples showing that the equations of motion can indeed be recast into Lax form.

Example 1. For any integrable system in the sense of Liouville, one may construct a Lax pair in a tautological way. Consider a finite dimensional Hamiltonian system, with n degrees of freedom, Poisson bracket $\{\ ,\ \}$ and Hamiltonian H. Suppose it is integrable in the sense of Liouville, which means that it possesses n independent integrals of the motion $F_i, i = 1, \ldots, n$, in involution. The Liouville theorem states that there exists, at least locally, a system of conjugate coordinates $I_i,\ \theta_i, i = 1, \ldots, n$, where the I_j are functions of the F's only. In these coordinates, the equations of motion take the very simple form

$$\dot{I}_j \;\;=\;\; 0, \quad \dot{\theta}_j = \frac{\partial H}{\partial I_j}. \tag{1.4}$$

Introduce the Lie algebra generated by $\{H_i, E_i,\ i = 1, \ldots, n\}$ with relations

$$[H_i, H_j] = 0, \quad [H_i, E_j] = 2\,\delta_{ij}E_j, \quad [E_i, E_j] = 0$$

This Lie algebra has a natural representation by $2n \times 2n$ matrices. Set:

$$L = \sum_{j=1}^{n} I_j\ H_j + 2I_j\theta_j E_j, \quad M = \sum_{j=1}^{n} \frac{\partial H}{\partial I_j} E_j$$

The equation $\dot{L} = [L, M]$ is then equivalent to eq.(1.4). Thus L, M form a Lax pair. However, this construction is useless since it requires the knowledge of the action angle variables to build the Lax pair, but if these are known, there is no need for a Lax pair anymore.

Example 2. As a second example we exhibit a Lax pair for the harmonic oscillator. Let:

$$L = \begin{pmatrix} p & \omega q \\ \omega q & -p \end{pmatrix}, \quad M = \begin{pmatrix} 0 & -1/2 \\ 1/2 & 0 \end{pmatrix}$$

We check immediately that the Lax equation eq.(1.3) is equivalent to the equations of motion $\dot{p} = -\omega q, \dot{q} = \frac{1}{\omega}p$. Let us observe that the Hamiltonian H can be written as $\frac{1}{4}\mathrm{Tr}L^2$. This example can be generalized to n independent harmonic oscillators by writing the Lax matrices L, M in a block diagonal form where each block is a two by two matrix as above. Now the conserved quantities are $\mathrm{Tr}L^{2p} = 2\sum(2F_i)^p$, $\mathrm{Tr}L^{2p+1} = 0$ so that they are equivalent to the collection of the F_i.

1.5 Existence of an r–matrix.

A Lax pair provides us with conserved quantities without referring to a Poisson structure. The notion of Liouville integrability requires the knowledge of a Poisson structure together with the involution property of the conserved quantities. We shall now describe the general form of Poisson structures which ensures the involution property for the conserved quantities given by the Lax pair.

Suppose we are given a Lax pair L, M, and suppose that the matrix L can be diagonalized:

$$L = U^{-1}\Lambda U \tag{1.5}$$

The matrix elements λ_k of the diagonal matrix Λ are the conserved quantities. We will not consider here the question of the independence of these quantities.

Let us introduce some notations. Let e_{ij} be a basis of the $N \times N$ matrices. We can write

$$L = \sum_{ij} L_{ij} e_{ij}$$

The components L_{ij} of the Lax matrix are functions on the phase space. We can evaluate the Poisson brackets $\{L_{ij}, L_{kl}\}$ and gather the results as follows. Set

$$L_1 = L \otimes 1 = \sum_{ij} L_{ij}(e_{ij} \otimes 1), \quad L_2 = 1 \otimes L = \sum_{ij} L_{ij}(1 \otimes e_{ij})$$

$$\{L_1, L_2\} = \sum_{ij,kl} \{L_{ij}, L_{kl}\} e_{ij} \otimes e_{kl}$$

The index 1 or 2 means that the matrix L sits in the first or second factor in the tensor product. Similarly, for α living in the tensor product of two copies of $N \times N$ matrices, we set

$$\alpha = \alpha_{12} = \sum_{ij,kl} \alpha_{ij,kl}\, e_{ij} \otimes e_{kl}, \quad \alpha_{21} = \sum_{ij,kl} \alpha_{ij,kl}\, e_{kl} \otimes e_{ij}$$

The Poisson brackets between the components of the Lax matrix L can be summarized in the following

Proposition. *The involution property of the eigenvalues of L is equivalent to the existence of a function, r_{12}, on the phase space such that:*

$$\{L_1, L_2\} = [r_{12}, L_1] - [r_{21}, L_2] \tag{1.6}$$

<u>Proof</u>. Assume first that the eigenvalues of L Poisson commute, $\{\lambda_i, \lambda_j\} = 0$. Recall that L is diagonalized by U, $L = U^{-1}\Lambda U$. Since U is a function on phase space, we compute directly the Poisson brackets $\{L_1, L_2\} = \{U_1^{-1}\Lambda_1 U_1, U_2^{-1}\Lambda_2 U_2\}$ using the Leibnitz rule. We get 9 terms. Out of these, 4 terms involve the Poisson brackets $\{U_1, U_2\}$. Introducing the quantity $k_{12} = U_1^{-1}U_2^{-1}\{U_1, U_2\}$, these terms can be written as $[[k_{12}, L_2], L_1] = \frac{1}{2}[[k_{12}, L_2], L_1] - \frac{1}{2}[[k_{21}, L_1], L_2]$. Four other terms involve $\{\Lambda_1, U_2\}$ and $\{U_1, \Lambda_2\}$. Introducing $q_{12} = U_1^{-1}U_2^{-1}\{U_1, \Lambda_2\}U_2$ we can write them as $-[q_{12}, L_1] + [q_{21}, L_2]$. Putting all this together, we get

$$\{L_1, L_2\} = U_1^{-1}U_2^{-1}\{\Lambda_1, \Lambda_2\}U_1U_2 + [r_{12}, L_1] - [r_{21}, L_2] \quad \text{with } r_{12} = -q_{12} + \tfrac{1}{2}[k_{12}, L_2]$$

This proves one part of the equivalence when $\{\Lambda_1, \Lambda_2\} = 0$. Conversely, suppose we have eq.(1.6). Then, in any matrix representation

$$\{L_1^n, L_2^m\} = [a_{12}^{n,m}, L_1] + [b_{12}^{n,m}, L_2] \tag{1.7}$$

with

$$\begin{aligned} a_{12}^{n,m} &= \sum_{p=0}^{n-1}\sum_{q=0}^{m-1} L_1^{n-p-1} L_2^{m-q-1} r_{12} L_1^p L_2^q \\ b_{12}^{n,m} &= -\sum_{p=0}^{n-1}\sum_{q=0}^{m-1} L_1^{n-p-1} L_2^{m-q-1} r_{21} L_1^p L_2^q. \end{aligned}$$

Taking the trace of eq(1.7), we get the desired involution property of $\mathrm{tr}(L^n)$, which is equivalent to the involution of the eigenvalues λ_k of L. ■

Although simple to prove, this proposition is important for developing formal aspects of integrable systems since it allows to control the Poisson brackets of the Lax matrix.

The Jacobi identity on the Poisson bracket yields the following constraint on r:

$$[L_1, [r_{12}, r_{13}] + [r_{12}, r_{23}] + [r_{32}, r_{13}] + \{L_2, r_{13}\} - \{L_3, r_{12}\}] + \text{cyclic permutations} = 0 \tag{1.8}$$

Notice that the first terms in eq.(1.8) are the only remaining ones if r happens to be constant. In a sense, solving this equation amounts to classifying integrable Hamiltonian systems.

Remark 1. The form of the bracket is preserved by gauge transformations. If

$$\{L_1, L_2\} = [r_{12}, L_1] - [r_{21}, L_2],$$

and $L' = g^{-1}Lg$, then there exists a function r'_{12} with values in $\mathcal{G}\otimes\mathcal{G}$ such that,

$$\{L'_1, L'_2\} = [r'_{12}, L'_1] - [r'_{21}, L'_2].$$

The function r'_{12} can be expressed in terms of the functions r_{12} and of the Poisson brackets between g and the Lax matrix L.

$$r'_{12} = g_1^{-1} g_2^{-1}\left(r_{12} - \{g_1, L_2\} g_1^{-1} + \frac{1}{2}[u_{12}, L_2]\right) g_1 g_2 \tag{1.9}$$

where $u_{12} = \{g_1, g_2\} g_1^{-1} g_2^{-1}$.

Remark 2. In the form (1.6) the antisymmetry property of the bracket is explicit, although r has no special symmetry property. Furthermore, we have the freedom to redefine r by

$$r_{12} \longrightarrow r_{12} + [\sigma_{12}, L_2], \tag{1.10}$$

where σ is symmetric, without changing the Poisson bracket.

Example. Let us give an example of this construction in our favorite simple example of harmonic oscillator. We have given above the Lax matrix L and introduced the action–angle coordinates ρ, θ. In these coordinates the matrix L is diagonalized by:

$$U = U^{-1} = \begin{pmatrix} \cos\theta/2 & \sin\theta/2 \\ \sin\theta/2 & -\cos\theta/2 \end{pmatrix}$$

Since $\{U_1, U_2\} = 0$, $r_{12} = -q_{12}$, which is easily computed to be:

$$r_{12} = \frac{\omega}{2\rho^2} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \otimes L$$

It is easy to verify that this r–matrix indeed satisfies eq.(1.6). Let us notice that it is a dynamical r–matrix, which means that it depends explicitly on the dynamical variables.

1.6 Commuting flows.

The Poisson brackets eq.(1.6) are equivalent to the involution of the eigenvalues of L. As an equivalent set of commuting Hamiltonians H_n we can choose the traces of the powers of the Lax matrix:

$$H_n = \mathrm{Tr}\,(L^n) \tag{1.11}$$

The Hamiltonians H_n are in involution, $\{H_n, H_m\} = 0$, since they are symmetric polynomials in the eigenvalues. Furthermore, we have

Proposition. *Suppose we have $\{L_1, L_2\} = [r_{12}, L_1] - [r_{21}, L_2]$. If we take $H_n = \mathrm{Tr}\,(L^n)$ as Hamiltonians, then the equations of motion admit a Lax representation:*

$$\frac{dL}{dt_n} \equiv \{H_n, L\} = [L, M_n]\,, \quad \text{with } M_n = n\;\mathrm{Tr}_1\,(L_1^{n-1} r_{21}) \tag{1.12}$$

Proof. Set $m = 1$ in eq.(1.7), take the trace of it over the first space, and get $dL/dt_n = [L, M_n]$ with $M_n = n\;\mathrm{Tr}_1(L_1^{n-1} r_{21})$. ■

Remark that the matrices M_n are unchanged under the transformation eq.(1.10). However, the matrices M_n are not unique since adding any matrix commuting with L does not change the equations of motion.

We close this chapter by presenting some of the few historical examples of integrable systems which were known at the end of the last century. Of course all systems with one degree of freedom (phase space of dimension 2) are integrable since the Hamiltonian H is conserved. We discuss below more sophisticated examples in higher dimension.

1.7 The Kepler problem.

The first system is the Kepler two–body problem. In the center of mass frame, the equations of motion take the form:

$$\frac{d^2 x_i}{dt^2} = -\frac{\partial V(r)}{\partial x_i}, \quad r = \sqrt{x_1^2 + x_2^2 + x_3^2}$$

In the traditional Kepler problem, $V(r) = C/r$ but we will consider any centrally symmetric potential $V(r)$. This is an Hamiltonian system with

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

and Poisson brackets $\{p_i, x_j\} = \delta_{ij}$. The phase space is of dimension 6, and we have to exhibit 3 commuting conserved quantities. Due to central symmetry, the angular momentum $J_{ij} = x_i p_j - x_j p_i$ is conserved, but the 3 components do not Poisson commute. However a set of 3 independent Poisson commuting quantities is provided by H, $J_3 \equiv J_{12}$, $J^2 \equiv J_{12}^2 + J_{23}^2 + J_{13}^2$. At this point, one may follow the standard solution, which takes advantage of the conservation of $\vec{J}$, and restrict ourselves to the plane perpendicular to $\vec{J}$ where the motion takes place. Here we prefer to show the Liouville theorem at work, and use only the 3 commuting conserved quantities. Due to the spherical symmetry of the problem, it is convenient to use spherical coordinates:

$$x_1 = r\sin(\theta)\cos(\phi), \quad x_2 = r\sin(\theta)\sin(\phi), \quad x_3 = r\cos(\theta)$$

We introduce the conjugate momenta p_r, p_θ, p_ϕ by writing the canonical one–form $\alpha = \sum p_i dx_i = p_r dr + p_\theta d\theta + p_\phi d\phi$. In these coordinates the conserved quantities read:

$$\begin{aligned} H &= \frac{1}{2}(p_r^2 + \frac{1}{r^2}p_\theta^2 + \frac{1}{r^2\sin^2\theta}p_\phi^2) + V(r) \\ J^2 &= p_\theta^2 + \frac{1}{\sin^2\theta}p_\phi^2 \\ J_3 &= p_\phi \end{aligned} \tag{1.13}$$

On the surface M_f corresponding to fixed values of the conserved quantities, we solve for the p's in terms of the the position variables, yielding:

$$p_r = \sqrt{2(H - V(r)) - \frac{J^2}{r^2}}, \quad p_\theta = \sqrt{J^2 - \frac{J_3^2}{\sin^2\theta}}, \quad p_\phi = J_3$$

The one–form α restricted on M_f is obviously closed, since all variables are *separated*, and the action S appearing in Liouville theorem reads:

$$S = \int^r \sqrt{2(H - V(r)) - \frac{J^2}{r^2}}dr + \int^\theta \sqrt{J^2 - \frac{J_3^2}{\sin^2\theta}}d\theta + \int^\phi J_3 d\phi$$

The angle variables corresponding to our action variables are given by

$$\psi_H = \frac{\partial S}{\partial H}, \quad \psi_{J^2} = \frac{\partial S}{\partial J^2}, \quad \psi_{J_3} = \frac{\partial S}{\partial J_3}$$

and have simple time evolution with respective frequencies $(1, 0, 0)$ by eq.(1.2). Hence ψ_{J^2} and ψ_{J_3} remain constant, while $\psi_H = t - t_0$. This gives the standard formula for the Kepler motion:

$$t - t_0 = \int^r \frac{dr}{\sqrt{2(H - V(r)) - \frac{J^2}{r^2}}}$$

Note that the constancy of ψ_{J_3} implies:

$$\dot{\phi} = \frac{J_3}{\sin^2\theta\sqrt{J^2 - \frac{J_3^2}{\sin^2\theta}}}\dot{\theta}$$

This, in turn, implies the constancy of:

$$J_1 = -J_3\cot\theta\cos\phi - \sin\phi\sqrt{J^2 - \frac{J_3^2}{\sin^2\theta}}, \quad J_2 = -J_3\cot\theta\sin\phi + \cos\phi\sqrt{J^2 - \frac{J_3^2}{\sin^2\theta}}$$

so that the motion takes place in the plane perpendicular to $\vec{J}$, as expected. It is worth noticing that the present approach is the one which prevails in Quantum Mechanics, where the 3 components of $\vec{J}$ cannot be measured simultaneously.

1.8 The Euler top.

We consider a rotating solid body attached to a fixed point. We first assume that there is no external force. It is very convenient to consider the equations of motion in a frame rotating with the body as discovered by Euler. We choose the moving frame with origin at the fixed point of the top (that is the point where the top is attached), and the axis being the principal inertia axis which diagonalize the inertia tensor computed with respect to the fixed point $I_{ij} = \int (\vec{x}^2 \delta_{ij} - x_i x_j)\rho(x) dx$. Let $\vec{J}$ be the angular momentum of the top seen in the moving frame. We have $\vec{J} = I.\vec{\omega}$ where $I = \mathrm{Diag}(I_1, I_2, I_3)$ and $\vec{\omega}$ is the rotation vector of the moving frame. We shall assume that the principal moments of inertia I_i are all different. Then the equation of motion reads:

$$\frac{d\vec{J}}{dt} = -\vec{\omega} \wedge \vec{J}$$

It expresses the conservation of $\vec{J}$ in the absolute frame. This can be recast into the Hamiltonian framework by defining the Poisson brackets:

$$\{J_i, J_j\} = \epsilon_{ijk} J_k$$

where ϵ_{ijk} is the usual antisymmetric tensor. The Hamiltonian reads:

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

This Poisson bracket is degenerate and one must choose a symplectic leaf to get a well–defined Hamiltonian system. This achieved by fixing the value of $\vec{J}^2$ to a numerical value. Then the phase space is of dimension 2, and the system is integrable with conserved quantity H. Note that the trajectories are immediately obtained as the intersection of the sphere $J_1^2 + J_2^2 + J_3^2 = \vec{J}^2$ and the ellipsoid $J_1^2/I_1 + J_2^2/I_2 + J_3^2/I_3 = 2H$. Using these relations to compute J_2 and J_3 in terms of J_1 and substituting in the equation of motion of J_1, i.e. $\dot{J}_1 = (I_3^{-1} - I_2^{-1}) J_2 J_3$, yields an equation of the form $\dot{J}_1 = \sqrt{\alpha + \beta J_1^2 + \gamma J_1^4}$, so that J_1 is an elliptic function of t.

1.9 The Lagrange top.

When the top is in a gravitational field its weight has to be taken into account and the problem is more complicated. Let us assume that the rotating frame has its origin at the fixed point. In general the problem is not integrable, and becomes integrable only in special cases. One case, found by Lagrange, is when two inertia moments are equal, for example $I_1 = I_2$ and the center of mass is located at a position $(x_1 = 0, x_2 = 0, x_3 = h)$ with respect to the rotating frame. This situation is achieved when the top has an axis of symmetry (around the third axis) and is attached to a point on this axis.

For any top in a gravitational field, the equations of motion in the rotating frame take the form:

$$\frac{d\vec{J}}{dt} = -\vec{\omega} \wedge \vec{J} + \vec{h} \wedge \vec{P}, \quad \frac{d\vec{P}}{dt} = -\vec{\omega} \wedge \vec{P} \tag{1.14}$$

where $\vec{P}$ is the weight of the top which is constant in the absolute frame and $\vec{h}$ is the vector from the fixed point to the center of mass which is constant in the rotating frame. From these equations one can check the conservation of three quantities: $\vec{P}^2$, $\vec{J} \cdot \vec{P}$ and the energy $H = \frac{1}{2}(\vec{J} \cdot I^{-1} \vec{J}) - \vec{P} \cdot \vec{h}$. In order to formulate these equations in a Hamiltonian framework let us introduce the following Poisson brackets between the dynamical quantities J and P:

$$\{J_i, J_j\} = \epsilon_{ijk} J_k, \quad \{J_i, P_j\} = \epsilon_{ijk} P_k, \quad \{P_i, P_j\} = 0$$

The Hamilton equations of motion are precisely eqs.(1.14). The Poisson structure is degenerate, namely the two conserved quantities $\vec{P}^2$, $\vec{J}\cdot\vec{P}$ are in the center. Hence the symplectic leaves are of dimension 4. The Hamiltonian H provides one conserved quantity defined on the leaves and that is all in general.

Using the particular hypothesis of the Lagrange top, namely the rotational symmetry around the third axis, it is easy to see that $\vec{J}\cdot\vec{h}$ is a second independent conserved quantity. Indeed multiplying the first eq.(1.14) by $\vec{h}$ we find

$$\frac{d\,\vec{h}\cdot\vec{J}}{dt} = \vec{h}\cdot(\vec{\omega}\wedge\vec{J}) = h_3(\omega_1 J_2 - \omega_2 J_1) = h_3\omega_1\omega_2(I_2 - I_1) = 0$$

where we have used that $\vec{h}$ is along the third axis. Since it is conserved, this quantity Poisson commutes with the Hamiltonian, hence the system is integrable.

To solve this system we describe the top by Euler angles (θ,ϕ,ψ). Recall that the rotation vector and the weight vector can be expressed in the moving frame as:

$$\begin{pmatrix}\omega_1\\ \omega_2\\ \omega_3\end{pmatrix} = \begin{pmatrix}\dot\phi\sin\theta\sin\psi + \dot\theta\cos\psi\\ \dot\phi\sin\theta\cos\psi - \dot\theta\sin\psi\\ \dot\phi\cos\theta + \dot\psi\end{pmatrix}, \qquad \begin{pmatrix}P_1\\ P_2\\ P_3\end{pmatrix} = \begin{pmatrix}P\sin\theta\sin\psi\\ P\sin\theta\cos\psi\\ P\cos\theta\end{pmatrix}$$

The two quantities in the center of the Poisson bracket are $\vec{P}^2 = P^2$ and $\vec{P}\cdot\vec{J} \equiv PJ_z = P(I_1\sin^2\theta + I_3\cos^2\theta)\dot\phi + PI_3\cos\theta\dot\psi$. Moreover the Lagrange conserved quantity reads $\vec{h}\cdot\vec{J} \equiv hJ_3 = hI_3(\dot\phi\cos\theta + \dot\psi)$. This allows to eliminate $\dot\phi$ and $\dot\psi$ in the Hamiltonian:

$$H = \frac{1}{2}I_1(\sin^2\theta\dot\phi^2 + \dot\theta^2) + \frac{1}{2}I_3(\dot\phi\cos\theta + \dot\psi)^2 - P\cos\theta$$

We get a one–dimensional system in the variable θ with Hamiltonian:

$$H = \frac{1}{2}I_1\,\dot\theta^2 + \frac{1}{2I_1}\frac{(J_z - J_3\cos\theta)^2}{\sin^2\theta} - P\cos\theta + \frac{1}{2}\frac{J_3^2}{I_3}$$

It follows that θ is an elliptic function of the time t. Note that the choice of the Euler angle coordinates has disentangled the dynamics, since ϕ and ψ no longer appear. This is related to the symmetry of the problem with respect to both the vertical axis and the axis of the top.

1.10 The Kowalevski top.

There is another much hidden case of integrability of the top, which was discovered by S. Kowalevski. Recall that we consider the motion of the top in a moving frame with origin at the fixed point of the top. Assume that the moments of inertia obey $I_1 = I_2 = 2I_3$. Assume further that the center of mass is on the plane $x_3 = 0$, but away from the origin, so that the top has no rotational symmetry. However we are free to choose the inertia axis up to a rotation around the third one, hence to assume that the fixed point is on the first axis. We introduce traditional notations

$$\vec{\omega} = \begin{pmatrix}p\\ q\\ r\end{pmatrix}, \quad \vec{P} = \begin{pmatrix}\gamma_1\\ \gamma_2\\ \gamma_3\end{pmatrix}, \quad \vec{h} = \begin{pmatrix}h\\ 0\\ 0\end{pmatrix}$$

and we write eqs.(1.14) in components, with $c_0 = h/I_3$:

$$\begin{aligned} 2\dot p &= qr & \dot\gamma_1 &= r\gamma_2 - q\gamma_3\\ 2\dot q &= -pr - c_0\gamma_3 & \dot\gamma_2 &= p\gamma_3 - r\gamma_1\\ \dot r &= c_0\gamma_2 & \dot\gamma_3 &= q\gamma_1 - p\gamma_2 \end{aligned}$$

The previously known conserved quantities are:

$$\vec{P}^2 = \gamma_1^2 + \gamma_2^2 + \gamma_3^2, \quad \vec{P}\cdot\vec{J} = I_3(2p\gamma_1 + 2q\gamma_2 + r\gamma_3), \quad H = \frac{I_3}{2}(2p^2 + 2q^2 + r^2) - h\gamma_1$$

The first two are in the center of the Poisson bracket, so the symplectic leaves are of dimension 4, and we need one further conserved quantity to prove integrability. To introduce it naturally, consider $z = p + iq$ and $\xi = \gamma_1 + i\gamma_2$. The equations of motion give:

$$2\dot{z} = -irz - ic_0\gamma_3, \quad \dot{\xi} = -ir\xi + i\gamma_3 z$$

We can eliminate γ_3 by considering the combination $z^2 + c_0\xi$ which obeys:

$$\frac{d}{dt}(z^2 + c_0\xi) = -ir(z^2 + c_0\xi), \quad \frac{d}{dt}(\overline{z}^2 + c_0\overline{\xi}) = ir(\overline{z}^2 + c_0\overline{\xi})$$

where the second equation is obtained from the first one by complex conjugation. It is then clear that $|z^2 + c_0\xi|^2$ is conserved. In terms of the original variables, we have obtained the Kowalevski conserved quantity:

$$K = (p^2 - q^2 + c_0\gamma_1)^2 + (2pq + c_0\gamma_2)^2 \tag{1.15}$$

Note that the conditions $I_1 = I_2 = 2I_3$ are essential in this calculation. The solution of this model has been obtained by S. Kowalevski and is considerably more involved than in the previous cases.

1.11 The Neumann model.

This model describes the motion of a particle on a sphere S^{N-1} submitted to harmonic forces with generically different frequencies in each direction. It was first introduced by Neumann. An easy description of the model is given in the Lagrangian formulation by introducing a Lagrange multiplier, Λ, and writing the Lagrangian:

$$\mathcal{L} = \sum_{k=1}^{N} \frac{1}{2}\big(\dot{x}_k^2 - a_k x_k^2\big) + \frac{1}{2}\Lambda\,\Big(\sum_{l=1}^{N} x_l^2 - 1\Big)$$

The equations of motion are:

$$\ddot{x}_k = -a_k x_k + \Lambda x_k, \quad \sum_l x_l^2 = 1$$

To compute Λ we multiply by x_k and sum over k, which yields $\Lambda = -\sum_k(\dot{x}_k^2 - a_k x_k^2)$ where we used that the constraint implies $\sum_k(x_k\ddot{x}_k + \dot{x}_k^2) = 0$. This leads to the non–linear Newton equations of motion for the particle:

$$\ddot{x}_k = -a_k x_k - x_k \sum_l (\dot{x}_l{}^2 - a_l x_l^2) \tag{1.16}$$

Conversely if we start with initial conditions satisfying $\sum_k x_k^2 = 1$ and $\sum_k x_k\dot{x}_k = 0$, these conditions are preserved by the time evolution.

It is important for us to cast this model in the Hamiltonian formulation. This is achieved by introducing a symmetry in a larger phase space and then reducing this symmetry. Consider a $2N$-dimensional free phase space $\{x_n, y_n, \quad n = 1\cdots N\}$ with canonical Poisson brackets, $\{x_n, y_m\} = \delta_{nm}$, and introduce the "angular momentum" antisymmetric matrix: $J_{kl} = x_k y_l - x_l y_k$ and the Hamiltonian:

$$H = \frac{1}{4}\sum_{k\neq l} J_{kl}^2 \; + \; \frac{1}{2}\sum_k a_k x_k^2 \tag{1.17}$$

We shall assume in the following that $a_1 < a_2 < \cdots < a_N$. The Hamiltonian equations are, with $X = (x_k), \quad Y = (y_k)$, and the diagonal constant matrix $D = (a_k\delta_{kl})$:

$$\dot{X} = -JX \qquad \dot{Y} = -JY - DX \tag{1.18}$$

The Hamiltonian has a symmetry $Y \to Y + \lambda X$ and $X \to X$, and we can perform a Hamiltonian reduction. The moment map of this action is $M = \frac{1}{2}\sum_k x_k^2$ since $\{M, Y_k\} = X_k, \quad \{M, X_k\} = 0$. We fix the moment to $M = \frac{1}{2}$. The reduced phase space is obtained by taking the quotient by the group of stability of the moment which is here the whole group of symmetry, see Chapter [14]. This amounts to considering functions of X and Y invariant under $Y \to Y + \lambda X$ (or alternatively to impose some gauge condition, i.e. $(X, Y = 0)$. Such invariant functions are provided by the matrix valued functions X and J. The reduced phase space has the correct dimension $2n - 2$ for a point on a sphere. Conversely given X and J an antisymmetric rank 2 matrix whose image contains X, we can always find a vector Y up to the above symmetry such that $J_{kl} = x_k y_l - x_l y_k$. The reduced equations of motion are equivalent to the equations of motion for the Neumann model. Indeed, the reduced system is characterized by the conditions ${}^tXX = 1$ and ${}^tXY = 0$, but the equations of motion eq.(1.18) do not preserve the second condition. We need to perform simultaneously a time dependent gauge transformation $Y \to Y + \lambda(t)X$ to keep the motion on the gauge surface. Writing:

$$0 = \frac{d}{dt}(X, Y + \lambda X) = (-JX, Y + \lambda X) + (X, -JY - DX + \dot{\lambda}X - \lambda JX) = -(X, DX) + \dot{\lambda}$$

since J is antisymmetric, gives $\dot{\lambda} = (X, DX)$. The new equation of motion for $Y' = (Y + \lambda X)$ on the gauge surface is thus:

$$\dot{Y}' = (-JY - DX) - \lambda JX + \dot{\lambda}X = -JY' - DX + (X, DX)X$$

Since $J = J'$ we have $\dot{X} = -J'X = Y'$ and $\dot{Y}' = -(Y', Y')X - DX + (X, DX)X$ so that eliminating Y' we finally get:

$$\ddot{X} = -DX - \Big((\dot{X}, \dot{X}) - (X, DX)\Big)X$$

which is identical to eq.(1.16).

The Liouville integrability of this system is a consequence of the existence of $(N - 1)$ independent quantities in involution, first found by K. Uhlenbeck:

$$F_k = x_k^2 + \sum_{l \neq k} \frac{J_{kl}^2}{a_k - a_l}, \quad \sum_k F_k = 1 \tag{1.19}$$

Notice that the Hamiltonian of the Neumann model can be expressed in terms of the F_k as

$$H = \frac{1}{2}\sum_k a_k F_k$$

The equations of motion can be written in terms of the two gauge invariant matrices $J = X\,{}^tY - Y\,{}^tX$ and $K = X\,{}^tX$

$$\dot{K} = -[J, K], \quad \dot{J} = [D, K]$$

That eq.(1.18) imply the above is a simple computation. Conversely, knowing K it is easy to compute X since K is a projector on X whose length is 1 and then one can compute Y knowing J, up to a gauge.

1.12 Geodesics on an ellipsoid.

The Neumann problem was shown by Moser to contain, in particular, the geodesic motion on an ellipsoid which was found to be integrable by Jacobi. Consider the Hamiltonian belonging to the Neumann conserved quantities, but different from the usual Neumann Hamiltonian:

$$H = \sum_k \frac{1}{a_k} F_k = Q(X,X) - Q(X,X)Q(Y,Y) + Q^2(X,Y), \quad Q(X,Y) = \sum_k \frac{1}{a_k} x_k y_k$$

Under the flow generated by H, H is of course conserved, so we can restrict ourselves on the surface $H = 0$. Remark that if one defines the quantity ξ, invariant under the gauge transformation $Y \to Y + \lambda X$,

$$\xi = Y - \frac{Q(X,Y)}{Q(X,X)} X$$

we have $H = Q(X,X)(1 - Q(\xi,\xi))$ so that the condition $H = 0$ is equivalent to the condition $Q(\xi,\xi) = 1$. The flow of H, for $H = 0$, leave ξ on the ellipsoid $Q(\xi,\xi) = 1$.

We want to show that the trajectories of ξ are geodesics on the ellipsoid. We compute the time evolution of ξ, and for this we remark that since ξ is gauge invariant, one can use the unreduced equations of motion:

$$\begin{aligned}
\dot{x}_k &= \frac{\partial H}{\partial y_k} = -2Q(X,X)\frac{y_k}{a_k} + 2Q(X,Y)\frac{x_k}{a_k} = -2Q(X,X)\frac{\xi_k}{a_k} \\
\dot{y}_k &= -\frac{\partial H}{\partial x_k} = -2(1 - Q(Y,Y))\frac{x_k}{a_k} - 2Q(X,Y)\frac{y_k}{a_k}
\end{aligned}$$

Note that, since X is gauge independent, so is its time derivative. Denoting $s = Q(X,Y)/Q(X,X)$ we then have:

$$\dot{\xi}_k = -2\left(1 - Q(Y,Y) + \frac{Q^2(X,Y)}{Q(X,X)}\right)\frac{x_k}{a_k} - \dot{s}x_k$$

On the surface $H = 0$ this becomes simply:

$$\frac{d\xi_k}{dt} = -\dot{s}x_k \Longrightarrow \frac{d\xi_k}{ds} = -x_k$$

Since $\vec{x}^2 = 1$ the length of the vector $d\vec{\xi}/ds$ is 1, which means that s is the length parameter on the trajectory of ξ. Next we compute:

$$\frac{d^2\xi_k}{ds^2} = \frac{1}{\dot{s}}\frac{dx_k}{dt} = -2\frac{Q(X,X)}{\dot{s}}\frac{\xi_k}{a_k}$$

The vector with components ξ_k/a_k is the gradient of $Q(\xi,\xi)$ hence is normal to the ellipsoid $Q(\xi,\xi) = 1$. The fact that the second derivative of $\vec{\xi}$ with respect to the length parameter s is normal to the surface characterizes geodesics, as we now show. We have to minimize the arc length, on a surface $f(\xi) = 0$, that is to minimize:

$$\int \left(\sqrt{\dot{\vec{\xi}} \cdot \dot{\vec{\xi}}} + \Lambda f(\xi)\right) dt$$

where Λ is a Lagrange parameter. The Euler–Lagrange equation reads:

$$\frac{d}{dt}\frac{\dot{\vec{\xi}}}{\sqrt{\dot{\vec{\xi}} \cdot \dot{\vec{\xi}}}} = \Lambda \nabla f(\xi)$$

So the derivative of the normalized velocity vector is perpendicular to the surface.

The geodesic motion on an ellipsoid was solved originally by Jacobi by introducing ellipsoidal coordinates which separate the variables of this problem. As they separate the variables of the Neumann model as well, we now explain this method for the Neumann model.

1.13 Separation of variables in the Neumann model.

Following Jacobi and Neumann, we introduce $(N-1)$ parameters $t_1, \cdots, t_{N-1}$ on the sphere. They are the roots of the equation:

$$u(t) \equiv \sum_k \frac{x_k^2}{t - a_k} = 0$$

Notice that this equation is invariant by $x_k \longrightarrow \lambda x_k$ so that $t_1 < t_2 < \cdots < t_{N-1}$ are indeed defined on the sphere. Conversely, by definition of the t_j we have for $x \in S^{(N-1)}$:

$$u(t) = \frac{\prod_j (t - t_j)}{\prod_k (t - a_k)} \quad \Longrightarrow \quad x_k^2 = \frac{\prod_j (a_k - t_j)}{\prod_{l \neq k} (a_k - a_l)} \tag{1.20}$$

Considering the graph of $u(t)$ it is easy to see that:

$$a_1 < t_1 < a_2 < t_2 < a_3 < \cdots < t_{N-1} < a_N$$

and we have a bijection of this domain $\mathcal{D}$ of the t_j's on the "quadrant" $x_k > 0 \ \forall k$ of the sphere.

Proposition. *The (t_j)'s define an orthogonal system of coordinates on the sphere.*

Proof. Let us consider, for each root t_j, the vector:

$$\vec{v}_j = \left(\frac{x_1}{t_j - a_1}, \cdots, \frac{x_N}{t_j - a_N} \right)$$

Since t_j solves $u(t) = 0$, we have $\vec{x} \cdot \vec{v}_j = 0$. Moreover

$$\vec{v}_j \cdot \vec{v}_{j'} = \sum_k \frac{x_k^2}{(t_j - a_k)(t_{j'} - a_k)} = -\frac{u(t_j) - u(t_{j'})}{t_j - t_{j'}} = 0, \quad \text{if } j \neq j'$$

Taking the derivatives of the equations $u(t_j) = 0$ with respect to $t_1, \cdots, t_{N-1}$, where the x_k are considered as functions of these variables, we obtain $\vec{v}_i \cdot \partial_{t_j} \vec{x} = 0$ if $i \neq j$ and $2\vec{v}_j \cdot \partial_{t_j} \vec{x} = \vec{v}_j^2$. Therefore:

$$\frac{\partial \vec{x}}{\partial t_j} = \frac{1}{2} \vec{v}_j. \tag{1.21}$$

are $(N-1)$ orthogonal vectors in the tangent plane to the sphere $S^{(N-1)}$ at the point $\vec{x}$. ∎

As a byproduct, since $\vec{v}_j^2 = -u'(t_j)$ we also get the metrical tensor:

$$g_{jj'} = \frac{\partial \vec{x}}{\partial t_j} \cdot \frac{\partial \vec{x}}{\partial t'_j} = -\frac{1}{4} \delta_{jj'} \frac{\prod_{i \neq j} (t_j - t_i)}{\prod_k (t_j - a_k)}$$

We now consider the canonical 1–form $\alpha = \sum_k y_k dx_k$ which reduces on the sphere to $\alpha = \sum_j p_j dt_j$ with

$$p_j = \sum_k y_k \frac{\partial x_k}{\partial t_j} = \frac{1}{2} \vec{y} \cdot \vec{v}_j$$

These $(N-1)$ equations determine $\vec{y}$ up to a vector proportional to $\vec{x}$ which does not affect the value of J_{kl}. We find $\vec{y} = 1/2 \sum_j g^{jj} p_j \vec{v}_j$ which easily gives:

$$J_{kl} = -\frac{1}{2} \sum_j (a_k - a_l) \, v_j^k v_j^l \, g^{jj} p_j$$

With this, we can compute the conserved quantities F_k, eq.(1.19), in terms of the new canonical coordinates t_i, p_j:

$$F_k = x_k^2 + \frac{1}{4}\sum_{jj'}\left(a_k \vec{v_j}\cdot\vec{v_{j'}} - \sum_l a_l v_j^l v_{j'}^l\right) v_j^k v_{j'}^k g^{jj} g^{j'j'} p_j p_{j'}$$

Remarking that $\vec{v_j}\cdot\vec{v_{j'}} = 4g_{jj}\delta_{jj'}$ and $\sum_l a_l v_j^l v_{j'}^l = 4t_j g_{jj}\delta_{jj'}$, we obtain:

$$F_k = x_k^2\left(1 - \sum_j \frac{g^{jj}\,p_j^2}{t_j - a_k}\right)$$

It is convenient to introduce the generating function for the F_k:

$$\mathcal{H}(\lambda) \equiv \sum_k \frac{F_k}{\lambda - a_k} = \frac{\prod_n(\lambda - b_n)}{\prod_k(\lambda - a_k)} \tag{1.22}$$

for appropriate b_n's, $n = 1, \cdots, N-1$, and we have used $\sum F_k = 1$ to normalize the leading coefficient in the numerator. By a simple calculation we find:

$$\mathcal{H}(\lambda) = u(\lambda)\left(1 - \sum_j \frac{g^{jj}p_j^2}{t_j - \lambda}\right) \tag{1.23}$$

Following the general strategy of the Liouville theorem, we express the momenta p_j in terms of the conserved quantities F_k and the t_j. We have:

$$g^{jj}p_j^2 = \lim_{\lambda\to t_j} \frac{\lambda - t_j}{u(\lambda)}\mathcal{H}(\lambda)$$

Taking into account the value of the metrical tensor and eq(1.20), we get $p_j^2 = -\frac{1}{4}\mathcal{H}(t_j)$. Then the function $S = \int \sum_j p_j dt_j$ reads:

$$S = \frac{1}{2}\sum_j \int^{t_j} \sqrt{-\mathcal{H}(t)}\; dt$$

Obviously, the variables have been separated in the ellipsoidal coordinates.

Choosing as independent action variables the $(N-1)$ independent quantities b_n instead of the N dependent quantities F_k, the conjugate variables ψ_n are:

$$\psi_n = \frac{\partial S}{\partial b_n} = -\frac{1}{4}\sum_j \int^{t_j} \sqrt{-\mathcal{H}(t)}\frac{dt}{t - b_n} \tag{1.24}$$

By the Liouville theorem, the time evolution of the ψ_n's under the Hamiltonian $\mathcal{H}(\lambda)$ is linear:

$$\dot{\psi}_n = \frac{\partial\mathcal{H}(\lambda)}{\partial b_n} = -\frac{\mathcal{H}(\lambda)}{\lambda - b_n} = -\frac{1}{4}\sum_j \frac{\dot{t}_j\sqrt{-\mathcal{H}(t_j)}}{t_j - b_n} \tag{1.25}$$

where the last equality is obtained by derivating eq.(1.24) with respect to time. This gives the evolution of the t_j's.

This can be rewritten in a more geometrical way. Denoting by τ the time variable, and introducing the polynomial of degree $(2N-1)$:

$$P(t) = \prod_{i=1}^{N}(t - a_i)\prod_{n=1}^{N-1}(t - b_n)$$

we can rewrite eq.(1.25) as the set of $N-1$ equations:

$$\sum_j \frac{Q_n(t_j)dt_j}{\sqrt{-P(t_j)}} = 4\frac{d\tau}{\prod_i(\lambda - a_i)}Q_n(\lambda), \quad Q_n(\lambda) = \prod_{m\neq n}(\lambda - b_m),\ n = 1, \cdots N-1$$

Since the Q_n are $N-1$ linearly independent polynomials of degree $N-2$ this system is equivalent to the following one:

$$\sum_j \frac{t_j^k dt_j}{\sqrt{-P(t_j)}} = 4\frac{\lambda^k}{\prod_i(\lambda - a_i)} d\tau, \quad k = 0, \cdots N-2 \tag{1.26}$$

The quantities $t^k dt/\sqrt{-P(t)}$, $k = 0, \cdots N-2$, are the Abelian differentials of first kind on the hyperelliptic Riemann surface of genus $g = N-1$ given by the equation $s^2 + P(t) = 0$. The sums appearing in the left–hand side are thus Abel sums, and describe a point in the Jacobian of the curve. Eq.(1.26) shows that this point moves linearly in time.

In these historical examples, we see a relation between integrability, separation of variables and Riemann surfaces. This is a phenomenon that will reappear in a much broader context in the following.

References

[1] J.L. Lagrange. **Mécanique analytique.** (1788) in. Oeuvres de Lagrange, Tome XII, Gauthier–Villars, (1889) Paris.

[2] C. Jacobi. **Vorlesungen über Dynamik**, Gesammelte Werke, Supplement band, (1884), Berlin.

[3] J. Liouville. *"Note sur l'intégration des équations différentielles de la dynamique"*. Journal de Mathématiques (Journal de Liouville) T. XX (1855), p. 137.

[4] C. Neumann. *"De problemate quodam mechanico, quod ad primam integralium ultraellipticorum classem revocatur."* Crelle Journal vol 56 (1859) p.46–63.

[5] Sophie Kowalevski *"Sur le problème de la rotation d'un corps solide autour d'un point fixe."* Acta Mathematica vol 12 (1889) p.177–232.

[6] C. S. Gardner, J. M. Greene, M. D. Kruskal, R. M. Miura. *"Method for solving the Korteweg-de Vries equation"*. Phys. Rev. Lett. 19 (1967), p. 1095.

[7] P. D. Lax. *"Integrals of non linear equations of evolution and solitary waves"*. Comm. Pure Appl. Math. 21 (1968), p. 467.

[8] L. Landau, E. Lifchitz **Mécanique.** MIR (1969) Moscou.

[9] K. Uhlenbeck. *Minimal 2–spheres and tori in* S^k. Preprint (1975).

[10] V. Arnold. **"Méthodes mathématiques de la mécanique classique"**. MIR (1976). Moscou.

[11] J. Moser. *Various aspects of integrable Hamiltonian systems.* Proc. CIME Bressanone, Progress in Mathematics, Vol8 Birkhauser (1978) p. 233

[12] E. K. Sklyanin. *"On the complete integrability of the Landau-Lifchitz equation"*. Preprint LOMI E-3-79. Leningrad 1979.

[13] M. Semenov Tian Shansky. *"What is a classical* r*-matrix"*. Funct. Anal. and its Appl. 17, 4 (1983), p. 17.

[14] D. Mumford. **Tata lectures on Theta II.** Progress in Mathematics. Vol. 43 (1984) Birkhäuser.

Chapter 2

Construction of integrable systems.

Given an integrable system, there is no useful known procedure to construct a Lax pair. Moreover, Lax pairs are not unique. In this chapter, we introduce families of Lax matrices $L(\lambda)$ depending analytically on a parameter λ. The study of the analytical properties of the Lax equation $\dot{L}(\lambda) = [L(\lambda), M(\lambda)]$ yields considerable insight into the structure of this equation, and in fact, specifies the possible form of the matrix $M(\lambda)$. Lax matrices with a spectral parameter are interpreted as coadjoint orbits. This introduces a symplectic structure into the framework and we show that the dynamical systems thus obtained are Liouville integrable systems. Many integrable systems are shown to fit into this framework, including the case of an infinite number of degrees of freedom.

2.1 Examples of Lax pairs with spectral parameter.

Example 1. Our first example will be provided by the Euler top, see section (1.8) in Chapter[1]. In this case, a Lax pair appears naturally. Let us introduce the 3×3 matrices $J_{ij} = \epsilon_{ijk} J_k$ and $\Omega_{ij} = \epsilon_{ijk}\omega_k$. Then the equation of motion $\frac{d\vec{J}}{dt} = -\vec{\omega} \wedge \vec{J}$ can be recast in matrix form:

$$\frac{dJ}{dt} = [\Omega, J]$$

This is a Lax pair with $L = J$, and $M = -\Omega$, but unfortunately the conserved quantities, $\mathrm{Tr}\, L^n$, either vanish or are functions of $\vec{J}^2$, and therefore the Hamiltonian is not included in this set of conserved quantities (recall that $\vec{J}^2$ is in the center of the Poisson bracket).

To cure this problem some modifications are needed. Let us introduce a diagonal matrix: $\mathcal{I} = \mathrm{Diag}(\mathcal{I}_1, \mathcal{I}_2, \mathcal{I}_3)$ with $\mathcal{I}_k = \frac{1}{2}(I_i + I_j - I_k)$, where (i, j, k) is a cyclic permutation of $(1, 2, 3)$. With these notations we have

$$J = \mathcal{I}\Omega + \Omega\mathcal{I}$$

We assume that all $\mathcal{I}_j$ are different and we set:

$$L(\lambda) = \mathcal{I}^2 + \frac{1}{\lambda} J, \quad M(\lambda) = -(\lambda\mathcal{I} + \Omega) \tag{2.1}$$

where λ is a free arbitrary parameter, later called the spectral parameter. To check that the Lax equation gives back the equations of motion, we compute:

$$\dot{L} - [L, M] = [J, \mathcal{I}] + [\mathcal{I}^2, \Omega] + \frac{1}{\lambda}(\dot{J} + [J, \Omega])$$

The first two terms cancel, while the vanishing of the $1/\lambda$–term gives the equations of motion. This Lax pair is much better than the previous one because

$$\mathrm{Tr}\, L^2 = \mathrm{Tr}\, \mathcal{I}^4 - \frac{2}{\lambda^2}\vec{J}^2$$

$$\mathrm{Tr}\, L^3 = \mathrm{Tr}\, \mathcal{I}^6 + \frac{3}{\lambda^2}\mathrm{Tr}\, \mathcal{I}^2 J^2 = \mathrm{Tr}\, \mathcal{I}^6 - \frac{3}{\lambda^2}\left(\frac{1}{4}(\mathrm{Tr}\, I)^2 \vec{J}^2 - I_1 I_2 I_3\, H\right)$$

hence we now do have the Hamiltonian amongst the conserved quantities of the form $\mathrm{Tr}\, L^n$.

The new important point is that the Lax matrix depends on a spectral parameter λ and this was necessary to generate the proper conserved quantities.

Example 2. As a second example we consider the Lagrange top. The matrices $L(\lambda)$ and $M(\lambda)$ are written as 4×4 matrices in block form.

$$L(\lambda) = \begin{pmatrix} 0 & I\, {}^t h + \lambda^{-2}\, {}^t P \\ Ih + \lambda^{-2} P & \lambda^{-1} J \end{pmatrix}, \quad M(\lambda) = \begin{pmatrix} 0 & -\lambda\, {}^t h \\ -\lambda h & -\Omega \end{pmatrix} \tag{2.2}$$

where the 3×3 matrices J and Ω are as in the previous example, and h and P are 3×1 matrices corresponding to the vectors $\vec{h}$ and $\vec{P}$ of the Lagrange top. Moreover I stands for the two equal moments of inertia of the top $I = I_1 = I_2$. Let us write the Lax equation:

$$\dot{L} - [L, M] = \begin{pmatrix} 0 & I\, {}^t h\Omega - {}^t hJ + \lambda^{-2}\left({}^t\dot{P} + {}^t P\Omega\right) \\ -I\Omega h + Jh + \lambda^{-2}(\dot{P} - \Omega P) & \lambda^{-1}\big(\dot{J} + [J, \Omega] + P\, {}^t h - h\, {}^t P\big) \end{pmatrix}$$

Due to the Lagrange condition $I = I_1 = I_2$ we have $I\Omega h = Jh$ and the vanishing of the other elements reduces to the equations of motion $\dot{J} = [\Omega, J]$, $\quad \dot{P} = \Omega P$.

Example 3. Finally consider the Neumann model. As we have seen in Chapter [1] the equations of motion are:

$$\dot{K} = -[J, K], \quad \dot{J} = [D, K]$$

To recast these two relations into the Lax form we introduce the matrices

$$L(\lambda) = D + \frac{1}{\lambda}J - \frac{1}{\lambda^2}K, \quad M(\lambda) = \frac{1}{\lambda}K \tag{2.3}$$

We then have:

$$\dot{L}(\lambda) - [L, M] = \frac{1}{\lambda}(\dot{J} - [D, K]) - \frac{1}{\lambda^2}(\dot{K} + [J, K])$$

The Lax equation with spectral parameter is equivalent to the vanishing of the two coefficients $\dot{J} - [D, K]$ and $\dot{K} + [J, K]$ since the first one is an antisymmetric matrix and the second one is a symmetric matrix. Hence the Lax equation is equivalent to the equations of motion of the Neumann model which are independent of λ.

2.2 The Zakharov-Shabat construction.

The modern developments of integrable systems rely on the notion of Lax pairs. Given an integrable system, there does not yet exists a useful algorithm to construct a Lax pair. There do exist however a general procedure to construct Lax pairs giving rise to integrable systems. Since integrable systems are rather scarce, we expect that this construction will not be a completely trivial matter. We will present a general method to construct matrices $L(\lambda)$ and $M(\lambda)$, depending on a spectral parameter λ and of a special form such that the Lax equation :

$$\partial_t L(\lambda) = [L(\lambda), M(\lambda)] \tag{2.4}$$

is equivalent to the equations of motion of an integrable system.

The method is due to Zakharov and Shabat. It consists in specifying the analytical properties of the matrices $L(\lambda)$ and $M(\lambda)$, $\lambda \in \mathbb{C}$. We will first consider systems with a finite number of degrees of freedom. The field theory case with an infinite number of degrees of freedom will be considered in a following section.

We introduce first a notation for any matrix valued rational function $f(\lambda)$ with poles of order n_k at points λ_k at finite distance. Let $f_k(\lambda)$ denotes its polar part a λ_k, so that we can write (if all poles are at finite distance)

$$f(\lambda) = f_0 + \sum_k f_k(\lambda), \quad \text{with} \quad f_k(\lambda) = \sum_{r=-n_k}^{-1} f_{k,r}(\lambda - \lambda_k)^r,$$

with f_0 constant. When there is no ambiguity about the pole we are considering, we will often use the alternative notation $f_-(\lambda) \equiv f_k(\lambda)$. Around one of the point λ_k, $f(\lambda)$ may be decomposed as follows:

$$f(\lambda) = f(\lambda)_+ + f(\lambda)_- \tag{2.5}$$

with $f(\lambda)_+$ regular at the point λ_k and $f(\lambda)_- = f_k(\lambda)$ is the singular part also called the principal or polar part of $f(\lambda)$ at λ_k.

Let us consider matrices $L(\lambda)$ and $M(\lambda)$ of dimension $N \times N$. We will assume that the matrices $L(\lambda)$ and $M(\lambda)$ are rational functions of the parameter λ. Let $\{\lambda_k\}$ be the set of their poles. With the above notations we can write quite generally

$$L = L_0 + \sum_k L_k, \quad \text{with} \quad L_k = \sum_{r=-n_k}^{-1} L_{k,r}(\lambda - \lambda_k)^r \tag{2.6}$$

and

$$M = M_0 + \sum_k M_k \quad \text{with} \quad M_k = \sum_{r=-m_k}^{-1} M_{k,r}(\lambda - \lambda_k)^r \tag{2.7}$$

Here n_k and m_k refer to the order of the poles at the corresponding point λ_k. The coefficients $L_{k,r}$ and $M_{k,r}$ are matrices. We will assume that the positions of the poles λ_k are constants independent of time.

The Lax equation (2.4), with $L(\lambda)$ and $M(\lambda)$ given by eqs.(2.6,2.7), holds identically in λ. Looking at eqs.(2.4) we see that the pole λ_k in the left hand side is a priori of order n_k while in the right hand side it is of order $n_k + m_k$. Hence we have two types of equations. The first type, does not contain time derivatives and comes from setting to zero the coefficients of the poles of order greater than n_k in the right hand side of the equation. This will be interpreted as m_k constraint equations on M_k. The equations of the second type are obtained by matching the coefficients of the poles of order less or equal to n_k on both sides of the equation. These equations contain time derivatives and are thus the true dynamical equations.

Proposition. *Assuming that $L(\lambda)$ has distinct eigenvalues in a neighbourhood of λ_k, one can perform a* regular *similarity transformation $g^{(k)}(\lambda)$ diagonalizing $L(\lambda)$ in a vicinity of λ_k.*

$$L(\lambda) = g^{(k)}(\lambda)\ A^{(k)}(\lambda)\ g^{(k)-1}(\lambda) \tag{2.8}$$

where $A^{(k)}(\lambda)$ is diagonal and has a pole of order n_k at λ_k. Moreover

$$M_k(\lambda) = (g^{(k)}(\lambda)\ B^{(k)}(\lambda)\ g^{(k)-1}(\lambda))_- \tag{2.9}$$

where $B^{(k)}(\lambda)$ is diagonal and has a pole of order m_k at λ_k.

Proof. Demanding that $L(\lambda)$ has distinct eigenvalues in a neighbourhood of λ_k means that $L_{k,-n_k}$ has distinct eigenvalues. Then the matrix $Q(\lambda) = (\lambda - \lambda_k)^{n_k} L(\lambda)$ which is regular at λ_k can be diagonalized in vicinity of λ_k with a regular matrix $g^{(k)}(\lambda)$. This proves eq.(2.8). Then defining $B^{(k)}(\lambda)$ by

$$M = g^{(k)}\ B^{(k)}\ g^{(k)-1} - \partial_t g^{(k)}\ g^{(k)-1} \tag{2.10}$$

the Lax equation becomes:

$$\dot{A}^{(k)}(\lambda) = [A^{(k)}(\lambda), B^{(k)}(\lambda)]$$

This implies $\dot{A}^{(k)} = 0$ as expected (because the commutator with a diagonal matrix has no element on the diagonal), and moreover if we assume that the diagonal elements of $A^{(k)}$ are all distinct this equation implies that $B^{(k)}$ is also *diagonal*. Finally the term $\partial_t g^{(k)}\ g^{(k)-1}$ is regular and does not contribute to the singular part M_k of M at λ_k. Hence $M_k = (g^{(k)}\ B^{(k)}\ g^{(k)-1})_-$ which only depends on $B^{(k)}_-$. It is worth noting that the first n coefficients of the expansion of $g^{(k)}(\lambda)$ only depend on the first n coefficients of the expansion of $Q(\lambda)$. The matrix $g^{(k)}(\lambda)$ is defined up to a right multiplication by an arbitrary analytic diagonal matrix. ∎

As a result, the decomposition of $L(\lambda)$ and $M(\lambda)$ in polar parts reads:

$$L = L_0 + \sum_k L_k, \quad \text{with} \quad L_k = \left(g^{(k)} A^{(k)} g^{(k)-1}\right)_- \tag{2.11}$$

and

$$M = M_0 + \sum_k M_k, \quad \text{with} \quad M_k = \left(g^{(k)} B^{(k)} g^{(k)-1}\right)_- \tag{2.12}$$

Only the singular parts of $A^{(k)}$ and $B^{(k)}$ contribute to L_k and M_k. The independent parameters in $L(\lambda)$ are thus L_0, the singular diagonal matrices $A^{(k)}_-$ of the form

$$A^{(k)}_- = \sum_{r=-n_k}^{-1} A_{k,r}(\lambda - \lambda_k)^r \tag{2.13}$$

and jets of regular matrices $\widehat{g}^{(k)}$ of order $n_k - 1$, defined up to right multiplication by a regular diagonal matrix $d^{(k)}(\lambda)$:

$$\widehat{g}^{(k)} = \sum_{r=0}^{n_k-1} g_{k,r}(\lambda - \lambda_k)^r \tag{2.14}$$

From these data we define $L = L_0 + \sum_k L_k$ with

$$L_k = \left(\widehat{g}^{(k)}\ A^{(k)}_-\ \widehat{g}^{(k)-1}\right)_- \tag{2.15}$$

Then around each λ_k, one can diagonalize $L(\lambda) = g^{(k)} A^{(k)} g^{(k)-1}$, yielding an extension of the matrices $A^{(k)}_-$ and $\widehat{g}^{(k)}$ to formal series $A^{(k)}$ and $g^{(k)}$ in $(\lambda - \lambda_k)$. Finally to define $M(\lambda) = M_0 + \sum_k M_k$, we choose a set of singular matrices $B^{(k)}(\lambda)$ and define M_k by eq.(2.9).

So, in the vicinity of a singularity, $L(\lambda)$ and $M(\lambda)$ can be simultaneously diagonalised if the Lax equation holds true. In this diagonal gauge, the Lax equation is trivial and simply states that the matrix $A^{(k)}(\lambda)$ is conserved. When we transform back into the original gauge however, we get a non trivial result

Proposition. *Let $L(\lambda)$ be a Lax matrix of the form eq.(2.6), $L = L_0 + \sum_k L_k$, The general solution $M(\lambda)$ of the constraints in the Lax equation is $M = M_0 + \sum_k M_k$ with*

$$M_k = \left(P^{(k)}(L,\lambda)\right)_- \tag{2.16}$$

where $P^{(k)}(L,\lambda)$ is a polynomial in L with coefficients rational in λ and $(\quad)_-$ denotes the singular part at $\lambda = \lambda_k$.

Proof. It is easy to show that this is indeed a solution. We have to check that the order of the poles is correct. Let us look at what happens around $\lambda = \lambda_k$. Using a beautiful argument first introduced by Gelfand and Dickey we write:

$$\begin{aligned} [L, M_k]_- &= \left[L, \left(P^{(k)}(L,\lambda)\right)_-\right]_- \\ &= \left[L, P^{(k)}(L,\lambda) - \left(P^{(k)}(L,\lambda)\right)_+\right]_- = -\left[L, \left(P^{(k)}(L,\lambda)\right)_+\right]_- \end{aligned}$$

where we used that a polynomial in L commutes with L. From this we see that the order of the pole at λ_k is less than n_k. To show that this is a general solution, recall eqs. (2.8, 2.10). Since $A^{(k)}(\lambda)$ is a diagonal $N \times N$ matrix with all its elements distinct in a vicinity of λ_k, its powers 0 up to $N-1$ span the space of diagonal matrices and one can write

$$B^{(k)} = P^{(k)}(A^{(k)},\lambda) \tag{2.17}$$

where $P^{(k)}(A^{(k)},\lambda)$ is a polynomial of degree $N-1$ in $A^{(k)}$. The coefficients of $P^{(k)}$ are rational combinations of the matrix elements of $A^{(k)}$ and $B^{(k)}$ hence admit Laurent expansions in $\lambda-\lambda_k$ in a vicinity of λ_k. Inserting eq. (2.17) into eq. (2.10) one gets $M = P^{(k)}(L,\lambda) - \partial_t g^{(k)}\, g^{(k)-1}$. Taking the singular part at λ_k the last term does not contribute and one gets $M_k = \left(P^{(k)}(L,\lambda)\right)_-$. Moreover in this formula the Laurent expansions of the coefficients of $P^{(k)}$ can be truncated at some positive power of $\lambda - \lambda_k$ since a high enough power cannot contribute to the singular part, yielding a polynomial with coefficients Laurent *polynomials* in $\lambda - \lambda_k$. ■

It is important to realize that the dynamical variables are the matrix elements of the Lax matrix, or the matrix elements of the $L_{k,r}$. Choosing the number and the order of the poles of the Lax matrix amounts to specifying a particular model. Choosing the polynomials $P^{(k)}(L,\lambda)$ amounts to specify the dynamical flows.

The above propositions give the general form of $M(\lambda)$ as far as the matrix structure and the λ–dependence is concerned. One should keep in mind however that the coefficients of the polynomials $P^{(k)}(L,\lambda)$ are also functions of the matrix elements of L and require further characterizations in order to get an integrable system which will be the subject of the next section.

Remark 1. If λ_k is a pole of $M(\lambda)$ and not a pole of $L(\lambda)$, one can redefine $M(\lambda)$ without changing the Lax equations of motion so as to eliminate the singularities of $M(\lambda)$ at λ_k. Indeed, redefining:

$$M(\lambda) \to M(\lambda) - P^{(k)}(L,\lambda)$$

does not change the Lax equation. The new $M(\lambda)$ is regular at λ_k. Of course we cannot eliminate the poles common to $L(\lambda)$ and $M(\lambda)$ by this procedure.

Remark 2. The Lax equation is invariant under similarity transformations,

$$L \to L' = g^{-1} L g M \to M' = g^{-1} M g + g^{-1}\partial_t g \tag{2.18}$$

If this similarity transformation is independent of λ, it will not spoil the analytic properties of $L(\lambda)$ and $M(\lambda)$. We can use the gauge freedom eq.(2.18) to diagonalize L_0,

$$L_0 = \mathrm{Diag}(a_1, \cdots, a_N)$$

Consistency of eq.(2.4) then requires M_0 to be also diagonal and thus $\dot{L}_0 = [L_0, M_0] = 0$. Hence M_0 is a polynomial P of L_0, so that replacing $M(\lambda) \to M(\lambda) - P(L(\lambda)$ gets rid of M_0.

We illustrate theses constructions on some examples and verify that the matrices $L(\lambda)$ and $M(\lambda)$ are indeed related as in eq.(2.16). We also compute the singular part of the matrix $A^{(k)}(\lambda)$ and find that they are independent of the dynamical variables in these examples. We will show in the next section that eq.(2.15) admits an important interpretation as a coadjoint orbit when $A_-^{(k)}$ are constant numerical matrices i.e. non dynamical matrices.

Example 1. Let us consider the Euler top. We see that $L(\lambda)$, eq.(2.1), has a pole at 0 and $M(\lambda)$ has a pole at ∞. Let us apply the above procedure to remove this pole. There exists a polynomial $P(x) = \alpha x^2 + \beta x + \gamma$ such that $P(\mathcal{I}^2) = \mathcal{I}$. We will need the coefficient $\alpha = -1/I_1 I_2 I_3$. Redefining $M(\lambda)$ to $M(\lambda) + \lambda P(L(\lambda))$ one gets $M = M_0 + (\alpha/\lambda) J^2$ with $M_0 = -\Omega + \alpha(\mathcal{I}^2 J + J\mathcal{I}^2) + \beta J$. One can check that $M_0 = 0$ (Hint: for $i \neq j$ compute $(\mathcal{I}_i - \mathcal{I}_j)(M_0)_{ij}$ using $P(\mathcal{I}_i^2) = \mathcal{I}_i$). Hence for the Euler top we can choose

$$M(\lambda) = \frac{\alpha}{\lambda} J^2 \tag{2.19}$$

We see that this new $M(\lambda)$ is such that $M(\lambda) = \alpha(\lambda L^2)_-$. The Lax matrix of the Euler top $L(\lambda) = \mathcal{I}^2 + \lambda^{-1} J$ is of the form $L_0 + L_-$ with L_0 diagonal non–dynamic. The eigenvalues of J are $(0, i\sqrt{\vec{J}^2}, -i\sqrt{\vec{J}^2})$, which is non–dynamical since $\vec{J}^2$ belongs to the center of the Poisson bracket and has been fixed to numerical value.

Example 2. For the Lagrange top $L(\lambda)$, eq.(2.2), has a pole at 0 and $M(\lambda)$ has a pole at ∞. One can remove this pole at ∞ by redefining $M(\lambda) \to M(\lambda) + I^{-1}\lambda L(\lambda)$. Notice however that since the eigenvalues of L_0 are degenerate one cannot express M_0 as a polynomial in L_0. We have $M = M_0 + M_-$ with $M_-(\lambda) = I^{-1}(\lambda L(\lambda))_-$. For the Lagrange top the Lax matrix is again of the form $L_0 + L_-$ where obviously L_0 is non–dynamical. For the singular part one gets: $A_- = \lambda^{-2}\mathrm{Diag}(\sqrt{\vec{P}^2}, -\sqrt{\vec{P}^2}, 0, 0)$ which again belongs to the center of the Poisson bracket and is non–dynamical.

Example 3. Finally, let us consider the Neumann model. We see from eq.(2.3) that we have $M(\lambda) = (-\lambda L(\lambda))_-$ where we project on the singular part at $\lambda = 0$. The Lax matrix of the Neumann model is of the form $L = L_0 + L_-$ with $L_0 = D$ which is a numerical diagonal matrix, and the singular part L_- at the only pole $\lambda = 0$ is given by: $L_- = \lambda^{-1} J - \lambda^{-2} K$. This is a rank two matrix whose image is spanned by the vectors x and y. It is easy to diagonalize in this subspace and one gets $A_- = \lambda^{-2}\mathrm{Diag}(1, 0, \cdots, 0)$. It is indeed a numerical matrix.

Remark 3. For Lax matrices $L(\lambda)$ and $M(\lambda)$ rational functions of λ, we can easily compare the number of variables to the number of equations contained in eq. (2.4). The variables are the matrices $L_0,\ L_{k,r}$ and $M_0,\ M_{k,r}$. A naive counting, assuming that $L(\lambda)$ and $M(\lambda)$ are generic and independent matrices, gives:

$$\begin{aligned} \#\ \text{variables} &= 2 + \sum_k n_k + \sum_k m_k = 2 + l + m \\ \#\ \text{equations} &= 1 + \sum_k (n_k + m_k) = 1 + l + m \end{aligned}$$

where l and m are the total order (degree of the divisor) of the poles of $L(\lambda)$ and $M(\lambda)$ respectively. Therefore there is one more variable as equations, which reflects the gauge invariance eq. (2.18) of the Lax equation.

If we assume however that λ belongs to a higher genus Riemann surface with genus g the situation is very different. Indeed suppose that $L(\lambda)$ and $M(\lambda)$ have poles of total multiplicity l and m respectively. Let us count the number of meromorphic functions on which $L(\lambda)$ and $M(\lambda)$ can be expanded with constant matrix coefficients. By the Riemann Roch theorem, $L(\lambda)$ can be expanded on a basis of $(l-g+1)$ independent meromorphic functions (in the generic case), and $M(\lambda)$ on $(m-g+1)$ functions. So we have

$$\# \text{ variables} \quad = \quad 2+l+m-2g$$

Similarly, the commutator $[L(\lambda), M(\lambda)]$ has poles of total multiplicity $l+m$ and can be expanded on a basis of $(l+m-g+1)$ independent functions. Therefore

$$\# \text{ equations} \quad = \quad 1+l+m-g$$

So $(\# \text{ equations} - \# \text{ variables}) = g-1$. Taking into account the gauge symmetry of the Lax equation, we see that the number of equations is always greater than the number of unknowns when $g>0$. This shows that if λ belongs to a Riemann surface of genus $g \geq 1$, (and such systems exist, at least for $g=1$, see Chapter [6]), one has to consider a non generic situation.

Remark 4. As we already mentioned, the dynamical variables are the matrix elements of the $L_{k,r}$. In most cases this is too general and we may try to impose more restrictions on the matrix elements of the $L_{k,r}$. For instance, assuming that the λ_k are all real, then clearly one can impose that all the matrices L_k and M_k are anti-Hermitian. Another simple example is provided by the Neumann model whose Lax matrices (2.3) are such that ${}^tL(-\lambda) = L(\lambda)$, ${}^tM(-\lambda) = -M(\lambda)$. More generally, we may define the action of a reduction group R and impose that $L(\lambda)$ and $M(\lambda)$ are invariant under R. This may be done for example by demanding that the Lax pair $L(\lambda)$ and $M(\lambda)$ satisfies:

$$\widehat{R}^{-1} L(r(\lambda)) \widehat{R} = L(\lambda), \quad \widehat{R}^{-1} M(r(\lambda)) \widehat{R} = M(\lambda)$$

for some representations $\widehat{R}$ and $r(\lambda)$ of R. This type of restriction is always compatible with the Lax equation. It provides a way to lower the number of degrees of freedom. An example of this procedure can be found in Chapter [12].

2.3 Coadjoint orbits and Hamiltonian formalism.

The building blocks L_k of the Lax matrix are of the form eq.(2.15). When $A_-^{(k)}$ is a constant independent of the dynamical variables, this very specific form admits a simple group theoretical interpretation that we now explain. This will allow us to introduce a natural symplectic structure for these systems.

We first recall some notions about adjoint and coadjoint actions of Lie algebras and Lie groups. Let G be a connected Lie group with Lie algebra $\mathcal{G}$. The group G acts on $\mathcal{G}$ by the adjoint action denoted Ad:

$$X \longrightarrow (\mathrm{Ad}\, g)(X) = gXg^{-1} \quad g \in G, \ X \in \mathcal{G}$$

Similarly the coadjoint action of G on the dual $\mathcal{G}^*$ of the Lie algebra $\mathcal{G}$ (i.e. the vector space of linear forms on the Lie algebra) is defined by:

$$(\mathrm{Ad}^*\, g.\Xi)(X) = \Xi(\mathrm{Ad}\, g^{-1}(X)), \quad g \in G, \ \Xi \in \mathcal{G}^*, \ X \in \mathcal{G}$$

The infinitesimal version of these actions provides an adjoint action of the Lie algebra $\mathcal{G}$ on $\mathcal{G}$ and $\mathcal{G}^*$ denoted ad and ad* given by:

$$\begin{aligned} \mathrm{ad}\, X(Y) &= [X,Y] \quad X,Y \in \mathcal{G}, \\ \mathrm{ad}^* X.\Xi(Y) &= -\Xi([X,Y]) \quad X,Y \in \mathcal{G}, \Xi \in \mathcal{G}^* \end{aligned}$$

To see how these notion relate to our problem, let us consider first a Lax matrix with only one polar singularity at $\lambda=0$:

$$L(\lambda) = \big(g(\lambda)\, A_-(\lambda)\, g^{-1}(\lambda)\big)_- \tag{2.20}$$

with $A_- = \sum_{r=-n}^{-1} A_r \lambda^r$, and $g(\lambda)$ has a regular expansion at $\lambda = 0$.

Let G be the formal loop group G of matrix valued formal power series expansion around $\lambda = 0$. The elements of G are regular series $g(\lambda) = \sum_{r=0}^{\infty} g_r \lambda^r$. The product law is the pointwise product: $(gh)(\lambda) = g(\lambda)h(\lambda)$. Formally, the Lie algebra $\mathcal{G}$ of G consists of elements of the form $X = \sum_{r=0}^{\infty} X_r \lambda^r$. Its Lie bracket is given by the pointwise commutator.

The dual $\mathcal{G}^*$ of $\mathcal{G}$ can be identified with the set of polar matrices $\Xi(\lambda) = \sum_{r\geq 1} \Xi_r \lambda^{-r}$, where the sum contains a finite but arbitrary large number of terms, by the pairing:

$$< \Xi, X >= \mathrm{Tr}\ \mathrm{Res}_{\lambda=0} \left(\Xi(\lambda)X(\lambda)\right) = \sum_r \mathrm{Tr}\left(\Xi_{r+1}X_r\right)$$

The coadjoint action of G on $\mathcal{G}^*$ is defined by $((\mathrm{Ad}^* g)\cdot \Xi)\ (X) = \Xi(g^{-1}\cdot X\cdot g)$ for $\Xi \in \mathcal{G}^*$ and any $X \in \mathcal{G}$. Using the above model for $\mathcal{G}^*$, and since $< \Xi, g^{-1}Xg >=< g\Xi g^{-1}, X >=< (g\Xi g^{-1})_-, X >$, we get

$$(\mathrm{Ad}^* g)\cdot \Xi(\lambda) = \left(g\cdot \Xi \cdot g^{-1}\right)_-$$

Generic coadjoint orbits consist of elements of the form $\Xi(\lambda) = \sum_{r=1}^{n} \Xi_r \lambda^{-r}$ with a fixed maximal order n of the pole. This is precisely eq.(2.20). This Lax matrix can thus be interpreted as the coadjoint orbit of the element A_- of $\mathcal{G}^*$ under the loop group G. In the generic case we are considering, the matrix $L(\lambda)$ fills the whole orbit.

Coadjoint orbits in $\mathcal{G}^*$ are equipped with the canonical Kostant-Kirillov symplectic structure. Before defining it we need some simple facts concerning the functions on $\mathcal{G}^*$. Denote the space of such functions by $\mathcal{F}(\mathcal{G}^*)$. The coadjoint action of G on $\mathcal{G}^*$ induces a coadjoint action of G on functions on $\mathcal{G}^*$, also denoted Ad^*. If $g \in G$, and $h \in \mathcal{F}(\mathcal{G}^*)$

$$\mathrm{Ad}^*\, g\, .\, h(\Xi) = h(\mathrm{Ad}^*\, g^{-1}(\Xi))$$

and a similar formula for the infinitesimal action ad^*. For a function h on $\mathcal{G}^*$ the differential dh may be viewed as an element of $\mathcal{G}$. This is because $\mathcal{G}^*$ is a vector space, the differential dh is a linear form on $\mathcal{G}^*$, i.e. an element of $\mathcal{G}^{**} \sim \mathcal{G}$, and one can write for $\delta\Xi \in \mathcal{G}^*$:

$$h(\Xi + \delta\Xi) = h(\Xi) + \delta\Xi(dh) + \mathcal{O}\left((\delta\Xi)^2\right)$$

From the Lie algebra structure on $\mathcal{G}$, we may construct a Poisson bracket on the space $\mathcal{F}(\mathcal{G}^*)$ of functions on $\mathcal{G}^*$. This is the Kostant–Kirillov bracket.

Definition. *If h_1 and $h_2 \in \mathcal{F}(\mathcal{G}^*)$ then the Kostant–Kirillov bracket is defined by:*

$$\{h_1, h_2\}(\Xi) = \Xi([dh_1, dh_2])$$

It is obvious that the bracket $\{\ ,\ \}$ is antisymmetric and verifies the Jacobi identity. Choosing the two linear functions $h_1(\Xi) = \Xi(X)$ and $h_2(\Xi) = \Xi(Y)$ with $X,\ Y\ \in \mathcal{G}$, so that $dh_1 = X$ and $dh_2 = Y$, the Kostant–Kirillov Poisson bracket reads:

$$\{\Xi(X), \Xi(Y)\} = \Xi([X, Y])$$

where the right–hand side is the linear function $\Xi \to \Xi([X, Y])$. This Poisson bracket is very natural but one has to be aware that it is a degenerate Poisson bracket.

Proposition. *The kernel of the Kostant–Kirillov bracket $\{\ ,\ \}$ is the set $\mathcal{I}(\mathcal{G}^*)$ of Ad^*–invariant functions.*

Proof. Let us first express the Ad^*–invariance property of a function h on $\mathcal{G}^*$. Performing an infinitesimal transformation we have: $h(\Xi + t\ \mathrm{ad}^*\, X\cdot \Xi) = h(\Xi) + t\ (\mathrm{ad}^*\, X\cdot \Xi)(dh) + \mathcal{O}(t^2)$ so h is invariant if $(\mathrm{ad}^*\, X\cdot$

$\Xi)(dh) = \Xi([dh, X]) = 0$ for all $\Xi \in \mathcal{G}^*$ and all $X \in \mathcal{G}$. Assuming now that $k \in \mathcal{F}(\mathcal{G}^*)$ is in the kernel of $\{\ ,\ \}$, we have $\{k, f\}(\Xi) = \Xi([dk, df]) = 0, \quad \forall f \in \mathcal{F}(\mathcal{G}^*)$. Thus k is ad*–invariant. The converse is obvious. ∎

This proposition means that the kernel of the Kostant–Kirrilov bracket $\{\ ,\ \}$ is the set of functions which are constant on the orbits of the coadjoint action of G. This implies that the restriction of this bracket to functions on a coadjoint orbit is non–degenerate, hence defines a symplectic structure.

In our situation, it is therefore natural to equip $L(\lambda)$ with the Kostant–Kirillov bracket associated to the loop group. This has the consequence that $A_-(\lambda)$ should be considered as being non dynamical matrices.

Proposition. *The diagonal matrix $A_-(\lambda)$, viewed as a function of $L(\lambda)$, is* Ad**–invariant. It characterizes the orbit and is non dynamical. The dynamical variables are the group elements $g(\lambda)$ modulo right multiplication by regular diagonal matrices.*

Proof. We have to check that A_- is invariant under the coadjoint action. Indeed

$$\mathrm{Ad}^*\, g \cdot L = (gLg^{-1})_- = gLg^{-1} - (gLg^{-1})_+$$

hence $A_-(\lambda)$, which is the singular part of the eigenvalues of $L(\lambda)$, is invariant under coadjoint action. ∎

The symplectic 2–form ω_K of the Kostant–Kirillov symplectic structure on coadjoint orbits is of course closed, but not exact. However, the coadjoint action defines a map φ from the group G to the orbit $\mathcal{O}_\Xi$ by $\varphi:\ g \to Ad^*g \cdot \Xi$. The pullback $\varphi^*\omega$ of ω on G is exact and one can write $\varphi^*\omega = d\alpha$ with $\alpha = \Xi(dgg^{-1})$. We will often meet this construction. Let us specialize it to our case. As above, we first identify $\mathcal{G}^*$ with series expansion singular at $\lambda = 0$ using the linear form induced by the trace of the residue at $\lambda = 0$. We then consider the orbit of the element $A_-(\lambda)$ and use the group element $g(\lambda)$ to parametrize a point in this orbit. The canonical one-form α is given by

$$\alpha = \mathrm{Tr}\ \mathrm{Res}_{\lambda=0} \left(A_- g^{-1}\delta g\right)$$

and the symplectic structure reads:

$$\omega = \delta\alpha = \mathrm{Tr}\ \mathrm{Res}_{\lambda=0}\Big(A_- g^{-1}\delta g \wedge g^{-1}\delta g\Big) \tag{2.21}$$

One way to check this is to remark that coadjoint action on $\mathcal{G}^*$ is inherited from left translation on the group, whose infinitesimal vector fields are $V_X = Xg$. The Hamiltonians generating these vector fields are $H_X = \alpha(V_X) = \Xi(X)$, because $dH_X = \Xi([X, dgg^{-1}]) = \omega(V_X, \cdot)$. Hence $\{\Xi(X), \Xi(Y)\} = \omega(V_X, V_Y) = \Xi([X, Y])$ i.e. we reproduce the Kostant–Kirillov bracket (see Chapter [14]).

In the multi pole case, we consider the direct sum of loop algebras $\mathcal{G}_k$, around $\lambda = \lambda_k$. An element of this Lie algebra has the form of a multiplet

$$X(\lambda) = (X_1(\lambda), X_2(\lambda), \cdots)$$

where $X_k(\lambda)$ admits an expansion of the form $X_k(\lambda) = \sum_{n\geq 0} X_{k,n}(\lambda - \lambda_k)^n$, and the Lie bracket is such that $[X_k(\lambda), X_l(\lambda)] = 0$ if $k \neq l$. The dual of this Lie algebra consists of multiplets

$$\Xi = (\Xi_1(\lambda), \Xi_2(\lambda), \cdots)$$

where $\Xi_k(\lambda)$ admits around λ_k an expansion of the form $\Xi_k(\lambda) = \sum_{r\geq 1} \Xi_{k,r}(\lambda - \lambda_k)^{-r}$. The pairing is simply

$$< \Xi, X >= \sum_k < \Xi_k, X_k >= \sum_k \mathrm{Tr}\ \mathrm{Res}_{\lambda_k}(\Xi_k(\lambda) X_k(\lambda))$$

The coadjoint action of G on $\mathcal{G}^*$ is given by the usual formula: if $g = (g_1, g_2, \cdots) \in G$ and $\Xi = (\Xi_1, \Xi_2, \cdots) \in \mathcal{G}^*$

$$(Ad^* g).\Xi(\lambda) = ((g_1 \Xi_1 g_1^{-1})_-, (g_2 \Xi_2 g_2^{-1})_-, \cdots)$$

A coadjoint orbit consists of elements Ξ_k with a fixed maximal order of the pole. Then, under the identification eq.(2.22), we can interpret eq.(2.11) as the coadjoint orbit of the element $((A_1)_-, (A_2)_-, \cdots)$.

Alternatively, we can consider the function on $\mathcal{G}^*$

$$L(\lambda) = L_0 + \sum_k \Xi_k \tag{2.22}$$

with poles at the points λ_k. Given this function we can recover the Ξ_k by extracting the polar parts. The constant matrix L_0 is added to match the formula for the Lax matrix eq.(2.11). By choice it is assumed to be invariant by coadjoint action. The pairing can be rewritten as

$$< L, X >= \sum_k \text{Tr Res}_{\lambda_k} L(\lambda) X_k(\lambda)$$

Only Ξ_k contributes to the residue at λ_k and the formula is compatible with the matrix L_0 being invariant by coadjoint action.

It is interesting to compute the dimension of this coadjoint orbit. In the formula $L_k = \left(g^{(k)}\, A_-^{(k)}\, g^{(k)-1}\right)_-$, the matrices $A_-^{(k)}$ characterize the orbit and are non–dynamical. The dynamical variables are the jets of order $(n_k - 1)$ of the $g^{(k)}$'s which gives $N^2 n_k$ parameters. But L_k is invariant under $g^{(k)} \to g^{(k)} d^{(k)}$ with $d^{(k)}$ a jet of diagonal matrices of the same order. Hence the dimension of the L_k orbit is $(N^2 - N) n_k$, and the dimension of the orbit is:

$$\dim \mathcal{M} = (N^2 - N) \sum_k n_k$$

In the multi pole case, the pullback of the symplectic form reads:

$$\omega = \sum_k \text{Tr Res}_{\lambda_k} \left(A_-^{(k)} g^{(k)-1} \delta g^{(k)} \wedge g^{(k)-1} \delta g^{(k)} \right) \tag{2.23}$$

Let us now use this symplectic form (2.23) to evaluate the Poisson brackets of the elements of the Lax matrix. To write them we use the matrix notation of previous Chapter [1] and show that they take the r–matrix form.

Proposition. *Under the genericity assumption of this section, the Poisson brackets of the matrix elements of $L(\lambda)$ can be written as:*

$$\{L_1(\lambda), L_2(\mu)\} = -\left[\frac{C_{12}}{\lambda - \mu}, L_1(\lambda) + L_2(\mu) \right] \tag{2.24}$$

with $C_{12} = \sum_{i,j} e_{ij} \otimes e_{ji}$ where the e_{ij} are the canonical basis matrices. The commutator in the right hand side of eq.(2.24) is the usual matrix commutator.

Proof. Let us first assume that we have only one pole and $L = (g A_- g^{-1})_-$. Because we are dealing with a Kostant-Kirillov bracket for the loop algebra of $gl(N)$, we can immediately write the Poisson bracket of the Lax matrix using the defining relation $\{L(X), L(Y)\} = L([X, Y])$. Since by definition, $L(X) = \text{Tr Res}_{\lambda=0}(L(\lambda) X(\lambda))$, this gives:

$$\{\text{Tr Res}_{\lambda=0} \left(L(\lambda) X(\lambda)\right)\ ,\ \text{Tr Res}_{\mu=0} \left(L(\mu) Y(\mu)\right)\} = \text{Tr Res}_{\lambda=0} \left(L(\lambda)[X(\lambda), Y(\lambda)]\right) \tag{2.25}$$

By definition of the notation $\{L_1, L_2\}$, we have:

$$\{\text{Tr Res}_{\lambda=0} \left(L(\lambda) X(\lambda)\right)\ ,\ \text{Tr Res}_{\mu=0} \left(L(\mu) Y(\mu)\right)\} =< \{L_1(\lambda), L_2(\mu)\}\, , \, X(\lambda) \otimes Y(\mu) >$$

where $<,>= \mathrm{Tr}_{12}\mathrm{Res}_\lambda \mathrm{Res}_\mu$. We need factorize $X(\lambda)\otimes Y(\mu)$ in eq.(2.25). To this end, we introduce a Casimir operator $\mathcal{C}_{12} = \sum_\alpha e_\alpha \otimes e^*_\alpha \in \mathcal{G}\otimes\mathcal{G}^*$ where e_α and e^*_α are two dual basis of $\mathcal{G}$ and $\mathcal{G}^*$ respectively. It is such that for $Y \in \mathcal{G}$, we have $Y_1 = \mathcal{C}_{12}(Y_2)$. Then we want to write $< L, [X,Y] >=< [\mathcal{C}_{12}, L_1], X\otimes Y >$, however this formula does not make sense as it stands because $[\mathcal{C}_{12}, L_1] = \sum_\alpha [e_\alpha, L]\otimes e^*_\alpha$ and $e_\alpha \in \mathcal{G}$ while $L \in \mathcal{G}^*$ and the commutator is not defined. To overcome this problem we embed $\mathcal{G}$ and its dual $\mathcal{G}^*$ into the full loop algebra $\widetilde{\mathcal{G}}$ generated by e^n_α, $n \in \mathbb{Z}$.

$$\widetilde{\mathcal{G}} = \mathcal{G} + \mathcal{G}^* \tag{2.26}$$

Note that in this sum, $\mathcal{G}$ and $\mathcal{G}^*$ do not commute. Using the evaluation representation of this full loop algebra the generators are represented by $\rho_\lambda(e^n_\alpha) = e_{ij}\lambda^n, \quad n \geq 0$ and $\rho_\lambda({e^*}^n_\alpha) = e_{ji}\lambda^{-n-1}, \quad n \geq 0$, so that under the pairing Tr Res we have $< {e^*}^n_{ij}, e^m_{kl} >= \delta_{ik}\delta_{jl}\delta_{nm}$. Let us compute the Casimir operator, assuming $|\lambda| < |\mu|$:

$$\mathcal{C}_{12}(\lambda,\mu) \equiv \rho_\lambda \otimes \rho_\mu \, \mathcal{C}_{12} = C_{12}\sum_{n=0}^{\infty} \frac{\lambda^n}{\mu^{n+1}} = -\frac{C_{12}}{\lambda-\mu}, \quad C_{12} = \sum_{i,j} e_{ij}\otimes e_{ji}$$

We can now write: $< L(\lambda)[X(\lambda), Y(\lambda)] >=< [\mathcal{C}_{12}(\lambda,\mu), L(\lambda)\otimes 1]\,, X(\lambda)\otimes Y(\mu) >$. Consider the rational function of λ: $\varphi(\lambda) = \{L_1(\lambda), L_2(\mu)\} - [\mathcal{C}_{12}(\lambda,\mu), L(\lambda)\otimes 1]$. By inspection φ contains only negative powers of μ, and we have $< \varphi, X(\lambda)\otimes Y(\mu) >= 0$. Hence φ contains only *positive* powers of λ and is regular at $\lambda = 0$. It has a pole at $\lambda = \mu$, due to the form of $\mathcal{C}(\lambda,\mu)$. We remove this pole by subtracting to φ the quantity $[\mathcal{C}_{12}(\lambda,\mu), 1\otimes L(\mu)]$ which contains only positive powers of λ and is therefore in the kernel of $< \cdot, X(\lambda)\otimes Y(\mu) >$. The pole at $\lambda = \mu$ disappears since $[C_{12}, L(\mu)\otimes 1 + 1\otimes L(\mu)] = 0$. The redefined φ is regular everywhere and vanishes for $\lambda \to \infty$ hence vanishes identically.

We can now study the multi pole situation occuring in eq.(2.11). Consider $L = L_0 + \sum_{k=1} L_k$. Each L_k lives in a coadjoint orbit as above equipped with its own symplectic structure. From eq.(2.23) they have vanishing mutual Poisson brackets $\{L_{j1}, L_{k2}\} = 0$ for $j,k = 0,\cdots,N$ and $j \neq k$. We assume further that L_0 does not contain dynamical variables

$$\{L_{01}, L_{02}\} = 0, \quad \{L_{01}, L_{k2}\} = 0$$

(here the indices 1 and 2 refer to the tensorial notation). Then since $C_{12}/(\lambda-\mu)$ is *independent* of the pole λ_k, it is obvious that the r-matrix relations at each orbit combine by addition to give a similar eq.(2.24) for the complete Lax matrix $L(\lambda)$. ■

This proposition shows that the generic Zakharov-Shabat system, equipped with this symplectic structure, is an *integrable Hamiltonian system*. It also gives us a very simple formula for the r-matrix specifying the Poisson bracket of $L(\lambda)$:

$$r_{12}(\lambda,\mu) = -r_{21}(\mu,\lambda) = -\frac{C_{12}}{(\lambda-\mu)}$$

The Jacobi identity is satisfied because this r-matrix verifies the Yang-Baxter equation (see eq.(1.8) in Chapter [1])

$$[r_{12}(\lambda_1,\lambda_2), r_{13}(\lambda_1,\lambda_3)] + [r_{12}(\lambda_1,\lambda_2), r_{23}(\lambda_2,\lambda_3)] + [r_{13}(\lambda_1,\lambda_3), r_{23}(\lambda_1,\lambda_3)] = 0$$

As in Chapter [1], these Poisson brackets for the Lax matrix ensure that one can define commuting quantities. The associated equations of motion can be written in the Lax form.

Proposition. *The quantities*

$$H^{(n)}(\lambda) = \mathrm{Tr}\,\Big(L^n(\lambda)\Big)$$

are in involution. The equations of motion associated to $H^{(n)}(\mu)$ can be written in the Lax form with $M = \sum_k M_k$:

$$M_k = n\left(\frac{L^{n-1}(\lambda)}{\lambda-\mu}\right)_k$$

Proof. The quantities $H^{(n)}(\lambda)$ are in involution because

$$\{\mathrm{Tr}\, L^n(\lambda), \mathrm{Tr}\, L^n(\mu)\} = \frac{nm}{\lambda-\mu}\mathrm{Tr}_{12}\left([C_{12}, L_1^n(\lambda)]L_2^{m-1}(\mu) + [C_{12}, L_2^m(\mu)]L_1^{n-1}(\lambda)\right) = 0$$

Similarly, we have:

$$\dot{L}(\lambda) = \{H^{(n)}(\mu), L(\lambda)\} = n\mathrm{Tr}_2\left[\frac{C_{12}}{\lambda-\mu}L_2^{n-1}(\mu), L_1(\lambda)\right]$$

The equations of motion generated by the Hamiltonian $H^{(n)}(\mu)$ take the Lax form:

$$\{H^{(n)}(\mu), L(\lambda)\} = [L(\lambda), M^{(n)}(\lambda,\mu)], \quad M^{(n)}(\lambda,\mu) = -n\frac{L^{n-1}(\mu)}{\lambda-\mu} \tag{2.27}$$

This $M^{(n)}(\lambda,\mu)$ has a pole at $\lambda = \mu$ and is otherwise regular. According to the general procedure we can remove this pole by subtracting some polynomial in $L(\lambda)$ without changing the equations of motion. Obviously one can redefine:

$$M^{(n)}(\lambda,\mu) \to M^{(n)}(\lambda,\mu) + n\frac{L^{n-1}(\lambda)}{\lambda-\mu} = n\frac{L^{n-1}(\lambda) - L^{n-1}(\mu)}{\lambda-\mu}$$

This new M has poles at all λ_k and is regular at $\lambda = \mu$. Decomposing into its polar parts, we write $M = \sum_k M_k$ with

$$M_k = n\left(\frac{L^{n-1}(\lambda)}{\lambda-\mu}\right)_k$$

This is of the form eq.(2.16) with $P^{(k)}(L,\lambda) = nL^{n-1}(\lambda)/(\lambda-\mu)$. ∎

Example 1. In the case of the Euler top $L(\lambda) = \mathcal{I}^2 + \frac{1}{\lambda}J$ we have already seen that the 3×3 matrix $J = 1/2\sum_{i,j} J_{ij}T_{ij}$, with $T_{ij} = e_{ij} - e_{ji}$, describes a coadjoint orbit of the orthogonal group $SO(3)$ characterized by the invariant $\mathrm{Tr}(J^2)$. The symplectic structure on this orbit is inherited from the usual Kostant–Kirillov bracket: $\{J_{ij}, J_{kl}\} = -1/2\mathrm{Tr}(J[T_{ij}, T_{kl}])$. Using the tensor notation we get

$$\{J_1, J_2\} = \frac{1}{2}\sum_{k,l,i} J_{kl}(T_{ki}\otimes T_{il} - T_{li}\otimes T_{ik})$$

Introducing the Casimir operator C_{12} of $so(3)$:

$$C_{12} = \frac{1}{2}\sum_{ij} T_{ij}\otimes T_{ji}$$

we can rewrite this Poisson bracket as $\{J_1, J_2\} = [C_{12}, J_1]$. This immediately leads to:

$$\{L_1(\lambda), L_2(\mu)\} = -\left[\frac{C_{12}}{(\lambda-\mu)}, L_1(\lambda) + L_2(\mu)\right]$$

since $L_0 = \mathcal{I}^2$ is a constant matrix.

Example 2. We consider next the Neumann model. Here we cannot avoid using a loop group. The Lax matrix $L(\lambda) = D + \frac{1}{\lambda}J - \frac{1}{\lambda^2}K$ has a singular part satisfying ${}^tL_-(-\lambda) = L_-(\lambda)$. This is *not* preserved by a general coadjoint action $L_-(\lambda) = g^{-1}(\lambda)L_-(\lambda)g(\lambda)$. To overcome this problem we consider the subgroup of matrices satisfying ${}^tg^{-1}(-\lambda) = g(\lambda)$ which may be called graded orthogonal. Its Lie algebra consists in matrices $X(\lambda)$ such that ${}^tX(-\lambda) = -X(\lambda)$. Its dual under the pairing Tr Res, consists in matrices $L(\lambda)$ such that ${}^tL(-\lambda) = L(\lambda)$, and having an expansion in a finite sum of strictly negative powers of λ. The matrix $L_-(\lambda)$ is an orbit under the coadjoint action. As a matter of fact let us compute:

$$\left(g^{-1}(\lambda)L_-(\lambda)g(\lambda)\right)_- = -\frac{1}{\lambda^2}\,{}^tg_0Kg_0 + \frac{1}{\lambda}\left({}^tg_0Jg_0 + {}^tg_1Kg_0 - {}^tg_0Kg_1\right)$$

with $g(\lambda) = g_0 + \lambda g_1 + \cdots$. Recalling that $K = X\,{}^tX$ and $J = X\,{}^tY - Y\,{}^tX$, we see that this is exactly of the same form as $L_-(\lambda)$ with

$$X \to {}^tg_0X, \quad Y \to {}^tg_0Y - {}^tg_1X$$

Let us compute the r–matrix of the Neumann model. This is similar to the previous computation except that the loop algebra being different, the Casimir operator has to be recomputed. A basis e_α of graded loop algebra is given by:

$$(e_{ij} - e_{ji})\lambda^{2n},\ i<j, \quad (e_{ij} + e_{ji})\lambda^{2n+1},\ i<j, \quad e_{ii}\lambda^{2n+1}, \quad n = 0, 1, \cdots$$

The dual basis e^*_α under Tr Res is given respectively by:

$$-\frac{1}{2}(e_{ij} - e_{ji})\lambda^{-2n-1},\ i<j, \quad \frac{1}{2}(e_{ij} + e_{ji})\lambda^{-2n-2},\ i<j, \quad e_{ii}\lambda^{-2n-2}, \quad n = 0, 1, \cdots$$

Then one gets for $\mathcal{C}_{12} = \sum_\alpha e_\alpha \otimes e^*_\alpha$:

$$\mathcal{C}_{12}(\lambda,\mu) = -\frac{1}{2}\frac{1}{\lambda-\mu}\sum_{ij} e_{ij}\otimes e_{ji} - \frac{1}{2}\frac{1}{\lambda+\mu}\sum_{ij} e_{ij}\otimes e_{ij}$$

This implies that poles at $\lambda = \pm\mu$ appear in $\varphi(\lambda) = \{L_1(\lambda), L_2(\mu)\} - [\mathcal{C}_{12}(\lambda,\mu), L(\lambda)\otimes 1]$. To cancel these poles we now subtract $[\mathcal{C}_{21}(\mu,\lambda), 1\otimes L(\mu)]$ which only contains positive powers of λ. Indeed the residue at $\lambda = \mu$ is $[C_{12}, L_1(\mu) + L_2(\mu)]$ and therefore vanishes as previously, while the residue at $\lambda = -\mu$ reads $[D_{12}, L_1(-\mu) - L_2(\mu)]$ with $D_{12} = \sum_{ij} e_{ij}\otimes e_{ij}$. Now $L(-\mu) = {}^tL(\mu)$ and one checks that for any matrix A the commutator $[D_{12}, A_1 - {}^tA_2]$ vanishes. Hence $\varphi(\lambda) - [\mathcal{C}_{21}(\mu,\lambda), 1\otimes L(\mu)] = 0$. Finally one gets:

$$\{L_1(\lambda)\,,\,L_2(\mu)\} = [r_{12}(\lambda,\mu), L_1(\lambda)] - [r_{21}(\mu,\lambda), L_2(\mu)], \quad r_{12}(\lambda,\mu) = \mathcal{C}_{12}(\lambda,\mu)$$

Note that this is a two–poles r–matrix with poles at $\lambda = \pm\mu$.

2.4 Elementary flows.

We have found that the time evolution of a Zakharov–Shabat system is given by matrices M of the form eq.(2.12). The general M can be reconstructed from elementary ones, that we shall call M_i, corresponding to diagonal matrices $B^{(k)}_-$ having a single pole of order n at λ_k in matrix entry α:

$$M_i = \left(g^{(k)}\xi_{(k,n,\alpha)}g^{(k)-1}\right)_k, \quad \xi_i \equiv \xi_{(k,n,\alpha)} = \frac{1}{(\lambda-\lambda_k)^n}E_{\alpha\alpha} \tag{2.28}$$

Here and in the following we use the multi index $i = (k, n, \alpha)$.

We call $t_i = t_{(k,n,\alpha)}$ the time variable associated to M_i through the Lax equation

$$\partial_{t_i}L = [L, M_i] \tag{2.29}$$

A general flow is a linear combination of these elementary ones. Remark that $M_i(\lambda)$, which is a priori defined around λ_k, is a rational fonction of λ, with only a polar part at λ_k, hence is defined in the whole λ–plane. The Lax equation eq.(2.29) has a meaning in the whole λ plane and defines the time evolution of the quantities locally defined at $\lambda_{k'}$ (such as $g_{k'}$) with respect to the times associated to λ_k.

We now verify directly that the flows ∂_{t_i} defined by eqs.(2.29) and (2.28) all commute. This amounts to showing that $[L, \partial_{t_i}M_j - \partial_{t_j}M_i + [M_i, M_j]] = 0$. We get even the stronger result:

Proposition. *The matrices M_i defining the time evolution t_i satisfy the zero curvature condition*

$$\partial_{t_i} M_j - \partial_{t_j} M_i + [M_i, M_j] = 0$$

As a consequence, the flows defined by eq.(2.29, 2.28) are all commuting.

Proof. Let $i = (k, n, \alpha)$ and $j = (k', n', \alpha')$. Diagonalizing $L = g^{(k')} A^{(k')} g^{(k')-1}$ around $\lambda_{k'}$ the Lax equation gives

$$\partial_{t_i} g^{(k')} = -M_i g^{(k')} + g^{(k')} d_i^{(k')} \tag{2.30}$$

where $d_i^{(k')}$ is an unknown diagonal matrix. This equation has to be interpreted as an expansion around $\lambda_{k'}$. If $i = (k, n, \alpha)$ and $k' = k$ this implies that $d_i^{(k)} = \xi_i + \text{regular}$, because $\partial_{t_i} g^{(k)}$ is regular, while if $k' \neq k$, we only conclude that $d_i^{(k')}$ is regular around k'. Note that $g^{(k)}$ is known only up to a right multiplication by a regular diagonal matrix $g^{(k)} \to g^{(k)} d^{(k)}$ and this changes $d_i^{(k')} \to d_i^{(k')} - d^{(k')-1} \partial_{t_i} d^{(k')}$.

From eqs.(2.28) and (2.30), we get

$$\partial_j M_i = -[M_j, g^{(k)} \xi_i g^{(k)-1}]_k + (g^{(k)} [d_j^{(k)}, \xi_i] g^{(k)-1})_k$$

Since the commutator in the second term involves diagonal matrices, it vanishes. Hence we have

$$\partial_j M_i = -[M_j, g^{(k)} \xi_i g^{(k)-1}]_k$$

Let us assume first that $\lambda_k \neq \lambda_{k'}$. Then M_j is regular at λ_k and only $(g^{(k)} \xi_i g^{(k)-1})_k = M_i$ contributes to the polar part of the above commutator yielding

$$\partial_j M_i = -[M_j, M_i]_k$$

Similarly we have

$$\partial_{t_i} M_j = -[M_i, M_j]_{k'}$$

The zero curvature condition follows because $[M_i, M_j]$ is a rational function with poles only at λ_k and $\lambda_{k'}$ and vanishing at infinity, so that

$$[M_i, M_j] = [M_i, M_j]_k + [M_i, M_j]_{k'}$$

Assume next that $\lambda_k = \lambda_{k'}$. We still have

$$\partial_j M_i = -[M_j, g^{(k)} \xi_i g^{(k)-1}]_k, \quad \partial_{t_i} M_j = -[M_i, g^{(k)} \xi_j g^{(k)-1}]_k$$

where now all projections are at λ_k. But $M_i - g^{(k)} \xi_i g^{(k)-1} = O(1)$ and $M_j - g^{(k)} \xi_j g^{(k)-1} = O(1)$ so that $[M_i - g^{(k)} \xi_i g^{(k)-1}, M_j - g^{(k)} \xi_j g^{(k)-1}]_k = 0$, or

$$[M_i, M_j] - [g^{(k)} \xi_i g^{(k)-1}, M_j]_k - [M_i, g^{(k)} \xi_j g^{(k)-1}]_k + [g^{(k)} \xi_i g^{(k)-1}, g^{(k)} \xi_j g^{(k)-1}]_k = 0$$

but the last term vanishes because $[\xi_i, \xi_j] = 0$. From this the zero curvature condition readily follows. ∎

When parametrizing L as in eq.(2.11), the dynamical variables are the $g^{(k)}(\lambda)$, modulo gauge transformations consisting of right multiplication by diagonal matrices. Let us write the equations of motion for these variables:

Proposition. *There exists a gauge choice such that the equations of motion read*

$$\partial_{t_i} g^{(k')} = -M_i g^{(k')} + g^{(k')} \partial_{t_i} \xi^{(k')}(\lambda, t), \quad \xi^{(k)}(\lambda, t) \equiv \sum_i \xi_i t_i, \quad i = (k, n, \alpha) \tag{2.31}$$

where the last sum runs over the indices i above the pole λ_k. The zero curvature condition also allows to introduce the "wave function" $\Psi(\lambda; t_1, t_2, \cdots)$, a matrix function depending on all the times simultaneously, and satisfying

$$\partial_{t_i}\Psi = -M_i\Psi \tag{2.32}$$

The choice of gauge above means that we can write locally around λ_k

$$\Psi = g^{(k)}(\lambda)e^{-\xi^{(k)}(\lambda,t)} \tag{2.33}$$

Proof. We have to return to eq.(2.30) and specify the matrix $d_i^{(k')}$. The zero curvature equation implies that $\partial_{t_i}d_j^{(k')} - \partial_j d_i^{(k')} + [d_i^{(k')}, d_j^{(k')}] = 0$ for any indices (i, j, k'). Since the matrices $d_i^{(k)}$ are diagonal, the commutator vanishes, and the condition implies that $d_i^{(k')} = \partial_{t_i}h^{(k')}$ for some diagonal matrix $h^{(k')}$. We can now use the freedom $g^{(k')} \to g^{(k')}d^{(k')}$ to suppress the regular part of $h^{(k')}$ around $\lambda_{k'}$. This is a choice of gauge. Since we already noticed that the singular part of $d_i^{(k')}$ around $\lambda_{k'}$ is fixed, $d_i^{(k')} = 0$ if $k \neq k'$ and $d_i^{(k')} = \xi_i$ if $k = k'$, this determines $d_i^{(k')}$ completely: $d_i^{(k)} = \partial_{t_i}\xi^{(k')}(\lambda, t)$. ■

2.5 Factorization problem.

We now show, in the case of one pole, that solving the hierarchy amounts to solving a factorization problem in a loop group.

The equations of the hierarchy are

$$\partial_{t_i}g = \left(g\xi_i g^{-1}\right)_+ g = -\left(g\xi_i g^{-1}\right)_- g + g\xi_i \tag{2.34}$$

DEFINIR LE PROBLEME DE FACTORISATION. REFERER AUX EXAMPLES

Proposition. *The solution of the linear system eq.(2.34) is obtained by solving the factorization problem*

$$g_-^{-1}(\lambda, t)g_+(\lambda, t) = e^{-\xi(\lambda,t)}g(\lambda, 0)e^{\xi(\lambda,t)} \tag{2.35}$$

Proof. Taking the time derivative of eq.(2.35) with respect to t_j, and multiplying on the left by g_- and on the right by g_+^{-1}, we get:

$$-\dot{g}_-g_-^{-1} + \dot{g}_+g_+^{-1} = -g_-\xi_j g_-^{-1} + g_+\xi_j g_+^{-1}$$

Identifying the $+$ and $-$ parts, we have:

$$\dot{g}_+g_+^{-1} = \left(g_+\xi_j g_+^{-1}\right)_+, \quad \dot{g}_-g_-^{-1} = g_-\xi_j g_-^{-1} - \left(g_+\xi_j g_+^{-1}\right)_-$$

The first equation is just eq.(2.34). ■

This is a remarkable result, as it shows that the solution of the integrable system is obtained by solving an algebraic factorization problem. This property can be promoted to an abstract algebraic setting for integrable systems, see Chapter [3].

Remark. The factorization problem eq.(2.35) can be stated in a slightly different but equivalent way. Since $\xi(\lambda)$ contains only a polar part, it can be reabsorbed in $g_-(t)$. Let us define $g_\pm(0)$ by factorizing $g(0) = g_-^{-1}(0)g_+(0)$, and $\widetilde{g}_-^{-1}(t)$ and $\widetilde{g}_+(t)$ by

$$\widetilde{g}_-^{-1}(t) = g_-(0)e^{\xi(t)}g_-^{-1}(t), \quad \widetilde{g}_+(t) = g_+(t)g_+^{-1}(0)$$

We can write

$$\widetilde{g}_-^{-1}(t)\widetilde{g}_+(t) = g_+(0)e^{-\xi(t)}g_+^{-1}(0)$$

Since $g_+(0)$ is the matrix which diagonalises the Lax matrix $L(0) = g_+(0)Ag_+^{-1}(0)$ and ξ_i is a polynomial in A, say $\xi_i = P_i(A)$, we can formulate the factorization problem in the form

$$\widetilde{g}_-^{-1}(t)\,\widetilde{g}_+(t) = e^{-\sum_i P_i(L)(0)\,t_i}$$

It is in this form that the factorization problem will reapear in Chapter [3].

2.6 Tau functions.

We now introduce the important concept of τ–functions. The elementary flows induce a coupling between the λ dependence and the corresponding time dependence, which is best expressed in terms of τ–functions.

Proposition. *Let $d = \sum_i dt_i \partial_{t_i}$. In the gauge of eq.(2.31), the one form*

$$\gamma = \sum_k \text{Tr Res}_{\lambda_k}\,(g^{(k)-1}\partial_\lambda g^{(k)} d\xi^{(k)})d\lambda \tag{2.36}$$

is closed $d\gamma = 0$. As a consequence we can define a function $\tau(t_1, t_2, \cdots)$ of the infinite series of times $t_{(k,n,\alpha)}$ for all k, n, α such that

$$\partial_{t_i} \log\tau = \sum_k \text{Tr Res}_{\lambda_k}\,(g^{(k)-1}\partial_\lambda g^{(k)}\xi_i) \tag{2.37}$$

Proof. The evolution equations eq.(2.31) can be written as $dg^{(k)} = g^{(k)}d\xi^{(k)} - Mg^{(k)}$ where we have set $M = \sum_i M_i dt_i$. Hence by differentiating eq.(2.36), we have

$$d\gamma \;=\; \sum_k \text{Tr Res}_{\lambda=\lambda_k}(d\partial_\lambda \xi^{(k)} \wedge d\xi^{(k)} - \partial_\lambda M \wedge g^{(k)}d\xi^{(k)}g^{(k)-1})d\lambda$$

The first terms vanish, Tr $\text{Res}_{\lambda=\lambda_k}(d\partial_\lambda\xi^{(k)} \wedge d\xi^{(k)}) = 0$, since the order of the pole is at least 3. Next $M = M^{(k)} + \sum_{l\neq k} M^{(l)}$, with $M^{(k)} = (g^{(k)}d\xi^{(k)}g^{(k)-1})_-$. So we can write $g^{(k)}d\xi^{(k)}g^{(k)-1} = M + X^{(k)}_{reg}$ where $X^{(k)}_{reg}$ is regular at λ_k, and $d\gamma = -\sum_k \text{Tr Res}_{\lambda=\lambda_k}(\partial_\lambda M \wedge (M + X^{(k)}_{reg})d\lambda$. On the other hand

$$\begin{aligned}\text{Tr Res}_{\lambda=\lambda_k}(\partial_\lambda(M + X^{(k)}_{reg}) \wedge (M + X^{(k)}_{reg}))d\lambda &= \text{Tr Res}_{\lambda=\lambda_k}(\partial_\lambda(g^{(k)}d\xi^{(k)}g^{(k)-1}) \wedge (g^{(k)}d\xi^{(k)}g^{(k)-1}))d\lambda \\ &= \text{Tr Res}_{\lambda=\lambda_k}(d\partial_\lambda\xi^{(k)} \wedge d\xi^{(k)})d\lambda = 0\end{aligned}$$

Since the residue of a derivative of a function of λ vanishes, Tr $\text{Res}_{\lambda=\lambda_k}(\partial_\lambda X^{(k)}_{reg} \wedge M)d\lambda = \text{Tr Res}_{\lambda=\lambda_k}(\partial_\lambda M \wedge X^{(k)}_{reg})d\lambda$ we get:

$$\text{Tr Res}_{\lambda=\lambda_k}(\partial_\lambda M \wedge X^{(k)}_{reg})d\lambda = -\frac{1}{2}\text{Tr Res}_{\lambda=\lambda_k}(\partial_\lambda M \wedge M)d\lambda$$

It follows that

$$d\gamma = -\sum_k \text{Tr Res}_{\lambda=\lambda_k}(\partial_\lambda M \wedge (M + X^{(k)}_{reg})d\lambda = -\frac{1}{2}\sum_k \text{Tr Res}_{\lambda=\lambda_k}(\partial_\lambda M \wedge M)d\lambda$$

Now $\text{Tr}(\partial_\lambda M \wedge M)d\lambda$ is a rational one–form on the λ Riemann sphere, hence the sum of its residues vanishes.
■

The τ–function has remarkable and beautiful implications in the theory of integrable systems. Here we shall only present its fundamental relation to the wave–function Ψ. Around a singularity λ_k, we can expand Ψ as $\Psi = g^{(k)}(\lambda)e^{-\xi^{(k)}(\lambda,t)}$. The matrix $g^{(k)}(\lambda)$ is regular at $\lambda = \lambda_k$ and can be written as $g^{(k)}(\lambda) = g_0^{(k)}\widetilde{g}^{(k)}(\lambda)$

with $\widetilde{g}^{(k)}(\lambda) = 1 + O(\lambda - \lambda_k)$. We will relate the matrix elements $\widetilde{g}^{(k)}_{\beta\alpha}(\lambda)$ to τ–functions. Let us consider the vicinity of a given pole which we assume to be at the origin $\lambda_k = 0$ (the case $\lambda_k \neq 0$ is simply recovered by translating λ in the formulae below). Introduce the differential operator $D^{(k)}_\alpha$

$$D^{(k)}_\alpha = \sum_{n>0} (\lambda - \lambda_k)^n \frac{\partial}{\partial t_{(k,n,\alpha)}} = \sum_{n>0} \lambda^n \frac{\partial}{\partial t_{(k,n,\alpha)}} \tag{2.38}$$

Proposition. *The action of $D^{(k)}_\alpha$ on the τ–function is given by:*

$$D^{(k)}_\alpha \log \tau = \lambda \mathrm{Tr}(E_{\alpha\alpha} \widetilde{g}^{(k)-1} \partial_\lambda \widetilde{g}^{(k)}) \tag{2.39}$$

Its action on $\widetilde{g}^{(k)}$ is given by:

$$\widetilde{g}^{(k)-1} D^{(k)}_\alpha \widetilde{g}^{(k)} = \lambda [\widetilde{g}^{(k)-1} \partial_\lambda \widetilde{g}^{(k)}, E_{\alpha\alpha}] + \widetilde{g}^{(k)-1} E_{\alpha\alpha} \widetilde{g}^{(k)} - E_{\alpha\alpha} \tag{2.40}$$

Proof. We first need simple results on formal series. Let $f(\lambda) = \sum_{j=0}^\infty f_j \lambda^j$. We have

$$\sum_{n=1}^\infty \lambda^n \left(\lambda^{-n} f(\lambda)\right)_+ = \lambda \partial_\lambda f(\lambda), \quad \sum_{n=1}^\infty \lambda^n \mathrm{Res}(\lambda^{-n} f(\lambda)) = \lambda f(\lambda) \tag{2.41}$$

Indeed,

$$\sum_{n=1}^\infty \lambda^n \left(\lambda^{-n} f(\lambda)\right)_+ = \sum_{n=1}^\infty \sum_{j=n}^\infty f_j \lambda^j = \sum_{j=1}^\infty \sum_{n=0}^{j} f_j \lambda^j = \sum_{j=1}^\infty j f_j \lambda^j = \lambda \partial_\lambda f(\lambda)$$

This shows the first equality. The second one is simpler: $\sum_{n=1}^\infty \lambda^n \mathrm{Res}(\lambda^{-n} f(\lambda)) = \sum_{n=1}^\infty \lambda^n f_{n-1} = \lambda f(\lambda)$.

In the definition of $\log \tau$, eq.(2.37), we can replace $g^{(k)}$ by $\widetilde{g}^{(k)}$ because $g_0^{(k)}$ is independent of λ. Then we get

$$\frac{\partial}{\partial t_{(k,n,\alpha)}} \log \tau = \mathrm{Tr}\ \mathrm{Res}_{\lambda=0} \, (\lambda^{-n} E_{\alpha\alpha} \widetilde{g}^{(k)-1} \partial_\lambda \widetilde{g}^{(k)})$$

Hence using eq.(2.41)

$$D^{(k)}_\alpha \log \tau = \sum_{n=1}^\infty \lambda^n \mathrm{Tr}\,\mathrm{Res}(\lambda^{-n} E_{\alpha\alpha} \widetilde{g}^{(k)-1} \partial_\lambda \widetilde{g}^{(k)}) = \lambda \mathrm{Tr}(E_{\alpha\alpha} \widetilde{g}^{(k)-1} \partial_\lambda \widetilde{g}^{(k)})$$

To prove the second formula, we start from the equations of the hierarchy:

$$\partial_{t_i} g^{(k)} = \left(g^{(k)} E_{\alpha\alpha} \lambda^{-n} g^{(k)-1} \right)_+ g^{(k)}, \quad i = (k, n, \alpha)$$

which decomposes into

$$\partial_{t_i} g_0^{(k)} = g_0^{(k)} \left(\widetilde{g}^{(k)} E_{\alpha\alpha} \lambda^{-n} \widetilde{g}^{(k)-1} \right)_0, \quad \partial_{t_i} \widetilde{g}^{(k)} = \left(\left(\widetilde{g}^{(k)} E_{\alpha\alpha} \lambda^{-n} \widetilde{g}^{(k)-1} \right)_+ - \left(\widetilde{g}^{(k)} E_{\alpha\alpha} \lambda^{-n} \widetilde{g}^{(k)-1} \right)_0 \right) \widetilde{g}^{(k)}$$

Using eq.(2.41) again and $\sum_{n=1}^\infty \lambda^n (\lambda^{-n} f(\lambda))_0 = f(\lambda) - f_0$, we get

$$\begin{aligned} D^{(k)}_\alpha \widetilde{g}^{(k)} &= \sum_{n=1}^\infty \lambda^n \left(\left(\widetilde{g}^{(k)} E_{\alpha\alpha} \lambda^{-n} \widetilde{g}^{(k)-1} \right)_+ - \left(\widetilde{g}^{(k)} E_{\alpha\alpha} \lambda^{-n} \widetilde{g}^{(k)-1} \right)_0 \right) \widetilde{g}^{(k)} \\ &= \left(\lambda \partial_\lambda (\widetilde{g}^{(k)} E_{\alpha\alpha} \widetilde{g}^{(k)-1}) - \widetilde{g}^{(k)} E_{\alpha\alpha} \widetilde{g}^{(k)-1} + E_{\alpha\alpha} \right) \widetilde{g}^{(k)} \end{aligned}$$

This is equivalent to eq.(2.40). ∎

Proposition. *There exists τ–functions $\tau_{\alpha\beta}^{(k)}(t)$ such that the matrix function $\widetilde{g}^{(k)}(t,\lambda)$ can be expressed in the form:*

$$\widetilde{g}_{\beta\alpha}^{(k)}(t,\lambda) = \delta_{\alpha\beta}\frac{\tau(t+[\lambda]_\alpha^{(k)})}{\tau(t)} + (1-\delta_{\alpha\beta})(\lambda-\lambda_k)\frac{\tau_{\beta\alpha}^{(k)}(t+[\lambda]_\alpha^{(k)})}{\tau(t)} \tag{2.42}$$

where the notation $t+[\lambda]_\alpha^{(k)}$ means the shift of the time $t_{(kn\alpha)}$

$$t+[\lambda]_\alpha^{(k)} \equiv t_{(k'n\gamma)} + \delta_{kk'}\delta_{\gamma\alpha}\frac{(\lambda-\lambda_k)^n}{n}$$

Proof. Let us take the matrix element β,α of eq.(2.40). If $\beta\neq\alpha$ we get $(\widetilde{g}^{(k)-1}D_\alpha^{(k)}\widetilde{g}^{(k)})_{\beta\alpha} = \lambda(\widetilde{g}^{(k)-1}\partial_\lambda\widetilde{g}^{(k)})_{\beta\alpha} + (\widetilde{g}^{(k)-1}E_{\alpha\alpha}\widetilde{g}^{(k)})_{\beta\alpha}$, while if $\beta=\alpha$, we get $(\widetilde{g}^{(k)-1}D_\alpha^{(k)}\widetilde{g}^{(k)})_{\alpha\alpha} = (\widetilde{g}^{(k)-1}E_{\alpha\alpha}\widetilde{g}^{(k)})_{\alpha\alpha} - 1$. Hence we can write

$$(\widetilde{g}^{(k)-1}D_\alpha^{(k)}\widetilde{g}^{(k)})_{\beta\alpha} = \lambda(\widetilde{g}^{(k)-1}\partial_\lambda\widetilde{g}^{(k)})_{\beta\alpha} + (\widetilde{g}^{(k)-1}E_{\alpha\alpha}\widetilde{g}^{(k)} - E_{\alpha\alpha})_{\beta\alpha} - \delta_{\beta\alpha}\lambda(\widetilde{g}^{(k)-1}\partial_\lambda\widetilde{g}^{(k)})_{\alpha\alpha}$$

Multiplying on the left by $\widetilde{g}^{(k)})$ and remembering that by eq.(2.39), $\lambda(\widetilde{g}^{(k)-1}\partial_\lambda\widetilde{g}^{(k)})_{\alpha\alpha} = D_\alpha^{(k)}\log\tau$, we obtain

$$D_\alpha^{(k)}\widetilde{g}_{\beta\alpha}^{(k)} = \lambda\partial_\lambda\widetilde{g}_{\beta\alpha}^{(k)} + (E_{\alpha\alpha}\widetilde{g}^{(k)})_{\beta\alpha} - (\widetilde{g}^{(k)}E_{\alpha\alpha})_{\beta\alpha} - D_\alpha^{(k)}\log\tau\,\widetilde{g}_{\beta\alpha}^{(k)}$$

If $\beta=\alpha$ we set $X_{\alpha\alpha}^{(k)} = \tau\widetilde{g}_{\alpha\alpha}^{(k)}$ and for $\beta\neq\alpha$, we set $X_{\beta\alpha}^{(k)} = \lambda^{-1}\tau\widetilde{g}_{\beta\alpha}^{(k)}$ The equation becomes $(D_\alpha^{(k)} - \lambda\partial_\lambda)X_{\beta\alpha}^{(k)} = 0$ and the equation becomes $(D_\alpha^{(k)} - \lambda\partial_\lambda)X_{\beta\alpha}^{(k)} = 0$. This means that $X_{\beta\alpha}^{(k)}(t,\lambda)$ is of the form $X_{\beta\alpha}^{(k)}(t,\lambda) = X_{\beta\alpha}^{(k)}(\cdots,t_{(k,n,\alpha)}+\lambda^n/n,\cdots)$. But $\widetilde{g}_{\alpha\alpha}^{(k)}(\lambda=0)=1$ so that the function $X_{\alpha\alpha}^{(k)}(t)$ is in fact equal to $\tau(t)$. ■

Eq.(2.42) is a famous formula discovered by Sato. Its main feature is that the λ dependence in the numerator is expressed by translation of the appropriate time variables. This is particularly stricking for the diagonal elements $\widetilde{g}_{\alpha\alpha}^{(k)}(t,\lambda)$

$$\widetilde{g}_{\alpha\alpha}^{(k)}(t,\lambda) = \frac{\tau(t+[\lambda]_\alpha^{(k)})}{\tau(t)}$$

More information on τ–functions, including relations between $\tau_{\beta\alpha}^{(k)}(t)$ and $\tau(t)$, will be given in Chapter [7].

2.7 Integrable field theories and monodromy matrix.

For a system with finite number of degrees of freedom, we have seen that a Lax matrix could be interpreted as a coadjoint orbit. It is possible to adapt this interpretation to field theory by properly choosing the Lie algebra involved. We shall consider two–dimensional field theory on a cylinder with space variable $x\in[0,2\pi]$ and time variable $t\in[-\infty,+\infty]$. To introduce the space variable x, we consider the loop algebra $\widetilde{\mathcal{G}}$ of (periodic) maps from the circle S^1 to the some Lie algebra $\mathcal{G}$, i.e. maps $x\in S^1\to\mathcal{G}$. The simplest case corresponds to choosing $\mathcal{G}$ to be the algebra of $N\times N$ matrices, but more frequently, it will be an element of a loop algebra with spectral parameter λ as in the finite dimensional case. So we are dealing with double loop algebras. In order to introduce some non triviality in the x direction, we consider the central extension of the x–loop algebra:

$$\widehat{\mathcal{G}} = \widetilde{\mathcal{G}} + \mathbb{C}e$$

The commutator of two elements $X_i = \widetilde{X}_i(x) + c_i e$ reads

$$[X_1,X_2] = [\widetilde{X}_1(x),\widetilde{X}_2(x)] + \int_0^{2\pi}(\partial_x\widetilde{X}_1(x)\widetilde{X}_2(x))dx\ e$$

where $(\ ,\ )$ is an invariant non degenerate bilinear form on $\mathcal{G}$ (such as Tr or Tr Res). The dual space $\widehat{\mathcal{G}}^*$ of $\widehat{\mathcal{G}}$ can be identified with the space of pairs of elements of the form $\Xi = (\widetilde{\Xi}(x), \xi)$ with pairing

$$\Xi(X) = \int_0^{2\pi} (\widetilde{\Xi}(x), \widetilde{X}(x))dx + \xi c$$

The coadjoint action is defined as usual $(\mathrm{ad}^* X \cdot \Xi)(Y) = -\Xi([X, Y])$ which takes the form

$$\begin{aligned}(\mathrm{ad}^* X \cdot \Xi)(Y) &= -\int_0^{2\pi} (\widetilde{\Xi}(x), [\widetilde{X}(x), \widetilde{Y}(x)])dx - \xi \int_0^{2\pi} (\partial_x \widetilde{X}(x) \widetilde{Y}(x))dx \\ &= \int_0^{2\pi} (-[\widetilde{\Xi}(x), \widetilde{X}(x)] - \partial_x \widetilde{X}(x), \widetilde{Y}(x))dx\end{aligned}$$

so that

$$\mathrm{ad}^* X \cdot \Xi = (-[\widetilde{\Xi}(x), \widetilde{X}(x)] - \partial_x \widetilde{X}(x), 0) \tag{2.43}$$

We see that the element ξ is invariant by coadjoint action, and we will choose orbits with $\xi = 1$. With this, the Lax equation $\partial_t L = -\mathrm{ad}^* M \cdot L$ reads, for $L = (U, 1)$ and $M = V$,:

$$\partial_t U - \partial_x V + [V, U] = 0 \tag{2.44}$$

This is nothing else than a zero curvature condition.

Alternatively, one can say that the variable x behaves like one of the times of finite dimensional systems. It is important however to realise that in the field theory case, the construction of commuting quantities is more complicated because we have to construct functions invariant under the coadjoint action eq.(2.43). For this, the right object to consider is the so called monodromy matrix which we now introduce.

The zero curvature (2.44) condition expresses the compatibility condition of the associated linear system

$$(\partial_x + U)\,\Psi = 0, \quad (\partial_t\ + V)\,\Psi = 0 \tag{2.45}$$

The matrices U and V can be thought of as the x and t components of a connection. This connection will be called the Lax connection. Given U and V, the linear system (2.45) determines the matrix Ψ up to multiplication on the right by a constant matrix. $\Psi \equiv \Psi(x, t)$ will be called the wave function.

The transport matrix is the matrix $T_\gamma(\lambda; x, t)$ representing the parallel transport of Ψ from the origin to the point (x, t) specified by the Lax connection, i.e. we solve the linear system along the path γ. We denote the transport matrix by:

$$T_\gamma(x, t) = \overleftarrow{\exp} \left[-\int_\gamma (U dx + V dt) \right] \tag{2.46}$$

Here γ is a path joining the origin to the space-time point (x, t). Since the Lax connection satisfies the zero curvature relation (2.44) the value of the transport matrix is independent of the choice of this path. In particular if γ is the path $x \in [0, 2\pi]$, with fixed time t we call the corresponding transport matrix the monodromy matrix.

$$T(\lambda, t) = \overleftarrow{\exp} \left[-\int_0^{2\pi} U(\lambda, x, t)dx \right] \tag{2.47}$$

Proposition. *Assume that all fields are periodic in x with period 2π. Let $T(\lambda, t)$ be the monodromy matrix and let*

$$\Im(\lambda) = \mathrm{Tr}\,(T(\lambda, t)) \tag{2.48}$$

Then, $\Im(\lambda)$ is independent of time. Hence the trace of the monodromy matrix is a generating function of conserved quantities.

Proof. We have

$$\begin{aligned}
\partial_t T(\lambda,t) &= -\int_0^{2\pi} dx e^{-\int_x^{2\pi} Udx}\dot{U}(\lambda,x)e^{-\int_0^x Udx} = -\int_0^{2\pi} dx e^{-\int_x^{2\pi} Udx}(\partial_x V + [U,V])e^{-\int_0^x Udx} \\
&= -\int_0^{2\pi} dx \partial_x \left(e^{-\int_x^{2\pi} Udx} V e^{-\int_0^x Udx} \right)
\end{aligned}$$

Performing the integral, we get

$$\partial_t T(\lambda,t) = -V(\lambda,2\pi,t)T(\lambda,t) + T(\lambda,t)V(\lambda,0,t) \tag{2.49}$$

So, if the fields are periodic, we have $V(\lambda,2\pi,t) = V(\lambda,0,t)$ and the relation becomes

$$\partial_t T(\lambda,t) = [T(\lambda,t), V(\lambda,0,t)]$$

This is a Lax equation. Taking the trace it implies that $\Im(\lambda)$ is time independent. Expanding in λ we obtain an infinite set of conserved quantities. ∎

2.8 Abelianization.

Thus we consider the linear system eq.(2.45) where $U(x,t,\lambda)$ and $V(x,t,\lambda)$ are matrices depending in a rational way in a parameter λ having poles at *constant* values λ_k.

$$U(\lambda) = U_0 + \sum_k U_k \qquad \text{with} \qquad U_k = \sum_{r=-n_k}^{-1} U_{k,r}(\lambda-\lambda_k)^r \tag{2.50}$$

$$V(\lambda) = V_0 + \sum_k V_k \qquad \text{with} \qquad V_k = \sum_{r=-m_k}^{-1} V_{k,r}(\lambda-\lambda_k)^r \tag{2.51}$$

The compatibility condition of the linear system (2.45) is the zero curvature condition (2.44). We demand that it holds identically in λ. These conditions are always compatible, since by the same naive counting argument as for finite dimensional systems there is one more variable as equations. The origin of this underdeterminacy is the same: the zero curvature condition is invariant by gauge transformations. If the gauge transformation is independent of λ, it will not spoil the analytic properties of U and V. Notice that eq.(2.44) implies that U_0 and V_0 are pure gauge, i.e. there exists a group valued function h such that

$$U_0 = h^{-1}\partial_x h \quad \text{and} \quad V_0 = h^{-1}\partial_t h \tag{2.52}$$

Using a λ independent gauge transformation *periodic* in x, we can always come to a gauge in which U_0 is constant diagonal and $V_0 = 0$. To show this we start from eq.(2.52). Writing that $U_0(x)$ is periodic implies that $\partial_x(h(x+2\pi)h^{-1}(x)) = 0$. We change basis so that $h(x+2\pi)h^{-1}(x)$ is diagonal and write it $\exp(2\pi P)$ with P diagonal. Hence we can write $h(x,t) = \exp(Px)\widetilde{h}(x,t)$ with $\widetilde{h}(x+2\pi,t) = \widetilde{h}(x,t)$ and we gauge transform under $\widetilde{h}$. Then we have $\widetilde{U}_0 = P$ and $\widetilde{V}_0 = x\partial_t P$. But $\widetilde{V}_0$ is periodic in x hence $\partial_t P = 0$. Once this gauge has been chosen, the number of equations is equal to the number of unknowns.

As for finite dimensional systems we first make a local analysis around each pole λ_k in order to understand solutions of eq.(2.44). We show that around each singularity λ_k, one can perform a gauge transformation bringing simultaneously U and V to a diagonal form.

Let us assume that the pole is located at $\lambda = 0$. The rational functions $U(\lambda)$, $V(\lambda)$ can be expanded in a Taylor series in a neighborhood of this pole

$$U(\lambda) = \sum_{r=-n}^{\infty} U_r(\lambda)^r, \quad V(\lambda) = \sum_{r=-m}^{\infty} V_r(\lambda)^r \tag{2.53}$$

We have

Proposition. *There exists a local, periodic, gauge transformation*

$$\partial_x + U = g(\partial_x + A)g^{-1}, \quad \partial_t + V = g(\partial_t + B)g^{-1} \tag{2.54}$$

where g, A and B are formal series in λ

$$g = \sum_{r=0}^{\infty} g_r \lambda^r, \quad A = \sum_{r=-n}^{\infty} A_r \lambda^r, \quad B = \sum_{r=-m}^{\infty} B_r \lambda^r$$

such that the matrices A and B are diagonal. Moreover $\partial_t A - \partial_x B = 0$.

Proof. Let g_0 be the matrix diagonalizing the leading term in eq.(2.53)

$$U_{-n} = g_0 \; A_{-n} \; g_0^{-1}$$

Let $\partial_x + U = g_0(\partial_x + \widetilde{U})g_0^{-1}$, $\widetilde{U} = \sum_{r=-n}^{\infty} \widetilde{U}_r \lambda^r$. Since $\lambda = 0$ is a pole of U and g_0 is regular we have $\widetilde{U}_{-n} = A_{-n}$. We set $g = g_0 \widetilde{g}$. The equation becomes $\partial_x + \widetilde{U} = \widetilde{g}(\partial_x + A)\widetilde{g}^{-1}$, or $\partial_x \widetilde{g} + \widetilde{U}\widetilde{g} - \widetilde{g}A = 0$. Expanding in powers of λ, we get

$$\partial_x \widetilde{g}_l + \sum_{r=-n}^{l} \left(\widetilde{U}_r \; \widetilde{g}_{l-r} - \widetilde{g}_{l-r} \; A_r \right) = 0, \quad l = -n, \cdots \infty \tag{2.55}$$

Of course $\widetilde{g}_l = 0$ if $l < 0$; $\widetilde{g}_0 = 1$. In the sum, we separate the first and the last term

$$\partial_x \widetilde{g}_l + [A_{-n}, \widetilde{g}_{l+n}] + \sum_{r=-n+1}^{l-1} \left(\widetilde{U}_r \widetilde{g}_{l-r} - \widetilde{g}_{l-r} A_r \right) + \widetilde{U}_l - A_l = 0, \quad l = -n, \cdots \infty \tag{2.56}$$

Projecting this equation on the diagonal matrices, we determine A_l in terms of A_k, $\; k < l$, and $\widetilde{g}_k$, $\; k < l+n$. (The term $[A_{-n}, \widetilde{g}_{l+n}]$ does not contribute to the diagonal since A_{-n} is itself diagonal). Similarly, the off diagonal part of this equation determines the off diagonal part of $\widetilde{g}_{l+n}$ in terms of the same variables. We can make the solution unique by requiring, for instance, diag $(\widetilde{g}_{l+n}) = 0$. Therefore $\widetilde{g}$ and A are determined recursively. Remark that this is a purely algebraic computation, so g and A are periodic functions of x and are algebraic functions of the matrix elements of U and their derivatives.

Under the gauge transformation g we obviously have $\partial_t A - \partial_x B = [A, B]$, or $\partial_t A_l - \partial_x B_l = \sum_{r=-n}^{\infty} [A_r, B_{l-r}]$. If B is regular at $\lambda = 0$ the expansion of B starts at B_0 and $\partial_t A_{-n} = [A_{-n}, B_0]$. Hence the off-diagonal part of the commutator is zero and therefore B_0 is diagonal. If however B is singular the expansion of B starts at B_{-m} and the most singular term in the commutator is $\lambda^{-n-m}[A_{-n}, B_{-m}]$, and this has to vanish because $n + m > \max(n, m)$. Hence B_{-m} is diagonal. We finish the proof by induction on l. Assume B_r is diagonal until B_{l+n-1}. Then $\partial_t A_l - \partial_x B_l = \sum_{r=-n}^{l+m} [A_r, B_{l-r}] = [A_{-n}, B_{l+n}]$, hence B_{l+n} is diagonal. It is important to notice that this procedure only requires local computations. There is no differential equation to be integrated when diagonalizing recursively the Lax connection around its poles. ∎

Of course if V is singular at $\lambda = 0$ but U is regular a similar argument starting from V allows to simultaneously diagonalize the two connections around $\lambda = 0$.

As for finite dimensional systems, we can reconstruct all the matrices U_k and V_k and therefore the Lax pair from simple data.

$$\begin{aligned} U &= U_0 + \sum_k U_k, \quad \text{with} \quad U_k \left(g^{(k)} A_-^{(k)} g^{(k)-1}\right)_- \\ V &= V_0 + \sum V_k, \quad \text{with} \quad V_k = \left(g^{(k)} B_-^{(k)} g^{(k)-1}\right)_- \end{aligned}$$

Remark. Let us check that the reconstruction formulae (2.57) for U and V are such that the order of the poles in the commutator appearing in the zero curvature condition matches the order of the poles in the derivative terms. Indeed the polar part at λ_k of the commutator is $[V, U_k]_- + [U, V_k]_-$ and can be higher than the order of the poles of the derivatives $\partial_t U_k$ and $\partial_x V_k$. So we have to show that the formula ensure that the poles of these commutators, which are naively of order $n_k + m_k$, are actually of order less than $\max(n_k, m_k)$. Indeed consider for exemple the commutator $[V, U_k]_-$. Similarly as for finite dimensional systems, we may write it as:

$$\begin{aligned} [V, U_k]_- &= \left[V, \left(g^{(k)} A_-^{(k)} g^{(k)-1}\right)_-\right]_- = \left[V, g^{(k)} A_-^{(k)} g^{(k)-1} - \left(g^{(k)} A_-^{(k)} g^{(k)-1}\right)_+\right]_- \\ &= \left[g^{(k)} \partial_t g^{(k)-1}, g^{(k)} A_-^{(k)} g^{(k)-1}\right]_- - \left[V, \left(g^{(k)} A_-^{(k)} g^{(k)-1}\right)_+\right]_- \end{aligned}$$

In the last line we use the fact that $\partial_t + V$ is diagonalized by $g^{(k)}$, ie. $V = g^{(k)} \partial_t g^{(k)-1} + g^{(k)} B^{(k)} g_k^{-1}$ with $B^{(k)}$ diagonal. All the terms of the last line have a pole of order at most n_k. Similarly one shows that the order of the pole of $[U, V_k]_-$ is at most m_k. This shows that the constraints hidden is eq.(2.44) are solved by the formulas (2.57).

Proposition. *The quantities $Q^{(k)}(\lambda) = \int_0^{2\pi} A^{(k)}(\lambda, x, t)\, dx$ are local conserved quantities of the field theory. They are related to eq.(2.48) by*

$$\Im(\lambda) = \text{Tr}\ \exp\left[-\int_0^{2\pi} A^{(k)}(\lambda, x, t) dx\right]$$

<u>Proof</u>. Around each pole, in the diagonal gauge, the zero curvature condition reduces to

$$\partial_t A^{(k)}(\lambda, x, t) - \partial_x B^{(k)}(\lambda, x, t) = 0$$

It is the equation of conservation of a current. The charge $Q^{(k)}(\lambda) = \int_0^{2\pi} A^{(k)}(\lambda, x, t)\, dx$ is conserved because

$$\partial_t Q_k(\lambda) = \int_0^{2\pi} \partial_x B^{(k)}(\lambda, x, t) = B^{(k)}(\lambda, 2\pi, t) - B^{(k)}(\lambda, 0, t) = 0$$

where we have used the fact that $A^{(k)}(\lambda, x, t)$ and $B^{(k)}(\lambda, x, t)$ are local in terms the coefficients of the Lax connection are therefore periodic in x. Expanding the conserved quantity $Q^{(k)}(\lambda)$ in $\lambda - \lambda_k$ produces an infinite number of conserved quantities. Under a gauge transformation $U \to {}^gU = g^{-1} U g + g^{-1} \partial_x g$ the monodromy matrix is transformed into ${}^gT(\lambda, t)$ with

$${}^gT(\lambda, t) = g^{-1}(2\pi, t) T(\lambda, t) g(0, t)$$

Thus if $g(x, t)$ is periodic, i.e $g(2\pi, t) = g(0, t)$, one has

$$\text{Tr}\ ({}^gT(\lambda, t)) = \text{Tr}\ (T(\lambda, t))$$

Around a pole $\lambda = \lambda_k$ we have seen that one can diagonalize the Lax connection by a periodic gauge transformation. In this diagonal gauge the monodromy matrix is easily computed and one gets:

$$\Im(\lambda) = \text{Tr}\ \exp\left[-\int_0^{2\pi} A^{(k)}(\lambda, x, t) dx\right]$$

There is no problem of ordering in the exponential since the matrices $A^{(k)}$ are diagonal. ∎

Example. Let us give an exemple of this general construction. For simplicity, we look for a Lax connection in which U and V have only one simple pole at two different points and $U_0 = V_0 = 0$. Choosing these points to be at $\lambda = \pm 1$, we can thus parametrize U and V as:

$$U = -\frac{1}{\lambda - 1} J_x, \quad V = \frac{1}{\lambda + 1} J_t$$

with J_x and J_t taking values in some Lie algebra. Decomposing the zero curvature condition $[\partial_x + U, \partial_t + V] = 0$ over its simple poles gives two equations:

$$\begin{aligned} \partial_t J_x - \tfrac{1}{2}[J_x, J_t] &= 0, \\ \partial_x J_t + \tfrac{1}{2}[J_x, J_t] &= 0. \end{aligned}$$

Taking the difference implies that $[\partial_t + J_t, \partial_x + J_x] = 0$. Thus J is a pure gauge and there exists g such $J_t = g^{-1}\partial_t g$ and $J_x = g^{-1}\partial_x g$. Taking now the sum of the two equations implies $\partial_t J_x + \partial_x J_t = 0$, or equivalently,

$$\partial_t(g^{-1}\partial_x g) + \partial_x(g^{-1}\partial_t g) = 0$$

This is the field equation of the so-called non-linear sigma model, with x, t as light-cone coordinates.

2.9 Poisson brackets of the monodromy matrix.

As we just saw the zero curvature equation leads to the construction of infinite set of conserved currents. We want to compute the Poisson brackets of the conserved charges associated to these conserved currents. For this we will compute the Poisson brackets of the metrix elements of the monodromy matrix.

In order to do it we assume the existence of a r–matrix relation:

$$\{U_1(\lambda, x), U_2(\mu, y)\} = [r_{12}(\lambda - \mu), U_1(\lambda, x) + U_2(\mu, y)] \;\; \delta(x - y) \tag{2.57}$$

We say that this Poisson bracket is ultralocal due to the presence of $\delta(x - y)$. This hypothesis actually covers a large class of integrable field theories.

Since we are computing Poisson brackets let us fix the time t, and consider the transport matrix from x to y

$$T(\lambda; y, x) = \overleftarrow{\exp}\,(-\int_x^y U(\lambda, z)dz)$$

In particular the monodromy matrix is $T(\lambda) = T(\lambda; 2\pi, 0)$. The matrix elements $[T]_{ij}$ of $T(\lambda; y, x)$ are functions on phase space. As in Chapter [1], the table of their Poisson brackets may be arranged in the following matrix form. Let M and N be two matrices and $M_1 = M \otimes 1$, $N_2 = 1 \otimes N$, define $\{M_1, N_2\}_{ij,kl} = \{M \stackrel{\otimes}{,} N\}_{ij,kl} = \{M_{ik}, N_{jl}\}$.

Proposition. *We have the fundamental relation for the transport matrix:*

$$\Big\{T_1(\lambda; y, x), T_2(\mu; y, x)\Big\} = -\Big[r_{12}(\lambda, \mu), T_1(\lambda; y, x)T_2(\mu; y, x)\Big] \tag{2.58}$$

As a consequence, the trace of the monodromy matrix $\Im(\lambda) = \mathrm{Tr}\ (T(\lambda))$, is a generating function a Poisson commuting quantities:

$$\Big\{\Im(\lambda), \Im(\mu)\Big\} = 0 \tag{2.59}$$

Proof. Let us first prove the relation (2.58) for the Poisson brackets of the transport matrices. Notice that λ is attached to T_1 and μ to T_2, so that there is no ambiguity if we do not write explicitly the λ and μ dependence. The transport matrix $T(y,x)$ verifies the differential equations

$$\begin{aligned} \partial_x T(y,x) - T(y,x)U(x) &= 0 \\ \partial_y T(y,x) + U(y)T(y,x) &= 0 \end{aligned} \tag{2.60}$$

Since Poisson brackets satisfy the Leibnitz rules, we have

$$\begin{aligned} \{T_1(y,x), T_2(y,x)\} &= \int_x^y \int_x^y dudv\ T_1(y,u)T_2(y,v)\{U_1(u),U_2(v)\}T_1(u,x)T_2(v,x) \\ &= \int_x^y \int_x^y dudv\ \delta(u-v)T_1(y,u)T_2(y,v).[r_{12}, U_1(u)+U_2(v)]T_1(u,x)T_2(v,x) \end{aligned} \tag{2.61}$$

Using again the differential equation satisfied by $T(y,x)$ this yields:

$$\begin{aligned} \{T_1(y,x), T_2(y,x)\} &= -\int_x^y \int_x^y dudv\ \delta(u-v).[T_1(y,u)T_2(y,v)\ r_{12}\ (\partial_u+\partial_v)T_1(u,x)T_2(v,x) \\ &\quad +(\partial_u+\partial_v)(T_1(y,u)T_2(y,v))\ r_{12}\ T_1(u,x)T_2(v,x)] \\ &= -\int_x^y dz\partial_z(T_1(y,z)T_2(y,z).r_{12}.T_1(z,x)T_2(z,x)) \end{aligned}$$

Integrating this exact differential gives the relation (2.58). Let us now show that the trace of the monodromy matrix $\Im(\lambda)$ generates Poisson commuting quantities. Take the trace of (2.58). In the left hand side we use the fact that $\mathrm{Tr}\,(A\otimes B) = \mathrm{Tr}\,(A)\mathrm{Tr}\,(B)$. The right hand side gives zero since it is the trace of a commutator. ■

The previous proposition shows that we may take as Hamiltonian any element of the family generated by $\Im(\mu)$. We will show that the equations of motion take the form of a zero curvature condition.

Proposition. *Taking $\Im(\mu)$ as Hamiltonian, we have*

$$\dot{U}(\lambda;x) \equiv \{\Im(\mu), U(\lambda;x)\} = \partial_x V(\lambda,\mu;x) - [V(\lambda,\mu;x), U(\lambda;x)] \tag{2.62}$$

where $V(\lambda,\mu;x) = \mathrm{Tr}_1\Big(T_1(\mu;2\pi,x)r_{12}(\mu,\lambda)T_1(\mu;x,0)\Big)$.

Proof. To simplify the notation, we do not explicitly write the λ,μ dependence as above, noting that μ is attached to index 1 and λ to index 2. We have:

$$\begin{aligned} \{T_1(2\pi,0), U_2(x)\} &= -\int_0^{2\pi} dy\ T_1(2\pi,y)\ \{U_1(y),U_2(x)\}\ T_1(y,0) \\ &= -T_1(2\pi,x)\ [r_{12}, U_1(x)+U_2(x)]\ T_1(x,0) \end{aligned}$$

Expanding the commutator we get four terms

$$\begin{aligned} \{T_1(2\pi,0), U_2(x)\} = &-T_1(2\pi,x)\cdot r_{12}\cdot \underbrace{U_1(x)T_1(x,0)}_{use\ equa.\ dif.} -T_1(2\pi,x)\cdot r_{12}\cdot \underbrace{U_2(x)\ T_1(x,0)}_{commute} \\ &+ \underbrace{T_1(2\pi,x)\ U_1(x)}_{use\ equa.\ dif.}\cdot r_{12}\cdot T_1(x,0) + \underbrace{T_1(2\pi,x)\ U_2(x)}_{commute}\cdot r_{12}\cdot T_1(x,0) \end{aligned}$$

Using the differential equations (2.60) and commuting factors as indicated gives

$$\{T_1(2\pi,0),U_2(x)\} = \partial_x V_{12}(x) - [V_{12}(x),U_2(x)]$$

where we have introduced $V_{12}(x) = T_1(2\pi,x)\cdot r_{12}\cdot T_1(x,0)$. Taking the trace over the first space, remembering that $\Im(\mu) = \mathrm{Tr}\,T(\mu)$ and setting $V(\lambda,\mu,x) = \mathrm{Tr}_1 V_{12}(x)$ we find eq.(2.62). ∎

2.10 The group of dressing transformations.

We now introduce a very important notion, the group of dressing transformations, which is related to the Zakharov-Shabat construction. These transformations provide a way to construct new solutions of the equations of motion from old ones. It defines a group action on the space of classical solutions of the model, and therefore on the phase space of the model.

In one sentence: Dressing transformations are non-local gauge transformations preserving the analytical structure of the Lax connection.

These transformations are intimetely related to the Riemann-Hilbert problem. This problem consists in factorizing a matrix defined on a contour in the complex plane into the product of two matrices respectively analytic inside or outside the contour. Let us first recall a few well known fact about the Riemann-Hilbert problem.

Theorem. *Let Γ be a closed contour in the λ-plane and $G(\lambda)$ a matrix defined on Γ. Then, there exist two matrices $\Theta_\pm(\lambda)$, with $\Theta_+(\lambda)$ analytic inside Γ and $\Theta_-(\lambda)$ analytic outside Γ such that $\det\Theta_\pm \neq 0$ and:*

$$G(\lambda) = \Theta_-^{-1}(\lambda)\Theta_+(\lambda) \tag{2.63}$$

This solution to the factorization problem is unique up to the multiplication by a constant matrix g: $\Theta_\pm(\lambda) \longrightarrow g\Theta_\pm(\lambda)$.

We choose a contour Γ in the λ-plane such none of the poles λ_k of the Lax connection are on Γ. So some of these poles are outside Γ and some are inside.

To define the transformation, we pick a group valued function $G(\lambda)$ on Γ. From the previous theorem, $G(\lambda)$ can be factorized as:

$$G(\lambda) = G_-^{-1}(\lambda)G_+(\lambda)$$

where, according to the Riemann-Hilbert problem, $G_+(\lambda)$ $(G_-(\lambda))$ are analytic inside (ouside) the contour Γ.

Let U, V be a solution of the zero curvature equation eq.(2.44) with the prescribed singularities specified in eqs.(2.50,2.51). Let $\Psi \equiv \Psi(\lambda; x,t)$ be the solution of the linear system (2.45) normalized by $\Psi(\lambda;0) = 1$. Let $G(\lambda) = {G_-}^{-1}(\lambda)G_+(\lambda)$ be an element of the loop group on Γ factorized according to the Riemann-Hilbert problem. We set:

$$\Theta(\lambda;x,t) = \Psi(\lambda;x,t)\cdot G(\lambda)\cdot \Psi(\lambda;x,t)^{-1} \tag{2.64}$$

At each space-time point (x,t), we may decompose $\Theta(\lambda,x,t)$ according to the Riemann-Hilbert problem as:

$$\Theta(\lambda;x,t) = \Theta_-^{-1}(\lambda;x,t)\cdot\Theta_+(\lambda;x,t) \tag{2.65}$$

with Θ_+ (Θ_-) analytic inside (outside) the contour Γ. Then,

Proposition. *The following function defined for λ on the contour Γ,*

$$\Psi^G(\lambda;x,t) = \Theta_\pm(\lambda;x,t)\cdot\Psi(\lambda;x,t)\cdot G_\pm^{-1}(\lambda) \tag{2.66}$$

extends to two functions $\Psi_\pm^G$ respectively defined inside and outside Γ such that $\Psi_-^{G\,-1}\Psi_+^G|_\Gamma = 1$.
It is solution of the linear system (2.45) with Lax connection U^G and V^G given by

$$\begin{aligned} U^G(\lambda;x,t) &= \Theta_\pm\cdot U\cdot\Theta_\pm^{-1} - \partial_x\Theta_\pm\cdot\Theta_\pm^{-1} \quad (2.67)\\ V^G(\lambda;x,t) &= \Theta_\pm\cdot V\cdot\Theta_\pm^{-1} - \partial_t\Theta_\pm\cdot\Theta_\pm^{-1} \quad (2.68)\end{aligned}$$

The matrices U^G and V^G, which satisfy the zero curvature equation (2.44), are meromorphic functions on the whole complex λ plane with the same analytic structure than the components U and V of the original Lax connection.

<u>Proof</u>. First it follows directly from the definitions of $G_\pm$ and $\Theta_\pm$ that for λ on Γ,

$$\Theta_+(\lambda;x,t)\cdot\Psi(\lambda;x,t)\cdot G_+^{-1}(\lambda) = \Theta_-(\lambda;x,t)\cdot\Psi(\lambda;x,t)\cdot G_-^{-1}(\lambda)$$

In other words, the two expressions of the right hand side of eq.(2.66) are equal and effectively define a unique function Ψ^G. It is clear that this function can be extended into two functions $\Psi_\pm^G$ respectively defined inside and outside this contour by:

$$\Psi_\pm^G = \Theta_\pm\cdot\Psi\cdot G_\pm^{-1}$$

By construction, they are such that $\Psi_-^{G\,-1}\Psi_+^G|_\Gamma = 1$.
Similarly we may use $\Psi_\pm^G$ to define the extension $U_\pm^G$ ($V_\pm^G$) of U^G (V^G) inside and ouside the contour Γ. Explicitely:

$$\begin{aligned} U_\pm^G &= -\partial_x\Psi_\pm\cdot\Psi_\pm^{-1} = -\partial_x\Theta_\pm\Theta_\pm^{-1} + \Theta_\pm U\Theta_\pm^{-1}\\ V_\pm^G &= -\partial_t\Psi_\pm\cdot\Psi_\pm^{-1} = -\partial_t\Theta_\pm\Theta_\pm^{-1} + \Theta_\pm V\Theta_\pm^{-1}\end{aligned}$$

We see that $U_+^G,\ V_+^G$ have the same singularities than $U,\ V$ in the region inside the contour Γ, while $U_-^G,\ V_-^G$ have the same singularities than $U,\ V$ in the region ouside Γ. Furthermore, since $\Psi_-^{G\,-1}\Psi_+^G|_\Gamma = 1$ we know that U_+ coincides with U_- on the contour Γ and similarly $V_+ = V_-$ for $\lambda\in\Gamma$. Therefore $U_\pm, V_\pm$ are defined in the whole λ plane and have the same singularities as $U,\ V$.
It is also possible to show that $U_\pm, V_\pm$ are globally defined by a direct calculation. Differentiating eq.(2.65) with respect to x we find

$$\begin{aligned} -\Theta_-^{-1}\cdot\left(\partial_x\Theta_-\Theta_-^{-1}\right)\cdot\Theta_+ + \Theta_-^{-1}\partial_x\Theta_+ &= \partial_x\Theta\\ &= (\partial_x\Psi)\cdot G\cdot\Psi^{-1} - \Psi\cdot G\cdot\Psi^{-1}\partial_x\Psi\\ &= -U\cdot\Theta_-^{-1}\Theta_+ - \Theta_-^{-1}\Theta_+\cdot U\end{aligned}$$

We have used the linear system (2.45) and again eq.(2.65) which defines $\Theta_\pm$ by the relation $\Psi^{-1}G\Psi = \Theta_-^{-1}\Theta_+$. Multiplying the previous equation by Θ_+^{-1} from the right and by Θ_- from the left gives:

$$-\partial_x\Theta_-\cdot\Theta_-^{-1} + \Theta_- U\Theta_-^{-1} = -\partial_x\Theta_+\cdot\Theta_+^{-1} + \Theta_+ U\Theta_+^{-1} \equiv U^G$$

Therefore, U^G is defined in all the λ-plane and has the same singularities as U since $\Theta_\pm$ and $\Theta_\pm^{-1}$ are analytical in their respective domains of definition. Of course, similarly V^G is defined in all the λ-plane and has the same singularities as V. ■

This proposition effectively tells us that the dressing transformations (2.66) map solutions of the equations of motion into new solutions. Indeed given a solution of the zero curvature equation with the prescribed singularities and an element of the loop group we produce a new solution of the zero curvature equation with same analytical structure. But since this analytic structure is the only information which specifies the model we have thus produced a new solution of the equations of motion.

Thus, the dressing transformations $\Psi \to \Psi^G$ act on the solution space. They form a group, called the dressing group. This group is not isomorphic to the loop group since its composition law is different that the composition law in the loop group. Indeed,

Proposition. *Let $G = G_-^{-1}G_+$ and $H = H_-^{-1}H_+$ be two elements of the dressing group. The composition law of dressing transformation is given by*

$$H \bullet G = (H_-G_-)^{-1}(H_+G_+) \tag{2.69}$$

Representing the elements of the dressing group by the pairs, (G_-, G_+) and (H_-, H_+) we may write the composition law as: $(H_-, H_+) \bullet (G_-, G_+) = (H_-G_-, H_+G_+)$. In particular the plus and minus components commute.

Proof. Consider two elements $G = G_-^{-1}G_+$ and $H = H_-^{-1}H_+$ and transform successively $\Psi \to \Psi^G \to (\Psi^G)^H$; we have :

$$\begin{aligned} \Psi^G &= \Theta_\pm^G \, \Psi \, G_\pm^{-1} \qquad \text{with} \qquad \Theta_\pm^G = (\Psi G \Psi^{-1})_\pm \\ (\Psi^G)^H &= \Theta_\pm^{HG} \, \Psi^G \, H_\pm^{-1} \qquad \text{with} \qquad \Theta_\pm^{HG} = (\Psi^G H \Psi^{G\,-1})_\pm \end{aligned} \tag{2.70}$$

The factorization of $(\Psi^G H \Psi^{G\,-1})$ can be written as follows:

$$(\Theta_-^{HG})^{-1}\Theta_+^{HG} \equiv \Psi^G H \Psi^{G\,-1} = \Theta_-^G \, \Psi \, (H_-G_-)^{-1}(H_+G_+) \, \Psi^{-1} \, \Theta_+^{G\,-1}$$

or, equivalently,

$$\Theta_\pm^{HG}\Theta_\pm^G = (\Psi \, (H_-G_-)^{-1}(H_+G_+) \, \Psi^{-1})_\pm$$

Inserting this formula into eq.(2.70) proves the multiplication law for the dressing transformations. ∎

We end up these general considerations on dressing transformations by clarifying their relation to the Poisson structure of the theory. We shall assume the Poisson bracket eq.(2.57) for the Lax connection with

$$r_{12}(\lambda, \mu) = \frac{C_{12}}{\lambda - \mu}, \quad C_{12} = \sum_{ij} e_{ij} \otimes e_{ji}$$

The Poisson bracket of the wave function is thus

$$\Big\{\Psi_1(\lambda; x), \Psi_2(\mu; x)\Big\} = -\Big[r_{12}(\lambda, \mu), \Psi_1(\lambda; x)\Psi_2(\mu; x)\Big] \tag{2.71}$$

The r–matrix is related to the factorization problem in the loop algebra.

Proposition. *The action of dressing transformations is a Lie-Poisson action. The non-Abelian generator is the monodromy matrix.*

Proof. Recall that for any $g = g_-^{-1}g_+ \in G$, a dressing transformation acts on the wave function $\Psi(x,t)$ by

$$\Psi^g = \Theta_\pm \, \Psi \, g_\pm^{-1} \tag{2.72}$$

where the group elements $\Theta_\pm$ are solutions of the factorization problem:

$$\Theta_-^{-1}\,\Theta_+ \;=\; \Psi\, g\, \Psi^{-1} \tag{2.73}$$

Notice that the two signs give the same result in eq.(2.72). Also this transformation preserves the normalization condition $\Psi(0) = 1$.
The infinitesimal form of eq.(2.72) is for any $X \in \mathcal{G}$, with $X = X_+ - X_-$:

$$\delta_X\, \Psi \;=\; \Theta_\pm\, \Psi - \Psi\, X_\pm \qquad \text{with} \qquad \Theta_\pm = (\Psi X \Psi^{-1})_\pm \tag{2.74}$$

Introduce the monodromy matrix $T(L) = \Psi(L)\Psi^{-1}(0)$, with 0 and L the two boundary values of the coordinate x. Using the ultralocality property, its Poisson bracket with the wave function is:

$$\Big\{\; \Psi_1(x), T_2(L)\; \Big\} \;=\; T_2(L)\Psi_2^{-1}(x)[r^\pm_{12}\;,\; \Psi_1(x)\Psi_2(x)]$$

Therefore we have:

$$\begin{aligned}
\mathrm{Tr}_2\Big(X_2T_2^{-1}(L)\Big\{\Psi_1(x),T_2(L)\Big\}\Big) \;&=\; \big(\Psi(x)X\Psi^{-1}(x)\big)_\pm\,\Psi(x) - \Psi(x)X_\pm \\
&=\; \delta_X\,\Psi(x)
\end{aligned}$$

for any $X \in \mathcal{G}$. This proves that $T(L)$ is the non-Abelian Hamiltonian and shows that the action is Lie-Poisson. ■

EXPLIQUER EN QUOI LE PROBLEME DE FACTORISATION EST RELIE AUX MATRICES $r_\pm$.

2.11 Soliton solutions.

In general, a matrix Riemann–Hilbert problem like eq.(2.63) cannot be solved explicitely by analytical methods. This statement applies however to the fundamental solution of the Riemann Hilbert problem i.e. the one satisfying the conditions $\det\Theta_\pm \neq 0$. Nevertheless, if we relax this condition, new solutions can easily be constructed from the fundamental one, and this can be used to produce new solutions to the equations of motion. Starting from a trivial vacuum solution, we obtain in this way the so called solitons solutions of integrable systems.

Definition. *We say that a matrix function* $\Theta(\lambda)$ *has a zero at the point* λ_0 *if* $\det\Theta(\lambda_0) = 0$ *and if in the vicinity of this point*

$$\Theta(\lambda) = F_0 + (\lambda-\lambda_0)F_1 + O(\lambda-\lambda_0)^2, \quad \Theta^{-1}(\lambda) = \frac{1}{\lambda-\lambda_0}C_0 + C_1 + O(\lambda-\lambda_0)$$

Since $\Theta(\lambda)\Theta^{-1}(\lambda) = \Theta^{-1}(\lambda)\Theta(\lambda) = Id$, we have $F_0C_0 = C_0F_0 = 0$, $C_0F_1 + C_1F_0 = Id$. In particular

$$\mathrm{Ker}\,F_0 = \mathrm{Im}\,C_0, \quad \mathrm{Ker}\,C_0 = \mathrm{Im}\,F_0$$

Let now Γ be a closed contour in the λ–plane, let $G(\lambda)$ be a matrix defined on Γ, and consider the Riemann–Hilbert problem

$$G(\lambda) = \Theta_-^{-1}(\lambda)\Theta_+(\lambda) \tag{2.75}$$

where $\Theta_+(\lambda)$ is analytic inside Γ and has N zeroes located at the points $\mu_1, \cdots, \mu_N$, and $\Theta_-^{-1}(\lambda)$ is analytic outside Γ and has N zeroes at the points $\lambda_1, \cdots, \lambda_N$ (we emphasize that it is $\Theta_-^{-1}(\lambda)$ which has zeroes, and not $\Theta_-(\lambda)$). Let us fix the two set of subspaces:

$$\mathcal{V}_n = \mathrm{Im}\, \Theta_-^{-1}(\lambda)|_{\lambda=\lambda_n}, \quad \mathcal{W}_n = \mathrm{Ker}\, \Theta_+(\lambda)|_{\lambda=\mu_n}$$

Then we have:

Proposition. *The choice of the subspaces $\mathcal{V}_n, \mathcal{W}_n$ specifies uniquely the solution of the Riemann–Hilbert problem, up to a left multiplication by a* constant *matrix.*

<u>Proof</u>. Suppose $\Theta_\pm$ and $\widetilde{\Theta}_\pm$ are two solutions of the Riemann problem. Then the function $\chi = \widetilde{\Theta}_- \Theta_-^{-1} = \widetilde{\Theta}_+ \Theta_+^{-1}$ is a meromorphic function in the whole complex plane. Its possible poles are located at λ_n, μ_n. But around such a pole, let us say μ_n, we have

$$\Theta_+(\lambda) = F_n + O(\lambda - \mu_n), \qquad \Theta_+^{-1}(\lambda) = \frac{1}{\lambda - \mu_n} C_n + O(1),$$

$$\widetilde{\Theta}_+(\lambda) = \widetilde{F}_n + O(\lambda - \mu_n), \qquad \widetilde{\Theta}_+^{-1}(\lambda) = \frac{1}{\lambda - \mu_n} \widetilde{C}_n + O(1)$$

Recalling that $\mathrm{Ker}\, F_n = \mathcal{W}_n = \mathrm{Im}\, C_n$, $\mathrm{Ker}\, \widetilde{F}_n = \mathcal{W}_n = \mathrm{Im}\, \widetilde{C}_n$, we see that around $\lambda = \mu_n$, $\chi = \frac{1}{\lambda - \mu_n} \widetilde{F}_n C_n + O(1)$ is regular because $\widetilde{F}_n C_n = 0$. The same analysis holds true at λ_n, and so $\chi(\lambda)$ is regular everywhere, hence a constant. ■

There is a simple way to construct the solution of the Riemann–Hilbert problem with zeroes, from its fundamental solution.

Proposition. *Let Θ_+ and Θ_-^{-1} be the solution of the Riemann–Hilbert problem eq.(2.75) with zeroes at μ_N and λ_N respectively such that $\mathcal{W}_N = \mathrm{Ker}\, \Theta_+(\mu_N)$ and $\mathcal{V}_N = \mathrm{Im}\, \Theta_-^{-1}(\lambda_N)$ are fixed. Let $\widetilde{\Theta}_\pm$ be a solution of the Riemann–Hilbert problem, without zeroes at these points. Then*

$$\Theta_+(\lambda) = \chi_0^{-1} \left(1 - \frac{\mu_N - \lambda_N}{\lambda - \lambda_N} P_N \right) \widetilde{\Theta}_+(\lambda), \quad \Theta_-^{-1}(\lambda) = \widetilde{\Theta}_-^{-1}(\lambda) \left(1 - \frac{\lambda_N - \mu_N}{\lambda - \mu_N} P_N \right) \chi_0 \tag{2.76}$$

where P_N is a projector such that

$$\mathrm{Im}\, P_N = \widetilde{\Theta}_+(\mu_N) \mathcal{W}_N, \quad \mathrm{Ker}\, P_N = \widetilde{\Theta}_-(\lambda_N) \mathcal{V}_N$$

and χ_0 is a constant matrix.

<u>Proof</u>. Introduce the matrix $\chi(\lambda)$ as above. It is a meromorphic function in the λ–plane with possible simple poles at λ_N and μ_N, so we can write

$$\chi = \chi_0 + \frac{1}{\lambda - \mu_N} \chi_1, \quad \chi^{-1} = \chi_2 + \frac{1}{\lambda - \lambda_N} \chi_3$$

From the condition $\chi(\lambda)\chi^{-1}(\lambda) = 1$ we find $\chi_2 = \chi_0^{-1}$ and the two relations

$$\chi_1 \left(\chi_0^{-1} + \frac{1}{\mu_N - \lambda_N} \chi_3 \right) = 0, \quad \left(\chi_0 + \frac{1}{\lambda_N - \mu_N} \chi_1 \right) \chi_3 = 0$$

Adding these two equations, we get $\chi_0 \chi_3 = -\chi_1 \chi_0^{-1}$. Let us define

$$P_N = \frac{\chi_0 \chi_3}{\lambda_N - \mu_N} = -\frac{\chi_1 \chi_0^{-1}}{\lambda_N - \mu_N}$$

This matrix is a projector since

$$P_N^2 = -\frac{1}{(\lambda_N - \mu_N)^2}\chi_1\chi_0^{-1}\chi_0\chi_3 = -\frac{1}{(\lambda_N - \mu_N)^2}(\lambda_N - \mu_N)\chi_1\chi_0^{-1} = P_N$$

We can rewrite $\chi(\lambda)$ in terms of P_N:

$$\chi(\lambda) = \left(1 - \frac{\lambda_N - \mu_N}{\lambda - \mu_N}P_N\right)\chi_0; \quad \chi^{-1}(\lambda) = \chi_0^{-1}\left(1 - \frac{\mu_N - \lambda_N}{\lambda - \mu_N}P_N\right)$$

from which eqs.(2.76) immediately follow. Next we have $\mathcal{W}_N = \mathrm{Ker}\,\Theta_+(\mu_N) = \mathrm{Ker}\left((1-P_N)\widetilde{\Theta}_+(\mu_N)\right)$ so that $(1-P_N)\widetilde{\Theta}_+(\mu_N)\mathcal{W}_N = 0$ or $\mathrm{Ker}\,(1-P_N) = \mathrm{Im}\,P_N = \widetilde{\Theta}_+(\mu_N)\mathcal{W}_N$. Similarly, $\mathcal{V}_N = \mathrm{Im}\,\Theta_-^{-1}(\lambda_N) = \mathrm{Im}\,\widetilde{\Theta}_-^{-1}(\lambda_N)(1-P_N)\chi_0$, or what is the same $\mathcal{V}_N = \widetilde{\Theta}_-^{-1}(\lambda_N)\mathrm{Ker}\,P_N$, so that $\mathrm{Ker}\,P_N = \widetilde{\Theta}_-(\lambda_N)\mathcal{V}_N$. ■

Repeated application of this result allows one to build a solution of the Riemann–Hilbert problem with zeroes at $\mu_1, \cdots, \mu_N$, $\lambda_1, \cdots, \lambda_N$.

$$\begin{aligned}
\Theta_+(\lambda) &= \chi_N^{-1}\left(1 - \frac{\mu_N - \lambda_N}{\lambda - \lambda_N}P_N\right)\cdots\chi_1^{-1}\left(1 - \frac{\mu_1 - \lambda_1}{\lambda - \lambda_1}P_1\right)\widetilde{\Theta}_+(\lambda) \\
\Theta_-^{-1}(\lambda) &= \widetilde{\Theta}_-^{-1}(\lambda)\left(1 - \frac{\lambda_1 - \mu_1}{\lambda - \mu_1}P_1\right)\chi_1\cdots\left(1 - \frac{\lambda_N - \mu_N}{\lambda - \mu_N}P_N\right)\chi_N
\end{aligned}$$

Here $\widetilde{\Theta}_\pm(\lambda)$ refers to the fundamental solution of the Riemann–Hilbert problem.

We now use these solutions with zeroes of the Riemann–Hilbert problem to construct new solutions of the zero curvature equations. We start with the linear system

$$(\partial_x + U(x,t,\lambda))\Psi = 0, \quad (\partial_t + V(x,t,\lambda))\Psi = 0$$

and dress it with a solution with zeroes of the Riemann–Hilbert problem, according to eqs.(2.67, 2.68):

$$\Psi^G = \Theta_\pm \Psi G_\pm^{-1}, \quad \Theta_-^{-1}\Theta_+ = \Psi^{-1}G\Psi$$

In general, the components of the dressed connexion will have simple poles at the points μ_n, λ_n.

$$U^G(\lambda; x, t) = \Theta_\pm \cdot U \cdot \Theta_\pm^{-1} - \partial_x\Theta_\pm \cdot \Theta_\pm^{-1}, \quad V^G(\lambda; x, t) = \Theta_\pm \cdot V \cdot \Theta_\pm^{-1} - \partial_t\Theta_\pm \cdot \Theta_\pm^{-1} \tag{2.77}$$

We must require that the residues of these poles vanish. At $\lambda = \mu_n$ we have

$$\Theta_+(\lambda) = F_n + O(\lambda - \mu_n),\ \Theta_+^{-1}(\lambda) = \frac{1}{\lambda - \mu_n}C_n + O(1)$$

and the vanishing of the residues read

$$F_n(\partial_x + U|_{\lambda=\mu_n})C_n = 0, \quad F_n(\partial_t + V|_{\lambda=\mu_n})C_n = 0$$

This means that the space $\mathcal{W}_n = \mathrm{Ker}\,F_n = \mathrm{Im}\,C_n$ should be invariant under the action of the operators $\partial_x + U|_{\lambda=\mu_n}$ and $\partial_t + V|_{\lambda=\mu_n}$. Similarly, at $\lambda = \lambda_n$ we have

$$\Theta_-^{-1}(\lambda) = F_n + O(\lambda - \lambda_n),\ \Theta_(\lambda) = \frac{1}{\lambda - \lambda_n}C_n + O(1)$$

and setting the residues to zero give

$$C_n(\partial_x + U|_{\lambda=\lambda_n})F_n = 0, \quad C_n(\partial_t + V|_{\lambda=\lambda_n})F_n = 0$$

This means that the space $\mathcal{V}_n = \mathrm{Im}\,F_n = \mathrm{Ker}\,C_n$ should be invariant under the action of the operators $\partial_x + U|_{\lambda=\lambda_n}$ and $\partial_t + V|_{\lambda=\lambda_n}$.

The simplest way to implement these invariance properties is to set

$$\mathcal{V}_n(x,t) = \Psi(x,t,\lambda_n)\mathcal{V}_n(0), \quad \mathcal{W}_n(x,t) = \Psi(x,t,\mu_n)\mathcal{W}_n(0)$$

where $\Psi(x,t,\lambda)$ is the solution of the undressed linear system.

So, from the knowledge of $\Psi(x,t,\lambda)$, $\mathcal{V}_n(0), \mathcal{W}_n(0)$, and the fundamental solution of the Riemann–Hilbert problem, we can produce new solutions of the equations of motion by introducing zeroes into the Riemann–Hilbert problem.

The interest of this procedure is that it yields non trivial results even if the Riemann–Hilbert problem is trivial i.e. $G(\lambda) = Id$. Then its fundamental solution is also trivial $\Theta_\pm(\lambda) = Id$, and the solutions with zeroes are constructed by purely algebraic means. The resulting $\Theta_\pm(\lambda)$ are rational functions of λ.

We also need a solution of the zero curvature condition

$$\partial_t U - \partial_x V + [U, V] = 0$$

to start with. Simple solutions can be found in the form

$$U = A(x,\lambda), \quad V = B(t,\lambda), \quad [A, B] = 0$$

then $\Psi = \exp\left(-\int_0^x A dx - \int_0^t B dt\right)$. The solutions obtained by dressing this simple type of solutions by the trivial Riemann–Hilbert problem with zeroes are called the soliton solutions.

Example. Let us illustrate this construction on the non linear sigma model. There, one has

$$U = -\frac{1}{\lambda - 1} J_x, \quad V = -\frac{1}{\lambda + 1} J_t$$

We consider the case of 2x2 matrices. The non–linear sigma model field is related to J_x, J_t, by $J_x = g^{-1}\partial_x g, J_t = g^{-1}\partial_t g$. The matrix g can be easily related to the solution Ψ of the linear system:

$$g = \Psi^{-1}(\lambda = 0, x, t)$$

A trivial solution of the equations of motion is to take

$$J_x = a\sigma_3, J_t = b\sigma_3, \quad \Psi_0 = \exp\left[\left(\frac{ax}{\lambda - 1} - \frac{bt}{\lambda + 1}\right)\sigma_3\right], \quad g_0 = \exp -(ax + bt)\sigma_3$$

We want to dress this solution, by solving a Riemann–Hilbert problem with zeroes at λ_1 and μ_1. According to the general formulae, we have

$$\Theta_+(\lambda) = \chi_0^{-1}\left(1 - \frac{\mu_1 - \lambda_1}{\lambda - \lambda_1} P\right), \quad \Theta_-^{-1}(\lambda) = \left(1 - \frac{\lambda_1 - \mu_1}{\lambda - \mu_1} P\right)\chi_0$$

The projector P can be parametrized by two vectors w and n

$$P_{ij} = \frac{w_i n_j}{w \cdot n}$$

The spaces $\mathcal{V}_1$ and $\mathcal{W}_1$ are defined by

$$\mathcal{V}_1 = \mathrm{Im}\,\Theta_-^{-1}(\lambda)|_{\lambda=\lambda_1} = \mathrm{Im}\,(1 - P)\chi_0 = \mathrm{Ker}\,P$$

$$\mathcal{W}_1 = \mathrm{Ker}\,\Theta_+(\lambda)|_{\lambda=\mu_1} = \mathrm{Ker}\,(1 - P)\chi_0 = \mathrm{Im}\,P$$

Hence $\mathcal{W}_1$ is spanned by the vector w, and $\mathcal{V}_1$ is spanned by the vector $n^\perp$ perpendicular to n. The invariance properties of these spaces are met if we set

$$w(x,t) = \Psi(x,t,\mu_1)w(0), \quad n^\perp(x,t) = \Psi(x,t,\lambda_1)n^\perp(0),$$

In components, this reads

$$w_1(x,t) = e^{\left(\frac{ax}{\mu_1-1}-\frac{bt}{\mu_1+1}\right)}w_1(0), \quad w_2(x,t) = e^{-\left(\frac{ax}{\mu_1-1}-\frac{bt}{\mu_1+1}\right)}w_2(0)$$

and

$$n_1(x,t) = e^{-\left(\frac{ax}{\lambda_1-1}-\frac{bt}{\lambda_1+1}\right)}n_1(0), \quad n_2(x,t) = e^{\left(\frac{ax}{\lambda_1-1}-\frac{bt}{\lambda_1+1}\right)}n_2(0)$$

these formulae completely determine the projector $P(x,t)$. The dressed field is then reconstructed by

$$g(x,t) = \Psi^{-1}(\lambda=0) = \Psi_0^{-1}(\lambda=0)\Theta_-^{-1}(\lambda=0) = g_0(x,t)\left(1+\frac{\lambda_1-\mu_1}{\mu_1}P(x,t)\right)\chi_0$$

We do not write the explicit formulae for the components of g as they are sufficiently cumbersome to show the non triviality of the procedure.

References

[1] L.D.Faddeev, *"Integrable Models in 1+1 Dimensional Quantum Field Theory"*. Les Houches Lectures 1982. Elsevier Science Publishers (1984).

[2] L. D. Faddeev, L. A. Takhtajan. **"Hamiltonian Methods in the Theory of Solitons"**. Springer 1986.

[3] V.E.Zakharov, A.B.Shabat *" Integration of Non Linear Equations of Mathematical Physics by the Method of Inverse Scattering.II"* Funct.Anal.Appl. 13,(1979) 166.

[4] M. Jimbo, T. Miwa, *"Monodromy preserving deformation of linear ordinary differential equations with rational coefficients. III"* Physica D 4 (1981) pp.26–46.

[5] L.A. Dickey **"Soliton Equations and Hamiltonian Systems."** World Scientific (1991)

[6] L.A. Dickey *"On the τ–function of matrix hierarchies of integrable equations."* J.Math.Phys 32 (1991) P.2996–3002.

[7] M. Semenov Tian Shansky. *"Dressing transformations and Poisson group actions"*. Publ. RIMS 21 (1985), p. 1237.

[8] J. Avan, M. Bellon *" Groups of Dressing Transformations for Integrable Models in Dimension Two."* Phys. Lett. B213 (1988) p.459-465.

Chapter 3

Algebraic Methods.

We discuss abstractly the group theoretical aspects of the theory. This covers both the case of a finite dimensional Lie group, and the case of an infinite dimensional one, such as a loop group. In this case the spectral parameter is interpreted as the loop parameter. Quite generally, the solution of the equations of motion is reduced to a factorization problem in Lie group theory. In the case of loop groups this problem often takes the form of a Riemann–Hilbert factorization. We illustrate these constructions in the case of finite dimensional Lie groups, with the open Toda chain model, which can be completely solved by algebraic methods. Finally, we show the effectiveness of these structures by constructing the Lax pair formulation of the Kowalevsky top.

3.1 The classical and modified Yang-Baxter equations.

The basic ingredient for the construction of integrable systems is the Lax matrix L. Moreover Liouville integrability is equivalent to the existence of an R–matrix, which then obeys a consistency condition coming from the Jacobi identity for the Poisson brackets. In many particular cases the R–matrix is non dynamical, i.e., all its entries are numbers. In this case the consistency condition has a nice algebraic interpretation, allowing further analysis. In particular it provides a general framework for reducing the solution of the equations of motion to a factorization problem in Lie group theory.

The r–matrix relation, eq.(1.6) in Chapter [1] can be presented in a dual form. Suppose $\mathcal{G}$ is a Lie algebra equipped with a non degenerate invariant scalar product $(\ ,\ )$. We will use a basis $\{t_a\}$ of $\mathcal{G}$, and denote $g_{ab} = (t_a, t_b)$ and g^{ab} its inverse. The Lax matrix L may be viewed as a linear form on $\mathcal{G}$, ie. as element of $\mathcal{G}^*$, the dual of $\mathcal{G}$:

$$X \in \mathcal{G} \longrightarrow L(X) = (L, X).$$

We may as well view r as a linear map $R : \mathcal{G} \longrightarrow \mathcal{G}$. If $r_{12} = \sum_{ab} r^{ab} t_a \otimes t_b$, then

$$R(X) = \sum_{ab} r^{ab} t_a (t_b, X) \tag{3.1}$$

In the dual form, the Poisson brackets eq.(1.6) reads

$$\{L(X), L(Y)\} = L([X, Y]_R) \quad \text{with} \quad [X, Y]_R = [R(X), Y] + [X, R(Y)] \tag{3.2}$$

No particular symmetry of R is assumed here. In this dual formalism, the Jacobi identity becomes an equation on R:

$$L\Big(\Big[X, \{L(Y), R(Z)\} - \{L(Z), R(Y)\} \quad + \quad [R(Y), R(Z)] - R\left([Y, Z]_R\right)\Big]\Big) \\ + \qquad cyclic\,perm. = 0 \tag{3.3}$$

The simplest example is provided by the case in which the R-matrix is proportional to the identity map. It corresponds to the so-called Kirillov bracket on $\mathcal{G}^*$. Indeed, suppose that $\mathcal{G}$ is a Lie algebra equipped with a non-degenerate invariant scalar product as above. The Kirillov Poisson brackets are defined on $\mathcal{G}^*$ by:

$$\{L(X), L(Y)\}_K = L([X, Y])$$

for any $X,\ Y$ in $\mathcal{G}$. It is of the form (1.6) with $r_K = \frac{1}{2}\sum_{ab} g^{ab} t_a \otimes t_b = \frac{1}{2}C_{12}$ with C_{12} the tensor Casimir of $\mathcal{G}$. Under dualization, it corresponds to $R_K = \frac{1}{2}$. The Kirillov bracket can be written more explicitely by using the basis $\{t_a\}$ and its dual $\{t^a\}$. For $L = \sum_a L_a t^a$ an element of $\mathcal{G}^*$, its components L_a are functions on $\mathcal{G}^*$ and their Kirillov Poisson brackets read:

$$\{L_a, L_b\} = f_{ab}^c\, L_c$$

with f_{ab}^c the structure constants of $\mathcal{G}$. The Lax matrix can be viewed as an element of $\mathcal{G}^*$, the dual of some Lie algebra $\mathcal{G}$. The Poisson brackets of the Lax matrix reads:

$$\{L(X), L(Y)\} = L\left([X, Y]_R\right), \quad \text{for} \quad X, Y \in \mathcal{G} \tag{3.4}$$

where the R–bracket is defined by:

$$[X, Y]_R = [R(X), Y] + [X, R(Y)] \tag{3.5}$$

The Jacobi identity takes the form of eq.(3.3) in Chapter [1]. If the R matrix is a constant on phase space (independent of the dynamical variables), the Jacobi identity becomes

$$L\Big(\Big[X, [R(Y), R(Z)] - R\left([Y, Z]_R\right)\Big]\Big) \quad + \quad \text{cyclic perm.} = 0$$

A particular way to fulfill this equation is to set

$$[R(X), R(Y)] - R\left([X, R(Y)] + [R(X), Y]\right) = -\frac{1}{4}[X, Y] \tag{3.6}$$

This equation is called the *modified* Yang-Baxter equation and will be extensively studied below.

Before doing this, we would like to explain its relation with another important equation appearing in this domain: the *classical* Yang-Baxter equation. For an r-matrix in $\mathcal{G} \otimes \mathcal{G}$ (not necessarily antisymmetric) this equation reads

$$[r_{12}, r_{13}] + [r_{12}, r_{23}] + [r_{13}, r_{23}] = 0 \tag{3.7}$$

This equation is important because it is the classical limit of the quantum Yang–Baxter equation, which is one of the main tools in the study of many quantum integrable models. In dualized form it reads

$$[R(X), R(Y)] - R([X, R(Y)] - [{}^tR(X), Y]) = 0 \tag{3.8}$$

Please, notice the subtle difference between the left-hand sides of eq.(3.6) and eq.(3.8). The two expressions agree if ${}^tR = -R$, i.e., if the r–matrix is antisymmetric.

The relation between the modified Yang-Baxter equation and the classical Yang-Baxter equation is as follows:

Proposition. *Let R be an antisymmetric solution of the modified Yang-Baxter equation, then $R^{\pm} = R \pm \frac{1}{2}$ both satisfy the classical Yang-Baxter equation.*

Proof. We compute

$$R^{\pm}([X,Y]_R) = R([X,Y]_R) \pm \tfrac{1}{2}[X,Y]_R = [R^{\pm}(X), R^{\pm}(Y)]$$

the statement follows from the observation that

$$[X,Y]_R = [R^{\mp}(X), Y] + [X, R^{\pm}(Y)]$$

and the fact that $R^{\mp} = -{}^tR^{\pm}$ when R is antisymmetric. ∎

Viewed as an element of $\mathcal{G} \otimes \mathcal{G}$, $R^{\pm}$ corresponds to $r_{12}^{\pm}$ with

$$r_{12}^{\pm} = r_{12} \pm \tfrac{1}{2} C_{12} \tag{3.9}$$

where r_{12} is the dualized form of the antisymmetric solution of the modified Yang–Baxter equation, and C_{12} is the tensor Casimir element in $\mathcal{G} \otimes \mathcal{G}$. Recall that if $\{t_a\}$ is a basis of $\mathcal{G}$, the tensor Casimir is $C_{12} = \sum_{a,b} g^{ab} t_a \otimes t_b$ with g^{ab} the inverse of $g_{ab} = (t_a, t_b)$, where this scalar product is given by an invariant non–degenerate bilinear form on $\mathcal{G}$.

3.2 Algebraic meaning of the Yang-Baxter equations.

We now undertake an abstract study of the bracket eq.(3.5), for a generic linear mapping $R : \mathcal{G} \longrightarrow \mathcal{G}$.

Proposition. *Let R be a solution of the modified Yang-Baxter equation, then the antisymmetric bracket $[\ ,\ ]_R$ satisfies the Jacobi identity. It thus defines another Lie algebra structure on $\mathcal{G}$.*

Proof. We have to prove that the bracket $[\ ,\ ]_R$ satisfies the Jacobi identity:

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

Expanding the external R–brackets, we get:

$$\begin{aligned}&[R(X),[Y,Z]_R] + [R(Z),[X,Y]_R] + [R(Y),[Z,X]_R] + \\ &\quad [X, R([Y,Z]_R)] + [Z, R([X,Y]_R)] + [Y, R([Z,X]_R)] \quad = \quad 0\end{aligned}$$

Developing the terms in the first line and using the Jacobi identity on the Lie bracket $[\ ,\ ]$, we can rewrite them as

$$-[X,[R(Y),R(Z)]] - [Y,[R(Z),R(X)]] - [Z,[R(X),R(Y)]]$$

They combine with the terms in the second line to yield:

$$[X,[R(Y),R(Z)] - R([Y,Z]_R)] + \text{cyclic permutations} = 0$$

This is satisfied if R is a solution of the modified Yang–Baxter equation because the original Lie bracket $[\ ,\ ]$ obeys the Jacobi identity. ∎

We are then in a very special situation where the vector space $\mathcal{G}$ is equipped with two Lie algebra structures defined by the two brackets $[\ ,\ ]$, and $[\ ,\ ]_R$. We will call $\mathcal{G}_R$, the Lie algebra with underlying vector space $\mathcal{G}$ but with Lie bracket $[\ ,\ ]_R$. The fact that $[\ ,\ ]_R$ is a Lie bracket implies the following interpretation of eq.(3.4). Namely, *the Poisson bracket structure of the Lax matrix is given by the Kirillov bracket for the algebra $\mathcal{G}_R$*. The relation between these two Lie algebra structures is described by the following:

Proposition. *Let us define* $R^{\pm} = R \pm \frac{1}{2}$, $\mathcal{K}_{\pm} = \mathrm{Ker}\,(R^{\mp})$ *and* $\mathcal{G}_{\pm} = \mathrm{Im}\,(R^{\pm})$, *then*
(i) $(R^{\pm}) : \mathcal{G}_R \longrightarrow \mathcal{G}$ *are Lie algebra homomorphisms.*
(ii) $\mathcal{G}_{\pm} \subset \mathcal{G}$ *are Lie subalgebras of* $\mathcal{G}$.
(iii) $\mathcal{K}_{\pm} \subset \mathcal{G}_R$ *are ideals of* $\mathcal{G}_R$ *and* $\mathcal{G}_{\pm} = \mathcal{G}_R/\mathcal{K}_{\mp}$.

Proof. This proposition is a straightforward consequence of the modified Yang-Baxter equation once we notice that $R^{\pm}[X,Y]_R = R[X,Y]_R \pm \frac{1}{2}[X,Y]_R = [R^{\pm}X, R^{\pm}Y]$. ∎

Given the maps R^+ and R^-, we thus construct two subalgebras $\mathcal{G}_{\pm} \in \mathcal{G}$. Since $R^+ - R^- = 1$, any element $X \in \mathcal{G}$ can be written as

$$X = X_+ - X_- \quad \text{with} \quad X_{\pm} = R^{\pm}(X) \in \mathcal{G}_{\pm} = \mathrm{Im}\,(R^{\pm}) \tag{3.10}$$

Written in term of this decomposition, the bracket $[\ ,\ ]_R$ takes the simple form:

$$[X,Y]_R = [X_+, Y_+] - [X_-, Y_-]$$

This means that $\mathcal{G}_R$ is a subalgebra of $\mathcal{G}_+ \oplus \mathcal{G}_-$. (In the direct sum we have $[\mathcal{G}_+, \mathcal{G}_-] = 0$). The question then arises to determine the image of $\mathcal{G}_R$ in $\mathcal{G}_+ \oplus \mathcal{G}_-$. This is the object of the next proposition.

Proposition.
(i) $\mathcal{K}_{\pm} \subset \mathcal{G}_{\pm}$ *are ideals in* $\mathcal{G}_{\pm}$, *hence* $\mathcal{G}_{\pm}/\mathcal{K}_{\pm}$ *are Lie algebras.*
(ii) The mapping $\theta : \mathcal{G}_+/\mathcal{K}_+ \longrightarrow \mathcal{G}_-/\mathcal{K}_-$ *defined by* $\theta : R^+X \longrightarrow R^-X$ *is a Lie algebra isomorphism.*

Proof. To prove the first part of the proposition, we remark that $\mathcal{K}_{\pm} \subset \mathcal{G}_{\pm}$ since for $X \in \mathcal{K}_{\pm}$ we have by definition $RX = \pm\frac{1}{2}X$, so that $X = \pm R^{\pm}X$. For the same reason, on $\mathcal{K}_{\pm}$ we have $[\ ,\ ]_R = \pm[\ ,\ ]$ and $\mathcal{K}_{\pm}$ are indeed subalgebras in $\mathcal{G}_{\pm}$. To prove that they are ideals, we consider $X \in \mathcal{K}_{\pm}$, $Y \in \mathcal{G}_{\pm}$. We can write $Y = R^{\pm}Z$ then

$$[X,Y] = [X, R^{\pm}Z] = [X, RZ] \pm \tfrac{1}{2}[X,Z] = [X,RZ] + [RX,Z] = [X,Z]_R$$

but $[X,Z]_R \in \mathcal{K}_{\pm}$ since $\mathcal{K}_{\pm}$ are ideals in $\mathcal{G}_R$.
To prove the second part let us denote by $\overline{X}_{\pm}$ the equivalence class

$$\overline{X}_{\pm} = R^{\pm}X \ \ [\mathrm{mod}\ \mathcal{K}_{\pm}].$$

First $\theta : \mathcal{G}_+/\mathcal{K}_+ \longrightarrow \mathcal{G}_-/\mathcal{K}_-$ is well defined, since an element of $\mathcal{K}_+$ is mapped to 0. The mapping θ is surjective because it is induced by the surjective mapping $\mathcal{G}_+ \to \mathcal{G}_-$ given by $R^+X \to R^-X$. It is injective because if $R^-X \in \mathcal{K}_-$ one has $R^+(X) = 0$ by definition of $\mathcal{K}_-$. Finally we prove that θ is a Lie algebra isomorphism, i.e: $[\theta(x), \theta(y)] = \theta[x,y]$ for any x, y of the form $x = \overline{R^+X}$, $y = \overline{R^+Y}$. Recalling that $[R^{\pm}X, R^{\pm}Y] = R^{\pm}[X,Y]_R$ and due to the definition of the Lie algebra bracket on $\mathcal{G}_{\pm}/\mathcal{K}_{\pm}$ we have:

$$\theta([\overline{R^+X}, \overline{R^+Y}]) = \theta(\overline{R^+[X,Y]_R}) = \overline{R^-[X,Y]_R} = [\overline{R^-X}, \overline{R^-Y}] = [\theta(\overline{R^+X}), \theta(\overline{R^+Y})]$$

which ends the proof of the proposition. ∎

Finally we have the following important result, which will allow to define the factorization problem:

Proposition. *Consider the two maps:*

$$\mathcal{G}_R \xrightarrow{R^+,R^-} \mathcal{G}_+ \oplus \mathcal{G}_- \xrightarrow{1,-1} \mathcal{G}$$

The first map is a Lie algebra injective homomorphism, and the second map is surjective. We call $\widetilde{\mathcal{G}}_R$ the image of $\mathcal{G}_R$ in $\mathcal{G}_+ \oplus \mathcal{G}_-$. It is characterized by the set of couples (R^+X, R^-Y) such that $\theta(\overline{R^+X}) = \overline{R^-Y}$.

Proof. The first map is injective and the second one is surjective because $R^+ - R^- = 1$. The elements of $\widetilde{\mathcal{G}}_R$ are of the form (R^+X, R^-X) hence satisfy the condition $\theta(\overline{R^+X}) = \overline{R^-X}$. Conversely let us start from a pair (R^+X, R^-Y) such that $\theta(\overline{R^+X}) = \overline{R^-Y}$. This means $\overline{R^-Y} = \overline{R^-X}$ hence there exists $K_- \in \mathcal{K}_-$ such that $R^-(X-Y) = K_- = -R^-K_-$, so that $(X - Y + K_-) = K_+$ belongs to $\mathrm{Ker}\, R^- = \mathcal{K}_+$. Then we have $R^+X = R^+\widetilde{X}$, $R^-Y = R^-\widetilde{X}$ with $\widetilde{X} = X + K_- = Y + K_+$ so that the point (R^+X, R^-Y) belongs to $\widetilde{\mathcal{G}}_R$. ∎

In the next section we give a simple example of this algebraic setting, thereby providing a powerful solution factory for the (modified) Yang–Baxter equation.

3.3 Adler-Kostant-Symes scheme.

A class of solutions of the modified Yang-Baxter equation is produced by the Adler-Kostant-Symes scheme. Let $\mathcal{G}$ be a Lie algebra, and assume we have two Lie subalgebras $\mathcal{A}$ and $\mathcal{B}$ such that, *as a vector space*, it is the direct sum

$$\mathcal{G} = \mathcal{A} + \mathcal{B}$$

Note that $\mathcal{A}$ and $\mathcal{B}$ are Lie subalgebras but they are not assumed to commute ($[\mathcal{A}, \mathcal{B}] \neq 0$). We denote by $P_\mathcal{A}$ the projection on $\mathcal{A}$ along $\mathcal{B}$ and $P_\mathcal{B} = 1 - P_\mathcal{A}$.

Proposition. *The linear map $R = \frac{1}{2}(P_\mathcal{A} - P_\mathcal{B})$ satisfies the modified Yang-Baxter equation.*

Proof. By definition we have $[X, Y]_R = [P_\mathcal{A}X, P_\mathcal{A}Y] - [P_\mathcal{B}X, P_\mathcal{B}Y]$, where we used $P_\mathcal{A} + P_\mathcal{B} = 1$, thus:

$$R([X, Y]_R) = \frac{1}{2}\left([P_\mathcal{A}X, P_\mathcal{A}Y] + [P_\mathcal{B}X, P_\mathcal{B}Y]\right)$$

Since

$$[RX, RY] = \frac{1}{4}\left([P_\mathcal{A}X, P_\mathcal{A}Y] - [P_\mathcal{A}X, P_\mathcal{B}Y] - [P_\mathcal{B}X, P_\mathcal{A}Y] + [P_\mathcal{B}X, P_\mathcal{B}Y]\right),$$

we obtain

$$[RX, RY] - R([X, Y]_R) = -\frac{1}{4}[X, Y].$$

This is the modified Yang-Baxter equation. ∎

This construction is a particular example of the general discussion. With the notations of the previous section, $R_+ = R + \frac{1}{2} = P_\mathcal{A}$ and $R_- = R - \frac{1}{2} = -P_\mathcal{B}$, and $\mathcal{G}_+ = \mathcal{K}_+ = \mathcal{A}$, $\mathcal{G}_- = \mathcal{K}_- = \mathcal{B}$. Thus $\mathcal{G}_\pm/\mathcal{K}_\pm = \{0\}$ so we have the decomposition of $\mathcal{G}_R$ as a direct sum *of Lie algebras*

$$\mathcal{G}_R = \mathcal{A} \oplus \mathcal{B}$$

In the Lie algebra $\mathcal{G}_R$, $\mathcal{A}$ and $\mathcal{B}$ commute. It is then clear that any element $X \in \mathcal{G}$ can be decomposed as $X = X_+ - X_-$ with $X_+ = P_\mathcal{A}X$ and $X_- = -P_\mathcal{B}X$.

Example. Let $\widetilde{\mathcal{G}}$ be the loop algebra $\widetilde{\mathcal{G}} = \mathcal{G} \otimes C[\lambda]$ with $\mathcal{G}$ a finite dimensional simple Lie algebra. Elements of $\widetilde{\mathcal{G}}$ are linear combinations of elements $X \otimes \lambda^n$ with $X \in \mathcal{G}$ and $n \in \mathbb{Z}$. The commutation relations in $\widetilde{\mathcal{G}}$ are:

$$[X \otimes \lambda^n, Y \otimes \lambda^m] = [X, Y] \otimes \lambda^{n+m}$$

The two subalgebras $\mathcal{A} = \mathcal{G}_+$, $\mathcal{B} = \mathcal{G}_-$ are spanned by elements of the form $X \otimes \lambda^n$ with $n \geq 0$ and $n < 0$ respectively. Clearly $[\mathcal{A}, \mathcal{B}] \neq 0$ for the Lie algebra structure of $\widetilde{\mathcal{G}}$. For any $X \in \mathcal{G}$ and any formal power series $f(\lambda) = \sum_{n \in \mathbb{Z}} f_n \lambda^n$, the maps $R^\pm$ are defined by:

$$\begin{aligned} R^+(X \otimes f(\lambda)) &= X \otimes f_+(\lambda) \\ R^-(X \otimes f(\lambda)) &= -X \otimes f_-(\lambda) \end{aligned}$$

where $f_+(\lambda) = \sum_{n \geq 0} f_n \lambda^n$ and $f_-(\lambda) = \sum_{n<0} f_n \lambda^n$ denote the regular and singular part of $f(\lambda)$ around $\lambda = 0$ respectively. Since by definition $f(\lambda) = f_+(\lambda) + f_-(\lambda)$, we have $R^+ - R^- = \mathbf{1}$. Elements of the subalgebras $\mathcal{G}_\pm$ are of the form $X \otimes f_\pm(\lambda)$ with $f_\pm(\lambda)$ regular (resp. singular) at the origin. We have $\mathcal{K}_\pm = \mathrm{Ker}\, R^\mp = \mathcal{G}_\pm$. Thus $\mathcal{G}_\pm/\mathcal{K}_\pm = \{0\}$ and the map θ is trivial. The decomposition eq.(3.10) for $X \otimes f(\lambda)$ consists in writing $X \otimes f(\lambda)$ into the sum of two functions, one analytic around the origin and the other one analytic around infinity. Introducing the scalar product in $\widetilde{\mathcal{G}}$:

$$(X(\lambda), Y(\lambda)) = \frac{1}{2i\pi} \oint d\lambda \mathrm{Tr}\,(X(\lambda)Y(\lambda))$$

where Tr ( ) is the invariant bilinear form on $\mathcal{G}$ we can write the operator R^+ in a more operatorial way: for any $X(\lambda)$ in $\widetilde{\mathcal{G}}$

$$R^+X(\lambda) = \frac{1}{2i\pi} \oint_{|\lambda|<|\mu|} \frac{\mathrm{Tr}_2(C_{12} 1 \otimes X(\mu))}{\mu - \lambda}$$

where C_{12} is the tensor Casimir for $\mathcal{G}$. This means that $r_{12}^+ = -C_{12}/(\lambda - \mu) = C_{12} \sum_{n=0}^\infty \lambda^n/\mu^{n+1}$, where we expand in powers of λ/μ because $|\lambda| < |\mu|$. A similar computation for r_{12}^- yields $r_{12}^- = -C_{12}/(\lambda - \mu) = -C_{12} \sum_{n=0}^\infty \mu^n/\lambda^{n+1}$. Finally we get $r_{12} = (r_{12}^+ + r_{12}^-)/2$ in the form:

$$r_{12}(\lambda, \mu) = -\frac{C_{12}}{\lambda - \mu}$$

3.4 Construction of integrable systems.

Starting from a Lie algebra structure on $\mathcal{G}$ one gets a Poisson bracket on its dual $\mathcal{G}^*$. Better if we have two Lie algebra structures on $\mathcal{G}$, namely $[\ ,\ ]$ and $[\ ,\ ]_R$ one obtains two Poisson brackets $\{\ ,\ \}$ and $\{\ ,\ \}_R$ on $\mathcal{G}^*$. It turns out that functions on $\mathcal{G}^*$ invariant under the action $[\ ,\ ]$ are in the center of the Poisson structure $\{\ ,\ \}$ and cannot provide useful Hamiltonians. However they are not in the center of the second Poisson structure $\{\ ,\ \}_R$ and here, are commuting Hamiltonians, and the equations of motion take Lax form.

MOVED IN THE PREVIOUS CHAPTER

We first recall some notions about adjoint and coadjoint actions of Lie algebras and Lie groups, see Chapter [14]. Let G be a connected Lie group with Lie algebra $\mathcal{G}$. The group G acts on $\mathcal{G}$ by the adjoint action denoted Ad:

$$X \longrightarrow \mathrm{Ad}\, g(X) = gXg^{-1} \quad g \in G,\ X \in \mathcal{G}$$

Similarly the coadjoint action of G on the dual $\mathcal{G}^*$ of the Lie algebra $\mathcal{G}$ (i.e. the vector space of linear forms on the Lie algebra) is defined by:

$$\mathrm{Ad}^*\, g.L(X) = L(\mathrm{Ad}\, g^{-1}(X)), \quad g \in G,\ L \in \mathcal{G}^*,\ X \in \mathcal{G}$$

The infinitesimal version of these actions provides an adjoint action of the Lie algebra $\mathcal{G}$ on $\mathcal{G}$ and $\mathcal{G}^*$ denoted ad and ad* given by:

$$\begin{aligned} \mathrm{ad}\, X(Y) &= [X, Y] \quad X, Y \in \mathcal{G}, \\ \mathrm{ad}^* X.L(Y) &= -L([X, Y]) \quad X, Y \in \mathcal{G}, L \in \mathcal{G}^* \end{aligned}$$

This induces a coadjoint action of G on functions on $\mathcal{G}^*$, also denoted Ad^*. Denote by $\mathcal{F}(\mathcal{G}^*)$ the functions on $\mathcal{G}^*$. If $g \in G$, and $h \in \mathcal{F}(\mathcal{G}^*)$

$$\mathrm{Ad}^*\, g \,.\, h(L) = h(Ad^*\ g^{-1}(L))$$

and a similar formula for the infinitesimal action ad^*. For a function h on $\mathcal{G}^*$ the differential dh may be viewed as an element of $\mathcal{G}$. This is because $\mathcal{G}^*$ is a vector space, the differential dh is a linear form on $\mathcal{G}^*$, i.e. an element of $\mathcal{G}^{**} \sim \mathcal{G}$, and one can write for $\delta L \in \mathcal{G}^*$:

$$h(L + \delta L) = h(L) + \delta L(dh) + \mathcal{O}\left((\delta L)^2\right)$$

Proposition. *A function h on $\mathcal{G}^*$ is ad^* invariant if and only if*

$$L([dh, X]) = 0 \ \ \forall L \in \mathcal{G}^*, \forall X \in \mathcal{G}$$

that is to say $\mathrm{ad}^*\, dh.L = 0$

Proof. Taking a particular tangent vector of the form $\delta L = \mathrm{ad}^*\, X \cdot L$ we have:

$$h(L + t\ \mathrm{ad}^*\, X \cdot L) = h(L) + t\ \mathrm{ad}^*\, X \cdot L(dh) + \mathcal{O}(t^2)$$

so the invariance property of h reads: $L([dh, X]) = 0 \ \ \forall L \in \mathcal{G}^*, \forall X \in \mathcal{G}$, or $\mathrm{ad}^*\, dh \cdot L = 0$. Conversely if h is infinitesimally invariant, it is G–invariant assuming G connected. ∎

From the Lie algebra structure on $\mathcal{G}$, we may construct a Poisson bracket on the space $\mathcal{F}(\mathcal{G}^*)$ of functions on $\mathcal{G}^*$. This is the Kirillov bracket.

Definition. *If h_1 and $h_2 \in \mathcal{F}(\mathcal{G}^*)$ then the Kirillov bracket is defined by:*

$$\{h_1, h_2\}(L) = L([dh_1, dh_2])$$

It is obvious that the bracket $\{\ ,\ \}$ is antisymmetric and verifies the Jacobi identity. Choosing the two linear functions $h_1(L) = L(X)$ and $h_2(L) = L(Y)$ with $X,\ Y\ \in \mathcal{G}$, so that $dh_1 = X$ and $dh_2 = Y$, the Kirillov Poisson bracket reads:

$$\{L(X), L(Y)\} = L([X, Y])$$

where the right–hand side is the linear function $L \to L([X, Y])$. This Poisson bracket is very natural but one has to be aware that it is a degenerate Poisson bracket.

Proposition. *The kernel of the Kirillov bracket $\{\ ,\ \}$ is the set $\mathcal{I}(\mathcal{G}^*)$ of Ad^*–invariant functions.*

Proof. Suppose $k \in \mathcal{F}(\mathcal{G}^*)$ is in the kernel of $\{\ ,\ \}$, we have

$$\{k, f\}(L) = L([dk, df]) = \mathrm{ad}^* dk \cdot L(df) = 0$$

for all $f \in \mathcal{F}(\mathcal{G}^*)$. Thus $\mathrm{ad}^* dk \cdot L = 0$, and k is invariant. The converse is obvious. ∎

This proposition means that the kernel of the bracket $\{\ ,\ \}$ is the set of functions which are constant on the orbits of the coadjoint action of G. In particular the restriction of the Kirillov bracket to functions on a coadjoint orbit is non–degenerate.

We now show that this setting can be used to construct integrable systems. We start from an R–matrix solution of the modified Yang–Baxter equation. On $\mathcal{G}$, we have two Lie algebra structures, one associated to the original Lie bracket $[\ ,\ ]$ and one corresponding to the bracket $[\ ,\ ]_R$. We thus have two adjoint actions

$$\mathrm{ad}X(Y) = [X, Y], \quad \mathrm{ad}_R X(Y) = [X, Y]_R = [RX, Y] + [X, RY]$$

We also have two coadjoint actions of $\mathcal{G}$ on $\mathcal{G}^*$:

$$\mathrm{ad}^* X \cdot L(Y) = -L([X,Y]), \quad \mathrm{ad}^*_R X \cdot L(Y) = -L([X,Y]_R)$$

One can express ad^*_R in terms of ad^*

$$\mathrm{ad}^*_R X \cdot L = \mathrm{ad}^*(RX) \cdot L + R^* \mathrm{ad}^* X \cdot L$$

We then have two Poisson brackets on $\mathcal{F}(\mathcal{G}^*)$, $\{\ ,\ \}$ and $\{\ ,\ \}_R$, which are the Kirillov brackets for the two Lie algbera structures, ie.

$$\{h_1, h_2\}\,(L) = L([dh_1, dh_2]), \quad \{h_1, h_2\}_R(L) = L([dh_1, dh_2]_R)$$

We also have two connected Lie groups, denoted G and G_R, with their adjoint and coadjoint actions.

Theorem.
(i) The Ad^**–invariant functions on* $\mathcal{G}^*$ *are in involution with respect to both Poisson brackets.*
(ii) Choosing an Ad^**–invariant Hamiltonian* h*, the equation of motion on* $\mathcal{G}^*$ *with respect to the Poisson bracket* $\{\ ,\ \}_R$ *may be written in the two equivalent forms*

$$\frac{dL}{dt} = -\mathrm{ad}^*_R dh \cdot L \tag{3.11}$$

$$\frac{dL}{dt} = -\mathrm{ad}^*\, M \cdot L \quad \text{with} \quad M = R(dh) \tag{3.12}$$

Proof. Let us choose two ad^*–invariant functions on $\mathcal{G}^*$, h_1 and h_2 i.e. $\mathrm{ad}^*\, dh_i \cdot L = 0 \quad (i = 1, 2)$. h_1 and h_2 are in the kernel of $\{\ ,\ \}$, and a fortiori $\{h_1, h_2\} = 0$. On the other hand,

$$\begin{aligned} \{h_1, h_2\}_R(L) = L([dh_1, dh_2]_R) &= L([Rdh_1, dh_2] + [dh_1, Rdh_2]) \\ &= \mathrm{ad}^* dh_2.L(Rdh_1) - \mathrm{ad}^*\, dh_1.L(Rdh_2) = 0 \end{aligned}$$

This proves the first part of the proposition. For a general function f on phase space, the equation of motion reads $\dot{f} = \{h, f\}_R$, and by definition of df we have $\dot{f} = \dot{L}(df)$. Thus:

$$\dot{L}(df)(L) = \{h, f\}_R(L) = L([dh, df]_R) = -\mathrm{ad}^*_R\, dh \cdot L(df)$$

Since this is true for any f we get eq.(3.11). We may also express ad^*_R with the help of ad^*, getting $\mathrm{ad}^*_R\, dh \cdot L = \mathrm{ad}^*\,(Rdh) \cdot L + R^*\, \mathrm{ad}^*\, dh \cdot L = \mathrm{ad}^*\, M \cdot L$ with $M = Rdh$, where we have used the invariance of h. This proves eq.(3.12). ■

This theorem realises what was announced at the beginning of this section and builds an integrable system from the algebraic data $\mathcal{G}$, $\mathcal{G}_R$. In the case where $\mathcal{G}$ is equipped with an invariant non degenerate quadratic form $(\ ,\ )$, $\mathcal{G}$ and its dual $\mathcal{G}^*$ may be identified, and ad^* becomes ad. The condition for a function on $\mathcal{G}^*$ to be ad^* invariant may then be written as $[dh, L] = 0$. Moreover we see that the equation of motion of L, eq.(3.12), takes the form of a Lax equation:

$$\frac{dL}{dt} = [L, M], \quad M = R(dh) \tag{3.13}$$

Remark 1. Notice that M is not uniquely defined, as we can add to it anything which commutes with L, in particular any polynomial in L.

Remark 2. We have used the bracket $\{\ ,\ \}_R$ to construct the evolution equation, so the phase space is really a coadjoint orbit of G_R on $\mathcal{G}^*$ on which the Poisson bracket is non–degenerate and defines a symplectic structure. However the Lax form of the equation of motion shows that the evolution takes place on an orbit of G. Hence the flow occurs on the intersection of orbits of G and G_R.

3.5 Solving by factorisation.

We consider a Hamiltonian system with a Lax matrix $L \in \mathcal{G}^*$ ($\mathcal{G}^*$ the dual of $\mathcal{G}$). We assume that $\mathcal{G}$ is equipped with an invaraint bilinear form, so that $\mathcal{G}^*$ and $\mathcal{G}$ are identified. We also assume that the Poisson brackets of L are given as above by $\{L(X), L(Y)\} = L([X, Y]_R)$ where R is a solution of the modified Yang-Baxter equation.

As explained in the previous section, the Hamiltonians h_n in involution $\{h_n, h_m\} = 0$ are chosen among the Ad*–invariant functions on $\mathcal{G}^* \sim \mathcal{G}$, for example, $h_n = \mathrm{Tr}\,(L^n)$. Let us choose one of these functions as the Hamiltonian h. We then know that the equations of motion have a Lax representation, eq(3.13). Since h is an invariant function, one has $[L, dh] = 0$ and we can write as well, using $R^{\pm} = R \pm \frac{1}{2}$:

$$\frac{\partial L}{\partial t} = [L, M_+] = [L, M_-]\,; \quad M_{\pm} = R^{\pm}(dh) \tag{3.14}$$

The fact that the equations of motion admit two Lax representations as in eq.(3.14) is the key point of the following discussion.

Theorem. *The solution of the equations of motion with Hamiltonian h is given by*

$$L(t) = g_+(t)\; L_0\; g_+^{-1}(t) = g_-(t)\; L_0\; g_-^{-1}(t)$$

where g_+ and g_- are the solutions of the factorization problem

$$\exp(-t dh(L_0)) = g_-^{-1}(t)\, g_+(t) \tag{3.15}$$

with $g_{\pm} \in G_{\pm}, \theta(\overline{g}_+) = \overline{g}_-$.

Proof. We want to solve

$$\frac{dL}{dt} = [L, M_+], \quad \frac{dL}{dt} = [L, M_-]$$

where $M_{\pm}$ are defined in eq.(3.14). Since $M_{\pm} \in \mathcal{G}_{\pm}$ we know that there are $g_+(t) \in G_+$ and $g_-(t) \in G_-$ such that $\dot{g}_+ g_+^{-1} = -M_+$ and $\dot{g}_- g_-^{-1} = -M_-$, with initial value $g_{\pm}(0) = 1$. Then we have:

$$L(t) = g_+(t).L_0.g_+^{-1}(t) = g_-(t).L_0.g_-^{-1}(t)$$

Using $M_+ - M_- = dh$ we get:

$$dh(L) = -\dot{g}_+ g_+^{-1} + \dot{g}_- g_-^{-1} = -g_-\left(\frac{d}{dt}(g_-^{-1} g_+)\right) g_+^{-1}$$

From the invariance of h by coadjoint action on L, we see that $dh(L) = dh(g_- \cdot L_0 \cdot g_-^{-1}) = g_- \cdot dh(L_0) \cdot g_-^{-1}$, which implies:

$$\frac{d}{dt}(g_-^{-1} g_+) \cdot (g_-^{-1} g_+)^{-1} = -dh(L_0)$$

yielding $g_-^{-1}(t) g_+(t) = \exp\left(-t dh(L_0)\right)$. We are led to the algebraic problem of decomposing the right–hand side as a product of elements in $G_{\pm}$, the Lie groups corresponding to $\mathcal{G}_{\pm}$, which can be done, at least in the vicinity of the unit element. The solution of this algebraic problem yields the solution of the evolution equation, at least for small t. ∎

Remark. Suppose that $\mathcal{G}$ is equipped with an invariant bilinear form which we can used to identify $\mathcal{G}^*$ and $\mathcal{G}$. Then, a set of Ad*–invariant functions is provided by the trace of the powers of L: $h_n = \mathrm{Tr}\,(L^n)$. In this case, we have:

$$dh_n(L) = n\,\mathrm{Tr}\,\left(L_2^{n-1} C_{12}\right) \quad \text{and} \quad M_{\pm} = n\,\mathrm{Tr}\,\left(L_2^{n-1} r_{12}^{\pm}\right)$$

with C_{12} the tensor Casimir.

3.6 The Lax pair of the Kowalevski top.

3.7 The open Toda Chains.

In this section we illustrate the previous general considerations by working out the example of the Toda chain. For the open chain that we consider here, the natural Lax matrix does not require a spectral parameter, and we only work with finite dimensional Lie algebras. The algebraic constructions presented above are powerful enough to provide the solution of the model. In contrast, the study of the closed Toda chain requires considering loop algebras, and will be presented in Chapter [5].

We will first review some notations for Lie algebras introduced in Chapter [16]. Let $\mathcal{G}$ be a finite dimensional simple Lie algebra of rank r, equipped with an invariant scalar product denoted $(\ ,\ )$. We choose a Cartan subalgebra with an orthonormal basis $\{H_i\}$. We have the Cartan decomposition of $\mathcal{G}$:

$$\mathcal{G} = \mathcal{N}_- \oplus \mathcal{H} \oplus \mathcal{N}_+, \quad \text{with} \quad \mathcal{N}_\pm = \bigoplus_{\beta \ \mathit{positive}} \mathcal{G}_{\pm\beta}$$

where $\mathcal{G}_{\pm\beta}$ are 1–dimensional, generated by the root vectors $E_{\pm\beta}$ for any positive root β. We will need the commutation relations:

$$\begin{aligned} [H_i, H_j] &= 0 \\ [H, E_{\pm\alpha}] &= \pm\alpha(H)\ E_{\pm\alpha} \\ [E_\alpha, E_{-\alpha}] &= (E_\alpha, E_{-\alpha})\ H_\alpha, \end{aligned}$$

where $H_\alpha = \sum_i \alpha(H_i) H_i$. We will often use the constants n_α such that

$$n_\alpha^2\ (E_\alpha, E_{-\alpha}) = 1$$

Toda chains are associated to any simple Lie algebra $\mathcal{G}$ as follows. Consider two elements in the Cartan subalgebra:

$$q = \sum_{i=1}^{r} q_i H_i, \quad p = \sum_{i=1}^{r} p_i H_i, \quad r = \operatorname{rank} \mathcal{G}$$

The coefficients $(q_i, p_i)\ \ i = 1, \ldots, r$ are the coordinates of a $(2r)$–dimensional phase space with Poisson brackets $\{p_i, q_j\} = \delta_{ij}$. The Hamiltonian is by definition:

$$h = \tfrac{1}{2}(p, p) + \sum_{\alpha \text{ simple}} \exp\left(2\alpha(q)\right)$$

The equations of motion are then:

$$\begin{aligned} \frac{dq}{dt} &= \{h, q\} = p \\ \frac{dp}{dt} &= \{h, p\} = -\sum_{\alpha \text{ simple}} 2H_\alpha \exp\left(2\alpha(q)\right). \end{aligned}$$

In the usual $sl(N+1)$ case, the Lie algebra of traceless $(N+1)\times(N+1)$ matrices, the Cartan algebra can be taken as traceless diagonal matrices. The simple root vectors E_{α_i} are the canonical matrices $E_{i,i+1}$, and $E_{-\alpha_i} = E_{i+1,i}$ for $i = 1, \cdots, N$. In this case it is not convenient to use an orthonormal basis of the Cartan algebra under the scalar product $(X, Y) = \operatorname{Tr}(XY)$. Instead we write an element in the Cartan subalgebra as $q = \sum_{i=1}^{N+1} q_i E_{ii}$ with the constraint $\sum q_i = 0$. In this description the simple roots are $\alpha_i(q) = q_i - q_{i+1}$. The dynamical variables of the open Toda chain are the two elements $q = \sum_{i=1}^{N+1} q_i E_{ii}$ and $p = \sum_{i=1}^{N+1} p_i E_{ii}$

with the constraints $\sum q_i = 0$ and $\sum p_i = 0$. These constaints not being compatible with the canonical Poisson brackets, we use instead the Dirac brackets:

$$\{p_i, q_j\} = \delta_{ij} - \frac{1}{N+1}, \quad \{p_i, p_j\} = \{q_i, q_j\} = 0$$

The Hamiltonian is:

$$H = \frac{1}{2}\sum_{i=1}^{N+1} p_i^2 + \sum_{i=1}^{N} e^{2(q_i - q_{i+1})}$$

and generates the equations of motion:

$$\begin{aligned}
\dot{q}_i &= p_i, \quad i = 1, \cdots, N+1 \\
\dot{p}_1 &= 2\, e^{2(q_1 - q_2)} \\
\dot{p}_i &= 2\Big(e^{2(q_i - q_{i+1})} - e^{2(q_{i-1} - q_i)}\Big), \quad i = 2, \cdots N \\
\dot{p}_{N+1} &= -2\, e^{2(q_N - q_{N+1})}
\end{aligned}$$

The particular form of the equations for $\dot{p}_1$ and $\dot{p}_{N+1}$ explains the terminology "open" Toda chain.

The equations of motion of the open Toda chain can be written in Lax form and yield integrable systems.

Proposition. *The equations of the open Toda chain admit a Lax pair representation $\dot{L} = [L, M]$ with:*

$$L = p + \sum_{\alpha \text{ simple}} n_\alpha\, e^{\alpha(q)}\, (E_\alpha + E_{-\alpha})\,, \quad M = \sum_{\alpha \text{ simple}} n_\alpha\, e^{\alpha(q)}\, (E_\alpha - E_{-\alpha}) \tag{3.16}$$

Proof. We compute succesively the time derivative of L and the commutator $[L, M]$. We have

$$\frac{dL}{dt} = \frac{dp}{dt} + \sum_\alpha n_\alpha\, e^{\,\alpha(q)} \alpha(\frac{dq}{dt})\, (E_\alpha + E_{-\alpha})$$

$$\begin{aligned}
[L, M] &= \sum_\alpha n_\alpha \alpha(p)\, e^{\alpha(q)}\ \ (E_\alpha + E_{-\alpha}) \\
&\quad + \sum_\alpha \sum_\beta n_\alpha n_\beta\, e^{(\alpha+\beta)(q)} [E_\alpha + E_{-\alpha}, E_\beta - E_{-\beta}].
\end{aligned}$$

Since $[E_{\pm\alpha}, E_{\pm\beta}]$ is antisymmetric in the exchange of α and β, the second term reduces to

$$\sum_\alpha \sum_\beta n_\alpha n_\beta\, e^{(\alpha+\beta)(q)}\, ([E_{-\alpha}, E_\beta] - [E_\alpha, E_{-\beta}])\,.$$

Since α and β are simple roots, we have $[E_\alpha, E_{-\beta}] = \delta_{\alpha\beta}(E_\alpha, E_{-\alpha})H_\alpha$ and therefore

$$[L, M] = -\sum_\alpha\ 2H_\alpha\ e^{2\alpha(q)} + \sum_\alpha n_\alpha\ \alpha(p)\ e^{\alpha(q)}\ \ (E_\alpha + E_{-\alpha})\,,$$

from which the result follows. ∎

It is important to notice that this calculation is performed using only the Lie algebra structure of $\mathcal{G}$ and never refers to a representation. It thus extends naturally to the case of a Kac-Moody algebra.

This Lax representation provides conserved quantities, namely any invariant polynomial in L. Recall the Chevalley theorem which states that, for a simple Lie algebra, the ring of invariant polynomials is freely generated by $r = \operatorname{rank}\mathcal{G}$ primitive polynomials which are of degree $n_i, i = 1, \cdots, r$, the so–called exponents of the Lie algebra. We have thus the right number of conserved quantities at our disposal. It remains to show that they are in involution. This is the aim of the next section.

3.8 The R–matrix of the Toda models.

Since the R-matrix of the Toda chain is the canonical example of solution of the classical Yang–Baxter equation associated to simple Lie algebras, and plays an important role in the following as well as in many integrable models, we present its computation in details.

Proposition. *Let L be the Lax matrix (3.16). Then, there exists a unique antisymmetric r–matrix $r_{12} \in \mathcal{G} \otimes \mathcal{G}$, independent of p and q, such that $\{L_1, L_2\} = [r_{12}, L_1 + L_2]$.*

$$r_{12} = \sum_{\alpha \text{ positive}} \frac{E_\alpha \otimes E_{-\alpha} - E_{-\alpha} \otimes E_\alpha}{(E_\alpha, E_{-\alpha})} \tag{3.17}$$

Proof. We have

$$\{L_1, L_2\} = \sum_{\alpha\ simple} n_\alpha e^{\alpha(q)} [H_\alpha \otimes (E_\alpha + E_{-\alpha}) - (E_\alpha + E_{-\alpha}) \otimes H_\alpha]$$

and

$$[r_{12}, L_1 + L_2] = [r_{12}, p \otimes 1 + 1 \otimes p] + \sum_{\alpha simple} n_\alpha e^{\alpha(q)} [r_{12}, (E_\alpha + E_{-\alpha}) \otimes 1 + 1 \otimes (E_\alpha + E_{-\alpha})].$$

Since r_{12} is independent of p and q, one can identify the coefficients of p and q in these equations, and we find for the r-matrix, the conditions:

$$\begin{aligned} [r_{12}, H_i \otimes 1 + 1 \otimes H_i] &= 0, \\ [r_{12}, (E_\alpha + E_{-\alpha}) \otimes 1 + 1 \otimes (E_\alpha + E_{-\alpha})] &= H_\alpha \otimes (E_\alpha + E_{-\alpha}) - (E_\alpha + E_{-\alpha}) \otimes H_\alpha \end{aligned}$$

when α is a simple root. Taking the commutator of the second relation with $H \otimes 1 + 1 \otimes H$, and using the first one, we see that it splits into two independent equations involving E_α and $E_{-\alpha}$ separately i.e:

$$\begin{aligned} [r_{12}, H_i \otimes 1 + 1 \otimes H_i] &= 0 \\ [r_{12}, E_{\pm\alpha} \otimes 1 + 1 \otimes E_{\pm\alpha}] &= H_\alpha \otimes E_{\pm\alpha} - E_{\pm\alpha} \otimes H_\alpha \quad (\alpha \text{ simple}) \end{aligned} \tag{3.18}$$

To solve these equations, we notice that the Casimir element C in $\mathcal{G} \otimes \mathcal{G}$ obeys $[C, 1 \otimes X + X \otimes 1] = 0$ for all $X \in \mathcal{G}$. Therefore, eq.(3.18) tell us that r is determined only up to the addition of a multiple of C. Next we observe that if r_{12} is a solution, then $-r_{21}$ is another solution, hence $(r_{12} - r_{21})/2$ is an antisymmetric solution. So we can assume without loss of generality that r_{12} is antisymmetric. Note that the Casimir element is symmetric, and therefore does not affect such a solution. To find it, we dualize the equations using $R(X) = \mathrm{Tr}_2\,(r_{12} \cdot 1 \otimes X)$. Denoting $(X, Y) = \mathrm{Tr}\,(XY)$, we obtain:

$$[R(X), H_i] - R([X, H_i]) = 0 \tag{3.19}$$

$$[R(X), E_{\pm\alpha}] - R\,([X, E_{\pm\alpha}]) = (E_{\pm\alpha}, X) H_\alpha - (H_\alpha, X) E_{\pm\alpha} \quad (\alpha \text{ simple}) \tag{3.20}$$

The first equation implies $R(\mathcal{H}) \subset \mathcal{H}$ and $R(\mathcal{G}_{\pm\beta}) \subset \mathcal{G}_{\pm\beta}$. The most general form of r_{12} compatible with these requirements and the antisymmetry property is:

$$r = \sum_{ij} r^{ij} H_i \otimes H_j + \sum_{\beta \text{ positive}} \frac{r_\beta}{(E_\beta, E_{-\beta})} (E_\beta \otimes E_{-\beta} - E_{-\beta} \otimes E_\beta).$$

If $X \in \mathcal{H}$, eq(3.20) reduces to

$$\pm\alpha(X)[r_\alpha - 1] E_{\pm\alpha} = \alpha(R(X)) E_{\pm\alpha}, \quad (\alpha \text{ simple})$$

which implies $R(X) = 0$ for all $X \in \mathcal{H}$, i.e. $r^{ij} = 0$, and $r_\alpha = 1$ for α a simple root. Finally for β any positive root, we have $R(E_\beta) = r_\beta E_\beta$, and eq(3.20) gives

$$r_\beta [E_\beta, E_\alpha] - r_{\beta+\alpha}[E_\beta, E_\alpha] = 0.$$

Therefore , we deduce that $r_\beta = 1$, for all positive roots β. Gathering all these information, we obtain:

$$R(H) = 0 \tag{3.21}$$

$$R(E_{\pm\beta}) = \pm E_{\pm\beta} \quad (\beta \text{ positive}) \tag{3.22}$$

which is equivalent to eq.(3.17). ■

This result proves the integrability of the Toda chains models. We have found a constant antisymmetric r–matrix, and we now show how it fits into the general framework explained above.

Proposition. *Define* $R(X) = \mathrm{Tr}_2\,(r_{12} \cdot 1 \otimes X)$, *it satisfies the modified Yang–Baxter equation:*

$$[R(X), R(Y)] - R([X,Y]_R) = -[X,Y] \tag{3.23}$$

Proof. This is checked by a case by case analysis, using equations (3.21,3.22). ■

The antisymmetric r-matrix $r = \frac{1}{2} r_{12}$ is a solution of the modified Yang-Baxter equation (3.6). Hence $r^{\pm}_{12} = \frac{1}{2}(r_{12} \pm C_{12})$ (C_{12} is the Casimir) are solutions of the classical Yang–Baxter equations:

$$[r^{\pm}_{12}, r^{\pm}_{13}] + [r^{\pm}_{12}, r^{\pm}_{23}] + [r^{\pm}_{13}, r^{\pm}_{23}] = 0$$

For simple Lie algebras, C_{12} reads:

$$C_{12} = \sum_i H_i \otimes H_i + \sum_{\alpha \text{ positive}} \frac{E_\alpha \otimes E_{-\alpha} + E_{-\alpha} \otimes E_\alpha}{(E_\alpha, E_{-\alpha})}$$

Thus, the matrices $r^{\pm}_{12}$ are given explicitly by:

$$r^{+}_{12} = \frac{1}{2} \sum_i H_i \otimes H_i + \sum_{\alpha \text{ positive}} \frac{E_\alpha \otimes E_{-\alpha}}{(E_\alpha, E_{-\alpha})} \tag{3.24}$$

$$r^{-}_{12} = -\frac{1}{2} \sum_i H_i \otimes H_i - \sum_{\alpha \text{ positive}} \frac{E_{-\alpha} \otimes E_\alpha}{(E_\alpha, E_{-\alpha})} \tag{3.25}$$

We illustrate the general results concerning the algebraic meaning of the classical Yang–Baxter equation in the case of the Toda r–matrix. The action of $R^{\pm}$ on the generators of $\mathcal{G}$ is given by

$$\begin{cases} R^+(H) = \frac{1}{2} H \\ R^+(E_\alpha) = E_\alpha \\ R^+(E_{-\alpha}) = 0 \end{cases} ; \quad \begin{cases} R^-(H) = -\frac{1}{2} H \\ R^-(E_\alpha) = 0 \\ R^-(E_{-\alpha}) = -E_{-\alpha} \end{cases} \tag{3.26}$$

for α any positive root. Therefore the subalgebras $\mathcal{G}_\pm = \mathrm{Im}\, R^{\pm}$, are just the two standard Borel subalgebras of $\mathcal{G}$, $\mathcal{G}_\pm = \mathcal{B}_\pm = \mathcal{H} \oplus \mathcal{N}_\pm$. The subalgebras $\mathcal{K}_\pm = \mathrm{Ker}\, R^{\mp}$ are $\mathcal{K}_\pm = \mathcal{N}_\pm$. Therefore $\mathcal{G}_\pm/\mathcal{K}_\pm = \mathcal{H}$, and the isomorphism $\theta : \mathcal{H} \to \mathcal{H}$ is just $\theta(H) = -H;\ H \in \mathcal{H}$. We see that in this example the isomorphism θ is non trivial. The image of $\mathcal{G}_R$ in $\mathcal{B}_+ \oplus \mathcal{B}_-$ consists in the couples (X_+, X_-) such that the components of X_+ and X_- on the Cartan subalgebra $\mathcal{H}$ are opposite. Any element $X \in \mathcal{G}$ can be decomposed uniquely as:

$$X = X_+ - X_- \quad \text{with} \quad X_\pm \in \mathcal{B}_\pm, \quad X_+|_{\mathcal{H}} = -X_-|_{\mathcal{H}}$$

In the simple Lie group G, the corresponding decomposition is a factorization problem, $g = g_-^{-1} g^+$ where $g_\pm \in \exp(\mathcal{B}_\pm)$ and the factors of $g_\pm$ on the Cartan torus are inverse to each other.

3.9 Solution of the open Toda chain.

To solve the equations of motion, we have to solve the factorization problem eq.(3.15). Namely,

$$\exp(-tdh(L_0)) = g_-(t)^{-1}g_+(t); \quad L(t) = g_+(t)L_0g_+^{-1}(t) = g_-(t)L_0g_-^{-1}(t) \tag{3.27}$$

with L_0 the Lax matrix at time $t=0$, and the factorization problem is specified by the r–matrix of the Toda chain, i.e. it is a Gauss decomposition with the diagonal part equally shared between the two factors. Even though this is a purely algebraic problem, it may not be an easy task to perform explicitly. Fortunately if $\mathcal{G}$ is a finite dimensional simple Lie algebra, there exists a suitable parametrization of the Lax operator, found by Kostant, in which the factorization can be done. We will explain this method in the case of $\mathcal{G} = sl(N+1)$.

We write the Lax matrix of the Toda chain as:

$$L_{\text{Toda}} = e^{-ad\, q}\mathcal{E}_- + p + e^{ad\, q}\mathcal{E}_+, \quad \text{with} \quad \mathcal{E}_\pm = \sum_{\alpha \text{ simple}} n_\alpha E_{\pm\alpha}$$

Remark that if we assign degree 0 to the generators of the Cartan subalgebra and degree 1 or -1 to the step generators E_α or $E_{-\alpha}$ with α simple, then $\mathcal{E}_-$ is of degree -1, p of degree 0 and $\mathcal{E}_+$ of $+1$.

The Lax matrix is a symmetric Jacobi matrix. We will parametrize the space of symmetric $(N+1)\times(N+1)$ Jacobi matrices, which is of dimension $2N$, with two diagonal matrices: $\Omega \in sl(N+1)$, and $\theta \in SL(N+1)$. To this end, we introduce the matrix w which is the unique matrix such that $\det w = 1$, $w\mathcal{H}w^{-1} \in \mathcal{H}$ and

$$wE_{i,i+1}w^{-1} = E_{N+2-i,N+1-i} \tag{3.28}$$

Explicitly, its matrix elements are $[w]_{ij} = (-1)^{\frac{N}{2}}\delta_{N+2-i,j}$. The action eq.(3.28) represents the action of the longest element in the Weyl group. Its main property in what follows is the relation

$$w^{-1}\mathcal{E}_-\, w = \mathcal{E}_+ \tag{3.29}$$

Lemma. *Let Ω be a traceless diagonal matrix, and θ be a diagonal matrix with determinant 1. There exists a unique element $u \in N_- = \exp(\mathcal{N}_-)$ which diagonalizes $\mathcal{E}_- + \Omega$, i.e.*

$$\mathcal{E}_- + \Omega = u\Omega u^{-1} \tag{3.30}$$

Introduce the unique solution $b_\pm \in B_\pm = \exp(\mathcal{B}_\pm)$ of the factorization problem

$$b_-^{-1}b_+ = w^{-1}u\theta u^{-1}$$

where we require that the diagonal is equally shared between $b_\pm$ as in the previous section. Then we have

$$b_+\left(\mathcal{E}_- + \Omega\right)b_+^{-1} = b_-\left(\mathcal{E}_+ + w^{-1}\Omega\, w\right)b_-^{-1}$$

Proof. From the factorization condition, we have $b_+ = b_-w^{-1}u\theta u^{-1}$. Then

$$\begin{aligned} b_+\left(\mathcal{E}_- + \Omega\right)b_+^{-1} &= b_-w^{-1}u\theta u^{-1}\left(\mathcal{E}_- + \Omega\right)u\theta^{-1}u^{-1}wb_-^{-1} \\ &= b_-w^{-1}\left(\mathcal{E}_- + \Omega\right)wb_-^{-1} = b_-\left(\mathcal{E}_+ + w^{-1}\Omega\, w\right)b_-^{-1} \end{aligned}$$

because $u\theta u^{-1}$ commutes with $\mathcal{E}_- + \Omega = u\Omega u^{-1}$. ∎

Note that the factorization problem is non trivial due to the presence of w which does not belong to N_-. As a straightforward consequence of this Lemma, we have the:

Proposition. *Choose Ω and θ as above, then the matrix L defined by*

$$L = b_+ (\mathcal{E}_- + \Omega) b_+^{-1} = b_- \left(\mathcal{E}_+ + w^{-1}\Omega\, w\right) b_-^{-1} \tag{3.31}$$

is a symmetric traceless Jacobi matrix.

Proof. The first form of L shows that L expands on degrees ≥ -1, while the second form shows that it expands on degrees ≤ 1, hence L is a Jacobi matrix. To show that it is symmetric write $b_\pm = h^{\pm 1} n_\pm$ with $n_\pm \in N_\pm$ and h in the Cartan torus. We implemented the condition that $b_\pm$ have inverse components on the Cartan torus. We write $h = \exp(-q)$ for $q \in \mathcal{H}$. The term in $E_{-\alpha}$ in L is given by $\exp(-\mathrm{ad}\, q)E_{-\alpha} = \exp(\alpha(q))E_{-\alpha}$. Similarly, the term in E_α is given by $\exp(\mathrm{ad}\, q)E_\alpha = \exp(\alpha(q))E_\alpha$. The two coefficients are equal. Finally since Ω is traceless, so is L. ■

The fact that we have at our disposal $2N$ parameters in Ω and θ indicates that we can parametrize all symmetric traceless Jacobi matrices in this way. The diagonal matrices Ω, θ form a system of coordinates on the space of symmetric Jacobi matrices.

In the $sl(N+1)$ case we have the useful formulae

$$\frac{h_{i+1}^2}{h_i^2} = e^{2(q_i - q_{i+1})} \tag{3.32}$$

Moreover, denoting $\Omega = \mathrm{diag}\,(\omega_i)$, we have explicitely

$$u_{ij} = \prod_{l=j+1}^{i} \frac{1}{\omega_j - \omega_l}, \quad u_{ij}^{-1} = \prod_{l=j}^{i-1} \frac{1}{\omega_i - \omega_l} \tag{3.33}$$

This is obtained by writing $u\ \Omega = (\mathcal{E}_- + \Omega)\ u$ or $(\omega_j - \omega_i)u_{ij} = u_{i-1,j}$. The solution of this recurrence relation with the boundary condition $u_{ii} = 1$ yields eq(3.33). To compute u^{-1} we proceed in the same way with the equation $\Omega u^{-1} = u^{-1}(\mathcal{E}_- + \Omega)$.

In the case of the open Toda chain, the Lax matrix $L(t)$ will be parametrized by two matrices $\Omega(t)$ and $\theta(t)$. We call $L_0 = L(t=0)$ the Lax matrix at time $t = 0$, and let Ω_0 and θ_0 denote its coordinates. Remark that the matrix elements of $\Omega(t)$ are the eigenvalues of $L(t)$, due to eqs.(3.30,3.31), so are *constant* in time. Hence $\Omega(t) = \Omega_0$. Similarly, the matrix $u(t) = u$ is independent of time because it depends only on the matrix elements of $\Omega(t)$.

Lemma. *In the above coordinates we have*

$$L(t) = g_+(t) L_0 g_+^{-1}(t) = g_-(t) L_0 g_-^{-1}(t), \quad \text{with} \quad g_+(t) = b_+(t) b_+^{-1}(0), \quad g_-(t) = b_-(t) b_-^{-1}(0) \tag{3.34}$$

Moreover $g_\pm$ have inverse components on the Cartan torus.

Proof. The result follows by writing in eq.(3.31) $(\mathcal{E}_- + \Omega_0) = b_+^{-1}(0) L_0 b_+(0)$ and $\left(\mathcal{E}_+ + w^{-1}\Omega_0 w\right) = b_-^{-1}(0) L_0 b_-(0)$. The components of $g_\pm$ on the Cartan torus are $(h(t)h^{-1}(0))^{\pm 1}$ and are clearly inverse to each other. ■

We now find the time evolution of the coordinate $\theta(t)$ by requiring that eq.(3.27) holds with the above $g_\pm(t)$. This yields the:

Proposition. *The time evolution of the Toda chain, in the coordinates* $\Omega(t), \theta(t)$, *reads*

$$\Omega(t) = \Omega_0, \quad \theta(t) = \theta_0 e^{-tdh(\Omega_0)} \tag{3.35}$$

Proof. The first assertion has already been proved. To prove the second one, we start from $\exp(-tdh(L_0)) = g_-(t)^{-1}g_+(t)$. The left hand side is easy to evaluate:

$$\exp(-tdh(L_0)) = b_+(0)u\exp(-tdh(\Omega_0))u^{-1}b_+^{-1}(0) \tag{3.36}$$

On the other hand using the explicit form of $g_\pm(t)$, the right hand side is equal to:

$$g_-^{-1}(t)g_+(t) = b_-(0)b_-^{-1}(t)b_+(t)b_+^{-1}(0) = b_-(0)w^{-1}u\theta(t)u^{-1}b_+^{-1}(0)$$

where in the second expression we used the factorization problem at time t for $b_\pm(t)$. However at $t=0$ the same problem yields $b_-(0) = b_+(0)u\theta_0^{-1}u^{-1}w$. Plugging into the previous expression we get:

$$g_-^{-1}(t)g_+(t) = b_+(0)u\theta_0^{-1}\theta(t)u^{-1}b_+^{-1}(0)$$

Comparing with eq.(3.36) we get $\theta_0^{-1}\theta(t) = \exp(-tdh(\Omega_0))$. ■

We can now give a compact form of the solution of the equations of motion in terms of determinants. To do that, consider the matrix $M = b_-^{-1}b_+ = n_-^{-1}h^2n_+ = w^{-1}u\theta u^{-1}$. We want to compute directly h^2. If we have a representation V with a highest weight vector $|v\rangle$ and its dual V^* with the dual highest weight vector $\langle v|$ we have $\langle v|M|v\rangle = \langle v|h^2|v\rangle$, because by definition $n_+|v\rangle = |v\rangle$ and $\langle v|n_- = \langle v|$. Since h lives in a space of dimension $r = \operatorname{rank} G$, using r such representations with independent highest weights, one completely reconstructs h. In the $sl(N+1)$ case we can take the N representations $\Lambda^k\mathbb{C}^{N+1}$, $k = 1,\cdots N$. Let $|\epsilon_1\rangle,\cdots,|\epsilon_{N+1}\rangle$ be the canonical basis of $\mathbb{C}^{N+1}$. Then the highest weight vector $|v_k\rangle$ and its dual $\langle v_k|$ are:

$$|v_k\rangle = |\epsilon_1\rangle\wedge\cdots\wedge|\epsilon_k\rangle, \quad \langle v_k| = \langle\epsilon_1|\wedge\cdots\wedge\langle\epsilon_k|$$

Since by definition, in the representation $\Lambda^k\mathbb{C}^{N+1}$, $M|v_k\rangle = M|\epsilon_1\rangle\wedge\cdots\wedge M|\epsilon_k\rangle$ it is easy to check that $n_+|v_k\rangle = |v_k\rangle$. Moreover, we have for any matrix M:

$$\langle v_k|M|v_k\rangle = \det\begin{pmatrix} m_{11} & \cdots & m_{1k} \\ \vdots & & \vdots \\ m_{k1} & \cdots & m_{kk}\end{pmatrix}$$

Proposition. *Let* $|v_k\rangle$ *be the highest weight vectors of* $sl(N+1)$ *defined above. Let* $M = w^{-1}u\theta(t)u^{-1}$ *and define the* N *functions of* t

$$\tau_k = \langle v_k|M|v_k\rangle = \det\begin{pmatrix} m_{11} & \cdots & m_{1k} \\ \vdots & & \vdots \\ m_{k1} & \cdots & m_{kk}\end{pmatrix}, \quad \text{with } M = (m_{ij})$$

then the solution of the Toda chain is given by

$$e^{2(q_i - q_{i+1})} = \frac{\tau_{i+1}\tau_{i-1}}{\tau_i^2} \tag{3.37}$$

Proof. Let h_i be the matrix elements of the diagonal $(N+1)\times(N+1)$ matrix h in $b_-^{-1}b_+ = n_-^{-1}h^2n_+$. The Toda position coordinates are given by eq.(3.32), $\exp 2(q_i - q_{i+1}) = h_{i+1}^2/h_i^2$. According to the previous analysis we have $\tau_k = \langle v_k|M|v_k\rangle = \langle v_k|h^2|v_k\rangle$, because $M = n_-^{-1}h^2n_+$, so that $\tau_k = h_1^2\cdots h_k^2$. Setting $\tau_0 = \tau_{N+1} = 1$ by convention, we deduce

$$h_i^2 = \frac{\tau_i}{\tau_{i-1}}$$

which proves the result. ■

Let us summarise the way to find the solution: given the initial data L_0 one computes the coordinates (Ω_0, θ_0) and we let these coordinates evolve according to eq.(3.35). Then, one computes u and evaluate the τ_i by the determinant formula. Finally, the Toda coordinates q_i are given by (3.37). All these steps are simple algebraic computations.

3.10 Toda system and Hamiltonian reduction.

A quite general and efficient way to produce integrable systems is to start from a free motion on a symplectic manifold with group action, and perform a suitable Hamiltonian reduction, see Chapter [14]. While this approach is efficient to solve models arising from finite–dimensional Lie groups (with Lax pair without spectral parameter), it is less effective when one has to start from a loop group. We apply this strategy to recover the Toda chain starting from the geodesic motion on a Lie group G.

Geodesics on the group G correspond to left translations of 1-parameter groups (the tangent vector is transported parallel to itself), therefore $\frac{d}{dt}(g^{-1}\dot{g}) = 0$. The solution is of the form $g(t) = g(0)\exp(Xt)$, for X in the Lie algebra of G. Of course this leaves open the problem of effectively computing the exponential, which was the main concern of the previous section.

Let G be a complex simple Lie group with Lie algebra $\mathcal{G}$. Let $\{H_i\}$ be the generators of a Cartan subalgebra $\mathcal{H}$, and let $\{E_{\pm\alpha}\}$ be the corresponding root vectors, chosen to form a Weyl basis, i.e. all the structure constants are real. The real normal form of $\mathcal{G}$ is the real Lie algebra $\mathcal{G}_0$ generated over $\mathbb{R}$ by the H_i and $E_{\pm\alpha}$. Let σ be the Cartan involution: $\sigma(H_i) = -H_i$, $\sigma(E_{\pm\alpha}) = -E_{\mp\alpha}$. The fixed points of σ form a Lie subalgebra $\mathcal{K}$ of $\mathcal{G}_0$ generated by $\{E_\alpha - E_{-\alpha}\}$. We have the decomposition:

$$\mathcal{G}_0 = \mathcal{K} \oplus \mathcal{M}$$

where $\mathcal{M}$ is the real vector space generated by the $\{E_\alpha + E_{-\alpha}\}$ and the $\{H_i\}$. Notice that due to the choice of the real normal form, $\mathcal{A} = \mathcal{H} \cap \mathcal{G}_0$ is a maximal abelian subalgebra of $\mathcal{G}_0$ and it is entirely contained in $\mathcal{M}$. Finally we need the real nilpotent subalgebras $\mathcal{N}_\pm$ generated respectively by the $\{E_{\pm\alpha}\}$.

Let G_0 be the connected Lie group corresponding to the Lie algebra $\mathcal{G}_0$, and similarly K corresponding to $\mathcal{K}$ and $N_\pm$ corresponding to $\mathcal{N}_\pm$. Notice that G_0/K is a symmetric space of the non–compact type. Finally the Cartan algebra $\mathcal{A}$ exponentiates to A.

The connected Lie group G_0 admits the following Iwasawa decomposition:

$$G_0 \simeq N_+ \times A \times K \text{ as a manifold}$$

that is any element g in G_0 can be written *uniquely* $g = nQk$. For example consider the case $G_0 = SL(N+1, \mathbb{R})$ and let $\sigma : g \to^t g^{-1}$, be the canonical automorphism of G. At the Lie algebra level, σ reads $\sigma(X) = -^tX$. The subgroup K is the group of orthogonal matrices with determinant 1, and the symmetric space G_0/K can be viewed as the space of symmetric matrices, with determinant 1. Finally A is the set of diagonal matrices, and N_+ is the group of upper triangular matrices with 1 on the diagonal. To show that any g can be written as nQk, recall the Gram–Schmidt orthogonalization procedure. Starting from a basis $(f_1, \cdots, f_{N+1})$ one constructs inductively an orthogonal basis by setting $u_1 = f_1$, $u_2 = f_2 - f_1\ (f_2, f_1)/(f_1, f_1)$, etc. , that is $(u_1, \cdots, u_{N+1}) = (f_1, \cdots, f_{N+1})n$ with $n \in N_+$. One then normalizes this basis by setting $(u'_1, \cdots, u'_{N+1}) = (u_1, \cdots, u_{N+1})Q$, with Q the diagonal matrix with entries $Q_j = 1/\sqrt{(u_j, u_j)}$. We apply this procedure to the basis f_j of columns of g^{-1}. Then we have $(f_1, \cdots, f_{N+1})g = (e_1, \cdots, e_{N+1})$ where e_j is the canonical basis of $\mathbb{R}^{N+1}$, while, on the other hand:

$$(f_1, \cdots, f_{N+1})n = (u'_1, \cdots, u'_{N+1}),\ (u_1, \cdots, u_{N+1})Q = (u'_1, \cdots, u'_{N+1}),\ (u'_1, \cdots, u'_{N+1})k = (e_1, \cdots, e_{N+1})$$

for a unique orthogonal matrix k. Comparing we get the unique decomposition $g = nQk$.

To get the Toda chain we shall perform the reduction of the geodesic motion on T^*G_0 by the action of the group N_+ on the left and K on the right.

We first recall some generalities about the symplectic structure on the cotangent bundle of a Lie group, see Chapter [14]. The tangent space of G at $g \in G$ consists of vectors $g \cdot X$ where $X \in \mathcal{G}$, the Lie algebra of G. The cotangent bundle $N = T^*G$ is identified to $G \times \mathcal{G}^*$ by $p \in T^*_g(G) \longrightarrow (g, \xi)$ where $p(g \cdot X) = \xi(X)$. Using the invariant bilinear form on $\mathcal{G}$, we can identify $\mathcal{G}$ and $\mathcal{G}^*$, so that $\xi(X) = (\xi, X)$.

If $(\delta g, \delta\xi)$ is a vector tangent to T^*G at the point (g, ξ), the canonical 1-form is given by:

$$\alpha\,(\delta g, \delta\xi) = -\xi\,(g^{-1} \cdot \delta g) \tag{3.38}$$

thereby providing the symplectic form $\omega = d\alpha$. Notice that α and ω are invariant under both left and right translations.

One can then attempt to reduce this phase space using Lie subgroups H_L and H_R of G with Lie algebras $\mathcal{H}_L$ and $\mathcal{H}_R$, acting respectively on the left and on the right on T^*G by:

$$(\ (h_L, h_R), (g, \xi)\ ) \rightarrow (h_L g h_R^{-1}, h_R \xi h_R^{-1})$$

Since we are interested in the computation of Poisson brackets of functions on phase space, it is important to remember that, if we have a group action on a manifold M, $(g, m) \rightarrow gm$, the action on functions on M reads $f \rightarrow g \cdot f$ where $g \cdot f(m) = f(g^{-1}m)$. In our case we can take as elementary functions the collection of the matrix elements of g in a faithful representation of G, and the components of ξ. Any other function can be expressed on these coordinates.

So it is enough to give the Poisson brackets of these elementary functions. The general theory, see Chapter [14], shows that these Poisson brackets are expressed as:

$$\{\xi(X), \xi(Y)\} = \xi([X, Y]), \quad \{\xi(X), g\} = g\,X, \quad \{g, g\} = 0 \tag{3.39}$$

In these equations g is an abuse of notation for the matrix $\rho(g)$ in some faithful representation. So eq.(3.39) really means $\{\xi(X), \rho(g)_{ij}\} = \rho(gX)_{ij})$, and $\{\rho(g)_{ij}, \rho(g)_{kl}\} = 0$. On these elementary *functions* g and $\xi(Y)$ the infinitesimal group action $\delta_{(X_L,X_R)}$ reads:

$$\delta_{(X_L,X_R)} g = -X_L g + g X_R, \quad \delta_{(X_L,X_R)} \xi(Y) = -([X_R, \xi], Y)$$

It is straightforward to check, using the Poisson brackets (3.39), that the generating Hamiltonians are:

$$H_{X_L} = -\xi(g^{-1} X_L g), \quad H_{X_R} = \xi(X_R)$$

where we recall that for any function f, we have by definition $\delta_X f = \{H_X, f\}$. This means that the moments are:

$$\mathcal{P}^L(g, \xi) = -P_{\mathcal{H}^*_L}\, g\xi g^{-1} \quad \mathcal{P}^R(g, \xi) = P_{\mathcal{H}^*_R}\, \xi \tag{3.40}$$

where we have introduced the projectors on $\mathcal{H}^*_{L,R}$ of forms in $\mathcal{G}^*$ induced by the restriction of these forms to $\mathcal{H}_{L,R}$, the Lie algebras of $H_{L,R}$.

The geodesic motion on G is a Hamiltonian system on T^*G whose Hamiltonian is: $H = \frac{1}{2}\,(\xi, \xi)$. Indeed one finds using eq.(3.39), $\dot{\xi} = 0$ and $g^{-1}\dot{g} = \xi$ as equations of motion. Notice that H is bi–invariant, so one can reduce by the action of the group N_+ on the left and K on the right. This is achieved by a suitable choice of the momentum. We take:

$$P_{\mathcal{K}^*}(\xi) = \mu^R = 0 \tag{3.41}$$

$$-P_{\mathcal{N}_+^*}(g\xi g^{-1}) = \mu^L = -\sum_{\alpha \text{ simple}} E_{-\alpha} \tag{3.42}$$

In the following we shall identify $\mathcal{N}_+^*$ with $\mathcal{N}_-$, and $\mathcal{K}^*$ with $\mathcal{K}$ through the invariant scalar product.

The isotropy group G_μ is $N_+ \times K$. This is obvious for the right component since $\mu^R = 0$. The isotropy group of μ^L is by definition the set of elements $g \in N_+$ such that:

$$(\mu^L, g^{-1}Xg) = (\mu^L, X) \; \forall X \in \mathcal{N}_+$$

Since μ^L only contains roots of degree -1, the only contribution to (μ^L, X) comes from $X^{(1)}$, the degree one component of X. But $(g^{-1}Xg)^{(1)} = X^{(1)} \; \forall g \in N_+$. This is because, for $Y \in \mathcal{N}_+$, $\exp(Y)X\exp(-Y) = X + [Y, X] + \cdots$, and Y is of degree ≥ 1. Hence the isotropy group of μ^L is N_+ itself.

Let us first compute the dimension of the reduced phase space. Let $d = \dim \mathcal{G}_0$ and $r = \dim \mathcal{A}$. Then we have:

$$\dim \mathcal{K} = \dim \mathcal{N}_+ = \frac{d-r}{2} \qquad \dim T^*G_0 = 2d$$

The dimension of the submanifold N_μ defined by the equations (3.41, 3.42) is $\dim N_\mu = 2d - \dim \mathcal{K} - \dim \mathcal{N}_+ = d + r$ and the dimension of the reduced phase space is $\dim \mathcal{F}_\mu = \dim N_\mu - \dim \mathcal{K} - \dim \mathcal{N}_+ = 2r$ which is the correct dimension of the phase space of the Toda chain.

Since the isotropy group is the whole of $N_+ \times K$ any point (g, ξ) of N_μ can be brought to the form (Q, L) with $Q \in A$ (due to the Iwasawa decomposition) by the action of the isotropy group. We shall identify $\mathcal{G}_0$ and $\mathcal{G}_0^*$ under the Killing form for convenience. Equation (3.41) implies that $L \in \mathcal{M}$ which is the orthogonal of $\mathcal{K}$. Thus we can write:

$$L = p + \sum_{\alpha \text{ positive}} l_\alpha (E_\alpha + E_{-\alpha}), \quad p \in \mathcal{A}$$

Inserting this form into equation (3.42) and setting $Q = \exp(q)$ we get:

$$P_{\mathcal{N}_-} QLQ^{-1} = \sum_{\alpha > 0} l_\alpha e^{-\alpha(q)} E_{-\alpha} = \sum_{\alpha \text{ simple}} E_{-\alpha}$$

hence $l_\alpha = \exp(\alpha(q))$ for α simple and $l_\alpha = 0$ otherwise. We have obtained the Lax matrix of the Toda chain:

$$L = p + \sum_{\alpha \text{ simple}} e^{\alpha(q)} (E_\alpha + E_{-\alpha}) \tag{3.43}$$

The set of the (Q, L) is obviously a $(2r)$–dimensional subvariety $\mathcal{S}$ of N_μ. For any point (g, ξ) in N_μ one can write uniquely $g = nQk$ and $\xi = k^{-1}Lk$.

One can compute the reduced symplectic form in the coordinates (Q, L). It is enough to evaluate the canonical 1–form:

$$\alpha = \xi(g^{-1}\delta g) = L(Q^{-1}\delta Q) = L(\delta q) = \sum_i p_i \delta q_i$$

This completes the identification of the reduced geodesic flow with the open Toda chain.

A nice feature of this approach is that it allows to compute naturally the r–matrix of the Toda model. We refer to Chapter [14] for the exposition of the general method. The function $L(X)$ (for any $X \in \mathcal{M}$) on the reduced phase space has a uniquely defined extension on T^*G_0, invariant under the action of the group $N_+ \times K$:

$$F_X(g, \xi) = (\xi, k^{-1}Xk) \tag{3.44}$$

where $k = k(g)$ is uniquely determined by the Iwasawa decomposition $g = nQk$. In this situation one has simply:

$$\{L(X), L(Y)\} = \{F_X, F_Y\} = \{(k\xi k^{-1}, X), (k\xi k^{-1}, Y)\}$$

We evaluate this with the help of equations (3.39), and then restrict to $n = k = e$.

For a function $f : G \to G$ we have:

$$\{\xi(X), f(g)\} = \frac{d}{dt} f(ge^{tX})|_{t=0}$$

Applying this to the function $k = k(g)$ we get:

$$\{\xi(X), k(g)\} = \nabla k(X), \quad \text{where } \nabla k(X) = \frac{\partial}{\partial t} k(Q\exp(tX))_{|t=0}$$

We then immediately get:

$$\{L(X), L(Y)\} = \Big(L\,,\,[X,Y] - [X, \nabla k(Y)] - [\nabla k(X), Y]\Big)$$

Notice that $[X, Y] \in \mathcal{K}$ because $X, Y \in \mathcal{M}$, hence $L([X, Y]) = 0$, because $L \in \mathcal{M}$. From this equation follows immediately an R–matrix structure for the Toda system given by:

$$R'X = -\nabla k(X) \in \mathcal{K}$$

We can compute the derivatives as follows: due to the Iwasawa decomposition we have uniquely

$$Qe^{tX} = e^{tX_+} \cdot Qe^{tX_a} \cdot e^{tX_K} \Rightarrow QX = X_+Q + X_aQ + QX_K, \quad X_+ \in \mathcal{N}_+,\ X_a \in \mathcal{A},\ X_K \in \mathcal{K}$$

and $\nabla k(X) = X_K$. Setting $X = \sum_{\alpha>0} x_\alpha(E_\alpha + E_{-\alpha}) + \sum_i x_i H_i = 2\sum_{\alpha>0} x_\alpha E_\alpha + \sum_i x_i H_i - \sum_{\alpha>0} x_\alpha(E_\alpha - E_{-\alpha})$ and multiplying by Q^{-1} on the left, we get, noticing that $Q^{-1}X_+Q \in \mathcal{N}_+$, $X_K = -\sum_{\alpha>0} x_\alpha(E_\alpha - E_{-\alpha})$.

In the dualized formalism defined by $R'X = \mathrm{Tr}_2(r'_{12}.\mathbf{1} \otimes X)$ this r–matrix reads:

$$r'_{12} = \frac{1}{2}\sum_{\alpha>0} \frac{(E_\alpha - E_{-\alpha}) \otimes (E_\alpha + E_{-\alpha})}{(E_\alpha, E_{-\alpha})} \tag{3.45}$$

We want to relate this to the r–matrix we computed in eq.(3.17), that is:

$$r_{12} = \sum_{\alpha>0} \frac{(E_\alpha \otimes E_{-\alpha} - E_{-\alpha} \otimes E_\alpha)}{(E_\alpha, E_{-\alpha})}$$

For this, we recall that L and r_{12} satisfy the equation $\{L_1, L_2\} = [r_{12}, L_1] - [r_{21}, L_2]$. Applying the automorphism $\sigma_1 = \sigma \otimes 1$ to this equation and noting that $\sigma(L) = -L$, we get $\{L_1, L_2\} = [\sigma_1 r_{12}, L_1] + [\sigma_1 r_{21}, L_2]$. Adding the two equations we get a similar r–matrix relation:

$$\{L_1, L_2\} = [\frac{1}{2}(r_{12} + \sigma_1 r_{12}), L_1] - [\frac{1}{2}(r_{21} - \sigma_1 r_{21}), L_2]$$

It is straightforward to check that:

$$\frac{1}{2}(r_{12} + \sigma_1 r_{12}) = r'_{12}, \quad \frac{1}{2}(r_{21} - \sigma_1 r_{21}) = r'_{21}$$

hence we recover eq.(3.45).

References

[1] M. Toda *"Vibration of a chain with non-linear interaction"* J. Phys. Soc. Jpn 22 (1967), P. 431.

[2] B. Kostant. *"The solution to the generalized Toda lattice and representation theory"*. Adv. Math. 34 (1979), pp. 195-338.

[3] M. Adler. *"On a trace functional for formal pseudodifferential operators and symplectic structure of the Korteweg-de Vries type equations"*. Inv. Math. 50 (1979), p. 219.

[4] E. K. Sklyanin. *"On the complete integrability of the Landau-Lifchitz equation"*. Preprint LOMI E-3-79. Leningrad 1979.

[5] W. Symes. *"Systems of Toda type, inverse spectral problems and representation theory"*. Inv. Math. 59 (1980), p. 13.

[6] M. Semenov Tian Shansky. *"What is a classical r-matrix"*. Funct. Anal. and its Appl. 17, 4 (1983), p. 17.

[7] Olshanetsky Perelomov.

Chapter 4

Analytical methods.

In this chapter, we present the general ideas for solving the Lax equations with spectral parameter. The spectral curve is the characteristic equation for the eigenvalues of the Lax matrix $\det(L(\lambda) - \mu) = 0$. Since the Lax equation $\dot{L}(\lambda) = [L(\lambda), M(\lambda)]$ is isospectral, the eigenvalues of $L(\lambda)$ are time independent and so is the spectral curve. At any point of the spectral curve there exists, by definition, an eigenvector of $L(\lambda)$ with eigenvalue μ. This defines an analytic line bundle $\mathcal{L}$ on the curve with prescribed Chern class. Its time evolution reduces to a linear flow on the jacobian of the spectral curve. We give three proofs of this result. The first one, by explicit construction of a meromorphic section of this line bundle. The second one, uses an unusual type of functions on the Riemann surface, the Baker–Akhiezer functions. Finally, the linearization property also follows very directly from the group theoretical factorization method in its Riemann–Hilbert incarnation. As a result, one can express the dynamical variables in terms of θ functions defined on the Jacobian of the spectral curve.

4.1 The spectral curve.

Let us consider a $N \times N$ Lax matrix $L(\lambda)$, depending as in the previous Chapter [2], rationally on a spectral parameter $\lambda \in \mathbb{C}$ with poles at points λ_k

$$L(\lambda) = L_0 + \sum_k L_k \tag{4.1}$$

with L_0 independent of λ and L_k the polar part of $L(\lambda)$ at λ_k, ie. $L_k = \sum_{r=-n_k}^{-1} L_{k,r}(\lambda - \lambda_k)^r$.

The analytical method is based on the analysis of the eigenvector equation:

$$(L(\lambda) - \mu \mathbf{1})\, \Psi(\lambda, \mu) = 0 \tag{4.2}$$

$\Psi(\lambda, \mu)$ is the eigenvector with eigenvalue μ. The characteristic equation for the eigenvalue problem (4.2) is:

$$\Gamma \quad : \quad R(\lambda, \mu) \equiv \det(L(\lambda) - \mu\ \mathbf{1}) = 0 \tag{4.3}$$

This defines an algebraic curve in $\mathbb{C}^2$ which is called the spectral curve. We are considering here the smooth compact curve obtained from this equation by the desingularization procedure explained in Chapter [15], even if we do not mention it explicitly. A point on Γ is a pair (λ, μ) satisfying eq.(4.3). If N is the dimension of the Lax matrix, the equation of the curve is of the form:

$$\Gamma \quad : \quad R(\lambda, \mu) \equiv (-\mu)^N + \sum_{q=0}^{N-1} r_q(\lambda) \mu^q = 0 \tag{4.4}$$

The coefficients $r_q(\lambda)$ are polynomials in the matrix elements of $L(\lambda)$ and therefore have poles at λ_k. Since the Lax equation $\dot{L} = [L, M]$ is isospectral, these coefficients are time-independent and are related to the action variables.

From eq.(4.4), we see that the spectral curve appears as an N-sheeted covering of the Riemann sphere. To a given point λ on the Riemann sphere there correspond N points on the curve whose coordinates are $(\lambda, \mu_1), \cdots (\lambda, \mu_N)$ where the μ_i are the solutions of the algebraic equation $R(\lambda, \mu) = 0$. By definition μ_i are the eigenvalues of $L(\lambda)$.

The goal of the method is to determine the analytical properties of the eigenvector $\Psi(\lambda, \mu)$ and see how much of $L(\lambda)$ can be reconstructed from them. The result is that one can reconstruct $L(\lambda)$ up to global (independent of λ) similarity transformations. This is not too surprising since the analytical properties of $L(\lambda)$ and the spectral curve are invariant under *global* gauge transformations. So from analyticity we can only hope to recover the system where global gauge transformations have been factored away.

In general, we may fix the gauge by diagonalizing $L(\lambda)$ for one value of λ. To be specific, we choose to diagonalize at $\lambda = \infty$, i.e. we diagonalize the coefficient L_0.

$$L_0 = \lim_{\lambda \to \infty} L(\lambda) = \mathrm{diag}(a_1, \cdots, a_N) \tag{4.5}$$

We assume for simplicity that all the a_i's are different. Then on the spectral curve, we have N points $Q_i \equiv (\lambda = \infty, \mu_i = a_i)$ above $\lambda = \infty$. In the gauge (4.5) there remains a residual action which consists in conjugating the Lax matrix by constant *diagonal* matrices. Generically, these transformations form a group of dimension $N - 1$ and we will have to factor it out.

Before doing complex analysis on Γ, one has to determine its genus. A general strategy is as follows. As we have seen, Γ is a N-sheeted covering of the Riemann sphere. There is a general formula expressing the genus g of an N-sheeted covering of a Riemann surface of genus g_0 (in our case $g_0 = 0$). It is the Riemann-Hurwitz formula:

$$2g - 2 = N(2g_0 - 2) + \nu \tag{4.6}$$

where ν is the branching index of the covering. Let us assume for simplicity that the branch points are all of order 2. To compute ν we observe that this is the number of values of λ where $R(\lambda, \mu)$ has a double root in μ. This is also the number of zeroes of $\partial_\mu R(\lambda, \mu)$ on the surface $R(\lambda, \mu) = 0$. But $\partial_\mu R(\lambda, \mu)$ is a meromorphic function on Γ and therefore the number of its zeroes is equal to the number of its poles. Hence it is enough to count the poles. These poles can only be located where the matrix $L(\lambda)$ itself has a pole. So we are down to a local analysis around the points of Γ such that $L(\lambda)$ has a pole.

Let us apply this idea to the matrix (4.1). Above a pole λ_k we have N branches of the form $\mu_j = l_j/(\lambda - \lambda_k)^{n_k} + \cdots$ where l_j are the eigenvalues of $L_{k,-n_k}$ that are assumed all distinct. On such a branch we have $\partial_\mu R(\lambda, \mu)|_{(\lambda, \mu_j(\lambda))} = \prod_{i \neq j} \big(\mu_j(\lambda) - \mu_i(\lambda)\big)$ which thus has a pole of order $(N-1)n_k$. Summing on all branches the total order of the poles over λ_k is $N(N-1)n_k$. Summing on all poles λ_k of $L(\lambda)$ we see that the total branching index is $\nu = N(N-1)\sum_k n_k$. This gives:

$$g = \frac{N(N-1)}{2} \sum_k n_k - N + 1$$

For completeness of the method it is important to observe that the genus is equal to the number of action variables occuring as independent parameters in the eq.(4.4) which should also be half the dimension of phase space. Let us count this last number in the case where $L(\lambda)$ is given by eq.(4.1). As we have seen in Chapter [2], $L_k = \big(g_k \cdot (A_k)_- \cdot g_k^{-1}\big)_-$ where the $(A_k)_-$ characterize the orbit and are non–dynamical. The dynamical variables are the jets of order $(n_k - 1)$ of the g_k's which gives $N^2 n_k$ parameters. But L_k is invariant under $g_k \to g_k d_k$ with d_k a jet of diagonal matrices of the same order. Hence the dimension of the L_k orbit is $(N^2 - N)n_k$. Moreover we have the residual global gauge invariance by diagonal matrices acting as $g_k \to dg_k$, or $L(\lambda) \to dL(\lambda)d^{-1}$. Notice that the orbits of this action are of dimension $(N-1)$, since the identity does not act. One performs an hamiltonian reduction by this action. First one fixes the momentum, yielding $(N-1)$ conditions, and then one takes the quotient by the stabilizer of the momentum which is here the whole group since it is abelian. Hence the dimension of the phase space is reduced by $2(N-1)$, yielding:

$$\dim \mathcal{M} = (N^2 - N)\sum_k n_k - 2(N-1) = 2g$$

Let us now count the number of independent coefficients in eq.(4.4). It is clear that $r_j(\lambda)$ is a rational function of λ. The value of r_j at ∞ is known since $\mu_j \to a_j$. Moreover $r_j(\lambda)$ has a pole of order jn_k at $\lambda = \lambda_k$. Hence it can be expressed on $j\sum_k n_k$ parameters namely the coefficients of all these poles. So we have altogether $\frac{1}{2}N(N+1)\sum_k n_k$ parameters. They are not all independent however. Indeed above $\lambda = \lambda_k$ the various branches can be written:

$$\mu_j = \sum_{n=1}^{n_k} \frac{c_n^{(j)}}{(\lambda - \lambda_k)^n} + \text{regular}$$

where all the coefficients $c_1^{(j)}, \cdots, c_{n_k}^{(j)}$ are fixed and non–dynamical because they are the matrix elements of the diagonal matrices $(A_k)_-$, while the regular part is dynamical. This implies Nn_k constraints on the coefficients of r_j (note that r_j is the symmetrical function $\sigma_j(\mu_1, \cdots, \mu_N)$ hence the n_k highest order terms in r_j are fixed). We are left with $\frac{1}{2}N(N-1)\sum_k n_k$ parameters, that is $g + N - 1$ parameters.

It remains to take the quotient by the action of constant diagonal matrices. Consider the hamiltonians $H_n = (1/n)\,\mathrm{res}_{\lambda=\infty}\mathrm{Tr}\,(L^n(\lambda))\,d\lambda$, i.e. the term in $1/\lambda$ in $\mathrm{Tr}\,(L^n(\lambda))$. These are functions of the $r_j(\lambda)$. We show that they are the generators of the diagonal action. First we have:

$$\mathrm{Res}_{\lambda=\infty}\mathrm{Tr}\,(L^n(\lambda))d\lambda = n\,\mathrm{Res}_{\lambda=\infty}\mathrm{Tr}\,(L_0^{n-1}\sum_k L_k(\lambda))d\lambda = n\,\mathrm{Res}_{\lambda=\infty}\mathrm{Tr}\,(L_0^{n-1}L(\lambda))d\lambda$$

since all L_k are of order $1/\lambda$ at ∞. Assuming that the system is equipped with the Poisson structure (2.24) of Chapter [2], we get

$$\{H_n, L(\mu)\} = -\,\mathrm{Res}_{\lambda=\infty}\mathrm{Tr}_1\, L_0^{n-1} \otimes 1 \left[\frac{C_{12}}{\lambda - \mu}, L(\lambda) \otimes 1 + 1 \otimes L(\mu)\right] d\lambda$$

The term $L(\lambda) \otimes 1$ in the commutator does not contribute because the L_0 part produces a vanishing contribution by cyclicity of the trace and all other terms are of order at least $1/\lambda^2$. The term $1 \otimes L(\mu)$ yields $-[L_0^{n-1}, L(\mu)]$ which is the coadjoint action of a diagonal matrix on $L(\mu)$. Since L_0 is generic the L_0^n generate the space of all diagonal matrices, so we get exactly $N-1$ generators $H_1, \cdots, H_{N-1}$. In the hamiltonian reduction procedure, the H_n are the moments of the group action and are to be set to fixed (non–dynamical) values. Hence when the system is properly reduced we are left with exactly g action variables.

Example. Let us consider the example of the Neumann model. Recall the Lax matrix (see Chapter [2]):

$$L(\lambda) = D + \frac{1}{\lambda}J - \frac{1}{\lambda^2}K$$

The spectral curve can be computed as follows:

$$\begin{aligned}
\det\,(L(\lambda) - \mu) &= \det\left((D-\mu) + \frac{1}{\lambda}J - \frac{1}{\lambda^2}K\right) \\
&= \det\,(D-\mu)\det\left(1 + (D-\mu)^{-1}(\frac{1}{\lambda}J - \frac{1}{\lambda^2}K)\right) \\
&= \prod_i (a_i - \mu)\left[1 + \frac{1}{\lambda^2}\sum_k \frac{F_k}{\mu - a_k}\right] \qquad (4.7)
\end{aligned}$$

where the conserved quantities F_k are given by:

$$F_k = x_k^2 + \sum_{j\neq k} \frac{J_{kj}^2}{a_k - a_j} \tag{4.8}$$

This is because J is a projector of rank 2 and K a projector of rank 1. The matrix $P = (D-\mu)^{-1}(\frac{1}{\lambda}J - \frac{1}{\lambda^2}K)$ is of rank two. Its image is spanned by the two vectors $v_1 = (D-\mu)^{-1}X$ and $v_2 = (D-\mu)^{-1}Y$ while its kernel is the $(N-2)$–dimensional space orthogonal to X and Y, which is generically supplementary to the image. In this basis we have $P = 0$ on the kernel, and:

$$\begin{aligned} Pv_1 &= \left(\frac{1}{\lambda}V - \frac{1}{\lambda^2}U\right) v_1 - \frac{1}{\lambda}Uv_2 \\ Pv_2 &= \left(\frac{1}{\lambda}(W+1) - \frac{1}{\lambda^2}V\right) v_1 - \frac{1}{\lambda}Vv_2 \end{aligned}$$

where the functions $U(\mu), V(\mu), W(\mu)$ are defined as:

$$U = \sum_i \frac{x_i^2}{a_i - \mu}, \quad V = \sum_i \frac{x_i y_i}{a_i - \mu}, \quad W = -1 + \sum_i \frac{y_i^2}{a_i - \mu} \tag{4.9}$$

From this it follows that:

$$\det(1+P) = 1 - \frac{1}{\lambda^2}(V^2 - UW) = 1 + \frac{1}{\lambda^2}\sum_k \frac{F_k}{\mu - a_k} \tag{4.10}$$

which yields the result. Incidentally this proves formula (1.19) in Chapter [1]. Since we have already found an r–matrix for the Neumann system this proves its integrability.

Of course the spectral curve can be written in polynomial form:

$$R(\lambda, \mu) \equiv \lambda^2 \prod_i (\mu - a_i) + \sum_k \prod_{i\neq k} (\mu - a_i) F_k = 0$$

Performing the birational transformation (see Chapter [15]) $\lambda' = \lambda \prod_i (\mu - a_i)$, we get:

$$\lambda'^2 = -\prod_i (\mu - a_i) \sum_k \prod_{i\neq k} (\mu - a_i) F_k = P_{2N-1}(\mu) = -\prod_{i=1}^{N} (\mu - a_i) \prod_{i=1}^{N-1} (\mu - \beta_i) \tag{4.11}$$

which is an hyperelliptic curve of genus $g = N-1$. Note that the phase space is of dimension $2(N-1)$ and that we have $(N-1)$ independent conserved quantities, namely the N quantities F_k modulo the relation $\sum_k F_k = 1$. Let us remark that in this case we do not quotient by the diagonal action. This is because the diagonal action does not preserve the particular form of the Lax matrix in terms of the vectors X and Y. Alternatively the moments $H_n = \mathrm{Tr}\, D^{n-1} J$ identically vanish in this case.

To illustrate the discussion of N–sheeted coverings, it is instructive to consider the covering projection $(\lambda, \mu) \to \lambda$ which allows to see the spectral curve as an N–sheeted covering of the Riemann sphere of the variable λ. To compute the branching index of this covering we have to find the total number of poles of $\partial_\mu R(\lambda, \mu)$. Such poles can only occur when $\lambda = \infty$ or $\mu = \infty$. First above $\lambda = \infty$ we have the N points $\mu = a_i$. These are not branch points and the local parameter is $1/\lambda$. Around such a point we have:

$$(\lambda, \mu) \to Q_i = (\infty, a_i), \quad \mu = a_i - F_i/\lambda^2 + O(1/\lambda^4) \tag{4.12}$$

by eq.(4.7), hence $\partial_\mu R = \lambda^2 \prod_{j\neq i}(a_i - a_j) + O(1)$. We thus have N double poles at these points. When $\mu \to \infty$ we have, by eq.(4.7), $\lambda^2 = -1/\mu \to 0$ and λ is again a local parameter.

$$(\lambda, \mu) \to (0, \infty), \quad \mu = -\frac{1}{\lambda^2} + O(1) \tag{4.13}$$

At this point we have $\partial_\mu R = -(-1/\lambda^2)^{N-2} + \cdots$, hence we have a pole of order $2N-4$. So the branching index is $\nu = 4N-4$. This yields $g = N-1$. Note that $R(\lambda, \mu)$ is a very non generic polynomial and the orbit to which $L(\lambda)$ belongs is a very low dimensional one, nevertheless all the numbers fit nicely.

4.2 The eigenvector bundle.

The aim of this Chapter is to present the general procedure for solving integrable models using analytical properties on the spectral curve. So, in order to simplify the exposition we shall assume that all functions or matrices, are generic.

We wish to examine how the Lax matrix can be reconstructed from the analytic data characterizing its eigenvectors. This analysis will show the special role played by the divisor D of the poles at finite distances of the eigenvector $\Psi(P)$. *This divisor contains all the dynamical information.*

Let P be a point on the spectral curve. We assume that $P = (\lambda, \mu)$ is not a branch point so that all eigenvalues of $L(\lambda)$ are distinct and the eigenspace at P is one–dimensional. Let $\Psi(P)$ be an eigenvector, and $\psi_j(P)$ its N components:

$$\Psi(P) = \begin{pmatrix} \psi_1(P) \\ \vdots \\ \psi_N(P) \end{pmatrix}$$

Since the normalization of the eigenvector $\Psi(\lambda, \mu)$ is arbitrary, one has to make a choice before making a statement about its analytical properties. We choose to normalize it such that its first component is equal to one, ie.

$$\psi_1(P) = 1, \quad \text{at any point} \quad P \in \Gamma.$$

It is then clear that the $\psi_j(P)$ depend locally analytically on P. As a matter of fact:

Proposition. *With the above normalization, the components of the eigenvectors $\Psi(P)$ at the point $P = (\lambda, \mu)$ are meromorphic functions on the spectral curve Γ.*

Proof. For a generic point P of the curve Γ, i.e. for a pair $P = (\lambda, \mu)$ satisfying equation (4.3), there exists a unique eigenvector $\Psi(P)$ of the matrix $L(\lambda)$ normalized by the condition $\psi_1(P) = 1$. The un–normalized components $\psi_i(P)$ can be taken as $\Delta_i(P)$ where $\Delta_i(P)$ are suitable minors of the matrix $L(\lambda) - \mu\mathbf{1}$, and are thus meromorphic functions on Γ. After normalizing the first component, all the other coordinates $\psi_j(P)$ are still meromorphic functions on Γ. ∎

To each point $P(\lambda, \mu)$ on Γ we associate a meromorphic eigenvector $\Psi(P)$. At a branch point however, special care must be taken since there could be several eigenvectors associated to that point. We show that for generic Lax matrix the eigenspaces are 1–dimensional even at a branch point P. Moreover the eigenspaces around P admit a unique analytic continuation at P, irrespective of the branch chosen.

Proposition. *To each point P in Γ we associate the eigenspace at P. This defines an analytic line bundle that we call the eigenvector bundle.*

Proof. Consider the matrix $A = L(\lambda) - \mu$. The fact that we are on a branch point of the curve Γ is expressed by two algebraic equations in the coefficients of A: $R(\lambda, \mu) = 0$ and $\partial_\mu R(\lambda, \mu) = 0$. If however the kernel of A at P is of dimension ≥ 2, this means that the dimension of its image is $\leq (N-2)$. Let us show that this implies three algebraic equations on the coefficients of A. Let $e_1, \cdots, e_N$ be the columns of A. We assume that $e_1, \cdots, e_{N-2}$ are independent. First we impose that $e_1, \cdots, e_N$ are linearly dependent, producing one condition $\det A = 0$. Then we impose that $e_1, \cdots, e_{N-1}$ are dependent on $e_1, \cdots, e_{N-2}$. This is expressed by the vanishing of two $(N-1) \times (N-1)$ minors, hence the variety of such matrices is of codimension 1 in the generic space. In view of the argument of the previous proposition it is clear that in the generic situation the eigenvector bundle is analytic with respect to the analytic parameter on Γ. ∎

We now compute the Chern class of the eigenvector bundle. To do that we consider the section $\Psi(P)$ and compute the number of poles of this section, noting that it never vanishes, because the first component is equal to 1.

The number of poles of the normalized eigenvectors cannot be deduced by simply counting the number of zeroes of minors. Indeed let $\widehat{\Delta}(\lambda,\mu)$ be the matrix of cofactors of $(L(\lambda)-\mu\mathbf{1})$, which, by definition, is such that $(L(\lambda)-\mu\mathbf{1})\widehat{\Delta} = R(\lambda,\mu)\mathbf{1}$. Therefore at $P=(\lambda,\mu)\in\Gamma$, the matrix $\widehat{\Delta}(P)$ is a matrix of rank one, since the kernel of $(L(\lambda)-\mu\mathbf{1})$ is of dimension one. Hence, for $P\in\Gamma$ the matrix elements of $\widehat{\Delta}(P)$ are of the form $\alpha_i(P)\beta_j(P)$ and the components of the normalised eigenvector are $\psi_i(P) = \frac{\alpha_i(P)\beta_1(P)}{\alpha_1(P)\beta_1(P)} = \frac{\alpha_i(P)}{\alpha_1(P)}$. We thus expect cancellations to occur when we take the ratio of the minors and we cannot deduce the number of poles of the normalised eigenvector by simply counting the number of zeroes of the first minor.

Proposition. *The number of poles of the eigenvectors $\Psi(P)$ is:*

$$m \quad = \quad g+N-1 \tag{4.14}$$

Proof. Let us introduce the function $W(\lambda)$ of the complex variable λ defined by:

$$W(\lambda) = \left(\det\widehat{\Psi}(\lambda)\right)^2$$

where $\widehat{\Psi}(\lambda)$ is the matrix of eigenvectors of $L(\lambda)$ defined as follows:

$$\widehat{\Psi}(\lambda) = \begin{pmatrix} \psi_1(P_1) & \psi_1(P_2) & \cdots & \psi_1(P_N) \\ \vdots & \vdots & \vdots & \vdots \\ \psi_N(P_1) & \psi_N(P_2) & \cdots & \psi_N(P_N) \end{pmatrix} \tag{4.15}$$

where the points P_i are the N points above λ. Changing the normalization of the eigenvectors $\Psi(P_j)$ amounts to multiplying $\widehat{\Psi}(\lambda)$ on the right by a diagonal matrix. By definition $\widehat{\Psi}(\lambda)$ is the matrix diagonalizing $L(\lambda)$. The function $W(\lambda)$ is well–defined as a rational function of λ on the Riemann sphere since the square of the determinant does not depend on the order of the P_j's. It has a double pole where $\Psi(P)$ has a simple pole. To count its poles, we count its zeroes.
We show that $W(\lambda)$ has a simple zero for values of λ corresponding to a branch–point of the covering, therefore $m=\nu/2$. Recall that from eq.(4.6) the number of branching points is $\nu = 2(N+g-1)$.
First notice that $W(\lambda)$ only vanishes on branch–points where there are at least two identical columns. Indeed, let $P_i = (\mu_i,\lambda)$ be the N points above λ. Then the $\Psi(P_i)$ are the eigenvectors of $L(\lambda)$ corresponding to the eigenvalues μ_i are thus linearly independent when all the μ_i's are different. Therefore $W(\lambda)$ cannot vanish at such a point. The other possibility for the vanishing of $W(\lambda)$ would be that the vector $\Psi(P)$ itself vanish at some point (all components have a common zero at this point), but this is impossible because the first component is always 1. Let us assume now that λ_0 corresponds to a branch point, which is generically of order 2. At such a point $W(\lambda)$ has a simple zero. Indeed let z be an analytical parameter on the curve around the branch point. The covering projection $P\to\lambda$ gets expressed as $\lambda = \lambda_0 + \lambda_1 z^2 + O(z^3)$. The determinant vanishes to order z, hence W vanishes to order z^2. This is precisely proportional to $\lambda-\lambda_0$. A similar analysis can be performed if the branch point is of higher order. ■

We now need to examine the behavior of the eigenvector around $\lambda=\infty$. At the N points Q_i above $\lambda=\infty$, the eigenvectors are proportional to the canonical vectors $e^{(i)}$, $(e^{(i)})_k = \delta_{ik}$, since $L(\lambda=\infty)$ is diagonal, cf eq.(4.5). While this is compatible with the normalization $\psi_1(P)=1$ at the point Q_1, it is not compatible at the points Q_i, $i\geq 2$, if the proportionality factor remains finite. The situation is described more precisely by the following:

Proposition. *The k^{th} component $\psi_k(P)$ of $\Psi(P)$ has a simple pole at Q_k and vanishes at Q_1 for $k = 2, 3, \cdots, N$.*

Proof. Around $Q_k(\lambda = \infty, \mu = a_k)$, $k = 1, \cdots, N$, the eigenspace of $L(\lambda)$ is spanned by a vector of the form $v_k(\lambda) = e^{(k)} + O(1/\lambda)$. The first component of v_k is $v_k^1 = \delta_{1k} + O(1/\lambda)$. To get the normalized Ψ on has to divide v_k by v_k^1. So we get:

$$\Psi(P)|_{P\sim Q_1} = \begin{pmatrix} 1 \\ O(1/\lambda) \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ O(1/\lambda) \end{pmatrix}, \quad \Psi(P)|_{P\sim Q_k} = \begin{pmatrix} 1 \\ O(1) \\ \vdots \\ b_k\lambda + O(1) \\ O(1) \\ \vdots \\ O(1) \end{pmatrix}, \quad k \geq 2 \tag{4.16}$$

where $b_k\lambda + O(1)$ is the announced pole of the $k^{\rm th}$ component of $\Psi(P)|_{P\sim Q_k}$. ∎

The previous proposition shows that fixing the gauge by imposing that $L(\lambda)$ is diagonal at $\lambda = \infty$ introduces $N-1$ poles at the positions Q_i, $i = 2, \cdots, N$. The location of these poles is really part of the choice of the gauge condition. They do not contain any dynamical information. Only the positions of the other g poles have a dynamical significance. Let D be the divisor of these dynamical poles. Recall that the vector $\Psi(P)$ possesses a pole if one of its components has a pole. Therefore the two previous propositions tell us that the divisor of the k^{th} components of the eigenvector $\Psi(P)$ is bigger than $(-D + Q_1 - Q_k)$. This information is enough to reconstruct the eigenvectors and the Lax matrix.

Proposition. *Let D be a generic divisor on Γ of degree g. Up to normalization, there is a unique meromorphic function $\psi_k(P)$ with divisor $(\psi_k) \geq -D + Q_1 - Q_k$.*

Proof. This is a direct application of the Riemann–Roch theorem, since ψ_k is required to have $g+1$ poles and one prescribed zero. Hence it is generically unique apart from multiplication by a constant $\psi_k \to d_k\psi_k$. ∎

Equipped with these functions $\psi_k(P)$ for $k = 2, \cdots, N$ we construct a vector function with values in $\mathbb{C}^N$:

$$\Psi(P) = \begin{pmatrix} 1 \\ \psi_2(P) \\ \vdots \\ \psi_N(P) \end{pmatrix}$$

The normalization ambiguity of the ψ_k translates into left multiplication of the vector $\Psi(P)$ by a *constant* diagonal matrix $d = \mathrm{diag}\,(1, d_2, \cdots, d_N)$. We have constructed a line bundle on the Riemann surface Γ, which is the line bundle associated to the divisor $-D - Q_2 - \cdots - Q_N$ of degree $-(g + N - 1)$ (see Chapter [15]). In fact we have constructed an embedding of this line bundle into $\Gamma \times \mathbb{C}^N$.

Theorem. *There exists a unique matrix $L(\lambda)$, rational in λ, such that*

$$(L(\lambda) - \mu\mathbf{1})\Psi(P) = 0$$

This matrix has prescribed poles at the points λ_k and satisfies the boundary condition: $\lim_{\lambda\to\infty} L(\lambda) = \mathrm{diag}(a_1, \cdots, a_N)$.

Proof. Consider the matrix $\widehat{\Psi}(\lambda)$ whose columns are the vectors $\psi(P_i)$ with $P_i = (\lambda, \mu_i)$, cf. eq.(4.15). This matrix depends on the ordering of the columns, i.e., on the ordering of the points P_i. However, the matrix

$$L(\lambda) = \widehat{\Psi}(\lambda) \cdot \widehat{\mu} \cdot \widehat{\Psi}^{-1}(\lambda) \tag{4.17}$$

does not depend on this ordering and is a well defined function on the base curve. Here $\widehat{\mu}$ is the diagonal matrix $\widehat{\mu} = \mathrm{diag}\,(\mu_1, \cdots, \mu_N)$. From its definition $\det\widehat{\Psi}(\lambda) \neq 0$ if λ is not a branch point. At a generic branch point two eigenvectors coalesce and the determinant has a simple zero with respect to the local parameter, but the matrix (4.17) is still regular. To see this recall that if $\widehat{\Delta}$ is the matrix of cofactors of $\widehat{\Psi}$ we have:

$$\widehat{\Psi} \cdot \widehat{\Delta} = \det\widehat{\Psi}\ 1, \quad L(\lambda) = \frac{1}{\det\widehat{\Psi}} \widehat{\Psi} \cdot \widehat{\mu} \cdot \widehat{\Delta} \tag{4.18}$$

At the branch point, $\widehat{\Psi}\widehat{\Delta} = 0$, thus $\mathrm{Im}\ \widehat{\Delta} = \mathrm{Ker}\ \widehat{\Psi}$ which is one–dimensional. We may assume without loss of generality that the two eigenvectors that coalesce are the first two columns of $\widehat{\Psi}$. So, at the branch point, the kernel of $\widehat{\Psi}$ is spanned by $e^{(1)} - e^{(2)}$, where $e^{(i)}$ are the canonical base vectors $(e^{(i)})_j = \delta_{ij}$. Since the first two diagonal elements of $\widehat{\mu}$ also become equal we see that $\widehat{\mu}$ acts scalarly on $\mathrm{Im}\ \widehat{\Delta}$, so $\widehat{\Psi}(\lambda) \cdot \widehat{\mu} \cdot \widehat{\Delta} = 0$. Hence the numerator of eq.(4.18) has a simple zero at the branch point, cancelling the simple pole of the determinant. Therefore the matrix $L(\lambda)$ is a rational function of the parameter λ. It has poles only at the projections of the points where μ has poles, i.e., at the points λ_k.
At $\lambda = \infty$, the leading part of $\widehat{\Psi}(\lambda)$ is diagonal since it is dominated by the functions $\psi_k(P)$ with P approaching Q_k. Therefore at infinity $L(\lambda)$ goes to $\widehat{\mu}|_{\lambda=\infty} = \mathrm{diag}(a_1, \cdots, a_N)$. ■

This theorem is a crucial step of this method of resolution. It tells us that, once the spectral curve has been given, which amounts to give the values of the integrals of motion, *all dynamical data are encoded into the divisor D*. In other words, this theorem teaches us that the dynamical variables are the points of this divisor.

It should be emphasized however that $\Psi(P)$ is defined up to left multiplication by diagonal matrices $\psi_k \to d_k\psi_k$. On the Lax matrix $L(\lambda)$ this amounts to a conjugation by a constant diagonal matrix. Hence the object we reconstruct is actually the hamiltonian reduction of the dynamical system by this group of diagonal matrices as emphasized at the beginning of this chapter.

Remark. It is worth comparing the reconstruction formula (4.17) for the Lax matrix with the local analysis of Chapter [2]. Recall that in this Chapter we explained that the Lax pair of matrices $L(\lambda)$ and $M(\lambda)$ could be diagonalized simultaneously, locally around each pole λ_k. Explicitly the diagonalization formula (2.8) for $L(\lambda)$ was $L(\lambda) = g_k A_k g_k^{-1}$ with g_k and A_k formal power series in $(\lambda - \lambda_k)$ and A_k diagonal. The expression (4.17) of $L(\lambda)$ in terms of the eigenvectors is simply a global version of the previous local statement.

Example. Let us illustrate the analytical properties of the eigenvector bundle on the example of the Neumann model. There is a simple description of the eigenvectors in this case. Indeed from eq.(2.3) in Chapter [2]

$$L(\lambda)\Psi = \mu\Psi \Longrightarrow \Psi = -(D-\mu)^{-1}\left(\frac{1}{\lambda}J\Psi - \frac{1}{\lambda^2}K\Psi\right) \tag{4.19}$$

Since $J\Psi = (Y \cdot \Psi)X - (X \cdot \Psi)Y$ and $K\Psi = (X \cdot \Psi)X$ we see that Ψ is known once we know its projections on X and Y. Projecting eq.(4.19) on X and Y one gets a 2×2 system:

$$\begin{pmatrix} 1 - \frac{1}{\lambda}V(\mu) - \frac{1}{\lambda^2}U(\mu) & \frac{1}{\lambda}U(\mu) \\ -\frac{1}{\lambda}(1 + W(\mu)) - \frac{1}{\lambda^2}V(\mu) & 1 + \frac{1}{\lambda}V(\mu) \end{pmatrix} \begin{pmatrix} X \cdot \Psi \\ Y \cdot \Psi \end{pmatrix} = 0$$

The vanishing of the determinant of this linear system is precisely the equation of the spectral curve:

$$\lambda^2 = V^2(\mu) - U(\mu)W(\mu)$$

Solving this system and inserting back into eq.(4.19) one gets:

$$\frac{\psi_k}{\psi_1} = \frac{a_1-\mu}{a_k-\mu}\,\frac{(\lambda - V(\mu))x_k + U(\mu)y_k}{(\lambda - V(\mu))x_1 + U(\mu)y_1}$$

Recalling the expansion (4.12) we see that $(a_1-\mu)/(a_k-\mu)$ has a double zero at $Q_1(\lambda=\infty, \mu=a_1)$ and a double pole at $Q_k(\lambda=\infty,\mu=a_k)$. Consider the meromorphic functions $\phi_k = (\lambda - V(\mu))x_k + U(\mu)y_k$. It has poles at the points Q_i. Using eq.(4.12) we see that it has double poles at $Q_i,\ i\neq k$ and a simple pole at Q_k. Indeed $(-V(\mu)x_k + U(\mu)y_k) = \lambda^2 x_i J_{ki}/F_i + O(1)$. This show that ψ_k/ψ_1 has a *simple* zero at Q_1 and a *simple* pole at Q_k.

To find the other poles of ψ_k/ψ_1 we study the zeroes of ϕ_k which has $(2N-1)$ poles and therefore $(2N-1)$ zeroes. Amongst them, N are common to all functions ϕ_k and cancel in ψ_k/ψ_1. Indeed a common zero is such that $\lambda - V(\mu) = U(\mu) = 0$. By eq.(4.10) the points satisfying these two equations are on the spectral curve. The equation $U(\mu)=0$ has $(N-1)$ roots, μ_j, at finite μ, and $\lambda_j = V(\mu_j)$ selects one of the two points above μ_j. In addition we have the point $\lambda=0,\ \mu=\infty$ which is a simple zero in view of eq.(4.13). Finally ψ_k/ψ_1 has a zero at Q_1 a pole at Q_k and $g=N-1$ poles at finite distance depending on the dynamical data, in agreement with the general result.

4.3 The adjoint linear system.

In view of eq.(4.17), it is important to compute $\widehat{\Psi}^{-1}(\lambda)$ in an efficient way. This is achieved by introducing the solution $\Psi^+(P)$ of the adjoint linear system:

$$\Psi^+(P)\,(L(\lambda) - \mu\mathbf{1}) \quad = \quad 0 \tag{4.20}$$

The precise relation between $\Psi^+(P)$ and the matrix $\widehat{\Psi}^{-1}(\lambda)$ is provided by the following:

Proposition. *Let $\Psi^+(P)$ be a solution of the adjoint system (4.20). The inverse of the matrix $\widehat{\Psi}(\lambda)$ defined in eq.(4.15) is the matrix whose lines are the N vectors*

$$\Psi^{(-1)}(P_j) = \Psi^+(P_j)/ < \Psi^+(P_j), \Psi(P_j) >$$

with P_j the N points above λ.

<u>Proof</u>. One has to show that

$$\sum_{k=1}^{N} \frac{\psi_k^+(P_j)}{< \Psi^+(P_j), \Psi(P_j) >}\psi_k(P_i) = \delta_{ij}$$

where P_i and P_j are two points of Γ above the same λ. This is obvious for $i=j$, and for $i\neq j$ $\Psi(P_i)$ and $\Psi^+(P_j)$ are orthogonal because they are associated to different eigenvalues $\mu(P_i)$ and $\mu(P_j)$. So we get:

$$\widehat{\Psi}^{-1}(\lambda) = \begin{pmatrix} \frac{\Psi^+(P_1)}{<\Psi^+(P_1),\Psi(P_1)>} \\ \vdots \\ \frac{\Psi^+(P_N)}{<\Psi^+(P_N),\Psi(P_N)>} \end{pmatrix} \tag{4.21}$$

■

We now use these relations between $\Psi^+(P)$ and $\widehat{\Psi}^{-1}(\lambda)$ to reconstruct them from their analyticity properties. One may perform on $\Psi^+(P)$ the same analysis as for the vector $\Psi(P)$. Normalizing the first component of $\Psi^+(P)$ to 1, one sees that the vector $\Psi^+(P)$ has g poles at a divisor D^+ and $(N-1)$ simple poles at $Q_2, \cdots, Q_N$ above $\lambda = \infty$. Moreover $\psi_k^+(P)$, $k \geq 2$, has a zero at Q_1. Our first task is to relate the divisor D^+ to D.

Proposition. *Let $\Psi^+(P)$ be the solution of eq.(4.20) normalized by $\psi_1^+(P) = 1$. The differential form*

$$\Omega \equiv \frac{d\lambda}{< \Psi^+(P)\Psi(P) >} \tag{4.22}$$

is an abelian differential of second kind with a double pole at Q_1 and zeroes at D and D^+. Conversely there is a unique differential Ω of the second kind with a double pole at the point Q_1 and having amongst its zeroes the g points of D. Its other g zeroes are therefore completely fixed and define D^+. Its image under the Abel map is given by:

$$\mathcal{A}(D^+) = -\mathcal{A}(D) + \mathcal{A}(B) - 2\sum_{j=2}^{N} \mathcal{A}(Q_j) \tag{4.23}$$

where B is the divisor of branch points of the covering $(\lambda, \mu) \to \lambda$.

Proof. Consider the meromorphic function $f(p) =< \Psi^+(P), \Psi(P) >$. It has $2(g+N-1)$ poles coming from the poles of Ψ and Ψ^+. Their divisor is $D + D^+ + 2\sum_{j=2}^{N} Q_j$. Therefore it has also $2(g+N-1)$ zeroes, which are in fact the branch points of the covering, as we next see (recall that the covering has $2(g+N-1)$ branch points by the Riemann–Hurwitz formula). So let $P = (\lambda_0, \mu_0)$ be a branch point and consider two points P_1 and P_2 above the same λ close to λ_0 on the two sheets of the covering that coalesce at P. Because $\Psi^+(P_1)$ and $\Psi(P_2)$ are dual eigenvectors corresponding to different eigenvalues, they are orthogonal:

$$< \Psi^+(P_1), \Psi(P_2) >= 0 \tag{4.24}$$

The assertion then follows by continuity, since $\Psi(P_2) \to \Psi(P)$ and $\Psi^+(P_1) \to \Psi^+(P)$ (recall that the line bundles are analytic at P).
At a branch point P we have $(\mu - \mu_0)^2 \sim (\lambda - \lambda_0)$ so $\mu - \mu_0$ is a local parameter and $d\lambda = (\mu - \mu_0)d\mu$ vanishes at P. Moreover $d\lambda$ has double poles at $Q_1, \cdots, Q_N$. We see that Ω is regular at the branch points and has a double pole at Q_1 and zeroes at $D + D^+$. Applying Abel's theorem to the meromorphic function $\Omega/d\lambda$ one gets eq.(4.23).
Conversely, assume that we have two such forms Ω and Ω' with divisor of zeroes $D + D^+$ and $D + D'^+$ respectively. Their quotient is a meromorphic function with a divisor of poles of degree g, i.e. a constant generically. Hence the differential Ω is unique. ∎

The outcome of this proposition is that Ω is uniquely characterized by its behaviour at infinity and the zeroes at D. Therefore, given the dynamical divisor D, we know the form Ω and we can find the divisor D^+ as the complementary set of zeroes of Ω.

This information on the divisor D^+ can now be used to reconstruct the vector $\Psi^+(P)$. Its components $\psi_k^+(P)$ are uniquely determined up to normalization once we know the divisor D^+. Here however we have no freedom on these normalizations since we must preserve the orthogonality conditions (4.24). Let $\widetilde{\psi}_k^+(P)$ any choice of such meromorphic function and let $\psi_k^+(P) = (1/c_k)\,\widetilde{\psi}_k^+(P)$. We want to determine the constant c_k. By hypothesis $\psi_k^+(P)$ satisfies an orthogonality relation of the form $< \Psi^+(P_j), \Psi(P_i) >= f(P_i)\delta_{ij}$. Then by uniqueness of the inverse matrix we also have:

$$\sum_k \psi_i^+(P_k)\psi_j(P_k)\frac{1}{f(P_k)} = \delta_{ij}, \quad \text{or} \quad \sum_k \widetilde{\psi}_i^+(P_k)\psi_j(P_k)\frac{1}{f(P_k)} = c_i\delta_{ij}$$

We have an independent characterization of $f(P)$ using eq.(4.22), i.e. $1/f(P) = \Omega/d\lambda(P)$ and therefore $f(P)$ is known from its analyticity properties. This allows us to compute c_i as follows.

Consider on the Riemann surface the form $\widetilde{\psi}_i^+(P)\psi_j(P)\Omega(P)$. For $i=j$, it has a double pole at Q_i and for $i \neq j$, it has simple poles at Q_i and Q_j and no other singularity. This is because the poles at D and D^+ in Ψ and Ψ^+ cancel against the zeroes of Ω, and the double pole of Ω at Q_1 combines with zeroes of Ψ and Ψ^+ at this point. Finally we define a form on the Riemann sphere of the variable λ:

$$\omega_{ij} = \sum_k \widetilde{\psi}_i^+(P_k)\psi_j(P_k)\Omega(P_k)$$

where the N points P_k are the points above λ. If there are no branch points amongst the P_k, λ is a local parameter around each P_k and $\omega_{ij} = g_{ij}(\lambda)d\lambda$. If there is a branch point a short computation shows that this still holds. Since ω_{ij} is regular for finite λ, the function $g_{ij}(\lambda)$ is in fact a polynomial of λ. Moreover ω_{ij} has poles of order at most 1 for $i \neq j$ and 2 for $i = j$ at $\lambda = \infty$. Since $d\lambda$ has a double pole at ∞, this implies $\omega_{ij} = 0$ for $i \neq j$ and $\omega_{ii} = c_i d\lambda$ for some *constants* c_i. We have obtained orthogonality relations:

$$\omega_{ij} = c_i\delta_{ij}d\lambda, \quad c_i = \lim_{P\to Q_i} \widetilde{\psi}_i^+(P)\psi_i(P)\frac{\Omega}{d\lambda}(P) \tag{4.25}$$

These orthogonality relations show that the inverse of the matrix $\widehat{\Psi}$ is given by:

$$\left(\widehat{\Psi}^{-1}\right)_{ij} = \frac{\Omega(P_i)}{d\lambda}\,\frac{1}{c_j}\widetilde{\psi}_j^+(P_i) \tag{4.26}$$

It is worth noticing that this expression is invariant under a change of normalization of the components of Ψ^+, i.e. $\psi_j^+ \to d_j\psi_j^+$ yields $c_j \to d_jc_j$ and d_j cancels. On the other hand it transforms appropriately under a change of normalization of the components of Ψ. One then gets $\Omega = d\lambda/ <\Psi^+,\Psi>$ if one sets the k^{th} component of Ψ^+ to the *invariant* value $1/c_k\ \widetilde{\psi}_k^+$. Let us summarize the situation in a proposition:

Proposition. *Given an effective divisor D of degree g, the functions $\psi_k(P)$ with divisor $\geq -(D+Q_k-Q_1)$ are unique up to normalization. There is a unique form Ω having a double pole at Q_1 and vanishing on D. It has g other zeroes at an effective divisor D^+ of degree g. There exists a* unique *set of functions $\psi_k^+(P)$ of divisor $\geq -(D^+ + Q_k - Q_1)$ such that $\sum_k \psi_i^+(P_k)\psi_j(P_k)\Omega(P_k) = \delta_{ij}d\lambda$. The inverse of the matrix $\widehat{\Psi}$ is $\widehat{\Psi}_{ij}^{-1} = \psi_j^+(P_i)\ \Omega(P_i)/d\lambda$.*

In the following we shall frequently denote $\Psi^{(-1)}(P)$ the line vector $\Psi^+(P)/<\Psi^+(P),\Psi(P)>$ appearing in eq.(4.21). This meromorphic section of the above line bundle has $2(g+N-1)$ poles at the branch points of the covering and $(g+N-1)$ zeroes at the points where $\Psi(P)$ has poles, in agreement with the value of the Chern class of this bundle.

4.4 Time evolution.

The aim of this section is to solve for the dynamics by looking at the time evolution of the dynamical divisor D. The outcome is the beautiful fact that *the dynamical flow linearizes on the Jacobian of the spectral curve.*

Recall that the time evolution is governed by the Lax equation,

$$\frac{d}{dt}L(\lambda) = [L(\lambda), M(\lambda)] \tag{4.27}$$

As we have seen in Chapter [2], the matrix $M(\lambda)$ is of the form

$$M = \sum_k M_k, \quad \text{with} \quad M_k(\lambda) = \left(P^{(k)}(L,\lambda)\right)_-$$

where $\left(P^{(k)}(L,\lambda)\right)$ is a polynomial in $L(\lambda)$ with rational coefficients in λ, and $\left(P^{(k)}(L,\lambda)\right)_-$ denotes its polar part at λ_k.

Suppose that at time t we made the analysis of the previous section and built the normalized eigenvector $\Psi(t,P)$. The Lax equation (4.27) induces the following time evolution of the eigenvector $\Psi(t,P)$.

$$\frac{d}{dt}\Psi(t,P) = -\left(M(\lambda) - C(t,P)\mathbf{1}\right)\ \Psi(t,P) \tag{4.28}$$

where $C(t,P)$ is an arbitrary function since $(L(\lambda)-\mu)(\frac{d\Psi}{dt}+M\Psi)=0$. Normalizing the eigenvector $\Psi(t,P)$ such that its first component equals 1 gives:

$$C(t,P) = \sum_j M_{1j}(\lambda)\psi_j(t,P)$$

By the analysis of the previous section, the normalized eigenvector $\Psi(t,P)$ has poles at the dynamical divisor $D(t)$ and at the $N-1$ points Q_k, $k=2,\cdots,N$. Consider the function $\mathcal{N}(t,dt,P)\equiv 1-C(t,P)dt$, with dt infinitesimal:

$$\mathcal{N}(t,dt,P) = 1 - dt\sum_j M_{1j}(\lambda)\psi_j(t,P)$$

One can rewrite equation (4.28) in the equivalent form:

$$\mathcal{N}(t,dt,P)\ \Psi(t+dt,P) = (\mathbf{1}-dtM(\lambda))\cdot\Psi(t,P) + O(dt^2) \tag{4.29}$$

We see that the meromorphic function $\mathcal{N}(t,dt,P)$ of $P\in\Gamma$ normalizes the eigenvector whose time evolution is naturally induced by the Lax equation. The divisor of this meromorphic function reads:

$$(\mathcal{N}) = D(t+dt) + \sum_{k,i}\sum_{\alpha=1}^{m_k} P^{\alpha}_{k,i} - D(t) - \sum_{k,i} m_k P_{k,i}$$

Indeed from eq.(4.29) we see that $\mathcal{N}$ cancels the poles of $\Psi(t+dt,P)$ at $D(t+dt)$ and produces the poles of $\Psi(t,P)$ at $D(t)$. The poles at $Q_2,\cdots,Q_N$ are the same on both sides and do not appear in $\mathcal{N}$. Moreover, since $M_k(\lambda)$ has a pole of order m_k at λ_k, $\mathcal{N}$ has poles of order m_k at the N points $P_{k,i}$ above λ_k. Finally $\mathcal{N}$ has extra zeroes $P^{\alpha}_{k,i}$. Since dt is small, $\mathcal{N}$ is close to 1, hence the zeroes are close to the poles, $D(t+dt)$ is close to $D(t)$, and on each sheet i of the covering there are exactly m_k zeroes $P^{\alpha}_{k,i}$ close to the m_k^{th} order pole $P_{k,i}$.

Theorem. *Let $\nu_j(t)$ with $j=1,\cdots,g$ be the points of the dynamical divisor $D(t)$. Let ω be any holomorphic differential on Γ. The time evolution of the points $\nu_j(t)$ induced by the Lax equation $\dot{L}(\lambda)=[L(\lambda),M(\lambda)]$ with $M(\lambda)=\sum_k\left(P^{(k)}(L,\lambda)\right)_-$ is such that:*

$$\frac{d}{dt}\sum_{j=1}^{g}\int^{\nu_j(t)}\omega = -\sum_k\sum_{i=1}^{N}\mathrm{Res}_{P_{k,i}}\left(\omega P^{(k)}(\mu,\lambda)\right) \tag{4.30}$$

where the points $P_{k,i}$ are the N points above λ_k. Notice that the right–hand side is independent *of time.*

<u>Proof</u>. Since $\mathcal{N}(t+dt,P)$ is a meromorphic function, Abel's theorem tells us that:

$$\sum_{j=1}^{g}\int_{\nu_j(t)}^{\nu_j(t+dt)}\omega + \sum_{k,i,\alpha}\int_{P_{k,i}}^{P_{k,\alpha}(dt)}\omega = 0 \tag{4.31}$$

for any holomorphic differential ω. This equation will give us the time evolution of the divisor as in eq.(4.30) if we can evaluate the second sum. For this we need the following lemma:

Lemma. *Consider a point $P \in \mathbb{C}$, a holomorphic differential ω in the neigbourhood V of P and an analytic function u in V with a pole of order m at P. Consider $\epsilon \in \mathbb{C}$ small enough, so that the m points $P_\alpha(\epsilon)$, where $u(P_\alpha(\epsilon)) + \epsilon^{-1} = 0$ belong to V. Then*

$$\lim_{\epsilon \to 0} \frac{1}{\epsilon} \sum_{\alpha=1}^{m} \int_{P}^{P_\alpha(\epsilon)} \omega = -Res_P(\omega u)$$

Proof. Let $\omega = d\sigma$ with $\sigma(P) = 0$. Then for any path π enclosing the zeroes $P_\alpha(\epsilon)$ of $u + \epsilon^{-1}$, we have

$$\frac{1}{\epsilon} \sum_{\alpha=1}^{m} \int_{P}^{P_i(\epsilon)} \omega \quad = \quad \frac{1}{\epsilon} \sum \sigma(P_i(\epsilon)) = \frac{1}{\epsilon} \sum_{\alpha} \mathrm{Res}_{P_\alpha(\epsilon)} \frac{u'}{u + \epsilon^{-1}} \sigma = \frac{1}{2i\pi} \int_{\pi} \frac{u'}{1 + \epsilon u} \sigma dz$$

When ϵ tends to zero, the right hand side tends to $\int_\pi u' \sigma \frac{dz}{2i\pi} = -\frac{1}{2i\pi} \int_\pi u\omega = -\mathrm{Res}_P(u\omega)$. ■

Decompose the second sum in eq.(4.31) as a sum of the terms associated to each pole of $M(\lambda)$, $P_{k,i}$ on Γ. The points $P^\alpha_{k,i}$ approaching $P_{k,i}$ are by definition solutions of $1 + dt u_k(P^\alpha_{k,i}) = 0$ with $u_k = -\sum_j (M_k)_{1j} \psi_j$. Thus, by the lemma one finds:

$$\sum_{j=1}^{g} \int_{\nu_j(t)}^{\nu_j(t+dt)} \omega = -dt \sum_{k,i} \mathrm{Res}_{P_{k,i}} \left(\omega \sum_j (M_k)_{1j} \psi_j(t, P) \right) \tag{4.32}$$

Recall that $M_k(\lambda)$ is the polar part of $P^{(k)}(L, \lambda)$, ie.

$$M_k(\lambda) = \left(P^{(k)}(L, \lambda) \right)_- = P^{(k)}(L, \lambda) - \left(P^{(k)}(L, \lambda) \right)_+ ,$$

with $\left(P^{(k)}(L, \lambda)\right)_+$ regular at λ_k, hence does not contribute to eq.(4.32). Therefore:

$$\begin{aligned} \mathrm{Res}_{P_{k,i}} \left(\omega \sum_j (M_k)_{1j} \psi_j(t, P) \right) &= \mathrm{Res}_{P_{k,i}} \left(\omega \left(P^{(k)}(L, \lambda) \Psi(t, P) \right)_1 \right) \\ &= \mathrm{Res}_{P_{k,i}} \left(\omega P^{(k)}(\mu, \lambda) \psi_1(t, P) \right) = \mathrm{Res}_{P_{k,i}} \left(\omega P^{(k)}(\mu, \lambda) \right) \end{aligned}$$

where we have used the fact that $\Psi(t, P)$ is an eigenvector and the normalization $\psi_1(t, P) = 1$. ■

Equation (4.30) can alternatively be written in terms of the Abel's map. Recall that given a point $P \in \Gamma$ or a divisor $D = (\nu_1, \cdots, \nu_g)$ of degree g on Γ, the Abel map with base point P_0 is defined by:

$$\mathcal{A}_j(P) = \int_{P_0}^{P} \omega_j \quad \text{and} \quad \mathcal{A}_j(D) = \sum_{j=1}^{g} \int_{P_0}^{\nu_j} \omega_j \tag{4.33}$$

with ω_j is a normalized basis of abelian differentials. Thus eq.(4.30) is equivalent to:

Theorem. *The flow induced by the Lax equation (4.27) on the eigenvector bundle is a linear flow on the Jacobian of the spectral curve:*

$$\mathcal{A}(D(t)) - \mathcal{A}(D_0) = tU^{(M)} \quad \text{with} \quad U_j^{(M)} = -\sum_{k,i} \mathrm{Res}_{P_{k,i}} \left(\omega_j P^{(k)}(\mu, \lambda) \right) \tag{4.34}$$

4.5 Theta functions formulae.

Since the motion linearizes on the Jacobian it is natural to express the solution in terms of Riemann's θ–functions. We use notations and results of Chapter [15], in particular $\mathcal{K}$ is the vector of Riemann's constant.

Recall that for e in an open set of the divisor $\Theta = \{z|\theta(z) = 0\}$ the function $\theta(e + \int_x^y \omega)$ does not vanish identically. By Riemann's theorem it has g zeroes $y_1, \cdots, y_g$ for given x. Since $e \in \Theta$ one of these zeroes is x, and we choose $y_1 = x$. We now show that $y_2, \cdots, y_g$ are independent of x. Indeed by Riemann's theorem we have: $\mathcal{A}(y_1) + \cdots + \mathcal{A}(y_g) = \mathcal{A}(x) - e - \mathcal{K}$, so that $y_2, \cdots, y_g$ are determined by $\mathcal{A}(y_2) + \cdots + \mathcal{A}(y_g) = -e - \mathcal{K}$. This equation is independent of x. As a side remark note that for such an y_j, $j \geq 2$, we have $\theta(-e + \int_{y_j}^x \omega) = 0$ for all x. This means that some translation of the curve embedded into the Jacobian is entirely contained in Θ, hence one has to be careful in the choice of the vector e. To use this result to construct meromorphic functions on the Riemann surface, notice that the building block $\theta(e + \int_{x_1}^y \omega)/\theta(e + \int_{x_2}^y \omega)$ has a zero at $y = x_1$ and a pole at $y = x_2$. The extra $(g-1)$ zeroes at the numerator and denominator cancel. We then assemble such blocks so that the product has no monodromy.

Explicit expressions for the matrix $\widehat{\Psi}(P)$ and its inverse are easily written. Let $D(t) = (\nu_1(t), \cdots, \nu_g(t))$ be the dynamical divisor. Let $U^{(M)}$ be the g-dimensional vector with components $U_j^{(M)} = -\sum_{k,i} \mathrm{Res}_{P_{k,i}} \left(\omega_j P^{(k)}(\mu, \lambda)\right)$ with ω_j the normalized abelian differentials. Then,

$$\psi_k(t, P) = d_k \; \frac{\theta(\mathcal{A}(P) - \mathcal{A}(Q_k) + \mathcal{A}(Q_1) - \zeta_{D(t)})}{\theta(\mathcal{A}(P) - \zeta_{D(t)})} \frac{\theta(e + \int_{Q_1}^P \omega)}{\theta(e + \int_{Q_k}^P \omega)} \tag{4.35}$$

In this equation $\zeta_D = \mathcal{A}(D) + \mathcal{K}$ with $\zeta_{D(t)} = \zeta_{D(0)} + tU^{(M)}$. The d_k are constants ($d_1 = 1$) related to the residual diagonal action of the gauge group. Note that the sum of the arguments of the theta functions in the numerator is equal to the sum of the arguments in the denominator so the whole expression has no monodromy, hence defines a meromorphic function with the correct zeroes and poles.

Similarly one has:

$$\psi_k^+(t, P) = \frac{1}{c_k} \; \frac{\theta(\mathcal{A}(P) - \mathcal{A}(Q_k) + \mathcal{A}(Q_1) - \zeta_{D^+(t)})}{\theta(\mathcal{A}(P) - \zeta_{D^+(t)})} \frac{\theta(e^+ + \int_{Q_1}^P \omega)}{\theta(e^+ + \int_{Q_k}^P \omega)} \tag{4.36}$$

Here D^+ is the divisor given by equation (4.23) and the normalization constants c_k are defined in eq.(4.25), and we now compute them. To do that we express the meromorphic function $\Omega/d\lambda$ in terms of theta functions. Let B be the set of branch points of the covering $(\lambda, \mu) \to \lambda$. We decompose the set of $2(g + N - 1)$ points of B into four subsets $B_0,\ B_0',\ B_1,\ B_1'$ such that $\# B_0 = \# B_0' = N - 1$ and $\# B_1 = \# B_1' = g$. This decomposition is arbitrary but does not affect the final formulae. Then we can write:

$$\frac{\Omega}{d\lambda}(P) = \frac{\prod_{j=2}^N \theta(e + \int_{Q_j}^P \omega)\theta(e^+ + \int_{Q_j}^P \omega)\ \theta(\mathcal{A}(P) - \zeta_{D(t)})\ \theta(\mathcal{A}(P) - \zeta_{D^+(t)})}{\prod_{b_k \in B_0} \theta(e + \int_{b_k}^P \omega)\ \prod_{b_k \in B_0'} \theta(e^+ + \int_{b_k}^P \omega)\ \theta(\mathcal{A}(P) - \zeta_{B_1})\ \theta(\mathcal{A}(P) - \zeta_{B_1'})} \tag{4.37}$$

This expression of $\Omega/d\lambda$ has the correct zeroes and poles, and has no monodromy in view of eq.(4.23).

To compute c_k we use eq.(4.25) yielding:

$$c_k = d_k \; \frac{\theta(\mathcal{A}(Q_1) - \zeta_{D(t)})\ \theta(\mathcal{A}(Q_1) - \zeta_{D^+(t)})\ \prod_{j \neq k} \theta(e + \int_{Q_j}^{Q_k} \omega)\theta(e^+ + \int_{Q_j}^{Q_k} \omega)}{\prod_{b_k \in B_0} \theta(e + \int_{b_k}^{Q_k} \omega) \prod_{b_k \in B_0'} \theta(e^+ + \int_{b_k}^{Q_k} \omega)\ \theta(\mathcal{A}(Q_k) - \zeta_{B_1})\ \theta(\mathcal{A}(Q_k) - \zeta_{B_1'})} \tag{4.38}$$

This gives an expression of the elements of the matrix $\widehat{\Psi}^{-1}$ as products of theta functions, through eq.(4.26).

Example Let us apply the above formalism to find the solution of the Neumann model. In this case since ${}^tL(\lambda) = L(-\lambda)$ the normalized $\Psi^+(\lambda, \mu)$ is equal to ${}^t\Psi(-\lambda, \mu)$. The transformation $(\lambda, \mu) \to (-\lambda, \mu)$

is just the hyperelliptic involution σ on the spectral curve of the Neumann model. The fixed points of σ are the $(2g+2)$ points (λ,μ) lying above $\lambda=\infty$, namely the Q_j, and above $\lambda=0$, namely the point $P_\infty(\lambda=0,\mu=\infty)$ and the $N-1$ points of coordinates $(\lambda=0,\mu=\beta_i)$, see eq.(4.11). We take the point P_∞ as base point of the Abel map, and note that the hyperelliptic involution changes the sign of the abelian differentials which are of the form $p(\mu)d\mu/\lambda$, so that for any point P we get $\mathcal{A}(\sigma(P))=-\mathcal{A}(P)$ modulo periods. The set B of branch points of the covering $(\lambda,\mu)\to\lambda$ are solutions of $\partial_\mu R(\lambda,\mu)=0$, i.e., $\sum_k F_k/(\mu-a_k)^2=0$. This equation has $(2N-2)$ roots at finite distance, each one giving rise to two points related by the hyperelliptic involution. Hence the set B is globally invariant under σ and one can choose $B_0'=\sigma(B_0)$ and $B_1'=\sigma(B_1)$ in the above construction. Considering eq.(4.23), we see that $D^+=\sigma(D)$. Indeed since $B=\sum_i(b_i+\sigma(b_i))$ with $\sigma(b_i)\neq b_i$ we have $\mathcal{A}(B)=\sum_i(\mathcal{A}(b_i)+\mathcal{A}(\sigma(b_i)))=0$ up to periods. Similarly since $\sigma(Q_j)=Q_j$ we have $\mathcal{A}(Q_j)=-\mathcal{A}(Q_j)$ so that $2\mathcal{A}(Q_j)=0$ modulo periods. This shows that $\mathcal{A}(Q_j)$ is a half–period on the Jacobian torus. Finally we get $\mathcal{A}(D^+)=-\mathcal{A}(D)$ modulo periods so that $D^+=\sigma(D)$ since D is generic.

From this we can understand that the requirement $\psi_k^+(t,P)=\psi_k(t,\sigma(P))$ fixes the constants d_k and e^+ in the expressions (4.35) and (4.36). Let us compare for instance the theta functions $\theta(\mathcal{A}(P)-\mathcal{A}(Q_k)+\mathcal{A}(Q_1)-\zeta_{D^+(t)})$ and $\theta(\mathcal{A}(\sigma(P))+\mathcal{A}(Q_k)-\mathcal{A}(Q_1)-\zeta_{D(t)})$. The sum of the two arguments vanishes modulo periods, because in the hyperelliptic case the vector of Riemann's constants $\mathcal{K}$ is some half–period, hence these two theta–functions have the same zeroes. Applying a similar argument to the other theta–functions and choosing $e^+=-e$, we see that:

$$\psi_k^+(t,P)=\frac{1}{c_k}\,\frac{\theta(-\mathcal{A}(P)+\mathcal{A}(Q_k)-\mathcal{A}(Q_1)-\zeta_{D(t)})}{\theta(-\mathcal{A}(P)-\zeta_{D(t)})}\,\frac{\theta(e-\int_{Q_1}^P\omega)}{\theta(e-\int_{Q_k}^P\omega)} \tag{4.39}$$

Starting from the expressions (4.35) for ψ_k, (4.39) for ψ_k^+ and eq.(4.37) for $\Omega/d\lambda$ one computes c_k according to eq.(4.25). Then the relation $\psi_k^+(t,P)=\psi_k(t,\sigma(P))$ fixes the constant d_k.

To compute the solution of the Neumann model remark that the diagonal element L_{kk} of the Lax matrix reads $L_{kk}(\lambda)=a_k-x_k^2/\lambda^2$. On the other hand using the reconstruction formula $L_{kk}(\lambda)=\sum_i\psi_k(P_i)\mu(P_i)(\psi^{(-1)})_k(P_i)$, where the P_i are the N points above λ. When $\lambda\to 0$ only $\mu(P_\infty)$ diverges as $-1/\lambda^2$ while all the other terms remain finite. Hence:

$$x_k^2=\psi_k(P_\infty)(\psi^{(-1)})_k(P_\infty)$$

Note that this formula implies $\sum_k x_k^2=1$ as it should be in the Neumann model, and moreover this expression is independent of the constant d_k. Inserting the above expressions we immediately find:

$$x_k^2(t)=\alpha_k\frac{\theta(\mathcal{A}(Q_k)-\mathcal{A}(Q_1)+\zeta_{D(t)})\,\theta(\mathcal{A}(Q_k)-\mathcal{A}(Q_1)-\zeta_{D(t)})}{\theta(\mathcal{A}(Q_1)-\zeta_{D(t)})\,\theta(\mathcal{A}(Q_1)+\zeta_{D(t)})}$$

where α_k is given by ratios of theta–functions *completely independent* of the dynamical divisor $D(t)$. It depends only on the geometry of the spectral curve.

It is more convenient to express this result in terms of theta–functions with characteristics. Note that $\mathcal{A}(Q_i)$ are half–periods. By redefining $\zeta_D\to\zeta_D+\mathcal{A}(Q_1)$ one gets rid of $\mathcal{A}(Q_1)$ in the theta–functions. Indeed the first factors in the numerator and denominator get translated by the period $2\mathcal{A}(Q_1)$. This produces exponential factors whose dependence in ζ_D cancel. Similarly in the formula relating ordinary theta–functions and theta–functions with characteristics there is an exponential factor linear in the argument. However due to the particular form of our expression all contributions from D cancel as well, this time between the two factors in the numerator, and separately between the two factors in the denominator. Finally $\mathcal{A}(Q_k)=\int_{P_\infty}^{Q_k}\omega=\Omega\eta_{2k-1}'+\eta_{2k-1}''$ is an even non singular characteristic. Hence one gets

$$x_k^2(t)=\alpha_k'\frac{\theta^2[\eta_{2k-1}](\zeta_{D(t)})}{\theta^2[0](\zeta_{D(t)})}$$

One could in principle evaluate directly the coefficients α'_k by using the very special properties of theta–functions on hyperelliptic curves. However we have a short cut by appealing to the Frobenius formula, only valid on hyperelliptic curves, which states that:

$$\sum_{k=1}^{g+1} \frac{\theta^2[\eta_{2k-1}](z)\ \theta^2[\eta_{2k-1}](0)}{\theta^2[0](z)\ \theta^2[0](0)} = 1$$

so we finally get:

$$x_k^2(t) = \frac{\theta^2[\eta_{2k-1}](z_0+Ut)\ \theta^2[\eta_{2k-1}](0)}{\theta^2[0](z_0+Ut)\ \theta^2[0](0)} \tag{4.40}$$

Here U is obtained by applying eq.(4.30) with $P^{(0)}(\mu,\lambda) = -\lambda\mu$ (recall that for the Neumann model there is only a singularity at $\lambda = 0$ and $M(\lambda) = (-\lambda L(\lambda))_-$, see Chapter [2]). This yields:

$$U_j = \sum_{P_i} \mathrm{Res}_{P_i}(\lambda\mu)\omega_j(P) = \mathrm{Res}_{P_\infty}(\lambda\mu)\omega_j(P) = -\omega_j(P_\infty)$$

where P_i are the N points above $\lambda = 0$, and we recall that the considered expression has only a pole at $P_\infty(\lambda = 0, \mu = \infty)$. The ω_j are the *normalized* abelian differentials.

4.6 Baker–Akhiezer functions.

Baker–Akhiezer functions are special functions with essential singularities on Riemann surfaces which may be used to express the eigenvectors of the linear system.

Definition. *Let $P_1, \cdots P_l$ be points on a Riemann surface Γ of genus g. Let $w_i(P)$, with $w_i(P_i) = 0$, be local parameters around these points. Let $S_i(P) = \sum_{r=-m_i}^{-1} S_{i,r} w_i^r$ be some singular parts around P_i. Let D be a divisor on Γ. A Baker–Akhiezer function $\Psi_{BA}(P)$ defined from these data is a function which is:*
1) meromorphic on Γ outside the points P_i with the divisor of its poles and zeroes satisfying $(\Psi_{BA}) + D > 0$
2) for $P \to P_i$ the product $\Psi_{BA}(P)e^{-S_i(w_i(P))}$ is analytic.

It is important to keep in mind the data involved in the definition of the Baker–Akhiezer functions. First we need a set of punctures P_i on the Riemann surface. Second a set of local parameters $w_i(P)$ allowing to define a set of singular parts S_i in the neighborhood of each puncture. Notice that this definition is not invariant under change of local parameters w_i. Assuming for the moment that such functions exist it is worth stating a few remarks.

Remark 1. If Baker–Akhiezer functions associated to a given set of data exist, they form a vector space. However the sum of two Baker–Akhiezer functions with different singular parts S_i is not a Baker–Akhiezer function.

Remark 2. The ratio of two Baker–Akhiezer functions associated to a given set of singular parts is a meromorphic function. This allows to use standard analysis on Riemann surfaces to study Baker–Akhiezer functions.

Remark 3. Even though Baker–Akhiezer functions are not meromorphic functions, they have the same number of poles and zeroes. This is because the differential form $d(\log f)$ for any Baker–Akhiezer function f is a meromorphic form, without residues at the singular points of f. The sum of its residues is the number of zeroes minus the number of poles of f and has to vanish.

We now give a fundamental formula expressing the Baker–Akhiezer functions in terms of Riemann theta functions. This in particular proves their existence.

Let $\Omega^{(S)}$ be the unique abelian differential of the second kind, normalized with vanishing a-periods, with singular part at the points P_i of the form $dS_i(w_i(P))$. Recall that a differential of second kind is a meromorphic differential with poles of order ≥ 2. See Chapter [15] for more details. Thus, near the points P_i,

$$\Omega^{(S)} = d\left(\sum_{r=-m_i}^{-1} S_{i,r}w_i^r + \text{regular}\right)$$

Denote by $2i\pi U^{(S)}$ the vector of b-periods of $\Omega^{(S)}$. Its g components $U_j^{(S)}$, $j = 1, \cdots, g$ are:

$$U_j^{(S)} = \frac{1}{2i\pi}\oint_{b_j} \Omega^{(S)} \tag{4.41}$$

Proposition. *If $D = \sum_{i=1}^{g} \nu_i$ is a generic divisor of degree g, the following expression defines a Baker–Akhiezer function with D as divisor of poles:*

$$\Psi_{BA}(P) = \text{const. } \exp\left(\int_{P_0}^{P} \Omega^{(S)}\right) \frac{\theta(\mathcal{A}(P) + U^{(S)} - \zeta)}{\theta(\mathcal{A}(P) - \zeta)} \tag{4.42}$$

Here $\zeta = \mathcal{A}(D) + \mathcal{K}$ with $\mathcal{K}$ is the vector of Riemann's constants and $\mathcal{A}$ denotes the Abel map with based point P_0, cf. eq.(4.33).
For a generic divisor D of degree $\geq g$, the dimension of the vector space of Baker–Akhiezer functions is equal to $deg(D) - g + 1$. In particular for $\deg D = g$ the above formula gives the unique Baker–Akhiezer functions having poles at D up to a constant.

Proof. It is enough to check that the function defined by the formula (4.42) is well-defined (i.e., it does not depend on the path of integration between P_0 and P) and has the desired analytical properties. Indeed when P describes some a–cycle, nothing happens because the theta functions are periodic under such shifts and $\Omega^{(S)}$ is normalized. If P describes the b_j–cycle the quotient of theta functions is multiplied by $\exp(-2i\pi U_j^{(S)})$ (see Chapter [15]) while the exponential factor changes by $\exp(2i\pi U_j^{(S)})$. To show the second part let us take the ratio of two such functions. We get a meromorphic function with $d = \deg D$ poles and by the Riemann-Roch theorem the dimension of the space of such functions is $d - g + 1$. ∎

It is worth noticing that generically we get a non trivial Baker–Akhiezer function with only g poles, while to get a non trivial meromorphic function we need generically $(g + 1)$ poles.

To understand why Baker–Akhiezer functions arise naturally in the construction of the eigenvectors, let us consider the unnormalized eigenvector $\Psi_{un}(t, P)$ whose time evolution is governed by the equation

$$\partial_t\Psi_{un}(t, P) = -M(\lambda)\Psi_{un}(t, P) \tag{4.43}$$

Of course the normalized eigenvector $\Psi(t, P)$ and the unnormalized eigenvector $\Psi_{un}(t, P)$ are related by multiplication by a scalar function: $\Psi_{un}(t, P) = f(t, P)\Psi(t, P)$ and $(\Psi_{un})_1(t, P) = f(t, P)$. Taking the first component of eq.(4.43) one gets:

$$\dot{f} = -Cf \quad \text{with} \quad C = \sum_j M_{1j}\psi_j$$

where C is the same object appearing in eq.(4.28).

Let us describe the singularities of $f(t, P)$. Remark that $C(P)$ has poles where Ψ has poles or at points above the poles λ_k of $M(\lambda)$ (recall that in general the poles of $M(\lambda)$ are a subset of the poles of $L(\lambda)$).

Consider first the points $P_{k,i}$, $i = 1, \cdots, N$ above a point λ_k. In the vicinity of $P_{k,i}$ we have:

$$\begin{aligned}\dot{f} &= -(M\Psi)_1 f = -\left(\left((P^{(k)}(L,\lambda))_- + \text{regular}\right)\Psi\right)_1 f \\ &= -\left(\left(P^{(k)}(L,\lambda) + \text{regular}\right)\Psi\right)_1 f = -\left(\left(P^{(k)}(\mu,\lambda)\right)_- + \text{regular}\right) f\end{aligned}$$

where we have used the fact that Ψ is an eigenvector of $L(\lambda)$ with eigenvalue μ and $\Psi_1 = 1$. The quantity $\dot{f}/f$ has poles at points (λ, μ) such that $\left(P^{(k)}(\mu,\lambda)\right)_- \neq 0$. The projection $()_-$ is computed using the local parameter λ. Notice that $\left(P^{(k)}(\mu(\lambda),\lambda)\right)_-$ is independent of time, therefore $f(t,P)$ has an essential singularity at $P_{k,i}$ of the form:

$$f(t,P) = e^{-t\left(P^{(k)}(\mu,\lambda)\right)_-} \times \text{regular}$$

Let us now consider the poles of C coming from Ψ. First at $\lambda = \infty$, while Ψ has poles, M vanishes so that C is regular, and nothing special happens for f. At a point $\nu(t)$ of the dynamical divisor $D(t)$ we have $\psi_i \sim \alpha_i(t)/(\lambda - \nu(t))$ and $C \sim r(t)/(\lambda - \nu(t))$. Comparing the second order pole in both sides of eq.(4.28) we find $r(t) = \dot{\nu}(t)$. Thus $\partial_t \log f = \partial_t \log(\lambda - \nu) + \text{regular}$, showing that $f(t,P)$ *vanishes* at the points of the dynamical divisor $D(t)$.

Finally let us remark that the poles of $f(t,P)$ are independent of time. Indeed assuming that $f(t,P)$ has a pole of order k at a point $\gamma(t)$, which by the previous argument is not a pole of C, the order of the poles on both sides of the equation $\dot{f} = -Cf$ are different if $\dot{\gamma} \neq 0$. Considering the solution such that $f(t=0) = 1$ we see that the divisor of its zeroes is the dynamical divisor $D(t)$ and the divisor of its poles is $D(0)$. Moreover f has essential singularities at the points $P_{k,i}$ with prescribed singular parts, hence f is the unique Baker–Akhiezer function with these essential singularities and the g poles corresponding to the divisor $D(0)$.

$$f(t,P) = \exp\left(\int_{P_0}^{P} t\,\Omega^{(M)}\right) \frac{\theta(\mathcal{A}(P) + tU^{(M)} - \zeta_{D(0)})}{\theta(\mathcal{A}(P) - \zeta_{D(0)})} \tag{4.44}$$

The linear t dependence in the theta–function in the numerator of this equation arises from the form of the singular exponential and the requirement that there is no monodromy. This provides another quick proof of the linearization of the flow on the Jacobian. Indeed the zeroes of $f(t,P)$ are at the divisor $D(t)$ such that:

$$\mathcal{A}(D(t)) - \mathcal{A}(D_0) = t\,U^{(M)} = \frac{t}{2i\pi}\oint_b \Omega^{(M)} \tag{4.45}$$

On the other hand we know that $\mathcal{A}(D(t))$ is given by eq.(4.34). These equations coincide since:

$$\frac{1}{2i\pi}\oint_{b_j} \Omega^{(M)} = -\sum_k \sum_{i=1}^{N} \mathrm{Res}_{P_{k,i}}\left(\omega_j P^{(k)}(\mu,\lambda)\right)$$

with ω_j the normalized abelian differentials. This is a consequence of Riemann's bilinear indentity, cf Chapter [15].

We finally give the expression of the components of the vector $\Psi_{un}(t,P) = f(t,P)\Psi(t,P)$.

$$(\Psi_{un})_k(t,P) = d_k\ \exp\left(t\int_{P_0}^{P}\Omega^{(M)}\right)\ \frac{\theta(\mathcal{A}(P) - \mathcal{A}(Q_k) + \mathcal{A}(Q_1) - \zeta_{D(t)})}{\theta(\mathcal{A}(P) - \zeta_{D(0)})}\frac{\theta(e + \int_{Q_1}^{P}\omega)}{\theta(e + \int_{Q_k}^{P}\omega)}$$

In the product $f(t,P)\psi_k(t,P)$ the zeroes of f cancel the dynamical poles of ψ_k which are replaced by the *constant* poles of f.

Remark. As explained in Chapter [2], different functions $P^{(k)}(L,\lambda)$ correspond to different dynamical flows. Therefore, different abelian differentials of second kind with poles at the points above λ_k correspond to different time evolutions. In other words all the different dynamics are encoded in the singular differentials $\Omega^{(M)}$.

4.7 Tau–functions.

We now whish to relate the formula for the Baker–Akhiezer function to the so-called tau function. Let us consider for simplicity the case of only one singular point P_∞, and let z be a local parameter for a point P in the vicinity of P_∞ such that $z(P_\infty) = 0$. The general case is similar but much more cumbersome to present.

In order to be able to describe at once all possible singular parts, we introduce an infinite set of formal variables t_k. Denote by $\Omega^{(k)}$ the normalized differential of the second kind with singular part $d(z^{-k})$ at P_∞ and denote by $U^{(k)}$ its b-periods,

$$U_j^{(k)} = \frac{1}{2i\pi} \oint_{b_j} \Omega_\infty^{(k)}$$

Proposition. *Let $\psi_{BA}(P)$ be the Baker–Akhiezer function with divisor D of poles of degree g and singular part $\xi(t,z) = \sum_k t_k z^{-k}$ at P_∞ (this should be understood in the sense of formal series), normalized such that:*

$$\psi_{BA}(P) = e^{\xi(t,z)}\,(1 + O(z)), \quad z \sim 0$$

Define $t - [z] = \left\{t_k - \frac{1}{k} z^k\right\}$. Then we have in the vicinity of P_∞:

$$\psi_{BA}(P) = e^{\xi(t,z)}\;\frac{\tau\,(t - [z])}{\tau(t)} \tag{4.46}$$

The function $\tau(t)$ may be expressed in terms of the Riemann's theta function,

$$\tau(t) = e^{\alpha(t)+\beta(t,t)}\;\theta(\mathcal{A}(P_\infty) + \sum_k t_k U^{(k)} - \zeta) \tag{4.47}$$

Here $\alpha(t)$ and $\beta(t,t)$ are a linear and a quadratic form in the times t respectively defined below, and $\zeta = \mathcal{A}(D) + \mathcal{K}$, with $\mathcal{A}(D)$ Abel's map and $\mathcal{K}$ the vector of Riemann's constants.

Proof. From formula (4.42) the normalized Baker–Akhiezer function can be written:

$$\psi(P) = e^{\xi(t,z)}\;\exp\left[\sum_k t_k \left(\int_{P_\infty}^{P} (\Omega^{(k)} - d\,z^{-k})\right)\right] \frac{\theta(\mathcal{A}(P) + \sum_k t_k U^{(k)} - \zeta)\theta(\mathcal{A}(P_\infty) - \zeta)}{\theta(\mathcal{A}(P) - \zeta)\theta(\mathcal{A}(P_\infty) + \sum_k t_k U^{(k)} - \zeta)} \tag{4.48}$$

Indeed eq.(4.42) contains $\int_{P_0}^P \Omega$ which can be written $\xi(t,z) + \int_{P_\infty}^P (\Omega - d\xi) + C^{\rm st}$ where $(\Omega - d\xi) = \sum_k t_k(\Omega^{(k)} - d\,z^{-k})$ is regular at P_∞, and our expression is obviously correctly normalized.

On the other hand if τ is assumed to be of the form (4.47),

$$\frac{\tau(t - [z])}{\tau(t)} = e^{-\alpha([z]) - 2\beta(t,[z]) + \beta([z],[z])}\,\frac{\theta(\mathcal{A}(P_\infty) + \sum_k (t_k - \frac{z^{-k}}{k})U^{(k)} - \zeta)}{\theta(\mathcal{A}(P_\infty) + \sum_k t_k U^{(k)} - \zeta)}$$

We need to compare these two equations. Recall that near P_∞, $\int_{P_\infty}^P (\Omega^{(k)} - d\,z^{-k}) = b_k(z)$ is regular. So we first choose $\alpha(t)$ and $\beta(t,t)$ such that:

$$\beta(t,[z]) = -\frac{1}{2}\sum_k t_k b_k(z) \quad \text{and} \quad \alpha([z]) = \beta([z],[z]) + \log\left(\frac{\theta(\mathcal{A}(P) - \zeta)}{\theta(\mathcal{A}(P_\infty) - \zeta)}\right)$$

This defines α and β uniquely. This choice takes care of the exponential prefactor and two of the theta functions in eq.(4.48). To deal with other two theta functions in (4.48) we Taylor expand the Abel map

$\mathcal{A}(P) - \mathcal{A}(P_\infty)$ around P_∞. Writing $\omega_j = \sum_{i=0}^\infty c_i^{(j)}(z)\,dz$ in the vicinity of the point P_∞ and Taylor expanding using Riemann bilinear identities, cf. Chapter [15], one deduces:

$$\mathcal{A}_j(P) - \mathcal{A}_j(P_\infty) = \sum_{i=1}^\infty c_{i-1}^{(j)} \frac{z^i}{i} = -\frac{1}{2\pi i} \sum_{k=1}^\infty \frac{z^k}{k} \oint_{b_j} \Omega^{(k)} = -\sum_{k=1}^\infty \frac{z^k}{k} U_j^{(k)}$$

Inserting this relation in the theta function gives:

$$\theta\left(\mathcal{A}(P) + \sum_k t_k U^{(k)} - \zeta\right) = \theta\left(\mathcal{A}(P_\infty) + \sum_k (t_k - \frac{z^k}{k}) U^{(k)} - \zeta\right) \tag{4.49}$$

Gathering all this we obtain formula (4.46). ■

The formula (4.46) relating Baker–Akhiezer functions to tau-functions is usually called Sato's formula. It may easily be generalized for several punctures, each puncture giving rise to one tau–function. We see that the Baker–Akhiezer function provides a global meaning to Sato's formula.
SYMMETRIE DES COEFFICIENTS b_{kl}.

4.8 Linearization and the factorization problem.

4.9 Symplectic form.

In Chapter [2] we showed that the Lax matrix can be interpreted as an element of a coadjoint orbit, and this allowed us to introduce a natural symplectic structure on it. More precisely, using the local diagonalization theorem, we showed that $L(\lambda) = L_0 + \sum_k L_k(\lambda)$ where each L_k may be written as:

$$L_k(\lambda) = \left(g_k \cdot (A_k)_- \cdot g_k^{-1}\right)_-$$

with $(A_k)_-$ diagonal matrices, cf. eq.(4.59). Moreover, locally around λ_k, $L(\lambda)$ can be diagonalized as:

$$L(\lambda) = g_k\ A_k\ g_k^{-1}$$

Both matrices A_k and g_k depend on λ, A_k is singular at λ_k but g_k is regular. The variables $(A_k)_-$ are not dynamical but specify the coadjoint orbit. The dynamical variables are the matrix elements of the matrices g_k. The symplectic form inherited from the coadjoint orbit structure is:

$$\omega = \sum_k \mathrm{Res}_{\lambda_k}\ \mathrm{Tr}\left((A_k)_- g_k^{-1}\delta g_k \wedge g_k^{-1}\delta g_k\right)\ d\lambda$$

We have seen that the Lax pair description of a dynamical system naturally provides coordinates on phase space, namely g independent action variables F_i which parametrize the spectral curve Γ, and g points $\nu_i = (\lambda_i, \mu_i)$ on the spectral curve, which we called the dynamical divisor. It is important to express the symplectic form in terms of these coordinates. The phase space appears as a fibered space whose base is the space of moduli of the spectral curve, explicitly described as coefficients of the equation $R(\lambda,\mu) = 0$ of the spectral curve, and the fiber at a given R is the Jacobian of the curve $R = 0$. On this space we introduce a differential δ which varies the dynamical variables F_i, λ_i, μ_i subjected to the constraint $R_{\{F_i\}}(\lambda_i, \mu_i) = 0$.

We will need an auxiliary fiber bundle above the same base whose fiber is $\Gamma \times \mathrm{Jac}(\Gamma)$. We extend δ to this space by keeping the previous definition on the $\mathrm{Jac}(\Gamma)$ part and on the Γ part, we differentiate any function of F_i, λ, μ with $R_{\{F_i\}}(\lambda, \mu) = 0$ keeping λ *constant.* This definition makes sense because the bundle of curves is given by the family of equations $R(\lambda, \mu) = 0$ where the coefficients of R depend on the moduli which parametrize the base space. This provides a universal definition of the meromorphic function λ on the whole family of curves. So differentiating on the bundle of curves keeping λ constant provides a horizontal direction, i.e. a connection. Explicitly for a function $f(P; F_i)$ we take λ as a local parameter, then $\delta f = \sum_i \partial_{F_i} f \delta F_i$. At a branch point however the local parameter is μ and we have:

$$\delta f = \partial_\mu f \delta\mu + \sum_i \partial_{F_i} f \delta F_i, \quad \text{with } \delta\mu = -\frac{1}{\partial_\mu R_{\{F_i\}}(\lambda,\mu)} \sum_i \partial_{F_i} R_{\{F_i\}}(\lambda,\mu)\delta F_i \tag{4.50}$$

Note that at a branch point of the covering $(\lambda, \mu) \to \lambda$ we have $\partial_\mu R_{\{F_i\}}(\lambda, \mu) = 0$, hence the differential δf acquires a pole even though f is regular. Remark however that if f depends rationally on λ and the F_i, δf is regular at the branch points.

Recall that at each point $P(\lambda, \mu)$ on Γ is defined a column eigenvector $\Psi(P)$ of the Lax matrix up to normalization, and that we have defined a dual line eigenvector $\Psi^{(-1)}(P)$ such that $< \Psi^{(-1)}(P), \Psi(P) >= 1$. This allows to define a three–form K on our extended fiber bundle. We regard it as a one-form on Γ whose coefficients are two–forms on phase space.

$$\begin{aligned} K &= K_1 + K_2 + K_3 \\ K_1 &= < \Psi^{(-1)}(P)\delta L(\lambda) \wedge \delta\Psi(P) > \, d\lambda \\ K_2 &= < \Psi^{(-1)}(P)\delta\mu \wedge \delta\Psi(P) > \, d\lambda \\ K_3 &= \delta\,(\log \partial_\mu R) \wedge \delta\mu\, d\lambda \end{aligned} \tag{4.51}$$

Of course $\Psi(P)$ is defined knowing the dynamical divisor up to multiplication by a diagonal matrix independent of P and we assume that some normalization has been chosen.

Proposition. *Let us define the two–form on phase space: $\omega = \sum_{k,i} \mathrm{Res}_{P_{k,i}} K$ where $P_{k,i}$ are the points above the poles λ_k of $L(\lambda)$. Then we have:*

$$\omega = 2 \sum_{i=1}^{g} \delta\lambda_i \wedge \delta\mu_i$$

hence ω is a symplectic form on phase space.

<u>Proof</u>. The sum of the residues of K seen as a form on Γ vanishes. The poles of K are located at four different places, first the dynamical poles of Ψ, then the poles at the $P_{k,i}$ coming from L and μ, next the poles above $\lambda = \infty$ coming from Ψ and $d\lambda$, and finally the poles at the branch points of the covering coming from the poles of $\Psi^{(-1)}$ and from eq.(4.50).

Let us compute the residues at the dynamical poles $(\nu_1, \cdots, \nu_g)$. We write the coordinates of these points as: $\nu_i = (\lambda_i, \mu_i)$ for $i = 1, \cdots g$. Near such a point we can choose λ as a universal local parameter and $\Psi = 1/(\lambda - \lambda_i) \times \Psi_{\mathrm{reg}}$ hence:

$$\delta\Psi = \frac{\delta\lambda_i}{\lambda - \lambda_i} (\Psi + O(1))$$

Since $(L - \mu)\Psi = 0$ and $\Psi^{(-1)}(L - \mu) = 0$, we have $(\delta L - \delta\mu)\Psi + (L - \mu)\delta\Psi = 0$. Multiplying by $\Psi^{(-1)}$ we get $\Psi^{(-1)}\delta L\Psi = \delta\mu$, therefore:

$$\mathrm{Res}_{\nu_i} K_1 = \delta\mu|_{\nu_i} \wedge \delta\lambda_i$$

Here $\delta\mu$ is to be seen as a meromorphic function on Γ given by eq.(4.50). However varying $R(\lambda_i, \mu_i) = 0$ we obtain:

$$\delta\mu|_{\nu_i} = \delta\mu_i + \left.\frac{\partial_\lambda R}{\partial_\mu R}\right|_{\nu_i} \delta\lambda_i$$

and the second term does not contribute to the wedge product.

The contribution of K_2 is exactly the same. Finally, K_3 is regular at ν_i and does not contribute to the residue at this point. So we finally get:

$$\mathrm{Res}_{\nu_i} K = 2\delta\mu_i \wedge \delta\lambda_i \tag{4.52}$$

We now show that there are no residues at the branch points due to the proper choice of K_2 and K_3. Let us look at the term K_1. At a branch point b, $\Psi^{(-1)}$ has a simple pole, δL is regular, $\delta\Psi$ has a simple pole due to eq.(4.50) and the form $d\lambda$ has a simple zero, hence the considered expression has a simple pole at b. To compute its residue it is enough to keep the polar part in $\delta\Psi$, i.e. to replace $\delta\Psi$ by $\partial_\mu \Psi \delta\mu$ (recall that μ is a good local parameter around b). We get:

$$\begin{aligned} \mathrm{Res}_b K_1 &= \mathrm{Res}_b < \Psi^{(-1)} \delta L \partial_\mu \Psi > \wedge \delta\mu \, d\lambda \\ &= \mathrm{Res}_b < \Psi^{(-1)} (\delta L - \delta\mu) \partial_\mu \Psi > \wedge \delta\mu \, d\lambda \end{aligned}$$

where in the last equation we have used the antisymmetry of the wedge product to replace δL by $\delta L - \delta\mu$. Using again the eigenvector equation $(L-\mu)\Psi = 0$, and varying the point (λ, μ) on the curve around b one gets

$$(L-\mu)\partial_\mu \Psi = \Psi - \frac{d\lambda}{d\mu} \frac{dL}{d\lambda} \Psi \tag{4.53}$$

where $d\lambda/d\mu$ vanishes at the branch point. We then differentiate with δ and get:

$$\begin{aligned} \mathrm{Res}_b < \Psi^{(-1)} (\delta L - \delta\mu) \partial_\mu \Psi > \wedge \delta\mu \, d\lambda &= \mathrm{Res}_b < \Psi^{(-1)} \delta\Psi > \wedge \delta\mu \, d\lambda \\ &\quad - \mathrm{Res}_b < \Psi^{(-1)} \delta \left(\frac{d\lambda}{d\mu} \frac{dL}{d\lambda} \Psi \right) > \wedge \delta\mu \, d\lambda \end{aligned}$$

The first term exactly cancels the term $\mathrm{Res}_b K_2$. The second term gives a non–vanishing contribution

$$\mathrm{Res}_b \frac{\delta\mu_b}{\mu - \mu_b} \wedge \delta\mu \, d\lambda$$

Indeed let us call $\zeta = (d\lambda/d\mu)(dL/d\lambda)\Psi$ which vanishes at $b = (\lambda_b, \mu_b)$. Then $\zeta = (\mu - \mu_b)\zeta_1$ hence $\delta\zeta = -\delta\mu_b/(\mu - \mu_b)\zeta + \delta\mu \, \zeta_2 + \zeta_3$ with ζ_3 regular. The second term does not contribute due to the antisymmetry of the wedge product and the third term has no residue. Using eq.(4.53) we have $< \Psi^{(-1)} \frac{d\lambda}{d\mu} \frac{dL}{d\lambda} \Psi >= 1$ yielding the above formula. This contribution is exactly cancelled by the contribution of K_3.

We now compute the residues above $\lambda = \infty$. Recall that we consider a reduced Hamiltonian system under the action of diagonal matrices. To fix this action on can normalize the eigenvectors at ∞ so that $\psi_i(Q_j) = \lambda\delta_{ij} + O(1)$ for $i, j = 2, \cdots, N$. Notice that $L = L_0 + O(1/\lambda)$ where L_0 is non–dynamical so $\delta L_0 = 0$, and that $\mu = a_i + O(1/\lambda)$ around Q_i hence $\delta L = \delta\mu = O(1/\lambda)$. Moreover $\Psi^{(-1)}$ vanishes at Q_i and $d\lambda$ has a double pole. Altogether K_1 and K_2 are regular at Q_i since $(\delta\Psi)(Q_i) = O(1)$ due to the normalization condition. Finally K_3 is also regular since on the sheet $\mu = \mu_i(\lambda)$ one can write $\partial_\mu R = \prod_{j \neq i}(\mu_i - \mu_j)$ yielding $\delta \log \partial_\mu R = O(1/\lambda)$. All this shows that K has no residues above $\lambda = \infty$. ■

Proposition. *The symplectic form ω is given by:*

$$\omega = 2 \sum_{i=1}^{g} \delta\lambda_i \wedge \delta\mu_i = 2 \sum_k \mathrm{Res}_{\lambda_k} \, \mathrm{Tr} \left((A_k)_- g_k^{-1} \delta g_k \wedge g_k^{-1} \delta g_k \right) d\lambda \tag{4.54}$$

where (λ_i, μ_i), $i = 1, \cdots g$, are the coordinates of the points of the dynamical divisor D.

<u>Proof</u>.

Let us compute the residues at the poles λ_k of K_1, where only L_k contributes. Recall the local diagonalization theorem of Chapter (2), cf. eq(2.8), which allows to write the Lax matrix as $L = g_k A_k g_k^{-1}$ around $\lambda = \lambda_k$. Thus locally around λ_k we may identify the matrix $\widehat{\Psi}(\lambda)$ with g_k. More precisely by eq.(4.17), we have $\widehat{\Psi}(\lambda) = g_k d_k$ and $\widehat{\Psi}^{-1}(\lambda) = d_k^{-1} g_k^{-1}$ with d_k a diagonal matrix. The residues are obtained by integrating over small circles surrounding each of the N points $P_{k,i}$ above λ_k. We can choose these small circles so that they project on the base λ on a single small circle surrounding λ_k. Then we get

$$
\begin{aligned}
\sum_{i=1}^{N} \mathrm{Res}_{P_{k,i}} K_1 &= \sum_{i=1}^{N} \frac{1}{2i\pi} \oint_{C_{k,i}} < \Psi^{(-1)}(P_i)\delta L(\lambda) \wedge \delta\Psi(P_i) > \, d\lambda \\
&= \frac{1}{2i\pi} \oint_{C_k} \mathrm{Tr} \left(\widehat{\Psi}^{-1}(\lambda)\delta L(\lambda) \wedge \delta\widehat{\Psi}(\lambda) d\lambda \right) \qquad (4.55)
\end{aligned}
$$

where we used the fact that $\widehat{\Psi}^{-1}(\lambda)$ is equal to the matrix whose lines are the vectors $\Psi^{(-1)}(P_i)$. The trace has been reconstructed in eq.(4.55) because $\Psi(P_i)$, $i = 1, \cdots N$, form a basis of eigenvectors. Using the identification of $\widehat{\Psi}(\lambda)$ in terms of g_k gives:

$$
\begin{aligned}
\mathrm{Res}_{\lambda_k} K_1 &= \mathrm{Res}_{\lambda_k} \mathrm{Tr} \left(d_k^{-1} g_k^{-1} \; \left(\delta g_k (A_k)_- g_k^{-1} - g_k (A_k)_- g_k^{-1} \delta g_k g_k^{-1} \right) \wedge (\delta g_k d_k + g_k \delta d_k) \right) \, d\lambda \\
&= -2 \, \mathrm{Res}_{\lambda_k} \mathrm{Tr} \left((A_k)_- g_k^{-1} \delta g_k \wedge g_k^{-1} \delta g_k \right) \, d\lambda \\
&\quad + \mathrm{Res}_{\lambda_k} \mathrm{Tr} \left(g_k^{-1} \delta g_k [(A_k)_-, \delta d_k d_k^{-1}] \right) \, d\lambda \qquad (4.56)
\end{aligned}
$$

The last term vanishes because it involves the commutator of two diagonal matrices. To compute the residues of K_2 at λ_k we remark that $\delta\mu$ is regular on all sheets above λ_k. This is because due to the form of $L(\lambda)$, we have $\widehat{\mu} = (A_k)_- + \text{regular}$. Since $(A_k)_-$ characterizes the coadjoint orbit and is not dynamical, one has to take $\delta(A_k)_- = 0$. It follows immediately that K_2 is regular. To compute the residue of K_3 we note that if $\partial_\mu R$ has a pole of some order m at $P_{k,i}$ and can be written $\partial_\mu R = c(\lambda)/(\lambda - \lambda_k)^m$ where $c(\lambda)$ is regular and non vanishing, we get $\delta \left(\log \partial_\mu R \right) = \delta \log c(\lambda)$ which is regular, since $\delta\lambda = 0$ and $\delta\lambda_k = 0$. Hence K_3 has no residue again because $\delta\mu$ is regular. ■

This proposition means that the coordinates (λ_i, μ_i) of the point ν_i of the dynamical divisor are canonical coordinates.

Remark 1. This result shows the nice interplay between the analytical and the group–theoretical approaches to integrable systems. We are able to show that (λ_i, μ_i) are canonical coordinates *only* using the fact that L parametrizes a coadjoint orbit, specified by *constant* matrices $(A_k)_-$ and L_0.

Remark 2. In practice, to perform this calculation, one has to compute at each pole of L the quantity $\omega_k = \mathrm{Res}_{\lambda_k} \mathrm{Tr} \left(\delta L \wedge \delta\widehat{\Psi} \, \widehat{\Psi}^{-1} \right) d\lambda$. In the residue at λ_k, only δL_k appears. From eq.(4.17) one has

$$
\delta L_k = [(\delta\widehat{\Psi} \, \widehat{\Psi}^{-1})_{\mathrm{jet}}, L_k] \qquad (4.57)
$$

where $()_{\mathrm{jet}}$ is the expansion to order $n_k - 1$. This equation determines $(\delta\widehat{\Psi} \, \widehat{\Psi}^{-1})_{\mathrm{jet}}$ up to a quantity commuting with L_k. It is easy to see that $\omega_k = \mathrm{Res}_{\lambda_k} \mathrm{Tr} \left([\delta\widehat{\Psi} \, \widehat{\Psi}^{-1}, L_k] \wedge \delta\widehat{\Psi} \, \widehat{\Psi}^{-1} \right) d\lambda$ is not affected by this ambiguity using the antisymmetry of the wedge product and the cyclicity of the trace.

Example. We consider the example of the Neumann model. The above analysis can be applied to this model, except for one feature. As we have already stressed there is no residual action of diagonal matrices on L, hence one has to pay special attention to the residues above $\lambda = \infty$. The sum of residues at the poles above $\lambda = \infty$ is $\mathrm{Res}_\infty \mathrm{Tr} \left(\delta L \wedge \delta \widehat{\Psi}\, \widehat{\Psi}^{-1} \right) d\lambda$. One can see that $\delta \widehat{\Psi}\, \widehat{\Psi}^{-1}$ is regular at ∞, while $d\lambda$ has a pole of order 2. Since δL has a zero of order 1, one generally gets a residue. However $\delta \widehat{\Psi}\, \widehat{\Psi}^{-1}$ is constrained by

$$\delta L = [\delta \widehat{\Psi}\, \widehat{\Psi}^{-1}, L] + \widehat{\Psi} \delta \widehat{\mu} \widehat{\Psi}^{-1}$$

At ∞ the second term vanishes and L tends to D, hence the order 0 term in $\delta \widehat{\Psi}\, \widehat{\Psi}^{-1}$ is diagonal. The leading term in δL is $1/\lambda\ \delta J$ and has no diagonal element, consequently the considered trace vanishes. Similarly the term involving K_2 has no residue because $\delta \widehat{\mu}$ vanishes to order 2 at $\lambda = \infty$.

To compute the residue at $\lambda = 0$ (a second order pole of L) we remark that the jet:

$$\delta \widehat{\Psi}\, \widehat{\Psi}^{-1} = \delta X\, {}^t X - X \delta^t X + \lambda \left(\delta Y\, {}^t X + X \delta^t Y - \delta X\, {}^t Y - Y \delta^t X \right) + O(\lambda^2)$$

solves eq.(4.57) and has the correct symmetry properties. One gets the expression of the symplectic form of the Neumann model in terms of the dynamical divisor:

$$\omega = 4 \sum_{i=1}^{N} \delta y_i \wedge \delta x_i = 2 \sum_{j=1}^{N-1} \delta \lambda_j \wedge \delta \mu_j$$

4.10 Separation of variables.

Let us call F_i, $i = 1, \cdots g$ the action variables which are also the moduli of the spectral curve. For fixed $\{F_i\}$, we have seen that the motion takes place on the Jacobian $\mathrm{Jac}\,(\Gamma(\{F_i\}))$. When varying initial conditions, the $\{F_i\}$ will eventually vary and we get a foliation of the (complexified) phase space in terms of the Jacobian tori of $\Gamma(\{F_i\})$. So we are back to the situation described in Liouville's theorem, cf Chapter(1). Let us check that the symplectic form do vanish when we restrict ourself to one of the tori of the foliation. We view $\mathrm{Jac}\,(\Gamma)$ as the g^{th} symmetric product Γ^g. Inverting the equation $R_{\{F_i\}}(\lambda_k, \mu_k) = 0$ one has $\mu_k = \mu_k(\{F_i\}, \lambda_k)$ which depends on λ_k only and not on the other λ's and the symplectic form can be written as:

$$\omega = \delta\alpha = \sum_{k=1}^{g} \delta \mu_k \wedge \delta \lambda_k = \sum_{i,k} (\frac{\partial \mu_k}{\partial F_i})\ \delta F_i \wedge \delta \lambda_k, \quad \alpha = \sum_{k=1}^{g} \mu_k \wedge \delta \lambda_k$$

So, if we restric ourselves to a level manifold $F_i = f_i$ we have $\partial F_i = 0$ and $\omega|_f = 0$, hence the level manifolds are Lagrangian subvarieties.

Let us explain why the variables (λ_k, μ_k) form a set of separated variables. The construction is similar to the method used for proving the Liouville theorem. Consider the function

$$S(\{F_i\}, \{\lambda_k\}) = \int_{m_0}^{m} \alpha = \sum_k \int_{\lambda_0}^{\lambda_k} \mu_k d\lambda$$

The integration contour is drawn on the level manifold $F_i = f_i$. Just as in the Liouville case, this function does not depend on local variations of the integration path. It is explicitly separated since it can be written as a sum of functions each depending on only one variable λ_k:

$$S(\{F_i\}, \{\lambda_k\}) = \sum_k S_k(\{F_i\}, \lambda_k)$$

Since $\frac{\partial S_k}{\partial \lambda_k} = \mu_k$ and since the point (λ_k, μ_k) belongs to the curve $\Gamma : R(\lambda, \mu) = 0$, each function S_k is solution of the differential equation,

$$R(\lambda_k, \frac{\partial S_k}{\partial \lambda_k}) = 0.$$

Of course the coefficients of the function $R(\lambda, \mu)$ depend on the values of the integrals of motion. This is an equation of the form

$$\left(\frac{\partial S_k}{\partial \lambda_k}\right)^N + \sum_{q=1}^{N-1} r_q(\lambda_k)\left(\frac{\partial S_k}{\partial \lambda_k}\right)^q = 0$$

where the coefficients $r_q(\lambda)$ are defined in eq.(4.4). This is just a Hamilton-Jacobi type equation in which the variables have been explicitely separated. In other words if one changes variables from the initial dynamical variables to the variables F_i and λ_k the Hamilton–Jacobi equation separates.

Remark. It is sometime advantageous to consider λ_k as a function of μ_k. Defining $S' = \sum_k \int_{\mu_0}^{\mu_k} \lambda_k d\mu_k$, we get $\lambda_k = \frac{\partial S'}{\partial \mu_k}$. The relation between S and S' is simply a Legendre transform: $S' = \sum_k \mu_k \lambda_k - S$.

4.11 Action-angle variables.

So far, we dealt with complexified dynamical sytems. We found that the phase space of this complexified system can be viewed as a fibration where the base is the moduli space of the spectral curve and the fiber is the Jacobian of the spectral curve corresponding to the specific values of the moduli parameters.

This is very similar to the situation depicted in the Liouville theorem, but the Liouville tori are real tori of dimension g, while the Jacobian has real dimension $2g$. We need to choose a real slice of this complex phase space. This can be done as follows. On Γ we choose a canonical basis of $2g$ cycles a_k, b_k. The cycles a_k are non intersecting. Once they are chosen, we can adapt the basis of Abelian differentials of the first kind ω_k such that they are normalized by $\oint_{a_k} \omega_l = \delta_{kl}$.

A connected component of the real slice can be defined by restricting the g points of the dynamical divisor D to move along these g non-intersecting cycles, each point on a different cycle. This obviously is a product of g real circles. One has to be aware that the real slice has several connected components, some of them are constructed by this recipe with other choices of basic cycles, but more subtle situation can arise. Finally explicit models correspond to specific cycles and not only to homology classes of cycles.

The angle variables are then given by

$$\theta_k = \sum_{i=1}^{g} \int_{\nu_0}^{\nu_i} \omega_k = \sum_{i=1}^{g} \int_{\lambda_0}^{\lambda_i} g_k(\lambda) d\lambda \tag{4.58}$$

where the integration paths are taken along the cycles a_k and the abelian forms ω_k ar written in terms of the local parameter λ in the form $\omega_k = g_k(\lambda)d\lambda$. With these assumptions the angles have real periods.

The angles being defined one may find the conjugated action variables.

Proposition. *The symplectic form reads*

$$\omega = \sum_{i=1}^{g} \delta\mu_i \wedge \delta\lambda_i = \sum_{i=1}^{g} \delta I_i \wedge \delta\theta_i$$

where the action variable I_k canonically conjugated to θ_k is

$$I_k = \oint_{a_k} \mu d\lambda \tag{4.59}$$

Proof. Let us define the quantity I_k by eq.(4.59). The main observation is that the 1-form $\frac{\partial}{\partial I_k}(\mu d\lambda)$ is an abelian differential of the first kind. Here we differentiate keeping λ constant and varying the moduli. This produces poles at the branch points in $\partial\mu/\partial I_k$ which are cancelled by the zeroes of $d\lambda$. Moreover $\mu d\lambda$ is a meromorphic 1-form on Γ with poles where μ has poles, i.e. above the poles of the Lax matrix $L(\lambda)$. In the vicinity of such a point μ can be expanded in terms of the diagonal matrix $A_k(\lambda)$ defined in eq.(2.8). But we already noticed that the singular part of the matrix $A_k(\lambda)$ is to be considered as a set of parameters defining the model, and not as dynamical variables. Hence their variations are zero, and so $\partial/\partial_{I_k}(\mu d\lambda)$ is regular. It follows that this form can be decomposed on the basis of holomorphic abelian differentials: $\frac{\partial}{\partial I_k}(\mu\partial\lambda) = \sum_i \alpha_{ki}\omega_i$. To find the coefficients α_{ki}, we integrate both sides on the cycles A_l. On the left hand side, we get

$$\oint_{A_l} \frac{\partial}{\partial I_k}\mu d\lambda = \frac{\partial}{\partial I_k}\oint_{A_l} \mu d\lambda = \frac{\partial}{\partial I_k} I_l = \delta_{kl}$$

On the right hand side we get α_{kl} because the first kind differentials ω_i are normalised. Hence

$$\frac{\partial}{\partial I_k}\,\mu = g_k(\lambda), \text{ or } \frac{\partial}{\partial I_k}\,(\mu d\lambda) = \omega_k, \quad \text{yielding} \quad \delta(\mu d\lambda) = \sum_k \delta I_k \wedge \omega_k$$

By definition of the angular variables given in eq.(4.58) one has:

$$\delta\theta_i = \sum_{j=1}^{g} g_i(\lambda_j)\delta\lambda_j + \sum_{j=1}^{g}\int_{\lambda_0}^{\lambda_j} \frac{\partial g_i(\lambda)}{\partial I_k} d\lambda\,\delta I_k$$

So we finally obtain with $\omega_i(\nu_j) = g_i(\lambda_j)\delta\lambda_j$:

$$\begin{aligned}
\omega &= \sum_i \delta(\mu_i d\lambda_i) = \sum_{i,j} \delta I_i \wedge \omega_i(\nu_j) \\
&= \sum_i \delta I_i \wedge \left(\delta\theta_i - \sum_j \int_{\lambda_0}^{\lambda_j} \sum_k \frac{\partial g_i(\lambda)}{\partial I_k} d\lambda\,\delta I_k \right) \\
&= \sum_i \delta I_i \wedge \left(\delta\theta_i - \sum_j \int_{\lambda_0}^{\lambda_j} \sum_k \frac{\partial^2 \mu}{\partial I_i \partial I_k} d\lambda\,\delta I_k \right) = \sum_i \delta I_i \wedge \delta\theta_i
\end{aligned}$$

This shows that the I_i are canonically conjugated to the θ_i. ∎

4.12 Infinite dimensional systems.

In the field theory case, we can use the previous constructions to find particular classes of solutions to the field equations, called finite zone solutions. The equations we have to solve are the first order differential system:

$$(\partial_x + U(\lambda))\Psi = 0 \tag{4.60}$$

$$(\partial_t + V(\lambda))\Psi = 0 \tag{4.61}$$

whose compatibility conditions are equivalent to the field equations. The situation is very different as compared to the finite dimensional case. As we saw previously in Chapter [2], the analog of the spectral curve is

$$\det(T(\lambda) - \mu) = 0 \tag{4.62}$$

where $T(\lambda)$ is the monodromy matrix of the linear system (4.60,4.61). This equation does not define an algebraic curve of finite genus. This had to be expected since in field theory we need an infinite number of action variables, which is incompatible with the finite genus of the spectral curve. Thus we cannot directly apply the previous construction.

However, if we restrict our ambition to find only particular solutions to eq.(4.60, 4.61), then the knowledge acquired in the previous sections of this chapter becomes very useful. In fact, these two equations are exactly of the type of eq.(4.28) whose solution was built in terms of Baker–Akhiezer functions. One can adapt this construction to solve simultaneously eqs.(4.60, 4.61).

The idea consists in interpreting the two equations (4.60,4.61) as evolution equations with respect to two different "times" for a system with finite number of degrees of freedom associated to some Lax matrix $L(\lambda)$, that we shall proceed to construct. This Lax matrix should satisfy:

$$\begin{aligned} [\partial_x + U(\lambda), L(\lambda)] &= 0 \\ [\partial_t + V(\lambda), L(\lambda)] &= 0 \end{aligned} \tag{4.63}$$

One first method to find such a Lax matrix is to consider the higher order flows as described in eq.(2.62) of Chapter [2]. One gets a family of compatible linear equations $(\partial_{t_i} + V_i)\Psi = 0$ for $i = 1, 2, 3, \cdots$ where we have identified $t_1 = x,\ V_1 = U$ and $t_2 = t,\ V_2 = V$. Since these equations are compatible they satisfy a zero–curvature condition:

$$F_{ij} \equiv \partial_{t_i} V_j - \partial_{t_j} V_i + [V_i, V_j] = 0, \quad \forall i, j = 1, \cdots, \infty$$

We now look for particular solutions which are stationary for some given time t_n, i.e. $\partial_{t_n} V_i = 0$ for all i. The zero curvature conditions F_{ni} reduce to a system of Lax equations:

$$\frac{dL}{dt_i} = [L, M_i], \quad i = 1, \cdots, \infty \quad \text{with } L = V_n,\ M_i = V_i$$

This is an integrable hierarchy for a finite–dimensional dynamical system described by the Lax matrix L. Taking n larger and larger the genus of the corresponding spectral curve usually increases and we get families of solutions involving more and more parameters.

An other way of constructing such Lax matrices consists in making use of the local abelianization theorem of Chapter [2]. This theorem tell us that, around each pole λ_k of $U(\lambda)$ and $V(\lambda)$, the solutions of eqs.(4.63) can be diagonalized as:

$$\begin{aligned} \partial_x + U(\lambda) &= g_k(\partial_x + A_k)g_k^{-1} \\ \partial_t + V(\lambda) &= g_k(\partial_t + B_k)g_k^{-1} \end{aligned} \tag{4.64}$$

with A_k and B_k diagonal matrices. Thus locally around λ_k, we may identify the wave function Ψ with g_k up to an exponential factor:

$$\Psi = g_k \ \exp(-xA_k - tB_k)$$

This allows to choose a Lax matrix of the general form described above having poles at the same points λ_k as $U(\lambda)$ and $V(\lambda)$, i.e.,

$$L(\lambda) = L_0 + \sum_k L_k(\lambda), \quad \text{with } L_k(\lambda) = \left(g_k (D_k)_- g_k^{-1}\right)_-$$

with $(D_k)_-$ some fixed diagonal matrix. Recall that $g_k(\lambda)$ can be expressed locally in terms of matrix elements of U and V hence $L(\lambda)$ is a local quantity. Varying the choices of the matrices $(D_k)_-$ will result in different spectral curves and a different set of solutions to eqs.(4.60, 4.61).

EST-CE VRAI ? On doit avoir $L(\lambda) = g_k A_k g_k^{-1}$ et donc

$$\left(g_k A_k g_k^{-1}\right)_+ = 0$$

References

[1] B. Dubrovin, V. Matveev, S. Novikov *"Non–linear equations of Korteweg–de Vries type, finite–zone linear operators, and abelian varieties"* Russian Math. Surveys 31 (1976) 59–146.

[2] P. van Moerbeke, D. Mumford. *" The Spectrum of difference operators and algebraic curves."* Acta.Math. 143 (1979) p.93-154.

[3] A.G. Reyman, M.A. Semenov-Tian-Shansky. *"Reduction of Hamiltonian Systems, Affine Lie Algebras and Lax Equations."* Inventiones Mathematicae 54, (1979) p. 81-100.

[4] M. Adler, P. van Moerbeke. *"Linearization of Hamiltonian Systems, Jacobi Varieties and Representation Theory."* Advances in Mathematics 38, (1980) p.318-379.

[5] A.G. Reyman, M.A. Semenov-Tian-Shansky. *"Reduction of Hamiltonian Systems, Affine Lie Algebras and Lax Equations II."* Inventiones Mathematicae 63, (1981) p. 425-432.

[6] D. Mumford **"Tata Lectures on Theta"** Vol. I and II Birkhauser (1983–1984).

[7] B.A. Dubrovin, I.M. Krichever, S.P. Novikov. *"Integrable Systems I"*. Encyclopedia of Mathematical Sciences, Dynamical Systems IV Springer (1990), p. 173-281.

[8] I. M. Krichever, D. H. Phong *"On the integrable geometry of soliton equations and N=2 supersymmetric gauge theories "* J.Diff.Geom. 45 (1997) 349–389.

[9] I. M. Krichever *"Elliptic solutions to difference non-linear equations and nested Bethe ansatz equations"* solv-int/9804016

Chapter 5

The closed Toda chain.

The aim of this chapter is to construct the general solution of the closed Toda chain by means of the analytical method. This example illustrates the general statement of Chapter [4] that the equations of motion are linearized on the Jacobian of the spectral curve.

We shall do this in two ways. The first one, based on a $(n+1)\times(n+1)$ Lax matrix, follows closely the general method explained in Chapter [4]. The second method, which is based on a 2×2 monodromy matrix, uses short cuts which in general are not available, and makes contact with Sklyanin's method of separation of variables.

We take advantage of this particular example to discuss the reality conditions, which are generally quite subtle.

5.1 The model.

We consider a chain of $(n+1)$ particles with positions q_i and momenta p_i with equations of motion given by:

$$\dot{q}_i = p_i, \quad \dot{p}_i = 2e^{2(q_{i-1}-q_i)} - 2e^{2(q_i - q_{i+1})}, \quad i = 1\cdots n+1$$

The fact that the chain is closed is implemented by setting $(p_0, q_0) = (p_{n+1}, q_{n+1})$ and $(p_{n+2}, q_{n+2}) = (p_1, q_1)$ which gives sense to the above equations for $i = 1$ and $i = n+1$. Alternatively one can view the particles as sitting on a circle. This is an Hamiltonian system with canonical Poisson brackets $\{p_i, q_j\} = \delta_{ij}$ and Hamiltonian:

$$H = \tfrac{1}{2}\sum_i p_i^2 + \sum_i \exp 2(q_i - q_{i+1})$$

Obviously the system has translational symmetry $q_i \to q_i + a$ and one can remove the center of mass motion by imposing the two conditions $\sum_{i=1}^{n+1} p_i = 0$ and $\sum_{i=1}^{n+1} q_i = 0$, so we are left with a phase space of dimension $2n$.

In contrast to the open Toda chain which is associated to the finite dimensional Lie algebra sl_{n+1}, the closed Toda chain is associated to the infinite dimensional loop algebra $\widetilde{sl}_{n+1} \equiv sl_{n+1} \otimes C(\lambda, \lambda^{-1})$, whose rank is $(n+1)$, cf. Chapter [16] for more details. The generators $E_{\pm\alpha_i}$ associated to the simple roots are represented by $E_{\alpha_i} = E_{i,i+1}$, $E_{-\alpha_i} = E_{i+1,i}$, for $i = 1, \cdots n$ together with $E_{\alpha_{n+1}} = \lambda E_{n+1,1}$, and $E_{-\alpha_{n+1}} = \lambda^{-1} E_{1,n+1}$, where $E_{j,k}$ are the $(n+1)\times(n+1)$ canonical matrices. The elements of the Cartan subalgebra are then represented by $(n+1)\times(n+1)$ diagonal traceless matrices.

We apply the general construction of the Toda chains, cf. Chapter [3]. The Lax matrix for the closed Toda chain in the loop algebra is:

$$
\begin{aligned}
L(\lambda) &= \sum_{i=1}^{n+1} p_i E_{ii} + \sum_{i=1}^{n+1} a_i (E_{\alpha_i} + E_{-\alpha_i}) \\
M(\lambda) &= \sum_{i=1}^{n+1} a_i (E_{\alpha_i} - E_{-\alpha_i})
\end{aligned}
\tag{5.1}
$$

where the coefficients a_i are given by $a_i = \exp{(q_i - q_{i+1})}$, $i = 1, \cdots, n$, and $a_{n+1} = \exp{(q_{n+1} - q_1)}$. Note that the quantity $(q_i - q_{i+1}) = \alpha_i(q)$ where q is the traceless diagonal matrix $\sum_i q_i E_{ii}$ and α_i is the simple root associated to the root vector $E_{i,i+1}$. The quantities a_i satisfy the condition $\mathcal{A} = \prod_{i=1}^{n+1} a_i = 1$.

Explicitly the Lax matrix reads:

$$
L(\lambda) = \begin{pmatrix}
p_1 & a_1 & 0 & & \cdots & & \lambda^{-1} a_{n+1} \\
a_1 & p_2 & a_2 & & \cdots & & 0 \\
\vdots & & \ddots & & & & \vdots \\
0 & & a_{i-1} & p_i & a_i & & 0 \\
\vdots & & & & \ddots & & \vdots \\
0 & & \cdots & & a_{n-1} & p_n & a_n \\
\lambda a_{n+1} & & \cdots & & 0 & a_n & p_{n+1}
\end{pmatrix}
\tag{5.2}
$$

$$
M(\lambda) = \begin{pmatrix}
0 & a_1 & 0 & & \cdots & & -\lambda^{-1} a_{n+1} \\
-a_1 & 0 & a_2 & & \cdots & & 0 \\
\vdots & & \ddots & & & & \vdots \\
0 & & -a_{i-1} & 0 & a_i & & 0 \\
\vdots & & & & \ddots & & \vdots \\
0 & & \cdots & & -a_{n-1} & 0 & a_n \\
\lambda a_{n+1} & & \cdots & & 0 & -a_n & 0
\end{pmatrix}
\tag{5.3}
$$

Note that the matrices $L(\lambda)$ and $M(\lambda)$ are regular except at $\lambda = 0$ or $\lambda = \infty$.

The equations of motion are indeed given by the Lax equation: $\dot{L}(\lambda) = [L(\lambda), M(\lambda)]$, as one can check easily by an explicit computation. Notice that the correct equations of motion for q_1 and q_{n+1} are obtained thanks to the λ–dependent terms in $L(\lambda)$ and $M(\lambda)$. As usual, the Lax equation ensures that $\mathrm{Tr}\,(L^p(\lambda))$ are conserved quantities, in particular the Hamiltonian reads $H = 1/2\,\mathrm{Tr}\,(L^2(\lambda))$ which is independent of λ.

Following the general construction of Toda models, the Poisson bracket structure of the Lax matrix can be written in terms of an r–matrix. This implies in particular that the closed Toda chain is integrable. There exists two natural r–matrices $r^{\pm}$ which can be computed according to the general formula (3.25) in Chapter [3]:

$$
r^+(\lambda, \lambda') = \rho(\lambda) \otimes \rho(\lambda') \left(\sum_i H_i \otimes H_i + 2 \sum_{\alpha > 0} E_\alpha \otimes E_{-\alpha} \right)
$$

where $\rho(\lambda)$ means the matricial representation as given above with spectral parameter λ.

The positive roots in $\widetilde{sl}_{n+1}$ are the $\lambda^n E_{ij}$ with $i < j$; $n \geq 0$ or $i \geq j$; $n > 0$, with corresponding negative roots the $\lambda^{-n} E_{ji}$. We immediately get:

$$
r^+(\lambda, \lambda') = \frac{\lambda + \lambda'}{\lambda' - \lambda} \sum_i E_{ii} \otimes E_{ii} + \frac{2}{\lambda' - \lambda} \left(\lambda' \sum_{i<j} E_{ij} \otimes E_{ji} + \lambda \sum_{i>j} E_{ij} \otimes E_{ji} \right)
$$

This expression is valid for $|\lambda| < |\lambda'|$. The formula for $r^-(\lambda, \lambda')$ is the same as for r^+ but valid for $|\lambda| > |\lambda'|$. So we consider in general the rational function $r_{12}(\lambda, \lambda') = r_{12}^{\pm}(\lambda, \lambda')$, and notice that $r_{12}(\lambda, \lambda') = -r_{21}(\lambda', \lambda)$ so that

$$\{L_1(\lambda), L_2(\lambda')\} = [r_{12}(\lambda, \lambda'), L_1(\lambda) + L_2(\lambda')]$$

Remark that in agreement with the general construction we have $M(\lambda) = M_+(\lambda) + M_-(\lambda)$ where $M_\pm = 1/2R_\pm(L)$ that is:

$$\begin{aligned} M_+ &= \frac{1}{2}\sum_i p_i H_i + \sum_{i=1}^{n+1} a_i E_{\alpha_i} \\ M_- &= -\frac{1}{2}\sum_i p_i H_i - \sum_{i=1}^{n+1} a_i E_{-\alpha_i} \end{aligned}$$

Of course $L(\lambda) = M_+(\lambda) - M_-(\lambda)$ so the Lax equation can be written:

$$\dot{L}(\lambda) = [L(\lambda), M(\lambda)] = [L(\lambda), 2M_+(\lambda)] = [L(\lambda), 2M_-(\lambda)]$$

5.2 The spectral curve.

The spectral curve Γ is the smooth algebraic curve defined by:

$$\Gamma \quad : \quad \det(L(\lambda) - \mu) = 0 \tag{5.4}$$

For a fixed λ, this is the equation for the eigenvalues of $L(\lambda)$. By expanding the determinant we see that it is of the form:

$$R(\lambda, \mu) = \mathcal{A}(\lambda + \lambda^{-1}) - \Delta(\mu) = 0 \tag{5.5}$$

where $\mathcal{A} = \prod_{i=1}^{n+1} a_i$, and $\Delta(\mu) = \mu^{n+1} + \cdots$ is a polynomial of degree $(n+1)$, hence the spectral curve is an hyperelliptic curve since it can be written as

$$\omega^2 = t^2(\mu) - 1, \quad \text{with} \quad \begin{cases} \omega = \lambda - t(\mu) \\ \\ t(\mu) = \Delta(\mu)/2\mathcal{A} \end{cases} \tag{5.6}$$

Let us compute the genus of the curve Γ. The polynomial $t^2(\mu)$ is of degree $2n+2$ and generically the equation $t^2(\mu) - 1 = 0$ has no double roots, so the genus of the curve Γ is $g = n$. This is equal to the number of degrees of freedom, i.e. half the dimension of phase–space. In the following, we shall always assume that we are in this generic situation.

Let us see how we can recover this result from the general analysis in Chapter [4]. We look at the curve as an $(n+1)$-sheeted cover of the λ plane. We recall the Riemann-Hurwitz formula for computing the genus:

$$2g - 2 = (2g_0 - 2)(n+1) + \nu$$

where $g_0 = 0$ and ν is the number of branch points. To find ν, we count the zeroes of $\frac{\partial}{\partial\mu}R(\lambda, \mu) = \frac{\partial}{\partial\mu}\Delta(\mu)$. This is a polynomial of degree n in μ, independent of λ. Hence we have n zeroes at finite distance. To each value of μ correspond two values of λ, and the total contribution of the branching points at finite distance is $2n$. We now look at $\mu = \infty$ which corresponds to $\lambda = \infty$ or $\lambda = 0$. For $\lambda = \infty$, we have $\lambda \simeq \mu^{n+1}$, which means that it is a branch point of order $n+1$, contributing n to ν. Similarly, $\lambda = 0$ is a branch point of order $n+1$, contributing also n to ν. Adding everything, we find $\nu = 4n$, and $2g - 2 = -2n - 2 + 4n = 2n - 2$, hence $g = n$.

We will call P^+ and P^- the two points above $\lambda = \infty$ and $\lambda = 0$ respectively. In the neighbourhood of $P^\pm$ the local parameter is μ^{-1} and the curve is parametrized by $(\mu, \lambda \simeq \mu^{\pm(n+1)})$. More precisely we have:

$$P^+ \quad : \quad \lambda = \frac{\mu^{n+1}}{\mathcal{A}} \Big(1 - \mu^{-1} \, (\sum_{j=1}^{n+1} p_j) + O(\mu^{-2}) \Big) \tag{5.7}$$

$$P^- \quad : \quad \lambda = \mathcal{A} \, \mu^{-n-1} \Big(1 + \mu^{-1} \, (\sum_{j=1}^{n+1} p_j) + O(\mu^{-2}) \Big) \tag{5.8}$$

The terms in the expansion of λ are found by direct expansion of eq.(5.5).

5.3 The eigenvectors.

The equation (5.4) is the condition for μ to be an eigenvalue of $L(\lambda)$. Therefore, to each point P on Γ one can associate an eigenvector $\Psi(P)$:

$$(L(\lambda) - \mu)\Psi(P) = 0, \quad P = (\lambda, \mu) \in \Gamma$$

here $\Psi(P)$ is a $(n+1)$ dimensional vector.

Writing explicitly the equation for $\Psi(P)$, we find the following system of linear equations:

$$\begin{aligned} p_1\psi_1 + a_1\psi_2 + \lambda^{-1}a_{n+1}\psi_{n+1} &= \mu\psi_1 \\ a_{i-1}\psi_{i-1} + p_i\psi_i + a_i\psi_{i+1} &= \mu\psi_i \\ \lambda a_{n+1}\psi_1 + a_n\psi_n + p_{n+1}\psi_{n+1} &= \mu\psi_{n+1} \end{aligned} \tag{5.9}$$

Extending the definition of the coefficients a_i, p_i by periodicity $a_{i+n+1} = a_i, p_{i+n+1} = p_i$, we can rewrite this set of equations as the eigenvalue problem of a second order difference operator $\mathcal{D}$ for quasi-periodic wave functions.

$$(\mathcal{D}\Psi)_i \equiv a_{i-1}\psi_{i-1} + p_i\psi_i + a_i\psi_{i+1}$$

This operator is a discrete version of a Schroedinger operator with periodic potential. Eqs.(5.9) are then equivalent to:

$$(\mathcal{D}\Psi)_i = \mu\psi_i, \quad \text{with} \quad \psi_{i+n+1} = \lambda\psi_i \tag{5.10}$$

Thus, ψ is a Bloch wave for the difference operator $\mathcal{D}$ with a Bloch momentum λ. In the following, we choose the normalization condition:

$$\psi_0(P) = 1 \quad \text{for all} \quad P \in \Gamma$$

or alternatively $\psi_{n+1}(P) = \lambda$. This is slightly different from the convention of Chapter [2] where we normalized $\psi_1(P) = 1$ but will prove more convenient in the following.

We need to know the analyticity properties of Ψ at infinity.

Proposition. *Let us normalize $\psi_0(P) = 1$. Then, at the points P^+ and P^- above $\lambda = \infty$ and $\lambda = 0$, the eigenvector $\Psi(P)$ behaves as:*

$$\psi_i(P) = e^{q_i - q_0}\mu^i \Big(1 - \mu^{-1} \, (\sum_{j=0}^{i-1} p_j) + O(\mu^{-2}) \Big), \quad \text{for} \quad P \sim P^+ \tag{5.11}$$

$$\psi_i(P) = e^{-q_i + q_0}\mu^{-i} \Big(1 + \mu^{-1} \, (\sum_{j=1}^{i} p_j) + O(\mu^{-2}) \Big), \quad \text{for} \quad P \sim P^- \tag{5.12}$$

where the q_i are the Toda position parameters ($q_0 = q_{n+1}$, $p_0 = p_{n+1}$).

Proof. It is easy to check that this is consistent with eq.(5.9). The result then follows by the uniqueness of the eigenvector. ∎

Comparing to the general theory, cf. eq.(4.14) in Chapter [4], we expect $g + (n+1) - 1 = 2n$ poles for the eigenvector. We see that we have a fixed pole of order n at P^+. There remains $g = n$ poles at finite distance. These are the *dynamical poles*. We shall denote by $\gamma_1, \cdots, \gamma_n$ their positions. Note that their number equals the number of degrees of freedom. Recall from Chapter [4], that these dynamical poles contain all the relevant information to reconstruct the eigenvector.

We now consider the time evolution of the eigenvector. It is given by the Lax equation

$$\frac{d}{dt}L(\lambda) = [L(\lambda), M(\lambda)]$$

This evolution leaves the spectral curve invariant, but the eigenvector evolves according to:

$$\frac{d}{dt}\Psi(t,P) = -\left(M(\lambda) - C(t,P)\,\mathbf{1}\right)\ \Psi(t,P)$$

with $C(t,P)$ some scalar, determined by imposing the condition $\psi_0(t,P) = 1$ or equivalently $\psi_{n+1}(t,P) = \lambda$ which yields $C(t,P) = \sum_j M_{n+1,j}\psi_j$.

Alternatively one can use the natural time evolution

$$\frac{d}{dt}\Psi(t,P) = -M(\lambda)\ \Psi(t,P) \tag{5.13}$$

but then, at $t \neq 0$ the eigenvector $\Psi(t,P)$ has an essential singularity at infinity, i.e. it is a Baker–Akhiezer function. It has n poles at finite distance independent of time, while the essential singularity at the points $P^\pm$ is given by:

Proposition. *Let $\Psi(t,P)$ be the eigenvector evolving according to eq.(5.13).*

$$\psi_i(t,\mu) \;=\; e^{q_i(t)}\, e^{-\mu t}\mu^i(1 + O(\mu^{-1})), \quad P \to P^+ \tag{5.14}$$

$$\psi_i(t,\mu) \;=\; e^{-q_i(t)}\, e^{\mu t}\mu^{-i}(1 + O(\mu^{-1})), \quad P \to P^- \tag{5.15}$$

Proof. The time evolution of the eigenvector is described by $(L\psi)_i = \mu\psi_i$ and

$$\frac{d}{dt}\psi_i = -\left(M\Psi\right)_i = \big((L - 2M_+)\Psi\big)_i = \big((-L - 2M_-)\Psi\big)_i$$

Since the matrix elements of the operator M are regular outside the points $P^\pm$, we see that the poles of ψ_i outside these points do not evolve with time.
Let us see what happens near P^+. Since $\psi_i(t,P)$ is also a solution of $(L(\lambda) - \mu)\Psi(t,P) = 0$, it has the form

$$\psi_i(t,P) = f(t,P)e^{q_i(t)-q_0(t)}\mu^i(1 + O(\mu^{-1}))$$

with $q_i(t)$ now depending on t. There can be an extra multiplicative factor $f(t,P)$ independent of n since we relax the condition $\psi_0 = 1$ for $t \neq 0$. Writing $\dot\psi_{n+1} = \dot f\lambda = -\big((L + 2M_-)\Psi\big)_{n+1} = -\mu\lambda f + p_{n+1}\lambda f + 2a_n\psi_n$ we get $\dot f = (-\mu + p_{n+1} + O(\mu^{-1}))f$ hence $f = \exp(-\mu t + q_0(t))(1 + O(\mu^{-1}))$ using $p_{n+1} = \dot q_0$. The analysis of the behavior near P^- is similar, using $M = 2M_+ - L$. ∎

5.4 Reconstruction formula.

Let the algebro–geometrical data be specified by the $g = n$ points $\gamma_1, \cdots, \gamma_g$ at finite distance together with the two points $P^{\pm}$ at infinity. From the previous section the divisor of the component $\psi_i(P)$ of $\Psi(P)$ is generically:

$$(\psi_i) = -\mathcal{D} = -\gamma_1 - \cdots - \gamma_g + i\,P^- - i\,P^+ \tag{5.16}$$

By the Riemann–Roch theorem there exists a unique meromorphic function, up to a proportionality constant, having divisor greater or equal to $-\mathcal{D}$ since $\deg \mathcal{D} = g$. This function is not constant since it vanishes at P^-. We can fix the proportionality constant (up to a sign) by requiring that the coefficients of $\mu^{\mp i}$ at $P^{\pm}$ are inverse to each other. Denote these coefficients by $e^{\pm(q_i - q_0)}$. The function $\psi_i(P)$ will have for all $i \geq 0$ the form:

$$\psi_i(\mu) = e^{q_i - q_0}\, \mu^i(1 - \mu^{-1}\, \xi_i^+ + O(\mu^{-2})), \quad P \to P^+ \tag{5.17}$$

$$\psi_i(\mu) = e^{-q_i + q_0}\, \mu^{-i}(1 - \mu^{-1}\, \xi_i^- + O(\mu^{-2})), \quad P \to P^- \tag{5.18}$$

Here $\xi_i^{\pm}$ are just Taylor coefficients. Of course, since $\psi_i(P)$ is a meromorphic function it also possesses g extra zeroes. We now show that these properties imply that these functions are solutions of the equation (5.10):

Proposition. *Let $\psi_i(P)$ be the unique meromorphic function having simple poles at the g points γ_i's, a pole of order i at P^+, and a zero of order i at P^-, and normalized as in eq. (5.17, 5.18). Then,*
i) The $\psi_i(P)$'s satisfy the Schroedinger equations $(\mathcal{D} - \mu)\Psi = 0$, eq.(5.10), with coefficients a_i, p_i given by:

$$a_i = \exp\left(q_i - q_{i+1}\right), \quad p_i = \xi_{i+1}^+ - \xi_i^+ \tag{5.19}$$

ii) The functions $\psi_i(P)$ are quasi–periodic:

$$\psi_{n+1+i}(P) = \lambda(P)\, \psi_i(P) \tag{5.20}$$

Therefore, the $\psi_i(P)$ are the components of the eigenvector of the Lax matrix $L(\lambda)$.

Proof. Let us consider the function $\widehat{\psi}_i(P) = \left((\mathcal{D} - \mu)\Psi\right)_i(P)$. Since the coefficients of $(\mathcal{D} - \mu)$ are regular outside $P^{\pm}$, $\widehat{\psi}_i$ possesses g poles at the γ_i's. Consider now the behavior at $P^{\pm}$. We have:

$$\begin{aligned}
\left((\mathcal{D} - \mu)\Psi\right)_i &= \mu^{i+1}\left(a_i e^{q_{i+1} - q_0} - e^{q_i - q_0}\right) \\
&\quad -\mu^i\left(a_i e^{q_{i+1} - q_0}\xi_{i+1}^+ - e^{q_i - q_0}(p_i + \xi_i^+)\right) + O(\mu^{i-1}), \quad P \to P^+, \\
\left((\mathcal{D} - \mu)\Psi\right)_i &= \mu^{-i+1}\left(a_{i-1}e^{q_0 - q_{i-1}} - e^{q_0 - q_i}\right) + O(\mu^{-i}), \qquad P \to P^-
\end{aligned}$$

If we choose the coefficients a_i and p_i as in eqs.(5.19), the function $(\mathcal{D} - \mu)\psi_i$ has a pole of order $(i - 1)$ at P^+, and a zero of order i at P^-. Therefore the divisor of $\widehat{\psi}_i(P)$ is greater than $-D' = -\gamma_1 - \cdots - \gamma_g + i\,P^- - (i - 1))P^+$ and we have $\deg D' = g - 1$. Thus, by the Riemann–Roch theorem, $(\mathcal{D} - \mu)\Psi = 0$. Notice that the coefficients a_i and p_i are choosen in order to decrease the order of the pole at P^+ by one unit.
Let us now prove the periodicity relation: $\psi_{n+1+i}(P) = \lambda\psi_i(P)$. We apply the Riemann-Roch theorem to each of these functions. The divisor of the function λ is $(\lambda) = (n + 1)P^- - (n + 1)P^+$. The poles at finite distance of ψ_{n+1+i} and ψ_i are located at the same positions $\gamma_1, \cdots, \gamma_g$. Their divisors are:

$$\begin{aligned}
(\psi_{n+1+i}) &\geq -\gamma_1 - \cdots - \gamma_g + (n + 1 + i)P^- - (n + 1 + i)P^+ \\
(\lambda\psi_i) = (\psi_i) + (\lambda) &\geq -\gamma_1 - \cdots - \gamma_g + i\,P^- - i\,P^+ + (n + 1)P^- - (n + 1)P^+
\end{aligned}$$

Notice that these divisors are both greater than $-D'' = -\gamma_1 - \cdots - \gamma_g + (n+1+i)P^- - (n+1+i)P^+$ and that $\deg D'' = g$. By the Riemann-Roch theorem, there is only one function satisfying this property and the two functions are therefore proportional. The proportionality constant is determined by comparing the behavior at P^+ or P^-, and is found to be equal to one. ∎

Let us now consider the functions at $t \neq 0$. We define for all integer $i \geq 0$ the function $\psi_i(t,P)$ as the Baker–Akhiezer function with g poles at the same positions $\gamma_1, \cdots, \gamma_n$ as above, and with essential singularities at $P^\pm$ given by:

$$\psi_i(t,\mu) = e^{q_i(t)}\, e^{-\mu t}\, \mu^i (1 - \mu^{-1}\, \xi_i^+(t) + \cdots), \quad P \to P^+ \tag{5.21}$$

$$\psi_i(t,\mu) = e^{-q_i(t)}\, e^{\mu t} \mu^{-i} (1 - \mu^{-1}\, \xi_i^-(t) + \cdots), \quad P \to P^- \tag{5.22}$$

with $q_i(t)$ normalisation parameters obtained by requiring that the expansions at P^+ and P^- start with inverse coefficients. Then the Taylor coefficients $\xi_i^\pm(t)$ are fixed. By Riemann-Roch theorem this function exists and is unique. Indeed, if two functions satisfy these properties, their ratio is a meromorphic function with g poles because the divisor (5.16) is of degree $-g$. Thus this ratio is a constant and the two functions are equal since they are both normalized by eqs.(5.21,5.22). Moreover we have $\psi_{n+1+i}(t,P) = \lambda \psi_i(t,P)$ by the same method as above. Taking into account the expansions eq.(5.7,5.8) of λ and the fact that $\sum_i p_i = 0$ in the center of mass system, it follows that the Taylor coefficients $\xi_i^\pm(t)$ are periodic, i.e. $\xi_{i+n+1}^\pm(t) = \xi_i^\pm(t)$.

Proposition. *The Baker-Akhiezer functions defined above satisfy the eigenvalue equation and the evolution equation:*

$$(L(\lambda) - \mu \mathbf{1})\Psi(t,P) = 0$$

$$\frac{d}{dt}\Psi(t,P) = -M(\lambda)\Psi(t,P)$$

with $M(\lambda)$ defined in eq.(5.1) with $a_i(t) = \exp\left(q_i(t) - q_{i+1}(t)\right)$.

Proof. The proof of the eigenvalue equation and the quasi–periodicity property (5.20) is the same as for the initial eigenvector at $t = 0$.
To prove the evolution equation let us consider for $i = 1, \cdots, n$ the expression:

$$E_i \equiv \frac{d}{dt}\psi_i + (a_i \psi_{i+1} - a_{i-1}\psi_{i-1})$$

Since the poles of the ψ_i's at finite distance are independent of time, and the a_i's are constant on the Riemann surface, E_i has the same g poles at finite distance as ψ_i. Its behavior at infinity is easily obtained as:

$$E_i = -e^{-\mu t} e^{q_i} \mu^i (\xi_{i+1}^+ - \xi_i^+ - \dot{q}_i) + O(\mu^{i-1}) \quad P \to P^+ \tag{5.23}$$

$$E_i = e^{\mu t} e^{-q_i} \mu^{-i} (\xi_{i-1}^- - \xi_i^- - \dot{q}_i) + O(\mu^{-i+1}), \quad P \to P^-$$

By Riemann-Roch theorem E_i is proportional to ψ_i, and we write $E_i = d_i(t)\psi_i$. Using the quasi–periodicity property eq.(5.20) we can restate this result as:

$$\dot{\Psi} = -M\Psi + d\Psi \tag{5.24}$$

where M is the matrix given in eq.(5.3) and d is the time–dependent diagonal matrix $d = \mathrm{Diag}\,(d_1, \cdots, d_{n+1})$ which is constant on the Riemann surface.
Deriving with respect to time the relation $(L(\lambda) - \mu)\Psi = 0$ we get using eq.(5.24) $(\dot{L} - [L,M] + [L,d])\Psi = 0$.

For any value of λ which is not a branch point of the covering $(\lambda,\mu)\to\lambda$ we have $n+1$ independent vectors $\Psi(\lambda,\mu_k)$ for which this equation is true, hence we get the matrix equation $\dot{L}(\lambda)-[L(\lambda),M(\lambda)]=-[L(\lambda),d]$, which remains true for all values of λ by analytic continuation. In particular it is true when $\lambda\to 1/\lambda$. Note that ${}^tL(\lambda)=L(1/\lambda)$, ${}^tM(\lambda)=-M(1/\lambda)$, so taking the transpose of the above equation evaluated at $1/\lambda$ and comparing it to the original equation we get $[L(\lambda),d]=0$. This implies that d is proportional to the identity matrix, $d=\delta(t)I$. Then the equations (5.23) are solved by $\dot{q}_i=\xi^+_{i+1}-\xi^+_i+\delta(t)$. In the center of mass system $\sum_{i=0}^n q_i=0$ so that $\delta(t)=0$. Notice that the matrix d disappears from the Lax equation. ■

This result shows that the definition of the eigenvectors $\Psi(t,P)$ as Baker-Akhiezer functions specified above solves the dynamics of the closed Toda chain. To get explicit expressions for $q_i(t)$ we need explicit formulae for the Baker–Akhiezer functions. This is done using Riemann's theta functions.

Let us fix a canonical set of cycles (a_i,b_j) on Γ, an basis point Q_0 on Γ and a set of holomorphic abelian differential ω_j, dual to the a-cycles (see Chapter [15]). Let $d\Omega^{(i)}$ be the meromorphic differential analytic on Γ outside the punctures $P^\pm$, with the following normalization conditions:

$$\begin{aligned}\oint_{a_k} d\Omega^{(i)} &= 0,\\ d\Omega^{(i)}(P) &= \pm(\mu^{i-1}+O(\mu^{-2}))d\mu,\quad P\to P^\pm\end{aligned}\tag{5.25}$$

The notation assumes that some multivalued primitives $\Omega^{(i)}(P)=\int_{Q_0}^P d\Omega^{(i)}$ have been chosen for these differentials. Since the local parameter around $P^\pm$ is $1/\mu$, $d\Omega^{(i)}$ has poles of order $(i+1)$ at the points $P^\pm$ at infinity. Remark that $d\Omega^{(i)}$, $i\geq 1$, are abelian differentials of the second kind, whereas $d\Omega^{(0)}$ is a normalized abelian differential of the third kind.

Let us also define for each i the g dimensional vectors $U^{(i)}$ whose components are the b-periods of the forms $d\Omega^{(i)}$:

$$U_k^{(i)}=\frac{1}{2\pi i}\oint_{b_k} d\Omega^{(i)}\tag{5.26}$$

Given these data, the Baker–Akhiezer functions defined above are expressed as follows:

Proposition. *Let $\mathcal{A}(P)$ be the Abel map, $\mathcal{A}_k(P)=\int_{Q_0}^P\omega_k$. Then the Baker-Akhiezer function $\psi_j(t,P)$ has the following expression:*

$$\psi_i(P)=r_i(t)\exp\left(i\,\Omega^{(0)}(P)-t\Omega^{(1)}(P)\right)\frac{\theta(\mathcal{A}(P)+i\,U^{(0)}-tU^{(1)}-\zeta_0)}{\theta(\mathcal{A}(P)-\zeta_0)}\tag{5.27}$$

where $r_i(t)$ is independent of P on Γ and

$$\zeta_0=\sum_{i=1}^g\mathcal{A}(\gamma_i)+\mathcal{K}\tag{5.28}$$

with $\mathcal{K}$ the vector of Riemann constants.

Proof. First, one checks that this function is well defined on Γ, i.e., it does not depend on the path of integration between Q_0 and P, using the formula of Chapter [15]. Then one checks that it has the right poles at finite distance. They are given by the zeroes of the theta function $\theta(A(P)-\zeta_0)$. By Riemann's theorem, they are located at the points $\gamma_1,\cdots,\gamma_g$ because we chose the vector ζ_0 according to eq.(5.28). Finally, we check that it has the right behavior in the neighbourhood of the points at infinity $P^\pm$. The theta functions are regular at infinity. Therefore, the behavior at infinity is governed by the differential $d\Omega^{(0)}$ and

$d\Omega^{(1)}$. From eq.(5.25), we deduce that when $P \to P^{\pm}$:

$$\begin{aligned}\Omega^{(0)}(P) &= \pm \log \mu + O(1) \\ \Omega^{(i)}(P) &= \pm \frac{\mu^i}{i} + O(1), \quad i \geq 1.\end{aligned}$$

The expressions above hold modulo periods, but as we have seen, this does not affect the global expression. Therefore, when $P \to P^{\pm}$:

$$\exp\left(i\,\Omega^{(0)}(P) - t\Omega^{(1)}(P)\right) = const.\ \mu^{\pm i} e^{\mp \mu t}\left(1 + O(\mu^{-1})\right)$$

This proves the result. ∎

In the expression (5.27) the coefficient $r_i(t)$ is fixed up to a sign by the requirement that the leading term in the expansions of $\psi_i(P)$ at $P^{\pm}$ are inverse to each other. In the following we avoid the use of this explicit value of $r_i(t)$ by considering ratios of the values of $\psi_i(P)$ at the vicinity of the points P^+ and P^-, so we don't give the value of $r_i(t)$.

Proposition. *Let $\tau_i(t)$ be the n tau–functions defined by:*

$$\tau_i(t) = e^{\frac{1}{2}\beta_0 i^2} \theta(iU^{(0)} - tU^{(1)} - \widehat{\zeta}_0) \tag{5.29}$$

where β_0 and $\widehat{\zeta}_0$ are some constants. Then, the solution of the equations of motion of the closed Toda chain is given by:

$$e^{2(q_i(t) - q_{i+1}(t))} = \frac{\tau_{i+1}(t)\tau_{i-1}(t)}{\tau_i^2(t)} \tag{5.30}$$

<u>Proof</u>. From eqs.(5.21, 5.22), we have:

$$\frac{\psi_i(P \to P^+)}{\psi_i(P \to P^-)} = e^{2q_i} \mu^{2i} e^{-2\mu t}(1 + O(\mu^{-1})), \quad \mu \to \infty$$

hence substituting the expression (5.27) of ψ_i we obtain an expression of the form:

$$e^{2q_i} = e^{\beta_0 i + \beta_1 t} \left(\frac{\theta(\mathcal{A}(P^+) + i\,U^{(0)} - tU^{(1)} - \zeta_0)}{\theta(\mathcal{A}(P^+) - \zeta_0)}\right)\left(\frac{\theta(\mathcal{A}(P^-) - \zeta_0)}{\theta(\mathcal{A}(P^-) + i\,U^{(0)} - tU^{(1)} - \zeta)}\right)$$

The exponential prefactor comes from $\lim_{P\to P^{\pm}}(\Omega^{(0)} \mp \log\mu) = \beta_0^{\pm}$ and $\lim_{P\to P^{\pm}}(\Omega^{(1)} \mp \mu) = \beta_1^{\pm}$. Then $\beta_k = \beta_k^+ - \beta_k^-$ and the singular part cancels with $\mu^{2i}e^{-2\mu t}$. Taking the quotient of these expressions for i and $i+1$ gives:

$$e^{2(q_i - q_{i+1})} = e^{-\beta_0} \frac{\theta(i\,U^{(0)} - tU^{(1)} - \zeta_0 + \mathcal{A}(P^+))\theta((i+1)U^{(0)} + -tU^{(1)} - \zeta_0 + \mathcal{A}(P^-))}{\theta(i\,U^{(0)} - tU^{(1)} - \zeta_0 + \mathcal{A}(P^-))\theta((i+1)U^{(0)} + -tU^{(1)} - \zeta_0 + \mathcal{A}(P^+))}$$

To end the proof of eq.(5.30), we show that:

$$\mathcal{A}(P^-) - \mathcal{A}(P^+) = U^{(0)}$$

This is a direct consequence of Riemann's bilinear relations which, for normalized abelien differential of the third kind as $d\Omega^{(0)}$ with residue -1 at P^+ and $+1$ at P^- implies:

$$U^{(0)} \equiv \frac{1}{2i\pi}\oint_{b_k} d\Omega^{(0)} \quad = \quad \int_{P^+}^{P^-} \omega_k = \mathcal{A}(P^-) - \mathcal{A}(P^+)$$

Then the result follows by defining $\widehat{\zeta}_0 = \zeta_0 - \mathcal{A}(P^+) - U^{(0)}$ and inserting the definition of the tau–function. ∎

The explicit formula of the Baker-Akhiezer function in terms theta functions also shows that Abel's map linearizes the dynamics. Indeed consider the component $\psi_0(t, P)$. It has poles at fixed position $\gamma_1, \cdots, \gamma_n$ with divisor ζ_0, and zeroes at n others points $\gamma_1(t), \cdots, \gamma_n(t)$. At $t = 0$, $\gamma_i(t = 0) = \gamma_i$ since $\psi_0(t = 0)$ is equal to one. The points $\gamma_i(t)$ are the zeroes of the theta function which is in the numerator of eq.(5.27). Thus, by Riemann's theorem,

$$\sum_{i=1}^{n} (\mathcal{A}(\gamma_i(t)) - \mathcal{A}(\gamma_i)) = tU^{(1)} \tag{5.31}$$

This flow is linear.

Remark. The formula (5.27) can be generalized by considering all flows associated to the other conserved Hamiltonians. If we denote by t_p the time associated to the Hamiltonian H_p, the generalized tau functions will then be given by replacing $tU^{(1)}$ by:

$$tU^{(1)} \quad \rightarrow \quad \sum_p t_p U^{(p)}$$

5.5 Symplectic structure.

We now want to prove that the coordinates (λ_i, μ_i) of the points of the dynamical divisor form a set of separated canonical coordinates. Recall that the Poisson bracket is the standard canonical bracket:

$$\{q_i, q_j\} = \{p_i, p_j\} = 0, \quad \{p_i, q_j\} = \delta_{ij}. \tag{5.32}$$

Proposition. *Let Γ be the spectral curve eq.(5.4). Let $(\lambda_i, \mu_i), i = 1 \cdots g$, be the g points of the dynamical divisor D. Then*

$$\omega = \sum_i \delta q_i \wedge \delta p_i = \frac{1}{2} \sum_i \frac{\delta \lambda_i}{\lambda_i} \wedge \delta \mu_i \tag{5.33}$$

Proof. According to the general procedure explained in Chapter [4], we start from the two form K on phase space with values in the 1-forms on Γ defined by:

$$\begin{aligned} K &= K_1 + K_2 + K_3 \\ K_1 &= < \Psi^{(-1)}(P) \delta L(\lambda) \wedge \delta \Psi(P) > \frac{d\lambda}{\lambda} \\ K_2 &= < \Psi^{(-1)}(P) \delta \mu \wedge \delta \Psi(P) > \frac{d\lambda}{\lambda} \\ K_3 &= \delta \left(\log \partial_\mu R \right) \wedge \delta \mu \, \frac{d\lambda}{\lambda} \end{aligned} \tag{5.34}$$

Here $R(\lambda, \mu) = 0$ is the equation of the spectral curve Γ. Notice the factor $1/\lambda$ as compared to the analysis of Chapter [4]. This is necessary to get the right Poisson bracket on the variables q_i, p_i.
We write that the sum of the residues of the form K seen as a 1–form on Γ vanishes. The poles of K are located at three different places, first the dynamical poles of Ψ, then the poles at $P^\pm$ coming from Ψ ,L and $d\lambda/\lambda$, and finally the poles at the branch points of the covering coming from the poles of $\Psi^{(-1)}$ and from

the fact that the δ differential is taken at fixed λ.
The evaluation of the residues at the dynamical poles and the branch points is exactly as in Chapter [4] and produces a sum of residues $2\sum_k \delta\mu_k \wedge d\lambda_k/\lambda_k$. So we are left with the computation of the residues at $P^{\pm}$. The residue at the point P^+ for instance is obtained by integrating over a small contour enclosing it. But the Riemann surface seen as a branched cover of the λ sphere has a branch point of order $n+1$ at P^+. So, a closed contour around P^+ runs on the $n+1$ sheets of the covering before returning to its starting point. Therefore, we can write

$$\mathrm{Res}_{P^+} K = \frac{1}{2i\pi}\oint K(P) = \frac{1}{2i\pi}\oint \sum_i K(P^i)$$

where in the last expression, the points P_i are the $n+1$ points of the contour over the point on the base with coordinate λ. This sum is independent of the order of the sheets and is therefore a 1-form on the base. The integral is also taken on the base.
The sum $\sum_i K_1(P_i)$ can be written as a trace since the vectors $\Psi(P_i)$, with P_i the $(n+1)$ points over λ, form a basis of eigenvectors of $L(\lambda)$. As in Chapter [4] it is convenient to introduce the $(n+1)\times(n+1)$ matrix $\widehat{\Psi}(P)$ whose columns are the $(n+1)$ vectors $\Psi(P_i)$. We thus have:

$$\mathrm{Res}_{P_+}(K_1) \quad = \quad -\frac{1}{2i\pi}\oint \mathrm{Tr}\left(\delta\widehat{\Psi}(\lambda)\widehat{\Psi}^{-1}(\lambda)\wedge\delta L(\lambda)\right)\frac{d\lambda}{\lambda}$$

The matrix $\delta\widehat{\Psi}(\lambda)\widehat{\Psi}^{-1}(\lambda)$ does not depend on the order of the sheets and is therefore a meromorphic function of λ which we now calculate.

Lemma.

$$\left(\delta\widehat{\Psi}\widehat{\Psi}^{-1}\right)_{ij}(\lambda) \quad = \quad -(\delta q_i-\delta q_0)\delta_{ij}+\delta\eta_i\left(\frac{1}{a_i}\delta_{i,j-1}+\frac{\lambda}{a_{n+1}}\delta_{i,n+1}\delta_{j,1}\right)+O(\lambda^2), \quad P\to P^- \tag{5.35}$$

$$\left(\delta\widehat{\Psi}\widehat{\Psi}^{-1}\right)_{ij}(\lambda) \quad = \quad (\delta q_i-\delta q_0)\delta_{ij}-\delta\xi_i\left(\frac{1}{a_{i-1}}\delta_{i,j+1}+\frac{1}{\lambda a_{n+1}}\delta_{i,1}\delta_{j,n+1}\right)+O(\lambda^{-2}), \quad P\to P^+ \tag{5.36}$$

where $\eta_i=\sum_{j=1}^{i}p_j$ *and* $\xi_i=\sum_{j=0}^{i-1}p_j$.

Proof. Let us first consider the behavior near P^-. From eq.(5.12), we have

$$\widehat{\Psi}_{ij}=\widehat{\Psi}_i(P_j) \quad = \quad e^{-q_i+q_0}\mu_j^{-i}\left(1+\eta_i\,\mu_j^{-1}+O(\mu_j^{-2})\right)$$

$$\widehat{\Psi}^{-1}_{ij} \quad = \quad \frac{1}{n+1}e^{q_j-q_0}\mu_i^j\left(1+O(\mu_j^{-1})\right)$$

where μ_j^{-1} is the local coordinate of the point P_j above λ, i.e. $\mu_j=\lambda^{\frac{-1}{n+1}}\alpha_j+\cdots$ where α_j is a $(n+1)$-th root of unity.
Taking variations, we limit ourselves to $O(\mu^{-2})$ so that we can set $\delta\mu=0$ using eq.(5.8) where $\sum_j p_j=0$, and obtain: $\delta\widehat{\Psi}_{ij}=-(\delta q_i-\delta q_0)\widehat{\Psi}_{ij}+\delta\eta_i\widehat{\Psi}_{ij}\mu_j^{-1}+O(\mu_j^{-2})$ and

$$\left(\delta\widehat{\Psi}\widehat{\Psi}^{-1}\right)_{ij}(\lambda)=-(\delta q_i-\delta q_0)\delta_{ij}+\delta\eta_i\sum_k\widehat{\Psi}_{ik}\mu_k^{-1}\widehat{\Psi}^{-1}_{kj}$$

Note that this shows that $\widehat{\Psi}^{-1}$ needs to be computed to leading order only. The last term is equal to:

$$\sum_k\widehat{\Psi}_{ik}\mu_k^{-1}\widehat{\Psi}^{-1}_{kj}=e^{q_j-q_i}\frac{1}{n+1}\sum_k\mu_k^{-i+j-1}=e^{q_j-q_i}\lambda^{\frac{i-j+1}{n+1}}\frac{1}{n+1}\sum_k\alpha_k^{-i+j-1}$$

The last sum is over the roots of unity. Hence, it vanishes unless $i-j+1\equiv 0 \bmod [n+1]$, which is possible only if $i=j-1$, $(i=1,\cdots,n)$, and $i=n+1, j=1$. Evaluating these two types of terms yields eq.(5.35). The analysis at P^+ is similar. ∎

We now finish the proof of the proposition, by first analysing the residue of K_1 at $P^\pm$. First, it is clear that only the terms written in eqs.(5.35, 5.36) contribute to the residue. Indeed if one keeps terms of order μ^{-2} the same computation as above yields contributions to $\delta\widehat{\Psi}\widehat{\Psi}^{-1}$ which are not vanishing only if $i-j \equiv \pm 2 \bmod [n+1]$ which cannot contribute to the trace with δL. The contributions to the residues at $P^\pm$ of the first terms in eqs.(5.35,5.36) add to $2\sum_i \delta p_i \wedge \delta q_i$ (remember that $d\lambda/\lambda$ has residue ∓ 1 at $P^\pm$). Similarly the contributions of the second terms at $P^\pm$ are also equal and add up to $2\sum_i \delta p_i \wedge \delta q_i$. Finally since $\delta\mu$ has a simple zero at $P^\pm$ the forms K_2 and K_3 are regular at these points. ∎

The coordinates $(\log\lambda_\alpha, \mu_\alpha)$ with $(\lambda_\alpha, \mu_\alpha)$ coordinates of the points of the dynamical divisors are therefore canonical coordinates. Since they are coordinates of points of the spectral curve, these n pairs of coordinates are related by the n separated equations, namely

$$R(\lambda_\alpha, \mu_\alpha) = 0, \quad \text{for} \quad \alpha = 1, \cdots, n$$

with $R(\lambda, \mu) = 0$ the equation of the spectral curve. As explained in Chapter [2], this implies that the variables are separated and furthermore, this allows to reconstruct the action–angle variables.

5.6 The Sklyanin approach.

We now introduce a description of the Toda chain by 2×2 matrices which is widely used. It is based on replacing the linear second order system eq.(5.10) by a linear 2×2 first order system. In fact we can rewrite it as:

$$\begin{pmatrix} \psi_i \\ \psi_{i+1} \end{pmatrix} = \widehat{T}_i(\mu) \begin{pmatrix} \psi_{i-1} \\ \psi_i \end{pmatrix}$$

where the $\widehat{T}_i$'s are given by:

$$\widehat{T}_i(\mu) = \begin{pmatrix} 0 & 1 \\ -a_{i-1}/a_i & -(p_i-\mu)/a_i \end{pmatrix} \tag{5.37}$$

The solution is (with $\psi_0 = \lambda^{-1}\psi_{n+1}$):

$$\begin{pmatrix} \psi_i \\ \psi_{i+1} \end{pmatrix} = \widehat{T}_i(\mu)\widehat{T}_{i-1}(\mu)\cdots\widehat{T}_1(\mu) \begin{pmatrix} \psi_0 \\ \psi_1 \end{pmatrix}$$

Hence the periodicity condition $\psi_{n+1+i} = \lambda\psi_i$ translates into:

$$\widehat{T}(\mu) \begin{pmatrix} \psi_0 \\ \psi_1 \end{pmatrix} = \lambda \begin{pmatrix} \psi_0 \\ \psi_1 \end{pmatrix}$$

where $\widehat{T}(\mu)$ is the monodromy matrix defined by:

$$\widehat{T}(\mu) = \widehat{T}_{n+1}(\mu)\widehat{T}_n(\mu)\cdots\widehat{T}_1(\mu) \equiv \begin{pmatrix} \widehat{A}(\mu) & \widehat{B}(\mu) \\ \widehat{C}(\mu) & \widehat{D}(\mu) \end{pmatrix} \tag{5.38}$$

Here $\widehat{A}, \widehat{B}, \widehat{C}, \widehat{D}$ are polynomials in μ of degrees $\deg\widehat{A} = n-1$, $\deg\widehat{B} = \deg\widehat{C} = n$, and $\deg\widehat{D} = n+1$. Since $\det\widehat{T}_i(\mu) = a_{i-1}/a_i$ one gets $\det\widehat{T}(\mu) = 1$, and the characteristic equation $\det(\widehat{T}(\mu) - \lambda) = 0$ reads

$$\lambda^2 - 2t(\mu)\lambda + 1 = 0, \tag{5.39}$$

with

$$2t(\mu) = \mathrm{Tr}\,(\widehat{T}(\mu)) = \widehat{A}(\mu) + \widehat{D}(\mu)$$

Clearly $t(\mu)$ is a polynomial in μ of degree $(n+1)$: $t(\mu) = \frac{1}{2\mathcal{A}}\left(\mu^{n+1} + \cdots\right)$. Moreover eq.(5.39) expresses the existence of an eigenvector of $L(\lambda)$ hence is proportional to eq.(5.4), so that $t(\mu) = \Delta(\mu)/\mathcal{A}$. By using $\widehat{T}(\mu)$ instead of $L(\lambda)$ we have exchanged the role payed by μ and λ. The spectral curve Γ is now presented as a two-sheeted covering of the μ-plane. Above each point μ there are two points corresponding to the two roots

$$\lambda^{\pm} = t(\mu) \pm \sqrt{t^2(\mu) - 1}$$

of eq.(5.39) such that $\lambda^+\lambda^- = 1$. In particular above $\mu = \infty$ we have the two points $P^+(\lambda = \infty)$ and $P^-(\lambda = 0)$.

In the following, instead of using the previous normalization condition $\psi_1 = 1$ we shall choose the natural normalization condition $\psi_0(P) = 1$. It is useful to introduce a standard basis of solutions of the linear system, which we denote as $\chi_i^{(0)}$, and $\chi_i^{(1)}$, specified by the boundary conditions:

$$\begin{pmatrix} \chi_0^{(0)} \\ \chi_1^{(0)} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad ; \quad \begin{pmatrix} \chi_0^{(1)} \\ \chi_1^{(1)} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{5.40}$$

These solutions are polynomials in μ with $\deg \chi_i^{(0)} = i - 2$ and $\deg \chi_i^{(1)} = i - 1$ and we have:

$$\psi_i = \chi_i^{(0)} + \psi_1 \chi_i^{(1)}.$$

The coefficient ψ_1 is determined by the periodicity condition $\psi_{n+1+i} = \lambda \psi_i$. This is equivalent to the eigenvalue equation $\widehat{T}(\mu)\begin{pmatrix} 1 \\ \psi_1 \end{pmatrix} = \lambda \begin{pmatrix} 1 \\ \psi_1 \end{pmatrix}$. It gives:

$$\psi_i^{\pm} = \chi_i^{(0)} + \left(\frac{\lambda^{\pm} - \widehat{A}(\mu)}{\widehat{B}(\mu)} \right) \chi_i^{(1)} \tag{5.41}$$

The two functions $\psi_i^{\pm}$ corresponding to the two Bloch waves are in fact the values of a unique meromorphic function $\psi_i(P)$ evaluated at the two points $(\mu, \lambda^{\pm})$ above μ, on Γ.

The poles at finite distance of ψ_i are thus the sames as those of ψ_1 and are located above the n zeroes of $\widehat{B}(\mu)$. When $\widehat{B}(\mu_k) = 0$ the two eigenvalues of $\widehat{T}(\mu_k)$ are $\widehat{A}(\mu_k)$ and $\widehat{D}(\mu_k)$ hence the numerator $\lambda - \widehat{A}(\mu)$ vanishes on one of the two points above μ_k. Therefore the function $\psi_i(P)$ has only one pole at $(\mu_k, \lambda_k = \widehat{D}(\mu_k))$. Hence the dynamical divisor is exactly the same as in the $(n+1) \times (n+1)$ matrix approach. As we know from previous section, the coordinates (μ_k, λ_k) of these points form a set of conjugated canonical variables. We shall present an alternative way of deriving this result in the following section.

5.7 The Poisson brackets.

We first establish an explicit formula for the Poisson brackets of the matrix elements of $\widehat{T}(\mu)$. It is actually more convenient to perform a gauge tranformation on $\widehat{T}(\mu)$ before computing these Poisson brackets. Since $\widehat{T}(\mu) = \widehat{T}_{n+1}(\mu) \cdots \widehat{T}_1(\mu)$ is the product of the matrices $\widehat{T}_i(\mu)$, under a gauge transformation of the $\widehat{T}_i(\mu)$,

$$\widehat{T}_i(\mu) \to T_i(\mu) = D_i \widehat{T}_i(\mu) D_{i-1}^{-1},$$

with D_i periodic, $D_{n+1+i} = D_i$, $\widehat{T}(\mu)$ gets conjugated by D_0: $\widehat{T}(\mu) \to T(\mu) = D_0 \widehat{T}(\mu) D_0^{-1}$. In particular, the spectral curve $\det(\widehat{T}(\mu) - \lambda) = 0$ is preserved by such gauge transformation.

We choose the matrices D_i such that the matrices $T_i(\mu)$ are local, i.e. $T_i(\mu)$ only depends on the canonical variables q_i and p_i. We take $D_i = \mathrm{Diag}\left(d_i, d_{i+1}^{-1}\right)$ with $d_i = \exp(q_i)$. Notice that $a_i = d_i/d_{i+1}$. The explicit expression for $T_i(\mu)$ are:

$$T_i(\mu) = \begin{pmatrix} 0 & e^{2q_i} \\ -e^{-2q_i} & (\mu - p_i) \end{pmatrix} \tag{5.42}$$

Remark that $\det T_i(\mu) = 1$ for all i. The matrix $T(\mu)$ is equal to $T(\mu) = T_{n+1}(\mu) \cdots T_1(\mu)$.

Proposition. *The Poisson brackets of the matrix elements of $T(\mu)$ are given by:*

$$\{T(\mu) \otimes 1, 1 \otimes T(\mu')\} = [r(\mu - \mu'), T(\mu) \otimes T(\mu')], \quad \text{with } r(\mu - \mu') = \frac{2C}{\mu - \mu'} \tag{5.43}$$

where $C = \sum_{ij} E_{ij} \otimes E_{ji}$ is the Casimir tensor.

Proof. We first prove this relation for each individual $T_i(\mu)$. Namely,

$$\{T_i(\mu) \otimes 1, 1 \otimes T_i(\mu')\} = [r(\mu - \mu'), T_i(\mu) \otimes T_i(\mu')]$$

This is shown by a direct computation using the explicit formula (5.42) for $T_i(\mu)$.
We then prove that if two Poisson–commuting matrices $T_1(\mu)$ and $T_2(\mu')$ satisfy (5.43) then so does their product. Indeed, using the fact that the Poisson bracket and the Lie bracket are both derivations

$$\begin{aligned}\{T_2(\mu)T_1(\mu) \otimes 1, 1 \otimes T_2(\mu')T_1(\mu')\} &= T_2(\mu) \otimes T_2(\mu')\{T_1(\mu) \otimes 1, 1 \otimes T_1(\mu')\} \\ &\quad + \{T_2(\mu) \otimes 1, 1 \otimes T_2(\mu')\}T_1(\mu) \otimes T_1(\mu') \\ &= [r(\mu - \mu'), T_2(\mu)T_1(\mu) \otimes T_2(\mu')T_1(\mu')].\end{aligned}$$

The claim then follows by induction. ∎

Remark that this also implies the integrability of the Toda chain.

Let $A(\mu)$, $B(\mu)$, $C(\mu)$ and $D(\mu)$ be the matrix elements of $T(\mu)$:

$$T(\mu) = \begin{pmatrix} A(\mu) & B(\mu) \\ C(\mu) & D(\mu) \end{pmatrix}.$$

In terms of the matrix elements of $\widehat{T}(\mu)$ introduced in the previous section, $A(\mu) = \widehat{A}(\mu)$, $D(\mu) = \widehat{D}(\mu)$ but $B(\mu) = (d_0 d_1)\widehat{B}(\mu)$ and $C(\mu) = \widehat{C}(\mu)/(d_0 d_1)$. In particular one finds:

$$B(\mu) = d_0^2 \left(\mu^n - \mu^{n-1}(\sum_{j=1}^n p_j) + \cdots \right).$$

This is a polynomial of degree n with n zeroes, which we denote by μ_ν, $\nu = 1, \cdots, n$:

$$B(\mu) = d_0^2 \prod_{\nu=1}^n (\mu - \mu_\nu).$$

As explained in the previous sections, these zeroes are the μ–coordinates of the poles of the Baker functions ψ_i.

Proposition. *Let μ_ν be the zeroes of $B(\mu)$ and λ_ν be the values of $D(\mu)$ at $\mu = \mu_\nu$. The n points (λ_ν, μ_ν) are the points of the dynamical divisor.*

$$B(\mu_\nu) = 0 \quad \text{and} \quad \lambda_\nu = D(\mu_\nu) \tag{5.44}$$

These parameters obey the following Poisson brackets, i.e. they form a set of canonical coordinates.

$$\{\mu_\nu, \mu_\alpha\} = \{\lambda_\nu, \lambda_\alpha\} = 0, \quad \{\lambda_\nu, \mu_\alpha\} = 2\lambda_\nu \; \delta_{\nu\alpha} \tag{5.45}$$

Proof. Recall that the matrix $T(\mu)$ is lower triangular at $\mu = \mu_\nu$, since $B(\mu_\nu) = 0$. Therefore, at $\mu = \mu_\nu$ we have: $1 = \det T(\mu_\nu) = A(\mu_\nu)D(\mu_\nu)$ and $2t(\mu_\nu) = A(\mu_\nu) + D(\mu_\nu)$, hence the points (λ_ν, μ_ν) are on the spectral curve:

$$\lambda_\nu + \lambda_\nu^{-1} = 2t(\mu_\nu) \tag{5.46}$$

The Poisson brackets (5.43) for $T(\mu)$ in particular implies the following:

$$\{A(\mu), A(\mu')\} = \{B(\mu), B(\mu')\} = \{C(\mu), C(\mu')\} = \{D(\mu), D(\mu')\}$$

$$\{A(\mu), B(\mu')\} = 2\frac{A(\mu)B(\mu') - B(\mu)A(\mu')}{\mu - \mu'} \tag{5.47}$$

$$\{A(\mu), D(\mu')\} = 2\frac{C(\mu)B(\mu') - B(\mu)C(\mu')}{\mu - \mu'} \tag{5.48}$$

$$\{B(\mu), D(\mu')\} = 2\frac{D(\mu)B(\mu') - B(\mu)D(\mu')}{\mu - \mu'} \tag{5.49}$$

$$\tag{5.50}$$

The first equation directly implies that the μ_ν Poisson commute.
Let $P(\mu)$ be a polynomial in μ with coefficients functions on phase space, and F be an arbitrary function on phase space. Then the Poisson bracket between F and the value of P at μ_ν is:

$$\{F, P(\mu_\nu)\} = \{F, P(\mu)\}\Big|_{\mu=\mu_\nu} + \{F, \mu_\nu\}\ \partial_\mu P(\mu_\nu) \tag{5.51}$$

We apply this to $F = D(\mu)$ and $P = B$. Since $B(\mu_\nu) = 0$ we get:

$$0 = \{D(\mu), B(\mu_\nu)\} = \{D(\mu), B(\mu')\}\Big|_{\mu'=\mu_\nu} + \{D(\mu), \mu_\nu\}\ \partial_\mu B(\mu_\nu)$$

Using eq.(5.49) we get

$$\{D(\mu), B(\mu_\nu)\} = \frac{2}{\mu - \mu_\nu}\frac{\lambda_\nu}{\partial_\mu B(\mu_\nu)}B(\mu)$$

Finally we apply the same formula to $F = \mu_\nu$ and $P = D$, evaluated at $\mu = \mu_{\nu'}$. One gets since $\{\mu_\nu, \mu_{\nu'}\} = 0$

$$\{\mu_\nu, \lambda_{\nu'}\} = \{\mu_\nu, D(\mu_{\nu'})\} = \{\mu_\nu, D(\mu)\}|_{\mu=\mu_{\nu'}}$$

So we have to evaluate $B(\mu)/(\mu-\mu_\nu)$ at $\mu = \mu_{\nu'}$ which yields $\delta_{\nu\nu'}\partial_\mu B(\mu_\nu)$. The equation $\{\lambda_{\nu'}, \mu_\nu\} = 2\delta_{\nu\nu'}\lambda_\nu$ follows.
Finally to prove $\{\lambda_\nu, \lambda_{\nu'}\} = 0$, one expands completely this expression using eq.(5.51), $\{D(\mu), D(\mu')\} = 0$ and $\{\mu_\nu, \mu_{\nu'}\} = 0$ getting:

$$\{\lambda_\nu, \lambda_{\nu'}\} = \partial_\mu D(\mu_{\nu'})\{D(\mu), \mu_{\nu'}\}|_{\mu=\mu_\nu} - \partial_\mu D(\mu_\nu)\{D(\mu), \mu_\nu\}|_{\mu=\mu_{\nu'}}$$

We have seen that each of these two terms vanishes when $\nu \neq \nu'$ and they obviously cancel each other when $\nu = \nu'$. ■

In particular, the symplectic two-form ω can be written as:

$$\omega = \frac{1}{2}\sum_{\nu=1}^{n} \frac{\delta\lambda_\nu}{\lambda_\nu} \wedge \delta\mu_\nu.$$

where δ denote the differential on the phase space. The canonical one-form α such that $\omega = \delta\alpha$ is $\alpha = 1/2\sum_\nu \log\lambda_\nu\delta\mu_\nu$.

The fact that the equations of motion linearize on the Jacobian variety of the spectral curve may also be derived in this framework. The generating function for the conserved Hamiltonians is the trace of the monodromy matrix $\mathrm{Tr}\,(T(\mu)) = 2t(\mu) = A(\mu) + D(\mu)$. Thus the equations of motion of any function F on phase space for this generic Hamiltonian $t(u)$ are:

$$\dot{F} \equiv \{t(u), F\}$$

Proposition. *The equations of motion of the zeroes μ_ν of $B(\mu)$ are:*

$$\dot{\mu}_\nu = \{t(u), \mu_\nu\} = (A(\mu_\nu) - D(\mu_\nu))\frac{B(u)}{(u-\mu_\nu)B'(\mu_\nu)} \tag{5.52}$$

Their linearization is given the following relations:

$$\sum_{\alpha=1}^{n} \frac{\mu_\alpha^k}{\sqrt{t^2(\mu_\alpha)-1}}\dot{\mu}_\alpha = 2u^k, \quad 0 \le k \le n-1 \tag{5.53}$$

Proof. The Poisson bracket $\{t(u), \mu_\nu\}$ can be computed in the same way as in previous proposition. One obtains:

$$\{t(u), \mu_\nu\} \quad = \quad \frac{B(u)}{(u-\mu_\nu)B'(\mu_\nu)}(A(\mu_\nu) - D(\mu_\nu))$$

Hence $\dot{\mu}_\nu$ is a polynomial in u of degree $n-1$ which vanishes at the points μ_α for $\alpha \neq \nu$ and takes the value $A(\mu_\nu) - D(\mu_\nu) = 2\sqrt{t^2(\mu_\nu)-1}$ for $\alpha = \nu$. On an hyperelliptic curve the abelian differentials are the

$$\omega_k = \frac{\mu^k}{\sqrt{t^2(\mu)-1}}\, d\mu, \quad 0 \le k \le n-1$$

so that the time differential of the Abel map $\mathcal{A}_k = \sum_\nu \int^{P_\nu} \omega_k$ is given by:

$$\dot{\mathcal{A}}_k = \sum_\alpha \frac{\mu_\alpha^k \dot{\mu}_\alpha}{\sqrt{t^2(\mu_\alpha)-1}}$$

This is also a polynomial in u of degree $n-1$. If one evaluates it at μ_ν, only the term $\alpha = \nu$ contributes in the sum and it gives $2\mu_\nu^k$. Hence the polynomial itself is $2u^k$. ∎

5.8 Reality conditions.

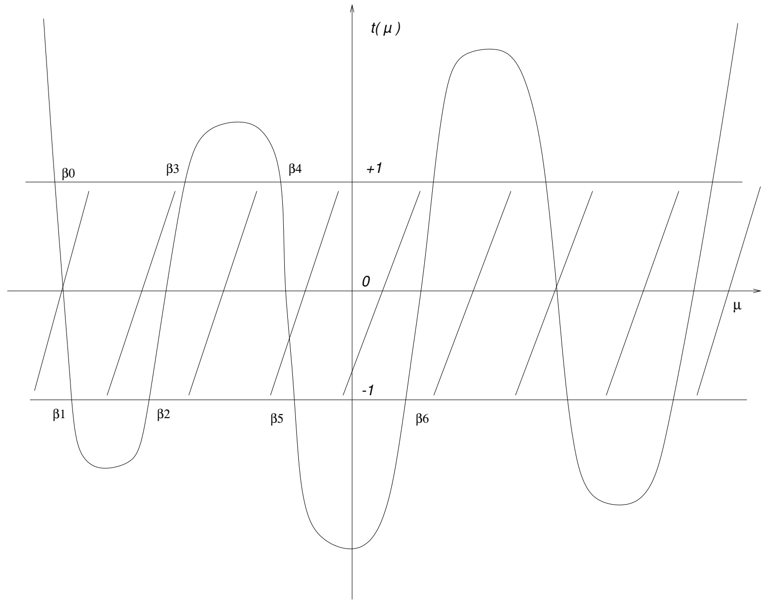
Up to now the dynamical variables of the Toda chain were complex valued. We now come to the important and generally difficult question of choosing a real slice in the complexified phase space. This complex phase space is a fibered space with basis the action variables and fiber the Jacobian variety of the corresponding spectral curve. While it is easy to take real action variables it is less trivial to choose a proper real slice on the Jacobian which will be identified with the Liouville torus.

Let us first remark that the equation of the spectral curve, eq.(5.5) shows that it is an hyperelliptic curve with hyperelliptic involution given by $(\lambda, \mu) \to (1/\lambda, \mu)$. Hence the branch points, which are invariant under this involution are located at $\lambda = 1/\lambda$, i.e. $\lambda = \pm 1$. For such values of λ the Lax matrix $L(\lambda)$, eq.(5.2) is symmetric. If moreover the dynamical variables are real the matrices $L(\lambda = \pm 1)$ are real symmetric matrices hence have $(n+1)$ real eigenvalues β_i. These values of μ are precisely the location of the branch points, so we get altogether $2(n+1)$ *real* branch points which completely characterize the curve. This implies in particular that the action variables, i.e. the coefficients of eq.(5.5) are real.

Writing the equation of the spectral curve as:

$$\lambda + \frac{1}{\lambda} = 2t(\mu)$$

where $t(\mu)$ is a polynomial of degree $(n+1)$ with real coefficients, the $2n+2$ branch points are located at $t(\mu) = \pm 1$. Hence the graph of $t(\mu)$ has the following form (here assuming $n+1$ even):



For $n+1$ odd, the left branch (where $\mu \to -\infty$) goes to $-\infty$. Remark that for real μ the two values of λ such that $\lambda^2 - 2t(\mu)\lambda + 1 = 0$ are either real or complex conjugated with modulus one. Borrowing the terminology of solid state physics the first case is called forbidden zone, while the second case is called allowed zone. We see that the forbidden zones correspond to the intervals $[\beta_i, \beta_{i+1}]$ for $i = 1, 3, \cdots, 2n-1$ (for both n even and odd), to which is added one extra zone passing through $\mu = \infty$ given by $\mu < \beta_0$ or $\mu > \beta_{2n+2}$. So there are n compact forbidden zones.

We now have to locate the points of the dynamical divisor in agreement with the reality conditions. Recall that they are given by the points $B(\mu) = 0$ and $\lambda = D(\mu)$. We now show that the zeroes of the polynomial $B(\mu)$ of degree n are real so that the corresponding λ's are also real since $D(\mu)$ has real coefficients. This immediately implies that the zeroes of $B(\mu)$ lie in the forbidden zones.

Proposition. *The points of the dynamical divisor have real coordinates, and there is exactly one such point in each of the n forbidden zones $[\beta_i, \beta_{i+1}]$ for $i = 1, 3, \cdots, 2n-1$.*

Proof. We first show that the zeroes of $B(\mu)$ are real. This is because the zeroes of $B(\mu)$ are also the poles of the eigenvector of $L(\lambda)$ normalized by $\psi_0 = 1$. The eigenvector equation being:

$$\begin{pmatrix} p_0-\mu & a_0 & \cdots & \lambda^{-1}a_n \\ a_0 & p_1-\mu & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ \lambda a_n & 0 & \cdots & p_n-\mu \end{pmatrix} \begin{pmatrix} \psi_0 \\ \psi_1 \\ \vdots \\ \psi_n \end{pmatrix} = 0$$

the poles of the normalized Ψ are included in the zero set of the minor of $(p_0 - \mu)$, see Chapter [4]. This minor is the characteristic equation of a real symmetric $n \times n$ matrix hence has n real eigenvalues μ_i.
To show that there is just one zero of $B(\mu)$ in each $[\beta_i, \beta_{i+1}]$, we show that $B(\mu)$ has the same sign as $a_0 dt(\mu)/d\mu$ in the *allowed* zones, hence changes sign between β_i and β_{i+1}, i odd, because so does $dt(\mu)/d\mu$. Since $B(\mu)$ has n zeroes the only possibility is that it has exactly one zero in each forbidden zone.
The analysis of the sign of $a_0 dt(\mu)/d\mu$ is quite involved. It rests on the analysis of the two fundamental solutions eq.(5.40) of the Schroedinger equation eq.(5.9). In particular recall that (χ_{n+1}, χ_{n+2}) is related to the initial values (χ_0, χ_1) by the monodromy matrix $\widehat{T}$ so that we have $\widehat{A}(\mu) = \chi^{(0)}_{n+1}(\mu)$, $\widehat{B}(\mu) = \chi^{(1)}_{n+1}(\mu)$, $\widehat{C}(\mu) = \chi^{(0)}_{n+2}(\mu)$, $\widehat{D}(\mu) = \chi^{(1)}_{n+2}(\mu)$, and $2t(\mu) = \chi^{(0)}_{n+1}(\mu) + \chi^{(1)}_{n+2}(\mu)$. Moreover since $\det \widehat{T} = 1$, we get the Wronskian relation:

$$\chi^{(0)}_{n+1}(\mu)\chi^{(1)}_{n+2}(\mu) - \chi^{(0)}_{n+2}(\mu)\chi^{(1)}_{n+1}(\mu) = 1 \tag{5.54}$$

We then consider two solutions $\chi_i^{(j)}(\mu)$ and $\chi_i^{(j')}(\mu')$ of the Schroedinger equation corresponding to different values μ and μ' of its parameter.

$$\begin{aligned} a_{i-1}\chi_{i-1}^{(j)}(\mu) + p_i\chi_i^{(j)}(\mu) + a_i\chi_{i+1}^{(j)}(\mu) &= \mu\chi_i^{(j)}(\mu) \\ a_{i-1}\chi_{i-1}^{(j')}(\mu') + p_i\chi_i^{(j')}(\mu') + a_i\chi_{i+1}^{(j')}(\mu') &= \mu'\chi_i^{(j')}(\mu') \end{aligned}$$

Here j, j' take the values $0, 1$. We multiply the first equation by $\chi_i^{(j')}(\mu')$, the second by $\chi_i^{(j)}(\mu)$ and subtract. Adding the resulting equations for $i = 1, 2, \cdots, n+1$ we get:

$$\begin{aligned} a_0\left(\chi_0^{(j)}(\mu)\chi_1^{(j')}(\mu') - \chi_1^{(j)}(\mu)\chi_0^{(j')}(\mu')\right) &- a_{n+1}\left(\chi_{n+1}^{(j)}(\mu)\chi_{n+2}^{(j')}(\mu') - \chi_{n+2}^{(j)}(\mu)\chi_{n+1}^{(j')}(\mu')\right) \\ &= (\mu - \mu')\left(\chi^{(j)}(\mu), \chi^{(j')}(\mu')\right) \end{aligned}$$

where we have denoted $(\chi^{(j)}(\mu), \chi^{(j')}(\mu')) = \sum_{i=1}^{n+1}\chi_i^{(j)}(\mu)\chi_i^{(j')}(\mu')$. Consider the four above equations for $(j, j') = (0,0), (1,1), (0,1), (1,0)$ and multiply them respectively by $\chi_{n+1}^{(1)}(\mu)$, $-\chi_{n+2}^{(0)}(\mu)$, $\chi_{n+2}^{(1)}(\mu)$ and $-\chi_{n+1}^{(0)}(\mu)$. Adding them and taking into account eq.(5.54) we get:

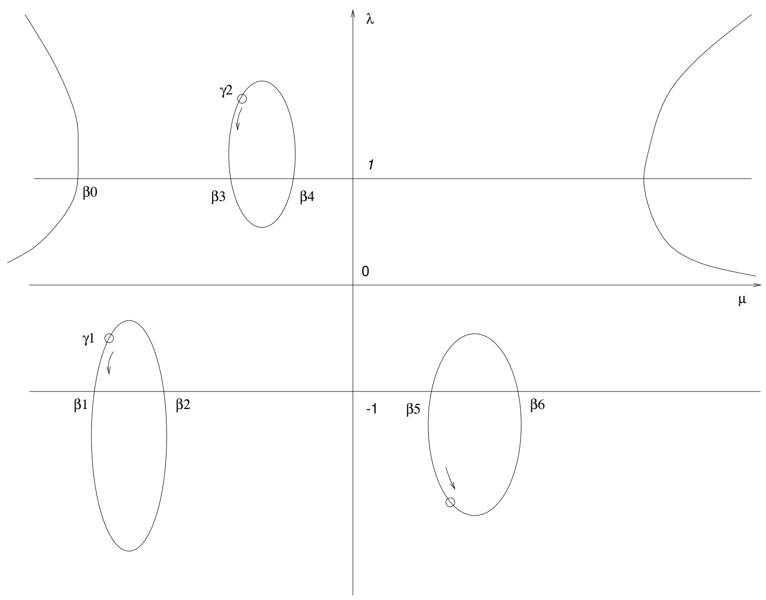
$$\begin{aligned} 2a_0\frac{t(\mu) - t(\mu')}{\mu - \mu'} &= \widehat{B}(\mu)(\chi^{(0)}(\mu), \chi^{(0)}(\mu')) - \widehat{C}(\mu)(\chi^{(1)}(\mu), \chi^{(1)}(\mu')) + \\ &\quad \widehat{D}(\mu)(\chi^{(0)}(\mu), \chi^{(1)}(\mu')) - \widehat{A}(\mu)(\chi^{(1)}(\mu), \chi^{(0)}(\mu')) \end{aligned}$$

We can now set $\mu = \mu'$ and obtain our final relation:

$$2a_0\frac{dt(\mu)}{d\mu} = \widehat{B}\left(\chi^{(0)}, \chi^{(0)}\right) - \widehat{C}\left(\chi^{(1)}, \chi^{(1)}\right) + (\widehat{D} - \widehat{A})\left(\chi^{(0)}, \chi^{(1)}\right) \tag{5.55}$$

The right–hand side is a sum (over $i = 1, \cdots, n+1$) of quadratic forms with discriminant $(\widehat{D} - \widehat{A})^2 + 4\widehat{B}\widehat{C} = 4(t^2(\mu) - 1)$. In the allowed zones we have $t^2(\mu) - 1 < 0$ hence the quadratic form has a definite sign, which is the sign of $\widehat{B}(\mu)$. Equation (5.55) shows that this is also the sign of $a_0 dt/d\mu$ as asserted above. ∎

We can now describe the real part of the spectral curve and the dynamical divisor. Real points on the spectral curve are obtained when $t^2(\mu) - 1 > 0$. Hence it has $n+1$ components, one of which extends to ∞. The dynamical divisor has n points γ_j with just one point in each of the components $\mathcal{C}_j$ at finite distance.



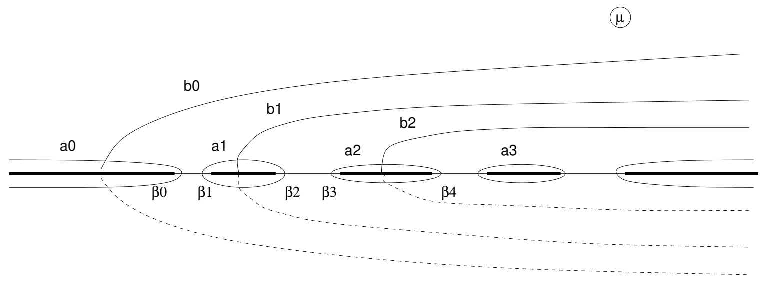
As time goes on, each of the points γ_j runs along the cycle $\mathcal{C}_j$ hence the whole motions lies on the real torus $\mathcal{C}_1 \times \mathcal{C}_2 \times \cdots \times \mathcal{C}_n$ which is the Liouville torus. It is easy to compute as above:

$$\dot{\lambda}_j = -2\frac{1}{u-\mu_j}\frac{B(u)}{B'(\mu_j)}\lambda_j t'(\mu_j)$$

(for the Hamiltonian $t(u)$) hence when the point γ_j hits the line $\lambda = \pm 1$ we have $\dot{\mu}_j = 0$ and $\dot{\lambda}_j \neq 0$ so the point γ_j continues in the same direction.

It is interesting to relate the Liouville torus to the real slice of the Jacobian variety of the spectral curve. Indeed one can define an antiholomorphic involution of the Jacobian by sending $\sum \gamma_j$ to $\sum \overline{\gamma}_j$ where if $\gamma = (\lambda, \mu)$ we define $\overline{\gamma} = (\overline{\lambda}, \overline{\mu})$ which also lies on the spectral curve since $t(\mu)$ has real coefficients. A real slice of the Jacobian can be defined as the set of fixed points of this involution. Note that $\sum \gamma_j$ is invariant if the γ_j are real or occur in complex conjugate pairs. The various combinations of these two possibilities define several connected components of the real slice of the Jacobian. The Liouville torus is just one of these components. Note that we have discussed the reality condition for the dynamical variables a_i and p_i, but one should further ensure that the variables a_i are positive to get real q_i's.

Altenatively one can view the Liouville torus as some choice of a–cycles on the spectral curve. Consider it as a two–sheeted covering of the complex μ–plane, with cuts between branch points as shown below:



Here the a–cycles are drawn on the first sheet, so in the limit where the ellipses go to the segments $[\beta_i, \beta_{i+1}]$ we see that λ goes to the two real roots of the spectral equation, thereof producing the cycles in the previous drawing. The b–cycles are also drawn, starting from ∞ they cut the corrresponding a–cycle once, then pass on the second sheet through the cut, and return to ∞ without cutting any more the a–cycles.

References

[1] H. Flaschka.

[2] P.van Moerbeke, *"The spectrum of Jacobi Matrices"*. Invent.Math. (1976) 45-81.

[3] P. van Moerbeke, D. Mumford, *"The spectrum of difference operators and algebraic curves."* Acta. Math. vol 143 (1979) pp.93-154.

[4] P. van Moerbeke *"About Isospectral deformations of discrete Laplacians."* Lecture Notes in Math. vol. 755 (1979) pp.313-370.

[5] B. Dubrovin, I. Krichever, S. Novikov, *"Integrable Systems. I."* Encyclopedia of Mathematical Sciences, Dynamical Systems IV. Springer (1990).

Chapter 6

The Calogero–Moser model.

The Calogero model consists of N identical particles on a line, at positions q_i and momenta p_i, with pairwise interactions and Hamiltonian:

$$H = \frac{1}{2}\sum_i p_i^2 + \frac{1}{2}\gamma^2 \sum_{\substack{i,j=1\\ i\neq j}}^{N} \frac{1}{(q_i - q_j)^2}$$

This dynamical system is integrable, moreover it remains integrable when the potential is replaced by an elliptic potential, i.e. $1/q^2 \to \wp(q)$. We are primarily interested by this elliptic model, which admits a Lax pair formulation with a spectral parameter lying on an elliptic curve. This provides an example of an integrable system whose spectral parameter belongs to a higher genus curve. Another interesting feature of the Calogero model is that its r–matrix explicitly contains dynamical variables. It is rewarding to consider a slightly generalized model, the spin Calogero model, which contains extra dynamical spin variables. Finally we show that the Calogero model is a particular instance of a more general construction due to Hitchin, relying on Hamiltonian reduction, which allows, in principle, to construct models with spectral parameter lying on higher genus Riemann surface.

6.1 The spin Calogero model.

Let us introduce a set of dynamical variables $(q_i, p_i)_{i=1\cdots N}$ and $(f_{ij})_{i,j=1\cdots N}$ together with the Poisson brackets:

$$\{p_i, q_j\} = \delta_{ij} \tag{6.1}$$

$$\{f_{ij}, f_{kl}\} = -\delta_{jk}\ f_{il} + \delta_{li}\ f_{kj} \tag{6.2}$$

and the Hamiltonian

$$H = \frac{1}{2}\sum_{i=1}^{N} p_i^2 - \frac{1}{2}\sum_{\substack{i,j=1\\ i\neq j}}^{N} f_{ij} f_{ji} V(q_i - q_j) \tag{6.3}$$

with potential $V(q) = \wp(q)$, where $\wp(q)$ is the Weierstrass function. This defines the spin Calogero model. The minus sign in front of the potential in H allows to simplify later formulae.

The equations of motion are easily derived:

$$\dot{q}_i = p_i, \qquad \dot{p}_i = \sum_{\substack{j=1\\ j\neq i}}^{N} f_{ij}\ f_{ji}\ V'(q_{ij}), \quad q_{ij} \equiv q_i - q_j \tag{6.4}$$

$$\dot{f}_{ii} = 0, \qquad \dot{f}_{ij} = \sum_{\substack{k=1 \\ k\neq i,j}}^{N} f_{ik}\, f_{kj}\; [V(q_{ik}) - V(q_{jk})] + (f_{ii} - f_{jj}) f_{ij} V(q_{ij}), \quad i \neq j \tag{6.5}$$

From (6.5) we see that f_{ii} are integrals of motion, and we can restrict the system to the submanifold

$$f_{ii} = \alpha \tag{6.6}$$

where α is a constant *independent* of i. In this case the last term in eq.(6.5) vanishes. We will see that it is this reduced system which admits a Lax pair and is integrable.

Let us count the number of degrees of freedom. This is not a completely trivial matter due to the degeneracy of the Poisson bracket eq.(6.2), and the necessary reduction to the manifold $f_{ii} = \alpha$.

The Poisson bracket eq.(6.2) is a Kirillov bracket for the coadjoint action of the group $GL(N)$. The symplectic leaves are the coadjoint orbits. The orbits are generically characterized by the eigenvalues of the matrix $F = (f_{ij})$ which are in the center of the Poisson bracket. Here we shall consider matrices F of rank l, with l different non–vanishing eigenvalues. Moreover we shall assume that the matrix F is diagonalizable, hence the orbit is of the form

$$\{C\nu C^{-1} | C \in GL(N)\}, \quad \text{with } \nu = \mathrm{Diag}(\nu_1, \ldots, \nu_l, 0, \ldots, 0) \tag{6.7}$$

The Hamiltonian eq.(6.3) is not invariant under the above $GL(N)$ but it is is preserved by special subgroups. First we have the discrete subgroup of permutation matrices, i.e. the Weyl group, which simply operates by permutation of the N indices i. More importantly we have the group of diagonal matrices, i.e. the Cartan torus, which operates by:

$$f_{ij} \to d_i^{-1} f_{ij} d_j \tag{6.8}$$

This action preserves the Hamiltonian which only depends on $f_{ij} f_{ji}$. We consider the dynamical system on a rank l orbit and reduced under this diagonal action.

Proposition. *The dimension of the reduced phase space $\mathcal{M}$ is*

$$\dim \mathcal{M} = 2\left[Nl - \frac{l(l+1)}{2} + 1\right] \tag{6.9}$$

<u>Proof</u>. The tangent space to the orbit (6.7) at $F = (f_{ij})$ is the set of matrices $U = [F, X]$ for any $X \in gl(N)$. In a basis where F is diagonal this equation reads $U_{ij} = (\nu_i - \nu_j) X_{ij}$, hence U_{ij} vanishes when $\nu_i = \nu_j$ but is otherwise arbitrary. So the dimension of the orbit is $2Nl - l^2 - l$.

The action of the toral subgroup of diagonal matrices induces a fibering of the orbits into fibers of dimension $N-1$ since multiple of the identity leave F invariant. The moment associated to this action is the collection of diagonal elements f_{ii}, that is $N-1$ non trivial moments since on the orbit the eigenvalues of F are fixed, hence so is $\mathrm{Tr}(F)$. Indeed under infinitesimal action $d_i = 1 + \epsilon_i$ the f_{ij} change as $\delta f_{ij} = (\epsilon_j - \epsilon_i) f_{ij}$, and if $\mathcal{P} = \sum_i \epsilon_i f_{ii}$ we have $\{\mathcal{P}, f_{ij}\} = (\epsilon_j - \epsilon_i) f_{ij}$, so $\mathcal{P}$ is the corresponding Hamiltonian. We consider the reduced dynamical system obtained by first fixing the moments to a common value $f_{ii} = \alpha$, and then quotienting by the stabilizer of this moment which is the whole diagonal group. This is the integrable system as considered above.

We can now count the number of degrees of freedom. We have $2N$ degrees for the q_i, p_i, plus $2Nl - l^2 - l$ for the orbit, minus $2(N-1)$ due to the Hamiltonian reduction which ends up to a phase space of dimension $2(Nl - l(l+1)/2 + 1)$. ∎

6.2 Lax pair.

The first step in proving that the reduced spin Calogero model is integrable consists in finding a Lax pair formulation. Here we will just give the Lax pair and check its correctness, but in later sections we will explain methods to construct it.

Proposition. *The equations of motion of the spin Calogero–Moser system are equivalent to the Lax equation*

$$\dot{L}(\lambda) = [L(\lambda), M(\lambda)] \tag{6.10}$$

where the Lax matrices with spectral parameter λ are given by:

$$\begin{aligned} L_{ij}(t,\lambda) &= \dot{q}_i\delta_{ij} + (1-\delta_{ij})f_{ij}\Phi(q_i - q_j, \lambda). && (6.11)\\ M_{ij}(t,\lambda) &= (1-\delta_{ij})f_{ij}\Phi'(q_i - q_j, \lambda). && (6.12)\end{aligned}$$

Here

$$\Phi(q,\lambda) = \frac{\sigma(\lambda - q)}{\sigma(\lambda)\sigma(q)}e^{\zeta(\lambda)q} \tag{6.13}$$

and the prime in eq.(6.12) refers to the derivative with respect to q.

Proof. The function Φ occurring in the Lax pair is known as the Lamé function and some of its poperties are described at the end of Chapter [15]. In particular it is an elliptic function of the spectral parameter λ and satisfies the equation

$$\left(\frac{d^2}{dx^2} - 2\wp(x)\right)\Phi(x,\lambda) = \wp(\lambda)\Phi(x,\lambda) \tag{6.14}$$

The Lax equation (6.10) reads:

$$\begin{aligned} \ddot{q}_i\delta_{ij} + (1-\delta_{ij})\left[\dot{f}_{ij}\Phi(q_{ij},\lambda) + f_{ij}\dot{q}_{ij}\Phi'(q_{ij},\lambda)\right] &= \\ (1-\delta_{ij})f_{ij}\dot{q}_{ij}\Phi'(q_{ij},\lambda) + \sum_{k\neq i,j} f_{ik}f_{kj}\left[\Phi(q_{ik},\lambda)\Phi'(q_{kj},\lambda) - \Phi'(q_{ik},\lambda)\Phi(q_{kj},\lambda)\right] \end{aligned}$$

This reduces straightforwardly to the equations of motion (in the case where $f_{ii} = \alpha$), if we use an identity satisfied by the Lamé function:

$$\Phi'(x,\lambda)\Phi(y,\lambda) - \Phi'(y,\lambda)\Phi(x,\lambda) = [\wp(y) - \wp(x)]\Phi(x+y,\lambda) \tag{6.15}$$

Taking the limit $y \to -x$ this also implies:

$$\Phi'(x,\lambda)\Phi(-x,\lambda) - \Phi'(-x,\lambda)\Phi(x,\lambda) = -\wp'(x)$$

To show this identity we compare the analyticity and monodromy properties of both members of the equation in the x variable, using the properties given in Chapter [15]. Both members tend to $x^{-2}\Phi(y,\lambda)$ when $x \to 0$, are multiplied by $\exp 2(\zeta(\lambda)\omega_l - \eta_l\lambda)$ when $x \to x + 2\omega_l$ and vanish when $x = y$. ■

By a similar method one shows that the Lamé function obeys the identity:

$$\Phi(q,\lambda)\Phi(-q,\lambda) = \wp(\lambda) - \wp(q) \tag{6.16}$$

hence $\frac{1}{2}\mathrm{Tr}\, L^2 = H + \wp(\lambda)\sum_{i<j} f_{ij}f_{ji}$, where H is the Hamiltonian of the spin Calogero model, and $\sum_{i\neq j} f_{ij}f_{ji} = \mathrm{Tr}\, F^2 - N\alpha^2$ is an orbit invariant.

Let us comment on the trigonometric and rational limits of the above formulae. The trigonometric limit is obtained when one of the periods $\omega \to \infty$. We choose the other one as $i\pi$. In this limit the function Φ becomes:

$$\Phi(q, \lambda) = (\coth\, q - \coth\, \lambda)\, e^{q \coth\, \lambda}$$

The exponential factor in $\Phi(q, \lambda)$ comes from the factor $\exp(\zeta(\lambda)q)$ which is necessary in the elliptic case to ensure the double periodicity of $\Phi(q, \lambda)$ in λ. In the trigonometric case however this exponential factor can be eliminated by performing a similarity transformation on $L(\lambda)$ without affecting the periodicity properties of the matrix elements of $L(\lambda)$. So we may define

$$L_{\text{trigo}}(\lambda) = \text{Diag}\,(e^{-q_i \coth\lambda}) \lim_{\omega\to\infty} L_{\text{elliptic}}(\lambda)\ \text{Diag}\,(e^{q_i \coth\lambda}) \tag{6.17}$$

The functions $V(q)$ and $\Phi(q, \lambda)$ become:

$$V(q) = \frac{1}{\sinh^2(q)}, \quad \Phi(q, \lambda) = \coth q - \coth \lambda, \tag{6.18}$$

The rational limit is obtained straightforwardly from the trigonometric limit by sending the second period $\omega' \to \infty$. The functions $V(q)$ and $\Phi(q, \lambda)$ become

$$V(q) = \frac{1}{q^2}, \quad \Phi(q, \lambda) = \frac{1}{q} - \frac{1}{\lambda}, \tag{6.19}$$

and of course $\coth\, \lambda \to 1/\lambda$ in eq (6.18).

6.3 The r–matrix

In this section we compute the r–matrix associated to the Lax matrix

$$L(\lambda) = \sum_{i=1}^{N} p_i e_{ii} + \sum_{\substack{i,j=1\\ i\neq j}}^{N} \Phi(q_{ij}, \lambda)\ f_{ij}\ e_{ij}. \tag{6.20}$$

thereby proving the integrability of the elliptic spin Calogero model. One should emphasize that it is only the reduced system ($f_{ii} = \alpha$) which is integrable. However the Lax matrix (6.20) is not a function on this reduced phase space and so is not expected to have an r–matrix in the usual sense. This accounts for the extra last term in eq.(6.21) below:

Proposition. *The Poisson brackets of the Lax matrix (6.20) are given by:*

$$\{L_1(\lambda), L_2(\mu)\} \quad = \quad [r_{12}(\lambda, \mu), L_1(\lambda)] - [r_{21}(\mu, \lambda), L_2(\mu)] + [\mathcal{D}, r_{12}(\lambda, \mu)] \tag{6.21}$$

where we have $\mathcal{D} = \sum_i f_{ii} \frac{\partial}{\partial q_i}$ and the r–matrix is expressed as:

$$r_{12}(\lambda, \mu) = a(\lambda, \mu) \sum_{i=1}^{N} e_{ii} \otimes e_{ii} - \sum_{\substack{i,j=1\\ i\neq j}}^{N} b(q_{ij}, \lambda, \mu)\ e_{ij} \otimes e_{ji} \tag{6.22}$$

with $a(\lambda, \mu) = \zeta(\lambda-\mu)-\zeta(\lambda)+\zeta(\mu)$ and $b(q, \lambda, \mu) = \frac{\sigma(\lambda-\mu-q)}{\sigma(\lambda-\mu)\sigma(q)} e^{(\zeta(\lambda)-\zeta(\mu))q}$. Here σ and ζ are the Weierstrass functions defined in Chapter [15].

Remark that $a(\lambda, \mu)$ and $b(q, \lambda, \mu)$ are true elliptic functions of both λ and μ. Note also that the r–matrix is antisymmetric, i.e. $r_{12}(\lambda, \mu) = -r_{21}(\mu, \lambda)$. The most remarkable feature of this r–matrix is that it depends on the dynamical variables q_{ij}, hence the name dynamical r–matrix.

Proof. We compute the various terms of eq.(6.21), and collect them according to the number of equal matrix indices. We have (all written indices are different, and sums are implied):

$$\begin{aligned}\{L_1(\lambda), L_2(\mu)\} &= \Phi'(q_{ij},\mu)f_{ij}(e_{ii}-e_{jj})\otimes e_{ij} - \Phi'(q_{ij},\lambda)f_{ij}e_{ij}\otimes(e_{ii}-e_{jj})\\ &+\Phi(q_{ij},\lambda)\Phi(q_{ki},\mu)f_{kj}e_{ij}\otimes e_{ki} - \Phi(q_{ij},\lambda)\Phi(q_{jk},\mu)f_{ik}e_{ij}\otimes e_{jk}\\ &-\Phi(q_{ij},\lambda)\Phi(q_{ji},\mu)(f_{ii}-f_{jj})e_{ij}\otimes e_{ji}\end{aligned}$$

Similarly one has, using the antisymmetry of the r–matrix:

$$\begin{aligned}[r_{12}, L_1(\lambda)+L_2(\mu)] &= \Big(a(\lambda,\mu)\Phi(q_{ij},\mu)+b(q_{ji},\lambda,\mu)\Phi(q_{ij},\lambda)\Big)f_{ij}(e_{ii}-e_{jj})\otimes e_{ij}\\ &+\Big(a(\lambda,\mu)\Phi(q_{ij},\lambda)+b(q_{ij},\lambda,\mu)\Phi(q_{ij},\mu)\Big)f_{ij}e_{ij}\otimes(e_{ii}-e_{jj})\\ &+\Big(b(q_{ij},\lambda,\mu)\Phi(q_{kj},\mu)-b(q_{ik},\lambda,\mu)\Phi(q_{kj},\lambda)\Big)f_{kj}e_{ij}\otimes e_{ki}\\ &+\Big(b(q_{kj},\lambda,\mu)\Phi(q_{ik},\lambda)-b(q_{ij},\lambda,\mu)\Phi(q_{ik},\mu)\Big)f_{ik}e_{ij}\otimes e_{jk}\end{aligned}$$

Finally we have:

$$[\mathcal{D}, r_{12}] = -b'(q_{ij},\lambda,\mu)(f_{ii}-f_{jj})e_{ij}\otimes e_{ji}$$

Using the symmetry properties $a(\mu,\lambda) = -a(\lambda,\mu)$ and $b(-q,\mu,\lambda) = -b(q,\lambda,\mu)$ we see that the eq.(6.21) boils down to the three identities:

$$\begin{aligned}\Phi'(q,\mu) &= a(\lambda,\mu)\Phi(q,\mu)+b(-q,\lambda,\mu)\Phi(q,\lambda)\\ \Phi(q,\lambda)\Phi(q',\mu) &= b(q,\lambda,\mu)\Phi(q+q',\mu)-b(-q',\lambda,\mu)\Phi(q+q',\lambda)\\ b'(q,\lambda,\mu) &= \Phi(q,\lambda)\Phi(-q,\mu)\end{aligned}$$

The first identity, written in terms of Weierstrass functions reads:

$$\frac{\sigma(\lambda-q)\sigma(\mu)\sigma(\lambda-\mu+q)}{\sigma(\lambda)\sigma(\mu-q)\sigma(\lambda-\mu)\sigma(q)} = \zeta(\mu-q)+\zeta(q)+\zeta(\lambda-\mu)-\zeta(\lambda)$$

One observes that both sides are elliptic functions of λ (and μ) and have same poles and residues, hence are equal. The second identity reads:

$$\sigma(q+q')\sigma(\lambda-\mu)\sigma(\lambda-q)\sigma(\mu-q') = \sigma(\lambda-\mu-q)\sigma(q')\sigma(\lambda)\sigma(\mu-q-q')+\sigma(\lambda-\mu+q')\sigma(\mu)\sigma(q)\sigma(\lambda-q-q')$$

which is true because both sides vanish at $\lambda = q$ and $\lambda = \mu$, are equal for $\lambda = 0$, and have the same monodromy properties when shifting λ by periods. Finally the third identity reduces to the first one. ∎

Remark. Eq.(6.21) holds for the non reduced dynamical system (6.3). We see that $\mathrm{Tr}\, L^n(\lambda)$ are *not* in involution for this system. Indeed we have:

$$\{\mathrm{Tr}\, L^n(\lambda), \mathrm{Tr}\, L^m(\mu)\} = -n\; m \sum_{\substack{i,j=1\\ i\neq j}}^{N} \Phi(q_{ij},\lambda)\Phi(q_{ji},\mu)(f_{ii}-f_{jj})\; [L^{n-1}(\lambda)]_{ij}\; [L^{m-1}(\mu)]_{ji} \qquad (6.23)$$

As we have already seen, the f_{ii} are the moments of the toral group, which acts on $L(\lambda)$ by conjugation. Hence the quantities $\mathrm{Tr}\, L^n(\lambda)$ are invariant under this action. It follows that one can compute very simply, see Chapter [14], their reduced Poisson bracket on the manifold $(f_{ii} = \alpha)_{i=1\cdots N}$: we just set $f_{ii} = \alpha$ in eq.(6.23) and therefore $\mathrm{Tr}\, L^n(\lambda)$ are in involution for the reduced system. This proves integrability of the spin Calogero model (we will count the number of action variables later on). ∎

Proposition. *The classical r-matrix satisfies the identity*

$$-\{L_1, r_{23}\} + \{L_2, r_{13}\} - \{L_3, r_{12}\} + [r_{12}, r_{13}] + [r_{12}, r_{23}] + [r_{13}, r_{23}] = 0 \tag{6.24}$$

Proof. It is done by direct calculation. It is more interesting however to see how this is related to the Jacobi identity. Let us call

$$Z_{12} = [\mathcal{D}, r_{12}], \quad \mathcal{D} = \sum_i f_{ii} \frac{\partial}{\partial q_i}$$

The Jacobi identity reads

$$\begin{aligned}
[L_1, [r_{12}, r_{23}] + [r_{12}, r_{13}] + [r_{32}, r_{13}] + \{L_2, r_{13}\} - \{L_3, r_{12}\}] + \text{cyclic permutations} \\
+[r_{23}, Z_{12}] + [r_{31}, Z_{23}] + [r_{12}, Z_{31}] - [r_{32}, Z_{13}] - [r_{13}, Z_{21}] - [r_{21}, Z_{32}] \\
+\{L_1, Z_{23}\} + \{L_2, Z_{31}\} + \{L_3, Z_{12}\} \qquad = 0
\end{aligned}$$

Note that the term commuted with L_1 is not quite equal to the left–hand side of eq.(6.24). We want to show that the missing term, $\{L_1, r_{23}\}$, is produced by all the Z contributions. Since $r_{12}(\lambda, \mu) = -r_{21}(\mu, \lambda)$, we find

$$\begin{aligned}
[r_{23}, Z_{12}] + [r_{31}, Z_{23}] + [r_{12}, Z_{31}] - [r_{32}, Z_{13}] - [r_{13}, Z_{21}] - [r_{21}, Z_{32}] \qquad = \\
-[\mathcal{D}, [r_{12}, r_{13}] + [r_{12}, r_{23}] + [r_{13}, r_{23}]]
\end{aligned}$$

Moreover, we have using $\{L_1, \mathcal{D}\} = -\sum_k [L_1, e_{kk}]\partial_{q_k}$:

$$\{L_1, Z_{23}\} = \{L_1, [\mathcal{D}, r_{23}]\} = [\mathcal{D}, \{L_1, r_{23}\}] + [\{L_1, \mathcal{D}\}, r_{23}] = [\mathcal{D}, \{L_1, r_{23}\}] - [L_1, \{L_1, r_{23}\}]$$

So that the Jacobi identity becomes

$$\begin{aligned}
[L_1, [r_{12}, r_{13}] + [r_{12}, r_{23}] + [r_{13}, r_{23}] - \{L_1, r_{23}\} + \{L_2, r_{13}\} - \{L_3, r_{12}\}] + c.p. \\
-[\mathcal{D}, [r_{12}, r_{13}] + [r_{12}, r_{23}] + [r_{13}, r_{23}] - \{L_1, r_{23}\} + \{L_2, r_{13}\} - \{L_3, r_{12}\}] \qquad = 0
\end{aligned}$$

As a result, the Jacobi identity is satisfied if eq.(6.24) holds, thereby giving a motivation to the computation which shows its vanishing. ■

Remark. Let us comment on the trigonometric limit of these formulae. Using the Lax matrix eq.(6.17) we find

$$\{L_1(\lambda), L_2(\mu)\} = [r_{12}(\lambda, \mu), L_1(\lambda)] - [r_{21}(\mu, \lambda), L_2(\mu)] + \sum_{\substack{i,j=1 \\ i \neq j}}^{N} \frac{f_{ii} - f_{jj}}{\sinh^2(q_i - q_j)}\, e_{ij} \otimes e_{ji} \tag{6.25}$$

where

$$r_{12}(\lambda, \mu) = \coth(\lambda - \mu)\mathcal{C} - \sum_{\substack{i,j=1 \\ i \neq j}}^{N} \coth(q_i - q_j) e_{ij} \otimes e_{ji} \tag{6.26}$$

and $\mathcal{C}$ is the Casimir element of $sl(N)$ $\mathcal{C} = \sum_{i,j=1}^{N} e_{ij} \otimes e_{ji}$. This r–matrix will reappear in the chapter on Toda field theory [11], and there, it will be clear that eq.(6.24) is satisfied.

6.4 The standard Calogero–Moser model.

We show here that the standard elliptic Calogero model and its r–matrix can be obtained from eq.(6.21) by a Hamiltonian reduction procedure.

Quite generally, we can parametrize the matrix $F = (f_{ij})$ of rank l as follows

$$f_{ij} = \sum_{r=1}^{l} b_i^r a_j^r \tag{6.27}$$

Moreover, the Poisson bracket eq.(6.2) is reproduced if we set

$$\{a_i^r, b_j^s\} = -\delta_{rs}\delta_{ij}$$

The equations of motion for the quantities a_i^r, b_i^r read

$$\dot{a}_i^r = \{H, a_i^r\} = -\sum_{k\neq i} V(q_{ik}) f_{ki} a_k^r, \quad \dot{b}_i^r = \{H, b_i^r\} = \sum_{k\neq i} V(q_{ik}) f_{ik} b_k^r$$

These equations of motion reproduce eq.(6.5) on the reduced manifold $f_{ii} = \alpha$.

To recover the scalar Calogero model, we choose $l = 1$ in eq.(6.27). So the vectors a_i and b_i have only one component. We simply call them a_i, b_i.

Recall the action of the toral Lie group eq.(6.8) which on the variables a_i, b_i reads:

$$a_i \longrightarrow d_i\, a_i, \quad b_i \longrightarrow d_i^{-1} b_i \tag{6.28}$$

The integrable system is obtained by applying the method of Hamiltonian reduction under this group. We fix the momentum to $f_{ii} = a_i b_i = \alpha = \sqrt{-1}\gamma$, which removes N degrees of freedom, and quotient by the isotropy subgroup G_α, so that at the end the $2N$ degrees of freedom $(a_i, b_i)_{i=1\cdots N}$ are eliminated, leaving as reduced Hamiltonian (6.3) the usual Calogero–Moser model (note that $f_{ij}f_{ji} = \alpha^2$)

$$H_{\mathrm{Cal}} = \frac{1}{2}\sum_{i=1}^{N} p_i^2 + \frac{1}{2}\,\gamma^2 \sum_{\substack{i,j=1\\ i\neq j}}^{N} V(q_i - q_j)$$

In order to perform the reduction at the level of the Lax matrix, we remark that the matrix

$$\begin{aligned} L^{Cal}(\lambda) &= g^{-1} L(\lambda)\, g \qquad \text{with } g = \mathrm{Diag}\,(a_i)_{i=1\cdots N} \\ &= \sum_{i=1}^{N} p_i\, e_{ii} + \alpha \sum_{\substack{i,j=1\\ i\neq j}}^{N} \Phi(q_{ij,\lambda})\, e_{ij} \end{aligned} \tag{6.29}$$

is invariant under the isotropy group G_α and we can compute the Poisson brackets of its matrix elements directly. We find a true r–matrix relation:

$$\{L_1^{Cal}(\lambda), L_2^{Cal}(\mu)\} = [r_{12}^{Cal}(\lambda,\mu), L_1^{Cal}(\lambda)] - [r_{21}^{Cal}(\mu,\lambda), L_2^{Cal}(\mu)] \tag{6.30}$$

with (see eq.(1.9) in Chapter [1]):

$$r_{12}^{Cal}(\lambda,\mu) = g_1^{-1} g_2^{-1} \left[\, r_{12}(\lambda,\mu) - \{g_1, L_2(\mu)\} g_1^{-1} + \frac{1}{2}[u_{12}, L_2(\mu)]\right] g_1 g_2$$

where $u_{12} = \{g_1, g_2\} g_1^{-1} g_2^{-1}$ is here equal to zero. We get

$$r_{12}^{Cal}(\lambda,\mu) = r_{12}(\lambda,\mu) - \sum_{i\neq j} \Phi(q_{ij},\mu) e_{ii} \otimes e_{ij}$$

Redefining

$$r_{12}^{Cal}(\lambda,\mu) \longrightarrow r_{12}^{Cal}(\lambda,\mu) + \left[\frac{1}{2\alpha}\sum_{i=1}^{N} e_{ii}\otimes e_{ii}, L_2^{Cal}(\mu)\right]$$

does not change eq.(6.30) and yields the r-matrix of the scalar Calogero model.

$$r_{12}^{Cal}(\lambda,\mu) = a(\lambda,\mu)\sum_i e_{ii}\otimes e_{ii} - \sum_{\substack{i,j=1\\ i\neq j}}^{N} b(q_{ij},\lambda-\mu)\; e_{ij}\otimes e_{ji} - \frac{1}{2}\sum_{\substack{i,j=1\\ i\neq j}}^{N} \Phi(q_{ij},\mu)\;(e_{ii}+e_{jj})\otimes e_{ij}$$

This is still a dynamical r–matrix but is not antisymmetric any more.

Finally, we give the formula for the matrix $M^{Cal} = g^{-1}Mg + g^{-1}\dot{g}$. We find $M_{ij}^{Cal} = \dot{a}_i/a_i\delta_{ij} + (1-\delta_{ij})\alpha\Phi'(q_i-q_j,\lambda)$. Using the equation of motion $\dot{a}_i = -\sum_{k\neq i}\alpha V(q_i-q_k)a_i$, we get

$$M^{Cal} = -\alpha\delta_{ij}\sum_{k\neq i} V(q_i-q_k) + (1-\delta_{ij})\alpha\Phi'(q_i-q_j,\lambda)$$

6.5 The spectral curve.

The spectral curve is defined as usual:

$$\Gamma : R(\lambda,\mu) \equiv \det\left(L(t,\lambda)-\mu I\right) = 0 \tag{6.31}$$

This defines a curve Γ which is time–independent due to the Lax equation eq.(6.10). Remark that Γ is invariant under the symmetries eq.(6.8).

Proposition. *The equation of the spectral curve takes the form:*

$$R(\lambda,\mu) \equiv \sum_{i=0}^{N} r_i(\lambda)\mu^i = 0 \tag{6.32}$$

where the $r_i(\lambda)$ are elliptic functions of λ, independent of t, which can be expanded on the Weierstrass $\wp$ function and its derivatives as:

$$r_i(\lambda) = I_i^0 + \sum_{s=0}^{N-i-2} I_{i,s}\partial_\lambda^s\wp(\lambda). \tag{6.33}$$

In a neighbourhood of $\lambda=0$, the function $R(\lambda,\mu)$ can be factorized as:

$$R(\lambda,\mu) = \prod_{i=1}^{N}(\mu-\mu_i(\lambda)); \quad \mu_i(\lambda) = (\alpha-\nu_i)\lambda^{-1} + h_i(\lambda); \tag{6.34}$$

where $h_i(\lambda)$ are regular functions of λ, and the ν_i are the eigenvalues of the matrix $F=(f_{ij})$, so that $\nu_i = 0,\quad i>l$.

Proof. The matrix elements $L_{ij}(t,\lambda)$ of the Lax matrix are elliptic functions of the variable λ having an essential singularity at $\lambda=0$, but the functions $r_i(\lambda)$ are meromorphic because the essential singularity can be gauged away near $\lambda=0$ since we can write:

$$L(t,\lambda) = G(t,\lambda)\widetilde{L}(t,\lambda)G^{-1}(t,\lambda),\; G_{ij} = \delta_{ij}\exp(\zeta(\lambda)q_i(t)), \tag{6.35}$$

where $\widetilde{L}_{ij}(t,\lambda)$ are meromorphic functions of λ in a neighbourhood of the point $\lambda=0$. Incorporating the constraint $f_{ii}=\alpha$ we have:

$$\widetilde{L}(t,\lambda) = \frac{1}{\lambda}(\alpha I - F(t)) + O(\lambda^0) \tag{6.36}$$

where $F(t)$ is the matrix of elements $f_{ij}(t)$. Therefore the elliptic functions $r_i(\lambda)$ have poles of degree at most $N-i$ at the point $\lambda = 0$, so that they can be expanded as a linear combinations of the function $\wp(\lambda)$ and its derivatives. The coefficients $I_i^0, I_{i,s}$ of this expansion are the integrals of motion of the spin Calogero model. They define the moduli of the algebraic curve Γ. We can always factorize the polynomial in μ, $R(\lambda, \mu)$ around $\lambda = 0$. The branches $\mu_i(\lambda)$ in eq.(6.34) have simple poles, of the stated form due to eq.(6.36). In particular since F is of rank l, the eigenvalue $\nu_i = 0$ has multiplicity $N-l$. ■

We are now in a position to compute the number of the action variables.

Proposition. *The number of the action variables is:* $Nl - l(l+1)/2 + 1 = \frac{1}{2}\dim \mathcal{M}$

Proof. The spectral equation $R(\lambda, \mu) = 0$ depends on $N(N+1)/2$ parameters I_i^0, I_{is}. However they are not all independent. The constraints come from the conditions $\nu_i = 0, \;\; i > l$ in eq.(6.34). To see how they translate on the parameters I_i^0, I_{is}, let us introduce the variable $\widetilde{\mu} = \mu - \alpha\lambda^{-1}$. Then we have $R(\lambda, \widetilde{\mu} + \alpha\lambda^{-1}) = \widetilde{R}(\lambda, \widetilde{\mu})$ which can be expanded as:

$$\widetilde{R}(\lambda, \widetilde{\mu}) = \sum_{i=0}^{N} \widetilde{R}_i(\widetilde{\mu})\lambda^{-i} + \mathcal{R}(\lambda, \widetilde{\mu}), \tag{6.37}$$

where $\widetilde{R}_i$ are polynomials in $\widetilde{\mu}$ and $\mathcal{R}(\lambda, \widetilde{\mu}) = O(\lambda)$ is regular at $\lambda = 0$. One can check easily that $\widetilde{R}_i(\widetilde{\mu})$ is of degree $\deg \widetilde{R}_i = N - i$ and has coefficients which are linear combinations of the parameters I_i^0, I_{is}.
The conditions $\nu_i = 0, \; i > l$ imply that

$$\widetilde{R}_i(\widetilde{\mu}) = 0, \quad i > l. \tag{6.38}$$

Altogether this is equivalent to a set of $(N-l)(N-l+1)/2$ linear equations on the parameters I_i^0, I_{is}. The total number of independent parameters is therefore equal to $Nl - l(l-1)/2$.
Next, remark that the expansion around $\lambda = 0$ of the branch μ_i for $i = 1, \cdots, l$ reads $\mu_i = (\alpha - \nu_i)\lambda^{-1} + O(1)$. This yields $l-1$ additional relations (since the ν_i are constants characterizing the orbit of F, not to be counted as dynamical variables). Note that this gives only $l-1$ constraints, and not l, because we have $N\alpha = \sum \nu_i = \operatorname{Tr} F$. This condition also accounts for the fact that the elliptic function $r_{N-1}(\lambda)$ is constant, since it cannot have a single pole. Hence the number of independent parameters is finally equal to $Nl - l(l-1)/2 - (l-1)$ which is exactly half the dimension of the reduced phase space. ■

We now compute the genus of the spectral curve Γ.

Proposition. *For generic values of the action variables the genus of the spectral curve is given by:*

$$g = Nl - \frac{l(l+1)}{2} + 1 \tag{6.39}$$

Proof. The idea of the proof is the same as in Chapter [4] and uses the Riemann–Hurwitz theorem. There is a difference however, because here the base curve is of genus 1. Equation (6.31) presents the compact Riemann surface Γ as an N-sheeted branched covering of the base curve of the variable λ. The sheets are the N roots in μ. By the Riemann–Hurwitz formula we have $2g - 2 = N(2g_0 - 2) + \nu$ where g_0 is the genus of the base curve, i.e., $g_0 = 1$, and ν is the number of branch points, i.e. the number of values of λ for which $R(\lambda, \mu)$ has a double root in μ. This is the number of zeroes of $\partial_\mu R(\lambda, \mu)$ on the surface $R(\lambda, \mu) = 0$. But $\partial_\mu R(\lambda, \mu)$ is a meromorphic function on the surface, hence it has as many zeroes as poles. The poles are located above $\lambda = 0$ or $\mu = \infty$ which is the same, and are easy to count.
Let P_i be the points of Γ lying on the different sheets over the point $\lambda = 0$. In the neighbourhood of P_i the

function μ has the expansion $\mu_i = (\alpha - \nu_i)\lambda^{-1} + h_i(\lambda)$. Hence, the function $\partial R/\partial\mu$ in the neighbourhood of P_i has the form

$$\partial R/\partial\mu = \prod_{j\neq i}[(\nu_j - \nu_i)\lambda^{-1} - (h_j(\lambda) - h_i(\lambda))].$$

From this we see that on each of the l sheets $(\lambda, \mu_i(\lambda))$ $(i = 1, \cdots, l)$ we have one pole of order $(N-1)$. On each of the $(N-l)$ sheets $(\lambda, \mu_i(\lambda))$ $(i = l+1, \cdots, N)$ we have one pole of order l. Finally $\nu = l(N-1) + (N-l)l$. Inserting this value in the Riemann–Hurwitz theorem yields the result. ∎

6.6 The eigenvector bundle.

As in Chapter [4] we consider at any point $P = (\lambda, \mu)$ of the spectral curve the unique eigenvector $\Psi(0, P)$ of $L(0, \lambda)$ with eigenvalue μ, normalized by $\psi_1(0, P) = 1$. We want to study the analyticity properties in λ of this eigenvector. A first difference with the rational case is that it has an essential singularity at the points P_i above $\lambda = 0$.

Proposition. *In the neighbourhood of the point P_i the component $\psi_j(0, P)$ has the form*

$$\psi_j(0, P) = \exp[\zeta(\lambda)(q_j(0) - q_1(0))](c_j^{(i)} + O(\lambda)), \tag{6.40}$$

where for $i = 1, \cdots, l$, $c_j^{(i)}$ is the eigenvector of the matrix F corresponding to the non–zero eigenvalue ν_i i.e.,

$$\sum_{j=1}^{N} f_{kj} c_j^{(i)} = \nu_i c_k^{(i)}$$

while for $i > l$ the $c^{(i)}$ form a basis of the kernel of F.

Proof. From equation (6.35), we have $\Psi(0, P) = G(0, \lambda)\widetilde{\Psi}(0, P)$, where $\widetilde{\Psi}(0, P)$ is an eigenvector of $\widetilde{L}(0, \lambda)$. Using equation (6.36), we have $\widetilde{\Psi}(0, P) = c^{(i)} + O(\lambda)$ where $c^{(i)}$ is an eigenvector of F. More precisely for $i = 1, \cdots, l$, $c^{(i)}$ is the unique eigenvector of F with eigenvalue $\nu_i \neq 0$ normalized by $c_1^{(i)} = 1$, while for $i > l$ it is determined by the limit for $P \to P_i$ of the normalized eigenvector of $\widetilde{L}(P)$ corresponding to the eigenvalue $\mu_i(P)$. This is an example of the lifting of degeneracy by higher order terms in λ. Therefore we have $\psi_j(0, P) = (c_j^{(i)} + O(\lambda)) \exp(\zeta(\lambda)q_j(0))$. Normalizing $\psi_1(0, P) = 1$ yields the result. ∎

We can now compute the number of poles of Ψ on Γ. The result is slightly different as in the rational case (where it was found to be $g + N - 1$).

Proposition. *The number of poles of $\Psi(0, P)$ is:*

$$m = Nl - \frac{l(l+1)}{2} = g - 1 \tag{6.41}$$

Proof. As in the rational case, let us introduce the function $W(\lambda)$ of the complex variable λ defined by:

$$W(\lambda) = (\mathrm{Det}\,|\psi_i(M_j)|)^2$$

where the M_j's are the N points above λ. It is well–defined on the base curve since the Det^2 does not depend on the order of the M_j's.
This function has an essential singularity at $\lambda = 0$ of the form $\exp\, 2\zeta(\lambda)\sum_i(q_i(0) - q_1(0))$. This does not affect

the property that the number of poles is equal to the number of zeroes, because this property is obtained by considering the sum of residues of W'/W on the fundamental torus, and noting that $\zeta'(z) = \wp(z) = \frac{1}{z^2} + O(z^2)$ is elliptic and has no residue. Clearly $W(\lambda)$ has a double pole where there exists a point M above λ at which $\Psi(M)$ has a simple pole.
As in the rational case, we show that $W(\lambda)$ has a simple zero for values of λ corresponding to a branch–point of the covering, hence $m = \nu/2 = g-1$, by Riemann–Hurwitz (here lies the difference with the rational case, because $g_0 = 1$).
First notice that $W(\lambda)$ only vanishes on branch–points, where there are at least two identical columns. Indeed, let $M_i = (\lambda, \mu_i)$ be the N points above λ. Then the $\Psi(M_i)$ are the non–vanishing eigenvectors of $L(\lambda)$ corresponding to the eigenvalues μ_i hence are linearly independent when all the μ_i's are different. Therefore $W(\lambda)$ cannot vanish at such a point. Let us assume now that λ corresponds to a branch point, which is generically of order 2. At such a point $W(\lambda)$ has a simple zero. Indeed let z be an analytical parameter on the curve around the branch point. The covering projection $M \to \lambda$ gets expressed as $\lambda = \lambda_0 + \lambda_1 z^2 + O(z^3)$. The determinant vanishes to order z^1 hence W vanishes to order z^2, but this is precisely proportional to $\lambda - \lambda_0$. ■

This is a very surprising result. In fact, from eq.(6.40), we see that the components of the eigenvector $\psi_i(0, P)$ are Baker-Akhiezer functions. But generically, such a function has at least g poles, not $g-1$. The only possible way to escape this contradiction is to admit that we are not in a generic situation. Indeed it was noted in Chapter [2] that when the genus g of the base curve is greater than or equal to 1, the equations of consistency of Lax pairs become overdetermined, preventing genericity of the Lax matrix. In fact, this will be a crucial ingredient in the solution of the Calogero model. But before explaining this, let us describe the time evolution.

6.7 Time evolution.

The next step is to compute the time evolution of the above eigenvectors. We let the eigenvector evolve according to the Lax equation:

$$\frac{d\Psi}{dt} = -M\Psi \tag{6.42}$$

We choose as initial condition the eigenvector $\Psi(0, P)$ normalized with its first component equal to 1. Of course at subsequent time this normalization will not hold any more. However we know that, in this setting, the poles of $\Psi(t, P)$ do not evolve with time, see Chapter [4].

Proposition. *The coordinates $\psi_j(t, P)$ of the vector-function $\Psi(t, P)$ are meromorphic functions on Γ except at the points P_i above $\lambda = 0$. Their poles $\gamma_1, \ldots, \gamma_{g-1}$ do not depend on t. In the neighbourhood of P_i they have the form*

$$\psi_j(t, P) = c_j^{(i)}(t, \lambda) \exp\left[\zeta(\lambda)(q_j(t) - q_1(0)) + m_i(\lambda)t\right], \tag{6.43}$$

where $c_j^{(i)}(t, \lambda)$ are regular functions of λ for $\lambda \sim 0$, $c_j^{(i)}(t, \lambda) = c_j^{(i)}(t) + O(\lambda)$. Here the $c^{(i)}(t)$ are eigenvectors of the matrix $F(t) = (f_{ij}(t))$ corresponding to the eigenvalues ν_i, and $m_i(\lambda)$ with:

$$m_i(\lambda) = (-\alpha + \nu_i)\lambda^{-2} - h_i(0)\lambda^{-1} \tag{6.44}$$

is the singular part of $-\lambda^{-1}\mu_i(\lambda)$ at $\lambda = 0$.

<u>Proof</u>. Let us consider the vector $\widetilde{\Psi}(t, P)$ defined as

$$\Psi(t, P) = G(t, \lambda)\widetilde{\Psi}(t, P), \tag{6.45}$$

where $G(t,\lambda)$ is defined in eq.(6.35) and $\Psi(t,P)$ evolves according to eq.(6.42). The vector $\widetilde{\Psi}(t,P)$ is an eigenvector of the matrix $\widetilde{L}(t,\lambda)$ and evolves according to the equation

$$(\partial_t + \widetilde{M}(t,\lambda))\widetilde{\Psi}(t,P) = 0, \quad \widetilde{M} = G^{-1}\partial_t G + G^{-1}MG. \tag{6.46}$$

From (6.11,6.12) it follows that:

$$\widetilde{M}_{ij} = \delta_{ij}\zeta(\lambda)\dot{q}_i + (1-\delta_{ij})f_{ij}\frac{\sigma(\lambda - q_{ij})}{\sigma(\lambda)\sigma(q_{ij})}[\zeta(q_{ij}-\lambda) - \zeta(q_{ij}) + \zeta(\lambda)]$$

so that collecting the coefficient of the $1/\lambda$ term we can write:

$$\widetilde{M}(t,\lambda) = \lambda^{-1}\widetilde{L}(t,\lambda) + O(\lambda^0). \tag{6.47}$$

Hence around P_i we have:

$$\partial_t\widetilde{\Psi}(t,\lambda) = (-\frac{1}{\lambda}\mu_i(\lambda) + O(1))\widetilde{\Psi}(t,P)$$

so that:

$$m_i(\lambda) = -(\lambda^{-1}\mu_i(\lambda))_- = (-\alpha + \nu_i)\lambda^{-2} - h_i(0)\lambda^{-1}$$

is independent of time. Multiplying by $G(t,\lambda)$ and normalizing $\psi_1(0,P) = 1$, we get the result. ∎

6.8 Reconstruction formulae.

We are now in a position to reconstruct the original dynamical variables in terms of the Riemann surface and the poles of the eigenvectors. From eq.(6.43), we see that the components $\psi_i(t,P)$ of the eigenvector are Baker-Akhiezer functions. Their behaviour above $\lambda = 0$ is

$$\psi_i(P) = \exp[(q_i(t) - q_1(0))\lambda^{-1} + m_j(\lambda)t](c_i^{(j)}(t) + O(\lambda)) \quad P \to P_j \tag{6.48}$$

As we already mentioned, there is a problem with the number of poles of the Baker–Akhiezer function. Generically, such a function has g poles. The function then exists and is unique up to normalization. In particular, its zeroes are completely determined. One way to construct a Baker–Akhiezer function with $g-1$ poles is to let one of the zeroes of the generic function cancel one of its poles. Clearly, this gives a relation between the parameters defining the function, namely the $q_i(t)$ defining the essential singularity and the moduli of the curve.

Let $\psi_i(t,P)$ the Baker-Akhiezer function which has the singularity eq.(6.48) at the points P_j, and a divisor of g poles $(\gamma_0, \gamma_1, \cdots, \gamma_{g-1})$. We will denote by $(\eta_0, \eta_1, \cdots, \eta_{g-1})$ the divisor of its zeroes.

By the general formula, we have:

$$\psi_i(t,P) = d_i(t)e^{[(q_i(t)-q_1(0))\int d\Omega_1 + t\int d\Omega_2]}\frac{\theta(\mathcal{A}(P) - U_1(q_i(t) - q_1(0)) - U_2 t - \zeta)}{\theta(\mathcal{A}(P) - \zeta)}$$

where Ω_1 and Ω_2 are normalized second kind abelian differentials with singularities $d\lambda^{-1}$ and $dm_j(\lambda)$ at the points P_j above $\lambda = 0$ respectively. Note that the differentials are independent of the index i of the component considered, and so are the vectors U_1 and U_2 of their b–periods. The functions $d_i(t)$ are arbitrary normalizations.

The divisor of the poles and zeroes satisfy according to Riemann's theorem:

$$\begin{aligned} \mathcal{A}(\eta_0) + \sum_{i=1}^{g-1} \mathcal{A}(\eta_i) - U_1(q_i(t) - q_1(0)) - U_2 t - \zeta &= -\mathcal{K} \\ \mathcal{A}(\gamma_0) + \sum_{i=1}^{g-1} \mathcal{A}(\gamma_i) - \zeta &= -\mathcal{K} \end{aligned}$$

Let us assume that the zero η_0 coincide with the pole γ_0. We then get

$$U_1 q_i(t) + U_2 t + V = -\mathcal{K} + \sum_{i=1}^{g-1} \mathcal{A}(\eta_i)$$

where we have denoted $V = -\mathcal{K} - U_1 q_1(0) + \sum_{i=1}^{g-1} \mathcal{A}(\gamma_i)$, a vector in $\mathrm{Jac}\,(\Gamma)$ *independent* of time. Now the expression in the right hand side of this equation belongs to the zero divisor of the θ function (see Chapter [15]), hence:

$$\theta(U_1 q_i(t) + U_2 t + V) = 0 \tag{6.49}$$

This is the constraint we were looking for. Note that all $q_i(t)$ satisfy the *same* equation which describes (for all values of time) the intersection of a straight line in the jacobian torus with the theta divisor. The slopes U_1 and U_2 are special enough to ensure that we have the right number of solutions.

To proceed to the reconstruction of the spin variables, it is enough to compute the matrix C of eigenvectors of f since we have $f = C\nu C^{-1}$. As we have shown the columns of C are the vectors $C^{(j)}$ with components given by the limit when $\lambda \to 0$ and $P = (\lambda, \mu_j(\lambda))$:

$$C_i^{(j)}(t) = \lim_{\lambda \to 0} d_i(t) e^{\left[(q_i(t)-q_1(0))\left(\int^P \Omega_1 - \frac{1}{\lambda}\right) + t\left(\int^P \Omega_2 - m_j(\lambda)\right)\right]} \frac{\theta(\mathcal{A}(P) - U_1(q_i(t) - q_1(0)) - U_2 t - \zeta)}{\theta(\mathcal{A}(P) - \zeta)}$$

The limits $\alpha_j = \lim_{\lambda \to 0}(\int^P \Omega_1 - \frac{1}{\lambda})$ and $\beta_j = \lim_{\lambda \to 0}(\int^P \Omega_2 - m_j(\lambda))$ are well–defined and depend on the point P_j. Note that if $C_i^{(j)} \to d_i C_i^{(j)}$ we have $f_{ij} \to d_i f_{ij} d_j^{-1}$ and we know that we must quotient by this toral action. Moreover if we change tha normalization of the vector $C(j)$, i.e. $C_i^{(j)} \to d'_j C_i^{(j)}$ the matrix F is invariant. Hence we can drop all factors that can be absorbed in these invariances, and we can choose:

$$C_i^{(j)}(t) = e^{\alpha_j q_i(t)}\, \theta(U_1 q_i(t) + U_2 t + V_j)$$

where we have introduced the constant vector $V_j = \zeta - U_1 q_1(0) - \mathcal{A}(P_j)$.

6.9 Symplectic structure.

To compute the symplectic form we need to consider the inverse $\widehat{\Psi}^{-1}(\lambda)$ of the matrix $\widehat{\Psi}(\lambda)$ whose columns are the eigenvectors at the N points above λ. The matrix $\widehat{\Psi}^{-1}(\lambda)$ is constructed by considering the adjoint system $\Psi^+(P)(L(\lambda) - \mu) = 0$. The line vector $\Psi^+(P)$ is reconstructed from its analytical properties, exactly like $\Psi(P)$. The lines of the matrix $\widehat{\Psi}^{-1}(\lambda)$ are the values of the line vector:

$$\Psi^{(-1)}(P) \equiv \frac{\Psi^+(P)}{< \Psi^+(P)\Psi(P) >}$$

at the N points P_j above λ. As before $< \Psi^+(P)\Psi(P) >$ is a meromorphic function on Γ with zeroes at the branch points of the covering $\mu \to \lambda$. Remark that this function is regular at $\lambda = 0$, in particular the essential singularities cancel. It follows that $\Psi^{(-1)}(P)$ has poles at the branch points, and zeroes at the divisor of poles of Ψ.

We remarked already several times that the components of the eigenvector can be written as ratios of suitable minors of the matrix $L(\lambda) - \mu$. In particular, the poles of $\Psi(P)$ are obtained as the zeroes of the first minor. Note that in this minor, the variables q_1, p_1 have disappeared. This corresponds to a reduction of the system by translational symmetry, which leaves a phase space of dimension $2(g-1)$, hence the $g-1$ poles of $\Psi(P)$. In particular one can choose the origin such that at initial time $q_1(0) = 0$.

Proposition. *The symplectic for of the spin Calogero model is*

$$\omega = \sum_{i=1}^{g-1} d\lambda_i \wedge d\mu_i = \sum_{i=1}^{N} dp_i \wedge dq_i - \omega_K \tag{6.50}$$

where ω_K is the Kirillov symplectic form on the orbit defined by $F = C\nu C^{-1}$, i.e. $\omega_K = \mathrm{Tr}\left(\nu C^{-1}\delta C \wedge C^{-1}\delta C\right)$.

Proof. As usual, we consider the form $K = K_1 + K_2 + K_3$ with:

$$\begin{aligned} K_1 &= \; < \Psi^{(-1)}(P)\delta L(\lambda) \wedge \delta\Psi(P) > \; d\lambda \\ K_2 &= \; < \Psi^{(-1)}(P)\delta\mu \wedge \delta\Psi(P) > \; d\lambda \\ K_3 &= \; \delta\left(\log \partial_\mu R\right) \wedge \delta\mu \, d\lambda \end{aligned}$$

and write that the sum of its residues vanishes to prove the equality of the two expressions for ω. The poles of K are the poles of $\Psi(P)$, $\{\gamma_i = (\lambda_i, \mu_i),\; i = 1\cdots, g-1\}$, the branch points, and the points P_j above $\lambda = 0$. At the points γ_i, the computation is exactly similar as in Chapter [4] yielding a sum of residues $2\sum_{i=1}^{g-1} \delta\mu_i \wedge \delta\lambda_i$ coming from equal contributions of K_1 and K_2. At the branch points the residues from K_1, K_2, K_3 cancel as in the rational case. The new features occur above $\lambda = 0$. It is easy to see that K_3 is regular above $\lambda = 0$ because $\delta\mu$ is *regular* at these points, since the coefficients of the polar part of μ_j are non–dynamical. The sum of residues of K_1 and K_2 at the P_j, can be written in matrix form as the residue of forms on the base as follows:

$$\begin{aligned} E_1 \equiv \sum_j \mathrm{Res}_{P_j} K_1 &= \mathrm{Res}_{\lambda=0} \mathrm{Tr}\left(\widehat{\Psi}^{-1}\delta L(\lambda) \wedge \delta\widehat{\Psi}\right) d\lambda \\ E_2 \equiv \sum_j \mathrm{Res}_{P_j} K_2 &= -\mathrm{Res}_{\lambda=0} \mathrm{Tr}\left(\delta\widehat{\Psi} \wedge \delta\widehat{\mu}\widehat{\Psi}^{-1}\right) d\lambda \end{aligned}$$

where $\widehat{\mu}$ is the diagonal matrix of the $\mu_j(\lambda)$. Since this is a local computation we first extract the essential singularities by setting $\widehat{\Psi} = G\widetilde{\Psi}$ and $L = G\widetilde{L}G^{-1}$ with G given in eq.(6.35). We get

$$\begin{aligned} E_1 &= \mathrm{Res}_{\lambda=0} \mathrm{Tr}\left([G^{-1}\delta G, \widetilde{L}] + \delta\widetilde{L}\right) \wedge \left(G^{-1}\delta G + \delta\widetilde{\Psi}\widetilde{\Psi}^{-1}\right) d\lambda \\ E_2 &= -\mathrm{Res}_{\lambda=0} \mathrm{Tr}\left(G^{-1}\delta G \wedge \widetilde{\Psi}\delta\widehat{\mu}\widetilde{\Psi}^{-1} + \delta\widetilde{\Psi} \wedge \delta\widehat{\mu}\widetilde{\Psi}^{-1}\right) d\lambda \end{aligned}$$

Note that $\mathrm{Tr}\left([G^{-1}\delta G, \widetilde{L}] \wedge G^{-1}\delta G\right) = 0$ since G is diagonal, and $\mathrm{Res}_{\lambda=0}\delta\widetilde{\Psi} \wedge \delta\widehat{\mu}\widetilde{\Psi}^{-1} = 0$ since $\delta\widehat{\mu}$ is regular at $\lambda = 0$. Collecting the remaining terms we have:

$$E_1 + E_2 = \mathrm{Res}_{\lambda=0} \mathrm{Tr}\left(\delta\widetilde{L} \wedge G^{-1}\delta G + \delta\widetilde{L} \wedge \delta\widetilde{\Psi}\widetilde{\Psi}^{-1} + [G^{-1}\delta G, \widetilde{L}] \wedge \delta\widetilde{\Psi}\widetilde{\Psi}^{-1} - G^{-1}\delta G \wedge \widetilde{\Psi}\delta\widehat{\mu}\widetilde{\Psi}^{-1}\right) d\lambda$$

Using $\widetilde{L}\widetilde{\Psi} = \widetilde{\Psi}\widehat{\mu}$, hence $\widetilde{\Psi}^{-1}\widetilde{L} = \widehat{\mu}\widetilde{\Psi}^{-1}$ we get

$$\mathrm{Tr}\left([G^{-1}\delta G, \widetilde{L}] \wedge \delta\widetilde{\Psi}\widetilde{\Psi}^{-1}\right) = \mathrm{Tr}\left(G^{-1}\delta G \wedge (\widetilde{L}\delta\widetilde{\Psi} - \delta\widetilde{\Psi}\widehat{\mu})\widetilde{\Psi}^{-1}\right) = \mathrm{Tr}\left(G^{-1}\delta G \wedge (\widetilde{\Psi}\delta\widehat{\mu} - \delta\widetilde{L}\widetilde{\Psi})\widetilde{\Psi}^{-1}\right)$$

where in the last step we have used $\widetilde{L}\delta\widetilde{\Psi} - \delta\widetilde{\Psi}\widehat{\mu} = \widetilde{\Psi}\delta\widehat{\mu} - \delta\widetilde{L}\widetilde{\Psi}$. Hence our expression simplifies to:

$$E_1 + E_2 = \mathrm{Res}_{\lambda=0}\mathrm{Tr}\left(2\delta\widetilde{L}\wedge G^{-1}\delta G + \delta\widetilde{L}\wedge\delta\widetilde{\Psi}\widetilde{\Psi}^{-1}\right)d\lambda$$

Note that $(G^{-1}\delta G)_{ij} = \lambda^{-1}\delta q_i\delta_{ij} + O(1)$ so that the first term contributes a residue $2\sum_i \delta p_i\wedge\delta q_i$. For the second term we note that $\widetilde{\Psi} = C + O(\lambda)$ (in the reduced system where $q_1(0) = 0$), and $\widetilde{L} = -\lambda^{-1}(F - \alpha) + O(1)$ hence the residue is $-\mathrm{Tr}\left(\delta F\wedge\delta CC^{-1}\right)$ since $\delta\alpha = 0$. Remembering that $F = C\nu C^{-1}$ this reads $-2\mathrm{Tr}\left(\nu C^{-1}\delta C\wedge C^{-1}\delta C\right)$. We recognize the Kirillov form on the coadjoint orbit of ν, see Chapter [2]. Writing that the sum of residues vanishes proves the proposition. ∎

6.10 Poles systems and double–Bloch condition.

In this section we present a natural construction of the Lax pair for the spin Calogero model, by relating it to the matrix KP equation. Before delving into this particular example it is illuminating to present the general context of this sort of connexion.

Let us consider a linear differential operator $\mathcal{D}$ in two variables x and t (there could be more than one time) and consider the differential equation $\mathcal{D}\Psi(x,t) = 0$. For example the KP operator reads $\mathcal{D} = \partial_t - \partial_x^2 + u(x,t)$. Assume moreover that the coefficients of $\mathcal{D}$ are doubly periodic meromorphic functions of the complex variable x with periods $2\omega_i$, $i = 1, 2$. We require nothing concerning the t–dependence. In general we know that for generic simply periodic potentials we can find a solution of the equation $\mathcal{D}\Psi(x,t) = 0$ which is quasi–periodic, i.e. $\Psi(x + 2\omega_1, t) = B_1\Psi(x,t)$. Such solutions are called Floquet or Bloch solutions.

In the case of elliptic potentials, since we have a second period, it is natural to require that Ψ be double Bloch, i.e.

$$\Psi(x + 2\omega_i, t) = B_i\Psi(x,t), \quad i = 1, 2$$

In contrast to the case of the one–period situation it turns out that this is in general *impossible*. Nevertheless one can find double Bloch solutions for very special potentials (from which the Calogero model will automagically spring out).

To understand the restrictions coming from the double Bloch condition let us assume that the function $x \to \Psi(x,t)$ has N poles at positions q_i on the torus with periods $2\omega_i$. Applying the Riemann–Roch theorem with $g = 1$ one sees that such functions form a vector space of dimension N. Indeed for any two such functions Ψ_1, Ψ_2 with the same Bloch multipliers, the quotient Ψ_2/Ψ_1 is a meromorphic function on the torus with N poles (since Ψ_1 has the same number of zeroes and poles, because Ψ_1'/Ψ_1 is elliptic), hence lives in a space of dimension $N - g + 1 = N$. The existence of such functions comes from their explicit construction using Lamé functions. Take any sum of the form:

$$\Psi(x,t,z) = \sum_{i=1}^{N} c_i(t,z)\Phi(x - q_i, z)e^{kx}$$

where $\Phi(q,z)$ is the Lamé function defined in eq.(6.13). Recall that we have $\Phi(x + 2\omega_i, z) = T_i(z)\Phi(x,z)$, with $T_i(z) = \exp(2\omega_i\zeta(z) - 2\eta_i z)$ (see Chapter [15]), so that $\Psi(x + 2\omega_i, t, z) = B_i\Psi(x,t,z)$ with $B_i = T_i(z)\exp(2k\omega_i)$. Given the two Bloch multipliers B_i we can adjust k and z to achieve this equality. So we have found an explicit form of the basis of the N–dimensional vector space of double Bloch functions. If we now require that this function Ψ obeys the equation $\mathcal{D}\Psi = 0$ we impose in fact more than N conditions. This is because $\mathcal{D}\Psi$ is a double Bloch function with the same multipliers, but differentiations and multiplication by potentials can only increase the degree of its divisor of poles. Hence $\mathcal{D}\Psi$ lives in a space of dimension greater than N, and its vanishing requires more than N linear conditions on the N coefficients c_i. This means that for a general operator $\mathcal{D}$ with elliptic coefficients there are no double Bloch solutions.

On the other hand, given a Riemann surface Γ of genus g, one can construct Baker–Akhiezer functions Ψ on it. A generic Baker–Akhiezer function depends on many parameters. It is defined first by the choice of the Riemann surface Γ, which depends on $3g-3$ moduli, i.e. depends on $3g-3$ complex parameters, and second by the choice of punctures P_α, local parameters z_α around P_α, and singular parts of order n_α at P_α. Only the first n_α coefficients in the expansion of z_α are relevant to the definition of the singular part, hence this produces a total of $(3g-3)+\sum_\alpha(1+n_\alpha)$ parameters. It is well known that such Baker–Akhiezer functions satisfy differential equations of the form $\mathcal{D}\Psi = 0$ (see an example below). In the case of one puncture the Baker–Akhiezer function is of the generic form:

$$\Psi(P,t) = e^{\sum_i t_i \int^P \Omega^{(i)}} \frac{\theta(\mathcal{A}(P) + \sum_i U^{(i)} t_i + \zeta)}{\theta(\mathcal{A}(P)+\zeta)}$$

We assume that x is t_1, but we could take for x any combination $x = \sum_i \alpha_i t_i$ of the elementary times t_i of the hierarchy. The condition for such a function Ψ to be double Bloch is that $2\omega_i U_1$ belongs to the lattice of periods of $\mathrm{Jac}\,(\Gamma)$. This means $2g$ conditions on the parameters of the Baker–Akhiezer functions. The dimension of the parameter space of Baker–Akhiezer functions is large enough to accomodate the $2g$ double Bloch conditions. This provides families of differential operators possessing double Bloch solutions. In this case the over–determined linear system on the coefficients $c_i(t,z)$ become compatible. The compatibility conditions may take the form of a Lax equation.

Let us apply this strategy to a simple example: we consider a smooth Riemann surface of genus g and l punctures P_β, $\beta = 1,\cdots,l$ with local parameters $w_\beta(P)$ around P_β ($w_\beta(P_\beta)=0$). Fix a divisor of degree $g+l-1$ in general position. Then there exists a unique Baker–Akhiezer function ψ_α having poles at this divisor and behaving in the neighborhood of each P_β as:

$$\psi_\alpha(x,t,P) = e^{w_\beta^{-1}x + w_\beta^{-2}t}\left(\delta_{\alpha\beta} + \sum_{s=1}^{\infty} \xi_s^{\alpha\beta}(x,t) w_\beta^s\right)$$

Indeed since the degree of the divisor of poles is $g+l-1$ we have vector space of functions of dimension l and we impose a system of l linear inhomogeneous normalization conditions.

Proposition. *Let $|\Psi(x,t,P)\rangle$ be the Dirac notation for the vector with components $\psi_\alpha(x,t,P)$. Then it satisfies the equation:*

$$(\partial_t - \partial_x^2 + u(x,t))|\Psi(x,t,P)\rangle = 0) \tag{6.51}$$

where the $l\times l$ matrix u is given by $u^{\alpha\beta}(x,t) = 2\partial_x \xi_1^{\alpha\beta}(x,t)$. Such potentials are called finite–zone potentials.

<u>Proof</u>. In the vicinity of each puncture P_β one can write:

$$(\partial_t - \partial_x^2)\psi_\alpha(x,t,P) = e^{w_\beta^{-1}x + w_\beta^{-2}t}(-2\partial_x \xi_1^{\alpha\beta} + O(w_\beta))$$

Since the left–hand side is meromorphic except at the P_j and has the same $g+l-1$ poles as Ψ, and has an appropriate essential singularity at each puncture, it can be expanded on the ψ_β so that one can write $(\partial_t - \partial_x^2)\psi_\alpha = -\sum_\beta u^{\alpha\beta}\psi_\beta$, for some $u^{\alpha\beta}(x,t)$ independent of $P\in\Gamma$, meromorphic with respect to x. Comparing to the right–hand side around P_β we find $u^{\alpha\beta} = 2\partial_x \xi_1^{\alpha\beta}$. ∎

We now express the condition that the potential u is elliptic and obtain its precise form.

Proposition. *If the finite–zone potential u is elliptic, it has the form:*

$$u(x,t) = \sum_{i=1}^{N} \rho_i(t)\wp(x - q_i(t))$$

where $\rho_i(t)$ is a $l \times l$ matrix of rank one.

Proof. We need the explicit form of $|\Psi\rangle$ as a Baker–Akhiezer function on some curve Γ. Let $P_1, \cdots, P_l$ be the punctures and $\gamma_1, \cdots, \gamma_{g+l-1}$ be the poles of the Baker–Akhiezer function. There exists a unique meromorphic function $h_\alpha(P)$ with poles at the γ_i and such that $h_\alpha(P_\beta) = \delta_{\alpha\beta}$. In particular it $l-1$ zeroes at the P_β, $\beta \neq \alpha$, and g other zeroes at a divisor D_α. Applying the Abel theorem it follows that $Z_0 + \mathcal{K} \equiv \mathcal{A}(P_\alpha) - \mathcal{A}(D_\alpha) = \sum_\beta \mathcal{A}(P_\beta) - \sum \mathcal{A}(\gamma_i)$ is independent of α. The Baker–Akhiezer function $\psi_\alpha(P)$ reads:

$$\psi_\alpha(P) = h_\alpha(P)\frac{\theta(\mathcal{A}(P) + U_1x + U_2t + Z_0 - \mathcal{A}(P_\alpha))\theta(Z_0)}{\theta(\mathcal{A}(P) + Z_0 - \mathcal{A}(P_\alpha))\theta(U_1x + U_2t + Z_0)}e^{x\int^P \Omega_1 + t\int^P \Omega_2}$$

In this formula the first theta function in the denominator cancels the extra g zeroes of $h_\alpha(P)$ at D_α. Then the first theta function in the numerator is obtained by requiring that there is no monodromy. Finally the two other theta functions are necessary to ensure the correct normalization of $\psi_\alpha(P)$ at P_α. It is now obvious to identify the poles in x of $|\Psi(x,t,P)\rangle$. They occur when:

$$\theta(U_1x + U_2t + Z_0) = 0$$

We recover in a very short way eq.(6.49), and this defines the functions $q_i(t)$ (but we still don't know precisely the curve Γ). Consider the residue of $\psi_\alpha(x,t,P)$ when $x = q_i(t)$. As a function of P it is a meromorphic function except at the punctures, and has the set of poles γ_i. Moreover in the neighborhood of each puncture P_β it has the form $\exp(w_\beta^{-1}q_i(t) + w_\beta^{-2}t)\,O(w_\beta)$, i.e. the coefficient of the exponential *vanishes*. This is because at P_β, $\beta \neq \alpha$ the function h_α vanishes, while at P_α the theta function $\theta(\mathcal{A}(P) + U_1q_i(t) + U_2t + Z_0 - \mathcal{A}(P_\alpha))$ vanishes, due to the definition of $q_i(t)$. In general such a Baker–Akhiezer function vanishes identically, however for the special values of the parameters $(x = q_i(t), t)$ it exists, and is a fortiori unique, up to a constant normalization. We call it $\sigma_i(t,P)$. This means that:

$$\psi_\alpha(x,t,P) = \frac{a_{i\alpha}(t)\sigma_i(t,P)}{x - q_i(t)} + O(1) \tag{6.52}$$

The potential is given by computing the expansion of $\psi_\alpha(x,t,P)$ around P_β. Using $\sigma_i(t,P) = (w_\beta b_i^\beta + O(w_\beta^2))\exp(w_\beta^{-1}q_i(t) + w_\beta^{-2}t)$ we find around $x = q_i$ and $P = P_\beta$:

$$\xi_1^{\alpha\beta}(x,t) = \frac{a_{i\alpha}(t)b_i^\beta(t)}{x - q_i(t)} + O(1)$$

Finally $u^{\alpha\beta} = 2\partial_x\xi_1^{\alpha\beta}$ has double poles at $x = q_i$ with coefficient $\rho_i^{\alpha\beta} = -2a_{i\alpha}(t)b_i^\beta(t)$. We impose the condition that u is elliptic. The double pole gives rise to a Weierstrass function $\wp(x - q_i(t))$ with a coefficient ρ_i of rank one. ■

Using again the Dirac notation $|a_i\rangle$ for the vector a_i^α and $\langle b_i|$ for the covector b_i^β, the rank–one matrix ρ_i reads $\rho_i = |a_i\rangle\langle b_i|$.

We now require that $|\Psi\rangle$ is double Bloch in x, hence can also be written as:

$$|\Psi\rangle = \sum_{i=1}^{N} |s_i(t,\lambda,\mu)\rangle\Phi(x - q_i(t),\lambda)e^{-\frac{\mu}{2}x + \frac{\mu^2}{4}} \tag{6.53}$$

In this formula λ and μ are free parameters, but the double Bloch condition will turn out to specify the Riemann surface on which (λ,μ) are coordinates. Moreover μ will become infinite at the punctures, and λ will be the local parameter required to define the essential singularity of the Baker–Akhiezer function. Finally $|s_i\rangle$ is, in view of eq.(6.52), of the form $\psi_i(t,\lambda,\mu)|a_i(t)\rangle$.

Proposition. *The equation*

$$\left(\partial_t - \partial_x^2 + \sum_{i=1}^{N} \rho_i(t)\wp(x - q_i(t))\right)|\Psi\rangle = 0 \tag{6.54}$$

has solutions $|\Psi\rangle$ of the form

$$|\Psi\rangle = \sum_{i=1}^{N} \psi_i(t,\lambda,\mu)\Phi(x - q_i(t),\lambda)e^{-\frac{\mu}{2}x+\frac{\mu^2}{4}t}|a_i(t)\rangle \tag{6.55}$$

if and only if $q_i(t)$ and the quantities $f_{ij} = \langle b_i|a_j\rangle$ satisfy the equations eqs.(6.4, 6.5) and the constraints $f_{ii} = 2$.

Proof. Inserting equation (6.55) into equation (6.54) we find the condition:

$$\begin{aligned}\mathcal{E} \quad &\equiv \quad \sum_{i=1}^{N} \Bigg\{ \Phi(x - q_i,\lambda)\frac{d(\psi_i|a_i\rangle)}{dt} - (\dot q_i - \mu)\Phi'(x - q_i,\lambda)\psi_i|a_i\rangle - \Phi''(x - q_i,\lambda)\psi_i|a_i\rangle \\ &+ \quad \sum_{j=1}^{N} f_{ji}\wp(x - q_j)\Phi(x - q_i,\lambda)\psi_i|a_j\rangle \Bigg\} = 0\end{aligned}$$

where $\Phi' = \partial_x\Phi$ and so on. The vanishing of the triple pole $(x - q_i)^{-3}$ gives the condition:

$$\langle b_i|a_i\rangle|a_i\rangle = 2|a_i\rangle \tag{6.56}$$

Using this condition and the Lamé equation (6.14), we can identify the double pole $(x - q_i)^{-2}$. Its vanishing gives the condition:

$$(\dot q_i - \mu)\psi_i|a_i\rangle + \sum_{j\neq i} f_{ij}\Phi(q_i - q_j,\lambda)\psi_j|a_i\rangle = 0 \tag{6.57}$$

We finally identify the residue of the simple pole and obtain the condition:

$$\left(\frac{d}{dt} - \wp(\lambda)\right)\psi_i|a_i\rangle + \sum_{j\neq i} f_{ji}\wp(q_i - q_j)\psi_i|a_j\rangle + \sum_{j\neq i} f_{ij}\Phi'(q_i - q_j,\lambda)\psi_j|a_i\rangle = 0. \tag{6.58}$$

Inserting back eqs.(6.56,6.57,6.58) into the expression of $\mathcal{E}$ one sees that $\mathcal{E}$ vanishes identically due to the functional equation (6.15). Hence the vector function $|\Psi\rangle$ given by eq.(6.55) satisfies eq.(6.54) if and only if the conditions (6.56,6.57,6.58) are fulfilled.

From eq.(6.56) it follows that the constraints $f_{ii} = 2$ should hold. Eq.(6.57) can then be rewritten as a matrix equation for the N–dimensional vector $\Psi = (\psi_i)$ (not to be confused with the l–dimensional object $|\Psi\rangle$)

$$(L(t,\lambda) - \mu I)\Psi = 0 \tag{6.59}$$

where the matrix $L(t,\lambda)$ is given by:

$$L_{ij}(t,\lambda) = \dot q_i\delta_{ij} + (1 - \delta_{ij})f_{ij}\Phi(q_i - q_j,\lambda) \tag{6.60}$$

We recognise the Lax matrix of the Calogero model, eq.(6.11), so that the N–vector Ψ identifies with the eigenvector considered in section (6.6).

We can rewrite equation (6.58) as:

$$|\dot a_i\rangle = -\Lambda_i|a_i\rangle - \sum_{j\neq i} f_{ji}\wp(q_i - q_j)|a_j\rangle \tag{6.61}$$

where we have defined:

$$\Lambda_i = \frac{\dot\psi_i}{\psi_i} - \wp(\lambda) + \sum_{j\neq i} f_{ij}\Phi'(q_i - q_j, \lambda)\frac{\psi_j}{\psi_i}$$

But this last equation can be rewritten:

$$(\partial_t - \wp(\lambda)I - \Lambda + M)\Psi = 0 \tag{6.62}$$

where $\Lambda = \mathrm{Diag}\,(\Lambda_i)$ and the matrix $M(t,\lambda)$ is given by:

$$M_{ij}(t,\lambda) = (1-\delta_{ij})f_{ij}\Phi'(q_i - q_j, \lambda). \tag{6.63}$$

We recognize the second matrix M of the Lax pair (6.12). The compatibility conditions of eq.(6.59,6.62) reads $\dot L = [L, M - \Lambda - \wp(\lambda)I]$. Of course the term $\wp(\lambda)I$ does not contribute to the commutator. Moreover we can get rid of the diagonal matrix Λ by performing a conjugation by a diagonal matrix on L, which amounts to quotient out the toral action. In this way we have exactly recovered the Lax pair of the Calogero model, eq.(6.10), hence the q_i and f_{ij} have to satisfy the Calogero equations of motion, in order for $|\Psi\rangle$ to be double Bloch. In the course of the proof λ and μ have been identified as coordinates on the spectral curve of the Calogero model, i.e. are related by the equation $\det\,(L(\lambda) - \mu) = 0$. One can see that the punctures P_β are l amongst the N points above $\lambda = 0$, and λ is a local parameter around each of them. ∎

The outcome of this analysis is that the double Bloch condition singles out very specific finite–zone potentials. It amounts to an overdetermined linear system on the coefficients of the expansion of $|\Psi\rangle$ which is equivalent to the Lax equation. In particular we have obtained in a simple and natural way the Lax matrices of the Calogero model in eq.(6.60,6.63). The method is clearly general and lends itself to extensions by Baker–Akhiezer functions with different patterns of essential singularities. Note that in our construction we have considered only two "times", x and t wich parametrize the singularity at each puncture. This provides a whole variety of integrable systems with spectral parameter lying on a genus one curve. In view of the counting of parameters in Lax equations in Chapter [2] this is a noticeable fact. Even more remarkable is a construction, due to Hitchin, which produces integrable systems with spectral parameter lying on an arbitrary genus curve. The Calogero model can also be seen as a particular case of this construction.

6.11 Hitchin systems.

Let Σ be a Riemann surface of genus g, and let G be a complex semi–simple Lie group. Let $\mathcal{A}$ be the space of type $(0,1)$ fields on Σ, i.e. of the form $A = A_{\overline{z}}(z,\overline{z})d\overline{z}$ for any local coordinate system z, with values in the Lie algebra of G. We define the "gauge group" $\mathcal{G}$ to be the space of maps from Σ to G, i.e. $h \in \mathcal{G}$, is a function $h(z,\overline{z})$ with values in G. The gauge group acts on $\mathcal{A}$ as follows:

$$A \longrightarrow A^h \equiv h^{-1}Ah + h^{-1}\overline{\partial}h \tag{6.64}$$

Note that the differences $A - A'$ form a vector space compatible with the gauge group action, hence $\mathcal{A}$ can be seen as an affine space with group action. We call $\mathcal{N} = \mathcal{A}/\mathcal{G}$ the orbit space of $\mathcal{A}$ under $\mathcal{G}$.

A tangent vector at the point $A \in \mathcal{A}$ is of the form $X = X_{\overline{z}}(z,\overline{z})d\overline{z}$ with values in the Lie algebra of G. A covector Φ at the point A is similarly of the form $\Phi = \Phi_z(z,\overline{z})dz$, the pairing between vectors and covectors at the point A being given by:

$$(\Phi, X) = \int_\Sigma \mathrm{Tr}\,(\Phi_z X_{\overline{z}})dz d\overline{z}$$

The gauge group acts on vectors and covectors by adjoint action, $X^h = h^{-1}Xh$, $\Phi^h = h^{-1}\Phi h$ and this leaves the pairing invariant.

The starting point of the construction is the cotangent bundle $T^*\mathcal{A}$ whose points are pairs (A, Φ) where Φ is a cotangent vector at $A \in \mathcal{A}$. The canonical symplectic form on this space reads:

$$\omega = \int_\Sigma \mathrm{Tr}\,(\delta\Phi_z \wedge \delta A_{\overline{z}}) dz d\overline{z}$$

Remark that this symplectic form is invariant under the gauge group action, so we can perform a Hamiltonian reduction by this group. To do that we need to compute the moment μ of this action. In the case of a cotangent bundle it is shown in Chapter [14] that the Hamiltonian generating the infinitesimal group action is given by $H_\epsilon \equiv (\mu, \epsilon) = \int \mathrm{Tr}\,(\mu_{z\overline{z}}\epsilon) dz d\overline{z} = \alpha(X_\epsilon(A, \Phi))$ where α is the canonical one–form $\alpha = \int_\Sigma \mathrm{Tr}\,(\Phi \delta A)$ and $X_\epsilon(A, \Phi)$ is the infinitesimal variation of the point (A, Φ) under gauge group action, namely:

$$X_\epsilon A = \overline{\partial}_A \epsilon \equiv \overline{\partial}\epsilon + [A, \epsilon], \quad X_\epsilon \Phi = [\Phi, \epsilon]$$

One gets after an integration by parts:

$$\mu_{z\overline{z}} dz d\overline{z} = -\left(\frac{\partial \Phi_z}{\partial \overline{z}} + [A_{\overline{z}}, \Phi_z]\right) dz d\overline{z}, \quad \text{or} \quad \mu = \overline{\partial}_A \Phi \equiv \overline{\partial}\Phi + A\Phi + \Phi A$$

The phase space $\mathcal{P}$ of the Hitchin system is obtained by choosing the moment equal to 0. The stability group of this moment is therefore the whole gauge group $\mathcal{G}$ so that we have:

$$\mathcal{P} = \mu^{-1}(0)/\mathcal{G}$$

Choosing $\mu = 0$ means that $\overline{\partial}_A \Phi = 0$, and this has a nice geometric interpretation. A cotangent vector at a point $n \in \mathcal{N}$, which is the class of $A \in \mathcal{A}$ under the $\mathcal{G}$ action, may be viewed as a linear form on $T^*_A \mathcal{A}$ vanishing on vectors tangent to the fiber, that is, such that $(\Phi, \overline{\partial}_A \epsilon) = 0$ for all ϵ. This condition is equivalent to $\overline{\partial}_A \Phi = 0$. This interpretation being covariant under the gauge group action, it follows that $\mathcal{P} = T^*\mathcal{N}$ where $\mathcal{N}$ is the orbit space (avoiding non generic orbits to obtain a good manifold). By a theorem of Narasimhan and Seshadri, the space $\mathcal{N}$ is known to be isomorphic to the moduli space of (stable) holomorphic G–bundles on Σ, and this shows that it is finite–dimensional.

Theorem. *The phase space $\mathcal{P} = T^*\mathcal{N}$ is of finite dimension: for $g > 1$ and G a semi–simple Lie group, we have*

$$\dim \mathcal{P} = 2 \dim \mathcal{N} = 2\,(g-1) \dim G$$

Proof. Let us sketch some ideas of the proof. The first step is to relate $\mathcal{N}$ to holomorphic G–bundles. Given $A \in \mathcal{A}$ and a sufficiently fine covering U_α of Σ one solves, for some C^∞ functions $h_\alpha \in G$ defined in U_α, the equation

$$h_\alpha^{-1} \overline{\partial} h_\alpha = A_\alpha \tag{6.65}$$

where $A_\alpha \equiv A|_{U_\alpha}$. We define a principal G–bundle by the transition functions $g_{\alpha\beta} = h_\alpha h_\beta^{-1}$ on $U_\alpha \cap U_\beta$. The action of $\mathcal{G}$ on the h_α reads $h_\alpha^h = h_\alpha h$, hence $g_{\alpha\beta}$ is gauge invariant, so the G–bundle is really attached to a point of $\mathcal{N}$. This bundle is holomorphic, $\overline{\partial} g_{\alpha\beta} = 0$, because $g_{\alpha\beta}^{-1} \overline{\partial} g_{\alpha\beta} = h_\beta (A_\alpha - A_\beta) h_\beta^{-1} = 0$ since $A_\alpha = A_\beta$ on $U_\alpha \cap U_\beta$. Of course h_α is defined up to right multiplication by a holomorphic function f_α, and this yields an equivalent presentation of the same bundle, with transition functions $g'_{\alpha\beta} = f_\alpha g_{\alpha\beta} f_\beta^{-1}$.
Next we remark that the associated determinant bundle has vanishing Chern class. Viewing G as a group of matrices, define the determinant bundle as a line bundle whose transition functions are the $\det g_{\alpha\beta}$. By definition of h_α we have $\overline{\partial} \log \det h_\alpha = \mathrm{Tr}\, A_\alpha = 0$ because we assume G semi–simple, so $\det h_\alpha$ is in fact holomorphic. Then $\det g_{\alpha\beta} = \det h_\alpha / \det h_\beta$ defines a trivial holomorphic line bundle, and its Chern class

vanishes.
To compute the dimension of $\mathcal{N}$, one computes the dimension of the cotangent space of $\mathcal{N}$ at a point $n \in \mathcal{N}$. Take a representative element $A \in \mathcal{A}$ of n. The cotangent space at n identifies to the space of forms Φ of type $(1,0)$ satisfying $\overline{\partial}_A \Phi = 0$. Consider the G–bundle attached to n, defined using some choice of functions h_α as in eq.(6.65). Define on each U_α the forms $\widehat{\Phi}_\alpha = h_\alpha \Phi h_\alpha^{-1}$, with values in the Lie algebra of G. By construction we have $\overline{\partial}\widehat{\Phi}_\alpha = 0$ and $\widehat{\Phi}_\alpha = g_{\alpha\beta}\widehat{\Phi}_\beta g_{\alpha\beta}^{-1}$, i.e. the $\widehat{\Phi}_\alpha$ define a *global holomorphic* section with values in 1–forms on Σ, of the associated bundle $\mathrm{Ad}\, P$, the bundle with fibers the Lie algebra of G, and adjoint group action. This shows that the cotangent space to $\mathcal{N}$ at n identifies to $T_n^*\mathcal{N} = H^0(\Sigma, \kappa \otimes \mathrm{Ad}\, P_n)$, where P_n is the the principal G–bundle attached to n, and κ is the canonical bundle.
The Riemann–Roch theorem has an extension to vector bundles which reads in our case:

$$\dim H^0(\Sigma, \kappa \otimes \mathrm{Ad}\, P) - \dim H^0(\Sigma, \mathrm{Ad}\, P) = (g-1)\mathrm{dimG} - c(\det \mathrm{Ad}\, P)$$

where the determinant bundle $\det \mathrm{Ad}\, P$ has transition functions $\det g_{\alpha\beta}$, and so has vanishing Chern class $c(\det \mathrm{Ad}\, P) = 0$. We proceed to show that generically, the vector bundle $\mathrm{Ad}\, P$ has no global holomorphic section, i.e. $\dim H^0(\Sigma, \mathrm{Ad}\, P) = 0$. Indeed if $\{f_\alpha\}$ with patching condition $f_\alpha = g_{\alpha\beta} f_\beta g_{\alpha\beta}^{-1}$ defines such a section, for any integer m the quantities $\mathrm{Tr}\, f_\alpha^m$ define a global holomorphic function on Σ, hence a constant. This means that the eigenvalues of f_α are constants, independent of α, and one can write $f_\alpha = u_\alpha \Lambda u_\alpha^{-1}$ for some u_α holomorphic on U_α and a constant diagonal matrix Λ. The patching condition reads $[\Lambda, u_\beta^{-1} g_{\alpha\beta} u_\alpha] = 0$. This implies that the matrices $u_\beta^{-1} g_{\alpha\beta} u_\alpha$ are block diagonal (blocks correspond to coincident eigenvalues of Λ). But this would mean that the fiber bundle P is decomposable, since the transition functions $u_\beta^{-1} g_{\alpha\beta} u_\alpha$ provide an equivalent description of our bundle. We exclude this situation because it is not generic. Finally the dimension of $T_n^*\mathcal{N}$ is $(g-1)\mathrm{dimG}$, and $\dim T^*\mathcal{N} = 2\,(g-1)\dim G$. Of course in the case of line bundles, we have $\dim H^0(\Sigma, \mathrm{Ad}\, P) = 1$, since we are here considering global holomorphic *functions* and $\dim T_n^*\mathcal{N} = g$.
∎

It will be useful to have an heuristic picture of the moduli. First, it is known that all vector bundles on a non compact Riemann surface are trivial. This implies that one can describe all vector bundles on a compact Riemann surface by using a covering with only two open sets U_0 and U_∞, where U_0 is a small disk around a point, say $z = 0$, and U_∞ is the Riemann surface with the point $z = 0$ removed. The bundles are then described by giving only one transition function $g_{0\infty}$ defined on the annulus $U_0 - \{z = 0\}$. We get an equivalent bundle by changing the transition function

$$g_{0\infty}(z) \to g'_{0\infty}(z) = f_0(z) g_{0\infty}(z) f_\infty^{-1}(z)$$

where f_0 is analytic non–vanishing on U_0, and f_∞ is analytic non–vanishing on U_∞, i.e. on the whole of $\Sigma - \{z = 0\}$. The moduli space of vector bundles is the space of transition functions $g_{0\infty}$ modulo the above redefinitions.

In the case of a line bundle, we can write quite generally

$$g_{0\infty}(z) = z^k e^{\sum_{-\infty}^{\infty} a_n z^n} \tag{6.66}$$

where k is an integer because $g_{0\infty}(z)$ is single valued on the annulus. Consider now on U_∞ the function $f_\infty(z) = \exp\left(\sum_{n=g+1}^{\infty} a_{-n}\varphi_{-n}(z)\right)$, where $\varphi_{-n}(z)$ is a meromorphic function on Σ such that around $z = 0$, we have $\varphi_{-n}(z) = z^{-n} + O(z^{-g})$ and is regular everywhere else. Notice that such a function exists and is unique for $n \geq g+1$. Using this f_∞, we can get rid of all negative powers $z^{-n}, n \geq g+1$ in eq.(6.66). Then using $f_0(z)$, we can get rid of all positive powers $z^n, n \geq 0$ as well. Hence, we are left with

$$g_{0\infty}(z) = z^k e^{\sum_1^g a_{-n} z^{-n}} \tag{6.67}$$

So the line bundles on Σ are holomorphically classified by an integer k, the Chern class of the bundle, and g continuous moduli, which describe in fact the Picard variety of Σ (the Jacobian for $k = 0$).

In the case of higher rank vector bundles, things are much more complicated. If Σ is a sphere, we know by a theorem of Birkhoff (Riemann–Hilbert factorization) that we can always decompose a matrix $g_{0\infty}(z)$ as

$$g_{0\infty}(z) = f_0(z)\lambda(z)f_-(z)^{-1} \tag{6.68}$$

where $\lambda(z)$ is a diagonal matrix with diagonal elements z^{k_i}, for some integers k_i, f_0 has an expansion in z, and f_- has an expansion in $1/z$. This exactly means that vector bundles on the sphere are classified by the integers k_i and have no continuous moduli. In other words on the sphere $\dim \mathcal{N} = 0$. The integers k_i are holomorphic invariants of the bundle. The Chern class of the corresponding determinant bundle is $\sum k_i$.

In the case of a general Riemann surface Σ, we can still use the Birkhoff theorem in the small disk around $z = 0$ to write the transition function on the annulus in the form eq.(6.68). Note that $f_-(z)$ is here only defined on the annulus and cannot in general be extended to $\Sigma - \{0\}$. However one can hope that a similar mechanism as in the case of line bundles allows to get rid of powers z^{-n} for $n \geq g+1$, so that we can write the transition function as in eq.(6.67), but with k a diagonal matrix with integer entries k_i, and $a_{-1}, \cdots, a_{-g}$ matrices. If in addition all the integers k_i vanish, one can use the freedom of redefinition of the bundle by constant matrices f_0 and f_∞ to diagonalize a_{-1}. We can still quotient away by conjugating by constant diagonal matrices while preserving this form. If $g > 1$, it remains a space of $(g-1)\dim G$ parameters (in fact $\operatorname{rank} G + [(g-1)\dim G - \operatorname{rank} G]$) which can be plausibly taken as coordinates on an open set of the moduli space. If $g = 1$, however we are left with

$$g_{0\infty}(z) = e^{a_{-1}/z} \tag{6.69}$$

where a_{-1} is diagonal, so that we have $\operatorname{rank} G$ parameters, which is known to be the correct dimension of the moduli space in that case. We will content ourselves with this interpretation in the following.

The construction of Hitchin integrable systems on $\mathcal{P}$ will now be done by defining a very simple set of Poisson commuting functions on $T^*\mathcal{A}$, invariant under the gauge group, and reducing them to $T^*\mathcal{N}$.

Let $P(X)$ be an invariant polynomial on the Lie algebra of G, i.e. $P(h^{-1}Xh) = P(X)$. Recall that the ring of such invariant polynomials is freely generated (Chevalley's theorem) by homogeneous polynomials P_i, $i = 1, \cdots, \operatorname{rank} G$, of degrees m_i, the so–called exponents of the Lie algebra. These numbers are such that:

$$\sum_{i=1}^{\operatorname{rank} G} (2m_i - 1) = \dim G$$

For any given invariant polynomial P of degree m, i.e. $P(X) = \operatorname{Tr} X^m$, consider the function on phase space taking values in differentials of type $(m, 0)$ (i.e. of the form $\omega(z, \overline{z})dz^m$):

$$(A, \Phi) \in T^*\mathcal{A} \rightarrow P(\Phi)$$

This differential is holomorphic, since $\overline{\partial}P(\Phi) = m \operatorname{Tr}(\overline{\partial}\Phi\Phi^{m-1}) = m \operatorname{Tr}(\overline{\partial}_A\Phi\Phi^{m-1}) = 0$, where we have used cyclicity of trace. Introducing a basis of holomorphic differentials of type $(m, 0)$, say $\omega_j^{(m)}$ we can write

$$P(\Phi) = \sum_j H_{P,j}(\Phi)\, \omega_j^{(m)}$$

The functions $H_{P,j}$ on phase space are $\mathcal{G}$–invariant, and define the Hamiltonians of the Hitchin systems. Note that the basis of differentials $\omega_j^{(m)}$ do not contain dynamical variables, since the Riemann surface Σ is a fixed parameter of the construction.

Proposition. *The functions $H_{P_m,j}$ associated to the primitive polynomials P_m, $m = 1, \cdots, \mathrm{rank}\, G$, which generate the ring of invariant polynomials, are in involution. Their number is $(g-1)\dim G$, so they define an Hamiltonian integrable system on $\mathcal{P} = T^*\mathcal{N}$.*

Proof. The functions $H_{P_m,j}$ seen as functions on $T^*\mathcal{A}$ are in involution, i.e. $\{H_{P_m,j}, H_{P_n,k}\} = 0$ for any two invariant polynomials, because they depend only on the momenta Φ. Since the polynomials P are G–invariant, the functions $H_{P,j}$ are gauge invariant. They are thus well defined on the symplectic quotient and in involution there, because one can compute directly their reduced Poisson brackets, see Chapter[14]. It remains to show that the number of the Hamiltonians is half the dimension of the phase space. Let us count them. We need the number of holomorphic differentials of type $(m, 0)$. By the Riemann–Roch theorem

$$\dim H^0(\Sigma, \kappa^m) - \dim H^0(\Sigma, \kappa^{1-m}) = c(\kappa^m) + 1 - g$$

where $c(\kappa^m) = m(2g-2)$ and $\dim H^0(\Sigma, \kappa^{1-m}) = 0$ because $1 - m < 0$, so that $\dim H^0(\Sigma, \kappa^m) = (2m - 1)(g-1)$. The total number of independent Hamiltonians $H_{P_i,j}$ is therefore:

$$\sum_{i=1}^{\mathrm{rank}\, G} \dim H^0(\Sigma, \kappa^{m_i}) = \sum_{i=1}^{\mathrm{rank}\, G} (2m_i - 1)(g-1) = (g-1)\dim G$$

This is half the dimension of the phase space. This counting works for $g > 1$. For $g = 0$ there are no regular differentials on the sphere, so that $\dim H^0(\Sigma, \kappa^m) = 0$. The Hitchin construction in this case yields no interesting system. For $g = 1$ one has $\dim H^0(\Sigma, \kappa^m) = 1$, since this space is spanned by dz^m, so that we find $\sum_{i=1}^{\mathrm{rank}\, G} 1 = \mathrm{rank}\, G$ Hamiltonians. ∎

In genus 0 and 1 the Hitchin construction does not provide useful dynamical systems. This is a motivation to generalize it to Riemann surfaces with marked points. Let $z_k \in \Sigma$ be N points on the Riemann surface Σ. At each of these points we associate an element u_k in the dual of the Lie algebra of G, which we shall identify with the Lie algebra itself using the invariant bilinear form.

Instead of choosing $\mu = 0$ we now choose a momentum:

$$\mu(A, \Phi) \equiv \overline{\partial}_A \Phi = 2i\pi \sum_k u_k \delta_{z_k} \tag{6.70}$$

where δ_{z_k} is the Dirac measure at the point z_k, namely represented locally around z_k by $\delta(z - z_k) dz d\overline{z}$. The stability group of this momentum is the subgroup of gauge transformations leaving u_k invariant: $\mathcal{G}_{z;u} = \{g \in \mathcal{G} \text{ s.t. } h(z_k) \in G_k\}$, where G_k is the stabilizer of u_k. The reduced phase space is as usual:

$$\mathcal{P}_{z;u} \equiv \mu^{-1}(2i\pi \sum_k u_k\, \delta_{z_k}) / \mathcal{G}_{z;u}$$

In other words, we have:

$$\mathcal{P}_{z;u} \equiv \{(A, \Phi) | \overline{\partial}_A \Phi = 2i\pi \sum_k u_k\ \delta_{z_k}\} / \mathcal{G}_{z;u} \tag{6.71}$$

The equation $\overline{\partial}_A \Phi = 2i\pi \sum_k u_k\ \delta_{z_k}$ specifies the behavior of Φ around the marked points. Indeed, let us parametrize A locally around z_k as $A = g_k^{-1} \overline{\partial} g_k$ with $g_k(z_k) = 1$. Then the condition that (A, Φ) belongs to $\mathcal{P}_{z;u}$ translates locally into the condition $\overline{\partial}(g_k \Phi g_k^{-1}) = 2i\pi u_k \delta_{z_k}$ so that $g_k \Phi g_k^{-1}$ is holomorphic in an open set around z_k excluding z_k, and behaving locally as

$$g_k \Phi g_k^{-1} = \frac{u_k dz}{z - z_k} + O(1)$$

since $\overline{\partial} \frac{dz}{z - z_0} = 2i\pi \delta_{z_0}$.

The construction of the commuting Hamiltonians then works as above. For P_i an invariant polynomial of degree m_i on the Lie algebra of G, one considers the function

$$(A, \Phi) \in \mathcal{P}_{z;u} \to P_i(\Phi) \tag{6.72}$$

which take values in the space of meromorphic forms of type $(m_i, 0)$ with poles at points z_k of order at most m_i. These functions are in involution and define an integrable system.

6.12 Examples of Hitchin systems.

Example 1. Let us illustrate this construction by considering the Riemann sphere with N marked points $z_1, \cdots, z_N$, and fixed parameters u_k in the dual of the Lie algebra of G. We take A's of the form

$$A = h^{-1}\overline{\partial} h \tag{6.73}$$

with $h(z, \overline{z}) \in G$ globally defined on the sphere, so that the attached principal bundle is trivial. The condition $\overline{\partial}_A \Phi = 2i\pi \sum_k u_k \delta_{z_k}$ imposes that $h\Phi h^{-1}$ possesses a simple pole at $z = z_k$ with residue $h(z_k)u_k h^{-1}(z_k)$. Since there are no holomorphic differentials on the sphere, the only solution is:

$$h\Phi h^{-1} = \sum_k \frac{\widehat{u}_k \, dz}{z - z_k} \quad \text{with} \quad \widehat{u}_k = h(z_k) u_k h^{-1}(z_k) \tag{6.74}$$

Requiring that $h\Phi h^{-1}$ is regular at infinity yields:

$$\sum_k \widehat{u}_k = 0 \tag{6.75}$$

The data (A, Φ) are parametrized by h through the equations (6.73,6.74), up to left multiplication $h \to lh$ with l holomorphic, hence constant, and together with the constraint eq.(6.75). Notice that this constraint is invariant under $h \to lh$. Gauge transformations act on h as $h \to hg$, where g is such that $g(z_k) \in G_k$, the stability group of u_k. Note that $\widehat{\Phi}$ and $\widehat{u}_k$ are invariant under this action. Quotienting by $\mathcal{G}_{z;u}$ allows us to gauge h away except at the marked points. At these points only a copy of G/G_k survives. This is equivalent to the orbit $\mathcal{O}_k$ of u_k under coadjoint action of G. Noting that $\widehat{u}_k$ describes the orbit $\mathcal{O}_k$ we see that:

$$\mathcal{P}_{z;u} = \{(\widehat{u}_k) \in \prod_k \mathcal{O}_k \text{ s.t. } \sum_k \widehat{u}_k = 0\}/G \tag{6.76}$$

where G acts on $\widehat{u}_k$ by $\widehat{u}_k \to l\widehat{u}_k l^{-1}$ for all k.

The reduced symplectic structure is the usual symplectic structure on coadjoint orbits. Indeed consider the canonical one–form on $T^*\mathcal{A}$, i.e. $2i\pi\alpha = \int_\Sigma \mathrm{Tr}\,(\Phi\delta A)$. Parametrizing A by $A = h^{-1}\overline{\partial} h$ one finds $2i\pi\alpha = -\int_\Sigma \mathrm{Tr}\left[(h^{-1}\delta h)\overline{\partial}_A \Phi\right]$ which, using eq.(6.70), reduces to:

$$\alpha = -\sum_k \mathrm{Tr}\,(u_k h_k^{-1} \delta h_k)$$

We recognize that $\delta\alpha$ is the Kirillov symplectic form on $\prod_k \mathcal{O}_k$, which means that the Poisson brackets read:

$$\{\mathrm{Tr}\,(\epsilon\widehat{u}_k), \mathrm{Tr}\,(\epsilon'\widehat{u}_{k'})\} = \delta_{kk'} \mathrm{Tr}\,([\epsilon, \epsilon']\widehat{u}_k)$$

We still have to quotient by the left action of G. Infinitesimal action is given by $\delta\widehat{u}_k = [\epsilon, \widehat{u}_k]$ and is generated by the Hamiltonian $H_\epsilon = \sum_k \mathrm{Tr}\,(\epsilon\widehat{u}_k)$, so the moment is $\mu = \sum_k \widehat{u}_k$ and $\mathcal{P}_{z;u} = \mu^{-1}(0)/G$.

Choosing the invariant polynomial $P_2(\Phi) = \mathrm{Tr}\,(\Phi^2)$ gives the Hamiltonian:

$$P_2(\Phi) = \sum_{k,l} \frac{\mathrm{Tr}\,(\widehat{u}_k \widehat{u}_l)}{(z - z_k)(z - z_l)} = \sum_k \frac{H_k}{z - z_k} \tag{6.77}$$

with

$$H_k = \sum_{l \neq k} \frac{\mathrm{Tr}\,(\widehat{u}_k \widehat{u}_l)}{z_k - z_l}$$

The H_k's form a family of commuting Hamiltonians usually called the Gaudin Hamiltonians, which are very closely related to the Neumann model.

Example 2. Let us explain how one may rederive the elliptic spin Calogero–Moser system from this general construction. Consider the torus $T_\tau \equiv \mathbb{C}/(\mathbb{Z} + \tau\mathbb{Z})$ of periods 1 and τ, $\mathrm{Im}\,\tau > 0$. We shall use a coordinate z on T_τ with the identification $z \sim z+1$ and $z \sim z+\tau$. We consider the case with one marked point at $z = 0$. Let u be an element of the (dual of the) Lie algebra attached to this marked point. We need a Cartan decomposition of the Lie algebra of G, i.e. consider the basis (H_i, E_α), with H_i in the Cartan subalgebra of the Lie algebra of G. We describe any fiber bundle on the torus T_τ using two open sets U_0, a small disk around $z = 0$, and $U_\infty = T_\tau - \{0\}$, using the transition function $g_{0\infty} = h_\infty h_0^{-1}$, where h_0 and h_∞ are C^∞ functions solving $A = h^{-1}\overline{\partial} h$ in the respective open sets U_0 and U_∞. It follows from the above discussion, eq.(6.69), that the transition function $g_{0\infty}$ is equivalent to one of the form

$$g_{0\infty} = e^{q/z}, \quad \text{with } q \text{ diagonal}$$

Note that in general h_0 and h_∞ are defined up to left multiplication by holomorphic functions, but the condition $h_\infty h_0^{-1} = g_{0\infty} = \exp(q/z)$ completely determines h_0 and h_∞ up to left multiplication by the same constant diagonal matrix.

The condition $\overline{\partial}_A \Phi = 2i\pi u \delta_0$ implies in U_0 that $\widehat{\Phi}_0 = h_0 \Phi h_0^{-1}$ has a simple pole at $z = 0$ with residue $\widehat{u} = h_0(0) u h_0(0)^{-1}$. Similarly in U_∞, $\widehat{\Phi}_\infty = h_\infty \Phi h_\infty^{-1}$ is regular. We have $\widehat{\Phi}_\infty = g_{0\infty} \widehat{\Phi}_0 g_{0\infty}^{-1}$ so that $\widehat{\Phi}$ defines an holomorphic section of our holomorphic vector bundle. Writing the patching condition when $z \to 0$ we see that $\widehat{\Phi}_\infty$ is regular on T_τ except at $z = 0$ where it behaves as (in the case where all holomorphic indices k_i vanish):

$$\widehat{\Phi}_\infty = e^{q/z} \left(\frac{\widehat{u}}{z} + O(1)\right) e^{-q/z}$$

In the following we drop the subscript ∞ in $\widehat{\Phi}_\infty$. Introducing the radicial decompositions:

$$\widehat{u} = \sum_i \widehat{u}_i H_i + \sum_\alpha \widehat{u}_\alpha E_\alpha, \quad \widehat{\Phi} = \left(\sum_i \widehat{\Phi}_i H_i + \sum_\alpha \widehat{\Phi}_\alpha E_\alpha \right) dz$$

we see that $\widehat{\Phi}_i \sim \frac{\widehat{u}_i}{z}$, and $\widehat{\Phi}_\alpha \sim \exp\left(\frac{\alpha(q)}{z}\right) \frac{\widehat{u}_\alpha}{z}$. The first condition means that $\widehat{\Phi}_i$ is an elliptic function with a single pole, hence $\widehat{u}_i = 0$ and $\widehat{\Phi}_i = p_i$ is constant. The second condition means that $\widehat{\Phi}_\alpha$ has an essential singularity of the Baker–Akhiezer type at $z = 0$ and the solution is the Lamé function:

$$\widehat{\Phi}_\alpha = -\widehat{u}_\alpha \frac{\sigma(z - \alpha(q))}{\sigma(z)\sigma(\alpha(q))} e^{\alpha(q)\zeta(z)}$$

The section $\widehat{\Phi}$ finally reads:

$$\widehat{\Phi} \equiv \widehat{\Phi}_\infty = \left(\sum_i p_i H_i + \sum_\alpha \widehat{\Phi}_\alpha E_\alpha \right) dz \tag{6.78}$$

which is exactly of the same form as the Lax matrix of the elliptic spin Calogero model (in the $sl(n)$ case). The variables (p_i, q_i) where $q = \sum_i q_i H_i$ are the momenta and positions of the particles of the Calogero model, while the $\widehat{u}_\alpha$ are the spin variables. Choosing now $P_2(\Phi) = \mathrm{Tr}\,(\Phi^2)$ and using the property of Lamé functions, eq.(6.16) we can relate $P_2(\Phi)$ to the Hamiltonian of the spin Calogero model.

$$P_2(\Phi) = H + \sum_\alpha (\widehat{u}_\alpha \widehat{u}_{-\alpha}) \wp(z), \quad \text{with } H = \sum_i p_i^2 - \sum_\alpha \wp(\alpha(q))\, \widehat{u}_\alpha \widehat{u}_{-\alpha}$$

As in the previous example, taking the quotient by the gauge group leaves only the variables q_i and p_i which parametrize respectively the moduli space $\mathcal{N}$, and the Cartan component of $\widehat{\Phi}$, and the spin variables $\widehat{u} = h_0(0)uh_0(0)^{-1}$ describing the G–orbit of u, restricted by $\widehat{u}_i = 0$. Moreover, since h is defined up to left multiplication by a constant diagonal matrix, and this leaves the constraint $\widehat{u}_i = 0$ invariant, the phase space is:

$$\mathcal{P}_u = \{(p, q, \widehat{u} \in \mathcal{O}_u) \text{ s.t. } \widehat{u}_i = 0\}/H$$

where H is the Cartan torus, acting on $\widehat{u}$ by conjugation.

It remains to show that our dynamical variables have the correct Poisson brackets. To do that we compute the canonical one–form:

$$2i\pi\alpha = \int_\Sigma \mathrm{Tr}\,(\Phi\delta A) = \int_{U_0} \mathrm{Tr}\,(\Phi\delta A) + \int_{\Sigma-U_0} \mathrm{Tr}\,(\Phi\delta A)$$

where we use the description of the bundles on T_τ with two open sets. Writing $A = h_0^{-1}\overline{\partial}h_0 = h_\infty^{-1}\overline{\partial}h_\infty$ in the respective open sets one computes:

$$\begin{aligned}\int_{U_0} \mathrm{Tr}\,(\Phi\delta A) &= -\int_{U_0} \mathrm{Tr}\,(h_0^{-1}\delta h_0\overline{\partial}_A\Phi) + \int_{\partial U_0} \mathrm{Tr}\,(\Phi h_0^{-1}\delta h_0)\\ \int_{\Sigma-U_0} \mathrm{Tr}\,(\Phi\delta A) &= -\int_{\Sigma-U_0} \mathrm{Tr}\,(h_\infty^{-1}\delta h_\infty\overline{\partial}_A\Phi) + \int_{\partial(\Sigma-U_0)} \mathrm{Tr}\,(\Phi h_\infty^{-1}\delta h_\infty)\end{aligned}$$

First we have $\int_{\partial(\Sigma-U_0)} = -\int_{\partial U_0}$. Second since $\overline{\partial}_A\Phi = u\delta_0$ the integral $\int_{\Sigma-U_0}$ vanishes, and we get:

$$\alpha = -\mathrm{Tr}\,(uh_0(0)^{-1}\delta h_0(0)) + \frac{1}{2i\pi}\int_{\partial U_0} \mathrm{Tr}\left(\Phi(h_0^{-1}\delta h_0 - h_\infty^{-1}\delta h_\infty)\right)$$

The contour integral can be rewritten as $\int_{\partial U_0} \mathrm{Tr}\left(\widehat{\Phi}_\infty(h_\infty h_0^{-1}\delta h_0 h_\infty^{-1} - \delta h_\infty h_\infty^{-1})\right)$. Varying the equation $h_\infty h_0^{-1} = g_{0\infty} = \exp{(q/z)}$ we have:

$$\frac{\delta q}{z} = \delta h_\infty h_\infty^{-1} - h_\infty h_0^{-1}\delta h_0 h_\infty^{-1}$$

so we finally get:

$$\alpha = -\mathrm{Tr}\,(uh_0(0)^{-1}\delta h_0(0)) - \frac{1}{2i\pi}\int_{\partial U_0} \mathrm{Tr}\left(\widehat{\Phi}_\infty\frac{\delta q}{z}\right) = -\sum_i p_i\delta q_i - \mathrm{Tr}\,(uh(0)^{-1}\delta h(0))$$

where to evaluate the contour integral, we noticed that since δq is diagonal, only the diagonal part of $\widehat{\Phi}$ in eq.(6.78) contributes to the trace. It is clear that the constraint $\widehat{u}_i = 0$ and the quotient by the Cartan torus is a symplectic quotient of the coadjoint orbit, so finally the symplectic form reads:

$$\omega = -\delta\alpha = \sum_i \delta p_i \wedge \delta q_i + \omega_K$$

Remark 1. The solution of the Hitchin systems can be viewed as the projection, under taking the quotient by the gauge group, of a straight line motion on $T^*\mathcal{A}$ since on this space the equations of motion are of the form $\dot{\Phi} = 0$ and $\dot{A} = P'(\Phi) = \text{Const}$.

Remark 2. Note that the constraints $\widehat{u}_i = 0$ are the same as the constraints $f_{ii} = \alpha$ in eq.(6.6), because the Cartan algebra for $sl(N)$ is generated by $e_{ii} - e_{jj}$. Moreover it is clear that $\widehat{\Phi}$ occuring in eq.(6.78) is the same as the Lax matrix in eq.(6.11). However this provides some insight in the true nature of the Lax matrix, which appears as a form–valued holomorphic section of a vector bundle (depending on the dynamical data).

Remark 3. The above construction naturally yields a Lax pair formulation of the equations of motion. Indeed $\widehat{\Phi} = h\Phi h^{-1}$, hence $d\widehat{\Phi}/dt = [M, \widehat{\Phi}]$ with $M = \dot{h}h^{-1}$. As has been emphasized above h is determined up to a constant (in z) diagonal matrix, the diagonal action on these data is given by $\widehat{\Phi} \to l\widehat{\Phi}l^{-1}$ and $M \to lMl^{-1} + \dot{l}$.

6.13 The trigonometric Calogero model.

In this section we show two remarkable properties of the trigonometric limit. The first one is the construction by Olshanetski and Perelomov of the trigonometric spin Calogero model by Hamiltonian reduction of the geodesic motion on a symmetric space, in the same spirit as the above construction of the elliptic Calogero model. The second one is the uncovering of a quite unexpected infinite symmetry of the trigonometric spin Calogero model.

Calogero model and symmetric spaces. Recall that the Lax matrix of the elliptic spin Calogero model, eq.(6.11) reads in the trigonometric case:

$$L_{ij} = p_i\delta_{ij} + (1-\delta_{ij})f_{ij}(\coth(q_{ij} - \coth(\lambda))e^{q_{ij}\coth(\lambda)}$$

Taking the limit $\lambda \to -\infty$ we get:

$$L_{ij} = p_i\delta_{ij} + (1-\delta_{ij})\frac{f_{ij}}{\sinh(q_i - q_j)}$$

This Lax matrix, without spectral parameter, is a good Lax matrix in the case of the scalar trigonometric Calogero model, as it produces N commuting conserved quantities. In the spin case, however, it is insufficient, and one has to work either with the Lax matrix with spectral parameter, or one can appeal to the Yangian symmetry introduced below, to provide an extra set of conserved commuting quantities. We will construct this Lax matrix by Hamiltonian reduction procedure, starting from a finite–dimensional space.

The construction is similar as in the case of the open Toda chain, see Chapter [3]. One starts from the symplectic space T^*G where G is a complex Lie group and reduce under the action on the left by a subgroup H_L and on the right by a subgroup H_R. The general theory, see Chapter [14], shows that the Poisson brackets in the coordinates (g,ξ) are expressed as:

$$\{\xi(X),\xi(Y)\} = \xi([X,Y]), \quad \{\xi(X),g\} = g\,X, \quad \{g,g\} = 0 \tag{6.79}$$

Recall that here g is the matrix of functions on G whose elements are the matrix elements of g in a faithful representation.

As in the case of the Toda chain we reduce the geodesic motion on G given by the Hamiltonian on T^*G, $H = \frac{1}{2}\,\mathrm{Tr}\,(\xi^2)$. Notice that H is bi–invariant, so one can attempt to reduce this dynamical system using Lie subgroups H_L and H_R of G with Lie algebras $\mathcal{H}_L$ and $\mathcal{H}_R$, acting respectively on the left and on the right on T^*G in order to obtain a non–trivial result. Recall that the corresponding moments are, see eq.(3.40) in Chapter [3]:

$$\mathcal{P}^L(g,\xi) = -P_{\mathcal{H}_L^*}\,g\xi g^{-1} \quad \mathcal{P}^R(g,\xi) = P_{\mathcal{H}_R^*}\,\xi \tag{6.80}$$

where we have introduced the projectors on $\mathcal{H}_{L,R}^*$ of forms in $\mathcal{G}^*$ induced by the restriction of these forms to $\mathcal{H}_{L,R}$, the Lie algebras of $H_{L,R}$.

Let us consider an involutive automorphism σ of G and let H be the subgroup of its fixed points. Then H acts on the right on G defining a principal fiber bundle of total space G and base G/H, which is a symmetric space. Moreover G acts on the left on G/H and in particular so does H itself. We consider the dynamical reduction under this left action of H on G/H. Equivalently we shall consider the reduction of the geodesic motion on G under the product $H_L \times H_R$ with $H_L = H_R = H$. The reduction is achieved with an adequate choice of the momentum $\mu = (-\mu_L, \mu_R)$ such that $(\mathcal{P}^L, \mathcal{P}^R) = \mu$. We take $\mu_R = 0$ so that the isotropy group of the right component is H_R itself. In this way we are in fact dealing with motions on G/H.

The derivative of σ at the unit element of G is an involutive automorphism of $\mathcal{G}$ also denoted σ. Let us consider its eigenspaces $\mathcal{H}$ and $\mathcal{K}$ associated with the eigenvalues $+1$ and -1 respectively. Thus we have a decomposition:

$$\mathcal{G} = \mathcal{H} \oplus \mathcal{K} \tag{6.81}$$

in which $\mathcal{H}$ is the Lie algebra of H. Note that $h\mathcal{K}h^{-1} = \mathcal{K}$, for $h \in \mathcal{H}$.

We shall consider the particular instance of the object G/H obtained by taking $G = SL(N, \mathbb{C})$ and $H = SU(N)$. Here we view $SL(N, \mathbb{C})$ as a *real* Lie algebra. Counting real parameters $\dim G = 2N^2 - 2$ (note that $\det M = 1$ yields 2 conditions), and $\dim H = N^2 - 1$. The automorphism σ is $\sigma(g) = (g^*)^{-1}$ where $*$ means transpose and complex conjugate, hence the set of its fixed points is $SU(N)$. The symmetric space $K = G/H$ may be identified with the space of positive Hermitean matrices. Indeed for any matrix $M \in SL(N, \mathbb{C})$ we can uniquely write the Cayley decomposition $M = KU$ with $U \in SU(N)$ and K positive Hermitean. This is because MM^* is Hermitean and strictly positive, so we can write $MM^* = VDV^{-1}$ with $V \in SU(N)$ and D diagonal real positive. We define $K = V\sqrt{D}V^{-1}$, where $\sqrt{D}$ is the positive square root. Then we check that $K^{-1}M$ is in $SU(N)$. At the Lie algebra level we have the decomposition (6.81) where $\mathcal{H}$ is the Lie algebra of antihermitean matrices, and $\mathcal{K}$ is the space of Hermitean matrices. Finally σ reads $X \to -X^*$, and is an automorphism for the structure of real Lie algebra. The dual of $\mathcal{G}$ is identified with $\mathcal{G}$ using the symmetric real invariant bilinear form $(X, Y) = \mathrm{Tr}\,(XY + X^*Y^*)$.

The choice of the moment μ_L is of course of crucial importance. We will consider μ_L in the (dual of the) Lie algebra of H, i.e. an antihermitean matrix, such that its isotropy subgroup H_μ is a maximal proper Lie subgroup of H_L, so that the phase space of the reduced system be of minimal dimension but non trivial. As a matter of fact the dimension of the reduced phase space is $2\dim G - \dim(H \times H) - \dim(H_\mu \times H)$. This is because $\dim T^*G = 2\dim G$, the constraint $(\mathcal{P}^L, \mathcal{P}^R) = (-\mu_L, 0)$ yields $\dim(H \times H)$ equations, and we still have to quotient by $H_\mu \times H$. In particular, if $H_\mu = H$ we get, in our case, a reduced phase space of dimension 0. To analyze the stabilizer of μ we have to solve $h^{-1}\mu h = \mu$. In a basis where μ is diagonal, we see that if μ has $N - l$ equal eigenvalues and the other l eigenvalues all different, the stabilizer is $U(1)^l \times U(N - l)/U(1)$. This yields a reduced phase space of dimension $2Nl - l(l + 1)$. Note that this is the dimension of the phase space of the spin Calogero model, eq.(6.9), apart from a discrepancy of 2 which corresponds to a reduction by the center of mass motion, as we shall see.

We now introduce coordinates on the group G. Using the Cayley decomposition we can write uniquely $g = KU$ with K Hermitean positive and U unitary. Diagonalizing K we write $K = h_L Q h_L^{-1}$ with Q diagonal with real positive entries, and $h_L \in SU(N)$. The columns of h_L form the orthonormal basis of eigenvectors of K, hence are defined up to a phase. So we can write $g = h_L Q h_R^{-1}$ where $h_L, h_R \in SU(N)$ are defined up to right multiplication by the same unitary diagonal matrix.

The set of points of T^*G with given momentum $(-\mu_L, 0)$ is defined by the equations:

$$P_{\mathcal{H}}\,(g\xi g^{-1}) = P_{\mathcal{H}}\,(h_L Q h_R^{-1}\xi h_R Q^{-1} h_L^{-1}) = \mu_L, \quad P_{\mathcal{H}}\,(\xi) = 0 \tag{6.82}$$

The second condition means that $\xi \in \mathcal{K}$ is an Hermitean traceless matrix. Setting $L = h_R^{-1}\xi h_R \in \mathcal{K}$, the first condition reads $(h_L QLQ^{-1} h_L^{-1}, X) = (\mu_L, X), \forall X \in \mathcal{H}$. Equivalently $(QLQ^{-1}, h_L^{-1} X h_L) = (h_L^{-1}\mu_L h_L, h_L^{-1} X h_L)$ yielding:

$$P_{\mathcal{H}}\,(QLQ^{-1}) = h_L^{-1}\mu_L h_L$$

Setting $Q = \mathrm{Diag}\,(e^{q_i})$ with $\sum_i q_i = 0$ (hence the reduction by the center of mass motion), we have $(QLQ^{-1})_{ij} = \exp\,(q_i - q_j) L_{ij}$. The projection on $\mathcal{H}$ amounts to take the antihermitean part of this matrix, giving the equation $\sinh\,(q_{ij}) L_{ij} = \widehat{\mu}_{ij}$ with $\widehat{\mu}_{ij} = (h_L^{-1}\mu_L h_L)_{ij}$. The solution of this equation is:

$$L_{ij} = \frac{\widehat{\mu}_{ij}}{\sinh\,(q_{ij})}, \; i \neq j, \quad L_{ii} = p_i, \;\; \sum p_i = 0, \quad \widehat{\mu}_{ii} = 0$$

where the p_i are arbitrary real numbers. Hence we have found:

$$L = \sum_i p_i E_{ii} + \sum_{i \neq j} \frac{\widehat{\mu}_{ij}}{\sinh\,(q_{ij})} E_{ij}$$

We recognize the Lax matrix L of the trigonometric spin Calogero model, with the spin variables $\widehat{\mu}_{ij}$ describing the coadjoint H–orbit of the momentum μ_L. In view of the definition of L the ambiguity on the definition of h_L, h_R amounts to quotienting by a conjugation by a diagonal unitary matrix.

We still have not explored the implications of the conditions $\widehat{\mu}_{ii} = 0$ on h_L. We do this for the case of the scalar model, i.e. μ_L has $N-1$ equal eigenvalues. This means that μ_L is of the form $\mu_L = i(VV^* - \alpha I)$ where V is an N–vector and α is determined so that μ_L is traceless. We then have $\widehat{\mu} = i(WW^* - \alpha I)$ where $W = h_L^{-1}V$. The conditions $\widehat{\mu}_{ii} = 0$ read $W_i\overline{W}_i = \alpha$ for all i. Since h_L is defined up to right multiplication by a diagonal unitary matrix, we can always assume that W_i is real positive, so that $W_i = \sqrt{\alpha}$. Of course there always exists some fixed $h_0^{-1} \in SU(N)$ mapping V to the vector $W = \sqrt{\alpha}(1, \cdots, 1)^t$, and any other solution h_L is of the form $h_L = h_0 h_W$ where h_W runs over the stability group of W, which is exactly $h_0^{-1} H_\mu h_0$ where H_μ is the stabilizer of μ. Here the element h_0 plays essentially no role, and one usually takes $\mu_L = i(WW^* - \alpha I)$. Then all solutions of the constraints $(\mathcal{P}_L, \mathcal{P}_R) = (-\mu_L, 0)$ are of the form $(h_L, h_R).(Q, L)$ with Q, L as above, and $h_L \in H_\mu$, as it should be. Quotienting by $H_\mu \times H$, according to the general procedure, yields the phase space of the Calogero model. A similar analysis can be performed in the case of the spin Calogero model.

Finally we identify the symplectic structure coming from Hamiltonian reduction to the symplectic structure of the spin Calogero model. It is enough to compute the canonical one–form:

$$\alpha = (\xi, g^{-1}\delta g) = (QLQ^{-1}, h_L^{-1}\delta h_L) + (L, Q^{-1}\delta Q) - (L, h_R^{-1}\delta h_R)$$

Using the constraint eq.(6.82) , the first term is $(\mu_L, \delta h_L h_L^{-1})$, namely the Kirillov structure on the spin variables, the second term is $\sum_i p_i \delta q_i$, and the last term vanishes because the constraint $P_{\mathcal{H}}\ (\xi) = 0$ also implies $P_{\mathcal{H}}\ (L) = 0$.

Yangian symmetry. We now show that the trigonometric limit of the spin Calogero model exhibits an unexpected symmetry, called a Yangian symmetry. This symmetry is induced by the parametrization (6.27) of f_{ij}. It is obvious that this parametrization introduces a $sl(l)$ symmetry into the theory: the transformation

$$a_i^r \longrightarrow \sum_{s=1}^{l} u^{rs} a_i^s, \quad b_i^r \longrightarrow \sum_{s=1}^{l} (u^{-1})^{rs} b_i^s$$

leaves the f_{ij} invariant and therefore also the Hamiltonians. The remarkable feature is that, this symmetry extends, in the trigonometric case, to an infinite symmetry algebra. The generators of the $sl(l)$ symmetry are the conserved charges

$$J_0^{rs} = \sum_i a_i^r b_i^s$$

We will construct a whole hierarchy of conserved charges J_n^{rs} that generate a Yangian symmetry. To do that, recall that the Lax matrix in the trigonometric case, eq.(6.17) reads:

$$L(\lambda) = L_0 - \coth(\lambda)(F - \alpha), \quad L_0 = \sum_{i=1}^{N} p_i e_{ii} + \sum_{\substack{i,j=1 \\ i \neq j}}^{N} \coth(q_{ij})\ f_{ij}\ e_{ij} \tag{6.83}$$

We consider the matrix:

$$L = \lim_{\lambda \to -\infty} L(\lambda) + \alpha = L_0 + F$$

and the subset $\mathrm{Tr}\,(L^n)$ of commuting quantities. Notice that th Hamiltonian H belongs to this subset, since a short computation shows that $H = \frac{1}{2}\mathrm{Tr}\, L^2 - \alpha\ \mathrm{Tr}\, L - \frac{1}{2}N\alpha^2$. We introduce the following quantities:

$$J_n^{rs} = \mathrm{Tr}\,(L^n F^{rs}), \quad r, s = 1, \cdots, l \quad n = 0, 1, \cdots, \infty \tag{6.84}$$

where F^{rs} is the $N \times N$ matrix of elements

$$(F^{rs})_{ij} = b_i^s a_j^r$$

We define the generating functional of the currents J_n^{rs}. It is the $l \times l$ matrix $T(z)$ with elements:

$$T^{rs}(z) = -\frac{1}{2}\delta_{rs} - \sum_{n \in \mathbb{N}} \frac{1}{z^{n+1}} J_n^{rs} = -\frac{1}{2}\delta_{rs} + \mathrm{Tr}\left(\frac{1}{L-z}F^{rs}\right) \tag{6.85}$$

Note that in this equation there appears the resolvent $\rho(z) = (L-z)^{-1}$ of the matrix L.

Proposition. *On the reduced manifold $f_{ii} = \alpha$ we have the following two properties:*
(i) The charges J_n^{rs} Poisson commute with the subset of Hamiltonians of the form $\mathrm{Tr}(L^n)$.
(ii) The generating functional $T(z)$ satisfies the defining relation of a (classical) Yangian algebra:

$$\{T_1(y), T_2(z)\} = -2[\frac{\Pi}{y-z}, T_1(y)T_2(z)], \quad \Pi = \sum_{r,s=1}^{l} e_{rs} \otimes e_{sr} \tag{6.86}$$

Proof. To prove this proposition we need the Poisson brackets

$$\{L_1, L_2\} = [r_{12}^0, L_1] - [r_{21}^0, L_2] + \sum_{\substack{i,j=1 \\ i \neq j}}^{N} (f_{ii} - f_{jj}) \frac{1}{\sinh^2(q_{ij})} e_{ij} \otimes e_{ji} \tag{6.87}$$

$$\{L_1, F_2^{rs}\} = [-r_{21}^0 + \mathcal{C}_{12}, F_2^{rs}] \tag{6.88}$$

$$\{F_1^{rs}, F_2^{tu}\} = (\delta_{ru} F_1^{ts} - \delta_{st} F_2^{ru})\, \mathcal{C}_{12} \tag{6.89}$$

where r_{12}^0 is obtained taking a limit in eq.(6.26) and is explicitly given by $r_{12}^0 = -\sum_{i \neq j} \coth(q_{ij}) e_{ij} \otimes e_{ji}$ and $\mathcal{C}_{12} = \sum_{ij} e_{ij} \otimes e_{ji}$ is the Casimir for $sl(N)$.

Remark that the currents J_n^{rs} and the Hamiltonians $\mathrm{Tr}(L^n)$ are invariant under the rescaling symmetry

$$b_i^r \longrightarrow \lambda_i \, b_i^r, \quad a_i^r \longrightarrow \lambda_i^{-1} a_i^r$$

Therefore we can compute their Poisson brackets on the reduced phase space straightforwardly; restricting ourselves to the manifold $f_{ii} = \alpha$, the last term in eq.(6.87) vanishes, and we will systematically drop its contribution in intermediate calculations.

We emphasize that in eq.(6.87,6.88) the same r-matrix appears. Moreover it is the term $[\mathcal{C}_{12}, F_2^{rs}]$ in eq.(6.88) which is responsible for the quadratic form of eq.(6.86), as we shall see in the following.

Introducing the generating functional $H(z) = \mathrm{Tr}\rho(z)$ of the Hamiltonians $\mathrm{Tr}(L^n)$ we compute

$$\begin{aligned} \{\rho_1(y)F_1^{rs}, \rho_2(z)\} &= -\left[\rho_2(z)\, r_{12}^0\, \rho_2(z)\ ,\ \rho_1(z)F_1^{rs}\right] + \left[\rho_1(y)\, r_{21}^0\, \rho_1(y)F_1^{rs}\ ,\ \rho_2(z)\right] \\ &\quad + \rho_1(y)\rho_2(z)\ [\mathcal{C}_{12}, F_1^{rs}]\ \rho_2(z) \end{aligned}$$

Taking the trace we obtain $\{T^{rs}(y), H(z)\} = \mathrm{Tr}\, \left(F^{rs}[\rho(y), \rho^2(z)]\right) = 0$. This proves the first part of the proposition. To prove the second part we evaluate

$$\begin{aligned} \{\rho_1(y)F_1^{rs}, \rho_2(z)F_2^{tu}\} &= -\left[\rho_2(z)\, r_{12}^0\, \rho_2(z)F_2^{tu}\ ,\ \rho_1(y)F_1^{rs}\right] + \left[\rho_1(y)\, r_{21}^0\, \rho_1(y)F_1^{rs}\ ,\ \rho_2(z)F_2^{tu}\right] \\ &\quad + \rho_1(y)\rho_2(z)\left\{(\delta_{ru}F_1^{ts} - \delta_{ts}F_2^{ru})\ \mathcal{C}_{12} + [\mathcal{C}_{12}, F_1^{rs}]\ \rho_2(z)F_2^{tu}\ -\ [\mathcal{C}_{12}, F_2^{tu}]\ \rho_1(y)F_1^{rs}\right\} \end{aligned}$$

Taking the trace we get:

$$\{T^{rs}(y), T^{tu}(z)\} = \mathrm{Tr}\, \left\{\left(\rho(y)\rho(z)(\delta_{ru}F^{ts} - \delta_{ts}F^{ru})\right) + \left(\rho(y)\left[\rho(z)F^{tu}\rho(z), F^{rs}\right]\right) - \left(\rho(z)\left[\rho(y)F^{rs}\rho(y), F^{tu}\right]\right)\right\}$$

Using the cyclicity of the trace and $\rho(y)\rho(z) = (\rho(y) - \rho(z))/(y-z)$ this becomes:

$$\begin{aligned} \{T^{rs}(y), T^{tu}(z)\} &= \frac{1}{y-z}\left(\delta_{ru}(T^{ts}(y) - T^{ts}(z)) - \delta_{ts}(T^{ru}(y) - T^{ru}(z))\right) \\ &\quad + \frac{2}{y-z}\mathrm{Tr}\, \left(\rho(y)F^{tu}\rho(z)F^{rs} - \rho(y)F^{rs}\rho(z)F^{tu}\right) \end{aligned}$$

Finally remarking that:

$$\begin{aligned}
\mathrm{Tr}\,\left(\rho(y)F^{tu}\rho(z)F^{rs}\right) &= \sum_{ijkl=1}^{N}\rho_{ij}(y)b_j^u a_k^t \rho_{kl}(z)b_l^s a_i^r = \left(\sum_{ij=1}^{N}\rho_{ij}(y)b_j^u a_i^r\right)\left(\sum_{kl=1}^{N}\rho_{kl}(z)b_l^s a_k^t\right) \\
&= \left(T^{ru}(y)+\frac{1}{2}\delta_{ru}\right)\left(T^{ts}(z)+\frac{1}{2}\delta_{ts}\right)
\end{aligned}$$

we prove the result (6.86). ∎

The proposition allows to organize the commuting conserved quantities into Yangian invariant ones, namely the $\mathrm{Tr}\,L^n$, and non Yangian invariant ones, namely $\mathrm{Tr}\,T^m(z)$ (for $m < l$, here the trace is over the l–dimensional internal space). Remark that the quadratic Poisson algebra eq.(6.86) also appears in the study of the closed Toda chain, and we have obtained here a very different representation of the same algebra.

The rational limit is obtained by applying the canonical transformation $p_i \longrightarrow \frac{1}{\epsilon}\,p_i, \quad q_i \longrightarrow \epsilon\, q_i$ and sending ϵ to zero. In this limit one gets a linear Poisson algebra

$$\{T_1(y), T_2(z)\} = [\frac{\Pi}{y-z}, T_1(y)+T_2(z)]$$

We see that going from the rational to the trigonometric case yields a deformation of this linear Poisson algebra. Wether there exists a further deformation of eq.(6.86) in the elliptic model is still unknown at the time of this writing.

References

[1] Calogero F., Exactly solvable one-dimensional many-body systems, Lett.Nuovo Cimento **13** (1975), 411-415.

[2] Moser J., Three integrable Hamiltonian systems connected with isospectral deformations, Adv.Math. **16** (1975), 441-416.

[3] Olshanetsky M.A., Perelomov A.M., Classical integrable finite dimensional systems related to Lie algebras. Physics Reports **71** (1981)

[4] Krichever I.M., Elliptic solutions of Kadomtsev-Petviashvilii equation and integrable systems of particles, Func. Anal. App **14** (1980), n 4, 282-290.

[5] Gibbons J., Hermsen T., A generalization of the Calogero-Moser system., Physica **11D** (1984), 337.

[6] Airault H., McKean H., Moser J., Rational and elliptic solutions of the KdV equation and related many-body problem, Comm. Pure and Appl. Math. **30** (1977), 95–125.

[7] Hitchin N., Stable bundles and integrable systems., Duke Math. Journ. **54** (1987) 91.

[8] Nekrasov N., Holomorphic bundles and many body systems. Comm. Math. Phys. **180** (1996) 587–604.

[9] Narasimhan M.S., Seshadri C.S., Holomorphic vector bundles on a compact Riemann surface, Math. Ann. **155** (1964) 69–80

[10] Gunning R.C., **Lectures on vector bundles over Riemann surfaces**, Princeton University Press (1967)

[11] Forster O., **Lectures on Riemann Surfaces**, Springer (1981)

Chapter 7

Isomonodromy deformations.

In this Chapter, the Lax framework is extended by replacing the Lax matrix by a first orded linear differential operator $L(\lambda) \to \partial_\lambda + M_\lambda(\lambda)$. Here the emphasis is on the study of monodromy properties of linear differential equations, and particularly to determine M_λ with a given monodromy. In general, the solution is not unique, and depends on a number of continuous parameters, the so–called isomonodromic deformation parameters. We show that the deformation equations with respect to these parameters are an integrable system. Ordinary integrable systems of the Lax type appear as particular cases of such systems, namely when the group generated by the monodromies is finite. However the generic situation is much more general, and this yields new integrable systems. In particular Painlevé equations appear in this setting. Just as solution of Lax equations were written in terms of θ functions, in the general case, the solutions are obtained in terms of new functions called τ functions. We show how to express the dynamical variables in terms of τ functions and their Schlesinger transforms, and show that, with this parametrization the equations of motion take the form of bilinear Hirota equations.

7.1 Introduction

In Chapter [2], we have seen that the equations

$$\partial_{t_i} w = -\left(g\xi_i g^{-1}\right)_- w; \quad w = g e^{-\sum_i \xi_i t_i} \tag{7.1}$$

define an integrable hierarchy. Here i is a multi index $i = (k, n, \alpha)$, and ξ_i is the diagonal matrix $\xi_i = (\lambda - \lambda_k)^{-n} E_{\alpha\alpha}$. We assume that the matrices are of size $N \times N$. The notation $(\,)_-$ means taking the polar part at λ_k. In eq.(7.1) the wave function $w(\lambda)$ was defined through the Lax equation $L(\lambda)w(\lambda) = w(\lambda)\widehat{\mu}$ where $\widehat{\mu}$ is the diagonal matrix of the eigenvalues of L. The flows defined by eq.(7.1) are all commuting and satisfy the zero curvature equations :

$$\partial_{t_i} M_j - \partial_{t_j} M_i + [M_i, M_j] = 0; \quad M_i = \left(g\xi_i g^{-1}\right)_- \tag{7.2}$$

In this Chapter we enlarge the framework by replacing the Lax equation by a differential equation in the spectral parameter λ. We further assume that this differential equation is linear with coefficients rational functions of λ.

$$\partial_\lambda w = -M_\lambda w \tag{7.3}$$

$$M_\lambda(\lambda) = \sum_{k=1}^{K} \Big(\frac{A_1^{(k)}}{(\lambda - \lambda_k)} + \cdots + \frac{A_{n_k+1}^{(k)}}{(\lambda - \lambda_k)^{n_k+1}} \Big) - A_0^{(\infty)} - \cdots - A_{n_\infty - 1}^{(\infty)} \lambda^{n_\infty - 1} \tag{7.4}$$

Again we will show that around each pole λ_k of M_λ the function $w(\lambda)$ admits an expansion of the form

$$w(\lambda) \sim g^{(k)}(\lambda)\, e^{-\xi^{(k)}(\lambda)}$$

where $g^{(k)}(\lambda) = g_0^{(k)}(1 + g_1^{(k)}(\lambda - \lambda_k) + \cdots)$ and

$$\begin{aligned}
\xi^{(k)} &= B_0^{(k)} \log(\lambda - \lambda_k) + \sum_{(n,\alpha)} \frac{t_{(k,n,\alpha)}}{(\lambda - \lambda_k)^n} E_{\alpha\alpha} \\
\xi^{(\infty)} &= B_0^{(\infty)} \log(\frac{1}{\lambda}) + \sum_{(n,\alpha)} t_{(\infty,n,\alpha)} \lambda^n E_{\alpha\alpha}
\end{aligned} \tag{7.5}$$

While the equations look similar as in the previous context, there are important differences. One of them is that the quantity $\xi^{(k)}$ includes now a logarithmic term, and another one is that the expansions are now only valid in the asymptotic sense (in general). Nevertheless this is sufficient to show that the polar part $M_\lambda^{(k)}$ of M_λ at λ_k is given by

$$M_\lambda^{(k)} = \left(g^{(k)} \partial_\lambda \xi^{(k)} g^{(k)\,-1}\right)_-$$

Again the integrable flows will be the evolutions with respect to the parameters $t_{(k,n,\alpha)}$ occurring in eq.(7.5) and these flows (which were isospectral flows in the Lax case) will be interpreted as isomonodromic deformations of the differential equation. Under these time evolutions we have

$$\partial_{t_i} w = -M_i w, \quad \text{with} \ \ M_i = M_i^{(k)} = \left(g^{(k)} \partial_{t_i} \xi^{(k)} g^{(k)\,-1}\right)_-, \quad i = (k, n, \alpha)$$

We will even generalize further the system by varying the positions λ_k of the poles.

$$\partial_{\lambda_k} w = -M_{\lambda_k} w, \quad M_{\lambda_k} = \left(g^{(k)} \partial_{\lambda_k} \xi^{(k)} g^{(k)\,-1}\right)_- \tag{7.6}$$

and show that this enlarges the set of commuting flows.

7.2 Monodromy data.

Our purpose is to clarify the relation between the globally defined function $w(\lambda)$ and the local expansions $w(\lambda) = g(\lambda) \exp(-\xi(\lambda))$ around each pole λ_k. We start from a wave–function $w(\lambda)$ solution of the linear differential equation:

$$\partial_\lambda w = -M_\lambda(\lambda) w$$

where $M_\lambda(\lambda)$ is a globally defined rational function of λ. Namely it is a sum of polar parts at given poles λ_k, each polar part being of the form $\sum_j M_j^{(k)}/(\lambda - \lambda_k)^j$. This is indeed of the form eq.(7.3) but here we consider the parameters $M_j^{(k)}$ as given and our aim is to reconstruct the asymptotic expansions from the differential equation, thereby obtaining $M_\lambda^{(k)} = \left(w \partial_\lambda \xi^{(k)} w^{-1}\right)_-$.

We begin by recalling some basic definitions and facts concerning linear differential equations of the form eq.(7.3) with singular points.

Definition. *The point λ_k is a regular singular point of the equation $\partial_\lambda w(\lambda) = -M_\lambda(\lambda) w(\lambda)$ if $M_\lambda(\lambda)$ has a simple pole at λ_k. When $M_\lambda(\lambda)$ has a pole of order $n_k + 1 \geq 2$ the point λ_k is called an irregular singular point of rank n_k.*

The point $\lambda = \infty$ is reduced to this definition by setting $\lambda' = 1/\lambda$. The equation becomes $\partial'_\lambda w = 1/\lambda'^2 M_\lambda(1/\lambda') w$. Assuming that $M_\lambda(\lambda)$ is a sum of poles at finite distance, we have $M_\lambda(\lambda) = O(1/\lambda)$ at ∞ so that $1/\lambda'^2 M_\lambda(1/\lambda')$ has a simple pole at ∞ which is thus a regular singular point.

The behaviour of solutions at these two types of singularities is very different. Assume first that the equation $\lambda\partial_\lambda w(\lambda) = -A(\lambda)w(\lambda)$ has a regular singularity at $\lambda = 0$, that is $A(\lambda)$ is analytic around $\lambda = 0$. We also assume that all eigenvalues of $A(0)$ are different and do not differ by integers.

Theorem. *There exists a fundamental matrix of solutions in some vicinity of $\lambda = 0$ of the form:*

$$w(\lambda) = g(\lambda)\, e^{-B\log\lambda}$$

where $g(\lambda)$ is analytic around $\lambda = 0$, and B is the diagonal form of $A(0)$.

The coefficients of the development of $g(\lambda)$ are obtained plugging a series expansion in the differential equation, and determining the coefficients recursively. The main point is that one gets a converging series.

If now we have an irregular singularity at $\lambda = 0$, i.e. the equation is of the form $\lambda^{n+1}\partial_\lambda w(\lambda) = -A(\lambda)w(\lambda)$ with $n \geq 1$, assume that the most singular term, i.e. $A(0)$ has distinct eigenvalues.

Theorem. *The differential equation has a fundamental system of formal solutions in the vicinity of $\lambda = 0$ of the form:*

$$w(\lambda) = g(\lambda)\, e^{-B_0\log\lambda}\, e^{-\xi(\lambda)} \tag{7.7}$$

where B and ξ are diagonal matrices, $\xi(\lambda)$ is a polynomial in $1/\lambda$ with dominant term D/λ^n (where D is the diagonal form of $A(0)$), and $\widetilde{g}(\lambda)$ has a formal expansion of the form $\widetilde{g}(\lambda) = g_0 + g_1\lambda + \cdots$.
In each angular sector *S of angle slightly bigger than π/n with vertex $\lambda = 0$, there exists a unique true solution $w(\lambda)$ which admits in S an* asymptotic expansion *given by the above formal series.*

In contrast to the previous case $g(\lambda)$ is not analytic at $\lambda = 0$. One can find a formal expansion $g(\lambda) = g_0 + g_1\lambda + \cdots$ by plugging it into the differential equation, but one finds a singular system of linear equations. One can nevertheless obtain formal solutions, but the resulting series is in general *divergent*.

Proof. We shall not give a complete proof of these theorems here (see the References particularly Wasow) but sketch a few of the ideas involved. For the proof it is simpler to assume that the singularity is at infinity so we set $z = 1/\lambda$. We consider the equation $z^{-q}w'(z) = A(z)w(z)$ around the singular point $z = \infty$ of order $(q+1)$. Here we assume $A(z) = A_0 + 1/zA_1 + \cdots$. We make the further simplifying assumption that all eigenvalues of A_0 are different, hence there is no restriction in assuming A_0 diagonal. We first show that there is a gauge transformation $w(z) = P(z)W(z)$ such that the resulting connection $z^qB(z) = P^{-1}(z)(z^qA(z))P(z) - P^{-1}(z)P'(z)$ is diagonal. The differential equation becomes $W'(z) = z^qB(z)W(z)$ which is readily solved yielding the fundamental system of solutions of our equation:

$$w(z) = P(z)e^{\int^z \zeta^q B(\zeta)d\zeta}$$

Note that this is of the form eq.(7.7).

So we need only to find one solution for P obeying the equation:

$$z^{-q}P'(z) = -P(z)B(z) + A(z)P(z) \tag{7.8}$$

Of course the transformation $P(z)$ is defined up to multiplication on the right by a diagonal matrix, since this would leave the connection B diagonal. We fix this ambiguity by requiring $P_{kk}(z) = 1$. Taking the diagonal element kk of eq.(7.8) one gets $B_{kk} = (A(z)P(z))_{kk}$. Plugging this into eq.(7.8) one gets a non–linear system for the $n(n-1)$ non–diagonal elements P_{ij} with $i \neq j$:

$$z^{-q}P'_{ij}(z) = (A(z)P(z))_{ij} - P_{ij}(A(z)P(z))_{jj} \tag{7.9}$$

We first show that this equation admits a formal solution of the form:

$$P(z) = 1 + \sum_{r\geq 1} \frac{1}{z^r} P_r, \quad (P_r)_{kk} = 0, \quad B(z) = A_0 + \sum_{r\geq 1} \frac{1}{z^r} B_r$$

Inserting into eq.(7.8), the coefficient of $1/z^r$ gives: $A_0 P_r - P_r A_0 = B_r + H_r$ for $r \geq 1$, where H_r depends on P_j, B_j for $j < r$. This is recursively solved by setting $(B_r)_{kk} = -(H_r)_{kk}$ (which reproduces the solution $B_{kk} = (AP)_{kk}$) and $(P_r)_{ij} = 1/(\lambda_i - \lambda_j)(H_r)_{ij}$ for $i \neq j$, where the λ_i are the diagonal elements of A_0. This formal solution is in general divergent and knows nothing about sectors.

We then show that there exists a true solution in the sector S having the above formal solution as asymptotics when $z \to \infty$. Equation (7.9) takes the form $z^{-q} P'(z) = f_0(z) + f_1(z,P) + f_2(z,P)$ where $f_1(z,P)$ is linear in P, and $f_2(z,P)$ is quadratic in P. Moreover when $z \to \infty$ $f_1(z,P) = [A_0, P] + \cdots$ hence this linear application is not singular. Finally we have a formal solution $\sum_{r\geq 1} P_r/z^r$ of this equation. It is known that there exists in the interior of any sector of angle less than 2π an analytic function $\Phi(z)$ with the given asymptotics $\sum_{r\geq 1} P_r/z^r$ when $z \to \infty$. We set $P(z) = U(z) + \Phi(z)$ and get a transformed equation for $U(z)$ of the same form $z^{-q} U'(z) = g_0(z) + g_1(z, U(z)) + g_2(z, U(z))$ but now $g_0(z)$ is asymptotic to 0 when $z \to \infty$ and the leading term when $z \to \infty$ of $g_1(z, U(z))$ is of the form $[A_0, U(z)]$. Finally our equation can be written:

$$z^{-q} U'(z) = [A_0, U(z)] + R(z, U)$$

where $R(z,U)$ is much smaller than U when $z \to \infty$ since U will be found to be asymptotic to 0. Hence the term $R(z,U)$ will be treated as a perturbation. The unperturbed equation is readily solved and has the fundamental solution:

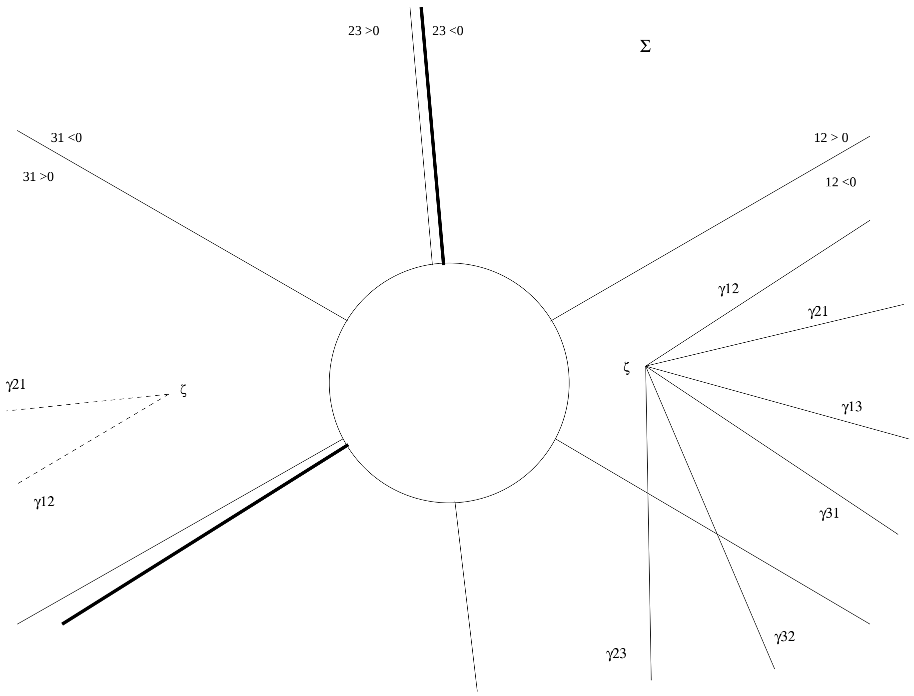
$$V(z) = e^{\frac{z^{q+1}}{q+1} A_0}\, V_0\, e^{-\frac{z^{q+1}}{q+1} A_0}$$

We now replace the differential equation by a system of coupled integral equations, here written in components:

$$U_{ij}(z) = \int_{\gamma_{ij}} e^{\frac{z^{q+1} - t^{q+1}}{q+1}(\lambda_i - \lambda_j)} t^q R_{ij}(t, U(t))\, dt, \quad i \neq j$$

where we allow the integration path γ_{ij} ending at z to depend on ij. The origins of the paths γ_{ij} represent the $N(N-1)$ integration constants of the problem. We need to specify the integration paths, and it is here that the sector S appears.

To be able to control the exponentials in the kernel of the integral equation one chooses paths γ_{ij} such that $\mathrm{Re}\,(z^{q+1} - t^{q+1})(\lambda_i - \lambda_j) < 0$. We examine now the conditions under which such a choice is possible. We use the variables $\zeta = z^{q+1}$ and $\tau = t^{q+1}$. In the τ plane we draw the lines $\mathrm{Re}\,\tau(\lambda_i - \lambda_j) = 0$ called Stokes lines. We shall here only describe the largest sectors in which we can bound the exponentials. Let Σ be a sector of angle slightly larger than π, such that each Stokes line intersects the interior of Σ on just one half–line. The sector S will be its image of angle slightly larger than $\pi/(q+1)$. Hence there exists a region in Σ, where $\mathrm{Re}\,\tau(\lambda_i - \lambda_j) > 0$. Then γ_{ij} will be taken as a straight line path from ∞ to ζ with appropriate slope achieving the inequality $\mathrm{Re}\,(\zeta - \tau)(\lambda_i - \lambda_j) < 0$.
The origins of the paths is chosen at ∞ so that $U_{ij}(z) \to 0$ when $z \to \infty$. Hence the solution $U_{ij}(z)$ is uniquely determined by this requirement, which subsequently yields uniqueness of the solution $w(\lambda)$ having the considered asymptotic expansion in the sector S.
Note that the paths occur in pairs γ_{ij} and γ_{ji} one on each side of the line $\mathrm{Re}\,(\zeta - \tau)(\lambda_i - \lambda_j) = 0$. Moreover we keep the slopes fixed when ζ varies in Σ, and this can be done through the whole sector Σ bounded by some Stokes lines. Note however that if ζ goes beyond the boundary, one of the pair of paths has to be modified (see the figure) yielding a different value of the integral.



Here the largest sector Σ is bounded by the Stokes rays, i.e. half Stoke lines, labelled 23 and 12, and when the point ζ crosses the boundary 12, the pair of paths γ_{12} and γ_{21} has to be modified as indicated by the dashed lines. Notice that this new path cannot be continuously deformed into the previous one. It is clear that the sector Σ is of angle greater than π, so that returning to the original variable z, one gets a sector S of size $\pi/(q+1)+\delta$ for some small positive δ.

Hence it is clear that for τ on the paths one can write:

$$\left| e^{\frac{z^{q+1}-t^{q+1}}{q+1}(\lambda_i-\lambda_j)} \right| \leq e^{-\alpha|x^{q+1}-t^{q+1}|}$$

for some positive fixed constant α (recall that $\lambda_i \neq \lambda_j$ for $i \neq j$).

From this point one can solve the integral equation by successive approximations and prove that this yields a convergent series U asymptotic to zero. We will not reproduce this analysis here. ∎

In general the true solution $w_1(\lambda)$ described in the theorem for the sector S_1, with asymptotic expansion $\widehat{w}(\lambda)$ here, can be analytically continued beyond the sector S_1, but it is very important to realize that its asymptotic expansion will be *different* there. If we consider another sector S_2 adjacent to S_1, there exists another true solution w_2 which has in S_2 the given asymptotic expansion $\widehat{w}(\lambda)$. Since S_1 and S_2 overlap there is a constant matrix $\mathcal{S}_1$ such that in this overlap $w_2 = w_1\mathcal{S}_1$. This relation remains true in $S_1 \cup S_2$ by analytic continuation. This phenomenon was first noticed by Stokes and the matrix $\mathcal{S}_1$ is called the matrix of Stokes multipliers.

Note that this Stokes matrix $\mathcal{S}_1$ is triangular up to some ordering of columns. Indeed in the smallest sector bounded by Stokes rays containing the overlap $S_1 \cap S_2$, one can order the columns such that for example $\mathrm{Re}\,(\lambda_1\tau) < \mathrm{Re}\,(\lambda_2\tau) < \cdots < \mathrm{Re}\,(\lambda_n\tau)$. Since w_1 and w_2 have the same asymptotics in the overlap, we have necessarily $w_2 = w_1\mathcal{S}_1$ where $\mathcal{S}_1$ is lower triangular. This is because $\exp(-\lambda_i\tau)$ is asymptotically negligible with respect to $\exp(-\lambda_j\tau)$ when $i > j$. Moreover the diagonal elements of $\mathcal{S}_1$ are equal to 1, as implied by the given asymptotics. Altogether the Stokes matrix $\mathcal{S}_1$ depends on $N(N-1)/2$ parameters.

The monodromy matrix around $\lambda = 0$ is the matrix $\mathcal{M}$ such that $w_1(e^{2i\pi}\lambda) = w_1(\lambda)\mathcal{M}$. We can easily relate the matrix $\mathcal{M}$ to the Stokes multipliers. Let us cover a neighborhood of the plane at $\lambda = 0$ by $2n$

sectors of angle $\pi/n+\delta$, denoted $S_1, S_2, \cdots, S_{2n}$. More precisely the sector S_j is defined by $(j-1)\pi/n \leq \arg(\lambda) \leq j\pi/n+\delta$. First in the sector S_2, we have $w_1(z) = w_2(z)\mathcal{S}_1^{-1}$ (where w_2 has also the asymptotic $\widehat{w}$) since this is true on the overlap. By recursion on the sector S_j we have $w_1(\lambda) = w_j(\lambda)\mathcal{S}_{j-1}^{-1}\cdots\mathcal{S}_1^{-1}$ for $\lambda \in S_j$. Making a complete 2π rotation around $\lambda=0$ we get a sector S_{2n+1} which projects over S_1, and we have on this sector $w_1(e^{2i\pi}\lambda) = w_{2n+1}(e^{2i\pi}\lambda)\mathcal{S}_{2n}^{-1}\cdots\mathcal{S}_1^{-1}$. But the asymptotic expansion of $w_{2n+1}(e^{2i\pi}\lambda)$ is by definition $\widehat{w}(e^{2i\pi}\lambda) = \widehat{w}(\lambda)\exp(-2i\pi B_0)$, so we see that:

$$w_1(e^{2i\pi}\lambda) \simeq \widehat{w}(\lambda)\, e^{-2i\pi B_0}\, \mathcal{S}_{2n}^{-1}\cdots\mathcal{S}_1^{-1}$$

By comparing the asymptotic expansions we have:

$$\mathcal{M} = \exp(-2i\pi B_0)\,\mathcal{S}_{2n}^{-1}\cdots\mathcal{S}_1^{-1}$$

We now consider the case of several singular points λ_k including the point ∞. Hence we consider the differential equation $\partial_\lambda w = -M_\lambda w$ where $M_\lambda(\lambda)$ is a general rational function of λ, see eq.(7.4). The polar parts $M_\lambda^{(k)}$ are readily obtained from this expression but some care must be taken concerning the polar part at ∞. To *define* $M_\lambda^{(\infty)}$ one chooses a local parameter $z=1/\lambda$ and one considers the polar part at 0. Changing variables in the differential equation one gets:

$$M_\lambda^{(\infty)} \equiv -\left(\frac{1}{z^2}M_\lambda(\frac{1}{z})\right)_- = A_{n_\infty-1}^{(\infty)}z^{-n_\infty-1}+\cdots+A_0^{(\infty)}z^{-2}-\sum_k A_1^{(k)}z^{-1} \tag{7.10}$$

To describe the monodromy data of this equation, we have to patch the local descriptions around each singularity. Around each singularity λ_k we have a formal solution

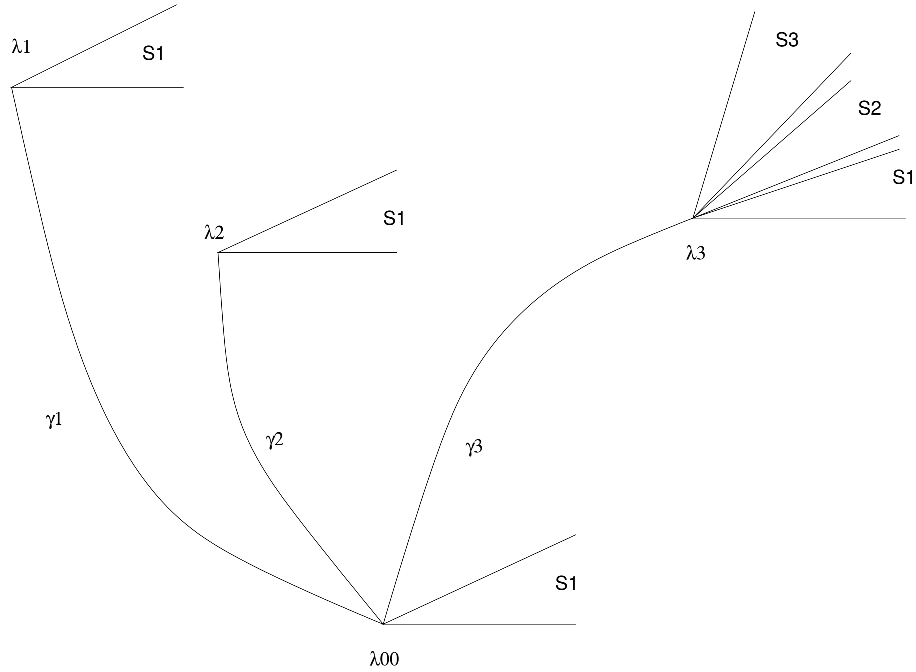
$$\widehat{w}^{(k)}(\lambda) = g^{(k)}(\lambda)\, e^{-\xi^{(k)}(\lambda)}, \quad g^{(k)}(\lambda) = g_0^{(k)}\Big(1+g_1^{(k)}(\lambda-\lambda_k)+\cdots\Big)$$

Since this obeys the differential equation one can readily identify the polar part of M_λ at λ_k, getting:

$$M_\lambda^{(k)} = \Big(\widehat{w}^{(k)}(\lambda)\partial_\lambda\xi^{(k)}\widehat{w}^{(k)\,-1}(\lambda)\Big)_- \tag{7.11}$$

Some care must be taken concerning $M_\lambda^{(\infty)}$. Changing variables to $z=1/\lambda$ we have:

$$M_\lambda^{(\infty)} = -\Big(\widehat{w}^{(\infty)}\frac{1}{z^2}\partial_z\xi^{(\infty)}\widehat{w}^{(\infty)\,-1}\Big)_- \tag{7.12}$$



To describe the global monodromy problem, we fix paths γ_k from a reference point which we choose to be ∞ to each λ_k. Around each λ_k there are sectors $S_j^{(k)}$ and corresponding solutions $w_j^{(k)}$ with the given asymptotics. Starting from $w_1^{(\infty)}$ we want to describe how it changes when analytically continued around singularities. We can continue $w_1^{(\infty)}$ along the path γ_k and compare the result to $w_1^{(k)}$. This defines matrices $\mathcal{C}^{(k)}$ such that $w_1^{(\infty)}(\lambda) = w_1^{(k)}(\lambda)\mathcal{C}^{(k)}$. Then the monodromy matrix of $w_1^{(\infty)}$ around λ_k is

$$\mathcal{M}^{(k)} = \mathcal{C}^{(k)\,-1}\, e^{-2i\pi B_0^{(k)}}\, \mathcal{S}_{2n}^{(k)\,-1} \cdots \mathcal{S}_1^{(k)\,-1}\mathcal{C}^{(k)} \tag{7.13}$$

Around $\lambda = \infty$ we have the same formula with $\mathcal{C}^{(\infty)} = 1$. The monodromy group is generated by these monodromy matrices, subject to the relation

$$\mathcal{M}^{(\infty)} \cdot \mathcal{M}^{(1)} \cdots \mathcal{M}^{(K)} = 1 \tag{7.14}$$

since the path around all these singularities is contractible.

Note that the determinant of each Stokes matrix is equal to 1 since there exists a permutation of the basis in which it is triangular with 1 on the diagonal. It follows that $\det(\mathcal{M}^{(k)}) = \exp(-2i\pi \mathrm{Tr}\, B_0^{(k)})$. Taking the determinant of eq.(7.14) we see that $\sum_k \mathrm{Tr}\, B_0^{(k)}$ is an integer, where the sum includes ∞. In fact there is a stronger compatibility condition for the existence of $w(\lambda)$ called the Fuchs condition, saying that this integer vanishes.

$$\sum_k \mathrm{Tr}\,(B_0^{(k)}) + \mathrm{Tr}\,(B_0^{(\infty)}) = 0 \tag{7.15}$$

To show this we take the trace of eq.(7.12) and we get:

$$\mathrm{Tr}\, M_\lambda^{(\infty)} = -\left(\mathrm{Tr}\, \frac{1}{z^2}\partial_z \xi^{(\infty)}\right)_- = \mathrm{Tr}\, B_0^{(\infty)} \frac{1}{z} + \cdots$$

Similarly at each singularity λ_k:

$$\mathrm{Tr}\, M_\lambda^{(k)} = \left(\mathrm{Tr}\, \partial_\lambda \xi^{(k)}\right)_- = \mathrm{Tr}\, B_0^{(k)} \frac{1}{\lambda - \lambda_k} + \cdots$$

Comparing with eq.(7.10) we get the Fuchs relation.

We want to study the correspondence between the coefficients of $M_\lambda(\lambda)$ and the monodromy data.

Definition. *We call the monodromy data the set of variables defined at each singular point λ_k including ∞:*

The $(2n_k)$ Stokes matrices $\mathcal{S}_j^{(k)}$

The connection matrices $\mathcal{C}^{(k)}$

Moreover we call singularity data at λ_k the coefficients of the singular terms $\xi^{(k)}$.

It is important to note that given M_λ the monodromy data are defined only up to a group of diagonal matrices at each λ_k. Indeed we have seen that these quantities are uniquely defined once we have chosen an asymptotic expansion around λ_k. More precisely once the most singular part of M_k has been diagonalized, there is a canonical asymptotic expansion of the form $\widehat{w}^{(k)}(\lambda) = (1 + O(\lambda - \lambda_k)) \exp(\xi^{(k)}(\lambda))$. Conjugating the matrix M_λ by a constant matrix one can assume that the most singular term at ∞ is diagonal. Then there is a canonical choice of $w_1^{(\infty)}$. At the other singular points however one has simply $\widehat{w}^{(k)}(\lambda) = g_0^{(k)}(1 + O(\lambda - \lambda_k)) \exp(\xi^{(k)}(\lambda))$ where $g_0^{(k)}$ is the matrix diagonalizing the leading singularity in M_k. It is obviously defined only up to multiplication on the right by a diagonal matrix $d^{(k)}$. Hence the $w_1^{(k)}$ are defined only up to right multiplication by $d^{(k)}$. This changes $\xi^{(k)} \to \xi^{(k)}$, $\mathcal{S}_j^{(k)} \to d^{(k)-1}\mathcal{S}_j^{(k)} d^{(k)}$, and $\mathcal{C}^{(k)} \to \mathcal{C}^{(k)} d^{(k)}$.

We are now in a position to compare the number of parameters in the matrix M_λ and in the reduced monodromy and singularity data set. Assuming that the leading singular coefficient at ∞ is diagonal, the matrix M_λ depends on $N^2(n_\infty + \sum_{k=1}^K n_k + K - 1) + N$ parameters. On the other hand each Stokes matrix depends on $N(N-1)/2$ parameters, and there are $2n_k$ such matrices at each singular point including ∞. Similarly each $\xi^{(k)}$ depends on $(n_k + 1)N$ parameters, while the matrix $\mathcal{C}^{(k)}$ (for $k = 1, \cdots, K$) depends on N^2 parameters. Altogether the set of monodromy and singularity data contains $N^2(n_\infty + \sum_{k=1}^K n_k + K) + N(K+1)$ parameters. These parameters are not all independent, since we have to take into account the relation eq.(7.14) between the monodromy matrices which removes N^2 parameters, and we have to quotient by the diagonal group action at each λ_k which removes NK parameters. One then gets exactly the number of free parameters in M_λ.

It is then reasonable to expect that these two sets of data are equivalent, that is to say, for any monodromy and singularity data satisfying the appropriate consistency conditions, one can find a unique differential equation of the above form with the given monodromy and singularity data. This is called the generalized Riemann problem, which was first studied and virtually solved by Birkhoff and more recently by Malgrange and Sibuya. Indeed the unicity is easy. Assume we have two differential equations with the same monodromy and singularity data, and consider their solutions $w_1(\lambda)$ and $w_2(\lambda)$. We can assume that they are canonically normalized at ∞. Let us consider $P(\lambda) = w_1(\lambda) w_2^{-1}(\lambda)$. We see that $P(\lambda)$ is single valued. The singular parts cancel at λ_k hence $P(\lambda)$ is holomorphic at λ_k. Since $P(\lambda = \infty) = 1$, we get $P(\lambda) = 1$. The same argument applied to $\frac{\partial w(\lambda)}{\partial \lambda} w^{-1}(\lambda)$ shows that M_λ is a rational function with a pole of order $n_k + 1$ at λ_k.

This characterization of the differential equation by its monodromy and singularity data is the subject of the following section.

7.3 Isomonodromy deformations.

As we have seen in the previous section, the matrix $w(\lambda)$ is determined by its monodromy data, and its singularity data $\xi^{(k)}$ at λ_k. These are two independent sets of quantities which can be specified arbitrarily. We will examine here the matrices $w(\lambda)$, with prescribed monodromy data $S_j^{(k)}, C^{(k)}$ and fixed parameters $B_0^{(k)}$, and varying the parameters t_i and λ_k.

We will show that $w(\lambda, \{t_i\}, \{\lambda_k\})$ satisfies a hierarchy of equations, eq.(7.1) with respect to the times t_i, and new equations, called Schlesinger's equations, with respect to the λ_k's. Notice that these are all the possible deformation parameters, since once they are fixed, together with the monodromy data, they determine the function $w(\lambda)$ uniquely.

Assume that we are given the positions λ_k of singularities, some Stokes matrices $\mathcal{S}_j^{(k)}$ at these singularities and some connection matrices $\mathcal{C}^{(k)}$. Then for any values of the times t_i one can find a matrix $M_\lambda(\lambda)$ of the form eq.(7.4) such that the differential equation $\partial_\lambda w = -M_\lambda w$ has a solution $w(\lambda)$ with the asymptotic expansion of the form $\widehat{w}^{(\infty)}(\lambda, t) = (1 + O(1/\lambda))\, e^{-\xi^{(\infty)}}$ in the first sector at ∞, and whose analytic continuation in the λ–plane has asymptotic expansions in the sector $S_j^{(k)}$ at λ_k of the form:

$$w(\lambda) \sim \widehat{w}^{(k)}(\lambda)\mathcal{S}_j^{(k)\,-1} \cdots \mathcal{S}_1^{(k)\,-1}\mathcal{C}^{(k)} \tag{7.16}$$

where

$$\widehat{w}^{(k)}(\lambda) = g_0^{(k)}(1 + g_1^{(k)}(\lambda - \lambda_k) + \cdots)\, e^{-\xi^{(k)}(\lambda,t)} \tag{7.17}$$

Denote by d the exterior differentiation with respect to the parameters t_i and λ_k.

Theorem. *The monodromy data for $w(\lambda)$ are independent of the deformation parameters if and only if the function $w(\lambda)$ satisfies differential equations with respect to the deformation parameters:*

$$dw = -Mw \tag{7.18}$$

where M is a rational *function of λ with values in one–forms.*

Proof. Let us consider the one–form $dw(\lambda)w^{-1}(\lambda)$, as a function of λ. From the assumption that the monodromy data are independent of the parameters, this one–form is single valued around λ_k because when we turn around λ_k the matrix $w(\lambda)$ gets multiplied by the monodromy matrix eq.(7.13), which is independent of the deformation parameters by assumption, hence cancels in $dw\, w^{-1}$. Moreover, in the vicinity of λ_k its asymptotic expansion can be computed in any sector S_j by inserting the asymptotic expansion eq.(7.16), yielding $dw\, w^{-1} \sim d\widehat{w}^{(k)}\, \widehat{w}^{(k)\,-1}$, where we use again the independence of the monodromy data from the deformation parameters, and the known fact that one can differentiate an asymptotic expansion valid in a sector. From the explicit form of $\widehat{w}^{(k)}$ we see that the singularity of the asymptotic expansion is polar at λ_k. Since this is true for all overlapping $2n$ sectors around λ_k it follows that the same is true for the function $dw\, w^{-1}$. More explicitly $\partial_{t_i} w(\lambda)\, w^{-1}(\lambda)$ has a pole of order n_k at λ_k, and $\partial_{\lambda_k} w(\lambda)w^{-1}(\lambda)$ has a pole of order $n_k + 1$ at λ_k. Hence $M(\lambda)$ is a rational function of λ.
Conversely assume that $M_\lambda(\lambda)$ is parametrized by some parameters τ_i and that one can write equations $\partial_{\tau_i} w = -M_i(\lambda)w$ where $M_i(\lambda)$ are rational functions of λ. Considering for example two adjacent sectors S_1, S_2 at a singularity, the solution $w(\lambda)$ has the asymptotic $\widehat{w}$ in S_1 and $\widehat{w}\mathcal{S}$ in S_2. Then $\partial_i w$ has the asymptotic $\partial_i \widehat{w} = -M_i \widehat{w}$ in S_1 (since M_i is rational). This relation on $\widehat{w}$ remains true in all sectors. In S_2 $\partial_i w$ has the asymptotic $\partial_i \widehat{w}\mathcal{S} + \widehat{w}\partial_i \mathcal{S}$ but this is also $-M_i w$ which has the asymptotic $-M_i \widehat{w}\mathcal{S}$. Hence $\partial_i \mathcal{S} = 0$ since $\mathcal{S}$ does not depend on λ. Similarly one shows that the connection matrices $\mathcal{C}^{(k)}$ are time–independent. ∎

So we are bound to study the system of equations

$$\partial_\lambda w = -M_\lambda w, \quad \partial_{\tau_i} w = -M_i(\lambda)w$$

with M_λ of the form eq.(7.4) and M_i rational matrices. This system of equations is compatible if the rational matrices M_λ, M_i obey the zero–curvature conditions:

$$\partial_{\tau_i} M_\lambda = \partial_\lambda M_i + [M_\lambda, M_i] \tag{7.19}$$

$$\partial_{\tau_i} M_j - \partial_{\tau_j} M_i + [M_i, M_j] = 0 \tag{7.20}$$

In this case the Frobenius theorem asserts that one can find solutions $w(\lambda)$ depending on the maximal number of parameters. In our situation the maximal number of isomonodromic parameters τ_i and compatible equations that one can introduce is just the set of times t_i and the λ_k.

We want to prove directly the compatibility of the system eqs.(7.19,7.20), with properly defined M_i's, yielding the existence of $w(\lambda;\tau_i)$. We start from the rational matrix M_λ and consider the differential equation $(\partial_\lambda + M_\lambda)w = 0$. Around λ_k there exists a formal solution of the form $\widehat{w}^{(k)} = g^{(k)} \exp(-\xi^{(k)})$. Hence $g^{(k)}$ obeys the equation:

$$\partial_\lambda g^{(k)} = -M_\lambda g^{(k)} + g^{(k)} \partial_\lambda \xi^{(k)} \tag{7.21}$$

Next we *define* the rational matrices:

$$M_i(\lambda) = \left(g^{(k)} \partial_{\tau_i} \xi^{(k)} g^{(k)\,-1} \right)_-$$

where τ_i stands for $t_{(k,n,\alpha)}$ and λ_k. From this we *define* some vector fields X_i acting on M_λ by:

$$X_i M_\lambda = \partial_\lambda M_i - [M_i, M_\lambda] \tag{7.22}$$

It is important to notice that the right–hand side of this equation has the same polar structure as $\partial_{\tau_i} M_\lambda$, so that eq.(7.22) actually defines vector fields on the coefficients of these polar expansions. To show that, one has to examine the order of the pole at λ_k in the commutator $[M_\lambda, M_i]$ (see a similar discussion in Chapter [2]). Around λ_k the dangerous term is:

$$[M_\lambda^{(k)}, M_i^{(k)}]_- = [g^{(k)} \partial_\lambda \xi^{(k)} g^{(k)\,-1} - \left(g^{(k)} \partial_\lambda \xi^{(k)} g^{(k)\,-1} \right)_+, g^{(k)} d\xi^{(k)} g^{(k)\,-1} - \left(g^{(k)} d\xi^{(k)} g^{(k)\,-1} \right)_-]_-$$

Next we remark that $[g^{(k)} \partial_\lambda \xi^{(k)} g^{(k)\,-1}, g^{(k)} d\xi^{(k)} g^{(k)\,-1}] = 0$, and the remaining terms are of the required order. This allows to consider the flows X_i as acting on the coefficients of M_λ and the λ_k's. In particular X_i and ∂_λ commute.

Theorem. *The flows X_i are all commuting, moreover we can identify $X_i = \partial_{\tau_i}$.*

Proof. The commutation of the flows X_i, X_j is expressed by $\partial_\lambda F_{ij} + [M_\lambda, F_{ij}] = 0$, where:

$$F_{ij} = X_i M_j - X_j M_i + [M_i, M_j]$$

We show in fact a stronger result, i.e. $F_{ij} = 0$. Our first task is to express the action of the flow X_i on the variables $g^{(k)}$ and $\xi^{(k)}$ around any pole λ_k. To do that we apply X_i to eq.(7.21) we get:

$$\partial_\lambda (X_i g^{(k)} + M_i g^{(k)}) = -M_\lambda (X_i g^{(k)} + M_i g^{(k)}) + (X_i g^{(k)} + M_i g^{(k)}) \partial_\lambda \xi^{(k)} + g^{(k)} \partial_\lambda X_i \xi^{(k)}$$

This equation is of the same form as eq.(7.21) so that writing $(X_i g^{(k)} + M_i g^{(k)}) = g^{(k)} h_i$ we get for h_i the equation:

$$\partial_\lambda h_i + [\partial_\lambda \xi^{(k)}, h_i] = \partial_\lambda X_i \xi^{(k)}$$

By definition $X_i \xi^{(k)}$ is diagonal. This is a linear equation for h_i. A particular solution is $h_i = X_i \xi^{(k)}$. The general solution is obtained by adding to it $e^{-\xi^{(k)}} D_i e^{\xi^{(k)}}$ for any matrix D_i independent of λ. Through its definition, we see that h_i has at most poles at λ_k hence D_i must be diagonal. Finally we get:

$$X_i g^{(k)} = -M_i g^{(k)} + g^{(k)} X_i \xi^{(k)} + g^{(k)} D_i$$

Looking at the polar part of this equation we see that $(g^{(k)} X_i \xi^{(k)} g^{(k)\,-1})_- = (g^{(k)} \partial_{\tau_i} \xi^{(k)} g^{(k)\,-1})_-$ from which it follows that $X_i \xi^{(k)} = \partial_{\tau_i} \xi^{(k)}$, assuming that $g_0^{(k)}$ is generic. Hence X_i identifies to ∂_{τ_i} on $\xi^{(k)}$.
We can now compute F_{ij}. It is simpler to use a compact notation: let M be the 1–form $\sum M_i d\tau_i$ and δ be

the differential vector field $\sum X_i d\tau_i$. Note that on $\xi^{(k)}$, δ identifies to d so that $\delta^2\xi^{(k)} = 0$. The conditions $F_{ij} = 0$ read $\delta M + M \wedge M = 0$, and the equation on $g^{(k)}$ becomes with $D = \sum D_i d\tau_i$:

$$\delta g^{(k)} = -Mg^{(k)} + g^{(k)}\delta\xi^{(k)} + g^{(k)}D$$

With the help of this equation one can compute the polar part of δM at λ_k:

$$\delta M^{(k)} = [g^{(k)}(\delta\xi^{(k)} - D)g^{(k)\,-1}, g^{(k)}\delta\xi^{(k)}g^{(k)\,-1}]_- - [M, g^{(k)}d\xi^{(k)}g^{(k)\,-1}]_-$$

The first term vanishes because it involves commutators of diagonal matrices. To evaluate the second term we remark that $M - g^{(k)}d\xi^{(k)}g^{(k)\,-1} = O(1)$ hence, squaring it, we get $M \wedge M - [M, g^{(k)}d\xi^{(k)}g^{(k)\,-1}] = O(1)$ since again the commutator of diagonal matrices vanishes. Taking the polar part of this relation, we conclude that $dM^{(k)} = -(M \wedge M)_-$. Since this is true at each pole λ_k and since M is a rational function of λ we see that $F_{ij} = 0$ for all i, j.
Since we have shown that $[X_i, X_j] = 0$ and knowing that $X_i = \partial_{\tau_i}$ on the $\xi^{(k)}$, the Frobenius theorem implies that one can simultaneously solve (locally) $\partial_{\tau_i} M_\lambda = X_i M_\lambda$ hereby obtaining rational matrices M_λ, M_i satisfying eqs.(7.19,7.20). ■

Example: The Schlesinger equations.
Consider a differential equation $\partial_\lambda w = -M_\lambda w$ where M_λ has only *regular* singularities, i.e.

$$M_\lambda(\lambda) = \sum_k \frac{A^{(k)}}{\lambda - \lambda_k}, \quad A^{(k)} = g_0^{(k)} B_0^{(k)} g_0^{(k)\,-1}$$

where we assume, according to the general analysis, that the matrices $A^{(k)}$ are diagonalizable. Note that there is a hidden regular singularity at ∞. The asymptotic expansions of $w(\lambda)$ at the λ_k are easily computed and found to be of the form:

$$\widehat{w}^{(k)}(\lambda) = g_0^{(k)}(1 + O(\lambda - \lambda_k))\, e^{-\xi^{(k)}(\lambda)}, \quad \xi^{(k)}(\lambda) = B_0^{(k)} \log{(\lambda - \lambda_k)}$$

This means that the times t_i are all set to 0 and that the only deformation parameters are the λ_k. The deformation equation $\partial_{\lambda_k} w = -M_{\lambda_k} w$ are constructed with the help of the general formula:

$$M_{\lambda_k} = \left(\widehat{w}^{(k)}\partial_{\lambda_k}\xi^{(k)}\widehat{w}^{(k)\,-1}\right)_- = -\frac{A^{(k)}}{\lambda - \lambda_k}$$

The zero curvature conditions read:

$$\begin{aligned} \partial_{\lambda_l} A^{(k)} &= -\frac{[A^{(k)}, A^{(l)}]}{\lambda_k - \lambda_l}, \quad l \neq k \\ \partial_{\lambda_k} A^{(k)} &= \sum_{l \neq k} \frac{[A^{(k)}, A^{(l)}]}{\lambda_k - \lambda_l} \end{aligned}$$

These equations are called Schlesinger equations. One can check easily that both these equations are contained in eq.(7.19), while eq.(7.20) is a direct consequence of them.

7.4 Isomonodromy and Riemann–Hilbert problem.

7.5 Schlesinger transformations

We have studied all continuous isomonodromic deformations, they are parametrized by the times t_i and the λ_k. However there remains discrete isomonodromic deformations. The basic remark is that although the matrices $B_0^{(k)}$ are not allowed to change continuously, a discrete change $B_0^{(k)} \to B_0^{(k)} + L^{(k)}$ where $L^{(k)}$ is a diagonal matrix with integer entries does not change the monodromy data, while the singularity data are modified by $e^{-\xi^{(k)}(\lambda)} \to (\lambda - \lambda_k)^{-L^{(k)}} e^{-\xi^{(k)}(\lambda)}$, i.e. we add extra zeroes or poles at the singularities. This is a discrete analogue to deformation of the t_i. These transformations of the singularity data are called Schlesinger transformations. For the continuous isomonodromic transformations we have deformation equations $dw = -Mw$ with M a rational matrix. We show that the same is true for discrete transformations. This means that one can consider that $w(\lambda)$ depends not only on the continuous variables t_i and λ_k but also on a set of integers, (an integer diagonal matrix above each singularity), and obeys difference equations with respect to these integers.

Proposition. *The two wave–functions $w(\lambda)$ associated with the data $B_0^{(k)}$ and $w'(\lambda)$ associated to the data $B_0^{(k)} + L^{(k)}$ are related by:*

$$w'(\lambda) = R(\lambda) w(\lambda) \tag{7.23}$$

where the matrix $R(\lambda)$ is a rational function of λ.

Proof. Consider the matrix $w'(\lambda)w^{-1}(\lambda)$. The essential singularities cancel and it has no monodromy. Hence, it is a rational function. ∎

In fact the integer matrices $L^{(k)}$ are restricted due to the Fuchs condition, eq.(7.15):

$$\sum_k \mathrm{Tr}\, L^{(k)} + \mathrm{Tr}\, L^{(\infty)} = 0$$

To study Schlesinger's tranformations, it is enough to concentrate on elementary ones:

Definition. *An elementary Schlesinger transformation is associated to the matrices*

$$L^{(l)} \begin{bmatrix} k & k' \\ \alpha & \alpha' \end{bmatrix} = \delta_{kl} E_{\alpha\alpha} - \delta_{k'l} E_{\alpha'\alpha'}$$

In other words we shift the α^{th} diagonal element by $+1$ above λ_k and the α'^{th} diagonal element by -1 above $\lambda_{k'}$ in order to fulfill the Fuchs condition $\sum_l \mathrm{Tr}(L^{(l)}) = 0$.

Proposition. *The matrix $R(\lambda)$ associated to the above elementary Schlesinger transformation is given by*

$$R(\lambda) = 1 + \frac{R_0}{\lambda - \lambda_k} \quad R^{-1}(\lambda) = 1 - \frac{R_0}{\lambda - \lambda_{k'}} \tag{7.24}$$

where the matrix R_0 is given by:

$$\begin{aligned} R_0 &= \frac{\lambda_k - \lambda_{k'}}{\left(g_0^{(k)-1} g_0^{(k')}\right)_{\alpha\alpha'}} \, g_0^{(k')} E_{\alpha'\alpha} g_0^{(k)-1}, \quad \text{if } k \neq k' \\ R_0 &= -\frac{1}{\left(g_1^{(k)}\right)_{\alpha\alpha'}} \, g_0^{(k)} E_{\alpha'\alpha} g_0^{(k)-1}, \quad \text{if } k = k' \end{aligned} \tag{7.25}$$

using the asymptotic expansion coefficients eq.(7.17).

Proof. The conditions determining $R(\lambda)$ are

$$R(\lambda) g^{(l)}(\lambda) = (g')^{(l)}(\lambda)(\lambda - \lambda_l)^{-L^{(l)}} \tag{7.26}$$

Suppose first $k \neq k'$. Let us write the condition eq.(7.26) in more details. If $R(\lambda) = 1 + R_0/(\lambda - \lambda_k)$ they read

$$\begin{aligned}\left(1 + \frac{1}{\lambda - \lambda_k} R_0\right) g^{(k)} &= (g')^{(k)} \left(\frac{1}{\lambda - \lambda_k} E_{\alpha\alpha} + (1 - E_{\alpha\alpha})\right)\\ \left(1 + \frac{1}{\lambda - \lambda_k} R_0\right) g^{(k')} &= (g')^{(k')} \Big((\lambda - \lambda_{k'}) E_{\alpha'\alpha'} + (1 - E_{\alpha'\alpha'})\Big)\end{aligned}$$

Projecting the first equation on $|\beta>, \beta \neq \alpha$, gives $R_0 g_0^{(k)} |\beta >= 0$. The solution of this equation is $R_0 = |X >< \alpha| g_0^{(k)-1}$. Similarly, projecting the second equation at $\lambda_{k'}$ on $|\alpha'>$ yields

$$g_0^{(k')} |\alpha' > + \frac{R_0}{\lambda_{k'} - \lambda_k} g_0^{(k')} |\alpha' >= 0$$

which gives

$$|X >= \frac{\lambda_k - \lambda_{k'}}{< \alpha| g_0^{(k)-1} g_0^{(k')} |\alpha' >} g_0^{(k')} |\alpha' >$$

which yields the first part of the proposition. Suppose next $k = k'$. The equation determining $R(\lambda)$ now reads

$$\left(1 + \frac{1}{\lambda - \lambda_k} R_0\right) g^{(k)} = (g')^{(k)} \left(\frac{1}{\lambda - \lambda_k} E_{\alpha\alpha} + (\lambda - \lambda_k) E_{\alpha'\alpha'} + 1 - E_{\alpha\alpha} - E_{\alpha'\alpha'}\right)$$

Comparing the terms of order $(\lambda - \lambda_k)^{-1}$ and projecting them on the vector $|\beta >, \beta \neq \alpha$, again yields $R_0 g_0^{(k)} |\beta >= 0$ so that $R_0 = |X >< \alpha| g_0^{(k)-1}$. Next considering the terms of order $(\lambda - \lambda_k)^0$, and projecting them on $|\alpha'>$ gives $g_0^{(k)} |\alpha' > + |X >< \alpha| g_1^{(k)} |\alpha' >= 0$, or

$$|X >= -\frac{1}{< \alpha| g_1^{(k)} |\alpha' >} g_0^{(k)} |\alpha' >$$

which yields the second part of the proposition. Since R_0 is a rank one projector, $R_0^2 = (\lambda_k - \lambda_{k'}) R_0$, the inverse of $R(\lambda)$ is readily computed. In particular we have $\det R(\lambda) = (\lambda - \lambda_{k'})/(\lambda - \lambda_k)$. ■

7.6 τ–functions

Consider a differential equation $\partial_\lambda w = -M_\lambda w$ where M_λ is a rational function of λ depending on isomonodromic deformation parameters. At each singularity λ_k we have asymptotic expansions of the form $w(\lambda) \sim g^{(k)}(\lambda)\, e^{-\xi^{(k)}(\lambda)}$. To any solution of the deformation equations, $dw = -Mw$, we can associate a one–form γ:

$$\gamma = \sum_k \mathrm{Res}_{\lambda=\lambda_k} \mathrm{Tr}(g^{(k)-1} \partial_\lambda g^{(k)} d\xi^{(k)}) d\lambda \qquad (7.27)$$

The sum is over all singularities including ∞.

Theorem. *The deformation equations imply that γ is closed: $d\gamma = 0$.*

Proof. We have already proved this equation in a more restricted setting in Chapter [2]. Let us repeat the proof of this important result in this more general context.

$$d\gamma = \sum_k \mathrm{Res}_{\lambda=\lambda_k} \mathrm{Tr}(-g^{(k)-1} dg^{(k)} g^{(k)-1} \partial_\lambda g^{(k)} d\xi^{(k)} + g^{(k)-1} \partial_\lambda dg^{(k)} d\xi^{(k)}) d\lambda$$

From the deformation equation we get:

$$dg^{(k)} = g^{(k)} d\xi^{(k)} - M g^{(k)}$$

so that

$$d\gamma = \sum_k \mathrm{Res}_{\lambda=\lambda_k} \mathrm{Tr}(d\partial_\lambda \xi^{(k)} \wedge d\xi^{(k)} - \partial_\lambda M \wedge g^{(k)} d\xi^{(k)} g^{(k)-1}) d\lambda$$

Now $\mathrm{Res}_{\lambda=\lambda_k} \mathrm{Tr}(d\partial_\lambda \xi^{(k)} \wedge d\xi^{(k)}) = 0$ since the order of the pole is at least 3. Next $M = M^{(k)} + \sum_{l\neq k} M^{(l)}$, $M^{(k)} = (g^{(k)} d\xi^{(k)} g^{(k)-1})_-$, so that we can write $g^{(k)} d\xi^{(k)} g^{(k)-1} = M + X^{(k)}_{reg}$ where $X^{(k)}_{reg}$ is regular at λ_k. On the other hand

$$\mathrm{Res}_{\lambda=\lambda_k} \mathrm{Tr}(\partial_\lambda (g^{(k)} d\xi^{(k)} g^{(k)-1}) \wedge (g^{(k)} d\xi^{(k)} g^{(k)-1})) d\lambda = \mathrm{Res}_{\lambda=\lambda_k} \mathrm{Tr}(d\partial_\lambda \xi^{(k)} \wedge d\xi^{(k)}) d\lambda = 0$$

hence

$$\mathrm{Res}_{\lambda=\lambda_k} \mathrm{Tr}(\partial_\lambda (M + X^{(k)}_{reg}) \wedge (M + X^{(k)}_{reg}) d\lambda = 0$$

Since the residue of a derivative of a function of λ vanishes we get:

$$\mathrm{Res}_{\lambda=\lambda_k} \mathrm{Tr}(\partial_\lambda M \wedge X^{(k)}_{reg}) d\lambda = -\frac{1}{2} \mathrm{Res}_{\lambda=\lambda_k} \mathrm{Tr}(\partial_\lambda M \wedge M) d\lambda$$

and it follows that

$$d\gamma = -\sum_k \mathrm{Res}_{\lambda=\lambda_k} \mathrm{Tr}(\partial_\lambda M \wedge (M + X^{(k)}_{reg}) d\lambda = -\frac{1}{2} \sum_k \mathrm{Res}_{\lambda=\lambda_k} \mathrm{Tr}(\partial_\lambda M \wedge M) d\lambda$$

Now $\mathrm{Tr}(\partial_\lambda M \wedge M) d\lambda$ is a rational one–form on the λ Riemann sphere, hence the sum of the residues vanishes.
∎

Example. Let us give the form γ in the Schlesinger case of regular singularities and deformation parameters λ_k. Since $d\xi = -\sum_k B_0^{(k)}/(\lambda - \lambda_k) d\lambda_k$, we have only to keep the constant term in $g^{(k)\,-1} \partial_\lambda g^{(k)}$ yielding:

$$\gamma = -\sum_k \mathrm{Tr}(g_1^{(k)} B_0^{(k)}) d\lambda_k$$

Starting from $\partial_\lambda w = -\sum_l (g^{(l)} \partial_\lambda \xi^{(l)} g^{(l)\,-1})_- w$ we get:

$$g_0^{(k)} (g_1^{(k)} + [B_0^{(k)}, g_1^{(k)}]) g_0^{(k)-1} = -\sum_{l\neq k} \frac{g_0^{(l)} B_0^{(l)} g_0^{(l)-1}}{\lambda_l - \lambda_k}$$

so that

$$\gamma = -\frac{1}{2} \sum_k \mathrm{Tr}(A_k A_l) \frac{d\lambda_l - d\lambda_k}{\lambda_l - \lambda_k}$$

We can verify that this form is closed using the Schlesinger equations.

By virtue of this result, to each solution of the deformation equations we can associate a τ–function:

Definition. *The τ–function is defined by*

$$\gamma = d \log \tau \tag{7.28}$$

To each solution of the deformation equations eq.(7.18), one can associate a τ–function. Hence to the Schlesinger transformed solution, we can associate a transformed τ–function. There is a simple relation between the original τ–function and its tranform by an elementary Schlesinger transformation:

Proposition.

$$\tau \begin{bmatrix} k & k' \\ \alpha & \alpha' \end{bmatrix} (t) = \tau(t) \begin{cases} -\frac{\left(g_0^{(k)-1} g_0^{(k')}\right)_{\alpha\alpha'}}{\lambda_k - \lambda_{k'}} & \text{if } k \neq k' \\ \left(g_1^{(k)}\right)_{\alpha\alpha'} & \text{if } k = k' \end{cases} \tag{7.29}$$

Proof. Using eq.(7.26) and the relation $\xi'^{(l)} = \xi^{(l)} + L^{(l)} \log(\lambda - \lambda_l)$, we can write:

$$\gamma' - \gamma = \sum_l < R^{-1}\partial_\lambda R g^{(l)} d\xi'^{(l)} g^{(l)-1} > + < g^{(l)-1}\partial_\lambda g^{(l)} d(\xi'^{(l)} - \xi^{(l)}) > + < L^{(l)}(\lambda - \lambda_l)^{-1} d\xi'^{(l)} >$$

The third term vanishes because the pole is of order at least 2. The second term is equal to:

$$< g^{(l)-1}\partial_\lambda g^{(l)} d(\xi'^{(l)} - \xi^{(l)}) >= -\sum_l \mathrm{Tr}(g_1^{(l)} L^{(l)}) d\lambda_l = -\left(g_1^{(k)}\right)_{\alpha\alpha} d\lambda_k + \left(g_1^{(k')}\right)_{\alpha'\alpha'} d\lambda_{k'}$$

Using the explicit form for R and computing $R^{-1}\partial_\lambda R$ one gets for the first term:

$$\sum_l < R^{-1}\partial_\lambda R g^{(l)} d\xi'^{(l)} g^{(l)-1} >= -\sum_l < \frac{1}{(\lambda - \lambda_k)(\lambda - \lambda_{k'})} R_0 g^{(l)} d\xi'^{(l)} g^{(l)-1} >$$

but, if $f(\lambda)$ is a function with an expansion at λ_l of the form $f(\lambda) = \sum_{i=-N}^{\infty} f_i (\lambda - \lambda_l)^i$ we have:

$$\mathrm{Res}_{\lambda_l}\left(\frac{1}{\lambda - \lambda_k} f(\lambda)\right) = \begin{cases} -f_-(\lambda)|_{\lambda=\lambda_k} & \text{if } \lambda_l \neq \lambda_k \\ f_0 & \text{if } \lambda_l = \lambda_k \end{cases}$$

hence looking at the case $k \neq k'$ one finds after some computation:

$$\begin{aligned} \sum_l < R^{-1}\partial_\lambda R g^{(l)} d\xi'^{(l)} g^{(l)-1} > &= \mathrm{Tr}\left(R_0 \frac{-\Theta^{(k)} + g_0^{(k)} g_1^{(k)} g_0^{(k)-1} d\lambda_k + \Theta^{(k')} - g_0^{(k')} g_1^{(k')} g_0^{(k')-1} d\lambda_{k'}}{\lambda_k - \lambda_{k'}} \right) \\ &+ \sum_l < R^{-1}\partial_\lambda R g^{(l)} (d\xi'^{(l)} - d\xi^{(l)}) g^{(l)-1} > \end{aligned}$$

where we have defined $\Theta^{(k)} = dg_0^{(k)} g_0^{(k)\,-1}$ which evaluates to:

$$\Theta^{(k)} = -\sum_{l \neq k} M^{(l)}|_{\lambda=\lambda_k} - \left(M^{(k)}(\lambda) - g^{(k)} d\xi^{(k)} g^{(k)\,-1}\right)|_{\lambda=\lambda_k} + g_0^{(k)} g_1^{(k)} g_0^{(k)-1} d\lambda_k$$

The last term reads:

$$\begin{aligned} \sum_l < R^{-1}\partial_\lambda R g^{(l)} (d\xi'^{(l)} &- d\xi^{(l)}) g^{(l)-1} >= -\sum_l < \frac{1}{(\lambda - \lambda_k)(\lambda - \lambda_{k'})} R_0 g^{(l)} \frac{1}{\lambda - \lambda_l} E^{(l)} g^{(l)-1} > d\lambda_l \\ = \frac{1}{\lambda_k - \lambda_{k'}} \Big\{ &- \sum_{l \neq k} \mathrm{Tr}(R_0 g_0^{(l)} \frac{E^{(l)}}{\lambda_k - \lambda_l} g_0^{(l)-1}) d\lambda_l + \mathrm{Tr}(R_0 g_0^{(k)} [g_1^{(k)}, E^{(k)}] g_0^{(k)-1}) d\lambda_k \\ &+ \sum_{l \neq k'} \mathrm{Tr}(R_0 g_0^{(l)} \frac{E^{(l)}}{\lambda'_k - \lambda_l} g_0^{(l)-1}) d\lambda_l - \mathrm{Tr}(R_0 g_0^{(k')} [g_1^{(k')}, E^{(k')}] g_0^{(k')-1}) d\lambda_{k'} \Big\} \end{aligned}$$

In the right hand side, only $l = k, k'$ contibute to the sums over l because $E^{(l)}$ vanishes otherwise. They produce the contribution

$$-\frac{d\lambda_k - d\lambda_{k'}}{\lambda_k - \lambda_{k'}}$$

The terms depending on g_1 give after substituting the explicit value of R_0:

$$\left(g_1^{(k)}\right)_{\alpha\alpha} d\lambda_k - \left(g_1^{(k')}\right)_{\alpha'\alpha'} d\lambda_{k'} - \frac{\left(g_1^{(k)} g_0^{(k)-1} g_0^{(k')}\right)_{\alpha\alpha'} d\lambda_k - \left(g_0^{(k)-1} g_0^{(k')} g_1^{(k')}\right)_{\alpha\alpha'} d\lambda_{k'}}{\left(g_0^{(k)-1} g_0^{(k')}\right)_{\alpha\alpha'}}$$

Hence putting everything together, we get:

$$\gamma' - \gamma = -\mathrm{Tr}\left(R_0 \frac{\Theta^{(k)} - \Theta^{(k')}}{\lambda_k - \lambda_{k'}}\right) - \frac{d\lambda_k - d\lambda_{k'}}{\lambda_k - \lambda_{k'}} = d \log \frac{\left(g_0^{(k)-1} g_0^{(k')}\right)_{\alpha\alpha'}}{\lambda_k - \lambda_{k'}}$$

The case $k = k'$ is proved similarly. ∎

7.7 Ricatti equation.

Notice that the right hand side of eqs.(7.29) is the leading term in the expansion of

$$G^{(kk')}(\lambda,\lambda') = \frac{\delta_{kk'} Id - g^{(k)-1}(\lambda) g^{(k')}(\lambda')}{\lambda - \lambda'} \tag{7.30}$$

in powers of $z_k = \lambda - \lambda_k$ and $z'_{k'} = \lambda' - \lambda_{k'}$. This double expansion has only *positive* powers of z_k and $z_{k'}$. This is clear when $k \neq k'$ and for $k = k'$ the zero in the denominator is cancelled by a zero in the numerator. The matrix elements of G are algebraic functions of the dynamical variables occuring in M_λ. We can recast the equations of motion of the hierarchy in terms of these new variables. They take a particularly simple form of the Ricatti type.

Proposition. *Let us consider the generating function for the flows associated to the pole λ_l:*

$$D^{(l)}_\alpha(\lambda) = \sum_{n>0} (\lambda - \lambda_l)^{n-1} \frac{\partial}{\partial t_{(l,n,\alpha)}} \tag{7.31}$$

Then G obeys the Ricatti type equation:

$$\begin{aligned} D^{(l)}_\alpha(\lambda'') G^{(kk')}(\lambda,\lambda') &= -G^{(kl)}(\lambda,\lambda'') E_{\alpha\alpha} G^{(lk')}(\lambda'',\lambda') \\ -\delta_{kl} E_{\alpha\alpha} \frac{G^{(kk')}(\lambda,\lambda') - G^{(kk')}(\lambda'',\lambda')}{\lambda - \lambda''} &+ \delta_{k'l} \frac{G^{(kk')}(\lambda,\lambda') - G^{(kk')}(\lambda,\lambda'')}{\lambda' - \lambda''} E_{\alpha\alpha} \end{aligned} \tag{7.32}$$

Similarly, the equation of motion relative to the position of the pole λ_l takes the form:

$$\begin{aligned} \partial_{\lambda_l} G^{(kk')}(\lambda,\lambda') &= -\mathrm{Res}_{\lambda''=\lambda_l}\, G^{(kl)}(\lambda,\lambda'') \partial_{\lambda_l} \xi^{(l)}(\lambda'') G^{(lk')}(\lambda'',\lambda') \\ -\delta_{kl} \Big(\partial_{\lambda_k} \xi^{(k)}(\lambda) G^{(kk')}(\lambda,\lambda') \Big)_+ &+ \delta_{k'l} \Big(G^{(kk')}(\lambda,\lambda') \partial_{\lambda_{k'}} \xi^{(k')}(\lambda') \Big)_+ \end{aligned} \tag{7.33}$$

where we take the positive power part in the expansions around λ_k and $\lambda_{k'}$ respectively.

Proof. We need the following identities: let $f(\lambda) = \sum_{j=1}^\infty f_j \lambda^j$, then we have:

$$\sum_{n=1}^\infty \lambda'^{n-1} (\lambda^{-n} f(\lambda))_+ = \frac{f(\lambda) - f(\lambda')}{\lambda - \lambda'}, \quad \sum_{n=1}^\infty \lambda'^{n-1} (\lambda^{-n} f(\lambda))_- = \frac{f(\lambda')}{\lambda - \lambda'}$$

Recalling the equations of the hierarchy expressed on the $g^{(k)}$:

$$\frac{\partial}{\partial t_{(l,n,\alpha)}} g^{(k)}(\lambda) = -(g^{(l)} E_{\alpha\alpha} (\lambda - \lambda_l)^{-n} g^{(l)-1})_- g^{(k)}(\lambda) + \delta_{kl} g^{(k)}(\lambda) E_{\alpha\alpha} (\lambda - \lambda_k)^{-n}$$

we see that they can be recast in the form:

$$D^{(l)}_\alpha(\lambda') g^{(k)}(\lambda) = g^{(k)}(\lambda) E_{\alpha\alpha} \frac{\delta_{kl}}{\lambda - \lambda'} - \frac{g^{(l)}(\lambda') E_{\alpha\alpha} g^{(l)-1}(\lambda') g^{(k)}(\lambda)}{\lambda - \lambda'}$$

which proves the first part of the proposition.
Similarly, using the identity for the function $f(\lambda) = \sum_{-\infty}^\infty f_n \lambda^n$:

$$f_-(\lambda) = \mathrm{Res}_{\lambda'=0} \left(\frac{1}{\lambda - \lambda'} f(\lambda') \right)$$

we can write the equation of motion for $g^{(k)}$ in the form:

$$\partial_{\lambda_l} g^{(k)}(\lambda) = -\mathrm{Res}_{\lambda''=\lambda_l} \frac{g^{(l)}(\lambda'') \partial_{\lambda_l} \xi^{(l)}(\lambda'') g^{(l)\,-1}(\lambda'')}{\lambda - \lambda''} \cdot g^{(k)}(\lambda) + \delta_{kl} g^{(k)}(\lambda) \partial_{\lambda_k} \xi^{(k)}(\lambda)$$

from which the second statement follows. ■

Taking the limits $l = k, \lambda'' \to \lambda$ and $l = k', \lambda'' \to \lambda'$ gives respectively if $k \neq k'$:

$$D_\alpha^{(k)}(\lambda) G^{(kk')}(\lambda,\lambda') = -G^{(kk)}(\lambda,\lambda) E_{\alpha\alpha} G^{(kk')}(\lambda,\lambda') - E_{\alpha\alpha}\partial_\lambda G^{(kk')}(\lambda,\lambda') \quad (7.34)$$

$$D_\alpha^{(k')}(\lambda') G^{(kk')}(\lambda,\lambda') = -G^{(kk')}(\lambda,\lambda') E_{\alpha\alpha} G^{(k'k')}(\lambda',\lambda') + \partial_{\lambda'} G^{(kk')}(\lambda,\lambda') E_{\alpha\alpha} \quad (7.35)$$

7.8 Sato's formula.

It is remarkable that the complete matrix $G^{(kk')}(\lambda,\lambda')$ can be reconstructed from τ and its elementary Schlesinger transforms. As a consequence we can express the matrix elements of G as quotients of $\tau-$functions, as in Sato's formula. We still denote $z_k = \lambda - \lambda_k$ and $z'_{k'} = \lambda' - \lambda_{k'}$ and introduce the notation:

$$t \to t - [z_k]_\alpha \text{ means } t_{(l,n,\gamma)} \to t_{(l,n,\gamma)} - \delta_{kl}\delta_{\gamma\alpha}\frac{z_k^n}{n}$$

Proposition. *We have:*

$$\tau(t)\, G_{\alpha\alpha'}^{(kk')}(\lambda,\lambda') = \begin{cases} \tau\begin{bmatrix} k & k' \\ \alpha & \alpha' \end{bmatrix}(t - [z_k]_\alpha + [z'_{k'}]_{\alpha'}) & \text{if } (k,\alpha) \neq (k',\alpha') \\ \\ \frac{1}{\lambda-\lambda'}(\tau(t) - \tau(t - [z_k]_\alpha + [z'_k]_{\alpha'})) & \text{if } (k,\alpha) = (k',\alpha') \end{cases} \quad (7.36)$$

where we have defined $z_k = \lambda - \lambda_k$ *and* $z'_{k'} = \lambda' - \lambda_{k'}$

Proof. This is a generalization of the proof eq.(2.42) in Chapter [2]. From the definition of the τ–function, we have

$$\frac{\partial}{\partial t_{(l,n,\alpha)}} \log \tau = \mathrm{Res}_{\lambda_l} \mathrm{Tr}\Big(g^{(l)-1}(\lambda)\partial_\lambda g^{(l)}(\lambda) E_{\alpha\alpha}(\lambda - \lambda_l)^{-n}\Big)$$

hence, using the identity eq.(2.41) in Chapter [2], we get

$$D_\alpha^{(k)}(\lambda)\log\tau = \mathrm{Tr}\Big(g^{(k)-1}(\lambda)\partial_\lambda g^{(k)}(\lambda)E_{\alpha\alpha}\Big) = G_{\alpha\alpha}^{kk}(\lambda,\lambda)$$

From this, it follows using eqs.(7.34,7.35) and the definition of G that:

$$(D_\alpha^{(k)}(\lambda) + \partial_\lambda) < \alpha|\tau(t)G^{(kk')}(\lambda,\lambda') = 0 \quad (7.37)$$

$$(D_\alpha^{(k')}(\lambda') - \partial_{\lambda'})\tau(t)G^{(kk')}(\lambda,\lambda')|\alpha> = 0 \quad (7.38)$$

These equations show that $\tau G^{(kk')}(\lambda,\lambda')$ is of the form

$$\tau(t)G_{\alpha\beta}^{(kk')}(\lambda,\lambda') = \tau_{\alpha\beta}(t_\gamma^{(l)} - \delta_{kl}\delta_{\gamma\alpha}[z_k] + \delta_{k'l}\delta_{\gamma\alpha'}[z'_{k'}])$$

To find what are the functions $\tau_{\alpha\beta}$, it is enough to compare the two sides of the equation at $z_k = z'_{k'} = 0$. But there we know that

$$\tau(t)G_{\alpha\beta}^{(kk')}(\lambda,\lambda') \to \tau\begin{bmatrix} k & k' \\ \alpha & \alpha' \end{bmatrix}(t)$$

this proves the proposition if $k \neq k'$. If $k = k'$, the right hand side of eq.(7.34) contains the extra term

$$\frac{G^{(kk)}(\lambda,\lambda') - G^{(kk)}(\lambda,\lambda)}{\lambda'-\lambda}E_{\alpha\alpha}$$

If $\alpha' \neq \alpha$ this does not affect eq.(7.37), but if $\alpha' = \alpha$, then it becomes

$$(D_\alpha^{(k)}(\lambda) - \partial_\lambda)\Big(\tau(t)\big((\lambda-\lambda')G_{\alpha\alpha}^{(kk}(\lambda,\lambda') - 1\big)\Big) = 0$$

Using the analogous equations for λ', and the fact that $\tau(t)\big((\lambda-\lambda')G_{\alpha\alpha}^{(kk}(\lambda,\lambda') - 1\big) \to -\tau$ when $\lambda \to \lambda_k$ and $\lambda' \to \lambda_k$, we get the second half of the proposition. ■

Many remarkable relations can be extracted from this result. In particular, setting $k = k'$, $z_k = 0$, we get denoting $g^{(k)}(\lambda) = g_0^{(k)}\widehat{g}^{(k)}(\lambda)$:

$$\begin{aligned}
\left(\widehat{g}^{(k)}\right)_{\alpha\alpha'}(\lambda') &= (\lambda' - \lambda_k)\frac{\tau\begin{bmatrix} k & k \\ \alpha & \alpha' \end{bmatrix}(t + [z'_k]_{\alpha'})}{\tau(t)}, \quad \alpha \neq \alpha' \\
\left(\widehat{g}^{(k)}\right)_{\alpha\alpha}(\lambda') &= \frac{\tau(t + [z'_k]_\alpha)}{\tau(t)}
\end{aligned} \tag{7.39}$$

Alternatively, we way consider the case $k = \infty, z_\infty = \lambda^{-1} = 0$. The proper definition of $G^{(\infty k')}(\lambda, \lambda')$ is then

$$G^{(\infty k')}(\lambda,\lambda') = \frac{\lambda g^{(\infty)-1}(\lambda) g^{(k')}(\lambda')}{\lambda - \lambda'} \tag{7.40}$$

Remembering that $g_0^{(\infty)} = 1$, we get

$$\left(g^{(k)}\right)_{\alpha\alpha'}(\lambda') = \frac{\tau\begin{bmatrix} \infty & k \\ \alpha & \alpha' \end{bmatrix}(t + [z'_k]_{\alpha'})}{\tau(t)},$$

Recalling the Sato equation, eq.(2.42) in Chapter [2], we see that we have completely identified the $\tau_{\alpha\beta}$ functions occurring in the numerator as Schlesinger transforms of the τ–function in the denominator.

7.9 Hirota's equations

Hirota noticed that integrable equations could be recast into a remarkable bilinear form in terms of τ–functions. Namely introducing the differential operator D with the definition:

$$D_i^n f \cdot g = \left(\frac{\partial}{\partial y_i}\right)^n f(x+y)g(x-y)|_{y=0} \tag{7.41}$$

the equations of motion take the form: $P(D)\tau \cdot \tau = 0$ where P is a polynomial in D. For instance, the equation

$$(D_1^4 + 3D_2^2 - 4D_1D_3)\tau \cdot \tau = 0$$

is the Hirota form of the KP equation. As a matter of fact setting

$$u = 2\frac{\partial^2}{\partial t_1^2}\log\tau \tag{7.42}$$

we get the Kadomtsev–Petviashvili equation:

$$3\frac{\partial^2 u}{\partial t_2^2} + \frac{\partial}{\partial t_1}\left(-4\frac{\partial u}{\partial t_3} + 6u\frac{\partial u}{\partial t_1} + \frac{\partial^3 u}{\partial t_1^3}\right) = 0 \tag{7.43}$$

We show in this section that this is a general phenomenon, and we have the:

Proposition. *In terms of the τ–function and its elementary Schlesinger transforms, the hierarchy equations take the Hirota bilinear form.*

Proof. Let us prove this proposition in a simple case (the other cases are similar). We assume for simplicity that l, k, k' are all different. Multiplying eq.(7.32) by $\tau^2(t)$ we get:

$$\tau^2(t)D_\beta^{(l)}(\lambda'')\frac{1}{\tau(t)}\tau\begin{bmatrix} k & k' \\ \alpha & \alpha' \end{bmatrix}(t-[z_k]_\alpha+[z'_{k'}]_{\alpha'}) \quad = \quad -\tau\begin{bmatrix} k & l \\ \alpha & \beta \end{bmatrix}(t-[z_k]_\alpha+[z''_l]_\beta)\,\tau\begin{bmatrix} l & k' \\ \beta & \alpha' \end{bmatrix}(t-[z''_l]_\beta+[z'_{k'}]_{\alpha'})$$

The left–hand side can be rewritten in terms of Hirota differential operator, using the identity:

$$f^2\frac{\partial}{\partial t}\frac{g}{f} = -(\dot{f}g-\dot{g}f) = -D_t\, f\cdot g$$

so that, introducing the generating function for Hirota differential operators :

$$D_{\mathrm{Hir},\alpha}^{(l)}(\lambda'') = \sum_{n>0}(\lambda''-\lambda_l)^{n-1}D_{t_{(l,n,\alpha)}}$$

and shifting the variables t by $t \to t+[z_k]_\alpha/2-[z'_{k'}]_{\alpha'}/2$, we get:

$$D_{\mathrm{Hir},\beta}^{(l)}(\lambda'')\tau(t+\frac{[z_k]_\alpha}{2}-\frac{[z'_{k'}]_{\alpha'}}{2})\cdot\tau\begin{bmatrix} k & k' \\ \alpha & \alpha' \end{bmatrix}(t-\frac{[z_k]_\alpha}{2}+\frac{[z'_{k'}]_{\alpha'}}{2}) \quad = \quad \tau\begin{bmatrix} k & l \\ \alpha & \beta \end{bmatrix}(t-\frac{[z_k]_\alpha}{2}-\frac{[z'_{k'}]_{\alpha'}}{2}+[z''_l]_\beta)\,\tau\begin{bmatrix} l & k' \\ \beta & \alpha' \end{bmatrix}(t+\frac{[z_k]_\alpha}{2}+\frac{[z'_{k'}]_{\alpha'}}{2}-[z''_l]_\beta)$$

We now expand this formula in powers of z_k, $z'_{k'}$ and z''_l using the equation

$$f(t+z)\,g(t-z) = e^{zD_t}\,f\cdot g$$

It is clear that the coefficients of any monomial in the considered variables have the form of Hirota bilinear equation. ■

Exactly the same method applies to the other Ricatti equations, and shows that they can be recast in Hirota form. This form is very remarkable as it allows for some easy particular solutions and moreover lends itself to a beautiful geometric interpretation that will be described in Chapter [8].

7.10 Tau–functions and Theta–functions.

In this section we show how the Lax matrix approach fits into the isomonodromy approach. In particular we show that, in this case, the τ–functions are essentially Riemann's theta–functions.

We start from a rational Lax matrix $L(\lambda)$ of size $N\times N$, with poles at λ_k. We consider the associated spectral curve, $\Gamma:\ \det(L(\lambda)-\mu)=0$, which is a compact Riemann surface presented as an N–sheeted ramified covering of the Riemann sphere. For any value of the spectral parameter λ one may consider the (multivalued) $N\times N$ matrix $w(\lambda)$ given by $(\psi(P_1),\cdots,\psi(P_N))$, where the P_j are the N points above λ in some order, and $\psi(P_j)$ is the eigenvector of $L(\lambda)$ for the corresponding eigenvalue μ_j. It has been explained in Chapter [4] how to choose the normalization of the eigenvector in such a way that its components, evolving according to the Lax equation, are Baker–Akhiezer functions on Γ. In general they have essential singularities at all points on Γ above the λ_k. Around each λ_k the matrix $w(\lambda)$ has an expansion of the form $w(\lambda) = g^{(k)}(\lambda)\exp(-\xi^{(k)}(\lambda))$ where $g^{(k)}$ is regular at $\lambda = \lambda_k$ and $\xi^{(k)}$ is a diagonal matrix singular at λ_k:

$$\xi^{(k)}(\lambda) = \sum_{n,\alpha} t_{k,n,\alpha} \frac{E_{\alpha\alpha}}{(\lambda-\lambda_k)^n}$$

Here the $t_{k,n,\alpha}$ are the times describing all the integrable flows of the hierarchy. Note that $w(\lambda)$ is precisely of the form considered in this Chapter. Hence it is natural to consider the matrix $M_\lambda = -\partial_\lambda w \cdot w^{-1}$.

Proposition. *The matrix M_λ is a rational function of λ. It has arbitrary order poles at the λ_k, simple poles at the projections of the branch points of the covering $(\lambda, \mu) \to \lambda$, and at the projections of the poles of $w(\lambda)$.*

Proof. The main point is that while $w(\lambda)$ is multivalued, namely its columns undergo permutations when one performs a loop around a branch point, $\partial_\lambda w \cdot w^{-1}$ is independent of the ordering of the columns of $w(\lambda)$, hence it is well defined as a function of λ. Using the asymptotic expansion around the punctures we see that the singular part of M_λ at λ_k is given by $(g^{(k)}\partial_\lambda \xi^{(k)} g^{(k)-1})_-$, hence is a finite order pole if there is a finite number of times.
Let us consider now a branch point, and for simplicity assume that it is of order 2, namely we assume that the first two columns of $w(\lambda)$ coalesce for $\lambda = \lambda_b$. The corresponding eigenvalues $(\mu_1, \mu_2, \cdots)$ are such that μ_1, μ_2 also coalesce. Locally the equation of Γ is of the form $\lambda - \lambda_b = (\mu - \mu_b)^2$. The first two eigenvectors are just the evaluation of one meromorphic vector function of μ at the two roots of this equation. Splitting the even and odd powers of $(\lambda - \lambda_b)^{\frac{1}{2}}$, we can write the matrix $w(\lambda)$ in the form:

$$\begin{aligned} w(\lambda) &= (\psi_e(z) + \sqrt{z}\psi_o(z), \psi_e(z) - \sqrt{z}\psi_o(z), \psi_3(z), \cdots) \\ &= (\psi_e(z), \psi_o(z), \psi_3(z), \cdots) \begin{pmatrix} 1 & 0 & 0 & \cdots \\ 0 & \sqrt{z} & 0 & \cdots \\ 0 & 0 & 1 & \\ \vdots & \vdots & & \ddots \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 & \cdots \\ 1 & -1 & 0 & \cdots \\ 0 & 0 & 1 & \\ \vdots & \vdots & & \ddots \end{pmatrix} \end{aligned}$$

where $z = \lambda - \lambda_b$ and the first matrix $g_{\mathrm{reg}}(z) = (\psi_e(z), \psi_o(z), \psi_3(z), \cdots)$ is regular around $z = 0$. The third matrix is an inessential invertible constant matrix, and the second matrix can be identified with $\exp -\xi(\lambda)$ with $\xi(\lambda) = -1/2 \cdot \log(\lambda - \lambda_b) E_{22}$. This produces in M_λ a polar part $(g\partial_\lambda \xi g^{-1})_- = 1/2 \cdot g_{\mathrm{reg}}(0) E_{22} g_{\mathrm{reg}}^{-1}(0) \cdot 1/(\lambda - \lambda_b)$.
At a pole of $w(\lambda)$ (above λ_c) at finite distance, one column (say the j^{th} one) has a pole. Hence one can write, up to right multiplication by a constant invertible matrix, $w(\lambda) = g(\lambda)\exp(-\xi(\lambda))$ with $\xi(\lambda) = \log(\lambda - \lambda_c)E_{jj}$. This again yields a simple pole in M_λ.
Finally above $\lambda = \infty$, we have N points $Q_1, \cdots, Q_N$ on Γ and in the setup of Chapter [4], $L(\lambda)$ tends to a diagonal matrix when $\lambda \to \infty$, moreover $w(\lambda)$ has simple poles at the $N-1$ points $Q_2, \cdots, Q_N$. The asymptotic form of $w(\lambda)$ when $\lambda \to \infty$ is $g_0^{(\infty)}(1 + O(1/\lambda))\exp(-B_0^{(\infty)}\log(1/\lambda))C^{(\infty)}$ where $C^{(\infty)}$ is a constant invertible matrix and $B_0^{(\infty)} = \sum_{i=2}^N E_{ii}$. This implies that around $\lambda = \infty$ the matrix M_λ has the form $-g_0^{(\infty)} B_0^{(\infty)} g_0^{(\infty)\,-1}/\lambda + O(1/\lambda^2)$. It is now clear that M_λ is a rational function of λ vanishing at ∞ and is the sum of its polar parts at finite distance. ∎

Remark. Note that the last statement of the proof implies that one must have

$$\sum_k \left(g^{(k)} \partial_\lambda \xi^{(k)} g^{(k)\,-1} \right)_- = -g_0^{(\infty)} B_0^{(\infty)} g_0^{(\infty)\,-1} + O\left(\frac{1}{\lambda^2}\right)$$

where the sum on k runs over poles at finite distance. In particular taking the trace one gets the Fuchs relation again. It is interesting to check it in this context. First we have $\operatorname{Tr} B_0^{(\infty)} = N-1$. The poles at finite distance are provided by the branch points, each one contributing $-1/2$ to the trace, and the g dynamical poles, each one contributing $+1$. Finally at the punctures, we have $\mathrm{Tr}\,(g^{(k)}(\lambda)\partial_\lambda \xi^{(k)} g^{(k)\,-1}(\lambda))_- = \mathrm{Tr}\,\partial_\lambda \xi^{(k)} = O(1/\lambda^2)$ at ∞, hence does not contribute to the Fuchs condition which therefore reads $g = \nu/2 - N + 1$. This is just the Riemann–Hurwitz formula.

It follows from the Proposition that the matrix $w(\lambda)$ constructed from the eigenvector bundle satisfies a differential equation $\partial_\lambda w(\lambda) + M_\lambda(\lambda) w(\lambda) = 0$ where M_λ is a rational function of λ of the type studied in this Chapter. However it is a very particular function of λ since the solution $w(\lambda)$ has no monodromy at the punctures and the poles of $L(\lambda)$. Its only non–trivial monodromy occurs around the branch points and is represented by permutations of columns. Globally the monodromy group is finite. Finally there are no Stokes matrices at any singularity. Hence the monodromy data, as defined above, only consist in the connexion matrices $\mathcal{C}^{(k)}$ which are time–independent matrices, function of the moduli of the spectral curve.

The general theory nevertheless applies to this very special situation and in particular τ–functions can be computed explicitly in terms of θ–functions, on which the action of Schlesinger transformations reduce to very simple translations.

Let $P_{k\alpha}$ be the N points above λ_k. We take $z_k = \lambda - \lambda_k$ as local parameter around each $P_{k\alpha}$. We introduce the singular parts $\xi_{k\alpha}(z_k) = l_{k\alpha} \log(z_k) \sum_{n\geq 1} t_{(k,n,\alpha)} z_k^{-n}$ and consider Baker–Akhiezer functions with these singular parts at each $P_{k\alpha}$, and poles at the $g + N - 1$ given points $D = (\nu_1, \cdots, \nu_g)$ (the dynamical divisor) and $Q_2, \cdots, Q_N$ (above ∞). The logarithmic terms with integer coefficients $l_{k\alpha}$ produce extra zeroes or poles of the Baker–Akhiezer functions at the punctures, and will be useful to understand Schlesinger transformations. Hence we choose these numbers such that $\sum_{k\alpha} l_{k\alpha} = 0$. The dimension of the space of such functions is N, by the Riemann–Roch theorem. We fix a particular λ_k and consider the N sheets above it. For each α one can define a unique Baker–Akhiezer function $\psi_\alpha^{(k)}(P)$ satisfying the N conditions:

$$\psi_\alpha^{(k)}(P) = e^{-\xi_{k\alpha}(z_k)}(\delta_{\alpha\beta} + O(z_k)), \quad \text{when } P \to P_{k\beta}$$

We put these N functions in a column vector $\psi^{(k)}(P)$ and form as usual the $N \times N$ matrix $w^{(k)}(\lambda) = (\psi^{(k)}(P_1), \cdots, \psi^{(k)}(P_N))$ where the P_i are the N points above λ. Remark that, around our particular λ_k, we have the expansion $w^{(k)}(\lambda) = \widehat{g}^{(k)}(\lambda) e^{-\xi^{(k)}(z_k)}$ with $\widehat{g}^{(k)}(\lambda) = 1 + O(z_k)$ and $\xi^{(k)}$ is the diagonal matrix $\sum \xi_{k\alpha} E_{\alpha\alpha}$, i.e. we have chosen the normalization $g_0^{(k)} = 1$ in the expansion (7.17). We are going to identify the τ–function by comparing the expression of the matrix $w^{(k)}$ in terms of theta–functions with the expression (7.39).

Following the general procedure of Chapter [4] we construct

$$\psi_\alpha^{(k)}(P) = C \cdot e^{-\int_{P_{k\alpha}}^{P} \Omega} \cdot \frac{\theta(\mathcal{A}(P) - V - \mathcal{K})}{\theta(\mathcal{A}(P) - \mathcal{A}(D) - \mathcal{K})}$$

where C is a normalization constant, Ω is an abelian differential chosen so that $\psi_\alpha^{(k)}$ has the required properties at the punctures and at the Q_j, the theta–function at the denominator has been introduced to take care of the poles at the dynamical divisor D, ($\mathcal{K}$ is the vector of Riemann's constants), and the vector V in the theta–function in the numerator is determined by requiring that the resulting function has no monodromy.

Let us first determine the form Ω. We write it as a sum of three pieces $\Omega = \Omega^{(t)} + \Omega^{(l)} + \Omega^{(q)}$ where $\Omega^{(t)}$ ensures that $\psi_\alpha^{(k)}(P)$ has the correct essential singularities at the punctures. One can write

$$\Omega^{(t)} = \sum_{k',n,\beta} t_{(k',n,\beta)} \Omega_{P_{k'\beta}}^{(n)}$$

where $\Omega_{P_{k'\beta}}^{(n)}$ is the normalized (i.e. the a–periods vanish) second kind abelian differential with just one singularity at $P_{k'\beta}$ and such that around this point $\Omega_{P_{k'\beta}}^{(n)} = d(z_{k'}^{-n}) +$ holomorphic. Note that the integral

$\int_{P_{k\alpha}}^{P}$ is ill–defined for the form $\Omega^{(n)}_{P_{k\alpha}}$. We take as its definition the unique primitive which around $P_{k\alpha}$ behaves as $z_k^{-n} + O(z_k)$ (no constant term). The form $\Omega^{(l)}$ is computed to produce a zero or pole of order $l_{k'\beta}$ at the puncture $P_{k'\beta}$. Noting that $\sum l_{k'\beta} = 0$, it is the unique normalized abelian differential of third kind with first order poles at these points with corresponding residues $-l_{k'\beta}$. As in the previous case the integral of this form with origin at $P_{k\alpha}$ is ill–defined. We take it to be the primitive behaving as $-l_{k\alpha}\log(z_k) + O(z_k)$, mod $2i\pi$. Finally the form $\Omega^{(q)}$ is introduced to get $N-1$ poles at the points $Q_2, \cdots, Q_N$ and $N-1$ zeroes at the points $P_{k\beta}$ with $\beta \neq \alpha$ (for the special k). It is given by the unique third kind differential with residues $+1$ at the Q_j and -1 at the $P_{k\beta}$.

It is now easy to compute the monodromy of the function $\exp(-\int_{P_{k\alpha}}^{P}\Omega)$. Since the differentials are normalized, there is no monodromy around the a–cycles. Around the cycle b_j the monodromy is given by $\exp(-\int_{b_j}\Omega)$. Using the monodromy property of theta–functions, eq.(15.12) in Chapter [15], we can cancel this monodromy by taking:

$$V_j = \frac{1}{2i\pi}\int_{b_j}\Omega + \mathcal{A}_j(D)$$

Decomposing Ω into its three components, we note that $\int_{b_j}\Omega^{(t)}$ is a linear form in the times with coefficients *independent* of the indices (k,α) since all punctures enter symmetrically in its definition. One can compute more precisely the other two contributions, by using Riemann's bilinear identity for third kind differentials:

$$\int_{b_j}\Omega_3 = 2i\pi\sum_l r_l\mathcal{A}_j(P_l)$$

where Ω_3 is a third kind differential with first order poles at the P_l with residue r_l ($\sum r_l = 0$). One gets:

$$\frac{1}{2i\pi}\int_{b_j}\Omega^{(l)} = -\sum_{k'\beta} l_{k'\beta}\mathcal{A}_j(P_{k'\beta})$$

where the last sum is over all punctures. Similarly we find:

$$\frac{1}{2i\pi}\int_{b_j}\Omega^{(q)} = -\sum_{\beta\neq\alpha}\mathcal{A}_j(P_{k\beta}) + \sum_{l=2}^{N}\mathcal{A}_j(Q_l) = \mathcal{A}_j(P_{k\alpha}) - \mathcal{A}_j(Q_1)$$

To get the last equation, note that if $P_1, \cdots, P_N$ are the N points above some λ_0 then the Abel sum $\sum_j \mathcal{A}(P_j)$ is a constant independent of λ_0. This is because the meromorphic function on Γ: $f(P) = (\lambda-\lambda_0)/(\lambda-\lambda_1)$ has zeroes at the points above λ_0 and poles at the points above λ_1, hence these two divisors are mapped to the same point by the Abel map due to the Abel theorem. In particular $\sum_\beta \mathcal{A}_j(P_{k\beta}) = \sum_{l=1}^{N}\mathcal{A}_j(Q_l)$.

There remains to compute the normalization constant C, such that $\psi^{(k)}_\alpha(P) = e^{-\xi_{k\alpha}(z_k)}(1 + O(z_k)$ when $P \to P_{k\alpha}$. Thanks to our definition of the primitive of the form Ω, we have $\int_{P_{k\alpha}}^{P}\Omega = \xi_{k\alpha}(z_k) + O(z_k)$. Hence we need only to normalize the quotient of theta–functions, and we find the final result:

$$\begin{aligned}
\psi^{(k)}_\alpha(P) &= e^{-\int_{P_{k\alpha}}^{P}\Omega}\cdot\frac{\theta(\mathcal{A}(P) - V - \mathcal{K})\theta(\mathcal{A}(P_{k\alpha}) - \mathcal{A}(D) - \mathcal{K})}{\theta(\mathcal{A}(P_{k\alpha}) - V - \mathcal{K})\theta(\mathcal{A}(P) - \mathcal{A}(D) - \mathcal{K})} \\
V &= \sum_{kn\alpha} t_{(k,n,\alpha)}U^{(k,n,\alpha)} - \sum_{k'\beta} l_{k'\beta}\mathcal{A}(P_{k'\beta}) + \mathcal{A}(P_{k\alpha}) - \mathcal{A}(Q_1) + \mathcal{A}(D)
\end{aligned} \tag{7.44}$$

To identify the tau–function, we have to expand this formula for $P \to P_{k\beta}$ (β may be equal to α) in powers of z_k and compare the result with eq.(7.39). We first look at $P \to P_{k\alpha}$. Note that, due to the explicit form of V, one can write for P close to $P_{k\alpha}$:

$$\begin{aligned}
\psi^{(k)}_\alpha(z_k) &= e^{-\xi_{k\alpha}(z_k)}\cdot e^{-\int_{P_{k\alpha}}^{P}(\Omega - d\xi_{k\alpha})}\frac{\theta(\mathcal{A}(P) - \mathcal{A}(D) - \mathcal{K})}{\theta(\mathcal{A}(P_{k\alpha}) - \mathcal{A}(D) - \mathcal{K})}\cdot\frac{\theta(\mathcal{A}(P) - \mathcal{A}(P_{k\alpha}) - U\cdot t - W)}{\theta(-U\cdot t - W)} \\
W &= -\sum_{k'\beta} l_{k'\beta}\mathcal{A}(P_{k'\beta}) - \mathcal{A}(Q_1) + \mathcal{A}(D) + \mathcal{K}
\end{aligned}$$

which compares to $\psi_\alpha^{(k)}(z_k) = e^{-\xi_{k\alpha}(z_k)}\widehat{g}_{\alpha\alpha}^{(k)}(z_k)$. The middle term is regular when $z_k \to 0$ and tends to one, so one can write it in the form $\exp(t.b(z_k) + a(z_k))$ where $t.b(z_k)$ is an expression linear in the times $t_{(k,n,\alpha)}$, and $a(z_k), b(z_k) = O(z_k)$. Considering the last term remark that using Riemann's bilinear identities, it can be rewritten as $\theta(U \cdot (t + [z]) + W)/\theta(U \cdot t + W)$. Indeed we Taylor expand the Abel map $\mathcal{A}(P) - \mathcal{A}(P_{k\alpha})$ around $P_{k\alpha}$. Writing the first kind differential $\omega_j = \sum_{i=0}^\infty c_i^{(j)} z_k^i dz_k$ one gets:

$$\mathcal{A}_j(P) - \mathcal{A}_j(P_{k\alpha}) = \sum_{n=1}^\infty c_{n-1}^{(j)} \frac{z_k^n}{n} = -\frac{1}{2\pi i} \sum_{n=1}^\infty \frac{z_k^n}{n} \oint_{b_j} \Omega_{P_{k\alpha}}^{(n)} = -\sum_{n=1}^\infty \frac{z_k^n}{n} U_j^{(k,n,\alpha)}$$

Inserting this relation in the theta function gives:

$$\theta\left(\mathcal{A}(P) - \mathcal{A}(P_{k\alpha}) - \sum_i t_i U^{(i)} - W\right) = \theta\left(\sum_i (t_i + \frac{z_k^n}{n}\delta_{kk'}\delta_{\alpha\beta})U^{(i)} + W\right)$$

where the multi–index i stands for (k', n, β). Hence one recognizes in the numerator the same theta–function as in the denominator but with the times appropriately shifted.

This leads us to look for a tau–function of the form:

$$\tau(t) = e^{\rho(t)+\sigma(t,t)}\, \theta(\sum_i t_i U^{(i)} + W)$$

where $\rho(t)$ is linear in t and $\sigma(t,t)$ is a quadratic form in t. One gets a condition on ρ and σ:

$$t \cdot b(z_k) + a(z_k) = \rho([z_k]_\alpha) + 2\sigma(t, [z_k]_\alpha) + \sigma([z_k]_\alpha, [z_k]_\alpha) \tag{7.45}$$

One can always choose ρ and σ satisfying this equation provided that b obeys a symmetry property stemming from the fact that the quadratic form σ is symmetric in its arguments. Indeed we have for $(k', \alpha') \neq (k, \alpha)$ (the equality case was treated in Chapter [4]) $b_{(k',n',\alpha')} = \int_{P_{k\alpha}}^P \Omega_{P_{k'\alpha'}}^{(n')} = \sum_n b_{(k',n',\alpha'),(k,n,\alpha)} z_k^n$, and the condition (7.45) implies $\sigma_{(k',n',\alpha'),(k,n,\alpha)} = 1/2 \cdot n b_{(k',n',\alpha'),(k,n,\alpha)}$. So we must have the relation $n b_{(k',n',\alpha'),(k,n,\alpha)} = n' b_{(k,n,\alpha),(k',n',\alpha')}$ (in this equation, the function $b_{(k,n,\alpha)}(z_{k'})$ is obtained by performing the same construction as above but starting from the privilegied point $P_{k'\alpha'}$). This is a consequence of Riemann's bilinear identities: apply the identity eq.(15.7) in Chapter [15] to the second kind differentials $\Omega_{P_{k\alpha}}^{(n)}$ and $\Omega_{P_{k'\alpha'}}^{(n')}$. Since these differentials are normalized their a–periods vanish and the left–hand side of the identity vanishes. We get $\sum \mathrm{Res}\,(b_{(k',n',\alpha')} db_{(k,n,\alpha)}) = 0$. There are two poles, one at $P_{k\alpha}$ and the other at $P_{k'\alpha'}$. Computing the residues yields the required relation. Altogether this shows that the quadratic form σ exists, is completely determined, and independent of the choice of the particular point $P_{k\alpha}$. The computation of the linear form ρ is straightforward. Hence we have proved the:

Proposition. *The τ–function associated to the algebro–geometric integrable system is given by:*

$$\tau(t) = e^{\sigma(t,t)+\rho(t)}\, \theta(U \cdot t + W) \tag{7.46}$$

We are now in position to discuss the effect of a Schlesinger transformation on this tau–function since this only amounts to changing the integers $l_{k\alpha}$. These integers only occur through the contribution $-\sum_{k'\beta} l_{k'\beta}\mathcal{A}(P_{k'\beta})$ to the vector W in eq.(7.46), and through a term linear in $l_{k\alpha}$ in $\rho(t)$ (appearing in the exponential prefactor). We see that an elementary Schlesinger transformation (which adds a pole at $P_{k\alpha}$ and a zero at $P_{k'\alpha'}$) is obtained by changing $l_{k\alpha} \to l_{k\alpha} - 1$ and $l_{k'\alpha'} \to l_{k'\alpha'} + 1$. The effect of such a transformation on the theta–function is remarkably simple and amounts to the translation:

$$\theta(U \cdot t + W) \to \theta(U \cdot t + W + \mathcal{A}(P_{k\alpha}) - \mathcal{A}(P_{k'\alpha'}))$$

To be complete note that the exponential prefactor picks up a term linear in the times under such a transformation.

This allows to perform an interesting check of the first equation in (7.39). Namely the matrix element $\widehat{g}^{(k)}_{\alpha\alpha'}$ is obtained by evaluating $\psi^{(k)}_\alpha(P)$ around the point $P_{k\alpha'}$. Considering eq.(7.44) we see that the relevant theta–function in the numerator is $\theta(U\cdot(t+[z_k]_{\alpha'})+W+\mathcal{A}(P_{k\alpha})-\mathcal{A}(P_{k'\alpha'}))$, which is exactly what we expect for $\tau\begin{bmatrix} k & k \\ \alpha & \alpha' \end{bmatrix}(t+[z_k]_{\alpha'})$, as predicted by eq.(7.39).

Hence we have found that the tau–function in the algebro–geometric situation is essentially the Riemann theta–function. Moreover the various matrix elements of $g(t,\lambda)$ are obtained by simple shifts of the arguments of the theta–function. The isomonodromic context of this Chapter is a generalization of the algebro–geometric context, so that the tau–functions can be viewed as generalizations of Riemann's theta–functions and should enjoy many of their remarkable properties. In Chapter [8] another framework is proposed allowing to directly define the tau–function, using the geometry of the infinite Grassmannian.

7.11 Painlevé equations

An important application of the theory of isomonodromic deformations concerns the Painlevé equations. The Painlevé property deals with singularities of solutions of differential equations. In this respect there is a striking difference between linear differential equations and non–linear ones. The solutions of linear differential equations have singularities (poles, branch points, essential singularities) only where the coefficients of the equation have singularities, i.e. the position of these singularities are fixed. In contrast the solutions of non–linear equations can develop singularities at arbitrary points, depending on initial conditions. For example $\dot{y}=y^2$ has the general solution $y=-1/(t-t_0)$. We call these singularities depending on the initial conditions movable singularities. In general one cannot avoid movable poles, however one can try to find equations whose solutions have no movable singularities other than poles, i.e. the branch points and essential singularities of all solutions are fixed. This is called the Painlevé property. In fact Painlevé and Gambier have classified all differential equations of the form:

$$\frac{d^2y}{dt^2}=R(t,y,\frac{dy}{dt})$$

where R is a rational function of its arguments, satisfying the Painlevé property. Up to trivial redefinitions they found 50 such equations. Of all these equations only 6 could not be integrated in terms of already known functions, and are listed below.

$$\begin{aligned}
(i)\quad \tfrac{d^2y}{dt^2} &= 6y^2+t \\
(ii)\quad \tfrac{d^2y}{dt^2} &= 2y^3+ty+\alpha \\
(iii)\quad \tfrac{d^2y}{dt^2} &= \frac{1}{y}\left(\frac{dy}{dt}\right)^2-\frac{1}{t}\frac{dy}{dt}+\frac{1}{t}(\alpha y^2+\beta)+\gamma y^3+\frac{\delta}{y} \\
(iv)\quad \tfrac{d^2y}{dt^2} &= \frac{1}{2y}\left(\frac{dy}{dt}\right)^2+\frac{3}{2}y^3+4ty^2+2(t^2-\alpha)y+\frac{\beta}{y} \\
(v)\quad \tfrac{d^2y}{dt^2} &= \left\{\frac{1}{2y}+\frac{1}{y-1}\right\}\left(\frac{dy}{dt}\right)^2-\frac{1}{t}\frac{dy}{dt}+\frac{(y-1)^2}{t^2}\left\{\alpha y+\frac{\beta}{y}\right\}+\frac{\gamma y}{t}+\frac{\delta y(y+1)}{y-1} \\
(vi)\quad \tfrac{d^2y}{dt^2} &= \frac{1}{2}\left\{\frac{1}{y}+\frac{1}{y-1}+\frac{1}{y-t}\right\}\left(\frac{dy}{dt}\right)^2-\left\{\frac{1}{t}+\frac{1}{t-1}+\frac{1}{y-t}\right\}\frac{dy}{dt}+ \\
&\qquad \frac{y(y-1)(y-t)}{t^2(t-1)^2}\left\{\alpha+\frac{\beta t}{y^2}+\frac{\gamma(t-1)}{(y-1)^2}+\frac{\delta t(t-1)}{(y-t)^2}\right\}
\end{aligned}$$

All these equations can be understood in the framework of isomonodromy deformations but we shall only consider equations (ii) and (vi) in order to illustrate various aspects of the method.
THIS IS NOT AN ACCIDENT. QUOTE MIWA MALGRANGE AND $\tau = \theta$.

Example 1. The Painlevé (ii) equation.

We apply the general construction to the case where all matrices are of size 2×2, with just one singularity at $\lambda = \infty$, the singularity data being given by:

$$\xi(\lambda) = \left(\frac{\lambda^3}{3} + \frac{t\lambda}{2} + \theta \log \left(\frac{1}{\lambda} \right) \right) \sigma_3, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Note that, since we are looking for an ordinary non–linear differential equation we have introduced only one time, i.e. the parameter t above, but clearly one could extend this construction to a whole hierarchy of times. Note also that we have introduced a branch point parametrized by θ (which will account for the parameter α in Painlevé (ii)), and that this is consistent with the Fuchs condition since $\mathrm{Tr}\,(\sigma_3) = 0$. Finally for irrational θ this prohibits any Lax pair interpretation of the differential equation since the monodromy is infinite.

We set $w(\lambda) = g(\lambda) e^{-\xi(\lambda)}$ and take

$$g(\lambda) = 1 + \frac{g_1}{\lambda} + \frac{g_2}{\lambda^2} + \frac{g_3}{\lambda^3} + \cdots$$

Note that there is no restriction in assuming that the leading term is 1. We also impose that $\det g(\lambda) = 1$, i.e. that $w(\lambda)$ belongs to $Sl(2)$. The g_i are such that the matrix $w(\lambda)$ satisfies the linear differential equation $\partial_\lambda w = -M_\lambda w$ with $M_\lambda = (g \partial_\lambda \xi g^{-1})_-$. Here the symbol $()_-$ means taking the polynomial part of the considered expression. Finally the isomonodromic deformation equation is $\partial_t w = -M_t w$ with $M_t = (g \partial_t \xi g^{-1})_-$. One gets:

$$\begin{aligned} M_\lambda &= \lambda^2 \sigma_3 + \lambda [g_1, \sigma_3] + \frac{t}{2} \sigma_3 + [g_2, \sigma_3] - [g_1, \sigma_3] g_1 \\ M_t &= \frac{1}{2} \lambda \sigma_3 + \frac{1}{2} [g_1, \sigma_3] \end{aligned}$$

We see that M_λ only depends on g_1, g_2, and $\partial_\lambda g = -M_\lambda g + g \partial_\lambda \xi$ determines the $g_3, g_4, \cdots$ in terms of these two matrices. One finds for $n \geq 3$:

$$[g_n, \sigma_3] = -(1 - \delta_{n,3}) g_{n-3} + \theta g_{n-3} \sigma_3 - \frac{t}{2} [g_{n-2}, \sigma_3] + [g_1, \sigma_3] g_{n-1} + [g_2, \sigma_3] g_{n-2} - [g_1, \sigma_3] g_1 g_{n-2} \qquad (7.47)$$

We parametrize g_i in the form:

$$g_i = \begin{pmatrix} \Delta_i + a_i & b_i \\ c_i & \Delta_i - a_i \end{pmatrix}$$

where Δ_i is obtained by requiring that $\det g(\lambda) = 1$, yielding $\Delta_1 = 0$, $\Delta_2 = (a_1^2 + b_1 c_1)/2$, $\Delta_3 = a_1 a_2 + (b_1 c_2 + b_2 c_1)/2$, etc. The left–hand side of eq.(7.47) has vanishing diagonal elements. This provides a constraint on the lower order g_n. On the other hand, the off–diagonal elements determine b_n and c_n. In the case $n = 3$, we find one constraint:

$$b_1 c_2 + b_2 c_1 = \frac{\theta}{2}$$

and get the off–diagonal elements:

$$\begin{aligned} -b_3 &= \frac{t}{2} b_1 + a_2 b_1 + a_1 b_2 + b_1 \Delta_2 \\ c_3 &= -\frac{t}{2} c_1 + a_2 c_1 + a_1 c_2 - c_1 \Delta_2 \end{aligned}$$

Similarly for $n = 4$ we get the constraint:

$$a_1 = (t + 4\Delta_2) b_1 c_1 - 2 b_2 c_2 - 2 a_1 (b_1 c_2 - b_2 c_1)$$

while a_2 is determined by the equation at order $n = 5$. Finally there are only 3 free parameters in g_1 and g_2.

The equations of motion read:

$$\dot{g}_1 = \frac{1}{2}[g_2, \sigma_3] - \frac{1}{2}[g_1, \sigma_3]g_1, \quad \dot{g}_2 = \frac{1}{2}[g_3, \sigma_3] - \frac{1}{2}[g_1, \sigma_3]g_2$$

Note that since $[g_3, \sigma_3]$ is known in terms of g_1, g_2 these equations close on these two matrices. In components we get:

$$\dot{a}_1 = b_1 c_1, \quad \dot{b}_1 = -b_2 - a_1 b_1, \quad \dot{c}_1 = c_2 - a_1 c_1$$
$$\dot{b}_2 = \frac{t}{2} b_1 + a_1 b_2 + 2 b_1 \Delta_2, \quad \dot{c}_2 = -\frac{t}{2} c_1 + a_1 c_2 - 2 c_1 \Delta_2$$

The equations for $\dot{\Delta}_2$ and $\dot{a}_2$ are in fact consequences of the equations of the first line. It is now natural to introduce the parameters:

$$y = a_1 + \frac{b_2}{b_1}, \quad z = b_1 c_1$$

With these parameters one has $a_1 = tz + 2z^2 + 2zy^2 - \theta y$. Then the equations of motion give $\dot{y} = t/2 + 2z + y^2$ and $\dot{z} = -2zy + \theta/2$. Eliminating z yields the equation:

$$\frac{d^2 y}{dt^2} = 2y^3 + yt + \frac{1}{2} + \theta$$

One recognizes the Painlevé (ii) equation with $\alpha = 1/2 + \theta$. Once the solution $y(t)$ of the Painlevé equation is known, it is easy to completely reconstruct the matrix $M_\lambda(\lambda; t)$. The peculiar nature of the t–dependence of M_λ ensuring the isomonodromy property is here particularly obvious! In particular it is known that solutions of the Painlevé equations may have poles at arbitrary positions.

The Schlesinger transformations take a particularly simple form on this example. They just amount to changing $\theta \to \theta_n = \theta + n$, so that the first column of w acquires a pole of order n at $\lambda = \infty$ while the second column acquires a zero of order n. Let w_n the wave–function constructed with the parameter θ_n, then we have for an elementary Schlesinger transformation $w_{n+1}(\lambda) = R_n(\lambda) w_n(\lambda)$ where $R_n(\lambda)$ is a matrix rational in λ. Plugging the asymptotic expansions at ∞, $w_n(\lambda) \sim g_n(\lambda) \exp(-\xi_n(\lambda))$, one gets $R_n(\lambda) g_n(\lambda) = g_{n+1}(\lambda)(\lambda E_{11} + \lambda^{-1} E_{22})$. Expanding to order λ^{-3} one gets:

$$R_n(\lambda) = \begin{pmatrix} \lambda - y_n & -b_{1,n} \\ 1/b_{1,n} & 0 \end{pmatrix}$$

where $b_{1,n}$ and y_n are the parameters occurring in $g_n(\lambda)$. Moreover one also gets:

$$a_{1,n+1} = -\frac{b_{2,n}}{b_{1,n}}, \quad b_{1,n+1} = b_{3,n} - y_n b_{2,n} - b_{1,n}(\Delta_{2,n} - a_{2,n}) \tag{7.48}$$

$$c_{1,n+1} = \frac{1}{b_{1,n}}, \quad c_{2,n+1} = \frac{a_{1,n}}{b_{1,n}} \tag{7.49}$$

From this we obtain the recursion relations:

$$z_{n+1} = -\frac{t}{2} - y_n^2 - z_n, \quad y_{n+1} = -y_n + \frac{\theta_{n+1}}{2 z_{n+1}}$$

It is now illuminating to introduce the tau–functions. First we compute the closed form $\gamma = \mathrm{Res}_\infty \, \mathrm{Tr}\,(g^{-1}(\lambda) \partial_\lambda d\xi) d\lambda$. Here we simply have $d\xi = \lambda \sigma_3 dt/2$, so that we find:

$$\gamma = \frac{1}{2} \mathrm{Tr}\,(g_1 \sigma_3) dt = a_1 dt$$

Hence τ_n is defined by $\dot{\tau}_n / \tau_n = a_{1,n}$. One can express the matrix elements of $g(\lambda)$ in terms of tau–functions according to the general results of the previous sections. Here things are so simple that it is even more straightforward to rederive the appropriate results. From eq.(7.48) we have $a_{1,n+1} = \dot\tau_{n+1}/\tau_{n+1} = -b_{2,n}/b_{1,n}$. It follows immediately that

$$y_n = a_{1,n} + \frac{b_{2,n}}{b_{1,n}} = \frac{\dot\tau_n}{\tau_n} - \frac{\dot\tau_{n+1}}{\tau_{n+1}} = \frac{d}{dt}\log\left(\frac{\tau_n}{\tau_{n+1}}\right)$$

On the other hands by the equations of motion we have $y_n = -d\log b_{1,n}/dt$. Moreover by eq.(7.49) we have $c_{1,n} = 1/b_{1,n-1}$. Hence we can take:

$$b_{1,n} = \frac{\tau_{n+1}}{\tau_n}, \quad c_{1,n} = \frac{\tau_{n-1}}{\tau_n}, \quad z_n = \frac{\tau_{n+1}\tau_{n-1}}{\tau_n^2}$$

The Hirota bilinear form of the equations of motion also follows straightforwardly. Let us write the three equations of motion:

$$\begin{aligned}
\dot a_{1,n} = z_n &\quad\rightarrow\quad \tau_n\ddot\tau_n - \dot\tau_n^2 - \tau_{n+1}\tau_{n-1} = 0\\
\dot z_n = -2y_n z_n - \theta_n &\quad\rightarrow\quad \tau_{n+1}\dot\tau_{n-1} - \tau_{n-1}\dot\tau_{n+1} + \theta_n\tau_n^2 = 0\\
\dot y_n = 2z_n + y_n^2 + \frac{t}{2} &\quad\rightarrow\quad \tau_{n+1}\ddot\tau_n + \tau_n\ddot\tau_{n+1} - 2\dot\tau_n\dot\tau_{n+1} + \frac{1}{2}t\tau_n\tau_{n+1} = 0
\end{aligned}$$

In the last equation we have used the first one to simplify the result.

Example 2. The Painlevé (vi) equation.

Here we consider the case of three regular singularities at finite distance which we put at $\lambda = 0, 1, t$, and will consider the isomonodromic deformation equations with respect to the parameter t. We choose the singularity data at these 3 points as:

$$\xi^{(0)}(\lambda) = \begin{pmatrix}\theta_0 & 0\\ 0 & 0\end{pmatrix}\log(\lambda), \quad \xi^{(1)}(\lambda) = \begin{pmatrix}\theta_1 & 0\\ 0 & 0\end{pmatrix}\log(\lambda-1), \quad \xi^{(t)}(\lambda) = \begin{pmatrix}\theta_t & 0\\ 0 & 0\end{pmatrix}\log(\lambda-t)$$

With this we construct the wave–function $w(\lambda)$ having the asymptotic expansion $w(\lambda) \sim g^{(k)}(\lambda)\exp(-\xi^{(k)}(\lambda))$, with $k = 0, 1, t$ and satisfying the differential equation $\partial_\lambda w = -M_\lambda w$. This implies:

$$M_\lambda(\lambda) = \frac{A^{(0)}}{\lambda} + \frac{A^{(1)}}{\lambda-1} + \frac{A^{(t)}}{\lambda-t}, \quad A^{(k)} = g_0^{(k)}\begin{pmatrix}\theta_k & 0\\ 0 & 0\end{pmatrix}g_0^{(k)\,-1}$$

In particular $A^{(k)}$ is a rank 1 projector, hence $\det A^{(k)} = 0$ and $\mathrm{Tr}\,A^{(k)} = \theta_k$. The equation of motion reads $\partial_t w = -M_t w$ with:

$$M_t = \left(g^{(t)}\partial_t\xi^{(t)}g^{(t)\,-1}\right)_- = -\frac{A^{(t)}}{\lambda-t}$$

We are free to make a global gauge transformation by a matrix constant in λ, and we use this freedom to diagonalize M_λ at ∞, so that we assume:

$$M_\lambda(\lambda) = \frac{1}{\lambda}\begin{pmatrix}\kappa_1 & 0\\ 0 & \kappa_2\end{pmatrix} + O(\lambda^{-2}), \quad \lambda\to\infty$$

This means that the differential equation $\partial_\lambda w = -M_\lambda w$ has a regular singularity at ∞, and the Fuchs condition reads $\kappa_1 + \kappa_2 = \theta_0 + \theta_1 + \theta_t$.

We can write M_λ as:

$$M_\lambda(\lambda) = \begin{pmatrix}m_{11}(\lambda) & m_{12}(\lambda)\\ m_{21}(\lambda) & m_{22}(\lambda)\end{pmatrix} = \frac{1}{\lambda(\lambda-1)(\lambda-t)}A(\lambda)$$

where $A_{ii}(\lambda)$ are second degree polynomials in λ with leading term $\kappa_i\lambda^2$, and $A_{12}(\lambda)$ and $A_{21}(\lambda)$ are degree one polynomials in λ. We set $A_{12}(\lambda) = \gamma(\lambda - y)$ introducing the important parameter y which will end up satisfying the Painlevé (vi) equation. Taking the trace of M_λ we have $A_{22}(\lambda) = \theta_0(\lambda-1)(\lambda-t) + \theta_1\lambda(\lambda-$

$t) + \theta_t\lambda(\lambda-1) - A_{11}(\lambda)$ which eliminates the parameters in A_{22}. Moreover we know that $\det A(\lambda)$ vanishes for $\lambda = 0, 1, t$, so we get three conditions on A_{21}.

$$A_{21}(\lambda) = \frac{A_{11}(\lambda)A_{22}(\lambda)}{\gamma(\lambda - y)}, \quad \lambda = 0, 1, t$$

The first two allow to express A_{21} in terms of A_{11} and A_{12}, while the last one provides, since $A_{21}(\lambda)$ is a linear function of λ, a constraint on the parameters in A_{11}:

$$A_{21}(t) = (1-t)A_{21}(0) + tA_{21}(1)$$

In order to write more explicitly this constraint, let us parametrize $A_{11}(\lambda)$ by the values it takes at $\lambda = t$ and $\lambda = y$, i.e. write the interpolation formula:

$$A_{11}(\lambda) = (\lambda - t)(\lambda - y)\kappa_1 + \frac{(\lambda - y)A_{11}(t) - (\lambda - t)A_{11}(y)}{t - y}$$

Substituting this into the previous equation one observes that the constraint is quadratic in $A_{11}(y)$ but, unexpectedly, is linear in $A_{11}(t)$. We solve $A_{11}(t)$ in term of $A_{11}(y)$ getting:

$$\begin{aligned} A_{11}(t) &= -\frac{\kappa_1^2}{\kappa_1 - \kappa_2}(y-t)(y-1+t) + \frac{\kappa_1}{\kappa_1 - \kappa_2}\Big((\theta_0 t + \theta_1(t-1))(y-t) + 2A_{11}(y)\Big) \\ &\quad - \frac{A_{11}(y)}{(\kappa_1 - \kappa_2)y(y-1)}\Big(A_{11}(y) + \theta_1 y(t-1) + \theta_0 t(y-1)\Big) \end{aligned}$$

The dynamical variables are now $y, \gamma, A_{11}(y)$ and we have to write the equations of motion $\partial_t M_\lambda = \partial_\lambda M_t + [M_\lambda, M_t]$. Taking the matrix element 12 of this equation and evaluating it at $\lambda = y$ gives:

$$\dot{y} = \frac{y(y-1)(y-t)}{t(t-1)}\left[-2m_{11}(y) + \frac{\theta_0}{y} + \frac{\theta_1}{y-1} + \frac{\theta_t + 1}{y-t}\right]$$

Taking the matrix element 11 evaluated at $\lambda = y$ gives:

$$\dot{m}_{11}(\lambda)|_{\lambda=y} = \frac{A_{11}(t)}{t(t-1)(y-t)^2} - \frac{1}{t(t-1)}\gamma m_{21}(y)$$

We introduce the dynamical variable $z = m_{11}(y) = A_{11}(y)/(y(y-1)(y-t))$ and compute $\dot{z} = \dot{m}_{11}(\lambda)|_{\lambda=y} + (\partial_\lambda m_{11})|_{\lambda=y} \cdot \dot{y}$. We find after some algebra:

$$\dot{z} = \frac{(3y^2 - 2(1+t)y + t)z^2 + \Big(-2(\kappa_1 + \kappa_2 + 1)y + 1 + \theta_0 + \theta_t + (\theta_0 + \theta_1)t\Big)z + \kappa_1(1 + \kappa_2)}{t(t-1)}$$

Eliminating z between the equations for $\dot{y}$ and $\dot{z}$ yields the Painlevé (vi) equation, with the following values of the parameters:

$$\alpha = \frac{1}{2}(\kappa_1 - \kappa_2 - 1)^2, \quad \beta = -\frac{1}{2}\theta_0^2, \quad \gamma = \frac{1}{2}\theta_1^2, \quad \delta = \frac{1}{2}(1 - \theta_t^2)$$

References

[1] E.L. Ince **Ordinary differential equations.** Dover Publications Inc. (1956)

[2] W. Wasow **Asymptotic expansions for ordinary differential equations.** Interscience Publishers (1965)

[3] Y. Sibuya **Linear differential equations in the complex domain: problems of analytic continuation.** AMS Providence (1990).

[4] B. Malgrange *Sur les déformations isomonodromiques.* Mathematics and Physics, (Paris, 1979/1982), Progr. Math., 37, Birkhäuser Boston, Mass. (1983).

[5] H. Flaschka, A. Newell *Monodromy and Spectrum preserving deformations I* Commun. Math. Phys. 76 (1980) 65–166

[6] M. Jimbo, T. Miwa, K. Ueno *Monodromy preserving deformations of ordinary differential equations with rational coefficients I* Physica 2D (1981) 306–352

[7] M. Jimbo, T. Miwa *Monodromy preserving deformations of ordinary differential equations with rational coefficients II* Physica 2D (1981) 407–448

[8] M. Jimbo, T. Miwa *Monodromy preserving deformations of ordinary differential equations with rational coefficients III* Physica 4D (1981) 26–46

Chapter 8

Vertex operators and integrable hierarchies.

The lessons of previous Chapters is that when several singularities are present, the situation decomposes as a product of local studies around each singularity. In this Chapter we consider the case of one singularity in all its generality, which amounts to the study of the Kadomtsev–Pteviashvili equation. This finds a natural presentation in Sato's Grassmannian approach. Hirota equations are interpreted as the Plücker equations of an infinite dimensional Grassmannian, on which the infinite dimensional group $GL(\infty)$ acts. The output is a remarkable formula expressing the τ function as $\tau(t) =< 0|e^{\sum_i H_i t_i} g|0 >$. Here, the time flows are induced by the action of one parameter groups. We give an explicit realization of the Grassmannian in a fermionic Fock space, allowing to obtain solitons through the action of vertex operators. Finally we show that the full KP hierarchy admits a Lax pair formulation in terms of pseudo–differential operators. This means that the Lie algebra underlying this formulation is induced by the associative algebra structure of pseudo–differential operators.

8.1 Tau functions and Hirota equations.

We have seen in Chapter [2] that one can simultaneously solve the hierarchy of equations:

$$\partial_{t_j} w = -(g\xi_j g^{-1})_- w \tag{8.1}$$

with g regular and B_j diagonal with constant singular part. Here w is a matrix of fundamental solutions of this system. The matrix elements $w_{\alpha\beta}$ admit a remarkable expression in terms of τ–functions,

$$w_{\alpha\beta}(z; t_1, t_2, \cdots) = \frac{\tau_{\alpha\beta}(t - [z]_\beta)}{\tau(t)} \exp\,(\sum_{n\geq 1} z^{-n} t_{n,\beta}) \tag{8.2}$$

IL FAUT MIEUX ECRIRE LA FORMULE DE SATO POUR $\widetilde{g}_{\alpha,\alpha}$

where $(t - [z]_\beta)$ means that only times $t_{n,\beta}$ are shifted by z^n/n. The nice feature of this expression is that the z dependence is folded into the t dependence. Moreover w obeys an infinite set of bilinear identities which are *equivalent* to the hierarchy. It follows that τ–functions also obeys an infinite set of bilinear identities. The local formula (8.2) was recovered in a global context in Chapter [4] where we found that the τ–functions were essentially Riemann's theta–functions.

The bilinear identities where introduced by Hirota who noticed that integrable equations could be recast into a remarkable form on τ–functions, namely $P(D)\tau\cdot\tau = 0$ where P is a polynomial of the differential operator D with the definition:

$$D_i^n f\cdot g = \left(\frac{\partial}{\partial y_i}\right)^n f(x+y)g(x-y)|_{y=0} \tag{8.3}$$

For instance, the equation

$$(D_1^4 + 3D_2^2 - 4D_1D_3)\tau\cdot\tau = 0$$

is the Hirota form of the KP equation. Indeed setting

$$u = 2\frac{\partial^2}{\partial t_1^2}\log\tau \tag{8.4}$$

we get the Kadomtsev–Petviashvili equation:

$$3\frac{\partial^2 u}{\partial t_2^2} + \frac{\partial}{\partial t_1}\left(-4\frac{\partial u}{\partial t_3} + 6u\frac{\partial u}{\partial t_1} + \frac{\partial^3 u}{\partial t_1^3}\right) = 0 \tag{8.5}$$

In this Chapter we elaborate on the relation between Sato's formula (8.2), Hirota's bilinear equations and representation theory of infinite dimensional Kac–Moody algebras. The main motivation is that Sato's formula has the following structure:

$$w(z,t) = \frac{V(z,t)\tau(t)}{\tau(t)}, \quad V(z,t) = \exp\,(\sum_{n\geq 1} z^{-n}t_n)\,\exp\,(-\sum_{n\geq 1}\frac{z^n}{n}\frac{\partial}{\partial t_n}) \tag{8.6}$$

where the operator $V(z,t)$ is a typical vertex operator which is known to play a central role in the study of representation theory of affine Kac–Moody algebras, see Chapter [16].

8.2 Bilinear identities and orbits.

The highest weight representations of Kac–Moody algebras by vertex operators automatically produce τ–functions obeying bilinear identities. The prototype of this construction is as follows.

Let $\mathcal{G}$ be an affine Kac-Moody algebra and $\mathcal{H}$ a Cartan subalgebra of $\mathcal{G}$. It is often convenient to choose an orthonormal basis t^a of $\mathcal{G}$, and also to consider a Cartan–Weyl basis (H_i, E_α) where H_i is a basis of $\mathcal{H}$, and E_α are the root vectors associated to the roots α, normalized by $(E_\alpha, E_{-\alpha}) = 1$. The tensor Casimir C_{12} of $\mathcal{G}$ is defined by:

$$C_{12} = \sum_a t^a\otimes t^a = \sum_i H^i\otimes H_i + \sum_\alpha E_\alpha\otimes E_{-\alpha}$$

Its fundamental properties are that it commutes with elements g of the group G, i.e. $C_{12}\; g\otimes g = g\otimes g\; C_{12}$, and its action on the tensor product of highest weight vectors $|\Lambda\rangle\otimes|\Lambda'\rangle$ is given by: $C_{12}|\Lambda\rangle\otimes|\Lambda'\rangle = (\Lambda,\Lambda')|\Lambda\rangle\otimes|\Lambda'\rangle$. This last property comes from the fact that highest weight vectors $|\Lambda\rangle$ are eigenstates of H_i and are annihilated by the generators E_α for α any positive root. These two relations imply the following identity:

$$C_{12}\; g|\Lambda\rangle\otimes g|\Lambda'\rangle = (\Lambda,\Lambda')\; g|\Lambda\rangle\otimes g|\Lambda'\rangle, \quad \forall\;\; g\in G \tag{8.7}$$

Suppose now that the representation of $\mathcal{G}$, with highest weight vector $|\Lambda\rangle$, admits a vertex operator construction in terms of bosonic operators α_n, $n\in\mathbb{Z}$, satisfying $[\alpha_n,\alpha_m] = n\delta_{n+m,0}$. Recall that a vertex operator is an operator acting on the Fock space generated by these bosonic oscillators and has the typical form $V =: \exp\left(\sum_{n\neq 0}(z^n/n)\alpha_n\right) : \quad \equiv \exp\left(\sum_{n<0}(z^n/n)\alpha_n\right)\exp\left(\sum_{n>0}(z^n/n)\alpha_n\right)$. Here the bosonic Fock space is generated by acting with the operators α_n on $|\Lambda\rangle$ with $\alpha_n|\Lambda\rangle = 0$ for $n > 0$. A basis of the Fock space is made of the states of the form $\alpha_{-n_1}\cdots\alpha_{-n_k}|\Lambda\rangle$. To have a vertex operator representation of $\mathcal{G}$ in that space means that the generators t^a of $\mathcal{G}$ are represented as polynomials in the α_n acting on the Fock space. Taking for $\alpha_{-n},\ n > 0$ the operator of multiplication by nt_n, and for $\alpha_n,\ n > 0$ the derivation operator $\partial/\partial t_n$ which satisfy the above commutation relations, the vertex operator appearing in eq.(8.6) is exactly recovered. Moreover the Fock space is the space of polynomials in the variables t_n.

Let us introduce the generating function, $H(t) = \sum_n \alpha_n t_n$, in an infinite number variables t_n. We *define* the τ–functions by:

$$\tau^g_\Lambda(t) = \langle\Lambda|e^{H(t)}g|\Lambda\rangle, \quad \text{with} \;\; H(t) = \sum_n \alpha_n t_n, \quad g \in G \tag{8.8}$$

Since any element $t^a \in \mathcal{G}$ can be represented by some polynomial in the α_n, there exist differential operators $\mathcal{D}^a_\Lambda(t,\partial_t)$ in the variables t_n such that:

$$\langle\Lambda|e^{H(t)}t^a g|\Lambda\rangle = \mathcal{D}^a_\Lambda(t,\partial_t)\langle\Lambda|e^{H(t)}g|\Lambda\rangle \tag{8.9}$$

Indeed, we have

$$\langle\Lambda|e^{H(t)}\alpha_m g|\Lambda\rangle = \begin{cases} \frac{\partial}{\partial t_m}\langle\Lambda|e^{H(t)}g|\Lambda\rangle \text{ for } m > 0 \\ mt_{-m}\langle\Lambda|e^{H(t)}g|\Lambda\rangle \text{ for } m < 0 \end{cases} \tag{8.10}$$

The first case is obvious, while for $m < 0$ one uses $\exp(H(t))\alpha_m\exp(-H(t)) = \alpha_m + mt_{-m}$ to bring α_m to the left where it is annihilated by the vacuum.

Taking the scalar product of the fundamental bilinear relation (8.7) with $\langle\Lambda|e^{H(t)} \otimes \langle\Lambda'|e^{H(t')}$ and using eq.(8.9) we obtain:

$$\left[\sum_a \mathcal{D}^a_\Lambda(t,\partial_t)\mathcal{D}^a_{\Lambda'}(t',\partial_{t'})\right]\tau^g_\Lambda(t)\tau^g_{\Lambda'}(t') = (\Lambda,\Lambda')\tau^g_\Lambda(t)\tau^g_{\Lambda'}(t')$$

These equations are of the Hirota type.

Clearly this construction can be applied to any vertex operator representation of an affine Kac-Mooody algebras $\mathcal{G}$. The hierarchy of differential equations we obtain in this way depends on the affine Kac-Moody algebra but also on the choice of the vertex operator representation. In the next section we make these ideas precise in the case of the group $GL(\infty)$.

8.3 Fermions and $GL(\infty)$

Let $gl(\infty)$ be the Lie algebra of infinite dimensional band matrices of the form M_{rs} with $r, s \in \mathbb{Z} + \frac{1}{2}$ (the $\frac{1}{2}$ is added for convenience) and such that $M_{rs} = 0$ for $|r-s|$ large. Note that the sum and the product of two such matrices is well–defined and of the same type. Finally $\widehat{gl}(\infty)$ is a central extension of $gl(\infty)$ that will be described later on, by displaying a representation of this algebra on some Hilbert space.

To do that we introduce fermionic operators $\psi_r,\ \psi^*_r,\ r \in \mathbb{Z} + \frac{1}{2}$, with the following anti-commutation relations:

$$\begin{aligned} \{\psi_r,\psi_s\} = \{\psi^*_r,\psi^*_s\} &= 0 \\ \{\psi_r,\psi^*_s\} &= \delta_{r+s,0} \end{aligned} \tag{8.11}$$

where $\{\ ,\ \}$ denotes the anticommutator, ie. $\{a,b\} = ab + ba$ for any operators a and b. By convention, we will say that the fermions ψ_r have charge $(+1)$ while the fermionic operators ψ^*_s have charge (-1).

These operators act on a fermionic Fock space built upon a vacuum vector $|0\rangle$ such that:

$$\psi_r|0\rangle = \psi_r^*|0\rangle = 0 \qquad \text{for } r \geq \tfrac{1}{2}$$

The Fock space is generated by acting successively on the vacuum with the fermionic operators. A basis of the fermionic Fock space is formed by the following set of states:

$$\psi_{-s_n}\cdots\psi_{-s_1}\cdot\psi^*_{-r_m}\cdots\psi^*_{-r_1}|0\rangle \tag{8.12}$$

with $s_j, r_j \geq \frac{1}{2}$. The Hilbert structure on the Fock space is defined by introducing the dual vacuum vector:

$$\langle 0|\psi_r = \langle 0|\psi_r^* = 0 \qquad \text{for } r \leq -\tfrac{1}{2}$$

and imposing $\langle 0|0\rangle = 1$ and that the adjoint of the fermionic operators are given by:

$$\psi_r^\dagger = \psi^*_{-r}, \quad (\psi_r^*)^\dagger = \psi_{-r}$$

This implies that the distinct states of the form (8.12) form an orthonormal basis of the Fock space. Since the fermions ψ_s have charge $(+1)$ and the fermions ψ^*_r charge (-1) the state (8.12) has charge $(n-m)$.

Let $|p\rangle$ be defined by:

$$\begin{aligned} |p\rangle &\equiv \psi_{-p+\frac{1}{2}}\cdots\psi_{-\frac{1}{2}}|0\rangle, \quad \text{for} \quad p>0, \\ &\equiv \psi^*_{p+\frac{1}{2}}\cdots\psi^*_{-\frac{1}{2}}|0\rangle, \quad \text{for} \quad p<0, \end{aligned} \tag{8.13}$$

These states have charge p and satisfy $\psi_{r-p}|p\rangle = \psi^*_{r+p}|p\rangle = 0$ for $r \geq \frac{1}{2}$. We call them charged vacuum vectors. The states with charge p are obtained by acting on the vacuum $|p\rangle$ with an equal number of ψ_s and ψ^*_r. In the following we will mostly restrict ourselves to neutral states which are linear conbinaisons of states of the form (8.12) with the same number of ψ and ψ^*.

Now we define the normal order on products of fermionic operators: this consists in writing all operators with positive indices on the right and multiplying by the signature of the considered permutation. The important property of normal ordered products is that their vacuum expectation value vanishes.

The relation between a monomial and its normal ordered form is given by the Wick theorem. For two fermions it reads:

$$\psi_r\psi^*_s \quad = \quad :\psi_r\psi^*_s: + \langle 0|\psi_r\psi^*_s|0\rangle$$

This is because to bring the left–hand side to its normal ordered form we have to perform at most one anticommutation, hence adding at most a c–number which can only be the vacuum expectation of the left–hand side since the normal product has vanishing vacuum expectation value. This scalar is called the contraction of the two operators. For more than two operators one proceeds by induction:

$$\phi_1 : \phi_2\cdots\phi_n :=: \phi_1\phi_2\cdots\phi_n : + \sum_{i=2}^{n}(-1)^i\langle 0|\phi_1\phi_i|0\rangle : \phi_2\cdots\widehat{\phi}_i\cdots\phi_n :$$

where ϕ_j is either some ψ_r or some ψ^*_s and the hat means omission.

We can now construct the algebra $\widehat{gl}(\infty)$. Consider neutral bilinear operators of the form $\psi_r\psi^*_s$, or charged bilinear operators of the form $\psi_r\psi_s$ or $\psi^*_r\psi^*_s$. They form a closed algebra. For example:

$$[\psi_r\psi^*_s, \psi_n\psi^*_m] = \delta_{s+n,0}(\psi_r\psi^*_m) - \delta_{r+m,0}(\psi_n\psi^*_s)$$

Considering neutral and charged bilinear operators together yields the infinite dimensional Lie algebra of orthogonal transformations, while restricting to neutral operators yields the infinite dimensional algebra of general linear transformations. If we were considering only a finite number of fermionic operators this restriction would correspond to the regular embedding of $su(n) \times u(1)$ in $so(2n)$. More precisely consider for an infinite dimensional band matrix M_{rs} the formal operator $X = \sum_{r,s} M_{rs}\psi_r\psi^*_{-s}$. The commutator of two such objects can be computed without ambiguity:

$$[X, Y] = \sum_{rsnm} M_{rs}N_{nm}[\psi_r\psi^*_{-s}, \psi_n\psi^*_{-m}] = \sum_{rm}[M, N]_{rm}\psi_r\psi^*_{-m}$$

hence reproduces the Lie algebra of the band matrices. However these objects cannot be represented on the Fock space since their matrix elements on this space may be infinite. For example $\langle 0|X|0\rangle = \sum_{r>0} M_{rr}$ which can be infinite. To overcome this problem we normal order the bilinear fermionic operator. Then all matrix elements of these objects are finite between states with finite number of particles. This induces a modification of the commutation rules:

$$\begin{aligned}[:\psi_r\psi^*_s:, :\psi_n\psi^*_m:] &= \delta_{s+n,0} :\psi_r\psi^*_m: - \delta_{r+m,0} :\psi_n\psi^*_s: \\ &\quad + \delta_{s+n,0}\langle 0|\psi_r\psi^*_m|0\rangle - \delta_{r+m,0}\langle 0|\psi_n\psi^*_s|0\rangle \end{aligned} \tag{8.14}$$

The additional c–number term is the central extension of $gl(\infty)$.

Definition. *The algebra $\widehat{gl}(\infty)$ is the infinite dimensional Lie algebra of elements of the form:*

$$\widehat{gl}(\infty) = \left\{ X = \sum_{r,s} M_{rs} :\psi_r\psi^*_s:, \quad M_{rs} = 0 \;\; if \;\; |r-s| >> 0 \right\} \oplus \mathbb{C} \tag{8.15}$$

equipped with the Lie bracket (8.14).

We will need to consider the group $GL(\infty)$ associated to this Lie algebra. The precise definition of this object is somewhat tricky. We will adopt a very naive point of view, and consider group elements of the form:

$$g = e^{X_1}e^{X_2}\cdots e^{X_k}$$

with $X_1, \cdots X_k \in \widehat{gl}(\infty)$ defined by *finite* sums of bilinears in the fermionic opeators. It is then clear that such group elements are well represented in the Fock space. Of course we will quickly encounter the need to extend this setting to infinite sums of bilinears, then one has to be very careful about convergence properties but we will not attempt to define a general framework and refer instead to the mathematical literature.

The group $GL(\infty)$ acts on the fermions by conjugation.

Proposition. *For any $g \in GL(\infty)$ we have*

$$g\psi_s g^{-1} = \sum_r \psi_r a_{rs} \tag{8.16}$$

$$g\psi^*_{-s}g^{-1} = \sum_r \psi^*_{-r}(a^{-1})_{sr} \tag{8.17}$$

for some $\mathbb{C}$-number matrix a_{rs}.

<u>Proof</u>. We use a matrix notation: ψ is an ∞ line vector, ψ^* is an ∞ column vector. We want to show that for a given g

$$\begin{aligned} g\psi g^{-1} &= \psi \cdot a \\ g^{-1}\psi^* g &= a \cdot \psi^* \end{aligned}$$

with the same matrix a. Remark that if $g_1\psi g_1^{-1} = \psi a_1$ and $g_2\psi g_2^{-1} = \psi a_2$ then $g_1 g_2\psi g_2^{-1}g_1^{-1} = \psi a_1 a_2$, and similarly if $g_1^{-1}\psi^* g_1 = b_1\psi^*$, $g_2^{-1}\psi^* g_2 = b_2\psi^*$, then $g_2^{-1}g_1^{-1}\psi^* g_1 g_2 = b_1 b_2\psi^*$. So, it is enough the verify that

$a = b$ on generators of the form $g = \exp \epsilon : \psi_r \psi_s^* :$. On these simple generators the normal ordering produces scalar factors which do not contribute to the adjoint action, and we can forget it. If $(r+s) \neq 0$, we have $g = e^{\epsilon \psi_r \psi_s^*} = 1 + \epsilon \psi_r \psi_s^*$, and therefore,

$$g \psi_k g^{-1} = \psi_k + \epsilon [\psi_r \psi_s^*, \psi_k] = \psi_l (\delta_{kl} + \epsilon \delta_{rl} \delta_{s+k,0})$$
$$g^{-1} \psi_{-l}^* g = \psi_{-l}^* - \epsilon [\psi_r \psi_s^*, \psi_{-l}^*] = (\delta_{kl} + \epsilon \delta_{rl} \delta_{s+k,0}) \psi_{-k}^*$$

If $(r+s) = 0$ then $g = e^{\epsilon \psi_r \psi_{-r}^*}$, and

$$g \psi_k g^{-1} = \psi_k e^{\epsilon \delta_{rk}} \quad , \quad g^{-1} \psi_{-k}^* g = \psi_{-k}^* e^{\epsilon \delta_{rk}}$$

This proves the result. ■

Notice that the definition we choose for $GL(\infty)$ ensures that the matrices a_{rs} only have a finite but arbitrary large number of non-zero entries.

It is convenient to work with generating functions $\psi(z)$ and $\psi^*(z)$ also called fermionic fields defined by:

$$\psi(z) = \sum_{r \in \mathbb{Z} + \frac{1}{2}} \psi_r z^{-r-1/2} \tag{8.18}$$
$$\psi^*(z) = \sum_{r \in \mathbb{Z} + \frac{1}{2}} \psi_r^* z^{-r-1/2} \tag{8.19}$$

Notice that $\psi(z)$ and $\psi^*(z)$ are single valued, ie. $\psi(z e^{2i\pi}) = \psi(z)$ and similarly for $\psi^*(z)$.

Proposition. *The fermionic fields satisfy the following "operator product expansion" :*

$$\psi(z)\psi^*(w) =: \psi(z)\psi^*(w) : + \frac{1}{z-w}, \quad \text{for} \quad |z| > |w| \tag{8.20}$$

Proof. It is enough to compute the vacuum expectation value of the left–hand side. This is given by $\sum_{r,s} \langle 0|\psi_r \psi_s^*|0\rangle z^{-r-\frac{1}{2}} w^{-s-\frac{1}{2}}$. Since $\langle 0|\psi_r \psi_s^*|0\rangle = 1$ when $r = -s > 0$ and vanishes otherwise we get $\langle 0|\psi(z)\psi^*(w)|0\rangle = 1/z \sum_{i=0}^{\infty} (w/z)^i = 1/(z-w)$ when $|z| > |w|$. ■

By Wick's theorem this can be generalized to any product of fields yielding:

$$\langle 0| \prod_{j=1}^{N} \psi(z_j) \prod_{j=1}^{N} \psi^*(w_j)|0\rangle = \det \left(\frac{1}{z_i - w_j} \right) \tag{8.21}$$

A direct consequence of eqs.(8.16, 8.17) is that there exists a fermionic analogue of the tensor Casimir. This fact is the crucial point of the fermionic construction of the hierarchy of differential equations.

Proposition. *Let S_{12} be defined by:*

$$S_{12} = \sum_{r \in Z + \frac{1}{2}} \psi_r \otimes \psi_{-r}^* = \oint \frac{dz}{2i\pi} \, \psi(z) \otimes \psi^*(z) \tag{8.22}$$

Then,

$$S_{12} \; g \otimes g = g \otimes g \; S_{12}, \qquad \forall g \in GL(\infty) \tag{8.23}$$

The operator $(S^\dagger S)$ is the tensor Casimir for $\widehat{gl}(\infty)$.

Proof. Eq.(8.23) is equivalent to :

$$\sum_{r\in Z+1/2} \psi_r g\otimes \psi^*_{-r} g = \sum_{r\in Z+1/2} g\psi_r \otimes g\psi^*_{-r}, \qquad \forall g\in GL(\infty)$$

From eq.(8.16,8.17), we have $\psi_r g = g\psi_s(a^{-1})_{sr}$ and $\psi^*_{-r}g = a_{rs}g\psi^*_{-s}$ so that

$$\begin{aligned}\sum_r \psi_r g\otimes \psi^*_{-r} g &= \sum_{r,s,k} g\psi_s \otimes g\psi^*_{-k}(a^{-1})_{sr}a_{rk}\\ &= \sum_s g\psi_s\otimes g\psi^*_{-s}\end{aligned}$$

This is equivalent to eq.(8.23). ■

Notice that the definition of the vacuum implies that $\psi_r|0\rangle\otimes\psi^*_{-r}|0\rangle = 0$, $\forall r$, since either $\psi_r|0\rangle = 0$ or $\psi^*_{-r}|0\rangle = 0$. Therefore

$$S_{12}|0\rangle\otimes|0\rangle = \sum_r \psi_r|0\rangle\otimes\psi^*_{-r}|0\rangle = 0 \tag{8.24}$$

for any $g\in GL(\infty)$. This relation is the fermionic analogue of the bosonic equation (8.7) and will be very important in the following.

Remark. The transformations (8.16,8.17) can alternatively be written in a dual form as transformations on functions of the variables z:

$$z^s \quad\rightarrow\quad A^g_s(z) = \sum_r z^r a_{rs}. \tag{8.25}$$

which is such that $(g\psi_s g^{-1}) = \oint dz\psi(z)A^g_s(z)z^{-1/2}$. These sums are finite sums as long as we use the previous definition for the group $GL(\infty)$. As pointed out above, extending this construction such that the definition of the functions $A^g_s(z)$ involves infinite sums requires choosing an appropriate completion of $GL(\infty)$.

8.4 Boson–fermion correspondence.

We now come to a truly remarkable result known as the boson–fermion correspondence. This will be the main technical tool used in the following sections.

We construct bosonic operators from the fermions and conversely. The bosonic operators H_n are bilinear in the fermionic operators.

Proposition. *Define the operators H_n by*

$$:\psi(z)\psi^*(z):= \sum_{n\in\mathbb{Z}} z^{-n-1}H_n \tag{8.26}$$

or equivalently $H_n = \sum_r :\psi_r\psi^*_{-r+n}:$. We then have the following commutation relations :

$$\begin{aligned}[][H_n,\psi(z)] &= z^n\psi(z)\\ [H_n,\psi^*(z)] &= -z^n\psi^*(z)\\ [H_n,H_m] &= n\delta_{n+m,0}\end{aligned} \tag{8.27}$$

Proof. The proof follows from standard manipulations using Wick's theorem. By definition, we have:

$$H_n = \oint \frac{dz}{2i\pi} z^n : \psi(z)\psi^*(z) :$$

Using Wick's theorem, we may write the product $H_n\, \psi(w)$ as:

$$\begin{aligned} H_n\psi(w) &= \oint \frac{dz}{2i\pi} z^n : \psi(z)\psi^*(z) : \psi(w) \\ &= \oint \frac{dz}{2i\pi} z^n : \psi(z)\psi^*(z)\psi(w) : - \oint_{|w|<|z|} \frac{dz}{2i\pi} z^n \frac{1}{w-z}\psi(z) \end{aligned}$$

Here the integration over z is performed on a circle C_1 around the origin enclosing w due to eq.(8.20). Similarly

$$\begin{aligned} \psi(w)H_n &= \oint \frac{dz}{2i\pi} z^n \psi(w) : \psi(z)\psi^*(z) : \\ &= \oint \frac{dz}{2i\pi} z^n : \psi(z)\psi^*(z)\psi(w) : - \oint_{|w|>|z|} \frac{dz}{2i\pi} z^n \frac{1}{w-z}\psi(z) \end{aligned}$$

The integral is over a circle C_2 around the origin *not* enclosing w. So the commutator $[H_n, \psi(w)]$ is:

$$\begin{aligned} [H_n, \psi(w)] &= -\oint_{C_1-C_2} \frac{dz}{2i\pi} z^n \frac{1}{w-z}\psi(z) \qquad (8.28) \\ &= \oint_{C_w} \frac{dz}{2i\pi} z^n \frac{1}{z-w}\psi(z) = w^n\psi(w) \end{aligned}$$

where C_w is a small circle around w. The other relations are proved similarly. ∎

From eq.(8.27) we see that the H_n obey bosonic commutation relations. Moreover we have $H_n^\dagger = H_{-n}$. In addition to the operator H_0 (which we will also call p, the momentum operator) which is both hermitian and in the center of the bosonic algebra, we introduce an operator q such that $[p, q] = -i$.

Let us now describe the bosonic Fock space. The vacuum vector $|l\rangle$ is such that:

$$H_n|l\rangle = 0, \quad \text{for} \quad n > 0, \quad H_0|l\rangle = l|l\rangle l \in \mathbb{Z}$$

Here we choose $l \in \mathbb{Z}$ because in terms of fermionic operators H_0 has integer spectrum. On these vacua, the operator e^{iq} acts as a translation operator: $e^{iq}|l\rangle = |l+1\rangle$.

Over each state $|l\rangle$ we construct a bosonic Fock space by acting with the creation operator H_n with $n < 0$. A basis of the bosonic Fock space is made of the states

$$H_{-n_1} \cdots H_{-n_k}|l\rangle$$

We will say that these states have charge l; so the charge is the H_0-eigenvalue. In particular, the neutral bosonic Fock space is spanned by the states obtained by acting with the H_n's on the vacuum $|0\rangle$.

We have constructed bosons from fermions, the unexpected result is that fermions can be reconstructed from bosons. We start with bosonic operators obeying commutation relations $[H_n, H_m] = n\delta_{n+m,0}$ and $[p, q] = -i$ with $p = H_0$:

Proposition. *Let* $\phi(z) = q - ip\log z - i\sum_{n\neq 0} \frac{H_n}{n} z^{-n}$, *with* $[p, q] = -i$. *Then, the fermionic fields are reconstructed as:*

$$\psi(z) =: \exp(i\phi(z)) : \quad , \quad \psi^*(z) =: \exp(-i\phi(z)) : \qquad (8.29)$$

where : : here denotes the bosonic normal ordering, which consists in moving the operators H_n with $n > 0$ to the right, and moving the e^{iq} factors to the left. Explicitly, the operators $:\exp(\pm i\phi(z)):$ read

$$:\exp(\pm i\phi(z)) := e^{\pm iq} z^{\pm p} \exp\left(\pm \sum_{n<0} \frac{z^{-n}}{n} H_n\right) \exp\left(\pm \sum_{n>0} \frac{z^{-n}}{n} H_n\right) \tag{8.30}$$

The operators $V_\pm(z)$ are called vertex operators.

Proof. The proof relies on the following relations:

$$\begin{aligned} :\exp(\pm i\phi(z))::\exp(\pm i\phi(w)): \; &= \; (z-w) :\exp(\pm i\phi(z))\exp(\pm i\phi(w)):, \quad |z|>|w| \\ \exp(\pm i\phi(z))::\exp(\mp i\phi(w)): \; &= \; \frac{1}{z-w} :\exp(\pm i\phi(z))\exp(\mp i\phi(w)):, \quad |z|>|w| \end{aligned} \tag{8.31}$$

which follow immediately from the standard Campbell–Hausdorff relation $e^A e^B = e^B e^A \; e^{[A,B]}$ for two operators A and B whose commutator $[A, B]$ is a c–number. To compute the anticommutator $\{\psi_r, \psi_s^*\}$ we represent the fermionic operators as contour integrals, $\psi_r = \oint dz z^{-r-1/2} :\exp(i\phi(z)):$ and similarly for ψ_s^*. We then use the above relations and apply contour integral manipulations as in the previous proposition:

$$\begin{aligned} \{\psi_r, \psi_s^*\} \; &= \; \Big[\oint_{|z|>|w|} dzdw - \oint_{|z|<|w|} dzdw\Big] \frac{z^{-r-1/2} w^{-s-1/2}}{z-w} : e^{i\phi(z)} e^{-i\phi(w)} : \\ &= \; \oint_0 dz \oint_z dw \, \frac{z^{-r-1/2} w^{-s-1/2}}{z-w} : e^{i\phi(z)} e^{-i\phi(w)} := \delta_{r+s,0} \end{aligned}$$

The proof that $\{\psi_r^*, \psi_s^*\} = \{\psi_r, \psi_s\} = 0$ is similar. ■

When describing the bosonic Fock space, we denoted by $|l\rangle$ states which are eigenstates of the momentum operator, $p|l\rangle = l|l\rangle$, and which satisfy $H_n|l\rangle = 0$ for $n > 0$. Using the bosonic realization of the fermions, one verifies that these states $|l\rangle$ also satisfy : $\psi_{r-l}|l\rangle = \psi_{r+l}^*|l\rangle = 0$ for $r \geq \frac{1}{2}$. Thus they coincide with the states that we defined in eq.(8.13), i.e. fermionic charged vacua of charge l.

The boson-fermion correspondence may also be viewed as a consequence of Cauchy's determinant formula. Indeed, Wick's theorem allows us to compute expectation values of vertex operators:

$$\langle 0| \prod_{j=1}^N : e^{i\phi(z_j)} : \prod_{j=1}^N : e^{-i\phi(w_j)} : |0\rangle = \frac{\prod_{i<j}(z_i - z_j)(w_i - w_j)}{\prod_{i,j}(z_i - w_j)} \tag{8.32}$$

Similarly we have:

$$\langle 0| \prod_{j=1}^N : e^{i\phi(z_j)} | -N >= \prod_{i<j}(z_i - z_j) \prod_i z_i^{-N} \tag{8.33}$$

The bosonization formula will then follows if the fermionic (8.21) and the bosonic (8.32) expectation values coincide; namely if we have:

$$\det\left(\frac{1}{z_i - w_j}\right) = \frac{\prod_{i<j}(z_i - z_j)(w_i - w_j)}{\prod_{i,j}(z_i - w_j)} \tag{8.34}$$

This is nothing but Cauchy's determinant formula.

8.5 τ–functions and Hirota bilinear identities.

The tau functions are functions on the orbit of the vacuum $|0\rangle$ under the action of the group $GL(\infty)$. To define them as in eq.(8.8), we again introduce the function $H(t)$ by :

$$H(t) \; = \; \sum_{n>0} t_n H_n$$

Definition. *We define the τ–function $\tau(t)$ as the vacuum expectation value:*

$$\tau(t;g) = \langle 0|e^{H(t)}g|0\rangle, \quad g \in GL(\infty) \tag{8.35}$$

We wish to prove that the τ–functions obey Hirota equations. From eq.(8.23) we get the important following relation satisfied by the states $g|0\rangle$ belonging to the orbit of the vacuum under $GL(\infty)$:

$$S_{12} \cdot g|0\rangle \otimes g|0\rangle = \oint \frac{dz}{2i\pi}\psi(z)g|0\rangle \otimes \psi^*(z)g|0\rangle = 0 \tag{8.36}$$

This is a fundamental identity on vectors in Fock space belonging to the orbit of the vacuum vector. It translates readily into bilinear identities on the τ–functions.

Proposition. *Let $\tau(t;g)$ be a τ–function as in eq.(8.35). It obeys the infinite set of bilinear identities:*

$$\oint \frac{dz}{2i\pi} \exp\left(-\sum_{n>0} 2z^n y_n\right) \exp\left(\sum_{n>0} \frac{z^{-n}}{n}\frac{\partial}{\partial y_n}\right) \tau(x+y;g)\tau(x-y;g) = 0 \tag{8.37}$$

<u>Proof</u>. Applying $e^{H(t')} \otimes e^{H(t")}$ to eq.(8.36), and taking the inner product with $\langle +1| \otimes \langle -1|$, we get :

$$\oint \frac{dz}{2i\pi} \langle +1|e^{H(t')}\psi(z)g|0\rangle \, \langle -1|e^{H(t")}\psi^*(z)g|0\rangle = 0$$

To compute $\langle +1|e^{H(t')}\psi(z)g|0\rangle$ and $\langle -1|e^{H(t")}\psi^*(z)g|0\rangle$ we use the bosonization formula. This gives:

$$\begin{aligned} \langle +1|e^{H(t')}\psi(z)g|0\rangle &= V_+(z,t')\langle 0|e^{H(t')}g|0\rangle = V_+(z,t')\cdot\tau(t';g) \\ \langle -1|e^{H(t")}\psi^*(z)g|0\rangle &= V_-(z,t")\langle 0|e^{H(t")}g|0\rangle = V_-(z,t")\cdot\tau(t";g) \end{aligned} \tag{8.38}$$

where $V_\pm(z;t)$ is the vertex operator in its differential operator representation:

$$V_\pm(z;t) = \exp\left(\mp\sum_{n>0} z^n t_n\right) \exp\left(\pm\sum_{n>0}\frac{z^{-n}}{n}\frac{\partial}{\partial t_n}\right)$$

Indeed inserting eq.(8.29) into the left–hand side of eq.(8.38) we get

$$\langle +1|e^{H(t')}\psi(z)g|0\rangle = \langle +1|e^{H(t')}e^q z^p \exp\left(\sum_{n<0}\frac{z^{-n}}{n}H_n\right)\exp\left(\sum_{n>0}\frac{z^{-n}}{n}H_n\right)g|0\rangle$$

Noticing that $\langle +1|e^q z^p = \langle 0|$, and commuting $\exp\left(\sum_{n<0}\frac{z^{-n}}{n}H_n\right)$ to the left using the Campbell–Hausdorff formula we get:

$$\langle 0|e^{H(t')}\exp\left(\sum_{n<0}\frac{z^{-n}}{n}H_n\right) = e^{-\xi(t,z)}\langle 0|e^{H(t')}, \quad \xi(t,z) = \sum_{n>0} t_n z^n$$

where we have used $\langle 0|H_n = 0$ for $n<0$. Hence the expression reads:

$$e^{-\xi(t,z)}\langle 0|e^{H(t')}\exp\left(\sum_{n>0}\frac{z^{-n}}{n}H_n\right)g|0\rangle = \exp\left(-\sum_{n>0} z^n t_n\right)\exp\left(\sum_{n>0}\frac{z^{-n}}{n}\frac{\partial}{\partial t_n}\right)\langle 0|e^{H(t')}g|0\rangle$$

The bilinear identity then becomes :

$$\oint \frac{dz}{2i\pi} V_+(z;t')\; V_-(z;t")\; \tau(t';g)\tau(t";g) = 0$$

Introducing the variables $t = \frac{1}{2}(t' + t")$ and $y = \frac{1}{2}(t' - t")$, we get the result. ∎

The bilinear identity (8.37) is the main equation of the vertex operator construction of integrable systems. It contains a hierarchy of integrable non-linear differential equations and much more.

The bilinear identity (8.37) contains an infinite set of bilinear Hirota equations. Let us introduce the Hirota operators for the infinite set of variables t_n :

$$D = \{D_1, D_2, \cdots, D_n, \cdots\}$$

The bilinear identity (8.37) for the τ-function can then be rewritten in terms the Hirota operators as:

$$\oint \frac{dz}{2i\pi} \exp\left(-\sum_{n>0} 2z^n y_n\right) \exp\left(\sum_{n>0} \frac{z^{-n}}{n} D_n\right) \exp\left(\sum_{n>0} y_n D_n\right) \tau(t;g)\cdot\tau(t;g) = 0$$

Using the definition of the elementary Schur polynomials:

$$e^{\xi(t,z)} = \exp\left(\sum_{n=1}^{\infty} t_n z^n\right) = \sum_{k=0}^{\infty} P_k(t) z^k \tag{8.39}$$

we can expand in z the quantity $\exp\left(-\sum_{n>0} 2z^n y_n\right)$. Integrating over z shows that the bilinear identity (8.37) is equivalent to:

$$\left(\sum_{j=0}^{\infty} P_j(-2y) P_{j+1}(\widehat{D}) \exp\sum_{n>0} y_n D_n\right) \tau(t;g)\cdot\tau(t;g) = 0, \qquad \forall y_n \tag{8.40}$$

where $\widehat{D} = \{D_1, \frac{1}{2}D_2, \cdots, \frac{1}{n}D_n, \cdots\}$.

These equations are Hirota equations for the KP hierarchy. They contain an infinite number of equations which are obtained by expanding this relation in power fo the y_n. For instance, the coefficient of y_1^4 gives the equation

$$(D_1^4 + 3D_2^2 - 4D_1D_3)\tau\cdot\tau = 0$$

This is the Hirota form of the KP equation.

Remark. Since the states $|l\rangle$ satisfy : $\psi_{r-l}|l\rangle = \psi^*_{r+l}|l\rangle = 0$ for $r \geq \frac{1}{2}$, we also have the relation :

$$S_{12}\cdot g|l\rangle \otimes g|l\rangle = \sum_r \psi_r g|l\rangle \otimes \psi^*_{-r} g|l\rangle = 0.$$

Thus we similarly prove the more general bilinear identities for the τ-functions $\tau_l(t;g) = \langle l|e^{H(t)}g|l\rangle$:

$$\oint \frac{dz}{2i\pi} z^{l-l'}\, V_-(z;t')\; V_+(z;t")\; \tau_l(t';g)\tau_{l'}(t";g) = 0, \qquad \forall t', t", l, l', \quad l \geq l' \tag{8.41}$$

8.6 The KP hierarchy and its soliton solutions.

To produce a solution of the K.P. hierarchy, we only have to choose an element $g \in GL(\infty)$ and calculate the τ-function. A particularly simple choice is to take

$$g_{1-sol.} = e^{a\psi(p)\psi^*(q)} = 1 + a\psi(p)\psi^*(q), \qquad \text{with} \quad |p| > |q|$$

This will correspond to the one-soliton solutions. The N-soliton solutions are obtained by choosing the group element to be product of N-elements corresponding to one soliton solutions.

Proposition. *The N-soliton solutions of the KP hierarchy are given by:*

$$g = g_1 g_2 \cdots g_N, \quad \text{with} \quad g_i = 1 + a_i \psi(p_i)\psi^*(q_i)$$

The N-soliton τ-functions are: $\tau_N(t;g) = \langle 0|e^{H(t)} \prod_{i=1}^N (1 + a_i\psi(p_i)\psi^*(q_i))\,|0\rangle$. They are equal to:

$$\tau_N(t;g) = 1 + \sum_{p=1}^{N} \sum_{\substack{I\subset\{1,\cdots,N\}\\ |I|=p}} \prod_{i<j\in I} \frac{(p_i-p_j)(q_i-q_j)}{(p_i-q_j)(q_i-p_j)} \cdot \prod_{i\in I} e^{\eta_i} \tag{8.42}$$

where $\eta_i = \xi(t,p_i) - \xi(t,q_i) + \log\left(\frac{a_i}{p_i-q_i}\right)$ with $\xi(t,p) = \sum_n p^n t_n$.

<u>Proof</u>. Using the commutation relation (8.27) of the fermions with $H(t)$, we rewrite the N-soliton τ-functions as:

$$\tau_N(t;g) = \langle 0| \prod_{i=1}^{N} \Big(1 + a_i e^{\xi(t,p_i)-\xi(t,q_i)} \psi(p_i)\psi^*(q_i)\Big)\,|0\rangle \tag{8.43}$$

Formula (8.42) immedialty follows by applying Wick's theorem to compute these expectation values

$$\tau_N(t;g) = 1 + \sum_i e^{\eta_i} + \sum_{i<j} e^{\eta_i+\eta_j} \frac{(p_i-p_j)(q_i-q_j)}{(p_i-q_j)(q_i-p_j)} + \cdots$$

■

8.7 Fermions and Grassmannian.

The previous construction can be described more geometrically by relating it to Grassmanians. Let V be a vector space of dimension N. The Grassmanian $Gr(M,V) \equiv Gr(M,N)$ is the space of M-dimensional hyperplanes in V. Since hyperplanes are uniquely determined by the data of M-dimensional frames in V modulo linear transformations, $Gr(M,V)$ may be defined as:

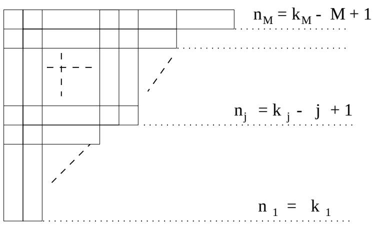
$$Gr(M,V) \equiv \{M - \text{frames in } V\}/GL(M) \tag{8.44}$$

Let $(e_1,\cdots,e_N)$ be a basis of V. Let $P(\Lambda^M V)$ be the projective space of dimension $C_M^N - 1$ on the M^{th} exterior power of V. A basis in $P(\Lambda^M V)$ is formed by the vectors $e_{k_1}\wedge\cdots\wedge e_{k_M}$. The vectors of this basis can be indexed by Young tableaux.

Recall that for any integer $p>0$ a partition of p is defined as a collection of integers $n_j,\ j=1,\cdots,m$ such that $p=\sum_{j=1}^m n_j$ and $1\le n_1\le n_2\le\cdots\le n_m$. To such a partition one associates a diagram of boxes with m lines, and n_m boxes in the first line, n_{m-1} boxes in the second line, up to n_1 boxes in the last line. This is called a Young tableau with p boxes.

To the set of number $(k_1,\cdots,k_M)$ with $1\le k_1<\cdots<k_M\le N$ (remark that $k_i\ne k_j$ for $i\ne j$) we associate the Young tableau $Y=[n_j]$ with j^{th} line of length $n_j = k_j - j + 1$ (now $1\le n_1\le\cdots\le n_M\le N$)

$\mathrm{Y} = [k_j - j + 1] =$



These Young tableaux have at most M lines and N columns. We will denote by e_Y the vector $e_{k_1} \wedge \cdots \wedge e_{k_M}$ corresponding to the Young tableau Y.

Grassmannian's are projective algebraic varieties. Indeed one can embed them into projective space by Plücker coordinates which satisfy algebraic relations as described in the:

Proposition. *Define the application from the Grassmannian $Gr(M,V)$ to $P(\Lambda^M V)$, which associates to a M-dimensional frame the exterior product of its basis vectors, modulo arbitrary rescaling:*

$$\begin{aligned} \Theta \; : \quad & Gr(M,V) \longrightarrow P(\Lambda^M V) \\ & (x_{(1)}, \cdots, x_{(M)}) \longrightarrow \widehat{x} = x_{(1)} \wedge \cdots \wedge x_{(M)} \end{aligned} \tag{8.45}$$

Any element of $\Lambda^M V$ can be written $\sum_Y X_Y \; e_Y \equiv \sum X^{k_1 \cdots k_M} e_{k_1} \wedge \cdots \wedge e_{k_M}$.
i) If $\widehat{x}$ is factorizable on the $x_{(j)} = \sum_i e_i \; x^i{}_j$, then $X_Y = \det \left(x^{k_i}_{\;j} \right)_{i,j=1,\cdots,M}$. The set of quantities X_Y for all Y (up to rescaling) are called the Plücker coordinates of the given M–space. Under change of the basis $(x_{(1)}, \cdots, x_{(M)})$ the wedge product $\widehat{x}$ is changed by a scalar, hence the point in the projective space uniquely characterizes the M–space.
ii) The image in the projective space of the Grassmannian is the set of points whose coordinates $X^{k_1 \cdots k_M}$ satisfy the bilinear Plücker relations: for all $k_1, \cdots, k_{M-1}$ and $h_1, \cdots, h_{M+1}$:

$$\sum_{j=1}^{M+1} X^{k_1 \cdots h_j \cdots k_{M-1}} \; X^{h_1 \cdots \widehat{h}_j \cdots h_{M+1}} = 0 \tag{8.46}$$

where the index $\widehat{h}_j$ has to be omitted.

Proof. The first part is obvious since one can write:

$$\widehat{x} = \sum_{k_1 < \cdots < k_M} X^{k_1 \cdots k_M} \; e_{k_1} \wedge \cdots \wedge e_{k_M} \quad \text{with} \quad X^{k_1 \cdots k_M} = \det \left(x^{k_i}_{\;j} \right)_{i,j=1,\cdots,M}$$

An element of $\Lambda^M V$ corresponds to some M–plane if and only if it can be written in a factorized form $\widehat{v} = v_1 \wedge \cdots \wedge v_M$. We now characterize decomposable vectors. To do that consider $\widehat{v} \in \Lambda^M V$ and let us determine the minimal subspace W of V such that $\widehat{v} \in \Lambda^M W \subset \Lambda^M V$. For any $v^* \in V^*$ the dual of V one can consider the $(M-1)$ form $i(v^*)\widehat{v}$ the interior product of v^* on $\widehat{v}$. Let $V' = \{v^* \in V^* | i(v^*)\widehat{v} = 0\}$ the set of v^* annihilating our given $\widehat{v}$. Then we have $W = (V')^\perp = \{v \in V | < v, V' \geq 0\}$.

Indeed, let $w_1, \cdots, w_l$ be a basis of W completed by $u_{l+1}, \cdots, u_M$ to a basis of V. If one considers the dual basis $w_1^*, \cdots, w_l^*, u_{l+1}^*, \cdots, u_M^*$ of V^*, then V' is just the span of $u_{l+1}^*, \cdots, u_M^*$ because this is of dimension $(N-M)$ and u_j^* vanishes on W. In this basis one can write $\widehat{v} = \widehat{v}_1 + \widehat{v}_2 + \cdots$ where $\widehat{v}_1 \in \Lambda^M W$, while the other terms can be expanded on increasing powers of the u_k. Since $u_j^* \in V'$ we have $i(u_j^*)\widehat{v} = 0$ and since $i(u_j^*)$ does not act on the *coefficients* of these terms, all these coefficients must vanish hence $\widehat{v} = \widehat{v}_1$ as was to be shown.

If W is of dimension M, then $\Lambda^M W$ is of dimension 1 and $\widehat{v} \in \Lambda^M W$ is factorizable and $w \wedge \widehat{v} = 0$ for all $w \in W$. Conversely if $\widehat{v} \in \Lambda^M V$ is factorizable hence of the form $w_1 \wedge \cdots \wedge w_M$ then W is the span of the w_j and is of dimension M. Finally if $w \wedge \widehat{v} = 0$, $\forall w \in W$, $\widehat{v}$ can be expressed as wedge products of all basis elements of W, and possibly more, hence $\dim W \leq M$, but $\widehat{v} \in \Lambda^M W$, hence $\dim W = M$ and $\widehat{v}$ is factorizable.

We write the condition $w \wedge \widehat{v} = 0$, $\forall w \in W$ in a more convenient way by introducing more general interior products. Let $\xi \in \Lambda^{M-1} V^*$ any $(M-1)$–covector. We define the vector $i(\xi)\widehat{v} \in V$ by giving its contraction with an arbitrary covector $v^* \in V^*$: $\langle i(\xi)\widehat{v}, v^* \rangle = \langle \widehat{v}, \xi \wedge v^* \rangle$. We have $W = \{i(\xi)\widehat{v} | \xi \in\in \Lambda^{M-1} V^*\}$. Indeed remark that $\langle i(\xi)\widehat{v}, v^*\rangle = \pm\langle i(v^*)\widehat{v}, \xi\rangle$ which vanishes for all ξ if and only if $i(v^*)\widehat{v} = 0$. In the case where $v^* \in V'$ we see that $i(\xi)\widehat{v} \in W$. Moreover the orthogonal of the space of $i(\xi)\widehat{v}$ is $\{v^*|i(v^*)\widehat{v} = 0\} = V' = W^\perp$. Finally the factorizability condition can be written:

$$\widehat{v} \text{ factorizable iff } i(\xi)\widehat{v} \wedge \widehat{v} = 0,\ \forall \xi \in \Lambda^{M-1}V^*$$

To express this condition it is sufficient to vary ξ over a basis of $\Lambda^{M-1}V^*$, yielding C^N_{M-1} quadratic equations on $\widehat{v}$. They are the equations of the Grassmannian embedded into projective space. To get the explicit form eq.(8.46) one takes the dual basis e^*_j of the basis of V and consider a set of $\xi = e^*_{k_1} \wedge \cdots e^*_{k_{M-1}}$ with $k1 < \cdots < k_{M-1}$ forming a basis of $\Lambda^{M-1}V^*$. We have $i(\xi)\widehat{v} = \sum_{h_j} X^{k_1\cdots k_{M-1}h_j} e_{h_j}$. Similarly one writes $\widehat{v} = \sum X^{h_1\cdots\widehat{h_j}\cdots h_{M+1}} e_{h_1} \wedge\cdots\wedge e_{h_{M+1}}$. Then $i(\xi)\widehat{v}\wedge\widehat{v} = \sum_j X^{k_1\cdots h_j\cdots k_{M-1}} X^{h_1\cdots\widehat{h_j}\cdots h_{M+1}} e_{h_1}\wedge\cdots\wedge e_{h_j}\wedge\cdots\wedge e_{h_{M+1}}$ yielding eq.(8.46). ∎

Remark. It is easy to see that the Plücker relations are necessary relations. Let us demand that a vector of the form $z = \sum_m z^{(m)} x_{(m)}$ belongs to the hyperplane spanned by the $x_{(j)}$. This means

$$0 = z \wedge \widehat{x} = \sum_{h_1<\cdots<h_{M+1}} Z^{h_1\cdots h_{M+1}}\ e_{h_1} \wedge \cdots e_{h_{M+1}}$$

with

$$Z^{h_1\cdots h_{M+1}} = \sum_{j=1}^{M+1}\sum_{m=1}^{M} (-)^{j-1} x^{h_j}_{m} z^{(m)} X^{h_1\cdots\widehat{h_j}\cdots h_{M+1}} \tag{8.47}$$

We next choose $z^{(m)}$ as the minor of the following determinant:

$$X^{k_1\cdots k_{M-1}h_j} = \det\begin{pmatrix} x^{k_1}_{1} & \cdots & x^{k1}_{M} \\ \vdots & & \vdots \\ x^{k_{M-1}}_{1} & \cdots & x^{k_{M-1}}_{M} \\ x^{h_j}_{1} & \cdots & x^{h_j}_{M} \end{pmatrix} = \sum_m z^{(m)} x^{h_j}_{m}$$

Expanding this determinant with respect to the last line shows that the coefficient $z^{(m)}$ only depends on the indices $k_1, \cdots, k_{M-1}$ and not on the index h_j. Inserting this formula into eq.(8.47) gives the Plücker relations.

The basis of the interpretation of the fermionic constructions of the preceding sections is the interpretation of the fermionic vacuum as a Dirac sea:

$$|0\rangle = \psi_{\frac{1}{2}}\psi_{\frac{3}{2}}\cdots|-\infty\rangle \tag{8.48}$$

More precisely, we need to consider the formal limit of the Grassmannian of an infinite dimensional space V_∞. A basis of V_∞ will be given by e_s with $s \in \mathbb{Z}+\frac{1}{2}$. We equip V_∞ with the scalar product $(e_s, e_r) = \delta_{r+s,0}$.

The definition of $Gr(V_\infty)$ is such that its image under the map Θ consists in the semi–infinite wedge products:

$$|Y\rangle = e_{-s_n} \wedge \cdots \wedge e_{-s_1} \wedge e_{\frac{1}{2}} \wedge \cdots \wedge \widehat{e}_{r_1} \wedge \cdots \wedge \widehat{e}_{r_m} \wedge \cdots \tag{8.49}$$

for $0 < s_1 < \cdots < s_n$ and $0 < r_1 < \cdots < r_m$, where $\widehat{e}_{r_j}$ have been omitted in the infinite wedge product. Note that the tail of these wedge products are of the form $\wedge e_s \wedge e_{s+1} \wedge \cdots$ for sufficiently large positive s. The states (8.49) are thus finite perturbations of the state $|0\rangle = e_{\frac{1}{2}} \wedge e_{\frac{3}{2}} \wedge \cdots$, which is called the Dirac vacuum or Dirac sea. We say that the state $|Y\rangle$ contains n particles and m holes.

Proposition. *The dictionary between the semi-infinite wedge product and the fermionic language is as follows:*

$$\begin{aligned} |0\rangle &= e_{\frac{1}{2}} \wedge e_{\frac{3}{2}} \wedge e_{\frac{5}{2}} \wedge \cdots \\ \psi_r &= e_r \wedge \\ \psi^*_s &= e_s \neg \end{aligned} \tag{8.50}$$

where $e_r \wedge$ means taking the wedge product with e_r, and $e_s\neg$ means taking the interior product with e_s (using the scalar product $(e_r, e_s) = \delta_{r+s,0}$), which also means removal of e_{-s}.

Proof. It is a direct consequence of exterior calculus that the definition (8.50) form a representation of the anticommutation relations (8.11). ■

The main observation is the identification of the infinite dimensional Grassmannian and the orbit of the vacuum of $GL(\infty)$. Indeed we have:

$$g|0\rangle = (g \cdot e_{\frac{1}{2}}) \wedge (g \cdot e_{\frac{3}{2}}) \wedge (g \cdot e_{\frac{5}{2}}) \wedge \cdots \tag{8.51}$$

where we denote as in eq.(8.16) $g \cdot e_s = \sum_r e_r a_{rs}$. This is because on the Dirac sea g acts as $g|0\rangle = g\psi_{\frac{1}{2}}g^{-1} \cdot g\psi_{\frac{3}{2}}g^{-1} \cdots |-\infty\rangle$ so that g acts on the ψ_r by conjugation. Since a_{rs} is a band matrix this produces a combination of finite perturbations of the vacuum state hence belonging to the space of semi–infinite wedge products. Finally this expression of $g|0\rangle$ is factorized, so belongs to the infinite Grassmannian. The converse is clear.

Since in equation (8.35) g acts on the vacuum state $|0\rangle$ many different elements of $GL(\infty)$ produce the same τ-function. The ones which have a non trivial effect are those such that $g|0\rangle \neq |0\rangle$. Actually only half of $GL(\infty)$ effectively acts on the τ-functions. Indeed, considering eqs.(8.16,8.17) we see that elements g which correspond to upper triangular matrices a act trivially on the vacuum state $|0\rangle$. This is because if a^+ is upper triangular, ie. $a^+_{rs} = 0$ for $r < s$, and using $\psi_r|0\rangle = 0$ for $r > 0$ we get for $s > 0$:

$$(g\psi_s g^{-1})|0\rangle = \sum_{r \geq s > 0} a_{rs}\psi_r|0\rangle = 0$$

Similarly one verifies that $(g\psi^*_s g^{-1})|0\rangle = 0$ for $s > 0$. Clearly we also have $(g\psi_s g^{-1})g|0\rangle = (g\psi^*_s g^{-1})g|0\rangle = 0$ for $s > 0$. This shows that $|0\rangle$ and $g|0\rangle$ coincide since the set of fermionic operators annihilating them coincides.

As usual with integrable models, this fact is related to a factorization problem. This factorization problem is to factorize the matrix a_{rs} specifying the action of g on the fermions, cf. eqs.(8.16,8.17), into the product of an upper triangular matrix a^+_{rs} by a lower triangular matrix a^-_{rs}. In $GL(\infty)$ this corresponds to a factorization of g as $g = g_- g_+$.

Thus it is enough to restrict ourselves to elements $g = g_-$ of $GL(\infty)$ corresponding to lower triangular matrices $a = a^-$. They act on the fermions as :

$$\psi_s \quad \rightarrow \quad \psi^g_s = \sum_{r \leq s} \psi_r \, a_{rs} = a_{ss}\psi_s + a_{s-1,s}\psi_{s-1} + \cdots \tag{8.52}$$

We can now identify the flows in $Gr(V_\infty)$ corresponding to the commuting flows of the KP hierarchy. For a given $g \in GL(\infty)$, the state $g|0\rangle$ is represented by the semi-infinite wedge product (8.54). The temporal evolution of this state is defined as $g(t)|0\rangle = e^{H(t)}ge^{-H(t)}|0\rangle = e^{H(t)}g|0\rangle$ where we use $H(t)|0\rangle = 0$ since $H_n|0\rangle = 0$ for $n > 0$ and $H(t) = \sum_{n>0} t_n H_n$. The infinitesimal evolution $\partial/\partial_{t_n} g(t)|0\rangle = H_n g(t)|0\rangle$ defines the action of H_n on such a state.

Proposition. *Let S be the map on V_∞ defined by*

$$S \; : \; e_s \;\; \rightarrow \;\; e_{s+1}$$

We extend S to infinite wedge products by

$$S\; e_{i_1} \wedge e_{i_2} \wedge \cdots = (Se_{i_1}) \wedge e_{i_2} \wedge \cdots \; + \; e_{i_1} \wedge (Se_{i_2}) \wedge \cdots \; + \; \cdots$$

The action of H_n on the the states (8.49) defining the infinite Grassmannian is given by the shift operator S^n acting on semi-infinite wedge product.

Proof. Recall the commutation relations (8.27) which can be written $[H_n, \psi_r] = \psi_{n+r}$. Any point in the Grassmannian can be written $\psi_{i_1} \cdots \psi_{i_m}|l\rangle$ where $|l\rangle$ is a semi–infinite wedge product without holes $e_l \wedge e_{l+1} \wedge \cdots$ and the indices $i_1 < i_2 < \cdots < i_m < l$ may be negative. The action of H_n on this state is a sum of commutators, yielding the finite sum $\sum_j \psi_{i_1} \cdots \psi_{i_j+n} \cdots \psi_{i_m}|l\rangle$ which is exactly the action of S on this state. ∎

Alternatively, we see the group as acting on functions of z as in eq.(8.25), that is:

$$z^s \quad \rightarrow \quad A^g_s(z) = \sum_{r \leq s} a_{rs}\; z^r = a_{ss} z^s + a_{s-1,s} z^{s-1} + \cdots \tag{8.53}$$

(we assume that the diagonal elements of a_{rs} are equal to one). We may represent the basis vector e_s by the function z^s so that the vacuum state is:

$$|0\rangle = z^{\frac{1}{2}} \wedge z^{\frac{3}{2}} \wedge z^{\frac{5}{2}} \wedge \cdots$$

The group $GL(\infty)$ then acts on the vacuum as in eq.(8.51):

$$g|0\rangle = A^g_{\frac{1}{2}}(z) \wedge A^g_{\frac{3}{2}}(z) \wedge A^g_{\frac{5}{2}}(z) \wedge \cdots \tag{8.54}$$

In this representation it is clear that elements of $GL(\infty)$ corresponding to an upper triangular matrix a_{rs} act trivially on the vacuum.

Representing the vector e_s by z^s we can rewrite the evolution equations (8.56,8.57) describing the KP flows as evolution equations on functions of z. This leads yet to another representation of the τ-function:

Proposition. *The τ–function may be rewritten as an infinite dimensional determinant:*

$$\tau(t; g) = \det \left[\oint dz e^{\xi(t,z)} z^{-r-1} A^g_s(z) \right]_{r,s \geq \frac{1}{2}} \tag{8.55}$$

Proof. Recall the definition of the τ-function as $\tau(t; g) = \langle 0|e^{H(t)} g|0\rangle$. Let $\xi(t, z) = \sum_n z^n t_n$. Then the fermionic operators satisfy the following "evolution equations":

$$e^{H(t)} \psi(z) e^{-H(t)} \;\; = \;\; e^{\xi(t,z)}\; \psi(z) \tag{8.56}$$

$$e^{H(t)} \psi^*(z) e^{-H(t)} \;\; = \;\; e^{-\xi(t,z)}\; \psi^*(z) \tag{8.57}$$

The time evolution of $g|0\rangle$ under $e^{H(t)}$ amounts to the substitution $z^r \rightarrow z^r e^{\xi(t,z)}$ in eq.(8.54), thus

$$e^{H(t)} g|0\rangle = \left(e^{\xi(t,z)} A^g_{\frac{1}{2}}(z) \right) \wedge \left(e^{\xi(t,z)} A^g_{\frac{3}{2}}(z) \right) \wedge \left(e^{\xi(t,z)} A^g_{\frac{5}{2}}(z) \right) \wedge \cdots$$

To compute the expectation value $\langle 0|e^{H(t)} g|0\rangle$ we have to extract the component of $e^{H(t)} g|0\rangle$ on $|0\rangle = z^{\frac{1}{2}} \wedge z^{\frac{3}{2}} \wedge z^{\frac{5}{2}} \wedge \cdots$. This is given by the determinant (8.55). ∎

8.8 Schur polynomials.

In this section we use the basis of semi–infinite wedge products indexed by Young diagrams to investigate the structure of τ–functions. In particular we get polynomial τ–functions which turn out to be the famous Schur polynomials. They form a special basis of symmetric polynomials in k variables. These polynomials can also be viewed as characters of the group $GL(k)$.

Consider variables $u_1, u_2, \cdots, u_k$ where k is arbitrarily large. Symmetric polynomials of degree d on these variables may be expressed on several basis. To construct one of them consider the elementary symmetric polynomials h_j of degree j defined by:

$$\prod_{i=1}^{k} \frac{1}{1-u_i s} = \sum_{j=0}^{\infty} h_j(u) s^j$$

To any partition of d into at most k parts, namely $\lambda = (\lambda_1 \geq \lambda_2 \geq \cdots \lambda_k \geq 0)$ such that $d = \sum_i \lambda_i$ we associate the symmetric polynomial of degree d, $h_\lambda = \prod_i h_{\lambda_i}$ and the Young diagram with at most k lines, whose i^{th} has λ_i boxes. It is known that the H_λ when λ describe the set of Young diagrams with d boxes form an integral basis of the space of symmetric polynomials of degree d.

To the same Young diagram λ one can associate another symmetric polynomial, known as the Schur polynomial:

$$S_\lambda = \frac{\det u_j^{\lambda_i+k-i}}{\Delta}, \quad \Delta = \prod_{i<j}(u_i - u_j)$$

where $\Delta = \det u_j^{k-i}$ is the Vandermonde determinant. Note that the numerator and the denominator are antisymmetric polynomials, and that the denominator divides the numerator so that S_λ is a symmetric polynomial of degree d. It can also be shown that the Schur polynomials form an integral basis of symmetric polynomials of degree d indexed by the Young diagrams. Hence we can express S_λ in terms of the h_μ and this relation is given by the formula $S_\lambda = \det h_{\lambda_i+j-i}$. As will be shown below this famous formula is equivalent to the boson–fermion correspondence.

We can also relate these definitions to the definition of "elementary Schur polynomials" introduced in eq.(8.39). Indeed we have

$$h_i(u) = P_i(t), \quad \text{for } t_n = -\frac{1}{n}\sum_j u_j^n$$

This may be seen by writing:

$$\sum_i h_i(u) s^i = \exp\left(\sum_j \log(1-u_j s)\right) = \exp\left(-\sum_{i=1}^{\infty}(\frac{1}{i}\sum_j u_j^i) s^j\right) = \sum P_i(t) s^i$$

Proposition. *Let $|Y\rangle$ be the neutral state $|Y\rangle \equiv \psi_{-s_m}\cdots\psi_{-s_1}\cdot\psi^*_{-r_1}\cdots\psi^*_{-r_m}|0\rangle$, with $s_m > \cdots > s_1$ and $r_m > \cdots > r_1$, then,*

$$\langle 0|e^{H(t)}|Y\rangle = \det\Big(P_{n_a+a-b}(t)\Big) = S_Y \tag{8.58}$$

with n_a the number of boxes in the a^{th} line, counted from the bottom, for a Young diagram $Y = [n_a]$ with $L = (r_m + \frac{1}{2})$ lines and $C = (s_m + \frac{1}{2})$ columns, enclosed in a hook of width m. This means that there is no box in the rectangle of size $(L-m)\times(C-m)$ in the lower right corner. The first m lines have length $(s_{m-k+1} + k - \frac{1}{2})$, and the first m columns have length $(r_{m-k+1} + k - \frac{1}{2})$ for $k = 1, \cdots, m$.

Proof. Let us give the details of the proof of eq.(8.58) in the simplest case $m = 1$. Thus we consider the state $|Y\rangle = \psi_{-s}\psi^*_{-r}|0\rangle$. Using the definition (8.13) for the charged vacuum $|q\rangle$ we can rewrite it as:

$$|Y\rangle = \psi_{-s}\psi_{\frac{1}{2}}\cdots\psi_{r-1}|-r-\tfrac{1}{2}\rangle$$

Representing the fermionic operators as contour integrals, $\psi_k = \oint dz z^{k-\frac{1}{2}}\psi(z)$, allows us to write

$$\langle 0|e^{H(t)}|Y\rangle = \oint \prod_{a=1}^{r+\frac{1}{2}} z_a^{r-m_a-\frac{1}{2}} \; \langle 0|e^{H(t)}\psi(z_{r+1/2})\cdots\psi(z_1)|-r-\tfrac{1}{2}\rangle$$

with $(m_1,\cdots,m_{r+1/2}) = (1,2\cdots,r-\frac{1}{2},r+s)$. Using the evolution equations (8.56) and the fact that $H(t)$ annihilates $|-r-\frac{1}{2}\rangle$, we get:

$$\langle 0|e^{H(t)}|Y\rangle = \oint \prod_{a=1}^{r+\frac{1}{2}} z_a^{r-m_a-\frac{1}{2}} e^{\xi(t,z_a)} \; \langle 0|\psi(z_{r+1/2})\cdots\psi(z_1)|-r-\tfrac{1}{2}\rangle$$

The expectation value can be evaluated using the bosonization formulae (8.29,8.31):

$$\langle 0|\psi(z_{r+1/2})\cdots\psi(z_1)|-r-\tfrac{1}{2}\rangle = \prod_{a<b}(z_a-z_b)\;\prod_a z_a^{-r-1/2}$$

Thus, because the product $\prod_{a<b}(z_a-z_b)$ is the Vandermonde determinant $\det(z_a^{b-1})$, we get:

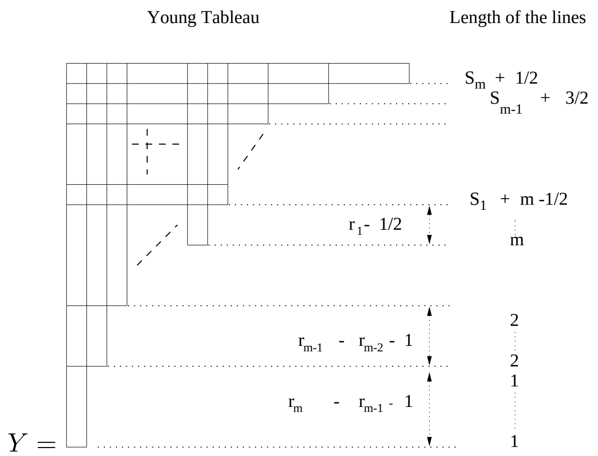
$$\begin{aligned}\langle 0|e^{H(t)}|Y\rangle \;&=\; \oint \prod_{a=1}^{r+\frac{1}{2}} z_a^{-m_a-1} e^{\xi(t,z_a)} \prod_{a<b}(z_a-z_b) \qquad (8.59)\\ &=\; \det\left(\oint dz z^{b-m_a-2}\; e^{\xi(t,z)}\right) = \det\Big(P_{m_a-b+1}(t)\Big)\end{aligned}$$

Since $m_a = n_a + a - 1$ this proves the result eq.(8.58) for $m=1$. The general cases are proved similarly starting from the representation of the states $|Y\rangle = \psi_{-s_1}\cdots\psi_{-s_m}\cdot\psi^*_{-r_1}\cdots\psi^*_{-r_m}|0\rangle$ as

$$|Y\rangle = \psi_{-s_1}\cdots\psi_{-s_m}\cdot\psi_{\frac{1}{2}}\cdots\widehat{\psi}_{r_k}\cdots\psi_{r_m-1}\widehat{\psi}_{r_m}|-r_m-\tfrac{1}{2}\rangle$$

where the operators $\widehat{\psi}_{r_k}$ have been omitted. ∎

This proposition allows to associate a Young diagram Y to the state $|Y\rangle \equiv \psi_{-s_m}\cdots\psi_{-s_1}\cdot\psi^*_{-r_1}\cdots\psi^*_{-r_m}|0\rangle$ in the following way:



For example in the simple case $m=1$ which has been considered in the proof of the proposition, Y is a hook diagram of width one, with the first line containing $s+1/2$ boxes, followed by $r-1/2$ lines with one box.

Proposition. *The Schur polynomial S_Y associated to the Young diagram Y can be written:*

$$S_Y \equiv \frac{\det\left(u_b^{n_a+a-1}\right)}{\Delta(u)} = \det\Big(P_{n_a+a-b}(t)\Big)$$

where $\Delta(u) = \prod_{a<b}(u_a - u_b)$ is the Vandermonde determinant.

Proof. We use the same technique as before, and limit ourselves to the case $m = 1$. In the expression (8.59) of $\langle 0|e^{H(t)}|Y\rangle$ we insert $\exp \xi(t,z) = \prod(1 - zu_i)^{-1}$. Next we use the Cauchy identity in the form:

$$\prod_{i,a} \frac{1}{1 - u_i z_a} = \frac{1}{\Delta(u)\Delta(z)} \det\left(\frac{1}{1 - z_a u_i}\right) \tag{8.60}$$

hence getting:

$$\langle 0|e^{H(t)}|Y\rangle = \frac{1}{\Delta(u)} \det\left(\oint dz \frac{z^{-m_a-1}}{1 - zu_i}\right) = \frac{\det\left(u_i^{m_a}\right)}{\Delta(u)} = S_Y(u)$$

since $m_a = n_a + a - 1$. ■

Remark. The polynomials $S_Y(u)$ introduced in the previous proposition are the characters of the group of linear transformations in the representation ρ_Y corresponding to the Young tableau Y. More precisely the character $\chi_Y(g) = \mathrm{Tr}\,(\rho_Y(g))$ for $g \in GL(k)$ is invariant under conjugation and only depends on the eigenvalues of g. One gets $\chi_Y(g) = S_Y(u)$ where $u_1, \cdots, u_k$ are the eigenvalues of g.

Any polynomial $f(t)$ can be seen as a syymetric polynomial of the variables u and conversely. This is because symmetric polynomials of u can be expressed on the Newton sums t. Since the Schur polynomials $S_Y(u)$ form a basis of symmetric polynomials, the function $f(t)$ can be expanded on the basis $\chi_Y(t)$ as $f(t) = \sum_Y \zeta_Y \chi_Y(t)$. The coefficients ζ_Y of the expansion can be recovered by:

$$\zeta_Y = \chi_Y(\widehat{\partial}_t) \cdot f(t)|_{t=0}$$

where $\chi_Y(\widehat{\partial}_t)$ are the differential operators obtained by replacing t_n by $\frac{1}{n}\frac{\partial}{\partial t_n}$ in the definition (8.58) of $\chi_Y(t_n)$.

To show that, we rewrite the Cauchy identity (8.60) in the form:

$$e^{H(t)} = \sum_Y S_Y(H)\chi_Y(t)$$

Indeed one expands each entry of the determinant $\det\,(1/(1 - u_i H_j))$ in powers of $u_i H_j$. One sees that for $l_1 > \cdots > l_k$ the coefficient of $H_1^{l_1} \cdots H_k^{l_k}$ is the determinant $\det\,(u_j^{l_i})$. By symmetry of the variables u and H one gets

$$\det\left(\frac{1}{1 - u_i H_j}\right) = \sum_l \det\,(u_j^{l_i}) \det\,(H_j^{l_i})$$

and one recovers the definition of Schur polynomials. Hence

$$\chi_Y(t) = \langle 0|e^{H(t)}|Y\rangle = \sum_{Y'} \chi_{Y'}(t)\langle 0|S_{Y'}(H)|Y\rangle$$

yielding $\langle 0|S_{Y'}(H)|Y\rangle = \delta_{YY'}$. We now compute:

$$\chi_Y(\widehat{\partial}_t) \cdot f(t) = \sum_{Y'} \langle 0|S_Y(H)e^{H(t)}|Y'\rangle \zeta_{Y'}$$

Letting $t \to 0$ yields the result.

In particular for a general τ–function of the form $\tau(t,g) = \langle 0|e^{H(t)}g|0\rangle$ we insert the closure relation $1 = \sum_Y |Y\rangle\langle Y|$ to get:

$$\tau(t;g) = \sum_Y \quad \zeta_Y(g) \quad \chi_Y(t) \tag{8.61}$$

with $\zeta_Y(g) = \langle Y|g|0\rangle$. The sum is over all Young diagrams. Conversely we have:

$$\zeta_Y(g) = \chi_Y(\widehat{\partial}_t) \cdot \tau(t;g)|_{t=0}$$

This allows to reconstruct the components of $g|0\rangle$ on the basis $|Y\rangle$ knowing the τ–function. We see that general τ–functions corresponding to any element of $GL(\infty)$ can be expressed on Schur polynomials. One gets just one Schur polynomial when g is choosen so that $g|0\rangle = |Y\rangle$.

Since $g|0\rangle$ belongs to the Grassmannian, i.e. is factorizable, the components $\zeta_Y(g)$ obey Plücker relations, and these relations characterize τ–functions among all functions of t. We can now identify the Plücker relation with eq.(8.36). We write formally an element of the infinite Grassmannian as:

$$|X\rangle = \sum_{k_1,k_2,\cdots,k_m,\cdots} X^{k_1,k_2,\cdots,k_m,\cdots} e_{k_1} \wedge e_{k_2} \wedge \cdots \wedge e_{k_m} \wedge \cdots$$

where the dots extend to infinity, but are regular after some index k_m (i.e. have no hole after k_m). The straight extension of the Plücker relation (8.46) to this setting reads:

$$\sum_j X^{k_1,k_2,\cdots,h_j,\cdots} X^{h_1,h_2,\cdots,\widehat{h_j},\cdots} = 0$$

In fact the sum over j is finite due to the regularity hypothesis. This can be rewritten using eqs.(8.50) in the elegant form:

$$\sum_j \psi_j \otimes \psi^*_{-j} |X\rangle \otimes |X\rangle = 0$$

which is exactly eq.(8.36). Of course this simple way of stating the Plücker relation relies on the infinite number of indices. As we have already shown the condition (8.36) translates into Hirota's equations on the τ–function which are therefore equivalent to the equations of the Grassmannian in the infinite wedge space.

8.9 The Segal–Wilson approach.

Up to now we have used the description of the Grassmannian imbedded into projective space by Plücker coordinates, using the fermionic language, which has the advantage of regularizing potential infinities by normal ordering. In this section we shall look at the Grassmannian as the space of appropriate linear subspaces of an infinite dimensional vector space. We start again from the space V_∞ with basis $z^n, n \in \mathbb{Z}$ which can be seen as the space of functions on the unit circle $|z| = 1$. To a function on the circle we associate its Fourier expansion $f(z) = \sum_n a_n z^n$. The space V_∞ is turned into a Hilbert space H by introducing the L^2 norm $\int |f|^2$ or what amounts to the same $\sum_n |a_n|^2$, and can be decomposed as direct sum of subspaces $H^\pm$, where H^+ is generated by $z^n, n \geq 0$ and H^- by $z^n, n < 0$. We consider the set of subspaces W which are comparable to H^+ in the following sense:

$$W \in Gr \quad \text{if and only if} \quad pr_+ : W \to H^+ \text{ is Fredholm and } pr_- : W \to H^- \text{ is compact} \tag{8.62}$$

The fact that the projection pr_- is compact means (for Hilbert spaces) that it is a limit (in the norm topology of maps) of operators with finite dimensional images. In our fermionic language these finite dimensional images correspond to states with a finite number of particles. The fact that the projection pr_+ is Fredholm means that the kernel and the cokernel of pr_+ are finite dimensional.

Let us illustrate these conditions by an example: assume that W is spanned by $e_{-2}, e_1 + e_{-1}, e_3, e_4, \cdots$. Then pr_+W is spanned by $e_1, e_3, e_4, \cdots$ and is of codimension 2 in H^+ since e_0 and e_2 are missing. Its kernel is spanned by e_{-2} and is of dimension 1. Similarly pr_-W is spanned by e_{-1}, e_{-2} and is of finite dimension 2. Under the Plucker embedding W goes to $e_{-2} \wedge (e_1 + e_{-1}) \wedge e_3 \wedge e_4 \cdots$ which expands on two semi–infinite wedge products. These two terms have the property $\#$holes$-\#$particles $= 1 =$ dim Coker $pr_+ -$dim Ker pr_+ which is called the index of the Fredholm operator pr_+. Notice that this is also the common charge of all the above states. This is a general fact expressing the stratification of the Grassmannian by the charge. In general we have Ker $pr_+ \in$ Im pr_-, but the first one is of finite dimension while the second one may become infinite dimensional under the limiting procedure yielding a general compact operator.

Finally recall some properties of compact and Fredholm operators. If u is compact and v is continuous (i.e. bounded) then uv and vu are compact. A Fredholm operator is defined to be continuous, hence its product by a compact operator is compact. Finally the sum of a Fredholm operator and a compact operator is Fredholm, and the product of two Fredholm operators is Fredholm. This allows to consider the group $GL(\infty)$ of matrices having the block form on the decomposition $H = H^+ \oplus H^-$:

$$h = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad a, d \text{ Fredholm and } b, c \text{ compact}$$

Obviously the product of two such elements is of the same form. Moreover a group element h acts on the Grassmannian, moving W to hW where the projections for hW are given by:

$$\begin{pmatrix} pr'_+ \\ pr'_- \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} pr_+ \\ pr_- \end{pmatrix}$$

and have therefore the required properties.

Two subgroups are of particular interest denoted $\Gamma^\pm$. An element of Γ^+ is given by the multiplication by an L^2 non vanishing function f on the unit circle with only positive powers of z, hence extending to an analytic function f in the unit disk normalized by $f(0) = 1$ It can be represented in the block form as:

$$h \in \Gamma^+, \quad h = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$$

where a and d are invertible, hence Fredholm and b is compact. Indeed let us write $f(z) = 1 + a_1 z + a_2 z^2 + \cdots$ and consider its action by multiplication on $g(z) = \sum_k b_k z^k$. If g has only positive powers, fg has only positive powers, moreover $1/f$ can be expanded on positive powers, hence a is invertible. It is continuous by Schwartz inequality so is Fredholm. If g has only negative powers of z, we remark that

$$f(z) z^{-n} = z^{-n} + a_1 z^{-n+1} + \cdots + a_{n-1} z^{-1} + \sum_{j \geq 0} a_{n+j} z^j$$

We see that the H^- part induces a triangular system with 1 on the diagonal hence is invertible, so that d is Fredholm. To show that b is compact consider the truncation $f_N = \sum_{j=0}^N a_j z^j$ which is such that $\dim\{pr_+(f_N g) | g \in H^-\} < N$ i.e. the corresponding b_N is of finite rank. We have $||f - f_N|| \to 0$ when $N \to \infty$ because $||(f - f_N) g|| \leq ||g|| \sum_{j=N}^\infty |a_j|^2$ by Schwartz inequality. Similarly we define Γ^- as the group of multiplication by non vanishing L^2 functions with only negative powers of z hence extending to $|z| > 1$ analytically.

In the following we restrict ourselves to spaces W which are transversal to H^-, i.e. such that $pr_+ : W \to H^+$ is an isomorphism. Such W have charge 0, and are "small" deformations of the vacuum $|0\rangle$. In this case one can define an operator $A : H^+ \to H^-$ as follows. To any $f \in H^+$ one can associate a unique element $w \in W$ such that $f = pr_+(w)$. We denote $w = pr_+^{-1}(f)$, and form $A(f) = pr_-(w)$ so that A is a compact operator.

We now want to understand the τ–function in this context. Let us fix an element W of the Grassmannian and vary h through Γ^+. In fact we are interested in the time evolution represented by multiplication by $h = \exp(\sum_{i\geq 0} t_i z^i)$. We define:

Definition. *Assuming that W is transverse to H^- so that $(pr_+|_W)^{-1}$ exists we set:*

$$\tau_W(h) = \det\left(h\, pr_+\, h^{-1}(pr_+|_W)^{-1}\right) = \det\left(1 + a^{-1}bA\right), \quad \text{where} \quad h^{-1} = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$$

As a matter of fact if $f \in H^+$ we have $(pr_+|_W)^{-1}(f) = w = f + Af$ with $Af \in H^-$. Then $pr_+\, h^{-1}(pr_+|_W)^{-1}(f) = af + bAf \in H^+$ and finally $h\, pr_+\, h^{-1}(pr_+|_W)^{-1}(f) = (1 + a^{-1}bA)f$. We have here an operator from H^+ to H^+ of the form $1 +$ compact and we want to take its determinant. This is possible if the operator $a^{-1}bA$ is of trace class, which is ensured if h is sufficiently regular on the unit circle.

It is important to show that this definition agrees with the previous construction in eq.(8.35). Recall the expression (8.54) of the Plücker embedding of some element W of the Grassmannian. Here the map $A^g_s(z)$ has to be identified with $(pr_+|_W)^{-1}(z^s)$ in the present language. To conform with the present notation the index s will be taken to be integer instead of half–integer and we shall count $0, 1, \cdots$ as the positive indices. Hence $A_0(z) \wedge A_1(z) \wedge \cdots$ represents the Plücker embedding of W. By definition the action of h^{-1} on W is $A_s \to e^{\xi(t,z)} A_s$ which under the Plücker embedding is represented by multiplication by $e^{H(t)}$ due to eq.(8.56). Next the projection P_+ becomes $|0\rangle\langle 0|$ and the multiplication by h is achieved by multiplying by $e^{-H(t)}$, so that:

$$h\, pr_+\, h^{-1}(pr_+|_W)^{-1}\,(H^+) \overset{\text{Plucker}}{\longrightarrow} e^{-H(t)}|0\rangle\langle 0|e^{H(t)} A_0(z) \wedge A_1(z) \wedge \cdots$$

Taking the scalar product with $\langle 0|$ produces the determinant of the operator $h\, pr_+\, h^{-1}(pr_+|_W)^{-1}$. Since $\langle 0|e^{-H(t)}|0\rangle = 1$ due to normal ordering of H, we reproduce exactly eq.(8.35). Note the similarity of this reasoning with eq.(8.55). Note that the operator $a : H^+ \to H^+$ is triangular with 1 on the diagonal hence one can set $\det a = 1$. If it was allowed to write $\det MN = \det M \cdot \det N$ for infinite dimensional matrices, one could content ourselves by defining $\tau_W = \det(a + bA)$, so the above definition performs a regularization of this too naive expression.

However when M, N are of the form $1 + m, 1 + n$ with m and n of trace class one is allowed to use this product formula. Moreover in this case we have $\tau_W(h) = \det(1 + bAa^{-1})$. In particular one obtains:

$$\tau_W(h_1 h_2) = \tau_W(h_1)\tau_{h_1^{-1}W}(h_2) \tag{8.63}$$

Indeed,

$$\begin{aligned} \tau_W(h_1)\tau_{h_1^{-1}W}(h_2) &= \det(h_2 pr_+ h_2^{-1}(pr_+|_{h_1^{-1}W})^{-1}) \cdot \det(pr_+ h_1^{-1}(pr_+|_W)^{-1} h_1) \\ &= \det\left(h_1 h_2 pr_+ h_2^{-1}(pr_+|_{h_1^{-1}W})^{-1} pr_+ h_1^{-1}(pr_+|_W)^{-1}\right) = \tau_W(h_1 h_2) \end{aligned}$$

This is because if $f \in H^+$ then $w = (pr_+|_W)^{-1} f \in W$ so that $h_1^{-1} w \in h_1^{-1} W$. Under pr_+ this gives some $g \in H^+$ which under $(pr_+|_{h_1^{-1}W})^{-1}$ reproduces $h_1^{-1} w = h_1^{-1}(pr_+|_W)^{-1} f$.

In this setting $h \in \Gamma^+$ and such that $h^{-1}W$ transverse to H^- it is very natural to define the Baker function as follows:

Definition. *The Baker function $\psi_W(h, z)$ is such that $h^{-1}\psi_W(h, z)$ is the inverse image under $pr_+|_{h^{-1}W}$ of the function $1 \in H^+$.*

This means that $h^{-1}\psi_W(h, z) \in h^{-1}W$ and can be written as $1 + \sum_{i=1}^{\infty} a_i(h) z^{-i}$, i.e. the Baker function is the unique function such that $\psi_W(h, \cdot) \in W$ having the form:

$$\psi_W(h, z) = h(z)\left(1 + \sum_{i=1}^{\infty} a_i(h) z^{-i}\right)$$

Taking for $h(z)$ the time–evolution operator $\exp(\sum_{i>0} t_i z^i$ we see that ψ_W has an essential singularity at $z = \infty$.

One can give an explicit expression of the Baker function in terms of the τ–function, namely this is Sato's formula:

Proposition. *We have*

$$\psi_W(t, \zeta) = e^{\xi(t,\zeta)} \frac{\tau_W(t - [\zeta^{-1}])}{\tau_W(t)} \tag{8.64}$$

where the group element $h \in \Gamma^+$ is written as $h = e^{\xi(t,z)}$ while the notation $t - [\zeta^{-1}]$ refers to the same group element with t_n replaced by $t_n - z^{-n}/n$.

Proof. The group element $h(t-[\zeta^{-1}])$ is given by $\exp(\sum_{n>0}(t_n - \zeta^{-n}/n)z^n = h(t)q_\zeta$ where $q_\zeta = 1 - z/\zeta \in \Gamma^+$. Since $\tau_W(hq_\zeta) = \tau_W(h)\tau_{h^{-1}W}(q_\zeta)$ by eq.(8.63) we have to show that $e^{-\xi(t,\zeta)}\psi_W(t,\zeta) = \tau_{h^{-1}W}(q_\zeta)$ As a function of ζ the left–hand side is $(pr_+|_{h^{-1}W})^{-1}(1)$ by definition of the Baker function. The right–hand side is a τ function of $h^{-1}W$, hence one can replace $h^{-1}W$ by W and we have to show the equality of the two functions of ζ: $(pr_+|_W)^{-1}(1) = \tau_W(q_\zeta) = \det(1 + a^{-1}bA)$ where the operators a and b appear in the block representation of q_ζ^{-1} and A is the operator induced by W as above. Since $q_\zeta^{-1}(z) = \sum_{n\geq 0} z^n/\zeta^n$, we have:

$$b(z^{-n}) = pr_+\left(\frac{z^{-n}}{(1 - z/\zeta)}\right) = \zeta^{-n} q_\zeta^{-1}(z)$$

while the action of a^{-1} on this element of H^+ is simply represented by its multiplication by $q_\zeta(z)$. Hence for any element $f \in H^-$ we have $a^{-1}bf = f(\zeta) \cdot 1 \in H^+$, i.e. the operator is of rank 1, so that $\det(1 + a^{-1}bA) = 1 + \mathrm{Tr}(a^{-1}bA)$. Since the image is spanned by the function 1 it is enough to compute its action on 1 to get the trace. But by definition of A it sends the basis element 1 on the negative power part of $(pr_+|_W)^{-1}(1)$ denoted by $f(z)$, so that $1 + \mathrm{Tr}(a^{-1}bA) = 1 + f(\zeta)$. ■

Let us see how the algebro–geometric construction fit in this setting. Let Σ be a compact Riemann surface of genus g and $\mathcal{L}$ be a line bundle on Σ, see chapter [15]. Fix a puncture x_∞ on Σ and let z^{-1} be a local parameter around x_∞ such that the disk D_∞ defined by $|z| > 1 - \epsilon$ is an open neighbourhood of x_∞ ($\epsilon > 0$ small) and the line bundle $\mathcal{L}$ admits a trivialization on D_∞. Hence sections of $\mathcal{L}$ locally appear as functions on D_∞. We shall also consider the open set Σ_0 which is the complement of the disk $|z| \leq 1 + \epsilon$, so that $\Sigma_0 \cap D_\infty$ is the annulus $1 - \epsilon < |z| < 1 + \epsilon$ and contains the unit circle S^1 defined by $|z| = 1$. The set of these data (puncture, local parameter, and trivialization of $\mathcal{L}$) is called a set of Krichever data. To this set one associates an element W of the Grassmannian: W is the L^2 closure of the set of analytic functions on the annulus, for some ϵ, which extend to holomorphic sections of $\mathcal{L}$ on Σ_0.

To show that pr_- is compact consider a function $f(z)$ on S^1 which belongs to W hence can be extended on Σ_0 and consider for $0 < \lambda < 1$ with λ close to 1 the function $f_\lambda(z) = f(\lambda z)$, and its values on S^1. Clearly $f_\lambda|_{S^1}$ depends continously on $f|_{S^1}$. We then project $f_\lambda|_{S^1}$ on H^- getting a function $(f_\lambda)_-$ which also depends continously on f. Finally one recovers $pr_-(f)$ by forming $(f_\lambda)_-(\lambda^{-1}z)$. But this last scaling is compact since only negative powers of z appear and $(1/\lambda)^{-n} \to 0$ when $n \to \infty$.

To show that pr_+ is Fredholm one shows in fact that $\mathrm{Ker}(pr_+ : W \to H^+) = H^0(\Sigma, \mathcal{L}_\infty)$ and that $\mathrm{Coker}(pr_+ : W \to H^+) = H^1(\Sigma, \mathcal{L}_\infty)$, where $\mathcal{L}_\infty = \mathcal{L} - [x_\infty]$ is the difference of the line bundle $\mathcal{L}$ and the point bundle at x_∞, i.e. a section of $\mathcal{L}_\infty$ arises from a section of $\mathcal{L}$ vanishing at x_∞. They are thus of finite dimension. Recall that $H^0(\Sigma, \mathcal{L}_\infty)$ is the set set of global holomorphic sections of $\mathcal{L}_\infty$. A function belongs to $\mathrm{Ker}(pr_+ : W \to H^+)$ if and only if it has only strictly negative powers of z hence extends to the interior of D_∞ and vanish at x_∞. By definition it extends to Σ_0 so providing a global section of $\mathcal{L}_\infty$.

We now show that $\mathrm{Coker}(pr_+ : W \to H^+) = H^1(\Sigma, \mathcal{L}_\infty)$. It is a well–known fact that the *non–compact* Riemann surface Σ_0 has no sheaf cohomology $H^1(\Sigma_0, \mathcal{L}_\infty) = 0$ and it is obviously the same for D_∞. Hence any non vanishing element in $H^1(\Sigma, \mathcal{L}_\infty)$ comes from some analytic function $\phi_{0\infty}$ on the annulus $\Sigma_0 \cap D_\infty$ which cannot be written as $\phi_0 - \phi_\infty$ where ϕ_0 extends to an analytic section of $\mathcal{L}$ on Σ_0 and ϕ_∞ is an analytic function on D_∞ vanishing at x_∞. But the part of $\phi_{0\infty}$ with strictly negative powers of z extends uniquely to a function ϕ_∞ vanishing at x_∞. So $H^1(\Sigma, \mathcal{L}_\infty)$ is isomorphic to the set of analytic functions on S^1 with only non–negative powers of z modulo those which extend to sections of $\mathcal{L}$. This is precisely the definition of $\mathrm{Coker}(pr_+ : W \to H^+)$. It is easy to see that this result remains true when one takes the limit $\epsilon \to 0$ and the L^2 closure of the considered functions on S^1.

Thus we arrived at the conclusion that the Fredholm index of pr_+ is given by:

$$\mathrm{Index}\,(W) = \dim H^0(\mathcal{L}_\infty) - \dim H^1(\mathcal{L}_\infty) = 1 - g + c(\mathcal{L}_\infty) = c(\mathcal{L}) - g$$

since $c(\mathcal{L}_\infty) = c(\mathcal{L}) - 1$ (see Chapter [15]), $c(\mathcal{L})$ is the Chern class of $\mathcal{L}$ and g is the genus of the Riemann surface. In particular in the interesting case where W is transverse to H^- there is no Kernel and no Cokernel, hence the index vanishes which needs $c(\mathcal{L}) = g$. Remark that in this case for generic position of the data W is indeed transverse to H^-. In practice one takes a puncture x_∞ and a set of g points in generic position and one considers meromorphic functions with poles at these g points and an essential singularity at x_∞ therefore Baker functions. These data define a unique W transverse to H^- in the Grassmannian by the above construction, and the Baker function identifies with the above definition.

8.10 From fermions to pseudo-differential operators.

The link between the vertex operator and the pseudo-differential operator approaches to the KP hierarchy can be established using the bilinear identity. We prefer to present it using the Grassmanian formulation and the determinant formula (8.55) for the τ-function. This connection is based on the fundamental differential property of the Schur polynomials:

$$\frac{\partial}{\partial t_n} P_k(t) = \partial^n P_k(t) = P_{k-n}(t) \tag{8.65}$$

with $\partial = \frac{\partial}{\partial t_1}$; It follows from their definition (8.39). In this section we will freely use the notion of pseudo-differential operators explained in Chapter [9].

Recall the formula (8.55) for the τ-function. Since $z^{-n} e^{\xi(t,z)} = \partial^{-n} e^{\xi(t,z)}$, it can be written as:

$$\tau(t; g) = \det \Big[F^g_{m;n}(t) \Big]_{m,n \geq 1}$$

with

$$F^g_{m;n}(t) = \partial^{-n} F^g_{m;0}(t) = \oint dz e^{\xi(t,z)} z^{-n-1/2} A^g_{m-1/2}(z)$$

Define now a pseudo-differential operator Φ by adding one line and one column to this determinant.

Proposition. *Let Φ be the pseudo-differential operator defined by:*

$$\Phi \cdot f \equiv \Big[1 + \sum_{n>0} w_n(t) \partial^{-n} \Big] \cdot f = \frac{1}{\tau(t; g)} \det \begin{pmatrix} \partial^{-n} f \\ \partial^{-n} F^g_{m;0}(t) \end{pmatrix}_{\substack{m \geq 1 \\ n \geq 0}} \tag{8.66}$$

Let

$$w(t,k) = \Phi e^{\xi(t,k)} = \widehat{w}(t,k) e^{\xi(t,k)}; \quad \widehat{w}(t,k) = 1 + \sum_{n>0} w_n(t) k^{-n}, \quad \xi(t,k) = \sum_{n=1}^{\infty} t_n k^n \tag{8.67}$$

then

$$w(t,k) = \frac{\tau(t-[k];g)}{\tau(t;g)} = \frac{\langle 1|e^{H(t)}\psi(k)g|0\rangle}{\tau(t;g)} \tag{8.68}$$

where $[k] = (\cdots, \frac{k^{-n}}{n}, \cdots)$. *Furthermore the KP flows acts on* Φ *as:*

$$\frac{\partial}{\partial t_n}\Phi = (\Phi\partial^n\Phi^{-1})_+\Phi - \Phi\partial^n = -(\Phi\partial^n\Phi^{-1})_-\Phi \tag{8.69}$$

Proof. By definition $\tau(t;g)\widehat{w}(t;k)$ is the following determinant:

$$\tau(t;g)\widehat{w}(t,k) = \det\begin{pmatrix} 1 & k^{-1} & k^{-2} & \cdots \\ F^g_{1,0} & \partial^{-1}F^g_{1,0} & \partial^{-2}F^g_{1,0} & \cdots \\ F^g_{2,0} & \partial^{-1}F^g_{2,0} & \partial^{-2}F^g_{2,0} & \cdots \\ \vdots & \vdots & \vdots & \cdots \end{pmatrix}$$

Substracting k^{-1} times the j^{th} column to the $(j+1)^{th}$ column allows to reduce the first line to $(1,\ 0,\ 0,\cdots)$. Expanding then the determinant with respect to this first line gives:

$$\tau(t;g)\widehat{w}(t;k) = \det\left[(1-k\partial)\partial^{-n}F^g_{m,0}\right]_{n,m\geq 1}$$

On the other hand, since $e^{\xi(t-[k],z)} = (1-zk^{-1})e^{\xi(t,z)}$ for $|z|<|k|$, we have:

$$F^g_{m,n}(t-[k]) = (1-k^{-1}\partial)F^g_{m,n}(t)$$

Inserting this in the previous formula proves the result.
Let us now prove the evolution equation (8.69). We will do it by looking at the evolution equation for an operator which provides a finite order approximation of Φ. Namely we introduce an order N differential operators D_N, ie $D_N = \partial^N + d_1(t)\partial^{N-1} + \cdots + d_N(t)$, such that:

$$\Phi = \lim_{N\to\infty} \quad D_N\partial^{-N}$$

D_N is defined by truncating the determinant (8.55) to a finite dimensional determinant, namely:

$$D_N\cdot f = \frac{1}{\widehat{\tau}_N(t;g)}\det\begin{pmatrix} \partial^N f & \partial^{N-1}f & \cdots & f \\ F^g_{1,0} & \partial^{-1}F^g_{1,0} & \cdots & \partial^{-N}F^g_{1,0} \\ \vdots & \vdots & & \vdots \\ F^g_{N,0} & \partial^{-1}F^g_{N,0} & \cdots & \partial^{-N}F^g_{N,0} \end{pmatrix}$$

with $\widehat{\tau}_N(t;g) = \det\Big(\partial^{-n}F^g_m\Big)_{n,m=1,\cdots,N}$. We are going to prove that:

$$\frac{\partial}{\partial t_n}D_N = (D_N\partial^n D_N^{-1})_+D_N - D_N\partial^n$$

Since both side of this equation are differential operators of order $(N-1)$, to prove this equality it is enough to check it on N linearly independent functions. We choose these functions to be the functions $F^g_{p,N} = \partial^{-N}F^g_{p,0}$, for $p=1,\cdots,N$. They span the kernel of D_N: $D_NF^g_{p,N}=0$. Thus,

$$\begin{aligned} \frac{\partial D_N}{\partial t_n}\cdot F^g_{p,N} &= -D_N\cdot\frac{\partial F^g_{p,N}}{\partial t_n} \\ &= \Big((D_N\partial^n D_N^{-1})_+D_N - D_N\partial^n\Big)\cdot F^g_{p,N} \end{aligned}$$

where to go to the last line we have used $\frac{\partial F^g_{p,N}}{\partial t_n} = \partial^n F^g_{p,N}$ and the fact that $F^g_{p,N}$ is on the kernel of D_N. This proves the evolution equation for D_N and hence, by taking the large N limit, it proves the evolution equation of Φ. ∎

References

[1] M.Jimbo, T.Miwa *"Solitons and Infinite Dimensional Lie Algebras."* RIMS 19 (1983) 943-1001.

[2] R.Hirota, *" Exact solution of the Korteweg-de Vries equation for multiple collisions of solitons."* Phys.Rev.Lett. 27 (1971) 1192.

[3] M. Sato, Y. Sato. *"Solitons equations as dynamical systems on infinite dimensional Grassman manifolds"*. Lect. Notes in Num. Appl. Anal. 5 (1982), p. 259. or Proc. U.S.-Japan Seminar Nonlinear PDE in Applied Science, Tokyo 1982, Ed. Lax, Fujita, pp.259-271, North Holland/Kinokuniya,

[4] G.Segal, G.Wilson, *"Loop Groups and Equations of KdV Type."* Publ.Math. I.H.E.S. 61 (1985) p.5-65.

[5] V.G. Kac and Wakimito, ...

[6] C. Itzykson and J.-B. Zuber, *"Combinatorics of the modular group II: the Kontsevich integrals."*, Int. J. Mod. Phys. A7 (1992) 5661-5705.

[7] Dickey, ..., J. Math. Phys....

Chapter 9

The KP hierarchy.

At the end of the last Chapter we showed that the equation of the KP hierarchy can be written as:

$$\partial_{t_n}\Phi = -(\Phi\partial^n\Phi^{-1})_-\Phi$$

where Φ is a pseudo–differential operator. This is identical to the standard form of the equations of an integrable hierarchy, but we are dealing here with the algebra of pseudo–differential operators, instead of a loop algebra. In this Chapter, we explain this setting and investigate the corresponding hierarchy. We also present the reduction of this hierarchy to the generalized KdV equations, and investigate their Poisson structures.

9.1 The algebra of pseudo–differential operators.

The algebra of differential operators is the algebra generated by $\mathbb{C}$–valued functions of one variable x and the derivation symbol ∂, with the usual Leibnitz rule, $\partial.a = a.\partial + (\partial a)$, where (∂a) means the function $(\partial_x a)(x)$. This defines the multiplication law between the symbol ∂ and the functions. An element in this algebra is a finite sum $A = \sum_{i=0}^N a_i\partial^i$, with N finite but arbitrary. The coefficients a_i are functions of x. To define the algebra of pseudo–differential operators we extend this algebra by introducing the "integration" symbol ∂^{-1} and its powers, with the following algebraic rules:

$$\begin{aligned}\partial^{-1}\partial &= \partial\partial^{-1} = 1\\ \partial^{-1}a &= \sum_{i=0}^{\infty}(-1)^i(\partial^i a)\partial^{-i-1}\end{aligned}\qquad(9.1)$$

These rules are taken as definition of the multiplication law for ∂^{-1}. They are motivated by the rule for integration by part. Indeed, symbolically we have :

$$\begin{aligned}\partial^{-1}(a.u) = \int dx\ au = \int dx\ a\partial(\partial^{-1}u) &= a\partial^{-1}u - \int dx\ (\partial a)\partial^{-1}u\\ &= a\partial^{-1}u - (\partial a)\partial^{-2}u + \int dx\ (\partial^2 a)\partial^{-2}u\end{aligned}$$

and so on. The algebra of pseudo–differential operators consists of elements which are semi-infinite sums of the form $A = \sum_{-\infty}^N a_i\partial^i$. They form an associative algebra with a unit. In fact the product of two such symbols can be uniquely brought to the same form (with ∂^i on the right) using the relations (9.1). Here we don't have to deal with the convergence questions involved in this reshuffling, since only a finite number of terms appear at each order in ∂^i. The algebra possesses a natural anti–homomorphism, which we called the formal adjoint, defined by: $(a\, \partial^i)^* = (-\partial)^i\, a$, for any function a. We summarize these facts in the following definition:

Let $\mathcal{P} = \left\{ A = \sum_{-\infty}^{N} a_i \partial^i \right\}$ be the set of formal pseudo-differential operators in one variable. We define the two subalgebras $\mathcal{P}_+ = \left\{ A = \sum_{i=0}^{N} a_i \partial^i \right\}$, the subalgebra of differential operators, and $\mathcal{P}_- = \left\{ A = \sum_{-\infty}^{-1} a_i \partial^i \right\}$, the subalgebra of integral operators. We have the direct sum decomposition of $\mathcal{P}$ as a vector space:

$$\mathcal{P} = \mathcal{P}_+ \oplus \mathcal{P}_-$$

Notice that $\mathcal{P}$ is naturally a Lie algebra. $\mathcal{P}_+$ and $\mathcal{P}_-$ are Lie subalgebras, but $\mathcal{P}_+$ and $\mathcal{P}_-$ do not commute. On $\mathcal{P}$ there exists a natural linear form:

Proposition. *The algebra $\mathcal{P}$ is equipped with a linear form, called the Adler trace, denoted $<\ \ >$, which is defined for any element $A = \sum_{-\infty}^{N} a_i \partial^i$ by:*

$$< A >= \int dx\ a_{-1}(x) \tag{9.2}$$

This linear form satisfies the fundamental trace property $< AB >=< BA >$, hence defines an ad*–invariant non–degenerate scalar product on $\mathcal{P}$ by: $< A, B >=< AB >$. Using this bilinear form we have the duality: $\mathcal{P}_+^* = \mathcal{P}_-$.*

<u>Proof</u>. There is a unique way of writing a pseudo–differential operator in the form $A = \sum_{-\infty}^{N} a_i \partial^i$, i.e. with all ∂^i at the right. We define the residue of A, denoted $\mathrm{Res}\, A$ as the coefficient of ∂^{-1}. This is a function of x and we have:

$$\mathrm{Res}\, A = a_{-1}(x), \quad < A >= \int dx\, \mathrm{Res}\, A$$

Let us prove the trace property. It is sufficient to verify it on operators of the form $A = a\, \partial^k$ and $B = b\, \partial^j$. If k and j are both positive or both strictly negative, we clearly have $< AB >=< BA >= 0$. Thus we take $A = a\, \partial^k$ and $B = b\, \partial^{-j-1}$ with $k,\ j \geq 0$. Using the relation:

$$\partial^{-j-1} a = \sum_{v=0}^{\infty} (-1)^v \begin{pmatrix} j+v \\ v \end{pmatrix} (\partial^v a)\, \partial^{-j-1-v}, \tag{9.3}$$

which is shown by induction, using the definition relation (9.1), and the identity

$$\begin{pmatrix} j+1+\mu \\ \mu \end{pmatrix} = \sum_{\nu=0}^{\mu} \begin{pmatrix} j+\nu \\ \nu \end{pmatrix}$$

we get:

$$< BA >= (-1)^{k-j} \begin{pmatrix} k \\ j \end{pmatrix} \int dx\ \left(b\, \partial^{k-j} a\right) = \begin{pmatrix} k \\ j \end{pmatrix} \int dx\ \left(a\, \partial^{k-j} b\right)$$

Similarly, using the Leibnitz rule, we compute: $AB = \sum_{v=0}^{\infty} \begin{pmatrix} k \\ v \end{pmatrix} (a\, \partial^v b)\, \partial^{k-j-v-1}$. Therefore, $< AB >$ is given by:

$$< AB >= \begin{pmatrix} k \\ j \end{pmatrix} \int dx\ \left(a\, \partial^{k-j} b\right)$$

Clearly, $< AB >$ and $< BA >$ coincide.
The invariance of the scalar product means $< A, [B, C] >=< [A, B], C >$. It follows from the trace property. We already noticed that $\mathcal{P}_+$ and $\mathcal{P}_-$ are isotropic with respect to the trace $< \mathcal{P}_\pm, \mathcal{P}_\pm >= 0$. To check

that $\mathcal{P}_+^\dagger = \mathcal{P}_-$ we consider a basis in $\mathcal{P}_+ : a_i\partial^i, i = 0, 1, \cdots, \infty$. The dual basis is of the form $\partial^{-i-1}b_i, i = 0, 1, \cdots, \infty$, since we have

$$< \partial^{-i-1}b_i, a_j\partial^j >= \delta_{ij} \int dx\ (a_ib_i)$$

and this is a basis for $\mathcal{P}_-$. ∎

As explained in Chapter [3], since $\mathcal{P}_-$ and $\mathcal{P}_+$ are two complementary Lie subalgebras of $\mathcal{P}$, the R-matrix

$$R = \frac{1}{2}(P_+ - P_-)$$

with P_+ and P_- the projections operators on $\mathcal{P}_+$ and $\mathcal{P}_-$, satisfies the modified Yang–Baxter equation. For any $V \in \mathcal{P}$ we will denote by $V_\pm$ its projection on $\mathcal{P}_\pm$. As usual, this R–matrix can be used to define a new Lie–algebra structure on $\mathcal{P}$ whose Lie bracket is denoted $[\ ,\ ]_R$. For any $V,\ W \in \mathcal{P}$ it is defined by:

$$[V, W]_R = [R(V), W] + [V, R(W)]$$

In particular, if $V = V_+ + V_-$ and $W = W_+ + W_-$, then $[V, W]_R = [V_+, W_+] - [V_-, W_-]$, so $\mathcal{P}_-$ and $\mathcal{P}_+$ commute for the R–bracket. We will denote by $\mathcal{P}_R$ the associated Lie algebra.

9.2 The KP hierarchy.

We introduce the KP flows by considering the action of the algebra $\mathcal{P}_+$, considered as a subalgebra of $\mathcal{P}_R$, on the formal group $G = \exp(\mathcal{P}_-)$. Note that, as a set we have $G \sim 1 + \mathcal{P}_-$, because powers of elements in $\mathcal{P}_-$ are in $\mathcal{P}_-$. Let Φ be an element of G. We write:

$$\Phi = 1 + \sum_1^\infty w_i\partial^{-i}\ \in (1 + \mathcal{P}_-) \tag{9.4}$$

Then $\Phi^{-1} = 1 + \sum_1^\infty w'_i\partial^{-i}$ and writing $\Phi^{-1}\Phi = 1$ one computes recursively the coefficients w'_i. One finds that they are of the form $w'_i = -w_i + p_i(w_1, \cdots, w_{i-1})$ where p_i is a polynomials in its arguments and their derivatives (up to order $i - 2$). For example one finds

$$\begin{aligned} w'_1 &= -w_1, \quad w'_2 = -w_2 + w_1^2 \\ w'_3 &= -w_3 - w_1(\partial w_1) + 2w_1w_2 - w_1^3 \end{aligned}$$

Recall, see Chapter [14], that one defines the infinitesimal dressing transformation of Φ by $V = V_+ + V_- \in \mathcal{P}_R$ by:

$$\delta_V\ \Phi = -\Phi\ V_+ + \left(\Phi \cdot V \cdot \Phi^{-1}\right)_+\ \Phi = \Phi\ V_- - \left(\Phi \cdot V \cdot \Phi^{-1}\right)_-\ \Phi \quad \in (1 + \mathcal{P}_-) \tag{9.5}$$

The two signs give the same result, and this defines a representation of $\mathcal{P}_R$: $[\delta_V, \delta_W] = \delta_{[V,W]_R}$ for any $V,\ W \in \mathcal{P}_R$. Since $\Phi \in (1+\mathcal{P}_-)$, the elements in $\mathcal{P}_-$ act trivially because for $V \in \mathcal{P}_-$ one has $\left(\Phi \cdot V \cdot \Phi^{-1}\right)_- = \Phi \cdot V \cdot \Phi^{-1}$, and only the elements of $\mathcal{P}_+$ generate non trivial flows. To get commuting flows we restrict ourselves to a commutative subalgebra of $\mathcal{P}_+$ namely the algebra generated by the ∂^k with $k > 0$. These are the flows of the KP hierarchy.

Definition. *For any $\Phi \in (1 + \mathcal{P}_-)$, define the k^{th} KP flow by:*

$$\partial_k\ \Phi = -\left(\Phi \cdot \partial^k \cdot \Phi^{-1}\right)_-\ \Phi \tag{9.6}$$

Note that $\partial_k\ \Phi = \delta_{\partial^k}\ \Phi$ since for $V = \partial^k$ one has $V_- = 0$. We know by construction that these flows commute, so that it is natural to write them as time evolutions. It is instructive to check directly this essential commutativity property.

Proposition. *The KP flows ∂_k commute.*

Proof. Consider the pseudo–differential operators $\Theta^k_\pm = \left(\Phi\partial^k\Phi^{-1}\right)_\pm$. One has:

$$\partial_k\partial_l\Phi - \partial_l\partial_k\Phi = \left([\Theta^l_-, \Theta^k] - [\Theta^k_-, \Theta^l] + [\Theta^k_-, \Theta^l_-]\right)_-\Phi$$

Replace $\Theta^k = \Theta^k_- + \Theta^k_+$ and similarly for Θ^l. One gets:

$$[\partial_k, \partial_l]\Phi = \left([\Theta^l_+, \Theta^k_-]_- + [\Theta^l_-, \Theta^k_+]_- + [\Theta^l_-, \Theta^k_-]_-\right)\Phi = [\Theta^l, \Theta^k]_-\Phi$$

The result follows because $[\Theta^l, \Theta^k] = 0$. ■

We now introduce the pseudo–differential operator:

$$Q = \Phi\cdot\partial\cdot\Phi^{-1}, \quad Q = \partial + \sum_{i=1}^{\infty} q_i\,\partial^{-i}$$

Note that Φ is the analogue of $g(\lambda)$ in the loop–algebra situation, and Q will be the analogue of the Lax matrix (see eq.(2.8) in Chapter [2]), but we will keep the traditional notations in this Chapter. In particular the evolution equation (9.6) is analogous to eq.(2.34) for g.

Proposition. *The coefficients q_i are of the form $q_i = -(\partial w_i) + h_i(w_1, \cdots, w_{i-1})$ where h_i is a differential polynomial in its arguments. The derivatives are at most of order $(i-1)$. Conversely, the w_i are recursively determined by the q_i up to constants.*

The above proposition is obvious, using the expression of Φ^{-1}. Note that the origin of the indeterminacy by constants in $q_i \to w_i$ is easy to understand. If $C = 1 + \sum_{i=1}^{\infty} c_i\partial^{-i}$ where the c_i are constants, the expression of Q is invariant when we replace Φ by ΦC.

Proposition. *Acting on $Q = \Phi\cdot\partial\cdot\Phi^{-1}$, the KP flows take the Lax form:*

$$\partial_k Q = [\ B_k\ ,\ Q\ ] \quad \text{with} \quad B_k = \left(Q^k\right)_+ \tag{9.7}$$

The role of the Lax matrix L is played by Q and the role of M_k is played by B_k, so that B_k is given by the the usual formula expressing M_k in terms of L. As in the Lax situation the zero curvature condition holds:

$$\partial_k B_l - \partial_l B_k = [B_k, B_l]$$

Proof. We have to compute $\partial_k Q$. Using the definition (9.6) so that $\partial_k\Phi = -(Q^k)_-\Phi$, we have:

$$\begin{aligned}
\partial_k Q = \partial_k(\Phi\partial\Phi^{-1}) &= (\partial_k\Phi)\ \partial\Phi^{-1} - \Phi\partial\Phi^{-1}\ (\partial_k\Phi)\ \Phi^{-1} \\
&= -\left(Q^k\right)_- Q + Q\left(Q^k\right)_- \\
&= -\left[\left(Q^k\right)_-, Q\right] = \left[\left(Q^k\right)_+, Q\right]
\end{aligned}$$

The zero–curvature condition follows from a direct computation:

$$\partial_k B_l - \partial_l B_k = \partial_k\left(Q^l\right)_+ - \partial_l\left(Q^k\right)_+ = \left([B_k, Q^l] - [B_l, Q^k]\right)_+$$

Therefore, using the decomposition $Q^k = \left(Q^k\right)_+ + \left(Q^k\right)_-$, we obtain:

$$\begin{aligned}
\partial_k B_l - \partial_l B_k - [B_k, B_l] &= \left([B_k, Q^l] - [B_l, Q^k] - [B_k, B_l]\right)_+ \\
&= -[B_k - Q^k, B_l - Q^l]_+ = -\left[\left(Q^k\right)_-, \left(Q^l\right)_-\right]_+ = 0
\end{aligned}$$

■

Proposition. *The quantities $H_k =< Q^k >$ are conserved.*

Proof. This is because $\partial_l H_k =< [B_l, Q^k] >$ vanishes due to the cyclicity of Adler's trace. ∎

Remark 1. The equations (9.7) are consistent in the sense that $[B_k, Q] \in \mathcal{P}_-$. This is because $[B_k, Q] = [(Q^k)_+, Q] = [Q^k - (Q^k)_-, Q] = -[(Q^k)_-, Q]$ which expands on the negative powers ∂^{-j}. Hence the Lax equations (9.7) produce non–linear equations of motion for the coefficients of Q, i.e. for the functions $\{q_i\}$. We denote by t_k the "KP–times" associated to these flows, so that the vector fields ∂_k are identified with $\partial_k = \frac{\partial}{\partial t_k}$.

Remark 2. The first KP–flow ∂_1 is identified with ∂, because we have $Q_+ = \partial$ so that the first flow reads:

$$\partial_1 \; Q = [(Q)_+, Q] = [\partial, Q] = \sum_{i=1}^{\infty} \; (\partial q_i)\partial^{-i}$$

Therefore, the KP–time t_1 is naturally identified with the variable x introduced in the definition of the algebra $\mathcal{P}$.

Example. To illustrate these formulae we compute the first few equations of motion. First we have:

$$(Q^2)_+ = \partial^2 + 2q_1, \quad (Q^3)_+ = \partial^3 + 3q_1\partial + 3(\partial q_1) + 3q_2$$

The time evolution with respect to t_2 reads:

$$\begin{aligned} \partial_2 q_1 &= \partial^2 q_1 + 2\partial q_2 \\ \partial_2 q_2 &= \partial^2 q_2 + 2\partial q_3 + 2q_1\partial q_1 \\ &\vdots \end{aligned}$$

Similarly the time evolution with respect to t_3 of q_1 is given by:

$$\partial_3 q_1 = \partial^3 q_1 + 3\partial^2 q_2 + 3\partial q_3 + 6q_1\partial q_1$$

Eliminating q_2 and q_3 between these equations one gets:

$$3\partial_2^2 q_1 = \partial(4\partial_3 q_1 - 12q_1\partial q_1 - \partial^3 q_1)$$

This reduces to the standard form of the KP equation setting $u = 2q_1$, see eq.(7.43) in Chapter [7]. It is the first of an infinite hierarchy of non–linear partial differential equations for $q_1, q_2, \cdots$.

9.3 The Baker–Akhiezer function of KP.

By analogy with the Lax situation, we consider eigenvectors of the operator Q, together with their time evolutions under the KP flows, namely we look for an eigenfunction ψ of Q such that:

$$\begin{aligned} (Q-k)\psi &= 0, \quad Q = \Phi\partial\Phi^{-1} \\ (\partial_m - B_m)\psi &= 0, \quad B_m = (Q^m)_+ \end{aligned} \tag{9.8}$$

In the algebro–geometric case such functions were shown to be Baker–Akhiezer functions and we will continue to call them by this name.

Proposition. *The Baker-Akhiezer function $\psi(x,k)$ obeying eqs.(9.8) can be written:*

$$\psi(x,k) = \Phi\ e^{\xi(t,k)} = (1 + w_1 k^{-1} + w_2 k^{-2} + \cdots)\ e^{\xi(t,k)} = \widehat{w}(x,k)\ e^{\xi(t,k)} \tag{9.9}$$

where $\xi(t,k) = \sum_{i=1}^{\infty}\ t_i k^i$.

Proof. We use the explicit expansion of Φ as $\Phi = 1 + \sum_{i=1}^{\infty}\ w_i \partial^{-i}$, and remark that $\partial^{-i}\ e^{\xi(t,k)} = k^{-i}\ e^{\xi(t,k)}$. Hence $\Phi e^{\xi(t,k)}$ is of the stated form. Then $Q\psi = (\Phi\partial\Phi^{-1})\Phi e^{\xi(t,k)} = k\psi$ Similarly, the evolution of ψ with respect to t_m is given by, using $\partial_m e^{\xi(t,k)} = \partial^m\ e^{\xi(t,k)}$:

$$\begin{aligned}
\partial_m \psi &= \partial_m(\Phi e^{\xi(t,k)}) = (\partial_m \Phi) e^{\xi(t,k)} + \Phi\left(\partial_m e^{\xi(t,k)}\right) \\
&= -\left(Q^m\right)_-\ \Phi e^{\xi(t,k)} + \Phi\left(\partial^m e^{\xi(t,k)}\right) = \left(-\left(Q^m\right)_- + Q^m\right) \Phi e^{\xi(t,k)} = \left(Q^m\right)_+ \psi
\end{aligned}$$

In the last equality, we have written $\Phi\partial^m e^{\xi} = (\Phi\partial^m\Phi^{-1})\Phi e^{\xi} = Q^m\psi$ and we used the decomposition $Q^m = \left(Q^m\right)_+ + \left(Q^m\right)_-$. ∎

It is useful to introduce the adjoint Baker–Akhiezer function by:

$$\psi^* = (\Phi^*)^{-1}\ e^{-\xi(t,k)}$$

where Φ^* is the formal adjoint of the pseudo–differential operator Φ. It satisfies the adjoint system:

Proposition. *The adjoint Baker–Akhiezer function obeys:*

$$\begin{aligned}
(\ Q^* - k\ )\,\psi^* &= 0, \quad Q^* = -(\Phi^*)^{-1}\partial\phi^* \\
(\ \partial_m + B_m^*\ )\,\psi^* &= 0, \quad B_m^* = (Q^m)_+
\end{aligned}$$

Proof. Recall that the formal adjoint is the antihomomorphism defined by $(a\ \partial^i)^* = (-\partial)^i\ a$. Therefore the formal adjoint of Q is $Q^* = \left(\Phi\cdot\partial\cdot\Phi^{-1}\right)^* = -(\Phi^*)^{-1}\cdot\partial\cdot\Phi^*$. As a consequence, we have:

$$Q^*\ \psi^* = -(\Phi^*)^{-1}\ \partial\ e^{-\xi(t,k)} = k\ \psi^*.$$

Similarly, we have:

$$\begin{aligned}
\partial_m\ \psi^* &= \left(\partial_m(\Phi^*)^{-1}\right)\ e^{-\xi(t,k)} + (\Phi^*)^{-1}\ \partial_m\ e^{-\xi(t,k)} \\
&= \left((Q^m)_-^* - (Q^m)^*\right)\ \psi^* = -B_m^*\ \psi^*
\end{aligned}$$

∎

Note that, compared to the algebro–geometric Akhiezer–Baker functions, the singularity here is at $k = \infty$. At this point the exponential factor $\exp\xi(t,k)$ has an essential singularity, while $\widehat{w}(x,k)$ is formally regular.

The fact that the Baker-Akhiezer functions are solutions of the linear system (9.8) implies that they satisfy an important bilinear identity:

Proposition. *The following bilinear identity holds for all $(i_1,\cdots,i_m), i_j \geq 0$*

$$\oint \frac{dk}{2i\pi}(\partial_1^{i_1}\cdots\partial_m^{i_m}\psi(t,k))\cdot\psi^*(t,k) = 0$$

This identity can formally be written as:

$$\oint \frac{dk}{2i\pi}\psi(t,k)\cdot\psi^*(t',k) = 0, \quad \forall t, t' \tag{9.10}$$

Proof. The integrals over k are residues around $k = \infty$, i.e. integrals on big circles around $k = \infty$. Notice that the integrands are meromorphic around ∞ because the essential singularities cancel. We first need a formula expressing the residue Res( ) of a product of two pseudo–differential operators, $D = \sum_i d_i \partial^i$ and $F = \sum_i f_i \partial^i$:

$$\oint \frac{dk}{2i\pi}(De^{kx})(Fe^{-kx}) = \text{Res}\,(DF^*) \tag{9.11}$$

The value of the residue is the coefficient of k^{-1} in the integrand.

$$\oint \frac{dk}{2i\pi}(De^{kx})(Fe^{-kx}) \quad = \quad \oint \frac{dk}{2i\pi}\sum_i d_i k^i \sum_j f_j(-k)^j = \sum_j (-1)^j d_{-j-1} f_j$$

Similarly, we compute the residue, i.e. the coefficient of ∂^{-1} in (DF^*) using $F^* = \sum_i \ (-\partial)^i f_i$:

$$\text{Res}\,(DF^*) \quad = \quad \text{Res}\,(\sum_{i,j} d_i \partial^i (-\partial)^j f_j) = \sum_j (-1)^j d_{-j-1} f_j$$

We can now prove the identity (9.10). Since $\partial_m \psi = B_m \psi$, and since B_m is a polynomial in ∂ (only a finite number of positive powers appear) it is sufficient to prove this equality for $(i, 0, \cdots, 0), i \geq 0$. In this case:

$$\begin{aligned}\oint \frac{dk}{2i\pi}(\partial^i \psi)\cdot\psi^* \quad &= \quad \oint \frac{dk}{2i\pi}(\partial^i \Phi e^{\xi(t,k)})\cdot(\Phi^*)^{-1}e^{-\xi(t,k)} \\ &= \quad \oint \frac{dk}{2i\pi}(\partial^i \Phi e^{kx})\cdot((\Phi^*)^{-1}e^{-kx}) = \text{Res}\,(\partial^i\Phi\cdot\Phi^{-1}) = \text{Res}\,(\partial^i) = 0\end{aligned}$$

■

In the previous proposition, we proved that the KP equations imply that the Baker-Akhiezer functions satisfy the bilinear identities (9.10). We now establish the converse statement.

Proposition. *Consider two formal series:*

$$\psi = \Phi e^{\xi(t,k)}, \quad \psi^* = \Phi^\star e^{-\xi(t,k)}$$

where Φ and $\Phi^\star$ are two pseudo–differential operators of the form:

$$\Phi = 1 + \sum_{i=1}^{\infty} w_i \partial^{-i}, \quad \Phi^\star = 1 + \sum_{i=1}^{\infty} w_i^\star (-\partial)^{-i}$$

where $\{w_i, w_i^\star\}$ are functions of the variables $\{t_i\}$. Let us assume that the bilinear identity (9.10) are satisfied. Then one has $\Phi^\star = (\Phi^*)^{-1}$. Moreover defining $Q = \Phi\partial\Phi^{-1}$ we have $\partial_m \Phi = -(Q^m)_-\Phi$. Hence ψ is the Baker–Akhiezer function of the KP hierarchy.

Proof. By definition the functions ψ and ψ^* are:

$$\psi = \left(1 + \sum_{i=1}^{\infty} w_i k^{-i}\right) e^{\xi(t,k)}, \quad \psi^* = \left(1 + \sum_{i=1}^{\infty} w_i^\star k^{-i}\right) e^{-\xi(t,k)}$$

and we assume that $\oint \frac{dk}{2i\pi}\partial^i\psi\cdot\psi^* = 0$ for any $i \geq 0$. We first prove that this implies that $\Phi^\star$ and Φ^* are inverse to each other. Indeed, using the previous formula for the Res ( ) of a product of two pseudo-differential operators, we have for any $i \geq 0$:

$$\text{Res}\,(\partial^i\Phi\ (\Phi^\star)^*) = \oint \frac{dk}{2i\pi}(\partial^i\Phi e^{\xi(t,k)})(\Phi^\star e^{-\xi(t,k)}) = \oint \frac{dk}{2i\pi}\partial^i\psi\cdot\psi^* = 0$$

where the hypothesis is used in the last step. But by construction $\Phi(\Phi^\star)^* = 1 + X$, with $X \in \mathcal{P}_-$, therefore the above equation implies $\mathrm{Res}\,(\partial^i X) = 0$, for all $i \geq 0$ and thus $X = 0$, so that $\Phi^\star = (\Phi^*)^{-1}$. Now let $Q = \Phi\partial\Phi^{-1}$, and $B_m = (Q^m)_+$. We show that $\partial_m \Phi = -(Q^m)_-\Phi$. First, observe that, using $\partial_m e^{\xi(t,k)} = \partial^m e^{\xi(t,k)}$ and $\Phi\partial^m = Q^m\Phi$, we have:

$$\begin{aligned}
\Big((\partial_m\Phi) + (Q^m)_-\Phi\Big)e^{\xi(t,k)} &= \Big(\partial_m \Phi e^{\xi(t,k)}\Big) - \Big(\Phi\partial_m - (Q^m)_-\Phi\Big)e^{\xi(t,k)} \\
&= \Big(\partial_m \Phi e^{\xi(t,k)}\Big) - \Big(\Phi\partial^m - (Q^m)_-\Phi\Big)e^{\xi(t,k)} \\
&= \Big(\partial_m - Q^m + (Q^m)_-\Big)\Phi\ e^{\xi(t,k)} = \Big(\partial_m - (Q^m)_+\Big)\Phi\ e^{\xi(t,k)}
\end{aligned}$$

By hypothesis ψ and ψ^* satisfy the bilinear identity (9.10). Therefore, since $(Q^m)_+$ is a differential *polynomial* we have, for any $i \geq 0$:

$$0 = \oint \frac{dk}{2i\pi}\partial^i(\partial_m - (Q^m)_+)\Phi e^{\xi(t,k)} \cdot \Phi^\star e^{-\xi(t,k)} = \oint \frac{dk}{2i\pi}\partial^i\Big((\partial_m\Phi) + (Q^m)_-\Phi\Big)e^{\xi(t,k)} \cdot \Phi^\star e^{-\xi(t,k)}$$

Equivalently, one can write:

$$0 = \mathrm{Res}\,\Big(\partial^i\Big((\partial_m\Phi) + (Q^m)_-\Phi\Big)\cdot(\Phi^\star)^*\Big) = \mathrm{Res}\,\Big(\partial^i\Big((\partial_m\Phi) + (Q^m)_-\Phi\Big)\Phi^{-1}\Big)$$

Since, this is true for any $i \geq 0$, it gives: $((\partial_m\Phi) + (Q^m)_-\Phi)\Phi^{-1} = 0$. Multiplying on the right by Φ gives the result. ∎

9.4 The tau function of KP.

The bilinear identity (9.10) allows to express the Baker–Akhiezer function in terms of a τ–function.

Proposition. *Assume that ψ and ψ^* are Baker–Akhiezer functions of the KP hierarchy satisfying eq.(9.10). Then there exists a function τ such that*

$$\psi(t,k) = \frac{\tau(t-[k])}{\tau(t)}\ e^{\xi(t,k)}, \quad \psi^*(t,k) = \frac{\tau(t+[k])}{\tau(t)}\ e^{-\xi(t,k)} \tag{9.12}$$

where $[k] = \left(\frac{1}{k}, \frac{1}{2k^2}, \frac{1}{3k^3}, \cdots\right)$ and $\xi(t,k) = \sum_{i=1}^\infty t_i k^i$.

Proof. Note that for $f(k) = 1 + \sum_{i=1}^\infty f_i k^{-i}$ we have, by the residue theorem:

$$\oint \frac{dk}{2i\pi}\frac{f(k)}{1-\frac{k}{k'}} = k'(f(k') - 1)$$

where k' is so big to be inside the integration contour around ∞. The bilinear identity applied to t and $t' = t - [k_1]$ yields:

$$0 = \oint \frac{dk}{2i\pi}\,\psi(t,k)\psi^*(t-[k_1],k) = \oint \frac{dk}{2i\pi}\,\frac{\widehat{w}(t,k)\widehat{w}^*(t-[k_1],k)}{1-k/k_1}$$

In the last step we used that

$$e^{-\xi(t-[k_1],k)} = e^{-\xi(t,k)}\,\frac{1}{1-k/k_1}$$

Notice that the Cauchy kernel in the right–hand side is produced by the very specific choice of shift we consider. Applying the above residue formula we get $\widehat{w}(t,k_1)\widehat{w}^*(t-[k_1],k_1) = 1$, i.e.

$$\widehat{w}^*(t-[k_1],k_1) = \frac{1}{\widehat{w}(t,k_1)} \tag{9.13}$$

Similarly applying the bilinear identity to t and $t' = t - [k_1] - [k_2]$ we get:

$$0 = \oint \frac{dk}{2i\pi}\, \psi(t,k)\psi^*(t-[k_1]-[k_2],k) = \oint \frac{dk}{2i\pi}\, \frac{\widehat{w}(t,k)\widehat{w}^*(t-[k_1]-[k_2],k)}{(1-k/k_1)(1-k/k_2)}$$

Since there is no residue at ∞ we get $\widehat{w}(t,k_1)\widehat{w}^*(t-[k_1]-[k_2],k_1) = \widehat{w}(t,k_2)\widehat{w}^*(t-[k_1]-[k_2],k_2)$. Eliminating $\widehat{w}^*$ using eq.(9.13) we obtain the functional equation:

$$\frac{\widehat{w}(t-[k_2],k_1)}{\widehat{w}(t,k_1)} = \frac{\widehat{w}(t-[k_1],k_2)}{\widehat{w}(t,k_2)}$$

We want to show that this equation implies:

$$\widehat{w}(t,k) = \frac{\tau(t-[k])}{\tau(t)} \tag{9.14}$$

It is trivial to verify that this solves the equation since both sides are equal to $(\tau(t-[k_1]-[k_2])\tau(t))/(\tau(t-[k_1])\tau(t-[k_2]))$.

We now proceed to show that this is the general solution. Taking the log of the functional equation we are led to study an equation of the form

$$f(t-[u],v) - f(t,v) = f(t-[v],u) - f(t,u)$$

where the function f is:

$$f(t,v) = \log\, \widehat{w}(t,v) = \frac{1}{v}w_1(t) + \frac{1}{v^2}(w_2(t) - \frac{1}{2}w_1^2(t)) + \cdots$$

Introducing the operator:

$$D_v = \frac{\partial}{\partial v} - \sum_{i=1}^{\infty} \frac{1}{v^{i+1}}\partial_i$$

we remark that $D_v\phi(t-[v]) = 0$ for any function $\phi(t)$. Applying D_v to the above equation we get $D_v f(t-[u],v) - D_v f(t,v) = -D_v f(t,u)$. Expanding in v^{-1} and setting $D_v f(t,v) = \sum_{i=1}^{\infty}\gamma_i(t)v^{-i-1}$ this reads:

$$\gamma_i(t-[u]) - \gamma_i(t) = \partial_i f(t,u) \tag{9.15}$$

Hence, considering $F_{ij}(t) = \partial_i\gamma_j(t) - \partial_j\gamma_i(t)$ we get the condition $F_{ij}(t-[u]) = F_{ij}(t)$. Expanding in powers of u^{-1} one sees immediately that $F_{ij}(t)$ is independent of all the time variables $t_1, t_2, \cdots$. But by construction, F_{ij} is a local differential polynomial in $w_1(t), w_2(t), \cdots$, for example $F_{12} = \partial_2 w_1 - 2\partial_1 w_2 + 2w_1\partial_1 w_1 - \partial_1^2 w_1$. Using the equations of motion of the KP hierarchy we can replace all the $\partial_k w_l$ for $k \geq 2$ by higher derivatives of w_l with respect to $t_1 = x$. Hence we can write F_{ij} as a polynomial in the $\partial^k w_l$. But the monomials in F_{ij} are independent, and we know that F_{ij} is constant, hence it reduces to its constant term. Since it vanishes for the particular solution $w = 0$ we see that $F_{ij} = 0$. So we can write $\gamma_i(t) = \partial_i \log\tau(t)$. Finally inserting this into eq.(9.15) we get eq.(9.14).

The formula for $\psi^*(t,k)$ is then a straightforward consequence of eq.(9.13) where one substitutes $t \to t + [k_1]$.

∎

The formula (9.12) is usually called Sato's formula. Using eq.(9.12) the bilinear identity (9.10) can be rewritten as bilinear identities for the τ-functions. Of course they coincide with the Hirota bilinear identities we obtained using vertex operators.

Remark 1. We recall that in terms of the fermionic description of the previous Chapter, the τ function and the Baker–Akhiezer function have a particularly elegant formulation:

$$\begin{aligned} \tau(t;g) &= <0|e^{H(t)}g|0> \\ \psi(x,k) &= \langle 1|e^{H(t)}\psi(k)g|0\rangle/\tau(x;g) \\ \psi^*(x,k) &= \langle -1|e^{H(x)}\psi^*(k)g|0\rangle/\tau(x;g) \end{aligned} \tag{9.16}$$

where g is an element of the group $GL(\infty)$, and $\psi(k)$ and $\psi^*(k)$ are the fermionic operators (not to be confused with the Baker–Akhiezer functions $\psi(t,k)$ and $\psi^*(t,k)$).

Remark 2. The Grassmannian formulation also shows that $\tau(t)$ is given by an infinite determinant. In particular interesting cases it degenerates to a finite determinant.

9.5 The generalized KdV equations.

The KP hiereachy is a system of evolution equations for the infinite set of functions $w_i(t)$ or equivalently the coefficients $q_i(t)$ appearing in Q. To reduce the sytem to a finite number of coefficients q_i we remark that Q obeys Lax equations $\partial_k Q = [(Q^k)_+, Q]$ hence Q^{n+1} also obeys the same Lax equation. The main remark is that one can impose consistently that Q^{n+1} is a differential operator, i.e. $(Q^{n+1})_- = 0$. This because one has then $[(Q^k)_+, Q^{n+1}]_- = 0$, so $\partial_k Q^{n+1}$ is a differential operator. Moreover the two sides of the Lax equation have the same order because one can also write $\partial_k Q^{n+1} = [-(Q^k)_-, Q^{n+1}]$, which is in fact of order ∂^{n-1}. So one can further impose that the coefficient of ∂^n in Q^{n+1} vanishes. Finally if we impose that $Q^{n+1} = L$ is a *differential* operator:

$$L = \partial^{n+1} + \sum_{i=0}^{n-1} u_i \partial^i \tag{9.17}$$

then L satisfies Lax equations which define a flow on the finite number of functions u_i. This flow is a closed system on the u_i because given L one can reconstruct Q such that $Q^{n+1} = L$.

Proposition. *Let L be a differential operator of the form $L = \partial^{n+1} + \sum_{i=0}^{n-1} u_i \partial^i$. There exists a unique pseudo differential operator of the form $Q = \partial + q_{-1}\partial^{-1} + \cdots$ such that $Q^{n+1} = L$. We will denote $Q = L^{\frac{1}{n+1}}$.*

Proof. If $Q = \partial + q_0 + q_{-1}\partial^{-1} + \cdots$ one first see that $Q^{n+1} = \partial^{n+1} + (n+1)q_0\partial^n + \cdots$. Since there is no term in ∂^n in L one has $q_0 = 0$. Then by recursion one shows that:

$$u_{n-1} = (n+1)q_{-1}, \quad u_{n-2} = (n+1)q_{-2} + \frac{n(n+1)}{2}\partial q_{-1}, \quad u_{n-i} = (n+1)q_{-i} + p_i(q_{-1}, \cdots, q_{-i+1})$$

where p_i is a differential polynomial in its arguments. Knowing the u_i this system uniquely determines the q_i. ■

Proposition. *Let L be a differential operator as eq.(9.17). Then the Lax equations*

$$\dot{L} = \left[\left(L^{\frac{k}{n+1}} \right)_+, L \right] \tag{9.18}$$

are consistent for all $k \in \mathbb{N}$.

Proof. We introduce the operator $Q = L^{\frac{1}{n+1}}$. It is useful to notice that Q^k, $\forall k \in N$ commutes with L since $LQ^k = Q^{n+1+k} = Q^k L$. Then, we have:

$$\left[\left(Q^k\right)_+, L \right] = \left[Q^k, L\right] - \left[\left(Q^k\right)_-, L \right] = -\left[\left(Q^k\right)_-, L \right]$$

Thus the differential operator $\left[\left(Q^k\right)_+, L \right]$ is of order less than or equal to $n-1$. ■

Example. Let us consider the case $n = 1$, then the operator L is the second order differential operator

$$L = \partial^2 + u$$

We first find Q such that $Q^2 = L$. One has $Q^2 = \partial^2 + 2q_{-1} + (2q_{-2} + \partial q_{-1})\partial^{-1} + \cdots$ so that $q_{-1} = \frac{1}{2}u$, $q_{-2} = -\frac{1}{4}\partial u$, etc. We again check on this simple example that all the q_{-j} are recursively determined in terms of u by requiring that no ∂^{-j} terms occur in Q^2. Then, computing the $(Q^k)_+$ we get for $k = 1$, $\partial_1 L = [\partial, L]$, and this reduces to the identification $\partial_1 = \partial$, for $k = 2$ we get the trivial equation $\partial_2 L = 0$, finally the first non trivial case is $k = 3$. We have $(Q^3)_+ = \partial^3 + \frac{3}{2}u\partial + \frac{3}{4}(\partial u)$ so the Lax equation reads $\partial_3 u = [(Q^3)_+, \partial^2 + u]$. This is the Korteweg–de Vries equation:

$$4\partial_3 u = \partial^3 u + 6u(\partial u)$$

This is the first of a hierarchy of equations obtained by taking $k = 3, 5, 7, \cdots$, called the KdV hierarchy (note that for k even one gets trivial equations), which will be studied in details in Chapter [10]. Finally this hierarchy is generalized by taking $n = 2, 3, \cdots$

We now show that the generalized KdV equations are Hamiltonian systems. The differential operator L is an element of $\mathcal{P}_+$. So we have to specify a Poisson structure on the space $\mathcal{F}(\mathcal{P}_+)$ of function on $\mathcal{P}_+$. If we see $\mathcal{P}_+$ as the dual of the Lie algebra $\mathcal{P}_-$ through the Adler trace, there is a natural Poisson bracket on $\mathcal{P}_+$, the Kirillov bracket. For any functions f and g on $\mathcal{P}_+$, is defined as usual by:

$$\{f, g\}_1(L) = < L\, ,\, [df, dg] > \quad \forall\, L \in \mathcal{P}_+ \tag{9.19}$$

where we understand that $df, dg \in \mathcal{P}_-$. In particular, if for any $X = \sum_{j=0}^{\infty} \partial^{-j-1} x_j \in \mathcal{P}_-$, we define the linear function $f_X(L)$ by:

$$f_X(L) = < L, X > = \int dx \sum_{i=0}^{n} (u_i x_i), \tag{9.20}$$

(note that this is a finite sum) we have $df_X = X \in \mathcal{P}_-$. Therefore $\{f_X, f_Y\} = f_{[X,Y]} = < L, [X, Y] >$ for any $X,\ Y \in \mathcal{P}_-$.

Proposition. *Let $L \in \mathcal{P}_+$ a differential operator of order $n+1$ as in eq.(9.17). Let*

$$H_k(L) = \left(\frac{n+1}{n+k+1}\right) \quad < L^{\frac{k}{n+1}+1} >$$

be the conserved quantities of the generalized KP hierarchy. Then:
(i) The quantities H_k are the Hamiltonians of the generalized KdV flows under the bracket (9.19):

$$\dot{L} = \{H_k, L\}_1 = \left[\left(L^{\frac{k}{n+1}}\right)_+, L\right]$$

(ii) The functions $H_k(L)$ are in involution with respect to this bracket.

<u>Proof</u>. Recall that $Q = L^{\frac{1}{n+1}}$ so that H_k is proportional to $< Q^{k+n+1} >$ which are the conserved quantities of the KP hierarchy. We first need to compute the differential of the Hamiltonian H_k. Let L and δL be differential operators of the form (9.17), then one has:

$$< (L + \delta L)^\nu > = < L^\nu > + \nu < L^{\nu-1} \delta L > + \cdots$$

which implies $d < L^\nu > = \nu (L^{\nu-1})_{(-n)}$ where the notation $(\ )_{(-n)}$ means projection on $\mathcal{P}_-$ truncated at the first n terms. This is because $\delta L = \delta u_{n-1} + \cdots + \delta u_0$ which is dual to elements of the form $\partial^{-1} x_0 + \cdots + \partial^{-n} x_n$ under the Adler trace. Hence:

$$dH_k(L) = \left(L^{\frac{k}{n+1}}\right)_{(-n)} = \left(Q^k\right)_{(-n)} \quad \in \mathcal{P}_- \tag{9.21}$$

We define the terms left over by the truncation as $\theta^{(k)}_{-(n+1)}$ by $\left(L^{\frac{k}{n+1}}\right)_- = dH_k + \theta^{(k)}_{-(n+1)}$.
We first prove the evolution equation. Consider as above the function $f_X(L) =< LX >$, then $\dot{f}_X = \{H_k, f_X\}(L) =< L, [dH_k, df_X] >=< [L, dH_k], X >$, where we used the invariance of the Adler trace. Therefore, the equation of motion under H_k are:

$$\dot{L} = [L, dH_k]_+ = \left[L, \left(L^{\frac{k}{n+1}}\right)_-\right]_+ - \left[L, \theta^{(k)}_{-(n+1)}\right]_+ = \left[\left(L^{\frac{k}{n+1}}\right)_+, L\right]$$

where we have used $[L^{\frac{k}{n+1}}, L] = 0$, and the fact that $[L, \theta^{(k)}_{-(n+1)}]_+ = 0$.
Next we show that the Hamiltonians H_k are in involution. We have:

$$\{H_k, H_l\}_1(L) =< L, [dH_k, dH_l] >=< [L, dH_k]_+ \, , \, dH_l >=< \left[\left(L^{\frac{k}{n+1}}\right)_+, L\right] \, , \, dH_l >$$

Using that $\left[\left(L^{\frac{k}{n+1}}\right)_+, L\right]$ is a differential operator of order $n-1$ and thus $< \left[\left(L^{\frac{k}{n+1}}\right)_+, L\right] \, , \, \theta^{(l)}_{-(n+1)} >= 0$, we have:

$$\{H_k, H_l\}_1(L) =< \left[\left(L^{\frac{k}{n+1}}\right)_+, L\right] \left(L^{\frac{l}{n+1}}\right)_- >=< \left[\left(L^{\frac{k}{n+1}}\right)_+, L\right] \; L^{\frac{l}{n+1}} >$$

In the last step we used that $< \mathcal{P}_+, \mathcal{P}_+ >= 0$ in order to replace $\left(L^{\frac{l}{n+1}}\right)_-$ by $L^{\frac{l}{n+1}}$. Finally, from the invariance of the trace, we obtain:

$$\{H_k, H_l\}_1(L) =< \left(L^{\frac{k}{n+1}}\right)_+ \left[L, L^{\frac{l}{n+1}}\right] >= 0$$

■

This proposition shows that the generalized KdV hierarchies are Hamiltonian systems. In the next section we show that there exists in fact another local Hamiltonian structure for the same hierarchy.

9.6 KdV Hamiltonian structures.

We can establish recursion relations between the equations of motion written with hamiltonians H_k and H_{k+n+1}. These relations are called Lenard recursion relations. This leads naturally to introduce two different Poisson brackets on $\mathcal{F}(\mathcal{P}_+)$ such that the generalized KdV equations are Hamilton equations for both brackets; such systems are called bihamiltonian systems.

Proposition. *Let L be the differential operator (9.17). Let us introduce two operators $\mathcal{D}_1$ and $\mathcal{D}_2$ acting on any $X \in \mathcal{P}_-$:*

$$\begin{aligned} \mathcal{D}_1(X) &= [L, X]_+ \\ \mathcal{D}_2(X) &= (LX)_+ L - L(XL)_+ - \frac{1}{n+1}[L, (\partial^{-1}[X, L]_{-1})] \end{aligned}$$

Then, the functions $H_k(L)$ satisfy the following recursion relation:

$$\mathcal{D}_1(dH_{k+n+1}) = \mathcal{D}_2(dH_k) \tag{9.22}$$

Proof. Note that for any $X \in \mathcal{P}_-$ we have $0 =< [X, L] >= \int dx\, [X, L]_{-1}$ so that $[X, L]_{-1}$ is a total derivative and the object $(\partial^{-1}[X, L]_{-1})$ appearing in the definition of $\mathcal{D}_2$ is by definition the primitive of this total

differential, i.e. it is local. For example, we have $[\partial^2+u,\sum_{i=1}^{\infty}x_i\partial^{-i}]_{-1}=\partial(\partial x_1+2x_2)$.
Using the decomposition $dH_{k+n+1}=Q_-^{k+n+1}-\theta_{-(n+1)}^{(k+n+1)}$ with $Q=L^{\frac{1}{n+1}}$, we have:

$$\mathcal{D}_1(dH_{k+n+1})=[L,dH_{k+n+1}]_+=[L,Q_-^{k+n+1}]_+-[L,\theta_{-(n+1)}^{(k+n+1)}]_+=-[L,Q_+^{k+n+1}]$$

We used the fact that $[L,\theta_{-(n+1)}^{(k+n+1)}]_+=0$ and the fact that $Q^{k+n+1}=Q_+^{k+n+1}+Q_-^{k+n+1}$ commutes with L. Now, the simple recursion relation: $Q^kL=Q^{k+n+1}$ implies that:

$$\left(Q^{k+n+1}\right)_+=\left(Q^kL\right)_+=\left(Q_+^kL\right)_++\left(Q_-^kL\right)_+=Q_+^kL+\left(Q_-^kL\right)_+$$

Thus we obtain:

$$\begin{aligned}\mathcal{D}_1(dH_{k+n+1}) &= -\left[L,\left(Q^{k+n+1}\right)_+\right]=-\left[L,Q_+^k\right]L+\left[L,\left(Q_-^kL\right)_+\right]\\ &= [L,dH_k]_+L-[L,(dH_kL)_+]-[L,(\theta_{-(n+1)}^{(k)}L)_+]\end{aligned}\tag{9.23}$$

where we used again the decomposition $dH_k=Q_-^k-\theta_{-(n+1)}^{(k)}$. The remarkable fact is that the last term involving $\theta_{-(n+1)}^{(k)}$ can also be expressed in terms of dH_k. Indeed, defining v_0 by $\theta_{-(n+1)}^{(k)}=v_0\partial^{-(n+1)}+\cdots$, we have $(\theta_{-(n+1)}^{(k)}L)_+=v_0$. Remarking that $[(L^{\frac{k}{n+1}})_-,L]_-=-[(L^{\frac{k}{n+1}})_+,L]_-=0$, and looking at the residue gives:

$$0=[(L^{\frac{k}{n+1}})_-,L]_{-1}=[dH_k,L]_{-1}+[\theta_{-(n+1)}^{(k)},L]_{-1}=[dH_k,L]_{-1}-(n+1)\partial v_0$$

Because the residue of a commutator is a total derivative, we can integrate the previous equation, and write consistently that:

$$v_0=\frac{1}{n+1}\left(\partial^{-1}[dH_k,L]_{-1}\right)$$

Inserting this result into eq(9.23) we get:

$$[L,dH_{k+n+1}]_+=(LdH_k)_+L-L(dH_kL)_+-\frac{1}{n+1}[L,(\partial^{-1}[dH_k,L]_{-1})]$$

which is the claimed statement. ∎

This remarkable relation lead Adler to conjecture that the operators $\mathcal{D}_1$ and $\mathcal{D}_2$ define two Poisson brackets on $\mathcal{F}(\mathcal{P}_+)$ denoted $\{\ ,\ \}_1$ and $\{\ ,\ \}_2$. These brackets are defined by:

$$\begin{aligned}\{f,g\}_1 &= <\mathcal{D}_1(df)\cdot dg>\\ \{f,g\}_2 &= <\mathcal{D}_2(df)\cdot dg>\end{aligned}$$

The first bracket is the Kirillov bracket (9.19), since $dg\in\mathcal{P}_-$ and the Adler trace is invariant. More explicitely, let us define as in eq.(9.20) the functions $f_X(L)=<LX>$ such that $df_X=X$. The explicit expressions for the two brackets are then:

$$\begin{aligned}\{f_X,f_Y\}_1(L) &= <LXY>-<LYX>\\ \{f_X,f_Y\}_2(L) &= <(LX)_+(LY)_->-<(XL)_+(YL)_->-\frac{1}{n+1}\int dx\left((\partial^{-1}[L,X]_{-1})\,[L,Y]_{-1}\right)\end{aligned}\tag{9.24}$$

We now prove that $\{\ ,\ \}_2$ is a Poisson bracket. The antisymmetry can easily be checked using the cyclicity of Adler trace and isotropy of $\mathcal{P}_\pm$. To check the Jacobi identity, we change variables, from the coefficients u_i of L, see eq.(9.17), to new variables p_j in which the Jacobi identity is obvious. This change of variables is called the Miura transformation.

Proposition. *Let $L = L_{n+1} L_n \cdots L_1$, with $L_i = \partial + p_i$, ie.*

$$L = (\partial + p_{n+1}) \cdot (\partial + p_n) \cdots (\partial + p_1), \quad \sum_{i=1}^{n+1} p_i = 0 \tag{9.25}$$

Let us define a Poisson bracket on the functions $p_k(x)$ by:

$$\{p_k(x), p_l(y)\} = \left(\delta_{kl} - \frac{1}{n+1} \right) \delta'(x-y) \tag{9.26}$$

Then we have $\{f_X, f_Y\}(L) = \{f_X, f_Y\}_2(L)$ and in particular $\{\ ,\ \}_2$ satisfies the Jacobi identity.

Proof. For any functions $f_X(L) =< LX >$, $X \in \mathcal{P}_-$, seen as functions of p_k, we have:

$$\{f_X, f_Y\}(L) = \sum_{k,l} \int dx\, dy \frac{\delta f_X}{\delta p_k(x)} \{p_k(x), p_l(y)\} \frac{\delta f_Y}{\delta p_l(y)}$$

We now check that inserting eq.(9.26) into this formula reproduces the second Poisson bracket structure. Define the operators:

$$L_{ij} = L_i L_{i-1} \cdots L_j, \quad \text{for } i \geq j, \quad L_{01} = 1, \ L_{n+1,n+2} = 1 \tag{9.27}$$

From the expression of f_X, we have: $\delta f_X / \delta p_k(x) = (L_{k-1,1} X L_{n+1,k+1})_{-1}(x)$. Thus,

$$\{f_X, f_Y\}(L) = \sum_{k,l} \int dx\, dy (L_{k-1,1} X L_{n+1,k+1})_{-1}(x)\ \{p_k(x), p_l(y)\}\ (L_{l-1,1} Y L_{n+1,l+1})_{-1}(y)$$

By definition of the Poisson bracket of the p_k, this is rewritten as:

$$\begin{aligned} \{f_X, f_Y\}(L) &= \sum_{k=1}^{n+1} \int dx (\partial (L_{k-1,1} X L_{n+1,k+1})_{-1}) (L_{k-1,1} Y L_{n+1,k+1})_{-1} \\ &- \frac{1}{n+1} \sum_{k=1}^{n+1} \sum_{l=1}^{n+1} \int dx (\partial (L_{k-1,1} X L_{n+1,k+1})_{-1}) (L_{l-1,1} Y L_{n+1,l+1})_{-1} \end{aligned} \tag{9.28}$$

We can write $\partial (L_{k-1,1} X L_{n+1,k+1})_{-1} = ([L_k, L_{k-1,1} X L_{n+1,k+1}])_{-1}$. This is because $L_k = \partial + p_k$ and one checks that for any pseudo differential $f = \sum_n f_n \partial^n$ one has $[\partial, f]_{-1} = (\partial f_{-1})$ and $[p_k, f]_{-1} = 0$. Thus $\{f_X, f_Y\}(L)$ is equal to:

$$\begin{aligned} &\sum_{k=1}^{n+1} \int dx \Big[(L_{k1} X L_{n+1,k+1})_{-1} (L_{k-1,1} Y L_{n+1,k+1})_{-1} - (L_{k-1,1} X L_{n+1,k})_{-1} (L_{k-1,1} Y L_{n+1,k+1})_{-1} \Big] \\ &- \frac{1}{n+1} \sum_{k,l=1}^{n+1} \int dx \Big[(L_{k1} X L_{n+1,k+1})_{-1} (L_{l-1,1} Y L_{n+1,l+1})_{-1} - (L_{k-1,1} X L_{n+1,k})_{-1} (L_{l-1,1} Y L_{n+1,l+1})_{-1} \Big] \end{aligned} \tag{9.29}$$

We use the identity, true for all p,

$$\int dx (U)_{-1} (V)_{-1} =< U_- (\partial + p) V_- >=< (\partial + p) U_- V_- > \tag{9.30}$$

to rewrite the first two terms in eq.(9.29) as:

$$\sum_{k=1}^{n+1} < (L_{k1} X L_{n+1,k+1})_-\ L_k\ (L_{k-1,1} Y L_{n+1,k+1})_- > - < L_k (L_{k-1,1} X L_{n+1,k})_-\ (L_{k-1,1} Y L_{n+1,k+1})_- >$$

Next replace all the terms U_- by $U - U_+$, to get

$$\sum_{k=1}^{n+1} \Big(- < (L_{k1} X L_{n+1,k+1})_+\ L_{k1} Y L_{n+1,k+1} > + < (L_{k-1,1} X L_{n+1,k})_+\ L_{k-1,1} Y L_{n+1,k} > \Big)$$

In the sum the terms cancel two by two, and we are left with $<(XL)_+YL> - <(LX)_+LY>$. The summation over k in the last two terms in eq.(9.29) reduces to:

$$
\begin{aligned}
\frac{1}{n+1}\sum_{l=1}^{n+1}\int dx\ [X,L]_{-1}(L_{l-1,1}YL_{n+1,l+1})_{-1} &= \frac{1}{n+1}\sum_{l=1}^{n+1} <[X,L]_{-1}L_{l-1,1}YL_{n+1,l+1}> \\
&= \frac{1}{n+1}\sum_{l=1}^{n+1} <L_{n+1,l+1}[X,L]_{-1}L_{l-1,1}Y> \qquad (9.31)
\end{aligned}
$$

We have $[L,f] = \sum_{l=1}^{n+1} L_{n+1,l+1}\partial f L_{l-1,1}$. Thus we finally obtain

$$
\frac{1}{n+1} <[L,\partial^{-1}([X,L]_{-1})]Y> = \frac{1}{n+1}\int dx\ (\partial^{-1}([X,L]_{-1})[Y,L]_{-1}
$$

This completes the proof that $\{\ ,\ \} = \{\ ,\ \}_2$. Since $\{\ ,\ \}$ obviously satisfies the Jacobi identity, so does $\{\ ,\ \}_2$. ∎

Proposition. *The Hamiltonians H_k are in involution with respect to both Poisson brackets.*

Proof. We already know that $\{H_k, H_l\}_1 = 0$. Using the recursion relation (9.22), we have:

$$
\{H_k, H_l\}_2 = <\mathcal{D}_2 dH_k, dH_l> = <\mathcal{D}_1 dH_{n+1+k}, dH_l> = \{H_{n+1+k}, H_l\}_1 = 0
$$

∎

The two Poisson brackets $\{\ ,\ \}_1$ and $\{\ ,\ \}_2$ have a remarkable compatibility property, called the Magri compatibility condition.

Proposition. *The two Poisson structures $\{\ ,\ \}_1$ and $\{\ ,\ \}_2$ are compatible, in the sense that for any λ_1, λ_2 the application $(f,g) \to \lambda_1\{f,g\}_1 + \lambda_2\{f,g\}_2$ is a Poisson bracket.*

Proof. The condition that the sum of two Poisson brackets $\{f,g\}_1 = \sum_{ij} P_{ij}\partial_i f\partial_j g$ and $\{f,g\}_2 = \sum_{ij} Q_{ij}\partial_i f\partial_j g$ satisfies the Jacobi identity reads:

$$
\Big(P_{il}\partial_l Q_{jk} + Q_{il}\partial_l P_{jk}\Big)\partial_i f\partial_j g\partial_k h + \text{ cyclic } = 0
$$

Since no second order differentials occur, it is sufficient to check it on the linear functions $f = f_X$, $g = f_Y$, and $h = f_Z$. Our condition becomes:

$$
<\mathcal{D}_1(X), d<\mathcal{D}_2(Y), Z>> + <\mathcal{D}_2(X), d<\mathcal{D}_1(Y), Z>> + \text{cyclic permutations} = 0
$$

We have:

$$
\begin{aligned}
d<\mathcal{D}_1(X), Y> &= [X,Y] \\
d<\mathcal{D}_2(X), Y> &= X(LY)_- - (YL)_-X + (XL)_-Y - Y(LX)_- + (YLX - XLY)_- \\
&\quad + \frac{1}{n+1}\Big[(\partial^{-1}[Y,L]_{-1}), X\Big] - \frac{1}{n+1}\Big[(\partial^{-1}[X,L]_{-1}), Y\Big]
\end{aligned}
$$

Plugging this into the above condition one checks that it is indeed satisfied. ∎

9.7 The Drinfeld Sokolov reduction.

The above Poisson brackets $\{\ ,\ \}_1$ and $\{\ ,\ \}_2$ have a nice Lie algebraic interpretation, which we now explain. Consider the loop algebra of traceless $(n+1)\times(n+1)$ matrices $U(x)$ with elements functions of x, and let $\mathcal{G}$ be the central extension of this loop algebra. It consists of pairs $(U(x),u)$ also denoted $U(x)+u\,c$ where u is a number, and c is called the central element. The commutator is defined as:

$$[(U(x),u),(V(x),v)] \equiv ([U(x),V(x)],\omega(U,V))$$

where ω is a bilinear antisymmetric form. The element $c=(0,1)$ commutes with everything. The Jacobi identity for this bracket becomes the cocycle condition:

$$\omega([U,V],W)+\omega([W,U],V)+\omega([V,W],U)=0$$

Note that this is a linear condition on ω so the sum of two cocycles is a cocycle. Trivial cocycles are given by $\Sigma([U,V])$ where Σ is linear form on the algebra of matrices. There is a natural invariant bilinear form on the loop algebra:

$$(U,V)=\int dx\,\mathrm{Tr}\,(U(x)V(x))$$

With this we can express the linear form Σ as $\Sigma(U)=(\Sigma,U)=\int dx\,\mathrm{Tr}\,(\Sigma(x)U(x))$. The standard non trivial cocycle is:

$$\omega_0(U,V)=\int dx\,\mathrm{Tr}\,(\partial_x U(x)V(x))$$

and we we can take for ω any linear combination

$$\omega(U,V)=\omega_0(U,V)+\Sigma([U,V])$$

The dual $\mathcal{G}^*$ can be identified with $\mathcal{G}$ using the non–degenerate bilinear form $(U+u\,c,V+v\,c)=(U,V)+uv$. Then the coadjoint action of $\mathcal{G}$ on $\mathcal{G}^*$ reads:

$$\mathrm{ad}^*_{(V,v)}\,(U,u)=(u\partial V+[U+\Sigma,V],0) \tag{9.32}$$

This is because:

$$\begin{aligned}\Big(\mathrm{ad}^*_{(V,v)}\,(U,u)\Big)(W,w) &= \Big((U,u),[(V,v),(W,w)]\Big)=(U,[V,W])+u\omega(V,W)\\ &= \int dx\,\mathrm{Tr}\,\Big([U,V]+u\partial V+[\Sigma,V]\Big)W\end{aligned}$$

Note that u is invariant by the coadjoint action (9.32) so in the following we take $u=1$. Then the coadjoint action of (V,v) becomes a gauge transformation on the operator $\partial+U+\Sigma$ namely:

$$\mathrm{ad}^*_{(V,v)}\,(U,1)=(U',1),\quad \text{with } U'=\partial V+[U+\Sigma,V] \tag{9.33}$$

Remark that the orbit of $(U,1)$ under this action is the set of all $(W,1)$ (up to boundary conditions). This is because one can always solve the equation $\delta U=\partial V+[U+\Sigma,V]$ for V, given any δU.

By construction any orbit of the gauge action in $\mathcal{G}^*$ is equipped with an invariant symplectic form, the Kirillov form. Explicitly, at the point U we have:

$$\{f,g\}(U)=\Big(\mathrm{ad}^*_{df}(U)\Big)(dg)=\Big(\partial df+[U,df],dg\Big)+\Big(\Sigma,[df,dg]\Big) \tag{9.34}$$

where df and dg are viewed as elements of $\mathcal{G}$. The two pieces obviously satisfy the Jacobi identity separately, and so does their sum, hence they define two compatible Poisson brackets. We show below that with a proper choice of Σ and an appropriate symplectic reduction, these two Poisson brackets reduce to the brackets $\{\ ,\ \}_1$ and $\{\ ,\ \}_2$ considered in the previous section, which are then compatible by construction.

We choose to reduce by a subgroup of the loop group, namely by the loop group N_- of lower triangular matrices with 1 on the diagonal. The dual of its Lie algebra can be identified with the algebra $\mathcal{N}_+$ of strictly upper triangular matrices. The Hamiltonian which generates the gauge action by (V, v) is simply $H_{(V,v)}(U, u) = (U, V)$ as in the general theory, see Chapter [14]. The moment at the point $(U, 1)$ is $\mathcal{P}(U) = P_{\mathcal{N}_+}U$. We take for the moment the matrix $\mu \in \mathcal{N}_+$ where $\mu_{k,k+1} = -1$ for $k = 1, \cdots, n$, and all other $\mu_{ij} = 0$. Note that:

$$\mu = -\sum_{i=1}^{n} E_{i,i+1} = -\mathcal{E}_+$$

where $\mathcal{E}_+$ is the sum of simple root vectors of the Lie algebra $sl(n+1)$.

The group of stability of μ is the whole group N_-. Indeed, the variation of the moment under the coadjoint action of N_- is given by $\delta\mu = P_{\mathcal{N}_+}([\mu + \Sigma, V] + \partial V)$, where $V \in \mathcal{N}_-$. We make the further choice that $\Sigma \in \mathcal{N}_-$. Then $[\Sigma, V] + \partial V \in \mathcal{N}_-$. Finally $[\mu, V]_{ij} = -V_{i+1,j} + V_{i,j-1}$. The projection $P_{\mathcal{N}_+}$ means that we keep the matrix elements with $i < j$, or $i + 1 \leq j$. Then the right hand side vanishes because $V \in \mathcal{N}_-$ means that $V_{ij} = 0$ for $i \leq j$. The reduced phase space is the set of $U(x)$ such that $P_{\mathcal{N}_+}U = \mu$, quotiented by the group N_-, with group action (9.33).

We describe the reduced phase space when:

$$\Sigma = \alpha\Sigma_0, \quad \text{with } \Sigma_0 = E_{n+1,1} \tag{9.35}$$

This choice will lead to the Poisson bracket $\{\ ,\ \}_1$. Then $[\Sigma, V] = 0$ for any $V \in \mathcal{N}_-$, and the coadjoint action of N_- reduces to a gauge action on $\partial + U$. The set $\mathcal{M}_\mu$ of points with given momentum μ can be identified with the set of differential operators $\partial + B + \mu$ where B is a lower triangular matrix (including the diagonal) with elements functions of x. The group N_- acts by gauge transformation on this operator, and leaves its form invariant.

One can use this gauge action to bring $\partial + B + \mu$ to either one of the two forms (note that B has $(n+1)$ more parameters than N_-):

$$D_p = \partial + \begin{pmatrix} p_1 & -1 & 0 & 0 & \cdots \\ 0 & p_2 & -1 & 0 & \cdots \\ \vdots & & \ddots & \ddots & \\ 0 & \cdots & & p_n & -1 \\ 0 & \cdots & \cdots & & p_{n+1} \end{pmatrix} \qquad D_u = \partial + \begin{pmatrix} 0 & -1 & 0 & 0 & \cdots \\ 0 & 0 & -1 & 0 & \cdots \\ \vdots & & \ddots & \ddots & \\ 0 & \cdots & & 0 & -1 \\ u_0 & u_1 & \cdots & & u_n \end{pmatrix} \tag{9.36}$$

Since we consider the loop algebra of traceless matrices we have $\sum_i p_i = 0$ and $u_n = 0$. The sets $(p_1, \cdots, p_{n+1})$ or $(u_1, \cdots, u_n)$ constitute two different coordinate systems on the reduced phase space $F_\mu = \mathcal{M}_\mu / N_-$.

Remark that to any point in $\mathcal{M}_\mu$, i.e. to any matrix differential operator $D = \partial + B + \mu$ one can associate a scalar differential operator of order $(n+1)$. To do that we consider the matrix differential equation $D\Psi = 0$ and write the differential equation of order $(n+1)$ induced on the first component ψ_1 of Ψ, which is of the form $L\psi_1 = 0$. Since the group action of N_- on Ψ leaves ψ_1 invariant, this differential equation is invariant under gauge transformations, so the coefficients of the equation are invariant functions under N_-. For the two particular forms D_p, D_u we get:

$$L = (\partial + p_{n+1}) \cdots (\partial + p_1), \quad L = \partial^{n+1} + u_{n-1}\partial^{n-1} + \cdots + u_1\partial + u_0$$

The coordinates u_i are invariant functions but beware that the p_i are not invariant.

It remains to express the reduced Poisson bracket in terms of these variables. We recall that the reduced bracket of invariant functions can be computed straightforwardly using the Poisson bracket of the unreduced phase space, see Chapter [14]. We take as invariant functions the functions $f_X = < LX >$, where X is any pseudo differential operator on $\mathcal{P}_-$, and $< \ >$ is the Adler trace. Hence we have at the point $(p+\mu)$ (where p is diagonal) $\{f_X, f_Y\}_{\mathrm{reduced}} = (\mathrm{ad}^*_{df_X}(p+\mu), df_Y)$. We first need to compute the variation of L when $p+\mu \to p+\mu+b$ where b is a small lower triangular matrix. Writing the system $(\partial + p + b + \mu)\Psi = 0$ in terms of ψ_1 and keeping only terms of first order in b we find, with the notations (9.27):

$$\delta L = \sum_{i \geq j} L_{n+1,i+1} b_{ij} L_{j-1,1}$$

This defines df_X by the relation $(df_X, b) = < \delta L, X >$ so that df_X is the upper triangular traceless matrix:

$$(df_X)_{ji} = \Big(L_{j-1,1} X L_{n+1,i+1}\Big)_{-1} - \frac{1}{n+1}\delta_{ij} \sum_k \Big(L_{k-1,1} X L_{n+1,k+1}\Big)_{-1}$$

We can now compute the coadjoint action:

$$\mathrm{ad}^*_{df_X}(p+\mu) = \partial df_X + [p + \mu + \Sigma, df_X]$$

Noting that df_Y is upper triangular, we need only to keep the lower triangular part in $\mathrm{ad}^*_{df_X}(p+\mu)$. Using eq.(9.34) and noting that the Σ independent term in this expression is upper triangular, we need only to keep the diagonal part in this term, that is $(df_X)_{kk}$. We get:

$$\{f_X, f_Y\}_{\mathrm{reduced}} = \sum_k (\partial (df_X)_{kk})(df_Y)_{kk} + ([\Sigma, df_X], df_Y)$$

Substituting the expressions of df_X and df_Y the Σ independent term immediately yields eq.(9.28), which shows that this part coincides with the bracket $\{\ ,\ \}_2$. This also shows that the Poisson bracket of the diagonal coordinates $p_i(x)$ is given by eq.(9.26).

We now look at the Σ dependent term and show that it reproduces the bracket $\{\ ,\ \}_1$. Due to the specific choice of Σ we have:

$$([\Sigma_0, df_X], df_Y) = \sum_{j=1}^{n+1} (df_X)_{1,j} (df_Y)_{j,n+1} - (df_Y)_{1,j}(df_X)_{j,n+1}$$

Inserting the expressions of df_X and df_Y and noting that the terms proportional to the identity do not contribute, we get:

$$([\Sigma_0, df_X], df_Y) = \sum_{j=1}^{n+1} \int dx\, (X L_{n+1,j+1})_{-1} (L_{j-1,1} Y)_{-1} - (Y L_{n+1,j+1})_{-1}(L_{j-1,1}X)_{-1}$$

Using eq.(9.30) this can be rewritten in terms of Adler traces:

$$([\Sigma_0, df_X], df_Y) = \sum_{j=1}^{n+1} < (X L_{n+1,j+1})_- L_j (L_{j-1,1} Y)_- > - < (Y L_{n+1,j+1})_- L_j (L_{j-1,1} X)_- >$$

Substituting everywhere $U_- = U - U_+$ and noting that $< U, V_+ > = < U_-, V >$ this reads:

$$([\Sigma_0, df_X], df_Y) = \sum_{j=1}^{n+1} < XLY > - < (X L_{n+1,j})_- L_{j-1,1} Y > - < (X L_{n+1,j+1})_+ L_{j,1} Y > \quad - \quad (X \leftrightarrow Y)$$

In the sum the first term yields $(n+1) < XLY >$, while the second and third terms regroup themselves as $-\sum_{j=2}^{n+1} < \Big((X L_{n+1,j})_- + (X L_{n+1,j})_+\Big) L_{j-1,1} Y >$, that is $-n < XLY >$, so that we finally get:

$$([\Sigma_0, df_X], df_Y) = < XLY > - < YLX > = - < L[X,Y] > = -\{f_X, f_Y\}_1$$

This construction, due to Drinfeld and Sokolov, can be generalized to Lie algebras other than $sl(n+1)$, replacing $\mathcal{E}_+$ by the sum of simple root vectors and Σ_0 by the root vector $E_{-\alpha}$ where α is the longest root. This also shows the nice interplay between different Lie algebra structures (the one induced by the algebra of pseudo differential operators, and the Kac Moody one) producing the same Kirillov Poisson brackets, after The crucial step of suitable Hamiltonian reduction.

References

[1] I.M. Gelfand, L.A. Dickey *Fractional powers of operators and Hamiltonian sytems.* Funkz. Anal. Priloz. 10 (1976) 13–29

[2] F. Magri *A simple model of integrable Hamiltonian equation.* J. Math. Phys. 19 (1978) 1156–1162

[3] M. Adler *On a trace functional for formal pseudo–differential operators and the symplectic structure of the Korteweg–de Vries equations.* Invent. Math. 50 (1979) 219–248

[4] D.R. Lebedev, Yu.I. Manin *Hamiltonian Gelfand–Dickey operator and coadjoint representation of the Volterra group.* Funkz. Analys. Priloz. 13 (1979) 40–46

[5] A.G. Reyman, M.A. Semenov–Tian–Shansky *Family of Hamiltonian structures, hierarchy of Hamiltonians, and reduction for matrix first order differential operators.* Funkz. Analys. Priloz. 14 (1980) 77–78

[6] V.G. Drinfeld, V.V. Sokolov *Equations of the Korteweg–de Vries type and simple Lie algebras.* Doklady AN SSSR 258 (1981) 11–16

[7] M.Jimbo, T.Miwa *"Solitons and Infinite Dimensional Lie Algebras."* RIMS 19 (1983) 943-1001.

[8] L.A. Dickey **"Soliton Equations and Hamiltonian Systems."** World Scientific (1991)

Chapter 10

The KdV hierarchy.

In this Chapter we study the Korteweg–de Vries equation, which occupies the central place in the modern theory of integrable systems. All the aspects of integrable systems discussed so far converge in this Chapter to draw the particularly rich landscape of this model. In particular the methods of pseudo–differential operators allow to easily discuss the formal aspects, the τ–functions yield soliton solutions, and the algebro–geometric methods yield finite zone solutions. The soliton solutions which we obtained in the Grassmannian setting by using vertex operators are also degenerate cases of these finite–zone solutions.

10.1 The KdV equation.

The Korteweg–de Vries (KdV) equation has been introduced historically as an approximation of the equations of hydrodynamics, describing unidimensional long waves in shallow water. Since then, it became the most central equation in the modern studies of integrable systems, following the pioneering work of Gardner, Greene, Kruskal and Miura who found an unexpected connection with the inverse scattering problem. This equation reads:

$$4\partial_t u = 6u\partial_x u + \partial_x^3 u \tag{10.1}$$

The numerical factors in front of each terms in eq.(10.1) can be modified by rescaling u, x and t. The integrable features of the KdV equation can be unveiled by introducing a connection A_x, A_t depending on a spectral parameter, whose zero curvature condition is eq.(10.1).

The connection for KdV consists of 2×2 matrix with a spectral parameter λ:

$$A_x = \begin{pmatrix} 0 & -1 \\ -\lambda + u & 0 \end{pmatrix}, \quad A_t = -\frac{1}{4} \begin{pmatrix} -\partial_x u & 4\lambda + 2u \\ 4\lambda^2 - 2\lambda u - \partial_x^2 u - 2u^2 & \partial_x u \end{pmatrix} \tag{10.2}$$

Remark that $\partial_x + A_x = D_u - \lambda \Sigma_0$ with the notations of eq.(9.35,9.36) in Chapter [9]. The zero curvature condition $F_{xt} \equiv \partial_x A_t - \partial_t A_x + [A_x, A_t] = 0$ yields the KdV equation (10.1).

Alternatively one can recast the KdV equation in the Lax form introducing L and M which are the following differential operators, see Chapter [9]:

$$L = \partial_x^2 + u, \quad M = -\frac{1}{4}(4\partial_x^3 + 3u\partial_x + 3\partial_x u) = -\frac{1}{4}(4\partial_x^3 + 6u\partial_x + 3(\partial_x u)) = -(L^{\frac{3}{2}})_+ \tag{10.3}$$

The Lax equation $\partial_t L = [L, M]$, with [,] the commutator of differential operators, reduces to the KdV equation.

Of course these two descriptions are not independent. To connect them, consider the wave function Ψ solution of the linear system:

$$(\partial_x + A_x)\Psi = 0, \quad (\partial_t + A_t)\Psi = 0 \tag{10.4}$$

Setting $\Psi = \begin{pmatrix} \psi \\ \chi \end{pmatrix}$, the x-equation reads $\chi = \partial_x \psi$ and

$$\Big(L - \lambda\Big)\psi = 0 \quad \text{with} \quad L = \partial_x^2 + u \tag{10.5}$$

The time evolution of ψ is given by $4\partial_t\psi = -\partial_x u \cdot \psi + (4\lambda + 2u)\,\partial_x\psi$. Using eq(10.5), this may be rewritten as:

$$\Big(\partial_t + M\Big)\psi = 0 \quad \text{with} \quad M = -\frac{1}{4}(3u\partial + 3\partial u + 4\partial^3) \tag{10.6}$$

The compatibility condition of eqs.(10.5,10.6) on ψ is the Lax equation $\partial_t L = [L, M]$, which is equivalent to the KdV equation.

As explained in Chapter [2], a general consequence of the zero curvature condition is the existence of non trivial conserved quantities. This requires however imposing appropriate boundary conditions. We shall consider here for definiteness either potentials $u(x)$ with fast decrease at $x \to \pm\infty$, or periodic boundary conditions on the interval $x \in [0, \ell]$, namely we require that A_x and A_t are periodic under $x \to x + \ell$. In this case conserved quantities are generated by $\mathrm{Tr}\, T_\ell(\lambda)$, where T_ℓ is the monodromy matrix associated to the linear system (10.4):

$$T_\ell(\lambda) = P \exp\left(-\int_0^\ell dt A_x(x,t)\right) \tag{10.7}$$

To compute this trace, we remark that it is invariant under periodic gauge transformations, and we build a gauge where the connection is diagonal.

We shall present this computation for general connections whose components A_x and A_t belong to the $sl(2)$ algebra, so that it can be applied to a wider class of systems. The commutation relations in the $sl(2)$ algebra are:

$$\begin{aligned} [H, E_\pm] &= \pm 2E_\pm \\ [E_+, E_-] &= H \end{aligned}$$

Its fundamental representation is given by:

$$H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad E_+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad E_- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

Proposition. *Let $A_x = A_h H + A_- E_- + A_+ E_+$ where $A_h(x,t)$, $A_\pm(x,t)$ are periodic functions of x. There exists a periodic gauge transformation $A_x \to^g A \equiv g^{-1}A_x g + g^{-1}\partial_x g$ such that ${}^gA_- = {}^gA_+ = 0$, and*

$${}^gA_h = \frac{1}{\ell}P_\ell(\lambda)H$$

where $P_\ell(\lambda)$ independent of x, is given by: $P_\ell(\lambda) = i\int_0^\ell dx\ v$. The function v is solution of the Ricatti equation:

$$v' + v^2 = V \quad \text{with} \quad V = \Phi' + \Phi^2 + A_- A_+ \quad \text{and} \quad \Phi = -A_h - \frac{1}{2}\frac{A_+'}{A_+} \tag{10.8}$$

The function $P_\ell(\lambda)$ is a generating functional for the conserved quantities.

Proof. The proof consists in implementing three gauge transformations: the first one annihilates the component along E_-, the second annihilates the component along E_+ and the third is choosen to ensure that

the component along H is constant.
Implement first a gauge transformation with $g = g_1 = \exp(f_- E_-)$ then

$$^gA_x = (A_h + A_+ f_-)H + (f'_- - 2A_h f_- - A_+ f_-^2 + A_-)E_- + A_+ E_+$$

where prime ($'$) means derivative with respect to x. The coefficient of E_- vanishes if f_- is a solution of the Ricatti equation

$$f'_- - 2A_h f_- - A_+ f_-^2 + A_- = 0$$

If one sets $f_- = -\frac{1}{A_+}(v - \Phi)$ and $\Phi = -A_h - \frac{1}{2}\frac{A'_+}{A_+}$ this equation becomes the Ricatti equation (10.8). As usual the substitution $v = y'/y$ linearizes the equation which then becomes the Schroedinger equation:

$$y'' - Vy = 0 \tag{10.9}$$

The potential V being periodic, one can take for y any one of the two quasi-periodic Bloch waves $y_\pm$:

$$y_\pm(x + \ell) = \exp(\pm P_\ell(\lambda))\ y_\pm(x) \tag{10.10}$$

$P_\ell(\lambda)$ is called the quasi-monentum. For definiteness, we shall take $v = y'_+/y_+$. We shall moreover assume that the Wronskian is normalized by $y_+ y'_- - y'_+ y_- = 1$. Notice that

$$P_\ell(\lambda) = \ln\left(\frac{y_+(\ell)}{y_+(0)}\right) = \int_0^\ell dx\ v \tag{10.11}$$

Similarly let us define $g_2 = \exp(f_+ E_+)$ and compute the connection gauge transformed with $g = g_1 g_2$:

$$^gA_x = (A_h + A_+ f_-)H + (f'_+ + 2(A_h + A_+ f_-)f_+ + A_+)E_+$$

We reduce ^gA_x to the diagonal form if we choose for f_+ the periodic solution of the equation $f'_+ + 2(A_h + A_+ f_-)f_+ + A_+ = 0$ whose solution is $f_+ = -A_+ y_+ y_-$.
Finally taking $g_3 = \exp(hH)$, the gauge transformed Lax connection ^gA_x with $g = g_1 g_2 g_3$ is:

$$^gA_x \quad = \quad (h' + A_h + A_+ f_-)\,H = \left(h' - v - \frac{1}{2}\frac{A'_+}{A_+}\right) H = \left(h' - \frac{y'_+}{y_+} - \frac{1}{2}\frac{A'_+}{A_+}\right) H$$

Choosing h the periodic function $h = \frac{1}{2}\ln(A_+ y_+^2 e^{-2P_\ell(\lambda)x/\ell})$ reduces the coefficient of H to a constant. ■

Conserved quantities are obtained by looking at the trace of the monodromy matrix $\mathrm{Tr}\,(T_\ell(\lambda))$. Once the connection has been diagonalized this trace is easy to compute. It follows from the previous proposition that the two eigenvalues of $T_\ell(\lambda)$ are $\exp(\pm P_\ell(\lambda))$:

$$\mathrm{Tr}\, T_\ell(\lambda) = 2\cosh P_\ell(\lambda), \quad P_\ell(\lambda) = \int_0^\ell dx\ v(x, \lambda)$$

The functional $P_\ell(\lambda)$ can serve as well as $\mathrm{Tr}\, T_\ell(\lambda)$ as a generating functional for integrals of motion. To construct them we have to solve the Ricatti equation (10.8) for $v(x, \lambda)$.

Let us now apply this proposition to the KdV equation. In view of the expression of the KdV connection (10.2), we have $A_h = 0$, $A_- = -\lambda + u$ and $A_+ = -1$. Thus $V = \lambda - u$ and the Ricatti equation reads:

$$v' + v^2 = \lambda - u \tag{10.12}$$

The Schroedinger equation associated to the Ricatti eq.(10.12) coincides with eq.(10.9). The quantity $P_\ell(\lambda)$ is the quasi–momentum of the Bloch eigenfunctions of the differential operator $L = \partial^2 + u$ with periodic potential u and eigenvalue λ.

To obtain *local* conserved quantities we expand around $\lambda = \infty$. When $\lambda \to \infty$, the solution of the Ricatti equation admits the asymptotic expansion:

$$v = \sqrt{\lambda} + \sum_{n\geq 0} \frac{(-1)^n v_n}{(\sqrt{\lambda})^n}, \quad 2v_{n+1} = v_n' + \sum_{p=0}^{n} v_p v_{n-p}, \quad v_0 = 0,\ 2v_1 = u$$

This gives a recursion relation for computing the coefficients v_n. Since it does not require any integration it leads to local integrals of motion. The first few coefficients are:

$$\begin{aligned}
v_1 &= \tfrac{1}{2}u \\
v_2 &= \tfrac{1}{4}u' \\
v_3 &= \tfrac{1}{8}(u^2 + u'') \\
v_4 &= \tfrac{1}{2}v_3' + \tfrac{1}{8}uu' \\
v_5 &= \tfrac{1}{2}v_4' + \tfrac{1}{32}u'^2 + \tfrac{1}{16}uu'' + \tfrac{1}{16}u^3
\end{aligned}$$

The conserved quantities are given by the integral

$$P_\ell(\lambda) = \int_0^\ell v\,dx = \sqrt{\lambda}\ell - \frac{1}{2\sqrt{\lambda}}\left(\int_0^\ell u\,dx\right) - \frac{1}{(2\sqrt{\lambda})^3}\left(\int_0^\ell u^2\,dx\right) - \frac{1}{(2\sqrt{\lambda})^5}\left(\int_0^\ell (u'^2 - 2u^3)dx\right) - \cdots \tag{10.13}$$

10.2 The KdV hierarchy.

We now particularize the formalism of pseudo–differential operators studied in Chapter [9] to the KdV situation. Namely we study the implications of the condition:

$$\Phi\partial^2\Phi^{-1} = L = \partial^2 + u, \quad \Phi = 1 + \sum_1^\infty w_i \partial^{-i}$$

A first consequence is that only odd times survive in the KdV hierarchy. Indeed, recalling the equations of motion of the KdV hierarchy $\partial_{t_i}\Phi = -(\Phi\partial^i\Phi^{-1})_-\Phi$, we see that for $i = 2j$, $(\Phi\partial^{2j}\Phi^{-1}) = L^j$ is a differential operator, so that its projection $(\ )_-$ vanishes.

Recall that we have defined two formal Baker–Akhiezer functions:

$$\psi(t,k) = \Phi e^{\xi(t,k)}, \quad \psi^*(t,k) = (\Phi^*)^{-1}e^{-\xi(t,k)}, \quad \xi(t,k) = \sum_{i=1}^\infty t_{2i-1}k^{2i-1}$$

where $t_1 = x$. The function ψ is an eigenfunction of L with the eigenvalue $\lambda = k^2$. Since L is obviously formally self–adjoint (recall that we have defined a $*$ operator such that $\partial^* = -\partial$), ψ^* is also an eigenfunction of L with the same eigenvalue.

$$(\partial^2 + u)\psi = k^2\psi, \quad (\partial^2 + u)\psi^* = k^2\psi^* \tag{10.14}$$

The Wronskian of these two solutions is a constant that we shall now compute.

Proposition. *The Wronskian of the two Baker–Akhiezer functions ψ and ψ^* is given by:*

$$W(\psi,\psi^*) \equiv \psi'(t,k)\psi^*(t,k) - \psi^{*\prime}(t,k)\psi(t,k) = 2k \tag{10.15}$$

where we have denoted $\psi' \equiv \partial_x \psi$.

Proof. From the definition of the Baker–Akhiezer functions we see that the essential singularities cancel in $W(\psi, \psi^*)$ and that this admits a power series development in k around ∞ of the form $W(k) = 2k + \alpha + \beta/k + \cdots$. We show that only the first term is present by computing the residue of $W(k)k^i$ for $i \geq -1$.

$$\oint \frac{dk}{2i\pi} k^i W(k) = \oint \frac{dk}{2i\pi} \Big(\partial\Phi\partial^i e^{kx}\Big)\Big((\Phi^*)^{-1} e^{-kx}\Big) - \Big(\Phi e^{kx}\Big)\Big(\partial(\Phi^*)^{-1}(-\partial)^i e^{-kx}\Big)$$

Note that the terms involving the times $t_3, t_5, \cdots$ in $\xi(t,k)$ cancel because there are only derivatives with respect to $x = t_1$. Using eq.(9.11) in Chapter [9] we can rewrite this as an Adler residue in the pseudo–differential algebra:

$$\oint \frac{dk}{2i\pi} k^i W(k) = \mathrm{Res}_\partial \Big(\partial\Phi\partial^i\Phi^{-1} + \Phi\partial^i\Phi^{-1}\partial\Big) = \mathrm{Res}_\partial \Big(\partial L^{\frac{i}{2}} + L^{\frac{i}{2}}\partial\Big)$$

For i even, this vanishes because $L^{\frac{i}{2}}$ is a differential operator. For i odd we are going to show that:

$$L^{\frac{i}{2}*} = -L^{\frac{i}{2}}, \quad i = \mathrm{odd}$$

so that $\partial L^{\frac{i}{2}} + L^{\frac{i}{2}}\partial$ is formally self–adjoint and so cannot have a residue. It is sufficient to show that $L^{\frac{1}{2}*} = -L^{\frac{1}{2}}$. But $L^{\frac{1}{2}} = \partial + \frac{1}{2}u\partial^{-1} - \frac{1}{4}u'\partial^{-2} + \cdots$ is the unique solution of the equation $(L^{\frac{1}{2}})^2 = L$ with leading term ∂. So its opposite is the unique solution of the same equation with leading term $-\partial$. Since $(L^{\frac{1}{2}*})^2 = L^* = L$ and $L^{\frac{1}{2}*} = -\partial + \cdots$ the result follows. ■

We introduce the quantity:

$$S(t,k) \equiv \psi^*(t,k)\psi(t,k) = 1 + \sum_{n=1}^{\infty} k^{-2n} S_{2n}(t)$$

Note that the essential singularities cancel in $S(t,k)$ and that $S(t,k) = S(t,-k)$. This is because, $\psi^*(t,k)$ and $\psi(t,-k)$ are solutions of the same eq.(10.14), and have the same behaviour at ∞ so we have $\psi^*(t,k) = c(k, t_3, \cdots)\psi(t,-k)$. Inserting this into eq.(10.15) we see that $c(k,t)$ is even in k. It follows that $S(t,k)$ is even in k. The aim of the following is to show that the whole KdV hierarchy can be written in terms of $S(t,k)$.

Proposition. *The Baker–Akhiezer functions can be expressed as:*

$$\psi(t,k) = S^{1/2}(t,k)e^{X(t,k)}, \quad \psi^*(t,k) = S^{1/2}(t,k)e^{-X(t,k)}, \quad \partial_x X(t,k) = \frac{k}{S(t,k)} \tag{10.16}$$

The quantity $X(t,k)$ is of the form $X(t,k) = \xi(t,k) + O(1/k)$.

Proof. This parametrization obviously satisfies $\psi^*\psi = S$ and this defines $X(t,k)$. Inserting it into eq.(10.15) yields the last equation. ■

As a consequence, eq.(10.14) translates into an equation on S. It is convenient to write it in the form of the Ricatti equation (10.12) with:

$$v = \frac{\partial_x \psi}{\psi} = \frac{1}{2}\partial_x \log S + \frac{k}{S} \tag{10.17}$$

Replacing ψ by ψ^* in this equation amounts to changing $k \to -k$. In particular this shows that the v_{2n} are derivatives with respect to x of local densities.

There is a simple expression of the coefficients S_{2j} as residues of fractional powers of L.

Proposition. *The coefficients S_{2j} are the local densities of the conserved quantities of the KdV hierarchy, as computed in Chapter [9].*

$$S_{2j} = \mathrm{Res}_\partial \left(L^{\frac{2j-1}{2}} \right) \tag{10.18}$$

Proof.

$$S_{2j} = \oint \frac{dk}{2i\pi} k^{2j-1} \psi^*(t,k)\psi(t,k) = \oint \frac{dk}{2i\pi} (\Phi^*)^{-1} e^{-kx} \Phi \partial^{2j-1} e^{kx} = \mathrm{Res}_\partial (\Phi \partial^{2j-1} \Phi^{-1}) = \mathrm{Res}_\partial \left(L^{\frac{2j-1}{2}} \right)$$

where we have used eq.(9.11) in Chapter [9]. ■

One can compute the coefficients S_{2k} by recursion as follows. Since ψ and ψ^* obey eq.(10.14) their product $S(k) = \psi\psi^*$ obeys the third order equation:

$$(\frac{1}{4}\partial^3 + \frac{1}{2}u' + u\partial - k^2 \partial) S(k) = 0$$

Expanding in k one gets $S'_{2n+2} = (\frac{1}{4}\partial^3 + \frac{1}{2}u' + u\partial) S_{2n}$. This recursion relation can also be understood as the Lenard recursion relation eq.(9.22) in Chapter [9]. Indeed since

$$H_k(L) = \frac{2}{k+2} < L^{\frac{k+2}{2}} >= \frac{2}{k+2} \int dx \mathrm{Res}_\partial L^{\frac{k}{2}+1} = \frac{2}{k+2} \int dx S_{k+3} \tag{10.19}$$

and using eq.(9.21) we get $dH_k = S_{k+1}\partial^{-1}$. The Lenard relation becomes $\mathcal{D}_1(S_{2n+2}\partial^{-1}) = \mathcal{D}_2(S_{2n}\partial^{-1})$ which is exactly the previous recursion relation. The first few values of S_{2j} are:

$$S_2 = \frac{1}{2}u, \quad S_4 = \frac{1}{8}u'' + \frac{3}{8}u^2, \quad S_6 = \frac{5}{16}u^3 - \frac{5}{32}u'^2 + \frac{5}{16}(uu')' + \frac{1}{32}u^{(iv)}$$

One can recast the equations of motion of the KdV hierarchy as equations on $S(t,k)$. It is convenient to introduce a generating function for all the time derivatives:

$$D(k) = \sum_{j \geq 1} k^{-2j} \partial_{2j-1} \tag{10.20}$$

Then we have:

Proposition. *The equations of the KdV hierarchy are equivalent to:*

$$D(k)S(t,k') = \frac{S(t,k)\cdot \partial_x S(t,k') - S(t,k')\cdot \partial_x S(t,k)}{k^2 - k'^2}, \quad |k| > |k'| \tag{10.21}$$

Proof. We shall first prove a similar equation on $\psi(t,k)$:

$$D(k)\psi(t,k') = \frac{2S(t,k)\partial_x - \partial_x S(t,k)}{2(k^2 - k'^2)} \psi(t,k'), \quad \text{for} \quad |k| > |k'| \tag{10.22}$$

Recall that, from eq.(9.8) in Chapter [9], the equation of motion for ψ is $\partial_{2j-1}\psi(t,k) = (L^{(2j-1)/2})_+ \psi(t,k)$. There is a simple recursion relation:

$$\left(L^{\frac{2j+1}{2}} \right)_+ = \left(L^{\frac{2j-1}{2}} \right)_+ L + S_{2j}\partial_x - \frac{1}{2}\partial_x S_{2j} \tag{10.23}$$

This is because, since L is a differential operator:

$$\left(L^{\frac{2j+1}{2}} \right)_+ = \left(L^{\frac{2j-1}{2}} L \right)_+ = \left(L^{\frac{2j-1}{2}} \right)_+ L + \left(\left(L^{\frac{2j-1}{2}} \right)_- L \right)_+$$

To compute the plus part in the last term we only have to keep the first two terms in the expansion of $(L^{(2j-1)/2})_-$ because L is a second order differential operator. We have:

$$\left(L^{\frac{2j-1}{2}}\right)_- = S_{2j}\partial^{-1} - \frac{1}{2}S'_{2j}\partial^{-2} + \cdots$$

where the first coefficient (which is the residue of the considered pseudo–differential operator) is determined by eq.(10.18), and the second coefficient is then fixed by the fact that the left hand–side is formally anti self–adjoint. The recursion relation eq.(10.23) follows immediately, and implies in turn:

$$\left(L^{\frac{2j-1}{2}}\right)_+ = \sum_{i=0}^{j-1}(S_{2i}\partial - \frac{1}{2}S'_{2i})L^{j-i-1} \tag{10.24}$$

Using $L\psi(t,k') = k'^2\psi(t,k')$, we have:

$$\partial_{2j-1}\psi(t,k') = \sum_{i=0}^{j-1}(S_{2i}\partial - \frac{1}{2}S'_{2i})k'^{2(j-i-1)}\psi(t,k')$$

The computation of $D(k)\psi(t,k')$ is then straightforward. The Baker–Akhiezer function ψ^* obeys the same equation (10.22) and eq.(10.21) follows immediately for $S = \psi^*\psi$. ■

We can give a more precise description of the decomposition eq.(10.16). Using eq.(10.22) we have:

$$D(k)X(k') = \frac{k'}{k^2 - k'^2}\left(\frac{S(k)}{S(k')}\right) = \frac{k'}{k^2 - k'^2}\left(\frac{S(k) - S(k')}{S(k')}\right) + \frac{k'}{k^2 - k'^2}, \quad \text{for} \quad |k| > |k'| \tag{10.25}$$

Of course $X(k) \equiv X(t,k)$ is a function of k but also of the times variables t_{2j-1}. Notice that we can expand $X(k)$ as:

$$X(t,k) = \xi(t,k) + \widetilde{X}(t,k), \quad \text{with} \quad \xi(t,k) = \sum_{j\geq 1} k^{2j-1}t_{2j-1} \tag{10.26}$$

with $\widetilde{X}(t,k)$ regular at $k = \infty$. This decomposition follows from the fact that $D(k)\xi(k',t) = \frac{k'}{k^2 - k'^2}$ for $|k| > |k'|$.

Using this formalism we now show that the conserved quantities given in eq.(10.13) are the same as the ones in eq.(10.19). This amounts to showing that v and S differ by the derivative in x of a local function.

Proposition. *We have the relation:*

$$\frac{dv}{dk} = S + \frac{1}{2}\partial_x\Big(\frac{d}{dk} + D(k)\Big)\log S \tag{10.27}$$

where $v = \partial_x\psi/\psi$.

Proof. From eq.(10.21) we have:

$$D(k)\log\, S(k') = \frac{S(k)}{k^2 - k'^2}\Big(\partial_x \log\, S(k') - \partial_x \log\, S(k)\Big)$$

We substitute eq.(10.17) in the previous equation and take the limit $k' \to k$. One gets:

$$D(k)\log\, S(k) = -\frac{S(k)}{k}\frac{dv}{dk} + \frac{1}{k} - \frac{d}{dk}\log\, S(k)$$

We derive this expression with respect to x and substitute $\partial_x v = k^2 - u - v^2$ to get the result. ■

Remark. As for the KP-hierarchy, there exists a tau-function $\tau(t)$ such that:

$$\psi(t,k) = e^{\xi(t,k)}\ \frac{\tau(t-[k])}{\tau(t)}, \quad \psi^*(t,k) = e^{-\xi(t,k)}\ \frac{\tau(t+[k])}{\tau(t)}$$

with $[k] = (\cdots, \frac{k^{-2j+1}}{2j-1}, \cdots)$. It is related to the generating function $S(t,k)$ by:

$$S(t,k) = \frac{\tau(t+[k])\tau(t-[k])}{\tau^2(t)} = 1 + \partial_x D(k) \log \tau(t) \tag{10.28}$$

This last formula follows by inserting the parametrization of ψ, ψ^* in terms of τ functions in eq.(10.27).

10.3 Hamiltonian structures.

Recall eq.(9.24) in Chapter [9] which define two Poisson structures:

$$\begin{aligned} \{f_X, f_Y\}_1(L) &= < LXY > - < LYX > \\ \{f_X, f_Y\}_2(L) &= < (LX)_+(LY)_- > - < (XL)_+(YL)_- > - \frac{1}{2}\int dx \left((\partial^{-1}[L,X]_{-1})\,[L,Y]_{-1} \right) \end{aligned}$$

Here $L = \partial^2 + u$, $X = X_{-1}\partial^{-1} + X_{-2}\partial^{-2} + \cdots$ and $f_X(L) =< LX >$, the Adler trace of the operator LX. To compute the Poisson brackets of u it is enough to take $X = X(x)\partial^{-1}$ so that $f_X(L) = \int dxu(x)X(x)$. The two Poisson brackets become:

$$\begin{aligned} \left\{ \int dxuX, \int dxuY \right\}_1 &= \int dx(XY' - X'Y) \\ \left\{ \int dxuX, \int dxuY \right\}_2 &= \int dxu(X'Y - XY') - \frac{1}{2}\int dxXY''' \end{aligned} \tag{10.29}$$

This means that we have:

$$\{u(x), u(y)\}_1 = (\partial_x - \partial_y)\delta(x-y) \tag{10.30}$$

$$\{u(x), u(y)\}_2 = -\frac{1}{2}(u(x)+u(y))(\partial_x - \partial_y)\delta(x-y) - \frac{1}{4}(\partial_x^3 - \partial_y^3)\delta(x-y) \tag{10.31}$$

These two Poisson brackets are local in x. Alternatively we can expand all fields in Fourier components. So we take $u(x) = \sum_k u_k e^{ikx}$, choosing $\ell = 2\pi$, and $X = e^{-inx}$, $Y = e^{-imx}$. The equations (10.29) become:

$$\begin{aligned} \{u_n, u_m\}_1 &= \frac{i}{\pi} n \delta_{n+m,0} \\ \{u_n, u_m\}_2 &= \frac{1}{2i\pi}\Big((n-m)u_{n+m} + \frac{n^3}{2}\delta_{n+m,0} \Big) \end{aligned} \tag{10.32}$$

The bracket $\{\ ,\ \}_1$ is called the Gardner–Faddeev–Zakharov bracket, while the bracket $\{\ ,\ \}_2$ is called the Magri–Virasoro bracket.

The Lax operator L can also be written in factorized form $L = (\partial - p)(\partial + p)$ where p is a function of x such that $u = p' - p^2$. This is the Miura transformation. The second Poisson bracket has a simple expression on this parametrization,

$$\{p(x), p(y)\}_2 = -\frac{1}{4}(\partial_x - \partial_y)\delta(x-y) \tag{10.33}$$

or in Fourier modes $\{p_n, p_m\} = -i\pi n \delta_{n+m,0}$.

It was shown in Chapter [9] that the Hamiltonians H_n are in involution under both Poisson brackets, and that these Poisson brackets are compatible in the sense that their sum is again a Poisson bracket. Let us give an exemple of the Hamiltonian equations of motion. Taking:

$$H_3 = \frac{1}{4}\int dxu^2, \quad H_5 = \frac{1}{8}\int dx(u'^2 - 2u^3)$$

one gets:

$$\partial_{t_3} u = \{H_5, u\}_1 = \{H_3, u\}_2 = \frac{1}{4}(6uu' + u''')$$

illustrating the fact that one finds the same equations of motion with the two Poisson brackets, but with different Hamiltonians.

10.4 Soliton solutions.

Considering the KP hierarchy with $Q = L^{\frac{1}{2}}$ we see that the equations of motion $\partial_{t_i}\Phi = -(L^{\frac{i}{2}})_-\Phi$ imply that Φ is stationnary with respect to the even times t_{2j}. Conversely any solution of the KP hierarchy which is stationnary for any even time is such that $(Q^{2j})_- = 0$, hence in particular $Q^2 = L$ is a differential operator. This solution is thus a solution of the KdV hierarchy. At the τ–function level, to obtain the decoupling of the even time variables it is sufficient to require that:

$$\tau_{KP}(t_{\text{even}}, t_{\text{odd}}) = e^{\sum_{n \text{ even}} c_n t_n} \tau_{KdV}(t_{\text{odd}}) \tag{10.34}$$

This is because the action of a Hirota differential operator with respect to even time variables on such a KP τ–function vanishes:

$$D^m_{t_{2k}} e^{c_{2k}t_{2k}} \tau_{KdV}(t_{\text{odd}}) \cdot e^{c_{2k}t_{2k}} \tau_{KdV}(t_{\text{odd}}) = \partial^m_y e^{c_{2k}(t_{2k}+y)} \tau_{KdV}(t_{\text{odd}}) e^{c_{2k}(t_{2k}-y)} \tau_{KdV}(t_{\text{odd}})|_{y=0} = 0$$

The Hirota equations of the KdV hierarchy are thus obtained from the Hirota equations of the KP hierarchy by simply erasing the even times. We get for instance the Hirota form of the KdV equation:

$$(D_1^4 - 4D_1D_3)\tau \cdot \tau = 0 \tag{10.35}$$

where D_{2k-1} are Hirota's differential operators, defined by eq.(8.3) in Chapter [8]. Setting

$$u = 2\frac{\partial^2}{\partial t_1^2} \log \tau = \frac{D_1^2 \tau \cdot \tau}{\tau^2} \tag{10.36}$$

we recover the KdV equation on u. Indeed one has:

$$u'' + 3u^2 = \frac{D_1^4 \tau \cdot \tau}{\tau^2}, \qquad D_1^4 \tau \cdot \tau = 2\tau^{(iv)}\tau - 8\tau'''\tau' + 6(\tau'')^2$$

$$\dot{u} = \partial_x \frac{D_1 D_3 \tau \cdot \tau}{\tau^2}, \qquad D_1 D_3 \tau \cdot \tau = 2(\dot{\tau}'\tau - \tau'\dot{\tau})$$

combining these expressions we see that the KdV equation is equivalent to the Hirota equation (10.35).

Recall that the KP tau–functions are constructed by choosing an element $g \in \widehat{GL(\infty)}$ (see eq.(8.35) in Chapter [9]):

$$\tau_{KP}(t) = \langle 0|e^{h(t)} g|0\rangle$$

with $h(t) = \sum_n h_n t_n$ and h_n bosonic oscillators. This τ–function satisfies the bilinear identity (8.37), which reduces to the KdV Hirota bilinear identity whenever $\tau_{KP}(t)$ is of the form eq.(10.34). The main problem is to find the group elements g such that this property holds.

Recall that the Lie algebra of $\widehat{GL(\infty)}$ consists in bilinear of fermions of the form $X = \sum_{rs} M_{rs} : \psi_r \psi^*_s :$ (see eq.(8.15) in Chapter [8]). So we have to find the subalgebra which commutes with the h_{2k}. One can then push the term $\exp\left(\sum_k t_{2k}h_{2k}\right)$ in τ_{KP} to the right where it hits the vacuum, and disappear since $h_{2k}|0>= 0$. In fact commutation up to a central element is sufficient since a central term would produce an exponential of a linear combination of the even times. Using eq.(8.27) in Chapter [8] we have:

$$[h_n, \psi_r] = \psi_{r+n}, \quad [h_n, \psi^*_s] = -\psi^*_{s+n}$$

so that the h_{2k} commute to X up to central terms if $M_{rs} = M_{r+2,s-2}$. This means that the infinite band matrix M_{rs} is made of diagonals whose elements reproduces themselves with period 2. This characterizes the Lie algebra $\widehat{sl(2)} \in \widehat{gl(\infty)}$. Hence we have found:

Proposition. *The τ–functions of the KdV hierarchy are given by:*

$$\tau_{KdV}(t) = \langle 0|e^{h(t)}g|0\rangle, \quad \text{with} \quad h(t) = \sum_{k>0} t_{2k-1}h_{2k-1} \text{ and } g \in \widehat{SL(2)}$$

As an application, we can find the KdV soliton solutions. Recall that the KP soliton solutions are obtained for $g = \prod g_i$ with $g_i = 1 + a_i\psi(p_i)\psi^*(q_i)$ where $\psi(p)$ and $\psi^*(q)$ are the fermionic fields. Since we have:

$$[h_n, \psi(z)\psi^*(w)] = (z^n - w^n) : \psi(z)\psi^*(w) : + \frac{z^n - w^n}{z - w}$$

we see that g_i will commute with h_{2k} if $p_i = -q_i$. Hence we have:

Proposition. *The n-soliton τ-functions of the KdV hierarchy are:*

$$\tau_n(t, g) = \langle 0|e^{h(t)} \prod_{i=1}^{n} (1 + a_i\psi(p_i)\psi^*(-p_i))\,|0\rangle \tag{10.37}$$

Explicitly they are equal to:

$$\begin{aligned}
\tau_n(X|p) &= 1 + \sum_{p=1}^{n} \sum_{\substack{I\subset\{1,\cdots,n\}\\ |I|=p}} \prod_{i<j\in I} \left(\frac{p_i - p_j}{p_i + p_j}\right)^2 \cdot \prod_{i\in I} X_i \\
&= 1 + \sum_i X_i + \sum_{i<j} X_iX_j \left(\frac{p_i - p_j}{p_i + p_j}\right)^2 + \cdots, \quad X_i = \frac{a_i}{2p_i} e^{2\xi(p_i,t)}
\end{aligned} \tag{10.38}$$

where $\xi(p, t) = \sum_{n\ odd} p^n t_n$. This can be written in compact form as:

$$\tau_n(X|p) = \det(1 + W), \quad \text{with} \quad W_{ij} = \sqrt{X_i}\frac{\sqrt{4p_ip_j}}{(p_i + p_j)}\sqrt{X_j} \tag{10.39}$$

<u>Proof</u>. We have $e^{h(t)}\psi(p_i)e^{-h(t)} = e^{\xi(t,p_i)}\psi(p_i)$ and $e^{h(t)}\psi^*(-p_i)e^{-h(t)} = e^{\xi(t,p_i)}\psi^*(-p_i)$ so that pushing $e^{h(t)}$ amounts to replacing $a_i \to (2p_i)X_i$. The expression (10.38) is obtained by simply using Wick's theorem.

$$\begin{aligned}
< 0|\prod_i(1 + 2p_iX_i\psi(p_i)\psi^*(-p_i))|0 > &= 1 + \sum_i 2p_iX_i < 0|\psi(p_i)\psi^*(-p_i)|0 > \\
&\quad + \sum_{i<j} 4p_ip_jX_iX_j < 0|\psi(p_i)\psi^*(-p_i)\psi(p_j)\psi^*(-p_j)|0 > + \cdots
\end{aligned}$$

Each of these vacuum expectation values is a determinant (see eq.(8.21) in Chapter [8]). We get:

$$< 0|\prod_i(1 + 2p_iX_i\psi(p_i)\psi^*(-p_i))|0 >= 1 + \sum_i X_i + \sum_{i<j} X_iX_j \det\left(\frac{1}{p_i + p_j}\right) + \cdots \tag{10.40}$$

By the Cauchy formula eq.(8.34) these determinants can be rewritten as in eq.(10.38). On the other hand, eq.(10.39) reduces to eq.(10.40) using the expansion formula:

$$\det(1 + W) = 1 + \sum_i W_{ii} + \sum_{i<j} \det\,[W_{(ij)}]_{2\times 2} + \sum_{i<j<k} \det\,[W_{(ijk)}]_{3\times 3} + \cdots$$

where $[W_{(i_1 i_2\cdots i_p)}]_{p\times p}$ is the $p \times p$ submatrix of W obtained by selecting out the rows and columns of indices $i_1i_2 \cdots i_p$, e.g. $\det\,[W_{(ij)}]_{2\times 2} = W_{ii}W_{jj} - W_{ij}W_{ji}$. ∎

Remark 1. Using the bosonization formula of Chapter [8], one recognizes that the operator

$$V(p) = p\, \psi(p)\psi^*(-p)$$

coincides with the vertex operator defining the level one vertex representation of the affine $\widehat{sl(2)}$ algebra, see Chapter [16]. In the bosonic representation, the group elements g_i can be written:

$$g_i \equiv 1 + a_i\psi(p_i)\psi^*(-p_i) = 1 + \frac{a_i}{p_i}V(p_i)$$

Hence the soliton solutions of KdV are directly related to vacuum expectation values of vertex operators.

The one soliton solution is obtained when $\tau = 1 + X = 1 + \exp\left(2p(x - x_0) + 2p^3 t\right)$. One gets:

$$u(x,t) = \frac{2p^2}{\cosh^2(2p(x - x_0) + 2p^3 t)}$$

It corresponds to a bump of height $2p^2$ propagating with velocity $-p^2$. In sharp contrast with the case of partial linear differential equation where wave packets spread out in time, here the bump preserves its shape for all times. Remark that the center of the bump is located at $X(x,t) = 1$.

Consider now the general n–soliton solution, eq.(10.38). We can analyze its shape asymptotically when $t \to \pm\infty$. Assume that $p_1 > p_2 > \cdots > p_n > 0$. We want to show that we have asymptotically n solitons moving from right to left with velocities $-(p_1)^2, \cdots - (p_n)^2$. Let us assume $t \to -\infty$ and consider what happens around $X_i = 1$, i.e. $x_i(t) = -p_i^2 t$. The values of the other X_j's when $x \sim x_i(t)$ are: $X_j \sim (a_j/\, 2p_j)\, \exp\left(2p_j(p_j^2 - p_i^2)t\right)$. Hence for large negative time, if $p_j^2 < p_i^2$, $X_j(x_i(t), t)$ is very large, while if $p_j^2 > p_i^2$, $X_j(x_i(t), t)$ is very small. So we can split the indices j into two subsets, relative to the index i. One subset $\mathcal{I}_+$ is such that $p_j^2 < p_i^2$, corresponding to X_j very large. The other one $\mathcal{I}_-$ is such that $p_j^2 > p_i^2$ and corresponds to X_j very small. We can evaluate the tau functions eq.(10.38) when $x \sim x_i(t)$. Of course one has to keep the terms containing the maximum number of X_j's, $j \in \mathcal{I}_+$. There are two such terms, yielding:

$$\tau(x,t)|_{x\sim x_i(t)} \sim \prod_{j\in\mathcal{I}_+} X_j \prod_{j<k\in\mathcal{I}_+} \frac{(p_j - p_k)^2}{(p_j + p_k)^2}\Big(1 + X_i \prod_{j\in\mathcal{I}_+} \frac{(p_i - p_j)^2}{(p_i + p_j)^2}\Big)$$

where we have to keep the second term because $X_i \sim 1$. When we compute the KdV field using eq.(10.36) the prefactor disappears because it involves an exponential linear in x, t. Hence the τ–function reduces to that of a 1 soliton τ–function, but with parameter

$$a_i \to a_i^{\mathrm{in}} = a_i \prod_{j\in\mathcal{I}_+} \frac{(p_i - p_j)^2}{(p_i + p_j)^2}$$

So around $x = x_i(t)$ the KdV field $u(x,t)$ looks like a one–soliton field, and the n–soliton solution appears as n widely separated solitons travelling to the left with velocities $-p_i^2$. At some finite times these solitons will collide and the above analysis becomes invalid. However for $t \to +\infty$ the solitons separate again, and we get the same picture with the roles of $\mathcal{I}_+$ and $\mathcal{I}_-$ reversed, so that:

$$a_i^{\mathrm{out}} = a_i \prod_{j\in\mathcal{I}_-} \frac{(p_i - p_j)^2}{(p_i + p_j)^2}$$

It follows that the interaction introduces a delay:

$$\delta_i = \log \frac{a_i^{\mathrm{out}}}{a_i^{\mathrm{in}}} = \log \prod_{j<i} \frac{(p_i - p_j)^2}{(p_i + p_j)^2} - \log \prod_{j>i} \frac{(p_i - p_j)^2}{(p_i + p_j)^2}$$

This formula shows that the delay is the sum of delays introduced by pairwise interactions. Moreover each individual soliton keeps its form and its velocity after interactions.

Knowing the picture of the n–soliton solutions in the regime $t \to \pm\infty$ as a superposition of n 1–soliton solutions, we can compute easily the conserved quantities. Since the quantities are conserved one can evaluate them at $t \to -\infty$. From eq.(10.28) we can write $S(x,t,k) = 1 + \sum_i \partial_x D(k) \log \tau_i(x,t)$ where τ_i is the 1–soliton τ–function with parameters $p_i, a_i^{\rm in}$, since for any given value of x only one soliton contributes. We have $\tau_i = 1 + a_i^{\rm in} \exp(2\xi(t,p_i))$ so that

$$\partial_{t_{2j-1}} \log \tau_i = 2p_i^{2j-1} \frac{a_i^{\rm in} e^{2\xi(t,p_i)}}{1 + a_i^{\rm in} e^{2\xi(t,p_i)}}$$

yielding $H(k) = \int dx S(x,t,k) = L + 2\sum_{j>1} k^{-2j} \sum_i p_i^{2j-1}$. So the complete set of conserved quantities is provided by:

$$H_j = \sum_{i=1}^{n} p_i^{2j-1}$$

Because the H_j are conserved for any j it follows that, in the scattering process of solitons, only delays and permutations of the p_i can occur.

10.5 Algebro–geometric solutions.

We wish to apply the analytical methods of Chapter [4] to construct solutions of the KdV equation. As explained in that Chapter one way to get a Lax matrix compatible with the equations of the KdV hierarchy is to seek for stationary solutions with respect to some higher time t_{2j-1}. Then the zero–curvature condition $\partial_i A_j - \partial_j A_i + [A_i, A_j] = 0$ becomes a Lax equation since the stationarity condition with respect to time j means $\partial_j A_i = 0$, for all i. The Lax matrix is A_j and its associated spectral curve is independent of all times t_i. A very simple example of this situation occurs when u is stationary with respect to $t_3 = t$. In this case the Lax matrix is A_t in eq.(10.2). The associated spectral curve is given by:

$$\Gamma: \quad \det(A_t - \mu) = \mu^2 - \frac{1}{4}\lambda^3 + \frac{1}{4}(u'' + 3u^2)\lambda + \frac{1}{16}(2uu'' - u'^2 + 4u^3) = 0$$

The zero–curvature condition becomes the Lax equation $\partial_x A_t = [A_t, A_x]$ and reduces to the stationary KdV equation $6uu' + u''' = 0$. Integrating one gets $3u^2 + u'' = C_1$ and $2u^3 + u'^2 = 2C_1 u + C_2$ for some constants C_1, C_2 so the spectral curve reads $\mu^2 = \lambda^3/4 - C_1\lambda/4 + C_2/16$, and is independent of x as it should be. This is the equation of a genus 1 curve, and u is given by an elliptic integral.

More interesting solutions will be obtained by assuming u to be independent of some higher time t_{2j-1}. Let us compute the matrices $A_{t_{2j-1}}$. Since $L^{\frac{2j-1}{2}}$ is anti self–adjoint, for either one of the two solutions ψ and ψ^* of the KdV hierarchy we have:

$$\partial_{t_{2j-1}} \begin{pmatrix} \psi \\ \partial_x \psi \end{pmatrix} = \begin{pmatrix} (L^{\frac{2j-1}{2}})_+ \psi \\ \partial_x (L^{\frac{2j-1}{2}})_+ \psi \end{pmatrix} = -A_{t_{2j-1}} \begin{pmatrix} \psi \\ \partial_x \psi \end{pmatrix}$$

Using the identity (10.24) and $\partial_x^2 \psi = (\lambda - u)\psi$ one gets:

$$A_{t_{2j-1}} = -\sum_{i=0}^{j-1} \lambda^{j-i-1} \begin{pmatrix} -\frac{1}{2}S'_{2i} & S_{2i} \\ (\lambda - u)S_{2i} - \frac{1}{2}S''_{2i} & \frac{1}{2}S'_{2i} \end{pmatrix}$$

Notice that $A_{t_{2j-1}}$ only depends on λ in agreement with the fact that ψ and ψ^* play the same role. In particular for $j = 1$ one finds A_x and for $j = 2$ one finds A_t. Writing $(L^{\frac{2j-1}{2}})_+ = (L^{\frac{2j-1}{2}}) - (L^{\frac{2j-1}{2}})_-$, we see that:

$$A_{t_{2j-1}} \begin{pmatrix} \psi \\ \partial_x \psi \end{pmatrix} = \lambda^{\frac{2j-1}{2}} \begin{pmatrix} \psi \\ \partial_x \psi \end{pmatrix} + O(1)$$

hence we have identified, asymptotically for $\lambda \to \infty$, the eigenvectors of $A_{t_{2j-1}}$ and the eigenvalues:

$$\mu = \lambda^{\frac{2j-1}{2}} + O(1) \tag{10.41}$$

The matrix $A_{t_{2j-1}}$ is traceless 2×2, so its associated spectral curve is an hyperelliptic curve of genus $(j-1)$ of the form $\mu^2 = R(\lambda)$ with $R(\lambda) = \prod_{j=0}^{2j-1}(\lambda - \lambda_j) = \det A_{t_{2j-1}}$. This curve is not a general hyperelliptic Riemann surface of genus $(j-1)$. In fact since $\mu = \sqrt{R(\lambda)}$ is an eigenvalue of $A_{t_{2j-1}}(\lambda)$, it has to be of the specific form eq.(10.41), showing that $R(\lambda)$ is very special (we have $R(\lambda) = \lambda^{2j-1} + C_1\lambda^{j-1} + C_2\lambda^{j-2} + \cdots$). To overcome this peculiarity, we notice that the stationarity condition can be relaxed by imposing the condition:

$$\sum_j c_j \partial_{t_{2j-1}} u = 0$$

for some constant coefficients c_j. Then the corresponding Lax matrix becomes $\sum_j c_j A_{t_{2j-1}}$. For any time t_{2i-1} the zero curvature condition implies the Lax equation (because the $A_{t_{2j-1}}$ only depend on u):

$$\partial_{t_{2i-1}}\Big(\sum_j c_j A_{t_{2j-1}}\Big) = \left[\Big(\sum_j c_j A_{t_{2j-1}}\Big), A_{t_{2i-1}}\right]$$

In the following we shall consider the hyperelliptic curve Γ constructed from such a Lax matrix. It is of the generic form:

$$\Gamma: \quad \mu^2 = R(\lambda) = \prod_{i=1}^{2g+1}(\lambda - \lambda_i) \tag{10.42}$$

We know that there exists a section Ψ of the eigenvector bundle on Γ, obeying $\partial_{t_{2i-1}}\Psi + A_{t_{2i-1}}\Psi = 0$ for all i. We want to construct it. For $i = 1$ we get:

$$\Psi = \begin{pmatrix} \psi \\ \partial_x\psi \end{pmatrix}, \quad (\partial_x^2 + u)\psi = \lambda\psi$$

so that it is enough to consider ψ. A consequence of eq.(10.41) is that ψ has the asymptotic behaviour at infinity $\psi = e^{\xi(t,k)}(1 + O(k^{-1}))$, where $k = \sqrt{\lambda}$.

We know (see Chapter [4]) that Ψ has $g+N-1 = g+1$ poles on Γ (N is the size of the Lax matrix). Here one of the poles is at ∞ because $\partial\psi \sim k\psi$ and $k = \sqrt{\lambda}$ is the local parameter at ∞. Hence we must require that ψ has g poles $(\nu_1, \cdots, \nu_g)$ at finite distance. Recall that these poles are independent of all times. With these data we construct the Baker–Akhiezer function on Γ which is the unique function with the following analytical properties:

- It has an essential singularity at the point P at infinity:

$$\psi(t,k) = e^{\xi(t,k)}(1 + \frac{\alpha(t)}{k} + O(1/k^2)) \quad \text{with} \quad k = \sqrt{\lambda}, \;\; \xi(t,k) = \sum_{i\geq 1} k^{2i-1} t_{2i-1} \tag{10.43}$$

- It has g simple poles, independent of all times. The divisor of these poles is $D = (\nu_1, \cdots, \nu_g)$.

This Baker–Akhiezer function solves the KdV hierarchy equations as the following two propositions show.

Proposition. *There exists a function $u(x,t)$ such that*

$$(\partial_x^2 + u)\;\psi = \lambda\psi \tag{10.44}$$

Proof. Consider on Γ the function $\partial_x^2\psi - \lambda\psi$. To define this object as a function on the curve, λ is assumed to be a meromorphic function on Γ. Remark that λ has only a double pole at ∞ and such a function exists only if Γ is hyperelliptic. So $\partial_x^2\psi - \lambda\psi$ has the same analytical properties as ψ itself at finite distance on Γ. At infinity we have by eq.(10.43):

$$\partial_x^2\psi - \lambda\psi = e^{\xi(t,k)}(2\partial_x\alpha + O(1/k))$$

hence it is a Baker–Akhiezer function, but with a normalization $2\partial_x\alpha$ instead of 1 at infinity. By the uniqueness theorem we have:

$$\partial_x^2\psi - \lambda\psi = -u(t)\psi \quad u(t) = -2\partial_x\alpha \tag{10.45}$$

■

Having found the potential u, we construct the differential operator $L = \partial^2 + u$ and show that the Baker–Akhiezer function ψ obeys all the equations of the associated KdV hierarchy.

Proposition. *The evolution of $\psi(t,\lambda)$ is given by:*

$$\partial_{t_{2i-1}}\psi = (L^{\frac{2i-1}{2}})_+\psi$$

where $L = \partial^2 + u$ is the KdV operator constructed above.

Proof. Consider the function $\partial_{t_{2i-1}}\psi - (L^{\frac{2i-1}{2}})_+\psi$. It has the same analytical properties as ψ at finite distance on Γ. At infinity we have $\partial_{t_{2i-1}}\psi = k^{2i-1}\psi + e^{\xi}O(1/k)$ and $(L^{\frac{2i-1}{2}})_+\psi = L^{\frac{2i-1}{2}}\psi - (L^{\frac{2i-1}{2}})_-\psi = k^{2i-1}\psi + e^{\xi}O(1/k)$, where we have used $L\psi = k^2\psi$. Summarizing we get:

$$\partial_{t_{2i-1}}\psi - (L^{\frac{2i-1}{2}})_+\psi = e^{\xi(t,k)}\,O(k^{-1}) \quad k \to \infty$$

By unicity, this Baker–Akhiezer function which vanishes at ∞ vanishes identically. ■

The Baker–Akhiezer function $\psi(t,k)$ can be written explicitly in terms of θ–functions, see Chapter [4]. Let $\Omega^{(2j-1)}$ be the unique normalized second kind differential (all the a–periods vanish) with a pole of order $2j$ at infinity, such that:

$$\Omega^{(2j-1)} = d\left(k^{2j-1} + \text{regular}\right), \quad \text{for} \quad k \to \infty$$

Let $U_k^{(2j-1)}$ be its b-periods:

$$U_k^{(2j-1)} = \frac{1}{2i\pi}\oint_{b_k}\Omega^{(2j-1)}$$

In terms of these data we have:

Proposition. *The Baker-Akhiezer function associated to the divisor D can be expressed as:*

$$\psi(t,P) = e^{\int_\infty^P \sum_j t_{2j-1}\Omega^{(2j-1)}}\,\frac{\theta(\mathcal{A}(P) + \sum_j t_{2j-1}U^{(2j-1)} - \zeta)\,\theta(\zeta)}{\theta(\mathcal{A}(P) - \zeta)\,\theta(\sum_j t_{2j-1}U^{(2j-1)} - \zeta)} \tag{10.46}$$

where $\mathcal{A}(P)$ is the Abel map on Γ with base point ∞, and $\zeta = \mathcal{A}(D) + \mathcal{K}$ with $\mathcal{K}$ the Riemann's constant vector. The KdV field u is then given by the Its–Matveev formula:

$$u(x,t) = 2\partial_x^2 \log\theta\left(\sum_j t_{2j-1}U^{(2j-1)} - \zeta\right) + \text{const.} \tag{10.47}$$

Proof. In eq.(10.46) the integral $\int_\infty^P$ has to be understood in the following sense: for k in a vicinity of ∞ one *defines* $\int_\infty^P \Omega^{(2j-1)}$ as the unique primitive of $\Omega^{(2j-1)}$ which behaves as $k^{2j-1} + O(1/k)$ (no constant term). Of course when this is analytically continued on the Riemann surface, b–periods will appear. However they will cancel out in eq.(10.46) due to the monodromy properties of θ–functions, leaving us with a well–defined normalized Baker-Akhiezer function. The formula for the KdV field is found by using:

$$\lambda - u = (\partial_x^2 \log \psi) + (\partial_x \log \psi)^2$$

Setting $\Omega^{(1)}(k) = k + \frac{\beta}{k} + O(k^{-2})$, where β does not depend on times, we have:

$$\partial_x \log \psi = k + \partial_x \log \theta \left(\mathcal{A}(P) + \sum_j t_{2j-1} U^{(2j-1)} - \zeta \right) - \partial_x \log \theta \left(\sum_j t_{2j-1} U^{(2j-1)} - \zeta \right) + \frac{\beta}{k} + O(k^{-2})$$

We evaluate this when $k \to \infty$. Using Riemann's bilinear identities, we can expand the Abel map $\mathcal{A}(P)$ around ∞ see eq.(4.49) in Chapter [4], and we have:

$$\theta \left(\mathcal{A}(P) + \sum_j t_{2j-1} U^{(2j-1)} - \zeta \right) = \theta \left(\sum_j (t_{2j-1} - \frac{k^{2j-1}}{2j-1}) U^{(2j-1)} - \zeta \right)$$

Hence we obtain:

$$\partial_x \log \psi = k - \frac{1}{k} \partial_x^2 \log \theta \left(\sum_j t_{2j-1} U^{(2j-1)} - \zeta \right) + \frac{\beta}{k} + O(\frac{1}{k^2})$$

It follows that $\partial_x^2 \log \psi = O(1/k)$, and $k^2 - u = k^2 - 2\partial_x^2 \log \theta + 2\beta + O(1/k)$ proving the result. ∎

Since the curve Γ is hyperelliptic one can easily express the solution ψ in terms of its zero divisor. It depends on all times t_{2j-1} but we shall first consider the x dependence. Let Z be the divisor of the zeroes of the Baker–Akhiezer function ψ. It is of degree g:

$$Z \equiv \left\{ \left(z_1(x), \sqrt{R(z_1(x))} \right), \cdots, \left(z_g(x), \sqrt{R(z_g(x))} \right) \right\}$$

In this formula one assumes that one sheet of the curve has been chosen for each point of the divisor. Let us define the polynomial of λ:

$$B(\lambda, x) = \prod_i (\lambda - z_i(x))$$

In terms of these data, we have:

Proposition. *The Baker-Akhiezer function is equal to:*

$$\frac{\psi(x, \lambda)}{\psi(x_0, \lambda)} = \sqrt{\frac{B(\lambda, x)}{B(\lambda, x_0)}} \exp \left(\int_{x_0}^x \frac{\sqrt{R(\lambda)}}{B(\lambda, x)} dx \right) \tag{10.48}$$

The KdV potential u is expressed as:

$$u = -2 \sum_{i=1}^g z_i(x) + \sum_{i=1}^{2g+1} \lambda_j \tag{10.49}$$

Proof. Let us first find the equations governing the x dependence of the divisor Z. Consider the function $\partial_x \psi / \psi$. It is a meromorphic function on Γ, it has poles at the points $(z_i(x), \sqrt{R(z_i(x))})$ and behaves like $k + O(1/k)$ at infinity. Hence we can write

$$\frac{\partial_x \psi}{\psi} = \frac{\sqrt{R(\lambda)} + Q(\lambda, x)}{\prod_{i=1}^g (\lambda - z_i(x))} \tag{10.50}$$

where Q is a polynomial of degree $g-1$ in λ. We determine Q by requiring that $\frac{\partial_x \psi}{\psi}$ has a pole above $\lambda = z_i(x)$ on the sheet $\sqrt{R(z_i)}$ and not on $-\sqrt{R(z_i)}$. Thus $Q(z_i(x), x) = \sqrt{R(z_i(x))}$ and

$$Q(\lambda, x) = \sum_i \sqrt{R(z_i(x))} \frac{\prod_{j \neq i}(\lambda - z_j(x))}{\prod_{j \neq i}(z_i(x) - z_j(x))}$$

On the other hand, in the vicinity of $z_i(x)$, we have:

$$\frac{\partial_x \psi}{\psi} = -\frac{\partial_x z_i(x)}{\lambda - z_i(x)} + O(1) \tag{10.51}$$

because around the zero we have $\psi(x, \lambda) \sim (\lambda - z_i)\widetilde{\psi}$. Comparing eq.(10.50) and eq.(10.51), we get the equations of motion for the divisor Z:

$$\partial_x z_i(x) = -2 \frac{\sqrt{R(z_i(x))}}{\prod_{j \neq i}(z_i(x) - z_j(x))} \tag{10.52}$$

One can now reconstruct the Baker-Akhiezer function itself. Indeed, inserting eq.(10.52) into (10.50) we get:

$$\frac{\partial_x \psi}{\psi} = \frac{\sqrt{R(\lambda)}}{B(\lambda, x)} - \frac{1}{2} \sum_i \frac{1}{\lambda - z_i(x)} \partial_x z_i(x)$$

Integrating this formula from x_0 to x gives eq.(10.48). To compute the potential $u(x)$, we insert eq.(10.48) into $(\partial_x^2 + u)\psi = \lambda\psi$. We get the polynomial identity

$$R = -\frac{1}{2} BB'' + \frac{1}{4} B'^2 + (-u + \lambda) B^2$$

where $' = \partial_x$. Comparing the terms λ^{2g} we obtain expression (10.49). ■

One can find a generalization of eq.(10.52) for any flow of the KdV hierarchy.

Proposition. *On the coordinates z_i of the divisor Z the equations of motion with respect to the time t_{2j-1} read:*

$$\partial_{t_{2j-1}} z_i = -2 \frac{P_j(z_i)\sqrt{R(z_i)}}{\prod_{l \neq i}(z_i - z_l)}, \quad P_j(\lambda) = \left(\frac{\prod_{i=1}^{g}(1 - z_i \lambda^{-1})}{\sqrt{\prod_{i=1}^{2g+1}(1 - \lambda_i \lambda^{-1})}} \lambda^{j-1} \right)_+ \tag{10.53}$$

where the $(\ )_+$ means taking the polynomial part in the expansion at $\lambda = \infty$. This can be used to give a direct proof of the linearization of the flows on the Jacobian of the hyperelliptic curve Γ.

Proof. The only difference with the previous discussion is that in the derivation of eq.(10.50), the behavior at ∞ is replaced by $\partial_{t_{2j-1}} \psi / \psi = k^{2j-1} + O(1/k)$, so that we have:

$$\frac{\partial_{t_{2j-1}} \psi}{\psi} = \frac{P_j(\lambda)\sqrt{R(\lambda)} + Q_j(\lambda)}{B(\lambda)}$$

where $P_j(\lambda)$ is a polynomial in λ of degree $(j-1)$, $Q_j(\lambda)$ is a polynomial in λ of degree $(g-1)$. The polynomial $P_j(\lambda) = \lambda^{j-1} + \cdots$ is uniquely determined by imposing the asymptotics at ∞ which gives $(j-1)$ linear conditions. The solution is given by eq.(10.53). The polynomial Q_j is determined as above, yielding $Q_j(z_i) = P_j(z_i)\sqrt{R(z_i)}$. One gets the equation of motion for the divisor Z:

$$\partial_{t_{2j-1}} z_i = -2 \frac{P_j(z_i)\sqrt{R(z_i)}}{\prod_{l \neq i}(z_i - z_l)}$$

To exhibit the linearization of the flows, it is sufficient to consider the time evolutions of the Abel sums

$$\frac{\partial}{\partial_{t_{2j-1}}}\sum_i \int^{z_i} \frac{\lambda^k}{\sqrt{R(\lambda)}} d\lambda = \sum_i \frac{z_i^k}{\sqrt{R(z_i)}} \partial_{t_{2j-1}} z_i = -2\sum_i \frac{z_i^k P_j(z_i)}{\prod_{l\neq i}(z_i - z_l)}$$

The right hand side is equal to the integral:

$$\frac{1}{2i\pi}\int_\gamma d\lambda \frac{\lambda^k P_j(\lambda)}{\prod_l(\lambda - z_l)}$$

where γ is a loop surrounding all the z_i. We deform the contour as a loop around ∞ so that:

$$\sum_i \frac{z_i^k P_j(z_i)}{\prod_{l\neq i}(z_i - z_l)} = -\mathrm{Res}_{\lambda=\infty} \frac{P_j(\lambda)}{B(\lambda)}\lambda^k d\lambda$$

To compute this residue at $\lambda = \infty$, we write:

$$2\mathrm{Res}_{\lambda=\infty} \frac{P_j(\lambda)}{B(\lambda)}\lambda^k d\lambda = \mathrm{Res}_\infty \frac{P_j(\lambda)\sqrt{R(\lambda)}}{B(\lambda)} \frac{\lambda^k d\lambda}{\sqrt{R(\lambda)}}$$

where the right–hand side is a residue computed on the curve Γ. The factor 2 appears since ∞ is a branch point of the covering $k \to \lambda = k^2$ around that point. The first factor behaves as $k^{2j-1} + O(1/k)$ by construction. So we have:

$$\frac{\partial}{\partial_{t_{2j-1}}}\sum_i \int^{z_i} \omega_k = 2\mathrm{Res}_{\lambda=\infty}\, (z^{-2j+1} + O(z))\omega_k$$

where $z = 1/k$ is the local parameter at ∞, and $\omega_k = \lambda^k d\lambda/\sqrt{R(\lambda)}$ is an unnormalized abelian differential. This shows that the flow linearizes on the Jacobian because the right–hand side is a constant. Since this equation is linear in ω_k it remains true for normalized abelian differentials. One can then use Riemann's bilinear identities to evaluate the right–hand side:

$$\frac{\partial}{\partial_{t_{2j-1}}}\mathcal{A}(Z) = \mathrm{Res}_\infty\, (z^{-2j+1} + O(z))\omega_k = \frac{1}{2i\pi}\oint_{b_k} \Omega^{(j)}$$

so the slope is exactly as expected. ■

Remark. The equation of motion of the divisor Z can be recast in compact form by introducing the generating functional of time derivatives eq.(10.20). For a function $f(\lambda) = \sum_{n=0}^\infty f_n \lambda^{-n}$ we have:

$$\sum_{j=1}^\infty \lambda'^{-j}(f(\lambda)\lambda^{j-1})_+ = \frac{f(\lambda')}{\lambda' - \lambda}$$

so that we get:

$$D(k)z_i = -2\frac{k}{k^2 - z_i}\frac{B(k^2)}{\sqrt{R(k^2)}}\frac{\sqrt{R(z_i)}}{\prod_{l\neq i}(z_i - z_l)} \tag{10.54}$$

Finally we show that eq.(10.49) can be generalized to a whole set of so–called trace identities.

Proposition. *The generating function $S(k)$ of the local quantities S_{2n} has a simple expression in terms of the divisor Z.*

$$S(k) = k\frac{B(\lambda, x)}{\sqrt{R(\lambda)}} = \frac{\prod_{i=1}^g (1 - z_i k^{-2})}{\sqrt{\prod_{j=1}^{2g+1}(1 - \lambda_i k^{-2})}}$$

Proof. Recall that we have defined $S(k) = \psi\psi^*$ where ψ and ψ^* are normalized such that their Wronskian is equal to $2k$. We can evaluate $S(k)$ using eq.(10.48) for ψ and a similar equation for ψ^* with the sign of the exponential reversed, provided that we normalize them to have the correct Wronskian. Using eq.(10.48) one gets $W = 2\psi(x_0,\lambda)\psi^*(x_0,\lambda)\sqrt{R(\lambda)}/B(\lambda,x_0)$, and the the expression of $S(k)$ follows. ∎

Taking the logarithm we get:

$$\sum_{n=0}^{\infty} k^{-2n}\frac{1}{n}\left(\sum_i z_i^n\right) = \frac{1}{2}\sum_{n=0}^{\infty} k^{-2n}\frac{1}{n}\left(\sum_i \lambda_i^n\right) - \log S(k)$$

hence we find:

$$\sum_i z_i = \frac{1}{2}\sum \lambda_i - \frac{u}{2}, \quad \sum_i z_i^2 = \frac{1}{2}\sum \lambda_i^2 - \frac{1}{4}u'' - \frac{1}{2}u^2$$

where the first equation is eq.(10.49) and we see that all the symmetric functions of the z_i have a local expression in terms of the potential u.

10.6 Finite zone solutions.

The Its–Matveev formula eq.(10.47) shows that, as a function of the *real* variable x, the potential $u(x,t)$ is almost periodic. Namely, the argument of the θ–function is a straight line $Y(x) = U^{(1)}x + Y_0$ which wraps densely around the Jacobian torus, and for ℓ sufficiently large $Y(x+\ell)$ returns arbitrarily close to $Y(x)$. This means that $Y(x+\ell) \sim Y(x) + n + \mathcal{B}m$ where $n, m \in \mathbb{Z}^g$ and $\mathcal{B}$ is the matrix of b periods. The effect of the translation by $n + \mathcal{B}m$ does not affect the potential because of the second order derivative in front of the logarithm. Hence $u(x+\ell) \sim u(x)$. One can choose the moduli of the curve Γ so that the potential is exactly periodic. This amounts to the condition:

$$\ell U^{(1)} = n + \mathcal{B}m \tag{10.55}$$

This gives $2g$ real conditions on the $2g+1$ complex parameters λ_i. If however the parameters λ_i are real (as will be the case in the following), these conditions quantize all the moduli of Γ (up to a translation $\lambda_i \to \lambda_i + \alpha$ which does not change the periods $U^{(1)}$). Under these conditions the potential u given by eq.(10.49) is periodic, and eq.(10.46) shows that the Baker-Akhiezer function $\psi(x)$ is a Bloch wave, i.e. $\psi(x+\ell) = e^{\mathcal{P}(\lambda)}\psi(x)$. Up to now, the potential u is complex and we analyze the conditions on λ_i and the divisor $Z(x)$ ensuring that u is real.

Consider the Schroedinger equation $\psi'' + u\psi = \lambda\psi$ with a real potential $u(x)$ with period ℓ. The space of solutions is spanned by the two solutions $y_1(x,\lambda)$ and $y_2(x,\lambda)$ with initial conditions:

$$y_1(0,\lambda) = 1,\ y_1'(0,\lambda) = 0, \quad y_2(0,\lambda) = 0,\ y_2'(0,\lambda) = 1$$

Because the initial conditions are independent of λ, y_1 and y_2 are entire functions of λ. Since u is periodic, $y_1(x+\ell,\lambda)$ and $y_2(x+\ell,\lambda)$ are two other solutions so we can write:

$$\begin{pmatrix} y_1(x+\ell,\lambda) \\ y_2(x+\ell,\lambda) \end{pmatrix} = T(\lambda)\begin{pmatrix} y_1(x,\lambda) \\ y_2(x,\lambda) \end{pmatrix}$$

where $T(\lambda)$ is a 2×2 matrix, the monodromy matrix. Because the Wronskian $W(y_1,y_2) = 1$ we have $\det T(\lambda) = 1$. Notice that for real λ, y_i are real so that $T(\lambda)$ is real, and in general $\overline{T(\lambda)} = T(\overline{\lambda})$. The Bloch waves are the two solutions which diagonalize the monodromy matrix. Denote $e^{\pm\mathcal{P}(\lambda)}$ its eigenvalues (with product 1). The two Bloch waves are such that $\psi_\pm(x+\ell) = e^{\pm\mathcal{P}(\lambda)}\psi_\pm(x)$ and we choose to normalize them by $\psi_\pm(0) = 1$. The quantity $\mathcal{P}(\lambda)$ is called the quasi–momentum. Expanding $\psi_\pm(x,\lambda) = y_1(x,\lambda) + \beta y_2(x,\lambda)$ and writing the Bloch conditions on $\psi_\pm$ and $\psi_\pm'$ gives:

$$\psi_\pm(x,\lambda) = y_1(x,\lambda) + \frac{e^{\pm\mathcal{P}(\lambda)} - y_1(\ell,\lambda)}{y_2(\ell,\lambda)}y_2(x,\lambda) \tag{10.56}$$

where $\mathcal{P}(\lambda)$ is obtained by solving the equation:

$$e^{\mathcal{P}(\lambda)} + e^{-\mathcal{P}(\lambda)} = y_1(\ell, \lambda) + y_2'(\ell, \lambda) \equiv t(\lambda) \tag{10.57}$$

Note that $t(\lambda) = \mathrm{Tr}\, T(\lambda)$ is an entire function of λ hence $e^{\pm\mathcal{P}(\lambda)}$ has no pole or zero at finite distance. It follows from eq.(10.56) that the poles in λ of $\psi_\pm(x, \lambda)$ are located where $y_2(\ell, \lambda) = 0$. One can discuss the nature of the solutions of eq.(10.57) when λ is real, so that $t(\lambda)$ is also real. When $\Delta(\lambda) \equiv t^2(\lambda) - 4$ is negative the quasi momentum $\mathcal{P}(\lambda)$ is pure imaginary. In this case $\psi(x)$ has an oscillatory behaviour in x at large scale corresponding to propagation of waves. This defines allowed zones in the spectrum. When $\Delta(\lambda)$ is positive, $\psi(x)$ has an exponential behaviour at large scale, so waves cannot propagate. This defines forbidden zones in the spectrum. Finally when $\Delta(\lambda) = 0$ we have $t(\lambda) = \pm 2$ and $e^{\mathcal{P}(\lambda)} = \pm 1$ (same sign). This means that $\psi(x + \ell) = \pm\psi(x)$ so that ψ is periodic or antiperiodic. This means that the periodic and antiperiodic levels are the boundaries of allowed and forbiden zones. When $\lambda \to \infty$ one can as a first approximation neglect the potential u. Then for $\lambda \to +\infty$ we have $\psi \sim e^{\pm\sqrt{\lambda}x}$ (forbidden zone) and $t(\lambda) \sim 2\cosh \sqrt{\lambda}\ell$. For $\lambda \to -\infty$ we have similarly $\psi \sim e^{\pm i\sqrt{-\lambda}x}$ (allowed zone) and $t(\lambda) \sim 2\cos \sqrt{-\lambda}\ell$. In fact when $u \neq 0$ one can show that for generic potentials there are an infinite number of forbidden zones of exponentially small sizes extending in the region $\lambda \to -\infty$. Finite–zone potentials are such that this phenomenon does not occur, i.e. there is a finite number of forbidden zones.

The classical theory of Sturm–Liouville equations gives a rather detailed information on the poles of the Baker-Akhiezer function and the boundaries of the zones. We recall the main facts. Consider a differential equation $y'' + uy = \lambda y$ with a real periodic potential u. We have:

$$\int_0^\ell f\partial^2 g = \int_0^\ell g\partial^2 f + [fg' - f'g]_0^\ell$$

so that we get a self–adjoint problem when the boundary conditions are such that the term $[fg' - f'g]_0^\ell$ vanishes. In this case the spectrum is real.

Proposition. *The poles of the Bloch wave ψ and the periodic and antiperiodic levels are all real. The periodic and antiperiodic levels form a sequence $\beta_1 > \beta_2 \geq \beta_3 > \beta_4 \geq \beta_5 > \cdots$ such that β_1 is a periodic level, β_2, β_3 are antiperiodic levels, β_4, β_5 are periodic, and so on. The allowed zones are the intervals $[\beta_{2j}, \beta_{2j-1}]$ and the forbidden zones are the intervals $[\beta_1, \infty]$ and the $[\beta_{2j+1}, \beta_{2j}]$. There is at least one pole of ψ in each forbidden zone, except $[\beta_1, +\infty]$.*

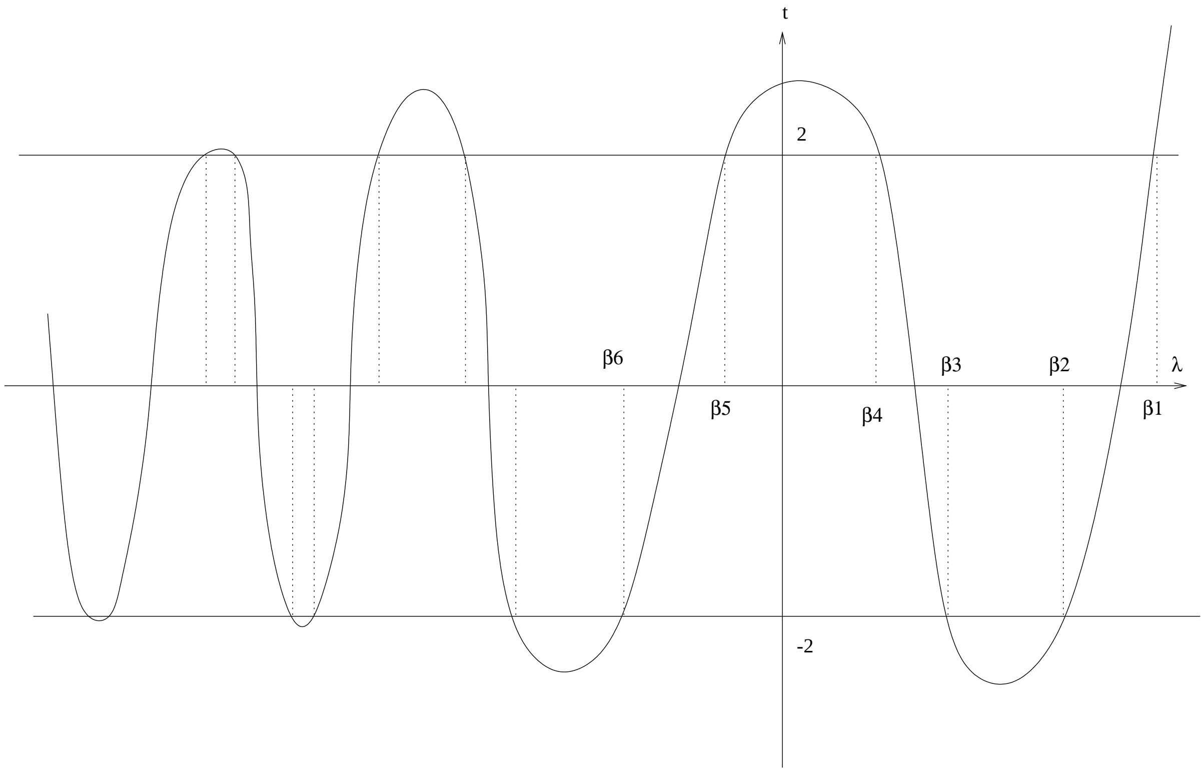
Proof. The poles of ψ are such that $y_2(\ell, \lambda) = 0$, so that y_2 is solution of the Sturm–Liouville problem with boundary conditions $y(0) = y(\ell) = 0$ which is self–adjoint. They are thus real. Similarly the periodic and antiperiodic levels correspond to the boundary conditions (valid in the case of a periodic potential) $y(\ell) = \pm y(0)$ and $y'(\ell) = \pm y'(0)$ which also lead to a self–adjoint problem. These levels are thus real. This shows that all the roots of $\Delta(\lambda)$ are real. We now show that $\partial_\lambda t(\lambda)$ has a definite sign in the allowed zones. We have $\partial_\lambda t(\lambda) = \partial_\lambda y_1(\ell, \lambda) + \partial_\lambda y_2'(\ell, \lambda)$. Since $z = \partial_\lambda y_j$ obeys the differential equation $z'' + (u - \lambda)z = y$ with initial conditions $z(0) = z'(0) = 0$ one has:

$$\begin{aligned}
\partial_\lambda y_j(x, \lambda) &= \int_0^x (y_1(x)y_2(\xi) - y_1(\xi)y_2(x))y_j(\xi)d\xi, \quad j = 1, 2 \\
\partial_\lambda y_j'(x, \lambda) &= \int_0^x (y_1'(x)y_2(\xi) - y_1(\xi)y_2'(x))y_j(\xi)d\xi
\end{aligned}$$

so that:

$$\partial_\lambda t(\lambda) = \int_0^\ell (A(\lambda)y_2^2(\xi) + B(\lambda)y_1(\xi)y_2(\xi) + C(\lambda)y_1^2(\xi))d\xi$$

where $A(\lambda) = y_1'(\ell, \lambda)$, $B(\lambda) = y_1(\ell, \lambda) - y_2'(\ell, \lambda)$, $C(\lambda) = -y_2(\ell, \lambda)$. The quadratic form appearing in the integrand has discriminant $B^2 - 4AC = (y_1 - y_2')^2 + 4y_2y_1' = t^2(\lambda) - 4 = \Delta(\lambda) < 0$ so it is of fixed sign in the whole domain of integration. It follows that $\partial_\lambda t(\lambda) \neq 0$ in an allowed zone, and that the sign of $\partial_\lambda t(\lambda)$ is the same as the sign of $C(\lambda) = -y_2(\ell, \lambda)$. In particular $y_2(\ell, \lambda)$ cannot vanish in an allowed zone. We see that $t(\lambda)$ either crosses the lines $t = \pm 2$ or is tangent to them, but cannot have an extremum in the region $|t(\lambda)| < 2$. From this follows that the periodic and antiperiodic levels are distributed as indicated in the Proposition. Obviously the sign of $\partial_\lambda t(\lambda)$ changes when one goes from one allowed zone to the next one, so that $y_2(\ell, \lambda)$ has at least one zero in each forbidden zone $[\beta_{2j+1}, \beta_{2j}]$ with $\beta_{2j+1} \neq \beta_{2j}$ and $j \geq 1$. ∎



If the periodic potential u is produced by the algebro–geometric construction above, the two Bloch waves $\psi_\pm$ are the two values of the Baker-Akhiezer function $\psi(x, (\lambda, \mu))$ at the two points $(\lambda, \pm\mu)$ above λ. At a branch point, we have $\psi_+ = \psi_-$ so that the quasi momentum satisfies $e^{\mathcal{P}(\lambda)} = \pm 1$ i.e. the branch points are zone boundaries and therefore real. Since the region $\lambda \to \infty$ is a forbidden zone, and corresponds to $R(\lambda) > 0$ we see, following sign changes, that allowed zones correspond to $R(\lambda) < 0$ and forbidden zones to $R(\lambda) > 0$. In particular the branch points form a sequence $\lambda_1 > \lambda_2 > \lambda_3 > \cdots > \lambda_{2g+1}$ with $\lambda_1 = \beta_1$. There may be degenerate forbidden zones with $\beta_{2j+1} = \beta_{2j}$ which do not correspond to branch points of the spectral curve. To compare eq.(10.48) with this discussion, we set $x_0 = 0$, and we see that for $x = 0$ the divisor $Z(x)$ of zeroes of ψ coincides with the divisor of poles. Hence the elements of $Z(0)$ are all real, and lie in forbidden zones. We know that the divisor $Z(0)$ is of degree g if Γ is of genus g, and we put one pole of ψ in each forbidden zone $[\lambda_{2j+1}, \lambda_{2j}]$, for $j = 1, \cdots, g$, and no pole in $[\lambda_1, +\infty]$ to get a periodic motion. The equation of motion (10.52) are thus regular and show that all points of $Z(x)$ stay real for all x, since $R(z_i(x)) > 0$ and remain in forbidden zones. Conversely, if all λ_j and $z_i(x)$ are real eq.(10.49) shows that u is real. We have also to choose the curve so that the potential is periodic. In this case the points of $Z(x)$ describe cyclic motions on g circles belonging to the real slice of Γ (i.e. λ and μ real, so $R(\lambda) > 0$) and this motion is periodic of period ℓ, so that $u(x)$ has period ℓ. It follows from eq.(10.48) that the quasi momentum is given by:

$$\mathcal{P}(\lambda) = \int_0^\ell \frac{\sqrt{R(\lambda)}}{B(\lambda, x)} dx \tag{10.58}$$

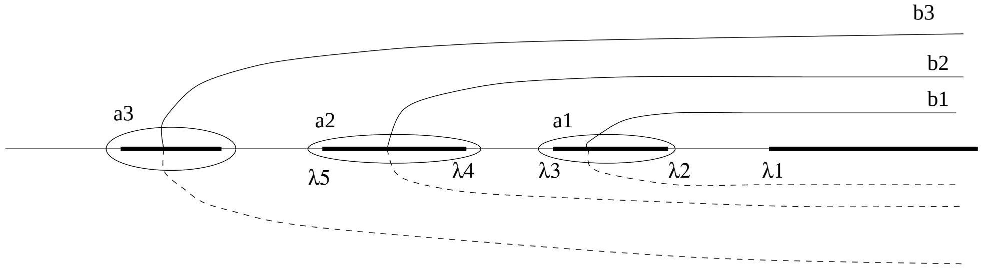
At a point β_j which is a boundary of a non degenerate forbidden zone, $\mathcal{P}(\lambda)$ changes from real to pure imaginary, which implies that $R(\lambda)$ changes sign. Hence all such zone boundaries appear amongst the branch points of Γ. All other periodic or antiperiodic levels are therefore points where $t(\lambda)$ is tangent to $t = \pm 2$, and there is an infinite number of them.

It is worth mentioning that, once a genus g Riemann surface Γ is chosen with moduli quantized to produce a periodic finite zone real potential, the KdV flows with respect to higher times preserves these conditions, and automatically produce a family of continuous deformations parameters.

This analysis also allows to compare the spectral curve of the monodromy matrix $T(\lambda)$ corresponding to a finite zone potential u with Γ. By definition, this spectral curve Γ_T is given by the equation $\mu^2 - t(\lambda)\mu + 1 = 0$. Setting $s = 2\mu - t(\lambda)$ we get the hyperelliptic equation $s^2 = \Delta(\lambda)$. The entire function $\Delta(\lambda)$ admits an infinite product representation:

$$\Delta(\lambda) = \prod_{i=1}^{2g+1} \left(1 - \frac{\lambda}{\lambda_i}\right) \prod_{j=1}^{\infty} \left(1 - \frac{\lambda}{\beta_j}\right)^2$$

where the β_j correspond to the degenerate forbidden zones and the λ_i are the branch points of Γ. This infinite product is convergent because since $t(\lambda) \sim 2\cos\sqrt{-\lambda}\ell$ we have $\beta_j \sim -j^2\pi^2/\ell$. Note that the potential u is determined by Γ so the β_j are really functions of the moduli λ_i of Γ. The birationnal transformation $s = s' \prod_{j=1}^{\infty}(1 - \frac{\lambda}{\beta_j})$ transforms the equation of Γ_T into the equation $s'^2 = \prod_{i=1}^{2g+1}(1 - \frac{\lambda}{\lambda_i})$, that is the equation of Γ. All points $(s = 0, \lambda = \beta_j)$ are singular points of Γ_T and Γ is the desingularization of Γ_T. Conversely Γ_T is obtained from Γ by identifying pair of points $(\lambda = \beta_j, s' = \pm s'(\beta_j))$ which accumulate at ∞.



We view the desingularized Riemann surface as a two–sheeted covering of the λ plane with cuts on the forbidden zones. We choose for a–cycles loops around the compact cuts on the above sheet. Then the b–cycles start from ∞ on the upper sheet, cross the a–cycles once, go to the lower sheet through the cut and return to ∞. The regular abelian differentials are of the form $\lambda^j d\lambda/s'$ for $j = 0, \cdots, g-1$. Their a–periods are real and their b–periods are pure imaginary. Hence normalized regular abelian differentials are real combinations of the above, so the matrix of b–periods is pure imaginary. Moreover the second kind abelian differential $\Omega^{(1)}$ is of the form $\lambda^g d\lambda/s'$ plus a combination of first kind differentials with real coefficients, so that $U^{(1)}$ is real. The periodicity condition of the potential eq.(10.55) can only be realized with $m = 0$ and becomes

$$\ell \int_{b_j} \Omega^{(1)} = 2i\pi n_j$$

We now show that the Bloch quasi momentum $\mathcal{P}(\lambda)$ is directly related to $\Omega^{(1)}$, indeed we have:

$$d\mathcal{P} = -\ell\Omega^{(1)}$$

To see that, remark that $\mu = e^{\mathcal{P}(\lambda)}$ satisfies $\mu + \mu^{-1} = t(\lambda)$ where $t(\lambda)$ is an entire function of λ. Hence μ has no pole or zero at finite distance. Moreover μ is a well–defined function on the smooth curve Γ since it is the quotient of Baker-Akhiezer functions of the point $P \in \Gamma$:

$$\mu(P) = \frac{\psi(P, x + \ell)}{\psi(P, x)}$$

and this implies the behavior at ∞: $\mu \sim e^{\ell\sqrt{\lambda}}$. Hence $d\mathcal{P} = d\mu/\mu$ is an abelian differential on Γ without singularities at finite distance, and having a double pole at ∞, i.e. $d\mu/\mu \sim -\ell dz/z^2$ where $z = 1/\sqrt{\lambda}$ is the local parameter. For real λ, μ is real of fixed sign on a forbidden zone, so that the a–periods of $d\mathcal{P} = d\mu/\mu$ vanish. This characterizes $d\mathcal{P} = -\ell\Omega^{(1)}$. On the other hand μ lives on the unit circle in an allowed zone. so the b–periods of $d\mu/\mu$ are pure imaginary, and of the form $2i\pi n_j$ since μ is well–defined on the curve, in agreement with the periodicity condition.

10.7 Action–angle variables.

We want to express the restriction of the symplectic forms corresponding to the two Poisson brackets $\{\ \}_{1,2}$ on the finite zone solutions of the KdV hierarchy. We start from the first Poisson bracket (10.30). The associated symplectic form reads:

$$\omega_1 = \int dx\, \delta u(x) \wedge \int^x \delta dy\, u(y)$$

This Poisson bracket is degenerate, and the kernel is $\int u(x)dx$. So this integral has to be kept constant under the variation δ. In the finite zone case one chooses a genus g hyperelliptic Riemann surface, and a divisor of degree g on it, which determine the potential u. The variations δu are expressed through the variations of the moduli of the curve, and the variations of the divisor.

Proposition. *The restriction of the symplectic form ω_1 on the finite zone solution constructed from the curve Γ : $\mu^2 = R(\lambda) = \prod_{i=1}^{2g+1}$ is expressed in terms of the divisor of poles $\nu_i = (z_i, \sqrt{R(z_i)})$ of the Baker–Akhiezer function as:*

$$\omega_1 = 2\sum_{i=1}^{g} \delta\mathcal{P}(\nu_i) \wedge \delta z_i$$

where $\mathcal{P}$ is the Bloch momentum such that $d\mathcal{P} = -\ell\Omega^{(1)}$ and $\mathcal{P} = k\ell + O(1/k)$ at infinity (which fixes the primitive of $\Omega^{(1)}$).

Proof. By analogy with the discussion in Chapter [4] we introduce a 1–form on the curve Γ with values 2–forms on phase space:

$$K = \frac{<\psi^*\delta L \wedge \delta\psi>}{W(\psi^*, \psi)} d\lambda$$

Here ψ is the Baker–Akhiezer function on Γ given in eq.(10.43), and ψ^* is its adjoint, i.e. is the composition of ψ with the hyperelliptic involution. They both satisfy the equation $L\psi = \lambda\psi$ with $L = \partial_x^2 + u(x)$. Their Wronskian $W(\psi^*, \psi)$ is independent of x. It is a meromorphic function on Γ with $2g+1$ poles located at the poles of ψ and ψ^* and at ∞ (where $W \sim 2k$) and $2g+1$ zeroes located at the branch points at finite distance (because at such a point $\psi^* = \psi$). The notation $<\ >$ means:

$$< f(x) >= \lim_{\ell\to\infty} \frac{1}{\ell}\int_0^\ell f(x)dx$$

Recalling that the potentials obtained from the finite–zone construction are almost periodic (and the Baker–Akhiezer functions are almost Bloch) this limit makes sense when $\ell \to \infty$. Finally the variation $\delta\psi$ is taken by keeping λ constant, and this induces poles at the branch points, see Chapter [4].

We write as usual that the sum of residues of K on Γ vanishes. The residue of K at ∞ is precisely the form ω_1. Indeed, using $\psi = e^{kx}(1 + \alpha/k + \cdots)$ with $u = -2\partial_x\alpha$ and $\psi^* = e^{-kx}(1 - \alpha/k + \cdots)$, so that $W(\psi^*, \psi) = 2k + O(1/k)$, we have:

$$\mathrm{Res}_\infty K = - < \delta u \wedge \delta\alpha >= \frac{1}{2\ell}\omega_1$$

The form K has poles at finite distance at the divisor of poles of ψ (the poles of ψ^* are cancelled by the Wronskian) and at the branch points of the covering $(\lambda,\mu)\to\lambda$. At a pole $\nu_i = (z_i, \sqrt{R(z_i)})$ of ψ we have, $\delta\psi = (\delta z_i/(\lambda - z_i))(\psi + O(1))$, since we keep λ constant in the variation. We get a contribution:

$$\mathrm{Res}_{\nu_i} K = \mathrm{Res}_{\nu_i} \frac{<\psi^*\delta L\psi>}{W(\psi^*,\psi)} \wedge \frac{\delta z_i}{\lambda - z_i} d\lambda$$

Varying $L\psi = \lambda\psi$ we have $<\psi^*\delta L\psi> = -<\psi^*(L-\lambda)\delta\psi>$. Integrating by parts, and using $(L-\lambda)\psi^* = 0$ we get:

$$<\psi^*(L-\lambda)\delta\psi> = \frac{1}{\ell}\Big[\psi^*\partial_x\delta\psi - \partial_x\psi^*\delta\psi\Big]_0^\ell$$

Assuming that ℓ is close enough to an almost period of u we write:

$$\psi(x+\ell) = e^{\mathcal{P}}\psi(x), \quad \psi^*(x+\ell) = e^{-\mathcal{P}}\psi^*(x)$$

where $\mathcal{P}$ depends on the point on Γ but not on x, and get:

$$<\psi^*\delta L\psi> = -\frac{1}{\ell}\delta\mathcal{P}\, W(\psi^*,\psi)$$

so that the contribution of the pole ν_i is:

$$\mathrm{Res}_{\nu_i} K = -\frac{1}{\ell}\delta\mathcal{P}(\nu_i)\wedge\delta z_i$$

We now show that the sum of the contributions to the residues of K at the branch points vanishes. At a branch point λ_i, $d\lambda$ has a zero, so does $W(\psi^*,\psi)$, and $\delta\psi$ has a pole. Since $d\lambda/W$ is regular it is enough to study $<\psi^*\delta L\wedge\delta\psi>$, which contains a simple pole, coming from $\delta\psi$. Since $\psi^* - \psi$ vanishes at the branch point, we can replace ψ^* by ψ in this expression. The operator δL being the multiplication by δu, we can write $\psi\delta L = \delta L\psi = -(L-\lambda)\delta\psi$. Finally using $\delta\psi\wedge\delta\psi = 0$ we arrive at

$$<\psi^*\delta L\wedge\delta\psi> = -<\delta\psi''\wedge\delta\psi> = <\delta\psi'\wedge\delta\psi'> - \frac{1}{\ell}[\delta\psi'\wedge\delta\psi]_0^\ell$$

Since $\delta\psi'\wedge\delta\psi' = 0$ we have:

$$\mathrm{Res}_{\lambda_i} K = -\mathrm{Res}_{\lambda_i}\left(\frac{1}{\ell}\left[\delta\psi'\wedge\delta\psi\right]_0^\ell \frac{d\lambda}{W(\psi^*,\psi)}\right)$$

To evaluate further this residue, we write $\psi = \psi_0 + \mu\psi_1 + \mu^2\psi_2 + \cdots$ and $\psi^* = \psi_0 - \mu\psi_1 + \mu^2\psi_2 + \cdots$ and noting that that $\delta\psi_k$ are finite coefficients, we get:

$$\mathrm{Res}_{\lambda_i} K = -\mathrm{Res}_{\lambda_i}\left(\frac{1}{\ell}\left[\delta\psi_0'\psi_1 - \delta\psi_0\psi_1'\right]_0^\ell \wedge\delta\mu\frac{d\lambda}{W(\psi^*,\psi)}\right)$$

It is easily checked that this residue at $(\lambda_i, 0)$ is identical to the residue at the same branch point of the form on Γ:

$$K_2 = -\left[\psi'^*\delta\psi - \psi'\delta\psi^*\right]_0^\ell \wedge \frac{\delta\mu}{\mu}\frac{d\lambda}{W(\psi^*,\psi)}$$

We write that the sum of residues of K_2 vanishes. The other potential poles of K_2 are at ∞ and the ν_i and ν_i^*. At infinity K_2 is regular because $d\lambda/W \sim dk$ has a double pole, but $\delta\mu/\mu = \delta R/(2R) = O(1/\lambda)$ has a double zero, while $[\psi'^*\delta\psi - \psi'\delta\psi^*]_0^\ell$ behaves as $[\delta\alpha]_0^\ell$ hence is regular. The other possible poles are at the poles of ψ and ψ^*. However, using that the Baker–Akhiezer functions are almost Bloch one can compute directly that $[\psi'^*\delta\psi - \psi'\delta\psi^*]_0^\ell = \delta\mathcal{P}(\psi'^*\psi + \psi'\psi^*)(0)$ so these poles disappear in K_2. As a consequence the sum of residues of K_2 at the branch points vanishes, and so does the sum of residues of K at the branch points. ∎

We can analyze the second symplectic structure in a similar way. The symplectic form ω_2 is expressed most simply in terms of the Miura variable p such that $u = p' - p^2$. Since the Poisson bracket is given by eq.(10.33) the symplectic form, in the variable p reads:

$$\omega_2 = \int dx\, \delta p(x) \wedge \int^x \delta dy\, p(y)$$

Note that, while ω_2 is not degenerate in the variable u, it has a kernel $\int_0^\ell p(x)dx$ in the variable p. So this quantity has to be kept fixed in the variations. With the same notations as in the previous Proposition:

Proposition. *The restriction of the symplectic form ω_2 on the finite zone solution is expressed as:*

$$\omega_2 = -\sum_{i=1}^{g} \delta\mathcal{P}(\nu_i) \wedge \frac{\delta z_i}{z_i}$$

Proof. Now we introduce the 1–form K on the curve Γ with values 2–forms on phase space:

$$K = \frac{< \psi^* \delta L \wedge \delta\psi >}{W(\psi^*, \psi)} \frac{d\lambda}{\lambda}$$

In contrast to the previous case, there is no pole at ∞ but a new pole appears at $\lambda = 0$. Let us compute the residue of K at $\lambda = 0$. Remarkably, the Miura variable p is related to ψ by

$$p = -\left.\frac{\partial_x \psi}{\psi}\right|_{\lambda=0}$$

and we can express $\psi(x) \equiv \psi(\lambda = 0, x)$ and $\psi^*(x) \equiv \psi^*(\lambda = 0, x)$ in terms of p as:

$$\psi(\lambda, x) = e^{-\int^x p(y)dy}, \quad \left.\frac{\psi^*(\lambda, x)}{W(\psi^*, \psi)}\right|_{\lambda=0} = -\psi(x) \int^x \psi^{-2}(y)dy + a\psi(x)$$

When u is periodic of period ℓ we can choose p periodic and ψ is a Bloch wave with Bloch momentum given by $\mathcal{P}(\lambda = 0) = -\int_0^\ell p(y)dy$, hence in the kernel of ω_2. The coefficient a in ψ^* is adjusted so that ψ^* satisfies the Bloch condition $\psi^*(x + \ell) = e^{-\mathcal{P}} \psi^*(x)$. The residue of K at $\lambda = 0$ is:

$$\begin{aligned}
\frac{< \psi^* \delta L \wedge \delta\psi >}{W} &= -\frac{< \psi^* \psi(\delta p' - 2p\delta p) \wedge \int^x \delta p >}{W} \\
&= \frac{< \psi^* \psi\, 2p\delta p \wedge \int^x \delta p >}{W} + \frac{< \delta p \wedge \partial_x \left(\psi^* \psi \int^x \delta p\right) >}{W} - \frac{\left[\psi^* \psi \delta p \wedge \int^x \delta p\right]_0^\ell}{\ell W} \\
&= -\frac{1}{\ell} \int \delta p \wedge \int^x \delta p
\end{aligned}$$

where we have used $(\psi^* \psi)' = -W + 2p\psi^* \psi$ and the fact that $\psi^* \psi$ and p are periodic, so the last term is proportional to $\delta \int_0^\ell p$ hence vanishes.

The analysis of the residues of K at the poles of ψ and ψ^* is exactly the same as in the previous case, replacing everywhere $d\lambda$ by $d\lambda/\lambda$ which accounts for the result $\delta\mathcal{P}(\nu_i) \wedge \delta z_i/z_i$. Finally at the branch points the analysis is also similar but the form K_2 involves $d\lambda/\lambda$ so has a no pole at ∞ but a pole at $\lambda = 0$. As above the residue of this pole is proportional to $\delta\mathcal{P}(\lambda = 0)$ which vanishes. ■

10.8 Analytical description of solitons.

The soliton solutions can be described as a singular limit of the finite zone solutions in which the branch points λ_j coincide in pairs. Consider the curve $s^2 = \lambda \prod_{i=1}^n (\lambda - \lambda_i)^2$ where we take $\lambda_i = p_i^2$ with $0 < p_1 < p_2 < \cdots$.

This singular curve is desingularized by setting $\lambda = k\prod(\lambda - \lambda_i)$, which leads to $\lambda = k^2$. Hence the desingularized curve is the Riemann sphere, which we identify to the complex k plane. The singular curve is obatined from the k plane by identifying the points $k = \pm p_i$ which are both mapped to the point $(s = 0, \lambda = \lambda_i)$. Any Baker-Akhiezer function ψ on the Riemann sphere is given by:

$$\frac{\psi(t,k)}{\psi(0,k)} = \prod_{i=1}^{n} \frac{k - \gamma_i(t)}{k - \gamma_i(0)} e^{\xi(k,t)} \tag{10.59}$$

If such a function comes from a function defined on the singular curve, and this is the case if it is a singular limit of finite zone solution, it must take the same value at identified points, so we must have:

$$\psi(t,p_i) = \psi(t,-p_i), \quad i = 1, \cdots, n, \quad \text{i.e.} \quad \prod_{j=1}^{n} \frac{p_i + \gamma_j(t)}{p_i - \gamma_j(t)} = a_i e^{2\xi(p_i,t)} \tag{10.60}$$

where a_i are time–independent constants. These equations are a linear system for the symmetric functions of the $\gamma_j(t)$ so they determine the time evolution of the divisor of zeroes of ψ.

On the other hand, the soliton solution is well known in terms of τ–functions, see eq.(10.38). From this we can construct the Baker-Akhiezer function through the Sato formula. We have:

$$\frac{\psi(t,k)}{\psi(0,k)} = \frac{\tau(t - [k])}{\tau(t)} \frac{\tau(t)|_{t=0}}{\tau(t - [k])|_{t=0}} e^{\xi(k,t)} \tag{10.61}$$

Recall that $\tau(t)$ depends on times only through the variables $X_i(t) = X_i(0)e^{2\xi(p_i,t)}$ so that:

$$X_i(t - [k]) = \frac{k - p_i}{k + p_i} X_i(t)$$

Plugging this into the formula for τ, $\tau(X) = 1 + \sum X_i + \cdots$, we find that:

$$\frac{\tau(t - [k])}{\tau(t)} = \frac{\prod(k - \gamma_i)}{\prod(k + p_i)} \tag{10.62}$$

so that the right–hand side of eq.(10.61) is exactly of the form eq.(10.59). We shall now show that the γ_i are the same. For this it is sufficient to show that the Baker-Akhiezer function in eq.(10.61) satisfies $\psi(t,p_i) = \psi(t,-p_i)$. Setting $k = \pm p_i$ in eq.(10.61) we see that the coefficient $(k - p_i)/(k + p_i)$ has either a pole or a zero, and this picks up either all the terms containing X_i, or all the terms not containing X_i in the τ–function. These two types of terms are easy to isolate by noting that:

$$\tau_n(X) = \tau_{n-1}(X) + X_i\ \tau_{n-1}\left(\frac{(p_i - p_k)^2}{(p_i + p_k)^2} X_k\right)$$

where τ_{n-1} is given by the formula for $(n-1)$ solitons (X_i removed), and in the coefficient of X_i all arguments X_k in τ_{n-1} are replaced as indicated. This formula is obvious from eq.(10.38). Then the verification of eq.(10.60) is immediate. The divisor $(\gamma_i(t))$ is determined by the zeroes of the function $k \to \tau(t - [k])$.

The equations of motion for the divisor Z are obtained by taking the limit of eq.(10.54). They read:

$$D(k)\gamma_i = -\frac{\prod_j(\gamma_i^2 - p_j^2)\prod_{j\neq i}(k^2 - \gamma_j^2)}{\prod_j(k^2 - p_j^2)\prod_{j\neq i}(\gamma_i^2 - \gamma_j^2)}$$

In particular taking the term in k^{-2} we find:

$$\partial_x\gamma_i = -\frac{\prod_j(\gamma_i^2 - p_j^2)}{\prod_{j\neq i}(\gamma_i^2 - \gamma_j^2)} \tag{10.63}$$

The solution of these equations is of course given by solving the algebraic system eq.(10.60).

We can also obtain the degenerate limit of the Baker–Akhiezer function given by eq.(10.48). Notice that:

$$\frac{\sqrt{R(\lambda)}}{B(\lambda,x)}\Big|_{\lambda=k^2} = k - \frac{1}{2}\sum_i \left(\frac{1}{k-\gamma_i} + \frac{1}{k+\gamma_i}\right)\partial_x\gamma_i \tag{10.64}$$

we can integrate eq.(10.48) to get eq.(10.59).

One can use the rationality of the spectral curve in this degenerate case to express the τ–function as a rational functions of the divisor Z.

Proposition. *The τ–function admits a simple expression in terms of the divisor Z.*

$$\tau(t) = (-1)^{\frac{n(n-1)}{2}} 2^n \left(\prod_{j=1}^n p_j\right) \frac{\prod_{i<j}(\gamma_i+\gamma_j)\prod_{i<j}(p_i+p_j)}{\prod_{i,j}(p_i-\gamma_j)}$$

Proof. We start from the logarithm of eq.(10.62) and apply the operator $\partial_k - D(k)$ which vanishes on $\log\tau(t-[k])$. We get:

$$D(k)\log\,\tau(t) = \sum_i \frac{1}{k-\gamma_i} - \sum_i \frac{1}{k+p_i} + \sum_i \frac{1}{k-\gamma_i} D(k)\gamma_i$$

Since the τ–function depends on times only through the $\gamma_i(t)$ we have $D(k)\log\,\tau(t) = \sum_i \partial_{\gamma_i}\log\,\tau(t)D(k)\gamma_i$. Identifying the residues of the poles in k at $k=\pm p_l$ and $k=\gamma_l$ one finds the n conditions:

$$\sum_i \partial_{\gamma_i}\log\tau(t)\frac{1}{p_l^2-\gamma_i^2}\partial_x\gamma_i = \sum_i \frac{1}{p_l-\gamma_i}\frac{1}{p_l^2-\gamma_i^2}\partial_x\gamma_i, \quad l=1,\cdots,n$$

This is a linear system for the quantities $\partial_{\gamma_i}\log\tau(t)\partial_x\gamma_i$, the solution of which requires the inversion of the Cauchy matrix $M_{ij} = 1/(p_i^2-\gamma_j^2)$. Recalling eq.(8.34) in Chapter [8] we find:

$$(M^{-1})_{ij} = (-1)^{n-1}\frac{\prod_{k\neq i}(p_j^2-\gamma_k^2)\prod_k(p_k^2-\gamma_i^2)}{\prod_{k\neq j}(p_j^2-p_k^2)\prod_{k\neq i}(\gamma_i^2-\gamma_k^2)} = \partial_x\gamma_i\,\frac{\prod_{k\neq i}(p_j^2-\gamma_k^2)}{\prod_{k\neq j}(p_j^2-p_k^2)}$$

This gives:

$$\partial_{\gamma_i}\log\,\tau = (-1)^{n-1}\sum_{j,l}\frac{\prod_{k\neq i}(p_j^2-\gamma_k^2)\prod_{k\neq j}(p_k^2-\gamma_l^2)}{\prod_{k\neq j}(p_j^2-p_k^2)\prod_{k\neq l}(\gamma_l^2-\gamma_k^2)}\frac{1}{p_j-\gamma_l}$$

We consider this expression as a rational function of γ_i. It has poles at $\gamma_i = p_j$ for any j and at $\gamma_i = \pm\gamma_j$ for $j\neq i$, and goes to 0 at $\lambda_i=\infty$. One sees easily that the residue at $\gamma_i = p_j$ is equal to -1, the residue at $\gamma_i=\gamma_j$ vanishes. We finally get:

$$\partial_{\gamma_i}\log\,\tau = \sum_j \frac{1}{p_j-\gamma_i} + \sum_{l\neq i}\frac{r_l}{\gamma_i+\gamma_l}, \quad r_l = \sum_j \frac{\prod_{k\neq j}(\gamma_l^2-p_k^2)\prod_{k\neq i,l}(p_j^2-\gamma_k^2)}{\prod_{k\neq j}(p_j^2-p_k^2)\prod_{k\neq l,i}(\gamma_l^2-\gamma_k^2)}$$

In fact we have $r_l=1$. This is equivalent to the identity:

$$\sum_j \frac{\prod_{k\neq i,l}(p_j^2-\gamma_k^2)}{\prod_{k\neq j}(p_j^2-p_k^2)}\frac{1}{\gamma_l^2-p_j^2} = \frac{\prod_{k\neq i,l}(\gamma_l^2-\gamma_k^2)}{\prod_k(\gamma_l^2-p_k^2)}$$

which is easily checked by comparing the residues in γ_l^2 in both sides. This gives the final simple result:

$$\partial_{\gamma_i}\log\,\tau = \sum_j \frac{1}{p_j-\gamma_i} + \sum_{l\neq i}\frac{1}{\gamma_i+\gamma_l}$$

Integrating this formula, we find:

$$\tau = C\,\frac{\prod_{i<j}(\gamma_i+\gamma_j)}{\prod_{i,j}(p_i-\gamma_j)} \tag{10.65}$$

where C does not depend on the γ_i hence is independent on times. To find it, remark that when the times are set to $-\infty$ we have $\tau = 1$. Eq.(10.62) implies that the $\gamma_i(-\infty)$ are equal to the $-p_i$ up to order. Inserting these values into eq.(10.65) gives

$$C = (-1)^{\frac{n(n-1)}{2}} 2^n \prod_j p_j \prod_{i<j} (p_i + p_j)$$

■

We would like to discuss the question of the Poisson structures of the KdV equation in these variables. As we know, we have a whole hierarchy of these structures. One can describe the restrictions of the symplectic forms to the manifold of finite zone solutions. Let ω_k, $k = 1, 2$ be the restricted k–th symplectic form. Then we have in terms of analytical variables:

$$\omega_k = \sum_{i=1}^{n} \delta\mathcal{P}(\nu_i) \wedge \frac{\delta z_i}{z_i^{k-1}} \tag{10.66}$$

where the Bloch momentum $\mathcal{P}(\lambda)$ is defined as:

$$\mathcal{P}(\lambda) = \log \frac{\psi(\lambda, x = \ell)}{\psi(\lambda, x = -\ell)}$$

To compute this Bloch momentum in the soliton limit, we use eq.(10.59) and we send $\ell \to \infty$. The eq.(10.60) shows that, up to order, we have $\gamma_i(\ell) \to p_i$, $\gamma_i(-\ell) \to -p_i$. Using eq.(10.58) and eq.(10.64), which is symmetric in the γ_i, we get:

$$\mathcal{P}(k) = 2\ell k - \log \prod_j \left(\frac{k + p_j}{k - p_j} \right) \quad \mathrm{mod} \quad i\pi$$

Hence,

$$\omega_k = 2 \sum_{i=1}^{n} \delta \log \prod_j \left(\frac{\gamma_i - p_j}{\gamma_i + p_j} \right) \wedge \frac{\delta\gamma_i}{\gamma_i^{2k-3}} = 4 \sum_{ij} \frac{\delta\gamma_i \wedge \delta p_j}{(\gamma_i^2 - p_j^2)\gamma_i^{2k-4}}$$

The form with $k = 2$ is the second Hamiltonian structure of KdV. It corresponds to the Magri–Virasoro Poisson bracket. As a check, we verify that we find the correct equation of motion for the divisor γ_i, eq.(10.63). We use the second symplectic structure ω_2 which gives the Poisson bracket:

$$\{\gamma_i, \gamma_j\}_2 = 0, \quad \{p_i, p_j\}_2 = 0, \quad \{\gamma_i, p_j\}_2 = \frac{\prod_{k \neq i}(p_j^2 - \gamma_k^2) \prod_k (\gamma_i^2 - p_k^2)}{\prod_{k \neq j}(p_j^2 - p_k^2) \prod_{k \neq i}(\gamma_i^2 - \gamma_k^2)}$$

where we have used the inverse of the Cauchy matrix $M_{ij} = 1/(p_i^2 - \gamma_j^2)$ computed above, to invert the symplectic matrix. We have to check that:

$$\partial_x \gamma_i = \{H_1, \gamma_i\}_2, \quad H_1 = \sum_j p_j$$

This reduces to the equation:

$$\sum_j \frac{\prod_{k \neq i}(p_j^2 - \gamma_k^2)}{\prod_{k \neq j}(p_j^2 - p_k^2)} = 1$$

Consider the left–hand side as a rational function of p_1^2. It tends to 1 when $p_1^2 \to \infty$ and the residues at the poles p_j^2 for $j \neq 1$, vanish.

10.9 Witham's equations.

References

[1] D.J. Korteweg and G. de Vries *" On the change of form of long waves advancing in a rectangular channel and on a new type of long stationary wave."* Philos. Mag. 39(1895) 422–443.

[2] C. S. Gardner, J. M. Greene, M. D. Kruskal, R. M. Miura. *"Method for solving the Korteweg–de Vries equation"*. Phys. Rev. Lett. 19 (1967), p. 1095.

[3] P.D. Lax *"Integrals of nonlinear equations of evolution and solitary waves."* Comm. Pure and Appl. Math. 21 (1968) 467–490.

[4] V.E. Zakharov, L.D. Faddeev *"Korteweg–de Vries equation: a completely integrable Hamiltonian system."* Funct. Anal. Appl. Vol. 8 N^0 4 (1971) 280–287

[5] S. Novikov, S.V. Manakov, L.P. Pitaevskii and V.E. Zakharov **"Theory of solitons. The inverse scattering method"** Consultants Bureau (1984)

[6] L. D. Faddeev, L. A. Takhtajan. **"Hamiltonian Methods in the Theory of Solitons"**. Springer 1986.

[7] L.A. Dickey **"Soliton Equations and Hamiltonian Systems."** World Scientific (1991)

[8] I. Krichever and D.H. Phong *"Symplectic forms in the theory of solitons"* Surveys in Differential geometry 4(1998)239–313 International press.

Chapter 11

The Toda field theories

In this Chapter, we study the Toda field theories which are generalizations of the Liouville equation. The equations of motion are both integrable, i.e. they admit a zero curvature representation, and conformally invariant. Therefore they allow to see the interplay between conformal symmetry and classical Yang–Baxter equation.

11.1 The Liouville equation.

In his studies of surfaces with constant curvature, Liouville introduced the equation $\nabla^2\varphi = 4e^{2\varphi}$, where $\nabla^2 = \partial_x^2 - \partial_t^2$. In the light cone coordinates $x_\pm = x \pm t$ the equation reads:

$$\partial_{x_+}\partial_{x_-}\varphi = e^{2\varphi} \tag{11.1}$$

What is remarkable about this non–linear partial differential equation is that Liouville was able to give its general solution:

$$e^{2\varphi} = -\frac{F'(x_+)G'(x_-)}{(F(x_+) - G(x_-))^2} \tag{11.2}$$

where F and G are arbitrary functions of the single variables x_+ and x_- respectively. Such functions are called chiral functions. A very important property of this equation is its conformal invariance. Namely changing coordinates

$$x_+ = f(x'_+), \quad x_- = g(x'_-), \quad \varphi = \varphi' - \frac{1}{2}\log(\partial f \partial g)$$

we see that the equation for the primed variables is the same as for the unprimed ones.

There is an important connection between the Liouville equation and the Schroedinger equation. The field $e^{-\varphi}$ satisfies two chiral equations:

$$\begin{aligned}
(\partial^2_{x_+} - \mathcal{U}(x_+))e^{-\varphi} &= 0, \quad \mathcal{U} = (\partial_{x_+}\varphi)^2 - \partial^2_{x_+}\varphi, \quad \partial_{x_-}\mathcal{U} = 0 \\
(\partial^2_{x_-} - \overline{\mathcal{U}}(x_-))e^{-\varphi} &= 0, \quad \overline{\mathcal{U}} = (\partial_{x_-}\varphi)^2 - \partial^2_{x_-}\varphi, \quad \partial_{x_+}\overline{\mathcal{U}} = 0
\end{aligned}$$

Indeed, the two Schroedinger equations are obtained readily by computing $\partial^2_{x_+}e^{-\phi}$ and the chirality of $\mathcal{U}$ and $\overline{\mathcal{U}}$ is then proved using the Liouville equation.

Conversely, starting from two arbitrary chiral potentials $\mathcal{U}$ and $\overline{\mathcal{U}}$, we construct solutions $\xi_i(x_+)$, $i = 1, 2$ and $\overline{\xi}_i(x_-)$, $i = 1, 2$ of the two Schroedinger equations, normalized such that their Wronskians are:

$$\xi_1\xi'_2 - \xi'_1\xi_2 = 1, \quad \overline{\xi}_1\overline{\xi}'_2 - \overline{\xi}'_1\overline{\xi}_2 = -1$$

and reconstruct the Liouville field by the formula:

$$e^{-\varphi(x_+,x_-)} = \xi_1(x_+)\overline{\xi}_1(x_-) + \xi_2(x_+)\overline{\xi}_2(x_-)$$

It is easy to check that φ satisfies the Liouville equation. One can relate this solution to eq.(11.2). Writing:

$$e^{-\varphi} = \frac{F-G}{\sqrt{-F'G'}} = \frac{F}{\sqrt{F'}}\frac{1}{\sqrt{-G'}} - \frac{1}{\sqrt{F'}}\frac{G}{\sqrt{-G'}}$$

one is lead to take:

$$\xi_1 = \frac{F}{\sqrt{F'}}, \quad \xi_2 = -\frac{1}{\sqrt{F'}}, \quad \overline{\xi}_1 = \frac{1}{\sqrt{-G'}}, \quad \overline{\xi}_2 = \frac{G}{\sqrt{-G'}}$$

One can check that the Wronskians are $W(\xi_1,\xi_2) = 1$ and $W(\overline{\xi}_1,\overline{\xi}_2) = -1$ hereby completing the identification. It follows that $F = -\xi_1/\xi_2$ and $G = \overline{\xi}_2/\overline{\xi}_1$. Finally one can use this identification to compute the potential $\mathcal{U}$ and $\overline{\mathcal{U}}$ in terms of F and G.

$$\mathcal{U} = -\frac{1}{2}S(F), \quad \overline{\mathcal{U}} = -\frac{1}{2}S(G), \quad S(f) = \frac{f'''}{f'} - \frac{3}{2}\frac{f''^2}{f'^2}$$

where $S(f)$ is the Schwarzian derivative. Remark that the two solutions ξ_1, ξ_2 are defined up to a linear transformation of determinant 1. This translates onto an homographic transformation on F which leaves the potential $\mathcal{U}$ invariant, and similarly for G and $\overline{\mathcal{U}}$. When the two homographic transformations of F and G are inverse to each other, φ is invariant.

Invariance of the Liouville equation under change of coordinates reflects itself into covariance properties of the Schroedinger equation. Changing $x = f(x')$, we have:

$$\xi'(x') = \frac{\xi(x)}{\sqrt{\partial f}}, \quad \mathcal{U}'(x') = \mathcal{U}(x)(\partial f)^2 + S(f), \quad (\partial'^2 - \mathcal{U}')\xi' = (\partial f)^{\frac{3}{2}}(\partial^2 - \mathcal{U})\xi$$

These equations exhibit the conformal properties of the objects involved, namely, ξ is a differential of weight $-\frac{1}{2}$ $\mathcal{U}$ is a Schwartzian connection, and $(\partial^2 - \mathcal{U})\xi$ is a differential of weight $\frac{3}{2}$.

In the next section, we generalize this setup to a large class of two–dimensional field theories called Toda field theories.

11.2 The Toda systems and their zero-curvature representations.

Toda field theories are two–dimensional analogs of the Toda chain studied in Chapter [3]. We first recall some notations concerning Lie algebras. Let $\mathcal{G}$ be a simple Lie algebra of rank n, equipped with an invariant scalar product denoted by $(\, , \,)$. We choose a Cartan subalgebra $\mathcal{H}$, with an orthonormal basis $\{H_i\}$. We also need the decomposition of $\mathcal{G}$ as:

$$\mathcal{G} = \mathcal{N}_- \oplus \mathcal{H} \oplus \mathcal{N}_+$$

We will denote by $\mathcal{B}_\pm = \mathcal{H} \oplus \mathcal{N}_\pm$ the two Borel subalgebras. We recall the commutation relations in a Cartan-Weyl basis:

$$[H_i, H_j] = 0, \quad [H, E_{\pm\alpha}] = \pm\alpha(H)\ E_{\pm\alpha}, \quad [E_\alpha, E_{-\alpha}] = H_\alpha$$

with $H_\alpha = \sum_i \alpha(H_i)H_i$ and $\beta(H_\alpha) = (\alpha,\beta)$. We adopt the normalisation $(E_\alpha, E_{-\alpha}) = 1$. The simple roots will be denoted by α_i with $i = 1,\cdots,n$. The fundamental highest weights, which we denote by $\lambda^{(i)}$, are defined by: $2(\lambda^{(i)},\alpha_j)/(\alpha_j,\alpha_j) = \delta_{ij}$. Both the roots and the weights belong to the dual of the Cartan subalgebra. The set of simple roots as well as the set of fundamental weights form a basis of $\mathcal{H}^*$. In particular, the roots can be decomposed onto the fundamental weights as $\alpha_i = \sum_j \lambda^{(j)}\ a_{ji}$ where $a_{ij} = 2(\alpha_i,\alpha_j)/(\alpha_i,\alpha_i)$ is the Cartan matrix of $\mathcal{G}$.

The Toda fields equations are obtained as follows: let $\varphi_i(x,t), i = 1, \cdots, n$ be n scalar fields. We define $\Phi = \frac{1}{2}\sum_i \varphi_i H_i$ which belongs to the Cartan subalgebra. In the light-cone variables $x_\pm = x \pm t$ and $\partial_{x_\pm} = \frac{1}{2}(\partial_x \pm \partial_t)$, the Toda field equations are:

$$\partial_{x_+}\partial_{x_-}\ \Phi = \frac{1}{2} \sum_{\alpha_j\ \text{simple}} H_\alpha \exp(2\alpha_j(\Phi)) \tag{11.3}$$

To write these equations in components, it is simpler to introduce $\varphi^i = \lambda^{(i)}(\Phi)$ for $i = 1, \cdots, n$ where $\lambda^{(i)}$ are the n fundamental weights. Since we have $\lambda^{(i)}(H_{\alpha_j}) = (\alpha_i, \alpha_i)\delta_{ij}/2$, and $\alpha_i(\Phi) = \sum_j \lambda^{(j)}(\Phi)\ a_{ji}$. Thus the Toda equations of motion can be written as:

$$\partial_{x_+}\partial_{x_-}\varphi^i = \frac{(\alpha_i, \alpha_i)}{4}\ e^{\,2\sum_j \varphi^j\ a_{ji}} \tag{11.4}$$

A very important property of these equations, whose proof is obvious, is their conformal invariance:

Proposition. *The Toda field equations are invariant under the transformations $x_+ \to x'_+$ with $x_+ = f(x'_+)$ and $x_- \to x'_-$ with $x_- = g(x'_-)$ and*

$$\Phi'(x'_+, x'_-) = \Phi(f(x'_+), g(x'_-)) + \tfrac{1}{2}H_\rho \ln(\partial f(x'_+)\partial g(x'_-))$$

where H_ρ is the Weyl vector characterized by the property $\alpha_j(H_\rho) = 1$ for all simple root α_j.

We introduce now a zero curvature representation for the Toda field equations.

Proposition. *Define the elements of $\mathcal{G}$:*

$$\mathcal{E}_\pm = \sum_{\alpha_j\ \text{simple}} E_{\pm\alpha_j}$$

and consider the connection $A_{x_\pm} \in \mathcal{B}_\pm$:

$$A_{x_+} = \partial_{x_+}\Phi + e^{\mathrm{ad}\,\Phi}\mathcal{E}_+, \quad A_{x_-} = -\partial_{x_-}\Phi + e^{-\mathrm{ad}\,\Phi}\mathcal{E}_- \tag{11.5}$$

The zero curvature condition, $F_{x_+x_-} = \partial_{x_+}A_{x_-} - \partial_{x_-}A_{x_+} + [A_{x_+}, A_{x_-}] = 0$ is equivalent to the field equations eq.(11.3).

Proof. The zero curvature condition can be worked out using only the Lie algebra structure of $\mathcal{G}$. We have:

$$\begin{aligned}
\partial_{x_+}A_{x_-} &= -\partial_{x_+}\partial_{x_-}\Phi - e^{-\mathrm{ad}\Phi}[\partial_{x_+}\Phi, \mathcal{E}_-] \\
\partial_{x_-}A_{x_+} &= \partial_{x_+}\partial_{x_-}\Phi + e^{\mathrm{ad}\Phi}[\partial_{x_-}\Phi, \mathcal{E}_+] \\
[A_{x_+}, A_{x_-}] &= e^{-\mathrm{ad}\Phi}[\partial_{x_+}\Phi, \mathcal{E}_-] + e^{\mathrm{ad}\Phi}[\partial_{x_-}\Phi, \mathcal{E}_+] + [e^{\mathrm{ad}\Phi}\mathcal{E}_+, e^{-\mathrm{ad}\Phi}\mathcal{E}_-]
\end{aligned}$$

so that $F_{x_+x_-} = -2\partial_{x_+}\partial_{x_-}\Phi + [e^{\mathrm{ad}\Phi}\mathcal{E}_+, e^{-\mathrm{ad}\Phi}\mathcal{E}_-]$. This can be simplified since

$$\left[e^{\mathrm{ad}\Phi}\mathcal{E}_+, e^{-\mathrm{ad}\Phi}\mathcal{E}_-\right] = \left[\sum_{\alpha\ \text{simple}} e^{\alpha(\Phi)}E_\alpha, \sum_{\beta\ \text{simple}} e^{\beta(\Phi)}E_{-\beta}\right] = \sum_{\alpha\ \text{simple}} e^{2\alpha(\Phi)}H_\alpha$$

∎

Example. Let us give the example associated to the Lie algebra $\mathcal{G} = sl_2$. Let E_+, E_-, H be its generators with commutation relations

$$[H, E_\pm] = \pm 2E_\pm, \quad [E_+, E_-] = H$$

We have $\Phi = \frac{1}{2}\varphi H$, and the connection reads in the fundamental representation:

$$A_{x_+} = \begin{pmatrix} \frac{1}{2}\partial_{x_+}\varphi & e^\varphi \\ 0 & -\frac{1}{2}\partial_{x_+}\varphi \end{pmatrix} \quad \text{and} \quad A_{x_-} = \begin{pmatrix} -\frac{1}{2}\partial_{x_-}\varphi & 0 \\ e^\varphi & \frac{1}{2}\partial_{x_-}\varphi \end{pmatrix}$$

The zero curvature condition is equivalent to the Liouville equation (11.1).

11.3 Solution of the Toda field equations.

The Toda field equtions being conformal invariant, one can solve them by splitting the chiralities, as in the Liouville equation. To do that, we use the zero–curvature representation. When Φ is a solution of Toda field equations, we have $F_{x_+x_-} = 0$ (and conversely), so we can write $A_{x_\pm}$ as a pure gauge

$$A_{x\pm} = -\left(\partial_{x_\pm}\Psi\right)\,\cdot\Psi^{-1} \tag{11.6}$$

Equivalently we have the linear system $(\partial_{x_\pm} + A_{x_\pm})\Psi = 0$.

Proposition. *Let $Q_\pm \in \exp\mathcal{B}_\pm$ be defined by two decompositions of Ψ as:*

$$\Psi = e^{\pm\Phi}\; N_\mp\; Q_\pm \quad \text{with} \quad N_\mp \in \exp\mathcal{N}_\mp, \qquad Q_\pm \in \exp\mathcal{B}_\pm \tag{11.7}$$

then, $Q_\pm$ satisfy the following equations of motion:

$$Q_-\partial_{x_+}Q_-^{-1} = 0, \qquad\qquad Q_-\partial_{x_-}Q_-^{-1} = -\overline{P} + \mathcal{E}_- \tag{11.8}$$

$$Q_+\partial_{x_-}Q_+^{-1} = 0, \qquad\qquad \partial_{x_+}Q_+\cdot Q_+^{-1} = P - \mathcal{E}_+ \tag{11.9}$$

where $P(x_+)$ and $\overline{P}(x_-)$ are chiral fields with values in the Cartan subalgebra.

Proof. We write Ψ in two different ways $\Psi = e^{-\Phi}\;G_1 = e^{\Phi}\;G_2$. Plugging this into eq.(11.6) we get:

$$G_1\cdot\partial_{x_+}G_1^{-1} = e^{2\mathrm{ad}\Phi}\mathcal{E}_+, \qquad\qquad \partial_{x_+}G_2\cdot G_2^{-1} = -2\partial_{x_+}\Phi - \mathcal{E}_+ \tag{11.10}$$

$$G_1\cdot\partial_{x_-}G_1^{-1} = -2\partial_{x_-}\Phi + \mathcal{E}_-, \qquad\qquad \partial_{x_-}G_2\cdot G_2^{-1} = -e^{-2\mathrm{ad}\Phi}\mathcal{E}_- \tag{11.11}$$

Let us prove the equations for Q_-. Performing the Gauss decomposition $G_1 = N_+Q_-$ we obtain:

$$N_+(Q_-\partial_{x_+}Q_-^{-1})N_+^{-1} + N_+\partial_{x_+}N_+^{-1} = e^{2\mathrm{ad}\Phi}\mathcal{E}_+$$

or, multiplying on the right by N_+ and on the left by N_+^{-1},

$$Q_-\partial_{x_+}Q_-^{-1} = -\partial_{x_+}N_+^{-1}N_+ + N_+^{-1}e^{2\mathrm{ad}\Phi}\mathcal{E}_+N_+$$

Since the left–hand side is in $\mathcal{B}_-$ and the right–hand side is in $\mathcal{N}_+$, they both vanish, and we get $Q_-\partial_{x_+}Q_-^{-1} = 0$ and $N_+\partial_{x_+}N_+^{-1} = e^{2\mathrm{ad}\Phi}\mathcal{E}_+$. This proves that Q_- only depends on x_-. Next, using eq.(11.11) and the decomposition $G_1 = N_+Q_-$ we obtain:

$$N_+(Q_-\partial_{x_-}Q_-^{-1})N_+^{-1} + N_+\partial_{x_-}N_+^{-1} = -2\partial_{x_-}\Phi + \mathcal{E}_- \tag{11.12}$$

The right–hand side has lowest height -1 given by the $\mathcal{E}_-$ term. So the lowest height term in $Q_-\partial_{x_-}Q_-^{-1}$ is also equal to $\mathcal{E}_-$. Since $Q_-\partial_{x_-}Q_-^{-1} \in \mathcal{B}_-$ it is necessarily of the form $-\overline{P} + \mathcal{E}_-$ with $\overline{P} \in \mathcal{H}$ only depending on x_-. The equations for Q_+ are proved similarly. ∎

One can reconstruct the Toda field Φ from the knowledge of $Q_\pm$. Let $|\lambda^{(i)}>$, $i = 1, \cdots, r = \mathrm{rank}\,\mathcal{G}$ be highest weight vectors for the fundamental representations of $\mathcal{G}$. Recall the main properties of $|\lambda^{(i)}>$:

$$H|\lambda^{(i)}>= \lambda^{(i)}(H)|\lambda^{(i)}>, \quad E_\alpha|\lambda^{(i)}>= 0 \;\; \text{for} \;\; \alpha > 0$$

This representation is obtained by taking the quotient of the universal enveloping algebra of $\mathcal{G}$ by the left ideal generated by the E_α and the $H - \lambda^{(i)}(H)1$. One can define on the universal enveloping algebra a $\star$ operation such that $(XY)^\star = Y^\star X^\star$ and $(\lambda X)^\star = \overline{\lambda}X^\star$. Indeed in a Cartan–Weyl basis, see Chapter [16], one can define $H^\star = H$, $E_\alpha^\star = E_{-\alpha}$ and this is compatible with the commutation relations because $C_{-\alpha,-\beta} = -C_{\alpha,\beta}$ and the $C_{\alpha,\beta}$ and the $\alpha(H)$ are real. In the highest weight representation, the operation $\star$ is just Hermitean conjugation. We denote $< \lambda^{(i)}|$ the conjugate highest weight which satisfies:

$$< \lambda^{(i)}| = \lambda^{(i)}(H) < \lambda^{(i)}|, \quad < \lambda^{(i)}|E_{-\alpha} = 0 \text{ for } \alpha > 0$$

From these rules one can compute any scalar product of the form $< \lambda^{(i)}|\mathcal{X}|\lambda^{(i)} >$ where $\mathcal{X}$ is an element of the universal enveloping algebra, by pushing all E_α to the right, all $E_{-\alpha}$ to the left using the commutation relations, and the fact that $|\lambda^{(i)} >$ is a common eigenvector of all H. Finally we normalize the scalar product by:

$$< \lambda^{(i)}|\lambda^{(i)} >= 1$$

Proposition. *For any fundamental representation, let $\lambda^{(i)}$ be its highest weight and define:*

$$\begin{aligned} \xi^{(i)} &= < \lambda^{(i)}|\, e^{-\Phi}\Psi = < \lambda^{(i)}|\, Q_+ \\ \overline{\xi}^{(i)} &= \Psi^{-1}e^{-\Phi}\,|\lambda^{(i)} > = Q_-^{-1}\,|\lambda^{(i)} > \end{aligned} \tag{11.13}$$

then the fields $\xi^{(i)}$ and $\overline{\xi}^{(i)}$ are chiral: $\partial_{x_-}\xi^{(i)} = 0$ and $\partial_{x_+}\overline{\xi}^{(i)} = 0$.
and the Toda field Φ can be recovered by the formula:

$$e^{-2\lambda^{(i)}(\Phi)} = \xi^{(i)} \cdot \overline{\xi}^{(i)} =< \lambda^{(i)}|Q_+Q_-^{-1}|\lambda^{(i)} > \tag{11.14}$$

Proof. First $e^{-\Phi}\Psi = N_-Q_+$ and $< \lambda^{(i)}|N_- =< \lambda^{(i)}|$ by our rules, so that $< \lambda^{(i)}|\, e^{-\Phi}\Psi =< \lambda^{(i)}|\, Q_+$ which depends only on x_+. Similarly $\Psi^{-1}e^{-\Phi}\,|\lambda^{(i)} >= Q_-^{-1}\,|\lambda^{(i)} >$ depends only on x_-. By the definition of $\xi^{(i)}$ and $\overline{\xi}^{(i)}$ we see that Ψ and Ψ^{-1} cancel in their scalar product, leaving $< \lambda^{(i)}|\, e^{-2\Phi}\,|\lambda^{(i)} >= e^{-2\lambda^{(i)}(\Phi)}$. The knowledge of these quantities for $i = 1, \cdots, r$ completely characterizes Φ. ■

Equations (11.9) determine $Q_\pm$ in terms of the two chiral fields P and $\overline{P}$ with values in $\mathcal{H}$. For example, consider the equation $\partial_{x_+}Q_+ = (P - \mathcal{E}_+)Q_+$ in the Liouville case $\mathcal{G} = sl_2$. We have:

$$Q_+ = \begin{pmatrix} q_{11} & q_{12} \\ 0 & q_{22} \end{pmatrix}, \quad (P - \mathcal{E}_+) = \begin{pmatrix} p & -1 \\ 0 & -p \end{pmatrix}, \quad \xi = (1, 0)Q_+ = (q_{11}, q_{12})$$

and the differential equation yields:

$$(\partial^2 - \mathcal{U})\xi = 0, \quad \mathcal{U} = p' + p^2$$

Hence the first order linear system for Q_+ is a matrix version of the Schroedinger equation, which is recovered as an equation for the first line of Q_+. The chiral fields P and $\overline{P}$ are the two arbitrary functions parametrizing the general solution of Toda field equations.

Let us describe more explicitly the case of $sl(n+1)$. The n fundamental representations are the vector representation and its wedge products. The vector representation acts on $\mathbb{C}^{n+1}$ with basis $|\epsilon_j\rangle$, $j = 1, \cdots, (n+1)$. The highest weight vector is $|\epsilon_1\rangle$. The elements of the Cartan algebra are the traceless diagonal matrices, and $\mathcal{E}_\pm = \sum E_{i,i\pm 1}$, where $E_{ij} = |\epsilon_i\rangle\langle\epsilon_j|$ are the canonical matrices acting on $\mathbb{C}^{n+1}$. In the vector representation the chiral fields are $\xi(x) = \langle\epsilon_1|Q_+$ and $\overline{\xi}(x) = Q_-^{-1}|\epsilon_1\rangle$. Let us decompose the fields P and $\overline{P}$ as $P = \sum_j P_jE_{jj}$ and $\overline{P} = \sum_j \overline{P}_jE_{jj}$, with $\sum P_j = \sum \overline{P}_j = 0$. Let ξ_j and $\overline{\xi}_j$ be the components of the chiral fields: $\xi(x) = \sum_j \langle\epsilon_j|\, \xi_j$, and $\overline{\xi}(x) = \sum_j \overline{\xi}_j|\epsilon_j\rangle$. The ξ_j's and $\overline{\xi}_j$' satisfy the order $(n + 1)$ differential equations:

$$\begin{aligned} (\partial_{x_-} + \overline{P}_{n+1})\cdots(\partial_{x_-} + \overline{P}_1)\, \overline{\xi}_j &= 0 \\ (\partial_{x_+} - P_{n+1})\cdots(\partial_{x_+} - P_1)\, \xi_j &= 0 \end{aligned} \tag{11.15}$$

Indeed, the first line of Q_+ is $(\xi_1, \cdots, \xi_{n+1})$ and the explicit form of $\mathcal{E}_+$ immediately yields the equation of order $(n + 1)$ [11.15].

Proposition. *The components of the Toda field along the fundamental weights $\lambda^{(p)}$ are given by:*

$$e^{-2\lambda^{(p)}(\Phi)} = \det\left(\partial_{x_+}^i \xi \cdot \partial_{x_-}^j \overline{\xi}\right), \quad i,j = 0, \cdots, p-1 \tag{11.16}$$

Proof. The highest weight vector of the $\lambda^{(p)}$ is the wedge product $|\lambda^{(p)}\rangle = |\epsilon_1\rangle \wedge \cdots \wedge |\epsilon_p\rangle$, and we have $\mathcal{E}_-^{p-1}|\epsilon_1\rangle = |\epsilon_p\rangle$ for $p = 1, \ldots, n$. Thus,

$$|\lambda^{(p)}\rangle = |\epsilon_1\rangle \wedge \mathcal{E}_-|\epsilon_1\rangle \wedge \cdots \wedge \mathcal{E}_-^{p-1}|\epsilon_1\rangle$$

Acting with Q_-^{-1} we obtain an expression for the chiral field $\overline{\xi}^{(p)}$:

$$\overline{\xi}^{(p)} = Q_-^{-1}|\lambda^{(p)}\rangle = Q_-^{-1}|\epsilon_1\rangle \wedge Q_-^{-1}\mathcal{E}_-|\epsilon_1\rangle \wedge \cdots \wedge Q_-^{-1}\mathcal{E}_-^{p-1}|\epsilon_1\rangle \tag{11.17}$$

We now use the equation of motion (11.9) to express $Q_-^{-1}\mathcal{E}_-^j|\epsilon_1\rangle$ in terms of the derivatives of the functions $\overline{f}_j$. Indeed using eq.(11.9) and differentiating it, we get:

$$\begin{aligned}
Q_-^{-1}\mathcal{E}_-|\epsilon_1\rangle &= \partial_{x_-} Q_-^{-1}|\epsilon_1\rangle + Q_-^{-1}\overline{P}|\epsilon_1\rangle \\
Q_-^{-1}\mathcal{E}_-^2|\epsilon_1\rangle &= \partial_{x_-}^2 Q_-^{-1}|\epsilon_1\rangle + Q_-^{-1}\left(\overline{P}\mathcal{E}_- + \mathcal{E}_-\overline{P} - \overline{P}^2 + \partial_{x_-}\overline{P}\right)|\epsilon_1\rangle \\
Q_-^{-1}\mathcal{E}_-^j|\epsilon_1\rangle &= \partial_{x_-}^j Q_-^{-1}|\epsilon_1\rangle + \cdots
\end{aligned}$$

The extra terms cancel in the wedge product (11.17). Therefore,

$$\begin{aligned}
\overline{\xi}^{(p)} &= Q_-^{-1}|\epsilon_1\rangle \wedge (\partial_{x_-}Q_-^{-1})|\epsilon_1\rangle \wedge \cdots \wedge (\partial_{x_-}^{p-1}Q_-^{-1})|\epsilon_1\rangle = \overline{\xi} \wedge \partial_{x_-}\overline{\xi} \wedge \cdots \wedge \partial_{x_-}^{p-1}\overline{\xi} \\
&= \sum_{j_1 < \cdots < j_p} |\epsilon_{j_1}\rangle \wedge \cdots \wedge |\epsilon_{j_p}\rangle \det \begin{pmatrix} \overline{\xi}_{j_1} & \cdots & \overline{\xi}_{j_p} \\ \partial_{x_-}\overline{\xi}_{j_1} & \cdots & \partial_{x_-}\overline{\xi}_{j_p} \\ \vdots & & \vdots \\ \partial_{x_-}^{p-1}\overline{\xi}_{j_1} & \cdots & \partial_{x_-}^{p-1}\overline{\xi}_{j_p} \end{pmatrix}
\end{aligned}$$

A similar expression for the chiral fields $\xi^{(p)} = \langle\lambda^{(p)}|Q_+ = \xi \wedge \partial_{x_+}\xi \wedge \cdots \partial_{x_+}^{p-1}\xi$ can be obtained. Equation (11.16) follows from $\exp(-2\lambda^{(p)}(\Phi)) = \xi^{(p)} \cdot \overline{\xi}^{(p)}$. Noting that $< \epsilon_{j_1}| \wedge \cdots \wedge < \epsilon_{j_p}|\epsilon_{k_1} > \wedge \cdots \wedge |\epsilon_{k_p} >= \epsilon_{j_1 \cdots j_p}^{k_1 \cdots k_p}$. we can also write

$$e^{-2\lambda^{(p)}(\Phi)} = \sum_{j_1 < \cdots < j_p} \det \begin{pmatrix} \xi_{j_1} & \cdots & \xi_{j_p} \\ \partial_{x_+}\xi_{j_1} & \cdots & \partial_{x_+}\xi_{j_p} \\ \vdots & & \vdots \\ \partial_{x_+}^{p-1}\xi_{j_1} & \cdots & \partial_{x_+}^{p-1}\xi_{j_p} \end{pmatrix} \det \begin{pmatrix} \overline{\xi}_{j_1} & \cdots & \overline{\xi}_{j_p} \\ \partial_{x_-}\overline{\xi}_{j_1} & \cdots & \partial_{x_-}\overline{\xi}_{j_p} \\ \vdots & & \vdots \\ \partial_{x_-}^{p-1}\overline{\xi}_{j_1} & \cdots & \partial_{x_-}^{p-1}\overline{\xi}_{j_p} \end{pmatrix}$$

This is the sl_{n+1} generalization of the Liouville solution. ∎

It is interesting to give a direct and algebraic proof that the formula we just obtained for Φ does indeed satisfy the Toda field equations of motion. Let us rewrite the equations of motion directly in terms of the quantities:

$$\tau_i(x_+, x_-) = \exp\Big(-2\lambda^{(i)}(\Phi) \Big) \tag{11.18}$$

Proposition. *In terms of the τ_i functions the equations of motion become:*

$$\tau_i(\partial_{x_-}\partial_{x_+}\tau_i) - (\partial_{x_+}\tau_i)(\partial_{x_-}\tau_i) = -\frac{(\alpha_i, \alpha_i)}{2}\left(\prod_{j \neq i} \tau_j^{-a_{ji}}\right) \tag{11.19}$$

where a_{ij} is the Cartan matrix of $\mathcal{G}$.

Proof. Recall the Toda equations (11.4) for the components $\varphi^i = \lambda^{(i)}(\Phi) = -\frac{1}{2}\log\tau_i$. We have $\exp\left(2\sum_j \varphi^j\, a_{ji}\right) = \prod_j \tau_j^{-a_{ji}}$. Eq.(11.19) for the τ-functions then follows using $a_{ii}=2$ and

$$\partial_{x_+}\partial_{x_-}\log\tau_i = \frac{1}{\tau_i^2}\;\left(\tau_i\,\partial_{x_-}\partial_{x_+}\tau_i - \partial_{x_+}\tau_i\,\partial_{x_-}\tau_i\right).$$

∎

Remark. Using Hirota's differential operators we can rewrite eq.(11.19) as:

$$D_+D_-\tau_i\cdot\tau_i = -(\alpha_i,\alpha_i)\prod_{j\neq i}\tau_j^{-a_{ji}}$$

In general $0 \le a_{ij}a_{ji} \le 4$ and $a_{ij} \le 0$ for $j \neq i$ so the right–hand side of this equation may be a polynomial of degree greater than 2 in the τ_j. However, in the sl_{n+1} case, we have $a_{ij} = -1$ for $j = i \pm 1$ (the other a_{ij} vanish) so this equation takes a bilinear form:

$$D_+D_-\tau_i\cdot\tau_i = -2\tau_{i-1}\tau_{i+1}$$

We now show that equations (11.19) can be derived directly from eqs.(11.9,11.8,11.14). By definition of the tau functions we have: $\tau_i =< \lambda^{(i)}|Q_+Q_-^{-1}|\lambda^{(i)} >$. Let us set $G = Q_+Q_-^{-1}$. Computing the left hand side of (11.19) we have:

$$\tau_i(\partial_{x_-}\partial_{x_+}\tau_i) - (\partial_{x_+}\tau_i)(\partial_{x_-}\tau_i) =< \lambda^{(i)}|\otimes < \lambda^{(i)}|G\otimes\partial_{x_-}\partial_{x_+}G - \partial_{x_+}G\otimes\partial_{x_-}G|\lambda^{(i)} > \otimes|\lambda^{(i)} >$$

Using the equations of motion of $Q_\pm$, we see that the fields P and $\overline{P}$ drop out because they always hit the highest weight vector giving the scalar factors $\lambda^{(i)}(P)$ or $\lambda^{(i)}(\overline{P})$, and we get:

$$\begin{aligned}\tau_i(\partial_{x_-}\partial_{x_+}\tau_i) - (\partial_{x_+}\tau_i)(\partial_{x_-}\tau_i) = &- < \lambda^{(i)}|G|\lambda^{(i)} >< \lambda^{(i)}|\mathcal{E}_+G\mathcal{E}_-|\lambda^{(i)} >\\ &+ < \lambda^{(i)}|\mathcal{E}_+G|\lambda^{(i)} >< \lambda^{(i)}|G\mathcal{E}_-|\lambda^{(i)} >\end{aligned}$$

This can be rewritten as

$$\tau_i(\partial_{x_-}\partial_{x_+}\tau_i) - (\partial_{x_+}\tau_i)(\partial_{x_-}\tau_i) = -\tfrac{1}{2} < \Psi^{(i)}|\;G\otimes G\;|\Psi^{(i)} > \tag{11.20}$$

where $|\Psi^{(i)} >= |\lambda^{(i)} > \otimes\mathcal{E}_-|\lambda^{(i)} > -\mathcal{E}_-|\lambda^{(i)} > \otimes|\lambda^{(i)} >= |\lambda^{(i)} > \wedge\,\mathcal{E}_-|\lambda^{(i)} >$. The proof of the equations of motion (11.19) now consists in comparing the right hand side of eqs.(11.19) and (11.20). This follows from the following lemma:

Lemma. *In the tensor product of two copies of the fundamental representation with highest weight $\lambda^{(i)}$, $|\Psi^{(i)} >$ is a highest weight vector with weight $-\sum_{j\neq i} a_{ji}\lambda^{(j)}$ and of norm $< \Psi^{(i)}|\Psi^{(i)} >= (\alpha_i,\alpha_i)$.*

Proof. We must show that E_α kills $|\Psi^{(i)} >$ for all positive roots α. It is enough to consider α simple. In the tensor product, E_α is represented by $E_\alpha\otimes I + I\otimes E_\alpha$. Hence we have to show that:

$$(E_\alpha\otimes I + I\otimes E_\alpha)|\Psi^{(i)} >= |\lambda^{(i)} > \otimes E_\alpha\mathcal{E}_-|\lambda^{(i)} > -E_\alpha\mathcal{E}_-|\lambda^{(i)} > \otimes|\lambda^{(i)} >= 0$$

But for α a simple positive root, $E_\alpha\mathcal{E}_-|\lambda^{(i)} >= \sum_{\beta\ \mathrm{simple}}[E_\alpha, E_{-\beta}]|\lambda^{(i)} >= H_\alpha|\lambda^{(i)} >= \lambda^{(i)}(H_\alpha)|\lambda^{(i)} >$, where we have used that for α β simple, $[E_\alpha, E_{-\beta}] = 0$ for $\beta\neq\alpha$. The first part of the lemma follows. For the second part, we must prove that $|\Psi^{(i)}>$ is an eigenstate of $(H_\alpha \otimes I + I \otimes H_\alpha)$ and find the eigenvalue. Let us compute:

$$\begin{aligned}(H_\alpha \otimes I + I \otimes H_\alpha)|\Psi^{(i)}> &= H_\alpha|\lambda^{(i)}> \otimes \mathcal{E}_-|\lambda^{(i)}> - H_\alpha\mathcal{E}_-|\lambda^{(i)}> \otimes |\lambda^{(i)}> \\ &\quad + |\lambda^{(i)}> \otimes H_\alpha\mathcal{E}_-|\lambda^{(i)}> - \mathcal{E}_-|\lambda^{(i)}> \otimes H_\alpha|\lambda^{(i)}>\end{aligned}$$

Next we notice that $\mathcal{E}_-|\lambda^{(i)}>= E_{-\alpha_i}|\lambda^{(i)}>$, because the norm of of $E_{-\beta}|\lambda^{(i)}>$ is equal to

$$<\lambda^{(i)}|E_\beta E_{-\beta}|\lambda^{(i)}>= \lambda^{(i)}(H_\beta)$$

and vanishes if $\beta \neq \alpha_i$ by definition of the fundamental weights. This implies $H_\alpha\mathcal{E}_-|\lambda^{(i)}\rangle = \left(-\alpha_i(H_\alpha) + \lambda^{(i)}(H_\alpha)\right) E_{-\alpha_i}|\lambda^{(i)}\rangle$, which gives:

$$(H_\alpha \otimes I + I \otimes H_\alpha)|\Psi^{(i)}>= [2\lambda^{(i)}(H_\alpha) - \alpha_i(H_\alpha)]|\Psi^{(i)}>$$

Thus the weight of $|\Psi^{(i)}>$ is $2\lambda^{(i)} - \alpha_i = -\sum_{j\neq i} a_{ji}\ \lambda^{(j)}$ because $\alpha_i = \sum_j \lambda^{(j)} a_{ji}$. Finally we have: $<\Psi^{(i)}|\Psi^{(i)}>= \lambda^{(i)}(H_{\alpha_i}) = (\alpha_i, \alpha_i)$. ■

We have to compute $<\Psi^{(i)}|\ G \otimes G\ |\Psi^{(i)}>$. Here $G \otimes G$ is the representation of G in the tensor product representation. Since this scalar product only depends on the values of the weights and the Lie algebra structure, we can compute it as well using the highest weight vector with the same weight and normalization:

$$|\Psi^{(i)}>= (\alpha_i, \alpha_i)^{\frac{1}{2}} \bigotimes_{j\neq i} \left(|\lambda^{(j)}>\right)^{\otimes(-a_{ij})}$$

which yields eq.(11.18).

11.4 Hamiltonian formalism.

The Toda field equations can be derived from the Lagrangian density:

$$\mathcal{L} = (\partial^\mu\Phi, \partial_\mu\Phi) - \sum_{\alpha_i\ simple} \exp(2\alpha(\Phi))$$

As usual we can introduce the canonical equal–time Poisson brackets between φ and its conjugate momentum π. In the following we compute Poisson brackets of the dynamical variables $\pi(x)$ and $\varphi(y)$ at $t = 0$.

$$\{\pi_i(x), \varphi_j(y)\} = \delta_{ij}\delta(x - y)$$

We define $\Pi(x) = \sum_i \pi_i(x)H_i$. Replacing $\partial_t\Phi(x)$ by $\Pi(x)$ in the expression of $A = A_{x_+} + A_{x_-}$, we get:

$$A = \Pi + \sum_{\alpha \text{ simple}} \exp(\alpha(\Phi))(E_\alpha + E_{-\alpha})$$

The Poisson brackets of A takes the familiar r–matrix form, whose proof is exactly the same as for the Toda chain in Chapter [3].

Proposition. *There exists a matrix $r \in \mathcal{G} \otimes \mathcal{G}$, independent of the fields π and φ, such that*

$$\{A_1(x), A_2(y)\} = [r_{12}, A_1(x) + A_2(y)]\delta(x - y) \tag{11.21}$$

with

$$r_{12} = \frac{1}{2} \sum_{\alpha \text{ positive}} \frac{E_\alpha \otimes E_{-\alpha} - E_{-\alpha} \otimes E_\alpha}{(E_\alpha, E_{-\alpha})} + \lambda \cdot C_{12} \tag{11.22}$$

where C_{12} the tensor Casimir:

$$C_{12} = \sum_i H_i \otimes H_i + \sum_{\alpha \text{ positive}} \frac{E_\alpha \otimes E_{-\alpha} + E_{-\alpha} \otimes E_\alpha}{(E_\alpha, E_{-\alpha})} \tag{11.23}$$

Recall that one can find two values of the parameter λ multiplying the tensor Casimir such that the corresponding matrices $r^\pm$ are solutions of the classical Yang-Baxter equation

$$[r^\pm_{12}, r^\pm_{13}] + [r^\pm_{12}, r^\pm_{23}] + [r^\pm_{13}, r^\pm_{23}] = 0$$

These two solutions are:

$$r^+ = \frac{1}{2} \sum_i H_i \otimes H_i + \sum_{\alpha \text{ positive}} \frac{E_\alpha \otimes E_{-\alpha}}{(E_\alpha, E_{-\alpha})} \tag{11.24}$$

$$r^- = -\frac{1}{2} \sum_i H_i \otimes H_i - \sum_{\alpha \text{ positive}} \frac{E_{-\alpha} \otimes E_\alpha}{(E_\alpha, E_{-\alpha})} \tag{11.25}$$

Recall that the wave function $\Psi(x)$ is the solution of the linear system $(\partial_x + A_x)\,\Psi(x) = 0$, normalized with the boundary condition: $\Psi(0) = 1$. Hence $\Psi(x)$ can be written as a path ordered exponential, and its Poisson brackets can be computed explicitly, as explained in Chapter [2].

Proposition. *The wave function $\Psi(x)$ satisfies the quadratic bracket:*

$$\{\Psi_1(x), \Psi_2(x)\} = -[r^\pm_{12}, \Psi_1(x)\Psi_2(x)] \tag{11.26}$$

We can use either r^+_{12} or r^-_{12}, since the difference is the tensor Casimir C_{12} which commutes with $\Psi(x)\otimes\Psi(x)$.

To each highest weight vector $|\lambda^{(r)}>$ we associate two chiral vectors $\xi^{(r)}$ and $\overline{\xi}^{(r)}$ by the formula (11.13). The Poisson brackets between these vectors can be computed and obey what is called an exchange algebra. Since we compute the Poisson brackets at equal time, $t = 0$ we do not distinguish the notations $x_+ = x_- = x$.

Proposition. *The exchange algebra of the chiral fields reads:*

$$\begin{aligned}
\{\xi^{(r)}_1(x), \xi^{(r')}_2(y)\} &= \xi^{(r)}_1(x)\xi^{(r')}_2(y)\; r^\pm_{12} && \text{for } x \gtrless y \\
\{\overline{\xi}^{(r)}_1(x), \overline{\xi}^{(r')}_2(y)\} &= r^\mp_{12}\; \overline{\xi}^{(r)}_1(x)\overline{\xi}^{(r')}_2(y) && \text{for } x \gtrless y \\
\{\xi^{(r)}_1(x), \overline{\xi}^{(r')}_2(y)\} &= -\xi^{(r)}_1(x) \cdot r^-_{12} \cdot \overline{\xi}^{(r')}_2(y) && \text{for } x \neq y \\
\{\overline{\xi}^{(r)}_1(x), \xi^{(r')}_2(y)\} &= -\xi^{(r')}_2(y) \cdot r^+_{12} \cdot \overline{\xi}^{(r)}_1(x) && \text{for } x \neq y
\end{aligned} \tag{11.27}$$

Proof. We have:

$$\begin{aligned}
\{\xi^{(r)}_1(x), \xi^{(r')}_2(y)\} &= <\lambda^{(r)}| \otimes <\lambda^{(r')}| \{e^{-\Phi_1(x)}\Psi_1(x), e^{-\Phi_2(y)}\Psi_2(y)\} \\
&= <\lambda^{(r)}| \otimes <\lambda^{(r')}| \exp[-\Phi_1(x)] \exp[-\Phi_2(y)] \cdot \\
&\quad \cdot \Big[\{\Psi_1(x), \Psi_2(y)\} - \{\Phi_1(x), \Psi_2(y)\}\Psi_1(x) - \{\Psi_1(x), \Phi_2(y)\}\Psi_2(y)\Big]
\end{aligned}$$

Thus we need to evaluate the Poisson brackets of the wave function at different points and the Poisson brackets between the wave function $\Psi(x)$ and the Toda field $\Phi(y)$. We first compute the latter. We have:

$$\{\Phi_1(x), \Psi_2(y)\} = \theta(y - x)\; \Psi_2(y)\Psi_2^{-1}(x)\; C^0_{12}\; \Psi_2(x) \tag{11.28}$$

where $C^0_{12} = \frac{1}{2}\sum_i H_i \otimes H_i$, and $\theta(x - y)$ is the step function. Indeed, from the defining equation of $\Psi(x)$ we have:

$$[\partial_y + A_2(y)]\{\Phi_1(x), \Psi_2(y)\} = -\{\Phi_1(x), A_2(y)\}\Psi_2(y)$$

but $\{\Phi_1(x), A_2(y)\} = -C^0_{12}\,\delta(x-y)$. The result follows by integrating this differential equation and imposing the boundary condition expressing the ultralocality of the model $\{\Phi_1(x), \Psi_2(y)\} = 0$ if $x > y$.
Let us now compute the Poisson brackets of the wave function at different points. For $x > y$ we can write $\Psi(x) = \Psi(x,y)\Psi(y)$, where $\Psi(x,y)$ is the transport matrix from y to x. Using the ultralocality property of the Poisson bracket we have, $\{\Psi_1(x), \Psi_2(y)\} = \Psi_1(x,y)\{\Psi_1(y), \Psi_2(y)\}$. Equation (11.26) yields:

$$\{\Psi_1(x), \Psi_2(y)\} = \Psi_1(x)\Psi_2(y)[r_{12} - \Psi_1^{-1}(y)\Psi_2^{-1}(y)r_{12}\Psi_1(y)\Psi_2(y)] \tag{11.29}$$

Using eqs.(11.28,11.29) we finally evaluate the Poisson brackets between the chiral fields as:

$$\begin{aligned}\{\xi^{(r)}_1(x), \xi^{(r')}_2(y)\} &= <\lambda^{(r)}|\otimes<\lambda^{(r')}|\exp[-\Phi_1(x)]\exp[-\Phi_2(y)]\,\cdot \\ &\quad\cdot[\Psi_1(x)\Psi_2(y)\cdot r_{12} - \Psi_1(x)\Psi_1^{-1}(y)\cdot[r_{12} - C^0_{12}].\Psi_1(y)\Psi_2(y)]\end{aligned}$$

In this formula, one can take $r_{12} = r^+_{12}$ and we have $1\otimes<\lambda^{(r')}|r^+_{12} = 1\otimes<\lambda^{(r')}|C^0_{12}$. We see that the last term vanishes. This proves the first equation for $x > y$. The other cases are proved similarly. ■

11.5 Conformal structure.

The solutions of Toda field theory were parametrized by two chiral fields P and $\overline{P}$. The transformation which relates the original fields to these chiral fields is highly non local. There exists however remarkable quantities which are local in terms of both sets of fields. The purpose of this section is to describe them and to find their Poisson bracket algebra. As we will see this algebra contains the Virasoro algebra as a subalgebra and is therefore the conformal symmetry algebra of the theory.

We consider only one chirality, and recall the linear systems eq.(11.9) and eq.(11.10):

$$[\partial_{x_+} - P + \mathcal{E}_+]Q_+ = 0 \tag{11.30}$$

$$[\partial_{x_+} + 2\partial_{x_+}\Phi + \mathcal{E}_+]N_-Q_+ = 0 \tag{11.31}$$

These equations have exactly the same structure, but one is expressed in terms of P while the other is expressed in terms of the original fields of the theory. The relation between the two equations is a gauge transformation $N_- \in \mathcal{N}_-$. Notice that the fields $\xi^{(r)}(x)$ are invariant under this gauge transformation because $\langle\lambda|N_- = \langle\lambda|$. For finite dimensional highest weight representations of $\mathcal{G}$ one can deduce from the first order linear systems eq.(11.30) and eq.(11.31) a single differential equation of higher order for the components of vector $\xi^{(r)}(x)$,

$$L\cdot\xi = 0 \tag{11.32}$$

Since the components of $\xi^{(r)}(x)$ are invariant under the gauge transformation N_-, the coefficients of this equation will also be invariant. Moreover these coefficients are local differential polynomials in terms of P or Φ.

In the following, we shall restrict ourselves to the $sl(n+1)$ case in the fundamental representation. In this representation, the vector $\xi(x)$ has $(n+1)$ components, ie. $\xi(x) = (\xi_i(x))$, $i = 1, \cdots, (n+1)$ and $\mathcal{E}_+ = \sum_i E_{i,i+1}$. where E_{ij} is the canonical basis of $(n+1)\times(n+1)$ matrices, finally $P = \sum_i P_i E_{ii}$ with $\sum_i P_i = 0$.

Using eq.(11.30) or eq.(11.31) one can write the operator L in eq.(11.32) explicitly:

$$L = \partial^{n+1} + \sum_{i=0}^{n-1} u_i\partial^i, \quad L = (\partial - P_{n+1})\cdots(\partial - P_1) \;\text{ with }\; \sum P_i = 0$$

or a similar formula with $P \to 2\partial_{x_+}\Phi$. It is easy to compute $u_{n-1} = -\mathcal{U}$:

$$\mathcal{U} = \tfrac{1}{2}(P,P) + (H_\rho, P') = 2(\partial_{x_+}\Phi, \partial_{x_+}\Phi) - 2(H_\rho, \partial^2_{x_+}\Phi) \tag{11.33}$$

where H_ρ is the element in $\mathcal{H}$ such that $[H_\rho, \mathcal{E}_+] = \mathcal{E}_+$, namely in our case $H_\rho = \mathrm{Diag}(\frac{n}{2}, \frac{n}{2}-1, \cdots, -\frac{n}{2}$, yielding $(H_\rho, P') = nP'_1 + (n-1)P'_2 + \cdots + P'_n$. Using the Toda equations of motion to eliminate the higher order time derivatives of Φ we also have $\mathcal{U} = \frac{1}{2}(\mathcal{H} + \mathcal{P})$ where:

$$\mathcal{H} = \frac{1}{2}(\Pi,\Pi) + \frac{1}{2}(\Phi_x, \Phi_x) + \sum_{\alpha \text{ simple}} e^{2\alpha(\Phi)} - 2(H_\rho, \Phi_{xx}), \quad \mathcal{P} = (\Pi, \Phi_x) - 2(H_\rho, \Pi_x)$$

are the energy and momentum densities respectively.

The function $\mathcal{U}$ in eq.(11.33) is the generalization of the function $\mathcal{U}$ introduced in the Liouville theory. We now compute its Poisson brackets.

Proposition. *Let $C^0_{12} = \frac{1}{2}\sum_i H_i \otimes H_i$. Then*

$$\begin{aligned} \{P_1(x), P_2(y)\} &= -(\partial_x - \partial_y)\delta(x-y)C^0_{12} \\ \{\mathcal{U}(x), \mathcal{U}(y)\} &= -\left(\mathcal{U}\partial_x + \partial_x\mathcal{U} + (H_\rho, H_\rho)\partial_x^3\right)\delta(x-y) \end{aligned} \tag{11.34}$$

Here $(H_\rho, H_\rho) = n(n+1)(n+2)/12$ and we recognize the standard Virasoro algebra, see Chapter [9].

Proof. We have to compute the Poisson brackets of P. To do this we start from the exchange algebra, and remark that the scalar product $\xi^{(r)}(x)|\lambda^{(r)} >$ satisfies:

$$\{\log \xi^{(r)}(x)|\lambda^{(r)} >, \log \xi^{(r')}(y)|\lambda^{(r')} >\} = \frac{1}{2}\sum_i \lambda^{(r)}(H_i)\lambda^{(r')}(H_i)\epsilon(x-y) \tag{11.35}$$

Note also that eq.(11.30) implies that

$$\xi^{(r)}(x)|\lambda^{(r)} >= e^{-\int^x \lambda^{(r)}(P)}\theta$$

where θ is some integration constant. Eq.(11.34) follows by differentiating eq.(11.35) with respect to x and y and using the independence of the fundamental weights. The Poisson bracket for $\mathcal{U}$ then follows readily, either from the expression of $\mathcal{U}$ in terms of P or in terms of Φ and Π. ■

We now explore the Poisson algebra of the invariant coefficients u_i. The quantity u_{n-1} is the generator of the conformal symmetry underlying the theory. Comparing with Chapter [9] we know that the Poisson bracket algebra of the u's coincides with the second Hamiltonian structure of the generalized KdV equation. We are going to rederive this result starting from the exchange algebra eq.(11.27). In the course this computation we will exhibit the conformal properties of the fields ξ_i.

Any component $\xi_i(x)$, $i = 1, \cdots, n+1$, of the vector $\xi(x)$ obeys the differential equation $L \cdot \xi_i = 0$ with L a differential operator of order $n+1$. The differential operator L can be written knowing the functions $\xi_i(x)$ by the formula:

$$L \cdot \xi = \det \begin{pmatrix} \xi & \xi_1 & \cdots & \xi_{n+1} \\ \xi' & \xi'_1 & \cdots & \xi'_{n+1} \\ \vdots & \vdots & & \vdots \\ \xi^{(n+1)} & \xi_1^{(n+1)} & \cdots & \xi_{n+1}^{(n+1)} \end{pmatrix} = 0 \tag{11.36}$$

Since the coefficients u_i of L are given by Wronskians of the ξ_i's, we can calculate directly their Poisson brackets knowing those of the ξ_i's. To express conveniently the result we need to recall some notations about pseudo differential operators, see Chapter [9]. We have introduced the derivation symbol $\partial = \partial_x$, with the usual Leibnitz rule $\partial.a = a.\partial + (\partial a)$ where (∂a) means $\partial_x a(x)$, and the integration symbol ∂^{-1} with the following rules

$$\partial^{-1}\partial = \partial\partial^{-1} = 1, \quad \partial^{-i-1} f = \sum_{v=0}^{\infty} (-1)^v \begin{pmatrix} i+v \\ v \end{pmatrix} (\partial^v f) \partial^{-i-1-v} \tag{11.37}$$

The elements $A = \sum_{-\infty}^{N} a_i \partial^i$ form an associative algebra with unit, called the algebra of formal pseudo differential operators in one variable. It is equipped with a linear form satisfying the fundamental trace property $< AB >=< BA >$, called the Adler trace and defined by:

$$< A >= \int dx a_{-1}$$

With these notations we have:

Proposition. *Let* $X = \sum_i \partial^{-i-1} X_i$ *and let* $f_X(L) =< LX >= \sum_i \int_0^\ell dx X_i(x) u_i(x)$. *Then one has:*

$$\{f_X(L), \xi_i(x)\} = -\left((XL)_+ + \frac{1}{n+1} \partial^{-1}([X, L]_{-1}) \right) \xi_i(x) \tag{11.38}$$

Proof. Using eq.(11.37) we find:

$$(XL)_+ = -\sum_{k>i\geq 0} \sum_{s=0}^{k-i-1} (-1)^{k-i-s} \begin{pmatrix} k-1-s \\ i \end{pmatrix} \partial^{k-i-s-1} (X_i u_k) \partial^s$$

and

$$[X, L]_{-1} = \sum_{k>i\geq 0} (-1)^{k+i} \begin{pmatrix} k \\ i \end{pmatrix} \partial^{k-i} (X_i u_k)$$

So we have to prove that:

$$\begin{aligned} \{f_X(L), \xi_q(x)\} \quad &= \quad -\frac{1}{(n+1)} \sum_{k>i\geq 0} (-1)^{k+i} \begin{pmatrix} k \\ i \end{pmatrix} \partial^{k-i-1} (X_i u_k) \xi_q(x) \\ &+ \sum_{k>i\geq 0} \sum_{s=0}^{k-i-1} (-1)^{k+i+s} \begin{pmatrix} k-s-1 \\ i \end{pmatrix} \partial^{k-i-s-1} (X_i u_k) \partial^s \xi_q \end{aligned} \tag{11.39}$$

From eq(11.36) the coefficients u_i in the operator L are given by $u_i = (-1)^i \det M_i$ where M_i is the matrix of elements

$$(M_i)_{ab} = \partial^a \xi_b \qquad \text{with} \quad \left\{ \begin{array}{lcl} a & = & 0, 1, \ldots n+1, \ a \neq i \\ b & = & 1, \ldots n+1 \end{array} \right.$$

Let $[\Delta_i(x)]_{kl}$ with $k = 0, \cdots n+1, \ k \neq i$ and $l = 1 \cdots n+1$, be the minors of the matrix M_i. By using the Leibnitz rules for the Poisson brackets, we obtain

$$\{f_X(L), \xi_q(x)\} = \int dz X_i(z) (-1)^i [\Delta_i(z)]_{kl} \partial_z^k \{\xi_l(z), \xi_q(x)\} \tag{11.40}$$

where all repeated indices are summed over. The Poisson brackets of the chiral fields were obtained in eq(11.27):

$$\{\xi_1(z), \xi_2(x)\} = \xi_1(z)\xi_2(x) r_{12}^- + \theta(z-x) \xi_1(z)\xi_2(x)(r_{12}^+ - r_{12}^-)$$

There are thus two contributions to eq.(11.40). The contribution of the term independent of $\theta(z-x)$ is

$$\int dz X_i(z) (-1)^i [\Delta_i(z)]_{kl} \partial^k \xi_m(z) \xi_n(x) r_{mn,lq}^-$$

Remembering that $\sum_{k\neq i}[\Delta_i(z)]_{kl}\partial^k\xi_m(z) = (\det M_i)\,\delta_{lm}$ since $[\Delta_i(z)]_{kl}$ are the minors of M_i, we see that the above expression vanishes since the r-matrices are such that: $\mathrm{Tr}_1 r^- = \mathrm{Tr}_2 r^- = \mathrm{Tr}_{12} r^- = 0$.
Calculating the $\theta(z-x)$ dependent term, we get

$$\{f_X(L), \xi_q(x)\} \quad = \quad \int dz \sum_{a\neq k} \begin{pmatrix} k \\ a \end{pmatrix} X_i(z)(-1)^i[\Delta_i(z)]_{kl}\partial_z^a\xi_m(z)\xi_n(x)\cdot(r^+ - r^-)_{mn,lq}\partial_z^{k-a}\theta(z-x)$$

The term $a = k$ in the sum may be excluded again due to the trace properties of the matrices $r^\pm$. To evaluate further this expression we remark that $r^+ - r^- = C$, where C is the Casimir element. To evaluate C for two vector representations of sl_{n+1} we first compute it in gl_{n+1} according to eq.(11.23). Here the E_α are the e_{ab} for $a < b$, and $E_{-\alpha} = e_{ba}$, while $H_i = e_{ii}$ so we get for C the permutation operator $P = \sum e_{ab}\otimes e_{ba}$, or $P_{ij,kl} = \delta_{il}\delta_{jk}$. We restrict this to sl_{n+1} by requiring that the partial traces of C vanish, yielding $C = -\frac{1}{n+1}1 + P$. Using this result we finally write the Poisson brackets (11.40) as:

$$\{f_X(A), \xi_q(x)\} = -\frac{1}{(n+1)}U_q(x) + V_q(x)$$

where $U_q(x)$ comes from the identity factor in C and $V_q(x)$ from the factor P.

$$U_q(x) \quad = \quad \int dz X_i(z)(-1)^i[\Delta_i(z)]_{kl}\sum_{a\neq k}\begin{pmatrix} k \\ a \end{pmatrix}\partial_z^a\xi_l(z)\xi_q(x)\partial_z^{k-a-1}\delta(z-x)$$

$$V_q(x) \quad = \quad \int dz X_i(z)(-1)^i[\Delta_i(z)]_{kl}\sum_{a\neq k}\begin{pmatrix} k \\ a \end{pmatrix}\partial_z^a\xi_q(z)\xi_l(x)\partial_z^{k-a-1}\delta(z-x)$$

The $U_q(x)$ term is easy to deal with. Noticing that

$$\sum_l [\Delta_i(z)]_{kl}\partial^a\xi_l(z) = (\det M_i)\,\delta_{ka} + (-1)^{k+i+1}(\det M_k)\,\delta_{ia} \quad \text{for} \quad \begin{cases} k\neq i \\ a\leq n+1 \end{cases} \tag{11.41}$$

we get

$$U_q(x) = \sum_{k>i}(-1)^{k+i}\begin{pmatrix} k \\ i \end{pmatrix}\partial^{k-i-1}(X_i u_k)\xi_q(x) \tag{11.42}$$

After integrating over z the expression of $V_q(x)$ becomes

$$V_q(x) = \sum_{a\neq k}(-1)^{i+k+a+1}\begin{pmatrix} k \\ a \end{pmatrix}\partial^{k-a-1}(X_i[\Delta_i]_{kl}\partial^a\xi_q)\xi_l(x)$$

The identity $(\partial^c A)B = \sum_{b=0}^c(-1)^b\begin{pmatrix} c \\ d \end{pmatrix}\partial^{c-b}(A\partial^b B)$ allows to rewrite this as

$$V_q(x) = \sum_{a\neq k}\sum_{b=0}^{k-a-1}(-1)^{k+a+i+b+1}\begin{pmatrix} k \\ a \end{pmatrix}\begin{pmatrix} k-a-1 \\ b \end{pmatrix}\partial^{k-a-b-1}(X_i[\Delta_i]_{kl}\partial^a\xi_q\partial^b\xi_l)$$

Suming over l using eq.(11.41) and performing the derivatives we obtain

$$V_q(x) = \sum_{s\geq 0}(-1)^{k+i}\begin{pmatrix} k-1-s \\ i \end{pmatrix}\left\{\sum_{a<k}(-1)^a\begin{pmatrix} k \\ a \end{pmatrix}\begin{pmatrix} k-1-a \\ k-1-s \end{pmatrix}\right\}\partial^{k-i-1-s}(X_i u_k)\partial^s\xi_q$$

The sum over a in the curly bracket is equal to $(-1)^s$ so that finally

$$V_q(x) = \sum_{s\geq 0}(-1)^{k+i+s}\begin{pmatrix} k-s-1 \\ i \end{pmatrix}\partial^{k-i-s-1}(X_i u_k)\partial^s\xi_q$$

which ends the proof. ∎

It is now very easy to derive the Poisson bracket algebra of the coefficient u_k

Proposition. *The Poisson brackets of the u_k reproduce the second Poisson bracket structure of the KdV hierarchy:*

$$\{f_X(L), f_Y(L)\} =< (XL)_+(YL)_- > - < (LX)_+(LY)_- > + \frac{1}{n+1}\int dx(\partial^{-1}[X,L]_{-1})[Y,L]_{-1} \qquad (11.43)$$

Proof. We start from $L\xi_i = 0$, so that $\{f_X(L), L\}\xi_i + L\{f_X(L), \xi_i\} = 0$ or equivalently

$$\{f_X(L), L\}\xi_i - L\left((XL)_+ + \frac{1}{n+1}\partial^{-1}[X,L]_{-1}\right)\xi_i = 0$$

Next, using again the differential equation for ξ, we rewrite this equation as:

$$\left\{\{f_X(L), L\} - \left(L(XL)_+ - (LX)_+L + \frac{1}{n+1}(L\partial^{-1}[X,L]_{-1} - \partial^{-1}[X,L]_{-1}L)\right)\right\}\xi_i = 0 \qquad (11.44)$$

Because $L(XL)_+ - (LX)_+L = (LX)_-L - L(XL)_-$, the differential operator in the right hand side of eq.(11.44) is of order at most n. Since it annihilates the $(n+1)$ linearly independent functions ξ_i, it is identically zero. The result then follows by multiplying eq.(11.44) by Y and taking the Adler trace. ∎

Remark. Let us apply eq.(11.38) with $X = \partial^{-n}X_{n-1}$, so that $f_X(L) = \int_0^{2\pi} dx X_{n-1}(x)u_{n-1}(x)$. We then get

$$\{f_X(L), \xi_i(x)\} = \left(-X_{n-1}\,\partial + \Delta^{(1)}\,X'_{n+1}\right)\xi^{(\lambda)}(x)$$

with $\Delta^{(1)} = \frac{n}{2} = \lambda^{(1)}(H_\rho)$ where $\lambda^{(1)}$ is the highest weight of the vector representation. Using $\xi^{(r)} = \xi^{(1)} \wedge \cdots \wedge \partial^{(r-1)}\xi^{(1)}$, we see that $\xi^{(r)}$ is also a primary field with conformal weight $\Delta^{(r)} = \lambda^{(r)}(H_\rho) = (r-1)(n-r)/2$, because each $\xi^{(1)}$ has conformal weight $n/2$ and each ∂ has weight -1.

11.6 Dressing transformations.

We now describe the group of dressing transformations in Toda field theories. In Chapter [2], dressing transformations were introduced as special gauge transformations preserving the form of the Lax connexion. On the wave function they read:

$$\Psi \quad\rightarrow\quad \Psi^g = \Theta_+\Psi g_+^{-1} = \Theta_-\Psi g_-^{-1} \quad \text{with} \quad \Theta_-^{-1}\Theta_+ = \Psi g_-^{-1}g_+\Psi^{-1} \qquad (11.45)$$

Here g is an element of G and $g_\pm$, $\Theta_\pm$ are obtained by solving a suitable factorization problem in G, which we now explain. The last condition in eq.(11.45) ensures that the two gauge transformations by $\Theta_\pm$ produce the same gauge transformed connection. We use this condition to determine $\Theta_\pm$ by a factorization problem in $G = \exp\mathcal{G}$. Of course the right factors $g_\pm$ play no role in the transformation of the connection, but are essential to understand the structure of the dressing group.

Proposition. *Let $g \in G = \exp\mathcal{G}$. Define a factorization problem*

$$g = g_-^{-1}g_+ \qquad (11.46)$$

by requiring that $g_\pm \in \exp\mathcal{B}_\pm$, and that the components of g_+ and g_- on $\exp(\mathcal{H})$ are inverse to each other. Then the gauge transformations $\Theta_\pm$ in eq.(11.45) preserve the form of the Toda connection.

Proof. We want that the gauge transformed connection $A^g_{x_\pm} \equiv -(\partial_{x_\pm}\Psi^g)\,\Psi^{g\,-1}$ be of the form:

$$A^g_{x_\pm} = \pm\partial_{x_\pm}\Phi^g + e^{\pm\mathrm{ad}\,\Phi^g}\mathcal{E}_\pm$$

for some field $\Phi^g \in \mathcal{H}$. In this case, the gauge transformation acts on the Toda field Φ. We have, for $\mu = x_+$ or $\mu = x_-$:

$$A^g_\mu = \Theta_\pm \partial_\mu \Theta_\pm^{-1} + \Theta_\pm A_\mu \Theta_\pm^{-1} \quad \text{with} \quad A_{x_\pm} = \pm\partial_{x_\pm}\Phi + e^{\pm \mathrm{ad}\,\Phi}\mathcal{E}_\pm$$

First, one has to see that $A^g_{x_+}$ decomposes on elements of height $(0,1)$ of the Lie algebra $\mathcal{G}$, and A_{x_-} on elements of height $(0,-1)$. The choice $\Theta_\pm \in \exp\mathcal{B}_\pm$ precisely ensures this condition. To show this, we use the fact that one can perform the gauge transformation using either Θ_+ or Θ_-. Using Θ_+ we find that $A^g_{x_+}$ contains only positive heights, while using Θ_- we see that its maximal height is 1, so finally $A^g_{x_+}$ decomposes on elements of height $0,1$. Similarly $A^g_{x_-}$ decomposes on elements of height $0,-1$.

Let us write

$$\begin{aligned}\Theta_+ &= e^{\Delta_+}M_+, \qquad \Delta_+ \in \mathcal{H}, \quad M_+ \in \exp\mathcal{N}_+ \\ \Theta_- &= e^{\Delta_-}M_-, \qquad \Delta_- \in \mathcal{H}, \quad M_- \in \exp\mathcal{N}_-\end{aligned}$$

Performing the gauge transformation Θ_+ we have:

$$A^g_{x_+} = e^{\Delta_+}M_+\partial_{x_+}(M_+^{-1}e^{-\Delta_+}) + e^{\Delta_+}M_+\left(\partial_{x_+}\Phi + e^{\mathrm{ad}\,\Phi}\mathcal{E}_+\right)M_+^{-1}e^{-\Delta_+}$$

Projecting this equation on $\mathcal{H}$, we obtain: $\partial_{x_+}\Phi^g = \partial_{x_+}\Phi - \partial_{x_+}\Delta_+$. Performing now the gauge transformation with Θ_- we get:

$$A^g_{x_+} = e^{\Delta_-}M_-\partial_{x_+}(M_-^{-1}e^{-\Delta_-}) + e^{\Delta_-}M_-\left(\partial_{x_+}\Phi + e^{\mathrm{ad}\,\Phi}\mathcal{E}_+\right)M_-^{-1}e^{-\Delta_-}$$

Therefore, comparing the terms of height $+1$ we get $e^{\mathrm{ad}\,\Phi^g}\mathcal{E}_+ = e^{\mathrm{ad}\,(\Phi+\Delta_-)}\mathcal{E}_+$, so we must have $\alpha_i(\Phi^g) = \alpha_i(\Phi+\Delta_-)$ for all α_i simple, which implies $\Phi^g = \Phi + \Delta_-$. The same analysis performed with A_{x_-} gives $\partial_{x_-}\Phi^g = \partial_{x_-}\Phi + \partial_{x_-}\Delta_-$ and $\Phi^g = \Phi - \Delta_+$. We see that these four equations are compatible if and only if

$$\Delta_- = -\Delta_+$$

i.e if Θ_+ and Θ_- have opposite components on $\exp\mathcal{H}$. ■

Proposition. *The induced action on the chiral fields $\xi(x_+)$ and $\overline{\xi}(x_-)$ is remarkably simple:*

$$\xi^g(x_+) = \xi(x_+)\; g_-^{-1} \qquad ; \qquad \overline{\xi}^g(x_-) = g_+\;\overline{\xi}(x_-) \tag{11.47}$$

Note that the factors $g_\pm$ of g act separately on the two chiral sectors. The dressed Toda field is:

$$e^{-2\lambda^{(r)}(\Phi^g(x,t))} = \xi^{(r)}(x_+)\; g_-^{-1}\cdot g_+\;\overline{\xi}^{(r)}(x_-)$$

<u>Proof</u>. We have

$$\xi^g = <\lambda|(e^{-\Phi^g}\Psi^g) = e^{-\lambda(\Phi^g)} <\lambda|\,(\Theta_-\Psi)\,g_-^{-1} = e^{-\lambda(\Phi^g-\Phi-\Delta_-)} <\lambda|\left(e^{-\Phi}\Psi\right)g_-^{-1} = \xi\cdot g_-^{-1} \tag{11.48}$$

In eq.(11.48) we used the gauge transformation by Θ_- and used the properties of the highest weight vector $<\lambda|$ The transformation law for $\overline{\xi}$ is found in a similar way. ■

Proposition. *Let $g = g_-^{-1} g_+$ and $h = h_-^{-1} h_+$ in G. Their multiplication in the dressing group is:*

$$(g_+, g_-) \bullet (h_+, h_-) \;=\; (g_+ h_+, g_- h_-) \tag{11.49}$$

In particular, the plus and minus components commute.

Proof. This follows from the general statement in Chapter [2] but is also obvious from the action on the chiral fields $\xi, \overline{\xi}$. ■

We denote by G^* the new group equipped with this multiplication law. For infinitesimal transformations, $g \simeq 1 + X$, and $g_\pm \simeq 1 + X_\pm$ with $X = X_+ - X_-$, the composition law corresponds to:

$$[X, Y]_R = [X_+, Y_+] - [X_-, Y_-] \tag{11.50}$$

This defines a new Lie algebra $\mathcal{G}^*$ which is the Lie algebra of G^*. As in Chapter [3], we introduce the operators $R_\pm$ such that $X_\pm = R_\pm X$. Then we have:

$$R_+ X = \mathrm{Tr}_2\,(r_{12}^+ X_2), \quad R_- X = \mathrm{Tr}_2\,(r_{12}^- X_2) \tag{11.51}$$

where $r_{12}^\pm$ are the matrices in eqs.(11.24,11.25). This exhibits the deep relation between the group of dressing transformations and the r–matrix formalism.

Using the action of dressing transformations on the chiral fields $\xi(x,t)$ and $\overline{\xi}(x,t)$, it is easy to understand the Poisson–Lie group nature of the group of dressing transformations, see Chapter [14]. Indeed the exchange algebra

$$\{\xi_1(x), \xi_2(y)\} = \xi_1(x)\xi_2(y) r_{12}^\pm$$

is not invariant by $\xi \to \xi g_-^{-1}$ since $r_{12}^\pm$ does not commute with $g_-^{-1} \otimes g_-^{-1}$. However, if we introduce Poisson brackets for the g_-^{-1} variables:

$$\{g_{-1}^{-1}, g_{-2}^{-1}\} = -[r_{12}^\pm, g_{-1}^{-1} g_{-2}^{-1}], \quad \{\xi_1(x), g_{-2}^{-1}\} = 0$$

then covariance of the Poisson brackets for ξ is recovered, because, viewing ξ^g as a function on the product of the group G^* by the phase space M equipped with the product Poisson structure, we have:

$$\begin{aligned}
\{\xi_1^g(x), \xi_2^g(y)\}_{G^* \times M} &= \{\xi_1(x), \xi_2(y)\}_M g_{-1}^{-1} g_{-2}^{-1} + \xi_1(x)\xi_2(y)\{g_{-1}^{-1}, g_{-2}^{-1}\}_{G^*} \\
&= \xi_1(x)\xi_2(y)\left[r_{12}^\pm g_{-1}^{-1} g_{-2}^{-1} + \{g_{-1}^{-1}, g_{-2}^{-1}\}_{G^*}\right] \\
&= \xi_1^g(x)\xi_2^g(y) r_{12}^\pm
\end{aligned}$$

More generally we have:

Proposition. *Let G^* be equipped with the Poisson brackets:*

$$\begin{aligned}
\{g_{+1}, g_{+2}\}_{G^*} &= [r_{12}^\pm, g_{+1} g_{+2}] \quad ; \quad \{g_{-1}, g_{-2}\}_{G^*} = [r_{12}^\mp, g_{-1} g_{-2}] \\
\{g_{-1}, g_{+2}\}_{G^*} &= [r_{12}^-, g_{-1} g_{+2}] \quad ; \quad \{g_{+1}, g_{-2}\}_{G^*} = [r_{12}^+, g_{+1} g_{-2}]
\end{aligned} \tag{11.52}$$

then the action $\xi \to \xi g_-^{-1}$ $\overline{\xi} \to g_+ \overline{\xi}$ leaves the exchange algebra (11.27) invariant.

Note that for the factorized element $g = g_-^{-1} g_+$ the Poisson bracket on G^* reads:

$$\{g_1, g_2\}_{G^*} = g_1 r_{12}^+ g_2 + g_2 r_{12}^- g_1 - g_1 g_2 r_{12}^\pm - r_{12}^\mp g_1 g_2$$

Since dressing transformations are not symplectic transformations, they are not generated by scalar hamiltonians. However, since they are Lie–Poisson actions, there exists a non-abelian analogue of the moment map. In other words, there exists a non-abelian hamiltonian generating the dressing transformations.

If we consider the system on the finite interval of length ℓ with periodic boundary conditions, then $\Psi(x+\ell) = \Psi(x) \cdot T$, where $T \equiv \Psi(\ell)$ is the monodromy matrix. Clearly, the Poisson brackets of T are: $\{T_1, T_2\} = -[r_{12}^{\pm}, T_1 T_2]$. The non-abelian hamiltonian turns out to be the monodromy matrix $T = \Psi(\ell)$. This means that the variation of the chiral fields are given by the following Poisson brackets:

Proposition. *The monodromy matrix generates dressing transformations:*

$$\begin{aligned} \delta_X \xi(x) &= -\xi(x) X_- = -\mathrm{Tr}_2 \left(X_2 T_2^{-1} \{\xi_1(x), T_2\} \right) \\ \delta_X \overline{\xi}(x) &= X_+ \overline{\xi}(x) = -\mathrm{Tr}_2 \left(X_2 T_2^{-1} \{\overline{\xi}_1(x), T_2\} \right) \end{aligned} \tag{11.53}$$

Proof. We have to compute the Poisson bracket of the monodromy matrix and the chiral fields. Using eqs.(11.28,11.29) we have:

$$\{T_1, e^{-\Phi_2(x)} \Psi_2(x)\} = T_1 e^{-\Phi_2(x)} \Psi_2(x) r_{12}^{\pm} + T_1 \Psi_1^{-1}(x) e^{-\Phi_2(x)} (C_{12}^0 - r_{12}^{\pm}) \Psi_1(x) \Psi_2(x)$$

Projecting on $1 \otimes < \lambda^{(r)}|$ and choosing the plus sign in $r_{12}^{\pm}$ we see that the second term cancels, yielding $\{T_1, \xi_2(x)\} = T_1 \xi_2 r_{12}^{+}$. Similarly we have:

$$\{\xi_1(x), T_2\} = \xi_1(x) T_2 \; r_{12}^{-} \quad \text{and} \quad \{\overline{\xi}_1(x), T_2\} = -T_2 r_{12}^{+} \overline{\xi}_1(x)$$

The proposition then follows from eq.(11.51). ■

We see that Toda field theories provide a particularly neat example of the action of dressing transformations and non abelian Hamiltonians.

11.7 Periodic solutions.

We now consider the Toda theory on a circle and study *periodic* solutions to the field equations. More precisely we assume that $\Phi(x + 2\pi, t) = \Phi(x, t)$. If the Toda field Φ is periodic, the potentials in the chiral equations:

$$L \cdot \xi = 0, \quad \overline{L} \cdot \overline{\xi} = 0 \tag{11.54}$$

are periodic, since they are local functions of Φ and its derivatives. Hence the natural basis of solutions are the quasi periodic solutions, the so called Bloch waves. This section is devoted to the construction and study of these Bloch waves. Instead of eq.(11.54) it is more convenient to work with eq.(11.30). Unfortunately, the relation between the coefficients of eq.(11.54) and the fields P and $\overline{P}$ is far from one to one. However eq.(11.30) implies, for example in the sl_{n+1} case, $(\partial - P_1)\xi_1 = 0$, $(\partial - P_2)(\partial - P_1)\xi_2 = 0$, etc., so that if the ξ_i are Bloch waves, we see recursively that the P_i are *periodic*. So one can always choose the chiral fields P and $\overline{P}$ periodic to parametrize solutions of the Toda field equations on the circle. Starting from eq.(11.30) with periodic field P we have:

$$Q_+(x + 2\pi) = Q_+(x) S$$

where the monodromy matrix S belongs to $\exp(\mathcal{B}_+)$. Here we note that Q_+ is a function of x_+ and the shift $(x, t) \to (x + 2\pi, t)$ is equivalent to $x_+ \to x_+ + 2\pi$.

To find the Bloch waves, we have to diagonalize this monodromy matrix:

$$S = g^{-1} e^{P_0} g \tag{11.55}$$

where $g \in \exp(\mathcal{N}_+)$. It is obvious from the triangularity of eq.(11.30) that $P_0 = \int_0^{2\pi} P(x)dx$ is the zero mode of the field $P(x)$.

The quasi periodic solutions are defined as

$$\psi^{(r)}(x) = \langle \lambda^{(r)} | Q_+(x) g^{-1} \theta, \quad \psi^{(r)}(x+2\pi) = \psi^{(r)}(x) e^{P_0} \tag{11.56}$$

where the diagonal matrix θ is a normalization matrix.

Doing a similar analysis on the other chirality, we have $Q_-^{-1}(x+2\pi) = \overline{S} Q_-^{-1}(x)$ with $\overline{S} = \overline{g}^{-1} e^{-\overline{P}_0} \overline{g}$ so that the Bloch wave is $\overline{\psi}^{(r)} = \overline{\theta}\, \overline{g} Q_-^{-1} |\lambda^{(r)} >$. The expression of the Toda field becomes:

$$e^{\lambda^{(r)}(\Phi)} = \psi^{(r)}(x_+) \overline{\psi}^{(r)}(x_-)$$

and will be periodic if $\overline{P}_0 = P_0$.

In this description, the Bloch waves are building blocks of the solution, so it is interesting to compute their Poisson brackets. With an adequate choice of the normalization θ, these brackets take a simple form.

Proposition. *There exists a choice of θ and $\overline{\theta}$ such that:*

$$\begin{aligned} \{\psi_1^{(r)}(x), \psi_2^{(r')}(y)\} &= \psi_1^{(r)}(x) \psi_2^{(r')}(y) r_{12}^{\pm}(P_0), \quad \text{for } x \gtrless y, \quad \{\psi_1^{(r)}(x), P_{02}\} = 2\psi_1^{(r)}(x) C_{12}^0 \\ \{\overline{\psi}_1^{(r)}(x), \overline{\psi}_2^{(r')}(y)\} &= -r_{12}^{\pm}(\overline{P}_0) \overline{\psi}_1^{(r)}(x) \overline{\psi}_2^{(r')}(y), \quad \text{for } x \gtrless y, \quad \{\overline{\psi}_1^{(r)}(x), \overline{P}_{02}\} = 2C_{12}^0 \overline{\psi}_1^{(r)}(x) \end{aligned} \tag{11.57}$$

where

$$r_{12}^{\pm}(P_0) = \pm \frac{1}{2} \mathcal{C}_{12} + \frac{1}{2} \sum_{\alpha \text{ positive}} \coth(\frac{1}{2} \alpha(P_0)) \left(\frac{E_\alpha \otimes E_{-\alpha} - E_{-\alpha} \otimes E_\alpha}{(E_\alpha, E_{-\alpha})} \right)$$

is the dynamical r–matrix of the trigonometric Calogero model.

In these formulae, P_0 and $\overline{P}_0$ as well as their conjugate variables θ and $\overline{\theta}$ are independent. In particular,

$$\{\psi_1^{(r)}(x), \overline{\psi}_2^{(r)}(y)\} = 0$$

The original phase space of the theory is recovered by performing a Hamiltonian reduction with the constraint $P_0 - \overline{P}_0 = 0$.

Proof. The proof is long but not difficult, so we shall only sketch the main steps. One first computes the Poisson brackets of the solution $Q_+(x)$ of eq.(11.30), normalized such that $Q_+(0) = 1$, which entails $S = Q_+(2\pi)$. We get for $0 < y < x < 2\pi$, using the Poisson bracket (11.34):

$$\{Q_{+1}(x), Q_{+2}(y)\} = Q_{+1}(x) Q_{+2}(y) \left[r_{12}^{\pm} - Q_{+1}^{-1}(y) Q_{+2}^{-1}(y) (r_{12}^{\pm} - C_{12}^0) Q_{+1}(y) Q_{+2}(y) \right] \tag{11.58}$$

$$\{Q_{+1}(x), S_2\} = Q_{+1}(x) C_{12}^0 S_2 + Q_{+1}(x) S_2 r_{12}^{\pm} - S_2 Q_{+2}^{-1}(x) (r_{12}^{\pm} + C_{12}^0) Q_{+1}(x) Q_{+2}(x) \tag{11.59}$$

$$\{S_1, S_2\} = -[r_{12}^{\pm}, S_1 S_2] - S_2 C_{12}^0 S_1 + S_1 C_{12}^0 S_2 \tag{11.60}$$

To show this, we view $Q_+(x)$ as a path ordered exponential so that we can write:

$$\{Q_{+1}(x), Q_{+2}(y)\} = Q_{+1}(x) Q_{+2}(y) \int_0^x dz \int_0^y dt\; Q_{+1}^{-1}(z) Q_{+2}^{-1}(t) \{P_1(z), P_2(t)\} Q_{+1}(z) Q_{+2}(t)$$

using eq(11.34), integrating by parts and writing all the terms explicitly we get:

$$\begin{aligned}\{Q_{+1}(x), Q_{+2}(y)\} &= Q_{+1}(x)Q_{+2}(y)\Big\{-\int_0^x dz \int_0^y dt\delta(z-t)(\partial_t - \partial_z)\left(Q_{+1}^{-1}(z)Q_{+2}^{-1}(t)C_{12}^0 Q_{+1}(z)Q_{+2}(t)\right) \\ &- \int_0^y dt\delta(x-t)Q_{+1}^{-1}(x)Q_{+2}^{-1}(t)C_{12}^0 Q_{+1}(x)Q_{+2}(t) + \int_0^y dt\delta(t)Q_{+2}^{-1}(t)C_{12}^0 Q_{+2}(t) \\ &+ \int_0^x dz\delta(z-y)Q_{+1}^{-1}(z)Q_{+2}^{-1}(y)C_{12}^0 Q_{+1}(z)Q_{+2}(y) - \int_0^x dz\delta(z)Q_{+1}^{-1}(z)C_{12}^0 Q_{+1}(z)\Big\}\end{aligned} \tag{11.61}$$

Since the field $P(x)$ is periodic, so is the delta function in eq(11.34), and we can write $\delta(t) = \frac{1}{2\pi}\sum_m e^{imt}$. Then one has for any continuous function $f(t)$:

$$\int_0^y dt\delta(t)f(t) = \begin{cases} \frac{1}{2}f(0) & \text{if} \quad 0 < y < 2\pi \\ 0 & \text{if} \quad y = 0 \\ \frac{1}{2}(f(0)+f(2\pi)) & \text{if} \quad y = 2\pi \end{cases} \tag{11.62}$$

The first equation follows from $\delta(-t) = \delta(t)$ and the last one follows from the classical theorem on Fourier series. The last four terms in eq.(11.61) are boundary terms, and straightforward to evaluate. The first term is equal for $x > y$ to:

$$\begin{aligned}&\int_0^x dz \int_0^y dt\delta(z-t)(\partial_t - \partial_z)\left(Q_{+1}^{-1}(z)Q_{+2}^{-1}(t)C_{12}^0 Q_{+1}(z)Q_{+2}(t)\right) = \\ &\quad \int_0^x dz \int_0^y dt\delta(z-t)\frac{1}{2}Q_{+1}^{-1}(z)Q_{+2}^{-1}(t)\left(\sum_i [H_i, \mathcal{E}_+]\otimes H_i - H_i \otimes [H_i, \mathcal{E}_+]\right)Q_{+1}(z)\otimes Q_{+2}(t) \\ &\quad = -r_{12}^{\pm} + Q_{+1}^{-1}(y)Q_{+2}^{-1}(y)r_{12}^{\pm}Q_{+1}(y)Q_{+2}(y)\end{aligned} \tag{11.63}$$

This is because, using eq.(3.18) in Chapter [3]

$$\frac{1}{2}\sum_i [H_i, \mathcal{E}_+]\otimes H_i - H_i\otimes[H_i, \mathcal{E}_+] = -[r_{12}^{\pm}, \mathcal{E}_+\otimes 1 + 1\otimes \mathcal{E}_+]$$

eq.(11.63) follows from the fact that $[r, \cdot]$ is a derivation, and eq.(11.58) is then obvious. The other two equations are also a matter of setting $x, y = 2\pi$ in eq.(11.61).

To proceed, we need to diagonalize S. This can be achieved with a strictly upper triangular matrix $g \in \exp \mathcal{N}_+$. The Poisson brackets of g and Q_+ can be computed. We write $S = g^{-1}e^{P_0}g$, and use the fact that P_0 commutes with everything because S is a function of P and P_0 is in the center of the bracket. Starting from eq.(11.59) we get, with $A = \frac{1}{2}\mathrm{ad}\, P_0$, $A_1 = A\otimes 1$, etc.:

$$\begin{aligned}\sinh(A_2)\{Q_{+1}(x), g_2\}g_2^{-1} &= \frac{1}{2}\cosh(A_2)Q_{+1}(x)g_2(r_{12}^- + C_{12}^0)g_2^{-1} + \frac{1}{2}\sinh(A_2)Q_{+1}(x)g_2(r_{12}^- - C_{12}^0)g_2^{-1} \\ &-\frac{1}{2}e^{A_2}g_2Q_{+2}^{-1}((r_{12}^- + C_{12}^0)Q_{+1}Q_{+2}\end{aligned}$$

Remark that $\sinh(A)E_\alpha = \sinh(\frac{1}{2}\alpha(P_0))E_\alpha$ does not vanish for a generic P_0 for any α. It follows that the operator on the left–hand side is invertible on the space $\mathcal{N}_+$. Remembering that $r_{12}^- + C_{12}^0 \in \mathcal{G}\otimes\mathcal{N}_+$, we see that the first and the third term in the right–hand side are already in $\mathcal{G}\otimes\mathcal{N}_+$. One can add $2Q_{+1}(x)C_{12}^0$, which is in the kernel of $\sinh(A_2)$ to the second term so that it belongs also to $\mathcal{G}\otimes\mathcal{N}_+$. One can then plainly divide by $\sinh(A_2)$ to get:

$$\begin{aligned}\{Q_{+1}(x), g_2\}g_2^{-1} &= \frac{1}{2}\coth(A_2)Q_{+1}(x)g_2(r_{12}^- + C_{12}^0)g_2^{-1} + \frac{1}{2}Q_{+1}(x)g_2(r_{12}^- - C_{12}^0)g_2^{-1} + Q_{+1}(x)C_{12}^0 \\ &-\frac{1}{2}\frac{e^{A_2}}{\sinh(A_2)}g_2Q_{+2}^{-1}((r_{12}^- + C_{12}^0)Q_{+1}Q_{+2}\end{aligned}$$

Note that ultimately we have to project this on $< \lambda^{(r)}| \otimes 1$ which annihilates the last term. Similarly eq.(11.60) yields:

$$2\sinh(A_1)\sinh(A_2)\{g_1, g_2\}g_1^{-1}g_2^{-1} = \sinh(A_1 + A_2)g_1g_2r_{12}^{\pm}g_1^{-1}g_2^{-1} + \sinh(A_1 - A_2)g_1g_2C_{12}^0g_1^{-1}g_2^{-1}$$

This can be solved in $\mathcal{N}_+ \otimes \mathcal{N}_+$ to get:

$$\begin{aligned} \{g_1, g_2\}g_1^{-1}g_2^{-1} &= \frac{1}{2}\coth(A_2)\Big[g_1g_2(r_{12}^{\pm} + C_{12}^0)g_1^{-1}g_2^{-1} - r_{12}^{\pm} - C_{12}^0 - 2(g_2C_{12}^0g_2^{-1} - C_{12}^0)\Big] \\ &+\frac{1}{2}\coth(A_1)\Big[g_1g_2(r_{12}^{\pm} - C_{12}^0)g_1^{-1}g_2^{-1} - r_{12}^{\pm} + C_{12}^0 + 2(g_1C_{12}^0g_1^{-1} - C_{12}^0)\Big] \end{aligned}$$

We now choose carefully the normalization θ in eq.(11.56). We set

$$\theta = e^K e^{Q_0}, \quad K = \sum_{n\neq 0} \frac{P_n}{in}, \quad \{P_{01}, Q_{01}\} = 2C_{12}^0 \tag{11.64}$$

The variable $Q_0 \in \mathcal{H}$ is simply the conjugate variable of P_0. To have a motivation for the choice of the variable K, consider the abelian case of eq.(11.30) i.e. $[\partial_{x_+} - P]Q_+ = 0$. The solution with the boundary condition $Q_+(0) = 1$ is $Q_+(z) = \exp(\int_0^z Pdz)$. If we take a periodic field P this is equal to $Q_+(z) = \exp(\frac{1}{2\pi}P_0z + \sum_n \frac{P_n}{in}e^{inz})e^{-K}$. Hence, with the choice eq.(11.64), we get

$$\psi(z) = \exp(Q_0 + \frac{1}{2\pi}P_0z + \sum_n \frac{P_n}{in}e^{inz})$$

We recognize the standard form of a vertex operator. The variable K has erased any reference to the origin. One can easily check that, with this normalization, the Bloch wave $\psi^{(r)}(x)$ is a differential of weight $-\lambda^{(r)}(H_\rho)$.

Using that $\{P_1(x), K_2\} = 2\left(\delta(x) - \frac{1}{2\pi}\right)C_{12}^0$ and $\{P_1(x), Q_{02}\} = \frac{1}{\pi}C_{12}^0$ we find $\{P_1(x), \theta_2\} = 2\delta(x)\theta_2C_{12}^0$. It follows that:

$$\begin{aligned} \{Q_{+1}(x), \theta_2\} &= 2Q_{+1}(x)\theta_2 \int_0^x dz\delta(z)Q_{+1}^{-1}(z)C_{12}^0Q_{+1}(z) \\ \{g_1, \theta_2\}g_1^{-1} &= \coth(A_1)\Big(g_1\theta_2C_{12}^0g_1^{-1} - \theta_2C_{12}^0\Big) \end{aligned}$$

We now have all the necessary information to compute the Poisson algebra of the Bloch waves, and get eq.(11.57). ■

Recall that up to now the two chiralities are independent, but we have to impose the constraint $P_0 = \overline{P}_0$ to get periodic solutions. We also have to get rid of one of the variables $Q_0, \overline{Q}_0$ to recover the original phase space of the theory. Noting that:

$$\{e^{-\lambda_{max}^{(r)}(\Phi)(x,t)}, P_0 - \overline{P}_0\} = 0$$

we see that the reconstructed field Φ does indeed live on the reduced phase space. Moreover, the Poisson algebra of the periodic fields Φ is local:

$$\left\{e^{-\lambda_{max}^{(r)}(\Phi)(x,t)}, e^{-\lambda_{max}^{(r')}(\Phi)(x',t)}\right\} = 0, \quad 0 \leq x, x' \leq 2\pi$$

We end up this section by writing the Yang-Baxter equation satisfied by $r^{\pm}(P_0)$. It is a direct consequence of the Jacobi identity on three ψ's, and the fact that ψ and P_0 do not Poisson commute. If we introduce the components of P_0 on the othonormal basis H_i of the Cartan subalgebra $P_0 = \sum_i P_0^iH_i$, we have

$$\{\psi^{(r)}(x), f(P_0)\} = \psi^{(r)}(x)\ \mathcal{D}(f)$$

where we define the operator $\mathcal{D}$ taking values in $\mathcal{H}$ as

$$\mathcal{D}(f) = \frac{1}{2\pi} \sum_i H_i \frac{\partial f}{\partial P_0^i}$$

Using this formula, we immediately see that the matrices $r^{\pm}(P_0)$ satisfy the dynamical Yang-Baxter equation

$$[r^{\pm}_{12}(P_0), r^{\pm}_{13}(P_0)] + [r^{\pm}_{12}(P_0), r^{\pm}_{23}(P_0)] + [r^{\pm}_{13}(P_0), r^{\pm}_{23}(P_0)] + \mathcal{D}_3 r^{\pm}_{12}(P_0) - \mathcal{D}_2 r^{\pm}_{13}(P_0) + \mathcal{D}_1 r^{\pm}_{23}(P_0) = 0$$

where the subscript on $\mathcal{D}$ refers to the label of the space in the tensor product $\mathcal{G} \otimes \mathcal{G} \otimes \mathcal{G}$ in which $\mathcal{D}$ takes its values.

References

[1] A.N.Leznov, M.V.Saveliev, *Representation of zero curvature for the system of non-linear partial differential equations* $x_{\alpha,z\overline{z}} = \exp(kx)_\alpha$ *and its integrability.* Lett .Math.Phys. 3 (1979) pp.489-494.

[2] J.L. Gervais, A. Neveu, *Novel triangle relation and absence of tachyons in Liouville string field theory.* Nucl. Phys. B238 (1984) 125.

[3] A.A. Belavin, A.M. Polyakov, A.B. Zamolodchikov, *Infinite conformal symmetry in two–dimensional quantum field theory.* Nucl. Phys. B241 (1984) 333.

[4] M. Semenov Tian Shansky. *"Dressing transformations and Poisson group actions"*. Publ. RIMS 21 (1985), p. 1237.

[5] O.Babelon, *"Extended Conformal Algebra and the Yang-Baxter Equation"*. Phys.Lett. B215 (1988) 523.

[6] J.H.Lu, *"Multiplicative and Affine Poisson Structures on Lie Groups"*. PhD Thesis, University of California at Berkeley (1990).

[7] O.Babelon, L.Bonora, F.Toppan *"Exchange algebra and the Drinfeld-Sokolov theorem"* Commun.Math.Phys. 140,(1991), 93.

[8] O.Babelon, D.Bernard, *"Dressing Transformations and the Origin of Quantum Group Symmetries"* Phys.Lett. B260 (1991) 81.

[9] V.A.Fateev,S.L.Lukyanov, *" Poisson-Lie Groups and Classical W-algebras."* Int. J. Mod. Phys. A 7 (1992) p.853-876.

Chapter 12

The sine–Gordon theory.

In this Chapter we study soliton and finite zone solutions of the sine–Gordon equation. In the next Chapter we will apply the inverse scattering method to discuss further solutions. Here, the sine–Gordon theory will be considered as a Toda field theory associated to the affine Lie algebra $\widehat{sl_2}$. In particular, soliton solutions are constructed by the action of very special elements of the dressing group on the vacuum solution. Compared to the KP situation, we have here an example with two singularities, each one accomodating one of the two relativistic light–cone variables $x \pm t$, and their corresponding hierarchy of times.

12.1 The sine–Gordon model.

The sine–Gordon equation reads:

$$\partial_t^2 \varphi - \partial_x^2 \varphi = -8m^2 \sin(2\varphi) \tag{12.1}$$

In this Chapter, we are concerned with the algebraic structure of this equation, so the field φ can be regarded as complex and it will be more convenient to replace the sine–Gordon equation by the sinh–Gordon equation:

$$\partial_t^2 \varphi - \partial_x^2 \varphi = -8m^2 \sinh(2\varphi) \tag{12.2}$$

This equation is the compatibility condition of a linear system:

$$(\partial_x + A_x)\, \Psi = 0, \quad (\partial_t \, + A_t)\, \Psi = 0 \tag{12.3}$$

where the matrices A_x, A_t belong to $sl_2 \otimes \mathbb{C}(\lambda, \lambda^{-1})$, where λ is the spectral parameter:

$$A_x = \begin{pmatrix} \frac{1}{2}\partial_t \varphi & m(\lambda e^{\varphi} + \lambda^{-1} e^{-\varphi}) \\ m(\lambda e^{-\varphi} + \lambda^{-1} e^{\varphi}) & -\frac{1}{2}\partial_t \varphi \end{pmatrix} \tag{12.4}$$

$$A_t = \begin{pmatrix} \frac{1}{2}\partial_x \varphi & m(\lambda e^{\varphi} - \lambda^{-1} e^{-\varphi}) \\ m(\lambda e^{-\varphi} - \lambda^{-1} e^{\varphi}) & -\frac{1}{2}\partial_x \varphi \end{pmatrix} \tag{12.5}$$

This means that the sinh–Gordon equation is equivalent to the zero curvature condition:

$$F_{xt} \equiv \partial_x A_t - \partial_t A_x + [A_x, A_t] = 0$$

Notice that this condition appeals only to the Lie algebra structure to which the Lax connexion belongs to.

It is often convenient to write the Lax connexion in the light cone coordinates $x_\pm = x \pm t$, $\partial_{x_\pm} = \frac{1}{2}(\partial_x \pm \partial_t)$. We find

$$A_{x_+} = \begin{pmatrix} \frac{1}{2}\partial_{x_+}\varphi & m\lambda e^{\varphi} \\ m\lambda e^{-\varphi} & -\frac{1}{2}\partial_{x_+}\varphi \end{pmatrix}, \quad A_{x_-} = \begin{pmatrix} -\frac{1}{2}\partial_{x_-}\varphi & m\lambda^{-1} e^{-\varphi} \\ m\lambda^{-1} e^{\varphi} & \frac{1}{2}\partial_{x_-}\varphi \end{pmatrix} \tag{12.6}$$

To understand the form of the connexion A_x, A_t, remark that they are matrices rational in the spectral parameter λ, with two poles, one at $\lambda = 0$, the other at $\lambda = \infty$. Moreover, in the light–cone coordinates, A_+ has a simple pole at ∞ and A_- a simple pole at 0. Let us look at the most general 2×2 system of this form:

$$\begin{aligned}(\partial_{x_+} + U)\Psi &= 0, \quad U = U_0 + \lambda U_1 \\ (\partial_{x_-} + V)\Psi &= 0, \quad V = V_0 + \frac{1}{\lambda} V_1\end{aligned}$$

and impose the zero curvature condition $[\partial_{x_+} + U, \partial_{x_-} + V] = 0$. The matrices U_i, V_i are taken to be traceless matrices, so contain 12 parameters. One can reduce this number by imposing a symmetry condition under a discrete group, as explained in Chapter [2]. Namely, we consider the group $\mathbb{Z}_2$ acting by:

$$\Psi(\lambda) \longrightarrow \sigma_z \Psi(-\lambda)\sigma_z^{-1}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

and we demand that Ψ be invariant by this action. This restriction means that the wave function belongs to the twisted loop group. It follows that: $\sigma_z U(-\lambda)\sigma_z = U(\lambda)$ and $\sigma_z V(-\lambda)\sigma_z = V(\lambda)$. We still have the possibility to perform a gauge transformation by an element g, independent of λ, in order to preserve the pole structure of the connexion, and commuting with the action of $\mathbb{Z}_2$, i.e. g diagonal. This gauge freedom can be used to set $(V_0)_{ii} = 0$. The symmetry condition then gives:

$$U = \begin{pmatrix} u_0 & \lambda u_1 \\ \lambda u_2 & -u_0 \end{pmatrix}, \quad V = \begin{pmatrix} 0 & \lambda^{-1} v_1 \\ \lambda^{-1} v_2 & 0 \end{pmatrix}$$

In this gauge, the zero curvature equation reduces to :

$$\partial_{x_-} u_0 + u_1 v_2 - v_1 u_2 = 0 \tag{12.7}$$

$$\partial_{x_-} u_1 = 0, \quad \partial_{x_-} u_2 = 0 \tag{12.8}$$

$$\partial_{x_+} v_1 + 2 v_1 u_0 = 0, \quad \partial_{x_+} v_2 - 2 v_2 u_0 = 0 \tag{12.9}$$

From eq.(12.8) we have $u_1 = \alpha(x_+)$, $u_2 = \beta(x_+)$. We set $u_0 = \partial_{x_+}\varphi$. Then, from eq(12.9) we have $v_1 = \gamma(x_-)\exp -2\varphi$ and $v_2 = \delta(x_-)\exp 2\varphi$. Finally eq(12.7) becomes :

$$\partial_{x_+}\partial_{x_-}\varphi + \beta(x_+)\gamma(x_-)e^{-2\varphi} - \alpha(x_+)\delta(x_-)e^{2\varphi} = 0$$

This is the sinh–Gordon equation. The functions $\alpha(x_+), \beta(x_+)$ and $\gamma(x_-), \delta(x_-)$ are irrelevant : they can be absorbed into a redefinition of the field φ and a change of the coordinates x_+, x_-. Taking them as constants, equal to m, we finally get

$$U = \begin{pmatrix} \partial_{x_+}\varphi & m\lambda \\ m\lambda & -\partial_{x_+}\varphi \end{pmatrix}; \quad V = \begin{pmatrix} 0 & m\lambda^{-1}e^{-2\varphi} \\ m\lambda^{-1}e^{2\varphi} & 0 \end{pmatrix} \tag{12.10}$$

This connexion is related to eq.(12.6) by a gauge transformation by the diagonal matrix $\exp(\frac{1}{2}\varphi\sigma_z)$. Hence the connexion A_{x_+}, A_{x_-} of the sinh–Gordon model is naturally recovered from two–poles systems with $\mathbb{Z}_2$ symmetry. This construction generalizes to other Lie algebras, the reduction group being generated by the Coxeter automorphism.

Remark. There is a relation between the linear system eq.(12.6) and what is called Bäcklund transformations. These are a way to produce new solutions of a non–linear partial differential equation from old ones. Assume that φ satisfies a second order non–linear partial differential equation. We require that the new function $\widehat{\varphi}$ is obtained by solving the first order system:

$$\partial_{x_+}\widehat{\varphi} = P(\widehat{\varphi}, \phi, \partial_{x_+}\varphi, \partial_{x_-}\varphi), \quad \partial_{x_-}\widehat{\varphi} = Q(\widehat{\varphi}, \phi, \partial_{x_+}\varphi, \partial_{x_-}\varphi)$$

Of course the compatibility condition of this system puts strong constraints on P and Q. We say the transformation is auto–Bäcklund if $\widehat{\varphi}$ satisfies the same equation as φ. In the sinh–Gordon case the transformation is defined by:

$$\partial_{x_+}(\widehat{\varphi}+\varphi) = 2m\lambda \sinh(\widehat{\varphi}-\varphi), \quad \partial_{x_-}(\widehat{\varphi}-\varphi) = 2m\lambda^{-1}\sinh(\widehat{\varphi}+\varphi) \tag{12.11}$$

where λ is an arbitrary parameter. The compatibility condition reads:

$$\partial_{x_+}\partial_{x_-}\widehat{\varphi} = -\partial_{x_+}\partial_{x_-}\varphi + 4m^2\cosh(\widehat{\varphi}-\varphi)\sinh(\widehat{\varphi}+\varphi) = \partial_{x_+}\partial_{x_-}\varphi + 4m^2\cosh(\widehat{\varphi}+\varphi)\sinh(\widehat{\varphi}-\varphi)$$

and reduces to the sinh–Gordon equation $\partial_{x_+}\partial_{x_-}\varphi = 2m^2\sinh(2\varphi)$. Moreover we then find $\partial_{x_+}\partial_{x_-}\widehat{\varphi} = 2m^2\sinh(2\widehat{\varphi})$, so if φ solves the sinh-Gordon equation, so does the transformed field $\widehat{\varphi}$.

The relation between eq.(12.11) and the linear system eq.(12.6) is obtained by setting $e^{\widehat{\varphi}} = \frac{u}{v}$. The Bäcklund transformation then reads:

$$\begin{aligned} u(-\partial_{x_+}v + \frac{1}{2}\partial_{x_+}\varphi\, v - m\lambda\, e^{-\varphi}u) + v(\partial_{x_+}u + \frac{1}{2}\partial_{x_+}\varphi\, u + m\lambda\, e^{\varphi}v) &= 0 \\ u(-\partial_{x_-}v - \frac{1}{2}\partial_{x_-}\varphi\, v - \frac{m}{\lambda}\, e^{\varphi}u) + v(\partial_{x_-}u - \frac{1}{2}\partial_{x_-}\varphi\, u + \frac{m}{\lambda}\, e^{-\varphi}v) &= 0 \end{aligned}$$

Requiring the vanishing of the four terms in the parenthesis yields exactly the linear system:

$$(\partial_{x_+} + A_{x_+})\begin{pmatrix} u \\ v \end{pmatrix} = 0, \quad (\partial_{x_-} + A_{x_-})\begin{pmatrix} u \\ v \end{pmatrix} = 0$$

where the connection is given in eq.(12.6). Conversely, if we have a solution (u, v) of the linear sytem associated to φ, then $e^{\widehat{\varphi}} = \frac{u}{v}$ is a solution of the Bäcklund transformation.

In general the relation between φ and $\widehat{\varphi}$ is non–local. However, when expanding $\widehat{\varphi}$ in formal power of either λ or $1/\lambda$ each term of the infinite series is a local function of φ. This remark can be used to deduce an infinite set of local conserved currents in the sinh–Gordon model. Indeed, from the defining relations (12.11) of the Bäcklund transformation, we see that the current $J_{x_+},\ J_{x_-}$ with components:

$$J_{x_+} = \lambda\cosh(\varphi-\widehat{\varphi}), \quad J_{x_-} = -\lambda^{-1}\cosh(\varphi+\widehat{\varphi})$$

is conserved: $\partial_{x_-}J_{x_+} + \partial_{x_+}J_{x_-} = 0$. Expanding it in power series of either λ or $1/\lambda$ gives two infinite series of local conserved quantities.

12.2 The affine sinh–Gordon model.

We now show that the sinh-Gordon theory belongs to the general class of Toda field theories. It is associated to the affine Lie algebra $\widehat{sl_2}$. The usual formula for the Lax connexion eq.(12.4, 12.5), is written in the loop representation of this algebra, with vanishing central charge. In the following we will consider non zero central charge and the benefit is that all the algebraic structure of Toda theories, studied in Chapter [11] then applies to the sinh–Gordon theory as well. In particular the existence of highest weight representations allows us to disentangle the action of the group of dressing transformations. As we will see, this leads directly to the N-soliton formula in terms of vertex operators.

The affine algebra $\widehat{sl_2}$ admits the Cartan decomposition: $\widehat{sl_2} = \widehat{\mathcal{N}}_- \oplus \widehat{\mathcal{H}} \oplus \widehat{\mathcal{N}}_+$. In the principal gradation a basis of $\widehat{sl_2}$ is given, see Chapter [16] by:

$$\begin{aligned} \widehat{\mathcal{H}} &= \{H, d, K\} \\ \widehat{\mathcal{N}}_+ &= \{E_+^{(2n-1)} = \lambda^{2n-1}E_+, E_-^{(2n-1)} = \lambda^{2n-1}E_-, H^{(2n)} = \lambda^{2n}H, n > 0\} \\ \widehat{\mathcal{N}}_- &= \{E_+^{(2n+1)} = \lambda^{2n+1}E_+, E_-^{(2n+1)} = \lambda^{2n+1}E_-, H^{(2n)} = \lambda^{2n}H, n < 0\} \end{aligned} \tag{12.12}$$

In particular, the simple root vectors can be taken as: $E_{\pm\alpha_1} = \lambda^{\pm 1} E_\pm$ and $E_{\pm\alpha_2} = \lambda^{\pm 1} E_\mp$. The commutation relations are:

$$\left[H^{(r)}, H^{(s)}\right] = Kr\ \delta_{r+s,0} \quad , \quad \left[H^{(r)}, E_\pm^{(s)}\right] = \pm 2 E_\pm^{(r+s)} \quad , \quad \left[E_+^{(r)}, E_-^{(s)}\right] = H^{(r+s)} + \frac{K}{2} r \delta_{r+s,0}$$

Following the general construction of Toda field theories of Chapter [11], we define a connexion $A_{x_\pm}$ by

$$A_{x_\pm} \quad = \quad \pm\partial_{x_\pm}\Phi + m e^{\pm ad \Phi}\mathcal{E}_\pm, \quad \text{with} \quad \mathcal{E}_\pm = \lambda^{\pm 1}(E_+ + E_-) = E_{\pm\alpha_1} + E_{\pm\alpha_2} \tag{12.13}$$

where the field Φ takes values in the Cartan subalgebra $\widehat{\mathcal{H}}$ of $\widehat{sl_2}$. Let us decompose it on the generators H, d and K of $\widehat{\mathcal{H}}$:

$$\Phi = \frac{1}{2} H\ \varphi + d\ \eta + \frac{1}{4} K\ \zeta$$

The zero curvature condition, $\partial_{x_+} A_{x_-} - \partial_{x_-} A_{x_+} + [A_{x_+}, A_{x_-}] = 0$, yields the equations of motion which can be worked out using only the Lie algebra structure of $\widehat{sl_2}$.

$$\partial_{x_+}\partial_{x_-}\varphi = m^2 e^{2\eta}(e^{2\varphi} - e^{-2\varphi}), \quad \partial_{x_+}\partial_{x_-}\eta = 0, \quad \partial_{x_+}\partial_{x_-}\zeta = m^2 e^{2\eta}(e^{2\varphi} + e^{-2\varphi}) \tag{12.14}$$

Thanks to the field η, the above equations are conformally invariant. This is in contrast with the sinh-Gordon equation which is not conformally invariant. Indeed, performing a change of coordinates $x_+ = f(x'_+)$ and $x_- = g(x'_-)$ the equations are invariant if we redefine the fields by:

$$\varphi'(x'_+, x'_-) = \varphi(f(x'_+), g(x'_-)), \quad \zeta'(x'_+, x'_-) = \zeta(f(x'_+), g(x'_-)), \quad \eta'(x'_+, x'_-) = \eta(f(x'_+), g(x'_-)) + \log(\partial f \partial g)$$

Remark that there are two real forms of these equations. One is when φ, ζ and η are all real, and the other one is when φ is pure imaginary, and ζ, η are real. These two forms correspond to the sinh–Gordon and sine–Gordon case respectively when one sets $\eta = 0$, which can be done consistently and decouples the equation for φ.

The benefit of having extended the loop algebra to the full affine algebra is that we have now at our disposal highest weight representations and the analysis of Chapter [11] can be applied straightforwardly, provided the action of "group" elements on the highest weight vector is defined. In the following we shall restrict ourselves to integrable highest weights and work freely with formal expressions. Final formulae will be evaluated explicitly in the level 1 representations of $\widehat{sl_2}$ where they will be seen to make sense.

As usual with Toda field theories, to a highest weight vector $|\Lambda\rangle$ one associates two sets of fields $\xi(x,t)$ and $\overline{\xi}(x,t)$ defined by :

$$\xi(x,t) = <\Lambda| e^{-\Phi}\Psi(x,t), \quad \overline{\xi}(x,t) = \Psi(x,t)^{-1} e^{-\Phi}|\Lambda> \tag{12.15}$$

These fields are chiral: $\partial_{x_-}\xi = 0$ and $\partial_{x_+}\overline{\xi} = 0$. From any highest weight Λ we can reconstruct $\Lambda(\Phi)$ by the formula:

$$\exp\left(-2\Lambda(\Phi)\right) = \xi(x_+) \cdot \overline{\xi}(x_-)$$

The affine $\widehat{sl_2}$ algebra has two fundamental highest weights, which we shall denote by Λ^- and Λ^+, see Chapter [16]. They are characterized by: $\Lambda^\pm(H) = \pm\frac{1}{2}$, $\Lambda^\pm(K) = 1$ and $\Lambda^\pm(d) = 0$, so that we have $\Lambda^\pm(\Phi) = \frac{1}{4}(\pm\varphi + \zeta)$. This is enough to reconstruct the fields φ and ζ. The field η cannot be obtained by these highest weight projections. This is not a problem if we are interested in the sinh–Gordon model, since then the free field η is equal to 0. Defining the tau–functions:

$$\tau_\pm = \exp\left(-2\Lambda^\pm(\Phi)\right)$$

we have:

$$e^{-\varphi} = \frac{\tau_+}{\tau_-} \quad \text{and} \quad e^{-\zeta} = \tau_+\tau_- \tag{12.16}$$

In terms of the τ–functions, the equations of motion take the form:

$$\tau_\pm(\partial_{x_-}\partial_{x_+}\tau_\pm) - (\partial_{x_+}\tau_\pm)(\partial_{x_-}\tau_\pm) = -m^2 \ e^{2\eta} \ \tau_\mp^2 \tag{12.17}$$

When $\eta = 0$ this is just the Hirota bilinear form of the sinh–Gordon equation.

We now describe the chiral fields ξ and $\overline{\xi}$ for the simplest solution, the vacuum solution of eq.(12.14):

$$\varphi_{vac} = 0, \quad \eta_{vac} = 0, \quad \zeta_{vac} = 2m^2 x_+ x_-$$

One can insert this solution into the linear system and compute the vacuum wave function $\Psi_{vac}(x,t)$. Indeed we have $\Phi_{vac} = \frac{1}{2}m^2 x_+ x_- K$ and the linear system becomes:

$$(\partial_{x_+} + \frac{1}{2}m^2 x_- K + m\mathcal{E}_+)\Psi = 0, \quad (\partial_{x_-} - \frac{1}{2}m^2 x_+ K + m\mathcal{E}_-)\Psi = 0$$

The first equation is readily solved by $\Psi = \exp(-\frac{1}{2}m^2 x_+ x_- K)\exp(-m x_+ \mathcal{E}_+)\widetilde{\Psi}(x_-)$. Inserting this in the second equation, and using $[\mathcal{E}_+, \mathcal{E}_-] = K$, which implies $\exp(m x_+ \mathrm{ad}\mathcal{E}_+)\mathcal{E}_- = \mathcal{E}_- + m x_+ K$, one gets $(\partial_{x_-} + m\mathcal{E}_-)\widetilde{\Psi} = 0$. We finally obtain:

$$\Psi_{vac}(x,t) = e^{-\frac{m^2}{2}x_+x_-K} e^{-mx_+\mathcal{E}_+} e^{-mx_-\mathcal{E}_-} = e^{\frac{m^2}{2}x_+x_-K} e^{-mx_-\mathcal{E}_-} e^{-mx_+\mathcal{E}_+} \tag{12.18}$$

where the two forms are equivalent thanks to the Campbell–Haussdorf formula. One can then compute the chiral fields:

$$\xi_{vac}(x_+) = <\Lambda| e^{-mx_+\mathcal{E}_+}, \quad \overline{\xi}_{vac}(x_-) = e^{mx_-\mathcal{E}_-}|\Lambda> \tag{12.19}$$

The reconstruction formula for Φ reads $\exp(-2\Lambda(\Phi)) = \xi_{vac}(x_+).\overline{\xi}_{vac}(x_-) = \exp\Big(-m^2 x_+ x_- \Lambda(K)\Big)$ as it should be. The τ–functions of the vacuum solution are:

$$\tau_+ = \tau_- = \tau_0 = \exp[-m^2 x_+ x_-] \tag{12.20}$$

Remark. Notice that we can embed the vacuum equations of motion $\partial_{x_+}\partial_{x_-}\zeta_{vac} = 2m^2$ into a larger hierarchy. Introduce the variables $x_\pm^{(r)}$ for r odd, and consider the following connexion, generalizing eqs.(12.13) evaluated at Φ_{vac},

$$A^{vac}_{x_\pm^{(r)}} \ = \ \pm\frac{1}{4}\partial_{x_\pm^{(r)}}\zeta \ K + m\mathcal{E}_\pm^{(r)}$$

where $\mathcal{E}_\pm^{(r)} = \lambda^{\pm r}(E_+ + E_-)$, with r odd. Since $[\mathcal{E}_+^{(r)}, \mathcal{E}_-^{(s)}] = K r \delta_{rs}$, the zero curvature condition reduces to $\partial_{x_+^{(r)}}\partial_{x_-^{(s)}}\zeta_{vac} = 2m^2 r\delta_{rs}$. A solution is $\zeta_{vac} = 2m^2\sum_r r \ x_+^{(r)} x_-^{(r)}$. In the same way as before, we can calculate $\Psi_{vac}(x_+^{(r)}, x_-^{(r)})$

$$\begin{aligned}
\Psi_{vac}(x_+^{(r)}, x_-^{(r)}) \ &= \ e^{-\frac{m^2}{2}\left(\sum_r r x_+^{(r)} x_-^{(r)}\right)K} e^{-m\sum_r x_+^{(r)}\mathcal{E}_+^{(r)}} e^{-m\sum_r x_-^{(r)}\mathcal{E}_-^{(r)}} \\
&= \ e^{\frac{m^2}{2}\left(\sum_r r x_+^{(r)} x_-^{(r)}\right)K} e^{-m\sum_r x_-^{(r)}\mathcal{E}_-^{(r)}} e^{-m\sum_r x_+^{(r)}\mathcal{E}_+^{(r)}}
\end{aligned}$$

The vacuum chiral fields are thus:

$$\xi_{vac}(x_+^{(r)}) = <\Lambda| e^{-m\sum_r x_+^{(r)}\mathcal{E}_+^{(r)}}, \quad \overline{\xi}_{vac}(x_-^{(r)}) = e^{m\sum_r x_-^{(r)}\mathcal{E}_-^{(r)}}|\Lambda> \tag{12.21}$$

The $x_\pm^{(r)}$ are the collection of elementary times defining the sinh–Gordon hierarchy, see Chapter [2]. Each chirality is attached separately to the poles at $\lambda = 0, \infty$ in the loop representation.

12.3 Dressing transformations and soliton solutions.

As shown in Chapter [11], dressing transformation by an element $g = g_-^{-1} g_+$ of the dressing group, act on the chiral fields ξ and $\overline{\xi}$ by:

$$\xi^g(x,t) = \xi(x,t) \cdot g_-^{-1}, \quad \overline{\xi}^g(x,t) = g_+ \cdot \overline{\xi}(x,t)$$

In our case the group G is an infinite dimensional group whose Lie algebra is the affine algebra $\widehat{sl_2}$, namely we consider G to be $G = (\exp \widehat{\mathcal{N}}_-)(\exp \widehat{\mathcal{H}})(\exp \widehat{\mathcal{N}}_+)$. The dressing transformations are defined by the factorization problem

$$g = g_-^{-1}\, g_+ \quad \text{with} \quad g_\pm \in B_\pm = (\exp \widehat{\mathcal{H}})(\exp \widehat{\mathcal{N}}_\pm)$$

and moreover we require that g_- and g_+ have inverse components on the Cartan torus.

When we dress the vacuum solution with an element $g = g_-^{-1} g_+$ we get a new solution Φ^g such that:

$$e^{-2\Lambda(\Phi^g)} = \langle \Lambda | e^{-mx_+ \mathcal{E}_+} g e^{mx_- \mathcal{E}_-} | \Lambda \rangle$$

NOTE: Rapport avec la Grassmannienne infinie, orbite du vide.

To construct new solutions of the sinh–Gordon equation, one has to choose particular elements g of the affine group. Remarkable elements can be constructed with vertex operators in the two level 1 representations of the affine $\widehat{sl_2}$ algebra. Let us recall here the main facts, and refer to Chapter [16] for more details. One introduces bosonic oscillators p_n for n odd, such that $[p_m, p_n] = m\delta_{n+m,0}$ and $p_n^\dagger = p_{-n}$. They generate a Fock space over the vacuum state $|0\rangle$ which is specified by $p_n|0\rangle = 0$ for $n > 0$. Let $Z(\lambda)$ be the generating function:

$$Z(\lambda) = -i\sqrt{2} \sum_{n\ odd} p_{-n} \frac{\lambda^n}{n}$$

The level one vertex operator representations with highest weight $\Lambda_\pm$ of the Lie algebra $\widehat{sl_2}$ are then obtained by setting:

$$\begin{aligned} \sum_{n\ odd} \lambda^{-n}(E_+^n + E_-^n) &= \frac{i}{\sqrt{2}} \lambda \frac{d}{d\lambda} Z(\lambda) \\ \sum_{n\ even} \lambda^{-n} H^n + \sum_{n\ odd} \lambda^{-n}(E_+^n - E_-^n) &= \pm V(\lambda) \end{aligned} \tag{12.22}$$

where $V(\lambda)$ denotes the vertex operator:

$$V(\lambda) = \tfrac{1}{2} : e^{-i\sqrt{2}Z(\lambda)} := \tfrac{1}{2} : e^{i\sqrt{2}Z(-\lambda)} :$$

The double dots means that the expressions are normal ordered, i.e. we write all p_n with $n > 0$ to the right. The representations with highest weight $\Lambda^\pm$ correspond to the plus and minus sign respectively, in eq.(12.22).

Recall the formula, see eq.(16.24):

$$V(\lambda)V(\mu) = \left(\frac{\lambda - \mu}{\lambda + \mu} \right)^2 : V(\lambda)V(\mu) : \quad |\lambda| > |\mu|$$

so, inside a correlation function $V^2(\mu) = 0$ and $1 + aV(\mu)$ represents an element g of the dressing group, i.e. in the representation ρ_{Λ^+} we have $\rho_{\Lambda^+}(g) = 1 + aV(\mu)$, while in the representation ρ_{Λ^-} we have $\rho_{\Lambda^-}(g) = 1 - aV(\mu)$ due to the sign change in eq.(12.22). More generally

$$\rho_{\Lambda^\pm}(g) = (1 \pm 2a_1 V(\mu_1))(1 \pm 2a_2 V(\mu_2)) \cdots (1 \pm 2a_N V(\mu_N)) \tag{12.23}$$

represents an element g of the dressing group, in the highest weight level one representations $\Lambda^\pm$ respectively.

Proposition. *Let $\tau_\pm^{(N)}$ be the tau–functions:*

$$\tau_\pm^{(N)}(x_\pm) = \langle \Lambda^\pm | e^{-mx_+\mathcal{E}_+} \prod_{i=1}^{N} (1 \pm 2a_i V(\mu_i)) e^{mx_-\mathcal{E}_-} |\Lambda^\pm\rangle$$

Then we have:

$$\frac{\tau_\pm^{(N)}}{\tau_0} = 1 + \sum_{p=1}^{N} (\pm)^p \sum_{k_1<k_2<\cdots<k_p} X_{k_1}\cdots X_{k_p} \prod_{k_i<k_j} \left(\frac{\mu_{k_i}-\mu_{k_j}}{\mu_{k_i}+\mu_{k_j}}\right)^2 \tag{12.24}$$

with $X_i = a_i \exp\left(2m(\mu_i x_+ + \mu_i^{-1}x_-)\right)$, and $\tau_0 = \exp\left(-m^2 x_+ x_-\right)$. The parameters μ_i are interpreted as the rapidities of the solitons and a_i are related to their positions.

Proof. The commutation relations between $\mathcal{E}_\pm$ and $V(\lambda)$ are $[\mathcal{E}_\pm, \mathcal{V}(\lambda)] = -\in\lambda^{\pm\infty}\mathcal{V}(\lambda)$ see eq.(16.26) in Chapter [16]. They imply that $e^{-mx_\pm\mathcal{E}_\pm}V(\mu_i)e^{mx_\pm\mathcal{E}_\pm} = e^{2m\mu_i^{\pm1}x_\pm}V(\mu_i)$. Commuting the $e^{-mx_+\mathcal{E}_+}$ to the right, and $e^{mx_-\mathcal{E}_-}$ to the left amounts to the replacement $a_i \to X_i$ and moreover produces the factor τ_0 since $[\mathcal{E}_+, \mathcal{E}_-] = K$, so that we get:

$$\tau_\pm^{(N)} = \tau_0 \langle\Lambda^\pm| \prod_{i=1}^{N}(1 + 2X_i V(\mu_i))|\Lambda^\pm\rangle$$

The explicit formula (12.24) then follows from Wick's theorem applied to the vertex operators $V(\mu_i)$. ■

Remark. One can incorporate the whole hierarchy of times, as in eq.(12.19) to this expression, using the commutation relations $[E_+^n + E_-^n, V(\lambda)] = -2\lambda^n V(\lambda)$. Then X_i is replaced by

$$X_i = a_i \exp\left(2m \sum_s (\mu_i^s x_+^{(s)} + \mu_i^{-s} x_-^{(s)})\right)$$

and τ_0 by $\exp\left(-m^2 \sum_s s x_+^{(s)} x_-^{(s)}\right.$.

The N-soliton tau functions can also be written as the determinant of a $N \times N$ matrix:

$$\frac{\tau_\pm^{(N)}}{\tau_0} = \det(1 \pm \mathcal{V}) \quad \text{with} \quad \mathcal{V}_{ij} = 2\frac{\sqrt{\mu_i\mu_j}}{\mu_i+\mu_j}\sqrt{X_i X_j} \tag{12.25}$$

This is proved by using the Cauchy identity:

$$\det\left(2\frac{\sqrt{\mu_i\mu_j}}{\mu_i+\mu_j}\right) = \prod_{i<j}\left(\frac{\mu_i-\mu_j}{\mu_i+\mu_j}\right)^2$$

and the expansion formula $\det(1+\mathcal{V}) = 1 + \sum_{p=1}^N \sum_{i_1<i_2<\cdots<i_p} \det \mathcal{V}(i_1, i_2, \cdots, i_p)$ where $\mathcal{V}(i_1, i_2, \cdots, i_p)$ is the submatrix of $\mathcal{V}$ with lines and columns $i_1, i_2, \cdots, i_p$.

For the sine–Gordon equation, the parameters a_i, μ_i, in eq.(12.23) satisfy specific reality conditions. For solitons or antisolitons, the rapidity μ is real and a is purely imaginary, i.e. $a = -i\epsilon e^\gamma$ with $\epsilon = 1$ for a soliton and $\epsilon = -1$ for an antisoliton. Indeed in this case, φ has to be replaced by $i\varphi$ so the one–soliton solution reads

$$e^{-i\varphi} = \frac{\tau_+}{\tau_-} = \frac{1+X}{1-X}$$

with $X = a\exp 2m((\mu+\mu^{-1})x + (\mu-\mu^{-1})t)$. For this to be of modulus 1, one takes $a = -i\epsilon e^\gamma$. At fixed t, when x goes from $-\infty$ to $+\infty$, φ goes from 0 to $\epsilon\pi$. The topological charge is defined as $Q = \frac{1}{\pi}(\varphi(+\infty) - \varphi(-\infty))$, so that solitons have charge +1 and antisolitons charge -1.

The “breathers” correspond to pairs of complex conjugated rapidities ($\mu_1 = \mu, \mu_2 = \overline{\mu}$) and positions ($X_1 = X, X_2 = -\overline{X}$). In this case we have:

$$e^{-i\varphi} = \frac{\tau_+}{\tau_-} = \frac{1 + X_1 + X_2 + \left(\frac{\mu_1-\mu_2}{\mu_1+\mu_2}\right)^2 X_1 X_2}{1 - X_1 - X_2 + \left(\frac{\mu_1-\mu_2}{\mu_1+\mu_2}\right)^2 X_1 X_2} = \frac{1 + 2i \operatorname{Im} X + \left(\frac{\operatorname{Im}\mu}{\operatorname{Re}\mu}\right)^2 |X|^2}{1 - 2i \operatorname{Im} X + \left(\frac{\operatorname{Im}\mu}{\operatorname{Re}\mu}\right)^2 |X|^2}$$

This shows that $\varphi(x = -\infty) = \varphi(+\infty) = 0$, moreover if we perform a Lorentz boost to go to the rest frame where $|\mu| = 1$ we see that φ is periodic in time, hence the denomination of this solution, of topological charge 0.

12.4 N–soliton dynamics.

The N–solitons solutions are a finite dimensional submanifold in the infinite dimensional phase space of sine–Gordon, to which the symplectic form of sine–Gordon can be restricted. This yields a finite dimensional symplectic manifold. The soliton solutions then become a relativistically invariant N–body problem first introduced by Ruijsenaars, which is a relativistic generalization of the Calogero model.

Consider the *sine*–Gordon model with equation of motion:

$$\partial_+ \partial_- \varphi = 2m^2 \sin(2\varphi) \tag{12.26}$$

The natural symplectic form is the canonical one:

$$\Omega_{SG} = \int_{-\infty}^{+\infty} dx \ \delta\pi(x) \wedge \delta\varphi(x)$$

where $\pi(x) = \partial_t \varphi(x)$ is the momentum conjugate to the field $\varphi(x)$, and δ denote the differential on the phase space. The N–soliton solution is given by $e^{-i\varphi} = \frac{\tau_+}{\tau_-}$ where $\tau_\pm$ are given by eq.(12.24). These solutions are parametrized by the $2N$ parameters a_i, μ_i which therefore provide coordinates on the N–soliton submanifold. In this section we will present the proofs in the *soliton* case. The analysis of antisolitons and breathers can be done with some technicalities. This means that the μ_i are real, and the a_i are pure imaginary. We now express the restriction of Ω_{SG} on this submanifold.

Proposition. *In the coordinates a_i and μ_i, the restriction of Ω_{SG} is given by:*

$$\omega = \sum_{i=1}^{N} \frac{\delta a_i}{a_i} \wedge \frac{\delta \mu_i}{\mu_i} + \sum_{i<j} \left(\frac{4\mu_i \mu_j}{\mu_i^2 - \mu_j^2} \right) \frac{\delta \mu_i}{\mu_i} \wedge \frac{\delta \mu_j}{\mu_j} \tag{12.27}$$

Proof. Consider first the one soliton solutions. In this case, the computation can be done directly using the formula $e^{-i\phi} = (1 + X)/(1 - X)$, which gives $\pi = 2i\dot{X}/(1 - X^2)$. One finds:

$$\Omega_{SG}\Big|_{\text{restricted}} = \left[-2 \int_{-\infty}^{+\infty} dx \partial_x \left(\frac{1}{1 - X^2} \right) \right] \cdot \frac{\delta\mu}{\mu} \wedge \frac{\delta a}{a} \tag{12.28}$$

Since X is pure imaginary, and $X(+\infty) = \infty$ and $X(-\infty) = 0$ the integral evaluates to an irrelevant numerical constant. So the formula is proved for 1 soliton. Consider next the general case with an arbitrary number of solitons with parameters (μ_i, a_i). Since the symplectic form can be computed at any time we evaluate it at $t \to \pm\infty$. In these limits the sine–Gordon field becomes asymptotically equal to the sum of one–soliton solutions with parameters $(\mu_i^{\text{in}}, a_i^{\text{in}})$ and $(\mu_i^{\text{out}}, a_i^{\text{out}})$:

$$\begin{aligned} \mu_i^{\text{in}} &= \mu_i^{\text{out}} = \mu_i \\ a_i^{\substack{\text{in} \\ \text{out}}} &= a_i \prod_{|\mu_j| \lessgtr |\mu_i|} \left(\frac{\mu_i - \mu_j}{\mu_i + \mu_j} \right)^2 \end{aligned}$$

These formulae are obtained as follows. Let us look at the quantities $X_i = a_i \exp(2m(\mu_i x_+ + \mu_i^{-1} x_-))$ entering the definition of the N soliton solution. When t is large, X_i is of order 1 when x is located at $x \sim x_i = -\frac{\mu_i - \mu_i^{-1}}{\mu_i + \mu_i^{-1}} t$. The values of the other X_j's when $x \sim x_i$ are $X_j \sim a_j e^{4m(\mu_j^2 - \mu_i^2)t/\mu_i\mu_j(\mu_i+\mu_i^{-1})}$. Hence for large positive time, if $\mu_j^2 > \mu_i^2$, $X_j(x_i, t)$ is very large, while if $\mu_j^2 < \mu_i^2$, $X_j(x_i, t)$ is very small. So we can split the N solitons into two subsets. One, $\mathcal{I}_+$, is such that $\mu_j^2 > \mu_i^2$, and the other one, $\mathcal{I}_-$, is such that $\mu_j^2 < \mu_i^2$. We can evaluate the tau functions eq.(12.25) when $x \sim x_i$. One has to keep the terms containing the maximum number of X_j's, $j \in \mathcal{I}_+$. These terms read:

$$\tau_\pm(x_i, t) \sim \prod_{j \in \mathcal{I}_+} X_j \prod_{k_j < k_i} \left(\frac{\mu_{k_i} - \mu_{k_j}}{\mu_{k_i} + \mu_{k_j}} \right)^2 \left[1 \pm X_i \prod_{j \in \mathcal{I}_+} \left(\frac{\mu_i - \mu_j}{\mu_i + \mu_j} \right)^2 \right]$$

where we have to keep the second term because $X_i \sim 1$. So we have

$$e^{-i\varphi} = \frac{\tau_+}{\tau_-} = \frac{1 + X_i \prod_{j \in \mathcal{I}_+} \left(\frac{\mu_i - \mu_j}{\mu_i + \mu_j} \right)^2}{1 - X_i \prod_{j \in \mathcal{I}_+} \left(\frac{\mu_i - \mu_j}{\mu_i + \mu_j} \right)^2}$$

which is indeed a one soliton formula, but with modified parameter a^{out}.

Since asymptotically the solitons decouple, in the symplectic form the crossed terms vanish (the overlap integrals are zero). Therefore the symplectic form reduces to the sum of the one-soliton expressions, but with the modified in and out parameters:

$$\omega = \sum_{i=1}^N \frac{\delta a_i^{\text{in}}}{a_i^{\text{in}}} \wedge \frac{\delta \mu_i}{\mu_i} = \sum_{i=1}^N \frac{\delta a_i^{\text{out}}}{a_i^{\text{out}}} \wedge \frac{\delta \mu_i}{\mu_i}$$

Both forms are equal to eq.(12.27), as can be checked easily. ■

The two-form (12.27) is non degenerate. It therefore defines a symplectic structure on the restricted phase space. The corresponding Poisson brackets are found by inverting the symplectic form (12.27):

$$\{\mu_i, \mu_j\} = 0, \quad \{\mu_i, a_j\} = \mu_i a_j \delta_{ij}, \quad \{a_i, a_j\} = \left(\frac{4\mu_i \mu_j}{\mu_i^2 - \mu_j^2} \right) a_i a_j \tag{12.29}$$

As a by product, we see that the transformation from the in–variables $(\mu_i^{\text{in}}, a_i^{\text{in}})$ to the out–variables $(\mu_i^{\text{out}}, a_i^{\text{out}})$ is symplectic.

Knowing the Poisson brackets between the parameters (μ_i, a_i), we can compute the Poisson brackets between the elements of the matrix $\mathcal{V}$ in eq.(12.25). Remarkably, they can be written in the r–matrix form:

$$\{\mathcal{V}_1, \mathcal{V}_2\} = [r_{12}, \mathcal{V}_1] - [r_{21}, \mathcal{V}_2] \tag{12.30}$$

where $r_{12} = \sum_{ij;kl} r_{ij;kl} E_{ij} \otimes E_{kl}$ and:

$$r_{ij;kl} = \frac{1}{8} \left(\frac{\mu_i + \mu_j}{\mu_i - \mu_j} \right) (\mathcal{V}_{jk}\delta_{il} + \mathcal{V}_{jl}\delta_{ik} + \mathcal{V}_{ik}\delta_{jl} + \mathcal{V}_{il}\delta_{jk})$$

Eq.(12.30) can be checked by straightforward computation.

From the definition of the sine–Gordon field in terms of the τ-functions, we see that it can be expressed only in terms of the eigenvalues Q_i of the matrix $\mathcal{V}$:

$$e^{-i\varphi} = \frac{\det(1 + \mathcal{V})}{\det(1 - \mathcal{V})} = \prod_{i=1}^N \left(\frac{1 + Q_i}{1 - Q_i} \right) \tag{12.31}$$

Recall that in the N–soliton case, we can assume $a_j = -ie^{\gamma_j}$ and $\mu_j > 0$ so that the matrix $\mathcal{V}$ is of the form $i\times$real symmetric matrix. It can be diagonalized with a real orthogonal matrix U. Moreover its eigenvalues Q_j are pure imaginary. Finally eq.(12.30) implies $\{Q_i, Q_j\} = 0$.

We want to express the evolution of solitons in terms of the variables Q_i. It is convenient to consider separately the evolution in the light–cone coordinates, and we consider for definiteness the evolution with respect to $x_+ = x + t$. Let U be the real orthogonal matrix which diagonalizes the symmetric matrix $\mathcal{V}$ and *define* a symmetric matrix L by:

$$\text{if} \quad \mathcal{V} = U^{-1}QU, \quad \text{then} \quad L = U\mu U^{-1} \tag{12.32}$$

where Q denotes the diagonal matrix diag(Q_i) and μ the diagonal matrix diag(μ_i) with μ_i the rapidities of the solitons. The matrix L plays the role of a Lax operator. Obviously, the quantities $tr(L^n) = \sum_{i=1}^N \mu_i^n$ are conserved during the evolution of the solitons. They are in involution under the Poisson bracket (12.29). The time evolution of L takes the Lax form, and we have:

Proposition. *The x_+ evolution of L is given by a Lax equation*

$$\dot{L} = [M, L] \quad \text{with} \quad M = \dot{U}U^{-1}$$

where the dot means $\frac{\partial}{\partial x_+}$. Moreover, L and M can be expressed in terms of the quantities Q_i and $\dot{Q}_i$:

$$L_{ij} = 2\frac{\sqrt{\dot{Q}_i\dot{Q}_j}}{Q_i + Q_j}, \quad \text{and} \quad M_{ij} = \frac{\sqrt{\dot{Q}_i\dot{Q}_j}}{Q_i - Q_j}(1 - \delta_{ij}) \tag{12.33}$$

If we introduce the coordinates $Q_j = ie^{q_j}$, the equations of motion read:

$$\ddot{q}_i = \sum_{k\neq i} \frac{2\dot{q}_i\dot{q}_k}{\sinh(q_i - q_k)} \tag{12.34}$$

Proof. The Lax form of the equation of motion follows from eq.(12.32). Let us prove eq.(12.33). We start from the relation $\mu\mathcal{V} + \mathcal{V}\mu = 2|e><e|$ where $|e>$ is the column vector with components $e_i = \sqrt{X_i\mu_i}$. Multiplying on the left by U and on the right by U^{-1} we get, using that U is orthogonal, $QL + LQ = 2|\tilde{e}><\tilde{e}|$ with $|\tilde{e}>= U\,|e>$. Since Q is diagonal, we obtain from this relation:

$$L_{ij} = 2\frac{\tilde{e}_i\tilde{e}_j}{Q_i + Q_j}$$

Next, since $\dot{\sqrt{X_i}} = \mu_i\sqrt{X_i}$, we have $\dot{\mathcal{V}} = \mu\mathcal{V} + \mathcal{V}\mu = |e><e| = U^{-1}\left(\dot{Q} + [Q, M]\right)U$, which implies $|\tilde{e}><\tilde{e}| = \dot{Q} + [Q, M]$. In components, this reads $\tilde{e}_i\tilde{e}_j = \dot{Q}_i\delta_{ij} + (Q_i - Q_j)M_{ij}$. If $i = j$, we find $\tilde{e}_i = \sqrt{\dot{Q}_i}$, and if $i \neq j$, we find the value of M_{ij} in terms of Q_i, $\dot{Q}_i$. This completely determines the antisymmetric matrix M. To write Q_j in the form ie^{q_j} remark that, since all μ_i are assumed positive, we have for real x_i:

$$\sum_{ij} \frac{x_i x_j}{\mu_i + \mu_j} = \int_0^\infty dt e^{-t(\mu_i+\mu_j)}x_i x_j = \int_0^\infty dt (\sum_j e^{-t\mu_j}x_j)^2 \geq 0$$

so that the eigenvalues Q_j of $\mathcal{V}$ are of the form $i\times$real positive number. The equations of motions eq.(12.34) are obtained by writing the Lax equation. ■

Instead of the variables $\dot{Q}_i$, let us introduce the variables $\rho_i = \dot{Q}_i/Q_i$. The Lax matrix becomes

$$L_{ij} = 2\frac{\sqrt{Q_i Q_j}}{Q_i + Q_j}\sqrt{\rho_i\rho_j}$$

Comparing the matrices $\mathcal{V}$ and L, on sees that L has exactly the same form as $\mathcal{V}$ with the change of variables $(\mu_i, a_i) \to (Q_i, \rho_i)$. This symmetry extends at the level of the symplectic structure.

Proposition. *The transformation $(\mu_i, a_i) \to (Q_i^{-1}, \rho_i)$ is a symplectic transformation.*

Proof. Let us consider

$$\omega_1 = \sum_{i=1}^{N} \frac{\delta a_i}{a_i} \wedge \frac{\delta\mu_i}{\mu_i} + \sum_{i<j} \frac{4\mu_i\mu_j}{\mu_i^2 - \mu_j^2} \frac{\delta\mu_i}{\mu_i} \wedge \frac{\delta\mu_j}{\mu_j}, \quad \omega_2 = \sum_{i=1}^{N} \frac{\delta\rho_i}{\rho_i} \wedge \frac{\delta Q_i}{Q_i} + \sum_{i<j} \frac{4Q_iQ_j}{Q_i^2 - Q_j^2} \frac{\delta Q_i}{Q_i} \wedge \frac{\delta Q_j}{Q_j}$$

We first express ω_1 using the independent variables μ_i, Q_i. We get $\omega_1 = \sum_{ij} A_{ij} \frac{\delta Q_i}{Q_i} \wedge \frac{\delta\mu_j}{\mu_j} + \sum_{i<j} B_{ij} \frac{\delta\mu_i}{\mu_i} \wedge \frac{\delta\mu_j}{\mu_j}$. Recall that, using the symplectic form ω_1 we have shown that $\mathcal{V}$ obeys an r–matrix relation, hence its eigenvalues Poisson commute. Noting that

$$\begin{pmatrix} 0 & A \\ -A^t & B \end{pmatrix}^{-1} = \begin{pmatrix} (A^t)^{-1}BA^{-1} & -(A^t)^{-1} \\ A^{-1} & 0 \end{pmatrix}$$

we see that the vanishing of the Poisson brackets $\{Q_i, Q_j\}$ requires $B_{ij} = 0$. To compute $A_{ij} = \frac{Q_i}{a_j}\frac{\partial a_j}{\partial Q_i}$ (at μ_k fixed) let us vary the matrix $\mathcal{V}$ in eq.(12.25) at μ fixed, getting $d\mathcal{V} = \frac{1}{2}(\mathcal{V}a^{-1}da + a^{-1}da\mathcal{V})$, where $a = \mathrm{Diag}(a_i)$. Similarly varying $\mathcal{V} = U^{-1}QU$ we get $d\mathcal{V} = U^{-1}(dQ + [Q, dUU^{-1}])U$ so that:

$$dQ + [Q, dUU^{-1}] = \frac{1}{2}\Big(Ua^{-1}daU^{-1}Q + QUa^{-1}daU^{-1}\Big)$$

Looking at the diagonal part of this equation, we get

$$\frac{dQ_i}{Q_i} = \sum_j U_{ij}U_{ji}^{-1}\frac{da_j}{a_j} = \sum_{j,k} U_{ij}U_{ji}^{-1}\frac{Q_k}{a_j}\frac{\partial a_j}{\partial Q_k}\frac{dQ_k}{Q_k}$$

so that, setting $\mathcal{M}_{ij} = U_{ji}^{-1}U_{ij}$, we have $A\mathcal{M} = 1$ or $A = \mathcal{M}^{-1}$.

We now express the form ω_2 in terms of the variables (Q_i, μ_i) exactly in the same way. In particular, starting from ω_2 one may deduce an r–matrix relation for L implying that $\{\mu_i, \mu_j\} = 0$ so that $\omega_2 = \sum_{ij} C_{ij} \frac{\delta Q_i}{Q_i} \wedge \frac{\delta\mu_j}{\mu_j}$ with $C_{ij} = -\frac{\mu_j}{\rho_i}\frac{\partial\rho_i}{\partial\mu_j}$ (at Q_k fixed). The computation of C_{ij} is the same as for A_{ij} except that U is replaced by U^{-1}, which leads to $\mathcal{M}C = -1$ so that $C = -A$, and $\omega_2 = -\omega_1$. Changing Q_i to Q_i^{-1} changes ω_2 to $-\omega_2$ thereby proving the symplecticity of the transformation. ■

It is convenient to introduce a new set of variables p_i by:

$$\rho_i = \exp(p_i)\prod_{k\neq i}\left(\frac{Q_k + Q_i}{Q_k - Q_i}\right) = e^{p_i}\prod_{k\neq i}\coth\left(\frac{q_i - q_k}{2}\right) \tag{12.35}$$

It is easy to check that the variables p_i, q_i are canonically conjugated:

$$\{q_i, q_j\} = 0, \quad \{p_i, p_j\} = 0, \quad \{p_i, q_j\} = \delta_{ij}$$

We have already noticed that L obeys an r–matrix relation, which implies that the function $\mathcal{T}(x) = \det(1 + xL)$ is a generating function of commuting Hamiltonians (which we already knew by construction of L)

$$\{\mathcal{T}(x), \mathcal{T}(y)\} = 0$$

Expanding $\mathcal{T}(x)$ in power of x, we find:

$$\mathcal{T}(x) = 1 + \sum_{p=1}^{N} x^p H_p = 1 + \sum_{p=1}^{N} x^p \sum_{k_1<\cdots<k_p} \rho_{\rho_{k_1}}\cdots\rho_{k_p}\prod_{k_i<k_j}\left(\frac{Q_{k_i} - Q_{k_j}}{Q_{k_i} + Q_{k_j}}\right)^2 \tag{12.36}$$

The Hamiltonian generating the evolution in the light cone coordinate x_+ is $H_+ = \mathrm{Tr}\,(L)$. The evolution in the other light cone coordinate x_- is generated by the inverse of the Lax matrix. Its Hamiltonian is $H_- = \mathrm{Tr}\,(L^{-1}) = H_{N-1}/\det(L)$.

$$H_\pm = \sum_j e^{\pm p_j} \prod_{k\neq j} \coth\left(\frac{q_j - q_k}{2}\right)$$

These are the Hamiltonians written by Ruijsenaars and Schneider. By construction, their corresponding flows are linearized in the variables (μ_i, X_i).

We end up this section by presenting the reality conditions corresponding to solitons, antisolitons and breathers. For the sine-Gordon field (12.31) and the Hamiltonians $H_\pm$ to be real, the coordinates (Q_i, p_i) have to come by pairs $(j, \overline{j})$, with $Q_{\overline{j}} = -\overline{Q}_j$ and $p_{\overline{j}} = \overline{p}_j$. The case $j = \overline{j}$ corresponds to a soliton or an antisoliton, and the case $j \neq \overline{j}$ to a breather. Therefore, in the coordinates q_j such that $Q_j = -ie^{q_j}$:

$$\begin{aligned} \mathrm{Im}\ q_s &= 0 && \text{for } s \text{ a soliton} \\ \mathrm{Im}\ q_{\overline{s}} &= \pi && \text{for } \overline{s} \text{ an antisoliton} \\ q_{\overline{b}} &= \overline{q}_b && \text{for } b \text{ a breather} \end{aligned}$$

Similarly the momenta p_s and $p_{\overline{s}}$ are real and p_b is complex with $p_{\overline{b}} = \overline{p}_b$.

12.5 Finite zone solutions.

The simplest example of finite zone solutions is the one–zone solution. It is obtained by looking at solutions $\varphi = \varphi(\eta)$ of the sinh–Gordon equation depending only on the variable $\eta = m(\mu x_+ + \mu^{-1}x_-)$. Then $\partial_{x_+}\partial_{x_-}\varphi = 2m^2\sinh(2\varphi)$ becomes $\varphi'' = e^{2\varphi} - e^{-2\varphi}$ where $'$ means the derivative with respect to η. This integrates to $\varphi'^2 = e^{2\varphi} + e^{-2\varphi} - 2E$, where E is a constant. Setting $e^{-2\varphi} = t$, this equation takes the form:

$$t'^2 = 4t(t^2 - 2Et + 1)$$

For $E = \pm 1$ the right–hand side is a perfect square and the equation is readily solved. For $E = +1$ one finds the soliton and antisoliton solutions. For general E, the equation can be compared with the equation for Weierstrass $\wp$ function:

$$\wp'(\eta)^2 = 4\wp^3(\eta) - g_2\wp(\eta) - g_3$$

and we have:

$$e^{-2\varphi} = t(\eta) = \wp(\eta - \eta_0) + \frac{2E}{3}, \quad g_2 = \frac{16}{3}E^2 - 4, \quad g_3 = \frac{8}{3}E(\frac{8}{9}E^2 - 1)$$

and η_0 is an arbitrary constant. Note that the discriminant $g_2^3 - 27g_3^2 = 64(E^2 - 1)$ vanishes only for the cases $E = \pm 1$.

The situation is simple enough to allow for a complete solution in terms of Weierstrass elliptic functions. Consider the elliptic curve

$$\Gamma : y^2 = 4x(x^2 - 2Ex + 1) \tag{12.37}$$

Let us assume $E > 1$ so that we can write $y^2 = 4x(x - a)(x - b)$ with $0 < a < b$ real. Then the periods $2\omega_1$ and $2\omega_2$ are real and pure imaginary respectively. The uniformizing variable is

$$z(x, y) = \int_{P(0,0)}^{P(x,y)} \frac{dx}{y}$$

The two half periods are

$$\omega_1 = \int_0^a \frac{dx}{\sqrt{4x(a-x)(b-x)}} \in \mathbb{R}^+; \quad \omega_2 = \int_0^\infty \frac{dx}{\sqrt{4x(a+x)(b+x)}} \in i\mathbb{R}^+;$$

The images of the four branch points are $z(0,0) = 0$, $z(\infty,\infty) = \omega_2$, $z(a,0) = \omega_1$, and $z(b,0) = \omega_1 + \omega_2$. The expressions of x and y in terms of z are

$$x = \wp(z-\omega_2) - \wp(\omega_2), \quad y = \wp'(z-\omega_2);$$

This is fixed from Weierstrass differential equation for the $\wp$ function, noting that $x = 0$ when $z = 0$ and $x = \infty$ when $z = \omega_2$. In particular, we find $\wp(\omega_2) = -\frac{2E}{3}$.

Proposition. *The one–zone solution is given by:*

$$e^{-2\varphi} = \wp(z-\omega_2) - \wp(\omega_2), \quad e^{2\varphi} = \wp(z) - \wp(\omega_2), \quad z = \eta - \eta_0$$

Proof. The formula for $e^{-2\varphi}$ is obtained by identifying $t = x$, $t' = y$ in eq.(12.37). Let us now prove the expression for $e^{2\varphi}$. Since $e^{-2\varphi}$ has a double pole at $z = \omega_2$ and a double zero at $z = 0$, the inverse $e^{2\varphi}$ has a double pole at $z = 0$ and a double zero at $z = \omega_2$. Hence $e^{2\varphi} = C^{\mathrm{st}}\,(\wp(z) - \wp(\omega_2))$. Writing $e^{-2\varphi}e^{2\varphi} = 1$ near $z = 0$, we find $C^{\mathrm{st}} = [\frac{1}{2}\wp''(\omega_2)]^{-1}$. In fact $\frac{1}{2}\wp''(\omega_2) = 1$. Indeed, $\wp''(\omega_2) = y'|_{z=0}$. Differentiating the equation of the curve Γ, we get $2yy' = 4x'((x-a)(x-b) + x(x-b) + x(x-a))$. When $z \to 0$, we have $x \to 0$ and remembering that $y = x'$, we find $y'(z=0) = 2$. ∎

To see how this simple one-zone solution fits into the general formalism of finite zone solutions, we need to construct a Lax matrix L satisfying the compatibility conditions

$$[\partial_{x_+} + A_{x_+}, L] = 0, \quad [\partial_{x_-} + A_{x_-}, L] = 0$$

As explained in Chapter [2], the general technique is to impose a stationnarity condition with respect to some time. The one–zone solution only depends on $\eta = m(\mu x_+ + \mu^{-1}x_-)$ hence satisfies the stationarity condition:

$$(\mu^{-1}\partial_{x_+} - \mu\partial_{x_-})\varphi = 0$$

So we choose as Lax matrix:

$$L(\lambda) = \frac{1}{m}(\mu^{-1}A_{x_+} - \mu A_{x_-}) = \begin{pmatrix} \frac{1}{2m}(\mu^{-1}\partial_{x_+}\varphi + \mu\partial_{x_-}\varphi) & \frac{\lambda}{\mu}e^{\varphi} - \frac{\mu}{\lambda}e^{-\varphi} \\ \frac{\lambda}{\mu}e^{-\varphi} - \frac{\mu}{\lambda}e^{\varphi} & -\frac{1}{2m}(\mu^{-1}\partial_{x_+}\varphi + \mu\partial_{x_-}\varphi \end{pmatrix}$$

The spectral curve is now easy to construct. By definition $\Gamma : \det(L(\lambda) - \Lambda) = 0$, or explicitly

$$\Lambda^2 = \frac{1}{4m^2}(\mu^{-1}\partial_{x_+}\varphi + \mu\partial_{x_-}\varphi)^2 - e^{2\varphi} - e^{-2\varphi} + \frac{\lambda^2}{\mu^2} + \frac{\mu^2}{\lambda^2}$$

Hence, for these special one–zone solutions, the quantity $\frac{1}{4m^2}(\mu^{-1}\partial_{x_+}\varphi + \mu\partial_{x_-}\varphi)^2 - e^{2\varphi} - e^{-2\varphi}$ is a constant independent of $x_\pm$. Indeed, noting that $\mu^{-1}\partial_{x_+}\varphi + \mu\partial_{x_-}\varphi = 2m\varphi'$, we find that this constant is $-2E$. So the curve finally reads, as an equation in the (Λ, λ) plane:

$$\Gamma : \quad \Lambda^2 = \frac{\lambda^2}{\mu^2} + \frac{\mu^2}{\lambda^2} - 2E$$

Besides the hyperelliptic involution, $\Lambda \to -\Lambda$, this curve has another involution $\lambda \to -\lambda$. Let Γ' be the curve obtained from Γ by factorizing this involution. This is done by setting

$$Y' = \mu\lambda^2\Lambda, \quad \lambda' = \lambda^2 \quad \text{or} \quad \Lambda = \frac{Y'}{\mu\lambda'}, \quad \lambda = \sqrt{\lambda'}$$

Notice that the transformation $\Gamma' \to \Gamma$ is one to two. Γ is a two sheeted unbranched covering of Γ'. The equation of the curve Γ' reads

$$\Gamma' : \quad Y'^2 = \lambda'(\lambda'^2 - 2E\mu^2\lambda' + \mu^4)$$

The points $P_+(\lambda' = 0)$ and $P_-(\lambda' = \infty)$ are branch points on Γ'. Their pre-images on Γ are two singular points which blow–up into four points on its desingularization. This makes the situation more complicated than on Γ', and this is why we will work on Γ', and eventualy lift the results to Γ if needed.

To proceed we need to describe the singularity structure of the wave function Ψ such that:

$$(\partial_{x_\pm} + A_{x_\pm})\Psi = 0, \quad (L(\lambda) - \Lambda)\Psi = 0 \tag{12.38}$$

These equations are naturally written on Γ and we need to port them on Γ'. This is achieved by performing a gauge transformation:

$$\Psi' = g\Psi = \begin{pmatrix} e^{\frac{\varphi}{2}} & 0 \\ 0 & \frac{1}{\lambda}e^{-\frac{\varphi}{2}} \end{pmatrix} \Psi$$

where the important factor is $1/\lambda$ and the factors $e^{\pm\varphi/2}$ have been introduced for later convenience. The gauge transformed connection $A'_{x_\pm}$ and the new Lax matrix $L' = gLg^{-1}$ still obey the Lax equations $[\partial_{x_\pm} + A'_{x_\pm}, L'] = 0$, but are now rational functions of λ'.

Explicitly the transformed connection is:

$$A'_{x_+} = \begin{pmatrix} 0 & m\lambda' e^{2\varphi} \\ me^{-2\varphi} & 0 \end{pmatrix}, \quad A'_{x_-} = \begin{pmatrix} -\partial_{x_-}\varphi & m \\ m\lambda'^{-1} & \partial_{x_-}\varphi \end{pmatrix} \tag{12.39}$$

moreover the transformed L' is still equal to $\frac{1}{m}(\mu^{-1}A'_{x_+} - \mu A'_{x_-})$ thanks to $\mu^{-1}\partial_{x_+}\varphi - \mu\partial_{x_-}\varphi = 0$. It is now straightforward to extract the singular parts of Ψ' at the branch points $\lambda' = 0$ and $\lambda' = \infty$.

In the neighborhood of the point $\lambda' = \infty$ of Γ' we can use the local parameter $\sqrt{\lambda'}$ and the corresponding eigenvalue of L' is $\frac{1}{\mu}\sqrt{\lambda'}$. The eigenvector is of the form $(\chi_1, \chi_2/\sqrt{\lambda'})$ with $\chi_1/\chi_2 = \exp(2\varphi) + O(1/\sqrt{\lambda'})$. Similarly, at the point $\lambda' = 0$ the eigenvalue of L' is $\mu/\sqrt{\lambda'} + O(\sqrt{\lambda'})$ and $\psi'_2/\psi'_1 \sim -1/\sqrt{\lambda'}$. As explained in Chapter [4], since Ψ' is an eigenvector of L', and $L' = \frac{1}{m}(\mu^{-1}A'_{x_+} - \mu A'_{x_-})$ we have $A'_{x_+}\Psi' = m\sqrt{\lambda'}\Psi' + O(1)$ around ∞. Similarly around $\lambda' = 0$ we have $A'_{x_-}\Psi' = -m/\sqrt{\lambda'}\Psi' + O(1)$. To get the precise form of Ψ' at the point $\lambda' = 0$, write $\Psi'_0 = \exp(\frac{m}{\sqrt{\lambda'}}x_-)(f, g/\sqrt{\lambda'})$ with $g = f + O(\sqrt{\lambda'})$ and plug into the equations of motion $(\partial_{x_\pm} + A'_{x_\pm})\Psi' = 0$. A careful analysis of these equations shows that f is a *constant*, independent of $x_\pm$ and so can be normalized to 1. We finally get:

$$\Psi'_0 \sim e^{\frac{m}{\sqrt{\lambda'}}x_-} \begin{pmatrix} 1 \\ \frac{1}{\sqrt{\lambda'}} \end{pmatrix}, \quad \Psi'_\infty \sim e^{-m\sqrt{\lambda'}x_+} \begin{pmatrix} \chi_1 \\ \frac{1}{\sqrt{\lambda'}}\chi_2 \end{pmatrix}, \quad \frac{\chi_1}{\chi_2} = e^{2\varphi} \tag{12.40}$$

We see that Ψ' has a pole at $\lambda' = 0$ on Γ'.

The previous analysis can be generalized immediately to provide finite zone solutions to the sinh–Gordon equation. Consider an arbitrary hyperelliptic curve Γ', of genus g, having $\lambda' = 0$ and $\lambda' = \infty$ as branch points:

$$Y'^2 = \lambda' P(\lambda'), \quad \deg P = 2g$$

On this curve we construct Baker–Akhiezer vectors Ψ' with $g + N - 1 = g + 1$ poles (N=2). We set one of these poles at $\lambda' = 0$. The other g poles at finite distance of Ψ' are independent of $x_\pm$. The essential singularities are set at $P_+(\lambda' = 0)$ and $P_-(\lambda' = \infty)$ with behavior specified in eq.(12.40), so ψ'_2 has a pole at P_+ and a zero at P_-. This characterise Ψ' uniquely.

Proposition. *We have:*

$$\begin{aligned} \left[\partial_{x_+} + \begin{pmatrix} 0 & m\lambda'\frac{\chi_1}{\chi_2} \\ m\frac{\chi_2}{\chi_1} & 0 \end{pmatrix}\right] \begin{pmatrix} \psi'_1 \\ \psi'_2 \end{pmatrix} &= 0 \\ \left[\partial_{x_-} + \begin{pmatrix} -\partial_{x_-}\log\chi_1 & m \\ \frac{m}{\lambda'} & -\partial_{x_-}\log\chi_2 \end{pmatrix}\right] \begin{pmatrix} \psi'_1 \\ \psi'_2 \end{pmatrix} &= 0 \end{aligned} \tag{12.41}$$

Proof. Consider $\partial_+\psi_1' + m\frac{\chi_1}{\chi_2}\lambda'\psi_2'$. It is a Baker-Akhiezer function. At P_-, it behaves like $e^{-m\sqrt{\lambda'}x_+}O(1)$, while at P_+ it behaves like $e^{-\frac{m}{\sqrt{\lambda'}}x_-}O(\sqrt{\lambda'})$ and therefore has a zero at this point. It has also g poles at finite distance inherited from Ψ'. So it is a Baker–Akhiezer function with g poles and a prescribed zero at $\lambda' = 0$ which therefore vanishes. We prove similarly that $\partial_{x_+}\psi_2' + \frac{\chi_2}{\chi_1}\psi_1' = 0$ and that $\partial_{x_-}\psi_1' + m\psi_2 - \frac{\partial_{x_-}\chi_1}{\chi_1}\psi_1' = 0$. Finally considering $\partial_{x_-}\psi_2' + \frac{m}{\lambda'}\psi_1' - \frac{\partial_{x_-}\chi_2}{\chi_2}\psi_2'$ we see that it has a simple pole at $\lambda' = 0$ but a double zero at $\lambda' = \infty$ hence also vanishes. ∎

We set $\chi_1/\chi_2 = \exp(2\varphi)$ and $\chi_1\chi_2 = \exp(2\zeta)$ and compute the zero curvature condition for the system (12.41). This yields $\partial_{x_+}\partial_{x_-}\zeta = 0$ and $\partial_{x_+}\partial_{x_-}\varphi = 2m^2\sinh(2\varphi)$. The free field ζ decouples, and can be set to 0. Then the connection appearing in eq.(12.41) is exactly the same as $A'_{x_\pm}$ in eq.(12.39).

We see that, starting from any hyperelliptic curve with branch points at 0 and ∞, and imposing the essential singularities eq.(12.40) at these two punctures yields readily a finite zone solution of the sinh–Gordon equation. Using two punctures allows to get independent equations of motion with respect to x_+ and x_- so producing relativistic generalizations of the finite zone solutions for the KdV equation.

References

[1] A.C.Scott, F.Y.F.Chu, D.W.McLaughlin, *"The Soliton: A New Concept in Applied Science."* Proceedings of the IEEE 61 (1973) 1443.

[2] R.F.Dashen, B.Hasslascher, A.Neveu, *"Particle Spectrum in Model Field Theories from Semi-classical Functional Integral Techniques."* Phys.Rev. D11 (1975) 3424.

[3] A.B.Zamolodchikov, Al.B.Zamolodchikov, *"Factorized S-Matrices in Two Dimensions as the Exact Solutions of Certain Relativistic Quantum Field Theory Models."* Annals of Physics 120 (1979) 253.

[4] S.N.M. Ruijsenaars, H. Schneider, *"A New Class of Integrable Systems and Its Relation to Solitons."* Annals of Physics 170, (1986) pp.370-405.

[5] L.D.Faddeev, L.A.Takhtajan, *"Hamiltonian Methods in the Theory of Solitons"*. Springer (1986)

[6] O.Babelon, D.Bernard, *"Affine Solitons: A Relation Between Tau Functions, Dressing and Bäcklund Transformations."* Int.J.Mod.Phys. A8 (1993) pp.507-543.

Chapter 13

Classical Inverse Scattering Method.

We introduce the basic tools of the Classical Inverse Scattering Method created by Gardner, Green, Kruskal and Miura. This well–known and very ingenious method is usually exploited to solve the KdV equation. Here we felt it was rewarding to show how it applies to a slightly more involved situation. The unique advantage of the inverse scattering method is that it allows to construct action–angle variables for the full infinite–dimensional phase space of field theory. Of course particular solutions like soliton solutions which depend on a finite number of parameters, are recovered this way.

13.1 The sine–Gordon equation.

Let us consider the sine–Gordon equation

$$\partial_t^2\varphi - \partial_x^2\varphi = -\frac{8m^2}{\beta}\sin(2\beta\varphi) \tag{13.1}$$

with $\varphi \equiv \varphi(x,t)$ a real scalar field. One can write this equation as the compatibility condition of a linear system

$$\begin{array}{c} (\partial_x + A_x)\,\psi = 0 \\ (\partial_t + A_t)\,\psi = 0 \end{array} \tag{13.2}$$

with:

$$A_x = i\begin{pmatrix} \frac{\beta}{2}\partial_t\varphi & m(\lambda e^{i\beta\varphi} - \lambda^{-1}e^{-i\beta\varphi}) \\ m(\lambda e^{-i\beta\varphi} - \lambda^{-1}e^{i\beta\varphi}) & -\frac{\beta}{2}\partial_t\varphi \end{pmatrix} \tag{13.3}$$

$$A_t = i\begin{pmatrix} \frac{\beta}{2}\partial_x\varphi & m(\lambda e^{i\beta\varphi} + \lambda^{-1}e^{-i\beta\varphi}) \\ m(\lambda e^{-i\beta\varphi} + \lambda^{-1}e^{i\beta\varphi}) & -\frac{\beta}{2}\partial_x\varphi \end{pmatrix} \tag{13.4}$$

As compared to eqs.(12.4,12.5) in Chapter [12], we have changed $\varphi \to i\varphi$ to get sine–Gordon, and $\lambda \to i\lambda$ so that $i(\partial_x + A_x)$ and $i(\partial_t + A_t)$ are formally self–adjoint for real λ.

In this Chapter, we will study the solutions of the sine–Gordon equation on the line, with boundary conditions:

$$\lim_{x\to-\infty}\varphi(x) = 0, \quad \lim_{x\to+\infty}\varphi(x) = \frac{Q\pi}{\beta}, \quad Q \in \mathbb{Z}$$

These asymptotic conditions are manifestly compatible with the equations of motion and correspond to a minimum of the sine–Gordon potential. We will assume that they are reached rapidly enough.

To solve a non linear partial differential equation by the classical inverse scattering method, we proceed in three steps.

- Given the initial data i.e. $\varphi(x,0)$ for the sine–Gordon case, we solve the direct scattering problem.

$$[\partial_x + A_x(\lambda, \varphi(x,0))]\psi = 0$$

  We thus determine some scattering data $a(\lambda,0)$, $b(\lambda,0)$ for the continuous spectrum, and discrete spectum.

- We determine the time evolution of the scattering data by using the second equation:

$$[\partial_t + A_t(\lambda, \varphi)]\psi = 0 \tag{13.5}$$

  in the regions $x = \pm\infty$ where the asympotics of the field $\varphi(x)$ is known to get $a(\lambda,t)$, $b(\lambda,t)$ ...

- Knowing the scattering data at time t, we reconstruct the matrix $A_x(\lambda,t)$ (and thus $\varphi(x,t)$) by solving the Gelfand-Levitan equation, which is a linear integral equation.

13.2 The Jost solutions.

In our case, the direct scattering problem is defined by the equation

$$\left[\frac{\partial}{\partial x} + i\begin{pmatrix} \frac{\beta}{2}\partial_t\varphi & m(\lambda e^{i\beta\varphi} - \lambda^{-1}e^{-i\beta\varphi}) \\ m(\lambda e^{-i\beta\varphi} - \lambda^{-1}e^{i\beta\varphi}) & -\frac{\beta}{2}\partial_t\varphi \end{pmatrix}\right]\psi = 0 \tag{13.6}$$

This linear system possesses a few simple properties:
For λ real, $A_x^* = \sigma_2 A_x \sigma_2$, (and more generally for complex λ we have $(A_x(\lambda,x))^* = \sigma_2 A_x(\lambda^*,x)\sigma_2$). Therefore if ψ is a solution of eq.(13.6), so is $\overline{\psi} = \sigma_2\psi^*$. This defines a conjugate solution:

$$\overline{\psi} = \begin{pmatrix} \overline{\psi}_1 \\ \overline{\psi}_2 \end{pmatrix} = \begin{pmatrix} -i\psi_2^* \\ i\psi_1^* \end{pmatrix} \tag{13.7}$$

Similarly, we have $A_x(-\lambda) = \sigma_3 A_x(\lambda)\sigma_3$, so that $\sigma_3\psi(x,-\lambda)$ is also a solution.
For two solutions $\xi = \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}$ and $\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}$, we may define the Wronskian $W(\xi,\psi)$ by:

$$W(\xi,\psi) = \xi_1\psi_2 - \xi_2\psi_1$$

Since $\mathrm{Tr}\,(A_x) = 0$, we have $\frac{d}{dx}W(\xi,\psi) = 0$, and the Wronskian is x-independent.
Assume that $\varphi \to 0$ for $x \to -\infty$, and $\varphi \to \frac{\pi Q}{\beta}$ for $x \to +\infty$. Then the linear system becomes respectively

$$\left[\frac{\partial}{\partial x} + ik(\lambda)\sigma_1\right]\psi = 0, x \to -\infty, \quad \left[\frac{\partial}{\partial x} + ie^{Qi\pi}k(\lambda)\sigma_1\right]\psi = 0, x \to +\infty \tag{13.8}$$

where $k(\lambda) = m(\lambda - \lambda^{-1})$. The solutions behave asymptotically for $x \to \pm\infty$ as ,

$$\psi|_{x\to-\infty} \sim c_1\begin{pmatrix} 1 \\ -1 \end{pmatrix} e^{ik(\lambda)x} + c_2\begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{-ik(\lambda)x}, \quad \psi|_{x\to+\infty} \sim c_1'\begin{pmatrix} 1 \\ -e^{i\pi Q} \end{pmatrix} e^{ik(\lambda)x} + c_2'\begin{pmatrix} 1 \\ e^{i\pi Q} \end{pmatrix} e^{-ik(\lambda)x}$$

The Jost solutions are solutions of the linear system with specific behavior at infinity:

Definition. *The Jost solutions f_1 and f_2 are defined by the following asymptotic behavior:*

$$f_1 \sim \begin{pmatrix} 1 \\ -e^{i\pi Q} \end{pmatrix} e^{ik(\lambda)x} \quad x \to +\infty; \quad f_2 \sim \begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{-ik(\lambda)x} \quad x \to -\infty; \tag{13.9}$$

These solutions are chosen in such way that the asymptotics remain bounded if we give a positive imaginary part to λ ($\mathrm{Im}\, k(\lambda) = (1+|\lambda|^{-2})\mathrm{Im}\lambda$).

Let us examine in details the asymptotic form of our Jost solutions when $|x| \to \infty$.

	$x = -\infty$	$x = +\infty$
f_1	$c_{11}(\lambda)\begin{pmatrix}1\\-1\end{pmatrix}e^{ik(\lambda)x} + c_{12}(\lambda)\begin{pmatrix}1\\1\end{pmatrix}e^{-ik(\lambda)x}$	$\begin{pmatrix}1\\-e^{iQ\pi}\end{pmatrix}e^{ik(\lambda)x}$
f_2	$\begin{pmatrix}1\\1\end{pmatrix}e^{-ik(\lambda)x}$	$c_{21}(\lambda)\begin{pmatrix}1\\-e^{iQ\pi}\end{pmatrix}e^{ik(\lambda)x} + c_{22}(\lambda)\begin{pmatrix}1\\e^{iQ\pi}\end{pmatrix}e^{-ik(\lambda)x}$

We get some relations between the coefficients $c_{ij}(\lambda)$ by calculating the Wronskians at $x = +\infty$ and $x = -\infty$:

	$x = -\infty$	$x = +\infty$
$W(f_1, f_2)$	$2c_{11}(\lambda)$	$2c_{22}(\lambda)e^{iQ\pi}$
$W(f_1, \overline{f}_2)$	$2ic_{12}(\lambda)$	$2ic_{21}^*(\lambda)$
$W(f_1, \overline{f}_1)$	$2i\Big(\vert c_{11}\vert^2 + \vert c_{12}\vert^2\Big)$	$2i$

Since the Wronskians are constant, we deduce that $c_{11}(\lambda) = c_{22}(\lambda)e^{iQ\pi}$ and $c_{12}(\lambda) = c_{21}^*(\lambda)$. We set $c_{11}(\lambda) = c_{22}(\lambda)e^{iQ\pi} = a(\lambda)$ and $c_{12}(\lambda) = c_{21}^*(\lambda) = -b(\lambda)$. We have

$$|a(\lambda)|^2 + |b(\lambda)|^2 = 1 \tag{13.10}$$

The asymptotic behavior of the Jost solutions can be rewritten as:

	$x = -\infty$	$x = +\infty$
f_1	$a(\lambda)\begin{pmatrix}1\\-1\end{pmatrix}e^{ik(\lambda)x} - b(\lambda)\begin{pmatrix}1\\1\end{pmatrix}e^{-ik(\lambda)x}$	$\begin{pmatrix}1\\-e^{iQ\pi}\end{pmatrix}e^{ik(\lambda)x}$
f_2	$\begin{pmatrix}1\\1\end{pmatrix}e^{-ik(\lambda)x}$	$-b^*(\lambda)\begin{pmatrix}1\\-e^{iQ\pi}\end{pmatrix}e^{ik(\lambda)x} + a(\lambda)\begin{pmatrix}e^{iQ\pi}\\1\end{pmatrix}e^{-ik(\lambda)x}$

Comparing the asymptotic expansions of $\overline{f}_i$ and $\sigma_3 f_i(-\lambda)$ we get, for real λ:

$$\overline{f}_1 = \frac{ib^*}{a}f_1 + \frac{i}{a}f_2, \qquad \overline{f}_2 = -\frac{i}{a}f_1 - \frac{ib^*}{a}f_2 \tag{13.11}$$

$$\sigma_3 f_1(-\lambda) = e^{-i\pi Q}\frac{b^*}{a}f_1 + e^{-i\pi Q}\frac{1}{a}f_2, \qquad \sigma_3 f_2(-\lambda) = \frac{1}{a}f_1 + \frac{b}{a}f_2 \tag{13.12}$$

From the last relation and $W(f_1, f_2) = 2a(\lambda)$ one gets the symmetry property $a(-\lambda) = e^{-i\pi Q}\, a^*(\lambda)$.

In the following we shall prove that f_1, f_2 and $a(\lambda)$ can be analytically continued to the upper half–plane. Then for λ in the upper half–plane, $-\lambda^*$ is also in the upper half–plane (it is the symmetric of λ with respect to the imaginary axis) and the relation $a(-\lambda) = e^{-i\pi Q}\, a^*(\lambda)$ for λ real extends to $a(-\lambda^*) = e^{-i\pi Q}\,(a(\lambda))^*$. Note also that f_1 and f_2 obey the symmetry property:

$$f_1(-\lambda^*, x) = -e^{i\pi Q}\sigma_1(f_1(\lambda, x))^*, \quad f_2(-\lambda^*, x) = \sigma_1(f_2(\lambda, x))^* \tag{13.13}$$

This is because $A_x(-\lambda^*, x) = \sigma_1(A_x(\lambda, x))^*\sigma_1$ so that $f(-\lambda^*, x)$ and $\sigma_1(f(x, \lambda))^*$ obey the same equation, and one can then compare the asymptotics at $x = \pm\infty$. Suppose that $a(\lambda)$ has some complex zeros λ_n, $\mathrm{Im}\,\lambda_n > 0$. From the symmetry property of $a(\lambda)$, these zeroes are either purely imaginary, or occur in pairs symmetric with respect to the imaginary axis. Because $a(\lambda) = -W(f_1, f_2)$, the two Jost solutions f_1 and f_2 become linearly dependent when $\lambda = \lambda_n$. Looking at the asymptotics, we see that this solution is normalizable. It corresponds to a bound state for the linear scattering problem (13.6). At λ_n, the solutions f_1, f_2 are proportional:

$$f_2(x, \lambda_n) = c_n f_1(x, \lambda_n), \quad a(\lambda_n) = 0 \text{ and } \mathrm{Im}(\lambda_n) > 0 \tag{13.14}$$

Equation (13.13) implies that, for λ_n pure imaginary, $c_n^* = -e^{i\pi Q}c_n$, and a similar relation when we have a pair of symmetric roots. The functions $a(\lambda)$ and $b(\lambda)$ and of the numbers λ_n and c_n constitute the scattering data.

Proposition. *The Jost solutions $f_1(x, \lambda)$ and $f_2(x, \lambda)$ for fixed x are analytic functions of λ in the upper half–plane* $\mathrm{Im}(\lambda) > 0$.

Proof. Denote $A_{\pm\infty} = \lim_{x\to\pm\infty} A_x$, i.e., $A_{-\infty} = ik(\lambda)\sigma_1$ and $A_\infty = ie^{i\pi Q}k(\lambda)\sigma_1$. We can write the linear systems for $f_{1,2}$ as $(\partial_x + A_{\pm\infty})f_{1,2} = (A_{\pm\infty} - A_x)f_{1,2}$. These equations plus the boundary conditions are equivalent to the integral equations

$$f_1(x, \lambda) = f_1^0(x, \lambda) - \int_{-\infty}^{+\infty} dy G_1(x, y)V_1(y, \lambda)f_1(y, \lambda) \tag{13.15}$$

$$f_2(x, \lambda) = f_2^0(x, \lambda) - \int_{-\infty}^{+\infty} dy G_2(x, y)V_2(y, \lambda)f_2(y, \lambda) \tag{13.16}$$

where we have set the explicitly λ dependent terms:

$$V_{1,2}(x, \lambda) = A_x(x, \lambda) - A_{\pm\infty}(x, \lambda), \qquad f_1^0 = \begin{pmatrix} 1 \\ -e^{i\pi Q} \end{pmatrix} e^{ik(\lambda)x}, \qquad f_2^0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{-ik(\lambda)x}$$

and the kernels G_1 and G_2 are:

$$\begin{aligned} G_1(x, y) &= -\frac{1}{2}\theta(y - x)\left[\begin{pmatrix} 1 & -e^{i\pi Q} \\ -e^{i\pi Q} & 1 \end{pmatrix} e^{ik(\lambda)(x-y)} + \begin{pmatrix} 1 & e^{i\pi Q} \\ e^{i\pi Q} & 1 \end{pmatrix} e^{-ik(\lambda)(x-y)}\right] \\ G_2(x, y) &= \frac{1}{2}\theta(x - y)\left[\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} e^{ik(\lambda)(x-y)} + \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} e^{-ik(\lambda)(x-y)}\right] \end{aligned}$$

To analyze the integral equation for f_1 we set $f_1 = e^{ik(\lambda)x}\widetilde{f}_1$ so that the equation becomes:

$$\begin{aligned} \widetilde{f}_1(x, \lambda) &= \begin{pmatrix} 1 \\ -e^{i\pi Q} \end{pmatrix} + \frac{1}{2}\int_x^\infty dy \left[\begin{pmatrix} 1 & -e^{i\pi Q} \\ -e^{i\pi Q} & 1 \end{pmatrix} + \begin{pmatrix} 1 & e^{i\pi Q} \\ e^{i\pi Q} & 1 \end{pmatrix} e^{2ik(\lambda)(y-x)}\right] V_1(y, \lambda)\widetilde{f}_1(y, \lambda) \\ &= \widetilde{f}_1^0 + \int_x^\infty dy K(x, y; \lambda)\widetilde{f}_1(y, \lambda) \end{aligned} \tag{13.17}$$

By iteration, we find the unique solution:

$$\widetilde{f}_1 = (1 + K_1 + K_1^2 + \cdots)\widetilde{f}_1^0$$

We show that this series converges uniformly, thereby proving the existence of the Jost solution f_1. Indeed, since in the upper half–plane $|\exp(2ik(\lambda)(y-x))| \leq 1$ we have $|K(x,y;\lambda)| \leq M(y;\lambda)$ where $M(y;\lambda) > 0$ is such that $\int^\infty dy M(y;\lambda)$ converges. This allows to bound the n^{th} iterate in the following way:

$$\begin{aligned} |(K^n \widetilde{f}_1^0)(x)| &= |\int_x^\infty dx_1 K(x,x_1)\cdots\int_{x_{n-1}}^\infty K(x_{n-1},x_n)\widetilde{f}_1^0| \\ &\leq |\widetilde{f}_1^0| \int_{x\leq x_1\leq\cdots\leq x_n} dx_1\cdots dx_n M(x_1)\cdots M(x_n) = \frac{1}{n!}\Big(\int_x^\infty dy M(y)\Big)^n |\widetilde{f}_1^0| \end{aligned}$$

Because of the $1/n!$, the series is absolutely and uniformly bounded by $\exp(\int_x^\infty dy M(y))$. This is the main observation in the theory of Volterra integral equations. Each term in the expansion is analytic in λ for $\operatorname{Im}\lambda > 0$, hence the uniform limit is also analytic. Note however that for $\lambda = 0$, $V(y,\lambda)$ has a pole so singularities are expected for f_1 and f_2. The equation for f_2 is treated in the same way. ∎

Remark 1. If we assume, in the previous proof, that the potential φ is such that V_1, V_2 have compact support (i.e. the potential reaches its limiting values $0, Q\pi/\beta$ at finite distance), then the integration domain in the integral equation is finite, and one can bound $e^{2ik(\lambda)(y-x)}$ even for $\operatorname{Im}(k) < 0$. It follows that $f_1(x,\lambda)$ and $f_2(x,\lambda)$ are analytic functions of λ in the whole λ plane. In this case, the scattering data $a(\lambda)$, $b(\lambda)$ can be analytically continued in the λ plane. This remains true if V_1, V_2 vanish at $\pm\infty$ rapidly enough to compensate for the growth of $e^{2ik(\lambda)(y-x)}$ when $\operatorname{Im}(k) < 0$. In particular, choosing λ to be a zero λ_n of $a(\lambda)$ in the upper half–plane, and looking at the asymptotic expansions of f_1, f_2 when $x = -\infty$ we see that $c_n = -1/b(\lambda_n)$ in eq.(13.14).

The next task is to compute the asymptotics of the Jost solutions when $\lambda \to 0$ and $\lambda \to \infty$.

Proposition. *The Jost solutions f_1, f_2 have the following asymptotics in λ, valid for any x:*

$$\begin{aligned} f_1 &= e^{ik(\lambda)x}\left(e^{\frac{i\beta\varphi}{2}\sigma_3}\begin{pmatrix} e^{-\frac{i\pi Q}{2}} \\ -e^{-\frac{i\pi Q}{2}}\end{pmatrix} + O(\frac{1}{|\lambda|})\right), \qquad f_2 = e^{-ik(\lambda)x}\left(e^{\frac{i\beta\varphi}{2}\sigma_3}\begin{pmatrix}1\\1\end{pmatrix} + O(\frac{1}{|\lambda|})\right), |\lambda|\to\infty \\ f_1 &= e^{ik(\lambda)x}\left(e^{-\frac{i\beta\varphi}{2}\sigma_3}\begin{pmatrix} e^{\frac{i\pi Q}{2}} \\ -e^{\frac{i\pi Q}{2}}\end{pmatrix} + O(|\lambda|)\right), \qquad f_2 = e^{-ik(\lambda)x}\left(e^{-\frac{i\beta\varphi}{2}\sigma_3}\begin{pmatrix}1\\1\end{pmatrix} + O(|\lambda|)\right), |\lambda|\to 0 \end{aligned} \tag{13.18}$$

Proof. It is easily checked that the right–hand sides of these equations have the correct asymptotics for f_2, f_1 when $x \to \pm\infty$ respectively, see eq.(13.9). Let us consider the case $\lambda \to \infty$ and analyze the asymptotics of f_1 for definiteness, the other cases being similar. We introduce $\widehat{f}_1$ by performing the gauge transformation:

$$f_1 = e^{ik(\lambda)x} e^{i\frac{\beta\varphi}{2}\sigma_3}\widehat{f}_1$$

It obeys the transformed equation $(\partial_x + ik(\lambda)(1+\sigma_1) + \widehat{A}_x)\widehat{f}_1 = 0$ with:

$$\widehat{A}_x = i\begin{pmatrix} \frac{\beta}{2}(\dot{\varphi}+\varphi') & m\lambda^{-1}(1-e^{-2i\beta\varphi}) \\ m\lambda^{-1}(1-e^{2i\beta\varphi}) & -\frac{\beta}{2}(\dot{\varphi}+\varphi') \end{pmatrix}$$

Notice that the boundary conditions on φ are such that $\widehat{A}_x$ rapidly vanishes when $x \to \pm\infty$. Equivalently, $\widehat{f}_1$ is solution of the integral equation:

$$\widehat{f}_1(x,\lambda) = \widehat{f}_1^0 + \int_x^\infty dy e^{ik(\lambda)(y-x)(1+\sigma_1)}\widehat{A}_x(y,\lambda)\widehat{f}_1(y,\lambda), \quad \widehat{f}_1^0 = e^{-i\frac{\pi Q}{2}}\begin{pmatrix}1\\-1\end{pmatrix}$$

We decompose this equation on the eigenvectors of the matrix $(1+\sigma_1)$ corresponding to the eigenvalues $0, 2$:

$$\widehat{f}_1(x,\lambda) = F(x,\lambda)\begin{pmatrix}1\\1\end{pmatrix} + G(x,\lambda)\begin{pmatrix}1\\-1\end{pmatrix}$$

This yields the two coupled scalar integral equations:

$$\begin{aligned}
F(x,\lambda) &= i\int_x^\infty dy e^{2ik(\lambda)(y-x)}\Big\{\frac{\beta}{2}(\dot\varphi+\varphi')G(y,\lambda) + m\lambda^{-1}\Big((1-\cos(2\beta\varphi))F(y,\lambda) - i\sin(2\beta\varphi)G(y,\lambda)\Big)\Big\}\\
G(x,\lambda) &= e^{-i\frac{\pi Q}{2}} + i\int_x^\infty dy\Big\{\frac{\beta}{2}(\dot\varphi+\varphi')F(y,\lambda) + m\lambda^{-1}\Big(-(1-\cos(2\beta\varphi))G(y,\lambda) + i\sin(2\beta\varphi)F(y,\lambda)\Big)\Big\}
\end{aligned}$$

This is a Volterra system, so the iteration procedure converges, exactly as above. Note that there is a Fourier exponential in the first equation, but not in the second one. Remark that for $h(x)$ rapidly decreasing at $+\infty$ we have:

$$\int_x^\infty dy e^{i\lambda y} h(y) = O(\lambda^{-1})$$

It follows that the iteration procedure yields $G(x,\lambda) = e^{-i\frac{\pi Q}{2}} + O(\lambda^{-1})$ and $F(x,\lambda) = O(\lambda^{-1})$. ■

Proposition. *The function $a(\lambda)$ is called the Jost function. It is analytic in the upper half–plane* $\mathrm{Im}\,(\lambda) > 0$. *Furthermore*

$$a(\lambda) = e^{-\frac{i\pi Q}{2}} + O\left(\frac{1}{|\lambda|}\right), \quad |\lambda|\to\infty, \quad a(\lambda) = e^{\frac{i\pi Q}{2}} + O\left(|\lambda|\right), \quad |\lambda|\to 0$$

Recall also that $a(-\lambda) = e^{-i\pi Q}a^*(\lambda)$ for real λ.

Proof. This follows from the relation $a(\lambda) = \frac{1}{2}W(f_1, f_2)$, and from the fact that f_1, f_2 are analytic in the upper half plane. ■

Remark 2. One can let $x\to-\infty$ in the above Volterra system for F, G. Comparing with the definitions of $a(\lambda)$ and $b(\lambda)$ we see that:

$$\widehat{f}_1 \sim a(\lambda)\begin{pmatrix}1\\-1\end{pmatrix} - b(\lambda)\begin{pmatrix}1\\1\end{pmatrix}e^{-2ik(\lambda)x}$$

so that we can directly identify:

$$\begin{aligned}
a(\lambda) &= e^{-i\frac{\pi Q}{2}} + i\int_{-\infty}^\infty dy\Big\{\frac{\beta}{2}(\dot\varphi+\varphi')F(y,\lambda) + m\lambda^{-1}\Big(-(1-\cos(2\beta\varphi))G(y,\lambda) + i\sin(2\beta\varphi)F(y,\lambda)\Big)\Big\}\\
b(\lambda) &= i\int_{-\infty}^\infty dy e^{2ik(\lambda)y}\Big\{\frac{\beta}{2}(\dot\varphi+\varphi')G(y,\lambda) + m\lambda^{-1}\Big((1-\cos(2\beta\varphi))F(y,\lambda) - i\sin(2\beta\varphi)G(y,\lambda)\Big)\Big\}
\end{aligned}$$

We see in the expression for $a(\lambda)$ that it can readily be extended in the upper half–plane, since F and G have such an extension. By contrast, since $b(\lambda)$ is a Fourier transform on the whole real axis, it cannot be extended outside the real λ axis in general. However, if the field φ attains its asymptotic values at finite distance, the integral is over a finite interval, and $b(\lambda)$ admits an analytic continuation in the whole λ plane. Similarly, if we assume that the field φ is C^∞ in x the functions F, G solutions of the integral equation will also be C^∞ in x, so that $\lambda^m b(\lambda)\to 0$ for $\lambda\to\infty$ for any $m>0$, and $\lambda^{-m}b(\lambda)\to 0$ for $\lambda\to 0$. This is because, integrating by parts,

$$\int_{-\infty}^\infty dx e^{i\lambda x}h(x) = -\frac{1}{i\lambda}\int_{-\infty}^\infty dx e^{i\lambda x}h'(x)$$

where the second integral exists by hypothesis. Iterating this formula, we see that the Fourier transform of h goes to 0 faster than any power of λ when $\lambda\to\infty$. Noting that $|a|^2+|b|^2=1$ we see that the module of a tends rapidly to 1 for $\lambda\to 0,\infty$ and so, all information is contained in its phase.

Remark 3. One can of course define Jost solutions f_3, f_4 analytic in the lower half plane. Just choose the appropriate boundary conditions

$$f_3 \sim \begin{pmatrix} 1 \\ e^{i\pi Q} \end{pmatrix} e^{-ik(\lambda)x} \quad x \to +\infty; \quad f_4 \sim \begin{pmatrix} 1 \\ -1 \end{pmatrix} e^{ik(\lambda)x} \quad x \to -\infty; \tag{13.19}$$

These solutions are linear conbinations of f_1, f_2. By comparing at $x = \pm\infty$, we find the relations valid for real λ:

$$f_3 = \frac{b^*(\lambda)}{a(\lambda)} e^{-iQ\pi} f_1 + \frac{e^{-iQ\pi}}{a(\lambda)} f_2, \quad f_4 = \frac{1}{a(\lambda)} f_1 + \frac{b(\lambda)}{a(\lambda)} f_2$$

We can write $b(\lambda) = W(a(\lambda)f_1, f_4)$. Since $a(\lambda)f_1$ and f_4 are not analytic in the same half plane, we recover the fact that $b(\lambda)$ cannot be extended outside the real axis in general. If we define the matrices $\Theta^\pm(\lambda)$, analytic in the upper and lower half plane respectively, by

$$(f_2, f_1) = \Theta^+(\lambda) e^{-ik(\lambda)\sigma_3 x}, \quad (\frac{1}{a^*(\lambda)} e^{iQ\pi} f_3, \frac{1}{a^*(\lambda)} f_4) = \Theta^-(\lambda) e^{-ik(\lambda)\sigma_3 x}$$

where the factors $e^{-ik(\lambda)x}$ are introduced so that $\Theta^\pm(\lambda)$ have finite limits when $\lambda \to 0$ and $\lambda \to \infty$ in their respective domains of analyticity, we see that we can write this relation as

$$(\Theta^-)^{-1}\Theta^+ = \begin{pmatrix} 1 & -b(\lambda)e^{-2ik(\lambda)x} \\ -b^*(\lambda)e^{2ik(\lambda)x} & 1 \end{pmatrix}$$

This is a Riemann–Hilbert problem, typical of a dressing transformation. Note however that the matrix Θ^+ is degenerate at the zeroes of $a(\lambda)$ in the upper half–plane and the matrix $(\Theta^-)^{-1}$ is degenerate at the zeroes of $a^*(\lambda)$ in the lower half–plane. We are thus led to a Riemann–Hilbert problem with zeroes, as discussed in Chapter [2]. In the following, we propose an alternative route to the solution of the inverse scattering problem, by transforming it to the Gelfand–Levitan linear integral equation.

In the following, we need alternative representations of the Jost solutions f_1 and f_2, in which the λ dependence is explicit. It is convenient to first get rid of the explicit φ dependence in the asymptotic expansions eq. (13.18), by defining:

$$f_1 = e^{\frac{i}{2}(\beta\varphi - Q\pi)\sigma_3} \widehat{f}_1, \quad f_2 = e^{\frac{i}{2}\beta\varphi\sigma_3} \widehat{f}_2$$

Proposition. *These functions admit the following Fourier representations:*

$$\begin{aligned} \widehat{f}_1(x,\lambda) &= e^{ik(\lambda)x} \widehat{f}_1^0 + \int_x^\infty dy \Big(\widehat{U}_1(x,y) + \lambda^{-1} \widehat{W}_1(x,y) \Big) e^{ik(\lambda)y} \\ \widehat{f}_2(x,\lambda) &= e^{-ik(\lambda)x} \widehat{f}_2^0 + \int_{-\infty}^x dy \Big(\widehat{U}_2(x,y) + \lambda^{-1} \widehat{W}_2(x,y) \Big) e^{-ik(\lambda)y} \end{aligned} \tag{13.20}$$

where $\widehat{U}_i(x,y), \widehat{W}_i(x,y)$ are two components vectors, and

$$\widehat{f}_1^0 = \begin{pmatrix} 1 \\ -e^{i\pi Q} \end{pmatrix}, \quad \widehat{f}_2^0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Proof. Consider a function $f(\lambda)$ analytic in the upper half–plane. Since $k(\lambda) = m(\lambda - 1/\lambda)$ the map $\lambda \to k$ covers twice the upper half k–plane, the two values λ and $-1/\lambda$ map to the same value of k. So the function $f(\lambda)$ can be written:

$$f(\lambda) = \frac{1}{2}\Big(f(\lambda) + f(-\frac{1}{\lambda})\Big) + \frac{1}{2}\Big(f(\lambda) - f(-\frac{1}{\lambda})\Big) = g_1(k) + \Big(\lambda + \frac{1}{\lambda}\Big) g_2(k)$$

where g_1 and g_2 are analytic functions of k. If f is bounded at $\lambda = 0$ and $\lambda = \infty$, this implies that $g_1(k)$ is bounded and $g_2(k)$ tends to zero at $k \to \infty$. Alternatively, since $\lambda + \frac{1}{\lambda} = 2\lambda^{\pm 1} \mp \frac{1}{m} k$ we can write

$f(\lambda) = h_1(k) + \lambda^{-1}h_2(k) = h_3(k) + \lambda h_4(k)$, and the functions h_1, h_3 are bounded when $k \to \infty$, while h_2, h_4 tend to zero. By the classical Paley–Wiener theorem, one can represent the h_i in the form:

$$h_i(k) = c_i + \int_0^\infty u_i(y) e^{iky} dy \tag{13.21}$$

where the functions $u_i(y)$ are sufficiently regular so that the Fourier integral tends to zero when $k \to \infty$, and c_i is then the limiting value of h_i. The analyticity of h_i in the upper half–plane is accounted for by the support of the Fourier transform on the positive half–line. To apply this to the Jost solutions, we remark that $e^{-ikx}\widehat{f}_1$ is analytic in the upper half–plane. Moreover $e^{-ikx}\widehat{f}_1$ is bounded at $\lambda \to 0, \infty$ by eq.(13.18), hence one can represent it as in eq.(13.21). Multiplying by e^{ikx} and changing variables $y + x \to y$ in the integral (13.21), we arrive at eq.(13.20). The other equation is obtained similarly. ∎

Remark. We can introduce as well functions $\check{f}_1$ $\check{f}_2$ by:

$$f_1 = e^{-\frac{i}{2}(\beta\varphi - Q\pi)\sigma_3}\check{f}_1, \quad f_2 = e^{-\frac{i}{2}\beta\varphi\sigma_3}\check{f}_2$$

which have the representation:

$$\begin{aligned}
\check{f}_1(x,\lambda) &= e^{ik(\lambda)x}\check{f}_1^0 + \int_x^\infty dy\Big(\check{U}_1(x,y) + \lambda\check{W}_1(x,y)\Big)e^{ik(\lambda)y} \\
\check{f}_2(x,\lambda) &= e^{-ik(\lambda)x}\check{f}_2^0 + \int_{-\infty}^x dy\Big(\check{U}_2(x,y) + \lambda\check{W}_2(x,y)\Big)e^{-ik(\lambda)y}
\end{aligned}$$

The kernels $\widehat{U}_i, \widehat{W}_i$ and $\check{U}_i, \check{W}_i$ are determined by the differential equation (13.6). Indeed $\widehat{f}_1$ obeys the gauge transformed equation $(\partial_x + \widehat{A}_x)\widehat{f}_1 = 0$ with the connection:

$$\widehat{A}_x = i\begin{pmatrix} \frac{\beta}{2}(\dot{\varphi} + \varphi') & me^{i\pi Q}(\lambda - \lambda^{-1}e^{-2i\beta\varphi}) \\ me^{i\pi Q}(\lambda - \lambda^{-1}e^{2i\beta\varphi}) & -\frac{\beta}{2}(\dot{\varphi} + \varphi') \end{pmatrix} = \lambda\widehat{A}_1 + \widehat{A}_0 + \lambda^{-1}\widehat{A}_{-1}$$

In the equation, there appears terms in λ and λ^{-2}. To rewrite them, we use:

$$\lambda^{-2}e^{ik(\lambda)y} = e^{ik(\lambda)y} - \frac{1}{im\lambda}\partial_y e^{ik(\lambda)y}, \quad \lambda e^{ik(\lambda)y} = \frac{1}{im}\partial_y e^{ik(\lambda)y} + \frac{1}{\lambda}e^{ik(\lambda)y}$$

and perform integration by parts. The differential equation then translates into the following relations on the kernels $\widehat{U}_1, \widehat{W}_1$:

$$\begin{aligned}
\left(\partial_x - \frac{1}{im}\widehat{A}_1(x)\partial_y\right)\widehat{U}_1(x,y) &= -\widehat{A}_0(x)\widehat{U}_1(x,y) - (\widehat{A}_1(x) + \widehat{A}_{-1}(x))\widehat{W}_1(x,y) \\
\left(\partial_x + \frac{1}{im}\widehat{A}_{-1}(x)\partial_y\right)\widehat{W}_1(x,y) &= -\widehat{A}_0(x)\widehat{W}_1(x,y) - (\widehat{A}_1(x) + \widehat{A}_{-1}(x))\widehat{U}_1(x,y)
\end{aligned}$$

and the boundary conditions:

$$\Big(1 + \frac{1}{im}\widehat{A}_1(x)\Big)\widehat{U}_1(x,x) = \widehat{A}_0(x)\widehat{f}_1^0, \quad \Big(1 - \frac{1}{im}\widehat{A}_{-1}(x)\Big)\widehat{W}_1(x,x) = \Big(\widehat{A}_{-1}(x) - im\Big)\widehat{f}_1^0$$

Alternatively, knowing $\widehat{U}_i, \widehat{W}_i$, these equations allow to reconstruct the connection $\widehat{A}_x(x,\lambda)$ and therefore the field $\varphi(x,t)$. In particular, the boundary conditions yield

$$e^{2i\beta\varphi(x)} = \frac{im - e^{i\pi Q}(\widehat{W}_1)_2(x,x)}{im + (\widehat{W}_1)_1(x,x)} \tag{13.22}$$

The Gelfand–Levitan equation which we will present below relates directly $\widehat{U}_i, \widehat{W}_i$ to the scattering data.

13.3 Time evolution of the scattering data.

In the previous section, from a field $\varphi(x,0)$ at time $t=0$, we have defined the scattering data $a(\lambda)$, $b(\lambda)$, λ_n and c_n. The second step in the classical inverse scattering method is to compute the time evolution of the scattering data, which turns out to be remarkably simple.

Proposition. *The time evolution of the sine–Gordon theory linearizes on the scattering data:*

$$\begin{aligned} \dot{a}(\lambda,t) &= 0, \quad \dot{b}(\lambda,t) = -2im(\lambda+\lambda^{-1})b(\lambda,t) \\ \dot{\lambda}_n &= 0, \quad \dot{c}_n = -2im(\lambda_n+\lambda_n^{-1})c_n \end{aligned} \tag{13.23}$$

<u>Proof</u>. Recall that for the Jost solution $f_1(x,\lambda)$ we have

$$f_1(x,\lambda)|_{x\to+\infty} \sim \begin{pmatrix} 1 \\ -e^{iQ\pi} \end{pmatrix} e^{ik(\lambda)x}, \quad f_1(x,\lambda)|_{x\to-\infty} \sim a(\lambda)\begin{pmatrix} 1 \\ -1 \end{pmatrix} e^{ik(\lambda)x} - b(\lambda)\begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{-ik(\lambda)x}$$

Consider now the time evolution of ψ given by the second member of the Lax pair $\frac{\partial\psi}{\partial t}+A_t\psi=0$. In the limit $x\to+\infty$, it reduces to

$$\frac{\partial\psi}{\partial t} + ie^{iQ\pi}m(\lambda+\lambda^{-1})\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\psi = 0 \tag{13.24}$$

Choose $\psi=\alpha(t)f_1(x,t,\lambda)$. Then, the asymptotic time evolution equation at $x\to+\infty$ gives $\dot{\alpha}=im(\lambda+\lambda^{-1})\alpha$ and for $x\to-\infty$

$$\begin{aligned} \frac{d}{dt}(\alpha a) - im(\lambda+\lambda^{-1})(\alpha a) &= 0 \\ \frac{d}{dt}(\alpha b) + im(\lambda+\lambda^{-1})(\alpha b) &= 0 \end{aligned}$$

Therefore $\dot{a}(\lambda,t)=0$ and $\dot{b}(\lambda,t)=-2im(\lambda+\lambda^{-1})b(\lambda,t)$ as claimed in the proposition.
For bound states, we have by definition $a(\lambda_n)=0$. So λ_n does not evolve. Consider now the wave function $f_n(x)\equiv f_2(x,\lambda_n)=c_nf_1(x,\lambda_n)$. We have

$$f_n(x)|_{x\to-\infty} \sim \begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{-ik(\lambda_n)x}, \quad f_n(x)|_{x\to+\infty} \sim c_n\begin{pmatrix} 1 \\ -e^{iQ\pi} \end{pmatrix} e^{ik(\lambda_n)x}$$

Take $\psi_n=\alpha(t)f_n$. For $x\to-\infty$ we have $\dot{\alpha}=-im(\lambda_n+\lambda_n^{-1})\alpha$ while for $x\to+\infty$, we have $\dot{c}_n-2im(\lambda_n+\lambda_n^{-1})c_n=0$. ■

So we have simple time evolutions of the scattering data:

$$\begin{aligned} a(\lambda,t) &= a(\lambda,0), \quad b(\lambda,t)=e^{-2im(\lambda+\lambda^{-1})t}b(\lambda,0) \\ \lambda_n(t) &= \lambda_n(0), \quad c_n(t)=e^{-2im(\lambda_n+\lambda_n^{-1})t}c_n(0) \end{aligned} \tag{13.25}$$

13.4 The Gelfand–Levitan–Marchenko equation.

We now explain the inverse problem which amounts to reconstructing the potential from the scattering data. The Gelfand–Levitan–Marchenko equation is a linear integral equation which determines $\widehat{U}_i(x,y),\widehat{W}_i(x,y)$ from the scattering data. Once these kernels are known, the local fields are reconstructed from their boundary conditions $\widehat{U}_i(x,x),\widehat{W}_i(x,x)$.

Recall eq.(13.11) that we write in the form:

$$\frac{f_2(x,\lambda)}{a(\lambda)} = r(\lambda) f_1(x,\lambda) - i\overline{f}_1(x,\lambda) \quad \text{with} \quad r(\lambda) = -\frac{b^*(\lambda)}{a(\lambda)} \tag{13.26}$$

To introduce the kernels $\widehat{U}_i(x,y), \widehat{W}_i(x,y)$ we need to rewrite this equation in terms of the $\widehat{f}_i$. Recall that $f_1 = g\widehat{f}_1$ with $g = \exp\left(\frac{i}{2}(\beta\varphi - Q\pi)\sigma_3\right)$. We multiply eq.(13.26) by g^{-1} and notice that $g^{-1}f_2 = e^{i\frac{\pi Q}{2}\sigma_3}\widehat{f}_2$ and $g^{-1}\overline{f}_1 = \overline{\widehat{f}}_1$, (because $g^{-1}\overline{f}_1 = g^{-1}\sigma_2 f_1^* = g^{-1}\sigma_2 g^* \widehat{f}_1^*$ and $g^{-1}\sigma_2 g^* = \sigma_2$). We get:

$$e^{i\frac{\pi Q}{2}\sigma_3}\frac{1}{a(\lambda)}\widehat{f}_2(x,\lambda) = r(\lambda)\widehat{f}_1(x,\lambda) - i\overline{\widehat{f}}_1(x,\lambda) \tag{13.27}$$

The Gelfand–Levitan equation is essentially the Fourier transform of this equation. To perform this transformation we need a Lemma:

Lemma. *We have the relations*

$$\int_{-\infty}^{+\infty} e^{ik(\lambda)x} d\lambda = \frac{2\pi}{m}\delta(x), \qquad \int_{-\infty}^{+\infty} e^{ik(\lambda)x}\frac{d\lambda}{\lambda} = 0, \qquad \int_{-\infty}^{+\infty} e^{ik(\lambda)x}\frac{d\lambda}{\lambda^2} = \frac{2\pi}{m}\delta(x) \tag{13.28}$$

Proof. Because of the singularities at $\lambda = 0, \infty$ one has to give a careful definition of these integrals on the real axis. We give a principal part definition, namely we set:

$$\int_{-\infty}^{\infty} d\lambda = \lim_{\epsilon\to 0}\left(\int_{-1/\epsilon}^{-\epsilon} d\lambda + \int_{\epsilon}^{1/\epsilon} d\lambda\right) \tag{13.29}$$

Let us prove the first formula. Recall that $k(\lambda) = m(\lambda - \lambda^{-1})$ so that $k(-\lambda^{-1}) = k(\lambda)$. In the integral from $-1/\epsilon$ to $-\epsilon$, change $\lambda \to -\lambda^{-1}$ to get

$$\int_{-\infty}^{+\infty} e^{ik(\lambda)x} d\lambda = \lim_{\epsilon\to 0}\int_{\epsilon}^{1/\epsilon} e^{ik(\lambda)x}(1+\lambda^{-2})d\lambda = \frac{1}{m}\int_{-\infty}^{+\infty} e^{ikx}dk = \frac{2\pi}{m}\delta(x)$$

The same technique applied to the second formula yields:

$$\int_{-\infty}^{+\infty} e^{ik(\lambda)x}\frac{d\lambda}{\lambda} = \lim_{\epsilon\to 0}\int_{\epsilon}^{1/\epsilon} e^{ik(\lambda)x}(-\frac{1}{\lambda}+\frac{1}{\lambda})d\lambda = 0$$

Finally the last equation is equivalent to the first one changing λ to $-1/\lambda$. ∎

Theorem. *The kernels $\widehat{U}_1(x,y), \widehat{W}_1(x,y)$, $y \geq x$, appearing in the Fourier transform of $\widehat{f}_1$, eq.(13.20), satisfy the linear integral equations:*

$$\begin{aligned} \frac{2i\pi}{m}\overline{\widehat{U}}_1(x,y) &= F_0(x+y)\widehat{f}_1^0 + \int_x^{\infty}\Big(F_0(y+z)\widehat{U}_1(x,z) + F_{-1}(y+z)\widehat{W}_1(x,z)\Big)dz \\ \frac{2i\pi}{m}\overline{\widehat{W}}_1(x,y) &= F_{-1}(x+y)\widehat{f}_1^0 + \int_x^{\infty}\Big(F_{-1}(y+z)\widehat{U}_1(x,z) + F_{-2}(y+z)\widehat{W}_1(x,z)\Big)dz \end{aligned} \tag{13.30}$$

where the functions $F_j(x)$ are given, in terms of the scattering data, by:

$$F_j(x) = \int_{-\infty}^{\infty} d\lambda \lambda^j e^{ik(\lambda)x} r(\lambda) - 2i\pi\sum_n e^{ik(\lambda_n)x}\lambda_n^j m_n, \quad m_n = \frac{c_n}{a'(\lambda_n)} \tag{13.31}$$

Proof. Remark that in eq.(13.30), $\widehat{U}_1$ and $\widehat{W}_1$ always appear with their second argument greater than the first one, in agreement with their definition. Hence the system of two equations (13.30) determines these two quantities in their domain of definition. We multiply eq.(13.27) by $\lambda^j e^{ik(\lambda)y}$ (or $j = 0, -1$) and integrate over λ from $-\infty$ to $+\infty$, with a principal part prescription, getting:

$$e^{i\frac{\pi Q}{2}\sigma_3}\int_{-\infty}^{\infty} d\lambda\lambda^j e^{ik(\lambda)y}\frac{1}{a(\lambda)}\widehat{f}_2(x,\lambda) = \int_{-\infty}^{\infty} d\lambda\lambda^j e^{ik(\lambda)y}\Big(r(\lambda)\widehat{f}_1(x,\lambda) - i\overline{\widehat{f}}_1(x,\lambda)\Big) \tag{13.32}$$

The right–hand side becomes an integral transform of the $\widehat{U}_i, \widehat{W}_i$ with kernel determined only by the spectral data. Recall the Fourier representations:

$$\begin{aligned}
\widehat{f}_1(x) &= e^{ikx}\widehat{f}_1^0 + \int_x^{\infty}(\widehat{U}_1(x,z) + \lambda^{-1}\widehat{W}_1(x,z))e^{ikz}dz \\
\overline{\widehat{f}}_1(x) &= e^{-ikx}\overline{\widehat{f}}_1^0 + \int_x^{\infty}(\overline{\widehat{U}}_1(x,z) + \lambda^{-1}\overline{\widehat{W}}_1(x,z))e^{-ikz}dz
\end{aligned}$$

where the second equation is derived from the first by complex conjugating and multiplying by σ_2. We evaluate the right–hand side of eq.(13.32) using the Lemma. We find for $j = 0, -1$ respectively:

$$\begin{aligned}
\int_{-\infty}^{\infty} d\lambda e^{iky}(r(\lambda)\widehat{f}_1(x) - i\overline{\widehat{f}}_1(x)) &= R_0(x+y)\widehat{f}_1^0 - i\frac{2\pi}{m}\delta(x-y)\overline{\widehat{f}}_1^0 - i\frac{2\pi}{m}\theta(y-x)\overline{\widehat{U}}_1(x,y) \\
&\quad + \int_x^{\infty} dz\Big(R_0(y+z)\widehat{U}_1(x,z) + R_{-1}(y+z)\widehat{W}_1(x,z)\Big)
\end{aligned}$$

and similarly:

$$\begin{aligned}
\int_{-\infty}^{\infty} d\lambda\lambda^{-1} e^{iky}(r(\lambda)\widehat{f}_1(x) - i\overline{\widehat{f}}_1(x)) &= R_{-1}(x+y)\widehat{f}_1^0 - i\frac{2\pi}{m}\theta(y-x)\overline{\widehat{W}}_1(x,y) \\
&\quad + \int_x^{\infty} dz\Big(R_{-1}(y+z)\widehat{U}_1(x,z) + R_{-2}(y+z)\widehat{W}_1(x,z)\Big)
\end{aligned}$$

where we have introduced the notation:

$$R_j(x) = \int_{-\infty}^{\infty} d\lambda\lambda^j e^{ik(\lambda)x} r(\lambda), \quad j = 0, -1, -2$$

Due to the analyticity properties of the Jost solutions, the left–hand side of eq.(13.32) can be evaluated with the residue theorem. Recall that the λ integrals are defined with the prescription (13.29). We close the contour in the upper half–plane by introducing a small half–circle C_ϵ of center 0 and radius ϵ and a large half–circle $C_{1/\epsilon}$ of center 0 and radius $1/\epsilon$. In the upper half–plane, the integrand has poles at the zeroes λ_n of $a(\lambda)$. So the residue theorem gives:

$$\int_{-\infty}^{\infty} d\lambda\lambda^j\frac{\widehat{f}_2(x,\lambda)}{a(\lambda)}e^{ik(\lambda)y} = 2i\pi\sum_n e^{ik(\lambda_n)y}\lambda_n^j\frac{\widehat{f}_2(x,\lambda_n)}{a'(\lambda_n)} + \oint_{C_\epsilon+C_{1/\epsilon}} d\lambda\lambda^j\frac{\widehat{f}_2(x,\lambda)}{a(\lambda)}e^{ik(\lambda)y} \tag{13.33}$$

where $a'(\lambda) = \frac{da}{d\lambda}$. To proceed we need to evaluate the integrals on the half–circles. Using the asymptotic expansions eq.(13.18) we see that on these circles $\widehat{f}_2(x,\lambda) = e^{-ik(\lambda)x} \times$ regular, and $a(\lambda)$ is regular. We consider first for $j = 0$, integrals of the form $\oint_C e^{ik(\lambda)(y-x)}d\lambda$, where k is large. The existence of such integrals *requires* $y \geq x$, otherwise the exponential explodes. Assuming in the following that this condition is satisfied we have $|e^{ik(\lambda)(y-x)}| \leq 1$ for λ in the upper half–plane, hence the integral on C_ϵ is bounded by $\pi\epsilon$ and can be neglected when $\epsilon \to 0$. On the other hand, the integral over $C_{1/\epsilon}$ reduces by the residue theorem to the integral on the real axis:

$$\oint_{C_{1/\epsilon}} e^{ik(\lambda)(y-x)}d\lambda \sim \int_{-\infty}^{\infty} e^{im\lambda(y-x)}d\lambda = \frac{2\pi}{m}\delta(x-y)$$

so that the last term in eq.(13.33) is equal to $\frac{2\pi}{m}\delta(x-y)e^{i\pi Q\sigma_3/2}\widehat{f}_2^0$. Taking into account the value of $\overline{\widehat{f}}_1^0$ this term precisely cancels the $\delta(x-y)$ term in the right–hand side of eq.(13.32), for $j = 0$. For $j = -1$ there is no such term in the right–hand side, but on the left–hand side we have to consider integrals of the form $\oint_{C_\epsilon} e^{ik(\lambda)x} d\lambda/\lambda$ with $x > 0$. Write $\lambda = \epsilon e^{i\theta}$ with $0 < \theta < \pi$, and fix $\eta > 0$ small enough. The integral takes the form:

$$\oint_{C_\epsilon} e^{ik(\lambda)x}\lambda^{-1}d\lambda \sim i\int_0^\pi d\theta e^{-\frac{mx}{\epsilon}\sin\theta - \frac{imx}{\epsilon}\cos\theta}$$

On the interval $[\eta, \pi-\eta]$ the integrand decays exponentially when $\epsilon \to 0$. On the intervals $[0, \eta]$ and $[\pi - \eta, \pi]$ the integrand is bounded by 1, and the integral by 2η. So the contribution on C_ϵ can be neglected, and a similar analysis shows that the contribution on $C_{1/\epsilon}$ can also be neglected.

We now evaluate $\widehat{f}_2(x, \lambda_n)$ appearing in eq.(13.33). Precisely at the zeroes λ_n of $a(\lambda)$ we have $f_2(x, \lambda_n) = c_n f_1(x, \lambda_n)$ for all x and c_n scalar, which translates into $\widehat{f}_2(x, \lambda_n) = c_n e^{-i\frac{1}{2}\pi Q\sigma_3}\widehat{f}_1(x, \lambda_n)$. Replacing $\widehat{f}_1(x, \lambda_n)$ by its expression, eq.(13.20), one gets:

$$e^{ik(\lambda_n)y}\lambda_n^j \frac{\widehat{f}_2(x,\lambda_n)}{a'(\lambda_n)} = e^{ik(\lambda_n)y}\frac{c_n\lambda_n^j}{a'(\lambda_n)}e^{-i\frac{1}{2}\pi Q\sigma_3}\left[e^{ik(\lambda_n)x}\widehat{f}_1^0 + \int_x^\infty dz e^{ik(\lambda_n)z}(\widehat{U}_1(x,z) + \lambda_n^{-1}\widehat{W}_1(x,z))\right]$$

Combining everything finally yields the Gelfand–Levitan equations. ∎

13.5 Soliton Solutions.

The solution of the Gelfand-Levitan equation (13.30) is particularly simple when we take $R(x) = 0$. This corresponds to $b(\lambda) = 0$ which means that there is no reflexion in the auxiliary scattering problem. Corresponding potentials are called reflectionless potentials. Then the kernels F_j are degenerate and the Gelfand–Levitan equations reduce to a finite linear system. The sine–Gordon solutions $\varphi(x, t)$ we get in this way are just the multi–soliton solutions.

The scattering data are λ_n and m_n. The Gelfand–Levitan kernels read:

$$F_j(x, y) = -2i\pi\sum_n m_n\lambda_n^j e^{ik(\lambda_n)(x+y)}$$

The Gelfand–Levitan equation then shows, looking at the y dependence, that $\widehat{U}_1(x, y)$ and $\widehat{W}_1(x, y)$ can be expanded as:

$$\widehat{U}_1(x,y) = \sum_n e^{-ik(\lambda_n^*)(x+y)}m_n^* u_n(x), \quad \widehat{W}_1(x,y) = \sum_n e^{-ik(\lambda_n^*)(x+y)}m_n^* w_n(x)$$

The y exponentials have been chosen in order that they remain bounded when $y \to \infty$ so that the z integrals converge. The factor $m_n^* e^{-ik(\lambda_n^*)x}$ has been introduced to simplify the latter formulae.

Inserting these forms into the Gelfand–Levitan equations and identifying the coefficient of $\exp(ik(\lambda_n)y)$ on both sides yields:

$$\begin{aligned} -\frac{1}{m}\overline{u}_n(x) &= \widehat{f}_1^0 + \sum_p \frac{im_p^*}{k(\lambda_n) - k(\lambda_p^*)}e^{-2ik(\lambda_p^*)x}(u_p(x) + \lambda_n^{-1}w_p(x)) \\ -\frac{\lambda_n}{m}\overline{w}_n(x) &= \widehat{f}_1^0 + \sum_p \frac{im_p^*}{k(\lambda_n) - k(\lambda_p^*)}e^{-2ik(\lambda_p^*)x}(u_p(x) + \lambda_n^{-1}w_p(x)) \end{aligned}$$

Since the right–hand sides of these equations are identical we get $w_n(x) = \frac{1}{\lambda_n^*}u_n(x)$. Substituting in the equation for u_n and noting that $k(\lambda_n) - k(\lambda_p^*) = m(\lambda_n - \lambda_p^*)(1 + 1/(\lambda_n\lambda_p^*))$ the equation simplifies to:

$$-\overline{u}_n(x) = m\widehat{f}_1^0 + \sum_p \frac{im_p^*}{\lambda_n - \lambda_p^*}e^{-2ik(\lambda_p^*)x}u_p(x)$$

In the following, we restrict ourselves to pure soliton and antisoliton solutions (no breathers), so that the λ_p are pure imaginary. To connect with the notations of Chapter [12] we set $\lambda_p = i/\mu_p$. The Gelfand–Levitan equation becomes, in matrix notation:

$$-\overline{u} = mf + Vu \tag{13.34}$$

where u is the vector with components $u_n = (u_{1n}, u_{2n})$, f is the vector with components $f_n = \widehat{f}_1^0 e_n$, where $e_n = 1$ for all n, and we have defined the matrix:

$$V_{np} = \frac{\mu_n \mu_p}{\mu_n + \mu_p} X_p, \quad X_p = m_p e^{-2m(\mu_p + \mu_p^{-1})x} \tag{13.35}$$

Note that, since $a(-\lambda^*) = e^{i\pi Q}(a(\lambda))^*$ and $c_n^* = -e^{i\pi Q} c_n$ in this case, we have $m_p^* = m_p$, and the matrix V is real. Remark that, if one includes the time dependence, i.e. $m_p \to m_p e^{-2m(\mu_p - \mu_p^{-1})t}$, the exponential appearing in V becomes $\exp\left(-2m[\mu_p(x+t) + \mu_p^{-1}(x-t)]\right)$ which is the familiar form encountered in Chapter [12].

Taking the bar of eq.(13.34) and using that $\overline{\overline{u}} = -u$ we get $u = m\overline{f} + V\overline{u}$ so that, eliminating $\overline{u}$:

$$(1 + V^2)u = m\overline{f} - mVf$$

Writing $(1 + V^2) = (1 + ie^{i\pi Q}V)(1 - ie^{i\pi Q}V)$ this is solved by:

$$u_1 = ime^{i\pi Q}(1 - ie^{i\pi Q}V)^{-1}e, \quad u_2 = im(1 + ie^{i\pi Q}V)^{-1}e$$

We can then compute the field φ by eq.(13.22). This is done in the:

Proposition. *We have:*

$$im + W_1(x,x) = im\frac{\det(1 + ie^{i\pi Q}V)}{\det(1 - ie^{i\pi Q}V)}, \quad im - e^{i\pi Q}W_2(x,x) = im\frac{\det(1 - ie^{i\pi Q}V)}{\det(1 + ie^{i\pi Q}V)}$$

so that $e^{-i\beta\varphi} = \frac{\tau_+}{\tau_-}$ as in eqs.(12.16,12.25) in Chapter [12].

Proof. Let us prove the first equation. We have

$$W_1(x,x) = i\sum_n \mu_n X_n (u_1)_n = -me^{i\pi Q}\,\mathrm{Tr}\,(M), \quad M \equiv (1 - ie^{i\pi Q}V)^{-1} e \otimes \mu X$$

where $\mu = \mathrm{Diag}\,(\mu_n)$ and X is the vector of components X_n, so that $im + W_1(x,x) = im\Big(1 + ie^{i\pi Q}\,\mathrm{Tr}\,(M)\Big)$. Remark that M is of rank 1, so $1 + \mathrm{Tr}\,(ie^{i\pi Q}\,M) = \det(1 + ie^{i\pi Q}M)$. Moreover:

$$1 + ie^{i\pi Q}M = (1 - ie^{i\pi Q}V)^{-1}\Big(1 - ie^{i\pi Q}(V - e \otimes \mu X)\Big)$$

and we finally remark that $(V - e \otimes \mu X) = -\mu^{-1}V\mu$. The second formula is obtained similarly. Equation (13.22) gives:

$$e^{2i\beta\varphi} = \frac{\det^2(1 - ie^{i\pi Q}V)}{\det^2(1 + ie^{i\pi Q}V)}$$

which identifies with the τ–function formula. ∎

13.6 Poisson brackets of the scattering data.

We now consider the sine–Gordon equation from the Hamiltonian point of view. The aim is to compute the Poisson brackets of the scattering data defined in the previous sections. We follow the method of the classical r–matrix introduced by Faddeev, Sklyanin and Takhtajan.

We start from the canonical Poisson brackets:

$$\{\pi(x), \varphi(y)\} = \delta(x-y)$$

and define the Hamiltonian on the interval $[-\ell, \ell]$ (ℓ will eventually tend to ∞)

$$H = \int_{-\ell}^{\ell} dx \left(\frac{1}{2}\pi^2(x) + \frac{1}{2}(\partial_x \varphi)^2(x) + \frac{4m^2}{\beta^2}(1 - \cos(2\beta\varphi)) \right) dx$$

The equations of motion read $\dot{\varphi} = \pi$ and $\dot{\pi} = \varphi'' - \frac{8m^2}{\beta}\sin(2\beta\varphi)$ which reproduces eq.(13.1).

Consider the auxiliary linear problem (13.6) on the interval $[-\ell, \ell]$. Let $\psi(-\ell)$ be the value of a solution at $x = -\ell$ and $\psi(\ell)$ its value at $x = \ell$. Then we can write:

$$\psi(\ell) = T_\ell(\lambda)\psi(-\ell)$$

where $T_\ell(\lambda)$ is the monodromy matrix.

Proposition. *The monodromy matrix is related to the scattering data by:*

$$T(\lambda) \equiv \begin{pmatrix} a(\lambda) & b(\lambda) \\ -b^*(\lambda) & a^*(\lambda) \end{pmatrix} = \frac{1}{2}\lim_{\ell\to\infty} \begin{pmatrix} e^{i\pi Q}e^{ik(\lambda)\ell} & e^{ik(\lambda)\ell} \\ e^{-ik(\lambda)\ell} & -e^{i\pi Q}e^{-ik(\lambda)\ell} \end{pmatrix} T_\ell(\lambda) \begin{pmatrix} e^{ik(\lambda)\ell} & e^{-ik(\lambda)\ell} \\ e^{ik(\lambda)\ell} & -e^{-ik(\lambda)\ell} \end{pmatrix} \tag{13.36}$$

Thus the scattering data are directly related to the monodromy matrix.

Proof. From the asymptotic form of the Jost solutions and the relation $|a|^2 + |b|^2 = 1$ we find

$$\frac{1}{a}f_1 + \frac{b}{a}f_2 \sim \begin{cases} \begin{pmatrix} 1 \\ -1 \end{pmatrix} e^{ik(\lambda)x} & x \to -\infty \\ a^* \begin{pmatrix} 1 \\ -e^{i\pi Q} \end{pmatrix} e^{ik(\lambda)x} + b \begin{pmatrix} e^{i\pi Q} \\ 1 \end{pmatrix} e^{-ik(\lambda)x} & x \to +\infty \end{cases}$$

Any solution $\psi(x)$ can be written as $\psi_1(\frac{1}{a}f_1 + \frac{b}{a}f_2) + \psi_2 f_2$ and behaves at $x = -\ell$ as:

$$\begin{pmatrix} e^{-ik(\lambda)\ell} & e^{ik(\lambda)\ell} \\ -e^{-ik(\lambda)\ell} & e^{ik(\lambda)\ell} \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}$$

while at $x = +\infty$ it behaves as:

$$\begin{pmatrix} a^* e^{ik(\lambda)\ell} + b e^{i\pi Q} e^{-ik(\lambda)\ell} & -b^* e^{ik(\lambda)\ell} + a e^{i\pi Q} e^{-ik(\lambda)\ell} \\ -a^* e^{i\pi Q} e^{ik(\lambda)\ell} + b e^{-ik(\lambda)\ell} & b^* e^{i\pi Q} e^{ik(\lambda)\ell} + a e^{-ik(\lambda)\ell} \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}$$

The result follows from writing $\psi(\ell) = T_\ell(\lambda)\psi(-\ell)$. ■

To compute the Poisson brackets of the scattering data we need the r–matrix relation:

Proposition.

$$\{T_{\ell,1}(\lambda), T_{\ell,2}(\mu)\} = -[r_{12}(\lambda,\mu), T_\ell(\lambda) \otimes T_\ell(\mu)] \tag{13.37}$$

where the matrix $r_{12}(\lambda,\mu)$ is given by

$$r_{12}(\lambda,\mu) = -\frac{\beta^2}{4}\left(\frac{\mu^2+\lambda^2}{\mu^2-\lambda^2} H \otimes H + \frac{4\lambda\mu}{\mu^2-\lambda^2}(E_+ \otimes E_- + E_- \otimes E_+) \right) \tag{13.38}$$

Proof. As we know from Chapter [2], see eqs.(2.57,2.58) it suffices to prove the much simpler local relation:

$$\{A_{x,1}(\lambda,x), A_{x,2}(\mu,y)\} = [r_{12}(\lambda,\mu), A_x(\lambda,x) \otimes I + I \otimes A_x(\mu,y)]\delta(x-y) \tag{13.39}$$

The r–matrix is obtained, up to a factor, by the general formula for Toda models, eq.(11.24) in Chapter [11] in which we insert the root decomposition, eq.(12.12) in Chapter [12]. We get in the loop representation:

$$r_{12} = \frac{1}{4} H \otimes H + \sum_{n>0} \Big(\lambda^{2n-1} E_+ \otimes \mu^{-2n+1} E_- + \lambda^{2n-1} E_- \otimes \mu^{-2n+1} E_+ + \frac{1}{2} \lambda^{2n} H \otimes \mu^{-2n} H \Big)$$

(we take as invariant bilinear form on $sl(2)$ the trace in the 2×2 representation, so that $(H, H) = 2$ and $(E_+, E_-) = 1$, this accounts for the relative factors). Summing the geometric series yields eq.(13.38). It is easy to check that the relation (13.39) holds true, and this allows to adjust the factor $-\beta^2$. ∎

Proposition. *The complete list of Poisson brackets of the scattering data is :*
1) Continuum–continuum :

$$\begin{aligned}
\{a(\lambda), a(\mu)\} &= 0, \quad \{a(\lambda), a^*(\mu)\} = 0, \quad \{b(\lambda), b(\mu)\} = 0 \\
\{a(\lambda), b(\mu)\} &= -\beta^2 \frac{\lambda\mu}{(\lambda+\mu)(\lambda-\mu+i0)} a(\lambda) b(\mu), \quad \{a(\lambda), b^*(\mu)\} = \beta^2 \frac{\lambda\mu}{(\lambda+\mu)(\lambda-\mu+i0)} a(\lambda) b^*(\mu) \\
\{b(\lambda), b^*(\mu)\} &= i\pi\beta^2 \, \lambda \, |a(\lambda)|^2 \, \delta(\lambda-\mu)
\end{aligned} \tag{13.40}$$

2) Continuum–discrete :

$$\begin{aligned}
\{a(\lambda), \lambda_n\} &= 0, \quad \{a(\lambda), \lambda_n^*\} = 0 \\
\{b(\lambda), \lambda_n\} &= 0, \quad \{b(\lambda), \lambda_n^*\} = 0 \\
\{b(\lambda), c_n\} &= 0, \quad \{b(\lambda), c_n^*\} = 0 \\
\{a(\lambda), c_n\} &= \frac{\beta^2 \lambda \lambda_n}{\lambda^2 - \lambda_n^2} a(\lambda) c_n, \quad \{a(\lambda), c_n^*\} = -\frac{\beta^2 \lambda \lambda_n^*}{\lambda^2 - \lambda_n^{*2}} a(\lambda) c_n^*
\end{aligned} \tag{13.41}$$

3) Discrete–discrete :

$$\begin{aligned}
\{\lambda_n, \lambda_m\} &= 0, \quad \{\lambda_n, \lambda_m^*\} = 0 \\
\{c_n, c_m\} &= 0, \quad \{c_n, c_m^*\} = 0 \\
\{\lambda_n, c_m\} &= -\frac{\beta^2}{2} \lambda_n c_m \delta_{nm}, \quad \{\lambda_n^*, c_m\} = 0 \ \text{ unless } \ \lambda_m = -\lambda_n^*
\end{aligned} \tag{13.42}$$

Proof. Call $E_L(\lambda)$ and $E_R(\lambda)$ the matrices:

$$E_L(\lambda) = \begin{pmatrix} e^{i\pi Q} e^{ik(\lambda)\ell} & e^{ik(\lambda)\ell} \\ e^{-ik(\lambda)\ell} & -e^{i\pi Q} e^{-ik(\lambda)\ell} \end{pmatrix}, \quad E_R(\lambda) = \begin{pmatrix} e^{ik(\lambda)\ell} & e^{-ik(\lambda)\ell} \\ e^{ik(\lambda)\ell} & -e^{-ik(\lambda)\ell} \end{pmatrix}$$

so that $T(\lambda) = \frac{1}{2} \lim_{\ell\to\infty} E_L(\lambda) T_\ell(\lambda) E_R(\lambda)$. Inserting this into eq.(13.37), we get the Poisson brackets of the elements of the matrix $T(\lambda)$.

$$\{T_1(\lambda), T_2(\mu)\} = -r_+(\lambda,\mu) T_1(\lambda) T_2(\mu) + T_1(\lambda) T_2(\mu) r_-(\lambda,\mu)$$

where we have defined:

$$\begin{aligned}
r_+(\lambda,\mu) &= E_L(\lambda) \otimes E_L(\mu) r(\lambda,\mu) E_L^{-1}(\lambda) \otimes E_L^{-1}(\mu), \quad \ell \to \infty \\
r_-(\lambda,\mu) &= E_R^{-1}(\lambda) \otimes E_R^{-1}(\mu) r(\lambda,\mu) E_R(\lambda) \otimes E_R(\mu), \quad \ell \to \infty
\end{aligned}$$

In order to evaluate these r–matrices, one first computes:

$$\begin{aligned}
E_L H E_l^{-1} &= e^{i\pi Q} \Big(e^{2ik\ell} E_+ + e^{-2ik\ell} E_- \Big), \quad E_L E_\pm E_l^{-1} = \frac{1}{2} \Big(e^{i\pi Q} H \mp e^{2ik\ell} E_+ \pm e^{-2ik\ell} E_- \Big) \\
E_R^{-1} H E_R &= e^{-2ik\ell} E_+ + e^{2ik\ell} E_-, \quad E_R^{-1} E_\pm E_R = \frac{1}{2} \Big(H \mp e^{-2ik\ell} E_+ \pm e^{2ik\ell} E_- \Big)
\end{aligned}$$

and then obtain $r_\pm$, before taking the $\ell \to \infty$ limit:

$$\begin{aligned}
-\frac{4}{\beta^2}\, r_\pm(\lambda,\mu) &= \frac{2\lambda\mu}{\mu^2-\lambda^2} H\otimes H \\
&\quad + \frac{\mu-\lambda}{\mu+\lambda} e^{\pm 2i(k(\lambda)+k(\mu))\ell} E_+\otimes E_+ + \frac{\mu-\lambda}{\mu+\lambda} e^{\mp 2i(k(\lambda)+k(\mu))\ell} E_-\otimes E_- \\
&\quad + \frac{\mu+\lambda}{\mu-\lambda} e^{\pm 2i(k(\lambda)-k(\mu))\ell} E_+\otimes E_- + \frac{\mu+\lambda}{\mu-\lambda} e^{\mp 2i(k(\lambda)-k(\mu))\ell} E_-\otimes E_+
\end{aligned}$$

We compute the limit $\ell \to \infty$ in the sense of distribution theory. This is done using the formula:

$$\lim_{\ell\to\infty} P\, \frac{1}{x} e^{\pm i\ell x} = \pm i\pi\delta(x)$$

Indeed, if f is analytic with slow growth at ∞ and $x > 0$ one can compute

$$\left(\int_{-\infty}^{-\epsilon} + \int_{\epsilon}^{\infty}\right) \frac{1}{x} e^{\pm i\ell x} f(x)dx = i\pi f(0)$$

by considering the closed contour obtained by adding a half–circle C_ϵ and a half–circle at infinity. The integral on this last circle vanishes in the limit $\ell \to \infty$, and the integral on C_ϵ gives $-i\pi f(0)$. In our case, we have principal values of the form:

$$\lim_{\ell\to\infty} P\, \frac{1}{\lambda-\mu} e^{\pm 2i(k(\lambda)-k(\mu))} = \pm i\pi\delta(\lambda-\mu)$$

which is obtained by the change of variables $\lambda-\mu \to 2(k(\lambda)-k(\mu))$ in the delta function. We have similar formulae for $\lambda+\mu$ but due to the symmetry properties under $\lambda \to -\lambda$, $a(-\lambda) = e^{i\pi Q} a^*(\lambda)$ and $b(-\lambda) =$???? one can restrict ourselves to $\lambda > 0$ and $\mu > 0$, and ignore the terms in $\delta(\lambda+\mu)$.

We can then take the limit $\ell \to \infty$ in $r_\pm$ getting:

$$-\frac{4}{\beta^2}\, r_\pm(\lambda,\mu) = \frac{1}{2}\Big(P\,\frac{\mu+\lambda}{\mu-\lambda} - \frac{\mu-\lambda}{\mu+\lambda}\Big) H\otimes H \pm i\pi(\lambda+\mu)\delta(\lambda-\mu)(E_+\otimes E_- - E_-\otimes E_+)$$

From this we compute the Poisson brackets of the scattering data, eq.(13.40). For instance:

$$\begin{aligned}
\{a(\lambda), b(\mu)\} &= \beta^2 \left[\frac{\lambda\mu}{\mu^2-\lambda^2} a(\lambda)b(\mu) + \frac{i\pi}{4}(\lambda+\mu)\delta(\lambda-\mu)b(\lambda)a(\mu)\right] \\
&= \frac{\beta^2}{4}\left[(\lambda+\mu)\Big(P\,\frac{1}{\mu-\lambda} + i\pi\delta(\mu-\lambda)\Big) - \frac{\mu-\lambda}{\mu+\lambda}\right] a(\lambda)b(\mu) \\
&= -\beta^2 \frac{\lambda\mu}{(\lambda-\mu+i0)(\lambda+\mu)} a(\lambda)b(\mu)
\end{aligned}$$

where in the last step we have used the identity:

$$\frac{1}{x\pm i0} = P\frac{1}{x} \mp i\pi\delta(x)$$

Note that the left–hand side and the right–hand side are analytic in the upper λ half–plane, as it should be. The other Poisson brackets in eq.(13.40) are computed similarly. There are 16 Poisson brackets in $\{T_1, T_2\}$ but the independent ones are listed in eq.(13.40).

The Poisson brackets for the discrete spectrum would require a detailed analysis, but they can be obtained quickly using the following trick. We insisted already on the fact that the function $b(\mu)$ cannot be analytically continued in the upper half plane. The situation changes however if the field $\varphi(x)$ is *compactly supported.* Then, $b(\mu)$ is analytic in the plane, and we have seen that:

$$c_n = -\frac{1}{b(\lambda_n)}$$

Assuming that we are in such a case, setting $\mu = \lambda_m$ into the equation for $\{a(\lambda), b(\mu)\}$ we get immediately $\{a(\lambda), c_m\} = \beta^2 \lambda\lambda_m/(\lambda^2 - \lambda_m^2) a(\lambda) c_m$. Letting further $a(\lambda) = (\lambda - \lambda_n)\widetilde{a}_n$ in this equation with $\lambda \to \lambda_n$ one gets $\{\lambda_n, c_m\} = -\frac{\beta^2}{2}\lambda_n c_m \delta_{nm}$. The remaining formulae in eqs.(13.41,13.42) are proved similarly. ■

13.7 Action–Angle variables.

Due to the boundary conditions of the field φ, which differ at $x = -\infty$ and $x = +\infty$, the generating function of conserved quantities is a modified trace of the monodromy matrix, namely $\mathrm{Tr}\,(T_\ell(\lambda)\rho)$ where ρ is the diagonal matrix $\rho = \mathrm{Diag}\,(e^{-i\frac{1}{2}\pi Q}, e^{i\frac{1}{2}\pi Q})$. This is a consequence of the zero curvature condition which implies, see eq.(2.49) in Chapter [2]:

$$\partial_t T_\ell(\lambda, t) = -A_t(\lambda, +\ell) T_\ell(\lambda, t) + T_\ell(\lambda, t) A_t(\lambda, -\ell)$$

and the explicit values $A_t(\lambda, +\ell) = im(\lambda + \lambda^{-1}) e^{i\pi Q}\sigma_1$ and $A_t(\lambda, -\ell) = im(\lambda + \lambda^{-1})\sigma_1$ for $\ell \to +\infty$. Computing T_ℓ from eq.(13.36) we express this generating function in terms of the scattering data:

$$\mathrm{Tr}\,(T_\ell(\lambda)\rho) \sim e^{i\frac{1}{2}\pi Q} e^{-2ik(\lambda)\ell} a(\lambda) + e^{-i\frac{1}{2}\pi Q} e^{2ik(\lambda)\ell} a^*(\lambda)$$

so that the generating functional for conserved quantities can be taken as $a(\lambda)$. Since $\{a(\lambda), a(\mu)\} = 0$ these conserved quantities Poisson commute.

Since

$$a(\lambda) = e^{-\frac{i\pi Q}{2}} + O\left(\frac{1}{|\lambda|}\right), \quad |\lambda| \to \infty, \quad a(\lambda) = e^{\frac{i\pi Q}{2}} + O\left(|\lambda|\right), \quad |\lambda| \to 0$$

we can expand $a(\lambda)$ around $\lambda = 0$ and $\lambda = \infty$:

$$\log\, a(\lambda) = -\frac{i\pi Q}{2} + \sum_{n=1}^{\infty} I_n^+ (i\lambda)^{-n} \quad |\lambda| \to \infty, \quad \log a(\lambda) = \frac{i\pi Q}{2} + \sum_{n=1}^{\infty} I_n^- (i\lambda)^n, \quad |\lambda| \to 0$$

We will now calculate the $I_n^\pm$ in two different ways. The first will give us $I_n^\pm$ in terms of the original field φ and the second will express $I_n^\pm$ in terms of the scattering data.

In order to compute I_n^+ in terms of φ ($\lambda \to \infty$), we perform a gauge transformation $\psi \to e^{i\frac{\beta}{2}\varphi\sigma_3}\widetilde{\psi}$, so that the gauge transformed connection $\widetilde{A}_x$ reads:

$$\widetilde{A}_x = i\frac{\beta}{2}(\dot{\varphi} + \varphi')H + im\lambda(1 - \lambda^{-2}e^{-2i\beta\varphi})E_+ + im\lambda(1 - \lambda^{-2}e^{2i\beta\varphi})E_-$$

Remark that $\widetilde{A}_x$ takes the same value at $x = \pm\infty$ and we have $\mathrm{Tr}\,(\widetilde{T}_\ell(\lambda)) = \mathrm{Tr}\,(T_\ell(\lambda)\rho)$. To compute this trace we can apply directly eq.(10.8) in Chapter [10] where we have found that $\mathrm{Tr}\,(\widetilde{T}_\ell(\lambda)) = e^{P_\ell(\lambda)} + e^{-P_\ell(\lambda)}$. For smooth ($C^\infty$) fields φ the quantity $P_\ell(\lambda)$ admits asymptotic expansions for $\lambda \to 0, \infty$ which can be computed using the Ricatti equation. The coefficients of this asymptotic expansion are integrals over local densities in the field φ containing higher and higher derivatives. Hence the smoothness condition is essential for their existence. On the other hand for such smooth fields we have seen that $b(\lambda)$ vanishes at $\lambda = 0, \infty$ as well as all its derivatives (since it vanishes faster than any inverse powers of $k(\lambda)$). This means that $b(\lambda)$ has zero asymptotic expansion at these points. Since $|a|^2 = 1 - |b|^2$ we see that $|a| = 1$ in the asymptotic sense, or $a(\lambda)^* \sim 1/a(\lambda)$. Using this fact we can compare the asymptotic expansions in both sides of the equation:

$$e^{P_\ell(\lambda)} + e^{-P_\ell(\lambda)} = e^{i\frac{1}{2}\pi Q} e^{-2ik(\lambda)\ell} a(\lambda) + e^{-i\frac{1}{2}\pi Q} e^{2ik(\lambda)\ell} a^*(\lambda)$$

and identify in the asymptotic sense:

$$P_\ell(\lambda) = i\frac{1}{2}\pi Q - 2ik(\lambda)\ell + \log\, a(\lambda), \quad \lambda \to \infty$$

To compute the left–hand side we recall that $P_\ell(\lambda)=\int_{-\ell}^{\ell}v(x,\lambda)dx$, where v is solution of the Ricatti equation $v'+v^2=V$ and V is determined in eq.(10.8) in terms of $\widetilde{A}_x$. We compute this expansion at the lowest non trivial order, so that $O(\lambda^{-2})$ terms are neglected. We get:

$$V=-m^2\lambda^2-\frac{\beta^2}{4}(\dot\varphi+\varphi')^2+2m^2\cos(2\beta\varphi)-i\frac{\beta}{2}(\dot\varphi+\varphi')'+O(\lambda^{-2})$$

One inserts $v=\pm im\lambda+\cdots$ and observe that an $O(1)$ term is imposssible. Up to a choice of sign, and substituting $k(\lambda)=m\lambda-m\lambda^{-1}$ we have:

$$v=-ik(\lambda)-im\lambda^{-1}\Big(\frac{\beta^2}{8m^2}(\dot\varphi+\varphi')^2+(1-\cos(2\beta\varphi))+\frac{i\beta}{4m^2}(\dot\varphi+\varphi')'\Big)+O(\lambda^{-2})$$

Since $(\dot\varphi+\varphi')$ vanishes at $x=\pm\infty$ we obtain:

$$P_\ell(\lambda)=-2ik(\lambda)\ell-im\lambda^{-1}\frac{\beta^2}{4m^2}(H+P)+O(\lambda^{-2})$$

where:

$$H=\int_{-\ell}^{\ell}(\frac{1}{2}\dot\varphi^2+\frac{1}{2}\varphi'^2+\frac{4m^2}{\beta^2}(1-\cos(2\beta\varphi)))dx,\quad P=\int_{-\ell}^{\ell}\dot\varphi\varphi'dx$$

Similarly to compute I_n^- $(\lambda\to 0)$ we perform the gauge transformation $\psi\to e^{-i\frac{\beta}{2}\varphi\sigma_3}\widetilde{\psi}$, so that the gauge transformed connection $\widetilde{A}_x$ reads:

$$\widetilde{A}_x=i\frac{\beta}{2}(\dot\varphi-\varphi')H-im\lambda^{-1}(1-\lambda^2e^{2i\beta\varphi})E_+-im\lambda^{-1}(1-\lambda^2e^{-2i\beta\varphi})E_-$$

Note however that, due to the sign change in the gauge transformation, we now have $\mathrm{Tr}\,\widetilde{T}=\mathrm{Tr}\,(T\rho^{-1})=e^{i\pi Q}\mathrm{Tr}\,(T\rho)$. The same computation as above yields:

$$P_\ell(\lambda)=-2ik(\lambda)\ell+im\lambda\frac{\beta^2}{4m^2}(H-P)+O(\lambda^2)$$

Comparing the asymptotic expansions in:

$$\mathrm{Tr}\,\widetilde{T}=e^{P_\ell(\lambda)}+e^{-P_\ell(\lambda)}=e^{-i\pi Q}(e^{i\frac{1}{2}\pi Q}e^{-2ik(\lambda)\ell}a(\lambda)+e^{-i\frac{1}{2}\pi Q}e^{2ik(\lambda)\ell}a^*(\lambda))$$

we find

$$P_\ell(\lambda)=-i\frac{1}{2}\pi Q-2ik(\lambda)\ell+\log a(\lambda),\quad \lambda\to 0$$

from which we identify:

$$I_1^{\pm}=\frac{\beta^2}{4m}(H\pm P)$$

Now we reconstruct $a(\lambda)$ from its analyticity properties, and get alternative expressions for the quantities $I_n^{\pm}$. Recall that the function $a(\lambda)$ is analytic in the upper half plane, behaves at $\lambda=0,\infty$ as:

$$e^{i\frac{1}{2}\pi Q}a(\lambda)=1+O\left(\frac{1}{|\lambda|}\right),\quad |\lambda|\to\infty,\quad e^{i\frac{1}{2}\pi Q}a(\lambda)=e^{i\pi Q}+O\left(|\lambda|\right),\quad |\lambda|\to 0$$

and obeys $(e^{i\frac{1}{2}\pi Q}a(-\lambda^*))^*=e^{i\frac{1}{2}\pi Q}a(\lambda)$. Therefore we can reconstruct $a(\lambda)$ if we know its modulus on the real axis and the position of its zeros.

$$e^{i\frac{1}{2}\pi Q}a(\lambda)=\prod_i\left(\frac{\lambda-\lambda_i}{\lambda-\lambda_i^*}\right)\ \exp\left[-\frac{1}{i\pi}\int_{-\infty}^{\infty}\frac{\log|a(\mu)|}{\lambda-\mu+i0}d\mu\right]\tag{13.43}$$

Indeed, the right–hand side is analytic in the *upper* half–plane. It is invariant under complex conjugation and changing $\lambda\to-\lambda^*$ because for each λ_i there is a $\lambda_j=-\lambda_i^*$ and one changes variables $\mu\to-\mu$ in the integral ($|a(-\mu)|=|a(\mu)|$). For $\lambda\to\infty$ one gets the correct asymptotic value 1, and finally for λ real, using:

$$\frac{1}{\lambda-\mu+i0}=P\frac{1}{\lambda-\mu}-i\pi\delta(\lambda-\mu)$$

the modulus of the right–hand side is $\exp\left(\int \log|a(\mu)|\delta(\lambda-\mu)d\mu\right) = |a(\lambda)|$. Using the asymptotic value at $\lambda = 0$, and noting that the integral vanishes for $\lambda = 0$ (by $\mu \to -\mu$) we must have $e^{i\pi Q} = \prod_i (\lambda_i/\lambda_i^*)$. Note that pair of roots symmetric with respect to the imaginary axis contribute a $+1$ in this product, but pure imaginary roots contribute a -1, so, modulo 2, Q must be equal to the number of pure imaginary roots, i.e. the total number of solitons and antisolitons must have the same parity as the topological charge Q. This is as it should be since a soliton has $Q = 1$ and an antisoliton has $Q = -1$.

On the real axis we have $|a(\lambda)|^2 = 1-|b(\lambda)|^2$ and we can replace $\frac{1}{\pi}\log|a(\lambda)|$ by $\rho(\lambda) = \frac{1}{2\pi}\log(1-|b(\lambda)|^2)$ in eq.(13.43), thus $a(\lambda)$ can be reconstructed from the knowledge of $|b(\lambda)|$ on the real axis and the zeroes λ_n.

For a smooth field φ, $\rho(\mu)$ decreases fastly at $\mu \to 0, \infty$ so that one can expand $1/(\lambda-\mu)$ in powers of λ/μ or μ/λ in the integral. This allows to identify the $I_n^\pm$: in fact we have the asymptotic expansions at $\lambda = 0$ and $\lambda = \infty$:

$$\begin{aligned}
\log a(\lambda) &= -\frac{i\pi Q}{2} + \sum_{n=1}^{\infty} \frac{1}{\lambda^n}\left(-\sum_j \frac{\lambda_j^n - \lambda_j^{*n}}{n} + i\int_{-\infty}^{\infty} \mu^{n-1}\rho(\mu)d\mu\right), \quad \lambda \to \infty \\
\log a(\lambda) &= \frac{i\pi Q}{2} + \sum_{n=1}^{\infty} \lambda^n \left(-\sum_j \frac{1}{n}\left(\frac{1}{\lambda_j^n} - \frac{1}{\lambda_j^{*n}}\right) - i\int_{-\infty}^{\infty} \mu^{-n-1}\rho(\mu)d\mu\right), \quad \lambda \to 0
\end{aligned}$$

From this we get $I_n^\pm = 0$ for n even, and for n odd we have $I_n^\pm = \pm I_{\pm n}$ where I_n is defined for $n \in \mathbb{Z}$ by:

$$I_n = (-1)^{(n-1)/2}\Big(\sum_j \frac{\lambda_j^n - \lambda_j^{*n}}{in} - \int_0^{\infty} \mu^{n-1}\rho(\mu)d\mu\Big) \in \mathbb{R}$$

In particular for $n = \pm 1$ we get:

$$H \pm P = \frac{4m}{\beta^2}\left(\pm\sum_j \frac{\lambda_j^{\pm 1} - \lambda_j^{*\pm 1}}{i} - \int_{-\infty}^{\infty} \mu^{\pm 1}\rho(\mu)\frac{d\mu}{\mu}\right)$$

If for simplicity we don't consider breathers and set $\lambda_j = i\xi_j$ we get:

$$\begin{aligned}
H &= \frac{4m}{\beta^2}\left(\sum_j (\xi_j + \xi_j^{-1}) + \int_0^{\infty} (\mu + \mu^{-1})|\rho(\mu)|\frac{d\mu}{\mu}\right) \\
P &= \frac{4m}{\beta^2}\left(\sum_j (\xi_j - \xi_j^{-1}) + \int_0^{\infty} (\mu - \mu^{-1})|\rho(\mu)|\frac{d\mu}{\mu}\right)
\end{aligned}$$

Setting $k_j = \frac{8}{\beta^2}\frac{m}{2}(\xi_j - \xi_j^{-1})$, $M = \frac{8}{\beta^2}m$ and $k = \frac{m}{2}(\mu - \mu^{-1})$ this can be written as:

$$\begin{aligned}
H &= \left(\sum_j \sqrt{k_j^2 + M^2} + \int_{-\infty}^{\infty} \sqrt{k^2+m^2}|\frac{8}{\beta^2}\rho(k)|\frac{dk}{\sqrt{k^2+m^2}}\right) \\
P &= \left(\sum_j k_j + \int_{-\infty}^{\infty} k|\frac{8}{\beta^2}\rho(k)|\frac{dk}{\sqrt{k^2+m^2}}\right)
\end{aligned}$$

which nicely exhibits the decomposition of the theory into a sum of relativistic modes. Note that the soliton j has mass M and momentum k_j, while the continuous spectrum is a superposition of modes of mass m.

Remark. One can extract the Poisson brackets of the solitonic modes from eq.(13.42) and recover eq.(12.29) in Chapter [12]. Indeed, the parameters a_n, μ_n in Chapter [12] are related to m_n, λ_n in this Chapter by $\lambda_n = \frac{i}{\mu_n}, m_n = a_n$ due to eq.(13.35). Moreover, from eqs.(13.31,13.43):

$$m_n = \frac{c_n}{a'(\lambda_n)}, \quad a(\lambda) = e^{-i\frac{\pi Q}{2}}\prod \frac{\lambda - \lambda_i}{\lambda + \lambda_i}$$

for purely solitonic solutions, so that:

$$m_n = 4ie^{i\frac{\pi Q}{2}} c_n \prod_{j\neq n} \frac{\mu_j + \mu_n}{\mu_j - \mu_n}$$

Then a strightforward computation using eqs.(13.42), which mean that $\log \mu_n$ and $\log c_n$ are canonically conjugated variables, yields:

$$\{\mu_i, \mu_j\} = 0, \quad \{\mu_i, m_j\} = \frac{\beta^2}{2}\mu_i m_j \delta_{ij}, \quad \{m_i, m_j\} = \frac{\beta^2}{2}\left(\frac{4\mu_i\mu_j}{\mu_i^2 - \mu_j^2}\right) m_i m_j$$

which identifies with eq.(12.29) up to the factor $\frac{\beta^2}{2}$.

With this we end up this Chapter on the classical inverse scattering method, thereby paying due tribute to Gardner, Green, Kruskal and Miura without whom this book would not exist.

References

[1] C. S. Gardner, J. M. Greene, M. D. Kruskal, R. M. Miura. *"Method for solving the Korteweg-de Vries equation"*. Phys. Rev. Lett. 19 (1967), p. 1095.

[2] E. K. Sklyanin. *"On the complete integrability of the Landau-Lifchitz equation"*. Preprint LOMI E-3-79. Leningrad 1979.

[3] S. Novikov, S.V. Manakov, L.P. Pitaevskii, V.E. Zakharov. **"Theory of solitons, the inverse scattering method."** Consultants Bureau. NY 1984.

[4] L. D. Faddeev, L. A. Takhtajan. **"Hamiltonian Methods in the Theory of Solitons"**. Springer 1986.

Chapter 14

Symplectic geometry.

14.1 Poisson manifolds and symplectic manifolds.

In this Chapter we investigate the formulation of classical mechanics using Poisson brackets. We work on a phase space M and consider the differentiable functions on M. We call $\mathcal{F}(M)$ the algebra of such functions. A Poisson bracket is a bilinear antisymmetric derivation of the algebra $\mathcal{F}(M)$: $\{\ ,\ \}: \mathcal{F}(M)\times\mathcal{F}(M) \to \mathcal{F}(M)$ such that:

$$\begin{aligned}
&\{f_1, f_2\} = -\{f_2, f_1\} \quad \text{antisymmetry} \\
&\{f_1, \alpha f_2 + \beta f_3\} = \alpha\{f_1, f_2\} + \beta\{f_1, f_3\}, \quad \alpha, \beta \text{ constants} \\
&\{f_1, f_2 f_3\} = \{f_1, f_2\}f_3 + f_2\{f_1, f_3\}, \quad \text{Leibnitz rule} \\
&\{f_1, \{f_2, f_3\}\} + \{f_3, \{f_1, f_2\}\} + \{f_2, \{f_3, f_1\}\} = 0, \quad \text{Jacobi identity}
\end{aligned} \tag{14.1}$$

Since the Poisson bracket is linear in f and obeys the Leibnitz rule, to any function $H \in \mathcal{F}(M)$ we can associate a vector field X_H on M by $X_H f = \{H, f\}$. When H is the Hamiltonian of the system, this vector field defines the time evolution by

$$\dot{f} = X_H f = \{H, f\} \quad \forall f \in \mathcal{F}(M)$$

Definition. *A Poisson manifold M is a manifold on which is defined a Poisson bracket.*

The important feature of Poisson brackets in classical mechanics is that if f_1 and f_2 are two conserved quantities, i.e. $\{H, f_1\} = \{H, f_2\} = 0$, then $\{f_1, f_2\}$ is also conserved due to the Jacobi identity, eq.(14.1).

In general the Poisson bracket is degenerate, which means that there are functions f on M such that $\{f, g\} = 0$ for all g. The set of such functions f is called the center of the Poisson algebra. If the center is non trivial, i.e. contains non constant functions, one can reduce the dynamical system by setting all functions of the center to constant values. This defines a foliation of M and the Poisson bracket is non–degenerate on the leaves.

In a local coordinate system x^j on M we can write:

$$\{\ f_1,\ f_2\}_M = \sum_{ij} \omega^{ij}(x) \frac{\partial f_1}{\partial x^i} \frac{\partial f_2}{\partial x^j}$$

due to the bilinearity and Leibnitz rule. Antisymmetry requires that $\omega^{ij}(x) = -\omega^{ji}(x)$. The Jacobi identity reads:

$$\omega^{is}\partial_s\omega^{jk} + \omega^{ks}\partial_s\omega^{ij} + \omega^{js}\partial_s\omega^{ki} = 0$$

Assume now that the matrix ω^{ij} is invertible, so that in particular the center of the Poisson algebra is trivial. Note that this can occur only when the dimension of M is even. Denote by ω_{ij} the inverse matrix of ω^{ij} so that $\partial_s\omega^{ij} = -\omega^{ia}\partial_s\omega_{ab}\omega^{bj}$. Plugging this into the Jacobi identity yields:

$$\sum_{s,a,b}\omega^{is}\omega^{ja}\omega^{kb}\Big(\partial_s\omega_{ab} + \partial_b\omega_{sa} + \partial_a\omega_{bs}\Big) = 0$$

Since ω^{ij} is invertible this is equivalent to the *linear* conditions:

$$\partial_a\omega_{bc} + \partial_c\omega_{ab} + \partial_b\omega_{ca} = 0$$

These conditions can be interpreted as the closedness of the two–form:

$$\omega = -\frac{1}{2}\sum_{ij}\omega_{ij}dx^i \wedge dx^j$$

that is $d\omega = 0$. This 2–form is invariant under changes of coordinates, and so is globally defined on M. We emphasize that the matrices entering the definition of the Poisson bracket and the symplectic form are inverse to each other.

Definition. *A symplectic manifold* (M, ω) *is a manifold* M *equipped with a non degenerate closed two form* ω, $d\omega = 0$.

Given a function H on a symplectic manifold, the Hamiltonian vector fields X_H is defined using the interior product i_X by:

$$dH = -i_{X_H}\omega \quad \text{i.e.} \quad dH = -\omega(X_H, \cdot)$$

Using local coordinates x^i on M we have $X_H = \sum_i X_H^i\partial_i$ with $X_H^i = -\omega^{ij}\partial_j H$. Knowing the symplectic form ω one can reconstruct the Poisson bracket as follows:

$$\{\ f_1,\ f_2\} = X_{f_1}(f_2) = -X_{f_2}(f_1) = \omega(X_{f_1}, X_{f_2}) \tag{14.2}$$

On a symplectic space, one can define the notion of symplectic transformations. Consider a bijection $\gamma : M \to M$ and the transform of the symplectic form ω under γ. This is the form $(\gamma^*\omega)_m(V, W) = \omega_{\gamma(m)}(\gamma_*V, \gamma_*W)$. We say that the transformation γ is symplectic if $\gamma^*\omega = \omega$. For an infinitesimal transformation, γ is specified by a vector field X on M, and this condition is equivalent to $\mathcal{L}_X\omega = 0$, where $\mathcal{L}_X$ is the Lie derivative, $\mathcal{L}_X = di_X + i_X d$.

We translate this property on Poisson brackets and show that it reads:

$$\gamma\{f, g\} = \{\gamma f, \gamma g\} \tag{14.3}$$

We recall that the action of γ on functions is given by $(\gamma f)(m) = f(\gamma^{-1}(m))$. Note that we have introduced the inverse of γ so that the property $\gamma_1 \cdot (\gamma_2 \cdot f) = (\gamma_1\gamma_2) \cdot f$ holds. If γ is symplectic we have:

$$X_{\gamma f}(m) = \gamma_*\Big(X_f(\gamma^{-1}(m))\Big)$$

This is because d commutes to pullbacks, i.e. $d(\gamma f) = \gamma^{-1*}df$, so that applying it to $V \in T_mM$ we get $d(\gamma f)_m(V) = df_{\gamma^{-1}(m)}(\gamma_*^{-1}V)$, that is:

$$\omega_m(X_{\gamma f}, V) = \omega_{\gamma^{-1}(m)}(X_f(\gamma^{-1}(m)), \gamma_*^{-1}V) = \omega_m(\gamma_*X_f \circ \gamma^{-1}(m), V)$$

where in the last step we have used that γ is symplectic. We use this to prove eq.(14.3). We have $\{\gamma f, \gamma g\}_m = \omega_m(X_{\gamma f}(m), X_{\gamma g}(m)) = \omega_{\gamma(\gamma^{-1}(m))}(\gamma_*X_f \circ \gamma^{-1}, \gamma_*X_g \circ \gamma^{-1})$. Using that γ is symplectic, this is equal to $\omega_{\gamma^{-1}(m)}(X_f \circ \gamma^{-1}(m), X_g \circ \gamma^{-1}(m)) = \{f, g\}(\gamma^{-1}(m)) = (\gamma\{f, g\})(m)$.

Proposition. *Any Hamiltonian flow is a symplectic transformation.*

Proof. Let H be the Hamiltonian with associated vector field X_H such that $dH = -i_{X_H}\omega$. We have:

$$\mathcal{L}_{X_H}\omega = (i_{X_H}d + di_{X_H})\omega = d(i_{X_H}\omega) = -d^2 H = 0$$

where we have used $d\omega = 0$. ∎

Example 1. A standard example of symplectic space is given by $M = \mathbb{R}^{2n}$ with coordinates (p_i, q_i), and symplectic form:

$$\omega = \sum_i dp_i \wedge dq_i$$

These coordinates are called canonical coordinates. The corresponding Poisson bracket reads:

$$\{q_i, q_j\} = 0, \quad \{p_i, p_j\} = 0, \quad \{p_i, q_j\} = \delta_{ij}$$

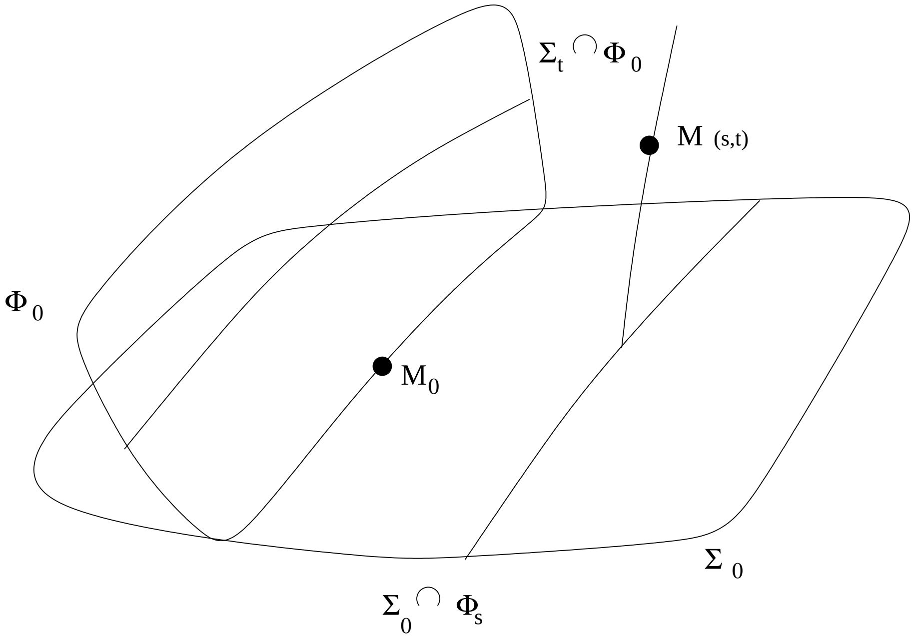
The Hamiltonian vector field corresponding to the function H is:

$$X_H = \sum_i -\frac{\partial H}{\partial q_i}\partial_{p_i} + \frac{\partial H}{\partial p_i}\partial_{q_i}$$

In fact this example is generic, at least locally. This is the Darboux theorem.

Theorem. *On any symplectic manifold (M, ω) one can introduce,* locally *around a point m_0, canonical coordinates (q_i, p_i) such that $\omega = \sum_i dp_i \wedge dq_i$. Moreover one can choose p_1 to be any given function H on M such that $dH(m_0) \neq 0$.*

Proof. The proof will be done by induction on the dimension $(2n)$ of M. We choose a coordinate system $y \in \mathbb{R}^{2n}$ around m_0, and define $p_1(y) = H(y)$. We can assume that $p_1(m_0) = 0$. We then introduce the Hamiltonian vector field X_{p_1} associated to p_1, which is non vanishing in the considered neighborhood of m_0. Choose an hypersurface Σ_0 passing through m_0, and transverse to X_{p_1}, i.e. such that X_{p_1} is not tangent to Σ_0. We want to define q_1 such that $\{p_1, q_1\} = 1$ and $q_1 = 0$ on Σ_0. This means that on any trajectory of the flow of X_{p_1} we want to achieve $\dot{q}_1 = \{p_1, q_1\} = 1$ hence $q_1(y) - q_1(z) = q_1(y) = t$, if z is the point at which the trajectory crosses Σ_0. By the assumption of transversality, for any y close to Σ_0 one defines $q_1(y)$ as the time needed to go from $z \in \Sigma_0$ to y along the trajectory of the flow X_{p_1}. The neighborhood of m_0 is foliated by the hypersurfaces Σ_t where $q_1(y) = t$. On the other hand, it is also foliated by the hypersurfaces Φ_s where $p_1(y) = s$.



We now show that one can simultaneously solve:

$$\partial_s y = -X_{q_1}(y), \quad \partial_t y = X_{p_1}(y)$$

hereby allowing to write $y = y(s,t,z)$ with $z \in \Sigma_0 \cap \Phi_0$ being the initial condition. This is because for any function $f(y)$ one has by definition $\partial_s f = -X_{q_1} f = \{-q_1, f\}$ and $\partial_t f = X_{p_1} f\{p_1, f\}$ so that, using the Jacobi identity:

$$\partial_s(\partial_t f) - \partial_t(\partial_s f) = -\{q_1, \{p_1, f\}\} + \{p_1, \{q_1, f\}\} = \{\{p_1, q_1\}, f\} = 0$$

since $\{p_1, q_1\} = 1$ is constant.

The vector field X_{p_1} is tangent to Φ_0 (in fact to any Φ_s, because $X_{p_1} p_1 = \{p_1, p_1\} = 0$), and similarly X_{q_1} is tangent to Σ_t. They are both transverse to $\Sigma_0 \cap \Phi_0$ and independent, because $\{p_1, q_1\} = \omega(X_{p_1}, X_{q_1}) = 1 \neq 0$. For any vector V tangent to $\Sigma_0 \cap \Phi_0$ we have $\omega(V, X_{p_1}) = \omega(V, X_{q_1}) = 0$ because $\omega(V, X_{p_1}) = V \cdot p_1$, $\omega(V, X_{q_1}) = V \cdot q_1$, and $p_1 = q_1 = 0$ are constant on this intersection. It follows that the restriction of ω to $\Sigma_0 \cap \Phi_0$ is non degenerate.

By the induction hypothesis we assume that we have found canonical coordinates p_i, q_i, $i \geq 2$ on the $(2n-2)$ dimensional symplectic variety $\Sigma_0 \cap \Phi_0$. We extend these coordinates as functions on M by setting $p_j(y) = p_j(z)$ and $q_j(y) = q_j(z)$ for $j \geq 2$ and $y = y(p_1, q_1, z)$ with $z \in \Sigma_0 \cap \Phi_0$. This amounts to keep them constant along the flows of X_{p_1} and X_{q_1} so that

$$\{p_1, p_j\} = \{p_1, q_j\} = \{q_1, p_j\} = \{q_1, q_j\} = 0$$

It remains to show that the symplectic form ω on M is equal to $\widetilde{\omega} = \sum_{i=1}^n dp_i \wedge dq_i$. We first show that this is true at any point $p_1 = q_1 = 0$ of $\Sigma_0 \cap \Phi_0$. Any vector V tangent to M at this point can be decomposed as a sum of a vector V_1 on the space spanned by X_{p_1} and X_{q_1}, and a vector V_2 tangent to $\Sigma_0 \cap \Phi_0$. Computing $\omega(V_1 + V_2, W_1 + W_2)$ we have seen that $\omega(V_1, W_2) = \omega(V_2, W_1) = 0$, while by the induction hypothesis $\omega(V_2, W_2) = (\sum_{j \geq 2} dp_j \wedge dq_j)(V_2, W_2)$. Finally $\omega(V_1, W_1) = (dp_1 \wedge dq_1)(V_1, W_1)$, because it is sufficient to compute this for $V_1 = X_{p_1}$ and $W_1 = X_{q_1}$, and both members are equal to $\{p_1, q_1\} = 1$. To show that the equality holds on M consider the Hamiltonian evolution $U_{(s,t)}$ under the flows of X_{p_1} and X_{q_1}. In the coordinates (p_i, q_i) it reads $(p_1, q_1, p_2, q_2 \cdots) \to (p_1 + s, q_1 + t, p_2, q_2, \cdots)$ hence the form $\widetilde{\omega} = \sum_i dp_i \wedge dq_i$ is obviously invariant. On the other hand the symplectic form ω is also invariant under Hamiltonian evolutions. Since we have shown that $\omega = \widetilde{\omega}$ for $p_1 = q_1 = 0$ these two forms coincide everywhere. ∎

Example 2. Another very natural example of symplectic space is the cotangent bundle $M = T^*N$ of any differentiable manifold N. On this bundle is defined a canonical one form α given by:

$$\alpha_x(X) = p(\pi_* X)$$

where $X \in T_x(T^*N),\ x \in T^*N$, so that one can write $x = (q, p)$, with $q = \pi(x) \in N$ (π is the projection on N), and p is a one form belonging to the fiber of q. Finally we define the symplectic form as the closed form $\omega = d\alpha$. To show that it is non degenerate we express it in terms of local coordinates. If $q = (q_1, \cdots, q_n)$ is a system of local coordinates on N, a basis of the tangent space of N at q is $(\partial_{q_1}, \cdots, \partial_{q_n})$ and a basis of the cotangent space at q is $(dq_1, \cdots, dq_n)$. In particular any point p in the fiber of T^*N above q is of the form $p = \sum_i p_i dq_i$. A tangent vector $X = (\delta q, \delta p)$ at the point (q, p) of T^*N has projection $\pi_* X = \delta q = \sum_i (\delta q)_i \partial_{q_i}$, so that $\alpha(X) = \sum_i p_i (\delta q)_i$, and the canonical form reads:

$$\alpha = p_1 dq_1 + \cdots + p_n dq_n$$

Then $\omega = d\alpha = \sum_i dp_i \wedge dq_i$ is non degenerate.

14.2 Symmetries and Hamiltonian reduction.

Let G be a Lie group and $\mathcal{G}$ its Lie algebra. Consider a symplectic manifold M on which G operates. We say that the action of G is symplectic if for any $g \in G$ the transformation $m \to gm$ is symplectic. In view of eq.(14.3) this means:

$$\{f_1(gm), f_2(gm)\} = \{f_1, f_2\}(gm)$$

For any $X \in \mathcal{G}$ consider the one–parameter group $g^t = \exp(tX)$. In the limit $t \to 0$ we define the action of X on functions by:

$$X \cdot f(m) = \frac{d}{dt} f(e^{-tX} \cdot m)|_{t=0} \tag{14.4}$$

This means that, considering the vector field $X \cdot m$ at $m \in M$ we have:

$$X \cdot f = -\mathcal{L}_{X \cdot m} f$$

Remark that by the general definition of the Lie derivative, we have:

$$\mathcal{L}_{X \cdot m} Y \cdot m \equiv [X \cdot m, Y \cdot m] = \frac{d}{dt}\Big(e^{-Xt} Y.(e^{Xt} m)\Big)|_{t=0} = -[X, Y] \cdot m$$

so that we get a representation on functions of the Lie algebra structure:

$$X \cdot (Y \cdot f) - Y \cdot (X \cdot f) = [X, Y] \cdot f \tag{14.5}$$

Finally the symplecticity condition reads:

$$\{X \cdot f_1, f_2\} + \{f_1, X \cdot f_2\} = X \cdot \{f_1, f_2\} \tag{14.6}$$

Proposition. *Let G be a Lie group acting on M by symplectic diffeomorphism. The action of any one parameter subgroup of G is locally Hamiltonian. This means that there exists a function H_X, locally defined on M, such that:*

$$X \cdot f \ = \ \{H_X, f\} \tag{14.7}$$

Proof. The condition eq.(14.6) is obviously necessary to have $X \cdot f = \{H_X, f\}$. It is also sufficient. Indeed, in terms of the canonical Darboux coordinates, we write $X \cdot m = \sum_i (X^{p^i}\partial_{p^i} + X^{q^i}\partial_{q^i})$ and we have:

$$\begin{aligned}\{X\cdot f, h\} + \{f, X\cdot h\} - X\cdot\{f,h\} &= \sum_{i,j}\left[(\partial_{p_j}X_{q_i} - \partial_{p_i}X_{q_j})\partial_{q_i}f\partial_{q_j}h + (\partial_{p_j}X_{p_i} + \partial_{q_i}X_{q_j})\partial_{p_i}f\partial_{q_j}h\right] \\ &\quad -\sum_{i,j}\left[(\partial_{q_j}X_{q_i} + \partial_{p_i}X_{p_j})\partial_{q_i}f\partial_{p_j}h + (\partial_{q_j}X_{p_i} - \partial_{q_i}X_{p_j})\partial_{p_i}f\partial_{p_j}h\right]\end{aligned}$$

If this vanishes identically for all f and h, this is equivalent to the condition $d\Omega_X = 0$ where $\Omega_X = -i_X\omega = \sum_i (X_{q_i}dp_i - X_{p_i}dq_i)$. So there exists, at least locally, a function H_X such that $\Omega_X = dH_X$ or:

$$X_{q_i} = \frac{\partial H_X}{\partial p_i} \qquad X_{p_i} = -\frac{\partial H_X}{\partial q_i}$$

Then

$$X\cdot f = \sum_i -\frac{\partial H_X}{\partial q_i}\frac{\partial f}{\partial p_i} + \frac{\partial H_X}{\partial p_i}\frac{\partial f}{\partial q_i} = \{H_X, f\}$$

This proves eq.(14.7). ■

If one knows that there is an invariant one–form α such that $\omega = d\alpha$ one can give an explicit formula for the function H_X, which is then globally defined:

$$H_X(m) = \alpha(X\cdot m) \tag{14.8}$$

Indeed, since α is invariant we have $\mathcal{L}_X\alpha = 0$. Then $0 = \mathcal{L}_X\alpha = (i_X d + d i_X)\alpha = \omega(X,\cdot) + d\alpha(X)$, so that comparing with $dH_X = -i_X\omega$ we see that $H_X = \alpha(X\cdot m)$.

Using eq.(14.7) and the Jacobi identity, we find that $X\cdot(Y\cdot f) - Y\cdot(X\cdot f) = \{\{H_X, H_Y\}, f\}$, so that eq.(14.5) yields $\{H_{[X,Y]} - \{H_X, H_Y\}, f\} = 0$. Because constants commute with any function we cannot conclude that $H_{[X,Y]} = \{H_X, H_Y\}$. This motivates the:

Definition. *Consider a Lie group G acting on a symplectic manifold M by symplectic action. This action is said to be Poissonnian if the Hamiltonians H_X of the one parameter subgroups are globally defined, depend linearly on X, and are such that*

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

In the previous case, where there exists an invariant one–form α, this property is always satisfied. In fact, we have $\{H_X, H_Y\} = \omega(X,Y) = d\alpha(X,Y) = X\cdot\alpha(Y) - Y\cdot\alpha(X) - \alpha([X,Y])$. By invariance of α we have $X\cdot\alpha(Y) = \mathcal{L}_X\alpha(Y) = \alpha([X,Y]) = -Y\cdot\alpha(X)$, so that $\{H_X, H_Y\} = \alpha([X,Y]) = H_{[X,Y]}$.

Example. This particular situation is important because it occurs in the case of a cotangent bundle. In this case we already know that the one–form α exists and is globally defined. We consider particular diffeomorphisms of $M = T^*N$, namely those which are induced by a diffeomorphism of the base N. We are going to show that α is invariant under such diffeomorphisms, and in particular under those which are induced by group actions on N. Indeed any diffeomorphism ϕ of N induces a transformation of T^*N as follows: a point (q,p) of T^*N is determined by a point $q \in N$ and a linear form p on the tangent space T_qN to N at q. The differential of ϕ at q, which we denote by ϕ_*, maps T_qN into $T_{\phi(q)}N$. Its transpose ϕ^* maps $T^*_{\phi(q)}N$ to T^*_qN hence ${\phi^*}^{-1}$ maps T^*_qN to $T^*_{\phi(q)}N$. The induced transformation $\widehat{\phi}$ on $M = T^*N$ is given by

$$\widehat{\phi}(q,p) = (\phi(q), {\phi^*}^{-1}(p))$$

Proposition. *The one–form α on T^*N is invariant under transformations induced by transformations of the base manifold N.*

Proof. The transformation $\widehat{\phi}$ of M induces a transformation $\widehat{\phi}_*$ on TM given by $\widehat{\phi}_*(\delta q, \delta p) = (\phi_* \delta q, {\phi^*}^{-1} \delta p)$. Recall the definition $\alpha(\delta q, \delta p) = p(\delta q)$. Quite generally we have:

$$(\widehat{\phi}^* \alpha)_{(q,p)}(\delta q, \delta p) = \alpha_{\widehat{\phi}(q,p)}(\widehat{\phi}_*(\delta q, \delta p)) = {\phi^*}^{-1}(p)(\phi_* \delta q) = p(\delta q) = \alpha(\delta q, \delta p)$$

■

In particular assuming that a Lie group G acts on the base manifold N, this action lifts to a Poissonnian action on $M = T^*N$, and for any $X \in \mathcal{G}$ we have $H_X(m) = \alpha(X \cdot m) = p(X \cdot q)$ for $m = (q, p)$.

When the action of a Lie group G on a symplectic manifold M is Poissonnian, to any $X \in \mathcal{G}$ is associated a function H_X such that $X \cdot f = \{H_X, f\}$, and $X \to H_X$ is linear. Hence there exists a function $\mathcal{P} : M \longrightarrow \mathcal{G}^*$ such that one can write $H_X(m) =< \mathcal{P}(m), X >$, where $<>$ is the pairing between $\mathcal{G}$ and its dual.

Definition. *The application $m \to \mathcal{P}(m) \in \mathcal{G}^*$ is called the moment map.*

The moment map has the following covariance property with respect to the action of G:

Proposition. *The value of the moment at the point $g \cdot m$ is related to its value at the point m by:*

$$\mathcal{P}(g \cdot m) = \mathrm{Ad}^*_g \mathcal{P}(m)$$

Proof. This is equivalent to $H_X(g \cdot m) = H_{g^{-1}Xg}(m)$. Since we asssume that G is connected it is sufficient to show this for an infinitesimal $g = 1 + \epsilon Y$. Then we have to show that $dH_X(Y \cdot m) = H_{[X,Y]}(m)$. Using eq.(14.4) we have $dH_X(Y \cdot m) = -Y \cdot H_X = -\{H_Y, H_X\}$ where we used eq.(14.7) in the last step. Since the group action is Poissonnian this is $H_{[X,Y]}$. ■

The moment map associates conserved quantities to symmetries of the Hamiltonian. This is the Noether theorem.

Theorem. *Let G be a Lie group acting by Poissonnian action on a symplectic manifold M, and let H be a Hamiltonian invariant under the action of G. Then the moment $\mathcal{P}$ is conserved under the flow of H.*

Proof. Let us fix $X \in \mathcal{G}$ and consider the function on M: $< \mathcal{P}(m), X >= H_X(m)$. Its time derivative under the flow of H is $\partial_t H_X = \{H, H_X\} = -X \cdot H = 0$ because H is invariant under the group action. ■

In the situation of the theorem, one can use the conserved quantities to reduce the number of degrees of freedom. This is called Hamiltonian reduction, because one is able to define a new symplectic variety of smaller dimension on which the reduced motion takes place.

Let M be a symplectic manifold and let G be a group acting on M by Poissonnian action. Let $\mathcal{P}$ be the moment map. Let us fix a particular value μ of the moment and consider the set of points of phase space where $\mathcal{P}(m) = \mu$. By the Noether theorem the motion takes place on this set:

$$M_\mu \equiv \mathcal{P}^{-1}(\mu) \qquad \mu \in \mathcal{G}^*$$

We have to assume that μ is not a critical value of $\mathcal{P}$, that is, at all $m \in M_\mu$ we have $d\mathcal{P}(m) \neq 0$. Hence there exists a tangent space at m to M_μ. Since M_μ is defined by $\dim G$ equations, we have:

$$\dim M_\mu = \dim M - \dim G \tag{14.9}$$

Note that M_μ is not in general a symplectic variety, and is not in general of even dimension. However there is a residual action of the group G on M_μ. Indeed we have seen that the action of G is transformed by $\mathcal{P}$ into the coadjoint action on $\mathcal{G}^*$. So the stabilizer G_μ of the fixed moment μ , that is the group of $g \in G$ such that $\mathrm{Ad}_g^* \mu = \mu$, preserves M_μ.

Definition. *The reduced phase space is the quotient:*

$$F_\mu \equiv M_\mu / G_\mu = \mathcal{P}^{-1}(\mu)/G_\mu$$

Here we assume that the quotient is well–defined as a differentiable manifold. In general there are particular values of μ where this quotient is ill–defined. However for generic situation we don't have to enter into such subtleties.

The nice feature of the reduced phase space is that it is naturally equipped with a symplectic structure, and in particular is of even dimension.

Proposition. *The following bilinear form defines a symplectic structure on F_μ: let ξ and η be two vectors tangent to F_μ at the point f. Consider at a point $m \in M_\mu$ above f any two tangent vectors to M_μ, ξ' and η' projecting on ξ, η respectively. We then set:*

$$\omega_f(\xi, \eta) = \omega_m(\xi', \eta')$$

This is independent of the chosen representatives m, ξ', η' and defines a symplectic form on F_μ.

Proof. We first show that given m, $\omega_m(\xi', \eta')$ is independant of the choices of representatives ξ', η'. This amounts to showing that $\omega_m(V, W) = 0$ for V vertical, i.e. tangent to the orbit of G_μ and W tangent to M_μ. Since V is vertical we can write $V = X \cdot m$ with $X \in \mathcal{G}_\mu$. We can consider the function H_X defined on M and note that $\omega_m(V, W) = -dH_X(W)$. But on M_μ we have $H_X = \mathcal{P}(X) = \mu(X)$. Since $dH_X(W)$ is the derivative of H_X in the direction of W which is tangent to M_μ this derivative vanishes.
The quantity $\omega_m(\xi', \eta')$ is independant of the choice of m since by invariance of ω we have for any $g \in G_\mu$ $\omega_{gm}(g\xi', g\eta') = \omega_m(\xi', \eta')$. This shows that ω_F, the form that we have defined on F_μ is well–defined, bilinear antisymmetric.
To show that ω_F is closed, note that the restriction $\omega|_{M_\mu}$ of ω to M_μ is obviously closed. On the other hand if π is the projection $M_\mu \to M_\mu/G_\mu$ we have just shown that $\omega|_{M_\mu} = \pi^* \omega_F$. Since d commutes with π^* we have $\pi^* d\omega_F = 0$. This means that $d\omega_F(\pi_* X, \pi_* Y) = 0$ for all X, Y. Since π_* is surjective we have $d\omega_F = 0$.
Finally ω_F is non–degenerate. We have seen that vertical vectors are in the kernel of $\omega|_{M_\mu}$. In fact they are the whole kernel of $\omega|_{M_\mu}$. Indeed we have seen that vertical vectors are orthogonal under ω to $T_m(M_\mu)$. Moreover we have

$$(\mathcal{G} \cdot m)^\perp = T_m(M_\mu)$$

because both sides have the same dimension $\dim M - \dim G$ since ω is non degenerate. But the kernel of the restriction of ω to M_μ is $\mathrm{Ker}\, \omega|_{M_\mu}(m) = T_m(M_\mu) \cap T_m^\perp(M_\mu) = \mathcal{G}_\mu \cdot m$. These vectors project to 0 when we take the quotient under G_μ, hence ω_F is non–degenerate. ■

We often need to compute the Poisson brackets of functions on the reduced phase space F_μ, knowing the Poisson brackets on M. Any function $\widetilde{f}$ on F_μ uniquely extends to a G_μ invariant function on M_μ. However to be able to compute Poisson brackets on M we have to extend further this function to a function f defined in a vicinity of M_μ in M. Even requiring complete G invariance is not sufficient to lift $\widetilde{f}$ to M. This is because while $\dim M = \dim M_\mu + \dim G$, the fiber along G at $m \in M_\mu$ is not transverse to M_μ, since G_μ leaves M_μ invariant. The general procedure consists in choosing arbitrary extensions f of $\widetilde{f}$ outside of

M_μ. Two extensions differ by a function vanishing on M_μ. Then we will show how to compute the reduced Poisson bracket as some modification of $\{f, g\}$ (computed on M) independent of the arbitrary choices (see eq.(14.10) below). The difference of the Hamiltonian vector fields of two extensions of the same function on F_μ is controlled by the following:

Lemma. *Let f be a function defined in a vicinity of M_μ and vanishing on M_μ. Then the Hamiltonian vector field X_f associated to f is tangent to the orbit $G \cdot m$ at any point $m \in M_\mu$.*

Proof. The subvariety M_μ is defined by the equations $\mathrm{H}_{X_i} = \mu_i$ for some basis X_i of $\mathcal{G}$. Since f vanishes on M_μ one can write $f = \sum (H_{X_i} - \mu_i) f_i$ for some functions f_i defined in the vicinity of M_μ. For any tangent vector V at a point $m \in M_\mu$ one has, using that $\sum (H_{X_i} - \mu_i)\, \partial f_i$ vanishes on M_μ:

$$df(m)(V) = \sum_i dH_{X_i}(m)(V) f_i(m) = -\omega\,(\sum_i f_i(m)\, X_i \cdot m\,,\, V\,)$$

since the Hamiltonian vector field associated to H_{X_i} is $X_i.m$. Hence $X_f = \sum f_i(m)\, X_i \cdot m \in \mathcal{G} \cdot m$.

As a consequence of this lemma we have a method to compute the reduced Poisson bracket. We take two functions defined on M_μ and invariant under G_μ and extend them arbitrarily. Then we compute their Hamiltonian vector fields on M. It turns out that they can be "projected" on the tangent space to M_μ by adding a vector tangent to the orbit $G \cdot m$. These projections are independent of the extensions and the reduced Poisson bracket is given by the value of the symplectic form on M acting on them.

Proposition. *Let f be a function defined in a vicinity of M_μ and G_μ–invariant on M_μ. At each point $m \in M_\mu$ one can choose a vector $V_f \cdot m \in \mathcal{G} \cdot m$ such that $X_f + V_f \cdot m \in T_m(M_\mu)$ and $V_f \cdot m$ is determined up to a vector in $\mathcal{G}_\mu \cdot m$.*

Proof. Recall that the symplectic orthogonal of $\mathcal{G} \cdot m$ is exactly $T_m(M_\mu)$, so we want to solve $\omega\,(X_f + V_f \cdot m, X \cdot m) = 0$, $\forall X \in \mathcal{G}$. Note that for $X, Y \in \mathcal{G}$, $\omega(X \cdot m, Y \cdot m) = \{H_X, H_Y\} = H_{[X,Y]} = \mathcal{P}\,([X, Y]) = \mu\,([X, Y])$ since the action is Poissonnian and $m \in M_\mu$. So the equation to be solved reads $\mathcal{L}_{X \cdot m} f = \mu\,([V_f, X])$. Both members are linear in $X \in \mathcal{G}$ and vanish when $X \in \mathcal{G}_\mu$ (the left–hand side because f is G_μ invariant, the right–hand side because G_μ stabilizes μ), so the equation can be seen as an equation on $\mathcal{G}/\mathcal{G}_\mu$. The mapping on this quotient $(\overline{X}, \overline{Y}) \to \mu([X, Y])$ is a skew–symmetric non degenerate (since we have quotiented by the kernel $\mathcal{G}_\mu$) bilinear form, hence the equation can be uniquely solved for $\overline{V_f}$. ∎

Proposition. *Let $\widetilde{f}$, $\widetilde{g}$ be two functions on the reduced phase space F_μ. We lift them to functions f, g defined on a vicinity of M_μ, and G_μ–invariant on M_μ. The reduced Poisson bracket is given by:*

$$\{\widetilde{f}, \widetilde{g}\}_{\text{reduced}} = \omega\,(X_f + V_f \cdot m, X_g + V_g \cdot m) = \{f, g\} - \mu([V_f, V_g]) \tag{14.10}$$

Proof. We want to show that the Hamiltonian vector field associated to $\widetilde{f}$ for the reduced symplectic form is $\pi_*(X_f(m) + V_f \cdot m)$ for $m \in M_\mu$. Note that this is independent of the choice of V_f modulo $\mathcal{G}_\mu$, and that $X_f(m) + V_f \cdot m$ is by construction tangent to M_μ. If V is an arbitrary tangent vector to M_μ at m, by definition of the reduced symplectic form, one has to check that:

$$d\widetilde{f}(\pi_* V) = -\omega_{\text{reduced}}(\pi_*(X_f(m) + V_f \cdot m), \pi_* V) = -\omega_m(X_f(m) + V_f \cdot m, V) \tag{14.11}$$

We have seen that the symplectic orthogonal of $\mathcal{G} \cdot m$ is $T_m(M_\mu)$ so that $\omega_m(V_f \cdot m, V) = 0$. On the other hand, since f is G_μ invariant we have $d\widetilde{f}(\pi_* V) = df(V) = -\omega_m(X_f, V)$. This proves eq.(14.11) and the first equality in eq.(14.10). To get the second form of the reduced Poisson bracket remark that $\omega(X_f, V_g \cdot m) = -\mu([V_f, V_g])$.

Indeed since $X_f + V_f \cdot m \in (\mathcal{G}\cdot m)^\perp$ one has $\omega(X_f, V_g\cdot m) = -\omega(V_f\cdot m, V_g\cdot m) = -H_{[V_f,V_g]}(m) = -\mu([V_f, V_g])$ for $m \in M_\mu$. ■

Remark. Note that if $f_{|\, M_\mu} = 0$ we have $X_f + V_f\cdot m \in \mathcal{G}_\mu \cdot m$ hence $\{\widetilde{f}, \widetilde{g}\}_{\text{reduced}} = \omega\,(X_f + V_f\cdot m, X_g + V_g\cdot m) = 0$ so that eq.(14.10) is independent of the arbitrary choices of f and g.

In the applications we have in mind, we are given functions f and g on M which are G–invariant. It is then obvious that $\{f, g\}$ is also G–invariant (invariance of ω) hence its restriction to M_μ is G_μ–invariant. Moreover the associated Hamiltonian vector fields X_f, X_g are tangent to M_μ since the G–invariance of f implies:

$$0 = df\,(X\cdot m) = -\omega\,(X_f, X\cdot m) = -dH_X(X_f),\ X\in\mathcal{G}$$

therefore the functions H_X are constant along X_f i.e. X_f is tangent to M_μ, so that one can take $V_f = 0$. It follows that for such functions the reduced Poisson bracket on F_μ is simply given by the ordinary Poisson bracket on M.

Proposition. *If f and g are G–invariant functions on M they define functions $\widetilde{f}$ and $\widetilde{g}$ on the reduced phase space F_μ and we have:*

$$\{\widetilde{f}, \widetilde{g}\}_{\text{reduced}} = \{f, g\}$$

where the right–hand side is also G–invariant and defines a function on F_μ.

14.3 The case $M = T^*G$.

Let us now apply these general considerations to the case where N is a Lie group G and $M = T^*G$. We have two actions of G on itself, namely

$$L_g:\ (g,n)\to gn,\quad R_g:\ (g,n)\to ng^{-1}$$

The first one is called left–action and the second one right–action. The differential L_{g*} of the map $n\to gn$ sends $T_n\,G\to T_{gn}\,G$. In particular if $n = e$ the unit element of G, this differential maps $T_e\,G = \mathcal{G}$ to $T_g\,G$. For any fixed $X\in\mathcal{G}$ we get a left–invariant vector field $g\cdot X$ on G. As above the action on functions on G is defined by $(gf)(n) = f(g^{-1}n)$ so that $(g(hf)) = (ghf)$. For an infinitesimal transformation $g = \exp(tX)$ with t small we have:

$$X\cdot f(n) = \frac{d}{dt} f(e^{-tX}\cdot n)|_{t=0} = -\mathcal{L}_{X\cdot n} f$$

To build a phase space we need the cotangent bundle $M = T^*\,G$. We can use left translations to identify M with $G\times\mathcal{G}^*$, where $\mathcal{G}^*$ is the dual of the Lie algebra.

$$\omega\in T^*_g\,G \longrightarrow (g,\xi)\quad\text{where}\quad \omega = L^*_{g^{-1}}\,\xi$$

Indeed $L_{g^{-1}*}$ maps $T_g\,G$ to $\mathcal{G}$, hence its transposed $L^*_{g^{-1}}$ maps $\mathcal{G}^*$ to $T^*_g\,G$. Explicitly $\omega_g(V) = \xi(g^{-1}V)$ which is simply $\xi(X)$ when $V = g\cdot X$ is a left–invariant vector field. Remark that if X_i is a basis of the Lie algebra $\mathcal{G}$ the left–invariant vector fields $g\cdot X_i$ provide a basis of each $T_g\,G$.

Since $M = T^*G$ there is a canonical one–form α defined as follows: if (v,κ) is a tangent vector to $T^*\,G$ at the point (g,ξ), so that $v\in T_g\,G$ and $\kappa\in\mathcal{G}^*$, we have:

$$\alpha(v,\kappa) = -\xi(g^{-1}v)$$

This canonical one–form is both left–invariant and right–invariant. This is because, according to its general construction it is invariant under any diffeomorphism of the base $M = G$ of the cotangent bundle. Hence the symplectic form $\omega = d\alpha$ is invariant under both actions. Note that we have introduced a minus sign to get the usual Kirillov bracket for the ξ coordinates below.

We can compute the Hamiltonians which generate infinitesimal left and right translations. We have seen that in the case of a cotangent bundle $H_X(m) = \alpha(X \cdot m)$ where $X \cdot m$ is the infinitesimal group action on M. For right translations $g \to gh^{-1}$ and $h = e^{tX_R}$ we get $H_{X_R}(g, \xi) = \alpha(-g \cdot X_R) = \xi(X_R)$. For left translations $g \to hg$ and $h = e^{tX_L}$ we get $H_{X_L}(g, \xi) = \alpha(X_L \cdot g) = -\xi(g^{-1} X_L g)$. From this, one sees that the moments for left and right actions are given by $\mathcal{P}_L(X_L) = -(g\xi g^{-1}, X_L)$ and $\mathcal{P}_R(X_R) = (\xi, X_R)$.

In the applications we consider the case when only subgroups H_L and H_R act by left and right translations on G. The moments live in the duals of the Lie algebra $\mathcal{H}_L$ and $\mathcal{H}_R$ hence require natural projections from $\mathcal{G}^*$, induced by restriction to the subalgebras. Namely if ξ is a linear form on $\mathcal{G}$ its restriction to $\mathcal{H}$ is an element of $\mathcal{H}^*$, and we denote it $P_{\mathcal{H}^*}\, \xi$. We have shown the:

Proposition. *Let H_L and H_R be two subgroups of G acting by left and right translations respectively on $T^*\, G$. The moments of these actions are:*

$$\mathcal{P}_R(g, \xi) = P_{\mathcal{H}_R^*}\, \xi \quad \mathcal{P}_L(g, \xi) = -P_{\mathcal{H}_L^*}\, g\xi g^{-1}$$

We often need to compute Poisson brackets of functions on $T^*\, G$. We have natural elementary functions on this phase space, namely the quantities $\xi(X)$ for any given $X \in \mathcal{G}$, and the matrix elements $\rho_{ij}(g)$ of g in any faithful representation of G. Any over function can be expressed as a function of these elementary ones. So it is enough to give the Poisson brackets of these elementary functions.

Proposition. *The Poisson brackets of the elementary functions on $T^*\, G$ reads:*

$$\{\rho_{ij}(g), \rho_{kl}(g)\} = 0, \quad \{\xi(X), \rho_{ij}(g)\} = \rho_{ij}(gX), \quad \{\xi(X), \xi(Y)\} = \xi([X, Y])$$

Proof. These relations are consequences of the fact that $H_X = \xi(X)$ is the Hamiltonian generating right translations. Indeed, since by the general theory, the action is Poissonnian we have $\{H_X, H_Y\} = H_{[X,Y]}$ and this gives the last equation. Moreover $\{H_X, \rho_{ij}(g)\}$ is the infinitesimal action of the right translation by X on the function $\rho_{ij}(g)$, namely the action $(h\rho_{ij})(g) = \rho_{ij}(gh)$ (note that we have here gh because this is an action on functions). So $\{H_X, \rho_{ij}(g)\} = \rho_{ij}(gX)$ proving the second equation. Finally the first equation is obvious since the $\rho_{ij}(g)$ only depend on the position variables and not the momenta. ∎

The Poisson bracket on ξ is the Kirillov bracket. We often drop the explicit reference to the representation ρ in the above formulae, but it is important to keep in mind that the g occurring in these equations is a function on phase space, and not a point on the base.

14.4 Poisson–Lie groups.

Consider two Poisson manifolds M_1 and M_2. The cartesian product $M_1 \times M_2$ is also equipped with a natural Poisson structure. Indeed the space of functions on $M_1 \times M_2$ is the tensor product of the space of functions on M_1 and of the space of functions on M_2. Namely, one can write any such function in the form $f(x, y) = \sum_i f_i^{(1)}(x) f_i^{(2)}(y)$ where the sum is in general infinite and requires some topology for its precise definition. We then define for two functions $f(x, y)$ and $g(x, y)$ the Poisson bracket:

$$\{f, g\}_{M_1 \times M_2} = \sum_{ij} \{f_i^{(1)}, g_j^{(1)}\}_{M_1} f_i^{(2)} g_j^{(2)} + \{f_i^{(2)}, g_j^{(2)}\}_{M_2} f_i^{(1)} g_j^{(1)}$$

This obeys all the properties of a Poisson bracket, and means that functions on M_1 Poisson commute with functions on M_2.

In particular, if G is a Lie group endowed with a Poisson structure, the product $G \times G$ has a Poisson structure and one may wonder whether the multiplication $(g, h) \to gh$ from $G \times G$ to G is compatible with the respective Poisson structures. More precisely, if we have two Poisson manifolds M and N and a map $\phi : M \to N$ this map is said to be Poisson if for any two functions f_1 f_2 on N we have $\{f_1 \circ \phi, f_2 \circ \phi\}_M = \{f_1, f_2\}_N \circ \phi$. In our case the multiplication is Poisson if we have:

$$\{f_1(gh), f_2(gh)\}_{G\times G} = \{f_1, f_2\}_G(gh)$$

where in the left–hand side, $f_{1,2}(gh)$ are to be viewed as functions on $G \times G$.

Definition. *A Poisson–Lie group G is a Lie group G equipped with a Poisson structure such that the multiplication in G, viewed as a map $G \times G \to G$, is a Poisson mapping.*

To describe the Poisson structure on G one can use the Lie algebra $\mathcal{G}$ to label the derivatives of any function at a point $g \in G$. We first choose a basis e_a of the Lie algebra $\mathcal{G}$

$$[e_a, e_b]_{\mathcal{H}} \;=\; f_{ab}^c \, e_c \tag{14.12}$$

and consider the right–invariant vector fields ∇_a^R defined as:

$$\nabla_a^R f(g) = \frac{d}{dt} f(e^{te_a} g)\Big|_{t=0}$$

which form a basis of the tangent space $T_g G$. The Poisson bracket of two functions f_1, f_2 on G can be written as a bilinear combination of derivatives with coefficients $\eta^{ab}(g)$ as:

$$\{f_1, f_2\}_G(g) = \sum_{a,b} \eta^{ab}(g)(\nabla_a^R f_1)(g)(\nabla_b^R f_2)(g) \tag{14.13}$$

The coefficients $\eta^{ab}(g)$ contain all the information on the Poisson structure, and we will express the Lie–Poisson condition on them. For this, it is convenient to introduce the element $\eta(g) \in \mathcal{G} \times \mathcal{G}$:

$$\eta(g) = \sum_{a,b} \eta^{ab}(g) e_a \otimes e_b$$

Any function on G can be expressed on elementary functions, namely the matrix elements of a faithful representation ρ of G. It is sufficient to express the Poisson brackets of such elementary functions. If we consider the function $g \to \rho(g)$ we have $\nabla_a^R \rho(g) = \rho(e_a g) = \rho(e_a)\rho(g)$. The Poisson bracket of two matrix elements then reads:

$$\{\rho_{ij}(g), \rho_{kl}(g)\} = \rho_{ij} \otimes \rho_{kl} \Big(\eta(g) g \otimes g\Big)$$

In the following we drop the explicit mention of the representation ρ and write this formula in the usual tensor notation:

$$\{g_1, g_2\} = \eta_{12}(g)\, g \otimes g \tag{14.14}$$

The antisymmetry of the Poisson bracket (14.13) requires $\eta_{12} = -\eta_{21}$.

Proposition. *The Lie–Poisson property is equivalent to the following cocycle property on η:*

$$\eta(gh) = \eta(g) + \mathrm{Ad}_g \cdot \eta(h)$$

where the adjoint action on $\mathcal{G} \otimes \mathcal{G}$ is defined as $\mathrm{Ad}_g \cdot \eta(h) = g \otimes g \, \eta(h) \, g^{-1} \otimes g^{-1}$.

Proof. We have $\{(gh)_1, (gh)_2\} = \{g_1, g_2\} h_1 h_2 + g_1 g_2 \{h_1, h_2\}$ since $\rho(gh) = \rho(g)\rho(h)$. This has to be equal to $\eta(gh)(gh)_1(gh)_2$. Comparing the two expressions the cocycle condition follows. ∎

This condition is called a cocycle condition because it means that η is a 1–cocycle for the Hochschild group cohomology of G with values in the representation $\mathcal{G}\otimes\mathcal{G}$. Looking at the vicinity of the identity e in G, that is writing $g=\exp(tX)$ and $h=\exp(t'Y)$, with t,t' small and $\eta(e^{tX})=\eta(e)+td_e\eta(X)+\cdots$, the cocycle condition implies at order 0 in t,t' that $\eta(e)=0$, and at second order that

$$d_e\eta([X,Y])=[\Delta X,d_e\eta(Y)]-[\Delta Y,d_e\eta(X)] \tag{14.15}$$

with $\Delta X=X\otimes 1+1\otimes X$. This means that the linear function $d_e\eta$ on $\mathcal{G}$ is a 1–cocycle for the Lie algebra cohomology with values in the same representation $\mathcal{G}\otimes\mathcal{G}$.

One can make use of $d_e\eta$ to introduce a Lie algebra structure on $\mathcal{G}^*$. For any function f on G the differential d_ef at the identity is a linear function on $\mathcal{G}$, hence an element of $\mathcal{G}^*$. Considering two functions f_1 and f_2 the eq.(14.13) shows that the differential $d_e\{f_1,f_2\}$ only depends on d_ef_1 and d_ef_2 and is proportional to $d_e\eta$, since $\eta(e)=0$.

Definition. *The Poisson bracket $\{,\}_G$ defines a Lie algebra structure on $\mathcal{G}^*$ by the following formula:*

$$[d_ef_1,d_ef_2]_{\mathcal{G}^*} \;=\; d_e\{f_1,f_2\}_G \tag{14.16}$$

The Jacobi identity for the Lie algebra structure is direct consequence of the Jacobi identity for the Poisson bracket on G.

Introducing a basis (e^a) in $\mathcal{G}^*$, dual to the basis (e_a) in $\mathcal{G}$, the differential at the identity can written as $d_ef=\sum_a e^a(\nabla_a f)\in\mathcal{G}^*$ where $\nabla_a f=\frac{d}{dt}f(e^{te_a})|_{t=0}$. Similarly

$$d_e\eta=e^c\nabla_c\eta=C_c^{ab}e^c\otimes e_a\otimes e_b$$

With these notations eq.(14.16) reads:

$$\left[e^a,e^b\right]_{\mathcal{G}^*} \;=\; C_c^{ab}\,e^c \tag{14.17}$$

so the structure constants are $C_c^{ab}=(\nabla_c\eta^{ab})$. We shall denote by G^* the connected Lie group with Lie algebra $\mathcal{G}^*$.

Let G be a Poisson–Lie group and let $\mathcal{D}=\mathcal{G}\oplus\mathcal{G}^*$. This is called the classical double. One can introduce a Lie algebra structure on $\mathcal{D}$ which extends the structure on $\mathcal{G}$ and $\mathcal{G}^*$ and such that the elements of $\mathcal{G}$ and $\mathcal{G}^*$ do not commute. In terms of a basis $(e_a)\in\mathcal{G}$ and its dual $(e^a)\in\mathcal{G}^*$, this structure reads:

$$\begin{aligned}
[e_a,e_b] &= f_{ab}^c\,e_c\\
[e^a,e_b] &= f_{bc}^a\,e^c-C_b^{ac}\,e_c\\
[e^a,e^b] &= C_c^{ab}\,e^c
\end{aligned}\tag{14.18}$$

Proposition. *The above brackets define a Lie algebra structure on $\mathcal{D}$.*

Proof. One defines also $[e_b,e^a]=-[e^a,e_b]$ so the antisymmetry is obvious. We need to verify the Jacobi identity. Due to the Jacobi identity on $\mathcal{G}$ and $\mathcal{G}^*$ one has only to verify the two cases $[e^a,[e^b,e_c]]+[e_c,[e^a,e^b]]+[e^b,[e_c,e^a]]=0$ and $[e^a,[e_b,e_c]]+[e_c,[e^a,e_b]]+[e_b,[e_c,e^a]]=0$. These relations reduce to:

$$C_d^{ab}f_{cl}^d-C_l^{bd}f_{dc}^a+C_l^{ad}f_{dc}^b+C_c^{db}f_{ld}^a-C_c^{da}f_{ld}^b=0$$

when using the Jacobi identity on the structure constants C_c^{ab} and f_{bc}^a. This is just the cocycle relation, eq.(14.15). ∎

Remark. With this bracket on the double, one can construct a solution of the Yang-Baxter equation:

$$r_{12} = \sum_a e_a \otimes e^a \quad \in \mathcal{G} \otimes \mathcal{G}^*$$

Indeed the Yang–Baxter equation reads:

$$[r_{12}, r_{13}] + [r_{12}, r_{23}] + [r_{13}, r_{23}] = [e_a, e_b] \otimes e^a \otimes e^b + e_a \otimes [e^a, e_b] \otimes e^b + e_a \otimes e_b \otimes [e^a, e^b] = 0$$

due to definition of the commutators on $\mathcal{D}$.

14.5 Action of a Poisson–Lie group on a symplectic manifold.

Let G be a Poisson–Lie group and M be a symplectic manifold. We assume that G acts on M. In this case the natural compatibility condition between the group action and the Poisson structures on G and M is:

Definition. *The action of a Poisson-Lie group on a symplectic manifold is a Lie–Poisson action if for any $g \in G$ and any functions f_1 and f_2 on M, we have:*

$$\{f_1(g \cdot m), f_2(g \cdot m)\}_{G\times M} = \{f_1, f_2\}_M(g \cdot m) \tag{14.19}$$

Here the Poisson structure on $G \times M$ is the product Poisson structure.

At the infinitesimal level, let $X \in \mathcal{G}$ and denote by $X \cdot m$ the vector field on M corresponding to the infinitesimal transformation generated by X. For any function f on M we define the action of X on f by:

$$(X.f)(m) = \frac{d}{dt} f(e^{-tX} \cdot m)|_{t=0} =< \zeta_f(m), X >$$

where $<,>$ denotes the pairing between $\mathcal{G}$ and $\mathcal{G}^*$. This defines a function $\zeta_f : \; m \to \zeta_f(m) \in \mathcal{G}^*$.

Proposition. *The infinitesimal form of eq.(14.19) for a Poisson–Lie action is:*

$$\{X \cdot f_1, f_2\}_M \; + \; \{f_1, X \cdot f_2\}_M \; + \; < [\zeta_{f_1}, \zeta_{f_2}]_{\mathcal{G}^*}, X > = \; X \cdot \{f_1, f_2\}_M \tag{14.20}$$

or equivalently, $\{\zeta_{f_1}, f_2\}_M \; + \; \{f_1, \zeta_{f_2}\}_M \; + \; [\zeta_{f_1}, \zeta_{f_2}]_{\mathcal{G}^*} \; = \; \zeta_{\{f_1,f_2\}_M}$.

Proof. The definition of the product Poisson bracket is equivalent to:

$$\{f_1(g \cdot m), f_2(g \cdot m)\}_{G\times M} = \{f_1(g \cdot m), f_2(g \cdot m)\}_M + \{f_1(g \cdot m), f_2(g \cdot m)\}_G$$

In the right–hand side of this formula, in the term $\{\ ,\ \}_M$ the functions $f_{1,2}(gm)$ are viewed as functions of m, and g is a parameter, while in the term $\{\ ,\ \}_G$ they are viewed as functions of g and m is a parameter. So the Lie–Poisson condition reads:

$$\{f_1(g \cdot m), f_2(g \cdot m)\}_M + \{f_1(g \cdot m), f_2(g \cdot m)\}_G = \{f_1, f_2\}_M(g \cdot m)$$

Taking $g = \exp(-tX)$ the first Poisson bracket becomes $-\{X \cdot f_1(m), f_2(m)\}_M - \{f_1(m), X \cdot f_2(m)\}_M$ and the right hand–side becomes $-X \cdot \{f_1, f_2\}_M(m)$. The second Poisson bracket is $- < d_e\{f_1(g \cdot m), f_2(g \cdot m)\}_G, X >$ where d_e is the differential of a function on G taken at $g = e$. By definition of the Lie bracket on $\mathcal{G}^*$ this is $- < [d_e f_1, d_e f_2]_{\mathcal{G}^*}, X >$. One gets eq.(14.20) noting that $\zeta_f(m) = -d_e f(g \cdot m)$. ∎

Introducing two dual basis of the Lie algebras $\mathcal{G}$ and $\mathcal{G}^*$, $e_a \in \mathcal{G}$ and $e^a \in \mathcal{G}^*$ with $< e^a, e_b >= \delta^a_b$, eq.(14.20) becomes, using $\zeta_f = \sum_a (e_a \cdot f) e^a$:

$$\{e_a \cdot f_1, f_2\}_M \; + \; \{f_1, e_a \cdot f_2\}_M \; - \; C_a^{bc}(e_b \cdot f_1)(e_c \cdot f_2) \; = \; e_a.\{f_1, f_2\}_M \tag{14.21}$$

It follows immediately from eq.(14.20) that a Lie–Poisson action cannot be symplectic if the algebra $\mathcal{G}^*$ is non Abelian. Hence we cannot expect that infinitesimal group actions are locally generated by Hamiltonians as in the symplectic case. There is however a generalization of this notion, in the Lie–Poisson case, by what is called non-Abelian Hamiltonians.

Proposition. *Assume that a Poisson–Lie group G acts on M by a Lie–Poisson action. Then, there exists a function Γ, locally defined on M and taking values in the group G^*, of Lie algebra $\mathcal{G}^*$, such that for any function f on M,*

$$X.f \;=<\; \Gamma^{-1}\;\{f,\Gamma\}_M, X > \quad , \quad \forall\; X \in \mathcal{G} \tag{14.22}$$

or equivalently $\zeta_f(m) \;=\; \Gamma^{-1}\;\{f,\Gamma\}_M(m)$. We will refer to Γ as the non-Abelian Hamiltonian of the Lie–Poisson action.

Proof. Introduce the Darboux coordinates (q^i, p^i) on the symplectic manifold M. For any $X \in \mathcal{G}$ expand $X \cdot m = X^{p^i}\partial_{p^i} + X^{q^i}\partial_{q^i}$ and introduce the form $\Omega_X = \sum_i (X^{q^i} dp^i - X^{p^i} dq_i)$. Finally we introduce the $\mathcal{G}^*$–valued one–form $\Omega = e^a\;\Omega_{e_a}$. As in the symplectic case, eq.(14.20) is then equivalent to the following zero–curvature condition for Ω:

$$d\Omega \;+ [\Omega,\Omega]_{\mathcal{G}^*} = 0$$

Therefore, locally on M, $\Omega = \Gamma^{-1}\; d\Gamma$ with $\Gamma \in G^*$. This proves eq.(14.22). ■

The converse is true: an action generated by a non–Abelian Hamiltonian as in eq.(14.22) is Lie–Poisson since we have:

$$X\cdot\{f_1,f_2\}_M - \{X\cdot f_1, f_2\}_M - \{f_1, X\cdot f_2\}_M =< \left[\Gamma^{-1}\{f_1,\Gamma\}_M\;,\;\Gamma^{-1}\{f_2,\Gamma\}_M\right]_{\mathcal{G}^*}, X >$$

In the Abelian case we have $\Gamma(m) = \exp\left(-\mathcal{P}(m)\right)$ where $\mathcal{P}$ is the momentum taking values in the Abelian Lie algebra $\mathcal{G}^*$. This is because, in the symplectic case, eq.(14.22) becomes $X \cdot f =< \{\mathcal{P}, f\}, X >= \{H_X, f\}$. Hence Γ is the non–abelian generalization of the moment map.

14.6 The groups G and G^*.

As a preparation to our study of dressing transformations, we apply the previous results to a more specific situation.

Let $\mathcal{G}$ be a Lie algebra with a bilinear invariant form denoted by Tr, with associated connected Lie group G. We choose an orthonormal basis $t^a \in \mathcal{G}$, $\mathrm{Tr}\,(t^a t^b) = \delta^{ab}$. In this basis the Lie bracket reads $\left[t^a, t^b\right] = f^{ab}_c t^c$ where f^{ab}_c are the structure constants. We denote by $\mathcal{C} = \sum_a t^a \otimes t^a$ the tensor Casimir in $\mathcal{G}\otimes\mathcal{G}$.

We equip G with a Lie–Poisson structure by choosing a cocycle $\eta(g)$. A simple way to fulfill the cocycle condition is to take for $\eta(g)$ a coboundary

$$\eta(g) = r - \mathrm{Ad}_g r$$

where r is a constant element in $\mathcal{G}\otimes\mathcal{G}$. Then eq.(14.14) becomes:

$$\{g_1, g_2\}_G = [r, g_1 g_2] \tag{14.23}$$

This is the Sklyanin bracket. In this case the Lie–Poisson property is easy to check. One has to show that:

$$\{(gh)_1, (gh)_2\}_{G\times G} = [r, (gh)_1 (gh)_2]$$

This is obvious when we notice that, in the product Poisson structure on $G \times G$, g and h Poisson commute. The computation then reduces to the computation we have done for the Sklyanin approach to the closed Toda chain, see Chapter [5]. Indeed, since in $G \times G$, g and h Poisson commute we have:

$$\begin{aligned} \{(gh)_1, (gh)_2\}_{G\times G} &= \{g_1, g_2\}_G h_1 h_2 + g_1 g_2 \{h_1, h_2\}_G \\ &= [r, g_1 g_2] h_1 h_2 + g_1 g_2 [r, h_1, h_2] = [r, g_1 h_1 g_2 h_2] \end{aligned}$$

Note that r is defined only up to the addition of a multiple of the Casimir element which drops out of $\eta(g)$. So we define $r_{12}^{\pm} = r \pm \frac{1}{2}\mathcal{C}$. Antisymmetry of the Poisson bracket is ensured by choosing r antisymmetric, so that $r_{12}^{\pm} = -r_{21}^{\mp}$. The Jacobi identity for the Poisson bracket is ensured by requiring that $r_{12}^{\pm}$ are solutions of the classical Yang–Baxter equation:

$$\left[r_{12}^{\pm}, r_{13}^{\pm}\right] + \left[r_{12}^{\pm}, r_{23}^{\pm}\right] + \left[r_{13}^{\pm}, r_{23}^{\pm}\right] = 0 \tag{14.24}$$

Eq.(14.24) is an equation in $\mathcal{G} \otimes \mathcal{G} \otimes \mathcal{G}$, and the indices on $r^{\pm}$ refer to the copies of $\mathcal{G}$ on which $r^{\pm}$ is acting. Using the bilinear form Tr to identify the vector spaces $\mathcal{G}^*$ and $\mathcal{G}$, the elements $r^{\pm}$ of $\mathcal{G} \otimes \mathcal{G}$ can be mapped into elements $R^{\pm} \in \mathcal{G} \otimes \mathcal{G}^* \cong \mathrm{End}\mathcal{G}$ defined by:

$$R^{\pm}(X) = \mathrm{Tr}_2\left(r_{12}^{\pm}(1 \otimes X)\right) \qquad ; \qquad \forall\, X \in \mathcal{G} \tag{14.25}$$

Note that we have $R^+ - R^- = Id$.

The Poisson bracket (14.23) on G induces a Lie algebra structure on $\mathcal{G}^*$ by eq.(14.16). Identifying the vector spaces $\mathcal{G}$ and $\mathcal{G}^*$ by Tr, the bracket on $\mathcal{G}^*$ is mapped to the R–bracket:

$$[\,X, Y\,]_R = \left[R^{\pm}(X), Y\right] + \left[X, R^{\mp}(Y)\right] = [R(X), Y] + [X, R(Y)] \tag{14.26}$$

with $R = \frac{1}{2}(R^+ + R^-)$. We gave a detailed analysis of this bracket in Chapter [3], and all the results apply here. We simply recall that because $R^+ - R^- = Id$, any $X \in \mathcal{G}$ admits a decomposition as:

$$X = X_+ - X_- \qquad \text{with} \qquad X_{\pm} = R^{\pm}(X) = \mathrm{Tr}_2\left(r_{12}^{\pm}(1 \otimes X)\right) \tag{14.27}$$

In terms of the components X_+ and X_-, the commutator in $\mathcal{G}^*$ becomes:

$$[\,X, Y\,]_R = [X_+, Y_+] - [X_-, Y_-]$$

In particular, the plus and minus components commute in $\mathcal{G}^*$. Moreover $R^{\pm}$ are Lie algebra homomorphisms so that $X_{\pm}$ live in two subalgebras $\mathcal{G}_{\pm}$ of $\mathcal{G}$. Recall also that the image of $\mathcal{G}^*$ in $\mathcal{G}_- \oplus \mathcal{G}_+$ is the set of $X = X_+ - X_-$ such that $\theta(X_+) = X_-$, see Chapter [3].

By exponentiation, the subalgebras $\mathcal{G}_{\pm}$ correspond to connected Lie subgroups $G_{\pm}$ of G, and the group G^* can be viewed as the set of couples (g_-, g_+) subjected to some condition $\theta(g_+) = g_-$ with product law:

$$(g_-, g_+) \cdot (h_-, h_+) = (g_- h_-, g_+ h_+) \tag{14.28}$$

Any element $g \in G$ (in a neighborhood of the identity) admits a unique factorization as:

$$g = g_-^{-1} g_+ \tag{14.29}$$

with $\theta(g_+) = g_-$, so that this associates to g a unique element of $G^* = (g_-, g_+)$ through a factorization problem. So as sets, G and G^* are identified, but they have different group structures.

The group G^* itself becomes a Poisson-Lie group if we introduce on it the Semenov–Tian–Shansky Poisson bracket:

$$\begin{aligned}
\{g_{+1}, g_{+2}\}_{G^*} &= -\left[r_{12}^{\pm}, g_{+1}\, g_{+2}\right] \\
\{g_{-1}, g_{-2}\}_{G^*} &= -\left[r_{12}^{\mp}, g_{-1}\, g_{-2}\right] \\
\{g_{-1}, g_{+2}\}_{G^*} &= -\left[r_{12}^{-}, g_{-1}\, g_{+2}\right] \\
\{g_{+1}, g_{-2}\}_{G^*} &= -\left[r_{12}^{+}, g_{+1}\, g_{-2}\right]
\end{aligned} \tag{14.30}$$

or, for the factorized element $g = g_-^{-1} g_+$:

$$\{g_1, g_2\}_{G^*} = -g_1\, r_{12}^{+}\, g_2 - g_2\, r_{12}^{-}\, g_1 + g_1\, g_2\, r_{12}^{\pm} + r_{12}^{\mp}\, g_1\, g_2 \tag{14.31}$$

The multiplication in G^* is a Poisson map for the brackets (14.30). The group G^* is therefore a Poisson–Lie group.

14.7 The group of dressing transformations.

We use the above results to understand the Poisson structure of the dressing transformations introduced in Chapter [2].

Definition. *Let G be a Poisson–Lie group associated to an r–matrix, and G^* its dual Poisson–Lie group. We define an action of G^* on G which we call a dressing transformation. We identify G^* to G as sets, via the factorization problem $g = g_-^{-1} g_+$. The dressing of $x \in G$ by $g \in G^*$ is defined by:*

$$(g = g_-^{-1} g_+ \in G^*,\ x \in G) :\to {}^g x = (xgx^{-1})_\pm\, x\, g_\pm^{-1} \ \in G \tag{14.32}$$

In this equation $(xgx^{-1})_\pm$ refer to the factorization $(xgx^{-1})_-^{-1}(xgx^{-1})_+ = (xgx^{-1})$ and this implies that the two signs give the same result for ${}^g x$.

Proposition. *The action $x \to {}^g x$ is a group action of G^* on G.*

<u>Proof</u>. We have to show that ${}^g({}^h x) = {}^{gh} x$, that is:

$$({}^h x\ g\ {}^h x^{-1})_\pm\ {}^h x\, g_\pm^{-1} = (x(gh)x^{-1})_\pm x(gh)_\pm^{-1} \tag{14.33}$$

Introducing the notation, for any $h \in G^*$: $\Theta_\pm^h = (xhx^{-1})_\pm$, so that ${}^h x = \Theta_\pm^h x h_\pm^{-1}$, and using the sign freedom to write the first ${}^h x$ in $({}^h x\ g\ {}^h x^{-1})_\pm$ with the minus sign, and the second one with the plus sign, the left–hand side of eq.(14.33) reads:

$$\left(\Theta_-^h\ x\ h_-^{-1} g h_+\ x^{-1}\, \Theta_+^{h\,-1}\right)_\pm \Theta_\pm^h\ x\ h_\pm^{-1} g_\pm^{-1}$$

Remark that since $g = g_-^{-1} g_+$, and due to the definition of the group law in G^*, $h_-^{-1} g h_+ = (gh)$. Moreover since by definition $x(gh)x^{-1} = \Theta_-^{(gh)\,-1}\, \Theta_+^{(gh)}$ we get:

$$\Theta_-^h\ x\ h_-^{-1} g h_+\ x^{-1}\, \Theta_+^{h\,-1} = \Theta_-^h\ \Theta_-^{(gh)\,-1}\, \Theta_+^{(gh)}\, \Theta_+^{h\,-1}$$

so that one reads the factorization:

$$\left(\Theta_-^h\ x\ h_-^{-1} g h_+\ x^{-1}\, \Theta_+^{h\,-1}\right)_\pm = \Theta_\pm^{(gh)} \Theta_\pm^{h\,-1}$$

From this the result follows immediately. ∎

The infinitesimal form of eq.(14.32) is for any $X \in \mathcal{G}$, with $X = X_+ - X_-$:

$$\delta_X \; x \;=\; Y_\pm \; x - x \; X_\pm \qquad \text{with} \qquad Y_\pm = (xXx^{-1})_\pm \tag{14.34}$$

One of the main property of this action is that it is a Lie–Poisson action of G^* on G if the groups G and G^* are equipped with the Poisson structures defined in eqs.(14.23) and (14.30), i.e. we have:

$$\{{}^g x_1, {}^g x_2\}_{G^* \times G} \;=\; \left[r^\pm_{12} \; , \; {}^g x_1 \, {}^g x_2\right]$$

We will prove this fact by exhibiting a non–Abelian Hamiltonian for dressing transformations, so that they are automatically Lie–Poisson.

Proposition. *The non-Abelian Hamiltonian of the dressing transformations (14.32), which is an element of $G^{**} \cong G$, is the identity function of G, i.e. it is x itself:*

$$X \cdot x = -\delta_X \, x \;=\; \mathrm{Tr}_2 \left(x_2^{-1} \; \{x_2, x_1\}_G X_2 \right) \quad \forall X \in \mathcal{G}, x \in G \tag{14.35}$$

Proof. First note that, as usual, the action on functions is defined with the inverse group element, so that the action of X on the function x is $-\delta_X x = -(xXx^{-1})_\pm x + xX_\pm$. From eq.(14.23), in which we take $\Gamma(x) = x$ we have:

$$\begin{aligned}
\mathrm{Tr}_2 \left(x_2^{-1} \; \{x_2, x_1\}_G X_2 \right) &= -\mathrm{Tr}_2 \left(x_2^{-1} [r^\pm_{12}, x_1 x_2] X_2 \right) \\
&= -\mathrm{Tr}_2 \left(r^\pm_{12} x_1 \, x_2 X_2 x_2^{-1} \right) + \mathrm{Tr}_2 \left(x_1 r^\pm_{12} X_2 \right) \\
&= -(xXx^{-1})_\pm \, x + x \, X_\pm \;=\; -\delta_X \, x
\end{aligned}$$

∎

It is a remarkable fact that there exists a non-Abelian Hamiltonian since the group G with the Sklyanin bracket is not a symplectic manifold, since the bracket is degenerate.

References

[1] V. Arnold. "**Méthodes mathématiques de la Mécanique classique.**" MIR 1976. Moscou.

[2] R. Abraham, J. Marsden. "**Foundations of Mechanics.**" Benjamin (1978) Reading, Massachusetts

[3] M. Semenov–Tian–Shansky. "*Dressing transformations and Poisson Group actions.*" Publ. RIMS 21 (1985) p. 1237.

[4] V. Drinfeld. "*Hamiltonian structures on Lie groups, Lie bialgebras and the geometric meaning of classical Yang–Baxter equations.*" Soviet Math. Dokl. 27 (1983), no. 1, pp.68–71.

Chapter 15

Riemann Surfaces

15.1 Smooth algebraic curves.

Riemann surfaces are compact smooth analytic varieties of dimension 1. This means that around each point p there is a neighbourhood and a local parameter $z(P)$ mapping it homeomorphically to an open ball $|z| < 1$ of the complex numbers. Moreover in the intersection of two such neighbourhoods the corresponding local parameters $z_1(P)$ and $z_2(P)$ must be related by an analytic bijection. Hence locally a smooth curve looks like the complex line. Finally a Riemann surface is compact hence it is a closed surface without boundary.

For our purposes it is very important to look at Riemann surfaces from an algebraic viewpoint, that is as the locus in $\mathbb{C}^2$ of an algebraic equation $P(x, y) = 0$, where P is a polynomial in the complex variables x and y. At a generic point of this locus one has $\partial_x P \neq 0$ and $\partial_y P \neq 0$ hence x as well as y can be taken as analytic local parameters in the vicinity of this point. If P has degree N in y we can obviously find N analytic solutions $y = f_j(x)$ for $j = 1, \cdots N$ defined in some open set of the x variable. These are the N branches of the curve presented as a N–fold covering of the complex line of the x variable. This situation can be analytically continued until one gets to a point where some branches meet, that is the equation in y, $P(x, y) = 0$, has a multiple root, so that $\partial_y P = 0$. Generically one still has $\partial_x P \neq 0$ at such a point and one can choose y instead of x as local parameter in the vicinity of this point which is a perfectly smooth point on the curve. The covering projection $(x, y) \to x$ however is branched at this point, that is several branches coalesce to one point. We say that $P(x, y) = 0$ expresses the curve as an N–fold branched covering of the complex line. Branch points occur when the discriminant of $P(x, y)$ viewed as a polynomial in y vanishes. This is a polynomial in x, hence has a finite number of roots above which the covering projection branches.

A more complicated situation arises if at some point both partial derivatives of P vanish. This means that locally the curve looks like the intersection of several lines hence is not smooth. In order to make contact with our general definition of a Riemann surface one has to perform an operation called desingularization, which basically consists in replacing the singular point by several ordinary points while leaving the analytic structure of the neighbourhood untouched. An easy way to understand how this can be achieved is to consider birational transformations. They are mappings of the complex plane to itself $(x, y) \to (x' = x'(x, y), y' = y'(x, y))$ where x' and y' are rational functions of x and y such that one can also express x and y as rational functions of x' and y'. At a non singular point of such a transformation it is clearly bijective and preserves the analytic structure. A simple example is the quadratic transformation $x' = 1/x$ and $y' = 1/y$ which is obviously bijective and analytic on its domain of definition. However when transforming the equation

$P(x,y)=0$ under such a mapping chosen to have its singular set precisely at a singular point of the curve one may blow the singular point into several ordinary points of the transformed curve.

Let us see how this works on a simple example. Consider the curve $y^2 = x^2 + x^3$, which is singular at $(0,0)$, and the birational transformation $x = x'$, $y = x'y'$ which can be inverted by $y' = y/x$. It is obviously bijective except at the singular point $(0,0)$ where y' is indeterminate. Substituting into the equation of the curve one gets ${y'}^2 = 1 + x'$ and we see that the two ordinary points $x' = 0$ and $y' = \pm 1$ project to the one singular point $x = y = 0$, while any other point of the transformed curve bijectively corresponds to just one point of the initial curve. This bijection preserves the analytic structure of both curves. A similar construction can be done around any singular point of an algebraic curve and then patched with the analytic structure around ordinary points to get a smooth Riemann surface associated to the equation $P = 0$. We say that it is the desingularized curve of the equation $P = 0$.

Finally to get a compact surface one has to consider points at infinity. We transform the equation $P = 0$ under the quadratic transformation $x = 1/x', \quad y = 1/y'$ so that a chart around ∞ becomes an analytic chart around 0. In general one gets a singular point at $(0,0)$ which has to be desingularized by the above method. We shall give a simple example of this procedure in the next section in which we study a frequent type of Riemann surface occuring in integrable models, the hyperelliptic curves.

15.2 Hyperelliptic curves.

Such a curve is given as the locus of the equation:

$$y^2 = P(x) \tag{15.1}$$

where P is a polynomial in x of degree N, written in the form $P = \prod_i^N (x - a_i)$. In order not to have singular points we shall assume $a_i \neq a_j$ for $i \neq j$. There is an analytic involution of the curve into itself $(x,y) \to (x,-y)$ which is called the hyperelliptic involution and the existence of such an automorphism in fact characterizes hyperelliptic curves. Of course the curve can be expressed as a two sheeted covering of the complex line branched above the points $x = a_i$. Around such a point $(x = a_i, y = 0)$ one can take y as a perfectly smooth local parameter. For example if $P(x) = xP_1(x)$ with $P_1(0) \neq 0$ one can express $x = y^2 + O(y^3)$ around $(0,0)$. Note that the situation would be entirely similar for an equation of the type $y^n = xP_1(x)$ around $(0,0)$ except that now n branches coalesce to the origin. The only tricky point is to consider the situation at infinity. Performing the transformation $x = 1/x', \quad y = 1/y'$ one gets

$$ {y'}^2 \prod_{i=1}^N (1 - x'a_i) - {x'}^N = 0 $$

and we see that the origin is *singular* for $N > 1$. It is now necessary to distinguish two cases, N odd or even.

When $N = 2g+1$, for some integer g, one can perform the birational transformation $x' = x''$ and $y' = ({x''}^g / \prod(1 - x''a_i))y''$ and the transformed equation reads ${y''}^2 - x'' \prod(1 - x''a_i) = 0$. We see that the singular point $(x' = 0, y' = 0)$ gives rise to a single point $(x'' = 0, y'' = 0)$ on the desingularised curve hence we say that there is just one point at infinity, but it is a branch point of the covering $(x,y) \to x$. The local parameter around ∞ is y'' and we can express locally:

$$ x = \frac{1}{{y''}^2}(1 + O(y'')), \quad y = \frac{1}{{y''}^{2g+1}}(1 + O(y'')) $$

Remark that x and y are meromorphic functions on the Riemann surface with poles only at ∞ where for example x has a pole of order 2. Finally one frequently says that $\sqrt{x}$ is a local parameter around ∞ since this is equivalent to $1/y''$.

When $N = 2g+2$ one performs the birational transformation $x' = x''$ and $y' = (x''^{(g+1)}/\prod(1-x''a_i))y''$ which yields $y''^2 = \prod(1-x''a_i)$. Hence we are in the generic situation, i.e. there are two points corresponding to the singular point $(x' = 0, y' = 0)$, namely $\infty_+ = (x'' = 0, y'' = +1)$ and $\infty_- = (x'' = 0, y'' = -1)$. In other words all the points of the desingularised curve bijectively and analytically correspond to points of the original curve except the two points $\infty_\pm$ which map to the same $\infty = (x' = 0, y' = 0)$. One also says that the singular curve is obtained from the non singular one by identifying two points. Around $\infty_\pm$ either x'' or y'' are good local parameters and one gets for example

$$x = \frac{1}{x''}, \quad y = \frac{1}{x''^{(g+1)}}(1 + O(x''))$$

We see that the meromorphic function x has two simple poles on the curve one at ∞_+ and the other at its hyperelliptic conjugate ∞_-. The curve can be seen as a two–sheeted branched covering of the Riemann sphere branched over the $2g+2$ points $(a_k, 0)$.

Let us remark that in both cases, $N = 2g+1$ and $N = 2g+2$, eq.(15.1) allows to present the hyperelliptic curve as a two–sheeted branched covering of the Riemann sphere with $2g+2$ branch points. In fact for $N = 2g+1$ we have the $2g+1$ points $(a_k, 0)$ plus the single point at infinity ∞, while for $N = 2g+2$ we have two points at infinity but they are not branching points of the covering. This will allow us to conclude that in both cases the underlying topological surface is of genus g.

15.3 The Riemann–Hurwitz formula.

Let us recall that for a triangulated surface where the triangulation has F faces, V vertices and A edges one defines the Euler–Poincaré characteristic

$$\chi = F - A + V = 2 - 2g$$

which is a topological invariant. Here g is called the genus of the Riemann surface and any surface of genus g is homeomorphic to a sphere with g handles.

Let us now assume that a Riemann surface is presented as an N–sheeted branched covering of some base space of Euler characteristic χ_0. Choose a *triangulation* of the base space such that all branch points are included as vertices. Now for each triangle consider its N preimages under the covering projection, and similarly the N preimages of each edge. However while ordinary vertices have N preimages each base point corresponding to a branch point has less than N preimages and the reduction is given by the order of the branch point, i.e. the number of branches which coalesce at the branch point minus 1. This is also called the index of the branch point. Altogether one gets a triangulation of the Riemann surface with NF triangles, NA edges and $NV - B$ vertices, where B is the total number of branch points counted with their index on the Riemann surface. Hence the Euler characteristic χ of the Riemann surface is related to χ_0 by $\chi = N\chi_0 - B$. This is the Riemann-Hurwitz formula:

$$2g - 2 = N(2g_0 - 2) + B \tag{15.2}$$

Let us apply this formula to the computation of the genus of an hyperelliptic curve. The base is the Riemann sphere for which there is a triangulation with 8 faces, 12 edges and 6 vertices, so that $\chi_0 = 2$ or $g_0 = 0$. We have seen that the covering has $B = 2g+2$ branch points of index 1 while $N = 2$, so the genus computed by the Riemann–Hurwitz formula is precisely the number g introduced above.

15.4 The field of meromorphic functions of a Riemann surface.

Consider a Riemann surface of genus g. A complex valued function on this surface is analytic around a point if its expression in terms of a local parameter is analytic, similarly it has a pole of order n if its expression in terms of a local parameter has a pole of order n, and so on. These definitions are invariant under analytic reparametrizations. Of course it is impossible to get an everywhere non constant analytic function on a compact Riemann surface, this is basically the Liouville theorem. We are generally interested in meromorphic functions which have a finite number of poles. Obviously the meromorphic functions form a field which is called the function field of the surface. When the curve is given by an equation $P(x, y) = 0$, any rational function of x and y is a meromorphic function on the curve with poles located at arbitrary points and one can show that the most general meromorphic function can be so constructed, that is the field of meromorphic functions is just the field of rational functions of x and y modulo the equation of the curve. This allows to give a completely algebraic description of Riemann surfaces.

Conversely, let f be a non trivial meromorphic function on a Riemann surface. The Cauchy theorem still holds true on a Riemann surface, hence integrating over a small circle whose interior is free of poles and zeroes of $f - a$ one gets:

$$0 = \frac{1}{2\pi i}\oint \frac{df}{f-a} = \#\,\text{zeroes of }\ (f-a) - \#\,\text{poles of }\ f$$

which shows that f takes each value the same number of times, say N times. This allows to present a general Riemann surface as an N–sheeted branched covering of the Riemann sphere by $p \to z = f(p)$. Now take any other meromorphic function g, and for any z on the Riemann sphere consider its N preimages p_j under f, and then the elementary symmetric functions of $g(p_j)$. They do not depend on the order of the sheets hence define meromorphic functions of z on the Riemann sphere, that is *rational* functions of z. It is then clear that g obeys a polynomial equation $P(f, g) = 0$. By similar arguments one can show that one can choose g so that the polynomial P is irreducible and then any other meromorphic function is a rational function of f and g. We see that any abstract Riemann surface may be viewed as a smooth compact algebraic curve.

Let us consider the example of an hyperelliptic curve $y^2 = P_{2g+1}(x)$ where P_{2g+1} is a polynomial of degree $2g+1$. Any meromorphic function can be written as $f(x, y) = (A(x)+yB(x))/C(x)$ with A, B, C polynomials in x since one can eliminate y^2 using the equation of the curve and eliminate y in the denominator. In this case we get a simple construction of a meromorphic function with $g + 1$ poles at $(x_1, y_1), \cdots, (x_{g+1}, y_{g+1})$. Indeed take

$$f(x, y) = \frac{y + Q(x)}{\prod_i (x - x_i)}$$

where Q is a polynomial of degree g determined by requiring that $Q(x_i) = y_i$ so that the numerator of f vanishes at the point $(x_i, -y_i)$. Note that f vanishes at ∞ and has g zeroes at finite distance. Also note that the meromorphic function x has a double pole at ∞ hence takes each value two times. The existence of such a function characterizes hyperelliptic curves.

15.5 Line bundles on a Riemann surface.

It is extremely convenient to generalize the notion of function on a Riemann surface by considering line bundles on the surface, and sections of these bundles. Moreover such line bundles occur naturally in the study of integrable systems. Let us consider a covering $\{U_\alpha\}$ of the Riemann surface by open sets U_α and assume that for each non void intersection $U_\alpha \cap U_\beta$ some continuous functions (called transition functions) $t_{\alpha\beta}$ are given which are neither vanishing nor ∞ on the intersection. Moreover we assume that on each non void triple intersection $U_\alpha \cap U_\beta \cap U_\gamma$ one has $t_{\alpha\beta}t_{\beta\gamma}t_{\gamma\alpha} = 1$ and that $t_{\alpha\beta}t_{\beta\alpha} = 1$. This defines a line bundle ξ. When the transition functions are differentiable it is a differentiable bundle, while if the transition functions are analytic it is an analytic bundle. We shall be concerned with analytic bundles on a Riemann surface. A section of the bundle ξ is a collection of functions f_α on each U_α such that on each intersection we have $f_\alpha = t_{\alpha\beta} f_\beta$. If all functions f_α are holomorphic this is a holomorphic section of ξ and the space of such analytic sections will be denoted $\Gamma(\xi)$. If all functions f_α are meromorphic it is called a meromorphic section. If f_α has a pole or a zero of order m at some point of $U_\alpha \cap U_\beta$, since $t_{\alpha\beta}$ is analytic without zero, it is the same for f_β and we say that the section has the zero or pole of order m.

Geometrically an analytic line bundle can be seen as a space locally analytically isomorphic to the cartesian product of U_α by the complex numbers seen as a vector space. Any point above U_α can be written (p, ϕ_α) with $p \in U_\alpha$ and ϕ_α a number. If moreover $p \in U_\beta$, the same point can be written (p, ϕ_β). The two descriptions patch if one can write $\phi_\alpha = t_{\alpha\beta}(p)\phi_\beta$ for some analytic non vanishing function $t_{\alpha\beta}$ defined on $U_\alpha \cap U_\beta$. The triple intersection condition is obviously obeyed. Moreover for any non vanishing analytic functions f_α defined on U_α an equivalent description of the line bundle is obtained by sending the point (p, ϕ_α) of the line bundle to $(p, f_\alpha(p)\phi_\alpha)$. In the new description the transition functions are now $t_{\alpha\beta}(p) f_\alpha(p)/f_\beta(p)$ so that transition functions differing by multiplication by a quotient f_α/f_β define the *same* line bundle ξ. A local or global section of the line bundle ξ can be viewed intrinsically as a map which associates to each point of the Riemann surface a point in the fiber above it. In a local trivialization $U_\alpha \times \mathbb{C}$ this point can be written $(p, f_\alpha(p))$. In the intersection $U_\alpha \cap U_\beta$ the two descriptions are related by $f_\alpha = t_{\alpha\beta} f_\beta$. We recover the above definition of sections. One can differentiate *differentiable* sections of ξ with respect to $\overline{z}$ since $\partial_{\overline{z}} f_\alpha = t_{\alpha\beta} \partial_{\overline{z}} f_\beta$ hence $\{\partial_{\overline{z}} f_\alpha\}$ is a section.

A very important example is provided by the *canonical bundle*. Consider a covering U_α such that each open set U_α is analytically isomorphic to a domain of the complex numbers by a coordinate $p \to z_\alpha(p)$. In each non trivial intersection $U_\alpha \cap U_\beta$ the coordinate z_β can be expressed as an analytic function of the coordinate z_α, and conversely. One can define:

$$\kappa_{\alpha\beta}(p) = \frac{dz_\beta}{dz_\alpha}(p)$$

which is holomorphic non vanishing (since $dz_\beta/dz_\alpha \cdot dz_\alpha/dz_\beta = 1$) in the intersection. The canonical bundle is defined by these transition functions. This definition is canonical since under change of local parameters $z_\alpha \to w_\alpha$ the transition functions transform according to $\kappa_{\alpha\beta} \to w'_\beta/w'_\alpha \cdot \kappa_{\alpha\beta}$ and the w'_α are analytic non vanishing on U_α. They therefore define the same bundle.

A global section of this bundle is a collection of analytic functions f_α such that $f_\alpha = \kappa_{\alpha\beta} f_\beta$ on intersections, that is $f_\alpha dz_\alpha = f_\beta dz_\beta$. It can be viewed as a globally defined holomorphic differential form of type $(1, 0)$ i.e. a form involving dz.

Another important example is provided by the so–called point bundles. Let the covering be as above and assume that p is a point on the Riemann surface belonging to some U_α and to no other U_β. One can assume that $z_\alpha(p) = 0$. Define the transition functions $t_{\alpha\beta} = z_\alpha$ on non trivial $U_\alpha \cap U_\beta$ and choose $t_{\beta\gamma} = 1$ on all other non trivial intersections. This defines a line bundle ξ_p. Note that our point bundle admits at least an analytic section, namely define $\sigma_\alpha(z_\alpha) = z_\alpha$ and $\sigma_\beta = 1$. The transition conditions $\sigma_\beta = t_{\beta\gamma}\sigma_\gamma$ are obeyed for any β, γ including α. This section has just one zero at p. We see that an analytic line bundle may have non trivial holomorphic sections while holomorphic functions on a compact Riemann surface are constant. Also note that a bundle ξ is the trivial bundle if and only if it admits an analytic non vanishing section since in this case one can write $t_{\alpha\beta} = f_\alpha/f_\beta$ for the given section f_α.

One can introduce a group structure on line bundles. Given two line bundles ξ and σ one can assume that they are defined on a common suitably refined covering U_α by transition functions $t_{\alpha\beta}$ and $s_{\alpha\beta}$. Then the product $\xi\sigma$ is defined by the transition functions $t_{\alpha\beta}s_{\alpha\beta}$ and the inverse ξ^{-1} is defined by $1/t_{\alpha\beta}$ while the neutral element is just the trivial bundle i.e. the cartesian product of the Riemann surface by the complex numbers which can be defined by $t_{\alpha\beta} = 1$. Note that the above definitions are coherent with the redefinitions $t_{\alpha\beta} = f_\alpha/f_\beta \cdot t_{\alpha\beta}$. This group law is commutative hence an *additive* notation is frequently used.

If $\{f_\alpha\}$ is a section of ξ and $\{g_\alpha\}$ a section of σ then $\{f_\alpha g_\alpha\}$ is a section of $\xi\sigma$ and $\{1/f_\alpha\}$ is a meromorphic section of ξ^{-1}. Moreover the quotient of two sections is a meromorphic *function* since the transition functions cancel, and the product of a section by a function is a section.

15.6 Divisors.

One can build more complicated bundles from point bundles at points $p_1, \cdots, p_k$ by taking their product, namely $\sum_j n_j \xi_{p_j}$. Here n_j are positive or negative integers and a formal sum of points p_j on the Riemann surface with multiplicities n_j is called a *divisor* denoted $\mathcal{D} = \sum_j n_j p_j$. The line bundle $\xi = \sum_j n_j \xi_{p_j}$ is the line bundle *associated* to the divisor $\mathcal{D}(\xi) = \sum_j n_j p_j$.

For any meromorphic function f on the Riemann surface with zeroes at p_j of order n_j and poles at q_k of order m_k one defines the divisor of the function f as the formal sum

$$\mathcal{D}(f) = \sum_j n_j p_j - \sum_k m_k q_k$$

The divisor of a section of a line bundle is similarly defined. A divisor is positive if all n_j are positive, hence a meromorphic section is analytic if and only if its divisor is positive.

Since ξ_p has a section with divisor p, the line bundle ξ associated to $\mathcal{D}$ has a section f_ξ with divisor $\mathcal{D}$. When the divisor $\mathcal{D}$ is the divisor of a meromorphic function f, the associated bundle ξ has a section f_ξ with divisor $\mathcal{D}$, hence a section $f^{-1} f_\xi$ which is analytic non vanishing so that ξ is the trivial bundle. Conversely if ξ associated to $\mathcal{D}$ is trivial, the section f_ξ gives rise to a meromorphic function of divisor $\mathcal{D}$ since we can dividef_ξ by an analytic non vanishing section. This allows to introduce an equivalence relation between divisors: two divisors $\mathcal{D}_1$ and $\mathcal{D}_2$ are equivalent if their associated bundles ξ_1 and ξ_2 are such that $\xi_1 - \xi_2$ is the trivial bundle, i.e. if $\mathcal{D}_1 - \mathcal{D}_2$ is the divisor of a meromorphic function.

To a divisor $\mathcal{D} = \sum_j n_j p_j$ is associated a number

$$\deg(\mathcal{D}) = \sum_j n_j$$

called the degree of the divisor. Two equivalent divisors have the same degree since a meromorphic function has the same number of zeroes and poles. Note that if the line bundle ξ has a meromorphic section σ of divisor $\mathcal{D}$ and η is the line bundle associated to $\mathcal{D}$ with meromorphic section τ with the same divisor $\mathcal{D}$, $\xi - \eta$ has an holomorphic non vanishing section σ/τ hence $\xi = \eta$.

15.7 Chern class.

It is well known from differential geometry that one can associate to differential vector bundles an integer called the Chern class which can be computed as the integral of some curvature form. For line bundles it can be shown that this integer classifies *differential* bundles. Moreover the Chern class is compatible with the group structure hence the Chern class of the "product" $\xi + \sigma$ is just the sum $c(\xi) + c(\sigma)$ of the Chern classes of ξ and σ. Simple proofs of these facts, adapted from the differential geometric case can be found in the references. This leads to an index theorem stating that for any meromorphic section f of a line bundle ξ one has:

$$\deg(\mathcal{D}(f)) = c(\xi)$$

As a matter of fact the trivial bundle has Chern class 0, hence this theorem boils down in this case to the above mentioned fact that a meromorphic function has the same number of zeroes and poles. This implies in particular that only bundles with positive Chern class may have holomorphic sections. Note that the point bundle ξ_p has Chern class 1 since it has a holomorphic section with just one zero at p.

15.8 Serre duality.

Having introduced these natural definitions we can now study the spaces of most interest to us, namely the space $\Gamma(\xi)$ of holomorphic sections of ξ. In order to do that it turns out to be very useful to introduce new spaces called $H^1(M,\mathcal{O}(\xi))$ associated to the Riemann surface M and the line bundle ξ. Their definition is very analogous to that of line bundles except that the multiplicative structure is replaced by an additive one. Namely consider a covering U_α (and we assume that each U_α is connected and simply connected) and for each non void intersection $U_\alpha\cap U_\beta$ an holomorphic local section $f_{\alpha\beta}$ of ξ (note that such sections have an intrinsic definition). Such sections are assumed to obey $f_{\alpha\beta}+f_{\beta\gamma}+f_{\gamma\alpha}=0$ on each non void triple intersection $U_\alpha\cap U_\beta\cap U_\gamma$. The space $H^1(M,\mathcal{O}(\xi))$ is the space of sets $\{f_{\alpha\beta}\}$ modulo trivial ones, i.e. modulo those of the form $f_{\alpha\beta}=f_\alpha-f_\beta$ for holomorphic sections f_α of ξ over U_α. Note that there is no non vanishing condition here. If trivialisations of the bundle are defined over the U_α one can represent these sections by complex valued functions related by transition functions. Let f_α and $f_{\alpha\beta}$ be represented over U_α by functions f^α_α and $f^\alpha_{\alpha\beta}$ while $f_{\alpha\beta}$ and f_β are represented on U_β by $f^\beta_{\alpha\beta}$ and f^β_β respectively. Then the equation expressing the triviality of $f_{\alpha\beta}$ reads on $U_\alpha\cap U_\beta$: $f^\alpha_{\alpha\beta}=f^\alpha_\alpha-t_{\alpha\beta}f^\beta_\beta$ or a similar expression for the β trivialization. Note that one can define a $\overline{\partial}$ operator on sections of ξ since $\overline{\partial}$ vanishes on transition functions.

Let us give an alternate description of $H^1(M,\mathcal{O}(\xi))$. One can always solve $f_{\alpha\beta}=f_\alpha-f_\beta$ with f_α *differentiable* sections of ξ on U_α by partition of unity arguments[1]. Consider the type $(0,1)$ forms (i.e. forms in $d\overline{z}$)

$$\overline{\partial}f_\alpha=\frac{\partial f_\alpha}{\partial\overline{z}}d\overline{z}$$

On an intersection we have $\overline{\partial}f_\alpha-\overline{\partial}f_\beta=\overline{\partial}f_{\alpha\beta}=0$ hence these forms patch to a globally defined section σ of ξ with values in type $(0,1)$ forms.

Conversely if such a differentiable section is given one can write it $\overline{\partial}f_\alpha$ on each U_α because by an application of Cauchy integral formula one can always solve $\overline{\partial}f=\sigma$ on a connected set with f a differentiable function. Then one can define on an intersection $f_{\alpha\beta}=f_\alpha-f_\beta$ which is analytic since $\overline{\partial}$ vanishes on it. Of course $f_{\alpha\beta}$ vanishes if and only if the set of all f_α define a section of ξ i.e. if $f_\alpha=f_\beta$. Hence $H^1(M,\mathcal{O}(\xi))$ can be identified with the set of differentiable sections of type $(0,1)$ of ξ, denoted $\Gamma^1_d(\xi)$, modulo the image by $\overline{\partial}$ of the set of differentiable sections of ξ.

$$H^1(M,\mathcal{O}(\xi))\equiv\Gamma^1_d(\xi)/\overline{\partial}\Gamma_d(\xi)$$

It can be shown that $H^1(M,\mathcal{O}(\xi))$ is finite–dimensional.

This description is useful to understand the Serre duality between the spaces $H^1(M,\mathcal{O}(\xi))$ and the space $\Gamma(\kappa-\xi)$, defined by a non singular pairing:

$$H^1(M,\mathcal{O}(\xi))\times\Gamma(\kappa-\xi)\to\mathbb{C}$$

As we have seen in the discussion of the canonical bundle a section of $\kappa-\xi$ can be viewed as a holomorphic section of $-\xi$ with values in type $(1,0)$ forms. Take a differentiable element of $\Gamma^1(\xi)$ locally of the form

[1]If $\sum_\alpha r_\alpha=1$ with $\mathrm{supp}(r_\alpha)\in U_\alpha$ define $f_\alpha=\sum_\gamma r_\gamma f_{\alpha\gamma}$ where $r_\gamma f_{\alpha\gamma}$ is extended to 0 in $U_\alpha-U_\gamma$ and note that $f_\alpha-f_\beta=\sum_\gamma r_\gamma(f_{\alpha\gamma}-f_{\beta\gamma})=\sum_\gamma r_\gamma f_{\alpha\beta}=f_{\alpha\beta}$

$f(z,\overline{z})d\overline{z}$ and an holomorphic section of $\kappa-\xi$ locally of the form $g(z)dz$ and consider their wedge product, locally $f(z,\overline{z})g(z)\,dz\wedge d\overline{z}$. In the product the transition functions $t_{\alpha\beta}$ of ξ and $t_{\alpha\beta}^{-1}$ of $-\xi$ cancel so that one ends up with a globally defined volume form on the Riemann surface that we integrate. This defines the pairing. Note that if f is identified to 0, that is if $f(z,\overline{z})d\overline{z}=\overline{\partial}h$ one can integrate by parts obtaining the integral of $h\overline{\partial}g\wedge dz$ which vanishes since g is holomorphic. Hence the pairing is well defined between the considered spaces. Finally the essential point is that the pairing is non singular. Indeed any linear form acting on $fd\overline{z}$ may be written $\int f(z,\overline{z})g(z,\overline{z})\,dz\wedge d\overline{z}$ for some distribution g, since this form vanishes when $fd\overline{z}=\overline{\partial}h$, the distribution g is weakly analytic hence is an analytic function by the classical Weyl lemma. It follows that $\Gamma(\kappa-\xi)$ and $H^1(M,\mathcal{O}(\xi))$ are finite dimensional spaces of the same dimension.

15.9 The Riemann-Roch theorem.

Let $\Gamma(\xi)$ be the finite dimensional space of holomorphic sections of the line bundle ξ and $c(\xi)$ its Chern class.

Theorem. *On a Riemann surface of genus g with canonical bundle κ we have for any line bundle ξ:*

$$\dim\Gamma(\xi)-\dim\Gamma(\kappa-\xi)=c(\xi)+1-g \tag{15.3}$$

Proof. We first show that $\chi(\xi)=\dim\Gamma(\xi)-\dim H^1(M,\mathcal{O}(\xi))-c(\xi)$ is independent of the line bundle ξ.
As a first step we show that $\chi(\xi+\xi_p)=\chi(\xi)$ for any point bundle ξ_p with analytic section σ_p vanishing at p.
Note that any local or global analytic section of ξ can be multiplied by σ_p to produce a section of $\xi+\xi_p$ vanishing at p. This is clearly an injective homomorphism of the space of sections. It fails to be surjective for global sections, if some global sections of $\xi+\xi_p$ do not vanish at p. Hence we have two cases:
a) There exists a global section of $\xi+\xi_p$ non vanishing at p. We have $\dim\Gamma(\xi+\xi_p)=\dim\Gamma(\xi)+1$.
b) All global sections of $\xi+\xi_p$ vanish at p. We have $\dim\Gamma(\xi+\xi_p)=\dim\Gamma(\xi)$.
Now let us consider local sections over intersections $U_\alpha\cap U_\beta$ (and we assume that p does not belong to such intersections and $p\in U_\alpha$). The homomorphism $f_{\alpha\beta}\to\sigma_p f_{\alpha\beta}$ is clearly bijective in this case since σ_p does not vanish outside p and trivial elements are mapped to trivial elements. This induces a surjective homomorphism $H^1(M,\mathcal{O}(\xi))\to H^1(M,\mathcal{O}(\xi+\xi_p))$ which fails to be injective if there exists a nontrivial set of sections $f_{\alpha\beta}$ of ξ mapping to a trivial set of sections $f'_\alpha-f'_\beta$ of $\xi+\xi_p$, that is $\sigma_p f_{\alpha\beta}=f'_\alpha-f'_\beta$.

In case a) above, we can substract the global non vanishing section to all f'_γ so as to achieve $f'_\alpha(p)=0$. Then one can divide all f'_γ by σ_p getting f_α such that $f_{\alpha\beta}=f_\alpha-f_\beta$. Hence the mapping is bijective and we have $\dim H^1(M,\mathcal{O}(\xi))=\dim H^1(M,\mathcal{O}(\xi+\xi_p))$.

In case b) we have $\dim H^1(M,\mathcal{O}(\xi))=\dim H^1(M,\mathcal{O}(\xi+\xi_p))+1$. Indeed we construct a set of sections $f_{\alpha\beta}$ which maps on a trivial set $f'_{\alpha\beta}=f'_\alpha-f'_\beta$ where we have choosen $f'_\alpha(p)\neq 0$. If we had $f_{\alpha\beta}=f_\alpha-f_\beta$ we would get $f'_\alpha-\sigma_p f_\alpha=f'_\beta-\sigma_p f_\alpha$ hence defining a global section of $\xi+\xi_p$ which by hypothesis vanishes at p, yielding $f'_\alpha(p)=0$, a contradiction. If however $f'_\alpha(p)=0$ one can divide it by σ_p and $f_{\alpha\beta}$ is trivial. This shows that non trivial elements $\{f_{\alpha\beta}\}$ are parametrized by $f'_\alpha(p)$ and form a space of dimension 1 modulo trivial elements.
Since $c(\xi+\xi_p)=c(\xi)+1$, in both cases we get $\chi(\xi+\xi_p)=\chi(\xi)$. Starting from $\xi-\xi_p$ we get $\chi(\xi)=\chi(\xi-\xi_p)$. Hence for any divisor $\mathcal{D}$ and line bundle η associated to $\mathcal{D}$ we have $\chi(\xi+\eta)=\chi(\xi)$.
The rest of the proof is easier. There exists $\mathcal{D}$ such that $\dim\Gamma(\xi+\eta)\neq 0$. Otherwise one gets $c(\eta)=C^t-\dim H^1(M,\mathcal{O}(\xi+\eta))\leq C^t$ with $C^t=\dim H^1(M,\mathcal{O}(\xi))-\dim\Gamma(\xi)$ is independent of η. This is impossible since $c(\eta)$ can be arbitrarily large. Let σ be a non trivial analytic section of $\xi+\eta$. Since η has a

meromorphic section σ_η of divisor $\mathcal{D}$ we see that ξ has a non trivial meromorphic section σ/σ_η. Hence ξ is the line bundle associated to this meromorphic section. We have obtained the:

Proposition. *Any line bundle on the Riemann surface M has a non trivial meromorphic section of divisor $\mathcal{D}$ and is the line bundle associated to this divisor.*

We can now take for ξ the trivial bundle in $\chi(\xi+\eta) = \chi(\xi)$ and we see that $\chi(\eta)$ is independent of η as previously claimed. Taking into account the Serre duality formula we get:

$$\chi = \dim\Gamma(\xi) - \dim\Gamma(\kappa-\xi) - c(\xi)$$

When ξ is the trivial bundle $\dim\Gamma(\xi) = 1$ since global analytic functions on M are constants, hence $\chi = 1 - \dim\Gamma(\kappa)$. When $\xi = \kappa$ one gets $\chi = \dim\Gamma(\kappa) - 1 - c(\kappa)$ hence $\chi = -(1/2)c(\kappa)$. To compute $c(\kappa)$ we view the Riemann surface M as an n–sheeted covering of the Riemann sphere using a meromorphic function f on M. We start from a meromorphic differential ω on the Riemann sphere and take its pull-back $f^*(\omega)$ on M. On the sphere we can take $\omega = dz$. Its divisor has degree -2, since it has a double pole at infinity. Hence its pull–back has n poles of order 2 above ∞. Moreover for each branch point of order ν the mapping f is locally $z = f(w) = w^\nu$, so that $f^*(dz) = \nu w^{\nu-1}dw$ has a zero of order $\nu-1$ which is the multiplicity index of the branch point. So the total multiplicity of the zeroes is precisely the total multiplicity B of the branch points of the covering, which by the Riemann–Hurwitz formula eq.(15.2) with $g_0 = 0$ is equal to $2g-2+2n$. The Chern class of κ is therefore $2g-2$. Hence $\chi = 1-g$ yielding the Riemann–Roch theorem. Moreover we see that $\dim\Gamma(\kappa) = g$ which means that the space of globally defined analytic one forms is of dimension g. ∎

In general the Riemann-Roch formula eq.(15.3) relates two unknown quantities. However if $c(\xi) > 2g-2$ we see that $c(\kappa-\xi) < 0$ hence $\dim\Gamma(\kappa-\xi) = 0$ and we get

$$\dim\Gamma(\xi) = c(\xi) + 1 - g, \quad if \quad c(\xi) > 2(g-1) \tag{15.4}$$

Consider now the meromorphic functions on M. Let $\Gamma(M,\mathcal{D})$ be the set of meromorphic functions on M whose divisor is bigger that $-\mathcal{D}$; i.e. $f \in \Gamma(M,\mathcal{D})$ iff the orders of its poles is less that those specified by $-\mathcal{D}$ and the order of its zeroes is greater than those specified by $-\mathcal{D}$. The space of holomorphic sections of the line bundle associated to $\mathcal{D}$ is isomorphic to $\Gamma(M,\mathcal{D})$, since this line bundle has a meromorphic section of divisor $\mathcal{D}$ and any section is obtained by multiplication by a meromorphic function. The section will be holomorphic iff the divisor of the function is greater than $-\mathcal{D}$.

Corollary. *The dimension of the space of meromorphic functions having at most k prescribed poles and at least l prescribed zeroes is greater or equal to $k-l+1-g$, and equal if $k-l \geq 2g-2$.*

If $k-l \geq g$ the equality is satisfied for generic positions of the prescribed zeroes and poles. Indeed let us take for simplicity $l = 0$. Then $\dim\Gamma(\kappa-\xi)$ is the dimension of the space of holomorphic differentials having k prescribed zeroes. But the space of holomorphic differentials is of dimension g and we want to impose k linear conditions. This is generically impossible for $k \geq g$.

Note that we have previously illustrated this situation by constructing a meromorphic function with $g+1$ prescribed poles on an hyperelliptic surface.

15.10 Abelian differentials.

Consider a Riemann surface M of genus g. Let a_i, b_i be a basis of cycles on M with canonical intersection matrix $(a_i \cdot a_j) = (b_i \cdot b_j) = 0,\ (a_i \cdot b_j) = \delta_{ij}$. This means that one can take differentiable loops $t \to a_i(t)$

and $t \to b_i(t)$ such that there is no intersection between loops a_i and a_j or b_j with $j \neq i$, while a_i and b_i intersect at just one point P. Moreover at P the tangent vectors a_i' and b_i' form a positively oriented (for the orientation given by the complex space structure) basis of the tangent space at P. One can then continuously deform these loops without changing the intersection index which is the sum of signs ± 1 at each intersection according to the orientation of the tangent vectors. In particular one can deform the loops a_i and b_i so that they have a common base point and then cut the Riemann surface along them getting a polygon with some edges identified. The boundary of this polygon can be described as $a_1 \cdot b_1 \cdot a_1^{-1} \cdot b_1^{-1} \cdots a_g \cdot b_g \cdot a_g^{-1} \cdot b_g^{-1}$ where the identifications are obvious. The common base point becomes all the vertices of the polygon.

The globally defined analytic one–forms on M are called *abelian differentials of first kind*. They form a space of dimension g over the complex numbers. Note that such a differential has $2g-2$ zeroes on the Riemann surface since $c(\kappa) = 2g-2$ and it has no pole. There is a natural pairing between these forms and loops obtained by integrating the form along the loop. It can be shown that the pairing between a-cycles and differentials is non degenerate (note they have the same dimension g). This is a consequence of the Riemann bilinear identities that we shall describe below. Hence we choose a basis of first kind abelian differentials, which we denote by ω_j, $j = 1, \cdots, g$, normalized with respect to the a-cycles :

$$\oint_{a_j} \omega_i = \delta_{ij}. \tag{15.5}$$

The matrix of b-periods is then defined as the matrix $\mathcal{B}$ with matrix elements :

$$\mathcal{B}_{ij} = \oint_{b_i} \omega_j \tag{15.6}$$

Taking the example of an hyperelliptic surface $y^2 = P_{2g+1}(x)$ where P is a polynomial of degree $2g+1$, a basis of regular abelian differentials is provided by the forms

$$\omega_j = x^j dx/y \quad \text{for } j = 0, \cdots, g-1$$

Indeed this is regular except perhaps at the branch points and at ∞. At a branch point the local parameter is y and we have $y^2 = a(x-b) + \cdots$ hence $x^j dx/y = (2b^j/a)(1+\cdots)dy$ which is regular. At ∞ we take $x' = 1/x$ and $y' = y/x^{(g+1)}$ so that $y'^2 = ax' + \cdots$ and $x^j dx/y = by'^{2(g-j-1)}dy'$ which is regular for $j \leq g-1$ since y' is the local parameter. Of course these forms are unnormalized. Note that their $(2g-2)$ zeroes are located at $x = 0$ and $x = \infty$.

Similarly *abelian differentials of second kind* are meromorphic differentials with poles of order greater than 2. Given a point P on M there exists an abelian differential of second kind whose only singularity is a pole of second order at p. Indeed by eq.(15.3) we see that $\dim \Gamma(\kappa + 2\xi_p) \geq c(\kappa + 2\xi_p) + 1 - g = g+1$. Such sections come from abelian differentials multiplied by σ_p^2 and since the space of regular differentials is of dimension g we see that there exists at least one section in that space whose division by σ_p^2 has a second order pole at p. Adding a proper combination of differentials of first kind one can always insure that all a–periods of the second kind differential vanish. Such a second kind differential is called normalized.

We can apply the Cauchy theorem to a meromorphic globally defined one–form of type (1,0), yielding the vanishing of the sum of residues of first order poles. Note that the residue of a meromorphic one–form is intrinsically defined since $\text{res} = (1/2\pi i)\oint \omega$ where the contour is a small loop around the given singularity. So we define *abelian differentials of third kind* as general meromorphic differentials with first order poles whose sum of residues vanish. Given two points p and q there exists a third kind differential whose only singularities are a pole of order 1 at p with residue 1, and a pole of order 1 at q with residue -1, and moreover normalized, i.e. such that all its a–periods vanish.

15.11 Riemann Bilinear Identities.

On Riemann surfaces provided with canonical cycles there is a pairing between meromorphic differentials. Namely, let Ω_1 and Ω_2 be two meromorphic differentials on M. The pairing $(\Omega_1 \bullet \Omega_2)$ is then defined by integrating them along the canonical cycles as follows :

$$(\Omega_1 \bullet \Omega_2) = \sum_{i=1}^{g} \left(\int_{a_j} \Omega_1 \int_{b_j} \Omega_2 - \int_{a_j} \Omega_2 \int_{b_j} \Omega_1 \right)$$

The Riemann bilinear identity expresses this quantity in terms of sum over residues:

Proposition. *Let g_1 be defined on the Riemann surface cut along the canonical cycles such that $dg_1 = \Omega_1$. We have:*

$$(\Omega_1 \bullet \Omega_2) = 2i\pi \sum_{\text{poles}} \text{res}(g_1 \Omega_2) \tag{15.7}$$

Proof. One computes $\oint g_1 \Omega_2$ on the boundary of the polygon representing the Riemann surface. On one hand this produces the sum of residues in eq.(15.7) on the other hand we compute this integral explicitely:

$$\int_{a_j a_j^{-1}} g_1 \Omega_2 = \int_{a_j} g_1 \Omega_2 - \int_{a_j} (g_1 + \int_{b_j} \Omega_1) \Omega_2$$

since g_1 is shifted by $\sum_{b_j} \Omega_1$ when crossing the cut a_j. Similarly, one gets:

$$\int_{b_j b_j^{-1}} g_1 \Omega_2 = \int_{b_j} g_1 \Omega_2 - \int_{b_j} (g_1 - \int_{a_j} \Omega_1) \Omega_2$$

Adiing the contributions for all j one gets $(\Omega_1 \bullet \Omega_2)$. ■

Corollary. *The matrix of b–periods $\mathcal{B}$ is symmetric.*

Indeed the pairing between the normalized holomorphic differentials is trivial : $(\omega_i \bullet \omega_j) = 0$ for $i, j = 1, \cdots, g$. This reads $\mathcal{B}_{ij} = \mathcal{B}_{ji}$.

Corollary. *If Ω_2 is a normalized differential of second kind with a pole of order n with principal part $z^{-n}dz$ at $z = 0$ for some local parameter z and $\Omega_1 = \omega_k$ is a normalized holomorphic differential expanded as $\omega_k = (\sum_{i=0}^{\infty} c_i z^i) dz$ around $z = 0$ one has:*

$$\int_{b_k} \Omega_2 = 2\pi i \frac{c_{n-2}}{n-1}$$

15.12 Jacobi Variety.

We have seen that differential line bundles on a Riemann surface are classified by their Chern class. It is not so for analytic line bundles. To describe their classification it is sufficient to consider the different analytic structures on line bundles of Chern class 0. The space of such structures is called the Jacobi variety. Since analytic line bundles are the same as equivalence classes of divisors on the Riemann surface, the Jacobi variety identifies with equivalence classes of divisors of degree 0. It is thus necessary to characterize the divisors of meromorphic functions.

In order to do that consider a divisor of degree 0 which can always be written $\mathcal{D} = \sum_i (p_i - q_i)$, with non necessarily distinct points. Choose paths γ_i from q_i to p_i and associate to $\mathcal{D}$ the point in $\mathbb{C}^g$ of coordinates:

$$\rho_k(\mathcal{D}) = \sum_i \int_{\gamma_i} \omega_k, \quad k = 1 \cdots g$$

Such sums are called Abel sums If the paths are homotopically deformed these integrals remain constant by the Cauchy theorem, but if one makes a further loop around a_k then $\rho_l \to \rho_l + \delta_{kl}$, while if one adds a loop around b_k, $\rho_l \to \rho_l + \mathcal{B}_{kl}$, hence this gives a well–defined point on the torus:

$$J(M) = \mathbb{C}^g / (\mathbb{Z}^g + \mathcal{B}\mathbb{Z}^g) \tag{15.8}$$

where $\mathcal{B}$ is the matrix of the b-periods. Moreover if one permutes independently the points p_i and q_i the point in the torus does not change. Indeed if γ_1' connects q_1 to p_2, γ_2' connects q_2 to p_1 and σ connects q_1 to q_2 one has $\int_{\gamma_1} \omega = \int_{\gamma_1'} \omega - \int_\sigma \omega$ up to periods and $\int_{\gamma_2} \omega = \int_{\gamma_2'} \omega + \int_\sigma \omega$ up to periods, so $\int_\sigma \omega$ cancels in the sum. Note that $J(M)$ is an additive group and that the above map from line bundles to points of $J(M)$ is an homomorphism.

The theorems of Abel and Jacobi state that this point on the torus $J(M)$ characterizes the divisor $\mathcal{D}$ up to equivalence, so that the Jacobian variety can be identified with the g–dimensional complex torus $J(M)$.

Theorem (Abel). *A divisor $\mathcal{D} = \sum_i (p_i - q_i)$ is the divisor of a meromorphic function if and only if for any first kind abelian differential ω the Abel sum $\sum_i \int_{\gamma_i} \omega$ vanishes modulo $\mathbb{Z} + \mathcal{B}\mathbb{Z}$ for any choice of paths γ_i from q_i to p_i.*

Theorem (Jacobi). *For any point $\lambda \in J(M)$ and a fixed reference divisor $\mathcal{D}_0 = \sum_{i=1}^g q_i$ one can find a divisor of g points $\mathcal{D} = p_1 + \cdots + p_g$ on M such that $\rho_k(\mathcal{D} - \mathcal{D}_0)$ maps to λ. Moreover for generic λ the divisor $\mathcal{D}$ is unique.*

Proof. Let f be a meromorphic function with divisor $\mathcal{D} = \sum_i (p_i - q_i)$ and consider for $\lambda \in \mathbb{C}$ the pencil of divisors $\mathcal{D}^\lambda = \sum_i (p_i^\lambda - q_i)$ of the meromorphic functions $f + \lambda$ (the poles are those of f, the zeroes vary analytically with λ). Finally let $\phi(\lambda)$ be the point in $\mathbb{C}/(\mathbb{Z}+\mathcal{B}\mathbb{Z})$ obtained by integrating ω along paths from q_i to p_i^λ. The map ϕ is obviously analytic and can be extended to an analytic map on the Riemann sphere. Indeed when $\lambda \to \infty$ we have $p^\lambda \to q$. But such a map from the Riemann sphere to a torus is necessarily constant, since dz is a regular differential on the torus, hence $\phi^*(dz)$ has to vanish on the Riemann sphere. Since $\phi(\infty) = 0$ one gets $\phi = 0$.
In order to prove the converse and the Jacobi theorem one has to exhibit particular functions or divisors. This will also be a consequence of the powerfull Riemann theorem that we will show later on. To show the generic uniqueness of the g points mapping to a given λ note that we have g equations for g unknowns, hence the space of solutions is generically of dimension 0. If we have two solutions, there exists a meromorphic function f whose poles and zeroes are respectively these solutions. Then $f + \lambda$ for $\lambda \in \mathbb{C}$ relates the divisor of poles to a one parameter family of equivalent divisors, hence the solution space would be of dimension 1, a contradiction. ∎

One can embed the Riemann surface M into its Jacobian $J(M)$ by the Abel map. Namely, choose a point $q_0 \in M$ and define the vector $\mathcal{A}(P)$ with coordinates $\mathcal{A}_k(P)$ modulo the above lattice:

$$\mathcal{A} : M \quad \longmapsto \quad J(M) \tag{15.9}$$

$$\mathcal{A}_k(P) \quad = \quad \int_{q_0}^P \omega_k \tag{15.10}$$

Clearly, the Abel map depends on the point q_0. But changing this point just amounts to a translation in $J(M)$. The Abel map is an analytic embedding of the Riemann surface into the g–dimensional torus $J(M)$ i.e. is injective.

15.13 Theta functions.

One can show using Riemann bilinear type identities that the imaginary part of the period matrix $\mathcal{B}$ is a positive definite quadratic form. This allows to define the Riemann theta-function:

$$\theta(z_1, \ldots, z_g) = \sum_{m \in \mathbb{Z}^g} e^{2\pi i(m,z)+\pi i(\mathcal{B}m,m)}. \tag{15.11}$$

since the series is convergent, as an analytic function on $\mathbb{C}^g$.

The theta function has simple automorphy properties with respect to the period lattice of the Riemann surface: for any $l \in \mathbb{Z}^g$ and $z \in \mathbb{C}^g$

$$\begin{aligned} \theta(z+l) &= \theta(z) \\ \theta(z+\mathcal{B}l) &= \exp[-i\pi(\mathcal{B}l, l) - 2i\pi(l,z)]\theta(z) \end{aligned} \tag{15.12}$$

The divisor of the theta function is the set of points in the Jacobian torus where $\theta(z) = 0$. Note that this is an analytic subvariety of dimension $g-1$ of the torus, well–defined due to the automorphy property.

The fundamental theorem of Riemann expresses the intersection of the image of the embedding of M into $J(M)$ with the divisor of the theta function.

Theorem. *Let $w = (w_1, \cdots, w_g) \in \mathbb{C}^g$ arbitrary. Either the function $\theta(\mathcal{A}(P) - w)$ vanishes identically for $P \in M$ or it has exactly g zeroes $P_1, \cdots, P_g$ such that:*

$$\mathcal{A}(P_1) + \cdots + \mathcal{A}(P_g) = w - \mathcal{K} \tag{15.13}$$

where $\mathcal{K}$ is the so–called vector of Riemann's constants, depending on the curve M and the point q_0 but independent of w.

Proof. We first dissect the Riemann surface M as explained above obtaining a polygon with boundary $a_1 \cdot b_1 \cdot a_1^{-1} \cdot b_1^{-1} \cdots a_g \cdot b_g \cdot a_g^{-1} \cdot b_g^{-1}$ in $\mathbb{C}$. Consider the analytic function on the polygon (or more precisely on the Riemann surface cut along the previous loops):

$$f(P) = \theta(\mathcal{A}(P) - w)$$

Assuming that f does not vanish identically it has discrete zeroes P_i. Since it has no pole the number of these zeroes is given by:

$$\#\,\text{zeroes} = \frac{1}{2\pi i} \oint \frac{df}{f}$$

where the integral is taken on the boundary of the polygon. This expresses as a sum of terms on the arcs a_k and b_k. The integrals on the arcs b_k and b_k^{-1} are related by a translation by the a_k period, hence cancel. Similarly the difference of the integrals on a_k and a_k^{-1} reduces by the automorphy property to $2i\pi \int_{a_k} \omega_k = 2i\pi$. Thus $\#\,\text{zeroes} = g$.
To prove the second identity we proceed similarly by considering the integral:

$$\frac{1}{2\pi i} \oint g_k \frac{df}{f}, \quad \text{with } dg_k = \omega_k, \text{ and } g_k(q_0) = 0$$

computed on the edge of the polygon, where g_k is well–defined on the polygon. On one hand this integral is equal to the sum of residues which occur at the zeroes of f, and produces:

$$\sum_{j=1}^{g} g_k(P_j) = \sum_j \mathcal{A}_k(P_j)$$

On the other hand it can be computed as a sum over arcs using the automorphy properties of the theta function:

$$\int_{a_j a_j^{-1}} g_k d(\log f) = -\int_{a_j} (g_k + \mathcal{B}_{jk})\big(d(\log f) - 2i\pi d\mathcal{A}_j(p)\big) + \int_{a_j} g_k d(\log f) \equiv 2i\pi \int_{a_j} g_k \omega_j$$

modulo $\mathcal{B}$ periods.

$$\int_{b_j b_j^{-1}} g_k d(\log f) = \int_{b_j} (g_k + \delta_{jk}) d(\log f) - \int_{b_j} g_k d(\log f) \equiv \delta_{jk}(2i\pi w_j - 2i\pi \mathcal{A}_j(q_1))$$

modulo periods, where q_1 is the base point of all the loops at the boundary. Putting all together one gets the Riemann formula with $\mathcal{K}$ given by a complicated expression, independent of w. ∎

Corollary (Jacobi's theorem). *Any point in the Jacobian $J(M)$ is the image of some degree g divisor $P_1 + \cdots + P_g$ on M.*

Proof. One has to find g points such that $\mathcal{A}(P_1) + \cdots + \mathcal{A}(P_g) = z$ modulo periods for given z. We find these points by solving the equation $\theta(\mathcal{A}(P) - \mathcal{K} - z) = 0$. ∎

The divisor of the zeroes of a θ-function has also a nice charaterization in terms of points on the Riemann surface.

Theorem. *The zero divisor of a theta function can be written as*

$$X = \mathcal{K} - \sum_{i=1}^{g-1} \mathcal{A}(\eta_i)$$

Proof. Let X be a point of the Θ divisor, i.e., $\theta(X) = 0$. Consider the function

$$\theta(\mathcal{A}(P) + X) \tag{15.14}$$

By Riemann's theorem, the zeroes of this function are such that (if it does not vanish identically)

$$\mathcal{A}(P_1) + \cdots + \mathcal{A}(P_g) + X = -\mathcal{K}$$

Among these zeroes, one has necessarily q_0, the base point of Abel's map (since $\theta(X) = 0$), say $P_g = q_0$. Hence we have $X = -\mathcal{K} - \sum_{i=1}^{g-1} \mathcal{A}(P_i)$.

Conversely if X is of this form, one solves eq.(15.14) producing a divisor of g points $\sum_{i=1}^{g} Q_i$ such that $\sum_{i=1}^{g-1} \mathcal{A}(P_i) = \sum_{i=1}^{g} \mathcal{A}(Q_i)$. A solution is obviously $P_i = Q_i$ for $i = 1, \cdots, g-1$ and $Q_g = q_0$. This solution is generically unique up to permutation due to the Jacobi theorem. Equation (15.14) for $P = q_0$ reads $\theta(X) = 0$. Note that the space of points of the form $X = -\mathcal{K} - \sum_{i=1}^{g-1} \mathcal{A}(P_i)$ and the solution of $\theta(X) = 0$ are both closed in the Jacobian, hence they are equal. ∎

The Riemann theorem can be used to express meromorphic functions in terms of θ–functions. Let f be a meromorphic function with g poles at points $\delta_1, \ldots, \delta_g$ and an additional pole at the point Q^+ and one of its zeroes at a specified point Q^-; i.e. with divisor greater than $D = -\delta_1 - \cdots - \delta_g - Q^+ + Q^-$. By the Riemann–Roch theorem there is a unique such function generically. Let w, w^+, w^-, w^0 be vectors defined by the formulae:

$$\begin{aligned} w &= \sum_{s=1}^{g} \mathcal{A}(\delta_s) + \mathcal{K} \\ w^+ &= \mathcal{A}(Q^+) + \sum_{s=2}^{g} \mathcal{A}(\delta_s) + \mathcal{K} \\ w^- &= \mathcal{A}(Q^-) + \sum_{s=2}^{g} \mathcal{A}(\delta_s) + \mathcal{K} \\ w^0 &= w + w^+ - w^-. \end{aligned}$$

Let us define the function

$$f(P) = \frac{\theta(\mathcal{A}(P) - w^-)\theta(\mathcal{A}(P) - w^0)}{\theta(\mathcal{A}(P) - w)\theta(\mathcal{A}(P) - w^+)} \tag{15.15}$$

From the Riemann theorem it follows that two factors in the denominator vanish generically at the points $\delta_1, \ldots, \delta_g$ and $Q^+, \delta_2, \ldots, \delta_g$ respectively. Similarly the two factors in the numerator vanish at $Q^-, \delta_2, \ldots, \delta_g$ and g other points. The zeroes at $\delta_2, \ldots, \delta_g$ cancel between the numerator and the denominator, thereby leaving us with the correct divisor of zeroes and poles. It remains to show that the function f is well–defined on M. This is because, due to the definition of w^0 the automorphy factors of the theta functions in equation (15.12) cancel between the numerator and the denominator when P describes b–cycles on M. The converse of Abel's theorem results from an analogous construction.

15.14 The genus one case.

The application of these ideas to the genus one case is the classical theory of elliptic functions. A genus one analytic surface can be viewed as the quotient of $\mathbb{C}$ by a lattice, whose periods are denoted by a classical convention $2\omega_1, 2\omega_2$. In this case the theorems of Abel and Jacobi identify the curve and its Jacobi variety. Note that dz is a well-defined regular analytic differential on this torus, and spans the one–dimensional space of abelian differentials. For any meromorphic function f on the torus, i.e. periodic with respect to the lattice, the differential fdz is a meromorphic differential, so the sum of its residues vanishes. Thus f has at least two poles or a pole of order two, and a meromorphic function with just one pole is impossible. The main example is the Weierstrass $\wp$ function:

$$\wp(z) = \frac{1}{z^2} + \sum_{m,n\neq 0} \left\{ \frac{1}{(z - 2m\omega_1 - 2n\omega_2)^2} - \frac{1}{(2m\omega_1 + 2n\omega_2)^2} \right\}$$

An analogue of the θ function is provided by the Weierstrass σ–function:

$$\sigma(z) = z \prod_{m,n\neq 0} \left(1 - \frac{z}{\omega_{mn}}\right) \exp\left[\frac{z}{\omega_{mn}} + \frac{1}{2}\left(\frac{z}{\omega_{mn}}\right)^2\right] \tag{15.16}$$

with $\omega_{mn} = 2m\omega_1 + 2n\omega_2$. This function is related to the $\wp$–function by the equations:

$$\zeta(z) = \frac{\sigma'(z)}{\sigma(z)}\,, \qquad \wp(z) = -\zeta'(z), \tag{15.17}$$

The $\wp$-function is doubly periodic, and the σ-function and ζ-functions transform according to:

$$\zeta(z + 2\omega_l) = \zeta(z) + 2\eta_l, \quad \sigma(z + 2\omega_l) = -\sigma(z)e^{2\eta_l(z+\omega_l)}$$

The Riemann bilinear identity applied to the forms $\Omega_1 = -\wp(z)dz$ and $\Omega_2 = dz$ yields:

$$2(\eta_1\omega_2 - \eta_2\omega_1) = i\pi$$

These functions have the symmetries

$$\sigma(-z) = -\sigma(z)\ , \qquad \zeta(-z) = -\zeta(z)\ , \qquad \wp(-z) = \wp(z). \tag{15.18}$$

Their behaviour at the neighbourhood of zero is

$$\sigma(z) = z + O(z^5)\ , \qquad \zeta(z) = z^{-1} + O(z^3)\ , \qquad \wp(z) = z^{-2} + O(z^2). \tag{15.19}$$

It is useful for the study of the elliptic Calogero model to introduce the Lamé function:

$$\Phi(x,z) = \frac{\sigma(z-x)}{\sigma(x)\ \sigma(z)} e^{\zeta(z)x} \tag{15.20}$$

It has the symmetry property:

$$\Phi(-x,z) = -\Phi(x,-z) \tag{15.21}$$

The function $\Phi(x,z)$ is a double-periodic function of the variable z:

$$\Phi(x, z+2\omega_l) = \Phi(x,z), \tag{15.22}$$

and has the expansion of the form:

$$\Phi(x,z) = (-z^{-1} + \zeta(x) + O(z))e^{\zeta(z)x} \tag{15.23}$$

at the point $z = 0$. As a function of x it has the following monodromy properties:

$$\Phi(x+2\omega_l, z) = \Phi(x,z)\exp 2(\zeta(z)\omega_l - \eta_l z). \tag{15.24}$$

and has a pole at the point $x = 0$.

$$\Phi(x,z) = x^{-1} + O(x), \tag{15.25}$$

The function Φ is a solution of the Lamé equation:

$$\big(\frac{d^2}{dx^2} - 2\wp(x)\big)\Phi(x,z) = \wp(z)\Phi(x,z). \tag{15.26}$$

Choosing the periods $\omega_1 = \infty$ and $\omega_2 = i\frac{\pi}{2}$ we obtain the hyperbolic case

$$\sigma(z) \to \sinh(z)\exp\left(-\frac{z^2}{6}\right)\ , \qquad \zeta(z) \to \coth(z) - \frac{z}{3}\ , \qquad \wp(z) \to \frac{1}{\sinh^2(z)} + \frac{1}{3} \tag{15.27}$$

and

$$\Phi(x,z) \to \frac{\sinh(z-x)}{\sinh(z)\sinh(x)} e^{x\coth z} \tag{15.28}$$

In the rational limit, we have

$$\sigma(z) \to z, \quad \zeta(z) \to \frac{1}{z}, \quad \wp(z) \to \frac{1}{z^2}, \quad \Phi(x,z) \to \left(\frac{1}{x} - \frac{1}{z}\right) e^{\frac{x}{z}}$$

References

[1] G. Springer *Introduction to Riemann surfaces.* Addison–Wesley (1957)

[2] R.C. Gunning *Lectures on Riemann surfaces.* Princeton University press (1966)

[3] P. Griffiths and J. Harris *Principles of algebraic geometry.* Wiley (1978)

[4] D. Mumford *Tata lectures on theta.* Volumes I and II Birkhauser (1983)

[5] J.P. Serre *Algebraic groups and class fields.* Springer

[6] E.T. Whittaker and G.N. Watson *A course of modern analysis.* Cambridge university press (1902)

[7] J. Fay *Theta functions on Riemann surfaces.* Springer lectures notes (1973)

Chapter 16

Lie algebras.

16.1 Lie groups and Lie algebras.

A Lie group is a group G which is at the same time a differentiable manifold, and such that the group operation $(g,h) \to gh^{-1}$ is differentiable. Due to a theorem of Montgomery and Zippin, the differentiable structure is automatically real analytic.

The maps $h \to gh$ and $h \to hg$ are called respectively left and right translations by g. Their differentials at the point h maps the tangent space $T_h(G)$ respectively to $T_{gh}(G)$ and $T_{hg}(G)$. We will denote by $g \cdot X$ and $X \cdot g$ the images of $X \in T_h(G)$ by these maps. This notation is coherent because, differentiating the associativity condition in G one gets $(g \cdot X) \cdot h = g \cdot (X \cdot h)$, and $g \cdot (h \cdot X) = (gh) \cdot X$.

In particular, this last relation shows that, for any X in the tangent space of G at the unit element e, the vector field with value $g \cdot X$ at g is invariant under any left translation. Conversely any such left invariant vector field is of the form $g \cdot X$. So in the following we identify left invariant vector fields on G and elements of $T_e(G)$. This vector space is called the Lie algebra of G and will be denoted by $\mathcal{G}$ and is finite dimensional. Alternatively one can define the vector field $X(g)$ by its action on any function f, namely $(X \cdot f)(g)$ is the derivative of f along the tangent vector X at g. This defines a new function $X \cdot f$. The left invariance of the vector field $X(g)$ means that $(X \cdot f)(hg) = X \cdot {}^h f(g)$ where ${}^h f : g \to f(hg)$. In other words the differential operator X commutes with left translations. More generally, for $X_1, X_2, \cdots \in \mathcal{G}$, one can consider linear combinations of differential operators (of any order) of the form $X_1 \cdot (X_2 \cdots (X_k \cdot f) \cdots)$ which obviously form an associative algebra of left invariant differential operators. The Lie algebra is embedded into this associative algebra as the set of first order differential operators. One defines the Lie bracket $[X,Y] = XY - YX$, in terms of the associative algebra product. The main point is that, while XY and YX are second order differential operators, their commutator is a first order differential operator, so that $[X,Y]$ belongs to the Lie algebra. It is clear from this definition that $(X,Y) \to [X,Y]$ is bilinear antisymmetric and obeys the Jacobi identity:

$$[[X,Y],Z] + [[Z,X],Y] + [[Y,Z],X] = 0 \tag{16.1}$$

The associative algebra of left–invariant differential operators on G is called the universal enveloping algebra of the Lie algebra $\mathcal{G}$ and will be denoted by $\mathcal{U}(\mathcal{G})$.

Finally there is a natural action of the Lie group G on its Lie algebra $\mathcal{G}$ called the adjoint action. Remark that for any X in the tangent space at e and any g, $g \cdot X \cdot g^{-1}$ is also in the tangent space at e to G. We define the adjoint action:

$$\mathrm{Ad}\,(g)(v) = g \cdot v \cdot g^{-1}, \quad v \in \mathcal{G} \tag{16.2}$$

and note that $\mathrm{Ad}_{gh} = \mathrm{Ad}_g \mathrm{Ad}_h$, so this is a group action of G on $\mathcal{G}$.

Let $X \in \mathcal{G}$ and consider the left invariant vector field $g \cdot X$. For small t we can solve the differential equation $\dot{g} = g \cdot X$ with initial condition $g(0) = e$. The solution $g(t) \in G$ is such that $g(s+t) = g(s)g(t)$ for s, t small. This is because both members solve the above differential equation with initial value $g(s)$ for $t = 0$. One can then use this property to show that the solution of the differential equation extends to a domain larger than initially defined (if we have a solution for $|s| \leq \epsilon$ then $g(s+t)$ is a solution for $|s+t| \leq 2\epsilon$ and still solves the equation there), and successively extends to all of $\mathbb{R}$. The solution, defined for all t is denoted $\exp(tX)$. In particular for $t = 1$ this defines the exponential map from $\mathcal{G}$ to G (obviously $\exp(tX)$ belongs to the connected component of e in G, so in the following we assume that G is connected).

The exponential map allows to relate subgroups of G to subalgebras of $\mathcal{G}$. Let $\mathcal{H}$ be a subalgebra of $\mathcal{G}$, and H be the smallest subgroup of G containing all the $\exp(X)$ for $X \in \mathcal{H}$. One can show that H is a Lie subgroup of G with Lie algebra $\mathcal{H}$. The precise definition of a Lie subgroup is quite tricky, in particular H may be embedded in a very complicated way in G, as shown by the simple example below. Conversely, given H its Lie algebra $\mathcal{H}$ is the set of $X \in \mathcal{G}$ such that $t \to e^{tX}$ is a continuous curve in H.

Example. Consider the torus $\mathbb{R}^2/\mathbb{Z}^2$ with the Abelian group law induced by the addition in $\mathbb{R}^2$. This is a Lie group G, with Abelian Lie algebra $\mathbb{R}^2$. Consider the one–parameter subgroup $H = \{\exp(tX) | t \in \mathbb{R}\}$ which is a Lie subgroup. When X has irrational slope, this subgroup is dense in G. In particular any neighborhood U of e contains infinititely many components of H.

A much nicer situation is obtained when H is closed in G, and fortunately this is the situation of interest for our purposes. Indeed we are mostly interested in the case when the Lie group G acts differentiably on a manifold M, and H is the stabilizer of a point $m \in M$. In this case, since the operation is continuous, H is automatically closed in G. Remarkably this is sufficient to ensure that H is a closed Lie subgroup of G, thanks to a theorem of E. Cartan.

Theorem. *If H is a closed subgroup of a Lie group G, there exists a unique analytic structure on H such that H is a Lie subgroup of G.*

Proof. We sketch the proof. The idea is to define

$$\mathcal{H} = \{X \in \mathcal{G} | \exp(tX) \in H, \ \forall t\}$$

and show that this is a subalgebra of $\mathcal{G}$. Easy computations show that:

$$\begin{aligned}
(\exp \frac{t}{n} X \exp \frac{t}{n} Y)^n &= \exp \Big(t(X+Y) + O(1/n) \Big) \\
(\exp \frac{-t}{n} X \exp \frac{-t}{n} Y \exp \frac{t}{n} X \exp \frac{t}{n} Y)^{n^2} &= \exp \Big(t^2 [X,Y] + O(1/n) \Big)
\end{aligned}$$

so that, if $X, Y \in \mathcal{H}$, both $(X+Y)$ and $[X,Y]$ are in $\mathcal{H}$ since H is closed. One then uses the closedness of H to show that the Lie subgroup of G of Lie subalgebra $\mathcal{H}$ is in fact equal to H (more precisely to the connected component of the identity in H). ∎

In the situation described above where G acts on M, one can show that the application $g \to g \cdot m$ of G onto the orbit $\mathcal{O}_m$ of m is open so that $\mathcal{O}_m$ is isomorphic to the homogeneous space G/H.

Examples. The most natural examples of Lie groups are provided by so–called algebraic subgroups of the general linear group $GL(n)$. These are subvarieties of $GL(n)$ defined by polynomial equations compatible with the multiplication law. For example the special linear group is defined by the equation $\det g = 1$, and the product of two such matrices has determinant one. Hence these are naturally closed Lie subgroups of $GL(n)$. The other standard examples are the subgroups of orthogonal and symplectic matrices.

For any Lie group G a representation ρ on a vector space V is a differentiable group homomorphism $G \to GL(V)$. For $g \in G$ and $v \in V$ one denotes $g \cdot v = \rho(g)(v)$ so that $(gh) \cdot v = g \cdot (h \cdot v)$. The differential of ρ at e maps the Lie algebra $\mathcal{G}$ to the Lie algebra $gl(V)$ of $GL(V)$. Similarly left invariant vector fields on G are mapped on left invariant vector fields on $GL(V)$, and so are their Lie brackets. Hence we get a representation of $\mathcal{G}$ on $gl(V)$, that we shall also denote ρ. In other words, we have $[\rho(X), \rho(Y)] = \rho([X, Y])$. Such a representation is faithful if $\rho : \mathcal{G} \to gl(V)$ is injective (i.e. for $X \neq 0$ there exists $v \in V$ such that $\rho(X)(v) \neq 0$). In this case $\mathcal{G}$ may be seen as a subalgebra of $gl(V)$. There is a natural representation of any Lie group on its Lie algebra, i.e. $V = \mathcal{G}$, given by the adjoint representation. This induces a representation of $\mathcal{G}$ on $\mathcal{G}$, also called the adjoint representation:

$$\mathrm{ad}_X(Y) = [X, Y] \tag{16.3}$$

It is easy to check that this is a representation of $\mathcal{G}$, using the Jacobi identity. Almost all results on Lie algebras are obtained by studying this representation.

16.2 Semi–simple Lie algebras.

Because there is such an interplay between Lie groups and Lie algebras we study here Lie algebras from an algebraic viewpoint. In this section we consider Lie algebras over the complex numbers, e.g. complexifications $\mathcal{G}_{\mathbb{R}} \otimes \mathbb{C}$ of real Lie algebras.

We will often use a basis (t_a), $a = 1, \cdots, \dim \mathcal{G}$, on the complex Lie algebra $\mathcal{G}$. The Lie bracket is then expressed as:

$$[t_a, t_b] = f_{ab}^c \; t_c$$

The coefficients f_{ab}^c are called structure constants. Remark that in this basis the matrix elements of the adjoint representation are $(\mathrm{ad}_{t_a})_b^c = f_{ab}^c$.

The adjoint representation (16.3) allows to define a natural bilinear form on $\mathcal{G}$, also called the Killing form by:

$$(X, Y) = \mathrm{Tr}\,(\mathrm{ad}_X \, \mathrm{ad}_Y)$$

This bilinear form is *invariant* in the sense that:

$$([X, Y], Z) = (X, [Y, Z])$$

This results immediately from the cyclic invariance of the trace and the fact that $X \to \mathrm{ad}_X$ is a representation. The invariance property also means that $(\mathrm{ad}_Y \, X, Z) + (X, \mathrm{ad}_Y \, Z) = 0$.

A Lie algebra is said to be semi–simple if it does not contain any non-trivial abelian ideal. The Cartan criterion says that this is the case if and only if the Killing form is non degenerate. In one direction this is easy. Indeed if X belongs to an abelian ideal $\mathcal{I}$, and Y is arbitrary, choose a basis t_a of $\mathcal{G}$ such that t_a is a basis of $\mathcal{I}$ for $i = 1, \cdots, r$. Then $\mathrm{ad}_Y \mathrm{ad}_X(t_a)$ vanishes for $a \leq r$ and belongs to $\mathcal{I}$ for $a > r$, hence $\mathrm{ad}_Y \, \mathrm{ad}_X$ has no diagonal element in this basis, and its trace vanishes. We see that any abelian ideal is in the kernel of the Killing form.

A Lie algebra $\mathcal{G}$ is called simple if it is semi–simple and its only ideals are either $\{0\}$ or the algebra $\mathcal{G}$ itself. For any ideal $\mathcal{I}$ in a semi–simple algebra, its orthogonal under the Killing form is also an ideal. Moreover by invariance and non degeneracy of the Killing form one sees that $\mathcal{I} \cap \mathcal{I}^\perp$ is an abelian ideal, (for $X \in \mathcal{I}$, $Y \in \mathcal{I}^\perp$, and Z arbitrary $([X, Y], Z) = (X, [Y, Z]) = 0$) hence vanishes. It follows that $\mathcal{G}$ is the direct sum of its simple ideals, this being an orthogonal decomposition.

We now introduce the concept of a Cartan subalgebra in a semi–simple Lie algebra. First an element X of $\mathcal{G}$ is called *semi–simple* if ad_X is a diagonalizable matrix in the adjoint representation. A Cartan subalgebra $\mathcal{H}$ of a semi–simple Lie algebra $\mathcal{G}$ is a maximal abelian subalgebra of $\mathcal{G}$ whose elements are all semi–simple. The existence of such an algebra is a very non trivial result. To construct it one starts with a *regular* element, that is an element of $\mathcal{G}$ such that ad_X has a maximal number of distinct eigenvalues (as a matrix in the adjoint representation). Then the subalgebra of $\mathcal{G}$ on which ad_X is nilpotent is a Cartan subalgebra. One can show that any two Cartan subalgebras are related by a Lie algebra automorphism, and their common dimension is called the rank of the Lie algebra and will be denoted as $\mathrm{rank}\,\mathcal{G}$. In the adjoint representation, the endomorphisms $\mathrm{ad}\,(H)$ for $H \in \mathcal{H}$ form a system of commuting diagonalizable endomorphisms. We can thus diagonalize them *simultaneously*.

Let $E_\alpha \in \mathcal{G}$ be the common eigenvectors:

$$\mathrm{ad}\,(H) \cdot E_\alpha = \alpha(H)\; E_\alpha$$

By construction the application

$$\alpha : H \in \mathcal{H} \to \alpha(H) \in R$$

is a linear form defined over $\mathcal{H}$. That is α belongs to the dual of the Cartan algebra: $\alpha \in \mathcal{H}^*$. These forms are called the *roots* of the Lie algebra. We shall denote their set by Δ. They satisfy a few simple properties: i) if α is a root then so is $-\alpha$, ii) a non zero root is non degenerate (i.e. the eigenspace is of dimension 1), iii) if α is a root and $t \in \mathbb{C}$, $t\alpha$ is not a root, except for $t = \pm 1$.

Let $\{H_i\}$ be a basis of the Cartan subalgebra. Then $\{H_i, E_\alpha\}$ form a basis of the Lie algebra $\mathcal{G}$ on which the Killing form has a very simple structure, namely:

$$(H_i, E_\alpha) = 0, \quad (E_\alpha, E_\beta) = 0 \quad \alpha + \beta \neq 0 \tag{16.4}$$

This is because $(H, [H', E_\alpha]) = \alpha(H')(H, E_\alpha) = ([H, H'], E_\alpha) = 0$, and $([H, E_\alpha], E_\beta) = \alpha(H)(E_\alpha, E_\beta) = (-([H, E_\beta], E_\alpha) = -\beta(H)(E_\alpha, E_\beta)$. As a consequence, the restriction of the Killing form to the Cartan subalgebra is non–degenerate, otherwise the Killing form would be degenerate on the Lie algebra.

It is convenient to introduce the isomorphism between $\mathcal{H}$ and its dual $\mathcal{H}^*$ induced by the Killing form:

$$\alpha \in \mathcal{H}^* \to H_\alpha \in \mathcal{H} \quad \text{with} \quad \alpha(H) = (H_\alpha, H), \quad \forall H \in \mathcal{H}$$

This defines, for any $\alpha \in \mathcal{H}^*$ an element $H_\alpha \in \mathcal{H}$ depending linearly on α. We may then define a bilinear form on $\mathcal{H}^*$ by:

$$(\alpha, \beta) = (H_\alpha, H_\beta) = \alpha(H_\beta), \quad \alpha, \beta \in \mathcal{H}^*$$

This form is non degenerate because the Killing form is non–degenerate on the Cartan subalgebra. Moreover the H_α for $\alpha \in \Delta$ span the Cartan subalgebra.

In the basis $\{H_i, E_\alpha\}$ the Lie bracket reads:

$$\begin{aligned}
[H_i, H_j] &= 0 \\
[H_i, E_\alpha] &= \alpha(H_i) E_\alpha \\
[E_\alpha, E_\beta] &= \begin{cases} (E_\alpha, E_{-\alpha}) H_\alpha, & \text{if } \alpha + \beta = 0; \\ C_{\alpha,\beta} E_{\alpha+\beta}, & \text{if } \alpha + \beta \in \Delta; \\ 0, & \text{if } \alpha + \beta \notin \Delta. \end{cases}
\end{aligned} \tag{16.5}$$

with $C_{\alpha,\beta}$ some structure constants. Here we remark that $[E_\alpha, E_\beta]$ either vanishes or is proportional to $E_{\alpha+\beta}$ if $\alpha + \beta \neq 0$ is a root, because $[H, [E_\alpha, E_\beta]] = [[H, E_\alpha], E_\beta] + [E_\alpha, [H, E_\beta]] = (\alpha(H) + \beta(H))[E_\alpha, E_\beta]$. If however $\alpha + \beta = 0$, this shows that $[E_\alpha, E_{-\alpha}]$ is in the Cartan subalgebra, and we have $(H, [E_\alpha, E_{-\alpha}]) = ([H, E_\alpha], E_{-\alpha}) = \alpha(H)(E_\alpha, E_{-\alpha}) = (H, (E_\alpha, E_{-\alpha}) H_\alpha)$.

For each root α the three generators $H_\alpha, E_\alpha, E_{-\alpha}$ form a $sl(2)$ subalgebra of $\mathcal{G}$. This allows to study the α–chain through β, that is the set of roots of the form $\beta + n\alpha$, using the commutation relations:

$$[H_\alpha, E_{\pm\alpha}] = \pm\alpha(H_\alpha)E_{\pm\alpha}, \quad [E_\alpha, E_{-\alpha}] = (E_\alpha, E_{-\alpha})H_\alpha$$

The vectors $\left(\mathrm{ad}E_{\pm\alpha}\right)^j E_\beta$ for $j \in \mathbb{N}$ are obviously linearly independent root vectors in $\mathcal{G}$, for the roots $\beta \pm j\alpha$, if they don't vanish. They span a representation space for the considered $sl(2)$ and since this representation is of finite dimension, the chain must be of finite length. Let $p \leq 0$ be the minimal index such that $\beta + p\alpha$ is a root, and $q \geq 0$ be the maximal index such that $\beta + q\alpha$ is a root. Let $\beta' = \beta + p\alpha$ and consider the vectors $v_j = \left(\mathrm{ad}E_\alpha\right)^j E_{\beta'}$ for $j \in \mathbb{N}$. By the minimality of p, we have $\mathrm{ad}\, E_{-\alpha} v_0 = 0$. Using this property, we can compute:

$$\mathrm{ad}H_\alpha v_j = \Big(\beta'(H_\alpha) + j\alpha(H_\alpha)\Big)v_j, \quad \mathrm{ad}E_{-\alpha} v_j = -j(E_\alpha, E_{-\alpha})\Big(\beta'(H_\alpha) + (j-1)\frac{\alpha(H_\alpha)}{2}\Big)v_{j-1} \tag{16.6}$$

Since v_{q-p+1} vanishes, but v_{q-p} does not, we have $\beta'(H_\alpha) + (q-p)\alpha(H_\alpha)/2 = 0$ or:

$$2\frac{(\beta, \alpha)}{(\alpha, \alpha)} = -(p+q) \in \mathbb{Z} \tag{16.7}$$

This result allows to show that the Killing form induces a *positive definite* scalar product on the *real* vector space $\sum_{\alpha\in\Delta} \mathbb{R} H_\alpha$. By duality this defines a positive definite scalar product on $\sum_{\alpha\in\Delta} \mathbb{R}\alpha$. Indeed computing the Killing form on the basis of $\mathcal{G}$ provided by the H_i and the $E_{\pm\alpha}$ we have

$$(\alpha, \beta) = (H_\alpha, H_\beta) = \mathrm{Tr}\,(\mathrm{ad}\, H_\alpha \mathrm{ad}\, H_\beta) = \sum_\gamma (\alpha, \gamma)(\beta, \gamma)$$

Taking $\alpha = \beta$ and dividing by $(\alpha, \alpha)^2$ we get:

$$\frac{1}{(\alpha, \alpha)} = \sum_\gamma \left(\frac{(\alpha, \gamma)}{(\alpha, \alpha)}\right)^2$$

so that (α, α) is a rational number. It follows that (α, β) is a rational number, hence is real. Then for any $x = \sum_\alpha x_\alpha \alpha$ with $x_\alpha \in \mathbb{R}$ we have $(x, x) = \sum_\gamma (x, \gamma)^2 \geq 0$ and this vanishes only if $x = 0$.

For later use let us write another consequence of eq.(16.6).

$$[E_{-\alpha}[E_\alpha, E_\beta]] = (E_\alpha, E_{-\alpha})q(1-p)\frac{(\alpha, \alpha)}{2}E_\beta \tag{16.8}$$

Indeed both members are homogeneous in the normalizations of $E_{\pm\alpha}$ and E_β so one can replace $E_\beta = v_{-p}$ with the notations of eq.(16.6), and then use eq.(16.7).

To any root α one can associate a reflexion w_α acting on $\mathcal{H}^*$ by:

$$w_\alpha(x) = x - 2\frac{(\alpha, x)}{(\alpha, \alpha)}\,\alpha$$

These orthogonal reflexions are called Weyl reflexions. They preserve the root system, i.e. if β is a root so is $w_\alpha(\beta)$. This is because, using eq.(16.7), $w_\alpha(\beta) = \beta + (p+q)\alpha$ is in the chain $\beta + p\alpha, \cdots, \beta + q\alpha$. The Weyl group is the discrete group generated by these reflexions.

While roots span $\mathcal{H}^*$, they are not linearly independent in general, and one can choose a subset of them which forms a basis. There exists a subset Π of roots α_i, $i = 1, \cdots r$, such that any other root α can be written $\alpha = \sum_i n_i \alpha_i$ where the n_i are integers all of the same sign. When all n_i are ≥ 0 we say that α is a positive root and otherwise α is called a negative root. The α_i are called *simple roots*. So they are positive roots which cannot be written as the sum of two positive roots. The choice of Π is not unique, but any two such choices are related by a unique transformation of the Weyl group.

To show the existence of a basis of simple roots, choose an hyperplane in $\mathcal{H}^*$ which does not contain any root. Half of the roots are then on one side of this hyperplane, and we call them positive roots. If a positive root can be written as a sum of two positive roots we call it decomposable, otherwise we call it simple. Obviously, any positive root can be written as a linear combination of simple roots with positive integer coefficients. In particular the simple roots span $\mathcal{H}^*$. We show that they are linearly independent. Remark that for two simple roots α and β their difference $(\beta-\alpha)$ is not a root, because if $(\beta-\alpha)$ is a positive root we can decompose $\beta=(\beta-\alpha)+\alpha$ as sum of positive roots, in contradiction with the simplicity of β, while if $(\alpha-\beta)$ is positive, we get similarly a contradiction with the simplicity of α. Hence the simple root condition means that $p=0$ in eq.(16.7) so that:

$$-2\frac{(\alpha,\beta)}{(\alpha,\alpha)}=n\in\mathbb{N}$$

In particular the scalar product (α,β) is negative for α,β simple roots. The α–chain through β consists of the roots of the form $\beta+n\alpha$ for $n=0,1,\cdots,-2(\alpha,\beta)/(\alpha,\alpha)$. Assume now that there is a linear relation between the simple roots that we can write in the form $\sum_s r_s\alpha_s=\sum_{s'} r_{s'}\alpha_{s'}$ with r_s and $r_{s'}$ real and positive, and the set $\{s\}$ disjoint from the set $\{s'\}$. From this equality one gets $(\sum_s r_s\alpha_s)^2=\sum_{ss'} r_s r_{s'}(\alpha_s,\alpha_{s'})$. The left–hand side is obviously positive, while the right–hand side is obviously negative because $s\neq s'$ so $(\alpha_s,\alpha_{s'})\leq 0$. It follows that $r_s=r_{s'}=0$, so the simple roots are linearly independent.

At the extreme opposite of the simple roots are *highest* roots. They are of the form $\theta=\sum_i n_i\alpha_i$ where the n_i are maximal ≥ 0 integers. For a *simple* Lie algebra one can show that the highest root is unique, and that all $n_i>0$.

Let α_i be a set of simple roots. One defines the *Cartan matrix*, which is independent of the choice of basis (since two basis are related by the Weyl group) by:

$$a_{ij}=\frac{2(\alpha_j,\alpha_i)}{(\alpha_i,\alpha_i)}\quad,\quad i,j=1,\cdots,\mathrm{rank}\,\mathcal{G} \tag{16.9}$$

It is such that $a_{ii}=2$ and $a_{ij}\leq 0$, $a_{ij}=0\Rightarrow a_{ji}=0$ for $i\neq j$. Moreover the a_{ij} for $i\neq j$ are negative integers such that $0\leq a_{ij}a_{ji}\leq 4$. This last condition comes from the fact that the scalar product is positive definite on $\mathcal{H}^*$.

The Cartan matrix is non–degenerate: $\det(a)\neq 0$ because $\det(a)$ is proportional to the determinant of the matrix of scalar products of simple roots which are linearly independent.

With the Cartan matrix, we can give a presentation of the Lie algebra $\mathcal{G}$, by generators and relations. For each simple root α_i the elements:

$$h_i=\frac{2}{(\alpha_i,\alpha_i)}H_{\alpha_i},\quad e_i^+=E_{\alpha_i},\quad e_i^-=\frac{1}{(E_{\alpha_i},E_{-\alpha_i})}E_{-\alpha_i}$$

generate a $sl(2)$ subalgebra with standard commutation relations. The Cartan matrix allows to reconstruct the Lie algebra from this set of $sl(2)$ subalgebras. Given a Cartan matrix a_{ij} satisfying the properties mentioned above, we may define $\mathcal{G}$ as the Lie algebra generated by the sets (h_i,e_i^+,e_i^-) with the relations (called the Serre relations):

$$\begin{aligned}
[h_i,h_j] &= 0\\
\left[h_i,e_j^{\pm}\right] &= \pm a_{ij}e_j^{\pm}\\
\left[e_i^+,e_j^-\right] &= \delta_{ij}h_i\\
(\mathrm{ad}\ e_i^{\pm})^{1-a_{ij}}\cdot e_j^{\pm} &= 0\quad\text{for}\quad i\neq j
\end{aligned} \tag{16.10}$$

The last condition is just the condition that the α_i–chain starting at α_j is of length $-a_{ij}$. The elements h_i generate the Cartan subalgebra $\mathcal{H}$. The fact that these relations yield a finite dimensional Lie algebra is the Serre theorem.

Let $\mathcal{N}_\pm$ be the subalgebra generated by the $e_i^\pm$. We have:

$$\mathcal{G} = \mathcal{N}_- \oplus \mathcal{H} \oplus \mathcal{N}_+$$

The subalgebras $\mathcal{B}_\pm = \mathcal{H} \oplus \mathcal{N}_\pm$ are called Borel subalgebras.

The classification of finite dimensional simple Lie algebras is then reduced to the classification of finite dimensional Cartan matrices satisfying the mentioned properties. This leads to four infinite series $A_n = sl(n+1)$, $B_n = so(2n+1)$, $C_n = sp(2n)$, $D_n = so(2n)$ and a few exceptional algebras called E_6, E_7, E_8 and F_4, G_2 (see the References).

A consequence of Serre's theorem is the existence of an involutive automorphism ω of the Lie algebra $\mathcal{G}$, called the Chevalley automorphism. To define it we give its action on the generators:

$$\omega(h_i) = -h_i, \quad \omega(e_i^+) = -e_i^-, \quad \omega(e_i^-) = -e_i^+$$

and check that it preserves the relations. Hence it extends to the whole Lie algebra. For any root α of the Lie algebra one can *choose* $E_{-\alpha} = -\omega(E_\alpha)$. Changing $E_\alpha \to \lambda E_\alpha$ we have $(E_\alpha, E_{-\alpha}) \to \lambda^2 (E_\alpha, E_{-\alpha})$ so that we can always achieve the condition $(E_\alpha, E_{-\alpha}) = 1$. Notice that in general λ will be a complex number.

16.3 Linear representations.

Recall that a linear representation on a vector space V of a finite dimensional Lie algebra $\mathcal{G}$ is an homomorphism ρ from $\mathcal{G}$ to $\mathrm{End}\, V$. We can define the sum and the product of two representations (ρ_1, V_1) and (ρ_2, V_2). The sum is the representation on the direct sum $V_1 \oplus V_2$ such that $\rho_{V_1 \oplus V_2}$ maps elements of $\mathcal{G}$ into block diagonal endomorphisms whose restrictions to $V_{1,2}$ coincide with their images under $\rho_{1,2}$, in other words $\rho_{V_1 \oplus V_2}$ is block diagonal. The product is the representation on the tensor product $V_1 \otimes V_2$ with $\rho_{V_1 \otimes V_2}(X) = \rho_1(X) \otimes 1 + 1 \otimes \rho_2(X)$ for any element $X \in \mathcal{G}$. A representation on a vector space V is said to be indecomposable if it cannot be decomposed into the sum of subrepresentations. Note that for a general algebra $\mathcal{A}$ the sum of two representations is a representation, but the tensor product is not. The Lie algebra case appears as very special, and this is due to the existence of the algebra homomorphism, called the coproduct:

$$\Delta : \mathcal{U}(\mathcal{G}) \to \mathcal{U}(\mathcal{G}) \otimes \mathcal{U}(\mathcal{G}), \quad \Delta(X) = X \otimes 1 + 1 \otimes X \tag{16.11}$$

The elements of the Cartan subalgebra are represented by a family of commuting diagonalizable endomorphisms. They can thus be simultaneously diagonalized. Let $|\lambda >$ be an eigenvector, and $\lambda(H)$ be the corresponding eigenvalues, which depend linearly on H. The various λ's are linear forms acting on the Cartan subalgebra $\mathcal{H}$, i.e. $\lambda \in \mathcal{H}^*$, and are called weights. The weights $\lambda(H)$ may have multiplicities, so we denote $|\lambda_a >$ the weight vectors with the same weight λ.

$$H|\lambda_a\rangle = \lambda(H)|\lambda_a\rangle$$

The weight vectors $|\lambda_a\rangle$ form a basis of the representation space V. Their number, degeneracy included, is the dimension of the representation. The representation space V contains representations of the $sl(2)$ subalgebras generated by $(H_\alpha, E_\alpha, E_{-\alpha})$ for any root α. From the knowledge of finite–dimensional representations of $sl(2)$ we get the basic integrality condition:

$$2\frac{(\lambda, \alpha)}{(\alpha, \alpha)} \in \mathbb{Z} \quad \text{for all} \quad \alpha \in \Delta$$

Note that the weight system of a representation is invariant under the action of the Weyl group. In other words, if λ is a weight, so is $w_\alpha(\lambda)$ for any root α.

The difference of two weights of a given irreducible representation always belongs to the root system. We may thus introduce an order between weights of a representation by $\lambda_1 > \lambda_2$ iff $\lambda_1 - \lambda_2 > 0$. Any finite dimensional representation possesses a highest weight since its number of weights is finite. This vector is unique for irreducible representations. Let Λ be this *highest weight*, which is thus non degenerate. The corresponding eigenvector $|\Lambda\rangle$ is called the highest weight vector of the representation. It is such that:

$$H|\Lambda\rangle = \Lambda(H)|\Lambda\rangle, \quad E_{\alpha_i}|\Lambda\rangle = 0$$

for $H \in \mathcal{H}$ and α_i any simple positive root. This follows from the fact that since Λ is a highest weight $\Lambda + \alpha_i$ is not a weight.

Given a representation with highest weight Λ one defines its Dynkin indices δ_i by:

$$\delta_i = 2\frac{(\Lambda, \alpha_i)}{(\alpha_i, \alpha_i)} \in \mathbb{N}$$

with α_i the simple roots. The proof that this is a positive integer is the same as in the case of the adjoint representation. By definition, the *fundamental weights* Λ_j are the highest weights with Dynkin indices δ_{ij}. They are specified by:

$$2\frac{(\Lambda_j, \alpha_i)}{(\alpha_i, \alpha_i)} = \delta_{ij}$$

The number of fundamental weights is equal to the rank of $\mathcal{G}$, and there exists representations with highest weights the fundamental weights (called fundamental repesentations). Any highest weight Λ of a finite dimensional representation may be decomposed on the fundamental weights: $\Lambda = \sum_j \delta_j \Lambda_j$ with δ_j the Dynkin indices. More generally, the weight lattice is the set of $\lambda \in \mathcal{H}^*$ such that $(\lambda, \alpha) \in \mathbb{Z}$ for any root α. Any weight of any representation is on the weight lattice. As a $\mathbb{Z}$–module, the weight lattice has a basis provided by the fundamental weights, such that $(\lambda_i, \alpha_j) = \delta_{ij}$, for any simple root α_j.

The highest weight representations may also be viewed as quotient of Verma modules. The Verma module associated to a highest weight vector $|\Lambda\rangle$ is the space $(U(\mathcal{N}_-)|\Lambda\rangle)$ with $U(\mathcal{N}_-)$ the envelopping algebra of $\mathcal{N}_-$. Then the irreducible representation with highest weight vector Λ is isomorphic to the quotient:

$$\left(U(\mathcal{N}_-)|\Lambda\rangle\right)/\mathcal{M}_\Lambda$$

where $\mathcal{M}_\Lambda$ the maximal submodule of $U(\mathcal{N}_-)|\Lambda\rangle$, which is shown to exist, to be unique (by maximality). The above quotient is finite–dimensional when Λ belongs to the weight lattice. The Verma module construction shows that any weight of the weight lattice is conjugated by the Weyl group to the highest weight of some representation. The root themselves generate a lattice called the root lattice. It is a sublattice of the weight lattice.

The Casimir operator C is the following operator, quadratic in the Lie algebra generators (t_a) forming a basis of the Lie algebra, hence living in the universal enveloping algebra of $\mathcal{G}$:

$$C = \sum_{a,b} t_a K^{ab} t_b$$

with K^{ab} the matrix inverse of the Killing form K_{ab}. Its main property is that it is in the center of the enveloping algebra, so that, in any given representation, the Casimir operator commutes with the endomorphisms representing the elements of the Lie algebra. It thus acts proportionally to the identity on irreducible representations. If Λ is the highest weight of the representation its value is:

$$C(\Lambda) = (\Lambda, \Lambda + \rho)$$

with ρ the Weyl vector equals to the sum of the fundamental weights:

$$\rho = \sum_j \Lambda_j$$

We frequently meet the tensor Casimir operator living in $\mathcal{G} \otimes \mathcal{G}$ given by $C_{12} = \sum_{a,b} K^{ab} t_a \otimes t_b$. Remark that we have, usinq eq.(16.11):

$$C_{12} = \frac{1}{2}(\Delta C - C \otimes 1 - 1 \otimes C)$$

The main property of the tensor Casimir is that $[C_{12}, \Delta(X)] = 0$ for any $X \in \mathcal{G}$.

16.4 Real Lie algebras.

Up to now, we considered complex Lie algebras. Examples are provided by complexification of real Lie algebras. More precisely let $\mathcal{G}$ be a real Lie algebra. This means that we have a basis X_a of the Lie algebra such that the structure constants are real, and we consider the linear combinations of the X_a with real coefficients. Its complexification $\mathcal{G}_{\mathbb{C}}$ is the set of elements $Z = X + iY$ with $X, Y \in \mathcal{G}$. On $\mathcal{G}_{\mathbb{C}}$ we define a conjugation $c : X + iY \to X - iY$. We have

$$c^2 = 1, \quad [c(Z_1), c(Z_2)] = c([Z_1, Z_2]), \quad c(\lambda Z) = \overline{\lambda} c(Z)$$

Conversely, given any such conjugation c we can write $\mathcal{G}_{\mathbb{C}} = \mathcal{G}_+ \oplus \mathcal{G}_-$ where $c|_{\mathcal{G}_\pm} = \pm 1$ and $\mathcal{G}_{\mathbb{C}}$ can be viewed as the complexification of $\mathcal{G}_+$ which is a real Lie algebra.

Different real Lie algebras may have the same complexification. For example $sl(2, \mathbb{C})$ is the common complexification of the two real Lie algebras $sl(2, \mathbb{R})$ and $su(2)$. The algebra $sl(2, \mathbb{R})$ is the Lie algebra of 2×2 traceless real matrices, with basis:

$$H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad E_+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad E_- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

and commutation relations $[H, E_\pm] = \pm 2 E_\pm$ and $[E_+, E_-] = H$. On the other hand $su(2)$ is the Lie algebra of antihermitean traceless 2×2 matrices, i.e. linear combinations with *real* coefficients of the matrices $t_k = i\sigma_k$, where σ_k are the Pauli matrices:

$$\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

The commutation relations are $[t_i, t_j] = -2\epsilon_{ijk} t_k$, and we see that the structure constants of $su(2)$ are *real*. The Lie algebras $sl(2, \mathbb{R})$ and $su(2)$ are referred to as the non–compact and compact real forms of $sl(2, \mathbb{C})$ respectively. Notice that although the algebra $su(2)$ is real, the matrices representing it have complex entries.

It is an important problem to classify the real forms of a given complex Lie algebra. This amounts to classifying the conjugations c. For this purpose remark that one can choose a basis of the Lie algebra $\mathcal{G}_{\mathbb{C}}$ such that all the structure constants are real. Indeed, choose the basis $E_{\pm\alpha}$, H_α such that $\omega(E_\alpha) = -E_\alpha$ (where ω is the Chevalley automorphism) and $(E_\alpha, E_{-\alpha}) = 1$. Then $[E_\alpha, E_{-\alpha}] = H_\alpha$, and $[H_\alpha, E_{\pm\beta}] = \pm\beta(H_\alpha) E_{\pm\beta}$ where $\beta(H_\alpha)$ are real. Setting $[E_\alpha, E_\beta] = C_{\alpha,\beta} E_{\alpha+\beta}$ and applying ω, we get the relation:

$$C_{-\alpha,-\beta} = -C_{\alpha,\beta} \tag{16.12}$$

To show that these structure constants are real, we compute:

$$[[E_\alpha, E_\beta], [E_{-\alpha}, E_{-\beta}]] = -C^2_{\alpha,\beta} H_{\alpha+\beta} = -C^2_{\alpha,\beta}(H_\alpha + H_\beta)$$

On the other hand, using the Jacobi identity and eq.(16.8) we get:

$$\begin{aligned}[][E_\alpha, E_\beta],[E_{-\alpha}, E_{-\beta}]] &= [E_{-\beta},[E_{-\alpha},[E_\alpha, E_\beta]]] + [E_{-\alpha},[E_{-\beta},[E_\beta, E_\alpha]]] \\ &= -q(1-p)\frac{(\alpha,\alpha)}{2}H_\alpha - q'(1-p')\frac{(\beta,\beta)}{2}H_\beta\end{aligned}$$

Here p' and q' refer to the β–chain through α. If $C_{\alpha,\beta} \neq 0$, i.e. $\alpha+\beta$ is a root, H_α and H_β are linearly independent, so identifying the coefficients we get:

$$C_{\alpha,\beta}^2 = q(1-p)\frac{(\alpha,\alpha)}{2}$$

and a similar formula with p' and q'. Recalling that $p < 0$ and $(\alpha,\alpha) > 0$ we see that $C_{\alpha,\beta}$ is real. The basis of the Lie algebra that we have constructed is called a Weyl basis.

The real Lie algebra $\mathcal{G}'$ spanned over $\mathbb{R}$ by $E_{\pm\alpha}$ and H_α is the analogue of the non–compact $sl(2,\mathbb{R})$ in the general case. We obviously have $\mathcal{G}_{\mathbb{C}} = \mathcal{G}' + i\mathcal{G}'$ and the real form is selected by the conjugation $c'(X+iY) = X - iY$.

Thanks to a theorem of H. Weyl, for any semi–simple complex Lie algebra there exists a real form which is the Lie algebra of a compact Lie group. The Lie algebra of a semi–simple compact Lie group is characterized by the fact that its Killing form is negative definite. Indeed if G is a compact Lie group, choose any positive definite scalar product on $\mathcal{G}$ its Lie algebra. Since G is compact, one can use the Haar integral on G to take the average of this bilinear form and obtain a positive definite *invariant* scalar product on $\mathcal{G}$. This means that the Lie group G acts by orthogonal matrices in the adjoint representation (for this scalar product). Hence in an orthonormal basis the matrices ad_X are antisymmetric and the Killing form $(X,X) = -\sum_{ij}(\mathrm{ad}_X)^2_{ij} \leq 0$ vanishes only when $\mathrm{ad}_X = 0$, i.e., when X is in the center of $\mathcal{G}$. We see that $\mathcal{G}$ decomposes as the orthogonal sum of its center and a semi–simple algebra $[\mathcal{G},\mathcal{G}]$ on which the Killing form is negative definite.

Conversely, starting from the Weyl basis, one can construct the compact form as follows: consider the generators

$$X_\alpha = (E_\alpha - E_{-\alpha}), \quad Y_\alpha = i(E_\alpha + E_{-\alpha}), \quad Z_\alpha = iH_\alpha \tag{16.13}$$

The real vector space spanned by these elements is a real Lie algebra thanks to eq.(16.12). Moreover it has a negative definite Killing form. To show this recall the ortogonality relations (16.4) so it is sufficient to look at each subspace $(X_\alpha, Y_\alpha, Z_\alpha)$ independently, where the check is simple, thereby proving the Weyl theorem. For instance $(X_\alpha, X_\alpha) = (E_\alpha - E_{-\alpha}, E_\alpha - E_{-\alpha}) = -2(E_\alpha, E_{-\alpha}) = -2$ and $(X_\alpha, Y_\alpha) = (E_\alpha - E_{-\alpha}, i(E_\alpha + E_{-\alpha})) = 0$, and so on. This compact real form corresponds to the conjugation:

$$c(E_{\pm\alpha}) = -E_{\mp\alpha}, \quad c(H_\alpha) = -H_\alpha \tag{16.14}$$

This conjugation selects the analog of $su(2)$ in the $sl(2)$ case.

In general the representations of a (real) compact Lie group G may be complex. Choosing on the representation space V an arbitrary sesquilinear form and averaging it by the group G one gets an invariant sesquilinear form, i.e. $< gv, gw > = < v, w >$. Hence all elements of G are represented by unitary matrices and elements of $\mathcal{G}$ by antihermitean matrices, hence the generators eq.(16.13) are such that $X_\alpha^+ = -X_\alpha$, $Y_\alpha^+ = -Y_\alpha$ and $Z_\alpha^+ = -Z_\alpha$. This also reads $E_\alpha^+ = E_{-\alpha}$, $H_\alpha^+ = H_\alpha$, or more abstractly for any X in the complexified Lie algebra $X^+ = -c(X)$.

In particular any maximal abelian subalgebra of $\mathcal{G}$ is a Cartan subalgebra, because antihermitean matrices are always diagonalizable with eigenvalues pure imaginary. Its image by the exponential map is the Weyl torus. One can choose a basis H_j of the Cartan algebra such that any element of the torus is of the form $h = \exp\left(\sum_j \theta_j H_j\right)$ with $\exp(2\pi H_j) = 1$ for all j. If $|\lambda>$ is a weight vector in V, we have $h|\lambda> = \chi_\lambda(h)|\lambda>$ where $\chi_\lambda(h)$ is a character, i.e. $\chi_\lambda(hh') = \chi_\lambda(h)\chi_\lambda(h')$. So we have $\chi_\lambda(h) = \exp\left(\sum_j \theta_j \lambda(H_j)\right)$. The condition $\exp(2\pi H_j) = 1$ gives $\lambda(H_j) \in i\mathbb{Z}$ for all j. This defines a lattice in $\mathcal{H}^*$ called the weight lattice of the *group* G, which is a sublattice of the weight lattice of the Lie algebra. Moreover, since the adjoint representation of G is well defined, the root lattice is a sublattice of the weight lattice of G.

In general the weight lattice of G is a sublattice of the weight lattice of $\mathcal{G}$, and is equal to it only when G is simply connected. This is because any positive weight of the weight lattice of $\mathcal{G}$ is the highest weight of some representation of $\mathcal{G}$ which can then be lifted to G. Hence, for a compact semi–simple simply connected Lie group, the weight lattices of the group and its Lie algebra are the same. Remark that $2\lambda(H_\alpha)/(\alpha, \alpha) \in \mathbb{Z}$ for any weight λ of $\mathcal{G}$, so in this case the elements H_j such that $\exp(2\pi H_j) = 1$ are of the form $H_j = i2H_{\alpha_j}/(\alpha_j, \alpha_j)$ where α_j are the simple roots.

When G is a semi–simple compact connected Lie group, its center is a finite group contained in all maximal tori of G. Assuming that the center of G is trivial, its weight lattice is equal to the root lattice. This is because in this case the adjoint representation is a faithful representation of G and so generates all representations of G, taking tensor products. It follows that the root lattice generates the weight lattice.

This allows to describe the various compact Lie groups G with Lie algebra $\mathcal{G}$. Starting from the universal cover of any one of them, which we call G, (and can be shown to be compact), having center Z the other Lie groups are of the form G/D where D is any subgroup of the discrete Abelian group Z. They have center Z/D isomorphic to the quotient of their weight lattice by the root lattice. Moreover their first homotopy group is isomorphic to the quotient of the weight lattice of $\mathcal{G}$ by their own weight lattice. We see that global topological properties of compact Lie groups are remarkably encoded in the structure of their tangent space at the unit element.

The classification of real forms of complex Lie algebras is also the basis of the study of symmetric spaces. We have obtained two conjugations c' and c which select non–compact and compact real forms of $\mathcal{G}_{\mathbb{C}}$. Remark that c' commutes with the conjugation c defined in eq.(16.14). Hence we can diagonalize c in the eigenspaces of c' and conversely. This yields the decompositions, called Cartan decompositions of the real Lie algebras $\mathcal{G}$ and $\mathcal{G}'$:

$$\mathcal{G} = t \oplus p, \quad \mathcal{G}' = t \oplus ip$$

where t is the Lie algebra $\mathcal{G} \cap \mathcal{G}'$. Indeed t is generated by the $(E_\alpha - E_{-\alpha})$, p is spanned by the $i(E_\alpha + E_{-\alpha})$ and the iH_α. Moreover we have the relations:

$$[t, t] \subset t, \quad [t, p] \subset p, \quad [p, p] \subset t$$

and similarly with $p \to ip$. For example in the case of $sl(n)$ this corresponds to the decomposition into symmetric and antisymmetric matrices. At the Lie group level to the algebra t corresponds a compact group K, and to the Lie algebras $\mathcal{G}$ and $\mathcal{G}'$ correspond appropriate Lie groups G and G' respectively compact and non–compact. One gets symmetric spaces G/K and G'/K of the compact and non–compact type respectively. This is the situation we have encountered in Chapter [6]. Many more conjugations exist, but we will not enter this subject.

16.5 Affine Kac–Moody algebras.

We start from a finite–dimensional simple Lie algebra $\mathcal{G}$ and construct the loop algebra which consists in formal Laurent polynomials $\widetilde{\mathcal{G}} = \mathcal{G} \otimes \mathbb{C}[\lambda, \lambda^{-1}]$ with Lie bracket:

$$[X \otimes \lambda^n, Y \otimes \lambda^m] = [X, Y] \otimes \lambda^{n+m}$$

The affine Kac–Moody algebra $\widehat{\mathcal{G}}$ is the central extension of the loop algebra $\widetilde{\mathcal{G}}$ by a central element denoted by K (this means that the formal element K commutes with all other elements). It is convenient to further extend this algebra by including the derivation $d = \lambda\partial_\lambda$. Thus, the affine Kac–Moody algebra is:

$$\widehat{\mathcal{G}} = \widetilde{\mathcal{G}} \oplus \mathbb{C}K \oplus \mathbb{C}d$$

and the Lie brackets are defined by (with (X, Y) the Killing form on $\mathcal{G}$):

$$\begin{aligned} [X \otimes \lambda^n, Y \otimes \lambda^m] &= [X, Y] \otimes \lambda^{n+m} + \tfrac{1}{2} n\delta_{m+n,0}(X, Y)K \\ [d, X \otimes \lambda^n] &= n\ X \otimes \lambda^n \\ [K, X \otimes \lambda^n] &= [K, d] = 0 \end{aligned} \tag{16.15}$$

Remark that the coefficient $\omega(X \otimes \lambda^n, Y \otimes \lambda^m) \equiv \frac{1}{2} n\delta_{m+n,0}(X, Y)$ of the central element K appearing in eq.(16.15) satisfies the cocycle condition $\omega([X, Y], Z) + \omega([Z, X], Y) + \omega([Y, Z], X) = 0$ ensuring that the Jacobi identity is satisfied.

An invariant bilinear form on the Kac–Moody algebra is given by:

$$(X \otimes \lambda^n, Y \otimes \lambda^m) = (X, Y)\delta_{n+m,0}, \quad (K, K) = (d, d) = 0, \quad (K, d) = 1$$

and $(K, X \otimes \lambda^n) = (d, X \otimes \lambda^n) = 0$. The fact that this form is invariant is easy to check by direct computation.

Alternatively, denoting a general element of $\widehat{\mathcal{G}}$ by $\widehat{X} = \widetilde{X}(\lambda) + X_K K + X_d d$ we have in the affine Kac–Moody algebra:

$$\begin{aligned} \left[\widehat{X}, \widehat{Y}\right] &= \left[\widetilde{X}, \widetilde{Y}\right] + \tfrac{1}{2} \oint \frac{d\lambda}{2i\pi} (\partial_\lambda \widetilde{X}(\lambda), \widetilde{Y}(\lambda))\ K \\ \left[K, \widehat{X}\right] &= [K, d] = 0 \\ \left[d, \widehat{X}\right] &= \lambda \frac{d}{d\lambda} \widetilde{X}(\lambda) \end{aligned} \tag{16.16}$$

Let α_i be the simple roots of the finite dimensional simple Lie algebra $\mathcal{G}$ and θ its highest root. The affine Kac-Moody algebra $\mathcal{G} \otimes \mathbb{C}(\lambda, \lambda^{-1}) \oplus \mathbb{C}K$ is generated by the following elements:

$$(E_{\alpha_i}, H_{\alpha_i}, E_{-\alpha_i}), \quad i, j = 1, \cdots, \mathrm{rank}\,\mathcal{G}, \quad (E_{-\theta} \otimes \lambda, K - H_\theta, E_\theta \otimes \lambda^{-1}) \tag{16.17}$$

Each triplet form a sl_2 subalgebra. These triplets are associated to the simple roots of the affine Kac-Moody algebra. The derivation $d = \lambda\partial_\lambda$ is not in the algebra generated by these elements and has to be added by hand. The λ dependence in this presentation corresponds to what is called the homogeneous gradation. The gradation is defined by the degree in λ, which is counted by d.

A slight modification of this construction allows to define the twisted affine Kac–Moody algebras. Assume that $\mathcal{G}$ has an automorphism τ of order N, i.e. $\tau^N = 1$. Let $\zeta = e^{2i\pi/N}$. One extends τ to an automorphism $\widehat{\tau}$ of the Kac–Moody algebra by setting:

$$\widehat{\tau}(X \otimes \lambda^n) = \tau(X) \otimes (\zeta\lambda)^n, \quad \widehat{\tau}(K) = K, \quad \widehat{\tau}(d) = d \tag{16.18}$$

Since $\widehat{\tau}$ is an automorphism, the set of its fixed points is a Lie algebra, which is called the twisted affine Kac–Moody algebra associated to τ, and denoted $\widehat{\mathcal{G}}_\tau$.

If the automorphism τ is an inner automorphism, $\widehat{\mathcal{G}}_\tau$ is isomorphic to the untwisted algebra $\widehat{\mathcal{G}}$. If however τ is not an inner automorphism one gets an essentially different algebra. It is known that this situation occurs only when $N = 2$ or $N = 3$ and only for particular simple Lie algebras, modulo quotienting by inner automorphisms.

Let us illustrate the use of an inner automorphism to obtain the presentation of the affine Kac–Moody algebra in the principal gradation. Consider the Weyl vector $\rho = \sum_i \Lambda_i$ where the Λ_i are the fundamental weights of $\mathcal{G}$. So we have $(\rho, \alpha_i) = 1$ for any simple root α_i of $\mathcal{G}$. Moreover, if θ is the highest root of $\mathcal{G}$ we define the dual Coxeter number h^* by $(\rho, \theta) = h^* - 1$. Let τ be the inner automorphism of $\mathcal{G}$:

$$\tau(X) = e^{-\frac{2i\pi}{h^*}H_\rho} X e^{\frac{2i\pi}{h^*}H_\rho}, \quad \tau(H) = H, \quad \tau(E_\alpha) = e^{-\frac{2i\pi}{h^*}(\rho,\alpha)} E_\alpha$$

and extend it to an automorphism $\widehat{\tau}$ of the affine algebra as in eq.(16.18) with $\zeta = e^{\frac{2i\pi}{h^*}}$. The algebra of its fixed points is isomorphic to our affine algebra, and is linearly generated by the $H \otimes \lambda^{mh^*}$ and the $E_\alpha \otimes \lambda^{mh^*+(\rho,\alpha)}$. It follows that the elements of degree $\pm 1, 0$ are:

$$(E_{\alpha_i} \otimes \lambda, H_{\alpha_i}, E_{-\alpha_i} \otimes \lambda^{-1}), \quad i,j = 1, \cdots, \mathrm{rank}\, \mathcal{G}, \quad (E_{-\theta} \otimes \lambda, K - H_\theta, E_\theta \otimes \lambda^{-1}) \tag{16.19}$$

These elements generate the whole fixed point algebra. This presentation differs from the presentation in eq.(16.17) by the way in which the degrees in λ are distributed.

Affine Kac-Moody algebra may also be presented by generators and relations using Cartan matrices and their associated set of generators. Namely an affine Cartan matrix is a finite dimensional matrix a_{ij} such that: $a_{ii} = 2$; $a_{ij} \leq 0$; $a_{ij} = 0 \Rightarrow a_{ji} = 0$; a_{ij} being negative integers for $i \neq j$, and such that the dimension of its kernel is one. Note that the only difference with the Cartan matrix of a semi–simple algebra is that its determinant vanishes. A classification of such Cartan matrices may be found in the References, where it is shown that irreducible ones yield exactly the standard and twisted algebras constructed above, for any simple Lie algebra $\mathcal{G}$. As for finite dimensional algebras, the affine Kac-Moody algebra with Cartan matrix a_{ij} is defined as the Lie algebra generated by the elements (e_i^+, e_i^-, h_i) with Serre relations:

$$\begin{aligned} [h_i, h_j] &= 0 \\ \left[h_i, e_j^\pm\right] &= \pm a_{ij} e_j^\pm \\ \left[e_i^+, e_j^-\right] &= \delta_{ij} h_i \\ (\mathrm{ad}\ e_i^\pm)^{1-a_{ij}} \cdot e_j^\pm &= 0 \quad \text{for} \quad i \neq j \end{aligned} \tag{16.20}$$

However one then gets an infinite algebra because $\det(a) = 0$. The elements h_i generate the Cartan subalgebra $\mathcal{H}$. Let n_j be such that $\sum_i n_i a_{ij} = 0$. Since by hypothesis the kernel of a_{ij} is one dimenional, such coefficients are unique up to a multiplicative constant. It is usually convenient to normalize them such that $\sum_j n_j = h^*$ with h^* the dual Coxeter number; the coefficient n_j are then all non-negative integers. By construction, the element K,

$$K = \sum_i n_i h_i$$

is central. The derivation d is not an element of the algebra generated by the (e_i^+, e_i^-, h_i). It has to be added by hand. Its commutation relations depends on the gradation one chooses. For example the principal gradation obtained in eq.(16.19) corresponds to the choice:

$$\left[d, e_i^+\right] = 1, \quad [d, h_i] = 0, \quad \left[d, e_i^-\right] = -1$$

In particular, the rank of the (untwisted) affine Kac–Moody algebra $\widehat{\mathcal{G}}$ is $(1 + \mathrm{rank}\, \mathcal{G})$ if one does not include the derivation in its definition and is $(2 + \mathrm{rank}\, \mathcal{G})$ if one does include it.

As for finite dimensional Lie algebras, one has the decomposition:

$$\widehat{\mathcal{G}} = \mathcal{N}_- \oplus \mathcal{H} \oplus \mathcal{N}_+$$

with $\mathcal{N}_\pm$ the subalgebra generated by the $e_i^\pm$ and $\mathcal{H}$ the Cartan subalgebra.

As for finite dimensional Lie algebras one may introduce roots, which are points in the dual $\mathcal{H}^*$ of the Cartan subalgebra $\mathcal{H}$, and systems of simple roots. However in contrast to the finite dimensional case, the number of root is infinite and roots may have multiplicities.

Weights are elements of $\mathcal{H}^*$. The fundamental weight vectors Λ_j are such that

$$\Lambda_j(h_i) = \delta_{ij} \tag{16.21}$$

By definition *integrable highest vectors* Λ are integer linear combinaisons of the fundamental weights: $\Lambda = \sum_j \delta_j \Lambda_j$ with δ_j integers. The coefficients δ_j are called Dynkin indices.

Unitary highest weight representations may be defined as for finite dimensional Lie algebras. Let Λ be an integrable highest weight vector and $|\Lambda\rangle$ be the corresponding highest weight vector. By definition one assumes that

$$h_i|\Lambda\rangle = \Lambda(h_i)|\Lambda\rangle \quad , \quad e_j^+|\Lambda\rangle = 0$$

The highest weight representation $V(\Lambda)$ with highest weight vector Λ is then defined as:

$$V(\Lambda) = \left(U(\mathcal{N}_-)|\Lambda\rangle\right)/\mathcal{M}_\Lambda$$

with the $U(\mathcal{N}_-)$ the universal enveloping algebra of $\mathcal{N}_-$ and $\mathcal{M}_\Lambda$ the maximal submodule of $U(\mathcal{N}_-)|\Lambda\rangle$. More concretely, vectors in $V(\Lambda)$ is are obtained by multiple action of the generators e_i^- on the highest weight vector $|\Lambda\rangle$. Remark that if $\Lambda = \sum_i \delta_i \Lambda_i$ then the central element $K = \sum_i n_i h_i$ acts on $V(\lambda)$ as the $\mathbb{C}$-number $K = \sum_i n_i \delta_i$. This number is called the level of the representation. Remark that the adjoint representation is not a highest weight representation.

Let us present the affine Kac-Moody algebra $\widehat{sl(2)}$ in more details. Let E_+, E_-, H be the three generators of the Lie algebra $sl(2)$

$$[H, E_\pm] = \pm 2E_\pm, \quad [E_+, E_-] = H$$

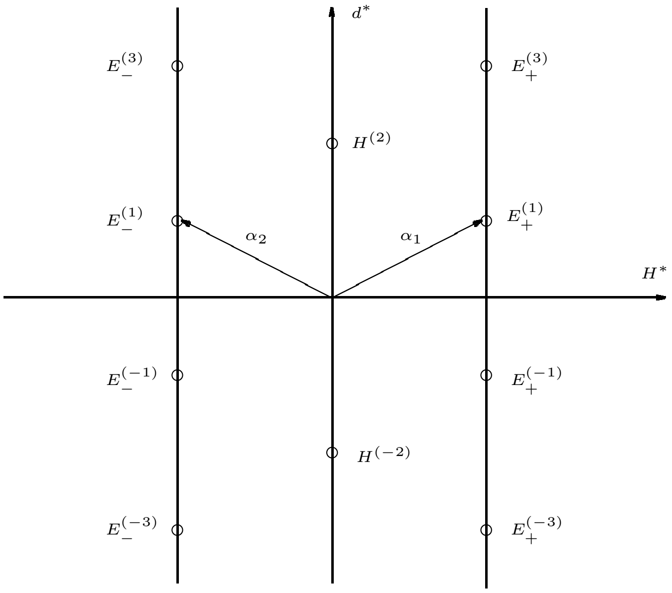
We normalize the Killing form on $sl(2)$ by $(H, H) = 2$, $(E_+, E_-) = 1$. The loop algebra $\widetilde{sl}(2)$ is the Lie algebra of traceless 2×2 matrices with entries Laurent polynomials in λ: $\widetilde{sl}(2) = sl(2) \otimes \mathbb{C}(\lambda, \lambda^{-1})$. The affine Lie algebra $\widehat{sl(2)}$ is the central extension of $\widetilde{sl}(2)$: $\widehat{sl(2)} = \widetilde{sl}(2) \oplus \mathbb{C}K \oplus \mathbb{C}d$, with K the central element and d the derivation $d = \lambda \frac{\partial}{\partial \lambda}$. Let us write the decomposition: $\widehat{sl(2)} = \widehat{\mathcal{N}}_- \oplus \widehat{\mathcal{H}} \oplus \widehat{\mathcal{N}}_+$. First one can choose, as in eq.(16.17) the simple root vectors $E_{\alpha_1} = E_+$, $E_{\alpha_2} = \lambda E_-$. Together with the Cartan algebra generators H, K, d and $E_{-\alpha_1} = E_-$, $E_{-\alpha_2} = \lambda^{-1} E_+$ they generate the whole algebra. The simple root vectors are of degree 0 and 1. This is called the homogeneous gradation. It is more convenient to define a gradation such that both simple root vectors have degree 1, the principal gradation. To do that we choose simple root vectors $E_{\alpha_1} = \lambda E_+$, $E_{\alpha_2} = \lambda E_-$. Together with the Cartan algebra generators H, K, d and $E_{-\alpha_1} = \lambda^{-1} E_-$, $E_{-\alpha_2} = \lambda^{-1} E_+$ they generate the algebra spanned by:

$$\begin{aligned}
\widehat{\mathcal{H}} &= \{H, d, K\} \\
\widehat{\mathcal{N}}_+ &= \{E_+^{(2n-1)} = E_+ \otimes \lambda^{2n-1}, E_-^{(2n-1)} = E_- \otimes \lambda^{2n-1}, H^{(2n)} = H \otimes \lambda^{2n}, n > 0\} \\
\widehat{\mathcal{N}}_- &= \{E_+^{(2n+1)} = E_+ \otimes \lambda^{2n+1}, E_-^{(2n+1)} = E_- \otimes \lambda^{2n+1}, H^{(2n)} = H \otimes \lambda^{2n}, n < 0\}
\end{aligned}$$

We can exhibit the isomorphism between the affine algebra $\widehat{sl(2)}$ in the homogeneous gradation, and this presentation which corresponds to the principal gradation. Indeed first replace the parameter λ by λ^2, in the homogeneous presentation. Then the simple root vectors of the homogeneous gradation are E_+ and $\lambda^2 E_-$. Then perform a conjugation by $\exp(\log(\lambda)H/2)$. This conjugation sends E_+ to λE_+ and E_- to $\lambda^{-1} E_-$ and extends to an isomorphism of the two algebras. Note that $d \to d - \frac{1}{2}H$. In the principal gradation, the commutation relations read:

$$\begin{aligned}
\left[H^{(r)}, H^{(s)}\right] &= r\,\delta_{r+s,0} K \\
\left[H^{(r)}, E_\pm^{(s)}\right] &= \pm 2 E_\pm^{(r+s)} \\
\left[E_+^{(r)}, E_-^{(s)}\right] &= H^{(r+s)} + \tfrac{1}{2} r\,\delta_{r+s,0} K
\end{aligned} \tag{16.22}$$

In the notation of eq.(16.20) we have $h_1 = H + \frac{1}{2}K$ and $h_2 = -H + \frac{1}{2}K$. Let H^*, K^*, d^* the dual basis of the basis H, K, d of the Cartan algebra. To the root vectors $E_{\pm}^{(r)}$ are associated the roots $\pm 2H^* + rd^*$, and to the root vectors $H^{(r)}$ are associated the roots rd^*. Let us draw the root diagram of $\widehat{sl}(2)$.



The affine $\widehat{sl(2)}$ algebra possesses two fundamental highest weights, denoted Λ^+ and Λ^-. They are characterized by eq.(16.21). Expanding on the dual basis H^*, K^* of H, K one gets $\Lambda^{\pm} = \pm\frac{1}{2}H^* + K^*$, or equivalently:

$$\Lambda^{\pm}(H) = \pm\tfrac{1}{2} \quad ; \quad \Lambda^{\pm}(K) = 1 \quad ; \quad \Lambda^{\pm}(d) = 0$$

Note that the levels of these fundamental representations are equal to one, i.e. K takes the value 1 on them.

16.6 Vertex operator representations.

We now recall the vertex operator construction of the level one representations of $\widehat{sl(2)}$, in the principal gradation. We introduce oscillators p_n for n *odd*, such that

$$[p_m, p_n] = m\delta_{n+m,0}$$

Note that the choice n odd ensures that there is no center in this algebra. Assume $p_n^+ = p_{-n}$. The vacuum $|0>$ is defined by $p_n|0>= 0$ for $n > 0$. Its dual $<0|$ is defined by by $<0|p_n = 0$ for $n < 0$, and the normalization condition $<0|0>= 1$. This allows to compute vacuum expectation values $<0|\mathcal{O}|0>$ of any operator $\mathcal{O}$. The representation space is the Fock space generated by the p_{-n} acting on the vacuum. We define the normal ordering on monomials of the p_n by putting p_n, $n > 0$ to the right. We denote it by : :. Define the operators acting on the Fock space:

$$Q(z) = -i\sqrt{2}\sum_{n \text{ odd}} p_{-n}\frac{z^n}{n}$$

We have

$$<0|Q(z_1)Q(z_2)|0>= \log\left(\frac{z_1+z_2}{z_1-z_2}\right) \qquad |z_1| > |z_2| \tag{16.23}$$

This is because

$$< 0|Q(z_1)Q(z_2)|0 >= -2 \sum_{n_1 n_2} < 0|p_{-n_1}p_{-n_2}|0 > \frac{z_1^{n_1} z_2^{n_2}}{n_1 n_2}$$

Here n_1, n_2 are odd integers, but the properties of the vacuum select $n_1 < 0$ and $n_2 > 0$ and we have $< 0|p_{-n_1}p_{-n_2}|0 >= -n_1 \delta_{n_1+n_2,0}$ using $p_{-n_1}p_{-n_2} = p_{-n_2}p_{-n_1} - n_1\delta_{n_1+n_2,0}$. The sum reduces to:

$$< 0|Q(z_1)Q(z_2)|0 >= 2 \sum_{n_2>0} \frac{1}{n_2} \left(\frac{z_2}{z_1}\right)^{n_2} = \log\left(\frac{z_1+z_2}{z_1-z_2}\right), \quad |z_1| > |z_2|$$

The vertex operator $V(r, z)$ is defined by:

$$V(r, z) = \frac{1}{2} : \exp(irQ(z)) :$$

Proposition. *The normal ordered form of a product of two vertex operators is given by:*

$$V(r, z_1)V(s, z_2) = \left(\frac{z_1 - z_2}{z_1 + z_2}\right)^{rs} : V(r, z_1)V(s, z_2) : \quad |z_1| > |z_2| \tag{16.24}$$

Proof. Let $Q_\pm = -i\sqrt{2}\sum_{\mp n>0} p_{-n}z^n/n$ so that $Q = Q_+ + Q_-$ and $Q_+|0 >= 0$. Then, by definition of the normal order $V(r, z) = \frac{1}{2}\exp(irQ_-(z))\exp(irQ_+(z))$. Hence to compute $: V(r, z_1)V(s, z_2) :$ we need to commute $\exp(irQ_+(z_1))$ to the right of $\exp(isQ_-(z_2))$. Now it is clear that the commutator of $Q_+(z_1)$ and $Q_-(z_2)$ is a $\mathbb{C}$–number. To evaluate this number one can take its vacuum expectation value. One has, using $Q_+|0 >= 0$ and $< 0|Q_- = 0$,

$$< 0|[Q_+(z_1), Q_-(z_2)]|0 >=< 0|Q_+(z_1)Q_-(z_2)|0 >=< 0|Q(z_1)Q(z_2)|0 >$$

which is given by eq.(16.23). Moreover if A and B are two operators such that $[A, B]$ is a $\mathbb{C}$–number, one has $e^A e^B = e^B e^A e^{[A,B]}$. So we arrive at eq.(16.24) since $e^{[A,B]} = ((z_1 - z_2)/(z_1 + z_2))^{rs}$. ∎

The level one vertex operator representations of the Lie algebra $\widehat{sl(2)}$ are obtained as follows:

$$\begin{aligned} \sum_{n \text{ odd}} z^{-n}(E_+^{(n)} + E_-^{(n)}) &= \frac{i}{\sqrt{2}} z \frac{d}{dz} Q(z) \equiv P(z) = \sum_n p_n z^{-n} \\ \sum_{n \text{ even}} z^{-n} H^{(n)} + \sum_{n \text{ odd}} z^{-n}(E_+^{(n)} - E_-^{(n)}) &= \pm V(z) \end{aligned} \tag{16.25}$$

where $V(z)$ denotes the vertex operator:

$$V(z) = V(-\sqrt{2}, z) = \frac{1}{2} : e^{-i\sqrt{2}Q(z)} := \frac{1}{2} : e^{i\sqrt{2}Q(-z)} :$$

Proposition. *The operators $E_\pm^{(n)}$ and $H^{(n)}$ defined in eq.(16.25) provide representations of the affine algebra (16.22) with $K = 1$. These representations correspond to the fundamental highest weights $\Lambda^\pm$ according to the sign in eq.(16.25). They are the fundamental level 1 representations of $\widehat{sl(2)}$.*

Proof. We derive eq.(16.24) with respect to z_1 and get for $|z_1| > |z_2|$:

$$\begin{aligned} \frac{1}{r}\frac{\partial}{\partial_{z_1}} V(r, z_1)V(s, z_2) &= \frac{i}{2} : \frac{dQ(z_1)}{dz_1} e^{irQ(z_1)} : V(s, z_2) \\ &= \frac{2sz_2}{z_1^2 - z_2^2}\left(\frac{z_1 - z_2}{z_1 + z_2}\right)^{rs} : V(r, z_1)V(s, z_2) : + \frac{i}{2}\left(\frac{z_1 - z_2}{z_1 + z_2}\right)^{rs} : \frac{dQ(z_1)}{dz_1} e^{irQ(z_1)} V(s, z_2) : \end{aligned}$$

Here we have used the fact that in the normal ordered product everything commutes so that one can derive the exponential straightforwardly. We then set $r = 0$. Defining:

$$\Gamma(z_1, z_2) = \frac{\sqrt{2}sz_1z_2}{z_1^2 - z_2^2} V(s, z_2) + : P(z_1)V(s, z_2) :$$

we get:

$$P(z_1)V(s, z_2) = \Gamma(z_1, z_2), \quad |z_1| > |z_2|$$

Similarly we derive eq.(16.24) with respect to z_2 and then exchange $(z_1, z_2, r) \to (z_2, z_1, s)$. We get:

$$V(s, z_2)P(z_1) = \Gamma(z_1, z_2), \quad |z_1| < |z_2|$$

Expanding $P(z) = \sum_n p_n z^{-n}$ where n is odd, so that

$$p_n = E_+^{(n)} + E_-^{(n)} = \oint_C \frac{dz}{2i\pi}\, z^{n-1} P(z)$$

we can write the commutator $[p_n, V(s, z_2)]$ as:

$$[p_n, V(s, z_2)] = \oint_{C_1 - C_2} \frac{dz_1}{2i\pi}\, z_1^{n-1} \Gamma(z_1, z_2)$$

where C_1 is a circle around the origin with $|z_1| > |z_2|$ while C_2 is a circle around the origin with $|z_1| < |z_2|$. This contour integral is given by the residues at the two poles $z_1 = \pm z_2$ and we finally obtain, setting $s = -\sqrt{2}$:

$$[E_+^{(n)} + E_-^{(n)}, V(z)] = -2z^n V(z) \tag{16.26}$$

Similarly, starting from eq.(16.24) and setting $V(z) = \sum_n V^{(n)} z^{-n}$, one gets:

$$[V_n, V(z_2)] = \oint_{C_1 - C_2} \frac{dz_1}{2i\pi}\, z_1^{n-1} \left(\frac{z_1 - z_2}{z_1 + z_2}\right)^2 : V(z_1)V(z_2) :$$

The residue is at $z_1 = -z_2$ and is easily computed remarking that $: V(z_2)V(-z_2) := 1/4$. One finds $[V_n, V(z)] = 2(-1)^n z^n P(z) + (-1)^n n z^n$. Separating n even and odd this reads (with ϵ the sign eq.(16.25)):

$$\epsilon[H^{(n)}, V(z)] = 2z^n P(z) + nz^n, \quad \epsilon[E_+^{(n)} + E_-^{(n)}, V(z)] = -2z^n P(z) - nz^n \tag{16.27}$$

From eqs.(16.26,16.27) one gets by expanding $V(z)$ into its components (note that ϵ cancels):

$$[H^{(n)}, E_\pm^{(m)}] = \pm 2E_\pm^{(n+m)} \tag{16.28}$$

$$[E_\pm^{(n)}, E_+^{(m)} - E_-^{(m)}] = -H^{(n+m)} \mp \frac{1}{2} n\delta_{n+m,0} \tag{16.29}$$

Finally we have $P(z) = \sum_n p_n z^{-n}$ so that $E_+^{(n)} + E_-^{(n)} = p_n$ and from the commutation relations of p_n we get $[E_+^{(n)} + E_-^{(n)}, E_+^{(m)} + E_-^{(m)}] = n\delta_{n+m,0}$. Combining with eq.(16.27) this gives:

$$[E_\pm^{(n)}, E_+^{(m)} + E_-^{(m)}] = \pm H^{(n+m)} + \frac{1}{2} n\delta_{n+m,0} \tag{16.30}$$

Equations (16.28,16.29,16.30) are equivalent to the commutation relations eq.(16.22) for $K = 1$. Hence we have obtained a level one repesentation of $\widehat{sl(2)}$. It remains to identify the highest weight. In fact, it is provided by the vacuum vector $|0>$. Indeed

$$V(z)|0>= \epsilon \frac{1}{2} e^{-i\sqrt{2}Q_-(z)}|0> \tag{16.31}$$

contains only positive odd powers of z. It follows that $(E_+^{(n)} - E_-^{(n)})|0>= 0$ and $H^{(n)}|0>= 0$ for $n > 0$. Moreover, since $E_+^{(n)} + E_-^{(n)} = p_n$ we have $(E_+^{(n)} + E_-^{(n)})|0>= 0$ for $n > 0$. This implies that the vacuum is annihilated by all positive root vectors, hence is a highest weight vector. Eq.(16.31) also gives $H^{(0)}|0>= \epsilon\frac{1}{2}|0>$ so that the corresponding weight is $\epsilon\frac{1}{2}$. It is known that this representation on Fock space is irreducible. ∎

References

[1] S. Helgason **"Differential geometry and symmetric spaces."** Academic Press (1962)

[2] J.E. Humphreys **"Introduction to Lie algebras and Representation theory."** Springer (1972)

[3] V. G. Kac **"Infinite dimensional Lie algebras."** Cambridge University Press (1985)

Index