A SHORT INTRODUCTION TO CLASSICAL AND QUANTUM INTEGRABLE SYTEMS.

O. Babelon

Laboratoire de Physique Théorique et Hautes Energies[1], (LPTHE)
Unité Mixte de Recherche UMR 7589
Université Pierre et Marie Curie-Paris6; CNRS; Université Denis Diderot-Paris7;

2007

[1]Tour 24-25, 5ème étage, Boite 126, 4 Place Jussieu, 75252 Paris Cedex 05.

DIRECTION DES SCIENCES DE LA MATIÈRE
SERVICE DE PHYSIQUE THÉORIQUE
Unité de recherche associée au CNRS

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

COURS DE PHYSIQUE THÉORIQUE DU SPHT
ANNÉE 2006-2007

Les vendredis de 14h30 à 16h30 au SPhT, Orme des Merisiers, Bat.774, Salle Itzykson

Petite introduction aux systèmes intégrables classiques et quantiques

Du 19 janvier au 16 février 2007

Organisé en commun avec l'École Doctorale de Physique de la Région Parisienne (ED107)

Olivier BABELON
Lab. de Phys. Théorique
et Hautes Energies
Universités Paris VI et VII

Le but de ce cours est d'introduire les méthodes analytiques de résolution des systèmes intégrables classiques et quantiques.

Systèmes *classiques*:

1- Paires de Lax, orbites coadjointes, matrice r classique.

2- Courbe spectrale, linéarisation du flot sur la jacobienne, variables séparées.

3- Théorie des champs, matrice de monodromie, solitons, solutions finite zones.

Systèmes *quantiques*:

4- Modèle de Gaudin, Ansatz de Bethe.

5- Des équations de Bethe a la courbe spectrale. Limite semi classique.

6- Chaine XXX, Ansatz de Bethe, équation de Baxter. Chaine de Toda...

Les cours sont de nature introductive et accessibles aux étudiants en deuxième année de troisième cycle. Ils sont ouverts aux physiciens de toute discipline et à toute personne intéressée.

Pour tout renseignement : **www-spht.cea.fr** ou **lectures@spht.saclay.cea.fr**

Contents

1 Integrable dynamical systems **7**
- 1.1 The Liouville theorem . . . 7
- 1.2 Lax pairs . . . 10
- 1.3 The Zakharov-Shabat construction. . . . 10
- 1.4 Coadjoint Orbits. . . . 14
- 1.5 Classical r-matrix. . . . 18
- 1.6 Examples. . . . 21
  - 1.6.1 The Jaynes-Cummings-Gaudin model. . . . 21
  - 1.6.2 The KdV hierarchy. . . . 25

2 Solution by analytical methods **31**
- 2.1 The spectral curve. . . . 31
- 2.2 Riemann surfaces. . . . 32
  - 2.2.1 Desingularisation . . . 32
  - 2.2.2 Riemann-Hurwitz theorem. . . . 34
  - 2.2.3 Riemann-Roch theorem. . . . 35
  - 2.2.4 Jacobi variety and Theta functions. . . . 36
- 2.3 Genus of the spectral curve. . . . 40
- 2.4 Dimension of the reduced phase space. . . . 41
- 2.5 The eigenvector bundle. . . . 43
- 2.6 Separated variables. . . . 46
- 2.7 Riemann surfaces and integrability. . . . 49
- 2.8 Solution of the Jaynes-Cummings-Gaudin model. . . . 52
  - 2.8.1 Degenerate case. . . . 53
  - 2.8.2 Non degenerate case . . . 56

3 Infinite dimensional systems. **67**
- 3.1 Integrable field theories and monodromy matrix. . . . 67
- 3.2 Abelianization. . . . 69
- 3.3 Poisson brackets of the monodromy matrix. . . . 72
- 3.4 Dressing transformations. . . . 74
- 3.5 Soliton solutions. . . . 76
  - 3.5.1 KdV solitons. . . . 78

3.6 Finite zones solutions. . . . 80
3.7 The Its-Matveev formula. . . . 81

4 The Jaynes-Cummings-Gaudin model. 89
4.1 Physical context. . . . 89
4.1.1 Rabi oscillations. . . . 89
4.1.2 Cold atoms condensates. . . . 92
4.2 Settings. . . . 93
4.3 Bethe Ansatz. . . . 95
4.4 Riccati equation. . . . 99
4.5 Baxter Equation. . . . 102
4.6 Bethe eigenvectors and separated variables. . . . 104
4.7 Quasi-Classical limit. . . . 108
4.8 Riemann surfaces and quantum integrability. . . . 110

5 The Heisenberg spin chain. 117
5.1 The quantum monodromy matrix . . . 117
5.2 The XXX spin chain. . . . 118
5.3 Algebraic Bethe Ansatz. . . . 123
5.3.1 Baxter equation. . . . 127
5.3.2 Separated variables. . . . 129

6 Nested Bethe Ansatz. 135
6.1 The R-matrix of the Affine sl_{n+1} algebra. . . . 135
6.2 sl_{n+1} generalization of the XXZ model. . . . 136
6.3 Commutation relations. . . . 137
6.4 Reference state. . . . 138
6.5 Bethe equations. . . . 141

Chapter 1

Integrable dynamical systems

The usually accepted definition of an integrable system in the sense of Liouville is a system with phase space of dimension $2n$ for which one knows n conserved quantities in involution. This a rather puzzling definition since by Darboux theorem one 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. However the generic dynamical system is certainely not what we mean by integrable system, so the hypothesis must be made more precise by requiring some global existence properties of the conserved quantities. A good starting point is to ask, following Moser, that the conserved quantities exist on an open domain of the phase space invariant under the dynamical flow, that is any trajectory starting in the domain stays in it.

In all the examples that we will consider the conserved quantities are even analytic functions of canonical coordinates on some open domain and the known solutions are similarly analytic.

1.1 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. In the canonical coordinates $\omega = \sum_j dp_j \wedge dq_j$. Let H be the hamiltonian of the system. For any function f on M, the equations of motion are Hamilton's equations:

$$\frac{df}{dt} \equiv \dot{f} = \{H, f\}$$

Here and in the following, a dot will refer to a time derivative.

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

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

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.

Theorem 1 (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:

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

Their solutions are:

$$\boxed{\begin{aligned} F_j(t) &= F_j(0) \\ \psi_j(t) &= \psi_j(0) + t\Omega_j. \end{aligned}}$$

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.

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 $d^2 S = 0$ 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.

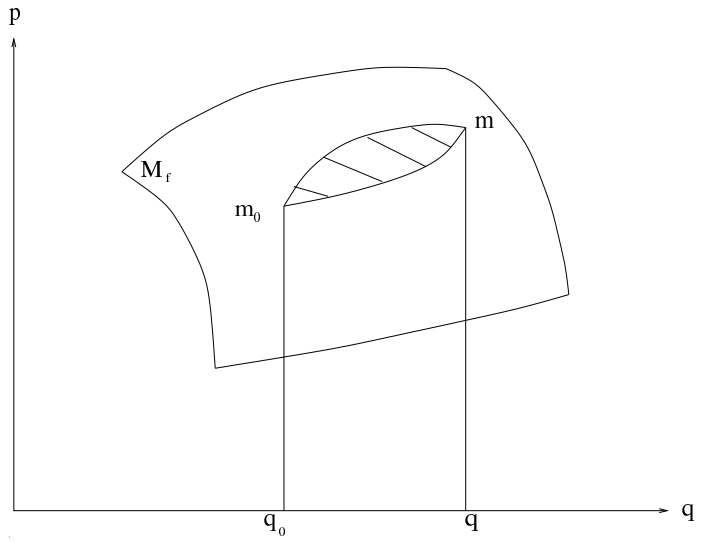


Figure 1.1: Integration path on the level manifold M_f.

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. ■

Remark. 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: $\Delta_{cycle}\psi_j = \frac{\partial}{\partial F_j}\Delta_{cycle} S$.

1.2 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 functions on the phase space of the system, with values in some Lie algebra $\mathcal{G}$, such that the hamiltonian evolution equations may be written as

$$\boxed{\frac{dL}{dt} \equiv \dot{L} = [M, L]} \tag{1.1}$$

Here, [,] denotes the bracket in the Lie algebra $\mathcal{G}$. We will denote by G the connected Lie group having $\mathcal{G}$ as a Lie algebra.

Although eq.(1.1) only requires a Lie algebra structure to be written down, we usually use finite dimensional representations of $\mathcal{G}$, so that L and M are matrices. However it will be important to keep in mind the more abstract formulation.

The interest in the existence of such a pair originates in the fact that it allows for an easy construction of conserved quantities. Indeed the solution of eq(1.1) is of the form

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

where $g(t) \in G$ is determined by the equation

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

It follows that the eigenvalues of L are conserved. We say that the evolution equation (1.1) is isopsectral which means that the spectrum of L is preserved by the time evolution. Alternatively the quantities

$$\boxed{H_n = \text{Tr } (L^n)}$$

are conserved.

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

The interest in the concept of Lax pairs relies on the existence of a tool allowing to produce such pairs fulfilling these constraints.

1.3 The Zakharov-Shabat construction.

Given an integrable system, there does not yet exists a useful algorithm to construct a Lax pair. There does exist however a general procedure, due to Zakharov and Shabat, to construct consistent Lax pairs giving rise to integrable systems. This is a general method

to construct matrices $L(\lambda)$ and $M(\lambda)$, depending on a spectral parameter $\lambda \in \mathbb{C}$, such that the Lax equation

$$\partial_t L(\lambda) = [M(\lambda), L(\lambda)] \tag{1.2}$$

is equivalent to the equations of motion of an integrable system. The main result is eq.(1.13) expressing the possible forms of the matrix M in the Lax pair.

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 $\{\epsilon_k\}$ be the set of their poles, namely the poles of $L(\lambda)$ and those of $M(\lambda)$. With the above notations, assuming no pole at infinity, we can write quite generally:

$$L(\lambda) = L_0 + \sum_k L_k(\lambda), \quad \text{with} \quad L_k(\lambda) \equiv \sum_{r=-n_k}^{-1} L_{k,r}(\lambda - \epsilon_k)^r \tag{1.3}$$

and

$$M(\lambda) = M_0 + \sum_k M_k(\lambda) \quad \text{with} \quad M_k(\lambda) \equiv \sum_{r=-m_k}^{-1} M_{k,r}(\lambda - \epsilon_k)^r \tag{1.4}$$

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

We now want to impose that the Lax equation (1.2), with $L(\lambda)$ and $M(\lambda)$ given by eqs.(1.3,1.4), holds identically in λ. It is important to realize that this is a very non trivial equation. Indeed looking at eqs.(1.2) we see that the pole at ϵ_k in the left hand side is a priori of order n_k while in the right hand side it is potentially 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. It turns out that one can solve the constraints equations.

We introduce a notation. For any matrix valued rational function $f(\lambda)$ with poles of order n_k at points ϵ_k at finite distance, we can decompose $f(\lambda)$ as

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

with f_0 a constant. The quantity $f_k(\lambda)$ is called the polar part at ϵ_k. 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 ϵ_k, $f(\lambda)$ may be decomposed as follows:

$$f(\lambda) = f(\lambda)_+ + f(\lambda)_- \tag{1.5}$$

with $f(\lambda)_+$ regular at the point ϵ_k and $f(\lambda)_- = f_k(\lambda)$ is the polar part.

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

$$L(\lambda) = g^{(k)}(\lambda)\ A^{(k)}(\lambda)\ g^{(k)-1}(\lambda) \tag{1.6}$$

where $A^{(k)}(\lambda)$ is diagonal and has a pole of order n_k at ϵ_k. Obviously, we can write the polar decomposition of $L(\lambda)$ as

$$\boxed{L = L_0 + \sum_k L_k, \quad \text{with} \quad L_k = \left(g^{(k)} A^{(k)} g^{(k)-1}\right)_-} \tag{1.7}$$

A first consequence of the Lax equation is that $M(\lambda)$ admits a similar polar decomposition

Proposition 1 *The decomposition of $M(\lambda)$ in polar parts reads*

$$\boxed{M = M_0 + \sum_k M_k, \quad \text{with} \quad M_k = \left(g^{(k)} B^{(k)} g^{(k)-1}\right)_-} \tag{1.8}$$

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

Proof. Defining $B^{(k)}(\lambda)$ by

$$M(\lambda) = g^{(k)}(\lambda)\ B^{(k)}(\lambda)\ g^{(k)-1}(\lambda) + \partial_t g^{(k)}(\lambda)\ g^{(k)-1}(\lambda) \tag{1.9}$$

the Lax equation becomes:

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

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 ϵ_k. Hence $M_k = (g^{(k)}\ B^{(k)}\ g^{(k)-1})_-$ which only depends on $B^{(k)}_-$. This simultaneous diagonalization of $L(\lambda)$ and $M(\lambda)$ works around any point where $L(\lambda)$ has distinct eigenvalues.

This proposition clarifies the structure of the Lax pair. 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

$$A^{(k)}_- = \sum_{r=-n_k}^{-1} A_{k,r}(\lambda - \epsilon_k)^r \tag{1.10}$$

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

$$\hat{g}^{(k)} = \sum_{r=0}^{n_k-1} g_{k,r}(\lambda - \epsilon_k)^r \tag{1.11}$$

From these data, we can reconstruct the Lax matrix $L(\lambda)$ by defining $L = L_0 + \sum_k L_k$ with

$$L_k \equiv \left(\hat{g}^{(k)}\, A_-^{(k)}\, \hat{g}^{(k)-1}\right)_- \tag{1.12}$$

Then around each ϵ_k, one can diagonalize $L(\lambda) = g^{(k)} A^{(k)} g^{(k)-1}$. This yields an extension of the matrices $A_-^{(k)}$ and $\hat{g}^{(k)}$ to complete series $A^{(k)}$ and $g^{(k)}$ in $(\lambda - \epsilon_k)$. Finally to define $M(\lambda) = M_0 + \sum_k M_k$, we choose a set of *diagonal* polar matrices $(B^{(k)}(\lambda))_-$ and use the series $g^{(k)}$ to define M_k by eq.(1.8).

In the vicinity of a singularity, $L(\lambda)$ and $M(\lambda)$ can be simultaneously diagonalized if the Lax equation holds true. In this diagonal gauge, the Lax equation simply states that the matrix $A^{(k)}(\lambda)$ is conserved and that $B^{(k)}(\lambda)$ is diagonal. When we transform these results into the original gauge, we get the general solution of the non dynamical constraints on $M(\lambda)$:

Proposition 2 *Let $L(\lambda)$ be a Lax matrix of the form eq.(1.3). The general form of the matrix $M(\lambda)$ such that the orders of the poles match on both sides of the Lax equation is $M = M_0 + \sum_k M_k$ with*

$$\boxed{M_k = \left(P^{(k)}(L,\lambda)\right)_-} \tag{1.13}$$

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

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 $\lambda = \epsilon_k$. Using a beautiful argument first introduced by Gelfand and Dickey we write:

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

where we used that a polynomial in L commutes with L. From this we see that the order of the pole at ϵ_k is less than n_k. To show that this is a general solution, recall eqs. (1.6,

1.8). Since $A^{(k)}(\lambda)$ is a diagonal $N \times N$ matrix with all its elements distinct in a vicinity of ϵ_k, its powers 0 up to $N-1$ span the space of diagonal matrices and one can write

$$B^{(k)} = P^{(k)}(A^{(k)}, \lambda) \tag{1.14}$$

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 $\lambda - \epsilon_k$ in a vicinity of ϵ_k. Inserting eq. (1.14) into eq. (1.8) 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 $\lambda - \epsilon_k$ since a high enough power cannot contribute to the singular part, yielding a polynomial with coefficients Laurent *polynomials* in $\lambda - \epsilon_k$.

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 a priori functions of the matrix elements of L and require further characterizations in order to get an integrable system. In the setting of the next section these coefficients will be *constants*.

Remark. The Lax equation is invariant under similarity transformations,

$$L \to L' = gLg^{-1}, \quad M \to M' = gMg^{-1} + \partial_t g g^{-1} \tag{1.15}$$

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.(1.15) to diagonalize L_0,

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

Consistency of eq.(1.2) then requires M_0 to be also diagonal and thus $\dot{L}_0 = [M_0, L_0] = 0$. 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.

1.4 Coadjoint Orbits.

In this section we show that the Zakharov–Shabat construction, when the matrices $A_-^{(k)}$ are non dynamical, can be interpreted as coadjoint orbits. This introduces a natural symplectic structure in the problem and gives a Hamiltonian interpretation to the Lax equation.

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 actions of the Lie algebra $\mathcal{G}$ on $\mathcal{G}$ and $\mathcal{G}^*$ denoted ad and ad* respectively and given by:

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

On the space $\mathcal{F}(\mathcal{G}^*)$ of functions on $\mathcal{G}^*$ there is a canonical Poisson bracket called the Kostant-Kirillov bracket. Let $\Xi \in \mathcal{G}^*$ and $X, Y \in \mathcal{G}$, we define

$$\boxed{\{\Xi(X), \Xi(Y)\} = \Xi([X,Y])} \tag{1.16}$$

If e_a is a basis of $\mathcal{G}$ and e^{*a} is the dual basis of $\mathcal{G}^*$, then we have

$$X = \sum_a X^a e_a \in \mathcal{G}, \quad \Xi = \sum_a \Xi_a e^{*a} \in \mathcal{G}^*$$

and

$$\Xi(X) = \sum \Xi^a X_a$$

Setting $X = e_a$, $Y = e_b$ in eq.(1.16) we find

$$\{\Xi_a, \Xi_b\} = f_{abc}\Xi_c$$

where we have introduced the structure constants of the Lie algebra $\mathcal{G}$

$$[e_a, e_b] = f_{abc} e_c$$

This formula define the Poisson bracket of the coordinates Ξ_a on $\mathcal{G}^*$. We can extend it on the functions on $\mathcal{G}^*$ (i.e. functions of the Ξ_a) in the usual way

$$\{F,G\}(\Xi) = \sum_{a,b} \frac{dF}{d\Xi_a}\frac{dG}{d\Xi_b}\{\Xi_a, \Xi_b\} = \sum_{a,b,c} \Xi_c f_{abc} \frac{dF}{d\Xi_a}\frac{dG}{d\Xi_b}$$

Introducing the differentials $dF, dG \in \mathcal{G}$ as

$$dF(\Xi) = \sum_a \frac{dF}{d\Xi_a} e_a, \quad dG(\Xi) = \sum_b \frac{dG}{d\Xi_b} e_b$$

the above formula can be rewritten in the more invariant way

$$\{F,G\}(\Xi) = \Xi([dF, dG])$$

The Kostant-Kirillov bracket is degenerate, meaning that there exists functions which Poisson commutes with everything. For instance if the basis e_a is chosen so that the structure constants f_{abc} are totally antisymmetric the function

$$\Xi^2 = \sum_a \Xi_a^2$$

is in the center of the Poisson bracket. Indeed

$$\{\Xi^2, \Xi_b\} = 2\sum_{a,c} f_{abc}\Xi_a\Xi_c = 0$$

In the context of Hamiltonian mechanics, it is important to identify all such functions and to set them to constants since they cannot contribute to the dynamics. This is where the notion of coadjoint orbit plays a very important role. The coadjoint orbit of an element $A \in \mathcal{G}^*$ is the set of elements of $\mathcal{G}^*$ defined as

$$\text{Orbit}(A) = \{\text{Ad}^*_g \cdot A, \forall g \in G\}$$

The center of the Kostant-Kirillov bracket consists of the functions which are Ad^*-invariant, i.e. which are constant on coadjoint orbits. In fact for such a function we have

$$F(\Xi) = F(\text{Ad}^*_g \cdot \Xi)$$

Taking an infinitesimal $g = 1 + \epsilon X$, this translates to

$$F(\Xi) = F(\Xi + \epsilon\ \text{ad}^* X \cdot \Xi) = F(\Xi) + \epsilon\ \text{ad}^* X \cdot \Xi(dF) + O(\epsilon^2)$$

hence, the Ad^*-invariance of $F(\Xi)$ can be written as

$$\text{ad}^* X \cdot \Xi(dF) = \Xi([dF, X]) = 0, \quad \forall X \in \mathcal{G}$$

On the other hand, if $F(\Xi)$ is in the kernel of the Kostant-Kirillov bracket we have

$$\{F(\Xi), \Xi(X)\} = \Xi([dF, X]) = 0, \quad \forall X \in \mathcal{G}$$

so that the Ad^*-invariant functions are in the kernel and vice versa. On coadjoint orbits, the Kostant-Kirillov bracket becomes non degenerate.

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

$$L(\lambda) = \Big(g(\lambda)\, A_-(\lambda)\, g^{-1}(\lambda)\Big)_- \tag{1.17}$$

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

Let G be the loop group of invertible matrix valued 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:

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

where $\mathrm{Res}_{\lambda=0}$ is defined to be the coefficient of λ^{-1}.

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

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

This is precisely eq.(1.17). The Lax matrix can thus be interpreted as belonging to the coadjoint orbit of the element $A_-(\lambda)$ of $\mathcal{G}^*$ under the loop group G.

$$L(\lambda) = \Big(g(\lambda)\, A_-(\lambda)\, g^{-1}(\lambda)\Big)_- = \mathrm{Orbit}(A_-(\lambda))$$

If we take any element of the orbit and try to diagonalize it, the singular part of the matrix of eigenvalues is precisely $A_-(\lambda)$ which is therefore an Ad*-invariant function and should be put to constants. This interpretation of $L(\lambda)$ as a coadjoint orbit therefore assumes that $A_-(\lambda)$ is not a dynamical variable. The coadjoint orbit is then equipped with the Kostant-Kirillov symplectic structure.

This construction can be extended to the multi–pole case. We consider the direct sum of loop algebras $\mathcal{G}_k$, around $\lambda = \epsilon_k$:

$$\mathcal{G} \equiv \bigoplus_k \mathcal{G}_k$$

An element of this Lie algebra has the form of a multiplet

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

where $X_k(\lambda)$, defined around ϵ_k, is of the form $X_k(\lambda) = \sum_{n\geq 0} X_{k,n}\ (\lambda - \epsilon_k)^n$. The Lie bracket is such that $[X_k(\lambda), X_l(\lambda)] = 0$ if $k \neq l$. The group G is the direct product of the groups G_k of regular invertible matrices at ϵ_k:

$$G \equiv (G_1, G_2, \cdots) \tag{1.18}$$

The dual $\mathcal{G}^*$ of this Lie algebra consists of multiplets

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

where $\Xi_k(\lambda)$ around ϵ_k is of the form $\Xi_k(\lambda) = \sum_{r\geq 1} \Xi_{k,r}(\lambda - \epsilon_k)^{-r}$. In this sum the number of terms is finite but arbitrary. The pairing is simply

$$\langle \Xi, X\rangle \equiv \sum_k \langle \Xi_k, X_k\rangle = \sum_k \text{Tr Res}_{\epsilon_k}(\Xi_k(\lambda)X_k(\lambda))$$

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, we can interpret eq.(1.7) as the coadjoint orbit of the element $((A_1)_-, (A_2)_-, \cdots)$.

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

$$L(\lambda) = L_0 + \sum_k \Xi_k \tag{1.19}$$

with poles at the points ϵ_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.(1.7). By choice it is assumed to be invariant by coadjoint action. The pairing can be rewritten as

$$\langle L, X\rangle = \sum_k \text{Tr Res}_{\epsilon_k} L(\lambda)X_k(\lambda)$$

Remark that only Ξ_k contributes to the residue at ϵ_k and the formula is compatible with the matrix L_0 being invariant by coadjoint action.

1.5 Classical r-matrix.

We can now use this symplectic form to evaluate the Poisson brackets of the elements of the Lax matrix and show that they take the r–matrix form. Let us first introduce some notations. Let E_{ij} be the canonical basis of the $N \times N$ matrices, $(E_{ij})_{kl} = \delta_{ik}\delta_{jl}$. We can write

$$L(\lambda) = \sum_{ij} L_{ij}(\lambda)E_{ij}$$

Let

$$L_1(\lambda) \equiv L(\lambda) \otimes 1 = \sum_{ij} L_{ij}(\lambda)(E_{ij} \otimes 1), \quad L_2(\mu) \equiv 1 \otimes L(\mu) = \sum_{ij} L_{ij}(\mu)(1 \otimes E_{ij})$$

The index 1 or 2 means that the matrix L sits in the first or second factor in the tensor product. More generally when we have tensor products with more copies, we denote by L_α the embedding of L in the α position, e.g. $L_3 = 1 \otimes 1 \otimes L \otimes 1 \otimes \cdots$. Finally, we define $\{L_1(\lambda), L_2(\mu)\}$ as the matrix of Poisson brackets between the elements of L:

$$\{L_1(\lambda), L_2(\mu)\} = \sum_{ij,kl} \{L_{ij}(\lambda), L_{kl}(\mu)\}E_{ij} \otimes E_{kl}$$

We assume that each $L_k(\lambda)$ is a generic element of an orbit of the loop group $GL(N)[\lambda]$ that L_0 and the $A_-^{(k)}$ are non–dynamical. Each orbit $L_k(\lambda)$ is equipped with the Kostant-Kirillov Poisson bracket and

$$\{L_k(X), L_{k'}(Y)\} = 0, \quad k \neq k' \tag{1.20}$$

Proposition 3 *With these assumptions, the Poisson brackets of the matrix elements of $L(\lambda)$ can be written as:*

$$\boxed{\{L_1(\lambda), L_2(\mu)\} = -\left[\frac{C_{12}}{\lambda-\mu}, L_1(\lambda) + L_2(\mu)\right]} \tag{1.21}$$

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.(1.21) is the usual matrix commutator.

Proof. Let us first assume that we have only one pole and $L = (gA_-g^{-1})_-$. 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 $\{L(X), L(Y)\} = L([X,Y])$. Using, $L(X) = \text{Tr Res}_{\lambda=0}(L(\lambda)X(\lambda))$, this gives:

$$\{L(X), L(Y)\} = \text{Tr Res}_{\lambda=0}\left(L(\lambda)[X(\lambda), Y(\lambda)]\right) \tag{1.22}$$

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

$$\{L(X), L(Y)\} = \langle \{L_1(\lambda), L_2(\mu)\}\, ,\, X(\lambda) \otimes Y(\mu) \rangle$$

where $\langle , \rangle = \text{Tr}_{12}\text{Res}_\lambda \text{Res}_\mu$. We need to factorize $X(\lambda) \otimes Y(\mu)$ in eq.(1.22). To this end, we introduce the operator, assuming $|\lambda| < |\mu|$,:

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

This operator is such that for $Y(\lambda) = \sum_{n=0}^{\infty} Y_n \lambda^n$ we have

$$Y_1(\lambda) = \text{Tr}_2 \text{Res}_\mu\, \mathcal{C}_{12}(\lambda,\mu) Y_2(\mu)$$

We can now write:

$$\langle L(\lambda)[X(\lambda), Y(\lambda)] \rangle = \langle [\mathcal{C}_{12}(\lambda,\mu), L(\lambda) \otimes 1]\, ,\, X(\lambda) \otimes Y(\mu) \rangle$$

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 $\langle \varphi, X(\lambda) \otimes Y(\mu) \rangle = 0$. 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 $\langle \cdot, X(\lambda) \otimes Y(\mu) \rangle$. 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. This proves eq.(1.21) in the one-pole case.

We can now study the multi–pole situation occuring in eq.(1.7). Consider $L = L_0 + \sum_k L_k$. Each L_k lives in a coadjoint orbit as above equipped with its own symplectic structure. From eq.(1.20) they have vanishing mutual Poisson brackets

$$\{(L_k(\lambda))_1, (L_{k'}(\mu))_2\} = 0, \quad k \neq k'$$

We assume further that L_0 does not contain dynamical variables

$$\{(L_0)_1, (L_0)_2\} = 0, \quad \{(L_0)_1, (L_k(\lambda))_2\} = 0$$

Then since $C_{12}/(\lambda - \mu)$ is *independent* of the pole ϵ_k, it is obvious that the r-matrix relations for each orbit combine by addition to give eq.(1.21) for the complete Lax matrix $L(\lambda)$. ∎

We have obtained a very simple formula for the r-matrix specifying the Poisson bracket of $L(\lambda)$:

$$\boxed{r_{12}(\lambda, \mu) = -r_{21}(\mu, \lambda) = -\frac{C_{12}}{(\lambda - \mu)}} \tag{1.23}$$

The Jacobi identity is satisfied because this r-matrix verifies the classical Yang–Baxter equation

$$\boxed{[r_{12}, r_{13}] + [r_{12}, r_{23}] + [r_{13}, r_{23}] = 0}$$

where r_{ij} stands for $r_{ij}(\lambda_i, \lambda_j)$. Note that r_{12} is antisymmetric: $r_{12}(\lambda_1, \lambda_2) = -r_{21}(\lambda_2, \lambda_1)$.

These Poisson brackets for the Lax matrix ensure that one can define commuting quantities. The associated equations of motion take the Lax form.

Proposition 4 *The functions on phase space:*

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

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

$$\boxed{M_k(\lambda) = -n \left(\frac{L^{n-1}(\lambda)}{\lambda - \mu}\right)_k} \tag{1.24}$$

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

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

where we have used that the trace of a commutator vanishes. Similarly, we have:

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

Performing the trace and remembering that $\mathrm{Tr}_2\,(C_{12}M_2) = M_1$, we get

$$\dot{L}(\lambda) = [M^{(n)}(\lambda,\mu), L(\lambda)], \quad M^{(n)}(\lambda,\mu) = n\frac{L^{n-1}(\mu)}{\lambda-\mu} \tag{1.25}$$

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:

$$M^{(n)}(\lambda,\mu) \to M^{(n)}(\lambda,\mu) - n\frac{L^{n-1}(\lambda)}{\lambda-\mu} = -n\frac{L^{n-1}(\lambda) - L^{n-1}(\mu)}{\lambda-\mu} \tag{1.26}$$

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

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

This is of the form eq.(1.13) with

$$P^{(k)}(L,\lambda) = -\frac{n}{\lambda-\mu}L^{n-1}(\lambda) \tag{1.27}$$

Notice that the coefficients of the polynomial $P^{(k)}(L,\lambda)$ are pure numerical constants.

This proposition shows that the generic Zakharov-Shabat system, equipped with this symplectic structure, is an *integrable Hamiltonian system* (the precise counting of independent conserved quantities will be done in Chapter [2]).

1.6 Examples.

1.6.1 The Jaynes-Cummings-Gaudin model.

We consider the following Hamiltonian

$$H = \sum_{j=0}^{n-1} 2\epsilon_j s_j^z + \omega\bar{b}b + g\sum_{j=0}^{n-1}\left(\bar{b}s_j^- + bs_j^+\right) \tag{1.28}$$

The $\vec{s}_j$ are spins variables, and $b, \bar{b}$ is a harmonic oscillator. The Poisson brackets read

$$\{s_j^a, s_j^b\} = -\epsilon_{abc} s_j^c, \quad \{b, \bar{b}\} = i \tag{1.29}$$

The $\vec{s}_j$ brackets are degenerate. We fix the value of the Casimir functions

$$\vec{s}_j \cdot \vec{s}_j = s^2$$

Phase space has dimension $2(n+1)$. The equations of motion read

$$\begin{aligned}
\dot{b} &= -i\omega b - ig \sum_{j=0}^{n-1} s_j^- && (1.30)\\
\dot{s}_j^z &= ig(\bar{b} s_j^- - b s_j^+) && (1.31)\\
\dot{s}_j^+ &= 2i\epsilon_j s_j^+ - 2ig\bar{b} s_j^z && (1.32)\\
\dot{s}_j^- &= -2i\epsilon_j s_j^- + 2igb s_j^z && (1.33)
\end{aligned}$$

We introduce the Lax matrices

$$\begin{aligned}
L(\lambda) &= \frac{2}{g^2}\lambda\sigma^z + \frac{2}{g}(b\sigma^+ + \bar{b}\sigma^-) - \frac{\omega}{g^2}\sigma^z + \sum_{j=0}^{n-1} \frac{\vec{s}_j \cdot \vec{\sigma}}{\lambda - \epsilon_j} && (1.34)\\
M(\lambda) &= -i\lambda\sigma^z - ig(b\sigma^+ + \bar{b}\sigma^-) && (1.35)
\end{aligned}$$

where σ^a are the Pauli matrices.

$$\sigma^\pm = \frac{1}{2}(\sigma^x \pm i\sigma^y), \quad [\sigma^z, \sigma^\pm] = \pm 2\sigma^\pm, \quad [\sigma^+, \sigma^-] = \sigma^z$$

It is not difficult to check that the equations of motion are equivalent to the Lax equation

$$\dot{L}(\lambda) = [M(\lambda), L(\lambda)] \tag{1.36}$$

Let

$$L(\lambda) = \begin{pmatrix} A(\lambda) & B(\lambda) \\ C(\lambda) & -A(\lambda) \end{pmatrix}$$

we have

$$\begin{aligned}
A(\lambda) &= \frac{2\lambda}{g^2} - \frac{\omega}{g^2} + \sum_{j=0}^{n-1} \frac{s_j^z}{\lambda - \epsilon_j}\\
B(\lambda) &= \frac{2b}{g} + \sum_{j=0}^{n-1} \frac{s_j^-}{\lambda - \epsilon_j}\\
C(\lambda) &= \frac{2\bar{b}}{g} + \sum_{j=0}^{n-1} \frac{s_j^+}{\lambda - \epsilon_j}
\end{aligned}$$

It is very simple to check that

$$\begin{aligned}
\{A(\lambda), A(\mu)\} &= 0 \\
\{B(\lambda), B(\mu)\} &= 0 \\
\{C(\lambda), C(\mu)\} &= 0 \\
\{A(\lambda), B(\mu)\} &= \frac{i}{\lambda-\mu}(B(\lambda)-B(\mu)) \\
\{A(\lambda), C(\mu)\} &= -\frac{i}{\lambda-\mu}(C(\lambda)-C(\mu)) \\
\{B(\lambda), C(\mu)\} &= \frac{2i}{\lambda-\mu}(A(\lambda)-A(\mu))
\end{aligned}$$

One can rewrite these equations in the usual form

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

where

$$P_{12} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

It follows that

$$\frac{1}{2}\mathrm{Tr}\,(\mathrm{L}^2(\lambda)) = \mathrm{A}^2(\lambda) + \mathrm{B}(\lambda)\mathrm{C}(\lambda)$$

generates Poisson commuting quantities. One has

$$\frac{1}{2}\mathrm{Tr}\; L^2(\lambda) = \frac{1}{g^4}(2\lambda-\omega)^2 + \frac{4}{g^2}H_n + \frac{2}{g^2}\sum_{j=0}^{n-1}\frac{H_j}{\lambda-\epsilon_j} + \sum_{j=0}^{n-1}\frac{\vec{s}_j\cdot\vec{s}_j}{(\lambda-\epsilon_j)^2} \tag{1.37}$$

where the $(n+1)$ Hamiltonians read

$$H_n = b\bar{b} + \sum_j s_j^z$$

and

$$H_j = (2\epsilon_j-\omega)s_j^z + g(bs_j^+ + \bar{b}s_j^-) + g^2\sum_{k\neq j}\frac{s_j\cdot s_k}{\epsilon_j-\epsilon_k}, \quad j=0,\cdots,n-1$$

The Hamiltonian eq.(1.28) is

$$H = \omega H_n + \sum_{j=0}^{n-1} H_j$$

Let us see how these formulae fit into our general scheme. The Lax matrix is a sum of simple poles at ϵ_j. The loop group at each one of these points is

$$g^{(j)}(\lambda) = g_0^{(j)} + (\lambda-\epsilon_j)g_1^{(j)} + \cdots$$

where $g_k^{(j)}$ are $SU(2)$ matrices. In fact the pole being simple, only $g_0^{(j)}$ contributes to the coadjoint orbit.

$$L_j(\lambda) = \left(g^{(j)}(\lambda) \, \frac{s\sigma^z}{\lambda - \epsilon_j} \, g^{(j)-1}(\lambda) \right)_- = \frac{s}{\lambda - \epsilon_j} \, g_0^{(j)} \sigma^z g_0^{(j)-1}$$

At infinity we consider the loop group

$$g^{(\infty)}(\lambda) = 1 + \lambda^{-1} g_{-1}^{(\infty)} + \cdots$$

and the coadjoint orbit

$$L_\infty(\lambda) = \frac{2}{g^2} \left(g^{(\infty)}(\lambda) \, \lambda\sigma^z \, g^{(\infty)-1}(\lambda) \right)_+ = \frac{2}{g^2} \lambda\sigma^z + \frac{2}{g^2} [g_{-1}^{(\infty)}, \sigma^z] \equiv \frac{2}{g^2} \lambda\sigma^z + \frac{2}{g} (b\sigma^+ + \bar{b}\sigma^-)$$

Identifying

$$L_0 = -\frac{\omega}{g^2} \sigma^z$$

we do have

$$L(\lambda) = L_0 + L_\infty(\lambda) + \sum_{j=0}^{n-1} L_j(\lambda)$$

To see how $M(\lambda)$ also fits into the scheme, we consider separately the evolution with respect to the Hamiltonians H_j, H_n. Since H_j is the coefficient of $\frac{2}{g^2(\mu - \epsilon_j)}$ in $\frac{1}{2}\mathrm{Tr}\, L^2(\mu)$ we just have to extract this coefficient in eq.(1.26), which for $n = 2$ reads (there is an extra factor i coming from the definition of the Poisson bracket)

$$M^{(2)}(\lambda, \mu) = -i \frac{L(\lambda) - L(\mu)}{\lambda - \mu}$$

We find

$$M_j(\lambda) = i \frac{g^2}{2} \frac{\vec{s}_j \cdot \vec{\sigma}}{\lambda - \epsilon_j}$$

Similarly, for ωH_n we have to extract the term in μ^0 in the same expression. We find

$$M_\infty(\lambda) = -i \frac{\omega}{2} \sigma^z \tag{1.38}$$

Hence

$$\begin{aligned} M(\lambda) = M_\infty(\lambda) + M_j(\lambda) &= i \frac{g^2}{2} \left(-\frac{\omega}{g^2} \sigma^z + \sum_{j=0}^{n-1} \frac{\vec{s}_j \cdot \vec{\sigma}}{\lambda - \epsilon_j} \right) \\ &= -i \Big(\lambda\sigma^z + g(b\sigma^+ + \bar{b}\sigma^-) \Big) + i \frac{g^2}{2} L(\lambda) \end{aligned}$$

Since the $L(\lambda)$ term does not contribute to the Lax equation, we have recovered the expression eq.(1.35) for the matrix $M(\lambda)$.

1.6.2 The KdV hierarchy.

In the one pole case, we have seen that the general structure of a Lax equation is

$$\dot{L}(\lambda) = \left[\Big(P(L,\lambda)\Big)_-, L(\lambda)\right] \tag{1.39}$$

The KdV hierarchies are constructed exactly on the same pattern but replacing loop algebras by the algebra of pseudo differential operators which we now describe.

The algebra of pseudo-differential operators is the algebra of elements of the form

$$A = \sum_{i=-\infty}^{N} a_i \partial^i$$

with N finite but arbitrary. The coefficients a_i are functions of x, ∂ is the usual derivation with respect to x and the "integration" symbol, ∂^{-1} is defined by the following algebraic rules:

$$\begin{aligned} \partial^{-1}\partial &= \partial\partial^{-1} = 1 \\ \partial^{-1} a &= \sum_{i=0}^{\infty} (-1)^i (\partial^i a)\partial^{-i-1} \end{aligned} \tag{1.40}$$

We denote by $\mathcal{P} = \left\{A = \sum_{-\infty}^{N} a_i\partial^i\right\}$ the set of formal pseudo-differential operators. Let $\mathcal{P}_+ = \left\{A = \sum_{i=0}^{N} a_i\partial^i\right\}$ be the subalgebra of differential operators, and let $\mathcal{P}_- = \left\{A = \sum_{-\infty}^{-1} a_i\partial^i\right\}$ be 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.

The formal group $G = \exp(\mathcal{P}_-)$ is called the Volterra group. 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:

$$\Phi = 1 + \sum_{i=1}^{\infty} w_i \partial^{-i} \ \in (1 + \mathcal{P}_-) \tag{1.41}$$

The coefficients of its inverse $\Phi^{-1} = 1 + \sum_1^\infty w_i'\partial^{-i}$ can be computed recursively from the relation $\Phi^{-1}\Phi = 1$.

For $A \in \mathcal{P}$ we define its residue, denoted $\mathrm{Res}_\partial A$, as the coefficient of ∂^{-1} in A:

$$\boxed{\mathrm{Res}_\partial A \equiv a_{-1}(x)} \tag{1.42}$$

On $\mathcal{P}$ there exists a natural linear form called the Adler trace, denoted $\langle\ \rangle$, defined by:

$$\boxed{\langle A\rangle = \int dx\, \mathrm{Res}_\partial A = \int dx\ a_{-1}(x)} \tag{1.43}$$

This linear form satisfies the fundamental trace property $\langle AB\rangle = \langle BA\rangle$.

Returning to integrable systems, we now let L be a differential operator of order n

$$L = \partial^n - \sum_{i=0}^{n-2} u_i \partial^i \tag{1.44}$$

In the algebra of pseudo differential operators its n-th root exists

$$Q = L^{\frac{1}{n}}, \quad Q = \partial + q_{-1}\partial^{-1} + \cdots$$

The generalized KdV hierarchies are defined by the Lax equations (compare with eq.(1.39). In both cases the projection is on the dual of the Lie algebras $\mathcal{G}[\lambda]$ and $\mathcal{P}_-$ respectively.)

$$\boxed{\partial_{t_k} L = \left[\left(L^{\frac{k}{n}}\right)_+, L\right]} \tag{1.45}$$

These equations are consistent for all $k \in \mathbb{N}$ in the sense that we have a differential operator of order $n-2$ on both side. To see it, we notice that Q^k, $\forall k \in N$ commutes with L since $LQ^k = Q^{n+k} = Q^k L$. Then, we have:

$$\left[\left(Q^k\right)_+, L\right] \ = \ \left[Q^k, L\right] - \left[\left(Q^k\right)_-, L\right] = -\left[\left(Q^k\right)_-, L\right]$$

From the last equality, it follows that the differential operator $\left[\left(Q^k\right)_+, L\right]$ is of order less or equal to $n-2$, so that the Lax equation eq.(1.45) is an equation on the coefficients of L. This is the original Gelfand-Dickey argument.

The differential operator L is an element of $\mathcal{P}_+$. If we view $\mathcal{P}_+$ as the dual of the Lie algebra $\mathcal{P}_-$ through the Adler trace, there is a natural Poisson bracket on $\mathcal{P}_+$: the Kostant–Kirillov bracket. For any functions f and g on $\mathcal{P}_+$, it is defined as usual by:

$$\{f, g\}(L) = \langle L\,,\, [df, dg]\rangle \quad \forall\ L \in \mathcal{P}_+ \tag{1.46}$$

where we understand that $df, dg \in \mathcal{P}_-$. In particular, for any $X = \sum_{j=0}^{\infty} \partial^{-j-1} x_j \in \mathcal{P}_-$, we define the linear function $f_X(L)$ by:

$$f_X(L) = \langle L, X \rangle \tag{1.47}$$

we have $df_X = X \in \mathcal{P}_-$. Therefore $\{f_X, f_Y\} = \langle L, [X, Y] \rangle = f_{[X,Y]}$ for any $X,\ Y \in \mathcal{P}_-$.

Proposition 5 *Let $L \in \mathcal{P}_+$ be the differential operator of order n as in eq.(1.44). Define the functions of L by*

$$\boxed{H_k(L) = \frac{1}{1 + \frac{k}{n}} \langle L^{\frac{k}{n}+1} \rangle}$$

(i) The functions $H_k(L)$ are the Hamiltonians of the generalized KdV flows under the bracket eq.(1.46):

$$\partial_{t_k} L = \{H_k, L\} = \left[\left(L^{\frac{k}{n}} \right)_+, L \right] \tag{1.48}$$

(ii) The functions $H_k(L)$ are in involution with respect to this bracket.

Proof. We first need to compute the differential of the Hamiltonian H_k. Let L and δL be differential operators of the form eq.(1.44). One has, using the cyclicity of Adler's trace:

$$\langle (L + \delta L)^\nu \rangle = \langle L^\nu \rangle + \nu \langle L^{\nu-1} \delta L \rangle + \cdots$$

which implies $d\langle L^\nu \rangle = \nu (L^{\nu-1})_{(-n)}$ where the notation $(\ )_{(-n)}$ means projection on $\mathcal{P}_-$ truncated at the first $n-1$ terms. This projection appears because $\delta L = -\delta u_{n-2} \partial^{n-2} - \cdots - \delta u_0$ which is dual to elements of the form $\partial^{-1} x_0 + \cdots + \partial^{-n+1} x_{n-1}$ under the Adler trace. Hence:

$$dH_k(L) = \left(L^{\frac{k}{n}} \right)_{(-n)} = \left(Q^k \right)_{(-n)} \quad \in \mathcal{P}_- \tag{1.49}$$

We call $\theta^{(k)}_{-(n)}$ the terms left over in the truncation:

$$\left(L^{\frac{k}{n}} \right)_- = dH_k + \theta^{(k)}_{-(n)} \tag{1.50}$$

We now prove eq.(1.48). Consider the function $f_X(L) = \langle LX \rangle$, then

$$\dot{f}_X = \langle \dot{L}, X \rangle = \{H_k, f_X\}(L) = \langle L, [dH_k, df_X] \rangle = \langle [L, dH_k], X \rangle$$

where we used the invariance of the Adler trace. Since $X \in \mathcal{P}_-$, only $[L, dH_k]_+$ contributes to this expression. But

$$[L, dH_k]_+ = \left[L, \left(L^{\frac{k}{n}} \right)_- \right]_+ - \left[L, \theta^{(k)}_{-(n)} \right]_+ = \left[\left(L^{\frac{k}{n}} \right)_+, L \right]$$

where we have used $[L^{\frac{k}{n}}, L] = 0$, and the fact that $[L, \theta^{(k)}_{-(n)}]_+ = 0$. So $[L, dH_k]_+$ is a differential operator of order at most $n-2$, and this proves eq.(1.48).

Next we show that the Hamiltonians H_k are in involution. We have:

$$\begin{aligned}\{H_k, H_{k'}\}(L) &= \langle L, [dH_k, dH_{k'}]\rangle = \langle [L, dH_k]_+ \, , \, dH_{k'}\rangle \\ &= \langle \left[\left(L^{\frac{k}{n}}\right)_+ , L\right] \, , \, dH_{k'}\rangle\end{aligned}$$

Using again the fact that $[L, dH_k]_+$ is of order at most $n-2$, we can replace $dH_{k'}$ by $\left(L^{\frac{k'}{n}}\right)_-$, and get:

$$\{H_k, H_{k'}\}(L) = \langle \left[\left(L^{\frac{k}{n}}\right)_+ , L\right] \left(L^{\frac{k'}{n}}\right)_- \rangle = \langle \left[\left(L^{\frac{k}{n}}\right)_+ , L\right] \; L^{\frac{k'}{n}} \rangle$$

In the last step we used that $\langle \mathcal{P}_+, \mathcal{P}_+\rangle = 0$ in order to replace $\left(L^{\frac{k'}{n}}\right)_-$ by $L^{\frac{k'}{n}}$. Finally, from the invariance of the trace, we obtain:

$$\{H_k, H_{k'}\}(L) = \langle \left(L^{\frac{k}{n}}\right)_+ \left[L, L^{\frac{k'}{n}}\right] \rangle = 0$$

These systems are called the generalized KdV hierarchies. They are field equations or infinite dimensional systems. They are integrable in the sense that they possess an infinite number of Poisson commuting conserved quantities but we are already beyond the strict framework of the Liouville theorem.

The KdV hierarchy corresponds to $n = 2$ and the generalized ones to $n = 3, 4, \cdots$. Let us consider the KdV case $n = 2$. 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...

$$Q = \partial - \frac{1}{2}u\partial^{-1} + \frac{1}{4}(\partial u)\partial^{-2} + \cdots$$

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. To obtain the KdV flows, we only have to compute $(Q^k)_+$, $k = 1, 2, \cdots$. For $k = 1$, we have $(Q)_+ = \partial$, and $\partial_1 L = [\partial, L]$. This reduces to the identification $\partial_{t_1} = \partial$. For $k = 2$, we have $(Q^2)_+ = L$ and we get the trivial equation $\partial_{t_2} L = 0$. 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 $\partial_{t_3} u = [(Q^3)_+, \partial^2 - u]$ which is the Korteweg–de Vries equation:

$$\boxed{4\partial_{t_3} u = \partial^3 u - 6u(\partial u)}$$

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 we get trivial equations).

Bibliography

[1] J. Liouville [1809-1882]. *"Note sur l'intégration des équations différentielles de la dynamique"*. Journal de Mathématiques (Journal de Liouville) T. XX (1855), p. 137.

[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 non linear equations of evolution and solitary waves"*. Comm. Pure Appl. Math. 21 (1968), p. 467.

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

[5] I.M. Gelfand, L.A. Dickey *"Fractional powers of operators and Hamiltonian systems"*. Funct. Anal. Appl. 10 (1976) 259.

[6] 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.

[7] 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.

[8] 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.

[9] 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.

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

[11] O.Babelon, D. Bernard, M. Talon, **Introduction to Classical Integrable Systems.** Cambridge University Press (2003).

Chapter 2

Solution by analytical methods

We present the general ideas for the solution of Lax equations when a spectral parameter is present. The method uses the geometry of the spectral curve and its complex analysis.

2.1 The spectral curve.

Let us consider a $N \times N$ Lax matrix $L(\lambda)$, depending rationally on a spectral parameter $\lambda \in \mathbb{C}$ with poles at points ϵ_k

$$L(\lambda) = L_0 + \sum_k L_k(\lambda) \tag{2.1}$$

As before L_0 a constant matrix is independent of λ and $L_k(\lambda)$ is the polar part of $L(\lambda)$ at ϵ_k, i.e.

$$L_k = \sum_{r=-n_k}^{-1} L_{k,r}(\lambda - \epsilon_k)^r$$

The analytical method of solution of integrable systems is based on the study of the eigenvector equation:

$$\boxed{(L(\lambda) - \mu \mathbf{1})\, \Psi(\lambda, \mu) = 0} \tag{2.2}$$

where $\Psi(\lambda, \mu)$ is the eigenvector with eigenvalue μ. The characteristic equation for the eigenvalue problem (2.2) is:

$$\boxed{\Gamma \quad : \quad \Gamma(\lambda, \mu) \equiv \det(L(\lambda) - \mu\ \mathbf{1}) = 0} \tag{2.3}$$

This defines an algebraic curve in $\mathbb{C}^2$ which is called the spectral curve. A point on Γ is a pair (λ, μ) satisfying eq.(2.3). Since the Lax equation $\dot{L} = [M, L]$ is isopsectral, this curve is independent of time.

If N is the dimension of the Lax matrix, the equation of the curve is of the form:

$$\Gamma \quad : \quad \Gamma(\lambda,\mu) \equiv (-\mu)^N + \sum_{q=0}^{N-1} r_q(\lambda)\mu^q = 0 \tag{2.4}$$

The coefficients $r_q(\lambda)$ are polynomials in the matrix elements of $L(\lambda)$ and therefore have poles at ϵ_k. The coefficients of these rational functions are independent of time.

From eq.(2.4), 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 $\Gamma(\lambda,\mu)=0$. By definition μ_i are the eigenvalues of $L(\lambda)$.

Our goal 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 consisting in similarity transformations by constant invertible matrices. 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.

$$L_0 = \lim_{\lambda\to\infty} L(\lambda) = \mathrm{diag}(a_1,\cdots,a_N) \tag{2.5}$$

We assume for simplicity that all the a_i's are different. Then on the spectral curve, we have N points above $\lambda=\infty$:

$$Q_i \equiv (\lambda=\infty, \mu_i = a_i)$$

In the gauge (2.5) 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.

2.2 Riemann surfaces.

2.2.1 Desingularisation

A Riemann surface is a compact analytic variety of dimension one. This means that there is a covering by open neighborhoods U_i and local homeomorphisms mapping them to the open disks $|z_i|<1$. On the intersection $U_i\cap U_j$ the local parameters z_i and z_j are related by an analytic bijection.

We will deal however with curves in $\mathbb{C}^2$ as in eq.(2.4). To relate it to the abstract definition we have to find around any point of Γ a local analytic parameter z.

A point λ_0,μ_0 is *regular* if none of the derivatives of $\Gamma(\lambda,\mu)$ vanish at that point.

$$\partial_\lambda\Gamma(\lambda_0,\mu_0)\neq 0, \quad \partial_\mu\Gamma(\lambda_0,\mu_0)\neq 0$$

Around such a point we can choose the local parameter z as either $\lambda - \lambda_0$ or $\mu - \mu_0$ and use the equation $\Gamma(\lambda, \mu) = 0$ to express the other one analytically in terms of z.

A point λ_0, μ_0 is a *branch point* if one of the derivatives of $\Gamma(\lambda, \mu)$ vanish at that point, but not both. Let us assume

$$\partial_\lambda \Gamma(\lambda_0, \mu_0) \neq 0, \quad \partial_\mu \Gamma(\lambda_0, \mu_0) = 0$$

Around such a point the curve looks like (for a branch point of order 2)

$$\Gamma(\lambda, \mu) \simeq \partial_\lambda \Gamma(\lambda_0, \mu_0)(\lambda - \lambda_0) + \frac{1}{2}\partial_\mu^2 \Gamma(\lambda_0, \mu_0)(\mu - \mu_0)^2 + \cdots$$

Clearly, one cannot use $\lambda - \lambda_0$ as a local parameter because then $\mu - \mu_0 \simeq a\sqrt{\lambda - \lambda_0} + \cdots$ is not analytic at $\lambda = \lambda_0$. However choosing $z = \mu - \mu_0$ is perfectly legal.

A point λ_0, μ_0 is a *singular point* if both derivatives of $\Gamma(\lambda, \mu)$ vanish at that point

$$\partial_\lambda \Gamma(\lambda_0, \mu_0) = 0, \quad \partial_\mu \Gamma(\lambda_0, \mu_0) = 0$$

An example of such a point is given by the curve

$$\mu^2 = a\lambda^2 + b\lambda^3$$

To give a meaning to the singular point we perform a birational transformation

$$\lambda = z, \quad \mu = zy \tag{2.6}$$

which can be inverted rationaly as long as $(\lambda, \mu) \neq (0, 0)$

$$z = \lambda, \quad y = \frac{\mu}{\lambda}$$

Under the transformation eq.(2.6) the curve becomes

$$y^2 = a + bz \tag{2.7}$$

Now, instead of the singular point $(0, 0)$, we get *two regular* points $(z = 0, y = \pm\sqrt{a})$ which are mapped to the singuar point $(0, 0)$ by eq.(2.6). We say that we have resolved (or blown up) the singularity by performing the birational transformation. It is always the desingularized curve like eq.(2.7) that we must consider to give a meaning to a singular point. We can use z as a local parameter on the desingularized curve.

Finally a special care must be taken for the points at ∞ where λ and (or) μ become infinite. Then we set

$$\lambda = 1/x, \quad \mu = 1/y \tag{2.8}$$

this brings it to the origin $(0, 0)$. In general we get a singular point and we have to blow it up with the above procedure.

Important examples are the cases of hyperelliptic curves. They are defined by

$$\mu^2 = P_{2n+2}(\lambda), \quad \text{or} \quad \mu^2 = P_{2n+1}(\lambda) \tag{2.9}$$

where $P_{2n+2}(\lambda)$ and $P_{2n+1}(\lambda)$ are polynomials of degree $2n+2$ and $2n+1$ respectively.

$$P_{2n+2}(\lambda) = a\lambda^{2n+2} + b\lambda^{2n+1} + \cdots, \quad \text{or} \quad P_{2n+1}(\lambda) = a\lambda^{2n+1} + b\lambda^{2n} + \cdots$$

To understand the point at ∞, we perform the transformation eq.(2.8). We get

$$y^2(a + bx + \cdots) = x^{2n+2}, \quad \text{or} \quad y^2(a + bx + \cdots) = x^{2n+1}$$

The point $(x, y) = (0, 0)$ is now a singular point. To blow it up we set

$$\begin{cases} x = x' \\ y = x'^{n+1}y' \end{cases} \quad \text{or} \quad \begin{cases} x = x' \\ y = x'^{n}y' \end{cases}$$

we obtain

$$y'^2 = \frac{1}{(a + bx' + \cdots)}, \quad \text{or} \quad y'^2 = \frac{x'}{(a + bx' + \cdots)}$$

In the first case the singularity has been resolved into two regular points, the local parameter can be taken to be $z = x'$. In the second case it has been resolved into a branch point, the local parameter can be taken to be $z = y'$. Summarizing, at infinity we have

$$\begin{cases} \lambda = \frac{1}{z} \\ \\ \mu = \pm\frac{\sqrt{a}}{z^{n+1}} + \cdots \end{cases} \quad \text{or} \quad \begin{cases} \lambda = \frac{1}{a}\frac{1}{z^2} + \cdots \\ \\ \mu = \frac{1}{a^n}\frac{1}{z^{2n+1}} + \cdots \end{cases}$$

Notice that the number of branch points in both cases is $2n + 2$.

2.2.2 Riemann-Hurwitz theorem.

This is the tool to compute the genus of a Riemann surface. Given a triangulation of the surface, the Euler-Poincaré characteristic is defined by

$$\boxed{\chi = F - A + V}$$

where F is the number of faces of the triangulation, A the number of edges and V the number of vertices. The Euler-Poincaré characteristic is a topological invariant. It is related to the genus by the formula

$$\boxed{\chi = 2 - 2g}$$

For instance in the case of the sphere, there is a triangulation with 8 faces, 12 edges and 6 vertices. Hence

$$\chi_0 = 2, \quad g_0 = 0$$

If Γ is a branched covering of Γ_0, we can lift a triangulation of Γ_0 to Γ. Let us choose the triangulation of Γ_0 such that the projections of the branch points are among its vertices.

Let F_0, A_0, V_0 be the number of faces, edges and vertices of this triangulation. Let N be the number of sheets of the covering $\Gamma \to \Gamma_0$. When we lift the triangulation of Γ_0 to Γ, we get a triangulation with $F = NF_0$ faces, $A = NA_0$ edges and $V = NV_0 - B$ vertices where B is the total index of the branch points. At a branch point the index is the number of sheets that coalesce minus one. Hence we find

$$\chi = N\chi_0 - B$$

or

$$\boxed{2g - 2 = N(2g_0 - 2) + B}$$

This is the Riemann-Hurwitz formula. Let us apply it to compute the genus of the hyperelliptic curves given by eqs.(2.9). They are coverings of the Riemann sphere with two sheets. Each branch point is of index 1. In both cases we have seen that there are exactly $2n + 2$ branch points. Therefore $2g - 2 = 2(2 \times 0 - 2) + 2n + 2$, that is $g = n$.

2.2.3 Riemann-Roch theorem.

This is the tool to count the number of meromorphic functions on a Riemann surface. A Divisor is a formal sum of points with multiplicites.

$$D = n_1 P_1 + n_2 P_2 + \cdots + n_r P_r, \quad n_i \in \mathbb{Z}$$

The degree of the divisor is

$$\deg D = \sum_i n_i$$

If f is a meromorphic function, we denote by (f) its divisor of zeroes and poles. Let $\mathcal{M}(D)$ the space of meromorphic functions whose is such that

$$(f) \geq D$$

that is $\mathcal{M}(D)$ is the det of meromorphic funtions whose order of the poles are at most the one specified by D and the order of zeroes are at least the one specified by D.

The Riemann-Roch theorem asserts that

$$\boxed{\dim \mathcal{M}(-D) = i(D) + \deg D - g + 1}$$

where $i(D)$ is the dimension of the space of meromorphic differentials ω such that

$$(\omega) \geq D$$

There are two cases where the theorem leads to simple answer. If $D = 0$, then $\mathcal{M}(D)$ is the set of *holomorphic* functions on the Riemann surface. We know that the only such function is the constant. Hence $\dim \mathcal{M}(D) = 1$. Similarly $i(D)$ is the dimension of the

space of holomorphic differentials. The theorem then give $i(D) = g$. We therefore get the very important result

$$\boxed{\text{dim (holomorphic differentials)} = g}$$

If $D < 0$, then dim $\mathcal{M}(D) = 0$ and this allows to count the meromorphic differentials

$$\boxed{i(D) = -\text{deg } D + g - 1, \quad \text{deg } D < 0}$$

Notice that it is the degree of the divisor which is relevant. The freedom gained by adding a pole is compensated by the restriction of adding a zero.

The next simple case is when deg $D \geq g$ where generically $i(D) = 0$. For instance if

$$D = P_1 + P_2 + \cdots + P_g$$

then $i(D)$ is the dimension of the space of holomorphic differentials with zeroes at the point of D. To construct such differentials we expand them on a basis ω_i of the g holomorphic differentials and try to impose the conditions

$$\sum_i \omega_i(P_j)c_i = 0, \quad j = 1, \cdots, g.$$

This system in general has no solution because for a generic set of g points P_j we have $\det \omega_i(P_j) \neq 0$. Hence $i(D) = 0$. Then the theorem gives

$$\boxed{\text{dim } \mathcal{M}(-D) = \text{deg } D - g + 1, \quad \text{deg } D \geq g}$$

The difficult case is when $0 < \text{deg } D < g$ and a careful investigation is needed.

2.2.4 Jacobi variety and Theta functions.

Consider a Riemann surface Γ of genus g. Let a_i, b_i be a basis of cycles on Γ with canonical intersection matrix $(a_i \cdot a_j) = (b_i \cdot b_j) = 0$, $(a_i \cdot b_j) = \delta_{ij}$. One can 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. We get 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 Γ are called *Abelian differentials of first kind*. They form a space of dimension g over the complex numbers. 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). 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 :

$$\oint_{a_j} \omega_i = \delta_{ij}. \tag{2.10}$$

The matrix of b-periods is then defined as the matrix $\mathcal{B}$ with matrix elements :

$$\mathcal{B}_{ij} = \oint_{b_i} \omega_j \tag{2.11}$$

Taking the example of an hyperelliptic surface $y^2 = P_{2g+1}(x)$ where $P(x)$ is a polynomial of degree $2g+1$, a basis of regular Abelian differentials is provided by the forms

$$\omega_j = \frac{x^{j-1}dx}{y} \quad \text{for } j = 1, \cdots, g$$

These forms are 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 $x^{j-1}dx/y = (2b^{j-1}/a)(1 + \cdots)dy$ which is regular. At ∞ we take $x' = 1/x$ and $y' = y/x^{(g+1)}$ so that $y'^2 = ax' + \cdots$ and $x^{j-1}dx/y = by'^{2(g-j)}dy'$ which is regular for $1 \leq j \leq g$ since y' is the local parameter. Of course these forms are unnormalized.

Similarly *Abelian differentials of second kind* are meromorphic differentials with poles of order greater than 2. Given a point p on Γ, there exists a unique normalized (all a–periods vanish) Abelian differential of second kind whose only singularity is a pole of second order at p. Indeed applying the Riemann-Roch theorem with deg $D = -2$, we find $i(D) = g+1$. The g comes from the first kind differentials which are included in this counting. Adding a proper combination of differentials of first kind one can always insure that all a–periods of the second kind differential vanish and the differential becomes uniquely determined.

We define *Abelian differentials of third kind* as general meromorphic differentials with first order poles whose sum of residues vanish (this condition results from the Cauchy theorem). Given two points p and q there exists a unique normalized (all a–periods vanish) 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.

On a Riemann surface on which we have chosen canonical cycles, there is a pairing between meromorphic differentials. Namely, let Ω_1 and Ω_2 be two meromorphic differentials on Σ. The pairing $(\Omega_1 \bullet \Omega_2)$ is defined by integrating them along the canonical cycles as follows:

$$(\Omega_1 \bullet \Omega_2) = \sum_{i=1}^{g} \left(\oint_{a_j} \Omega_1 \oint_{b_j} \Omega_2 - \oint_{a_j} \Omega_2 \oint_{b_j} \Omega_1 \right)$$

The Riemann bilinear identity expresses this quantity in terms of residues:

Proposition 6 *Let g_1 be a function defined on the Riemann surface, cut along the canonical cycles, and such that $dg_1 = \Omega_1$. We have:*

$$(\Omega_1 \bullet \Omega_2) = 2i\pi \sum_{\text{poles}} \text{res}(g_1 \Omega_2) \tag{2.12}$$

Corollary 2 *The matrix of b–periods $\mathcal{B}$ is symmetric.*

Corollary 3 *Let Ω_2 be 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. Let $\Omega_1 = \omega_k$ be a normalized holomorphic differential expanded as*

$$\omega_k = (\sum_{i=0}^{\infty} c_i z^i) dz$$

around $z = 0$. One has:

$$\oint_{b_k} \Omega_2 = 2\pi i \frac{c_{n-2}}{n-1}$$

By linearity, if $\Omega^{(P)}$ is a normalized second kind differential with principal part $dP(z)$ where $P(z) = \sum_{n=1}^{N} p_n z^{-n}$, then we have

$$\frac{1}{2i\pi} \oint_{b_k} \Omega^{(P)} = -\text{Res}\,(\omega_k P) \tag{2.13}$$

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:

$$\rho_k(\mathcal{D}) = \sum_i \int_{\gamma_i} \omega_k, \quad k = 1 \cdots g$$

where the ω_i are holomorphic differentials. Such sums are called Abel sums. If the paths are homotopically deformed these integrals remain constant by the Cauchy theorem. If one makes a loop around a_k then $\rho_l \to \rho_l + \delta_{kl}$. If one makes a loop around b_k, then $\rho_l \to \rho_l + \mathcal{B}_{kl}$. Hence the maps ρ_k give a well–defined point on the torus:

$$J(\Gamma) = \mathbb{C}^g / \left(\mathbb{Z}^g + \mathcal{B}\mathbb{Z}^g\right) \tag{2.14}$$

where $\mathcal{B}$ is the matrix of the b-periods. If one permutes independently the points p_i and q_i the point in the torus does not change. To see it, let the paths γ_1' connect q_1 to p_2, γ_2' connect q_2 to p_1 and σ connect 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.

The theorems of Abel and Jacobi state that the point on the torus $J(\Gamma)$ characterizes the divisor $\mathcal{D}$ up to equivalence.

Theorem 4 (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 5 (Jacobi) *For any point $u \in J(\Gamma)$ 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 Σ such that $\rho_k(\mathcal{D} - \mathcal{D}_0)$ maps to u. Moreover for generic u the divisor $\mathcal{D}$ is unique.*

One can embed the Riemann surface Γ into its Jacobian $J(\Gamma)$ by the Abel map. Namely, choose a point $q_0 \in \Gamma$ and define the vector $\mathcal{A}(p)$ with coordinates $\mathcal{A}_k(p)$ modulo the lattice of periods:

$$\mathcal{A} : \Gamma \longmapsto J(\Gamma) \tag{2.15}$$

$$\mathcal{A}_k(p) = \int_{q_0}^{p} \omega_k \tag{2.16}$$

Clearly, the Abel map depends on the point q_0. But changing this point just amounts to a translation in $J(\Gamma)$.

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:

$$\theta(z_1, \ldots, z_g) = \sum_{m \in \mathbb{Z}^g} e^{2\pi i (m,z) + \pi i (\mathcal{B}m, m)}. \tag{2.17}$$

Since the series is convergent, it defines 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 $z \in \mathbb{C}^g$

$$\begin{aligned} \theta(z + l) &= \theta(z) \\ \theta(z + \mathcal{B}l) &= \exp[-i\pi(\mathcal{B}l, l) - 2i\pi(l, z)]\theta(z) \end{aligned} \tag{2.18}$$

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 Γ into $J(\Gamma)$ with the divisor of the theta function.

Theorem 6 (Riemann) *Let $w = (w_1, \cdots, w_g) \in \mathbb{C}^g$ arbitrary. Either the function $\theta(\mathcal{A}(p) - w)$ vanishes identically for $p \in \Gamma$ or it has exactly g zeroes $p_1, \cdots, p_g$ such that:*

$$\mathcal{A}(p_1) + \cdots + \mathcal{A}(p_g) = w - \mathcal{K} \tag{2.19}$$

where $\mathcal{K}$ is the so–called vector of Riemann's constants, depending on the curve Γ and the point q_0 but independent of w.

2.3 Genus of the spectral curve.

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:

$$2g - 2 = N(2g_0 - 2) + B \tag{2.20}$$

where B is the branching index of the covering. Let us assume for simplicity that the branch points are all of order 2. To compute B we observe that this is the number of values of λ where $\Gamma(\lambda, \mu)$ has a double root in μ. This is also the number of zeroes of $\partial_\mu \Gamma(\lambda, \mu)$ on the surface $\Gamma(\lambda, \mu) = 0$. But $\partial_\mu \Gamma(\lambda, \mu)$ is a meromorphic function on Γ, and therefore the number of its zeroes is equal to the number of its poles and 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. Around such a point the curve reads

$$\left(\mu - \frac{l_1}{(\lambda - \epsilon_k)^{n_k}} + \cdots\right) \cdots \left(\mu - \frac{l_N}{(\lambda - \epsilon_k)^{n_k}} + \cdots\right) = 0$$

where l_j are the eigenvalues of $L_{k,-n_k}$ that are assumed all distinct. When λ tends to ϵ_k, μ tends to infinity. We bring this point to the origin by setting

$$\mu = \frac{1}{y}, \quad \lambda - \epsilon_k = z$$

Around the point $(0, 0)$ the curve reads

$$(l_1 y - z^{n_k} + \cdots) \cdots (l_N y - z^{n_k} + \cdots) = 0$$

Clearly the point $(0, 0)$ is a singular point. To desingularise the curve, we set $y = z^{n_k} y'$ and we find

$$\left(l_1 y' - 1 + \cdots\right) \cdots \left(l_N y' - 1 + \cdots\right) = 0$$

The singular point has blown up to the N points $z = 0, y' = l_j^{-1}$. Hence, above a pole ϵ_k, we have N branches of the form

$$\mu_j = \frac{l_j}{z^{n_k}} + \cdots, \quad \lambda - \epsilon_k = z$$

On such a branch we have $\partial_\mu \Gamma(\lambda, \mu)|_{(\lambda, \mu_j(\lambda))} = \prod_{i \neq j} (\mu_j(\lambda) - \mu_i(\lambda))$ which thus has a pole of order $(N-1)n_k$. Summing on all branches the total order of the poles over ϵ_k is $N(N-1)n_k$. Summing on all poles ϵ_k of $L(\lambda)$ we see that the total branching index is $B = N(N-1)\sum_k n_k$. This gives:

$$\boxed{g = \frac{N(N-1)}{2} \sum_k n_k - N + 1}$$

2.4 Dimension of the reduced phase space.

For consistency of the method it is important to observe that the genus is related to the dimension of the phase space and to the number of action variables occuring as independent parameters in eq.(2.4), which should also be half the dimension of phase space. The original phase space $\mathcal{M}$ is the coadjoint orbit

$$L = L_0 + \sum_k L_k, \quad L_k = (g^{(k)}\, A_-^{(k)}\, g^{(k)-1})_-$$

Let us compute its dimension. 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 the even number:

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

The reduced phase space $\mathcal{M}_{\text{reduced}}$ is obtained by performing the quotient by the residual global diagonal gauge transformations.

Proposition 7 *The reduced phase space $\mathcal{M}_{\text{reduced}}$ has dimension $2g$ and there are g proper action variables in eq.(2.4).*

Proof. The residual global gauge transformations act by diagonal matrices as $g_k \to d g_k$, or $L(\lambda) \to d L(\lambda) d^{-1}$. This preserves the diagonal form of L_0. The orbits of this action are of dimension $(N-1)$, since the identity does not act. The action of this diagonal group is Hamiltonian and its generators are given just below. The phase space $\mathcal{M}_{\text{reduced}}$ is obtained by 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. As a result, the dimension of the phase space is reduced by $2(N-1)$, yielding:

$$\boxed{\dim \mathcal{M}_{\text{reduced}} = (N^2 - N) \sum_k n_k - 2(N-1) = 2g}$$

Let us now count the number of independent coefficients in eq.(2.4). It is clear that $r_j(\lambda)$ is a rational function of λ. The value of $r_j(\lambda)$ at ∞ is known since $\mu_j \to a_j$. Note that $r_j(\lambda)$ is the symmetrical function $\sigma_j(\mu_1, \cdots, \mu_N)$ where μ_i are the eigenvalues of $L(\lambda)$. Above $\lambda = \epsilon_k$, they can be written as

$$\mu_j = \sum_{n=1}^{n_k} \frac{c_n^{(j)}}{(\lambda - \epsilon_k)^n} + \text{regular} \tag{2.21}$$

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. We see that $r_j(\lambda)$ has a pole of order jn_k at $\lambda = \epsilon_k$, and so can be expressed on $j\sum_k n_k$ parameters, namely the coefficients of all these poles. Summing over j we have altogether a pool of $\frac{1}{2}N(N+1)\sum_k n_k$ parameters. They are not all independent however, because in eq.(2.21) the coefficients $c_n^{(j)}$ are non dynamical. This implies that the n_k highest order terms in $r_j(\lambda)$ are fixed and yields Nn_k constraints on the coefficients of $r_j(\lambda)$. 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 symplectic quotient by the action of constant diagonal matrices. We assume that the system is equipped with the Poisson bracket (1.21). 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 $r_j(\lambda)$ in eq.(2.4). We show that they are the generators of the diagonal action. First we have:

$$\begin{aligned}\mathrm{Res}_{\lambda=\infty}\mathrm{Tr}\,(L^n(\lambda))d\lambda &= n\,\mathrm{Res}_{\lambda=\infty}\mathrm{Tr}\,(L_0^{n-1}\sum_k L_k(\lambda))d\lambda \\ &= n\,\mathrm{Res}_{\lambda=\infty}\mathrm{Tr}\,(L_0^{n-1}L(\lambda))d\lambda \qquad (2.22)\end{aligned}$$

since all $L_k(\lambda)$ are of order $1/\lambda$ at ∞. Using the Poisson bracket 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. Setting

$$\mu_j(\lambda) = a_j + \frac{b_j}{\lambda} + \cdots \qquad (2.23)$$

around the point $Q_j = (\infty, a_j)$, we have $H_n = \sum_j a_j^{n-1}b_j$. So, both a_i (by definition) and b_i are non dynamical. On the functions $r_j(\lambda)$ this implies that their expansion at infinity starts as $r_j(\lambda) = r_j^{(0)} + \frac{r_j^{(-1)}}{\lambda} + \cdots$, with $r_j^{(0)}$ and $r_j^{(-1)}$ non dynamical. Hence when the system is properly reduced we are left with exactly g action variables.

The constraints eqs.(2.21,2.23) can be summarized in an elegant way. Introduce the differential δ with respect to the dynamical moduli. Then our constraints mean that the differential $\delta\mu d\lambda$ is regular everywhere on the spectral curve because the coefficients of the various poles being non dynamical, they are killed by δ:

$$\boxed{\delta\mu\; d\lambda = \text{ holomorphic}}$$

Since the space of holomorphic differentials is of dimension g, the right hand side of the above equation is spanned by g parameters which shows that the space of dynamical moduli is g dimensional. Notice that these action variables are coefficients in the pole expansions of the functions $r_j(\lambda)$, and thus appear *linearly* in the equation of Γ. Hence eq.(2.4) can be written in the form

$$\boxed{\Gamma:\ R(\lambda,\mu)\equiv R_0(\lambda,\mu)+\sum_{j=1}^{g}R_j(\lambda,\mu)H_j=0} \tag{2.24}$$

2.5 The eigenvector bundle.

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, i.e.

$$\psi_1(P)=1,\quad \text{at any point}\quad P\in\Gamma.$$

It is then clear that the $\psi_j(P)$ depend locally analytically on P. As a matter of fact:

Proposition 8 *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. Let $\hat{\Delta}(\lambda,\mu)$ be the matrix of cofactors of $(L(\lambda)-\mu\mathbf{1})$, which, by definition, is such that $(L(\lambda)-\mu\mathbf{1})\hat{\Delta}=\Gamma(\lambda,\mu)\mathbf{1}$. Therefore at $P=(\lambda,\mu)\in\Gamma$, each column of the matrix $\hat{\Delta}$ is proportional to the eigenvector $\Psi(P)$. Hence we have

$$\psi_i(P)=\frac{\Delta_{ij}(\lambda,\mu)}{\Delta_{1j}(\lambda,\mu)}$$

which is a meromorphic function on Γ.

In fact the matrix $\hat{\Delta}(P)$ 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 $\hat{\Delta}(P)$ 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 normalized eigenvector by simply counting the number of zeroes of the first minor.

Proposition 9 *We say that the vector $\Psi(P)$ possesses a pole if one of its components has a pole. The number of poles of the normalized vector $\Psi(P)$ is:*

$$\boxed{m = g + N - 1} \tag{2.25}$$

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 $\hat{\Psi}(\lambda)$ is the matrix of eigenvectors of $L(\lambda)$ defined as follows:

$$\widehat{\Psi}(\lambda) = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \psi_2(P_1) & \psi_2(P_2) & \cdots & \psi_2(P_N) \\ \vdots & \vdots & \vdots & \vdots \\ \psi_N(P_1) & \psi_N(P_2) & \cdots & \psi_N(P_N) \end{pmatrix} \tag{2.26}$$

where the points P_i are the N points above λ. 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. 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 $\lambda - \lambda_0$. Hence $W(\lambda)$ has a simple zero for values of λ corresponding to a branch–point of the covering, therefore $m = B/2$. Recall that from eq.(2.20) the number of branch points is $B = 2(N + g - 1)$.

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.(2.5). 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 10 *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 one has to divide V_k by V_k^1. So we get:

$$\Psi(P)|_{P\sim Q_1}=\begin{pmatrix}1\\ O(1/\lambda)\\ \vdots\\ \vdots\\ \vdots\\ \vdots\\ O(1/\lambda)\end{pmatrix},\quad \Psi(P)|_{P\sim Q_k}=\begin{pmatrix}1\\ O(1)\\ \vdots\\ O(\lambda)\\ O(1)\\ \vdots\\ O(1)\end{pmatrix},\quad k\geq 2 \tag{2.27}$$

where $O(\lambda)$ is the announced pole of the k^{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 independent of time, and is really part of the choice of the gauge condition. These poles 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. We call it the dynamical divisor. 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 11 *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 $\mathbb{C}^N$:

$$\Psi(P)=\begin{pmatrix}1\\ \psi_2(P)\\ \vdots\\ \psi_N(P)\end{pmatrix}$$

Consider the matrix $\widehat{\Psi}(\lambda)$ whose columns are the vectors $\psi(P_i)$ with $P_i=(\lambda,\mu_i)$ are the N points above λ, cf. eq.(2.26). This matrix depends on the ordering of the columns, i.e., on the ordering of the points P_i. However, the matrix

$$L(\lambda)=\widehat{\Psi}(\lambda)\cdot\widehat{\mu}\cdot\widehat{\Psi}^{-1}(\lambda) \tag{2.28}$$

does not depend on this ordering and is a well defined function on the base curve. Here $\widehat{\mu}$ is the diagonal matrix $\widehat{\mu}=\operatorname{diag}(\mu_1,\cdots,\mu_N)$.

It should be emphasized however that $\Psi(P)$ is defined up to normalizations $\psi_k \to d_k\psi_k$. In fact the normalization ambiguity of the ψ_k translates into left multiplication of the vector $\Psi(P)$ by a *constant* diagonal matrix $d = \mathrm{diag}\,(1, d_2, \cdots, d_N)$. 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.

2.6 Separated variables.

We show here in the example of the Jaynes-Cummings-Gaudin model that the coordinates of the dynamical poles of the eigenvectors form a set of separated variables. A general proof exists but it is cumbersome.

Let us write the Lax matrix as

$$L(\lambda) = \begin{pmatrix} A(\lambda) & B(\lambda) \\ C(\lambda) & -A(\lambda) \end{pmatrix}$$

The spectral curve reads $\det(L(\lambda) - \mu) = 0$ or

$$\Gamma : \ \mu^2 - A^2(\lambda) - B(\lambda)C(\lambda) = 0$$

Clearly we have

$$A^2(\lambda) + B(\lambda)C(\lambda) = \frac{Q_{2n+2}(\lambda)}{\prod_j (\lambda - \epsilon_j)^2}$$

where $Q_{2n+2}(\lambda)$ is a polynomial of degree $2n+2$. Defining $y = \mu \prod_j (\lambda - \epsilon_j)$, the equation of the curve becomes

$$y^2 = Q_{2n+2}(\lambda)$$

which is an hyperelliptic curve of genus n. The dimension of the phase space of the model is $2(n+1)$. However, we have to reduce that model by the action of the group of conjugation by diagonal matrices which in our case is of dimension 1. Hence we confirm that

$$g = \frac{1}{2}\dim \mathcal{M}_{\text{reduced}} = n$$

The generator of the group is the Hamiltonian H_n as shown by eq.(1.38). We can compute easily

$$\delta\mu d\lambda = \frac{1}{\sqrt{Q_{2n+2}(\lambda)}} \prod_{j=0}^{n-1} (\lambda - \epsilon_j) \left(\frac{4}{g^2}\delta H_n + \frac{2}{g^2} \sum_{j=0}^{n-1} \frac{\delta H_j}{\lambda - \epsilon_j} \right)$$

The coefficients of δH_j are polynomials of degree $n - 1 = g - 1$, hence give rise to holomorphic differentials. The coefficient of δH_n is a polynomial of degree $n = g$ and lead to a singular differential. It is absent in the reduced model where $\delta H_n = 0$.

At each point of the spectral curve, we can solve the equation

$$\Big(L(\lambda)-\mu\Big)\Psi=0$$

Normalizing the second component (instead of the first for later convenience) of Ψ to be 1, we find

$$\Psi=\begin{pmatrix}\psi_1\\1\end{pmatrix},\quad \psi_1=\frac{A(\lambda)+\mu}{C(\lambda)}$$

Hence the poles of Ψ are located above the zeroes of $C(\lambda)$. Note that if $C(\lambda_k)=0$, then the points on Γ above λ_k have coordinates $\mu_k=\pm A(\lambda_k)$. The pole of Ψ is at the point $\mu_k=A(\lambda_k)$, since at the other point $\mu_k=-A(\lambda_k)$ the numerator of ψ_1 has a zero.

Recalling that

$$C(\lambda)=\frac{2\bar{b}}{g}+\sum_{j=0}^{n-1}\frac{s_j^+}{\lambda-\epsilon_j}=\frac{2\bar{b}}{g}\frac{\prod_{k=1}^{n}(\lambda-\lambda_k)}{\prod_{j=0}^{n-1}(\lambda-\epsilon_j)} \tag{2.29}$$

This shows that indeed the eigenvector has $n=g$ dynamical poles.

At infinity, we have two points

$$Q_\pm:\mu=\pm\frac{1}{g^2}(2\lambda-\omega)(1+O(\lambda^{-2}))$$

Remembering that

$$\begin{aligned}A(\lambda)&=\frac{1}{g^2}(2\lambda-\omega)+O(\lambda^{-1})\\C(\lambda)&=\frac{2\bar{b}}{g}+O(\lambda^{-1})\end{aligned}$$

we find the following behavior of the function ψ_1 at the two points $Q_\pm$

$$\begin{aligned}Q_+&:\quad \psi_1=\frac{1}{g\bar{b}}(2\lambda-\omega)+O(\lambda^{-1})\\Q_-&:\quad \psi_1=O(\lambda^{-1})\end{aligned}$$

showing that the eigenvector has a pole at Q_+ and a zero at Q_- in agreement with the general result.

We can now compute the symplectic form in the coordinates λ_k,μ_k. From the constraint $(s_j^z)^2+s_j^+s_j^-=s^2$ we we can eliminate s_j^-. Remembering the Poisson bracket $\{s_j^z,s_j^+\}=is_j^+$, we can write the symplectic form as

$$\Omega=-i\delta b\wedge\delta\bar{b}+i\sum_j\frac{\delta s_j^+}{s_j^+}\wedge\delta s_j^z$$

From eq.(2.29), we see that s_j^+ is the residue of $C(\lambda)$ at $\lambda = \epsilon_j$, so that

$$s_j^+ = \frac{2\bar{b}}{g} \frac{\prod_k (\epsilon_j - \lambda_k)}{\prod_{i \neq j} (\epsilon_j - \epsilon_i)}$$

it follows that

$$\frac{\delta s_j^+}{s_j^+} = \frac{\delta \bar{b}}{\bar{b}} + \sum_k \frac{\delta \lambda_k}{\lambda_k - \epsilon_j}$$

therefore

$$\Omega = -i\delta b \wedge \delta \bar{b} + i \frac{\delta \bar{b}}{\bar{b}} \wedge \sum_j \delta s_j^z + i \sum_k \sum_j \frac{\delta \lambda_k \wedge \delta s_j^z}{\lambda_k - \epsilon_j}$$

But

$$\delta A(\lambda_k) = \frac{2}{g^2} \delta \lambda_k + \sum_j \frac{\delta s_j^z}{\lambda_k - \epsilon_j} - \frac{s_j^z}{(\lambda_k - \epsilon_j)^2} \delta \lambda_k$$

Therefore

$$\begin{aligned} \Omega &= -i\delta b \wedge \delta \bar{b} + i \frac{\delta \bar{b}}{\bar{b}} \wedge \sum_j \delta s_j^z + i \sum_k \delta \lambda_k \wedge \delta A(\lambda_k) \\ &= i \frac{\delta \bar{b}}{\bar{b}} \wedge \left[\delta(b\bar{b}) + \sum_j \delta s_j^z \right] + i \sum_k \delta \lambda_k \wedge \delta A(\lambda_k) \end{aligned}$$

Finally

$$\boxed{\Omega = i\delta \log \bar{b} \ \wedge \delta H_n + i \sum_k \delta \lambda_k \wedge \delta \mu_k}$$

This shows that the variables λ_k, μ_k are canonically conjugate. Remark that the separated variables are invariant under the diagonal group action

$$\{H_n, \lambda_k\} = 0, \quad \{H_n, \mu_k\} = 0 \tag{2.30}$$

so that they are really coordinates on the reduced phase space.

Let us explain why the variables (λ_k, μ_k) form a set of separated variables. Consider the function

$$S(\{F_j\}, \{\lambda_k\}) = \int_{m_0}^{m} \alpha = \sum_k \int_{\lambda_0}^{\lambda_k} \mu_k d\lambda_k$$

The integration contour is done on the level manifold $F_j = f_j$, and μ is obtained from the spectral curve. Just as in the proof of the Liouville theorem, 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 one depending on only one variable λ_k:

$$S(\{F_j\},\{\lambda_k\}) = \sum_k S_k(\{F_j\},\lambda_k)$$

Since $\sum_k \mu_k d\lambda_k$ is the canonical form, this function is a reduced action and satisfies the Hamilton-Jacobi equation. It depends on n arbitrary constants H_j and is thus the *complete* integral of the Hamilton-Jacobi equation. The variables have been explicitly separated.

2.7 Riemann surfaces and integrability.

Consider a curve in $\mathbb{C}^2$

$$\Gamma:\ R(\lambda,\mu) \equiv R_0(\lambda,\mu) + \sum_{j=1}^{g} R_j(\lambda,\mu) H_j = 0 \tag{2.31}$$

where the H_i are the only dynamical moduli, so that $R_0(\lambda,\mu)$ and $R_i(\lambda,\mu)$ do not contain any dynamical variables. If things are set up so that Γ is of genus g and there are exactly g Hamiltonian H_j, then the curve is completely determined by requiring that it passes through g points (λ_i,μ_i), $i=1,\cdots,g$. Indeed, the moduli H_j are determined by solving the linear system

$$\sum_{j=1}^{g} R_j(\lambda_i,\mu_i) H_j + R_0(\lambda_i,\mu_i) = 0, \quad i=1,\cdots,g \tag{2.32}$$

whose solution is

$$H = -B^{-1}V \tag{2.33}$$

where

$$H = \begin{pmatrix} H_1 \\ \vdots \\ H_i \\ \vdots \\ H_g \end{pmatrix}, \quad B = \begin{pmatrix} R_1(\lambda_1,\mu_1) & \cdots & R_g(\lambda_1,\mu_1) \\ \vdots & & \vdots \\ R_1(\lambda_i,\mu_i) & \cdots & R_g(\lambda_i,\mu_i) \\ \vdots & & \vdots \\ R_1(\lambda_g,\mu_g) & \cdots & R_g(\lambda_g,\mu_g) \end{pmatrix}, \quad V = \begin{pmatrix} R_0(\lambda_1,\mu_1) \\ \vdots \\ R_0(\lambda_i,\mu_i) \\ \vdots \\ R_0(\lambda_g,\mu_g) \end{pmatrix}$$

Here, of course, we assume that generically $\det B \neq 0$.

Theorem 7 *Suppose that the variables (λ_i,μ_i) are separated i.e. they Poisson commute for $i \neq j$:*

$$\{\lambda_i,\lambda_j\} = 0, \quad \{\mu_i,\mu_j\} = 0, \quad \{\lambda_i,\mu_j\} = p(\lambda_i,\mu_i)\delta_{ij} \tag{2.34}$$

Then the Hamiltonians H_i, $i = 1\cdots g$, defined by eq.(2.33) Poisson commute

$$\{H_i,H_j\} = 0$$

Proof. Let us compute

$$\begin{aligned} B_1B_2\{(B^{-1}V)_1,(B^{-1}V)_2\} &= \{B_1,B_2\}(B^{-1}V)_1(B^{-1}V)_2 \\ &\quad -\{B_1,V_2\}(B^{-1}V)_1 - \{V_1,B_2\}(B^{-1}V)_2 + \{V_1,V_2\} \end{aligned}$$

Taking the matrix element i,j of this expression, we get

$$\begin{aligned} \Big(B_1B_2\{(B^{-1}V)_1,(B^{-1}V)_2\}\Big)_{ij} &= \delta_{ij}\sum_{k,l}\{B_{ik},B_{il}\}(B^{-1}V)_k(B^{-1}V)_l \\ -\delta_{ij}\sum_k\{B_{ik},V_i\}(B^{-1}V)_k - \delta_{ij}\sum_l\{V_i,B_{il}\}(B^{-1}V)_l + \delta_{ij}\{V_i,V_i\} &= 0 \end{aligned}$$

where δ_{ij} occurs because the variables are separated. ∎

It can hardly be simpler. The only thing we use is that the Poisson bracket vanishes between different lines of the matrices, and then the antisymmetry. We did not even need to specify the Poisson bracket between λ_i and μ_i. The Hamiltonian are in involution whatever this Poisson bracket is. This is the root of the multihamiltonian structure of integrable systems.

Lax matrices built with the help of coadjoint orbits of loop groups lead to spectral curves of the very special form eq.(2.31) where the H_j are the Poisson commuting Hamiltonians. The coefficients $R_j(\lambda,\mu)$ have a simple geometrical meaning. As we have seen varying the moduli H_i at λ constant one has

$$\delta\mu\, d\lambda = \text{holomorphic} \tag{2.35}$$

Since

$$\delta\mu\, d\lambda = -\sum_j \delta H_j \frac{R_j(\lambda,\mu)}{\partial_\mu R(\lambda,\mu)} d\lambda$$

we see that the coefficients $R_j(\lambda,\mu)$ are in fact the numerators of a basis of holomorphic differentials on Γ:

$$\omega_j = \frac{R_j(\lambda,\mu)}{\partial_\mu R(\lambda,\mu)} d\lambda = \sigma_j(\lambda,\mu)d\lambda \tag{2.36}$$

Define the angles as the images of the divisor (λ_k,μ_k) by the Abel map:

$$\boxed{\theta_j = \sum_k \int^{\lambda_k} \sigma_j(\lambda,\mu)d\lambda}$$

where $\sigma_j(\lambda,\mu)d\lambda$ is any basis of holomorphic differentials. This maps the dynamical divisor $\{\lambda_k,\mu_k\}$ to a point on the Jacobian of Γ .

Theorem 8 *Under the above map, the flows generated by the Hamiltonians H_i are linear on the Jacobian.*

Proof. We want to show that the velocities $\partial_{t_i}\theta_j$ are constants, or

$$\partial_{t_i}\theta_j = \sum_k \partial_{t_i}\lambda_k \; \sigma_j(\lambda_k, \mu_k) = C_{ij}^{ste}$$

One has (no summation over k)

$$\begin{aligned} \partial_{t_i}\lambda_k = \{H_i, \lambda_k\} &= -\{B_{il}^{-1}V_l, \lambda_k\} \\ &= B_{ir}^{-1}\{B_{rs}, \lambda_k\}B_{sl}^{-1}V_l - B_{il}^{-1}\{V_l, \lambda_k\} \\ &= -B_{ik}^{-1}\Big[\{B_{ks}, \lambda_k\}H_s + \{V_k, \lambda_k\}\Big] \end{aligned}$$

where in the last line, we used the separated structure of the matrix B and the vector V. Explicitly ($p(\lambda, \mu) = 1$ in eq.(2.34) for coadjoint orbits)

$$\partial_{t_i}\lambda_k = B_{ik}^{-1}\Big[\partial_\mu R_s(\lambda_k, \mu_k)H_s + \partial_\mu R_0(\lambda_k, \mu_k)\Big]$$

or

$$\boxed{\partial_{t_i}\lambda_k = B_{ik}^{-1}\partial_\mu R(\lambda_k, \mu_k)} \tag{2.37}$$

which we rewrite as (remember that $B_{kj} = R_j(\lambda_k, \mu_k)$)

$$\partial_{t_i}\lambda_k \; \frac{R_j(\lambda_k, \mu_k)}{\partial_\mu R(\lambda_k, \mu_k)} = B_{ik}^{-1}B_{kj} \tag{2.38}$$

Recalling eq.(2.36), we have shown

$$\sum_k \partial_{t_i}\lambda_k \; \sigma_j(\lambda_k, \mu_k) = \delta_{ij} \tag{2.39}$$

■

There are cases where the condition eq.(2.35) is modified. This happens for instance when coadjoint orbits are non generic, or as we will see in the next Chapter when the Lax matrix belongs to the group G^* instead of the Lie algebra $\mathcal{G}^*$. In those cases the generalized condition reads

$$\frac{\delta\mu}{f(\lambda, \mu)}\, d\lambda = \text{holomorphic} \tag{2.40}$$

Obviously the counting argument still works in this case. Moreover by adapting the function $p(\lambda, \mu)$ entering the Poisson bracket, eq.(2.34), we can preserve the condition

that the flows linearize on the Jacobian. The condition eq.(2.40) defines the holomorphic differentials as

$$\sigma_j(\lambda,\mu)d\lambda = \frac{R_j(\lambda,\mu)}{f(\lambda,\mu)\partial_\mu R(\lambda,\mu)}d\lambda$$

Then eq.(2.39) becomes

$$\sum_k \partial_{t_i}\lambda_k \,\frac{f(\lambda_k,\mu_k)}{p(\lambda_k,\mu_k)}\sigma_j(\lambda_k,\mu_k) = \delta_{ij} \tag{2.41}$$

which produces a linear flow on the Jacobian when $f(\lambda,\mu) = p(\lambda,\mu)$.

2.8 Solution of the Jaynes-Cummings-Gaudin model.

In the Jaynes-Cummings-Gaudin model, the spectral curve reads

$$\mu^2 = \frac{1}{g^4}(2\lambda-\omega)^2 + \frac{4}{g^2}H_n + \frac{2}{g^2}\sum_j \frac{H_j}{\lambda-\epsilon_j} + \sum_j \frac{\vec{s}_j\cdot\vec{s}_j}{(\lambda-\epsilon_j)^2}$$

so that

$$R_j(\lambda,\mu) = \frac{2}{g^2}\frac{1}{\lambda-\epsilon_j}, \quad R_0(\lambda,\mu) = -\mu^2 + \frac{1}{g^4}(2\lambda-\omega)^2 + \frac{4}{g^2}H_n + \sum_j \frac{\vec{s}_j\cdot\vec{s}_j}{(\lambda-\epsilon_j)^2}$$

Imposing that the points (λ_k,μ_k) belong to the spectral curve, we get the set of equations

$$\sum_j \frac{H_j}{\lambda_k-\epsilon_j} = \frac{g^2}{2}\mu_k^2 - \frac{1}{2g^2}(2\lambda_k-\omega)^2 - 2H_n - \frac{g^2}{2}\sum_j \frac{\vec{s}_j\cdot\vec{s}_j}{(\lambda_k-\epsilon_j)^2}$$

The matrix B_{kj} is the Cauchy matrix

$$B_{kj} = \frac{1}{\lambda_k-\epsilon_j} \tag{2.42}$$

Since the Hamiltonian is $H = \omega H_n + \sum_j H_j$ the equation of motion takes the form, using eq.(2.37)

$$\dot{\lambda}_k = \sum_j \partial_{t_j}\lambda_k = -ig^2\mu_k \sum_j B^{-1}_{jk} \tag{2.43}$$

or

$$\boxed{\dot{\lambda_k} = ig^2\mu_k\frac{\prod_j(\lambda_k-\epsilon_j)}{\prod_{l\neq k}(\lambda_k-\lambda_l)} = ig^2\frac{\sqrt{Q_{2n+2}(\lambda_k)}}{\prod_{l\neq k}(\lambda_k-\lambda_l)}} \tag{2.44}$$

Let us check that eq.(2.44) is of the form eq.(2.39). In fact, multiplying eq.(2.43) by $\mu_k^{-1}B_{ki}$ with B_{ki} given by eq.(2.42), and summing over k, we find

$$\sum_k \frac{B_{ki}}{\mu_k}\dot{\lambda_k} = -ig^2$$

or equivalently

$$\sum_k \frac{\prod_{j\neq i}(\lambda_k-\epsilon_j)}{\sqrt{Q_{2n+2}(\lambda_k)}}\dot{\lambda}_k = -ig^2$$

The coefficients of $\dot{\lambda}_k$ are precisely a basis of holomorphic differentials. In general such an equation is solved by the Abel transformation and θ-functions (see e.g. [5]).

2.8.1 Degenerate case.

Let us consider the degenerate case where

$$Q_{2n+2}(\lambda) = \frac{4}{g^4}\prod_{l=1}^{n+1}(\lambda - E_l)^2$$

This happens for instance when we start from an initial condition where

$$b = \bar{b} = 0, \quad s_j^{\pm} = 0, \quad s_j^z = e_j s, \quad e_j = \pm 1 \tag{2.45}$$

The energy of this configuration is

$$H = 2s\sum_j \epsilon_j e_j$$

At time t_i we have $B(\lambda)|_{t=t_i} = C(\lambda)|_{t=t_i} = 0$ and

$$\frac{Q_{2n+2}^2(\lambda)}{\prod_j(\lambda-\epsilon_j)^2} = A^2(\lambda)|_{t=t_i} = \frac{4}{g^4}\left(\lambda - \frac{\omega}{2} + \frac{sg^2}{2}\sum_{j=0}^{n-1}\frac{e_j}{\lambda-\epsilon_j}\right)^2$$

The zeroes of $Q_{2n+2}(\lambda)$ are located at the zeroes of $A(\lambda)|_{t=t_i}$ and are thus the roots of the equation

$$E = \frac{\omega}{2} - \frac{sg^2}{2}\sum_{j=0}^{n-1}\frac{e_j}{E-\epsilon_j} \tag{2.46}$$

This equation has also a remarkable interpretation. Its solutions E_l are the eigen frequencies of the small fluctuations around the configuration eq.(2.45). To perform the analysis of the small fluctuations around this configuration, we assume that $b, \bar{b}, s_j^{\pm}$ are first order and $s_j^z = se_j + \delta s_j^z$. Then δs_j^z is determined by saying that the spin is of length s and is of second order.

$$\delta s_j^z = -\frac{e_j}{2s}s_j^- s_j^+$$

This is compatible with eq.(1.31). The linearized equations of motion are

$$\dot{b} = -i\omega b - ig\sum_j s_j^- \tag{2.47}$$

$$\dot{s}_j^- = -2i\epsilon_j s_j^- + 2isge_j b \tag{2.48}$$

and their complex conjugate. We look for eigenmodes of the form

$$b(t) = b(0)e^{-2iEt}, \quad s_j^- = s_j^-(0)e^{-2iEt}, \quad \forall j$$

We get from eq.(2.48)

$$s_j^- = -sg\frac{e_j}{E-\epsilon_j}b$$

Inserting into eq.(2.47), we obtain the self-consistency equation for E

$$E = \frac{\omega}{2} - \frac{sg^2}{2}\sum_j \frac{e_j}{E-\epsilon_j} \tag{2.49}$$

which is exactly eq.(2.46).

Let us assume first that $\epsilon_j > 0$. The minimal energy state among the configurations eq.(2.45) corresponds to all the spins down: $e_j = -1$, $j = 0, \cdots, n-1$. The graph of the curves $y = \lambda - \omega/2$ and $y = \frac{sg^2}{2}\sum_{j=0}^{n-1}\frac{e_j}{\epsilon_j - \lambda}$ when all spins are down is presented in Fig.[2.1]. We see that we have $n+1$ *real* roots, meaning that the state is localy stable.

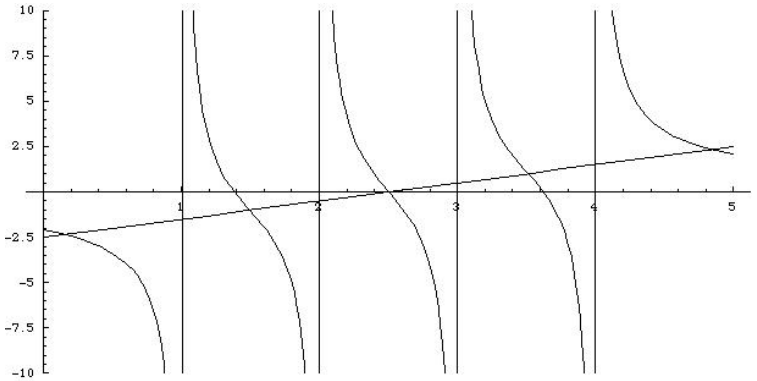


Figure 2.1: The solutions of eq.(2.49) when $\epsilon_j > 0$, $e_j = -1$.

Let us assume next that $\epsilon_j < 0$. The minimal energy state among the configurations eq.(2.45) now corresponds to all the spins up: $e_j = 1$, $j = 0, \cdots, n-1$. The graphs now look like Fig[2.2]. We see that we have $n+1$ real roots if $\omega > \omega_{sup}$ or $\omega < \omega_{inf}$. In between we have $n-1$ real roots and a *pair of complex conjugate roots*, which means that an instability develops. The question then arises to determine the time evolution of the non linear system.

The equations of motion of the separated variables are in that case

$$\dot{\lambda_i} = 2i\frac{\prod_k(\lambda_i - E_k)}{\prod_{j\neq i}(\lambda_i - \lambda_j)} \tag{2.50}$$

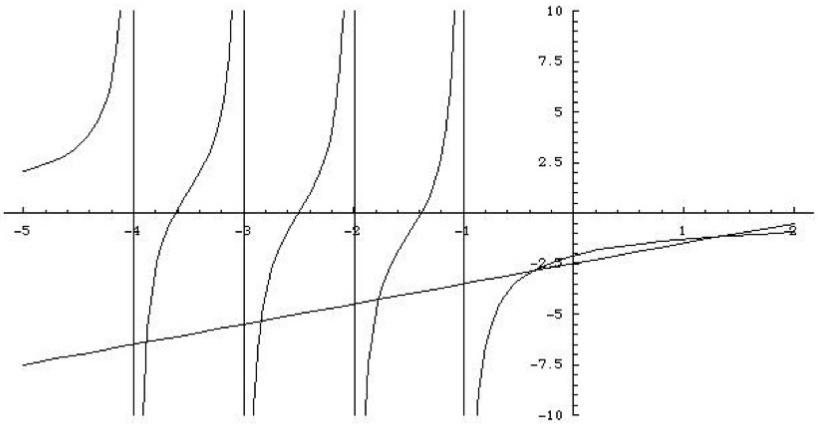


Figure 2.2: The solutions of eq.(2.49) when $\epsilon_j < 0$, $e_j = 1$.

One can solve explicitly these equations, hence finding the solution of the reduced model on the unstable surface. From eq.(2.50), we have

$$\frac{\dot{\lambda}_i}{\lambda_i - E_l} = 2i \frac{\prod_{k \neq l}(\lambda_i - E_k)}{\prod_{j \neq i}(\lambda_i - \lambda_j)}$$

Therefore

$$\sum_i \frac{\dot{\lambda}_i}{\lambda_i - E_l} = 2i \int_{C_\infty} \frac{dz}{2i\pi} \frac{\prod_{k \neq l}(z - E_k)}{\prod_j (z - \lambda_j)} = 2i(\Sigma_1 - \sigma_1(E) + E_l)$$

where C_∞ is a big circle at infinity surrounding all the λ_j. Hence

$$\log \prod_j (E_l - \lambda_j) = 2iE_l t + 2i \int^t dt(\Sigma_1(t) - \sigma_1(E))$$

where $\Sigma_1(t) = \sum_j \lambda_j(t)$. Define

$$\dot{\gamma} = 2i(\Sigma_1(t) - \sigma_1(E))\gamma$$

so that

$$\log \left[\gamma^{-1}(t) \prod_j (E_l - \lambda_j) \right] = 2iE_l t + C_l$$

Introducing

$$\mathcal{P}(\lambda, t) = \prod_j (\lambda - \lambda_j)$$

the above equation becomes (we parametrize $e^{C_l} = A_l \prod_{k\neq l}(E_l - E_k)$ for convenience)

$$\mathcal{P}(E_l, t) = \gamma(t) A_l \prod_{k\neq l}(E_l - E_k)\, e^{2iE_l t}, \quad l = 1, \cdots n+1.$$

These are $n+1$ conditions on the polynomial $\mathcal{P}(\lambda, t)$ of degree n in λ. It can then be reconstructed by the Lagrange interpolation formula

$$\boxed{\mathcal{P}(\lambda, t) = \gamma(t) \sum_l A_l e^{2iE_l t} \prod_{k\neq l}(\lambda - E_k)} \tag{2.51}$$

Imposing that the term of degree n should be normalized to λ^n, we get

$$\boxed{\gamma(t) = \frac{1}{\sum_l A_l e^{2iE_l t}}} \tag{2.52}$$

From

$$\begin{aligned} \mathcal{P}(\lambda, t) &= \gamma(t) \left(\sum_l A_l e^{2iE_l t} \right) \lambda^n - \gamma(t) \left(\sum_l (\sigma(E) - E_l) A_l e^{2iE_l t} \right) \lambda^{n-1} + \cdots \\ &= \lambda^n - \Sigma_1 \lambda^{n-1} + \cdots \end{aligned}$$

we deduce that

$$\frac{\sum_l A_l E_l e^{2iE_l t}}{\sum_l A_l e^{2iE_l t}} = \sigma_1(E) - \Sigma_1(t) = -\frac{1}{2i}\frac{\dot{\gamma}}{\gamma}$$

which shows the consistency of the construction. The original variables are given by

$$\begin{aligned} b(t) &= \gamma(t) \\ s_j^-(t) &= \frac{2}{g} \frac{\gamma(t) \mathcal{P}(\epsilon_j, t)}{\prod_{k\neq j}(\epsilon_j - \epsilon_k)} \\ s_j^z(t) &= -s + \frac{4}{g^2} \frac{\mathcal{P}(\epsilon_j, t)}{\prod_{k\neq j}(\epsilon_j - \epsilon_k)} \left[\epsilon_j + \Sigma_1(t) - \sigma_1(\epsilon) - \frac{\omega}{2} \right] \end{aligned}$$

This is an exact solution of the Jaynes-Cummings-Gaudin model. It can be viewed as a non linear superposition of the eigenmodes of the system.

2.8.2 Non degenerate case

In the non degenerate case, we can find the generalization of eq.(2.51). The result is eq.(2.66). Consider the hyperelliptic curve

$$y^2 = P_{2g+2}(x) = \sum_{i=0}^{2g+2} p_i x^i \tag{2.53}$$

We choose a partition of the branch points into $g+1$ disjoint pairs. The branch points at which the cuts start are called a_i and the end points of the cuts are called b_i, $i = 0, 1, \cdots g$.

A basis of holomorphic differentials is

$$\boxed{\omega_i = \frac{x^{i-1}dx}{y}, \quad i = 1, \cdots, g}$$

The dual basis of second kind differentials is

$$\boxed{\eta_j = \sum_{k=j}^{2g+1-j} (k+1-j) p_{k+1+j} \frac{x^k dx}{4y} = \frac{x^j}{4y}\frac{d}{dx}(P_{2g+2}(x)x^{-2j})_+, \quad j = 1, \cdots, g}$$

where $()_+$ means the polynomial part in the expansion at $x = \infty$. The first thing to check is that η_j is a second kind differential, i.e. has no residue at infinity. For this, we remark that

$$\begin{aligned} \frac{x^j}{4y}\frac{d}{dx}(P_{2g+2}(x)x^{-2j})_+ &= \frac{x^j}{4y}\frac{d}{dx}\Big(P_{2g+2}(x)x^{-2j} - (P_{2g+2}(x)x^{-2j})_-\Big) \\ &= \frac{1}{2}\frac{d}{dx}(yx^{-j}) - \frac{x^j}{4y}\frac{d}{dx}(P_{2g+2}(x)x^{-2j})_- \end{aligned} \tag{2.54}$$

The first term is a total derivative and cannot have a residue. The second term is regular at infinity. To compute $(\eta_j \bullet \omega_i)$, let us apply eq.(2.12). Using eq.(2.54) we see that we can take $g_1 = \frac{1}{2}(yx^{-j}) + \text{regular}$. Hence

$$(\eta_j \bullet \omega_i) = \frac{1}{2}\text{res}_\infty(x^{i-j-1}dx + \text{regular}) = \frac{1}{2}\delta_{ij}$$

The periods of these differentials are denoted as follows ($i, j = 1, \cdots, g$)

$$\begin{aligned} (2\omega)_{ij} &= \int_{a_j} \omega_i, \qquad (2\omega')_{ij} = \int_{b_j} \omega_i \\ (2\eta)_{ij} &= \int_{a_j} \eta_i, \qquad (2\eta')_{ij} = \int_{b_j} \eta_i \end{aligned}$$

The Riemann bilinear identities imply

$$\omega' \, {}^t\omega - \omega \, {}^t\omega' = 0, \quad \eta' \, {}^t\omega - \eta \, {}^t\omega' = -\frac{i\pi}{2}\text{Id}, \quad \eta' \, {}^t\eta - \eta \, {}^t\eta' = 0$$

Let us introduce the symmetric function

$$\begin{aligned} F(x_1, x_2) &= P_{2g+2}(\sqrt{x_1 x_2}) + P_{2g+2}(-\sqrt{x_1 x_2}) + \frac{x_1 + x_2}{2\sqrt{x_1 x_2}}(P_{2g+2}(\sqrt{x_1 x_2}) - P_{2g+2}(-\sqrt{x_1 x_2})) \\ &= 2\sum_{i=0}^{g+1} p_{2i} x_1^i x_2^i + (x_1 + x_2)\sum_{i=0}^{g} p_{2i+1} x_1^i x_2^i \end{aligned}$$

With it, we construct the fundamental symmetrical bi-differential

$$\boxed{\Omega(x_1,x_2)=\frac{2y_1y_2+F(x_1,x_2)}{4(x_1-x_2)^2}\frac{dx_1}{y_1}\frac{dx_2}{y_2}}$$

If $x_2 \to x_1+\epsilon$, we have

$$F(x_1,x_2)=2P_{2g+2}(x_1)+\epsilon P'_{2g+2}(x_1)+O(\epsilon^2)=2y(x_1)y(x_2)+O(\epsilon^2)$$

from which it follows that

$$\Omega(x_1,x_2)=\left(\frac{1}{(x_1-x_2)^2}+O(1)\right)dx_1dx_2 \tag{2.55}$$

Alternatively, we can write

$$\Omega(x_1,x_2) = \frac{\partial}{\partial x_2}\left(\frac{y_1+y_2}{2y_1(x_1-x_2)}\right)dx_1dx_2+\sum_{i=1}^{g}\omega_i(x_1)\eta_i(x_2) \tag{2.56}$$

$$= -\frac{\partial}{\partial x_1}\left(\frac{y_1+y_2}{2y_2(x_1-x_2)}\right)dx_1dx_2+\sum_{i=1}^{g}\eta_i(x_1)\omega_i(x_2) \tag{2.57}$$

In general the fundamental Kleinian σ-function is defined by

$$\boxed{\sigma(z)=\left(\frac{1}{D(\Gamma)}\right)^{\frac{1}{4}}\sqrt{\frac{\pi^g}{\det(2\omega)}}e^{\frac{1}{2}\langle z|\eta\omega^{-1}|z\rangle}\theta[\epsilon_R](z,\omega,\omega')}$$

where $[\epsilon_R]$ is the characteristic of the vector of Riemann constants. The function $D(\Gamma)$ is the discriminant of the equation $R(\lambda,\mu)$ defining the curve Γ. Its main property is that it is invariant under modular transformations

$$\sigma(z,\omega,\omega')=\sigma(z,\hat{\omega},\hat{\omega}')$$

where $(\hat{\omega},\hat{\omega}')$ is related to (ω,ω') by a $Sp(2g,\mathbb{Z})$ transformation. From this we define the Kleinian ζ and $\wp$-functions by

$$\zeta_i(z)=\frac{\partial\log\sigma(z)}{\partial z_i},\quad i=1,\cdots g$$

and

$$\wp_{ij}(z)=-\frac{\partial^2\log\sigma(z)}{\partial z_i\partial z_j},\quad i,j=1,\cdots g$$

The functions $\wp_{ij}(z)$ are automorphic functions with respect to the $Sp(2g,\mathbb{Z})$ transformations.

The functions $\sigma(z)$, $\zeta_i(z)$ and $\wp_{ij}(z)$ have the following periodicity properties:

$$\sigma(z+|\Omega(m,m')\rangle) = e^{\langle E(m,m')|z+\Omega(m,m')\rangle} e^{-i\pi\langle m|m\rangle + 2i\pi(\langle m|q'\rangle - \langle m'|q\rangle)}\sigma(z)$$

$$\zeta_i(z+\Omega(m,m')) = \zeta_i(z) + E_i(m,m')$$

$$\wp_{ij}(z+\Omega(m,m')) = \wp_{ij}(z)$$

where we have introduced the vectors of periods

$$|E(m,m')\rangle = 2\eta|m\rangle + 2\eta'|m'\rangle, \quad |\Omega(m,m')\rangle = 2\omega|m\rangle + 2\omega'|m'\rangle$$

For an hyperelliptic curve, things can be made more explicit. The function $\sigma(z)$ reads

$$\boxed{\sigma(z) = e^{\langle z|(2\omega)^{-1}\eta|z\rangle + 4i\pi\langle q'|(2\omega)^{-1}|z\rangle + i\pi\langle q'|\tau|q'\rangle - 2i\pi\langle q'|q\rangle}\theta((2\omega)^{-1}z - K_{a_0})}$$

where K_{a_0} is the vector of Riemann constants

$$K_{a_0} = \sum_{k=1}^{g}\int_{a_0}^{a_k}(2\omega)^{-1}\omega(x)$$

Because a_0 is a branch point, K_{a_0} is a half period and can be written as $K_a = q + \tau q'$ with half integers q and q'. The characteristic $[\epsilon_R]$ is

$$[\epsilon_R] = \begin{bmatrix} {}^t q \\ {}^t q' \end{bmatrix}$$

Notice that if

$$z_j = \sum_{k=1}^{g}\int_{a_0}^{x_k}\omega_j(x)$$

we have $(2\omega)^{-1}z - K_{a_0} = (2\omega)^{-1}u$ where

$$\boxed{u_i = \sum_{k=1}^{g}\int_{a_k}^{x_k}\frac{x^{i-1}dx}{y}} \tag{2.58}$$

Theorem 9 *Let $(a_0, y(a_0))$, (x,y) and (μ,ν) be arbitrary points on Γ. Let*

$$\{(x_1,y_1),\cdots,(x_g,y_g)\}, \quad \{(\mu_1,\nu_1),\cdots,(\mu_g,\nu_g)\}$$

be arbitrary distinct points. Then the following relation is valid

$$\int_\mu^x \sum_{i=1}^{g}\int_{\mu_i}^{x_i}\Omega(x,x_i) \;=\; \log\left\{\frac{\sigma\left(\int_{a_0}^{x}\omega - \sum_{i=1}^{g}\int_{a_i}^{x_i}\omega\right)\sigma\left(\int_{a_0}^{\mu}\omega - \sum_{i=1}^{g}\int_{a_i}^{\mu_i}\omega\right)}{\sigma\left(\int_{a_0}^{x}\omega - \sum_{i=1}^{g}\int_{a_i}^{\mu_i}\omega\right)\sigma\left(\int_{a_0}^{\mu}\omega - \sum_{i=1}^{g}\int_{a_i}^{x_i}\omega\right)}\right\}$$

Proof. Using eq.(2.55), the left hand side in this formula behaves like $\log(x-x_i)$ when $x \to x_i$ and $-\log(x-\mu_i)$ when $x \to \mu_i$ and these are the only singularities. Moreover it vanishes when $x \to \mu$. This behaviour is clearly reproduced by the right hand side. Then using eq.(2.56) one can check that the periods on both sides are the same. Hence the two expressions are identical. ∎

Theorem 10 (Bolza)

$$\boxed{\zeta_j\left(\int_{a_0}^{x}\omega+u\right)=-\sum_{k=0}^{g}\int_{a_k}^{x_k}\eta_j(x)+\frac{1}{2}\sum_{k=0}^{g}y_k\frac{\left(\frac{R(z)}{z-x_k}z^{-j}\right)_+\Big|_{z=x_k}}{R'(x_k)}} \tag{2.59}$$

Proof. Setting $\mu_i = a_i$, we get

$$\int_{\mu}^{x}\sum_{k=1}^{g}\int_{a_k}^{x_k}\Omega(x,x_k) \;=\; \log\left\{\frac{\sigma\left(\int_{a_0}^{x}\omega-u\right)\sigma\left(\int_{a_0}^{\mu}\omega\right)}{\sigma\left(\int_{a_0}^{x}\omega\right)\sigma\left(\int_{a_0}^{\mu}\omega-u\right)}\right\}$$

We now take the derivative of both sides with respect to u_j. We obtain

$$\int_{\mu}^{x}\sum_{k=1}^{g}\Omega(x,x_k)\frac{\partial x_k}{\partial u_j}=-\zeta_j\left(\int_{a_0}^{x}\omega-u\right)+\zeta_j\left(\int_{a_0}^{\mu}\omega-u\right)$$

With the help of eq.(2.57) we find

$$\begin{aligned}\zeta_j\left(\int_{a_0}^{x}\omega-u\right)-\zeta_j\left(\int_{a_0}^{\mu}\omega-u\right) \;&=\; \frac{1}{2}\sum_{k=1}^{g}\frac{1}{y_k}\frac{\partial x_k}{\partial u_j}\left[\frac{y+y_k}{x-x_k}-\frac{\nu+y_k}{\mu-x_k}\right]\\ &\quad-\int_{\mu}^{x}dx\sum_i\eta_i(x)\sum_k\omega_i(x_k)\frac{\partial x_k}{\partial u_j}\end{aligned}$$

From eq.(2.58) we have

$$\sum_k\omega_i(x_k)\frac{\partial x_k}{\partial u_j}=\delta_{ij} \tag{2.60}$$

Hence

$$\zeta_j\left(\int_{a_0}^{x}\omega-u\right)-\zeta_j\left(\int_{a_0}^{\mu}\omega-u\right)=-\int_{\mu}^{x}dx\;\eta_j(x)+\frac{1}{2}\sum_{k=1}^{g}\frac{1}{y_k}\frac{\partial x_k}{\partial u_j}\left[\frac{y+y_k}{x-x_k}-\frac{\nu+y_k}{\mu-x_k}\right]$$

Applying the hyperelliptic involution $\sigma(x,y)=(x,-y)$, $\sigma(\mu,\nu)=(\mu,-\nu)$, we get as well

$$\zeta_j\left(\int_{a_0}^{x}\omega+u\right)-\zeta_j\left(\int_{a_0}^{\mu}\omega+u\right)=-\int_{\mu}^{x}dx\;\eta_j(x)+\frac{1}{2}\sum_{k=1}^{g}\frac{1}{y_k}\frac{\partial x_k}{\partial u_j}\left[\frac{y-y_k}{x-x_k}-\frac{\nu-y_k}{\mu-x_k}\right] \tag{2.61}$$

Let us introduce the Vandermonde matrix

$$V_{ki} = x_k^{i-1}$$

then eq.(2.60) gives

$$\frac{1}{y_k}\frac{\partial x_k}{\partial u_j} = V_{jk}^{-1} = \frac{(P(z)z^{-j})_+|_{z=x_k}}{P'(x_k)} \tag{2.62}$$

where

$$P(z) = \prod_{k=1}^{g}(z - x_k)$$

The last formula is easy to prove

$$\sum_j V_{lj}V_{jk}^{-1} = \frac{1}{P'(x_k)}\sum_j (P(z)z^{-j})_+|_{z=x_k} x_l^{j-1}$$

but

$$\sum_{j=1}^{g}(P(z)z^{-j})_+ x_l^{j-1} = \sum_j x_l^{j-1}\sum_{r\geq j} p_r z^{r-j} = \sum_r p_r z^r \sum_{1\leq j\leq r} x_l^{j-1}z^{-j} = \frac{P(z)-P(x_l)}{z-x_l}$$

If $z \to x_k \neq x_l$, we get zero because $P(x_k) = P(x_l) = 0$, while if $z \to x_l$, we get $P'(x_l)$. Hence we have shown that $\sum_j V_{lj}V_{jk}^{-1} = \delta_{kl}$.

Let

$$R(z) = (z - x)P(z)$$

By simple polar decomposition, we have

$$\frac{(P(z)z^{-j})_+|_{z=x}}{R'(x)} = \sum_k \frac{(P(z)z^{-j})_+|_{z=x_k}}{P'(x_k)}\frac{1}{x-x_k} = \sum_k V_{jk}^{-1}\frac{1}{x-x_k} \tag{2.63}$$

Next we have the identity

$$\frac{\left(\frac{R(z)}{z-x_k}z^{-j}\right)_+\Big|_{z=x_k}}{R'(x_k)} - \frac{\left(\frac{P(z)}{z-x_k}z^{-j-1}\right)_+\Big|_{z=x_k}}{P'(x_k)} = -\frac{(P(z)z^{-j})_+|_{z=x_k}}{P'(x_k)}\frac{1}{x-x_k} \tag{2.64}$$

This is because

$$\begin{aligned}\left(\frac{R(z)}{z-x_1}z^{-j}\right)_+ &= \sum_{r\leq g-j} z^{g-j-r}(-1)^r[x\sigma_{r-1}(x_2,\cdots,x_g) + \sigma_r(x_2,\cdots,x_g)]\\ \left(P(z)z^{-j}\right)_+ &= \sum_{r\leq g-j} z^{g-j-r}(-1)^r[x_1\sigma_{r-1}(x_2,\cdots,x_g) + \sigma_r(x_2,\cdots,x_g)]\end{aligned}$$

so that

$$\begin{aligned}\left(\frac{R(z)}{z-x_1}z^{-j}\right)_+ - \left(P(z)z^{-j}\right)_+ &= (x-x_1)\sum_{r\leq g-j} z^{g-j-r}(-1)^r\sigma_{r-1}(x_2,\cdots,x_g)\\ &= -(x-x_1)\left(\frac{P(z)}{z-x_1}z^{-j}\right)_+\end{aligned}$$

It follows that

$$\frac{\left(\frac{R(z)}{z-x_k}z^{-j}\right)_+\Big|_{z=x_k}}{R'(x_k)} - \frac{\left(\frac{P(z)}{z-x_k}z^{-j-1}\right)_+\Big|_{z=x_k}}{P'(x_k)} = -V^{-1}_{jk}\frac{1}{x-x_k} \tag{2.65}$$

We can now evaluate the sums appearing in the right hand side of eq.(2.61). We have

$$\begin{aligned}\sum_{k=1}^{g}\frac{1}{y_k}\frac{\partial x_k}{\partial u_j}\frac{y-y_k}{x-x_k} &= y\sum_k V^{-1}_{jk}\frac{1}{x-x_k} - \sum_k y_k V^{-1}_{jk}\frac{1}{x-x_k}\\ &= y\frac{(P(z)z^{-j})_+|_{z=x}}{R'(x)} + \sum_k y_k\left(\frac{\left(\frac{R(z)}{z-x_k}z^{-j}\right)_+\Big|_{z=x_k}}{R'(x_k)} - \frac{\left(\frac{P(z)}{z-x_k}z^{-j-1}\right)_+\Big|_{z=x_k}}{P'(x_k)}\right)\end{aligned}$$

Noticing that $P(z) = R(z)/(z-x)$ and defining $(x_0,y_0)=(x,y)$, we get

$$\sum_{k=1}^{g}\frac{1}{y_k}\frac{\partial x_k}{\partial u_j}\frac{y-y_k}{x-x_k} = \sum_{k=0}^{g} y_k\frac{\left(\frac{R(z)}{z-x_k}z^{-j}\right)_+\Big|_{z=x_k}}{R'(x_k)} - \sum_{k=1}^{g} y_k\frac{\left(\frac{P(z)}{z-x_k}z^{-j-1}\right)_+\Big|_{z=x_k}}{P'(x_k)}$$

Defining $(x_{g+1},y_{g+1})=(\mu,\nu)$, $\tilde{R}(z)=(z-\mu)P(z)$, $a_{g+1}=a_0$, we obtain

$$\sum_{k=1}^{g}\frac{1}{y_k}\frac{\partial x_k}{\partial u_j}\left[\frac{y-y_k}{x-x_k}-\frac{\nu-y_k}{\mu-x_k}\right] = \sum_{k=0}^{g} y_k\frac{\left(\frac{R(z)}{z-x_k}z^{-j}\right)_+\Big|_{z=x_k}}{R'(x_k)} - \sum_{k=1}^{g+1} y_k\frac{\left(\frac{\tilde{R}(z)}{z-x_k}z^{-j-1}\right)_+\Big|_{z=x_k}}{\tilde{R}'(x_k)}$$

Hence

$$\begin{aligned}\zeta_j\left(\int_{a_0}^{x}\omega+u\right) + \sum_{k=0}^{g}\int_{a_k}^{x_k}\eta_j(x) - \frac{1}{2}\sum_{k=0}^{g} y_k\frac{\left(\frac{R(z)}{z-x_k}z^{-j}\right)_+\Big|_{z=x_k}}{R'(x_k)} &=\\ \zeta_j\left(\int_{a_0}^{\mu}\omega+u\right) + \sum_{k=1}^{g+1}\int_{a_k}^{x_k}\eta_j(x) - \frac{1}{2}\sum_{k=1}^{g+1} y_k\frac{\left(\frac{\tilde{R}(z)}{z-x_k}z^{-j}\right)_+\Big|_{z=x_k}}{\tilde{R}'(x_k)}&\end{aligned}$$

The left hand side is a symmetric function of $x, x_1, \cdots, x_g$, but the right hand side does not depend on x. Therefore the whole expression is a constant independent of $x, x_1, \cdots, x_g$. Applying the hyperelliptic involution, we see that this constant vanishes. ■

The main interest in the Kleinian functions is that they give a very explicit solution to the Jacobi problem.

Theorem 11 *Let Γ as in eq.(2.53). Then the preimage of the point $u \in \mathrm{Jac}(\Gamma)$ is given by the set of points $(P_1, P_2, \cdots, P_g)$ where $(x_1, x_2, \cdots, x_g)$ are the zeroes of the polynomial*

$$\boxed{\mathcal{P}(x,u) = x^g - \mathcal{P}_g(u)x^{g-1} - \mathcal{P}_{g-1}(u)x^{g-2} - \cdots - \mathcal{P}_1(u)} \tag{2.66}$$

where

$$\boxed{\mathcal{P}_i(u) = \frac{1}{\sqrt{p_{2g+2}}}\left\{\zeta_i\left(u + \int_{a_0}^{P_+}\omega\right) - \zeta_i\left(u + \int_{a_0}^{P_-}\omega\right) - c_i\right\}}$$

The coordinates $(y_1, y_2, \cdots, y_g)$ are given by

$$y_k = -\left.\frac{\partial \mathcal{P}(x,u)}{\partial u_g}\right|_{x=x_k} \tag{2.67}$$

The constants c_i are given by eq.(2.69).

Proof. We now take the limit $(x,y) \to P_\pm$ in eq.(2.59). When $x \to \infty$ we have, by eq.(2.65)

$$\lim_{x\to\infty} \frac{\left(\frac{R(z)}{z-x_k}z^{-j}\right)_+\Big|_{z=x_k}}{R'(x_k)} = \frac{\left(\frac{P(z)}{z-x_k}z^{-j-1}\right)_+\Big|_{z=x_k}}{P'(x_k)}$$

On the other hand

$$y(x)\frac{\left(\frac{R(z)}{z-x}z^{-j}\right)_+\Big|_{z=x}}{R'(x)} = y(x)\frac{\left(P(x)z^{-j}\right)_+}{P(x)} = y(x)\frac{P(x)x^{-j} - \left(P(x)z^{-j}\right)_-}{P(x)}$$

When $x \to \infty$, the first term behaves as $y(x)x^{-j} \simeq (y(x)x^{-j})_+$, while in the second term, $\left(P(x)z^{-j}\right)_- \simeq (-1)^{g-j+1}\sigma_{g-j+1}(x_1, x_2, \cdots, x_g)x^{-1} + O(x^{-2})$ and we find

$$y(x)\frac{\left(\frac{R(z)}{z-x}z^{-j}\right)_+\Big|_{z=x}}{R'(x)} \simeq (y(x)x^{-j})_+ - \sqrt{p_{2g+2}}(-1)^{g-j+1}\sigma_{g-j+1}(x_1, x_2, \cdots, x_g)$$

Since the limits of the left hand side of eq.(2.59) are finite, so are the limits

$$c_j^{(\pm)} = \lim_{(x,y)\to P_\pm} \left(-\int_{a_0}^{x} \eta_j + \frac{1}{2}(yx^{-j})_+ \right) \tag{2.68}$$

and they are obviously independent of $x_1, \cdots, x_g$. Therefore we arrive at

$$\zeta_j \left(\int_{a_0}^{P_+} \omega + u \right) - \zeta_j \left(\int_{a_0}^{P_-} \omega + u \right) = c_j + \sqrt{p_{2g+2}}(-1)^{g-j}\sigma_{g-j+1}(x_1, x_2, \cdots, x_g)$$

where we have set

$$c_j = c_j^{(+)} - c_j^{(-)} \tag{2.69}$$

Multiplying by x^{j-1} and summing over $j = 1, \cdots, g$, we get

$$\sum_{j=1}^{g} x^{j-1} \left(\zeta_j \left(\int_{a_0}^{P_+} \omega + u \right) - \zeta_j \left(\int_{a_0}^{P_-} \omega + u \right) - c_j \right) = -\sqrt{p_{2g+2}} \left(\prod_{k=1}^{g} (x - x_k) - x^g \right)$$

Hence we have shown

$$\prod_{k=1}^{g} (x - x_k) = x^g - \sum_{j=1}^{g} \mathcal{P}_{g-j}(u) x^{j-1}$$

which is eq.(2.66). From eq.(2.62), we have

$$\frac{\partial x_k}{\partial u_g} = \frac{y_k}{\prod_{l\neq k}(x_k - x_l)}$$

and

$$\frac{\partial}{\partial u_g} \prod_l (x - x_l) = -\sum_k \frac{\partial x_k}{\partial u_g} \prod_{l\neq k} (x - x_l)$$

so that

$$\left. \frac{\partial}{\partial u_g} \prod_l (x - x_l) \right|_{x=x_k} = -y_k$$

This proves eq.(2.67). ∎

For applications to the Jaynes-Cummings-Gaudin model, $u \equiv u(t)$ is chosen to be a linear function of time. Once the $(x_k(t), y_k(t))$ are found by solving the inversion Jacobi problem, it is easy to reconstruct the original variables of the model.

Bibliography

[1] H. Baker, **Abelian functions: Abel's theorem and allied theory including the theory of theta functions.** Cambridge University Press (1897) (Reprinted 1995)

[2] 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.

[3] P. van Moerbeke, D. Mumford. *" The spectrum of difference operators and algebraic curves."* Acta.Math. 143 (1979) p.93-154.

[4] M. Adler, P. van Moerbeke. *"Linearization of Hamiltonian Systems, Jacobi Varieties and Representation Theory."* Advances in Mathematics 38, (1980) p.318-379.

[5] D. Mumford **"Tata Lectures on Theta"** Vol. I and II Birkhauser (1983–1984).

[6] B.A. Dubrovin, I.M. Krichever, S.P. Novikov. *"Integrable Systems I"*. Encyclopedia of Mathematical Sciences, Dynamical Systems IV Springer (1990), p. 173-281.

[7] V. Buchstaber, V. Enolskii, D. Leikin, *Hyperelliptic Kleinian Functions and Applications.* Amer. Math. Soc. Transl. Vol. 179, (1997), pp.1-33.

[8] J. Eilbeck, V. Enolskii, H. Holden *The hyperelliptic ζ-function and the integrable massive Thirring model.* Proc. R. Soc. Lond. A vol. 459 (2003) pp. 1581-1610.

[9] O.Babelon, D. Bernard, M. Talon, **Introduction to Classical Integrable Systems.** Cambridge University Press (2003).

Chapter 3

Infinite dimensional systems.

3.1 Integrable field theories and monodromy matrix.

For a system with a finite number of degrees of freedom, we have seen that a Lax matrix could be interpreted as a coadjoint orbit. In field theory we have seen that one possibility is to consider the algebra of pseudo differential operators. We obtained in this way the KdV hierarchies. A more general approach consists is starting from the zero curvature representation which assumes that the field equations can be recast in the form

$$\boxed{\partial_t U - \partial_x V - [V, U] = 0} \tag{3.1}$$

As before the matrices U and V will in general depend on an extra parameter λ. The zero curvature (3.1) condition expresses the compatibility condition of the associated linear system

$$\boxed{(\partial_x - U)\,\Psi = 0, \quad (\partial_t - V)\,\Psi = 0} \tag{3.2}$$

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 (3.2) determines the matrix Ψ up to multiplication on the right by a constant matrix, which we can fix by requiring $\Psi(\lambda, 0, 0) = 1$. This Ψ will be called the wave function.

Choosing a path γ from the origin to the point (x, t) the wave function can be written symbolically as

$$\boxed{\Psi(x,t) = \overleftarrow{\exp}\left[\int_\gamma (U dx + V dt)\right]} \tag{3.3}$$

where $\overleftarrow{\exp}$ denotes the path–ordered exponential. This is just the parallel transport along the curve γ with the connection (U, V). Since the Lax connection satisfies the zero curvature relation (3.1) the value of the path–ordered exponential 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 $\Psi(2\pi, t)$ the monodromy matrix $T(\lambda, t)$:

$$\boxed{T(\lambda,t) \equiv \overleftarrow{\exp}\left[\int_0^{2\pi} U(\lambda,x,t)dx\right]} \tag{3.4}$$

where we assume that $U(\lambda, x, t)$ and $V(\lambda, x, t)$ depend on a spectral parameter λ.

It is the monodromy matrix which plays the role of the Lax matrix in the field theoretical context as the following proposition shows

Proposition 12 *Assume that all fields are periodic in x with period 2π. Let $T(\lambda, t)$ be the monodromy matrix. Its time evolution is given by the Lax equation*

$$\boxed{\partial_t T(\lambda,t) = [V(\lambda,0,t), T(\lambda,t)]} \tag{3.5}$$

As a consequence the quantities

$$\boxed{H^{(n)}(\lambda) = \mathrm{Tr}\,(T^n(\lambda,t))} \tag{3.6}$$

are independent of time. Hence traces of powers of the monodromy matrix generate conserved quantities.

Proof. Thinking of the path–ordered exponential on $[0, 2\pi]$ as

$$\overleftarrow{\exp}\left[\int_0^{2\pi} U(x)dx\right] \sim (1+\delta x U(x_n))\cdots(1+\delta x U(x_1))$$

with a subdivision $x_1 = 0 < x_2 < \cdots < x_n = 2\pi$ such that $x_{i+1} - x_i = \delta x \to 0$, we get (all exponentials are path–ordered exponentials):

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

Performing the integral,

$$\partial_t T(\lambda,t) = V(\lambda,2\pi,t)T(\lambda,t) - T(\lambda,t)V(\lambda,0,t) \tag{3.7}$$

So, if the fields are periodic, we have $V(\lambda, 2\pi, t) = V(\lambda, 0, t)$ and we obtain eq.(3.5). This is a Lax equation. It implies that $H^{(n)}(\lambda)$ is time independent. Expanding in λ we obtain an infinite set of conserved quantities.

3.2 Abelianization.

We consider the linear system eq.(3.2) where $U(\lambda, x, t)$ and $V(\lambda, x, t)$ are matrices depending in a rational way on a parameter λ having poles at *constant* values λ_k.

$$U = U_0 + \sum_k U_k \qquad \text{with} \qquad U_k = \sum_{r=-n_k}^{-1} U_{k,r}(\lambda - \lambda_k)^r \tag{3.8}$$

$$V = V_0 + \sum_k V_k \qquad \text{with} \qquad V_k = \sum_{r=-m_k}^{-1} V_{k,r}(\lambda - \lambda_k)^r \tag{3.9}$$

The compatibility condition of the linear system (3.2) is the zero curvature condition (3.1). We demand that it holds identically in λ.

As for finite dimensional systems we first make a local analysis around each pole λ_k in order to understand solutions of eq.(3.1). Around each singularity λ_k, one can perform a gauge transformation bringing simultaneously $U(\lambda)$ and $V(\lambda)$ to a diagonal form. The important new feature that must be to emphasized, as compared to the finite dimensional case, is that this construction is *local* in x. We have

Proposition 13 *There exists a local, periodic, gauge transformation*

$$\partial_x - U = g^{(k)}(\partial_x - A^{(k)})g^{(k)-1}, \quad \partial_t - V = g^{(k)}(\partial_t - B^{(k)})g^{(k)-1} \tag{3.10}$$

where $g^{(k)}(\lambda)$, $A^{(k)}(\lambda)$ and $B^{(k)}(\lambda)$ are formal series in $\lambda - \lambda_k$

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

such that the matrices $A^{(k)}(\lambda)$ and $B^{(k)}(\lambda)$ are diagonal. Moreover $\partial_t A^{(k)}(\lambda) - \partial_x B^{(k)}(\lambda) = 0$.

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

$$U = U_0 + \sum_k U_k, \quad \text{with} \quad U_k \equiv \left(g^{(k)} A_-^{(k)} g^{(k)-1}\right)_- \tag{3.11}$$

$$V = V_0 + \sum V_k, \quad \text{with} \quad V_k \equiv \left(g^{(k)} B_-^{(k)} g^{(k)-1}\right)_- \tag{3.12}$$

In this diagonal gauge it is easy to compute the conserved quantities:

Proposition 14 *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.(3.6) by*

$$H^{(n)}(\lambda) = \mathrm{Tr}\, \exp\left[n \int_0^{2\pi} A^{(k)}(\lambda, x, t) dx\right] = \mathrm{Tr}\, \exp\left[n Q^{(k)}(\lambda)\right]$$

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

We now give some examples of 2 dimensional field theories having a zero curvature representation.

Example 1. The first example is the non–linear σ model. 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:

$$U = \frac{1}{\lambda - 1} J_x, \quad V = -\frac{1}{\lambda + 1} J_t \tag{3.13}$$

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:

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

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.

Example 2. Another important example is the sinh–Gordon model. It also has a two poles Lax connection, one pole at $\lambda = 0$, the other at $\lambda = \infty$. Moreover, we require that in the light cone coordinates, $x_\pm = x \pm t$, $U(\lambda, x_\pm)$ has a simple pole at $\lambda = 0$ and $V(\lambda, x_\pm)$ a simple pole at $\lambda = \infty$. The most general 2×2 system of this form is:

$$\begin{aligned} (\partial_{x_+} - U)\Psi &= 0, \quad U = U_0 + \lambda^{-1} U_1 \\ (\partial_{x_-} - V)\Psi &= 0, \quad V = V_0 + \lambda V_1 \end{aligned}$$

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, Namely, we consider the group Z_2 acting by:

$$\Psi(\lambda) \longrightarrow \sigma_z \Psi(-\lambda)\sigma_z^{-1}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{3.14}$$

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 connection, and commuting with the action of Z_2, i.e. g diagonal. This gauge freedom can be used to set $(V_0)_{ii} = 0$. The symmetry condition then gives:

$$U = \begin{pmatrix} u_0 & \lambda^{-1}u_1 \\ \lambda^{-1}u_2 & -u_0 \end{pmatrix}, \quad V = \begin{pmatrix} 0 & \lambda v_1 \\ \lambda v_2 & 0 \end{pmatrix}$$

In this gauge, the zero curvature equation reduces to :

$$\partial_{x_-} u_0 - u_1 v_2 + v_1 u_2 = 0 \tag{3.15}$$

$$\partial_{x_-} u_1 = 0, \quad \partial_{x_-} u_2 = 0 \tag{3.16}$$

$$\partial_{x_+} v_1 - 2 v_1 u_0 = 0, \quad \partial_{x_+} v_2 + 2 v_2 u_0 = 0 \tag{3.17}$$

From eq.(3.16) we have $u_1 = \alpha(x_+)$, $u_2 = \beta(x_+)$. We set $u_0 = \partial_{x_+}\varphi$. Then, from eq(3.17) we have $v_1 = \gamma(x_-)\exp 2\varphi$ and $v_2 = \delta(x_-)\exp -2\varphi$. Finally eq(3.15) 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 arbitrary 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

$$U = \begin{pmatrix} \partial_{x_+}\varphi & m\lambda^{-1} \\ m\lambda^{-1} & -\partial_{x_+}\varphi \end{pmatrix}; \quad V = \begin{pmatrix} 0 & m\lambda e^{2\varphi} \\ m\lambda e^{-2\varphi} & 0 \end{pmatrix} \tag{3.18}$$

Hence the Lax connection of the sinh–Gordon model is naturally recovered from two–poles systems with Z_2 symmetry. This construction generalizes to other Lie algebras, the reduction group being generated by the Coxeter automorphism and yields the Toda field theories.

Example 3. The KdV equation reads:

$$4\partial_t u = -6u\partial_x u + \partial_x^3 u \tag{3.19}$$

The KdV equation can be written as the zero curvature condition

$$F_{xt} \equiv \partial_x V - \partial_t U - [U, V] = 0$$

with the connection U, V, depending on a spectral parameter λ:

$$U = \begin{pmatrix} 0 & 1 \\ \lambda + u & 0 \end{pmatrix}, \quad V = \frac{1}{4}\begin{pmatrix} \partial_x u & 4\lambda - 2u \\ 4\lambda^2 + 2\lambda u + \partial_x^2 u - 2u^2 & -\partial_x u \end{pmatrix} \tag{3.20}$$

Alternatively one can recast the KdV equation in the Lax form $\partial_t L = [M, L]$, where L and M are the following *differential* operators:

$$\begin{aligned} L &= \partial^2 - u \\ M &= \frac{1}{4}(4\partial^3 - 6u\partial - 3(\partial_x u)) = (L^{\frac{3}{2}})_+ \end{aligned} \tag{3.21}$$

The operator ∂ acts as ∂_x, and the notation $(L^{\frac{3}{2}})_+$ refers to the pseudo differential operator formalism.

Of course these two descriptions are not independent. To relate them, consider the linear system:

$$(\partial_x - U)\begin{pmatrix}\Psi\\ \chi\end{pmatrix} = 0, \quad (\partial_t - V)\begin{pmatrix}\Psi\\ \chi\end{pmatrix} = 0 \tag{3.22}$$

The x-equation yields $\chi = \partial_x \Psi$ and

$$(L - \lambda)\Psi = 0 \quad \text{with} \quad L = \partial_x^2 - u \tag{3.23}$$

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(3.23), this may be rewritten as:

$$(\partial_t - M)\Psi = 0 \quad \text{with} \quad M = \frac{1}{4}(4\partial^3 - 3u\partial - 3\partial u) \tag{3.24}$$

The compatibility condition of eqs.(3.23,3.24) is the Lax equation $\partial_t L = [M, L]$, which is equivalent to the KdV equation.

Eq.(3.23) is the Schroedinger equation with potential u. The parameter λ gets an interpretation as a point of the spectrum of this operator. This is the origin of the terminology "spectral parameter".

3.3 Poisson brackets of the monodromy matrix.

As we just saw, the zero curvature equation leads to the construction of infinite set of local conserved quantities. We want to compute their Poisson brackets. For this we will compute the Poisson brackets of the matrix elements of the monodromy matrix.

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

$$\{U_1(\lambda, x), U_2(\mu, y)\} = [r_{12}(\lambda - \mu), U_1(\lambda, x) + U_2(\mu, y)]\delta(x - y) \tag{3.25}$$

We assume that $r_{12}(\lambda - \mu)$ is a r–matrix as in eq.(1.23). We say that the Poisson bracket eq.(3.25) is ultralocal due to the presence of $\delta(x-y)$ only. This hypothesis actually covers a large class of interesting integrable field theories, but certainly not all of them.

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

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

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. We use the usual tensor notation to arrange the table of their Poisson brackets.

Proposition 15 *If eq.(3.25) holds, we have the fundamental Sklyanin relation for the transport matrix:*

$$\{T_1(\lambda;y,x),T_2(\mu;y,x)\}=[r_{12}(\lambda,\mu),T_1(\lambda;y,x)T_2(\mu;y,x)] \tag{3.26}$$

As a consequence, the traces of powers of the monodromy matrix $H^{(n)}(\lambda)=\mathrm{Tr}\ (T^n(\lambda))$, generate Poisson commuting quantities:

$$\{H^{(n)}(\lambda),H^{(m)}(\mu)\}=0 \tag{3.27}$$

Proof. Let us first prove the relation (3.26) 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

$$\begin{aligned}\partial_x T(y,x)+T(y,x)U(x) &= 0\\ \partial_y T(y,x)-U(y)T(y,x) &= 0\end{aligned} \tag{3.28}$$

Since Poisson brackets satisfy the Leibnitz rules, we have

$$\{T_1(y,x),T_2(y,x)\}= \int_x^y\int_x^y dudv\ T_1(y,u)T_2(y,v)\{U_1(u),U_2(v)\}T_1(u,x)T_2(v,x) \tag{3.29}$$

Replacing $\{U_1(u),U_2(v)\}$ by eq.(3.25), and using the differential equation satisfied by $T(y,x)$ this yields:

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

Integrating this exact derivative gives the relation (3.26). Let us now show that the trace of the monodromy matrix $H^{(n)}(\lambda)$ generates Poisson commuting quantities. Eq.(3.26) implies

$$\{T_1^n(\lambda),T_2^m(\mu)\}=[r_{12}(\lambda,\mu),T_1^n(\lambda)T_2^m(\mu)]$$

We take the trace of this relation. In the left hand side we use the fact that $\mathrm{Tr}_{12}\,(A\otimes B)=\mathrm{Tr}\,(A)\mathrm{Tr}\,(B)$ and get $\{H^{(n)}(\lambda),H^{(m)}(\mu)\}$. The right hand side gives zero because it is the trace of a commutator.

Let us emphasis that it is the integration process involved in the transport matrix which leads from the linear Poisson bracket eq.(3.25) to the quadratic Sklyanin Poisson bracket eq.(3.26).

The proposition shows that we may take as Hamiltonian any element of the family generated by $H^{(n)}(\mu)$. We show that the corresponding equations of motion take the form of a zero curvature condition.

Proposition 16 *Taking $H^{(n)}(\mu)$ as Hamiltonian, we have*

$$\dot{U}(\lambda,x) \equiv \{H^{(n)}(\mu), U(\lambda,x)\} = \partial_x V^{(n)}(\lambda,\mu,x) + [V^{(n)}(\lambda,\mu,x), U(\lambda,x)] \tag{3.30}$$

where

$$V^{(n)}(\lambda,\mu;x) = n\mathrm{Tr}_1\Big(T_1(\mu;2\pi,x)r_{12}(\mu,\lambda)T_1(\mu;x,0)T_1^{n-1}(\mu,2\pi,0)\Big)$$

This provides the equations of motion for a hierarchy of times, when we expand in μ.

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

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

Using the differential equations (3.28) 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 $V_{12}(x) = T_1(2\pi,x)\cdot r_{12}\cdot T_1(x,0)$. From this we get $\{T_1^n(2\pi,0), U_2(x)\} = \partial_x V_{12}^{(n)}(x) + [V_{12}^{(n)}(x), U_2(x)]$ with $V_{12}^{(n)}(x) = \sum_i T_1^{n-i-1}V_{12}(x)T_1^i$. Taking the trace over the first space, remembering that $H^{(n)}(\mu) = \mathrm{Tr}\,T^n(\mu)$, and setting $V^{(n)}(\lambda,\mu,x) = \mathrm{Tr}_1 V_{12}^{(n)}(x)$, we find eq.(3.30).

3.4 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 field 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.

Dressing transformations are special non-local gauge transformations preserving the analytical structure of the Lax connection. These transformations are intimately related to the Riemann–Hilbert problem which we have discussed in the section on factorization.

We choose a contour Γ in the λ-plane such that none of the poles λ_k of the Lax connection are on Γ. We will take for Γ the sum of contours $\Gamma^{(k)}$ each one surrounding a pole λ_k as in the factorization problem.

To define the dressing transformation, we pick a group valued function $g(\lambda) \in \tilde{G}$ on Γ. From the Riemann–Hilbert problem, $g(\lambda)$ can be factorized as:

$$g(\lambda) = g_-^{-1}(\lambda) g_+(\lambda) \tag{3.31}$$

where $g_+(\lambda)$ and $g_-(\lambda)$ are analytic inside and ouside the contour Γ respectively. In the following discussion we assume that $g(\lambda)$ is close enough to the identity so that there are no indices.

Let U, V be a solution of the zero curvature equation eq.(3.1) with the prescribed singularities specified in eqs.(3.8,3.9). Let $\Psi \equiv \Psi(\lambda; x, t)$ be the solution of the linear system (3.2) normalized by $\Psi(\lambda; 0, 0) = 1$. We set:

$$\theta(\lambda; x, t) = \Psi(\lambda; x, t) \cdot g(\lambda) \cdot \Psi(\lambda; x, t)^{-1} \tag{3.32}$$

At each space-time point (x, t), we perform a λ decomposition of $\theta(\lambda, x, t)$ according to the Riemann-Hilbert problem as:

$$\theta(\lambda; x, t) = \theta_-^{-1}(\lambda; x, t) \cdot \theta_+(\lambda; x, t) \tag{3.33}$$

with θ_+ and θ_- analytic inside and outside the contour Γ respectively. Then,

Proposition 17 *The following function, defined for λ on the contour Γ,*

$$\Psi^g(\lambda; x, t) = \theta_\pm(\lambda; x, t) \cdot \Psi(\lambda; x, t) \cdot g_\pm^{-1}(\lambda) \tag{3.34}$$

extends to a function Ψ_+^g defined inside Γ except at the points λ_k where it has essential singularities and a function Ψ_-^g defined outside Γ. On Γ we have $\Psi_-^{g\,-1}\Psi_+^g|_\Gamma = 1$. So $\Psi_\pm^g$ define a unique function Ψ^g which is normalized by $\Psi^g(\lambda, 0) = 1$ and is solution of the linear system (3.2) with Lax connection U^g and V^g given by

$$U^g(\lambda; x, t) = \theta_\pm \cdot U \cdot \theta_\pm^{-1} + \partial_x \theta_\pm \cdot \theta_\pm^{-1} \tag{3.35}$$

$$V^g(\lambda; x, t) = \theta_\pm \cdot V \cdot \theta_\pm^{-1} + \partial_t \theta_\pm \cdot \theta_\pm^{-1} \tag{3.36}$$

The matrices U^g and V^g, which satisfy the zero curvature equation (3.1), are meromorphic functions on the whole complex λ plane with the same analytic structure as the components $U(\lambda)$ and $V(\lambda)$ of the original Lax connection.

Proof. First it follows directly from the definitions of $g_\pm$ and $\theta_\pm$ that for λ on Γ,

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

so that, the two expressions of the right hand side of eq.(3.34) with the $+$ and $-$ signs are equal, and effectively define a unique function Ψ^g on Γ. It is clear that this function can

be extended into two functions $\Psi^g_\pm$ respectively defined inside and outside this contour by:

$$\Psi^g_\pm = \theta_\pm \cdot \Psi \cdot g_\pm^{-1}$$

These functions have the same essential singularities as Ψ at the points λ_k. By construction, they are such that ${\Psi^g_-}^{-1}\Psi^g_+|_\Gamma = 1$.
We may use $\Psi^g_\pm$ to define the Lax connection $U^g_\pm$, $V^g_\pm$ inside and ouside the contour Γ. Explicitly:

$$\begin{aligned} U^g_\pm &= \partial_x \Psi^g_\pm \cdot \Psi^{g-1}_\pm = \partial_x \theta_\pm \theta_\pm^{-1} + \theta_\pm U \theta_\pm^{-1} \\ V^g_\pm &= \partial_t \Psi^g_\pm \cdot \Psi^{g-1}_\pm = \partial_t \theta_\pm \theta_\pm^{-1} + \theta_\pm V \theta_\pm^{-1} \end{aligned}$$

Since ${\Psi^g_-}^{-1}\Psi^g_+|_\Gamma = 1$ we see that U_+ coincides with U_- on the contour Γ and similarly $V_+ = V_-$ for $\lambda \in \Gamma$ and hence the pairs $U^g_\pm$, $V^g_\pm$ define a conection U^g, V^g on the whole λ–plane. Since $\theta_\pm$ are regular in their respective domains of definition, we see that U^g, V^g have the same singularities as U, V.

This proposition effectively states that the dressing transformations (3.34) map solutions of the equations of motion into new solutions. Given a solution U, V of the zero curvature equation with the prescribed pole structure and an element of the loop group $\tilde{G}$, we produce a new solution of the zero curvature equation with same analytical structure. But since this analytic structure is the main information which specifies the model we have produced a new solution of the equations of motion.

3.5 Soliton solutions.

In general, a matrix Riemann–Hilbert problem like eq.(3.31) cannot be solved explicitly by analytical methods. This statement applies to the fundamental solution of the Riemann Hilbert problem i.e., the one satisfying the conditions $\det \theta_\pm \neq 0$. However, once the fundamental solution is known, new solutions "with zeroes" can easily be constructed from it. 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 soliton solutions.

Let $\tilde{\theta}_\pm(\lambda)$ be the fundamental solution of a Riemann–Hilbert problem. A solution of the Riemann–Hilbert problem with zeroes at $\mu_1, \cdots, \mu_N$, $\lambda_1, \cdots, \lambda_N$ is.

$$\boxed{\theta_+(\lambda) = \chi_N^{-1}\left(1 - \frac{\mu_N - \lambda_N}{\lambda - \lambda_N} P_N\right) \cdots \chi_1^{-1}\left(1 - \frac{\mu_1 - \lambda_1}{\lambda - \lambda_1} P_1\right) \tilde{\theta}_+(\lambda)}$$

$$\boxed{\theta_-^{-1}(\lambda) = \tilde{\theta}_-^{-1}(\lambda)\left(1 - \frac{\lambda_1 - \mu_1}{\lambda - \mu_1} P_1\right) \chi_1 \cdots \left(1 - \frac{\lambda_N - \mu_N}{\lambda - \mu_N} P_N\right) \chi_N}$$

where P_i are projectors $P_i^2 = P$. We assume that the P_i and the χ_i are independent of λ. When $\lambda = \mu_i$, which we assume inside Γ, then $\theta_+(\lambda)$ contains as a factor the projector $(1 - P_i)$ so that $\det \theta(\mu_i) = 0$. Similarly, if $\lambda = \lambda_i$, which we assume outside Γ, then $\theta_-^{-1}(\lambda)$ contains as a factor the projector $(1 - P_i)$ so that $\det \theta(\lambda_i) = 0$. Hence the name Riemann-Hilbert problem with zeroes. We now extend the method of dressing transformations to the case of a Riemann–Hilbert problem with zeroes.

Let

$$\Theta_+^{(n)} = \chi_{n-1}^{-1}\left(1 - \frac{\mu_{n-1} - \lambda_{n-1}}{\lambda - \lambda_{n-1}} P_{n-1}\right) \cdots \chi_1^{-1}\left(1 - \frac{\mu_1 - \lambda_1}{\lambda - \lambda_1} P_1\right) \tilde{\theta}_+(\lambda)\Big|_{\lambda=\mu_n}$$

$$\Theta_-^{(n)-1} = \tilde{\theta}_-^{-1}(\lambda)\left(1 - \frac{\lambda_1 - \mu_1}{\lambda - \mu_1} P_1\right) \chi_1 \cdots \left(1 - \frac{\lambda_{n-1} - \mu_{n-1}}{\lambda - \mu_{n-1}} P_{n-1}\right) \chi_{n-1}\Big|_{\lambda=\lambda_n}$$

Proposition 18 *Given a Lax connection satisfying the zero curvature condition and the associated wave function $\Psi(\lambda, x, t)$, and given vector spaces $\mathcal{V}_n(0), \mathcal{W}_n(0)$ we define uniquely the projectors P_n by*

$$\boxed{\text{Ker } P_n(x,t) = \Theta_-^{(n)}\ \Psi(\lambda_n, x, t)\mathcal{V}_n(0)} \tag{3.37}$$

$$\boxed{\text{Im } P_n(x,t) = \Theta_+^{(n)}\ \Psi(\mu_n, x, t)\mathcal{W}_n(0)} \tag{3.38}$$

Then for any $g(\lambda) = g_-^{-1}(\lambda)g_+(\lambda)$ on Γ, the transformation $\Psi \to \Psi^g$,

$$\Psi^g = \theta_\pm \Psi g_\pm^{-1}, \quad \theta_-^{-1}\theta_+ = \Psi^{-1} g \Psi$$

is a dressing transformation, i.e. preserves the analytic structure of the Lax connection.

Proof. We start with the linear system

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

and dress it with a solution with zeroes of the Riemann–Hilbert problem, according to eqs.(3.35, 3.36):

$$U^g = \theta_\pm \cdot U \cdot \theta_\pm^{-1} + \partial_x \theta_\pm \cdot \theta_\pm^{-1}, \quad V^g = \theta_\pm \cdot V \cdot \theta_\pm^{-1} + \partial_t \theta_\pm \cdot \theta_\pm^{-1}$$

In general, the components of the dressed connection will have simple poles at the points μ_n, λ_n. We must require that the residues of these poles vanish. At $\lambda = \mu_n$, isolating the terms containing P_n, we have

$$\theta_+(\lambda) \simeq M_n(1 - P_n)\Theta_+^{(n)}, \quad \theta_+^{-1}(\lambda) \simeq \frac{\mu_n - \lambda_n}{\lambda - \mu_n}\Theta_+^{(n)-1} P_n M_n^{-1}$$

so that

$$\theta_+(\partial_x - U)\theta_+^{-1} \simeq \frac{\mu_n - \lambda_n}{\lambda - \mu_n} M_n\ (1 - P_n)\ \Theta_+^{(n)}(\partial_x - U|_{\mu_n})\Theta_+^{(n)-1}\ P_n\ M_n^{-1}$$

To kill the pole at $\lambda = \mu_n$, we choose

$$\text{Im } P_n = \Theta_+^{(n)}\Psi(\mu_n, x, t)\mathcal{W}(0)$$

A vector in Im P_n is a linear combination, with coefficients possibly depending on x, t, of the column of the matrix in the right hand side. Note that the factor $1 - P_n$ is important to kill the terms coming from this x, t dependence of the coefficients. Clearly the same argument works for $\partial_t - V$. Similarly, when $\lambda \simeq \lambda_n$ we have

$$\theta_-^{-1}(\lambda) \simeq \Theta_-^{(n)-1}(1 - P_n)N_n, \quad \theta_-(\lambda) \simeq \frac{\lambda_n - \mu_n}{\lambda - \lambda_n} N_n^{-1}\ P_n\ \Theta_-^{(n)}$$

so that

$$\theta_-(\partial_x - U)\theta_-^{-1} \simeq \frac{\lambda_n - \mu_n}{\lambda - \lambda_n} N_n^{-1}\ P_n\ \Theta_-^{(n)}(\partial_x - U|_{\lambda_n})\Theta_-^{(n)-1}(1 - P_n)N_n$$

To kill the pole we choose

$$\text{Ker } P_n = \text{Im } (1 - P_n) = \Theta_-^{(n)}\Psi(\lambda_n, x, t)\mathcal{V}(0)$$

Clearly the same argument works with $\partial_t - V$.

The interest of this procedure is that it yields non trivial results even if the Riemann–Hilbert problem is trivial i.e. $g(\lambda) = \text{Id}$. Then its fundamental solution is also trivial $\tilde{\theta}_\pm(\lambda) = \text{Id}$, and the solutions with zeroes are constructed by purely algebraic means. The resulting $\theta_\pm(\lambda)$ are rational functions of λ.

To make this method effective, we need a simple solution of the zero curvature condition $\partial_t U - \partial_x V - [V, U] = 0$ to start with. Simple solutions can be found in the form

$$U = A(\lambda, x), \quad V = B(\lambda, t), \quad [A, B] = 0$$

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 soliton solutions.

3.5.1 KdV solitons.

Let us illustrate this construction for one soliton solution of the KdV equation. Recall the famous KdV soliton:

$$\boxed{u_{\text{soliton}}(x, t) = -\frac{2k^2}{\cosh(kx + k^3 t)}} \tag{3.39}$$

We obtain it by dressing the vacuum solution $u(x,t)=0$. The solution of the vacuum linear system reads

$$\Psi(x,t)=e^{(x+\lambda t)U(\lambda)},\quad U(\lambda)=\begin{pmatrix}0&1\\ \lambda&0\end{pmatrix}$$

We consider the trivial Riemann-Hilbert problem ($g(\lambda)=1$).

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

Let

$$\theta_-^{-1}(\lambda)=\left(1-\frac{k^2}{\lambda}P\right)\chi,\quad \theta_+(\lambda)=\chi^{-1}\left(1+\frac{k^2}{\lambda-k^2}P\right)$$

Hence $\theta_-^{-1}(\lambda)$ has a zero at $\lambda=k^2$ and $\theta_+(\lambda)$ at $\lambda=0$. Since

$$\Psi(0,x,t)=\begin{pmatrix}0&x\\ 0&0\end{pmatrix},\quad \Psi(k^2,x,t)=\begin{pmatrix}\cosh(kx+k^3t)&k^{-1}\sinh(kx+k^3t)\\ k\sinh(kx+k^3t)&\cosh(kx+k^3t)\end{pmatrix}$$

we can choose

$$\text{Im } P=\begin{pmatrix}1\\0\end{pmatrix},\quad \text{Ker } P=\begin{pmatrix}k^{-1}\tanh(kx+k^3t)\\1\end{pmatrix}$$

so that

$$P=\begin{pmatrix}1&-k^{-1}\tanh(kx+k^3t)\\0&0\end{pmatrix}$$

Then, we find

$$U_{\text{soliton}}=\chi^{-1}\begin{pmatrix}-k\tanh(kx+k^3t)&1\\ \lambda-k^2&k\tanh(kx+k^3t)\end{pmatrix}\chi-\chi^{-1}\partial_x\chi$$

We see at this stage that the λ dependence is already essentially correct. Choosing

$$\chi=\begin{pmatrix}1&0\\ k\tanh(kx+k^3t)&1\end{pmatrix}$$

we find

$$U_{\text{soliton}}=\begin{pmatrix}0&1\\ \lambda-u_{\text{soliton}}(x,t)&0\end{pmatrix}$$

with $u_{\text{soliton}}(x,t)$ given by eq.(3.39). We can repeat the procedure, and reach finally the N-solitons solution

$$\boxed{u_{\text{N-solitons}}(x,t)=-2\frac{\partial^2}{\partial x^2}\log\tau_N}$$

where

$$\boxed{\tau_N=\det(1+W),\quad W_{ij}=\frac{\sqrt{X_iX_j}}{k_i+k_j},\quad X_i=a_ie^{2(k_ix+k_i^3t)}}$$

3.6 Finite zones solutions.

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:

$$(\partial_x - U(\lambda))\Psi = 0 \tag{3.40}$$
$$(\partial_t - V(\lambda))\Psi = 0 \tag{3.41}$$

whose compatibility conditions are equivalent to the field equations. The situation is very different as compared to the finite dimensional case. As we saw, the analog of the spectral curve is

$$\det(T(\lambda) - \mu) = 0 \tag{3.42}$$

where $T(\lambda)$ is the monodromy matrix of the linear system (3.40,3.41). This equation defines an algebraic curve of *infinite genus* and the analytical tools must be carefully adapted.

If however, we restrict our goal to find only particular solutions to eqs.(3.40, 3.41), then we can look at situations where the curve eq.(3.42) is infinitely degenerate leaving only a finite genus curve.

One common way to do that is as follows. The two equations (3.40, 3.41) are exactly of type of eq.(**??**) whose solution was built using the usual analytical tools.

To interprete the two equations (3.40,3.41) as evolution equations with respect to two different "times" for a system with finite number of degrees of freedom we need a Lax matrix $L(\lambda)$ satisfying

$$[\partial_x - U(\lambda), L(\lambda)] = 0 \tag{3.43}$$
$$[\partial_t - V(\lambda), L(\lambda)] = 0$$

To exhibit such Lax matrices, we consider the higher order flows associated to higher Hamiltonians. They provide 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 *stationnary* for some given time t_n, i.e. $\partial_{t_n} V_i = 0$ for all i. The zero curvature conditions $F_{ni} = 0$ reduce to a system of Lax equations:

$$\frac{dL}{dt_i} = [M_i, L], \quad i = 1, \cdots, \infty \quad \text{with } L = V_n,\ M_i = V_i$$

This is an integrable hierarchy for a finite–dimensional dynamical system described by the Lax matrix $L(\lambda)$. 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.

3.7 The Its-Matveev formula.

Let us apply these ideas to construct solutions of the KdV equation. As explained, 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_j.

A very simple example of this situation occurs when u is stationary with respect to the first time $t_3 = t$. In that case the Lax matrix is V given in eq.(3.20). The associated spectral curve is:

$$\Gamma : \quad \det(V-\mu) = \mu^2 - \frac{1}{4}\lambda^3 + \frac{1}{4}(3u^2 - u'')\lambda + \frac{1}{16}(2uu'' - u'^2 - 4u^3) = 0$$

The zero–curvature condition becomes the Lax equation $\partial_x V = [U, V]$ 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

$$\mu^2 = \lambda^3/4 - C_1\lambda/4 - C_2/16$$

It is independent of x as it should be. This is a genus 1 curve and u is given by the Weierstrass function.

$$u(x) = 2\wp(x + \zeta)$$

For higher times the matrices $V_{t_{2j-1}}$ are defined by

$$\partial_{t_{2j-1}} \begin{pmatrix} \Psi \\ \partial_x \Psi \end{pmatrix} = \begin{pmatrix} (L^{\frac{2j-1}{2}})_+ \Psi \\ \partial_x (L^{\frac{2j-1}{2}})_+ \Psi \end{pmatrix} = V_{t_{2j-1}} \begin{pmatrix} \Psi \\ \partial_x \Psi \end{pmatrix}$$

The matrices $V_{t_{2j-1}}$ are not hard to compute but it is quite clear that they are 2×2 traceless matrices. Stationarity with respect to any higher time always lead to an hyperelliptic spectral curve and we will just retain this feature.

Hence we consider an hyperelliptic curve of the generic form:

$$\Gamma : \quad \mu^2 = R(\lambda) = \prod_{i=1}^{2g+1} (\lambda - \lambda_i) \tag{3.44}$$

The point at ∞ is a branch point, and a local parameter around that point is $z = \sqrt{\lambda}$. Our goal is to construct a function Ψ on Γ satisfying the equations

$$\boxed{L\Psi = (\partial_x^2 - u)\Psi = \lambda\Psi, \quad \partial_{t_{2j-1}}\Psi = (L^{\frac{2j-1}{2}})_+ \Psi} \tag{3.45}$$

for some potential u. If we succeed to do it, $u(x, t_3, t_5, \cdots)$ will have to satisfy the KdV hierarchy equations by consistency.

Baker–Akhiezer functions are special functions with essential singularities on Riemann surfaces. It is a fact that there exists a *unique* function Ψ on Γ with the following analytic properties

- It has an essential singularity at the point P at infinity:

$$\Psi(t, z) = e^{\xi(t,z)}\Big(1 + \frac{\alpha(t)}{z} + O(1/z^2)\Big) \tag{3.46}$$

  where $z = \sqrt{\lambda}$ and $\xi(t, z) = \sum_{i\geq 1} z^{2i-1} t_{2i-1}$.

- It has g simple poles, independent of all times. The divisor of these poles is denoted $D = (\gamma_1, \cdots, \gamma_g)$.

This function is called a Baker–Akhiezer function. Even though Baker–Akhiezer functions are not meromorphic functions, they have the same number of poles and zeroes. In fact the differential form $d(\log \Psi)$ is a meromorphic form. The sum of its residues is the number of zeroes minus the number of poles of Ψ and this has to vanish. The essential singularity does not contribute because around P we have $d(\log \Psi) = d\xi(t, z) + \text{regular}$ and $d\xi(t, z)$ has no residue (remember that $1/z$ is the local parameter around P).

The uniqueness is then clear. If we have two such functions with the same singularity structure and divisor D, their ratio is a meromorphic function on Γ with g poles which can only be a constant.

The existence will be proved by giving an explicit formula for Ψ in terms of θ functions.

But before that, we will prove that this Baker-Akhiezer function solves the KdV hierarchy equations. Let Ψ be the Baker-Akhiezer function as above with the following behaviour at infinity:

$$\boxed{\Psi = e^{\xi(t,z)}(1 + O(z^{-1})), \quad \xi(t, z) = \sum_{j=1}^{\infty} z^{2j-1} t_{2j-1}}$$

Then we have

Proposition 19 *There exists a function $u(x, t)$ such that*

$$(\partial_x^2 - u)\ \Psi = \lambda\Psi \tag{3.47}$$

Proof. Consider on Γ the function $\partial_x^2\Psi - \lambda\Psi$. To define this object as a function on the curve, λ is viewed as a meromorphic function on Γ. Remark that λ has only a double pole at ∞ and such a function exists only if Γ is hyperelliptic. We see that $\partial_x^2\Psi - \lambda\Psi$ has the same analytical properties as Ψ itself at finite distance on Γ. At infinity we have by eq.(3.46):

$$\partial_x^2\Psi - \lambda\Psi = e^{\xi(t,z)}(2\partial_x\alpha + O(1/z)), \quad z = \sqrt{\lambda}$$

So it is a Baker–Akhiezer function, but with a normalization $2\partial_x\alpha$ instead of 1 at infinity. By the uniqueness theorem of such functions, we have:

$$\partial_x^2\Psi - \lambda\Psi = u\Psi, \quad u = 2\partial_x\alpha \tag{3.48}$$

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 20 *The evolution of Ψ 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 $\partial_{t_{2i-1}}\Psi = z^{2i-1}\Psi + e^{\xi}O(1/z)$ and $(L^{\frac{2i-1}{2}})_+\Psi = L^{\frac{2i-1}{2}}\Psi - (L^{\frac{2i-1}{2}})_-\Psi = z^{2i-1}\Psi + e^{\xi}O(1/z)$, where we have used $L\Psi = z^2\Psi$. Hence we get:

$$\partial_{t_{2i-1}}\Psi - (L^{\frac{2i-1}{2}})_+\Psi = e^{\xi(t,z)}\,O(z^{-1}) \quad z \to \infty$$

By unicity, this Baker–Akhiezer function which vanishes at ∞ vanishes identically.

We now give a fundamental formula expressing the Baker–Akhiezer functions in terms of Riemann theta functions. Recall that a differential of second kind is a meromorphic differential with poles of order ≥ 2. 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:

$$\Omega^{(2j-1)} = d\left(z^{2j-1} + \text{regular}\right), \quad \text{for} \quad z \to \infty$$

Let $U_k^{(2j-1)}$ be its b-periods:

$$U_k^{(2j-1)} = \frac{1}{2i\pi}\oint_{b_k}\Omega^{(2j-1)}$$

Define

$$\Omega = \sum_j \Omega^{(2j-1)}t_{2j-1}$$

and denote by $2i\pi U$ the vector of b-periods of Ω.

Proposition 21 *If $D = \sum_{i=1}^{g} \gamma_i$ is a generic divisor of degree g, the following expression defines a Baker–Akhiezer function with D as divisor of poles:*

$$\boxed{\Psi(P) = \text{const. } \exp\left(\int_{P_0}^{P} \Omega\right) \frac{\theta(\mathcal{A}(P) + U - \zeta)}{\theta(\mathcal{A}(P) - \zeta)}} \tag{3.49}$$

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.

Proof. It is enough to check that the function defined by the formula (3.49) 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 a–periodic and Ω is normalized. If P describes the b_j–cycle the quotient of theta functions is multiplied by $\exp(-2i\pi U_j$ while the exponential factor changes by $\exp(2i\pi U_j)$, so that Ψ is well–defined. Clearly it has the right poles if deg $D = g$. ∎

As a consequence, the normalized Baker-Akhiezer function with the divisor of poles $D = (\gamma_1, \cdots, \gamma_g)$ can be expressed as:

$$\boxed{\Psi(t, P) = e^{\int_{\infty}^{P} \Omega} \frac{\theta(\mathcal{A}(P) + U - \zeta)\, \theta(\zeta)}{\theta(\mathcal{A}(P) - \zeta)\, \theta(U - \zeta)}} \tag{3.50}$$

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 potential, u, is given by the Its–Matveev formula:

$$\boxed{u(x,t) = -2\partial_x^2 \log \theta\Big(\sum_j t_{2j-1} U^{(2j-1)} - \zeta\Big) + \text{const.}} \tag{3.51}$$

In fact, in eq.(3.50) the integral $\int_{\infty}^{P}$ has to be understood in the following sense: for z in a vicinity of ∞, one *defines* $\int_{\infty}^{P} \Omega^{(2j-1)}$ as the unique primitive of $\Omega^{(2j-1)}$ which behaves as $z^{2j-1} + O(1/z)$ (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.(3.50) 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 $\Omega^{(1)}(z) = d(z + \frac{\beta}{z} + O(z^{-2}))$, where β does not depend on times, we have:

$$\begin{aligned}\partial_x \log \Psi &= z + \partial_x \log \theta\Big(\mathcal{A}(P) + \sum_j t_{2j-1} U^{(2j-1)} - \zeta\Big) - \\ &\quad -\partial_x \log \theta\Big(\sum_j t_{2j-1} U^{(2j-1)} - \zeta\Big) + \frac{\beta}{z} + O(z^{-2})\end{aligned}$$

We evaluate this expression when $z \to \infty$. Using Riemann's bilinear identities, we can expand the Abel map $\mathcal{A}(P)$ around ∞, and we have:

$$\begin{aligned}\theta\left(\mathcal{A}(P) + \sum_j t_{2j-1} U^{(2j-1)} - \zeta\right) &= \theta\left(\sum_j \left(t_{2j-1} - \frac{z^{-2j+1}}{2j-1}\right) U^{(2j-1)} - \zeta\right) \\ &= \theta\left(\left(x - \frac{1}{z}\right) U^{(1)} + \left(t_3 - \frac{1}{3z^3}\right) U^{(3)} + \cdots - \zeta\right)\end{aligned}$$

Keeping the $1/z$ terms, we obtain:

$$\partial_x \log \Psi = z - \frac{1}{z} \partial_x^2 \log \theta\Big(\sum_j t_{2j-1} U^{(2j-1)} - \zeta\Big) + \frac{\beta}{z} + O(\frac{1}{z^2})$$

Differentiating once more with respect to x, we also get $\partial_x^2 \log \Psi = O(1/z)$. It follows that $z^2 + u = z^2 - 2\partial_x^2 \log \theta + 2\beta + O(1/z)$ proving the result.

Bibliography

[1] L.D.Faddeev, *"Integrable Models in 1+1 Dimensional Quantum Field Theory"*. Les Houches Lectures 1982. Elsevier Science Publishers (1984).

[2] M. Semenov-Tian-Shansky. *"Dressing transformations and Poisson group actions"*. Publ. RIMS 21 (1985), p. 1237.

[3] L. D. Faddeev, L. A. Takhtajan. **"Hamiltonian Methods in the Theory of Solitons"**. Springer 1986.

[4] L.A. Dickey **"Soliton Equations and Hamiltonian Systems."** World Scientific (1991)

[5] O.Babelon, D. Bernard, M. Talon, **Introduction to Classical Integrable Systems.** Cambridge University Press (2003).

Chapter 4

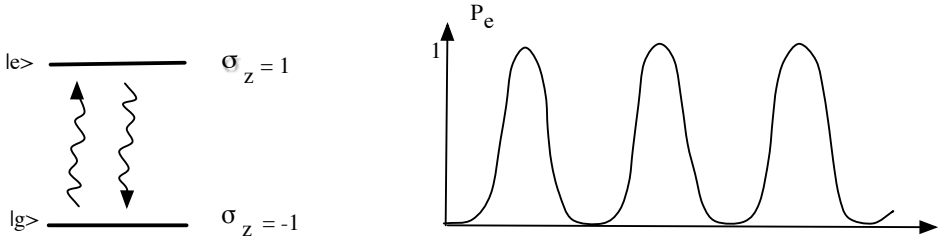
The Jaynes-Cummings-Gaudin model.

4.1 Physical context.

4.1.1 Rabi oscillations.

This is the interaction of a two levels atom and photons in a single cavity mode. When the electromagnetic field is classical, the system is described by the Hamiltonian

$$H = \hbar\omega_0 \frac{\sigma^z}{2} + \hbar\Omega[\sigma^+ b + \sigma^- b^\dagger]$$



We assume

$$b(t) = e^{-i\omega t} b_0$$

The resonnance condition is $\Delta = \omega_0 - \omega = 0$ but we keep $\Delta \neq 0$ for a while. Let us denote

$$|\psi\rangle = e^{-i\omega b_0^\dagger b_0 t} \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}$$

The Schroedinger equation becomes

$$\begin{pmatrix} \dot{\psi}_1 \\ \dot{\psi}_2 \end{pmatrix} = -i \begin{pmatrix} \frac{\omega_0}{2} & \Omega b_0 e^{-i\omega t} \\ \Omega b_0^\dagger e^{i\omega t} & -\frac{\omega_0}{2} \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}$$

From this we get a second order equation for ψ_1

$$\ddot{\psi}_1 - i\omega\dot{\psi}_1 + \left(\Omega^2 b_0^\dagger b_0 + \frac{1}{4}(\omega_0^2 - 2\omega_0\omega)\right)\psi_1 = 0$$

If the atom is in its fundamental state at time $t = 0$, i.e. $\psi_1(0) = 0$, then the solution is

$$\psi_1 = \alpha e^{-i\frac{\omega}{2}t}(e^{i\gamma t} - e^{-i\gamma t}), \quad \psi_2 = -\frac{\alpha}{2\Omega b_0}e^{i\frac{\omega}{2}t}((2\gamma + \Delta)e^{i\gamma t} + (2\gamma - \Delta)e^{-i\gamma t})$$

where we introduced the Rabi frequency

$$\Omega_{\text{Rabi}}^2 = \Omega^2(\kappa^2 + b_0^\dagger b_0), \quad \kappa = \frac{\Delta}{2\Omega}$$

The norm of the state is $\langle\psi|\psi\rangle = \psi_1^*\psi_1 + \psi_2^*\psi_2$. The probability to find the atom in the excited state at time t is

$$P_e(t) = \frac{\psi_1^*(t)\psi_1(t)}{\langle\psi|\psi\rangle} = \Omega^2 b_0^\dagger b_0 \frac{\sin^2\Omega_{\text{Rabi}}t}{\Omega_{\text{Rabi}}^2}$$

These are the famous Rabi oscillations, the amplitude is maximal at the resonnance $\Delta = 0$. Notice that

$$\frac{\langle\psi|\sigma^z|\psi\rangle}{\langle\psi|\psi\rangle} = \frac{\psi_1^*\psi_1 - \psi_2^*\psi_2}{\langle\psi|\psi\rangle} = 2\frac{\psi_1^*\psi_1}{\langle\psi|\psi\rangle} - 1$$

so that

$$\boxed{\frac{\langle\psi|\sigma^z|\psi\rangle}{\langle\downarrow|\sigma^z|\downarrow\rangle} = 1 - 2\Omega^2 b_0^\dagger b_0 \frac{\sin^2\Omega_{\text{Rabi}}t}{\Omega_{\text{Rabi}}^2}} \tag{4.1}$$

What happens if the electromagnetic field is quantum? and in particular if the number of photons is small ($5 \leq \bar{n} \leq 40$)? The system now is described by the Jaynes-Cummings Hamiltonian [1]

$$H = \hbar\omega_0\frac{\sigma^z}{2} + \hbar\omega b^\dagger b + \hbar\Omega[\sigma^+ b + \sigma^- b^\dagger]$$

Where we recall the usual commutation relations

$$[\sigma^z, \sigma^\pm] = \pm 2\sigma^\pm, \quad [\sigma^+, \sigma^-] = \sigma^z, \quad [b, b^\dagger] = 1$$

It turns out that the model is still exactly solvable. The key is the existence of an extra conserved quantity. Let

$$H_1 = b^\dagger b + \frac{1}{2}\sigma^z, \quad H_0 = \kappa\sigma^z + b^\dagger\sigma^- + \sigma^+ b, \quad \kappa = \frac{\Delta}{2\Omega}$$

We have

$$H = \omega H_1 + \Omega H_0, \quad [H_1, H_0] = 0$$

The Heisenberg equations of motion are

$$i\dot{s^z} = [H, s^z] = \Omega\Big(H_0 - 2\kappa s^z - 2bs^+\Big)$$

But for the spin 1/2 representation, we have

$$H_0 = \begin{pmatrix} \kappa & b \\ b^\dagger & -\kappa \end{pmatrix}$$

so that

$$bs^+ = (H_0 + \kappa)\left(\frac{1}{2} - s^z\right)$$

Hence, we get

$$\boxed{\dot{s^z} = i\Omega\Big(\kappa - 2H_0 s^z\Big)}$$

Since H_0 is conserved the solutions is extremely simple (keep the order of the operators)

$$\boxed{s^z(t) = \frac{\kappa}{2}H_0^{-1} + e^{-2i\Omega t\, H_0}\left(s^z(0) - \frac{\kappa}{2}H_0^{-1}\right)}$$

We now introduce the states

$$|n,\uparrow\rangle = (b^\dagger)^n|0\rangle \otimes |\uparrow\rangle, \quad |n,\downarrow\rangle = (b^\dagger)^n|0\rangle \otimes |\downarrow\rangle$$

We have

$$H_0|n,\uparrow\rangle = \kappa|n,\uparrow\rangle + |n+1,\downarrow\rangle, \quad H_0|n,\downarrow\rangle = n|n-1,\uparrow\rangle - \kappa|n,\downarrow\rangle$$

which implies

$$H_0^2|n,\uparrow\rangle = (\kappa^2 + n + 1)|n,\uparrow\rangle$$

Let us define

$$\Omega_n^2 = \Omega^2(\kappa^2 + n + 1)$$

It is then simple to show that

$$\langle m,\uparrow|s^z|n,\uparrow\rangle = \frac{1}{2}\delta_{n,m}\langle n|n\rangle\left[1 - 2\Omega^2(n+1)\frac{\sin^2\Omega_n t}{\Omega_n^2}\right]$$

We introduce the coherent states $b(0)|\alpha\rangle = \alpha|\alpha\rangle$.

$$|\alpha\rangle = e^{-\frac{1}{2}|\alpha|^2}\sum_{n=0}^{\infty}\frac{\alpha^n}{\sqrt{n!}}|n\rangle$$

For such states, the average number of photons in the cavity is $\bar{n} = |\alpha|^2$. We obtain

$$\boxed{\frac{\langle \alpha, \uparrow |\sigma^z(t)|\alpha, \uparrow\rangle}{\langle \alpha, \downarrow |\sigma^z(0)|\alpha, \downarrow\rangle} = e^{-|\alpha|^2} \sum_{n=0}^{\infty} \frac{|\alpha|^{2n}}{n!} \left(1 - 2\Omega^2(n+1)\frac{\sin^2 \Omega_n t}{\Omega_n^2}\right)}$$

Drawing this quantity as a function of time, we observe a collapse and resurgence phenomenon of the Rabi oscillations.

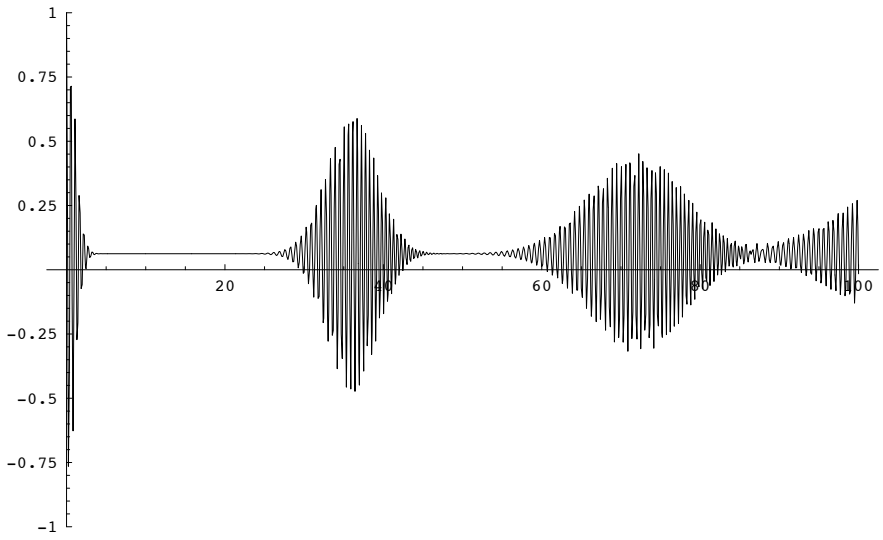


Figure 4.1: The collapses and revivals of Rabi oscillations. $\bar{n} = 30$, $\Delta = 2\sqrt{2}$, $\Omega = 1$.

4.1.2 Cold atoms condensates.

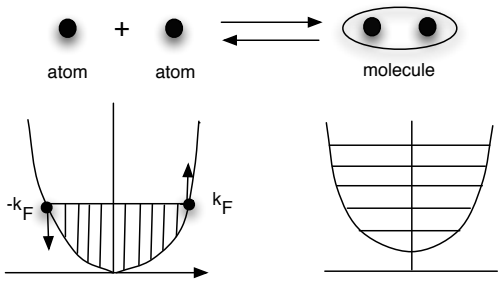
Consider alcali atoms like Li, K, Na, etc... Let $\vec{I}$ denotes the magnetic moment of the nucleus and $\vec{S}$ the spin of the electron. By choosing the isotope, we can arrange that the atom is a fermion or a boson. We will consider the case of a fermion. The Hamiltonian of a single atom in a magnetic field is

$$H = g\vec{I} \cdot \vec{S} + g_B \, \vec{S} \cdot \vec{B}$$

For two atoms far apart, the Hamiltonian is simply the sum of the Hamiltonians of the idividual atoms.

$$H = H_1 + H_2$$

When they come closer together however they start to interact. We consider a situation where we can have formation bound states or "molecules" between two atoms. Atoms are fermions, molecules are bosons.



Let $c^\dagger_{j\sigma}$ and $c_{j\sigma}$ be creation and annihilation operators for fermions in the state $\sigma =\uparrow$ or $\downarrow$ in an orbital of energy ϵ_j. Let $b^\dagger$ and b be creation annihilation operators of a molecule at zero momentum. The Hamiltonian of the boson-fermion condensate is

$$H = \sum_{j,\sigma} \epsilon_j c^\dagger_{j\sigma} c_{j\sigma} + \omega b^\dagger b + g \sum_j \left(b^\dagger c_{j\downarrow} c_{j\uparrow} + b c^\dagger_{j\uparrow} c^\dagger_{j\downarrow} \right) \tag{4.2}$$

This can be rewritten in terms of pseudo spins

$$2s^z_j = \sum_\sigma c^\dagger_{j\sigma} c_{j\sigma} - 1, \quad s^-_j = c_{j\downarrow} c_{j\uparrow}, \quad s^+_j = c^\dagger_{j\uparrow} c^\dagger_{j\downarrow} \tag{4.3}$$

we get

$$H = \sum_{j=0}^{n-1} 2\epsilon_j s^z_j + \omega b^\dagger b + g \sum_{j=0}^{n-1} \left(b^\dagger s^-_j + b s^+_j \right) \tag{4.4}$$

This is the Jaynes-Cummings-Gaudin Hamiltonian.

In the Born-Openheimer approximation the energy levels become a function of the distance $E \to E(r)$ as shown in Fig.[4.2]. A Feshbach resonnance occurs when a bound state becomes degenerate with a scattering state. By tuning the magnetic field, one can adjust the molecule state to be just above or below the atomic state. We can thus induce a transition in the system at will. A particularly interesting situation is the case of a soudain perturbation. At $t = 0$ the system is in an atomic state, and at $t > 0$ molecules start forming in the fundamental state (at zero temperature). What is the dynamical evolution of the system? What is the rate of formation of "Cooper" pairs?

4.2 Settings.

Let $b, b^\dagger$ and s_j be quantum operators

$$[b, b^\dagger] = \hbar, \quad [s^+_j, s^-_j] = 2\hbar s^z_j, \quad [s^z_j, s^\pm_j] = \pm\hbar s^\pm_j$$

We assume that s_j acts on a spin s representation

$$s^z_j |m_j\rangle = \hbar m_j |m_j\rangle, \quad s^\pm_j |m_j\rangle = \hbar\sqrt{s(s+1) - m_j(m_j \pm 1)}\, |m_j \pm 1\rangle, \quad m_j = -s, -s+1, \cdots, s-1, s$$

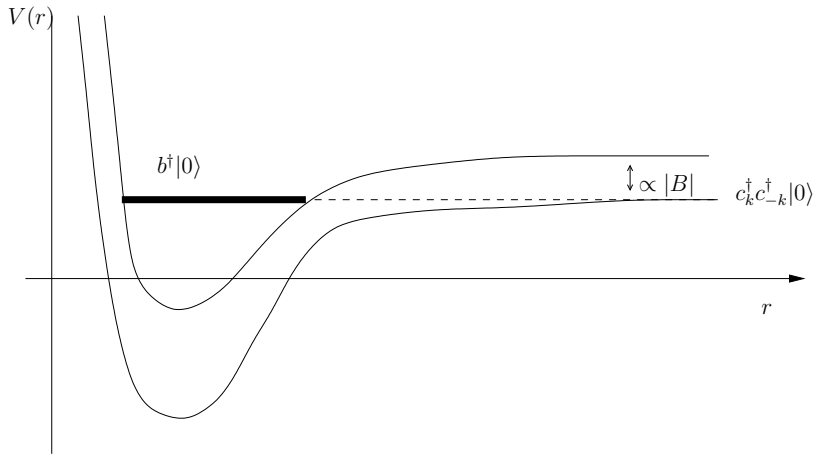


Figure 4.2: Feshbach resonnance.

where s is integer or half integer. Notice that

$$(s^+)^r \mid -s\rangle = \sqrt{\frac{2s!r!}{(2s-r)!}} \mid -s+r\rangle$$

so that

$$||(s^+)^r \mid -s\rangle||^2 = \frac{2s!\ r!}{(2s-r)!} = \frac{\Gamma(2s+1)\Gamma(r+1)}{\Gamma(2s-r+1)} \tag{4.5}$$

Hence, if $r > 2s$ the norm is automatically zero because the Γ function in the denominator has a pole. Similarly, we assume that $b, b^\dagger$ act on the Fock space $b^{\dagger n}|0\rangle$.

Instead of the representations above, we will work with the Bargman spaces. For the oscillator $b, b^\dagger$ this is the space

$$\mathcal{B}_b = \left\{ f(z), \text{entire function of } z \;\Big|\; \int |f(z)|^2 e^{-\frac{|z|^2}{\hbar}} dz d\bar{z} < \infty \right\}$$

On this space we have

$$b = \hbar \frac{d}{dz}, \quad b^\dagger = z \tag{4.6}$$

For the spin operators, following Sklyanin [5], we set

$$\begin{aligned} s^z &= \hbar \left(w \frac{d}{dw} - s \right) \\ s^+ &= \hbar\, w \\ s^- &= \hbar \left(-w \frac{d^2}{dw^2} + 2s \frac{d}{dw} \right) \end{aligned}$$

Notice that the value of the Casimir is

$$(s^z)^2 + \frac{1}{2}(s^+s^- + s^-s^+) = \hbar^2 s(s+1)$$

The lowest weight vector corresponds to the constant function 1. Other states are obtained by applying $s^+ \simeq w$. Let $\mathcal{I}$ be the ideal in the set of polynomials $C[w]$ generated by w^{2s+1}. The above operators are well defined on $C[w]/\mathcal{I}$. This amounts to showing that $s^a\mathcal{I} \subset \mathcal{I}$. We have only to check the dangerous case

$$s^- w^{2s+1} = -(2s+1)2sw^{2s} - 2s(2s+1)w^{2s} = 0$$

The representation space is $C[w]/\mathcal{I}$ and is of dimension $2s+1$.

4.3 Bethe Ansatz.

On this Hilbert space acts the Hamiltonian

$$H = \sum_{j=0}^{n-1} 2\epsilon_j s_j^z + \omega b^\dagger b + g\sum_{j=0}^{n-1}(b^\dagger s_j^- + b s_j^+)$$

Let

$$L(\lambda) = \begin{pmatrix} A(\lambda) & B(\lambda) \\ C(\lambda) & -A(\lambda) \end{pmatrix}$$

where

$$\begin{aligned} A(\lambda) &= \frac{2\lambda}{g^2} - \frac{\omega}{g^2} + \sum_{j=0}^{n-1} \frac{s_j^z}{\lambda - \epsilon_j} \\ B(\lambda) &= \frac{2b}{g} + \sum_{j=0}^{n-1} \frac{s_j^-}{\lambda - \epsilon_j} \\ C(\lambda) &= \frac{2b^\dagger}{g} + \sum_{j=0}^{n-1} \frac{s_j^+}{\lambda - \epsilon_j} \end{aligned}$$

are now quantum operators. It is very simple to check that

$$\begin{aligned} [A(\lambda), A(\mu)] &= 0 \\ [B(\lambda), B(\mu)] &= 0 \\ [C(\lambda), C(\mu)] &= 0 \\ [A(\lambda), B(\mu)] &= \frac{\hbar}{\lambda - \mu}(B(\lambda) - B(\mu)) \\ [A(\lambda), C(\mu)] &= -\frac{\hbar}{\lambda - \mu}(C(\lambda) - C(\mu)) \\ [B(\lambda), C(\mu)] &= \frac{2\hbar}{\lambda - \mu}(A(\lambda) - A(\mu)) \end{aligned}$$

One can rewrite these equations in the usual form

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

where

$$P_{12} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

We now show that

$$\boxed{\mathrm{Tr}\,(\mathrm{L}^2(\lambda)) = 2\mathrm{A}^2(\lambda) + \mathrm{B}(\lambda)\mathrm{C}(\lambda) + \mathrm{C}(\lambda)\mathrm{B}(\lambda)}$$

generate commuting quantities.

Proposition 22 *we have*

$$\boxed{[\mathrm{Tr}\,\mathrm{L}^2(\lambda), \mathrm{Tr}\,\mathrm{L}^2(\mu)] = 0}$$

Proof. First one has

$$[\mathrm{Tr}\,(\mathrm{L}^2(\lambda)), \mathrm{L}(\mu)]_{\mathrm{q}} = [\mathrm{M}(\lambda, \mu), \mathrm{L}(\mu)]_{\mathrm{aux}}$$

with

$$M(\lambda, \mu) = -2\hbar \frac{L(\lambda) - L(\mu)}{\lambda - \mu}$$

where we distinguished the commutators in the quantum space and the commutator in the auxiliary space. Alternatively, we have

$$[\mathrm{Tr}\,(\mathrm{L}^2(\lambda)), \mathrm{L}_2(\mu)]_{\mathrm{q}} = -\frac{2\hbar}{\lambda - \mu}[\mathrm{L}_2(\lambda), \mathrm{L}_2(\mu)]_{\mathrm{aux}} = -\frac{2\hbar}{\lambda - \mu}\mathrm{Tr}_1[\mathrm{P}_{12}\mathrm{L}_1(\lambda), \mathrm{L}_2(\mu)] \quad (4.7)$$

It follows that

$$[\mathrm{Tr}\,(\mathrm{L}^2(\lambda)), \mathrm{L}_2^2(\mu)]_{\mathrm{q}} = -\frac{2\hbar}{\lambda - \mu}[\mathrm{L}_2(\lambda), \mathrm{L}_2^2(\mu)]_{\mathrm{aux}} = -\frac{2\hbar}{\lambda - \mu}\mathrm{Tr}_1[\mathrm{P}_{12}\mathrm{L}_1(\lambda), \mathrm{L}_2^2(\mu)]$$

and therefore

$$\begin{aligned} [\mathrm{Tr}\,(\mathrm{L}^2(\lambda)), \mathrm{Tr}\,(\mathrm{L}^2(\mu))]_{\mathrm{q}} &= -\frac{2\hbar}{\lambda - \mu}\mathrm{Tr}_{12}P_{12}[L_1(\lambda), L_2^2(\mu)] \\ &= \frac{2\hbar^2}{(\lambda - \mu)^2}\mathrm{Tr}_{12}P_{12}\Big([P_{12}, L_1(\lambda) + L_2(\mu)]L_2(\mu) + L_2(\mu)[P_{12}, L_1(\lambda) + L_2(\mu)]\Big) \end{aligned}$$

Expanding the four terms and using $P_{12}^2 = 1$ and the cyclicity of the trace (for P_{12} only) we arrive at

$$[\mathrm{Tr}\,(\mathrm{L}^2(\lambda)), \mathrm{Tr}\,(\mathrm{L}^2(\mu))] = \frac{2\hbar^2}{(\lambda-\mu)^2}\mathrm{Tr}_{12}\Big([\mathrm{L}_1(\lambda), \mathrm{L}_2(\mu)] - \mathrm{P}_{12}[\mathrm{L}_1(\lambda), \mathrm{L}_2(\mu)]\Big)$$

The first term in the right hand side vanishes because Tr $L(\lambda) = 0$. The second term vanishes too because

$$\begin{aligned}\mathrm{Tr}_{12}\, P_{12}[L_1(\lambda), L_2(\mu)] &= \frac{\hbar}{\lambda-\mu}\mathrm{Tr}\,\mathrm{P}_{12}[\mathrm{P}_{12}, \mathrm{L}_1(\lambda) + \mathrm{L}_2(\mu)] \\ &= \frac{\hbar}{\lambda-\mu}\mathrm{Tr}\,(\mathrm{L}_1(\lambda) + \mathrm{L}_2(\mu))[\mathrm{P}_{12}, \mathrm{P}_{12}] = 0\end{aligned}$$

■

We are now in a position to write the Bethe Ansatz. Let

$$|0\rangle = |0\rangle \otimes |-s_1\rangle \otimes \cdots \otimes |-s_n\rangle, \quad b|0\rangle = 0, \quad s_j^-|0, -s_j\rangle = 0$$

We have

$$\boxed{B(\lambda)|0\rangle = 0}$$

and

$$\boxed{A(\lambda)|0\rangle = a(\lambda)|0\rangle}$$

$$\boxed{a(\lambda) = \frac{2\lambda}{g^2} - \frac{\omega}{g^2} - \sum_j \frac{\hbar s_j}{\lambda - \epsilon_j}}$$

Since

$$[B(\lambda), C(\lambda)] = 2\hbar A'(\lambda)$$

we also have

$$B(\lambda)C(\lambda)|0\rangle = 2\hbar a'(\lambda)|0\rangle$$

With all this we have

$$\frac{1}{2}\mathrm{Tr}\,\mathrm{L}^2(\lambda)|0\rangle = (\mathrm{a}^2(\lambda) + \hbar \mathrm{a}'(\lambda))|0\rangle$$

Let

$$\boxed{\Omega(\mu_1, \mu_2, \cdots, \mu_M) = C(\mu_1)C(\mu_2)\cdots C(\mu_M)|0\rangle}$$

For the following, we need the

$$[\mathrm{Tr\,L}^2(\lambda), \mathrm{C}(\mu)] = \frac{4\hbar}{\lambda-\mu}(\mathrm{C}(\mu)\mathrm{A}(\lambda) - \mathrm{C}(\lambda)\mathrm{A}(\mu))$$

which is obtained from eq.(4.7) by pushing the A's to the right. We recall also that

$$[A(\lambda), C(\mu)] = -\frac{\hbar}{\lambda-\mu}(C(\lambda) - C(\mu)) \tag{4.8}$$

Proposition 23 *One has*

$$\begin{aligned}\frac{1}{2}\mathrm{Tr\,L}^2(\lambda)\Omega(\mu_1,\mu_2,\cdots,\mu_\mathrm{M}) &= \Lambda(\lambda,\mu_1,\cdots,\mu_M)\Omega(\mu_1,\mu_2,\cdots,\mu_M) &(4.9)\\ &+\sum_i \Lambda_i(\lambda,\mu_1,\cdots,\mu_M)\Omega(\lambda,\mu_1,\cdots,\hat{\mu_i},\cdots,\mu_M) &(4.10)\end{aligned}$$

The first term is called the wanted term, and the other ones are called the unwanted terms. Their coefficients are respectively

$$\Lambda(\lambda,\mu_1,\cdots,\mu_M) = a^2(\lambda) + \hbar a'(\lambda) + \sum_i \frac{2\hbar}{\lambda-\mu_i}a(\lambda) + 2\sum_i\sum_{j>i}\frac{\hbar}{\lambda-\mu_i}\frac{\hbar}{\lambda-\mu_j} \tag{4.11}$$

$$\Lambda_i(\lambda,\mu_1,\cdots,\mu_M) = \frac{2\hbar}{\lambda-\mu_i}\left(a(\mu_i) + \sum_{j\neq i}\frac{\hbar}{\mu_i-\mu_j}\right) \tag{4.12}$$

Proof. We have

$$\begin{aligned}\frac{1}{2}\mathrm{Tr\,L}^2(\lambda)\Omega(\mu_1,\mu_2,\cdots,\mu_\mathrm{M}) &= (a^2(\lambda)+\hbar a'(\lambda))\Omega(\mu_1,\mu_2,\cdots,\mu_M) &(4.13)\\ &+\frac{1}{2}[\mathrm{Tr\,L}^2(\lambda), \mathrm{C}(\mu_1)\mathrm{C}(\mu_2)\cdots\mathrm{C}(\mu_\mathrm{M})]|0\rangle\end{aligned}$$

But

$$\begin{aligned}\frac{1}{2}[\mathrm{Tr\,L}^2(\lambda), \mathrm{C}(\mu_1)\mathrm{C}(\mu_2)\cdots\mathrm{C}(\mu_\mathrm{M})]|0\rangle &= &(4.14)\\ = \sum_i \frac{2\hbar}{\lambda-\mu_i}C(\mu_1)\cdots\Big(C(\mu_i)A(\lambda) - C(\lambda)A(\mu_i)\Big)C(\mu_{i+1})\cdots C(\mu_M)|0\rangle\end{aligned}$$

We now push $A(\lambda)$ and $A(\mu_i)$ to the right, using eq.(4.8). Clearly, when we do so we will generate terms only of the form (remember that $|0\rangle$ is an eigenvector of both $A(\lambda)$ and $A(\mu_i)$)

$$\begin{aligned}\Omega(\mu_1,\mu_2,\cdots,\mu_M) &= C(\mu_1)C(\mu_2)\cdots C(\mu_M)|0\rangle\\ \Omega(\lambda,\mu_1,\cdots,\hat{\mu_i},\cdots,\mu_M) &= C(\lambda)C(\mu_1)\cdots\widehat{C(\mu_i)}\cdots C(\mu_M)|0\rangle\end{aligned}$$

The wanted term cannot come from the term $C(\lambda)A(\mu_i)$ in the above formula because one of the $C(\mu_i)$ has already replaced its argument by λ and there is no way to recover it. Hence one has to use the first term $C(\mu_i)A(\lambda)$ and push $A(\lambda)$ to the right using only the second term in eq.(4.8). We get in this unique way

$$\left(\sum_i \frac{2\hbar}{\lambda-\mu_i}a(\lambda) + 2\sum_i\sum_{j>i}\frac{\hbar}{\lambda-\mu_i}\frac{\hbar}{\lambda-\mu_j}\right)\Omega(\mu_1,\mu_2,\cdots,\mu_M)$$

Adding the wanted contribution, eq.(4.13), we obtain eq.(4.11). Let us see now how to get the first unwanted term, the one where μ_1 has been replaced by λ. Clearly this term has to come from the term $C(\lambda)A(\mu_1)$ in eq.(4.14). Then one has to push $A(\mu_1)$ to the right using only the second term in eq.(4.8). We get in this unique way

$$\frac{2\hbar}{\lambda-\mu_1}\left(a(\mu_1)+\sum_{j\neq 1}\frac{\hbar}{\mu_1-\mu_j}\right)\Omega(\lambda,\mu_2,\cdots,\mu_M)$$

Since $\Omega(\mu_1,\mu_2,\cdots,\mu_M)$ is completely symmetrical in the μ_i, we have proved eq.(4.12). ■

The unwanted terms vanish if the Bethe equations are satisfied

$$\boxed{a(\mu_i)+\sum_{j\neq i}\frac{\hbar}{\mu_i-\mu_j}=0} \tag{4.15}$$

Taking into account these conditions, the eigenvalue can be rewritten as

$$\boxed{\Lambda(\lambda,\mu_1,\cdots,\mu_M)=a^2(\lambda)+\hbar a'(\lambda)+2\hbar\sum_i\frac{a(\lambda)-a(\mu_i)}{\lambda-\mu_i}} \tag{4.16}$$

4.4 Riccati equation.

We now analyse the Bethe equations eqs.(6.8). We introduce the function

$$\boxed{S(z)=\sum_i\frac{1}{z-\mu_i}}$$

Proposition 24 *The Bethe equations (6.8) imply the following Riccati equation on $S(z)$*

$$\boxed{S'(z)+S^2(z)+\frac{2}{\hbar g^2}\Big((2z-\omega)S(z)-2M\Big)=\sum_j 2s_j\frac{S(z)-S(\epsilon_j)}{z-\epsilon_j}} \tag{4.17}$$

Proof. The Bethe equations read

$$\frac{2\mu_i}{g^2} - \frac{\omega}{g^2} - \sum_j \frac{\hbar s_j}{\mu_i - \epsilon_j} + \sum_{j\neq i} \frac{\hbar}{\mu_i - \mu_j} = 0$$

we multiply by $1/(z-\mu_i)$ to get

$$\frac{2}{g^2}\frac{\mu_i}{z-\mu_i} - \frac{\omega}{g^2}\frac{1}{z-\mu_i} - \sum_j \frac{\hbar s_j}{\mu_i - \epsilon_j}\frac{1}{z-\mu_i} + \sum_{j\neq i} \frac{\hbar}{\mu_i - \mu_j}\frac{1}{z-\mu_i} = 0$$

We now sum over i. We have

$$\sum_{i=1}^{M} \frac{\mu_i}{z-\mu_i} = \sum_i \frac{\mu_i - z}{z-\mu_i} + \frac{z}{z-\mu_i} = -M + zS(z)$$

$$\sum_{i=1}^{M} \frac{1}{\mu_i - \epsilon_j}\frac{1}{z-\mu_i} = \frac{1}{z-\epsilon_j}\sum_i \left(\frac{1}{z-\mu_i} + \frac{1}{\mu_i - \epsilon_j}\right) = \frac{S(z) - S(\epsilon_j)}{z-\epsilon_j}$$

$$\begin{aligned}
\sum_{i=1}^{M}\sum_{j\neq i} \frac{1}{\mu_i - \mu_j}\frac{1}{z-\mu_i} &= \frac{1}{2}\sum_{i=1}^{M}\sum_{j\neq i} \frac{1}{\mu_i - \mu_j}\left(\frac{1}{z-\mu_i} - \frac{1}{z-\mu_j}\right) = \\
\frac{1}{2}\sum_{i=1}^{M}\sum_{j\neq i} \frac{1}{(z-\mu_i)(z-\mu_j)} &= \frac{1}{2}\left(\sum_{i,j} \frac{1}{(z-\mu_i)(z-\mu_j)} - \sum_i \frac{1}{(z-\mu_i)^2}\right) \\
&= \frac{1}{2}(S^2(z) + S'(z))
\end{aligned}$$

■

In equation (4.17) the $S(\epsilon_j)$ appear as parameters. They can be determined as follows. Suppose first that $s_j = 1/2$. We let $z \to \epsilon_i$ into eq.(4.17) getting

$$S'(\epsilon_i) + S^2(\epsilon_i) + \frac{2}{\hbar g^2}\Big((2\epsilon_i - \omega)S(\epsilon_i) - 2M\Big) = S'(\epsilon_i) + \sum_{j\neq i} \frac{S(\epsilon_i) - S(\epsilon_j)}{\epsilon_i - \epsilon_j}$$

The remarkable thing is that $S'(\epsilon_i)$ cancel in this equation and we get a set of closed algebraic equations determining the $S(\epsilon_j)$.

Proposition 25 *Let $s_j = 1/2$. In that case the constants $S(\epsilon_j)$ are determined by the set of closed algebraic equations:*

$$\boxed{S^2(\epsilon_i) + \frac{2}{\hbar g^2}\Big((2\epsilon_i - \omega)S(\epsilon_i) - 2M\Big) = \sum_{j\neq i} \frac{S(\epsilon_i) - S(\epsilon_j)}{\epsilon_i - \epsilon_j}, \quad i = 1, \cdots, n}$$

Suppose next that $s = 1$. We expand the Riccati equation around $z = \epsilon_i$:

$$(z-\epsilon_i)^0 \quad : \quad S'(\epsilon_i) + S^2(\epsilon_i) + \frac{2}{\hbar g^2}((2\epsilon_i - \omega)S(\epsilon_i) - 2M) = 2S'(\epsilon_i) + 2\sum_{j\neq i}\frac{S(\epsilon_i) - S(\epsilon_j)}{\epsilon_i - \epsilon_j}$$

$$(z-\epsilon_i)^1 \quad : \quad S''(\epsilon_i) + 2S(\epsilon_i)S'(\epsilon_i) + \frac{2}{\hbar g^2}((2\epsilon_i - \omega)S'(\epsilon_i) + 2S(\epsilon_i))$$
$$= S''(\epsilon_i) - 2\sum_{j\neq i}\frac{S(\epsilon_i) - S(\epsilon_j)}{(\epsilon_i - \epsilon_j)^2} - \frac{S'(\epsilon_i)}{\epsilon_i - \epsilon_j}$$

We see that in the second equation $S''(\epsilon_i)$ cancel. The first equation allows to compute $S'(\epsilon_i)$ and the second equation then gives a set of closed equations for the $S(\epsilon_i)$. We have shown the

Proposition 26 *Let $s_j = 1$. In that case the constants $S'(\epsilon_j)$ and $S(\epsilon_j)$ are determined by the of closed algebraic equations:*

$$\begin{aligned} S^2(\epsilon_i) + \frac{2}{\hbar g^2}((2\epsilon_i - \omega)S(\epsilon_i) - 2M) &= S'(\epsilon_i) + 2\sum_{j\neq i}\frac{S(\epsilon_i) - S(\epsilon_j)}{\epsilon_i - \epsilon_j} \\ 2S(\epsilon_i)S'(\epsilon_i) + \frac{2}{\hbar g^2}((2\epsilon_i - \omega)S'(\epsilon_i) + 2S(\epsilon_i)) &= -2\sum_{j\neq i}\frac{S(\epsilon_i) - S(\epsilon_j)}{(\epsilon_i - \epsilon_j)^2} - \frac{S'(\epsilon_i)}{\epsilon_i - \epsilon_j} \end{aligned}$$

The general mechanism is clear. For a spin s, we expand

$$S'(z) - 2s\frac{S(z) - S(\epsilon)}{z - \epsilon} = \sum_m \frac{m - 2s}{m!}S^{(m)}(\epsilon)(z-\epsilon)^{m-1}$$

and we see that the coefficient of $S^{(2s)}(\epsilon)$ vanishes in the term $m = 2s$. The equations coming from $(z-\epsilon)^{m-1}$ for $m = 1, \cdots, 2s-1$ allow to compute

$$S'(\epsilon), \cdots, S^{(2s-1)}(\epsilon)$$

by solving at each stage a linear equation. Plugging into the equation for $m = 2s$, we obtain a closed equation of degree $2s+1$ for $S(\epsilon)$.

$$P_{2s+1}(S(\epsilon)) = 0 \tag{4.18}$$

Notice that if $M < 2s$, the system will truncate at level M because there always exists a relation of the form $S^{(M)} = P(S, S', \cdots S^{(M-1)})$.

The $S(\epsilon_j)$ also determine the eigenvalues as well. Recall that

$$\Lambda(\lambda, \mu_1, \cdots, \mu_M) = a^2(\lambda) + \hbar a'(\lambda) + 2\hbar\sum_i \frac{a(\lambda) - a(\mu_i)}{\lambda - \mu_i}$$

so that

$$\begin{aligned}\frac{a(\lambda)-a(\mu_i)}{\lambda-\mu_i} &= \frac{2}{g^2}-\sum_j \frac{\hbar s_j}{\lambda-\mu_i}\left(\frac{1}{\lambda-\epsilon_j}-\frac{1}{\mu_i-\epsilon_j}\right)\\ &= \frac{2}{g^2}+\sum_j \frac{\hbar s_j}{\lambda-\epsilon_j}\frac{1}{\mu_i-\epsilon_j}\end{aligned}$$

hence

$$\sum_i \frac{a(\lambda)-a(\mu_i)}{\lambda-\mu_i} = \frac{2M}{g^2}-\sum_j \frac{\hbar s_j S(\epsilon_j)}{\lambda-\epsilon_j}$$

and

$$\boxed{\Lambda(\lambda)=a^2(\lambda)+\hbar a'(\lambda)+2\hbar\left(\frac{2M}{g^2}-\sum_j \frac{\hbar s_j S(\epsilon_j)}{\lambda-\epsilon_j}\right)} \tag{4.19}$$

Expanding $a(\lambda)$ we find

$$\Lambda(\lambda)=\frac{1}{g^4}(2\lambda-\omega)^2+\frac{4}{g^2}H_n+\frac{2}{g^2}\sum_j\frac{H_j}{\lambda-\epsilon_j}+\sum_j\frac{\hbar^2 s_j(s_j+1)}{(\lambda-\epsilon_j)^2}$$

with

$$H_n=\hbar M-\hbar\sum_j s_j+\frac{\hbar}{2}$$

and

$$H_j=\frac{2\omega}{g^2}\hbar s_j-\frac{4}{g^2}s_j\epsilon_j-2\hbar^2 s_j S(\epsilon_j)+2\sum_{i\neq j}\frac{\hbar^2 s_j s_i}{\epsilon_j-\epsilon_i}$$

4.5 Baxter Equation.

We can linearize this Riccati equation by setting

$$S(z)=\frac{\psi'(z)}{\psi(z)} \tag{4.20}$$

Obviously

$$\boxed{\psi(z)=\prod_{i=1}^{M}(z-\mu_i)} \tag{4.21}$$

The linearized equation reads

$$\boxed{\psi''(z)+\frac{2}{\hbar}a(z)\psi'(z)+\frac{2}{\hbar}\left(-\frac{2M}{g^2}+\sum_j\frac{\hbar s_j S(\epsilon_j)}{z-\epsilon_j}\right)\psi(z)=0} \tag{4.22}$$

Here, we should understand that $S(\epsilon_j)$ are determined by the procedure explained in the previous section. For such values of the parameters, the equation has the following remarkable property.

Proposition 27 *For the special values of the parameters $S(\epsilon_j)$ coming from the Bethe equations, the solutions of eq.(4.22) have trivial monodromy.*

Proof. This is clear for the solution $\psi_1(z)$ defined by eq.(4.21) since it is a polynomial. A second solution can be constructed as usual

$$\psi_2(z) = \psi_1(z) \int^z \exp\left(-\frac{2}{\hbar}\int^y a(t)dt - 2\log\psi_1(y)\right) dy = \psi_1(z) \int^z \frac{\prod_j (y-\epsilon_j)^{2s}}{\prod_i (y-\mu_i)^2} e^{-\frac{2}{\hbar g^2}(y^2-\omega y)} dy$$

The monodromy will be trivial if the pole at $y = \mu_i$ has no residue preventing the apparition of logarithms. Expanding around μ_i, we have

$$\begin{aligned}\exp\left(-\frac{2}{\hbar}\int^y a(t)dt - 2\log\psi_1(y)\right) &= \frac{e^{\left(-\frac{2}{\hbar}\int^{\mu_i} a(t)dt - 2\sum_{j\neq i}(\mu_i-\mu_j)\right)}}{(y-\mu_i)^2} \times \\ &\exp\left(-\frac{2}{\hbar}(y-\mu_i)\left[a(\mu_i) + \sum_{j\neq i}\frac{\hbar}{\mu_i-\mu_j}\right] + O(z-\mu_i)^2\right)\end{aligned}$$

but the coefficient of the dangerous $(y-\mu_i)$ term vanishes by virtue of the Bethe equations. ∎

Next we set

$$\boxed{\psi(z) = \exp\left(-\frac{1}{\hbar}\int^z a(y)dy\right) Q(z)} \tag{4.23}$$

We obtain for $Q(z)$ the equation

$$\boxed{\hbar^2 Q''(z) - \left(a^2(z) + \hbar a'(z) + \frac{4\hbar M}{g^2} - 2\hbar^2 \sum_j \frac{s_j S(\epsilon_j)}{z-\epsilon_j}\right) Q(z)}$$

Comparing with eq.(4.19), we obtain Baxter's equation

$$\boxed{\hbar^2 Q''(z) - \Lambda(z) Q(z)} \tag{4.24}$$

Notice that

$$\exp\left(-\frac{1}{\hbar}\int^z a(y)dy\right) = e^{-\frac{1}{\hbar g^2}(z^2-\omega z)} \prod_j (z-\epsilon_j)^{s_j}$$

so that

$$Q(z) = \frac{e^{\frac{1}{\hbar g^2}(z^2-\omega z)}}{\prod_j (z-\epsilon_j)^{s_j}}\psi(z) = \frac{e^{\frac{1}{\hbar g^2}(z^2-\omega z)}}{\prod_j (z-\epsilon_j)^{s_j}} \prod_{i=1}^M (z-\mu_i) \tag{4.25}$$

4.6 Bethe eigenvectors and separated variables.

We recall that

$$\Omega(\mu_1, \mu_2, \cdots, \mu_M) = C(\mu_1) \cdots C(\mu_M)|0\rangle$$

By definition of the separated variables, we have

$$\boxed{C(\lambda) = \frac{2z}{g} \frac{\prod_{i=1}^{n}(\lambda - \lambda_i)}{\prod_{j=1}^{n}(\lambda - \epsilon_j)}}$$

Inserting into the Bethe state and remembering that

$$\psi(z) = \prod_i (z - \mu_i)$$

we get

$$\boxed{\Omega(\mu_1, \mu_2, \cdots, \mu_M) = \left(\prod_j \frac{1}{\psi(\epsilon_j)} \right) \left(\frac{2z}{g} \right)^M \prod_i \psi(\lambda_i)|0\rangle}$$

Proposition 28 *In the separated variables, the Hamiltonians read*

$$H_i = \frac{\prod_k (\epsilon_i - \lambda_k)}{\prod_{k \neq i}(\epsilon_i - \epsilon_k)} \sum_j \frac{\prod_{k \neq i}(\lambda_j - \epsilon_k)}{\prod_{k \neq j}(\lambda_j - \lambda_k)} \left(\frac{d^2}{d\lambda_j^2} + \frac{2}{\hbar} a(\lambda_j) \frac{d}{d\lambda_j} - \frac{M}{\hbar} \right) \tag{4.26}$$

Proof. Write eq.(4.22) for each separated variable as

$$\sum_j \frac{1}{\lambda_i - \epsilon_j} H_j \psi(\lambda_i) = -\left(\frac{d^2}{d\lambda_i^2} + \frac{2}{\hbar} a(\lambda_i) \frac{d}{d\lambda_i} - \frac{M}{\hbar} \right) \psi(\lambda_i)$$

Since this formula holds for a basis of eigenvectors, we can "factor" by $\psi(\lambda_i)$. Inverting the Cauchy matrix $B_{ij} = 1/(\lambda_i - \epsilon_j)$, and taking care of the order of operators we obtain the Hamiltonians H_i in terms of the separated variables

$$\boxed{H_i = -B^{-1}_{ij} V_j, \quad B_{ij} = \frac{1}{\lambda_i - \epsilon_j}, \quad V_j = \frac{d^2}{d\lambda_j^2} + \frac{2}{\hbar} a(\lambda_j) \frac{d}{d\lambda_j} - \frac{M}{\hbar}}$$

explicitly, they are just eqs.(4.26). ∎

Proposition 29 *The Hamiltonians eq.(4.26) commute.*

This is a general result and we will prove it in section 4.8.

We now introduce the scalar product

$$||\Omega||^2 = \int d\lambda_i d\bar{\lambda}_i \; W\bar{W} \; \rho(x_1, x_2, \cdots, x_n)|\psi(\lambda_i)|^2 \tag{4.27}$$

where

$$W = \prod_{i \neq j} (\lambda_i - \lambda_j)$$

and

$$x_i = \frac{1}{\hbar} \frac{\prod_j |\epsilon_i - \lambda_j|^2}{\prod_{k \neq i} (\epsilon_i - \epsilon_k)^2}$$

The rationale for this is that we want to change variables from the spin variables $s_j^+ = \hbar w_j$ to the separated variables λ_j, that is

$$w_j = \frac{2z}{\hbar g} \frac{\prod_i (\epsilon_j - \lambda_i)}{\prod_{k \neq j} (\epsilon_j - \epsilon_k)}, \quad w_j \bar{w}_j = \frac{z\bar{z}}{\hbar} x_j$$

The factor $W\bar{W}$ just comes from the Jacobian of this transformation. In the reduced model however, the variables $z, \bar{z}$ are integrated out. The measure $\rho(x_1, x_2, \cdots, x_n)$ is determined by requiring that the Hamiltonian H_j are Hermitian.

Proposition 30 *The Hamiltonians H_j are Hermitian with respect to the scalar product eq.(4.27) with*

$$\rho(x_1, x_2, \cdots, x_n) = \int_0^\infty dy e^{-y} y^{M+n-\sum_i (s_i+1/2)} \prod_i \frac{J_{2s_i+1}(2\sqrt{yx_i})}{x_i^{s_i+1/2}} \tag{4.28}$$

where $J_{2s_i+1}(x)$ is the Bessel function. For $n = 1$, the formula for $\rho(x)$ can be simplified giving

$$\rho(x) = \partial_x^{M+1} \left[e^{-x} x^{M-2s} \right] = e^{-x} P_{M-2s}(x)$$

where $P_{M-2s}(x)$ is a Laguerre polynomial of degree $M - 2s$.

Proof. We have to show that

$$\int d\lambda_k d\bar{\lambda}_k |\psi(\lambda_k)|^2 \sum_j \left(-\frac{d^2}{d\lambda_j^2} + \frac{2}{\hbar} \frac{d}{d\lambda_j} \cdot a(\lambda_j) + \frac{M}{\hbar} \right) B_{ij}^{-1} |W|^2 \rho(x_1, \cdots, x_n) \tag{4.29}$$

is real. Now

$$B_{ij}^{-1} = \Delta^{-1} \Delta_{ji}$$

where Δ_{ji} is the minor of the element B_{ji}. It is clearly independent of λ_j. Hence

$$\frac{d}{d\lambda_j} B_{ij}^{-1} = B_{ij}^{-1} \left(\frac{d}{d\lambda_j} - \Delta^{-1} \frac{d}{d\lambda_j} \Delta \right)$$

We have

$$\Delta = \frac{\prod_{i\neq j}(\lambda_i - \lambda_j)\prod_{i\neq j}(\epsilon_i - \epsilon_j)}{\prod_{i,j}(\lambda_i - \epsilon_j)} = \prod_{j\neq i}(\epsilon_i - \epsilon_j)^{-1} W \prod_{i=1}^{n} z_i^{-1}$$

where we introduced

$$z_i = \frac{\prod_j(\epsilon_i - \lambda_j)}{\prod_{j\neq i}(\epsilon_i - \epsilon_j)}, \quad x_i = \frac{z_i \bar{z}_i}{\hbar}$$

Remark the important formula

$$\frac{d}{d\lambda_j} z_k = B_{jk} z_k$$

hence

$$\Delta^{-1}\frac{d}{d\lambda_j}\Delta = W^{-1}\frac{d}{d\lambda_j}W - \sum_k B_{jk}$$

so that

$$\frac{d}{d\lambda_j} B_{ij}^{-1}|W|^2 = B_{ij}^{-1}|W|^2 \left(\frac{d}{d\lambda_j} + \sum_k B_{jk}\right)$$

$$\frac{d^2}{d\lambda_j^2} B_{ij}^{-1}|W|^2 = B_{ij}^{-1}|W|^2 \left(\frac{d^2}{d\lambda_j^2} + 2\sum_k B_{jk}\frac{d}{d\lambda_j} + 2\sum_k B_{jk}\sum_{l\neq k}\frac{1}{\epsilon_k - \epsilon_l}\right)$$

Next, we have

$$\frac{d}{d\lambda_j}\rho(x_1, \cdots, x_n) = \sum_k B_{jk} x_k \frac{\partial}{\partial x_k}\rho(x_1, \cdots, x_n)$$

$$\begin{aligned}
\frac{d^2}{d\lambda_j^2}\rho(x_1, \cdots, x_n) &= \sum_{k,l} B_{jk} B_{jl} x_k x_l \frac{\partial^2}{\partial x_k \partial x_l}\rho(x_1, \cdots, x_n) \\
&= \sum_k B_{jk}^2 x_k^2 \frac{\partial^2}{\partial x_k^2}\rho(x_1, \cdots, x_n) + 2\sum_{k,l} B_{jk}\frac{1}{\epsilon_k - \epsilon_l} x_k x_l \frac{\partial^2}{\partial x_k \partial x_l}\rho(x_1, \cdots, x_n)
\end{aligned}$$

Putting everything together eq.(4.29) becomes

$$\int d\lambda_k d\bar{\lambda}_k |\psi(\lambda_k)|^2 \sum_j B_{ij}^{-1}|W|^2 \left\{ -\sum_k B_{jk}^2 x_k \mathcal{D}_k + \sum_k B_{jk}\mathcal{O}_k + \frac{1}{\hbar}\mathcal{D}_0 \right\} \rho(x_1, \cdots, x_n) \tag{4.30}$$

where

$$\mathcal{D}_k = x_k \partial_{x_k}^2 + 2(s_k + 1)\partial_{x_k}$$

$$\mathcal{D}_0 = \left(\sum_k x_k \partial_{x_k} + M + n + 1\right)$$

and

$$\mathcal{O}_k = \frac{1}{\hbar}\left(\epsilon_k - \frac{\omega}{2}\right) - 2\sum_{l\neq k}\frac{1}{\epsilon_k-\epsilon_l}\Big((x_k\partial_{x_k}+s_k+1)(x_l\partial_{x_l}+s_l+1) - s_k s_l - 1\Big)$$

The conditions on $\rho(x_1,\cdots,x_n)$ are that the quantity in the curly bracket in eq.(4.30) should be equal to its complex conjugate. When we do the sum over j, we have

$$\sum_{j,k} B^{-1}_{ij}B_{jk}\mathcal{O}_k = \mathcal{O}_i$$

which is real and gives no condition. Next we have the identities

$$\begin{aligned}
\sum_j B^{-1}_{ij} &= -z_i \\
\sum_j B^{-1}_{ij}B^2_{jk} &= -\frac{1}{\epsilon_i-\epsilon_k}\frac{z_i}{z_k}, \quad i\neq k \\
\sum_j B^{-1}_{ij}B^2_{ji} &= -\frac{1}{z_i} - \sum_{k\neq i}\frac{1}{\epsilon_i-\epsilon_k}\frac{z_k}{z_i}
\end{aligned}$$

The conditions on $\rho(x_1,\cdots,x_n)$ then read

$$-(z_i-\bar{z}_i)\left[\mathcal{D}_i\rho + \mathcal{D}_0\rho\right] + \sum_{k\neq i}\frac{z_k\bar{z}_i - \bar{z}_k z_i}{\epsilon_i-\epsilon_k}\left[\mathcal{D}_i\rho - \mathcal{D}_k\rho\right] = 0$$

Finally find the n conditions

$$\mathcal{D}_i\rho - \mathcal{D}_k\rho = 0, \quad k\neq i \tag{4.31}$$

$$\mathcal{D}_i\rho + \mathcal{D}_0\rho = 0 \tag{4.32}$$

Notice that eq.(4.32) is independent of i if the conditions eq.(4.31) are satisfied. A solution of eq.(4.31) is

$$\rho(x_1,\cdots,x_n) = \sum_{p=0}^{\infty} C_p \sum_{q_1+\cdots+q_n=p}\prod_{i=1}^{n}\frac{x_i^{q_i}}{q_i!(2s_i+1+q_i)!}$$

Then eq.(4.32) gives

$$C_{p+1} + (M+n+p+1)C_p = 0$$

the solution of which is

$$C_p = (-1)^p\begin{pmatrix} M+n+p \\ p\end{pmatrix}p!$$

Hence we have found

$$\rho(x_1,x_2,\cdots,x_n) = \sum_{p=0}^{\infty}(-1)^p\begin{pmatrix} M+n+p \\ p\end{pmatrix}p!\sum_{q_1+\cdots+q_n=p}\prod_{i=1}^{n}\frac{x_i^{q_i}}{q_i!(2s_i+1+q_i)!} \tag{4.33}$$

This is just eq.(4.28) ∎

This important formula should be further studied. In particular, for $\psi(z)$ being a Bethe state, one should be able to compute it exactly because we know that by Gaudin formula

$$||\Omega(\mu_1, \mu_2, \cdots, \mu_M)||^2 = M! \det \Delta$$

where Δ is the Jacobian matrix of Bethe's equations. This is still very mysterious.

4.7 Quasi-Classical limit.

The exact formula relating $Q(z)$ and $\psi(z)$, eq.(4.25), allows to study the properties of the solutions of Bethe roots μ_i in the quasi-classical limit $\hbar \to 0$.

Let us set

$$y(z) = \hbar \frac{Q'(z)}{Q(z)}$$

Then Baxter's equation, eq.(4.36), becomes

$$\hbar\, y'(z) + y^2(z) = \Lambda(z) \tag{4.34}$$

where we recall that

$$\Lambda(z) = a^2(z) + \hbar a'(z) + 2\hbar \left(\frac{2M}{g^2} - \sum_j \frac{\hbar s_j S(\epsilon_j)}{z - \epsilon_j} \right), \quad a(z) = \frac{2z}{g^2} - \frac{\omega}{g^2} - \sum_j \frac{\hbar s_j}{z - \epsilon_j}$$

In the semi-classical limit eq.(4.34) becomes the equation of the spectral curve of the model (in that limit $\hbar s_j = O(\hbar^0)$):

$$y^2(z) = \Lambda(z)$$

This also means that we can write in the semi-classical limit, as expected,

$$Q(z) = e^{\frac{1}{\hbar}\int^z y(\lambda) d\lambda} \simeq e^{\frac{1}{\hbar}\int^z \sqrt{\Lambda(\lambda)} d\lambda + O(\hbar^0)}$$

From eq.(4.23) we deduce that

$$y(z) = a(z) + \sum_i \frac{\hbar}{z - \mu_i}$$

so that we expect in the semi-classical limit

$$\sum_i \frac{\hbar}{z - \mu_i} \simeq \sqrt{\Lambda(z)} - a(z)$$

This is a remarkable formula. It gives us the distribution of Bethe roots μ_i in the semi-classical limit, as we now show.

Let $\sqrt{\Lambda(z)}$ be represented as a meromorphic function in the cut z-plane. Let us put the cuts so that

$$\sqrt{\Lambda(z)} = \frac{2z}{g^2} - \frac{\omega}{g^2} + O(z^{-1}), \quad |z| \to \infty$$

and (we neglect terms of order $\hbar$ which do not contribute in the leading $\hbar$ approximation).

$$\sqrt{\Lambda(z)} = -\frac{\hbar s_j}{z - \epsilon_j} + O(1), \quad z \to \epsilon_j$$

By the Cauchy theorem, we have

$$\sqrt{\Lambda(z)} = \int_{\mathcal{C}} \frac{dz'}{2i\pi} \frac{\sqrt{\Lambda(z')}}{z' - z}$$

where $\mathcal{C}$ is a big circle C_0 at infinity, minus small circles C_j around $z = \epsilon_j$, minus contours A_i around the cuts of $\sqrt{\Lambda(z)}$. Hence

$$\sqrt{\Lambda(z)} = \int_{C_0} \frac{dz'}{2i\pi} \frac{\sqrt{\Lambda(z')}}{z' - z} - \sum_j \int_{C_j} \frac{dz'}{2i\pi} \frac{\sqrt{\Lambda(z')}}{z' - z} - \sum_i \int_{A_i} \frac{dz'}{2i\pi} \frac{\sqrt{\Lambda(z')}}{z' - z}$$

But

$$\int_{C_0} \frac{dz'}{2i\pi} \frac{\sqrt{\Lambda(z')}}{z' - z} = (\sqrt{\Lambda(z)})_+ = \frac{2z}{g^2} - \frac{\omega}{g^2}$$

and

$$\int_{C_j} \frac{dz'}{2i\pi} \frac{\sqrt{\Lambda(z')}}{z' - z} = \int_{C_j} \frac{dz'}{2i\pi} \frac{-\hbar s_j}{(z' - \epsilon_j)(z' - z)} = -\frac{\hbar s_j}{z - \epsilon_j}$$

so that we arrive at

$$\sqrt{\Lambda(z)} = \frac{2z}{g^2} - \frac{\omega}{g^2} - \sum_j \frac{\hbar s_j}{z - \epsilon_j} - \sum_i \int_{A_i} \frac{dz'}{2i\pi} \frac{\sqrt{\Lambda(z')}}{z' - z}$$

hence

$$\sqrt{\Lambda(z)} - a(z) = -\sum_i \int_{A_i} \frac{dz'}{2i\pi} \frac{\sqrt{\Lambda(z')}}{z' - z}$$

and therefore

$$\sum_i \frac{\hbar}{z - \mu_i} = \sum_i \int_{A_i} \frac{dz'}{2i\pi} \frac{\sqrt{\Lambda(z')}}{z - z'} + O(\hbar)$$

Comparing both members of this formula suggests that the Bethe roots μ_i accumulate in the semiclassical limit on curves A_i along which the singularities of both side should match. To determine these curves we assume that the Bethe roots μ_i tend to a continuous function $\mu(t)$ when $\hbar \to 0$ ($t = \hbar\, i$ and $i = O(\hbar^{-1})$).

$$\sum_i \frac{\hbar}{z - \mu_i} = \sum_i \frac{\hbar}{z - \mu(i)} = \int \frac{dt}{z - \mu(t)} = \int_{\mathcal{A}} d\mu \left(\frac{dt}{d\mu}\right) \frac{1}{z - \mu}$$

Here $\mathcal{A} = \sum A_i$. Hence, comparing with the semi-classical result, we conclude that the function $\mu(t)$ should satisfy the differential equation (we rename the function $\mu(t)$ to $z(t)$.)

$$\frac{dz}{dt} = \frac{2i\pi}{\sqrt{\Lambda(z)}} \tag{4.35}$$

The boundary condition is that the integral curve $z(t)$ should start (and end !) at a branch point of the spectral curve $y^2 = \Lambda(z)$.

This result can be checked by numerical calculation. We consider the one spin-s system. A typical situation is shown in Fig.(4.3). The agreement is spectacular.

We can say a word on how the Bethe equations were solved. We first determine $S(\epsilon)$ by solving the polynomial equation eq.(4.18) and then determine $\psi(z)$, eq.(4.21), by solving eq.(4.22). The Bethe roots are then obtained by solving the polynomial equation $\psi(z) = 0$.

4.8 Riemann surfaces and quantum integrability.

Let us consider a set of separated variables

$$[\lambda_i, \lambda_j] = 0, \quad [\mu_i, \mu_j] = 0, \quad [\lambda_i, \mu_j] = p(\lambda_i, \mu_i)\delta_{ij}$$

We want Baxter's equation, so we start from the linear system

$$\sum_j R_j(\lambda_i, \mu_i)H_j + R_0(\lambda_i, \mu_i) = 0 \tag{4.36}$$

Here the H_j are on the right, and in $R_j(\lambda_i, \mu_i)$, $R_0(\lambda_i, \mu_i)$, we assume some order between λ_i, μ_i, but the coefficients in these functions are non dynamical. Hence we start from the linear system

$$BH = -V \tag{4.37}$$

We notice that we can define unambiguously the left inverse of B. First, the determinant D of B is well defined because it never involves a product of elements on the same line. The same is true for the cofactor Δ_{ij} of the element B_{ij} (we include the sign $(-1)^{i+j}$ in the definition of Δ_{ij}). Define

$$B^{-1}_{ij} \equiv (B^{-1})_{ij} = D^{-1}\Delta_{ji}$$

We have

$$(B^{-1}B)_{ij} = \sum_k D^{-1}\Delta_{ki}B_{kj}$$

But Δ_{ki} does *not* contain any element B_{kl}, hence the product $\Delta_{ki}B_{kj}$ is commutative, and the usual construction of the inverse of B is still valid. If right inverse of B exists, and it exists at least classically, it coincides with the left inverse in an associative algebra with unit. So we have the identities

$$(BB^{-1})_{ij} = \sum_k B_{ik}B^{-1}_{kj} = \sum_k B_{ik}D^{-1}\Delta_{jk} = \delta_{ij} \tag{4.38}$$

We write the solution of eq.(4.37) as

$$H = -B^{-1}V \tag{4.39}$$

Theorem 12 *The quantities H_i defined by eq.(4.39), which solve Baxter's equations eqs.(4.36), are all commuting*

$$[H_i, H_j] = 0$$

Proof. Using that V_k and V_l commute, $[V_k, V_l] = 0$, we compute

$$\begin{aligned} [H_i, H_j] &= \sum_{k,l} [B_{ik}^{-1}V_k, B_{jl}^{-1}V_l] \\ &= \sum_{k,l} [B_{ik}^{-1}, B_{jl}^{-1}]V_kV_l - B_{ik}^{-1}[B_{jl}^{-1}, V_k]V_l + B_{jl}^{-1}[B_{ik}^{-1}, V_l]V_k \end{aligned} \tag{4.40}$$

Using

$$[A^{-1}, B^{-1}] = A^{-1}B^{-1}[A, B]B^{-1}A^{-1} = B^{-1}A^{-1}[A, B]A^{-1}B^{-1}$$

so that

$$\begin{aligned} [B_{ik}^{-1}, B_{jl}^{-1}] &= \sum_{rs,r's'} B_{ir}^{-1}B_{jr'}^{-1}[B_{rs}, B_{r's'}]B_{s'l}^{-1}B_{sk}^{-1} \\ &= \sum_{rs,r's'} B_{jr'}^{-1}B_{ir}^{-1}[B_{rs}, B_{r's'}]B_{sk}^{-1}B_{s'l}^{-1} \end{aligned}$$

the first term can be written

$$\begin{aligned} \sum_{k,l} [B_{ik}^{-1}, B_{jl}^{-1}]V_kV_l &= \sum \frac{1}{2}B_{ir}^{-1}B_{jr'}^{-1}[B_{rs}, B_{r's'}]\Big(B_{s'l}^{-1}B_{sk}^{-1} + B_{s'k}^{-1}B_{sl}^{-1}\Big)V_kV_l \\ &= \sum \frac{1}{2}B_{jr'}^{-1}B_{ir}^{-1}[B_{rs}, B_{r's'}]\Big(B_{sk}^{-1}B_{s'l}^{-1} + B_{sl}^{-1}B_{s'k}^{-1}\Big)V_kV_l \end{aligned}$$

Using that $[B_{rs}, B_{r's'}] = \delta_{rr'}[B_{rs}, B_{rs'}]$ and is therefore antisymmetric in ss', and setting

$$K_{ss'} = \sum_{k,l} \Big(B_{s'l}^{-1}B_{sk}^{-1} + B_{s'k}^{-1}B_{sl}^{-1} - B_{sl}^{-1}B_{s'k}^{-1} - B_{sk}^{-1}B_{s'l}^{-1}\Big)V_kV_l$$

we get

$$\begin{aligned} \sum_{k,l} [B_{ik}^{-1}, B_{jl}^{-1}]V_kV_l &= \sum_{rss'} \frac{1}{4}B_{ir}^{-1}B_{jr}^{-1}[B_{rs}, B_{rs'}]K_{ss'} \\ &= -\sum_{rss'} \frac{1}{4}B_{jr}^{-1}B_{ir}^{-1}[B_{rs}, B_{rs'}]K_{ss'} \\ &= \sum_{rss'} \frac{1}{8}[B_{ir}^{-1}, B_{jr}^{-1}][B_{rs}, B_{rs'}]K_{ss'} \end{aligned}$$

The last two terms in eq.(4.40) are simpler, we get

$$\sum_{k,l} B^{-1}_{jl}[B^{-1}_{ik}, V_l]V_k - B^{-1}_{ik}[B^{-1}_{jl}, V_k]V_l = \sum_{rsk}[B^{-1}_{ir}, B^{-1}_{jr}][B_{rs}, V_r]B^{-1}_{sk}V_k$$

The quantities H_i will commute if

$$[B^{-1}_{ir}, B^{-1}_{jr}] = 0, \quad \forall i, j, r \tag{4.41}$$

This is true as shown in the next Lemma. ■

The condition eq.(4.41) says that the elements on the same column of B^{-1} commute among themselves. In a sense this is a condition dual to the one on B. It is true semiclassically because

$$\{B^{-1}_{ir}, B^{-1}_{jr}\} = \sum_{a,a',b,b'} B^{-1}_{ia}B^{-1}_{ja'}\{B_{ab}, B_{a'b'}\}B^{-1}_{br}B^{-1}_{b'r} = \sum_{a,b,b'} B^{-1}_{ia}B^{-1}_{ja}\{B_{ab}, B_{ab'}\}B^{-1}_{br}B^{-1}_{b'r} = 0$$

where in the last step we use the antisymmetry of the Poisson bracket. We show that it is also true quantum mechanically

Lemma 1 *Let B be a matrix whose elements commute if they do not belong to the same line*

$$[B_{ik}, B_{jl}] = 0 \quad \text{if } i \neq j$$

Then the left inverse B^{-1} of B is defined without ambiguity and moreover elements on a same column of B^{-1} commute

$$[B^{-1}_{ir}, B^{-1}_{jr}] = 0$$

Proof. We want to show that

$$\Delta_{ri}B^{-1}_{jr} = \Delta_{rj}B^{-1}_{ir}$$

denote by $\beta^{(r)}_i$ the vector with components B_{ki}, $k \neq r$. Then we have (with $j > i$)

$$\begin{aligned}
\Delta_{ri}B^{-1}_{jr} &= (-1)^{r+i}\beta^{(r)}_1 \wedge \beta^{(r)}_2 \wedge \cdots \widehat{\beta^{(r)}_i} \wedge \cdots \beta^{(r)}_j \wedge \cdots \beta^{(r)}_g B^{-1}_{jr} \\
&= (-1)^{r+i+g-j}\beta^{(r)}_1 \wedge \beta^{(r)}_2 \wedge \cdots \widehat{\beta^{(r)}_i} \wedge \cdots \widehat{\beta^{(r)}_j} \wedge \cdots \beta^{(r)}_g \wedge \beta^{(r)}_j B^{-1}_{jr} \\
&= (-1)^{r+i+g-j+1}\beta^{(r)}_1 \wedge \beta^{(r)}_2 \wedge \cdots \widehat{\beta^{(r)}_i} \wedge \cdots \widehat{\beta^{(r)}_j} \wedge \cdots \beta^{(r)}_g \wedge \sum_{k\neq j} \beta^{(r)}_k B^{-1}_{kr} \\
&= (-1)^{r+i+g-j+1}\beta^{(r)}_1 \wedge \beta^{(r)}_2 \wedge \cdots \widehat{\beta^{(r)}_i} \wedge \cdots \widehat{\beta^{(r)}_j} \wedge \cdots \beta^{(r)}_g \wedge \beta^{(r)}_i B^{-1}_{ir} \\
&= (-1)^{r+j}\beta^{(r)}_1 \wedge \beta^{(r)}_2 \wedge \cdots \beta^{(r)}_i \wedge \cdots \widehat{\beta^{(r)}_j} \wedge \cdots \beta^{(r)}_g B^{-1}_{ir} \\
&= \Delta_{rj}B^{-1}_{ir}
\end{aligned}$$

where in the third line we used eq.(4.38). In the above manipulations, we never have two operators B_{ij} on the same line so we can use the usual properties of the wedge product. Moreover it is important that the line r is absent in the definition of $\beta^{(r)}$. Remark that this equation can also be written $\Delta_{ri} D^{-1} \Delta_{rj} = \Delta_{rj} D^{-1} \Delta_{ri}$ which is a Yang–Baxter type equation. ∎

With this Lemma, we have completed the proof of our theorem. It is remarkable that, again, only the separated nature of the variables λ_i, μ_i is used in this construction, but the precise commutation relations between λ_i, μ_i does not even need to be specified. This is the origin of the multi Hamiltonian structure of integrable systems, here extended to the quantum domain.

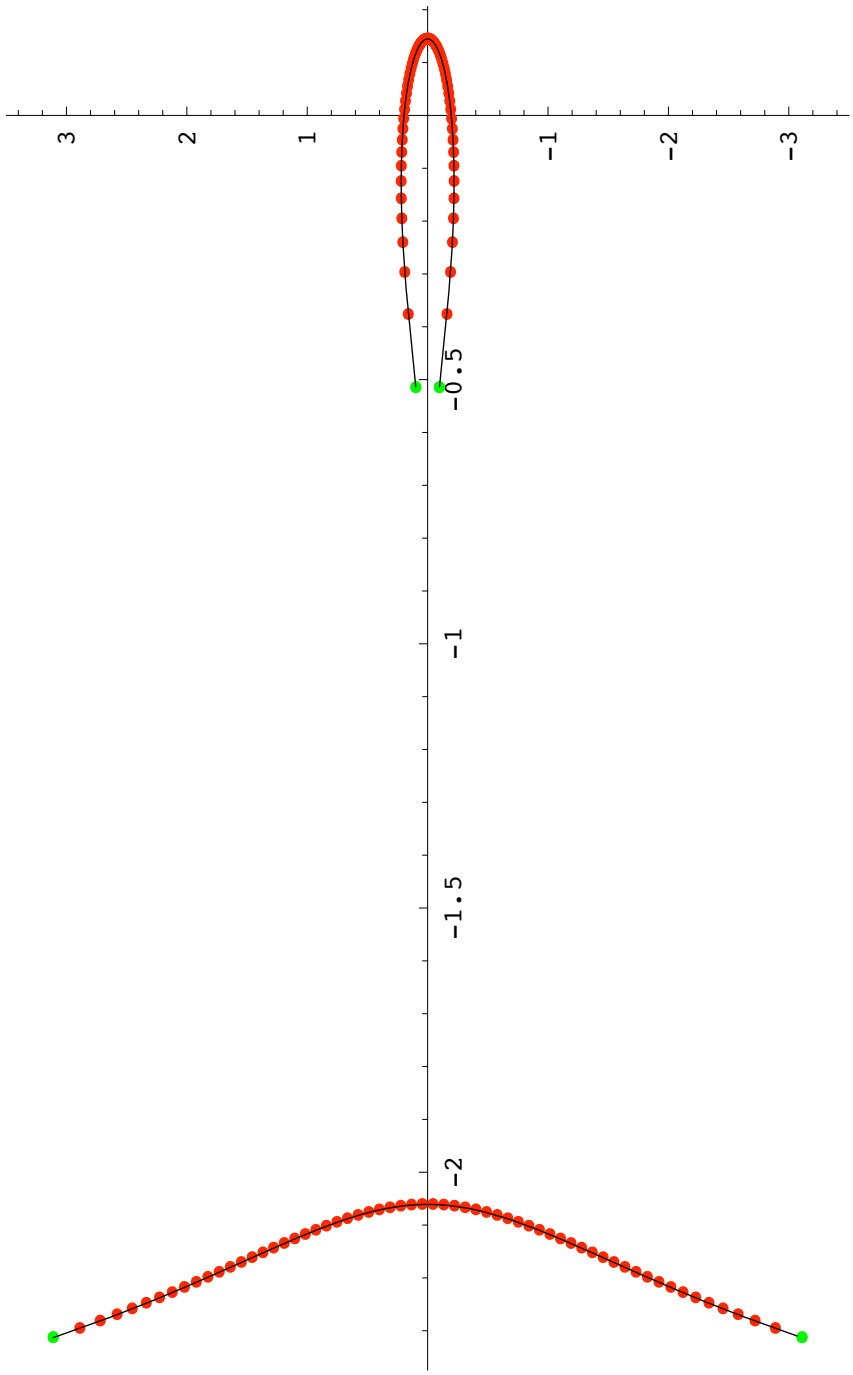


Figure 4.3: Red dots are the Bethe roots μ_i for the one spin system. Green dots are the branch points. The thin black curve is the solution of eq.(4.35). ($\hbar = 1/30$, $s = 1/\hbar$, $M = 4/\hbar$, highest energy state).

Bibliography

[1] E. Jaynes, F. Cummings, Proc. IEEE vol. 51 (1963) p. 89.

[2] J. Ackerhalt, K. Rzazewski, *Heisenberg-picture operator perturbation theory.* Phys. Rev. A, vol. 12, (1975) 2549.

[3] N. Narozhny, J. Sanchez-Mondragon, J. Eberly, *Coherence versus incoherence: Collapse and revival in a simple quantum model.* Phys. Rev. A Vol. 23, (1981) p. 236.

[4] E. Yuzbashyan, V. Kuznetsov, B. Altshuler, *Integrable dynamics of coupled Fermi-Bose condensates.* cond-mat/0506782

[5] E. Sklyanin, *Separation of variables in the Gaudin model.* J. Soviet Math., Vol. 47, (1979) pp. 2473-2488.

[6] M. Gaudin, **La Fonction d'Onde de Bethe.** Masson, (1983).

[7] F. V. Atkinson, *Multiparameter spectral theory.* Bull. Amer. Math. Soc. 74:1-27 (1968). **Multiparameter Eigenvalue Problems.** Academic, 1972.

[8] B. Enriquez, V. Rubtsov, *Commuting families in skew fields and quantization of Beauville's fibration.* math.AG/0112276.

[9] O. Babelon, M. Talon, *Riemann surfaces, separation of variables and classical and quantum integrability.* hep-th/0209071. Phys. Lett. A. 312 (2003) 71-77.

Chapter 5

The Heisenberg spin chain.

5.1 The quantum monodromy matrix

We have seen that in the case of two dimensional integrable field theories, the analog of the Lax matrix is the monodromy matrix

$$T_L(\lambda) = P \exp\left[\int_0^L U dx\right] \tag{5.1}$$

In order to define this quantity at the quantum level, it is convenient to start from the discretized version of the theory. Thus, we consider a lattice with N sites and lattice spacing $\Delta = L/N$. To each site n of the lattice, we attach the local transport matrix $L_n(\lambda)$ (we retain traditional notations). The matrix elements of $L_n(\lambda)$ are functions of the local quantum fields of the model. Over each site of the lattice, there is a local Hilbert space $\mathcal{H}_n$ on which the field operators act non trivially. The total Hilbert space of the discretized system is the tensor product of all these local Hilbert spaces.

$$\mathcal{H} = \bigotimes_{n=1}^{N} \mathcal{H}_n$$

We now define the quantum monodromy matrix $T_N(\lambda)$ by the discretized version of eq.(5.1)

$$\boxed{T_N(\lambda) = \prod_{n=1}^{N} L_n(\lambda).} \tag{5.2}$$

This is our basic object of study in this Chapter.

5.2 The XXX spin chain.

We now provide the canonical example of this construction: the XXX spin chain. Consider a lattice with N sites. On each lattice site n, we attach a Hilbert space $\mathcal{H}_n = \mathbb{C}^2$. The total Hilbert space is of dimension 2^N.

$$\mathcal{H} = \mathbb{C}^2_1 \otimes \mathbb{C}^2_2 \otimes \cdots \mathbb{C}^2_N$$

On each site we introduce the spin operators s^i_n acting non trivially on $\mathcal{H}_n$ only where they are represented by Pauli matrices

$$\vec{s}_n = 1 \otimes 1 \otimes \cdots 1 \otimes \frac{\hbar}{2}\,\vec{\sigma} \otimes 1 \cdots \otimes 1$$

where the non trivial innsertion is on the n-th position. We recall that

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

We will also use $s^\pm_n = s^1_n \pm i s^2_n$. We define

$$P_{12} = \frac{1}{2}(\mathrm{Id} \otimes \mathrm{Id} + \vec{\sigma} \otimes \vec{\sigma}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

On each lattice site we introduce the local L_n-operator

$$\boxed{L_n(\lambda) = \begin{pmatrix} \lambda + i s^3_n & i s^-_n \\ i s^+_n & \lambda - i s^3_n \end{pmatrix} = \lambda\,\mathrm{Id} + i\,\vec{\sigma}\,\vec{s}_n} \tag{5.3}$$

Define now the following matrices of operators,

$$L_{1n}(\lambda) = L_n(\lambda) \otimes 1, \quad L_{2n}(\lambda) = 1 \otimes L_n(\lambda)$$

where the tensor product now refers to the auxiliary matrix space. It is a simple exercise (multiplication of 4×4 matrices) to check that it satisfies the relation

$$\boxed{R_{12}(\lambda - \mu) L_{1n}(\lambda) L_{2n}(\mu) = L_{2n}(\mu) L_{1n}(\lambda) R_{12}(\lambda - \mu)} \tag{5.4}$$

with

$$\boxed{R_{12}(\lambda) = \frac{1}{\lambda + i\hbar}(\lambda\,\mathrm{Id} + i\hbar P_{12})}$$

or explicitly

$$R_{12}(\lambda) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & c(\lambda) & b(\lambda) & 0 \\ 0 & b(\lambda) & c(\lambda) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

with

$$b(\lambda) = \frac{i\hbar}{\lambda + i\hbar}, \quad c(\lambda) = \frac{\lambda}{\lambda + i\hbar}$$

Note that the so called ultralocality condition holds

$$\boxed{L_{1n}(\lambda)L_{2m}(\mu) = L_{2m}(\mu)L_{1n}(\lambda) \quad n \neq m}$$

We now construct the monodromy matrix $T_N(\lambda)$ for a lattice of N sites

$$T_N(\lambda) = \prod_{n=1}^{N} L_n(\lambda)$$

As in the classical case, one can go from the local formula eq.(5.4) to a global one for the quantum monodromy matrix.

Proposition 31 *If $L_n(\lambda)$ is such that (5.4) and the ultralocality condition holds, then*

$$\boxed{R_{12}(\lambda - \mu)T_1(\lambda)T_2(\mu) = T_2(\mu)T_1(\lambda)R_{12}(\lambda - \mu)} \tag{5.5}$$

Proof. We use ultralocality to write

$$T_1(\lambda)T_2(\mu) = \prod_{n=1}^{N} [L_{1n}(\lambda)L_{2n}(\mu)]$$

Then, using (5.4), we find

$$\begin{aligned} R_{12}(\lambda - \mu)T_1(\lambda)T_2(\mu) &= \prod_{n=1}^{N} [L_{2n}(\mu)L_{1n}(\lambda)]R_{12}(\lambda - \mu) \\ &= T_2(\mu)T_1(\lambda)R_{12}(\lambda - \mu) \end{aligned}$$

In the last step we have used again the ultralocality property. ∎

Remark. This is a quantum analog of the classical formula

$$\{T_1(\lambda), T_2(\mu)\} = [r_{12}(\lambda, \mu), T_1(\lambda)T_2(\mu)]$$

to which it reduces in the semi classical limit

$$R_{12}(\lambda,\mu) \to 1 + \hbar\, r_{12}(\lambda,\mu) + O(\hbar^2) \tag{5.6}$$

Remark. Since $R_{12}(\lambda-\mu)$ depends only on one variable, $\lambda-\mu$, the same argument works for inhomogenous models. defined by

$$T(\lambda;\alpha_1,\cdots,\alpha_N) = L_1(\lambda-\alpha_1)L_2(\lambda-\alpha_2)\cdots L_N(\lambda-\alpha_N)$$

From these commutation relations we can extract a family of commuting Hamiltonians. Define

$$\boxed{t(\lambda) = \mathrm{Tr}\; T(\lambda)}$$

then

$$\boxed{[t(\lambda), t(\mu)] = 0}$$

This is because

$$T_1(\lambda)T_2(\mu) = R_{12}^{-1}(\lambda-\mu)T_2(\mu)T_1(\lambda)R_{12}(\lambda-\mu)$$

then take the trace and the cyclicity property to bring $R^{-1}(\lambda-\mu)$ to the right where it cancels the $R(\lambda-\mu)$ factor. Then use the fact that

$$\mathrm{Tr}_{12}T_1(\lambda)T_2(\mu) = t(\lambda)t(\mu), \quad \mathrm{Tr}_{12}T_2(\mu)T_1(\lambda) = t(\mu)t(\lambda)$$

Local Hamiltonians are obtained by expanding around a point $\lambda = \lambda_0$ such that $L_{12}(\lambda_0) = P_{12}$, the permutation operator. To see that we get local quantities by expanding around λ_0, we write $L_{12}(\lambda) = \sum_{i,j} e_{ij} \otimes L^{ij}(\lambda)$. Then

$$t(\lambda) = \sum_{i_1,\cdots,i_N} L_1^{i_N i_1} \otimes L_2^{i_1 i_2} \otimes \cdots \otimes L_N^{i_{N-1} i_N}$$

The point λ_0 is such that $L^{ij}(\lambda_0) = e_{ji}$. Replacing $L^{ij}(\lambda)|_{\lambda_0} = e_{ji}$, we get

$$\begin{aligned}
t(\lambda)|_{\lambda_0} &= \sum_{i_1,\cdots,i_N} e_1^{i_1 i_N} \otimes e_2^{i_2 i_1} \otimes \cdots \otimes e_N^{i_N i_{N-1}} \\
t^{-1}(\lambda)|_{\lambda_0} &= \sum_{j_1,\cdots,j_N} e_1^{j_N j_1} \otimes e_2^{j_1 j_2} \otimes \cdots \otimes e_N^{j_{N-1} j_N} \\
t'(\lambda)|_{\lambda_0} &= \sum_{i_1,\cdots,i_N} e_1^{i_1 i_N} \otimes e_2^{i_2 i_1} \otimes \cdots e_{n-1}^{i_{n-1} i_{n-2}} \otimes (L_n^{i_{n-1} i_n})' \otimes e_{n+1}^{i_{n+1} i_n} \otimes \cdots e_N^{i_N i_{N-1}}
\end{aligned}$$

so that

$$\boxed{\frac{d}{d\lambda}\log t(\lambda)|_{\lambda_0} = t^{-1}(\lambda)t'(\lambda)|_{\lambda_0} = \sum_n \sum_{ijk} e_n^{ij}(L_n^{ik})' \otimes e_{n+1}^{jk}}$$

In this expression, we see that only nearest neighbours sites are coupled. More generally, a derivative of order p couples $p+1$ sites. In our case

$$\boxed{L_{1n}(\lambda) = \left(\lambda - \tfrac{i\hbar}{2}\right) \mathrm{Id} + i\hbar P_{1n} = \left(\lambda + \tfrac{i\hbar}{2}\right) R_{1n}\left(\lambda - \tfrac{i\hbar}{2}\right)} \tag{5.7}$$

so that

$$L_{1n}\left(\tfrac{i\hbar}{2}\right) = i\hbar P_{1n}$$

Expanding around that point we get a local Hamiltonian.

$$H_{XXX} = \sum_n \vec{s}_n \vec{s}_{n+1}$$

This is the Heisenberg spin chain Hamiltonian.

Another remark obvious from eq.(5.7) is that $L(\lambda)$ and $R(\lambda)$ are essentially the same thing and eq.(5.4) is identical to the Yang-Baxter equation for $R(\lambda)$.

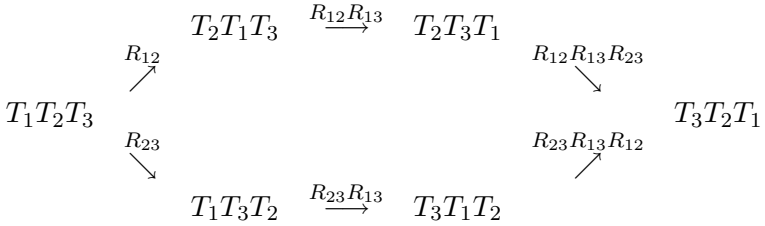
The product of operators is associative. Writing this condition on eq.(5.5) puts a constraint on the matrix $R(\lambda - \mu)$.

Proposition 32 *A sufficient condition for associativity is that the matrix $R(\lambda - \mu)$ satisfies the Yang-Baxter equation*

$$\boxed{R_{12}(\lambda_1 - \lambda_2)R_{13}(\lambda_1 - \lambda_3)R_{23}(\lambda_2 - \lambda_3) = R_{23}(\lambda_2 - \lambda_3)R_{13}(\lambda_1 - \lambda_3)R_{12}(\lambda_1 - \lambda_2)} \tag{5.8}$$

This is an equation in $End(V_1 \otimes V_2 \otimes V_3)$. The notation R_{12} means as usual that it acts non trivially on the space $V_1 \otimes V_2$, and is the identity on V_3 ...

Proof. One should check the equation by a direct calculation. However there is a good reason for this equation to hold. Consider the product $T_1(\lambda_1)T_2(\lambda_2)T_3(\lambda_3)$. There are two ways to bring it to the form $T_3(\lambda_3)T_2(\lambda_2)T_1(\lambda_1)$ with the help of eq.(5.5):



Following the upper path of the diagram produces the combination of R-matrices in the left hand side of eq.(5.8), and the lower path produces the right hand side of this equation. The Yang-Baxter equation ensures the commutativity of the above diagram, which itself reflects the associativity of the product of the T's. ∎

Remark. In the classical limit eq.(5.6), the Yang-Baxter equation becomes the classical Yang-Baxter equation

$$[r_{12}, r_{13}] + [r_{12}, r_{23}] + [r_{13}, r_{23}] = 0$$

We end this section by listing a few supplementary important properties of the monodromy matrix $T_N(\lambda)$ of the XXX spin chain. Writing

$$T_N(\lambda) = \begin{pmatrix} A(\lambda) & B(\lambda) \\ C(\lambda) & D(\lambda) \end{pmatrix}$$

we see that $A(\lambda)$ and $D(\lambda)$ are polynomials of degree N in λ, while $B(\lambda)$ and $C(\lambda)$ are polynomials of degree $N-1$. In fact we have

$$\boxed{T_N(\lambda) = \lambda^N \mathrm{Id} + i\lambda^{N-1}\ \vec{\sigma}\cdot\vec{S} + O(\lambda^{N-2})}$$

where $\vec{S}$ is the total spin operator

$$\vec{S} = \sum_n \vec{s}_n$$

Also, we have the important conjugation property (for real λ)

$$\boxed{B^\dagger(\lambda) = -C(\lambda)}$$

We note the obvious relation ($SU(2)$ symmetry)

$$[L_n(\lambda), \frac{\hbar}{2}\sigma^a + s_n^a] = 0$$

which implies as well

$$\boxed{[T_N(\lambda), \frac{\hbar}{2}\sigma^a + S^a] = 0} \tag{5.9}$$

Finally, we have the relation (quantum determinant)

$$\sigma^2 L_n(\lambda)\sigma^2 L_n^t(\lambda - i\hbar) = (\lambda(\lambda - i\hbar) + \hbar^2 s(s+1))\mathrm{Id}$$

which implies the same type of relation on $T(\lambda)$

$$\boxed{\sigma^2 T(\lambda)\sigma^2 T^t(\lambda - i\hbar) = (\lambda(\lambda - i\hbar) + \hbar^2 s(s+1))^N \mathrm{Id}} \tag{5.10}$$

5.3 Algebraic Bethe Ansatz.

One can use the fundamental commutation relations

$$R_{12}(\lambda,\mu)T_1(\lambda)T_2(\mu) = T_2(\mu)T_1(\lambda)R_{12}(\lambda,\mu) \tag{5.11}$$

to diagonalize $t(\lambda) = \mathrm{Tr}T(\lambda)$. The method applies to the case of 2×2 matrices. Larger matrices will be treated in the next Chapter. Let

$$T(\lambda) = \begin{pmatrix} A(\lambda) & B(\lambda) \\ C(\lambda) & D(\lambda) \end{pmatrix}$$

and let us write the R-matrix in the form

$$R(\lambda,\mu) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & c(\lambda,\mu) & b(\lambda,\mu) & 0 \\ 0 & b(\lambda,\mu) & c(\lambda,\mu) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

so that the formulae will apply to cases more genaral than the XXX spin chain. Writing explicitly eq.(5.11) we obtain the following set of commutation relations:

$$\begin{aligned}
[A(\lambda),A(\mu)] &= 0 \\
[D(\lambda),D(\mu)] &= 0 \\
[B(\lambda),B(\mu)] &= 0 \\
[C(\lambda),C(\mu)] &= 0 \\
B(\lambda)A(\mu) &= b(\lambda,\mu)B(\mu)A(\lambda) + c(\lambda,\mu)A(\mu)B(\lambda) \\
B(\mu)D(\lambda) &= b(\lambda,\mu)B(\lambda)D(\mu) + c(\lambda,\mu)D(\lambda)B(\mu) \\
C(\lambda)A(\mu) &= c(\mu,\lambda)A(\mu)C(\lambda) + b(\mu,\lambda)C(\mu)A(\lambda) \\
C(\mu)D(\lambda) &= c(\mu,\lambda)D(\lambda)C(\mu) + b(\mu,\lambda)C(\lambda)D(\mu) \\
c(\lambda,\mu)[C(\lambda),B(\mu)] &= -b(\lambda,\mu)(A(\lambda)D(\mu) - A(\mu)D(\lambda)) \\
c(\lambda,\mu)[A(\lambda),D(\mu)] &= -b(\lambda,\mu)(C(\lambda)B(\mu) - C(\mu)B(\lambda))
\end{aligned}$$

Although all these relations are useful, we will need here only three of them which we rewrite in the convenient form:

$$[B(\lambda),B(\mu)] = 0 \tag{5.12}$$

$$A(\lambda)B(\mu) = \frac{1}{c(\mu,\lambda)}B(\mu)A(\lambda) - \frac{b(\mu,\lambda)}{c(\mu,\lambda)}B(\lambda)A(\mu) \tag{5.13}$$

$$D(\lambda)B(\mu) = \frac{1}{c(\lambda,\mu)}B(\mu)D(\lambda) - \frac{b(\lambda,\mu)}{c(\lambda,\mu)}B(\lambda)D(\mu) \tag{5.14}$$

The idea of the algebraic Bethe Ansatz is to use the operator $B(\mu)$ as a creation operator to generate eigenstates of the operator

$$t(\lambda) = A(\lambda) + D(\lambda)$$

We recall that this operator is a generating function for the commuting Hamiltonians.

We need a reference state to start with. On each site, introduce a local vacuum $|0\rangle_n$ in the local Hilbert space $\mathcal{H}_n$.

$$|0\rangle_n = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

The action of the local operator $L_n(\lambda)$ eq.(5.3) on $|0\rangle_n$ is given by

$$L_n(\lambda)|0\rangle_n = \begin{pmatrix} \alpha(\lambda)|0\rangle_n & (L_n(\lambda))_{12}|0\rangle_n \\ 0 & \delta(\lambda)|0\rangle_n \end{pmatrix}$$

The action of the monodromy matrix $T(\lambda)$ on the state

$$|0\rangle = |0\rangle_1 \otimes |0\rangle_2 \otimes \ldots \otimes |0\rangle_N$$

is obtained by taking the product of these local triangular matrices. Therefore, it is also triangular and the diagonal elements are just the products of the local diagonal elements

$$T(\lambda)|0\rangle = \begin{pmatrix} \alpha(\lambda)^N|0\rangle & B(\lambda)|0\rangle \\ 0 & \delta(\lambda)^N|0\rangle \end{pmatrix}$$

Thus we have

$$\begin{aligned} A(\lambda)|0\rangle &= \alpha(\lambda)^N|0\rangle \\ D(\lambda)|0\rangle &= \delta(\lambda)^N|0\rangle \\ C(\lambda)|0\rangle &= 0 \end{aligned}$$

Notice that $|0\rangle$ is an eigenstate of $\Im(\lambda) = A(\lambda) + D(\lambda)$. We shall now generate new eigenstates by applying operators $B(\mu_i)$ to $|0\rangle$.

Proposition 33 *The vectors $\Omega(\mu_1, \ldots, \mu_M)$ defined by:*

$$\Omega(\mu_1, \ldots, \mu_M) = B(\mu_1)B(\mu_2)\ldots B(\mu_M)|0\rangle \tag{5.15}$$

are eigenstates of $t(\lambda)$ if the following relations are satisfied :

$$\left(\frac{\alpha(\mu_j)}{\delta(\mu_j)}\right)^N = \prod_{\substack{l=1 \\ l\neq j}}^{M} \frac{c(\mu_l, \mu_j)}{c(\mu_j, \mu_l)} \tag{5.16}$$

The corresponding eigenvalue of $\Im(\lambda)$ is given by

$$t(\lambda; \{\mu_j\}) = \alpha^N(\lambda) \prod_{l=1}^{M} \frac{1}{c(\mu_l, \lambda)} + \delta^N(\lambda) \prod_{l=1}^{M} \frac{1}{c(\lambda, \mu_l)} \tag{5.17}$$

Proof. To evaluate the action of the operator $A(\lambda)$ on the vector $\Omega(\mu_1,\ldots,\mu_M)$, we use the commutation relation eq.(5.13). When pushing $A(\lambda)$ to the right of the chain of $B(\mu_j)$ operators, we generate 2^M terms, coming naturally in $M+1$ groups

$$\begin{aligned} A(\lambda)B(\mu_1)B(\mu_2)\ldots B(\mu_M)|0\rangle &= \Lambda(\lambda;\mu_1,\ldots,\mu_M)\prod_{l=1}^{M}B(\mu_l)|0\rangle \\ &\quad+\sum_{j=1}^{M}\Lambda_j(\lambda;\mu_1,\ldots,\mu_M)B(\lambda)\prod_{\substack{l=1\\l\neq j}}^{M}B(\mu_l)|0\rangle \end{aligned}$$

where Λ and Λ_j are numerical coefficients. The first term is of the form required for an eigenvector, and we call it a wanted term, the other ones are the unwanted terms.
The wanted term is obtained by using only the first term in eq.(5.13) for the commutation of $A(\lambda)$ through the $B(\mu_j)$. Otherwise one of the B operators would carry the argument λ. Since this is the only way to obtain it, this gives immediately

$$\Lambda(\lambda;\mu_1,\ldots,\mu_M)=\alpha^N(\lambda)\prod_{l=1}^{M}\frac{1}{c(\mu_l,\lambda)}$$

To obtain the coefficient $\Lambda_j(\lambda;\mu_1,\ldots,\mu_M)$ of the unwanted term, we first bring the operator $B(\mu_j)$ in the first position in the chain of B's (the B operators commute by eq.(5.12)). In the first commutation of $A(\lambda)$, we use the second term of eq.(5.13). The operator A carries now the argument μ_j and must keep it until hitting the reference state $|0\rangle$. Hence we commute it using only the first term in eq.(5.13). This gives uniquely

$$\Lambda_j(\lambda;\mu_1,\ldots,\mu_M)=-\frac{b(\mu_j,\lambda)}{c(\mu_j,\lambda)}\alpha^N(\mu_j)\prod_{\substack{l=1\\l\neq j}}^{M}\frac{1}{c(\mu_l,\mu_j)}$$

Similarly, we have

$$\begin{aligned} D(\lambda)B(\mu_1)\ldots B(\mu_M)|0\rangle &= \Lambda'(\lambda;\mu_1,\ldots,\mu_M)\prod_{l=1}^{M}B(\mu_l)|0\rangle \\ &\quad+\sum_{j=1}^{M}\Lambda'_j(\lambda;\mu_1,\ldots,\mu_M)B(\lambda)\prod_{\substack{l=1\\l\neq j}}^{M}B(\mu_l)|0\rangle \end{aligned}$$

with

$$\begin{aligned} \Lambda'(\lambda;\mu_1,\ldots,\mu_M) &= \delta^N(\lambda)\prod_{l=1}^{M}\frac{1}{c(\lambda,\mu_l)} \\ \Lambda'_j(\lambda;\mu_1,\ldots,\mu_M) &= -\frac{b(\lambda,\mu_j)}{c(\lambda,\mu_j)}\delta^N(\mu_j)\prod_{\substack{l=1\\l\neq j}}^{M}\frac{1}{c(\mu_j,\mu_l)} \end{aligned}$$

The vector $\Omega(\mu_1, \ldots, \mu_M)$ will be an eigenvector of $\mathfrak{F}(\lambda)$ if all the unwanted terms vanish. This gives the conditions $\Lambda_j + \Lambda'_j = 0$. Since the function $b(\lambda)/c(\lambda)$ is an odd function of λ, these conditions reduce to eq.(5.16). The corresponding eigenvalue is simply $\Lambda(\lambda; \mu_1, \ldots, \mu_M) + \Lambda'(\lambda; \mu_1, \ldots, \mu_M)$. ■

Let us write these equations explicitly in the case of the XXX spin chain. The Bethe equations take the form

$$\boxed{\left(\frac{\mu_j + \frac{i\hbar}{2}}{\mu_j - \frac{i\hbar}{2}}\right)^N = \prod_{k \neq j} \frac{\mu_j - \mu_k + i\hbar}{\mu_j - \mu_k - i\hbar}} \tag{5.18}$$

and the corresponding eigenvalue is

$$\boxed{t(\lambda; \{\mu_j\}) = \left(\lambda + \frac{i\hbar}{2}\right)^N \prod_k \frac{\lambda - \mu_k - i\hbar}{\lambda - \mu_k} + \left(\lambda - \frac{i\hbar}{2}\right)^N \prod_k \frac{\lambda - \mu_k + i\hbar}{\lambda - \mu_k}} \tag{5.19}$$

A special property of the XXX spin chain is that the Bethe eigenvectors are all highest weights vectors for the total spin operator $\vec{S}$. To prove it we start from eq.(5.9) which implies

$$[S^3, B(\lambda)] = -\hbar B(\lambda), \quad [S^+, B(\lambda)] = \frac{\hbar}{2}(A(\lambda) - D(\lambda))$$

The first relation implies immediately

$$\boxed{S^3 \Omega(\mu_1, \cdots, \mu_M) = \left(\frac{N}{2} - M\right) \Omega(\mu_1, \cdots, \mu_M)}$$

To prove that

$$\boxed{S^+ \Omega(\mu_1, \cdots, \mu_M) = 0}$$

we use the second relation to get

$$S^+ \Omega(\mu_1, \cdots, \mu_M) = \frac{1}{2} \sum_j B(\mu_1) \cdots B(\mu_{j-1})(A(\mu_j) - D(\mu_j)) B(\mu_{j+1}) \cdots B(\mu_M)|0\rangle$$

when we push the oprators A and D to the right, we get a sum of terms of the form

$$S^+ \Omega(\mu_1, \cdots, \mu_M) = \frac{1}{2} \sum_j \Gamma_j(\mu_1, \cdots, \mu_M) B(\mu_1) \cdots \widehat{B(\mu_j)} \cdots B(\mu_M)|0\rangle$$

Let us compute $\Gamma_1(\mu_1, \cdots, \mu_M)$. The only way to get this term is to start from

$$(A(\mu_1) - D(\mu_1))B(\mu_2)\cdots B(\mu_M)|0\rangle$$

and push $A(\mu_1)$ and $D(\mu_1)$ to the right using only the first terms in eq.(5.13,5.14). Hence we find

$$\Gamma_1(\mu_1, \cdots, \mu_M) = \alpha^N(\mu_1) \prod_{k\neq 1} \frac{1}{c(\mu_k, \mu_1)} - \delta^N(\mu_1) \prod_{k\neq 1} \frac{1}{c(\mu_1, \mu_k)}$$

This vanishes if the Bethe equations are satisfied. Clearly the same is true for the other $\Gamma_j(\mu_1, \cdots, \mu_M)$ by symmetry.

5.3.1 Baxter equation.

Let us introduce the polynomial

$$Q(\lambda) = \prod_{m=1}^{M} (\lambda - \mu_m)$$

Then the Bethe equations eq.(5.18) ca be rewritten as

$$\left(\mu_k + \frac{i\hbar}{2}\right)^N Q(\mu_k - i\hbar) + \left(\mu_k - \frac{i\hbar}{2}\right)^N Q(\mu_k + i\hbar) = 0$$

This means that the polynomial of degree $N + M$

$$\left(\lambda + \frac{i\hbar}{2}\right)^N Q(\lambda - i\hbar) + \left(\lambda - \frac{i\hbar}{2}\right)^N Q(\lambda + i\hbar)$$

is divisible by $Q(\lambda)$. Hence there exists a polynomial $t(\lambda)$ of degree N such that

$$\boxed{\left(\lambda + \frac{i\hbar}{2}\right)^N Q(\lambda - i\hbar) + \left(\lambda - \frac{i\hbar}{2}\right)^N Q(\lambda + i\hbar) = t(\lambda)Q(\lambda)}$$

This is Baxter's equation. The polynomial $t(\lambda)$ is the same as in eq.(5.19) because that equation can be rewritten as

$$t(\lambda; \{\mu_j\}) = \left(\lambda + \frac{i\hbar}{2}\right)^N \frac{Q(\lambda - i\hbar)}{Q(\lambda)} + \left(\lambda - \frac{i\hbar}{2}\right)^N \frac{Q(\lambda + i\hbar)}{Q(\lambda)}$$

hence the coefficients of this polynomial are just the eigenvalues of the set of commuting Hamiltonians.

Just as in the Gaudin model, it is interesting to introduce the Riccati version of this equation. We set (do not confuse this S with the total spin !!)

$$\boxed{S(\lambda) = \frac{Q(\lambda - i\hbar)}{Q(\lambda)}}$$

Then Baxter equation becomes

$$\boxed{\left(\lambda+\frac{i\hbar}{2}\right)^N S(\lambda)+\left(\lambda-\frac{i\hbar}{2}\right)^N S^{-1}(\lambda+i\hbar)=t(\lambda)}$$

This equation determines both $S(\lambda)$ and $t(\lambda)$. To find the equation for $t(\lambda)$, we expand around $\lambda=-i\hbar/2$ getting

$$(\epsilon-i\hbar)^N S^{-1}(\epsilon+i\hbar/2)=t(\epsilon-i\hbar/2)-\epsilon^N S(\epsilon-i\hbar/2)$$

Similarly, expanding around $\lambda=i\hbar/2$ we get

$$(\epsilon+i\hbar)^N S(\epsilon+i\hbar/2)=t(\epsilon+i\hbar/2)-\epsilon^N S^{-1}(\epsilon+3i\hbar/2)$$

Multiplying the two, we find

$$\boxed{t\left(\epsilon+\frac{i\hbar}{2}\right)t\left(\epsilon-\frac{i\hbar}{2}\right)=(\hbar^2+\epsilon^2)^N+O(\epsilon^N)} \tag{5.20}$$

This is a system of N equations for the $N+1$ coefficients of $t(\lambda)$ which determines it completely if we remember that $t(\lambda)=2\lambda^N+O(\lambda^{N-1})$. In fact the sub leading coefficient is also easy to compute

$$t(\lambda)=2\lambda^N-\hbar^2\left(\frac{N(N-1)}{4}+M(M-N-1)\right)\lambda^{N-2}+O(\lambda^{N-3})$$

For chains with a small number of sites N, we find by solving directly eq.(5.20) (we set $\hbar=1$):

N	$t(\lambda)$	M	S^3	mult	dim
2	$2\lambda^2-\frac{1}{2}$	0	1	3	
2	$2\lambda^2+\frac{3}{2}$	1	0	1	$2^2=3+1$
3	$2\lambda^3-\frac{3}{2}\lambda$	0	$\frac{3}{2}$	4	
3	$2\lambda^3+\frac{3}{2}\lambda-\frac{\sqrt{3}}{2}$	1	$\frac{1}{2}$	2	
3	$2\lambda^3+\frac{3}{2}\lambda+\frac{\sqrt{3}}{2}$	1	$\frac{1}{2}$	2	$2^3=4+2\times 2$
4	$2\lambda^4-3\lambda^2+\frac{1}{8}$	0	2	5	
4	$2\lambda^4+\lambda^2-\frac{7}{8}$	1	1	3	
4	$2\lambda^4+\lambda^2-2\lambda+\frac{1}{8}$	1	1	3	
4	$2\lambda^4+\lambda^2+2\lambda+\frac{1}{8}$	1	1	3	
4	$2\lambda^4+3\lambda^2-\frac{3}{8}$	2	0	1	
4	$2\lambda^4+3\lambda^2+\frac{13}{8}$	2	0	1	$2^4=5+3\times 3+2\times 1$

This construction can be generalized to the case of a spin-s chain. The Baxter equation then reads

$$\boxed{(\lambda+i\hbar s)^N Q(\lambda-i\hbar)+(\lambda-i\hbar s)^N Q(\lambda+i\hbar)=t(\lambda)Q(\lambda)}$$

The Riccati equations reads

$$\boxed{(\lambda+i\hbar s)^N S(\lambda)+(\lambda-i\hbar s)^N S^{-1}(\lambda+i\hbar)=t(\lambda)} \tag{5.21}$$

For $s=1$ for instance we expand around $\lambda=i\hbar$, $\lambda=0$, $\lambda=-i\hbar$ to get

$$\begin{aligned}(\epsilon+2i\hbar)^N S(\epsilon+i\hbar) &= t(\epsilon+i\hbar)+O(\epsilon^N)\\ (\epsilon+i\hbar)^N S(\epsilon)+(\epsilon-i\hbar)^N S^{-1}(\epsilon+i\hbar) &= t(\epsilon)\\ (\epsilon-2i\hbar)^N S^{-1}(\epsilon) &= t(\epsilon-i\hbar)+O(\epsilon^N)\end{aligned}$$

from which we deduce ($s=1$)

$$\boxed{t(\epsilon+i\hbar)t(\epsilon)t(\epsilon-i\hbar)=(\epsilon-i\hbar)^N(\epsilon+2i\hbar)^N t(\epsilon-i\hbar)+(\epsilon+i\hbar)^N(\epsilon-2i\hbar)^N t(\epsilon+i\hbar)+O(\epsilon^N)}$$

Clearly, for a spin-s, $s\geq 0$, the degree of the equation is $2s+1$. If however $s<0$ the equations generically do not lead to a finite degree equation.

5.3.2 Separated variables.

Suppose for a while that the spin in eq.(5.3) is classical. Then the spectral curve reads

$$\det(T(\lambda)-\mu)=\mu^2-t(\lambda)\mu+(\lambda^2+s^2)^N=0 \tag{5.22}$$

where $t(\lambda)=A(\lambda)+D(\lambda)=2\lambda^N+O(\lambda^{N-2})$. This is a hyperelliptic curve. To put it in canonical form, we set $y=\mu-\frac{1}{2}t(\lambda)$ so that

$$y^2=\frac{1}{4}t^2(\lambda)-(\lambda^2+s^2)^N\simeq(t_2-Ns^2)\lambda^{2N-2}+\cdots$$

from what we see that the genus is $N-2$. Remark that the dynamical moduli H_i appear linearly in $t(\lambda)$. If we vary them, we get

$$\frac{\delta\mu}{\mu}d\lambda=\frac{\delta t(\lambda)d\lambda}{\mu-(\lambda^2+s^2)^N\mu^{-1}}=\frac{\delta t(\lambda)}{2y}d\lambda=\text{holomorphic}$$

Hence the natural (and correct !) Poisson bracket for the separated variables is

$$\{\lambda_k,\mu_{k'}\}=\delta_{kk'}\mu_k \tag{5.23}$$

Setting otherwise

$$\mu = (\lambda + is)^N S$$

the spectral curve becomes

$$\boxed{(\lambda + is)^N S + (\lambda - is)^N S^{-1} = t(\lambda)}$$

We see that the Riccati equation eq.(5.21) can be viewed as a deformation of the spectral curve. The fact that this deformation involves difference operators instead of differential operators can be understood because the natural quantization of eq.(5.23) is that μ_k is a shift operator

$$\boxed{\mu_k \lambda_k = (\lambda_k - i\hbar)\mu_k} \tag{5.24}$$

We now prove this formula in the quantum XXX spin chain. Following Sklyanin, we introduce the quantum separated variables as the zeroes of $B(\lambda)$.

$$\boxed{B(\lambda) = iS^- \prod_{k=1}^{N-1} (\lambda - \lambda_k)}$$

Because $[B(\lambda), B(\lambda')] = 0$, we have

$$\boxed{[S^-, \lambda_k] = 0, \quad [\lambda_k, \lambda_{k'}] = 0}$$

For an operator $X(\lambda) = \sum_n X_n \lambda^n$, we define

$$\mathcal{N}[X(\lambda_k)] = \sum_n \lambda_k^n X_n$$

where the operator λ_k is ordered on the left. Substituting $\mu \to \lambda_k$ according to this rule in eq.(5.13) which we write in the form

$$(\lambda - \mu)A(\mu)B(\lambda) = (\lambda - \mu + i\hbar)B(\lambda)A(\mu) - i\hbar B(\mu)A(\lambda)$$

we get

$$\mathcal{N}[(\lambda - \lambda_k)A(\lambda_k)]B(\lambda) = \mathcal{N}[(\lambda - \lambda_k + i\hbar)B(\lambda)A(\lambda_k)] - i\hbar \mathcal{N}[B(\lambda_k)]A(\lambda)$$

Now $\mathcal{N}[B(\lambda_k)] = B(\lambda_k) = 0$ because all coefficients of $B(\lambda)$ commute with λ_k. Next, obviously $\mathcal{N}[(\lambda - \lambda_k)A(\lambda_k)] = (\lambda - \lambda_k)\mathcal{N}[A(\lambda_k)]$. Finally again because $B(\lambda)$ commutes with λ_k, we have $\mathcal{N}[(\lambda - \lambda_k + i\hbar)B(\lambda)A(\lambda_k)] = (\lambda - \lambda_k + i\hbar)B(\lambda)\mathcal{N}[A(\lambda_k)]$ Hence, setting

$$\mu_k^{(+)} = \mathcal{N}[A(\lambda_k)]$$

we get

$$(\lambda - \lambda_k)\ \mu_k^{(+)}\ B(\lambda) = (\lambda - \lambda_k + i\hbar) B(\lambda)\ \mu_k^{(+)}$$

and therefore

$$(\lambda - \lambda_k)\ \mu_k^{(+)}\ S^- \prod_j (\lambda - \lambda_j) = (\lambda - \lambda_k + i\hbar) S^- \prod_j (\lambda - \lambda_j) \mu_k^{(+)}$$

By eq.(5.9), we have

$$[S^-, A(\lambda)] = \frac{1}{2} B(\lambda)$$

which implies also

$$[S^-, \mu_k^{(+)}] = 0$$

Simplifying on the left by $(\lambda - \lambda_k) S^-$, we get

$$\mu_k^{(+)} \prod_j (\lambda - \lambda_j) = (\lambda - \lambda_k + i\hbar) \prod_{j \neq k} (\lambda - \lambda_j) \mu_k^{(+)}$$

this means

$$\mu_k^{(+)} \lambda_k = (\lambda_k - i\hbar) \mu_k^{(+)}$$

at least when $\mu_k^{(+)}$ acts on symmetric functions of the λ_k's. By exactly the same argument, defining

$$\mu_k^{(-)} = \mathcal{N}[D(\lambda_k)]$$

we show that

$$\mu_k^{(-)} \lambda_k = (\lambda_k + i\hbar) \mu_k^{(-)}$$

From their definitions, we have

$$\mu_k^{(+)} - (\mathcal{N}[A(\lambda_k) + D(\lambda_k))] + \mu_k^{(-)} = 0 \tag{5.25}$$

To finish the identification of $\mu_k^{(\pm)}$ with $\mu^{\pm 1}$ in eq.(5.22), we show that

$$\boxed{\mu_k^{(+)} \mu_k^{(-)} = (\lambda_k^2 - i\hbar\lambda_k + \hbar^2 s(s+1))^N} \tag{5.26}$$

We start with the identity eq.(5.10) which implies in particular

$$A(\lambda) D(\lambda - i\hbar) - B(\lambda) C(\lambda - i\hbar) = (\lambda(\lambda - i\hbar) + \hbar^2 s(s+1))^N$$

Now, we have

$$\begin{aligned} \mathcal{N}[A(\lambda_k) D(\lambda_k - i\hbar)] &= \sum_n (\lambda_k - i\hbar)^n \mathcal{N}[A(\lambda_k)] D_n \\ &= \sum_n (\lambda_k - i\hbar)^n \mu_k^{(+)} D_n = \sum_n \mu_k^{(+)} \lambda_k^n D_n = \mu_k^{(+)} \mathcal{N}[D(\lambda_k)] = \mu_k^{(+)} \mu_k^{(-)} \end{aligned}$$

More simply, we have $\mathcal{N}[B(\lambda_k)C(\lambda_k - i\hbar)] = 0$. This completes the proof of eq.(5.26).

We can solve eq.(5.26) by redefining

$$\mu_k^{(+)} = (\lambda_k + i\hbar s)^N S_k, \quad \mu_k^{(-)} = (\lambda_k - i\hbar s)^N S_k^{-1}$$

where S_k is the shift operator for the variable λ_k

$$S_k \lambda_k = (\lambda_k - i\hbar) S_k$$

Then eq.(5.25) becomes

$$\boxed{(\lambda_k + i\hbar s)^N S_k + (\lambda_k - i\hbar s)^N S_k^{-1} = \mathcal{N}(t(\lambda_k))}$$

As in the Gaudin model, the Bethe states can be expressed in terms of the separated variables. The formula is

$$\Omega(\mu_1, \cdots, \mu_M) = (iS^-)^M \prod_k Q(\lambda_k)|0\rangle$$

At this stage, what seems to be still missing is the scalar product in this representation.

Bibliography

[1] C. N. Yang. *" Some exact results for the many body problem in one dimension with repulsive delta function interaction."* Phys. Rev. Lett. 19 (1967) pp. 1312-1315.

[2] E.K.Slyanin, L.A.Takhtadjan, L.D.Faddeev, *"Quantum Inverse Problem Method.I."* Theor.Math.Phys. 40 (1979) pp.194-220.

[3] L.D.Faddeev, *"Integrable Models in 1+1 Dimensional Quantum Field Theory"*. Les Houches Lectures 1982. Elsevier Science Publishers (1984).

[4] E. Sklyanin, *"The quantum Toda chain"*. Lecture notes in Physics, Vol 226 (1985) p. 196.

[5] L.D. Faddeev *Algebraic aspects of Bethe Ansatz.* Int.J.Mod.Phys. A10 (1995) 1845-1878, hep-th/9404013.

[6] L.D. Faddeev, *How Bethe Ansatz works for integrable models.* Les Houches Lectures 1996. hep-th/9605187

Chapter 6

Nested Bethe Ansatz.

There exists a non trivial generalization of Bethe equations to the case of $SU(n)$ spins. We present here the algebraic version of this construction.

6.1 The R-matrix of the Affine sl_{n+1} algebra.

Our starting point will be the R-matrix associated to the quantum loop algebra $\mathcal{U}_q(sl_{n+1} \otimes C(\lambda, \lambda^{-1})$. It can be written as

$$\boxed{R(\lambda) = \frac{\lambda R - \lambda^{-1}\overline{R}}{\lambda q - \lambda^{-1}q^{-1}}}$$

with

$$\begin{aligned} R &= \sum_{i \neq j} e_{ii} \otimes e_{jj} + q \sum_i e_{ii} \otimes e_{ii} + (q - q^{-1}) \sum_{i<j} e_{ij} \otimes e_{ji} \\ \overline{R} &= \sum_{i \neq j} e_{ii} \otimes e_{jj} + q^{-1} \sum_i e_{ii} \otimes e_{ii} - (q - q^{-1}) \sum_{i>j} e_{ij} \otimes e_{ji} \end{aligned}$$

Here e_{ij} denotes the matrix with elements $[e_{ij}]_{\alpha\beta} = \delta_{i\alpha}\delta_{j\beta}$. The matrices R and $\overline{R}$ are the R-matrices of $\mathcal{U}_q(sl_{n+1})$.

The matrix $R(\lambda)$ satisfies the Yang-Baxter equation.

$$R_{12}(\lambda/\mu)R_{13}(\lambda)R_{23}(\mu) = R_{23}(\mu)R_{13}(\lambda)R_{12}(\lambda/\mu)$$

Let us write explicitly the matrix elements of $R(\lambda)$. They read

$$R_{ab}^{\alpha\beta}(\lambda) = \frac{\lambda - \lambda^{-1}}{\lambda q - \lambda^{-1}q^{-1}} \delta_a^\alpha \delta_b^\beta + \left[\left(1 - \frac{\lambda - \lambda^{-1}}{\lambda q - \lambda^{-1}q^{-1}} \right) \delta_{ab} + \frac{q - q^{-1}}{\lambda q - \lambda^{-1}q^{-1}} \lambda^{\epsilon(a-b)} (1 - \delta_{ab}) \right] \delta_b^\alpha \delta_a^\beta$$

6.2 sl_{n+1} generalization of the XXZ model.

We use R to define the matrix of statistical weights of a vertex model, and we construct as usual the transfer matrix

$$T_{1N}(\lambda) = R_{11}(\lambda)R_{12}(\lambda)\cdots R_{1N}(\lambda)$$

This matrix satisfies

$$R_{12}(\lambda/\mu)T_{1N}(\lambda)T_{2N}(\mu) = T_{2N}(\mu)T_{1N}(\lambda)R_{12}(\lambda/\mu) \tag{6.1}$$

As usual, the traces $\mathcal{T}_N(\lambda) = tr_1 T_{1N}(\lambda)$ commute

$$[\mathcal{T}_N(\lambda), \mathcal{T}_N(\mu)] = 0$$

Since $R(\lambda)|_{\lambda=1} = P$, the permutation operator, we can construct local commuting quantities

Proposition 34 *Let*

$$H = -\frac{q-q^{-1}}{2}\lambda\frac{\partial \log \mathcal{T}_N(\lambda)}{\partial\lambda}\bigg|_{\lambda=1} + \frac{1}{2}N\,(q+q^{-1})$$

we have

$$\boxed{H = -\sum_{n=1}^{N}\left\{\sum_{i\neq j} e_n^{ij}e_{n+1}^{ji} + \frac{(q+q^{-1})}{2}\sum_i e_n^{ii}e_{n+1}^{ii} - \frac{(q-q^{-1})}{2}\sum_{i,j}\epsilon(i-j)e_n^{ii}e_{n+1}^{jj}\right\}}$$

Proof. We write $R_{12}(\lambda) = \sum_{i,j} e_{ij}\otimes L^{ij}(\lambda)$ with

$$L^{ii}(\lambda) = \frac{\lambda q^{e_{ii}} - \lambda^{-1}q^{-e_{ii}}}{\lambda q - \lambda^{-1}q^{-1}};\quad L^{ij}(\lambda) = \frac{q-q^{-1}}{\lambda q - \lambda^{-1}q^{-1}}\lambda^{\epsilon(j-i)}e_{ji}$$

then remark that $L^{ij}(\lambda)|_{\lambda=1} = e_{ji}$, so that the general construction applies:

$$\mathcal{T}_N^{-1}(\lambda)\mathcal{T}_N'(\lambda)|_{\lambda=1} \;=\; \sum_n\sum_{ijk} e_n^{ij}(L_n^{ik})'\otimes e_{n+1}^{jk}$$

■

This Hamiltonian acts on the Hilbert space

$$\mathcal{H}_N = \prod_{j=1}^{N}\otimes h_j;\quad h_j = \mathbb{C}^{n+1}$$

For $n = 1$, we recover the XXZ model.

6.3 Commutation relations.

We can arrange the matrix elements of $R(\lambda)$ as follows

$$R(\lambda) \;=\; \begin{array}{c} \\ 11 \\ 1i \\ i1 \\ ii' \end{array} \begin{array}{c} \begin{array}{cccc} 11 & 1j & j1 & jj' \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & c(\lambda) & b(\lambda) & 0 \\ 0 & \tilde{b}(\lambda) & c(\lambda) & 0 \\ 0 & 0 & 0 & R^{(1)} \end{pmatrix} \end{array}$$

where

$$[b(\lambda)]_{ij} = \frac{q-q^{-1}}{\lambda q - \lambda^{-1}q^{-1}}\lambda\delta_{ij}; \quad [\tilde{b}(\lambda)]_{ij} = \frac{q-q^{-1}}{\lambda q - \lambda^{-1}q^{-1}}\lambda^{-1}\delta_{ij}; \quad [c(\lambda)]_{ij} = \frac{\lambda-\lambda^{-1}}{\lambda q - \lambda^{-1}q^{-1}}\delta_{ij}$$

Since $b, \tilde{b}, c$ are all proportionnal to the identity matrix, we will treat them as ordinary $\mathbb{C}$-numbers in what follows.

To take advantage of the block structure of the R-matrix, we also decompose the matrix $T(\lambda)$ in a similar way

$$T(\lambda) = \begin{pmatrix} A(\lambda) & B(\lambda) \\ C(\lambda) & D(\lambda) \end{pmatrix}$$

where

$$A(\lambda) = T_{11}(\lambda); \quad B_i(\lambda) = T_{1i}(\lambda); \quad C_i(\lambda) = T_{i1}(\lambda); \quad D_{ij}(\lambda) = T_{ij}(\lambda)$$

Thus $B(\lambda)$ is a line vector and $C(\lambda)$ is a column vector of dimension n. $D(\lambda)$ is a $n^2 \times n^2$ matrix. The fundamental equation eq.(6.1) yields the following relations

$$A(\lambda)B(\mu) \;=\; \frac{1}{c(\mu,\lambda)}B(\mu)A(\lambda) - \frac{b(\mu,\lambda)}{c(\mu,\lambda)}B(\lambda)A(\mu) \tag{6.2}$$

$$D_1(\lambda)B_2(\mu) \;=\; \frac{1}{c(\lambda,\mu)}B_2(\mu)D_1(\lambda)R^{(1)}_{12}(\lambda,\mu) - \frac{\tilde{b}(\lambda,\mu)}{c(\lambda,\mu)}B_1(\lambda)D_2(\mu) \tag{6.3}$$

$$B_1(\lambda)B_2(\mu) \;=\; B_2(\mu)B_1(\lambda)R^{(1)}_{12}(\lambda,\mu) \tag{6.4}$$

We will use the second relation eq.(6.3) in the form

$$D_1(\lambda)B_2(\mu) \;=\; \frac{1}{c(\lambda,\mu)}B_2(\mu)D_1(\lambda)R^{(1)}_{12}(\lambda,\mu) - \frac{\tilde{b}(\lambda,\mu)}{c(\lambda,\mu)}B_2(\lambda)D_1(\mu)P_{12}$$

where P_{12} is the permutation operator $P^{nl}_{jk} = \delta^n_k\delta^l_j$.

6.4 Reference state.

To apply the algebraic Bethe Ansatz, the first step is to find a reference state to start with. In our case this is provided by

Proposition 35 *Let* $|1\rangle$ *be the vector*

$$|1\rangle = |1\rangle_1 \otimes |1\rangle \cdots \otimes |1\rangle_N, \quad \text{where } |1\rangle_n = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

The action of operatorial entries of $T(\lambda)$ *on* $|1\rangle$ *is given by*

$$\begin{aligned} A(\lambda)|1\rangle &= \alpha^N(\lambda)|1\rangle \\ D(\lambda)|1\rangle &= \delta^N(\lambda) Id|1\rangle \\ C(\lambda)|1\rangle &= 0 \end{aligned}$$

where

$$\alpha(\lambda) = 1; \quad \delta(\lambda) = c(\lambda) = \frac{\lambda - \lambda^{-1}}{\lambda q - \lambda^{-1} q^{-1}}$$

Proof. We have

$$\begin{aligned} L_n^{i1}(\lambda)|1\rangle_n &= 0 \\ L_n^{11}(\lambda)|1\rangle_n &= \alpha(\lambda)|1\rangle_n \\ L_n^{ii}(\lambda)|1\rangle_n &= \delta(\lambda)|1\rangle_n \\ L_n^{ij}(\lambda)|1\rangle_n &= 0, \quad i \neq j \end{aligned}$$

The result is obtained by multiplying these triangular matrices at each site of the lattice. ■

The idea of the algebraic Bethe Ansatz is to look for eigenstates of the form

$$\Psi_X(\{\mu\}) = \sum_{i_1 \cdots i_p} X^{i_1 \cdots i_p} B_{i_1}(\mu_1) \cdots B_{i_p}(\mu_p)|1\rangle$$

Since B is seen as a line vector, we consider X as a column vector. We can rewrite our state in a tensor notation

$$\Psi_X(\{\mu\}) = B_1(\mu_1) \cdots B_p(\mu_p)|1\rangle X$$

We want to find the action of

$$\mathcal{T}(\lambda) = A(\lambda) + \text{Tr } D(\lambda)$$

on this vector using eqs.(6.2,6.3,6.4).

Proposition 36 *The action of* $\mathcal{T}(\lambda) = A(\lambda) + \mathrm{Tr}\ D(\lambda)$ *on the vector* $\Psi_X(\{\mu\})$ *is of the form*

$$\begin{aligned} \mathcal{T}(\lambda)\Psi_X(\{\mu\}) &= B_1(\mu_1)\cdots B_p(\mu_p)|1\rangle Y_0 && (6.5)\\ &\quad + \sum_{i=1}^{p} B_i(\lambda)B_{i+1}(\mu_{i+1})\cdots B_p(\mu_p)\cdots B_{i-1}(\mu_{i-1})|1\rangle Y_i && (6.6)\end{aligned}$$

The first term is called the wanted term, Y_0 *is given by*

$$\boxed{Y_0 = \alpha^N(\lambda)\prod_{i=1}^{p}\frac{1}{c(\mu_i,\lambda)}\ X + \delta^N(\lambda)\prod_{i=1}^{p}\frac{1}{c(\lambda,\mu_i)}\ \mathcal{T}(\lambda,\{\mu\})X} \tag{6.7}$$

The other terms are called the unwanted terms. Setting them to zero yields the equations

$$\boxed{\alpha^N(\mu_i)\prod_{j\neq i}^{p}\frac{1}{c(\mu_j,\mu_i)}X - \delta^N(\mu_i)\prod_{j\neq i}\frac{1}{c(\mu_i,\mu_j)}\mathcal{T}(\mu_i;\{\mu\})X = 0} \tag{6.8}$$

The matrix $\mathcal{T}(\lambda,\{\mu\})$ *appearing in these equations is defined as*

$$\boxed{\mathcal{T}(\lambda,\{\mu\}) = \mathrm{Tr}_0 R_{01}^{(1)}(\lambda/\mu_1)\cdots R_{0p}^{(1)}(\lambda/\mu_p)}$$

This is the transfer matrix of a spin model where the spin takes n *values instead of* $n+1$*.*

Proof. *Wanted terms.* Let us start with $A(\lambda)$. The wanted term is obtained by commuting $A(\lambda)$ through the B's using only the first term in the commutation relation eq.(6.2). We get

$$A(\lambda)\Psi_X(\{\mu\})|_{\text{wanted}} = \alpha^N(\lambda)\prod_{i=1}^{p}\frac{1}{c(\mu_i,\lambda)}\ \Psi_X(\{\mu\})$$

Let us now evaluate $\mathrm{Tr}\ D(\lambda)\Psi_X(\{\mu\})|_{\text{wanted}}$. This term is obtained by commuting $D(\lambda)$ with the B's using only the first term in eq.(6.3). We have

$$\begin{aligned} D_0(\lambda)\Psi_X(\{\mu\}) &= \prod_{i=1}^{p}\frac{1}{c(\lambda,\mu_i)}B_1(\mu_1)\cdots B_p(\mu_p)D_0(\lambda)|1\rangle R_{01}^{(1)}(\lambda,\mu_1)\cdots R_{0p}^{(1)}(\lambda,\mu_p)X\\ &= \delta^N(\lambda)\prod_{i=1}^{p}\frac{1}{c(\lambda,\mu_i)}B_1(\mu_1)\cdots B_p(\mu_p)|1\rangle R_{01}^{(1)}(\lambda,\mu_1)\cdots R_{0p}^{(1)}(\lambda,\mu_p)X\end{aligned}$$

and taking the trace, we get

$$\text{Tr } D(\lambda)\Psi_X(\{\mu\})|_{\text{wanted}} = \delta^N(\lambda)\prod_{i=1}^{p}\frac{1}{c(\lambda,\mu_i)}B_1(\mu_1)\cdots B_p(\mu_p)|1\rangle\mathcal{T}(\lambda;\{\mu\})X$$

Adding these two contributions, we obtain exactly the wanted term and the expression for Y_0.

Unwanted terms. The unwanted term Y_1 has a contribution proportional to $\alpha^N(\mu_1)$ coming from the commutation of $A(\lambda)$, and a contribution proportionnal to $\delta^N(\mu_1)$ coming from the commutation of Tr $D(\lambda)$.

To calculate the first contribution, we commute first $A(\lambda)$ with $B(\mu_1)$ using the second term in eq.(6.2), and then we commute $A(\mu_1)$ with the other B's using only the first term in eq.(6.2). We get

$$A(\lambda)\Psi_X(\{\mu\})|_{\text{unwa.}} = -\alpha^N(\mu_1)\frac{b(\mu_1,\lambda)}{c(\mu_1,\lambda)}\prod_{i=2}^{p}\frac{1}{c(\mu_i,\mu_1)}B_1(\lambda)B_2(\mu_2)\cdots B_p(\mu_p)|1\rangle X$$

Let us now evaluate the term proportional to $\delta^N(\mu_1)$ It is obtained by commuting first $D(\lambda)$ and $B(\mu_1)$ using the second term in eq.(6.3) and then pushing $D(\mu_1)$ through the right using only the first term in eq.(6.3). We get

$$\begin{aligned} D_0(\lambda)\Psi_X(\{\mu\}) \quad = \quad & -\delta^N(\mu_1)\frac{\tilde{b}(\lambda,\mu_1)}{c(\lambda,\mu_1)}\prod_{i=2}^{p}\frac{1}{c(\mu_1,\mu_i)}B_1(\lambda)B_2(\mu_2)\cdots B_p(\mu_p)|1\rangle \\ & R_{02}^{(1)}(\mu_1,\mu_2)\cdots R_{0p}^{(1)}(\mu_1,\mu_p)P_{01}X \end{aligned}$$

taking the trace Tr_0 and identifying P_{01} with $R_{01}^{(1)}(\mu_1,\mu_1)$, we get

$$trD(\lambda)\Psi_X(\{\mu\})|_{\text{unwa.}} = -\delta^N(\mu_1)\frac{\tilde{b}(\lambda,\mu_1)}{c(\lambda,\mu_1)}\prod_{i=2}^{p}\frac{1}{c(\mu_1,\mu_i)}B_1(\lambda)B_2(\mu_2)\cdots B_p(\mu_p)|1\rangle\mathcal{T}(\mu_1;\{\mu\})X$$

Putting all this together we obtain

$$Y_1 = -\alpha^N(\mu_1)\frac{b(\mu_1,\lambda)}{c(\mu_1,\lambda)}\prod_{i=2}^{p}\frac{1}{c(\mu_i,\mu_1)}\ X - \delta^N(\mu_1)\frac{\tilde{b}(\lambda,\mu_1)}{c(\lambda,\mu_1)}\prod_{i=2}^{p}\frac{1}{c(\mu_1,\mu_i)}\mathcal{T}(\mu_1;\{\mu\})X$$

It remains to examine what happens for an unwanted term where λ takes the place of $\mu_i, i \geq 2$. It is enough to assume that it is μ_2. To compute it, we first push $B_1(\mu_1)$ in the last position using eq.(6.4). Thus, we write

$$\Psi_X(\{\mu\}) = \prod_{i\neq 1} B_2(\mu_2)\cdots B_p(\mu_p)B_1(\mu_1)|1\rangle R_{12}^{(1)}(\mu_1,\mu_2)\cdots R_{1p}^{(1)}(\mu_1,\mu_p)X$$

Using the relation $\text{Tr}_0 P_{01}M_{02} = M_{12}$ we can rewrite this formula as

$$\Psi_X(\{\mu\}) = \prod_{i\neq 1} B_2(\mu_2)\cdots B_p(\mu_p)B_1(\mu_1)|1\rangle\mathcal{T}(\mu_1;\{\mu\})X$$

Performing exactly the same analysis as before, the unwanted term takes the form

$$B_2(\lambda)B_3(\mu_3)\cdots B_p(\mu_p)B_1(\mu_1)|1\rangle Y_2$$

where

$$\begin{aligned} Y_2 &= -\Bigg[\alpha^N(\mu_2)\frac{b(\mu_2,\lambda)}{c(\mu_2,\lambda)}\prod_{i\neq 2}\frac{1}{c(\mu_i,\mu_2)}\mathcal{T}(\mu_1;\{\mu\})X \\ &\qquad +\delta^N(\mu_2)\frac{\tilde{b}(\lambda,\mu_2)}{c(\lambda,\mu_2)}\prod_{i\neq 2}\frac{1}{c(\mu_2,\mu_i)}\mathcal{T}(\mu_2;\{\mu\})\mathcal{T}(\mu_1;\{\mu\})X\Bigg] \end{aligned}$$

and since $\mathcal{T}(\mu_2;\{\mu\})$ commutes with $\mathcal{T}(\mu_1;\{\mu\})$ we can rewrite Y_2 as

$$\begin{aligned} Y_2 &= -\mathcal{T}(\mu_1;\{\mu\})\Bigg[\alpha^N(\mu_2)\frac{b(\mu_2,\lambda)}{c(\mu_2,\lambda)}\prod_{i\neq 2}\frac{1}{c(\mu_i,\mu_2)}X \\ &\qquad +\delta^N(\mu_2)\frac{\tilde{b}(\lambda,\mu_2)}{c(\lambda,\mu_2)}\prod_{i\neq 2}\frac{1}{c(\mu_2,\mu_i)}\mathcal{T}(\mu_2;\{\mu\})X\Bigg] \end{aligned}$$

Notice that when we set the unwanted terms to zero, the λ dependence drops out because

$$\frac{b(\mu_i,\lambda)}{c(\mu_i,\lambda)} = -\frac{\tilde{b}(\lambda,\mu_i)}{c(\lambda,\mu_i)}$$

■

6.5 Bethe equations.

The vector X must be a simultaneous eigenvector of the matrices $\mathcal{T}(\lambda,\{\mu\})$ for arbitrary values of λ. These equations are compatible since

$$[\mathcal{T}(\lambda,\{\mu\}),\mathcal{T}(\lambda',\{\mu\})] = 0$$

We are back to the original problem of diagonalizing a tranfer matrix for a model where the spin takes n *values instead of* $n+1$. Therefore, the solution will be obtained by repeating n-times the procedure.

At each step we introduce a set of μ's

$$\{\mu^{(m)}\} = \{\mu_1^{(m)},\cdots,\mu_{p_m}^{(m)}\} \quad m = 1,\cdots,n$$

To the set $\{\mu^{(m)}\}$ is associated the transfer matrix of an inhomogeneous vertex model

$$\mathcal{T}^{(m)}(\lambda,\{\mu^{(m)}\}) = \mathrm{Tr}_0 R_{01}^{(m)}(\lambda/\mu_1^{(m)})R_{02}^{(m)}(\lambda/\mu_2^{(m)})\cdots R_{0p_m}^{(m)}(\lambda/\mu_{p_m}^{(m)})$$

In the matrix $R^{(m)}$ the spin takes the values $m+1, \cdots, n+1$. In particular, $R^{(n)}(\lambda) = 1$ is a C-number.

The eigenvalues $\Lambda^{(m)}(\lambda, \{\mu^{(m)}\})$ of $\mathcal{T}^{(m)}(\lambda, \{\mu^{(m)}\})$ depend on all the subsequent $\{\mu^{(k)}\}, k = m+1, \cdots n$.

Proposition 37 *The sets of numbers $\{\mu^{(m)}\}, m = 1, \cdots, n$ are determined by Bethe's equations*

$$\begin{aligned}
\left(\frac{\alpha(\mu_i^{(1)})}{\delta(\mu_i^{(1)})}\right)^N \prod_{j=1}^{p_2} c(\mu_j^{(2)}, \mu_i^{(1)}) &= \prod_{j\neq i}^{p_1} \frac{c(\mu_j^{(1)}, \mu_i^{(1)})}{c(\mu_i^{(1)}, \mu_j^{(1)})} \\
&\vdots \\
\prod_{j=1}^{p_{k+1}} c(\mu_j^{(k+1)}, \mu_i^{(k)}) \prod_{j=1}^{p_{k-1}} \frac{1}{c(\mu_i^{(k)}, \mu_j^{(k-1)})} &= \prod_{j\neq i}^{p_k} \frac{c(\mu_j^{(k)}, \mu_i^{(k)})}{c(\mu_i^{(k)}, \mu_j^{(k)})} \\
&\vdots \\
\prod_{j=1}^{p_{n-1}} \frac{1}{c(\mu_i^{(n)}, \mu_j^{(n-1)})} &= \prod_{j\neq i}^{p_k} \frac{c(\mu_j^{(n)}, \mu_i^{(n)})}{c(\mu_i^{(n)}, \mu_j^{(n)})}
\end{aligned}$$

The eigenvalues are given by solving the set of recursion relations

$$\begin{aligned}
\Lambda^{(k)}(\lambda, \{\mu^{(k)}\}) &= \prod_{i=1}^{p_{k+1}} \frac{1}{c(\mu_i^{(k+1)}, \lambda)} \\
&\quad + \prod_{j=1}^{p_k} c(\lambda, \mu_j^{(k)}) \prod_{i=1}^{p_{k+1}} \frac{1}{c(\lambda, \mu_i^{(k+1)})} \Lambda^{(k+1)}(\lambda, \{\mu^{(k+1)}\}) \qquad (6.9) \\
&\vdots \\
\Lambda^{(n-1)}(\lambda, \{\mu^{(n-1)}\}) &= \prod_{i=1}^{p_n} \frac{1}{c(\mu_i^{(n)}, \lambda)} + \prod_{j=1}^{p_{n-1}} c(\lambda, \mu_j^{(n-1)}) \prod_{i=1}^{p_n} \frac{1}{c(\lambda, \mu_i^{(n)})}
\end{aligned}$$

these equations completely determine the $\Lambda^{(k)}$'s once the $\{\mu^{(k)}\}$'s are known.

Proof. Eqs.(6.9) are just coming from the expression of Y_0 eq.(6.7) in the k^{th} step and remembering that $\alpha = 1$ and $\delta^N(\lambda)$ should be replaced by $\prod_{j=1}^{p_k} c(\lambda, \mu_j^{(k)})$ since we are dealing with an inhomogeneous model. The expression for $\Lambda^{(n-1)}$ takes into account the fact that in the last step, $R^{(n)}(\lambda) = 1$ is a $\mathbb{C}$-number.

The $\{\mu^{(k)}\}$'s are solutions of Bethe's equations eq.(6.8), which in the k^{th} step read

$$\prod_{j\neq i}^{p_{k+1}} \frac{1}{c(\mu_j^{(k+1)}, \mu_i^{(k+1)})} = \prod_{j=1}^{p_k} c(\mu_i^{(k+1)}, \mu_j^{(k)}) \prod_{j\neq i}^{p_{k+1}} \frac{1}{c(\mu_i^{(k+1)}, \mu_j^{(k+1)})} \Lambda^{(k+1)}(\mu_i^{(k+1)}, \{\mu^{(k+1)}\})$$

Now, since $c(\lambda, \lambda) = 0$, from eq.(6.9), we have

$$\Lambda^{(k)}(\mu_i^{(k)}, \{\mu^{(k)}\}) = \prod_{j=1}^{p_{k+1}} \frac{1}{c(\mu_j^{(k+1)}, \mu_i^{(k)})}$$

so that Bethe's equations become

$$\prod_{j\neq i}^{p_{k+1}} \frac{1}{c(\mu_j^{(k+1)}, \mu_i^{(k+1)})} = \prod_{j=1}^{p_k} c(\mu_i^{(k+1)}, \mu_j^{(k)}) \prod_{j\neq i}^{p_{k+1}} \frac{1}{c(\mu_i^{(k+1)}, \mu_j^{(k+1)})} \prod_{j=1}^{p_{k+2}} \frac{1}{c(\mu_j^{(k+2)}, \mu_i^{(k+1)})}$$

■

Bibliography

[CNY67] C. N. Yang. *" Some exact results for the many body problem in one dimension with repulsive delta function interaction."* Phys. Rev. Lett. 19 (1967) pp. 1312-1315.

[Ch80] I. Cherednik, *"On a method of constructing factorized S matrices in elementary functions."* Theor. Math. Phys. 43 (1980) pp. 356-358.

[BVV81] O.Babelon, H.J.de Vega, C.M.Viallet *"Solutions of the factorization equations from Toda field theory"* Nuclear Physics B190 FS3 (1981) 542-552.

[PS81] J. Perk, C. L. Schultz, *"New families of commuting transfer matrices in q-state vertex models."* Phys. Lett. 84A (1981) pp.407-410.