Preprint typeset in JHEP style - HYPER VERSION

Student Seminar:
Classical and Quantum Integrable Systems

Gleb Arutyunov[a]

[a] *Institute for Theoretical Physics and Spinoza Institute, Utrecht University*
3508 TD Utrecht, The Netherlands

Abstract: The students will be guided through the world of classical and quantum integrable systems. Starting from the famous Liouville theorem and finite-dimensional integrable models, the basic aspects of integrability will be studied including elements of the modern classical and quantum soliton theory, the Riemann-Hilbert factorization problem and the Bethe ansatz.

Delivered at Utrecht University, 20 September 2006- 24 January 2007

Contents

1. Liouville Theorem **2**
- 1.1 Dynamical systems of classical mechanics 2
- 1.2 Harmonic oscillator 5
- 1.3 The Liouville theorem 7
- 1.4 Action-angle variables 9

2. Examples of integrable models solved by Liouville theorem **11**
- 2.1 Some general remarks 11
- 2.2 The Kepler two-body problem 12
  - 2.2.1 Central fields in which all bounded orbits are closed. 15
  - 2.2.2 The Kepler laws 17
- 2.3 Rigid body 20
  - 2.3.1 Moving coordinate system 20
  - 2.3.2 Rigid bodies 21
  - 2.3.3 Euler's top 23
  - 2.3.4 On the Jacobi elliptic functions 27
  - 2.3.5 Mathematical pendulum 29
- 2.4 Systems with closed trajectories 31

3. Lax pairs and classical r-matrix **32**
- 3.1 Lax representation 32
- 3.2 Lax representation with a spectral parameter 34
- 3.3 The Zakharov-Shabat construction 36

4. Two-dimensional integrable PDEs **41**
- 4.1 General remarks 42
- 4.2 Soliton solutions 43
  - 4.2.1 Korteweg-de-Vries cnoidal wave and soliton 43
  - 4.2.2 Sine-Gordon cnoidal wave and soliton 45
- 4.3 Zero-curvature representation 47
- 4.4 Local integrals of motion 49

5. Quantum Integrable Systems **55**
- 5.1 Coordinate Bethe Ansatz (CBA) 56
- 5.2 Algebraic Bethe Ansatz 68
- 5.3 Nested Bethe Ansatz (to be written) 79

6. Introduction to Lie groups and Lie algebras **80**

7. Homework exercises **95**
7.1 Seminar 1 95
7.2 Seminar 2 96
7.3 Seminar 3 97
7.4 Seminar 4 98
7.5 Seminar 5 100
7.6 Seminar 6 102
7.7 Seminar 7 105
7.8 Seminar 8 107

---

1. Liouville Theorem

1.1 Dynamical systems of classical mechanics

To motivate the basic notions of the theory of Hamiltonian dynamical systems consider a simple example.

Let a point particle with mass m move in a potential $U(q)$, where $q = (q^1, \dots q^n)$ is a vector of n-dimensional space. The motion of the particle is described by the Newton equations

$$m\ddot{q}^i = -\frac{\partial U}{\partial q^i}$$

Introduce the momentum $p = (p_1, \dots, p_n)$, where $p_i = m\dot{q}^i$ and introduce the energy which is also know as the *Hamiltonian* of the system

$$H = \frac{1}{2m}p^2 + U(q)\,.$$

Energy is a conserved quantity, i.e. it does not depend on time,

$$\frac{dH}{dt} = \frac{1}{m}p_i\dot{p}_i + \dot{q}^i\frac{\partial U}{\partial q^i} = \frac{1}{m}m^2\dot{q}_i\ddot{q}_i + \dot{q}^i\frac{\partial U}{\partial q^i} = 0$$

due to the Newton equations of motion.

Having the Hamiltonian the Newton equations can be rewritten in the form

$$\dot{q}^j = \frac{\partial H}{\partial p_j}\,, \qquad \dot{p}_j = -\frac{\partial H}{\partial q^j}\,.$$

These are the fundamental Hamiltonian equations of motion. Their importance lies in the fact that they are valid for arbitrary dependence of $H \equiv H(p, q)$ on the dynamical variables p and q.

The last two equations can be rewritten in terms of the single equation. Introduce two $2n$-dimensional vectors

$$x = \begin{pmatrix} p \\ q \end{pmatrix}, \qquad \nabla H = \begin{pmatrix} \frac{\partial H}{\partial p_j} \\ \frac{\partial H}{\partial q^j} \end{pmatrix}$$

and $2n \times 2n$ matrix J:

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

Then the Hamiltonian equations can be written in the form

$$\dot{x} = J \cdot \nabla H\,, \qquad \text{or} \qquad J \cdot \dot{x} = -\nabla H\,.$$

In this form the Hamiltonian equations were written for the first time by Lagrange in 1808.

Vector $x = (x^1, \ldots, x^{2n})$ defines a state of a system in classical mechanics. The set of all these vectors form a *phase space* $M = \{x\}$ of the system which in the present case is just the $2n$-dimensional Euclidean space with the metric $(x, y) = \sum_{i=1}^{2n} x^i y^i$.

The matrix J serves to define the so-called *Poisson brackets* on the space $\mathcal{F}(M)$ of differentiable functions on M:

$$\{F, G\}(x) = (\nabla F, J\nabla G) = J^{ij}\partial_i F \partial_j G = \sum_{j=1}^{n} \Big(\frac{\partial F}{\partial p_j}\frac{\partial G}{\partial q^j} - \frac{\partial F}{\partial q^j}\frac{\partial G}{\partial p_j} \Big)\,.$$

Problem. *Check that the Poisson bracket satisfies the following conditions*

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

for arbitrary functions F, G, H.

Thus, the Poisson bracket introduces on $\mathcal{F}(M)$ the structure of an infinite-dimensional Lie algebra. The bracket also satisfies the Leibnitz rule

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

and, therefore, it is completely determined by its values on the basis elements x^i:

$$\{x^j, x^k\} = J^{jk}$$

which can be written as follows

$$\{q_i, q_j\} = 0\,, \quad \{p^i, p^j\} = 0\,, \quad \{p^i, q_j\} = \delta^i_j\,.$$

The Hamiltonian equations can be now rephrased in the form

$$\dot{x}^j = \{H, x^j\} \qquad \Leftrightarrow \qquad \dot{x} = \{H, x\} = X_H\,.$$

A Hamiltonian system is characterized by a triple $(M, \{,\}, H)$: a phase space M, a Poisson structure $\{,\}$ and by a Hamiltonian function H. The vector field X_H is called the *Hamiltonian vector field* corresponding to the Hamiltonian H. For any function $F = F(p, q)$ on phase space, the evolution equations take the form

$$\frac{dF}{dt} = \{H, F\}$$

Again we conclude from here that the Hamiltonian H is a time-conserved quantity

$$\frac{dH}{dt} = \{H, H\} = 0\,.$$

Thus, the motion of the system takes place on the subvariety of phase space defined by $H = E$ constant.

In the case under consideration the matrix J is non-degenerate so that there exist the inverse

$$J^{-1} = -J$$

which defines a skew-symmetric bilinear form ω on phase space

$$\omega(x, y) = (x, J^{-1}y)\,.$$

In the coordinates we consider it can be written in the form

$$\omega = \sum_j dp_j \wedge dq^j\,.$$

This form is closed, i.e. $d\omega = 0$.

A non-degenerate closed two-form is called symplectic and a manifold endowed with such a form is called a symplectic manifold. Thus, the phase space we consider is the symplectic manifold.

Imagine we make a change of variables $y^j = f^j(x^k)$. Then

$$\dot{y}^j = \underbrace{\frac{\partial y^j}{\partial x^k}}_{A^j_k}\dot{x}^k = A^j_k J^{km}\nabla^x_m H = A^j_k J^{km}\frac{\partial y^p}{\partial x^m}\nabla^y_p H$$

or in the matrix form

$$\dot{y} = AJA^t \cdot \nabla_y H\,.$$

The new equations for y are Hamiltonian if and only if

$$AJA^t = J$$

and the new Hamiltonian is $\tilde{H}(y) = H(x(y))$.

Transformation of the phase space which satisfies the condition

$$AJA^t = J$$

is called canonical. In case A does not depend on x the set of all such matrices form a Lie group known as the real symplectic group $\mathrm{Sp}(2n, \mathbb{R})$. The term "symplectic group" was introduced by Herman Weyl. The geometry of the phase space which is invariant under the action of the symplectic group is called *symplectic geometry.* Symplectic (or canonical) transformations do not change the symplectic form ω:

$$\omega(Ax, Ay) = -(Ax, JAy) = -(x, A^t JAy) = -(x, Jy) = \omega(x, y)\,.$$

In the case we considered the phase space was Euclidean: $M = \mathbb{R}^{2n}$. This is not always so. The generic situation is that the phase space is a manifold. Consideration of systems with general phase spaces is very important for understanding the structure of the Hamiltonian dynamics.

1.2 Harmonic oscillator

Historically it is proved to be difficult to find a dynamical system such that the Hamiltonian equations could be solved exactly. However, there is a general framework where the explicit solutions of the Hamiltonian equations can be constructed. This construction involves

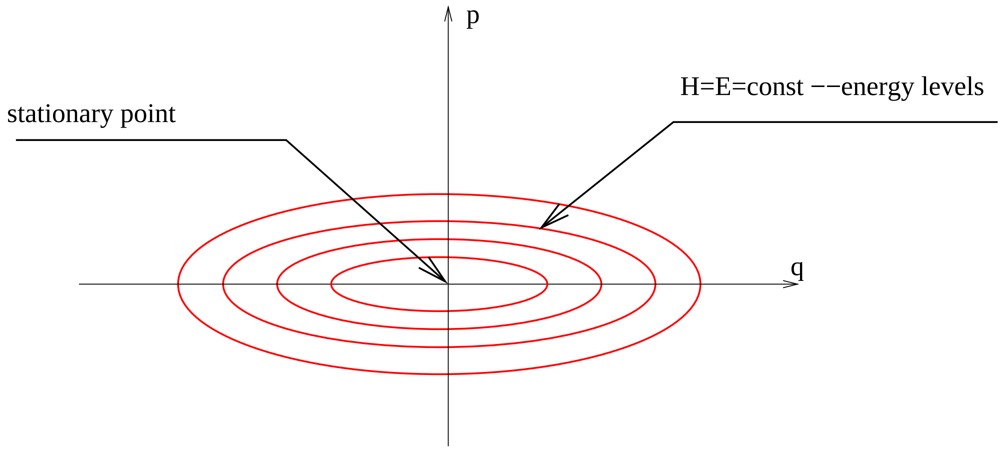
- solving a finite number of algebraic equations
- computing finite number of integrals.

If this is the way to find a solution then one says it is obtained by *quadratures.* The dynamical systems which can be solved by quadratures constitute a special class which is known as the *Liouville integrable systems* because they satisfy the requirements of the famous Liouville theorem. The Liouville theorem essentially states that if for a dynamical system defined on the phase space of dimension $2n$ one finds n independent functions F_i which Poisson commute with each other: $\{F_i, F_j\} = 0$ then this system van be solved by quadratures.

To get more insight on the Liouville theorem let us consider the simplest example – harmonic oscillator. The phase space has dimension 2 and the Hamiltonian is

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

while the Poisson bracket is $\{p, q\} = 1$. Energy is conserved, therefore, the phase space is fibred into ellipses $H = E$.



HARMONIC OSCILLATOR −− PROTOTYPE OF LIOUVILLE INTEBRABLE SYSTEM

Problem. Rewrite the Poisson bracket $\{p, q\} = 1$ and the Hamiltonian in the new coordinate system

$$p = \rho\cos(\theta)\,, \qquad q = \frac{\rho}{\omega}\sin(\theta)\,.$$

The answer is

$$\{\rho, \theta\} = \frac{\omega}{\rho}\,.$$

The hamiltonian is

$$H = \frac{1}{2}\rho^2 \to \rho = \sqrt{2H}\,.$$

We see that ρ is an integral of motion. Equation for θ:

$$\dot{\theta} = \{H, \theta\} = \rho\{\rho, \theta\} = \omega \qquad \Rightarrow \qquad \theta(t) = \omega t + \theta_0.$$

This means that the flow takes place on the ellipsis with the fixed value of ρ.

Generalization to n harmonic oscillators is easy:

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

Commuting integrals

$$F_i = \frac{1}{2}(p_i^2 + \omega_i^2 q_i^2)\,.$$

Define the common level manifold

$$M_f = \{x \in M:\ F_i = f_i, \quad i = 1, \ldots, M\}$$

This manifold is isomorphic to n-dimensional real torus which is a cartesian product of n topological circles. These tori foliate the phase space and can be parametrized with n angle variables θ_i which evolve linearly in time with frequencies ω_i. This motion is conditionally periodic: if all the periods $T_i = \frac{2\pi}{\omega_i}$ are rationally dependent:

$$\frac{T_i}{T_j} = \text{rational number}$$

the motion is periodic, otherwise the flow is dense on the torus.

1.3 The Liouville theorem

The system is Liouville integrable if it possesses n independent conserved quantities F_i, $i = 1, \ldots, n$, $\{H, F_i\}$ which are in involution

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

The Liouville theorem. Suppose that we are given n functions in involution on a symplectic $2n$-dimensional manifold

$$F_1, \ldots, F_n\,, \qquad \{F_i, F_j\} = 0\,.$$

Consider a level set of the functions F_i:

$$M_f = \{x \in M:\ F_i = f_i, \quad i = 1, \ldots, n\}$$

Assume that the n functions F_i are independent on M_f. In other words, the n-forms dF_i are linearly independent at each point of M_f. Then

1. M_f is a smooth manifold, invariant under the flow with $H = H(F_i)$.

2. If the manifold M is compact and connected then it is diffeomorphic to the n-dimensional torus
$$T^n = \{(\psi_1, \ldots, \psi_n) \bmod 2\pi\}$$

3. The phase flow with the Hamiltonian function H determines a conditionally periodic motion on M_f, i.e. in angular variables
$$\frac{d\psi_i}{dt} = \omega_i\,, \qquad \omega_i = \omega_i(F_j)\,.$$

4. The equations of motion with Hamiltonian H can be integrated by quadratures.

Let us outline the proof. Consider the level set of the integrals

$$M_f = \{x \in M : \; F_i = f_i, \;\; i = 1, \ldots, M\} \,.$$

By assumptions, the n one-forms dF_i are linearly independent at each point of M_f; by the implicit function theorem, M_f is an n-dimensional submanifold on the 2n-dimensional phase space M. Moreover, the n linearly-independent vector fields

$$\xi_{F_i} = \{F_i, \ldots\}$$

are tangent to M_f and commute with each other.

Let $\alpha = \sum_i p_i dq_i$ be the canonical 1-form and $\omega = d\alpha = \sum_i dp_i \wedge dq_i$ is the symplectic form on the phase space M. Consider a canonical transformation

$$(p_i, q_i) \rightarrow (F_i, \psi_i)$$

i.e.

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

such that F_i are treated as the new momenta. If we found this transformation then equations of motion read as

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

Thus, ω_i are constant in time. In these coordinates equations of motion are solved trivially

$$F_j(t) = F_j(0) \,, \qquad \psi_j(t) = \psi_j(0) + t\omega_j \,.$$

Thus, we see that the basic problem is to construct a canonical transformation $(p_i, q_i) \rightarrow (F_i, \psi_i)$. This is usually done with the help of the so-called generating function S. Consider M_f: $F_i(p, q) = f_i$ and solve for p_i: $p_i = p_i(f, q)$. Consider the function

$$S(f, q) = \int_{m_0}^{m} \alpha = \int_{q_0}^{q} \sum_i p_i(f, \tilde{q}) d\tilde{q}_i$$

We see that

$$p_j = \frac{\partial S}{\partial q_j}$$

and we further define

$$\psi_j = \frac{\partial S}{\partial f_j}$$

Thus, we have

$$dS = \frac{\partial S}{\partial q_j} dq_j + \frac{\partial S}{\partial f_j} df_j = p_j dq_j + \psi_j df_j$$

Since $d^2 S = 0$ we get

$$\sum_j dp_j \wedge dq_j = \sum_j df_j \wedge d\psi_j \,,$$

i.e. the transformation is canonical.

The next point is to show that S exists, i.e. it does not depend on the path. If we have a closed path from m_0 to m and from m to m_0 and assume that M_f does not have non-trivial cycles then by the Stokes theorem we get

$$\Delta S = \int_{m_0}^{m_0} \alpha = \int d\alpha = \int \omega = 0$$

because the form ω vanishes on M_f:

$$\omega(\xi_{F_i}, \xi_{F_j}) = \{F_i, F_j\} = 0 \,.$$

In case the manifold M_f has non-trivial cycles the situation changes and one gets the change of S given by integral of α over a cycle

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

which is a function of F_i only! This tells us that in this case the variables ψ_j are multi-valued.

Mention Darboux

1.4 Action-angle variables

As follows from the Liouville theorem under suitable assumptions of compactness and connectedness motion of a dynamical system in the $2n$-dimensional phase space happens on a n-dimensional torus T^n being a common level of n commuting integrals of motion. The torus has n fundamental cycles C_j which allow to introduce the "normalized" action variables

$$I_j = \frac{1}{2\pi} \oint_{C_j} p_i(q,f) dq_i \equiv \frac{1}{2\pi} \oint_{C_j} \alpha \,,$$

where f_i define the common level T^n of the commuting integrals F_i. The variables I_j are functions of f_i only and therefore they are constants of motion. The angle variables are introduced as independent angle coordinates on the cycles

$$\frac{1}{2\pi} \oint_{C_j} d\theta_i = \delta_{ij} \,.$$

Let us show that the variables (I_i, θ_i) are canonically conjugate. For that we need to construct a canonical transformation $(p_i, q_i) \to (I_i, \theta_i)$. Consider a generating function depending on I_i and q_i:

$$S(I,q) = \int_{m_0}^{m} \alpha = \int_{q_0}^{q} p_i(q', I) dq'_i \,.$$

We see that

$$p_j = \frac{\partial S}{\partial q_j} \qquad \Longrightarrow \qquad p = p(q, I).$$

Let us introduce

$$\theta_j = \frac{\partial S}{\partial I_j} \qquad \Longrightarrow \qquad \theta = \theta(q, I).$$

and show that θ_j are indeed coincide with the properly normalized angle variables. We have

$$\frac{1}{2\pi}\oint_{C_j} d\theta_i = \frac{1}{2\pi}\oint_{C_j} d\frac{\partial S}{\partial I_i} = \frac{\partial}{\partial I_i}\Big(\frac{1}{2\pi}\oint_{C_j} dS\Big) = \frac{\partial}{\partial I_i}\Big(\frac{1}{2\pi}\oint_{C_j} \frac{\partial S}{\partial q_k}dq_k + \underbrace{\frac{\partial S}{\partial I_k}dI_k}_{=0 \;\; \text{on} \;\; C_j}\Big)$$

$$= \frac{\partial}{\partial I_i}\Big(\frac{1}{2\pi}\oint_{C_j} \alpha\Big) = \delta_{ij} \,.$$

Furthermore,

$$dI_i \wedge d\theta_i = -d(\theta_i dI_i) = -d\Big(\frac{\partial S}{\partial I_i} dI_i\Big) = -d\Big(dS - \frac{\partial S}{\partial q_i}dq_i\Big) = d(p_i dq_i) = dp_i \wedge dq_i \,.$$

Problem. Find action-angle variables for the harmonic oscillator.
We have

$$E = \frac{1}{2}(p^2 + \omega^2 q^2) \qquad \Longrightarrow \qquad p(E,q) = \pm\sqrt{2E - \omega^2 q^2} \,.$$

and, therefore,

$$I = \frac{1}{2\pi}\oint_E dq\sqrt{2E - \omega^2 q^2} = \frac{2}{2\pi}\int_{-\frac{\sqrt{2E}}{\omega}}^{\frac{\sqrt{2E}}{\omega}} dq\sqrt{2E - \omega^2 q^2} = \frac{E}{\omega} \,.$$

The generating function of the canonical transformation reads

$$S(I,q) = \omega \int^q dx\sqrt{2I - x^2} \,,$$

while for the angle variables we obtain

$$\theta = \frac{\partial S}{\partial I} = \omega \int^q \frac{dx}{\sqrt{2I - x^2}} = \omega \arctan \frac{q}{\sqrt{2I - q^2}} \qquad \Longrightarrow \qquad q = \sqrt{2I}\sin\frac{\theta}{\omega} \,.$$

Finally, we explicitly check that the transformation to the action-angle variables is canonical

$$dp \wedge dq = \omega\Big(\frac{dI}{\sqrt{2I - q^2}} - \frac{qdq}{\sqrt{2I - q^2}}\Big) \wedge dq = \frac{\omega}{\sqrt{2I - q^2}} dI \wedge d\big(\sqrt{2I}\sin\frac{\theta}{\omega}\big) = dI \wedge d\theta \,.$$

2. Examples of integrable models solved by Liouville theorem

2.1 Some general remarks

Problem. Consider motion in the potential

$$V(q) = \frac{g^2}{\sin^2 q}\,, \qquad E > g^2\,.$$

Solve eoms and find a period of oscillations. One has

$$t - t_0 = \int_{q_0}^{q} \frac{dq}{\sqrt{2(E - \frac{g^2}{\sin^2 q})}} = -\int_{q_0}^{q} \frac{d\cos q}{\sqrt{2E}\sqrt{\frac{(E-g^2)}{E} - \cos^2 q}} = -\int_{\arccos q_0}^{\arccos q} \frac{dx}{\sqrt{2E}\sqrt{\frac{(E-g^2)}{E} - x^2}}\,.$$

Thus, motion happens on the interval $q_0 < q < \pi - q_0$ and taking $q_0 = \arcsin\sqrt{\frac{g^2}{E}}$ one gets

$$t = -\frac{1}{\sqrt{2E}}\Big(\arcsin\frac{x}{\sqrt{\frac{E-g^2}{E}}}\Big)\Big|^{x=\cos q}_{x=\sqrt{\frac{E-g^2}{E}}}$$

We see from here that

$$\cos\underbrace{\sqrt{2E}}_{\omega}t = \cos\Big(\frac{\pi}{2} - \arcsin\frac{x}{\sqrt{1-\frac{g^2}{E}}}\Big) = \frac{x}{\sqrt{1-\frac{g^2}{E}}} = \frac{1}{\sqrt{1-\frac{g^2}{E}}}\cos q\,.$$

Period is

$$T = \frac{2\pi}{\omega} = \frac{2\pi}{\sqrt{2E}}\,.$$

It does not depend on g^2!!!

Problem. Consider a one-dimensional harmonic oscillator with the frequency ω and compute the area surrounded by the phase curve corresponding to the energy E. Show that the period of motion along this phase curve is given by $T = \frac{dS}{dE}$.

A curve is an ellipsis

$$\Big(\frac{x}{a}\Big)^2 + \Big(\frac{y}{b}\Big)^2 = 1$$

with the area

$$S = 2b\int_{-a}^{a} \mathrm{d}x\sqrt{1 - x^2/a^2} = 2ba\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \mathrm{d}\phi\cos\phi\sqrt{1-\sin^2\phi} = 2ab\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \mathrm{d}\phi\cos^2\phi = \pi ab\,.$$

We have to identify $a = \rho$, $b = \frac{\rho}{\omega}$ so that

$$S = \pi ab = \pi\frac{\rho^2}{\omega} = \frac{2\pi}{\omega}E\,.$$

From here we see that

$$\frac{dS}{dE} = \frac{2\pi}{\omega} = T\,,$$

where T is a period of motion. The last expression has the same form as the first law of thermodynamics $dE = \frac{1}{T}dS$ provided that $1/T$ is the temperature (the period $\equiv$ the inverse temperature).

Problem. Let E_0 be the value of the potential at a minimum point ξ. Find the period $T_0 = \lim_{E\to E_0} T(E)$ of small oscillations in a neighborhood of the point ξ.

We have

$$H = \frac{p^2}{2} + V(x) = \frac{p^2}{2} + \underbrace{V(\xi)}_{\text{const}} + \underbrace{V'(\xi)}_{=0}(x-\xi) + \frac{1}{2}V''(\xi)(x-\xi)^2 + \cdots$$

Effectively we have motion described by the harmonic oscillator with the Hamiltonian

$$H_{\text{eff}} = \frac{p^2}{2} + \frac{1}{2}V''(\xi)q^2$$

whose frequency is $\omega = \sqrt{V''(\xi)}$. Therefore the period of small oscillations is

$$T_0 = \frac{2\pi}{\sqrt{V''(\xi)}}\,.$$

2.2 The Kepler two-body problem

Here we consider one of the historically first examples of integrable systems solved by the Liouville theorem: The Kepler two-body problem of planetary motion.

In the center of mass frame eoms are

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

In the original Kepler problem $V(r) = -\frac{k}{r}$, $k > 0$. The Hamiltonian

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

and the bracket $\{p_i, x_j\} = \delta_{ij}$.

Problem . Show that the angular momentum

$$\vec{J} = (J_1, J_2, J_3)\,, \qquad J_{ij} = x_i p_j - x_j p_i = \epsilon_{ijk} J_k$$

is conserved.

$$\dot{J}_{ij} = \dot{x}_i p_j - x_i \dot{p}_j - (i \leftrightarrow j) = p_i p_j + \frac{\partial V}{\partial r} x_i \frac{\partial r}{\partial x_j} - (i \leftrightarrow j) = \frac{\partial V}{\partial r}\Big(x_i \frac{\partial r}{\partial x_j} - x_j \frac{\partial r}{\partial x_i}\Big) = 0$$

Note that this is a consequence of the central symmetry.

Problem. Compute the Poisson brackets

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

Show that there are three commuting quantities

$$H, \quad J_3, \quad J^2 = J_1^2 + J_2^2 + J_3^2$$

Rewrite the canonical one form in the polar coordinates

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

We find

$$\alpha = \sum_i p_i dx_i = p_r dr + p_\theta d\theta + p_\phi d\phi \,,$$

where the original momenta are expressed as

$$\begin{aligned} p_1 &= \frac{1}{r}\Big(rp_r\cos\phi\sin\theta + p_\theta\cos\theta\cos\phi - p_\phi\frac{\sin\phi}{\sin\theta}\Big)\,,\\ p_2 &= \frac{1}{r}\Big(rp_r\sin\phi\sin\theta + p_\theta\cos\theta\sin\phi + p_\phi\frac{\cos\phi}{\sin\theta}\Big)\,,\\ p_3 &= p_r\cos\theta - \frac{1}{r}p_\theta\sin\theta\,. \end{aligned}$$

Conserved quantities

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

To better understand the physics we note that the motion happens in the plane orthogonal to the vector $\vec{J}$. Without loss of generality we can rotate our coordinate system such that in a new system $\vec{J}$ has only the third component: $\vec{J} = (0, 0, J_3)$. This simply accounts in putting in our previous formulae $\theta = \frac{\pi}{2}$. Then we note that

$$\dot{\phi} = \{H, \phi\} = \{\frac{p_\phi^2}{2r^2\sin^2\theta}, \phi\} = \frac{p_\phi}{r^2\sin^2\theta}$$

that for $\theta = \frac{\pi}{2}$ expresses the integral of motion p_ϕ as

$$p_\phi = r^2\dot{\phi}.$$

This is the conservation law of angular momentum discovered by Kepler through observations of the motion of Mars. The quantity $p_\phi = J$ has a simple geometric meaning. Kepler introduced the *sectorial velocity* C:

$$C = \frac{dS}{dt}\,,$$

where ΔS is an area of the infinitezimal sector swept by the radius-vector $\vec{r}$ for time Δt:

$$\Delta S = \frac{1}{2}r\cdot r\dot{\phi}\Delta t + \mathcal{O}(\Delta t^2) \approx \frac{1}{2}r^2\dot{\phi}\Delta t\,.$$

This is the (second) law discovered by Kepler: *in equal times the radius vector sweeps out equal areas, so the sectorial velocity is constant.* This is one of the formulations of the conservation law of angular momentum.[1]

[1]Some satellites have very elongated orbits. According to Kepler's law such a satellite spends most of its time in the distant part of the orbit where the velocity $\dot{\phi}$ is small.

We can now see how the solution can be found by using the general approach based on the Liouville theorem. The expressions for the momenta on the surface of constant energy and $J = J_3$ are

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

We can thus construct the generating function of the canonical transformation from from the Liouville theorem

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

and the associated angle variables

$$\psi_H = \frac{\partial S}{\partial H}\,, \qquad \psi_J = \frac{\partial S}{\partial J}$$

We have eoms

$$\dot{\psi}_H = 1\,, \qquad \dot{\psi}_J = 0\,.$$

Integrating the first one we obtain

$$\psi_H = t - t_0$$

and, therefore,

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

The equation for ψ_J gives

$$\psi_J = -\int^r \frac{J dr}{r^2\sqrt{2(H-V) - \frac{J^2}{r^2}}} + \phi = 0\,,$$

so that

$$\phi = \int^r \frac{J dr}{r^2\sqrt{2\big(E - V(r) - \frac{J^2}{2r^2}\big)}}\,.$$

Generically, equation which defines the values of r at which $\dot{r} = 0$:

$$E - V(r) - \frac{J^2}{2r^2} = 0$$

has two solutions: $r_{\min}$ and $r_{\max}$, they are called *pericentum and apocentrum* respectively[2]. When $\dot{r} = 0$, $\dot{\phi} \neq 0$. The r oscillates monotonically between $r_{\min}$ and

[2]If the earth is the center then $r_{\min}$ and $r_{\max}$ are called perigee and apogee, if the sun – perihelion and apohelion, if the moon – perilune and apolune.

$r_{\max}$ while ϕ changes monotonically. The angle between neighboring apocenter and pericenter is given by

$$\Delta\phi = \int_{\mathrm{r_{min}}}^{r_{\max}} \frac{J dr}{r^2\sqrt{2\big(E - V(r) - \frac{J^2}{2r^2}\big)}}\,.$$

Generic orbit is not closed! It is closed only if $\Delta\phi = 2\pi\frac{m}{n}$, $m, n \in \mathbb{Z}$, otherwise it is everywhere dense in the annulus. The annulus might degenerate into a circle.

2.2.1 Central fields in which all bounded orbits are closed.

Determination of a central potential for which all bounded orbits are closed is called *the I.L.F. Bertrand problem.*

There are only two cases for which bounded orbits are closed

$$\begin{aligned} V(r) &= ar^2\,, \qquad a \geq 0\,,\\ V(r) &= -\frac{k}{r}\,, \qquad k \geq 0\,. \end{aligned}$$

To show this we have to solve several problems.

Problem. Show that the angle ϕ between the pericenter and apocenter is equal to the half-period of an oscillation in the one dimensional system with potential energy $W(x) = V(J/x) + \frac{x^2}{2}$.

Substitution $r = \frac{J}{r}$ gives

$$\Delta\phi = \int_{x_{\min}}^{x_{\max}} \frac{dx}{\sqrt{2(E - W(x))}}\,.$$

Problem. Find the angle ϕ for an orbit close to the circle of radius r.

Effectively the angle ϕ is described by half-period of oscillation

$$\Delta\phi = \int_{x_{\min}}^{x_{\max}} \frac{dx}{\sqrt{2(E - W(x))}}\,.$$

We have

$$\Delta\phi = \frac{\pi}{\omega}\,, \qquad \omega = \sqrt{W''(x)}\,,$$

where $x = \frac{J}{r}$. We find

$$\begin{aligned} W'(x) &= \partial_x V(J/x) + x = -\frac{J}{x^2}V'(J/x) + x\,,\\ W''(x) &= 2\frac{J}{x^3}V'(J/x) + \frac{J^2}{x^4}V''(J/x) + 1\,. \end{aligned}$$

We have to take

$$-\frac{J}{x^2}V'(J/x) + x = 0 \quad\Longrightarrow\quad \frac{x^3}{J} = V'(J/x) \quad\Longrightarrow\quad \frac{J}{r^{3/2}} = \sqrt{V'(r)}\,.$$

Thus,

$$W''(x) = \frac{J}{x^3}\Big(3V'(r) + rV''(r)\Big) = \frac{3V'(r) + rV''(r)}{V'(r)}$$

and, therefore, the half-period is

$$\Delta\phi_{\rm circ} = \pi\sqrt{\frac{V'(r)}{3V'(r) + rV''(r)}}\,.$$

Problem. Find the potentials V for which the magnitude of $\Delta\phi_{\rm circ}$ does not depend on the radius.

We have to require

$$\Big(\frac{3V'(r) + rV''(r)}{V'(r)}\Big)' = \Big(\frac{rV''(r)}{V'(r)}\Big)' = \Big(r\big(\log V'(r)\big)'\Big)' = 0\,,$$

i.e.

$$\log V'(r) = {\rm const}\int\frac{1}{r} = s\log r + m\,, \qquad s, m = {\rm const}\,.$$

Further,

$$V'(r) = {\rm const}\; r^s\,, \qquad \Longrightarrow \qquad V(r) = ar^\alpha\,,$$

or $V(r) = b\log r$ if $s = -1$. Finally, the expression $\frac{V'(r)}{3V'(r)+rV''(r)}$ should be positive. If we take $V(r) = ar^\alpha$ we will get

$$\frac{V'(r)}{3V'(r) + rV''(r)} = \frac{\alpha}{3\alpha + \alpha(\alpha-1)} = \frac{1}{2+\alpha} > 0 \qquad \Longrightarrow \quad \alpha > -2\,.$$

Finally, we also have

$$\Delta\phi_{\rm circ} = \frac{\pi}{\sqrt{2+\alpha}}\,.$$

Here the logarithmic case correspond to $\alpha = 0$. Particular cases are $\alpha = 2$ which gives $\Delta\phi_{\rm circ} = \frac{\pi}{2}$ and $\alpha = -1$ which gives $\Delta\phi_{\rm circ} = \pi$.

Problem. Let $V(r) \to \infty$ as $r \to \infty$. Find

$$\lim_{E\to\infty}\Delta\phi_{\rm circ}(E, J)$$

Let us make a substitution $x = yx_{\rm max}$, we get

$$\Delta\phi_{\rm circ} = \int_{y_{\rm min}}^{1}\frac{dy}{\sqrt{2(Q(1) - Q(y))}}$$

where $Q(y) = \frac{y^2}{2} + \frac{1}{x_{\rm max}^2}V\Big(\frac{J}{yx_{\rm max}}\Big)$. As $E \to \infty$ we have $x_{\rm max} \to \infty$ and $y_{\rm min} \to 0$ and the second term in Q can be discarded. Thus, we get

$$\Delta\phi_{\rm circ} = \int_0^1\frac{dy}{\sqrt{1-y^2}} = \frac{\pi}{2}\,.$$

Problem. Let $V(r) = -kr^{-\beta}$, where $0 < \beta < 2$. Find

$$\lim_{E\to-0}\Delta\phi_{\rm circ}(E, J)$$

One has

$$\Delta\phi_{\rm circ} = \int_{x_{\rm min}}^{x_{\rm max}} \frac{dx}{\sqrt{2E + \frac{2k}{J^\beta}x^\beta - x^2}} \quad \overset{E\to 0}{\longrightarrow} \quad \int_{x_{\rm min}}^{x_{\rm max}} \frac{dx}{\sqrt{\frac{2k}{J^\beta}x^\beta - x^2}}$$

Rescale $x = \alpha y$ with α satisfying the relation $\frac{2k}{J^\beta}\alpha^\beta = \alpha^2$, then we get

$$\Delta\phi_{\rm circ} = \int_0^1 \frac{dy}{\sqrt{y^\beta - y^2}} = \frac{\pi}{2-\beta}\,.$$

We note that the result does not depend on J.

Now we are ready to find the potentials for which all bounded orbits are closed. If all bounded orbits are closed, then, in particular, $\Delta\phi_{\rm circ} = 2\pi\frac{m}{n} = {\rm const}$. That means that $\Delta\phi_{\rm circ}$ should not depend on the radius, which is the case for the potentials

$$V(r) = ar^\alpha\,, \quad \alpha > -2 \qquad \text{and} \qquad V(r) = b\log r\,.$$

In both cases $\Delta\phi_{\rm circ} = \frac{\pi}{\sqrt{2+\alpha}}$. If $\alpha > 0$ then $\lim_{E\to\infty}\Delta\phi_{\rm circ}(E,J) = \frac{\pi}{2}$ and therefore $\alpha = 2$. If $\alpha < 0$ then $\lim_{E\to 0}\Delta\phi_{\rm circ}(E,J) = \frac{\pi}{2+\alpha}$. Then we have an equality $\frac{\pi}{2+\alpha} = \frac{\pi}{\sqrt{2+\alpha}}$ which gives $\alpha = -1$. In the case $\alpha = 0$ we find $\Delta\phi_{\rm circ} = \frac{\pi}{\sqrt{2}}$ which is not commensurable with 2π. Therefore all bounded orbits are closed only for $V = ar^2$ and $U = -\frac{k}{r^2}$.

2.2.2 The Kepler laws

For the original Kepler problem we have

$$V(r) = -\frac{k}{r} + \frac{J^2}{2r^2}\,.$$

and

$$\phi = \int \frac{Jdr}{r^2\sqrt{2(E + \frac{k}{r} - \frac{J^2}{2r^2})}}$$

Integrating we get

$$\phi = \arccos \frac{\frac{J}{r} - \frac{k}{J}}{\sqrt{2E + \frac{k^2}{J^2}}}\,.$$

An integration constant is chosen to be zero which corresponds to the choice of an origin of reference for the angle ϕ at the pericenter. Introduce the notation

$$\frac{J^2}{k} = p\,, \qquad \sqrt{1 + \frac{2EJ^2}{k^2}} = e\,,$$

This leads to

$$r = \frac{p}{1 + e\cos\phi}$$

This is the so-called *focal equation of a conic section*. When $e < 1$, i.e. $E < 0$, the conic section is an ellipse. The number p is called a parameter of the ellipse and e the *essentricity*. The motion is bounded for $E < 0$.

The semi-axis a is determined as

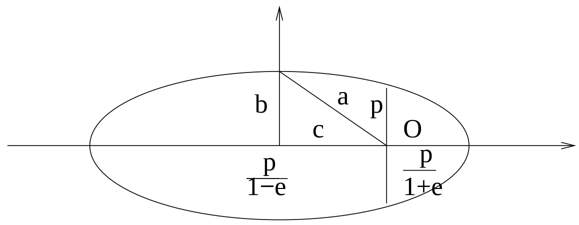
$$2a = \frac{p}{1-e} + \frac{p}{1+e} = \frac{2p}{1-e^2}\,.$$

We also have

$$c = a - \frac{p}{1+e} = \frac{1}{2}\Big(\frac{p}{1-e} - \frac{p}{1+e}\Big) = \frac{ep}{1-e^2}\,.$$

Thus,

$$\frac{c}{a} = e\,.$$



Keplerian ellipse

Obviously, we have three distinguished points

$$\begin{aligned} \phi = 0: &\qquad r = \frac{p}{1+e}\,, \\ \phi = \frac{\pi}{2}: &\qquad r = p\,, \\ \phi = \pi: &\qquad r = \frac{p}{1-e}\,. \end{aligned}$$

We can now formulate the Kepler laws:

1. *The first law: Planets describe ellipses with the Sun at one focus.*
2. *The second law: The sectorial velocity is constant.*
3. *The third law: The period of revolution around an elliptical orbit depends only on the size of the major semi-axes. The squares of the revolution periods of two planets on different elliptical orbits have the same ratio as the cubes of their major semi-axes.*

Let us prove the third law. Let T be a revolutionary period and S be the area swept out by the radius vector over the period. We have

$$S = \pi ab = \pi a^2\sqrt{1-e^2} = \pi\frac{p^2}{(1-e^2)^2}\sqrt{1-e^2} = \pi\frac{p^2}{(1-e^2)^{\frac{3}{2}}} = \frac{\pi kJ}{(\sqrt{2|E|})^3}\,,$$

while

$$a = \frac{p}{1-e^2} = \frac{k}{2|E|}\,.$$

On the other hand, since the sectorial velocity C is constant we have

$$\int_0^T C = \int_0^T \mathrm{d}t\frac{dS}{dt} = S\,, \qquad \Longrightarrow \qquad CT = \frac{J}{2}T = S\,,$$

i.e.

$$T = \frac{2S}{J} = \frac{2\pi k}{(\sqrt{2|E|})^3} = \frac{2\pi}{\sqrt{k}}a^{\frac{3}{2}}\,.$$

It is interesting to note that the total energy depends only on the major semi-axis a and it is the same for the whole set of elliptical orbits from a circle of radius a to a line segment of length $2a$. The value of the second semi-axis do depend on the angular momentum.

The Runge-Lenz vector and the Liouville torus. The phase space of the motion in the central field is $\mathrm{T}^*\mathbb{R}^3$, i.e. it is six-dimensional. There are four conserved integrals: three components of the angular momentum J_i and the energy E. This shows that the motion happens on the two-dimensional manifold. In case of the bounded motion it is the two-dimensional Liouville torus. Thus, there are two frequencies associated and when they are not rationally commensurable the orbits are not closed but rather dense on the torus. For the specific Kepler motion (with any sign of k) there is one more non-trivial conserved quantity appears which is absent for a generic central potential: The Runge-Lenz vector (for definiteness we assume that $k > 0$):

$$\vec{R} = \vec{v}\times\vec{J} - k\frac{\vec{r}}{r}\,.$$

Problem. Show that the Runge-Lenz vector is conserved.

Indeed, we have

$$\dot{\vec{R}} = \dot{\vec{v}}\times \underbrace{\vec{J}}_{m\ \vec{r}\times\vec{v}} - k\frac{\vec{v}}{r} + k\frac{\vec{r}(\vec{v}\vec{r})}{r^3} = m\dot{\vec{v}}\times(\vec{r}\times\vec{v}) - k\frac{\vec{v}}{r} + k\frac{\vec{r}(\vec{v}\vec{r})}{r^3}\,.$$

On the other hand,

$$m\dot{\vec{v}} = -\frac{\partial U}{\partial r}\frac{\vec{r}}{r} = -k\frac{\vec{r}}{r^3}$$

and, therefore,

$$\dot{\vec{R}} = -k\frac{1}{r^3}\vec{r}\times(\vec{r}\times\vec{v}) - k\frac{\vec{v}}{r} + k\frac{\vec{r}(\vec{v}\vec{r})}{r^3}$$

Further one has to use the formula

$$\vec{r}\times(\vec{r}\times\vec{v}) = (\vec{v}\vec{r})\vec{r} - r^2\vec{v}$$

to show that $\dot{\vec{R}} = 0$. The last formula can be proved by noting that the vector

$$\vec{r} \times (\vec{r} \times \vec{v}) = \alpha\vec{r} + \beta\vec{v}$$

is orthogonal to $\vec{r}$. Thus, multiplying both sides by $\vec{r}$ we get

$$0 = \alpha r^2 + \beta(\vec{v}\vec{r})\,.$$

On the other hand, multiplying both sides by $\vec{v}$ we get

$$(\vec{v}, \vec{r} \times (\vec{r} \times \vec{v})) = \alpha(\vec{v}\vec{r}) + \beta v^2$$

which gives

$$\begin{aligned}(\vec{v}, \vec{r} \times (\vec{r} \times \vec{v})) &= -(\vec{r} \times \vec{v}, \vec{r} \times \vec{v} = -r^2 v^2 \sin\phi = -r^2 v^2 (1 - \cos^2\phi) \\ &= -r^2 v^2 + (\vec{v}\vec{r})^2 = \alpha(\vec{v}\vec{r}) + \beta v^2\,.\end{aligned}$$

These two equations allows one to find

$$\alpha = (\vec{v}\vec{r})\,, \qquad \beta = -r^2\,.$$

2.3 Rigid body

2.3.1 Moving coordinate system

Let K and k will be two oriented Euclidean spaces. A *motion* of K relative to k is a mapping smoothly depending on t:

$$D_t : \quad K \to k\,,$$

which preserves the metric and orientation. Every motion can be uniquely written as the composition of a rotation (D_t which maps the origin of K into the origin of k, i.e. D_t is linear mapping) and a translation C_t: $k \to k$. Let call K and k moving and stationary coordinate systems respectively. Let $q(t)$ and $Q(t)$ will be the radius-vector of a point in a stationary and moving coordinate systems respectively. Then

$$q(t) = D_t Q(t) = \underbrace{B_t Q(t)}_{\text{rotation}} + \underbrace{r(t)}_{\text{translation}}\,.$$

Differentiating we get an addition formula for velocities

$$\dot{q} = \underbrace{\dot{B}Q}_{\text{transferred rotation}} + B\dot{Q} + \dot{r}\,.$$

Suppose a point does not move w.r.t. to the moving frame, i.e. $\dot{Q} = 0$ and also that $r = \dot{r} = 0$. Then

$$\dot{q} = \dot{B}Q = \dot{B}B^{-1}q = Aq\,,$$

where $A: \; k \to k$ is a linear operator on k. Since B is a rotation, it is an orthogonal transformation: $BB^t = 1$. Differentiating w.r.t to t we get

$$\dot{B}B^t + B\dot{B}^t = 0 \quad \Longrightarrow \quad \dot{B}B^{-1} + (\dot{B}B^{-1})^t = 0\,,$$

i.e. A is skew-symmetric. On the other hand, every skew-symmetric operator from $\mathbb{R}^3$ to $\mathbb{R}^3$ is the operator of vector multiplication by a fixed vector ω:

$$\dot{q} = \omega \times q\,.$$

Generically ω depends on t. Thus, in the case of purely rotational motion with $\dot{Q} \neq 0$ we will have

$$\dot{q} = \omega \times q + B\dot{Q} = \underbrace{\omega \times q}_{\text{transferred velocity}} + \underbrace{v'}_{\text{relative velocity}}\,.$$

2.3.2 Rigid bodies

A rigid body is a system of point masses, constrained by holonomic relations expressed by the fact that the distance between points is constant

$$|x_i - x_j| = r_{ij} = \text{const}\,.$$

If a rigid body moves freely then its center of mass moves uniformly and linearly. A rigid body rotates about its center of mass as if the center of mass were fixed at a stationary point O. In this way the problem is reduced to a problem with *three degrees of freedom* – motion of a rigid body around a fixed point O. The problem of rotation of rigid body can be studied in more generality without assuming that the fixed point coincides with the center of mass of a body. Since the Lagrangian function is invariant under all rotations around O by Noether theorem the components of the angular momentum M are conserved: $\dot{M} = 0$. The total energy which is equal to the kinetic energy is also conserved. Thus, we see that

In the problem of motion of rigid body around a fixed point, in the absence of outside forces, there are four integrals of motion: three components on M and the energy. Thus, motion happens on a two-dimensional space inside the six-dimensional phase-space (three rotation angles plus three velocities) :

$$M_f = \{M_x = f_1\,, \quad M_y = f_2\,, \quad M_z = f_3\,, \quad E = f_4 > 0\,\}.$$

The phase space is a cotangent bundle to SO(3). The manifold M_f is invariant: if the initial conditions of motion give a point on M_f then for all time of motion the point in TSO(3) corresponding to the position and velocity of the body remains in M_f. The two-dimensional manifold M_f admits a globally defined vector field (this is the field of velocities of the motion on TSO(3)), it is orientable and compact (E

is the bounded kinetic energy). According to the known theorem in topology, a two-dimensional compact orientable manifold admitting globally defined vector field is isomorphic to a torus. This is our Liouville torus[3]. According to the Liouville theorem motion on the torus will be characterized by two frequencies ω_1 and ω_2. If their ratio is not a rational number then the body never returns to its original state of motion.

Consider a rigid body rotation around a fixed point O and denote by K a coordinate system rotating with the body around O: in K the body is at rest. Every vector in K is carried to k by an operator B. By definition of the angular momentum we have

$$M = q \times m\dot{q} = m\ q \times (\omega \times q)\,.$$

Denote by J and by Ω the angular momentum and angular velocity in the moving frame K. We have

$$J = m\ Q \times (\Omega \times Q)\,.$$

This defines a linear map A: $K \to K$ such that $A\Omega = J$. This operator is symmetric:

$$(AX, Y) = (m\ Q \times (X \times Q), Y) = m(Q \times X, Q \times Y)$$

because the r.h.s. is symmetric function of X, Y. The operator A is called *the inertia tensor*. We see that taking $X = Y = \Omega$ we get

$$E = T = \frac{1}{2}(A\Omega, \Omega) = \frac{1}{2}(J, \Omega) = \frac{m}{2}(Q \times \Omega, Q \times \Omega) = \frac{m}{2}\dot{Q}^2 = \frac{m}{2}\dot{q}^2\,.$$

being a symmetric operator A is diagonalizable and it defines three mutually orthogonal characteristic directions. In the basis where A is diagonal the inertia operator and the kinetic energy take a very simple form

$$\begin{aligned} J_i &= I_i\Omega_i\,, \\ T &= \frac{1}{2}\sum_{i=1}^{3} I_i\Omega_i^2\,. \end{aligned}$$

The axes of this particular coordinate system are called *the principle inertia axes.*

Problem. Rewrite expression for energy via the quantities of the stationary frame k.

We have

$$\begin{aligned} E = \frac{1}{2}(A\Omega, \Omega) = \frac{1}{2}(J, \Omega) &= \frac{1}{2}(M, \omega) = \frac{m}{2}(q \times (q \times \omega), \omega) = \frac{m}{2}(q \times \omega, q \times \omega) \\ &= \frac{1}{2}m\Big(\omega^2 q^2 - (\omega q)^2\Big) = \frac{1}{2}\omega_i\omega_j\ \underbrace{m\Big(x_i^2\delta_{ij} - x_i x_j\Big)}_{\text{inertia tensor}}\,. \end{aligned}$$

[3]We cannot use the Liouville theorem to derive this result, because the integrals M_i do not commute with each other and, therefore, the Frobenious theorem cannot be applied to deduce that the level set is a smooth manifold. Nevertheless we can identify the Liouville torus by different means.

2.3.3 Euler's top

Consider the motion of a rigid body around a fixed point O. Let J and Ω will be the vector of angular momentum and the angular momentum *in the body*, i.e. in the moving coordinate system K. We have $A\Omega = J$, where A is the inertia tensor. The angular momentum $M = B_t J$ of the body in space is preserved. Thus, we have

$$0 = \dot{M} = \dot{B}J + B\dot{J} = \dot{B}B^{-1}M + B\dot{J} = \omega \times M + B\dot{J} = B\Big(\Omega \times J + \dot{J}\Big) .$$

From here we find

$$\frac{dJ}{dt} = J \times \Omega = J \times A^{-1} J .$$

These are the famous *Euler equations* which describe the motion of the angular momentum insider the rigid body. If one takes the coordinate adjusted to the principle axes then one gets the following system of equations

$$\frac{dJ_1}{dt} = a_1 J_2 J_3 ,$$
$$\frac{dJ_2}{dt} = a_2 J_3 J_1 ,$$
$$\frac{dJ_3}{dt} = a_3 J_1 J_2 .$$

Here

$$a_1 = \frac{I_2 - I_3}{I_2 I_3} , \qquad a_2 = \frac{I_3 - I_1}{I_1 I_3} , \qquad a_3 = \frac{I_1 - I_2}{I_1 I_2} .$$

In this way the Euler equations can be viewed as equations for the components of the angular momentum insider the body.

Consider the energy

$$H = \frac{1}{2}(J, A^{-1}J) = \frac{1}{2}\sum_{i=1}^{3} \frac{J_i^2}{I_i} .$$

It is easy to verify explicitly that it is conserved due to eoms:

$$\dot{H} = \sum_{i=1}^{3} J_i \frac{J_i}{I_i} = J_1 J_2 J_3 \Big(\frac{a_1}{I_1} + \frac{a_2}{I_2} + \frac{a_3}{I_3}\Big) = 0 .$$

Verify the conservation of the length of the angular momentum

$$\dot{J}^2 = \sum_{i=1}^{3} J_i \dot{J}_i = J_1 J_2 J_3 \Big(a_1 + a_2 + a_3\Big) = 0 .$$

This is of course agrees with the fact that M is conserved and that $M^2 = J^2$. Thus, we have proved that the Euler equations have two quadratic integrals: the energy and $M^2 = J^2$. Thus, J lies on the intersection of an ellipsoid and a sphere:

$$2E = \frac{J_1^2}{I_1} + \frac{J_2^2}{I_2} + \frac{J_3^2}{I_3} , \qquad J^2 = J_1^2 + J_2^2 + J_3^2 .$$

One can study the structure of the curves of intersection by fixing the ellipsoid $E > 0$ and changing the radius J of the sphere.

Note that alternatively the Euler equations can be rewritten as the equations for the angular velocity Ω:

$$\begin{aligned}
\frac{d\Omega_1}{dt} + \frac{I_3 - I_2}{I_1}\Omega_2\Omega_3 &= 0\,, \\
\frac{d\Omega_2}{dt} + \frac{I_1 - I_3}{I_2}\Omega_3\Omega_1 &= 0\,, \\
\frac{d\Omega_3}{dt} + \frac{I_2 - I_1}{I_3}\Omega_1\Omega_2 &= 0.
\end{aligned}$$

We could express Ω_1 and Ω_3 from the conservation laws

$$\begin{aligned}
\Omega_1^2 &= \frac{1}{I_1(I_3 - I_1)}\Big((2EI_3 - J^2) - I_2(I_3 - I_2)\Omega_2^2\Big)\,, \\
\Omega_3^2 &= \frac{1}{I_3(I_3 - I_1)}\Big((J^2 - 2EI_1) - I_2(I_2 - I_1)\Omega_2^2\Big)\,.
\end{aligned}$$

Then plugging this into the Euler equation for Ω_2 we obtain

$$\frac{d\Omega_2}{dt} = \frac{1}{I_2\sqrt{I_1 I_3}}\sqrt{\Big((2EI_3 - J^2) - I_2(I_3 - I_2)\Omega_2^2\Big)\Big((J^2 - 2EI_1) - I_2(I_2 - I_1)\Omega_2^2\Big)}\,.$$

We assume that $I_3 > I_2 > I_1$ and further that $M^2 > 2EI_2$. Then making the substitutions

$$\tau = t\sqrt{\frac{(I_3 - I_2)(J^2 - 2EI_1)}{I_1 I_2 I_3}}\,, \qquad s = \Omega_2\sqrt{\frac{I_2(I_3 - I_2)}{2EI_3 - J^2}}$$

and introducing the positive parameter $k^2 < 1$ by

$$k^2 = \frac{(I_2 - I_1)(2EI_3 - J^2)}{(I_3 - I_2)(J^2 - 2EI_1)}$$

we obtain

$$\tau = \int_0^s \frac{\mathrm{d}s}{\sqrt{(1 - s^2)(1 - k^2 s^2)}}\,.$$

The initial time $\tau = 0$ is chosen such that for $s = 0$ one has $\Omega_2 = 0$. Inverting the last integral one gets the Jacobi elliptic function[4]

$$s = \operatorname{sn}\tau\,.$$

Using two other elliptic functions

$$\operatorname{cn}^2\tau + \operatorname{sn}^2\tau = 1\,, \qquad \operatorname{dn}^2\tau + k^2\operatorname{sn}^2\tau = 1$$

[4]Elliptic functions were first applied to this problem in Rueb, Specimen inaugural, Utrecht, 1834.

we obtain the solution

$$\Omega_1 = \sqrt{\frac{2EI_3 - J^2}{I_1(I_3 - I_1)}}\,\mathrm{cn}\,\tau\,,$$
$$\Omega_2 = \sqrt{\frac{2EI_3 - J^2}{I_2(I_3 - I_1)}}\,\mathrm{sn}\,\tau\,,$$
$$\Omega_3 = \sqrt{\frac{J^2 - 2EI_1}{I_3(I_3 - I_1)}}\,\mathrm{dn}\,\tau\,.$$

Period of all these three elliptic functions is given by 4K, where K is the *complete elliptic integral of the first kind*:

$$\mathrm{K} = \int_0^1 \frac{\mathrm{d}s}{\sqrt{(1-s^2)(1-k^2s^2)}}\,.$$

Period in time t is therefore given by

$$T = 4\mathrm{K}\sqrt{\frac{I_1 I_2 I_3}{(I_3 - I_2)(J^2 - 2EI_1)}}\,.$$

After this time both Ω and J will return to their original values. Thus, Ω or J perform a strictly periodic motion. What is remarkable, is that the top itself *does not return in its original position in the stationary coordinate system k.*

We have obtained that the angular momentum J moves periodically with the period T. On the other hand, we know that the Liouville torus has the dimension two! This means that the actual motion of the body should be parameterized by two frequencies $\omega_{1,2}$. Let us express the angular velocity Ω via the Euler angles and their derivatives. Let x_1, x_2, x_3 be the axes of the moving frame k. Components of $\dot{\theta}$ on x_1 are

$$\dot{\theta}_1 = \dot{\theta}\cos\psi\,, \quad \dot{\theta}_2 = -\dot{\theta}\sin\psi\,, \quad \dot{\theta}_3 = 0\,.$$

The velocity $\dot{\phi}$ is directed along Z. Its projections are

$$\dot{\phi}_1 = \dot{\phi}\sin\theta\sin\psi\,, \quad \dot{\phi}_2 = \dot{\phi}\sin\theta\cos\psi\,, \quad \dot{\phi}_3 = \dot{\phi}\cos\theta\,.$$

Finally, the velocity $\dot{\psi}$ is directed along x_3. Thus, we can write the components of the angular velocity in the moving frame as

$$\Omega_1 = \dot{\phi}\sin\theta\sin\psi + \dot{\theta}\cos\psi$$
$$\Omega_2 = \dot{\phi}\sin\theta\cos\psi - \dot{\theta}\sin\psi$$
$$\Omega_3 = \dot{\phi}\cos\theta + \dot{\psi}\,.$$

Substituting these formula into the expression for the kinetic energy $T = \frac{1}{2}I_i\Omega_i^2$ we obtain the kinetic energy in terms of the Euler angles.

Problem. By using Euler angles relate the angular momenta in the moving and the stationary coordinate systems. The momentum M is directed along the Z axis of the stationary coordinate system.

We have

$$\begin{aligned} M\sin\theta\sin\psi &= I_1\Omega_1\,, \\ M\sin\theta\cos\psi &= I_2\Omega_2\,, \\ M\cos\theta &= I_3\Omega_3\,. \end{aligned}$$

From here

$$\cos\theta = \frac{I_3\Omega_3}{M}\,, \qquad \tan\psi = \frac{I_1\Omega_1}{I_2\Omega_2}\,.$$

Solution of the last problem allows one to find

$$\begin{aligned} \cos\theta &= \sqrt{\frac{I_3(M^2-2EI_1)}{M^2(I_3-I_1)}}\,\mathrm{dn}\,\tau\,, \\ \tan\psi &= \sqrt{\frac{I_1(I_3-I_2)}{I_2(I_3-I_1)}}\frac{\mathrm{cn}\,\tau}{\mathrm{sn}\,\tau}\,. \end{aligned}$$

Thus, both angles θ and ψ are periodic functions of time with the period T (the same period as for Ω!). However, the angle ϕ *does not appear* in the formulas relating the angular momenta in the moving and the stationary coordinate systems. We can find it from

$$\begin{aligned} \Omega_1 &= \dot\phi\sin\theta\sin\psi + \dot\theta\cos\psi \\ \Omega_2 &= \dot\phi\sin\theta\cos\psi - \dot\theta\sin\psi\,. \end{aligned}$$

Solving we get

$$\dot\phi = \frac{\Omega_1\sin\psi + \Omega_2\cos\psi}{\sin\theta}\,.$$

This leads to the differential equation

$$\frac{d\phi}{dt} = M\frac{I_1\Omega_2^2 + I_2\Omega_2^2}{I_1^2\Omega_1^2 + I_2^2\Omega_2^2}\,.$$

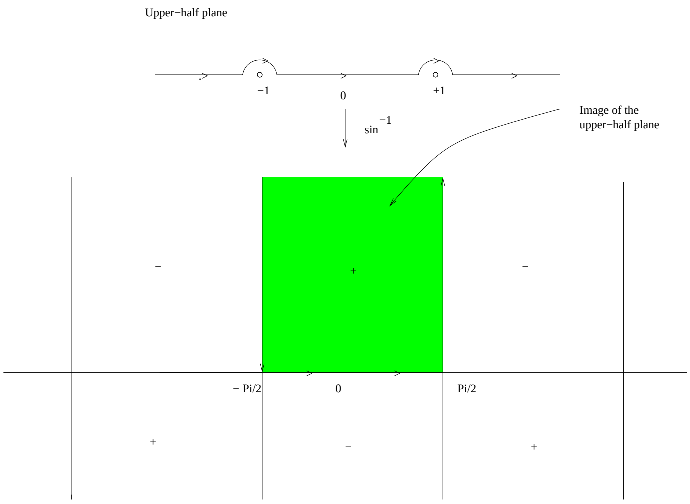
Thus, solution is given by quadrature but the integrand contains elliptic functions in a complicated way. One can show that the period of ϕ, which is T' is not comparable with T. This leads to the fact that the top never returns to its original state. The periods T and T' are the periods of motion over the Liouville torus.

2.3.4 On the Jacobi elliptic functions

Consider a *trigonometric integral*

$$y = \sin^{-1} x = \int_0^x \frac{\mathrm{d}y}{\sqrt{1-y^2}} = \int_0^{\arcsin\phi} \frac{\mathrm{d}\sin\phi}{\sqrt{1-\sin^2\phi}}.$$

If $-1 \le \mathrm{Re} x \le 1$ this integral coincides with the function $y = \arcsin x$.



This integral maps the punctured (at ± 1)upper-half plane one-to-one onto the shaded strip. The integral is inverted by the function sin which is period with the period

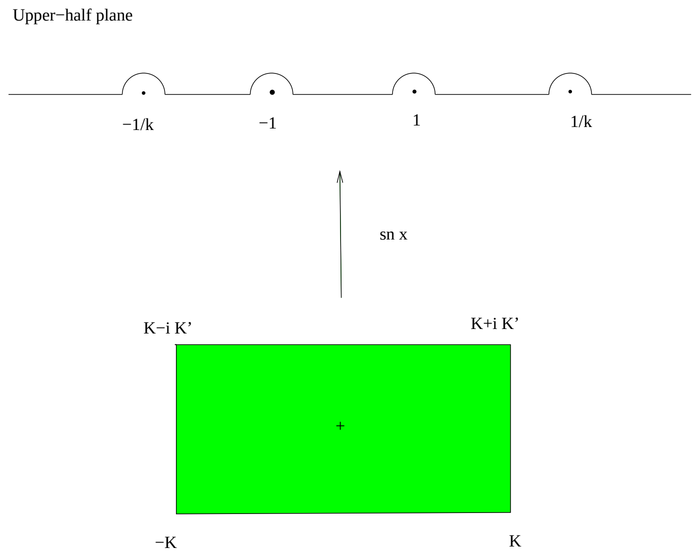
$$2\pi = 4 \times \text{complete integral} \int_0^1 \frac{\mathrm{d}y}{\sqrt{1-y^2}}.$$

Thus, sin can be viewed as the function on the complex cylinder $\mathbf{X} = \mathbb{C}/\mathbb{L}$ with $\mathbb{L} = 2\pi\mathbb{Z}$. It also gives a Riemann map of the strip $|x| < \frac{\pi}{2}$, $y > 0$ to the upper-half plane, standardized by the values $0, 1, \infty$ at $0, \frac{\pi}{2}, i\infty$.

It is remarkable discovery of Gauss and Abel that the same picture holds for the incomplete integral of the first kind:

$$x \to \int_0^x \frac{\mathrm{d}y}{\sqrt{(1-y^2)(1-k^2y^2)}}.$$

The case $k = 0$ is trigonometric we have discussed above. The novel point is that for $k^2 \ne 0, 1$ the inversion of the integral now leads to an elliptic function, that is a single-valued function having not just one but two independent complex periods.



The rectangle region is mapped by the Jacobi function $\operatorname{sn} x$ one-to-one onto the upper-half-plane with four punctures.

The mapping of the upper half-plane onto the rectangle is such that the points $0, 1, 1/k, \infty, -1/k, -1$ have the images $0, \mathrm{K}, \mathrm{K} + i\mathrm{K}', i\mathrm{K}', \mathrm{K} - i\mathrm{K}', -\mathrm{K}$ respectively. The function $\operatorname{sn} x$ repeats in congruent blocks of four rectangles and, therefore, is invariant under translations by $\omega_1 = 4\mathrm{K}(k)$ and $\omega_3 = 2i\mathrm{K}'(k)$. Here K and K′ are complete elliptic integrals (K′ is called complementary)

$$\mathrm{K} = \int_0^1 \frac{\mathrm{d}y}{\sqrt{(1-y^2)(1-k^2y^2)}}$$

$$\mathrm{K}' = \int_1^{1/k} \frac{\mathrm{d}y}{\sqrt{(1-y^2)(1-k^2y^2)}} = \int_0^1 \frac{\mathrm{d}y}{\sqrt{(1-y^2)(1-k'^2y^2)}} \,,$$

where $k = \sqrt{1-k^2}$ is the complementary modulus.

Writing

$$x = \int_0^{\operatorname{sn} x} \frac{\mathrm{d}y}{\sqrt{(1-y^2)(1-k^2y^2)}}$$

and differentiating over x we will get

$$1 = \frac{\operatorname{sn}' x}{\sqrt{(1-y^2)(1-k^2y^2)}}$$

or

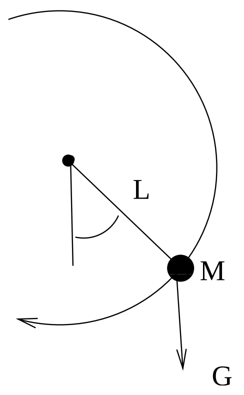
$$(\operatorname{sn}' x)^2 = (1-y^2)(1-k^2y^2)\,.$$

This is differential equation satisfied by the Jacobi elliptic function $\operatorname{sn} x$.

2.3.5 Mathematical pendulum

The theory of elliptic functions finds beautiful applications in many classical problems. One of them is the motion of the mathematical pendulum in the gravitational field of the Earth.

Consider the mathematical pendulum (of mass M) in the gravitational field of the Earth.



A pendulum in the gravitational field of the Earth. Here L is its length and G is the gravitational constant.

First we derive the eoms. The radius-vector and the velocity are is

$$\vec{r}(t) = (\underbrace{L\sin\theta}_{x}, \underbrace{L\cos\theta}_{y})\,, \qquad \vec{v}(t) = (L\cos\theta\dot{\theta}, -L\sin\theta\dot{\theta})\,.$$

Projecting the Newton equations of the axes x and y we find

$$L\frac{d^2\cos\theta}{dt^2} = mg\,, \qquad L\frac{d^2\sin\theta}{dt^2} = 0\,.$$

Differentiating we get

$$-L(\cos\theta\dot{\theta}^2 + \sin\theta\ddot{\theta}) = mg\,, \qquad -\sin\theta\dot{\theta}^2 + \cos\theta\ddot{\theta} = 0\,.$$

Excluding from these equations $\dot{\theta}^2$ we obtain the equations of motion

$$L\ddot{\theta} = -mg\sin\theta\,.$$

This equation can be integrated once by noting that

$$\frac{d\dot{\theta}^2}{dt} = 2\dot{\theta}\ddot{\theta} = 2\dot{\theta}\big(-\frac{mg}{L}\sin\theta\big) = -\frac{2mg}{L}\sin\theta\dot{\theta} = \frac{2mg}{L}\frac{d}{dt}\cos\theta\,,$$

i.e. that

$$\frac{d}{dt}\left(\dot{\theta}^2 - \frac{2mg}{L}\cos\theta\right) = 0\,,$$

Thus, the combination $\dot{\theta}^2 - \frac{2mg}{L}\cos\theta$ is an integral of motion. In fact, this is nothing else as the total energy. Indeed, the total energy is (up to an additive constant which can be always added)

$$E = \frac{m\vec{v}^2}{2} + U = \frac{mL^2}{2}\dot{\theta}^2 + mgL(1-\cos\theta)\,.$$

We rewrite the conservation law in the form

$$L^2\dot{\theta}^2 = 2gh - 4gL\sin^2\frac{\theta}{2}\,,$$

where h is an integration constant. Making the change of variables $y = \sin\frac{\theta}{2}$ we arrive at

$$\dot{y}^2 = \frac{g}{L}(1-y^2)\Big(\frac{h}{2L} - y^2\Big)\,.$$

We have now several cases to consider

- Under the oscillatory motion the point does not reach the top of a circle. This means that $\dot{y}$ terns to zero for some $y < 1$. Thus, $\frac{h}{2L} < 1$. Denoting $h = 2Lk^2$, where k is a positive constant less then one we obtain

  $$\dot{y}^2 = \frac{gk^2}{L}\left(1 - k^2\frac{y^2}{k^2}\right)\left(1 - \frac{y^2}{k^2}\right)\,.$$

  Solution to this equation is

  $$y = k\,\mathrm{sn}\left(\sqrt{\frac{g}{L}}(t-t_0), k\right)\,.$$

  The integration constants are t_0 and k, they are determined from the initial conditions. the period is $T = \sqrt{\frac{L}{g}}\mathrm{K}(k)$.

- Rotatory motion. Here $h > 2L$. Thus, taking $2L = hk^2$ we will have $k^2 < 1$. Equation becomes

  $$\dot{y}^2 = \frac{g}{Lk^2}(1-y^2)(1-k^2y^2)$$

  whose solution is

  $$y = \mathrm{sn}\left(\sqrt{\frac{g}{L}}\frac{t-t_0}{k}, k\right)\,.$$

- The point reaches the top. Here $h = 2L$ and we get

  $$\dot{y}^2 = \frac{g}{L}(1-y^2)^2 \quad\rightarrow\quad \dot{y} = \sqrt{\frac{g}{L}}(1-y^2)\,.$$

  Solution is

  $$y = \tanh\left(\sqrt{\frac{g}{L}}(t-t_0)\right)\,.$$

2.4 Systems with closed trajectories

The Liouville integrable systems of phase space dimension $2n$ are characterized by the requirement to have n globally defined integrals of motion $F_j(p,q)$ Poisson commuting with each other. Taking the level set

$$M_f = \{F_j = f_j, \quad j = 1, \dots n\}$$

we obtain (in the compact case) the n-dimensional torus. In general frequencies of motion ω_j on the Liouville torus are not rationally comparable and, as the result, the corresponding trajectories are not closed.

A special situation arises if at least two frequencies become rationally comparable. Such a motion is called *degenerate*. Here we will be interested in the situation of the completely degenerate motion, i.e. when all n frequencies ω_j are comparable. In this case the classical trajectory is a closed curve and the number of global integrals raises to $2n-1$.[5] They cannot Poisson-commute with each other because the maximal possible number of commuting integrals can be n only. Below we will give already accounted examples of degenerate motion.

Two-dimensional harmonic oscillator. The Hamiltonian is

$$H = \frac{1}{2}(p_1^2 + p_2^2) + \frac{1}{2}(\omega_1^2 q_1^2 + \omega_2^2 q_2^2)\,.$$

There are two independent and mutually commuting integrals

$$F_1 = \frac{1}{2}p_1^2 + \frac{1}{2}\omega_1^2 q_1^2\,, \qquad F_2 = \frac{1}{2}p_2^2 + \frac{1}{2}\omega_2^2 q_2^2\,,$$

such that $H = F_1 + F_2$. If the ratio ω_1/ω_2 is irrational the trajectories are everywhere dense on the Liouville torus. However, if

$$\frac{\omega_1}{\omega_2} = \frac{r}{s}\,,$$

where r, s are relatively prime integers then there is a new additional integral of motion

$$F_3 = \bar{a}_1^s a_2^r\,,$$

where

$$\bar{a}_1 = \frac{1}{\sqrt{2\omega_1}}(p_1 + i\omega_1 q_1)\,, \qquad a_2 = \frac{1}{\sqrt{2\omega_2}}(p_2 - i\omega_2 q_2)\,.$$

Indeed, we have

$$\dot{F}_3 = \bar{a}_1^{s-1} a_2^{r-1}\big(s a_2 \dot{\bar{a}}_1 + r \bar{a}_1 \dot{a}_2\big)\,.$$

[5] In quantum mechanics we have in this case the degenerate levels.

Then using the eoms $\dot{q} = p$ and $\dot{p} = -\omega^2 q$ we find

$$\dot{\bar{a}}_1 = i\omega_1 \bar{a}_1 \,, \qquad \dot{a}_2 = -i\omega_2 a_2 \,,$$

Thus,

$$\dot{F}_3 = i\bar{a}_1^s a_2^r \big(s\omega_1 - r\omega_2\big) = 0 \,.$$

This integral is homogenous function of degree $r + s$ both over the coordinates and momenta. The trajectories are closed. They are the so-called *Lissajous figures*. Find the Poisson brackets between F and $F_i = \frac{1}{2}(p_i^2 + \omega_i^2 q_i^2)$.

The Kepler problem. We know that the orbits in the Keplerian problem are closed for $E < 0$. There exists an additional conserved Runge-Lenz vector:

$$\vec{R} = \vec{v} \times \vec{J} - k\frac{\vec{r}}{r} \,.$$

This vector is othogonal to the angular momentum:

$$(\vec{J}, \vec{R}) = (\vec{J}, \vec{v} \times \vec{J}) - \frac{k}{r}(\vec{J}, \vec{r}) = 0 - 0 = 0 \,.$$

Thus, there are five independent integrals of motion in the system with six phase-space degrees of freedom. The Kepler Hamiltonian can be expressed via these five quantities. Thus, the motion is completely degenerate.

The Euler top. The phase space has dimension six. We found four globally defined conserved quantities: the Hamiltonian and three components of the angular momentum. That is the reason why the Liouville torus has dimension two instead of three. Since $6 - 4 = 2 \neq 1$ the motion is partially, but not completely degenerate.

3. Lax pairs and classical r-matrix

In this section we will study the cornerstone concepts of the modern theory of integrable systems: the Lax pairs and classical r-matrix.

3.1 Lax representation

Let L, M be two matrices which are also functions on the phase space, i.e. $L \equiv L(p, q)$ and $M = M(p, q)$, such that the Hamiltonian equations of motion can be written in the form

$$\dot{L} = [M, L] \,.$$

This is the Lax representation (the Lax pair) of the Hamiltonian equations. The importance of this representation lies in the fact that it provides a straightforward construction of the conserved quantities:

$$I_k = \mathrm{tr} L^k \,.$$

Indeed,

$$\dot{I}_k = k\mathrm{tr}(L^{k-1}\dot{L}) = k\mathrm{tr}(L^{k-1}[M,L]) = \mathrm{tr}[M,L^k] = 0\,.$$

In fact solution of the Lax equation is

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

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

$$M(t) = \dot{g}g^{-1}\,.$$

By the Newton theorem, the integrals I_k are the functions of the eigenvalues of the matrix L. The evolution of the system is called *isospectral* because the eigenvalues of the matrix L are preserved in time. A Lax pair is not uniquely defined.

Problem. Show that if g is any invertible matrix then

$$\mathbf{L} = gLg^{-1}\,, \qquad \mathbf{M} = gMg^{-1} + \dot{g}g^{-1}$$

also defines a Lax pair. We have

$$\dot{\mathbf{L}} = \dot{g}Lg^{-1} + g[M,L]g^{-1} - gLg^{-1}\dot{g}g^{-1} = [gMg^{-1} + \dot{g}g^{-1}, gLg^{-1}] = [\mathbf{M},\mathbf{L}]\,.$$

A simple example of a dynamical system which possesses the Lax pair is provided by the harmonic oscillator. One can take

$$L = \begin{pmatrix} p & \omega q \\ \omega q & -p \end{pmatrix}, \qquad M = \begin{pmatrix} 0 & -\frac{1}{2}\omega \\ \frac{1}{2}\omega & 0 \end{pmatrix}.$$

Indeed,

$$\begin{pmatrix} \dot{p} & \omega\dot{q} \\ \omega\dot{q} & -\dot{p} \end{pmatrix} = \begin{pmatrix} 0 & -\frac{1}{2}\omega \\ \frac{1}{2}\omega & 0 \end{pmatrix}\begin{pmatrix} p & \omega q \\ \omega q & -p \end{pmatrix} - \begin{pmatrix} p & \omega q \\ \omega q & -p \end{pmatrix}\begin{pmatrix} 0 & -\frac{1}{2}\omega \\ \frac{1}{2}\omega & 0 \end{pmatrix} = \begin{pmatrix} -\omega^2 q & \omega p \\ \omega p & \omega^2 q \end{pmatrix}$$

and we get the eoms of the harmonic oscillator $\dot{q} = p$ and $\dot{p} = -\omega^2 q$. The Hamiltonian is $H = \frac{1}{4}\mathrm{tr}L^2$.

Obviously the Lax representation makes no reference to the Poisson structure. We can find however the general form of the Poisson bracket between the matrix elements of L which ensures that the conserved eigenvalues of L are in involution. Suppose that L is diagonalizable

$$L = U\Lambda U^{-1}\,.$$

One has

$$\begin{aligned}
&\{L_1,L_2\} = \{U_1\Lambda_1U_1^{-1}, U_2\Lambda_2U_2^{-1}\} = \\
&= \underbrace{\{U_1,U_2\}\Lambda_1U_1^{-1}\Lambda_2U_2^{-1}} + U_1\{\Lambda_1,U_2\}U_1^{-1}\Lambda_2U_2^{-1} - \underbrace{U_1\Lambda_1U_1^{-1}\{U_1,U_2\}U_1^{-1}\Lambda_2U_2^{-1}} \\
&+ U_2\{U_1,\Lambda_2\}\Lambda_1U_1^{-1}U_2^{-1} - U_1\Lambda_1U_2U_1^{-1}\{U_1,\Lambda_2\}U_1^{-1}U_2^{-1} \\
&- \underbrace{U_2\Lambda_2U_2^{-1}\{U_1,U_2\}U_2^{-1}\Lambda_1U_1^{-1}} - U_2\Lambda_2U_2^{-1}U_1\{\Lambda_1,U_2\}U_2^{-1}U_1^{-1} + \underbrace{U_1\Lambda_1U_1^{-1}U_2\Lambda_2U_2^{-1}\{U_1,U_2\}U_1^{-1}U_2^{-1}}\,,
\end{aligned}$$

where we have assumed that the eigenvalues commute $\{\Lambda_1,\Lambda_2\}=0$. Introducing

$$k_{12}=\{U_1,U_2\}U_1^{-1}U_2^{-1}\,,\qquad q_{12}=U_2\{U_1,\Lambda_2\}U_1^{-1}U_2^{-1}\,,\qquad q_{21}=U_1\{U_2,\Lambda_1\}U_1^{-1}U_2^{-1}$$

we could write

$$\begin{aligned}\{L_1,L_2\} &= k_{12}L_1L_2+L_1L_2k_{12}-L_1k_{12}L_2-L_2k_{12}L_1\\ &\quad -q_{21}L_2+q_{12}L_1-L_1q_{12}+L_2q_{21}\,.\end{aligned}$$

This bracket can be further written as

$$\begin{aligned}\{L_1,L_2\} &= [k_{12}L_2-L_2k_{12},L_1]+[q_{12},L_1]-[q_{21},L_2]\\ &= \frac{1}{2}[[k_{12},L_2],L_1]-\frac{1}{2}[[k_{21},L_1],L_2]+[q_{12},L_1]-[q_{21},L_2]\\ &= [r_{12},L_1]-[r_{21},L_2]\,,\end{aligned}$$

where we have introduced the so-called *r-matrix*

$$r_{12}=q_{12}+\frac{1}{2}[k_{12},L_2]\,.$$

Finally, the Jacobi identity for the bracket yields the following constraint on r:

$$[L_1,[r_{12},r_{13}]+[r_{12},r_{23}]+[r_{32},r_{13}]+\{L_2,r_{13}\}-\{L_3,r_{12}\}]+\text{cycl. perm}=0$$

Solving this equation for r is equivalent to classifying integrable systems. If r is constant, i.e. independent of the dynamical variables, then only the first term is left. In particular, the Jacobi identity is satisfied if

$$[r_{12},r_{13}]+[r_{12},r_{23}]+[r_{32},r_{13}]=0\,.$$

If r-matrix here is antisymmetric: $r_{12}=-r_{21}$ then the corresponding equation is called *the classical Yang-Baxter equation.*

3.2 Lax representation with a spectral parameter

Here we introduce the Lax matrices $L(\lambda)$, $M(\lambda)$ which depend analytically on a parameter λ called a spectral parameter. We start by considering example of the Euler top. Introduce two 3×3 anti-symmetric matrices

$$J=\begin{pmatrix}0 & -J_3 & -J_2\\ -J_3 & 0 & J_1\\ J_2 & -J_1 & 0\end{pmatrix}\,,\qquad \Omega=\begin{pmatrix}0 & -\Omega_3 & -\Omega_2\\ \Omega_3 & 0 & \Omega_1\\ \Omega_2 & -\Omega_1 & 0\end{pmatrix}\,.$$

Then we can see that the Euler equations are equivalent to the following Lax representation

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

i.e. $L = J$ and $M = \Omega$. However, $\mathrm{tr} L^n$ either vanish or are functions of J^2 and, therefore, they do not contain the Hamiltonian. This can be cured by introducing the diagonal matrix $\mathcal{I}$:

$$\mathcal{I} = \begin{pmatrix} \frac{1}{2}(I_2 + I_3 - I_1) & 0 & 0 \\ 0 & \frac{1}{2}(I_1 + I_3 - I_2) & 0 \\ 0 & 0 & \frac{1}{2}(I_1 + I_2 - I_3) \end{pmatrix} .$$

One can see that

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

Assuming that all I_i are different we introduce

$$L(\lambda) = \mathcal{I}^2 + \frac{1}{\lambda} J , \qquad M(\lambda) = \lambda\mathcal{I} + \Omega .$$

Then we write the equation

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

which reduces to

$$\frac{1}{\lambda}\dot{J} = [\lambda\mathcal{I} + \Omega, \mathcal{I}^2 + \frac{1}{\lambda}J] = [\Omega, \mathcal{I}^2] + [\mathcal{I}, J] + \frac{1}{\lambda}[\Omega, J]$$

We see that

$$[\Omega, \mathcal{I}^2] + [\mathcal{I}, J] = \Omega\mathcal{I}^2 - \mathcal{I}^2\Omega + \mathcal{I}(\mathcal{I}\Omega + \Omega\mathcal{I}) - (\mathcal{I}\Omega + \Omega\mathcal{I})\mathcal{I} = 0 .$$

Thus, vanishing of the $1/\lambda$-term gives the Euler equations of motion. This Lax pair produces the Hamiltonian among the conserved quantities. We have

$$\begin{aligned} \mathrm{tr} L(\lambda)^2 &= \mathrm{tr}\mathcal{I}^4 - \frac{2}{\lambda^2} J^2 \\ \mathrm{tr} L(\lambda)^3 &= \mathrm{tr}\mathcal{I}^6 - \frac{3}{\lambda^2}\Big(\frac{1}{4}(\mathrm{tr}\mathcal{I})^2 J^2 - I_1 I_2 I_3 H\Big) . \end{aligned}$$

The Euler-Arnold equations. The three-dimensional Euler top admits natural generalization to the $\mathfrak{so}(n)$ Lie algebra. Let $\Omega \in \mathfrak{so}(n)$ and $\mathcal{I}$ is a diagonal matrix. Then

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

is also skew-symmetric matrix: $J^t = -J$. Assuming that all eigenvalues of $\mathcal{I}$ are different we introduce

$$L(\lambda) = \mathcal{I}^2 + \frac{1}{\lambda} J , \qquad M(\lambda) = \lambda\mathcal{I} + \Omega .$$

Equations

$$\dot{J} = [J, \Omega] , \qquad J = \mathcal{I}\Omega + \Omega\mathcal{I}$$

are called the *Euler-Arnold* equations. They are equivalent to the spectral-dependent Lax equations

$$\frac{d}{dt}\Big(\mathcal{I}^2+\frac{1}{\lambda}J\Big)=[\lambda\mathcal{I}+\Omega,\mathcal{I}^2+\frac{1}{\lambda}J]\,.$$

The later are known as the *Manakov equations.*

The Kepler problem. Another interesting Lax pair can be found for the Kepler problem (M.Antonowicz and S.Rauch-Wojciechowski). Introduce the following L and M matrices which depend on three different parameters $\lambda_1,\lambda_2,\lambda_3$:

$$L=\frac{1}{2}\begin{pmatrix}-\sum_{i=k}^{3}\frac{x_k\dot{x}_k}{\lambda-\lambda_k} & \sum_{i=k}^{3}\frac{x_kx_k}{\lambda-\lambda_k}\\ -\sum_{i=k}^{3}\frac{\dot{x}_k\dot{x}_k}{\lambda-\lambda_k} & \sum_{i=k}^{3}\frac{x_k\dot{x}_k}{\lambda-\lambda_k}\end{pmatrix}\,,\qquad M=\begin{pmatrix}0 & 1\\ \frac{k}{r^3} & 0\end{pmatrix}\,,$$

where $r=\sqrt{x_1^2+x_2^2+x_3^2}$ and x_k are coordinates of the particle, while $p_k=\dot{x_k}$ are the corresponding conjugate momenta. Newton's equation for x_k arises as the condition of vanishing of the residue of the pole $\lambda=\lambda_k$.

3.3 The Zakharov-Shabat construction

There is no general algorithm how to construct a Lax pair for a given integrable system. However, there is a general procedure of how to construct consistent Lax pairs giving rise to integrable systems. This is a general method how to construct the spectral dependent matrices $L(\lambda)$ and $M(\lambda)$ such that

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

are equivalent to the eoms of an integrable system.

The basic idea of the Zakharov-Shabat construction is to specify the analytic properties of the matrices $L(\lambda)$ and $M(\lambda)$ for $\lambda\in\mathbb{C}$.

Let $f(\lambda)$ be a matrix-valued function which has poles at $\lambda=\lambda_k\neq\infty$ of order n_k. We can write

$$f(\lambda)=\underbrace{f_0}_{\text{const}}+\sum_k f_k(\lambda),\qquad \underbrace{f_k(\lambda)}_{\text{polar part}}=\sum_{r=-n_k}^{-1}f_{k,r}(\lambda-\lambda_k)^r\,.$$

Around any λ_k this function can be decomposed as

$$f(\lambda)=f_+(\lambda)+f_-(\lambda)\,,$$

where $f_+(\lambda)$ is regular at $\lambda=\lambda_k$ and $f_-(\lambda)=f_k(\lambda)$ is the polar part.

Assume that $L(\lambda)$ and $M(\lambda)$ are rational functions of λ. Let $\{\lambda_k\}$ be the set of poles of $L(\lambda)$ and $M(\lambda)$. Assuming *no* poles at infinity we can write

$$L(\lambda) = L_0 + \sum_k L_k(\lambda)\,, \qquad L_k(\lambda) = \sum_{r=-n_k}^{-1} L_{k,r}(\lambda-\lambda_k)^r$$

$$M(\lambda) = M_0 + \sum_k M_k(\lambda)\,, \qquad M_k(\lambda) = \sum_{r=-m_k}^{-1} M_{k,r}(\lambda-\lambda_k)^r\,.$$

Here $L_{k,r}$ and $M_{k,r}$ are matrices and *we assume that* λ_k *do not depend on time.*

Looking at the Lax equation we see that at $\lambda=\lambda_k$ the l.h.s. has a pole of order n_k, while the r.h.s. has a potential pole of the order n_k+m_k. Hence there are two type of equations. The first type does not contain the time derivatives and comes from setting to zero the coefficients of the poles of order greater than n_k on the r.h.s. of the equation. This gives m_k constraints on the matrix 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. These are equations for the dynamical variables because they involve time derivatives.

Consider the matrix $L(\lambda)$ around $\lambda=\lambda_k$. Then the matrix $Q(\lambda)=(\lambda-\lambda_k)^{n_k}L(\lambda)$ is regular around λ_k, i.e.

$$Q(\lambda) = (\lambda-\lambda_k)^{n_k}L(\lambda) = Q_0 + (\lambda-\lambda_k)Q_1 + (\lambda-\lambda_k)^2 Q_2 + \cdots$$

Such a matrix can be always diagonalized by means of a regular similarity transformation

$$g(\lambda)Q(\lambda)g(\lambda)^{-1} = D(\lambda) = D_0 + (\lambda-\lambda_k)D_1 + \cdots\,.$$

Indeed, regularity means that

$$\begin{aligned} g(\lambda) &= g_0 + (\lambda-\lambda_k)g_1 + (\lambda-\lambda_k)^2 g_2 + \cdots \\ g(\lambda)^{-1} &= h_0 + (\lambda-\lambda_k)h_1 + (\lambda-\lambda_k)^2 h_2 + \cdots \end{aligned}$$

and, therefore,

$$\begin{aligned} I &= g(\lambda)g(\lambda)^{-1} = \\ &= \Big(g_0 + (\lambda-\lambda_k)g_1 + (\lambda-\lambda_k)^2 g_2 + \cdots\Big)\Big(h_0 + (\lambda-\lambda_k)h_1 + (\lambda-\lambda_k)^2 h_2 + \cdots\Big) \\ &= g_0h_0 + (\lambda-\lambda_k)(g_0h_1 + g_1h_0) + \cdots \end{aligned}$$

This allows to determine recurrently the inverse element

$$h_0 = g_0^{-1}\,, \qquad h_1 = -g_0^{-1}g_1g_0^{-1}\,, \qquad \text{etc.}$$

Thus,

$$
\begin{aligned}
& g(\lambda)Q(\lambda)g(\lambda)^{-1} = \\
& = \Big(g_0 + (\lambda-\lambda_k)g_1 + \cdots\Big)\Big(Q_0 + (\lambda-\lambda_k)Q_1 + \cdots\Big)\Big(g_0^{-1} - (\lambda-\lambda_k)g_0^{-1}g_1g_0^{-1} + \cdots\Big) \\
& = g_0Q_0g_0^{-1} + (\lambda-\lambda_k)\Big(g_0Q_1g_0^{-1} + g_1Q_0g_0^{-1} - g_0Q_0g_0^{-1}g_1g_0^{-1}\Big) + \cdots
\end{aligned}
$$

Thus, we see that g_0 must diagonalize Q_0:

$$D_0 = g_0Q_0g_0^{-1}$$

and g_1 is found from the condition that

$$g_0Q_1g_0^{-1} + g_1Q_0g_0^{-1} - g_0Q_0g_0^{-1}g_1g_0^{-1} = g_0Q_1g_0^{-1} + [g_1g_0^{-1}, D_0]$$

is diagonal. The commutator of a diagonal matrix with any matrix is off-diagonal. Thus, $[g_1g_0^{-1}, D_0]$ is off-diagonal and the matrix g_1 is found from the condition that $[g_1g_0^{-1}, D_0]$ kills the off-diagonal elements of $g_0Q_1g_0^{-1}$. Thus,

$$D(\lambda) = D_0 + (\lambda-\lambda_k)\big(g_0Q_1g_0^{-1}\big)_{ii}E_{ii} + \cdots$$

Thus, we have shown that by means of a regular similarity transformation around the pole $\lambda = \lambda_k$ the Lax matrix can be brought to the diagonal form

$$A(\lambda) = \sum_{r=-n_k}^{-1} \underbrace{A_{k,r}}_{\text{diag}}(\lambda-\lambda_k)^r + \text{regular}$$

The diagonalizing matrix $g(\lambda)$ is defined up to right multiplication by an arbitrary analytic diagonal matrix.

Define the matrix $B(\lambda)$ as

$$M(\lambda) = g(\lambda)B(\lambda)g(\lambda)^{-1} + \dot{g(\lambda)}g(\lambda)^{-1}\,,$$

where $g(\lambda)$ is a regular matrix which diagonalizes $L(\lambda)$ around $\lambda = \lambda_k$. The Lax representation implies that

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

Since $A(\lambda)$ is diagonal then $\dot{A}(\lambda) = 0$, i.e., $A(\lambda)$ comprises integrals of motion! Further, the consistency of the Lax equation implies that $B(\lambda)$ is a diagonal matrix as well.

We have

$$L_k = (g^{(k)}A^{(k)}g^{(k)-1})_-\,, \qquad M_k = (g^{(k)}B^{(k)}g^{(k)-1})_-\,.$$

We see that because $g^{(k)}$ is regular the matrices L_k and M_k depend only on the *singular* part of $A^{(k)}$ and $B^{(k)}$. Also expanding

$$g^{(k)} = \sum_{r=0}^{n_k-1} g_{k,r}(\lambda-\lambda_k)^r + \text{higher powers}$$

we see that only terms with $r = 0, \cdots, n_k - 1$ contribute to the singular parts of L_k and M_k.

The discussion above allows one to establish the independent degrees of freedom of the Lax pair For every pole λ_k these are two singular diagonal matrices

$$A_-^{(k)} = \sum_{r=-n_k}^{-1} \underbrace{A_{k,r}}_{\text{diag}}(\lambda-\lambda_k)^r, \qquad B_-^{(k)} = \sum_{r=-m_k}^{-1} \underbrace{B_{k,r}}_{\text{diag}}(\lambda-\lambda_k)^r$$

and a regular matrix $G^{(k)}$ of the order $n_k - 1$, defined up to right multiplication by a regular diagonal matrix

$$G^{(k)} = \sum_{r=0}^{n_k-1} g_{k,r}(\lambda-\lambda_k)^r,$$

plus, in addition two constant matrices L_0 and M_0. The L and M matrices are reconstructed from these data as

$$L(\lambda) = L_0 + \sum_k L_k(\lambda), \qquad L_k(\lambda) = (G^{(k)} A_-^{(k)} G^{(k)-1})_-$$
$$M(\lambda) = M_0 + \sum_k M_k(\lambda), \qquad M_k(\lambda) = (g^{(k)} B_-^{(k)} g^{(k)-1})_- .$$

Note that $g^{(k)}$ is determined by $G^{(k)}$. In other words, with $G^{(k)}$ one constructs $L(\lambda)$ and then diagonalize it around pole λ_k which produces the whole series $g^{(k)}$. These series is then used to build M_k.

Since $L(\lambda)$ and $M(\lambda)$ are rational functions we can easily count the number of independent variables and the number of equations. The independent variables contained in L are L_0 and $L_{k,r}$, $r = 1, \cdots, n_k$ (i.e. for each k there are n_k matrices). The independent variables contained in M are M_0 and $M_{k,r}$, $r = 1, \cdots, m_k$ (i.e. for each k there are m_k matrices). Thus, a counting in units of N^2, which is the size of matrices, gives

$$\text{number of variables} = \underbrace{2}_{L_0, M_0} + \sum_k n_k + \sum_k m_k = 2 + l + m$$
$$\text{number of equations} = \underbrace{1}_{\text{constant part}} + \underbrace{\sum_k (n_k + m_k)}_{\text{number of poles}} = 1 + l + m .$$

We see that the there is one more variable than the number of equations which reflects the gauge invariance of the Lax equation. On the Riemann surfaces of the higher genus the situation changes and the number of equations is always bigger than the number of independent variables.

The general solution of the non-dynamic constraints on $M(\lambda)$ has the form

$$M = M_0 + \sum_k M_k\,, \qquad M_k = P^{(k)}(L,\lambda)_-\,,$$

where $P^{(k)}(L,\lambda)$ is a polynomial in $L(\lambda)$ with coefficients rational in λ and $P^{(k)}(L,\lambda)_-$ is its singular part at $\lambda=\lambda_k$. Indeed, assuming that this is a solution we have

$$[M_k,L]_- = [P^{(k)}(L,\lambda)_-,L]_- = [P^{(k)}(L,\lambda)-P^{(k)}(L,\lambda)_+,L]_- = -[P^{(k)}(L,\lambda)_+,L]_-$$

but the r.h.s. here has poles of degree n_k and less. Let us show that this is the general solution. Recall that $A^{(k)}(\lambda)$ is a diagonal $N\times N$ matrix with all its matrix elements distinct at $\lambda=\lambda_k$. Its powers

$$\Big(A^{(k)}(\lambda)\Big)^0, \qquad \cdots, \qquad \Big(A^{(k)}(\lambda)\Big)^{N-1}$$

span the space of all diagonal matrices. Thus,

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

where $P^{(k)}(A^{(k)})$ is a polynomial of degree $N-1$ in $A^{(k)}$. Substituting this into the formula for M_k we get

$$M_k = (g^{(k)}B^{(k)}_-g^{(k)-1})_- = (g^{(k)}P^{(k)}(A^{(k)}(\lambda),\lambda)g^{(k)-1})_- = P^{(k)}(L,\lambda)_-\,.$$

The coefficients of $P^{(k)}$ are rational functions of the matrix elements of $A^{(k)}$ and $B^{(k)}$ and therefore they admit the Laurent expansion in $\lambda-\lambda_k$.

The following situation takes place:

- Dynamical variables are the elements of L. Choosing the number and the order of poles of the Lax matrix amounts to specifying a particular model.
- Choosing the polynomials $P^{(k)}(L,\lambda)$ is equivalent to specifying the the dynamical flows (one of the Hamiltonians).

The Euler top. For the Euler top we have

$$L(\lambda) = \mathcal{I}^2 + \frac{1}{\lambda}J\,, \qquad M(\lambda) = \lambda\mathcal{I} + \Omega\,.$$

We can add to M a polynomial of L to shift the pole in λ from infinity to the zero point. In fact one has to take

$$P(L) = \lambda(\alpha L^2 + \beta L + \gamma)\,,$$

where

$$\begin{aligned}
\alpha &= -\frac{1}{I_1 I_2 I_3}\,,\\
\beta &= \frac{I_1^2 + I_2^2 + I_3^3}{2 I_1 I_2 I_3}\,,\\
\gamma &= \frac{(I_1 + I_2 + I_3)(I_2 + I_3 - I_1)(I_1 + I_2 - I_2)(I_1 + I_2 - I_3)}{16 I_1 I_2 I_3}\,.
\end{aligned}$$

With this choice we get

$$M(\lambda) \to \lambda\mathcal{I} + \Omega - P(L) = \underbrace{\Omega - \alpha(\mathcal{I}^2 J + J\mathcal{I}^2) - \beta J}_{=0} - \frac{\alpha}{\lambda} J^2\,.$$

Thus we have a new Lax pair

$$L(\lambda) = \mathcal{I}^2 + \frac{1}{\lambda} J\,, \qquad M(\lambda) = -\frac{\alpha}{\lambda} J^2\,.$$

Check

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

Thus, we should get

$$\dot{J} = -\frac{1}{I_1 I_2 I_3}[\mathcal{I}^2, J^2]\,.$$

These are precisely the Euler equations

$$\frac{dJ_1}{dt} = a_1 J_2 J_3\,, \qquad \frac{dJ_2}{dt} = a_2 J_3 J_1\,, \qquad \frac{dJ_3}{dt} = a_3 J_1 J_2.$$

Here

$$a_1 = \frac{I_2 - I_3}{I_2 I_3}\,, \qquad a_2 = \frac{I_3 - I_1}{I_1 I_3}\,, \qquad a_3 = \frac{I_1 - I_2}{I_1 I_2}\,.$$

The eigenvalues of J are $(0, i\sqrt{\vec{J}^2}, -i\sqrt{\vec{J}^2})$ and they are non-dynamical since $\vec{J}^2$ belongs to the center of the Poisson structure.

4. Two-dimensional integrable PDEs

Here we introduce some interesting examples of infinite-dimensional Hamiltonian systems which appear to be integrable.

4.1 General remarks

Remarkably, there exist certain differential equations for functions depending on two variables (x, t) which can be treated as integrable Hamiltonian systems with infinite number of degrees of freedom. This is an (incomplete) list of such models

- The Korteweg-de-Vries equation

$$\frac{\partial u}{\partial t} = 6uu_x - u_{xxx}\,.$$

- The non-linear Schrodinger equation

$$i\frac{\partial \psi}{\partial t} = -\psi_{xx} + 2\kappa|\psi|^2\psi\,,$$

  where $\psi = \psi(x, t)$ is a complex-valued function.

- The Sine-Gordon equation

$$\frac{\partial^2 \phi}{\partial t^2} - \frac{\partial^2 \phi}{\partial x^2} + \frac{m^2}{\beta}\sin\beta\phi = 0$$

- The classical Heisenberg magnet

$$\frac{\partial \vec{S}}{\partial t} = \vec{S} \times \frac{\partial^2 \vec{S}}{\partial x^2}\,,$$

  where $\vec{S}(x, t)$ lies on the unit sphere in $\mathbb{R}^3$.

The complete specification of each model requires also boundary and initial conditions. Among the important cases are

1. *Rapidly decreasing case.* We impose the condition that

$$\psi(x, t) \to 0 \quad \text{when} \quad |x| \to \infty$$

   sufficiently fast, i.e., for instance, it belongs to the Schwarz space $\mathscr{L}(\mathbb{R}^1)$, which means that ψ is differentiable function which vanishes faster than any power of $|x|^{-1}$ when $|x| \to \infty$.

2. *Periodic boundary conditions.* Here we require that ψ is differentiable and satisfies the periodicity requirement

$$\psi(x + 2\pi, t) = \psi(x, t)\,.$$

The soliton was first discovered by accident by the naval architect, John Scott Russell, in August 1834 on the Glasgow to Edinburg channel.[6] The modern theory originates from the work of Kruskal and Zabusky in 1965. They were the first ones to call Russel's solitary wave a solition.

4.2 Soliton solutions

Here we discuss the simplest cnoidal wave type (periodic) and also one-soliton solutions of the KdV and SG equations For the discussion of the cnoidal wave and the one-soliton solution of the non-linear Schrodinger equation see the corresponding problem in the problem set.

4.2.1 Korteweg-de-Vries cnoidal wave and soliton

By rescaling of t, x and u one can bring the KdV equation to the canonical form

$$u_t + 6uu_x + u_{xxx} = 0\,.$$

We will look for a solution of this equation in the form of a single-phase periodic wave of a permanent shape

$$u(x,t) = u(x - vt)\,,$$

where $v = \text{const}$ is the phase velocity. Plugging this ansatz into the equation we obtain

$$-vu_x + 6uu_x + u_{xxx} = \frac{d}{dx}\Big(- vu + 3u^2 + u_{xx}\Big) = 0\,.$$

We thus get

$$-vu + 3u^2 + u_{xx} + e = 0\,,$$

where e is an integration constant. Multiplying this equation with an integrating factor u_x we get

$$-vuu_x + 3u^2u_x + u_xu_{xx} + eu_x = \frac{d}{dx}\Big(- \frac{v}{2}u^2 + u^3 + \frac{1}{2}u_x^2 + eu\Big) = 0\,,$$

[6] Russel described his discovery as follows: "I believe I shall best introduce this phenomenon by describing the circumstances of my own first acquaintance with it. I was observing the motion of a boat which was rapidly drawn along a narrow channel by a pair of horses, when the boat suddenly stopped-not so the mass of the water in the channel which it had put in motion; it accumulated round the prow of the vessel in a state of violent agitation, then suddenly leaving it behind, rolled forward with great velocity, assuming the form of a large solitary elevation, a rounded, smooth and well-defined heap of water, which continued its course along the channel apparently without change of form or diminution of speed. I followed it on horseback, and overtook it still rolling on at a rate of some eight or nine miles an hour, preserving its original figure some thirty feet along and a foot or foot and a half in height. Its height gradually diminished, and after a chase of one or two miles I lost it in the windings of the channel. Such, in the month of August 1834, was my first chance interview with that singular and beautiful phenomenon which I have called the Wave of Translation, a name which it now very generally bears.

We thus obtain

$$u_x^2 = k - 2eu + vu^2 - 2u^3 = -2(u-b_1)(u-b_2)(u-b_3)\,,$$

where k is another integration constant. In the last equation we traded the integration constants e, k for three parameters $b_3 \geq b_2 \geq b_1$ which satisfy the relation

$$v = 2(b_1 + b_2 + b_3)\,.$$

Equation

$$u_x^2 = -2(u-b_1)(u-b_2)(u-b_3)\,,$$

describes motion of a "particle" with the coordinate u and the time x in the potential $V = 2(u-b_1)(u-b_2)(u-b_3)$. Since $u_x^2 \geq 0$ for $b_2 \leq u \leq b_3$ the particle oscillates between the end points b_2 and b_3 with the period

$$\ell = 2\int_{b_2}^{b_3} \frac{\mathrm{d}u}{\sqrt{-2(u-b_1)(u-b_2)(u-b_3)}} = \frac{2\sqrt{2}}{(b_3-b_2)^{1/2}}\mathrm{K}(m)\,,$$

where m is an elliptic modulus $0 \leq m = \frac{b_3-b_2}{b_3-b_1} \leq 1$.

The equation

$$u_x^2 = -2(u-b_1)(u-b_2)(u-b_3)\,,$$

can be integrated in terms of Jacobi elliptic cosine function $\mathrm{cn}(x,m)$ to give

$$u(x,t) = b_2 + (b_3 - b_2)\ \mathrm{cn}^2\Big(\sqrt{(b_3-b_1)/2}(x - vt - x_0), m\Big)\,,$$

where x_0 is an initial phase. This solution is often called as *cnoidal wave.* When $m \to 1$, i.e. $b_2 \to b_1$ the cnoidal wave turns into a solitary wave

$$u(x,t) = b_1 + \frac{A}{\cosh^2\Big(\sqrt{\frac{A}{2}}(x - vt - x_0)\Big)}\,.$$

Here the velocity $v = 2(b_1+b_2+b_3) = 2(2b_1+b_3) = 2(3b_1+b_3-b_1)$ is connected to the amplitude $A = b_3 - b_1$ by the relation

$$v = 6b_1 + 2A\,.$$

Here $u(x,t) = b_1$ is called a background flow because $u(x,t) \to b_1$ as $x \to \pm\infty$. One can further note that the background flow can be eliminated by a passage to a moving frame and using the invariance of the KdV equation w.r.t. the Galilean transformation $u \to u + d$, $x \to x - 6dt$, where d is constant.

To sum up the cnoidal waves form a three-parameter family of the KdV solutions while solitons are parametrized by two independent parameters (with an account of the background flow).

4.2.2 Sine-Gordon cnoidal wave and soliton

Consider the Sine-Gordon equation

$$\phi_{tt} - \phi_{xx} + \frac{m^2}{\beta}\sin\beta\phi = 0\,,$$

where we assume that the functions $\phi(x,t)$ and $\phi(x,t) + 2\pi/\beta$ are assumed to be equivalent. Make an ansatz

$$\phi(x,t) = \phi(x - vt)$$

which leads to

$$(v^2-1)\phi_{xx} + \frac{m^2}{\beta}\sin\beta\phi = 0\,.$$

This can be integrated once

$$C = \frac{v^2-1}{2}\phi_x^2 - \frac{m^2}{\beta^2}\cos\beta\phi = \frac{v^2-1}{2}\phi_x^2 + \frac{2m^2}{\beta^2}\sin^2\frac{\beta\phi}{2} - \frac{m^2}{\beta^2}\,.$$

where C is an integration constant. This is nothing else as the conservation law of energy for the mathematical pendulum in the gravitational field of the Earth! We further bring equation to the form

$$\phi_x^2 = \frac{2}{v^2-1}\Big(C + \frac{m^2}{\beta^2} - \frac{2m^2}{\beta^2}\sin^2\frac{\beta\phi}{2}\Big)\,. \tag{4.1}$$

As in the case of the pendulum we make a substitution $y = \sin\frac{\beta\phi}{2}$ which gives

$$(y')^2 = \frac{m^2}{(v^2-1)}(1-y^2)\left(\frac{C+\frac{m^2}{\beta^2}}{\frac{2m^2}{\beta^2}} - y^2\right)\,.$$

This leads to solutions in terms of elliptic functions which are analogous to the cnoidal waves of the KdV equation. However, as we know the pendulum has three phases of motion: oscillatory (elliptic solution), rotatory (elliptic solution) and motion with an infinite period. The later solution is precisely the one that would correspond to the Sine-Gordon soliton we are interested in. Assuming $v^2 < 1$ we see[7] that such a solution would arise from (4.1) if we take $C = -\frac{m^2}{\beta^2}$. In this case equation (4.1) reduces to

$$\phi_x = \frac{2m}{\beta\sqrt{1-v^2}}\sin\frac{\beta\phi}{2}\,.$$

This can be integrated to[8]

$$\phi(x,t) = -\epsilon_0\frac{4}{\beta}\arctan\exp\left(\frac{m(x-vt-x_0)}{\sqrt{1-v^2}}\right)\,.$$

[7] Restoring the speed of light c this condition for the velocity becomes $v^2 < c^2$, i.e., the center of mass of the soliton cannot propagate faster than light.

[8] From the equation above we see that if $\phi(x,t)$ is a solution then $-\phi(x,t)$ is also a solution.

Here $\epsilon_0 = \pm 1$. This solution can be interpreted in terms of relativistic particle moving with the velocity v. The field $\phi(x,t)$ has an important characteristic – topological charge

$$Q = \frac{\beta}{2\pi}\int \mathrm{d}x \frac{\partial\phi}{\partial x} = \frac{\beta}{2\pi}(\phi(\infty) - \phi(-\infty))\,.$$

On our solutions we have

$$Q = \frac{\beta}{2\pi}\Big(-\epsilon_0\frac{4}{\beta}\Big)\big(\frac{\pi}{2} - 0\big) = -\epsilon_0\,,$$

because $\arctan(\pm\infty) = \pm\frac{\pi}{2}$ and $\arctan 0 = 0$. In addition to the continuous parameters v and x_0, the soliton of the SG model has another important discrete characteristic – topological charge $Q = -\epsilon_0$. Solutions with $Q = 1$ are called solitons (kinks), while solutions with $Q = -1$ are called ani-solitons (anti-kinks).

Here we provide another useful representation for the SG soliton, namely

$$\phi(x,t) = \epsilon_0 \frac{2i}{\beta}\log\frac{1 + ie^{\frac{m(x-vt-x_0)}{\sqrt{1-v^2}}}}{1 - ie^{\frac{m(x-vt-x_0)}{\sqrt{1-v^2}}}}\,.$$

Indeed, looking at the solution we found we see that we can cast it in the form $\arctan\alpha = z \equiv -\frac{\beta}{4\epsilon_0}\phi(x,t)$ or $\alpha = \tan z = -i\frac{e^{2iz}-1}{e^{2iz}+1}$, where $\alpha = e^{\frac{m(x-vt-x_0)}{\sqrt{1-v^2}}}$. From here $z = \frac{1}{2i}\log\frac{1+i\alpha}{1-i\alpha}$ and the announced formula follows.

Remark. The stability of solitons stems from the delicate balance of "nonlinearity" and "dispersion" in the model equations. Nonlinearity drives a solitary wave to concentrate further; dispersion is the effect to spread such a localized wave. If one of these two competing effects is lost, solitons become unstable and, eventually, cease to exist. In this respect, solitons are completely different from "linear waves" like sinusoidal waves. In fact, sinusoidal waves are rather unstable in some model equations of soliton phenomena.

Sine-Gordon model has even more sophisticated solutions. Consider the following

$$\phi(x,t) = \frac{4}{\beta}\arctan\frac{\omega_2}{\omega_1}\frac{\sin\Big(\frac{m\omega_1(t-vx)}{\sqrt{1-v^2}} + \phi_0\Big)}{\cosh\Big(\frac{m\omega_2(x-vt-x_0)}{\sqrt{1-v^2}}\Big)}\,.$$

This is solution of the SG model which is called a *double-soliton or breaser*. Except motion with velocity v corresponding to a relativistic particle the breaser oscillates both in space and in time with frequencies $\frac{mv\omega_1}{\sqrt{1-v^2}}$ and $\frac{m\omega_1}{\sqrt{1-v^2}}$ respectively. The parameter ϕ_0 plays a role of the initial phase. In particular, if $v = 0$ the breaser is a time-periodic solution of the SG equation. It has zero topological charge and can be interpreted as the bound state of the soliton and anti-soliton.

4.3 Zero-curvature representation

The inverse scattering method (the method of finding certain class of solutions of a non-linear integrable PDE) is based on the following remarkable observation. A two-dimensional PDE appears as the consistency condition of the overdetermined system of equations

$$\frac{\partial\Psi}{\partial x} = U(x,t,\lambda)\Psi\,,$$
$$\frac{\partial\Psi}{\partial t} = V(x,t,\lambda)\Psi\;.$$

for a proper choice of the matrices $U(x,t,\lambda)$ and $V(x,t,\lambda)$. The consistency condition arises upon differentiation the first equation w.r.t. t and the second w.r.t. x:

$$\frac{\partial^2\Psi}{\partial t\partial x} = \partial_t U(x,t,\lambda)\Psi + U(x,t,\lambda)\partial_t\Psi = \Big(\partial_t U(x,t,\lambda) + U(x,t,\lambda)V(x,t,\lambda)\Big)\Psi\,,$$
$$\frac{\partial^2\Psi}{\partial x\partial t} = \partial_x V(x,t,\lambda)\Psi + V(x,t,\lambda)\partial_x\Psi = \Big(\partial_x V(x,t,\lambda) + V(x,t,\lambda)U(x,t,\lambda)\Big)\Psi\,,$$

which implies the fulfilment of the following relation

$$\partial_t U - \partial_x V + [U,V] = 0\,.$$

If we introduce a gauge field $\mathscr{L}_\alpha$ with components $\mathscr{L}_x = U$, $\mathscr{L}_t = V$, then the last relation is the condition of vanishing of the curvature of $\mathscr{L}_\alpha$:

$$F_{\alpha\beta}(\mathscr{L}) \equiv \partial_\alpha\mathscr{L}_\beta - \partial_\beta\mathscr{L}_\alpha - [\mathscr{L}_\alpha,\mathscr{L}_\beta] = 0\,.$$

Example: KdV equation. Introduce the following 2×2 matrices

$$U = \begin{pmatrix} 0 & 1 \\ \lambda+u & 0 \end{pmatrix}, \qquad V = \begin{pmatrix} u_x & 4\lambda - 2u \\ 4\lambda^2 + 2\lambda u + u_{xx} - 2u^2 & -u_x \end{pmatrix}.$$

Show by direct computation that

$$\partial_t U - \partial_x V + [U,V] = \begin{pmatrix} 0 & 0 \\ u_t + 6uu_x - u_{xxx} & 0 \end{pmatrix}.$$

Example: Sine-Gordon equation. Introduce the following 2×2 matrices

$$U = \frac{\beta}{4i}\phi_t\sigma_3 + \frac{k_0}{i}\sin\frac{\beta\phi}{2}\sigma_1 + \frac{k_1}{i}\cos\frac{\beta\phi}{2}\sigma_2$$
$$V = \frac{\beta}{4i}\phi_x\sigma_3 + \frac{k_1}{i}\sin\frac{\beta\phi}{2}\sigma_1 + \frac{k_0}{i}\cos\frac{\beta\phi}{2}\sigma_2\,,$$

where σ_i are the Pauli matrices[9] and

$$k_0 = \frac{m}{4}\Big(\lambda + \frac{1}{\lambda}\Big)\,, \qquad k_1 = \frac{m}{4}\Big(\lambda - \frac{1}{\lambda}\Big)\,.$$

Show by direct computation that the condition of zero curvature is equivalent to the Sine-Gordon equation.

The one-parameter family of the flat connections allows one to define the monodromy matrix $\mathrm{T}(\lambda)$ which is the path-ordered exponential of the Lax component $U(\lambda)$:

$$\mathrm{T}(\lambda) = \mathscr{P} \exp \int_0^{2\pi} \mathrm{d}x U(\lambda)\,. \tag{4.3}$$

Let us derive the time evolution equation for this matrix. We have

$$\begin{aligned} \partial_t \mathrm{T}(\lambda) &= \int_0^{2\pi} \mathrm{d}x \;\; \mathscr{P}e^{\int_x^{2\pi} \mathrm{d}y U} \,(\partial_t U)\, \mathscr{P}e^{\int_0^x \mathrm{d}y U} \\ &= \int_0^{2\pi} \mathrm{d}x \;\; \mathscr{P}e^{\int_x^{2\pi} \mathrm{d}y U} \,(\partial_x V + [V, U])\, \mathscr{P}e^{\int_0^x \mathrm{d}y U}\,, \end{aligned} \tag{4.4}$$

where in the last formula we used the flatness of $\mathscr{L}_\alpha \equiv (U, V)$. The integrand of the expression we obtained is the total derivative

$$\partial_t \mathrm{T}(\lambda) = \int_0^{2\pi} \mathrm{d}x \; \partial_x \left(\mathscr{P}e^{\int_x^{2\pi} \mathrm{d}y U} \; V \; \mathscr{P}e^{\int_0^x \mathrm{d}y U} \right)\,. \tag{4.5}$$

Thus, we obtained the following evolution equation

$$\partial_t \mathrm{T}(\lambda) = [V(2\pi, t, \lambda), \mathrm{T}(\lambda)]\,. \tag{4.6}$$

This formula shows that the eigenvalues of $\mathrm{T}(\lambda)$ generate an infinite set of integrals of motion upon expansion in λ. Thus, the spectral properties of the model are encoded into the monodromy matrix.

The wording "monodromy" comes from the fact that $\mathrm{T}(t)$ represents the monodromy of a solution of the fundamental linear problem:

$$\Psi(2\pi, t) = \mathrm{T}(t)\Psi(0, t)\,.$$

---

[9]The Pauli matrices are

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\,, \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}\,, \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\,. \tag{4.2}$$

Indeed, if we differentiate this equation over t we get

$$\partial_t \Psi(2\pi, t) = \partial_t \mathrm{T}\Psi(0,t) + \mathrm{T}\partial_t\Psi(0,t)\,,$$

which, according to the fundamental linear system, gives

$$\mathscr{L}_t(2\pi, t)\mathrm{T}\Psi(0,t) = \partial_t \mathrm{T}\Psi(0,t) + \mathrm{T}\mathscr{L}_t(0,t)\Psi(0,t)\,.$$

This leads to the same equation for the time evolution of the monodromy matrix as found before:

$$\partial_t \mathrm{T} = [\mathscr{L}_t, \mathrm{T}]\,.$$

4.4 Local integrals of motion

The Lax representation of the two-dimensional PDE allows one to exhibit an infinite number of conservation laws. The procedure to derive the conservation laws from the Lax representation is a direct analogue of the Zakharov-Shabat construction for the finite-dimensional case. It is called the *abelianization procedure.*

Once again we start from the zero-curvature condition

$$\partial_t U - \partial_x V - [V, U] = 0\,.$$

We assume that the matrices $U(x,t,\lambda)$ and $V(x,t,\lambda)$ depend on the spectral parameter λ in a rational way and they have poles at constant, i.e. x,t-independent, values of λ_k. Thus, we can write

$$U = U_0 + \sum_k U_k\,, \qquad U_k = \sum_{r=-n_k}^{-1} U_{k,r}(x,t)(\lambda-\lambda_r)^r\,,$$

$$V = V_0 + \sum_k V_k\,, \qquad V_k = \sum_{r=-m_k}^{-1} V_{k,r}(x,t)(\lambda-\lambda_r)^r\,.$$

The same counting as in the finite-dimensional case shows that the zero-curvature equations are always compatible: there is one more variable than the number of equations, but there is a gauge transformation which leads the zero-curvature condition invariant.

To understand solutions of the zero-curvature condition we will perform a local analysis around a pole $\lambda = \lambda_k$. Our aim is to show that around each singularity one can perform a gauge transformation which brings the matrices $U(\lambda)$ and $V(\lambda)$ to a diagonal form. Finally, to make the consideration as simple as possible we assume that the pole is located at zero.

In the neighbourhood of $\lambda = 0$ the functions U and V can be expanded into Laurent series

$$U(x,t,\lambda) = \sum_{r=-n}^{\infty} U_r(x,t)\lambda^r\,, \qquad V(x,t,\lambda) = \sum_{r=-m}^{\infty} V_r(x,t)\lambda^r\,.$$

Let $g \equiv g(x,t,\lambda)$ be a regular gauge transformation around $\lambda = 0$ that is

$$g = \sum_{r=0}^{\infty} g_r \lambda^r \,, \qquad g^{-1} = \sum_{r=0}^{\infty} h_r \lambda^r \,.$$

Consider the gauge transformation

$$\begin{aligned} \tilde{U} &= gUg^{-1} + \partial_x g g^{-1} \,, \\ \tilde{V} &= gVg^{-1} + \partial_t g g^{-1} \,. \end{aligned}$$

Consider the transition matrix $\mathrm{T}(x,y,\lambda)$ which is a solution of the differential equation

$$\Big(\partial_x - U(x,\lambda)\Big)\mathrm{T}(x,y,\lambda) = 0$$

satisfying the initial condition $\mathrm{T}(x,x,\lambda) = I$. Formally such a solution is given by the path-ordered exponent

$$\mathrm{T}(x,y,\lambda) = \mathscr{P} e^{\int_y^x \mathrm{d}z U(z,\lambda)} \,.$$

Under the gauge transformation we have

$$g(x,\lambda)\Big(\partial_x - U(x,\lambda)\Big)g^{-1}(x,\lambda)\mathrm{T}_g(x,y,\lambda) = 0 \,,$$

where $\mathrm{T}_g(x,y,\lambda)$ is the transition matrix for the gauged-transformed connection which also obeys the condition $\mathrm{T}_g(x,x,\lambda) = I$. Thus, we obtain

$$\mathrm{T}_g(x,y,\lambda) = g(x,\lambda)\mathrm{T}(x,y,\lambda)g^{-1}(y,\lambda) \,.$$

This formula shows how the transition matrix transforms under the gauge transformations of the Lax connection. By means of a regular gauge transformation the transition matrix can be diagonalized around every pole of the matrix U:

$$\mathrm{T}(x,y,\lambda) = g(x,\lambda) \exp(D(x,y,\lambda)) g^{-1}(y,\lambda) \,,$$

where

$$D(x,y,\lambda) = \sum_{r=-n}^{\infty} D_r(x,y)\lambda^r$$

is the diagonal matrix. Below we consider a concrete example which illustrates the abelianization procedure as well as the technique of constructing local integrals of motion.

Example: The Heisenberg model. We start with the definition of the model classical Heisenberg model. Consider a spin variable $S(x)$:

$$S(x) = \sum_i S^i(x)\sigma_i \,.$$

Clearly, $S^i(x)^2 = s^2$. Here σ^i are the standard Pauli matrices obeying the relations

$$[\sigma_i, \sigma_j] = 2i\epsilon_{ijk}\sigma_k\,, \qquad \mathrm{tr}(\sigma_j\sigma_k) = 2\delta_{ij}\,.$$

The spins $S^i(x)$ are the dynamical variables subject to the Poisson structure

$$\{S^i(x), S^j(y)\} = \epsilon^{ijk}S^k(x)\delta(x-y)\,.$$

The phase space is thus infinite-dimensional. Check in the class the Jacobi identity! The Hamiltonian of the model is

$$H = -\frac{1}{4}\int_0^{2\pi} \mathrm{d}x\ \mathrm{tr}(\partial_x S\partial_x S)$$

Let us derive equations of motion (in the class!). We have

$$\begin{aligned}
&\partial_t S(x) = \{H, S(x)\} = -\frac{1}{4}\int_0^{2\pi} \mathrm{d}y\ \{\mathrm{tr}(\partial_y S\partial_y S), S(y)\} \\
&= -\int_0^{2\pi} \mathrm{d}y\ \partial_y S^j(y)\{\partial_y S^j(y), S^k(x)\}\sigma_k = -\int_0^{2\pi} \mathrm{d}y\ \partial_y S^j(y)\epsilon^{jki}\partial_y(S^i(y)\delta(y-x))\sigma_k = \\
&= \epsilon^{jki}\partial_x^2 S^j(x)S^i(x)\sigma_k = \epsilon^{ijk}S^i(x)\partial_x^2 S^j(x)\sigma_k = \frac{1}{2i}[S^i(x)\sigma_i, \partial_x^2 S^j(x)\sigma_j]
\end{aligned}$$

Thus, equations of motion read

$$\partial_t S = -\frac{i}{2}[S, \partial_x^2 S] = -\frac{i}{2}\partial_x[S, \partial_x S]\,.$$

If we introduce the non-abelian $\mathfrak{su}(2)$-current J with components

$$J_x = S\,, \qquad J_t = -\frac{i}{2}[S, \partial_x S]$$

then the equations of motion take the form of the current conservation law:

$$\partial_t J_x - \partial_x J_t = 0\,,$$

which is $\epsilon^{\alpha\beta}\partial_\alpha J_\beta = 0$. Equations of motion

$$\partial_t S = -\frac{i}{2}[S, \partial_x^2 S]$$

called the *Landau-Lifshitz equations*. In this form these equations can be generalized to any Lie algebra. The integrability of the model relies on the fact that equations of motion can be obtained from the condition of zero curvature:

$$(\partial_\alpha - \mathscr{L}_\alpha)\Psi(x,t) = 0\,.$$

Here

$$\begin{aligned}
\mathscr{L}_x &= -\frac{i}{\lambda}S(x)\,, \\
\mathscr{L}_t &= -\frac{2is^2}{\lambda^2}S(x) - \frac{1}{2\lambda}[S(x), \partial_x S(x)]\,.
\end{aligned}$$

Indeed,

$$\partial_t \mathscr{L}_x - \partial_x \mathscr{L}_t + [\mathscr{L}_x, \mathscr{L}_t] = -\frac{i}{\lambda}\partial_t S(x) + \frac{2is^2}{\lambda^2}\partial_x S(x) + \frac{1}{2\lambda}\partial_x[S(x), \partial_x S(x)] + \frac{i}{2\lambda^2}[S(x), [S(x), \partial_x S(x)]] = 0\,.$$

Now one can compute the Poisson bracket between the components $\mathscr{L}_x \equiv U(x, \lambda)$ of the Lax connection. We have

$$\{U(x,\lambda), U(y,\mu)\} = -\frac{1}{\lambda\mu}\{S^i(x), S^j(y)\}\sigma_i \otimes \sigma_j = -\frac{1}{\lambda\mu}\epsilon^{ijk}S^k(x)\sigma_i \otimes \sigma_j \delta(x-y)\,.$$

On the other hand, let us compute

$$\Big[\frac{\sigma_i \otimes \sigma_i}{\lambda - \mu}, U(x,\lambda)\otimes \mathbb{I} + \mathbb{I} \otimes U(y,\mu)\Big]\delta(x-y) = -\Big[\frac{\sigma_i \otimes \sigma_i}{\lambda-\mu}, \frac{i}{\lambda}S(x)\otimes \mathbb{I} + \mathbb{I}\otimes \frac{i}{\mu}S(y)\Big]\delta(x-y)$$
$$= -\frac{i}{\lambda-\mu}S^k(x)\Big(\frac{1}{\lambda}[\sigma_i, \sigma_k]\otimes \sigma_i + \frac{1}{\mu}\sigma_i \otimes [\sigma_i, \sigma_k]\Big)\delta(x-y) =$$
$$= \frac{2}{\lambda-\mu}\Big(\frac{1}{\lambda} - \frac{1}{\mu}\Big)\epsilon^{ijk}S^k(x)\sigma_i\otimes\sigma_j\delta(x-y) = -\frac{2}{\lambda\mu}\epsilon^{ijk}S^k(x)\sigma_i \otimes \sigma_j \delta(x-y)\,.$$

We thus proved that the Poisson bracket between the components of the Lax connection can be written in the form

$$\{U(x,\lambda), U(y,\mu)\} = \Big[r(\lambda,\mu), U(x,\lambda)\otimes \mathbb{I} + \mathbb{I}\otimes U(y,\mu)\Big]\delta(x-y)\,,$$

where the classical r-matrix appears to be

$$r(\lambda,\mu) = \frac{1}{2}\frac{\sigma_i \otimes \sigma_i}{\lambda - \mu}\,.$$

This form of the brackets between the components of the Lax connection implies that the Poisson bracket between the components of the monodromy matrix

$$\mathrm{T}(\lambda) = \mathscr{P}\exp\Big[\int_0^{2\pi} \mathrm{d}x\; U(x,\lambda)\Big]$$

is

$$\{\mathrm{T}(\lambda) \otimes \mathrm{T}(\mu)\} = \Big[r(\lambda,\mu), \mathrm{T}(\lambda)\otimes \mathrm{T}(\mu)\Big]\,.$$

This is the famous *Sklyanin bracket.* It is quadratic in the matrix elements of the monodromy matrix.

From the definition, $\mathrm{T}(\lambda)$ is analytic (entire)[10] in λ with an essential singularity at $\lambda = 0$[11] It is easy to find the expansion around $\lambda = \infty$:

$$\mathrm{T}(\lambda) = \mathbb{I} + \frac{i}{\lambda}\int_0^{2\pi} \mathrm{d}x \; S(x) - \frac{1}{\lambda^2}\int_0^{2\pi} \mathrm{d}x \; S(x) \int_0^x \mathrm{d}y \; S(y) + \cdots$$

The development in $1/\lambda$ has an infinite radius of convergency.

To find the structure of $\mathrm{T}(\lambda)$ around $\lambda = 0$ is more delicate but very important as it provides the local conserved charges in involution. Let us introduce the so-called partial monodromy

$$\mathrm{T}(x, \lambda) = \mathscr{P} \exp\Big[\int_0^x \mathrm{d}y \; U(y, \lambda)\Big] .$$

The main point is to note that there exists a *local* gauge transformation, *regular* at $\lambda = 0$, such that

$$\mathrm{T}(x, \lambda) = g(x) D(x) g^{-1}(0) ,$$

where $D(x) = \exp(id(x)\sigma_3)$ is a diagonal matrix. We can choose g to be unitary, and, since g is defined up to to a diagonal matrix, we can require that it has a real diagonal part:

$$g = \frac{1}{(1 + v\bar{v})^{\frac{1}{2}}} \begin{pmatrix} 1 & v \\ -\bar{v} & 1 \end{pmatrix} .$$

Then the differenial equation for the monodromy

$$\partial_x \mathrm{T} = U\mathrm{T} = -\frac{i}{\lambda} S\mathrm{T}$$

[10] In complex analysis, an entire function is a function that is holomorphic everywhere on the whole complex plane. Typical examples of entire functions are the polynomials, the exponential function, and sums, products and compositions of these. Every entire function can be represented as a power series which converges everywhere. Neither the natural logarithm nor the square root function is entire. Note that an entire function may have a singularity or even an essential singularity at the complex point at infinity. In the latter case, it is called a transcendental entire function. As a consequence of Liouville's theorem, a function which is entire on the entire Riemann sphere (complex plane and the point at infinity) is constant.

[11] Consider an open subset U of the complex plane $\mathbb{C}$, an element a of U, and a holomorphic function f defined on $U - a$. The point a is called an essential singularity for f if it is a singularity which is neither a pole nor a removable singularity. For example, the function $f(z) = \exp(1/z)$ has an essential singularity at $z = 0$. The point a is an essential singularity if and only if the limit

$$\lim_{z \to a} f(z)$$

does not exist as a complex number nor equals infinity. This is the case if and only if the Laurent series of f at the point a has infinitely many negative degree terms (the principal part is an infinite sum). The behavior of holomorphic functions near essential singularities is described by the Weierstrass-Casorati theorem and by the considerably stronger Picard's great theorem. The latter says that in every neighborhood of an essential singularity a, the function f takes on every complex value, except possibly one, infinitely often.

becomes a differential equation for g and d:

$$g^{-1}\partial_x g + i\partial_x d\sigma_3 + \frac{i}{\lambda}g^{-1}Sg = 0\,.$$

We project this equation on the Pauli matrices and get

$$\begin{aligned}\partial_x v &= -\frac{i}{\lambda}(S_- + 2vS_3 - S_+v^2)\\ \partial_x d &= \frac{1}{2\lambda}(-2S_3 + vS_+ + \bar{v}S_-)\,.\end{aligned}$$

The first of these equations is a Riccati equation for $v(x)$. Expanding in λ the functions $v(x)$ and $d(x)$ as

$$\begin{aligned}\partial_x d &= -\frac{s}{\lambda} + \sum_{n=0}^{\infty}\rho_n(x)\lambda^n\\ v(x) &= \sum_{n=0}^{\infty} v_n(x)\lambda^n\,, \qquad\qquad v_0 = \frac{S_3 - s}{S_+}\,,\end{aligned}$$

we rewrite the Riccati equation in the form

$$2isv_{n+1} = -v_n' + iS_+\sum_{m=1}^{n} v_{n+1-m}v_m$$

and

$$\rho_n = \frac{1}{2}(v_{n+1}S_+ + \bar{v}_{n+1}S_-)\,.$$

Note that $v(x)$ is regular at $\lambda = 0$. Equations above recursively determine the functions $v_n(x)$ and $\rho_n(x)$ as *local* functions of the dynamical variables $S^i(x)$. This describes the asymptotic behavior of $\mathrm{T}(\lambda)$ around $\lambda = 0$. The asymptotic series become convergent if we regularize the model by discretizing the space interval!

Concerning the monodromy matrix $\mathrm{T}(\lambda)$, since $g(x)$ is local and if we assume periodic boundary conditions, we can write

$$\mathrm{T}(\lambda) = \cos p(\lambda)\mathbb{I} + i\sin p(\lambda)M(\lambda)\,,$$

where $M(\lambda) = g(0)\sigma_3 g(0)^{-1}$ and

$$p(\lambda) = \int_0^{2\pi} \mathrm{d}x\ \partial_x d\,.$$

The trace of the monodromy matrix, called *the transfer matrix*, is

$$\mathrm{tr}\mathrm{T}(\lambda) = 2\cos p(\lambda)\,.$$

Thus, $p(\lambda)$ is the generating function for the commuting local conserved quantities

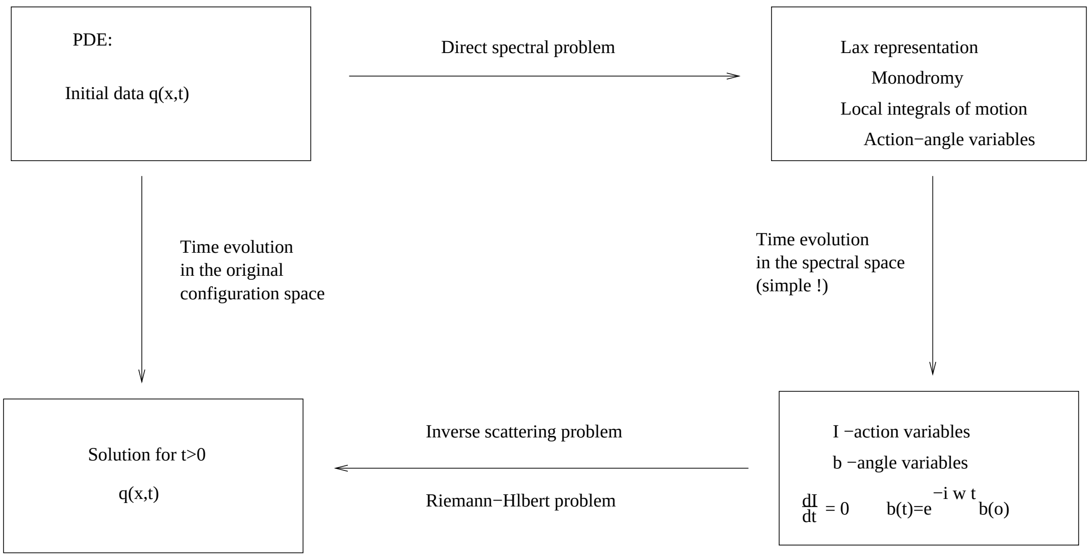
$$I_n = \int_0^{2\pi} \mathrm{d}x \;\rho_n(x)\,.$$

The first three integrals are

$$\begin{aligned} I_0 &= \frac{i}{4s}\int_0^{2\pi} \mathrm{d}x \;\log\Big(\frac{S_+}{S_-}\Big)\partial_x S_3\,, \\ I_1 &= -\frac{1}{16s^3}\int_0^{2\pi} \mathrm{d}x \;\mathrm{tr}\Big(\partial_x S\partial_x S\Big)\,, \\ I_2 &= \frac{i}{64s^5}\int_0^{2\pi} \mathrm{d}x \;\mathrm{tr}\Big(S[\partial_x S, \partial_x^2 S]\Big)\,. \end{aligned}$$

The integrals I_0 and I_1 correspond to momentum and energy respectively.

We conclude this section by outlining a general scheme known as Inverse Scattering Method which allows one to construct explicitly the multi-soliton solutions of integrable PDE's.



INVERSE SCATTERING TRANSFORM −− NON−LINEAR ANALOG OF THE FOURIER TRANSFORM

5. Quantum Integrable Systems

In this section we consider certain quantum integrable systems. The basis tool to solve them is known under the generic name "Bethe Ansatz". There are several different constructions of this type. They are

- *Coordinate Bethe ansatz.* This technique was originally introduced by H. Bethe to solve the XXX Heisenberg model.

- *Algebraic Bethe ansatz.* It was realized afterwards that the Bethe ansatz can be formulated in such a way that it can be understood as the quantum analogue of the classical inverse scattering method. Thus, "Algebraic Bethe ansatz" is another name for "Quantum inverse scattering method".

- *Functional Bethe ansatz.* The algebraic Bethe ansatz is not the only approach to solve the spectral problems of models connected with the Yang-Baxter algebra. It is only applicable if there exists a pseudo-vacuum. For models like the Toda chain, which has the same R-matrix as XXX Heisenberg magnet (spin-$\frac{1}{2}$ chain), but has no pseudo-vacuum, the algebraic Bethe ansatz fails. For these types of models another powerful technique, the method of "sepration of variables" was devised by E. Sklyanin. It is also known as "Functional Bethe Ansatz"

- *Nested Bethe ansatz.* The generalization of the Bethe ansatz to models with internal degrees of freedom proved to be very hard, because scattering involves changes of the internal states of scatters. This problem was eventually solved by C.N. Yang and M. Gaudin by means of what is nowadays called "nested Bethe ansatz".

- *Asymptotic Bethe anzatz* Many integrable systems in the finite volume cannot be solved by the Bethe ansatz methods. However, the Bethe ansatz provides the leading finite-size correction to the wave function, energy levels, etc. for systems in infinite volumes. Introduced and extensively studied by B. Sutherland.

- *Thermodynamic Bethe ansatz.* This method allows to investigate the thermodynamic properties of integrable systems.

5.1 Coordinate Bethe Ansatz (CBA)

Here we will demonstrate how CBA works at an example of the so-called one-dimensional spin-$\frac{1}{2}$ XXX Heisenberg model of ferromagnetism.

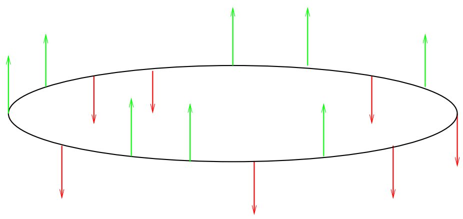
Consider a discrete circle which is a collection of ordered points labelled by the index n with the identification $n \equiv n + L$ reflecting periodic boundary conditions. Here L is a positive integer which plays the role of the length (volume) of the space. The numbers $n = 1, \ldots, L$ form a fundamental domain. To each integer n along the chain we associate a two-dimensional vector space $V = \mathbb{C}^2$. In each vector space we pick up the basis

$$|\uparrow\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad |\downarrow\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

We will call the first element "spin up" and the second one "spin down". We introduce the spin algebra which is generated by the spin variables S_n^α, where $\alpha = 1, 2, 3$, with commutation relations

$$[S_m^\alpha, S_n^\beta] = i\hbar\epsilon^{\alpha\beta\gamma}S_n^\gamma\delta_{mn}\,.$$

The spin operators have the following realization in terms of the standard Pauli matrices: $S_n^\alpha = \frac{\hbar}{2}\sigma^\alpha$ and the form the Lie algebra $\mathfrak{su}(2)$. Spin variables are subject to the periodic boundary condition $S_n^\alpha \equiv S_{n+L}^\alpha$.



Spin chain. A state of the spin chain can be represented as $|\psi\rangle = |\uparrow\uparrow\downarrow\uparrow\cdots\downarrow\uparrow\rangle$

The Hilbert space of the model has a dimension 2^L and it is

$$\mathscr{H} = \prod_{n=1}^{L} \otimes V_n = V_1 \otimes \cdots \otimes V_L$$

This space carries a representation of the global spin algebra whose generators are

$$S^\alpha = \sum_{n=1}^{L} \mathbb{I} \otimes \cdots \otimes \underbrace{S_n^\alpha}_{n-\text{th place}} \otimes \cdots \otimes \mathbb{I}\,.$$

The Hamiltonian of the model is

$$H = -J\sum_{n=1}^{L} S_n^\alpha S_{n+1}^\alpha\,,$$

where J is the coupling constant. More general Hamiltonian of the form

$$H = -J\sum_{n=1}^{L} J^\alpha S_n^\alpha S_{n+1}^\alpha\,,$$

where all three constants J^α are different defines the so-called XYZ model. In what follows we consider only XXX model. The basic problem we would like to solve is to find the spectrum of the Hamiltonian H.

The first interesting observation is that the Hamiltonian H commutes with the spin operators. Indeed,

$$
\begin{aligned}
[H,S^\alpha] &= -J\sum_{n,m=1}^{L}[S_n^\beta S_{n+1}^\beta,S_m^\alpha] = -J\sum_{n,m=1}^{L}[S_n^\beta,S_m^\alpha]S_{n+1}^\beta + S_n^\beta[S_{n+1}^\beta,S_m^\alpha] \\
&= -i\hbar\sum_{n,m=1}^{L}\left(\delta_{nm}\epsilon^{\alpha\beta\gamma}S_n^\beta S_{n+1}^\gamma - \delta_{n+1,m}\epsilon^{\alpha\beta\gamma}S_n^\beta S_{n+1}^\gamma\right) = 0\,.
\end{aligned}
$$

In other words, the Hamiltonian is central w.r.t all $\mathfrak{su}(2)$ generators. Thus, the spectrum of the model will be degenerate – all states in each $\mathfrak{su}(2)$ multiplet have the same energy.

In what follows we choose $\hbar = 1$ and introduce the raising and lowering operators $S_n^\pm = S_n^1 \pm iS_n^2$. They are realized as

$$
S^+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\,, \qquad S^- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}\,.
$$

The action of these spin operators on the basis vectors are

$$
\begin{array}{lll}
S^+|\uparrow\rangle = 0\,, & S^+|\downarrow\rangle = |\uparrow\rangle\,, & S^3|\uparrow\rangle = \tfrac{1}{2}|\uparrow\rangle\,, \\
S^-|\downarrow\rangle = 0\,, & S^-|\uparrow\rangle = |\downarrow\rangle\,, & S^3|\downarrow\rangle = -\tfrac{1}{2}|\downarrow\rangle\,.
\end{array}
$$

This indices the action of the spin operators in the Hilbert space

$$
\begin{array}{lll}
S_k^+|\uparrow_k\rangle = 0\,, & S_k^+|\downarrow_k\rangle = |\uparrow_k\rangle\,, & S_k^3|\uparrow_k\rangle = \tfrac{1}{2}|\uparrow_k\rangle\,, \\
S_k^-|\downarrow_k\rangle = 0\,, & S_k^-|\uparrow_k\rangle = |\downarrow_k\rangle\,, & S_k^3|\downarrow_k\rangle = -\tfrac{1}{2}|\downarrow_k\rangle\,.
\end{array}
$$

The Hamiltonian can be then written as

$$
H = -J\sum_{n=1}^{L}\tfrac{1}{2}(S_n^+S_{n+1}^- + S_n^-S_{n+1}^+) + S_n^3S_{n+1}^3\,,
$$

For $Lb = 2$ we have

$$
H = -J\Big(S^+\otimes S^- + S^-\otimes S^+ + 2S^3\otimes S^3\Big) = -J\begin{pmatrix} \frac{1}{2} & 0 & 0 & 0 \\ 0 & -\frac{1}{2} & 1 & 0 \\ 0 & 1 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 & \frac{1}{2} \end{pmatrix}\,.
$$

This matrix has three eigenvalues which are equal to $-\frac{1}{2}J$ and one which is $\frac{3}{2}J$. Three states

$$
\underbrace{v_{s=1}^{\rm hw} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}}_{\rm h.w.}\,, \qquad \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}\,, \qquad \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}
$$

corresponding to equal eigenvalues form a representation of $\mathfrak{su}(2)$ with spin $s = 1$ and the state

$$v^{\rm hw}_{s=0} = \underbrace{\begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \end{pmatrix}}_{\rm h.w.}$$

which corresponds to $\frac{3}{2}J$ is a singlet of $\mathfrak{su}(2)$. Indeed, the generators of the global $\mathfrak{su}(2)$ are realized as

$$S^+ = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad S^- = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix}, \qquad S^3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

The vectors $v^{\rm hw}_{s=1}$ and $v^{\rm hw}_{s=0}$ are the highest-weight vectors of the $s = 1$ and $s = 0$ representations respectively, because they are annihilated by S^+ and are eigenstates of S^3. In fact, $v^{\rm hw}_{s=0}$ is also annihilated by S^- which shows that this state has zero spin. Thus, we completely understood the structure of the Hilbert space for $L = 2$.

In general, the Hamiltonian can be realized as $2^L \times 2^L$ symmetric matrix which means that it has a complete orthogonal system of eigenvectors. The Hilbert space split into sum of irreducible representations of $\mathfrak{su}(2)$. Thus, for L being finite the problem of finding the eigenvalues of H reduces to the problem of diagonalizing a symmetric $2^L \times 2^L$ matrix. This can be easily achieved by computer provided L is sufficiently small. However, for the physically interesting regime $L \to \infty$ corresponding to the *thermodynamic limit* new analytic methods are required.

In what follows it is useful to introduce the following operator:

$$P = \frac{1}{2}\Big(\mathbb{I} \otimes \mathbb{I} + \sum_\alpha \sigma^\alpha \otimes \sigma^\alpha\Big) = 2\Big(\frac{1}{4}\mathbb{I} \otimes \mathbb{I} + \sum_\alpha S^\alpha \otimes S^\alpha\Big)$$

which acts on $\mathbb{C}^2 \otimes \mathbb{C}^2$ as the permutation: $P(a \otimes b) = b \otimes a$. Indeed, we have

It is appropriate to call S^3 the operator of the total spin. On a state $|\psi\rangle$ with M spins down we have

$$S^3|\psi\rangle = \Big(\frac{1}{2}(L - M) - \frac{1}{2}M\Big)|\psi\rangle = \Big(\frac{1}{2}L - M\Big)|\psi\rangle\,.$$

Since $[H, S^3] = 0$ the Hamiltonian can be diagonalized within each subspace of the full Hilbert space with a given total spin (which is uniquely characterized by the number of spins down).

Let $M < L$ be a number of overturned spins. If $M = 0$ we have a unique state

$$|F\rangle = |\uparrow \cdots \uparrow\rangle.$$

This state is an eigenstate of the Hamiltonian with the eigenvalue $E_0 = -\frac{JL}{4}$:

$$H|F\rangle = -J\sum_{n=1}^{L} S^3_n S^3_{n+1}|\uparrow\cdots\uparrow\rangle = -\frac{JL}{4}|\uparrow\cdots\uparrow\rangle\,.$$

Let M be arbitrary. Since the M-th space has the dimension $\frac{L!}{(L-M)!M!}$ one should find the same number of eigenvectors of H in this subspace. So let us write the eigenvectors of H in the form

$$|\psi\rangle = \sum_{1\leq n_1<\cdots<n_M\leq L} a(n_1,\ldots,n_M)|n_1,\ldots,n_M\rangle$$

with some unknown coefficients $a(n_1,\ldots,n_M)$. Here

$$|n_1,\ldots,n_M\rangle = S^-_{n_1}S^-_{n_2}\ldots S^-_{n_M}|F\rangle$$

and non-coincident integers describe the positions of the overturned spins. Obviously, the coefficients $a(n_1,\ldots,n_M)$ must satisfy the following requirement of periodicity:

$$a(n_2,\ldots,n_M,n_1+N) = a(n_1,\ldots,n_M)\,.$$

The coordinate Bethe ansatz postulates the form of these coefficients (Hans Bethe, 1931)

$$a(n_1,\ldots,n_M) = \sum_{\pi\in S_M} A_\pi \exp\Big(i\sum_{j=1}^{M} p_{\pi(j)}n_j\Big)\,.$$

Here for each of the M overturned spins we introduced the variable p_j which is called *pseudo-momentum* and S_M denotes the permutation group over the labels $\{1,\ldots,M\}$. To determine the coefficients A_π as well as the set of pseudo-momenta $\{p_j\}$ we have to use the eigenvalue equation for H and the periodicity condition for $a(n_1,\ldots,n_M)$. It is instructive to work in detail the cases $M=1$ and $M=2$ first.

For $M=1$ case we have

$$|\psi\rangle = \sum_{n=1}^{L} a(n)|n\rangle\,, \qquad a(n) = Ae^{ipn}\,.$$

Thus, in this case

$$|\psi\rangle = A\sum_{n=1}^{L} e^{ipn}|n\rangle$$

is nothing else but the Fourier transform. The periodicity condition leads to determination of the pseudo-momenta

$$a(n+L) = a(n) \qquad \Longrightarrow \qquad e^{ipL} = 1\,,$$

i.e. the $\frac{L!}{(L-1)!1!} = L$ allowed values of the pseudo-momenta are

$$p = \frac{2\pi k}{L} \qquad \text{with} \qquad k = 0, \cdots, L-1\,.$$

Further, we have the eigenvalue equation

$$H|\psi\rangle = -\frac{JA}{2}\sum_{m,n=1}^{L} e^{ipm}\Big[S_n^+S_{n+1}^- + S_n^-S_{n+1}^+ + 2S_n^3S_{n+1}^3\Big]\,|m\rangle = E(p)|\psi\rangle\,.$$

To work out the l.h.s. we have to use the formulae

$$S_n^+S_{n+1}^-|m\rangle = \delta_{nm}|m+1\rangle\,, \qquad S_n^-S_{n+1}^+|m\rangle = \delta_{n+1,m}|m-1\rangle$$

as well as

$$\begin{aligned} 2S_n^3S_{n+1}^3|m\rangle &= \frac{1}{2}|m\rangle\,, \qquad \text{for} \quad m \neq n, n+1\,,\\ 2S_n^3S_{n+1}^3|m\rangle &= -\frac{1}{2}|m\rangle\,, \quad \text{for} \quad m=n, \quad \text{or} \quad m = n+1\,. \end{aligned}$$

Taking this into account we obtain

$$\begin{aligned} H|\psi\rangle \;=\; -\frac{JA}{2}\Big[\sum_{n=1}^{L}\Big(e^{ipn}|n+1\rangle + e^{ip(n+1)}|n\rangle\Big) + \frac{1}{2}\sum_{m=1}^{L}\Big(\sum_{\substack{n=1\\ n\neq m,m-1}}^{L}\Big)e^{ipm}|m\rangle \\ -\frac{1}{2}\sum_{n=1}^{L} e^{ipn}|n\rangle - \frac{1}{2}\sum_{n=1}^{L} e^{ip(n+1)}|n+1\rangle\Big]\,. \end{aligned}$$

Using periodicity conditions we finally get

$$H|\psi\rangle \;=\; -\frac{JA}{2}\sum_{n=1}^{L}\Big(e^{ip(n-1)} + e^{ip(n+1)} + \frac{L-4}{2}e^{ipn}\Big)|n\rangle = -\frac{J}{2}\Big(e^{-ip} + e^{ip} + \frac{L-4}{2}\Big)|\psi\rangle\,.$$

From here we read off the eigenvalue

$$E - E_0 = J(1-\cos p) = 2J\sin^2\frac{p}{2}\,,$$

where $E_0 = -\frac{JL}{4}$. Excitation of the spin chain around the pseudo-vacuum $|F\rangle$ carrying the pseudo-momentum p is called a *magnon*[12]. Thus, magnon can be viewed

[12]The concept of a magnon was introduced in 1930 by Felix Bloch in order to explain the reduction of the spontaneous magnetization in a ferromagnet. At absolute zero temperature, a ferromagnet reaches the state of lowest energy, in which all of the atomic spins (and hence magnetic moments) point in the same direction. As the temperature increases, more and more spins deviate randomly from the common direction, thus increasing the internal energy and reducing the net magnetization. If one views the perfectly magnetized state at zero temperature as the vacuum state of the ferromagnet, the low-temperature state with a few spins out of alignment can be viewed as a gas of quasiparticles, in this case magnons. Each magnon reduces the total spin along the direction of magnetization by one unit of and the magnetization itself by , where g is the gyromagnetic ratio. The quantitative theory of quantized spin waves, or magnons, was developed further by Ted Holstein and Henry Primakoff (1940) and Freeman Dyson (1956). By using the formalism of second quantization they showed that the magnons behave as weakly interacting quasiparticles obeying the Bose-Einstein statistics (the bosons).

as the pseudo-particle with the momentum $p = \frac{2\pi k}{L}$, $k = 0, \ldots, L-1$ and the energy

$$E = 2J \sin^2 \frac{p}{2} \, .$$

The last expression is the dispersion relation for one-magnon states.

Let us comment on the sign of the coupling constant. If $J < 0$ then $E_k < 0$ and $|F\rangle$ is not the ground state, i.e. a state with the lowest energy. In other words, in this case, $|F\rangle$ is not a vacuum, but rather a pseudo-vacuum, or "false" vacuum. The true ground state in non-trivial and needs some work to be identified. The case $J < 0$ is called the anti-ferromagnetic one. Oppositely, if $J > 0$ then $|F\rangle$ is a state with the lowest energy and, therefore, is the true vacuum. Later on we will see that the anti-ferromagnetic ground state corresponds $M = \frac{1}{2}L$ and, therefore, it is spinless. The ferromagnetic ground state corresponds to $M = 0$ and, therefore, carries maximal spin $S^3 = \frac{1}{2}L$.[13]

Let us now turn to the more complicated case $M = 2$. Here we have

$$|\psi\rangle = \sum_{1 \le n_1 < n_2 \le L} a(n_1, n_2)|n_1, n_2\rangle \, ,$$

where

$$a(n_1, n_2) = A e^{i(p_1 n_1 + p_2 n_2)} + B e^{i(p_2 n_1 + p_1 n_2)} \, .$$

The eigenvalue equation for H imposes conditions on $a(n_1, n_2)$ analogous to the $M = 1$ case. Special care is needed, however, when two overturned spins are sitting next to each other. Thus, we are led to consider

$$\begin{aligned}
H|\psi\rangle &= -\frac{J}{2} \sum_{1 \le n_1 < n_2 \le L} a(n_1, n_2) \sum_{m=1}^{L} \Big[S_m^+ S_{m+1}^- + S_m^- S_{m+1}^+ + 2 S_m^3 S_{m+1}^3 \Big] |n_1, n_2\rangle \\
&= \Bigg\{ -\frac{J}{2} \Big[\sum_{\substack{1 \le n_1 < n_2 \le L \\ n_2 > n_1 + 1}} a(n_1, n_2) \Big(|n_1 + 1, n_2\rangle + |n_1, n_2 + 1\rangle + |n_1 - 1, n_2\rangle + |n_1, n_2 - 1\rangle \Big) \\
&\qquad + \frac{L-4}{2} \sum_{\substack{1 \le n_1 < n_2 \le L \\ n_2 > n_1 + 1}} a(n_1, n_2)|n_1, n_2\rangle - \frac{1}{2} 4 \sum_{\substack{1 \le n_1 < n_2 \le L \\ n_2 > n_1 + 1}} a(n_1, n_2)|n_1, n_2\rangle \Big] \Bigg\} + \\
&+ \Bigg\{ -\frac{J}{2} \sum_{1 \le n_1 \le L} a(n_1, n_1 + 1) \Big[|n_1, n_1 + 2\rangle + |n_1 - 1, n_1 + 1\rangle + \big(\frac{L-2}{2} - 1\big) |n_1, n_1 + 1\rangle \Big] \Bigg\} \, .
\end{aligned}$$

[13]Many crystals possess the ordered magnetic structure. This means that in absence of external magnetic field the averaged quantum-mechanical magnetic moment in each elementary crystal cell is different from zero. In the ferromagnetic crystals (Fe, Ni, Co) the averaged values of magnetic moments of all the atoms have the same orientation unless the temperature does not exceed a certain critical value called the Curie temperature. Due to this, ferromagnets have a spontaneous magnetic moment, i.e. a macroscopic magnetic moment different from zero in the vanishing external field. In more complicated anti-ferromagnetic crystals (carbons, sulfates, oxides) the averaged values of magnetic moments of individual atoms compensate each other within every elementary crystal cell.

Here in the first bracket we consider the terms with $n_2 > n_1+1$, while the last bracket represents the result of action of H on terms with $n_2 = n_1 + 1$. Using periodicity conditions we are allowed to make shifts of the summation variables n_1, n_2 in the first bracket to bring all the states to the uniform expression $|n_1, n_2\rangle$. We therefore get

$$H|\psi\rangle = -\frac{J}{2}\Bigg\{\sum_{n_2>n_1} a(n_1-1,n_2)|n_1,n_2\rangle + \sum_{n_2>n_1+2} a(n_1,n_2-1)|n_1,n_2\rangle$$
$$+\sum_{n_2>n_1+2} a(n_1+1,n_2)|n_1,n_2\rangle + \sum_{n_2>n_1} a(n_1,n_2+1)|n_1,n_2\rangle + \frac{L-8}{2}\sum_{n_2>n_1+1} a(n_1,n_2)|n_1,n_2\rangle\Bigg\}$$
$$-\frac{J}{2}\Bigg\{\sum_{1\le n_1\le L} a(n_1,n_1+1)\Big[|n_1,n_1+2\rangle + |n_1-1,n_1+1\rangle + \frac{L-4}{2}|n_1,n_1+1\rangle\Big]\Bigg\}\,.$$

Now we complete the sums in the first bracket to run the range $n_2 > n_1$. This is achieved by adding and subtracting the missing terms. As the result we will get

$$H|\psi\rangle =$$
$$-\frac{J}{2}\Bigg\{\sum_{n_2>n_1}\Big(a(n_1-1,n_2) + a(n_1,n_2-1) + a(n_1+1,n_2) + a(n_1,n_2+1) + \frac{L-8}{2}a(n_1,n_2)\Big)|n_1,n_2\rangle$$
$$-\sum_{1\le n_1\le L}\Big(a(n_1,n_1)|n_1,n_1+1\rangle + a(n_1+1,n_1+1)|n_1,n_1+1\rangle +$$
$$+\underbrace{a(n_1,n_1+1)|n_1,n_1+2\rangle} + \underbrace{a(n_1,n_1+2)|n_1,n_1+2\rangle} + \frac{L-8}{2}a(n_1,n_1+1)|n_1,n_1+1\rangle\Big)\Bigg\}$$
$$-\frac{J}{2}\Bigg\{\sum_{1\le n_1\le L} a(n_1,n_1+1)\Big[\underbrace{|n_1,n_1+2\rangle + |n_1-1,n_1+1\rangle} + \frac{L-4}{2}|n_1,n_1+1\rangle\Big]\Bigg\}\,.$$

The underbraced terms cancel out and we finally get

$$H|\psi\rangle =$$
$$-\frac{J}{2}\Bigg\{\sum_{n_2>n_1}\Big(a(n_1-1,n_2) + a(n_1,n_2-1) + a(n_1+1,n_2) + a(n_1,n_2+1) + \frac{L-8}{2}a(n_1,n_2)\Big)|n_1,n_2\rangle\Bigg\}$$
$$+\frac{J}{2}\Bigg\{\sum_{1\le n_1\le L}\Big(a(n_1,n_1) + a(n_1+1,n_1+1) - 2a(n_1,n_1+1)\Big)|n_1,n_1+1\rangle\Bigg\}\,.$$

If we impose the requirement that

$$a(n_1,n_1) + a(n_1+1,n_1+1) - 2a(n_1,n_1+1) = 0 \tag{5.1}$$

then the second bracket in the eigenvalue equation vanishes and the eigenvalue problem reduces to the following equation

$$2(E-E_0)a(n_1,n_2) = J\Big[4a(n_1,n_2) - \sum_{\sigma=\pm1} a(n_1+\sigma,n_2) + a(n_1,n_2+\sigma)\Big]\,. \tag{5.2}$$

Substituting in eq.(5.1) the Bethe ansatz for $a(n_1, n_2)$ we get

$$Ae^{(p_1+p_2)n} + Be^{i(p_1+p_2)n} + Ae^{(p_1+p_2)(n+1)} + Be^{i(p_1+p_2)(n+1)}$$
$$-2\Big(Ae^{i(p_1n+p_2(n+1))} + Be^{i(p_2n+p_1(n+1))}\Big) = 0\,.$$

This allows one to determine the ratio

$$\frac{B}{A} = -\frac{e^{i(p_1+p_2)} + 1 - 2e^{ip_2}}{e^{i(p_1+p_2)} + 1 - 2e^{ip_1}}\,.$$

Problem. Show that for real values of momenta the ratio $\frac{B}{A}$ is the pure phase:

$$\frac{B}{A} = e^{i\theta(p_2,p_1)} \equiv S(p_2,p_1)\,.$$

This phase is called the S-matrix. We further note that it obeys the following relation

$$S(p_1,p_2)S(p_2,p_1) = 1\,.$$

Thus, the two-magnon Bethe ansatz takes the form

$$a(n_1,n_2) = e^{i(p_1n_1+p_2n_2)} + S(p_2,p_1)e^{i(p_2n_1+p_1n_2)}\,,$$

where we factored out the unessential normalization coefficient A.

Let us now substitute the Bethe ansatz in eq.(5.2). We get

$$\begin{aligned}
2(E-E_0)\Big(Ae^{i(p_1n_1+p_2n_2)} + Be^{i(p_2n_1+p_1n_2)}\Big) &= J\Big[4\Big(Ae^{i(p_1n_1+p_2n_2)} + Be^{i(p_2n_1+p_1n_2)}\Big) - \\
&- \Big(Ae^{i(p_1n_1+p_2n_2)}e^{ip_1} + Be^{i(p_2n_1+p_1n_2)}e^{ip_2}\Big) - \Big(Ae^{i(p_1n_1+p_2n_2)}e^{-ip_1} + Be^{i(p_2n_1+p_1n_2)}e^{-ip_2}\Big) \\
&- \Big(Ae^{i(p_1n_1+p_2n_2)}e^{ip_2} + Be^{i(p_2n_1+p_1n_2)}e^{ip_1}\Big) - \Big(Ae^{i(p_1n_1+p_2n_2)}e^{-ip_2} + Be^{i(p_2n_1+p_1n_2)}e^{-ip_1}\Big)\Big]\,.
\end{aligned}$$

We see that the dependence on A and B cancel out completely and we get the following equation for the energy

$$E - E_0 = J\Big(2 - \cos p_1 - \cos p_2\Big) = 2J\sum_{k=1}^{2}\sin^2\frac{p_k}{2}\,.$$

Quite remarkably, the energy appears to be additive, i.e. the energy of a two-magnon state appears to be equal to the sum of energies of one-magnon states! This shows that magnons essentially behave themselves as free particles in the box.

Finally, we have to impose the periodicity condition $a(n_2, n_1+L) = a(n_1,n_2)$. This results into

$$e^{i(p_1n_2+p_2n_1)}e^{ip_2L} + \frac{B}{A}e^{ip_1L}e^{i(p_2n_2+p_1n_1)} = e^{i(p_1n_1+p_2n_2)} + \frac{B}{A}e^{i(p_2n_1+p_1n_2)}$$

which implies

$$e^{ip_1L} = \frac{A}{B} = S(p_1,p_2)\,, \qquad e^{ip_2L} = \frac{B}{A} = S(p_2,p_1)\,.$$

The last equations are called "Bethe equations". *They are nothing else but the quantization conditions for momenta p_k.*

Let us note the following useful representation for the S-matrix.

We have

$$
\begin{aligned}
S(p_2,p_1) &= -\frac{e^{ip_2}\big(e^{ip_1}-1\big)+1-e^{ip_2}}{e^{ip_1}\big(e^{ip_2}-1\big)+1-e^{ip_1}} = -\frac{e^{ip_2}e^{\frac{i}{2}p_1}\big(e^{\frac{i}{2}p_1}-e^{-\frac{i}{2}p_1}\big)+e^{\frac{i}{2}p_2}\big(e^{-\frac{i}{2}p_2}-e^{\frac{i}{2}p_2}\big)}{e^{ip_1}e^{\frac{i}{2}p_2}\big(e^{\frac{i}{2}p_2}-e^{-\frac{i}{2}p_2}\big)+e^{\frac{i}{2}p_1}\big(e^{-\frac{i}{2}p_1}-e^{\frac{i}{2}p_1}\big)} \\
&= -\frac{e^{\frac{i}{2}p_2}\sin\frac{p_1}{2}-e^{-\frac{i}{2}p_1}\sin\frac{p_2}{2}}{e^{\frac{i}{2}p_1}\sin\frac{p_2}{2}-e^{-\frac{i}{2}p_2}\sin\frac{p_1}{2}} = \frac{\big(\cos\frac{p_2}{2}+i\sin\frac{p_2}{2}\big)\sin\frac{p_1}{2}-\big(\cos\frac{p_1}{2}-i\sin\frac{p_1}{2}\big)\sin\frac{p_2}{2}}{\big(\cos\frac{p_1}{2}+i\sin\frac{p_1}{2}\big)\sin\frac{p_2}{2}-\big(\cos\frac{p_2}{2}-i\sin\frac{p_2}{2}\big)\sin\frac{p_1}{2}} \\
&= -\frac{\cos\frac{p_2}{2}\sin\frac{p_1}{2}-\cos\frac{p_1}{2}\sin\frac{p_2}{2}+2i\sin\frac{p_1}{2}\sin\frac{p_2}{2}}{\cos\frac{p_1}{2}\sin\frac{p_2}{2}-\cos\frac{p_2}{2}\sin\frac{p_1}{2}+2i\sin\frac{p_1}{2}\sin\frac{p_2}{2}} = \frac{\frac{1}{2}\cot\frac{p_2}{2}-\frac{1}{2}\cot\frac{p_1}{2}+i}{\frac{1}{2}\cot\frac{p_2}{2}-\frac{1}{2}\cot\frac{p_1}{2}-i}\,.
\end{aligned}
$$

Thus, we obtained

$$
S(p_1,p_2) = \frac{\frac{1}{2}\cot\frac{p_1}{2}-\frac{1}{2}\cot\frac{p_2}{2}+i}{\frac{1}{2}\cot\frac{p_1}{2}-\frac{1}{2}\cot\frac{p_2}{2}-i}\,.
$$

It is therefore convenient to introduce the variable $\lambda = \frac{1}{2}\cot\frac{p}{2}$ which is called *rapidity* and get

$$
S(\lambda_1,\lambda_2) = \frac{\lambda_1-\lambda_2+i}{\lambda_1-\lambda_2-i}\,.
$$

Hence, on the rapidity plane the S-matrix depends only on the difference of rapidities of scattering particles.

Taking the logarithm of the Bethe equations we obtain

$$
Lp_1 = 2\pi m_1 + \theta(p_1,p_2)\,, \qquad Lp_2 = 2\pi m_2 + \theta(p_2,p_1)\,,
$$

where the integers $m_i \in \{0,1,\dots,L-1\}$ are called *Bethe quantum numbers*. The Bethe quantum numbers are useful to distinguish eigenstates with different physical properties. Furthermore, these equations imply that the total momentum is

$$
P = p_1 + p_2 = \frac{2\pi}{L}(m_1+m_2)\,.
$$

Writing the equations in the form

$$
p_1 = \underbrace{\frac{2\pi m_1}{L}} + \frac{1}{L}\theta(p_1,p_2)\,, \qquad p_2 = \underbrace{\frac{2\pi m_2}{L}} + \frac{1}{L}\theta(p_2,p_1)\,,
$$

we see that the magnon interaction is reflected in the phase shift θ and in the deviation of the momenta p_1, p_2 from the values of the underbraced one-magnon wave numbers. What is very interesting, as we will see, that *magnons either scatter off each other or form the bound states.*

The first problem is to find all possible Bethe quantum numbers (m_1, m_2) for which Bethe equations have solutions. The allowed pairs (m_1, m_2) are restricted to

$$0 \leq m_1 \leq m_2 \leq L - 1\,.$$

This is because switching m_1 and m_2 simply interchanges p_1 and p_2 and produces the same solution. There are $\frac{1}{2}L(L+1)$ pairs which meet this restriction but only $\frac{1}{2}L(L-1)$ of them yield a solution of the Bethe equations. Some of these solutions have real p_1 and p_2, the others yield the complex conjugate momenta $p_2 = p_1^*$.

The simplest solutions are the pairs for which one of the Bethe numbers is zero, e.g. $m_1 = 0$, $m = m_2 = 0, 1, \dots, L-1$. For such a pair we have

$$Lp_1 = \theta(p_1, p_2)\,, \qquad Lp_2 = 2\pi m + \theta(p_2, p_1)\,,$$

which is solved by $p_1 = 0$ and $p_2 = \frac{2\pi m}{L}$. Indeed, for $p_1 = 0$ the phase shift vanishes: $\theta(0, p_2) = 0$. These solutions have the dispersion relation

$$E - E_0 = 2J \sin^2 \frac{p}{2}\,, \qquad p = p_2$$

which is the same as the dispersion for the one-magnon states. These solutions are nothing else but $\mathfrak{su}(2)$-descendants of the solutions with $M = 1$.

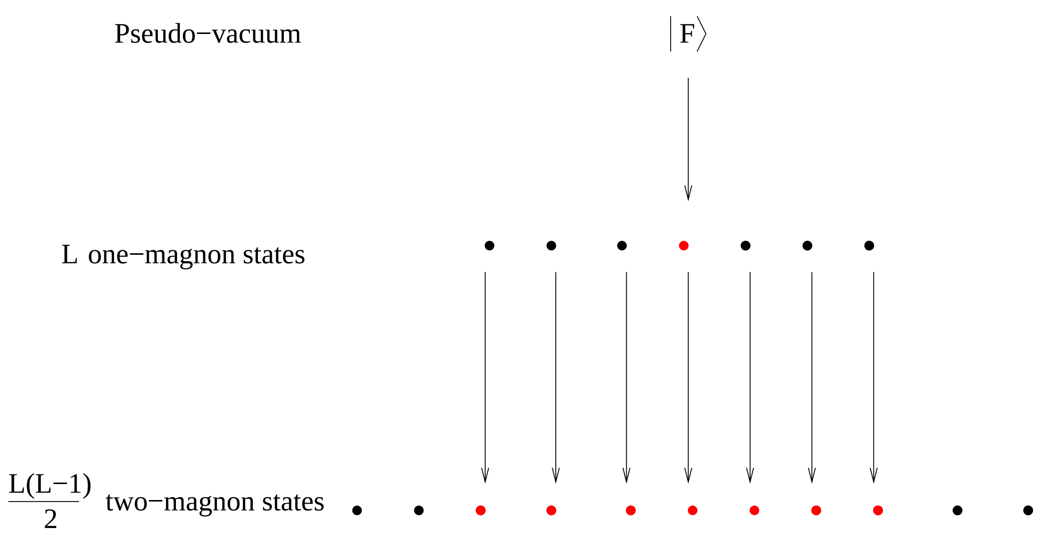
One can show that for $M = 2$ all solutions are divided into three distinct classes

$$\underbrace{\text{Descendents}}_{L}\,, \qquad \underbrace{\text{Scattering States}}_{\frac{L(L-5)}{2}+3}\,, \qquad \underbrace{\text{Bound States}}_{L-3}$$

so that

$$L + \frac{L(L-5)}{2} + 3 + L - 3 = \frac{1}{2}L(L-1)$$

gives a complete solution space of the two-magnon problem.



The $\mathfrak{su}(2)$-multiplet structure of the $M=0,1,2$ subspaces.

The most non-trivial fact about the Bethe ansatz is that many-body (multi-magnon) problem reduces to the two-body one. It means, in particular, that the multi-magnon S-matrix appears to be expressed as the product of the two-body ones. Also the energy is additive quantity. Such a particular situation is spoken about as "Factorized Scattering". In a sense, factorized scattering for the quantum many-body system is the same as integrability because it appears to be a consequence of existence of additional conservation laws. For the M-magnon problem the Bethe equations read

$$e^{ip_k L} = \prod_{\substack{j=1 \\ j\neq k}}^{M} \mathrm{S}(p_k, p_j)\,.$$

The most simple description of the bound states is obtained in the limit when $L\to\infty$. If p_k has a non-trivial positive imaginary part then e^{ip_kL} tends to ∞ and this means that the bound states correspond in this limit to poles of the r.h.s. of the Bethe equations. In particular, for the case $M=2$ the bound states correspond to poles in the two-body S-matrix. In particular, we find such a pole when

$$\tfrac{1}{2}\cot\tfrac{p_1}{2} - \tfrac{1}{2}\cot\tfrac{p_2}{2} = i\,.$$

This state has the total momentum $p = p_1 + p_2$ which must be real. These conditions can be solved by taking

$$p_1 = \frac{p}{2} + iv\,, \qquad p_2 = \frac{p}{2} - iv\,.$$

The substitution gives

$$\cos\tfrac{1}{2}\Big(\tfrac{p}{2}+iv\Big)\sin\tfrac{1}{2}\Big(\tfrac{p}{2}-iv\Big) - \cos\tfrac{1}{2}\Big(\tfrac{p}{2}-iv\Big)\sin\tfrac{1}{2}\Big(\tfrac{p}{2}+iv\Big) = 2i\sin\tfrac{1}{2}\Big(\tfrac{p}{2}+iv\Big)\sin\tfrac{1}{2}\Big(\tfrac{p}{2}-iv\Big)\,,$$

which is

$$\cos\frac{p}{2} = e^{v}\,.$$

The energy of such a state is

$$E = 2J\Big(\sin^2\frac{p_1}{2} + \sin^2\frac{p_2}{2}\Big) = 2J\Big(\sin^2\Big(\frac{p}{4}+i\frac{v}{2}\Big) + \sin^2\Big(\frac{p}{4}-i\frac{v}{2}\Big)\Big)\,.$$

We therefore get

$$E = 2J\Big(1-\cos\frac{p}{2}\cosh v\Big) = 2J\Big(1-\cos\frac{p}{2}\frac{\cos^2\frac{p}{2}+1}{2\cos\frac{p}{2}}\Big) = J\sin^2\frac{p}{2}\,.$$

Thus, the position of the pole uniquely fixes the dispersion relation of the bound state.

5.2 Algebraic Bethe Ansatz

Here we will solve the Heisenberg model by employing this time a new method called the Algebraic Bethe ansatz. This method allows one to reveal the integrable structure of the model as well as to study its properties in the thermodynamic limit.

Fundamental commutation relation. Suppose we have a periodic chain of length L. The basic tool of the algebraic Bethe ansatz approach is the so-called Lax operator. The definition of the Lax operator involves the local "quantum" space V_i, which for the present case is chosen to be a copy of $\mathbb{C}^2$. The Lax operator $L_{i,a}$ acts in $V_i \otimes V_a$:

$$L_{i,a}(\lambda): \qquad V_i \otimes V_a \to V_i \otimes V_a \,.$$

Explicitly, it is given by

$$L_{i,a}(\lambda) = \lambda \mathbb{I}_i \otimes \mathbb{I}_a + i \sum_\alpha S_i^\alpha \otimes \sigma^\alpha \,,$$

where $\mathbb{I}_i$, S_i^α act in V_i, while the unit $\mathbb{I}_a$ and the Pauli matrices σ^α act in an another Hilbert space $\mathbb{C}^2$ called "auxiliary". The parameter λ is called the *spectral parameter*. Another way to represent that the Lax operator is to write it as 2×2 matrix with operator coefficients

$$L_{i,a}(\lambda) = \begin{pmatrix} \lambda + iS_i^3 & iS_i^- \\ iS_i^+ & \lambda - iS_i^3 \end{pmatrix} \,.$$

Introducing the permutation operator

$$\mathrm{P} = \frac{1}{2}\Big(\mathbb{I} \otimes \mathbb{I} + \sum_{\alpha=1}^{3} \sigma^\alpha \otimes \sigma^\alpha\Big)$$

we can write the Lax operator in the alternative form

$$L_{i,a}(\lambda) = \Big(\lambda - \frac{i}{2}\Big)\mathbb{I}_{i,a} + i\mathrm{P}_{i,a} \,.$$

The most important property of the Lax operator is the commutation relations between its entries. Consider two Lax operators, $L_{i,a}(\lambda_1)$ and $L_{i,b}(\lambda_2)$, acting in the same quantum space but in two different auxiliary spaces. The products of these two operators $L_{i,a}(\lambda_1)L_{i,b}(\lambda_2)$ and $L_{i,b}(\lambda_2)L_{i,a}(\lambda_1)$ are defined in the triple tensor product $V_i \otimes V_a \otimes V_b$. Remarkably, it turns out that these two product are related by a similarity transformation which acts non-trivially in the tensor product $V_a \otimes V_b$ only. Namely, there exists an intertwining operator $R_{a,b}(\lambda_1, \lambda_2) = R_{ab}(\lambda_1 - \lambda_2)$ such that the following relation is true

$$R_{ab}(\lambda_1 - \lambda_2)L_{ia}(\lambda_1)L_{ib}(\lambda_2) = L_{ib}(\lambda_2)L_{ia}(\lambda_1)R_{ab}(\lambda_1 - \lambda_2) \,. \tag{5.3}$$

This intertwining operator is called *quantum R-matrix* and it has the following explicit form

$$R_{ab} = \lambda \mathbb{I}_{ab} + i\mathrm{P}_{ab}$$

The form of the L-operator and the R-matrix is essentially the same.

We check

$$\Big((\lambda_1-\lambda_2)\mathbb{I}_{ab}+i\mathrm{P}_{ab}\Big)L_{ia}(\lambda_1)L_{ib}(\lambda_2)=L_{ib}(\lambda_2)L_{ia}(\lambda_1)\Big((\lambda_1-\lambda_2)\mathbb{I}_{ab}+i\mathrm{P}_{ab}\Big)\,,$$

which leads to

$$(\lambda_1-\lambda_2)\Big(L_{ia}(\lambda_1)L_{ib}(\lambda_2)-L_{ib}(\lambda_2)L_{ia}(\lambda_1)\Big)=i\mathrm{P}_{ab}\Big(L_{ia}(\lambda_2)L_{ib}(\lambda_1)-L_{ia}(\lambda_1)L_{ib}(\lambda_2)\Big)\,,$$

It is easy to see that

$$L_{ia}(\lambda_1)L_{ib}(\lambda_2)-L_{ib}(\lambda_2)L_{ia}(\lambda_1)=\mathrm{P}_{ib}\mathrm{P}_{ia}-\mathrm{P}_{ia}\mathrm{P}_{ib}$$

and

$$i\mathrm{P}_{ab}\Big(L_{ia}(\lambda_2)L_{ib}(\lambda_1)-L_{ia}(\lambda_1)L_{ib}(\lambda_2)\Big)=(\lambda_1-\lambda_2)\mathrm{P}_{ab}(\mathrm{P}_{ib}-\mathrm{P}_{ia})=(\lambda_1-\lambda_2)(\mathrm{P}_{ib}\mathrm{P}_{ai}-\mathrm{P}_{ia}\mathrm{P}_{ib})$$

This proves the statement.

The relation (5.3) is called the fundamental commutation relation.

Yang-Baxter equation. It is convenient to suppress the index of the quantum space and write the fundamental commutation relation as

$$R_{ab}(\lambda_1-\lambda_2)L_a(\lambda_1)L_b(\lambda_2)=L_b(\lambda_2)L_a(\lambda_1)R_{ab}(\lambda_1-\lambda_2)\,.$$

We can think about L as being 2×2 matrix whose matrix elements are generators of a certain associative algebra (operators). Relations (5.3) define the then the commutation relations between the generators of this algebra. Substituting the indices a and b for 1 and 2 we will write the general form of the fundamental commutation relations

$$R_{12}(\lambda_1,\lambda_2)L_1(\lambda_1)L_2(\lambda_2)=L_2(\lambda_2)L_1(\lambda_1)R_{12}(\lambda_1,\lambda_2)\,.$$

What the R-matrix does is that it interchange the position of the matrices L_1 and L_2. Consider a triple product

$$\begin{aligned}L_1L_2L_3&=R_{12}^{-1}L_2L_1R_{12}L_3=R_{12}^{-1}L_2L_1L_3R_{12}=\\&=R_{12}^{-1}R_{13}^{-1}L_2L_3L_1R_{13}R_{12}=R_{12}^{-1}R_{13}^{-1}R_{23}^{-1}L_3L_2L_1R_{23}R_{13}R_{12}\,.\end{aligned}$$

Essentially, we brought the product $L_1L_2L_3$ to the form $L_3L_2L_1$. However, we can reach the same effect by changing the order of permutations

$$\begin{aligned}L_1L_2L_3&=R_{23}^{-1}L_1L_3L_2R_{23}=R_{23}^{-1}R_{13}^{-1}L_3L_1L_2R_{13}R_{12}=\\&=R_{12}^{-1}R_{13}^{-1}L_2L_3L_1R_{13}R_{12}=R_{23}^{-1}R_{13}^{-1}R_{12}^{-1}L_3L_2L_1R_{12}R_{13}R_{23}\,.\end{aligned}$$

Thus, if we require that we do not generate new triple relations between the elements of L we should impose the following condition on the R-matrix:

$$R_{12}R_{13}R_{23}=R_{23}R_{13}R_{12}\,. \tag{5.4}$$

This is the *quantum Yang-Baxter equation.*

Semi-classical limit and quantization. Why the quantum Yang-Baxter equation is called "quantum"? Assume that the R-matrix depends on the additional parameter $\hbar$ and when $\hbar \to 0$ it expands into the power series starting with the unit:

$$R_{12} = \mathbb{I}_{12} + \hbar r_{12} + \cdots$$

Expanding quantum Yang-Baxter equation we see that the leading terms as well as the terms proportional to $\hbar$ cancel out. At order $\hbar^2$ we find

$$\hbar^2\Big([r_{12}, r_{13}] + [r_{12}, r_{23}] + [r_{13}, r_{23}]\Big) + \mathcal{O}(\hbar^3) = 0\,.$$

At order $\hbar^2$ we find the *classical Yang-Baxter equation*. Thus, the quantum Yang-Baxter equation can be considered as the deformation (or quantization) of the classical Yang-Baxter equation. Further we recall the relation between the Poisson bracket of classical observables and the Poisson brachet of their quantum counterparts

$$\{A, B\} = \lim_{\hbar\to 0}\frac{1}{\hbar}[\hat{A}, \hat{B}]$$

Now we notice that the fundamental computation relations can be written in the equivalent form

$$[L_1(\lambda_1), L_2(\lambda_2)] = \frac{i\hbar}{\lambda_1 - \lambda_2}[\mathrm{P}_{12}, L_1(\lambda_1)L_2(\lambda_2)]\,.$$

This formula allows for the semi-classical limit

$$\underbrace{L}_{\text{quantum}} \to \underbrace{\mathscr{L}}_{\text{classical}}$$

and it defines the Poisson bracket on the space of classical $\mathscr{L}$-operators

$$\{\mathscr{L}_1(\lambda_1), \mathscr{L}_2(\lambda_2)\} = \lim_{\hbar\to 0}\frac{1}{\hbar}[\mathscr{L}_1(\lambda_1), \mathscr{L}_2(\lambda_2)] = \Big[\frac{i}{\lambda_1 - \lambda_2}\mathrm{P}_{12}, \mathscr{L}_1(\lambda_1)\mathscr{L}_2(\lambda_2)\Big]\,.$$

We see that $r = \frac{i}{\lambda}\mathrm{P}$ appears to be the classical r-matrix. Thus, the semi-classical limit of the fundamental commutation relations is nothing else as the Sklyanin bracket for classical Heisengerg magnetic. Inversely, we can think about the fundamental commutation relations as quantization of the Poisson algebra of the classical L-operators.

Monodromy and transfer matrix. For a chain of length L define the monodromy as the ordered product of L-operators along the chain[14]

$$T_a(\lambda) = L_{L,a}(\lambda)\ldots L_{1,a}(\lambda).$$

[14]Recall the definition of the monodromy as the path-ordered exponent in the classical case.

The monodromy is an operator on $V_L \otimes V_{L-1} \otimes \ldots \otimes V_1 \otimes V_a$. It we take the trace of the monodromy w.r.t. to its matrix part acting in the auxiliary space we obtain an object which is called the *transfer matrix* and it is denoted as $\tau(\lambda) = \mathrm{tr}_a T_a(\lambda)$. Denote $L = L_{i,a}$ and $L' = L_{i+1,a}$

$$R_{12}L'_1L_1L'_2L_2 = R_{12}L'_1L'_2L_1L_2 = L'_2L'_1R_{12}L_1L_2 = L'_2L'_1L_2L_1R_{12} = L'_2L_2L'_1L_1R_{12}\,.$$

This is because L_1 and L'_2 commute – they act both in different auxiliary spaces and different quantum spaces. Thus, we deduce the commutation relation between the components of the monodromy

$$R_{12}(\lambda-\mu)T_1(\lambda)T_2(\mu) = T_2(\mu)T_1(\lambda)R_{12}(\lambda-\mu)\,.$$

Now we can proof the fundamental fact about the commutation relations above. Rewrite them in the form

$$T_1(\lambda)T_2(\mu) = R_{12}(\lambda-\mu)^{-1}T_2(\mu)T_1(\lambda)R_{12}(\lambda-\mu)$$

and takes the trace over the first and the second space. We will get

$$\tau(\lambda)\tau(\mu) = \mathrm{tr}_{1,2}\Big(R_{12}(\lambda-\mu)^{-1}T_2(\mu)T_1(\lambda)R_{12}(\lambda-\mu)\Big) = \tau(\mu)\tau(\lambda)\,.$$

Thus, the transfer matrices commute with each other for different values of the spectral parameter

$$[\tau(\lambda), \tau(\mu)] = 0\,.$$

Hence, $\tau(\lambda)$ generates an abelian subalgebra. If we find the Hamiltonian of the model among this commuting family then we can call our model quantum integrable. The Hamiltonian must be

$$H = \sum_{a,k} c_{ka}\frac{d^k}{d\lambda^k}\ln\tau(\lambda)|_{\lambda=\lambda_a}\,.$$

for some coefficients c_{ka}. This will ensute that the Hamiltonian belongs to the family of commuting quantities. Since all the integrals from this family mutually commute they can be simultaneously diagonalized.

Represent the monodromy as the 2×2 matrix in the auxiliary space

$$T(\lambda) = \begin{pmatrix} A(\lambda) & B(\lambda) \\ C(\lambda) & D(\lambda) \end{pmatrix}\,,$$

where the entries are operators acting in the space $\otimes_{i=1}^L V_i$. From the definition of the monodromy and the L-operator it is clear that T is a polynomial in λ and

$$T(\lambda) = \lambda^L + i\lambda^{L-1}\sum_{n=1}^{L} S_n^\alpha \otimes \sigma^\alpha + \cdots$$

Thus, the transfer matrix is also polynomial of degree L:

$$\tau(\lambda) = \mathrm{tr}_a T_a(\lambda) = A(\lambda) + D(\lambda) = 2\lambda^L + \sum_{j=0}^{L-2} Q_j \lambda^j \,.$$

Note that the subleading term of order λ^{L-1} is absent because Pauli matrices are traceless. The coefficients Q_j mutually commute

$$[Q_i, Q_j] = 0\,.$$

Hamiltonian and Momentum. It remains to find the Hamiltonian among the commuting family generated by the transfer matrix. The L-operator has two special points on the spectral parameter plane.

- $\lambda = \frac{i}{2}$, where $L_{i,a}(i/2) = iP_{ia}$.
- $\lambda = \infty$. We see that

  $$\frac{1}{i}\mathrm{Res}\frac{T(\lambda)}{\lambda^L} = \sum_{n=1}^{L} S^\alpha \otimes \sigma^\alpha = \underbrace{S^\alpha}_{\mathfrak{su}(2)} \otimes \sigma^\alpha \,.$$

  This point will be related to the realization of the global $\mathfrak{su}(2)$ symmetry of the model.

Let us investigate the first point. We have

$$\begin{aligned} T_a(i/2) &= i^L \mathrm{P}_{L,a}\mathrm{P}_{L-1,a}\cdots\mathrm{P}_{1,a} = i^L \mathrm{P}_{L-1,L}\mathrm{P}_{L-2,L}\cdots\mathrm{P}_{1,L}\mathrm{P}_{L,a} = \\ &= i^L \mathrm{P}_{L-2,L-1}\mathrm{P}_{L-3,L-1}\cdots\mathrm{P}_{1,L-1}\mathrm{P}_{L-1,L}\mathrm{P}_{L,a} = \cdots = i^L\mathrm{P}_{12}\mathrm{P}_{23}\cdots\mathrm{P}_{L-1,L}\mathrm{P}_{L,a}\,. \end{aligned}$$

Thus, we have managed to isolate a single permutation carrying the index of the auxiliary subspace. Taking the trace and recalling that $\mathrm{tr}_a \mathrm{P}_{L,a} = \mathbb{I}_L$ we obtain the transfer matrix

$$\tau(i/2) = i^L \mathrm{P}_{12}\mathrm{P}_{23}\cdots\mathrm{P}_{L-1,L} = \mathcal{U} \qquad \leftarrow \qquad \text{shift operator}$$

Operator $\mathcal{U}$ is unitary $\mathcal{U}^\dagger\mathcal{U} = \mathcal{U}\mathcal{U}^\dagger = \mathbb{I}$ and it generates a shift along the chain:

$$\mathcal{U}^{-1} X_n \mathcal{U} = X_{n-1}\,.$$

By definition an operator of the infinitezimal shift is the momentum and on the lattice it is introduced as

$$\mathcal{U} = e^{ip}\,.$$

Now we differentiate the logarithm of the transfer matrix

$$\frac{dT_a(\lambda)}{d\lambda}|_{\lambda=i/2} = i^{L-1}\sum_n \mathrm{P}_{L,a}\cdots\underbrace{\mathrm{P}_{n,a}}_{\text{absent}}\cdots\mathrm{P}_{1,a} = i^{L-1}\sum_n \mathrm{P}_{12}\mathrm{P}_{23}\cdots\mathrm{P}_{n-1,n+1}\cdots\mathrm{P}_{L-1,L}.$$

This allows to establish that

$$\frac{d\tau(\lambda)}{d\lambda}\tau(\lambda)^{-1}|_{\lambda=i/2} =$$

$$= i^{-1}\Big(\sum_n \mathrm{P}_{12}\mathrm{P}_{23}\cdots \mathrm{P}_{n-1,n+1}\cdots \mathrm{P}_{L-1,L}\Big)\Big(\mathrm{P}_{L,L-1}\mathrm{P}_{L-1,L-2}\cdots \mathrm{P}_{2,1}\Big) = \frac{1}{i}\sum_{n,n+1}^{L} \mathrm{P}_{n,n+1}\,.$$

On the other hand we see that

$$H = -J\sum_{n=1}^{L} S_n^\alpha S_{n+1}^\alpha = -\frac{J}{4}\sum_{n=1}^{L}\sigma_n^\alpha\sigma_{n+1}^\alpha = -J\Big(\frac{1}{2}\sum_{n=1}^{L}\mathrm{P}_{n,n+1} - \frac{L}{4}\Big)\,.$$

Hence,

$$H = -J\Big(\frac{i}{2}\frac{d\tau(\lambda)}{d\lambda}\tau(\lambda)^{-1} - \frac{L}{4}\Big)|_{\lambda=i/2}\,,$$

i.e. the Hamiltonian belongs to the family of $L-1$ commuting integrals. To obtain L commuting integrals we can add the operator S^3 to this family.

The spectrum of the Heisenberg model. Here we compute the eigenvalues of H by using the algebraic Bethe ansatz. First we derive the commutation relations between the operators A, B, C, D. The form of the R-matrix is

$$R(\lambda-\mu) = \begin{pmatrix} \lambda-\mu+i & 0 & 0 & 0 \\ 0 & \lambda-\mu & i & 0 \\ 0 & i & \lambda-\mu & 0 \\ 0 & 0 & 0 & \lambda-\mu+i \end{pmatrix}.$$

We compute

$$T_a(\lambda) = \begin{pmatrix} A(\lambda) & 0 & B(\lambda) & 0 \\ 0 & A(\lambda) & 0 & B(\lambda) \\ C(\lambda) & 0 & D(\lambda) & 0 \\ 0 & C(\lambda) & 0 & D(\lambda) \end{pmatrix}, \qquad T_b(\lambda) = \begin{pmatrix} A(\mu) & B(\mu) & 0 & 0 \\ C(\mu) & D(\mu) & 0 & 0 \\ 0 & 0 & A(\mu) & B(\mu) \\ 0 & 0 & C(\mu) & D(\mu) \end{pmatrix}.$$

Plugging this into the fundamental commutation relation we get

$$\begin{pmatrix} (\alpha+i)A_\lambda A_\mu & (\alpha+i)A_\lambda B_\mu & (\alpha+i)B_\lambda A_\mu & (\alpha+i)B_\lambda B_\mu \\ \alpha A_\lambda C_\mu + iC_\lambda A_\mu & \alpha A_\lambda D_\mu + iC_\lambda B_\mu & \alpha B_\lambda C_\mu + iD_\lambda A_\mu & \alpha B_\lambda D_\mu + iD_\lambda B_\mu \\ iA_\lambda C_\mu + \alpha C_\lambda A_\mu & iA_\lambda D_\mu + \alpha C_\lambda B_\mu & iB_\lambda C_\mu + \alpha D_\lambda A_\mu & iB_\lambda D_\mu + \alpha D_\lambda B_\mu \\ (\alpha+i)C_\lambda C_\mu & (\alpha+i)C_\lambda D_\mu & (\alpha+i)D_\lambda C_\mu & (\alpha+i)D_\lambda D_\mu \end{pmatrix} =$$

$$= \begin{pmatrix} (\alpha+i)A_\mu A_\lambda & \alpha B_\mu A_\lambda + iA_\mu B_\lambda & iB_\mu A_\lambda + \alpha A_\mu B_\lambda & (\alpha+i)B_\mu B_\lambda \\ (\alpha+i)C_\mu A_\lambda & \alpha D_\mu A_\lambda + iC_\mu B_\lambda & iD_\mu A_\lambda + \alpha C_\mu B_\lambda & (\alpha+i)D_\mu B_\lambda \\ (\alpha+i)A_\mu C_\lambda & \alpha B_\mu C_\lambda + iA_\mu D_\lambda & iB_\mu C_\lambda + \alpha A_\mu D_\lambda & (\alpha+i)B_\mu D_\lambda \\ (\alpha+i)C_\mu C_\lambda & \alpha D_\mu C_\lambda + iC_\mu D_\lambda & iD_\mu C_\lambda + \alpha C_\mu D_\lambda & (\alpha+i)D_\mu D_\lambda \end{pmatrix}.$$

To write down the fundamental commutation relations we have used the shorthand notations $A_\lambda \equiv A(\lambda)$ and $\alpha = \lambda - \mu$. The relevant commutation relations are

$$\begin{aligned}
[B(\lambda), B(\mu)] &= 0\,, \\
A(\lambda)B(\mu) &= \frac{\lambda-\mu-i}{\lambda-\mu}B(\mu)A(\lambda) + \frac{i}{\lambda-\mu}B(\lambda)A(\mu)\,, \\
D(\lambda)B(\mu) &= \frac{\lambda-\mu+i}{\lambda-\mu}B(\mu)D(\lambda) - \frac{i}{\lambda-\mu}B(\lambda)D(\mu)\,.
\end{aligned} \tag{5.5}$$

The main idea of the algebraic Bethe ansatz is that there exists a pseudo-vacuum $|0\rangle$ such that $C(\lambda)|0\rangle = 0$ and the eigenvectors of $\tau(\lambda)$ with M spins down have the form

$$|\lambda_1, \lambda_2, \cdots, \lambda_M\rangle = B(\lambda_1)B(\lambda_2)\cdots B(\lambda_M)|0\rangle\,,$$

where $\{\lambda_i\}$ are "Bethe roots" which we will compare later on with the pseudo-momenta p_i of the magnons in the coordinate Bethe ansatz approach. One can see that the pseudo-vacuum can be identified with the state

$$|0\rangle = \otimes_{n=1}^{L} |\uparrow_n\rangle\,.$$

Indeed, since we have

$$L_n(\lambda)|\uparrow_n\rangle = \begin{pmatrix} (\lambda+\frac{i}{2})|\uparrow_n\rangle & i|\downarrow_n\rangle \\ 0 & (\lambda-\frac{i}{2})|\uparrow_n\rangle \end{pmatrix}$$

we find that

$$T(\lambda)|0\rangle = \begin{pmatrix} (\lambda+\frac{i}{2})^L|0\rangle & * \\ 0 & (\lambda-\frac{i}{2})^L|0\rangle \end{pmatrix}\,,$$

where $*$ stands for irrelevant terms. Thus, we indeed have

$$C(\lambda)|0\rangle = 0\,, \qquad A(\lambda)|0\rangle = \Big(\lambda+\frac{i}{2}\Big)^L|0\rangle\,, \qquad D(\lambda)|0\rangle = \Big(\lambda-\frac{i}{2}\Big)^L|0\rangle\,.$$

Comparing with the coordinate Bethe ansatz we see that $|0\rangle \equiv |F\rangle$. We also see that $|0\rangle$ is an eigenstate of the transfer matrix. The algebraic Bethe ansatz states that the other eigenstates are of the form

$$|\lambda_1, \lambda_2, \cdots, \lambda_M\rangle = B(\lambda_1)B(\lambda_2)\cdots B(\lambda_M)|0\rangle$$

provided the Bethe roots $\{\lambda_i\}$ satisfy certain restrictions. Let us now find these restrictions.

We compute

$$\begin{aligned}
A(\lambda)B(\lambda_1)B(\lambda_2)\cdots B(\lambda_M)|0\rangle &= \Big(\lambda+\frac{i}{2}\Big)^L\Big(\prod_{n=1}^{M}\frac{\lambda-\lambda_n-i}{\lambda-\lambda_n}\Big)B(\lambda_1)B(\lambda_2)\cdots B(\lambda_M)|0\rangle \\
&\quad + \sum_{n=1}^{M} W_n^A(\lambda,\{\lambda_i\})B(\lambda)\prod_{\substack{j=1\\ j\neq n}}^{M} B(\lambda_j)|0\rangle\,.
\end{aligned}$$

Here the coefficients $W_n^A(\lambda, \{\lambda_i\})$ depend on λ and the set $\{\lambda_i\}_{i=1}^M$. To determine this coefficient we note that since the operators $B(\lambda)$ commute with each other we can write

$$|\lambda_1, \lambda_2, \cdots, \lambda_M\rangle = B(\lambda_n) \prod_{\substack{j=1 \\ j \neq n}}^{M} B(\lambda_j)|0\rangle .$$

Thus,

$$A(\lambda)|\lambda_1, \lambda_2, \cdots, \lambda_M\rangle = \frac{\lambda - \lambda_n - i}{\lambda - \lambda_n} B(\lambda_n)A(\lambda) \prod_{\substack{j=1 \\ j \neq n}}^{M} B(\lambda_j)|0\rangle + \frac{i}{\lambda - \lambda_n} B(\lambda)A(\lambda_n) \prod_{\substack{j=1 \\ j \neq n}}^{M} B(\lambda_j)|0\rangle .$$

From this equation we see that only the second term on the r.h.s. will contribute to W_n^A since this term does not contain $B(\lambda_n)$. If we now mover $A(\lambda)$ past $B(\lambda_j)$ we see that the only way to avoid the appearance of $B(\lambda_n)$ is to use only the first term on the r.h.s. of eq.(5.5). So the resulting term should have the form

$$\frac{i}{\lambda - \lambda_n} \Big(\lambda_n + \frac{i}{2}\Big)^L \prod_{\substack{i=1 \\ i \neq n}}^{M} \frac{\lambda_n - \lambda_i - i}{\lambda_n - \lambda_i} B(\lambda) \prod_{\substack{j=1 \\ j \neq n}}^{M} B(\lambda_j)|0\rangle ,$$

i.e.

$$W_n^A(\lambda, \{\lambda_i\}) = \frac{i}{\lambda - \lambda_n} \Big(\lambda_n + \frac{i}{2}\Big)^L \prod_{\substack{j=1 \\ j \neq n}}^{M} \frac{\lambda_n - \lambda_j - i}{\lambda_n - \lambda_j} .$$

In the same way we obtain

$$D(\lambda)B(\lambda_1)B(\lambda_2)\cdots B(\lambda_M)|0\rangle = \Big(\lambda - \frac{i}{2}\Big)^L \Big(\prod_{n=1}^{M} \frac{\lambda - \lambda_n + i}{\lambda - \lambda_n}\Big) B(\lambda_1)B(\lambda_2)\cdots B(\lambda_M)|0\rangle$$
$$+ \sum_{n=1}^{M} W_n^D(\lambda, \{\lambda_i\}) B(\lambda) \prod_{\substack{j=1 \\ j \neq n}}^{M} B(\lambda_j)|0\rangle$$

and

$$W_n^D(\lambda, \{\lambda_i\}) = -\frac{i}{\lambda - \lambda_n} \Big(\lambda_n - \frac{i}{2}\Big)^L \prod_{\substack{j=1 \\ j \neq n}}^{M} \frac{\lambda_n - \lambda_j + i}{\lambda_n - \lambda_j} .$$

Thus, we will solve the eigenvalue problem

$$\tau(\lambda)|\lambda_1, \cdots, \lambda_M\rangle = \Lambda(\lambda, \{\lambda_n\})|\lambda_1, \cdots, \lambda_M\rangle$$

with

$$\Lambda(\lambda, \{\lambda_n\}) = \Big(\lambda + \frac{i}{2}\Big)^L \prod_{n=1}^{M} \frac{\lambda - \lambda_n - i}{\lambda - \lambda_n} + \Big(\lambda - \frac{i}{2}\Big)^L \prod_{n=1}^{M} \frac{\lambda - \lambda_n + i}{\lambda - \lambda_n}$$

provided $W_n^A + W_n^D = 0$ for all n, which means that

$$\Big(\lambda_n + \frac{i}{2}\Big)^L \prod_{\substack{j=1\\ j\neq n}}^{M} \frac{\lambda_n - \lambda_j - i}{\lambda_n - \lambda_j} = \Big(\lambda_n - \frac{i}{2}\Big)^L \prod_{\substack{j=1\\ j\neq n}}^{M} \frac{\lambda_n - \lambda_j + i}{\lambda_n - \lambda_j}\,.$$

We write the last equations in the form

$$\Big(\frac{\lambda_n + i/2}{\lambda_n - i/2}\Big)^L = \prod_{\substack{j=1\\ j\neq n}}^{M} \frac{\lambda_n - \lambda_j + i}{\lambda_n - \lambda_j - i}\,.$$

These are *the Bethe equations.* Introducing $\lambda_j = \cot p_j$ the Bethe equations take precisely the same form as derived in the coordinate Bethe ansatz approach:

$$e^{ip_i L} = \prod_{\substack{j=1\\ j\neq i}}^{M} S(p_i, p_j)\,.$$

Note that the parametrization $\lambda_j = \cot p_j$ has a singularity at $k_j = 0$. From the experience with the coordinate Bethe ansatz we know that all the eigenvectors for which $k_j \neq 0$ are the highest weight states of the global spin algebra $\mathfrak{su}(2)$. Thus, we expect that the eigenvectors obtained in the algebraic Bethe ansatz approach have the same property. Now we are going to investigate this issue in mode detail.

Realization of the symmetry algebra. Let us consider the fundamental commutation relations in the limitimg case $\mu \to \infty$. We get

$$\Big((\lambda - \mu) + \frac{i}{2}(\mathbb{I}_a \otimes \mathbb{I}_b + \sum_\alpha \sigma_a^\alpha \otimes \sigma_b^\alpha)\Big) T_a(\lambda) \Big(\mu^L + i\mu^{L-1} \sum_{n,\alpha} S_n^\alpha \otimes \sigma_b^\alpha + \cdots\Big) =$$
$$= \Big(\mu^L + i\mu^{L-1} \sum_{n,\alpha} S_n^\alpha \otimes \sigma_b^\alpha + \cdots\Big) T_a(\lambda) \Big((\lambda - \mu) + \frac{i}{2}(\mathbb{I}_a \otimes \mathbb{I}_b + \sum_\alpha \sigma_a^\alpha \otimes \sigma_b^\alpha)\,.$$

The leading term of the order μ^{L+1} cancel out. The subleading term of order μ^L gives

$$-iT_a(\lambda) \sum_{n,\alpha} S_n^\alpha \otimes \sigma_b^\alpha + \frac{i}{2} T_a(\lambda) + \frac{i}{2}\Big(\sum_\alpha \sigma_a^\alpha \otimes \sigma_b^\alpha\Big) T_a(\lambda) =$$
$$= \frac{i}{2} T_a(\lambda) + \frac{i}{2} T_a(\lambda) \Big(\sum_\alpha \sigma_a^\alpha \otimes \sigma_b^\alpha\Big) - i \sum_{n,\alpha} S_n^\alpha \otimes \sigma_b^\alpha T_a(\lambda)\,.$$

Simplifying we get

$$\sum_\alpha [T_a(\lambda), S^\alpha + \frac{1}{2}\sigma_a^\alpha] \otimes \sigma_b^\alpha = 0\,.$$

This results into the following equation which describes how the components of the monodromy transform under the global symmetry generators

$$[S^\alpha, T_a(\lambda)] = \frac{1}{2}[T_a(\lambda), \sigma_a^\alpha]\,.$$

Thus, we end up with three separate equations

$$[S^3, T_a(\lambda)] = \frac{1}{2}[T_a(\lambda), \sigma_a^3] = \frac{1}{2}\Big[\begin{pmatrix} A(\lambda) & B(\lambda) \\ C(\lambda) & D(\lambda)\end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1\end{pmatrix}\Big] = \begin{pmatrix} 0 & -B(\lambda) \\ C(\lambda) & 0\end{pmatrix},$$

$$[S^+, T_a(\lambda)] = \frac{1}{2}[T_a(\lambda), \sigma_a^+] = \Big[\begin{pmatrix} A(\lambda) & B(\lambda) \\ C(\lambda) & D(\lambda)\end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0\end{pmatrix}\Big] = \begin{pmatrix} -C(\lambda) & A(\lambda)-D(\lambda) \\ 0 & C(\lambda)\end{pmatrix},$$

and

$$[S^-, T_a(\lambda)] = \frac{1}{2}[T_a(\lambda), \sigma_a^-] = \Big[\begin{pmatrix} A(\lambda) & B(\lambda) \\ C(\lambda) & D(\lambda)\end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0\end{pmatrix}\Big] = \begin{pmatrix} B(\lambda) & 0 \\ D(\lambda)-A(\lambda) & -B(\lambda)\end{pmatrix}.$$

Essentially, we need the following commutation relations

$$[S^3, B] = -B\,, \qquad [S^+, B] = A - D\,.$$

The action of the symmetry generators on the pseudo-vacuum have been already derived

$$S^+|0\rangle = 0\,, \qquad S^3|0\rangle = \frac{L}{2}|0\rangle\,.$$

So the state $|0\rangle$ is the highest weight state of the symmetry algebra. Further, we find

$$S^3|\lambda_1, \cdots, \lambda_M\rangle = \Big(\frac{L}{2} - M\Big)|\lambda_1, \cdots, \lambda_M\rangle$$

and

$$\begin{aligned} S^+|\lambda_1, \cdots, \lambda_M\rangle &= \sum_j B(\lambda_1)\dots B(\lambda_{j-1})(A(\lambda_j) - D(\lambda_j))B(\lambda_{j+1})\dots B(\lambda_M)|0\rangle \\ &= \sum_j O_j B(\lambda_1)\dots B(\lambda_{j-1})\hat{B}(\lambda_j)B(\lambda_{j+1})\dots B(\lambda_M)|0\rangle\,. \end{aligned}$$

The coefficients O_j are unknown for the moment. To calculate O_j we will use the arguments similar to those for computing W_j^A and W_j^D. The only contributions to O_j will come from

$$B(\lambda_1)\dots B(\lambda_{k-1})(A(\lambda_k) - D(\lambda_k))B(\lambda_{k+1})\dots B(\lambda_M)|0\rangle \qquad \text{with} \qquad k \le j.$$

If $k = j$ this contribution will be

$$\prod_{k_j+1}^{M} \frac{\lambda_j - \lambda_k - i}{\lambda_j - \lambda_k}\Big(\lambda_j + \frac{i}{2}\Big)^L - \prod_{k_j+1}^{M} \frac{\lambda_j - \lambda_k + i}{\lambda_j - \lambda_k}\Big(\lambda_j - \frac{i}{2}\Big)^L$$

and if $k < j$ the contribution will be

$$W_j^A(\lambda_k, \{\lambda\}_{k+1}^M) + W_j^D(\lambda_k, \{\lambda\}_{k+1}^M)\,.$$

Thus, adding up we obtain

$$\begin{aligned}
O_j &= \prod_{k_j+1}^{M} \frac{\lambda_j - \lambda_k - i}{\lambda_j - \lambda_k}\Big(\lambda_j + \frac{i}{2}\Big)^L + \sum_{k=1}^{j-1} W_j^A(\lambda_k, \{\lambda\}_{k+1}^M) \\
&- \prod_{k_j+1}^{M} \frac{\lambda_j - \lambda_k + i}{\lambda_j - \lambda_k}\Big(\lambda_j - \frac{i}{2}\Big)^L + \sum_{k=1}^{j-1} W_j^D(\lambda_k, \{\lambda\}_{k+1}^M) = \\
&= \prod_{k=j+1}^{M} \frac{\lambda_j - \lambda_k - i}{\lambda_j - \lambda_k}\Big(\lambda_j + \frac{i}{2}\Big)^L \left(1 + \sum_{k=1}^{j-1} \frac{i}{\lambda_k - \lambda_j} \prod_{p=k+1}^{j-1} \frac{\lambda_j - \lambda_p - i}{\lambda_j - \lambda_p}\right) \\
&- \prod_{k=j+1}^{M} \frac{\lambda_j - \lambda_k + i}{\lambda_j - \lambda_k}\Big(\lambda_j - \frac{i}{2}\Big)^L \left(1 - \sum_{k=1}^{j-1} \frac{i}{\lambda_k - \lambda_j} \prod_{p=k+1}^{j-1} \frac{\lambda_j - \lambda_p + i}{\lambda_j - \lambda_p}\right).
\end{aligned}$$

Let us now note the useful identity

$$t_n \equiv 1 + \sum_{k=n}^{j-1} \frac{i}{\lambda_k - \lambda_j} \prod_{p=k+1}^{j-1} \frac{\lambda_j - \lambda_p - i}{\lambda_j - \lambda_p} = \prod_{k=n}^{j-1} \frac{\lambda_j - \lambda_k - i}{\lambda_j - \lambda_k}\,.$$

We will prove this by induction over n. For $n = j-1$ and $n = j-2$ we have

$$\begin{aligned}
t_{j-1} &= 1 + \frac{i}{\lambda_{j-1} - \lambda_j} = \frac{\lambda_j - \lambda_{j-1} - i}{\lambda_j - \lambda_{j-1}}\,, \\
t_{j-2} &= 1 + \frac{i}{\lambda_{j-1} - \lambda_j} + \frac{i}{\lambda_{j-2} - \lambda_j}\frac{\lambda_j - \lambda_{j-1} - i}{\lambda_j - \lambda_{j-1}} = \frac{\lambda_j - \lambda_{j-1} - i}{\lambda_j - \lambda_{j-1}}\frac{\lambda_j - \lambda_{j-2} - i}{\lambda_j - \lambda_{j-2}}\,.
\end{aligned}$$

Now we suppose that the formula holds for $n = l$, then we have

$$t_{l-1} = t_l + \frac{i}{\lambda_{l-1} - \lambda_j} \prod_{p=l}^{j-1} \frac{\lambda_j - \lambda_p - i}{\lambda_j - \lambda_p} = \prod_{p=l-1}^{j-1} \frac{\lambda_j - \lambda_p - i}{\lambda_j - \lambda_p}\,,$$

which proves our assumption. With this formula at hand we therefore find

$$1 + \sum_{i=1}^{j-1} \frac{i}{\lambda_i - \lambda_j} \prod_{p=i+1}^{j-1} \frac{\lambda_j - \lambda_p - i}{\lambda_j - \lambda_p} = \prod_{k=1}^{j-1} \frac{\lambda_j - \lambda_k - i}{\lambda_j - \lambda_k}\,.$$

In the same way one can show that

$$1 - \sum_{i=1}^{j-1} \frac{i}{\lambda_i - \lambda_j} \prod_{p=i+1}^{j-1} \frac{\lambda_j - \lambda_p + i}{\lambda_j - \lambda_p} = \prod_{k=1}^{j-1} \frac{\lambda_j - \lambda_k + i}{\lambda_j - \lambda_k}\,.$$

It follows now from the Bethe equations that

$$O_j = \Big(\lambda_j + \frac{i}{2}\Big)^L \prod_{\substack{k=1\\k\neq j}} \frac{\lambda_j-\lambda_k-i}{\lambda_j-\lambda_k} - \Big(\lambda_j - \frac{i}{2}\Big)^L \prod_{\substack{k=1\\k\neq j}} \frac{\lambda_j-\lambda_k+i}{\lambda_j-\lambda_k} = 0\,.$$

This proves that the eigenvectors obtained from the algebraic Bethe ansatz are the highest weight vectors of the spin algebra.

Finally, we can compute the eigenvalues of the corresponding Bethe eigenvectors. We obtain

$$E = -i\Big(\frac{i}{2}\frac{d\tau(\lambda)}{d\lambda}\tau(\lambda)^{-1}|_{\lambda=i/2} - \frac{L}{4}\Big) = E_0 + \frac{J}{2}\sum_{j=1}^{L}\frac{1}{\lambda_j^2+\frac{1}{4}}\,.$$

If we now use the parametrization $\lambda_j = \frac{1}{2}\cot\frac{p_j}{2}$ we get

$$E - E_0 = J\sum_{j=1}^{L}\frac{2}{1+\cot^2\frac{p_j}{2}} = 2\sum_{j=1}^{L}\sin^2\frac{p_j}{2}\,.$$

This expression agrees with the one obtained in the coordinate Bethe ansatz framework.

Let us summarize some important observations about the Bethe ansatz. First of all, the Heisenberg model has $\mathfrak{su}(2)$ symmetry which results into the fact that the eigenvectors calculated by using the Bethe ansatz procedure splits into irreducible representations of $\mathfrak{su}(2)$. For finite values of λ_j the eigenvectors of the algebraic Bethe ansatz are the always the highest weight states of $\mathfrak{su}(2)$. Descendents of the highest weight vectors correspond to roots at infinity, correspondingly $p_j = 0$. A second observation is that the algebraic Bethe ansatz enables us to prove integrability of the model and it gives an explicit construction of the Hilbert space of states in terms of simultaneous eigenvectors of commuting integrals of motion. Comparing to the classical inverse scattering method one can see that $\tau(\lambda)$ resembles the classical action variables, while $B(\lambda)$ corresponds to the angle variables.

5.3 Nested Bethe Ansatz (to be written)

Let g be an element from S_M, the permutation group of of the integers 1 to M. Obviously, there are $M!$ permutations. Any such permutation is a collection of integers

$$g = \Big(g1, g2, \cdots, gM\Big)\,.$$

In other words, g puts $g1$ on the first place, etc. Every of M particles is characterized by its position x_i. We choose the fundamental region

$$x_1 \leq x_2 \leq \cdots \leq x_M\,.$$

Now we have to specify which of M particles has a coordinate x_1, which – coordinate x_2, etc. This is specified by fixing a permutation Q. The Bethe ansatz for the wave function states that we look it in the form

$$\Psi(x|Q) = \sum_{\pi \in S_M} a(Q|\pi) e^{i \sum_{j=1}^{M} x_j p_{\pi(j)}} \,.$$

6. Introduction to Lie groups and Lie algebras

To introduce a concept of a Lie group we need two notions: the notion of a group and the notion of a smooth manifold.

Definition of a group. A set of elements G is called a group if it is endowed with two operations: for any pair g and h from G there is a third element from G which is called the product gh, for any element $g \in G$ there is the inverse element $g^{-1} \in G$. The following properties must be satisfied

- $(fg)h = f(gh)$
- there exists an identity element $\mathbb{I} \in G$ such that $\mathbb{I}g = g\mathbb{I} = g$
- $gg^{-1} = \mathbb{I}$

Definition of a smooth manifold. Now we introduce the notion of a differentiable manifold. Any set of points is called a differentiable manifold if it is supplied with the following structure

- M is a union: $M = \cup_q U_q$, where U_q is homeomorphic (i.e. a continuous one-to-one map) to the n-dimensional Euclidean space
- Any U_q is supplied with coordinates x_q^α called the *local coordinates*. The regions U_q are called coordinate *charts*.
- any intersection $U_q \cap U_p$, if it is not empty, is also a region of the Euclidean space where two coordinate systems x_q^α and x_p^α are defined. It is required that any of these two coordinate systems is expressible via the other by a differentiable map:

$$\begin{aligned} x_p^\alpha &= x_p^\alpha(x_q^1, \cdots x_q^n)\,, \qquad && \alpha = 1, \cdots n \\ x_q^\alpha &= x_q^\alpha(x_p^1, \cdots x_p^n)\,, \qquad && \alpha = 1, \cdots n \end{aligned} \tag{6.1}$$

  Then the Jacobian $\det\left(\frac{\partial x_p^\alpha}{\partial x_q^\beta}\right)$ is different from zero. The functions (6.1) are called *transition functions* from coordinates x_q^α to x_p^α and vice versa. If all the transition functions are infinitely differentiable (i.e. have all partial derivatives) the corresponding manifold is called *smooth*.

Definition of a Lie group: A smooth manifold G of dimension n is called a Lie group if G is supplied with the structure of a group (multiplication and inversion) which is compatible with the structure of a smooth manifold, i.e., the group operations are smooth. In other words, a Lie group is a group which is simultaneously a smooth manifold and the group operations are smooth.

The list of basic matrix Lie groups

- The group of $n \times n$ invertible matrices with complex or real matrix elements:

$$A = a_i^j\,, \qquad \det A \neq 0$$

It is called *the general linear group* $GL(n,\mathbb{C})$ or $GL(n,\mathbb{R})$. Consider for instance $GL(n,\mathbb{R})$. Product of two invertible matrices is an invertible matrix is invertible; an invertible matrix has its inverse. Thus, $GL(n,\mathbb{R})$ is a group. Condition $\det A \neq 0$ defines a domain in the space of all matrices $M(n,\mathbb{R})$ which is a linear space of dimension n^2. Thus, the general linear group is a domain in the linear space $\mathbb{R}^{n^2}$. Coordinates in $M(n,\mathbb{R})$ are the matrix elements a_i^j. If A and B are two matrices then their product $C = AB$ has the form

$$c_i^j = a_i^k b_k^j$$

It follows from this formula that the coordinates of the product of two matrices is expressible via their individual coordinates with the help of smooth functions (polynomials). In other words, the group operation which is the map

$$GL(n,\mathbb{R}) \times GL(n,\mathbb{R}) \to GL(n,\mathbb{R})$$

is smooth. Matrix elements of the inverse matrix are expressible via the matrix elements of the original matrix as no-where singular rational functions (since $\det A \neq 0$) which also defines a smooth mapping. Thus, the general Lie group is a Lie group.

- *Special linear group* $SL(n,\mathbb{R})$ or $SL(n,\mathbb{C})$ is a group of real or complex matrices satisfying the condition

$$\det A = 1\,.$$

- *Special orthogonal group* $SO(n,\mathbb{R})$ or $SO(n,\mathbb{C})$ is a group or real or complex matrices satisfying the conditions

$$AA^t = \mathbb{I}\,, \qquad \det A = 1\,.$$

- *Pseudo-orthogonal groups* $SO(p,q)$. Let g will be pseudo-Euclidean metric in the space $\mathbb{R}^n_{p,q}$ with $p+q=n$. The group $SO(p,q)$ is the group of real matrices which preserve the form g:

$$AgA^t = g\,, \qquad \det A = 1\,.$$

- *Unitary group* $U(n)$ – the group of unitary $n\times n$ matrices:

$$UU^\dagger = \mathbb{I}\,.$$

- *Special unitary group* $SU(n)$ – the group of unitary $n\times n$ matrices with the unit determinant

$$UU^\dagger = \mathbb{I}\,, \qquad \det U = 1\,.$$

- *Pseudo-unitary group* $U(p,q)$:

$$AgA^\dagger = g\,,$$

  where g is the pseudo-Euclidean metric. Special pseudo-unitary group requires in addition the unit determinant $\det A = 1$.

- *Symplectic group* $Sp(2n,\mathbb{R})$ or $Sp(2n,\mathbb{C})$ is a group or real or complex matrices satisfying the condition

$$AJA^t = J$$

  where J is $2n\times 2n$ matrix

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

  and $\mathbb{I}$ is $n\times n$ unit matrix.

  Question to the class: What are the eigenvalues of J? Answer:

$$J = \mathrm{diag}(\mathrm{i},\cdots \mathrm{i}; -\mathrm{i},\cdots,-\mathrm{i}).$$

  Thus, the group $Sp(2n)$ is really different from $SO(2n)$!

The powerful tool in the theory of Lie groups are the Lie algebras. Let us see how they arise by using as an example $SO(3)$. Let A be "close" to the identity matrix

$$A = \mathbb{I} + \epsilon a$$

is an orthogonal matrix $A^t = A^{-1}$. Therefore,

$$\mathbb{I} + \epsilon a^t = (\mathbb{I}+\epsilon a)^{-1} = \mathbb{I} - \epsilon a + \epsilon^2 a^2 + \cdots$$

From here $a^t = -a$. The space of matrices a such that $a^t = -a$ is denoted as $so(3)$ and called the Lie algebra of the Lie group $SO(3)$. The properties of this Lie algebra: $so(3)$ is a linear space, in $so(3)$ the commutator is defined: if $a, b \in so(3)$ then $[a, b]$ also belongs to $so(3)$. A linear space of matrices is called a Lie algebra if the commutator does not lead out of this space. Commutator of matrices naturally arises from the commutator in the group:

$$\begin{aligned} ABA^{-1}B^{-1} &= (\mathbb{I} + \epsilon a)(\mathbb{I} + \epsilon b)(\mathbb{I} + \epsilon a)^{-1}(\mathbb{I} + \epsilon b)^{-1} \\ &= (\mathbb{I} + \epsilon a)(\mathbb{I} + \epsilon b)(\mathbb{I} - \epsilon a + \epsilon^2 a^2 + \cdots)(\mathbb{I} - \epsilon b + \epsilon^2 b^2 + \cdots) = \\ &= \mathbb{I} + \epsilon(a + b - a - b) + \epsilon^2(ab - a^2 - ab - ba - b^2 + ab + a^2 + b^2) + \cdots = \\ &= \mathbb{I} + \epsilon^2 [a, b] + \cdots \end{aligned}$$

The algebra and the Lie group in our example are related as

$$\exp a = \sum_{n=0}^{\infty} \frac{a^n}{n!} = A \in SO(3)$$

Exponential of matrix. The exponent $\exp a$ of the matrix a is the sum of the following series

$$\exp a = \sum_{m=0}^{\infty} \frac{a^m}{m!} \, .$$

This series shares the properties of the usual exponential function, in particular it is convergent for any matrix A. The following obvious properties are

- If matrices X and Y commute then

$$\exp(X + Y) = \exp(X) \exp(Y)$$

- The matrix $A = \exp X$ is invertible and $A^{-1} = \exp(-X)$.
- $\exp(X^t) = (\exp X)^t$.

Definition of a Lie algebra: A linear vector space $\mathcal{J}$ (over a field $\mathbb{R}$ or $\mathbb{C}$) supplied with the multiplication operation (this operation is called *the commutator*) $[\xi, \eta]$ for $\xi, \eta \in \mathcal{J}$ is called a Lie algebra if the following properties are satisfied

1. The commutator $[\xi, \eta]$ is a bilinear operation, i.e.

$$[\alpha_1 \xi_1 + \alpha_2 \xi_2, \beta_1 \eta_1 + \beta_2 \eta_2] = \alpha_1 \beta_1 [\xi_1, \eta_1] + \alpha_2 \beta_1 [\xi_2, \eta_1] + \alpha_1 \beta_2 [\xi_1, \eta_2] + \alpha_2 \beta_2 [\xi_2, \eta_2]$$

2. The commutator is skew-symmetric: $[\xi, \eta] = -[\eta, \xi]$

3. The Jacobi identity

$$[[\xi,\eta],\zeta]+[[\eta,\zeta],\xi]+[[\zeta,\xi],\eta]=0$$

Let $\mathcal{J}$ be a Lie algebra of dimension n. Choose a basis $e_1,\cdots,e_n\in\mathcal{J}$. We have

$$[e_i,e_j]=C_{ij}^k e_k$$

The numbers C_{ij}^k are called *structure constants* of the Lie algebra. Upon changing the basis these structure constants change as the tensor quantity. Let $e'_i=A_i^j e_i$ and $[e'_i,e'_j]=C'^k_{ij}e'_k$ then

$$C'^k_{ij}A_k^m e_m=A_i^r A_j^s[e_r,e_s]=A_i^r A_j^s C_{rs}^m e_m$$

Thus, the structure constants in the new basis are related to the constants in the original basis as

$$C'^k_{ij}=A_i^r A_j^s C_{rs}^m (A^{-1})_m^k\,. \tag{6.2}$$

Skew-symmetry and the Jacobi identity for the commutator imply that the tensor C_{ij}^k defines the Lie algebra if and only if

$$C_{ij}^k=-C_{ij}^k\,,\qquad C_{p[i}^m C_{jk]}^p=0\,.$$

Classify all Lie algebras means in fact to find all solutions of these equations modulo the equivalence relation (6.2).

Example. The Lie algebra $so(3,\mathbb{R})$ of the Lie group $SO(3,\mathbb{R})$. It consists of 3×3 skew-symmetric matrices. We can introduce a basis in the space of these matrices

$$X_1=\begin{pmatrix}0&0&0\\0&0&-1\\0&1&0\end{pmatrix},\qquad X_2=\begin{pmatrix}0&0&1\\0&0&0\\-1&0&0\end{pmatrix},\qquad X_3=\begin{pmatrix}0&-1&0\\1&0&0\\0&0&0\end{pmatrix}.$$

In this basis the Lie algebra relations take the form

$$[X_1,X_2]=X_3\,,\qquad [X_2,X_3]=X_1\,,\qquad [X_3,X_1]=X_2\,.$$

These three relation can be encoded into one

$$[X_i,X_j]=\epsilon_{ijk}X_k\,.$$

Example. The Lie algebra $su(2)$ of the Lie group $SU(2)$. It consists of 2×2 skew-symmetric matrices. The basis can be constructed with the help of the so-called Pauli matrices σ_i

$$\sigma_1=\begin{pmatrix}0&1\\1&0\end{pmatrix},\qquad \sigma_2=\begin{pmatrix}0&-i\\i&0\end{pmatrix},\qquad \sigma_3=\begin{pmatrix}1&0\\0&-1\end{pmatrix}.$$

These matrices satisfy the relations

$$[\sigma_i, \sigma_j] = 2i\epsilon_{ijk}\sigma_k\,, \qquad \{\sigma_i, \sigma_j\} = 2\delta_{ij}\,.$$

If we introduce $X_i = -\frac{i}{2}\sigma_i$ which are three linearly independent anti-hermitian matrices then the $su(2)$ Lie algebra relations read

$$[X_i, X_j] = \epsilon_{ijk}X_k$$

Note that the structure constants are real! Comparing with the previous example we see that the Lie algebra $su(2)$ is isomorphic to that of $so(3, \mathbb{R})$:

$$su(2) \approx so(3, \mathbb{R})\,.$$

With every matrix group we considered above one can associate the corresponding matrix Lie algebra. The vector space of this Lie algebra is the tangent space at the identity element of the group. For this case the operation "commutator" is the usual matrix commutator. The tangent space to a Lie group at the identity element naturally appears in this discussion. To understand why let us return to the case of the Lie group $GL(n, \mathbb{R})$. Consider a one-parameter curve $A(t) \in GL(n, \mathbb{R})$, i.e, a family of matrices $A(t)$ from $GL(n, \mathbb{R})$ which depend on the parameter t. Let this curve to pass though the identity at $t = 0$, i.e., $A(0) = \mathbb{I}$. Then the tangent vector (the velocity vector!) at $t = 0$ is the matrix $\dot{A}(t)|_{t=0}$. Other way around, let X be an arbitrary matrix. Then the curve $A(t) = \mathbb{I} + tX$ for t sufficiently closed to zero lies in $GL(n, \mathbb{R})$. It is clear that

$$A(0) = \mathbb{I}\,, \qquad \dot{A}(0) = X\,.$$

In this way we demonstrated that the space of vectors which are tangent to the group $GL(n, \mathbb{R})$ at the identity coincide with the space of all $n \times n$ matrices. This example of $GL(n, \mathbb{R})$ demonstrates a universal connection between Lie group G and its Lie algebra: *The tangent space to G at the identity element is the Lie algebra w.r.t. to the commutator. This Lie algebra is called the Lie algebra of the group G.*

Exercise to do in the class: making infinitesimal expansion of a group element close to the identity compute the Lie algebras for the classical matrix groups discussed above. The answer is the following list:

The list of basic matrix Lie algebras

- The general Lie group $GL(n, \mathbb{R})$ or $GL(n, \mathbb{C})$ has the matrix Lie algebra which is $M(n, \mathbb{R})$ or $M(n, \mathbb{C})$, where $M(n)$ is the space of all real or complex matrices.

- Special linear group $SL(n,\mathbb{R})$ or $SL(n,\mathbb{C})$ has the Lie algebra $sl(n,\mathbb{R})$ or $sl(n,\mathbb{C})$ which coincides with the space of all real or complex matrices with zero trace.

- Special orthogonal group $SO(n,\mathbb{R})$ or $SO(n,\mathbb{C})$ has the Lie algebra $so(n,\mathbb{R})$ or $so(n,\mathbb{C})$ which are real or complex matrices satisfying the condition

$$X^t = -X\,.$$

- Pseudo-orthogonal group $SO(p,q)$ has the Lie algebra which is the algebra of matrices X satisfying the condition

$$Xg + gX^t = 0\,.$$

  We see that if we introduce the matrix $u = Xg$ then the relation defining the Lie algebra reads

$$u + u^t = 0\,.$$

  Thus, the matrix u is skew-symmetric $u^t + u = 0$. This map establishes the isomorphism between $so(p,q)$ and the space of all skew-symmetric matrices.

- Unitary group $U(n)$ has the Lie algebra which is the space of all anti-hermitian matrices

$$X^\dagger = -X\,.$$

- Special unitary group $SU(n)$ has the Lie algebra which is the space of all anti-hermitian matrices with zero trace

$$X^\dagger = -X\,, \qquad \mathrm{tr}X = 0\,.$$

- Pseudo-unitary group $U(p,q)$ has the Lie algebra which is the space of all matrices obeying the relation

$$Xg + gX^\dagger = 0\,.$$

  The space $u(p,q)$ is isomorphic to the space of anti-hermitian matrices. The isomorphism is established by the formula $u = Xg$. Finally the Lie algebra of the special pseudo-unitary group is defined by further requirement of vanishing trace for X.

- The symplectic group $Sp(2n,\mathbb{R})$ or $Sp(2n,\mathbb{C})$ has the Lie algebra which comprises all is the is a group or real or complex matrices satisfying the condition

$$XJ + JX^t = 0$$

where J is $2n \times 2n$ matrix

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

and $\mathbb{I}$ is $n \times n$ unit matrix.

Linear representations of Lie groups Consider an action of a Lie group a n-dimensional vector space $\mathbb{R}^n$. This action is called *a linear representation* of Lie group G on $\mathbb{R}^n$ if for any $g \in G$ the map

$$\rho : \quad g \to \rho(g)$$

is a linear operator on $\mathbb{R}^n$. In other words, by a linear representation of G on $\mathbb{R}^n$ we call the homomorphism ρ which maps G into $GL(n, \mathbb{R})$, the group of linear transformations of $\mathbb{R}^n$. The homomorphism means that under this map the group structure is preserved, i.e.

$$\rho(g_1 g_2) = \rho(g_1)\rho(g_2)\,.$$

Any Lie group G has a distinguished element – $g_0 = \mathbb{I}$ and the tangent space T at this point. Transformation

$$G \to G : \qquad g \to hgh^{-1}$$

is called *internal automorphism* corresponding to an element $h \in G$. This transformation leaves unity invariant: $h\mathbb{I}h^{-1} = \mathbb{I}$ and it transforms the tangent space T into itself:

$$\mathrm{Ad}(h) : \quad T \to T\,.$$

This map has the following properties:

$$\mathrm{Ad}(h^{-1}) = (\mathrm{Ad}h)^{-1}\,, \qquad \mathrm{Ad}(h_1 h_2) = \mathrm{Ad}h_1 \mathrm{Ad}h_2\,.$$

In other words, the map $h \to \mathrm{Ad}h$ is a *linear representation* of G:

$$\mathrm{Ad} : \quad G \to GL(n, \mathbb{R})\,,$$

where n is the dimension of the group.

Generally, one-parameter subgroups of a Lie group G are defined as parameterized curves $F(t) \subset G$ such that $F(0) = \mathbb{I}$ and $F(t_1 + t_2) = F(t_1)F(t_2)$ and $F(-t) = F(t)^{-1}$. As we have already discussed for matrix groups they have the form

$$F(t) = \exp(At)$$

where A is an element of the corresponding Lie algebra. In an abstract Lie group G for a curve $F(t)$ one defines the t-dependent vector

$$F^{-1}\dot{F} \in T\,.$$

If this curve $F(t)$ is one-parameter subgroup then this vector does not depend on t! Indeed,

$$\dot{F} = \frac{dF(t+\epsilon)}{d\epsilon}|_{\epsilon=0} = F(t)\Big(\frac{dF(\epsilon)}{d\epsilon}\Big)_{\epsilon=0}\,,$$

i.e. $\dot{F} = F(t)\dot{F}(0)$ and $F^{-1}(t)\dot{F}(t) = \dot{F}(0) = \text{const}$. Oppositely, for any non-zero $a \in T$ there exists a unique one-parameter subgroup with

$$F^{-1}\dot{F} = a\,.$$

This follows from the theorem about the existence and uniqueness of solutions of usual differential equations.

It is important to realize that even for the case of matrix Lie groups there are matrices which are not images of any one-parameter subgroup. The exercise to do in the class: Consider the following matrix:

$$g = \begin{pmatrix} -2 & 0 \\ 0 & -3 \end{pmatrix} \in GL^{+}(2,\mathbb{R})\,,$$

where $GL^{+}(2,\mathbb{R})$ is a subgroup of $GL(2,\mathbb{R})$ with positive determinant. Show that there does not exist any real matrix ξ such that

$$e^{\xi} = g\,.$$

The answer: it is impossible because since the matrix ξ is real the eigenvalues $\lambda_{1,2}$ of ξ must be either real of complex conjugate. The eigenvalues of e^{ξ} are e^{λ_1} and e^{λ_2}. If λ_i are real then $e^{\lambda_i} > 0$. If λ_i are complex conjugate then e^{λ_i} are also complex conjugate.

It is also important to realize that different vectors ξ under the exponential map can be mapped on the one and the same group element. As an example, consider the matrices of the form

$$\xi = \alpha \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \beta \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\,,$$

where $\alpha, \beta \in \mathbb{R}$. Exponent e^{ξ} can be computed by noting that

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}^2 = -\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\,.$$

Then we have

$$e^{\xi} = e^{\alpha}\Big[\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\cos\beta + \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\sin\beta\Big]\,.$$

It is clear that

$$\alpha \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \beta \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad \alpha \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + (\beta + 2\pi k) \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

has the the same image under the exponential map. In the sufficiently small neighbourhood of 0 in $M(n, \mathbb{R})$ the map $\exp A$ is a diffeomorphism. The inverse map is constructed by means of series

$$\ln x = (x - \mathbb{I}) - \frac{1}{2}(x - \mathbb{I})^2 + \frac{1}{3}(x - \mathbb{I})^3 - \cdots$$

for x sufficiently close to the identity.

Linear representation of a Lie algebra. Adjoint representation. Let $\mathcal{J}$ be a Lie algebra. We say that a map

$$\rho : \quad \mathcal{J} \to M(n, \mathbb{R})$$

defines a representation of the Lie algebra $\mathcal{J}$ is the following equality is satisfied

$$\rho[\zeta, \eta] = [\rho(\eta), \rho(\zeta)]$$

for any two vectors $\zeta, \eta \in \mathcal{J}$.

Let $F(t)$ be a one-parameter subgroup in G. Then $g \to FgF^{-1}$ generates a one-parameter group of transformations in the Lie algebra

$$\mathrm{Ad}F(t) : \quad T \to T\,.$$

The vector $\frac{d}{dt}\mathrm{Ad}F(t)|_{t=0}$ lies in the Lie algebra. Let $a \in T$ and let $F(t) = \exp(bt)$ then

$$\frac{d}{dt}\mathrm{Ad}F(t)|_{t=0}\; a = \frac{d}{dt}\Big(\exp(bt) a \exp(-bt)\Big)|_{t=0} = [b, a]$$

Thus to any element $b \in \mathcal{J}$ we associate an operator ad_b which acts on the Lie algebra:

$$\mathrm{ad}_b : \quad \mathcal{J} \to \mathcal{J}, \qquad \mathrm{ad}_b a = [b, a]\,.$$

This action defines a representation of the Lie algebra on itself. This representation is called *adjoint*. To see that this is indeed representation we have to show that it preserves the commutation relations, i.e. that from $[x, y] = z$ it follows that

$$[\mathrm{ad}x, \mathrm{ad}y] = \mathrm{ad}z\,.$$

We compute

$$[\mathrm{ad}x, \mathrm{ad}y]w = \mathrm{ad}x\ \mathrm{ad}y w - \mathrm{ad}y\ \mathrm{ad}x w = [x, [y, w]] - [y, [x, w]] = [x, [y, w]] + [y, [w, x]] = \\ -\ [w, [x, y]] = [[x, y], w] = [z, w] = \mathrm{ad}z w\,.$$

Here the Jacobi identity has been used.

Semi-simple and simple Lie algebras. General classification of Lie algebras is a very complicated problem. To make a progress simplifying assumptions about the structure of the algebra are needed. The class of the so-called simple and semi-simple Lie algebras admits a complete classification.

A Lie subalgebra $\mathcal{H}$ of a Lie algebra $\mathcal{J}$ is a linear subspace $\mathcal{H} \subset \mathcal{J}$ which is closed w.r.t. to the commutation operation. An *ideal* $\mathcal{H} \subset \mathcal{J}$ is a subspace in $\mathcal{J}$ such that for any $x \in \mathcal{J}$ the following relation holds

$$[x, \mathcal{H}] \subset \mathcal{H} .$$

A Lie algebra $\mathcal{J}$ which does not have any ideals except the trivial one and the one coincident with $\mathcal{J}$ is called *simple*. A Lie algebra which have no commutative (i.e. abelian) ideals is called semi-simple. One can show that any semi-simple Lie algebra is a sum of simple Lie algebras. Consider for instance the Lie algebra $u(n)$ which is the algebra of anti-hermitian matrices

$$u + u^{\dagger} = 0 .$$

The Lie algebra $su(n)$ is further distinguished by imposing the condition of vanishing trace: $\mathrm{tr} u = 0$. The difference between $u(n)$ and $su(n)$ constitute all the matrices which are proportional to the identity matrix $i\mathbb{I}$. Since

$$[\lambda i\mathbb{I}, u] = 0$$

the matrices proportional to $i\mathbb{I}$ form an ideal in $u(n)$ which is abelian. Thus, $u(n)$ has the abelian ideal and, therefore, $u(n)$ is not semi-simple. In opposite, $su(n)$ has no non-trivial ideals and therefore it is the simple Lie algebra.

A powerful tool in the Lie theory is the so-called Cartan-Killing from on a Lie algebra. Consider the adjoint representation of $\mathcal{J}$. The Cartan-Killing form on $\mathcal{J}$ is defined as

$$(a, b) = -\mathrm{tr}(\mathrm{ad}_a \mathrm{ad}_b)$$

for any two $a, b \in \mathcal{J}$. The following central theorem in the Lie algebra theory can be proven: *A Lie algebra is semi-simple if and only if its Cartan-Killing form is non-degenerate.*

For a simple Lie algebra $\mathcal{J}$ of a group G the internal automorphisms $\mathrm{Ad}g$ constitute the linear irreducible representation (i.e. a representation which does not have invariant subspaces) of G in $\mathcal{J}$. Indeed, if $\mathrm{Ad}(g)$ has an invariant subspace $\mathcal{H} \subset \mathcal{J}$, i.e. $g\mathcal{H}g^{-1} \subset \mathcal{H}$ for any g then sending g to the identity we will get

$$[\mathcal{J}, \mathcal{H}] \subset \mathcal{H}$$

i.e. $\mathcal{H}$ is an ideal which contradicts to the assumption that $\mathcal{J}$ is the semi-simple Lie algebra.

Cartan subalgebra. To demonstrate the construction of the adjoint representation and introduce the notion of the Cartan subalgebra of the Lie algebra we use the concrete example of $su(3)$. The Lie algebra $su(3)$ comprises the matrices of the form iM, where M is traceless 3×3 hermitian matrix. The basis consists of eight matrices which we chose to be the Gell-Mann matrices:

$$\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$
$$\lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \qquad \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \qquad \lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$
$$\lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & 0 & 0 \end{pmatrix}, \qquad \lambda_8 = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}.$$

There are two diagonal matrices among these: λ_3 and λ_8 which we replace by $T_z = \frac{1}{2}\lambda_3$ and $Y = \frac{1}{\sqrt{3}}\lambda_8$. We introduce the following linear combinations of the generators

$$t_\pm = \frac{1}{2}(\lambda_1 \pm i\lambda_2)\,, \qquad v_\pm = \frac{1}{2}(\lambda_4 \pm i\lambda_5)\,, \qquad u_\pm = \frac{1}{2}(\lambda_6 \pm i\lambda_y)\,.$$

One can easily compute, e.g.,

$$[t_+, t_+] = 0\,, \quad [t_+, t_-] = 2t_z, \quad [t_+, t_z] = -t_+\,, \quad [t_+, u_+] = v_+\,, \quad [t_+, u_-] = 0\,,$$
$$[t_+, v_+] = 0\,, \quad [t_+, v_-] = -u_-\,, \quad [t_+, y] = 0\,.$$

Since the Lie algebra of $su(3)$ is eight-dimensional the adjoint representation is eight-dimensional too. Picking up $(t_+, t_-, t_z, u_+, u_-, v_+, v_-, y)$ as the basis we can realize the adjoint action by 8×8 matrices. For instance,

$$\mathrm{ad}_{t_+}\begin{pmatrix} t_+ \\ t_- \\ t_z \\ u_+ \\ u_- \\ v_+ \\ v_- \\ y \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}}_{\text{matrix realization of } t_+}\begin{pmatrix} t_+ \\ t_- \\ t_z \\ u_+ \\ u_- \\ v_+ \\ v_- \\ y \end{pmatrix}$$

Note that both ad_{t_z} and ad_y are diagonal. Thus, if $x = at_z + by$ then ad_x is also diagonal. Explicitly we find

$$\mathrm{ad}_x = \begin{pmatrix} a & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -a & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{1}{2}a+b & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2}a-b & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{2}a+b & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{2}a-b & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

In other words, the basis elements $(t_+, t_-, t_z, u_+, u_-, v_+, v_-, y)$ are all eigenvectors of ad_x with eigenvalues $a, -a, 0, -\frac{1}{2}a+b, \frac{1}{2}a-b, -\frac{1}{2}a-b$ and 0 respectively. The procedure we followed in crucial for analysis of other (larger) Lie algebras. We found a two-dimensional subalgebra generated by t_z and y which is abelian. Further, we have chosen a basis for the rest of the Lie algebra such that each element of the basis is an eigenvector of ad_x if x is from this abelian subalgebra. This abelian subalgebra is called *the Cartan subalgebra.*

In general the Cartan subalgebra H is determined in the following way. An element $h \in H$ is called *regular* if ad_h has as simple as possible number of zero eigenvalues (i.e. multiplicity of zero eigenvalue is minimal). For instance, for $su(3)$ the element ad_{t_z} has two zero eigenvalues, while ad_y has for zero eigenvalues. Thus, the element ad_{t_z} is regular, while ad_y is not. *A Cartan subalgebra is a maximal commutative subalgebra which contains a regular element.* In our example the subalgebra generated by t_z and y is commutative and its maximal since there is no other element we can add to it which would not destroy the commutativity.

Roots. It is very important fact proved in the theory of Lie algebras that any simple Lie algebra has a Cartan subalgebra and it admits a basis where each basis vector is an eigenstate of all Cartan generators; the corresponding eigenvalues depend of course on a Cartan generator. In our example of $su(3)$ for an element $x = at_z + by$

we have

$$\begin{aligned}
\mathrm{ad}_x t_+ &= a t_+ \\
\mathrm{ad}_x t_- &= a t_- \\
\mathrm{ad}_x t_z &= 0 t_z \\
\mathrm{ad}_x u_+ &= (-\frac{1}{2}a + b)u_+ \\
\mathrm{ad}_x u_- &= (\frac{1}{2}a - b)u_- \\
\mathrm{ad}_x v_+ &= (\frac{1}{2}a + b)v_+ \\
\mathrm{ad}_x v_- &= (-\frac{1}{2}a - b)v_- \\
\mathrm{ad}_x y &= 0 y\,.
\end{aligned}$$

We see that all eigenvalues are *linear* functions of the Cartan element x, in other words, if we denote by e_α the six elements $t_\pm, v_\pm, u_\pm$ and by h_i the two Cartan elements t_z, y we can write all the relations above as

$$[h_i, h_j] = 0$$

$$[h_i, e_\alpha] = \alpha(h_i)e_\alpha\,,$$

where $\alpha(h_i)$ is a linear function of h_i. The generators e_α, which are eigenstates of the Cartan subalgebra, are called *root vectors*, while the corresponding linear functions $\alpha(h)$ are called *roots*. To every root vector e_α we associate the root α which is a linear function on the Cartan sualgebra H. Linear functions on H, by definition, form the dual space H^* to the Cartan subalgebra H.

The Cartan-Weyl basis. Now we can also investigate what is the commutator of the root vectors. By using the Jacobi identity we find

$$[h, [e_\alpha, e_\beta]] = -[e_\alpha, [e_\beta, h]] - [e_\beta, [h, e_\alpha]] = (\alpha(h) + \beta(h))[e_\alpha, e_\beta]\,.$$

This clearly means that there are three distinct possibilities

- $[e_\alpha, e_\beta]$ is zero
- $[e_\alpha, e_\beta]$ is a root vector with the root $\alpha + \beta$
- $\alpha + \beta = 0$ in which case $[e_\alpha, e_\beta]$ commutes with every $h \in H$ and, therefore, is an element of the Cartan subalgebra.

  Thus,

$$[e_\alpha, e_\beta] = N_{\alpha\beta} e_{\alpha+\beta}$$

if $\alpha + \beta$ is a root,

$$[e_\alpha, e_{-\alpha}] \sim h_\alpha$$

and $[e_\alpha, e_\beta] = 0$ if $\alpha + \beta$ is not a root. The numbers $N_{\alpha\beta}$ depend on the normalization of the root vectors. The basis (h_i, e_α) of a Lie algebra with the properties described above is called *the Cartan-Weyl basis*.

7. Homework exercises

7.1 Seminar 1

Exercise 1. Consider a point particle moving in the potential U of the form depicted in figure 1.

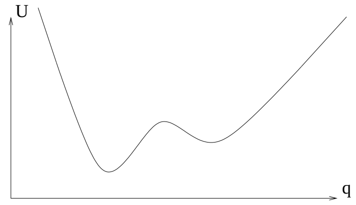


Fig. 1. Potential energy of a particle

Draw the phase curve of this particle. Hint: consult the case of the harmonic oscillator.

Exercise 2. Consider a point particle moving in the potential U of the forms depicted in figure 2.

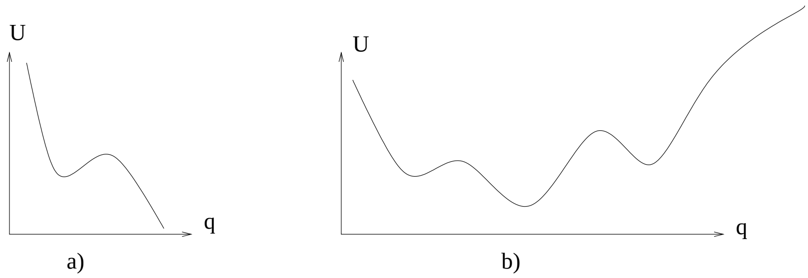


Fig. 2. Potential energies of a particle

Draw the corresponding phase curves.

Exercise 3. Consider a point particle of unit mass m which moves in one dimension (the coordinate q and the momentum p) in the potential $U(q)$, where

- case 1:
$$U(q) = \frac{g^2}{q^2}\,, \qquad E > 0$$
and g^2 is a (coupling) constant.
- case 2:
$$U(q) = \frac{g^2}{\sinh^2 q}\,, \qquad E > 0$$

- case 3:

$$U(q) = -\frac{g^2}{\cosh^2 q}\,, \qquad -g^2 < E < 0$$

- case 4:

$$U(q) = -\frac{g^2}{\cosh^2 q}\,, \qquad E > 0$$

Solve equations of motion for each of these potentials by quadratures. In which case the motion if finite?

Exercise 3. Consider a linear space M with coordinates x^k, $k = 1, \ldots, n$. Show that the expression

$$\{F(x), G(x)\} = C_l^{jk} x^l \partial_j F \partial_k G$$

defines the Poisson bracket provided the constants C_l^{jk} coincide with the structure constants of a Lie algebra.

7.2 Seminar 2

Exercise 1. Work out following the book three integrable tops: Euler, Lagrange and Kowalewski tops. Work out equations of motion and check the conservation laws.

7.3 Seminar 3

Exercise 1. Consider a motion of a one-dimensional system. Let $S(E)$ be the area enclosed by the closed phase curve corresponding to the energy level E. Show that the period of motion along this curve is equal to

$$T = \frac{dS(E)}{dE}\,.$$

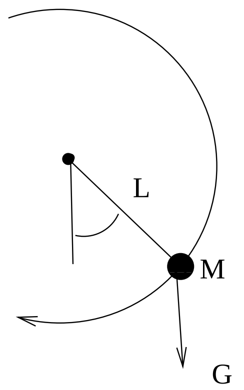
Exercise 2. At the entry of the satellite into a circular orbit at a distance 300km from the Earth the direction of its velocity deviates from the intended direction by 1° towards the earth. How is the perigee of the orbit changed?

Exercise 3. Find the principle axes and moments of inertia of the uniform planar plate $|x| \leq q$, $|y| \leq b$, $z = 0$ with respect to 0.

Exercise 4. Find the inertia tensor of the uniform ellipsoid with the semi-axes a, b, c.

Exercise 5. Solve the Euler equations for the symmetric top: $I_1 = I_2$.

Exercise 6. Consider the mathematical pendulum (of mass M) in the gravitational field of the Earth. Integrate equations of motion in terms of Jacobi elliptic functions. If the second (imaginary) period has any physical meaning? What is the elliptic modulus k^2? Consider the limits $k = 0^+$ and $k = 1^-$.



A pendulum in the gravitational field of the Earth. Here L is its length and G is the gravitational constant.

7.4 Seminar 4

Exercise 1 (K. Bohlin). Consider the Kepler problem. Let x, y be the Cartesian coordinates on the plane of motion. Introduce a complex variable $z = x + iy$ and show that the non-linear change of variables $z \to u^2$, $t \to \tau$ given by

$$z = u^2\,, \qquad \frac{dt}{d\tau} = 4|u^2| = 4|z|$$

maps the Kepler orbits with the constant energy $E < 0$ into the ones of the harmonic oscillator with the complex amplitude u (a two-dimensional oscillator). Find the period of oscillations.

Exercise 2 (Lissajous figures). Consider the two-dimensional harmonic oscillator. Show that if

$$\frac{\omega_1}{\omega_2} = \frac{r}{s}\,,$$

where r, s are relatively prime integers then there is a new additional integral of motion

$$F = \bar{a}_1^s a_2^r\,,$$

where

$$\bar{a}_1 = \frac{1}{\sqrt{2\omega_1}}(p_1 + i\omega_1 q_1)\,, \qquad a_2 = \frac{1}{\sqrt{2\omega_2}}(p_2 - i\omega_2 q_2)\,.$$

The corresponding closed trajectories of the two-dimensional harmonic oscillator are called the Lissajous figures. Find the Poisson brackets between F and $F_i = \frac{1}{2}(p_i^2 + \omega_i^2 q_i^2),\ \ i = 1, 2$.

Exercise 3. Consider the Kepler problem. Show that the components of the angular momentum J_i and the components of the Runge-Lenz vector R_i form w.r.t. the Poisson bracket the Lie algebra $\mathfrak{so}(4)$. Recall that a 4×4 matrix X belongs to the Lie algebra $\mathfrak{so}(4)$ if it is skew-symmetric, i.e. $X^t + X = 0$. Express the Kepler Hamiltonian in terms of the conserved quantities J_i and R_i.

Exercise 4. Prove that the Poisson bracket

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

between the components of the matrix L implies that the quantities $I_k = \mathrm{tr} L^k$ are in involution, i.e. that $\{I_k, I_m\} = 0$.

Exercise 5 (Calogero model). Consider a dynamical system of n particles with the coordinates q_j and momenta p_j, where $j = 1, \ldots n$. The Hamiltonian of the system is

$$H = \frac{1}{2}\sum_{j=1}^{n} p_j^2 + g^2 \sum_{i<j} \frac{1}{q_{ij}^2}\,.$$

Show that the matrices L and M with the matrix elements

$$L_{jk} = p_j\delta_{jk} + \frac{ig}{q_{jk}}(1-\delta_{jk})\,, \qquad q_{ij} \equiv q_i - q_j$$

$$M_{jk} = -ig\left(\delta_{jk}\sum_{s\neq j}\frac{1}{q_{js}^2} - (1-\delta_{jk})\frac{1}{q_{jk}^2}\right)\,.$$

provide the Lax representation for this dynamical system. Compute $\mathrm{tr}L^2$.

Reading. Study sections 16.1 -16.3 (about Lie algebras and Lie groups) from the book. Read section 3.3 on the coadjoint orbits. There will be 15 minutes writing test on the theoretical notions discussed in these sections.

7.5 Seminar 5

Exercise 1 (Open Toda chain)
Consider a system of n interacting particles described by coordinates q_j and the corresponding conjugate momenta p_j, where $j = 1, \ldots, n$. The Hamiltonian of the system has the form

$$H = \frac{1}{2}\sum_{1}^{n} p_j^2 + \sum_{j=1}^{n-1} \exp[2(q_j - q_{j+1})] \, .$$

Show that equations of motion are equivalent to the Lax equation $\dot{L} = [L, M]$, where

$$L = \sum_{j=1}^{n} p_j E_{jj} + \sum_{j=1}^{n-1} \exp[(q_j - q_{j+1})](E_{j,j+1} + E_{j+1,j}) \, ,$$
$$M = \sum_{j=1}^{n-1} \exp[(q_j - q_{j+1})](E_{j,j+1} - E_{j+1,j}) \, .$$

Here E_{jk} is a matrix which has only one non-zero matrix element equal to 1 standing in the intersection of j's row with k's column.

Exercise 2 (Differential equations for Jacobi elliptic functions).
Using the differential equation for the Jacobi elliptic function $\mathrm{sn}(x, k)$:

$$(\mathrm{sn}'(x, k))^2 = (1 - \mathrm{sn}(x, k)^2)(1 - k^2 \, \mathrm{sn}(x, k)^2) \, .$$

and the identities relating $\mathrm{sn}(x, k)$ with other two functions $\mathrm{cn}(x, k)$ and $\mathrm{dn}(x, k)$ derive the differential equations for $\mathrm{cn}(x, k)$ and $\mathrm{dn}(x, k)$.

Exercise 3 (The cnoidal wave and soliton of the NLS equation)
Consider the non-linear Schrodinger (NLS) equation

$$i\frac{\partial \psi}{\partial t} = -\psi_{xx} + 2\kappa |\psi|^2 \psi \, ,$$

where $\psi = \psi(x, t)$ is a complex-valued function and we assume that $\kappa < 0$. By making single-wave propagating ansatze for the modulus and the phase of $\psi(x, t)$ determine the cnoidal wave solution of the NLS equation. Show that upon degeneration of the elliptic modulus the cnoidal wave turns into a soliton solution of the NLS equation.

Exercise 4 (Hamiltonian formulation of KdV equation)
Consider the following two Poisson brackets on the space of Schwarzian functions $u(x)$:

$$\{u(x), u(y)\} = -\partial_x \delta(x - y) \, .$$

Show that the KdV equation can be viewed as the Hamiltonian equation

$$u_t = \{H, u\} \, ,$$

where

$$H = \int_{-\infty}^{\infty} \left(\tfrac{1}{2} u_x^2 + u^3 \right) .$$

Show that the Poisson structure is degenerate and

$$Q = \int_{-\infty}^{\infty} \mathrm{d}x \; u(x)$$

is the central element of the Poisson bracket.

Exercise 5 (Sine-Gordon Lagrangian)
Consider the Sine-Gordon model with the Lagrangian density

$$\mathscr{L} = \frac{1}{2} \partial_\mu \phi \partial^\mu \phi + \frac{m^2}{\beta^2} (1 - \cos \beta\phi)$$

over two-dimensional Minkowski space-time. Using the canonical formalism construct the Hamiltonian (the generator of time translations) of the model. Using the Noether theorem construct the momentum P (the generator of space translations) and the generator K of Lorentz rotations.

Remark. The generators H, P, K form the Poincáre algebra of two-dimensional space-time.

7.6 Seminar 6

Exercise 1

Prove that the commutators

$$[e_2, e_3] = e_1, \quad [e_1, e_5] = 2e_1\,, \quad [e_2, e_5] = e_2 + e_3 \quad [e_3, e_5] = e_3 + e_4 \qquad [e_4, e_5] = e_4$$

endows the space $\mathbb{R}^5$ with the structure of a Lie algebra. Find the structure tensor (structure constants) of this Lie algebra.

Exercise 2

In the space $\mathbb{R}^3$ define a multiplication

$$[e_a, e_b] = 0\,, \qquad [e_3, e_a] = B_a^b e_b\,,$$

where $a, b = 1, 2$ and B_a^b is a 2×2 matrix. Show that this commutator table endows the space $\mathbb{R}^3$ with the structure of a Lie algebra. Show that this construction allows one to obtain any three-dimensional Lie algebra.

Exercise 3 (The exponential map)

Let X be an element of the Lie algebra $sl(2, \mathbb{R})$. Show that

- if $\det X < 0$ then

$$e^X = \cosh\sqrt{-\det X}\ \mathbb{I} + \frac{\sinh\sqrt{-\det X}}{\sqrt{-\det X}}\ X\,.$$

- if $\det X > 0$ then

$$e^X = \cos\sqrt{\det X}\ \mathbb{I} + \frac{\sin\sqrt{\det X}}{\sqrt{\det X}}\ X\,.$$

Exercise 4 (The exponential map)

Prove an equality

$$\exp\begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & & & & 1 \\ 0 & & \cdots & & \lambda \end{pmatrix} = \begin{pmatrix} e^\lambda & e^\lambda & \frac{e^\lambda}{2!} & \cdots & \frac{e^\lambda}{(n-1)!} \\ 0 & e^\lambda & e^\lambda & \cdots & \frac{e^\lambda}{(n-2)!} \\ \vdots & & & & \vdots \\ 0 & & & & \frac{e^\lambda}{1!} \\ 0 & & \cdots & & e^\lambda \end{pmatrix}.$$

Exercise 5
Prove that for any matrix A the following identity is valid

$$\det(\exp A) = \exp(\mathrm{tr} A)\,,$$

or, equivalently,

$$\exp(\mathrm{tr}\ln A) = \det A\,.$$

Remark. This is very important identity which enters into the proofs of many formulas from various branches of mathematics and theoretical physics. It must always stay with you. Learn it by heart by repeating the magic words "exponent trace of log is determinant".

Exercise 6
Let

$$A = \begin{pmatrix} 0 & -c_3 & c_2 \\ c_3 & 0 & -c_1 \\ -c_2 & c_1 & 0 \end{pmatrix}.$$

Show that the matrices

$$O(c_1, c_2, c_3) = (\mathbb{I} + A)(\mathbb{I} - A)^{-1}$$

belong to the Lie group $SO(3)$. Show that the multiplication operation in $SO(3)$ written in coordinates (c_1, c_2, c_3) takes the form

$$O(c)O(c') = O(c'')$$

where

$$c'' = (c + c' + c \times c')/(1 - (c, c'))\,.$$

Here $c = (c_1, c_2, c_3)$ is viewed as three-dimensional vector.

Exercise 7
Let $G = SU(2)$ and H^{2j} is the space of all homogenious polynomials of degree $2j$, $j = 0, 1/2, 1, \cdots$

$$f(z_1, z_2) = \sum_{n=-j}^{n=j} a_n z_1^{j-n} z_2^{j+n}$$

with $a_n \in \mathbb{C}$. Show that

$$T_j(g) f(z_1, z_2) = f(\alpha z_1 + \gamma z_2, \beta z_1 + \delta z_2)$$

is a representation of the group $SU(2)$. Here

$$g = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

with $\alpha = \bar{\delta}$ and $\gamma = -\bar{\beta}$ is a group element of $SU(2)$.

Exercise 8
Prove that

$$\alpha(\phi) = -\mathbb{I} + \frac{2}{1+c^2t^2} \begin{pmatrix} 1+c_1^2t^2 & c_1c_2t^2 - c_3t & c_1c_3t^2 + c_2t \\ c_2c_1t^2 + c_3t & 1+c_2^2t^2 & c_2c_3t^2 - c_1t \\ c_3c_1t^2 - c_2t & c_3c_2t^2 + c_1t & 1+c_3^2t^2 \end{pmatrix}$$

is a one-parameter subgroup in $SO(3)$, where $\tan\frac{\phi}{2} = ct$, $c^2 = c_1^2 + c_2^2 + c_3^2$ and $\vec{c} = (c_1, c_2, c_3)$ is a constant vector.

Exercise 9
Let

$$B_\varphi = \begin{pmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad C_\theta = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}.$$

Show that any matrix $A \in SO(3)$ can be represented in the form

$$A = B_\varphi C_\theta B_\psi\,.$$

Write the one-parameter subgroup from the exercise 8 in the coordinates (φ, θ, ψ).

7.7 Seminar 7

Exercise 1

Consider the classical Heisenberg model. Show that the formula for the Poisson brackets between the components of the Lax matrix

$$\{U(x,\lambda),U(y,\mu)\} = \Big[r(\lambda,\mu),U(x,\lambda)\otimes\mathbb{I}+\mathbb{I}\otimes U(y,\mu)\Big]\delta(x-y)\,,$$

with the classical r-matrix

$$r(\lambda,\mu)=\frac{1}{2}\frac{\sigma_i\otimes\sigma_i}{\lambda-\mu}\,.$$

implies that the Poisson bracket between the components of the monodromy matrix

$$\mathrm{T}(\lambda)=\mathscr{P}\exp\Big[\int_0^{2\pi}\mathrm{d}x\;U(x,\lambda)\Big]$$

is of the form

$$\{\mathrm{T}(\lambda)\otimes\mathrm{T}(\mu)\}=\Big[r(\lambda,\mu),\mathrm{T}(\lambda)\otimes\mathrm{T}(\mu)\Big]\,.$$

Exercise 2

Show that the Jacobi identity for the Poisson bracket

$$\{\mathrm{T}(\lambda)\otimes\mathrm{T}(\mu)\}=\Big[r(\lambda,\mu),\mathrm{T}(\lambda)\otimes\mathrm{T}(\mu)\Big]\,.$$

implies the classical Yang-Baxter equation for the r-matrix sckew-symmetric $r_{12}(\lambda,\mu)=-r_{21}(\mu,\lambda)$:

$$[r_{12}(\lambda,\mu),r_{13}(\lambda,\nu)]+[r_{12}(\lambda,\mu),r_{13}(\mu,\nu)]+[r_{13}(\lambda,\nu),r_{23}(\mu,\nu)]=0$$

Check (e.g. by using Mathematica) that the r-matrix

$$r(\lambda,\mu)=\frac{1}{2}\frac{\sigma_i\otimes\sigma_i}{\lambda-\mu}\,.$$

solves the classical Yang-Baxter equation.

Exercise 3

Consider the zero-curvature representation for the KdV equation:

$$U=\begin{pmatrix}0 & 1\\ \lambda+u & 0\end{pmatrix}\,,\qquad V=\begin{pmatrix}u_x & 4\lambda-2u\\ 4\lambda^2+2\lambda u+u_{xx}-2u^2 & -u_x\end{pmatrix}\,.$$

Using abelianization procedure around the pole $\lambda=\infty$ find the first four integrals of motion.

Exercise 4

Consider the non-linear Schrodinger equation:

$$i\frac{\partial\psi}{\partial t} = -\frac{\partial^2\psi}{\partial x^2} + 2\kappa|\psi|^2\psi\,,$$

where $\psi \equiv \psi(x,t)$ is a complex function. Show that this equation admits the following zero-curvature representation

$$U = U_0 + \lambda U_1\,, \qquad V = V_0 + \lambda V_1 + \lambda^2 V_2\,,$$

where

$$U_0 = \sqrt{\kappa}(\bar{\psi}\sigma_+ + \psi\sigma_-)\,, \qquad U_1 = \frac{1}{2i}\sigma_3$$

and

$$V_0 = i\kappa|\psi|^2\sigma_3 - i\sqrt{\kappa}(\partial_x\bar{\psi}\sigma_+ - \partial_x\psi\sigma_-)\,, \qquad V_1 = -U_0\,, \qquad V_2 = -U_1\,.$$

Using the abelianization procedure around $\lambda = \infty$ find the first four local integrals of motion. What is the physical meaning of the first three integrals?

7.8 Seminar 8

Exercise 1

Consider XXX Heisenberg model. For the chain of length $L = 3$ find the matrix form of the Hamiltonian as well as its eigenvalues. Construct the corresponding matrix representation of the global $\mathfrak{su}(2)$ generators. How many $\mathfrak{su}(2)$ multiplets the Hilbert space of the $L = 3$ model contains?

Exercise 2

Carry out an explicit construction of the Bethe wave-function $a(n_1, n_2, n_3)$ for three-magnon states of the Heisenberg model. Derive the corresponding Bethe equations.

Exercise 3

Show that on the rapidity plane $\lambda = \frac{1}{2}\cot\frac{p}{2}$ the S-matrix of the Heisenberg model takes the form

$$S(\lambda_1, \lambda_2) = \frac{\lambda_1 - \lambda_2 + i}{\lambda_1 - \lambda_2 - i}\,.$$

Hence, it depends only on the difference of rapidities of scattering particles.

Exercise 4

Show that L two-magnon states of the Heisenberg model with $p_1 = 0$ and $p_2 = \frac{2\pi m}{L}$ with $m = 0, 1, \ldots, L-1$ are $\mathfrak{su}(2)$-descendants of the one-magnon states.