Notes on

Topics in Analysis and PDE: Integrable systems and nonlinear Fourier transforms

Herbert Koch
Universität Bonn
Winter Term 2022/2023

Correction are welcome and should be sent to `koch@math.uni-bonn.de` or told me during office hours. The notes are only for participants of the course V5B1 *Topics in PDE and Analysis* at the University of Bonn, winter term 2022/23. A current version can be found at

`http://www.math.uni-bonn.de/ag/ana/WeSe2223/V5B1_WS_22.html`.

Contents

1 Introduction

Integrable systems and equations are connecting several very different areas in mathematics. The notion itself is only vaguely defined and refers to two different but related properties.

1. Liouville integrable systems are Hamiltonian ODEs for which the general solution can by expressed in terms of integrals and algebraic manipulations. In general the integrals cannot explicitly evaluated and he importance lies in some qualitaive consequences: Typical solutions are dense on tori of half the dimension of the space, and the tori foliate a large set in the space.

   This structure is somewhat stable under perturbation. This is the content of the KAM (Kolmogorov, Arnold, Moser) theory. A related question is the question whether the solar system is stable. If one neclects the attraction between planets one gets an uncoupled system of two body Kepler systems, but it is a delicate question whether the KAM theorie applies, and if it does, what the implicatons for the solar system are.

2. Integrability refers to solvability via formulas. It is not quite clear what that means. However examples show that integrable systems, despite being rare, occur at important places. Two prototypical integrable PDES are the nonlinear Schrödinger equation

   $$i\partial_t u + u_{xx} = \pm 2|u|^2 u$$

   and the Korteweg-de Vries equation

   $$\partial_t u + u_{xxx} - 6u\partial_x u = 0$$

   which are universal asymptotic equaions for wave propagation, for essentially every nonlinear wsystem describing wave propagation.

3. While one may consider integrable systems as very exceptional and particular - even small perturbations destroy integrability - there are not many really different 'integrable structures'. Again this is a vage statement which I cannot make precise. However both NLS and KdV are

linked to a structure which is relevant in algebraic geometry, representation theory, PDEs, wave propagation, random processes and random matrices and statistical physics.

One feature of many integrable equatons is the Lax-Pair structure. Let

$$L\psi = (-\partial^2 + u) \tag{1.1}$$

and

$$B\psi = -4\partial^3\psi + 3(u\partial\psi + \partial(u\psi)) \tag{1.2}$$

Then

$$L_t = [B, L]$$

is formally equivalent to the KdV equation.

$$\begin{aligned}
[-4\partial^3 + 3(\partial u + u\partial), -\partial^2 + u] &= -4[\partial^3, u] + 3([2u\partial + u', -\partial^2] + [2u\partial + u', u]) \\
&= -4u^{(3)} - 12u''\partial - 12u'\partial^2 + 12u'\partial^2 + 6u''\partial + 3u^{(3)} + 6u''\partial + 6uu' \\
&= -u^{(3)} + 6uu'.
\end{aligned}$$

The remarkable feature is that the commutator of the differential operators is a multiplication operator.

Consider the system

$$L(u)\psi = z^2\psi, \qquad \partial_t\psi = B(u)\psi.$$

Suppose that the initial value problem

$$\psi_t = B(u(t))\psi, \qquad \psi(0) = \psi_0$$

is uniquely solvable and defines a map

$$\psi_0 \to U(t,0)\psi_0 = \psi(t)$$

on L^2. Then, at least formally

$$\begin{aligned}
\frac{d}{dt}\Big(L(u(t)) - U(0,t)L(u(0))U(0,t)^{-1}\Big)& \\
= [B(u(t), L(u(t))] &- B(u(t))(U(0,t)L(u(0))U(0,t)^{-1} \\
&+ U(0,t)L(u(0))(U(0,t)^{-1}B(u(t)) \\
= [B(u(t)), L(u(t)) &- U(0,t)L(u(0))U(0,t)^{-1}]
\end{aligned}$$

and since the identity holds for $t = 0$

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

For matrices we would say that $L(u_0)$ and $L(u(t))$ are similar, i.e. they are the same up to the choice of a basis. In particular the spectrum does not change under the evolution.

We may rewrite the pair of equations

$$L\psi = z^2\psi, \quad \psi_t = B\psi$$

as a two by two system of the form

$$\partial_x \phi + A(z)\phi = 0, \qquad \partial_t \Phi + B(z)\phi = 0.$$

The Cauchy problem for linear equations can always be solved. We can solve the general initial value problem consistently and simultaneously iff a compatibility condition is satisfied,

$$0 = [\partial_x + A, \partial_t + B] = \partial_x B - \partial_t A + [A, B].$$

This can be understood as a vanishing curvature condition for a connection of a two dimensional vector bundle, which is again equivalent to the Korteweg-de Vries equation.

The study of the KdV equation is strongly related to the spectral properties of the Lax operator, which is a linear object.

Outline:

1. Symplectic structures ad Hamiltonian equations [1]
2. Examples of integrable ODEs ([6, 22, 2])
3. The Korteweg-de Vries hierarchy ([2, 8],,
4. The KdV equation at H^{-1} after Killip and Visan [16]
5. If time permits: Global solutions to the defocusing Nonlinear Schrödinger equation

2 Hamiltonian equations

Reference: Arnold [1] .

2.1 Vector fields and flows

Definition 2.1. *A Lie algebra is a vector space E with an skew symmetric bilinear map called Lie bracket*

$$E \times E \ni (a,b) \to [a,b] \in E$$

which satisfies the Jacobi identity

$$[[a,b],c] + [[b,c],a] + [[c,a],b] = 0.$$

Skew symmetry means

$$[a,b] + [b,a] = 0$$

Examples:

1. Matrices with the commutator $[A,B] = AB - BA$.
2. Bounded operators on Banach spaces.
3. The Virasoro algebra. Let $X = \{f\partial\}$ be the smooth complex vector fields on the unit circle. The vector fields together with the commutator form Lie algebra by the calculation above. We define on $X \times \mathbb{C}$ a Lie bracket
$$[(\xi\partial, \alpha), (\zeta\partial, \beta)] = \left([\xi\partial, \zeta\partial], \frac{1}{24\pi} \int_0^{2\pi} \xi' \zeta'' dx \right)$$

We consider smooth vector fields X on open sets $U \subset \mathbb{R}^d$,

$$X = (a_j(x))_{1 \le j \le d}$$

which we identify with the differential operator

$$f \to f_X = \partial_X f = \sum_{j=1}^d a_j \partial_j f.$$

We use the same notation for vector valued functions.

The commutator is defined by

$$[X,Y]f = \partial_X \partial_Y f - \partial_Y \partial_X f = \sum_{j,k=1}^d [a_j(\partial_j b_k) - b_j \partial_j a_k)] \partial_k f$$

and it is again a smooth vector field. It satisfies

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

and it is a derivation (it satisfies the Leibniz formula)

$$\partial_X(fg) = f\partial_X g + g\partial_X f.$$

We denote the set of smooth vector fields by $C^\infty(U;\mathbb{R}^d)$. We have the obvious bilinear map

$$C^\infty(U) \times C^\infty(U;\mathbb{R}^d) \ni f \times X \to fX = (fa_j)_{1\le j\le d} \in C^\infty(U,\mathbb{R}^d).$$

Consider the ODE

$$\frac{d}{dt}x(t) = X(x(t)), \quad x(0) = x_0.$$

It has unique solution $x(t) = \Phi(t,x_0) = \Phi_X(t,x_0)$ on a maximal times interal $0 \in (t_-,t_+)$. Then

$$\frac{\partial}{\partial t}\Phi(t,x_0) = X(\Phi(t,x_0))$$

and

$$\frac{\partial}{\partial t}\Big(\frac{\partial}{\partial x_0^j}\Phi(t,x_0)\Big) = DX(\Phi_X(t,x_0))\Big(\frac{\partial}{\partial x_0^j}\Phi(t,x_0)\Big)$$

where the first bracket denotes the argument of DX.

We will ignore the restriction of t to this interval whenever it is unimportant.

Lemma 2.2. *The flows commute, i.e.*

$$\Phi_X(t,\Phi_Y(s,x)) = \Phi_Y(s,\Phi_X(t,x)) \tag{2.1}$$

for all x and $|s|,|t|$ small if and only if the vector fields commute, i.e.

$$[X,Y] = 0. \tag{2.2}$$

Proof. Suppose the flows commute (2.1). We differentiate with respect to t:

$$\begin{aligned} 0 &= \frac{d}{dt}(\Phi_X(t,\Phi_Y(s,x)) - \Phi_Y(s,\Phi_X(t,x))) \\ &= X(\Phi_X(t,\Phi_Y(s,x))) - D_x\Phi(s,\Phi_X(t,x))X(\Phi_X(t,x)). \end{aligned}$$

We evaluate at $t=0$

$$0 = X(\Phi_Y(s,x)) - D_2\Phi(s,x)X(x). \tag{2.3}$$

We differentiate with respect to s (denoting D_2 the differentiation with respect to the second variable)

$$0 = DX(\Phi(s,x))Y(\Phi_Y(s,x))) - DY(\Phi(s,x))D_2\Phi(s,x)X(x). \tag{2.4}$$

The evaluation at $s = 0$ gives (2.2). Vice versa: We assume that the vector fields commute (2.2). We first prove that then (2.3) holds. First using commutation for the first term

$$\begin{aligned} &\frac{d}{ds}\Big[X(\Phi_Y(s,x) - D_2\Phi(s,x)X(x)\Big] \\ &\qquad = DX(\Phi_Y(s,x))Y(\Phi_Y(s,x)) - DY(\Phi_Y(s,x))D_2\Phi_Y(s,x)X(x) \\ &\qquad = DY(\Phi_Y(s,x))\Big[X(\Phi_Y(s,x)) - D_2\Phi_Y(s,\Phi_X(t,x)X(x)\Big] \end{aligned}$$

where D_2 denotes the derivative with respect to the second variable. Hence, with $f(s) = |(X(\Phi_Y s, x) - D_2\Phi_Y(s,\Phi_X(t,x)X(x)|$

$$\Big|\frac{d}{ds}f(s)\Big| \le \sup |DY|_{L(\mathbb{R}^n,\mathbb{R}^n)}|f(s)|$$

(if $f(s) > 0$) where the supremum is taken over the evaluation over

$$\{\Phi_Y(s,x) : s \in [-T,T]\}$$

for some $T > 0$. Now Gronwall's inequality implies

$$f(t) \le e^{C|t|}f(0) = 0$$

hence (2.4). We fix s. Then, using (2.4),

$$\begin{aligned} &\frac{d}{dt}\Big[\Phi_X(t,\Phi_Y(s,x)) - \Phi_Y(s,\Phi_X(t,x))\Big] \\ &\qquad = \Big[X(\Phi(t,\Phi_Y(s,x))) - D_2\Phi_Y(s,\Phi_X(t,x))X(\Phi_X(t,x))\Big] \\ &\qquad = X(\Phi(t,\Phi_Y(s,x))) - X(\Phi_y(s,\Phi_x(t,x)) \end{aligned}$$

and with

$$f(t) = |\Phi_X(t,\Phi_Y(s,x)) - \Phi_Y(s,\Phi_X(t,x))|$$

and deduce $f(0) = 0$ and

$$|\frac{d}{dt}f(t)| \le Cf(t)$$

where C is the supremum of the Lipschitz constant of X over a suitable set. This implies (2.4) and concludes the proof. □

Let X_j be N vector fields which pairwise commute. Given $(t_j)_{1\le j\le N}$. We define

$$\Phi((t_j),x) = \Phi_{X_N}(t_N,\Phi_{X_{N-1}}(t_{N-1},\dots\Phi_{X_1}(t_1,x))$$

Let σ be a permutation of the indices. By an iterative use (more precisely a double induction) of Lemma 2.2

$$\Phi((t_j),x) = \Phi_{X_{\sigma(N)}}(t_{\sigma(N)},\Phi_{X_{\sigma(N-1)}}(t_{\sigma(N-1)},\dots\Phi_{X_{\sigma(1)}}(t_{\sigma(1)},x)).$$

Theorem 2.3 (Frobenius)**.** *Let X_j, $1 \le j \le N$ be vector fields which are linearly independent at every point. Then the following is equivalent.*

1. *The commutators evaluated at any point ly in the span of the vector fields, i.e. we can uniquely write*

$$[X_j, X_k] = \sum_{l=1}^{N} a_{jkl}(x) X_l.$$

2. *There is a foliation of a neighorhood of any point by smooth manifolds of dimension N so that the vector fields are tangent. Equivalently there are coordinates so that all the vector fields have values in $\mathbb{R}^N \times \{0\} \subset \mathbb{R}^n$.*

Proof. We assume $0 \in U$ and argue in a neighborhood of zero. After a linear transform we may assume that

$$X_j(0) = e_j(0).$$

We write $X_j(x) = \sum_{k=1}^n c_{jk} e_k$. Then $C = (c_{jk})_{1 \le j,k \le N}$ is invertible near 0. We define

$$\tilde X_j = \sum_{k=1}^{N} (C^{-1})_{jk} X_k$$

so that the $X_j = \sum_{k=1}^N C_{jk}(x) \tilde X_k$. We may replace the X_k by $\tilde X_k$ since neither assumption nor conclusion changes. We have

$$\tilde X_j(x) = e_j + \sum_{k=N+1}^{n} c_{jk} e_k. \tag{2.5}$$

We claim that the vector fields $\tilde X_j$ commute: We can write

$$[\tilde X_j, \tilde X_k] = \sum_{l=N+1}^{n} a_{jkl} e_l = \sum_{l=1}^{N} b_{jkl} \tilde X_l.$$

The first equality follows from (2.5) and the second holds since the commutators are in the span of the vector fields. Since the first N components of the commutator vanish (by the middle term) the b_{jkl} vanish and the vector fields commute.

We define a map

$$\mathbb{R}^n = \mathbb{R}^N \times \mathbb{R}^{n-N} \subset V \ni ((t_j)_{1 \le k \le N}, (y_j)_{1 \le k \le n-N}) \to \Phi(t, (0, y)) \in \mathbb{R}^n$$

which is smooth. By construction $D\Phi(0) = 1_{\mathbb{R}^n}$, hence it is a local diffeomorphism. In V we have the trivial foliation with the leafs $S_\alpha = \{(t,\alpha) \in V\}$. The submanifolds $\Phi(S_\alpha)$ foliate a neighborhood of zero.

Vice versa: If there exists such a map then if $x = \Phi(t,(0,y))$

$$X_j(x) = \partial_{t_j}\Phi(t,(0,y)).$$

$\square$

2.2 The symplectic structure

2.2.1 Symplectic vector spaces

Let $\mathbb{K}$ be $\mathbb{R}$ or $\mathbb{C}$. E be an n dimensional $\mathbb{K}$ vector space.

Lemma 2.4. *Let $\omega : E \times E \to \mathbb{K}$ by an antisymmetric bilinear map. There exists basis e_i so that for some $N \le d/2$*

$$\omega\Big(\sum_{j=1}^n a_j e_j, \sum_{j=1}^n b_j e_j\Big) = \sum_{j=1}^N a_j b_{j+N} - a_{j+N} b_j \tag{2.6}$$

Proof. Let e_1 e some vector so that there is another vector f so that $\omega(e_1, f) \ne 0$. We define

$$e_2 = \frac{1}{\omega(e_1,f)} f.$$

Suppose we have constructed $(e_j)_{1 \le j \le 2J}$. Suppose there exists e_{2J+1} and f so that

$$\omega(e_j, e_{2J+1}) = \omega(e_j, f) = 0, \quad \omega(e_{2j+1}, f) \ne 0$$

for $1 \le j \le 2J$. Then we continue recursively until the procedure stops because

$$\omega(e_j, e) = \omega(e_j, f) = 0 \quad \text{for } 1 \le j \le 2J \quad \Longrightarrow \omega(e,f) = 0$$

We complement the basis and claim that then (2.6) with $J = N$holds. $\square$

Definition 2.5. *We call ω a sympletic form if $2N = d$ in the construction above.*

Let ω be an antisymmetric bilinear map. We define the linear map

$$B : E \to E^*$$

by

$$B(x)(y) = \omega(x,y).$$

Then ω is a symplectic form if and only if B is invertible. In that case we denote the inverse by $J : E^* \to E$.

On $\mathbb{K}^{2n} \ni (p, q)$ there is the cannonical symplectic form

$$\omega_0((p_1, q_1), (p_2, q_2)) = p_1 \cdot q_2 - p_2 \cdot q_1$$

In this case

$$J = \begin{pmatrix} 0 & 1_{\mathbb{R}^n} \\ -1_{\mathbb{R}^n} & 0 \end{pmatrix}, \quad J^{-1} = \begin{pmatrix} 0 & -1_{\mathbb{R}^n} \\ 1_{\mathbb{R}^n} & 0 \end{pmatrix}.$$

Definition 2.6. *Let (E, ω) be a $d = 2N$ dimensonal symplectic vector space. A subvector space $F \subset E$ of dimension N is called Lagrangian if $\omega|_{F\times F} = 0$.*

In $(\mathbb{R}^{2N}, \omega_0)$ we see that Lagranigian subspaces exist. There cannot be a subspace of higher dimension so that the restricton vanishes.

Definition 2.7. *Let (E, ω) and (F, μ) be a symplectic vector spaces. An invertible linear map $A : E \to F$ is called a symplectic if*

$$\omega(z_1, z_2) = \mu(Az_1, Az_2).$$

The symplectic group is the group of symplectic linear maps from $(\mathbb{R}^{2n}, \omega_0)$ to itself.

It is not hard to make the condition for a matrix to be symplectic precise. Let $z_j = (x_j, x_j)$ ánd consider the matrix

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

We want to check under which conditions it is symplectic. First every symplectic operator is invertible. Since

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} Ax + By \\ Cx + Dy \end{pmatrix}$$

and

$$\omega_0 \left(\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} \right) = x_1^T y_2 - y_1^T x_2,$$

$$\omega_0 \left(\begin{pmatrix} Ax_1 + By_1 \\ Cx_1 + Dy_1 \end{pmatrix}, \begin{pmatrix} Ax_2 + By_2 \\ Cx_2 + Dy_2 \end{pmatrix} \right)$$
$$= (Ax_1 + By_1)^T (Cx_2 + Dy_2) - (Cx_1 + Dy_1)^T (Ax_2 + By_2)$$

It is symplectic iff

- A^TC, B^TD are symmetric and $A^TD - C^TB = 1_{\mathbb{K}^N}$
- AB^T , CD^T are symmetric and $AD^T - BC^T = 1_{\mathbb{K}^N}$

A 2×2 matrix is symplectic iff its determinant is 1.

We will be almost exclusively be concerned with submanifolds of $\mathbb{R}^d$. Here we are even more restrictive and consider only open subsets of $\mathbb{R}^{2N}$. Using local coordinates we obtain similar constructions on manifolds.

2.2.2 Primer on differential forms

Consider skew k linear maps $(\mathbb{R}^n)^k \to \mathbb{R}$, which we denote by Λ^k. They are k linear maps T

$$(v_1, \dots, v_k) \to T(v_1, \dots, v_k) \in \mathbb{R}$$

so that the value changes sign whenever we exchange to vectors. Equivalently T is zero if the vectors are linearly dependent. A basis is given by

$$dx_{j_1} \wedge dx_{j_2} \cdots \wedge dx_{j_k}$$

with

$$1 \le j_1 < j_2 < \dots j_k.$$

They are defined by

$$dx_{j_1} \wedge dx_{j_2} \cdots \wedge dx_{j_k}(e_{i_1}, e_{i_2} \dots e_{i_k}) = \begin{cases} \pm 1 & \text{if } (i_k) \text{ is a permutation of } (j_k) \\ & \text{and the sign is the sign of} \\ & \text{the permutation} \\ 0 & \text{otherwise} \end{cases}$$

We define the smash product of a k_1 resp. k_2 form by the basis. The smash product of two basis forms vanishes unless all indices are different. We then permute step by step. If $X \in \mathbb{R}^n$ and ω is an skew k linear map we define the $k-1$ form $i_X\omega$ by

$$(i_X\omega)(X_2, \dots X_k) = \omega(X, X_2, \dots X_k).$$

A differential k form on an open set $U \subset \mathbb{R}^n$ is a (continuous / smooth /analytic) map U to the skew k linear maps. Typically we express them with respect to the basis above.

We define the exterior derivative of a C^1 k form, which is a continuous $k+1$ form via

$$d(f dx_{j_1} \wedge dx_{j_2} \cdots \wedge dx_{j_k}) = df \wedge dx_{j_1} \wedge dx_{j_2} \cdots \wedge dx_{j_k}.$$

We say a form is closed, if its exterior derivative vanishes, and exact if it is an exterior derivative. On easily sees that always

$$dd\omega = 0$$

and hence exact implies closed. The Poincaré lemma say that on star shaped closed forms are exact.

If $\phi : U \to V$ is C^1 map and if ω is a continuous k form on V we define the pull back by

$$\phi^*\omega(x; X_1, \dots X_k) = \omega(\phi(x); D\Phi(x)X_1, \dots, D\Phi(x)X_2).$$

The operations are all compatible:

$$\phi^*(\omega \wedge \nu) = \phi^*\omega \wedge \phi^*\nu$$

$$d\phi^*\omega = \phi^* d\omega.$$

If X is a vector field and Φ_X is the flow we define the Lie derivative

$$\mathcal{L}_X\omega = \frac{d}{dt}\Phi^*(t)\omega\Big|_{t=0}$$

Then

$$\mathcal{L}_x\omega = i_X d\omega + d(i_X\omega) \tag{2.7}$$

2.2.3 Definition of symplectic manifolds and first properties

We work in local coordinates and hence argue on $\mathbb{R}^n$, even when we formulate statements about manifolds.

Definition 2.8. *A symplectic form ω is a closed 2 form on a $2N$ dimensional manifold M^{2N} of maximal rank at every point. (M^{2N}, ω) is called symplectic.*

By the Poincaré lemma locally any sympectic form is exact and there exists a 1 form α so that $\omega = d\alpha$.

We again may ask: Can we choose coordinates so that the symplectic form becomes simple? The answer is yes.

Theorem 2.9 (Darboux theorem)**.** *Let (M^{2n}, ω) be a symplectic manifold. For $z \in M$ there exists a neighborhood $U \subset M$ and a diffeomorphism*

$$\phi : U \to \mathbb{R}^{2n}$$

so $\omega = \phi^*\omega_0$.

Proof. It suffices to prove the statement for a symplectic form on an open set $U \subset \mathbb{R}^{2N}$. We use an elegant argument which goes back to Moser. Arnold [1], Section 43 B provides a different proof.

Let ω by a symplectic from on $B_r(0) \subset \mathbb{R}^{2N}$. By a linear change of coordinates and Lemma 2.4 we may assume that $\omega(0) = \omega_0$ and, decreasing r if necessary

$$\omega_t = (1-t)\omega_0 + t\omega$$

is a family of symplectic forms parametrized over $[0,1]$.

We want to construct a family of vector fields $X(t,x)$ with $X(t,0) = 0$ and maps $\Psi(t,x)$ defined by

$$\Psi(0,x) = x, \quad \frac{\partial}{\partial t}\Psi(t,x) = X(t, \Psi(t,x))$$

so that $\Psi^*(t)\omega_t = \omega_0$. We differentiate the left hand side. Then

$$\begin{aligned}
0 &= \frac{d}{dt}\omega_0(V,W) = \frac{d}{dt}(\Psi^*(t)\omega_t)(x)(V,W) \\
&= \frac{d}{dt}\omega_t(\Psi(t,x))(D\Psi(t,x)V, D\Psi(t,x)W) \\
&= \Big(\Psi^*(t)(\omega - \omega_0) + \Psi^*\mathcal{L}_X\omega_t\Big)(V,W) \\
&= \Psi^*(t)(\omega - \omega_0 + di_X\omega_t + i_X d\omega_t)(V,W) \\
&= \Psi^*(t)d(\alpha + i_X\omega_t)(V,W)
\end{aligned}$$

The symplectic forms are closed and we write $\omega_2 - \omega_1 = d\alpha$. We observe $i_X\omega_t(x) = J_t^{-1}(x)X(x)$ hence $X(x) = -J_t(x)\alpha(X)$ where J_t is the map connected to the sumplectic form $\omega_t(x)$. Now we read the equalities in the opposite direction. □

There is an extension with the same proof.

Theorem 2.10 (Darboux-Weinstein theorem)**.** *Let (M^{2n}, ω) be a symplectic manifold, $N^d \subset M^{2n}$ a submanifold. Let $\tilde\omega$ be second symplectic forms which coincides with ω on N. Then there exist two neighborhood U_1 and U_2 of N and a diffeomorphism $\phi : U_1 \to U_2$ so that*

$$\omega = \phi^*\tilde\omega.$$

Definition 2.11. *Let (M^{2n}, ω_M) and (N^{2n}, ω_N) be symplectic manifolds. A symplectomorphism is a diffeomorphism $\phi : M^{2n} \to N^{2n}$ so that*

$$\omega_M = \phi^*\omega_N.$$

2.3 Hamiltonians and Hamiltonian vector fields

Let (M^{2n}, ω) be an open set with a symplectic form and $H \in C^1(M^{2n})$. We define the Hamiltonian vector field as the unique vector field ∇_H so that for all V in the tangent space at x

$$dH(x)(V) = \omega(x)(V, \nabla_H(x)).$$

We can work this out for ω_0 where

$$\nabla_H = -J\nabla H = \begin{pmatrix} -\frac{\delta H}{\partial q_1} \\ \vdots \\ -\frac{\partial H}{\partial q_n} \\ \frac{\partial H}{\partial p_1} \\ \dots \\ \frac{\partial H}{\partial p_n} \end{pmatrix}.$$

The Hamiltonian equations are

$$\frac{d}{dt} p_j = -\frac{\partial H}{\partial q_j}, \quad \frac{d}{dt} q_j = \frac{\partial H}{\partial p_j} \tag{2.8}$$

In the sequel we consider smooth Hamiltonians, even if the statement and definition require only a finite number of derivatives. We define the Poisson bracket of two smooth functions by

$$\{f, g\} = \omega(Jdg, Jdf) \tag{2.9}$$

with with ω_0 becomes

$$\{f, g\} = \omega_0 \left(\begin{pmatrix} -\partial_{q_1} g \\ \vdots \\ -\partial_{q_n} g \\ \partial_{p_1} g \\ \vdots \\ \partial_{p_n} g \end{pmatrix}, \begin{pmatrix} -\partial_{q_1} f \\ \vdots \\ -\partial_{q_n} f \\ \partial_{p_1} f \\ \vdots \\ \partial_{p_n} f \end{pmatrix} \right) = \sum_{j=1}^n (\partial_{p_j} g \partial_{q_j} f - \partial_{q_j} g \partial_{p_j} f).$$

The definition of the Poisson bracket is independent of the coordinates. It satisfies

$$\nabla_g f = \{f, g\} \tag{2.10}$$

which we check for $(\mathbb{R}^{2m}, \omega_0)$ and $\begin{pmatrix} p \\ q \end{pmatrix} \in \mathbb{R}^{2n}$,

$$\nabla_g f = \sum_{j=1}^n -\partial_{q_j} f \partial_{p_j} g + \partial_{p_j} f \partial_{q_j} g,$$

skew symmetry

$$\{f, g\} + \{g, f\} = 0, \tag{2.11}$$

which is clear from the definition and the Jacobi identity

$$\{\{f, g\}, h\} + \{\{g, h\}, f\} + \{\{h, f\}, g\} = 0. \tag{2.12}$$

To verify the Jacobi identity we observe that schematically

$$\{\{f, g\}, h\} = Dh(DfD^2g + DgD^2f).$$

We will show that we can write the cyclic Poisson bracked above so that there are no second derivatives on f, which by symmetry implies the Jacobi identity.

We compute using (2.10) and skew symmetry

$$\{\{f, g\}, h\} + \{h, f\}, g\}) = \nabla_h \nabla_g f - \nabla_g \nabla_h f = [\nabla_h, \nabla_g] f$$

The commutator is a first order operator, which proves the claim. The Leibniz formula holds

$$\{f, gh\} = \{f, g\}h + \{f, h\}g. \tag{2.13}$$

There are compatibility equation between the Poisson bracket for functions and the commutator of vector fields

$$\nabla_{\{f,g\}} = [\nabla_g, \nabla_f]. \tag{2.14}$$

since

$$\begin{aligned}
\nabla_{\{f,g\}} h &= \{h, \{f, g\}\} = -\{\{f, g\}, h\} \\
&= \{\{g, h\}, f\} + \{\{h, f\}, g\} \\
&= -\nabla_f \nabla_g h + \nabla_g \nabla_f h
\end{aligned}$$

Lemma 2.12. *Suppose that $\{H, g\} = 0\}$. Then g is constant along the flow of H. In particular H is constant on the Hamiltonian flow of itself. The Hamiltonian flow is a (local) symplectomorphism.*

Proof.

$$\frac{d}{dt} g(\Phi_{\nabla_H}(t, x)) = dg(\nabla_H)(\Phi_{\nabla_H}(t) = \{H, g\} \circ \Phi_{\nabla_H}(t) = 0.$$

We have to prove

$$\mathcal{L}_{\nabla H} \omega = 0.$$

This is a local statment and it sufices to prove it in $(\mathbb{R}^{2n}, \omega)$, $(p,q) \in \mathbb{R}^{2n}$,

$$\omega_0 \left(\begin{pmatrix} p_1 \\ q_1 \end{pmatrix}, \begin{pmatrix} p_2 \\ q_2 \end{pmatrix} \right) = (p_1^T, q_1^T) J \begin{pmatrix} p_2 \\ q_2 \end{pmatrix}.$$

We compute

$$\begin{aligned} &\frac{d}{dt} (p_1^T, q_1^T)(D\Phi)^T(t,x) J D\phi(t,x) \begin{pmatrix} p_2 \\ q_2 \end{pmatrix} \Big|_{t=0} \\ &\qquad = (p_1^T, q_1^T) \Big((D\nabla_H)^T J + J D\nabla_H \Big) \begin{pmatrix} p_2 \\ q_2 \end{pmatrix} \Big|_{t=0} \end{aligned}$$

and

$$\begin{aligned} ((D\nabla_H)^T J + J D\nabla_H) &= \begin{pmatrix} (D^2_{pq}H)^T & -D^2_{pp}H \\ D^2_{qq}H & -D^2_{pq}H \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \\ &\quad + \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} D^2_{pq}H & D^2_{qq}H \\ -D^2_{pp}H & -(D^2_{pq}H)^T \end{pmatrix} \\ &= \begin{pmatrix} D^2_{pp}H & (D^2_{pq}H)^T \\ D^2_{pq}H & D^2_{qq}H \end{pmatrix} + \begin{pmatrix} -D^2_{pp}H & -(D^2_{pq}H)^T \\ -D^2_{pq}H & -D^2_{qq}H \end{pmatrix} \\ &= 0 \end{aligned}$$

$\square$

Examples:

1. The free motion. $H(p,q) = \frac{1}{2m}|p|^2$ where m is the mass. The Hamiltonian equations are

$$\frac{d}{dt} p_j = 0, \quad \frac{d}{dt} q_j = \frac{1}{m} p_j.$$

Here q is the position, p the momentum which is m times the velocity of the free particle. The general solution is

$$q(t) = q(0) + \frac{t}{m} p.$$

2. The harmonic oscillator, $n = 1$, $H(p,q)) = \frac{1}{2}(|p|^2 + |q|^2)$. The Hamiltonian equations are

$$\frac{d}{dt} q = p, \qquad \frac{d}{dt} p = -q$$

The general solution is

$$\begin{pmatrix} p(t) \\ q(t) \end{pmatrix} = \begin{pmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{pmatrix} \begin{pmatrix} p(0) \\ q(0) \end{pmatrix}.$$

3. The anharmonic oscillator, $n = 1$, $H(p, q) = \frac{1}{2m}|p|^2 + V(q)$ $V(q) = -L\cos(q)$. The Hamiltonian equations are

$$\frac{d}{dt}p = -L\sin(q), \quad \frac{d}{dt}q = \frac{1}{m}p.$$

We cannot solve explicitly but the orbits are on the level sets of H.

4. Consider, $I_j = \frac{1}{2}(p_j^2 + q_j^2)$

$$H(p, q) = h\big((\frac{1}{2}(I_j)_{1\le j\le n})\big)$$

Then

$$\{H.I_j\} = 0$$

and the Hamiltonian equations are

$$\dot{p}_j = -\frac{\partial H}{\partial I_j}\frac{\partial I_j}{\partial q_j} = -\frac{\partial H}{I_j}q_j$$

$$\dot{q}_j = \frac{\partial H}{\partial I_j}p_j.$$

The action variables are I_j and the angle variables are given by polar coordinates. In polar coordinates

$$dp_j \wedge dq_j = \frac{1}{2}dr^2 d\theta.$$

Definition 2.13. *We say two functions f and g Poisson commute if*

$$\{f, g\} = 0.$$

Definition 2.14. *We call a Hamiltonian function H Liouville integrable on U if there exists n functions F_j which Poisson commute with one another, and also with H so that the rank of $\{dF_j\}$ is n at every point.*

Of course H may be one of the vector fields.

The span of $(\nabla_{F_j})_{1\le j\le n}$ is a Lagrangian subspace, and there cannot be more Poisson commuting functions so that the rank of the Hamiltonian vector fields is the number of the functions.

By the previous considerations the functions are conserved under the flow, and hence the flow lines stay on the level set.

Theorem 2.15 (Liouville-Arnold integrability)**.** *Let H be Liouville integrable on an open set M^{2n} with the Poisson commuting functons $(F_j)_{1\le j\le n}$ which also Poisson commute with H. Let $\alpha \in \mathbb{R}^n$ so that the level set*

$$\mathcal{T} := \{z \in U : F_j(z) = \alpha_j \quad \textit{for } 1 \le j \le n\}$$

is nondegenerate, connected and compact. Then there exists a ball $B_r(\alpha)$ and an open set $V \subset \mathbb{R}^n$ so that the following holds: We equipp $V \times (\mathbb{S}^1)^n \ni (I,\theta)$ with the symplectic form

$$\omega_0 = d \sum_{j=1}^n I_j d\theta_j$$

and a symplectomorphism

$$\Phi : \{z : (F_j(z)) \in B_r(\alpha) \to V \times (\mathbb{S}^1)^n.$$

The $I_j \circ \Phi$ Poisson commute with the F_j and with H. There exists a map $\phi : B_r \to V$ so that $\phi(\alpha)$ is the first component of $\Phi(z)$ iff $F(z) = \alpha$,

We consider the I_j as functions on U. The are called action variables and θ are called angles. The actions I_j Poisson commute. In particular level sets $I = I_0$ are level set $F = \alpha$ and the map $I \to \alpha$ is a diffeomorphism. As a consequence the level ses of F are n dimensional tori.

The Hamiltonian H can be written as a function of the action variables

$$H = h(I)$$

and the Hamiltonian equations become

$$\frac{d}{dt} I_j = 0, \quad \frac{d}{dt}\theta_j = \frac{\partial h}{\partial I_j} =: \omega_j$$

which we can solve as

$$\theta_j(t) = \theta_j(0) + t\omega_j$$

modulo 2π.

Proof. Let

$$\Theta : \mathbb{R}^n \ni t \to \Phi_F(t, z_0).$$

By assumption

$$d\Theta = \sum \nabla_{F_i} dt_i$$

has rank n since the vector fields ∇_{F_i} are linearly independent. The range is open in $\mathcal{T}$ since it has dimension n. It is also closed: Let $z \in \mathcal{T}$ be in the closure of the range. Then the flows map a neighborhood of 0 to a neighborhood of z in $\mathcal{T}$. Thus there exists $s, t \in \mathbb{R}^n$ so that

$$\Phi(t, z_0) = \Phi(s, z)$$

hence

$$\Phi(t - s, z_0) = z.$$

Then

$$T = \{t : \Phi(t) = z_0\}$$

is a discrete additive subgroup of $\mathbb{R}^n$. It does not depend on z_0

Lemma 2.16. *Discrete additive subgroups $T \subset \mathbb{R}^n$ are generated by at most n vectors in $\mathbb{R}^n$.*

Proof. No ball contains more than a finite number of elements of the discrete subgroup. Let $e_1 \neq 0$ be one of the vectors with the smallest norm and let X_1 be its span. Then $T \cap X_1 = \{ne_1 : n \in \mathbb{Z}\}$. Otherwise we would find an element of T closer to the origin.

We define e_j recursively and denote by X_j the span of the first j vectors and assume that $T \cap X_j$ is generated by the e_j. Suppose that we have defined X_j. We claim

$$d(T \backslash X_j, X_j) > 0.$$

If not there exists a sequence $g_k \in G \backslash X_j$ so that

$$d(g_k, X_j) \to 0$$

Adding a linear integer linear combination of the e_k, $1 \leq k \leq j$ we may assume that the sequence is bounded, and taking a subsequence converges to some $g_0 \in G$ since G is discrete. But this contradicts the discreteness. We take $e_{j+1} \in G$ closest to X_j and claim that $(e_k)_{k \leq j+1}$ generate $G \cap X_{j+1}$. Suppose not. Then there exists $g \in G \cap X_{j+1}$ with distance to X_j at most half the distance of e_{j+1} to X_j. This is a contradiction. This process has to stop at latest at $j = n$. □

The map Θ is a diffeomorphism of $\mathbb{R}^n/T \to \mathcal{T}$. Then T has to have rank n since $\mathcal{T}$ is compact. As a consequence level sets in a neighborhood consist of n dimensional tori. Decreasing the manifold we may assume that it is foliated by tori invariant under the flow.

We fix a smooth maps

$$B_r^{\mathbb{R}^n}(\alpha_0) \ni \alpha \to z_\alpha \in \mathcal{T}(\alpha)$$

and $e_n(\alpha)$, $T(\alpha)$ the stabilizer of Φ_α generated by $e_n(\alpha)$ and

$$B_r^{\mathbb{R}^n}(\alpha_0) \times (\mathbb{S}^1)^n \to \Phi(\sum s_j e_j(\alpha), z(\alpha)) \in W \subset M^{2n}$$

which is a diffeomorphism. The tangent spaces of $(\mathbb{S}^1)^n$ are Lagrangian. The symplectic form is independent of the t variables, hence

$$\omega = \sum_{i,j=1}^{n} \omega_{ij}(\alpha) d\alpha_j \wedge dt_i + \gamma_{ij}(\alpha) d\alpha_i \wedge d\alpha_j.$$

The form is closed, hence

$$0 = d\omega = \sum_{ijk=1}^{n} \frac{\partial \omega_{ij}(\alpha)}{\partial \alpha_k} d\alpha_j \wedge d\alpha_k \wedge dt_i + \frac{\partial \gamma_{ij}}{\partial \alpha_k} d\alpha_k \wedge d\alpha_i \wedge d\alpha_j$$

Thus

$$d(\omega_{ij}(\alpha) d\alpha_j) = 0 \quad \text{ and } d\gamma_{ij}(\alpha) d\alpha_i d\alpha_j = 0.$$

By the Poincaré lemma there exist functions f_i so that $dI_i = \omega_{ij}(\alpha) d\alpha_j$. We claim that the map

$$\alpha \to (f_i)_{1 \le i \le n}$$

is a diffeomorphism. Suppose not. There is a point z_0 and a non zero vector $v \in \mathbb{R}^n$ so that $D_\alpha(f_i)(v) = 0$. But then $\omega(z_0)$ vanishes on the span of $\{0\} \times \mathbb{R}^n \cup \{v, 0\}$, a contradiction to the assumptions that ω has rank n. Let $I_i := \phi_i(\alpha) = f_i$ be the action variables. Then F_j and H are functions of the action variables, $F_j = F_j(I)$ and $H = h(I)$ and ω can be written as

$$\omega = \sum_{j=1}^{n} dI_j \wedge d\theta_j + \tilde{\gamma}_{ij}(I) dI_i \wedge dI_j.$$

Again there exists a one form $\sum_{j=1}^{n} \beta_j(I) dI_j$ so that

$$\sum_{j=1}^{n} d\beta_j(I) dI_j = \sum_{i,j=1}^{n} \tilde{\gamma}_{ij} dI_i \wedge dI_j.$$

We define

$$\theta = \tilde{\theta} - \beta(I)$$

so that

$$\sum_j I_j d\theta_j = \sum_{j=1}^{n} I_j d\tilde{\theta}_j + d\sum_{j=1}^{n} (\beta_j I_j) + \sum_{j=1}^{n} \beta_j dI_j$$

which implies

$$\sum_{j=1}^{n} I_j d\tilde{\theta}_j = \sum_{j=1}^{n} I_j d\theta_j + \sum_{j=1}^{n} \beta_j dI_j.$$

By construction and the chain rule

$$\{I_i, F_j\} = \{I_i(F), F_j\} = 0$$

and similarly $\{I_i, H\} = 0$. □

This proof has a draw-back from Arnold's point of view: It involves solving ODE's and it is possible to be more explicit. We do that in the next section.

2.3.1 Action-angle variables in the case $2n = 2$

We begin with a discussion of symplectomorphisms

$$\Phi : (\mathbb{R}^{2n}, \omega_0) \ni (p, q) \to (P, Q) \in (\mathbb{R}^{2n}, \omega_0).$$

Since

$$\omega_0 = d(pdq) = d(PdQ)$$

we see

$$d(\sum_{j=1}^{n} p_j dq_j - P_j dQ_j) = 0$$

hence there exists a potential S with

$$\sum_{j=1}^{n} p_j dq_j - P_j dQ_j = dS$$

Now suppose that

$$\mathbb{R}^{2n} \ni (p, q) \to (q, Q)$$

is a diffeomorphism and let $S(q, Q)$ be a function so that

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

Define $P_j = -\frac{\partial S}{\partial Q_j}$. Then

$$dS = \sum_{j=1}^{n} \frac{\partial S}{\partial q_j} dq_j + \frac{\partial S}{\partial Q_j} dQ_j = \sum p_j dq_j - P_j dQ_j,$$

and the map

$$(p, q) \to (P, Q)$$

is a local symplectomorphism.

This is used to complement certain sets of coordinate (Q) to P so that the symplectic form is $d(PdQ)$. There are obvious variants if

$$(p, q) \to (q, P) \text{ or } (p, q) \to (p, Q) \text{ or } (p, q) \to (p, P)$$

are (local) diffeomorphisms: One expresses S in terms of the coordinates and complements the coordinates in capital letters so that one obtains a diffeomorphism.

We apply this to the illustrative case of $n = 1$, $(\mathbb{R}^2, \omega_0)$. This is always integrable since we may take $F = H$. We assume that $H_\alpha = \{(p, q) : H(p, q) = \alpha\}$ is connected, compact and nondegenerate. We observe

$$dp \wedge dq = d(pdq).$$

Let $\gamma \in C^1([0, 1]; H_\alpha)$ with $\gamma(0) = \gamma(1)$ be a parametrization of the level set in positive orientation. Let $D(\alpha)$ be the encircled set. We define

$$I(\alpha) = \int_\gamma pdq = Area(D_\alpha)$$

by Green's formula and hence $\frac{dI}{d\alpha} \neq 0$ and thus $\alpha = h(I)$. We define

$$S(I, q) = \int_{z(\alpha)}^{z} pdq$$

for z on the level set with the path of integration is on the level set in positive orientation. Suppose that we can write $p = p(I, q)$. Then we consider S as a function of I and q. Since

$$dS = \frac{\partial S}{\partial q} dq + \frac{\partial S}{\partial I} dI$$

we see that by construction $p = \frac{\partial S}{\partial q}$. We define

$$\theta = \frac{\partial S}{\partial I}$$

so that

$$dS = pdq + \theta dI = pdq - Id\theta + d(\theta I)$$

and

$$\omega_0 = d(Id\theta).$$

The map

$$(p, q) \to (I, \theta)$$

is thus a symplectomorphism. To complete the argument we observe that the period ΔS of S is I so that

$$\frac{d\Delta S}{dI} = 1$$

hence $\theta \in \mathbb{R}/2\pi\mathbb{Z}$.

2.4 The two body problem

The two body problem describes two point masses in $\mathbb{R}^3$ with mass $m_1, m_2 > 0$, which attract each other by gravitation. The kinetic energ is with $p_1, q_1 \in \mathbb{R}^3$

$$\frac{1}{2m_1}|p_1|^2 + \frac{1}{2m_2}|p_2|^2$$

and the potential energy

$$-\frac{g}{|q_1 - q_2|}$$

where $g > 0$ is a gravitational constant. The Hamiltonian is the sum of the two,

$$H = \frac{1}{2m_1}|p_1|^2 + \frac{1}{2m_2}|p_2|^2 - \frac{g}{|q_1 - q_2|}$$

The symplectic form is

$$d\left(\sum_{j=1}^{3} p_1^j dq_1^j + p_2^j dq_2^j\right)$$

and the Hamiltonian equations are

$$\frac{d}{dt}q_{1,2} = \frac{1}{m_{1,2}}p_{1,2}, \quad \frac{d}{dt}p_1 = -g\frac{q_1 - q_2}{|q_1 - q_2|^2}, \quad \frac{d}{dt}p_2 = -g\frac{q_2 - q_1}{|q_2 - q_1|^2}$$

The total momentum is $P = p_1 + p_2$ which Poisson commutes with H,

$$[P^k, H] = -\sum_k \frac{\partial H}{\partial q_1^k} + \frac{\partial H}{\partial q_2^k} = 0$$

The center of mass is

$$Q = \frac{m_1 q_1 + m_2 q_2}{m_1 + m_2 2}$$

and

$$\frac{d}{dt}Q^j = \frac{P^j}{m_1 + m_2}$$

so that

$$Q(t) = q(0) + t\frac{P}{m_1 + m_2}.$$

We define the reduced mass as

$$m = \frac{m_1 m_2}{m_1 + m_2}$$

Let $q = q_2 - q_1$. Then

$$\begin{aligned} \frac{d^2}{dt^2}q &= \frac{d^2}{dt^2}q_2 - \frac{d^2}{dt^2}q_1 \\ &= \frac{d}{dt}\frac{p_2}{m_2} + \frac{d}{dt}\frac{p_1}{m_1} \\ &= -g\left(\frac{1}{m_2} + \frac{1}{m_2}\right)\frac{1}{|q|^3}q \end{aligned}$$

so with $p = m\dot{q}$ we have

$$\frac{d}{dt}q = \frac{1}{m}p, \quad \frac{d}{dt}p = -\frac{g}{|q|^3}q.$$

These are the Hamiltonian equations for the Hamiltonian

$$H = \frac{1}{2m}|p|^2 - \frac{g}{|q|} \tag{2.15}$$

of a particle in a radial potential. The map

$$(p_1, p_2, q_1, q_2) \to (P, p, Q, q)$$

is a symplectomorphism. We have reduced the two body problem to a uniform motion and the problem of a particle in a central force field.

The angular momentum is $L = q \times p$. Its components Poisson commute with H, which we check for the third component.

$$\begin{aligned} \{L^3, H\} &= \sum_{j=1}^{3} \partial_{q_j} L^3 \partial_{p_j} H - \partial_{p_j} L^3 \partial_{q_j} H \\ &= \frac{1}{m}(p_2 p_1 - p_1 p_2 + (-q_2(-\frac{g q_1}{|q|^3}) + q_1(-\frac{g q_2}{|q|^3})) = 0. \end{aligned}$$

However the components of angular momentum do not Poisson commute with another. However H, L^3 and $|L|^2$ Poisson commute. We have $n = 3$ and three Poisson commuting functions.

We proceed in a more geometric fashion. The vector L is conserved, and

$$L = r \times q$$

hence r and q stay in the plane perpendicular to L. Without loss assume that $L = le_3$ and we have reduced the problem to the planar Kepler problem (see [27])

$$\frac{d}{dt}q^1 = p^1, \quad \frac{d}{dt}q^2 = p_2, \frac{d}{dt}p_1 = -\frac{g}{|q|^3}q_1, \frac{d}{dt}p_2 = -\frac{g}{|q|^3}q_2 \tag{2.16}$$

with Hamiltonian

$$H = \frac{1}{2m}(p_1^2 + p_2^2) - \frac{g}{\sqrt{q_1^2 + q_2^2}}.$$

There are two Poisson commuting functions, H and the angular momentum

$$l = q_1 p_2 - q_2 p_1.$$

The differentials are

$$dH = \frac{1}{m}(p_1 dp_1 + p_2 dp_2) + \frac{g}{|q|^3}(q_1 dq_1 + q_2 dq_2)$$

$$dl = q_1 dp_2 + p_2 dq_1 - q_2 dp_1 - p_1 dq_2$$

which are easily seen to be linearly independent whenever $q \neq 0$ and $l \neq 0$. If $H < 0$ and $l \neq 0$ then level sets are bounded: $\frac{g}{|q|} \geq -H$ implies $|q|$ is bounded on the level set. Boundedness of the momentae is harder. We choose polar coordinates

$$q = r \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}$$

and $\omega = \frac{d\theta}{dt}$ so that

$$m\frac{d^2}{dt^2}r = -\frac{g}{r^2} + \frac{1}{m}\Big(\frac{|p|^2}{r} - \frac{(p_1 q_1 + p_2 q_2)^2}{r^3}\Big)$$

and

$$|p|^2|q|^2 - (p_1 q_1 + p_2 q_2)^2 = (q_1 p_2 - q_2 p_1)^2$$

hence

$$m\frac{d^2 r}{dt^2} - mr\omega^2 = -\frac{g}{r^2}. \tag{2.17}$$

We obtain Kepler's second law from the conservation of $l = mr^2\omega$ (the line $[0, q]$ sweeps out an area propotional to time), and

$$\frac{d}{dt} = \frac{l}{mr^2}\frac{d}{d\theta}$$

which allows rewrite (2.17) as

$$\frac{l}{r^2}\frac{d}{d\theta}\left(\frac{l}{mr^2}\frac{dr}{d\theta}\right) - \frac{l^2}{mr^3} = -\frac{g}{r^2}.$$

Let $u = \frac{1}{r}$. Then

$$\frac{d^2u}{d\theta^2} + u = \frac{gm}{l^2}$$

which are one dimensional Hamiltonian equations which can be solved:

$$u = \frac{gm}{l^2}(1 + e\cos(\theta - \theta_0))$$

At $\theta = \theta_0$ we have $\dot{u} = 0$ hence $\dot{r} = 0$. Evaluating the Hamiltonian of (2.17) and hence the Hamitonian of Kepler's problem gives

$$e = \sqrt{1 + \frac{2HL^2}{g^2m}}$$

is the eccentricity and θ_0 the phase offset. Then $e = 0$ is a circle, $e < 1$ an ellipse, $e = 1$ a parabola and $e > 1$ a hyperbola. Hence $H < 0$ gives an ellipse (Kepler's first law) , if $H = 0$ it is a parabola, and if $H > 0$ a hyperbola. To solve the motion we use

$$l = mr^2\frac{d\theta}{dt}$$

and solve the scalar first order ODE via separation of variables,

$$\frac{d\theta}{dt} = \frac{l}{mr^2} = \frac{l}{m}u^2 = \frac{g^2m}{ml^3}(1 + e\cos(\theta - \theta_0))^2.$$

We see that the level sets are bounded if $H < 0$ and $l \neq 0$. We missing yet the action angle variables.

For that we return to the three dimensional problem with central force. We follow [2] and write it in polar coordinates

$$q_1 = r\sin\theta\cos\phi, q_2 = r\cos\theta\cos\phi, x_3 = r\sin\phi$$

We write

$$\alpha = \sum_{j=1}^{3} p_j dq_j = p_r dr + p_\theta d\theta + p_\phi d\phi$$

so that

$$(p, q) \to (r, \theta, \phi, p_r, p_\theta, p_\phi)$$

is a symplectomorphism. Then

$$H = \frac{1}{2}(p_r^2 + \frac{1}{r^2}p_\theta^2 + \frac{1}{r^2 \sin^2\theta}p_\phi^2$$

$$|L|^3 = p_\theta^2 + \frac{1}{\sin^2\theta}p_\phi^2$$

$$L^3 = p_\phi$$

This completes the change of variables.

On the level set

$$p_r = \sqrt{2(H + \frac{g}{r}) - \frac{|L|^2}{r^2}}, p_\theta = \sqrt{|L|^2 - \frac{(L^3)^2}{\sin^2\theta}}, p_\phi = L^3 \tag{2.18}$$

The 1 form α restricted to the level set is obviously closed and has a potential S on the level sets (locally) , which we normalize by chosen points on the level sets where S vanishes depending smoothly on the Poisson commuting functions. The angle variables are as in the one dimensional case

$$\psi_H = \frac{DS}{dH}, \psi_{|L|^2} = \frac{\partial S}{\partial |L|^2}, \psi_{L^3} = \frac{\partial S}{\partial L^3}$$

By the same argument as there with the generating function S this defines a symplectomorphism.

We write the Hamiltonian has a function of the action variables $H, |L|^2, L^3$, $H = h(H, |L|^2, L^3)$ with $h = H$. Hence $\frac{d}{dt}\Psi_H = \frac{\partial h}{\partial H} = 1$ and $\Psi_{|L|^2}$ and Ψ_{L^3} are constant for the flow of H.

2.5 Poisson Geometry

We follow Weinstein [5, 26] in this section.

Definition 2.17. *Let M^N be a manifold. A Poisson structure is a bilinear map*

$$C^\infty(M) \times C^\infty(M) \ni (f, g) \to \{f, g\} \in C^\infty(M)$$

which satisfies

$$\{f,g\} + \{g,f\} = 0 \qquad \textit{skew symmetry}$$

$$\{f,\{g,h\}\} + \{g,\{h,f\}\} + \{h,\{f,g\}\} = 0 \qquad \textit{Jacobi identity}$$

$$\{f,gh\} = \{f,g\}h + \{f,h\}g \qquad \textit{Leibniz rule.}$$

Let M^m and N^n by Poisson manifolds. A map $\phi : M^m \to N^n$ is a Poisson map if

$$\{f \circ \phi, g \circ \phi\}_M = \{f,g\}_N \circ \phi.$$

We define the pull back by this formula which we can use to define a Poisson structure on M from a Poisson structure of N and this formula.

Remarks and definitions.

1. A Casimir is a function f so

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

for all g. Examples are constant functions since

$$\{f,1\} = \{f,1*1\} = 2\{f,1\} = 0$$

2. If $f = gh$ and g and h vanish at a point then $\{f,gh\} = 0$ at this point by the Leibniz rule. Consider a Poisson structure on a subset U of $\mathbb{R}^n$. Let $x_0 \in U$ and

$$f(x) = f(x_0) + \sum \partial_{x_j} f(x_0)(x-x_0)_j + T_2 f(x)$$

Then

$$\{g,f\}(x_0) = \sum_{j=1}^n \partial_{x_j} f(x_0)\{g, x_j - x_j^0\}(x^0)$$

and as a consequence

$$f \to (g \to \{f,g\})$$

is a first order operator, which is a vector field which we call Hamiltonian vector field and denote by ∇_f. We obtain again by the same argument as for symplectic manifolds

$$[\nabla_g, \nabla_f] = \nabla_{\{f,g\}}.$$

3. The Poisson bracket is a sum over products of first order derivatives and in local coordinates resp. in $\mathbb{R}^n$ there exists a skew symmetric matrix Π_{ij} so for some smooth skew symmetric Π_{ij} called the Schouten tensor
$$\{f, g\}(x) = \sum_{i,j} \Pi_{ij}(x)\partial_i f(x)\partial_j g(x) \tag{2.19}$$
where
$$\Pi_{ij} = \{x_i, x_j\}.$$
The Hamiltonian vector field is
$$\nabla_f = \Pi df.$$
It generates a flow.

4. If Π is skew and constant then the Jacobi identity is automatically satisfied. After a linear coordinate transform
$$\Pi = \begin{pmatrix} 0 & -1_k & 0 \\ 1_k & 0 & 0 \\ 0 & 0 & 0_l \end{pmatrix}.$$
Casimirs are functions of $(c_j)_{1\le j\le l}$ with coordinates $(p, q, c) \in \mathbb{R}^k \times \mathbb{R}^k \times \mathbb{R}^l$.

2.5.1 The Lie-Poisson (or Kirrilov-Kostant) bracket on the dual of a Lie algebra

Let $\mathfrak{g}$ be a finite dimensional Lie algebra and let $\mathfrak{g}^*$ be its dual. If $f \in C^1(\mathfrak{g}^*)$ then $df(\mu) \in \mathfrak{g}^{**} = \mathfrak{g}$ and we define the Lie-Poisson bracket using the bracket of the Lie algebra
$$\{f, g\}(\mu) = \mu([df(\mu), dg(\mu)]).$$
It is clearly bilinear, skew symmetric and satisfies the Leibniz rule.

Lemma 2.18. *The Lie-Poisson bracket satisfies the Jacobi identity.*

Proof. We define the jacobiator (or Schouten-Nijemhuis bracket) on $C^\infty(\mathfrak{g}^*)$ as
$$J(f, g, h) = \{\{f, g\}, h\} + \{\{g, h\}, f\} + \{\{h, f\}, g\}.$$
It is skew symmetric and a derivation in each argument:
$$\begin{aligned} J(f, g, hl) &= \{\{f, g\}, hl\} + \{\{g, h\}l + \{g, l\}h, f\} + \{\{h, f\}l + \{l, f\}h, g\} \\ &= (\{\{f, g\}, h\} + \{\{g, h\}, f\} + \{\{h, f\}, g\})l \\ &\quad + (\{\{f, g\}, l\} + \{\{g, l\}, f\} + \{\{l, f\}, g\})h \\ &\quad + \{g, h\}\{l, f\} + \{g, l\}\{h, f\} + \{h, f\}\{l, g\} + \{l.f\}\{h, g\} \end{aligned}$$

Thus

$$J(f,g,h)(\mu_0) = J((\mu \to df(\mu_0)(\mu), (\mu \to dg(\mu_0)(\mu)), (\mu \to dh(\mu_0)(\mu)))$$

and it suffices to check the Jacobi identity for linear functions. Suppose

$$f(\mu) = \mu(A), g(\mu) = \mu(B), h(\mu) = \mu(C)$$

are linear functions, $A, B, C \in \mathfrak{g}$. Then

$$\{f,g\}(\mu) = \mu([A,B])$$

and

$$\{f,\{g,h\}\}+\{g,\{h,f\}\}+\{h,\{f,g\}\} = \mu([A,[B,C]]+[B,[C,A]]+[C,[A,B]]) = 0$$

and the Jacobi identity for linear functions for the Lie-Poisson bracket is equivalent to the Jacobi identity of the Lie algebra. This implies the claim. □

Let $A \in \mathfrak{g}$. It defines a linear map on $\mathfrak{g}$ called the adjoint represenation of the Lie agebra

$$ad_X Y = [X,Y],$$

for which Lie bracket and commutator are compatible:

$$\begin{aligned} ad_{[X,Y],Z} &= [[X,Y],Z] \\ &= -[[Y,Z],X] - [[Z,X],Y] \\ &= ad_X[Y,Z] - ad_Y[Z,X] \\ &= ad_X ad_Y Z - ad_Y ad_X Z \end{aligned}$$

The coadjoint representation on $\mathfrak{g}^*$ is defined by

$$ad_X^* \mu(Y) = -\mu(ad_X Y).$$

We compute

$$\nabla_H f(\mu) = \mu([dH, df]) = \mu(ad_{dH}(df)) = -ad_{dH}^*(m)((df)(m))$$

and

Lemma 2.19.

$$\nabla_H = -\operatorname{ad}_{dH}^* . \tag{2.20}$$

2.5.2 The splitting theorem of Weinstein

At a point $x_0 \in U \subset \mathbb{R}^N$ the Hamiltonian vector fields span a subspace of the tangent space $\mathbb{R}^n$. Its dimension is the rank of the skew symmetric $\Pi(x_0)$, which is even.

Theorem 2.20 (Splitting theorem of Weinstein). *Let M^N be a Poisson manifold and $x_0 \in M$. There exist coordinates $((p_j)_{1\le j\le n}, (q_j)_{1\le j\le n}, (y_j)_{1\le j\le M})$ so that*

$$\Pi = \begin{pmatrix} 0 & -1 & 0 & 0 & \dots & 0 \\ 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & -1 & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & \tilde\Pi(y) \end{pmatrix}$$

where $\tilde\Pi$ defines a local Poisson structure on $\mathbb{R}^{N-2n}$ which vanishes for $y = 0$.

There is an important consequence: $N = \mathbb{R}^{2n} \times \{0_{N-2n}\}$ is a Poisson manifold with a Poisson structure of maximal rank, since

$$\{f, g\}|_N = \sum_{j=1}^n \partial_{q_j} f|_N \partial_{p_j} g|_N - \partial_{p_j} f|_N \partial_{q_j} g|_N.$$

As a consequence on the submanifold $\{y = 0\}$ there is a natural Poisson structure of full rank which is the Poisson structure of the canonical symplectic form on this manifold. In particular Darboux's theorem is a special case.

Through every point there is symplectic submanifold defined by the Poisson structure so that the Hamiltonian vector fields are tangent. As an immediate consequence the Hamiltonian flows define Poisson mappings, since we know that in the symplectic case.

Proof. We assume that there is Hamiltonian vector field ∇_{p_1} which does not vanish at x_0 - otherwise we take $n = 0$. So we assume that there is another function q so that $\{p, q\}(x_0) = 1$. We choose coordinates $(p_1, q_1, (\tilde y_j)_{1\le j\le N-2})$ so that

$$\{p_1, q_1\} = \nabla_{p_1} q_1 = 1.$$

This is a linear ordinary differential equation which has a local solution.

Then ∇_{p_1} and ∇_{q_1} are linearly independent and

$$[\nabla_{p_1}, \nabla_{q_1}] = \nabla_1 = 0.$$

We apply the Frobenius theorem wand choose coordinates $(p_1, q_1, (y_j))$ so that the all components besides the first two of the vector fields ∇_{p_1} and ∇_{q_1} vanish. Then the dy_j are linearly independent, $\nabla_{p_1}(y_j)) = \nabla_{q_1}(y_j) = 0$ since they are constant on the leafs. Then by the Jacobi identity

$$\{\{y_i, y_j\}, p_1\} = \{\{y_i, y_j\}, q_1\} = 0$$

and the Poisson brackets $\{y_i, y_j\}$ are constant on the leafs, and hence a function of the ys. In this coordinates M is a locally a product of $\mathbb{R}^2$ with the canonical symplectic structure, and an $N-2$ dimensional Poisson manifold. We obtain the claim by induction of j. □

3 Integrable ODEs

3.1 Euler's equation on Lie groups: The spinning tops of Euler, Lagrange and Manakov and the Korteweg-de Vries equation

Most of the material of this subsection is contained in Khesin and Wendt [15], see also Marsden and Ratiu [18]. Let $SO(n)$ be the set of orthogonal $n \times n$ matrices of determinant 1. It is a smooth manifold in the space of $n \times n$ matrices since

$$O \in SO(n) \quad \Longleftrightarrow \det O = 1, O^T O = 1$$

is a nondegenerate level set. The tangent space at the identity is the set $so(n)$ of traceless skew symmetric matrices. The matrix product is smooth since it is bilinear in the set of matrices. Inversion is smooth by the implicit function theorem (or linearity of taking transposeds).

Define

$$\exp(tX) = \sum_{k=0}^{\infty} \frac{(tX)^k}{k!} \in GL(n)$$

Then

$$\frac{d}{dt} \exp(tX) = X \exp(tX) = \exp(tX) X$$

$$\exp((s+t)X) = \exp(sX) \exp(tX)$$

$$\exp(0X) = 1$$

If $X \in so(n)$ then $\exp(tX) \in SO(n)$ which is rotation in n dimensions.

There is a canonical representation of $SO(n)$ on $so(n)$, the adjoint representation

$$\operatorname{Ad}_O X = OXO^{-1} = OXO^T.$$

We can differentiate

$$\frac{d}{dt}\exp(tX)Y\exp(-tX)|_{t=0} = XY - YX = ad_X Y$$

and obtain the adjoint representation of the Lie algebra on itself. The coadjoint representation is defined by

$$\operatorname{Ad}^*_O m(X) := m(\operatorname{Ad}_{O^{-1}}) = m(O^{-1}XO)$$

and

$$\operatorname{ad}^*_Y m(X) := -m(\operatorname{ad}_Y X).$$

The left resp right multiplication defines a map

$$O : T_e G \to T_O G$$

$$X \to OX \quad \text{resp. } X \to XO.$$

The inner product

$$so(n) \times so(n) \ni \langle X, Y \rangle = -\frac{1}{2}\operatorname{tr} XY = \frac{1}{2}\operatorname{tr} X^T Y$$

is invariant under the adjoint representation:

$$\langle \operatorname{Ad}_O X, \operatorname{Ad}_O Y \rangle = \frac{1}{2}\operatorname{tr}((OXO^T)^T OYO) = \frac{1}{2}\operatorname{tr} X^T Y = \langle X, Y \rangle.$$

On the level of the Lie algebra

$$\begin{aligned} -\langle \operatorname{ad}_Y X, Z \rangle &= \langle \operatorname{ad}_X Y, Z \rangle \\ &= \operatorname{tr}[X, Y]^T Z \\ &= \operatorname{tr}[Y, X] Z \\ &= \operatorname{tr}(YXZ - XYZ) \\ &= \operatorname{tr}(XZY - XYZ) \\ &= \langle X, \operatorname{ad}_Y Z \rangle. \end{aligned}$$

The Killing form $B(X, Y)$ is the trace of $\operatorname{ad}_X \operatorname{ad}_Y$ in the adjoint representation. One can calculate for $SO(n)$

$$B(X, Y) = -(n-2)\langle X, Y \rangle$$

Theorem 3.1. *The leaves of the Lie-Poisson bracket are the orbits under the co adjoint representation. A function is a Casimir if and only if it is invariant under the coadjoint representation.*

Proof. The tangent space of any leaf is spanned by the Hamiltonian vector fields. If $v \in \mathfrak{g}$ and $m \in \mathfrak{g}^*$ then there is a Hamiltonian function which has this derivative at m. Suppose that f is smooth and invariant under the coadjoint representation Ad^*. Using the exponential map and differentiating we see that

$$\mathrm{ad}_v\, df = 0$$

for every v hence $\{f, g\} = 0$ for all g by Lemma 2.19. Thus f is a Casimir. All arguments are reversible.

Similarly, since the tangent space of the leaf consists of the evaluation of Hamiltonian vector fields this argument shows that $\mathrm{ad}_v^* m$ always tangent. Using the matrix exponential and differentiating with respect to t shows that the leaf is the orbit.

□

Let K be a rigid body with density $\rho \geq 0$ and mass

$$\int_\rho dx > 0.$$

We write

$$X(t) = O^{-1}(t)\dot{O}(t) \in \mathfrak{so}(n).$$

Then

$$\begin{aligned} \frac{1}{2}\int_K \left|\frac{d}{dt}O(t)x\right|^2 \rho dx &= \frac{1}{2}\int |O(t)X(t)x|^2 \rho(x) dx \\ &= \frac{1}{2}\int |X(t)x|^2 dy \end{aligned}$$

We define the inertia matrix resp bilinear form

$$(X, Y) := \frac{1}{2}\int_K Xx \cdot Yx\rho dx$$

which is a nondegenerate inner product on $so(n)$. We define the Hamiltonian

$$H(O, \dot{O}) := \frac{1}{2}(O^{-1}(t)\dot{O}(t), O^{-1}(t)\dot{O}(t)) = \frac{1}{2}(X(t), X(t)).$$

The principle of the least action says that a path is a critical point of the action functional

$$\int_0^T \frac{1}{2}(O^{-1}\dot{O}(t), O^{-1}\dot{O}(t))dt.$$

The inertia defines an invertible linear map $A : so \to so^*$.

Theorem 3.2. *Let $O(t)$ be a critical point of the action functional in the sense above. Let*

$$m(t) = A\dot{O}(t).$$

Then m satisfies the Hamiltonian equation (with respect to the Lie-Poisson bracket)

$$\frac{d}{dt}m(t) = -ad^*_{A^{-1}m(t)}m(t)$$

with Hamiltonian

$$H = \frac{1}{2}(A^{-1}m, A^{-1}m).$$

Suppose that $m(t)$ is a solution to the Hamiltonian equations. Let $h(t) = Am(t)$ and let $O(t)$ satisfy

$$\dot{O}(t) = O(t)h(t), O(0) \in SO(n).$$

Then O is a critical point of the action functional.

Proof. We consider a smooth function $O(s;t) \in SU(n)$ and assume that $t \to O(0,t)$ is a critical point of the action functional and $O(s,0) = O(0,0)$, $O(s,T) = O(0,T)$. We differential the action with respect to s. Then

$$\begin{aligned} 0 &= \frac{d}{ds}\int \frac{1}{2}(O^{-1}(s,t)\dot{O}(s,t), O^{-1}(s,t)\dot{O}(s,t)dt \\ &= \int (\frac{\partial}{\partial s}O^{-1}(s,t)\dot{O}(s,t), O^{-1}(s,t))dt \end{aligned}$$

Let $X(t) = O^{-1}(0,t)\dot{O}(0,t)$.

$$\begin{aligned} \frac{\partial}{\partial s}O^{-1}\dot{O} &= O^{-1}\frac{\partial}{\partial t}\frac{\partial}{\partial s}O - O^{-1}\partial_s O O^{-1} \\ &= \frac{d}{dt}(O^{-1}\partial_s O) + [X, O^{-1}\partial_s O] \end{aligned}$$

Thus by the definition of A

$$\begin{aligned} 0 &= \int -(g^{-1}\partial_s g, \dot{X}) + ([v, O^{-1}\partial_s O], X)dt\Big|_{s=0} \\ &= \int -\frac{d}{dt}(Av)(g^{-1}\partial_s g) - ad^*_X(Av)(g^{-1}\partial_s g)dt\Big|_{s=0} \end{aligned}$$

hence, with $m = AX$

$$\frac{d}{dt}m = -ad^*_{A^{-1}m}m.$$

□

In the considerations above we can replace $SO(n)$ with any Lie group, with the exeption of the form $\langle , \rangle$. The Killing form is nondegenerate for any semisimple Lie group, and negative definite for semisimple compact Lie groups. Semisimple compact Lie groups can be realized as matrix groups.

3.1.1 Euler's spinning top

Here we specialize our consideration to the case $n = 3$. The group $SU(3)$ is not simply connected. Its universal covering space is $SU(2)$, which is a the group of the quanternonian multiplicative quaternonian multiplication on the three dimensional sphere. The covering map $SU(2) \to SO(3)$ is given by the adjoint representation.

Elements in $\mathfrak{so}(3)$ can be written as

$$\begin{pmatrix} 0 & -x_1 & x_3 \\ x_1 & 0 & -x_2 \\ -x_3 & x_2 & 0 \end{pmatrix}$$

The commutator is given by the negative of the cross product

$$\begin{aligned} &\left[\begin{pmatrix} 0 & -x_1 & x_3 \\ x_1 & 0 & -x_2 \\ -x_3 & x_2 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -y_1 & y_3 \\ y_1 & 0 & -y_2 \\ -y_3 & y_2 & 0 \end{pmatrix} \right] \\ &\qquad = \begin{pmatrix} 0 & -(x_2 y_3 - x_3 y_2) & x_1 y_2 - x_2 y_1 \\ -(x_2 y_3 - x_3 y_2) & 0 & x_3 y_1 - x_1 y_3 \\ -(x_2 y_3 - x_3 x_2) & x_3 y_1 - x_1 y_3 & 0 \end{pmatrix}. \end{aligned}$$

The adjoint representation becomes $\mathrm{ad}_X \vec{y} = x \times y$ hence $Ad_O \vec{x} = O\vec{x}$. An easy calculation shows that

$$\frac{1}{2} \operatorname{tr} \begin{pmatrix} 0 & -x_1 & x_3 \\ x_1 & 0 & -x_2 \\ -x_3 & x_2 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -y_1 & y_3 \\ y_1 & 0 & -y_2 \\ -y_3 & y_2 & 0 \end{pmatrix} = x\dot{y}$$

and we identify $so^*(3)$ and $so(3)$ so that

$$\mathrm{Ad}_O^* \vec{m} = O^{-1} \vec{m}.$$

The group orbits of $SO(n)$ are the spheres and the origin. The Hamiltonian equations become

$$\frac{d}{dt} m = m \times A^{-1} m.$$

They define two dimensional Hamiltonian equations on the orbits. Two dimensional Hamiltonian equations are integrable.

Rotating the coordinates we may assume that A is diagonal with positive entries I_j. The Hamiltonian is

$$\frac{1}{2}m^T A^{-1} m = \frac{1}{2}(I_1^{-1}m_1^2 + I_2^{-1}m_2^2 + I_3^{-1}m_3^2).$$

The trajectories are given generically by the intersection of the sphere with the ellipsoid,

$$|m|^2 = R^2, \quad I_1^{-1}m_1^2 + I_2^{-1}m_2^2 + I_3^{-1}m_3^2 = H. \tag{3.1}$$

Then

$$\dot{m}_1 = (m \times A^{-1}m)_1 = (I_3^{-1} - I_2^{-1})m_2 m_3$$

and for some coefficients (solving the two equations (3.1) for m_2 and m_3 in terms of m_1, H and R)

$$\dot{m}_1 = \sqrt{\alpha + \beta m_1^2 + \gamma m_1^4}.$$

which can be integrated by separation of variables.

Alternatively we may count Poisson commuting functions

$$H, |L|^2, L^3$$

on a six dimensional space. However, we would have to rewrite the equations in Hamiltonian form on the cotangent bundle of $SO(3)$.

3.2 Lagrange's spinning top

This material is from [2].

We consider a spinning top in gravity in $\mathbb{R}^3$ which is fixed at the origin which is not the center of mass. The center of mass is

$$q = \frac{1}{M}\int x\rho dx, \quad M = \int \rho dx.$$

The kinetic energy is again

$$\frac{1}{2}(O^{-1}\frac{d}{dt}O, O^{-1}\frac{d}{dt}O)$$

and the potential energy is

$$q_3 = (O(t)q(0))_3.$$

We assume that two of the eigen values of the intertia A are the same, $I_1 = I_2 = \lambda$ and the third is $\mu \neq \nu$. Then, with L the angular momentum (again by the least action principle)

$$\frac{d}{dt}L = -(\frac{d}{dt}O) \times L - e_3 \times q, \qquad \frac{d}{dt}q = -(\frac{d}{dt}O) \times q.$$

where

$$O^{-1}L = IO^{-1}\dot{O}$$

The Hamiltonian is

$$H = \frac{1}{2}L \cdot I^{-1}L + q_3$$

and the Poisson bracket (basically giving Π)

$$\{L_i, L_{i+1}\} = L_{i+3}, \quad \{L_i, q_{i+1}\} = q_{i+2}, \quad \{q_i, q_{i+1}\} = 0$$

with indices mod 3. There are two Casimirs, $|q|^2$ and $L \cdot q$

$$L \cdot q = O^T L \cdot OL$$

for every orthogonal matrix and the dynamics is reduced to a four dimensional sub manifold.

We write

$$q = r \begin{pmatrix} \sin\theta \sin\psi \\ \sin\theta\cos\psi \\ \cos\theta \end{pmatrix}$$

$$\dot{O} = \begin{pmatrix} \dot\phi \sin\theta\sin\psi + \dot\theta\cos\psi \\ \dot\phi\sin\theta\cos\psi - \dot\theta\sin\psi \\ \dot\phi\cos\theta + \dot\psi \end{pmatrix}$$

The Hamiltonian is

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

The Casimirs are r and

$$q \cdot L = (Oq \cdot O^T L) = r\Big[(I_1 \sin^2\theta + I_3\cos\theta)\dot\phi + I_3\cos\theta\dot\psi\Big] =: rL_z.$$

The third component of the angular momentum Poisson commutes with H since H is independent of ψ,

$$L_3 = I_3(\dot\phi\cos\theta + \dot\psi)$$

We use these identities to eliminate $\dot\phi$ and $\dot\psi$ from the Hamiltonian,

$$H = \frac{1}{2}I_1\dot\theta^2 + \frac{1}{2I_1}\frac{(L_z - L_3\cos\theta)^2}{\sin^2\theta} - r\cos\theta + \frac{1}{2}\frac{L_3^2}{I_3}.$$

The Hamiltonian H is constant and we consider this as on ODE for θ.

3.2.1 The case of $\mathfrak{so}(n)$

The quadratic form on $\mathfrak{so}(n)$

$$\langle X, Y\rangle = -\frac{1}{2}\,\mathrm{tr}(XY)$$

defines the linear invertible map $J^{-1} : \mathfrak{so}(n) \to \mathfrak{so}^*(n)$. We define the gradient of $f \in C^1(\mathfrak{so}(n))$

$$\nabla f(X) = J^{-1} df(X).$$

We pull back the Lie-Poisson structure to $\mathfrak{so}(n)$. Let $f(X) = F(J^{-1}x)$, $g(x) = G(J^{-1}x)$. Then

$$\begin{aligned}\{f, g\}^{\mathfrak{so}(n)}(X) = \{F, G\}^{\mathfrak{so}^*(n)}(J^{-1}X) &= J^{-1}X([dF(J^{-1}X), dG(J^{-1}X)]) \\ &= \langle X, [Jdf(X), Jdg(X)]\rangle \\ &= \langle X, [\nabla f(X), \nabla g(X)]\rangle \\ &= -\langle [\nabla f(X), X], \nabla g(X)\rangle\end{aligned}$$

hence

$$\nabla_f g(X) = -\langle [\nabla f(X), X], \nabla g(X)\rangle$$

and

$$\nabla_f(X) = -[\nabla f(X), X]. \tag{3.2}$$

The Hamiltonian of Euler's spinning top is the kinetic energy

$$H(X) = \frac{1}{2} J^{-1}X(A^{-1}J^{-1}X)$$

and the equation for the spinning top on $\mathfrak{so}(n)$ become

$$\frac{d}{dt}X = -[\nabla H(X), X]. \tag{3.3}$$

Let

$$I_{ij} = \frac{1}{2}\int_K x^i x^j \rho dx$$

The inertia form is

$$\sum_{i,j,k=1}^{n} X^{ik} I_{kl} X^{jl} = AX(Y).$$

with the summation convention and the Hamiltonian is

$$H(X) = \frac{1}{2} AX(Y).$$

After an orthogonal change of coordinates (using Schur' s decomposition I is a diagonal matrix with diagonal entries I_j, $1 \le j \le n$. We denote the invertible map $\tilde{I}X := \frac{1}{2}(IX + XI)$.

Lemma 3.3. *The following formular holds*

$$\sum_{i,k,l=1}^{n} Y^{ik} I_{kl} X^{il} = -\frac{1}{2} \operatorname{tr} Y(XI + IX)$$

Proof. Let e_{ij} be the matrix with an 1 in the ith row and jth column. A basis of $\mathfrak{so}(n)$ is given by $e_{ij} - e_{ij}$, $i < j$. Let $Y = e_{ij} - e_{ji}$. On the left hand side we obtain

$$\sum_{l=1}^{n} (I_{jl} X^{il} - I_{il} X^{jl})$$

and on the right hand side

$$-\frac{1}{2} \sum_{l=1}^{n} (e_{ij} - e_{ji})(X^{il} I_{lj} + X^{jl} I_{li}) = I_{lj} X^{il} - I_{li} X^{jl}.$$

□

Theorem 3.4. *The equations of the spinning top are*

$$\frac{d}{dt} X = [X, (\tilde{I})^{-1} X] \tag{3.4}$$

and the Hamiltonian is

$$H(X) = -\frac{1}{4} \operatorname{tr}(X \tilde{I}^{-1} X).$$

Let $\Omega = \tilde{I}^{-1} X$. Then we can write the equation as

$$\dot{X} = [X, \Omega] = X\Omega - \Omega X$$

resp. as

$$I\dot{\Omega} = [I\Omega, \Omega]. \tag{3.5}$$

3.3 Euler's top in any dimension: Manakov's idea

We reduced the problem of the n dimensional spinning top to the Hamiltonian equation

$$\frac{d}{dt} m(t) = -\operatorname{ad}^*_{A^{-1} m(t)} m(t)$$

on $\mathfrak{so}^*(n)$. We have seen that the solution stay on leaves of the coadjoint action. The inner product allows to identify $\mathfrak{so}(n)$ and $\mathfrak{so}^*(n)$ as vector spaces. This allowed write the ODE as a Hamiltonian differential equation on $\mathfrak{so}(n)$

$$\dot{X} = [X, \tilde{I}^{-1}X]$$

with the Hamiltonian

$$H(X) = -\frac{1}{4}\operatorname{tr} X\tilde{I}^{-1}X.$$

We first discuss the orbits of the coadjoint representation. Their dimension is even. The rank of any skew symmetric matrix X is even and every even number occurs as rank. The rank is constant under the coadjoint action. Also the spectrum is conserved.

Lemma 3.5. *Let X be a skew adjoint $n \times n$ matrix of rank $2k$. Then there exists an orthogonal matrix so that*

$$O^{-1}XO = \begin{pmatrix} \begin{pmatrix} 0 & \lambda_1 \\ -\lambda_1 & 0 \end{pmatrix} & 0 & \dots & 0 & \dots \\ 0 & \begin{pmatrix} 0 & \lambda_2 \\ -\lambda_2 & 0 \end{pmatrix} & \dots & 0 & \dots \\ \vdots & \vdots & \ddots & \vdots & \ddots \\ 0 & 0 & \dots & \begin{pmatrix} 0 & \lambda_k \\ -\lambda_k & 0 \end{pmatrix} & \dots \\ 0 & 0 & \dots & 0 & O_{\mathbb{R}^{(n-2k)\times(n-2k)}} \end{pmatrix}$$

where the λ_j are nonzero real numbers.

Proof. We take an orthonormal basis in the range and complement it. This reduces the problem to $n = 2k$. The matrix X^2 is negativ definit and has a Schur decomposition. Retricting to eigenspaces we reduce the problem to $X^2 = -\lambda^2 1$, $\lambda > 0$. The eigenvalues are $\pm i\lambda$. Let $E_{i\lambda}$ be the eigen space to the eigen value $i\lambda$ in $\mathbb{C}^n$. We write $x \in E_{i\lambda}$ as $x = x_r + ix_i$ and compute

$$X(x_r - ix_i) = \overline{Xx} = -i\lambda(x_r - ix_i)$$

and

$$X(x_r) = \frac{1}{2}(X(x_r + ix_i) + (x_r - ix_i)) = \frac{1}{2}(i\lambda(x_r + ix_i) - i\lambda(x_r - ix_i)) = \lambda(ix_i)$$

$$X(x_i i) = -\lambda x_r.$$

Taking the real resp. complex part we obtain the decomposition. The matrix X is skewsymmetric hence normal in $\mathbb{C}^n$ and hence has an orthogonal eigen space decomposition. $\square$

3.3.1 The group GL(n) and its Lie algebra

We denote the group of matrices with determinant 1 by GL(n) and its Lie algebra by $\mathfrak{gl}(n)$ which consists of all matrices of trace 0. Then

$$\begin{aligned} \langle A, B\rangle : &= -\frac{1}{2}\operatorname{tr} AB \\ &= -\frac{1}{4}(\operatorname{tr}(A + A^T)B + \operatorname{tr}(A - A^T)B) \\ &= -\frac{1}{4}(\operatorname{tr}(A + A^T)(B + B^T)) - \frac{1}{4}((A - A^T)(B - B^T)) \end{aligned} \tag{3.6}$$

since if A is symmetric and B is skew symmetric then

$$\operatorname{tr} AB = \operatorname{tr} B^T A = -\operatorname{tr} BA = -\operatorname{tr} AB$$

hence $\operatorname{tr} AB = 0$. The adjoint representation on GL(n) is

$$\operatorname{Ad}_G A = GAG^{-1}.$$

The quadratic form is nondegenerate: It is positive definit on skew adjoint matrices and negative definit on symmetric matrices. It defines a unique map $\mathfrak{gl}(n) \to \mathfrak{gl}^*(n)$. Let

$$\tilde{I} : A \to \frac{1}{2}(IA + AI)$$

which is again diagonal and invertible. The Hamiltonian

$$H(A) = \frac{1}{4}\operatorname{tr}(A\tilde{I}^{-1}A)$$

leads again to the Hamiltonian equations

$$\dot{A} = [A, \tilde{I}^{-1}A]$$

on $SL(n)$.

3.3.2 The Lax-Pair

Equality (3.4) can be understood is a Lax-pair. We can write it even as

$$\frac{d}{dt}(X + \lambda I^2) = [X + \lambda I^2, (\tilde{I})^{-1}X + \lambda I]. \tag{3.7}$$

for $\lambda \in C$:

$$[X + \lambda I^2, (\tilde{I})^{-1}X + \lambda I] = [X, (\tilde{I})^{-1}X] + \lambda([X, I] + [I^2, \tilde{I}^{-1}X].$$

We continue

$$\begin{aligned}[X,I]+[I^2,\tilde{I}^{-1}X] &= \frac{1}{2}[I\Omega+\Omega I,I]+[I^2,\Omega]\\ &= I\Omega I+\Omega I^2-I^2\Omega-I\Omega I+I^2\Omega-\Omega I^2\\ &= 0.\end{aligned}$$

Equation (3.7) is remarkable, but we want to rewrite it.

Theorem 3.6. *Then the Lax pair becomes*

$$\frac{d}{dt}(X+\lambda I^2)=[X+\lambda I^2,\tilde{I}^{-1}(X+\frac{1}{2}\lambda I^2)]. \tag{3.8}$$

It is a Hamiltonian equation with Hamiltonian

$$H(X)=-\frac{1}{4}\operatorname{tr}(X+\frac{1}{2}\lambda I^2)\tilde{I}^{-1}(X+\frac{1}{2}\lambda I^2)$$

and the shifted Poisson structure

$$\{f,g\}=\langle X+\lambda I^2,[\nabla f,\nabla g]\rangle$$

in $\mathrm{GL}(n)$.

Proof. The left hand side is $\frac{d}{dt}X$. The identity matrix commutes with everything, and the terms on the right hand side of (3.8) differ from the previous term by the addition of constants. □

- It is the basis of the proof of integrability.
- It is a blue print for the proof of integrability of many other problems.

Clearly

$$\operatorname{tr}\frac{1}{2k}(X+\lambda I^2)^k$$

is constant on the orbits of the translated coadjoint action. Hence it is a Casimir for the translated Poisson structure and conserved under the Hamiltonian flow.

We expand the trace in λ. The coefficient of λ^j vanishes if $k-j$ is odd. $\operatorname{tr} I^{2k}$ is constant. Moreover $\operatorname{tr} X^k$ is constant on the orbit. let $d_{k,j}$ be the coefficient of λ^j. We obtain $[(k-1)/2]$ conserved quantities, if we ignore the trivial ones. In total there are up to

$$\sum_{k=2}^n[(k-1)/2]=\frac{1}{2}(\frac{n(n-1)}{2}-[n/2])$$

nontrivial conserved quantities. The dimension of the generic orbit (of rank n resp $n-1$, with $[n/2]$ pairwise different eigenvalues of $-X^2$) is

$$\frac{n(n-1)}{2} - [n/2]$$

since the dimension of $\mathfrak{so}(n)$ is $\frac{n(n-1)}{2}$ and there are $[n/2]$ conditions. so this is the correct number of conserved quantities.

Theorem 3.7. *1. the d_{jk} Poisson commute with the Hamiltonian.*

2. the d_{jk} Poisson commute

3. The quantities $(dd_{jk})_{jk}$ have maximal rank if rank $X = 2[n/2]$ and the eigenvalues are all different.

The first claim is a consequence of our considerations. We do not prove the last claim.

3.4 BiHamiltonian structure and integrability

We consider the following setting: Let $\{,\}_0$ and $\{,\}_1$ be two compatible different Poisson structures on $\mathbb{R}^n$. Compatible means that $\{,\}_t = \{,\}_0 + t\{,\}_1$ satisfies the Jacobi identity for all t. Suppose that Q_t is a Casimir for all $t \in [0,1]$, which depends analytically on t uniformly on compact sets. Let

$$Q_t(x) = \sum_{j=0}^{N} H_n(x)t^n + O(t^{N+1}).$$

Then the H_n Poisson commute with respect to both Poisson srtuctures.

We expand

$$0 = \{H_n, Q_t\}_t = \sum_{m=0}^{N} (t^m\{H_n, H_m\}_0 + t^{m+1}\{H_n, H_m\}_1) + O(t^{N+1})$$

hence

$$\{H_n, H_0\}_0 = 0, \quad \{H_n, H_{m+1}\}_0 + \{H_n, H_m\}_1 = 0.$$

Since by skew symmetry

$$\{H_n, H_n\}_0 = \{H_n, H_n\}_1 = 0$$

we obtain the statement by induction. Now suppose that

$$P_t = \sum_{j=0}^{N} L_j(x)t^j$$

is a second Casimir. Then

$$0 = \{L_n, H_0\}_0 = \{L_0, H_m\}$$

and

$$0 = \{L_n, H_{m+1}\}_0 + \{L_n, H_m\}_1 = \{L_{n+1}, H_m\}_0.$$

Thus

$$0 = \{L_{n+m-1}, H_0\}_0 = \{L_n, H_{m+1}\}_0 = \{L_n, H_m\}_1.$$

We apply this consideration to

$$P = \mathrm{tr}((X + t\Lambda)^k), Q = \mathrm{tr}((X + t\Lambda)^l).$$

By the same type of argument

$$\nabla^0_{H_n} = -\nabla^1_{H_{n+1}}.$$

3.5 The Virasoro-Bott group

Here I follow Khesin and Wendt [15]. See also [9]. The Virasoro-Bott group $\mathfrak{VIR}$ and its Lie algebra $\mathfrak{vir}$ are of central importance in conformal field theory, which is only tangentially related to our interests here. We will see that the Korteweg-de Vries equation and other integrable equations arise as geodesic equations for a metric on the $\mathfrak{vir}^*$. We obtain an infinite number of Poisson commuting Hamiltonians in the same fashion as for the general top.

The main step is the computation of the coadjoint action of $\mathfrak{vir}$ but also the coadjoint action of $\mathfrak{VIR}$ is of interest.

3.5.1 The diffeomorphism group $\mathfrak{D}\mathrm{IFF}(\mathbb{S}^1)$

Let $\mathfrak{D}\mathrm{IFF}(M)$ be the group of diffeomormphisms on a manifold M which we assume to be an open subset of $\mathbb{R}^n$ since we want to argue with local coordinates. The diffeomorphism group $\mathfrak{D}\mathrm{IFF}(M)$ acts on smooth functions by

$$(\phi, f) \to f \circ \phi^{-1}$$

so that

$$(\phi \circ \psi, f) \to f \circ (\phi \circ \psi)^{-1}$$

If $\psi(t)$ is smooth family of diffeomorphisms with $\phi(0, x) = x$ then

$$\frac{d}{dt}\psi(t))|_{t=0} = X$$

where X is smooth vector field. We obtain the adjoint representation by

$$\frac{d}{dt}(\phi \circ \psi(t) \circ \phi^{-1})|_{t=0} = (D\phi \circ X \circ \phi^{-1})$$

and

$$\mathrm{Ad}_\phi(X) = D\Phi(\phi^{-1}(x))X(\phi^{-1}(x)).$$

Differentiating once more we obtain as before

$$\mathrm{ad}_Y(X) = -[X, Y].$$

If M is a compact smooth manifold there is an exponential map defined by

$$\dot{x} = X(x), x(0) = 0.$$

In general it is neither locally surjectiv nor injectiv.

We write smooth vector fields resp. elements of $\mathfrak{D}\mathrm{iff}(\mathbb{S}^1)$ as $v\partial$. The adjoint representation is

$$\mathrm{Ad}_\phi v\partial = \phi'(\phi^{-1}(x))v(\phi^{-1}(x))\partial$$

$$\mathrm{ad}_u v = -uv' + v'u.$$

We define the smooth dual as quadratic differentials

$$\{udx^2 : u \in C^{(}\mathbb{S}^1)\}$$

where dx^2 is a symbol suggesting the coadjoint action

$$\mathrm{Ad}^*_{\phi^{-1}} udx^2 = u \circ \phi(\phi')^2 dx^2.$$

We define the duality map

$$udx^2(v\partial) =: \int uvdx$$

so that

$$\mathrm{Ad}^*_{\phi^{-1}} udx^2(v\partial) = udx^2(\mathrm{Ad}_\phi v\partial = \int u\phi'(\phi^{-1})v(\phi^{-1})dx = \int u(\phi(y))|\phi'(y)|^2 v(y)dy$$

and udx^2 is a suggestive notation. The Lie bracket is the negative of the commutator and

$$\mathrm{ad}^*_{v\partial}(udx^2)(w\partial) = udx^2([v\partial, w\partial]) = \int u(vw' - v'w)dx = -\int w((uv)' + uv')dx = ((uv)' + uv')dx^2($$

In particular

$$\sqrt{|u|}dx$$

transforms as a measure and

$$\int \sqrt{|u|}dx$$

is a Casimir. It is not hard to see that it is the only Casimir if u does never vanish. If it vanishes at two points x_1 and x_2 where the derivative does not vanish then this structure is preserved under small perturbation by the implicit function theorem and

$$\int_{x_1}^{x_1} \sqrt{|u|}dx$$

is another Casimir. One can show that these two span the space of all Casimirs. There is a striking consequence:

- If u never vanishes then the coadjoint orbit has codimension 1.
- If u has two zeros with nonvanishing derivative then the codimension is 2.

This is in striking constrast to the finite dimensional case, where the dimension of the coadjoint orbits is always even, and hence the codimenions are all even or all odd.

3.5.2 Definition of the Virasoro-Bott group

We slightly change the notation at define $\mathfrak{D}\mathrm{IFF}(\mathbb{S}^1)$ as the group of orientation preserving diffeomorphisms.

The Virasoro-Bott is a so called universal central extension of $\mathfrak{D}\mathrm{IFF}(\mathbb{S}^1)$. It is essentially unique.

Definition 3.8. *We define the Bott cocycle as*

$$B : \mathfrak{D}\mathrm{IFF}(\mathbb{S}^1) \times \mathfrak{D}\mathrm{IFF}(S^1) \ni (\phi, \psi) \to \frac{1}{2} \int \log((\phi \circ \psi)') \frac{d}{dx} \log \psi' dx \tag{3.9}$$

and the group multiplication on $\mathfrak{V}\mathrm{IR} = \mathfrak{D}\mathrm{IFF}(\mathbb{S}^1) \times \mathbb{R}$ *by*

$$(\phi, s)(\psi, t) = (\phi \circ \psi, s + t + B(\phi, \psi) \tag{3.10}$$

We have to prove that we obtain a group operation. First we claim that the Bott cocycle satisfies the defining relation of a cocycle

$$B(\phi\circ\psi,\eta)+B(\phi,\psi)=B(\phi,\psi\circ\eta)+B(\psi,\eta)).$$

To verify this we calculate

$$\begin{aligned}2B(\phi\circ\psi,\eta) &= \int \log((\phi\circ\psi\circ\eta)')\partial_x\log\eta' dx\\ &= \int_{\mathbb{S}^1}(\log(\phi'\circ\psi\circ\eta)\partial_x\log\eta' + \log((\psi\circ\eta)')\partial_x\log\eta' dx\\ &= \int_{\mathbb{S}^1}(\log(\phi'\circ\psi\circ\eta)\partial_x\log\eta' + 2B(\psi,\eta)\end{aligned}$$

$$\begin{aligned}2B(\phi,\psi\circ\eta) &= \int \log((\phi\circ\psi\circ\eta)')\partial_x\log((\psi\circ\eta)')dx\\ &= \int_{\mathbb{S}^1}(\log(\phi'\circ\psi\circ\eta)\partial_x\log\eta' dx + 2B(\phi,\psi).\end{aligned}$$

Now we check associativity

$$(\phi,s)((\psi,t)(\eta,u)) = (\phi,s)(\psi\circ\eta,t+u+B(\psi,\eta)) = (\phi\circ\psi\circ\eta,s+t+u+B(\phi,\psi\circ\eta)+B(\psi,\eta))$$

$$((\phi,s)(\psi,t))(\eta,u) = (\phi\circ\psi,s+t+B(\phi,\psi))(\eta,u) = (\phi\circ\psi\circ\eta,s+t+u+B(\phi,\psi)+B(\phi\circ\psi,\eta)),$$

identity

$$(1,0)(\phi,s)=(\phi,s+B(1,\phi)=(\phi,s)$$

and inverse

$$(\phi,t)(\phi^{-1},-t)=(1,B(\phi,\phi^{-1}))=(1,0).$$

We proceed with the adjoint representation. Let $\psi(t,x)$ be a family of diffeomorphisms with $\psi(1,x)=x$, $v=\partial_t\psi(x,t)|_{t=0}$ and $\eta(t)$ with $\eta(0)=0$. We compute

$$\begin{aligned}2\frac{d}{dt}\big(B(\psi(t),\phi^{-1})\Big|_{t=0} &= \frac{d}{dt}\int \log((\psi(t,\phi^{-1}(x))')\frac{d}{dx}\log(\phi^{-1})'dx|_{t=}\\ &= \frac{d}{dt}\int \log((\psi'(t,\phi^{-1}(x))(\phi^{-1})')\frac{d}{dx}\log(\phi^{-1})'dx|_{t=0}\\ &= -\int v'(x)\frac{d}{dx}\log\phi' dx\\ &= \int v\frac{d^2}{dx^2}\log\phi' dx\\ &= \int v\frac{\phi'''\phi'-(\phi'')^2}{\phi'^2}dx.\end{aligned}$$

$$I = 2\frac{d}{dt}\big(B(\phi,\psi(t)\phi^{-1})\Big|_{t=0} = \int \Big\{\frac{d}{dt}\log(\phi(\psi(t,\phi^{-1}))')\Big|_{t=0}\Big\}\frac{d}{dx}\log((\phi^{-1})')dx$$

since $\phi(\phi^{-1}(x)) = x$ and hence the first factor vanishes unless the time derivative falls on it. The derivative of the product is product of the derivagtives. We do a substitution $y = \phi^{-1}(x)$,

$$\begin{aligned}
I &= \int \Big\{\frac{d}{dt}\log(\phi'(\psi(t,y))\psi'(t,y)(\phi')^{-1})\Big|_{t=0}\Big\}\frac{d}{dy}(\log((\phi')^{-1})dy \\
&= -\int \Big\{\frac{\phi''v}{\phi'} + v'\Big\}\frac{\phi''}{\phi'}dy \\
&= -\int v\Big(\frac{\phi''}{\phi'}\Big)^2 - v\partial_x\frac{\phi''}{\phi'}dy \\
&= \int v\frac{\phi'''\phi' - 2(\phi'')^2}{(\phi')^2}dy
\end{aligned}$$

hence

$$\mathrm{Ad}_{(\phi,s)}(v\partial,t) = \Big(\phi'(\phi^{-1})v(\phi^{-1})\partial, t + \int v\frac{\phi'''\phi' - \frac{3}{2}(\phi'')^2}{\phi'}dx.\Big) \tag{3.11}$$

Equally important, let with $\partial_t\phi(t)|_{t=0} = u$,

$$\begin{aligned}
\omega(u\partial,v\partial) &:= -\frac{d}{dt}\int v\Big(\frac{3}{2}\frac{(\phi'')^2}{(\phi')^2} + \frac{\phi'''}{\phi'}dx\Big)\Big|_{t=0} \\
&= \int_{\mathbb{S}^1} -vu'''dx = \int_{\mathbb{S}^1} v''u'dx
\end{aligned} \tag{3.12}$$

and the Lie bracket on $\mathfrak{vir}$ becomes

$$[(u\partial,t),(v\partial,s)] = \Big(-(uv_x - vu_x)\partial, \int u'v''dx\Big). \tag{3.13}$$

We write elements of $(udx^2,a) \in \mathfrak{vir}^*$ with the duality map

$$(udx^2,a)(v\partial,t) = \int uvdx + at \tag{3.14}$$

Proposition 3.1. *The coadjoint representation is given by*

$$\mathrm{Ad}^*_{(\phi^{-1},s)}(udx^2,a) = \Big(u(\phi)\phi'^2dx^2 + a\frac{\phi'\phi''' - \frac{3}{2}(\phi'')^2}{(\phi')^2}dx^2, a\Big)$$

and

$$\mathrm{ad}^*_{(v\partial,b)}(udx^2,a) = \Big(-(2uv' + u'v + av''')dx^2, 0\Big).$$

Proof. By definition

$$\begin{aligned}
\mathrm{Ad}^*_{\phi^{-1},-s}(udx^2,a)(v\partial_x,b) &= (udx^2,a)\Big(\mathrm{Ad}_{(\phi,s)}(v\partial_x,b)\Big)\\
&= (udx^2,a)\left(\phi'(\phi^{-1})v(\phi^{-1})\partial, b+\int v\frac{\phi'''\phi'-\frac{3}{2}(\phi'')^2}{(\phi')^2}dx\right)\\
&= \int u(x)\phi'(\phi^{-1}(x))v(\phi^{-1}x)+av\frac{\phi'''\phi'-\frac{3}{2}(\phi'')^2}{(\phi')^2}dx+ab\\
&= \int v\Big(u(\phi(x)(\phi'(x))^2+ab+a\frac{\phi'\phi'''-\frac{3}{2}(\phi'')^2}{(\phi')^2}dx\Big)
\end{aligned}$$

and

$$\begin{aligned}
\mathrm{ad}^*_{(v\partial,b)}(udx^2,a)(w\partial,c) &= -(udx^2,a)([(v\partial,b),(w\partial,c)])\\
&= -(udx^2,a)\Big(-(vw'-v'w),\int v'w''dx\Big)\\
&= -\int u(v'w-vw')+av'w''dx\\
&= -\int((uv)'+uv'+av''')wdx.
\end{aligned}$$

□

The central variable a is fixed under the coadjoint action. The coadjoint action depends only on the diffeomorphism!

3.5.3 Hill's operator

It is convenient to write elements $(u\,dx^2,a)\in\mathfrak{vir}^*$ as Hill's operator

$$2a\partial^2+u\in\mathfrak{vir}^*$$

Recall that if

$$(2a\partial^2+u)f=(2a\partial^2+u)g=0$$

then the Wronskian $W(f,g)=fg'-gf'$ is constant,

$$\frac{d}{dx}(fg'-gf')=0$$

Definition 3.9. *We define the Schwarzian derivative of η as*

$$S(\eta)=\frac{\eta'\eta'''-\frac{3}{2}(\eta'')^2}{(\eta')^2}$$

Proposition 3.2. *Suppose that f and g are linearly independent, equivalently $W(f,g) \neq 0$ and $g \neq 0$. Then*

$$u = aS(f/g).$$

Proof. Let $\eta = f/g$. Then

$$\eta' = -\frac{W(f,g)}{g^2}, \qquad \eta'' = 2\frac{W(f,g)g'}{g^3}, \qquad \eta''' = 2\frac{W(f,g)g''}{g^3} - 6\frac{W(f,g)(g'')^2}{g^4}$$

and

$$\begin{aligned} aS(\eta) &= a\frac{\eta'\eta''' - \frac{3}{2}(\eta'')^2}{(\eta')^2} \\ &= ag^4\Big(-2\frac{g''}{g^5} + 6\frac{(g')^2}{g^6} - 6\frac{(g')^2}{g^6}\Big) \\ &= u. \end{aligned}$$

□

Theorem 3.10. *Let $a \neq 0$,*

$$(Udx^2, a) := \mathrm{Ad}^*_{(\phi^{-1},0)}(udx^2, a) = ([u \circ \phi|\phi'|^2 + aS(\phi)]dx^2, a).$$

Let f, g be as above. Then

$$(2a\partial^2 + U)\big[(f \circ \phi)|\phi'|^{-1/2})\big] = 0 \tag{3.15}$$

and

$$aS((f/g) \circ \phi) = U. \tag{3.16}$$

Proof. We begin with

$$S(\phi \circ \psi) = S(\phi) \circ \psi(\psi')^2 + S(\psi) \tag{3.17}$$

which implies (3.16). We compute

$$(\phi \circ \psi)' = \phi' \circ \psi\psi'$$

$$(\phi \circ \psi)'' = \phi'' \circ \psi(\psi')^2 + \phi' \circ \psi\psi''$$

$$(\phi \circ \psi)''' = \phi''' \circ \psi(\psi')^3 + 3\phi'' \circ \psi\psi'\psi'' + \phi' \circ \psi\psi'''.$$

In order to lighten the notation we omit the $\circ\phi$ in the sequel. Then

$$\begin{aligned}
&\frac{\phi'\phi'''(\psi')^4 + 3\phi'\phi''(\psi')^2\psi'' + \phi'^2\psi'\psi''' - \frac{3}{2}(\phi''(\psi')^2 + \phi'\psi'')^2}{(\phi'\psi')^2} \\
&\qquad = \frac{\phi'\phi'''(\psi')^4 + \phi'^2\psi'\psi''' - \frac{3}{2}(\phi'')^2(\psi')^4 - \frac{3}{2}(\phi')^2(\psi'')^2}{(\phi'\psi')^2} \\
&\qquad = \frac{\phi'\phi''' - \frac{3}{2}(\phi'')^2}{(\phi')^2}(\psi')^2 + \frac{\psi'\psi''' - \frac{3}{2}(\psi'')^2}{(\psi')^2}.
\end{aligned}$$

Finally

$$\begin{aligned}
&(a\partial^2 + 2u\circ\phi(\phi')^2 + aS(\phi))(f\circ\phi|\phi'|^{-1/2}) \\
&\quad = (af'' + uf)\circ\phi\phi'^{3/2} \\
&\qquad - a\frac{\phi'\phi''' - \frac{3}{2}(\phi'')^2}{(\phi')^2} f\circ\phi\phi'^{-1/2} + S(\phi)(f\circ\phi)(\phi')^{-1/2}
\end{aligned}$$

□

3.5.4 The Korteweg-de Vries equation

We want to use a Lie-Poisson bracket on quadratic differentials for fixed a and we need to define nonlinear functions on $\mathfrak{vir}*$. We insist that we always consider smooth functionals

$$F : \mathfrak{vir}^* \to \mathbb{R}$$

by which we mean that there exists N so that F extends to a smooth map

$$F : \{udx^2 : u \in C^N\} \to \mathbb{R}.$$

We define the variational derivative

$$\frac{\delta F}{\delta u}(u)(v) = \frac{d}{dt}F(u+tv)|_{t=0}$$

and we consider $\frac{\delta F}{\delta u}(u) \in (C^N(\mathbb{S}))^*$.

We define the inner product

$$\langle (v\partial, a), (u\partial, b)\rangle = -\frac{1}{2}\int uvdx - \frac{1}{2}ab,$$

in $\mathfrak{vir}^*$. The tangent space at $\eta \in \mathfrak{VIR}$ is given by

$$(v\circ\eta\partial, a) \in T_\eta\,\mathfrak{VIR}\,.$$

We equipp it with the right invariant metric for v and a,

$$-\frac{1}{2}\int |\frac{d}{dt}\eta\circ\eta^{-1}|^2dx - \frac{1}{2}a_t^2.$$

Arnold's reduction to $\mathfrak{vir}^*$ - with the same calculation as for the top - is

$$\frac{d}{dt}(udx^2, a) = \operatorname{ad}^*_{-2u,0} u$$

with the + sign since we use the right invariant metric, meaning that we consider the velocity as a function of the actual position, not material coordinates.The relation between the two is

$$v^R(t,x) = v^L(t, \phi^{-1}(t)).$$

Theorem 3.11. *The Euler equation is the Korteweg-de Vries equation*

$$u_t - 2au_{xxx} - 6uu_x = 0$$

and $a_t = 0$

Proof. The Euler equation for the Hamiltonian

$$H((udx^2, a)) = -\int |u|^2dx + a^2$$

is

$$\frac{d}{dt}m = -2\operatorname{ad}^*_{A^{-1}m(t)} m(t)$$

which becomes

$$(u_tdx^2, a_t) = -2\operatorname{ad}^*_{(u\partial,a)}(udx^2, a) = ((6uu_x + 2au_{xxx})dx^2, 0).$$

The sign convention may be irritating. □

We observe that the central a becomes a parameter which we can choose. It acts trivially and we essentially reduce the consideration the hyper plane $a = -1/2$, for which we obtain the Korteweg-de Vries equation:

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

Theorem 3.12. *There exists a sequence of Poisson commuting integrals*

$$H_{-1}(u) = \int u, H_0(u) = \frac{1}{2}\int u^2dx, H_1(u) = \int \frac{1}{2}u_x^2 + u^3dx$$

$$H_2 = \frac{1}{2}\int u_{xx}^2 + 10uu_x^2 + 5u^4dx.$$

Proof. **Step 1.** We define the family of Lie-Poisson brackets

$$\{F,G\}^s(wdx^2,-\frac{1}{2}) = ((w+s)dx^2,-\frac{1}{2})\left(\left[(\frac{\delta F}{\delta u},\frac{\partial F}{\partial a}),(\frac{\delta G}{\delta u},\frac{\partial G}{\partial a})\right]\right)(wdx^2,a)$$
$$= \int w(\partial_x\delta F\delta G - \delta F\partial_x\delta G\partial_x\delta F - \frac{1}{2}\big(\partial_x\delta F\partial_x^2\delta G - (\partial_x^2\delta F)\partial_x\delta G\big)dx$$
$$+ 2s\int \delta_x F\delta G - \delta F\delta_x\delta G.$$

for $t=0$. The bracket

$$\{F,G\}^{Gardner} = \int \partial_x\delta F\delta G - \delta F\partial_x\delta G\,dx$$

is the Gardner bracket, which satisfies the Jacobi identity since it is a constant skew symmetric form. The Jacobi identity is clearly satisfied.
Step 2. The KdV equation is a Hamiltonian equation for all these Poisson structures. Consider the Hamiltonian

$$H(udx^2,-\frac{1}{2}) = \int -u^2+4sudx$$

so that

$$\delta H = -2u+4s$$

and

$$-\frac{1}{2}\partial_{xxx}(2u+4s)+(u+s)\partial(2u-4s)+\partial((u+s)(2u-4s) = -u_{xxx}+6uu_x$$

Step 3: The Casimirs
Let f,g satisfy with $\lambda\in\mathbb{R}$

$$-\partial^2 f+(u+\lambda^2)f=0$$

and $f^\eta = f\circ\eta(\eta')^{-1/2}$, and $g^\eta = g\circ\eta(\eta')^{-\frac{1}{2}}$.

Lemma 3.13. *The Wronskian is a Casimir in the sense that*

$$W(f^\eta,g^\eta) = W(f,g).$$

and it is a Casimir for the s shifted Lie-Poisson bracket,

$$(-\partial^3+u\partial+\partial u+2\lambda^2\partial)\frac{\delta}{\delta u}W(f,g) = 0.$$

In the periodic case, which we consider here, there is another important Casimir, related to the monodromy. We write $-\phi'' + u\phi = -\lambda^2\phi$ as a system

$$\phi_0\prime = \phi_1, \phi_1' = (u+\lambda^2)\phi_0$$

The monodromy matrix $M = X(1)$ where

$$X(0) = 1_{\mathbb{R}^2}, \quad \dot{X} = \begin{pmatrix} 0 & 1 \\ u+\lambda^2 & 0 \end{pmatrix}. \tag{3.18}$$

A change of coordinates leads to a similar monodromy matrix hence

$$\log \operatorname{tr} M$$

is a Casimir.
Step 4: Asymptotics of the Casimir. Suppose that λ is large. The functions $e^{\lambda x}$ and $e^{-\lambda x}$ are solutions to

$$-\phi'' + \lambda^2\phi = 0.$$

Then

$$X^0(x) = \begin{pmatrix} \frac{1}{2}(e^{\lambda x} + e^{-\lambda x}) & \frac{1}{2\lambda}(e^{\lambda x} - e^{-\lambda x}) \\ \frac{\lambda}{2}(e^{\lambda x} - e^{-\lambda x}) & \frac{1}{2}(e^{\lambda x} + e^{-\lambda x}) \end{pmatrix}$$

$$M = \begin{pmatrix} \frac{1}{2}(e^{\lambda x} + e^{-\lambda x}) & \frac{1}{2\lambda}(e^{\lambda} - e^{-\lambda}) \\ \frac{\lambda}{2}(e^{\lambda} - e^{-\lambda}) & \frac{1}{2}(e^{\lambda} + e^{-\lambda}) \end{pmatrix} = e^{\lambda} + e^{-\lambda}$$

and

$$\log(\operatorname{tr} M) - \lambda = O(\lambda^{-\infty}).$$

Now we consider (3.18) which we can solve by a fixed point argument,

$$X(x) = X^0(x) + \int_0^x X^0(x)(X^0(s))^{-1} \begin{pmatrix} 0 & 0 \\ u & 0 \end{pmatrix} X(s) ds$$

and it is an exercise to show that with some $c \neq 0$

$$\log \operatorname{tr} X(1) - \lambda = \frac{c}{2\lambda} \int_0^1 u dx$$

One obtains an asympotic expansion

$$\frac{1}{2c} M(\lambda) \sim \sum_{j=-1}^{\infty} H_j (2\lambda)^{-3-2j}$$

with $H_{-1} = \frac{1}{2} \int u dx$.

Step 5: Lenard recursion We have seen that there is a Casimir $\log \operatorname{tr} M(\lambda) - \lambda$ for $\{,\}^{\lambda^2}$. Then

$$(-\partial 12\partial^3 + u\partial + \partial u + 2\lambda^2\partial) \sum_{n=-1}^{N} H_n (-1)^n \lambda^{-3-2n}$$

with $H_{-1} = \frac{1}{2}\int u dx$, $\delta H_{-1} = -1$,

$$(-\partial^3 + u\partial + \partial u)\frac{1}{2} = u'.$$

In general

$$(\partial^3 - 2u\partial - 2\partial u)\delta H_n = -\partial \delta H_{n+1}.$$

Thus

$$u' = \partial \delta H_0, \quad H_0 = \frac{1}{2}\int u^2 dx$$

$$(\partial^3 - 2u\partial - 2\partial u)\delta H_0 = u_{xxx} - 6uu_x = -\partial \delta H_1$$

$$H_1 = \int \frac{1}{2}u_x^2 + u^3 dx$$

$$\begin{aligned}(\partial^3 - 2u\partial - 2\partial u)(-u_{xx} + 3u^2) &= -u^{(5)} + 2uu_{xxx} + 2\partial(uu_{xx}) + 3\partial^3 u^2 - 30u^2 u_x \\ &= -\partial(u^{(4)} - 4uu_{xx} + u_x^2 - 3\partial^2 u^2 + 10u^3).\end{aligned}$$

hence

$$H_2 = \frac{1}{2}\int u_{xx}^2 + 10uu_x^2 + 5u^4 dx.$$

□

3.5.5 The case of KdV on the line

Let $\mathfrak{V}\mathrm{IR}(\mathbb{R})$ be the set of monotone diffeomorphisms of the real line so that

$$\partial_x \eta - 1 \in H^N(\mathbb{R})$$

for all N. The whole setup generalizes to this setting.

There are special solutions, the left and right Jost functions, if λ is sufficiently large. They are characterized by

$$\lim_{x\to-\infty} e^{-\lambda x} f_l(x) = \lim_{x\to+\infty} e^{\lambda x} f_r = 1$$

The operator

$$(L + \lambda^2)\psi = (-\partial^2 + (u + \lambda^2))\psi = f$$

is invertible with a compact inverse on the circle. Formally we can write

$$(-\partial^2+u+\lambda^2) = (\partial+\lambda)(-\partial+\lambda)+u = (\partial+\lambda)(1+(\partial+\lambda)^{-1}u(-\partial+\lambda))^{-1}(-\partial+\lambda)$$

and we can rewrite the equation as

$$(1 + (\partial + \lambda)^{-1}u(-\partial + \lambda)^{-1}\tilde{\psi} = \tilde{f}$$

and

$$\tilde{\psi} = \sum_{n=0}^{\infty}(-1)^n(\partial + \lambda)^{-1}u(-\partial + \lambda)^{-1}\tilde{f}$$

We can also use the left and right Jost function to realize the resolvent as an operator with an integral kernel,

$$L^{-1}f = \int g(x,y)f(y)dy$$

with

$$g(x,y) = (W(\psi_l,\psi_r))^{-1}\left\{\begin{array}{ll} \psi_l(x)\psi_r(y) & \text{if } x<y \\ \psi_r(x)\psi_l(y) & \text{if } y<x \end{array}\right.$$

Lemma 3.14. *The function*

$$-\frac{1}{2\lambda}W_\lambda(\psi_{l,\lambda},\psi_{r,\lambda})$$

is a Casimir which satisfies

$$\lim_{\lambda\to\infty}\log\frac{1}{2\lambda}W_\lambda(f_{r,\lambda},f_{l,\lambda} = \frac{c}{\lambda}\int udx + O(\lambda^{-3}).$$

3.6 Geodesics on Ellipsoids

This is a classical problem where the techinique of contrained Hamiltonians can be used. The close connection to geometry allows to see the setting of Liouville integrability in action. The approach follows Moser [22].

3.6.1 Constrained Hamiltonian systems

We consider Hamiltonian equations on $\mathbb{R}^{2n} \ni (p,q)$ with the cannonical symplectic form. Let $M^{2(n-r)}$ be a submanifold given as nondegenerate level set of the smooth vector valued function $(G_j)_{1\le j\le 2r}$,

$$M = \{x : G_j(x) = 0 \qquad \text{for } 1 \le j \le 2r\}$$

and $\operatorname{rk} DG(x) = 2r$ for $x \in M$. We require even

$$\det(\{G_i, G_j\})_{1\le i,j\le 2r} \neq 0$$

which makes M a symplectic manifold (the restriction of the cannonical form remains closed. This reduces the question to the linear question of chosing a basis). We consider the Hamiltonian equations

$$\frac{d}{dt}\begin{pmatrix} p \\ q \end{pmatrix} = J\nabla H(p,q)$$

resp.

$$\dot p_j = -\frac{\partial H}{\partial q_j}, \quad \dot q_j = \frac{\partial H}{\partial p_j}.$$

Orbits stay on M if $\{H, G_j\} = 0$ on M, which is not true in general.

We constrain it to M by replacing ∇_H by the vector field

$$\nabla_H - \sum_{j=1}^{2r} \lambda(p,q)\nabla_{G_r}$$

so that this vector field is tangential. This requires

$$\{H, G_k\} - \sum_{j=1}^{2r} \lambda_j(p,q)\{G_j, G_k\} = 0$$

for $1 \le k \le 2r$ which can be solved for λ_j by the assumption of nondegeneracy. Then $H|_M$ is the Hamiltonian of this vector field.

There is no reason to expect that the reduced system is integrable, even if the original system has been integrable. There is however one situation when it is integrable.

Suppose that F_j, $1 \le j \le n$ Poisson commute with H and with themselves, assume that in the setting above $G_{r+j} = F_j$, $1 \le j \le r$ and in addition

$$\det\{G_j, F_k\} \neq 0.$$

We will see that then the Hamiltonian

$$H^* = H - \sum \lambda_j F_j - \sum \mu_j G_j$$

and F_j, $r < j \le n$ Poisson commute on M and hence the constraint equation is integrable (we again neglect nondegeneracy and compactness of the level sets). To see this we argue as follows. First for $1 \le k \le r$

$$0 = \{H^*, F_k\} = \sum_{j=1}^{r} \mu_j \{G_j, F_k\}$$

hence $\mu_j = 0$ and

$$H^* = H - \sum_{j=1}^{r} \lambda_j F_j$$

where the λ_j are defined by $\{H^*, G_j\} = 0$. Then

$$\{H^*, F_j\}|_M = \{F_j, F_k\}_M = 0.$$

In particular all the F_j are constant under the flow of H^* which stays on M.

3.6.2 Geodesic flow on the ellipsoid

Let A be a nondegenerate symmetric $n \times n$ matrix. We consider the nondegenerate level set

$$E = \{q \in \mathbb{R}^n : \langle A^{-1}q, q\rangle = 1\}.$$

It is an ellipsoid if A is positive definite. The constraint second order differential equations are given by

$$\ddot{q} = -\nu A^{-1} q$$

with ν determined by

$$0 = \frac{d}{dt}^2 \langle A^{-1}q, q\rangle = \langle A^{-1}q, \ddot{q}\rangle + \langle A^{-1}\dot{q}, \dot{q}\rangle = -\nu |A^{-1}q|^2 + \langle A^{-1}\dot{q}, \dot{q}\rangle$$

so that

$$\nu = |A^{-1}q|^{-2} \langle A^{-1}\dot{q}, \dot{q}\rangle.$$

In a Hamiltonian formulation we constrain the free Hamiltomian

$$\frac{1}{2}|p|^2$$

to the manifold (the cotangent bundle of E which we can identity with the tangent bundle). The cotangent bundle carries a cannonical one form which in local coordinates in $U \subset \mathbb{R}^n_q \times \mathbb{R}^n_p$ is given by

$$\sum_j p_j \wedge dq_j.$$

$$M = \{(p, q) : \langle A^{-1}q, q\rangle = 0, \langle A^{-1}q, p\rangle = 0\}.$$

The 1 form on the cotangent bundle is simple the restriction of the cannonical 1 form in $\mathbb{R}^{2n}$. We find

$$H^* = \frac{1}{2}|p|^2 - \lambda_1(\langle A^{-1}q, q\rangle - 1) - \lambda_2 \langle A^{-1}q, p\rangle$$

where

$$\lambda_1 = -\frac{1}{2}|A^{-1}q|^{-2}\langle A^{-1}p, p\rangle, \quad \lambda_2 = |A^{-1}q|^{-2}\langle A^{-1}q, p\rangle$$

which we write as

$$H^* = \frac{1}{2}|p|^2 + \frac{\mu}{2}\Phi(p,q) - \frac{\mu}{2}\langle A^{-1}q, p\rangle^2$$

with

$$\mu = |A^{-1}q|^{-2}, \quad \Phi(p,q) = (\langle A^{-1}q, q\rangle - 1)\langle A^{-1}p, p\rangle - \langle A^{-1}q, p\rangle^2.$$

The last term in H^* vanishes quadratically and we can drop it without harm. Clearly

$$\{(q,p) : q \in E, \Phi(p,q) = 0\}$$

is the (co)tangent space of the ellipse. We arrive at the Hamiltonian

$$H^* = \frac{1}{2}|p|^2 + \frac{\mu}{2}\Phi$$

with the contrained Hamiltonian equations

$$\begin{aligned} \dot{q} &= y \\ \dot{p} &= -\mu\langle A^{-1}p, p\rangle A^{-1}q. \end{aligned} \tag{3.19}$$

Theorem 3.15. *The Hamiltonian equation defined by the Hamiltonian $H|_M*$ on the (co)tangent bundle M of E is Liouville integrable on an open set with a complement of codimension* 1*.*

A short calculation shows that $\{\Phi, \frac{1}{2}|p|^2\} = 0$ and we are in the setting of the integrable constrained equations provided we find functions which Poisson commute with $\frac{1}{2}|p|^2$ and Φ.

3.6.3 Construction of Poisson commuting functions

Let α_j by the eigenvalues of A. We may assume that A is diagonal to make the calculations more concrete, but we prefer to formulate in a coordinate independent fashion. We consider the (confocal) quadrics E_z for $z \neq \alpha_k$

$$\{q : \langle (z1_{\mathbb{R}^n} - A)^{-1}q, q\rangle + 1 = 0\},$$

We define analogously to above

$$Q_z(p,q) = \langle (z1 - A)^{-1}q, p\rangle; \quad Q_z(x) = Q_z(x,x)$$

$$
\begin{aligned}
\Phi_z(p,q) &= (1+Q_z(q))Q_z(p) - Q_z^2(p,q) \\
&= \sum \frac{p_j^2}{z-\alpha_j} + \sum_{j,k} \frac{q_j^2 p_k^2 - p_j q_j p_k q_k}{(z-\alpha_j)(z-\alpha_k)}
\end{aligned}
$$

We do a partial fraction expansion, assuming that all α_k are different: We multiple by $z-\alpha_i$ and evaluate at $z=\alpha_i$,

$$
\Phi_z(p,q) = \sum_{k=1}^{n} \frac{F_k(p,q)}{z-\alpha_k}
$$

with

$$
F_k(p,q) = p_k^2 + \sum_{j\neq k} \frac{(q_j p_k - q_k p_j)^2}{\alpha_k - \alpha_j}
$$

Lemma 3.16. *For all* $z_1, z_2 \neq \alpha_k$

$$
\{\Phi_{z_1}, \Phi_{z_2}\} = 0
$$

and also

$$
\begin{aligned}
\{F_j, F_k\} &= 0 \\
\{F_j, |p|^2\} &= 0.
\end{aligned}
$$

Proof. Only the first identity need to be proven, since it implies the second. The third equality follows by

$$
|p|^2 = \sum_{k=1}^{n} F_k.
$$

The proof of the first identity is a direct calculation using

$$
\begin{aligned}
\{Q_z(q), Q_z(p)\} &= 4\langle A^{-1}q, A^{-1}p\rangle, \\
\{Q_z(q), Q_z(p,q)\} &= 2|A^{-1}q|^2, \\
\{Q_z(p,q), Q_z(p)\} &= 2|A^{-1}p|^2
\end{aligned}
$$

which is left as a tedious exercise. □

3.6.4 The Lax-pair

Let

$$
L = \left(1_{\mathbb{R}^n} - \frac{1}{|p|^2} p\otimes p\right)(A - x\otimes x)\left(1_{\mathbb{R}^n} - \frac{1}{|p|^2} p\otimes p\right)
$$

and

$$
B = -\left(\frac{p_i q_j - p_i q_j}{\alpha_i \alpha_j}\right)_{1\leq i,j\leq n}.
$$

Theorem 3.17.

$$\dot{L} = [B.L]$$

is equivalent to

$$\dot{q}_j = \frac{\partial\Phi}{\partial p_j}, \quad \dot{p}_j = -\frac{\partial\Phi}{\partial q_j}$$

Proof. We spell out

$$\dot{q} = 2(\langle A^{-1}q, q\rangle - 1)A^{-1}p - 2\langle A^{-1}q, p\rangle A^{-1}p$$

$$\dot{p} = -2\langle A^{-1}p, p\rangle A^{-1}q - 2\langle A^{-1}q, p\rangle A^{-1}p$$

Since $\{|p|^2, \phi(p,q)\} = 0$ we may restrict to $|p| = 1$. Then

$$L = (1 - p \otimes p)(A - x \otimes x)(1 - p \otimes p)$$

and

$$B = A^{-1}p \otimes A^{-1}q - A^{-1}q \otimes A^{-1}p$$

so that

$$\begin{aligned}
(1 - p \otimes p)(A - q \otimes q)(1 - p \otimes p)A^{-1}p \otimes A^{-1}q \\
&= (1 - p \otimes p)(A - q \otimes q(A^{-1}p \otimes A^{-1}q - \langle p, A^{-1}p\rangle p \otimes A^{-1}q) \\
&(1 - p \otimes p)\Big(p \otimes A^{-1}q - \langle q, A^{-1}p\rangle q \otimes A^{-1}q \\
&\quad - \langle p, A^{-1}p\rangle Ap \otimes A^{-1}p + \langle q, p\rangle\langle p, A^{-1}p\rangle q \otimes A^{-1}q\Big)
\end{aligned}$$

The only point here is to show that the calculation is doable. It is left as a tedious exercise. □

Lemma 3.18.

$$\frac{|p|^2}{z}\frac{\det(z1 - L)}{\det(z1 - A)} = \sum_{k=1}^{n} \frac{F_k}{z - \alpha_k}$$

Proof. For simplicity we assume $|p| = 1$. Then

$$\begin{aligned}
z1 - L &= z1 - (1 - p \otimes p)(A - q \otimes q)(1 - p \otimes p) \\
&= (z1 - A) + \langle p, Ap\rangle p \otimes p - Ap \otimes p - p \otimes Ap - q' \otimes q'
\end{aligned}$$

where $q' = (1 - p \otimes p)q$. Then

$$\frac{\det(z1 - L)}{\det(z1 - A)} = \det\Big(1 + (z1 - A)^{-1}\Big(\langle p, Ap\rangle p \otimes p - Ap \otimes p - p \otimes Ap - q' \otimes q'\Big)\Big)$$

where the right hand side is the determinant of matrix 1+ soemthing of rank 3.This can be expanded. □

The eigenvalues β_j of L are all different. They can be considered as functions of the F_k and hence they Poisson commute, and also with $|p|^2$.

There is a geometric interpretation:

$$E_0 = \{(p,q) : \langle A^{-1}q, q\rangle = 1, \Phi(p,q) = 0\}$$

and

$$\Phi(p,q) = 0$$

iff the line

$$\{q + tp\}$$

is tangent to E_0 and the Hamiltonian ϕ describes an evolution of tangent lines. With this interpretation one can describe the tori as the set of simultaneous tangent lines to to the E_{β_j}.

p is an eigen function to the eigenvalue 0. There is a second eigen function to the eigenvalue 0 on E since Φ vanishes. There remain $n-2$ nonzero eigenvalues of L which are zeros of

$$\sum_{k=1}^n \frac{F_k}{z-\alpha_k}.$$

The eigen vectors are the normals of the E_{β_j} to the tangent and hence they are orthogonal. We have

$$\sum \alpha_k^{-1} F_k = -\Phi(p,q) = 0$$

which is a simple zero.

The level sets are compact. Suppose the α_j are pairwise disjoint.

3.7 The Calogero-Moser system

Here we consider the Hamiltonian

$$H(p,q) = \frac{1}{2}|p|^2 + \sum_{i\neq j} \frac{1}{(q_j - q_k)^2} \tag{3.20}$$

which describes n particle on the real line with the repelling potential so that particles never collide.

Theorem 3.19. *On an open set whose complement has Hausdorff dimenionsion at most* 1 *the Hamiltonian equation above is Liouville integrable.*

The point of view in this subsection is due to Kazhdan, Kostant, and Sternberg [12] and Etingof [7]. Part of the presentation follows [15].

3.7.1 Moment maps and Poisson group action

Let M be a Poisson manifold and, $\mathfrak{g}$ the Lie algebra of the connected compact Lie (matrix) group G. We recall that $\mathfrak{g}^*$ carries a Poisson structure. A map $J : M \to \mathfrak{g}^*$ is called a moment map if it is a Poisson map. This defines a (anti) Lie algebra map fron $\mathfrak{g}$ to the vector fields of M by

$$\nabla_X : \mathfrak{g} \ni X \to \nabla_{J(x)(X)} =: \nabla_X.$$

We have

$$\nabla_{[X,Y]} = [\nabla_Y, \nabla_X].$$

We obtain an action of G by mapping $\exp(tX)$ to the map defined as time t map of the ODE

$$\dot{x} = \nabla_X(x).$$

Examples are

- $M = T^*(G)$,

  $$J((g, mg^t)) = m \in \mathfrak{g}^*$$

  One can check that

  $$\{F \circ J, G \circ J\}^{T^*G} = \{F, G\}^{\mathfrak{g}^*} \circ J.$$

- $M = \mathfrak{su}^*(n) \times \mathfrak{su}(n)$. Consider the special unitary group $\mathrm{SU}(n)$ of complex unitary $n \times n$ matrices of determinant 1. They are a Lie group with Lie algebra $\mathfrak{su}(n)$, the skew adjoint matrices with trace 0.

  We consider on $\mathfrak{su}(n)$ the inner product

  $$\langle A, B\rangle := -\operatorname{tr} AB^*$$

  which gives $\mathfrak{su}(n)$ the structure of a complex Euclidean vector space. On $\mathfrak{su}(n) \times su(n) = \mathfrak{su}^*(n) \times \mathfrak{su}(n) = T^* \mathfrak{su}(n)$ we define the symplectic form

  $$\omega((A_1, A_2), (B_1, B_2)) = \langle A_1, B_2\rangle - \langle A_2, B_1\rangle = -\operatorname{tr} A_1 B_2^* + \operatorname{tr} A_2 B_1^*$$

  The group $\mathrm{SU}(n)$ acts on $\mathfrak{su}(n)$ by the adjoint representation, and on $\mathfrak{su}^*(n)$ by

  $$\mathrm{Ad}_U^*(m)(A) = m(U^{-1}AU) = -\operatorname{tr}(mU^{-1}AU) = -\operatorname{tr} UmU^{-1}A = UmU^{-1}(A)$$

and the coadjoint action representation agrees with the adjoint representation. There is the obvious action resp. representation of $\mathrm{SU}(n)$ on $\mathfrak{su}(n) \times \mathfrak{su}(n)$. As for $SO(n)$

$$\langle \operatorname{ad}_A B, C\rangle + \langle B, \operatorname{ad}_A C\rangle = -\frac{1}{2}\Big(\operatorname{tr}(AB - BA)C^* + \operatorname{tr}(B(AC - CA)^*) \Big) = 0$$

There is the cannonical map

$$\mathfrak{su}^*(n) \times \mathfrak{su}(n) \ni (A, B) \to \phi(A, B) = [A, B] \in \mathfrak{su}^*(n).$$

We claim that it is a Poisson map:

$$\{F \circ \phi, G \circ \phi\}^{\mathfrak{su}^*(n) \times \mathfrak{su}(n)}(A, B) = \{F, G\}^{\mathfrak{su}^*(n)}([A, B]).$$

It suffices to verify this for linear functions

$$\begin{aligned} - \operatorname{tr}(F(AB - BA)), &- \operatorname{tr}(G(AB - BA)) \\ &+ \{\operatorname{tr}(F(AB - BA)), \operatorname{tr}(G(AB - BA))\} \\ &= \operatorname{tr}(BF - FB)(GA - AG)^* - \operatorname{tr}(FA - AF)(GB - BG)^* \\ &= - \operatorname{tr}[A, B][F, G]^*. \end{aligned}$$

It is compatible with the action of $\mathrm{SU}(n)$

$$[UAU^{-1}, UBU^{-1}] = U[A, B]U.$$

In the second setting let $N \in \mathfrak{su}(n)$, G the stabilizer group. Then $\mu^{-1}(N)$ is a union of G orbits. Suppose that N is a regular value of the moment map and that $\Phi^{-1}(D)/G$ is smooth then it has a natural symplectic structure.

To see this we observe that if f and g are G invariant then so is $\{f, g\}$ since $g \in G$ is a symplectomorphism. We obtain a Poisson structure on $\mathfrak{su}(n) \times \mathfrak{su}(n)/G$ if this is smooth.

In this setting let M/G be the equivalence classes and assume that M/G is a smooth manifold. If f, h are invariant under G then also $\{f, h\}$ is invariant under the action of G. Since G invariant functions are in one to one correspondence to functions on M/G we obtain a Poisson structure on M/G.

3.7.2 The Calogero-Moser system

Let $N \in \mathfrak{su}(n)$ with the stablizer group G,

$$G = \{U \in \mathrm{SU}(n) : \operatorname{Ad}_U N = UNU^{-1} = N\}.$$

Then G acts on

$$M = \{(A, B) \in \mathfrak{su}(n) \times \mathfrak{su}(n) : [A, B] = N\}.$$

We are interested in the quotient

$M/G = \{(A_1, A_2) \sim (B_1, B_2) \in M : \text{ There exists } U \in G : UB_1U^{-1} = A_1, UB_2U^{-1} = A_2\}.$

Lemma 3.20. *Let*

$$N = i \left(-1_{\mathbb{R}^n} + \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} (1, \dots, 1) \right)$$

and $[Q, P] = D$. *Then there exists* $g \in G$ *so that* gQg^{-1} *is the diagonal matrix with diagonal* iq_j, $q_1 \leq q_2 \dots$. *The off diagonal entries of* P *are*

$$p_{jk} = -\frac{i}{q_j - q_k}.$$

Proof. Any matrix $X \in \mathfrak{su}(n)$ can be diagonalized. We choose $g \in SU(n)$ so that

$$D = gQg^{-1}$$

is diagonal. Let $E = gPg^{-1}$ and $w = g \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$. Then

$$[D, E] = i(1 - w \otimes \bar{w})$$

with 0 on the diagonal (since D is diagonal). Thus $w_j \bar{w}_j = 1$ and $w_j = e^{it_j}$ for some t_j. Thus the product of the diagonal matrix with ientries e^{-it_j} and g is in the stabilitzer of N. We permute the coordinates so that the diagonal entries iq_j of Q are ordered. □

Now P has the off diagonal entries

$$p_{jk} = \frac{-i}{q_j - q_k}.$$

while the diagonal entries are arbitrary. In particular all the q_j are different.

Lemma 3.21. *The value* N *is a regular value for the moment map.* G *acts freely on* $\Phi^{-1}(N)$

Proof. This is a consequence of the fact that $\mathfrak{su}(n)$ is simple (has no proper ideals). If

$$gQg^{-1} = \tilde{g}Q\tilde{g}^{-1}$$

then

$$\tilde{g}^{-1}gQg^{-1}\tilde{g} = Q$$

which implies $\tilde{g} = g$. $\square$

In particular M/G is smooth.

Let

$$\mathcal{A} = \{X \in \mathfrak{su}(n) : X - i = iw \otimes w \text{ with } |w| = \sqrt{n}\}$$

This is an orbit of the coadjoint action of $\mathrm{SU}(n)$. If f, g are $\mathrm{SU}(n)$ invariant functions on $\mathfrak{su}(n) \times \mathfrak{su}(n)$ then $\{f, g\}$ is clear $\mathrm{SU}(n)$ invariant. Let

$$\tilde{M} = \{(X, Y) \in \mathfrak{su}(n) \times \mathfrak{su}(n) : [X, Y] \in \mathcal{A}\}$$

Given $X \in \mathcal{A}$ and

$$f_X(A, B) = \langle X, [A, B]\rangle.$$

Then

$$\{f_X, f_Y\}(A, B) = \langle X + Y, [A, B]\rangle$$

and the span J of these functions is invariant under the Poisson bracket. The quotient

$$C^\infty(\mathfrak{su}(n) \times \mathfrak{su}(n))/J$$

can be understood as functions on the quotient $\tilde{M}/\mathrm{SU}(n)$.

The group $\mathrm{SU}(n)$ acts on $\tilde{M}$. Moreover

$$\tilde{M}/\mathrm{SU}(n) = M/G$$

is a manifold because the second is a manifold.

If f and g are functions on $\mathfrak{su}(n)\times\mathfrak{su}(n)$ which are invariant under $SU(n)$ and Poisson commute they define Poisson commuting functions on M/G.

Consider the Hamiltonian

$$H(Q, P) = -\frac{1}{2} \operatorname{tr} P^2$$

which is invariant under the action of $SU(n)$ with the symplectic form

$$\sigma((A_1, B_1), (A_2, B_2)) = -\operatorname{tr} A_1 B_1 + \operatorname{tr} A_2 B_2.$$

It defines a function $\tilde{H}$ on $\Phi^{-1}(M)/G_M$,

$$\tilde{H} = \frac{1}{2}\sum_{j=1}^{n} p_{jj}^2 + \sum_{j>k} \frac{1}{(q_j - q_k)^2}.$$

First integrals are $\operatorname{tr} P^k$, $2 \leq k \leq n$.

There is a very simple fairly explicit formula for orbits. The Hamiltonian $H(P,Q) = -\frac{1}{2} \operatorname{tr} P^2$ defines the evolution

$$Q(t) = Q(0) + tP(0), \qquad P(t) = P(0).$$

We obtain the orbit by choosing initial data in $\mathcal{A}$ and projecting to $\tilde{M}/\operatorname{SU}(n)$.

This can be generalized to general simple Lie groups, to loop groups (which I don't define) for the Hamiltonian

$$H(p,q) = \frac{1}{2}|p|^2 + \sum_{j \neq k} \frac{a^2}{\sin^2(a(q_j - q_k))}$$

$$H(p,q) = \frac{1}{2}|p|^2 + \sum_{j \neq k} \frac{a^2}{\sinh^2(a(q_j - q_k))}$$

$$H(p,q) = \frac{1}{2}|p|^2 + \sum_{j \neq k} \frac{a^2}{\mathcal{P}(a(q_j - q_k))}$$

where $\mathcal{P}$ is teh Weierstrass $\mathcal{P}$ function with periods 1 and τ.

3.8 Jacobi operators and the Toda latice

We consider the Toda differential equation with Hamiltonian

$$H = \frac{1}{2} \sum_{k=1}^{N+1} p_k^2 + \sum_{k=1}^{N} \exp(q_k - q_{k+1}).$$

The Toda lattice is again related to simple Lie group. There are variants for other simple Lie groups and centrally extented loup groups, where the Toda differential equation becomes a periodic equation. There are semi infinite and infinite variants. The purpose of this section is the introduction of a general structure leading to integrable systems with strong similarities to the Korteweg-de Vries equation.

The Hamiltonian equations are

$$\dot{q}_k = p_k, \quad \dot{q}_k = e^{q_{k-1} - q_k} - e^{q_k - q_{k+1}}.$$

(using $e^{q_0 - q_1} = e^{q_{N+1} - q_{N+2}} = 0$)

Flaschka and Manakov introduced the variables

$$a_k = -p_k/2, b_k = \frac{1}{2} \exp((q_k - q_{k+1})/2)$$

so that the differential equation becomes equivalent to the Lax pair

$$\frac{d}{dt}L = [B, L]$$

where

$$L = \begin{pmatrix} a_1 & b_1 & 0 & \dots \\ b_1 & a_2 & b_3 & \dots \\ 0 & b_2 & a_3 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{3.21}$$

$$B = \begin{pmatrix} 0 & b_1 & 0 & \dots \\ -b_1 & 0 & b_2 & \dots \\ 0 & -b_2 & 0 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{3.22}$$

Definition 3.22. *A Jacobi matrix is a real matrix of the form*

$$T = \begin{pmatrix} a_0 & b_0 & 0 & \dots & 0 & \\ b_0 & a_1 & b_1 & \dots & 0 & 0 \\ 0 & b_1 & a_2 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \\ 0 & 0 & 0 & \dots & a_{n-1} & b_{n-1} \\ 0 & 0 & 0 & \dots & b_{n-1} & a_n \end{pmatrix}$$

with positive off-diagonal entries. We also consider the (semi-)infinite dimensional case where we assume in addition $\sup_j |a_j|, \sup_j |b_j| \leq c < \infty$.

A direct calculation give the Lax equations.

Jacobi matrices occur in the theory of orthogonal polynomials: Let μ be a compactly supported probility measure. Recursing orthogonalisation leads to orthogonal monic polynomials

$$P_n = x^n + \sum_{j=0}^{n-1} c_j x^j.$$

The eigen values of a Jacobi matrix are all real and distinct, a fact which we will discuss on Thursday.

As usual the symmetry of the problem can be better understood for Lie groups and Lie algebras.

3.8.1 The Lie group $SL(N+1)$

The Lie group SL$(N+1)$ consists of all real $(N+1)\times(N+1)$ matrices of determinant 1. Its Lie algebra $\mathfrak{sl}(N+1)$ consists of the matrices with trace 0. The Lie algebra is simple: The only ideals in $\mathfrak{sl}(N+1)$ are $\mathfrak{sl}(N+1)$ and $\{0\}$. It has a nondegenerate bilinear form

$$\langle A, B\rangle = \operatorname{tr} AB$$

(no negative sign) which allows to identity $\mathfrak{sl}^*(N+1)$ and $\mathfrak{sl}(N+1)$.

More generally, let $\mathfrak{g}$ be a finite dimensional Lie algebra. Suppose that the linear map $R : \mathfrak{g} \to \mathfrak{g}$ satisfies the modified Yang Baxter equation

$$[RX, RX] - R([X, RY]) + [RX, Y]) = -\frac{1}{4}[X, Y] \tag{3.23}$$

where the factor $\frac{1}{4}$ is convention. An example is a multiple of the identity. Assume that R satisfies (3.23). Then

$$[X, Y]_R = [RX, Y] + [X, RY] \tag{3.24}$$

is a second Lie algebra structure:

$$[X, Y]_R + [Y, X]_R = [RX, Y] + [X, RY] + [RY, X] + [Y, RX] = 0$$

$$\begin{aligned}
&[X, [Y, Z]_R]_R + [Y, [Z, X]_R]_R + [Z, [X, Y]_R]_R \\
&\quad = [X, [RY, Z] + [Y, RZ]]_R + [Y, [RZ, X] + [Z, RX]]_R \\
&\qquad + [Z, [RX, Y] + [X, RY]]_R \\
&\quad = [RX, [RY, Z]] + [X, R[RY, Z]] + [RX, [Y, RZ]] + [X, R[Y, RZ]] \\
&\quad + [RY, [RZ, X]] + [Y, R[RZ, X]] + [RY, [Z, RX]] + [Y, R[Z, RX]] \\
&\quad + [RZ, [RX, Y]] + [Z, R[RX, Y]] + [RZ, [X, RY]] + [Z, R[X, RY]] \\
&\quad = [X, [RY, RZ] - R[[Y, Z]_R] + [Y, [RZ, RX] - R[[Z, X]_R] \\
&\qquad + [Z, [RX, RY] - R[[X, Y]_R] \\
&\quad = \frac{1}{4}([X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]]
\end{aligned}$$

Proposition 3.3. *Suppose that $\mathfrak{A}$ and $\mathfrak{B}$ are Lie subalgebras of a Lie algebra $\mathfrak{g}$ and as a vector space*

$$\mathfrak{g} = \mathfrak{A} + \mathfrak{B}$$

and let $P_{\mathfrak{A}}$ be the projection to $\mathfrak{A}$ along $\mathfrak{B}$ and similarly $P_{\mathfrak{B}}$ the projection along $\mathfrak{A}$. Then $R = \frac{1}{2}(P_{\mathfrak{A}} - P_{\mathfrak{B}})$ satisfies the modified Yang-Baxter equation.

Suppose that R satisfies the modified Yang-Baxter equation and denote the two Lie brackets by $[.,.]$ and $[.,.]_R$. We obtain the codajoint representation for $L \in \mathfrak{g}^*$

$$\operatorname{ad}_X^* L(Y) = -L([X,Y]), \quad \operatorname{ad}_{R,X}^* Y = -L([X,Y]_R) \tag{3.25}$$

Theorem 3.23. *The Casimirs on $\mathfrak{g}^*$ Poisson commute for $[.,.]_R$. If H is a Casimir then the Hamiltonian equations with respect to $[.,.]_R$ can be written as*

$$\frac{d}{dt}L = -\operatorname{ad}_{R,dH}^* L$$
$$\frac{d}{dt}L = \operatorname{ad}_{-R(dH)}^* L$$

Proof. Let H_1, H_2 be Casimirs. Then

$$\begin{aligned}\{H_1, H_2\}(L) &= L([dH_1, dH_2]_R) \\ &= L([RdH_1, dH_2] + [dH_1, RdH_2]) \\ &= \operatorname{ad}_{dH_2}^* L(RdH_1) - \operatorname{ad}_{dH_1}^* LRdH_2 \\ &= 0.\end{aligned}$$

$$\{H, f\}_R(L) = L([dH, df]_R) = -\operatorname{ad}_{R,dH}^* Ldf.$$

□

In the case of an invariant non-degenerate quadratic form on $\mathfrak{g}$ we can identify $\mathfrak{g}^*$ and $\mathfrak{g}$ and the equations take the form

$$\frac{d}{dt}L = [M, L], \qquad M = -R(\nabla H).$$

We obtain an algorithmus for solving these equations. We rewrite the equations as

$$\frac{d}{dt}L = [M_+, L] = [M_-, L] \quad \text{with } M_\pm = -R^\pm dH \tag{3.26}$$

We specialize to $\mathrm{SL}(N+1)$. There is the decomposition

$$\mathfrak{sl}(N+1) = N_- \times H \times N_+$$

where N_- consists of the strict lower triangular matrices, N_+ of the strict upper triangular matrices, and H are the traceless diagonal matrices. We define

$$R(h) = 0, R(n_-) = -\frac{1}{2}n_-, R(n_+) = \frac{1}{2}n_+$$

for $h \in H$, $n_\pm \in N_\pm$. It satisfies the modified Yang-Baxter equation

$$[RX, RY] - R([X,Y]_R) = -\frac{1}{4}[X,Y]$$

which we check case by case

1. If $X, Y \in H$ all terms vanish.
2. If $X, Y \in N_\pm$ then

$$[RX, RY] - R([X, RY] + [RX, Y]) = -\frac{1}{4}[X,Y]$$

3. If $X \in N_+$, $Y \in H$ then

$$[RX, RY] - R([X,Y]_R) = -R[RX, Y] = -\frac{1}{4}[X,Y]$$

4. $X \in N_+$, $Y \in N_-$

$$\begin{aligned}[RX, RY] - R([RX, Y] + [X, RY] &= -\frac{1}{4}[X,Y] - \frac{1}{2}R([X,Y] - [X,Y]) \\ &= -\frac{1}{4}[X,Y].\end{aligned}$$

We define linear map $R_\pm$ by $R \pm \frac{1}{2}1$,

$$R_+((n_-, h, n_+)) = (0, \frac{1}{2}h, n_+)$$

and

$$R_-(n_-, h, n_+)) = (-n_-, -\frac{1}{2}h, 0).$$

Let $G_\pm$ the group of upper/lower triangular matrices with determinant 1.

Proposition 3.4.

$$L(t) = \theta_+(t)L(0)\theta_+^{-1}(t) \tag{3.27}$$

where

$$\exp(-tdH(L(0)) = \theta_-^{-1}(t)\theta_+(t). \tag{3.28}$$

$$\theta_\pm \in G_\pm$$

and the diagonals of θ_+ *and* θ_- *are inverses.*

Proof. We want to solve

$$\frac{d}{dt}L = [M_\pm, L] \qquad \text{with } M_\pm = -R^\pm dH(L)$$

Since $M_\pm \in \mathfrak{g}_\pm$ we can solve

$$\dot\theta_\pm = M_\pm \theta_\pm \qquad \text{with } \theta_\pm(0) = 1.$$

Then

$$L(t) = \theta_+(t)L(0)\theta_+^{-1}(t) = \theta_-(t)L(0)\theta_-^{-1}(t).$$

Since $M_+ - M_- = -dH(L(t))$ we get

$$dH(L(t)) = -\dot\theta_+\theta_+^{-1} + \dot\theta_-\theta_-^{-1}(t) = -\theta_-\left(\frac{d}{dt}(\theta_-^{-1}(t)\theta_+(t))\right)\theta_+^{-1}$$

We recall that H is a Casimir. Hence

$$dH(\theta_- L(0)\theta_-^{-1}) = \theta_- dH(L(0)\theta_-^{-1}$$

and

$$\frac{d}{dt}(\theta_-^{-1}\theta_+)(\theta_-^{-1}\theta_+)^{-1} = -dH(L(0))$$

hence

$$\theta_-^{-1}(t)\theta_+(t) = \exp(-tdH(L(0)))$$

There is a unique solution to this factorization problem for small t since there is a unique decomposition at the level of the Lie algebra. □

3.8.2 Jacobi matrices and orthogonal polynomials

We will obtain a much more explicit expression for the Toda flow, which we relate to this section. The strongest results are due to Kostant, [17]. Our presentation follows Deift [6] and Moser [23] contains much more interesting and reasonably elementary material. The survey von van Moerbecke [19] puts this part into a far reaching and much more demanding algbraic context.

Let H be a Hilbert space, T a bounded self adjoint operator on H, $x_0 \in H$, $\|x_0\| = 1$. Let recursively

$$\hat x_{n+1} = T\hat x_n$$

and assume that the $\hat x_n$ are linearly independent. x_0 is called cyclic if the closure of the span of the $\hat x_n$ is H. We apply the Gram-Schmidt procedure to obtain an orthogonal sequence x_n,

$$\langle x_n, x_m\rangle = \delta_{nm}$$

Clearly we can write

$$Tx_n = \sum_{j=0}^{n+1} c_j x_j.$$

We observe that

$$\langle Tx_n, x_m \rangle = \langle x_n, Tx_m \rangle = 0$$

unless $|n-m| \leq 1$. Let $a_k = \langle Tx_k, x_k \rangle$ and $b_k = \langle Tx_k, x_{k-1} \rangle$ for $k \geq 1$ $(b_0 = 0)$. Then

$$Tx_k = b_{k+1}x_{k+1} + a_k x_k + b_k x_{k-1} \tag{3.29}$$

and T has a representation as Jacobi matrix.

Lemma 3.24. *Let T be a $(N+1) \times (N+1)$ Jacobi matrix. Then T has $N+1$ simple real eigenvalues.*

Proof. We claim that if $Tx = \lambda x$, $x \neq 0$ then $x_0 \neq 0$ and $x_N \neq 0$. It is not hard to see that otherwise x is trivial. Now suppse $Tx = \lambda x$ and $Ty = \lambda y$. If $a_1 x_0 + a_2 y_0 = 0$ then

$$a_1 x + a_2 y = 0$$

and hence the two vectors are linearly independent. $\square$

Let $\tilde{x}_n$ be linearly independent, and x_0 cylic. The spectral theorem states that there is a probability measure μ with bounded support on $\mathbb{R}$ so that the map

$$H \ni x_n \to t^n \in L^2(\mu)$$

is an isometry. The Jacobi flow becomes a flow on the probability measures

$$\mu(t) = \frac{1}{\int e^{ty} d\mu(y)} e^{tx} \mu.$$

A particular case is if

$$\mu = \sum_{j=0}^{N} \beta_j^2 \delta_{x_j}$$

so that $L^2(\mu)$ is a vector space of dimension $N+1$. In this case we consider

$$Tf = tf$$

and we start with $x_0 = 1$ and $\tilde{x}_n = t^n$.

Let $|x_0| = 1$ and x_n be the orthogonal bases constructed through $T^n x_0$. Let

$$G(z) := \langle x_0, (T-z)^{-1} x_0 \rangle$$

where we chose the convention that the inner product is complex linear in the second compenent. We compute for $\operatorname{Im} z > 0$

$$\begin{aligned} \operatorname{Im} G(z) &= \frac{1}{2i}\Big(\langle x_0, (T-z)^{-1} x_0\rangle - \langle x_0, (T-\bar z)^{-1} x_0 \rangle \Big) \\ &= \langle x_0, \operatorname{Im} z (T-z)^{-1}(z-T)^{-1} x_0\rangle > 0. \end{aligned}$$

We diagonalize T and obtain

$$G(z) = \sum_{j=0}^N \beta_j^2 \frac{1}{z_j - z}$$

for some $\beta_j \ge 0$ and

$$1 = \lim_{y\to\infty} (x+iy) G(x-iy) = \sum \beta_j^2$$

Theorem 3.25. *Let $\mathcal{T}(\mathbb{R}^{N-1})$ be the set of Jacobi operators. Then*

$$\mathcal{T}(\mathbb{R}^{N+1}) \ni T \to (z,\beta) \in \{ z_0 < z_1 < \dots z_N, 0 < \beta_j < 1\}$$

is a diffeomorphismus. The Toda flow is equivalent to

$$\frac{d}{dt} z_j = 0, \qquad \frac{d}{dt}\beta_j = -(z_j - \sum_{k=0}^N z_k r_k^2)\beta_j.$$

We begin with describing the inverse of the map. Let μ be a compactly supported probablitiy measure (a sum on $N+1$ Diracs in our case) and we define

Definition 3.26. *The moment matrix is for $n \le N$*

$$M_n = \Big(\int t^{i+j} d\mu \Big)_{0 \le i,j \le n}$$

and

$$D_n = \det M_n$$

$$D_n(x) = \det \begin{pmatrix} \int t^0 d\mu & \int t d\mu & \dots & \int t^n d\mu \\ \int t d\mu & \int t^2 d\mu & \dots & \int t^{n+1} d\mu \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x & \dots & x^n \end{pmatrix}.$$

We observe

$$0 \leq \| \sum_{j=0}^{n} s_j t^j \|^2_{L^2(\mu)} = (s, Ms)$$

and hence M is positive definite for $n \leq N$ and $D_n > 0$. We claim the the n th orthogonal polynomial are

$$p_n(x) = \frac{1}{\sqrt{D_{n-1} D_n}} D_n(x).$$

First

$$\int x^j D_n(x) dx = 0$$

if $j < n$ since two rows of the marix coincide and similarly

$$\int x^n D_n(x) d\mu = D_n.$$

Moreover

$$D_n(x) = D_{n-1} x^n + \dots$$

and we complete the argument with

$$\int D_n(x)^2 dx = D_{n-1} \int x^n D_n(x) dx = D_n D_{n-1}.$$

We compare to the three term recursion (3.29) and find the formulary

$$b_n^2 = \frac{D_{n-1} D_{n+1}}{D_n^2}.$$

With some more effort

$$a_n = \partial_t \log(D_{n+1}/D_n).$$

This completes the contruction of an explicit inverse.

By the Lax equation the eigenvalues of T are independent of time and it remains to deduce the differential equations for β_j.

Let $A_k(z)$ be the lower right $k \times k$ submatrix of $z - T$ and $\Delta_k(z) = \det(A_k(z))$. We expand the first column to obtain a recursion formula

$$\Delta_k = (z - a_{N+1-k}) \Delta_{k-1} - b^2_{N+2-k} \Delta_{k-2} \tag{3.30}$$

with the obvious modification for $k = 1, 2$. We claim that

$$G(z) = \frac{\Delta_N(z)}{\Delta_{N+1}(z)} \tag{3.31}$$

We determine the componente of $(z-T)^{-1}e_0$:

$$((z-T)^{-1}x_0)_k = \begin{cases} \frac{\Delta_N}{\Delta_{N+1}} & \text{if } k=0 \\ \frac{\Delta_{N-k}}{\Delta_{N+1}} b_1 \dots b_k & \text{if } k=1,\dots,N \end{cases}$$

using the recursion formula for $w=(z-T)^{-1}e_0$ for $1 \le j \le N-1$

$$b_j w_{j-1} - (z-a_j)w_j + b_{j+1}w_{j+1} = 0$$

and

$$(\lambda - a_0)w_0 - b_1 w_1 = 1$$

and comparing it to (3.30).

We compute the time derivative of β_n. Let $R=(z-T)^{-1}$. Then

$$\frac{d}{dt}R = R\frac{d}{dt}TR = R(BT-TB)R = RB - BR$$

and

$$\frac{d}{dt}G(z) = \langle e_0, (RB-BR)e_0 \rangle = 2b_1 R_{01} = 2b_1^2 \frac{\Delta_{N-1}}{\Delta_{N+1}}.$$

Clearly

$$\frac{d}{dt}G(z) = \sum_{k=1}^{n} \frac{\beta_k \dot\beta_k}{z-z_k}.$$

We compare the residue at $z=z_k$ in both expressions using again the recursion formula (3.30) and $\Delta_{N+1}(z_k)=0$

$$\Delta_{N-1}(z_k) = \frac{z_k - a_0}{b_1^2}\Delta_N(z_k)$$

and the residue is the same as the one of

$$2(z-a_0)\frac{\Delta_N}{\Delta_{N+1}}.$$

and

$$2\beta_k\dot\beta_k = 2(z_k - a_0)r_k^2.$$

Since

$$0 = \sum_k \beta_k \dot\beta_k = \sum z_k \beta_k^2 - a_0$$

we obtain

$$a_0 = \sum z_k \beta_k^2.$$

This completes the proof.

4 The Korteweg-de Vries equation

A large part of this section is motivated by Killip and Visan [16]. There is a huge literature on the KdV equation. The original papers are still of interest. An interesting perspective is Segal's contribution in [10].

4.1 The Schrödinger operator

4.1.1 Sobolev spaces

We denote the Sobolev space

$$H^k(\mathbb{R}) = \{f \in L^2 : \partial_x^j f \in L^2 \text{ for } j \le k\}$$

with norm (for $\tau > 0$)

$$\|f\|^2_{H^k_\tau} = \sum_{j=0}^k \tau^{2(k-j)} \|f^{(j)}\|^2_{L^2} = \|(\tau^2 + |\xi|^2)^{k/2} \hat f\|^2_{L^2}.$$

The second equality is a consequence of the theorem of Plancherel for the Fourier transform. We use it to define the norm (and the space) for $k \in \mathbb{R}$. The dual space is

$$(H^k_\tau)^* = H^{-k}(\mathbb{R})$$

with the formal duality map

$$H^k \times H^{-k} \ni (f,g) \to \int fg dx.$$

The fundamantal theorem of calculus yields the Sobolev esimate

$$\|f\|^2_{L^\infty} \le |f|_{L^2} \|f_x\|_{L^2} \le \frac{1}{2}\|f\|^2_{H^1}$$

by

$$|f(0)|^2 \le 2\int_0^\infty |f f_x| dx \le 2\|f\|_{L^2(0,\infty)} \|f_x\|_{L^2(0,\infty)}$$

by taking the smaller value on $(-\infty, 0)$ resp $(0,\infty)$. Even more is true: Functions on H^1 are Hölder continuous with exponent $\frac{1}{2}$. We obtain the embeddings (with $C_b(\mathbb{R})$ the space of bounded continuous functions)

$$H^1 \subset C_b(\mathbb{R}) \qquad L^1(\mathbb{R}) \subset H^{-1}(\mathbb{R}).$$

As a consequence H^1 is an algebra,

$$\|fg\|^2_{H^1} = \|fg\|^2_{L^2} + \|\partial_x(fg)\|^2_{L^2} \le \|f\|_{L^\infty}(\|g\|_{L^2} + \|g_x\|_{L^2}) + \|f_x\|_{L^2}\|g\|_{L^\infty} \le 2\|f\|_{H^1}\|g\|_{H^1}$$

By duality

$$\|fg\|_{H^{-1}} \le 2\|f\|_{H^1}\|g\|_{H^{-1}}.$$

The map

$$H^1_\tau \ni f \to \pm f_x + \tau f \in L^2$$

is an isometry which is seen by the Fourier transform. It is invertible with

$$(\pm\partial + \tau)^{-1} f = \left\{ \begin{array}{cc} -e^{-\tau x} & \text{if } + \\ -e^{\tau x} & \text{if } - \end{array} \right\} * f.$$

By duality

$$L^2 \ni f \to \mp f_x + \tau f \in H^{-1}_\tau$$

also an isometric isomorphism and we can write $f \in H^{-1}_\tau$ in the form

$$f = \mp g_x + \tau g$$

with $g \in L^2$ and $\|g\|_{L^2} = \|f\|_{H^{-1}_\tau}$.

4.1.2 Definition of the Schrödinger operator

The key object is the Schrödinger operator

$$L\phi = (-\partial^2 + u)\phi$$

where $u \in H^{-1}(\mathbb{R})$. The operator L defines formally a symmetric quadratic form on H^1

$$B_u(\phi, \psi) = \int \phi_x \psi_x + u\phi\psi dx. \tag{4.1}$$

which we define first and understand L as an operator

$$L : H^1 \to H^{-1}.$$

We claim that the quadratic form of $L + \tau^2 : H^1 \to H^{-1}$ has an inverse which defines a bounded self adjoint map $L^2 \to H^2 \subset L^2$ is τ is sufficiently large. The first claim is a consequence of the lemma of Lax-Milgram and the second follows from the calculation below. Suppose that $u = v_x + \tau v$. Then, with $B_\tau(\phi) = B(\phi, \phi) + \tau^2 \|\phi\|^2_{L^2}$,

$$B_\tau(\phi, \phi) = \|\phi_x^2\|^2_{L^2} + \tau^2\|\phi\|^2_{L^2} + \int (v_x + \tau v)\phi^2 dx$$

and

$$\begin{aligned}\int(\partial_x v+\tau v)\phi^2 dx+\int h\phi dx\\ &\le 2\|v\|_{L^2}\|\phi\|_{L^\infty}\|\phi\|_{H^1_\tau}+\|h\|_{L^2}\|\phi\|_{L^2}\\ &\le 2\tau^{-1/2}\|u\|_{H^{-1}_\tau}\|\phi\|^{\frac12}_{L^2}\|\phi\|^{\frac32}_{H^1_\tau}+\|h\|^2_{L^2}+\frac14\|\phi\|^2_{L^2}\\ &\le \left(\tau^{-1/2}\|u\|_{H^{-1}}\right)^4\|\phi\|^2_{L^2}+\|h\|^2_{L^2}+\frac34\|\phi\|^2_{H^1}\end{aligned}$$

hence, if τ is sufficiently large the form B_τ is coercive and

$$L+\tau^2 H^1\to H^{-1}$$

is invertible. As a byproduct

$$\|\phi\|_{H^1}\le c(\|u\|_{H^{-1}})\|\phi\|_{L^2}+\|L\phi\|_{H^{-1}}).$$

It is not hard to deduce more regularity if u and $L\phi$ are more regular.

Lemma 4.1.

$$\|\phi\|_{H^2}\le c(1+\|u\|_{L^2})\|\phi\|_{L^2}+\|L\phi\|_{L^2}. \tag{4.2}$$

4.1.3 Eigenvalues and Eigenfunctions

Lemma 4.2. *Let $u\in H^{-1}$ and $\tau>0$. Then there is at most a finite number of eigen values below $-\tau^2$.*

Proof. We will verify that there is a subspace V of H^1 of finite codimension so that the quadratic form B_τ is nonnegative in it. This implies the claim. We write

$$L+\tau^2=(\partial_x+\tau)(1+(\partial+\tau)^{-1}u(-\partial+\tau)^{-1})(-\partial_x+\tau).$$

It suffices to find a subspace of finite codimension in L^2 so that the inner bracket is nonnegative on it. This follows once we prove that the inner operator is identity plus compact. We even prove that

$$(\partial+\tau)^{-1}u(-\partial+\tau)^{-1}$$

it is Hilbert-Schmidt (i.e. it integral kernel is square integrable. We write $u=v_x+\tau v$. The operator becomes

$$v(-\partial+\tau)^{-1}$$

with integral kernel

$$k(x,y) = -v(x)\chi_{x>y}\exp(\tau(y-x))$$

with (by Fubini)

$$\|k\|_{L^2(\mathbb{R}\times\mathbb{R})} = \frac{1}{\sqrt{2\tau}}\|v\|_{L^2}$$

□

As a consequence negative eigen functions can only accumulate at 0. There may of may not be negative eigenvalues. If there are there is a lowest eigen value, called the ground state energy with an eigenfunction called the ground state. It is the minimizer of

$$\frac{B(\phi,\phi)}{\|\phi\|_{L^2}^2}.$$

This function is bounded from below by the arguments for Lax-Milgram. Let $\phi \in H^1$ with $\|\phi\|_{L^2} = 01$ be a minimizing sequence. It is not hard to see that there is a converging subsequence, and hence there exists a minimizer ϕ_0. Then also $|\phi_0|$ is a minimizer which does not change sign.

4.1.4 The Sturm oscillation argument

Let $\lambda \in \mathbb{C}\backslash[0,\infty)$ and suppose that $\phi \in H^1$ is an eigen function

$$L\phi = \lambda\phi.$$

Then $\lambda \in (-\infty, 0)$ and $\phi \in H^2 \subset C^1(\mathbb{R})$.

In the sequel we neglect the regularity of the potential u and other functions under consideration for some time and pretent that it is always sufficiently smooth. Other the arguments remain essentially the same, but they are more technical.

Lemma 4.3. *Suppose that $\phi, \psi \in H^1(I;\mathbb{R})$, $\psi(a) = \psi(b) = 0$. We assume*

$$\frac{L\psi}{\psi} \le \frac{L\phi}{\phi}$$

whenever $\phi \neq 0$ resp. $\psi \neq 0$. Then either ϕ has a zero in (a,b) or ϕ and ψ are linearly dependent.

Proof. Without loss of generality (restricting the interval) we may assume that $\psi > 0$ in the interior and, by arguing by constradiction, also that $\phi > 0$ in the interior.

Let

$$W = \psi'\phi - \psi\phi'.$$

Then

$$\frac{d}{dx} W = \psi''\phi - \psi\phi'' \geq \psi\phi \geq 0.$$

Since $\psi(a) = 0$, $\psi'(a) \geq 0$ and $\phi(a) \geq 0$ we see that $W(a) \geq 0$ and $W(b) \leq 0$, which can only be true if $W = 0$ and the functions are linearly dependent. This argument also works in the case $b = \infty$.

$\square$

As a consequence we obtain

Theorem 4.4. *If ϕ_1 and ϕ_2 are eigenfunctions to the eigenvalues $-\tau_1^2$ and $-\tau_2^2$ with $\tau_2 < \tau_1$ then there is a zero of ϕ_2 between two zeros of ϕ_1 (by an abuse of notation we allow semiinfinite nodal intervals.) The negative eigenvalues are simple. The ground state, the lowest eigenfunction has no zero. If we order the eigenvalues*

$$-\tau_0^2 < -\tau_1^2 < \cdots < -\tau_N^2 < 0$$

then the eigenfunctions to $-\tau_n^2$ have exactly n zeros.

Proof. The first statement, the interlacing follows from Lemma 4.3. Let ϕ and ψ be eigenfunctions to the eigenvalue $-\tau^2$. By the interlacing property they are either linearly dependent, or the zeros are interlaced, but this is not possible since the number of nodal intervals has to be the same.

We turn to the ground state. It is the minimizer of

$$\int \phi_x^2 + u\phi^2 dx, \qquad \text{under the constraint } \|\phi\|_{L^2}^2$$

Using a minimizing sequence one can prove existence of a minimizer unless the functional is nonnegative. If ϕ is a minimizer then also $|\phi|$ is a minimzer and we may assume that ϕ is nonnegative. It satisfies the Euler Lagrange equation with the Lagrangian multiplier λ

$$-\phi'' + u\phi = \lambda\phi.$$

Multiplication by ϕ and integration shows that λ is the ground state energy.

Let $\mathcal{V}(n)$ the set of n dimensional subspaces of H^1. Then

$$-\tau_n^2 = \inf_{V\in\mathcal{V}(n)} \sup_{v\in V} \frac{B(v)}{|v|^2}$$

assuming that this number is negative.

We complain that the nth eigenfunction ϕ has exactly $n+1$ nodal intervals. By interlacing it has at least $n+1$ nodal intervals. Suppose ϕ has $m > n+1$ nodal intervals I_k $0 \le k \le m$. Then we find constants c_j with $c_0 = 0$ so that

$$\tilde\phi = \sum c_j \chi_{I_k} \phi$$

is orthogonal to all the previous eigenfunctions and has norm 1. This contradicts the simplicity of the eigenvalue.

□

4.1.5 Bounds on the number of eigenvalues

Lemma 4.5. *Let $a \in \mathbb{R}$, $b \in (a,\infty]$, $a < b$ and $\tau \ge 0$. Suppose that $\phi \in H^1(a,b)$ satisfies $\phi \ge 0$, $\phi(a) = \phi(b) = 0$ (if $b = \infty$, we assume $\phi(a) = 0$ and $\tau > 0$) and*

$$-\phi'' + (u+\tau^2)\phi = 0 \qquad \text{on } (a,b).$$

Then

$$\int_a^b (x-a)u_- dx > 1. \tag{4.3}$$

Proof. We have seen that ϕ is the ground state and it minimizes

$$\frac{\int \phi_x^2 + u\phi^2 + \tau^2\phi^2 dx}{\|\phi\|_{L^2}^2}$$

in a suitable function spaces. This minimum is a continuous strictly monotonically decreasing function of b (including $b \to \infty$ if $\tau > 0$) which tends to ∞ as $b \to a$. If we replace u by $-u_-$ we decrease the functional and by decreasing b we may assume that $u = -u_-$ and $\phi'(a) > 0$. Then $\phi'' \le 0$

$$\phi'(a) = 1 + \int_a^x \phi''(y)dy = 1 + \int_a^x u\phi dy \ge 1 - \int_a^x u_-(y-a)dy$$

which leads to a contradiction unless (4.3) holds. □

Theorem 4.6. *Suppose that*

$$\int u_-|x|dx \le N$$

Then L has at most $N+1$ negative eigen values.

Proof. Suppose that the eigenfunction ϕ has M zeros. It then has $M+1$ nodal intervals. On M of them we apply the previous lemma. Then

$$\int u_-|x|dx > M$$

$\square$

Theorem 4.7. *Suppose that*

$$u \in L^1 \subset H^{-1}.$$

then there is no positive eigenvalue.

Proof. Suppose that

$$-\phi_{xx} + u\phi = \tau^2\phi$$

Let $g(x,y)$ be the Green's function of $-\partial^2 - \tau^2$ supported in $x \le y$.

$$g(x,y) = \begin{cases} 0 & \text{if } x > y \\ \frac{\sin(\tau(x-y))}{\tau} & \end{cases}$$

Then

$$\phi(x) = \int_x^\infty g(x,y)u(y)\phi(y)dy$$

hence

$$\begin{aligned} &\|\phi\|_{L^\infty(x,\infty)} \\ &\qquad \le \int_x^\infty |g(x,y)u(y)|dy\|\phi\|_{L^\infty(x,\infty)} \\ &\qquad \le \sup \frac{1}{\tau}\int_x^\infty |u(y)|dy\|\phi\|_{L^\infty(x,\infty)}. \end{aligned}$$

We choose x large so that $\|u\|_{L^\infty(x,\infty)}$ is small and see that ϕ vanishes for large arguments. We solve the Cauchy problem for the ODE starting with large values for x and obtain that $\phi = 0$. $\square$

Example: Let $\kappa > 0$. Then

$$(-\partial^2 - 2\kappa\delta_0)e^{-\kappa|x|} = -\kappa^2 e^{-\kappa|x|}$$

4.2 Jost solutions and the Miura map

Let $\operatorname{Im} z > 0$. We study solutions to

$$-\phi_{xx} + (-2izf + f_x)\phi = z^2\phi.$$

which we rewrite as system with

$$\psi := -\phi_x + f\phi - iz\phi$$

$$\phi_x = -\psi + f\phi - iz\phi$$

$$\begin{aligned}\psi_x &= -\phi_{xx} + f_x\phi + f\phi_x - iz\phi_x \\ &= 2izf\phi + z^2\phi - iz\phi_x - f\psi - izf\phi + f^2\phi \\ &= (iz - f)\psi + f^2\phi\end{aligned}$$

Let

$$\Phi = e^{izx - \int_0^x f\,dy} \left(\begin{array}{c} \phi \\ \psi \end{array} \right)$$

Then

$$\Phi_x = \begin{pmatrix} 0 & -1 \\ f^2 & 2iz - 2f \end{pmatrix} \Phi. \tag{4.4}$$

We search

$$\Phi = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \tilde{\Phi}.$$

The equation

$$\psi_x = 2(iz - f)\psi + f^2\phi$$

can be solved for ψ, using $\psi(x) \to 0$ as $x \to -\infty$,

$$\psi(x) = \int_{-\infty} \exp\left(2iz(x-y) - 2\int_y^x f\,dt\right) f^2\phi(y)dy$$

and ϕ satisfies

$$\phi(\tilde{x}) = S\phi(\tilde{x}) := \int_{y<x<\tilde{x}} \exp\left(2iz(x-y) - 2\int_y^x f\,dt\right) f^2\phi(y)dy.$$

Lemma 4.8.

$$\|S\|_{C_b -> C_b} \leq \exp\left((\operatorname{Im} z)^{-1/2}\|f\|_{L^2}\right)(\operatorname{Im} z)^{-1}\|f\|_{L^2}^2$$

Proof. The norm is given by

$$\sup_{\tilde{x}} \|k(\tilde{x}, y)\|_{L^1(-\infty,x)}$$

where

$$k(\tilde{x}, y) = \int_y^{\tilde{x}} \exp(2iz(x-y) - 2\int_y^x f dt) f^2(y) dx$$

and

$$2iz(x-y) - 2\int_y^x f dt \le -2 \operatorname{Im} z|x-y| + 2|x-y|^{\frac{1}{2}}\|f\|_{L^2} \le -\operatorname{Im} z|x-y| + (\operatorname{Im} z)^{-1}\|f\|_{L^2}^2.$$

and, with $C = \exp\Big(\operatorname{Im} z^{-1}\|f\|_{L^2}^2\Big)$

$$|k(\tilde{x}, y)| \le C \int_y^{\tilde{x}} \exp(-\operatorname{Im} z(x-y)) dx f^2(y) \le C(\operatorname{Im} z)^{-1} f^2(y)$$

□

Assuming that $\|f\|_{L^2} \le \operatorname{Im} z^{-1}$ we obtain a expansion

$$\phi(x) = \sum_{n=0}^{\infty} \phi_{2n}(x)$$

with

$$\phi_{2n}(\tilde{x}) = \int_{y_1<x_1<y_2<x_2\cdots<\tilde{x}} \prod k(x_j, y_j) dy_x dx_j$$

and

$$\phi_{2n} \le (C \operatorname{Im} z^{-1}\|f\|_{L^2}^2)^n.$$

There are a number of consequences. The Jost solution satisfies

$$\phi_l \sim e^{-izx + \int_0^x f(y)dy}$$

near $-\infty$ for all $\operatorname{Im} z > 0$, since we can always argue on an interval of the type $(-\infty, a)$ on which the $|f\|_{L^2}$ norm is small. It depends holomophically on z and smootly on f.

A calculation shows that

$$\phi = \phi_l \int_b^x \phi_l^{-2} dx$$

is a second solution, which grows exponentially near $-\infty$. We obtain a basis for space of solutions, similarly on the right hand side.

We obtain a representation

$$\phi_l = c_1\phi_r + c_2\phi_r \int_a^x \phi_r^{-2} dy$$

and we define

$$a(z) = \lim_{x\to\infty} e^{izx - \int_0^x f(y)dy} \phi_l(x)$$

which is holomorphic. It vanishes exactly at the eigenvalues.

We define for either $\operatorname{Im} z > 0$ or, $\operatorname{Re} z = 0$ and z^2 is below the ground state the function w

$$w = \partial_x \log \phi_l + iz.$$

This is possible since either condition ensures that ϕ_l never vanishes. There is no way to define the logarithm uniquely, but we choose a branch. We calculate

$$w_x + w^2 - 2izw = u. \tag{4.5}$$

It is related to a factorization

$$-\partial^2 + u - z^2 = (\partial + w - iz)(-\partial + w - iz)$$

and in particular, if $z = i\tau$

$$\int \phi(-\partial^2+u+\tau^2)\phi dx = \int \phi(\partial+w+\tau)(-\partial+w+\tau)\phi dx = \|(-\partial+w+\tau)\phi\|_{L^2}^2.$$

Lemma 4.9. *Let $\tau > 0$. The Miura map*

$$M : L^2(\mathbb{R}) \ni w \to M(w) = w_x+2\tau w+w^2 \in \{u \in H^{-1} : L+\tau^2 \text{ is positive definite }\}.$$

is a diffeomorphism.

Proof. The map is quadratic and smooth. Let $u \in H^{-1}$ with $L + \tau^2$ positive definite. Then ϕ_l is real without zeros and

$$\Big\| \frac{\partial_x \phi}{\phi} - \tau \Big\|_{L^2(x-1,x+1)} \to 0$$

as $x \to \pm\infty$. We write $u = f_x + 2\tau f$. Then, by (4.5)

$$w(x) = \int_{-\infty}^x \exp\Big(- 2\tau(x-y) - 2\int_y^x w dt\Big)(f' + 2\tau f) dy.$$

Then

$$-2\tau(x-y) - \int_y^x w dt \le -\tau(x-y) + C.$$

Schur's lemma gives the bound for the 'f' part. We integrate by parts to remove the derivative from f and we have to bound

$$-\int_{-\infty}^{x} \exp\Big(-2\tau(x-y) - 2\int_y^x w dt\Big)((w+2\tau)f)dy.$$

in L^2, which is left as exercise. Let w_j be solutions for u_j and $w = w_2 - w_1$. Then

$$\partial_x w + 2\tau w + (w_1 + w_2)w = u_2 - u_1$$

and we argue as above. Similarly we invert the linearization and obtain differentiability of the inverse. □

4.2.1 Creating eigenvalues: The Bäcklund transform

Suppose $L + \tau^2$ is positive definit, ϕ_r resp ϕ_l the left and right Jost solution, $c_1, c_2 > 0$ and

$$\phi = c_1\phi_l + c_2\phi_r.$$

Both ϕ_l and ϕ_r are positive, hence the same is true for ϕ. Moreover $\phi \to \infty$ as $x \to \pm\infty$. Let

$$w = \partial_x \log \phi$$

Then again

$$-\partial^2 + u + \tau^2 = (\partial + w)(-\partial + w).$$

We obtain a map u to $\tilde{u}$ by replacing w by $-w$, or equivalently, by changing the order. Then

$$\tilde{u} = -w_x + w^2 - \tau^2.$$

Lemma 4.10. *We have*

$$(-\partial^2 + u)(\partial + w) = (\partial + w)(-\partial^2 + \tilde{u})$$

Moreover $-\tau^2$ is an eigenvalue of $-\partial^2 + \tilde{u}$ with eigenfunction ϕ^{-1}. The remaining spectrum does not change.

There is an instructive example: $u = 0$, $\psi_l = \frac{1}{2}e^{\tau x}$ $\psi_r = \frac{1}{2}e^{-\tau x}$, $\phi = \cosh(\tau x)$, $\phi^{-1} = \operatorname{sech}(\tau x)$,

$$w = \partial_x(\cosh(\tau x)) = \tau \tanh(\tau x)$$

$$\begin{aligned}\tilde{u} &= -w_x + w^2 \\ &= -\tau^2 \operatorname{sech}^2(\tau x) + \tau^2 \tanh^2(\tau x) - \tau^2 \\ &= -2\tau^2 \operatorname{sech}^2(\tau x).\end{aligned}$$

4.3 The Green's function

Let $u \in H^{-1}$, $\operatorname{Im} z > 0$ and either $\operatorname{Re} z \neq 0$ or $L + \tau^2$ positive definite.

Lemma 4.11. *The operator $(L - z^2)^{-1}$ has an integral kernel given by*

$$g(x,y) = -\frac{1}{2iz} \begin{cases} \phi_l(y)\phi_r(x) & \text{if } y < x \\ \phi_r(y)\phi_l(x) & \text{if } y > x \end{cases}$$

Proof. We obverse the the integral kernel is symmetric and continous. For fixed y $\mathbb{R}\backslash\{y\} \ni x \to k(x,y)$ is clearly a solution. We have exponential decay away from the diagonal. The jump of $\partial_x k(x,y)$ on the diagonal is the Wronskian $\frac{1}{2iz} W(\phi_l, \phi_r)$. The Wronskian is constant and we check at $\pm\infty$ that the jump is -1. $\square$

Since

$$-\partial^2 + u - z^2 = (\partial + w - iz)(-\partial + w - iz)$$

and since we can invert the first operator operators explicitly we obtain an expression of the integral kernel in terms of w. First

$$(-\partial + w - iz)f = g$$

can be inverted by

$$f(x) = -\int_x^\infty \exp\left(-iz(x-y) - \int_x^y w(t)dt\right) g(y)dy$$

and

$$(\partial + w - iz)f = g$$

we invert by

$$f(x) = \int_{-\infty}^x \exp\left(iz(x-y) - \int_y^x w(t)dt\right) g(y)dy$$

hence we invert

$$(-\partial^2 + u - z^2)f = g$$

$$f(x) = \int \int_{\max\{x,y\}}^\infty \exp\left(-iz(x+y-2t) - \int_x^t w(t)dt - \int_y^t w(t)dt\right) dtdy$$

so that

$$g(x,y) = \int_{\max\{x,y\}}^\infty \exp\left(-iz(x+y-2t) - \int_x^t w(s)ds - \int_y^t w(s)\right) dt$$

and

$$g(x) := g(x,x) = \int_x^\infty \exp(\left(-2iz(x-t) - 2\int_x^t w\right) dt$$

hence

$$\partial_x g + 2(iz - w)g = -1 \tag{4.6}$$

We substitute $g = -\frac{1}{2iz(v+1)}$ and obtain

$$2iz = -\frac{v'}{(v+1)^2} + \frac{2iz}{v+1} - 2\frac{w}{v+1}.$$

Lemma 4.12.

$$-\frac{1}{2}\partial_x \log(v+1) - izv = w \tag{4.7}$$

4.3.1 Regularized Fredholm determinant

Let K be a compact operator on a Hilbert space which does not have an eigenvalue 1. Then by Lidskii's theorem one can define a determinant so that

$$\operatorname{tr}(1+K) = \prod_j (1+\lambda_j)$$

where λ_j are the eigenvalues. We want to apply it to

$$\begin{aligned} -\partial^2 + u - z^2 &= (-\partial - iz)(1 + (-\partial - iz)^{-1}(\partial f + izf)(\partial - iz)^{-1})(\partial - iz) \\ &= (-\partial - iz)(1 - f(\partial - iz)^{-1} - (-\partial - iz)^{-1} f\partial(\partial - iz)^{-1})(\partial - iz) \end{aligned}$$

where the bracket has the form $1 + K$, K Hilbert-Schmidt.

We will be interested in something like $\log\det L$ up to constants independent of u. So it suffices to try to define

$$\log\det(1 + (-\partial + iz)^{-1} u(\partial + iz)^{-1}).$$

This is still not good since the operator is only Hilbert-Schmidt, but lets ignore this point for a moment. For matrices we would have (for diagonalizable matrices which are dense we check that by diagonalization)

$$\log\det(1+A) = -\sum_{n=1}^\infty \frac{(-1)^{n+1}}{n} \operatorname{tr} A^n$$

provided $\|A\| < 1$. This can be justified for trace class operators as well, where it is a consequence of Lidskii's theorem

$$\log\det(1+K) = \sum_{n=1}^\infty \frac{(-1)^{n+1}}{n} \operatorname{tr} K^n.$$

This is almost good, since $\operatorname{tr} K^n$ is defined for K Hilbert-Schmidt if $n \geq 2$ (we have AB is trace class of both A and B are Hilbert-Schmidt).

Definition 4.13. *Let K be a trace class operator. We define*

$$\det_2(1+K) = \det(1+K)\exp(-\operatorname{tr} K)$$

Theorem 4.14. *There is a unique continuous extension of $\det_2(1+K)$ to Hilbert-Schmidt operators K. Moreover, if K is Hilbert-Schmidt and satisfies $\|K\|_{H\to H} < 1$ then*

$$\log\det_2(1+K) = \sum_{n\geq 2} \frac{(-1)^{n+1}}{n} \operatorname{tr}(K^n).$$

Let

$$R_\pm = (\pm\partial - iz)^{-1}.$$

Then

$$\begin{aligned} \log\det_2(1+R_-uR_+) &= \sum_{n\geq 2} \frac{(-1)^{n+1}}{n} \operatorname{tr}((R_-uR_+)^n \\ &= \sum_{n\geq 2} \frac{(-1)^{n+1}}{n} \operatorname{tr}(((-\partial^2 - z^2)^{-1}u)^n. \end{aligned} \tag{4.8}$$

We calculate

$$\begin{aligned} &\frac{d}{ds}\log\det(1+R_-(u+s\phi)R_+) - s\operatorname{tr}(R_-\phi R_+)|_{s=0} \\ &\quad= \frac{d}{ds}\left[\log\det(1+R_-uR_+) + \log\det(1+s(1+R_-uR_+)^{-1}R_-\phi R_+) - s\operatorname{tr}(R_-\phi R_+)\right]\Big|_{s=0} \\ &\quad= \operatorname{tr}(1+R_-uR_+)^{-1}R_-vR_+) - \operatorname{tr}(R_-vR_+) \\ &\quad= \operatorname{tr}(\partial - iz)(-\partial^2 - z^2 + u)^{-1}(-\partial - iz)(R_-\phi R_+)) + \frac{1}{2iz}\int \phi dx \\ &\quad= \operatorname{tr}\left[(-\partial^2 - z^2 + u)^{-1}\phi\right] + \frac{1}{2iz}\int \phi dx \\ &\quad= \int \Big(g(x,x) + \frac{1}{2iz}\Big)\phi dx. \end{aligned}$$

We recall that the integral kernel of $-\partial^2 - z^2$ is

$$g_0(x,y) = -\frac{1}{2iz}e^{iz|x-y|}$$

The summands are not defined since the functions u, v are not assumed to be in L^1, and the operators are Hilbert-Schmidt but not trace class. This is dealt with either by approximation, or the definition via the series: Let $\phi \in L^2$ and $\|R_- u R_+\|_{L^2 \to L^2} < 1$. Then

$$
\begin{aligned}
&\frac{d}{ds} \log \det{}_2(1 + R_-(u+sv)R_+)\Big|_{s=0} \\
&\qquad = \sum_{n=2}^{\infty} (-1)^{n+1} \operatorname{tr}\Big((R_- u R_+)^{n-1} R_- \phi R_+\Big) \\
&\qquad = \operatorname{tr}\Big(\sum_{n=1}^{\infty} (-1)^n (-\partial^2 - z^2)^{-1} u)^{n-1} (-\partial^2 - z^2)^{-1} \phi\Big) \\
&\qquad = \operatorname{tr}\Big((1 - (-\partial^2 - z^2)^{-1} u)^{-1} (-\partial^2 + z^2)^{-1} v\Big) - \operatorname{tr}(-\partial^2 + z^2)^{-1} v \\
&\qquad = \operatorname{tr}(-\partial^2 - u - z^2) u)^{-1} \phi - (-\partial^2 + z^2)\phi\Big) \\
&\qquad = \int \Big(g(x,x) + \frac{1}{2iz}\Big) \phi dx
\end{aligned}
$$

Theorem 4.15. *The following identities hold:*

$$a(z) = \det{}_2(1 + R_- u R_+), \tag{4.9}$$

$$\frac{\delta}{\delta u} a = g(x,x) + \frac{1}{2iz}. \tag{4.10}$$

Proof. We observe that $a(z, 0) = 1 = \det_2(1 + R_- 0 R_+)$. Both a and $\log_2 T$ are defined by some expansion, from which it is not hard to see that the derivatives above are continuous in $\phi \in L^2$.

We want to calculate $\frac{d}{ds} a(z, u + s\phi)|_{s=0}$ which is defined by the Jost solution ϕ_l which satisfies

$$-\phi_l'' + u\phi_l = z^2 \phi_l$$

We denote the derivative of ϕ_l with respect to s at $s = 0$ by $\dot{\phi}_l$. It satisfies

$$-\dot{\phi}_l'' + u\dot{\phi}_l - z^2 \dot{\phi}_l = -\phi \phi_l$$

and

$$\dot{\phi}_l = c\phi_l - \frac{1}{2iz} \phi_l(x) \int \phi_r(y) \phi_l(y) \phi(y) dy + \frac{1}{2iz} \phi_r(x) \int \phi_l^2(y) \phi(y) dy.$$

Then

$$
\begin{aligned}
&\lim_{x \to \infty} \exp\Big(izx - \int_0^x f dy\Big)\Big(-\frac{1}{2iz} \phi_l \int \phi_r \phi_l \phi dy + \frac{1}{2iz} \phi_r(x) \int \phi_l(y) \phi(y) dy\Big) \\
&\qquad = a(z) \int g(x,x) \phi(x) dx.
\end{aligned}
$$

Differentiating (recall $\partial f + 2izf = u$ hence $2iz\int f = \int u$)

$$\lim_{x\to-\infty} e^{izx-\int_0^x f dx}\phi_l(x) = 1$$

we obtain

$$0 = \lim_{x\to-\infty} e^{izx-\int_0^x f dx}\dot\phi_l + \int_{-\infty}^0 \dot f dy$$

and

$$c = -\int_{-\infty}^0 \dot v dx$$

We obtain

$$\frac{d}{ds}\log a(z,u+sv)|_{s=0} = \int \left(g(x,x) + \frac{1}{2iz}\right) v(y)dy$$

Since

$$a(z,0) = \det_2(1) = 1$$

and

$$\partial_s \log(a(z,su)) = \partial_s \det_2(1 + s(-\partial - iz)^{-1}u(\partial - iz)^{-1})$$

(4.9) follows from the fundamental theorem of calculus. The identity (4.10) is a consequence of the calculations. □

The expansion consist of the summands $\frac{(-1)^{n+1}}{n}\operatorname{tr}\Big((-\partial - iz)^{-1}u(\partial - iz)^{-1}\Big)^n$ and

$$\operatorname{tr}\Big((-\partial-iz)^{-1}u(\partial-iz)^{-1}\Big)^n = (-2iz)^{-n}\frac{(-1)^{n+1}}{n}\int_{\mathbb{R}^n}\prod_{j=1}^n e^{iz|x_{j+1}-x_j|}u_j(x_j)dx_j$$

where $x_{n+1} = x_1$ where $u_j = u$,

Lemma 4.16.

$$\left|\int_{\mathbb{R}^n}\prod_{j=1}^n e^{iz|x_{j+1}-x_j|}u_j(x_j)dx_j\right| \le (\operatorname{Im} z/2)^{-n/2}\prod_{j=1}^n \|u_j\|_{L^2} \tag{4.11}$$

and, if $\sum \frac{1}{p_j} = 1$

$$\left|\int_{\mathbb{R}^n}\prod_{j=1}^n e^{iz|x_{j+1}-x_j|}u_j(x_j)dx_j\right| \le (\operatorname{Im} z/2)^{1-n}\prod_{j=3}^n \|u_j\|_{L^{p_j}} \tag{4.12}$$

Proof. We estimate by taking the absolute value of the integrand and omit the exponential term with $|x_{n+1} - x_n|$. It suffices to bound for $\tau > 0$

$$\left| \int_{\mathbb{R}^n} \prod_{j=1}^n e^{-\tau|x_{j+1}-x_j|} u_j(x_j) dx_j \right| \leq (\tau)^{-n/2} \prod_{j=1}^n \|u_j\|_{L^2}.$$

Similarly we deal with the second estimate. If $n = 2$ we interpret the estimate as the inner product of the an operator applied to the first function. By Schur's lemma we obtain the bound

$$\left| \int_{\mathbb{R}^2} e^{-\tau|x_2-x_1|} u_1(x_1) u_2(x_2) dx_1 dx_2 \right| \leq \sup_x \int e^{-\tau|x-y|} dy \|u_1\|_{L^p} \|u_2\|_{L^q}.$$

Similar we write

$$\int \mathbb{R}^n \prod_{j=1}^{n-1} e^{-\tau|x_{j+1}-x_j|} |u_j(x_j)| dx_j = \int \Big(\int_{\mathbb{R}^{n-2}} \prod_{j=1}^{n-2} e^{-\tau|x_{j+1}-x_j|} u_j(x_j)(Tu_n)(x_{n-1}) dx_{n-1}$$

where T is the integral operator defined above. Then, by Cauchy-Schwarz

$$\|Tu_n\|_{L^\infty} \leq \sqrt{\frac{2}{\tau}} \|u_n\|_{L^2}.$$

For the second estimate we let

$$\frac{1}{p'_{n-1}} = \frac{1}{p_n} + \frac{1}{p_{n-1}}$$

and

$$u'_{n-1} = u_{n-1} T u_n$$

with

$$\|u'_{n-1}\|_{L^{p'_{n-1}}} \leq \|u_{n-1}\|_{L^{p_{n-1}}} \|Tu_n\|_{L^{p_n}}$$

which we estimate by Schur's lemma. Induction yields the full estimate. □

Theorem 4.17. *There is an asymptotic expansion*

$$-\frac{i}{2} \log a(z) \sim \sum_{n=0}^\infty H_n (2z)^{-3-2n}.$$

where H_n are functions on H^n given as integrals over differential polynomials

$$H_N = \int \frac{1}{2} |u^{(n)}|^2 + \text{ cubic and higer }.$$

To be more precise: There exists $\delta > 0$ so that if $(\operatorname{Im} z)^{-3/2} \|u\|_{L^2} < \delta$ Then

$$\left| (2z)^{2n+3} (-i/2) \log a(z) - \sum_{j=0}^{n-1} H_n (2z)^{2(n-j)} \right| \leq C(1 + \|u\|_{L^2}^n) \|u\|_{H^n}^2 \quad (4.13)$$

A consequence of the estimate is that

$$(2z)^{2n+3}(-i/2)\log a(z) - \sum_{j=0}^{n-1} H_n (2iz)^{2(n-j)} \to H_n$$

for $u \in H^n$ and $\operatorname{Im} z \to \infty$.

We will later see that

$$H_0 = \frac{1}{2}\int u^2 dx, \quad H_1 = \int \frac{1}{2}u_x^2 + u^3 dx, \quad H_2 = \int \frac{1}{2}u_{xx}^2 - 10uu_x^2 + 5u^4 dx$$

Proof. By the previous argument we it suffices to show the expansion for a finite number of terms. We begin with the first

$$\frac{1}{\xi^2 - 4z^2} = \frac{1}{-4z^2} + \frac{1}{4z^2}\frac{\xi^2}{\xi^2 - 4z^2}$$

hence iteratively

$$\begin{aligned}
&\frac{1}{(2iz)^2}\int e^{2iz|x_1-x_2|}u(x_1)u(x_2)dx_1dx_2 \\
&\quad = -\frac{1}{iz}\int (\xi^2 - 4z^2)^{-1}\hat{u}(\xi_1)\overline{\hat{u}(\xi)}d\xi \\
&\quad = \sum_{j=1}^{n-1} \|u^{(j)}\|_{L^2}^2 (2z)^{-2j-3} - \frac{1}{(2z)^{2(n-1)}}\int \frac{1}{\xi^2 - 4z^2}|\hat{u}^{(n)}|^2 d\xi
\end{aligned}$$

For $n = 3$ we use Fubini and integrate by parts

$$\begin{aligned}
&\int_{\mathbb{R}^3} u(x_1)u(x_2)u(x_3)e^{iz(|x_1-x_2|+|x_2-x_3|+|x_1-x_3|}dx_1dx_2dx_3 \\
&\quad = \frac{1}{iz}\int_{\mathbb{R}^3} u(x_1)u(x_2)u(x_3)\left(\frac{x_1-x_2}{|x_1-x_2|} + \frac{x_1-x_3}{|x_1-x_3|}\right)\partial_{x_1}e^{iz(|x_1-x_2|+|x_2-x_3|+|x_1-x_3|)}dx_1dx_2dx_3 \\
&\quad = -\frac{1}{iz}\int_{\mathbb{R}^3} u'(x_1)u(x_2)u(x_3)\frac{x_1-x_2}{|x_1-x_2|}e^{iz(|x_1-x_2|+|x_2-x_3|+|x_1-x_3|)}dx_1dx_2dx_3 \\
&\qquad - \frac{1}{iz}\int_{\mathbb{R}^3} u'(x_1)u(x_2)u(x_3)\frac{x_1-x_3}{|x_1-x_3|}e^{iz(|x_1-x_2|+|x_2-x_3|+|x_1-x_3|)}dx_1dx_2dx_3 \\
&\qquad + \frac{1}{iz}\int_{\mathbb{R}^2} 2(u(x_2)+u(x_3))u(x_2)u(x_3)e^{2iz(|x_2-x_3|}dx_2dx_3
\end{aligned}$$

We iterate that until we either obtain one dimenionsal integrals over differential polynomials or sufficient decay, since we gain a power $\frac{1}{iz}$ in each integration by parts or we reduce the integration dimension by 1. It remain to do the proper counting. $\square$

We will use also the epansion for the variational derivatives. Again we reduce matters to a finite number of terms, and to variational derivatives of integrals over differential polynomials. The Hamiltonians H_n and their variation derivatives can be computed by the Lenard recursion. Let ϕ, ψ be solutions to ¸

$$(-\partial^2 - z^2 + u)\phi = 0.$$

Then

$$\partial(\phi\psi) = \phi'\psi + \phi\psi'$$
$$\partial^2(\phi\psi) = \phi''\psi + 2\phi'\psi' + \phi\psi'' = 2(-z^2+u)(\phi\psi) + 2\phi'\psi'$$
$$\partial^3(\phi\psi) = 4(-z^2+u)\partial(\phi\psi) + 2u'(\phi\psi)$$

and we arrive at the crucial equation

$$\partial^3(\phi\psi) + 4(z^2 - u)\partial(\phi\psi) - 2u'\phi\psi = 0 \tag{4.14}$$

Now, by Lemma 4.10 and (4.10)

$$\frac{\delta \log a}{\delta u} = -\frac{1}{2iz}(\phi\psi + 1)$$

so that

$$-(\partial^3 + 4u\partial + 2u')\frac{\delta \log a}{\delta u} = \partial\Big(z^2\frac{\delta \log a}{\delta u} - \frac{1}{iz}u\Big) \tag{4.15}$$

We expand and we obtain the Lenard recursion

$$\partial\frac{\delta H_{n+1}}{\delta u} = (-\partial^3 + 4u\partial + u')\frac{\delta H_n}{\delta u}$$

with

$$H_0 = \frac{1}{2}\int u^2 dx, \frac{\delta H_1}{\delta u} = u$$

and

$$\partial\frac{\delta}{\delta u}H_1 = (-\partial^3 + 4u\partial + 2u')u = -u''' + 6uu' = \partial(-u_{xx} + 3u^2)$$

$$H_1 = \int \frac{1}{2}u_x^2 + u^3 dx$$

$$\partial\frac{\delta}{\delta u}H_2 = (-\partial^3+4u\partial+2u')(-u_{xx}+3u^2) = \partial_x\Big(u^{(4)}-3\partial^2u^2-10\partial_x(uu_x)+5u_x^2+10u^3\Big)$$

$$H_2 = \frac{1}{2}\int u_{xx}^2 + 10uu_x^2 + 5u^4 dx.$$

The amazing story of Lenards contribution is told by Praught and Smirnov in [24].

4.4 Computing Poisson brackets

On nice functions on $\mathcal{S}(\mathbb{R})$ (including integrals over differential polynomials, and $a(z)$) we define the Gardner-Poisson bracket

Definition 4.18.

$$\{F, G\} = \int \frac{\delta F}{\delta u} \partial_x \frac{\delta G}{\delta u} dx$$

The Hamiltonian vector field of F is then

$$\partial_x \frac{\delta F}{\delta u}.$$

Then

$$\{u(x), F\} = \int \delta_x \partial_x \frac{\delta F}{\delta u}$$

in the distributional sense for test functions

$$\left\{ \int \phi u dx, F \right\} = -\int \frac{\delta F}{\delta u} \partial_x \phi dx.$$

In particular

$$\{u, H_0\} = \partial_x u$$

$$\{u, H_1\} = -\partial_{xxx} u + 6 u u_x.$$

Lemma 4.19. *The Gardner Poisson structure and the Magri structure are compatible in the sense that for* $F(w) = f(w_x - 2izw + w^2)$

$$\int \frac{\delta F}{\delta w} \partial_x \frac{\delta G}{\delta w} dx = \int \frac{\delta f}{\delta u}|_{u=w_x - i\tau w + w^2} (-\partial^3 + \partial(u - z^2) + (u - z^2)\partial) \frac{\delta g}{\delta u}|_{u=w_x - i\tau w + w^2} dx.$$

Proof. This is a consequence of the chain rule. Let

$$u = (w + s\omega)_x - 2iz(w + +s\omega) + (w + s\omega)^2$$

Then

$$\begin{aligned}
\frac{d}{ds} F(w + s\omega)|_{s=0} &= \frac{d}{ds} f(\partial_x(w + s\omega) - 2iz(w + s\omega) + (w + s\omega)^2)|_{s=0} \\
&= \int \frac{\delta f}{\delta u} (\omega_x - 2iZ\omega + 2w\omega) dx \\
&= \int (-\partial - 2iz + 2w) \frac{\delta f}{\delta u} \omega dx
\end{aligned}$$

and

$$
\begin{aligned}
&\int \frac{\delta}{\delta w} F \partial_x \frac{\delta}{\delta w} G dx \\
&\quad = \int (-\partial_x - 2iz + 2w) \frac{\delta f}{\delta u} \partial_x (-\partial_x - 2iz + 2w) \frac{\delta g}{\delta u} dx \\
&\quad = \int \frac{\delta f}{\delta u} (\partial - 2iz + 2w) \partial (-\partial - 2iz + 2w) \frac{\delta g}{\delta u} dx \\
&\quad = \int \frac{\delta f}{\delta u} (-\partial^3 + 4(w_x - 2izw + w^2 - 4z^2)\partial + 2(\partial(w_x - 2izw + w^2)) \frac{\delta g}{\delta u} dx \\
&\quad = \frac{\delta f}{\delta u} (-\partial^3 + 4u\partial + 2u_x) \frac{\delta g}{\delta u} dx
\end{aligned}
$$

□

Theorem 4.20. *The functions $a(z_1, u)$, $a(z_2, u)$ and H_n all Poisson commute.*

Proof.

$$
\begin{aligned}
&- 8z_1 z_2 \int \frac{\delta}{\delta u} \log a(z_1) \frac{\delta}{\delta u} \log a(z_2) \\
&\quad = 2 \int (\phi_l(x, z_1)\phi_r(x, z_1) + \frac{1}{2iz_1}) \partial_x \Big(\phi_l(x, z_2)\phi_r(x, z_2) + \frac{1}{2iz_2} \Big) dx \\
&\quad = \int \Big(\phi_l(z_1)\phi_r(z_1) + \frac{1}{2iz_1} \Big) \partial_x \Big(\phi_l(z_2)\phi_r(z_2)) + \frac{1}{2iz_2} \Big) \\
&\qquad\quad - \Big(\phi_l(z_2)\phi_r(z_2) + \frac{1}{2iz_2} \Big) \partial_x \Big(\phi_l(z_1)\phi_r(z_1) + \frac{1}{2iz_1} \Big) dx \\
&\quad = \lim_{X\to\infty} \int_X^X (z_1^2 - z_2^2) \partial_x \Big(W(\phi_l(z_1)\phi_r(z_2)) W(\phi_l(z_1)\phi_r(z_2)) \Big) \\
&\qquad\quad + \partial_x \Big(\frac{1}{2iz_2} \Big(\phi_l(z_1)\phi_r(z_1) + \frac{1}{2iz_1} \Big) - \frac{1}{2iz_1} \Big(\phi_l(z_z)\phi_r(z_2) + \frac{1}{2iz_2} \Big) \Big) dx \\
&\quad = 0
\end{aligned}
$$

□

Formula (4.14) shows that formally $\det(1+R_u R_+)$ is a Casimir for a linear combination of the Gardner and the Magri Poisson bracket.

Theorem 4.21. *The following identities hold*

$$
\Big\{ u, -iz \ln a \Big\} = \frac{1}{2} \partial_x \frac{v}{v(z) + 1}
$$

$$\{u, H_0\} = u_x$$

$$\{u, H_1\} = -u_{xxx} + 6uu_x$$

$$\{v(\tilde z), -iz \ln a\} = \frac{1}{4z^2 - 4\tilde z^2} \partial \frac{v(\tilde z) - v(z)}{v(z) + 1}$$

$$\{v, H_0\} = v_x$$

$$\begin{aligned} \{v, H_1\} &= 2\partial_x((1+v)u) \\ &= \partial_x \Big[- v_{xx} + \frac{3}{2} \frac{v_x^2}{v+1} + 2\tau^2 v^3 + 6\tau^2 v^2 \Big] \end{aligned}$$

$$\{w, H_0\} = w_x$$

$$\{w, H_1\} = -w_{xxx} + \partial_x (2w^3 + 6\tau w^2)$$

Proof. We begin with (recall $g = -\frac{1}{2iz(v+1)}$)

$$\begin{aligned} \{u, -iz \log a\} &= -iz \partial_x \frac{\delta}{\delta u} \log a \\ &= -iz \partial_x \Big(g(x,x) + \frac{1}{2iz} \Big) \\ &= -\frac{1}{2} \partial_x \Big(\frac{1}{v+1} - 1 \Big) \\ &= \frac{1}{2} \partial \frac{v}{v+1} \end{aligned}$$

$$\Big\{ u, \frac{1}{2} \int u^2 dx \Big\} = \partial_x u$$

$$\Big\{ u, \int \frac{1}{2} u_x^2 + u^3 \Big\} = \partial_x (-\partial^2 u + 6uu_x).$$

By translation invariance

$$\{w, H_0\} = w_x, \{v, H_0\} = v_x.$$

Next

$$
\begin{aligned}
\{v(z_1), -2iz_2 \log a(z_2)\} &= -\frac{2iz_2}{2iz_1}\Big\{\frac{1}{g(z_1,x)}, \log a(z_2)\Big\} \\
&= \frac{z_2}{z_1}\frac{1}{g^2(z_1,x)}\{g(z_1,x), \log a(z_2)\} \\
&= -\frac{z_2}{z_1}(g(z_1,x))^{-2}\Big(L_{z_1}^{-1}\{u, \log a(z_2)\}L_{z_1}^{-1}\delta_x\Big)(x) \\
&= -\frac{z_2}{z_1}\frac{1}{g^2(z_1)}\int g(z_1,x,y)\partial g(z_2,y)g(z_1,y,x)dy \\
&= -\frac{z_2/z_1}{4z_1^2-4z_2^2}\frac{1}{g^2(z_1)}\int g(z_1,x,y) \\
&\qquad \times(-\partial^3+2\partial u+2u\partial-4z_1^2\partial)g(z_2,y)g(z_1,y,x)dy \\
&= -\frac{z_2/z_1}{4z_1^2-4z_2^2}\frac{1}{g^2(z_1)}\int g(z_1,x,y)\Big\{L_{z_1}g'(z_2)+g'(z_2)L_{z_1} \\
&\qquad -2L_{z_1}g(z_2)\partial+2\partial g(z_2)L_{z_1}\Big\}g(z_1,y,x)dy \\
&= -2\frac{z_2/z_1}{4z_1^2-4z_2^2}g^{-2}(z_1)(g'(z_2)g(z_1)-g(z_2)g'(z_1)) \\
&= 2\frac{z_2/z_1}{4z_2^2-4z_1^2}\partial\frac{g(z_2)}{g(z_1)} \\
&= 2\frac{1}{4z_2^2-4z_1^2}\partial\frac{v(z_1)+1}{v(z_2)+1} \\
&= 2\frac{1}{4z_2^2-4z_1^2}\partial\frac{v(z_1)-v(z_2)}{v(z_2)+1}.
\end{aligned}
$$

We recall

$$
-\frac{i}{2}(2z)^5a(z)-(2z)^2\frac{1}{2}\int u^2dx \to \frac{1}{2}\int u_x^2+2u^3dx
$$

at least for smooth u. We are interested in the limit

$$
\begin{aligned}
\lim_{\operatorname{Im} z\to\infty}&\Big\{v(z), (2\tilde z)^5\frac{i}{2}a(\tilde z)-(2\tilde z)^2\frac{1}{2}\int u^2dx\Big\} \\
&= -\lim_{\operatorname{Im} z\to\infty}(2\tilde z)^2\partial\Big(\frac{v(z)-v(\tilde z)-v(z)(1+v(\tilde z)}{1+v(\tilde z)}\Big) \\
&= \partial_x((1+v(z))u)
\end{aligned}
$$

We recall

$$
w_x-2izw+w^2=u
$$

and

$$-\frac{1}{2}\partial_x \log(1+v) + izv = w$$

The map

$$L^2 \ni w \to w_x - 2izw + w^2 =: u \in H^{-1}$$

is a diffeomorphism with derivative (see Lemma 4.9

$$\dot{w} \to \dot{w}_x - 2iz\dot{w}$$

at $w = 0$ which we can explicitly invert. Similarly

$$\{v \in H^2 : v \neq -1\} \ni v \to -\frac{1}{2}\partial \log(1+v) - izv \in L^2$$

is a diffeomorphism with derivative

$$\dot{v} \to -\frac{1}{2}\dot{v}_x - iz\dot{v}$$

at $v = 1$ which is again invertible. Thus $w \to u$ is invertible near zero and the deirvative at 0 is

$$\dot{u} \to \dot{v} = 2(\partial^2 - 4z^2)^{-1}\dot{u}$$

Lemma 4.22. *We have*

$$\lim_{\tau \to \infty} 2\tau^2 v(i\tau) = u$$

in H^{-1}.

We compute

$$\begin{aligned}
\partial_x(1+v)u &= \partial_x(1+v)(w_x - 2izw + w^2) \\
&= \partial_x\Big[(1+v)\Big(-\frac{1}{2}\partial_x \frac{v_x}{1+v} + izv_x - 2iz\frac{v_x}{1+v} + z^2 v \\
&\qquad + \frac{1}{4}\frac{v_x^2}{(1+v)^2} - iz\frac{vv_x}{1+v} - z^2v^2\Big)\Big] \\
&= \partial_x\Big[-v_{xx} + \frac{3}{2}\frac{v_x}{1+v} - 2z^2v^3 - 6z^2v^2\Big]
\end{aligned}$$

and

$$\begin{aligned}
\{w, H_1\} &= \{-\partial_x \log(1+v) - izv, H_1\} - \partial_x\{\log(1+v), H_1\} - iz\{v, H_1\} \\
&= -\partial_x \frac{1}{1+v}\{v, H_1\} - iz\{v, H_1\}
\end{aligned}$$

and by the chain rule one arrives at the formula for w after a massive calculation.

$\square$

4.5 Wellposedness of KdV in H^{-1}

4.5.1 The τ flow

We consider the τ flow

$$u_t = \partial_x \frac{\delta \tau a(i\tau)}{\delta u} = \frac{1}{2}\partial_x \frac{v}{v+1} \tag{4.16}$$

where we omit $i\tau$ in the argument.

Proposition 4.1. *There exists $\delta > 0$ so that for $n \geq -1$ the map*

$$B_{\delta\tau^{1/2}}(0)^{H^{-1}} \cap H^n \ni u \to v \in H^{n+2}$$

is smooth and satisfies

$$c_n^{-1}\|u\|_{H^n_\tau} \leq \|v\|_{H^{n+2}_\tau} \leq c_n \|u\|_{H^n_\tau}$$

and

$$c_n^{-1}\|u_2 - u_1\|_{H^n_\tau} \leq \|v_2 - v_1\|_{H^{n+2}_\tau} \leq c_n \|u_2 - u_1\|_{H^n_\tau}.$$

Proof. By Lemma 4.9 the map $w \to u$ is a diffeomorphism in suitable spaces. By the triangle inequality

$$\begin{aligned} \Big| \|u\|_{H^{-1}_{2\tau}} - \|w\|_{L^2} \Big| &\leq \|w^2\|_{H^{-1}_\tau} \\ &\leq (2\tau)^{-1/2}\|w^2\|_{L^1} \\ &\leq (2\tau)^{-1/2}\|w\|^2_{L^2} \end{aligned}$$

and

$$\begin{aligned} \Big| \|w\|_{L^2} - \frac{1}{2}\|v\|_{H^1_\tau} \Big| &\leq \|\frac{1}{2} v_x \frac{v}{v+1}\|_{L^2} \\ &\leq \|v\|_{H^1_{2\tau}} \frac{1}{2}\|\frac{v}{v+1}\|_{L^\infty} \\ &\leq \|v\|_{H^1_{2\tau}}\|v\|_{L^\infty} \\ &\leq \tau^{-1/2}\|v\|^2_{H^1_{2\tau}} \end{aligned}$$

assuming

$$\|v\|_{L^\infty} \leq \tau^{-1/2}\|v\|_{H^1_\tau} \leq \frac{1}{2}$$

Similarly we estimate differences. The arguments for larger n is the same, but it requires some interpolation. $\square$

As a consequence we obtain a local solution in H^n for $n \geq -1$ to the Cauchy problem for (4.16) by the Cauchy-Lipschitz theorem which holds in Banach spaces. The functions $a(z)$ Poisson commute and hence they are preserved by the τ flow.

Lemma 4.23. *Let τ be sufficiently large. Then*

$$-\tau \log a(i\tau) = \int w^2 dx$$

Proof.

$$\begin{aligned} -\tau \log a(i\tau) &= -\tau \lim_{X\to\infty} \left(\log \frac{\phi_l(X)e^{-\tau X}}{\phi_l(-X)e^{\tau X}} - \frac{1}{2\tau}\int_{-X}^X u dx \right) \\ &= -\tau \lim_{X\to\infty} \left(\int_{-X}^X w - \frac{1}{2\tau} u dx \right) \\ &= \lim_{X\to\infty} \left(-\tau \int_{-X}^X w - \frac{1}{2\tau}(w_x + 2\tau w + w^2) dx \right) \\ &= \frac{1}{2}\int w^2 dx. \end{aligned}$$

□

We obtain by the triangle inequality which implies a uniform bound by Proposition 4.1 (or its proof)

$$\|u(t)\|_{H^{-1}_\tau} \leq 2\|u_0\|_{H^{-1}_\tau} \tag{4.17}$$

is δ is sufficiently small. In particular the flow is global in H^n, $n \geq -1$.

Definition 4.24. *Let X be a translation invariant Banach space. We call a subset $Q \subset X$ equicontinuous if for all ε there exists h_0 so that*

$$\|f(.+h) - f\|_X < \varepsilon \qquad \text{for } |h| \leq h_0$$

Lemma 4.25. *A bounded set $Q \subset H^{-1}$ is equicontinuous if and only if*

$$\lim_{\tau\to\infty} \sup_{f\in Q} \|f\|_{H^{-1}_\tau} = 0$$

Proof. Suppose that Q is equicontinuous and $\varepsilon > 0$. Let h_0 be as in the definition and $j \in C_c^\infty(-h_0, h_0)$ with $\int j = 1$. Then

$$\|f - f * j\|_{H^{-1}} \leq \varepsilon$$

$$
\begin{aligned}
\|f\|_{H^{-1}_\tau} &\le \|f - f * j\|_{H^{-1}} + \|(\xi^2+1)^{-1/2}\hat{j}(\xi)\hat{f}\|_{L^2} \\
&= \varepsilon + \sup \frac{(\xi^2+1)\hat{j}}{\xi^2+\tau^2}\|f\|_{H^{-1}}
\end{aligned}
$$

which implies the uniform convergence. Vice versa, suppose that

$$
\lim_{\tau\to\infty} \sup_{u\in Q} \|u\|_{H^{-1}_\tau} = 0
$$

and let $\varepsilon > 0$. There exists τ so that

$$
\|u\|_{H^{-1}_\tau} \le \varepsilon.
$$

We write $u = u_{<\tau} + u_{>\tau}$ with

$$
u_{<\tau} = \mathcal{F}^{-1}\Big(\chi_{|\xi|\le\tau}\hat{u}\Big)
$$

so that $\||\xi|^{-1}u_{>\tau}\|_{L^2} < \varepsilon$ and $\|u_{<\tau}\|_{H^{-1}} < C$. Then

$$
\|u_{>\tau}(.+h) - u_{>\tau}(.)\|_{H^{-1}} \le 2\varepsilon
$$

and

$$
\|u_{<\tau}(.+h) - u_{<\tau}\|_{H^{-1}} \le h\|u_{<\tau}\|_{L^2} \le h\tau\|u\|_{H^{-1}}.
$$

□

Let $Q \subset H^{-1}$ be a bounded and equicontinuous set of initial data, let $\varepsilon > 0$ and $\tau \ge \tau_0$ so for $u \in Q$

$$
\|u\|_{H^1_\tau} < \varepsilon
$$

and suppose that τ_0 is sufficiently large. Let $u(t;\tau)$ be the τ flow applied to $u_0 \in Q$. Then

$$
\|u(t;\tau)\|_{H^{-1}_\tau} \le 2\varepsilon
$$

for all $\tau \ge \tau_0$ and $t \in \mathbb{R}$. With Lemma 4.25 and Propositon 4.1 we see that

$$
\{u(t;\tau) : u_0 \in Q\} \subset H^{-1}
$$

is bounded and equicontinuous and the corresponding set $Q_v \in H^1$ is also equicontinuous.

4.5.2 Wellposedness in H^{-1}

Theorem 4.26 (Wellposedness of KdV in H^{-1})**.** *The Korteweg-de Vries equation is wellposed in H^1. The flow has a unique continuous extension to H^{-1}*

The key is

Proposition 4.2. *The map*

$$B^{H^{-1}}_{\delta\tau_0^{1/2}} \times [\tau_0, \infty) \times \mathbb{R} \ni (u_0, \tau, t) \to u(t, \tau)$$

has a unique continuous extension to

$$B^{H^{-1}}_{\delta\tau_0^{1/2}} \times [\tau_0, \infty] \times \mathbb{R}$$

Proof. We have seen that the τ flows are global in time and the orbits are equicontinuous. In the v coordinates we obtain solutions to

$$v_t = \partial_x \Big(\frac{1}{4\tau_0^2 - 4\tau^2} \frac{v - v(i\tau)}{1 + v(i\tau)} - v \Big) \tag{4.18}$$

The corresponding us form a bounded and equicontinuous set in H^{-1}, hence $\{v(t,\tau)\}$ is bounded (and small) and equicontinuous in H^1.

Lemma 4.27. *Let $\tilde{Q} \subset H^1_\tau \cap \{v > -\frac{1}{2}\}$ be a bounded and equicontinuous set. Then*

$$\Big(\frac{1}{4\tau_0^2 - 4\tau^2} \frac{v - v(i\tau)}{1 + v(i\tau)} - v \Big) \to 2(1+v)u \in H^{-1}$$

uniformly for $v \in \tilde{Q}$.

Proof. This is a quantitative version of the arguments for Theorem 4.21 using the estimates of Propositon 4.1. □

As a consequence (compare the Hamiltonian vector fields in Theorem 4.21)

$$v(t, \tau) \to v(t) \in H^{-2}$$

uniformly for compact time intervals where $v(t)$ is uniformly bounded in H^1. We obtain convergence in H^1.

Lemma 4.28. *Let $v_n(t)$, $t \in I$ be a uniformly bounded and equicontinuous sequence in H^1 which converges uniformly to $v(t)$ in H^{-2}. Then uniformly $v(t) \in H^1$ and $v_n \to v$ in H^1.*

In particular v is a solution to the KdV equation in the v coordinates,

$$v_t = 2\partial_x\Big((1+v)u\Big).$$

Proof. Given t there exists a weakly converging subsequence of $v_n(t)$ in H^1 satisfying

$$|v(t)\|_{H^1} = \| \lim_{n\to\infty} v_n(t)\|_{H^1} \le \liminf_{n\to\infty} \|v_n(t)\|_{H^1} < \infty$$

uniformly in t. We subtract v and reduce the problem to $v=0$. Let h_0 be such that

$$\|v_n(t,.+h) - v_n(t)\|_{H^1} < \varepsilon \qquad \text{for } |h| \le h_0.$$

we decompose $v_n = v_{n,>\tau} + v_{n,<\tau}$. Then as above

$$\|v_{n,<\tau}\|_{H^1} \le \tau^{-3}\|v_{n,<\tau}\|_{H^{-2}} \to 0 \qquad \text{as } n \to \infty$$

and

$$\|v_{n,>\tau}(t)\|_{H^1} < \varepsilon.$$

$\square$

Now suppose that $u_0 \in H^2$. Then $H_0 = \frac{1}{2}\|u\|_{L^2}^2$ is conserved and the L^2 is uniformly bounded. Moreover

$$\|u_x\|_{L^2}^2 \le 2H_1 + \|u\|_{L^3}^3 \le \|u_x\|_{H^1}^{\frac12}\|u\|_{L^2}^{\frac52}$$

and hence $\|u(t)\|_{H^1}$ is uniformly bounded. Similarly

$$\|u_{xx}\|_{L^2}^2 \le H_2 + \|u\|_{L^4}^4 + |u\|_{L^\infty}\|u_x\|_{L^2}^2$$

hence $\|u(t)\|_{H^2}$ is uniformly bounded.

By the diffeomorphismus property $\|v(t)\|_{H^2}$ is uniformly bounded. We first compute

$$w_t = \partial_t(-\frac{1}{2}\frac{v_x}{1+v} + \tau v) = -w_{xxx} + \partial_x(2w^3 + 6\tau w^2)$$

in H^{-1} and

$$u_t = \partial_t(w_x + 2\tau w + w^2) = -u_{xxx} - 6uu - x$$

in H^{-2}. Let u_1 and u_2 be two solutions. Then, formally,

$$\begin{aligned}
\frac{d}{dt}\|u_1-u_2\|_{L^2}^2 &= 3\left|\int (u_2^2-u_1^2)(u_2-u_1)_x dx\right| \\
&= \frac{3}{2}\left|\int (u_2+u_1)_x(u_2-u_1)^2 dx\right| \\
&\le \|u_2-u_1\|_{L^2}^2(\|u_2'\|_{L^\infty} + \|u_1'\|_{L^\infty}) \\
&\le \Big(\|u_2\|_{H^2} + \|u_1\|_{H^2}\Big)\|u_2-u_1\|_{L^2}^2
\end{aligned}$$

where we used the Sobolev inequality. Now Grönwall's lemma implies uniqueness.

4.5.3 Results on wellposedness of the KdV equation

1. Inverse scattering methods: Sufficiently regular and decaying initial data. Schuur [25] gives a fairly precise description of general solutions.

2. Bona-Smith [3] use energy estimates (integration by parts) to prove wellposedness in H^2 . The essential part is the uniqueness argument we used above.

3. Kenig. Ponce and Vega [14] use dispersive techniques to lower the regularity.

4. Bourgain introduced a large number of new ideas to deal with initial data in L^2, [4].

5. Kenig, Ponce and Vega [13] use bilinear estimates to push wellposedness to negative regularity $s > -\frac{3}{4}$. This type of argument cannot be pushed below $s = -\frac{3}{4}$.

6. Kappeler and Topalov [11] proved a similar result for periodic solution. Their argument depends on complex algebraic geometry, more precisely on Riemann surfaces.

7. Molinet showed that no wellposedness can hold in H^s $s < -1$. [20, 21].

8. Higher order KdV equations?

$\square$

4.6 The inverse scattering technique for KdV

Consider the Schrödinger equation

$$i\partial_t \phi + \phi_{xx} - u\phi = 0$$

assuming $\|(1+|x|)u\|_{L^1} < \infty$ and $u \in L^2$. Then there are finitely many eigenvalues whose eigenfunctions span a finite dimensional space $H_p \subset L^2$ corresponding to the point spectrum. The orthogonal complement is H_c.

The operator $H = -\partial^2 + u$ is self adjoint with domain H^2 and generates by Stone's theorem a unitary semigroup

$$\exp(-itH)\exp(-it(-\partial^2 + V))$$

One defines the wave operators via $H_0 = -\partial^2$ by

$$\Omega_\pm = \lim_{t\to\pm\infty} e^{it(H)} e^{-itH_0}$$

Then $\Omega_\pm$ is a unitary map $L^2 \to H_c$, they intertwine the evolutions

$$\Omega_\pm e^{-itH_0} = e^{-itH}\Omega_\pm$$

and the scattering operator is defined by

$$S = \Omega_+^{-1}\Omega_- : L^2 \to L^2.$$

A Fourier transform in t allows to describe $\Omega_\pm$ in terms of 'eigen functions' $\phi_{l,r}(x,\xi)$, $\xi \in \mathbb{R}$. At the boundary $z = \xi \in \mathbb{R}$ the equation $L\phi = \xi^2\phi$ is invariant under complex conjugation. We can write

$$\phi_l(x,\xi) = a(\xi)\overline{\phi_r(x,\xi)} + b(\xi)\phi_r(x,\xi)$$

where we normalize by

$$\lim_{x\to-\infty} e^{i\xi x}\phi_l(x,\xi) = 1$$

The interpretation is that $\overline{\phi_r}$ describes the incoming wave, ϕ_r the reflected and ϕ_l the transmitted wave.

In our situation the Jost solutions can be defined in the closed upper halfplane with the normalization

$$\lim_{x\to-\infty} e^{i\xi x}\phi_l(x,\xi) = 1$$

The Wronskian $W(\phi_l, \overline{\phi_l})$ is constant and has the limit $2i\xi$ at $-\infty$. Thus

$$\begin{aligned} 2i\xi &= \lim_{x\to\infty} W(a(\xi)\overline{\phi_r(x,\xi)} + b(\xi)\phi_r(x,\xi), \overline{a(\xi)\overline{\phi_r(x,\xi)} + b(\xi)\phi_r(x,\xi)}) \\ &= 2i\xi(|a(\xi)|^2 - |b(\xi)|^2) \end{aligned}$$

hence

$$1 = |a(\xi)|^2 - |b(\xi)|^2.$$

We call $T(z) = a(z)^{-1}$ the transmission coefficient and

$$R(\xi) = a(\xi)^{-1}b(\xi)$$

the reflection coefficient. We obtain the relation

$$|T(\xi)|^2 + |R(\xi)|^2 = 1$$

or equivalently the matrix

$$\begin{pmatrix} a(\xi) & \bar{b}(\xi) \\ b(\xi) & \bar{a}(\xi) \end{pmatrix} \in SU(1,1).$$

4.6.1 The Lax pair and the KdV equation

The Lax pair for the KdV equation is

$$L\phi = (-\partial^2 + u) = z^2\phi$$

$$P\phi = (-4\partial^3 + 3(\partial u + u\partial))\phi$$

so that the Korteweg-de Vries equation arises as compatibility condition for

$$L\phi = z^2\phi \qquad L_t = [P, L] \quad \Longleftrightarrow \quad [\partial_t - P, L] = 0. \tag{4.19}$$

If we want to solve the two equation

$$L\phi = z^2\phi \quad (\partial_t - P)\phi = 0$$

simultaneously we need a more flexible variant of the Jost solutions. Recall that we assume $\int(1 + |x|)|u|dx < \infty$ so that we can set

$$\phi^l(z, t, x) = \kappa(t, z)\phi_l(z, t, x)$$

with $\kappa(0, z) = 1$ where we normalize

$$\lim_{x\to-\infty} e^{izx}\phi_l(z, t, x) = 0.$$

We assume fast decay of u or even compact support so that

$$\phi^l(z, t, x) = \kappa(t, z)e^{-izx}$$

for x close to $-\infty$. Similarly

$$\phi^r(z, t, x) = \kappa_r(t, z)e^{izx}$$

for x near ∞ with $\kappa_r(0, z) = 1$. Then

$$\kappa_t + 4(-iz)^3\kappa = 0$$

and

$$\kappa = e^{-4iz^3t}.$$

Similarly

$$\kappa_r = e^{4iz^3t}$$

Now we turn to $z = \xi \in \mathbb{R}$. Then, as above, with the standard normalization

$$\phi_l(t, x; \xi) = a(\xi)\overline{\phi_r(t, x, \xi)} + be^{8i\xi^3t}\phi_r(t, x, \xi).$$

Similarly at eigenvalues $-\tau_j^2$

$$\phi_r(t,x,i\tau_j) = \gamma_j(t)\phi_r(t,x,i\tau_j)$$

where

$$\gamma_j(t) = \gamma_j(0)e^{-8i(i\tau_j)^3 t} = \gamma_j(0)e^{-8\tau_j^3 t}.$$

The inverse scattering approach consists in

1. Study the map
$$u_0 \to (R(0,\xi), \tau_j, \gamma_j(0))$$

2. evolve by the linear equations

3. Study
$$(R(t,\xi), \tau_j(t), \gamma_j(t)) \to u(t)$$

Observe that b determines $|T|$ and hence $|a|$. If there is no eigenvalue then $\log a(z)$ is a holomorphic function with real part $\log|a|$. Then on the real line

$$\operatorname{Im}\log a = H \operatorname{Re}\log a$$

where H is the Hilbert transform and hence a is determined by $|b|$ on the real line. Slightly more work is needed for the general case.

4.6.2 Scattering for the Lax operator

Recall

$$H = -\partial^2 + u$$

as usual and $H_0 = -\partial^2$. It is a selfadjoint operator, which by Stone's theorem defines a unitary group e^{-itH} by

$$i\partial_t \phi = H\phi.$$

The Möller resp. wave operators are motivated by the following question: Let $\phi_0 \in L^2$ and $\phi(t) = e^{-tH_0}\phi_0$. Does there exists ψ_0 so that

$$\lim_{t\to\pm\infty} \|e^{-itH}\psi_0 - e^{-itH_0}\phi_0\|_{L^2} = 0?$$

The answer is yes and it is given by

$$\Omega_\pm \phi = \lim_{t\to\pm\infty} e^{itH}e^{-itH_0}\phi_0$$

The Möller operators

$$\Omega_\pm : L^2 \to L^2$$

are isometric and the range is the orthogonal complement of the span of the eigenvalues. The scattering operator is

$$S\phi = \Omega_+^{-1}\Omega_-\phi.$$

At least formally

$$\frac{d}{dt}e^{itH}e^{-itH_0} = ie^{itH}(H - H_0)e^{itH_0} = ie^{itH}ue^{-itH_0}$$

hence

$$\Omega_+ = \phi(t) + i\int_0^\infty e^{itH}v\phi(t)dt = \phi(t) + i\lim_{0<\varepsilon\to 0}\int_0^\infty e^{itH-\varepsilon t}v\phi(t)dt.$$

By the inverse Fourier transform we write

$$\phi = \frac{1}{\sqrt{2\pi}}\int a(p)e^{ipx}dp$$

and

$$\phi(t) = \frac{1}{\sqrt{2\pi}}\int a(p)e^{ipx-ip^2t}dp$$

so that

$$\begin{aligned}\phi &= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} a(p)\Big(e^{ixp} + i\lim_{\varepsilon\to 0}\int_0^\infty e^{it(H-|p|^2)-\varepsilon t}ue^{ipx}\Big)dtdp\\ &= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} a(p)\Big(e^{ixp} - i(H - p^2 + i0)^{-1}ue^{ixp}\Big)dp\end{aligned}$$

we arrive at (formally, but this can be justified)

$$\psi(p) = e^{ixp} - \lim_{\varepsilon\to 0} i(H - p^2 + i\varepsilon)^{-1}ue^{ixp}$$

which can be seen by looking at the assumptotics to be

$$\phi_r(p) = a^{-1}(p)\overline{\phi_l(p)} - \frac{b(p)}{a(p)}\overline{\phi_r(p)}$$

It is a interpretated as a wave ϕ_r coming in from the right, with $\overline{\phi_l}$ being the transmitted partd and $\overline{\phi_r(p)}$ the reflected part.

References

[1] V. I. Arnold. *Mathematical methods of classical mechanics*. Vol. 60. Graduate Texts in Mathematics. Translated from the 1974 Russian original by K. Vogtmann and A. Weinstein, Corrected reprint of the second (1989) edition. Springer-Verlag, New York, [1989?], pp. xvi+516.

[2] Olivier Babelon, Denis Bernard, and Michel Talon. *Introduction to classical integrable systems*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, Cambridge, 2003, pp. xii+602.

[3] J. L. Bona and R. Smith. "The initial-value problem for the Korteweg-de Vries equation". In: *Philos. Trans. Roy. Soc. London Ser. A* 278.1287 (1975), pp. 555–601.

[4] J. Bourgain. "Fourier transform restriction phenomena for certain lattice subsets and applications to nonlinear evolution equations. II. The KdV-equation". In: *Geom. Funct. Anal.* 3.3 (1993), pp. 209–262.

[5] Ana Cannas da Silva and Alan Weinstein. *Geometric models for noncommutative algebras*. Vol. 10. Berkeley Mathematics Lecture Notes. American Mathematical Society, Providence, RI; Berkeley Center for Pure and Applied Mathematics, Berkeley, CA, 1999, pp. xiv+184.

[6] P. A. Deift. *Orthogonal polynomials and random matrices: a Riemann-Hilbert approach*. Vol. 3. Courant Lecture Notes in Mathematics. New York University, Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI, 1999, pp. viii+273.

[7] Pavel Etingof. *Calogero-Moser systems and representation theory*. Zurich Lectures in Advanced Mathematics. European Mathematical Society (EMS), Zürich, 2007, pp. x+92.

[8] Ludwig D. Faddeev and Leon A. Takhtajan. *Hamiltonian methods in the theory of solitons*. English. Classics in Mathematics. Translated from the 1986 Russian original by Alexey G. Reyman. Springer, Berlin, 2007, pp. x+592.

[9] Laurent Guieu and Claude Roger. *L'Algèbre et le Groupe de Virasoro*. Vol. PM027. Les Publications CRM (Centre de Recherches Mathématiques de Montréal), 2007.

[10] N. J. Hitchin, G. B. Segal, and R. S. Ward. *Integrable systems*. Vol. 4. Oxford Graduate Texts in Mathematics. Twistors, loop groups, and Riemann surfaces, Lectures from the Instructional Conference held at the University of Oxford, Oxford, September 1997. The Clarendon Press, Oxford University Press, New York, 1999, pp. x+136.

[11] T. Kappeler and P. Topalov. “Global wellposedness of KdV in $H^{-1}(\mathbb{T},\mathbb{R})$”. In: *Duke Math. J.* 135.2 (2006), pp. 327–360.

[12] D. Kazhdan, B. Kostant, and S. Sternberg. “Hamiltonian group actions and dynamical systems of Calogero type”. In: *Comm. Pure Appl. Math.* 31.4 (1978), pp. 481–507.

[13] Carlos E. Kenig, Gustavo Ponce, and Luis Vega. “A bilinear estimate with applications to the KdV equation”. In: *J. Amer. Math. Soc.* 9.2 (1996), pp. 573–603.

[14] Carlos E. Kenig, Gustavo Ponce, and Luis Vega. “Well-posedness of the initial value problem for the Korteweg-de Vries equation”. In: *J. Amer. Math. Soc.* 4.2 (1991), pp. 323–347.

[15] Boris Khesin and Robert Wendt. *The geometry of infinite-dimensional groups*. Vol. 51. Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics]. Springer-Verlag, Berlin, 2009, pp. xii+304.

[16] Rowan Killip and Monica Vişan. “KdV is well-posed in H^{-1}”. In: *Ann. of Math. (2)* 190.1 (2019), pp. 249–305.

[17] Bertram Kostant. “The solution to a generalized Toda lattice and representation theory”. In: *Adv. in Math.* 34.3 (1979), pp. 195–338.

[18] Jerrold E. Marsden and Tudor S. Ratiu. *Introduction to mechanics and symmetry*. Second. Vol. 17. Texts in Applied Mathematics. A basic exposition of classical mechanical systems. Springer-Verlag, New York, 1999, pp. xviii+582.

[19] Pierre van Moerbeke. “Random matrix theory and integrable systems”. In: *The Oxford handbook of random matrix theory*. Oxford Univ. Press, Oxford, 2011, pp. 198–230.

[20] Luc Molinet. “A note on ill posedness for the KdV equation”. In: *Differential Integral Equations* 24.7-8 (2011), pp. 759–765.

[21] Luc Molinet. “Sharp ill-posedness results for the KdV and mKdV equations on the torus”. In: *Adv. Math.* 230.4-6 (2012), pp. 1895–1930.

[22] J. Moser. *Integrable Hamiltonian systems and spectral theory*. Lezioni Fermiane. [Fermi Lectures]. Scuola Normale Superiore, Pisa, 1983, pp. iv+85.

[23] Jürgen Moser. “Finitely many mass points on the line under the influence of an exponential potential–an integrable system”. In: *Dynamical systems, theory and applications (Rencontres, Battelle Res. Inst., Seattle, Wash., 1974)*. Lecture Notes in Phys., Vol. 38. Springer, Berlin, 1975, pp. 467–497.

[24] Jeffery Praught and Roman G. Smirnov. “Andrew Lenard: a mystery unraveled”. In: *SIGMA Symmetry Integrability Geom. Methods Appl.* 1 (2005), Paper 005, 7.

[25] Peter Cornelis Schuur. *Asymptotic analysis of soliton problems*. Vol. 1232. Lecture Notes in Mathematics. An inverse scattering approach. Springer-Verlag, Berlin, 1986, pp. viii+180.

[26] Alan Weinstein. “The local structure of Poisson manifolds”. In: *J. Differential Geom.* 18.3 (1983), pp. 523–557.

[27] Wikipedia. *Kepler problem — Wikipedia, The Free Encyclopedia.* `http://en.wikipedia.org/w/index.php?title=Kepler%20problem&oldid=1040898601`. [Online; accessed 08-November-2022]. 2022.