Finite and Infinite Dimensional Lie Groups And Evolution Equations

(Symmetries, Conservation Laws,
And Integrable Systems)

Michael E. Taylor

Contents

Chapter IV. The Korteweg-deVries Equation

Chapter V. The Camassa-Holm equation

Introduction

These notes are based on a semester course on integrable systems given at UNC. They explore various ways in which Lie groups can be used to study evolution equations arising in mathematical physics. We treat a variety of equations, both ODE and PDE, but the emphasis is on PDE, and in such cases the Lie groups involved are infinite dimensional.

In Chapter I we consider equations of geodesic motion on a Lie group, finite or infinite dimensional, endowed with a right (or left) invariant metric tensor. A general program is set out to describe the geodesic in terms of the solution to an evolution equation for a curve on the Lie algebra of the group. Quite a number of very interesting PDE arise in this setting. Perhaps the most fundamental is the Euler equation for ideal incompressible fluid flow. Various other examples are also produced in this chapter, including the Korteweg-deVries equation and the Camassa-Holm equation.

Chapter II discusses Hamiltonian systems. The equations derived in Chapter I can be recast in Hamiltonian form, yielding an evolution equation for a curve in the dual of the Lie algebra. For this discussion it is useful to work with Poisson structures, which are more general than symplectic structures. The dual of a Lie algebra has a natural Lie-Poisson structure, and also a variety of other Poisson structures, including particularly "frozen Poisson structures," choices of which lead to "Poisson pairs." The equations arising from Chapter I have a Hamiltonian form with respect to the Lie-Poisson structure, and in addition a number of them are Hamiltonian with respect to another element of a Poisson pair, yielding a "bi-Hamiltonian" structure. When this holds, there is a program that yields a sequence of conservation laws, which frequently presents the evolution equation as an "integrable system." As is well known, the full Euler equations for ideal incompressible fluid flow do not form an integrable system, but a number of other equations derived in Chapter I do.

Chapter III is to some degree a digression. We extend the setting of Chapter I to Lie groups equipped with both a right invariant metric tensor and a potential. One family of physical examples is considered, namely spinning tops. One very classical family of symmetric spinning tops is shown to yield integrable Hamiltonian systems, though without invoking the bi-Hamiltonian technology mentioned above.

Chapters IV and V explore in further detail two of the PDE that arose in Chapter I, namely the Korteweg-deVries (KdV) equation and the Camassa-Holm (CH) equation. Both have a bi-Hamiltonian form, which can be seen in terms of the Lie-Poisson structure and a frozen Poisson structure (different in the two cases) on the dual of the Virasoro algebra, which is a central extension of the Lie algebra of vector fields on the circle. We also discuss various different (though ultimately related) methods to construct sequences of conservation laws, involving Lax pairs and isospectral families of linear operators. In the case of KdV we also discuss the method of Gel'fand and Dickii to produce systematically heirarchies of Hamiltonian systems in involution, a program that has been used to produce many additional such heirarchies of integrable systems. We make use of conservation laws to establish global existence of smooth solutions to KdV. We also discuss how such global existence sometimes works and sometimes breaks down for CH.

Analytical and geometric background for these notes can be found in the author's PDE text [T3]. This includes basic geometric background, given in Chapters 1–2 and Appendices B and C of [T3], as well as techniques for analyzing nonlinear evolution equations developed in the third volume, and material on pseudodifferential operators presented in Chapter 7 of [T3]. Basic concepts regarding Lie groups and their Lie algebras, in the finite-dimensional case, can be found in [T4].

Chapter 1: Geodesic Flows on Groups

Introduction

This first chapter is devoted to the equations of geodesic flow on a Lie group, equipped with a right (or left) invariant metric tensor. One of its focal points involves the curious phenomenon that a remarkable number of partial differential equations that have arisen to describe fluid motion also arise as equations for geodesics on various infinite-dimensional Lie groups, endowed with right-invariant metric tensors.

The premier example is the Euler equation for ideal incompressible fluid flow, which turns out to be the geodesic equation on the group of volume-preserving diffeomorphisms of a Riemannian manifold M, with a right-invariant metric whose value at the tangent space to the identity element of this group, which consists of divergence-free vector fields on M, is given by the L^2-inner product. Other examples involve equations that arise to describe approximate fluid motions, via formal nonlinear geometrical optics. These include the Korteweg-deVries equation and the Camassa-Holm equation, and also the cubic nonlinear Schrödinger equation, the latter via a transform of a special case of the Landau-Lifschitz equation.

We begin with a general study of the geodesic equation on a Lie group G, with a right-invariant metric tensor. This leads to an "Euler equation" for a curve in the Lie algebra $\mathfrak{g}$ of G. We show how this yields a familiar treatment of the free motion of a rigid body in $\mathbb{R}^n$, in terms of a geodesic motion on $SO(n)$. We then move on to infinite-dimensional Lie groups, starting with $\mathrm{Diff}(M)$, the group of diffeomorphisms of a compact manifold M, then considering subgroups, in particular the group of volume-preserving diffeomorphisms. We proceed to look at various general constructions, involving central extensions, semidirect products, and gauge groups. We see that special cases of these constructions give rise to various PDE mentioned in the previous paragraph.

Of these derivations, only the derivation of the equation of ideal incompressible flow and that of rigid body motion are directly parallel to the standard physical derivation. The main interest in the fact that these other PDE arise as geodesic equations on Lie groups rests on the resulting mathematical structure. This is manifested most clearly in terms of a transformation to an evolution equation for a curve in $\mathfrak{g}^*$. Two phenomena have the potential to come into play, though usually one is dominant. One is that these curves in $\mathfrak{g}^*$ lie in coadjoint orbits. The other is that the evolution equation might be Hamiltonian in two different ways. Either one of these phenomena has the potential to generate important conservation laws. These ideas will be developed further in subsequent chapters.

1. Geodesic equation for right (or left) invariant metrics

Let G be a Lie group (later we will allow G to be infinite dimensional), endowed with a right-invariant metric tensor. We want to study the geodesic equation. A geodesic $u : [a, b] \to G$ is a stationary point of the energy functional

$$I(u) = \int_a^b L(u(t), u'(t))\, dt, \tag{1.1}$$

where

$$L(p, v) = B(vp^{-1}, vp^{-1}), \quad p \in G,\ v \in T_pG, \tag{1.2}$$

$B(\cdot,\cdot)$ being an inner product on $T_eG = \mathfrak{g}$, the Lie algebra of G. Here we (somewhat informally) picture G as a group of elements in an algebra $\mathcal{A}$ of linear transformations on a vector space, so that right multiplication by $p^{-1} \in G$ maps T_pG to T_eG, e being the identity element. All the products below are products in $\mathcal{A}$. Clearly matrix groups have a natural structure of this sort. In other cases, one can have G act on a sufficiently rich space of functions. This approach serves to lighten the notation in the calculations below.

The standard Lagrange equation for a stationary point of (1.1) is

$$\frac{d}{dt} D_vL(u, u_t) = D_pL(u, u_t). \tag{1.3}$$

When $L(p, v)$ is given by (1.2), we have

$$\begin{aligned} D_pL(p, v)W &= -2B(vp^{-1}Wp^{-1}, vp^{-1}), \\ D_vL(p, v)W &= 2B(Wp^{-1}, vp^{-1}), \end{aligned} \tag{1.4}$$

with $W \in T_pG$. We see that

$$\frac{d}{dt} D_vL(u, u_t)W = -2B(Wu^{-1}u_tu^{-1}, u_tu^{-1}) + 2B(Wu^{-1}, u_{tt}u^{-1} - u_tu^{-1}u_tu^{-1}). \tag{1.5}$$

Thus the Lagrange equation (1.3) becomes

$$\begin{aligned} -B(Wu^{-1}u_tu^{-1}, u_tu^{-1}) &+ B(Wu^{-1}, u_{tt}u^{-1} - u_tu^{-1}u_tu^{-1}) \\ &= -B(u_tu^{-1}Wu^{-1}, u_tu^{-1}), \quad \forall\, W \in T_uG. \end{aligned} \tag{1.6}$$

We now set

$$v(t) = u_t(t)u(t)^{-1}, \quad v : (a, b) \to \mathfrak{g}. \tag{1.7}$$

Hence $u_t = vu$, $u_{tt} = v_t u + v u_t$, and, with $Y = Wu^{-1} \in \mathfrak{g}$, the equation (1.6) yields

$$-B(Yv, v) + B(Y, v_t) = -B(vY, v), \quad \forall\, Y \in \mathfrak{g}, \tag{1.8}$$

or equivalently we obtain the following equation for $v(t)$:

$$B(Y, v_t) + B([v, Y], v) = 0, \quad \forall\, Y \in \mathfrak{g}. \tag{1.9}$$

We call (1.9) the Euler equation for the geodesic flow. Note that if one solves (1.9) for $v(t)$, then the geodesic $u(t)$ is obtained as a solution to the (generally non-autonomous) linear equation

$$u_t(t) = v(t)u(t). \tag{1.10}$$

Another way to write (1.9) is

$$v_t + \mathrm{ad}_B^* v\,(v) = 0, \tag{1.11}$$

where we refer to the adjoint of the linear map $\mathrm{ad}\, v : \mathfrak{g} \to \mathfrak{g}$, taken with respect to the inner product $B(\cdot, \cdot)$. Here $\mathrm{ad}\, v(X) = [v, X]$.

Remark 1. One has a similar calculation for a left-invariant metric. Say instead of (1.2) we take

$$L(p, v) = B(p^{-1}v, p^{-1}v). \tag{1.12}$$

Analogues of (1.4)–(1.6) hold, and if we replace (1.7) by

$$v(t) = u(t)^{-1}u_t(t), \tag{1.13}$$

then a slight variant of (1.8) holds and yields the Euler equation

$$B(Y, v_t) - B([v, Y], v) = 0, \quad \forall\, Y \in \mathfrak{g}, \tag{1.14}$$

which differs from (1.9) only in a sign.

Remark 2. As is well known, a critical point of (1.1)–(1.2) is a *constant-speed* geodesic. We can directly verify this from (1.9) as follows. If $v : I \to \mathfrak{g}$ solves (1.9), then

$$\frac{d}{dt} B(v, v) = 2B(v_t, v) = -2B([v, v], v) = 0, \tag{1.15}$$

the second identity obtained by taking $Y = v$ in (1.9). Hence, given a solution to (1.9),

$$B(v(t), v(t)) = C_1, \tag{1.16}$$

with C_1 independent of t.

2. The reductive case, and the coadjoint picture

Suppose $\mathfrak{g}$ has a non-degenerate, Ad-invariant, symmetric bilinear form $Q(\cdot,\cdot)$, so

$$Q(\operatorname{ad} v\ X, Y) = -Q(X, \operatorname{ad} v\ Y), \quad \forall\ X, Y, v \in \mathfrak{g}. \tag{2.1}$$

Given another bilinear form $B(\cdot,\cdot)$, as in §1, we can write

$$B(u,v) = Q(u, Av), \quad A : \mathfrak{g} \to \mathfrak{g}. \tag{2.2}$$

Then the Euler equation (1.9) becomes

$$\begin{aligned} Q(Y, Av_t) &= -Q(\operatorname{ad} v(Y), Av) \\ &= Q(Y, [v, Av]), \end{aligned} \tag{2.3}$$

for all $Y \in \mathfrak{g}$, or simply

$$Av_t = [v, Av]. \tag{2.4}$$

Alternatively, if we set

$$w(t) = Av(t), \tag{2.6}$$

we get the equation

$$w_t = [A^{-1}w, w]. \tag{2.7}$$

One implication of (2.7) is that, if $t_0 \in (a,b)$, then

$$w(t) \in \mathcal{O}_{w(t_0)} = \{\operatorname{Ad}(g)\, w(t_0) : g \in G\}, \quad \forall\ t \in (a,b). \tag{2.8}$$

In other words, a solution to (2.7) must lie in an adjoint orbit.

Example. Consider $\mathfrak{g} = \mathrm{so}(n)$, the space of skew-symmetric, real, $n \times n$ matrices, with

$$Q(u,v) = \operatorname{Tr}(uv^t) = -\operatorname{Tr}(uv). \tag{2.9}$$

Let us consider

$$B(u,v) = \operatorname{Tr}(uRv^t), \quad R : \mathbb{R}^n \to \mathbb{R}^n, \tag{2.10}$$

with R positive-definite. Then (2.2) holds with

$$A(u) = \frac{1}{2}(uR + Ru), \tag{2.11}$$

the skew-symmetric part of uR.

While compact Lie groups and semisimple Lie groups have non-degenerate, symmetric bilinear forms satisfying (2.1), a vast array of Lie groups do not. Hence it is useful to have the following more general reformulation of the Euler equation (1.9). The inner product $B(\cdot,\cdot)$ used in §1 produces an isomorphism

$$\beta : \mathfrak{g} \longrightarrow \mathfrak{g}^*, \quad B(u,v) = \langle u, \beta v\rangle. \tag{2.12}$$

Then the Euler equation (1.9) is equivalent to

$$w_t - \mathrm{ad}^*(v)\, w = 0, \tag{2.13}$$

for

$$w(t) = \beta v(t), \tag{2.14}$$

and with the coadjoint representation of $\mathfrak{g}$ on $\mathfrak{g}^*$ given by $\mathrm{ad}^*(v)w = -(\mathrm{ad}\, v)^* w$. In other words, we get the differential equation for a path in $\mathfrak{g}^*$:

$$w_t - \mathrm{ad}^*(\beta^{-1} w) w = 0. \tag{2.15}$$

Parallel to (2.8), we have

$$w(t) \in \mathcal{O}^*_{w(t_0)} = \{\mathrm{Ad}^*(g) w(t_0) : g \in G\}, \quad \forall\, t \in (a,b). \tag{2.16}$$

In other words, a solution to (2.15) must lie in a coadjoint orbit.

Remark. The result (2.8) and its more general counterpart (2.16) amount to conservation laws. One particular conservation law that holds when (2.8) applies is that

$$Q(w(t), w(t)) = C_2, \tag{2.17}$$

independent of t. In other words,

$$Q(Av(t), Av(t)) = C_2. \tag{2.18}$$

Compare this conservation law with (1.16), which in this context is

$$Q(Av(t), v(t)) = C_1. \tag{2.19}$$

3. Rigid body motion in $\mathbb{R}^n$ and geodesics on $SO(n)$

Suppose there is a rigid body in $\mathbb{R}^n$, with a mass distribution at $t = 0$ given by a function $\rho(x)$, which we will assume is piecewise continuous and has compact support. Suppose the body moves, subject to no external forces, only the constraint of being rigid; we want to describe the motion of the body. According to the Lagrangian approach to mechanics, we seek an extremum of the integrated kinetic energy, subject to this constraint.

If $\xi(t, x)$ is the position in $\mathbb{R}^n$ at time t of the point on the body whose position at time 0 is x, then we can write the Lagrangian as

$$I(\xi) = \frac{1}{2} \int_{t_0}^{t_1} \int\limits_{\mathbb{R}^n} \rho\big(\xi(t,x)\big) |\dot{\xi}(t,x)|^2 \; dx \; dt. \tag{3.1}$$

Here, $\dot{\xi}(t,x) = \partial\xi/\partial t$.

Using center of mass coordinates, we will assume that the center of mass of the body is at the origin, and its total linear momentum is zero, so

$$\xi(t,x) = W(t)x, \quad W(t) \in SO(n), \tag{3.2}$$

where $SO(n)$ is the group of rotations of $\mathbb{R}^n$. Thus, describing the motion of the body becomes the problem of specifying the curve $W(t)$ in $SO(n)$. We can write (3.1) as

$$\begin{aligned} I(\xi) &= \frac{1}{2} \int_{t_0}^{t_1} \int\limits_{\mathbb{R}^n} \rho\big(W(t)x\big) |W'(t)x|^2 \; dx \; dt \\ &= \frac{1}{2} \int_{t_0}^{t_1} \int\limits_{\mathbb{R}^n} \rho(y) |W'(t)W(t)^{-1}y|^2 \; dy \; dt \\ &= J(W). \end{aligned} \tag{3.3}$$

We look for an extremum, or other critical point, where we vary the family of paths $W : [t_0, t_1] \to SO(n)$ (keeping the endpoints fixed).

Let us reduce the formula (3.3) for $J(W)$ to a single integral, over t. In fact, we have the following.

Lemma 3.1. *If A and B are real $n \times n$ matrices, i.e., belong to $M(n, \mathbb{R})$, then*

$$\int \rho(y) \; (Ay, By) \; dy = Tr(B^t A \mathcal{I}_\rho) = Tr(A \mathcal{I}_\rho B^t), \tag{3.4}$$

where

$$\mathcal{I}_\rho = \int \rho(y)\ y \otimes y\ dy \in \bigotimes^2 \mathbb{R}^n \approx M(n, \mathbb{R}). \tag{3.5}$$

Proof. It suffices to note that

$$(Ay, By) = \operatorname{Tr}\left(A(y \otimes y)B^t\right).$$

Recall that we are assuming that the body's center of mass is at 0 and its total linear momentum vanishes (at $t = 0$), i.e.,

$$\int \rho(y) y\ dy = 0. \tag{3.6}$$

By (3.4), we can write the Lagrangian (3.3) as

$$\begin{aligned} J(W) &= \frac{1}{2} \int_{t_0}^{t_1} \operatorname{Tr}\left(W'(t)W(t)^{-1}\mathcal{I}_\rho(W'(t)W(t)^{-1})^t\right)\ dt \\ &= \frac{1}{2} \int_{t_0}^{t_1} \operatorname{Tr}\left(Z(t)\mathcal{I}_\rho Z(t)^t\right) dt, \end{aligned} \tag{3.7}$$

where

$$Z(t) = W'(t)W(t)^{-1}. \tag{3.8}$$

Note that, if $W : (t_0, t_1) \to SO(n)$ is smooth, then

$$Z(t) \in \operatorname{so}(n), \quad \forall\ t \in (t_0, t_1). \tag{3.9}$$

where $\operatorname{so}(n)$ is the set of real antisymmetric $n \times n$ matrices, the Lie algebra of $SO(n)$; in particular $\operatorname{so}(n)$ is the tangent space to $SO(n) \subset M(n, \mathbb{R})$ at the identity element.

Clearly we are in the situation discussed in §§1–2, with

$$Q(u, v) = \operatorname{Tr}(uv^t), \quad B(u, v) = \operatorname{Tr}(u\mathcal{I}_\rho v^t). \tag{3.10}$$

Hence $W(t)$ is a geodesic on $SO(n)$, and the Euler equation (via (2.4)) takes the form

$$\mathcal{L}_\rho Z' = [Z, \mathcal{L}_\rho Z], \tag{3.11}$$

where (as in (2.11))

$$\mathcal{L}_\rho : \operatorname{so}(n) \to \operatorname{so}(n), \quad \mathcal{L}_\rho(U) = \frac{1}{2}(U\mathcal{I}_\rho + \mathcal{I}_\rho U). \tag{3.12}$$

Alternatively, if we set

$$M(t) = \mathcal{I}_\rho Z(t) + Z(t)\mathcal{I}_\rho. \tag{3.13}$$

Then the Euler equation is equivalent to

$$\frac{dM}{dt} = -[M, Z]. \tag{3.14}$$

In case $n = 3$, we have an isomorphism $\kappa : \mathbb{R}^3 \to \mathrm{so}(3)$ given by

$$\kappa(\omega_1, \omega_2, \omega_3) = \begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix}, \tag{3.15}$$

so that the cross product on $\mathbb{R}^3$ satisfies $\omega \times x = Ax,\ A = \kappa(\omega)$. Then the vector-valued function

$$\omega(t) = -\kappa^{-1} Z(t) \tag{3.16}$$

is called the angular velocity of the body.

The isomorphism κ has the properties

$$\kappa(x \times y) = [\kappa(x), \kappa(y)], \quad \mathrm{Tr}\,\big(\kappa(x)\kappa(y)^t\big) = 2x \cdot y. \tag{3.17}$$

Furthermore, a calculation shows that, if $\mathcal{L}_\rho$ is given by (3.12), then

$$\kappa^{-1}\mathcal{L}_\rho \kappa(\omega) = \frac{1}{2}\mathcal{J}_\rho \omega, \quad \mathcal{J}_\rho = (\mathrm{Tr}\ \mathcal{I}_\rho)I - \mathcal{I}_\rho. \tag{3.18}$$

Hence, with $M(t) \in \mathrm{so}(3)$ defined by (3.13), we have

$$\mu(t) = -\kappa^{-1} M(t) = \mathcal{J}_\rho \omega(t), \tag{3.19}$$

the first identity defining $\mu(t) \in \mathbb{R}^3$. The equation (3.14) is then equivalent to

$$\frac{d\mu}{dt} = -\omega \times \mu. \tag{3.20}$$

The vector $\mu(t)$ is called the angular momentum of the body, and $\mathcal{J}_\rho$ is called the inertia tensor. The equation (3.20) is the standard form of Euler's equation for the free motion of a rigid body in $\mathbb{R}^3$.

Note that $\mathcal{J}_\rho$ is a positive definite 3×3 matrix. Let us choose a positively oriented orthonormal basis of $\mathbb{R}^3$ consisting of eigenvectors of $\mathcal{J}_\rho$, say $\mathcal{J}_\rho e_j = J_j e_j$. Then, if $\omega = (\omega_1, \omega_2, \omega_3)$, we have $\mu = (J_1\omega_1, J_2\omega_2, J_3\omega_3)$, and

$$\omega \times \mu = \big((J_3 - J_2)\omega_2\omega_3, (J_1 - J_3)\omega_1\omega_3, (J_2 - J_1)\omega_1\omega_2\big).$$

Hence, (3.20) takes the form

$$
\begin{aligned}
J_1\dot\omega_1 + (J_3 - J_2)\omega_2\omega_3 &= 0,\\
J_2\dot\omega_2 + (J_1 - J_3)\omega_1\omega_3 &= 0,\\
J_3\dot\omega_3 + (J_2 - J_1)\omega_1\omega_2 &= 0.
\end{aligned} \tag{3.21}
$$

If we multiply the ℓth line in (3.21) by $\dot\omega_\ell$ and sum over ℓ, we get $(d/dt)(J_1\omega_1^2+J_2\omega_2^2+J_3\omega_3^2) = 0$, while if instead we multiply by $J_\ell\dot\omega_\ell$ and sum, we get $(d/dt)(J_1^2\omega_1^2 + J_2^2\omega_2^2 + J_3^2\omega_3^2) = 0$. Thus we have the conserved quantities

$$
\begin{aligned}
J_1\omega_1^2 + J_2\omega_2^2 + J_3\omega_3^2 &= C_1,\\
J_1^2\omega_1^2 + J_2^2\omega_2^2 + J_3^2\omega_3^2 &= C_2.
\end{aligned} \tag{3.22}
$$

If any of the quantities J_ℓ coincide, the system (3.21) simplifies. If, on the other hand, we assume that $J_1 < J_2 < J_3$, then we can write the system (3.21) as

$$
\dot\omega_2 = \beta^2\omega_1\omega_3, \quad \dot\omega_1 = -\alpha^2\omega_2\omega_3, \quad \dot\omega_3 = -\gamma^2\omega_1\omega_2, \tag{3.23}
$$

where

$$
\alpha^2 = \frac{J_3 - J_2}{J_1}, \quad \beta^2 = \frac{J_3 - J_1}{J_2}, \quad \gamma^2 = \frac{J_2 - J_1}{J_3}. \tag{3.24}
$$

If we then set

$$
\zeta_1 = \alpha\omega_2, \quad \zeta_2 = \beta\omega_1, \quad \zeta_3 = \alpha\beta\omega_3, \tag{3.25}
$$

this system becomes

$$
\dot\zeta_1 = \zeta_2\zeta_3, \quad \dot\zeta_2 = -\zeta_1\zeta_3, \quad \dot\zeta_3 = -\gamma^2\zeta_1\zeta_2. \tag{3.26}
$$

For this system we have conserved quantities

$$
\zeta_1^2 + \zeta_2^2 = c_1, \quad \gamma^2\zeta_1^2 + \zeta_3^2 = c_2, \tag{3.27}
$$

a fact which is equivalent to (3.22). (We mention that arranging that $J_1 < J_2 < J_3$ might change the orientation, hence the sign in (3.20).)

Note that we can use (3.27) to decouple the system (3.26), obtaining

$$
\begin{aligned}
\dot\zeta_1 &= \big[(c_1 - \zeta_1^2)(c_2 - \gamma^2\zeta_1^2)\big]^{1/2}\\
\dot\zeta_2 &= -\big[(c_1 - \zeta_2^2)(c_2 - \gamma^2 c_1 + \zeta_2^2)\big]^{1/2}\\
\dot\zeta_3 &= -\big[(c_2 - \zeta_3^2)(c_1 - \gamma^{-2}c_2 + \gamma^{-2}\zeta_3^2)\big]^{1/2}
\end{aligned} \tag{3.28}
$$

Thus ζ_j are given by elliptic integrals; cf. [Lawd].

Remark. Note that the conservation laws in (3.22) are special cases of (2.18)–(2.19).

4. Geodesics on Diff(M)

Here we see how the Euler equation derived in §1 looks when $G = \mathrm{Diff}(M)$, the group of diffeomorphisms of M. We assume M is a compact manifold, endowed with a Riemannian metric. The Lie algebra is

$$\mathfrak{g} = \mathrm{Vect}(M), \tag{4.1}$$

the space of real vector fields on M. The natural Lie bracket on $\mathfrak{g}$ is the *negative* of the standard lie bracket of vector fields on M, i.e., $-[u, v]$.

For $B(\cdot,\cdot)$, we consider bilinear forms of the form

$$B(u, v) = (Au, v)_{L^2} = \int_M g(Au, v)\, dV. \tag{4.2}$$

The operator A is taken to be some positive-definite, self-adjoint operator acting on vector fields (usually a differential operator), $g(\cdot,\cdot)$ is the metric tensor on M, and dV its volume element.

The Euler equation (1.9) has the form

$$B(v_t, Y) = B(v, \mathcal{L}_v Y), \quad \forall\, Y \in \mathrm{Vect}(M), \tag{4.3}$$

where $\mathcal{L}_v$ is the Lie derivative (cf. [T3], Chapter 1, §8). Using (4.2), we rewrite this as

$$(Av_t, Y)_{L^2} = (Av, \mathcal{L}_v Y)_{L^2} = (\mathcal{L}_v^* Av, Y)_{L^2},$$

so the Euler equation becomes

$$Av_t = \mathcal{L}_v^* Av, \tag{4.4}$$

or

$$v_t = A^{-1}\mathcal{L}_v^* Av. \tag{4.5}$$

Here $\mathcal{L}_v^*$ is the L^2-adjoint of $\mathcal{L}_v$, which can be computed as follows. Since $\int g(w, u)\, dV$ is invariant when all relevant objects are pulled back appropriately by diffeomorphisms $M \mapsto M$, we have

$$\begin{aligned} \int_M g(w, \mathcal{L}_v u)\, dV = &-\int_M g(\mathcal{L}_v w, u)\, dV - \int_M (\mathcal{L}_v g)(u, w)\, dV \\ &- \int_M g(w, u)\, (\mathcal{L}_v\, dV). \end{aligned} \tag{4.6}$$

Note that

$$(\mathcal{L}_v g)(u,w) = 2\, g(\mathrm{Def}(v)u, w), \tag{4.7}$$

where $\mathrm{Def}(v)$ is the deformation tensor associated with a vector field v, and

$$\mathcal{L}_v(dV) = (\mathrm{div}\, v)\, dV, \tag{4.8}$$

where $\mathrm{div}\, v$ is the divergence of v (cf. [T3], Chapter 2, §3). Hence (as is well known),

$$\mathcal{L}_v^* = -\mathcal{L}_v - Tv, \tag{4.9}$$

with

$$Tv = 2\ \mathrm{Def}\ v + (\mathrm{div}\, v)I. \tag{4.10}$$

Going back to the Euler equation (4.5), we can rewrite it as

$$v_t + A^{-1}(\mathcal{L}_v + Tv)Av = 0, \tag{4.11}$$

or, alternatively, using $A^{-1}BA = B + A^{-1}[B, A]$, and $\mathcal{L}_v v = 0$, we have

$$v_t + (Tv)v + A^{-1}[\mathcal{L}_v, A]v + A^{-1}[Tv, A]v = 0, \tag{4.12}$$

with Tv given by (4.10).

Example. Suppose A is the identity operator in (4.2), i.e., $B(u,v)$ is simply the L^2-inner product. Then all the commutators in (4.12) vanish, and one has the following first-order, quasi-linear equation:

$$v_t + 2(\mathrm{Def}\, v)v + (\mathrm{div}\, v)v = 0. \tag{4.13}$$

Here is another way to write the Euler equation. Since

$$\mathcal{L}_v w = \nabla_v w - \nabla_w v, \quad 2(\mathrm{Def}\, v)w = \nabla_w v + (\nabla v)^t \cdot w, \tag{4.14}$$

we have

$$(\mathcal{L}_v + Tv)w = \nabla_v w + (\nabla v)^t \cdot w + (\mathrm{div}\, v)w, \tag{4.15}$$

so the Euler equation (4.4) becomes

$$Av_t + \nabla_v Av + (\nabla v)^t \cdot Av + (\mathrm{div}\, v)Av = 0. \tag{4.16}$$

5. The Camassa-Holm equation

Here we specialize the setting of §4 to $M = S^1$, the circle, where we have a canonical identification

$$\mathrm{Vect}(S^1) \approx C^\infty(S^1), \quad f \leftrightarrow f(x)\,\partial_x, \tag{5.1}$$

the Lie bracket given as usual by

$$[f\,\partial_x, g\,\partial_x] = (fg' - f'g)\partial_x. \tag{5.2}$$

We take

$$A = I - \partial_x^2, \tag{5.3}$$

so the inner product $B(u,v)$ is hence

$$B(u,v) = \int_{S^1} (uv + u'v')\,dx. \tag{5.4}$$

In this case the calculation of $\mathcal{L}_v^*$ is elementary; one has

$$\begin{aligned}(\mathcal{L}_v^* u, w)_{L^2} = (u, \mathcal{L}_v w)_{L^2} &= (u, vw_x - v_x w)_{L^2} \\ &= (-v_x u - (vu)_x, w)_{L^2},\end{aligned} \tag{5.5}$$

and hence

$$\mathcal{L}_v^* u = -vu_x - 2v_x u, \tag{5.6}$$

which is seen to be a special case of (4.9)–(4.10). Hence the Euler equation (4.4) takes the form

$$Av_t + v(Av)_x + 2v_x(Av) = 0, \tag{5.7}$$

or, using (5.3),

$$(1 - \partial_x^2)v_t + 3vv_x - 2v_x v_{xx} - vv_{xxx} = 0. \tag{5.8}$$

Another way to write this is as

$$(1 - \partial_x^2)(v_t + vv_x) + \partial_x\Big(v^2 + \frac{1}{2}v_x^2\Big) = 0. \tag{5.9}$$

This is the Camassa-Holm equation (cf. [CH]). Similar derivations are given in [Ko], [Mis2].

6. Geodesics on $\mathrm{Diff}_{\mathfrak{H}}(M)$

Here we look at the Euler equation for geodesics on a subgroup of $\mathrm{Diff}(M)$, whose Lie algebra $\mathfrak{H}$ is a Lie subalgebra of $\mathrm{Vect}(M)$. We consider

$$B(u,v) = (Au,v)_{L^2} = \int_M g(Au,v)\,dV, \tag{6.1}$$

as in (4.2). Let P denote the orthogonal projection of (the L^2 completion of) $\mathrm{Vect}(M)$ onto (the L^2 completion of) $\mathfrak{H}$. Then the Euler equation becomes

$$(Av_t, Y)_{L^2} = (Av, \mathcal{L}_v Y)_{L^2} = (P\mathcal{L}_v^* Av, Y)_{L^2}, \tag{6.2}$$

for all $Y \in \mathfrak{H}$, or

$$Av_t = P\mathcal{L}_v^* Av. \tag{6.3}$$

Assume A and P *commute*, so (6.3) becomes

$$v_t = PA^{-1}\mathcal{L}_v^* Av, \tag{6.4}$$

or, parallel to (4.12),

$$v_t + P\Big((Tv)v + A^{-1}[\mathcal{L}_v, A]v + A^{-1}[Tv, A]v\Big) = 0, \tag{6.5}$$

where, as in (4.10), we have

$$Tv = 2\ \mathrm{Def}\, v + (\mathrm{div}\, v)I. \tag{6.6}$$

An alternative formula for $(Tv)v$ is

$$(Tv)v = \nabla_v v + \frac{1}{2}\nabla|v|^2 + (\mathrm{div}\, v)v. \tag{6.7}$$

Example. Take $A = I$, so (6.1) is simply the L^2-inner product. Then the Euler equation (6.5) becomes

$$v_t + P\Big(\nabla_v v + \frac{1}{2}\nabla|v|^2 + (\mathrm{div}\, v)v\Big) = 0. \tag{6.8}$$

For another form of the Euler equations, we can use (4.9) and (4.14) to write (6.3) as

$$Av_t + P\big(\nabla_v Av + (\nabla v)^t \cdot Av + (\operatorname{div} v)Av\big) = 0. \tag{6.9}$$

7. Ideal incompressible fluid flow

An incompressible fluid flow on a compact Riemannian manifold M defines a one-parameter family of volume-preserving diffeomorphisms

$$F(t,\cdot) : M \longrightarrow M. \tag{7.1}$$

The flow can be described in terms of its velocity field

$$u(t,y) = F_t(t,x), \quad y = F(t,x), \tag{7.2}$$

where $F_t(t,x) = (\partial/\partial t)F(t,x)$. We assume the fluid has uniform density and derive Euler's equation for the dynamics of the fluid flow.

If we suppose there are no external forces acting on the fluid, the dynamics are determined by the constraint condition, that $F(t,\cdot)$ perserve volume, or equivalently that $\operatorname{div} u(t,\cdot) = 0$ for all t. The Lagrangian involves the kinetic energy alone, so we seek to find critical points of

$$L(F) = \int_a^b \int_M g(F_t(t,x), F_t(t,x))\, dV\, dt, \tag{7.3}$$

on the space of maps $F : (a,b) \times M \to M$ with the volume-preserving property.

To compare (7.3) with the Lagrangian (1.1)–(1.2) for geodesics in $\operatorname{Diff}_{\mathfrak{H}}(M)$ when

$$\mathfrak{H} = \{v \in \operatorname{Vect}(M) : \operatorname{div} v = 0\}, \tag{7.4}$$

and $B(u,v) = (u,v)_{L^2}$, let us note that for $F(t) : M \to M$ we have

$$F'(t)F(t)^{-1} = \frac{d}{ds} F(t+s) \circ F(t)^{-1}\big|_{s=0}, \tag{7.5}$$

and in particular, for $y \in M$, $F'(t)F(t)^{-1}(y) \in T_yM$ is given by

$$F'(t)F(t)^{-1}(y) = \frac{d}{ds} F(t+s,x)\big|_{s=0,F(t,x)=y} = u(t,y), \tag{7.6}$$

for $u(t,y)$ as in (7.2). Hence the Lagrangian (1.1) takes the form

(7.7)
$$\begin{aligned} I(F) &= \int_a^b \int_M g(F'(t)F(t)^{-1}, F'(t)F(t)^{-1})\, dV(y)\, dt \\ &= \int_a^b \int_M g(u(t,y), u(t,y))\, dV(y)\, dt \\ &= \int_a^b \int_M g(F_t(t,x), F_t(t,x))\, dV(x)\, dt, \end{aligned}$$

the last identity using the volume-preserving property of $F(t)$. Hence the problem of finding critical points of (7.3) is precisely the problem of finding geodesics on the group $\mathrm{Diff}_{\mathfrak{H}}(M)$, with $\mathfrak{H}$ given by (7.4) and with $A = I$ in (6.1). Hence we have the following special case of (6.8):

(7.8)
$$v_t + P\nabla_v v = 0, \quad \operatorname{div} v = 0.$$

Here P is the Helmholtz projection of vector fields onto divergence-free vector fields. This is the standard Euler equation for ideal incompressible fluid flow.

7B. Lagrange averaged Euler equations

Here we take $\mathfrak{H}$ as in (7.4) but consider the Euler equations from §6 when A is not the identity but rather

(7B.1)
$$A = I - \alpha^2 \Delta.$$

For simplicity, take $M = \mathbb{T}^n$, and let Δ act componentwise on vector fields. Then A and P, the Helmholtz projection defined above, commute. Furthermore, (6.9) becomes

(7B.2)
$$Av_t + P\big(\nabla_v Av + (\nabla v)^t \cdot Av\big) = 0,$$

since $\operatorname{div} v = 0$. See [Sh2] for a discussion and further references.

8. Geodesics on central extensions

Let G be a Lie group (perhaps infinite dimensional) and $\widetilde{G}$ a central extension, with Lie algebra

(8.1)
$$\tilde{\mathfrak{g}} = \mathfrak{g} \oplus \mathbb{R},$$

having a Lie bracket of the form

$$[(u,a),(v,b)] = ([u,v],\gamma(u,v)). \tag{8.2}$$

Here

$$\gamma : \mathfrak{g} \times \mathfrak{g} \longrightarrow \mathbb{R} \tag{8.3}$$

is a skew-symmetric bilinear form. In order for (8.2) to be a Lie bracket, it must satisfy the Jacobi identity, which imposes on γ the following "cocycle" condition:

$$\gamma([u,v],w) + \gamma([v,w],u) + \gamma([w,u],v) = 0, \quad \forall\ u,v,w \in \mathfrak{g}. \tag{8.4}$$

Not all Lie groups have nontrivial central extensions, but a variety of important examples arise, as we will see.

Let us put on $\tilde{\mathfrak{g}}$ an inner product of the form

$$\mathcal{B}((u,a),(v,b)) = B(u,v) + ab, \tag{8.5}$$

with $B(\cdot,\cdot)$ as in §1. The Euler equation for geodesic flow on $\widetilde{G}$ is the following for $(v,c) : I \to \tilde{\mathfrak{g}}$:

$$\mathcal{B}((v_t,c_t),(Y,y)) + \mathcal{B}((v,c),[(v,c),(Y,y)]) = 0, \quad \forall (Y,y) \in \tilde{\mathfrak{g}}. \tag{8.6}$$

Using (8.2) and (8.5), we can rewrite this as

$$B(v_t,Y) + c_t y + B(v,[v,Y]) + c\gamma(v,Y) = 0, \quad \forall\ (Y,y) \in \tilde{\mathfrak{g}}. \tag{8.7}$$

Setting $Y = 0$ in (8.7) yields $c_t = 0$, i.e., $c \equiv c_0$, a constant. Then the equation for $v(t)$ becomes

$$B(v_t,Y) + B(v,[v,Y]) = -c_0\gamma(v,Y), \quad \forall\ Y \in \mathfrak{g}. \tag{8.8}$$

Equivalently, if as in §2 we use

$$w(t) = \beta v(t), \quad \beta : \mathfrak{g} \to \mathfrak{g}^*, \quad B(u,v) = \langle u, \beta v\rangle, \tag{8.9}$$

and we also take

$$\kappa : \mathfrak{g} \to \mathfrak{g}^*, \quad \gamma(v,Y) = \langle Y, \kappa(v)\rangle, \tag{8.10}$$

we obtain

$$w_t + \mathrm{ad}^*(v)w = -c_0\kappa(v), \tag{8.11}$$

or

$$w_t + \mathrm{ad}^*(\beta^{-1}w)w = -c_0\,\kappa\circ\beta^{-1}w. \tag{8.12}$$

The curve $w(t)$ does not lie in a coadjoint orbit of $\mathfrak{g}^*$, typically, though of course $(w(t), c_0)$ does lie in a coadjoint orbit of $\tilde{\mathfrak{g}}^* = \mathfrak{g}^* \oplus \mathbb{R}$.

Example. Let $G = \mathbb{R}^{2n}$, with the standard additive structure, so $\mathfrak{g} = \mathbb{R}^{2n}$, with the trivial Lie bracket. There is a central extension of G known as the Heisenberg group, whose Lie algebra $\mathfrak{h}^n = \mathbb{R}^{2n} \oplus \mathbb{R}$ has the Lie bracket

$$[(v_1, v_2, a), (w_1, w_2, b)] = (0, v_1\cdot w_2 - v_2\cdot w_1), \tag{8.13}$$

where $v_j, w_j \in \mathbb{R}^n$ and $v_j \cdot w_k$ is the standard dot product of vectors in $\mathbb{R}^n$. In this case, since G is abelian, $\mathrm{ad}^* \equiv 0$, and (8.12) takes the form

$$w_t = -c_0\,\kappa\circ\beta^{-1}w, \tag{8.14}$$

with solution

$$w(t) = e^{-c_0 t E}w(0), \quad E = \kappa\circ\beta^{-1} \in \mathrm{End}(\mathfrak{g}^*) = \mathrm{End}(\mathbb{R}^{2n}). \tag{8.15}$$

Note that if $B(\cdot,\cdot)$ is a positive-definite inner product on $\mathfrak{g} = \mathbb{R}^{2n}$, then E is skew-adjoint with respect to the inner product induced on $\mathfrak{g}^* = \mathbb{R}^{2n}$.

9. The Virasoro group and KdV

The group $\mathrm{Diff}(S^1)$ has a central extension called the Virasoro group, whose Lie algebra

$$\mathrm{Vir}(S^1) = \mathrm{Vect}(S^1) \oplus \mathbb{R} \approx C^\infty(S^1) \oplus \mathbb{R} \tag{9.1}$$

is given by (8.2), with the cocycle

$$\gamma(u, v) = (u', v'')_{L^2} = \int_{S^1} u'(x)v''(x)\,dx. \tag{9.2}$$

It is straightforward to verify the cocycle condition (8.4) in this case. The Euler equation for a curve in $\mathrm{Vir}(S^1)$ yields $(v(t), c_0)$, with $c_0 \in \mathbb{R}$ and $v(t)$ satisfying (8.8). If we take

$$B(u, v) = (u, v)_{L^2} = \int_{S^1} u(x)v(x)\,dx, \tag{9.3}$$

the equation (8.8) becomes

$$(v_t, Y)_{L^2} - (\mathcal{L}_v^* v, Y)_{L^2} = c_0(v''', Y)_{L^2}, \quad \forall\, Y \in C^\infty(S^1), \tag{9.4}$$

which in view of the formula

$$\mathcal{L}_v^* v = -3vv_x, \tag{9.5}$$

derived in (5.6), yields the equation

$$v_t + 3vv_x - c_0 v_{xxx} = 0, \tag{9.6}$$

known as the Korteweg-deVries equation.

10. Geodesics on semidirect products

Here we consider the semidirect product $H = G \times_\varphi V$ of a group G with a vector space V, on which there is a G-action, $\varphi : G \to \mathrm{Aut}(V)$. We have a Cartesian product, $H = G \times V$, with the group law

$$(g_1, v_1)(g_2, v_2) = (g_1 g_2, v_1 + \varphi(g_1)v_2). \tag{10.1}$$

The Lie algebra of H is $\mathfrak{h} = \mathfrak{g} \oplus V$, as a vector space, with Lie bracket

$$[(u, x), (v, y)] = ([u, v], \psi(u)y - \psi(v)x), \tag{10.2}$$

where $\psi = d\varphi$ is the derived representation of $\mathfrak{g}$ on V.

Let us take an inner product on $\mathfrak{h}$ of the form

$$\mathcal{B}((u, x), (v, y)) = B(u, v) + Q(x, y), \tag{10.3}$$

where $B(\cdot, \cdot)$ is an inner product on $\mathfrak{g}$, as in §1, and $Q(\cdot, \cdot)$ an inner product on V. Then the Euler equation associated with the geodesic flow on H, with right invariant metric, is the following equation for $(v, x) : I \to \mathfrak{g} \oplus V$:

$$\mathcal{B}((v_t, x_t), (Y, y)) + \mathcal{B}((v, x), [(v, x), (Y, y)]) = 0, \quad \forall\, (Y, y) \in \mathfrak{g} \oplus V. \tag{10.4}$$

Using (10.2) and (10.3), we can rewrite this as

$$B(v_t, Y) + B(v, [v, Y]) = -Q(x_t, y) - Q(x, \psi(v)y - \psi(Y)x), \tag{10.5}$$

for all $(Y, y) \in \mathfrak{g} \oplus V$. Setting $Y = 0$ yields

$$Q(x_t, y) + Q(x, \psi(v)y) = 0, \quad \forall\, y \in V, \tag{10.6}$$

or

$$x_t = -\psi(v)^* x, \tag{10.7}$$

where $\psi(v)^* : V \to V$ is the adjoint of $\psi(v)$ with respect to the inner product $Q(\cdot,\cdot)$. We then reduce (10.5) to (10.7) plus

$$B(v_t, Y) + B(v, [v, Y]) = Q(x, \psi(Y)x), \quad \forall\, Y \in \mathfrak{g}. \tag{10.8}$$

Now we can define

$$\psi_Q^B(x) \in \mathfrak{g}, \quad B(\psi_Q^B(x), Y) = Q(x, \psi(Y)x), \tag{10.9}$$

a quadratic function of $x \in V$. Then the equation (10.8) can be rewritten as

$$B(v_t, Y) + B(v, [v, Y]) = B(\psi_Q^B(x), Y), \quad \forall\, Y \in \mathfrak{g}, \tag{10.10}$$

which is coupled to (10.7). In the terminology of (1.11), we can write this coupled system as

$$\begin{aligned} v_t + \mathrm{ad}_B^*\, v\,(v) &= \psi_Q^B(x), \\ x_t &= -\psi(v)^* x. \end{aligned} \tag{10.11}$$

Example. Take $V = \mathfrak{g}^*$, $\varphi = \mathrm{Ad}^*$, i.e.,

$$H = G \times_{\mathrm{Ad}^*} \mathfrak{g}^*. \tag{10.12}$$

Also, having an inner product $B(\cdot,\cdot)$ on $\mathfrak{g}$, let $Q(\cdot,\cdot)$ be the inner product induced on $\mathfrak{g}^*$, i.e.,

$$Q(x, y) = \langle \beta^{-1} x, y \rangle, \tag{10.13}$$

where, as in (2.12),

$$\beta : \mathfrak{g} \to \mathfrak{g}^*, \quad B(u, v) = \langle u, \beta v \rangle. \tag{10.14}$$

We also use

$$u = \beta^{-1} x, \quad z = \beta v, \tag{10.15}$$

with (v, x) as in (10.11). The equation (10.7) for x is equivalent to

$$\langle \beta^{-1} x_t, y \rangle = -\langle \beta^{-1} x, (\mathrm{ad}^*\, v) y \rangle,$$

for all $y \in \mathfrak{g}^*$, hence to

$$u_t = [v, u]. \tag{10.16}$$

Meanwhile we have

$$\begin{aligned} \langle \beta\psi_Q^B(x), Y\rangle &= B(\psi_Q^B(x), Y) = Q(x, \psi(Y)x) \\ &= \langle \beta^{-1}x, \mathrm{ad}^*(Y)x\rangle = -\langle [\beta^{-1}x, Y], x\rangle \\ &= -\langle Y, \mathrm{ad}^*(\beta^{-1}x)x\rangle, \end{aligned} \tag{10.17}$$

or

$$\beta\psi_Q^B(x) = -\,\mathrm{ad}^*(\beta^{-1}x)x. \tag{10.18}$$

Hence (10.10) can be written

$$\langle Y, \beta v_t\rangle + \langle Y, (\mathrm{ad}^*\, v)\beta v\rangle = \langle Y, \mathrm{ad}^*(\beta x)x\rangle, \quad \forall\, Y \in \mathfrak{g}, \tag{10.19}$$

so coupled to (10.16) we have

$$z_t + (\mathrm{ad}^*\, v)z = (\mathrm{ad}^*\, u)x. \tag{10.20}$$

It is also useful to realize that when H has the form (10.12), its Lie algebra $\mathfrak{h} = \mathfrak{g} \oplus \mathfrak{g}^*$, with Lie bracket

$$[(u,\xi),(u,\eta)] = ([u,v], \mathrm{ad}^*(u)\,\eta - \mathrm{ad}^*(v)\,\xi), \tag{10.21}$$

has a nondegenerate, Ad-invariant bilinear form:

$$\mathcal{Q}((u,\xi),(v,\eta)) = \langle u,\eta\rangle + \langle v,\xi\rangle, \tag{10.22}$$

where $\langle\cdot,\cdot\rangle$ is the natural $\mathfrak{g}\times\mathfrak{g}^*$ pairing. Indeed, straightforward computations give

$$\begin{aligned} \mathcal{Q}(\mathrm{ad}(w,\zeta)(u,\xi),(v,\eta)) &= -\langle u,\psi(w)\eta\rangle + \langle v,\psi(w)\xi\rangle - \langle v,\psi(u)\zeta\rangle, \\ \mathcal{Q}((u,\xi),\mathrm{ad}(w,\zeta)(v,\eta)) &= \langle u,\psi(w)\eta)\rangle - \langle u,\psi(v)\zeta\rangle - \langle v,\psi(w)\xi\rangle, \end{aligned} \tag{10.23}$$

with $\psi = \mathrm{ad}^*$, and the fact that the first quantity in (10.23) is equal to the negative of the second comes down to

$$-\langle v,\psi(u)\zeta\rangle = \langle \mathrm{ad}(u)v,\zeta\rangle = -\langle \mathrm{ad}(v)u,\zeta\rangle = \langle u,\psi(v)\zeta\rangle. \tag{10.24}$$

11. Ideal incompressible MHD

Here we take

$$G = \mathrm{Diff}_{\mathfrak{H}}(M), \quad \mathfrak{g} = \mathfrak{H} = \{v \in \mathrm{Vect}(M) : \mathrm{div}\, v = 0\}, \tag{11.1}$$

as in (7.4), and consider the Euler equation for geodesic flow on

$$H = G \times_{\mathrm{Ad}^*} \mathfrak{g}^*. \tag{11.2}$$

This is an evolution equation for $(v, \xi) : I \to \mathfrak{g} \oplus \mathfrak{g}^*$. It turns out to be equivalent to the system of equations for a velocity field v and a magnetic field B, knows as the equations of magnetohydrodynamics, for an ideal incompressible fluid interacting with a magnetic field.

The inner product we place of $\mathfrak{g}$ is the L^2-inner product:

$$B(u, v) = (u, v)_{L^2} = \int_M g(u, v)\, dV. \tag{11.3}$$

We use this inner product to "identify" $\mathfrak{g}^*$ with $\mathfrak{g}$. Denote by B the image of ξ under this identification. Formally, parallel to (10.15), we have $B = \beta^{-1}\xi$.

We write out the coupled system (10.7), (10.10) in this context. The equation (10.10) becomes

$$(v_t, Y)_{L^2} - (v, \mathcal{L}_v Y)_{L^2} = -(\psi_Q^B(\xi), Y)_{L^2}, \quad \forall\, Y \in \mathfrak{g}, \tag{11.4}$$

or

$$v_t - P\mathcal{L}_v^* v = -P\psi_Q^B(\xi), \tag{11.5}$$

where P is the Helmholtz projection, as used in (7.5). Using the formulas (4.9)–(4.10) and (6.7), we can rewrite (11.5) as

$$v_t + P\nabla_v v = -P\psi_Q^B(\xi), \quad \mathrm{div}\, v = 0, \tag{11.6}$$

in parallel with (7.5). This is coupled to the following incarnation of (10.16):

$$B_t + \mathcal{L}_v B = 0, \quad \mathrm{div}\, B = 0. \tag{11.7}$$

As in (10.18), we have

$$\beta\psi_Q^B(\xi) = -\,\mathrm{ad}^*(B)\,\xi. \tag{11.8}$$

Alternatively, we have

$$\psi_Q^B(\xi) = \mathcal{L}_B^* B = -\nabla_B B - \frac{1}{2}\nabla|B|^2, \tag{11.9}$$

the first identity by (11.8) and the second by (4.9)–(4.10), plus (6.7) and the fact that $\operatorname{div} B = 0$.

The vector calculus identity $\nabla|u|^2 = 2\nabla_u u + 2u \times \operatorname{curl} u$ yields yet another alternative, in case $\dim M = 3$:

$$\psi_Q^B(\xi) = B \times \operatorname{curl} B - \nabla|B|^2. \tag{11.10}$$

Then the incompressible MHD equations take the standard form:

$$\begin{aligned} v_t + P\nabla_v v &= -P(B \times \operatorname{curl} B), \\ B_t + \mathcal{L}_v B &= 0, \\ \operatorname{div} v = \operatorname{div} B &= 0. \end{aligned} \tag{11.11}$$

12. Geodesics on gauge groups and the Landau-Lifschitz equation

Let G be a compact Lie group and M a compact Riemannian manifold. We consider the "gauge group" $C^\infty(M, G)$, which acts as a group of automorphisms on the trivial principal G-bundle $M \times G \to M$. The Lie algebra of $C^\infty(M, G)$ is

$$\mathfrak{h} = C^\infty(M, \mathfrak{g}). \tag{12.1}$$

Let us take a bi-invariant inner product $Q(\cdot, \cdot)$ on $\mathfrak{g}$ and a self-adjoint operator A, acting on real valued functions on M. We pick a basis of $\mathfrak{g}$ and let A act componentwise on $\mathfrak{g}$-valued functions. Then we set

$$B(u, v) = \int_M Q(Au(x), Av(x))\, dV(x). \tag{12.2}$$

The Euler equation (1.9) is

$$B(v_t, Y) + B(v, [v, Y]) = 0, \quad \forall\, Y \in \mathfrak{h}. \tag{12.3}$$

Here

$$\begin{aligned} B(v_t, Y) &= \int_M Q(Av_t(x), AY(x))\, dV(x) \\ &= \int_M Q(A^2 v_t, Y)\, dV, \end{aligned} \tag{12.4}$$

and

$$B(v,[v,Y]) = \int_M Q(Av, A[v,Y])\,dV$$
$$(12.5) \qquad = \int_M Q(A^2v,[v,Y])\,dV$$
$$= -\int_M Q([v,A^2v],Y)\,dV.$$

Hence the Euler equation for geodesic flow on $C^\infty(M,G)$ is

$$A^2 v_t = [v, A^2 v]. \tag{12.6}$$

Equivalently, for

$$w = A^2 v, \tag{12.7}$$

we have

$$w_t = [A^{-2}w, w]. \tag{12.8}$$

Example. A particularly important case of A is

$$A = (1-\Delta)^{-1/2}. \tag{12.9}$$

Then $A^{-2} = 1-\Delta$, and since $[(1-\Delta)w,w] = -[\Delta w, w]$, the equation (12.8) becomes

$$w_t = -[\Delta w, w]. \tag{12.10}$$

This is known as the Landau-Lifschitz equation.

13. From Landau-Lifschitz to cubic NLS

Let us specialize the setting of §12 to $M = S^1$. The Landau-Lifschitz equation (12.10) becomes

$$w_t = [w, w_{xx}], \tag{13.1}$$

for $w : \mathbb{R}\times S^1 \to \mathfrak{g}$, the Lie algebra of a compact Lie group G. We next specialize to $G = SO(3)$ and use the isomorphism (3.15) of so(3) with $\mathbb{R}^3$, with the cross product. Then the Landau-Lifschitz equation becomes

$$w_t = w \times w_{xx}, \tag{13.2}$$

for $w : \mathbb{R} \times S^1 \to \mathbb{R}^3$. Note that (13.2) can be written

$$w_t = \partial_x (w \times w_x). \tag{13.3}$$

The equation (13.2) is also known as the Heisenberg magnet equation.

Note that whenever $\gamma(t,x)$ solves

$$\gamma_t = \gamma_x \times \gamma_{xx} \tag{13.4}$$

then $w(t,x) = \partial_x \gamma(t,x)$ solves (13.2). It is interesting to regard $\gamma(t,x)$ as a 1-parameter family of curves $x \mapsto \gamma(t,x)$, parametrized by t, and (13.4) defines an evolution of this family of curves. Note that if (13.4) holds then

$$\partial_t (\gamma_x \cdot \gamma_x) = 2\gamma_x \cdot \gamma_{tx} = 2w \cdot w_t = 0, \tag{13.5}$$

so the "speed" $|\gamma_x(t,x)|$ is independent of t. In particular, if $x \mapsto \gamma(0,x)$ is a unit-speed curve, then so is $x \mapsto \gamma(t,x)$ for all t (for which (13.4) holds). Let us restrict attention to such an evolution of unit-speed curves. Thus it is suggestive to set $T(t,x) = \partial_x \gamma(t,x)$, to denote the unit tangent to $x \mapsto \gamma(t,x)$. We recall the Frenet-Serret formulas:

$$\begin{aligned} T_x &= \kappa N \\ N_x &= -\kappa T + \tau B \\ B_x &= - \tau N, \end{aligned} \tag{13.6}$$

with (T,N,B) an orthonormal frame such that $B = T \times N$. In particular, since

$$\gamma_x = T, \quad \gamma_{xx} = \kappa N, \tag{13.7}$$

the equation (13.4) is equivalent to

$$\gamma_t = \kappa B. \tag{13.8}$$

This is known as the "filament equation." It provides a crude model of the motion of a curve on which vorticity is concentrated in a 3D incompressible fluid.

R. Hasimoto made the following remarkable connection between the filament equation and the cubic nonlinear Schrödinger equation.

Proposition 13.1. *Let $\gamma(t,x)$ solve (13.4) and satisfy $|\gamma_x| \equiv 1$, and set*

$$\psi(t,x) = \kappa(t,x)\, e^{iM(t,x)}, \quad M(t,x) = \int_0^x \tau(t,s)\, ds + \mu(t), \tag{13.9}$$

where $\mu(t)$ will be specified below. Then ψ solves

$$i\psi_t + \psi_{xx} + \frac{1}{2}|\psi|^2\psi = 0. \tag{13.10}$$

To verify (13.10), we begin with $\psi_t = (\kappa_t + i\kappa M_t)e^{iM}$, etc., obtaining

$$i\psi_t + \psi_{xx} + \frac{1}{2}|\psi|^2\psi = F\psi, \tag{13.11}$$

with

$$F = -M_t + \frac{\kappa_{xx}}{\kappa} - \tau^2 + \frac{1}{2}\kappa^2 + \Big(\frac{\kappa_t}{\kappa} + 2\frac{\kappa_x\tau}{\kappa} + \tau_x\Big)i. \tag{13.12}$$

Thus (13.10) is equivalent to the following two identities:

$$\kappa_t + 2\kappa_x\tau + \tau_x\kappa = 0, \tag{13.13}$$

and

$$M_t = \frac{\kappa_{xx}}{\kappa} - \tau^2 + \frac{1}{2}\kappa^2. \tag{13.14}$$

If we specify $\mu(t)$ to satisfy

$$\mu'(t) = \frac{\kappa_{xx}(t,0)}{\kappa(t,0)} - \tau(t,0)^2 + \frac{1}{2}\kappa(t,0)^2,$$

then (13.14) is equivalent to

$$\tau_t = \partial_x\Big(\frac{\kappa_{xx}}{\kappa} - \tau^2 + \frac{1}{2}\kappa^2\Big). \tag{13.15}$$

To establish (13.13), we will compare the identity

$$\gamma_{tt} = \kappa_t B + \kappa B_t, \quad (B_t \perp B), \tag{13.16}$$

which follows from (13.8), with the identity

$$\gamma_{tt} = \gamma_{tx} \times \gamma_{xx} + \gamma_x \times \gamma_{txx}, \tag{13.17}$$

which follows from (13.4). To establish (13.15), we will compare the identity

$$B_{xt} = -\tau_t N - \tau N_t \quad (N_t \perp N), \tag{13.18}$$

which follows from the third formula in (13.6), with another formula for B_{xt}, which in turn is derived from the identity for B_t produced by comparing (13.16) and (13.17).

To compute the ingredients in (13.17), we have from (13.4) that

$$\begin{aligned} \gamma_{tx} &= \gamma_x \times \gamma_{xxx}, \\ \gamma_{txx} &= \gamma_{xx} \times \gamma_{xxx} + \gamma_x \times \gamma_{xxxx}. \end{aligned} \tag{13.19}$$

To obtain the ingredients in (13.19), we can differentiate (13.7), using the Frenet-Serret formulas (13.6) to obtain

$$\begin{aligned} \gamma_{xxx} &= -\kappa^2 T + \kappa_x N + \kappa\tau B, \\ \gamma_{xxxx} &= -(2\kappa_x + \kappa^2)T + (\kappa_{xx} - \kappa^3 - \kappa\tau^2)N + (2\kappa_x\tau + \kappa\tau_x)B. \end{aligned} \tag{13.20}$$

This leads to

$$\gamma_{tx} = -\kappa\tau N + \kappa_x B, \tag{13.21}$$

and

$$\gamma_{txx} = \kappa^2\tau T - (\kappa_x\tau + \kappa\tau_x)N + (\kappa_{xx} - \kappa\tau^2)B. \tag{13.22}$$

Plugging these formulas plus (13.7) into (13.17) gives

$$\gamma_{tt} = -\kappa\kappa_x T + (\kappa\tau^2 - \kappa_{xx})N - (2\kappa_x\tau + \kappa\tau_x)B. \tag{13.27}$$

Comparison with (13.16) yields the equation (13.13), from comparing coefficients of B, and it also yields

$$\kappa B_t = -\kappa\kappa_x T + (\kappa\tau^2 - \kappa_{xx})N. \tag{13.24}$$

Dividing by κ and taking the x-derivative yields

$$B_{tx} = -\kappa\tau^2 T - \Bigl(\frac{\kappa_{xx}}{\kappa} - \tau^2 + \frac{1}{2}\kappa^2\Bigr)_x N + \Bigl(\tau^3 - \frac{\tau}{\kappa}\kappa_{xx}\Bigr)B. \tag{13.25}$$

Comparing the coefficients of N in (13.18) and (13.25) gives the identity (13.15), and completes the proof of Proposition 13.1.

Remark. The integral of $\tau(t,x)$ over $x \in [0,2\pi]$ is not necessarily an integral multiple of 2π, so $\psi(t,x)$ is not generally periodic in x, even if $\gamma(t,x)$ is.

14. Geodesics on FIO(M)

We consider FIO(M), the group of unitary, zero-order, Fourier integral operators on a compact Riemannian manifold M. Its Lie algebra is

$$\Psi^1(M) = \{A \in OPS^1_{1,0}(M) : A^* = -A\}, \tag{14.1}$$

with the commutator as Lie bracket. Here $OPS^1_{1,0}(M)$ is the space of pseudo-differential operators of order 1 and type $(1,0)$. We use notation and results of [T].

We consider the following sort of inner product on $\Psi^1(M)$. Pick $k > n+2$, where $n = \dim M$, and pick a positive-definite elliptic operator $Q \in OPS^{-k}_{1,0}(M)$. Then set

$$B(A_1, A_2) = \mathrm{Tr}(A_1 Q A_2^*), \quad A_j \in \Psi^1(M). \tag{14.2}$$

Note that $A_1 Q A_2^* \in OPS^{-k+2}_{1,0}(M)$ is trace class. Then the Euler equation (1.9) takes the form

$$\mathrm{Tr}(v_t Q Y^*) + \mathrm{Tr}(vQ(Y^* v^* - v^* Y^*)) = 0, \quad \forall\, Y \in \Psi^1(M). \tag{14.3}$$

Let us temporarily assume $k > n+3$. Under this hypothesis on Q, vQY^*v^* and vQv^*Y^*, members of $OPS^{-k+3}_{1,0}(M)$, are both trace class. We claim that

$$\mathrm{Tr}(vQY^*v^*) = \mathrm{Tr}(v^*vQY^*). \tag{14.4}$$

In fact this is a special case of the following result.

Lemma 14.1. *If $A_j \in OPS^{m_j}_{1,0}(M)$ and $m_1 + m_2 < -n$ ($n = dim\, M$), then*

$$\mathrm{Tr}(A_1 A_2) = \mathrm{Tr}(A_2 A_1). \tag{14.5}$$

Proof. Let us assume $m_2 < 0$; if this does not hold then we must have $m_1 < 0$, and a similar argument will work with the roles of A_1 and A_2 reversed. Decomposing A_2 into self-adjoint and skew-adjoint parts, we see it suffices to prove (14.5) when A_2 is self-adjoint.

Taking an appropriate $\widetilde{A}_2$, positive definite and elliptic in $OPS^{m_2}_{1,0}(M)$, and replacing A_2 by $A_2 + a\widetilde{A}_2$, with large positive a, we can assume A_2 is elliptic, and positive-definite, so

$$A_2^r \in OPS^{rm_2}_{1,0}(M), \quad \text{for} \quad r \in \mathbb{R}. \tag{14.6}$$

Given that $m_2 < 0$, we can pick $k \in \mathbb{Z}^+$ and write $A_1 A_2 = A_1 A_2^{1-1/k} A_2^{1/k}$ with $m_1 + (1-1/k)m_2 < -n$. Then

$$\mathrm{Tr}(A_1 A_2) = \mathrm{Tr}(A_2^{1/k} A_1 A_2^{1-1/k}), \tag{14.7}$$

by Theorem 3.1 of [Si]. Now applying this sort of argument to $(A_2^{1/k} A_1) A_2^{1-1/k}$, we have

$$\mathrm{Tr}(A_1 A_2) = \mathrm{Tr}(A_2^{2/k} A_1 A_2^{1-2/k}),$$

and iterating this argument we arrive at the conclusion (14.5).

This lemma suffices to prove (14.4), under the hypothesis that $k > n+3$. Now we improve this result.

Lemma 14.2. *Assuming* $k > n+2$, $Q \in OPS^{-k}_{1,0}$, $v, Y \in OPS^1_{1,0}$, *we have*

$$\mathrm{Tr}(vQ[Y^*, v^*]) = \mathrm{Tr}([v^*, vQ]Y^*). \tag{14.8}$$

Proof. Let J_ε be a Friedrichs mollifier, $0 < \varepsilon \le 1$, and set $Q_\varepsilon = J_\varepsilon Q J_\varepsilon$. Then, by (14.4), (14.8) holds, with Q replaced by Q_ε, for each $\varepsilon > 0$. Now, as $\varepsilon \to 0$,

$$vQ_\varepsilon[Y^*, v^*] \to vQ[Y^*, v^*], \quad [v^*, vQ_\varepsilon]Y^* \to [v^*, vQ]Y^*, \tag{14.9}$$

in $OPS^{-k+2+\delta}_{1,0}(M)$, for each $\delta > 0$, and hence we have convergence in trace norm, provided $k > n+2$. This yields (14.8).

Using (14.8), we can rewrite the Euler equation (14.3) as

$$\mathrm{Tr}\big((v_t Q + [v^*, vQ])Y^*\big) = 0, \quad \forall\, Y \in \Psi^1(M), \tag{14.10}$$

which yields

$$v_t Q - Q v_t^* + [v^*, vQ] - [v^*, vQ]^* = 0, \tag{14.11}$$

or

$$v_t Q + Q v_t = [v, vQ + Qv]. \tag{14.12}$$

Note the formal similarity to the Euler equation (3.13)–(3.14) for geodesics on $SO(n)$. This is to be expected; it just takes a little more analysis to produce (14.12) in the current situation.

Chapter 2: Poisson Brackets and Hamiltonian Vector Fields

Introduction

Poisson structures are a class of structures that include symplectic structures as a special case. A Poisson structure on a manifold M gives rise to a Poisson bracket on functions on M, in turn yielding an association $f \mapsto H_f$ of a Hamiltonian vector field H_f to such a function. The dual $\mathfrak{g}^*$ of a Lie algebra gets a natural Poisson structure, called a Lie-Poisson structure, and this plays a central part in the role of Lie groups in differential equations.

We define Poisson structures in §1 and establish some basic properties, including a frequently useful criterion for testing whether the crucial Jacobi identity holds for a candidate for a Poisson structure. We proceed in §2 to discuss the Lie-Poisson structure. Section 3 discusses the special case of symplectic structures. The reader has likely come across symplectic structures before, and our treatment here is brief. More can be found in Chapter I of [T3], for example. In §4 we show how right-invariant Lagrangians on a Lie group lead to Hamiltonian equations on the dual to its Lie algebra, with its Lie-Poisson structure. This provides a generalization of the setting of Chapter I.

In §5 we discuss how the Lie-Poisson structure on the dual $\mathfrak{g}^*$ of a Lie algebra gives rise to other Poisson structures, namely shifted Poisson structures and frozen Poisson structures. From these we get "Poisson pairs," pairs of Poisson structures with a crucial compatibility property. Some Hamiltonian systems on $\mathfrak{g}^*$, with its Lie-Poisson structure are also Hamiltonian with respect to a frozen Poisson structure. Such systems are special cases of a class called bi-Hamiltonian systems, which we discuss in §6. Bi-Hamiltonian systems are often integrable, with a string of conservation laws generated by the Lenard scheme. We give a general discussion of this here. As we will see in Chapters IV and V, the Korteweg-deVries equation and the Camassa-Holm equation each have a bi-Hamiltonian structure, and one obtains an infinite sequence of conservation laws in each of these cases.

1. Poisson structures

A Poisson structure on a manifold M is a map assigning to $f, g \in C^\infty(M)$ a function $\{f, g\} \in C^\infty(M)$, satisfying the following four identities:

(1.1) (Anti-symmetry) $\{f, g\} = -\{g, f\},$

(1.2) ($\mathbb{R}$-linearity) $\{f, c_1 g_1 + c_2 g_2\} = c_1\{f, g_1\} + c_2\{f, g_2\},$

(1.3) (Leibniz rule) $\{f, g_1 g_2\} = g_1\{f, g_2\} + g_2\{f, g_1\},$

(1.4) (Jacobi identity) $\{\{f, g\}, h\} + \{\{g, h\}, f\} + \{\{h, f\}, g\} = 0,$

for all $f, g, h, g_1, g_2 \in C^\infty(M)$, $c_1, c_2 \in \mathbb{R}$.

Properties (1.1)–(1.3) imply that there is a vector field (called a Hamiltonian vector field) H_f associated to each $f \in C^\infty(M)$, such that

$$\{f, g\} = H_f g, \tag{1.5}$$

and that $f \mapsto H_f$ is $\mathbb{R}$-linear. Given this, (1.4) is equivalent to the identity

$$H_{\{f,g\}} = [H_f, H_g], \tag{1.6}$$

for all $f, g \in C^\infty(M)$.

The standard example is $M = \mathbb{R}^{2n}$, with coordinates $(x_1, \dots, x_n, \xi_1, \dots, \xi_n)$, and with Poisson bracket given by

$$\{f, g\} = \sum_{j=1}^{n} \Big(\frac{\partial f}{\partial \xi_j} \frac{\partial g}{\partial x_j} - \frac{\partial f}{\partial x_j} \frac{\partial g}{\partial \xi_j} \Big). \tag{1.7}$$

In this case it is routine to verify (1.1)–(1.4).

There are more subtle constructions of Poisson structures, such as the one we will see in §2, where it will be easy to verify properties (1.1)–(1.3) and more of a job to verify (1.4), or equivalently (1.6). The following result will prove to be a useful tool.

Proposition 1.1. *Assume (1.1)–(1.3) hold for* $\{\cdot, \cdot\}$. *Let* $\mathcal{L}$ *be a linear subspace of* $C^\infty(M)$. *Suppose (1.6) holds for all* $f, g \in \mathcal{L}$. *Then (1.6) holds for all* $f, g \in \mathcal{A}$, *the algebra generated by* $\mathcal{L}$.

Proof. It suffices to show that

$$H_{\{f,g_j\}} = [H_f, H_{g_j}] \Rightarrow H_{\{f,g_1 g_2\}} = [H_f, H_{g_1 g_2}]. \tag{1.8}$$

We are assuming (1.1)–(1.3). Note that $H_{g_1 g_2} u = \{g_1 g_2, u\} = g_1\{g_2, u\} + g_2\{g_1, u\}$, and hence

$$H_{g_1 g_2} = g_1 H_{g_2} + g_2 H_{g_1}. \tag{1.9}$$

Also, we generally have

$$[X, uY] = u[X, Y] + (Xu)Y, \tag{1.10}$$

for vector fields X, Y and $u \in C^\infty(M)$. Hence

$$[H_f, H_{g_1 g_2}] = g_1[H_f, H_{g_2}] + \{f, g_1\} H_{g_2} + g_2[H_f, H_{g_1}] + \{f, g_2\} H_{g_1}. \tag{1.11}$$

Meanwhile, using (1.3) on $\{f, g_1 g_2\}$ and applying (1.9) gives

$$\begin{aligned} H_{\{f,g_1 g_2\}} &= H_{g_1\{f,g_2\}} + H_{g_2\{f,g_1\}} \\ &= g_1 H_{\{f,g_2\}} + \{f, g_2\} H_{g_1} + g_2 H_{\{f,g_1\}} + \{f, g_1\} H_{g_2}. \end{aligned} \tag{1.12}$$

Now the identity of (1.11) and (1.12) follows from the hypothesis in (1.8), so the implication in (1.8) is established.

The following result implies that a Poisson structure is defined by a second-order tensor field.

Proposition 1.2. *If* $f_1, f_2 \in C^\infty(M)$ *and* $df_1(p) = df_2(p)$, *then* $H_{f_1}(p) = H_{f_2}(p) \in T_pM$.

Proof. Consider $f = f_1 - f_2$, so $df(p) = 0$. It follows easily from Taylor's formula with remainder that one can write, locally near p,

$$f(x) - f(p) = \sum_{j=1}^{n} g_j h_j, \quad g_j(p) = h_j(p) = 0, \tag{1.13}$$

with $g_j, h_j \in C^\infty(M)$. That $H_f = 0$ at p then follows from the analogue of (1.9).

Consequently, given a Poisson structure on M, there exists a linear map

$$\mathcal{J} : T_p^*M \longrightarrow T_pM, \tag{1.14}$$

depending smoothly on p, such that

$$H_f(p) = \mathcal{J} df(p), \quad \forall\, p \in M. \tag{1.15}$$

Thus $\mathcal{J}$ is a contravariant tensor field of order 2; we call it the Poisson tensor. We see from (1.1) that

$$\mathcal{J}^* = -\mathcal{J}. \tag{1.16}$$

Maps on M that preserve the Poisson bracket are called Poisson maps. The following records an important source of Poisson maps.

Proposition 1.3. *Let* M *have a Poisson structure. The flow* $\mathcal{F}_t$ *generated by a Hamiltonian vector field* $X = H_u$ *preserves this Poisson structure, i.e.,*

$$\mathcal{F}_t^*\{f, g\} = \{\mathcal{F}_t^* f, \mathcal{F}_t^* g\}, \tag{1.17}$$

where $\mathcal{F}_t^* f(x) = f(\mathcal{F}_t x)$.

Proof. The definition of $\mathcal{F}_t^* f$ yields

$$\frac{d}{dt}\mathcal{F}_t^* f = X\mathcal{F}_t^* f = \mathcal{F}_t^* X f. \tag{1.18}$$

Hence

$$\begin{aligned} &\frac{d}{dt}\mathcal{F}_{-t}^*\{\mathcal{F}_t^* f, \mathcal{F}_t^* g\} \\ &\quad = -\mathcal{F}_{-t}^* X\{\mathcal{F}_t^* f, \mathcal{F}_t^* g\} + \mathcal{F}_{-t}^*\{X\mathcal{F}_t^* g, \mathcal{F}_t^* g\} + \mathcal{F}_{-t}^*\{\mathcal{F}_t^* f, X\mathcal{F}_t^* g\}. \end{aligned} \tag{1.19}$$

Now, by (1.4),

$$X = H_u \Longrightarrow X\{\tilde{f}, \tilde{g}\} = \{X\tilde{f}, \tilde{g}\} + \{\tilde{f}, X\tilde{g}\}, \tag{1.20}$$

so the quantity (1.19) vanishes. This proves (1.17).

Generally we say a vector field X is an infinitesimal Poisson map if the flow it generates satisfies (1.17). The content of Proposition 1.3 is that any Hamiltonian vector field is an infinitesimal Poisson map. As the proof shows, a (more general) sufficient condition for X to be an infinitesimal Poisson map is that

$$X\{f,g\} = \{Xf,g\} + \{f,Xg\}, \tag{1.21}$$

for all $f,g \in C^\infty(M)$. This condition is also seen to be necessary.

2. Lie-Poisson structure on $\mathfrak{g}^*$

Let $\mathfrak{g}$ be a Lie algebra, so we have a bracket $[\cdot,\cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$, satisfying

$$[X,Y] = -[Y,X], \tag{2.1}$$

$$[X, c_1Y_1 + c_2Y_2] = c_1[X,Y] + c_2[X,Y_2], \tag{2.2}$$

$$[[X,Y],Z] + [[Y,Z],X] + [[Z,X],Y] = 0, \tag{2.3}$$

for all $X,Y,Z,Y_1,Y_2 \in \mathfrak{g}$, $c_1,c_2 \in \mathbb{R}$. The identity (1.3) is the Jacobi identity in this context. Another way to put it is the following. Given $X \in \mathfrak{g}$, define $\operatorname{ad} X : \mathfrak{g} \to \mathfrak{g}$ by

$$\operatorname{ad} X\,(Y) = [X,Y]. \tag{2.4}$$

Then, given (2.1)–(2.2), the identity (2.3) is equivalent to the identity

$$\operatorname{ad}[X,Y] = [\operatorname{ad} X, \operatorname{ad} Y], \tag{2.5}$$

where the right side of (2.5) denotes the commutator $(\operatorname{ad} X)(\operatorname{ad} Y) - (\operatorname{ad} Y)(\operatorname{ad} X)$.

Let $\mathfrak{g}^*$ denote the dual space of $\mathfrak{g}$. We define a Poisson bracket of functions in $C^\infty(\mathfrak{g}^*)$ as follows. Given $f : \mathfrak{g}^* \to \mathbb{R}$, $\xi \in \mathfrak{g}^*$, we have $df(\xi) : \mathfrak{g}^* \to \mathbb{R}$, linear, i.e., $df(\xi) \in \mathfrak{g}$. Using this, we set

$$\{f,g\}(\xi) = \langle [df(\xi), dg(\xi)], \xi \rangle, \tag{2.6}$$

$[\cdot,\cdot]$ denoting the Lie bracket on $\mathfrak{g}$.

We need to verify the properties (1.1)–(1.4) when $\{f,g\}$ is defined by (2.6). The identities (1.1)–(1.2) are obvious consequences of (2.1)–(2.2), and (1.3) follows readily from the identity

$$d(g_1g_2) = g_1\,dg_2 + g_2\,dg_1. \tag{2.7}$$

The proof of the Jacobi identity (1.4), or equivalently (1.6), is somewhat more subtle, but we can establish it fairly cleanly, using Proposition 1.1.

First, note that if $f : \mathfrak{g}^* \to \mathbb{R}$ is *linear*, then, for all $\xi \in \mathfrak{g}^*$, $df(\xi) = f \in \mathfrak{g}$, so

$$\begin{aligned} f, g \text{ linear} &\Longrightarrow \{f, g\}(\xi) = \langle [f, g], \xi \rangle \\ &\Longrightarrow \{f, g\} = [f, g]. \end{aligned} \tag{2.8}$$

Hence the fact that (1.4) holds for f, g, h linear follows directly from (2.3). This implies

$$f, g \text{ linear} \Longrightarrow H_{\{f,g\}} = [H_f, H_g], \tag{2.9}$$

at least as vector fields acting on linear functions h. But a vector field on a linear space is determined by its action on linear functions, so we have (2.9). Now we can apply Proposition 1.1 to deduce that (1.6) holds whenever f and g are polynomials. But the space of polynomials is dense in $C^\infty(\mathfrak{g}^*)$, so the validity of (1.6), hence of (1.4), for all $f, g, h \in C^\infty(\mathfrak{g}^*)$ follows by a limiting argument.

We now consider integral curves of a vector field H_f, defined by (1.5), when $\{\cdot, \cdot\}$ is the Poisson bracket on $C^\infty(\mathfrak{g}^*)$ given by (2.6). An integral curve $\xi(t)$ of H_f satisfies

$$\xi'(t) = H_f(\xi(t)). \tag{2.10}$$

Equivalently, given $X \in \mathfrak{g}$, defining a linear map $X : \mathfrak{g}^* \to \mathbb{R}$, we have

$$\begin{aligned} \frac{d}{dt} \langle X, \xi(t) \rangle &= \{f, X\}(\xi(t)) \\ &= \langle [df(\xi), X], \xi(t) \rangle \\ &= -\langle X, \operatorname{ad}^* \, df(\xi(t))\, \xi(t) \rangle, \end{aligned} \tag{2.11}$$

the last identity incorporating the definition of the coadjoint representation of $\mathfrak{g}$ on $\mathfrak{g}^*$:

$$\langle \operatorname{ad} Y(X), \xi \rangle = -\langle X, \operatorname{ad}^* Y(\xi) \rangle. \tag{2.12}$$

Hence the differential equation (2.10) takes the form

$$\xi'(t) = -\operatorname{ad}^*\bigl(df(\xi(t))\bigr)\xi(t). \tag{2.13}$$

From this we deduce that if (2.10) holds on an interval $I = (a, b)$ and $t_0 \in I$, then, for all $t \in I$,

$$\xi(t) \in \mathcal{O}^*_{\xi(t_0)} = \{\operatorname{Ad}^*(g)\, \xi(t_0) : g \in G\}, \tag{2.14}$$

where G is a Lie group with Lie algebra $\mathfrak{g}$, and Ad^* the coadjoint representation of G on $\mathfrak{g}^*$.

Example. Suppose

$$\beta : \mathfrak{g} \to \mathfrak{g}^*, \quad \langle X, \beta Y \rangle = B(X, Y), \quad \text{positive-definite,} \tag{2.15}$$

and set

$$f : \mathfrak{g}^* \to \mathbb{R}, \quad f(\xi) = \frac{1}{2} \langle \beta^{-1} \xi, \xi \rangle. \tag{2.16}$$

A calculation gives

$$df(\xi) = \beta^{-1} \xi, \tag{2.17}$$

so in this case the differential equation (2.10) becomes

$$\xi_t = -\,\mathrm{ad}^*(\beta^{-1} \xi)\, \xi. \tag{2.18}$$

It follows from (2.14) that H_f, defined by (2.6), is tangent to each coadjoint orbit $\mathcal{O}^*_{\xi_0}$. The following implies that $H_f\big|_{\mathcal{O}^*_{\xi_0}}$ is determined by $f\big|_{\mathcal{O}^*_{\xi_0}}$.

Proposition 2.1. *If $f_j \in C^\infty(\mathfrak{g}^*)$ and $f_1 = f_2$ on $\mathcal{O}^*_{\xi_0}$, then $H_{f_1} = H_{f_2}$ on $\mathcal{O}^*_{\xi_0}$.*

Proof. Consider $f = f_1 - f_2$, so $f = 0$ on $\mathcal{O}^*_{\xi_0}$. We are claiming that

$$g \in C^\infty(\mathfrak{g}^*) \Longrightarrow H_f g = 0 \ \text{ on } \ \mathcal{O}^*_{\xi_0}. \tag{2.19}$$

Indeed, $H_f g = -H_g f$, and by the reasoning above H_g is tangent to $\mathcal{O}^*_{\xi_0}$, so this is clear.

In other words, each coadjoint orbit in $\mathfrak{g}^*$ gets a Poisson structure. We will see in §3 that this is actually a symplectic structure.

Remark. In case $\mathfrak{g}$ has a non-degenerate, Ad-invariant quadratic form, we can rewrite (2.13) in terms of a differential equation for a curve in $\mathfrak{g}$. In fact, such a quadratic form yields an isomorphism

$$Q : \mathfrak{g} \longrightarrow \mathfrak{g}^*, \quad Q \circ \mathrm{ad} = \mathrm{ad}^* \circ Q. \tag{2.20}$$

Hence if we set

$$v(t) = Q^{-1} \xi(t), \tag{2.21}$$

the differential equation (2.13) is equivalent to

$$v'(t) = -[df(Qv(t)), v(t)]. \tag{2.22}$$

In case f is given by (2.16), this becomes $v_t = -[\beta^{-1}Qv, v]$. This can be compared to (2.7) of Chapter I.

Another formula equivalent to (2.22) arises by taking

$$F : \mathfrak{g} \to \mathbb{R}, \quad F(v) = f(Qv). \tag{2.23}$$

The $dF(v)w = df(Qv)Qw$, so $dF(v) \in \mathfrak{g}^*$ satisfies

$$dF(v) = Q(df(Qv)). \tag{2.24}$$

Thus (2.22) is equivalent to

$$v' = -[Q^{-1}(dF(v)), v]. \tag{2.25}$$

Note how this generalizes the situation giving rise to (2.7) in Chapter I. It is natural to set

$$\nabla F(v) = Q^{-1}(dF(v)), \tag{2.26}$$

and write (2.25) as

$$v' = H_F^Q(v), \quad H_F^Q(v) = -\operatorname{ad}(\nabla F(v))\, v. \tag{2.27}$$

3. Symplectic structures

A Poisson structure on a manifold M is said to be *symplectic* provided

$$\{H_f(p) : f \in C^\infty(M)\} = T_pM, \quad \forall\, p \in M. \tag{3.1}$$

In view of Proposition 2.1, this is equivalent to the statement that

$$\mathcal{J} : T_p^*M \longrightarrow T_pM \tag{3.2}$$

is an isomorphism, where

$$H_f(p) = \mathcal{J} df(p), \quad p \in M. \tag{3.3}$$

In such a case, since $\mathcal{J}^* = -\mathcal{J}$, we define a 2-form σ on M by

$$\sigma(X, Y) = \langle X, \mathcal{J}^{-1}Y\rangle. \tag{3.4}$$

Then σ is non-degenerate, and we have, for all smooth functions f and vector fields X,

$$\sigma(X, H_f) = \langle X, df\rangle = Xf. \tag{3.5}$$

Conversely, given a non-degenerate 2-form σ, we can use (3.5) to define $f \mapsto H_f$, uniquely, and then set $\{f, g\} = H_f g$. Such a bracket clearly satisfies (1.2)–(1.3). Also

$$H_f g = \sigma(H_f, H_g) = -H_g f, \tag{3.6}$$

so we easily have (1.1). The following specifies when the Jacobi identity holds.

Proposition 3.1. *If σ is a non-degenerate 2-form on M and $f \mapsto H_f$ is defined by (3.5), then the Jacobi identity holds if and only if σ is closed, i.e., $d\sigma = 0$.*

Proof. Since (3.1) holds in this setting, it suffices to show that

$$d\sigma(H_{f_0}, H_{f_1}, H_{f_2}) = 0, \tag{3.7}$$

for all $f_j \in C^\infty(M)$, if and only if the Jacobi identity holds. A standard formula for the 3-form $d\sigma$ (cf. [T2], Chapter I, (13.15)) gives the left side of (3.7) as

$$\sum_{\ell=0}^{2} (-1)^\ell H_{f_\ell} \sigma(H_{f_j}, H_{f_k}) + \sum_{0 \le \ell < j \le 2} (-1)^{j+\ell} \sigma([H_{f_\ell}, H_{f_j}], H_{f_k}). \tag{3.8}$$

In the first sum, $\{\ell, j, k\} = \{0, 1, 2\}$ and $j < k$. In the second sum also $\{\ell, j, k\} = \{0, 1, 2\}$. We can write the first sum as

$$\{f_0, \{f_1, f_2\}\} - \{f_1, \{f_0, f_2\}\} + \{f_2, \{f_0, f_1\}\}, \tag{3.9}$$

and the second sum as

$$\begin{aligned} &- [H_{f_0}, H_{f_1}]f_2 + [H_{f_0}, H_{f_2}]f_1 - [H_{f_1}, H_{f_2}]f_0 = \\ &- 2\{f_0, \{f_1, f_2\}\} - 2\{f_1, \{f_2, f_0\}\} - 2\{f_2, \{f_0, f_1\}\}. \end{aligned} \tag{3.10}$$

Thus the left side of (3.7) is equal to

$$-\{f_0, \{f_1, f_2\}\} - \{f_1, \{f_2, f_0\}\} - \{f_2, \{f_0, f_1\}\}, \tag{3.11}$$

which shows that $d\sigma = 0$ if and only if the Jacobi identity holds.

REMARK. A closed, non-degenerate 2-form is called a symplectic form.

The standard example of a symplectic structure is the Poisson structure on $\mathbb{R}^{2n}$ defined by (1.7). In this case, we have

$$\sigma = \sum_{j=1}^{n} d\xi_j \wedge dx_j. \tag{3.12}$$

We also have $\sigma = d\kappa$, where κ is the following 1-form on $\mathbb{R}^{2n}$:

$$\kappa = \sum_{j=1}^{n} \xi_j \, dx_j. \tag{3.13}$$

This is called the contact form. We will present a more general construction of contact forms below.

Another family of examples of symplectic structures arises on coadjoint orbits. Recall the Poisson structure on the dual $\mathfrak{g}^*$ of a Lie algebra, defined by (2.6). As shown in Proposition 2.1, this induces a Poisson structure on each coadjoint orbit $\mathcal{O}^*_{\xi_0} \subset \mathfrak{g}^*$. We claim that each of these is a symplectic structure, i.e., that (3.1) holds, for $M = \mathcal{O}^*_{\xi_0}$. Note that, given $\xi \in \mathcal{O}^*_{\xi_0}$,

$$T_\xi \mathcal{O}^*_{\xi_0} = \{\mathrm{ad}^*\, X\, \xi : X \in \mathfrak{g}\}. \tag{3.14}$$

On the other hand, as seen in (2.11),

$$H_f(\xi) = -\,\mathrm{ad}^*(df(\xi))\,\xi, \tag{3.15}$$

so it remains to show that for each $X \in \mathfrak{g}$, $\xi \in \mathfrak{g}^*$, there exists $f \in C^\infty(\mathfrak{g}^*)$ such that $df(\xi) = X$. Indeed, $f(\xi) = \langle X, \xi\rangle$ works, so (3.1) is verified.

We now discuss another very important family of symplectic manifolds, generalizing (3.12). Namely, we consider the natural symplectic structure that arises on the cotangent bundle T^*M of a smooth manifold M. In fact, the symplectic form is given by $\sigma = d\kappa$, where κ is the contact form on T^*M, constructed as follows. Let

$$\pi : T^*M \longrightarrow M \tag{3.16}$$

be the natural projection. Given $p \in M$ and $z \in T^*_p M$, we want to define $\kappa(z) \in T^*_z(T^*M)$ by its action on $v \in T_z(T^*M)$. The formula is

$$\langle v, \kappa(z)\rangle = \langle (D\pi)v, z\rangle, \tag{3.17}$$

where

$$D\pi : T_z(T^*M) \longrightarrow T_p M \tag{3.18}$$

is the derivative of π in (3.16).

We examine κ in local coordinates $z = (x, \xi) = (x_1, \dots, x_n, \xi_1, \dots \xi_n)$ on T^*M arising from local coordinates $x = (x_1, \dots, x_n)$ on M. Given $v \in T_z(T^*M)$, we write $v = (v_1, \dots, v_n, w_1, \dots, w_n)$, so

$$(D\pi)v = (v_1, \dots, v_n), \quad \langle (D\pi)v, z\rangle = \sum_{j=1}^n v_j \xi_j.$$

Hence (3.17) gives

$$\kappa = \sum_{j=1}^n \xi_j\, dx_j,$$

as in (3.13), so σ is given by (3.12). In particular σ is non-degenerate, as well as closed, so it is a symplectic form.

4. Right (and left) invariant Lagrangians on Lie groups

We study the equation for a critical point of a Lagrangian integral

$$I(u) = \int_a^b L(u(t), u'(t))\, dt, \tag{4.1}$$

for a path $u : [a, b] \to G$, where G is a Lie group, and we assume $L(p, v)$ has the form

$$L(p, v) = F(vp^{-1}), \tag{4.2}$$

for some smooth $F : \mathfrak{g} \to \mathbb{R}$. The calculations here generalize those done for geodesics on Lie groups with right (or left) invariant metric tensors in Chapter I, §1.

The Lagrange equation for a critical path for (4.1) is

$$\frac{d}{dt} D_v L(u, u_t) = D_p L(u, u_t). \tag{4.3}$$

When $L(p, v)$ is given by (4.2), we have

$$\begin{aligned} D_p L(p, v)W &= -DF(vp^{-1})vp^{-1}Wp^{-1}, \\ D_v L(p, v)W &= DF(vp^{-1})Wp^{-1}, \end{aligned} \tag{4.4}$$

with $W \in T_pG$. We see that

$$\begin{aligned} \frac{d}{dt} D_v L(u, u_t)W = D^2F(u_tu^{-1})&(u_{tt}u^{-1} - u_tu^{-1}u_tu^{-1}, Wu^{-1}) \\ &- DF(u_tu^{-1})Wu^{-1}u_tu^{-1}. \end{aligned} \tag{4.5}$$

Thus the Lagrange equation (4.3) becomes

$$\begin{aligned} D^2F(u_tu^{-1})(u_{tt}u^{-1} - u_tu^{-1}u_tu^{-1}, Wu^{-1}) - DF(u_tu^{-1})Wu^{-1}u_tu^{-1} \\ = -DF(u_tu^{-1})u_tu^{-1}Wu^{-1}, \end{aligned} \tag{4.6}$$

for all $W \in T_uG$.

We now set

$$v(t) = u_t(t)u(t)^{-1}, \quad v : [a, b] \to \mathfrak{g}. \tag{4.7}$$

Hence $u_t = vu,\ u_{tt} = v_t u + v u_t$, and, with $Y = Wu^{-1} \in \mathfrak{g}$, the equation (4.6) yields

$$D^2F(v_t, Y) - DF(v)Yv = -DF(v)vY, \quad \forall\, Y \in \mathfrak{g}, \tag{4.8}$$

or equivalently we obtain the following equation for $v(t)$:

$$D^2F(v)(v_t, Y) + DF(v)[v, Y] = 0, \quad \forall\, Y \in \mathfrak{g}. \tag{4.9}$$

Note that if one can solve (1.9) for $v(t)$, then the solution $u(t)$ to (4.3) is obtained as a solution to the (generally non-autonomous) linear equation

$$u_t(t) = v(t)u(t). \tag{4.10}$$

We can derive from (4.9) an equation for a curve $\xi(t)$ in $\mathfrak{g}^*$, upon setting

$$\xi = DF(v). \tag{4.11}$$

This is a "Legendre transform." We have $\langle \xi_t, Y\rangle = D^2F(v_t, Y)$, so (4.9) is equivalent to

$$\langle \xi_t, Y\rangle = -DF(v)[v, Y] = -\langle \xi, [v, Y]\rangle, \quad \forall\, Y \in \mathfrak{g}, \tag{4.12}$$

hence to

$$\xi_t = \mathrm{ad}^*(v)\, \xi. \tag{4.13}$$

If we use a left-invariant Lagrangian, i.e., replace (4.2) by

$$L(p, v) = F(p^{-1}v), \tag{4.14}$$

we get similar formulas, involving $v(t) = u(t)^{-1}u_t(t)$ in place of (4.7). We wind up with a sign change in the second term of (4.9). Thus setting $\xi = F(v)$ as in (4.11) yields

$$\xi_t = -\,\mathrm{ad}^*(v)\, \xi, \tag{4.15}$$

in place of (4.13).

We will concentrate on (4.15), which is reminiscent of the equation (2.13):

$$\xi_t = -\,\mathrm{ad}^*(df(\xi))\, \xi, \tag{4.16}$$

itself equivalent to (2.10):

$$\xi_t = H_f(\xi), \tag{4.17}$$

where $f \mapsto H_f$ is determined by the Lie-Poisson structure (2.6) on $\mathfrak{g}^*$. We claim that, if the Legendre transform (4.11) is a difeomorphism from $\mathfrak{g}$ to $\mathfrak{g}^*$, then (4.15) is equivalent to (4.16), with $f(\xi)$ defined by

$$f(\xi) = DF(v)v - F(v), \quad \xi = DF(v). \tag{4.18}$$

Equivalently,

$$f(\xi) = \langle v, \xi \rangle - F(v). \tag{4.19}$$

In fact, defining $f(\xi)$ in this way, with $\xi = \xi(v)$ given by (4.11), we have

$$\begin{aligned} Df(\xi)D\xi(v)W &= \langle v, D\xi(v)W \rangle + \langle W, \xi \rangle - DF(v)W \\ &= \langle v, D\xi(v)W \rangle, \end{aligned} \tag{4.20}$$

for all $W \in \mathfrak{g}$. Hence, given that $D\xi(v)$ is invertible, we have

$$Df(\xi) = v, \tag{4.21}$$

showing that (4.15) and (4.16) coincide.

Note that conversely, if we are given $f : \mathfrak{g}^* \to \mathbb{R}$ and if (4.21) defines a diffeomorphism $Df : \mathfrak{g}^* \to \mathfrak{g}$, then we can define the associated Lagrangian by

$$F(v) = \langle v, \xi \rangle - f(\xi) = Df(\xi)\xi - f(\xi). \tag{4.22}$$

5. Shifted and frozen Poisson structures on $\mathfrak{g}^*$

Here we consider variants of the Lie-Poisson structure on the dual $\mathfrak{g}^*$ of a Lie algebra $\mathfrak{g}$, which was treated in §2. Recall that, for $f, g \in C^\infty(\mathfrak{g}^*)$,

$$\{f, g\}(\xi) = \langle [df(\xi), dg(\xi)], \xi \rangle. \tag{5.1}$$

One way to modify this Poisson structure is simply to translate it. That is to say, fix $\xi_0 \in \mathfrak{g}^*$, and define $\tau_{\xi_0} : C^\infty(\mathfrak{g}^*) \to C^\infty(\mathfrak{g}^*)$ by

$$\tau_{\xi_0} f(\xi) = f(\xi + \xi_0). \tag{5.2}$$

Then set

$$\{f, g\}_{\xi_0} = \tau_{\xi_0}^{-1}\{\tau_{\xi_0} f, \tau_{\xi_0} g\}. \qquad \text{(Shifted Poisson structure)} \tag{5.3}$$

It is obvious that (1.1)–(1.4) hold for $\{\cdot, \cdot\}_{\xi_0}$, given that it holds for $\{\cdot, \cdot\}$.

To relate these Poisson structures in another way, we calculate

$$\begin{aligned}\{\tau_{\xi_0} f, \tau_{\xi_0} g\}(\xi - \xi_0) &= \langle [d\tau_{\xi_0} f(\xi - \xi_0), d\tau_{\xi_0} g(\xi - \xi_0)], \xi - \xi_0\rangle \\ &= \langle [df(\xi), dg(\xi)], \xi - \xi_0\rangle,\end{aligned} \tag{5.4}$$

i.e.,

$$\{f, g\}_{\xi_0} = \{f, g\} - \{f, g\}^{\xi_0}, \tag{5.5}$$

with

$$\{f, g\}^{\xi_0}(\xi) = \langle [df(\xi), dg(\xi)], \xi_0\rangle. \qquad \text{(Frozen Poisson structure)} \tag{5.6}$$

As the label suggests, $\{\cdot, \cdot\}^{\xi_0}$ also satisfies (1.1)–(1.4). Of course, (1.1)–(1.2) are obvious. Also (1.3) readily follows from (2.7). Hence there is a correspondence $f \mapsto H_f^{\xi_0}$ such that $\{f, g\}^{\xi_0} = H_f^{\xi_0} g$. We need to verify (1.4), or its equivalent

$$H^{\xi_0}_{\{f,g\}^{\xi_0}} = [H_f^{\xi_0}, H_g^{\xi_0}]. \tag{5.7}$$

Retracing an argument in §2, we first note that

$$f, g \ \text{ linear} \Longrightarrow \{f, g\}^{\xi_0}(\xi) = \langle [f, g], \xi_0\rangle. \tag{5.8}$$

In this case the bracket yields a function that is independent of ξ. Thus when f, g, and h are linear, all three terms in (1.4) (with $\{\cdot, \cdot\}$ replaced by $\{\cdot, \cdot\}^{\xi_0}$) are zero. Hence both sides of (5.7) are vector fields that have the same action on linear functions, so these vector fields are equal, whenever f and g are linear. From here, as in §2, we can apply Proposition 1.1 to verify (5.7) for general f and g, so (5.6) does define a Poisson structure on $\mathfrak{g}^*$.

For an alternative approach, note that if we replace ξ_0 by ξ_0/ε in (5.5), we obtain

$$\varepsilon\{f, g\} - \{f, g\}^{\xi_0} = \varepsilon\{f, g\}_{\xi_0/\varepsilon}, \quad \forall\, \varepsilon > 0.$$

Since the right side clearly gives a Poisson structure, so does the left side, for all $\varepsilon > 0$. Taking $\varepsilon \to 0$ yields the Jacobi identity for $\{\cdot, \cdot\}^{\xi_0}$.

We use the notions introduced above to produce rich families of functions on $\mathfrak{g}^*$ that Poisson commute. Denote by $I(\mathfrak{g}^*)$ the space of Ad^*-invariant functions on $\mathfrak{g}^*$:

$$I(\mathfrak{g}^*) = \{f \in C^\infty(\mathfrak{g}^*) : f(\mathrm{Ad}^*(g)\xi) = f(\xi), \forall\, g \in G\}. \tag{5.9}$$

Clearly

$$g \in I(\mathfrak{g}^*) \Longrightarrow \{f, g\} = 0, \quad \forall\ f \in C^\infty(\mathfrak{g}^*), \tag{5.10}$$

since we have seen that H_f is tangent to coadjoint orbits. We now establish the following.

Proposition 5.1. *Fix $\xi_0 \in \mathfrak{g}^*$. Then*

$$f, g \in I(\mathfrak{g}^*), \quad s, t \in \mathbb{R} \Longrightarrow \{\tau_{s\xi_0} f, \tau_{t\xi_0} g\} = 0. \tag{5.11}$$

Proof. For notational simplicity, set $f_{s\xi_0} = \tau_{s\xi_0} f$, etc. First note that, under the hypotheses of (5.11),

$$\begin{aligned} \{f_{s\xi_0}, g_{t\xi_0}\}_{-t\xi_0} &= \tau_{t\xi_0}\{\tau_{(s-t)\xi_0} f, g\} = 0, \\ \{f_{s\xi_0}, g_{t\xi_0}\}_{-s\xi_0} &= \tau_{s\xi_0}\{f, \tau_{(t-s)\xi_0} g\} = 0, \end{aligned} \tag{5.12}$$

the first identity on each line by (5.3) and the second by (5.10). Hence we obtain

$$0 = s\{f_{s\xi_0}, g_{t\xi_0}\}_{-t\xi_0} - t\{f_{s\xi_0}, g_{t\xi_0}\}_{-s\xi_0} = (s-t)\{f_{s\xi_0}, g_{t\xi_0}\}, \tag{5.13}$$

upon applying (5.5) to the quantities on the left sides of (5.12), and observing a cancellation of $\pm\{f_{s\xi_0}, g_{t\xi_0}\}^{st\xi_0}$. This yields the conclusion in (5.11) when $s \neq t$. The case $s = t$ follows by continuity.

6. Poisson pairs and bi-Hamiltonian vector fields

Let $\{\cdot,\cdot\}_0$ and $\{\cdot,\cdot\}_1$ be two Poisson structures on a manifold M. They are said to form a *Poisson pair* provided

$$s\{\cdot,\cdot\}_0 + t\{\cdot,\cdot\}_1 = \{\cdot,\cdot\}_{s,t} \tag{6.1}$$

is a Poisson structure on M for each $s, t \in \mathbb{R}$. (Clearly this need only be checked for $s = 1,\ t \in \mathbb{R}$.) It is clear that $\{\cdot,\cdot\}_{s,t}$ always satisfies (1.1)–(1.3). In particular there is a correspondence $f \mapsto H_f^{s,t}$, such that $H_f^{s,t} g = \{f, g\}_{s,t}$. With obvious notation,

$$H_f^{s,t} = sH_f^0 + tH_f^1. \tag{6.2}$$

The condition that we have a Poisson pair is that (1.4) holds, i.e.,

$$[H_f^{s,t}, H_g^{s,t}] = H_{\{f,g\}_{s,t}}^{s,t}. \tag{6.3}$$

Expanding both sides and making obvious cancellations, we see that the Jacobi condition is equivalent to

$$[H_f^1, H_g^0] + [H_f^0, H_g^1] = H_{\{f,g\}_0}^1 + H_{\{f,g\}_1}^0. \tag{6.4}$$

Equivalently, the condition to have a Poisson pair is

$$\begin{aligned}(6.5)\qquad &\{f,\{g,h\}_0\}_1+\{g,\{h,f\}_0\}_1+\{h,\{f,g\}_0\}_1\\ &+\{f,\{g,h\}_1\}_0+\{g,\{h,f\}_1\}_0+\{h,\{f,g\}_1\}_0=0,\end{aligned}$$

for all $f,g,h\in C^\infty(M)$.

Significant examples of Poisson structures arise for $M=\mathfrak{g}^*$, the dual to a Lie algebra $\mathfrak{g}$. In fact, fix $\xi_0\in\mathfrak{g}^*$, let $\{\cdot,\cdot\}_0=\{\cdot,\cdot\}$ be the Lie-Poisson bracket, and let $\{\cdot,\cdot\}_1=\{\cdot,\cdot\}^{\xi_0}$ be the frozen Poisson bracket, given by (5.6). In view of the computation (5.4)–(5.5), we have

$$(6.6)\qquad \{f,g\}+t\{f,g\}^{\xi_0}=\{f,g\}_{-t\xi_0},$$

which as seen in §5 is a Poisson bracket, for each $t\in\mathbb{R}$ (i.e., each $-t\xi_0\in\mathfrak{g}^*$).

Suppose we have a Poisson pair $\{\cdot,\cdot\}_0,\ \{\cdot,\cdot\}_1$ on M. A vector field X on M is said to be bi-Hamiltonian (with respect to this pair of Poisson structures) if there exist $f_j\in C^\infty(M)$ such that

$$(6.7)\qquad X=H^0_{f_0}=H^1_{f_1}.$$

The following result is a useful tool in the study of integrable systems.

Proposition 6.1. *Assume X is a bi-Hamiltonian vector field with respect to Poisson structures $\{\cdot,\cdot\}_j,\ j=0,1$. Then $H^0_{f_1}$ is also an infinitesimal Poisson map for $\{\cdot,\cdot\}_1$.*

Proof. The claim is that $Y=H^0_{f_1}$ satisfies

$$(6.8)\qquad Y\{g,h\}_1=\{Yg,h\}_1+\{g,Yh\}_1;$$

cf. Proposition 1.3 and the subsequent discussion. Equivalently, we need to show that

$$(6.9)\qquad \{f_1,\{g,h\}_1\}_0-\{\{f_1,g\}_0,h\}_1-\{g,\{f_1,h\}_0\}_1=0.$$

Now (6.5) implies that the left side of (6.9) together with the following sums to zero:

$$(6.10)\qquad \{f_1,\{g,h\}_0\}_1-\{\{f_1,g\}_1,h\}_0-\{g,\{f_1,h\}_1\}_0.$$

But if (6.7) holds, then (6.10) is equal to

$$(6.11)\qquad \{f_0,\{g,h\}_0\}_0-\{\{f_0,g\}_0,h\}_0-\{g,\{f_0,h\}_0\}_0,$$

whixch is 0 by (1.4). This proves the proposition.

The way this result is used is as follows. Given the setting of Proposition 6.1, then *under "favorable" circumstances,*

$$\text{There exists } f_2 \in C^\infty(M) \text{ such that } H^0_{f_1} = H^1_{f_2}. \tag{6.12}$$

That is to say, the infinitesimal Poisson map $Y = H^0_{f_1}$ for $\{\cdot,\cdot\}_1$ discussed in Proposition 6.1 is actually a Hamiltonian vector field for this Poisson structure. Thus we have an analogue of (6.7), to which Proposition 6.1 applies. If this "favorable circumstance" persists, we can write

$$H^0_{f_2} = H^1_{f_3}, \tag{6.13}$$

and continue, obtaining f_j such that

$$H^0_{f_j} = H^1_{f_{j+1}}, \quad j \ge 0. \tag{6.14}$$

When this works, the sequence is said to follow the *Lenard scheme* (cf. [AK], p. 310). An important corollary is that, when this program works, these functions are in involution. This is a consequence of the following result.

Proposition 6.2. *Assume $\{\cdot,\cdot\}_0$ and $\{\cdot,\cdot\}_1$ are a Poisson pair on M, and X is a bi-Hamiltonian vector field, of the form (6.7). Then*

$$\{f_0, f_1\}_0 = \{f_0, f_1\}_1 = 0. \tag{6.15}$$

Furthermore, if the sequence $\{f_j : j \ge 0\}$ exists, satisfying (6.14), then

$$\{f_j, f_k\}_0 = \{f_j, f_k\}_1 = 0, \quad \forall\, j,k \ge 0. \tag{6.16}$$

Proof. First we prove (6.15). In fact (6.7) gives $\{f_0, f_1\}_0 = H^0_{f_0} f_1 = H^1_{f_1} f_1 = 0$, and similarly $\{f_0, f_1\}_1 = -H^1_{f_1} f_0 = -H^0_{f_0} f_0 = 0$.

Now suppose $\{f_j : j \ge 0\}$ exists, satisfying (6.14). The same reasoning used to establish (6.15) immediately gives

$$\{f_j, f_{j+1}\}_0 = \{f_j, f_{j+1}\}_1 = 0, \quad \forall\, j \ge 0. \tag{6.17}$$

Hence (6.16) is true whenever $|j-k| \le 1$. To treat the general case, assume $j < k$ and note the following:

$$\begin{aligned} \{f_j, f_{k+1}\}_0 &= H^0_{f_j} f_{k+1} = \{f_{j+1}, f_{k+1}\}_1, \\ \{f_j, f_{k+1}\}_1 &= -H^1_{f_{k+1}} f_j = \{f_j, f_k\}_0. \end{aligned} \tag{6.18}$$

The transformations here decrease $k+1-j$ to $k-j$, so (6.16) follows by induction.

The following gives an important "favorable circumstance" for invoking (6.14):

Proposition 6.3. *In the setting of Proposition 6.1, assume also that* $\{\cdot,\cdot\}_1$ *is a symplectic structure on* M*, and that* M *is simply connected. Then (6.12) holds, and so does (6.14).*

Proof. The result of Proposition 6.1 that $Y = H^0_{f_1}$ is an infinitesimal Poisson map for $\{\cdot,\cdot\}_1$ implies that its flow preserves the associated symplectic form (call it σ_1), or equivalently that $\mathcal{L}_Y\sigma_1 = 0$. Now use Cartan's formula to get

$$0 = \mathcal{L}_Y\sigma_1 = d(\sigma_1\rfloor Y) + (d\sigma_1)\rfloor Y \Longrightarrow d(\sigma_1\rfloor Y) = 0. \tag{6.19}$$

Since M is simply connected, this implies

$$\sigma_1\rfloor Y = -df_2, \tag{6.20}$$

for some $f_2 \in C^\infty(M)$, which in turn gives (6.12).

Chapter 3: Motion on a Lie Group With a Potential

Introduction

Here we have a brief variation on the theme of Chapter I. Namely, we consider motion of a particle on a Lie group, endowed with a right-invariant metric tensor, and also equipped with a potential, giving rise to a force field. We examine one family of physical problems, associated with the "heavy top," or the "spinning top." We first derive equations of motion for such a top (spinning on a table) in $\mathbb{R}^n$, formulated as motion on $SO(n)$, equipped with such a metric tensor and potential. We then note some "miracles" that occur when $n = 3$, arising from the unique isomorphism of the Lie algebra so(3) with $\mathbb{R}^3$, equipped with the cross product. Even specialized to three dimensions, most spinning top equations are not integrable. We discuss one family of integrable examples, discovered by Lagrange back in the dawn of the theory of analytical dynamics.

1. Lie groups with potentials

Let G be a Lie group, endowed with a right-invariant metric tensor. We want to study the motion of a particle on G, under the influence of a force arising from a potential V. The path $u : [a, b] \to G$ followed by such a particle is a stationary point of the functional

$$I(u) = \int_a^b L(u(t), u'(t))\, dt, \tag{1.1}$$

where

$$L(p, v) = \frac{1}{2} B(vp^{-1}, vp^{-1}) - V(p), \quad p \in G,\ v \in T_pG, \tag{1.2}$$

$B(\cdot, \cdot)$ being an inner product on $T_eG = \mathfrak{g}$.

The equation of motion is produced via calculations parallel to those done in Chapter I, §1. The standard Lagrange equation for such a stationary point is

$$\frac{d}{dt} D_v L(u, u_t) = D_p L(u, u_t). \tag{1.3}$$

When $L(p, v)$ is given by (1.2), we have

$$\begin{aligned} D_p L(p, v)W &= -B(vp^{-1}Wp^{-1}, vp^{-1}) - DV(p), \\ D_v L(p, v)W &= B(Wp^{-1}, vp^{-1}), \end{aligned} \tag{1.4}$$

with $W \in T_pG$. We see that

(1.5)
$$\frac{d}{dt} D_v L(u, u_t) W = -B(Wu^{-1}u_t u^{-1}, u_t u^{-1}) + B(Wu^{-1}, u_{tt}u^{-1} - u_t u^{-1} u_t u^{-1}).$$

Thus the Lagrange equation (1.3) becomes

(1.6)
$$\begin{aligned} -B(Wu^{-1}u_t u^{-1}, u_t u^{-1}) + B(Wu^{-1}, u_{tt}u^{-1} - u_t u^{-1} u_t u^{-1}) \\ = -B(u_t u^{-1} W u^{-1}, u_t u^{-1}) - DV(u)W, \quad \forall\, W \in T_uG. \end{aligned}$$

We now set

$$v(t) = u_t(t)u(t)^{-1}, \quad v : (a, b) \to \mathfrak{g}. \tag{1.7}$$

Hence $u_t = vu$, $u_{tt} = v_t u + v u_t$, and, with $Y = Wu^{-1} \in \mathfrak{g}$, the equation (1.6) yields

$$-B(Yv, v) + B(Y, v_t) = -B(vY, v) - DV(u)Yu, \quad \forall\, Y \in \mathfrak{g}, \tag{1.8}$$

or equivalently we obtain the following equation for $v(t)$:

$$B(v_t, Y) + B(v, [v, Y]) = -DV(u)Yu, \quad \forall\, Y \in \mathfrak{g}. \tag{1.9}$$

Unlike the geodesic equation (1.9) of Chapter I, this is not an equation for v alone, but it is completed by coupling it to

$$u_t = vu. \tag{1.10}$$

For solutions to the system (1.9)–(1.10), the total energy $(1/2)B(v, v) + V(u)$ is conserved. This is verified by the following calculation:

$$\begin{aligned} \frac{d}{dt}\Big(\frac{1}{2}B(v, v) + V(u)\Big) &= B(v_t, v) + DV(u)u_t \\ &= -B(v, [v, v]) - DV(u)vu + DV(u)u_t \\ &= 0, \end{aligned} \tag{1.11}$$

the second identity by using $Y = v$ in (1.9) and the third via (1.10).

We can transform (1.9) to an ODE for a curve $\xi(t)$ in $\mathfrak{g}^*$ by the usual device. Take

$$\beta : \mathfrak{g} \longrightarrow \mathfrak{g}^*, \quad B(u, v) = \langle u, \beta v \rangle, \tag{1.12}$$

and set $\xi(t) = \beta v(t)$. Then (1.9) yields

$$\langle \xi_t, Y \rangle + \langle \xi, \mathrm{ad}(v)Y \rangle = -DV(u)Yu, \quad \forall\, Y \in \mathfrak{g}. \tag{1.13}$$

We can write

$$DV(u)Yu = \langle DV(u)u, \mathrm{Ad}(u^{-1})Y \rangle, \tag{1.14}$$

with $DV(u)u : T_eG \to \mathbb{R}$, linear, i.e., $DV(u)u \in \mathfrak{g}^*$. Then the system (1.9)–(1.10) becomes

$$\begin{aligned} \xi_t - \mathrm{ad}^*(v)\xi &= -\,\mathrm{Ad}^*(u)(DV(u)u), \quad v = \beta^{-1}\xi, \\ u_t &= vu. \end{aligned} \tag{1.15}$$

2. The heavy top

Suppose there is a rigid body in $\mathbb{R}^n$, with a mass distribution at $t = 0$ given by a function $\rho(x)$, which we will assume is piecewise continuous and has compact support. Suppose the body moves, subject to the force of gravity, and with the constraint that one point remains fixed, say at the origin. This situation models the motion of a heavy top, spinning about while sitting on a table. We want to describe the motion of such a body. The derivation of these equations can be compared with the derivation of the equations of motion of a free rigid body, done in §3 of Chapter I.

According to the Lagrangian approach to mechanics, we seek an extremum of the following Lagrangian, subject to this constraint. If $\xi(t,x)$ is the position in $\mathbb{R}^n$ at time t of the point on the body whose position at time 0 is x, then we can write the Lagrangian as

$$I(\xi) = \frac{1}{2}\int_a^b \int_{\mathbb{R}^n} \rho\big(\xi(t,x)\big)|\dot{\xi}(t,x)|^2\,dx\,dt - g\int_a^b \int_{\mathbb{R}^n} \rho\big(\xi(t,x)\big)\,\gamma_0 \cdot x\,dx\,dt. \tag{2.1}$$

Here, $\dot{\xi}(t,x) = \partial\xi/\partial t$, γ_0 is the unit vector pointing in the vertical direction, opposite to the direction of the force of gravity, and g is the gravitational acceleration.

Our rigidity assumption plus the assumption that the point at the origin remains fixed allows us to write

$$\xi(t,x) = W(t)x, \quad W(t) \in SO(n), \tag{2.2}$$

where $SO(n)$ is the group of rotations of $\mathbb{R}^n$. Thus, describing the motion of the body becomes the problem of specifying the curve $W(t)$ in $SO(n)$. We can write (2.1) as

$$\begin{aligned} I(\xi) &= \frac{1}{2}\int_a^b \int_{\mathbb{R}^n} \rho\big(W(t)x\big)|W'(t)x|^2\,dx\,dt - g\int_a^b \int_{\mathbb{R}^n} \rho\big(W(t)x\big)\,\gamma_0 \cdot x\,dx\,dt \\ &= \frac{1}{2}\int_a^b \int_{\mathbb{R}^n} \rho(y)|W'(t)W(t)^{-1}y|^2\,dy\,dt - g\int_a^b \int_{\mathbb{R}^n} \rho(y)\,\gamma_0 \cdot W(t)^{-1}y\,dy\,dt \\ &= J(W). \end{aligned} \tag{2.3}$$

We look for an extremum, or other critical point, where we vary the family of paths $W : [a,b] \to SO(n)$ (keeping the endpoints fixed).

Let us reduce the formula (2.3) for $J(W)$ to a single integral, over t. As shown in Chapter I, §3, we have

$$\int \rho(y)\ (Ay, By)\ dy = \mathrm{Tr}\,(B^t A \mathcal{I}_\rho) = \mathrm{Tr}\,(A \mathcal{I}_\rho B^t), \tag{2.4}$$

where

$$\mathcal{I}_\rho = \int \rho(y)\ y \otimes y\ dy \in \bigotimes{}^2\ \mathbb{R}^n \approx M(n, \mathbb{R}). \tag{2.5}$$

Let us also set

$$\sigma = \int \rho(y) y\, dy \in \mathbb{R}^n. \tag{2.6}$$

Then we can write the Lagrangian (2.3) as

$$J(W) = \frac{1}{2} \int_a^b \mathrm{Tr}\,\big(W'(t)W(t)^{-1}\mathcal{I}_\rho (W'(t)W(t)^{-1})^t\big)\, dt - g \int_a^b \sigma \cdot W(t)\gamma_0\, dt. \tag{2.7}$$

This has the general form

$$I(u) = \int_a^b L(u(t), u'(t))\, dt, \tag{2.8}$$

considered in §1, i.e.,

$$L(p, v) = \frac{1}{2} B(vp^{-1}, vp^{-1}) - V(p), \quad V(p) = g\sigma \cdot p\,\gamma_0, \quad p \in G,\ v \in T_pG, \tag{2.9}$$

$B(\cdot,\cdot)$ being an inner product on $T_eG = \mathfrak{g}$, namely

$$B(v, w) = \mathrm{Tr}(v \mathcal{I}_\rho w^t). \tag{2.10}$$

Thus a critical path $W(t)$ for (2.7) satisfies the following Euler-Lagrange equation (from (1.9)), involving $W(t)$ and $Z(t) = W'(t)W(t)^{-1}$:

$$B(Z_t, Y) + B(Z, [Z, Y]) = -g\sigma \cdot YW\gamma_0, \quad \forall\ Y \in \mathrm{so}(n), \tag{2.11}$$

since in this case $DV(W)X = g\sigma \cdot X\gamma_0$ and we use $X = YW$. This equation is coupled to

$$W_t = ZW. \tag{2.12}$$

Here $W : [a,b] \to SO(n)$ and $Z : [a,b] \to \mathrm{so}(n)$, while $\sigma, \gamma \in \mathbb{R}^n$.

Note that the conserved energy for this system is

$$\begin{aligned} \mathcal{E} &= \frac{1}{2} B(Z,Z) + V(W) \\ &= \frac{1}{2} \operatorname{Tr}(Z \mathcal{I}_\rho Z^t) + g\sigma \cdot W\gamma_0. \end{aligned} \tag{2.13}$$

Note that if we set

$$M(t) = \mathcal{I}_\rho Z(t) + Z(t)\mathcal{I}_\rho, \tag{2.14}$$

then $2B(Z,Y) = Q(M,Y) = \operatorname{Tr}(MY^t)$, and since $Q(\cdot,\cdot)$ is Ad-invariant, (2.11) becomes

$$Q(M_t, Y) + Q([M,Z], Y) = -2g\sigma \cdot YW\gamma_0, \quad \forall\, Y \in \mathrm{so}(n). \tag{2.15}$$

3. The heavy top in 3D

We specialize the study of the equations for a heavy top spinning on a table in $\mathbb{R}^n$, studied in the last section, to the case $n = 3$. We use the isomorphism $\kappa : \mathbb{R}^3 \to \mathrm{so}(3)$, given by (3.15) of Chapter I, and we set

$$\omega(t) = -\kappa^{-1} Z(t), \quad \mu(t) = -\kappa^{-1} M(t) = \mathcal{J}_\rho \omega(t), \quad \gamma(t) = W(t)\gamma_0. \tag{3.1}$$

Here $\mathcal{J}_\rho = (\operatorname{Tr} \mathcal{I}_\rho) I - \mathcal{I}_\rho$, as in (3.18) of Chapter I. Recall that

$$\kappa(\omega)x = \omega \times x, \quad \kappa(x \times y) = [\kappa(x), \kappa(y)], \quad \operatorname{Tr}(\kappa(x)\kappa(y)^t) = 2x \cdot y. \tag{3.2}$$

Then (2.15) takes the following form (with $y = \kappa^{-1}(Y)$):

$$-\mu_t \cdot y + (\mu \times \omega) \cdot y = -g\sigma \cdot (y \times \gamma), \quad \forall\, y \in \mathbb{R}^3. \tag{3.3}$$

This leads to

$$\mu_t = \mu \times \omega + g\gamma \times \sigma. \tag{3.4}$$

This equation is coupled to

$$\gamma_t = \gamma \times \omega, \tag{3.5}$$

which follows from (2.12).

Together (3.4)–(3.5) form the heavy top equations in 3D. We note that the conserved energy (2.13) takes the form

$$\mathcal{E} = \frac{1}{2}\mu \cdot \omega + g\sigma \cdot \gamma, \quad \mu = \mathcal{J}_\rho \omega. \tag{3.6}$$

Another conserved quantity is $\mu \cdot \gamma$. Indeed, we have

$$\frac{d}{dt}(\mu \cdot \gamma) = (\mu \times \omega) \cdot \gamma + \mu \cdot (\gamma \times \omega) = 0, \tag{3.7}$$

making use of (3.4) and (3.5). It also follows from (3.5) that

$$\frac{d}{dt}(\gamma \cdot \gamma) = 0, \tag{3.8}$$

but this is also a trivial consequence of the definition $\gamma(t) = W(t)\gamma_0$ plus the fact that $W(t) \in SO(3)$.

It is instructive to rewrite (3.4)–(3.5) in a form (which is a variant of (2.15)) involving M and Z, given by (3.1), and also

$$\Gamma(t) = -\kappa\,\gamma(t), \quad \Sigma = -\kappa\,\sigma, \tag{3.9}$$

with κ as in (3.1)–(3.2). We obtain the system

$$\begin{aligned} M_t &= [Z, M] + g[\Sigma, \Gamma], \\ \Gamma_t &= [Z, \Gamma]. \end{aligned} \tag{3.10}$$

The conserved quantities (3.6)–(3.8) take the form

$$\mathcal{E} = \frac{1}{4}\operatorname{Tr}(MZ^t) + \frac{g}{2}\operatorname{Tr}(\Sigma\Gamma^t), \quad \operatorname{Tr}(M\Gamma^t), \quad \operatorname{Tr}(\Gamma\Gamma^t). \tag{3.11}$$

It is a remarkable fact that the system (3.10) can be put in the commutator form as a differential equation for a curve in the Lie algebra of the Euclidean group

$$E(3) \approx G \times_{\mathrm{Ad}} \mathfrak{g}, \quad G = SO(3), \tag{3.12}$$

whose Lie algebra $\mathrm{e}(3) = \{(X, u) : X, u \in \mathrm{so}(3)\}$ has Lie bracket

$$[(X, u), (Y, v)] = ([X, Y], \operatorname{ad} X(v) - \operatorname{ad} Y(u)); \tag{3.13}$$

cf. Chapter I, §10. By this recipe we have

$$[(Z, g\Sigma), (\Gamma, M)] = ([Z, \Gamma], [Z, M] + g[\Sigma, \Gamma]), \tag{3.14}$$

so the system (3.10) is equivalent to

$$\frac{d}{dt}(\Gamma, M) = [(Z, g\Sigma), (\Gamma, M)]. \tag{3.15}$$

As usual, it follows from (3.15) that a solution $(\Gamma(t), M(t))$ to (3.10) lies in an orbit

$$\mathcal{O}_{\Gamma_0, M_0} = \{\mathrm{Ad}(g)(\Gamma_0, M_0) : g \in E(3)\}. \tag{3.16}$$

Of course this implies that $F(\Gamma(t), M(t))$ is independent of t whenever $F : \mathrm{e}(3) \to \mathbb{R}$ is Ad-invariant, i.e., $F(\mathrm{Ad}(g)X) = F(X)$ for all $g \in E(3),\ X \in \mathrm{e}(3)$. We note that the functions

$$F_1(\Gamma, M) = \mathrm{Tr}(M\Gamma^t), \quad F_2(\Gamma, M) = \mathrm{Tr}(\Gamma\Gamma^t) \tag{3.17}$$

have this property. To see this, note that to verify Ad-invariance of a function F it suffices to show that, for each $X, Y \in \mathrm{e}(3)$,

$$\frac{d}{dt} F(e^{t\ \mathrm{ad}\, Y} X)\big|_{t=0} = DF(X)[Y, X] = 0. \tag{3.18}$$

Verifying that F_1 and F_2 have this property is straightforward.

As we have seen, in general the Lie algebra of $G \times_{\mathrm{Ad}^*} \mathfrak{g}^*$ has an Ad-invariant biliear form given by

$$\mathcal{B}((X, \xi), (Y, \eta)) = \langle X, \eta \rangle + \langle Y, \xi \rangle. \tag{3.19}$$

Cf. Chapter I, §10. If also $\mathfrak{g}$ has a G-Ad-invariant quadratic form, say $Q(\cdot, \cdot)$, as is always the case if G is compact, then the Lie algebra of $G \times_{\mathrm{Ad}} \mathfrak{g}$ has an Ad-invariant bilinear form, naturally related to (3.19), namely

$$\mathcal{B}((X, A), (Y, B)) = Q(X, B) + Q(Y, A). \tag{3.20}$$

This induces a linear isomorphism of the Lie algebra $\mathfrak{h}$ of $G \times_{\mathrm{Ad}} \mathfrak{g}$ onto its dual $\mathfrak{h}^*$. This takes adjoint orbits in $\mathfrak{h}$ to coadjoint orbits in $\mathfrak{h}^*$. In particular the standard Poisson structure on $\mathfrak{h}^*$ is transferred to a Poisson structure on $\mathfrak{h}$, yielding a symplectic structure on each adjoint orbit.

This observation applies in particular to (3.16). It naturally follows that the equation (3.15) is of Hamiltonian type, more precisely, of the form

$$\frac{d}{dt}(\Gamma, M) = H_{\mathcal{E}}(\Gamma, M), \tag{3.21}$$

where $\mathcal{E}(\Gamma, M)$ is given by (3.11), with M and Z related as in (2.14). In light of the computations (2.20)–(2.25) of Chapter II, if we use (3.20) with $Q(X, B) = (1/2)\,\mathrm{Tr}(XB^t)$, verifying (3.21) comes down to showing that

$$D\mathcal{E}(\Gamma, M)\,(A, B) = \frac{1}{2}\,\mathrm{Tr}(ZB^t) + \frac{g}{2}\,\mathrm{Tr}(\Sigma A^t), \tag{3.22}$$

which is a straightforward computation from the formula (3.11) for $\mathcal{E}(\Gamma, M)$.

Typical adjoint orbits of the form (3.16) are symplectic manifolds of dimension 4. It follows from general theory that the system (3.20) is integrable if one can find a second function $\mathcal{F}$, functionally independent of $\mathcal{E}$, such that $\{\mathcal{E}, \mathcal{F}\} = 0$. Typically the system (3.20) is not integrable. One classical case where it is integrable, due to Lagrange, will be discussed in the next section.

4. Lagrange's symmetric top

Here we discuss a special class of 3D tops, which were shown by Lagrange to yield integrable systems. Namely, we assume that $\mathcal{I}_\rho$, defined by (2.5), has a double eigenvalue, say a, and another eigenvalue, say b, and that the b-eigenspace is spanned by σ, defined by (2.6). Equivalently, $\mathcal{J}_\rho$ has a double eigenvalue $\tilde{a}$ and an eigenvalue $\tilde{b}$, and its $\tilde{b}$-eigenspace is spanned by σ. Since $\mu(t) = \mathcal{J}_\rho \omega(t)$, this implies that $\mu(t) - \tilde{a}\omega(t)$ is parallel to σ, i.e., there exists $\alpha(t)$ such that

$$\mu(t) - \tilde{a}\omega(t) = \alpha(t)\sigma. \tag{4.1}$$

Then the equation (3.4) implies

$$\mu_t = -\alpha\omega \times \sigma + g\gamma \times \sigma, \tag{4.2}$$

and hence

$$\frac{d}{dt}(\mu \cdot \sigma) = \mu_t \cdot \sigma = 0. \tag{4.3}$$

In other words, $\mu(t) \cdot \sigma$ is another conserved quantity for the Lagrange top. Equivalently,

$$\mathcal{F}(\Gamma, M) = \mathrm{Tr}(M\Sigma^t) \tag{4.4}$$

is conserved for the system (3.15). Another equivalent formulation is that

$$\omega \cdot \sigma = \mathcal{J}_\rho^{-1}\mu \cdot \sigma = \mu \cdot \mathcal{J}_\rho^{-1}\sigma = \tilde{b}^{-1}\,\mu \cdot \sigma \tag{4.5}$$

is conserved.

As we have discussed in §3, (3.15) is a Hamiltonian system of the form

$$\frac{d}{dt}(\Gamma, M) = H_{\mathcal{E}}(\Gamma, M), \quad \mathcal{E}(\Gamma, M) = \frac{1}{4}\,\mathrm{Tr}(MZ^t) + \frac{g}{2}\,\mathrm{Tr}(\Sigma\Gamma^t), \tag{4.6}$$

on the 4-dimensional symplectic manifolds $\mathcal{O}_{\Gamma_0, M_0}$, with Z and M related by (2.14). The fact that $\mathcal{F}$ is constant on integral curves of $H_{\mathcal{E}}$ is equivalent to

$$\{\mathcal{F}, \mathcal{E}\} = 0. \tag{4.7}$$

Thus the system (4.6) is integrable for Lagrange's symmetric top.

Let us take a further look at integrating (3.4)–(3.5) in this case. Rotating coordinates and scaling, we can assume $\sigma = (0, 0, 1)^t$, so the conservation law (4.5) gives

$$\omega = (\omega_1, \omega_2, c_3)^t. \tag{4.8}$$

Then the equation (3.5) takes the form

$$\begin{aligned} \gamma_1' &= c_3\gamma_2 - \omega_2\gamma_3, \\ \gamma_2' &= -c_3\gamma_1 + \omega_1\gamma_3, \\ \gamma_3' &= \omega_2\gamma_1 - \omega_1\gamma_2. \end{aligned} \tag{4.9}$$

Also the conservation law (3.7), i.e., $\mu \cdot \gamma = k$, takes the form

$$\tilde{a}(\omega_1\gamma_1 + \omega_2\gamma_2) + \tilde{b}c_3\gamma_3 = k, \tag{4.10}$$

which implies

$$(k - \tilde{b}c_3\gamma_3)^2 = \tilde{a}^2(\omega_1^2\gamma_1^2 + \omega_2^2\gamma_2^2 + 2\omega_1\omega_2\gamma_1\gamma_2). \tag{4.11}$$

Meanwhile, the last equation in (4.8) implies

$$(\gamma_3')^2 = \omega_2^2\gamma_1^2 + \omega_1^2\gamma_2^2 - 2\omega_1\omega_2\gamma_1\gamma_2, \tag{4.12}$$

and we can use (4.11) to eliminate all the terms containing $\gamma_1\gamma_2$. We get

$$(\gamma_3')^2 = (\omega_1^2 + \omega_2^2)(\gamma_1^2 + \gamma_2^2) - \tilde{a}^{-2}(k - \tilde{b}c_3\gamma_3)^2. \tag{4.13}$$

To streamline this further, we use $|\gamma|^2 \equiv 1$ and the conservation of energy, which yields

$$\omega_1^2 + \omega_2^2 + c_3^2 = 2\mathcal{E} - 2g\gamma_3. \tag{4.14}$$

Thus (4.13) becomes

$$(\gamma_3')^2 = (2\mathcal{E} - c_3^2 - 2g\gamma_3)(1 - \gamma_3^2) - \tilde{a}^{-2}(k - \tilde{b}c_3\gamma_3)^2 = p(\gamma_3), \tag{4.15}$$

where p is a cubic polynomial, whose coefficients involve $\mathcal{E}, c_3, g, k, \tilde{a}$, and $\tilde{b}$. Separating variables, we obtain an elliptic integral:

$$\int \frac{d\gamma_3}{\sqrt{p(\gamma_3)}} = t. \tag{4.16}$$

Chapter 4: The Korteweg-deVries Equation

Introduction

In §9 of Chapter I we derived the Korteweg-deVries equation as the equation of geodesic motion on the Virasoro group, a central extension of the group of diffeomorphisms of the circle S^1. Here we study the KdV equation further. In §1 we recast KdV as a Hamiltonian system on the dual Vir* of the Virasoro algebra, with its natural Lie-Poisson structure. We show how KdV is a bi-Hamiltonian system, also Hamiltonian with respect to a certain frozen Poisson structure on Vir*, which together with the Lie-Poisson structure forms a Poisson pair, as a special case of material covered in §6 of Chapter II. In §2 we recast all this in terms of Poisson structures on the dual to the Lie algebra Vect$(S^1) \approx C^\infty(S^1)$, and in §3 we apply the Lenard scheme, introduced in §6 of Chapter II, and show how it produces a sequence of conservation laws for solutions to KdV.

In §4 we apply the first 3 of these conservation laws to give a demonstration of global existence of smooth solutions to KdV. In fact, these conservation laws yield bounds on the H^2-norm of a solution, which are more than adequate for global existence, in view of the local existence and persistence results we obtain in Appendix A. The results of Appendix A are parallel to familiar results for hyperbolic PDE, such as obtained in [T2].

In fact, there is an infinite string of conservation laws, said to guarantee "integrability" of KdV. We do not give a direct discussion of what this integrability means; that is front and center in many treatments of the subject, mentioned in the references. (A concise characterization of integrability can be found on p. 638 of [Mc2].) However, we proceed to discuss further facets of the production of these conservation laws, all intimately connected to integrability. In §5 we discuss how Lax pairs arise in KdV, and lead to families of isospectral Schrödinger operators, and how the conservation of the spectrum is related to the previously constructed conservation laws, via "heat asymptotics," making use of some results on such asymptotics established in Appendix B.

In §6 we discuss the Gel'fand-Dickii approach to the production and analysis of Lax pairs, which leads to a systematic production of conservation laws, via a "residue" calculation. This relies on a technical result, which eventually got a neat treatment by G. Wilson, whose argument we give in §7. Actually, the material discussed here was developed by these authors in a more general context, to produce further classes of integrable systems, such as the "KP-heirarchy." We have confined the scope here to KdV, but the reader who gets through this material might be well prepared to read about these more general matters, in the papers we cite on this material.

1. KdV as a bi-Hamiltonian system

In Chapter I, §9, we produced the Korteweg-deVries equation as a geodesic equation on the Virasoro group, in particular as an evolution equation for a curve on its Lie algebra,

$$\mathrm{Vir} = C^\infty(S^1) \oplus \mathbb{R}, \tag{1.1}$$

with Lie bracket given by

$$[(u,a),(v,b)] = (u'v - uv', \gamma(u,v)), \tag{1.2}$$

where

$$\gamma(u,v) = (u_x, v_{xx})_{L^2}, \tag{1.3}$$

and with inner product on Vir given by

$$\mathcal{B}((u,a),(v,b)) = (u,v)_{L^2} + ab. \tag{1.4}$$

We complement the discussion in §9 of Chapter I with a sketch of the coadjoint formulation. We use the inner product (1.4) to identify Vir and Vir^*:

$$\mathrm{Vir}^* = C^\infty(S^1) \oplus \mathbb{R}. \tag{1.5}$$

We define $\kappa : \mathrm{Vir} \to \mathrm{Vir}^*$ so that

$$\gamma(u,v) = \langle v, \kappa(u)\rangle, \quad \text{i.e., } \kappa(u) = -u_{xxx}. \tag{1.6}$$

Starting with

$$\langle (v,b), \mathrm{ad}^*(u,a)(w,c)\rangle = -\langle ([u,v], \gamma(u,v)), (w,c)\rangle, \tag{1.7}$$

and using (1.2)–(1.4), we obtain

$$\mathrm{ad}^*(u,a)\,(w,c) = (\mathrm{ad}^*(u)w - c\kappa(u), 0). \tag{1.8}$$

Now the general set-up for geodesic flow, as derived in (2.13) of Chapter I, gives the evolution for a curve $(v(t), c(t))$ in Vir^*:

$$\begin{aligned}(v_t, c_t) &= \mathrm{ad}^*(v,c)\,(v,c) \\ &= (\mathrm{ad}^*(v)v - c\kappa(v), 0).\end{aligned} \tag{1.9}$$

Hence $c_t = 0$, so $c \equiv c_0$. Recall from (5.6) of Chapter I that

$$\mathrm{ad}^*(v)v = \mathcal{L}_v^* v = -3vv_x, \tag{1.10}$$

so the formula (1.6) for $\kappa(v)$ yields for v the Korteweg-deVries equation

$$v_t + 3vv_x - c_0 v_{xxx} = 0, \tag{1.11}$$

in agreement with (9.6) of Chapter I.

We can write this equation as a Hamiltonian equation with respect to the Lie-Poisson structure on Vir^*, given as usual by

$$\{f, g\}(u, a) = \langle [df(u, a), dg(u, a)], (u, a)\rangle. \tag{1.12}$$

By the standard set-up (cf. (2.13) of Chapter II), a Hamiltonian vector field with respect to this Poisson structure is given by

$$H_{F_0}(v, c) = -\operatorname{ad}^*(dF_0(v, c))\,(v, c). \tag{1.13}$$

Hence (1.9) is equivalent to $(v_t, c_t) = H_{F_0}(v, c)$, with

$$F_0(v, c) = -\frac{1}{2}(v, v)_{L^2} - \frac{1}{2}c^2. \tag{1.14}$$

Now, given $(u_0, a_0) \in \mathrm{Vir}^*$, there is also the frozen Poisson structure on Vir^*:

$$\{f, g\}_1(u, a) = \langle [df(u, a), dg(u, a)], (u_0, a_0)\rangle. \tag{1.15}$$

With respect to this Poisson structure, a Hamiltonian vector field is given by

$$H^1_{F_1}(v, c) = -\operatorname{ad}^*(dF_1(v, c))\,(u_0, a_0). \tag{1.16}$$

For the particular frozen Poisson structure we will use here, we take

$$(u_0, a_0) = (1, 0). \tag{1.17}$$

Furthermore, we set

$$F_1(v, c) = -\frac{1}{4}\int_{S^1} (v^3 + cv_x^2)\,dx. \tag{1.18}$$

Note that

$$\begin{aligned} dF_1(v, c)(u, a) &= -\frac{1}{4}\frac{\partial}{\partial t}\int_{S^1} \left[(v + tu)^3 + (c + ta)(v_x + tu_x)^2\right] dx\Big|_{t=0} \\ &= -\frac{1}{4}\int_{S^1} (3v^2 u + av_x^2 + 2cv_x u_x)\,dx, \end{aligned} \tag{1.19}$$

and hence

$$dF_1(v,c) = -\frac{1}{2}\Big(\frac{3}{2}v^2 - cv_{xx}, \frac{1}{2}\|v_x\|^2_{L^2}\Big). \tag{1.20}$$

Then, by (1.8) and (1.10),

$$\begin{aligned} -\operatorname{ad}^*(dF_1(v,c))(1,0) &= \frac{1}{2}\Big(\operatorname{ad}^*\Big(\frac{3}{2}v^2 - cv_{xx}\Big)\cdot 1, 0\Big) \\ &= \frac{1}{2}\Big(\frac{3}{2}\mathcal{L}^*_{v^2}1 - c\mathcal{L}^*_{v_{xx}}1, 0\Big) \\ &= (-3vv_x + cv_{xxx}, 0), \end{aligned} \tag{1.21}$$

since $\mathcal{L}^*_v w = -vw_x - 2v_x w$ (cf. (5.6) of Chapter I).

Thus we have a bi-Hamiltonian structure:

$$H_{F_0} = H^1_{F_1}. \tag{1.22}$$

2. Poisson structures induced on $C^\infty(S^1)$

In §1 we produced two Poisson structures on $\text{Vir}^* = C^\infty(S^1) \oplus \mathbb{R}$, yielding Hamiltonian vector fields of the following form:

$$\begin{aligned} H^0_{F_0}(v,c) &= -\operatorname{ad}^*(dF_0(v,c))\,(v,c) \\ &= -(\operatorname{ad}^*(d_vF_0(v,c))v - c\kappa(d_vF_0(v,c)), 0), \end{aligned} \tag{2.1}$$

and

$$\begin{aligned} H^1_{F_1}(v,c) &= -\operatorname{ad}^*(dF_1(v,c))\,(1,0) \\ &= -(\operatorname{ad}^*(d_vF_1(v,c))1, 0). \end{aligned} \tag{2.2}$$

Recall that

$$\kappa(u) = -\partial_x^3 u, \quad \operatorname{ad}^*(u)v = \mathcal{L}^*_u v = -uv_x - 2u_x v. \tag{2.3}$$

It is clear that the vector fields $H^0_{F_0}$ and $H^1_{F_1}$ are tangent to each hyperplane $c = c_0$. Thus, for each $c_0 \in \mathbb{R}$, we have a pair of Poisson structures on $C^\infty(S^1)$ (in fact, a Poisson pair), yielding vector fields

$$\begin{aligned} H^0_{f_0}(v) &= -\operatorname{ad}^*(df_0(v))v - c_0\kappa(df_0(v)), \\ H^1_{f_1}(v) &= -\operatorname{ad}^*(df_1(v))1, \end{aligned} \tag{2.4}$$

for $f_j : C^\infty(S^1) \to \mathbb{R}$. Here, given $v \in C^\infty(S^1)$, $df_j(v) \in C^\infty(S^1)$ is determined by the identity

$$(df_j(v), w)_{L^2} = \frac{\partial}{\partial t} f_j(v + tw)\big|_{t=0}. \tag{2.5}$$

For example, if $\varphi_j = \varphi_j(s_0, s_1, \dots, s_\ell)$, then

$$f_j(v) = \varphi_j(v, \partial_x v, \dots, \partial_x^\ell v) \Rightarrow df_j(v) = \sum_{k=0}^{\ell} (-1)^k \, \partial_x^k \frac{\partial \varphi_j}{\partial s_k}(v, \dots, \partial_x^\ell v). \tag{2.6}$$

For some further formulas, let us set

$$V_j = df_j(v), \tag{2.7}$$

and bring in (2.3), to write

$$\begin{aligned} H^0_{f_0}(v) &= (-c_0 \partial_x^3 + 2v\partial_x + v_x) V_0 = \mathcal{E} V_0, \\ H^1_{f_1}(v) &= 2\partial_x V_1 = \mathcal{D} V_1. \end{aligned} \tag{2.8}$$

Here

$$\mathcal{E}, \mathcal{D} : C^\infty(S^1) \longrightarrow C^\infty(S^1) \tag{2.9}$$

are the Poisson tensors, defined in (1.14)–(1.15) of Chapter II. Parallel to (1.16) of Chapter II, we have $\mathcal{D}$ and $\mathcal{E}$ skew-adjoint, with respect to the L^2-inner product. Parallel to (1.14) and (1.18), we consider

$$f_0(v) = -\frac{1}{2} \int_{S^1} v^2 \, dx, \quad f_1(v) = -\frac{1}{4} \int_{S^1} (v^3 + c_0 v_x^2) \, dx, \tag{2.10}$$

yielding

$$V_0 = -v, \quad V_1 = \frac{1}{2}\Big(-\frac{3}{2} v^2 + c_0 v_{xx}\Big), \tag{2.11}$$

and hence

$$\begin{aligned} H^0_{f_0}(v) &= \mathcal{E} V_0 = -3vv_x + c_0 v_{xxx}, \\ H^1_{f_1}(v) &= \mathcal{D} V_1 = -3vv_x + c_0 v_{xxx}. \end{aligned} \tag{2.12}$$

Thus the bi-Hamiltonian structure of KdV is re-stated.

It follows immediately that, when $v(t, x)$ is a sufficiently smooth solution to the Korteweg-deVries equation, then $f_0(v)$ and $f_1(v)$, which a priori are functions of

t, are actually independent of t, and hence give conservation laws for solutions to KdV. The following is also seen to be conserved:

$$f_{-1}(v) = \int_{S^1} v\,dx. \tag{2.13}$$

In §3 we will produce an infinite list of additional conserved quantities.

3. Conservation laws and the KdV heirarchy

Here we will construct a sequence $\{f_j : j \geq 0\}$ of functions $f_j : C^\infty(S^1) \to \mathbb{R}$, such that f_0 and f_1 are given by (2.10) and

$$H^0_{f_j} = H^1_{f_{j+1}}, \quad j \geq 0, \tag{3.1}$$

where $f \mapsto H^j_f$ are the Poisson structures on $C^\infty(S^1)$ given by (2.4), or equivalently by (2.8). That is to say, we want to produce $V_j = df_j(v)$ such that

$$\mathcal{E}V_j = \mathcal{D}V_{j+1}, \tag{3.2}$$

where, as in (2.8),

$$\mathcal{E}V = (-c_0\partial_x^3 + 2v\partial_x + v_x)V, \quad \mathcal{D}V = 2\partial_x V. \tag{3.3}$$

The formulas for V_0 and V_1 are given in (2.11).

Thus, to get V_2, we compute that

$$\mathcal{E}V_1 = (-c_0\partial_x^3 + 2v\partial_x + v_x)\Big(-\frac{3}{2}v^2 + c_0\partial_x^2 v\Big) = \mathcal{D}V_2, \tag{3.4}$$

with

$$V_2 = -\frac{c_0^2}{2}\partial_x^4 v + \frac{5}{4}c_0(v_x^2 + 2vv_{xx}) - \frac{5}{4}v^3, \tag{3.5}$$

which satisfies $V_2 = df_2(v)$ with

$$f_2(v) = -\frac{1}{4}\int_{S^1} \Big[c_0^2(v_{xx})^2 + 5c_0 v v_x^2 + \frac{5}{4}v^4\Big]\,dx. \tag{3.6}$$

Note that the form of $\mathcal{D}$ implies that the Poisson structure $\{\cdot,\cdot\}_1$ is symplectic, so the fact that one can continue indefinitely with (3.2) is suggested by Proposition 6.3 of Chapter II. However, that result applies literally only in the finite-dimensional case. The following furnishes justification here.

Lemma 3.1. *For each $j \geq 2$, there exists V_{j+1} such that (3.2) holds. One has*

$$V_j = \alpha_j c_0^j \partial_x^{2j} v + \Phi_j(v, \dots, \partial_x^{2j-2} v). \tag{3.7}$$

Proof. Let us set

$$\mathcal{R} = \mathcal{D}^{-1}\mathcal{E}, \tag{3.8}$$

so we desire to show that

$$V_j = \mathcal{R} V_{j-1}, \tag{3.9}$$

for all j, and so far we have it for $j = 1, 2$. To provide an inductive proof, we need to show that, as long as (3.9) holds, $\mathcal{E}V_j$ belongs to $R(\mathcal{D})$, the range of $\mathcal{D} = 2\partial_x$. In view of the formula (3.3) for $\mathcal{E}$, it suffices to show that $v\mathcal{D}V_j \in R(\mathcal{D})$. Note that, by (2.11),

$$v\mathcal{D}V_j = -v\mathcal{D}\mathcal{R}^j v. \tag{3.10}$$

Now formal integration by parts produces the identity

$$(\mathcal{D}^{-1}\mathcal{E})^* = \mathcal{E}\mathcal{D}^{-1}, \quad \text{hence} \;\; \mathcal{R}^* = \mathcal{D}\mathcal{R}\mathcal{D}^{-1}. \tag{3.11}$$

Thus we have, for some A_j,

$$\begin{aligned} v\mathcal{D}V_j &= ((\mathcal{R}^*)^j \mathcal{D}v)v + \partial_x A_j \\ &= (\mathcal{D}\mathcal{R}^j v)v + \partial_x A_j, \end{aligned} \tag{3.12}$$

hence

$$v\mathcal{D}V_j = \frac{1}{2}\partial_x A_j, \tag{3.13}$$

proving the lemma.

From here we have, as in §6 of Chapter II:

Proposition 3.2. *For each $j \geq 0$, there exist $f_j : C^\infty(S^1) \to \mathbb{R}$ such that (3.1) holds. Furthermore,*

$$\{f_j, f_k\}_0 = \{f_j, f_k\}_1 = 0, \quad \forall\, j, k \geq 0. \tag{3.14}$$

The sequence of Hamiltonian vector fields $\{H^0_{f_j} : j \geq 0\}$ is called the KdV heirarchy. It follows from (3.14) that each f_j provides a *conservation law* for sufficiently smooth solutions to the KdV equation (1.11).

4. Global existence of solutions to KdV

As shown in Appendix A to this chapter, the initial value problem

$$v_t + 3vv_x - c_0 v_{xxx} = 0, \quad v(0,x) = u(x), \tag{4.1}$$

has a short-time solution, given $u \in C^\infty(S^1)$, and this solution does not break down as long as one has a bound on $\|v(t)\|_{C^1(S^1)}$. Here we will show that, for any given nonzero, real c_0, there is such a bound, and hence we have global existence of a smooth solution to (4.1). We will make use of the following conservation laws, established in §§2–3:

$$\begin{aligned} E_0(v) &= \int_{S^1} v^2\,dx, \\ E_1(v) &= \int_{S^1} (v_x^2 + c_0^{-1} v^3)\,dx, \\ E_2(v) &= \int_{S^1} \Big(v_{xx}^2 + 5c_0^{-1} v v_x^2 + \frac{5}{4} c_0^{-2} v^4\Big)\,dx. \end{aligned} \tag{4.2}$$

Each $E_j(v)$ is a constant multiple of $f_j(v)$, given in (2.10) and (3.6).

We compare these quantities with

$$H_j(v) = \int_{S^1} (\partial_x^j v)^2\,dx. \tag{4.3}$$

Clearly $E_0(v) = H_0(v)$. We can obtain upper bounds on $H_1(v)$ and $H_2(v)$ as follows. First we have

$$\|v\|_{L^\infty}^2 \le K^2 H_0(v) + K^2 H_1(v), \tag{4.4}$$

and hence

$$\Big|\int v^3\,dx\Big| \le \|v\|_{L^\infty} H_0(v) \le K H_0(v)^{3/2} + K H_0(v) H_1(v)^{1/2}. \tag{4.5}$$

Thus

$$H_1(v) \le E_1(v) + K|c_0^{-1}| E_0(v)^{3/2} + K|c_0^{-1}| E_0(v) H_1(v)^{1/2}. \tag{4.6}$$

Writing this as

$$y^2 \le B + 2Ay, \quad y^2 = H_1(v),$$

and hence $y^2 - 2Ay \le B$, we have $y \le A + \sqrt{B + A^2}$, hence $y^2 \le 4A^2 + 2B$, i.e.,

$$H_1(v) \le 2E_1(v) + 2K|c_0^{-1}|E_0(v)^{3/2} + K^2 c_0^{-2} E_0(v)^2. \tag{4.7}$$

Having this, we next immediately obtain

$$H_2(v) \le E_2(v) + 5|c_0^{-1}|\big(E_0(v)^{1/2} + H_1(v)^{1/2}\big)H_1(v). \tag{4.8}$$

Since $\|v\|_{C^1}^2 \le CH_2(v) + CH_0(v)$, the asserted bound on solutions to (4.1) follows from the conservation of the three quantities in (4.2), and this leads to global existence. We give a formal statement of such a result.

Proposition 4.1. *Given $u \in H^k(S^1)$, $k \ge 2$, there exists a unique global solution*

$$v \in C(\mathbb{R}, H^k(S^1)) \tag{4.9}$$

to the initial value problem (4.1). The quantities $E_j(v(t))$ for $0 \le j \le 2$ given by (4.2) are independent of t.

Proof. As we have advertised, global existence of v satisfying (4.1) and (4.9) follows from the results of Appendix A once one has a bound on $\|v(t)\|_{H^2}$, and by (4.8) such a bound follows from the fact that $E_j(v(t))$ are independent of t, for $0 \le j \le 2$.

Now in the derivation of these conservation laws we assumed v was "smooth." We need to show that these laws hold when v has limited regularity. We can analyze

$$\begin{aligned} \frac{d}{dt}E_2(v(t)) &= (v_{txx}, v_{xx}) + \frac{5}{c_0}(v_t, v_x^2) + \frac{10}{c_0}(v_{tx}, vv_x) + \frac{5}{c_0^2}(v_t, v^3) \\ &= c_0(\partial_x^5 v, \partial_x^2 v) + \cdots, \end{aligned} \tag{4.10}$$

and show this is equal to zero, given $v \in C(I, H^k(S^1))$ satisfying (4.1), with $k \ge 4$. Then we readily have the stated results of Proposition 4.1 in case $k \ge 4$. It remains to show that Proposition 4.1 holds for $k = 2$ and 3.

We can establish this as follows. Using a Friedrichs mollifier J_ε, set $u_k = J_\varepsilon u \in C^\infty(S^1)$, with $\varepsilon = 2^{-k}$, and solve

$$\partial_t v_k + 3v_k \partial_x v_k - c_0 \partial_x^3 v_k = 0, \quad v_k(0, x) = u_k(x), \tag{4.11}$$

obtaining $v_k \in C^\infty(\mathbb{R} \times S^1)$, by the results established above. Furthermore, for each k, we have

$$E_j(v_k(t)) = E_j(u_k) = e_{jk}, \tag{4.12}$$

independent of t, for $j = 0, 1, 2$. This has the following implications:

$$\begin{aligned} v_k &\text{ bounded in } L^\infty(\mathbb{R}, H^2(S^1)), \\ \partial_t v_k &\text{ bounded in } L^\infty(\mathbb{R}, H^{-1}(S^1)). \end{aligned} \tag{4.13}$$

We deduce that, upon passing to a subsequence (still denoted v_k), we have, for each $T < \infty,\ \delta > 0$,

$$\begin{aligned} v_k \to v \ \text{ in } \ L^\infty(\mathbb{R}, H^2(S^1)),\ \text{weak}^*, \\ v_k \to v \ \text{ in } \ C([-T,T], H^{2-\delta}(S^1)), \ \text{ in norm}, \\ v_k(t) \to v(t) \ \text{ in } \ H^2(S^1),\ \text{weak}^*,\ \forall\, t \in \mathbb{R}. \end{aligned} \tag{4.14}$$

It readily follows that v is the unique solution in $L^\infty(\mathbb{R}, H^2(S^1))$ to (4.1). Furthermore, we have norm convergence

$$\begin{aligned} v_k(\partial_x v_k)^2 \to v v_x^2 \ \text{ in } \ C([-T,T], H^{1-\delta}(S^1)), \\ v_k^4 \to v^4 \ \text{ in } \ C([-T,T], H^{2-\delta}(S^1)). \end{aligned} \tag{4.15}$$

Hence we have

$$E_2(v_k) - H_2(v_k) \longrightarrow E_2(v) - H_2(v), \tag{4.16}$$

locally uniformly in t, as $k \to \infty$. Furthermore,

$$E_2(v_k(t)) \equiv e_{2k} \to E_2(u), \tag{4.17}$$

so, for each t, as $k \to \infty$,

$$H_2(v_k(t)) \to H_2(v(t)) + E_2(u) - E_2(v(t)). \tag{4.18}$$

On the other hand, the third result in (4.14) implies

$$\limsup_{k\to\infty}\ H_2(v_k(t)) \ge H_2(v(t)), \tag{4.19}$$

for each t, so we deduce that

$$E_2(v(t)) \le E_2(u) = E_2(v(0)), \tag{4.20}$$

for each $t \in \mathbb{R}$.

This bound suffices to complete the global existence result in Proposition 4.1 when $k \ge 2$. As for the conservation of $E_2(v(t))$, we can get this from (4.20) by the following simple device. We can start the evolution at an arbitrary time, say t_0, and then the reasoning leading to (4.20) gives

$$E_2(v(t)) \le E_2(v(t_0)), \quad \forall\, t, t_0 \in \mathbb{R}. \tag{4.21}$$

Then reversing the roles of t and t_0 gives the reverse inequality, hence equality. The task of establishing conservation of $E_j(v)$, for $j = 0, 1$, is more elementary.

5. The Lax pair approach to conservation laws

Given $v = v(t,x)$, defined for $t \in I,\ x \in S^1$, and given $b \in \mathbb{R}$, consider

$$L(t) = \partial_x^2 - bv(t,x), \tag{5.1}$$

a one-parameter family of differential operators on S^1, which are self-adjoint, with discrete spectrum. As observed by P. Lax, one can form

$$M(t) = a\partial_x^3 + v\partial_x + \partial_x v = a\partial_x^3 + 2v\partial_x + v_x, \tag{5.2}$$

and the equation

$$\frac{\partial L}{\partial t} = [cM, L], \tag{5.3}$$

for certain choices of constants a, b, c, holds precisely when v solves the Korteweg-deVries equation

$$v_t + 3vv_x - c_0 v_{xxx} = 0. \tag{5.4}$$

From this it can be deduced that all the operators $L(t)$ are unitarily equivalent, and hence have the same spectrum. This gives rise to conservation laws, providing an alternative route to that described in §3. The pair (L, cM) is called a *Lax pair.*

To see how this works out, we compute that, when L and M are given by (5.1)–(5.2),

$$[M, L]f = -(3ab + 4)(v_x\partial_x^2 f + v_{xx}\partial_x f) - [2bvv_x + (ab + 1)v_{xxx}]f. \tag{5.5}$$

We take $ab = -4/3$, and then

$$[M, L]f = -\Big(2bvv_x - \frac{1}{3}v_{xxx}\Big)f. \tag{5.6}$$

Then the equation (5.3) holds if and only if

$$v_t = 2cvv_x - \frac{c}{3b}v_{xxx}. \tag{5.7}$$

Thus (5.3) is equivalent to (5.4), provided

$$a = \frac{8}{3}c_0, \quad b = -\frac{1}{2c_0}, \quad c = -\frac{3}{2}. \tag{5.8}$$

To continue, $M(t)$ in (5.2) is skew-adjoint for each t, the solution operator $U(t)$ to

$$\frac{\partial w}{\partial t} = cM(t)w, \quad w(t) = U(t)w(0), \tag{5.9}$$

is unitary on $L^2(S^1)$ for each t, and we have

$$L(t) = U(t)L(0)U(t)^*. \tag{5.10}$$

Hence, for all t in an interval on which we have a solution to (5.4),

$$\operatorname{Spec} L(t) = \operatorname{Spec} L(0). \tag{5.11}$$

In particular, for each $s > 0$,

$$\operatorname{Tr} e^{s(\partial_x^2 - bv(t))} = \operatorname{Tr} e^{s(\partial_x^2 - bv(0))}. \tag{5.12}$$

Now, as shown in Appendix B, there is an asymptotic expansion as $s \searrow 0$:

$$\operatorname{Tr} e^{s(\partial_x^2 - bv)} \sim \sum_{k \geq 0} s^{-1/2+k} \widetilde{E}_k(bv), \tag{5.13}$$

with coefficients $\widetilde{E}_k(bv)$ described by (B.30). These are hence conserved quantities for solutions to (5.4), conservation laws that can be compared to those produced in §3.

We next present a direct proof, adapted from [Lax2], that the conserved quantities in (5.13) Poisson commute, at least for the Poisson structure $\{\cdot,\cdot\}_1$ given by the second formula of (2.4), or equivalently (2.8), i.e.,

$$\{f, g\}_1(v) = 2 \int_{S^1} (\partial_x V) W \, dx, \tag{5.14}$$

where

$$V = df(v), \ W = dg(w) \in C^\infty(S^1). \tag{5.15}$$

To begin, say

$$\operatorname{Spec}(\partial_x^2 - bv) = \{-\lambda_j(v)\}, \quad \lambda_1(v) \leq \lambda_2(v) \leq \cdots. \tag{5.16}$$

The result in Theorem 6.4 of [Lax2] is:

Lemma 5.1. *We have, for each j,k,*

$$\{\lambda_j, \lambda_k\}_1 = 0. \tag{5.17}$$

Proof. Let w_j be the real-valued λ_j-eigenfunction of $\partial_x^2 - bv$, uniquely determined up to a factor of ± 1 (as long as λ_j is a simple eigenvalue) by the normalization $\|w_j\|_{L^2} = 1$. A calculation gives

$$d\lambda_j(v) = bw_j^2. \tag{5.18}$$

Another calculation gives

$$Hw_j^2 = -4\lambda_j\,\partial_x w_j^2, \tag{5.19}$$

where

$$H = \partial_x^3 - 4bv\partial_x - 2bv_x. \tag{5.20}$$

Hence we have

$$\begin{aligned} \{\lambda_j, \lambda_k\}_1(v) &= 2(\partial_x\, d\lambda_j(v), d\lambda_k(v))_{L^2} \\ &= 2b^2(\partial_x w_j^2, w_k^2) \\ &= -\frac{b^2}{2\lambda_j}(Hw_j^2, w_k^2), \end{aligned} \tag{5.21}$$

the last identity by (5.19). Note that $H^* = -H$, so this yields

$$\{\lambda_j, \lambda_k\}_1(v) = \frac{b^2}{2\lambda_j}(Hw_k^2, w_j^2). \tag{5.22}$$

Now interchanging the roles of j and k in (5.21) gives

$$\{\lambda_k, \lambda_j\}_1(v) = -\frac{b^2}{2\lambda_k}(Hw_k^2, w_j^2). \tag{5.23}$$

But the left sides of (5.22) and (5.23) are negatives of each other, so we obtain

$$\lambda_j\{\lambda_j, \lambda_k\}_1 = \lambda_k\{\lambda_j, \lambda_k\}_1, \tag{5.24}$$

which implies (5.17) whenever $\lambda_j \neq \lambda_k$.

Now Lemma 5.1 is neat as far as it goes, but it depends upon the (unstated) hypothesis of simple spectrum, which we do not always have. To fix this up, let us set, for each $T \in (0, \infty)$,

$$\mathcal{O}_T = \{v \in C^\infty(S^1) : \text{All } \lambda_j \text{ such that } |\lambda_j(v)| \le T \text{ are simple}\}. \tag{5.25}$$

It is easy to show that $\mathcal{O}_T$ is open in $C^\infty(S^1)$, and it is also known that $\mathcal{O}_T$ is dense, for each $T < \infty$. Certainly (5.17) holds on $\mathcal{O}_T$, provided $|\lambda_j|, |\lambda_k| < T$. From here we proceed as follows. Let

$$\widetilde{\mathcal{S}} = \{\psi \in C^\infty(\mathbb{R}) : \psi(t) \text{ is rapidly decreasing as } t \to -\infty\}, \tag{5.26}$$

and for $\psi \in \widetilde{\mathcal{S}}$ set

$$f_\psi : C^\infty(S^1) \to \mathbb{R}, \quad f_\psi(v) = \operatorname{Tr}\psi(\partial_x^2 - bv). \tag{5.27}$$

We have:

Corollary 5.2. *Given* $T \in (0, \infty)$,

$$\varphi, \psi \in C_0^\infty((-T,T)) \Longrightarrow \{f_\varphi, f_\psi\}_1 = 0 \quad \text{on } \mathcal{O}_T. \tag{5.28}$$

Hence

$$\varphi, \psi \in C_0^\infty(\mathbb{R}) \Longrightarrow \{f_\varphi, f_\psi\}_1 \equiv 0. \tag{5.29}$$

Hence

$$\varphi, \psi \in \widetilde{\mathcal{S}} \Longrightarrow \{f_\varphi, f_\psi\}_1 \equiv 0. \tag{5.30}$$

Proof. The discussion above gives (5.28). Then (5.29) follows from the denseness of $\mathcal{O}_T$ in $C^\infty(S^1)$, and then (5.30) follows by another approximation argument, plus the fact that

$$\operatorname{Spec}(\partial_x^2 - bv) \subset (-\infty, A(v)], \quad A(v) = \sup_x (-bv(x)). \tag{5.31}$$

From here it readily follows that, if we set

$$E_t(v) = \operatorname{Tr} e^{t(\partial_x^2 - bv)}, \tag{5.32}$$

then

$$s, t > 0 \Longrightarrow \{E_s, E_t\}_1 \equiv 0. \tag{5.33}$$

Then the analysis yielding (5.13) also gives

$$\{E_j^b, E_k^b\}_1 \equiv 0, \quad \forall\, j, k, \tag{5.34}$$

where $E_j^b(v) = \widetilde{E}_j(bv)$.

6. The Gel'fand-Dickii approach

Here we discuss an approach taken in [GD] (also pursued in [Ad], and in [SeW] and [Wi]), involving the following algebra Ψ of "formal pseudodifferential operators" on S^1. An element of Ψ has the form

$$P = \sum_{k=-\infty}^{m} p_k(x) \partial^k, \tag{6.1}$$

where $p_k \in C^\infty(S^1)$, and m (the order of P) is an integer; we also say $P \in \Psi^m$. The *symbol* of P is

$$p = \sum_{k=-\infty}^{m} p_k(x)\xi^k. \tag{6.2}$$

If also $Q = \sum_{j\le n} q_j(x)\partial^j$, we define $PQ \in \Psi^{m+n}$ in such a fashion that the derivation identity holds:

$$\partial f = f\partial + f_x. \tag{6.3}$$

This leads to the identity $PQ = R$, with symbol

$$r = p \circ q = \sum_{\ell\ge 0} \frac{1}{\ell!}(\partial_\xi^\ell p)(\partial_x^\ell q). \tag{6.4}$$

For example, if $m \in \mathbb{N}$, $q = q(x)$,

$$\xi^{-m} \circ q = \sum_{\ell\ge 0} (-1)^\ell \binom{m+\ell-1}{\ell} q^{(\ell)}(x)\,\xi^{-m-\ell}. \tag{6.5}$$

The algebra $\mathrm{DO}(S^1) = \cup_{m\ge 0} \mathrm{DO}^m(S^1)$ of differential operators on S^1, consisting of elements of the form $P = \sum_{k=0}^m p_k(x)\partial^k$, is a subalgebra of Ψ.

We mention that there is an algebra $\Psi^*(S^1)$ of *operators* on $\mathcal{D}'(S^1)$,

$$\Psi^*(S^1) \subset OPS^*(S^1) = \bigcup_{m\in\mathbb{Z}} OPS^m(S^1), \tag{6.6}$$

and a surjective homomorphism

$$\sigma : \Psi^*(S^1) \longrightarrow \Psi, \tag{6.7}$$

whose kernel consists of smoothing operators on $\mathcal{D}'(S^1)$. See, e.g., Chapter 2 of [T]. However, this algebra of pseudodifferential operators will not be used here, just the algebra Ψ described above.

The algebra Ψ will be used to construct a sequence of differential operators

$$P_k \in \mathrm{DO}^{2k+1}(S^1), \tag{6.8}$$

such that

$$[P_k, L] \in \mathrm{DO}^0(S^1), \quad L = \partial^2 - v. \tag{6.9}$$

Then the equations

$$\frac{\partial L}{\partial t} = [P_k, L] \tag{6.10}$$

are equivalent to various PDE for v. The case $k = 1$ will be seen to be (essentially) KdV. Furthermore, a construction involving the "residue" of an element of Ψ (to be defined below) will produce an infinite sequence of conservation laws, valid simultaneously for all of these PDEs.

The analysis begins with the construction of a "square root" of L, of the form

$$L^{1/2} = \partial + Q, \quad Q \in \Psi^{-1}. \tag{6.11}$$

The symbol q of Q is uniquely determined, as follows. We want

$$\begin{aligned} \xi^2 - v = (\xi + q) \circ (\xi + q) &= \xi^2 + \xi \circ q + q \circ \xi + q \circ q \\ &= \xi^2 + 2q\xi + q_x + q \circ q. \end{aligned} \tag{6.12}$$

Write

$$q = \sum_{j \ge 1} q_j(x)\xi^{-j}. \tag{6.13}$$

We then have

$$2q_1 = -v, \quad 2q_2 = -(\partial_x q_1), \tag{6.14}$$

and, for $k \ge 3$,

$$2q_k = -(\partial_x q_{k-1}) - (q \circ q)_{k-1}, \tag{6.15}$$

where $(q \circ q)_{k-1}$ denotes the coefficient of $\xi^{-(k-1)}$ in $q \circ q$, a coefficient which is determined by $\{q_1, \dots, q_{k-2}\}$, via (6.4).

Having constructed $L^{1/2}$, we set

$$P_k = (L^{k+1/2})_+, \tag{6.16}$$

where in general for $P \in \Psi$ given by (6.1), with $m \ge 0$,

$$P_+ = \sum_{k=0}^{m} p_k(x)\partial^k. \tag{6.17}$$

To take one example, note that

$$\begin{aligned} L^{1+1/2} &= (\partial^2 - v)\Big(\partial - \frac{1}{2}v\partial^{-1} + \frac{1}{4}v_x\partial^{-2} + \cdots\Big) \\ &= -\partial^3 - v\partial - \frac{1}{2}(v\partial + 2v_x) + \frac{1}{4}v_x, \quad \text{mod } \Psi^{-1}, \end{aligned} \tag{6.18}$$

i.e.,

$$P_1 = \partial^3 - \frac{3}{2} v\partial - \frac{3}{4} v_x. \tag{6.19}$$

We next note that

$$[P_k, L] = -[(L^{k+1/2})_-, L] = M_k \in \mathrm{DO}^0(S^1), \tag{6.20}$$

where, parallel to (6.17), for general $P \in \Psi$ as in (6.1), we set

$$P_- = P - P_+ = \sum_{k \le -1} p_k(x)\partial^k. \tag{6.21}$$

The first identity in (6.20) arises from the elementary identity

$$[L^{k+1/2}, L] = 0. \tag{6.22}$$

Clearly the left side of (6.20) belongs to $\mathrm{DO}(S^1)$, while the right side belongs to Ψ^0, so the commutator belongs to $\mathrm{DO}^0(S^1)$, as asserted, i.e., it is a multiplication operator, say multiplication by $M_k(x)$. From the left side of (6.20), we see that $M_k(x)$ is a polynomial in $v(x)$ and its derivatives, of order $\le 2k+1$. Note that if we take $b = 1$ in (5.1) and $a = -4/3$ in (5.2), then the computation (5.5) yields

$$[P_1, L] = M_1(x) = \frac{1}{2} v v_x - \frac{1}{4} v_{xxx}. \tag{6.23}$$

As advertised in (6.10), the kth PDE in the heirarchy produced by this construction is defined by

$$\frac{\partial L}{\partial t} = [P_k, L], \quad P_k = (L^{k+1/2})_+. \tag{6.24}$$

In view of the calculation (6.20), this becomes

$$\frac{\partial v}{\partial t} = -M_k, \tag{6.24A}$$

the right side involving v and its derivatives of order $\le 2k+1$. For $k = 1$, we have from (6.23) the equation

$$\frac{\partial v}{\partial t} = -\frac{1}{2} v v_x + \frac{1}{4} v_{xxx}. \tag{6.25}$$

This is a variant of the KdV equation (5.4). It can be converted to (5.4) via a simple transformation.

The next result provides one of the keys for producing conserved quantities by this method. We will present the proof of this result in §7.

Lemma 6.1. *If $L = \partial^2 - v(t,x)$ solves (6.24), then, for each $m \in \mathbb{Z}^+$,*

$$\frac{\partial}{\partial t} L^{m+1/2} = [P_k, L^{m+1/2}]. \tag{6.26}$$

The mechanism given in [Ad] for producing conserved quantities involves the *residue* of an element $P \in \Psi$, defined by

$$\operatorname{Res} P = p_{-1}(x), \tag{6.27}$$

when P has the form (6.1). We will establish the following.

Proposition 6.2. *If $v = v(t,x)$ and $L = \partial^2 - v$ solves (6.24), then, for each $m \in \mathbb{Z}^+$,*

$$\int_{S^1} (\operatorname{Res} L^{m+1/2})\, dx \tag{6.28}$$

is a conserved quantity.

Proof. The identity (6.26) implies that the t-derivative of (6.28) is equal to

$$\int_{S^1} \operatorname{Res}([P_k, L^{m+1/2}])\, dx. \tag{6.29}$$

That this vanishes is a consequence of the following general result.

Lemma 6.3. *Given $A, B \in \Psi$, we have*

$$\int_{S^1} \operatorname{Res}\,[A,B]\, dx = 0. \tag{6.30}$$

Proof. One readily verifies that it suffices to establish this for

$$A = a\partial^k, \quad B = b\partial^{-m}, \quad k,m > 0,\ k \geq m. \tag{6.31}$$

We have

$$AB = \sum_{j=0}^{k-1} \binom{k}{j} a b^{(j)}\, \partial^{k-m-j}, \tag{6.32}$$

and, using (6.5), we have

$$BA = \sum_{j=0}^{\infty} (-1)^j \binom{m+j-1}{j} b a^{(j)}\, \partial^{k-m-j}. \tag{6.33}$$

It follows that

$$\operatorname{Res}[A, B] = \binom{k}{k-m+1} \big[ab^{(k-m+1)} + (-1)^{k-m} ba^{(k-m+1)}\big]. \tag{6.34}$$

An integration by parts shows that the integral over S^1 of this quantity vanishes, so we have (6.30).

Note that $\operatorname{Res} L^{1/2} = -(1/2)v$, so the case $m = 0$ of Proposition 6.2 asserts that $\int v\,dx$ is conserved for solutions to the KdV equation (6.25) (which is quite elementary) and all the other PDE of the form (6.24A). Looking at $m = 1$, we extend the calculation (6.18) a bit, to obtain

$$\begin{aligned} L^{1+1/2} &= \Big(\partial - \frac{1}{2}v\partial^{-1} + \frac{1}{4}v_x\partial^{-2} + \frac{1}{8}(2v^2 - v_{xx})\partial^{-3} + \cdots\Big)(\partial^2 - v) \\ &= \partial^3 - \frac{1}{2}v\partial + \frac{1}{4}v_x + \frac{1}{8}(2v^2 - v_{xx})\partial^{-1} + \cdots \\ &\quad - v\partial - v_x + \frac{1}{4}v^2\partial^{-1} + \cdots \\ &= \partial^3 - \frac{3}{2}v\partial - \frac{3}{4}v_x + \frac{1}{8}(4v^2 - v_{xx})\partial^{-1} + \cdots . \end{aligned} \tag{6.35}$$

Hence

$$\operatorname{Res} L^{1+1/2} = \frac{1}{8}(4v^2 - v_{xx}), \tag{6.36}$$

and hence

$$\int_{S^1} v^2\,dx \tag{6.37}$$

is conserved for solutions to KdV (as we have seen before) and the other equations (6.24A). Computing $\operatorname{Res} L^{m+1/2}$ for $m = 2, 3$ will yield other conservation laws of the form (4.2), and higher values of m will yield still more conservation laws, equivalent to those that arise via (5.13).

7. A graded algebra of formal pseudodifferential operators

Here we set up another algebra, slightly different from the algebra Ψ used in §6, and use it to prove Lemma 6.1. The treatment here is adapted from [Wi].

Let $\mathcal{B}$ be the algebra with unit over $\mathbb{R}$ generated by $\{v^{(j)} : j \in \mathbb{Z}^+\}$, with $v^{(0)} = v$, and let $\partial : \mathcal{B} \to \mathcal{B}$ be the derivation satisfying $\partial v^{(j)} = v^{(j+1)}$. Let $\Psi_{\mathcal{B}}$ consist of symbols of the form

$$p = \sum_{k=-\infty}^{m} p_k\,\xi^k, \quad p_k \in \mathcal{B}. \tag{7.1}$$

We say such p belongs to $\Psi^m_{\mathcal{B}}$. If also $q=\sum_{k\le n} q_k\xi^k\in\Psi^n_{\mathcal{B}}$, we define $p\circ q\in\Psi^{m+n}_{\mathcal{B}}$ by the formula (6.4), with $\partial_x=\partial$.

We also define the following "gradings" on $\mathcal{B}$ and on $\Psi_{\mathcal{B}}$. We set

$$\deg v^{(j)}=j+2,\quad \deg v^{(j_1)}\cdots v^{(j_k)}=(j_1+2)+\cdots+(j_k+2), \tag{7.2}$$

and define $\mathcal{B}^\alpha$ to consist of sums of monomials of degree α, for $\alpha\in\mathbb{Z}^+$. We also set $\deg\xi=1$, and define $\Psi^{;\alpha}_{\mathcal{B}}$ to consist of elements of the form (7.1) with $p_k\in\mathcal{B}^{\alpha-k}$ (so $p_k\ne 0\Rightarrow \alpha\ge k$). We also set

$$\Psi^{m,\alpha}_{\mathcal{B}}=\Psi^m_{\mathcal{B}}\cap\Psi^{;\alpha}_{\mathcal{B}}. \tag{7.3}$$

It is readily verified that

$$p\in\Psi^{;\alpha}_{\mathcal{B}},\ q\in\Psi^{;\beta}_{\mathcal{B}}\Longrightarrow p\circ q\in\Psi^{;\alpha+\beta}_{\mathcal{B}}, \tag{7.4}$$

and more precisely,

$$p\in\Psi^{m,\alpha}_{\mathcal{B}},\ q\in\Psi^{n,\beta}_{\mathcal{B}}\Longrightarrow p\circ q\in\Psi^{m+n,\alpha+\beta}_{\mathcal{B}}. \tag{7.5}$$

One of the principal objects of our attention is

$$L=\xi^2+v\in\Psi^{2,2}_{\mathcal{B}}. \tag{7.6}$$

Compared with the formula for L in (6.9), we are here sticking to the "symbol" notation (using ξ instead of ∂), and we have made a (harmless) change in sign in v. Computations as in (6.11)–(6.15) yield

$$L^{1/2}=\xi+\frac{1}{2}v\xi^{-1}+\frac{1}{4}v^{(1)}\xi^{-2}+\cdots\in\Psi^{1,1}_{\mathcal{B}}. \tag{7.7}$$

More generally,

$$L^{k/2}\in\Psi^{k,k}_{\mathcal{B}}, \tag{7.8}$$

for $k\in\mathbb{Z}^+$. In fact, as is easily verified, (7.8) holds for all $k\in\mathbb{Z}$, positive or negative.

We denote by $\mathcal{Z}(L)$ the set of elements of $\Psi_{\mathcal{B}}$ that commute with L. The following characterization of $\mathcal{Z}(L)$ will be very useful.

Lemma 7.1. *Given $P\in\mathcal{Z}(L)$, there exist integers $m_1>m_2>\cdots$ and $a_j\in\mathbb{R}$ such that*

$$P=a_1L^{m_1/2}+a_2L^{m_2/2}+\cdots. \tag{7.9}$$

Proof. Say $P \in \Psi^m_{\mathcal{B}}$ has the form (7.1), with $p_m \neq 0$. Thus $[P, L] \in \Psi^{m+1}_{\mathcal{B}}$ has the property

$$\begin{aligned} [P, L] &= P\xi^2 - \xi^2 \circ P \qquad \mod \Psi^m_{\mathcal{B}} \\ &= 2(\partial p_m)\xi^{m+1} \qquad \mod \Psi^m_{\mathcal{B}}, \end{aligned} \tag{7.10}$$

so if $[L, P] = 0$ then necessarily $\partial p_m = 0$, so $p_m \in \mathcal{B}^0$, i.e., p_m is a *constant.* So set $m_1 = m,\ a_1 = p_m$, and we have

$$P_2 = P - a_1 L^{m_1/2} \in \Psi^{m_2}_{\mathcal{B}}, \quad m_2 < m_1, \quad [P_2, L] = 0. \tag{7.11}$$

An inductive argument yields (7.9).

Given $P \in \mathcal{Z}(L)$, we set

$$\partial_P v = \text{ coefficient of } \ \xi^0 \ \text{ in } \ [P_+, L], \tag{7.12}$$

where, as usual, if P has the form (7.1),

$$P_+ = \sum_{k=0}^{m} p_k\, \xi^k. \tag{7.13}$$

Having (7.12), we then define $\partial_P : \mathcal{B} \to \mathcal{B}$ in such a fashion that it is a derivation, commuting with ∂. That is, we set

$$\partial_P v^{(j)} = \partial^j(\partial_P v), \tag{7.14}$$

and then define ∂_P on $v^{(j_1)} \cdots v^{(j_k)}$ to act as a derivation. We then define

$$\partial_P : \Psi_{\mathcal{B}} \longrightarrow \Psi_{\mathcal{B}} \tag{7.15}$$

to act componentwise, so

$$\partial_P(q_k\, \xi^k) = (\partial_P q_k)\xi^k, \quad q_k \in \mathcal{B}.$$

It readily follows that ∂_P is a derivation on $\Psi_{\mathcal{B}}$, i.e.,

$$\partial_P(Q_1 \circ Q_2) = (\partial_P Q_1) \circ Q_2 + Q_1 \circ (\partial_P Q_2), \quad \forall\ Q_j \in \Psi_{\mathcal{B}}. \tag{7.16}$$

The following is an elementary precursor to the main result of this section.

Lemma 7.2. *Given $P \in \mathcal{Z}(L)$, P_+ as in (7.13),*

$$\partial_P L = [P_+, L]. \tag{7.17}$$

Proof. It is clear from the definitions that the left side of (7.17) is equal to $\partial_P v$. As for the right side of (7.17), we have, with notations parallel to (6.8)–(6.9),

$$P_+ \in \mathrm{DO}^m_{\mathcal{B}},\ L \in \mathrm{DO}^2_{\mathcal{B}} \Longrightarrow [P_+, L] \in \mathrm{DO}^{m+1}_{\mathcal{B}}, \tag{7.18}$$

while

$$P_- = P - P_+ \in \Psi^{-1}_{\mathcal{B}} \Longrightarrow [P_+, L] = -[P_-, L] \in \Psi^0_{\mathcal{B}}, \tag{7.19}$$

hence $[P_+, L] \in \mathrm{DO}^0_{\mathcal{B}}$. Hence the definition (7.12) yields the identity (7.17).

The following extension of Lemma 7.2 is the central result of this section. The case $P = L^{k+1/2},\ Q = L^{m+1/2}$ implies Lemma 6.1.

Proposition 7.3. *Given $P, Q \in \mathcal{Z}(L)$, we have*

$$\partial_P Q = [P_+, Q]. \tag{7.20}$$

Proof. It suffices to treat the cases where P and Q are homogeneous, i.e.,

$$P \in \Psi_{\mathcal{B}}^{;m}, \quad Q \in \Psi_{\mathcal{B}}^{;n}. \tag{7.21}$$

Lemma 7.1 then implies that

$$P = aL^{m/2} \in \Psi_{\mathcal{B}}^{m,m}, \quad Q = bL^{n/2} \in \Psi_{\mathcal{B}}^{n,n}. \tag{7.22}$$

We can assume $m > 0$. Applying the derivation ∂_P to the identity $[Q, L] = 0$, we obtain

$$[\partial_P Q, L] + [Q, \partial_P L] = 0. \tag{7.23}$$

Now (7.17) gives

$$[Q, \partial_P L] = [Q, [P_+, L]] = [[Q, P_+], L], \tag{7.24}$$

so we have $[\partial_P Q - [P_+, Q], L] = 0$, i.e.,

$$\partial_P Q - [P_+, Q] \in \mathcal{Z}(L). \tag{7.25}$$

Now from (7.22) plus the fact that the coefficient of ξ^n in Q is constant, it follows that

$$\partial_P Q \in \Psi_{\mathcal{B}}^{n-1,n+m}, \tag{7.26}$$

and meanwhile

$$[P_+, Q] = -[P_-, Q] \in \Psi_{\mathcal{B}}^{n-2} \cap \Psi_{\mathcal{B}}^{;n+m} = \Psi_{\mathcal{B}}^{n-2,n+m}. \tag{7.27}$$

Hence

$$\partial_P Q - [P_+, Q] \in \Psi_{\mathcal{B}}^{n-1,n+m}. \tag{7.28}$$

However, given $m > 0$, it follows from Lemma 7.1 that an element of $\mathcal{Z}(L)$ satisfying (7.28) must vanish. This proves Proposition 7.3.

To obtain Lemma 6.1 from this result, note that (with our current sign convention), by (7.17) the equation $\partial_t L = [P_+, L]$ is equivalent to $\partial_t v = \partial_P v$. Now if $w \in \mathcal{B}^\alpha$ is a polynomial in v and its derivatives (x-derivatives, that is), the construction of ∂_P is just such that we obtain $\partial_t v = \partial_P v \Rightarrow \partial_t w = \partial_P w$. Hence, for $Q = L^{m+1/2}$, we obtain $\partial_t Q = \partial_P Q$.

Regarding the conserved quantities produced in §6, as a consequence of Lemma 6.1, note that

$$P \in \Psi^{\alpha}_{\mathcal{B}} \Longrightarrow \operatorname{Res} P \in \mathcal{B}^{\alpha+1}, \tag{7.29}$$

so in particular, for $m \in \mathbb{Z}^+$,

$$\operatorname{Res} L^{m+1/2} \in \mathcal{B}^{2m+2}. \tag{7.30}$$

By way of comparison, note that the integrands of $E_j(v)$ in (4.2) belong to $\mathcal{B}^{2j+4}$. We can also see from (B.13)–(B.14) that the conserved quantities $\widetilde{E}_k(bv)$ in (5.13) have integrands in $\mathcal{B}^{2k+4}$, since

$$b_k \in \mathcal{B}^{k+2}. \tag{7.31}$$

We define the space of *conserved densities*:

$$\mathcal{C}^k = \{f \in \mathcal{B}^k : \partial_P f \in \operatorname{Range} \partial, \forall\, P \in \mathcal{Z}(L)\}/\partial\mathcal{B}^{k-1}. \tag{7.32}$$

It seems likely that, for each $m \in \mathbb{Z}^+$,

$$\dim \mathcal{C}^{2m+2} = 1, \tag{7.33}$$

and that $\mathcal{C}^{2m+2}$ is spanned by $\operatorname{Res} L^{m+1/2}$, and also that it is spanned by b_{2m}. A somewhat related issue is discussed in [Wi2] and in [Fl].

A. Local existence of solutions to generalized KdV equations

Here we establish short-time existence of unique solutions to nonlinear evolution equations of the form

$$\partial_t v = Lv + g(v)v_x, \quad v(0,x) = u(x), \tag{A.1}$$

where $g \in C^\infty(\mathbb{R})$ and L is a constant-coefficient, skew-adjoint differential operator, e.g., $L = \partial_x^3$. We take $x \in S^1$, though higher-dimensional cases can be similarly treated. Our technique is adapted from a treatment of quasi-linear hyperbolic equations, as presented in Chapter 5 of [T2] and in Chapter 16 of [T3].

To begin, we let $\{J_\varepsilon : 0 < \varepsilon \le 1\}$ be a Friedrichs mollifier, and consider the evolution equations

$$\partial_t v_\varepsilon = J_\varepsilon L J_\varepsilon v_\varepsilon + J_\varepsilon g(v_\varepsilon) J_\varepsilon \partial_x v_\varepsilon, \quad v_\varepsilon(0,x) = J_\varepsilon u(x). \tag{A.2}$$

For each $\varepsilon > 0$, this is a Banach space ODE, whose local solvability follows by standard contraction mapping principle arguments. In order to show that solutions v_ε exist on an interval independent of ε and that there is a limit v solving (A.1), we need estimates, which are obtained as follows. We have

$$\begin{aligned} \frac{d}{dt}(\partial_x^k v_\varepsilon, \partial_x^k v_\varepsilon)_{L^2} &= 2(\partial_x^k \partial_t v_\varepsilon, \partial_x^k v_\varepsilon)_{L^2} \\ &= 2(\partial_x^k g(v_\varepsilon) J_\varepsilon \partial_x v_\varepsilon, J_\varepsilon \partial_x^k v_\varepsilon)_{L^2}. \end{aligned} \tag{A.3}$$

The term containing L disappears since L is skew-adjoint and commutes with ∂_x^k. To proceed, we establish the following.

Lemma A.1. *We have*

$$\big|(\partial_x^k g(v)\partial_x w, \partial_x^k w)_{L^2}\big| \le C\|w\|_{H^k}^2\|g(v)\|_{C^1} + C\|g(v)\|_{H^k}\|w\|_{C^1}\|w\|_{H^k}. \tag{A.4}$$

Proof. Write the inner product on the left as

$$([\partial_x^k, g(v)]\partial_x w, \partial_x^k w)_{L^2} + (g(v)\partial_x w_k, w_k)_{L^2}, \quad w_k = \partial_x^k w. \tag{A.5}$$

To estimate the first term in (A.5), we we use the *Moser estimate*:

$$\|[\partial_x^k, g(v)]\partial_x w\|_{L^2} \le C\|g(v)\|_{C^1}\|w\|_{H^k} + C\|g(v)\|_{H^k}\|\partial_x w\|_{L^\infty}; \tag{A.6}$$

cf. (3.6.1) of [T2], or (3.22) in Chapter 13 of [T3]. To treat the second term in (A.5), note that

$$2(g(v)\partial_x w_k, w_k)_{L^2} = -(g'(v)v_x\, w_k, w_k)_{L^2}, \tag{A.7}$$

which is bounded in absolute value by $\|g'(v)v_x\|_{L^\infty}\|w_k\|_{L^2}^2$. Since $\|g'(v)v_x\|_{L^\infty} \le \|g(v)\|_{C^1}$, we have (A.4).

Returning to (A.3), we have

$$\frac{d}{dt}\|\partial_x^k v_\varepsilon\|_{L^2}^2 \le C\|g(v_\varepsilon)\|_{C^1}\|J_\varepsilon v_\varepsilon\|_{H^k}^2 + C\|J_\varepsilon v_\varepsilon\|_{C^1}\|g(v_\varepsilon)\|_{H^k}\|J_\varepsilon v_\varepsilon\|_{H^k}. \tag{A.8}$$

Another Moser-type estimate applies:

$$\|g(v)\|_{H^k} \le C\big(\|v\|_{L^\infty}\big)\,(1 + \|v\|_{H^k}); \tag{A.9}$$

cf. (3.1.20) of [T2], or (3.30) in Chapter 13 of [T3]. Hence we obtain

$$\frac{d}{dt}\|v_\varepsilon\|_{H^k}^2 \le C\big(\|v_\varepsilon\|_{C^1}\big)\,(1 + \|v_\varepsilon\|_{H^k}^2). \tag{A.10}$$

Now $\|v\|_{C^1} \le C\|v\|_{H^2}$ for functions on S^1 (more generally, $\|v\|_{C^1} \le C\|v\|_{H^k}$ for $k > n/2 + 1$, for functions on a compact n-dimensional manifold). Hence (A.10) yields

$$\frac{d}{dt}\|v_\varepsilon\|_{H^k}^2 \le \Phi_k\big(\|v_\varepsilon\|_{H^k}^2\big), \tag{A.11}$$

for solutions to (A.2), as long as $k \ge 2$. Then Gronwall's inequality yields an estimate

$$\|v_\varepsilon(t)\|_{H^k} \le \psi_k\big(\|u\|_{H^k}\big), \quad |t| \le T = T(u), \tag{A.12}$$

for solutions to (A.2). These estimates are independent of ε, and they imply solvability of (A.2) on an interval independent of ε, as well as such estimates on this interval. Then we have

$$
\begin{aligned}
\|\partial_t v_\varepsilon\|_{H^{k-3}} &\le C\|Lv_\varepsilon\|_{H^{k-3}} + C\|g(v_\varepsilon)\partial_x J_\varepsilon v_\varepsilon\|_{H^{k-3}} \\
&\le C\|v_\varepsilon\|_{H^k} + C\big(\|v_\varepsilon\|_{L^\infty}\big)\big(1+\|v_\varepsilon\|_{H^k}\big)\,\|v_\varepsilon\|_{H^k},
\end{aligned}
\tag{A.13}
$$

(if say $L = \partial_x^3$), assuming $k \ge 2$. Consequently, if the initial data u belongs to $H^k(S^1)$, we deduce the existence of a subsequence v_{ε_ν} converging to

$$
v \in L^\infty(I, H^k(S^1)) \cap \mathrm{Lip}(I, H^{k-3}(S^1)),
\tag{A.14}
$$

with convergence in the weak* topology in these function spaces, and hence in the strong topology in $C(I, H^{k-\delta}(S^1))$, for example. It follows that such a limit satisfies (A.1), so we have a local existence result.

To treat uniqueness, suppose that w satisfies the conditions of (A.14) and also solves (A.1). Then

$$
\begin{aligned}
\frac{d}{dt}\|v-w\|_{L^2}^2 &= 2(v_t - w_t, v-w)_{L^2} \\
&= 2(g(v)v_x - g(w)w_x, v-w)_{L^2} \\
&= 2((g(v)-g(w))v_x, v-w)_{L^2} + (g(w)(v_x-w_x), v-w)_{L^2}.
\end{aligned}
\tag{A.15}
$$

We have

$$
((g(v)-g(w))v_x, v-w)_{L^2} \le K(\|v\|_{C^1} + \|w\|_{L^\infty})\|v-w\|_{L^2}^2,
\tag{A.16}
$$

and

$$
\begin{aligned}
2(g(w)\partial_x(v-w), v-w)_{L^2} &= (g(w)_x(v-w), v-w)_{L^2} \\
&\le \|g(w)\|_{C^1}\|v-w\|_{L^2}^2,
\end{aligned}
\tag{A.17}
$$

hence

$$
\frac{d}{dt}\|v-w\|_{L^2}^2 \le \Big[K\|v\|_{C^1} + K\|w\|_{L^\infty} + \|g(w)\|_{C^1}\Big]\|v-w\|_{L^2}^2.
\tag{A.18}
$$

Under the hypothesis that v and w satisfy (A.14), we can estimate $\|v\|_{C^1}$, $\|w\|_{L^\infty}$, and $\|g(w)\|_{C^1}$, and then apply Gronwall's inequality to deduce uniqueness.

We next show that if $u \in H^k(S^1)$, with $k \ge 2$, then for the solution $v(t)$ to (A.1) we have a bound on $\|v(t)\|_{H^k}$ as long as $\|v(t)\|_{C^1}$ is bounded. This result together with the local existence result established above will enable us to obtain the global existence result for KdV in §4, as a consequence of conservation laws derived in §§2–3.

Estimates to establish this persistence result are variants of those used in (A.3)–(A.10). We assume v solves (A.1), with $u \in H^k(S^1)$, and start with

$$\begin{aligned}\frac{d}{dt}(\partial_x^k J_\varepsilon v, \partial_x^k J_\varepsilon v)_{L^2} &= 2(\partial_x^k J_\varepsilon \partial_t v, \partial_x^k J_\varepsilon v)_{L^2} \\ &= 2(\partial_x^k J_\varepsilon g(v)\partial_x v, \partial_x^k J_\varepsilon v)_{L^2}.\end{aligned} \tag{A.19}$$

We write this last term as

$$2(g(v)\partial_x \partial_x^k J_\varepsilon v, \partial_x^k J_\varepsilon v)_{L^2} + 2([\partial_x^k J_\varepsilon, g(v)]\partial_x v, \partial_x^k J_\varepsilon v)_{L^2}. \tag{A.20}$$

The first term in (A.20) is

$$= (g'(v)v_x \, \partial_x^k J_\varepsilon v, \partial_x^k J_\varepsilon v) \le \|g'(v)v_x\|_{L^\infty} \|\partial_x^k J_\varepsilon v\|_{L^2}^2. \tag{A.21}$$

As for the second term in (A.20), we have a Moser-type estimate similar to (A.6):

$$\|[\partial_x^k J_\varepsilon, g(v)]\partial_x w\|_{L^2} \le C\|g(v)\|_{C^1}\|w\|_{H^k} + C\|g(v)\|_{H^k}\|\partial_x w\|_{L^\infty}, \tag{A.22}$$

which also follows from (3.6.1) of [T2]. Thus the second term in (A.20) is

$$\le C\|g(v)\|_{C^1}\|v\|_{H^k}\|\partial_x^k J_\varepsilon v\|_{L^2} + C\|g(v)\|_{H^k}\|v\|_{C^1}\|\partial_x^k J_\varepsilon v\|_{L^2}. \tag{A.23}$$

We deduce that

$$\frac{d}{dt}\|\partial_x^k J_\varepsilon v\|_{L^2}^2 \le C\|g(v)\|_{C^1}\|v\|_{H^k}^2 + C\big(\|v\|_{C^1}\big)\|v\|_{H^k}(1 + \|v\|_{H^k}), \tag{A.24}$$

and hence

$$\frac{d}{dt}\|J_\varepsilon v\|_{H^k}^2 \le C_2\big(\|v\|_{C^1}\big)(1 + \|v\|_{H^k}^2). \tag{A.25}$$

Integration yields

$$\|J_\varepsilon v(t)\|_{H^k}^2 \le \|J_\varepsilon u\|_{H^k}^2 + \int_0^t C_2\big(\|v(s)\|_{C^1}\big)\,(1 + \|v(s)\|_{H^k}^2)\,ds, \tag{A.26}$$

and letting $\varepsilon \to 0$ gives

$$\|v(t)\|_{H^k}^2 \le \|u\|_{H^k}^2 + \int_0^t C_2\big(\|v(s)\|_{C^1}\big)\,(1 + \|v(s)\|_{H^k}^2)\,ds. \tag{A.27}$$

As long as we have a bound on $\|v(s)\|_{C^1}$, a Gronwall argument applied to (A.27) gives an estimate on $\|v(t)\|_{H^k}$, which establishes our persistence result.

We summarize the results established here.

Proposition A.2. *Given $k \geq 2$ and initial data $u \in H^k(S^1)$, the equation (A.1) has a unique solution*

$$v \in C(I, H^k(S^1)), \tag{A.28}$$

for some open interval I about $t = 0$. The solution persists as long as $\|v(t)\|_{C^1}$ does not blow up.

At this point we have proven everything stated in this proposition except (A.28); what we have at this point is (A.14), which implies that $v(t)$ is a continuous function of $t \in I$ with values in $H^k(S^1)$, with the weak* topology. To conclude that it is continuous when $H^k(S^1)$ has the strong topology, it suffices to show that the function $t \mapsto \|v(t)\|_{H^k}$ is continuous. To see this, note that a time-reversal argument allows us to bound the *absolute value* of the left side of (A.25). This implies that $N_\varepsilon(t) = \|J_\varepsilon v(t)\|^2_{H^k}$ is Lipschitz continuous in t, uniformly in ε. Since $J_\varepsilon v(t) \to v(t)$ in H^k-norm for each $t \in I$, it follows that $\|u(t)\|^2_{H^k} = N_0(t) = \lim_{\varepsilon\to 0} N_\varepsilon(t)$ has the same Lipschitz continuity. The proof is complete.

Remark. Proposition A.2 can be extended, to allow k to be any real number satisfying $k > 3/2$. In fact, the only technical points involve treating (A.6) and (A.22) for nonintegral k, and in such cases (3.6.1) of [T2] (then called a Kato-Ponce estimate), still applies.

A comparable result was obtained in [Kat], using different techniques. A stronger result, valid for somewhat rougher initial data, is obtained in [KPV].

B. Trace asymptotics of $e^{t(\partial_x^2 - V)}$

Here we derive results on the asymptotic behavior (as $t \searrow 0$) of the trace of $e^{t(\partial_x^2+V)}$, acting on functions on S^1. There results are applied in §5 to produce conservation laws for solutions to KdV.

We begin with an accurate approximate solution to

$$\partial_t u = u_{xx} - Vu, \quad u(0,x) = f(x), \tag{B.1}$$

with the following Fourier integral representation:

$$\begin{aligned} u(t,x) &= \int a(t,x,\xi) e^{ix\xi} \hat{f}(\xi)\, d\xi \\ &= \frac{1}{2\pi} \iint a(t,x,\xi) e^{i(x-y)\xi} f(y)\, dy\, d\xi. \end{aligned} \tag{B.2}$$

The natural setting for (B.2) is $x \in \mathbb{R}$, but the standard method of images allows one to pass to the case $x \in S^1$. Using $\partial_x^2(fg) = f_{xx}g + 2f_xg_x + fg_{xx}$, we see that

$$\partial_x^2(ae^{ix\xi}) = (-\xi^2 a + 2i\xi a_x + a_{xx})e^{ix\xi}. \tag{B.3}$$

Hence, if $u(t,x)$ has the form (B.2),

$$\partial_t u - \partial_x^2 u + Vu = \int (a_t + \xi^2 a - 2i\xi a_x - a_{xx} + Va) e^{ix\xi} \hat{f}(\xi)\, d\xi. \tag{B.4}$$

Thus we want $a(t,x,\xi)$ to satisfy, in an appropriate sense (to be specified below),

$$\partial_t a \sim -\xi^2 a + 2i\xi \partial_x a + (\partial_x^2 - V)a. \tag{B.5}$$

The way we achieve this is to set

$$a \sim \sum_{k\ge 0} a_k(t,x,\xi), \tag{B.6}$$

and require

$$\begin{aligned} \partial_t a_0 &= -\xi^2 a_0, & a_0(0,x,\xi) = 1, \\ \partial_t a_1 &= -\xi^2 a_1 + 2i\xi \partial_x a_0, & a_1(0,x,\xi) = 0, \end{aligned} \tag{B.7}$$

and, for $k \ge 2$,

$$\partial_t a_k = -\xi^2 a_k + 2i\xi \partial_x a_{k-1} + (\partial_x^2 - V) a_{k-2}, \quad a_k(0,x,\xi) = 0. \tag{B.8}$$

Note that (B.7) yields

$$a_0(t,x,\xi) = e^{-t\xi^2}, \quad a_1(y,x,\xi) = 0. \tag{B.9}$$

It is convenient to set

$$a_k(t,x,\xi) = b_k(t,x,\xi) e^{-t\xi^2}. \tag{B.10}$$

Then we have

$$b_0 = 1, \quad b_1 = 0, \tag{B.11}$$

and, for $k \ge 2$,

$$\partial_t b_k = 2i\xi \partial_x b_{k-1} + (\partial_x^2 - V) b_{k-2}, \quad b_k(0,x,\xi) = 0, \tag{B.12}$$

i.e.,

$$b_k(t,x,\xi) = 2i\xi \int_0^t \partial_x b_{k-1}(s,x,\xi)\, ds + \int_0^t (\partial_x^2 - V) b_{k-2}(s,x,\xi)\, ds. \tag{B.13}$$

Here are a few more explicit formulas:

$$
\begin{aligned}
b_2 &= -tV, \\
b_3 &= -i\xi t^2 \partial_x V, \\
b_4 &= \frac{2}{3}\xi^2 t^3 \partial_x^2 V - \frac{t^2}{2}(\partial_x^2 V - V^2), \\
b_5 &= \frac{1}{3} i\xi^3 t^4 \partial_x^3 V - \frac{1}{3} i\xi t^3 (2\partial_x^3 V - V\partial_x V).
\end{aligned} \tag{B.14}
$$

An induction demonstrates that, for $k \geq 2$,

$$b_k(t, x, \xi) = t^{k/2}\, B_k(x, t^{1/2}\xi), \tag{B.15}$$

where $B_k(x, \zeta)$ is a polynomial in ζ, of degree $k-2$, which is *even* in ζ if k is even and *odd* in ζ if k is odd. The coefficients of these polynomials are smooth functions of x, in fact polynomials in $V(x)$ and its derivatives. We will comment more on this a little later on.

From (B.15) we can deduce such estimates as

$$
\begin{aligned}
|a_k(t, x, \xi)| &\leq C_k t^{k/2}\, e^{-t\xi^2} \\
&\leq C_k' (1 + |\xi|)^{-k}\, e^{-t\xi^2/4}.
\end{aligned} \tag{B.16}
$$

Similar estimates hold for x and t derivatives of $a_k(t, x, \xi)$. We can use these estimates to see how well

$$E_N(t)f(x) = \int A_N(t, x, \xi) e^{ix\xi} \hat{f}(\xi)\, d\xi, \quad A_N = \sum_{j \leq N} a_j \tag{B.17}$$

approximates $e^{t(\partial_x^2 + V)} f(x)$. Note that, by a calculation similar to (B.4),

$$
\begin{aligned}
(\partial_t - \partial_x^2 + V) E_N(t) f(x) &= \int R_N(t, x, \xi) e^{ix\xi} \hat{f}(\xi)\, d\xi = R_N(t) f(x), \\
R_N(t, x, \xi) &= 2i\xi \partial_x a_N + (\partial_x^2 - V)(a_{N-1} + a_N).
\end{aligned} \tag{B.18}
$$

Hence, by DuHamel's principle,

$$E_N(t) f(x) = e^{t(\partial_x^2 - V)} f(x) + \int_0^t e^{(t-s)(\partial_x^2 - V)} R_N(s) f(x)\, ds, \tag{B.19}$$

and estimates on $R_N(t, x, \xi)$ from (B.16) give a sense in which $e^{t(\partial_x^2 - V)} \sim E_N(t)$.

This allows us to justify the following trace result:

$$\operatorname{Tr} e^{t(\partial_x^2 - V)} \sim \frac{1}{2\pi} \sum_{k \geq 0} \iint a_k(t, x, \xi)\, d\xi\, dx. \tag{B.20}$$

The other ingredient used here is that, if

$$\begin{aligned} Kf(x) &= \int k(x,\xi)e^{ix\xi}\hat{f}(\xi)\,d\xi \\ &= \frac{1}{2\pi}\iint k(x,\xi)e^{i(x-y)\xi}f(y)\,dy\,d\xi, \end{aligned} \tag{B.21}$$

then

$$\operatorname{Tr} K = \frac{1}{2\pi}\iint k(x,\xi)\,d\xi\,dx. \tag{B.22}$$

To analyze the individual terms in (B.20), write

$$\begin{aligned} \iint a_k(t,x,\xi)\,d\xi\,dx &= \iint b_k(t,x,\xi)e^{-t\xi^2}\,d\xi\,dx \\ &= t^{k/2}\iint B_k(x,t^{1/2}\xi)e^{-t\xi^2}\,d\xi\,dx \\ &= t^{(k-1)/2}\iint B_k(x,\zeta)e^{-\zeta^2}\,d\zeta\,dx. \end{aligned} \tag{B.23}$$

If k is odd the integrand is an odd function of ζ, and hence the integral vanishes, so the nonzero contributions come only from even k.

In summary, we have

$$\operatorname{Tr} e^{t(\partial_x^2 - V)} \sim \sum_{k\ge 0} t^{-1/2+k}\,B_{2k}, \tag{B.24}$$

with

$$B_k = \int_{S^1} b_k(x)\,dx, \tag{B.25}$$

and

$$B_k(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} B_k(x,\zeta)e^{-\zeta^2}\,d\zeta, \tag{B.26}$$

where $B_k(x,\zeta)$ are given as in (B.15). Note also that

$$\int_{-\infty}^{\infty} \zeta^{2\ell}e^{-\zeta^2}\,d\zeta = \int_0^\infty e^{-s}s^{\ell-1/2}\,ds = \Gamma\Big(\ell+\frac{1}{2}\Big). \tag{B.27}$$

Given the inductive procedure (B.12) for constructing $b_k(t,x,\xi)$, we can say that $B_k(x)$ has the following form:

$$B_{k+2}(x) = \beta_k\partial_x^k V + Q(\partial_x^j V), \tag{B.28}$$

where Q is a sum of monomials involving $(\partial_x^{j_1}V)\cdots(\partial_x^{j_\nu}V)$ with $j_1+\cdots+j_\nu\le k-2$. Note that, by (B.13),

$$\int_{S^1} b_k(t,x,\xi)\,dx = \int_{S^1}\int_0^t V\,b_{k-2}(s,x,\xi)\,ds\,dx. \tag{B.29}$$

Via an analogue of (B.28) for $B_k(x)$ and an integration by parts, we have, for $k\ge 1$,

$$B_{2k+2} = \int_{S^1}\big[\alpha_k(\partial_x^{k-1}V)^2+\widetilde{Q}_k(\partial_x^j V)\big]\,dx, \tag{B.30}$$

where $\widetilde{Q}_k$ is a sum of monomials involving $(\partial_x^{j_1}V)\cdots(\partial_x^{j_\nu}V)$, with $j_1+\cdots+j_\nu\le 2k-4$, and each $j_\mu\le k-2$. Furthermore, it can be shown that each $\alpha_k\ne 0$.

Chapter 5: The Camassa-Holm Equation

Introduction

The Camassa-Holm equation was derived in §5 of Chapter I as the equation for geodesic flow on the diffeomorphism group of the circle S^1, whose Lie algebra $\mathrm{Vect}(S^1) \approx C^\infty(S^1)$ is equipped with the square H^1-norm. We study the equation further here. In §1 we note its Hamiltonian form on the dual of this Lie algebra, with the natural Lie-Poisson structure. In fact, this equation has a bi-Hamiltonian structure, but the second member of the Poisson pair does not arise as a frozen Poisson structure on the dual of $\mathrm{Vect}(S^1)$. Rather, one needs to go to the Virasoro algebra to see the appropriate Poisson pair arise in this fashion.

In §2 we use this bi-Hamiltonian structure and the Lenard scheme to produce a string of conserved quantities. The nature of these conserved quantities is different from those that arise for KdV. An interesting one is

$$\int \sqrt{w(t,x)}\, dx,$$

where $w = v - v_{xx}$. One implication of this conservation law is that if $w(0,x) \ge 0$, and v solves CH for t in an interval I about 0, then $w(t,x) \ge 0$ for $t \in I$. One says the solution has positive momentum density. In §3 we produce an isospectral family of self-adjoint operators associated to a solution to CH, a related string of conservation laws, and another perspective on this positivity-preserving result.

In §4 we discuss global existence of solutions to CH with positive momentum density. Our treatment follows [CoE], but is more streamlined, partly because we have available cleaner local existence results, discussed in Appendix A.

In more general cases, singularities can develop in the CH equations. We discuss some simple examples in §5; more general examples can be found in [CH] and [CoE]. Section 6 gives a brief description of some of the results on global weak solutions to CH, due to [XZ] and [Mc2]. Section 7 briefly discusses a family of Lipschitz solutions to CH, known as peakons and multi-peakons, whose central importance was first pointed out in [CH].

1. CH as a bi-Hamiltonian system

In §5 of Chapter I the Camassa-Holm equation was produced in the following form:

$$Av_t + v(Av)_x + 2v_x(Av) = 0, \tag{1.1}$$

for $v \in \mathrm{Vect}(S^1) \approx C^\infty(S^1)$, with

$$A = I - \partial_x^2. \tag{1.2}$$

By the approach taken there, this gives a curve in $\mathrm{Vect}(S^1)$ describing a geodesic flow on $\mathrm{Diff}(S^1)$, when the inner product on $\mathrm{Vect}(S^1) \approx C^\infty(S^1)$ is given by

$$B(u,v) = \int_{S^1} (uv + u'v')\, dx. \tag{1.3}$$

Let us recast this in Hamiltonian form on $\mathrm{Vect}^* \approx C^\infty(S^1)$, with the Lie-Poisson structure, defined by

$$H^0_{f_0}(w) = -\operatorname{ad}^*(df_0(w))w. \tag{1.4}$$

Recall from Chapter I that

$$\operatorname{ad}^*(u)w = \mathcal{L}_u^* w = -uw_x - 2u_x w. \tag{1.5}$$

Now we take

$$f_0(w) = -\frac{1}{2} \int_{S^1} w\, A^{-1} w\, dx, \tag{1.6}$$

so

$$df_0(w) = -A^{-1}w, \tag{1.7}$$

and the evolution equation $\partial_t w = H^0_{f_0}(w)$ becomes

$$\partial_t w = \mathcal{L}^*_{A^{-1}w} w = -(A^{-1}w)w_x - 2(A^{-1}w_x)w. \tag{1.8}$$

This is equivalent to (1.1), with

$$w = Av. \tag{1.9}$$

Note that it is natural to regard

$$A : \mathrm{Vect} \longrightarrow \mathrm{Vect}^*, \tag{1.10}$$

as indeed it arises from the inner product (1.3). The quantity w defined by (1.9) is called the *momentum density.*

There is a second Poisson structure on $C^\infty(S^1)$ with respect to which the CH equation is Hamiltonian, and which together with (1.4) forms a Poisson pair. However, it does not arise as a frozen Poisson structure on Vect^*. To relate this Poisson pair to a Lie-Poisson structure plus a frozen Poisson structure, we go to the Virasoro algebra

$$\mathrm{Vir} = \mathrm{Vect}(S^1) \oplus \mathbb{R} \approx C^\infty(S^1) \oplus \mathbb{R}, \tag{1.11}$$

as described in §9 of Chapter I, and in §1 of Chapter IV. As derived in (1.8) of Chapter IV, we have

$$\mathrm{ad}^*(u,a)(w,c) = (\mathrm{ad}^*(u)w - c\kappa(u), 0), \tag{1.12}$$

for $(u,a) \in \mathrm{Vir}$, $(w,c) \in \mathrm{Vir}^*$, with $\mathrm{ad}^*(u)w$ as in (1.5) and

$$\kappa(u) = -\partial_x^3 u. \tag{1.13}$$

As in §1 of Chapter IV, we have the Lie-Poisson structure on Vir^* defined by

$$H^0_{F_0}(w,c) = -\big(\mathrm{ad}^*(d_w F_0(w,c))w - c\kappa(d_w F_0(w,c)), 0\big). \tag{1.14}$$

Then the evolution equation $(w_t, c_t) = H^0_{F_0}(w,c)$ gives $c_t = 0$, i.e., $c \equiv c_0$, and then

$$\partial_t w = -\,\mathrm{ad}^*(d_w F_0(w,c_0)) - c_0\kappa(d_w F_0(w,c_0)). \tag{1.15}$$

If

$$F_0(w,c) = f_0(w) = -\frac{1}{2}\int_{S^1} w\, A^{-1} w\, dx, \tag{1.16}$$

then we obtain

$$\partial_t w = -(A^{-1}w)w_x - 2(A^{-1}w_x)w - c_0\, \partial_x^3 A^{-1} w, \tag{1.17}$$

and hence, for $v = A^{-1}w$ we obtain

$$Av_t + v(Av)_x + 2v_x(Av) = -c_0\partial_x^3 v, \tag{1.18}$$

which for $c_0 = 0$ is (1.1).

The second Poisson structure on Vir^* we use here is the following frozen Poisson structure:

$$\begin{aligned} H^1_{F_1}(w,c) &= -\,\mathrm{ad}^*\big(dF_1(w,c)\big)\big(\tfrac{1}{2}, -1\big) \\ &= \big(-\tfrac{1}{2}\,\mathrm{ad}^*(d_w F_1(w,c))1 + \kappa(d_w F_1(w,c)), 0\big). \end{aligned} \tag{1.19}$$

With $V = d_w F_1(w,c)$, we have $\mathrm{ad}^*(V)1 = -2\partial_x V$, so

$$H^1_{F_1}(w,c) = (\partial_x V - \partial_x^3 V, 0) = (\partial_x AV, 0), \quad V = d_w F_1(w,c). \tag{1.20}$$

Now we set

$$(1.21)\qquad F_1(w,c) = f_1(w) = -\frac{1}{2}\int_{S^1} (v^3 + v v_x^2)\,dx, \quad v = A^{-1}w.$$

We have

(1.22)

$$d_w F_1(w,c) = df_1(w) = -\frac{1}{2}A^{-1}\big[3(A^{-1}w)^2 + (A^{-1}w_x)^2 - 2\partial_x\big((A^{-1}w)(A^{-1}w_x)\big)\big].$$

The evolution equation $(w_t, c_t) = H^1_{F_1}(w,c)$ again gives $c \equiv c_0$, and then

$$(1.23)\qquad \partial_t w = -\frac{1}{2}\partial_x\big[3v^2 + v_x^2 - 2\partial_x(v v_x)\big], \quad v = A^{-1}w,$$

or

$$(1.24)\qquad (1-\partial_x^2)v_t = -3 v v_x + 2 v_x v_{xx} + v v_{xxx},$$

which is equivalent to (1.1); cf. (5.8) of Chapter I.

Note that H^0_F and H^1_F are always tangent to each space $c =$ const. in Vir^*. Thus we have a Poisson pair on such subspaces, which are linearly isomorphic to $C^\infty(S^1)$. The case c_0 gives us a Poisson pair on $C^\infty(S^1)$ and the bi-Hamiltonian structure of CH. As the calculations above show, the two Poisson structures are

$$(1.25)\qquad \begin{aligned} H^0_{f_0}(w) &= \mathcal{E}V_0 = (2w\partial_x + w_x)V_0, \\ H^1_{f_1}(w) &= \mathcal{D}V_1 = (\partial_x - \partial_x^3)V_1, \end{aligned}$$

where $V_j = df_j(w)$.

2. Conservation laws

It follows from the computations of §1 that the quantities

$$(2.1)\qquad \begin{aligned} f_0(w) &= -\frac{1}{2}\int_{S^1} w\,A^{-1}w\,dx = -\frac{1}{2}\int_{S^1}(v^2 + v_x^2)\,dx \\ f_1(w) &= -\frac{1}{2}\int_{S^1}(v^3 + v v_x^2)\,dx, \end{aligned}$$

with $v = A^{-1}w$, are constant, independent of t, for sufficiently smooth solutions of the CH equation (1.1), and they realize this equation as a bi-Hamiltonian system $w_t = H^0_{f_0}(w) = H^1_{f_1}(w)$, with

$$(2.2)\qquad \begin{aligned} H^0_f(w) &= \mathcal{E}\,df(w), & H^1_f(w) &= \mathcal{D}\,df(w), \\ \mathcal{E}V &= (2w\partial_x + w_x)V, & \mathcal{D}V &= (\partial_x - \partial_x^3)V. \end{aligned}$$

Somewhat in parallel with the calculations in §3 of Chapter IV, we will be able to implement the Lenard scheme (described abstractly in §6 of Chapter II) to produce further conservation laws, by solving inductively

$$\mathcal{E}V_j = \mathcal{D}V_{j+1}, \quad V_j = df_j(w). \tag{2.3}$$

So far we have

$$V_0 = -A^{-1}w, \quad V_1 = -\frac{1}{2}A^{-1}\big[3v^2 + v_x^2 - 2\partial_x(vv_x)\big]. \tag{2.4}$$

This time, we go "backwards," and solve for V_{-1}:

$$\mathcal{E}V_{-1} = \mathcal{D}V_0 = -\partial_x AA^{-1}w = -w_x, \tag{2.5}$$

obtaining

$$V_{-1} = -1, \tag{2.6}$$

and giving

$$f_{-1}(w) = -\int_{S^1} w\,dx \tag{2.7}$$

as a conserved quantity. The conservation of this quantity is of course also immediate from (1.23). Note also that

$$\int_{S^1} w\,dx = \int_{S^1} (v - v_{xx})\,dx = \int_{S^1} v\,dx, \tag{2.8}$$

the conservation of which is also immediate from the following variant of (1.1):

$$v_t + \partial_x\Big(\frac{1}{2}v^2 + A^{-1}\Big(v^2 + \frac{1}{2}v_x^2\Big)\Big) = 0; \tag{2.9}$$

cf. (5.9) of Chapter I.

Next we solve for V_{-2}:

$$\mathcal{E}V_{-2} = \mathcal{D}V_{-1} = 0. \tag{2.10}$$

Of course $V_{-2} = 0$ would solve (2.10), but a more interesting solution is

$$V_{-2} = w^{-1/2}, \tag{2.11}$$

giving

$$f_{-2}(w) = 2\int_{S^1} \sqrt{w}\,dx \tag{2.12}$$

as a conserved quantity.

Since f_{-2} is apparently not smooth at functions w that vanish somewhere on S^1 (and perhaps change sign), one might perhaps wonder about the validity of (2.12) as a conserved quantity. In fact, we have the following result, established in [CoE]. Set $w_+ = w$ on $E(t) = \{x : w(t,x) \geq 0\}$, 0 on the complement, and set $w_- = -w$ on $\{x : w(t,x) \leq 0\}$, 0 on the complement.

Proposition 2.1. *Given $v \in C(I, H^3(S^1))$, solving (1.1), the quantities*

$$\int_{S^1} \sqrt{w_+(t,x)}\, dx, \quad \int_{S^1} \sqrt{w_-(t,x)}\, dx, \tag{2.13}$$

are independent of $t \in I$.

Proof. We have, for $\varepsilon > 0$,

$$\frac{d}{dt} \int_{S^1} \sqrt{\varepsilon + w_+}\, dx = \frac{1}{2} \int_{E(t)} \frac{w_t}{\sqrt{\varepsilon + w_+}}\, dx. \tag{2.14}$$

The equation (1.1) implies

$$w_t = -v w_x - 2 v_x w, \tag{2.15}$$

so the right side of (2.14) is

$$\begin{aligned} &= \int_{E(t)} \frac{w v_x}{\sqrt{\varepsilon + w_+}}\, dx - \frac{1}{2} \int_{E(t)} \frac{w_x v}{\sqrt{\varepsilon + w_+}}\, dx \\ &= - \int_{E(t)} \sqrt{\varepsilon + w_+}\, v_x\, dx + \varepsilon \int_{E(t)} \frac{v_x}{\sqrt{\varepsilon + w_+}}\, dx - \frac{1}{2} \int_{E(t)} \frac{w_x v}{\sqrt{\varepsilon + w_+}}\, dx. \end{aligned} \tag{2.16}$$

Now integration by parts yields

$$\int_{E(t)} \sqrt{\varepsilon + w_+}\, v_x\, dx = -\frac{1}{2} \int_{E(t)} \frac{w_x v}{\sqrt{\varepsilon + w_+}}\, dx + \sum_J \big[v \sqrt{\varepsilon + w_+} \big]_J, \tag{2.17}$$

where J runs over all the intervals $J_\nu = (a_\nu, b_\nu)$ into which the interior of $E(t)$ decomposes, and $[f]_J = f(b_\nu) - f(a_\nu)$. Thus

$$\Big| \sum_J \big[v \sqrt{\varepsilon + w_+} \big]_J \Big| = \sqrt{\varepsilon} \Big| \sum_J (v(b_\nu) - v(a_\nu)) \Big| \le \sqrt{\varepsilon} \int_{S^1} |v_x|\, dx.$$

Hence, if $0, t \in I$, we have

$$\begin{aligned} \int_{S^1} \sqrt{\varepsilon + w_+(t,x)}\, dx = &\int_{S^1} \sqrt{\varepsilon + w_+(0,x)}\, dx \\ &+ \varepsilon \int_0^t \int_{S^1} \frac{v_x}{\sqrt{\varepsilon + w_+}}\, \chi_{w \ge 0}\, dx\, dt + O(\sqrt{\varepsilon}). \end{aligned} \tag{2.18}$$

Since

$$\frac{\varepsilon}{\sqrt{\varepsilon + w_+}} \le \sqrt{\varepsilon}, \tag{2.19}$$

we have upon taking $\varepsilon \to 0$ the result

$$\int_{S^1} \sqrt{w_+(t,x)}\, dx = \int_{S^1} \sqrt{w_+(0,x)}\, dx, \tag{2.20}$$

which handles the first term in (2.13). The other term is treated similarly.

Remark. A more natural regularity hypothesis to make on v in Proposition 2.1 is

$$v \in C(I, H^2(S^1)). \tag{2.21}$$

In fact, the extension of Proposition 2.1 to this case can be accomplished by an approximation argument. We omit the details. (We might provide them in a more polished version of these notes.)

One corollary of Lemma 3.1 is that if $v \in C(I, C^3(S^1))$ solves (1.1) and $w_0 = (1-\partial_x^2)v(0,x) \ge 0$, then $w = (1-\partial_x^2)v \ge 0$ on $I \times S^1$. This observation will play an important role in §4.

One can continue applying the Lenard scheme, producing further conservation laws. The next one is

$$f_{-3}(w) = \int_{S^1} \left(\tfrac{1}{4} w_x^2 w^{-5/2} - 2w^{-1/2}\right) dx. \tag{2.22}$$

Another method of producing conserved quantities will be discussed in §3.

3. Related isospectral family

Here we look at the operator $S : L^2(S^1) \to L^2(S^1)$, defined by

$$S = K^* M_w K, \tag{3.1}$$

where

$$K = \left(\tfrac{1}{2} - \partial_x\right)^{-1}, \quad K^* = \left(\tfrac{1}{2} + \partial_x\right)^{-1}, \quad M_w f = wf. \tag{3.2}$$

We show that if $w = Av$ and $v = v(t)$ evolves via the CH-equation (1.1), then the spectrum of $S = S(t)$ is invariant. Our treatment is adapted from [Mc2].

Note that $K : L^2(S^1) \to H^1(S^1)$, and

$$w \in H^{-1}(S^1) \Longrightarrow M_w : H^1(S^1) \to H^{-1}(S^1), \tag{3.3}$$

so

$$w \in H^{-1}(S^1) \Longrightarrow S : L^2(S^1) \to L^2(S^1). \tag{3.4}$$

Note also that S is self-adjoint (we are taking w to be real valued). We claim that S is actually Hilbert-Schmidt. In fact S^2 is a bounded, positive semi-definite operator, and we have

$$\operatorname{Tr} S^2 = \operatorname{Tr} K^* M_w K K^* M_w K = \operatorname{Tr} T^2, \tag{3.5}$$

where

$$T = R M_w, \quad R = K K^* = \Big(\frac{1}{4} - \partial_x^2\Big)^{-1}. \tag{3.6}$$

We can write

$$R f(x) = \int_{S^1} r(x-y) f(y)\, dy. \tag{3.7}$$

Then

$$\begin{aligned} \operatorname{Tr} S^2 = \operatorname{Tr} T^2 &= \iint r(x-z) w(z) r(z-x) w(x)\, dx\, dz \\ &= \iint r(x-z)^2 w(x) w(z)\, dx\, dz, \end{aligned} \tag{3.8}$$

since $r(x) = r(-x)$. Furthermore, if we consider

$$A^{-1} f(x) = (1 - \partial_x^2)^{-1} f(x) = \int_{S^1} k(x-y) f(y)\, dy, \tag{3.9}$$

the fact that $r(x)$ is a linear combination of $e^{x/2}$ and $e^{-x/2}$ on $(0, L)$ while $k(x)$ is a linear combination of e^x and e^{-x} on $(0, L)$, and both functions are invariant under $x \mapsto L - x$, allows us to deduce that

$$r(x)^2 = \alpha k(x) + \beta, \tag{3.10}$$

for some constants α and β (whose computation we leave to the reader). We obatin

$$\begin{aligned} \operatorname{Tr} S^2 &= \alpha (A^{-1} w, w) + \beta \Big(\int w\, dx \Big)^2 \\ &= \alpha \int_{S^1} (v^2 + v_x^2)\, dx + \beta \Big(\int_{S^1} v\, dx \Big)^2. \end{aligned} \tag{3.11}$$

Note that the terms on the right side of (3.11) are among the conserved quantities identified in §2.

We deduce that S is a compact, self-adjoint operator, so $L^2(S^1)$ has an orthonormal basis $\{F_j\}$ of eigenfunctions of S, with eigenvalue λ_j. The following result can be compared with Lemma 5.1 of Chapter IV.

Lemma 3.1. *We have, for each j, k,*

$$\{\lambda_j, \lambda_k\}_0 = 0, \tag{3.12}$$

where $\{f, g\}_0 = H^0_f g$ is the first Poisson structure described in (1.25).

Proof. With $\lambda_j = \lambda_j(w)$, a computation gives

$$d\lambda_j(w) = f_j^2, \quad f_j = KF_j, \tag{3.13}$$

and

$$\mathcal{E} f_k^2 = (2w\partial_x + w_x) f_k^2 = \frac{1}{2}\lambda_k(\partial_x - \partial_x^3) f_k^2. \tag{3.14}$$

Then

$$\begin{aligned} \frac{1}{\lambda_k}\{\lambda_j, \lambda_k\}_0 &= \frac{1}{2}\int f_j^2(\partial_x - \partial_x^3) f_k^2\, dx \\ &= -\frac{1}{2}\int f_k^2(\partial_x - \partial_x^3) f_j^2\, dx \\ &= -\frac{1}{\lambda_j}\{\lambda_k, \lambda_j\}_0, \end{aligned} \tag{3.15}$$

which gives (3.12) whenever $\lambda_j \neq \lambda_k$.

This result suffers from the same limitations as Lemma 5.1 of Chapter IV, due to the possibility of eigenvalues coalescing. (Actually, in the setting of [Mc2], with S^1 replaced by $\mathbb{R}$, it is the case that S always has simple spectrum, but in the current setting this need not hold.) A variant of the remedy used there is also effective here. Namely, set

$$\mathcal{F} = \{\psi \in C^\infty(\mathbb{R}) : \psi(s) = 0 \text{ near } s = 0\}, \tag{3.16}$$

and set

$$f_\psi(w) = \operatorname{Tr} \psi(K^* M_w K). \tag{3.17}$$

Then we have

$$\varphi, \psi \in \mathcal{F} \Longrightarrow \{f_\varphi, f_\psi\}_0 = 0. \tag{3.18}$$

Then a limiting argument gives

$$\{\tau, f_\varphi\}_0 = 0, \quad \forall\, \varphi \in \mathcal{F}, \tag{3.19}$$

where

$$\tau(w) = \operatorname{Tr} S^2 = \operatorname{Tr}(K^* M_w K)^2. \tag{3.20}$$

By (3.11) we have

$$\tau(w) = -2\alpha f_0(w) + \beta f_{-1}(w)^2, \tag{3.21}$$

where $f_0(w)$ is given by (2.1) and $f_{-1}(w)$ by (2.7). Note that, since $df_{-1}(w) = -1$,

$$\{f_{-1}, \lambda_k\}_0 = -\frac{1}{2}\lambda_k \int_{S^1} (\partial_x - \partial_x^3) f_k^2 \, dx = 0. \tag{3.22}$$

Hence from (3.19) and (3.21) we deduce

$$\{f_0, f_\psi\}_0 = 0, \quad \forall\, \psi \in \mathcal{F}, \tag{3.23}$$

or, loosely stated,

$$\{f_0, \lambda_j\}_0 = 0, \quad \forall\, j. \tag{3.24}$$

Thus the quantities λ_j are all constants of motion for a solution to the CH equation, i.e., $S = K^* M_w K$ is isospectral.

Remark. Note that

$$(Sf, f)_{L^2} = \int_{S^1} \big(Kf(x)\big)^2 \, w(x) \, dx. \tag{3.25}$$

In particular,

$$S \geq 0 \Longleftrightarrow w \geq 0 \quad \text{on} \quad S^1. \tag{3.26}$$

Thus the fact that S is isospectral, when $w(t) = (1 - \partial_x^2) v(t)$ and $v(t)$ solves CH, implies that

$$w_0 \geq 0 \Longrightarrow w(t) \geq 0, \tag{3.27}$$

by a different method than that used in §2 (the conservation of (2.13)) to reach such a conclusion.

4. Global existence of solutions with positive momentum density

In this section we establish some results on global solutions to the initial value problem

$$Av_t + v(Av)_x + 2v_x(Av) = 0, \quad v(0,x) = v_0(x), \tag{4.1}$$

under the hypothesis that v_0 satisfies the condition

$$w_0(x) = (1 - \partial_x^2) v_0(x) \geq 0 \quad \text{on} \quad S^1. \tag{4.2}$$

We start with the following result of [CoE].

Proposition 4.1. *Assume $v_0 \in H^k(S^1)$, $k \geq 2$, and that (4.2) holds. Then the equation (4.1) has a unique global solution*

$$v \in C(\mathbb{R}, H^k(S^1)) \cap C^1(\mathbb{R}, H^{k-1}(S^1)). \tag{4.3}$$

Proof. From Proposition A.1 we have a solution v of the type (4.3) on some interval I about $t = 0$, and we can guarantee $I = \mathbb{R}$ if we can show that $\|v(t)\|_{C^1}$ does not blow up. This follows from conservation laws established in §2. In particular the conservation of $\int \sqrt{w_\pm}\, dx$ shows that if (4.2) holds then $w(t,x) = (1-\partial_x^2)v(t,x) \geq 0$ for all $t \in I$. We also have conservation of

$$\int_{S^1} v(t,x)\, dx = \int_{S^1} [v(t,x) - \partial_x^2 v(t,x)]\, dx. \tag{4.4}$$

But if the integrand on the right is ≥ 0, this yields

$$\|(1-\partial_x^2)v(t)\|_{L^1(S^1)} \leq K, \tag{4.5}$$

which immediately implies

$$\|v(t)\|_{C^1(S^1)} \leq K_2, \tag{4.6}$$

and proves the proposition.

The following result on more singular global solutions was also established in [CoE].

Proposition 4.2. *Assume $(1-\partial_x^2)v_0$ is a positive measure on S^1. Then (4.1) has a global solution, satisfying, for all $\delta > 0$,*

$$v \in C(\mathbb{R}, H^{3/2-\delta}(S^1)) \cap Lip(\mathbb{R} \times S^1). \tag{4.7}$$

Proof. Using a mollifier on S^1, we can take $v_{0,k} \in C^\infty(S^1)$ such that

$$w_{0,k} = (1-\partial_x^2)v_{0,k} \geq 0, \quad \|w_{0,k}\|_{L^1} \leq K, \tag{4.8}$$

and

$$v_{0,k} \to v_0 \text{ in } H^{3/2-\delta}(S^1), \quad \forall\, \delta > 0. \tag{4.9}$$

We also have convergence in various other function spaces, which we need not specify here. By Proposition 4.1 there is a unique $v_k \in C^\infty(\mathbb{R} \times S^1)$, satisfying

$$A\partial_t v_k + v_k(Av_k)_x + 2(\partial_x v_k)(Av_k) = 0, \quad v_k(0,x) = v_{0,k}(x), \tag{4.10}$$

or equivalently

$$\partial_t v_k + v_k \partial_x v_k + A^{-1}\partial_x\Big(v_k^2 + \frac{1}{2}(\partial_x v_k)^2\Big) = 0, \quad v_k(0,x) = v_{0,k}(x). \tag{4.11}$$

We have

$$(1-\partial_x^2)v_k(t,x) \geq 0, \quad \forall\ t, x, \tag{4.12}$$

and

$$\int\limits_{S^1} (1-\partial_x^2)v_k(t,x)\,dx = \int\limits_{S^1} w_{0,k}(x)\,dx, \tag{4.13}$$

for all $t \in \mathbb{R}$. Hence

$$\|(1-\partial_x^2)v_k(t)\|_{L^1(S^1)} \leq K < \infty, \quad \text{so } \|v_k(t)\|_{C^1(S^1)} \leq K' < \infty. \tag{4.14}$$

Furthermore, by the Sobolev imbedding theorem, the first result in (4.14) implies that

$$\|v_k(t)\|_{H^{3/2-\delta}(S^1)} \leq K_\delta < \infty, \quad \forall\ \delta > 0. \tag{4.15}$$

Thus we have $\{v_k\}$ uniformly bounded in $L^\infty(\mathbb{R}, H^{3/2-\delta}(S^1))$, for each $\delta > 0$. Also, since $\{\partial_x v_k\}$ is bounded in $L^\infty(\mathbb{R}\times S^1)$, we have (by (4.11)) $\{\partial_t v_k\}$ uniformly bounded in $L^\infty(\mathbb{R}\times S^1)$, so $\{v_k\}$ is uniformly bounded in $\mathrm{Lip}(\mathbb{R}\times S^1)$. It follows that $\{v_k\}$ has a subsequence v_{k_ν}, converging to

$$v \in L^\infty(\mathbb{R}, H^{3/2-\delta}(S^1)) \cap \mathrm{Lip}(\mathbb{R}\times S^1), \tag{4.16}$$

in the weak* topology, hence strongly in $C([-T,T], H^{3/2-2\delta}(S^1))$, for all $T < \infty$, $\delta > 0$. It follows that v satisfies

$$\partial_t v + v\partial_x v + A^{-1}\partial_x\Big(v^2 + \frac{1}{2}v_x^2\Big) = 0, \quad v(0,x) = v_0(x), \tag{4.17}$$

i.e., v solves (4.1).

Remark 1. From (4.12)–(4.14) we see that the solution v produced in Proposition 4.2 also has the properties

$$(1-\partial_x^2)v(t) \in \mathcal{M}^+(S^1), \quad \|(1-\partial_x^2)v(t)\|_{\mathrm{TV}} \leq K < \infty, \quad \forall\ t \in \mathbb{R}, \tag{4.18}$$

where $\mathcal{M}^+(S^1)$ denotes the space of positive Borel measures on S^1 and $\|\cdot\|_{\mathrm{TV}}$ the total variation norm.

Remark 2. Propositions 4.1–4.2 have obvious analogues with "w_0 positive" replaced by "w_0 negative."

Remark 3. The proof of Proposition 4.1 used a conservation law that was stated in §3 to hold for $v_0 \in H^2(S^1)$ but demonstrated there only for $v_0 \in H^3(S^1)$. (In fact Proposition 4.1 was demonstrated in [CoE] only for $k \geq 3$.) To complete the proof of Proposition 4.1 in the case $k = 2$, we can proceed as follows. Given $v_0 \in H^2(S^1)$ such that (4.2) holds, let $v_{0,k} \in C^\infty(S^1)$ approximate v_0 as in (4.8), with (4.9) strengthened to $v_{0,k} \to v_0$ in $H^2(S^1)$. (Sorry for the double use of the letter k.) Then Proposition 4.1 applies to $v_{0,k}$, yielding $v_k \in C^\infty(\mathbb{R} \times S^1)$, satisfying (4.11), with limit v satisfying (4.16). But also $v(0, x) = v_0$, and using (A.11), with $k = 2$, we can deduce that $v \in C(I, H^2(S^1))$ for an interval I about $t = 0$. Since, by (4.16), $\|v(t)\|_{C^1}$ does not blow up, this solution persists, satisfying (4.3) with $k = 2$.

Uniqueness of solutions to (4.1) produced by Proposition 4.2 was established in [CoM]. Here we prove the following uniqueness result.

Proposition 4.3. *In the setting of Proposition 4.2, the solution to (4.1) produced by mollifying the initial data as in (4.8)–(4.9) is unique.*

Proof. Let u be another solution to (4.1), satisfying (4.7) and (4.18). We will estimate $w(t) = u(t) - v(t)$. To begin, we temporarily assume $u(0)$ and $v(0)$ are smooth, and estimate

$$
\begin{aligned}
&\frac{d}{dt} \int_{S^1} \big[|w(t)| + |w_x(t)|\big]\, dx \\
&\quad = \int_{S^1} \big[w_t(t,x)\ \operatorname{sgn} w(t,x) + w_{tx}(t,x)\ \operatorname{sgn} w_x(t,x)\big]\, dx.
\end{aligned} \tag{4.19}
$$

We concentrate on estimating the integral of the second term on the right side of (4.19), since the estimate on the first term is similar but simpler.

By (4.17) we have

$$
w_{tx} = -\partial_x\big(uu_x - vv_x + KF(u,u_x) - KF(v,v_x)\big), \tag{4.20}
$$

with

$$
K = \partial_x(1 - \partial_x^2)^{-1}, \quad F(u,u_x) = u^2 + \frac{1}{2}u_x^2. \tag{4.21}
$$

Note that $\partial_x K = \partial_x^2 (1-\partial_x^2)^{-1} = -1 + (1-\partial_x^2)^{-1}$, so

$$\begin{aligned}\int_{S^1} \big|\partial_x K(F(u,u_x) - F(v,v_x))\big|\,dx \\ &\le C \int_{S^1} |F(u,u_x) - F(v,v_x)|\,dx \\ &\le C \int_{S^1} \Big[|u+v|\cdot|u-v| + \frac{1}{2}|u_x+v_x|\cdot|u_x-v_x|\Big]\,dx \\ &\le C' \int_{S^1} \big[|w| + |w_x|\big]\,dx,\end{aligned} \tag{4.22}$$

where we have used (4.7) to bound $|u+v|$ and $|u_x+v_x|$. It remains to estimate

$$-\int_{S^1} (\operatorname{sgn} w_x)\, \partial_x(uu_x - vv_x)\,dx. \tag{4.23}$$

We write

$$\begin{aligned}\partial_x(uu_x - vv_x) &= \partial_x(uu_x - vu_x + vu_x - vv_x) \\ &= \partial_x(wu_x + vw_x) \\ &= w_x u_x + wu_{xx} + v_x w_x + vw_{xx},\end{aligned} \tag{4.24}$$

and separately estimate the resulting four terms. First,

$$\int_{S^1} |w_x u_x|\,dx \le C \int_{S^1} |w_x|\,dx, \tag{4.25}$$

since $\|u_x\|_{L^\infty} \le C$. Next,

$$\begin{aligned}\int_{S^1} |wu_{xx}|\,dx &\le C_1 \sup_x |w| \\ &\le C_2 \int_{S^1} \big[|w| + |w_x|\big]\,dx,\end{aligned} \tag{4.26}$$

the first estimate by (4.18), with u in place of v. Next,

$$\int_{S^1} |v_x w_x|\,dx \le C \int_{S^1} |w_x|\,dx, \tag{4.27}$$

since $\|v_x\|_{L^\infty} \le C$. Finally,

$$
\begin{aligned}
\Bigl|\int\limits_{S^1} (\operatorname{sgn} w_x) v w_{xx}\,dx\Bigr| &= \Bigl|\int\limits_{S^1} v\,\partial_x |w_x|\,dx\Bigr| \\
&= \Bigl|\int\limits_{S^1} v_x\,|w_x|\,dx\Bigr| \\
&\le C\int\limits_{S^1} |w_x|\,dx.
\end{aligned}
\tag{4.28}
$$

These estimates and their counterparts for the first term on the right side of (4.19) yield

$$
\frac{d}{dt}\int\limits_{S^1} \bigl[|w| + |w_x|\bigr]\,dx \le C\int\limits_{S^1} \bigl[|w| + |w_x|\bigr]\,dx. \tag{4.29}
$$

It follows from Gronwall's inequality that

$$
\|u(t) - v(t)\|_{H^{1,1}(S^1)} \le e^{C|t|}\,\|u(0) - v(0)\|_{H^{1,1}(S^1)}, \tag{4.30}
$$

for $t \ge 0$, with a similar estimate for $t \le 0$. Here

$$
C = C_0(\|(1-\partial_x^2)u(0)\|_{\mathrm{TV}} + \|(1-\partial_x^2)v(0)\|_{\mathrm{TV}}). \tag{4.31}
$$

So far we have (4.30)–(4.31) when $u(0)$ and $v(0)$ are smooth. Now we take general $u(0)$ and $v(0)$ satisfying the conditions of Proposition 4.2, and approximate by smooth $u_{0,k},\ v_{0,k}$, as in (4.8)–(4.9). Then as in (4.30) we get

$$
\|u_k(t) - v_k(t)\|_{H^{1,1}(S^1)} \le e^{C|t|}\|u_{0,k} - v_{0,k}\|_{H^{1,1}(S^1)}, \tag{4.32}
$$

and the analogue of (4.31) holds, so C can be taken independent of k in (4.32). Passing to the limit $k \to \infty$, we obtain (4.30) for the general solutions produced by Proposition 4.2. This proves uniqueness.

Remark 4. A stronger uniqueness result is proven in [CoM]. Namely, solutions to (4.1) satisfying (4.7) and (4.18) are unique. The proof involves estimating

$$
\frac{d}{dt}\int\limits_{S^1} \bigl[|\rho_\varepsilon * w|^2 + |\rho_\varepsilon * w_x|^2\bigr]\,dx, \tag{4.33}
$$

where ρ_ε is a mollifier. The techniques are parallel to but more elaborate than those used above.

5. Breakdown of smooth solutions

In §4 we had some global existence results for solutions to (1.1) with positive momentum density, including global existence of smooth solutions, given smooth initial data. It turns out that the solutions can develop singularities in cases where this hypothesis fails. One simple family of examples arises as follows.

Proposition 5.1. *Assume* $v_0 \in C^\infty(S^1)$, $v_0(-x) = -v_0(x)$, *and* $\partial_x v_0(0) < 0$. *Then there exists* $T \in (0,\infty)$ *such that* $\partial_x v(t,0) \to -\infty$ *as* $t \nearrow T$.

Proof. Let us rewrite (1.1) as

$$v_t = -\partial_x\Bigl(\frac{1}{2}v^2 + A^{-1}\Bigl(v^2 + \frac{1}{2}v_x^2\Bigr)\Bigr), \quad v(0,x) = v_0(x), \tag{5.1}$$

where as usual $A = 1 - \partial_x^2$. Then we have

$$\begin{aligned} \partial_t(v_x) &= -\partial_x^2\Bigl(\frac{1}{2}v^2 + A^{-1}\Bigl(v^2 + \frac{1}{2}v_x^2\Bigr)\Bigr) \\ &= -\frac{1}{2}v_x^2 + v^2 - vv_x - A^{-1}\Bigl(v^2 + \frac{1}{2}v_x^2\Bigr). \end{aligned} \tag{5.2}$$

The hypothesis that $v_0(x)$ is odd and that (5.1) holds on $I \times S^1$ implies that $v(t,x)$ is odd in x for all $t \in I$. In particular $v(t,0) \equiv 0$. Hence $s(t) = v_x(t,0)$ satisfies

$$s'(t) \le -\frac{1}{2}s(t)^2, \tag{5.3}$$

since $A^{-1}(v^2 + v_x^2/2) \ge 0$. Now clearly (5.3) plus $s(0) < 0$ forces $s(t) \to -\infty$ for finite $t > 0$.

In [CH] a more general mechanism for breakdown is discussed, which involves an inflection point in $v_0(x)$ to the right of its maximum. (Here $x \in \mathbb{R}$.)

The paper [CoE] discusses some other conditions that guarantee breakdown of smooth solutions to (5.1). Such breakdown is demonstrated under the hypothesis that $v_0 \ne 0$ and

$$\int_{S^1} v_0(x)\, dx = 0. \tag{5.4}$$

Also breakdown is demonstrated when $v_0 \ne 0$ and

$$\int_{S^1} (v_0^3 + v_0(\partial_x v_0)^2)\, dx = 0. \tag{5.5}$$

In §6 we discuss some progress in treating weak solutions in cases where such breakdown is possible.

6. Remarks on weak solutions

The fact that $\int(v^2+v_x^2)\,dx$ is conserved for solutions to (1.1) makes it natural to think that this equation has a global weak solution in $L^\infty(\mathbb{R},H^1(S^1))$, given initial data $v_0\in H^1(S^1)$. One attempt to get this might involve the following variant of the method used in Appendix A. We start with the approximating equation

$$\partial_t v_\varepsilon = -J_\varepsilon\big((J_\varepsilon v_\varepsilon)(J_\varepsilon v_\varepsilon)_x\big) - J_\varepsilon\partial_x A^{-1}\Big((J_\varepsilon v_\varepsilon)^2 + \frac{1}{2}(\partial_x J_\varepsilon v_\varepsilon)^2\Big). \tag{6.1}$$

We estimate the H^1 norm:

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}(Av_\varepsilon,v_\varepsilon) &= (A\partial_t v_\varepsilon, v_\varepsilon)\\ &= -\big((J_\varepsilon v_\varepsilon)(J_\varepsilon v_\varepsilon)_x, AJ_\varepsilon v_\varepsilon\big) + \big((J_\varepsilon v_\varepsilon)^2+\tfrac{1}{2}(\partial_x J_\varepsilon v_\varepsilon)^2, \partial_x J_\varepsilon v_\varepsilon\big)\\ &= 0.\end{aligned} \tag{6.2}$$

Thus $\{v_\varepsilon : 0<\varepsilon\le 1\}$ is uniformly bounded in $L^\infty(\mathbb{R},H^1(S^1))$. We then have $\partial_t v_\varepsilon$ bounded in $L^\infty(\mathbb{R},L^2(S^1))$. Hence there is a sequence $\varepsilon_\nu\to 0$ with v_{ε_ν} converging to

$$v\in L^\infty(\mathbb{R},H^1(S^1))\cap \mathrm{Lip}(\mathbb{R},L^2(S^1)), \tag{6.3}$$

in the weak* topology. The left side of (6.1) converges to $\partial_t v$ weakly. We also have no problem with convergence of two of the three terms on the right side of (6.1), to $-vv_x$ and $-\partial_x A^{-1}(v^2)$, respectively. Furthermore, we have convergence

$$(\partial_x J_{\varepsilon_\nu} v_{\varepsilon_\nu})^2 \longrightarrow w\in L^\infty(\mathbb{R},\mathcal{M}^+(S^1)), \tag{6.4}$$

where $\mathcal{M}^+(S^1)$ denotes the space of positive measures on S^1. However, it is not clear that $w=v_x^2$.

Another approach is taken in [XZ], using

$$v_t = -vv_x - A^{-1}\partial_x\Big(v^2+\frac{1}{2}v_x^2\Big) + \varepsilon v_{xx}. \tag{6.5}$$

Using the theory of Young measures, the authors establish existence of weak solutions to CH, with initial data in $H^1(S^1)$, as a limit of solutions to (6.5). The issue of uniqueness is still open.

We also mention [Mc2], which "integrates" the CH equation in terms of the flows generated by $H^0_{\lambda_j}$, with $\lambda_j(w)$ as in §3. This paper has a discussion of how singularities can arise in the solution to CH.

7. Peakons and multi-peakons

The CH equation (1.1) has a family of solutions on $\mathbb{R}\times\mathbb{R}$,

$$v(t,x)=ce^{-|x-ct|}, \tag{7.1}$$

called peakons. There are also peakon solutions on $\mathbb{R}\times S^1$, with $S^1=\mathbb{R}/(L\mathbb{Z})$, such as

$$v(t,x)=\sum_{k=-\infty}^{\infty} e^{-|x-t+k|}, \tag{7.2}$$

in case $L=1$. There are also "multi-peakon" solutions of (1.1) on $\mathbb{R}\times\mathbb{R}$, of the form

$$v(t,x)=\sum_{j=1}^{N}\eta_j(t)\,e^{-|x-y_j(t)|}. \tag{7.3}$$

Plugging (7.3) into (1.1) yields a system of ODE for $(y,\eta)=(y_1,\dots,y_N,\eta_1,\dots,\eta_N)$ of Hamiltonian type:

$$(\dot y,\dot\eta)=H_F(y,\eta), \tag{7.4}$$

with

$$F(y,\eta)=\frac{1}{2}\sum_{j,k=1}^{N}\eta_j\eta_k\,e^{-|y_j-y_k|}. \tag{7.5}$$

This is precisely the equation for geodesic motion on $\mathbb{R}^N$, equipped with the metric tensor (g_{jk}) whose inverse matrix (g^{jk}) is given by

$$g^{jk}(y)=e^{-|y_j-y_k|}. \tag{7.6}$$

This is an integrable system of ODE. See [CH] and [ACHM] for a study.

A. Local existence of solutions to equations of CH-type

Here we establish local existence of solutions to the equation (1.1) with initial data in $H^k(S^1),\ k\ge 2$. It is convenient to rewrite (1.1) as

$$v_t+vv_x+A^{-1}\partial_x\Bigl(v^2+\frac{1}{2}v_x^2\Bigr)=0,\quad A=1-\partial_x^2. \tag{A.1}$$

Compare (5.9) of Chapter I. We consider more generally the initial value problem

$$v_t=g(v)v_x+KF(v,v_x),\quad v(0,x)=u(x), \tag{A.2}$$

given $K=K(D)\in OPS^{-1}(S^1)$. We will establish the following result.

Proposition A.1. *Given $k \geq 2$ and initial data $u \in H^k(S^1)$, the equation (A.2) has a unique solution*

$$v \in C(I, H^k(S^1)) \cap C^1(I, H^{k-1}(S^1)), \tag{A.3}$$

for some interval I about $t = 0$. The solution persists as long as $\|v(t)\|_{C^1}$ does not blow up.

The proof of this result is similar to that presented in Appendix A of Chapter IV for the class of PDE considered there (related to KdV). Since there are some significant differences, we go over the proof in this case.

As in Chapter IV, we let $\{J_\varepsilon : 0 < \varepsilon \leq 1\}$ be a Friedrichs mollifier. Now we consider the evolution equations

$$\partial_t v_\varepsilon = J_\varepsilon g(v_\varepsilon) J_\varepsilon \partial_x v_\varepsilon + J_\varepsilon K J_\varepsilon F(v_\varepsilon, \partial_x v_\varepsilon), \quad v_\varepsilon(0, x) = J_\varepsilon u(x). \tag{A.4}$$

For each $\varepsilon > 0$, this is a Banach space ODE, whose local solvability is standard. In order to show that solutions v_ε exist on an interval independent of ε and that there is a limit v solving (A.2), we need estimates. To start, we have

(A.5)
$$\begin{aligned} \frac{d}{dt}(\partial_x^k v_\varepsilon, \partial_x^k v_\varepsilon)_{L^2} &= 2(\partial_x^k \partial_t v_\varepsilon, \partial_x^k v_\varepsilon) \\ &= 2(\partial_x^k g(v_\varepsilon) \partial_x J_\varepsilon v_\varepsilon, \partial_x^k J_\varepsilon v_\varepsilon) + 2(\partial_x^k K J_\varepsilon F(v_\varepsilon, \partial_x v_\varepsilon), \partial_x^k J_\varepsilon v_\varepsilon). \end{aligned}$$

Lemma A.1 of Chapter IV applies to the first term in the last sum, bounding its absolute value by

$$C \|J_\varepsilon v_\varepsilon\|_{H^k}^2 \|g(v_\varepsilon)\|_{C^1} + C \|g(v_\varepsilon)\|_{H^k} \|J_\varepsilon v_\varepsilon\|_{C^1} \|J_\varepsilon v_\varepsilon\|_{H^k}, \tag{A.6}$$

and, as in (A.9) of Chapter IV, we have a Moser estimate

$$\|g(v_\varepsilon)\|_{H^k} \leq C(\|v\|_{L^\infty})\,(1 + \|v_\varepsilon\|_{H^k}). \tag{A.7}$$

As for the last term in (A.5), since $K \in OPS^{-1}(S^1)$, we easily bound its absolute value by

$$C \|F(v_\varepsilon, \partial_x v_\varepsilon)\|_{H^{k-1}} \|J_\varepsilon v_\varepsilon\|_{H^k} \leq C(\|v_\varepsilon\|_{C^1})\,(1 + \|v_\varepsilon\|_{H^k}) \|J_\varepsilon v_\varepsilon\|_{H^k}, \tag{A.8}$$

the last inequality by a Moser estimate parallel to (A.7).

From (A.5)–(A.8) we have

$$\frac{d}{dt}\|v_\varepsilon\|_{H^k}^2 \leq C(\|v_\varepsilon\|_{C^1})\,(1 + \|v_\varepsilon\|_{H^k}^2), \tag{A.9}$$

and hence

$$\frac{d}{dt}\|v_\varepsilon\|_{H^k}^2 \leq \Phi_k\big(\|v_\varepsilon\|_{H^k}^2\big), \tag{A.10}$$

for solutions to (A.4), as long as $k \geq 2$. Then Gronwall's inequality yields an estimate

$$\|v_\varepsilon(t)\|_{H^k} \leq \psi_k\big(\|u\|_{H^k}\big), \quad |t| \leq T = T(u), \tag{A.11}$$

for solutions to (A.4). These estimates are independent of ε, and they imply solvability of (A.4) on an interval independent of ε, as well as such estimates on this interval. Then we have

$$\begin{aligned} \|\partial_t v_\varepsilon\|_{H^{k-1}} &\leq C\|g(v_\varepsilon)\partial_x J_\varepsilon v_\varepsilon\|_{H^{k-1}} + C\|F(v_\varepsilon, \partial_x v_\varepsilon)\|_{H^{k-2}} \\ &\leq C(\|v_\varepsilon\|_{C^1})\,(1 + \|v_\varepsilon\|^2_{H^k}), \end{aligned} \tag{A.12}$$

assuming $k \geq 2$. Consequently, if the initial data u belongs to $H^k(S^1)$, we deduce the existence of a subsequence v_{ε_ν} converging to

$$v \in L^\infty(I, H^k(S^1)) \cap \operatorname{Lip}(I, H^{k-1}(S^1)), \tag{A.13}$$

with convergence in the weak* topology in these function spaces, and hence in the strong topology in $C(I, H^{k-\delta}(S^1))$. It follows that such a limit satisfies (A.2). This gives the local existence result asserted in Proposition A.1, except for the additional regularity stated in (A.3), which we will address below.

To establish uniqueness, we begin as in (A.15) of Chapter IV. If w satisfies the conditions of (A.13) and also solves (A.2), then

$$\begin{aligned} \frac{d}{dt}\|v - w\|^2_{L^2} &= 2(v_t - w_t, v - w)_{L^2} \\ &= 2(g(v)v_x - g(w)w_x, v - w) + 2(KF(v, v_x) - KF(w, w_x), v - w). \end{aligned} \tag{A.14}$$

As in (A.16)–(A.17) of Chapter IV, we bound the first term in the last sum by

$$\big[C\|v\|_{C^1} + C\|w\|_{L^\infty} + \|g(w)\|_{C^1}\big]\,\|v - w\|^2_{L^2}. \tag{A.15}$$

However, the last term in (A.14) is more recalcitrant. We can bound it by

$$C(\|v\|_{C^1}, \|w\|_{C^1})\,\big(\|v - w\|^2_{L^2} + \|v - w\|_{L^2}\|v_x - w_x\|_{L^2}\big), \tag{A.16}$$

but the appearance of the quantity $\|v_x - w_x\|_{L^2}$ obliges us to do more work.

We therefore complement (A.14) with

$$\begin{aligned} \frac{d}{dt}\|v_x - w_x\|^2_{L^2} &= 2(v_{tx} - w_{tx}, v_x - w_x) \\ &= 2\big(\partial_x(g(v)v_x) - \partial_x(g(w)w_x), v_x - w_x\big) \\ &\quad + 2\big(\partial_x KF(v, v_x) - \partial_x KF(w, w_x), v_x - w_x\big). \end{aligned} \tag{A.17}$$

To estimate the first term in the last sum, we write

$$\begin{aligned}(A.18)\qquad \big(\partial_x(g(v)V)-\partial_x(g(w)W),V-W\big) &= \big(\partial_x\big(g(v)(V-W)\big),V-W\big)\\ &\quad+\big(\partial_x\big((g(v)-g(w))W\big),V-W\big).\end{aligned}$$

An integration by parts applied to the first term on the right side of (A.18) bounds its absolute value by

$$(A.19)\qquad \frac{1}{2}\|g'(v)\|_{L^\infty}\|v\|_{C^1}\|V-W\|^2_{L^2}.$$

Meanwhile the last term in (A.18) is bounded in absolute value by

$$(A.20)\qquad \big(\|\partial_x g(v)-\partial_x g(w)\|_{L^2}\|W\|_{L^\infty}+\|g(v)-g(w)\|_{L^\infty}\|\partial_x W\|_{L^2}\big)\|V-W\|_{L^2}.$$

Setting $V=v_x$ and $W=w_x$, we bound the absolute value of the first term on the right side of (A.17) by

$$(A.21)\qquad C(\|v\|_{C^1},\|w\|_{H^2})\,\|v-w\|^2_{H^1}.$$

We turn to the last term in (A.17). Since $\partial_x K\in OPS^0(S^1)$, it is bounded in absolute value by

$$(A.22)\qquad C\|F(v,v_x)-F(w,w_x)\|_{L^2}\|v_x-w_x\|_{L^2}\le C(\|v\|_{C^1},\|w\|_{C^1})\,\|v-w\|^2_{H^1}.$$

Putting together (A.14)–(A.22), we have

$$(A.23)\qquad \frac{d}{dt}\|v-w\|^2_{H^1}\le C(\|v\|_{C^1},\|w\|_{H^2})\,\|v-w\|^2_{H^1},$$

which yields the uniqueness result of Proposition A.1.

As for the persistence result stated in Proposition A.1, and also the result (A.3) on the regularity of the solution, which improves (A.13), the proofs of these facts use the same sorts of estimates that appear in (A.19)–(A.27) of Chapter IV. Similar arguments are also given on pp. 1084–1085 of [Mis2].

References

[ACHM] M. Ablowitz, R. Camassa, D. Holm, and J. Marsden, The geometry of peaked solitons and billiard solutions of a class of integrable PDE's, Lett. Math. Phys. 32 (1994), 137–151.

[Ad] M. Adler, On a trace functional for formal pseudo-differential operators and the symplectic structure for Korteweg-deVries type equations, Invent. Math. 50 (1979), 219–248.

[Ar] V. Arnold, Sur la géometrie différentielle des groupes de Lie de dimension infinie et ses applications à l'hydrodynamique des fluides parfaits, Ann. Inst. Fourier 16 (1966), 319–361.

[AK] V. Arnold and B. Khesin, Topological methods in hydrodynamics, Springer-Verlag, New York, 1998.

[Aud] M. Audin, Spinning Tops, Cambridge Univ. Press, Cambridge, 1996.

[CH] R. Camassa and D. Holm, An integrable shallow water equation with peaked solitons, Phys. Rev. Lett. 71 (1993), 1661–1664.

[CHH] R. Camassa, D. Holm, and M. Hyman, A new integrable shallow-water equation, Adv. Appl. Mech. 31 (1994), 1–33.

[Che] I. Cherednik, Basic Methods of Soliton Theory, World Scientific, 1996.

[Che2] I. Cherednik, Loop Groups in Soliton Theory, Preprint, 2003.

[CoE] A. Constantin and J. Escher, Well-posedness, global existence, and blowup phenomena for a periodic quasi-linear hyperbolic equation, CPAM 51 (1998), 475–504.

[CoM] A. Constantin and L. Molinet, Global weak soulutions for a shallow water equation, Commun. Math. Phys. 211 (2000), 45–61.

[Dol] L. Dolan, Kac-Moody algebra is hidden symmetry of chiral models, Phys. Rev. Lett. 47 (1981), 1371–1374.

[DJ] P. Drazin and R. Johnson, Solitons: an introduction, Cambridge Univ. Press, 1989.

[Eb] P. Eberlein, Left invariant geometry of Lie groups, Cubo 6 (2004), 427–510.

[EM] D. Ebin and J. Marsden, Groups of diffeomorphisms and the motion of an incompressible fluid, Ann. Math. 92 (1970), 102–163.

[Fl] H. Flashka, Construction of conservation laws for Lax equations: comments on a paper by G. Wilson, Quart. J. Math. 34 (1983), 61–65.

[FNR] H. Flashka, A. Newell, and T. Ratiu, Kac-Moody Lie algebras and soliton equations, Physica D 9:2 (1983), 300–332.

[FF] B. Fuchssteiner and A. Fokas, Symplectic structures, their Bäcklund transformations and hereditary symmetries, Physica D 4 (1981/82), 47–66.

[Ga] C. Gardner, Korteweg-deVries equation and generalizations. IV. The Koretweg-deVries equation as a Hamiltonian system, J. Math. Phys. 12 (1971), 1548–1551.

[GD] I. Gel'fand and L. Dickii, Fractional powers of operators and Hamiltonian systems, Funct. Anal. and Appl. 10 (1976), 259–273.

[Gu] M. Guest, Harmonic Maps, Loop Groups, and Integrable Systems, Cambridge Univ. Press, 1997.

[Ham] R. Hamilton, The Inverse function theorem of Nash and Moser, Bull. AMS 7 (1982), 65–222.

[Kac] V. Kac, Infinite Dimensional Lie Algebras, Birkhauser, Boston, 1983.

[Kac2] V. Kac (ed.), Infinite Dimensional Groups with Applications, MSRI Publ. #4, Springer-Verlag, New York, 1985.

[Kat] T. Kato, On the Cauchy problem for the (generalized) Korteweg-deVries equation, Adv. in Math. 8 (1983), 93–128.

[KPV] C. Kenig, G. Ponce, and L. Vega, Well-posedness of the initial-value problem for the Korteweg-deVries equation, JAMS 4 (1991), 323–347.

[KM] B. Khesin and G. Misiolek, Euler equations on homogeneous spaces and Virasoro orbits, Advances in Math. 176 (2003), 116–144.

[Kic] S. Kichenassamy, Nonlinear Wave Equations, Marcel Dekker, 1995.

[Ko] S. Kouranbaeva, The Camassa-Holm equation as a geodesic flow on the diffeomorphism group, J. Math. Phys. 40 (1999), 857–868.

[Ku] S. Kumar, Kac-Moody Groups, Their Flag Varieties, and Representation Theory, Birkhauser, Boston, 2002.

[Lax] P. Lax, Integrals of nonlinear equations of evolution and solitary waves, CPAM 21 (1968), 467–490.

[Lax2] P. Lax, Periodic solutions of the KdV equation, CPAM 28 (1975), 141–188.

[Lawd] D. Lawden, Elliptic Functions and Applications, Springer-Verlag, New York, 1989.

[Man] Yu. Manin, Algebraic aspects of non-linear differential equations, J. Soviet Math. 11 (1979), 1–122.

[MaR] J. Marsden and T. Ratiu, Introduction to Mechanics and Symmetry, Springer-Verlag, New York, 1994.

[Mc] H. McKean, Shallow water and the diffeomorphism group, Applied and Industrial Math., Venice-2, 1998, Kluwer Acad. Publ., Dordrecht (2000), 135–143.

[Mc2] H. McKean, Fredholm determinants and the Camassa-Holm heirarchy, CPAM 56 (2003), 638–680.

[MiR] P. Michor and T. Ratiu, On the geometry of the Virasoro-Bott group, J. Lie Theory 8 (1998), 293–309.

[Mis1] G. Misiolek, A shallow water equation as a geodesic flow on the Bott-Virasoro group, J. Geom. Phys. 24 (1998), 203–208.

[Mis2] G. Misiolek, Classical solutions of the periodic Camassa-Holm equation, GAFA 12 (2002), 1080–1104.

[Miu] R. Miura (ed.), Bäcklund Transformations, LNM #515, Springer-Verlag, New York, 1976.

[Mum] D. Mumford, An algebro-geometric construction of commuting operators and of solutions to the Toda lattice equation, Korteweg-deVries equation and related non-linear equations, Proc. Kyoto Conf., 1978.

[New] A. Newell, Solitons in Mathematics and Physics, SIAM, Philadelphia, 1985.

[Olv] P. Olver, Applications of Lie Groups to Differential Equations, Springer-Verlag, New York, 1986

[Pal] R. Palais, The symmetries of solitons, Bull. AMS 34 (1997), 339–403.

[Per] A. Perelomov, Integrable Systems of Classical Mechanics and Lie Algebras, Vol. 1, Birkhauser, Boston, 1990.

[PS] A. Pressley and G. Segal, Loop Groups, Oxford Univ. Press, 1986.

[S] G. Segal, The geometry of the KdV equation, Int. J. Mod. Phys. A6 (1991), 2859–2869.

[SeW] G. Segal and G. Wilson, Loop groups and equations of KdV type, Publ. Math. IHES 61 (1985), 5–65.

[STS] M. Semenev-Tian-Shanski, What a classical r-matrix is, Funct. Anal. Appl. 17 (1983), 259–272.

[Sh] S. Shkoller, Geometry and curvature of diffeomorphism groups with H^1 metric and mean hydrodynamics, J. Funct. Anal. 160 (1998), 337–365.

[Sh2] S. Shkoller, On incompressible averaged Lagrangian hydrodynamics, arXiv:math/9908109v4.

[Si] B. Simon, Trace Ideals and Their Applications, Cambridge Univ. Press, 1979.

[SS] C. Sulem and P.-L. Sulem, The Nonlinear Schrödinger Equation, Springer-Verlag, New York, 1999.

[Sy] W. Symes, Hamiltonian group actions and integrable systems, Physica D1 (1980), 523–557.

[T] M. Taylor, Pseudodifferential Operators, Princeton Univ. Press, Princeton, N.J., 1981.

[T2] M. Taylor, Pseudodifferential Operators and Nonlinear PDE, Birkhauser, Boston, 1991.

[T3] M. Taylor, Partial Differential Equations, Vols. 1–3, Springer-Verlag, New York, 1996; second ed., 2011.

[T4] M. Taylor, Lectures on Lie Groups, Lecture Notes, available at http://www.unc.edu/math/Faculty/met/lieg.html

[Te] C.-L. Terng, Soliton equations and differential geometry, J. Diff. Geom. 45 (1997), 407–445.

[TU] C.-L. Terng and K. Uhlenbeck, Surveys in Differential Geometry: Integrable Systems, International Press, Boston, 1998.

[TU2] C.-L. Terng and K. Uhlenbeck, Geometry of solitons, Notices AMS 47 (2000), 17–25.

[U] K. Uhlenbeck, Harmonic maps into Lie groups (classical solutions of the chiral model), J. Diff. Geom. 30 (1989), 1–50.

[Wi] G. Wilson, Commuting flows and conservation laws for Lax equations, Math. Proc. Camb. Phil. Soc. 86 (1979), 131–143.

[Wi2] G. Wilson, On the construction of conservation laws for Lax equations, Quart. J. Math. 32 (1981), 491–512.

[XZ] Z. Xin and P. Zhang, On the weak solution to a shallow water equation, CPAM 53 (2000), 1411–1433.

[ZS] V. Zakharov and A. Shabat, A scheme for integrating the non-linear equations of mathematical physics by the inverse scattering method, Funct. Anal. and Appl. 8 (1974), 226–235.

[Zh] P. Zhidkov, Korteweg-deVries and Nonlinear Schrödinger Equations: Qualitative Theory, LNM #1756, Springer-Verlag, New York, 2001.

Index