INTEGRABLE SYSTEMS

Maciej Dunajski

Department of Applied Mathematics and Theoretical Physics
University of Cambridge
Wilberforce Road, Cambridge CB3 0WA, UK

May 10, 2012

Contents

Introduction

Integrable systems are nonlinear differential equations which 'in principle' can be solved analytically. This means that the solution can be reduced to a finite number of algebraic operations and integrations. Such systems are very rare - most nonlinear differential equations admit chaotic behaviour and no explicit solutions can be written down. Integrable systems nevertheless lead to a very interesting mathematics ranging from differential geometry and complex analysis to quantum field theory and fluid dynamics. The main reference for the course is [6]. There are other books which cover particular topics treated in the course:

- **Integrability of ODEs** [4] (Hamiltonian formalism, Arnold–Liouville theorem, action–angle variables). The integrability of ordinary differential equations is a fairly clear concept (i.e. it can be defined) based on existence of sufficiently many well behaved first integrals, or (as a physicist would put it) constant of motions.

- **Integrability of PDEs** [15], [5](Solitons, Inverse Scattering Transform). The universally accepted definition of integrability does not exist in this case. The phase space is infinite dimensional but having 'infinitely many' first integrals may not be enough - we could have missed every second one. Here one focuses on properties of solutions and solutions generation techniques. We shall study solitons - solitary non-linear waves which preserve their shape (and other characteristics) in the evolution. These soliton solutions will be constructed by means of an inverse problem: recovering a potential from the scattering data.

- **Lie symmetries** [9], [16] (Group invariant solutions, vector fields, symmetry reduction, Painlevé equations). The powerful symmetry methods can be applied to ODEs and PDEs alike. In case of ODEs a knowledge of sufficiently large symmetry group allows a construction of the most general solution. For PDEs the knowledge of symmetries is not sufficient to construct the most general solution, but it can be used to find new solutions from given ones and to reduce PDEs to more tractable ODEs. The PDEs integrable by inverse problems reduce to equations with Painlevé property.

Chapter 1

Integrability in classical mechanics

In this Chapter we shall introduce the integrability of ordinary differential equations. It is a fairly clear concept based on existence of sufficiently many well behaved first integrals.

1.1 Hamiltonian formalism

Motion of a system with n degrees of freedom is described by a trajectory in a $2n$ dimensional phase space M (locally think of an open set in $\mathbb{R}^{2n}$ but globally it can be topologically non-trivial manifold - e.g. a sphere or a torus. See Appendix A) with local coordinates

$$(p_j, q_j), \quad j = 1, 2, \ldots, n.$$

The dynamical variables are functions $f : M \times \mathbb{R} \longrightarrow \mathbb{R}$, so that $f = f(p, q, t)$ where t is called 'time'. Let $f, g : M \times \mathbb{R} \longrightarrow \mathbb{R}$. Define a Poisson bracket of f, g to be a function

$$\{f, g\} := \sum_{k=1}^{n} \frac{\partial f}{\partial q_k}\frac{\partial g}{\partial p_k} - \frac{\partial f}{\partial p_k}\frac{\partial g}{\partial q_k}. \tag{1.1}$$

It satisfies

$$\{f, g\} = -\{g, f\}, \qquad \{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = 0.$$

The second property is called the Jacobi identity. The coordinate functions (p_j, q_j) satisfy the canonical commutation relations

$$\{p_j, p_k\} = 0, \quad \{q_j, q_k\} = 0, \quad \{q_j, p_k\} = \delta_{jk}.$$

Given a Hamiltonian $H = H(p, q, t)$ (usually $H(p, q)$) the dynamics is determined by

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \{f, H\}, \qquad \text{for any} \qquad f = f(p, q, t).$$

Setting $f = p_j$ or $f = q_j$ yields Hamilton's equations of motion

$$\dot{p}_j = -\frac{\partial H}{\partial q_j}, \qquad \dot{q}_j = \frac{\partial H}{\partial p_j}. \tag{1.2}$$

The system (1.2) of $2n$ ODEs is *deterministic* in a sense that $(p_j(t), q_j(t))$ are uniquely determined by $2n$ initial conditions $(p_j(0), q_j(0))$. Equations (1.2) also imply that volume elements in phase space are conserved. This system is essentially equivalent to Newton's equations of motion. The Hamiltonian formulation allows a more geometric insight to classical mechanics. It is also the starting point to quantisation.

Definition 1.1.1 *A function $f = f(p_j, q_j, t)$ which satisfies $\dot{f} = 0$ when equations* (1.2) *hold is called a first integral or a constant of motion. Equivalently,*

$$f(p(t), q(t), t) = const$$

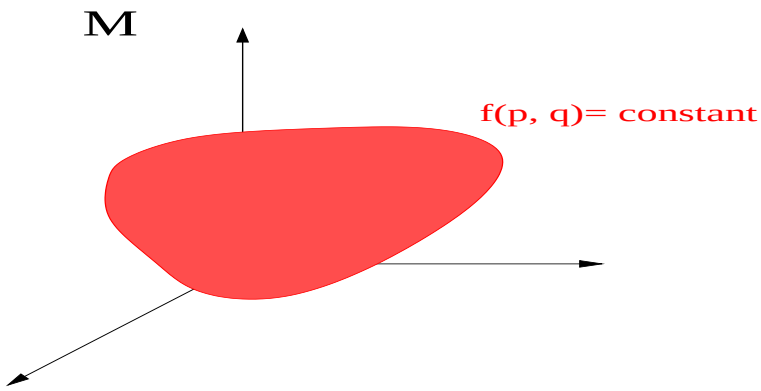
if $p(t), q(t)$ are solutions of (1.2).

In general the system (1.2) will be solvable if it admits 'sufficiently many' first integrals and the reduction of order can be applied. This is because any first integral eliminates one equation.

- **Example.** Consider a system with one degree of freedom with $M = \mathbb{R}^2$ and the Hamiltonian
$$H(p, q) = \frac{1}{2}p^2 + V(q).$$
Hamilton's equations (1.2) give
$$\dot{q} = p, \quad \dot{p} = -\frac{dV}{dq}.$$
The Hamiltonian itself is a first integral as $\{H, H\} = 0$. Thus
$$\frac{1}{2}p^2 + V(q) = E$$
where E is a constant called energy. Now
$$\dot{q} = p, \quad p = \pm\sqrt{2(E - V(q)}$$
and one integration gives a solution in the implicit form
$$t = \pm \int \frac{dq}{\sqrt{2(E - V(q))}}.$$
The explicit solution could be found if we can perform the integral on the RHS and invert the relation $t = t(q)$ to find $q(t)$. These two steps are not always possible to take but nevertheless we would certainly regard this system as integrable.

It is useful to adopt a more geometric approach. Assume that a first integral f does not explicitly depend on time, and that it defines a hypersurface $f(p, q) = \text{const}$ in M (Figure.1.1). Two hypersurfaces corresponding to two independent first integrals generically intersect in a surface of co–dimension 2 in M. In general the trajectory lies on a surface of dimension $2n - L$ where L is the number of independent first integrals. If $L = 2n - 1$ this surface is a curve - a solution to (1.2).

Figure 1.1: Level surface



How to find first integrals? Given two first integrals which do not explicitly depend on time their Poisson bracket will also be a first integral if it is not zero. This follows from the Jacobi identity and the fact all first integrals Poisson commute with the Hamiltonian. More generally, the Noether theorem gives some first integrals (they correspond to symmetries Hamilton's equation (1.2) may possess e.g. time translation, rotations) but not enough. The difficulty with finding the first integrals has deep significance. For assume we use some existence theorem for ODEs and apply it to (1.2). Now solve the algebraic equations

$$q_k = q_k(p^0, q^0, t), \quad p_k = p_k(p^0, q^0, t),$$

for the initial conditions (p^0, q^0) thus giving

$$q^0{}_k = q^0{}_k(p, q, t), \quad p^0{}_k = p^0{}_k(p, q, t).$$

This gives $2n$ first integrals as obviously (p^0, q^0) are constants which we can freely specify. One of these integrals determines the time parametrisations and others could perhaps be used to construct the trajectory in the phase space. However for some of the integrals the equations

$$f(p, q) = \text{const}$$

may not define a 'nice' surface in the phase space. Instead it defines a pathological (at least from the applied mathematics point of view) set which densely covers the phase space. Such integrals do not separate points in M.

One first integral - energy - always exist for Hamiltonian systems giving the energy surface $H(p, q) = E$, but often it is the only first integral. Sufficiently complicated, deterministic, systems may behave according to the laws of thermodynamics: probability that the system is contained in some element of the energy surface is proportional to the normalised volume of this element. This means that the time evolution covers uniformly the entire region of the constant energy surface in the phase space. It is not known wether this ergodic postulate can be derived from Hamilton's equations.

Early computer simulations in the 1960s revealed that some nonlinear systems (with infinitely many degrees of freedom!) are not ergodic. Soliton equations

$$u_t = 6uu_x - u_{xxx}, \quad u = u(x,t), \qquad KdV$$

or

$$\phi_{xx} - \phi_{tt} = \sin\phi, \quad \phi = \phi(x,t), \qquad Sine-Gordon$$

are examples of such systems. Both posses infinitely many first integrals. We shall study them in Chapter 2.

1.2 Integrability and action–angle variables

Given a system of Hamilton's equations (1.2) it is often sufficient to know n (rather than $2n-1$) first integrals as each of them reduces the order of the system by two. This underlies the following definition of an *integrable system.*

Definition 1.2.1 *An integrable system consists of a $2n$-dimensional phase-space M together with n independent functions (in a sense that the gradients ∇f_j are linearly independent vectors on a tangent space to any point in M) $f_1, \ldots, f_n : M \to \mathbb{R}$ such that*

$$\{f_j, f_k\} = 0, \qquad j,k = 1, \ldots, n. \tag{1.3}$$

The vanishing of Poisson brackets (1.3) means that the first integrals are in involution. We shall show that integrable systems lead to completely solvable Hamilton's equations of motion. Let us first explore the freedom in (1.2) given by a coordinate transformation of a phase-space

$$Q_k = Q_k(p,q), \qquad P_k = P_k(p,q).$$

This transformation is called *canonical* if it preserves the Poisson bracket

$$\sum_{k=1}^{n} \frac{\partial f}{\partial q_k}\frac{\partial g}{\partial p_k} - \frac{\partial f}{\partial p_k}\frac{\partial g}{\partial q_k} = \sum_{k=1}^{n} \frac{\partial f}{\partial Q_k}\frac{\partial g}{\partial P_k} - \frac{\partial f}{\partial P_k}\frac{\partial g}{\partial Q_k}$$

for all $f, g : M \longrightarrow \mathbb{R}$. Canonical transformations preserve Hamilton's equation (1.2).

Given a function $S(q,P,t)$ such that

$$\det\left(\frac{\partial^2 S}{\partial q_j \partial P_k}\right) \neq 0$$

we can construct a canonical transformation by setting

$$p_k = \frac{\partial S}{\partial q_k}, \quad Q_k = \frac{\partial S}{\partial P_k}, \quad \widetilde{H} = H + \frac{\partial S}{\partial t}.$$

The function S is an example of a generating function [4, 11, 19]. The idea behind the following Theorem is to seek a canonical transformation such that in the new variables $H = H(P_1, \ldots, P_n)$ so that

$$P_k(t) = P_k(0) = \text{const}, \qquad Q_k(t) = Q_k(0) + t\frac{\partial H}{\partial P_k}.$$

Finding a generating function for such canonical transformation is in practise very difficult, and deciding whether a given Hamiltonian system is integrable (without a priori knowledge of n Poisson commuting integrals) is still an open problem.

Theorem 1.2.2 (Arnold, Liouville) *Let*

$$(M, f_1, \ldots, f_n)$$

be an integrable system with a Hamiltonian $H = f_1$, and let

$$M_f := \{(p,q) \in M; f_k(p,q) = c_k\}, \qquad c_k = const, \qquad k = 1, \ldots, n$$

be an n-dimensional level surface of first integrals f_k. Then

- *If M_f is compact and connected then it is diffeomorphic to a torus*

$$T^n := S^1 \times S^1 \times \ldots \times S^1,$$

and (in a neighbourhood of this torus in M) one can introduce the 'action-angle' coordinates

$$I_1, \ldots, I_n, \phi_1, \ldots, \phi_n, \qquad 0 \leq \phi_k \leq 2\pi,$$

such that angles ϕ_k *are coordinates on M_f and* actions $I_k = I_k(f_1, \ldots, f_n)$ *are first integrals.*

- *The canonical equations of motion* (1.2) *become*

$$\dot{I}_k = 0, \qquad \dot{\phi}_k = \omega_k(I_1, \ldots, I_n), \qquad k = 1, \ldots, n \tag{1.4}$$

and so the integrable systems are solvable by quadratures (a finite number of algebraic operations, and integrations of known functions).

Proof. We shall follow the proof given in [4], but try to make it more accessible by avoiding the language of differential forms

- The motion takes place on the surface

$$f_1(p,q) = c_1, f_2(p,q) = c_2, \ldots, f_n(p,q) = c_n$$

of dimension $2n - n = n$. The first part of the Theorem says that this surface is a torus[1]. For each point in M there exists precisely one torus T^n passing through that point. This means that M admits a foliation by n–dimensional leaves. Each leaf is a torus and different tori correspond to different choices of the constants $c_1, \ldots, c_n$.

Assume

$$\det\Big(\frac{\partial f_j}{\partial p_k}\Big) \neq 0$$

so that the system $f_k(p,q) = c_k$ can be solved for the momenta p_i

$$p_i = p_i(q, c)$$

[1]This part of the proof requires some knowledge of Lie groups and Lie algebras. It is given in Appendix A.

and the relations $f_i(q, p(q, c)) = c_i$ hold identically. Differentiate these identities with respect to q_j

$$\frac{\partial f_i}{\partial q_j} + \sum_k \frac{\partial f_i}{\partial p_k}\frac{\partial p_k}{\partial q_j} = 0$$

and multiply the resulting equations by $\partial f_m/\partial p_j$

$$\sum_j \frac{\partial f_m}{\partial p_j}\frac{\partial f_i}{\partial q_j} + \sum_{j,k} \frac{\partial f_m}{\partial p_j}\frac{\partial f_i}{\partial p_k}\frac{\partial p_k}{\partial q_j} = 0.$$

Now swap the indices and subtract $(mi) - (im)$. This yields

$$\{f_i, f_m\} + \sum_{j,k} \Big(\frac{\partial f_m}{\partial p_j}\frac{\partial f_i}{\partial p_k}\frac{\partial p_k}{\partial q_j} - \frac{\partial f_i}{\partial p_j}\frac{\partial f_m}{\partial p_k}\frac{\partial p_k}{\partial q_j}\Big) = 0.$$

The first term vanishes as the first integrals are in involution. Rearranging the indices in the second term gives

$$\sum_{j,k} \frac{\partial f_i}{\partial p_k}\frac{\partial f_m}{\partial p_j}\Big(\frac{\partial p_k}{\partial q_j} - \frac{\partial p_j}{\partial q_k}\Big) = 0$$

and, as the matrices $\partial f_i/\partial p_k$ are invertible,

$$\frac{\partial p_k}{\partial q_j} - \frac{\partial p_j}{\partial q_k} = 0. \tag{1.5}$$

This condition implies that

$$\oint \sum_j p_j dq_j = 0$$

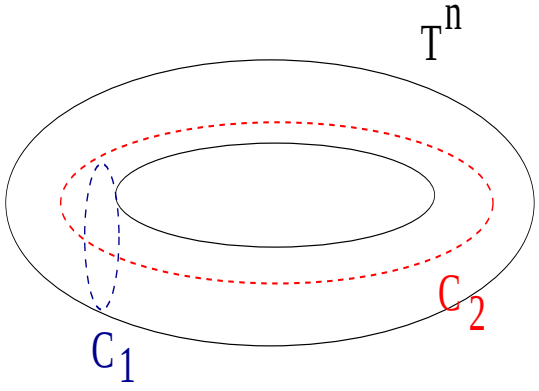
for any closed contractible curve on the torus T^n. This is a consequence of the Stokes theorem. To see it recall that in $n = 3$

$$\oint_{\delta D} \mathbf{p} \cdot d\mathbf{q} = \int_D (\nabla \times \mathbf{p}) \cdot d\mathbf{q}$$

where δD is a boundary of a surface D and

$$(\nabla \times \mathbf{p})_m = \frac{1}{2}\epsilon_{jkm}\Big(\frac{\partial p_k}{\partial q_j} - \frac{\partial p_j}{\partial q_k}\Big).$$

- There are n closed curves which can not be contracted down to a point, so that the corresponding integrals do not vanish.



Cycles on a Torus

Therefore we can define the action coordinates

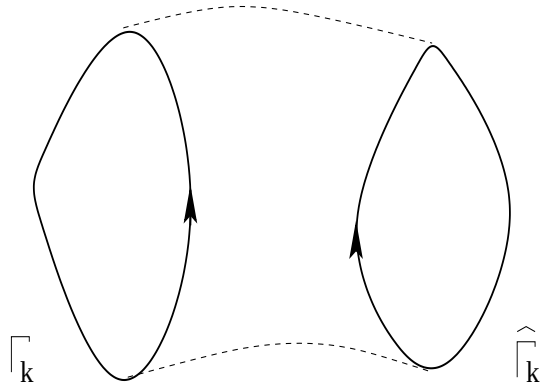
$$I_k := \frac{1}{2\pi} \oint_{\Gamma_k} \sum_j p_j dq_j, \tag{1.6}$$

where the closed curve Γ_k is the k-th basic cycle (the term 'cycle' in general means 'submanifold without boundary') of the torus T^n

$$\Gamma_k = \{(\tilde{\phi}_1, \ldots, \tilde{\phi}_n) \in T^n; 0 \leq \tilde{\phi}_k \leq 2\pi, \tilde{\phi}_j = \text{const} \ \text{ for } j \neq k\},$$

where $\tilde{\phi}$ are some coordinates[2] on T^n.

The Stokes theorem implies that the actions (1.6) are independent on the choice of Γ_k.



Stokes Theorem

This is because

$$\oint_{\Gamma_k} \sum_j p_j dq_j + \oint_{\hat{\Gamma}_k} \sum_j p_j dq_j = \int \Big(\frac{\partial p_i}{\partial q_j} - \frac{\partial p_j}{\partial q_i}\Big) dq_j \wedge dq_i = 0$$

where we have chosen Γ and $\hat{\Gamma}$ to have opposite orientations.

- The actions (1.6) are also first integrals as $\oint p(q,c)dq$ only depends on $c_k = f_k$ and f_ks are first integrals. The actions are Poisson commuting

$$\{I_i, I_j\} = \sum_{r,s,k} \frac{\partial I_i}{\partial f_r}\frac{\partial f_r}{\partial q_k}\frac{\partial I_j}{\partial f_s}\frac{\partial f_s}{\partial p_k} - \frac{\partial I_i}{\partial f_r}\frac{\partial f_r}{\partial p_k}\frac{\partial I_j}{\partial f_s}\frac{\partial f_s}{\partial q_k} = \sum_{r,s} \frac{\partial I_i}{\partial f_r}\frac{\partial I_j}{\partial f_s}\{f_r, f_s\} = 0$$

and in particular $\{I_k, H\} = 0$.

The torus M_f can be equivalently represented by

$$I_1 = \tilde{c}_1, \qquad \ldots, \qquad I_1 = \tilde{c}_n.$$

for some constants $\tilde{c}_1, \ldots, \tilde{c}_n$ (We might have been tempted just to define $I_k = f_k$ but then the transformation $(p,q) \to (I,\phi)$ would not be canonical in general.)

[2]This is a non-trivial step. In practice it is unclear how to explicitly describe the n–dimensional torus and the curves Γ_k in $2n$ dimensional phase space. Thus, to some extend the Arnold–Liouville theorem has a character of the existence theorem.

- We shall construct the angle coordinates ϕ_k canonically conjugate to the actions using a generating function

$$S(q, I) = \int_{q_0}^{q} \sum_j p_j dq_j,$$

where q_0 is some chosen point on the torus. This definition does not depend on a path joining q_0 and q as a consequence of (1.5) and Stokes's theorem. Choosing a different q_0 just adds a constant to S thus leaving the *angles*

$$\phi_i = \frac{\partial S}{\partial I_i}$$

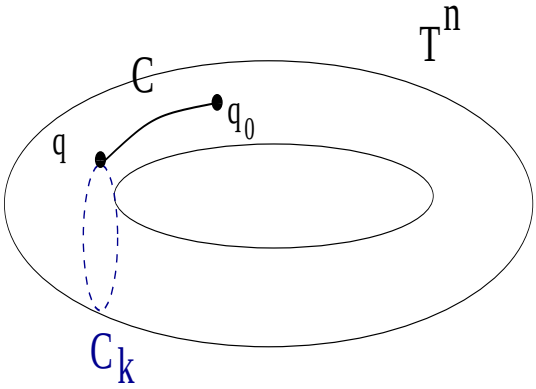
invariant.

- The angles are periodic coordinates with a period 2π. To see it consider two paths C and $C \cup C_k$ (where C_k represents the kth cycle) between q_0 and q and calculate

$$S(q, I) = \int_{C \cup C_k} \sum_j p_j dq_j = \int_C \sum_j p_j dq_j + \int_{C_k} \sum_j p_j dq_j = S(q, I) + 2\pi I_k$$

so

$$\phi_k = \frac{\partial S}{\partial I_k} = \phi_k + 2\pi.$$



Generating Function

- The transformations

$$q = q(\phi, I), \quad p = p(\phi, I), \qquad \text{and} \qquad \phi = \phi(q, p), \quad I = I(q, p)$$

are canonical (as they are defined by a generating function) and invertible. Thus

$$\{I_j, I_k\} = 0, \quad \{\phi_j, \phi_k\} = 0, \quad \{\phi_j, I_k\} = \delta_{jk}$$

and the dynamics is given by

$$\dot{\phi}_k = \{\phi_k, \widetilde{H}\}, \quad \dot{I}_k = \{I_k, \widetilde{H}\},$$

where

$$\widetilde{H}(\phi, I) = H(q(\phi, I), p(\phi, I)).$$

The I_ks are first integrals, therefore

$$\dot{I}_k = -\frac{\partial \widetilde{H}}{\partial \phi_k} = 0$$

so $\widetilde{H} = \widetilde{H}(I)$ and

$$\dot{\phi}_k = \frac{\partial \widetilde{H}}{\partial I_k} = \omega_k(I)$$

where the ω_ks are also first integrals. This proves (1.4). Integrating these canonical equations of motion yields

$$\phi_k(t) = \omega_k(I)t + \phi_k(0), \quad I_k(t) = I_k(0). \tag{1.7}$$

These are n circular motions with constant angular velocities.

□

The trajectory (1.7) may be closed on the torus or it may cover it densely. That depends on the values of the angular velocities. If $n = 2$ the trajectory will be closed if ω_1/ω_2 is rational and dense otherwise.

Interesting things happen to the tori under a small perturbation of the integrable Hamiltonian

$$H(I) \longrightarrow H(I) + \epsilon K(I, \phi).$$

In some circumstances the motion is still periodic and most tori do not vanish but become deformed. This is governed by the Kolmogorov–Arnold–Moser (KAM) theorem - not covered in this course. Consult the popular book by Schuster [18], or read the complete account given by Arnold [4].

- **Example.** All time-independent Hamiltonian system with two-dimensional phase spaces are integrable. Consider the harmonic oscillator with the Hamiltonian

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

  Different choices of the energy E give a foliation of M_f by ellipses

  $$\frac{1}{2}(p^2 + \omega^2 q^2) = E.$$

  For a fixed value of E we can take $\Gamma = M_f$. Therefore

  $$I = \frac{1}{2\pi} \oint_{M_f} pdq = \frac{1}{2\pi} \int \int_S dpdq = \frac{E}{\omega}$$

  where we used the Stokes's theorem to express the line integral in terms of the area enclosed by M_f.

The Hamiltonian expressed in the new variables is $\widetilde{H} = \omega I$ and

$$\dot{\phi} = \frac{\partial \widetilde{H}}{\partial I} = \omega, \qquad \phi = \omega t + \phi_0.$$

To complete the picture we need to express (I, ϕ) in terms of (p, q). We already know

$$I = \frac{1}{2}\Big(\frac{1}{\omega}p^2 + \omega q^2\Big).$$

Thus the generating function is

$$S(q, I) = \int p dq = \pm \int \sqrt{2I\omega - \omega^2 q^2} dq$$

and (choosing a sign)

$$\phi = \frac{\partial S}{\partial I} = \int \frac{\omega dq}{\sqrt{2I\omega - \omega^2 q^2}} = \arcsin\Big(q\sqrt{\frac{\omega}{2I}}\Big) - \phi_0.$$

This gives

$$q = \sqrt{\frac{2I}{\omega}} \sin(\phi + \phi_0)$$

and finally we recover the familiar solution

$$p = \sqrt{2E}\cos(\omega t + \phi_0), \qquad q = \sqrt{2E/\omega^2}\sin(\omega t + \phi_0).$$

- **Example.** The Kepler problem is another doable example. Here the four–dimensional phase space is coordinatised by $(q_1 = \phi, q_2 = r, p_1 = p_\phi, p_2 = p_r)$ and the Hamiltonian is

$$H = \frac{{p_\phi}^2}{2r^2} + \frac{{p_r}^2}{2} - \frac{\alpha}{r}$$

where $\alpha > 0$ is a constant. One readily verifies that

$$\{H, p_\phi\} = 0$$

so the system is integrable in the sense of Definition 1.2.1. The level set M_f of first integrals is given by

$$H = E, \quad p_\phi = \mu$$

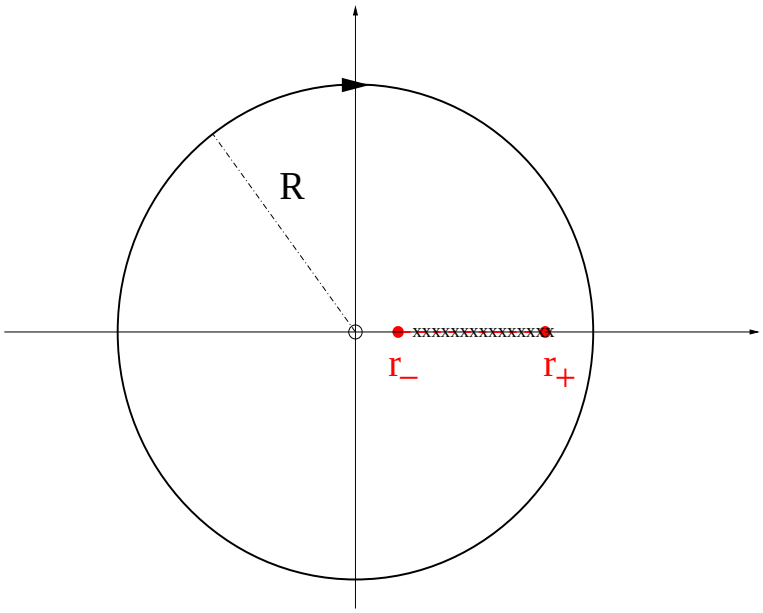
which gives

$$p_\phi = \mu, \quad p_r = \pm\sqrt{2E - \frac{\mu^2}{r^2} + \frac{2\alpha}{r}}.$$

This leaves ϕ arbitrary and gives one constraint on (r, p_r). Thus ϕ and one function of (r, p_r) parametrise M_f. Varying ϕ and fixing the other coordinate gives one cycle $\Gamma_\phi \subset M_f$ and

$$I_\phi = \frac{1}{2\pi}\oint_{\Gamma_\phi} p_\phi d\phi + p_r dr = \frac{1}{2\pi}\int_0^{2\pi} p_\phi d\phi = p_\phi.$$

Figure 1.2: Branch cut for the Kepler integral.



To find the second action coordinate fix ϕ (on top of H and p_ϕ). This gives another cycle Γ_r and

$$\begin{aligned} I_r &= \frac{1}{2\pi}\oint_{\Gamma_r} p_r dr \\ &= 2\frac{1}{2\pi}\int_{r_-}^{r_+}\sqrt{2E-\frac{\mu^2}{r^2}+\frac{2\alpha}{r}}dr \\ &= \frac{\sqrt{-2E}}{\pi}\int_{r_-}^{r_+}\frac{\sqrt{(r-r_-)(r_+-r)}}{r}dr \end{aligned}$$

where the periodic orbits have $r_- \leq r \leq r_+$ and

$$r_\pm = \frac{-\alpha \pm \sqrt{\alpha^2+2\mu^2 E}}{2E}.$$

The integral can be performed using the residue calculus and choosing a contour with a branch cut from r_- to r_+ on the real axis [3]. Consider a branch of

$$f(z) = \sqrt{(z-r_-)(r_+-z)}$$

defined by a branch cut from r_- to r_+ with $f(0) = i\sqrt{(}r_+r_-)$ on the top side of the cut. We evaluate the integral over a large circular contour $|z| = R$ integrating the Laurent

[3]The following method is taken from Max Born's *The Atom* published in 1927. I thank Gary Gibbons for pointing out this reference to me.

expansion

$$\begin{aligned}\int_{|z|=R} z^{-1}f(z)dz &= \int_0^{2\pi}\sqrt{-1}\Big(1-\frac{r_-}{R}e^{-i\theta}\Big)^{1/2}\Big(1-\frac{r_+}{R}e^{-i\theta}\Big)^{1/2}iRe^{i\theta}d\theta\\ &= \pi(r_++r_-) \qquad \text{when} \quad R\to\infty,\end{aligned}$$

since all terms containing powers of $\exp(i\theta)$ are periodic and do not contribute to the integral. The same value must arise from a residue at 0 and collapsing the contour onto the branch cut (when calculating the residue remember that $z=0$ is on the left hand side of the cut and thus $\sqrt{-1}=-i$. Integration along the big circle is equivalent to taking a residue at ∞ which is on the right side of the cut where $\sqrt{-1}=i$). Thus

$$\pi(r_++r_-)=2\pi\sqrt{r_+r_-}+\int_{r_-}^{r_+}\frac{\sqrt{(r-r_-)(r_+-r)}}{r}dr-\int_{r_+}^{r_-}\frac{\sqrt{(r-r_-)(r_+-r)}}{r}dr$$

a rational number (here it is equal to 1). The orbits are therefore closed - a remarkable result known lue must arise from a residue at 0 and collapsing the contour onto the branch cut (when calculating the residue remember that $z=0$ is on the left hand side of the cut and thus $\sqrt{-1}=-i$. Integration along the big circle is equivalent to taking a residue at ∞ which is on the right side of the cut where $\sqrt{-1}=i$). Thus

$$\pi(r_++r_-)=2\pi\sqrt{r_+r_-}+\int_{r_-}^{r_+}\frac{\sqrt{(r-r_-)(r_+-r)}}{r}dr-\int_{r_+}^{r_-}\frac{\sqrt{(r-r_-)(r_+-r)}}{r}dr$$

and

$$\begin{aligned}I_r &= \frac{\sqrt{-2E}}{\pi}\,\frac{\pi}{2}(r_++r_--2\sqrt{r_+r_-})\\ &= \alpha\sqrt{\frac{1}{2|E|}}-\mu.\end{aligned}$$

The Hamiltonian becomes

$$\widetilde{H}=-\frac{\alpha^2}{2(I_r+I_\phi)^2}$$

and we conclude that the absolute values of frequencies are equal and given by

$$\frac{\partial\widetilde{H}}{\partial I_r}=\frac{\partial\widetilde{H}}{\partial I_\phi}=\frac{\alpha^2}{(I_r+I_\phi)^3}=\Big(\frac{r_++r_-}{2}\Big)^{-3/2}\sqrt{\alpha}.$$

This is a particular case when the ratio of two frequencies is a rational number (here it is equal to 1). The orbits are therefore closed - a remarkable result known to Kepler.

1.3 Poisson structures

There is a natural way to extend the Hamiltonian formalism by generalising the notion of Poisson bracket (1.1). A geometric approach is given by symplectic geometry [4]. We shall take a lower level (but a slightly more general) point of view and introduce the Poisson structures. The phase space M is m dimensional with local coordinates $(\xi^1,\ldots,\xi^m)$. In particular we do not distinguish between the positions and momenta.

Definition 1.3.1 *A skew–symmetric matrix $\omega^{ab} = \omega^{ab}(\xi)$ is called a Poisson structure if the Poisson bracket defined by*

$$\{f, g\} = \sum_{a,b=1}^{m} \omega^{ab}(\xi) \frac{\partial f}{\partial \xi^a} \frac{\partial g}{\partial \xi^b} \tag{1.8}$$

satisfies

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

The second property is called the Jacobi identity. It puts restrictions on $\omega^{ab}(\xi)$ which can be seen noting that

$$\omega^{ab}(\xi) = \{\xi^a, \xi^b\}$$

and evaluating the Jacobi identity on coordinate functions.

Given a Hamiltonian $H : M \times \mathbb{R} \longrightarrow \mathbb{R}$ the dynamics is governed by

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \{f, H\}$$

and the Hamilton's equations generalising (1.2) become

$$\dot{\xi}^a = \sum_{b=1}^{m} \omega^{ab}(\xi) \frac{\partial H}{\partial \xi^b}. \tag{1.9}$$

- **Example**. Let $M = \mathbb{R}^3$ and $\omega^{ab} = \sum_{c=1}^{3} \varepsilon^{abc} \xi^c$, where ε^{abc} is the standard totally anti-symmetric tensor. Thus

  $$\{\xi^1, \xi^2\} = \xi^3, \quad \{\xi^3, \xi^1\} = \xi^2, \quad \{\xi^2, \xi^3\} = \xi^1.$$

  This Poisson structure admits a Casimir - any function $f(r)$ where

  $$r = \sqrt{(\xi^1)^2 + (\xi^2)^2 + (\xi^3)^2}$$

  Poisson commutes with the coordinate functions

  $$\{f(r), \xi^a\} = 0.$$

  This is independent on the choice of the Hamiltonian. With a choice

  $$H = \frac{1}{2}\Big(\frac{(\xi^1)^2}{a_1} + \frac{(\xi^2)^2}{a_2} + \frac{(\xi^3)^2}{a_3}\Big)$$

  where a_1, a_2, a_3 are constants, the Hamilton's equations (1.9) become the equations of motion of a rigid body fixed at its centre of gravity

  $$\dot{\xi}^1 = \frac{a_3 - a_2}{a_2 a_3} \xi^2 \xi^3, \qquad \dot{\xi}^2 = \frac{a_1 - a_3}{a_1 a_3} \xi^1 \xi^3, \qquad \dot{\xi}^3 = \frac{a_2 - a_1}{a_1 a_2} \xi^1 \xi^2.$$

Assume that $m = 2n$ is even and the matrix ω is invertible with $W_{ab} := (\omega^{-1})_{ab}$. The Jacobi identity implies that the antisymmetric matrix $W_{ab}(\xi)$ is closed, i.e.

$$\partial_a W_{bc} + \partial_c W_{ab} + \partial_b W_{ca} = 0, \qquad \forall a, b, c = 1, \ldots, m.$$

In this case W_{ab} is called a symplectic structure. The Darboux theorem states that in this case there locally exists a coordinate system

$$\xi^1 = q_1, \cdots, \xi^n = q_n, \xi^{n+1} = p_1, \cdots, \xi^{2n} = p_n$$

such that

$$\omega = \begin{pmatrix} 0 & 1_n \\ -1_n & 0 \end{pmatrix}$$

and the Poisson bracket reduces to the standard form (1.1). A simple proof can be found in [4]. One constructs a local coordinate system (p, q) by induction w.r.t half of the dimension of M. Choose a function p_1, and find q_1 by solving the equation $\{q_1, p_1\} = 1$. Then consider a level set of (p_1, q_1) in M which is locally a symplectic manifold. Now look for (p_2, q_2) etc.

- **Example.** The Poisson structure in the last example is degenerate as the matrix ω^{ab} is not invertible. This degeneracy always occurs if the phase space is odd dimensional or/and there exists a non-tivial Casimir. Consider the restriction of $\omega^{ab} = \sum_{c=1}^{3} \varepsilon^{abc}\xi^c$ to a two-dimensional sphere $r = C$. This gives a symplectic structure on the sphere given by

  $$\{\xi^1, \xi^2\} = \sqrt{C^2 - (\xi^1)^2 - (\xi^2)^2}$$

  or

  $$W = \frac{1}{\sqrt{C^2 - (\xi^1)^2 - (\xi^2)^2}} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

  This of course has no Casimir functions apart from constants. It is convenient to choose a different parametrisation of the sphere: if

  $$\xi^1 = C \sin\theta \cos\phi, \quad \xi^2 = C \sin\theta \sin\phi, \quad \xi^3 = C \cos\theta$$

  then in the local coordinates (θ, ϕ) the symplectic structure is given by $\{\theta, \phi\} = \sin^{-1}\theta$ or

  $$W = \sin\theta \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

  which is equal to the volume form on the two–sphere. The radius C is arbitrary. Therefore the Poisson phase space $\mathbb{R}^3$ is foliated by symplectic phase spaces S^2 as there is exactly one sphere centred at the origin through any point of $\mathbb{R}^3$. This is a general phenomenon: fixing the values of the Casimir functions on Poisson spaces gives the foliations by symplectic spaces. The local Darboux coordinates on S^2 are given by $q = -\cos\theta, p = \phi$ as then

  $$\{q, p\} = 1.$$

The Poisson generalisation is useful to set up the Hamiltonian formalism in the infinite–dimensional case. Formally one can think of replacing the coordinates on the trajectory $\xi^a(t)$ by a dynamical variable $u(x,t)$. Thus the discrete index a becomes the continuous independent variable x (think of m points on a string versus the whole string). The phase space $M = \mathbb{R}^m$ is replaced by a space of smooth functions on a line with appropriate boundary conditions (decay or periodic). The whole formalism may be set up making the following replacements

$$\begin{aligned} \text{ODEs} &\longrightarrow \text{PDEs} \\ \xi^a(t), a = 1, \dots, m &\longrightarrow u(x,t), x \in \mathbb{R} \\ \sum_a &\longrightarrow \int_{\mathbb{R}} dx \\ \text{function } f(\xi) &\longrightarrow \text{functional } F[u] \\ \frac{\partial}{\partial \xi^a} &\longrightarrow \frac{\delta}{\delta u}. \end{aligned}$$

The functionals are given by integrals

$$F[u] = \int_{\mathbb{R}} f(u, u_x, u_{xx}, \dots) dx$$

(we could in principle allow the t derivatives but we will not for the reasons to become clear shortly). Recall that the functional derivative is

$$\frac{\delta F}{\delta u(x)} = \frac{\partial f}{\partial u} - \frac{\partial}{\partial x}\frac{\partial f}{\partial (u_x)} + \Big(\frac{\partial}{\partial x}\Big)^2 \frac{\partial f}{\partial (u_{xx})} + \dots$$

and

$$\frac{\delta u(y)}{\delta u(x)} = \delta(y - x)$$

where the δ on the RHS is the Dirac delta which satisfies

$$\int_{\mathbb{R}} \delta(x) dx = 1, \qquad \delta(x) = 0 \text{ for } x \neq 0.$$

The presence of the Dirac delta will constantly remind us that we have entered a territory which is rather slippery from a pure mathematics perspective. We should rest reassured that the formal replacements made above can nevertheless be given a solid functional-analytic foundation. This will not be done in this course.

The analogy with finite dimensional situation (1.8) suggests a following definition of a Poisson bracket

$$\{F, G\} = \int_{\mathbb{R}^2} \omega(x, y, u) \frac{\delta F}{\delta u(x)} \frac{\delta G}{\delta u(y)} dx dy$$

where the Poisson structure $\omega(x, y, u)$ should be such that the bracket is anti–symmetric and the Jacobi identity holds. A canonical (but not the only) choice is

$$\omega(x, y, u) = \frac{1}{2}\frac{\partial}{\partial x}\delta(x - y) - \frac{1}{2}\frac{\partial}{\partial y}\delta(x - y).$$

This is analogous to the Darboux form in which ω^{ab} is a constant and antisymmetric matrix and the Poisson bracket reduces to (1.1). This is because the differentiation operator $\partial/\partial x$ is anti–self–adjoint with respect to an inner product

$$< u, v >= \int_{\mathbb{R}} u(x)v(x)dx$$

which is analogous to a matrix being anti–symmetric. With this choice

$$\{F, G\} = \int_{\mathbb{R}} \frac{\delta F}{\delta u(x)} \frac{\partial}{\partial x} \frac{\delta G}{\delta u(x)} dx \tag{1.10}$$

and the Hamilton's equations become

$$\begin{aligned} \frac{\partial u}{\partial t} &= \{u, H[u]\} = \int_{\mathbb{R}} \frac{\delta u(x)}{\delta u(y)} \frac{\partial}{\partial y} \frac{\delta H}{\delta u(y)} dy \\ &= \frac{\partial}{\partial x} \frac{\delta H[u]}{\delta u(x)}. \end{aligned} \tag{1.11}$$

- **Example.** The KdV equation mentioned earlier is a Hamiltonian system with the Hamiltonian given by the functional

$$H[u] = \int_{\mathbb{R}} \Big(\frac{1}{2} u_x^2 + u^3 \Big) dx.$$

It is assumed that u belongs to the space of functions decaying sufficiently fast at when $x \to \pm\infty$.

Chapter 2

Soliton equations and Inverse Scattering Transform

The universally accepted definition of integrability does not exist for partial differential equations. The phase space is infinite dimensional but having 'infinitely many' first integrals may not be enough - we could have missed every second one. One instead focuses on properties of solutions and solutions generation techniques. We shall study solitons - solitary non-linear waves which preserve their shape (and other characteristics) in the evolution. These soliton solutions will be constructed by means of an inverse problem: recovering a potential from the scattering data.

2.1 History of two examples

Soliton equations originate in the 19th century. Some of them appeared in the study of non-linear wave phenomena and other arose in differential geometry of surfaces in $\mathbb{R}^3$

- The KdV equation

$$u_t - 6uu_x + u_{xxx} = 0, \qquad \text{where} \quad u = u(x,t) \tag{2.1}$$

has been written down, and solved in the simplest case, by Korteweg and de-Vires in 1895 to explain the following account of J. Scott Russell. Russell observed a soliton while ridding on horseback beside a narrow barge channel. The following passage has been taken from J . Scott Russell. Report on waves, Fourteenth meeting of the British Association for the Advancement of Science, 1844. *'I was observing the motion of a boat which was rapidly drawn along a narrow channel by a pair of horses, when the boat suddenly stopped - not so the mass of water in the channel which it had put in motion; it accumulated round the prow of the vessel in a state of violent agitation, then suddenly leaving it behind, rolled forward with great velocity, assuming the form of a large solitary elevation, a rounded, smooth and well-defined heap of water, which continued its course along the channel apparently without change of form or diminution of speed. I followed it on horseback, and overtook it still rolling on at a rate of some eight or nine miles an hour, preserving its original figure some thirty feet long and a foot to a foot and a half in height. Its height gradually diminished, and after a chase of one or two miles I lost it*

in the windings of the channel. Such, in the month of August 1834, was my first chance interview with that singular and beautiful phenomenon which I have called the Wave of Translation'.

- The Sine–Gordon equation

$$\phi_{xx} - \phi_{tt} = \sin\phi \qquad \text{where} \quad \phi = \phi(x,t) \tag{2.2}$$

locally describes the isometric embeddings of surfaces with constant negative Gaussian curvature in the Euclidean space $\mathbb{R}^3$. The function $\phi = \phi(x,t)$ is the angle between two asymptotic directions $\tau = (x+t)/2$ and $\rho = (x-t)/2$ on the surface along which the second fundamental form is zero. If the first fundamental form of a surface parametrised by (ρ, τ) is

$$ds^2 = d\tau^2 + 2\cos\phi\, d\rho d\tau + d\rho^2, \qquad \text{where} \quad \phi = \phi(\tau, \rho)$$

then the Gaussian curvature is constant and equal to -1 provided that

$$\phi_{\tau\rho} = \sin\phi.$$

which is (2.2).

The integrability of the Sine–Gordon equation have been used by Bianchi, Bäcklund, Lie and other classical differential geometers to construct new embeddings.

2.1.1 Physical derivation of KdV

Consider the linear wave equation

$$\Psi_{xx} - \frac{1}{v^2}\Psi_{tt} = 0$$

where $\Psi_{xx} = \partial_x^2\Psi$ etc. which describes a propagation of waves travelling with a constant velocity v. Its derivation is based on three simplifying assumptions:

- There is no dissipation i.e. the equation is invariant with respect to time inversion $t \to -t$.
- The amplitude of oscillation is small and so the nonlinear terms (like Ψ^2) can be omitted.
- There is no dispersion, i.e. the group velocity is constant.

In the derivation of the KdV we follow [15] and relax these assumptions.

The general solution of the wave equation is a superposition of two waves travelling in opposite directions

$$\Psi = f(x - vt) + g(x + vt)$$

where f and g are arbitrary functions of one variable. Each of these two waves is characterised by a linear 1st order PDE, e.g.

$$\Psi_x + \frac{1}{v}\Psi_t = 0 \qquad \longrightarrow \qquad \Psi = f(x - vt).$$

To introduce the dispersion consider a complex wave

$$\Psi = e^{i(kx - \omega(k)t)}$$

where $\omega(k) = vk$ and so the group velocity $d\omega/dk$ equals to the phase velocity v. We change this relation by introducing the dispersion

$$\omega(k) = v(k - \beta k^3 + \ldots)$$

where the absence of even terms in this expansion guarantees real dispersion relations. Let us assume that the dispersion is small and truncate this series keeping only the first two terms. The equation satisfied by

$$\Psi = e^{i(kx - v(kt - \beta k^3 t))}$$

is readily found to be

$$\Psi_x + \beta \Psi_{xxx} + \frac{1}{v}\Psi_t = 0.$$

This can be rewritten in a form of a conservation law

$$\rho_t + j_x = 0,$$

where the density ρ and the current j are given by

$$\rho = \frac{1}{v}\Psi, \qquad j = \Psi + \beta \Psi_{xx}.$$

To introduce nonlinearity modify the current

$$j = \Psi + \beta \Psi_{xx} + \frac{\alpha}{2}\Psi^2.$$

The resulting equation is

$$\frac{1}{v}\Psi_t + \Psi_x + \beta \Psi_{xxx} + \alpha \Psi \Psi_x = 0.$$

The non–zero constants (v, β, α) can be eliminated by a simple change of variables $x \to x - vt$ and rescaling Ψ. This leads to the standard form of the KdV equation

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

The simplest 1–soliton solution found by Korteweg and de-Vires is

$$u(x,t) = -\frac{2\chi^2}{\cosh^2 \chi(x - 4\chi^2 t - \phi_0)}. \tag{2.3}$$

The KdV is not a linear equation therefore multiplying this solution by a constant will not give another solution. The constant ϕ_0 determines the location of the extremum at $t = 0$. We should therefore think of a one–parameter family of solutions labelled by $\chi \in \mathbb{R}$.

The one–soliton (2.3) was the only regular solution of KdV such that $u, u_x \to 0$ as $|x| \to \infty$ known until 1965 when Gardner, Green, Kruskal and Miura analysed KdV numerically. They took two waves with different amplitudes as their initial profile. The computer simulations revealed that the initially separated waves approached each-other distorting their shapes, but eventually the larger wave overtook the smaller wave and both waves re-emerged with their sizes and shapes intact. The relative phase shift was the only result of the non–linear interaction. This behaviour resembles what we usually associate with particles and not waves.

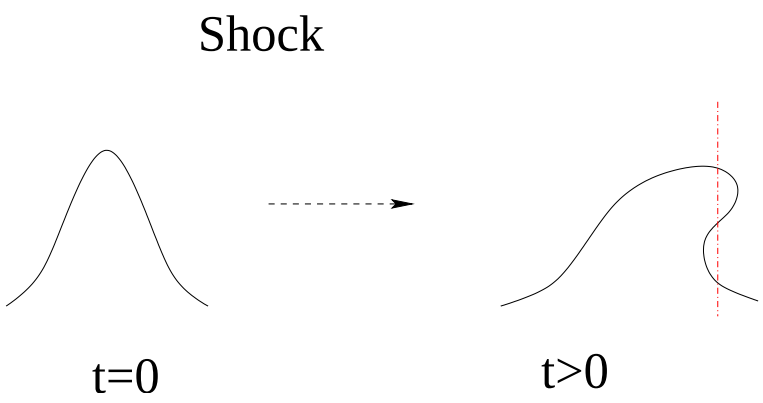
Thus Zabruski and Kruskal named these waves 'solitons' (like electrons, protons, barions and other particles ending with 'ons'). In this Chapter we shall construct more general N–soliton solutions describing the interactions of 1–solitons.

To this end we note that the existence of a stable solitary wave is a consequence of cancellations of effects caused by non–linearity and dispersion.

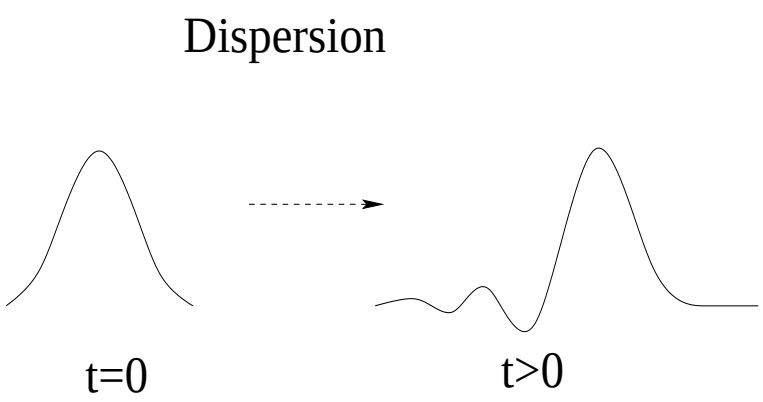
- If the dispersive term were not present the equation would be

$$u_t - 6uu_x = 0$$

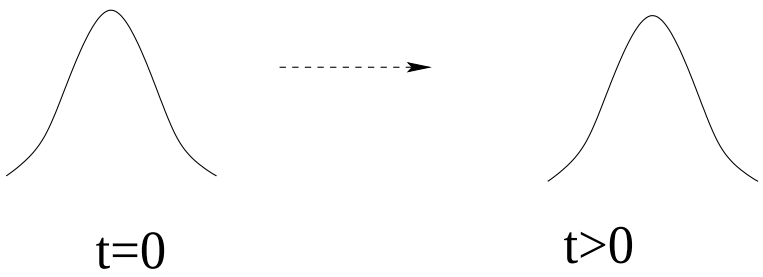
and the resulting solution would exhibit a discontinuity of first derivatives at some $t_0 > 0$ (shock, or 'breaking the wave'). This solution can be easily found using the method of characteristics.



- If the nonlinear term were not present the initial wave profile would disperse in the evolution $u_t + u_{xxx} = 0$.



- The presence of both terms allows smooth localised soliton solutions



of which (2.3) is an example (the plot gives $-u(x,t)$).

2.1.2 Bäcklund transformations for the Sine–Gordon equation

Let us consider the Sine–Gordon equation - the other soliton equation mentioned in the introduction to this Chapter. The simplest solution generating technique is the Bäcklund transformation. Set $\tau = (x+t)/2, \rho = (x-t)/2$ so that the equation (2.2) becomes

$$\phi_{\tau\rho} = \sin\phi.$$

Now define the Bäcklund relations

$$\partial_\rho(\phi_1 - \phi_0) = 2b\sin\Big(\frac{\phi_1+\phi_0}{2}\Big), \qquad \partial_\tau(\phi_1+\phi_0) = 2b^{-1}\sin\Big(\frac{\phi_1-\phi_0}{2}\Big), \qquad b = \text{const.}$$

Differentiating the first equation w.r.t τ, and using the second equation yields

$$\begin{aligned}\partial_\tau\partial_\rho(\phi_1-\phi_0) &= 2b\,\partial_\tau\sin\Big(\frac{\phi_1+\phi_0}{2}\Big) = 2\sin\Big(\frac{\phi_1-\phi_0}{2}\Big)\cos\Big(\frac{\phi_1+\phi_0}{2}\Big)\\ &= \sin\phi_1 - \sin\phi_0.\end{aligned}$$

Therefore ϕ_1 is a solution to the Sine–Gordon equation if ϕ_0 is. Given ϕ_0 we can solve the first order Bäcklund relations for ϕ_1 and generate new solutions form the ones we know. The trivial solution $\phi_0 = 0$ yields 1–soliton solution of Sine–Gordon

$$\phi_1(x,t) = 4\arctan\Big(\exp\Big(\frac{x-vt}{\sqrt{1-v^2}} - x_0\Big)\Big)$$

where v is a constant with $|v| < 1$. This solution is called a kink (Figure 2.1). A static kink corresponds to a special case $v = 0$.

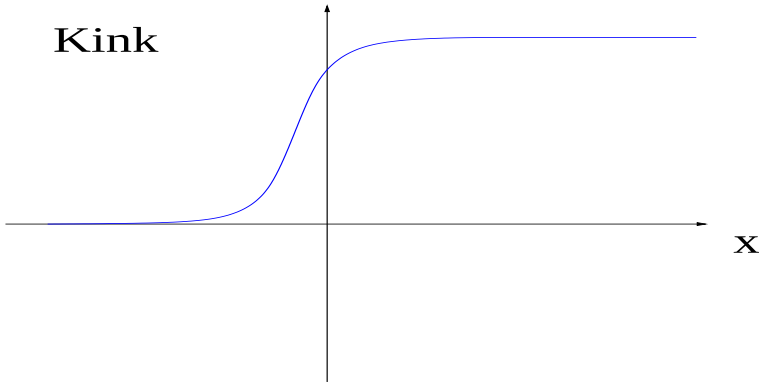
One can associate a topological charge

$$N = \frac{1}{2\pi}\int_{\mathbb{R}} d\phi = \frac{1}{2\pi}\Big(\phi(x=\infty,t) - \phi(x=-\infty,t)\Big)$$

with any solution of the Sine Gordon equation. It is an integral of a total derivative which depends only on boundary conditions. It is conserved if one insists on finiteness of the energy

$$E = \int_{\mathbb{R}}\Big(\frac{1}{2}\Big(\phi_t^2+\phi_x^2\Big) + (1-\cos(\phi))\Big)dx.$$

Figure 2.1: Sine–Gordon Kink



Note that the Sine–Gordon equations didn't enter the discussion at this stage. Topological charges, like N, are in this sense different form first integrals like E which satisfy $\dot{E} = 0$ as a consequence of (2.2). For the given kink solution $N(\phi) = 1$ and the kink is stable as it would take infinite energy to change this solution into a constant solution $\phi = 0$ with $E = 0$.

There exist interesting solutions with $N = 0$: a soliton–antisoliton pair has $N = 0$ but is non–trivial

$$\phi(x,t) = 4\arctan\Big(\frac{v\cosh\frac{x}{\sqrt{1-v^2}}}{\sinh\frac{vt}{\sqrt{1-v^2}}}\Big).$$

At $t \to -\infty$ this solution represents widely separated pair of kink and anti–kink approaching each-other with velocity v. A non–linear interaction takes place at $t = 0$ and as $t \to \infty$ kink and anti-kink reemerge unchanged.

2.2 Inverse scattering transform for KdV

One of the most spectacular methods of solving soliton equations comes from quantum mechanics. It is quite remarkable, as the soliton equations we have discussed so far have little to do with the quantum world.

Recall that the mathematical arena of quantum mechanics is the infinite–dimensional complex vector space $\mathcal{H}$ of functions [17]. Elements Ψ of this space are referred to as wave functions, or state vectors. In case of one–dimensional quantum mechanics we have $\Psi : \mathbb{R} \to \mathbb{C}$, $\Psi = \Psi(x) \in \mathbb{C}$. The space $\mathcal{H}$ is equipped with a unitary inner product

$$(\Psi, \Phi) = \int_{\mathbb{R}} \overline{\Psi(x)}\Phi(x)dx. \tag{2.4}$$

The functions which are square integrable, i.e. $(\Psi, \Psi) < \infty$ like $\Psi = e^{-x^2}$, are called bound states. Other functions, like e^{-ix}, are called the scattering states.

Given a real valued function $u = u(x)$ called the potential, the time independent Schrödinger equation

$$-\frac{\hbar^2}{2m}\frac{d^2\Psi}{dx^2} + u\Psi = E\Psi$$

determines the x–dependence of a wave function. Here $\hbar$ and m are constants which we shall not worry about and E is the energy of the quantum system. The energy levels can be discrete for bound states or continuous for scattering states. This depends on the potential $u(x)$. We shall regard the Schrödinger equation as an eigen–value problem and refer to Ψ and E as eigenvector and eigenvalue respectively.

According to the Copenhagen interpretation of quantum mechanics the probability density for the position of a quantum particle is given by $|\Psi|^2$, where Ψ is a solution to the the Schrödinger equation. The time evolution of the wave function is governed by a time dependent Schrödinger equation

$$i\hbar\frac{\partial\Psi}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\Psi}{\partial x^2} + u\Psi.$$

This equation implies that for bound states the quantum–mechanical probability is conserved in a sense that

$$\frac{d}{dt}\int_{\mathbb{R}} |\Psi|^2 dx = 0.$$

The way physicists discover new elementary particles is by scattering experiments. Huge accelerators collide particles through targets and, by analysing the changes to momenta of scattered particles, a picture of a target is built[1]. Given a potential $u(x)$ one can use the Schrödinger equation to find Ψ, the associated energy levels and the scattering data in the form of so called reflection and transmission coefficients. Experimental needs are however different: the scattering data is measured in the accelerator but the potential (which gives the internal structure of the target) needs to be recovered. This comes down to the following mathematical problem

- Recover the potential from the scattering data.

This problem was solved in the 1950s by the Gelfand, Levitan and Marchenko [8, 14] who gave a linear algorithm for reconstructing $u(x)$. Gardner, Green, Kruskal and Miura [7] used this algorithm to solve the Cauchy problem for the KdV equation. Their remarkable idea was to regard the initial data in the solution of KdV as a potential in the Schrödinger equation.

Set $\hbar^2/(2m) = 1$ and write the 1-dimensional Schrödinger equation as an eigenvalue problem

$$\Big(-\frac{d^2}{dx^2} + u(x)\Big)\Psi = E\Psi.$$

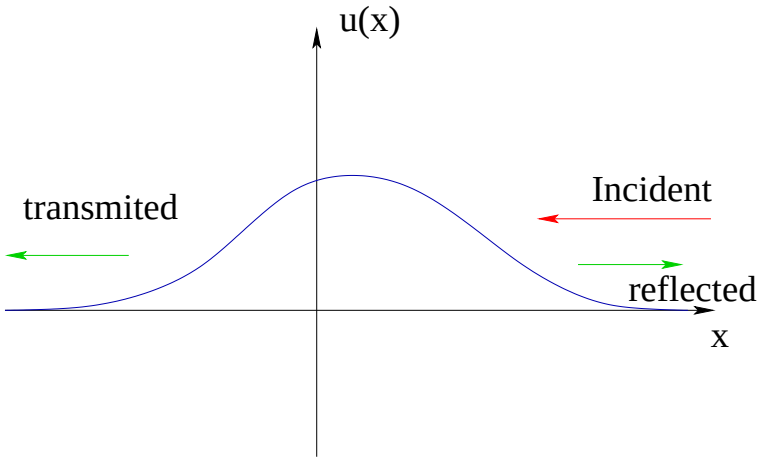
We allow u to depend on x as well as t which at this stage should be regarded as a parameter.

In the scattering theory one considers the beam of free particles incident from $+\infty$. Some of the particles will be reflected by the potential (which is assumed to decay sufficiently fast as $|x| \to \infty$) and some will be transmitted. There may also be a number of bound states with discrete energy levels. The Gelfand–Levitan–Marchenko theory shows that given

- energy levels E,
- transmission probability T,

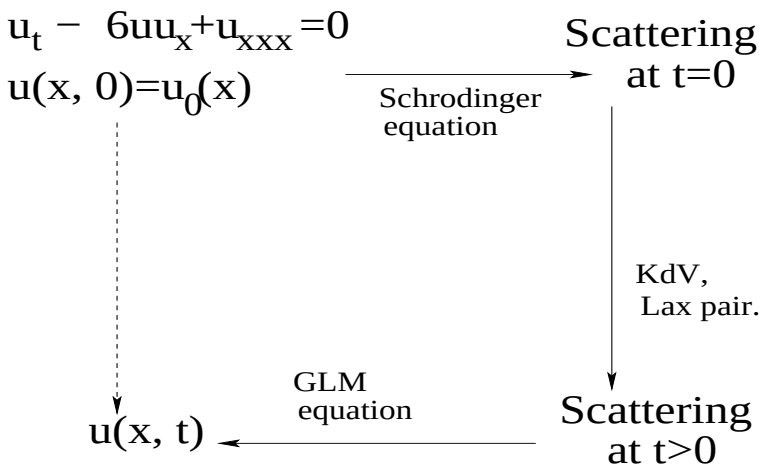
[1]These kind of experiments will take place in the Large Hadron Collider LHC opened in September 2008 at CERN. The LHC is located in a 27km long tunnel under the Swiss/French border outside Geneva. It is hoped that the elusive Higgs particle and a whole bunch of other exotic form of matter will be discovered.

Figure 2.2: Reflection and Transmission



- reflection probability R,

one can find the potential u: Given $u_0(x)$ one finds the scattering data at $t = 0$. If $u(x,t)$ is a solution to the KdV equation (2.1) with $u(x,0) = u_0(x)$ then the scattering data $(E(t), T(t), R(t))$ satisfies simple linear ODEs determining their time evolution. In particular E does not depend on t. Once this has been determined, $u(x,t)$ is recovered by solving a linear integral equation. The Gardner, Green, Kruskal and Miura scheme for solving KdV is summarised in the following table



We should stress that in this method the time evolution of the scattering data is governed by the KdV and not by the time dependent Schrödinger equation. In fact the time dependent Schrödinger equation will not play any role in the following discussion.

2.2.1 Direct scattering

The following discussion summarises the basic one–dimensional quantum mechanics of a particle scattering on a potential [17, 15].

- Set $E = k^2$ and rewrite the Schrödinger equation as

$$Lf := \Big(-\frac{d^2}{dx^2} + u(x) \Big) f = k^2 f \tag{2.5}$$

where L is called the Schrödinger operator. Consider the class of potentials $u(x)$ such that

$$\int_{\mathbb{R}} (1 + |x|)|u(x)|dx < \infty$$

which of course implies that $|u(x)| \to 0$ as $x \to \pm\infty$. This condition guaranties that there exists only a finite number of discrete energy levels (thus it rules out both the harmonic oscillator and the hydrogen atom).

- At $x \to \pm\infty$ the problem (2.5) reduces to a 'free particle'

$$f_{xx} + k^2 f = 0$$

with the general solution

$$f = C_1 e^{ikx} + C_2 e^{-ikx}.$$

The pair of constants (C_1, C_2) is in general different at $+\infty$ and $-\infty$.

- For each $k \neq 0$ the set of solutions to (2.5) forms a 2–dimensional complex vector space G_k. The reality of $u(x)$ implies that if f satisfies (2.5) then so does $\overline{f}$.

Consider two bases $(\psi, \overline{\psi})$ and $(\phi, \overline{\phi})$ of G_k determined by the asymptotic

$$\psi(x,k) \cong e^{-ikx}, \quad \overline{\psi}(x,k) \cong e^{ikx} \qquad \text{as} \quad x \longrightarrow \infty$$

and

$$\phi(x,k) \cong e^{-ikx}, \quad \overline{\phi}(x,k) \cong e^{ikx} \qquad \text{as } x \longrightarrow -\infty.$$

Any solution can be expanded in the first basis, so in particular

$$\phi(x,k) = a(k)\psi(x,k) + b(k)\overline{\psi}(x,k).$$

Therefore, if $a \neq 0$, we can write

$$\frac{\phi(x,k)}{a(k)} = \begin{cases} \frac{e^{-ikx}}{a(k)}, & \text{for } x \to -\infty \\ e^{-ikx} + \frac{b(k)}{a(k)} e^{ikx}, & \text{for } x \to \infty. \end{cases} \tag{2.6}$$

- Consider a particle incident from ∞ with the wave function e^{-ikx} (Figure 2.2). The transmission coefficient $t(k)$ and the reflection coefficient $r(k)$ are given by

$$t(k) = \frac{1}{a(k)}, \quad r(k) = \frac{b(k)}{a(k)}.$$

They satisfy

$$|t(k)|^2 + |r(k)|^2 = 1 \tag{2.7}$$

which is intuitively clear as the particle is 'either reflected or transmitted'. To prove it recall that given the Wronskian

$$W(f,g) = fg_x - gf_x$$

of any two functions we have

$$W_x = fg_{xx} - gf_{xx} = 0$$

if f, g both satisfy the Schrödinger equation (2.5). Thus $W(\phi, \overline{\phi})$ is a constant which can be calculated for $x \to -\infty$

$$W(\phi, \overline{\phi}) = e^{-ikx}(e^{ikx})_x - e^{ikx}(e^{-ikx})_x = 2ik.$$

Analogous calculation at $x \to \infty$ gives $W(\psi, \overline{\psi}) = 2ik$. On the other hand

$$\begin{aligned} W(\phi, \overline{\phi}) &= W(a\psi + b\overline{\psi}, \overline{a}\overline{\psi} + \overline{b}\psi) \\ &= |a|^2 W(\psi, \overline{\psi}) + a\overline{b}W(\psi, \psi) + b\overline{a}W(\overline{\psi}, \overline{\psi}) - |b|^2 W(\psi, \overline{\psi}) \\ &= 2ik(|a|^2 - |b|^2). \end{aligned}$$

Thus $|a(k)|^2 - |b(k)|^2 = 1$ or equivalently (2.7) holds.

2.2.2 Properties of the scattering data

Assume that $k \in \mathbb{C}$. In scattering theory (see e.g.[15]) one proves the following

- $a(k)$ is holomorphic in the upper half plane $Im(k) > 0$.
- $\{Im(k) \geq 0, |k| \to \infty\} \quad \longrightarrow \quad |a(k)| \to 1.$
- Zeroes of $a(k)$ in the upper half plane lie on the imaginary axis. The number of these zeroes is finite if

  $$\int_{\mathbb{R}} (1 + |x|)|u(x)| < \infty.$$

  Thus $a(i\chi_1) = \cdots = a(i\chi_N) = 0$ where $\chi_n \in \mathbb{R}$ can be ordered as

  $$\chi_1 > \chi_2 > \cdots > \chi_N > 0.$$

- Consider the asymptotics of ϕ at these zeroes. Formula (2.6) gives

  $$\phi(x, i\chi_n) = \begin{cases} e^{-i(i\chi_n)x}, & \text{for } x \to -\infty \\ a(i\chi_n)e^{-i(i\chi_n)x} + b(i\chi_n)e^{i(i\chi_n)}, & \text{for } x \to \infty. \end{cases}$$

  Thus

  $$\phi(x, i\chi_n) = \left\{ \begin{array}{ll} e^{\chi_n x}, & \text{for } x \to -\infty \\ b_n e^{-\chi_n x}, & \text{for } x \to \infty. \end{array} \right\}. \tag{2.8}$$

where $b_n = b(i\chi_n)$. We should stress that this considerations apply to the discrete part of the spectrum so b_n should be regarded as independent from b which appears in the reflection coefficient. One can show that $b_n \in \mathbb{R}$ and that it satisfies

$$b_n = (-1)^n |b_n|$$

and $ia'(i\chi_n)$ has the same sign as b_n.

Moreover

$$\Big(-\frac{d^2}{dx^2} + u(x) \Big) \phi(x, i\chi_n) = -\chi_n^2 \phi(x, i\chi_n)$$

so ϕ is square integrable with energy $E = -\chi_n^2$.

2.2.3 Inverse Scattering

We want to recover the potential $u(x)$ from the scattering data which consists of the reflection coefficients and the energy levels

$$r(k), \{\chi_1, \ldots, \chi_N\}$$

so that $E_n = -\chi_n^2$ and

$$\phi(x, i\chi_n) = \begin{cases} e^{\chi_n x}, & \text{for } x \to -\infty \\ b_n e^{-\chi_n x}, & \text{for } x \to \infty. \end{cases}$$

The inverse scattering transform Gelfand–Levitan–Marchenko consist of the following steps

- Set
$$F(x) = \sum_{n=1}^{N} \frac{b_n e^{-\chi_n x}}{ia'(i\chi_n)} + \frac{1}{2\pi} \int_{-\infty}^{\infty} r(k) e^{ikx} dk. \tag{2.9}$$

- Consider the GLM integral equation
$$K(x, y) + F(x + y) + \int_x^{\infty} K(x, z) F(z + y) dz = 0 \tag{2.10}$$
and solve it for $K(x, y)$.

- Then
$$u(x) = -2 \frac{d}{dx} K(x, x) \tag{2.11}$$
is the potential in the corresponding Schrödinger equation.

These formulae are given in the t–independent way, but t can be introduced as a parameter. If the time dependence of the scattering data is known, the solution of the GLM integral equation $K(x, y, t)$ will also depend on t and so will the potential $u(x, t)$.

2.2.4 Lax formulation

If the potential $u(x)$ in the Schrödinger equation depends on a parameter t, its eigenvalues will in general change with t. The inverse scattering transform is an example of an isospectral problem, when this does not happen

Proposition 2.2.1 *If there exist a differential operator A such that*

$$\dot{L} = [L, A] \tag{2.12}$$

where

$$L = -\frac{d^2}{dx^2} + u(x,t),$$

then the spectrum of L does not depend on t.

Proof. Consider the eigenvalue problem

$$Lf = Ef.$$

Differentiating gives

$$L_t f + Lf_t = E_t f + Ef_t.$$

Note that $ALf = EAf$ and use the representation (2.12) to find

$$(L - E)(f_t + Af) = E_t f. \tag{2.13}$$

Take the inner product (2.4) of this equation with f and use the fact that L is self–adjoint

$$E_t ||f||^2 =< f, (L - E)(f_t + Af) >=< (L - E)f, f_t + Af >= 0.$$

Thus $E_t = 0$. This derivation also implies that if $f(x,t)$ is an eigenfunction of L with eigenvalue $E = k^2$ then so is $(f_t + Af)$.

□

What makes the method applicable to KdV equation (2.1) is that KdV is equivalent to (2.12) with

$$L = -\frac{d^2}{dx^2} + u(x,t), \quad A = 4\frac{d^3}{dx^3} - 3\Big(u\frac{d}{dx} + \frac{d}{dx}u\Big). \tag{2.14}$$

To prove this statement it is enough to compute both sides of (2.12) on a function and verify that $[L, A]$ is the multiplication by $6uu_x - u_{xxx}$ (also $\dot{L} = u_t$). This is the Lax representation of KdV [12]. Such representations (for various choices of operators L, A) underlie integrability of PDEs and ODEs.

2.2.5 Evolution of the scattering data

We will now use the Lax representation to determine the time evolution of the scattering data. Assume that the potential $u(x,t)$ in the Schrödinger equation satisfies the KdV equation (2.1). Let $f(x,t)$ be an eigenfuction of the Schrödinger operator $Lf = k^2 f$ defined by its asymptotic behaviour

$$f = \phi(x,k) \longrightarrow e^{-ikx}, \qquad \text{as } x \to -\infty.$$

Equation (2.13) implies that if $f(x,t)$ is an eigenfunction of L with eigenvalue k^2 then so is $(f_t + Af)$. Moreover $u(x) \to 0$ as $|x| \to \infty$ therefore

$$\dot{\phi} + A\phi \longrightarrow 4\frac{d^3}{dx^3}e^{-ikx} = 4ik^3 e^{-ikx} \qquad \text{as } x \to -\infty.$$

Thus $4ik^3\phi(x,k)$ and $\dot{\phi} + A\phi$ are eigenfunctions of the Schrödinger operator with the same asymptotic and we deduce that they must be equal: Their difference is in the kernel of $L - k^2$ and so must be a linear combination of ψ and $\overline{\psi}$. But this combination vanishes at ∞ so, using the independence of ψ and $\overline{\psi}$, it must vanish everywhere. Thus the ODE

$$\dot{\phi} + A\phi = 4ik^3\phi$$

holds for all $x \in \mathbb{R}$. We shall use this ODE and the asymptotics at $+\infty$ to find ODEs for $a(k)$ and $b(k)$. Recall that

$$\phi(x,k) = a(k,t)e^{-ikx} + b(k,t)e^{ikx} \quad \text{as } x \to \infty.$$

Substituting this to the ODE gives

$$\begin{aligned} \dot{a}e^{-ikx} + \dot{b}e^{ikx} &= \Big(-4\frac{d^3}{dx^3} + 4ik^3\Big)(ae^{-ikx} + be^{ikx}) \\ &= 8ik^3 be^{ikx}. \end{aligned}$$

Equating the exponentials gives

$$\dot{a} = 0, \qquad \dot{b} = 8ik^3 b$$

and

$$a(k,t) = a(k,0), \qquad b(k,t) = b(k,0)e^{8ik^3 t}.$$

In the last Section we have shown that k does not depend on t and so the zeroes $i\chi_n$ of a are constant. Using formula (2.8) and following the method given above it can be shown that $b_n(t) = b_n(0)e^{8\chi_n^3 t}$. The evolution of the scattering data is thus given by the following

$$\begin{aligned} a(k,t) &= a(k,0), \\ b(k,t) &= b(k,0)e^{8ik^3 t}, \\ r(k,t) &= \frac{b(k,t)}{a(k,t)} = r(k,0)e^{8ik^3 t}, \\ \chi_n(t) &= \chi_n(0), \\ b_n(t) &= b_n(0)e^{8\chi_n^3 t}, \\ a_n(t) &= 0, \\ \beta_n(t) &= \frac{b_n(t)}{ia'(i\chi_n)} = \beta_n(0)e^{8\chi_n^3 t}. \end{aligned} \tag{2.15}$$

2.3 Reflectionless potentials and solitons

The formula (2.15) implies that if the reflection coefficient is initially zero, it is zero for all t. In this case the inverse scattering procedure can be carried out explicitly. The resulting solutions are called N-solitons, where N is the number of zeroes $i\chi_1, \ldots, i\chi_N$ of $a(k)$. These solutions describe collisions of 1–solitons (2.3) without any non-elastic effects. The 1-solitons generated after collisions are 'the same' as those before the collision. This fact was discovered numerically in the 1960s and boosted the interest in the whole subject.

Assume $r(k,0) = 0$ so that (2.15) implies

$$r(k,t) = 0.$$

2.3.1 One soliton solution

We shall first derive the 1-soliton solution. The formula (2.9) with $N = 1$ gives

$$F(x,t) = \beta(t)e^{-\chi x}.$$

This depends on x as well as t because $\beta(t) = \beta(0)e^{8\chi^3 t}$ from (2.15). We shall suppress this explicit t dependence in the following calculation and regard t as a parameter. The GLM equation (2.10) becomes

$$K(x,y) + \beta e^{-\chi(x+y)} + \int_x^\infty K(x,z)\beta e^{-\chi(z+y)}dz = 0.$$

Look for solutions in the form

$$K(x,y) = K(x)e^{-\chi y}.$$

This gives

$$K(x) + \beta e^{-\chi x} + K(x)\beta \int_x^\infty e^{-2\chi z}dz = 0,$$

and after a simple integration

$$K(x) = -\frac{\beta e^{-\chi x}}{1 + \frac{\beta}{2\chi}e^{-2\chi x}}.$$

Thus

$$K(x,y) = -\frac{\beta e^{-\chi(x+y)}}{1 + \frac{\beta}{2\chi}e^{-2\chi x}}.$$

This function also depends on t because β does. Finally the formula (2.11) gives

$$\begin{aligned} u(x,t) &= -2\frac{\partial}{\partial x}K(x,x) = -\frac{4\beta\chi e^{-2\chi x}}{(1 + \frac{\beta}{2\chi}e^{-2\chi x})^2} \\ &= -\frac{8\chi^2}{\hat{\beta}^{-1}e^{\chi x} + \hat{\beta}e^{-\chi x}}, \quad \text{where} \quad \hat{\beta} = \sqrt{\beta/(2\chi)} \\ &= -\frac{2\chi^2}{\cosh\left(\chi(x - 4\chi^2 t - \phi_0)\right)^2}, \quad \phi_0 = \frac{1}{2\chi}\log\left(\frac{\beta_0}{2\chi}\right) \end{aligned}$$

which is the 1-soliton solution (2.3).

The energy of the corresponding solution to the Schrödinger equation determines the amplitude and the velocity of the soliton. The soliton is of the form $u = u(x - 4\chi^2 t)$ so it represents a wave moving to the right with velocity $4\chi^2$ and phase ϕ_0.

2.3.2 N–soliton solution

There are N energy levels which we order $\chi_1 > \chi_2 > \ldots > \chi_N > 0$. The function (2.9) is

$$F(x) = \sum_{n=1}^{N} \beta_n e^{-\chi_n x}$$

and the GLM equation (2.10) becomes

$$K(x,y) + \sum_{n=1}^{N} \beta_n e^{-\chi_n(x+y)} + \int_x^\infty K(x,z) \sum_{n=1}^{N} \beta_n e^{-\chi_n(z+y)} dz = 0.$$

The kernel of this integral equation is degenerate in a sense that

$$F(z+y) = \sum_{n=1}^{N} k_n(z) h_n(y),$$

so we seek solutions of the form

$$K(x,y) = \sum_{n=1}^{N} K_n(x) e^{-\chi_n y}.$$

After one integration this gives

$$\sum_{n=1}^{N} (K_n(x) + \beta_n e^{-\chi_n x}) e^{-\chi_n y} + \sum_{n=1}^{N} \Big(\beta_n \sum_{m=1}^{N} \frac{K_m(x)}{\chi_m + \chi_n} e^{-(\chi_n + \chi_m)x} \Big) e^{-\chi_n y} = 0.$$

The functions $e^{-\chi_n y}$ are linearly independent, so

$$K_n(x) + \beta_n e^{-\chi_n x} + \sum_{m=1}^{N} \beta_n K_m(x) \frac{1}{\chi_m + \chi_n} e^{-(\chi_n + \chi_m)x} = 0.$$

Define a matrix

$$A_{nm}(x) = \delta_{nm} + \frac{\beta_n e^{-(\chi_n + \chi_m)x}}{\chi_n + \chi_m}.$$

The linear system becomes

$$\sum_{m=1}^{N} A_{nm}(x) K_m(x) = -\beta_n e^{-\chi_n x},$$

or

$$AK + B = 0,$$

where B is a column vector

$$B = [\beta_1 e^{-\chi_1 x}, \beta_2 e^{-\chi_2 x}, \cdots, \beta_n e^{-\chi_n x}]^T.$$

The solution of this system is

$$K = -A^{-1}B.$$

Using the relation

$$\frac{dA_{mn}(x)}{dx} = -B_m e^{-\chi_n x}$$

we can write

$$\begin{aligned} K(x,x) &= \sum_{m=1}^{N} e^{-\chi_m x} K_m(x) = -\sum_{m,n=1}^{N} e^{-\chi_m x}(A^{-1})_{mn} B_n \\ &= \sum_{m,n=1}^{N} (A^{-1})_{mn} \frac{dA_{nm}(x)}{dx} = \mathrm{Tr}\Big(A^{-1}\frac{dA}{dx}\Big) \\ &= \frac{1}{\det A}\frac{d}{dx}\det A. \end{aligned}$$

Finally we reintroduce the explicit t-dependence to write the N–soliton solution as

$$u(x,t) = -2\frac{\partial^2}{\partial x^2}\ln(\det A(x)) \qquad \text{where} \qquad A_{nm}(x) = \delta_{nm} + \frac{\beta_n e^{-(\chi_n+\chi_m)x}}{\chi_n + \chi_m}. \tag{2.16}$$

2.3.3 Two-soliton asymptotics

Let us analyse a two-soliton solution with $\chi_1 > \chi_2$ in more detail. Set

$$\tau_k = \chi_k x - 4\chi_k^3 t, \qquad k = 1, 2$$

and consider the determinant

$$\det A = \Big(1 + \frac{\beta_1(0)}{2\chi_1}e^{-2\tau_1}\Big)\Big(1 + \frac{\beta_2(0)}{2\chi_2}e^{-2\tau_2}\Big) - \frac{\beta_1(0)\beta_2(0)}{(\chi_1+\chi_2)^2}e^{-2(\tau_1+\tau_2)}.$$

We first analyse the case $t \to -\infty$. In the limit $x \to -\infty$ we have $\det A \sim e^{-2(\tau_1+\tau_2)}$ so

$$\log(\det A) \sim \text{const} - 2(\tau_1 + \tau_2)$$

and $u \sim 0$ which we already knew. Now move along the x axis and consider the leading term in $\det A$ when $\tau_1 = 0$ and then when $\tau_2 = 0$. We first reach the point $\tau_1 = 0$ or

$$x = 4\chi_1^2 t.$$

In the neighbourhood of this point $\tau_2 = 4t\chi_2(\chi_1^2 - \chi_2^2) \ll 0$ and

$$\det A \sim \frac{\beta_2(0)}{2\chi_2}e^{-2\tau_2}\Big(1 + \frac{\beta_1(0)}{2\chi_1}\Big(\frac{\chi_1-\chi_2}{\chi_1+\chi_2}\Big)^2 e^{-2\tau_1}\Big).$$

Differentiating the logarithm of $\det A$ yields

$$u \sim -2\frac{\partial^2}{\partial x^2}\Big(1+\frac{\beta_1(0)}{2\chi_1}\Big(\frac{\chi_1-\chi_2}{\chi_1+\chi_2}\Big)^2 e^{-2\chi_1(x-4\chi_1^2t^2)}\Big)$$

which looks like a one soliton solution with a phase

$$(\phi_1)_- = \frac{1}{2\chi_1}\log\Big(\frac{\beta_1(0)}{2\chi_1}\Big(\frac{\chi_1-\chi_2}{\chi_1+\chi_2}\Big)^2\Big).$$

We now move along the x axis until we reach $\tau_2 = 0$. Repeating the above analysis shows that now $\tau_1 = 4\chi_1(\chi_2^2-\chi_1^2)t \gg 0$ and around the point $x = 4\chi_2^2 t$ we have

$$\det A \sim 1 + \frac{\beta_2(0)}{2\chi_2}e^{-2\tau_2}.$$

Therefore the function u looks like a one–soliton solution with a phase

$$(\phi_2)_- = \frac{1}{2\chi_2}\log\Big(\frac{\beta_2(0)}{2\chi_2}\Big).$$

As t approaches 0 the two solitons coalesce and the exact behaviour depends on the ratio χ_1/χ_2.

We perform analogous analysis as $t \to \infty$. If $x \to \infty$ then $\det A \sim 1$ and $u \sim 0$. We move along the x axis to the left until we reach $\tau_1 = 0$ where $\tau_2 \gg 0$ and the profile of u is given by one–soliton with the phase

$$(\phi_1)_+ = \frac{1}{2\chi_1}\log\Big(\frac{\beta_1(0)}{2\chi_1}\Big).$$

Then we reach the point $\tau_2 = 0, \tau_1 \ll 0$ where there is a single soliton with the phase

$$(\phi_2)_+ = \frac{1}{2\chi_2}\log\Big(\frac{\beta_2(0)}{2\chi_1}\Big(\frac{\chi_1-\chi_2}{\chi_1+\chi_2}\Big)^2\Big).$$

Thus the larger soliton has overtaken the smaller one. This asymptotic analysis shows that the solitons have preserved their shape but their phases have changed

$$\begin{aligned}\Delta\phi_1 &= (\phi_1)_+ - (\phi_1)_- = -\frac{1}{\chi_1}\log\frac{\chi_1-\chi_2}{\chi_1+\chi_2},\\ \Delta\phi_2 &= (\phi_2)_+ - (\phi_2)_- = -\frac{1}{\chi_2}\log\frac{\chi_1-\chi_2}{\chi_1+\chi_2}.\end{aligned}$$

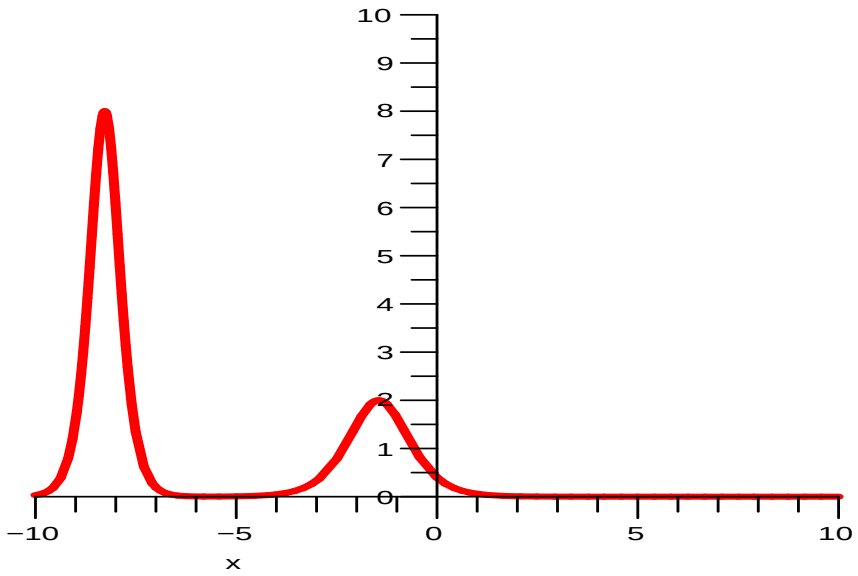
The only result of the interaction can be measured by

$$-\log\frac{\chi_1-\chi_2}{\chi_1+\chi_2}$$

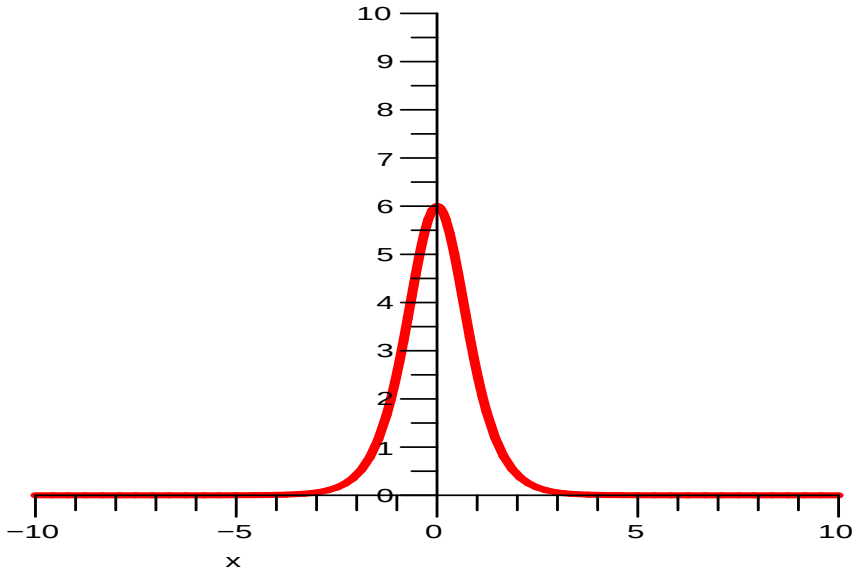
which is large if the difference between the velocities χ_1 and χ_2 is small.

The figures show the two–soliton solution at $t = -1, t = 0$ and $t = 1$ (for the chosen parameters $t = -1$ is considered to be a large negative time when the two solitons are separated). It should be interpreted as a passing collision of fast and slow soliton. The larger, faster soliton has amplitude 8, and the slower, smaller soliton has amplitude 2. Its velocity is one half of that of the fast soliton. The solitons are separated at $t = -1$. At $t = 0$ the collision takes place. The wave amplitude becomes smaller than the sum of the two waves. At $t = 1$ the larger soliton has overtaken the smaller one. The amplitudes and shapes have not changed.

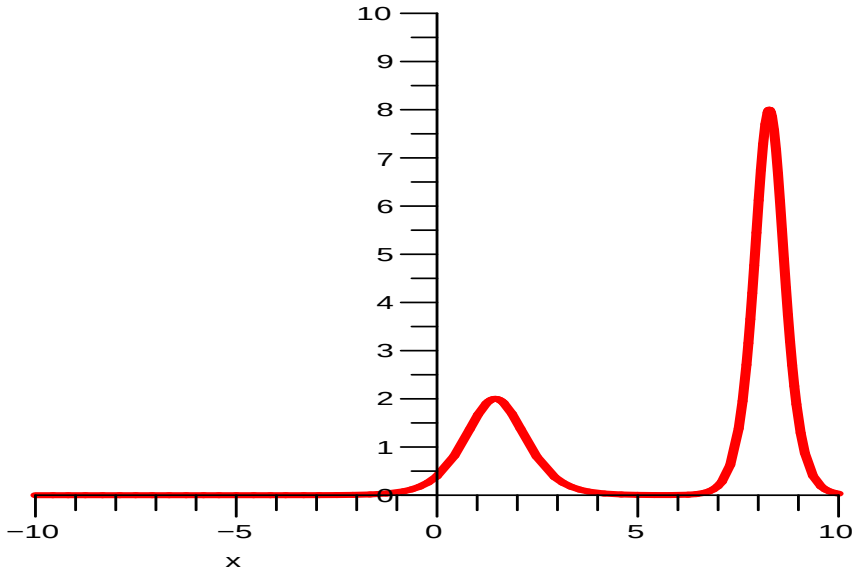
2-soliton solution at $t = -1$.



2-soliton solution at $t = 0$. The total amplitude is smaller than the sum of the two amplitudes.



2-soliton solution at $t = 1$. Amplitudes and shapes preserved by the collision.



This picture generalises to $N > 2$. The general solution (2.16) asymptotically represents N separate solitons ordered accordingly to their speed. The tallest (and therefore fastest) soliton is at the front, followed by the second tallest etc. At $t = 0$ the 'interaction' takes place and then the individual solitons re-emerge in the opposite order as $t \to \infty$. The total phase–shift is the sum of pairwise phase–shifts [15].

The number of the discrete eigenvalues N in the Schrödinger operator is equal to the number of solitons at $t \to \pm\infty$. This number is of course encoded in the initial conditions. To see it consider

$$u(x,0) = u_0(x) = -\frac{N(N+1)}{\cosh^2(x)}, \quad N \in \mathbb{Z}^+.$$

Substituting $\xi = \tanh(x) \in (-1,1)$ in the Schrödinger equation

$$-\frac{d^2 f}{dx^2} + u_0(x) f = k^2 f$$

yields the associated Legendre equation

$$\frac{d}{d\xi}\Big((1-\xi^2)\frac{df}{d\xi}\Big) + \Big(N(N+1) + \frac{k^2}{1-\xi^2}\Big) f = 0.$$

Analysis of the power series solution shows that the square integrable solutions exist if $k^2 = -\chi^2$ and $\chi = 1, 2, \ldots, N$. Therefore $F(x)$ in the GLM equation is given by

$$F(x) = \sum_{n=1}^{N} \beta_n e^{-\chi_n x},$$

and the earlier calculation applies leading to a particular case of the N–soliton solution (2.16). See the more complete discussion of this point in [5].

Chapter 3

Hamiltonian formalism and the zero curvature representation

3.1 First integrals

We shall make contact with the Definition 1.2.1 of finite–dimensional integrable systems and show that KdV has infinitely many first integrals. Rewrite the expression (2.6)

$$\phi(x,k) = \begin{cases} e^{-ikx}, & \text{for } x \to -\infty \\ a(k,t)e^{-ikx} + b(k,t)e^{ikx}, & \text{for } x \to \infty, \end{cases}$$

when the time dependence of the scattering data has been determined using the KdV equation. The formula (2.15) gives

$$\frac{\partial}{\partial t}a(k,t) = 0, \quad \forall k$$

so the scattering data gives infinitely many first integrals provided that they are non–trivial and independent. We aim to express these first integrals in the form

$$I[u] = \int_{\mathbb{R}} P(u, u_x, u_{xx}, \ldots)dx$$

where P is a polynomial in u and its derivatives.

Set

$$\phi(x,t,k) = e^{-ikx + \int_{-\infty}^{x} S(y,t,k)dy}.$$

For large x the formula (2.6) gives

$$e^{ikx}\phi \cong a(k) + b(k,t)e^{2ikx}.$$

If we assume that k is in the upper half plane $Im(k) > 0$ the second term on the RHS goes to 0 as $x \to \infty$. Thus

$$\begin{aligned} a(k) &= \lim_{x\to\infty} e^{ikx}\phi(x,t,k) = \lim_{x\to\infty} e^{\int_{-\infty}^{x} S(y,t,k)dy} \\ &= e^{\int_{-\infty}^{\infty} S(y,t,k)dy}, \end{aligned} \tag{3.1}$$

where the above formula also holds in the limit $Im(k) \to 0$ because of the real analyticity. Now we shall use the Schrödinger equation with t regarded as a parameter

$$-\frac{d^2\phi}{dx^2} + u\phi = k^2\phi$$

to find an equation for S. Substituting

$$\frac{d\phi}{dx} = (-ik + S(x,k))\phi, \quad \frac{d^2\phi}{dx^2} = \frac{dS}{dx}\phi + (-ik + S(x,k))^2\phi$$

gives the Riccati type equation

$$\frac{dS}{dx} - 2ikS + S^2 = u, \tag{3.2}$$

(we stress that both S and u depend on x as well as t). Look for solutions of the form

$$S = \sum_{n=1}^{\infty} \frac{S_n(x,t)}{(2ik)^n}.$$

Substituting this to (3.2) yields a recursion relation

$$S_1(x,t) = -u(x,t), \quad S_{n+1} = \frac{dS_n}{dx} + \sum_{m=1}^{n-1} S_m S_{n-m} \tag{3.3}$$

which can be solved for the first few terms

$$S_2 = -\frac{\partial u}{\partial x}, \quad S_3 = -\frac{\partial^2 u}{\partial x^2} + u^2, \quad S_4 = -\frac{\partial^3 u}{\partial x^3} + 2\frac{\partial}{\partial x}u^2,$$

$$S_5 = -\frac{\partial^4 u}{\partial x^4} + 2\frac{\partial^2}{\partial x^2}u^2 + \left(\frac{\partial u}{\partial x}\right)^2 + 2\frac{\partial^2 u}{\partial x^2}u - 2u^3.$$

Now using the time independence (2.15) of $a(k)$ for all k and combining it with (3.1) implies that

$$\int_{\mathbb{R}} S_n(x,t)dx$$

are first integrals of the KdV equation. Not all of these integrals are non-trivial. For example S_2 and S_4 given above are total x derivatives so they integrate to 0 (using the boundary conditions for u). The same is true for all even terms S_{2n}. To see it set

$$S = S_R + iS_I$$

where S_R, S_I are real valued functions and substitute this to (3.2). Taking the imaginary part gives

$$\frac{dS_I}{dx} + 2S_R S_I - 2kS_R = 0$$

which integrates to

$$S_R = -\frac{1}{2}\frac{d}{dx}\log(S_I - k).$$

The even terms

$$\frac{S_{2n}(x)}{(2ik)^{2n}}, \qquad n = 1, 2, \dots$$

in the expansion of a are real. Comparing this with the expansion of S_R in k shows that S_{2n} are all total derivatives and therefore

$$\int_{\mathbb{R}} S_{2n} dx = 0.$$

Let us now concentrate on the remaining non–trivial first integrals. Set

$$I_{n-1}[u] = \frac{1}{2} \int_{\mathbb{R}} S_{2n+1}(x,t) dx, \qquad n = 0, 1, 2, \cdots . \tag{3.4}$$

Our analysis shows

$$\frac{dI_n}{dt} = 0.$$

The first of these is just the integral of u itself. The next two are known as momentum and energy respectively

$$I_0 = \frac{1}{2} \int_{\mathbb{R}} u^2 dx, \quad I_1 = -\frac{1}{2} \int_{\mathbb{R}} (u_x^2 + 2u^3) dx,$$

where in the last integral we have isolated the total derivative in

$$S_5 = -\frac{\partial^4}{\partial x^4} u + 2\frac{\partial^2}{\partial x^2} u^2 + 2\frac{\partial}{\partial x}\Big(u\frac{\partial u}{\partial x}\Big) - \Big(\frac{\partial u}{\partial x}\Big)^2 - 2u^3$$

and eliminated it using the integration by parts and boundary conditions. These two first integrals are associated, via Noether's theorem, with the translational invariance of KdV: if $u(x,t)$ is a solution then $u(x+x_0,t)$ and $u(x,t+t_0)$ are also solutions. The systematic way of constructing such symmetries will be presented in Chapter 4.

3.2 Hamiltonian formalism

We can now cast the KdV in the Hamiltonian form with the Hamiltonian functional given by the energy integral $H[u] = -I_1[u]$. First calculate.

$$\frac{\delta I_1[u]}{\delta u(x)} = -3u^2 + u_{xx}, \quad \frac{\partial}{\partial x}\frac{\delta I_1[u]}{\delta u(x)} = -6uu_x + u_{xxx}.$$

Recall that the Hamilton canonical equations for PDEs take the form (1.11)

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\frac{\delta H[u]}{\delta u(x)}.$$

Therefore

$$\frac{\partial u}{\partial t} = -\frac{\partial}{\partial x}\frac{\delta I_1[u]}{\delta u(x)}, \tag{3.5}$$

is the KdV equation. With some more work (see [15]) it can be shown that

$$\{I_m, I_n\} = 0$$

where the Poisson bracket is given by (1.10) so that KdV is indeed integrable in the Arnold–Liouville sense. For example

$$\begin{aligned}
\{I_n, I_1\} &= \int_{\mathbb{R}} \frac{\delta I_n}{\delta u(x)} \frac{\partial}{\partial x} \frac{\delta I_1}{\delta u(x)} dx = -\int_{\mathbb{R}} \frac{\delta I_n}{\delta u(x)} u_t \, dx \\
&= -\frac{1}{2} \int_{\mathbb{R}} \sum_{k=0}^{2n} (-1)^k \Big(\Big(\frac{\partial}{\partial x} \Big)^k \frac{\partial S_{2n+1}}{\partial u^{(k)}} \Big) u_t \, dx \\
&= \frac{1}{2} \int_{\mathbb{R}} \sum_{k=0}^{2n} \frac{\partial S_{2n+1}}{\partial u^{(k)}} \frac{\partial}{\partial t} u^{(k)} \, dx \\
&= \frac{d}{dt} I_n[u] = 0
\end{aligned}$$

where we used integration by parts and the boundary conditions.

3.2.1 Bi–Hamiltonian systems

Most systems integrable by the inverse scattering transform are Hamiltonian in two distinct ways. This means that for a given evolution equation $u_t = F(u, u_x, \ldots)$ there exist two Poisson structures $\mathcal{D}$ and $\mathcal{E}$ and two functionals $H_0[u]$ and $H_1[u]$ such that

$$\frac{\partial u}{\partial t} = \mathcal{D} \frac{\delta H_1}{\delta u(x)} = \mathcal{E} \frac{\delta H_0}{\delta u(x)}. \tag{3.6}$$

One of these Poisson structures can be put in a form $\mathcal{D} = \partial/\partial x$ and corresponds to the standard Poisson bracket (1.10), but the second structure $\mathcal{E}$ gives a new Poisson bracket.

In the finite–dimensional context discussed in Section 1.2 this would correspond to having two skew-symmetric matrices ω, Ω which satisfy the Jacobi Identity. The Darboux theorem implies the existence of a local coordinate system (p, q) in which one of these, say ω, is a constant skew-symmetric matrix. The matrix components of second structure Ω will however be non–constant functions of (p, q). Using (1.9) we write the bi–Hamiltonian condition as

$$\omega^{ab} \frac{\partial H_1}{\partial \xi^a} = \Omega^{ab} \frac{\partial H_0}{\partial \xi^a},$$

where $\xi^a, a = 1, \ldots, 2n$ are local coordinates on the phase space M, and H_0, H_1 are two distinct functions on M. The matrix valued function

$$R_a{}^c = \Omega^{bc} (\omega^{-1})_{ab}$$

is called a recursion operator. It should be thought of as an endomorphism $R = \Omega \circ \omega^{-1}$ acting on the tangent space T_pM, where $p \in M$. This endomorphism smoothly depends on a point p. The existence of such recursion operator is, under certain technical assumptions, equivalent

to Arnold–Liouville integrability in a sense of Theorem 1.2.2. This is because given one first integral H_0 the remaining $(n-1)$ integrals $H_1, \dots, H_{n-1}$ can be constructed recursively by

$$\omega^{ab}\frac{\partial H_i}{\partial \xi^a} = R^i\Big(\omega^{ab}\frac{\partial H_0}{\partial \xi^a}\Big) \qquad i = 1, 2, \dots, n-1.$$

The extension of this formalism to the infinite dimensional setting provides a practical way of constructing first integrals. In the case of KdV the first Hamiltonian formulation (3.5) has $\mathcal{D} = \partial/\partial x$ and

$$H_1[u] = \int_{\mathbb{R}} \Big(\frac{1}{2}{u_x}^2 + u^3\Big)dx.$$

The second formulation can be obtained taking

$$H_0[u] = \frac{1}{2}\int_{\mathbb{R}} u^2 dx, \qquad \mathcal{E} = -\partial_x^3 + 4u\partial_x + 2u_x.$$

In general it is required that a pencil of Poisson structures $\mathcal{D} + c\mathcal{E}$ is also a Poisson structure (i.e. satisfies the Jacobi identity) for any constant $c \in \mathbb{R}$. If this condition is satisfied, the bi–Hamiltonian formulation gives an effective way to construct first integrals. The following result is proved in the book of Olver [16]

Theorem 3.2.1 *Let* (3.6) *be a bi–Hamiltonian system, such that the Poisson structure* $\mathcal{D}$ *is non–degenerate*[1]*, and let*

$$R = \mathcal{E} \circ \mathcal{D}^{-1}$$

be the corresponding recursion operator. Assume that

$$R^n\Big(\mathcal{D}\frac{\delta H_0}{\delta u(x)}\Big)$$

lies in the image of $\mathcal{D}$ *for each* $n = 1, 2, \dots$ *. Then there exists conserved functionals*

$$H_1[u], H_2[u], \dots$$

which are in involution, i. e.

$$\{H_m, H_n\} := \int_{\mathbb{R}} \frac{\delta H_m}{\delta u(x)}\mathcal{D}\frac{\delta H_n}{\delta u(x)}dx = 0.$$

The conserved functionals $H_n[u]$ are constructed recursively from H_0 by

$$\mathcal{D}\frac{\delta H_n}{\delta u(x)} = R^n\Big(\mathcal{D}\frac{\delta H_0}{\delta u(x)}\Big), \qquad n = 1, 2, \dots \ . \tag{3.7}$$

In the case of the KdV equation the recursion operator is

$$R = -\partial_x^2 + 4u + 2u_x{\partial_x}^{-1}, \tag{3.8}$$

where ${\partial_x}^{-1}$ is formally defined as integration with respect to x, and formula (3.7) gives an alternative way of constructing the first integrals (3.4).

[1]A differential operator $\mathcal{D}$ is degenerate is there exists a non–zero differential operator $\hat{\mathcal{D}}$ such that the operator $\hat{\mathcal{D}} \circ \mathcal{D}$ is identically zero.

3.3 Zero curvature representation

We shall discuss a more geometric form of the Lax representation where integrable systems arise as compatibility conditions of overdetermined system of matrix PDEs. Let $U(\lambda)$ and $V(\lambda)$ be matrix valued functions of (ρ, τ) depending on the auxiliary variable λ called the spectral parameter. Consider a system of linear PDEs

$$\frac{\partial}{\partial \rho} v = U(\lambda) v, \quad \frac{\partial}{\partial \tau} v = V(\lambda) v \tag{3.9}$$

where v is a column vector whose components depend on (ρ, τ, λ). This is an overdetermined system as there are twice as many equations as unknowns. The compatibility conditions can be obtained by cross–differentiating and commuting the partial derivatives

$$\frac{\partial}{\partial \tau}\frac{\partial}{\partial \rho} v - \frac{\partial}{\partial \rho}\frac{\partial}{\partial \tau} v = 0$$

which gives

$$\frac{\partial}{\partial \tau}(U(\lambda)v) - \frac{\partial}{\partial \rho}(V(\lambda)v) = \Big(\frac{\partial}{\partial \tau}U(\lambda) - \frac{\partial}{\partial \rho}V(\lambda) + [U(\lambda), V(\lambda)]\Big)v = 0.$$

This has to hold for all characteristic initial data so the linear system (3.9) is consistent iff the nonlinear equation

$$\frac{\partial}{\partial \tau}U(\lambda) - \frac{\partial}{\partial \rho}V(\lambda) + [U(\lambda), V(\lambda)] = 0 \tag{3.10}$$

holds. The whole scheme is known as the zero curvature representation[2]. Most non-linear integrable equation admit a zero–curvature representation analogous to (3.10).

- **Example.** If

$$U = \frac{i}{2}\begin{pmatrix} 2\lambda & \phi_\rho \\ \phi_\rho & -2\lambda \end{pmatrix}, \qquad V = \frac{1}{4i\lambda}\begin{pmatrix} \cos(\phi) & -i\sin(\phi) \\ i\sin(\phi) & -\cos(\phi) \end{pmatrix} \tag{3.11}$$

  where $\phi = \phi(\rho, \tau)$ then (3.10) is equivalent to the Sine–Gordon equation

$$\phi_{\rho\tau} = \sin(\phi).$$

- **Example.** Consider the zero curvature representation with

$$\begin{aligned} U &= i\lambda\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + i\begin{pmatrix} 0 & \overline{\phi} \\ \phi & 0 \end{pmatrix}, \\ V &= 2i\lambda^2\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + 2i\lambda\begin{pmatrix} 0 & \overline{\phi} \\ \phi & 0 \end{pmatrix} + \begin{pmatrix} 0 & \overline{\phi}_\rho \\ -\phi_\rho & 0 \end{pmatrix} - i\begin{pmatrix} |\phi|^2 & 0 \\ 0 & -|\phi|^2 \end{pmatrix}. \end{aligned} \tag{3.12}$$

  The condition (3.10) holds if the complex valued function $\phi = \phi(\tau, \rho)$ satisfies the non-linear Schrödinger equation

$$i\phi_\tau + \phi_{\rho\rho} + 2|\phi|^2\phi = 0.$$

  This is another famous soliton equation which can be solved by inverse scattering transform.

[2]The terminology, due to Zaharov and Shabat, comes from differential geometry where (3.10) means that the curvature of a connection $Ud\rho + Vd\tau$ is zero.

There is a freedom in the matrices $U(\lambda), V(\lambda)$ known as the gauge invariance. Let $g = g(\tau, \rho)$ be an arbitrary invertible matrix. The transformation

$$\widetilde{U} = gUg^{-1} + \frac{\partial g}{\partial \rho} g^{-1}, \quad \widetilde{V} = gVg^{-1} + \frac{\partial g}{\partial \tau} g^{-1} \tag{3.13}$$

maps solutions to the zero curvature equation into new solutions: if the matrices (U, V) satisfy (3.10) then so do the matrices $(\widetilde{U}, \widetilde{V})$. To see it assume that $v(\rho, \tau, \lambda)$ is a solution to the linear system (3.9), and demand that $\tilde{v} = g(\rho, \tau)v$ be another solution for some $(\widetilde{U}, \widetilde{V})$. This leads to the gauge transformation (3.13).

One can develop a version of inverse scattering transform which recovers $U(\lambda)$ and $V(\lambda)$ from a linear scattering problem (3.9). The representation (3.10) can also be an effective direct method of finding solutions if we know n linearly independent solutions $v_1, \ldots, v_n$ to the linear system (3.9) at the first place. Let $\Phi(\rho, \tau, \lambda)$ be a fundamental matrix solution to (3.9). The columns of Φ are the n linearly independent solutions $v_1, \ldots, v_n$. Then (3.9) holds with v replaced by Φ and we can write

$$U(\lambda) = \frac{\partial \Phi}{\partial \rho} \Phi^{-1}, \quad V(\lambda) = \frac{\partial \Phi}{\partial \tau} \Phi^{-1}.$$

In practice one assumes a simple λ dependence in Φ, characterised by a finite number of poles with given multiplicities. One general scheme of solving (3.10), known as the dressing method, is based on the Riemann–Hilbert problem which we shall review next.

3.3.1 The Riemann–Hilbert problem

Let $\lambda \in \overline{\mathbb{C}} = \mathbb{C} + \{\infty\}$ and let Γ be a closed contour in the extended complex plane. In particular we can consider Γ to be a real line $-\infty < \lambda < \infty$ regarded as a circle in $\overline{\mathbb{C}}$ passing through ∞. Let $G = G(\lambda)$ be a matrix valued function on the contour Γ. The Riemann–Hilbert problem is to construct two matrix valued functions $G_+(\lambda)$ and $G_-(\lambda)$ holomorphic respectively inside and outside the contour such that on Γ

$$G(\lambda) = G_+(\lambda) G_-(\lambda). \tag{3.14}$$

In the case when Γ is the real axis G_+ is required to be holomorphic in the upper half–plane and G_- is required to be holomorphic in the lower half–plane. If (G_+, G_-) is a solution of the Riemann–Hilbert problem, then

$$\widetilde{G}_+ = G_+ \, g^{-1}, \qquad \widetilde{G}_- = g \, G_-$$

will also be a solution for any constant invertible matrix g. This ambiguity can be avoided by fixing a values of G_+ or G_- at some point in their domain, for example by setting $G_-(\infty) = I$. If the matrices $G_\pm$ are everywhere invertible then this normalisation guarantees that the solution to (3.14) is unique.

Solving a Riemann–Hilbert problem comes down to an integral equation. Choose a normalisation $G_+(\lambda_0) = I$ and set $G_-(\lambda_0) = g$ for some $\lambda_0 \in \mathbb{C}$. Assume that the Riemann–Hilbert problem has a solution of the form

$$(G_+)^{-1} = h + \oint_\Gamma \frac{\Phi(\xi)}{\xi - \lambda} d\xi$$

inside the contour Γ, and

$$G_- = h + \oint_\Gamma \frac{\Phi(\xi)}{\xi - \lambda} d\xi$$

outside Γ, where h is determined by the normalisation condition to be

$$h = g - \oint_\Gamma \frac{\Phi(\xi)}{\xi - \lambda_0} d\xi.$$

The Plemelj formula [3] can be used to determine $(G_+)^{-1}$ and G_- on the contour: If $\lambda \in \Gamma$ then

$$\begin{aligned} (G_+)^{-1}(\lambda) &= h + \oint_\Gamma \frac{\Phi(\xi)}{\xi - \lambda} d\xi + \pi i \Phi(\lambda) \\ G_-(\lambda) &= h + \oint_\Gamma \frac{\Phi(\xi)}{\xi - \lambda} d\xi - \pi i \Phi(\lambda), \end{aligned}$$

where the integrals are assumed to be defined by the principal value. Substituting these expressions to (3.14) yields the integral equation for $\Phi = \Phi(\lambda)$. If the normalisation is canonical, so that $h = g = 1$, the equation is

$$\frac{1}{\pi i}\Big(\int_\Gamma \frac{\Phi(\xi)}{\xi - \lambda} d\xi + I \Big) + \Phi(\lambda)(G + I)(G - I)^{-1} = 0.$$

The simplest case is the scalar Riemann–Hilbert problem where G, G_+, G_- are ordinary functions. In this case the solution can be written down explicitly as

$$\begin{aligned} G_+ &= \exp\Big(-\Big(\frac{1}{2\pi i} \int_{-\infty}^{\infty} \frac{\log G(\xi)}{\xi - \lambda} d\xi \Big) \Big), \quad \mathrm{Im}(\lambda) > 0 \\ G_- &= \exp\Big(\frac{1}{2\pi i} \int_{-\infty}^{\infty} \frac{\log G(\xi)}{\xi - \lambda} d\xi \Big), \quad \mathrm{Im}(\lambda) < 0. \end{aligned}$$

This is verified by taking a logarithm of (3.14)

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

and applying the Cauchy integral formulae.

3.3.2 Dressing method

We shall assume that the matrices (U, V) in the zero curvature representation (3.10) have rational dependence on the spectral parameter λ. The complex analytic data for each of these matrices consist of a set of poles (including poles at $\lambda = \infty$) with the corresponding multiplicities. Define the divisors to be the sets

$$S_U = \{\alpha_i, n_i, n_\infty\}, \quad S_V = \{\beta_j, m_j, m_\infty\}, \qquad i = 1, \dots, n, \qquad j = 1, \dots, m$$

so that

$$\begin{aligned} U(\rho,\tau,\lambda) &= \sum_{i=1}^{n}\sum_{r=1}^{n_i}\frac{U_{i,r}(\rho,\tau)}{(\lambda-\alpha_i)^r}+\sum_{k=0}^{n_\infty}\lambda^k U_k(\rho,\tau) \\ V(\rho,\tau,\lambda) &= \sum_{j=1}^{m}\sum_{r=1}^{m_i}\frac{V_{j,r}(\rho,\tau)}{(\lambda-\beta_j)^r}+\sum_{k=0}^{m_\infty}\lambda^k V_k(\rho,\tau). \end{aligned} \tag{3.15}$$

The zero curvature condition (3.10) is a system of non–linear PDEs on coefficients

$$U_{i,r}, \quad U_k, \quad V_{j,r}, \quad V_k$$

of U and V. Consider a trivial solution to (3.10)

$$U = U_0(\rho,\lambda), \qquad V = V_0(\tau,\lambda)$$

where U_0, V_0 are any two commuting matrices with divisors S_U and S_V respectively.

Let Γ be a contour in the extended complex plane which does not contain any points from $S_U \cup S_V$, and let $G(\lambda)$ be a smooth matrix–valued function defined on Γ. The dressing method [20] is a way of constructing a non–trivial solution with analytic structure specified by divisors S_U, S_V out of the data

$$(U_0,\, V_0,\, \Gamma,\, G).$$

It consists of the following steps

1. Find a fundamental matrix solution to a linear system of equations

$$\frac{\partial}{\partial\rho}\Psi_0 = U_0(\lambda)\Psi_0, \quad \frac{\partial}{\partial\tau}\Psi_0 = V_0(\lambda)\Psi_0. \tag{3.16}$$

   This overdetermined system is compatible as U_0, V_0 satisfy (3.10).

2. Define a family of smooth functions $G(\rho,\tau.\lambda)$ parametrised by (ρ,τ) on Γ

$$G(\rho,\tau,\lambda) = \Psi_0(\rho,\tau,\lambda)G(\lambda){\Psi_0}^{-1}(\rho,\tau,\lambda). \tag{3.17}$$

   This family admits a factorisation

$$G(\rho,\tau,\lambda) = G_+(\rho,\tau,\lambda)G_-(\rho,\tau,\lambda) \tag{3.18}$$

   where $G_+(\rho,\tau,\lambda)$ and $G_-(\rho,\tau,\lambda)$ are solutions to the Riemann–Hilbert problem described in the last subsection, and are holomorphic respectively inside and outside the contour Γ.

3. Differentiate (3.18) with respect to ρ and use (3.16) and (3.17). This yields

$$\frac{\partial G_+}{\partial\rho}G_- + G_+\frac{\partial G_-}{\partial\rho} = U_0G_+G_- - G_+G_-U_0.$$

   Therefore we can define

$$U(\rho,\tau,\lambda) := \Big(\frac{\partial G_-}{\partial\rho} + G_-U_0\Big){G_-}^{-1} = -{G_+}^{-1}\Big(\frac{\partial G_+}{\partial\rho} - U_0G_+\Big)$$

which is holomorphic in $\overline{\mathbb{C}}/S_U$. The Liouville Theorem (stating that every bounded holomorphic function is constant) applied to the extended complex plane implies that $U(\rho,\tau,\lambda)$ is rational in λ and has the same pole structure as U_0.

Analogous argument leads to

$$V(\rho,\tau,\lambda) := \Big(\frac{\partial G_-}{\partial\tau} + G_- V_0\Big) {G_-}^{-1} = -{G_+}^{-1}\Big(\frac{\partial G_+}{\partial\tau} - V_0 G_+\Big)$$

which has the same pole structure as V_0.

4. Define two matrix valued functions

$$\Psi_+ = {G_+}^{-1}\Psi_0, \qquad \Psi_- = {G_-}^{-1}\Psi_0.$$

Equations (3.16) and the definitions of (U,V) imply that these matrices both satisfy the overdetermined system

$$\frac{\partial}{\partial\rho}\Psi_\pm = U(\lambda)\Psi_\pm, \quad \frac{\partial}{\partial\tau}\Psi_\pm = V(\lambda)\Psi_\pm.$$

We can therefore deduce that $U(\rho,\tau,\lambda)$ and $V(\rho,\tau,\lambda)$ are of the form (3.15) and satisfy the zero curvature relation (3.10).

This procedure is called 'dressing' as the bare, trivial solution (U_0,V_0) has been dressed by an application of a Riemann–Hilbert problem to a non-trivial (U,V). Now, given another matrix valued function $G = G'(\lambda)$ on the contour we could repeat the whole procedure and apply it to (U,V) instead of (U_0,V_0). This would lead to another solution (U',V') with the same pole structure. Thus dressing transformations act on the space of solutions to (3.10) and form a group. If $G = G_+G_-$ and $G' = {G'}_+G'-$ then

$$(G\circ G') = G_+{G'}_+{G'}_-G_-.$$

The solution to the Riemann–Hilbert problem (3.18) is not unique. If $G_\pm$ give a factorisation of $G(\rho,\tau,\lambda)$ then so do

$$\widetilde{G}_+ = G_+ g^{-1}, \qquad \widetilde{G}_- = gG_-$$

where $g = g(\rho,\tau)$ is a matrix valued function. The corresponding solutions $(\widetilde{U},\widetilde{V})$ are related to (U,V) by the gauge transformation (3.13). Fixing the gauge is therefore equivalent to fixing the value of G_+ or G_- at one point of the extended complex plane, say $\lambda=\infty$. This leads to a unique solution of the Riemann–Hilbert problem with $G_\pm(\infty) = G(\infty) = I$.

The dressing method leads to a general form of U and V with prescribed singularities, but more work is required to make contact with specific integrable models when additional algebraic constraints need to be imposed on U and V. For example in the Sine–Gordon case (3.11) the matrices are anti-Hermitian. The anti-Hermticity condition gives certain constraints on the contour Γ and the function G. Only if these constraints hold, the matrices resulting from the dressing procedure will be given (in some gauge) in terms of the solution to the Sine–Gordon equation.

3.3.3 From Lax representation to zero curvature

The zero curvature representation (3.10) is more general than the scalar Lax representation but there is a connection between the two. First similarity is that the Lax equation (2.12) also arises as a compatibility condition for two overdetermined PDEs. To see it take f to be an eigenfunction of L with a simple eigenvalue $E = \lambda$ and consider the relation (2.13) which follows from the Lax equations. If $E = \lambda$ is a simple eigenvalue then

$$\frac{\partial f}{\partial t} + Af = C(t)f$$

for some function C which depends on t but not on x. Therefore one can use an integrating factor to find a function $\hat{f} = \hat{f}(x, t, \lambda)$ such that

$$L\hat{f} = \lambda\hat{f}, \qquad \frac{\partial \hat{f}}{\partial t} + A\hat{f} = 0, \tag{3.19}$$

where L is the Schrödinger operator and A is some differential operator (for example given by (2.14)). Therefore the Lax relation

$$\dot{L} = [L, A]$$

is a compatibility of an overdetermined system (3.19).

Consider a general scalar Lax pair

$$\begin{aligned} L &= \frac{\partial^n}{\partial x^n} + u_{n-1}(x,t)\frac{\partial^{n-1}}{\partial x^{n-1}} + \ldots + u_1(x,t)\frac{\partial}{\partial x} + u_0(x,t) \\ A &= \frac{\partial^m}{\partial x^m} + v_{m-1}(x,t)\frac{\partial^{m-1}}{\partial x^{m-1}} + \ldots + v_1(x,t)\frac{\partial}{\partial x} + v_0(x,t) \end{aligned}$$

given by differential operators with coefficients depending on (x, t). The Lax equations

$$\dot{L} = [L, A]$$

(in general there will be more than one) are non–linear PDEs for the coefficients

$$(u_0, \ldots, u_{n-1}, v_0, \ldots, v_{m-1}).$$

The linear nth order scalar PDE

$$L\hat{f} = \lambda\hat{f} \tag{3.20}$$

is equivalent to the first order matrix PDE

$$\frac{\partial F}{\partial x} = U_L F$$

where $U_L = U_L(x, t, \lambda)$ is an n by n matrix

$$U_L = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 0 & 1 \\ \lambda - u_0 & -u_1 & -u_2 & \ldots & -u_{n-2} & -u_{n-1} \end{pmatrix}$$

and F is a column vector

$$F = (f_0, f_1, \ldots, f_{n-1})^T, \qquad \text{where} \quad f_k = \frac{\partial^k \hat{f}}{\partial x^k}.$$

Now consider the second equation in (3.19)

$$\frac{\partial \hat{f}}{\partial t} + A\hat{f} = 0$$

which is compatible with (3.20) if the Lax equations hold. We differentiate this equation with respect to x and use (3.20) to express $\partial_x^n \hat{f}$ in terms of λ and lower order derivatives. Repeating this process $(n-1)$ times gives an action of A on components of the vector F. We write it as

$$\frac{\partial F}{\partial t} = V_A F$$

using the method described above. This leads to a pair of first order linear matrix equations with the zero curvature compatibility conditions

$$\frac{\partial U_L}{\partial t} - \frac{\partial V_A}{\partial x} + [U_L, V_A] = 0.$$

These conditions hold if the operators (L, A) satisfy the Lax relations $\dot{L} = [L, A]$.

- **Example.** Let us apply this procedure to the KdV Lax pair (2.14). Set

  $$f_0 = \hat{f}(x,t,\lambda), \qquad f_1 = \partial_x \hat{f}(x,t,\lambda).$$

  The eigenvalue problem $L\hat{f} = \lambda \hat{f}$ gives

  $$(f_0)_x = f_1, \qquad (f_1)_x = (u - \lambda) f_0.$$

  The equation $\partial_t \hat{f} + A\hat{f} = 0$ gives

  $$\begin{aligned} (f_0)_t &= -4(f_0)_{xxx} + 6uf_1 + 3u_x f_0 \\ &= -u_x(f_0) + (2u + 4\lambda) f_1. \end{aligned}$$

  We differentiate this equation with respect to x and eliminate the second derivatives of $\hat{f}$ to get

  $$(f_1)_t = ((2u + 4\lambda)(u - \lambda) - u_{xx}) f_0 + u_x f_1.$$

  We now collect the equations in the matrix form $\partial_x F = U_L F, \partial_t F = V_A F$ where $F = (f_0, f_1)^T$ and

  $$U_L = \begin{pmatrix} 0 & 1 \\ u - \lambda & 0 \end{pmatrix}, \quad V_A = \begin{pmatrix} -u_x & 2u + 4\lambda \\ 2u^2 - u_{xx} + 2u\lambda - 4\lambda^2 & u_x \end{pmatrix}. \tag{3.21}$$

  We have therefore obtained a zero curvature representation for KdV.

3.4 Hierarchies and finite gap solutions.

We shall end our discussion of the KdV equation with a description of KdV hierarchy. Recall that KdV is a Hamiltonian system (3.5) with the Hamiltonian given by the first integral $-I_1[u]$. Now choose a (constant multiple of) a different first integral $I_n[u]$ as a Hamiltonian and consider the equation

$$\frac{\partial u}{\partial t_n} = (-1)^n \frac{\partial}{\partial x} \frac{\delta I_n[u]}{\delta u(x)} \tag{3.22}$$

for a function $u = u(x, t_n)$. This leads to an infinite set of equations known as higher KdVs. The first three equations are

$$\begin{aligned} u_{t_0} &= u_x, \\ u_{t_1} &= 6uu_x - u_{xxx}, \\ u_{t_2} &= 10uu_{xxx} - 20u_x u_{xx} - 30u^2 u_x - u_{xxxxx}. \end{aligned}$$

Each of these equations can be solved by inverse scattering method we have discussed, and the functionals $I_k, k = -1, 0, \cdots$ are first integrals regardless which one of them is chosen as a Hamiltonian. In the associated Lax representation L stays unchanged, but A is replaced by a differential operator of degree $(2n + 1)$. One can regard the higher KdVs as a system of overdetermined PDEs for

$$u = u(t_0 = x, t_1 = t, t_2, t_3, \cdots),$$

where we have identified t_0 with x using the first equation in (3.22).

This system is called a hierarchy and the coordinates $(t_2, t_3, \cdots)$ are known as higher times. The equations of the hierarchy are consistent as the flows generated by time translations commute

$$\begin{aligned} \frac{\partial}{\partial t_m}\frac{\partial}{\partial t_n}u - \frac{\partial}{\partial t_n}\frac{\partial}{\partial t_m}u &= (-1)^n \frac{\partial}{\partial t_m}\frac{\partial}{\partial x}\frac{\delta I_n[u]}{\delta u(x)} - (-1)^m \frac{\partial}{\partial t_n}\frac{\partial}{\partial x}\frac{\delta I_m[u]}{\delta u(x)} \\ &= \{u, \partial_m I_n - \partial_n I_m + \{I_m, I_n\}\} = 0 \end{aligned}$$

where we used the Jacobi identity and the fact that $I_n[u]$ Poisson commute.

The concept of the hierarchy leads to a beautiful method of finding solutions to KdV with periodic initial data, i.e.

$$u(x, 0) = u(x + X_0, 0)$$

for some period X_0. The method is based on the concept of stationary (i.e. time independent) solutions, albeit applied to a combination of the higher times.

Consider the first $(n+1)$ higher KdVs and take $(n+1)$ real constants $c_0, \ldots, c_n$. Therefore

$$\sum_{k=0}^{n} c_k \frac{\partial u}{\partial t_k} = \sum_{k=0}^{n} (-1)^k c_k \frac{\partial}{\partial x}\frac{\delta I_k[u]}{\delta u(x)}.$$

The stationary solutions correspond to u being independent on the combination of higher times on the LHS. This leads to an ODE

$$\sum_{k=0}^{n} (-1)^k c_k \frac{\delta I_k[u]}{\delta u(x)} = c_{n+1}, \qquad c_{n+1} = \text{const.} \tag{3.23}$$

The recursion relations (3.3) for S_ks imply that this ODE is of order $2n$. Its general solution depends on $2n$ constants of integration as well as $(n+1)$ parameters c as we can always divide (3.23) by $c_n \neq 0$. Altogether one has $3n+1$ parameters. The beauty of this method is that the ODE is integrable in the sense of Arnold–Liouville theorem and its solutions can be constructed by hyper–elliptic functions. The corresponding solutions to KdV are known as finite–gap solutions. Their description in terms of a spectral data is rather involved and uses Riemann surfaces and algebraic geometry - see Chapter 2 of [15].

We shall now present the construction of the first integrals to equation (3.23) (we stress that (3.23) is an ODE in x so the first integrals are functions of u and its derivatives which do not depend on x when (3.23) holds). The higher KdV equations (3.22) admit a zero curvature representation

$$\frac{\partial}{\partial t_n}U - \frac{\partial}{\partial x}V_n + [U, V_n] = 0$$

where

$$U = \begin{pmatrix} 0 & 1 \\ u-\lambda & 0 \end{pmatrix}$$

is the matrix obtained for KdV in section (3.3.3) and $V_n = V_n(x,t,\lambda)$ are traceless 2 by 2 matrices analogous to V_A which can be obtained using (3.22) and the recursion relations for (3.2). The components of V_n depend on (x,t) and are polynomials in λ of degree $(n+1)$. Now set

$$\Lambda = c_0 V_0 + \ldots + c_n V_n$$

where c_k are constants and consider solutions to

$$\frac{\partial}{\partial T}U - \frac{\partial}{\partial x}\Lambda + [U, \Lambda] = 0,$$

such that

$$\frac{\partial}{\partial T}U = 0, \qquad \text{where} \qquad \frac{\partial}{\partial T} = c_0\frac{\partial}{\partial t_0} + \ldots + c_n\frac{\partial}{\partial t_n}.$$

This gives rise to the ODE

$$\frac{d}{dx}\Lambda = [U, \Lambda]$$

which is the Lax representation of (3.23). This representation reveals existence of many first integrals for (3.23) . We have

$$\frac{d}{dx}\mathrm{Tr}(\Lambda^p) = \mathrm{Tr}(p[U,\Lambda]\Lambda^{p-1}) = p\mathrm{Tr}(-\Lambda U\Lambda^{p-1} + U\Lambda^p) = 0, \quad p = 2, 3, \ldots$$

by the cyclic property of trace. Therefore all the coefficients of the polynomials $\mathrm{Tr}(\Lambda(\lambda)^p)$ for all p are conserved (which implies that the whole spectrum of $\Lambda(\lambda)$ is constant in x). It turns out [15] that one can find n independent non-trivial integrals in this set which are in involution thus guaranteeing the integrability of (3.23) is a sense of the Arnold–Liouville theorem 1.2.2.

The resulting solutions to KdV are known as 'finite gap' potentials. Let us justify this terminology. The spectrum of Λ does not depend on x and so the coefficients of the characteristic polynomial

$$\det\left(\mathbf{1}\mu - \Lambda(\lambda)\right) = 0$$

also do not depend on x. Using the fact that $\Lambda(\lambda)$ is trace free we can rewrite this polynomial as

$$\mu^2 + R(\lambda) = 0 \tag{3.24}$$

where

$$\begin{aligned} R(\lambda) &= \lambda^{2n+1} + a_1\lambda^{2n} + \ldots + a_{2n}\lambda + a_{2n+1} \\ &= (\lambda - \lambda_0)\ldots(\lambda - \lambda_{2n}). \end{aligned}$$

Therefore the coefficients $a_1, \ldots, a_{2n+1}$ (or equivalently $\lambda_0, \ldots, \lambda_{2n}$) do not depend on x. However $(n+1)$ of those coefficients can be expressed in terms of the constants c_k and thus the corresponding first integrals are trivial. This leaves us with n first integrals for an ODE (3.23) of order $2n$.

It is possible to show [15] that

- All solutions to the KdV equation with periodic initial data arise from (3.23).
- For each λ the corresponding eigenfunctions of the Schrödinger operator $L\psi = \lambda\psi$ can be expanded in a basis $\psi_\pm$ such that

  $$\psi_\pm(x + X_0) = e^{\pm ipX_0}\psi_\pm(x)$$

  for some $p = p(\lambda)$ ($\psi_\pm$ are called Bloch functions). The set of real λ for which $p(\lambda) \in \mathbb{R}$ is called the permissible zone. The roots of the polynomial $R(\lambda)$ are the end-points of the permissible zones

  $$(\lambda_0, \lambda_1), \quad (\lambda_2, \lambda_3), \quad \ldots, \quad (\lambda_{2n-2}, \lambda_{2n-1}), \quad (\lambda_{2n}, \infty).$$

The equation (3.24) defines a Riemann surface Γ of genus n. The number of forbidden zones (gaps) is therefore finite for the periodic solutions as the Riemann surface (3.24) has finite genus. This justifies the name 'finite gap' potentials.

- **Example.** If $n = 0$ the Riemann surface Γ has topology of the sphere and the corresponding solution to the KdV is a constant. If $n = 1$ Γ is called an elliptic curve (it has topology of a two-torus) and the ODE (3.23) is solvable by elliptic functions: the stationary condition

  $$c_0\frac{\partial u}{\partial t_0} + c_1\frac{\partial u}{\partial t_1} = 0$$

  yields

  $$c_0 u_x + c_1(6uu_x - u_{xxx}) = b, \qquad b = \text{const}$$

  where we used the fist two equations of the hierarchy. We can set $c_1 = 1$ redefining the other two constants. This ODE can be integrated and the general solution is a Weierstrass elliptic function

  $$\int \frac{du}{\sqrt{2u^3 + c_0u^2 - 2bu + d}} = x - x_0.$$

  The stationary condition implies that $u = u(x - c_0t)$ where we have identified $t_0 = x, t_1 = t$. Thus $x_0 = c_0t$. These solutions are called cnoidal waves because the corresponding elliptic function is often denoted 'cn'.

If $n > 1$ the Riemann surface Γ is a hyper-elliptic curve and the corresponding KdV potential is given in terms of Riemann's Theta function [15].

Chapter 4

Lie symmetries and reductions

4.1 Lie groups and Lie algebras

Phrases like 'the unifying role of symmetry in ...' feature prominently in the popular science literature. Depending on the subject, the symmetry may be 'cosmic', 'Platonic', 'perfect', 'broken' or even 'super'[1].

The mathematical framework used to define and describe the symmetries is group theory. Recall that a group is a set G with a map

$$G \times G \to G, \quad (g_1, g_2) \to g_1 g_2$$

called the group multiplication which satisfies the following properties:

- Associativity
$$(g_1 g_2) g_3 = g_1 (g_2 g_3) \qquad \forall g_1, g_2, g_3 \in G.$$
- There exist an identity element $e \in G$ such that
$$eg = ge = g, \qquad \forall g \in G.$$
- For any $g \in G$ there exists an inverse element $g^{-1} \in G$ such that
$$gg^{-1} = g^{-1} g = e.$$

A group G acts on a set X if there exists a map $G \times X \to X$, $(g, p) \to g(p)$ such that

$$e(p) = p, \quad g_2(g_1(p)) = (g_2 g_1)(p)$$

for all $p \in X$, and $g_1, g_2 \in G$. The set $\mathrm{Orb}(p) = \{g(p),\, g \in G\} \subset X$ is called the orbit of p. Groups acting on sets are often called groups of transformations.

[1] Supersymmetry is a symmetry between elementary particles known as bosons and fermions. It is a symmetry of equations underlying the current physical theories. Supersymmetry predicts that each elementary particle has its supersymmetric partner. No one has yet observed supersymmetry. Perhaps it will be found in the LHC. See a footnote on page 26.

In this Chapter we shall explore the groups which act on solutions to differential equations. Such group actions occur both for integrable and non-integrable systems so the methods we shall study are quite universal[2]. In fact all the techniques of integration of differential equations (like separation of variables, integrating factors, homogeneous equations, ...) students have encounter in their education are special cases of the symmetry approach. See [16] for a very complete treatment of this subject and [9] for an elementary introduction at an undergraduate level.

The symmetry programme goes back to a 19th century Norwegian mathematician Sophus Lie who developed a theory of continuous transformations now known as Lie groups. One of most important of Lie's discoveries was that a continuous group G of transformations is easy to describe by infinitesimal transformations characterising group elements close (in a sense of Taylor's theorem) to the identity element. These infinitesimal transformations are elements of Lie algebra $\mathfrak{g}$. For example a general element of the rotation group $G = SO(2)$

$$g(\varepsilon) = \begin{pmatrix} \cos\varepsilon & -\sin\varepsilon \\ \sin\varepsilon & \cos\varepsilon \end{pmatrix}$$

depends on one parameter ε. The group $SO(2)$ is a Lie group as g, its inverse and the group multiplication depend on ε in a continuous and differentiable way. This Lie group is one–dimensional as one parameter - the angle of rotation - is sufficient to describe any rotation around the origin in $\mathbb{R}^2$. A rotation in $\mathbb{R}^3$ depends on three such parameters - the Euler angles used in classical dynamics - so $SO(3)$ is a three dimensional Lie group. Now consider the Taylor series

$$g(\varepsilon) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \varepsilon \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} + O(\varepsilon^2).$$

The antisymmetric matrix

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

represents an infinitesimal rotation as $A\mathbf{x} = (-y, x)^T$ are components of the vector tangent to the orbit of $\mathbf{x}$ at $\mathbf{x}$. The one dimensional vector space spanned by A is called a Lie algebra of $SO(2)$.

The following definition is not quite correct (Lie groups should be defined as manifolds - see the Definition A.0.3 in Appendix A) but it is sufficient for our purposes.

Definition 4.1.1 *An m–dimensional Lie group is a group whose elements depend continuously of m parameters such that the maps $(g_1, g_2) \to g_1 g_2$ and $g \to g^{-1}$ are smooth (infinitely differentiable) functions of these parameters.*

The infinitesimal description of Lie groups is given by Lie algebras.

Definition 4.1.2 *A Lie algebra is a vector space $\mathfrak{g}$ with an anti–symmetric bilinear operation called a Lie bracket $[\,,\,]_{\mathfrak{g}} : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ which satisfies the Jacobi identity*

$$[A,[B,C]] + [C,[A,B]] + [B,[C,A]] = 0, \qquad \forall A, B, C \in \mathfrak{g}.$$

[2]It is however the case that integrable systems admit 'large' groups of symmetries and non-integrable systems usually do not.

If the vectors $A_1, \ldots, A_{\dim \mathfrak{g}}$ span $\mathfrak{g}$, the algebra structure is determined by the structure constants $f^\gamma_{\alpha\beta}$ such that

$$[A_\alpha, A_\beta]_\mathfrak{g} = \sum_\gamma f^\gamma_{\alpha\beta} A_\gamma, \qquad \alpha, \beta, \gamma = 1, \ldots, \dim \mathfrak{g}.$$

The Lie bracket is related to non-commutativity of the group operation as the following argument demonstrates. Let $a, b \in G$. Set

$$a = I + \varepsilon A + O(\varepsilon^2), \qquad b = I + \varepsilon B + O(\varepsilon^3)$$

for some A, B and calculate

$$aba^{-1}b^{-1} = (I + \varepsilon A + \ldots)(I + \varepsilon B + \ldots)(I - \varepsilon A + \ldots)(I - \varepsilon B + \ldots) = I + \varepsilon^2 [A, B] + O(\varepsilon^2)$$

where $\ldots$ denote terms of higher order in ε and we used the fact $(1 + \varepsilon A)^{-1} = 1 - \varepsilon A + \ldots$ which follows from the Taylor series. Some care needs to be taken with the above argument as we have neglected the second order terms in the group elements but not in the answer. The readers should convince themselves that these terms indeed cancel out.

- **Example.** Consider the group of special orthogonal transformations $SO(n)$ which consist of n by n matrices a such that

  $$aa^T = I, \qquad \det a = 1.$$

  These conditions imply that only $n(n-1)/2$ matrix components are independent and $SO(n)$ is a Lie group of dimension $n(n-1)/2$. Setting $a = I + \varepsilon A + O(\varepsilon^2)$ shows that infinitesimal version of the orthogonal condition is anti-symmetry

  $$A + A^T = 0.$$

  Given two anti-symmetric matrices their commutator is also anti-symmetric as

  $$[A, B]^T = B^T A^T - A^T B^T = -[A, B].$$

  Therefore the vector space of antisymmetric matrices is a Lie algebra with a Lie bracket defined to be the matrix commutator. This Lie algebra, called $\mathfrak{so}(n)$, is a vector space of dimension $n(n-1)/2$. This is equal to the dimension (the number of parameters) of the corresponding Lie group $SO(n)$.

- **Example.** An example of a three–dimensional Lie group is given by the group of 3 by 3 upper triangular matrices

  $$g(m_1, m_2, m_3) = \begin{pmatrix} 1 & m_1 & m_3 \\ 0 & 1 & m_2 \\ 0 & 0 & 1 \end{pmatrix}. \tag{4.1}$$

  Note that $g = \mathbf{1} + \sum_\alpha m_\alpha T_\alpha$, where the matrices T_α are

  $$T_1 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad T_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad T_3 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{4.2}$$

This Lie group is called **Nil**, as the matrices T_α are all nilpotent. These matrices span the Lie algebra of the group **Nil** and have the commutation relations

$$[T_1, T_2] = T_3, \quad [T_1, T_3] = 0, \quad [T_2, T_3] = 0. \tag{4.3}$$

This gives the structure constants $f^3_{12} = -f^3_{21} = 1$ and all other constants vanish.

A three-dimensional Lie algebra with these structure constants is called the Heisenberg algebra because of its connection with Quantum Mechanics - think of T_1 and T_2 as position and momentum operators respectively, and T_3 as $i\hbar$ times the identity operator.

In the above example the Lie algebra of a Lie group was represented by matrices. If the group acts on a subset X of $\mathbb{R}^n$, its Lie algebra is represented by vector fields[3] on X. This approach underlies the application of Lie groups to differential equations so we shall study it next.

4.2 Vector fields and one parameter groups of transformations

Let X be an open set in $\mathbb{R}^n$ with local coordinates $x^1, \ldots, x^n$ and let $\gamma : [0,1] \longrightarrow X$ be a parametrised curve, so that $\gamma(\varepsilon) = (x^1(\varepsilon), \ldots, x^n(\varepsilon))$. The tangent vector $V|_p$ to this curve at a point $p \in X$ has components

$$V^a = \dot{x}^a|_p, \quad a = 1, \ldots, n, \qquad \text{where } \doteq \frac{d}{d\varepsilon}.$$

The collection of all tangent vectors to all possible curves through p is an n-dimensional vector space called the tangent space T_pX. The collection of all tangent spaces as x varies in X is called a tangent bundle $TX = \cup_{x\in M} T_xX$. The tangent bundle is a manifold of dimension $2n$ (see Appendix A).

A vector field V on X assigns a tangent vector $V|_p \in T_pX$ to each point in X. Let $f : X \longrightarrow \mathbb{R}$ be a function on X. The rate of change of f along the curve is measured by a derivative

$$\begin{aligned} \frac{d}{d\varepsilon} f(x(\varepsilon))|_{\varepsilon=0} &= V^a \frac{\partial f}{\partial x^a} \\ &= V(f) \end{aligned}$$

where

$$V = V^1 \frac{\partial}{\partial x^1} + \ldots + V^n \frac{\partial}{\partial x^1}.$$

Thus vector fields can be thought of as first order differential operators. The derivations $\{\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^n}\}$ at the point p denote the elements of the basis of T_pX.

An integral curve γ of a vector field V is defined by $\dot{\gamma}(\varepsilon) = V|_{\gamma(\varepsilon)}$or equivalently

$$\frac{dx^a}{d\varepsilon} = V^a(x). \tag{4.4}$$

[3]The structure constants $f^\gamma_{\alpha\beta}$ do not depend on which of these representation is used.

This system of ODEs has a unique solution for each initial data, and the integral curve passing through p with coordinates x^a is called a flow $\tilde{x}^a(\varepsilon, x^b)$. The vector field V is called a generator of the flow, as

$$\tilde{x}^a(\varepsilon, x) = x^a + \varepsilon V^a(x) + O(\varepsilon^2).$$

Determining the flow of a given vector field comes down to solving a system of ODEs (4.4).

- **Example.** Integral curves of the vector field

$$V = x\frac{\partial}{\partial x} + \frac{\partial}{\partial y}$$

on $\mathbb{R}^2$ are found by solving a pair of ODEs $\dot{x} = x$, $\dot{y} = 1$. Thus

$$(x(\varepsilon), y(\varepsilon)) = (x(0)e^{\varepsilon}, y(0) + \varepsilon).$$

There is one integral curve passing through each point in $\mathbb{R}^2$.

The flow is an example of one–parameter group of transformations, as

$$\tilde{x}(\varepsilon_2, \tilde{x}(\varepsilon_1, x)) = \tilde{x}(\varepsilon_1 + \varepsilon_2, x), \qquad \tilde{x}(0, x) = x.$$

An invariant of a flow is a function $f(x^a)$ such that $f(x^a) = f(\tilde{x}^a)$ or equivalently

$$V(f) = 0$$

where V is the generating vector field.

- **Example.** The one parameter group $SO(2)$ of rotations on the plane

$$(\tilde{x}, \tilde{y}) = (x\cos\varepsilon - y\sin\varepsilon, x\sin\varepsilon + y\cos\varepsilon)$$

is generated by

$$\begin{aligned} V &= \Big(\frac{\partial \tilde{y}}{\partial \varepsilon}|_{\varepsilon=0}\Big)\frac{\partial}{\partial y} + \Big(\frac{\partial \tilde{x}}{\partial \varepsilon}|_{\varepsilon=0}\Big)\frac{\partial}{\partial x} \\ &= x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}. \end{aligned}$$

The function $r = \sqrt{x^2 + y^2}$ is an invariant of V.

A Lie bracket of two vector fields V, W is a vector field $[V, W]$ defined by its action on functions

$$[V, W](f) := V(W(f)) - W(V(f)). \tag{4.5}$$

The components of the Lie bracket are

$$[V, W]^a = V^b\frac{\partial W^a}{\partial x^b} - W^b\frac{\partial V^a}{\partial x^b}.$$

From its definition the Lie bracket is bi-linear, antisymmetric and it satisfies the Jacobi identity

$$[V, [W, U]] + [U, [V, W]] + [W, [U, V]] = 0. \tag{4.6}$$

A geometric interpretation of the Lie bracket is the infinitesimal commutator of two flows. To see it consider $\tilde{x}_1(\varepsilon_1, x)$ and $\tilde{x}_2(\varepsilon_2, x)$ which are the flows of vector fields V_1 and V_2 respectively. For any $f : X \to \mathbb{R}$ define

$$F(\varepsilon_1, \varepsilon_2, x) := f(\tilde{x}_1(\varepsilon_1, (\tilde{x}_2(\varepsilon_2, x)))) - f(\tilde{x}_2(\varepsilon_2, (\tilde{x}_1(\varepsilon_1, x)))).$$

Then

$$\frac{\partial^2}{\partial\varepsilon_1\partial\varepsilon_2} F(\varepsilon_1, \varepsilon_2, x)|_{\varepsilon_1=\varepsilon_2=0} = [V_1, V_2](f).$$

- **Example.** Consider the three–dimensional Lie group **Nil** of 3 by 3 upper triangular matrices

$$g(m_1, m_2, m_3) = \begin{pmatrix} 1 & m_1 & m_3 \\ 0 & 1 & m_2 \\ 0 & 0 & 1 \end{pmatrix}$$

acting on $\mathbb{R}^3$ by matrix multiplication

$$\tilde{\mathbf{x}} = g(m_1, m_2, m_3)\mathbf{x} = (x + m_1 y + m_3 z, y + m_2 z, z).$$

The corresponding vector fields[4] V_1, V_2, V_3 are

$$V_\alpha = \Big(\frac{\partial\tilde{x}}{\partial m_\alpha}\frac{\partial}{\partial\tilde{x}} + \frac{\partial\tilde{y}}{\partial m_\alpha}\frac{\partial}{\partial\tilde{y}} + \frac{\partial\tilde{z}}{\partial m_\alpha}\frac{\partial}{\partial\tilde{z}}\Big)|_{(m_1,m_2,m_3)=(0,0,0)}$$

which gives

$$V_1 = y\frac{\partial}{\partial x}, \quad V_2 = z\frac{\partial}{\partial y}, \quad V_3 = z\frac{\partial}{\partial x}.$$

The Lie brackets of these vector fields are

$$[V_1, V_2] = -V_3, \quad [V_1, V_3] = 0, \quad [V_2, V_3] = 0.$$

Thus we have obtained the representation of the Lie algebra of **Nil** by vector fields on $\mathbb{R}^3$. Comparing this with the commutators of the matrices (4.2) we see that the structure constants only differ by an overall sign. The Lie algebra spanned by the vector fields V_α is isomorphic to Lie algebra spanned by the matrices M_α.

- **Example.** A driver of a car has two transformation at his disposal. These are generated by vector fields

$$\text{STEER} = \frac{\partial}{\partial\phi}, \qquad \text{DRIVE} = \cos\theta\frac{\partial}{\partial x} + \sin\theta\frac{\partial}{\partial y} + \frac{1}{L}\tan\phi\frac{\partial}{\partial\theta}, \qquad L = \text{const.}$$

where (x, y) are coordinates of the center of the rear axle, θ specifies the direction of the car, and ϕ is the angle between the front wheels and the direction of the car. These two flows don't commute, and

$$[\text{STEER}, \text{DRIVE}] = \text{ROTATE},$$

[4]Note that the lower index labels the vector fields while the upper index labels the components. Thus $V_\alpha = V_\alpha^a \partial/\partial x^a$

where the vector field

$$\text{ROTATE} = \frac{1}{L\cos^2\phi}\frac{\partial}{\partial\theta}$$

generates the manoeuvre steer, drive, steer back, drive back. This manoeuvre alone doesn't guarantee that the driver parks his car in a tight space. The commutator

$$[\text{DRIVE}, \text{ROTATE}] = \frac{1}{L\cos^2\phi}\Big(\sin\theta\frac{\partial}{\partial x} - \cos\theta\frac{\partial}{\partial y}\Big) = \text{SLIDE}$$

is the key to successful parallel parking. One needs to perform the following sequence steer, drive, steer back, drive, steer, drive back, steer back, drive back!

In general the Lie bracket is a closed operation in a set of the vector fields generating a group. The vector space of vector fields generating the group action gives a representation of the corresponding Lie algebra. The structure constants $f^{\gamma}_{\alpha\beta}$ do not depend on which of the representations (matrices or vector fields) is used.

4.3 Symmetries of differential equations

Let $u = u(x,t)$ be a solution to the KdV equation (2.1). Consider a vector field

$$V = \xi(x,t,u)\frac{\partial}{\partial x} + \tau(x,t,u)\frac{\partial}{\partial t} + \eta(x,t,u)\frac{\partial}{\partial u}$$

on the space of dependent and independent variables $\mathbb{R}\times\mathbb{R}^2$. This vector field generates a one–parameter group of transformations

$$\tilde{x} = \tilde{x}(x,t,u,\varepsilon), \qquad \tilde{t} = \tilde{t}(x,t,u,\varepsilon), \qquad \tilde{u} = \tilde{u}(x,t,u,\varepsilon).$$

This group is called a symmetry of the KdV equation if

$$\frac{\partial\tilde{u}}{\partial\tilde{t}} - 6\tilde{u}\frac{\partial\tilde{u}}{\partial\tilde{x}} + \frac{\partial^3\tilde{u}}{\partial\tilde{x}^3} = 0.$$

The common abuse of terminology is to refer to the corresponding vector field as a symmetry, although the term infinitesimal symmetry is more appropriate.

- **Example.** An example of a symmetry of the KdV is given by

  $$\tilde{x} = x, \qquad \tilde{t} = t + \varepsilon, \qquad \tilde{u} = u.$$

  It is a symmetry as there is no explicit time dependence in the KdV. Its generating vector field is

  $$V = \frac{\partial}{\partial t}.$$

Of course there is nothing special about KdV in this definition and the concept of a symmetry applies generally to PDEs and ODEs.

Definition 4.3.1 *Let $X = \mathbb{R}^n \times \mathbb{R}$ be the space of independent and dependent variables in a PDE. A one–parameter group of transformations of this space*

$$\tilde{u} = \tilde{u}(x^a, u, \varepsilon), \qquad \tilde{x}^b = \tilde{x}^b(x^a, u, \varepsilon)$$

is called a Lie point symmetry (or symmetry for short) group of a PDE

$$F[u, \frac{\partial u}{\partial x^a}, \frac{\partial^2 u}{\partial x^a \partial x^b}, \ldots] = 0 \tag{4.7}$$

if its action transforms solutions to other solutions i.e.

$$F[\tilde{u}, \frac{\partial \tilde{u}}{\partial \tilde{x}^a}, \frac{\partial^2 \tilde{u}}{\partial \tilde{x}^a \partial \tilde{x}^b}, \ldots] = 0.$$

This definition naturally extends to multi–parameter groups of transformation. A Lie group G is a symmetry of a PDE if any of its one–parameter subgroups is a symmetry in a sense of Definition 4.3.1.

A knowledge of Lie point symmetries is useful for the following reasons

- It allows to use known solutions to construct new solutions.

  Example. The Lorentz group

  $$(\tilde{x}, \tilde{t}) = \Big(\frac{x - \varepsilon t}{\sqrt{1-\varepsilon^2}}, \frac{t - \varepsilon x}{\sqrt{1-\varepsilon^2}}\Big), \qquad \varepsilon \in (-1, 1)$$

  is the symmetry group of the Sine–Gordon equation (2.2). Any t–independent solution $\phi_S(x)$ to (2.2) can be used to obtain a time dependent solution

  $$\phi(x,t) = \phi_S\Big(\frac{x - \varepsilon t}{\sqrt{1-\varepsilon^2}}\Big), \qquad \varepsilon \in (-1, 1).$$

  In physics this procedure is known as 'Lorentz boost'. The parameter ε is usually denoted by v and called velocity. For example the Lorentz boost of a static kink is a moving kink.

- For ODEs each symmetry reduces the order by 1. So a knowledge of a sufficiently many symmetries allows a construction of the most general solution.

  Example. An ODE

  $$\frac{du}{dx} = F\Big(\frac{u}{x}\Big)$$

  admits a scaling symmetry

  $$(x, u) \longrightarrow (e^\varepsilon x, e^\varepsilon u), \qquad \varepsilon \in \mathbb{R}.$$

  This one–dimensional group is generated by a vector field

  $$V = x\frac{\partial}{\partial x} + u\frac{\partial}{\partial u}.$$

Introduce the invariant coordinates

$$r = \frac{u}{x}, \qquad s = \log|x|$$

so that

$$V(r) = 0, \qquad V(s) = 1.$$

If $F(r) = r$ the general solution is $r = \text{const}$. Otherwise

$$\frac{ds}{dr} = \frac{1}{F(r) - r}$$

and the general implicit solution is

$$\log|x| + c = \int^{\frac{u}{x}} \frac{dr}{F(r) - r}.$$

- For PDEs the knowledge of the symmetry group is not sufficient to construct the most general solution, but it can be used to find special solutions which admit symmetry.

  Example. Consider the one–parameter group of transformations

  $$(\tilde{x}, \tilde{t}, \tilde{u}) = (x + c\varepsilon, t + \varepsilon, u)$$

  where $c \in \mathbb{R}$ is a constant. It is straightforward to verify that this group is a Lie point symmetry of the KdV equation (2.1). It is generated by the vector field

  $$V = \frac{\partial}{\partial t} + c\frac{\partial}{\partial x}$$

  and the corresponding invariants are u and $\xi = x - ct$. To find the group invariant solutions assume that a solution of the KdV equation is of the form

  $$u(x, t) = f(\xi).$$

  Substituting this to the KdV yields a third order ODE which easily integrates to

  $$\frac{1}{2}\Big(\frac{df}{d\xi}\Big)^2 = f^3 + \frac{1}{2}cf^2 + \alpha f + \beta$$

  where (α, β) are arbitrary constants. This ODE is solvable in terms of an elliptic integral, which gives all group invariant solutions in the implicit form

  $$\int \frac{df}{\sqrt{f^3 + \frac{1}{2}cf^2 + \alpha f + \beta}} = \sqrt{2}\xi.$$

  Thus we have recovered the cnoidal wave which in Section 3.4 arose from the finite gap integration. In fact the one–soliton solution (2.3) falls into this category: if f and its first two derivatives tend to zero as $|\xi| \to \infty$ then α, β are both zero and the elliptic integral reduces to an elementary one. Finally we obtain

  $$u(x, t) = -\frac{2\chi^2}{\cosh^2 \chi(x - 4\chi^2 t - \phi_0)}$$

  which is the one–soliton solution (2.3) to the KdV equation.

4.3.1 How to find symmetries

Some of them can be guessed. For example if there is no explicit dependence of independent coordinates in the equation then the translations $\tilde{x}^a = x^a + c^a$ are symmetries. All translations form an n–parameter abelian group generated by n vector fields $\partial/\partial x^a$.

In the general case of (4.7) we could substitute

$$\tilde{u} = u + \varepsilon\eta(x^a, u) + O(\varepsilon^2), \qquad \tilde{x}^b = x^b + \varepsilon\xi^b(x^a, u) + O(\varepsilon^2)$$

into the equation (4.7) and keep the terms linear in ε. A more systematic method is given by the *prolongation* of vector field. Assume that the space of independent variables is coordinatised by (x, t) and the equation (4.7) is of the form

$$F(u, u_x, u_{xx}, u_{xxx}, u_t) = 0.$$

(for example KdV is of that form). The prolongation of the vector field

$$V = \xi(x, t, u)\frac{\partial}{\partial x} + \tau(x, t, u)\frac{\partial}{\partial t} + \eta(x, t, u)\frac{\partial}{\partial u}$$

is

$$\mathbf{pr}(V) = V + \eta^t\frac{\partial}{\partial u_t} + \eta^x\frac{\partial}{\partial u_x} + \eta^{xx}\frac{\partial}{\partial u_{xx}} + \eta^{xxx}\frac{\partial}{\partial u_{xxx}},$$

where $(\eta^t, \eta^x, \eta^{xx}, \eta^{xxx})$ are certain functions of (u, x, t) which can be determined algorithmically in terms of (ξ, τ, η) and their derivatives (we will do it in the next Section). The prolongation $\mathbf{pr}(V)$ generates a one–parameter group of transformations on the 7–dimensional space with coordinates

$$(x, t, u, u_t, u_x, u_{xx}, u_{xxx}).$$

(This is an example of a jet space. The symbols $(u_t, u_x, u_{xx}, u_{xxx})$ should be regarded as independent coordinates and not as derivatives of u.). The vector field V is a symmetry of the PDE if

$$\mathbf{pr}(V)(F)|_{F=0} = 0. \tag{4.8}$$

This condition gives a linear system of PDEs for (ξ, τ, η). Solving this system yields the most general symmetry of a given PDE. The important point is that (4.8) is only required to hold when (4.7) is satisfied ('on shell' as a physicist would put it).

4.3.2 Prolongation formula

The first step in implementing the prolongation procedure is to determine the functions

$$\eta^t, \eta^x, \eta^{xx}, \dots$$

in the prolonged vector field. For simplicity we shall assume that we want to determine a symmetry of Nth order ODE

$$\frac{d^N u}{dx^N} = F\Big(x, u, \frac{du}{dx}, \cdots, \frac{d^{N-1}u}{dx^{N-1}}\Big).$$

Consider a vector field

$$V = \xi \frac{\partial}{\partial x} + \eta \frac{\partial}{\partial u}.$$

Its prolongation

$$\mathbf{pr}(V) = V + \sum_{k=1}^{N} \eta^{(k)} \frac{\partial}{\partial u^{(k)}}$$

generates a one–parameter transformation group

$$\tilde{x} = x + \varepsilon\xi + O(\varepsilon^2), \qquad \tilde{u} = u + \varepsilon\eta + O(\varepsilon^2), \qquad \tilde{u}^{(k)} = u^{(k)} + \varepsilon\eta^{(k)} + O(\varepsilon^2)$$

of the $(N+2)$ dimensional jet space with coordinates $(x, u, u', \ldots, u^N)$.

The prolongation is an algorithm for the calculation of the functions $\eta^{(k)}$. Set

$$D_x = \frac{\partial}{\partial x} + u' \frac{\partial}{\partial u} + u'' \frac{\partial}{\partial u'} + \cdots + u^{(N)} \frac{\partial}{\partial u^{(N-1)}}.$$

The chain rule gives

$$\tilde{u}^{(k)} = \frac{d\tilde{u}^{(k-1)}}{d\tilde{x}} = \frac{D_x \tilde{u}^{(k-1)}}{D_x \tilde{x}},$$

so

$$\tilde{u}^{(1)} = \frac{D_x \tilde{u}}{D_x \tilde{x}} = \frac{\frac{du}{dx} + \varepsilon D_x(\eta) + \ldots}{1 + \varepsilon D_x(\xi) + \ldots} = \frac{du}{dx} + \varepsilon(D_x\eta - \frac{du}{dx} D_x\xi) + O(\varepsilon^2).$$

Thus

$$\eta^{(1)} = D_x\eta - \frac{du}{dx} D_x\xi.$$

The remaining prolongation coefficients can now be constructed recursively: The relation

$$\tilde{u}^{(k)} = \frac{u^{(k)} + \varepsilon D_x(\eta^{(k-1)})}{1 + \varepsilon D_x(\xi)}$$

yields the general prolongation formula

$$\eta^{(k)} = D_x(\eta^{(k-1)}) - \frac{d^k u}{dx^k} D_x\xi. \tag{4.9}$$

The procedure is entirely analogous for PDEs, where $u = u(x^a)$ but one has to keep track of the index a labelling the independent variables. Set

$$D_a = \frac{\partial}{\partial x^a} + (\partial_a u)\frac{\partial}{\partial u} + (\partial_a^2 u)\frac{\partial}{\partial(\partial_a u)} + \cdots + (\partial_a^N u)\frac{\partial}{\partial(\partial_a^{N-1} u)},$$

where

$$\partial_a^k u = \frac{\partial^k u}{\partial (x^a)^k}.$$

The first prolongation is

$$\eta^{(a)} = D_a\eta - \sum_{b=1}^{n} (D_a\xi^b)\frac{\partial u}{\partial x^b}$$

and the higher prolongations are given recursively by the formula

$$\eta^{A,a} = D_a \eta^A - \sum_{b=1}^{n} (D_a \xi^b) \frac{\partial u^A}{\partial x^b}$$

where $A = (a_1, \ldots, a_k)$ is a multi-index and

$$u^A = \frac{\partial^k u}{\partial x^{a_1} \partial x^{a_2} \ldots \partial x^{a_k}}.$$

- **Example.** Let us follow the prolongation procedure to find the most general Lie–point symmetry of the second order ODE

$$\frac{d^2 u}{dx^2} = 0.$$

We first need to compute the second prolongation

$$\mathbf{pr}(V) = \xi \frac{\partial}{\partial x} + \eta \frac{\partial}{\partial u} + \eta^x \frac{\partial}{\partial u_x} + \eta^{xx} \frac{\partial}{\partial u_{xx}}.$$

This computation does not depend on the details of the equation but only on the prolongation formulae (4.9). The result is

$$\eta^x = \eta_x + (\eta_u - \xi_x) u_x - \xi_u u_x^2,$$

$$\eta^{xx} = \eta_{xx} + (2\eta_{xu} - \xi_{xx}) u_x + (\eta_{uu} - 2\xi_{xu}) u_x^2 - \xi_{uu} u_x^3 + (\eta_u - 2\xi_x) u_{xx} - 3\xi_u u_x u_{xx}.$$

Now we substitute this, and the ODE to the symmetry criterion (4.8)

$$\mathbf{pr}(V)(u_{xx}) = \eta^{xx} = 0.$$

Thus

$$\eta_{xx} + (2\eta_{xu} - \xi_{xx}) u_x + (\eta_{uu} - 2\xi_{xu}) u_x^2 - \xi_{uu} u_x^3 = 0$$

where we have used the ODE to set $u_{xx} = 0$. In the second order equation the value of u_x can be prescribed in an arbitrary way at each point (initial condition). Therefore the coefficients of u_x, u_x^2 and u_x^3 all vanish

$$\eta_{xx} = 0, \quad 2\eta_{xu} - \xi_{xx} = 0, \quad \eta_{uu} - 2\xi_{xu} = 0, \quad \xi_{uu} = 0.$$

The general solution of these linear PDEs is

$$\begin{aligned} \xi(x,u) &= \varepsilon_1 x^2 + \varepsilon_2 xu + \varepsilon_3 x + \varepsilon_4 u + \varepsilon_5, \\ \eta(x,u) &= \varepsilon_1 xu + \varepsilon_2 u^2 + \varepsilon_6 x + \varepsilon_7 u + \varepsilon_8. \end{aligned}$$

Therefore the trivial ODE in our example admits an eight dimensional group of symmetries.

Let $V_\alpha, \alpha = 1, \ldots, 8$ be the corresponding vector fields obtained by setting $\varepsilon_\alpha = 1$ and $\varepsilon_\beta = 0$ if $\beta \neq \alpha$

$$V_1 = x^2\frac{\partial}{\partial x} + xu\frac{\partial}{\partial u}, \quad V_2 = xu\frac{\partial}{\partial x} + u^2\frac{\partial}{\partial u}, \quad V_3 = x\frac{\partial}{\partial x}, \quad V_4 = u\frac{\partial}{\partial x},$$

$$V_5 = \frac{\partial}{\partial x}, \quad V_6 = x\frac{\partial}{\partial u}, \quad V_7 = u\frac{\partial}{\partial u}, \quad V_8 = \frac{\partial}{\partial u}.$$

Each of the eight vector fields generates a one–parameter group of transformations. Calculating the Lie brackets of these vector fields verifies that they form a Lie algebra of $PGL(3, \mathbb{R})$.

It is possible to show that the Lie point symmetry group of a general second order ODE has dimension at most 8. If this dimension is 8 then the ODE is equivalent to $u_{xx} = 0$ by a coordinate transformation $u \to U(u, x), x \to X(u, x)$.

This example shows that the process of prolonging the vector fields and writing down the linear PDEs characterising the symmetries is tedious but algorithmic. It is worth doing a few examples by hand to familiarise oneself with the method but in practice it is best to use computer programmes like MAPLE or MATHEMATICA to do symbolic computations.

- **Example.** Lie point symmetries of KdV. The vector fields

$$V_1 = \frac{\partial}{\partial x}, \quad V_2 = \frac{\partial}{\partial t}, \quad V_3 = \frac{\partial}{\partial u} - 6t\frac{\partial}{\partial x}, \quad V_4 = x\frac{\partial}{\partial x} + 3t\frac{\partial}{\partial t} - 2u\frac{\partial}{\partial u}$$

generate a four–parameter symmetry group of KdV. The group is non–abelian as the structure constants of the Lie algebra spanned by V_α are non–zero:

$$[V_2, V_3] = -6V_1, \quad [V_1, V_4] = V_1, \quad [V_2, V_4] = 3V_2, \quad [V_3, V_4] = -2V_3$$

and all other Lie brackets vanish.

One can use the prolongation procedure to show that this is in fact the most general symmetry group of KdV. One needs to find a third prolongation of a general vector field on $\mathbb{R}^3$ - this can be done 'by hand' but it is best to use `MAPLE` package `liesymm` with the command `determine`. Type `help(determine);` and take it from there.

4.4 Painlevé equations

In this Section we shall consider ODEs in complex domain. This means that both the dependent and independent variables are complex numbers. Let us first discuss the linear ODEs of the form

$$\frac{d^N w}{dz^N} + p_{N-1}(z)\frac{d^{N-1}w}{dz^{N-1}} + \cdots + p_1(z)\frac{dw}{dz} + p_0(z)w = 0 \tag{4.10}$$

where $w = w(z)$. If the functions $p_0, \ldots, p_{N-1}$ are analytic at $z = z_0$, then z_0 is called a regular point and for a given initial data there exist a unique analytic solution in the form of a power series

$$w(z) = \sum_k a_k (z - z_0)^k.$$

The singular points of the ODE (4.10) can be located only at the singularities of p_k. Thus the singularities are fixed – their location does not depend on the initial conditions. Nonlinear ODEs lose this property.

- **Example.** Consider a simple nonlinear ODE and its general solution

$$\frac{dw}{dz} + w^2 = 0, \qquad w(z) = \frac{1}{z - z_0}.$$

The location of the singularity depends on the constant of integration z_0. This is a movable singularity.

A singularity of a nonlinear ODE can be a pole (of arbitrary order), a branch point or an essential singularity.

- **Example.** The ODE with the general solution

$$\frac{dw}{dz} + w^3 = 0, \quad w(z) = \frac{1}{\sqrt{2(z - z_0)}}.$$

has a movable singularity which is a branch point. Another example with a movable logarithmic branch point is

$$\frac{dw}{dz} + e^w = 0, \quad w(z) = \ln(z - z_0).$$

Definition 4.4.1 *The ODE*

$$\frac{d^N w}{dz^N} = F\Big(\frac{d^{N-1}w}{dz^{N-1}}, \cdots, \frac{dw}{dz}, w, z\Big)$$

where F is rational in w and its derivatives has Painlevé property if its movable singularities are at worst poles.

In 19th century Painlevé, Gambier and Kowalewskaya aimed to classify all second order ODEs with Painlevé property up to a change of variables

$$\tilde{w}(w, z) = \frac{a(z)w + b(z)}{c(z)w + d(z)}, \quad \tilde{z}(z) = \phi(z)$$

where the functions a, b, c, d, ϕ are analytic in z. There exist 50 canonical types 44 of which are solvable in terms of 'known' functions (sinus, cosinus, elliptic functions or in general solutions to linear ODEs) [10]. The remaining 6 equations define new transcendental functions

$$\begin{aligned}
\frac{d^2w}{dz^2} &= 6w^2 + z \quad \text{PI}, \\
\frac{d^2w}{dz^2} &= 2w^3 + wz + \alpha \quad \text{PII}, \\
\frac{d^2w}{dz^2} &= \frac{1}{w}\Big(\frac{dw}{dz}\Big)^2 - \frac{1}{z}\frac{dw}{dz} + \frac{\alpha w^2 + \beta}{z} + \gamma w^3 + \frac{\delta}{w} \quad \text{PIII},
\end{aligned} \tag{4.11}$$

$$
\begin{aligned}
\frac{d^2w}{dz^2} &= \frac{1}{2w}\Big(\frac{dw}{dz}\Big)^2 + \frac{3}{2}w^3 + 4zw^2 + 2(z^2-\alpha)w + \frac{\beta}{w} \quad \text{PIV},\\
\frac{d^2w}{dz^2} &= \left(\frac{1}{2w} + \frac{1}{w-1}\right)\Big(\frac{dw}{dz}\Big)^2 - \frac{1}{z}\frac{dw}{dz} + \frac{(w-1)^2}{z^2}\left(\alpha w + \frac{\beta}{w}\right)\\
&\quad + \frac{\gamma w}{z} + \frac{\delta w(w+1)}{w-1} \quad \text{PV},\\
\frac{d^2w}{dz^2} &= \frac{1}{2}\left(\frac{1}{w} + \frac{1}{w-1} + \frac{1}{w-z}\right)\Big(\frac{dw}{dz}\Big)^2 - \left(\frac{1}{z} + \frac{1}{z-1} + \frac{1}{w-z}\right)\frac{dw}{dz}\\
&\quad + \frac{w(w-1)(w-z)}{z^2(z-1)^2}\left(\alpha + \beta\frac{z}{w^2} + \gamma\frac{z-1}{(w-1)^2} + \delta\frac{z(z-1)}{(w-z)^2}\right) \quad \text{PVI}.
\end{aligned}
$$

Here $\alpha, \beta, \gamma, \delta$ are constants. Thus PVI belongs to a four–parameter family of ODEs but PI is rigid up to coordinate transformations.

How to check the Painlevé property for a given ODE? If a second order equation possesses the Painlevé property then it is either linearisable or can be put into one of the six Painlevé types by appropriate coordinate transformation. Exhibiting such transformation is often the most straightforward way of establishing the Painlevé property.

Otherwise, especially if we suspect that the equation does not have Painlevé property, the singular point analysis may be performed. If a general Nth order ODE possesses the PP then the general solution admits a Laurent expansion with a finite number of terms with negative powers. This expansion must contain N arbitrary constants so that the initial data consisting of w and its first $(N-1)$ derivatives can be specified at any point. Assume that a leading term in the expansion of the solution is of the form

$$
w(z) \sim a(z-z_0)^p, \qquad a \neq 0, \quad a, p \in \mathbb{C}
$$

as $z \to z_0$. Substitute this to the ODE and require the maximal balance condition. This means that two (or more) terms must be of equal maximally small order as $(z-z_0) \to 0$. This should determine a and p and finally the form of a solution around z_0. If z_0 is a singularity we should also be able to determine if it is movable or fixed.

- **Example.** Consider the ODE
$$
\frac{dw}{dz} = w^3 + z.
$$
The maximal balance condition gives
$$
ap(z-z_0)^{p-1} \sim a^3(z-z_0)^{3p}.
$$
Thus $p = -1/2, a = \pm i\sqrt{2}^{-1}$ and
$$
w(z) \sim \pm i\frac{\sqrt{2}}{2}(z-z_0)^{-1/2}
$$
possesses a movable branch point as z_0 depends on the initial conditions. The ODE does not have Painlevé property.

- **Example.** Consider the first Painlevé equation

$$\frac{d^2w}{dz^2} = 6w^2 + z.$$

The orders of the three terms in this equations are

$$p-2, \qquad 2p, \qquad 0.$$

Balancing the last two terms gives $p = 0$ but this is not the maximal balance as the first term is then of order -2. Balancing the first and last terms gives $p = 2$. This is a maximal balance and the corresponding solution is analytic around z_0. Finally balancing the first two terms gives $p = -2$ which again is the maximal balance: the 'balanced' terms behave like $(z - z_0)^{-4}$ and the remaining term is of order 0. Now we find that $a = a^2$ and so $a = 1$ and

$$w(z) \sim \frac{1}{(z - z_0)^2}.$$

Thus the movable singularity is a second order pole.

This singular point analysis is good to rule out Painlevé property, but does not give sufficient conditions (at least not in the heuristic form in which we presented it), as some singularities may have been missed or the Laurent series may be divergent. The analysis of sufficient conditions is tedious and complicated - we shall leave it out.

The Painlevé property guarantees that the solutions of six Painlevé equations are single valued thus giving rise to proper functions. The importance of the Painlevé equations is that they define new transcendental functions in the following way: Any sufficiently smooth function can be defined as a solution to certain DE. For example we can define the exponential function as a general solution to

$$\frac{dw}{dz} = w$$

such that $w(0) = 1$. Similarly we define the function PI from the general solution of the first Painlevé equation. From this point of view the exponential and PI functions are on equal footing. Of course we know more about the exponential as it possesses simple properties and arises in a wide range of problems in natural sciences. We know much less about PI, but it is largely because we have not yet been bothered to reveal its properties.

The irreducibility of the Painlevé equations is a more subtle issue. It roughly means the following. One can define a field of classical functions by starting off with the rational functions $Q[z]$ and adjoining those functions which arise as solutions of algebraic or linear differential equations with coefficients in $Q[z]$. For example the exponential, Bessel function, hyper-geometric function are all solutions of linear DEs, and thus are classical. A function is called irreducible (or transcendental) if it is not classical. Painlevé himself anticipated that the Painlevé equations define irreducible functions but the rigorous proofs for PI and PII appeared only recently. They use a far reaching extension of Galois theory from the number fields to differential fields of functions. The irreducibility problem is analogous to existence of non–algebraic numbers (numbers which are not roots of any polynomial equations with rational coefficients). Thus the the appearance of Galois theory is not that surprising.

4.4.1 Painlevé test

How to determine whether a given PDE is integrable? It is to large extend an open problem as the satisfactory definition of integrability of PDEs is still missing. The following algorithm is based on the observation of Ablowitz, Ramani and Segur [1] (see also [2]) that PDEs integrable by inverse scattering transform reduce (when the solutions are required to be invariant under some Lie symmetries) to ODEs with Painlevé property.

- **Example**. Consider a Lie point symmetry

$$(\tilde{\rho}, \tilde{\tau}) = (c\rho, \frac{1}{c}\tau), \qquad c \neq 0$$

of the Sine–Gordon equation

$$\frac{\partial^2 \phi}{\partial \rho \partial \tau} = \sin \phi.$$

The group invariant solutions are of the form $\phi(\rho, \tau) = F(z)$ where $z = \rho\tau$ is an invariant of the symmetry. Substituting $w(z) = \exp(iF(z))$ to the Sine–Gordon yields

$$\frac{d^2 w}{dz^2} = \frac{1}{w}\Big(\frac{dw}{dz}\Big)^2 - \frac{1}{z}\frac{dw}{dz} + \frac{1}{2}\frac{w^2}{z} - \frac{1}{2z},$$

which is the third Painlevé equation PIII with the special values of parameters

$$\alpha = \frac{1}{2}, \quad \beta = -\frac{1}{2}, \quad \gamma = 0, \quad \delta = 0.$$

- **Example.** Consider the modified KdV equation

$$v_t - 6v^2 v_x + v_{xxx} = 0,$$

and look for a Lie point symmetry of the form

$$(\tilde{v}, \tilde{x}, \tilde{t}) = (c^\alpha v, c^\beta x, c^\gamma t), \qquad c \neq 0.$$

The symmetry condition will hold if all three terms in the equation have equal weight

$$\alpha - \gamma = 3\alpha - \beta = \alpha - 3\beta.$$

This gives $\beta = -\alpha, \gamma = -3\alpha$ where α can be chosen arbitrarily. The corresponding symmetry group depends on one parameter c^α and is generated by

$$V = v\frac{\partial}{\partial v} - x\frac{\partial}{\partial x} - 3t\frac{\partial}{\partial t}.$$

This has two independent invariants

$$z = (3)^{-1/3}\, x\, t^{-1/3}, \qquad w = (3)^{1/3}\, v\, t^{1/3}$$

where the constant factor $(3)^{-1/3}$ has been added for convenience. The group invariant solutions are of the form $w = w(z)$ which gives

$$v(x,t) = (3t)^{-1/3} w(z).$$

Substituting this to the modified KdV equation leads to a third order ODE for $w(z)$

$$w_{zzz} - 6w^2 w_z - w - z w_z = 0.$$

Integrating this ODE once shows that $w(z)$ satisfies the second Painlevé equation PII with general value of the parameter α.

The general Painlevé test comes down to the following algorithm: Given a PDE

1. Find all its Lie point symmetries.
2. Construct ODEs characterising the group invariant solutions.
3. Check for Painlevé property.

This procedure only gives necessary conditions for integrability. If all reduction possess Painlevé property the PDE does not have to be integrable in general.

Appendix A

Manifolds

This course is intended to give an elementary introduction to the subject and the student is expected only to be familiar with basic real and complex analysis, algebra and dynamics as covered in the undergraduate syllabus. In particular no knowledge of differential geometry is assumed. One obvious advantage of this approach is that the course is suitable for advanced undergraduate students.

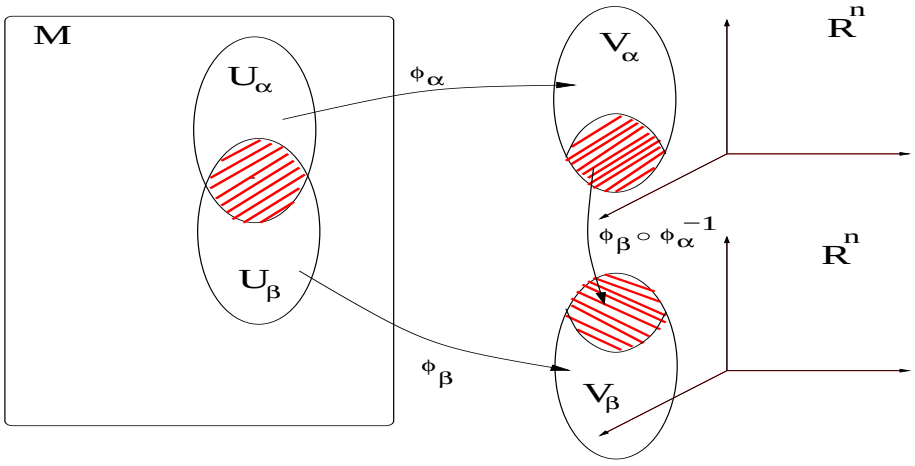
The disadvantage is that the discussion of Hamiltonian formalism and continuous groups of transformations in earlier chapters used phrases like 'spaces coordinatised by (p, q)', 'open sets in $\mathbb{R}^n$' or 'groups whose elements smoothly depend on parameters' instead calling these object by their real name - manifolds. This Appendix is intended to fill this gap.

Definition A.0.2 *An n–dimensional smooth manifold is a set M together with a collection of open sets U_α called the coordinate charts such that*

- *The open sets U_α labeled by a countable index α cover M.*
- *There exist one-to-one maps $\phi_\alpha : U_\alpha \to V_\alpha$ onto open sets in $\mathbb{R}^n$ such that for any pair of overlapping coordinate charts the maps*

$$\phi_\beta \circ \phi_\alpha^{-1} : \phi_\alpha(U_\alpha \cap U_\beta) \longrightarrow \phi_\beta(U_\alpha \cap U_\beta)$$

are smooth (i.e. infinitely differentiable) functions from $\mathbb{R}^n$ to $\mathbb{R}^n$.



Thus a manifold is a topological space together with additional structure which makes the local differential calculus possible. The space $\mathbb{R}^n$ itself is of course a manifold which can be covered by one coordinate chart.

- **Example.** A less trivial example is the unit sphere

$$S^n = \{\mathbf{r} \in \mathbb{R}^{n+1}, |\mathbf{r}| = 1\}.$$

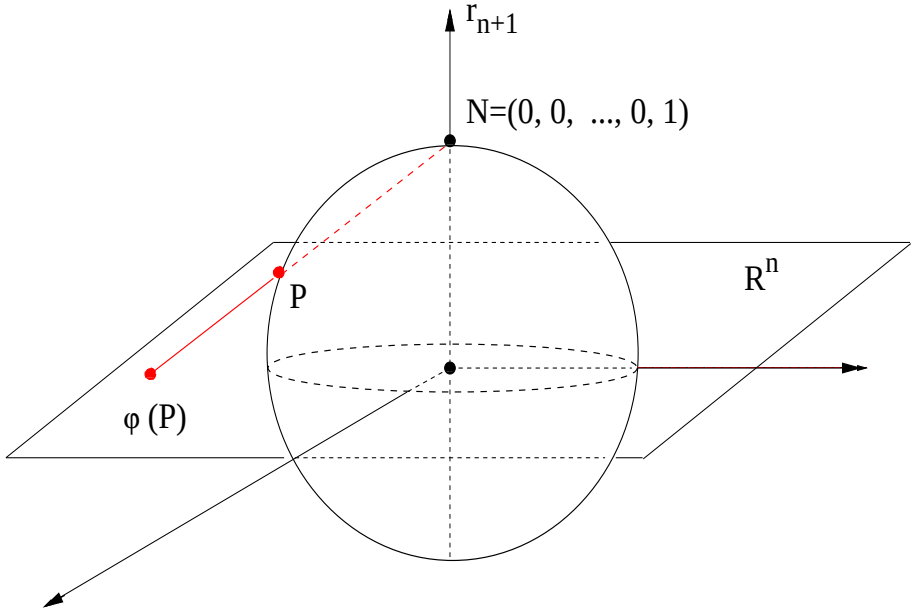
To verify that it is indeed a manifold, cover S^n by two open sets $U_1 = U$ and $U_2 = \widetilde{U}$

$$U = S^n/\{0, \ldots, 0, 1\}, \quad \widetilde{U} = S^n/\{0, \ldots, 0, -1\}$$

and define the local coordinates by stereographic projections

$$\phi(r_1, r_2, \ldots, r_{n+1}) = \left(\frac{r_1}{1 - r_{n+1}}, \ldots, \frac{r_n}{1 - r_{n+1}}\right) = (x_1, \ldots, x_n) \in \mathbb{R}^n,$$

$$\tilde{\phi}(r_1, r_2, \ldots, r_{n+1}) = \left(\frac{r_1}{1 + r_{n+1}}, \ldots, \frac{r_n}{1 + r_{n+1}}\right) = (x_1, \ldots, x_n) \in \mathbb{R}^n.$$



Using

$$\frac{r_k}{1 + r_{n+1}} = \left(\frac{1 - r_{n+1}}{1 + r_{n+1}}\right) \frac{r_k}{1 - r_{n+1}}, \qquad k = 1, \ldots, n$$

where $r_{n+1} \neq \pm 1$ shows that on the overlap $U \cap \widetilde{U}$ the transition functions

$$\phi \circ \tilde{\phi}^{-1}(x_1, \ldots, x_n) = \left(\frac{x_1}{x_1^2 + \ldots + x_n^2}, \ldots, \frac{x_n}{x_1^2 + \ldots + x_n^2}\right),$$

are smooth.

A cartesian product of manifolds is also a manifold. For example the n–torus arising in the Arnold–Liouville theorem 1.2.2 is a cartesian product on n one–dimensional spheres.

Another way to obtain interesting manifolds is to define them as surfaces in $\mathbb{R}^n$ by a vanishing condition for a set of functions. If $f_1, \dots, f_k : \mathbb{R}^n \to \mathbb{R}$ then the set

$$M_f := (x \in \mathbb{R}^n, \quad f_i(x) = 0, \quad i = 1, \dots, k) \tag{A1}$$

is a manifold if the rank of the k by n matrix of gradients ∇f_i is constant in a neighborhood of M_f in $\mathbb{R}^n$. If this rank is maximal and equal to k then $\dim M_f = n - k$. The manifold axioms can be verified using the implicit function theorem. For example the sphere S^{n-1} arises this way with $k = 1$ and $f_1 = 1 - |\mathbf{x}|^{\mathbf{2}}$. There is a theorem which says that every manifold arises as some surface in $\mathbb{R}^n$ for sufficiently large n. If the manifold is m dimensional then n is at most $2m + 1$. This useful theorem is now nearly forgotten - differential geometers like to think of manifolds as abstract objects defined by a collections of charts as in Definition A.0.2.

Lie groups

We can now give a proper definition of a Lie group

Definition A.0.3 *A Lie group G is a group and, at the same time, a smooth manifold such that the group operations*

$$G \times G \to G, \quad (g_1, g_2) \to g_1 g_2, \qquad \textit{and} \qquad G \to G, \quad g \to g^{-1}$$

are smooth maps between manifolds.

- **Example.** The general linear group $G = GL(n, \mathbb{R})$ is an open set in $\mathbb{R}^{n^2}$ defined by the condition $\det g \neq 0, g \in G$. It is therefore a Lie group of dimension n^2. The special orthogonal group $SO(n)$ is defined by (A1), where the $n(n+1)/2$ conditions in $\mathbb{R}^{n^2}$ are

  $$gg^T - \mathbf{1} = 0, \quad \det g = 1.$$

  The determinant condition just selects a connected component in the set of orthogonal matrices, so it does not count as a separate condition. It can be shown that the corresponding matrix of gradients has constant rank and thus $SO(n)$ is an $n(n-1)/2$ dimensional Lie group.

In Chapter 4 a Lie algebra $\mathfrak{g}$ was defined as a vector space with an anti–symmetric bilinear operation which satisfies the Jacobi identity (4.6).

A Lie algebra of a Lie group G is the tangent space to G at the identity element, $\mathfrak{g} = T_e G$ with the Lie bracket defined by a commutator of vector fields at e.

Proof of the first part of Arnold–Liouville's theorem 1.2.2.

The gradients ∇f_k are independent, thus the set

$$M_f := \{(p, q) \in M; f_k(p, q) = c_k\}$$

where $c_1, c_2, \ldots, c_n$ are constant defines a manifold of dimension n. Let $\xi^a = (p, q)$ be local coordinates on M such that the Poisson bracket is

$$\{f, g\} = \omega^{ab} \frac{\partial f}{\partial \xi^a} \frac{\partial g}{\partial \xi^b}, \qquad a, b = 1, 2, \ldots, 2n$$

where ω is a constant anti–symmetric matrix

$$\begin{pmatrix} 0 & 1_n \\ -1_n & 0 \end{pmatrix}.$$

The vanishing of Poisson brackets $\{f_j, f_k\} = 0$ implies that each Hamiltonian vector field

$$X_{f_k} = \omega^{ab} \frac{\partial f_k}{\partial \xi^b} \frac{\partial}{\partial \xi^a}$$

is orthogonal (in the Euclidean sense) to any of the gradients $\partial_a f_j, a = 1, \ldots, 2n, j, k = 1, \ldots, n$. The gradients are perpendicular to M_f, thus the Hamiltonian vector fields are tangent to M_f. They are also commuting as

$$[X_{f_j}, X_{f_k}] = -X_{\{f_j, f_k\}} = 0,$$

so the vectors generate an action of the abelian group $\mathbb{R}^n$ on M. This action restricts to an $\mathbb{R}^n$ action on M_f. Let $p_0 \in M_f$, and let Γ be a lattice consisting of all vectors in $\mathbb{R}^n$ which fix p_0 under the group action. Then Γ is a finite subgroup of $\mathbb{R}^n$ and (by an intuitively clear modification of the orbit–stabiliser theorem) we have

$$M_f = \mathbb{R}^n / \Gamma.$$

Assuming that M_f is compact, this quotient space is diffeomorphic to a torus T^n.

□

In fact this argument shows that we get a torus for any choice of the constants c_k. Thus, varying the constants, we find that the phase–space M is foliated by n–dimensional tori.

Bibliography

[1] Ablowitz, M. J., Ramani, A., & Segur, H. (1980) A connection between nonlinear evolution equations and ordinary differential equations of P-type. I, II. J. Math. Phys. **21** 715–721 and 1006–1015.

[2] Ablowitz, M.J. & Clarkson, P.A. (1992) *Solitons, Nonlinear evolution equations and inverse scattering*, L.M.S. Lecture note series, **149**, CUP.

[3] Ablowitz, M. J. & Fokas, A. S. (2003) Introduction and Applications of Complex Variables, Cambridge University Press, second edition.

[4] Arnold, V. I. (1989) *Mathematical Methods of Classical Mechanics.*, second edition. Graduate Texts in Mathematics, **60**, Springer.

[5] Drazin, P. G., & Johnson, R. S. (1989) *Solitons: an introduction.* Cambridge Texts in Applied Mathematics. Cambridge University Press, Cambridge.

[6] Dunajski, M. (2009) *Solitons, Instantons and Twistors*, Oxford Graduate Texts in Mathematics **19**, OUP, Oxford.

[7] Gardner, C. Green, J., Kruskal, M. & Miura, R. (1967) Method for Solving the Korteweg-deVries Equation Phys. Rev. Lett. **19**, 1095.

[8] Gelfand, I. M & Levitan, B. M. (1951) On the determination of a differential equation from its spectral function Izv. Akad. Nauk SSSR, Ser. Mat. **15**, 309.

[9] Hydon P. E. (2000) *Symmetry Methods for Differential Equations: A Beginner's Guide*, CUP.

[10] Ince, E. L. (1956) *Ordinary Differential Equations*, Dover.

[11] Landau, L. D., & Lifshitz, E. M. (1995) *Course of Theoretical Physics, Vol I, II.* Butterworth-Heinemann.

[12] Lax, P. (1968), Integrals of nonlinear equations of evolution and solitary waves, Comm. Pure Applied Math. **21** 467-490.

[13] Magri, F. (1978) A simple model of the integrable Hamiltonian equation, J. Math. Phys., **19**, 1156-1162.

[14] Marchenko, V. A. (1955) Reconstruction of the potential energy from the phases of scattered waves Dokl. Akad. Nauk SSSR **104.**

[15] Novikov, S., Manakov, S. V., Pitaevskii, L. P. & Zakharov V. E. (1984) *Theory of Solitons: The Inverse Scattering Method*, Consultants Bureau, New York.

[16] Olver, P. J. (1993) *Applications of Lie groups to differential equations.* Springer-Verlag, New–York.

[17] Schiff, L, I (1969) *Quantum Mechanics*, 3rd ed. McGraw-Hill

[18] Schuster, H.G. (1988) *Deterministic Chaos: An Introduction*, second edition, VCH Publishers, New York.

[19] Woodhouse, N. M. J. (1987) *Introduction to analytical dynamics* OUP.

[20] Zaharov, V. E., & Shabat, A. B. (1979) Integration of the nonlinear equations of mathematical physics by the method of the inverse scattering problem. II. Funct. Anal. Appl. **13**, 13–22.