Solitons III
(2023-24)

Patrick Dorey (Michaelmas)
p.e.dorey@durham.ac.uk

Stefano Cremonesi (Epiphany)
stefano.cremonesi@durham.ac.uk

Department of Mathematical Sciences, Durham University

March 12, 2024

Contents

1 Introduction **2**
1.1 What is a soliton? . . . 2
1.2 The ball-and-box model . . . 9

2 Waves, dispersion and dissipation **13**
2.1 Dispersion . . . 13
2.2 Example: the Gaussian wave packet . . . 16
2.3 Dissipation . . . 19
2.4 Summary . . . 20

3 Travelling waves **21**
3.1 The KdV soliton . . . 22
3.2 The sine-Gordon kink . . . 25
3.3 A mechanical model for the sine-Gordon equation . . . 27
3.4 Travelling wave solutions and 1d point particles (bonus material) . . . 30

4 Topological lumps and the Bogomol'nyi bound **35**
4.1 The sine-Gordon kink as a topological lump . . . 35
4.2 The Bogomol'nyi bound . . . 37
4.3 Summary . . . 40

5 Conservation laws **42**
5.1 The basic idea . . . 43
5.2 Example: conservation of energy for sine-Gordon . . . 43
5.3 Conserved quantities for the KdV equation . . . 44
5.4 The Gardner transform . . . 46
5.5 Extra conservation laws for relativistic field equations . . . 50

6 Bäcklund transformations **57**
6.1 Definition . . . 58
6.2 A simple example . . . 58
6.3 The Bäcklund transformation for sine-Gordon . . . 60

6.4 First example: the sine-Gordon kink from the vacuum . . . 60
6.5 The theorem of permutability . . . 63
6.6 The two-soliton solution . . . 67
6.7 Asymptotics of multisoliton solutions . . . 68
6.8 The breather . . . 75

7 The Hirota method 79
7.1 Motivations . . . 79
7.1.1 Series solutions . . . 79
7.1.2 Some hints . . . 80
7.2 KdV equation in bilinear form . . . 81
7.2.1 The quadratic form of the KdV equation . . . 81
7.2.2 Hirota's bilinear operator . . . 82
7.3 Solutions . . . 84
7.3.1 Example: 1-soliton . . . 84
7.3.2 The N-soliton solution (sketch) . . . 85
7.4 Asymptotics of 2-soliton solutions and phase shifts . . . 89

8 Overview of the inverse scattering method 93
8.1 Initial value problems . . . 93
8.2 Linear initial value problems . . . 95
8.3 Outline of the method to solve the IVP for KdV . . . 97
8.3.1 The KdV-Schrödinger connection . . . 98
8.3.2 Recipe for constructing the KdV solution (in 3 steps) . . . 99

9 The basics of scattering theory 101
9.1 Overview: the physical interpretation . . . 101
9.2 Examples . . . 106
9.2.1 Zero potential: $V(x) = 0$ . . . 106
9.2.2 Delta function potential: $V(x) = a\delta(x)$ . . . 108
9.3 Reflectionless potentials . . . 113

10 Time evolution of the scattering data 119
10.1 Scattering data for general potentials . . . 119
10.2 Examples of scattering data . . . 121
10.3 The idea of a Lax pair . . . 123
10.3.1 The Lax pair for KdV . . . 126
10.4 Time evolution of the scattering data . . . 128

11 Interlude: the KdV hierarchy and conservation laws 131
11.1 Deriving the KdV equation (and generalising it) . . . 131
11.2 Connection with conservation laws . . . 136
11.2.1 The functional derivative . . . 137

12 Inverse scattering (or "reassembly") 141

12.1 The recipe for inverse scattering: the Marchenko equation 142
12.2 Example 1: the single KdV soliton 144
12.3 Example 2: the N-soliton solution 146

13 Integrable systems in classical mechanics 149
13.1 The Lax pair for the simple harmonic oscillator 155
13.2 The Lax pair for the Toda lattice 157

Chapter 1

Introduction

1.1 What is a soliton?

To a **first approximation**, solitons are very special solutions of a special class of **non-linear** partial differential equations (PDEs), or **'wave equations'**. (We will provide a more technical definition shortly.)
You might know that field theories, or the partial differential equations (PDEs) that describe their equations of motion, have solutions which look like **waves**. Solitons are special solutions which are localised in space and therefore look like **particles**. That's the reason for suffix -on, as in electron, proton or photon.

The **historical discovery** of solitons occurred in 1834, when a young Scottish civil engineer named **John Scott Russell** was conducting experiments to improve the design of canal barges at the Union Canal in Hermiston, near Edinburgh, see figure 1.1. Accidentally, a rope pulling a barge snapped, and here is what happened next in the words of John Scott Russell himself [John Scott Russell, 1845]:

> “ 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

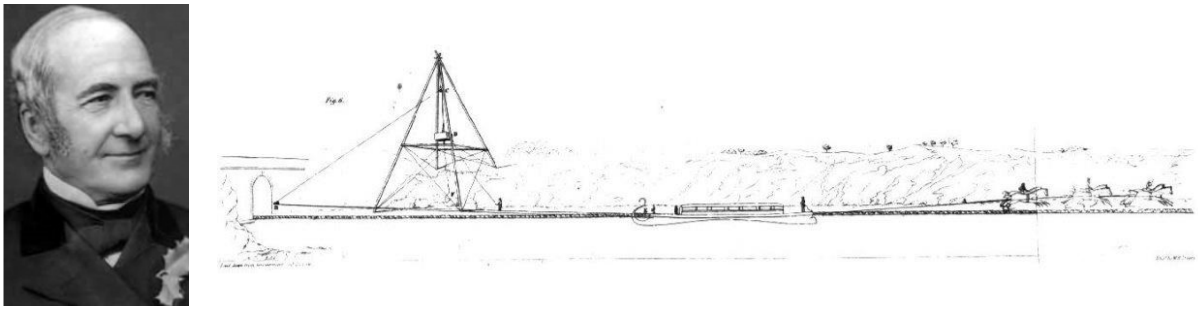


Figure 1.1: John Scott Russell, portrayed at a later time, and an artist's impression of the initial condition of his observation in 1834 (with a liberal interpretation of a 'pair of horses').

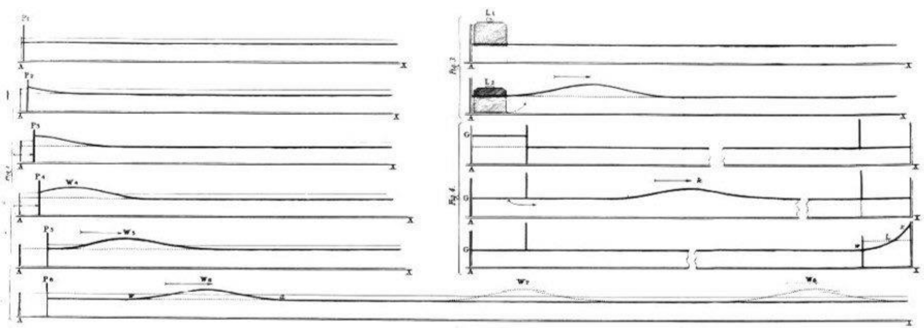


Figure 1.2: A depiction of two experiments carried out by John Scott Russell to recreate the Wave of Translation and study its properties.

> 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.
>
> *John Scott Russell*

As we will appreciate in the coming chapters, this solitary Wave of Translation behaves very differently from the ordinary waves which solve linear differential equations, which are a good approximation when interactions are small. Different linear waves can be added up ("superimposed") to obtain any wave profile, but these different linear waves travel at different speeds which depend on their wavelengths. As a result, any localised wave profile which is the superposition of various linear waves will "disperse" and lose its shape over time, because it consists of several linear waves which travel at different speeds. Russell's **"Wave of Translation"**, which is now called a **"soliton"** using a term coined by [Zabusky and Kruskal, 1965], behaved very differently, maintaining its shape unaltered over a surprisingly long time. Convinced that his observation was very important, John Scott Russell followed it up by a number of experiments in which he recreated his waves of translation and studied their properties, see figure 1.2. His results were published ten years later in the report [John Scott Russell, 1845],

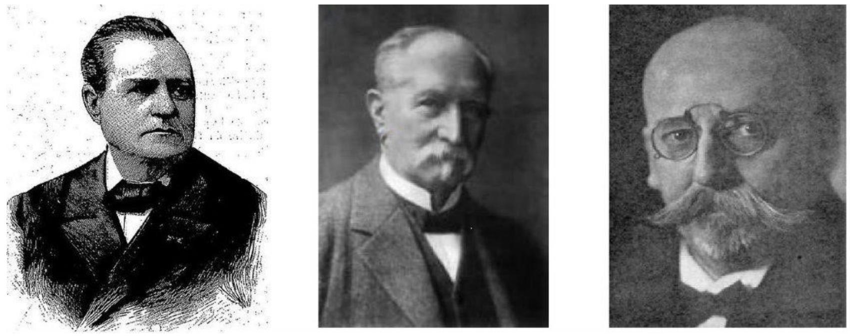


Figure 1.3: From left to right: Joseph Valentin Boussinesq, Diederik Korteweg and Gustav de Vries.

but much to his chagrin the scientific community paid little attention.

It took a few decades before a mathematical equation that describes shallow water waves and captures the peculiar phenomenon observed by John Scott Russell was introduced. The equation was first written down by the French mathematician and physicist Joseph Valentin Boussinesq [Boussinesq, 1877], and was then independently rediscovered by the Dutch mathematicians Diederik Korteweg and Gustav de Vries [Korteweg and Vries, 1895], see figure 1.3. According to the principle that things in science are named after the last people to discover them, this equation is now known as the

- **KORTEWEG-DE VRIES (KdV) EQUATION (1895):**

$$\boxed{u_t + 6uu_x + u_{xxx} = 0}\,. \tag{1.1}$$

  This is a short-hand for the partial differential equation

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

  for the real 'field' $u(x, t)$, which represents the height of a wave (measured from the surface of water at rest) in one space dimension x at time t. This equation:

  - describes long wavelength shallow water waves propagating in one space dimension;

  - captures the properties observed by John Scott Russell;

  - is a subtle limit of the equation describing real water waves propagating in one space dimension, in coordinates moving with the wave (see [Drazin and Johnson, 1989] for details if you are interested).

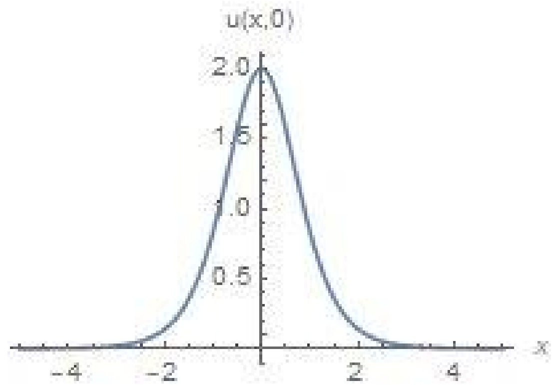


Figure 1.4: Plot of the initial condition $u(x,0) = 2\,\mathrm{sech}^2 x$ for the KdV equation.

REMARKS on the KdV equation:

1. Non-linear $\Longrightarrow$ Superposition principle fails. (Superposition principle: if u_1 and u_2 are solutions then so is $a_1 u_1 + a_2 u_2$ for all constants a_1, a_2)

2. 1st order in $t \Longrightarrow$ Solution determined by **initial condition** $u(x, 0)$.

3. Looks simple, but hides a rich mathematical structure.

We'll start by investigating the time evolution of the localised initial condition plotted in figure 1.4,

$$\boxed{u(x,0) = \frac{2}{\cosh^2(x)}}\,, \tag{1.2}$$

with the help of a computer. To gain some intuition, let's look at the KdV equation (1.1) piece by piece:

1. Drop the non-linear term $6uu_x$, to obtain the **LINEARISED KdV EQUATION**:

$$u_t + u_{xxx} = 0\ . \tag{1.3}$$

See an animation of the time evolution here. The initial localised wave **disperses**, *i.e.* it spreads out to the left, and $u \to 0$ as $t \to +\infty$ for any fixed x.

2. Drop the dispersive term u_{xxx}, to obtain the **DISPERSIONLESS KdV EQUATION**:

$$u_t + 6uu_x = 0\ . \tag{1.4}$$

In this case non-linearity causes the wave to **pile up** and **break** after a finite time: $|u_x| \to \infty$ as $t \to \sqrt{3}/16 \simeq 0.108$, which can be computed using the method of characteristics. Read this if you are interested in the calculation of the breaking time and see an animation of the time evolution here (the high frequency oscillations near the breaking point are an artifact of the numerical approximation).

3. Keep all terms to recover the **KdV EQUATION**:

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

The two previous effects cancel and we get a **"travelling wave"**, which keeps its form and just moves to the right, as you can see here.

Admittedly, the initial condition that we chose in (1.2) was very special. Generic solutions of KdV have a much more complicated behaviour (indeed equations (1.3)-(1.4) and their solutions are recovered in certain limits). Let us then experiment with a slightly more general class of initial conditions:

$$\boxed{u(x,0) = \frac{N(N+1)}{\cosh^2(x)}} \, , \qquad N > 0, \tag{1.5}$$

which reduces to the previous initial condition (1.2) if $N = 1$. Animations of the time evolution of the initial condition (1.5) under the KdV equation, for N ranging from 0.25 to 4, are here. [1]

These numerical experiments indicate that:[2]

- N **integer**:

  the initial wave splits into N **solitons** moving to the right with no dispersion.

- N **not integer**:

  the initial wave splits into $\lceil N \rceil$ **solitons** moving to the right plus **dispersing waves**, where $\lceil N \rceil$ denotes the least integer greater than or equal to N (this is called the *ceiling function*) .

- Either way, the different solitons move at different speeds. It can be checked that

$$\begin{array}{rcl} \text{SPEED} & \propto & \text{HEIGHT} \\ \text{WIDTH} & \propto & (\text{HEIGHT})^{-1/2} \end{array}$$

  in agreement with John Scott Russell's empirical observations.[3]

[1] Note: in this animation space has been compactified to a circle using periodic boundary conditions $u(10,t) = u(-10,t)$. This allows us to investigate what happens when two solitons collide. This will be briefly discussed below, and we will return to this specific feature in greater detail later.

[2] We will derive these results analytically later.

[3] Roughly, KdV solitons only move to the right because the limit of the physical wave equation that leads

One more feature is visible if one works with periodic spatial boundary conditions (BC), in which space is a circle, as was assumed in the previous animations: faster solitons catch up with and overtake slower solitons, with seemingly no final effect on their shapes! This is very surprising for a non-linear equation, for which the superposition principle does not hold. Note also that something funny happens during the overtaking: the height of the wave decreases, unlike for linear equations where different waves add up. This unusual behaviour was first observed in experiments by John Scott Russell, who was convinced that this was very important. It took a long time for the mathematics necessary to understand this phenomenon to develop and for the scientific community to fully come on board with John Scott Russell.[4]

We can summarize the previous observations in the following working definition of a soliton, that we will use in the rest of the course:

A **<u>SOLITON</u>** is a solution of a non-linear wave equation (or PDE) which:

1. IS LOCALISED
   (It's a "lump" of energy)

2. KEEPS ITS LOCALISED SHAPE OVER TIME
   (It moves with constant shape and velocity in isolation)

3. IS PRESERVED UNDER COLLISIONS WITH OTHER SOLITONS
   (If two or more solitons collide, they re-emerge from the collision with the same shapes and velocities.)

Watch this video (tip: turn down the volume) of water solitons created in a lab, which obey the previous defining properties to a very good approximation.

---

to the KdV equation involves switching to a reference frame which moves together with the fastest possible left-moving waves. Relative to that reference frame, all other waves move to the right.

[4]The modern revival of solitons was kickstarted by the numerical and analytical results of [Zabusky and Kruskal, 1965], who built on the earlier important numerical work of Fermi, Pasta, Ulam and Tsingou [Fermi et al., 1955]. (The paper of Fermi et al. was based on the first ever computer-aided numerical experiment, done on the MANIAC computer at Los Alamos [Porter et al., 2009]. Mary Tsingou's role in coding the problem was neglected for a long time and has only received the attention it deserves in recent years [Dauxois, 2008].)

It was universally expected at the time that in any non-linear physical system and for any initial conditions, interactions would spread the energy of the system evenly among all its degrees of freedom over time ('thermalisation' and 'equipartition of energy') and cause the system to explore all its available configurations ('ergodicity'). This process is what makes thermodynamics and statistical mechanics work.

Fermi et al. set out to study a system of non-linearly coupled oscillators numerically, with the aim of observing how thermalisation occurs. The system initially appeared to thermalise as expected, but to their great surprise they observed that it developed close-to-periodic (rather than ergodic) behaviour over longer time scales. A decade later, Zabusky and Kruskal showed that the system studied by Fermi et al. is approximated in a certain limit by the KdV equation, whose very special properties can explain the surprising behaviour of the system.

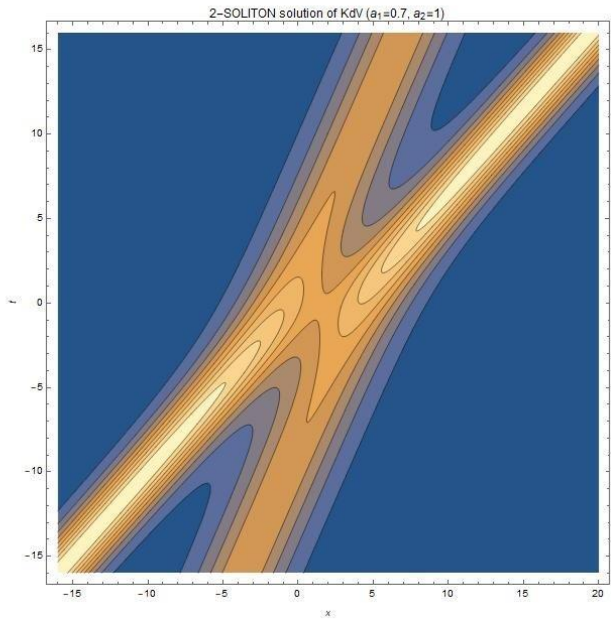


Figure 1.5: Contour plot of the energy density of two colliding KdV solitons, as a function of space and time. Lighter regions have higher energy density and correspond to the cores of the two solitons. We can see the trajectories of the two solitons and the phase shift induced by the collision: the faster soliton is advanced, while the slower soliton is retarded by the collision.

Solitons are not just objects of purely academic interest. They can appear in nature under a variety of circumstances. For instance, here is a video of the Severn bore taken on the 2019 spring equinox: as the high tide coming from the Atlantic Ocean enters the funnel-shaped estuary of the Severn, water surges forming highly localised waves which travel (and can be surfed!) for several miles into the Bristol Channel.

REMARKS:

- Property 3 does not mean that nothing happens to solitons which collide: as we will study towards the end of the term, the effect of the collision is to advance or retard the solitons by a so-called "phase shift". As an example, in figure 1.5 we can see the trajectories of two colliding KdV solitons and the phase shifts resulting from their interaction.

- Only very special field theories (or equivalently, wave equations) admit solitons as defined above. They are called **integrable** and are usually defined in 1 space + 1 time dimensions. Property 3 is the key. (Some people use the term "integrable soliton" for the above definition, but we will stick with "soliton" in this course.)

Solitons have been studied in depth since the 1960s in relation to many contexts:

- **Applied Maths**: water waves, optical fibres, electronics, biological systems...
- **High Energy Physics**: particle physics, gauge theory, string theory...
- **Pure Maths**: special functions, algebraic geometry, spectral theory, group theory...

We will consider two main examples of integrable soliton equations in this course:

$$\textbf{KdV}: \qquad \boxed{u_t + 6uu_x + u_{xxx} = 0} \tag{1.6}$$

$$\textbf{sine}-\textbf{Gordon}: \qquad \boxed{u_{tt} - u_{xx} = -\sin u} \tag{1.7}$$

- **THIS TERM**: we will (mostly) study simple **pure solitons with no dispersion**.
- **NEXT TERM**: you will study **"inverse scattering"**, a powerful formalism that allows an analytical understanding of the time evolution of **generic initial conditions**.[5]

To get a better feel for solitons before we start, let's consider a discrete model which displays solitons but no dispersion. This is an example of a "cellular automaton", a zero-player game where the rules for time evolution are fixed and the only freedom is in the choice of initial condition, but in which surprisingly rich patterns can develop.[6]

1.2 The ball-and-box model

This term we will learn several analytic methods to generate single and multiple soliton solutions of non-linear differential equations like KdV, and study the properties of these solutions.

As we have seen, experimenting with these equations on a computer can be very useful to develop intuition about the properties of solitons. The trouble is that you need a big-ish computer for most of these numerical experiments.

[5] The inverse scattering formalism was designed for equations in which space is the real line, but it is also useful if space is a finite interval or a circle (periodic bc). *E.g.* a sinusoidal initial condition on a circle evolves into a train of solitons [Zabusky and Kruskal, 1965], see this animation. Here is a contour plot of the energy density, showing the trajectories of the various solitons, which after a while recombine into a sinusoidal wave, leading to the periodic behaviour discussed in footnote4.

[6] The most famous cellular automaton is perhaps John Conway's Game of Life. Read about it here if you have never heard of it. If you search Conway's game of life or cellular automata on YouTube you will enter a rabbit hole of cool videos, often accompanied by an electronic music soundtrack. Too bad that we won't study those cellular automata further in this course, apart from the simple model which is the subject of next section.

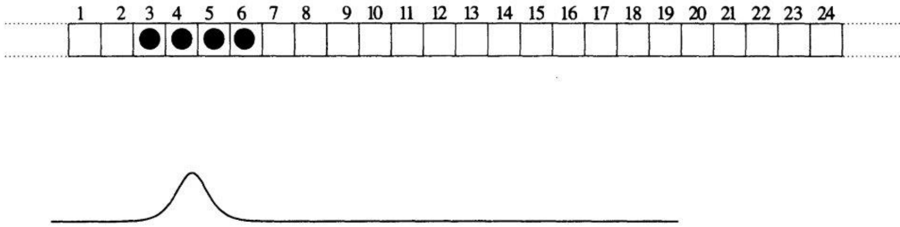


Figure 1.6: A localised configuration of the ball and box model and its continuous analogue.

Fortunately, it was realised around 1990 that many properties of continuous solitons can be mimicked by **much simpler discrete models**, which can be studied by drawing pictures with **pen and paper**. A nice and simple example is the **BALL AND BOX MODEL** of [Takahashi and Satsuma, 1990]. In this model, **space and time** are **discrete**. In particular:

- Continuous space is replaced by an infinite line of boxes, labelled by a position $i \in \mathbb{Z}$
- At any instant $t \in \mathbb{Z}$, the configuration of the system is specified by filling a number of boxes with one ball each, as in figure 1.6.
- Time evolution $t \to t + 1$ is governed by the

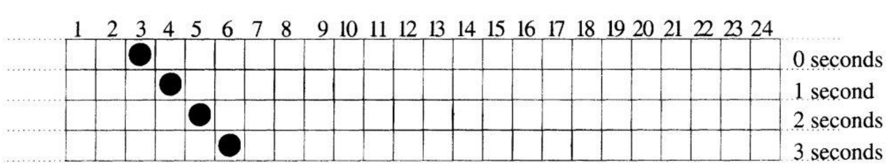
<u>BALL AND BOX RULE</u>:

Move the leftmost ball to the next empty box to its right. Repeat the process until all balls have been moved exactly once. When you are done, the system has been evolved forward one unit in time.

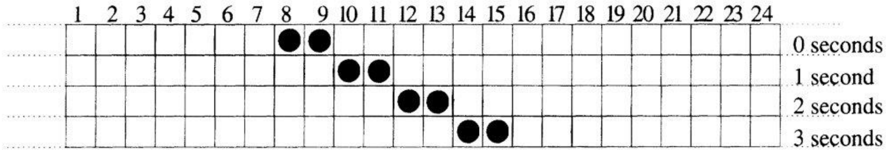
The ball and box rule plays the role of the PDE for continuous solitons, *e.g.* $u_t = -6uu_x - u_{xxx}$ in the case of the KdV equation.

<u>EXAMPLES</u>:

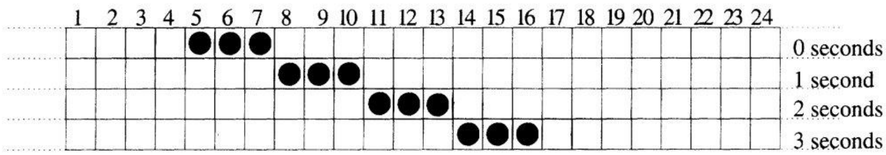
- **1 ball**:



- **2 consecutive balls**:



- **3 consecutive balls**:

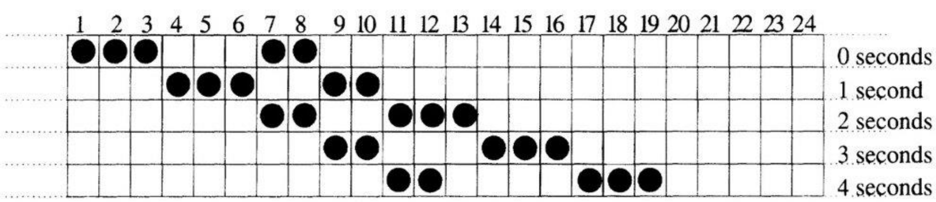


We see that a sequence of n **consecutive balls** behaves like a soliton: it keeps its shape and translates by n boxes in one unit of time. So for this class of solitons

$$\text{SPEED} \quad = \quad \text{LENGTH}\,,$$

where we define the speed as the length travelled per unit time.

So far we have only checked that the defining properties 1 and 2 of a soliton are obeyed by a sequence of consecutive balls. To check the remaining property 3, let us consider what happens when a longer (=faster) soliton is behind(=to the left of) a shorter (=slower) soliton. After a while the faster soliton will catch up and collide with the slower soliton. What happens next? Let's look at an example with a length-3 soliton following a length-2 soliton:



The two solitons keep the **same shape** after the collision, but their order is reversed: the faster soliton has overtaken the slower one. If we look carefully, we can also notice that the positions of the two solitons are **delayed/advanced** by a finite amount compared to the positions that each soliton would have had in the absence of the other soliton. This spatial advance or delay is an example of a **"phase shift"**; it is for a soliton which is advanced and negative for a soliton which is retarded. In the previous example the length-3 soliton has a phase shift of $+4$ and the length-2 soliton has a phase shift of -4. [Make sure that you understand how this phase shift is computed from the previous figure!] This is analogous to the phase shift visible in figure 1.5 in the scattering of continuous KdV solitons.

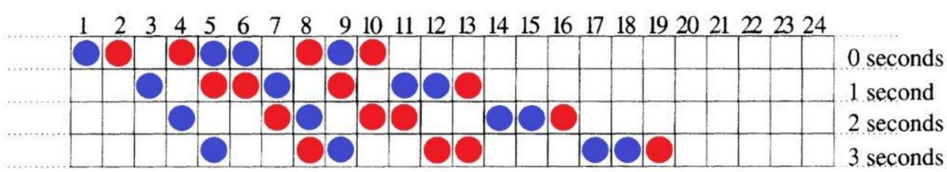
∗ EXERCISE: Generalize the previous example to a length m soliton overtaking a length n soliton (with $m > n$) and find a general rule for what happens. (Start with separation $l \geqslant n$ between the two solitons, that is, there are l empty boxes between the two solitons in the initial configuration.) **[Ex 4]**

The ball and box model can be generalized by introducing balls of different colours. For instance, in the **2-COLOUR BALL AND BOX MODEL**, balls come in two colours (say **BLUE** and **RED**), and again each box can be filled by at most one ball, of either colour.[7] The time evolution $t \to t + 1$ is governed by the

2-COLOUR BALL AND BOX RULE:

Move the leftmost **BLUE** ball to the next empty box to its right. Repeat the process until all **BLUE** balls have been moved exactly once. Then do the same for the **RED** balls. When all the **BLUE** and **RED** balls have been moved, the system has been evolved forward by one unit of time.

EXAMPLE:



∗ EXERCISE: Can you classify solitons in the 2-colour ball and box model? **[Ex 5]**
What happens when solitons collide? **[Ex 7*]**
(Starred exercises are for the bravest.)

Next, we will return to continuous wave equations and aim to make the phenomenon of **dispersion** more precise.

[7] If you happen to be colour blind and this part of the notes is not accessible, please let me know and I'll replace the two colours by different symbols.

Chapter 2

Waves, dispersion and dissipation

The main reference for this chapter is §1.1 of the book [Drazin and Johnson, 1989].

2.1 Dispersion

Usually, localised waves **spread out** (**"disperse"**) as they travel. This prevents them from being solitons. Let's understand this phenomenon first.

EXAMPLES:

1. **ADVECTION EQUATION** (linear, 1st order):

$$\boxed{\frac{1}{v}u_t + u_x = 0} \tag{2.1}$$

$\longrightarrow$ Solution

$$u(x,t) = f(x - vt) \qquad \text{for any function } f \,,$$

i.e. a wave moving with velocity v (right-moving if $v > 0$, left-moving if $v < 0$). The wave keeps a fixed profile $f(\xi)$ and moves rigidly at velocity v (indeed $\xi = x - vt$):

So in this case there is no dispersion, but nothing else happens either.

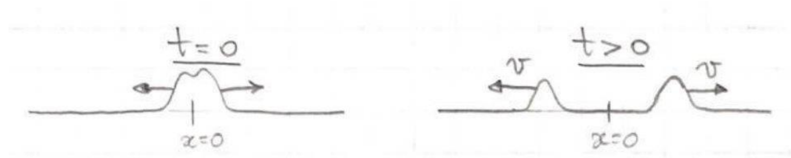
2. "THE" **WAVE EQUATION** or **D'ALEMBERT EQUATION** (linear, 2nd order):

$$\boxed{\frac{1}{v^2}u_{tt} - u_{xx} = 0} \qquad (v > 0 \text{ wlog}) \tag{2.2}$$

$\longrightarrow$ Solution

$$u(x,t) = f(x - vt) + g(x + vt) \qquad \text{for any functions } \ f,\ g\ ,$$

i.e. the superposition of a right-moving and a left-moving wave with velocities $\pm v$:



All waves move at the **same** speed, so there is no dispersion, but there is no interaction either, so this is also not very interesting for our purposes.

3. **KLEIN-GORDON EQUATION**[1] (linear, 2nd order):

$$\boxed{\frac{1}{v^2}u_{tt} - u_{xx} + m^2 u = 0}\ , \tag{2.3}$$

where we take $v > 0$ wlog.
This is a more interesting equation. Let us try a complex **"plane wave"** solution[2]

$$\boxed{u(x,t) = e^{i(kx-\omega t)}}\ . \tag{2.4}$$

Substituting the plane wave (2.4) in the Klein-Gordon equation (2.3), we find:

$$\begin{aligned} -\frac{\omega^2}{v^2}e^{i(kx-\omega t)} + k^2 e^{i(kx-\omega t)} + m^2 e^{i(kx-\omega t)} &= 0 \\ \Longrightarrow \quad -\frac{\omega^2}{v^2} + k^2 + m^2 &= 0\ . \end{aligned}$$

[1]This is the first relativistic wave equation (with v the speed of light). It was introduced independently by Oskar Klein [Klein, 1926] and Walter Gordon [Gordon, 1926], who hoped that their equation would describe electrons. It doesn't, but it describes massive elementary particles without spin, like the pion or the Higgs boson.

[2]This is called a "plane wave" because its three-dimensional analogue $u(\vec{x},t) = \exp[i(\vec{k}\cdot\vec{x} - \omega t)]$ has constant u along a plane $\vec{k}\cdot\vec{x} = \text{const}$ at fixed t. Unless specified, in this course we are interested in **real fields** u. It is nevertheless convenient to use complex plane waves (2.4) and eventually take the real or imaginary part to find a real solution, rather than working with the real plane waves $\cos(kx - \omega t)$ and $\sin(kx - \omega t)$ from the outset.

So the plane wave (2.4) is a solution of the Klein-Gordon equation (2.3) provided that ω satisfies

$$\boxed{\omega = \omega(k) = \pm\, v\,\sqrt{k^2+m^2}}\ . \tag{2.5}$$

We will usually ignore the sign ambiguity and only consider the $+$ sign in (2.5) and similar equations.[3]

VOCABULARY:

k	**wavenumber**	$\lambda = \frac{2\pi}{k}$	**wavelength** (periodicity in x)
ω	**angular frequency**	$\tau = \frac{2\pi}{\omega}$	**period** (periodicity in t)

A formula like (2.5) relating ω to k: **dispersion relation**.

The maxima of a real plane wave, like for instance $\mathrm{Re}\, e^{i(kx-\omega(k)t)}$ or $\mathrm{Im}\, e^{i(kx-\omega(k)t)}$, are called **"wave crests"**. By a slight abuse of terminology, we will refer to the wave crests of the real or imaginary part of a complex plane wave like (2.4) simply as the wave crests of the complex plane wave.

By rewriting the complex plane wave solution (2.4) of the Klein-Gordon equation as $e^{ik(x-c(k)t)}$, we see that its wave crests move at the velocity

$$c(k) = \frac{\omega(k)}{k} = v\sqrt{1+\frac{m^2}{k^2}}\ \mathrm{sign}(k)\ .$$

Plane waves with **different wavenumbers** move at **different velocities**, so if we try to make a lump of real Klein-Gordon field by superimposing different plane waves

$$\boxed{u(x,t) = \ \mathrm{Re}\int_{-\infty}^{+\infty} dk\ f(k)\ e^{i(kx-\omega(k)t)}}\ , \tag{2.6}$$

it will **disperse**.

In fact, there are two different notions of velocity for a wave:

- **PHASE VELOCITY**

$$\boxed{c(k) = \frac{\omega(k)}{k}}\ , \tag{2.7}$$

which is the velocity of wave crests.

[3]We do not lose generality here, since we can obtain the plane wave solution with opposite ω by taking the complex conjugate plane wave solution and sending $k \to -k$.

- **GROUP VELOCITY**

$$\boxed{c_g(k) = \frac{d\omega(k)}{dk}}\,, \tag{2.8}$$

which is the velocity of the lump of field while it disperses.

We will understand better the relevance of the group velocity in the next section.

<u>**REMARK**</u>:

The energy (and information) carried by a wave travels at the **group velocity**, not at the phase velocity. For a **relativistic wave equation** with **speed of light** v, no signals can be transmitted faster than the speed of light. So it should be the case that $\boxed{|c_g(k)| \leqslant v}$ for all wavenumbers k, but there is no analogous bound on the phase velocity. For example, for the Klein-Gordon equation (2.3), we can calculate

- |Group velocity|:

$$|c_g(k)| = \left|\frac{d\omega(k)}{dk}\right| = \frac{v}{\sqrt{1+\frac{m^2}{k^2}}} \leqslant v$$

consistently with the principles of relativity.

- |Phase velocity|:

$$|c(k)| = \left|\frac{\omega(k)}{k}\right| = v\,\sqrt{1+\frac{m^2}{k^2}} \geqslant v\,,$$

which is faster than the speed of light v for all k, but this is not a problem.

2.2 Example: the Gaussian wave packet

The simplest example of a localised field configuration obtained by superposition of plane waves is the "GAUSSIAN WAVE PACKET", which is obtained by choosing a Gaussian

$$f(k) = e^{-a^2(k-\bar{k})^2} \qquad (a > 0,\ \bar{k} \in \mathbb{R})$$

in the general superposition (2.6). This represents a lump of field with

average wavenumber	$\bar{k}$
spread of wavenumber	$\sim 1/a$,

see fig. 2.1.

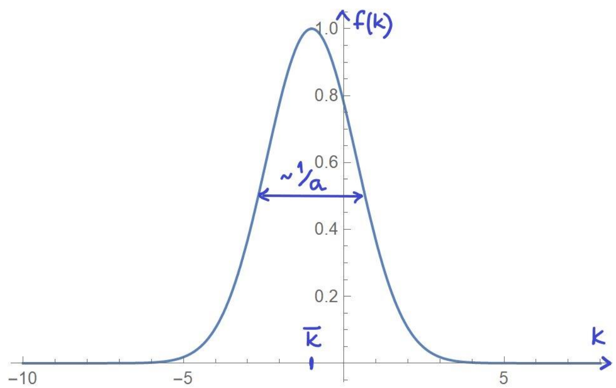


Figure 2.1: Gaussian wavepacket in Fourier space.

Then $u(x,t) = \operatorname{Re} z(x,t)$ is a real solution of the Klein-Gordon equation, where

$$\boxed{z(x,t) = \int_{-\infty}^{+\infty} dk\, e^{-a^2(k-\bar{k})^2} e^{i(kx-\omega(k)t)}}\,, \tag{2.9}$$

provided that $\omega(k) = v\,\sqrt{k^2+m^2}$.[4]

Since most of the integral (2.9) comes from the region $k \approx \bar{k}$, we can obtain a good approximation to (2.9) by Taylor expanding $\omega(k)$ about $k = \bar{k}$. Expanding to first order in $(k-\bar{k})$ we obtain

$$\begin{aligned}
\omega(k) &= \omega(\bar{k}) + \omega'(\bar{k}) \cdot (k-\bar{k}) + \mathcal{O}((k-\bar{k})^2) \\
&= \omega(\bar{k}) + c_g(\bar{k}) \cdot (k-\bar{k}) + \mathcal{O}((k-\bar{k})^2) \\
&\approx \omega(\bar{k}) + c_g(\bar{k}) \cdot (k-\bar{k})\ ,
\end{aligned}$$

where in the second line we used (2.8) and in the third line we introduced a short-hand $\approx$ to

[4] $z(x,t)$ is a complex solution of the Klein-Gordon equation. Since the Klein-Gordon equation is a linear equation with real coefficients, the complex conjugate $z(x,t)^*$ is also a solution of the Klein-Gordon equation, as are $\operatorname{Re} z(x,t)$ and $\operatorname{Im} z(x,t)$.

avoid writing $\mathcal{O}((k-\bar{k})^2)$ every time. Substituting in (2.9), we find

$$
\begin{aligned}
z(x,t) &\approx \int_{-\infty}^{+\infty} dk\; e^{-a^2(k-\bar{k})^2} e^{i\{kx-[\omega(\bar{k})+c_g(\bar{k})\cdot(k-\bar{k})]t\}} \\
&= e^{i[\bar{k}x-\omega(\bar{k})t]} \int_{-\infty}^{+\infty} dk\; e^{-a^2(k-\bar{k})^2} e^{i(k-\bar{k})[x-c_g(\bar{k})t]} \\
&\underset{k\to k+\bar{k}}{=} e^{i[\bar{k}x-\omega(\bar{k})t]} \int_{-\infty}^{+\infty} dk\; e^{-a^2k^2+ik[x-c_g(\bar{k})t]} \\
&\underset{\substack{\text{complete}\\ \text{the square}}}{=} e^{i[\bar{k}x-\omega(\bar{k})t]} e^{-\frac{1}{4a^2}[x-c_g(\bar{k})t]^2} \int_{-\infty}^{+\infty} dk\; e^{-a^2\left\{k-\frac{i}{2a^2}[x-c_g(\bar{k})t]\right\}^2} \\
&\underset{\substack{\text{Gaussian}\\ \text{integral}}}{=} \underbrace{e^{i[\bar{k}x-\omega(\bar{k})t]}}_{\text{CARRIER WAVE}} \cdot \underbrace{\frac{\sqrt{\pi}}{a} e^{-\frac{1}{4a^2}[x-c_g(\bar{k})t]^2}}_{\text{ENVELOPE}} ,
\end{aligned}
$$

where in the second line we factored out a plane wave with $k = \bar{k}$, in the third line we changed integration variable replacing k by $k+\bar{k}$, in the fourth line we completed the square $Ak^2 + Bk = A(k+\frac{B}{2A})^2 - \frac{B^2}{4A}$, and in the last line we used the Gaussian integral formula

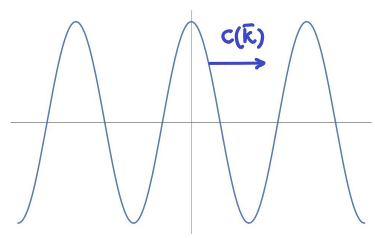
$$
\int_{-\infty+ic}^{+\infty+ic} e^{-Ak^2} = \sqrt{\frac{\pi}{A}} ,
$$

which holds for all $A > 0$ and $c \in \mathbb{R}$. The final result is the product of a:

1. **"CARRIER WAVE"**:

   a plane wave moving at the **phase velocity**

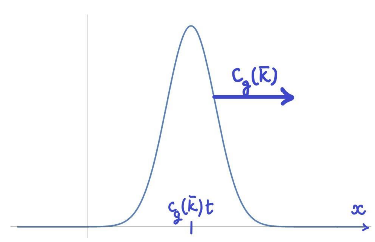
   $$c(\bar{k}) = \frac{\omega(\bar{k})}{\bar{k}}$$



2. **"ENVELOPE"**:

   a localised profile (or "wave packet") moving at the **group velocity**

   $$c_g(\bar{k}) = \omega'(\bar{k}) .$$



Click here to see an animation of a Gaussian wavepacket with a (Gaussian) envelope and a carrier wave moving at different velocities. In the animation the phase velocity is much larger than the group velocity.

To this order of approximation, the spatial **width** of the lump has the *parametric dependence*

$$\text{WIDTH} \quad \sim \quad a\,,$$

meaning that the width doubles if a is doubled, and is constant in time. (Indeed, a simultaneous rescaling of $x - c_g(\bar{k})t$ and a by the same constant λ leaves the envelope invariant.)

*** EXERCISE**: Improve on the previous approximation by including the 2nd order in $k - \bar{k}$. Show that **[Ex 10]**

$$\text{WIDTH}^2 \quad \sim \quad a^2 + \tfrac{\omega''(\bar{k})}{4a^2}t^2$$

and that the amplitude of the wave packet also decreases as time increases.

This leads to the phenomenon of **DISPERSION**, whereby the profile of the wave packet changes as it propagates. In particular, starting from a localised wave packet, dispersion makes the wave packet spread out: the width of the initial wave packet grows and the amplitude decreases as time increases. See this animation of the time evolution of the Gaussian wave-packet up to second order in $(k - \bar{k})$.

2.3 Dissipation

So far we have considered wave equations which lead to a real dispersion relation, so $\omega(k) \in \mathbb{R}$. If instead $\omega(k) \in \mathbb{C}$, then a new phenomenon occurs: **DISSIPATION**, where the **amplitude** of the wave **decays (or grows) exponentially in time**. For a plane wave

$$u(x,t) = e^{i(kx - \omega(k)t)} = e^{i(kx - \operatorname{Re}\omega(k)\cdot t))}e^{\operatorname{Im}\omega(k)\cdot t} \tag{2.10}$$

and we have two cases:

- $\operatorname{Im}\omega(k) < 0$: **“PHYSICAL DISSIPATION”**
  The amplitude **decays** exponentially with time.

- $\operatorname{Im}\omega(k) > 0$: **“UNPHYSICAL DISSIPATION”**
  The amplitude **grows** exponentially with time (physically unacceptable).

EXAMPLES:

1.
$$\boxed{\frac{1}{v}u_t + u_x + \alpha u = 0} \qquad (\alpha > 0,\ v > 0) \tag{2.11}$$

Sub in a plane wave $u = e^{i(kx-\omega t)}$:

$$-i\frac{\omega}{v} + ik + \alpha = 0 \quad\Longrightarrow\quad \omega(k) = v(k - i\alpha)\,,$$

leading to a **complex** dispersion relation. The plane wave solution is therefore

$$u(x,t) = e^{ik(x-vt)}e^{-\alpha vt}$$

and the wave decays exponentially, or **"dissipates"**, to zero as $t \to +\infty$. This is an example of physical dissipation. ($\alpha v < 0$ would have led to unphysical dissipation.)

2. **HEAT EQUATION:**

$$\boxed{u_t - \alpha u_{xx} = 0} \qquad (\alpha > 0) \tag{2.12}$$

*** EXERCISE**: Sub in a plane wave and derive the dispersion relation $\omega(k) = -i\alpha k^2$.

So the plane wave solution of the heat equation is

$$u(x,t) = e^{ikx}e^{-\alpha k^2 t}$$

and the waves dissipates as time passes.

2.4 Summary

- **Linear** wave equation $\longrightarrow$ (Complex) **plane wave** solutions $u = e^{i(kx-\omega t)}$.

  Sub in to get $\omega = \omega(k)$ **dispersion relation**.

- Wave crests move at $c(k) = \omega(k)/k$ **phase velocity**.

  (If $\omega(k) \in \mathbb{C}$, then we define the phase velocity as $c(k) = \operatorname{Re}\omega(k)/k$.)

- Lumps of field move at $c_g(k) = \omega'(k)$ **group velocity**.
  /wave packets

  (If $\omega(k) \in \mathbb{C}$, then we define the group velocity as $c_g(k) = \operatorname{Re}\omega'(k)$.)

- **Dispersion** (real ω, width increases and amplitude decreases) and **dissipation** (complex ω, amplitude decreases exponentially) smooth out and destroy localised lumps of energy in linear wave (or field) equations.

- **Non-linearity** can have an opposite effect (steepening and breaking, see chapter 0).

- For **solitons** the competing effects counterbalance one another precisely, leading to stable lumps of energy, unlike for ordinary waves.

Chapter 3

Travelling waves

The main references for this chapter are §2.1-2.2 of [Drazin and Johnson, 1989] and §2.1 of [Dauxois and Peyrard, 2006].

A **“TRAVELLING WAVE”** is a solution of a wave equation of the form

$$\boxed{u(x,t) = f(x - vt)}\ ,$$

where f is a function of a single variable, which we will typically denote by $\xi := x - vt$. The **velocity** v of the travelling wave could either be:

1. **Fixed** in terms of a **parameter** appearing in the **wave equation**, as in d'Alembert's general solution
$$u(x,t) = f(x - vt) + g(x + vt)$$
of the wave equation
$$\frac{1}{v^2} u_{tt} - u_{xx} = 0\ ,$$
which is the linear superposition of two travelling waves with velocities $\pm v$.

2. A **free parameter of the solution**, as in the KdV soliton that we will derive shortly.

REMARK:
In some cases (*e.g.* “the” wave equation or the sine-Gordon equation) there will be both a velocity parameter appearing in the equation (*e.g.* the speed of light) and a different velocity parameter appearing in the travelling wave solution (namely, the speed of the wave). To avoid confusion, from now on the velocity parameter appearing in the wave equation will be set to

1 by an appropriate choice of units, and v will be reserved for the velocity of the travelling wave. For example, we will write "the" wave equation as $u_{tt} - u_{xx} = 0$ and d'Alembert's general solution as $u(x,t) = f(x-t) + g(x+t)$, which is the superposition of two travelling waves with velocities $v = \pm 1$.

3.1 The KdV soliton

We would like to find a travelling wave solution of the KdV equation

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

with **boundary conditions** (BC's)

$$\text{BC's}: \qquad u, u_x, u_{xx} \xrightarrow[x\to\pm\infty]{} 0$$

for all finite values of t. Let us accept these BC's for the time being; we will derive them later.

Substituting in the KdV equation the travelling wave ansatz $u(x,t) = f(x-vt) \equiv f(\xi)$ where $\xi = x - vt$, using the chain rule to express partial derivatives wrt x and t in terms of ordinary derivatives wrt ξ as follows,

$$\frac{\partial}{\partial x} = \frac{\partial \xi}{\partial x}\frac{d}{d\xi} = \frac{d}{d\xi}\,, \qquad \frac{\partial}{\partial t} = \frac{\partial \xi}{\partial t}\frac{d}{d\xi} = -v\frac{d}{d\xi}\,,$$

and using primes to denote derivatives wrt ξ, we obtain an ODE which we can integrate twice:

$$\begin{aligned} & -vf' + 6ff' + f''' = 0 \\ \underset{\int d\xi}{\Longrightarrow} \quad & -vf + 3f^2 + f'' = A \\ \underset{\int d\xi\, f'}{\Longrightarrow} \quad & -\frac{v}{2}f^2 + f^3 + \frac{1}{2}(f')^2 = Af + B\,, \end{aligned}$$

where A and B are integration constants. The second integration used an integrating factor f', as denoted by the short-hand $\int d\xi f'$.

We can determine the integration constants A and B by imposing the BC's, which imply that $f, f', f'' \to 0$ as $\xi \to \pm\infty$. Sending $\xi \to \pm\infty$ in the second and third line above we find[1]

$$\begin{aligned} \text{BC's:} \quad & A = B = 0 \\ \Longrightarrow \quad & (f')^2 = f^2(v - 2f) \\ \Longrightarrow \quad & f' = \pm f\sqrt{v - 2f} \\ \Longrightarrow \quad & \int \frac{df}{f\sqrt{v-2f}} = \pm\xi \equiv \pm(x - vt)\,. \qquad (*) \end{aligned}$$

[1] Always impose the boundary conditions carefully and keep in mind that they don't always imply that the integration constants vanish. This is a major source of mistakes in homework and exams.

where we note that we need $f \leqslant v/2$ to ensure that $f, f' \in \mathbb{R}$.

To calculate the integral obtained by separation of variables, we change integration variable

$$
\begin{aligned}
& & f &= \frac{v}{2}\mathrm{sech}^2\vartheta & (**)\\
\implies & & df &= -v\frac{\sinh\vartheta}{\cosh^3\vartheta}d\vartheta\ , \\
& & \sqrt{v-2f} &= \sqrt{v}\sqrt{1-\frac{1}{\cosh^2\vartheta}} = \pm\sqrt{v}\frac{\sinh\vartheta}{\cosh\vartheta} \\
\implies & & \frac{df}{f\sqrt{v-2f}} &= \mp\frac{v\frac{\sinh\vartheta}{\cosh^3\vartheta}\,d\vartheta}{\frac{v}{2}\frac{1}{\cosh^2\vartheta}\sqrt{v}\frac{\sinh\vartheta}{\cosh\vartheta}} = \mp\frac{2}{\sqrt{v}}d\vartheta\ . & (***)
\end{aligned}
$$

Substituting $(***)$ in $(*)$ and keeping in mind that the sign ambiguities arising from taking square roots in the two equations are unrelated (and therefore only the relative sign ambiguity matters), we find

$$
\begin{aligned}
-\frac{2}{\sqrt{v}}\int d\vartheta &= \pm(x-vt) \\
\implies \qquad \vartheta &= \pm\frac{\sqrt{v}}{2}(x-x_0-vt)\ ,
\end{aligned}
$$

where x_0 is an integration constant. Substituting in $(**)$ we find the travelling wave solution

$$
\boxed{u(x,t) = f(x-vt) = \frac{v}{2}\mathrm{sech}^2\left[\frac{\sqrt{v}}{2}(x-x_0-vt)\right]} \tag{3.1}
$$

where the sign ambiguity has disappeared because sech^2 is an even function.

The travelling wave solution (3.1) of the KdV equation is the **KdV SOLITON**. See 3.1 for a snapshot of the KdV soliton.

REMARKS:

- For a **real non-singular** solution we need $v \geqslant 0$, which means that KdV solitons only travel to the right.[2]

[2]For $v < 0$ the travelling wave solution just found is

$$
-\frac{|v|}{2}\sec^2\left[\frac{\sqrt{|v|}}{2}(x-x_0+|v|t)\right]\ ,
$$

which moves to the left with speed $|v|$. However it **diverges** wherever $[\dots] = \left(n+\frac{1}{2}\right)\pi$ with $n \in \mathbb{Z}$. We are always after real bounded solutions, so we discard this singular (or divergent) solution; it also fails to satisfy the given boundary conditions.

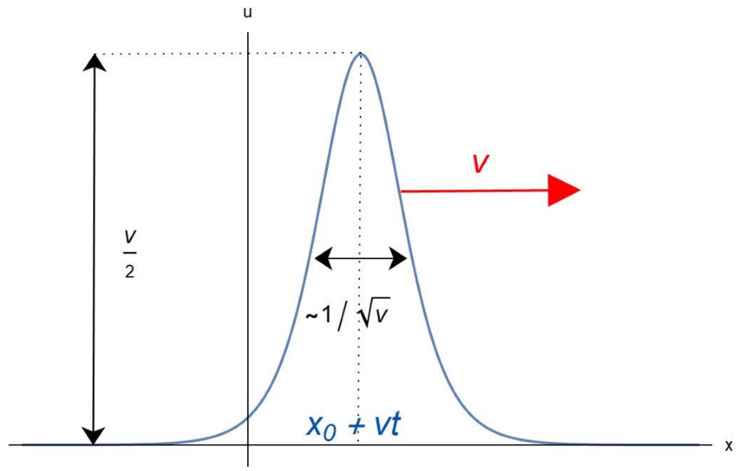


Figure 3.1: Snapshot of the KdV soliton.

- **PROPERTIES** of the KdV soliton:

VELOCITY	v
HEIGHT	$v/2$
WIDTH	$\sim \frac{1}{\sqrt{v}}$
CENTRE	$x_0 + vt$

Clarification:

What do I mean by WIDTH $\sim 1/\sqrt{v}$? A **possible** definition of the width of the soliton is as the distance between the two points where the value of u is reduced by a factor of e from its maximum, that is WIDTH $= |x_+ - x_-| \equiv 2\Delta x$ where $u(x_\pm) = v/(2e)$. For $\sqrt{v}\Delta x \gg 1$, we can approximate $\text{sech}^2\left(\frac{\sqrt{v}}{2}\Delta x\right) \approx 4e^{-\sqrt{v}\Delta x}$, therefore this definition of width would give

$$\text{WIDTH} = 2\Delta x \approx \frac{2}{\sqrt{v}}(1 + \log 4) \approx \frac{4.77}{\sqrt{v}} \ .$$

(Without the approximation one finds $4.34.../\sqrt{v}$.) However the above definition of width was somewhat arbitrary: for instance we could have looked at points where the value u is reduced by a factor of 2, or 3, or else, from its maximum. Given a precise definition of width, one can determine the precise coefficient of $1/\sqrt{v}$ above, but fixating on a precise definition would be somewhat absurd given the arbitrariness in the definition. It is better to say that **"the width is of the order of"** (or equivalently **"goes like"**) $1/\sqrt{v}$. This is independent of the precise definition of width and captures the essential point that the spatial coordinate x appears multiplied by $\sqrt{v}$ in the KdV soliton solution (3.1). We use $\sim$ to denote this **parametric dependence**. This is not to be confused with $\approx$, which means "is approximately equal to".

A final comment: if the BC's are changed to allow $A, B \neq 0$ (*e.g.* if we impose periodic boundary conditions, which is equivalent to solving the KdV equation on a circle), then the ODE for the travelling wave solution can still be integrated exactly using elliptic functions. See §2.4, 2.5 of [Drazin and Johnson, 1989] if you are interested.

3.2 The sine-Gordon kink

Let us seek a travelling wave solution the **sine-Gordon** equation

$$u_{xx} - u_{tt} = \sin u \ ,$$

where u is an angular variable u defined modulo 2π, subject to the boundary conditions

$$\text{BC's :} \qquad u \bmod 2\pi, \ u_x \xrightarrow[x\to\pm\infty]{} 0$$

for every finite t. (More about these BC's later.)

Substituting the travelling wave ansatz $u(x,t) = f(x-vt) \equiv f(\xi)$ in the sine-Gordon equation, we find

$$\begin{aligned}
& (1-v^2)f'' = \sin f \\
\Longleftrightarrow \qquad & f'' = \gamma^2 \sin f \ , \qquad \text{where } \gamma := \frac{1}{\sqrt{1-v^2}} \\
\underset{\int d\xi\ f'}{\Longrightarrow} \qquad & \frac{1}{2}(f')^2 = A - \gamma^2 \cos f \\
\text{BC's:} \qquad & A = \gamma^2 \\
\Longrightarrow \qquad & f' = \pm\sqrt{2}\gamma\sqrt{1-\cos f} = \pm 2\gamma \sin\frac{f}{2} \\
\Longrightarrow \qquad & \int \frac{df}{2\sin\frac{f}{2}} = \pm\gamma(x - x_0 - vt) \\
\Longrightarrow \qquad & \log\tan\frac{f}{4} = \pm\gamma(x - x_0 - vt)
\end{aligned}$$

where x_0 is an undetermined integration constant.

We find therefore the following travelling wave solution of the sine-Gordon equation

$$\boxed{u(x,t) = f(x-vt) = 4\arctan\left(e^{\pm\gamma(x-x_0-vt)}\right)} \ , \tag{3.2}$$

which goes by the name of **"KINK"** (+ sign) or **"ANTI-KINK"** (− sign).

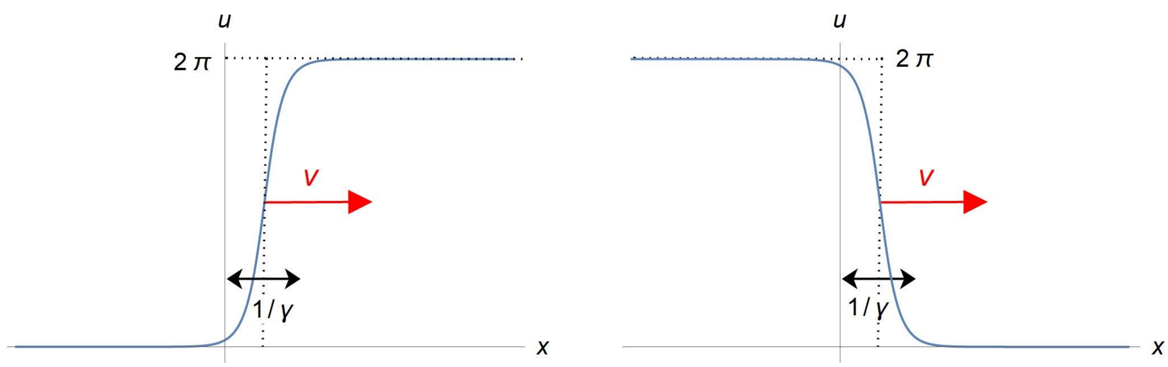
Note that the BC required that as $\xi \to \pm\infty$

$$f(\xi) \to 2\pi n_\pm \ , \qquad f'(\xi) \to 0 \qquad (\Rightarrow f''(\xi) \to 0) \ ,$$

where the two integers $n_\pm \in \mathbb{Z}$ can be different. Indeed they are different for a kink (/antikink) solution. Choosing the branch of the $\arctan$ such that

$$\arctan(0^\pm) = 0^\pm \;, \qquad \arctan(\pm\infty) = \pm\left(\frac{\pi}{2}\right)^{\mp} \;,$$

we find that the kink and the anti-kink solution look as in fig. 3.2 at a fixed time t:



a) Kink b) Anti-kink

Figure 3.2: Snapshots of the sine-Gordon kink and anti-kink.

REMARKS:

1. Choosing a different branch of the $\arctan$[3] shifts the whole solution $u(x,t)$ by a multiple of 2π. This is inconsequential. What matters is:

$$\begin{aligned} u(+\infty,t) - u(-\infty,t) &= +2\pi \qquad \text{KINK} \\ u(+\infty,t) - u(-\infty,t) &= -2\pi \qquad \text{ANTI-KINK} \end{aligned}$$

2. The velocity of the kink/anti-kink could be

$$\begin{aligned} v > 0 &: \quad \text{RIGHT-MOVING} \\ v = 0 &: \quad \text{STATIC} \\ v < 0 &: \quad \text{LEFT-MOVING} \end{aligned}$$

3. For a **real** solution we need

$$\gamma^2 \geqslant 0 \quad \Longrightarrow \quad |v| \leqslant 1 = \text{speed of light}$$

4. The kink/antikink is a **localised** lump centred at $x_0 + vt$ and with

$$\text{WIDTH} \;\sim\; \frac{1}{\gamma} = \sqrt{1-v^2}\;.$$

[3]along with reversing the sign and adjusting the integration constant if the multiple is odd. Check for yourself.

So faster kinks/antikinks are narrower. This phenomenon is known as "Lorentz contraction" and is a feature of special relativity. γ is called the **"Lorentz factor"**.

NOTE: It might be confusing to state that the kink/antikink is localised, when u interpolates between different values as $x \to \pm\infty$. The key point is that u is an angular variable which is only defined modulo addition of 2π. To define the width it is better to look at single-valued objects like e^{iu} or $\partial_x u$, which do not suffer from the above ambiguity. This point will become more concrete later when we calculate the energy density of the kink, which is a single-valued and everywhere positive function, which achieves a maximum at the centre of the kink and approaches zero far away from the centre, see figure 4.2.

3.3 A mechanical model for the sine-Gordon equation

Consider a chain of infinitely many identical pendulums hanging from a straight wire which cannot be stretched but can be twisted. Each identical pendulum consists of a massless[4] rod of length L, with a weight of mass M at the end of the rod. The pivot of the n-th pendulum at position na along the line, where $n \in \mathbb{Z}$ and a is the separation, and the configuration of the n-th pendulum at time t is encoded by $\theta_n(t)$, the angle between the pendulum and the downward pointing vertical at time t. See figure 3.3.

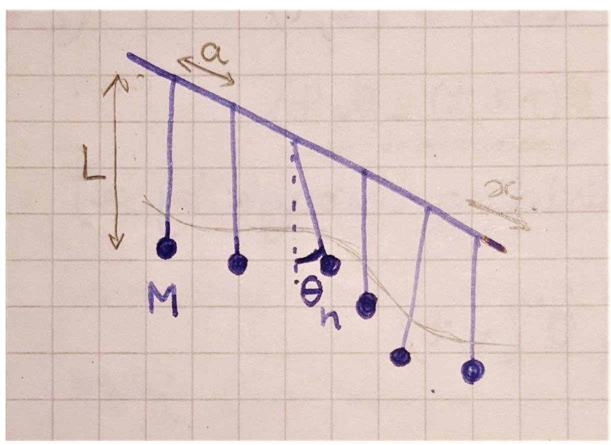


Figure 3.3: Section of an infinite chain of pendulums separated by distance a.

The pendulums are subject to two kinds of forces: a gravitational force due to the attraction between the Earth and the weights, which favours downward pointing pendulums; and a twisting force between neighbouring pendulums due to the wire, which favours a straight untwisted wire and therefore the alignment of neighbouring pendulums.[5] The **equations of**

[4]This assumption can be easily relaxed, leading to no qualitative difference in what follows.

[5]This is a slight lie. If you have studied rigid bodies you will recognise that these are "torques" rather than forces. The equation of motion (3.3) is not the standard Newton's law $F = ma$, but rather its rotational analogue, which states that the total torque equals the product of the moment of inertia and the angular acceleration.

motion (the analogue of Newton's equation $F = ma$) for this physical system are a coupled system of infinitely many ODE's labelled by the integer n, one for each pendulum, which take the form

$$ML^2\ddot{\theta}_n(t) = \underbrace{-MgL \cdot \sin\theta_n(t)}_{\text{net gravitational force}} + \underbrace{\frac{k}{a}\Big(\theta_{n+1}(t) - \theta_n(t)\Big) + \frac{k}{a}\Big(\theta_{n-1}(t) - \theta_n(t)\Big)}_{\text{twisting forces exerted by neighbouring pendulums}}, \quad n \in \mathbb{Z} \tag{3.3}$$

where a dot denotes a time derivative, g is the gravitational acceleration and k is an elastic constant that parametrizes the strength of the twisting force.

Now we are going to take the so called **"continuum limit"** of this infinite-dimensional discrete system, in which the separation between consecutive pendulums becomes infinitesimally small and the average mass density (*i.e.* the mass per unit length) along the line is kept fixed:

$$a \to 0\ , \qquad m = M/a \ \text{ fixed}\ .$$

In the continuum limit, the position $x = na$ of the n-th pendulum along the line effectively becomes a continuous real variable, which replaces the discrete index $n \in \mathbb{Z}$. Identifying $\theta_n(t) \equiv \theta(x = na, t)$, the collection $\{\theta_n(t)\}_{n\in\mathbb{Z}}$ of angular coordinates of the infinitely many pendulums at time t is replaced in the limit by a single function $\theta(x,t)$ of two continuous variables, space and time. By the definition of the derivative as a limit, we also have that

$$\frac{\theta_{n+1}(t) - \theta_n(t)}{a} \to \theta'(x,t)\ ,$$
$$\frac{1}{a}\left(\frac{\theta_{n+1}(t) - \theta_n(t)}{a} - \frac{\theta_n(t) - \theta_{n-1}(t)}{a}\right) \to \theta''(x,t)\ .$$

where a prime denotes an x-derivative.

Dividing the equations of motion (3.3) by $ML^2 = amL^2$ and taking the continuum limit we find the single equation of motion

$$\ddot{\theta} = -\frac{g}{L}\sin\theta + \frac{k}{mL^2}\theta''$$

for the **"field"** $\theta(x,t)$. We can get rid of the constants by rescaling x and t[6], and rearrange to get the equation

$$\ddot{\theta} - \theta'' = -\sin\theta\ ,$$

which is nothing but the sine-Gordon equation $\theta_{tt} - \theta_{xx} = -\sin\theta$ for the field θ! We say therefore that the sine-Gordon equation is the continuum limit of (3.3).

We can use this mechanical model to gain some intuition about the possible configurations of the sine-Gordon field:

[6] Send $x \mapsto \sqrt{\frac{k}{mgL}}\, x$ and $t \mapsto \sqrt{\frac{L}{g}}\, t$.

- The **lowest energy state** (or **"ground state"**, or **"vacuum"**) of the system is the configuration with all pendulums pointing downwards,

$$\theta(x,t) = 0 \pmod{2\pi} \qquad \forall\, x \,,$$

which is a configuration of stable equilibrium.[7] See figure 3.4.

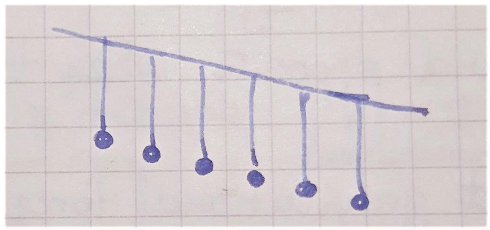


Figure 3.4: Configuration of stable equilibrium for the chain of pendulums.

- By a continuous perturbation of the vacuum, we can obtain configuration which represents a "small wave", which satisfies the same boundary conditions of the vacuum, $\theta \to 0$ as $x \to \pm\infty$:[8]

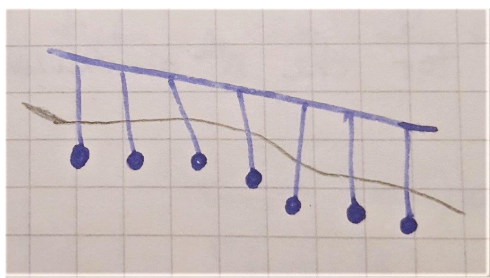


Figure 3.5: A small wave going through the chain of pendulums.

- There are also configurations in which the chain of pendulums twists around the line. If they twist once in the direction of increasing angles, so that θ increases by 2π from $x \to -\infty$ to $x \to +\infty$, this describes a kink or a continuous deformation thereof:

  If instead they twist once in the direction of decreasing angles, so that θ decreases by 2π from $x \to -\infty$ to $x \to +\infty$, this describes an anti-kink or a continuous deformation thereof.

- The **limiting values** of the sine-Gordon field θ as $x \to \pm\infty$ are **fixed**: changing them would require twisting infinitely many pendulums by 360 degrees, which would cost energy.

[7] We will confirm this intuition later when we study the energy of the sine-Gordon field.

[8] We will see later that this "small wave" does not need to be small, in fact. For instance it could look like a kink followed by an antikink.

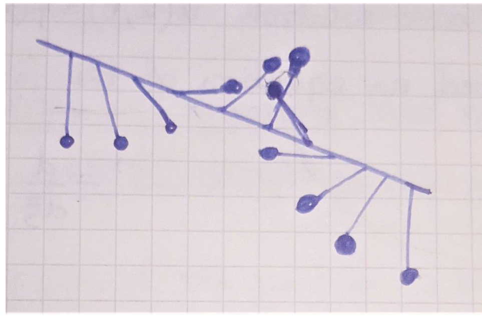


Figure 3.6: A kink going through the chain of pendulums.

If

$$\theta(+\infty, t) - \theta(-\infty, t) = 2m\pi \ , \qquad \text{with } \ m \neq 0 \ \text{ integer} \ ,$$

then the configuration of the system **cannot be deformed continuously to the vacuum** where all pendulums point downwards, unlike the "small wave" mentioned above. This tells us that the kink (or the antikink) **cannot disperse/dissipate into the vacuum**. This is related to the notion of **topological stability**, which we will discuss in the next chapter.

I invite you to play with this Wolfram demonstration of the chain of coupled pendulums, using Mathematica (which should be available on university computers – let me know if it isn't) or the free Wolfram Player. Play with the parameters and visualise a kink, the scattering of two kinks or of a kink and an anti-kink, and the breather, a bound state of a kink and an anti-kink. We will study all of these configurations in the continuum limit later in the term, using the sine-Gordon equation.

3.4 Travelling wave solutions and 1d point particles (bonus material)

Looking for a travelling wave solutions $u(x,t) = f(x-vt) \equiv f(\xi)$ of the KdV and sine-Gordon equation, we encountered equations of the form

$$\boxed{f'' = \hat{F}(f)}$$

where a prime denotes a derivative with respect to ξ. We integrated this equation to

$$\boxed{\frac{1}{2}(f')^2 + \hat{V}(f) = \hat{E} = \text{const}} \qquad (*)$$

where

$$\hat{V}(f) = -\int df \ \hat{F}(f) \ .$$

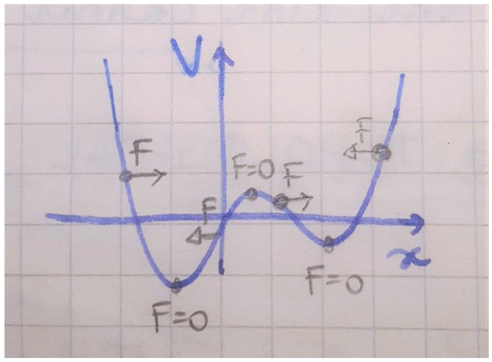


Figure 3.7: Example of a potential energy $V(x)$ and force $F(x) = -V'(x)$.

By tuning the integration constant in this indefinite integral and absorbing it in $\hat{E}$, we can set $\hat{E}$ to zero or to any value we wish.

The previous equations are **analogous** to the **classical mechanics** of a **point particle** moving in **one space dimension**. Let $x(t)$ be the position of the point particle at time t and dots denote time derivatives. The equation of motion (EoM) of the point particle is **Newton's equation**

$$\boxed{m\ddot{x} = F(x)}$$

(mass $\times$ acceleration $=$ force) can be integrated to the **energy conservation** law

$$\boxed{\frac{1}{2}m\dot{x}^2 + V(x) = E = \text{const}}$$

(kinetic energy $+$ potential energy $=$ total energy, which is constant in time), where the force and the potential energy are related by

$$\boxed{F(x) = -\frac{d}{dx}V(x)}\,.$$

The potential energy and the total energy can be shifted by a common constant with no physical change. See figure 3.7 for an example of a potential energy $V(x)$ and the associated force $F(x) = -V'(x)$.

It may be useful to think of x as the horizontal coordinate of a point particle (think of an infinitesimal ball) moving on a hill of vertical height $V(x)$ at coordinate x, subject only to the gravitational force and the reaction of the ground (which is equal and opposite when the ground is flat). Even if you are not very familiar with classical mechanics, you will hopefully have some intuition of what will happen to the ball.[9]

[9] You can also model this by riding a brakeless bike in hilly Durham. It's a good idea to develop some intuition about this physical system without running the experiment yourself, which I don't recommend. (This is one of a number of reasons why theoretical physics is superior to experimental physics.)

The mathematical correspondence between the equations for a travelling wave in one space and one time dimension and for a classical point particle in one space dimension is

$$\boxed{\begin{array}{rcl} \xi & \longleftrightarrow & t \\ f & \longleftrightarrow & x \\ 1 & \longleftarrow & m \\ \hat{F}(f) & \longleftrightarrow & F(x) \\ \hat{E}-\hat{V}(f) & \longleftrightarrow & E-V(x) \end{array}}$$

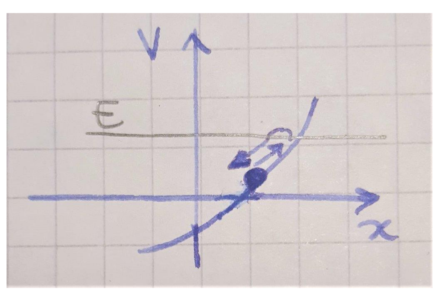
This correspondence allows us to understand the qualitative behaviour of travelling waves even when we cannot integrate equation $(*)$ exactly, using elementary facts from classical mechanics, which are encoded in the the mathematics of the previous equations:

1. The total **energy is conserved** and can only be converted from kinetic energy (which is non-negative!) to potential energy and vice versa. The velocity $\dot{x}$ of the point particle is zero if and only if the kinetic energy is zero, which means that all the energy is stored in potential energy:
$$\dot{x}=0 \qquad \Longleftrightarrow \qquad V(x)=E\ .$$

2. When the point particle reaches one of the special values of x such that $V(x) = E$, either of two things happens depending on the acceleration of the particle:

   (a) $F(x) = -\frac{d}{dx}V(x) \neq 0$:

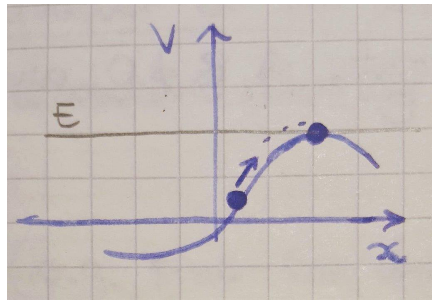
   The acceleration is non-vanishing, therefore the particle **reverses its direction of motion**:



   These values of x are known as **"turning points"**.

   (b) $F(x) = -\frac{d}{dx}V(x) = 0$:

   The acceleration vanishes and the particle **stops**.



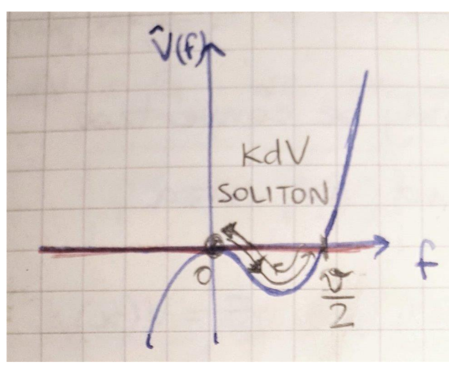
These values of x are known as **"equilibrium points"**. The approach to equilibrium takes an infinite time.

*** EXERCISE**: Derive the previous statements by Taylor expanding the potential energy about a point where $V(x) = E$ and substituting the expansion in the energy conservation law.

Now let us translate this discussion to the context of travelling waves. We will focus on the examples of the KdV and the sine-Gordon equation here, but more examples are available in **[Ex 13]** in the problems set.

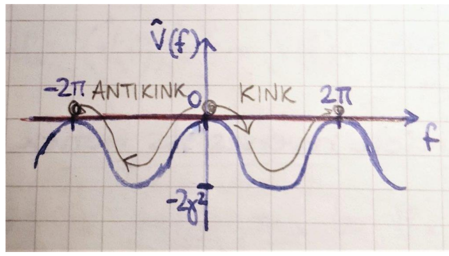
EXAMPLES:

1. **KdV**: $\hat{E} = 0\,, \quad \hat{V}(f) = f^2\left(f - \frac{v}{2}\right) \qquad (v > 0)$



From a graphical analysis of $\hat{V}(f)$ and the analogy between travelling waves and point particles in one dimension, we see that there exists a travelling wave solution that starts at $f = 0^+$ at $\xi \to -\infty$, increases until the 'turning point' $f = v/2$, and decreases to $f = 0^+$ at $\xi \to +\infty$. This is nothing but the KdV soliton (3.1) that we found in section 3.1. If instead the travelling wave solution starts at $f = 0^-$ at $\xi \to -\infty$, then it will fall down the cliff and reach $f \to -\infty$, leading to a singular solution, that we discard. Note that if $v < 0$ we have that $\hat{V}(0) = 0$, but $\hat{V}(f) > 0$ for small $f \neq 0$. Therefore the only real solution obeying the boundary conditions is the constant zero solution $f(\xi) = 0$ for all ξ. If $v = 0$, in addition to the trivial solution there is also a singular real travelling wave solution that we discard on physical grounds.

2. **sine-Gordon**: $\hat{E} = 0\,,\quad \hat{V}(f) = \gamma^2\,(\cos f - 1)$



From a graphical analysis of $\hat{V}(f)$, we see that two classes of travelling wave solutions exist: one where f interpolates between $2n\pi$ at $x \to -\infty$ and $2(n+1)\pi$ $x \to -\infty$, and another where f interpolates between $2n\pi$ at $x \to -\infty$ and $2(n-1)\pi$ $x \to -\infty$. We identify these solutions with the kink and anti-kink (3.2) of section 3.2.

*** EXERCISE**: Using the analogy with a one-dimensional point particle, determine the qualitative behaviour of a travelling wave solution of the KdV equation on a circle (*i.e.* with periodic boundary conditions). [**Hint**: allow integration constants $A, B \neq 0$ and look at $\hat{V}(f)$.] **[Ex 14*]**

Chapter 4

Topological lumps and the Bogomol'nyi bound

The main references for this chapter are §5.3, 5.1 of [Manton and Sutcliffe, 2004] and §2.1 of [Dauxois and Peyrard, 2006].

4.1 The sine-Gordon kink as a topological lump

In chapter 3 the **topological properties** of the **sine-Gordon kink** were mentioned briefly – they ensure that it cannot disperse or dissipate to the vacuum. Let us understand these topological properties better. As a reminder, the sine-Gordon equation for the field u is

$$u_{tt} - u_{xx} + \sin u = 0 \ .$$

Starting from the **discrete** mechanical model involving pendulums of section 3.3, rescaling x and t as in footnote 6 so as to eliminate all constants, and taking the continuum limit $a \to 0$, it is not hard to see that the **kinetic energy** T and the **potential energy** V of the sine-Gordon field are **[Ex 15]**

$$T = \int_{-\infty}^{+\infty} dx \ \frac{1}{2} u_t^2 \tag{4.1}$$

$$V = \int_{-\infty}^{+\infty} dx \left[\underbrace{\frac{1}{2} u_x^2}_{\text{twisting}} + \underbrace{(1 - \cos u)}_{\text{gravity}} \right] . \tag{4.2}$$

<u>**REMARK**</u>:
The kinetic and potential energies of the sine-Gordon field are the continuum limits of the kinetic and potential energies of the infinite chain of pendulums. They should not be confused with $\frac{1}{2}(f')^2$ and $\hat{V}(f)$ for the one-dimensional point particle in the analogy of section 3.4.

We can use this result to deduce the boundary conditions that we anticipated in section 3.2. The boundary conditions follow from requiring that **all field configurations** have **finite (total) energy** $E = T + V$. Since the total energy is the integral over the real line of the sum of three non-negative terms, the limits of all three terms as $x \to \pm\infty$ must be zero to ensure the convergence of the integral. So the finiteness of the energy requires the boundary conditions

$$u_t \ , \quad u_x \ , \quad 1 - \cos u \quad \xrightarrow[x\to\pm\infty]{} \quad 0 \qquad \forall \, t \ .$$

Since $1 - \cos u = 0$ iff u is an integer multiple of 2π, we need

$$\boxed{u(-\infty, t) = 2\pi n_- \ , \qquad u(+\infty, t) = 2\pi n_+ \ ,} \tag{4.3}$$

for some integers $n_\pm$. (This means that pendulums are at rest, pointing downwards, as $x \to \pm\infty$.)

<u>REMARKS</u>:[1]

1. The overall value of $n_\pm$ has no meaning, since u is defined modulo 2π. A shift of the field $u \mapsto u + 2\pi k$ is unphysical, but it shifts $n_\pm \mapsto n_\pm + k$. What really matters is the difference $n_+ - n_-$, which is invariant under this ambiguity:

$$\frac{1}{2\pi}[u(+\infty,t) - u(-\infty,t)] = n_+ - n_- = \#\text{ of “twists”/“kinks”}$$

2. The integer $n_+ - n_-$ is **"TOPOLOGICAL"**, *i.e.* it does not change under any **continuous** changes of the field u (and of the energy E). In particular, it cannot change under time evolution, since time is continuous. Therefore it is a **constant of motion** or a **"conserved charge"** (more about this in the next chapter). Since the conservation of $n_+ - n_-$ is due to a topological property, we call this a **"TOPOLOGICAL CHARGE"**.[2] Solutions with the same topological charge are said to belong to the same **"TOPOLOGICAL SECTOR"**.

3. Dispersion and dissipation occur by time evolution, a continuous process which cannot change the value of the integer $n_+ - n_-$. Since the vacuum has $n_+ - n_- = 0$, any configuration with $n_+ - n_- \neq 0$ **cannot disperse/dissipate to the vacuum**.

[1] Some of these remarks were made for kinks and antikinks in the previous chapter. Now that we derive them from the BC's, we see that they hold more generally for all solutions.

[2] [Advanced remark for those who know some topology – if you don't, you can safely ignore this:] Mathematically, $n_+ - n_-$ is a "winding number", the topological invariant which characterises maps $S^1 \to S^1$. The first S^1 is the compactification of the spatial real line, with the points at infinity identified, and the second S^1 is the circle parametrised by u mod 2π. The winding number counts how many times u winds around the circle as x goes from $-\infty$ to $+\infty$.

<u>**VOCABULARY**</u>:

- **"TOPOLOGICAL CONSERVATION LAW":**
  The conservation (in time) of a topological charge, that is $\frac{d}{dt}(\text{topological charge}) = 0$.

- **"TOPOLOGICAL LUMP":**
  A localised field configuration which cannot dissipate or disperse to the vacuum by virtue of a topological conservation law.

So the **sine-Gordon kink** is a **topological lump**. It is also a **soliton**, but to see that we will have to check property 3, which concerns scattering.

Topological lumps also exist in higher dimensions. A notable example is the "magnetic monopole", a magnetically charged localised object that exists in certain generalizations of electromagnetism in three space and one time dimensions. Another example is the "vortex", which is a topological lump if space is $\mathbb{R}^2$.[3]

4.2 The Bogomol'nyi bound

Among the kink solutions found in (3.2) using the travelling wave *ansatz*, there was a **STATIC KINK** with zero velocity. Topology tells us that it cannot disperse or dissipate completely to the vacuum. But **is its precise shape "stable"** under small perturbations? This would be **guaranteed if** we could show that **it minimises the energy** amongst all configurations with the same topological charge. The reason is that any perturbation near a minimum of the energy would increase the energy, which however is conserved upon time evolution.[4]

A useful analogy to keep in mind is with a point particle on a hilly landscape under the force of gravity, as in figure 4.1: if the point particle is sitting still at a local miminum of the height, minimising the energy (locally), it is in a configuration of stable equilibrium. Any perturbation would necessarily move the particle up the hill, but this is not allowed under time evolution as it would increase the total energy.

So we will seek a lower bound for the total energy $E = T + V$ in the topological sector of the kink, which has topological charge $n_+ - n_- = 1$. The energy is the integral of a non-negative

[3]Indeed there is a topological charge, the 'vortex number', which is conserved and can be non-vanishing if space is $\mathbb{R}^2$. On the other hand, topology implies that the vortex number vanishes on the two-sphere S^2: this is fortunate, because if it were non-vanishing there would always be hurricanes going around the surface of Earth.

[4]We will in fact show that the static kink is a global minimum of the energy amongst configurations with unit topological charge. This ensures its stability even when one includes quantum effects, which we are not concerned with in this course.

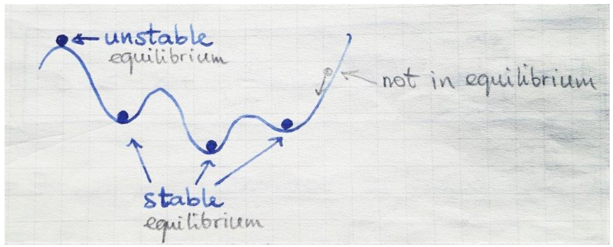


Figure 4.1: A point particle on a hilly landscape is stable if it locally minimises the energy. This happens when it is sitting still at a minimum of the potential energy.

energy density, so immediately find the lower bound $E \geqslant 0$, but we can do better than that:

$$
\begin{aligned}
E = T + V \quad &= \quad \int_{-\infty}^{+\infty} dx \left[\frac{1}{2} u_t^2 + \frac{1}{2} u_x^2 + (1 - \cos u) \right] \\
&\underset{(u_t^2 \geqslant 0)}{\geqslant} \quad \int_{-\infty}^{+\infty} dx \left[\frac{1}{2} u_x^2 + (1 - \cos u) \right] \\
&= \quad \int_{-\infty}^{+\infty} dx \left[\frac{1}{2} u_x^2 + 2 \sin^2 \frac{u}{2} \right] \\
&\underset{\text{“Bogomol'nyi trick”}}{=} \quad \int_{-\infty}^{+\infty} dx \left[\frac{1}{2} \left(u_x \pm 2 \sin \frac{u}{2} \right)^2 \mp 2 \sin \frac{u}{2} \cdot u_x \right] \\
&= \quad \int_{-\infty}^{+\infty} dx \, \frac{1}{2} \left(u_x \pm 2 \sin \frac{u}{2} \right)^2 \pm 4 \left[\cos \frac{u}{2} \right]_{-\infty}^{+\infty} . \qquad (*)
\end{aligned}
$$

A few comments are in order:

1. The inequality in the second line follows from omitting the non-negative term $\frac{1}{2} u_t^2$. It is **“saturated”** (that is, it becomes an equality) for **static field configurations**, such that $u_t = 0$;

2. In the third line we used a half-angle formula;

3. In the fourth line we used the so called **“Bogomol'nyi trick”** to replace a sum of squares by the square of a sum plus a correction term which is a total x-derivative;

4. In the fifth line we integrated the total derivative, leading to a **“boundary term”** (or **“surface term”**) which only depends on the limiting values of the field at spatial infinity.

If u satisfies the 1-**kink BC's**

$$
\boxed{u(-\infty, t) = 0 \, , \qquad u(+\infty, t) = 2\pi} \, ,
$$

then the boundary term evaluates to

$$4\left[\cos\frac{u}{2}\right]_{-\infty}^{+\infty} = 4(-1-1) = -8\,.$$

Picking the lower (*i.e.* $-$) signs in ($*$), we obtain the lower bound

$$\boxed{E \geqslant \int_{-\infty}^{+\infty} dx\,\frac{1}{2}\left(u_x - 2\sin\frac{u}{2}\right)^2 + 8 \geqslant 8} \tag{4.4}$$

for the energy, where the second inequality is saturated if the expression in brackets vanishes.[5] Equation (4.4) is known as the **"BOGOMOL'NYI BOUND"**.

The **Bogomol'nyi bound** (4.4) is **saturated** (*i.e.* $E = 8$) if and only if the field configuration is **static**, that is

$$\boxed{u_t = 0}\,,$$

and satisfies the **"BOGOMOL'NYI EQUATION"**

$$\boxed{u_x = 2\sin\frac{u}{2}}\,. \tag{4.5}$$

So we can find the **least energy field configurations** in the "1**-kink topological sector"** (*i.e.* with $n_+ - n_- = 1$) by looking for static solutions $u = u(x)$ of the Bogomol'nyi equation:

$$u_x = 2\sin\frac{u}{2} \qquad \Longrightarrow \qquad \int dx = \int \frac{du}{2\sin\frac{u}{2}} = \log\tan\frac{u}{4}\,,$$

whose general solution is

$$\boxed{u(x) = 4\arctan\left(e^{x-x_0}\right)}\,. \tag{4.6}$$

This is nothing but the **static kink**, which we obtained in section 3.2 as a special case of a travelling wave solution of the sine-Gordon equation with $v = 0$.

REMARK:
Note that the Bogomol'nyi equation, being a first order differential equation (in fact an ODE once we impose $u_t = 0$), is much easier to solve than the full equation of motion, the sine-Gordon equation, which is a second order PDE.

$*$ **EXERCISE**: Check that a field configuration that saturates the Bogomol'nyi bound is automatically a solution of the sine-Gordon equation.

[5] Picking the upper (*i.e.* $+$) signs in ($*$) we obtain the lower bound $E \geqslant -8$, which is weaker than the trivial bound $E \geqslant 0$ therefore not very useful. The Bogomol'nyi trick always has a sign ambiguity. The choice of sign that leads to the stricter inequality depends on the sign of the boundary term.

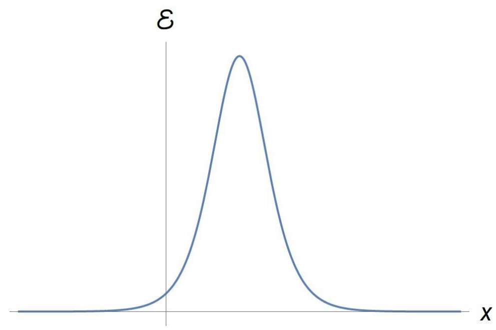


Figure 4.2: The energy density of a static kink.

So we learned that **amongst all solutions with topological charge** $n_+ - n_- = 1$, the **static kink has the least energy**, hence it is **stable**. Indeed, topology in principle allows the kink to disperse to other solutions with $n_+ - n_- = 1$, but the dispersing waves would carry some of the energy away. Since the static kink has the least energy in the $n_+ - n_- = 1$ topological sector, it can't lose energy, hence it's stable. This notion of stability which originates from minimising the energy in a given topological sector is called **"TOPOLOGICAL STABILITY"**.

Using staticity and the Bogomol'nyi equation, we now have a shortcut to compute the **energy density** $\mathcal{E}$ of the **static kink**, namely the integrand of the total energy $E = \int_{-\infty}^{+\infty} dx\, \mathcal{E}$:

$$\mathcal{E} \;=\; \frac{1}{2}u_t^2 + \frac{1}{2}u_x^2 + 2\sin^2\frac{u}{2} \underset{\substack{u_t=0\\ u_x = 2\sin\frac{u}{2}}}{=} u_x^2 \;=\; 4\mathrm{sech}^2(x - x_0)\ ,$$

which shows that the energy density of the kink is localised near x_0, see figure 4.2.

*** EXERCISE**: Think about how to generalise the Bogomol'nyi bound for higher topological charge, for instance $n_+ - n_- = 2$. This is not obvious! **[Ex 17]**

4.3 Summary

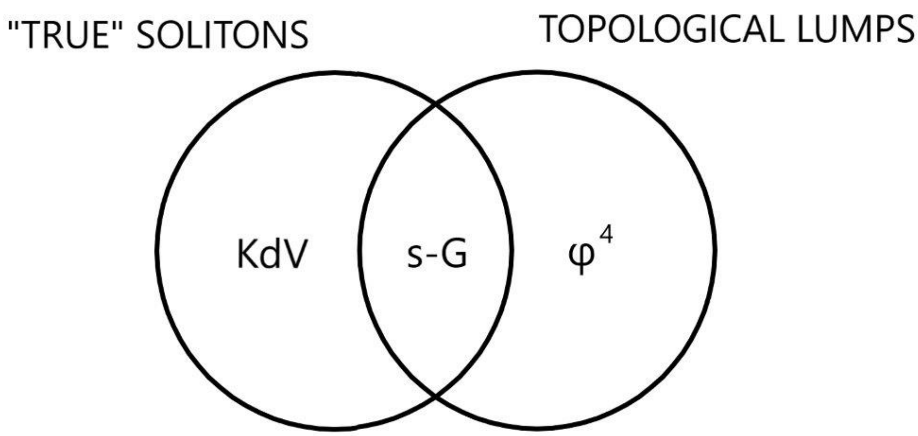
There are two ways for a lump to be long-lived:

1. by **INTEGRABILITY** (infinitely many conservation laws, more about this next)
   $\longrightarrow$ **"TRUE" (or "INTEGRABLE") SOLITONS**

2. by **TOPOLOGY** (topological conservation law)
   $\longrightarrow$ **TOPOLOGICAL LUMPS.**[6]

[6] Some people use the term solitons for both integrable solitons and topological lumps, but in this course we will only refer to the former as "solitons").

It is important to note that these two mechanisms are not mutually exclusive: there are some lumps, like the sine-Gordon kink, which are both topological lumps and true solitons. The various possibilities and some examples are summarised in the following Venn diagram:



Chapter 5

Conservation laws

The main references for this chapter are §5.1.1 and §5.1.2 of [Drazin and Johnson, 1989].

Conservation laws provide the most fundamental characterisation of a physical system: they tell us which quantities don't change with time. For the purpose of this course, they play a key role because they explain why the motion of "true" solitons is so restricted that they scatter without changing their shapes.

The idea of a conservation law is to construct spatial integrals of functions of the field u and its derivatives

$$\boxed{Q = \int_{-\infty}^{+\infty} dx\ \rho(u, u_x, u_{xx}, \ldots, u_t, u_{tt}, \ldots)} \tag{5.1}$$

which are constant in time (in physics parlance, they are "**constants of motion"**)

$$\boxed{\frac{d}{dt}Q = 0} \tag{5.2}$$

when u satisfies its **equation of motion (EoM)**, such as the sine-Gordon equation or the KdV equation. The constant of motion (5.1) is called a **"CONSERVED CHARGE" or "CONSERVED QUANTITY"** and the equation (5.2) stating its time-independence is called a **"CONSERVATION LAW"**.

For the KdV and the sine-Gordon equation, it turns out that there exist **infinitely many conserved quantities**. This makes them **"integrable systems"** (more about this next term) and explains many of their special properties.

5.1 The basic idea

The standard method for constructing a conserved charge like (5.1) involves finding two functions ρ and j of u and its derivatives, such that the EoM for u implies the **"LOCAL CONSERVATION LAW"** or **"CONTINUITY EQUATION"**

$$\boxed{\frac{\partial \rho}{\partial t} + \frac{\partial j}{\partial x} = 0} \tag{5.3}$$

and the boundary conditions imply

$$j \to C \quad \text{as } x \to \pm\infty \tag{5.4}$$

with the *same* constant C at $-\infty$ and $+\infty$. Then

$$\frac{d}{dt}\int_{-\infty}^{+\infty} dx\ \rho = \int_{-\infty}^{+\infty} dx\ \frac{\partial \rho}{\partial t} \underset{(5.3)}{=} -\int_{-\infty}^{+\infty} dx\ \frac{\partial j}{\partial x} = -[j]_{-\infty}^{+\infty} \underset{(5.4)}{=} 0\ .$$

Hence

$$Q = \int_{-\infty}^{+\infty} dx\ \rho \tag{5.5}$$

is a **conserved CHARGE**. ρ is called the conserved **"CHARGE DENSITY"**, and j is called the conserved **"CURRENT DENSITY"** (or just **"CURRENT"**, by a common abuse of terminology.)

5.2 Example: conservation of energy for sine-Gordon

Is the total energy

$$\boxed{E = \int_{-\infty}^{+\infty} dx\ \mathcal{E}}$$

conserved for the sine-Gordon field, where the energy density is

$$\boxed{\mathcal{E} = \frac{1}{2}u_t^2 + \frac{1}{2}u_x^2 + (1 - \cos u)}\ ? \tag{5.6}$$

The energy density $\mathcal{E}$ plays the role of ρ here. Can we show then that $\rho = \mathcal{E}$ obeys a continuity equation (5.3) for some function j that obeys the limit condition (5.4), when the sine-Gordon equation (EoM)

$$u_{tt} - u_{xx} + \sin u = 0$$

holds? Let's compute:

$$\begin{aligned}\frac{\partial \mathcal{E}}{\partial t} &= u_t u_{tt} + u_x u_{xt} + \sin u \cdot u_t \\ &= u_t (u_{tt} + \sin u) + u_x u_{xt} \\ &\underset{\text{EoM}}{=} u_t u_{xx} + u_x u_{xt} = \frac{\partial}{\partial x}(u_t u_x) \equiv \frac{\partial}{\partial x}(-j) \ ,\end{aligned}$$

and since the BC's for the sine-Gordon equation imply that $u_t u_x \to 0$as $x \to \pm\infty$, we deduce that **energy is conserved**:

$$\boxed{\frac{dE}{dt} = 0} \ .$$

5.3 Conserved quantities for the KdV equation

Let us return to the KdV equation

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

We can rewrite the KdV equation as a continuity equation

$$\frac{\partial}{\partial t} u + \frac{\partial}{\partial x}(3u^2 + u_{xx}) = 0$$

and since the BC's appropriate for KdV on the line $\mathbb{R}$ are that $u, u_x, u_{xx}, \cdots \to 0$ as $x \to \pm\infty$, we deduce that

$$\boxed{Q_1 = \int_{-\infty}^{+\infty} dx \ u} \ , \tag{5.7}$$

is **conserved**. For the canal, this is the conservation of water[1].

Next, we can ask whether $\rho = u^2$ is a conserved charge density. Let us compute

$$\begin{aligned}(u^2)_t = 2uu_t &\underset{\text{KdV}}{=} -12u^2 u_x - 2uu_{xxx} = -4(u^3)_x - 2uu_{xxx} \\ &= (-4u^3 - 2uu_{xx})_x + 2u_x u_{xx} = (-4u^3 - 2uu_x + u_x^2)_x \ ,\end{aligned}$$

where to go from the first to the second line we used the trick familiar from integration by parts, $fg_x = (fg)_x - f_x g$. (We say that fg_x and $-f_x g$ are equal up to a total x-derivative.) Hence we deduce that

$$\boxed{Q_2 = \int_{-\infty}^{+\infty} dx \ u^2} \ , \tag{5.8}$$

[1](5.7) is the (net) area under the profile of the wave, taking $u = 0$ (flat water surface) as zero. Assuming that water has constant density (mass per unit area) and choosing units so that the density is 1, (5.7) is also the mass of the wave.

which is interpreted as the **momentum** of the wave, is **conserved**.

Next, what about $\rho = u^3$? Using the notation “ $=$ ” to mean “equal up to a total x-derivative” and striking out terms which are total derivatives (t.d.), we find

$$
\begin{aligned}
(u^3)_t = 3u^2 u_t \underset{\text{KdV}}{=} -18\cancelto{\text{t.d.}}{u^3 u_x} - 3u^2 u_{xxx} \text{ “}=\text{” } 6u u_x u_{xx} \\
\underset{\text{KdV}}{=} -u_t u_{xx} - \cancelto{\text{t.d.}}{u_{xxx} u_{xx}} \text{ “}=\text{” } u_{tx} u_x = \frac{1}{2}(u_x^2)_t \,,
\end{aligned}
$$

so rearranging we find a third **conserved** charge

$$
\boxed{Q_3 = \int_{-\infty}^{+\infty} dx \, \left(u^3 - \frac{1}{2} u_x^2 \right)} \,, \tag{5.9}
$$

which is interpreted as the **energy** of the wave.

It turns out that the conservation laws (5.7)-(5.9) of **mass, momentum** and **energy** follow by Noether’s theorem from the “obvious” symmetries

$$
\begin{aligned}
u &\mapsto u + c & &\Longrightarrow & &\text{mass conservation} \\
x &\mapsto x + c' & &\Longrightarrow & &\text{momentum conservation} \\
t &\mapsto t + c'' & &\Longrightarrow & &\text{energy conservation}
\end{aligned}
$$

of the KdV equation, so they are expected. But then surprisingly [Miura et al., 1968] found (by hand!) **eight more conserved charges**, all (but one, see **[Ex 23]**) of the form

$$
\boxed{Q_n = \int_{-\infty}^{+\infty} dx \, (u^n + \dots)} \,,
$$

e.g.

$$
\begin{aligned}
Q_4 &= \int_{-\infty}^{+\infty} dx \, \left(u^4 - 2u u_x^2 + \frac{1}{5} u_{xx}^2 \right) \\
Q_5 &= \int_{-\infty}^{+\infty} dx \, \left(u^5 - 5u^2 u_x^2 + u u_{xx}^2 - \frac{1}{14} u_{xxx}^2 \right) \\
&\vdots \\
Q_{10} &= \int_{-\infty}^{+\infty} dx \, \left(u^{10} - 60 u^7 u_x^2 + (29 \text{ terms}) + \frac{1}{4862} u_{xxxxxxxxx}^2 \right) \,.
\end{aligned} \tag{5.10}
$$

*** EXERCISE**: Calculate $Q_6, \dots, Q_9$ as well and the 29 missing terms in Q_{10}. [2]

This surprising result raises **two natural questions**:

[2] Just kidding.

1. Are there infinitely many more conserved charges?
2. If so, is there a systematic way to find them?

5.4 The Gardner transform

The answer to both questions is affirmative, and is based on a very clever (though at first sight unintuitive) method devised by **Gardner** [Miura et al., 1968].

/First, let us suppose that the KdV field $u(x,t)$ can be expressed in terms of another function $v(x,t)$ as

$$\boxed{u = \lambda - v^2 - v_x}\,, \tag{5.11}$$

where λ is a real parameter. Substituting (5.11) in the KdV equation we find

$$\begin{aligned} 0 &= (\lambda - v^2 - v_x)_t + 6(\lambda - v^2 - v_x)(\lambda - v^2 - v_x)_x + (\lambda - v^2 - v_x)_{xxx} \\ &= \dots \qquad \textbf{[Ex 24]} \\ &= -\boxed{\left(2v + \frac{\partial}{\partial x}\right)\left[v_t + 6(\lambda - v^2)v_x + v_{xxx}\right] = 0}\,. \end{aligned} \tag{5.12}$$

So

$$\boxed{\text{KdV for } u \quad \Longleftrightarrow \quad \text{(5.12) for } v}\,,$$

and in particular, if v solves

$$\boxed{v_t + 6(\lambda - v^2)v_x + v_{xxx} = 0}\,, \tag{5.13}$$

then u given by (5.11) solves KdV.

For $\lambda = 0$, (5.13) is the "wrong sign" mKdV equation that you encountered in **[Ex 13 (b)]**, and

$$\boxed{u = -v^2 - v_x} \tag{5.14}$$

is known as the **Miura transform**, found by Miura earlier in 1968 [Miura, 1968].

Gardner's idea was to change Miura's transformation by setting

$$\boxed{\begin{aligned} &v = \epsilon w + \frac{1}{2\epsilon} \\ &\lambda = \frac{1}{4\epsilon^2} \end{aligned}} \tag{5.15}$$

for some non-vanishing real constant ϵ. Then

$$\lambda - v^2 = \frac{1}{4\epsilon^2} - \left(\epsilon w + \frac{1}{2\epsilon}\right)^2 = -w - \epsilon^2 w^2 \ ,$$

which implies that u and w are related by the **Gardner transform (GT)**

$$\boxed{u = -w - \epsilon w_x - \epsilon^2 w^2} \ . \tag{5.16}$$

We will use the free parameter ϵ to great advantage below.

In terms of w, the **KdV equation for u**, or equivalently equation (5.12) for v becomes

$$\left(2\epsilon w + \frac{1}{\epsilon} + \frac{\partial}{\partial x}\right)\left[\epsilon w_t - 6(w + \epsilon^2 w^2)\epsilon w_x + \epsilon w_{xxx}\right] = 0 \ ,$$

or equivalently

$$\boxed{\left(1 + \epsilon\frac{\partial}{\partial x} + 2\epsilon^2 w\right)\left[w_t - 6(w + \epsilon^2 w^2)w_x + w_{xxx}\right] = 0} \ . \tag{5.17}$$

In particular, any w that solves the simpler equation

$$\boxed{w_t - 6(w + \epsilon^2 w^2)w_x + w_{xxx} = 0} \tag{5.18}$$

produces a u that solves the KdV equation by the Gardner transform (5.16).

Now we are going to think about this backwards: let's view u as a **fixed solution of KdV**, while w **varies with** ϵ so that (5.16) holds. Then

- For $\underline{\epsilon = 0}$, equation (5.17) is nothing but the KdV equation with a reversed middle term. Indeed the Gardner transform reduces to $u = -w$ in this case.

- For $\underline{\epsilon \neq 0}$, we encounter **two problems**:

  1. To obtain w in terms of u, we need to solve a **differential equation** (5.16);

  2. The **differential operator** $1 + \epsilon\frac{\partial}{\partial x} + 2\epsilon^2 w$ in (5.17) is **non-trivial**. It might have a non-vanishing kernel, so we can't immediately conclude that (5.18) holds.

Gardner's intuition was that we can **solve both problems at once** by viewing w as a **formal**

power series in ϵ:[3]

$$\boxed{w(x,t) = \sum_{n=0}^{\infty} w_n(x,t)\epsilon^n = w_0(x,t) + w_1(x,t)\epsilon + w_2(x,t)\epsilon^2 + \dots} \tag{5.19}$$

1. To solve the first problem, we substitute (5.19) in the Gardner transform (5.16)

$$\begin{aligned} u &= -(w_0 + w_1\epsilon + w_2\epsilon^2 + \dots) - \epsilon(w_0 + w_1\epsilon + w_2\epsilon^2 + \dots)_x \\ &\quad - \epsilon^2(w_0 + w_1\epsilon + w_2\epsilon^2 + \dots)^2 \\ &= \begin{array}{lllll} -w_0 & -\epsilon w_1 & -\epsilon^2 w_2 & -\epsilon^3 w_3 & + \dots \\ & -\epsilon w_{0,x} & -\epsilon^2 w_{1,x} & -\epsilon^3 w_{2,x} & + \dots \\ & & -\epsilon^2 w_0^2 & -\epsilon^3 2w_0 w_1 & + \dots \end{array} \end{aligned}$$

and invert it to determine w in terms of u. Since u is fixed, it is of order ϵ^0. Comparing order by order we obtain:

$$\epsilon^0: \qquad w_0 = -u \tag{5.20}$$
$$\epsilon^1: \qquad w_1 = -w_{0,x} = u_x \tag{5.21}$$
$$\epsilon^2: \qquad w_2 = -w_{1,x} - w_0^2 = -u_{xx} - u^2 \tag{5.22}$$
$$\epsilon^3: \qquad w_3 = -w_{2,x} - 2w_0 w_1 = u_{xxx} + 4uu_x \tag{5.23}$$
$$\vdots$$

which in principle determines recursively all the coefficients w_n of the formal power series (5.19) in terms of u.

2. Since w is a formal power series in ϵ, so is the expression inside the square brackets in (5.17):

$$\left[w_t - 6(w + \epsilon^2 w^2)w_x + w_{xxx}\right] \equiv z(x,t) = \sum_{n=0}^{\infty} z_n(x,t)\epsilon^n = z_0 + z_1\epsilon + z_2\epsilon^2 + \dots$$

The same applies to the differential operator

$$A \equiv \mathbb{1} + \epsilon\frac{\partial}{\partial x} + 2\epsilon^2 w \equiv \mathbb{1} + \sum_{n=1}^{\infty} A_n \epsilon^n ,$$

where $\mathbb{1}$ is the identity operator, and A_n are linear (differential) operators:

$$A_1 = \frac{\partial}{\partial x}, \qquad A_2 = 2w_0 \cdot , \qquad A_3 = 2w_1 \cdot , \qquad A_4 = 2w_2 \cdot , \qquad \dots$$

[3] By a formal power series we mean that we don't worry about the convergence of the series. (5.19) is actually an asymptotic expansion, for those who know what that is.

where I wrote the dots to make clear which operators act by multiplication by a function. Then (5.17) becomes the formal power series equation

$$
\begin{aligned}
0 &= \left(1+\sum_{n=1}^{\infty} A_n \epsilon^n\right)\left(\sum_{k=0}^{\infty} z_k \epsilon^k\right) \\
&= z_0 \quad +\epsilon z_1 \quad +\epsilon^2 z_2 \quad +\epsilon^3 z_3 \quad +\dots \\
&\qquad\quad +\epsilon A_1 z_0 \quad +\epsilon^2 A_1 z_1 \quad +\epsilon^3 A_1 z_2 \quad +\dots \\
&\qquad\qquad\qquad +\epsilon^2 A_2 z_0 \quad +\epsilon^3 A_2 z_1 \quad +\dots \\
&\qquad\qquad\qquad\qquad\qquad +\epsilon^3 A_3 z_0 \quad +\dots \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad +\dots
\end{aligned}
$$

which we can solve order by order as follows:

$$
\begin{aligned}
\epsilon^0 : &\qquad z_0 = 0 \\
\epsilon^1 : &\qquad z_1 = -A_1 z_0 \\
\epsilon^2 : &\qquad z_2 = -A_1 z_1 - A_2 z_0 = 0 \\
\epsilon^3 : &\qquad z_3 = -A_1 z_2 - A_2 z_1 - A_3 z_0 = 0 \\
&\qquad \vdots
\end{aligned}
\tag{5.24}
$$

Thus we have shown that, **order by order** in the formal power series **in ϵ, equation (5.18) holds**! But – punchline ahead – (5.18) is a continuity equation

$$
\boxed{\frac{\partial}{\partial t} w + \frac{\partial}{\partial x}\left(-3w^2 - 2\epsilon^2 w^3 + w_{xx}\right) = 0} \, . \tag{5.25}
$$

Since $w, w_x, w_{xx}, \dots \to 0$ as $x \to \pm\infty$ order by order in powers of ϵ, this means that the charge

$$
\boxed{\tilde{Q} = \int_{-\infty}^{+\infty} dx \; w} \tag{5.26}
$$

is **conserved**.

Now comes the important point: since $w = \sum_{n=0}^{\infty} w_n \epsilon^n$ is a formal power series in ϵ, so is the conserved charge $\tilde{Q}$:[4]

$$
\tilde{Q} = \int_{-\infty}^{+\infty} dx \sum_{n=0}^{\infty} w_n \epsilon^n = \sum_{n=0}^{\infty} \epsilon^n \int_{-\infty}^{+\infty} dx \; w_n \equiv \sum_{n=0}^{\infty} \epsilon^n \tilde{Q}_n \, .
$$

And since $\tilde{Q}$ is a conserved charge for all values of the free parameter ϵ, it must be that the **charges**

$$
\boxed{\tilde{Q}_n = \int_{-\infty}^{+\infty} dx \; w_n} \qquad (n = 0, 1, 2, \dots) \tag{5.27}
$$

[4]Strictly speaking the middle equality assumes convergence, but we are working with a formal expansion, so we don't need to worry about this subtlety.

are **all separately conserved!**

Going back to (5.22), we find that the first few conserved charges are

$$\boxed{\begin{aligned} \tilde{Q}_0 &= -\int_{-\infty}^{+\infty} dx\; u \equiv -Q_1 \\ \tilde{Q}_1 &= +\int_{-\infty}^{+\infty} dx\; u_x = [u]_{-\infty}^{+\infty} = 0 \\ \tilde{Q}_2 &= -\int_{-\infty}^{+\infty} dx\; (u_{xx} + u^2) = -\int_{-\infty}^{+\infty} dx\; u^2 \equiv -Q_2 \\ \tilde{Q}_3 &= +\int_{-\infty}^{+\infty} dx\; (u_{xxx} + 4uu_x) = [u_{xx} + 2u^2]_{-\infty}^{+\infty} = 0 \\ &\vdots \end{aligned}} \tag{5.28}$$

As you might have guessed, the **general pattern** is as follows:

$$\begin{aligned} \tilde{Q}_{2n-1} &= \int_{-\infty}^{+\infty} dx\; (\text{total derivative}) = 0 \\ \tilde{Q}_{2n-2} &= \text{const} \times Q_n = \text{const} \times \int_{-\infty}^{+\infty} dx\; (u^n + \dots) \neq 0\,. \end{aligned}$$

See [Drazin and Johnson, 1989] for a general proof.

The existence of **infinitely many conserved charges** makes the **KdV** equation **"integrable"**. As you'll see in the exercises for this chapter, these unexpected conservation laws give us a lot of information about multi-soliton solutions of the KdV equation, see **[Ex 23]** and **[Ex 25]**.

5.5 Extra conservation laws for relativistic field equations

Let's return to our other main example, the sine-Gordon model. We've already seen that energy is conserved, but this is not particularly surprising. In fact for any **relativistic field theory** in 1 space (x) + 1 time (t) dimensions (*e.g.* Klein-Gordon, sine-Gordon, "ϕ^4", ...),

$$\boxed{E = \int_{-\infty}^{+\infty} dx\; \mathcal{E} = \int_{-\infty}^{+\infty} dx \left[\frac{1}{2}u_t^2 + \frac{1}{2}u_x^2 + V(u)\right]} \tag{5.29}$$

is **conserved**, provided the equation of motion (**EoM**)

$$\boxed{u_{tt} - u_{xx} = -V'(u)} \tag{5.30}$$

is satisfied.

*** EXERCISE**: Check this statement.

The **"scalar potential"** $V(u)$ determines the theory. For instance

$$V(u) = \begin{cases} \frac{1}{2}m^2u^2 & \text{(Klein-Gordon)} \\ 1-\cos u & \text{(sine-Gordon)} \\ \frac{\lambda}{2}(u^2-a^2)^2 & \text{("}\phi^4\text{")} \\ \ldots \end{cases}$$

A deep theorem due to Emmy Noether, already mentioned in passing above, shows that the **conservation of energy** follows from the **invariance** of the theory under arbitrary **time translations** $t \mapsto t+c$. Similarly, **invariance under space translations** $x \mapsto x+c'$ implies the **conservation of momentum** P.

We will not delve into Noether's theorem, but you might encounter it in other courses. In any case, it is of limited help for our purposes: our main interest will be in more surprising, 'bonus', charges, similar to those already seen for the KdV equation in the last section. The **question** that we would like to answer is:

Can there be more conserved quantities, in addition to energy and momentum?

We will answer this question constructively.

The **first step** is to switch to **light-cone coordinates**

$$\boxed{x^\pm = \frac{1}{2}(t \pm x)} \Longleftrightarrow \begin{cases} t &= x^+ + x^- \\ x &= x^+ - x^- \end{cases}, \tag{5.31}$$

which are so called because the trajectories of light rays are $x^+ = \text{const}$ or $x^- = \text{const}$ for left-moving or right-moving rays respectively. By the chain rule we calculate

$$\partial_\pm \equiv \frac{\partial}{\partial x^\pm} = \frac{\partial t}{\partial x^\pm}\frac{\partial}{\partial t} + \frac{\partial x}{\partial x^\pm}\frac{\partial}{\partial x} = \frac{\partial}{\partial t} \pm \frac{\partial}{\partial x} \equiv \partial_t \pm \partial_x$$
$$\Longrightarrow \quad \partial_+\partial_- = \partial_t^2 - \partial_x^2 \,,$$

so the EoM can be written as

$$\boxed{u_{+-} = -V'(u)}\,, \tag{5.32}$$

where we used the shorthand notation $f_\pm \equiv \frac{\partial f}{\partial x^\pm} \equiv \partial_\pm f$.

Now suppose that a couple of densities T and X can be found such that given the equation of motion (5.32),

$$\boxed{\partial_- T = \partial_+ X}\,. \tag{5.33}$$

Converted back to the original space and time coordinates x and t, this is nothing but the continuity equation (5.3)

$$\partial_t \underbrace{(T-X)}_{\rho} - \partial_x \underbrace{(T+X)}_{-j} = 0 \,.$$

with $\rho = T - X$ and $j = -T - X$. Provided that the limiting values of $-T - X$ as $x \to \pm\infty$ agree so that (5.4) holds, this means that $\int_{-\infty}^{\infty}(T - X)dx$ will be a conserved quantity.

The goal is to construct examples of such (T, X) pairs, and to simplify life I'll suppose that T is a polynomial in x^+-derivatives of u: this means we are looking for **"polynomial conserved densities"**. We will also (mostly) disregard total x^+-derivatives in T, or in other words consider two polynomial conserved densities which differ by a total x^+-derivative to be equivalent: if (T, X) solves (5.33) and $T' = T + \partial_+ U$, then

$$\partial_- T' = \partial_- T + \partial_- \partial_+ U = \partial_+ X'$$

where $X' = X + \partial_- U$. Hence (T', X') is another solution to (5.33), but so long as the limits of U as $x \to \pm\infty$ are equal, it leads to exactly the same conserved quantity as before:

$$\int_{-\infty}^{\infty} (T' - X')\, dx - \int_{-\infty}^{\infty} (T - X)\, dx = \int_{-\infty}^{\infty} (\partial_+ U - \partial_- U)\, dx = \int_{-\infty}^{\infty} 2\partial_x U\, dx = [2U]_{-\infty}^{\infty} = 0\,.$$

One more concept is useful: the **"rank"**, or **"Lorentz spin"** of a single term in a general polynomial in u and its light-cone derivatives is the **number of ∂_+ derivatives minus the number of ∂_- derivatives**. For instance $(u_+)^3 u_- u_{++-}$ has Lorentz spin $3 - 1 + (2 - 1) = 3$. According to the theory of special relativity, objects of different spins transform differently under the "Lorentz group" of symmetries of relativistic field equations. If you would like to know more about Lorentz transformations and Lorentz spin, you can read this optional note. Terms with different Lorentz spins will never cancel against each other in (5.33), since using the equation of motion (5.32) to convert an occurance of u_{+-} into $-V'(u)$ does not affect the rank. As a result, each spin can be considered separately and so, for $s = 0, 1, 2 \ldots$, we will look for solutions (T_{s+1}, X_{s-1}) to (5.33), where T_{s+1} is a polynomial in the x^+-derivatives of u with Lorentz spin $s + 1$. Via (5.33), X_{s-1} must then have spin $s-1$. The corresponding conserved charge will be written as Q_s:

$$\boxed{Q_s = \int_{-\infty}^{+\infty} dx\, (T_{s+1} - X_{s-1})} \tag{5.34}$$

As $x \to \pm\infty$ we'll assume that all derivatives of u tend to zero, but (to allow for topological lumps) u itself might tend to other, possibly unequal, values. Notice also that for each pair (T_{s+1}, X_{s-1}) the roles of x^+ and x^- can be swapped throughout to find a partner pair (T_{-s-1}, X_{-s+1}) where T_{-s-1} is a polynomial in x^- derivatives, with Lorentz spin $-s-1$.

Proceeding spin by spin:

$\boxed{s=0}$ $T_1 = u_+$

is the unique polynomial density of spin 1, up to an irrelevant multiplicative factor which can be absorbed in the normalisation of the charge. It solves (5.33) with $X_{-1} = u_-$, since $\partial_- u_+ = u_{-+} = u_{+-} = \partial_+ u_-$. The corresponding spin zero conserved charge is the **topological charge**

$$Q_0 = \int_{-\infty}^{+\infty} dx\ (u_+ - u_-) = 2\int_{-\infty}^{+\infty} dx\ u_x = 2[u]_{-\infty}^{+\infty}\ .$$

Note: T_1 differs from zero by a total x^+-derivative, $T_1 = 0 + \partial_+ U$ with $U = u$, so by the rules above we might want to discard it. That would be too hasty, since this U could have different limits as $x \to \pm\infty$, in fact, this happens precisely in those cases where the topological charge is non-trivial.

$\boxed{s=1}$ $T_2 \supset u_{++},\ u_+^2,$

which is a shorthand for: T_2 is a linear combination of u_{++} and u_+^2. However $u_{++} = (u_+)_+$ is a total derivative, and since $u_+ \to 0$ as $x \to \pm\infty$ we can disregard this term without loss of generality, and consider $T_2 = u_+^2$. Then

$$\partial_- T_2 = \partial_- u_+^2 = 2u_+ u_{+-} \underset{\text{EoM}}{=} -2V'(u)u_+ = -2\partial_+ V(u) \equiv \partial_+ X_0$$

with $X_0 = -2V(u)$. Therefore

$$\boxed{Q_1 = \int_{-\infty}^{+\infty} dx\ (T_2 - X_0) = \int_{-\infty}^{+\infty} dx\ [u_+^2 + 2V(u)]} \tag{5.35}$$

is conserved, for any V. Swapping x^+ and x^-, $T_{-2} = u_-^2$ is another conserved density, with the same X_0, leading to

$$\boxed{Q_{-1} = \int_{-\infty}^{+\infty} dx\ (T_2 - X_0) = \int_{-\infty}^{+\infty} dx\ [u_-^2 + 2V(u)]} \tag{5.36}$$

Taking the sum and difference and choosing a convenient normalization, we find two conserved charges

$$\frac{1}{4}(Q_1 + Q_{-1}) = \int_{-\infty}^{+\infty} dx \left[\frac{1}{4}(u_+^2 + u_-^2) + V(u)\right]$$
$$\equiv \boxed{E = \int_{-\infty}^{+\infty} dx \left[\frac{1}{2}u_t^2 + \frac{1}{2}u_x^2 + V(u)\right]} \tag{5.37}$$
$$\frac{1}{4}(Q_{-1} - Q_1) = \int_{-\infty}^{+\infty} dx \, \frac{1}{4}(u_-^2 - u_+^2)$$
$$\equiv \boxed{P = -\int_{-\infty}^{+\infty} dx \, u_t u_x} , \tag{5.38}$$

which are interpreted as the **energy** E and the **momentum** P.

$\boxed{s = 2}$ $T_3 \supset u_{+++}, \; u_{++}u_+, \; u_+^3,$

but $u_{+++} = (u_{++})_+$ and $u_{++}u_+ = \frac{1}{2}(u_+^2)_+$ are total derivatives of functions which vanish at spatial infinity, hence they can be disregarded. So without loss of generality we can take $T_3 = u_+^3$ and then

$$\partial_- T_3 = \partial_- u_+^3 = 3u_+^2 u_{+-} \underset{\text{EoM}}{=} -3V'(u)u_+^2 \; .$$

The RHS of the previous equation **cannot be a total x^+-derivative**, because the highest x^+ derivative of u (in this case u_+) does not appear linearly.

*** EXERCISE**: Think about it and convince yourself that this statement is correct. Suppose that $\partial_+^n u$ is the highest x^+-derivative of u appearing in a function Y of u and its x^+-derivatives. How does the highest x^+-derivative of u appear in $\partial_+ Y$ then?

We learn therefore that there is **no conserved charge Q_2 of spin 2** built out of polynomial conserved densities.

$\boxed{s = 3}$ $T_4 \supset u_{++++}, \; u_{+++}u_+, \; u_{++}^2, \; u_{++}u_+^2, \; u_+^4,$

but we can drop the first and fourth term as they are total derivatives of functions which vanish at spatial infinity. Moreover $u_{+++}u_+ = -u_{++}^2 + (u_{++}u_+)_+$, so we can also disregard one of $u_{+++}u_+$ and u_{++}^2 without loss of generality. The most general expression for T_4 up to an irrelevant total x^+-derivative is therefore

$$\boxed{T_4 = u_{++}^2 + \frac{1}{4}\lambda^2 u_+^4} , \tag{5.39}$$

where λ is a constant to be determined below and the factor of $1/4$ was inserted for later convenience.[5] Then

$$\begin{aligned}\partial_- T_4 &= 2u_{++}u_{++-} + \lambda^2 u_+^3 u_{+-} \\ &\underset{\text{EoM}}{=} -2u_{++}\left(V'(u)\right)_+ - \lambda^2 u_+^3 V'(u) \\ &= -2u_{++}u_+ V''(u) - \lambda^2 u_+^3 V'(u)\,.\end{aligned}$$

This may not seem very promising, but the highest derivative in the first term occurs linearly, allowing a total derivative to be extracted using the trick familiar from integration by parts:

$$\begin{aligned}&= -(u_+^2 V''(u))_+ + u_+^3 V'''(u) - \lambda^2 u_+^3 V'(u) \\ &= -(u_+^2 V''(u))_+ + u_+^3\left[V'''(u) - \lambda^2 V'(u)\right]\,. \end{aligned} \tag{5.40}$$

We are hoping to obtain a total x^+-derivative. The first term in (5.40) *is* a total x^+-derivative, but in the second term the highest derivative, which is u_+, does not appear linearly but rather to the third power. By the previous argument which was the topic of the exercise, the second term is a total x^+-derivative if and only if

$$\boxed{V'''(u) - \lambda^2 V'(u) = 0}\,. \tag{5.41}$$

If (5.41) holds, we have $X_2 = -u_+^2 V''(u)$ and

$$\boxed{Q_3 = \int_{-\infty}^{+\infty} dx (T_4 - X_2) = \int_{-\infty}^{+\infty} dx\,\left[u_{++}^2 + \frac{1}{4}\lambda^2 u_+^4 + u_+^2 V''(u)\right]} \tag{5.42}$$

is a **conserved charge of spin 3**. If instead (5.35) does not hold, there is no extra (polynomial) conserved charge of spin 3.

To summarize, the **relativistic field theories** which have an **extra conserved charge of spin 3** are those with a scalar potential $V(u)$ which satisfies equation (5.41) for some value of the constant λ. Let us examine the various possibilities:

1. $\boxed{\lambda^2 = 0}$: $V(u) = A + B(u - u_0)^2$,

   where A and B are constants. Up to a linear redefinition of u, this scalar potential leads to the **Klein-Gordon equation**. This is a linear equation which describes a free field (*i.e.* a field free from interactions) and is therefore not interesting from the point of view of solitons.

[5] To be precise, T_4 should be written as a linear combination of u_{++}^2 and u_+^4. It turns out that the coefficient of u_{++} must be non-vanishing, hence we can normalise it to 1.

2. $\boxed{\lambda^2 \neq 0}$: $\qquad V(u) = A + Be^{\lambda u} + Ce^{-\lambda u}$,

   where A, B and C are constants.

   a) If only one of B, C is non-vanishing, the EoM is either

   $$\underline{C = 0:} \quad u_{+-} = -B\lambda e^{\lambda u} \qquad \text{or} \qquad \underline{B = 0:} \quad u_{+-} = C\lambda e^{-\lambda u} \ .$$

   By a linear redefinition of u, we can always rewrite the EoM as the **Liouville equation**

   $$\boxed{u_{+-} = e^u} \ . \tag{5.43}$$

   b) If neither B or C vanish, then by a linear redefinition of u we can write the EoM as the **sine-Gordon equation**

   $$\boxed{u_{+-} = -\sin u} \tag{5.44}$$

   if $\lambda^2 < 0$, or as the **sinh-Gordon equation**

   $$\boxed{u_{+-} = -\sinh u} \tag{5.45}$$

   if $\lambda^2 > 0$.

The equations (5.43)-(5.45) are special: they have "hidden" conservation laws that generic interacting relativistic field equations $u_{+-} = -V'(u)$ lack. More can be done in this direction – in particular, it is possible to show that the extra charge just found for Sine-Gordon is the first of an infinite sequence, just like for KdV – but instead the next chapter will look into how the sine-Gordon kinks scatter against each other.

Chapter 6

Bäcklund transformations

The main reference for this chapter is §5.4 of [Drazin and Johnson, 1989].

So far, we have constructed solutions for moving solitons only as travelling waves, which describe the propagation of a single soliton. Our next goal will be to construct analytic solutions for multiple colliding solitons. In these cases it won't be possible to reduce the partial differential equation to an ordinary differential equation, so the existence of such exact solutions is much more surprising. The method that we will use in this chapter is a solution-generating technique called the **Bäcklund transformation**.

The method was introduced in the late 19th century by the Swedish mathematician **Albert Victor Bäcklund** and by the Italian mathematician **Luigi Bianchi**[1] in the 1880s to map between pairs of surfaces in three-dimensional space. The sine-Gordon equation appears in this context when one considers hyperboloids, which are surfaces of negative curvature.

There are two main uses of the Bäcklund transformation:

1. To generate **solutions of a difficult PDE from solutions of a simpler PDE**;

2. To generate **new solutions of a given PDE from already known solutions of the same PDE**.

We will mostly be interested in use 2, but you will see examples of use 1 in Ex 26-28 in the

[1] who, notably, was born Parma, the hometown of next term's lecturer. This is the same Bianchi after whom the Bianchi identities in differential geometry and general relativity are named.

problem sheet. Our final goal in this chapter will be to obtain multi-soliton solutions of the sine-Gordon equation.

6.1 Definition

Consider two functions u and v, and two differential equations

$$\boxed{P[u] = 0} \tag{6.1}$$

$$\boxed{Q[v] = 0} \tag{6.2}$$

where P and Q are two differential operators.

If there is a pair of relations (which could be differential equations)

$$\boxed{R_1[u, v] = 0\,, \quad R_2[u, v] = 0} \tag{6.3}$$

between u and v such that

- If $P[u] = 0$, *i.e.* (6.1), then (6.3) can be solved for v, to give a solution of (6.2), $Q[v] = 0$;

- If $Q[v] = 0$, *i.e.* (6.2), then (6.3) can be solved for u, to give a solution of (6.1), $P[u] = 0$;

then (6.3) is called a **Bäcklund transformation (BT)**. If furthermore $P = Q$, so that the two differential equations are identical, then (6.3) is called an **auto-Bäcklund transformation (a-BT)**.

This is useful if (6.3) is easier to solve than (6.1) or (6.2). Then we can use (6.3) to generate solutions of the harder equation from solutions of the easier equation. If $P = Q$, we can start from a simple seed solution (*e.g.* $u = 0$) to generate new non-trivial solutions.

Vocabulary:
- (6.1) and (6.2) are **"integrability conditions"** for the Bäcklund transformation (6.3).
- (6.3) can be integrated for v if the integrability condition $P[u] = 0$ is satisfied.
- (6.3) can be integrated for u if the integrability condition $Q[v] = 0$ is satisfied.

6.2 A simple example

Take the two-dimensional Laplace operator $P = Q = \partial_x^2 + \partial_y^2$ in (6.1) and (6.2):

$$P[u] = u_{xx} + u_{yy} = 0 \tag{6.4}$$

$$Q[v] = v_{xx} + v_{yy} = 0 \tag{6.5}$$

and for the Bäcklund transformation (6.3)

$$\begin{aligned} R_1[u,v] &= u_x - v_y = 0 \\ R_2[u,v] &= u_y + v_x = 0 \; . \end{aligned} \tag{6.6}$$

Let us check that (6.4)-(6.5) are integrability conditions for (6.6). Differentiating (6.6) with respect to x and y and adding or subtracting we find

$$\begin{aligned} 0 &= +\partial_x R_1 + \partial_y R_2 = +u_{xx} - v_{yx} + u_{yy} + v_{xy} = u_{xx} + u_{yy} \\ 0 &= -\partial_y R_1 + \partial_x R_2 = -u_{xy} + v_{yy} + u_{yx} + v_{xx} = v_{xx} + v_{yy} \; , \end{aligned}$$

therefore the relations (6.6) imply (6.4) and (6.5).[2] This shows that (6.6) is an auto-Bäcklund transformation for the two-dimensional Laplace equation.

EXAMPLE:

$v(x,y) = 2xy$ solves the Laplace equation (6.5). Let us use the a-BT to find another solution u of the same equation:

$$\begin{cases} u_x = v_y = 2x \\ u_y = -v_x = -2y \end{cases} \quad \Longrightarrow \quad \begin{cases} u = x^2 + f(y) \\ f'(y) = -2y \quad \Rightarrow \quad f(y) = -y^2 + c \; , \end{cases}$$

so we find the function $u(x,y) = x^2 - y^2 + c$, where c is a constant. It is immediate to check that this u solves the Laplace equation (6.4).

The equations $R_1[u,v] = R_2[u,v] = 0$ in (6.6) are nothing but the **Cauchy-Riemann equations** for the **holomorphic** (= complex analytic) **function** $w = u + iv$ of the complex variable $z = x + iy$. In the example above, $w(z) = z^2 + c$. The equations $P[u] = 0$ and $Q[v] = 0$ in (6.4)-(6.5) simply state that the real and imaginary parts of a holomorphic function are harmonic, that is, they solve the Laplace equation. Two such functions u and v are often called **harmonic conjugate** of each other.

REMARKS:

1. Given v, the Bäcklund transformation (6.6) is a system of two equations for u. Generically there won't be any solutions for the system (6.6). For example, if we pick $v = x^2$, then the system
$$\begin{cases} u_x = v_y = 0 \\ u_y = -v_x = -2x \end{cases}$$
has no solutions for u. But $v = x^2$ doesn't solve (6.5)! The integrability condition (6.5) is what guarantees that the system (6.6) can be consistently solved for u.

[2]Note: in this example we don't even need to use the other differential equation. This is not always the case.

2. This auto-Bäcklund transformation generates a new solution to the Laplace equation from a seed solution, but if we apply it a second time we get back the original seed solution (up to an irrelevant integration constant that we can ignore). So this auto-Bäcklund transformation is an involution. To get further solutions we will need to introduce a parameter.

6.3 The Bäcklund transformation for sine-Gordon

Recall that the sine-Gordon equation written in light-cone coordinates $x^{\pm} = \frac{1}{2}(t \pm x)$ is

$$\boxed{u_{+-} = -\sin u}. \tag{6.7}$$

Let us try the Bäcklund transformation

$$\boxed{\begin{aligned} (u-v)_+ &= \frac{2}{a}\sin\left(\frac{u+v}{2}\right) \\ (u+v)_- &= -2a\sin\left(\frac{u-v}{2}\right) \end{aligned}} \tag{6.8}$$

where a is a (non-zero) parameter. Cross-differentiating, and recalling that $\sin(A \pm B) = \sin A \cos B \pm \cos A \sin B$, which implies $\sin(A+B) + \sin(A-B) = 2\sin A\cos B$,

$$\begin{aligned} (u-v)_{+-} &= \frac{1}{a}\cos\left(\frac{u+v}{2}\right) \cdot (u+v)_- = -2\cos\left(\frac{u+v}{2}\right)\sin\left(\frac{u-v}{2}\right) \\ &= -\sin u + \sin v \\ (u+v)_{-+} &= -a\cos\left(\frac{u-v}{2}\right) \cdot (u-v)_+ = -2\cos\left(\frac{u-v}{2}\right)\sin\left(\frac{u+v}{2}\right) \\ &= -\sin u - \sin v\ . \end{aligned}$$

Adding and subtracting, we find that both u and v obey the sine-Gordon equation:

$$\boxed{u_{+-} = -\sin u} \tag{6.9}$$

$$\boxed{v_{+-} = -\sin v} \tag{6.10}$$

Therefore (6.8) is an auto-Bäcklund transformation for the sine-Gordon equation, for any non-zero value of a. The extra parameter will allow us to generate multi-soliton solutions. We will start in the next section by rederiving the one-kink solution.

6.4 First example: the sine-Gordon kink from the vacuum

Let us take the vacuum solution

$$\boxed{v = 0} \tag{6.11}$$

as our initial (seed) solution. Then the auto-Bäcklund transformation (6.8) is

$$\begin{aligned} u_+ &= \frac{2}{a}\sin\frac{u}{2} \\ u_- &= -2a\sin\frac{u}{2}\ . \end{aligned} \tag{6.12}$$

We can integrate both equations by separation of variables, using the indefinite integral

$$\int \frac{du}{\sin\frac{u}{2}} = 2\log\tan\frac{u}{4}$$

up to an integration constant. We get

$$\begin{cases} \frac{2}{a}x^+ = 2\log\tan\frac{u}{4} + f(x^-) \\ -2ax^- = 2\log\tan\frac{u}{4} + g(x^+) \end{cases} \tag{6.13}$$

where the functions f and g are "constants" of integration. They are only constant with respect to the variable that is integrated, but they can (and do!) depend on the other variable.

Subtracting and rearranging, we get

$$\frac{2}{a}x^+ + g(x^+) = -2ax^- + f(x^-)\ . \tag{6.14}$$

The left-hand-side is only a function of x^+, while the right-hand-side is only a function of x^-. Since the two sides are equal, they must therefore be equal to a constant, which we set to be $-2c$ for future convenience. Hence

$$\begin{aligned} f(x^-) &= 2ax^- - 2c \\ g(x^+) &= -\frac{2}{a}x^+ - 2c \end{aligned}$$

and so

$$2\log\tan\frac{u}{4} = \frac{2}{a}x^+ - 2ax^- + 2c\ ,$$

that is

$$\boxed{u = 4\arctan\left(e^{\frac{1}{a}x^+ - ax^- + c}\right)}\ . \tag{6.15}$$

Finally, we convert to (x,t) coordinates:

$$\frac{1}{a}x^+ - ax^- = \frac{1}{2a}(t+x) - \frac{a}{2}(t-x) = \frac{1}{2}\left[\left(a+\frac{1}{a}\right)x - \left(a-\frac{1}{a}\right)t\right] = \frac{1+a^2}{2a}\left(x - \frac{a^2+1}{a^2-1}t\right).$$

Defining

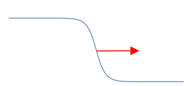
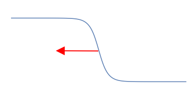
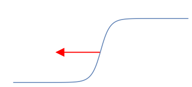
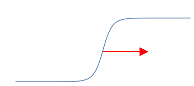
$$\boxed{\begin{aligned} v &:= \frac{a^2-1}{a^2+1} \\ \epsilon &:= \operatorname{sign}(a) \\ \gamma &:= \frac{1}{\sqrt{1-v^2}} \underset{*\text{Ex}}{=} \frac{1+a^2}{2|a|} \end{aligned}}\,, \tag{6.16}$$

the solution (6.15) generated by an auto-Bäcklund transformation of the vacuum is

$$\boxed{u(x,t) = 4\arctan\left(e^{\epsilon\gamma(x-x_0-vt)}\right)}\,, \tag{6.17}$$

where we traded the integration constant c for x_0. This solution describes a kink or an anti-kink moving at velocity v.

Properties: $a > 0$: kink $\quad |a| > 1$: right-moving
$a < 0$: anti-kink $\quad |a| < 1$: left-moving

$a < -1$:	$-1 < a < 0$	$0 < a < 1$	$a > 1$
Right-moving anti-kink	Left-moving anti-kink	Left-moving kink	Right-moving kink
			

So the auto-Bäcklund transformation creates a kink/anti-kink from the vacuum! By varying the parameter $a \in \mathbb{R}\backslash\{0\}$ and the integration constant x_0 or c, we reproduce all the kink and anti-kink solutions derived in section 3.2 as travelling waves.

The amazing fact is that this holds more generally: the auto-Bäcklund transformation (almost) *always* adds a kink or an anti-kink to the seed solution.[3] (The only exception is if one tries to add a soliton with the same velocity as one already present.) Therefore we can think of the auto-Bäcklund transformation as a **solution-generating technique** which “adds” kinks or anti-kinks.

We will use the following graph to denote the action of a Bäcklund transformation on with parameter a and integration constant c on a seed solution u_1, which adds a kink or anti-kink and generates the new solution u_2:

[3]Which of the two is added depends on the seed. More about this later.

$$u_1 \xrightarrow[c]{a} u_2$$

We can add a kink/anti-kink wherever we like (by choosing c) and with whatever velocity we like (by choosing a). For example

$$u_0 \xrightarrow[c_1]{a_1} u_1 \xrightarrow[c_2]{a_2} u_2 \xrightarrow[c_3]{a_3} u_3$$

adds three kinks/anti-kinks to the seed solution u_0.

The problem with this is that the integrations get harder and harder as we keep adding solitons. Luckily, a nice theorem tells us that, having found one-soliton solutions, we can obtain multi-soliton solutions without doing any further integrals.

6.5 The theorem of permutability

Let's apply the Bäcklund transformation twice, with parameters a_1 and a_2, in the two possible orders:

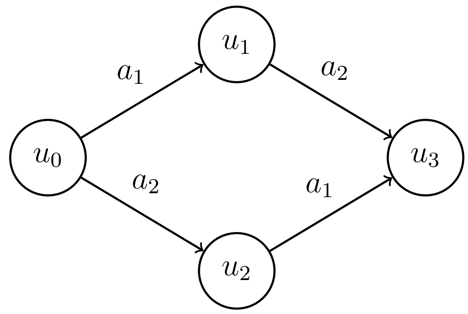
$$\begin{tikzcd} & u_1 \arrow[r, "a_2"] & u_3 \\ u_0 \arrow[ur, "a_1"] \arrow[dr, "a_2"'] & & \\ & u_2 \arrow[r, "a_1"] & u_4 \end{tikzcd}$$

The final results u_3 and u_4 both look like the seed solution u_0 with two added solitons, with parameters a_1 and a_2. Could they actually be the same solution? The answer is yes, according to the following theorem:

THEOREM (Bianchi 1902):

For any u_1 and u_2, the integration constants in the second Bäcklund transformations, which generate u_3 and u_4, can be arranged such that u_3 and u_4 are equal.

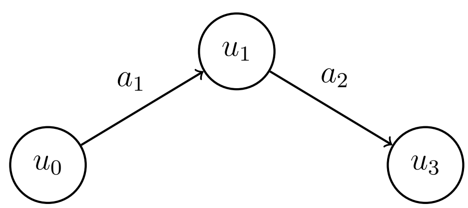
In other words, the a_1 and a_2 BT's can be made to **commute**. Diagrammatically:



I will spare you the proof of the theorem, which is a bit involved. Hopefully the statement makes intuitive sense, given the soliton content of u_3 and u_4.

This result has a nice application. We have two ways of getting to u_3 from u_0: either through u_1 or through u_2. By comparing these two ways we will be able to get rid of all derivatives in the Bäcklund transformations and thereby obtain an algebraic relation between the four solutions u_0, u_1, u_2, u_3.

Let's start by considering the ∂_+ parts of the transformations, and let's look at the upper route first:



We have

$$\begin{aligned} (u_1 - u_0)_+ &= \frac{2}{a_1} \sin \frac{u_1 + u_0}{2} \\ (u_3 - u_1)_+ &= \frac{2}{a_2} \sin \frac{u_3 + u_1}{2} \ . \end{aligned} \tag{6.18}$$

Adding the two equations to cancel u_1 out in the left-hand side, we get

$$\boxed{(u_3 - u_0)_+ = \frac{2}{a_1} \sin \frac{u_1 + u_0}{2} + \frac{2}{a_2} \sin \frac{u_3 + u_1}{2}} \ . \tag{6.19}$$

For the lower route

$$\begin{tikzcd}
u_0 \arrow[dr, "a_2"] & & u_3 \\
& u_2 \arrow[ur, "a_1"] &
\end{tikzcd}$$

we swap $a_1 \leftrightarrow a_2$, $u_1 \leftrightarrow u_2$ and get

$$\boxed{(u_3 - u_0)_+ = \frac{2}{a_2} \sin \frac{u_2 + u_0}{2} + \frac{2}{a_1} \sin \frac{u_3 + u_2}{2}} \,. \tag{6.20}$$

We have found two different expressions for $(u_3 - u_0)_+$. Equating them, we obtain an algebraic relation between u_0, u_1, u_2, u_3:

$$\boxed{\frac{1}{a_1} \sin \frac{u_1 + u_0}{2} + \frac{1}{a_2} \sin \frac{u_3 + u_1}{2} = \frac{1}{a_2} \sin \frac{u_2 + u_0}{2} + \frac{1}{a_1} \sin \frac{u_3 + u_2}{2}} \,. \tag{6.21}$$

This is very useful: for example, starting from u_0 equal to the vacuum and two one-soliton solutions u_1, u_2, we can generate a 2-soliton solution u_3 algebraically. We can then iterate the procedure and get a 3-soliton solution, then a 4-soliton solution, and so on and so forth. What we have found is akin to a **"non-linear superposition principle"**: the Bäcklund transformation and the permutability theorem provide us with a machinery to "add" solutions of a non-linear equation!

To check that this procedure is consistent, let's see what happens for the ∂_- part of the Bäcklund transformations. For the upper route

$$\begin{tikzcd}
& u_1 \arrow[dr, "a_2"] & \\
u_0 \arrow[ur, "a_1"] & & u_3
\end{tikzcd}$$

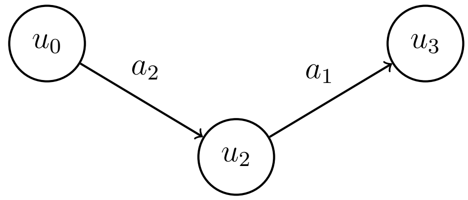
we have

$$\begin{aligned}
(u_1 + u_0)_- &= -2a_1 \sin \frac{u_1 - u_0}{2} \\
(u_3 + u_1)_- &= -2a_2 \sin \frac{u_3 - u_1}{2} \,.
\end{aligned} \tag{6.22}$$

Subtracting the two equations we get

$$\boxed{(u_0 - u_3)_- = 2a_2 \sin \frac{u_3 - u_1}{2} - 2a_1 \sin \frac{u_1 - u_0}{2}} \,. \tag{6.23}$$

For the lower route



we swap again $a_1 \leftrightarrow a_2$, $u_1 \leftrightarrow u_2$ and get

$$\boxed{(u_0 - u_3)_- = 2a_1 \sin\frac{u_3 - u_2}{2} - 2a_2 \sin\frac{u_2 - u_0}{2}} \,. \tag{6.24}$$

Equating (6.23) and (6.24), we find the algebraic relation

$$\boxed{a_2 \sin\frac{u_3 - u_1}{2} - a_1 \sin\frac{u_1 - u_0}{2} = a_1 \sin\frac{u_3 - u_2}{2} - a_2 \sin\frac{u_2 - u_0}{2}} \,. \tag{6.25}$$

Consistency requires that the two algebraic relations (6.21) and (6.25) agree. To see that, let's first rewrite (6.21) in the following form:

$$\frac{1}{a_1}\left(\sin\frac{u_1 + u_0}{2} - \sin\frac{u_3 + u_2}{2}\right) = \frac{1}{a_2}\left(\sin\frac{u_2 + u_0}{2} - \sin\frac{u_3 + u_1}{2}\right) \,.$$

Multiplying by $a_1 a_2/2$ and using the identity $\sin A \pm \sin B = 2\sin\frac{A \pm B}{2}\cos\frac{A \mp B}{2}$, this becomes

$$\boxed{\begin{aligned} &a_2 \sin\frac{u_1 + u_0 - u_3 - u_2}{4} \cancel{\cos\frac{u_1 + u_0 + u_3 + u_2}{4}} \\ &\qquad = a_1 \sin\frac{u_2 + u_0 - u_3 - u_1}{4} \cancel{\cos\frac{u_2 + u_0 + u_3 + u_1}{4}} \end{aligned}} \tag{6.26}$$

where we are allowed to simplify the common cosine factor in the two sides because the argument is a function of x and t which is generically different from $\pi/2$ modulo π.

Similarly, (6.25) can be rearranged as

$$a_1\left(\sin\frac{u_3 - u_2}{2} + \sin\frac{u_1 - u_0}{2}\right) = a_2\left(\sin\frac{u_3 - u_1}{2} + \sin\frac{u_2 - u_0}{2}\right) \,,$$

which upon using the same trigonometric identity as above becomes

$$\boxed{\begin{aligned} &a_1 \sin\frac{u_3 - u_2 + u_1 - u_0}{4} \cancel{\cos\frac{u_3 - u_2 - u_1 + u_0}{4}} \\ &\qquad = a_2 \sin\frac{u_3 - u_1 + u_2 - u_0}{4} \cancel{\cos\frac{u_3 - u_1 - u_2 + u_0}{4}} \end{aligned}} \tag{6.27}$$

which agrees with equation (6.26) upon simplification. So everything is consistent.

To conclude this discussion, let's manipulate (the simplified version of) equation (6.26) a bit further, with the aim of determining u_3 given u_0, u_1 and u_2. Letting $A = (u_0 - u_3)/4$ and $B = (u_1 - u_2)/4$, (6.26) becomes

$$\begin{aligned} & a_1 \sin(A - B) = a_2 \sin(A + B) \\ \implies \quad & a_1(\sin A \cos B - \sin B \cos A) = a_2(\sin A \cos B + \sin B \cos A) \; . \end{aligned}$$

Dividing through by $\cos A \ \cos B$, we find

$$\begin{aligned} & a_1(\tan A - \tan B) = a_2(\tan A + \tan B) \; . \\ \implies \quad & (a_1 - a_2) \tan A = (a_1 + a_2) \tan B \; . \end{aligned}$$

In terms of u_0, u_1, u_2, u_3, this reads

$$\boxed{\tan \frac{u_0 - u_3}{4} = \frac{a_1 + a_2}{a_1 - a_2} \tan \frac{u_1 - u_2}{4}} \; , \tag{6.28}$$

which is an improvement on (6.26) since u_3 appears only once. Equivalently, we can write

$$\boxed{\tan \frac{u_3 - u_0}{4} = \frac{a_2 + a_1}{a_2 - a_1} \tan \frac{u_1 - u_2}{4}} \; . \tag{6.29}$$

Either of (6.28) or (6.29) allow us to express u_3 in terms of u_0, u_1, u_2.

6.6 The two-soliton solution

Finally a payoff. Take the vacuum as the seed solution, *i.e.* $u_0 = 0$. Then u_1 and u_2 are known from before: they are single kinks or antikinks. Equation (6.29) gives the double Bäcklund transformed u_3 as

$$\tan \frac{u_3}{4} = \frac{a_2 + a_1}{a_2 - a_1} \tan \frac{u_1 - u_2}{4} = \frac{a_2 + a_1}{a_2 - a_1} \frac{\tan \frac{u_1}{4} - \tan \frac{u_2}{4}}{1 + \tan \frac{u_1}{4} \tan \frac{u_2}{4}} \; , \tag{6.30}$$

where we used the trigonometric identity

$$\tan(A - B) = \frac{\tan A - \tan B}{1 + \tan A \cdot \tan B}$$

for the second equality. The 1-soliton (*i.e.* kink or antikink) solutions are

$$\boxed{\tan \frac{u_i}{4} = e^{\theta_i}} \quad (i = 1, 2) \tag{6.31}$$

where

$$\boxed{\theta_i = \frac{x^+}{a_i} - a_i x^- + c_i = \epsilon_i \gamma_i (x - \bar{x}_i - v_i t)}\ , \tag{6.32}$$

as seen in section 6.4. Here $\bar{x}_{1,2}$ are the centres of the two solitons at $t = 0$. Substituting equation (6.31) in equation (6.30) we find the **2-soliton solution**xx

$$\boxed{\tan \frac{u_3}{4} = \mu \frac{e^{\theta_1} - e^{\theta_2}}{1 + e^{\theta_1 + \theta_2}}} \tag{6.33}$$

where

$$\boxed{\mu = \frac{a_2 + a_1}{a_2 - a_1}} \tag{6.34}$$

REMARK:
If the two solitons have the same velocity $v_1 = v_2$, which means

$$\frac{a_1^2 - 1}{a_1^2 + 1} = \frac{a_2^2 - 1}{a_2^2 + 1} \qquad \Longrightarrow \qquad a_1 = \pm a_2\ ,$$

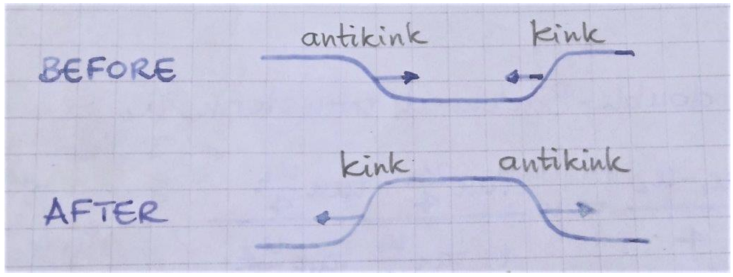
then $\mu = 0$ or ∞ and the 2-soliton solution (6.33) breaks down. In particular, there is no static 2-soliton solution! As we will see later, this is because the two solitons exert a force on one another.

But this is too fast. We haven't confirmed yet that equation (6.33) contains two solitons. Let's understand that next.

6.7 Asymptotics of multisoliton solutions

We will focus here on the 2-soliton solution of the sine-Gordon equation, but the method applies more generally to any multi-soliton solutions of integrable equations (*e.g.* the KdV equation).

Our goal will be to study the new solution (6.33) and identify two solitons hidden in its asymptotics for $t \to \mp\infty$, namely BEFORE and AFTER the collision. Here is an example of what the solution may look like at early times (before the collision) and at late times (after the collision) in the case of a collision of a kink and an anti-kink:



It is not completely obvious how to find the early time and late time asymptotics analytically. If we just take $t \pm \infty$ with x fixed, the two solitons will be at spatial infinity and we will miss them (unless one of the two has zero velocity, in which case we will see that soliton). We should instead follow one or the other soliton by letting

$$\boxed{t \to \pm\infty \qquad \text{with} \qquad X_V = x - Vt \quad \text{fixed}}\,, \tag{6.35}$$

for some appropriate constant velocity V. If there is a soliton moving at velocity V in the original (x, t) coordinates, it will appear stationary in the (X_V, t) coordinates. For this reason (X_V, t) is called a **"comoving frame"**: they are coordinates for a reference frame which moves together with an object (*e.g.* a soliton) of velocity V.

Let us try this for the solution (6.33) which we obtained from a double Bäcklund transformation of the vacuum. We will now use u to denote the field in the resulting solution, which reads

$$\tan\frac{u}{4} = \mu\frac{e^{\theta_1} - e^{\theta_2}}{1 + e^{\theta_1+\theta_2}}$$

with

$$\mu = \frac{a_2 + a_1}{a_2 - a_1}\,, \qquad \theta_i = \epsilon_i\gamma_i(x - v_i t - \bar{x}_i)\,.$$

If we switch to a comoving frame with velocity V, the exponents read

$$\begin{aligned}\theta_i &= \epsilon_i\gamma_i(x - Vt + Vt - v_i t - \bar{x}_i)\\ &= \epsilon_i\gamma_i(X_V - (v_i - V)t - \bar{x}_i)\,,\end{aligned} \tag{6.36}$$

where we see the appearance of the "relative velocity" $v_i - V$, that is the velocity in the comoving frame.

For each soliton we now have three cases for the limit (6.35), corresponding to a positive, zero or negative relative velocity for the soliton:

Case	$t \to -\infty$	$t \to +\infty$
$V < v_i$	$\theta_i \to +\epsilon_i\infty$	$\theta_i \to -\epsilon_i\infty$
$V = v_i$	θ_i finite	θ_i finite
$V > v_i$	$\theta_i \to -\epsilon_i\infty$	$\theta_i \to +\epsilon_i\infty$

Recall that $\epsilon_i = \pm 1$ is a sign, and $\gamma_i > 0$ so it does not affect the sign of θ_i in the limit.

This tells us that if $\underline{V \neq v_1, v_2}$, then $\theta_1, \theta_2 \to \pm\infty$ as $|t| \to \infty$. This implies that[4]

$$\tan\frac{u}{4} = \mu\frac{e^{\theta_1} - e^{\theta_2}}{1 + e^{\theta_1+\theta_2}} \to \pm\infty \text{ or } 0 \; .$$

So $u/4$ tends to an integer multiple of $\pi/2$, which means that u tends to an integer multiple of 2π: the field is in the vacuum. The conclusion is that if we go off to infinity in the original (x, t) plane in any direction apart from $\frac{dx}{dt} = v_1, v_2$, then $u \to 2\pi n$ for some $n \in \mathbb{Z}$.

If instead $\underline{V = v_1 \text{ or } v_2}$, we need to study the limit more carefully. We will consider a single case $\underline{a_1, a_2 > 0}$, leaving the other cases for the exercises. Since $a_1 \neq a_2$ for the solution to exist, let us take without loss of generality

$$\boxed{a_2 > a_1 > 0} \qquad \Longrightarrow \qquad v_2 > v_1 \; , \quad \epsilon_1 = \epsilon_2 = 1 \; , \quad \mu > 0 \; .$$

Consider $\underline{V = v_1}$ first, or "let's ride the slower soliton". In the comoving frame the exponents θ_i read

$$\boxed{\begin{aligned} \theta_1 &= \gamma_1(x - v_1 t - \bar{x}_1) = \gamma_1(X_{v_1} - \bar{x}_1) \\ \theta_2 &= \gamma_2(x - v_2 t - \bar{x}_2) = \gamma_2(X_{v_1} - (v_2 - v_1)t - \bar{x}_2) \end{aligned}} \tag{6.37}$$

so θ_1 stays finite, whereas $\theta_2 \to \mp\infty$ as $t \to \pm\infty$ with X_{v_1} fixed (I used that $v_2 > v_1$).

One of the two limits is easier to analyse, so let's start with that:

1. $\underline{t \to +\infty}$:

   In this limit $\theta_2 \to -\infty$, so $e^{\theta_2} \to 0$ and

$$\begin{aligned} \tan\frac{u}{4} &= \mu\frac{e^{\theta_1} - e^{\theta_2}}{1 + e^{\theta_1+\theta_2}} \\ &\to \mu e^{\theta_1} \\ &= \mu e^{\gamma_1(X_{v_1} - \bar{x}_1)} \\ &= e^{\gamma_1\left(x - v_1 t - \bar{x}_1 + \frac{1}{\gamma_1}\log\mu\right)} \; , \end{aligned}$$

[4] According to the signs of the limits of θ_1 and θ_2, the limit of $\tan(u/4)$ is as follows:

$$\begin{aligned} ++ :& \qquad \tan(u/4) \to 0 \\ +- :& \qquad \tan(u/4) \to +\infty \\ -+ :& \qquad \tan(u/4) \to -\infty \\ -- :& \qquad \tan(u/4) \to 0 \; . \end{aligned}$$

where in the last line we have expressed the finite limit in the comoving coordinates in terms of the original (x, t) coordinates.

This is a kink, the centre of which moves with velocity v_1 along the trajectory

$$\boxed{x = v_1 t + \bar{x}_1 - \frac{1}{\gamma_1} \log \frac{a_2 + a_1}{a_2 - a_1}} \,. \tag{6.38}$$

The last term is negative and represents a backward shift in space of the slower soliton compared to where it would have been at the same time in the absence of the faster soliton. (Equivalently, we can view this as a time delay for reaching a fixed value of x.)

2. $\underline{t \to -\infty}$:

   In this limit $\theta_2 \to +\infty$, so $e^{\theta_2} \to +\infty$ and

$$\begin{aligned} \tan \frac{u}{4} &= \mu \frac{e^{\theta_1} - e^{\theta_2}}{1 + e^{\theta_1 + \theta_2}} \\ &\to -\mu e^{-\theta_1} \,. \end{aligned}$$

   Recalling that $\tan\left(A \pm \frac{\pi}{2}\right) = -\frac{1}{\tan A}$, this means that

$$\begin{aligned} \tan\left(\frac{u}{4} \pm \frac{\pi}{2}\right) &\to \mu^{-1} e^{\theta_1} \\ &= e^{\gamma_1\left(x - v_1 t - \bar{x}_1 - \frac{1}{\gamma_1}\log\mu\right)} \,. \end{aligned}$$

   Therefore

$$u\Big|_{t\to-\infty,\ X_{v_1}\ \text{finite}} \approx \pm 2\pi + 4 \arctan e^{\gamma_1\left(x - v_1 t - \bar{x}_1 - \frac{1}{\gamma_1}\log\mu\right)} \,.$$

   (The $\pm$ sign ambiguity can be fixed by continuity. It turns out that -2π is correct.)

   This is a kink, the centre of which moves with velocity v_1 along the trajectory

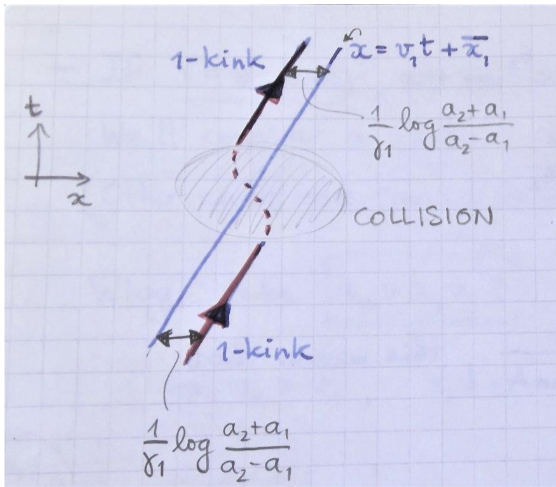
$$\boxed{x = v_1 t + \bar{x}_1 + \frac{1}{\gamma_1} \log \frac{a_2 + a_1}{a_2 - a_1}} \,. \tag{6.39}$$

   The last term is positive and represents a forward shift of the slower soliton compared to where it would have been at the same time in the absence of the faster soliton. (Equivalently, we can view this as a time advancement.)

Comparing the trajectories at early times ($t \to -\infty$) and at late times ($t \to +\infty$), we see that the collision with the faster soliton shifts the slower soliton backwards by

$$\frac{2}{\gamma_1} \log \frac{a_2 + a_1}{a_2 - a_1} \,,$$

as exemplified by this figure:



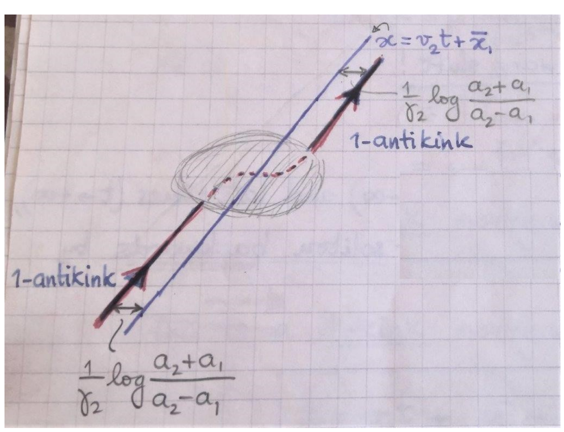
We say that the slower soliton has a negative phase shift:

$$\boxed{\text{PHASE SHIFT}_{\text{slower}} = -\frac{2}{\gamma_1} \log \frac{a_2 + a_1}{a_2 - a_1}} \tag{6.40}$$

We conclude that the slower kink emerges from the collision with the same shape and velocity, but delayed by a finite phase shift.

Now consider $\underline{V = v_2}$, or "let's ride the faster soliton". The calculation is similar to what we did above, so I'll let you work out the details in **[Ex 30]**. If you do this exercise you will find a surprise: even though $a_2 > 0$, so that acting on the vacuum with the a_2-Bäcklund transformation produces a kink, the component of the two-soliton solution (6.33) that moves at velocity v_2 is actually an anti-kink! So, even though the Bäcklund transformation always adds a soliton, the nature of the added soliton depends on what is already there.

The shifts have opposite signs to before, as exemplified by this figure:



This results in a positive phase shift:

$$\boxed{\text{PHASE SHIFT}_{\text{faster}} = +\frac{2}{\gamma_2}\log\frac{a_2+a_1}{a_2-a_1}}\,. \tag{6.41}$$

Summarising, we have the following picture for the collision of the anti-kink and the kink:

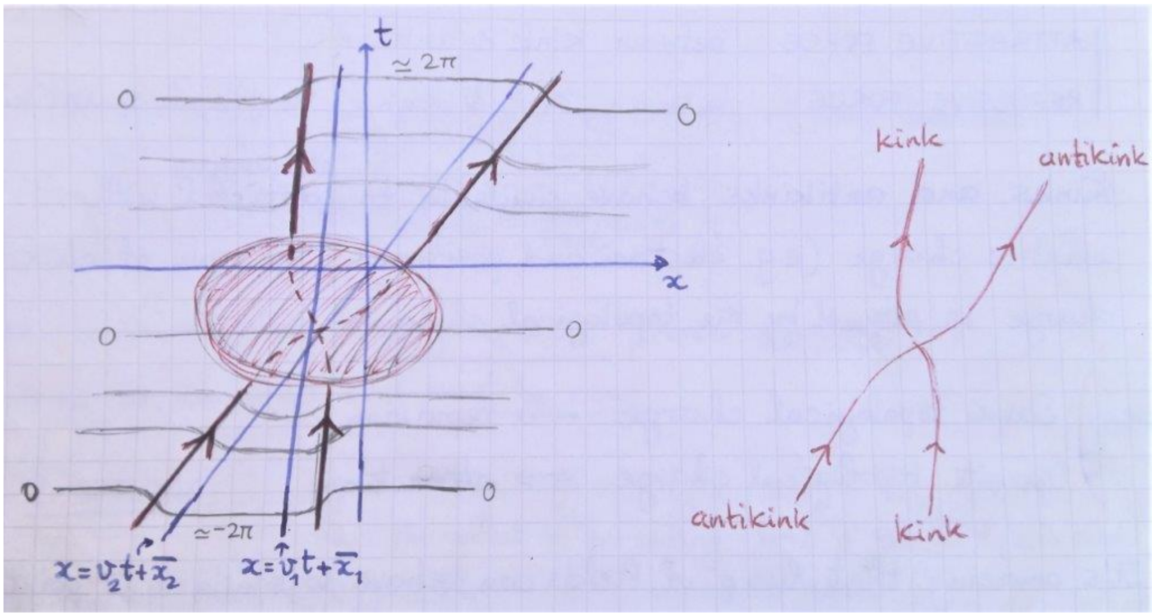


Figure 6.1: Schematic summary of the kink-antikink solution.

See also here for the plot of the kink-antikink solution with parameters $a_1 = 1.1$ and $a_2 = 2$, here for a contour plot of its energy density, which clearly shows the trajectories of the kink and the anti-kink, and here for an animation of the time evolution.

REMARK:
From the plot of the exact solution or the contour plot of its energy density we see that the kink and the anti-kink attract each other. Indeed we observe that they get closer during the interaction.

The remaining cases for the signs of a_1 and a_2 can be analysed similarly, see **[Ex 31]** and **[Ex 32]**. In particular, the 2-soliton solution that contains two kinks is depicted in figure 6.7.[5] (See also here for a plot of the kink-kink solution with parameters $a_1 = 0.6$ and $a_2 = -1.5$, here for a contour plot of its energy density, which clearly shows the trajectories of the two kinks, and here for an animation of the time evolution.)

[5]The solution that contains two anti-kinks can be obtained by sending $u \mapsto -u$.

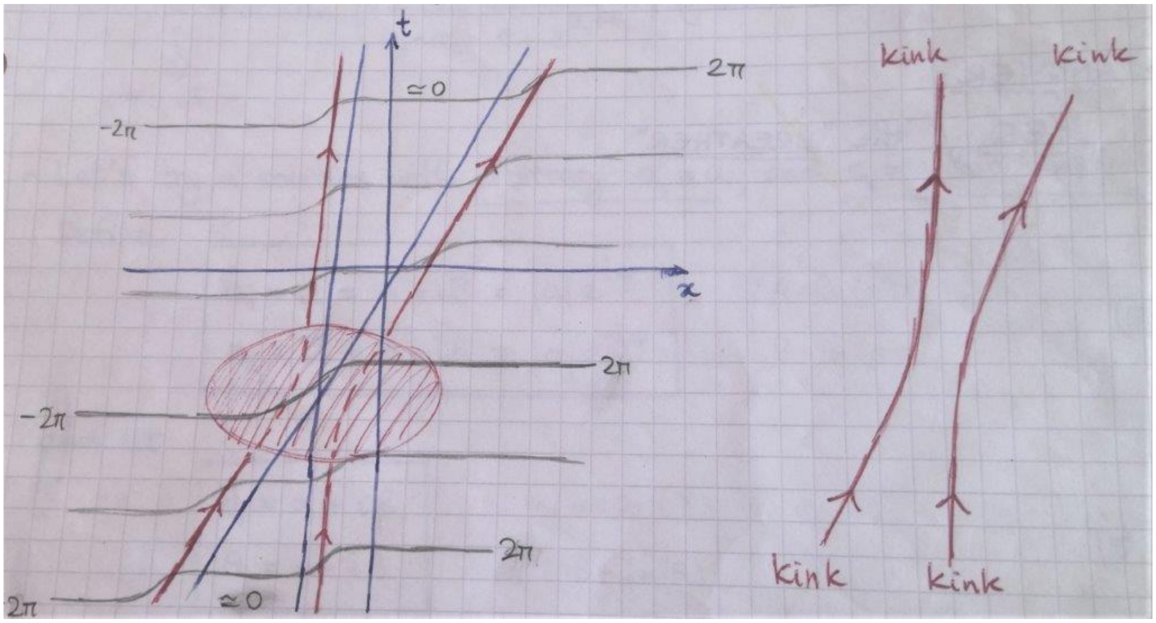


Figure 6.2: Schematic summary of the kink-kink solution.

From the plot of the exact solution or the contour plot of its energy density we see that the two kinks repel each other. Indeed they get further apart during the interaction. Curiously, they also seem to swap their identities!

INTERPRETATION:

ATTRACTIVE FORCE	between	kink and anti-kink
REPULSIVE FORCE	between	kink and kink
REPULSIVE FORCE	between	anti-kink and anti-kink

So kinks and anti-kinks behave similarly to elementary particles with electric charge, such as the electron and the positron. The role of electric charge is played here by the topological charge:

> Solitons with like topological charges repel
>
> Solitons with opposite topological charges attract.

It is quite amazing that lump of fields can behave so similarly to pointlike elementary particles. In the 1950's and 1960's, Tony Skyrme used versions of kinks (and anti-kinks) in four spacetime dimensions to model the behaviour of protons and neutrons in atomic nuclei. This is a very far-reaching idea, which unfortunately we don't have time to investigate further in this module.

We have seen that kinks and anti-kinks attract each other. This raises a natural question: can they stick together, or in physics parlance "form a bound state"? The answer is yes. The resulting bound state of a kink and an anti-kink is the "breather", which we now turn to.

6.8 The breather

Recall the general 2-soliton solution (6.33) of the sine-Gordon equation, that we rewrite here for convenience:

$$\boxed{u = 4\arctan\left(\frac{a_2 + a_1}{a_2 - a_1}\frac{e^{\theta_1} - e^{\theta_2}}{1 + e^{\theta_1+\theta_2}}\right)}\ .$$

This is a solution of the sine-Gordon equation for any values of the Bäcklund parameters a_1 and a_2 (and integration constants c_1 and c_2), even complex values. However, the sine-Gordon field u is an angle and so it must be real. There are essentially two options to achieve this:[6]

1. $\underline{a_1, a_2 \text{ (and } c_1, c_2) \in \mathbb{R}}$: this is what we have considered so far;

2. $\underline{a_2 = a_1^* \text{ (and } c_2 = c_1^*)}$: this is what we will consider next. But let's first check that the corresponding u is real:

$$\begin{aligned} u^* &= \left[4\arctan\left(\frac{a_2 + a_1}{a_2 - a_1}\frac{e^{\theta_1} - e^{\theta_2}}{1 + e^{\theta_1+\theta_2}}\right)\right]^* \\ &= 4\arctan\left(\frac{a_2^* + a_1^*}{a_2^* - a_1^*}\frac{e^{\theta_1^*} - e^{\theta_2^*}}{1 + e^{\theta_1^*+\theta_2^*}}\right) \\ &= 4\arctan\left(\frac{a_1 + a_2}{a_1 - a_2}\frac{e^{\theta_2} - e^{\theta_1}}{1 + e^{\theta_2+\theta_1}}\right) \\ &= 4\arctan\left(\frac{a_2 + a_1}{a_2 - a_1}\frac{e^{\theta_1} - e^{\theta_2}}{1 + e^{\theta_1+\theta_2}}\right) = u\ . \end{aligned}$$

   To get to the second line we used the fact that $\arctan(z)$ and e^z are complex analytic functions, therefore $[\arctan(z)]^* = \arctan(z^*)$ and $[e^z]^* = e^{z^*}$. To get to the third line we used $\theta_2 = \theta_1^*$, which follows from $a_2 = a_1^*$ and $c_2 = c_1^*$.

Let us then consider option 2 and try a solution with arbitrary $a_1 = a_2^* \equiv a$ and with $c_1 =$

[6] To be precise, one can also add to the integration constants c_1 and c_2 an integer multiple of πi. This has the effect of permuting the two solitons if the multiple is odd, and has no effect if the multiple is even.

$c_2 = 0$ for simplicity. Define

$$\boxed{\begin{aligned} a_1 &= a = A + iB = |a|e^{i\varphi} \\ a_2 &= \bar{a} = A - iB = |a|e^{-i\varphi} \end{aligned}} \tag{6.42}$$

where $A = \operatorname{Re}(a)$, $B = \operatorname{Im}(a)$, $\varphi = \arg(a)$, and let

$$\boxed{\begin{aligned} \theta_1 &= \alpha + i\beta \\ \theta_2 &= \alpha - i\beta \end{aligned}}\,, \tag{6.43}$$

with α and β real functions of x, t to be determined below. Then

$$\begin{aligned} \tan\frac{u}{4} &= \frac{|a|(e^{-i\varphi} + e^{i\varphi})}{|a|(e^{-i\varphi} - e^{i\varphi})} \cdot \frac{e^{\alpha+i\beta} - e^{\alpha-i\beta}}{1 + e^{2\alpha}} \\ &= \frac{2\cos\varphi}{-2i\sin\varphi} \cdot \frac{2i\sin\beta}{2\cosh\alpha} \end{aligned}$$

which simplifies to

$$\boxed{\tan\frac{u}{4} = -\frac{\cos\varphi}{\sin\varphi}\frac{\sin\beta}{\cosh\alpha}}\,. \tag{6.44}$$

To finish the calculation, let's determine the functions α, β in terms of the coordinates x, t and the parameters $|a|$ and φ:

$$\begin{aligned} \alpha + i\beta = \theta_1 &= \frac{1}{a}x^+ - ax^- \\ &= \frac{\bar{a}}{|a|^2}x^+ - ax^- = \frac{A - iB}{|a|^2}x^+ - (A + iB)x^-\,. \end{aligned} \tag{6.45}$$

Therefore

$$\begin{aligned} \alpha = \operatorname{Re}(\theta_1) &= \frac{A}{|a|^2}x^+ - Ax^- \\ &= \frac{A}{|a|}\left(\frac{1}{|a|}x^+ - |a|x^-\right)\,. \end{aligned}$$

We can now do similar manipulations to those after equation (6.15) to find

$$\boxed{\alpha = \frac{A}{|a|}\gamma(x - vt) \underset{(6.42)}{=} \cos\varphi \cdot \gamma(x - vt)}\,, \tag{6.46}$$

where

$$\boxed{\begin{aligned} v &= \frac{|a|^2 - 1}{|a|^2 + 1} \\ \gamma &= \frac{1}{\sqrt{1 - v^2}} = \frac{1 + |a|^2}{2|a|} \end{aligned}}\,. \tag{6.47}$$

*** EXERCISE**: Show that similarly **[Ex 33]**

$$\boxed{\beta = \frac{B}{|a|}\gamma(vx-t) \underset{(6.42)}{=} \sin\varphi \cdot \gamma(vx-t)}\ . \tag{6.48}$$

Substituting these expressions in (6.44) we find the **breather** solution

$$\boxed{\tan\frac{u}{4} = -\cot\varphi \cdot \frac{\sin(\sin\varphi \cdot \gamma(vx-t))}{\cosh(\cos\varphi \cdot \gamma(x-vt))}}\ . \tag{6.49}$$

REMARK:

- The ratio of the prefactor and the denominator in the RHS,

$$\frac{-\cot\varphi}{\cosh(\cos\varphi \cdot \gamma(x-vt))}\ ,$$

  defines an envelope which moves at the group velocity v. Recall that $|v| < 1$, where 1 is the speed of light, so this is consistent with the laws of special relativity.

- The numerator

$$\sin(\sin\varphi \cdot \gamma(x-vt))$$

  defines a carrier wave which moves at the phase velocity $1/v$.

To see why the solution (6.49) is called a breather, let us set $|a| = 1$, or equivalently $v = 0$. (This can be achieved by switching to a comoving frame if $v \neq 0$.) Then the breather simplifies to

$$\boxed{\tan\frac{u}{4} = \cot\varphi \cdot \frac{\sin(\sin\varphi \cdot t)}{\cosh(\cos\varphi \cdot x)}} \tag{6.50}$$

and the field looks like a bouncing (or "breathing") bound state of a kink and an anti-kink, with time period

$$\boxed{\tau = \frac{2\pi}{|\sin\varphi|}}\ . \tag{6.51}$$

See figure (6.3) for a summary of the $v = 0$ breather solution, this for a plot of the breather solution with $v = 0$ and $\varphi = \pi/10$, this for a contour plot of its energy density, which clearly shows the trajectories of the breathing pair of kink and anti-kink, and this for an animation of the time evolution.

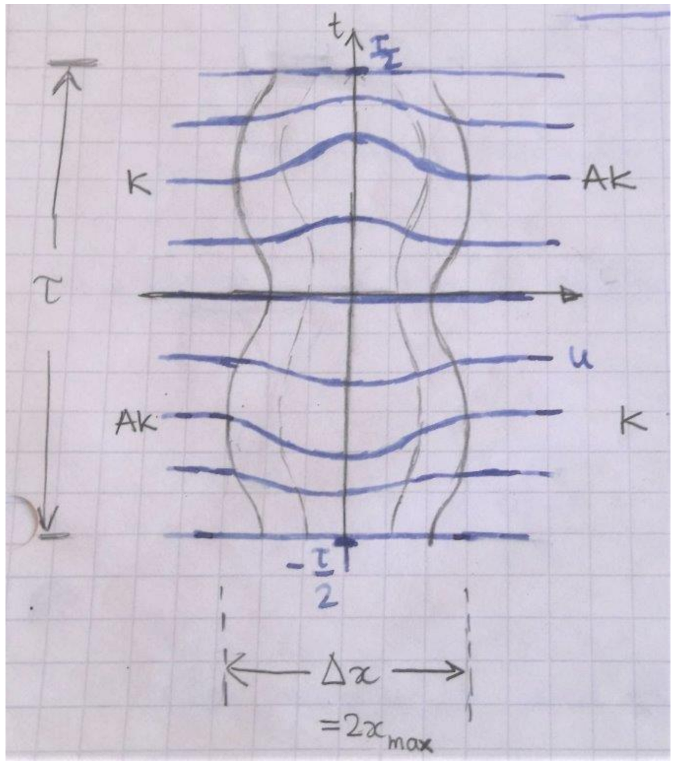


Figure 6.3: Summary of the $v = 0$ breather solution.

One can show[7] that the $v = 0$ breather has energy $E_{\text{breather}} = 16 \cos\varphi$. Since a static kink and a static anti-kink have energy $E_{\text{kink}} = E_{\text{antikink}} = 8$, the binding energy of the kink and the anti-kink in the breather is

$$E_{\text{binding}} = E_{\text{breather}} - E_{\text{kink}} - E_{\text{antikink}} = -16(1 - \cos\varphi) \ .$$

This is negative as expected: the binding lowers the energy of the solution.

As $\varphi \to 0$, the binding energy tends to zero. It is immediate to see from equation (6.51) that the time period of the bounce diverges: $\tau \sim 1/|\varphi| \to \infty$. The spatial size of the breather also diverges like **[Ex 34]**

$$\boxed{x_{\text{max}} \sim -\log|\varphi|} \to \infty \ .$$

In this limit the kink and the antikink become more and more loosely bound. The resulting solution

$$u = 4 \arctan\left(t \cdot \text{sech}(x)\right)$$

describes a kink and an anti-kink starting infinitely far away from one another and doing half an oscillation. Since $\text{sech}(x) \approx 2e^{-|x|}$ as $|x| \to \infty$, the kink and the anti-kink do not follow linear trajectories as $t \to \pm\infty$. Rather, the asymptotic trajectories of the kink and the anti-kink are given by $|x| \sim \log|t|$.

[7] This is a good but technical exercise, which is not in the problem sheet.

Chapter 7

The Hirota method

The main reference for this chapter is §5.3 of [Drazin and Johnson, 1989].

This is an alternative to the Bäcklund transformationas a way to generate multi-soliton solutions, which is sometimes available when the Bäcklund transformation is not. It was devised by **Hirota** [Hirota, 1971] to write N-soliton solutions of the KdV equation, and was then generalised to a large class of equations. We will focus on the KdV equation in this chapter.

7.1 Motivations

7.1.1 Series solutions

Let us substitute

$$\boxed{u = w_x} \tag{7.1}$$

in the KdV equation

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

We find the equation

$$w_{xt} + 6w_x w_{xx} + w_{xxxx} = 0 \ ,$$

which we can integrate with respect to x:

$$w_t + 3w_x^2 + w_{xxx} = g(t) \ .$$

We will drop the integration "constant" (with respect to x) $g(t)$ in what follows, since it can be absorbed in a redefinition of w that does not change $u = w_x$:

$$w_{\text{old}}(x,t) = w_{\text{new}}(x,t) + \int_{t_0}^{t} dt' \ g(t') \ .$$

Using the new w (and dropping the subscript "new"), we have the following equation:

$$\boxed{w_t + 3w_x^2 + w_{xxx} = 0} \, . \tag{7.2}$$

For w small, the w_x^2 term is negligible and the equation is linear – and hence, easier to solve. To be more systematic, we can look for a series solution

$$w = \epsilon w_1 + \epsilon^2 w_2 + \ldots \, .$$

Substituting in and solving order by order in ϵ:

$$\begin{array}{lll} \epsilon^1 : & w_{1t} + w_{1xxx} = 0 & \text{the linear equation} \\ \epsilon^2 : & w_{2t} + 3{w_{1x}}^2 + w_{2xxx} = 0 & \text{the first ‘correction’} \end{array}$$

and so on. In principle we can solve these equations in turn, rather as we did for the Gardner transform.

Bad news: We'd need to continue infinitely far to find an *exact* formula for w.

Good news: The method would be saved if it happened that $w_m = 0$ for all $m > n$ for some n. Then the *approximate* solution up to order n would turn out the be *exact*.

Bad news: This phenomenon does not happen for the simple scheme just described. Something more subtle will be needed, which is exactly what Hirota discovered.

7.1.2 Some hints

A close relative of KdV is **Burger's equation**:

$$u_t + uu_x - \lambda u_{xx} = 0 \, ,$$

where λ is a parameter. Substituting $u = -2\lambda v_x/v = -2\lambda \frac{\partial}{\partial x}(\log v)$ (exercise!) turns this into the linear **heat equation**

$$v_t = \lambda v_{xx} \, .$$

Further evidence that logarithmic derivatives might have a role to play comes if we recall the one-soliton solution of KdV:

$$u = 2\mu^2 \operatorname{sech}^2 \left[\mu(x - x_0 - 4\mu^2 t)\right]$$

with

$$\mu = \frac{\sqrt{v}}{2} \, .$$

This one-soliton solution can be written as $u = w_x$ with

$$w = 2\mu \, \tanh \left[\mu(x - x_0 - 4\mu^2 t)\right] \, .$$

We can integrate the right-hand side once more, using $\tanh y = \frac{d}{dy}\log\cosh y$ to find

$$u = 2\frac{\partial^2}{\partial x^2}\log\cosh\left[\mu(x - x_0 - 4\mu^2 t)\right] .$$

This can be simplified further. Letting $X = x - x_0 - 4\mu^2 t$,

$$\begin{aligned} u &= 2\frac{\partial^2}{\partial x^2}\log\frac{e^{-\mu X}(1+e^{2\mu X})}{2} \\ &= 2\frac{\partial^2}{\partial x^2}\left[-\mu X - \log 2 + \log\left(1+e^{2\mu X}\right)\right] \\ &= 2\frac{\partial^2}{\partial X^2}\log\left(1+e^{2\mu X}\right) . \end{aligned}$$

In terms of the original variables,

$$\boxed{u(x,t) = 2\frac{\partial^2}{\partial x^2}\log\left(1+e^{2\mu(x-x_0-4\mu^2 t)}\right)} .$$

This is the form of the one-soliton solution of KdV that we will refer to in the following.

7.2 KdV equation in bilinear form

7.2.1 The quadratic form of the KdV equation

Inspired by the rewritten form of the one-soliton solution, let's substitute

$$\boxed{w = 2\frac{\partial}{\partial x}\log f = \frac{f_x}{f}} \iff \boxed{u = 2\frac{\partial^2}{\partial x^2}\log f} \tag{7.3}$$

in equation (7.2).[1] Then

$$\begin{aligned} \frac{1}{2}w_t &= \frac{f_{xt}f - f_x f_t}{f^2} , \\ \frac{1}{2}w_x &= \frac{f_{xx}f - f_x^2}{f^2} , \\ \frac{1}{2}w_{xx} &= \dots \qquad \textbf{[Ex 35]} \\ \frac{1}{2}w_{xxx} &= \frac{f_{xxxx}}{f} - 4\frac{f_{xxx}f_x}{f^2} - 3\frac{f_{xx}^2}{f^2} + 12\frac{f_{xx}f_x^2}{f^3} - 6\frac{f_x^4}{f^4} , \end{aligned} \tag{7.4}$$

and equation (7.2) for w becomes **[Ex 35]**

$$\frac{f_{xt}}{f} - \frac{f_x f_t}{f^2} + 3\frac{f_{xx}^2}{f^2} - 4\frac{f_{xxx}f_x}{f^2} + \frac{f_{xxxx}}{f} = 0$$

[1] In the literature on integrable systems, the function f is now called the τ-function.

for f.

Multiplying by f^2, we find the so called **quadratic form of the KdV equation**:

$$\boxed{f f_{xt} - f_x f_t + 3f_{xx}^2 - 4f_x f_{xxx} + f f_{xxxx} = 0}\ . \tag{7.5}$$

Some cancellations have taken place to get to the quadratic form (7.5) of the KdV equation, but at first sight this might not seem progress on the initial equation (7.2). But (7.5) is quadratic in f and it can be rewritten in a neat way. A hint for that is that

$$\frac{\partial}{\partial x}\frac{\partial}{\partial t}\left(\frac{1}{2}f^2\right) = \frac{\partial}{\partial x}(f f_t) = f f_{xt} + f_x f_t\ .$$

This is almost like the first two terms in (7.5), except for the relative sign. We will fix this sign problem shortly.

7.2.2 Hirota's bilinear operator

Hirota defined a bilinear differential operator D which maps a pair of functions (f, g) into a single function $D(f \cdot g)$. If we work on C^∞ functions, then

$$\begin{aligned} D\ :\qquad & C^\infty \times C^\infty \to C^\infty \\ & (f, g) \mapsto D(f \cdot g)\ , \end{aligned}$$

and bilinearity means that

$$\begin{aligned} D(a_1 f_1 + a_2 f_2 \cdot g) &= a_1 D(f_1 \cdot g) + a_2 D(f_2 \cdot g) \\ D(f \cdot b_1 g_1 + b_2 g_2) &= b_1 D(f \cdot g_1) + b_2 D(f \cdot g_2) \end{aligned}$$

for any constants a_1, a_2, b_1, b_2.

REMARK:
This is unlike the usual linear differential operators that you are familiar with, such as $\left(\frac{\partial}{\partial x}\right)^n$, which maps a single function f to a single function $\frac{\partial^n f}{\partial x^n}$.

For any integers $m, n \geqslant 0$, we define **Hirota's bilinear differential operator** $D_t^m D_x^n$ by

$$\boxed{[D_t^m D_x^n(f \cdot g)](x,t) := \left(\frac{\partial}{\partial t} - \frac{\partial}{\partial t'}\right)^m \left(\frac{\partial}{\partial x} - \frac{\partial}{\partial x'}\right)^n f(x,t) g(x',t')\bigg|_{\substack{x'=x \\ t'=t}}}\ . \tag{7.6}$$

Let us look at a few examples. We start with

$$\begin{aligned}
[D_t(f\cdot g)](x,t) &= \left(\frac{\partial}{\partial t}-\frac{\partial}{\partial t'}\right) f(x,t)g(x',t')\Big|_{\substack{x'=x\\t'=t}} \\
&= f_x(x,t)g(x',t') - f(x,t)g_{t'}(x',t')\Big|_{\substack{x'=x\\t'=t}} \\
&= f_t(x,t)g(x,t) - f(x,t)g_t(x,t)\ ,
\end{aligned} \tag{7.7}$$

so

$$D_t(f\cdot g) = f_t g - f g_t \qquad \text{and} \qquad D_t(f,f) = 0\ ,$$

and similarly for D_x. Next we look at

$$\begin{aligned}
[D_tD_x(f\cdot g)](x,t) &= \left(\frac{\partial}{\partial t}-\frac{\partial}{\partial t'}\right)\left(\frac{\partial}{\partial x}-\frac{\partial}{\partial x'}\right) f(x,t)g(x',t')\Big|_{\substack{x'=x\\t'=t}} \\
&= \left(\frac{\partial}{\partial t}-\frac{\partial}{\partial t'}\right)\left(f_x(x,t)g(x',t') - f(x,t)g_{x'}(x',t')\right)\Big|_{\substack{x'=x\\t'=t}} \\
&= f_{xt}(x,t)g(x,t) - f_t(x,t)g_x(x,t) - f_x(x,t)g_t(x,t) + f(x,t)g_{xt}(x,t)\ ,
\end{aligned}$$

so

$$D_tD_x(f\cdot g) = f_{xt}g - f_tg_x - f_xg_t + fg_{xt} \qquad \text{and} \qquad D_tD_x(f\cdot f) = 2(ff_{tx} - f_tf_x)\ . \tag{7.8}$$

This is promising, because the right-hand-side of the last expression reproduces the first two terms in the quadratic form of the KdV equation (7.5), up to an overall factor of 2. Let's proceed and compute

$$D_x^2(f\cdot g) = f_{xx}g - 2f_xg_x + fg_{xx}\ , \tag{7.9}$$

which implies

$$D_x^2(f\cdot f) = 2(ff_{xx} - f_x^2)\ .$$

REMARK:
Note that $D_x^2(f\cdot f) \neq 0$ even though $D_x(f\cdot f) = 0$. This is not inconsistent, because $D_x^2(f\cdot f) \neq D_x\left(D_x(f\cdot f)\right)$. In fact, the right-hand side of this last expression is meaningless, since the outer D_x must act on a pair of functions, but $D_x(f\cdot f)$ is a single function.

Finally, we can calculate

$$\begin{aligned}
D_x^4(f\cdot g) &= \dots \qquad\qquad \textbf{[Ex 36]} \\
&= f_{xxxx}g - 4f_{xxx}g_x + 6f_{xx}g_{xx} - 4f_xg_{xxx} + fg_{xxxx}\ .
\end{aligned}$$

Note that the result is like $\partial_x^4(fg)$, but with alternating signs! So

$$D_x^4(f\cdot f) = 2(ff_{xxxx} - 4f_xf_{xxx} + 3f_{xx}^2)\ . \tag{7.10}$$

Here is the miracle: the KdV equation in its quadratic form (7.5) can be recast as

$$\boxed{(D_t D_x + D_x^4)(f \cdot f) = 0} \tag{7.11}$$

where the bilinear operator $D_t D_x + D_x^4$ is defined by linearity on the space of operators of the type (7.6), namely $(D_t D_x + D_x^4)(f \cdot g) = D_t D_x(f \cdot g) + D_x^4(f \cdot g)$. Equation (7.11) is the so called **bilinear form of the KdV equation**.

REMARK:
Observe that we can formally factor the Hirota operator as

$$D_t D_x + D_x^4 = (D_t + D_x^3) D_x \ ,$$

which is a short-hand for

$$(D_t D_x + D_x^4)(f, g) = (\partial_t - \partial_{t'} + (\partial_x - \partial_{x'})^3)(\partial_x - \partial_{x'}) f(x,t) g(x',t') \Big|_{\substack{x'=x \\ t'=t}} .$$

This is not an accident. It is related to the fact that the differential operator $\partial_t + \partial_x^3$ appears in the linearised KdV equation for u, and therefore the differential operator $(\partial_t + \partial_x^3)\partial_x$ appears in the linearisation of the equation for w (before integration with respect to x).

7.3 Solutions

We will need two ideas to find multi-soliton solutions. The first is inspired by a rather basic observation: if we take $f = 1$, then the KdV field is the vacuum $u = 0$; if instead we take

$$f = 1 + e^{2\mu(x - x_0 - 4\mu^2 t)} \ ,$$

then the u is the one-soliton (travelling wave) solution of KdV. Since (7.11) is a bilinear equation, this suggests that multi-soliton solutions might be obtained from an f which is a sum of exponentials of linear functions of x and t, with $1 = e^0$ as the trivial case. But before we get to the general case, let us check the Hirota formalism by rederiving this one-soliton solution.

7.3.1 Example: 1-soliton

Let's try

$$\boxed{f = 1 + e^{\theta}} \tag{7.12}$$

with

$$\theta = ax + bt + c \ ,$$

where a, b, c are constants.

Lemma 1. *If* $\theta_i = a_i x + b_i t + c_i$ $(i = 1, 2)$*, then* **[Ex 38]**

$$\boxed{D_t^m D_x^n (e^{\theta_1} \cdot e^{\theta_2}) = (b_1 - b_2)^m (a_1 - a_2)^n e^{\theta_1 + \theta_2}}\ . \tag{7.13}$$

In particular

$$\begin{aligned} &D_t^m D_x^n (e^{\theta} \cdot e^{\theta}) = 0 \qquad\qquad \text{(unless } m = n = 0\text{)} \\ &D_t^m D_x^n (e^{\theta} \cdot 1) = (-1)^{m+n} D_t^m D_x^n (1 \cdot e^{\theta}) = b^m a^n e^{\theta}\ . \end{aligned} \tag{7.14}$$

Therefore the bilinear form of the KdV equation for $f = 1 + e^{\theta}$ is

$$\begin{aligned} 0 &= (D_t D_x + D_x^4)(1 + e^{\theta} \cdot 1 + e^{\theta}) \\ &\underset{\text{bilinearity}}{=} (D_t D_x + D_x^4)\left[(1 \cdot 1) + (1 \cdot e^{\theta}) + (e^{\theta} \cdot 1) + (e^{\theta} \cdot e^{\theta})\right] \\ &\underset{(7.14)}{=} 2(D_t D_x + D_x^4)(e^{\theta} \cdot 1) \\ &\underset{(7.14)}{=} 2(ba + a^4)e^{\theta} = 2a(b + a^3)e^{\theta}\ . \end{aligned}$$

There are two ways to solve this algebraic equation:

1. $\underline{a = 0}$: then f is independent of x, and $u = 0$.

2. $\underline{b = -a^3}$: then
$$f = 1 + e^{ax - a^3 t + c},$$
and
$$\boxed{u = 2\frac{\partial^2}{\partial x^2} \log\left(1 + e^{ax - a^3 t + c}\right)}\ , \tag{7.15}$$
which is nothing but the one-soliton solution with velocity $v = a^2$, up to redefinitions of the constants.

7.3.2 The N-soliton solution (sketch)

The second idea is to look for a power series solution (or a so-called "perturbative expansion" in an *auxiliary parameter* ϵ,

$$\boxed{f(x,t) = \sum_{n=0}^{\infty} \epsilon^n f_n(x,t) \quad \text{with} \quad f_0 = 1}\ , \tag{7.16}$$

and hope that the series terminates at some value of n, so that we can take ϵ to be finite and eventually set it to 1.

We will write the bilinear form of KdV as

$$\boxed{B(f\cdot f)=0} \qquad \text{with} \qquad \boxed{B=D_tD_x+D_x^4}\,.$$

Substituting in (7.16), we find

$$\begin{aligned} 0 &= B\Big(\sum_{n_1=0}^{\infty}\epsilon^{n_1}f_{n_1}\cdot\sum_{n_2=0}^{\infty}\epsilon^{n_2}f_{n_2}\Big)\\ &= \sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty}\epsilon^{n_1+n_2}B(f_{n_1}\cdot f_{n_2}) \end{aligned}$$

where in the second line we used the bilinearity of the Hirota operator B. Gathering terms of the same degree $n=n_1+n_2$ in ϵ, we can rewrite this as

$$\boxed{0=\sum_{n=0}^{\infty}\epsilon^n\sum_{m=0}^{n}B(f_{n-m}\cdot f_m)\underset{B(1\cdot 1)=0}{=}\sum_{n=1}^{\infty}\epsilon^n\sum_{m=0}^{n}B(f_{n-m}\cdot f_m)}\,. \tag{7.17}$$

Let's solve this equation order by order in ϵ. We find that

$$\boxed{\sum_{m=0}^{n}B(f_{n-m}\cdot f_m)=0 \quad \forall\, n=1,2,\ldots} \tag{7.18}$$

with $f_0=1$. Writing (7.18) as

$$\boxed{B(f_n\cdot 1)+B(1\cdot f_n)=(\text{expression in } f_1,f_2,\ldots,f_{n-1})}\,, \tag{7.19}$$

makes it clear that we can solve (7.18) recursively to determine the Taylor coefficients of f. We will need another lemma:

Lemma 2. **[Ex 39]** *For any function f,*

$$\boxed{D_t^mD_x^n(f\cdot 1)=(-1)^{m+n}D_t^mD_x^n(1\cdot f)=\frac{\partial^m}{\partial t^m}\frac{\partial^n}{\partial x^n}f}\,.$$

Using this lemma, we can write the recursion relation (7.19) more explicitly as

$$\boxed{\frac{\partial}{\partial x}\left(\frac{\partial}{\partial t}+\frac{\partial^3}{\partial x^3}\right)f_n=-\frac{1}{2}\sum_{m=1}^{n-1}B(f_{n-m}\cdot f_m)}\,, \tag{7.20}$$

which is valid for all $n=1,2,\ldots$. In the following this recursion relation, which determines f_n in terms of all the f_m with $m<n$, will be referred to as A_n.

For $n = 1$ this reduces to

$$\frac{\partial}{\partial x}\left(\frac{\partial}{\partial t} + \frac{\partial^3}{\partial x^3}\right) f_1 = 0$$

or, with appropriate boundary conditions,

$$\boxed{\left(\frac{\partial}{\partial t} + \frac{\partial^3}{\partial x^3}\right) f_1 = 0}\ ,$$

which is a linear equation. A simple solution is

$$\boxed{f_1 = \sum_{i=1}^{N} e^{a_i x - a_i^3 t + c_i} \equiv \sum_{i=1}^{N} e^{\theta_i}}\ , \tag{7.21}$$

where a_i and c_i are, as usual, constants.

The higher f_n can then be determined recursively using A_n (7.20). The amazing fact is that with f_1 as in equation (7.21), the expansion (7.16) terminates at order N. All the higher equations $A_{n>N}$ are solved with $f_{n>N} = 0$! This is quite non-trivial: it requires that $f_1, \ldots, f_N$ satisfy the consistency conditions that the RHS of A_n vanish for $n = N+1, \ldots, 2N$.

The N-soliton solution of KdV is then given by

$$\boxed{f = 1 + f_1 + f_2 + \cdots + f_N}\ ,$$

where we set $\epsilon = 1$ (or absorbed it in the constants c_i).

<u>EXAMPLES</u>:

$\boxed{N = 1}$ In this case

$$f_1 = e^{a_1 x - a_1^3 t + c_1} \equiv e^{\theta_1}$$

and A_2 reads

$$\partial_x(\partial_t + \partial_x^3) f_2 = -\frac{1}{2} B(e^{\theta_1} \cdot e^{\theta_1}) \underset{(7.14)}{=} 0\ .$$

So we can take $f_2 = 0$ (and $f_3 = f_4 = \cdots = 0$ as well). Setting $\epsilon = 1$, or absorbing ϵ in c_1, we get

$$f = 1 + e^{\theta_1}\ ,$$

the one-soliton solution as we found in (7.15).

$\boxed{N = 2}$ In this case

$$f_1 = e^{\theta_1} + e^{\theta_2}$$

and A_2 reads

$$\begin{aligned}
\partial_x(\partial_t + \partial_x^3) f_2 &= -\frac{1}{2} B(e^{\theta_1} + e^{\theta_2} \cdot e^{\theta_1} + e^{\theta_2}) \\
&\underset{\substack{\text{bilinearity}\\ +(7.13)}}{=} -B(e^{\theta_1} \cdot e^{\theta_2}) \\
&\underset{\substack{B = D_t D_x + D_x^4 \\ +(7.13)}}{=} -(a_1 - a_2)[-a_1^3 + a_1^3 + (a_1 - a_2)^3] e^{\theta_1 + \theta_2} \\
&= 3 a_1 a_2 (a_1 - a_2)^2 e^{\theta_1 + \theta_2} \, .
\end{aligned}$$

So let's try

$$f_2 = A e^{\theta_1 + \theta_2}$$

for some constant A to be determined. Substituting in the previous equation we find

$$\begin{aligned}
(a_1 + a_2)[-a_1^3 - a_2^3 + (a_1 + a_2)^3] A e^{\theta_1 + \theta_2} &= 3 a_1 a_2 (a_1 - a_2)^2 e^{\theta_1 + \theta_2} \\
\Rightarrow \quad 3 a_1 a_2 (a_1 + a_2)^2 A &= 3 a_1 a_2 (a_1 - a_2)^2 \\
\Rightarrow \quad A &= \left(\frac{a_1 - a_2}{a_1 + a_2} \right)^2 \, .
\end{aligned}$$

So we get

$$\boxed{f = 1 + e^{\theta_1} + e^{\theta_2} + \left(\frac{a_1 - a_2}{a_1 + a_2} \right)^2 e^{\theta_1 + \theta_2}} \tag{7.22}$$

for the 2-soliton solution of KdV, where again we set $\epsilon = 1$ or absorbed it into shifts of c_1 and c_2.

∗ EXERCISE: Show that $B(f_1 \cdot f_2) = 0$ and $B(f_2 \cdot f_2) = 0$, so that one can consistently set $f_3 = f_4 = \cdots = 0$. **[Ex 40]**

General N — Let's first rewrite the 2-soliton solution (7.22) that we have just found:

$$\begin{aligned}
f &= (1 + e^{\theta_1})(1 + e^{\theta_2}) - e^{\theta_1 + \theta_2} + \left(\frac{a_1 - a_2}{a_1 + a_2} \right)^2 e^{\theta_1 + \theta_2} \\
&= (1 + e^{\theta_1})(1 + e^{\theta_2}) - \frac{4 a_1 a_2}{(a_1 + a_2)^2} e^{\theta_1 + \theta_2} \\
&= \begin{vmatrix} 1 + e^{\theta_1} & \frac{2a_1}{a_1 + a_2} e^{\theta_2} \\ \frac{2a_2}{a_1 + a_2} e^{\theta_1} & 1 + e^{\theta_2} \end{vmatrix} \, .
\end{aligned}$$

So we can write

$$\boxed{f = \det(S)} \, , \qquad \text{where} \qquad \boxed{S_{ij} = \delta_{ij} + \frac{2a_i}{a_i + a_j} e^{\theta_j}} \, ,$$

where here $i, j \in \{1, 2\}$.[2]

It turns out that this formula generalises to higher N, with S an $N \times N$ matrix of the same form but with $i, j \in \{1, \dots, N\}$, giving the N-soliton solution of KdV. This can be proved by induction. One can also show that

$$f_n = \sum_{1 \leqslant i_1 < i_2 < \cdots < i_n \leqslant N} e^{\theta_{i_1} + \theta_{i_2} + \cdots + \theta_{i_n}} \prod_{1 \leqslant j < k \leqslant n} \left(\frac{a_{i_j} - a_{i_k}}{a_{i_j} + a_{i_k}} \right)^2 .$$

7.4 Asymptotics of 2-soliton solutions and phase shifts

To see that the $N = 2$ solution (7.22) does indeed involve two solitons, we can follow the same logic as in section 6.7, where we studied the asymptotics of 2-soliton solutions of the sine-Gordon equation. Namely, we switch to an appropriate comoving frame and only then take $t \to \pm\infty$.

Recall that

$$f = 1 + e^{\theta_1} + e^{\theta_2} + A e^{\theta_1 + \theta_2}$$

where

$$\theta_i = a_i x - a_i^3 t + c_i \ , \qquad A = \left(\frac{a_1 - a_2}{a_1 + a_2} \right)^2 .$$

We can take $\underline{0 < a_1 < a_2}$ without loss of generality[3] so $v_1 = a_1^2 < v_2 = a_2^2$. Let's follow the slower soliton first:

$$\boxed{t \to \pm\infty \quad \text{with} \quad X_{a_1^2} = x - a_1^2 t \quad \text{fixed}} \ .$$

Then

$$\begin{aligned} \theta_1 &= a_1 X_{a_1^2} + c_1 \\ \theta_2 &= a_2 \left(X_{a_1^2} - (a_2^2 - a_1^2) t \right) + c_2 \ . \end{aligned}$$

Let us consider the two limits in turn.

1. $\underline{t \to +\infty}$: in this limit θ_1 stays fixed and $\theta_2 \to -\infty$, so

   $$f \to 1 + e^{\theta_1} \ .$$

   This describes a KdV soliton centred at

   $$\boxed{x_{\text{centre}}(t) = a_1^2 t - \frac{c_1}{a_1}} \ .$$

[2] Note that using e^{θ_i} instead of e^{θ_j} in the definition of the matrix element S_{ij} produces the same determinant.
[3] Convince yourself of this statement.

2. $\underline{t \to -\infty}$: in this limit θ_1 stays fixed and $\theta_2 \to +\infty$, so

$$f \to e^{\theta_2}(1 + Ae_1^{\theta}) \ .$$

The prefactor e^{θ_2} does not matter, because

$$\begin{aligned} u &= 2\frac{\partial^2}{\partial x^2} \log f \equiv 2\frac{\partial^2}{\partial x^2}\left[\theta_2 + \log(1 + Ae^{\theta_1})\right] \\ &= 2\frac{\partial^2}{\partial x^2} \log(1 + Ae^{\theta_1}) \\ &= 2\frac{\partial^2}{\partial x^2} \log\left(1 + e^{a_1 x - a_1^3 t + c_1 + \log A}\right) \ . \end{aligned}$$

where in the second line we used that θ_2 is linear in x, and in the third line we expressed the result in the original (x, t) coordinates. This describes a KdV soliton centred at

$$\boxed{x_{\text{centre}}(t) = a_1^2 t - \frac{c_1 + \log A}{a_1}} \ .$$

Therefore the slower soliton has a negative phase shift:

$$\boxed{\text{PHASE SHIFT}_{\text{slower}} = \frac{1}{a_1} \log A = -\frac{2}{a_1} \log\left|\frac{a_2 + a_1}{a_2 - a_1}\right| < 0} \ .$$

Next, let's follow the faster soliton:

$$\boxed{t \to \pm\infty \quad \text{with} \quad X_{a_2^2} = x - a_2^2 t \quad \text{fixed}} \ .$$

Then

$$\begin{aligned} \theta_1 &= a_1\left(X_{a_2^2} - (a_1^2 - a_2^2)t\right) + c_1 \\ \theta_2 &= a_2 X_{a_2^2} + c_2 \ . \end{aligned}$$

Again, we consider the two limits in turn.

1. $\underline{t \to -\infty}$: in this limit $\theta_1 \to -\infty$ and θ_2 stays fixed, so

$$f \to 1 + e^{\theta_2} \ .$$

This describes a KdV soliton centred at

$$\boxed{x_{\text{centre}}(t) = a_2^2 t - \frac{c_2}{a_2}} \ .$$

2. $\underline{t \to +\infty}$: in this limit $\theta_1 \to +\infty$ and θ_2 stays fixed, so

$$f \to e^{\theta_1}(1 + Ae^{\theta}_2)\ ,$$

which describes a KdV soliton centred at

$$\boxed{x_{\text{centre}}(t) = a_2^2 t - \frac{c_2 + \log A}{a_2}}\ .$$

Therefore the faster soliton has a positive phase shift:

$$\boxed{\text{PHASE SHIFT}_{\text{faster}} = -\frac{1}{a_2}\log A = \frac{2}{a_2}\log\left|\frac{a_2 + a_1}{a_2 - a_1}\right| > 0}\ .$$

Summarising, from the analysis of the asymptotics of the 2-soliton solution we obtain the picture in Fig. 7.1. We have therefore verified that KdV solitons satisfy the third defining

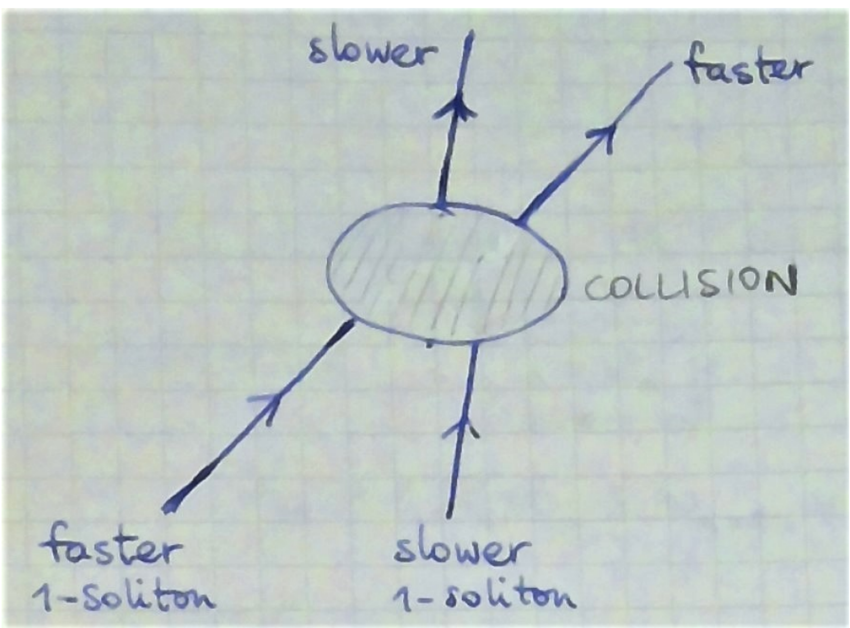


Figure 7.1: Schematic summary of the 2-soliton solution of KdV.

property of a soliton 3: when two KdV solitons collide, they emerge from the collision with the same shapes and velocities that they had before the collision. The effect of the interaction is in the phase shifts of the two solitons, which capture the advancement of the faster soliton and the delay of the slower soliton.

We can also look at the exact 2-soliton solution encoded in (7.3) and (7.22) to get a better feel for what happens during the collision. Here is a plot of the 2-soliton solution with parameters

$a_1 = 0.7$ and $a_2 = 1$. The contour plot below clearly shows the trajectories of the two KdV solitons and how they repel each other and swap identities when they get close, resulting in a phase shift. Finally, here is an animation of their time evolution.

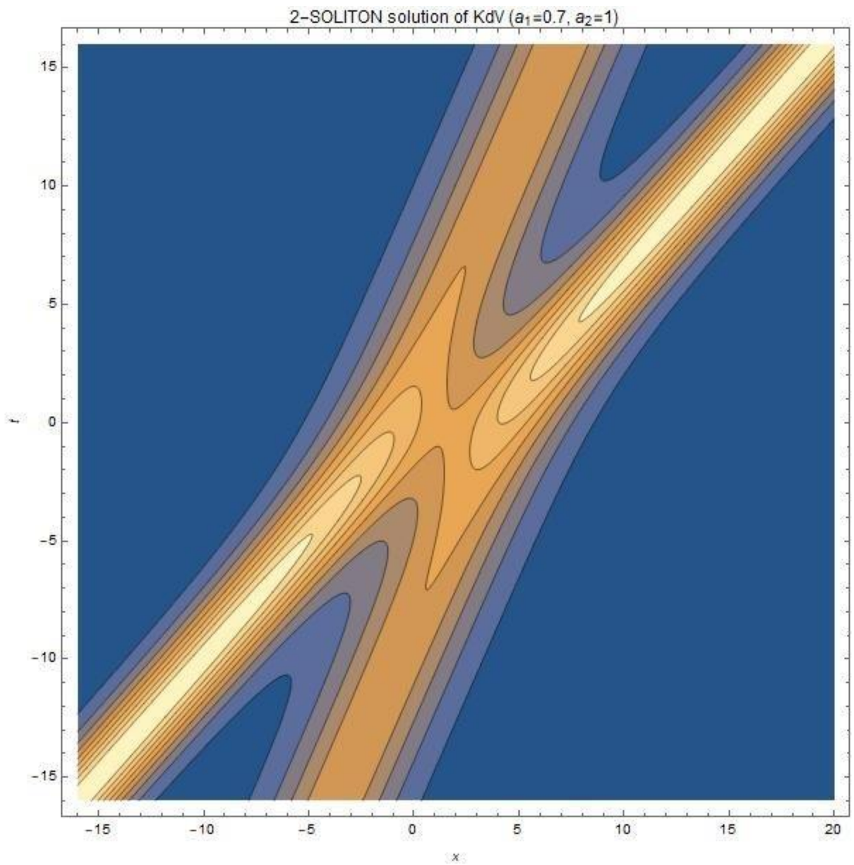


Figure 7.2: A two-soliton solution of the KdV equation.

Chapter 8

Overview of the inverse scattering method

For this chapter, see section 4.2 of [Drazin and Johnson, 1989] and section III of [Aktosun, 2009].

8.1 Initial value problems

So far we have seen a variety of methods to construct **particular solutions** to integrable PDEs.

Question: can we find a **general solution** to these PDEs?

In more detail, we want to solve the following **Initial Value Problem** (IVP):

> Given a wave equation and 'enough' initial data at an initial time $t = 0$, find $u(x, t)$ at all later times $t > 0$.

For there to be a unique solution, **sufficient initial data** must be specified:

- If the PDE is **1st order in time**, *e.g.* KdV, we must specify $\underline{u(x, 0)}$;
- If the PDE is **2nd order in time**, *e.g.* sine-Gordon, we must specify $\underline{u(x, 0),\, u_t(x, 0)}$;

- etc.

[Why stop there? This is because given these initial data we can use the PDE to solve for higher t derivatives at $t = 0$. *E.g.* for KdV, if I tell you $u(x, 0)$, you can use the PDE to find out what $u_t(x, 0)$ must be: it's not an independent datum.]

But given that information, can we actually construct $u(x, t)$ for all $t > 0$, analytically if possible? So far, the answer is no, unless the initial condition happens to be a snapshot of one of the special solutions seen before.

E.g. in KdV, what if

(a) $u(x, 0) = 2\,\mathrm{sech}^2(x)$

(b) $u(x, 0) = 2.001\,\mathrm{sech}^2(x)$

(c) $u(x, 0) = 6\,\mathrm{sech}^2(x)$?

Case (a) is a snapshot of a one-soliton solution at $t = 0$, so, assuming the uniqueness of solutions,[1] the answer to (a) at all later times is

$$u(x, t > 0) = 2\mathrm{sech}^2(x - 4t)\,,$$

which describes a single soliton moving to the right with velocity $v = 4$. But what about (b) and (c)?

It turns out that

(b) $\rightarrow$ { 2 solitons, 1 very small, both moving right, + some dispersing junk moving left }

(c) $\rightarrow$ { 2 solitons, both moving right, and that's all }

In fact, the initial condition for (c) is a snapshot of a "pure" 2-soliton solution.

Inverse scattering will allow us to understand situations like (b), and give a much more complete understanding of when things like (a) and (c) occur. As you might remember seeing "experimentally" at the start of last term, whenever the height h of the $u(x, 0) = h\mathrm{sech}^2(x)$ initial condition of KdV is $h = N(N+1)$, with $N = 1, 2, 3, \ldots$, we are in a situation like (a) or (c), where this is precisely an N-soliton solution. But why? Inverse scattering gives analytic

[1] Uniqueness of the solution follows if the IVP is *well-posed.* It can be shown that the KdV initial value problem is locally/globally well-posed if one works in a suitable function space [Kenig et al., 1991]

insight into this question.

To get an idea of how this might work, let's first look at a simpler setting: linear wave equations.

8.2 Linear initial value problems

For a **linear wave equation**, the **general solution** is a **linear transformation** of the **initial data**.

Examples

1. The heat equation

$$\boxed{u_t = u_{xx}} \qquad (x \in \mathbb{R}, t > 0)\,. \tag{8.1}$$

Given the initial data $u(x,0) \equiv u_0(x)$, the general solution $u(x,t)$ at $t \geqslant 0$ is

$$\boxed{u(x,t) = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{4\pi t}}\, e^{-(x-x')^2/(4t)} u_0(x')\, dx'} \tag{8.2}$$

and this *is* a linear transform of $u_0(x)$ (it's actually a "Green's function" solution).

* **EXERCISE**: Check that (8.2) solves (8.1) and reduces to $u(x,0) = u_0(x)$ at $t = 0$.

2. The Klein-Gordon equation

$$\boxed{u_{tt} - u_{xx} + m^2 u = 0} \qquad (x \in \mathbb{R}, t > 0)\,. \tag{8.3}$$

This is second-order in t, so we need to specify $u(x,0)$ ***and*** $u_t(x,0)$:

$$u(x,0) = \alpha(x)\,, \qquad u_t(x,0) = \beta(x)\,. \tag{8.4}$$

With luck, { (8.3) + (8.4) } is a well-posed initial value problem. t can be solved using a **Fourier transform** with respect to x, seen in AMV. Recall that this is like the Fourier series, but for functions on a infinite line. See the handout for a reminder of the key properties of the Fourier transform that we'll need.

Given $u(x,t)$, set

$$\boxed{\begin{aligned} \tilde{u}(k,t) &= \int_{-\infty}^{+\infty} dx\, e^{-ikx}\, u(x,t) \\ u(x,t) &= \int_{-\infty}^{+\infty} \frac{dk}{2\pi}\, e^{ikx}\, \tilde{u}(k,t) \end{aligned}} \tag{8.5}$$

where the second equation (‘*inverse Fourier transform*’) shows how to get u back from $\widetilde{u}$. Note that t is a spectator in the (inverse) Fourier transform, which trades x for the conjugate variable k.

Working with $\widetilde{u}(k,t)$ instead of $u(x,t)$ is a good move, because (8.3) for u implies

$$\boxed{\widetilde{u}_{tt} + (k^2+m^2)\widetilde{u} = 0} \tag{8.6}$$

for $\widetilde{u}$, and this equation is easier to solve: there are only t derivatives, so it is as an *ordinary* differential equation (ODE) in t for each value of k, rather than a PDE in x,t.

The general solution of (8.6) is

$$\begin{aligned} \widetilde{u}(k,t) &= A(k)\,e^{i\omega t} + B(k)\,e^{-i\omega t}\,, \qquad \text{where} \\ \omega = \omega(k) &\equiv \sqrt{k^2+m^2} \qquad \text{(dispersion relation)} \end{aligned} \tag{8.7}$$

and the integration constants A and B can be fixed by matching with the initial condition at $t=0$:

$$\begin{aligned} \widetilde{\alpha}(k) &\equiv \widetilde{u}_(k,0) = A(k) + B(k) \\ \widetilde{\beta}(k) &\equiv \widetilde{u}_t(k,0) = i\omega\,(A(k) - B(k))\ . \end{aligned} \tag{8.8}$$

Solving for A and B and simplifying the resulting expression for the Fourier transformed field $\widetilde{u}(k,t)$, we obtain

$$\boxed{\begin{aligned} \widetilde{u}(k,t) &= \frac{1}{2}\left(\widetilde{\alpha}(k) + \frac{\widetilde{\beta}(k)}{i\omega}\right)e^{i\omega t} + \frac{1}{2}\left(\widetilde{\alpha}(k) - \frac{\widetilde{\beta}(k)}{i\omega}\right)e^{-i\omega t} \\ &= \widetilde{\alpha}(k)\cos(\omega t) + \frac{1}{\omega}\widetilde{\beta}(k)\sin(\omega t)\ . \end{aligned}} \tag{8.9}$$

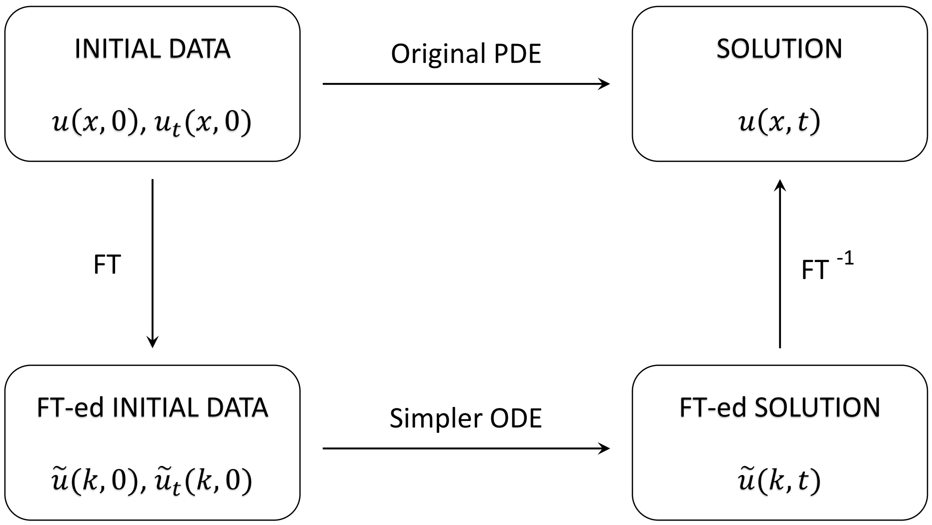
Finally, an **inverse Fourier transform** allows $u(x,t)$ to be found:

$$\boxed{\begin{aligned} u(x,t) &= \int_{-\infty}^{+\infty}\frac{dk}{2\pi}\left[\widetilde{\alpha}(k)\cos(\omega t) + \frac{1}{\omega}\widetilde{\beta}(k)\sin(\omega t)\right]e^{ikx} \\ &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\frac{dk\,dx'}{2\pi}\left[u(x',0)\cos(\omega t) + \frac{1}{\omega}u_t(x',0)\sin(\omega t)\right]e^{ik(x-x')} \end{aligned}} \tag{8.10}$$

with $\omega = \omega(k) = \sqrt{k^2+m^2}$. Again, this is a **linear** function of the initial data $u(x,0)$ and $u_t(x,0)$. (This won’t be true for KdV.)

<u>**KEY FEATURE:**</u> the **Fourier transformed data** $\widetilde{u}(k,t)$ for each value of k **evolved separately** and in a **simple** way in the “transformed” equation (8.6). (Something like this *will* still be true for KdV.)

GENERAL PICTURE for solving **linear Initial Value Problems**:[2]

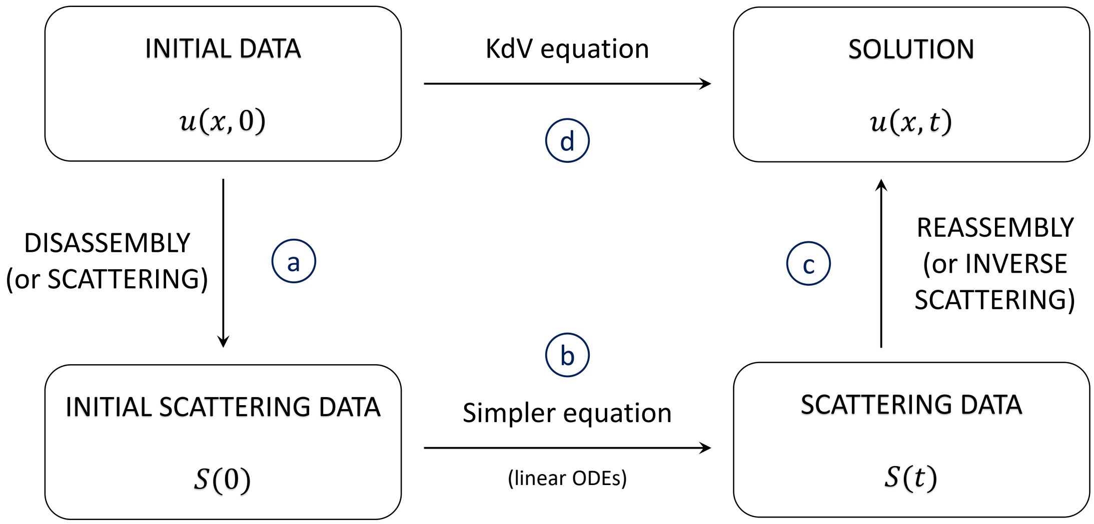


and we follow the indirect path $\downarrow {}_{\rightarrow} \uparrow$ to solve the direct but harder problem $^{\rightarrow}$.

8.3 Outline of the method to solve the IVP for KdV

The above logic will turn out to be the correct "big idea" for KdV also, but in a much more subtle way since the KdV equation is nonlinear, therefore the Fourier transform is of no help. The method goes as follows:

[2]The initial data are written for 2nd order in time equations such as the Klein-Gordon equation, and need to be modified appropriately according to the degree in time of the differential equation.



Again, instead of doing step (d) directly, we will go the roundabout route of (a) → (b) → (c). All these three steps will turn out to be nontrivial, even though simpler than (d).

8.3.1 The KdV-Schrödinger connection

We will follow the route of the original discoverers of the method, **Gardner, Greene, Kruskal and Miura (GGKM)**, in 1967 [Gardner et al., 1967]. Their aim was to solve the **KdV equation**

$$\boxed{u_t + 6uu_x + u_{xxx} = 0} \tag{8.11}$$

for $t > 0$ on the full line $-\infty < x < \infty$, with the initial condition

$$\boxed{u(x,0) = f(x)}\ , \tag{8.12}$$

where $f(x)$ is sufficiently localised in space.[3]

Recall first the (generalised) Miura transformation: if $v(x,t)$ satisfies

$$v_t + 6(\lambda - v^2)v_x + v_{xxx} = 0\ , \tag{8.13}$$

then

$$\boxed{u = \lambda - v^2 - v_x} \tag{8.14}$$

[3]Technically, we demand

$$\int_{-\infty}^{+\infty} dx\ (1 + |x|)|f(x)| < \infty\ .$$

solves the KdV equation (8.11). Now think about this backwards: take u to be known, and try to solve (8.14) for v. (We can ignore (8.13) from now on.) Equation (8.14) is a so called Riccati equation (a 1st order ODE quadratic in the unknown), and there is a standard trick to solve such an equation be rewriting it as a linear 2nd order ODE: write

$$\boxed{v = \psi_x/\psi} \tag{8.15}$$

for some other function ψ, and try to find ψ first. With a small amount of rearrangement (8.14) becomes

$$\boxed{\psi_{xx} + u\psi = \lambda\psi}\,. \tag{8.16}$$

Now (8.16) is interesting (and this is what attracted GGKM's attention) because it's a well-known equation: the **time-independent Schrödinger equation**, the quantum-mechanical equation for a point particle moving in a potential $V(x) = -u(x)$.

The QM interpretation of the equation won't be too important here, apart from the fact that a great deal was known about its solutions, and GGKM were able to exploit this. The important thing is that any **field profile** u can be associated with another family of **functions** ψ by solving (8.16), which is sometimes called the **associated linear problem**.

8.3.2 Recipe for constructing the KdV solution (in 3 steps)

Let's begin to unpack the diagram 8.3.

a **DISASSEMBLY (or SCATTERING)**

Note that t appears in (8.16) only as a parameter, in $u(x,t)$ (the differential equation is in x). Start at $\underline{t=0}$: the initial data $u(x,0)$ plays the role of a potential. For each eigenvalue λ, ψ is a different "eigenfunction" of the differential operator $L = \frac{\partial^2}{\partial x^2} + u$.[4] This is the **associated linear problem** (at $t = 0$). The function $\psi(x)$ describes the **scattering** of a quantum-mechanical particle (or wave) off the potential, with certain **reflection/transmission coefficients**, which capture the asymptotic behaviour of ψ as $x \to \pm\infty$. The set of these coefficients, for different values of λ, are the **initial scattering data** $S(0)$.[5] The scattering data is analogous to $\widetilde{\alpha}(k)$ and $\widetilde{\beta}(k)$ in the linear (Klein-Gordon) case, while λ is like k.

b **TIME EVOLUTION**

Next we'll have to **evolve the eigenvalues** λ and the **scattering data** S forward in time t. Here we are helped by an amazing fact: if $u(x,t)$ solves the KdV equation, then the eigenvalues

[4] ψ is an eigenvector of the operator L (in an infinite dimensional vector space, in fact a Hilbert space). But L is a differential operator in x and f is therefore a function of x, so it's called an "eigenfunction".

[5] More precisely, we will need the reflection coefficients plus some extra data. More about this later.

$\lambda(t)$ are independent of t. Therefore we only need to **evolve the scattering data** $S(0)$ into $S(t)$. It turns out that this **time evolution** is **simple**: it is governed by **linear ODEs**. This uses an ingenious idea of Peter Lax.

c **REASSEMBLY (or INVERSE SCATTERING)**

The final step is to **reconstruct the potential** $u(x,t)$ **from the scattering data** $S(t)$ at time t. This is called **"inverse scattering"**. It may be surprising that one can do it – 'Can one hear the shape of a drum?', as Marc Kac put in a related context –, but that this is possible for these Schroedinger problems was already known at the time of GGKM.

This will be a long story, so it will be good to keep this "roadmap" in mind as we go, starting with step (a).

Chapter 9

The basics of scattering theory

A reference for this chapter are sections 3.1 and 3.2 of [Drazin and Johnson, 1989]. Any Quantum Mechanics book will also cover the relevant material.

Our aim in this chapter is to analyse the possible solutions to the eigenvalue problem $L\psi = \lambda\psi$, that is

$$\left(\frac{d^2}{dx^2} + u(x)\right)\psi(x) = \lambda\psi(x) \tag{9.1}$$

with $\psi(x)$ **bounded** for all x (which restricts the possible values of λ). Note that this relaxes slightly the requirement that $\int_{-\infty}^{+\infty}|\psi|^2 dx < \infty$, *i.e.* $\psi \in L^2(\mathbb{R})$, which is made in the context of Quantum Mechanics.

Note: in this chapter the **KdV time** t just appears as a **parameter** in $u(x,t)$, and stays fixed (therefore will be dropped from the notation). The operator L in the eigenvalue problem (9.1) should then be viewed as a (second order) ordinary differential operator in x.

9.1 Overview: the physical interpretation

FACT: the equation

$$\boxed{i\frac{\partial}{\partial\tau}\Psi(x,\tau) = \left(-\frac{\partial^2}{\partial x^2} + V(x)\right)\Psi(x,\tau)\,,} \tag{9.2}$$

known as the **time-dependent Schrödinger equation**, describes a particle (of mass $\frac{1}{2}$) moving on a line in a potential $V(x)$ in quantum mechanics (in natural units, where $\hbar = 1$). The *wavefunction* Ψ tells you where the particle is likely to be: $|\Psi(x,\tau)|^2 dx$ is the probability to find it in the interval $[x, x + dx]$ at time τ. (Note: this time τ is not to be confused with the KdV time t.)

To solve (9.2), separate variables

$$\Psi(x,\tau) = \psi(x)\phi(\tau) \tag{9.3}$$

and substitute in and rearrange to find

$$i\frac{\dot{\phi}}{\phi} = \frac{-\psi'' + V\psi}{\psi} = \text{constant} \equiv k^2 \ , \tag{9.4}$$

where the dot denotes $\frac{d}{d\tau}$, the dash $\frac{d}{dx}$, and the constant was called k^2 for later convenience. Solving first the equation for ϕ,

$$\dot{\phi} = -ik^2\phi \quad \Rightarrow \quad \boxed{\phi(\tau) = e^{-ik^2\tau}} \ , \tag{9.5}$$

while $\psi(x)$ satisfies the **time independent Schrödinger equation**

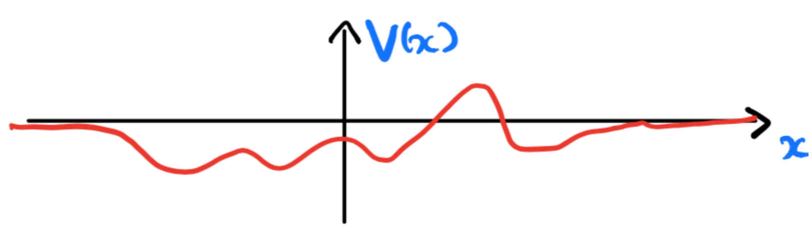
$$\boxed{\left(-\frac{d^2}{dx^2} + V(x)\right)\psi(x) = k^2\psi(x)} \ , \tag{9.6}$$

which is the same as (9.1) with the identifications

$$\boxed{\begin{aligned} u &= -V \\ \lambda &= -k^2 \end{aligned}} \ . \tag{9.7}$$

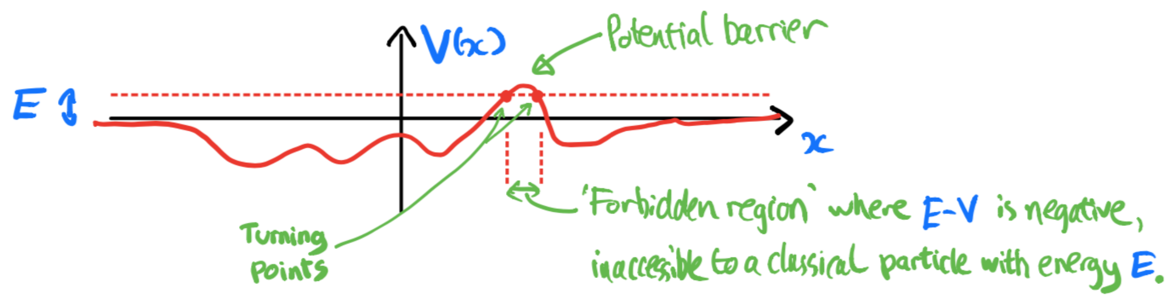
In quantum mechanics, (9.6) then describes a particle with energy $E = k^2 = -\lambda$ moving in the potential $V(x) = -u(x)$.

With the link to the KdV equation in mind, we'll consider potentials which tend to zero (sufficiently fast, as in footnote 3) as $x \to \pm\infty$:



In classical mechanics, a particle with total (kinetic plus potential) energy $E = T + V$ is localised, and bounces off the potential at the "turning points" x_* where $V(x_*) = E$.

By contrast, in quantum mechanics, there's a non-zero chance to find the particle anywhere (where V is finite), and the particle can 'tunnel' through potential barriers which are impenetrable in classical mechanics.



The **scattering data** will be encoded in the **asymptotics** (limiting behaviour) of $\psi(x)$ as $x \to \pm\infty$.

Since $V(x) \to 0$ as $x \to \pm\infty$, (9.6) in these regions reduces to

$$\boxed{-\frac{d^2}{dx^2}\psi = k^2\psi} \tag{9.8}$$

with two independent (plane wave) solutions $e^{\pm ikx}$. (Hence $k = \sqrt{E} = \sqrt{-\lambda}$ is the absolute value of the wave number of these plane waves.)

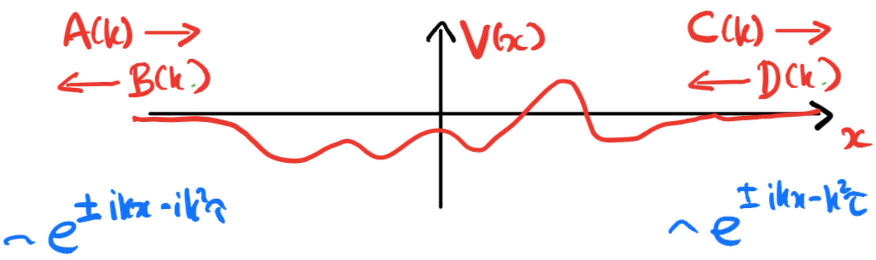
So the **general solution** with eigenvalue $E = k^2$ has the asymptotics

$$\boxed{\begin{array}{ll} \psi(x) \approx A(k)e^{ikx} + B(k)e^{-ikx}\,, & x \to -\infty \\ \psi(x) \approx C(k)e^{ikx} + D(k)e^{-ikx}\,, & x \to +\infty \end{array}} \tag{9.9}$$

and, restoring the τ-dependence,

$$\boxed{\begin{array}{ll} \Psi(x,\tau) \approx A(k)e^{ikx-ik^2\tau} + B(k)e^{-ikx-ik^2\tau}\,, & x \to -\infty \\ \Psi(x,\tau) \approx C(k)e^{ikx-ik^2\tau} + D(k)e^{-ikx-ik^2\tau}\,, & x \to +\infty \end{array}} \tag{9.10}$$

showing that for real $k > 0$ the 'A' and 'C' terms correspond to right-moving waves , while the 'B' and 'D' terms correspond to left-moving waves, taking $k > 0$ without loss of generality:



This solution will be bounded for any values for A, B, C and D if $E = k^2 > 0$.

As we'll see in examples, solving (9.6) in the middle region where $V(x) \neq 0$ interpolates between the two asymptotic regions and imposes two relations among A, B, C and D, leaving two undetermined coefficients, as expected for a 2nd order ODE.

To fix these remaining coefficients, for $k^2 > 0$ we will impose

$$\boxed{A(k) = 1\,, \qquad D(k) = 0} \tag{9.11}$$

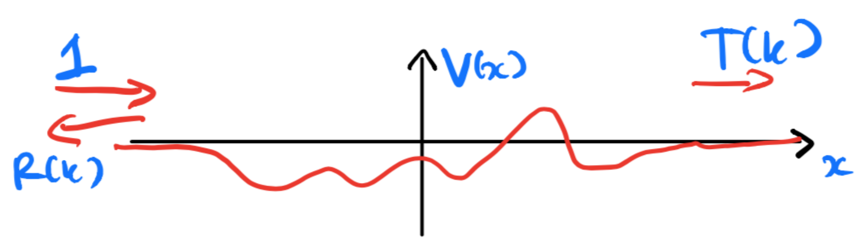
and write

$$\begin{aligned} B(k) &\equiv R(k) \qquad \text{(the \textbf{reflection coefficient})} \\ C(k) &\equiv T(k) \qquad \text{(the \textbf{transmission coefficient})} \end{aligned} \tag{9.12}$$

so that the resulting **scattering solution** has asymptotics

$$\boxed{\begin{aligned} \psi(x) &\approx e^{ikx} + R(k)e^{-ikx}\,, \qquad & x \to -\infty \\ \psi(x) &\approx T(k)e^{ikx}\,, \qquad & x \to +\infty \end{aligned}} \tag{9.13}$$

and represents a unit flux (since $A(k) = 1$) of incoming particles from the left, partially reflected from the potential and partially transmitted through it:



It can be shown (Ex 52) that

$$\boxed{|R(k)|^2 + |T(k)|^2 = 1} \tag{9.14}$$

meaning that with probability 1 the particle is either reflected or transmitted (conservation of probability).

Aside: the Wronskian

Results such as $|R(k)|^2 + |T(k)|^2 = 1$, derived in exercise 52, are proved using a gadget called the *Wronskian*. For two functions $f(x)$ and $g(x)$, their Wronskian is the function

$$\boxed{W[f,g](x) = f'(x)g(x) - f(x)g'(x)\,.}$$

Two facts about W:

1) If f and g are linearly dependent, the $W[f,g] = 0$ identically.

(It's easy to see that W is antisymmetric, and linear in each of its arguments. Then if, say, $f(x) = \alpha g(x)$ with α a constant, $W[f,g] = W[\alpha g, g] = \alpha W[g,g] = 0$.)

2) The converse statement, that $W[f,g] = 0$ implies that f and g are linearly dependent, is more tricky. The following is easily proved: if
  a) $W[f,g](x) = 0$ on some interval, and
  b) one or other of f and g is nonzero on that interval,
then f and g are linearly dependent on that interval.

(Say it's g that is nonzero. Dividing $W[f,g](x) = 0$ through by g^2 shows that $\frac{d}{dx}(f/g) = 0$, so $f/g =$ constant, and f and g are linearly dependent.)

Notes:

• Some sort of extra condition such as b) *is* needed: consider, as suggested by Peano in 1889,

$$f(x) = x^2\,, \quad g(x) = x|x| = x^2\mathrm{sign}(x)\,.$$

Then f and g are not linearly dependent on $\mathbb{R}$, even though $W[f,g] = 0$ everywhere. (Exercise: check this!)

• In fact, though it won't be proved here, the result that $W[f,g](x) = 0$ everywhere implies f and g are linearly dependent *does* hold if both f and g are analytic. This is true of solutions to the ODEs we are dealing with here, so we will therefore assume that the converse statement to **1)** does hold in all cases we will need.

Now back to the time independent Schrödinger equation

$$\left(-\frac{d^2}{dx^2} + V(x)\right)\psi(x) = E\psi(x) = k^2\psi(x)\,.$$

So far we have looked at cases with $k^2 = E > 0$. For $\underline{k^2 < 0}$, let $k = i\mu$ with $\mu > 0$ real, so $E = -\mu^2$. Then the asymptotics of the general solution (9.9) become

$$\boxed{\begin{array}{ll} \psi(x) \approx a(\mu)e^{-\mu x} + b(\mu)e^{\mu x}\,, & x \to -\infty \\ \psi(x) \approx c(\mu)e^{-\mu x} + d(\mu)e^{\mu x}\,, & x \to +\infty \end{array}} \tag{9.15}$$

and it follows that

$$\boxed{\psi \text{ bounded}} \quad \Leftrightarrow \quad \boxed{a(\mu) = d(\mu) = 0} \tag{9.16}$$

In such cases ψ is not only bounded, it also tends to zero at $\pm\infty$ and satisfies $\int_{-\infty}^{+\infty} |\psi|^2 dx < 0$.

Note that there might be no values for μ at which this happens. But if it does, the corresponding ψ is called a **bound state solution**.

Fact: Given a potential $V(x)$ tending to zero at $\pm\infty$ (and $> -\infty$ except possibly for discrete values of x), bound state solutions only exist for a finite (possibly empty) set of μ's:

$$\{\mu_k\}_{k=1}^{N} = \{\mu_1, \mu_2, \dots \mu_N\}, \qquad \mu_1 < \mu_2 < \dots < \mu_N\,. \tag{9.17}$$

Summary

Bounded solutions to

$$\left(-\frac{d^2}{dx^2} + V(x)\right)\psi(x) = E\psi(x) = k^2\psi(x),$$

or equivalently $\left(\frac{d^2}{dx^2} + u(x)\right)\psi(x) = \lambda\psi(x)$ with $u(x) = -V(x)$ and $\lambda = -E$, come in two flavours when $V(x) \to 0$ as $x \to \pm\infty$:

1. $E = k^2 = -\lambda \in (0, +\infty)$: the **"continuous spectrum"**, leading to **scattering solutions** which are **bounded**, and have **oscillatory asymptotics**;

2. $E = -\mu^2 = -\lambda \in \{-\mu_1^2, -\mu_2^2, \dots - \mu_N^2\}$: the **"discrete spectrum"**, leading to **bound state solutions** which are **square integrable** (*i.e.* $\int_{-\infty}^{+\infty} |\psi(x)|^2 dx < \infty$), and have **damped asymptotics**.

(Note: for some rather-special, slowly-decaying potentials, there may also be some square integrable solutions with $k^2 > 0$. These so-called 'bound states in the continuum' (BICs) crop up in a number of physical applications, but won't be relevant for the current discussion.)

9.2 Examples

9.2.1 Zero potential: $V(x) = 0$

This was already done, essentially, when looking at the asymptotics for general V. We must solve $-\frac{d^2}{dx^2}\psi = k^2\psi$ for all $x \in \mathbb{R}$. There are **two cases** to consider.

(a) $k^2 > 0$:

The general solution, valid for all x, not just asymptotically, is

$$\psi(x) = Ae^{ikx} + Be^{-ikx} \,. \tag{9.18}$$

Restoring the τ dependence, it describes a left or right moving wave, bounded for all real values of k.

Comparing with (9.9) shows that in this case $C(k) = A(k)$ and $D(k) = B(k)$. Imposing $A(k) = 1$ and $D(k) = 0$ then gives us the scattering solution:

$$\psi(x) = e^{ikx} \,,$$

from which it follows that

$$\boxed{R(k) = 0 \,, \quad T(k) = 1} \tag{9.19}$$

If you think about it this should seem reasonable: with no potential, a particle incident from the left is transmitted through to the right with probability 1.

(b) $k^2 = -\mu^2 < 0$:

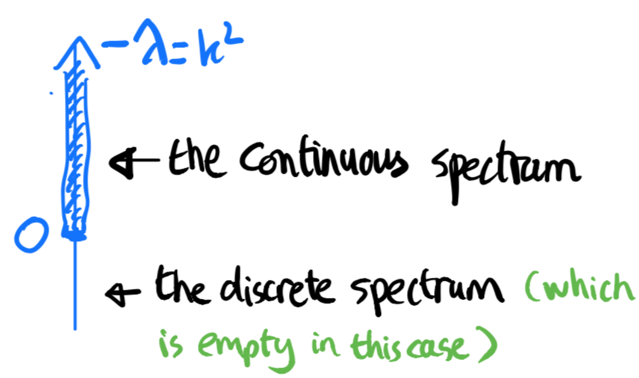
The general solution from part (a) turns into

$$\psi(x) = ae^{-\mu x} + be^{\mu x} \,. \tag{9.20}$$

and the only way to keep this bounded as $x \to \pm\infty$ is to set $a = b = 0$. Thus there are no bound state solutions for this problem.

Summary

For $u = 0$, the problem $L(u)\psi = \lambda\psi$, ψ bounded, has a 'scattering' solution for all real $\lambda < 0$, and no solutions for $\lambda > 0$:



9.2.2 Delta function potential: $V(x) = a\delta(x)$

Here a is a real constant and $\delta(x)$ is the Dirac delta function. Recall that $\delta(x)$ can be viewed as the limit of a sequence (a 'delta sequence') of unit-area functions which are increasingly concentrated at the origin, so that for any continuous[1] (test) function $f(x)$,

$$\int_{-\infty}^{+\infty} \delta(x) f(x)\, dx = f(0)\,.$$

We seek a single solution $\psi(x)$, solving the equation in regions (1) and (2), and also consistent with the potential at $x = 0$.

(a) $\underline{k^2 > 0}$:

In regions (1) and (2), $V(x) = 0$, so ψ satisfies $-\frac{d^2}{dx^2}\psi = k^2\psi$ and as in example 1, the solutions in the two regions are

$$\boxed{\psi(x) = \begin{cases} A(k)e^{ikx} + B(k)e^{-ikx}\,, & x < 0 \\ C(k)e^{ikx} + D(k)e^{-ikx}\,, & x > 0 \end{cases}} \tag{9.21}$$

To finish, we must match the two parts of the solution at $x = 0$, and this will determine the relation(s) between A, B, C and D.

To find the matching conditions, let us integrate the time-independent Schroedinger equation

$$-\psi''(x) + a\delta(x)\psi(x) = k^2\psi(x) \tag{9.22}$$

in an infinitesimal neighbourhood $[-\epsilon, +\epsilon]$ of $x = 0$:

$$\int_{-\epsilon}^{+\epsilon} dx\,[-\psi''(x) + a\delta(x)\psi(x)] = k^2 \int_{-\epsilon}^{+\epsilon} dx\,\psi(x)$$

[1] If $f(x)$ is discontinuous, then the result of the integral is the average of the right and left limits $\lim_{x\to 0^\pm} f(x)$. We won't need to worry about such subtleties.

$$\Rightarrow \qquad -\left[\psi'(x)\right]_{-\epsilon}^{+\epsilon} + a\psi(0) = k^2 \int_{-\epsilon}^{+\epsilon} dx\, \psi(x)\,. \tag{9.23}$$

Provided that ψ is bounded (which it is), the RHS of this equation tends to 0 as $\epsilon \to 0$, and taking this limit implies $-\left[\,\psi'(x)\,\right]_{0^-}^{0^+} + a\psi(0) = 0$, or

$$\boxed{\left[\,\psi'(x)\,\right]_{0^-}^{0^+} = a\psi(0)}\,. \tag{9.24}$$

The attentive reader will have noticed that we have implicitly assumed that $\psi(x)$ is continuous at $x = 0$. It is not hard to relax that assumption to deduce a modified version of (9.24), which can then be used to deduce that $\psi(x)$ is continuous:

$$\boxed{\left[\,\psi(x)\,\right]_{0^-}^{0^+} = 0}\,. \tag{9.25}$$

(Alternatively, the same conclusion can be derived by integrating (9.22) twice. We recommend this exercise to the motivated reader.)

Applying the **matching conditions** (9.25) and (9.24) to (9.21) we have

$$\begin{cases} A + B = C + D \\ ik(C - D) - ik(A - B) = a(A + B) = a(C + D) \end{cases}$$

which in turn implies

$$\begin{cases} A + B = C + D \\ A - B = \left(1 - \frac{a}{ik}\right) C - \left(1 - \frac{a}{ik}\right) D\,. \end{cases}$$

Adding and subtracting,

$$\begin{aligned} A &= \left(1 - \frac{a}{2ik}\right) C - \frac{a}{2ik} D \\ B &= \frac{a}{ik} C + \left(1 + \frac{a}{2ik}\right) D\,. \end{aligned} \tag{9.26}$$

Substituting into (9.21) gives the general solution, with two undetermined constants as expected.

To get to the **scattering solution**, set $D = 0$ and then divide through so that $A = 1$:

$$\boxed{\psi(x) = \begin{cases} e^{ikx} + \frac{a}{2ik-a}\, e^{-ikx}\,, & x < 0 \\ \frac{2ik}{2ik-a}\, e^{ikx}\,, & x > 0 \end{cases}} \tag{9.27}$$

and from this the **reflection and transmission coefficients** can be read off:

$$\begin{aligned} R(k) &= \frac{a}{2ik - a} \\ T(k) &= \frac{2ik}{2ik - a} \end{aligned} \tag{9.28}$$

and it's easy to see that

$$|R(k)|^2 + |T(k)|^2 = 1 \tag{9.29}$$

as expected.

(b) $\underline{k^2 = -\mu^2 < 0}$:

Setting $k = i\mu$ in (9.21), (9.26) with $\mu > 0$ we obtain the general solution in this regime:

$$\psi(x) = \begin{cases} A(i\mu)e^{-\mu x} + B(i\mu)e^{\mu x}\,, & x < 0 \\ C(i\mu)e^{-\mu x} + D(i\mu)e^{\mu x}\,, & x > 0 \end{cases} \tag{9.30}$$

Given that we chose $\mu > 0$, this is **bounded** as $x \to \pm\infty$ iff

$$\boxed{A(i\mu) = D(i\mu) = 0}\,. \tag{9.31}$$

Substituting into (9.26),

$$\begin{cases} 0 = \left(1 + \frac{a}{2\mu}\right)C \\ B = -\frac{a}{\mu}C \end{cases}$$

giving two options:

1. $A = B = C = D = 0$ (trivial);

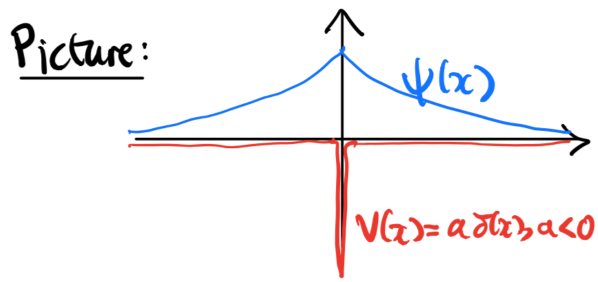
2. $$\boxed{\mu = -\frac{a}{2}\,, \quad B = C}\,. \tag{9.32}$$

Given that we took $\mu > 0$, option 2 means that there is a bounded solution with $k^2 < 0$ only for $\underline{a < 0}$. The bound state solution is then

$$\boxed{\psi(x) = e^{\frac{a}{2}|x|} = \begin{cases} e^{\frac{a}{2}x}\,, & x < 0 \\ e^{-\frac{a}{2}x}\,, & x > 0 \end{cases}} \tag{9.33}$$

and for this case

$$\boxed{k^2 = -\frac{a^2}{4}}\,. \tag{9.34}$$



NOTE: we can obtain bound state solutions by an alternative method. First, observe that we need $D(k) = 0$ for the general solution to be bounded as $x \to \pm\infty$. Hence we could substitute $k = i\mu$ directly in the scattering solution (9.27):

$$\boxed{\psi(x) = \begin{cases} e^{-\mu x} + \frac{a}{-2\mu - a}\, e^{\mu x} \,, & x < 0 \\ \frac{-2\mu}{-2\mu - a}\, e^{-\mu x} \,, & x > 0 \end{cases}} \,, \tag{9.35}$$

which looks hopelessly unbounded as $x \to -\infty$ due to the $e^{-\mu x}$ term. There is a **trick** though: let's change normalization by dividing through by $T(i\mu) = \frac{2\mu}{2\mu+a}$, which gives

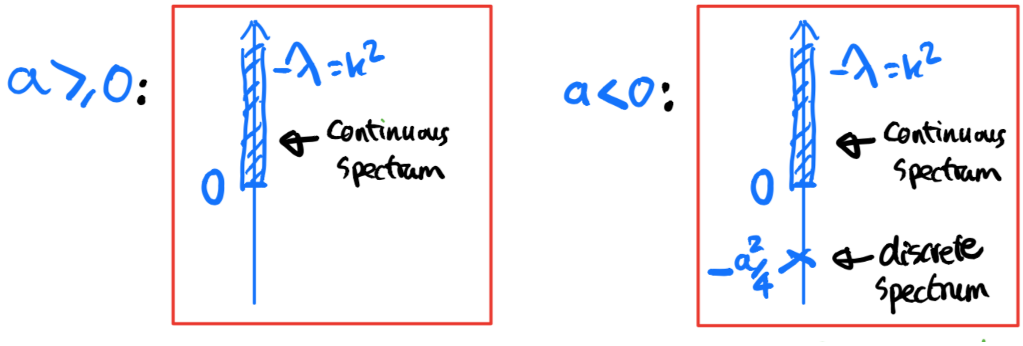
$$\boxed{\psi(x) = \begin{cases} \frac{2\mu+a}{2\mu} e^{-\mu x} - \frac{a}{2\mu}\, e^{\mu x} \,, & x < 0 \\ e^{-\mu x} \,, & x > 0 \end{cases}} \,. \tag{9.36}$$

This re-normalized solution is now bounded as $x \to -\infty$ if (and only if) $\mu = -\frac{a}{2}$, in which case we recover the bound state solution (9.33).[2]

Summary

For $V(x) = -u(x) = a\delta(x)$, the eigenvalue problem $L(u)\psi = \lambda\psi$, with ψ bounded, has a scattering solution for all real $\lambda < 0$, and either no solutions for $\lambda > 0$ if $a \geqslant 0$, or one solution for $\lambda > 0$ if $a < 0$:

[2]We could have alternatively divided by $R(i\mu)$ and reached the same conclusion.



- For **all** $k^2 > 0$, a **scattering solution**

$$\boxed{\psi(x) = \begin{cases} e^{ikx} + R(k)\, e^{-ikx}\,, & x < 0 \\ T(k)\, e^{ikx}\,, & x > 0 \end{cases}} \tag{9.37}$$

  exists with reflection and transmission coefficients

$$\boxed{R(k) = \frac{a}{2ik - a}\,, \qquad T(k) = \frac{2ik}{2ik - a}}\,. \tag{9.38}$$

- For **isolated** $k^2 = -\mu^2 < 0$, a **bound state solution**

$$\boxed{\psi(x) = \begin{cases} \frac{R(i\mu)}{T(i\mu)} e^{\mu x}\,, & x < 0 \\ e^{-\mu x}\,, & x > 0 \end{cases}} \tag{9.39}$$

  exists if $\mu = -a/2$, such that

$$\boxed{\frac{1}{T(i\mu)} = 0}\,. \tag{9.40}$$

The general story

For potentials $V(x)$ which tend to zero as $x \to \pm\infty$, **bound state solutions** can be obtained **from scattering solutions** by

1. dividing the scattering solution through by $T(k)$;

2. setting

$$\boxed{k = i\mu = \text{pole of } T(k) \text{ on the positive imaginary axis}}\,.$$

The above condition determines the **discrete spectrum** of $-\frac{d^2}{dx^2} + V(x)$.

Note: the transmission coefficient $T(k)$ can be equivalently replaced by the reflection coefficient $R(k)$ above. Their poles coincide because they satisfy (9.14).

For more examples, see Ex 53 and 54 in the problem sheet.

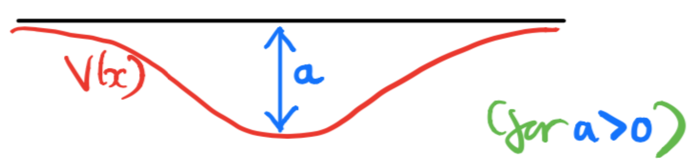
9.3 Reflectionless potentials

We now return to the initial field configurations $u(x,0) = a\,\mathrm{sech}^2(x)$ that were tried for the KdV field earlier. These seemed to lead to interesting field evolutions whenever a was equal to $n(n+1)$ with n a positive integer, and it's natural to wonder whether this interesting behaviour is also apparent in the correpsonding scattering problem.

The relevant potential is

$$\boxed{V(x) = -a\,\mathrm{sech}^2(x)} \tag{9.41}$$

as illustrated below:



The time independent Schödinger equation (T.I.S.E.) to be solved is

$$\boxed{-\psi''(x) - a\,\mathrm{sech}^2(x)\,\psi(x) = k^2\psi(x)} \tag{9.42}$$

and we're after **bounded** solutions to this eigenvalue problem.

Substituting

$$\boxed{y = \tanh(x)} \in (-1, 1) \tag{9.43}$$

so that

$$\frac{d}{dx} = \mathrm{sech}^2(x)\frac{d}{dy} = (1-y^2)\frac{d}{dy} \tag{9.44}$$

the T.I.S.E. becomes

$$\boxed{\frac{d}{dy}\left[(1-y^2)\frac{d\psi}{dy}\right] + \left(\frac{k^2}{1-y^2} + a\right)\psi = 0} \tag{9.45}$$

and putting

$$\boxed{k^2 = -m^2\,, \quad a = n(n+1)} \tag{9.46}$$

this becomes the **general (or associated) Legendre equation**:

$$\boxed{\frac{d}{dy}\Big[(1-y^2)\frac{d\psi}{dy}\Big] + \Big(n(n+1) - \frac{m^2}{1-y^2}\Big)\psi = 0} \tag{9.47}$$

This equation has been much studied, and its solutions are known in general in terms of certain special functions.

Fact 1:

If $n = 1, 2, 3, \dots$ (*i.e.* $\underline{n \in \mathbb{Z}_{\geqslant 0}}$) and $\underline{m = 0}$ (so $k = 0$), then (9.47) becomes the **Legendre equation** and its bounded solution for $y \in [-1, 1]$ is

$$\psi = P_n(y) = \frac{1}{n!\,2^n}\frac{d^n}{dy^n}(y^2-1)^n\,, \tag{9.48}$$

the n^{th} **Legendre polynomial of the first kind**. The first few examples are:

$$\begin{aligned}
P_1(y) &= y \\
P_2(y) &= -\tfrac{1}{2} + \tfrac{3}{2}y^2 \\
P_3(y) &= -\tfrac{3}{2}y + \tfrac{5}{2}y^3 \\
P_4(y) &= \tfrac{3}{8} - \tfrac{15}{4}y^2 + \tfrac{35}{8}y^4\,.
\end{aligned}$$

In general, $P_j(-y) = (-1)^j P_j(y)$ and $P_j(1) = 1$. Since $y = \pm 1$ corresponds to $x = \pm\infty$, this means that these are <u>bounded</u> solutions to the Schrödinger equation (tending to 1 or maybe -1 as $x \to \pm\infty$) but they are not <u>bound states</u> (for which ψ would have to tend to zero as $x \to \pm\infty$).

(The second solutions, the *Legendre functions of the second kind*, $Q_n(y)$, blow up at $y = \pm 1$.)

Fact 2:

If $n \in \mathbb{Z}_{\geqslant 0}$, bounded solutions to (9.47) only exist for

$$\boxed{m = 0, 1, 2 \dots n} \tag{9.49}$$

and are

$$P_n^m(y) = (-1)^m(1-y^2)^{m/2}\frac{d^m}{dy^m}P_n(y)\,. \tag{9.50}$$

These are the **associated Legendre ‘polynomials’ of the first kind** (the word polynomials is in quotes since for m odd, $m/2$ is not an integer so they aren’t strictly speaking polynomials).

Fact 3:

Even when m and n are not integers (and in fact even when they are complex), solutions to (9.47) can be written explicitly using certain special functions. We have that

$$\boxed{P_n^m(y) = \frac{1}{\Gamma(1-m)}\left(\frac{1+y}{1-y}\right)^{m/2} {}_2F_1\big(-n, n+1; 1-m; \tfrac{1-y}{2}\big)} \tag{9.51}$$

solves (9.47), and reduces to (9.50) if $n \in \mathbb{Z}_{\geqslant 0}$ and $m = 0, 1, \ldots n$.

Here

$$\boxed{\Gamma(z) = \int_0^\infty dt\, t^{z-1} e^{-t}} \tag{9.52}$$

is **Euler's Gamma function**, which satisfies

$$\Gamma(z+1) = z\Gamma(z)\;, \tag{9.53}$$

which along with $\Gamma(1) = 1$ implies that

$$\Gamma(N+1) = N! \;\text{ if } N \in \mathbb{Z}_{\geqslant 0}\;. \tag{9.54}$$

Other key properties are that

$$\begin{aligned} \Gamma(z) &\neq 0 \;\; \forall z \\ \frac{1}{\Gamma(z)} &= 0 \;\text{ iff } z \in \{0, -1, -2, \ldots\} \\ \Gamma(z)\Gamma(1-z) &= \frac{\pi}{\sin(\pi z)}\;. \end{aligned} \tag{9.55}$$

${}_2F_1$ is the **hypergeometric function** and has the Taylor expansion

$$\boxed{{}_2F_1(a, b; c; z) = \frac{\Gamma(c)}{\Gamma(a)\Gamma(b)} \sum_{k=0}^{\infty} \frac{\Gamma(k+a)\Gamma(k+b)}{\Gamma(k+c)} \frac{z^k}{k!}} \tag{9.56}$$

for $|z| < 1$, and is defined by analytic continuation elsewhere. The first few terms are

$${}_2F_1(a, b; c; z) = 1 + \frac{ab}{c} z + \frac{a(a+1)b(b+1)}{c(c+1)} \frac{z^2}{2!} + \ldots\;.$$

So, up to normalisation, a potentially bounded solution to (9.45) is

$$\boxed{\psi(x) = P_n^m(y = \tanh(x))} \tag{9.57}$$

with

$$\boxed{m = ik\,, \quad n = \frac{\sqrt{1+4a}}{2} - \frac{1}{2}}\;. \tag{9.58}$$

Note that in (9.58) we picked the roots which give the scattering solution with particles incident from the left. Observe also that n is real (and $n \geqslant -\frac{1}{2}$) if $a \geqslant -\frac{1}{4}$, and $n \geqslant 0$ if $a \geqslant 0$.

(a) $k^2 > 0$ – the continuous spectrum

- $\underline{x \to +\infty:}$ In this limit $y = \tanh(x) \approx 1 - 2e^{-2x} \to 1^-$ and so

$$ {}_2F_1(\dots;\tfrac{1-y}{2}) \to {}_2F_1(\dots;0) = 1\,; \qquad \frac{1+y}{1-y} \approx e^{2x}\,. $$

Putting these bits together,

$$ \boxed{\psi \approx \frac{1}{\Gamma(1-ik)}\,e^{ikx}} \tag{9.59} $$

as $x \to +\infty$.

- $\underline{x \to -\infty:}$ In this limit $y = \tanh(x) \approx -1 + 2e^{2x} \to -1^+$ and $\frac{1+y}{1-y} \approx e^{2x}$, and it turns out that

$$ \frac{1}{\Gamma(1{-}m)}\,{}_2F_1(-n,n{+}1;1{-}m;\tfrac{1-y}{2}) \approx \frac{\Gamma(-m)}{\Gamma(1{-}m{+}n)\Gamma(-m{-}n)} + \frac{\Gamma(m)}{\Gamma(-n)\Gamma(n{+}1)}\,e^{-2mx}\,. $$

This asymptotic can be proved using the already-mentioned properties of the hypergeometric function together with the identity

$$ \begin{aligned} \frac{\sin(\pi(c{-}a{-}b))}{\pi}\,{}_2F_1(a,b;c;z) = {}& \frac{{}_2F_1(a,b;c;1{-}z)}{\Gamma(c{-}a)\Gamma(c{-}b)\Gamma(a{+}b{-}c{+}1)} \\ & - (1{-}z)^{c-a-b}\,\frac{{}_2F_1(c{-}a,c{-}b;c{-}a{-}b{+}1;1{-}z)}{\Gamma(a)\Gamma(b)\Gamma(c{-}a{-}b{+}1)}\,. \end{aligned} $$

Hence

$$ \boxed{\psi \approx \frac{\Gamma(-ik)}{\Gamma(1{-}ik{+}n)\Gamma(-ik{-}n)}\,e^{ikx} + \frac{\Gamma(ik)}{\Gamma(-n)\Gamma(n{+}1)}\,e^{-ikx}} \tag{9.60} $$

as $x \to -\infty$.

Normalising this scattering solution so that the coefficient of e^{ikx} at $-\infty$ is 1, we can read off the values of $R(k)$ and $T(k)$:

$$ \boxed{\begin{aligned} R(k) &= \frac{\Gamma(ik)\Gamma(1{-}ik{+}n)\Gamma(-ik{-}n)}{\Gamma(-ik)\Gamma(1{+}n)\Gamma(-n)} = -\frac{\sin(\pi n)}{\pi}\,\frac{\Gamma(ik)\Gamma(1{-}ik{+}n)\Gamma(-ik{-}n)}{\Gamma(-ik)} \\ T(k) &= \frac{\Gamma(1{-}ik{+}n)\Gamma(-ik{-}n)}{\Gamma(1{-}ik)\Gamma(-ik)}\,. \end{aligned}} \tag{9.61} $$

Note: The $\sin(\pi n)$ factor in $R(k)$ means that it vanishes for all k if n is an integer. The corresponding potentials

$$ \boxed{V(x) = -n(n{+}1)\,\mathrm{sech}^2(x)} \tag{9.62} $$

with $n \in \mathbb{Z}_{\geqslant 0}$ (without loss of generality since a is unchanged if we flip sign to $n + 1/2$) are called **reflectionless**: no particles are reflected for **any** value of k.

(b) $k^2 < 0$ **– the discrete spectrum**

To find the discrete spectrum, set $k = i\mu$, where $\mu > 0$, and divide the scattering solution through by $T(i\mu)$ to find a possible eigenfunction

$$\boxed{\psi(x) \approx \begin{cases} \frac{1}{T(i\mu)} e^{-\mu x} + \frac{R(i\mu)}{T(i\mu)} e^{\mu x} & x \to -\infty \\ e^{-\mu x} & x \to +\infty \end{cases}} \tag{9.63}$$

This is automatically bounded as $x \to +\infty$; it will be bounded as $x \to -\infty$ if (and only if) $\mu \geqslant 0$ is such that $1/T(i\mu) = 0$. (In fact we'll require $\mu > 0$, since $\int_{-\infty}^{+\infty} |\psi|^2 dx$ should be finite for the discrete spectrum.) This in turn requires

$$\frac{1}{T(i\mu)} = \frac{\Gamma(1+\mu)\Gamma(\mu)}{\Gamma(1+\mu+n)\Gamma(\mu-n)} = 0\,.$$

Given that μ must be a positive real number and that $\Gamma(z)$ has no zeros, there are two options:

(1) $1 + \mu + n = -j$, with $j \in \mathbb{Z}_{\geqslant 0}$;

(2) $\mu - n = -h$, with $h \in \mathbb{Z}_{\geqslant 0}$.

- If $n \notin \mathbb{R}$ then there are no real solutions for μ;
- if $n \in \mathbb{R}$ we can take $n \geqslant -1/2$ without losing generality, since **(1)** ↔ **(2)** when $n \mapsto -1-n$.

Then **(1)** never holds, while solutions for positive μ do exist for option **(2)** provided $n \geqslant 0$:

$$\boxed{\mu = n, n-1, n-2 \ldots n - \lfloor n \rfloor} \tag{9.64}$$

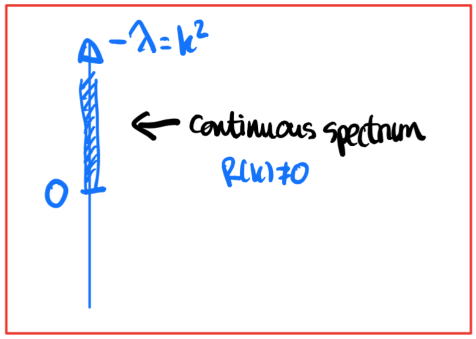
where $\lfloor n \rfloor$ = 'floor' of n = {largest integer $\leqslant n$}. So

$$\boxed{\text{Total number of bound states} = \lceil n \rceil} \tag{9.65}$$

where $\lceil n \rceil$ = 'ceiling' of n = {smallest integer $\geqslant n$}. (If n is an integer, then the last eigenvalue, for $\mu = 0$, should be discarded as the corresponding ψ is not square integrable and so is not a bound state. It's in the continuous spectrum instead.)

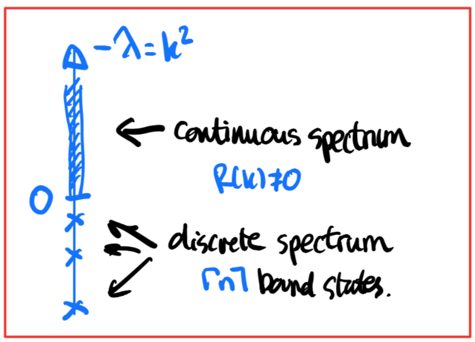
Summary for $V(x) = -a\,\mathrm{sech}^2(x) = -n(n+1)\,\mathrm{sech}^2(x)$**:**

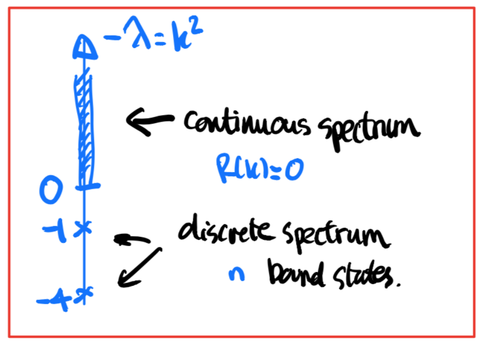
- $a < 0$:



- $a = n(n+1) > 0$:

(n not an integer (say $n = 2.5$) on the left, $n \in \mathbb{Z}_{>0}$ (say $n = 2$) on the right)





Chapter 10

Time evolution of the scattering data

See section 5.2 [Drazin and Johnson, 1989] and [Aktosun, 2009] for this chapter.

10.1 Scattering data for general potentials

So far we've seen that for any localised initial data $u(x,0)$ for the KdV equation, the auxiliary time-independent Schrödinger equation

$$-\psi''(x) + V(x)\psi(x) = k^2\psi(x) \tag{10.1}$$

with potential $V(x) = -u(x,0)$ has

1. A *continuous spectrum* of non-negative eigenvalues $E = k^2 \geqslant 0$ and eigenfunctions

$$\boxed{\psi(x) \approx \begin{cases} e^{ikx} + R(k)\, e^{-ikx}\,, & x \to -\infty \\ T(k)\, e^{ikx}\,, & x \to +\infty \end{cases}} \tag{10.2}$$

   normalised so that the incoming flux is one;

2. A (maybe empty) *discrete spectrum* of negative eigenvalues $E = k^2 = -\mu_n^2 < 0$, indexed by $n = 1, 2 \ldots N$. These look like

$$\boxed{\psi_n(x) \approx \begin{cases} c_n\, e^{\mu_n x}\,, & x \to -\infty \\ d_n\, e^{-\mu_n x}\,, & x \to +\infty \end{cases}} \tag{10.3}$$

   So far the ψ_n's we have found have been normalised so that $d_n = 1$, but now we will instead normalise them so that

$$\boxed{\langle \psi_n, \psi_n \rangle = \int_{-\infty}^{+\infty} |\psi_n(x)|^2 \, dx = 1} \, .$$

Once ψ_n has been normalised in this way, the number c_n is called the **normalising coefficient** and it will be needed later to reconstruct $V(x) = -u(x)$. More precisely, to reconstruct $V(x)$ we will need to know the eigenvalues and the asymptotics of the eigenfunctions as $x \to -\infty$:

$$\boxed{S = \Big\{ R(k) \, , \quad \{\mu_n, c_n\}_{n=1}^{N} \Big\}} \tag{10.4}$$

This is called the **scattering data**, refining the notion of scattering data given earlier.

- Clearly, u (or $V = -u$) determines the scattering data completely (this was step **(a)**, disassembly, of the roadmap).

- Amazingly, the converse also holds: u (or $V = -u$) can be reconstructed uniquely from the scattering data (step **(c)**, reassembly).

- The next major task is to return to step **(b)**, time evolution, to see precisely how the scattering data evolves.

Before going there, let's make precise the scattering data for two sets of potentials studied earlier.

10.2 Examples of scattering data

1) $\underline{V(x) = a\,\delta(x)}$:

For all values of a we have

$$\boxed{R(k) = \frac{a}{2ik - a}}\ .$$

• For $a \geqslant 0$ that's all.

• For $a < 0$ there is also a single bound state $\psi(x) = Ae^{-\mu|x|}$ with $\mu = -a/2 > 0$. Normalising determines $A^2/\mu = 1$ so $A = \sqrt{\mu} = \sqrt{-a/2}$, up to an irrelevant sign ambiguity.

Thus the general scattering data for $u(x,0) = -a\delta(x)$, $V(x) = a\delta(x)$, is

$$\boxed{S(0) = \begin{cases} \left\{ R(k) = \frac{a}{2ik-a} \right\} & \text{if } a \geqslant 0 \\ \left\{ R(k) = \frac{a}{2ik-a}\ ,\ \left\{ \mu_1 = -a/2, c_1 = \sqrt{-a/2} \right\} \right\} & \text{if } a < 0 \end{cases}} \tag{10.5}$$

2) $\underline{V(x) = -n(n{+}1)\,\mathrm{sech}^2(x)}$, $n \in \mathbb{Z}_{\geqslant 0}$:

(a) Scattering states: $R(k) = 0$, since the potential is reflectionless.

(b) Bound states: we have $\psi_m(x) = A\,P_n^m(\tanh(x))$, $m = 1, 2 \ldots n$, where A is a normalisation constant that can be fixed by imposing

$$1 = \int_{-\infty}^{+\infty} |\psi_m(x)|^2\,dx = A^2 \int_{-1}^{1} P_n^m(y)^2\,\frac{dy}{1-y^2} = A^2\,\frac{(n{+}m)!}{m(n{-}m)!}$$

where the last equality makes use of one of the standard properties of P_n^m.

In addition P has the asymptotics

$$P_n^m(\tanh(x)) \approx (-1)^n \frac{(n{+}m)!}{m!(n{-}m)!}\,e^{mx}\,, \quad x \to -\infty\,.$$

Hence the asymptotics of the normalised bound state is

$$\psi_m(x) \approx (-1)^n \frac{1}{m!}\sqrt{\frac{m(n{+}m)!}{(n{-}m)!}}\,e^{mx}\,, \quad x \to -\infty\,.$$

The full scattering data is

$$\boxed{S(0) = \left\{ R(k) = 0\,,\ \left\{ \mu_m^{(n)} = m\,,\ c_m^{(n)} = (-1)^n \tfrac{1}{m!}\sqrt{\tfrac{m(n+m)!}{(n-m)!}} \right\}_{m=1}^{n} \right\}}\ . \tag{10.6}$$

3) $\underline{V(x) = -n'(n'+1)\,\mathrm{sech}^2(x)}$, $n' = n + \epsilon$, $n \in \mathbb{Z}_{\geqslant 0}$, $|\epsilon| \ll 1$: This is a small perturbation of the previous case, hence the potential is almost reflectionless.

The discrete eigenvalues are

$$\mu = (\epsilon,)\, 1 + \epsilon, 2 + \epsilon, \ldots, n + \epsilon\ .$$

(The first eigenvalue $\mu = \epsilon$ is there only if $\epsilon > 0$. Let's assume that $\epsilon < 0$ and not worry about it.)

Compared to the previous case, we just need to replace factorials by gamma functions:

$$\Gamma(n') = \Gamma(n)\left(1 + \epsilon\psi(n) + (\epsilon^2)\right) \tag{10.7}$$

where

$$\boxed{\psi(z) := \frac{\Gamma'(z)}{\Gamma(z)}} \tag{10.8}$$

is the *digamma function*. Expanding to first order in ϵ one finds

$$\boxed{c_m^{(n')} approx c_m^{(n)}\left[1 + \epsilon\left(\frac{1}{m} + \frac{1}{2}\psi(n+m+1) - \psi(m+1)\right)\right]} \tag{10.9}$$

for the discrete eigenvalues

$$\boxed{\mu_m^{(n')} = m + \epsilon}\,, \qquad m = 1, \ldots, n. \tag{10.10}$$

The reflection coefficients is easier and more interesting. From (9.61) we obtain

$$\boxed{\begin{aligned} R(k) &= -\frac{\sin(\pi(n+\epsilon)}{\pi}\,\frac{\Gamma(ik)\Gamma(1-ik+n+\epsilon)\Gamma(-ik-n-\epsilon)}{\Gamma(-ik)} \\ &\approx \epsilon\cdot(-1)^{1+n}\frac{\Gamma(ik)\Gamma(1-ik+n)\Gamma(-ik-n)}{\Gamma(-ik)} \end{aligned}}\,. \tag{10.11}$$

Equations (**??**) are the scattering data for $n' = n + \epsilon$ with $0 < -\epsilon \ll 1$.

Now that we have worked out the scattering data for some initial values of the KdV field, we'd like to understand how to evolve the scattering data forward in (KdV) time t, when $u = -V$ evolves according to the KdV equation $u_t + 6uu_x + u_{xxx} = 0$.

10.3 The idea of a Lax pair

We want to solve the initial value problem for a PDE

$$\boxed{u_t = N(u)} \tag{10.12}$$

where $N(u)$ is a function of $u, u_x, u_{xx}, \ldots$ (but no t derivatives), and with the boundary conditions $u, u_x, u_{xx}, \cdots \to 0$ as $x \to \pm\infty$. For the KdV equation $N(u) = -6uu_x - u_{xxx}$, but we can be more general.

We'll think of $\psi_{xx} + u\psi = \lambda\psi$ at some fixed time t as an **eigenvalue problem**:

$$L(u)\psi = \lambda\psi \tag{10.13}$$

where $L(u)$ is the following differential operator:

$$L(u) = \frac{\partial^2}{\partial x^2} + u(x,t)\,. \tag{10.14}$$

Notes:

1. You should think of differential operators such as L, or $\partial/\partial x$, or whatever, as acting on all functions sitting to their right.

2. (10.13) *does* pick out "special" values of λ (the *eigenvalues*) since we require that $\psi(x)$ is square integrable (*i.e.* $\int_{-\infty}^{+\infty} |\psi(x)|^2 dx < \infty$) which in particular means $\psi(x) \to 0$ both as $x \to -\infty$ and as $x \to +\infty$. (Later, we will relax this a little to allow solutions ψ that are merely *bounded*, but for now we will require that the stronger condition holds.)

3. The "t" in (10.14) has nothing to do with the time in the time-dependent Schrödinger equation you might see in quantum mechanics; rather, it's the KdV time.

Since L depends on u, and u depends on t, the *eigenfunctions* ψ and (in principle) the *eigenvalues* λ might be different at different (KdV) times.

But, we have two remarkable facts:

THEOREM:
(i) If $u = u(x,t)$ evolves by the KdV equation, then the set of eigenvalues $\{\lambda\}$ of $L(u)$ (the **spectrum** of $L(u)$) is independent of t;
(ii) There is a set of eigenfunctions ψ of $L(u)$ which evolves in t simply, as

$$\psi_t = M(u)\psi\ , \tag{10.15}$$

where $M(u)$ is another differential operator.

The result (i) is particularly striking – it says that the spectra of $\partial^2/\partial x^2 + u(x,0)$ and $\partial^2/\partial x^2 + u(x,t)$ are the same, which is very unexpected since $u(x,0)$ and $u(x,t)$ might look very different.

PROOF:

First, we'll assume that an operator $M(u)$ can be found such that the time evolution of $L(u(x,t))$ is given by

$$\begin{aligned} L(u)_t &= M(u)L(u) - L(u)M(u) \\ &= [M(u), L(u)] \end{aligned} \tag{10.16}$$

when u evolves by KdV (we'll find M later).

Here, $[M,L] := ML - LM$ is called the **commutator** of the two operators M and L. Since M and L can both involve x derivatives, $ML \neq LM$ is possible – see later for examples.

Now let λ and ψ be an eigenvalue and eigenfunction of L, so that $L\psi = \lambda\psi$. Taking $\partial/\partial t$ of this equation,

$$L_t\psi + L\psi_t = \lambda_t\psi + \lambda\psi_t \, .$$

Rearranging,

$$\begin{aligned} \lambda_t\psi &= \lambda_t\psi + L\psi_t - \lambda\psi_t \\ &= (ML - LM)\psi + (L-\lambda)\psi_t \qquad \text{(using (10.16))} \\ &= (M\lambda - LM)\psi + (L-\lambda)\psi_t \qquad \text{(using } L\psi = \lambda\psi\text{)} \\ &= (L-\lambda)\big(\psi_t - M\psi\big) \, . \end{aligned} \tag{10.17}$$

Now consider the **inner product** on square integrable (complex) functions of x:

$$\boxed{\langle\psi_1,\psi_2\rangle := \int_{-\infty}^{+\infty} dx\, \overline{\psi_1(x)}\psi_2(x)} \,, \qquad \psi_1,\psi_2 \in L^2(\mathbb{R}) \, . \tag{10.18}$$

Lemma: L is **self-adjoint**, *i.e.*

$$\boxed{\langle\psi_1, L\psi_2\rangle = \langle L\psi_1,\psi_2\rangle} \qquad \forall\psi_1,\psi_2 \in L^2(\mathbb{R}) \, . \tag{10.19}$$

Proof of the lemma: integrate by parts to find

$$\begin{aligned} \langle\psi_1, L\psi_2\rangle &= \int_{-\infty}^{+\infty} dx\, \overline{\psi_1(x)} \left(\frac{\partial^2}{\partial x^2} + u(x,t)\right)\psi_2(x) \\ &= \int_{-\infty}^{+\infty} dx\, \overline{\left[\left(\frac{\partial^2}{\partial x^2} + u(x,t)\right)\psi_1(x)\right]}\,\psi_2(x) = \langle L\psi_1,\psi_2\rangle \, . \end{aligned}$$

where we used that the boundary terms vanish because $\psi_1, \psi_1' \to 0$ as $x \to \pm\infty$, and that $u \in \mathbb{R}$. $\square$

Using this lemma and the fact that the eigenvalues λ of a self-adjoint operator (like L) are real (proof left as an exercise), (10.17) implies

$$\begin{aligned}\lambda_t\langle\psi,\psi\rangle &= \langle\psi,(L-\lambda)(\psi_t - M\psi)\rangle \\ &= \langle(L-\lambda)\psi,(\psi_t - M\psi)\rangle = 0\ ,\end{aligned}$$

where in the last equality we used that $L\psi = \lambda\psi$. Since $0 < \langle\psi,\psi\rangle < \infty$, we deduce that

$$\boxed{\lambda_t = 0}\ , \tag{10.20}$$

which is result (i).[1] $\square$

For result (ii), we need to show that $(L-\lambda)\psi = 0$ continues to be true if ψ changes according to $\psi_t = M\psi$. Calculating,

$$\begin{aligned}\frac{\partial}{\partial t}\big((L-\lambda)\psi\big) &= L_t\psi + L\psi_t - \lambda_t\psi - \lambda\psi_t \\ &= L_t\psi + L\psi_t - \lambda\psi_t \qquad \text{(since we already know } \lambda_t = 0\text{)} \\ &= L_t\psi + LM\psi - \lambda M\psi \qquad \text{(using } \psi_t = M\psi\text{)} \\ &= L_t\psi + LM\psi - M\lambda\psi \qquad \text{(since } \lambda \text{ is a number)} \\ &= L_t\psi + LM\psi - ML\psi \qquad \text{(using } L\psi = \lambda\psi\text{)} \\ &= \big(L_t - [M,L]\big)\psi \\ &= 0 \qquad \text{(using (10.16))}\end{aligned} \tag{10.21}$$

This shows that if $\psi_t = M\psi$ and ψ starts off as an eigenfunction at $t = 0$, then it stays that way:

$$(L-\lambda)\psi = (\text{constant with respect to } t) = (L-\lambda)\psi\big|_{t=0} = 0\ . \tag{10.22}$$

This is result (ii).[2] $\square$

L and M are called a **Lax pair**. All that remains now is to find a suitable $M(u)$.

---

[1] Note that this derivation holds for the discrete spectrum, so that the inner product is finite. But we already know that the continuous spectrum consists of all $\lambda \leqslant 0$ for all t, so that part is trivially constant.

[2] We may also start from (10.17), which tells us that $\lambda_t = 0$ (part (i) of the theorem) implies

$$(L-\lambda)(\psi_t - M\psi) = 0\ .$$

Hence $\psi_t - M\psi$ is in the kernel of the operator $L - \lambda$ (at any given time t). If λ belongs to the discrete spectrum of L, there is a unique bound state ψ, which means that $\ker(L-\lambda) = \mathbb{R}\,\psi$ is one-dimensional and is generated by ψ. Therefore

$$\psi_t - M\psi = A\psi$$

for a constant A, or equivalently $e^{At}(\partial_t - M)(e^{-At}\psi) = 0$. If $A = 0$, (ii) follows for $\psi(t)$. If $A \neq 0$, the statement does not follow for $\psi(t)$, but we can construct $\tilde{\psi}(t) := e^{-At}\psi(t)$, which is an eigenfunction of L with

10.3.1 The Lax pair for KdV

We've already decided that $L(u) = \frac{\partial^2}{\partial x^2} + u(x,t)$, where u evolves according to the KdV equation. We want to find $M(u)$ such that

$$u_t = N(u) \equiv -6uu_x - u_{xxx} = 0 \iff L(u)_t = [M(u), L(u)] \,. \tag{10.23}$$

Here $L(u)_t$ denotes the time derivative of the operator $L(u) = \frac{\partial^2}{\partial x^2} + u$. Since the operator $\frac{\partial^2}{\partial x^2}$ does not depend on time, we have $L(u)_t = u_t$. Hence we need

$$[M(u), L(u)] = N(u) \equiv -6uu_x - u_{xxx} \,. \tag{10.24}$$

For now we'll just make an inspired guess for $M(u)$, and check that it works; in the next chapter a more systematic approach will be explained. The guess is to set

$$\boxed{M(u) = -\left(4D^3 + 6uD + 3u_x\right)} \tag{10.25}$$

where to save ink the notation $D \equiv \frac{\partial}{\partial x}$, $D^2 \equiv \frac{\partial^2}{\partial x^2}$, *etc.* has been adopted.

Notice that operators like D act on *everything* to their right, and that differential operators are defined by their actions on functions. So for example $[D, u]$ is defined by how it would act on any (say smooth) function $f(x)$. Calculating,

$$\begin{aligned} [D,u]f &= (Duf - uD)\,f \\ &= D(uf) - u(Df) = (Du)f + u(Df) - u(Df) = u_x f \,. \end{aligned}$$

Thus the effect of $[D, u]$ on $f(x)$ is to multiply it by $u_x(x)$. Since this is true for all functions $f(x)$ we have that

$$[D, u] = u_x \tag{10.26}$$

as an identity between differential operators. Perhaps more usefully, this can be rephrased as

$$Du = uD + u_x$$

which shows how to "shuffle" Ds past other functions. This can be used to rewrite expressions in a form where all Ds are on the right in all terms, making cancellations easier to spot. More generally (see problem 58)

$$\boxed{[D^n, u] = \sum_{m=0}^{n-1} \binom{n}{m} u_{\underbrace{x\ldots x}_{n-m\ \text{times}}} D^m} \,. \tag{10.27}$$

eigenvalue λ for all t, coincides with ψ at $t = 0$, and satisfies the time evolution equation $\tilde{\psi}_t = M\tilde{\psi}$. This proves (ii) for $\tilde{\psi}$. For the continuous part of the spectrum there are two linearly independent eigenfunctions for any eigenvalue $\lambda = -k^2 < 0$, but only one satisfies the boundary condition of a scattering solution with unit flux of right-moving plane waves coming in from $x = -\infty$, hence the same argument applies.

We will also need $[D^n, D^m] = 0$ for all n, m, and $[g(x), h(x)] = 0$ for all functions g and h. Finally, for all operators A, B, C we have

$$\begin{aligned} [A, BC] &= ABC - BCA \\ &= ABC - BAC + BAC - BCA \\ &= [A, B]C + B[A, C] , \end{aligned} \tag{10.28}$$

and similarly

$$[AB, C] = A[B, C] + [A, C]B . \tag{10.29}$$

Now just calculate! We have

$$L = D^2 + u , \quad M = -(4D^3 + 6uD + 3u_x)$$

so

$$\begin{aligned} -[M(u), L(u)] &= [4D^3 + 6uD + 3u_x, D^2 + u] \\ &= 4[D^3, u] + 6[uD, D^2] + 6[uD, u] + 3[u_x, D^2] \\ &= 4[D^3, u] + 6[u, D^2]D + 6u[D, u] + 3[u_x, D^2] \\ &= 4(u_{xxx} + 3u_x xD + 3u_x D^2) - 6(u_{xx} + 2u_x D)D \\ &+ 6uu_x - 3(u_{xxx} + 2u_{xx}D) \\ &= u_{xxx} + 6uu_x , \end{aligned} \tag{10.30}$$

and somewhat surprisingly all of the Ds have gone, reproducing (10.24) as promised. This completes the proof that

$$\boxed{\text{KdV for } u} \iff \boxed{L_t = [M, L]}$$

with the Lax pair

$$\boxed{\begin{aligned} L &= D^2 + u \\ M &= -(4D^3 + 6uD + 3u_x) \end{aligned}} \tag{10.31}$$

Notes:

1. L and M were both **differential** operators, since they involved $D = \frac{d}{dx}$, but in some senses $[L, M]$ wasn't: $[L, M]$ acting on some function $f(x)$ doesn't do any differentiating, but just multiplies f pointwise by $(u_{xxx} + 6uu_x)$. For this reason the operator $[L, M]$ is called **multiplicative**.

2. The equation for the time evolution of ψ, $\psi_t = M(u)\psi$, is linear (good news!), but since M depends on $u(x, t)$, the thing we're trying to find, it's not yet clear we have made too much progress on step (b) (bad news). We will fix this later, once we have developed a better understanding of the scattering data.

10.4 Time evolution of the scattering data

We have seen that if u evolves by the KdV equation, then:

1. the eigenvalues λ of $L(u) = D^2 + u$ remain constant in t;

2. the eigenfunctions ψ evolve by $\psi_t = M(u)\psi$.

Question: how does the scattering data associated to $V = -u$ evolve in time?

Answer: We need to look at the asymptotics of the time-evolution equation $\psi_t = M(u)\psi$ as $x \to \pm\infty$. Recall that for KdV

$$M(u) = -(4D^3 + 6uD + 3u_x)$$

and so, since u, $u_x \to 0$ as $x \to \pm\infty$ for all t, as follows from the boundary conditions on u,

$$\boxed{M(u) \sim -4D^3 \quad \text{as } x \to \pm\infty} \tag{10.32}$$

and is **independent of** $u(x,t)$. This is the **key point**: we can evolve the scattering data forward in t without knowing in advance what u evolves to!

[You might worry about the bound state normalisation condition $(\psi_m, \psi_m) = 1$: is this preserved under time evolution? It turns out that the answer is yes: this follows, with a little work, from the antisymmetry of B, that is $M(u)^\dagger = -M(u)$. See problem 61.]

Next, we need to work out explicitly the t evolution of the asymptotics of the scattering and bound state solutions.

(a) The continuous spectrum ($-\lambda = k^2 > 0$)

Start with an un-normalised scattering solution:

$$\boxed{\psi_k(x;t) \approx \begin{cases} A(k;t)\, e^{ikx} + B(k;t)\, e^{-ikx}\ , & x \to -\infty \\ C(k;t)\, e^{ikx}\ , & x \to +\infty \end{cases}} \tag{10.33}$$

Imposing

$$\frac{\partial}{\partial t}\psi_k(x;t) = M(u)\psi_k(x;t) \sim -4D^3\psi_k(x;t)$$

as $x \to \pm\infty$, we have

$$A_t(k;t)\, e^{ikx} + B_t(k;t)\, e^{-ikx} = 4ik^3 \left[A(k;t)\, e^{ikx} - B(k;t)\, e^{-ikx}\right]$$
$$C_t(k;t)\, e^{ikx} = 4ik^3 C(k;t)\, e^{ikx}$$

and, hence, equating coefficients of $e^{\pm ikx}$,

$$\begin{aligned} A_t(k;t) &= 4ik^3 A(k;t) \\ B_t(k;t) &= -4ik^3 B(k;t) \\ C_t(k;t) &= 4ik^3 C(k;t) \end{aligned} \tag{10.34}$$

Solving,

$$\begin{aligned} A(k;t) &= A(k;0)\, e^{4ik^3 t} \\ B(k;t) &= B(k;0)\, e^{-4ik^3 t} \\ C(k;t) &= C(k;0)\, e^{4ik^3 t}\,. \end{aligned} \tag{10.35}$$

Dividing the un-normalised solution at time t through by $A(k;t)$ so that it continues to be correctly normalised with unit incoming flux, $R(k;t)$ and $T(k;t)$ can be read off as follows:

$$\begin{aligned} R(k;t) &= R(k;0)\, e^{-8ik^3 t} \\ T(k;t) &= T(k;0)\,. \end{aligned} \tag{10.36}$$

This can be summed up in the asymptotics of the normalised scattering solution:

$$\boxed{\psi_k(x;t) \approx \begin{cases} e^{ikx} + R(k;0)\, e^{-ik(x+8k^2 t)} & x \to -\infty \\ T(k;0)\, e^{ikx} & x \to +\infty \end{cases}} \tag{10.37}$$

As we will see later, the reflected waves for ψ_k, encoded in $R(k;t)$, translate into a **dispersive component** of $u(x,t)$, moving to the left as t increases.

(b) The discrete spectrum $(-\lambda = -\mu_n^2 < 0)$

The n^{th} bound state wave function has asymptotics

$$\psi_n(x;t) \approx \begin{cases} c_n(t)\, e^{\mu_n x}\,, & x \to -\infty \\ d_n(t)\, e^{-\mu_n x}\,, & x \to +\infty \end{cases} \tag{10.38}$$

Imposing

$$\frac{\partial}{\partial t}\psi_n(x;t) = M(u)\psi_k(x;t) \approx -4D^3\psi_n(x;t)$$

as $x \to \pm\infty$, we have

$$\frac{\partial}{\partial t}\, c_n(t) = -4\mu_n^3\, c_n(t)$$

$$\frac{\partial}{\partial t}\, d_n(t) = +4\mu_n^3\, d_n(t)$$

and, solving,

$$\boxed{\begin{aligned} c_n(t) &= c_n(0)\, e^{-4\mu_n^3 t} \\ d_n(t) &= d_n(0)\, e^{+4\mu_n^3 t} \end{aligned}} \tag{10.39}$$

Again, this can be summarised as

$$\psi_n(x;t) \approx \begin{cases} c_n(0)\, e^{\mu_n(x-4\mu_n^2 t)}\,, & x \to -\infty \\ d_n(0)\, e^{-\mu_n(x-4\mu_n^2 t)}\,, & x \to +\infty \end{cases} \tag{10.40}$$

This will translate into a soliton for $u(x,t)$, moving to the right with velocity $4\mu_n^2$.

These results describe the time evolution of the scattering data, completing step **(b)** of the inverse scattering method.

Chapter 11

Interlude: the KdV hierarchy and conservation laws

11.1 Deriving the KdV equation (and generalising it)

It's natural to ask whether there are any other evolution equations for $u(x, t)$ such that the eigenvalues of

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

(acting on bounded functions of x) are constant in t. In more fancy language, we're looking for equations such that the $L(u)$'s at different times are **isospectral**; these are called **isospectral flows**.

The answer is yes, there are more such equations, and the Lax pair idea allows us to find them.

Key point: the proof in section 10.3 only used the equivalence

$$\boxed{u_t = N(u)} \iff \boxed{L(u)_t = [M(u), L(u)]}\,. \tag{11.1}$$

No other details of M were needed, so some other $M(u)$ should work just as well (leading to other functions $N(u)$ of $u, u_x, u_{xx}, \ldots$ in the KdV-like equation on the left). However, $M(u)$ is not completely arbitrary: since $L_t = u_t$, and is a multiplicative operator, $[M, L]$ must also be multiplicative. This means all the D's must cancel out when computing the commutator. If they *do* cancel, what's left in $[M, L]$ will be a polynomial in u, u_x, u_{xx} etc, and setting this equal to u_t will give us the desired evolution equation.

Which other conditions, if any, should $M(u)$ satisfy? To answer that, let's remind ourselves of some technology.

(i) (Hermitian) inner product

For two functions $\phi(x)$ and $\chi(x)$, we define

$$\langle\phi,\chi\rangle = \int_{-\infty}^{+\infty} dx\,\overline{\phi(x)}\chi(x)\,dx\,.$$

(The complex conjugation on the first term ensures $(\phi,\phi) > 0$ for $\phi \neq 0$ even when ϕ is complex.)

In this notation, the key property of $L = D^2 + u$ used in the Lax proof was that

$$\langle\phi, L\chi\rangle = \langle L\phi,\chi\rangle$$

for all ϕ and χ.

(ii) The adjoint of an operator

If A is a differential operator, define $A^\dagger$ ("A dagger") to be the operator such that

$$\langle\phi, A\chi\rangle = \langle A^\dagger\phi,\chi\rangle \tag{11.2}$$

for all ϕ and χ. $A^\dagger$ is called the **adjoint** of A; it's a bit like a matrix transpose and, like the matrix transpose, satisfies

$$(A^\dagger)^\dagger = A\,,\quad (AB)^\dagger = B^\dagger A^\dagger\,,\quad [A,B]^\dagger = [B^\dagger, A^\dagger]$$

(exercise: check!). The key property of $L(u)$ (acting on the space L^2 of square integrable functions) was

$$\boxed{L(u)^\dagger = L(u)} \tag{11.3}$$

and such operators are called **self-adjoint** (**/symmetric/hermitian**). Other important operators have

$$A^\dagger = -A$$

and are called **skew-adjoint** (**/antisymmetric/antihermitian**).

Now if A is just multiplication by a real function, then $A^\dagger = A$. (Exercise: why?) This must be true of $[L,M]$ as it is supposed to be a (real) multiplicative operator, so M must be such that that $[L,M]^\dagger = [L,M]$.

What can we deduce about M from this? $[L, M] = [L, M]^\dagger$ and $L = L^\dagger$ imply that

$$\begin{aligned} [L, M] &= [L, M]^\dagger \\ &= [M^\dagger, L^\dagger] \\ &= [M^\dagger, L] \\ &= -[L, M^\dagger] \\ \Rightarrow &\boxed{[L, M + M^\dagger] = 0} \end{aligned} \tag{11.4}$$

Otherwise stated, the symmetric/hermitian part of M must commute with L. (As with matrices, any M can be written as

$$M = \frac{1}{2}(M + M^\dagger) + \frac{1}{2}(M - M^\dagger)$$

where the first term is the symmetric (/hermitian) part of M, and the second the antisymmetric (/antihermitian) part.)

Since it's only the bit of M which doesn't commute with L that makes a difference to the equation $L_t + [L, M] = 0$, this means that M can be assumed to be antisymmetric (/antihermitian):

$$\boxed{M(u)^\dagger = M(u)} \,. \tag{11.5}$$

Note: this guarantees that $\langle \psi, \psi \rangle$ is constant under time evolution $\psi_t = M(u)\psi$ (problem 61).

In summary, we need an $M(u)$ such that:

1. $M(u)^\dagger = M(u)$, *i.e.* M is antisymmetric;

2. $[M(u), L(u)]$ is multiplicative.

How to write such an M? Being a differential operator in x, we can write M

$$M = \sum_{j=0}^{m} \alpha_j(x) D^j \tag{11.6}$$

where $\alpha_j(x)$ are functions of x (and in principle of t, but we'll suppress that dependence in our notation). we'll choose a different basis by writing

$$M = \sum_{j=0}^{m} (\beta_j(x) D^j + D^j \beta_j(x)) \,.$$

(It can be checked that this is always possible.)

Now if $\alpha(x)$ is real, $\alpha(x)^\dagger = \alpha(x)$, and also

$$D^\dagger = -D \tag{11.7}$$

(this is proved by integration by parts), which implies

$$\begin{aligned} (D^{2j})^\dagger &= D^{2j} && \text{(self-adjoint)} \\ (D^{2j-1})^\dagger &= -D^{2j-1} && \text{(skew-adjoint)} \,. \end{aligned} \tag{11.8}$$

Replacing M by its antisymmetric part $\frac{1}{2}(M - M^\dagger)$, it becomes

$$M = \sum_{0<2j-1\leqslant m} \left(\beta_{2j-1}(x)D^{2j-1} + D^{2j-1}\beta_{2j-1}(x)\right) . \tag{11.9}$$

It can also be checked that $[L, M]$ being multiplicative forces the coefficient of the leading term in D to be a constant, so the general guess is

$$M_n(u) = D^{2n-3} + \sum_{j=1}^{n-2} \left(\beta_j(x)D^{2j-1} + D^{2j-1}\beta_j(x)\right) . \tag{11.10}$$

where $\beta_j(x)$ are real functions.

Notes:

- the degree $2n - 3$ of the leading term was picked for later convenience;
- the β_j's have been relabelled going from (11.9) to (11.10);
- setting the coefficient of the leading term to 1 in (11.10) does not lose any generality, since an overall rescaling of $M(u)$ can be "undone" in $L_t + [L, M] = 0$ by rescaling time.

There's now no alternative but to calculate. When the dust settles, $N_n(u) \equiv [M_n, L]$ will be a polynomial in u, u_x, u_{xx} etc, and setting $L_t + [L, M_n] = 0$, that is $u_t = N_n(u)$, will give a KdV-like equation with x derivatives up to order $2n - 3$.

The first few cases:

$\boxed{\text{n=1}}$ We have $M(u) = 0$, therefore $[M(u), L(u)] = 0 = N_1(u)$. The PDE ofor u is then

$$\boxed{u_t = 0} \,, \tag{11.11}$$

which makes $L(u) = D^2 + u$ isospectral at different t, but trivially: the whole operator $L(u)$ does not depend on t, not just its eigenvalues.

$\boxed{\text{n=2}}$ We have $M(u) = D$, therefore

$$[M(u), L(u)] = [D, D^2 + u] = u_x \equiv N_2(u) \ . \tag{11.12}$$

The PDE for u is then the **advection equation**

$$\boxed{u_t = u_x} \ . \tag{11.13}$$

Its general solution is

$$u(x,t) = u(x+t, 0) \ ,$$

a travelling wave moving at constant velocity -1 (we could change the velocity by changing the coefficient of the leading term in $M_2(u)$). Again $L(u) = D^2 + u$ is isospectral at different t, but still quite trivially: the profile of u translates rigidly at a fixed speed, and the same applies to its eigenfunctions, hence the eigenvalues remain the same.

$\boxed{\text{n=3}}$ We have $M(u) = D^3 + \beta_1 D + D\beta_1$, therefore

$$\begin{aligned}
[M(u), L(u)] &= [D^3 + \beta_1 D + D\beta_1, D^2 + u] = [D^3 + 2\beta_1 D + \beta_{1,x}, D^2 + u] \\
&= [D^3, u] - 2[D^2, \beta_1]D + 2\beta_1[D, u] - [D^2, \beta_{1,x}] \\
&= (3u_x D^2 + 3u_{xx} D + u_{xxx}) - 2(2\beta_{1,x} D + \beta_{1,xx})D \\
&+ 2\beta_1 u_x - (2\beta_{1,xx} D + \beta_{1,xxx}) \\
&= (3u_x - 4\beta_{1,x})D^2 + (3u_{xx} - 4\beta_{1,xx})D + (u_{xxx} + 2\beta_1 u_x - \beta_{1,xxx}) \ .
\end{aligned} \tag{11.14}$$

If we require that this operator be multiplicative, the (left) coefficients of D^2 and D must vanish. Setting to zero the D^2 term we find that

$$3u_x - 4\beta_{1,x} = 0 \implies \beta_1 = \frac{3}{4}u + k$$

for a constant k.[1] The D term then vanishes too, and the multiplicative D^0 term becomes

$$u_{xxx} - \beta_{1,xxx} + 2\beta_1 u_x = \frac{1}{4}u_{xxx} + \frac{3}{2}uu_x + 2ku_x \equiv N_3(u) \ .$$

The PDE for u is then

$$\boxed{u_t = \frac{1}{4}u_{xxx} + \frac{3}{2}uu_x + 2ku_x} \ . \tag{11.15}$$

Rescaling $t \mapsto -4t$ and taking $k = 0$, this becomes nothing but the KdV equation. Alternatively, it's easy to check that if $u(x,t)$ satisfies (11.15), then

$$\boxed{\widetilde{u}(x,t) := u(x + 8kt, -4t)} \tag{11.16}$$

satisfies the KdV equation.

[1]To be precise, k need only be constant with respect to x. However, any t-dependence can be absorbed by a redefinition of t, x, u, so we can take k to be a constant without loss of generality.

This shows that the KdV equation is the third member of a hierarchy of partial differential equations $u_t = N_n(u) = [M_n(u), L(u)]$:

$$\boxed{\begin{array}{ll} n=1: & u_t = 0 \\ n=2: & u_t + u_x = 0 \\ n=3: & u_t + 6uu_x + u_{xxx} = 0 \\ n=4: & u_t + 30u^2u_x + 20u_xu_{xx} + 10uu_{xxx} + u_{xxxxx} = 0 \end{array}} \tag{11.17}$$

These are the first equations of the **KdV hierarchy**, and in each case, they evolve $u(x,t)$ forward in time in such a way as to leave the spectrum of $L(u) = D^2 + u$ unchanged.

We normalized the n-th member of the KdV hierarchy (11.17) in such a way that the $u_{x\ldots x}$ term with $2n-3$ derivatives has coefficient 1. The corresponding $M_n(u)$ operators are

$$\begin{array}{ll} n=1: & M_1(u) = 0 \\ n=2: & M_2(u) = -D \\ n=3: & M_3(u) = -4D^3 - (3uD + Du) \\ n=4: & M_4(u) = -16D^5 - 20(uD^3 + D^3u) - 5\left((3u^2 - u_{xx})D + D(3u^2 - u_{xx})\right) \ . \end{array}$$

11.2 Connection with conservation laws

Recall from last term that the KdV equation has an infinite sequence of conserved charges:

$$\boxed{Q_n = \int_{-\infty}^{+\infty} dx\ \rho_n} \tag{11.18}$$

where the conservation of Q_n, $\frac{dQ_n}{dt} = 0$, is proved by showing that

$$\boxed{\frac{\partial}{\partial t}\rho_n + \frac{\partial}{\partial x j_n} = 0} \tag{11.19}$$

when the KdV equation holds, for some current density j_n with

$$\boxed{[j_n]_{-\infty}^{\infty} = 0}\ . \tag{11.20}$$

Normalising the charge densities ρ_n's as

$$\boxed{\rho_n = u^n + \ldots}\ . \tag{11.21}$$

the first few examples are

$$\begin{aligned}
\rho_1 &= u \\
\rho_2 &= u^2 \\
\rho_3 &= u^3 - \tfrac{1}{2}u_x^2 \\
\rho_4 &= u^4 - 2uu_x^2 + \tfrac{1}{5}u_{xx}^2 \\
\rho_5 &= u^5 - \tfrac{105}{21}u^2u_x^2 + uu_{xx}^2 - \tfrac{1}{21}u_{xxx}^2 \\
&\ \vdots
\end{aligned} \tag{11.22}$$

So we now have two infinite sequences:

- For the KdV equation itself, the sequence Q_1, Q_2, Q_3,... of conserved charges;
- Going beyond KdV, an infinite sequence N_1, N_2, N_3,... of polynomials in u and its x derivatives such that setting $u_t = N_n(u)$ leaves the eigenvalues of $D^2 + u(x,t)$ constant.

How do these two sequences tie together, if at all?

The most boring possibility: each evolution equation $u_t = K_n(u)$ has its "own" set of Q_n's, conserved charges for that equation alone. In fact the answer, found by Gardner, is more clever. To explain it, a new concept is needed.

11.2.1 The functional derivative

(Also known as the variational, or Fréchet, derivative.)

Suppose f is some function of u and its x derivatives. Then

$$\boxed{F[u] = \int_{-\infty}^{+\infty} dx f(u, u_x, u_{xx}\dots)} \tag{11.23}$$

is an example of a *functional* of u: it takes a function $u(x)$ and yields a number $F[u]$. In practice u might also depend on the time t, in which case the formula should be taken at fixed t, which is not integrated over. Since t is a spectator for most of the following considerations, for now we won't write it explicitly in formulae.

Now consider a small variation $\delta u(x)$ of $u(x)$, such that $u(x) \to u(x) + \delta u(x)$, with $\delta u(x) \to 0$ as $x \to \pm\infty$. We demand similarly that $\delta u_x \equiv (\delta u)_x, \delta u_{xx} \equiv (\delta u)_{xx}, \dots \to 0$ as $x \to \pm\infty$.

This changes $F[u]$ to

$$
\begin{aligned}
F[u+\delta u] &= \int_{-\infty}^{+\infty} dx f(u+\delta u, (u+\delta u)_x, (u+\delta u)_{xx}, \dots) \\
&= \int_{-\infty}^{+\infty} dx f(u+\delta u, u_x+\delta u_x, u_{xx}+\delta u_{xx}, \dots) \\
&= \int_{-\infty}^{+\infty} dx \left(f(u, u_x, u_{xx}, \dots) + \frac{\partial f}{\partial u}\,\delta u + \frac{\partial f}{\partial u_x}\,\delta u_x + \frac{\partial f}{\partial u_{xx}}\,\delta u_{xx} + \dots \right) \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\text{(Taylor expanding)} \\
&= F[u] + \int_{-\infty}^{+\infty} dx \left(\frac{\partial f}{\partial u}\,\delta u + \frac{\partial f}{\partial u_x}\,\delta u_x + \frac{\partial f}{\partial u_{xx}}\,\delta u_{xx} + \dots \right) + O((\delta u)^2) \\
&= F[u] + \int_{-\infty}^{+\infty} dx \left(\frac{\partial f}{\partial u} - \frac{\partial}{\partial x}\frac{\partial f}{\partial u_x} + \frac{\partial^2}{\partial x^2}\frac{\partial f}{\partial u_{xx}} + \dots \right) \delta u + O((\delta u)^2) \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\text{(integrating by parts)}
\end{aligned}
\tag{11.24}
$$

and the term multiplying $\delta u(x)$ in the last line is called the **functional derivative** of $F[u]$, written as $\frac{\delta F[u]}{\delta u}$. More precisely, $\frac{\delta F[u]}{\delta u}$ is defined by

$$
\boxed{F[u+\delta u] = F[u] + \underbrace{\int_{-\infty}^{+\infty} dx\, \frac{\delta F[u]}{\delta u}\,\delta u}_{=:\ \delta F[u]} + O(\delta u^2)}
\tag{11.25}
$$

which is like $f(x+\delta x) = f(x) + \frac{df}{dx}\,\delta x + O((\delta x)^2)$ for ordinary functions.

For functionals defined as in (11.23) the calculation just completed shows that

$$
\boxed{\frac{\delta F[u]}{\delta u} = \frac{\partial f}{\partial u} - \frac{\partial}{\partial x}\frac{\partial f}{\partial u_x} + \frac{\partial^2}{\partial x^2}\frac{\partial f}{\partial u_{xx}} + \dots}
\tag{11.26}
$$

Aside: The concept of functional derivative is central in the modern understanding of classical physics, including classical field theory as we are studying here. One can obtain the equations of motion (or 'Euler-Lagrange equations') for a classical field $u(x,t)$ by requiring that the action

$$S[u] = \int dx\, dt\, \mathcal{L}(u, u_t, u_x, \dots)$$

is stationary under all infinitesimal variations δu of the field consistent with the boundary conditions.

Examples

(a) $f = u \quad \Rightarrow \quad \frac{\delta F[u]}{\delta u} = 1$

(b) $f = u^3 \quad \Rightarrow \quad \frac{\delta F[u]}{\delta u} = 3u^2$

(c) $f = u_x^2 \quad \Rightarrow \quad \frac{\delta F[u]}{\delta u} = -2u_{xx}$

(Exercise: check these results.)

The conserved quantities $Q_n = \int dx \rho_n$ are examples of functionals of u, and so we can also calculate their functional derivatives:

$$\begin{aligned}
\frac{\delta Q_1}{\delta u} &= \frac{\delta}{\delta u} \int_{-\infty}^{+\infty} dx \; u = 1 \\
\frac{\delta Q_2}{\delta u} &= \frac{\delta}{\delta u} \int_{-\infty}^{+\infty} dx \; u^2 = 2u \\
\frac{\delta Q_3}{\delta u} &= \frac{\delta}{\delta u} \int_{-\infty}^{+\infty} dx \; (u^3 - \frac{1}{2} u_x^2) = 3u^2 + u_{xx} \\
\frac{\delta Q_4}{\delta u} &= \frac{\delta}{\delta u} \int_{-\infty}^{+\infty} dx \; (u^4 - 2uu_x^2 + \frac{1}{5} u_{xx}^2) \\
&= 4u^3 + 4u_x^2 + 4uu_{xx} - 2u_x^2 + \frac{2}{5} u_{xxxx} \\
&\vdots
\end{aligned} \tag{11.27}$$

Taking $\frac{\partial}{\partial x}$ of each of these,

$$\frac{\partial(1)}{\partial x} = 0 \, , \qquad \frac{\partial(2u)}{\partial x} = 2u_x \, , \qquad \frac{\partial(3u^2 + u_{xx})}{\partial x} = 6uu_x + u_{xxx}$$

and these match, up to an overall scale, the first three equations of the KdV hierarchy:

$$u_t = 0 \, , \qquad u_t = -u_x \, , \qquad u_t = -6uu_x - u_{xxx} \, .$$

The normalisations of the charges, or else the scale of t, can be adjusted to make these matches precise. They are the first three examples of Gardner's general result:

$$\boxed{u_t = \frac{\partial}{\partial x} \left(\frac{\delta Q_n}{\delta u} \right)} \quad \longleftrightarrow \quad \boxed{u_t = N_n(u)} \tag{11.28}$$

connecting the n^{th} KdV conservation law with the n^{th} equation of the KdV hierarchy. Thus the two sequences are the same!

Furthermore:

1. If $u_m(x,t)$ evolves by the m^{th} equation in the KdV hierarchy, all the Q_n's are conserved densities for it;

2. Imagine we have one "time" for each equation in the hierarchy, so that instead of $u_m(x,t)$ with $\frac{\partial}{\partial t}u_m = N_m(u)$ we have $u(x, t_1, t_2, t_3 \dots)$ with

$$\boxed{\frac{\partial}{\partial t_m}u_m = N_m(u) \equiv \frac{\partial}{\partial x}\frac{\delta Q_m}{\delta u}} \quad \forall m = 1, 2, \dots \tag{11.29}$$

   Then if we evolve (or 'flow') $u(x, t_1, t_2, t_3 \dots)$ for a while in t_i, then for a while in t_j, we end up with the <u>same</u> function of x as if we'd evolved in t_j first followed by t_i. This is the idea of **commuting flows**: it's very important in "modern" soliton theory.

Chapter 12

Inverse scattering (or "reassembly")

To conclude the inverse scattering method, we need to reassemble the KdV field $u(x,t)$, or equivalently the Schrödinger potential $V(x;t) = -u(x,t)$, from the time-evolved scattering data. This is step **(c)**: "reassembly / inverse scattering".

This touches on a general question: if all you were allowed to do was sit at infinity and chuck particles at your potential, and measure how they come back, could you deduce the form of $V(x)$?

This question is of practical importance, for example when looking for oil using seismic reflection, or in medicine (one example there being deducing the shape of the inner ear from reflected sound waves). It belongs to the category of "inverse problems": deducing the form of an operator (here $D^2 + u$) from information about its spectrum (μ_i, c_n and so on): "can you hear the shape of a drum?"

For this one-dimensional (Schrödinger) case, the result was already known, found by Marchenko (following earlier work by Gelfand and Levitan), some years before GGKM.

In fact you don't need to know $T(k)$, just $R(k)$ for real k, together with the N discrete eigenvalues $-\mu_j^2$, $j = 1, \dots N$, and the normalising coefficients c_j, $j = 1, \dots N$. The full set $\left\{ R(k),\ \ \{\mu_n, c_n\}_{n=1}^{N} \right\}$ is precisely the scattering data that we evolved forward in time in the chapter before last.

There are two important special cases:

1. $N = 0$: $V(x)$ has no bound states;
2. $R(k) = 0 \ \forall k$: $V(x)$ is reflectionless, but there is still information about $V(x)$ hidden in the bound state eigenvalues and normalisation coefficients.

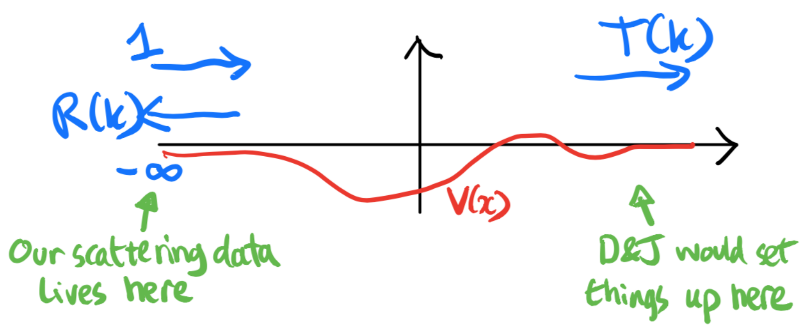
It turns out that:

1. $\Rightarrow$ initial data contains no solitons;
2. $\Rightarrow$ initial data contains only solitons.

12.1 The recipe for inverse scattering: the Marchenko equation

We want to solve the inverse scattering problem for given scattering data at $x = -\infty$ to determine the potential $V(x)$, and hence the KdV field $u(x) = -V(x)$, at any fixed KdV time t.

The derivation is long and we'll skip it here – see for example section 3.3 of Drazin and Johnson [Drazin and Johnson, 1989]. But a warning: everything in Drazin and Johnson is phrased in terms of scattering solutions with waves arriving from $+\infty$, and asymptotics also at $+\infty$, while we do the opposite:



Once the not inconsiderable quantity of dust has settled, the upshot is the following recipe:

1. Construct the function

$$\boxed{F(\xi) = \int_{-\infty}^{+\infty} \frac{dk}{2\pi}\, R(k)\, e^{-ik\xi} + \sum_{n=1}^{N} c_n^2\, e^{\mu_n \xi}} \tag{12.1}$$

   from the scattering data

$$\boxed{S = \Big\{ R(k)\,, \quad \{\mu_n, c_n\}_{n=1}^{N} \Big\}} \tag{12.2}$$

2. Solve the *Marchenko equation*

$$\boxed{K(x,z) + F(x+z) + \int_{-\infty}^{x} dy\, K(x,y)\, F(y+z) = 0} \tag{12.3}$$

   to determine the unknown function $K(x,z)$ for all $z \leqslant x$ (and set $K(x,z) = 0$ for $x < z$).

3. Finally determine the Schrödinger potential from

$$\boxed{V(x) = 2\frac{d}{dx} K(x,x)} \tag{12.4}$$

   The KdV field is then given by $u = -V$.

   Note: $K(x,x)$ is defined by demanding one-sided continuity of $K(x,z)$, as the left-sided limit of $K(x,z)$ at $z = x$:

$$K(x,x) := \lim_{z \to x^-} K(x,z)\ .$$

This all applies at one fixed KdV time t. But using the results of the last section of the last chapter, we know that

$$\begin{aligned} R(k;t) &= R(k;0)\, e^{-8ik^3 t} \\ c_n(t) &= c_n(0)\, e^{-4\mu_n^3 t} \end{aligned} \tag{12.5}$$

while k^2 and μ_n^2 are independent of time.

So to find the field at time t, we just apply the above recipe starting from

$$\begin{aligned} F(\xi;t) &= \int_{-\infty}^{+\infty} \frac{dk}{2\pi}\, R(k;t)\, e^{-ik\xi} + \sum_{n=1}^{N} c_n(t)^2\, e^{\mu_n \xi} \\ &= \int_{-\infty}^{+\infty} \frac{dk}{2\pi}\, R(k;0)\, e^{-ik(\xi + 8k^2 t)} + \sum_{n=1}^{N} c_n(0)^2\, e^{\mu_n(\xi - 8\mu_n^2 t)} \end{aligned} \tag{12.6}$$

At least in principle, this solves the problem! In practice the term involving R in the definition of F, with the integral over k, makes the calculation of F hard when $t > 0$. But for reflectionless potentials this term is absent, and $F(\xi, t)$ can be read off at any time t. This turns out to yield the 'pure' multisoliton solutions that can also be found via Bäcklund or Hirota. Even when R is nonzero, it can be shown that the term involving R goes to zero as $t \to \infty$. All of which leads to the following 'big picture':

(A) $\{\mu_n, c_n\}_{n=1}^{N} \leftrightarrow N$ right-moving solitons hidden inside the initial data:

(B) $R(k) \leftrightarrow$ a superposition of dispersive left-moving waves hidden inside the initial data:

The net result is a sort of "nonlinear Fourier analysis" (which reduces to the usual Fourier solution in the limit of small-amplitude waves).

12.2 Example 1: the single KdV soliton

Consider a reflectionless potential, so $R(k) = 0$, with just one bound state encoded in $\{\mu_1, c_1\} \equiv \{\mu, c\}$. Then (at fixed t)

$$\boxed{F(\xi) = c^2 e^{\mu\xi}} \tag{12.7}$$

and the Marchenko equation (12.3) reads

$$\boxed{K(x,z) + c^2 e^{\mu(x+z)} + \int_{-\infty}^{x} dy\, K(x,y)\, c^2 e^{\mu(y+z)} = 0} \tag{12.8}$$

This needs to be solved for $z \leqslant x$. As a first step, factorise $e^{\mu z}$ from the last two terms:

$$K(x,z) + e^{\mu z} \left(c^2 e^{\mu x} + \int_{-\infty}^{x} dy\, K(x,y)\, c^2 e^{\mu y} \right) = 0\,, \tag{12.9}$$

and note that the terms in brackets are independent of z, meaning that

$$\boxed{K(x,z) = h(x)\, e^{\mu z}} \tag{12.10}$$

for some $h(x)$. Substituting back into (12.9) and dividing through by $e^{\mu z}$, $h(x)$ must satisfy

$$0 = h(x) + c^2 e^{\mu x} + c^2 \int_{-\infty}^{x} dy\, h(x)\, e^{2\mu y} = h(x) \left(1 + c^2 \int_{-\infty}^{x} dy\, e^{2\mu y} \right) + c^2 e^{\mu x}$$

and hence

$$h(x) = -\frac{c^2 e^{\mu x}}{1 + \frac{c^2}{2\mu} e^{2\mu x}} \,. \tag{12.11}$$

If we set

$$c^2 = 2\mu\, e^{-2\mu x_0} \tag{12.12}$$

(thereby trading c for x_0) we obtain

$$h(x) = -2\mu \frac{e^{\mu(x-2x_0)}}{1 + e^{2\mu(x-x_0)}} \tag{12.13}$$

and so (for $z \leqslant x$)

$$K(x,z) = -2\mu \frac{e^{\mu(x+z-2x_0)}}{1 + e^{2\mu(x-x_0)}} \,. \tag{12.14}$$

Hence

$$\boxed{V(x) = 2\frac{d}{dx} K(x,x) = -2\mu^2 \mathrm{sech}^2(\mu(x-x_0))} \tag{12.15}$$

and $u = -V$ is indeed a snapshot of a single KdV soliton, at a time (say $t = 0$) when its centre is at $x = x_0$.

Time evolution is easily included using

$$c(t)^2 = c^2(0)\, e^{-8\mu^3 t} = 2\mu\, e^{-2\mu(x_0 - 4\mu^2 t)} \tag{12.16}$$

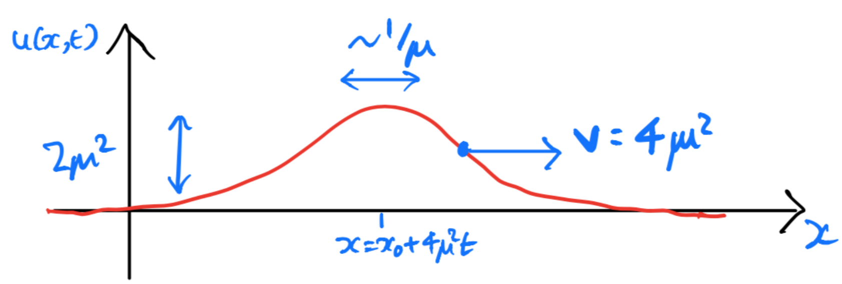
which has the effect of translating the centre of the soliton as

$$x_0 \to x_0 + 4\mu^2 t \tag{12.17}$$

and the KdV field at time t is

$$\boxed{u(x,t) = -V(x,t) = 2\mu^2 \mathrm{sech}^2(\mu(x - x_0 - 4\mu^2 t))} \tag{12.18}$$

which is a single moving soliton just as found earlier in the course:



12.3 Example 2: the N-soliton solution

NOTE: this section was not taught in 2023-24. It’s nice stuff which you are welcome to read, but it won’t be examined.

Now let’s consider a situation with $R(k) = 0$ but with N bound states, encoded in $\{\mu_n, c_n\}_{n=1}^{N}$. Then

$$F(\xi) = \sum_{n=1}^{N} c_n^2\, e^{\mu_n \xi}\,. \tag{12.19}$$

Since

$$F(x{+}z) = \sum_{n=1}^{N} c_n^2\, e^{\mu_n x} e^{\mu_n z}$$

is a sum of factorised terms, we will look for a solution where $K(x, z)$ is also a sum of factorised terms. This is best encoded using a vector and matrix notation, setting

$$E(x) = \begin{pmatrix} e^{\mu_1 x} \\ \vdots \\ e^{\mu_N x} \end{pmatrix}, \quad L(x) = \begin{pmatrix} c_1^2 e^{\mu_1 x} \\ \vdots \\ c_N^2 e^{\mu_N x} \end{pmatrix}, \quad H(x) = \begin{pmatrix} h_1(x) \\ \vdots \\ h_N(x) \end{pmatrix}, \tag{12.20}$$

where $H(x)$ is yet to be determined. With this notation set up, we have

$$F(x{+}z) = E^T(x) L(z) \tag{12.21}$$

(where the T superscript denotes a transpose) and we’ll look for a $K(x, z)$ of the form

$$K(x, z) = H^T(x) L(z)\,. \tag{12.22}$$

Substituting into the Marchenko equation, we find

$$\begin{aligned} 0 &= K(x,z) + F(x{+}z) + \int_{-\infty}^{x} dy\, K(x,y)\, F(y{+}z) \\ &= H^T(x)L(z) + E^T(x)L(z) + H^T(x) \int_{-\infty}^{x} dy\, L(y)E^T(y)L(z) \\ &= \Big(H(x) + E(x) + \int_{-\infty}^{x} dy\, E(y)L^T(y)H(x)\Big)^T L(z)\,. \end{aligned} \tag{12.23}$$

If the term in brackets on the last line can be made to vanish, we’ll have a solution. In turn this will be true if

$$\Gamma(x)\, H(x) = -E(x) \tag{12.24}$$

where $\Gamma(x)$ is not the gamma function seen earlier, but rather the $N \times N$ matrix

$$\Gamma(x) = \mathbb{1}_{N\times N} + \int_{-\infty}^{x} dy\, E(y)L^T(y) \tag{12.25}$$

with matrix elements

$$\begin{aligned}\Gamma(x)_{mn} &= \delta_{mn} + \int_{-\infty}^{x} dy\, e^{\mu_m y} c_n^2 e^{\mu_n y} \\ &= \delta_{mn} + c_n^2 \frac{e^{(\mu_m+\mu_n)y}}{\mu_m+\mu_n}\,.\end{aligned} \tag{12.26}$$

Note also we have

$$\frac{d}{dx}\Gamma(x) = E(x)L^T(x)\,, \tag{12.27}$$

a formula that will be useful shortly.

From (12.24) we have

$$H(x) = -\Gamma(x)^{-1}E(x) \tag{12.28}$$

and so

$$\begin{aligned}K(x,z) = L^T(z)H(x) &= -L^T(z)\Gamma(x)^{-1}E(x) \\ &= -\operatorname{tr}\big(\Gamma(x)^{-1}E(x)L^T(z)\big)\,.\end{aligned} \tag{12.29}$$

Therefore

$$\begin{aligned}K(x,x) &= -\operatorname{tr}\big(\Gamma(x)^{-1}E(x)L^T(x)\big) \\ &= -\operatorname{tr}\Big(\Gamma(x)^{-1}\frac{d}{dx}\Gamma(x)\Big) \\ &= -\operatorname{tr}\Big(\frac{d}{dx}\log\Gamma(x)\Big) \\ &= -\frac{d}{dx}\operatorname{tr}\big(\log\Gamma(x)\big) \\ &= -\frac{d}{dx}\log\big(\det\,\Gamma(x)\big)\end{aligned} \tag{12.30}$$

using the matrix identities

$$\frac{d}{dx}\log\Gamma = \Gamma^{-1}\frac{d}{dx}\Gamma\,, \quad \operatorname{tr}(\log\Gamma) = \log(\det\,\Gamma)\,. \tag{12.31}$$

This implies that the KdV field is

$$u = -2\frac{d}{dx}K(x,x) = 2\frac{d^2}{dx^2}\log(\det\,\Gamma(x)) \tag{12.32}$$

or, putting back the t-dependence hidden in Γ (through the c_n),

$$u(x,t) = 2\frac{\partial^2}{\partial x^2}\log(\det\,\Gamma(x;t)) \tag{12.33}$$

with

$$\Gamma(x;t)_{mn} = \delta_{mn} + c_n^2(t)\frac{e^{(\mu_m+\mu_n)x}}{\mu_m+\mu_n}\,. \tag{12.34}$$

These formulae are very similar to the N-soliton KdV solutions found by Hirota. To see that they are in fact exactly the same, we can use Sylvester's determinant theorem, which states that

$$\det(\mathbb{1}_{N\times N} + AB) = \det(\mathbb{1}_{N\times N} + BA) \tag{12.35}$$

for any pair of $N \times N$ matrices A, B.

Taking

$$A_{mn} = e^{\mu_m x}\delta_{mn} \; , \quad B_{mn} = \frac{c_n^2 e^{\mu_n x}}{\mu_m + \mu_n}$$

we have

$$(AB)_{mn} = \frac{c_n^2 e^{(\mu_m+\mu_n)x}}{\mu_m + \mu_n} \; , \quad (BA)_{mn} = \frac{c_n^2 e^{2\mu_n x}}{\mu_m + \mu_n} \; ,$$

and so we can equivalently write

$$\boxed{u(x,t) = 2\frac{\partial^2}{\partial x^2}\log(\det S(x;t))} \tag{12.36}$$

with

$$\begin{aligned} S(x;t)_{mn} &= \delta_{mn} + \frac{1}{\mu_m + \mu_n}\, c_n^2(t)\, e^{2\mu_n x} \\ \Longrightarrow \boxed{S(x;t)_{mn} = \delta_{mn} + \frac{2\mu_n}{\mu_m + \mu_n}\, e^{2\mu_n(x - x_{0,n} - 4\mu_n^2 t)}} \end{aligned} \tag{12.37}$$

where, just as done above for the one-soliton solution, we traded $c_n(0)$ for $x_{0,n}$ by setting

$$c_n(0)^2 = 2\mu_n e^{-2\mu_n x_{0,n}} \; . \tag{12.38}$$

These equations give the general form of the N-soliton solution of the KdV equation.

Chapter 13

Integrable systems in classical mechanics

So far, we've (secretly) been looking at infinite-dimensional systems: classical field theories in one space and one time dimension, though these can often be thought of as the continuum limits (see last term) of systems with discrete sets of degrees of freedom.

Many of the methods we've seen, in particular the idea of a Lax pair, can also apply to finite-dimensional systems, and more precisely to **finite-dimensional classical integrable Hamiltonian systems**. To understand what these words mean, some definitions are needed.

A **finite-dimensional Hamiltonian system** is defined by:

1. A set of (generalised) coordinates $q_{i=1\ldots n}$ and momenta $p_{i=1\ldots n}$, which completely specify the configuration of the system at time t (the $2n$-dimensional space parametrised by these so-called **canonical coordinates** q, p is called the **phase space** of the system);

2. A function $H(q, p)$ defined on phase space called the **Hamiltonian**. (The Hamiltonian may depend explicitly on time, in which case we write $H = H(q, p, t)$, but this won't be needed for our purposes.)

3. The time evolution equations are then **Hamilton's equations**.

$$\boxed{\begin{aligned} \dot{q}_i &= \frac{\partial H}{\partial p_i} \\ \dot{p}_i &= -\frac{\partial H}{\partial q_i} \end{aligned}} \tag{13.1}$$

with the dots denoting time derivatives.

<u>**Note:**</u> we can take $n \to \infty$ without taking a continuum limit, and get an **infinite-dimensional discrete Hamiltonian system**. Most (if not all) of what we'll see in the following applies to that case as well (with some extra care about limits and convergence).

<u>**Example:**</u>
for n point particles with masses m_i moving in one dimension under conservative forces associated with a potential energy $V(q_1, \dots q_n)$, the Hamiltonian is

$$H(q,p) = \sum_{i=1}^{n} \frac{p_i^2}{2m_i} + V(q_1, \dots q_n) \tag{13.2}$$

and Hamilton's equations are

$$\dot{q}_i = \frac{p_i}{m_i}\ , \quad \dot{p}_i = -\frac{\partial V(q_1, \dots q_n)}{\partial q_i}\ . \tag{13.3}$$

These are the same as Newton's equations,

$$m_i \ddot{q}_i = -\frac{\partial V(q_1, \dots q_n)}{\partial q_i}\ , \tag{13.4}$$

put into a first-order form.

One can associate to a Hamiltonian system a **Poisson bracket** $\{\ ,\ \}$, a bilinear antisymmetric form on the space of differentiable functions of q and p:

$$\boxed{\{f,g\} := \sum_{i=1}^{n} \left(\frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i} - \frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i} \right)} \tag{13.5}$$

Clearly $\{f,g\} = -\{g,f\}$ and $\{f,f\} = 0$.

Hamilton's equations (13.1) imply that any $f(q,p)$ which does not depend explicitly on time, but only implicitly via $q(t)$ and $p(t)$, evolves as

$$\begin{aligned} \frac{d}{dt} f(q(t),p(t)) &= \sum_{i=1}^{n} \left(\dot{q}_i \frac{\partial f}{\partial q_i} + \dot{p}_i \frac{\partial f}{\partial p_i} \right) \\ &= \sum_{i=1}^{n} \left(\frac{\partial H}{\partial p_i}\frac{\partial f}{\partial q_i} - \frac{\partial H}{\partial q_i}\frac{\partial f}{\partial p_i} \right) \end{aligned}$$

That is,

$$\boxed{\frac{d}{dt}f(q,p) = \{H(q,p), f(q,p)\}} \tag{13.6}$$

(If f also depends explicitly on t, so $f = f(q(t), p(t), t)$, then

$$\frac{d}{dt}f(q,p,t) = \frac{\partial}{\partial t}f(q,p,t) + \{H(q,p), f(q,p,t)\}\,,$$

by the chain rule and Hamilton's equations.)

Functions $F(q,p)$ which don't depend explicitly on time and have vanishing Poisson bracket with the Hamilton $H(q,p)$ are conserved:

$$\boxed{\frac{d}{dt}F(q(t),p(t)) = \{H(q,p), F(q,p)\} = 0} \tag{13.7}$$

In particular, the antisymmetry of the Poisson bracket means that the Hamiltonian is conserved, as long as it doesn't depend explicitly on time (which we always assume):

$$\frac{d}{dt}H(q(t),p(t)) = \{H(q(t),p(t)), H(q(t),p(t))\} = 0\,. \tag{13.8}$$

Hence

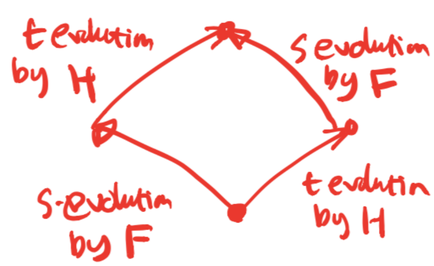
$$\boxed{H(q(t),p(t)) = E = \text{constant}} \tag{13.9}$$

which is nothing but the conservation of energy.

Note: If $\{F, H\} = 0$, then not only is $F(q,p)$ conserved under the time evolution (13.1), but also $H(q,p)$ is conserved under a different time evolution with a different time, s say, and Hamiltonian $F(q,p)$:

$$\left\{\begin{aligned} \frac{d}{ds}q_i &= \frac{\partial F}{\partial p_i} \\ \frac{d}{ds}p_i &= -\frac{\partial F}{\partial q_i} \end{aligned}\right\} \quad \Rightarrow \quad \frac{d}{ds}H(q,p) = \{F(q,p), H(q,p)\} = 0\,. \tag{13.10}$$

It also means (via the Jacobi identity for the Poisson bracket) that we can evolve along the two times, along t and then s, or vice versa, and we will end up at the same point in phase space:



In fancy language, F and H such that $\{F, H\} = 0$ are said to be **in involution** and they generate **commuting flows**, where one flow is t-evolution with Hamiltonian H, and the other flow is s-evolution with Hamiltonian F. We saw this idea earlier when discussing the KdV hierarchy. (The Poisson bracket was not introduced there, but it is possible to do so.)

Definition: A finite-dimensional Hamiltonian system $\{\, q_{i=1\dots n}, p_{i=1\dots n}, H(q_i, p_i) \,\}$ is called **completely integrable** if it has n independent conserved quantities $Q_i(q,p)$ satisfying $\{Q_i, H\} = 0$, which are mutually *in involution*, that is

$$\boxed{\{Q_i, Q_j\} = 0 \quad \forall i, j = 1 \dots n} \tag{13.11}$$

One of these conserved quantities is always the original Hamiltonian H.

For such systems it is possible to find a new set of coordinates φ_i and momenta Q_i on phase space such that the Hamiltonian only depends of the Q_i and not on the φ_i:

$$H = H(Q) \quad \Rightarrow \quad \begin{cases} \dot{\varphi} = \dfrac{\partial H}{\partial Q_i} \\ \dot{Q} = -\dfrac{\partial H}{\partial \varphi_i} = 0 \end{cases} \tag{13.12}$$

These are called **action-angle variables** (φ_i: angle variables; Q_i: action variables). The name is because if the surfaces of constant H are compact, then the φ_i parametrise periodic orbits and can therefore be thought of as angular variables.

The n conserved quantities Q_i are the finite-dimensional analogues of the infinitely-many conserved charges of the KdV hierarchy discussed in section 11.2.

What is interesting for us here is that the integrability of such classical systems can be established by constructing a Lax pair L, M, satisfying

$$\boxed{\dot{L} = [M, L]} \tag{13.13}$$

This is as we saw for KdV, but now L and M will be $n \times n$ matrices instead of differential operators. We'll see that the n conserved quantities are the eigenvalues $\lambda_{i=1\dots n}$ of the Lax matrix L (though as we'll also see, it may be more convenient sometimes to use some functions of those eigenvalues instead, such as the sums of their powers).

(To show that the conservation laws are in involution is a bit more tricky, and won't be discussed here.)

In general, if there are n q's, $q_{i=1\dots n}$, L and M will be $n \times n$ matrices and the n conserved quantities will be coded up in the n eigenvalues $\lambda_1 \dots \lambda_n$ of the Lax matrix L.

The Lax equation (13.13), with L and M functions of time, can be solved formally by

$$\boxed{L(t) = U(t)L(0)U(t)^{-1}} \tag{13.14}$$

where the time evolution operator $U(t)$ is the unique solution of the following (matrix) ordinary differential equation:

$$\boxed{\begin{aligned} \dot{U}(t) &= M(t)U(t) \\ U(0) &= \mathbb{1} \end{aligned}} \tag{13.15}$$

This can be proved as follows:

$$\begin{aligned} \dot{L} &= \frac{d}{dt}\left(UL(0)U^{-1}\right) \\ &= \dot{U}L(0)U^{-1} + UL(0)(\dot{U^{-1}}) \\ &= \dot{U}L(0)U^{-1} - UL(0)U^{-1}\dot{U}U^{-1} \\ &= \dot{U}U^{-1}UL(0)U^{-1} - UL(0)U^{-1}\dot{U}U^{-1} \\ &= ML - ML \\ &= [M, L] \end{aligned} \tag{13.16}$$

where the result $(\dot{U^{-1}}) = -U^{-1}\dot{U}U^{-1}$ used in going from the second line to the third can be proved by differentiating $UU^{-1} = \mathbb{1}$, and we used $M = \dot{U}U^{-1}$ in the penultimate equality.

Note that the time evolution operator U is unitary (that is, $U^\dagger = U^{-1}$) if M is anti-hermitian (that is, $M^\dagger = -M$).

The formal solution (13.14) can be used to prove that the eigenvalues of the Lax matrix L do not depend on time, just as was the case for KdV in infinitely-many dimensions. To see this, consider the **characteristic polynomial** of L:

$$\boxed{P_L(\lambda) = \det(\lambda\mathbb{1} - L)} \tag{13.17}$$

This is a degree n monic polynomial ("monic": $\lambda^n + \ldots$) whose roots are the n eigenvalues $\lambda_{i=1\ldots n}$ of L. Now L is going to be a Hermitian – often real – matrix which can be diagonalised by conjugating it with some unitary matrix V:

$$L = V\Lambda V^{-1}, \quad \Lambda = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} \tag{13.18}$$

Thus (in a sequence of equalities that you might have seen before)

$$\begin{aligned}
P_L(\lambda) &= \det(\lambda \mathbb{1} - L) \\
&= \det(\lambda \mathbb{1} - V\Lambda V^{-1}) \\
&= \det(\lambda V V^{-1} - V\Lambda V^{-1}) \\
&= \det(V(\lambda \mathbb{1} - \Lambda)V^{-1}) \\
&= \det(V)\det(\lambda \mathbb{1} - \Lambda)\det(V^{-1}) \\
&= \det(\lambda \mathbb{1} - \Lambda) \\
&= \prod_{i=1}^{n}(\lambda - \lambda_i) \\
&= \lambda^n - c_1\lambda^{n-1} + c_2\lambda^{n-2} - \cdots + (-1)^n \prod_{i=1}^{n} \lambda_i \,.
\end{aligned} \tag{13.19}$$

(The signs of the coefficients on the last line are chosen for later convenience.)

Since time evolution is also given by conjugation (this time by $U(t)$ instead of V), the same argument shows that

$$\begin{aligned}
P_{L(t)}(\lambda) &= \det(\lambda \mathbb{1} - U(t)L(0)U(t)^{-1}) \\
&= \det(\lambda \mathbb{1} - L(0)) \\
&= P_{L(0)}(\lambda)
\end{aligned} \tag{13.20}$$

which implies that the eigenvalues λ_i of $L(t)$ are independent of time, as claimed.

Equivalently, we can take the n conserved quantities to be the coefficients c_k of the characteristic polynomial, also known as elementary symmetric polynomials,

$$c_k = \sum_{1 \leqslant i_1 < i_2 \cdots < i_k \leqslant n} \lambda_{i_1}\lambda_{i_2}\ldots\lambda_{i_k} \,, \quad k = 1 \ldots n \,, \tag{13.21}$$

or equivalently as the so-called power sum symmetric polynomials

$$s_k = \sum_{i=1}^{n} \lambda_i^k = \mathrm{tr}(L^k) \,, \quad k = 1 \ldots n \tag{13.22}$$

The two sets of symmetric polynomials are related by the Girard-Newton identities

$$k c_k = \sum_{i=1}^{k} (-1)^{i-1} c_{k-i} s_i \,.$$

Note that the conservation of s_k can be proved directly, taking d/dt of $\mathrm{tr}(L^k)$, expanding out, and using the Lax pair and then the cyclic property of the trace.

As a final remark about the general formalism, note that the eigenvalue equation for $L(t)$, namely

$$L(t)\psi(t) = \lambda\psi(t) \tag{13.23}$$

is solved formally by

$$\psi(t) = U(t)\psi(0) \tag{13.24}$$

where $\psi(0)$ is an eigenfunction at $t = 0$:

$$\begin{aligned} L(t)\psi(t) &= U(t)L(0)U(t)^{-1}U(t)\psi(0) \\ &= U(t)L(0)\psi(0) \\ &= U(t)\lambda\psi(0) \\ &= \lambda U(t)\psi(0) \\ &= \lambda\psi(t)\,. \end{aligned} \tag{13.25}$$

13.1 The Lax pair for the simple harmonic oscillator

The Hamiltonian for the simple harmonic oscillator (S.H.O.), which has $n = 1$, is

$$H(q,p) = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 q^2\,. \tag{13.26}$$

Hamilton's equations are then

$$\boxed{\dot{q} = \frac{p}{m}\,, \quad \dot{p} = -m\omega^2 q\,.} \tag{13.27}$$

These equations are equivalent to a Lax equation of the form (13.13) with

$$L = \begin{pmatrix} p & m\omega q \\ m\omega q & -p \end{pmatrix}\,, \quad M = \frac{\omega}{2}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\,. \tag{13.28}$$

Indeed

$$\dot{L} = \begin{pmatrix} \dot{p} & m\omega\dot{q} \\ m\omega\dot{q} & -\dot{p} \end{pmatrix}\,, \quad [M,L] = \begin{pmatrix} -m\omega^2 q & \omega p \\ \omega p & m\omega^2 q \end{pmatrix} \tag{13.29}$$

and so $\dot{L} = [M,L] \quad \leftrightarrow \quad$ (13.27).

Since in this case M is independent of t, the time evolution operator defined by (13.15) is simply

$$\boxed{U(t) = e^{Mt}} \tag{13.30}$$

where the exponential of the matrix Mt is defined by its Taylor expansion:

$$e^{Mt} = \sum_{n=0}^{\infty} \frac{t^n}{n!} M^n\,. \tag{13.31}$$

This can be calculated explicitly, noting that

$$M^2 = -\left(\frac{\omega}{2}\right)^2 \mathbb{1}$$

and so

$$M^{2k} = (-1)^k \left(\frac{\omega}{2}\right)^{2k} \mathbb{1}\,, \quad M^{2k+1} = (-1)^k \left(\frac{\omega}{2}\right)^{2k} M = (-1)^k \left(\frac{\omega}{2}\right)^{2k+1} \begin{pmatrix} 0 & -1 \\ 1 & 0\,, \end{pmatrix} \tag{13.32}$$

and so

$$\begin{aligned} U(t) &= \sum_{k=0}^{\infty} \frac{t^{2k}}{(2k)!} M^{2k} + \sum_{k=0}^{\infty} \frac{t^{2k+1}}{(2k+1)!} M^{2k+1} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} \left(\frac{\omega t}{2}\right)^{2k} + \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} \left(\frac{\omega t}{2}\right)^{2k+1} \\ &= \cos\left(\frac{\omega t}{2}\right) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \sin\left(\frac{\omega t}{2}\right) \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} \cos(\omega t/2) & -\sin(\omega t/2) \\ \sin(\omega t/2) & \cos(\omega t/2) \end{pmatrix} \end{aligned} \tag{13.33}$$

where we used the Taylor series for sine and cosine in the third equality. Note that the time evolution operator is nothing by a rotation matrix by the angle $\omega t/2$.

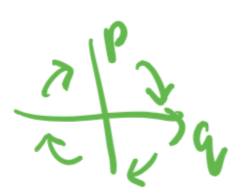
Hence

$$\begin{aligned} L(t) &= \begin{pmatrix} p(t) & m\omega q(t) \\ m\omega q(t) & -p(t) \end{pmatrix} = U(t) L(0) U(t)^{-1} \\ &= \begin{pmatrix} \cos(\omega t/2) & -\sin(\omega t/2) \\ \sin(\omega t/2) & \cos(\omega t/2) \end{pmatrix} \begin{pmatrix} p(0) & m\omega q(0) \\ m\omega q(0) & -p(0) \end{pmatrix} \begin{pmatrix} \cos(\omega t/2) & \sin(\omega t/2) \\ -\sin(\omega t/2) & \cos(\omega t/2) \end{pmatrix} \\ &= \dots \\ &= \begin{pmatrix} p(0)\cos(\omega t) - m\omega q(0)\sin(\omega t) & p(0)\sin(\omega t) + m\omega q(0)\cos(\omega t) \\ p(0)\sin(\omega t) + m\omega q(0)\cos(\omega t) & -p(0)\cos(\omega t) + m\omega q(0)\sin(\omega t) \end{pmatrix} \end{aligned} \tag{13.34}$$

and hence

$$\begin{aligned} q(t) &= q(0)\cos(\omega t) + \frac{p(0)}{m\omega}\sin(\omega t) \\ p(t) &= p(0)\cos(\omega t) - m\omega\, q(0)\, \sin(\omega t) \end{aligned} \tag{13.35}$$

This shows that, up to a scaling of the axes, the time evolution is a uniform rotation in the S.H.O. phase space:



In this case $n = 1$, and there is just one nontrivial conserved quantity, which should be the Hamiltonian. Indeed $\mathrm{tr}(L) = 0$ (so this is trivially conserved) while

$$\mathrm{tr}(L^2) = \mathrm{tr}\begin{pmatrix} p^2 + m^2\omega^2 q^2 & 0 \\ 0 & p^2 + m^2\omega^2 q^2 \end{pmatrix} = 2(p^2 + m^2\omega^2 q^2) = 4m\,H(q,p) \tag{13.36}$$

is the only independent conserved quantity. While this case is a bit easy, it does illustrate the general point that it's simpler to work with traces of powers of the Lax matrix, rather than with the individual eigenvalues themselves.

13.2 The Lax pair for the Toda lattice

The last example was a bit trivial. Much less trivial, and still the subject of research, is the **finite Toda lattice** which describes n particles on a line, each one interacting with its nearest neighbours through an exponential potential. Let's take the particles to have equal masses, $m_i = 1$. Toda's Hamiltonian is

$$H(p,q) = \sum_{i=1}^{n} \left(\frac{p_i^2}{2} + e^{-(q_i - q_{i-1})} \right) \tag{13.37}$$

where, at least at $t = 0$,

$$q_0 \equiv -\infty < q_1 < q_2 \cdots < q_n < q_{n+1} \equiv +\infty\,.$$

Hamilton's equations for this system are:[1]

$$\boxed{\begin{aligned} \dot{q}_i &= p_i \\ \dot{p}_i &= e^{-(q_i - q_{i-1})} - e^{-(q_{i+1} - q_i)} \end{aligned}} \tag{13.38}$$

which is a system of coupled differential equations.

[1]We can avoid formally setting $q_0 \equiv -\infty$ and $q_{n+1} \equiv +\infty$ if we write

$$H(p,q) = \sum_{i=1}^{n} \frac{p_i^2}{2} + \sum_{i=1}^{n-1} e^{-(q_{i+1} - q_i)}\,.$$

Then Hamilton's equations are $\dot{q}_i = p_i$ for all i and

$$\dot{p}_1 = -e^{-(q_2 - q_1)}\,,\ \dot{p}_2 = e^{-(q_2 - q_1)} - e^{-(q_3 - q_2)}\,,\ \ldots,\ \dot{p}_{n-1} = e^{-(q_{n-1} - q_{n-2})} - e^{-(q_n - q_{n-1})}\,,\ \dot{p}_n = e^{-(q_n - q_{n-1})}\,.$$

This agrees with (13.38) if we set $q_0 \equiv -\infty$ and $q_{n+1} \equiv +\infty$ in the latter, thus removing the positive term in the first equation and the negative term in the last equation.

Note that it follows from these equations that $\frac{d}{dt}\sum_{i=1}^n p_i = 0$, so $P := \sum_{i=1}^n p_i =$ constant, say, and $\frac{d}{dt}\sum_{i=1}^n q_i = P$. This in turn implies that $Q := \sum_{i=1}^n q_i = Pt + Q(0)$, thus solving the time evolution for the centre of mass Q of the system.

The Lax pair is most simply formulated in terms of **Flaschka's variables**:

$$\boxed{a_i = \tfrac{1}{2}\, e^{-(q_{i+1}-q_i)/2}\,, \quad b_i = -\tfrac{1}{2}\, p_i} \tag{13.39}$$

which satisfy

$$\begin{aligned} \dot{a}_i &= \tfrac{1}{4}\, e^{-(q_{i+1}-q_i)/2}(p_{i+1}-p_i) = a_i(b_{i+1}-b_i) \\ \dot{b}_i &= -\tfrac{1}{2}\,(e^{-(q_i-q_{i-1})} - e^{-(q_{i+1}-q_i)}) = 2(a_i^2 - a_{i-1}^2) \end{aligned} \tag{13.40}$$

(It might be objected that Flaschka's variables only encode the differences of the q_is, but given the note above, we already know their overall sum, so the differences are all that we need.)

Then the Lax pair is

$$\begin{aligned} L &= \begin{pmatrix} b_1 & a_1 & & & & \\ a_1 & b_2 & a_2 & & & \\ & a_2 & b_3 & a_3 & & \\ & & \ddots & \ddots & \ddots & \\ & & & a_{n-2} & b_{n-1} & a_{n-1} \\ & & & & a_{n-1} & b_n \end{pmatrix} \\ M &= \begin{pmatrix} 0 & a_1 & & & & \\ -a_1 & 0 & a_2 & & & \\ & -a_2 & 0 & a_3 & & \\ & & \ddots & \ddots & \ddots & \\ & & & -a_{n-2} & 0 & a_{n-1} \\ & & & & -a_{n-1} & 0 \end{pmatrix} \end{aligned} \tag{13.41}$$

(Exercise: check for yourself that $\dot{L} = [M, L] \Rightarrow$ (13.40).)

This implies that the eigenvalues of L, or equivalently the traces of the powers of L, are all conserved! This gives us n conserved quantities,

$$\boxed{Q_k = \mathrm{tr}(L^k)\,, \quad k = 1\ldots n\,.} \tag{13.42}$$

The first few are

$$
\begin{aligned}
Q_1 &= \operatorname{tr}(L) \\
&= \sum_{i=1}^{n} b_i = -\tfrac{1}{2}\sum_{i=1}^{n} p_i \qquad \text{(total momentum)} \\
Q_2 &= \operatorname{tr}(L^2) \\
&= \sum_{i=1}^{n} b_i^2 + 2\sum_{i=1}^{n-1} a_i^2 \\
&= \tfrac{1}{2}\left(\tfrac{1}{2}\sum_{i=1}^{n} p_i^2 + \sum_{i=1}^{n-1} e^{-(q_{i+1}-q_i)}\right) \qquad \text{(the Hamiltonian, or total energy)} \\
Q_3 &= \operatorname{tr}(L^3) \\
&= \sum_{i=1}^{n} b_i^3 + 3\sum_{i=1}^{n-1} a_i^2(b_i + b_{i+1}) \\
&= \tfrac{1}{8}\left(\sum_{i=1}^{n} p_i^3 - 3\sum_{i=1}^{n-1} e^{-(q_{i+1}-q_i)}(p_i + p_{i+1})\right)
\end{aligned}
\tag{13.43}
$$

Interestingly, the limit $n \to \infty$ yields the **infinite Toda lattice**, which describes an infinite number of particles on a line, and this system has solitons.

The index $i \in \mathbb{Z}$ for the infinite Toda lattice is analogous to $x \in \mathbb{R}$ for KdV, while $q_i(t) \in \mathbb{R}$ corresponds to $u(x,t) \in \mathbb{R}$. Thus space has been discretised, while time remains continuous, as does the field value. (In the ball and box model the process of discretisation goes two steps further, with both time and the field values also becoming discrete.)

The solitons of the infinite Toda lattice can be derived in a number of ways, including inverse scattering. The following turns out to be a solution, for any γ, $k > 0$:

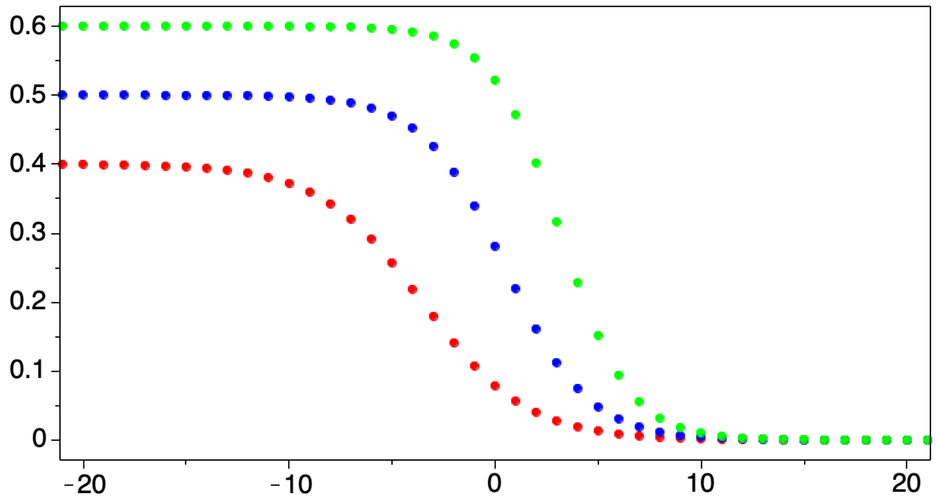
$$
\boxed{q_l(t) = q_0 - \log \frac{1 + \gamma\, e^{-2kl \pm 2\sinh(k)t}}{1 + \gamma\, e^{-2k(l-1) \pm 2\sinh(k)t}}}
\tag{13.44}
$$

This is a single soliton moving through $\mathbb{Z}$ with

$$
\begin{aligned}
\text{velocity} &= \pm\sinh(k)/k\,, \\
\text{width} &\sim 1/k\,.
\end{aligned}
\tag{13.45}
$$

As for KdV, the faster a soliton is moving, the narrower it becomes.

Here's a plot comparing three of these solitons at $t = 0$, taking the '+' option with $q_0 = 0$ in (13.44), with $(k, \gamma) = (0.2, 0.2)$ (red), $(k, \gamma) = (0.25, 1)$ (blue) and $(k, \gamma) = (0.3, 5)$ (green):



Note that the horizontal axis here is the index l, while in the sketch between equations (13.37) and (13.38) it was the 'field value' q_l.

It is also possible to find N-soliton solutions, which turn out to have a form similar to those we found earlier for the KdV equation:

$$\boxed{q_l(t) = q_0 - \log \frac{\det(\mathbb{1}_{N\times N} + C_l(t))}{\det(\mathbb{1}_{N\times N} + C_{l-1}(t))}} \tag{13.46}$$

where $\mathbb{1}_{N\times N}$ is the $N \times N$ identity matrix, and $\{C_l(t)\}$ is a family of $N \times N$ matrices which depend on the space coordinate l and the time coordinate t as follows:

$$\boxed{(C_l(t))_{ij} = \frac{\sqrt{\gamma_i \gamma_j}}{1 - e^{-(k_i+k_j)}}\, e^{-(k_i+k_j)l - (\sigma_i \sinh(k_i) + \sigma_j \sinh(k_j))t}} \tag{13.47}$$

with $k_i, \gamma_i > 0$ and $\sigma_i = \pm 1$.

The end

Bibliography

Aktosun, T. (2009). *Inverse Scattering Transform and the Theory of Solitons*, page 4960–4971. Springer New York.

Boussinesq, J. (1877). *Essai sur la théorie des eaux courantes.* Impr. nationale.

Dauxois, T. (2008). Fermi, Pasta, Ulam and a mysterious lady. *Physics Today*, 61(1):55–57. arXiv: 0801.1590.

Dauxois, T. and Peyrard, M. (2006). *Physics of Solitons.* Cambridge University Press.

Drazin, P. G. and Johnson, R. S. (1989). *Solitons: An Introduction.* Cambridge University Press.

Fermi, E., Pasta, J. R., Ulam, S. M., and Tsingou, M. (1955). Studies of nonlinear problems. *No. LA-1940. Los Alamos Scientific Lab., N. Mex.*

Gardner, C. S., Greene, J. M., Kruskal, M. D., and Miura, R. M. (1967). Method for solving the korteweg-devries equation. *Phys. Rev. Lett.*, 19:1095–1097.

Gordon, W. (1926). Der Comptoneffekt nach der Schrödingerschen Theorie. *Zeitschrift für Physik*, 40(1):117–133.

Hirota, R. (1971). Exact solution of the Korteweg—de Vries equation for multiple collisions of solitons. *Phys. Rev. Lett.*, 27:1192–1194.

John Scott Russell (1845). *Report on Waves: Made to the Meetings of the British Association in 1842-43.*

Kenig, C. E., Ponce, G., and Vega, L. (1991). Well-posedness of the initial value problem for the korteweg-de vries equation. *Journal of the American Mathematical Society*, 4(2):323–347.

Klein, O. (1926). Quantentheorie und fünfdimensionale Relativitätstheorie. *Zeitschrift für Physik*, 37(12):895–906.

Korteweg, D. D. J. and Vries, D. G. d. (1895). On the change of form of long waves advancing in a rectangular canal, and on a new type of long stationary waves. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 39(240):422–443.

Manton, N. and Sutcliffe, P. (2004). *Topological Solitons*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, Cambridge.

Miura, R. M. (1968). Korteweg-de Vries Equation and Generalizations. I. A Remarkable Explicit Nonlinear Transformation. *Journal of Mathematical Physics*, 9(8):1202–1204.

Miura, R. M., Gardner, C. S., and Kruskal, M. D. (1968). Korteweg – de Vries Equation and Generalizations. II. Existence of Conservation Laws and Constants of Motion. *Journal of Mathematical Physics*, 9(8):1204–1209.

Porter, M. A., Zabusky, N. J., Hu, B., and Campbell, D. K. (2009). Fermi, Pasta, Ulam and the birth of experimental mathematics: a numerical experiment that Enrico Fermi, John Pasta, and Stanislaw Ulam reported 54 years ago continues to inspire discovery. *American Scientist*, 97(3):214–221.

Takahashi, D. and Satsuma, J. (1990). A soliton cellular automaton. *Journal of the Physical Society of Japan*, 59:3514–3519.

Zabusky, N. J. and Kruskal, M. D. (1965). Interaction of "solitons" in a collisionless plasma and the recurrence of initial states. *Physical Review Letters*, 15(6):240.