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

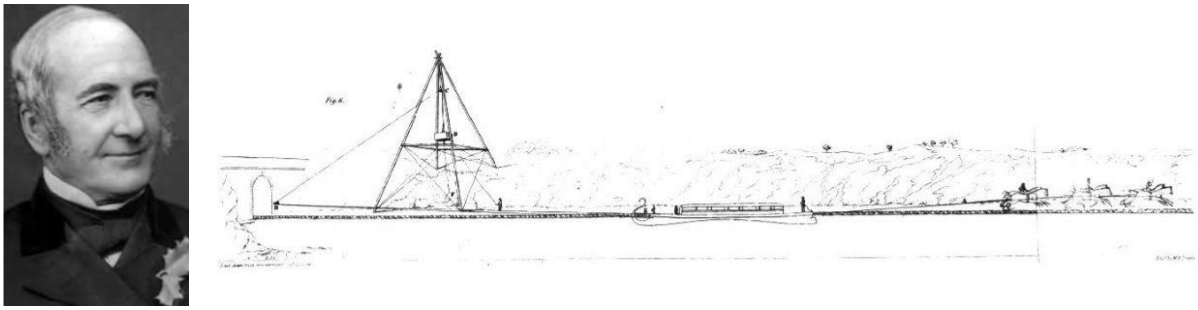


Figure 1.1: John Scott Russell, portrayed at a later time, and an artist's impression of the initial condition of his observation in 1834 (with a liberal interpretation of a 'pair of horses').

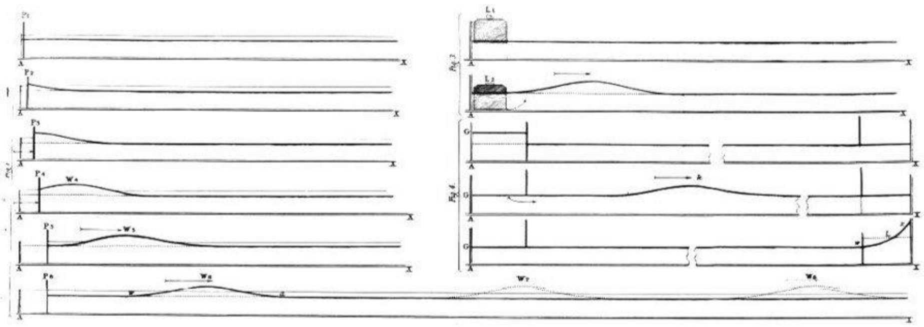


Figure 1.2: A depiction of two experiments carried out by John Scott Russell to recreate the Wave of Translation and study its properties.

> 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],

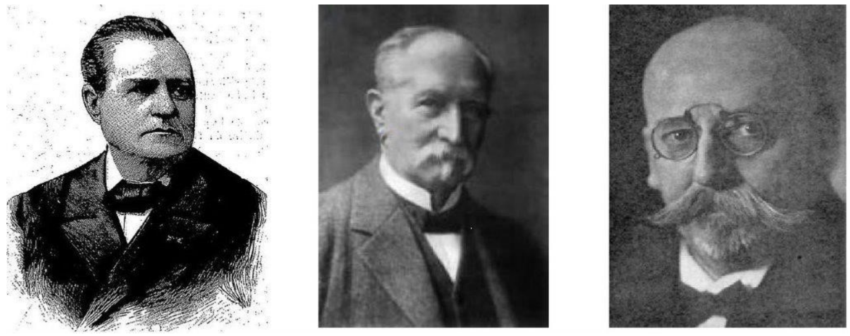


Figure 1.3: From left to right: Joseph Valentin Boussinesq, Diederik Korteweg and Gustav de Vries.

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).

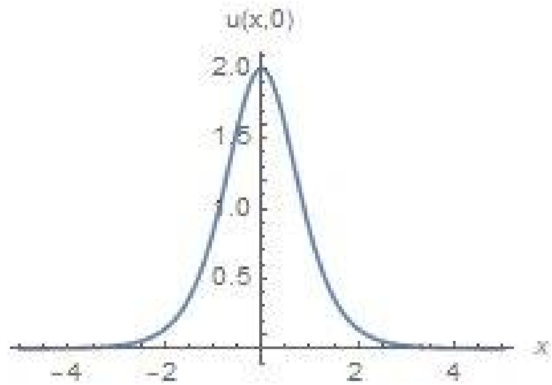


Figure 1.4: Plot of the initial condition $u(x,0) = 2\,\mathrm{sech}^2 x$ for the KdV equation.

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 $a_1u_1 + a_2u_2$ 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 $t \to +\infty$, uniformly in 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 the nonlinearity causes the wave to pile up and break after a finite time: $\max|u_x(x)| \to \infty$ 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 exactly cancel, at all points $x \in \mathbb{R}$, and we get a **"travelling wave"**, which keeps its form and just moves to the right, as you can see here.

The initial condition that we chose in (1.2) was very special. Generic solutions of KdV have a much more complicated behaviour, and indeed equations (1.3)-(1.4) and their solutions are recovered in certain limits. Let's 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

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

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 **SOLITON** 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.

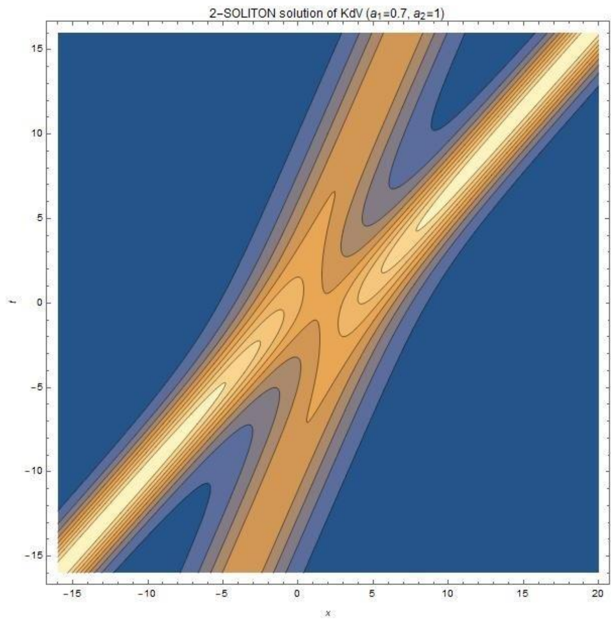


Figure 1.5: Contour plot of the energy density of two colliding KdV solitons, as a function of space and time. Lighter regions have higher energy density and correspond to the cores of the two solitons. We can see the trajectories of the two solitons and the phase shift induced by the collision: the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

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

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

- **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.

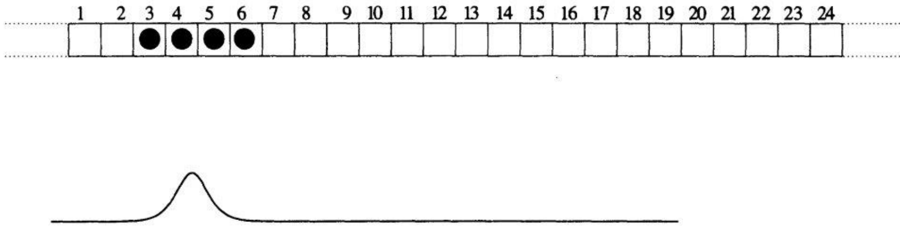


Figure 1.6: A localised configuration of the ball and box model and its continuous analogue.

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

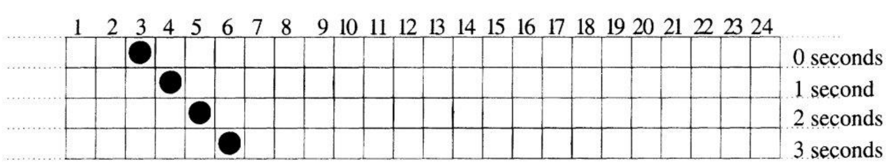
BALL AND BOX RULE:

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.

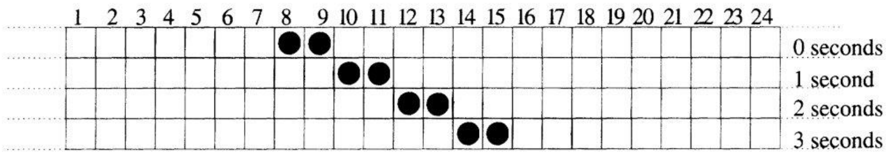
The ball and box rule plays the role of the PDE for continuous solitons, *e.g.* $u_t = -6uu_x - u_{xxx}$ in the case of the KdV equation.

EXAMPLES:

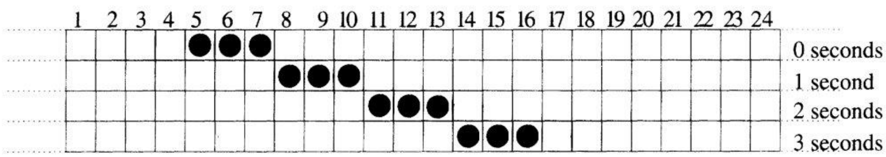
- **1 ball**:



- **2 consecutive balls**:



- **3 consecutive balls**:

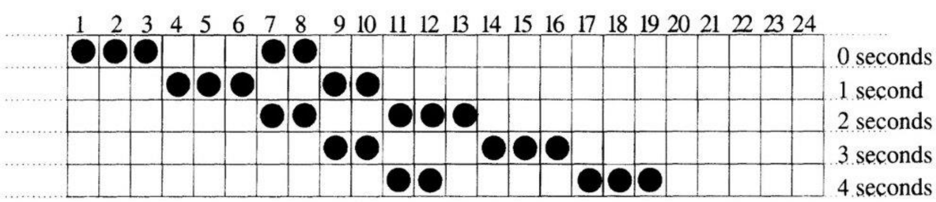


We see that a sequence of n **consecutive balls** behaves like a soliton: it keeps its shape and translates by n boxes in one unit of time. So for this class of solitons

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

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.

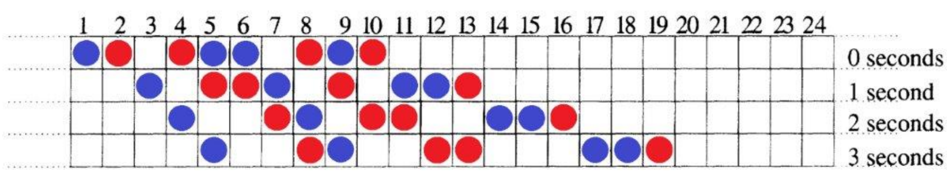
*** EXERCISE**: Generalize the previous example to a length m soliton overtaking a length n soliton (with $m > n$) and find a general rule for what happens. (Start with separation $l \geqslant n$ between the two solitons, that is, there are l empty boxes between the two solitons in the initial configuration.)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