Lecture Notes in Nonlinear Physics

Ivan Bonamassa
ivan.bms.2011@gmail.com
Dipartimento di Matematica e Fisica "Ennio de Giorgi"

January 18, 2013

Contents

1 Introduction to Soliton Theory **3**
1.1 Nonlinearity, dispersion and dissipation . . . 3
1.2 Some NLEE in $(1+1)$, $(2+1)$–dimensions . . . 5
1.3 The discovery of solitary waves . . . 9
1.4 The Fermi–Pasta–Ulam problem . . . 11
1.4.1 The solution and the discovery of solitons . . . 12

2 Miura Transform and KdV solutions by Inverse Scattering **14**
2.1 Conservation laws and Miura transform . . . 14
2.2 Exact solutions by Inverse Scattering . . . 17
2.3 Gel'fand–Levitan–Marchenko and KdV soliton solutions . . . 19
2.4 The Hamiltonian formulation . . . 21
2.4.1 KdV first integrals and (bi–)Hamiltonian structure . . . 25

3 Lax pairs and AKNS method **28**
3.1 Lax pairs for the KdV equation and for other integrable NLEEs . . . 28
3.2 The AKNS formalism and applications to the ZS spectral problem . . . 30

4 The classical Bäcklund Transformation (BT) **34**
4.1 Bianchi's Permutability Theorem . . . 36
4.2 BT in other setting and the connection with IST . . . 37
4.2.1 The connection with multi–solitons solutions . . . 42

5 The gauge–Bäcklund Transform (gBT) **44**
5.1 Gauge–Bäcklund transforms and the permutability theorem . . . 44
5.2 Applications to the ZS spectral problem . . . 47
5.2.1 Gauge–Bäcklund Transforms for the ZS spectral problem . . . 49
5.3 Permutability theorem and nonlinear superposition formula . . . 53

6 Painlevé Transcendents **54**
6.1 Lax pairs and Hamiltonian structure . . . 56
6.2 Connection with IST–solvable NLEEs . . . 58

7 Asymptotic perturbative methods and universality **60**
7.1 Asymptotic series and perturbations . . . 60
7.2 Singular perturbation methods: the Duffing equation . . . 62
7.3 The Poincaré method . . . 63
7.4 Multiple–scale method . . . 64

7.5 Multiple–scale approach to NLEEs 65
7.5.1 Nonlinear wave modulation and the NLS equation 66
7.5.2 Long–wave approximation and the KdV equation 69

A Bäcklund transform for the second Painlevé equation **I**
References III

Abstract

Intro ch. III ablowitz. As a summary.

1 Introduction to Soliton Theory

1.1 Nonlinearity, dispersion and dissipation. Wave phenomena are abundant in Physics. They could be described first as waves on strings or perhaps on water surfaces or on stretched membranes. They also could be related to the propagation of sound and electromagnetic radiation. In all these areas it's common practise to develop wave propagation concepts from the well known *d'Alembert equation*

$$\Box u(\mathbf{x},t)=0, \qquad \Box := \tfrac{1}{c^2}\partial_t^2-\partial_x^2-\partial_y^2-\partial_z^2, \tag{1.1}$$

where $u:\mathbb{R}^{3\times 1}\to\mathbb{R}$ describes the wave profile, whilst $c\in\mathbb{R}_0^+$ (e.g., c is the speed of light in vacuum for e.m. waves). Notice that (1.1) is linear. In this introductory part we'll work in $(1+1)$–dimension: we restrict to equation $(\partial_t^2-c^2\partial_x^2)u(x,t)=0$, whose general solution (expressed in variables $x\pm ct$) is

$$u(x,t)=f(x-ct)+g(x+ct), \tag{1.2}$$

being f,g (not necessary differentiable) functions, determined once initial conditions $u(x,0)$, $u_t(x,0)$ are assigned. Solution (1.2) describe two distinct waves moving in opposite directions with the same velocity c; those waves do not interact and can be overlapped in light of the **principle of linear superposition**. What follows is that these waves *don't* change their shape during the propagation[1].

Consider now unidirectional propagating waves and introduce the partial differential equation (PDE)

$$u_t+cu_x=0, \tag{1.3}$$

having invoked short–hand notation for partial derivatives and chosen $c=1$, for simplicity. Equation (1.3) is an example of *evolution equation*, in light of the following

Definition 1.1 - *An* ***evolution equation*** *is a PDE for an unknown function $u=u(x,t)$ of the form*

$$u_t=\mathcal{D}(u,u_x,u_{xx},\dots), \tag{1.4}$$

being $\mathcal{D}$ involves only u and its space derivatives. If $\mathcal{D}$ is nonlinear then equation (1.4) is called a ***nonlinear evolution equation****, hereafter abbreviated as* ***NLEE****.*

To avoid technical arguments we'll assume solutions of (1.4) to *decay sufficiently rapidly*, i.e. that $u(x,t)\in \mathrm{C}^\infty(\mathbb{R})$ and $u_{x^k}\to 0$ for $|x|\to\infty$, $\forall\, k\in\mathbb{N}_0$ (actually $k=3$ will suffice).

Let's now discuss some important examples of evolution equation. First, consider the following one

$$u_t+u_x+u_{xxx}=0, \tag{1.5}$$

which represents a **dispersive** wave equation. To show up this property, impose *harmonic wave–type* solutions of (1.5), i.e. that $u(x,t)=e^{\imath(kx-\omega t)}$; the latter solves (1.5) if and only if (iff)

$$\omega(k)=k(1-k^2), \qquad k\in\mathbb{R}, \tag{1.6}$$

where $k\in\mathbb{R}$ lets $u(x,t)$ to oscillate at $t=0$. Thus u describes a wave propagating with *phase velocity*

$$v_f:=\frac{\omega(k)}{k}=1-k^2 \qquad (k\in\mathbb{R}), \tag{1.7}$$

which means that waves with different wave number (equivalently, wave length $\lambda=2\pi/k$) propagate with different velocities. This is the fundamental property of dispersive waves. Notice that the idea can be extended, by integration, to how many components you desire; in fact

$$u(x,t)=\int_{-\infty}^{+\infty}\mathcal{A}(k)e^{i[kx-\omega(k)t]}\,\mathrm{d}\,k, \tag{1.8}$$

[1] Take the f component of (1.2) and choose the new coordinate $\xi=x-ct$: then $f=f(\xi)$ doesn't change form during propagation, for every fixed ξ; thus $f(x,t)$ has the same shape of $f(x,0)$, $\forall\, t\in\mathbb{R}^+$.

being $\mathcal{A}(k)$ the *Fourier transform* of $u(x,0)$.

Till now we have tacitly assumed the dispersion function $\omega = \omega(k)$ to be a real valued function for $k \in \mathbb{R}$. This is true as far as odd derivatives are added to equation (1.5); the picture changes dramatically if we add even derivatives. Let's try, for example, the following PDE

$$u_t + u_x - u_{xx} = 0. \tag{1.9}$$

Again, after imposing harmonic solutions of (1.9), we get the expression

$$\omega(k) = k(1 - ik) \qquad \Longrightarrow \qquad u(x,t) = e^{-k^2 t} e^{ik(x-t)}. \tag{1.10}$$

In this case $v_f(k) = 1 \; \forall \, k \in \mathbb{R}^+$ and $u = u(x,t)$ describes a wave whose profile is *damped* by the factor $e^{-k^2 t}$ as t increases. The *exponential decay* exhibited in (1.10) is usually called **dissipation**[2].

Finally, let's discuss a key concept in wave propagation, namely **nonlinearity**. For this purpose, consider the NLEE

$$u_t + (1 + u)u_x = 0. \tag{1.11}$$

A comparison of the latter with equation (1.3) suggests that the nonlinearity uu_x could be obtained by replacing c with $(1 + u)$ in the solution. Actually, after solving equation (1.11) (e.g., with the *method of characteristics*), the suggested substitution holds true and the general solution is [DrJ993]

$$u(x,t) = f\big[x - (1 + u)t\big], \tag{1.12}$$

where f is ad arbitrary function. The problem is then to solve equation (1.12) for u once the initial wave profile $u(x,0)$ is assigned. This may be a difficult task since solution of equation (1.12) (with $f > 0$ for some $x \in \mathbb{R}$) is a *single–valued* function for a *finite time*. In fact, it exists a time at which the solution exhibits non–uniqueness as a wave which has ""broken" (see Figure 1). To fix this problem one usually insert a *discontinuity* in the solution, simulating a *shock*.

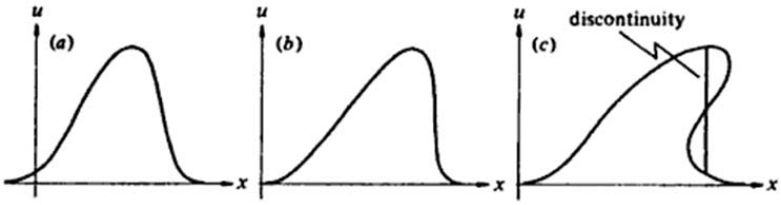


Figure 1: Temporal evolution of a nonlinear wave profile. (**a**) $t = t_1$; (**b**) $t = t_2 > t_1$; (**c**) $t = t_3 > t_2$. The wave becomes vertical at $t = t_2$ and thereafter the solution is three–valued in a region [DrJ993].

Another complication is that the principle of linear superposition *doesn't generally* hold true for nonlinear equations[3]. However, we'll see that a related principle *can* be formulated for certain NLEE.

Of course one can imagine to combine nonlinearity with dissipative and/or dispersive wave properties. So, for example, we might derive NLEE like the following ones:

$$u_t + (1 - u)u_x + u_{xxx} = 0 \qquad \text{or} \qquad u_t + (1 - u)u_x - u_{xx} = 0.$$

The first one is known as *Kortweg–de Vries equation*[4] (hereafter abbreviated as KdV equation) and embodies dispersion, whilst the second one, known as *Burgers equation*, has dissipative properties. Our main concern will be with the method of solution (and the properties of) the KdV equation and other related ""exactly integrable" equations, briefly discussed in the next section.

[2] Notice that a linear combination of odd and even derivatives could describes both dispersive and dissipative waves.
[3] Verify that, given solutions $u_1 = u_1(x,t)$, $u_2 = u_2(x,t)$ of (1.11), then $u \equiv u_1 + u_2$ doesn't solve equation (1.11).
[4] To simplify notation, we'll drop from now on the saxon genitive in writing equations' names.

1.2 Some NLEE in (1+1), (2+1)–dimensions. We start with evolution equation in (1+1)–dimensions.

• Burgers Equation

$$u_t + (1+u)u_x - \nu u_{xx} = 0, \qquad \nu \in \mathbb{R}_0^+. \tag{1.13}$$

Equation (1.13) is a dissipative NLEE deduced in 1906 by J. M. Burgers (1895–1981). It is widely used in Physics, e.g., to describe gas dynamics and traffic flux. The parameter $\nu \in \mathbb{R}_0^+$ is related to the *viscosity* of the fluid of interest (case $\nu = 0$ is called *inviscid Burgers equation*). Equation (1.13) can be written in a more common form: using the transformation $\alpha\phi \mapsto 1 + u$, $\alpha \in \mathbb{R} \setminus \{0\}$, we get

$$\phi_t + \alpha\phi\phi_x - \nu\phi_{xx} = 0. \tag{1.14}$$

When solving the nonlinear initial value problem, equation (1.14) can be linearized by choosing

$$\phi \equiv -2\,\nu\frac{\varphi_x}{\varphi}, \tag{1.15}$$

known as *Hopf–Cole transformation*. Applying (1.15), one recovers the *heat equation*:

$$\phi_t + \phi\phi_x = \nu\phi_{xx} \qquad \Longrightarrow \qquad \frac{\varphi_x}{\varphi}\left(\varphi_t - \nu\varphi_{xx}\right) = \left(\varphi_t - \nu\varphi_{xx}\right)_x,$$

$\alpha = 1$ for simplicity. Then if $\varphi = \varphi(x,t)$ solves the heat equation, $\phi = \phi(x,t)$ solves Burgers one.

• Korteweg–de Vries equation (KdV)

$$u_t + (1+u)u_x + u_{xxx} = 0 \qquad \xrightarrow{\alpha\phi \mapsto 1+u,\ \alpha\in\mathbb{R}\setminus\{0\}} \qquad \phi_t + \alpha\phi\phi_x + \phi_{xxx} = 0. \tag{1.16}$$

Equation (1.16) was introduced in 1895 by D. J. Korteweg (1848–1941) and G. de Vries (1866–1934) to describe the phenomenon observed by Lord S. Russell in 1834. The KdV equation has many physical applications: for example, in can be used to describe the propagation of unidirectional waves in *shallow waters*, as the continuum limit of anharmonic non–linear lattices, for ion–acoustic and magnetohydrodynamic waves in cold plasma [New985]. The KdV equation is a dispersive NLEE; also (as we'll see soon) it is *Galilei–invariant*, it admits *infinite non–trivial constants of motion* and has *soliton solutions*, as N. J. Zabusky (1929) and M. D. Kruskal (1925–2006) discovered in 1965 [ZaK965]. For now, let's point out the following property about uniqueness of KdV solutions [Kar998]

Lemma 2.1 - *KdV solutions that decay sufficiently rapid are uniquely determined by initial data.*

Proof. Let u, v be two different KdV solutions belonging to the same initial condition; let then be $w = u - v$. Substituting in (1.16) one finds $w_t + \alpha(uw_x + wv_x) + w_{xxx} = 0$. Multiplying both size of the latter by w, integrating by parts and using the fact that w is a rapidly decreasing function, we obtain

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}\,t}\int_{\mathbb{R}} w^2\,\mathrm{d}\,x + \alpha\int_{\mathbb{R}}\left(v_x - \tfrac{1}{2}u_x\right)w^2\,\mathrm{d}\,x = 0.$$

Calling $E(t) \equiv \frac{1}{2}\int_{\mathbb{R}} w^2\,\mathrm{d}\,x$ and $M \equiv \sup|\alpha(v_x - \frac{u_x}{2})| < \infty$ we get $E(t) \le E(0)e^{-\alpha M t}$. Since $E(0) = \frac{1}{2}\int_{\mathbb{R}}\omega^2(0,t)\,\mathrm{d}\,x = 0$ we have $E(t) = 0$ and then $w = 0$, so $u(x,t) = v(x,t)\ \forall\, t \in \mathbb{R}_0^+$. □

There are many variants of the KdV equation; among others we point out the following one

$$\phi_t - \alpha\phi^2\phi_x + \phi_{xxx} = 0, \tag{1.17}$$

known as *modified KdV* (mKdV). In particular, equation (1.17) is called *focusing* (*defocusin*) mKdV if $\alpha > 0$ ($\alpha < 0$). In §2.1 we'll see that mKdV and KdV are related: in fact, it will be shown that if ϕ is a solution of mKdV with $\alpha = 6$, then (and only then) $\psi \equiv \phi_x + \phi^2$ is a solution of KdV (with $\alpha = 6$).

• Sine–Gordon equation (SG)

$$\phi_{tt} - \phi_{xx} + \sin\phi = 0 \qquad \xrightarrow{u \mapsto \frac{1}{2}(x+t),\ v \mapsto \frac{1}{2}(x-t)} \qquad \phi_{uv} = \sin\phi. \tag{1.18}$$

Equation (1.18) is an example of a *hyperbolic*[5] NLEE, whose name comes from the *Klein–Gordon* (KG) *equation* $(\Box + \mu^2)\psi = 0$. Notice that equations in (1.18) are equivalent: the first one is expressed in spacetime coordinates (x, t), whilst the second one in light–cone coordinates (u, v). The latter is the original form derived, independently, by L. Bianchi (1856–1928) and A. V. Bäcklund (1845–1922) when studying auto–transformations of *pseudospherical surfaces* (i.e., surfaces with negative curvature constant); their works conduct to the introduction of the so called *Bäcklund (auto)–transformations*, whose details will be discussed later (see §4). The SG equation admits soliton solutions.

Among many generalizations of the SG equation, we mention the *shine–Gordon equation* (ShG)

$$\phi_{tt} - \phi_{xx} + \sinh\phi = 0, \tag{1.19}$$

which shares all properties satisfied by SG with the exception of the existence of soliton solutions.

• Boussinesq equation

$$\phi_{xx} - \phi_{tt} + 6\left(\phi^2\right)_{xx} + \phi_{x^4} = 0. \tag{1.20}$$

This dispersive NLEE was introduced by J. V. Boussinesq (1842–1929) in 1872 in response to Russell's observations [Bou871]. His basic idea was to eliminate the vertical component of the *velocity flux*[6], studying only the planar propagation. This is known as *Boussinesq approximation* and it can be applied to the flux with performing a Taylor expansion of the velocity potential $\varphi = \varphi(x, z; t)$ around, e.g., the bed level $z = -h$, being h the mean water depth. For an *incompressible*, *conservative* flux, one has

$$\varphi = \varphi_f - \tfrac{1}{2}(z+h)^2\left(\varphi_f\right)_{xx} + \tfrac{1}{24}(z+h)^4\left(\varphi_f\right)_{x^4} + \dots, \tag{1.21}$$

where $\varphi_b(x; t) \equiv \varphi(x, -h; t)$. Then the expansion is truncated to a *finite* number of terms. Taking into account the boundary conditions of the wave profile at the free surface elevation $z = \eta(x, t)$

$$\begin{cases} \eta_t + u\eta_x - w = 0, \\ \varphi_t + \frac{1}{2}\left(u^2 + w^2\right) + g\eta = 0, \end{cases} \qquad u \equiv \varphi_x, \quad w \equiv \varphi_z,$$

and applying approximation (1.21), we gets finally the *coupled Boussinesq equations*

$$\begin{cases} \eta_t + \left[(h+\eta)u_b\right]_x = \frac{h^3}{6}\left(u_b\right)_{xxx}, \\ \left(u_b\right)_t + u_b\left(u_b\right)_x + g\eta_x = \frac{h^2}{2}\left(u_b\right)_{xxt}, \end{cases} \tag{1.22}$$

having defined with u_b the horizontal component of the velocity flux at the bed. Equations (1.22) holds iff h is a constant independent of position x; they also reduce to the so called *shallow water equations* with the additional assumption $(u_b)_{xxx} = 0 = (u_b)_{xxt}$. Other approximations allow to reduce equations (1.22) in a single NLEE (conveniently normalized), formally equivalent to expression (1.20).

Many wave phenomena are well described by the Boussinesq equation. Furthermore, under suitable assumptions, equation (1.20) reduces to KdV for waves on the line, to the *Kadomtsev–Petviashvili* equation (1.25) for planar waves and to the *nonlinear Schrödinger* equation (1.23) for narrowband waves.

[5] Second order PDEs have form $A\psi_{xx} + 2B\psi_{xy} + C\psi_{yy} + D\psi_x + E\psi_y + F = 0$ and can be classified like conics: assuming $\psi \in \mathrm{C}^2(\mathbb{R})$, if $B^2 - AC \lesseqgtr 0$ equations are called *elliptic* (level off noises, e.g. Laplace equation), *parabolic* (maintain noises, without magnifying them, e.g. heat equation) and *hyperbolic* (amplify noises, e.g. d'Alembert equation).

[6] We remind that fluid properties can be specified by flux velocity's ones: defining with $\boldsymbol{\mathcal{J}} : \Omega \subseteq \mathbb{R}^3 \to \mathbb{R}^3$ (Ω an open set) the flux vector field, we'll say the fluid is *stationary* iff $\boldsymbol{\mathcal{J}}_t = \mathbf{0}$, *incompressible* iff $\nabla \cdot \boldsymbol{\mathcal{J}} = 0$ (i.e. iff $\boldsymbol{\mathcal{J}}$ is solenoidal) and *irrotational* iff $\nabla \wedge \boldsymbol{\mathcal{J}} = \mathbf{0}$. Notice that if $\boldsymbol{\mathcal{J}}$ is irrotational and Ω is simply connected, then $\boldsymbol{\mathcal{J}}$ is conservative and can be written as gradient of a given *velocity potential* $\varphi : A \subseteq \mathbb{R}^3 \to \mathbb{R}$, whilst if $\boldsymbol{\mathcal{J}} = \nabla\varphi$, then $\boldsymbol{\mathcal{J}}$ is irrotational.

• Nonlinear Schrödinger equation (NLS)

$$\psi_{xx} + i\psi_t + \alpha\,|\psi|^2\,\phi = 0. \tag{1.23}$$

Unlike linear Schrödinger equation, the NLS never describes the evolution of a quantum state. As we shall see, the NLS is *completely integrable* [ZaM974] and can be solved by means of the *Inverse Scattering* (or *Spectral*) *Transform* (IST), which allows to transform equation (1.23) into a set of linear equations known as *Zakharov–Shabat system* (ZS); from this property we'll show that NLS has *soliton solutions*. The NLS is also *Galilei–invariant*, i.e. if $\phi = \phi(x,t)$ solves (1.23), also $\phi_{[v]}(x,t) \equiv \phi(x-vt,t)e^{-\frac{iv}{2}(2x+vt)}$ does. The NLS has many physical applications: in Nonlinear Optics, for example, it models many nonlinear effects (due to the propagation of light pulses) in fiber optics like self–trapping, auto–modulations of monochromatic waves or stimulated–Raman dispersion; it describes the propagation of Langevin waves in plasma, of heat pulses in solids and it thought to be a good model for explaining the formation of *killer waves*. Also it's gauge equivalent to the so called isotropic *Landau–Lifšitz equation* (also known as Heisenberg ferromagnet equation) $\mathbf{S}_t = \mathbf{S} \wedge \mathbf{S}_{xx}$, where $\mathbf{S}$ is the spin vector operator.

Notice that expression (1.23) for NLS is not unique in literature. There exist different forms for NLS, each one ascribable to a perturbation of the free particle linear Schrödinger equation. A general class of such equations, known as *derivative–NLS* (d–NLS) is of the form (in $(3+1)$–dimensions)

$$i\psi_t + \nabla^2\psi = f\big(\psi, \psi^*, \psi_{x,y,z}, \psi^*_{x,y,z}\big), \tag{1.24}$$

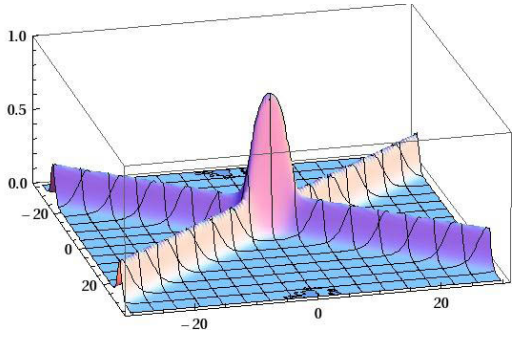
where f is an analytic function of ψ, its spatial derivatives $\psi_{x,y,z}$ and their complex conjugates. Expression (1.23) has a relevant role among equations (1.24): in fact, we'll see that it is possible to put a generic dispersive NLEE in NLS form under (an appropriate choice of) a *multiple scale limit*[7].

We close this section with a brief discussion about two important NLEE in $(2+1)$–dim.

• Kadomtsev–Petviashvili equation (KP)

$$\left(\phi_t + 6\phi\phi_x + \phi_{xxx}\right)_x + 3\sigma^2\phi_{yy} = 0, \qquad \sigma = (\imath, -1). \tag{1.25}$$

These two equations where originally written in 1970 by B. B. Kadomtsev (1928–1998) and V. Petviashvili (1936–1993) with the aim of describing ion–acoustic wave propagation in plasma in presence of strong, transversal perturbations. Equations (1.25) generalize KdV in two spatial dimensions, the latter being recovered for no transversal dynamics, i.e. if $\phi_{yy} = 0$. It's common choice to distinguish equations (1.25) in KPI for $\sigma = \imath$ and KPII for $\sigma = 1$. KP equations represent (*universal*) *completely integrable* bidimensional systems (they can be solved with IST), meaning that many bidimensional integrable systems can be obtained as particular cases of KP; they also admit *n–soliton solutions*. Next figure shows a 2–solitons interaction from KPII (left) and a picture of sea waves interaction taken in Oregon (right).



[7] Although we don't have a general theorem, till now there are no counterexamples to this "universal" property.

Furthermore KP equations has *periodic solutions* such as *cnoidal waves*[8], i.e. periodic unidimensional *stable* solutions (see Figure 2), or *hexagonal waves solutions*, which correspond to biphasic solutions characterized by an hexagonal periodic cell structure (see Figure 3).

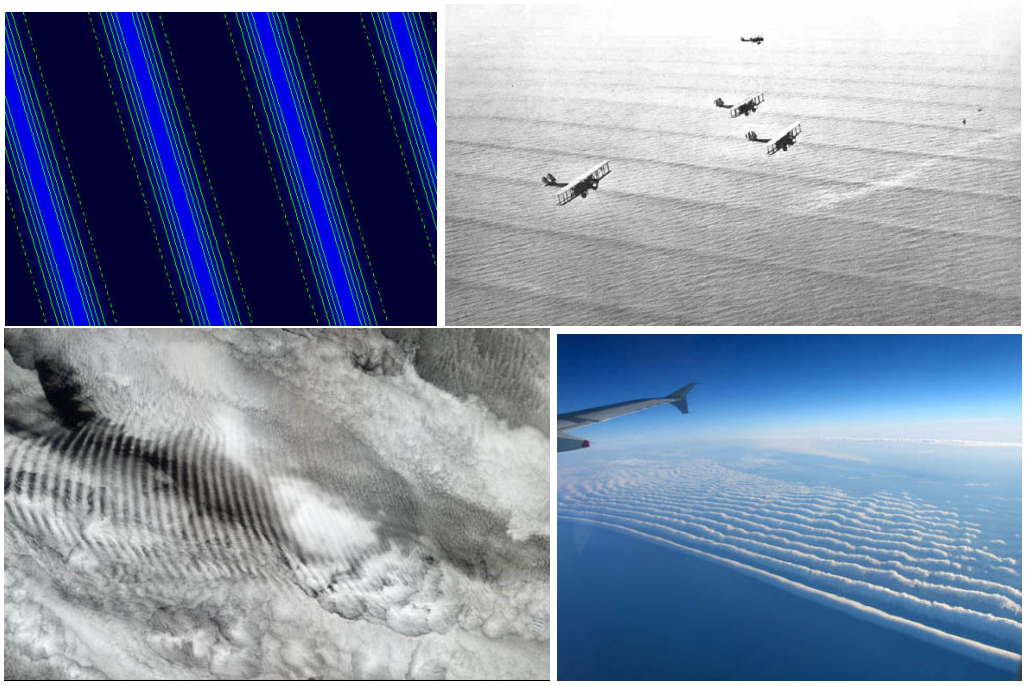


Figure 2: KP cnoidal wave solution (top left), cnoidal waves in shallow waters near the Panama coast (top right); cloudy cnoidal waves over Atlantic (bottom left) and Indian (bottom right) Ocean.

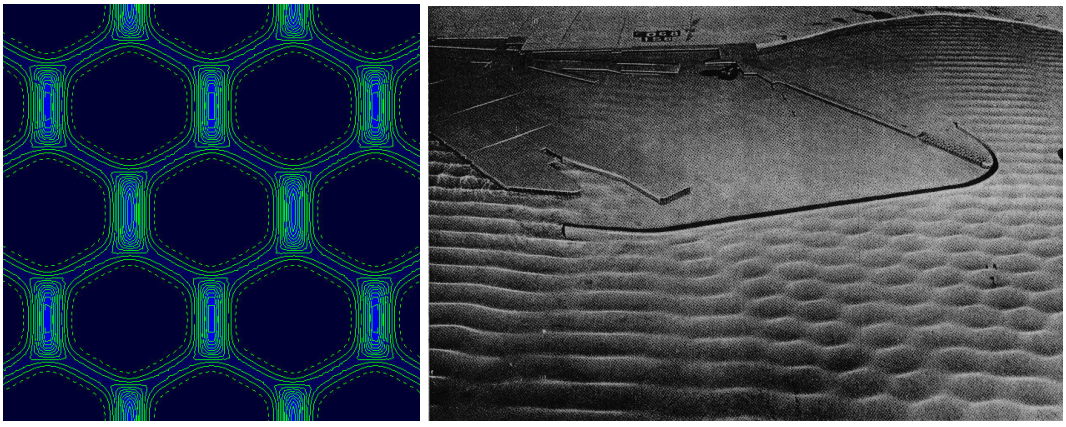


Figure 3: KP biphasic solution (left), interaction of two wave trains with the walls of an harbor model (right).

[8]The term ""cnoidal" originates from the symbol ""cn" of a special class of *elliptic functions* introduced by C. G. J. Jacobi (1804–1851) in 1829. These functions can be defined by means of their Lambert series expansion as [AbS972]

$$cn(x) = \frac{2\pi}{\mathcal{K}\sqrt{m}} \sum_{n=1}^{\infty} \frac{q^{n+1/2}}{1+q^{2n+1}} \cos\left[\frac{\pi x}{2\mathcal{K}}(2n+1)\right], \qquad q := e^{-\frac{\pi\mathcal{K}'}{\mathcal{K}}}, \quad \mathcal{K}(m) := \int_0^{\pi/2} \frac{\mathrm{d}\,\theta}{\sqrt{1-m\sin^2\theta}}, \quad \imath\mathcal{K}'(m) := \imath\mathcal{K}(1-m),$$

where m is a parameter, whilst $q, \mathcal{K}, \mathcal{K}'$ are *special functions* known as *nome*, *quarter real period*, *quarter imaginary period*.

• Davey–Stewartson equation (DS)

$$\begin{cases} \imath u_t + \frac{1}{2}\left(\sigma^2 u_{xx} + u_{yy}\right) + \alpha\,|u|^2\,u - \rho\phi u = 0, \\ \phi_{xx} - \sigma^2\phi_{yy} = 2\alpha\big(\,|u|^2\,\big)_{xx} = 0, \end{cases} \qquad \sigma = (1,\imath), \quad \alpha = (1,-1), \quad \rho \in \mathbb{R}. \tag{1.26}$$

Introduced by D. J. Benney and G. J. Roskes in 1969 to describe the propagation of tridimensional water waves with finite depth [BeR969], re–derived and written in the form (1.26) by A. Davey and K. Stewartson in 1974 [DeS974], DS equations are a system of coupled NLEE for a complex field $u = u(x,y;t)$ (wave profile) and a real field $\phi = \phi(x,y;t)$ (mean–flow). As for the KP equations, DS equations are called DSI for $\sigma = 1$ and DSII for $\sigma = \imath$; furthermore the parameter α distinguishes DS(I,II) in *focusing* (or *attractive*) for $\alpha = 1$ and *de–focusing* (or *repulsive*) for $\alpha = -1$. The parameter ρ in (1.26) is defined by the field of application of DS: $\rho < 0$ in optics, whereas $\rho > 0$ for water waves. Notice that the system is decoupled when $\rho = 0$ and can be reduced to the classical NLS equation.

Interestingly, DSI and DSII equations *integrable* by IST [AbC991] and (as the NLS equation) have *universal* character. Notice also that DS(I,II) equations admit four types of *soliton solutions*: *linear solutions*, characterized by a unidimensional structure, *algebraic*, *periodic* and *reticular solutions*, characterized instead by a *bidimensional localized structure*, i.e. exponentially decreasing in all directions. The existence of such solutions, i.e. of localized multi–dimensional solitons (also known as **dromions**), was proved for the first time in 1988 by the group Boiti–Léon–Martina–Pempinelli [BLMP88, BMP993], using DS as mathematical model. This result open the way to new applications in Nonlinear Physics, showing that solitons are not just a unidimensional phenomenon.

1.3 The discovery of solitary waves. In previous sections some important NLEEs were introduced. Here we'll give a brief historical introduction to the KdV equation and its solitary wave solution. First let's describe some empirical properties of solitary waves, as they were collected by Russell in reproducing the observed phenomenon in the laboratory [DrJ993]. With simple instruments, he deduced the following key properties: *i*) solitary waves are *stable* and can travel over very long distances (normal ones tend to flatten out); *ii*) they never merge after a superposition and re–emerge with the same shape, but translated along the spatial coordinate[9]; *iii*) with reference to Figure 4, if the amplitude a of the wave is greater than undisturbed water depth h, then it splits in two waves with amplitudes a_1, a_2 (with $a_1 \neq a_2$) s.t. $a = a_1 + a_2$; *iv*) the speed c of the wave depends on a and h through the relation

$$c^2 = g(h + a), \tag{1.27}$$

being g the gravity acceleration. Thus solitary waves are *gravity waves* and higher waves travel faster.

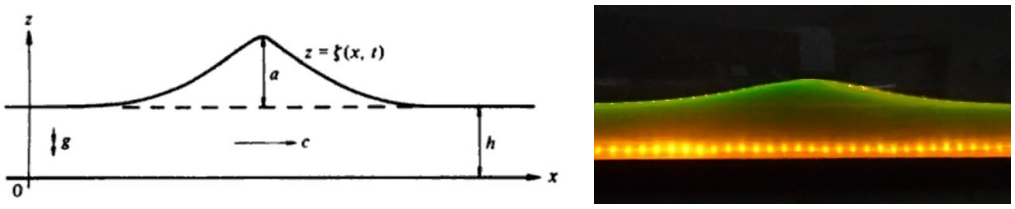


Figure 4: Parametrization of a solitary wave (left) and real model experiment (right).

To put formula (1.27) on a firmer footing, both Boussinesq (1871) and Lord Rayleigh (1876) started from the equations for an inviscid, incompressible fluid and, assuming a solitary wave to have length scale much greater than water depth (i.e., $a \ll h$), they showed that the wave profile $z = \zeta(x;t)$ is

$$\zeta(x,t) = a\,\mathrm{sech}^2\big[\beta(x - ct)\big], \qquad \beta^2 \equiv \frac{3a}{4h^2(h+a)}, \quad \forall a > 0. \tag{1.28}$$

[9] In other words, they don't re–emerge in the same position as they will do if they moved at constant velocity.

However they didn't find an equation for $z = \zeta(x,t)$ having (1.28) as solution. This step was completed in 1895 by Kortweg and de Vries: they proved that the wave profile satisfies the NLEE [DrJ993]

$$\zeta_t = \frac{3}{2}\left(\frac{g}{h}\right)^{\frac{1}{2}}\left(\frac{2}{3}\epsilon\zeta_\chi + \zeta\zeta_\chi + \frac{\sigma}{3}\zeta_{\chi\chi\chi}\right), \qquad \chi \equiv x - (gh)^{\frac{1}{2}}\left(1 - \frac{\epsilon}{h}\right)t, \quad \sigma \equiv \frac{h^3}{3} - \frac{Th}{g\rho}, \tag{1.29}$$

where ϵ is arbitrary, σ incorporates the surface tension T of a fluid with density ρ and χ is a coordinate chosen to be moving (almost) with the wave. Notice that $|\epsilon|, |\sigma| \ll 1$. Using then the new variable $\chi \equiv \chi + \epsilon\,(g/h)^{1/2}\,t$, equation (1.29) gets the usual KdV–form $\zeta_t = \gamma\zeta\zeta_\chi + \delta\zeta_{\chi\chi\chi}$, where $\gamma \equiv \frac{3}{2}\left(\frac{g}{h}\right)^{\frac{1}{2}}$, $\delta \equiv \frac{\gamma\sigma}{3}$.

Let's now verify that expression (1.28) is a solitary wave solution of the KdV equation. To do this, suppose that the solution of (1.29) is stationary in the χ–frame, then $\zeta = \zeta(\chi)$ and so

$$\tfrac{2\epsilon}{3}\zeta_\chi + \zeta\zeta_\chi + \tfrac{\sigma}{3}\zeta_{\chi\chi\chi} = 0, \tag{1.30}$$

which is an ordinary differential equation (ODE) in χ[10]. Direct integration of equation (1.30) gets $\frac{\sigma}{3}\eta_{\chi\chi} + \frac{1}{2}\zeta^2 + \frac{2\epsilon}{3}\zeta = a$, being a an integration constant; multiplying by ζ_χ and integrating once more, one finds $\sigma(\zeta_\chi)^2 + \zeta^3 + 2\epsilon\zeta^2 = 6(a\zeta + b)$, for another constant of integration b. This equation can be solved using elliptic integrals, but we're considering rapidly decaying solutions, then $a, b = 0$ and so

$$\sigma\big(\zeta_\chi\big)^2 + \zeta^3 + 2\epsilon\zeta^2 = 0.$$

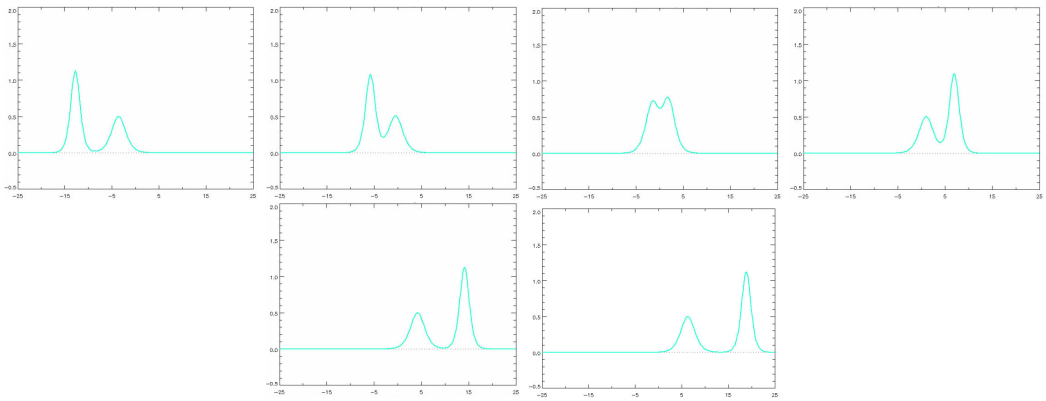
This equation can be integrated once again [DrJ993], but it's simpler to verify directly that $\zeta(\chi) = a\,\mathrm{sech}^2(\beta\chi)$ solves it: after substitution, we find that $a = 4\sigma\beta^2$ and $\epsilon = -2\sigma\beta^2$. Having then in mind the definition of the moving variable χ, the solitary wave–type solution becomes

$$\zeta(x,t) = a\,\mathrm{sech}^2\left\{\frac{1}{2}\left(\frac{a}{\sigma}\right)^{\frac{1}{2}}\left[x - (gh)^{\frac{1}{2}}\left(1 + \frac{a}{2h}\right)t\right]\right\}. \tag{1.31}$$

This agrees with equations (1.27) and (1.28) iff we neglect surface tension so that $\sigma = \frac{1}{3}h^3$) and assume that the water depth is much bigger than the maximum wave amplitude (i.e. $a/h \ll 1$), for then

$$c \sim (gh)^{1/2}\Big(1 + \frac{1}{2}\frac{a}{h}\Big), \qquad \beta \sim \frac{1}{2}\Big(\frac{3a}{h^3}\Big)^{1/2}.$$

Thus velocity is proportional to the wave amplitude and width (defined as the distance between points of height $a/2$, say) is inversely proportional to $\sqrt{a}$, i.e. taller waves travel faster and are narrower.



We are left with stability properties of KdV solitary wave solutions; in fact, we need to explain why they re–emerge with the same form, up to a *phase change*, after interactions (see figure). The first intuition there was something unusual about this phenomenon came in 1953, in Los Alamos, where E. Fermi (1901–1954), J. Pasta (1918–1984) and S. Ulam (1909–1984) were working on a new problem.

[10] Equivalently, we can deduce equation (1.30) directly from (1.16). In fact, stationary wave solutions have form $u(x,t) = f(x - ct)$: substituting this in (1.16) one gets the ODE $f''' + (\alpha f - c)f' = 0$, which is nothing but equation (1.30).

1.4 The Fermi–Pasta–Ulam problem. During the summer of 1953, at the Los Alamos National Laboratory, E. Fermi, J. Pasta and S. Ulam[11] where performing some simulations for the oscillations in nonlinear lattices, in order to study the thermalization process of a solid [FPU955]. In particular, they wish to analyze how a system with many degrees of freedom evolves from a near–equilibrium state towards *thermalization* after adding a weak (nearest neighbor) *nonlinear interaction*.

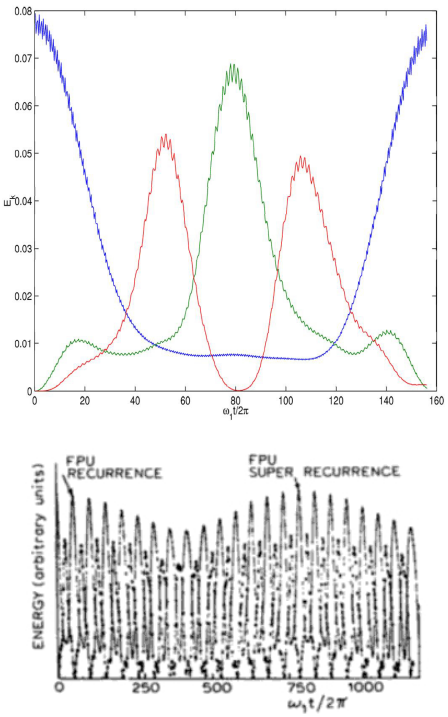
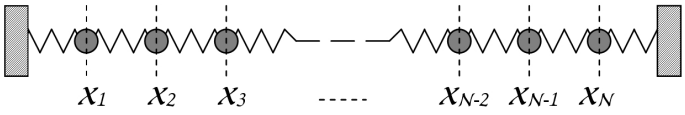


Figure 5: k–mode energy $E_k = E_k(t)$ (equation (1.35)) for the three lowest normal modes [DPR005] (up); super–recurrence of $k = 1$ mode [TuM972] (bottom).

Their original idea was to simulate the one–dimensional analogue of atoms in a crystals: a long chain of linear oscillators with a weak nonlinear correction (quadratic for the FPU–α and cubic for the FPU–β one). In light of the *equipartition theorem* (a consequence of *ergodicity*), Fermi, Pasta and Ulam thought that, due to nonlinear correction, the energy introduced into the lowest normal mode (i.e. $k = 1$) should have slowly drifted to other modes, eventually reaching thermalization [DaR008]. First calculations indeed suggested that this was the case; however, letting the program run longer, they discovered (with great surprise) that the system departed from the near–equipartition state, evolving towards a state where almost all the energy (up to 3%) was back to mode $k = 1$ (see Figure 5, up). Further investigations, performed later with faster computers, showed that the same phenomenon repeats many times and that a *super–recurrence* occurs at very late times, at which the 100% of the initial energy is recovered by the initial mode (see Figure 5, bottom). Therefore, the system behaved in a completely different way with respect to what was expected. This highly remarkable result, known as the ""*FPU paradox*", shows that nonlinearity is not enough to guarantee the equipartition of energy [DaR008].

In order to show how the FPU paradox was solved, let's introduce the system more formally[12]. Let ℓ be the length of a string of mass M, aligned to the x–axis with origin fixed in the left extreme of the string itself. Divide then the interval $[0, \ell]$ in $N + 1$ segments of length $h \equiv {}^{\ell}/_{N+1}$, each one reduced to a point with mass $m \equiv {}^{M}/_{N+2}$; define $x_n := nh$, $n = 1, 2, \dots, N$, the equilibrium position of the n–th oscillator so that $x_0 \equiv 0$ and $x_{N+1} \equiv \ell$ are the coordinates of the string's endings (see figure).



We add now a small perturbation to the system. Let's then identify by η_n the displacement of the n–th oscillator from the equilibrium position (such that $\eta_0 = 0 = \eta_{N+1}$) and define with $x'_n(t) \equiv x_n + \eta_n(t)$ the position of the mass point ad time t; also, assume the (nearest neighbor) interacting potential of the form $U(\delta_n) = \frac{1}{2}k\delta_n^2 + \frac{1}{3}k\beta\delta_n^3$, where $\beta(\ll 1)$ is a coefficient measuring the intensity of the nonlinear contribution and $\delta_n \equiv [(x'_{n+1} - x'_n) - h] = (\eta_{n+1} - \eta_n)$. If we chose as generalized coordinated the

[11] With the fundamental contribution of M. Tsingou (1928), one of the very first computational physicist [Dau008].
[12] For matter of choice, we're going to consider the FPU–β model; similar arguments follows for the α one.

(mutually independent) displacements $\{\eta_n\}_{n=1,2,\dots,N}$, we finally can write the lagrangian

$$L = \frac{m}{2}\sum_{n=1}^{N}\dot{\eta}_n^2 - \frac{k}{2}\sum_{n=1}^{N}(\eta_{n+1}-\eta_n)^2 - \frac{k\beta}{3}\sum_{n=1}^{N}(\eta_{n+1}-\eta_n)^3 . \tag{1.32}$$

Consequently, the Euler–Lagrange equation of motion for the n–th oscillator is

$$m\ddot{\eta}_n = k(\eta_{n+1}+\eta_{n-1}-2\eta_n)\big[1+\beta(\eta_{n+1}-\eta_{n-1})\big], \qquad \eta_0(t) = 0 = \eta_{N+1}(t),\ \forall\, t\in\mathbb{R}_0^+. \tag{1.33}$$

Being interested in studying the lattice motion starting from the equilibrium state, Fermi, Pasta and Ulam imposed the additional condition $\dot{\eta}_n(0) = 0$ for $n = 1, 2, \dots, N$. The Cauchy problem for equation (1.33) is then completely determined by the initial displacements $\{\eta_n(0)\}_{n=1,2,\dots,N}$ of the N oscillators.

It has to be noted that in absence of the nonlinear interaction (i.e. for $\beta = 0$), the solutions of (1.33) can be written in terms of the normal modes[13] ($m = k = 1$ for sake of simplicity)

$$Q_k(t) = \sqrt{\tfrac{2}{N+1}}\sum_{n=1}^{N}\eta_n(t)\sin\left(\frac{nk\pi}{N+1}\right), \qquad k = 1, 2, \dots, N, \tag{1.34}$$

with frequencies and energies respectively given by [For992]

$$\omega_k = 2\sin\left(\frac{\pi}{2}\frac{k}{N+1}\right), \qquad E_k = \frac{1}{2}\left(\dot{Q}_k^2 + \omega_k^2 Q_k^2\right). \tag{1.35}$$

Since the system is composed by N uncoupled harmonic oscillators, the energies $\{E_k\}_{k=1,2,\dots,N}$ are constant of motion for every pair of initial condition and so no thermalization process occurs. If instead the β is small but not zero, then equation (1.33) can be written in mode representation as

$$\ddot{Q}_k + \omega_k^2 Q_k = \beta\sum_{i,j=1}^{N} C_{ij}Q_iQ_j\omega_i\omega_j, \tag{1.36}$$

where the coefficients C_{ij} define the complicate dependence of the motion from the nonlinear interaction. The presence of the last term led Fermi, Pasta and Ulam to think that energy would have slowly drifted to other modes, reaching thermalization[14].

1.4.1 The solution and the discovery of solitons. The solution of the FPU paradox came in 1965 and is two–fold: from one side it's due to *deterministic chaos* and from another to the existence of *solitons*. The first line of thought was developed in 1965 by B. V. Chirikov (1928–2008) and F. M. Izrailev starting from an important result obtained some years earlier by A. N. Kolmogorov (1903–1987), J. K. Moser (1928–1999) and V. I. Arnol'd (1937–2010), known today as KAM theorem. Invoking the last one, Chirikov and Izrailev proved that if the perturbation in the FPU system is sufficiently strong (at least to induce a ""superposition" of the nonlinear resonances), then the observed recurrences brakes up and the system state goes rapidly towards thermalization [ChI965].

The second approach, instead, was proposed by N. J. Zabusky and M. D. Kruskal using the so–called *continuum limit* [ZaK965]; they succeed to relate the periodic behavior observed by Fermi, Pasta and Ulam to the dynamics of *localized excitations*, which they called ""solitons", obtained as solution of the KdV equation, which turned out to be nothing but the continuum limit of equation (1.33). In order to discuss the last line of thought, let's rewrite equation (1.33) in terms of the density ρ and the elasticity coefficient μ of the string, defined s.t. $m = \rho h$, $k = \mu/h$. Then equation (1.33) takes the form

$$\ddot{\eta}_n = c^2\left(\frac{\eta_{n+1}+\eta_{n-1}-2\eta_n}{h^2}\right)\big[1+\beta(\eta_{n+1}-\eta_{n-1})\big], \qquad c \equiv \sqrt{\mu/\rho}. \tag{1.37}$$

[13] I.e., the Fourier transform of the displacements $\{\eta_n(t)\}_{n=1,2,\dots,N}$ with initial conditions $\eta_0 = 0 = \eta_{N+1}$.

[14] Notice that to prove this assumption, Fermi, Pasta and Ulam integrated the equations of motion (1.36) for different values of β, fixing the initial shape of the string as sinusoidal one.

To study the continuum limit of the latter equation, we need to find a function $u = u(x;t)$, with $x \in [0,\ell]$, that should spatially interpolate the discretized shape of the string, i.e. such that

$$u(x_n,t) = u(nh,t) = \eta_n(t), \qquad \forall\, n = 1,2,\dots,N. \tag{1.38}$$

Notice that $\ddot{\eta}_n = u_{tt}(x_n;t)$ and $\eta_{n\pm1} = u(x_n \pm h;t)$, for $n = 1,2,\dots,N$. Substituting in (1.37), we find

$$u_{tt}(x,t) = c^2 \frac{u(x+h,t) + u(x-h,t) - 2u(x,t)}{h^2} \left\{1 + \beta\big[u(x+h,t) - u(x-h,t)\big]\right\}. \tag{1.39}$$

Assuming the function $u = u(x;t)$ to be *analytic*, we can expand it in Taylor series

$$\frac{1}{c^2}u_{tt} = 2\sum_{n\in\mathbb{N}} \frac{h^{2n-2}u_{x^{2n}}}{(2n)!}\left[1 + 2\beta\sum_{k\in\mathbb{N}}\frac{h^{2k-1}u_{x^{2k-1}}}{(2k-1)!}\right] = 2\left[\frac{u_{xx}}{2!} + \frac{h^2}{4!}u_{x^4} + \frac{h^4}{6!}u_{x^6} + \mathcal{O}(h^6)\right] +$$
$$+ 4\beta h\left[\frac{u_{xx}}{2!} + \frac{h^2}{4!}u_{x^4} + \frac{h^4}{6!}u_{x^6} + \mathcal{O}(h^6)\right]\left[u_x + \frac{h^2}{3!}u_{x^3} + \frac{h^4}{5!}u_{x^5} + \mathcal{O}(h^6)\right].$$

Neglecting infinitesimals of order higher than h^6 and performing the scaling transformation $(x,t) \mapsto (\xi,\tau)$, with $\xi \equiv x - ct$ and $\tau \equiv (\beta h)ct$, one finds the equation[15]

$$u_{\xi\tau} - \frac{1}{2}\beta h u_{\tau\tau} = -\frac{1}{24}\frac{h}{\beta}u_{\xi^4} - \frac{1}{840}\frac{h^3}{\beta}u_{\xi^6} - 2\left(\frac{u_{\xi\xi}}{2} + \frac{h^2}{4!}u_{\xi^4} + \frac{h^4}{6!}u_{\xi^6}\right)\left(u_\xi + \frac{h^2}{3!}u_{\xi^3} + \frac{h^4}{5!}u_{\xi^5}\right).$$

Taking then the continuum limit (i.e. $h \to 0$), reminding that the nonlinear perturbation is small (i.e., $\beta \to 0$) and assuming that h,β tend to zero with the *same velocity* (meaning that $h\beta = \mathcal{O}(h^2)$), we finally find the PDE $u_{\xi\tau} + \delta^2 u_{\xi^4} + u_\xi u_{\xi^2} = 0$, where $\delta^2 \equiv \lim_{h,\beta\to0}\sqrt{h/24\beta}$. Considering then the substitution $w \equiv u_\xi$, one recovers the usual form of the KdV equation, that is

$$w_\tau + ww_\xi + \delta^2 w_{\xi\xi\xi} = 0. \tag{1.40}$$

Having discovered the existence of a relation between the FPU equation and KdV, Zabusky and Kruskal considered the initial value problem of (1.40) with periodic boundary conditions: they chose $\delta = 0.022$ and assumed a cosine initial shape $u(x,0) = \cos(\pi x)$ with $x \in [0,2]$ (see Figure 6) [DrJ993]. From their simulations, it was observed that at first the nonlinear term ww_ξ of equation (1.40) dominated over the dispersive one $\delta^2 w_{\xi\xi\xi}$ (being $\delta \ll 1$) and, as usual, the wave increase in those regions where it has negative slope[16]. At later times this dominance regime starts to invert itself and the dispersive terms balances the nonlinearity, avoiding the formation of a discontinuity; then the solution develops a train of eight well defined waves (full line) each one with small wavelength (of order δ). The amplitudes of the oscillations grow and finally each oscillation achieves an almost steady amplitude (increasing linearly from left to right) and has a shape almost identical to that of an individual KdV–solitary–wave solution (i.e. $\sinh^2$–like). «*Finally*», as Zabusky and Kruskal reported in their paper [ZaK965], «*each such ""solitary–wave pulse" or ""**soliton**" begins to move uniformly at a rate which is linearly proportional to its amplitude. Thus the solitons spread apart*». However, having assumed periodic boundary conditions, «*two or more solitons eventually overlap spatially and interact nonlinearly*» and, shortly after the interaction, «*they reappear virtually unaffected in size or shape*».

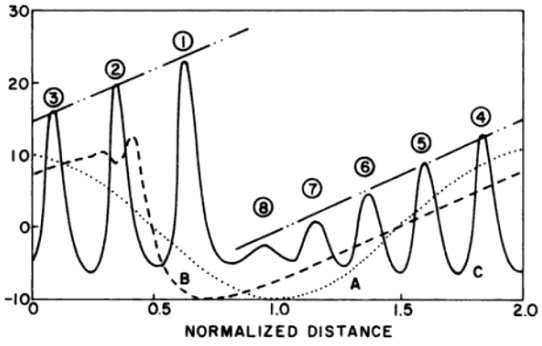
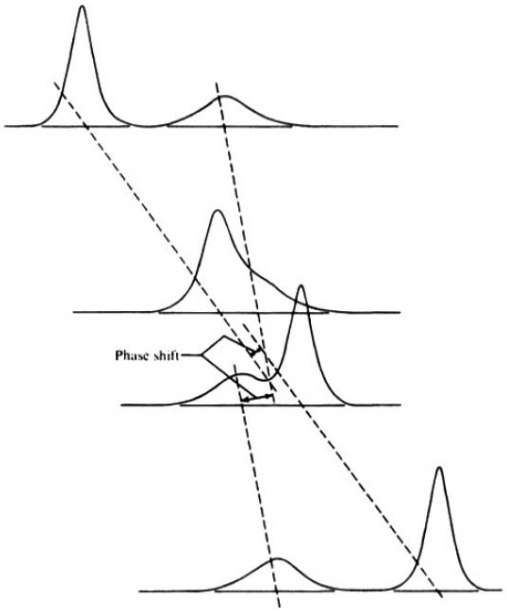


Figure 6: ZK solution: $t = 0$ (dotted line), $t_b = 1/\pi$ (broken line) and $t = 3.6/\pi$ (full line) [ZaK965].

[15] Under re–scaling, differential operators transform simultaneously as follows: $\partial_x^k = \partial_\xi^k$ and $\partial_t = -c(\partial_\xi - \beta h\partial_\tau)$, so $\partial_t^2 = c^2(\partial_\xi^2 - 2\beta h\partial_{\xi\tau} + \beta^2h^2\partial_\tau^2)$. Thus in coordinates (ξ,τ) the d'Alembertian has form $\Box = -2\beta h\partial_{\xi\tau} + (\beta h)^2\partial_\tau^2$.

[16] Assuming $u(x,0)$ has negative slope, it's possible to prove that higher parts of the wave travel faster than lower ones. The global effect is that the wave *steepens* till a characteristic time t_b (called ""*breaking time*") at which the graph of the function $u = u(x;t)$ has a vertical tangent; function $u(x;t)$ becomes multi–valued for $t > t_b$ and a *shock wave* arises.



Actually, the only difference after the nonlinear interaction is a small *phase change* (see figure on the left [AbS981]), i.e. a shifting of the center of each wave with respect to where it would have been if it traveled alone. This property inspired Zabusky and Kruskal to call these waves ""solitons", in order to emphasize their particle–like character (e.g. like *protons, photons* or *phonons*). Therefore FPU recurrences can be understood as (nonlinear) superpositions of solitons.

This discovery leaded to an intese study of the subject, with the aim to clarify the novel aspects that appeared. For example, it was not clear if (and if so, what) other NLEE had soliton solutions and how to obtain them; answering these questions is the main purpose of the present work. Of course, in order to detect and calculate soliton solutions of NLEE, we need a precise definition of a *soliton*. This is not easy task; for now we shall associate the term with any NLEE solution which represents a stationary, localized (i.e. decays or approach a constant value at infinity) wave that can interact strongly with other solitons and retain its identity[17] [DrJ993].

2 Miura Transform and KdV solutions by Inverse Scattering

2.1 Conservation laws and Miura transform. In the previous section we've learned that the mechanism behind the formation of solitons is (roughly speaking) a delicate balance between nonlinear and dispersive wave effects. From the beginning this observation attracted the attention of many researchers and, few years after the discovery of Zabusky and Kruskal, it was understood that the phenomenon could be explained in terms of one or more *conservation laws*, which we shall presently define.

Definition 2.1 - *A* ***conservation law*** *associated to* $u_t = \mathcal{D}(u, u_x, \dots)$ *is an expression of the form*

$$\mathcal{T}_t + \mathcal{X}_x = 0, \tag{2.1}$$

where $\mathcal{T} = \mathcal{T}(x, t; u, u_x, u_t, \dots)$ *is the* ***conserved density*** *and* $-\mathcal{X} = -\mathcal{X}(x, t; u, u_x, u_t, \dots)$ *its* ***flux****.*

Notice that a *local* conservation law depends *only* on u and its derivatives. For these conservation laws it's possible to integrate equation (2.1) to see that $\mathcal{I}[u] \equiv \int_{\mathbb{R}} \mathcal{T}(u, u_x, u_t, \dots)\, \mathrm{d}x$ is a constant, i.e. $\mathcal{I}_t = 0$. We then say that $\mathcal{I}$ is a *constant of the motion* of an *integral* of the differential equation[18].

Consider now the KdV equation. Two conservation laws follow by direct examination:

$$u_t + \left(\tfrac{\alpha}{2}u^2 + u_{xx}\right)_x = 0, \qquad \left(\tfrac{1}{2}u^2\right)_t + \left(\tfrac{\alpha}{3}u^3 + uu_{xx} - \tfrac{1}{2}u_x^2\right)_x = 0,$$

where the latter follows from the former by multiplying by u. Besides these ""trivial" conservation laws, others were discovered by R. M. Miura [Miu968] and it was conjectured that there was an infinite

[17] Notice that in the context of the KdV equation and other similar equations, it's usual to refer to a single–soliton solution as the *solitary–wave*, but when more then one of them appear in a solution, they are called solitons. Another way to express this is that the soliton becomes a solitary wave when it is infinitely separated from any other soliton. Also, it has to be pointed out that there exist NLEE that have solitary–wave solutions that are *not* solitons, whereas others (like the KdV equation) have solitary–waves which *are* also solitons.

[18] Notice that not all constants of the motions arise from conservation laws.

number of them. In studying KdV conservation laws and those associated to the mKdV (see equation (1.17)), Miura discovered an important transformation: if v is a solution of (1.17) then (*and only then*)

$$u = -\left(v^2 + v_x\right) \tag{2.2}$$

is a KdV solution. In particular, renaming mKdV by $\mathcal{M}[v]$ and KdV by $\mathcal{K}[u]$, the following holds true[19]

$$\mathcal{K}[u] = -\left(2v + \partial_x\right)\mathcal{M}[v]. \tag{2.3}$$

The operator on the r.h.s. guarantees the single valuedness of the transformation in one direction only.

The transformation (2.3) leaded to the discovery of other important properties of the KdV equation. Originally (1968) it was the basis of a proof, presented by C. S. Gardner, Kruskal and Miura, that the KdV equation has indeed an infinite number of *conserved quantities* [MGK968]. The basic idea follows from the property of *Galilei–invariance* of the KdV equation, i.e. the transformation

$$\begin{cases} x' = x + \frac{6}{\varepsilon^2}t, \\ t' = t, \end{cases} \qquad u(x,t) = u'(x',t') - \frac{1}{\varepsilon^2}, \tag{2.4}$$

leaves the KdV equation invariant. The proof goes as follows: equation (2.4) implies the differential operator transformations $\partial_x = \partial_{x'}$, $\partial_t = \frac{6}{\varepsilon^2}\partial_{x'} + \partial_{t'}$, from which we obtain

$$\mathcal{K}[u(x,t)] \xrightarrow{(x,t)\mapsto(x',t')} u_{t'} + \frac{6}{\varepsilon^2}u_{x'} + 6uu_{x'} + u_{x'x'x'} = u'_{t'} + 6u'u'_{x'} + u'_{x'x'x'} = \mathcal{K}[u'(x',t')].$$

Analogously it's possible to prove that, under Galilei transformations and the rescaling

$$v(x,t) = -\varepsilon w(x',t') + \tfrac{1}{\varepsilon}, \tag{2.5}$$

the mKdV equation transforms itself into a conservation law equation of the form [AbS981]

$$\mathcal{M}[v(x,t)] = 0 \qquad \xrightarrow{(x,t)\mapsto(x',t')} \qquad w_{t'} + \left(6w^2 - 2\varepsilon^2 w^3 + w_{x'x'}\right)_{x'} = 0. \tag{2.6}$$

Thus w is a conserved density and $\int w\,\mathrm{d}\,x'$ is a constant of the motion for the mKdV equation. Substitution of the rescaling expressions (2.4), (2.5) in the Miura transformation (2.2), yields

$$u' = 2w + \varepsilon w_{x'} - \varepsilon^2 w^2. \tag{2.7}$$

Thinking of $\varepsilon \ll 1$, we can solve equation (2.7) iteratively[20]: after reordering in powers of ε, one finds

$$w = w_0 + \varepsilon w_1 + \varepsilon^2 w_2 + \mathcal{O}(\varepsilon^3) \equiv \frac{u'}{2} - \frac{\varepsilon}{4}u'_{x'} + \frac{\varepsilon^2}{8}\left[(u')^2 + u'_{x'x'}\right] + \mathcal{O}(\varepsilon^3). \tag{2.8}$$

Being w a conserved density, it must be independent of the parameter ε: thus also $(w_0, w_1, w_2, \dots)$ are (distinct, nontrivial) conserved densities, which implies that KdV and mKdV have an infinite number of constants of the motion. We shall give later an alternative proof of this property for the KdV equation.

Let's go back to the Miura transformation. Note that equation (2.2) can be interpreted as a *Riccati equation*, a second order ODE with non–constant coefficients of the form

$$y_x(x) = \mathcal{P}(x) + \mathcal{Q}(x)y(x) + \mathcal{R}(x)y^2(x), \tag{2.9}$$

[19] From the expression (2.2) we find the following transformations: $u_t = -\left(2vv_t + v_{xt}\right)$, $u_x = -\left(2vv_x + v_{xx}\right)$ and so $u_{xxx} = -\left(6v_{xx}v_x + 2vv_{xxx} + v_{xxxx}\right)$; substitutions in the equation for $\mathcal{K}[u]$ with $\alpha = 6$ complete the proof.

[20] Notice that $w = \frac{u'}{2} - \frac{\varepsilon}{2}w_{x'} + \frac{\varepsilon^2}{2}w^2$, thus $w_{x'} = \frac{u'_{x'}}{2} - \frac{\varepsilon}{2}w_{x'x'} + \mathcal{O}(\varepsilon^2)$, $w_{x'x'} = \frac{u'_{x'x'}}{2} + \mathcal{O}(\varepsilon)$ and so on; therefore

$$w = \tfrac{u'}{2} - \tfrac{\varepsilon}{2}\left[\tfrac{u'_{x'}}{2} - \tfrac{\varepsilon}{2}\left(\tfrac{u'_{x'x'}}{2} - \dots\right)\right] + \tfrac{\varepsilon^2}{2}\left(\tfrac{u'}{2} + \dots\right)^2 = \tfrac{u'}{2} - \tfrac{\varepsilon}{4}u'_{x'} + \tfrac{\varepsilon^2}{8}\left[u'_{x'x'} + (u')^2\right] + \mathcal{O}(\varepsilon^3).$$

begin $\mathcal{P}(x)$, $\mathcal{Q}(x)$, $\mathcal{R}(x)$ smooth, non–zero, real valued functions. Equation (2.9) is generalization of the Bernoulli equation (which is recovered assuming $\mathcal{P} = 0$): it was introduced by Count J. F. Riccati (1676–1754) and can be linearized applying the following transformation

$$y \equiv -\frac{\psi_x}{\mathcal{R}\psi} \qquad \Longrightarrow \qquad y_x = \frac{\psi_x\left(\psi\mathcal{R}_x + \psi_x\mathcal{R}\right) - \psi_{xx}\psi\mathcal{R}}{\left(\mathcal{R}\psi\right)^2}. \tag{2.10}$$

Substituting the latter in (2.9), we obtain the second order ODE with non–constant coefficients

$$\mathcal{R}\psi_{xx} - \left(\mathcal{R}_x + \mathcal{Q}\mathcal{R}\right)\psi_x + \mathcal{R}^2\mathcal{P}\psi = 0. \tag{2.11}$$

Therefore, if ψ is a solution of the linearized Riccati equation, then y is a solution of the Riccati equation. Comparing equation (2.2) with (2.9) we see that the Miura transformation is a Riccati equation for v in terms of u, with $\mathcal{P}(x) = -u(x)$, $\mathcal{Q}(x) = 0$, $\mathcal{R}(x) = -1$. Thus equation (2.2) can be linearized by the transformation $v = {}^{\psi_x}/_{\psi}$, from which one obtains $\psi_{xx} + u\psi = 0$. However, since KdV must preserve Galilei–invariance, Miura, Gardner and Kruskal considered [MGK968]

$$\psi_{xx} + (\lambda + u)\,\psi = 0, \tag{2.12}$$

which is exactly the stationary linear Schrödinger equation in one dimension. This result has great relevance: if ψ is a solution of equation (2.12), then $u = -(v^2 + v_x)$ with $v = {}^{\psi_x}/_{\psi}$ is a KdV solution. In this sense we can think the linear Schrödinger equation as an *implicit linearization* for KdV.

In 1967 the team Gardner, Greene, Kruskal and Miura (hereafter abbreviated GGKM) discovered a method for exact solving the initial value problem for the KdV equation [GGKM67, GGKM74]. Before describing the original procedure, let's discuss the following, slightly different, method [AbS981]. Associate to equation (2.12) the time evolution equation

$$\psi_t = \mathcal{A}\psi + \mathcal{B}\psi_x, \tag{2.13}$$

where $\mathcal{A}, \mathcal{B}$ are scalar functions independent of ψ. Consider now the system of equations (2.12), (2.13): forcing the so–called compatibility condition $\psi_{xxt} = \psi_{txx}$, yields

$$\begin{cases} \psi_{xx} = -\left(\lambda + u\right)\psi, \\ \psi_t = \mathcal{A}\psi + \mathcal{B}\psi_x, \end{cases} \xrightarrow{\psi_{xxt} = \psi_{txx}} \begin{cases} \lambda_t + u_t + \mathcal{A}_{xx} - 2\mathcal{B}_x(\lambda + u) - \mathcal{B}u_x = 0, \\ 2\mathcal{A}_x + \mathcal{B}_{xx} = 0. \end{cases} \tag{2.14}$$

Choosing now $\mathcal{A} = u_x$ and $\mathcal{B} = 2(2\lambda - u)$, we get the non–trivial expression

$$\lambda_t + u_t + u_{xxx} + 6uu_x = \lambda_t + \mathcal{K}[u] = 0 \qquad \Longrightarrow \qquad \mathcal{K}[u] = 0 \iff \lambda_t = 0. \tag{2.15}$$

Hence KdV is satisfied iif the eigenvalues of the associated Schrödinger equation are time–independent (i.e. they are constants of the motion in the KdV equation). There exist some deductive procedures to calculate the expressions for $\mathcal{A}$ and $\mathcal{B}$; also, it can be proved that there are infinitely many equations of the form (2.13) associated to (2.12), corresponding to different choices of functions $\mathcal{A}, \mathcal{B}$. In subsequent sections we shall discuss how results (2.12)–(2.15) can be used to reconstruct $u(x,t)$ given $u(x,0)$.

Remark 2.1 – The sufficiency of condition (2.15) can be proved by following an alternative procedure. Consider the system of equations (2.14) and suppose that $\lambda_t = 0$; furthermore, let $\mathcal{A}, \mathcal{B}$ be two first order polynomials in λ, i.e. $\mathcal{A} = \mathcal{A}_0 + \lambda\mathcal{A}_1$ and $\mathcal{B} = \mathcal{B}_0 + \lambda\mathcal{B}_1$. Substitution in system (2.14) yields

$$\begin{cases} 2\lambda^2\mathcal{B}_{1x} - \lambda^1\left(\mathcal{A}_{1xx} - 2\mathcal{B}_{0x} - 2\mathcal{B}_{1x}u - \mathcal{B}_1u_x\right) + \lambda^0\left(\mathcal{A}_{0xx} - 2\mathcal{B}_{0x}u - \mathcal{B}_0u_x + u_t\right) = 0, \\ \lambda^1\left(\mathcal{B}_{1xx} + 2\mathcal{A}_{1x}\right) + \lambda^0\left(\mathcal{B}_{0xx} + 2\mathcal{A}_{0x}\right) = 0. \end{cases}$$

Now force to zero the coefficients belonging to the same powers of λ: from λ^2 one gets $\mathcal{B}_{1x} = 0$ and then $\mathcal{B}_1(x,t) = b_1(t)$ (with $b_1 = b_1(t)$ a real function), whilst coefficients of λ^1, λ^0 give respectively

$$\lambda^1 : \begin{cases} \mathcal{A}_{1xx} - 2\mathcal{B}_{0x} - 2\mathcal{B}_{1x}u - \mathcal{B}_1u_x = 0, \\ \mathcal{B}_{1xx} + 2\mathcal{A}_{1x} = 0, \end{cases} \qquad \lambda^0 : \begin{cases} \mathcal{A}_{0xx} - 2\mathcal{B}_{0x}u - \mathcal{B}_0u_x + u_t = 0, \\ \mathcal{B}_{0xx} + 2\mathcal{A}_{0x} = 0. \end{cases}$$

Solutions of the first system are $\mathcal{A}_1(x,t) = a_1(t)$ and $\mathcal{B}_0(x,t) = -\frac{1}{2}b_1(t)u(x,t) + b_2(t)$, which combined with the first equation of the second system yields $\mathcal{A}_0(x,t) = \frac{1}{4}b_1(t)u_x(x,t) + a_2(t)$. Substituting now the expressions for $\mathcal{A}_0, \mathcal{A}_1, \mathcal{B}_1$ and $\mathcal{B}_0$ in the second equation of the second system, we find

$$\frac{1}{4}b_1 u_{xxx} + \frac{3}{2}b_1 u u_x - b_2 u_x + u_t = 0 \qquad \xrightarrow[\forall\, t\in\mathbb{R}_0^+]{b_1(t)\stackrel{!}{=}4,\ b_2(t)\stackrel{!}{=}0} \qquad \mathcal{K}[u] = 0.$$

Note that a_1, a_2 are not involved in the last expression, so it's possible to set $a_1(t) = 0 = a_2(t)\ \forall\, t \in \mathbb{R}_0^+$ without loss of generality. Inserting what obtained in $\mathcal{A}$ and $\mathcal{B}$ gives $\mathcal{A} = u_x$ and $\mathcal{B} = 2(2\lambda - u)$, which are exactly the positions that GGKM made when passing from equation (2.14) to equation (2.15).

2.2 Exact solutions by Inverse Scattering. This paragraph contains a brief, somehow historical, introduction to the Inverse Scattering Transform method for the KdV equation as was originally introduced by GGKM[21] in 1967. Again, we consider only initial data that decay sufficiently rapidly.

Let's recall some properties of the stationary Schrödinger equation. Assuming $u : \mathbb{R} \to \mathbb{R}$ to be a smooth real potential function, it's known that equation (2.12) admits two kind of solutions: *bound states*, characterized by negative eigenvalues and a countable set of $L_2(\mathbb{R})$ eigenfunctions and *unbound states*, corresponding to positive eigenvalue and generalized eigenfunctions behaving asymptotically as periodic waves. Under the additional hypothesis that $u(x,t)$ decays sufficiently rapidly as $|x| \to \infty$, one can show that there is in fact only a finite number N of negative eigenvalues. Our purpose is to solve the KdV equation by determining the time evolution of $\lambda = \lambda(t)$ and $\psi = \psi(x,t)$ in equation (2.12).

Consider first bound state solutions. Solving equation (2.12) for u and substituting in KdV, gives

$$\lambda_{nt}\psi_n^2 + \big(\psi_n \mathcal{Q}_x - \psi_{nx}\mathcal{Q}\big)_x = 0, \qquad \mathcal{Q} \equiv \psi_{nt} + \psi_{nxxx} - 3(\lambda - u)\psi_{nx}. \tag{2.16}$$

Equation (2.16) can be integrated: since eigenfunctions are square integrable, one finds $\lambda_{nt} = 0$, consistently with result (2.15). Therefore equation (2.16) can be directly integrated [Kar998]

$$\frac{\psi_{nxx}}{\psi_n} = \frac{\mathcal{Q}_{xx}}{\mathcal{Q}} \qquad \Longrightarrow \qquad \mathcal{Q} = \mathcal{C}(t)\psi_n + \mathcal{D}(t)\psi_n \int_0^x \frac{\mathrm{d}\, s}{\psi_n^2(s)}, \tag{2.17}$$

and eigenfunctions can be normalized with respect to the $L_2(\mathbb{R})$–norm, i.e. $\int_{\mathbb{R}} \psi_n^2(x,t)\,\mathrm{d}\, x = 1$. Having negative eigenvalues, rewrite λ_n as $(\imath\kappa_n)^2$ with $\kappa_n \in \mathbb{R}^+$. Note that function $\phi_n(x) \equiv \psi_n \int_0^x \psi_n^{-2}(s)\,\mathrm{d}\, s$ is a solution of (2.16); but we know that, asymptotically, solutions of (2.12) look like linear combinations of $e^{\pm\kappa_n x}$ and, being L_2–functions, $\psi_n \to 0$ as $|x| \to \infty$. Thus ϕ_n blows up and we must have $\mathcal{D} = 0$. Substituting in equation (2.17), multiplying by ψ_n and integrating, one funds immediately that $\mathcal{C} = 0$ for discrete eigenvalues. Now introduce the functions $c_n = c_n(t)$ for $n = 1, 2, \ldots, N$, defined such that

$$\psi_n \sim c_n(t) e^{-\kappa_n x} \qquad \text{as} \qquad |x| \to \infty.$$

Substituting in equation (2.17), using the fact that $u \to 0$ as $|x| \to \infty$ and $\lambda_t = 0$, one finds

$$\frac{\mathrm{d}\, c_n(t)}{\mathrm{d}\, t} = 4\kappa_n^3 c_n(t). \tag{2.18}$$

Thus a trivial calculations allows to determine the time evolution of the coefficients $\{c_n\}_{n=1,2\ldots,N}$, which constitute the discrete part of the so–called *scattering data* for the eigenvalue problem in equation (2.12).

We are left with the case of unbound states. Having in mind that solutions of equation (2.12) belonging to positive eigenvalues $\lambda \equiv k^2 \in \mathbb{R}^+$ are asymptotically linear combinations of $e^{\pm\imath kx}$ (being $u \to 0$ as $|x| \to \infty$) and following the same arguments as before, one can prove that $\lambda_t = 0$, as expected. Thus, in order to solve the KdV equation we need to complete the calculation of the scattering data for

[21] We'll mainly follow S. Karigiannis, "" *The Inverse Scattering Transform and integrability of NLEE*" [Kar998].

the eigenvalue problem in equation (2.12). With this aim, let's calculate the *reflection* and *transmission coefficients* associated to the following asymptotic boundary conditions

$$\psi \sim \begin{cases} e^{-\imath kx} + \rho(k,t)e^{\imath kx} & \text{as} \quad x \to +\infty, \\ \tau(k,t)e^{-\imath kx} & \text{as} \quad x \to -\infty. \end{cases} \tag{2.19}$$

Actually we only need the first boundary condition to solve KdV; in fact, forcing the latter in (2.16) gives us [Kar998] $\mathcal{C}(k) = 4\imath k^3$, $\mathcal{D}(k) = 0$, $\forall\, k \in \mathbb{R}^* := \mathbb{R} \setminus \{0\}$ and substitution of (2.19) in (2.16) yields

$$\frac{\mathrm{d}\,\rho(k,t)}{\mathrm{d}\,t} = 8\imath k^3 \rho(k,t), \qquad \frac{\mathrm{d}\,\tau(k,t)}{\mathrm{d}\,t} = 0, \tag{2.20}$$

which are again trivially integrable. Using equations (2.18) and (2.20), it's possible to completely determine the time evolution of the **scattering data**, which we define formally as

$$\mathcal{S}(t) := \Big\{ \big\{ c_n(t), \kappa_n(t) \big\}_{n=1,2,\dots,N};\ \rho(k,t),\ \tau(k,t),\ k \in \mathbb{R}^* \Big\}.$$

Once scattering data are known, one can finally solve the Cauchy problem for KdV. In fact the initial condition $u(x,0) = u_0$ yields the initial scattering data $\mathcal{S}(0)$ and we can use equations (2.18), (2.20) to calculate $\mathcal{S}(t)$, $\forall\, t \in \mathbb{R}^+$. So what remains to solve the KdV initial value problem is to invert $\mathcal{S}(t)$ to get $u(x,t)$. This procedure, whose key point lies in the fact that the temporal variable t plays the role of a parameter, involves a *linear integral equation* (LIE), known as **Gel'fand–Levitan–Marchenko equation** (GLM). We will discuss this inversion method in the next paragraph.

Thus we have in principle the inverse scattering method, summarized in the following diagram.

$$\begin{array}{ccc} u(x;0) & \xrightarrow{\text{direct scattering, linear ODE at } t=0} & \mathcal{S}(0) \\ \Big\downarrow \text{solution to NLEE} & & \Big\downarrow \text{time evolution of } \mathcal{S} \\ u(x;t) & \xleftarrow[\text{inverse scattering, linear IE at } t \in \mathbb{R}^+]{} & \mathcal{S}(t) \end{array}$$

i) Map the KdV–solution $u = u(x,t)$ to a potential in a stationary Schrödinger equation; *ii*) once the initial condition $u = u(x,0)$ is given, calculate the initial scattering data $\mathcal{S}(0)$ and their time evolution $\mathcal{S}(t)$; *iii*) invert the process to calculate the potential $u(x,t)$. Note that the power of the method consists in reducing a *nonlinear* initial value problem to the solution of linear ODEs (for the time evolution of the scattering data) and a linear integral equation (where t is nothing but a parameter).

Remark 2.4. The IST method can be interpreted as a *nonlinear analogue* of the Fourier transform. To see this, consider the linearized KdV equation: $u_t + u_{xxx} = 0$. Fourier transforming both sides yields

$$u_t(x,t) + u_{xxx}(x,t) = 0 \qquad \xrightarrow{\hat{u} \equiv \mathcal{F}[u]} \qquad \hat{u}_t(k,t) = \imath k^3 \hat{u}(k,t),$$

since $u \to 0$ as $|x| \to \infty$. Solving the linear ODE with the initial condition $\hat{u}(k,0) = \mathcal{F}[u(x,0)]$, one finds $\hat{u} = \hat{u}(k,t)$ $\forall\, t \in \mathbb{R}^+$. Inversion of the Fourier–transform gives the solution $u = u(x,t)$ for the linearized KdV. Note that in the last step t is nothing more than a parameter (as for GLM).

We finish with summarizing our results in a way that will become useful in §3. Combine equations (2.12) and (2.17): from the former one gets $\lambda\psi_x = -\psi_{xxx} - u_x\psi - u\psi_x$ and substitution in equation (2.12) yields $\psi_t = -4\psi_{xxx} - 3u_x\psi - 6u\psi_x + \mathcal{C}\psi$. Therefore we are left with the system

$$\begin{cases} \mathscr{L}\psi \equiv \big[-\partial_x^2 - u(x,t)\big]\,\psi = \lambda\psi, \\ \mathscr{M}\psi \equiv \big[-4\partial_x^3 - 3\big(u\partial_x + \partial_x u\big) + \mathcal{C}\big]\,\psi = \psi_t. \end{cases} \tag{2.21}$$

In what follows we'll call $\mathscr{L}$ the operator of the **spectral problem** and $\mathscr{M}$ the operator of an associated **time–evolution equation**. Note that for discrete eigenvalues (where $\mathcal{C}(t) = 0$) it is possible to recover relation $\lambda_t = 0$ by forcing the *compatibility condition* $\psi_{xxt} = \psi_{txx}$ in equations (2.21).

2.3 Gel'fand–Levitan–Marchenko and KdV soliton solutions. Introduced in 1950's by I. M. Gel'fand (1913–2009), B. L. Levitan (1914–2004) and V. A. Marchenko (1922) [GeL951, Mar955], the inverse scattering method for the Schrödinger equation goes sketchily as follows [Akt009]: first form from the time–evolved scattering data $\mathcal{S}(t)$ the kernel $\mathscr{B} : \mathbb{R} \times \mathbb{R}_0^+ \to \mathbb{R}$, defined as

$$\mathscr{B}(\xi;t) := \sum_{n=1}^{N} c_n^2(t) e^{-\kappa_n \xi} + \frac{1}{2\pi} \int_{\mathbb{R}} \rho(k,t) e^{\imath k \xi} \,\mathrm{d}\, k, \tag{2.22}$$

then solve the corresponding Gel'fand–Levitan–Marchenko integral equation [New983]

$$\mathscr{K}(x,z;t) + \mathscr{B}(x+z;t) + \int_{\mathbb{R}} \mathscr{K}(x,y;t)\mathscr{B}(y+z;t) \,\mathrm{d}\, y = 0, \tag{2.23}$$

for function $\mathcal{K} : \mathbb{R}^2 \times \mathbb{R}_0^+ \to \mathbb{R}$ which is univocally defined by equation (2.23) for $x \le z < +\infty$, is zero for $x > z$ and is subjected to the asymptotic boundary condition $\mathscr{K}(y,z;t) \to 0$ as $y \to +\infty$. Then the potential $u = u(x,t)$ that gives rise to this scattering data is given by [Kar998]

$$u(x,t) = 2\partial_x \mathscr{K}(x,x;t). \tag{2.24}$$

Remark 2.2 – Let $\mathfrak{C}(\mathbb{R}^2 \times \mathbb{R}_0^+)$ denote the Banach space of bounded, continuous real–valued functions on $\mathbb{R}^2 \times \mathbb{R}_0^+$ with the sup norm. Define the functional $\mathscr{F}_{\mathscr{B}} : \mathfrak{C}(\mathbb{R}^2 \times \mathbb{R}_0^+) \to \mathfrak{C}(\mathbb{R}^2 \times \mathbb{R}_0^+)$ as [Pal000]

$$\mathscr{F}_{\mathscr{B}}\big(\mathscr{K}\big)(x,z;t) = -\mathscr{B}(x+z;t) - \int_{\mathbb{R}} \mathscr{K}(x,y;t)\mathscr{B}(y+z;t) \,\mathrm{d}\, y.$$

Then $\mathscr{K}$ satisfies GLM iif it's a fixed–point of $\mathscr{F}_{\mathscr{B}}$. Note that $\mathscr{F}_{\mathscr{B}}$ is Lipschitz with constant $\|\mathscr{B}\|_{L_1}$, so if $\|\mathscr{B}\|_{L_1} < 1$ then $\mathscr{F}_{\mathscr{B}}$ is a *contraction* on $(\mathfrak{C}, \mathfrak{d})$, being $\mathfrak{d} : \mathfrak{C} \times \mathfrak{C} \to \mathbb{R}$ the metric induced by the sup norm. Then the *Banach fixed point theorem* implies that the GLM equation has a **unique solution**, obtained by taking the limit of the sequence $\{\mathscr{K}_n\}_{n\in\mathbb{N}}$, with $\mathscr{K}_1 \equiv -\mathscr{B}$ and $\mathscr{K}_{n+1} = \mathscr{F}_{\mathscr{B}}(\mathscr{K}_n)$.

Remark 2.3 – Suppose that $\rho(k,t) = 0$ identically and that $\mathscr{B}$ is ""separable", i.e. it satisfy an identity of the form $\mathscr{B}(x+z;t) = \sum_{n=1}^{N} \mathscr{X}_n(x)\mathscr{Z}_n(z;t)$, begin $\mathscr{X}_n, \mathscr{Z}_n$ the n–th components of the row and column N–dimensional vectors $\mathcal{X}, \mathcal{Z}$, respectively. Then the GLM equation takes the form

$$\mathscr{K}(x,z;t) + \sum_{n=1}^{N} \mathscr{X}_n(x)\mathscr{Z}_n(z;t) + \sum_{n=1}^{N} \mathscr{Z}_n(z;t) \int_x^{+\infty} \mathscr{K}(x,y;t)\mathscr{X}_n(y) \,\mathrm{d}\, y = 0 \tag{2.25}$$

and it follows that $\mathscr{K}$ must have the form $\mathscr{K}(x,z;t) = \mathscr{Y}(x;t)\mathscr{Z}(z;t)$, where $\mathscr{Y}(x;t)$ is an N–dimensional row–vector whose entries are unknown functions of x and t. Substitution in (2.25) yields

$$\mathscr{Y}_n(x;t) + \mathscr{X}_n(x) + \sum_{m=1}^{N} a_{nm}(x;t)\mathscr{Y}_m(x;t) = 0, \qquad a_{nm}(x;t) \equiv \int_x^{+\infty} \mathscr{X}_n(y)\mathscr{Z}_m(y;t) \,\mathrm{d}\, y;$$

setting $\mathrm{A}(x;t) \equiv \big(a_{nm}(x;t)\big)_{n,m=1,2,\dots,N}$ gives $\mathscr{X} + (\mathbb{1}_N + \mathrm{A})\mathscr{Y} = \mathbb{O}_N$; therefore $\mathscr{Y} = -\mathscr{A}^{-1}\mathscr{X}$, having defined $\mathscr{A}(x;t) \equiv \mathbb{1}_N + \mathrm{A}(x;t)$ and supposed it non–singular $\forall\,(x,t) \in \mathbb{R} \times \mathbb{R}_0^+$. Finally we have

$$\mathscr{K}(x,z;t) = -\sum_{n=1}^{N} \mathscr{X}_n(x) \sum_{m=1}^{N} \mathscr{A}^{-1}_{nm}(x;t)\mathscr{Z}_m(z;t) = -\mathscr{X}(x)\mathscr{A}^{-1}(x;t)\mathscr{Z}(z;t). \tag{2.26}$$

Hence, under the hypothesis of a *reflectionless potential* $\rho(k,t) = 0$ and a separable kernel $\mathscr{B}(x+z;t) \equiv \mathscr{X}(x)\mathscr{Z}(z;t)$, the GLM equation is reduced to N linear algebraic equations.

The presence of a reflectionless potential has a central role in the search for soliton solutions to NLEEs. In fact, it turns out that if $\rho(x,0)=0$, then the solution of the GLM equation describes the interaction of N solitons[22], each one with amplitude and speed characterized by the κ_n's and position characterized by the c_n's. To show this property for the KdV equation, integrate expression (2.20) for the reflection coefficient to get $\rho(k,t)=\rho(k,0)e^{8\imath k^3 t}\ \forall\,(k,t)\in\mathbb{R}^*\times\mathbb{R}^+$; thus, if $\rho(k,0)=0$ it follows that $\rho(k,t)=0\ \forall\,(k,t)\in\mathbb{R}^*\times\mathbb{R}_0^+$ which means that if $u=u(x,t)$ evolves by KdV and it's reflectionless at $t=0$, then it stays reflectionless for all t. Therefore the kernel (2.22) is separable and we have

$$\mathscr{X}^{(\mathrm{KdV})}(x)=\left(e^{-\kappa_1 x},e^{-\kappa_2 x},\dots,e^{-\kappa_N x}\right),\quad \mathscr{Z}^{(\mathrm{KdV})}(z;t)=\left(c_1(t)e^{-\kappa_1 z},c_2(t)e^{-\kappa_2 z},\dots,c_N(t)e^{-\kappa_N z}\right)^{\mathsf{T}}.$$

From the above expressions one finds also immediately that, for all $n,m=1,2,\dots,N$, it is

$$a_{nm}^{(\mathrm{KdV})}(x;t)=\int_x^{+\infty}\mathscr{X}_n(y)\mathscr{Z}_m(y;t)\,\mathrm{d}\,y=c_n^2(t)\frac{e^{-(\kappa_n+\kappa_m)x}}{\kappa_n+\kappa_m};$$

recalling now that $\mathscr{A}_{nm}(x;t)\equiv\delta_{nm}+a_{nm}(x;t)$ for $n,m=1,2,\dots,N$, we get the relation

$$\mathscr{A}^{(\mathrm{KdV})}(x;t)=\left\{\delta_{nm}+c_n^2(0)\frac{e^{-(\kappa_n+\kappa_m)x+8\kappa_n^3 t}}{\kappa_n+\kappa_m}\right\}_{n,m=1,2,\dots,N},\tag{2.27}$$

having used equation (2.18). Now $\partial_x\mathscr{A}_{mn}^{(\mathrm{KdV})}=-c_m^2e^{-(\kappa_n+\kappa_m)x}$, thus expression (2.26) for $z=x$ yields

$$\begin{aligned}\mathscr{K}^{(\mathrm{KdV})}&=-\sum_{n,m=1}^{N}\left[\mathscr{A}_{(\mathrm{KdV})}^{-1}\right]_{nm}c_m^2e^{-(\kappa_n+\kappa_m)x}=\sum_{n,m=1}^{N}\left[\mathscr{A}_{(\mathrm{KdV})}^{-1}\right]_{nm}\partial_x\mathscr{A}_{mn}^{(\mathrm{KdV})}\\&=\mathrm{Tr}\left[\mathscr{A}_{(\mathrm{KdV})}^{-1}\partial_x\mathscr{A}^{(\mathrm{KdV})}\right]=\partial_x\mathrm{Tr}\left[\ln\left(\mathscr{A}^{(\mathrm{KdV})}\right)\right]=\partial_x\ln\left[\det\left(\mathscr{A}^{(\mathrm{KdV})}\right)\right],\end{aligned}$$

having used the well–known identity $\det\left(e^A\right)=e^{\mathrm{Tr}(A)}$ with $A=\ln\mathscr{A}$ in the last passage. Substitution in (2.24) gives finally the expression for *N–solitons solution* of the KdV equation, namely

$$u_{N\text{–sol}}^{(\mathrm{KdV})}(x;t)=2\partial_x^2\ln\left\{\det\left[\mathscr{A}^{(\mathrm{KdV})}(x;t)\right]\right\}.\tag{2.28}$$

Consider the case $N=1$ (see note 18, pg. 15): substitution of expression (2.27) in (2.28) yields

$$u_{1\text{–sol}}^{(\mathrm{KdV})}(x;t)=2\partial_x^2\ln\left[1+\tfrac{c^2(t)}{2\kappa}e^{-2\kappa x}\right]=2\kappa^2\mathrm{sech}^2\left[\kappa(x-4\kappa^2t)+\theta_0\right],\qquad\theta_0\equiv\ln\tfrac{\sqrt{2\kappa}}{c(0)},$$

being $\kappa\equiv\kappa_1$, $c\equiv c_1$; therefore, we've recovered our earlier formula (1.31) for the KdV solitary–wave solution. In the same fashion, one finds for the KdV 2–solitons solution the expression [Pal000]

$$u_{2\text{-sol}}^{(\mathrm{KdV})}=2\frac{\kappa_1^2z_1+\kappa_2^2z_2+2(\kappa_1+\kappa_2)^2z_1z_2+\alpha(\kappa_1^2z_2+\kappa_2^2z_1)z_1z_2}{(1+z_1+z_2+\alpha z_1z_2)^2},\tag{2.29}$$

where $\sqrt{\alpha}\equiv\frac{\kappa_1-\kappa_2}{\kappa_1+\kappa_2}$ and $z_i(x;t)\equiv\frac{c_i^2(0)}{2\kappa_i}e^{-2\tau_i}$, being $\tau_i\equiv\kappa_i(x-4\kappa_i^2t)$ for $i=1,2$; note that

$$\det\left(\mathscr{A}_{2\text{-sol}}^{(\mathrm{KdV})}\right)=1+z_1+z_2+\alpha z_1z_2.$$

Assuming $\kappa_1>\kappa_2$, one can show that solution (2.29) actually behaves as a superposition of two traveling wave solutions of the form (1.31). With this aim, let's analyze the 2–*solitons asymptotics* [Dun010], starting with the case $t\to-\infty$: in the limit $x\to-\infty$ one has $\det(\mathscr{A})\sim e^{-2(\tau_1+\tau_2)}$ and then $u\sim0$, as expected. Now move along the x–axis to the right and consider the leading term in $\det(\mathscr{A})$ when

[22] Note that soliton solutions always correspond to eigenvalue problems with *reflectionless potentials*.

$\tau_1 = 0$ and then when $\tau_2 = 0$. We first reach the point $x = 4\kappa_1^2 t$ where $\tau_1 = 0$: in the neighborhood of this point one has $\tau_2 = 4\kappa_2(\kappa_1^2 - \kappa_2^2) \ll 0$, which means that

$$\det\left(\mathscr{A}_{2\text{–sol}}^{(\mathrm{KdV})}\right) \sim \frac{c_2^2(0)}{2\kappa_2} e^{-2\tau_2}\left(1 + \frac{\alpha c_1^2(0)}{2\kappa_1} e^{-2\tau_1}\right) \qquad \Longrightarrow \qquad u_{2\text{–sol}}^{(\mathrm{KdV})} \sim 2\left(1 + \frac{\alpha c_1^2(0)}{2\kappa_1} e^{-2\tau_1}\right)_{xx},$$

which has the form of a 1–soliton solution with phase $\theta_1^- \equiv \frac{1}{2}\ln {}^{2\kappa_1}/_{\alpha c_1^2(0)}$. Now move to the left until reaching the point $x = 4\kappa_2^2 t$ where $\tau_2 = 0$: around this point $\tau_1 = 4\kappa_1(\kappa_2^2 - \kappa_1^2) \gg 0$ and then

$$\det\left(\mathscr{A}_{2\text{–sol}}^{(\mathrm{KdV})}\right) \sim 1 + \frac{c_2^2(0)}{2\kappa_2} e^{-2\tau_2}.$$

Thus the wave profile is a KdV 1–soliton with phase $\theta_2^- = \frac{1}{2}\ln {}^{2\kappa_2}/_{c_2^2(0)}$. Note that at $t = 0$ the two solitons coalesce and the behavior depends on κ_1/κ_2. Consider now the $t \to +\infty$ limit: if $x \to +\infty$ then $\det(\mathscr{A}) \sim 1$ and $u \sim 0$. Moving along the x–axis to the left one first reaches the point $\tau_1 = 0$, where $\tau_2 \gg 0$ and the wave–profile is a KdV 1–soliton with phase $\theta_1^+ = \frac{1}{2}\ln {}^{2\kappa_1}/_{c_1^2(0)}$, and then the point $\tau_2 = 0$, where $\tau_1 \ll 0$ and the single soliton has phase $\theta_2^+ = \frac{1}{2}\ln {}^{2\kappa_2}/_{\alpha c_2^2(0)}$. Comparing asymptotic phases θ_1^-, θ_1^+ and θ_2^-, θ_2^+ we can deduce that the taller soliton has overtaken the lower one and that they haven't changed in shape but just in phase, with variations given by the expressions

$$\Delta\theta_1 \equiv \theta_1^+ - \theta_1^- = \frac{1}{2}\ln\alpha, \qquad \Delta\theta_2 \equiv \theta_2^+ - \theta_2^- = -\frac{1}{2}\ln\alpha;$$

note that the total phase–shift is zero. Hence we conclude that the KdV 2–soliton solution (2.29) describes the interaction between two separate solitons, whose only result can be measured by $\ln\left(\frac{\kappa_1-\kappa_2}{\kappa_1+\kappa_2}\right)$, which is large iif $\kappa_1 - \kappa_2$ is small, i.e. iff the two solitons have similar velocities [Dun010].

Remark 2.4 – The picture above generalized to the case $N > 2$. As $t \to \mp\infty$ the general (so–called *pure*) N–solitons solution (2.28) represents N separate solitons ordered accordingly to their speed: the tallest (and therefore faster) one is at the front, followed by the second tallest and so on. At $t = 0$ they all coalesce and then individual solitons re–emerge in the opposite order as $t \to \pm\infty$.

Remark 2.5 – The number N of discrete eigenvalues for the Schrödinger operator equals the number of KdV solitons at $t \to \pm\infty$. This number is encoded in the initial conditions; to see this, consider

$$u_{\text{N–sol}}^{(\mathrm{KdV})}(x;0) \equiv u_0(x) = 2N(N+1)\mathrm{sech}^2(x), \qquad N \in \mathbb{N}_0$$

and make the change of variable $\mathbb{R} \ni x \mapsto \xi \equiv \tanh(x) \in (-1,1)$ in the Schrödinger equation:

$$\psi_{xx} + u_0(x)\psi = k^2\psi \qquad \Longrightarrow \qquad \frac{\mathrm{d}}{\mathrm{d}\,\xi}\left[(1-\xi^2)\frac{\mathrm{d}\,\psi(\xi)}{\mathrm{d}\,\xi}\right] + \left[N(N+1) + \frac{k^2}{1-\xi^2}\right]\psi(\xi) = 0,$$

which is nothing but the associated Legendre equation. Analysis of the power series solution of the latter shows that $\psi \in L_2\big((-1,1)\big)$ iif $k^2 = -\kappa^2$ with $\kappa = 1, 2, \ldots, N$ [Dun010, DrJ993].

2.4 The Hamiltonian formulation. In the last paragraph we have seen that the KdV equation is IST–solvable and admits N–solitons solutions. However, proceeding in the same fashion as for the KdV case, it can be shown that *every known IST–solvable NLEE admits multi–soliton solutions*[23]. Then a question arises naturally: does *every* IST–solvable NLEE admits soliton–solutions? As far as we know there is no definite answer to this problem nor a theorem which states it; however, we can try to answer a (somehow) less ambitious question, namely why solitary–wave type solutions for IST–solvable NLEEs are *stable*. There are different ways to tackle this problem: here we're going to give some clues about

[23] This is, for example, the case of NLS & SG in $(1+1)$–dimensions and of KP & DS in $(2+1)$–dimensions.

the connection with **completely integrable Hamiltonian systems** [AbS972, NMPZ84].

First of all, recall some basics of Hamiltonian system. Let's consider a system with n degrees of freedom, whose motion is described by a trajectory in a $2n$–dimensional *phase space* $\mathcal{M}$, parametrized with local coordinates (p_i, q_i) for $i = 1, 2, \dots, n$. The *dynamical variables* of the system are real, smooth functions $f : \mathcal{M} \times \mathbb{R} \to \mathbb{R}$ s.t. $f = f(q, p; t)$ and the phase space $\mathcal{M}$ is a *Poisson manifold*, being equipped with a bilinear map $\{\cdot, \cdot\}_{\mathcal{M}} : \mathrm{C}^\infty(\mathcal{M}) \times \mathrm{C}^\infty(\mathcal{M}) \to \mathrm{C}^\infty(\mathcal{M})$ with respect to which the pair $(\mathrm{C}^\infty(\mathcal{M}), \{\cdot, \cdot\}_{\mathcal{M}})$ forms an algebra. That maps is called *Poisson bracket* and is defined as

$$\{f, g\}_{\mathcal{M}} := \sum_{i=1}^{n} \Big(f_{q_i} g_{p_i} - f_{p_i} g_{q_i} \Big), \qquad \forall\, f, g \in \mathrm{C}^\infty\big(\mathcal{M} \times \mathbb{R}\big).$$

Note that $\{\cdot, \cdot\}_{\mathcal{M}}$ is an *antisymmetric derivation*, satisfying the *Jacobi identity*. If $\{f, g\}_{\mathcal{M}} = 0$ then f, g are said to be in *involution*. The pair (q, p) satisfies the *canonical commutation relations*

$$\big\{q_i, q_j\big\}_{\mathcal{M}} = 0, \qquad \big\{p_i, p_j\big\}_{\mathcal{M}} = 0, \qquad \big\{q_i, p_j\big\}_{\mathcal{M}} = \delta_{ij}, \quad \forall\, i, j = 1, 2, \dots, n.$$

Once the Hamiltonian function of the system $\mathcal{H} = \mathcal{H}(q, p; t)$ is given, the evolution of a dynamical variable f is determined by the well–known **Hamilton–Liouville** (HL) **equation** [GPS001]

$$\frac{\mathrm{d}\, f}{\mathrm{d}\, t} = \big\{f, \mathcal{H}\big\}_{\mathcal{M}} + \frac{\partial f}{\partial t}.$$

In particular, assuming that the canonical coordinates (q, p) doesn't depend explicitly on time, one finds the HL *equations of motion* $\dot{p}_i = -\mathcal{H}_{q_i}$ and $\dot{q}_i = -\mathcal{H}_{p_i}$ for $i = 1, 2, \dots, n$. Since we'll soon be interested in infinite dimensional systems, let's rewrite the latter equations in **symplectic form** as

$$\dot{u} := \begin{pmatrix} \dot{p} \\ \dot{q} \end{pmatrix} = \begin{pmatrix} \mathbb{O}_n & -\mathbb{1}_n \\ \mathbb{1}_n & \mathbb{O}_n \end{pmatrix} \begin{pmatrix} \mathcal{H}_p \\ \mathcal{H}_q \end{pmatrix} =: \mathcal{J}_0 \mathcal{H}_u, \tag{2.30}$$

being $\mathcal{H}_\chi \equiv (\partial_{\chi_1} \mathcal{H}, \partial_{\chi_2} \mathcal{H}, \dots, \partial_{\chi_n} \mathcal{H})^\intercal$. Within this formulation, one can give a different definition of the Poisson bracket, namely in terms of the standard scalar product on $\mathbb{R}^{2n}$:

$$\big\{f, g\big\}_{\mathcal{M}} = \big\langle f_u, \mathcal{J}_0 g_u \big\rangle_{\mathbb{R}^{2n}}. \tag{2.31}$$

An arbitrary system $\dot{v} = \mathcal{J} \mathcal{G}_v$ can be put in Hamiltonian form iff $\exists \mathcal{T} \in \mathrm{GL}(\mathbb{R}^{2n})$ s.t. $\mathcal{J}_0 = \mathcal{T} \mathcal{J} \mathcal{T}^{-1}$; this holds iff $\mathcal{J}$ is non–singular and skew–adjoint [Lax978]. Having in mind the infinite dimensional generalization, let's drop here the non–singular requirement for the following [Kar998]

Definition 2.2 - *A system of differential equations is called Hamiltonian iff it is of the form $\dot{u} = \mathcal{J} \mathcal{H}_u$, for some skew–adjoint linear operator $\mathcal{J}$ and some Hamiltonian function $\mathcal{H}$.*

Consider now a dynamical variable $f \in \mathrm{C}^\infty(\mathcal{M})$, constant along the Hamiltonian flow. From the HL equations it follows that $\{f, \mathcal{H}\}_{\mathcal{M}} = 0$ and it's also possible to prove the converse [Lax978]. We'll say that f is an *integral of the motion* for the system iff $\{f, \mathcal{H}\}_{\mathcal{M}} = 0$. Constants of the motion are important since each of them allows to reduce the order of the system by a factor of 2; this property underlines the following

Definition 2.3 - *An Hamiltonian system with n degrees of freedom is said to be **completely integrable** iff it admits n independent[24] constant of the motion $\{f_i\}_{i=1,2,\dots,n}$, with $f_1 \equiv \mathcal{H}$, in involution, i.e. s.t.*

$$\big\{f_i, f_j\big\}_{\mathcal{M}} = 0, \qquad \forall\, i, j = 1, 2, \dots, n.$$

[24] Meaning that the gradients $\{\nabla f_i\}_{i=1,2,\dots,n}$ are linearly independent vectors on a tangent space to any point in $\mathcal{M}$.

Integrable systems lead to *completely solvable* HL equations of the motion; in order to show this, let's remind the freedom in the choice of canonical coordinates. Then, consider the coordinate transformation

$$(q,p) \longmapsto \Big(\mathcal{Q}(q,p), \mathcal{P}(q,p)\Big).$$

This transformation is called *canonical* iff it leaves invariant the Poisson bracket, i.e. iff

$$\big\{f(q,p), g(q,q)\big\}_{\mathcal{M}(q,p)} = \big\{f(\mathcal{Q},\mathcal{P}), g(\mathcal{Q},\mathcal{P})\big\}_{\mathcal{M}(\mathcal{Q},\mathcal{P})}, \qquad \forall\, f,g \in \mathrm{C}^\infty(\mathcal{M}).$$

Consequently, canonical commutation relations and HL equations are invariant under canonical transformations. To construct canonical transformations, one introduces the so–called *generating functions* [Arn989]; for example, given a function $S = S(q,\mathcal{P};t)$ which satisfies the non–degenerate condition $\det(S_{q_i\mathcal{P}_j})_{i,j=1,2\ldots,n} \neq 0$, one can construct a canonical transformations by setting $p_i = S_{q_i}$, $\mathcal{Q}_i = S_{\mathcal{P}_i}$ and $\mathcal{H}' = \mathcal{H} + S_t$. For completely integrable systems, the idea is to seek for a canonical transformations s.t. $\mathcal{H}' = \mathcal{H}'(\mathcal{P}_1,\ldots,\mathcal{P}_n)$ and then $\mathcal{P}_i(t) = \mathcal{P}_i(0) = const$ and $\mathcal{Q}_i(t) = \mathcal{Q}_i(0) + \mathcal{H}'_{\mathcal{P}_i} t$ for $i = 1,2,\ldots,n$. The latter would then represent the solutions to the HL equations of motion and the integrable systems would then be *completely solvable by quadratures*[25]. The well–known **Arnol'd–Liouville integrability theorem** applies in this contest [Arn989].

THEOREM 2.1 - *Let $(\mathcal{M}, \{f_i\}_{i=1,2,\ldots,n})$ be a completely integrable Hamiltonian system and let*

$$\mathcal{M}_f := \big\{(q,p) \in \mathcal{M} \,:\, f_i(q,p) = c_k,\, k = 1,2,\ldots,n\big\},$$

with c_k a constant $\forall\, k = 1,2,\ldots,n$, be an n–dimensional level surface of first integrals. Then

- ▷ *if $\mathcal{M}$ is compact and connected then it is diffeomorphic to an n–torus $\mathbb{T}^n$;*
- ▷ *exists a canonical transformation $(q,p) \mapsto (I,\phi)$, where*

$$I_i := \frac{1}{2\pi} \oint_{\Gamma_i} \sum_{j=1}^n p_j \,\mathrm{d}\, q_j, \quad \phi_i \in [0,2\pi) \qquad i = 1,2,\ldots,n,$$

  *Γ_i the i–th basic cycle of $\mathbb{T}^n$, s.t. **angles** ϕ_i are coordinates on $\mathbb{T}^n$ and **actions** I_i are first integrals;*

- ▷ *in action–angle coordinates, the canonical equation of the motion become*

$$\dot{I}_i(t) = 0, \quad \dot{\phi}_i(t) = \omega_i\big(I_1, I_2, \ldots, I_n\big), \qquad i = 1,2,\ldots,n, \tag{2.32}$$

  where $\omega_i(I) := \partial_{I_i}\mathcal{H}'$ are frequencies and $\mathcal{H}'(I,\phi) \equiv \mathcal{H}\big(q(I,\phi), p(I,\phi)\big)$.

Proof. See V. I. ARNOL'D – *Mathematical Methods of Classical Mechanics*, ch. 9, pgg: 258–285. □

REMARK 2.6 – The actions I_i are independent on the choice of Γ_i: in fact, given a cycle Γ'_i with opposite orientation with respect to Γ_i, the Stokes theorem implies that

$$\oint_{\Gamma_i} \sum_{j=1}^n p_j \,\mathrm{d}\, q_j + \oint_{\Gamma'_i} \sum_{j=1}^n p_j \,\mathrm{d}\, q_j = \iint_{\Omega_i} \Big(\partial_{q_k} p_j - \partial_{q_j} p_k\Big) \,\mathrm{d}\, q_j \wedge \mathrm{d}\, q_k = 0,$$

being Ω_i a cylindroid whose top and bottom surfaces have Γ_i and Γ'_i as boundaries. Note that actions I_i are also *first integrals*: assuming $\det(\partial_{p_j} f_i) \neq 0$, the system $f(q,p) = \mathbf{c}$ can be solved for the momenta $p = p(q,\mathbf{c})$ and thus $\oint_{\Gamma_i} p(q,\mathbf{c}) \cdot \mathrm{d}\, q$ only depends on $\mathbf{c}$. Moreover the I_i's are in involution:

$$\big\{I_i, I_j\big\}_{\mathcal{M}} = \sum_{\ell,m=1}^n \frac{\partial I_i}{\partial f_\ell} \frac{\partial I_j}{\partial f_m} \big\{f_\ell, f_m\big\}_{\mathcal{M}} = 0, \qquad \forall\, i,j = 1,2,\ldots,n.$$

[25] In other words, performing a finite number of algebraic operations and integrations of known functions.

We can now generalize the discussion to infinite dimensional systems and describe the relation with IST–solvable NLEEs. Formally one can think to replace our original phase space $\mathcal{M}$ with the space of all real valued C^∞ rapidly decreasing functions, i.e. the *Schwartz space* $\mathrm{S}(\mathbb{R})$. Thus coordinates are given by functions $u = u(x)$, being x the continuous analogue of the index i in the finite dimensional case; dynamical variables are represented with *functionals* $\mathcal{F} : \mathrm{S}(\mathbb{R}) \to \mathbb{R}$ given by integrals $\mathcal{F}[u] = \int_{\mathbb{R}} f(u, u_x, u_{xx}, \dots)\,\mathrm{d}\,x$ and[26] gradients $\mathcal{F}_u$ are replaced by functional derivatives of $\mathcal{F}$, defined as

$$\int_{\mathbb{R}} \frac{\delta \mathcal{F}[u]}{\delta u(x)} v(x)\,\mathrm{d}\,x := \lim_{\varepsilon\to 0} \frac{\mathcal{F}[u+\varepsilon v] - \mathcal{F}[u]}{\varepsilon} = \frac{\mathrm{d}\,\mathcal{F}[u+\varepsilon v]}{\mathrm{d}\,\varepsilon}\bigg|_{\varepsilon=0}, \qquad \forall\, v \in \mathrm{S}(\mathbb{R}), \tag{2.33}$$

where $v \in \mathrm{S}(\mathbb{R})$ has the role of a *test function*. In particular, the Hamiltonian is given by the integral

$$\mathcal{H}[u] = \int_{\mathbb{R}} \mathcal{H}(u, u_x, u_{xx}, \dots)\,\mathrm{d}\,x,$$

begin $\mathcal{H}$ the *Hamiltonian density* function. From definition (2.33), we can deduce an explicit expression for the functional derivative. Take e.g. the Hamiltonian functional $\mathcal{H}$: assuming $|\varepsilon| \ll 1$, one finds

$$\mathcal{H}\big[u+\varepsilon v\big] = \mathcal{H}[u] + \varepsilon \int_{\mathbb{R}} \Big(\mathcal{H}_u v + \mathcal{H}_{u_x} v_x + \mathcal{H}_{u_{xx}} v_{xx} + \cdots \Big)\,\mathrm{d}\,x.$$

Now observe that $\mathcal{H}_{u_x} v_x = \partial_x(\mathcal{H}_{u_x} v) - v\partial_x \mathcal{H}_{u_x}$ and $\mathcal{H}_{u_{xx}} v_{xx} = \partial_x(\mathcal{H}_{u_{xx}} v_x - v_{xx}\partial_x \mathcal{H}_{u_{xx}}) + v\partial_x^2 \mathcal{H}_{u_{xx}}$: since $v \in \mathrm{S}(\mathbb{R})$, all boundary terms goes to zero and we are left with

$$\frac{\mathcal{H}\big[u+\varepsilon v\big] - \mathcal{H}[u]}{\varepsilon} = \int_{\mathbb{R}} \left(\mathcal{H}_u - \frac{\partial \mathcal{H}_{u_x}}{\partial x} + \frac{\partial^2 \mathcal{H}_{u_{xx}}}{\partial x^2} - \frac{\partial^3 \mathcal{H}_{u_{xxx}}}{\partial x^3} + \cdots \right) \mathrm{d}\,x.$$

Finally, the *fundamental lemma of variational calculus* leads to expression [GeF963]

$$\frac{\delta \mathcal{H}[u]}{\delta u(x)} = \mathcal{H}_u - \frac{\partial \mathcal{H}_{u_x}}{\partial x} + \frac{\partial^2 \mathcal{H}_{u_{xx}}}{\partial x^2} - \frac{\partial^3 \mathcal{H}_{u_{xxx}}}{\partial x^3} + \cdots. \tag{2.34}$$

Observe that if $\mathcal{H} = \mathcal{H}(u_x, u_{xx}, u_{xxx}, \dots)$, i.e. if $u = u(x)$ is *cyclic*, then equation (2.34) becomes

$$\frac{\delta \mathcal{H}[u_x]}{\delta u(x)} = -\frac{\partial}{\partial x}\left(\mathcal{H}_{u_x} - \frac{\partial \mathcal{H}_{u_{xx}}}{\partial x} + \frac{\partial^2 \mathcal{H}_{u_{xxx}}}{\partial x^2} - \cdots \right) = -\frac{\partial}{\partial x} \frac{\delta \mathcal{H}[u_x]}{\delta u_x(x)}. \tag{2.35}$$

In order to introduce the equations of the motion, we need $\mathrm{S}(\mathbb{R})$ to be a Poisson manifold; thus one has to define a Poisson bracket $\{\cdot,\cdot\}_{\mathrm{S}(\mathbb{R})} : \mathrm{C}^\infty(\mathrm{S}(\mathbb{R}))^{\times 2} \to \mathrm{C}^\infty(\mathrm{S}(\mathbb{R}))$ s.t. $(\mathrm{C}^\infty(\mathrm{S}(\mathbb{R})), \{\cdot,\cdot\}_{\mathrm{S}(\mathbb{R})})$ is an algebra. The analogy with the finite dimensional case suggests the following definition

$$\big\{\mathcal{F}, \mathcal{G}\big\}_{\mathrm{S}(\mathbb{R})} := \big\langle \delta_u \mathcal{F}, \mathfrak{J} \delta_u \mathcal{G} \big\rangle_{\mathrm{S}(\mathbb{R})} := \int_{\mathbb{R}} \frac{\delta \mathcal{F}[u]}{\delta u(x)} \mathfrak{J}(x) \frac{\delta \mathcal{G}[u]}{\delta u(x)}\,\mathrm{d}\,x, \qquad \forall\, \mathcal{F}, \mathcal{G} \in \mathrm{C}^\infty\big(\mathrm{S}(\mathbb{R})\big), \tag{2.36}$$

where $\mathfrak{J}$ is a (possibly singular) skew–adjoint linear operator. In what follows we will chose $\mathfrak{J} = \partial_x$, which is skew–adjoint with respect to the inner product of $\mathrm{S}(\mathbb{R})$, since $\forall\, u, v \in \mathrm{S}(\mathbb{R})$ one has

$$\big\langle \mathfrak{J} u, v \big\rangle_{\mathrm{S}(\mathbb{R})} = \int_{\mathbb{R}} u_x(x) v(x)\,\mathrm{d}\,x = u(x)v(x)\Big|_{-\infty}^{+\infty} - \int_{\mathbb{R}} u(x) v_x(x)\,\mathrm{d}\,x = \big\langle u, -\mathfrak{J} v \big\rangle_{\mathrm{S}(\mathbb{R})}.$$

Choosing $\mathcal{F}[u] = u$ and $\mathcal{G}[u] = \mathcal{H}[u]$ in (2.36), the HL equation of the motion becomes [Lax978]

$$\frac{\partial u(x)}{\partial t} = \big\{u, \mathcal{H}[u]\big\}_{\mathrm{S}(\mathbb{R})} = \int_{\mathbb{R}} \frac{\delta u(x)}{\delta u(y)} \mathfrak{J}(y) \frac{\delta \mathcal{H}[u]}{\delta u(y)}\,\mathrm{d}\,y = \mathfrak{J}(x) \frac{\delta \mathcal{H}[u]}{\delta u(x)} \qquad \Longrightarrow \qquad u_t = \mathfrak{J} \delta_u \mathcal{H}, \tag{2.37}$$

[26] We could in principle allow the t derivatives in f, but is unnecessary for the reasons to become clear shortly.

which defines an infinite dimensional Hamiltonian system, referring to our previous definition. Note that in (2.37) we have used the relation $\frac{\delta f(x)}{\delta f(y)} = \delta(x-y)$, where the δ on the r.h.s. is the Dirac delta.

REMARK 2.7 – Equation (2.37) can be written in a form closely resembling the usual one, namely $\dot p = -\mathcal{H}_q$, $\dot q = \mathcal{H}_p$. To do this, replace the standard derivatives with functional derivatives, so that

$$q_t = \delta_p \mathcal{H}[q,p], \qquad p_t = -\delta_q \mathcal{H}[q,p], \tag{2.38}$$

for a pair $\big(q(x,t), p(x,t)\big)$ of *conjugated variables*. Within this formulation the Poisson bracket has form

$$\big\{\mathcal{F},\mathcal{G}\big\}_{\mathrm{S}(\mathbb{R})} = \int_{\mathbb{R}} \left(\frac{\delta \mathcal{F}[q,p]}{\delta q(x)} \frac{\delta \mathcal{G}[q,p]}{\delta p(x)} - \frac{\delta \mathcal{F}[q,p]}{\delta p(x)} \frac{\delta \mathcal{G}[q,p]}{\delta q(x)} \right) \mathrm{d}\, x, \qquad \forall\, \mathcal{F},\mathcal{G} \in \mathrm{C}^\infty\big(\mathrm{S}(\mathbb{R})\big), \tag{2.39}$$

and the canonical commutation relations between the conjugated variables become $\forall\, x,y \in \mathbb{R}$

$$\big\{q(x),q(y)\big\}_{\mathrm{S}(\mathbb{R})} = 0, \qquad \big\{p(x),p(y)\big\}_{\mathrm{S}(\mathbb{R})} = 0, \qquad \big\{q(x),p(y)\big\}_{\mathrm{S}(\mathbb{R})} = \delta(x-y). \tag{2.40}$$

Note that if $q = q(x,t)$ is cyclic, then (from the equations (2.38), combined with (2.35), it follows that) we may set $p = q_x$ and use (2.35) to recover the expression (2.36) from (2.39) [AbS981].

2.4.1 KdV first integrals and (bi–)Hamiltonian structure. We are now ready to explain the connections between the KdV equation and the Hamiltonian formalism. In §2.2 we have obtained for the transmission coefficient the relation $\tau_t(k;t) = 0$, $\forall\, k \in \mathbb{R}^*$. Therefore, one may interpret $\{\tau(k,t), k \in \mathbb{R}^*\}$ as a set of infinitely many first integrals, provided that they are non–trivial and independent. Since we wish to express these first integrals in functional form, set [MGK968, ZaF971]

$$\psi(x,k;t) = e^{-\imath k x + \int_x^{+\infty} \mathcal{T}(y,k;t)\,\mathrm{d}\, y}, \tag{2.41}$$

where $\mathcal{T}$ is a smooth function to be determined. Now consider the behavior of the wave function $\psi = \psi(x,k;t)$ as $x \to -\infty$: equation (2.19) gives expression $e^{-\imath k x}\psi(x,k;t) \sim \tau(k;t)$ and then

$$\tau(k;t) = \lim_{x\to-\infty} e^{\imath k x}\psi(x,k;t) = e^{\int_{\mathbb{R}} \mathcal{T}(y,k;t)\,\mathrm{d}\, y}, \tag{2.42}$$

which holds for $\Im k < 0$, if one assumes k to be in the lower half plane, and then in the limit $\Im k \to 0$, because of real analyticity [Dun010]. Having cast first integrals in functional form, let's find an equation for $\mathcal{T}$. Inserting the expression (2.41) in the Schrödinger equation (here rewritten with $\lambda = k^2$) yields $\mathcal{T}_x\psi + (\mathcal{T}+\imath k)^2\psi + u\psi = -k^2\psi$, which holds $\forall\, \psi \in L_2(\mathbb{R})$; therefore we find [ZaF971]

$$\mathcal{T}_x(x,k;t) + \mathcal{T}^2(x,k;t) + 2\imath k \mathcal{T}(x,k;t) = -u(x,t). \tag{2.43}$$

Note that equation (2.43) is a Riccati–type equation (see formula (2.9)) and one can search for solutions of (2.43) of the form $\mathcal{T}(x,k;t) = \sum_{n\in\mathbb{N}} (2\imath k)^{-n} \mathcal{T}_n(x;t)$, leading to the *recursion formula* [AbS981]

$$\mathcal{T}_1(x;t) = -u(x;t), \qquad \mathcal{T}_{n+1}(x;t) = \partial_x \mathcal{T}_n(x;t) + \sum_{m=1}^{n-1} \mathcal{T}_m(x;t)\mathcal{T}_{n-m}(x;t), \tag{2.44}$$

which can be solved for the few first terms, giving the following expressions:

$$\mathcal{T}_2 = -u_x, \qquad \mathcal{T}_3 = -u_{xx} + u^2, \qquad \mathcal{T}_4 = -u_{xxx} + 2\big(u^2\big)_x,$$
$$\mathcal{T}_5 = -u_{x^4} + 2\big(u^2\big)_{xx} + \big(u_x\big)^2 + 2uu_{xx} - 2u^3.$$

Combining then the time dependence of $\tau = \tau(k;t)$ given in (2.20) with expression (2.42), implies $\frac{\mathrm{d}}{\mathrm{d}\, t}\int_{\mathbb{R}} \mathcal{T}_n(x;t)\,\mathrm{d}\, x = 0$ so that $I_n[u] \equiv \int_{\mathbb{R}} \mathcal{T}_n(x;t)\,\mathrm{d}\, x$ are *first integrals for the KdV equation* $\forall\, n \in \mathbb{N}$.

REMARK 2.8 – From the above results it follows that the KdV equation admits a *countable infinite* set of first integrals of the motion, as already noted in §2.1 following Miura's approach [Miu968]; indeed, the expressions obtained here for the *conserved densities* $\mathcal{T}_1, \mathcal{T}_2, \mathcal{T}_3$ coincide respectively (up to irrelevant multiplicative factors) with the relations for w_0, w_1, w_2 given in (2.8). Note also that not all the first integrals in the set $\{I_n\}_{n\in\mathbb{N}}$ are non–trivial, e.g. $I_2[u] = \int_{\mathbb{R}} \mathcal{T}_2(x;t)\,\mathrm{d}\,x = -u\big|_{-\infty}^{+\infty} = 0$, being $u \in \mathrm{S}(\mathbb{R})$ and the same holds for I_4. Actually, one can prove that all even terms I_{2n} are trivially zero, since all the expressions for $\mathcal{T}_{2n}$ with $n \in \mathbb{N}$ have the form of an exact x–derivative [NMPZ84].

Therefore, we are left with odd first integrals of the KdV equation, given by the expression

$$I_{n-1}[u] \equiv \frac{1}{2}\int_{\mathbb{R}} \mathcal{T}_{2n+1}(u, u_x, u_{xx}, \dots)\,\mathrm{d}\,x, \qquad n \in \mathbb{N}. \tag{2.45}$$

The first of these is just the integral of u itself, whilst for the next one gets

$$I_0[u] = \frac{1}{2}\int_{\mathbb{R}} u^2\,\mathrm{d}\,x, \qquad I_1[u] = \frac{1}{2}\int_{\mathbb{R}} \left(u_x^2 - 2u^3\right)\mathrm{d}\,x.$$

These two first integrals are associated with the *translational invariance* of KdV and thus, via Noether's theorem, with the conservation of total momentum and energy, respectively. Now we can cast the KdV equation in Hamiltonian form: introducing the Hamiltonian functional $\mathcal{H}_{\mathrm{KdV}}[u] = I_1[u]$, with Hamiltonian density given by $\mathcal{H}_{\mathrm{KdV}}(u, u_x) = \frac{1}{2}u_x^2 - u^3$, the HL equation (2.37) becomes

$$u_t = \frac{\partial}{\partial x}\frac{\delta \mathcal{H}_{\mathrm{KdV}}}{\delta u} = \frac{\partial}{\partial x}\left(\frac{\partial \mathcal{H}_{\mathrm{KdV}}}{\partial u} - \frac{\partial}{\partial x}\frac{\partial \mathcal{H}_{\mathrm{KdV}}}{\partial u_x}\right) = -\left(3u^2 + u_{xx}\right)_x.$$

Thus the KdV is an infinite dimensional Hamiltonian system. Note that I_0 commutes with $\mathcal{H}_{\mathrm{KdV}}$, since

$$\left\{I_0, \mathcal{H}_{\mathrm{KdV}}\right\}_{\mathrm{S}(\mathbb{R})} = \int_{\mathbb{R}} \frac{\delta I_0[u]}{\delta u(x;t)}\frac{\partial}{\partial x}\frac{\delta \mathcal{H}_{\mathrm{KdV}}[u]}{\delta u(x;t)}\,\mathrm{d}\,x \overset{\delta_u I_0[u]=u}{=\!=\!=} -\int_{\mathbb{R}} u\left(6uu_x + u_{xxx}\right)\mathrm{d}\,x = 0,$$

being u an $\mathrm{S}(\mathbb{R})$ function. Indeed, all the non–trivial first integrals in (2.45) commute with $\mathcal{H}_{\mathrm{KdV}}$:

$$\left\{I_n, \mathcal{H}_{\mathrm{KdV}}\right\}_{\mathrm{S}(\mathbb{R})} = \int_{\mathbb{R}} \frac{\delta I_n[u]}{\delta u(x;t)}\frac{\partial u(x;t)}{\partial t}\,\mathrm{d}\,x = \frac{\mathrm{d}\,I_n[u]}{\mathrm{d}\,t} = 0, \qquad \forall\, n \in \mathbb{N}_0.$$

With a bit of more work [NMPZ84] one can show also that $\{I_n, I_m\}_{\mathrm{S}(\mathbb{R})} = 0$, $\forall\, n, m \in \mathbb{N}_0$.

Therefore, one may ask whether or not $\mathcal{H}_{\mathrm{KdV}}$ is also completely integrable in the Arnol'd–Liouville sense. The answer is affermative, but rather involved and it won't be discussed here in detail[27]; we shall give just a clue to the main reasons behind it. First note that the existence of a countable infinite set of first integrals in involution indeed *suggests* that the KdV system is completely integrable, but is not sufficient for the infinite dimensional case since (roughly speaking) one doesn't know ""how many" functionals in involution are required to assure complete integrability [AbS981]. From another point of view, one notes that expressions $\tau(k,t) = \tau(k,0)$ and $\ln[\rho(k,t)] = \ln[\rho(k,0)] + \imath 8k^3 t$ $\forall\,(t,k) \in \mathbb{R}\times\mathbb{R}^*$, are formally equivalent with equations (2.32), suggesting that $(\ln\rho, \tau)$ can be somehow regarded as action–angle variables for $\mathcal{H}_{\mathrm{KdV}}$, thereby identifying the KdV system as a completely integrable one in a precise sense. Although $(\ln\rho, \tau)$ are not themselves canonical, it has been proved in 1971 by Zakharov & L. D. FADDEEV (1934) and independently in 1975 by H. FLASCHKA (1945) & A. C. NEWELL (1941) [FlN975], that certain functions of τ and ρ satisfy the Poisson commutation relations, namely

$$\big(\mathcal{P}(k), \mathcal{Q}(k)\big) \equiv \big(-\tfrac{k}{\pi}\ln\left[1 - |\rho(k)|^2\right], \arg\left[\rho(k)\right]\big), \qquad \big(\mathcal{P}_n, \mathcal{Q}_n\big) \equiv \big(-2\kappa_n^2, \ln c_n\big),$$

for the unbounded and the discrete components of the spectrum of $\mathcal{H}_{\mathrm{KdV}}$, respectively. Moreover, it can be shown that $\mathcal{H}_{\mathrm{KdV}} = \mathcal{H}_{\mathrm{KdV}}(\mathcal{P}(k), \{\mathcal{P}_n\}_{n=1,2,\dots,N})$, which is the defining property of action–angle variables. Thus the mapping $u \mapsto \left\{\mathcal{P}(k), \mathcal{Q}(k), \{\mathcal{P}_n, \mathcal{Q}_n\}_{n=1,2,\dots,N}\right\}$ is canonical, meaning that the *IST is a canonical transformation* for the KdV system to a set of action–angle variables.

[27] See e.g. M. J. ABLOWITZ, H. SEGUR – *Solitons and the Inverse Scattering Transform*, § 1.6.b, pgg. 58–67.

Therefore, we can reassume the main points of the above arguments[28] as follows:

- the time independence property of the transmission coefficients $\tau = \tau(k,t)$ related to the Schrödinger scattering problem, implies the existence of an infinite set $\{I_n\}_{n\in\mathbb{N}_0}$ of conserved quantities;
- the IST is a canonical transformation for the KdV equation to a set of action–angle variables;
- the KdV is an infinite dimensional Hamiltonian system, with Hamilton functional given by $\mathcal{H}_{\mathrm{KdV}}[u] = \frac{1}{2}\int_{\mathbb{R}}(u_x^2 - 2u^3)\,\mathrm{d}\,x$, completely integrable in the Arnol'd–Liouville sense.

The depicted scenario gives then a clue to understand the existence of multi–solitons solutions in the KdV (and in general in every known IST–solvable NLE) equation. The complete integrability property for the KdV system may be plausibly regarded as a reasons for the stability of its stationary solutions; in fact, the relatively simple picture that emerges in physical variables (e.g. solitons with pairwise interactions) is a direct consequence of the existence of a set of action–angle variables. In particular, no stochastic motion can occur within an IST–solvable initial value problem[29] [AbS981].

We end this section with one last property that is shared by *most of* the IST–solvable systems, namely the so–called bi–Hamiltonian structure [FaT007]. Let's therefore give the following

Definition 2.4 - *A differential manifold $\mathcal{M}$ over a field $\mathbb{K}$ is a **bi–Poisson** manifold iff it is endowed with* two *Poisson brackets* $\{\cdot\,,\cdot\}_{\mathcal{M}}, \{\cdot\,,\cdot\}'_{\mathcal{M}} : \mathrm{C}^\infty(\mathcal{M})^{\times_2} \to \mathrm{C}^\infty(\mathcal{M})$ *s.t. the linear combination*

$$\big\{f,g\big\}^{(\mu)}_{\mathcal{M}} := \big\{f,g\big\}'_{\mathcal{M}} - \mu\big\{f,g\big\}_{\mathcal{M}}, \tag{2.46}$$

defines a Poisson bracket $\forall\,\mu\in\mathbb{K}$ *and* $\forall\, f,g\in\mathrm{C}^\infty(\mathcal{M})$*; expression* (2.46) *defines a **pencil** of Poisson brackets and is known as the* compatibility condition *between the two Poisson structures.*

In particular, for the infinite dimensional case the compatibility condition (2.46) has the form

$$\big\{\mathcal{F},\mathcal{G}\big\}^{(\mu)}_{\mathrm{S}(\mathbb{R})} = \big\langle \delta_u\mathcal{F}, \mathfrak{J}_\mu\delta_u\mathcal{G}\big\rangle_{\mathrm{S}(\mathbb{R})} := \big\langle \delta_u\mathcal{F}, (\mathfrak{J}' - \mu\mathfrak{J})\delta_u\mathcal{G}\big\rangle_{\mathrm{S}(\mathbb{R})}, \qquad \forall\mu\in\mathbb{K},\ \ \forall\,\mathcal{F},\mathcal{G}\in\mathrm{C}^\infty\big(\mathrm{S}(\mathbb{R})\big),$$

being $\mathfrak{J},\mathfrak{J}'$ two (possibly singular) skew–adjoint linear operators; following the definition 2.4, the one parameter family of linear operators $\{\mathfrak{J}_\mu,\ \mu\in\mathbb{K}\}$ is called a *pencil of Poisson operators.*

Suppose now to endow an Hamiltonian system with a bi–Poisson manifold as phase–space; this leads to the next definition, here formulated for the infinite dimensional case.

Definition 2.5 - *Let $\{\mathfrak{J}' - \mu\mathfrak{J},\ \mu\in\mathbb{R}\}$ be a real pencil of Poisson operators. A system of differential equations is called* **bi–Hamiltonian** *if there exist two Hamilton functionals $\mathcal{H}_0,\mathcal{H}_1 : \mathrm{S}(\mathbb{R}) \to \mathbb{R}$ s.t.*

$$u_t = \mathfrak{J}\delta_u\mathcal{H}_0[u] = \mathfrak{J}'\delta_u\mathcal{H}_1[u], \qquad u\in\mathrm{S}(\mathbb{R}). \tag{2.47}$$

The KdV is a well–known example of a bi–Hamiltonian system [Mag978, FOR996]: the two forms correspond respectively to the choices $\mathcal{H}_1[u] = I_1[u]$, $\mathfrak{J} = \partial_x$ and $\mathcal{H}_0[u] = I_0[u]$, $\mathfrak{J}' = -\partial_x^3 - u\partial_x - \partial_x u$.

The bi–Hamiltonian formulation gives a very effective way to construct first integrals. Assuming that $\mathfrak{J},\mathfrak{J}'$ satisfy (2.46) and that $\mathfrak{J}$ is non–degenerate[30], one can define the *recursion operator* $\mathfrak{R} := \mathfrak{J}'\mathfrak{J}^{-1}$ and invoke a fundamental result, proved by P. J. Olver, which states that if $\mathfrak{R}^n(\mathfrak{J}\delta_u\mathcal{H}_0)$ lies in the image of $\mathfrak{J}$ $\forall\, n\in\mathbb{N}_0$, then there exists a set $\{\mathcal{H}_n\}_{n\in\mathbb{N}_0}$ of conserved functionals given recursively by

$$\mathfrak{J}\delta_u\mathcal{H}_n[u] = \mathfrak{R}^n\big(\mathfrak{J}\delta_u\mathcal{H}_0[u]\big), \qquad n\in\mathbb{N}_0, \tag{2.48}$$

which are all in involution [Olv993]. For the KdV system $\mathfrak{R}_{\mathrm{KdV}} = -\partial_x^2 - 4u - 2u_x\partial_x^{-1}$ (where ∂_x^{-1} formally defines the x–integration) and all first integrals in (2.45) are recovered by (2.48).

[28] Note that the discussion in indeed quite general, since the same properties also holds in presence of a very large class of NLEEs, from which the KdV is recovered as a particular case [AbS981].

[29] This doesn't suggest that the problem has become trivial; the inverse method is highly non–trivial and, in general, one isn't even guaranteed that its solution are unique and well–behaved. In fact, there are cases of IST–solvable systems (e.g. the ShG) in which singular solutions may develop, provided they don't violate any conservation laws .

[30] A differential operator $\mathfrak{J}$ is degenerate iff there exists a non–zero differential operator $\tilde{\mathfrak{J}}$ s.t. $\mathfrak{J}\circ\tilde{\mathfrak{J}} = 0$ identically.

3 Lax pairs and AKNS method

3.1 Lax pairs for the KdV equation and for other integrable NLEEs. As briefly described, the IST was first developed and applied to the KdV equation by GGKM in 1968; however it was unclear if it would apply to other (physically significant) NLEE [AbS981]. In 1972, V. E. Zakharov and A. B. Shabat showed that indeed the method was not a fluke [ZaS972]; using a technique first introduced by Lax [Lax968], they showed that the NLS equation is related to a linear scattering problem and succeeded in solving this way the associated initial value problem. Shortly after, using these ideas M. Wadaty introduced a method for solving the mKdV equation [Wad972] and in 1973 the team of M. J. Ablowitz, D. J. Kaup, A. C. Newell and H. Segur (hereafter abbreviated AKNS) did the same for the Sine–Gordon equation [AKNS73], developing a general method to obtain, given a suitable linear eigenvalue problem, NLEEs solvable by IST which also keep the spectrum invariant.

Let's start with considering the essential ideas behind Lax's approach. For this purpose, consider the system of equations (2.21): let $\mathscr{L}$ and $\mathscr{M}$ be two $L_2(\mathbb{R})$–differential operator respectively related to the spectral problem and to the associated time evolution equation. Lax observed that the property of time–invariance of the spectrum $\sigma(\mathscr{L})$ could be shown to be equivalent to the statement that $\mathscr{L}(0)$ and $\mathscr{L}(t)$ (both self–adjoint in this case) with $t \in \mathbb{R}^+$ are *unitarily–equivalent*, i.e. there exists a continuous uniparametric family of unitary operators $\{\mathcal{U}(t), t \in \mathbb{R}_0^+\}$ such that

$$\mathscr{L}(t) = \mathcal{U}(t)\mathscr{L}(0)\mathcal{U}^{-1}(t), \qquad \forall\, t \in \mathbb{R}_0^+. \tag{3.1}$$

The proof is straightforward: let $\phi = \phi(x, 0; \lambda)$ be an eigenfunction of $\mathscr{L}(0)$ with eigenvalue λ; evolution under the action of $\mathcal{U}(t)$ gives $\phi(x, t; \lambda) \equiv \mathcal{U}(t)\phi(x, 0; \lambda)$ and it trivially follows that

$$\mathscr{L}(t)\phi(x,t;\lambda) = \mathcal{U}(t)\mathscr{L}(0)\big[\mathcal{U}^{-1}(t)\mathcal{U}(t)\big]\phi(x,0;\lambda) \xrightarrow{\mathcal{U}^*\mathcal{U}=\mathbb{1}=\mathcal{U}\mathcal{U}^*} \lambda\big[\mathcal{U}(t)\phi(x,0;\lambda)\big] = \lambda\phi(x,t;\lambda).$$

Taking now the time derivative of both sides of equation (3.1), one gets

$$\mathscr{L}_t(t)\mathcal{U}(t) + \mathscr{L}(t)\mathcal{U}_t(t) = \mathcal{U}_t(t)\mathscr{L}(0) \qquad \xrightarrow{\mathscr{M}\equiv\mathcal{U}_t\mathcal{U}^*} \qquad \mathscr{L}_t(t) = \big[\mathscr{M}(t), \mathscr{L}(t)\big], \tag{3.2}$$

known as **Lax equation**; operators $\mathscr{L}, \mathscr{M}$ are then called a **Lax pair**. Note that $\big(\mathscr{M}, L_2(\mathbb{R})\big)$ is *skew–adjoint*, i.e. $\mathscr{M}^* = -\mathscr{M}$, which follows differentiating equation $\mathcal{U}\mathcal{U}^* = \mathbb{1}$ with respect to t.

Actually, the skew–adjointness is also necessary to preserve unitarity, as shown in the next [Kar998]

Lemma 3.1 - *Let $\{\mathcal{U}(t), t \in \mathbb{R}_0^+\}$ be a one–parameter family of linear operators s.t. $\mathcal{U}_t = \mathscr{M}\mathcal{U}$, for some skew–adjoint one–parameter family $\{\mathscr{M}(t), t \in \mathbb{R}_0^+\}$. Then if $\mathcal{U}(0)$ is unitary, also $\mathcal{U}(t)$ is $\forall\, t \in \mathbb{R}_0^+$.*

Proof. Let $\mathbf{v}_1, \mathbf{v}_2 \in \mathcal{V}$, $\mathcal{V}$ a vector space over $\mathbb{R}$; define $\mathbf{w}_i(t) \equiv \mathcal{U}(t)\mathbf{v}_i$, $\forall\, i = 1, 2$, $\forall\, t \in \mathbb{R}_0^+$. Then

$$\big\langle \mathbf{w}_1, \mathbf{w}_{2t}\big\rangle_V = \big\langle \mathbf{w}_1, \mathscr{M}\mathbf{w}_2\big\rangle_V = \big\langle \mathscr{M}^*\mathbf{w}_1, \mathbf{w}_2\big\rangle_V = -\big\langle \mathbf{w}_{1t}, \mathbf{w}_2\big\rangle_V \qquad \Longrightarrow \qquad \partial_t\big\langle \mathbf{w}_1, \mathbf{w}_2\big\rangle_V = 0,$$

having observed that $\mathbf{w}_{it} = \mathscr{M}\mathbf{w}_i$; note that $\langle\,\cdot,\cdot\,\rangle_V : V \times V \to \mathbb{R}$ is a scalar product over V. Thus

$$\big\langle \mathbf{w}_1(t), \mathbf{w}_2(t)\big\rangle_V = \big\langle \mathcal{U}(t)\mathbf{v}_1, \mathcal{U}(t)\mathbf{v}_2\big\rangle_V = \big\langle \mathcal{U}(0)\mathbf{v}_1, \mathcal{U}(0)\mathbf{v}_2\big\rangle_V = \big\langle \mathbf{v}_1, \mathbf{v}_2\big\rangle_V,$$

having used unitarity of $\mathcal{U}(0)$. Then $\mathcal{U}(t)$ is an isometry $\forall\, t \in \mathbb{R}_0^+$ and $\mathcal{U}(t)\mathcal{U}^*(t) = \mathbb{1} = \mathcal{U}^*(t)\mathcal{U}(t)$ $\forall\, t \in \mathbb{R}_0^+$, i.e. $\mathcal{U}(t)$ is unitary $\forall\, t \in \mathbb{R}_0^+$. □

Lax showed that if equation (3.2) holds, then the spectrum $\sigma(\mathscr{L})$ is invariant in time.

Theorem 3.1 - *Let $\big\{\big(\mathscr{L}(t), H\big), t \in \mathbb{R}_0^+\big\}$ be a one–parameter family of self–adjoint operators on the Hilbert H. Suppose eigenvalues and eigenfunctions of equation $\mathscr{L}\phi = \lambda\phi$ continuously differentiable with respect to t; if $\exists\,\{\mathscr{M}(t), t \in \mathbb{R}_0^+\}$ s.t. Lax equation holds, then $\lambda_t = 0$. Also if λ is simple then $\phi_t = (\mathscr{M} + \mathcal{C})\phi$ for some continuous function $\mathcal{C} = \mathcal{C}(t)$ and if $(\mathscr{M} + \mathcal{C})$ is skew–adjoint then $\partial_t\|\phi\|_H = 0$.*

Proof. Differentiating $\mathscr{L}\phi = \lambda\phi$ yields $\mathscr{L}_t\phi + \mathscr{L}\phi_t = \lambda_t\phi + \lambda\phi_t$; from (3.1) and $\mathscr{M}\mathscr{L}\phi = \lambda\mathscr{M}\phi$, one gets

$$\lambda_t\phi\mathbb{1}_H = \big(\lambda\mathbb{1}_H - \mathscr{L}\big)\big(\mathscr{M}\phi - \phi_t\mathbb{1}_H\big). \tag{3.3}$$

Now take the inner product of both sides with ϕ and use self–adjointness property of $\mathscr{L}$ to get

$$\lambda_t\big\|\phi\big\|_H^2 = \big\langle\big(\lambda\mathbb{1}_H - \mathscr{L}\big)\big(\mathscr{M}\phi - \phi_t\mathbb{1}_H\big), \phi\big\rangle_H = \big\langle\big(\mathscr{M}\phi - \phi_t\mathbb{1}_H\big), \big(\lambda^*\mathbb{1}_H - \mathscr{L}^*\big)\big\rangle_H = 0 \qquad \Longrightarrow \qquad \lambda_t = 0.$$

Now return to equation (3.3): if λ is simple, since $\mathscr{M}\phi - \phi_t$ belongs to the same eigenspace, we must have $\phi_t = (\mathscr{M} + \mathcal{C})\phi$, for some $\mathcal{C} = \mathcal{C}(t)$. Then it follows that

$$\partial_t\big\|\phi\big\|_H^2 = \big\langle\phi, (\mathscr{M} + \mathcal{C})\phi\big\rangle_H + \big\langle(\mathscr{M} + \mathcal{C})\phi, \phi\big\rangle_H$$

and hence if $\mathscr{M}+\mathcal{C}$ is skew–adjoint, then $\|\phi\|_H^2$ is independent from the time variable t. □

For the KdV equation we have already shown in §2.2 that $\mathcal{C}(t) = 0\ \forall\, t \in \mathbb{R}_0^+$ and that operators $\mathscr{L}_{\mathrm{KdV}}, \mathscr{M}_{\mathrm{KdV}}$ defined in equation (2.21) constitute a Lax pair for the KdV equation. Indeed one can easily compute that the Lax equation for $\mathscr{L}_{\mathrm{KdV}}$ and $\mathscr{M}_{\mathrm{KdV}}$ gives exactly the KdV equation (1.16).

Thus in general we try to associate to an evolution equation of the form $u_t = \mathcal{D}(u)$ a *self–adjoint* operator $\mathscr{L}$ and a *skew–adjoint* operator $\mathscr{M}$ that satisfy the Lax equation. By the above remarks, this condition guarantees that the spectrum of $\mathscr{L}$ is a set of integrals for $u_t = \mathcal{D}(u)$ and that the last one is solvable by IST. The difficulties with this method are that one must ""guess" a suitable $\mathscr{L}$ and then find the associated $\mathscr{M}$ in order to satisfy equation (3.2). On the contrary, an advantage is that once $\mathscr{L}$ is given, there is a somewhat systematic (but complicated) way to find a sequence of evolution equations for which $\mathscr{L}$ constitutes an *isospectral flow* [Kar998]. To show how this method applies take the Schrödinger operator $\mathscr{L} = -\partial_{xx} - u(x,t)$; stationarity yields equation $\mathscr{L}_t = -u_t$. Thus the problem is reduced in finding a skew–adjoint operator $\mathscr{M}$ s.t. $[\mathscr{M}, \mathscr{L}] = -u_t$. If we try, e.g., $\mathscr{M}_0 = \partial_x$ we get isospectral flow for the evolution equation $u_t + u_x = 0$ (since $[\mathscr{M}_0, \mathscr{L}] = u_x$); trying instead the skew–adjoint operator $\mathscr{M}_1 = a\partial_{xxx} + b\partial_x + \partial_x b$, with a, b to be determined, one gets

$$\big[\mathscr{M}_1, \mathscr{L}\big] = 3au_x\partial_{xx} + 3au_{xx}\partial_x + 2bu_x - 4b_x\partial_{xx} - 4b_{xx}\partial_x - b_{xxx}.$$

Choosing then $a = 4$, $b = 3u(x,t)$ one recovers the KdV equation and expression[31] (2.21) for $\mathscr{M}_{\mathrm{KdV}}$.

REMARK 3.1 – Note that the expression for $\mathscr{M}_{\mathrm{KdV}}$ obtained by following the Lax formalism is equivalent to the form used by GGKM in (2.14) [GGKM67], i.e. the associated time evolution equations

$$\psi_t = \mathscr{M}_{\mathrm{KdV}}^{(\mathrm{Lax})}\psi = 4\psi_{xxx} - 6u\psi_x - 3u_x\psi, \qquad \psi_t = \mathscr{M}_{\mathrm{KdV}}^{(\mathrm{GGKM})}\psi = u_x\psi + 2\big(2\lambda - u\big)\psi_x,$$

are equivalent. To prove this, invoke the spectral problem: differentiation of both sides with respect to x gives $\psi_{xxx} + \lambda\psi_x + u_x\psi + u\psi_x = 0$. Now isolate the term $u_x\psi$ and substitute it into the GGKM expression to get $\psi_t = -\psi_{xxx} + 3\lambda\psi_x - 3u\psi_x$; doing the same for $\lambda\psi_x$ leads to equation $\psi_t = -4\psi_{xxx} - 6u\psi_x - 3u_x\psi$. Note that implication Lax → GGKM simply follows by inverting the previous argument.

As written at the beginning, other NLEE where found to be IST–solvable by means of the Lax formalism. In particular, in 1972 Zakharov and Shabat found a Lax pair for a slightly generalized version of the NLS equation (1.23) with the factor α replaced by $\frac{2}{1-p^2}$, for some $p \neq \pm 1$ [ZaS972]. The Lax pair they discovered consisted in the following pair of 2×2 differential operators

$$\mathscr{L}_{p\text{–NLS}} = \imath\begin{pmatrix} 1+p & 0 \\ 0 & 1-p \end{pmatrix}\frac{\partial}{\partial x} + \begin{pmatrix} 0 & u^* \\ u & 0 \end{pmatrix}, \qquad \mathscr{M}_{p\text{–NLS}} = \imath p\mathbb{1}_2\frac{\partial^2}{\partial x^2} + \begin{pmatrix} \frac{|u|^2}{1+p} & \imath u_x^* \\ -\imath u_x & -\frac{|u|^2}{1-p} \end{pmatrix}. \tag{3.4}$$

[31] The constant $\mathcal{C} = \mathcal{C}(t)$ in expression (2.21) must be set to zero, since $\mathscr{M}_{\mathrm{KdV}}$ is already skew–adjoint.

With these choices, the Lax equation (3.2) is satisfied and equals the ZS–form of the NLS [Kar998].

REMARK 3.2 – Note that if $p = 0$ one recovers the NLS equation (1.23) with $\alpha = 2$. In this case it's easy to see that $\mathscr{L}_{\text{NLS}}$ is self–adjoint and $\mathscr{M}_{\text{NLS}}$ is skew–adjoint, although is *not* true in general for $p \neq 0$.

In their paper, ZS applied the IST method to a general 2×2 eigenvalue problem obtained applying to equation $\mathscr{L}_{p\text{–NLS}}\boldsymbol{\phi} = \lambda\boldsymbol{\phi}$, where $\boldsymbol{\phi} = (\phi_1, \phi_2)^{\intercal}$ is a 2×1 column vector and ϕ_1, ϕ_2 are smooth functions with respect to spatial and time variables, the following change of variables [ZaS972]

$$\phi_1 = \sqrt{1-p}\, e^{-\imath\left(\frac{\lambda}{1-p^2}\right)x}\varphi_2, \quad \phi_2 = \sqrt{1+p}\, e^{-\imath\left(\frac{\lambda}{1-p^2}\right)x}\varphi_1, \quad q \equiv \frac{\imath u}{\sqrt{1-p^2}}, \quad \zeta \equiv \frac{\lambda p}{1-p^2}.$$

Performing these changes of variables leads to the eigenvalue problem [Kar998]

$$\begin{cases} \varphi_{1x} + \imath\zeta\varphi_1 = q(x,t)\varphi_2, \\ \varphi_{2x} - \imath\zeta\varphi_2 = -q^*(x,t)\varphi_1. \end{cases}$$

Replacing now q, $-q^*$ with the arbitrary (still smooth and slowly decreasing) functions $v = v(x,t)$, $u = u(x,t)$ respectively, we finally obtain the so–called **Zakharov–Shabat equations** [AbS981]

$$\begin{cases} \varphi_{1x} + \imath\zeta\varphi_1 = u(x,t)\varphi_2, \\ \varphi_{2x} - \imath\zeta\varphi_2 = v(x,t)\varphi_1, \end{cases}$$

written equivalently in matrix form as

$$\boldsymbol{\varphi}_x = \begin{pmatrix} \varphi_{1x} \\ \varphi_{2x} \end{pmatrix} = \begin{pmatrix} -\imath\zeta & u(x,t) \\ v(x,t) & \imath\zeta \end{pmatrix}\begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix} \equiv \mathcal{X}\boldsymbol{\varphi}. \tag{3.5}$$

REMARK 3.3 – Note that functions u, v play in equation (3.5) the same role that the *potential* $u = u(x,t)$ has in the Schrödinger equation. Also, if we choose $v(x,t) = -1$ in equations (3.5) and let $\varphi_2 \equiv \psi$, then after some simplifications, the Zakharov–Shabat system of equations reduces to $\psi_{xx} + \zeta^2\psi + u(x,t)\psi = 0$, which is nothing but the linear Schrödinger equation we used in §2.2 to solve the KdV equation.

3.2 The AKNS formalism and applications to the ZS spectral problem. We'll now follow an alternative method introduced by AKNS [AKNS74] in 1974 and inspired by the matrix form of the ZS equations. The procedure can be formulated as follows: consider the following two linear equations

$$\boldsymbol{\varphi}_x = \mathcal{X}\boldsymbol{\varphi}, \tag{3.6}$$
$$\boldsymbol{\varphi}_t = \mathcal{T}\boldsymbol{\varphi}, \tag{3.7}$$

where $\boldsymbol{\varphi} \neq \mathbf{0}$ is an n–dimensional vector and $\mathcal{X}, \mathcal{T}$ are $n \times n$ matrices. Cross differentiating equations (3.6), (3.7) and imposing the *compatibility condition* $\boldsymbol{\varphi}_{xt} = \boldsymbol{\varphi}_{tx}$, leads to $\big(\mathcal{X}_t - \mathcal{T}_x + [\mathcal{X}, \mathcal{T}]\big)\boldsymbol{\varphi} = \mathbf{0}$ which holds $\forall\, \boldsymbol{\varphi} \neq \mathbf{0}$. Therefore, we are left with the following equation

$$\mathcal{X}_t - \mathcal{T}_x + \big[\mathcal{X}, \mathcal{T}\big] = \mathbb{O}, \tag{3.8}$$

where $\mathbb{O}$ is the null $n \times n$ matrix. Equation (3.8) is equivalent to Lax equation and it's a matter of convention to refer to (3.6), (3.7) as to the **principal** and the **auxiliary spectral problem**, respectively.

It turns out that, given $\mathcal{X}$, there is a simple deductive procedure to find a $\mathcal{T}$ s.t. equation (3.8) contains a NLEE; of course, in order for (3.8) to be effective, the operator $\mathcal{X}$ should have a parameter which plays the role of an eigenvalue, say ζ, obeying condition $\zeta_t = 0$. Although the method is very general, we'll concentrate on the case of a 2×2 eigenvalue problem[32]. Then consider the ZS equations

[32] For a more general treatment see e.g. M. J. ABLOWITZ, H. SEGUR – *Solitons and the Inverse Scattering Transform.*

(3.5) and associate to them the most general form of a linear (but local) time evolution equation for the 2×1 column vector $\boldsymbol{\varphi}$ of functions φ_1, φ_2, i.e. [AbS981]

$$\begin{pmatrix} \varphi_{1x} \\ \varphi_{2x} \end{pmatrix} = \begin{pmatrix} -\imath\zeta & u(x,t) \\ v(x,t) & \imath\zeta \end{pmatrix} \begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix}, \qquad \begin{pmatrix} \varphi_{1t} \\ \varphi_{2t} \end{pmatrix} = \begin{pmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{pmatrix} \begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix}, \tag{3.9}$$

where $\mathcal{A}, \mathcal{B}, \mathcal{C}$ and $\mathcal{D}$ are scalar functions independent of $\boldsymbol{\varphi}$, but can depend on u, v and their derivatives. To simplify notations, let's rewrite equations (3.9) in terms of the Pauli σ_3 matrix as

$$\boldsymbol{\varphi}_x = \big(\mathcal{Q} - \imath\zeta\sigma_3\big)\boldsymbol{\varphi}, \qquad \boldsymbol{\varphi}_t = \mathcal{T}\boldsymbol{\varphi},$$

being $\mathcal{Q}$ a 2×2 matrix with entries $\mathcal{Q}_{12} = u$, $\mathcal{Q}_{21} = v$ and zeros on the main diagonal. Note that if $v = -1$, we get the linear Schrödinger equation (where ζ^2 plays the role of λ) as pointed out in §3.1 and if $v = \pm u^*$ (or $v = \pm u$, if u is real) one can also get some physically significant NLEE.

Since we desire to derive NLEEs for u and v that leave the spectrum of equations (3.9) invariant under their flow, we make use of condition $\zeta_t = 0$ and proceed in cross–differentiating equations (3.9); forcing the compatibility condition $\boldsymbol{\varphi}_{xt} = \boldsymbol{\varphi}_{tx}\ \forall\, \boldsymbol{\varphi} \neq \mathbf{0}$ yields (after some calculations) equation

$$\mathcal{Q}_t - \mathcal{T}_x = \big[\mathcal{T}, \mathcal{Q}\big] - \imath\zeta\big[\mathcal{T}, \sigma_3\big],$$

explicitly

$$\begin{pmatrix} -\mathcal{A}_x & u_t - \mathcal{B}_x \\ v_t - \mathcal{C}_x & -\mathcal{D}_x \end{pmatrix} = \begin{pmatrix} \mathcal{B}v - \mathcal{C}u & \big(\mathcal{A} - \mathcal{D}\big)u \\ -\big(\mathcal{A} - \mathcal{D}\big)v & -\big(\mathcal{B}v - \mathcal{C}u\big) \end{pmatrix} + 2\imath\zeta \begin{pmatrix} 0 & \mathcal{B} \\ -\mathcal{C} & 0 \end{pmatrix}.$$

Note that $\mathcal{A}_x = \mathcal{C}u - \mathcal{B}v = -\mathcal{D}_x$ thus, without loss of generality, we can take $\mathcal{A} = -\mathcal{D}$ in what follows. With this position, we restrict to the system of three differential equations with unknowns $\mathcal{A}, \mathcal{B}$ and $\mathcal{C}$

$$\begin{cases} \mathcal{A}_x &= \mathcal{C}u - \mathcal{B}v, \\ u_t &= \mathcal{B}_x + 2\mathcal{A}u + 2\imath\zeta\mathcal{B}, \\ v_t &= \mathcal{C}_x - 2\mathcal{A}v - 2\imath\zeta\mathcal{C}. \end{cases} \tag{3.10}$$

We want to solve the system (3.10) so that equations (3.6), (3.7) are compatible. Doing this, we'll see that still another condition must to be satisfied, which turns out to be an evolution equation. Various methods are feasible to solve system (3.10): here we're going to adopt an expansion procedure [AbS981].

Since ζ is a free parameter (it might be small), we try for a polynomial solution to (3.10) in powers of ζ. Let's start with a first order polynomial, i.e. suppose that $\mathcal{A} = \mathcal{A}_0 + \zeta\mathcal{A}_1$, $\mathcal{B} = \mathcal{B}_0 + \zeta\mathcal{B}_1$ and $\mathcal{C} = \mathcal{C}_0 + \zeta\mathcal{C}_1$ are solutions of the set of equations (3.10). This yields the system

$$\begin{cases} \big(\mathcal{A}_{0x} - \mathcal{C}_0 u + \mathcal{B}_0 v\big) + \big(\mathcal{A}_{1x} - \mathcal{C}_1 u + \mathcal{B}_1 v\big)\zeta = 0, \\ u_t = \big(\mathcal{B}_{0x} + 2\mathcal{A}_0 u\big) + \big(\mathcal{B}_{1x} + 2\mathcal{A}_1 u + 2\imath\mathcal{B}_0\big)\zeta + 2\imath\mathcal{B}_1\zeta^2, \\ v_t = \big(\mathcal{C}_{0x} - 2\mathcal{A}_0 v\big) + \big(\mathcal{C}_{1x} - 2v\mathcal{A}_1 - 2\imath\mathcal{C}_0\big)\zeta - 2\imath\mathcal{C}_1\zeta^2. \end{cases}$$

Now equate coefficients belonging to same powers of ζ. For ζ^2 one finds $\mathcal{B}_1 = 0 = \mathcal{C}_1$, whilst ζ gives

$$\begin{cases} \mathcal{A}_{1x} &= \mathcal{C}_1 u - \mathcal{B}_1 v, \\ \mathcal{B}_{1x} &= -2\imath\mathcal{B}_0 - 2\mathcal{A}_1 u, \\ \mathcal{C}_{1x} &= 2\imath\mathcal{C}_0 + 2\mathcal{A}_1 v, \end{cases} \xrightarrow{\mathcal{B}_1 = 0 = \mathcal{C}_1} \begin{cases} \mathcal{A}_1(x,t) &= a_1(t), \\ \mathcal{B}_0(x,t) &= \imath a_1(t) u(x,t), \\ \mathcal{C}_0(x,t) &= \imath a_1(t) v(x,t); \end{cases}$$

finally, equating ζ^0–coefficients and substituting what previously obtained, we get

$$\begin{cases} \mathcal{A}_{0x} &= \mathcal{C}_0 u - \mathcal{B}_0 v, \\ u_t &= \mathcal{B}_{0x} + 2\mathcal{A}_0 u, \\ v_t &= \mathcal{C}_{0x} - 2\mathcal{A}_0 v, \end{cases} \xrightarrow[\mathcal{A}_0(x,t) = a_0(t)]{\mathcal{A}_{0x} = \imath a_1[v,u] = 0} \begin{cases} u_t(x,t) = \imath a_1(t) u_x(x,t) + 2a_0(t) u(x,t), \\ v_t(x,t) = \imath a_1(t) v_x(x,t) - 2a_0(t) v(x,t). \end{cases} \tag{3.11}$$

Therefore first order polynomial solutions of (3.10) doesn't yield interesting evolution equations for ""potentials" u and v, since the ones we've found are both linear in their arguments.

Let's try second order polynomials, i.e. suppose that $\mathcal{A} = \mathcal{A}_0 + \zeta\mathcal{A}_1 + \zeta^2\mathcal{A}_2$, $\mathcal{B} = \mathcal{B}_0 + \zeta\mathcal{B}_1 + \zeta^2\mathcal{B}_2$ and $\mathcal{C} = \mathcal{C}_0 + \zeta\mathcal{C}_1 + \zeta^2\mathcal{C}_2$ are solutions to (3.10). Substitution yields the following system

$$\begin{cases} \big(\mathcal{A}_{0x} - \mathcal{C}_0 u + \mathcal{B}_0 v\big) + \big(\mathcal{A}_{1x} - \mathcal{C}_1 u + \mathcal{B}_1 v\big)\zeta + \big(\mathcal{A}_{2x} - \mathcal{C}_2 u + \mathcal{B}_2 v\big)\zeta^2 = 0, \\ u_t = \big(\mathcal{B}_{0x} + 2u\mathcal{A}_0\big) + \big(\mathcal{B}_{1x} + 2u\mathcal{A}_1 + 2\imath\mathcal{B}_0\big)\zeta + \big(\mathcal{B}_{2x} + 2u\mathcal{A}_2 + 2\imath\mathcal{B}_1\big)\zeta^2 + 2\imath\mathcal{B}_2\zeta^3, \\ v_t = \big(\mathcal{C}_{0x} - 2v\mathcal{A}_0\big) + \big(\mathcal{C}_{1x} - 2v\mathcal{A}_1 - 2\imath\mathcal{C}_0\big)\zeta + \big(\mathcal{C}_{2x} - 2v\mathcal{A}_2 - 2\imath\mathcal{C}_1\big)\zeta^2 - 2\imath\mathcal{C}_2\zeta^3. \end{cases}$$

Equating ζ^3–coefficients immediately gives $\mathcal{B}_2 = 0 = \mathcal{C}_2$, whilst from ζ^2 and ζ one gets respectively

$$\zeta^2 : \begin{cases} \mathcal{A}_{2x} = \mathcal{C}_2 u - \mathcal{B}_2 v, \\ \mathcal{B}_{2x} = -2u\mathcal{A}_2 - 2\imath\mathcal{B}_1, \\ \mathcal{C}_{2x} = 2v\mathcal{A}_2 + 2\imath\mathcal{C}_1, \end{cases} \xrightarrow{\mathcal{B}_2 = 0 = \mathcal{C}_2} \begin{cases} \mathcal{A}_2(x,t) = a_2(t), \\ \mathcal{B}_2(x,t) = \imath a_2(t) u(x,t), \\ \mathcal{C}_2(x,t) = \imath a_2(t) v(x,t). \end{cases}$$

$$\zeta : \begin{cases} \mathcal{A}_{1x} = \mathcal{C}_1 u - \mathcal{B}_1 v, \\ \mathcal{B}_{1x} = -2u\mathcal{A}_1 - 2\imath\mathcal{B}_0, \\ \mathcal{C}_{1x} = 2v\mathcal{A}_1 + 2\imath\mathcal{C}_0, \end{cases} \xrightarrow[\mathcal{C}_1 = \imath a_2 v]{\mathcal{B}_1 = \imath a_2 u} \begin{cases} \mathcal{A}_1(x,t) = a_1(t), \\ \mathcal{B}_0(x,t) = \imath a_1(t) u(x,t) - \frac{1}{2} a_2(t) u_x(x,t), \\ \mathcal{C}_0(x,t) = \imath a_1(t) v(x,t) + \frac{1}{2} a_2(t) v_x(x,t). \end{cases}$$

Note that ζ^3,ζ^2–coefficients for second order polynomials gives formally the same results obtained from ζ^2,ζ–coefficients for first order polynomials. Finally, equating ζ^0–coefficients yields

$$\begin{cases} \mathcal{A}_{0x} = \mathcal{C}_0 u - \mathcal{B}_0 v, \\ u_t = \mathcal{B}_{0x} + 2u\mathcal{A}_0, \\ v_t = \mathcal{C}_{0x} - 2v\mathcal{A}_0, \end{cases} \xrightarrow[\mathcal{A}_0 = \frac{a_2}{2} uv + a_0]{\mathcal{A}_{0x} = \imath a_1 [v,u] + \frac{a_2}{2}(uv)_x = \frac{a_2}{2}(uv)_x} \begin{cases} u_t = \imath a_1 u_x - \frac{a_2}{2} u_{xx} + a_2 u^2 v + 2a_0 u, \\ v_t = \imath a_1 v_x + \frac{a_2}{2} v_{xx} - a_2 v^2 u - 2a_0 v. \end{cases} \tag{3.12}$$

Therefore, following the AKNS formalism we have obtained a general family of NLEEs associated to the ZS problem. In particular, taking for simplicity $a_1(t) = 0 = a_0(t)$ (i.e. eliminating from equations (3.12) the LEEs (3.11) obtained from first order polynomials), the coupled NLEEs (3.12) get the form

$$\begin{cases} u_t - a_2 u^2 v + \frac{1}{2} a_2 u_{xx} = 0, \\ v_t + a_2 u v^2 - \frac{1}{2} a_2 v_{xx} = 0, \end{cases} \tag{3.13}$$

which are reminiscent of NLS (1.23). Indeed the latter equation results if we let $v = \varepsilon u^*$ with $\varepsilon = \pm 1$:

$$\begin{cases} u_t - \varepsilon a_2 u|u|^2 + \frac{1}{2} a_2 u_{xx} = 0, \\ u_t^* + \varepsilon a_2 u^* |u|^2 - \frac{1}{2} a_2 u_{xx}^* = 0. \end{cases} \tag{3.14}$$

Equations (3.14) are compatible iif $a_2^* = -a_2$, i.e. iff $a_2 = \imath\alpha$ with $\alpha \in \mathbb{R}$; it's a common choice to set $\alpha = 2$. Thus system (3.14) reduces to the *attractive* (+ sign) or *repulsive* (− sign) NLS equation

$$\imath u_t = u_{xx} \pm 2u|u|^2. \tag{3.15}$$

Note that for the attractive NLS equation soliton solutions can be found, but no soliton solutions exists in the repulsive one for rapidly decaying potentials [AbS981].

In summary, considering the ZS spectral problem $\boldsymbol{\varphi}_x = (Q - \imath\zeta\sigma_3)\boldsymbol{\varphi}$ and the associated time–evolution equation $\boldsymbol{\varphi}_t = \mathcal{T}\boldsymbol{\varphi}$, the compatibility condition is given by equations (3.8). Searching for second order polynomial solutions, we've shown that condition (3.10) holds iff potentials u, v, with $v = \pm u^*$, satisfy a NLS equation (3.15) and that, with this choice, the operator $\mathcal{T}$ has the form

$$\mathcal{T}_{\text{NLS}} = \begin{pmatrix} \imath\varepsilon|u|^2 + 2\imath\zeta^2 & -\imath u_x - 2\zeta u \\ \imath\varepsilon u_x^* - 2\varepsilon\zeta u^* & -\imath\varepsilon|u|^2 - 2\imath\zeta^2 \end{pmatrix}.$$

Of course one could guess if there are other interesting NLEEs for higher order polynomials. Indeed this is the case for third order polynomials, which is the last example we're going to treat in this

paragraph. Thus assume that $\mathcal{A} = \mathcal{A}_0 + \zeta\mathcal{A}_1 + \zeta^2\mathcal{A}_2 + \zeta^3\mathcal{A}_3$ and analogue expressions for $\mathcal{B}$ and $\mathcal{C}$ are solutions of system (3.10). Therefore, one finds the following results [AbS981]

$$\begin{aligned}
\mathcal{A} &= a_3\zeta^3 + a_2\zeta^2 + \tfrac{1}{2}(a_3uv + a_1)\zeta + \tfrac{1}{2}a_2uv - \tfrac{\imath}{4}a_3(uv_x - u_xv) + a_0,\\
\mathcal{B} &= \imath a_3u\zeta^2 + \big(\imath a_2u - \tfrac{1}{2}a_3u_x\big)\zeta + \big(\imath a_1u + \tfrac{\imath}{2}a_3u^2v - \tfrac{1}{2}a_2u_x - \tfrac{\imath}{4}a_3u_{xx}\big),\\
\mathcal{C} &= \imath a_3v\zeta^2 + \big(\imath a_2v + \tfrac{1}{2}a_3v_x\big)\zeta + \big(\imath a_1v + \tfrac{\imath}{2}a_3v^2u + \tfrac{1}{2}a_2v_x - \tfrac{\imath}{4}a_3v_{xx}\big),
\end{aligned} \tag{3.16}$$

and the nonlinear evolution equations obtained from the compatibility condition (3.10) are

$$\begin{cases}
u_t + \tfrac{\imath}{4}a_3\big(u_{xxx} - 6uvu_x\big) + \tfrac{1}{2}a_2\big(u_{xx} - 2u^2v\big) - \imath a_1u_x - 2a_0u = 0,\\
v_t + \tfrac{\imath}{4}a_3\big(v_{xxx} - 6uvv_x\big) - \tfrac{1}{2}a_2\big(v_{xx} - 2v^2u\big) - \imath a_1v_x + 2a_0v = 0.
\end{cases} \tag{3.17}$$

Evolution equations of physical interest are obtained from (3.17) as special cases; in particular, we find

$$\begin{aligned}
&\begin{pmatrix} a_0 = a_1 &= a_2 = 0\\ a_3 &= -4\imath \end{pmatrix} \quad \Longrightarrow \quad \mathcal{K}[u] = 0 \quad \text{if} \quad v = -1,\\
&\begin{pmatrix} a_0 = a_1 = a_3 &= 0\\ a_2 = 2\imath, \quad v &= \mp u^* \end{pmatrix} \quad \Longrightarrow \quad \mathcal{M}[u] = 0 \quad \text{if} \quad v = \mp u^*;\\
&\qquad\qquad\qquad\qquad\quad \Longrightarrow \quad \imath u_t = u_{xx} \pm u|u|^2 \quad \text{(NLS)}.
\end{aligned}$$

Until now we have considered only polynomial expansions of $\mathcal{A}, \mathcal{B}$ and $\mathcal{C}$ corresponding to *positive* powers of ζ. However, we may also find interesting NLEEs corresponding to expansion in *inverse powers* of ζ (or both) [AbS981]. Considerer, as an example, the following choice

$$\mathcal{A}(x,t) = \frac{a(x,t)}{\zeta}, \qquad \mathcal{B}(x,t) = \frac{b(x,t)}{\zeta}, \qquad \mathcal{C}(x,t) = \frac{c(x,t)}{\zeta}.$$

Substituting in the compatibility condition (3.10) and equating then ζ^0, ζ^{-1}–coefficients yields

$$\begin{aligned}
\zeta^0 \ : &\begin{cases} a_x = cu - bv,\\ u_t = 2\imath b,\\ v_t = -2\imath c, \end{cases}\\
\zeta^{-1} \ : &\begin{cases} b_x = -2au,\\ c_x = 2av, \end{cases}
\end{aligned} \quad \xrightarrow[c(x,t)=\frac{\imath}{2}v_t(x,t)]{b(x,t)=-\frac{\imath}{2}u_t(x,t)} \quad \begin{cases} u_{tx} = -4\imath au,\\ v_{tx} = -4\imath av,\\ a_x = \frac{\imath}{2}(uv)_t. \end{cases} \tag{3.18}$$

System (3.18) has other important NLEEs as special cases; well–known examples are

$$\begin{aligned}
&\begin{pmatrix} a(x,t) = \frac{\imath}{4}\cos\big[w(x,t)\big]\\ v(x,t) = -u(x,t) = \frac{1}{2}w_x(x,t) \end{pmatrix} \quad \Longrightarrow \quad w_{xt}(x,t) = \sin\big[w(x,t)\big] \quad \text{(SG)};\\
&\begin{pmatrix} a(x,t) = \frac{\imath}{4}\cosh\big[w(x,t)\big]\\ u(x,t) = v(x,t) = \frac{1}{2}w_x(x,t) \end{pmatrix} \quad \Longrightarrow \quad w_{xt}(x,t) = \sinh\big[w(x,t)\big] \quad \text{(ShG)}.
\end{aligned}$$

Note that the presence of the factor $1/2$ in the choices made for u, v in both above cases, is inserted in order to satisfy the third equation in system (3.18), i.e. equation $a_x = \frac{\imath}{2}(uv)_t$.

Indeed many other (somehow less famous) NLEEs can be derived by taking combinations of previous procedures or different expansions from above; of course, all these procedures gives exactly the same results as does Lax approach (3.2), but are simpler because of their *algebraic* nature [AbS981].

4 The classical Bäcklund Transformation (BT)

The previous sections were mainly devoted to introduce the basics of Soliton Theory. In particular, we have underlined the importance of the IST method and its connection with multi–solitons solutions of NLEEs, though stressing its limits and the difficulties related to the inverse problem. Furthermore, we have emphasized that we still don't have a definite criterion to characterize IST–solvable NLEEs, meaning that we still ignore if any of the common characteristics encountered so far (i.e. multi–solitons solutions, Arnol'd–Liouville integrability, (bi–)Hamiltonian structure, Lax pairs) is both necessary and sufficient for a system to be IST–solvable. In this section we're going to discuss a technique that is generally believed to be a *sufficient condition* for IST in $(1+1)$–dimensions, that is the *existence of a Bäcklund Transform* (BT). As we shall see, if this method holds then one can deduce *without quadratures* the multi–solitons solutions for the NLEEs under scrutiny.

The BT has its origins in a contest apparently far from Soliton Theory, namely in differential geometry. It was discovered in 1875 by A. V. Bäcklund (1845–1922) while studying some particular transformations between surfaces with constant negative total curvature, often referred as *pseudospherical surfaces* [RoS002]; some years later, it was shown by L. Bianchi (1856–1928) that this particular transformation was associated to an elegant invariance of the sine–Gordon equation[33] [Bia885, Bia879]. To introduce BT, consider a suitable system of coordinates $(u,v)\in\mathbb{R}^2$, parametrizing a surface $\mathcal{S}$ imbedded in $\mathbb{R}^3$. The line element of $\mathcal{S}$ is given by the first fundamental quadratic form $\mathrm{d}\,s^2 = g_{\mu\nu}\,\mathrm{d}\,u^\mu\,\mathrm{d}\,v^\nu$ for $\mu,\nu=1,2$, being $g_{\mu\nu}=g_{\nu\mu}$ the symmetric metric tensor associated to $\mathcal{S}$ [Eis947]. For surfaces of constant negative *total* curvature, the coordinates can be chosen s.t. $g_{11}=g_{22}=\rho^2$ and $g_{21}=\rho^2\cos\omega$, where $\rho\in\mathbb{R}^*$ is related to the total curvature $\mathcal{K}$ via the relation $\mathcal{K}=-\rho^{-2}\in\mathbb{R}^-$. Therefore, the line element of a pseudospherical surfaces with total curvature $\mathcal{K}$ may be written as

$$\mathrm{d}\,s^2 = -\tfrac{1}{\mathcal{K}}\big(\mathrm{d}\,u^2 + 2\cos\omega\,\mathrm{d}\,u\,\mathrm{d}\,v + \mathrm{d}\,v^2\big),$$

where $\omega(u,v)\in[0,2\pi)$ is the angle between two asymptotic lines and satisfies the well–known *Gauss-Codazzi–Mainardi equations*. Invoking the Gauss' *theorema egregium*, the latter equations simply reduce to the sine–Gordon equation (on the light–cone) $\omega_{uv}=-\mathcal{K}\sin\omega$ [RoS002]; therefore, each solution $\omega=\omega(u,v)$ will corresponds to a pseudospherical surface with curvature $\mathcal{K}=-\rho^{-2}$. Well known solutions are illustrated in Figure 7 and in Figure 8 below [McL994].

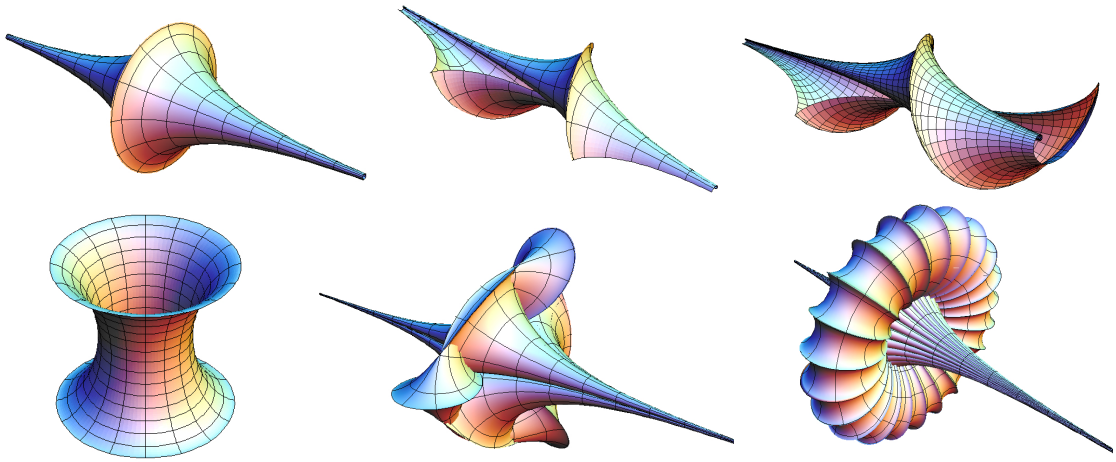


Figure 7: *Pseudosphere* obtained as surface of revolution of the tractrix (upper left), *Dini's surface* or twisted pseudosphere (upper center), *Kink* surface (upper right), *hyperboloid* of one sheet (lower left), *breather* surface (lower center) and *parametric breather* (lower right); from R. Palais' Virtual Math Museum.

[33] Which was introduced in 1862 by the engineer E. Bour (1832–1866), still in the contest of differential geometry.

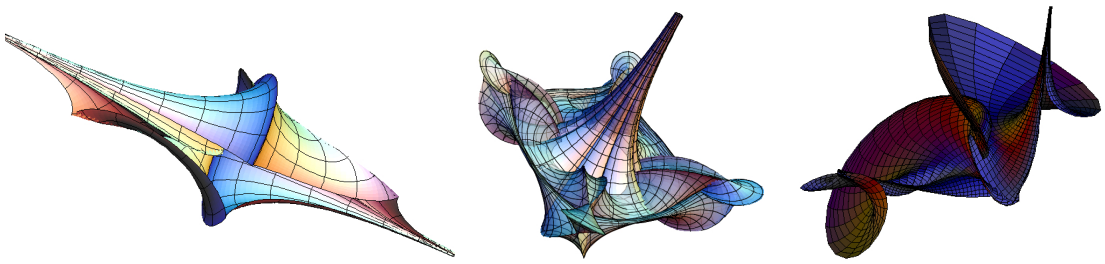


Figure 8: Two–solitons (left), breather–soliton (center) and three–solitons (right); from Virtual Math Museum.

Now note that the SG equation is invariant under the scaling (Lie point symmetry) $u' = \beta u$, $v' = \beta^{-1} v$ with $\beta \in \mathbb{R}^*$, so each solution $\omega = \omega(u, v)$ generates a one–parameter class of solutions $\omega'(u', v') = \omega(\beta u, \beta^{-1} v)$. It was the quest for additional techniques for generating such surfaces that led Bäcklund to introduce the transformation bearing nowadays his name: he founded that if ω_0 is a solution of the SG, then a new solution ω_1 can be obtained by means of the following relations [RoS002, Lam976]:

$$\mathbb{B}_\beta \;:\; \begin{cases} \dfrac{\partial}{\partial u}\left(\dfrac{\omega_1 - \omega_0}{2}\right) = \dfrac{\rho}{\beta}\sin\left(\dfrac{\omega_1 + \omega_0}{2}\right), \\ \dfrac{\partial}{\partial v}\left(\dfrac{\omega_1 + \omega_0}{2}\right) = \dfrac{1}{\rho\beta}\sin\left(\dfrac{\omega_1 - \omega_0}{2}\right), \end{cases} \tag{4.1}$$

which represent the standard form of the **Bäcklund transform** (hereafter indicated as $\mathbb{B}_\beta$) for the Sine–Gordon equation[34]. The parameter $\beta \in \mathbb{R}^*$ is called the *BT parameter*. Thus $\mathbb{B}_\beta$ defines a mapping between pseudospherical surfaces (or, equivalently, SG solutions) which preserves the scalar curvature; in particular, if $\mathbb{B}_\beta$ is a correspondence between solutions of the *same* equation (which is of course the case for the SG), then $\mathbb{B}_\beta$ is commonly termed an **auto–Bäcklund transformation** (aBT).

Remark 4.1 – The BT wouldn't be useful if a first solution ω_0 could not be obtained; fortunately, the so–called "*seed vacuum*" solution $\omega_0 = 0$ provides a basis for constructing further solutions. As an example, let's calculate the second non–trivial SG–solution: setting $\omega_0 \equiv 0$ in (4.1) yields

$$\begin{cases} \omega_{1u} = \dfrac{2\beta}{\rho}\sin\omega_1, \\ \omega_{1v} = \dfrac{2}{\beta\rho}\sin\omega_1, \end{cases} \xrightarrow{\int \frac{\mathrm{d}\chi}{\sin\chi} = \ln\left|\tan\frac{\chi}{2}\right| + \mathrm{cost}} \begin{cases} \ln\left|\tan\dfrac{\omega_1}{4}\right| = \dfrac{\beta}{\rho}u + \bar{c}(u), \\ \ln\left|\tan\dfrac{\omega_1}{4}\right| = \dfrac{1}{\beta\rho}v + \tilde{c}(v). \end{cases}$$

The last two expression implies that $\bar{c}(u) = \tilde{c}(v) \equiv c$, leading to the *one–soliton*[35] SG–solution

$$\omega_1(u, v) = 4\arctan\left(e^{\frac{\beta}{\rho}u + \frac{1}{\beta\rho}v + c}\right). \tag{4.2}$$

Analytic expressions for multi–solitons solutions which encapsulate their nonlinear interaction may be obtained by a purely algebraic procedure. This is the consequence of an elegant theorem proved in 1892 by L. Bianchi and derived from the a–BT $\mathbb{B}_\beta$, known nowadays as the Bianchi's Permutability Theorem.

---

[34] Since we're not going to show how these equations are obtained [RoS002], let's simply verify that indeed ω_0, ω_1 satisfy the SG equation. Cross differentiation of equations (4.1) and the compatibility conditions $\partial_{uv}\omega_i = \partial_{vu}\omega_i$ yield

$$\begin{array}{l} \left(\frac{\omega_1-\omega_0}{2}\right)_{uv} = \frac{\beta}{\rho}\cos\left(\frac{\omega_1+\omega_0}{2}\right)\left(\frac{\omega_1+\omega_0}{2}\right)_v = \frac{1}{\rho^2}\cos\left(\frac{\omega_1+\omega_0}{2}\right)\sin\left(\frac{\omega_1-\omega_0}{2}\right), \\ \left(\frac{\omega_1+\omega_0}{2}\right)_{vu} = \frac{\beta}{\rho}\cos\left(\frac{\omega_1-\omega_0}{2}\right)\left(\frac{\omega_1-\omega_0}{2}\right)_u = \frac{1}{\rho^2}\cos\left(\frac{\omega_1-\omega_0}{2}\right)\sin\left(\frac{\omega_1+\omega_0}{2}\right), \end{array} \xrightarrow[\text{add}]{\text{sub}} \begin{array}{l} \omega_{0uv} = \frac{1}{\rho^2}\sin\omega_0, \\ \omega_{1uv} = \frac{1}{\rho^2}\sin\omega_1. \end{array}$$

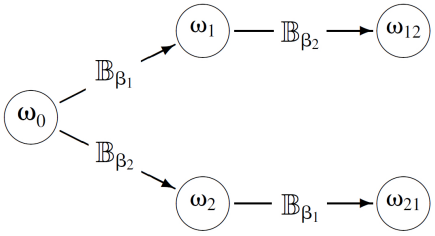
[35] Actually, it should be noted that (in light of the identity $\sin[2\tan^{-1}(e^{\chi+y})] = \mathrm{sech}(\chi + y)$) it is the quantities

$$\omega_{1u} = \frac{2\beta}{\rho}\mathrm{sech}\left(\frac{\beta}{\rho}u + \frac{1}{\beta\rho}v + c\right), \qquad \omega_{1v} = \frac{2}{\beta\rho}\mathrm{sech}\left(\frac{\beta}{\rho}u + \frac{1}{\beta\rho}v + c\right),$$

which have the characteristic hump shape associated with a soliton solution of a NLEE [RoS002].

4.1 Bianchi's Permutability Theorem. Suppose that ω_0 is a seed solution of the SG equation and let ω_1, ω_2 be the BT of ω_0 obtained respectively via relations $\omega_1(u,v) \equiv \mathbb{B}_{\beta_1}[\omega_0](u,v)$ and $\omega_2(u,v) \equiv \mathbb{B}_{\beta_2}[\omega_0](u,v)$, being $\beta_1\beta_2 \in \mathbb{R}^*$; therefore, recalling the expression (4.2), one has

$$\omega_i(u,v) = 4\arctan\left(e^{\alpha_i}\right), \qquad \alpha_i \equiv \frac{\beta_i}{\rho}u + \frac{1}{\beta_i\rho}v + c_i, \quad i = 1,2.$$



Now define $\omega_{12} := \mathbb{B}_{\beta_2}[\omega_1]$ and $\omega_{21} := \mathbb{B}_{\beta_1}[\omega_2]$; the situation is schematically represented with the non–commutative *Bianchi diagram* on the left [RoS002]. It's natural to ask if there are any conditions under which this Bianchi diagram becomes *commutative*, i.e. under what circumstances the identity

$$\mathbb{B}_{\beta_1}\mathbb{B}_{\beta_2} = \mathbb{B}_{\beta_2}\mathbb{B}_{\beta_1} \tag{4.3}$$

holds. To investigate this matter, let us consider the u and v–components of the BTs for the non–commutative Bianchi diagram, respectively given by expressions

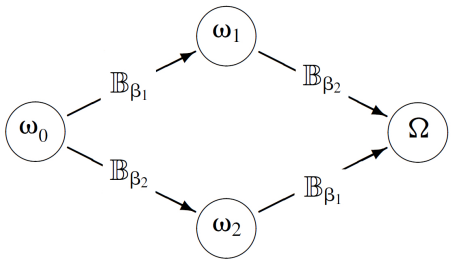
$$\omega_{1u} = \omega_{0u} + \frac{2\beta_1}{\rho}\sin\left(\frac{\omega_1+\omega_0}{2}\right), \qquad \omega_{1v} = -\omega_{0u} + \frac{2\beta_1}{\rho}\sin\left(\frac{\omega_1-\omega_0}{2}\right),$$
$$\omega_{12u} = \omega_{1u} + \frac{2\beta_2}{\rho}\sin\left(\frac{\omega_{12}+\omega_1}{2}\right), \qquad \omega_{12v} = -\omega_{1u} + \frac{2\beta_2}{\rho}\sin\left(\frac{\omega_{12}-\omega_1}{2}\right),$$
$$\omega_{2u} = \omega_{0u} + \frac{2\beta_2}{\rho}\sin\left(\frac{\omega_2+\omega_0}{2}\right), \qquad \omega_{2v} = -\omega_{0u} + \frac{2\beta_2}{\rho}\sin\left(\frac{\omega_2-\omega_0}{2}\right),$$
$$\omega_{21u} = \omega_{2u} + \frac{2\beta_1}{\rho}\sin\left(\frac{\omega_{21}+\omega_2}{2}\right). \qquad \omega_{21v} = -\omega_{2u} + \frac{2\beta_1}{\rho}\sin\left(\frac{\omega_{21}-\omega_2}{2}\right).$$

Inserting the first and third equations in the second and forth ones and setting then $\omega_{12} = \omega_{21} \equiv \Omega$, one obtains respectively from the u and v components, the following relations

$$\beta_1\left[\sin\left(\frac{\omega_1+\omega_0}{2}\right) - \sin\left(\frac{\Omega+\omega_2}{2}\right)\right] = \beta_2\left[\sin\left(\frac{\omega_2+\omega_0}{2}\right) - \sin\left(\frac{\Omega+\omega_1}{2}\right)\right],$$
$$\beta_1\left[\sin\left(\frac{\Omega-\omega_1}{2}\right) + \sin\left(\frac{\omega_2-\omega_0}{2}\right)\right] = \beta_2\left[\sin\left(\frac{\Omega-\omega_2}{2}\right) + \sin\left(\frac{\omega_1-\omega_0}{2}\right)\right].$$

Now subtract the first equation from the second one, invoke the addition and subtraction formulas for the sine function and the identity $\tan(\frac{\alpha-\beta}{2}) = \frac{\sin\alpha-\cos\beta}{\sin\alpha+\cos\beta}$ and rearrange properly the terms of the equation; one finally arrives to the formula

$$\tan\left(\frac{\Omega-\omega_0}{4}\right) = \frac{\beta_2+\beta_1}{\beta_2-\beta_1}\tan\left(\frac{\omega_2-\omega_1}{4}\right), \tag{4.4}$$



which is the main result of the *Bianchi's permutability theorem* [Bia923]. Therefore, if the condition (4.3) holds, a new SG's solution Ω can be constructed from equation (4.4), for a given seed ω_0, and the Bianchi diagram becomes commutative, as indicated in the figure on the right.

REMARK 4.2 – Since the permutability theorem (4.4) allows to construct new solutions for the SG equation from older ones, it's usually referred in the literature as the SG–**nonlinear superposition principle**, in the sense that if $\omega_0 = \omega_0(u,v)$ is a seed solution of the SG equation and Ω is obtained via the BT with the relation $\Omega = \mathbb{B}_{\beta_1}\left[\mathbb{B}_{\beta_2}[\omega_0]\right]$, then $\Omega = \Omega(u,v)$ is also a solution of the SG equation.

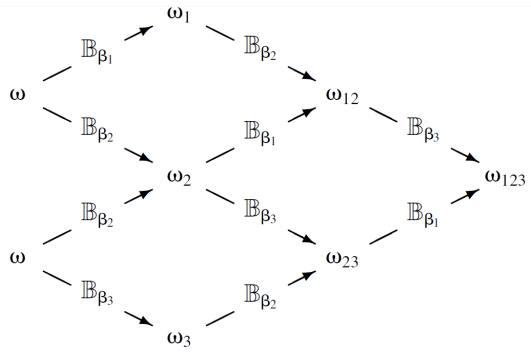


Figure 9: Three–folded Bianchi lattice [RoS002].

The commutative property now allows to construct a *Bianchi lattice*, obtained by iterated applications of the permutability theorem (4.4) (see Figure 9). Therefore *N–solitons solutions* of the SG equation may be generated by purely *algebraic* procedures, each corresponding to a nonlinear superposition of N single soliton solutions, with Bäcklund parameters $\beta_1, \beta_2, \ldots, \beta_N$. As an example, let's calculate the two–solitons solution for the SG. The expression follows immediately from (4.4) with ω_1, ω_2 given by formula (4.2): thus we have

$$\omega_{2\text{–sol}}^{(\mathrm{SG})} = 4 \arctan \left[\frac{\beta_2 + \beta_1}{\beta_2 - \beta_1} \tan \left(\frac{\omega_2 - \omega_1}{4} \right) \right].$$

Invoking therefore the trigonometric subtraction formula $\tan(\alpha - \beta) = \frac{\tan\alpha - \tan\beta}{1 + \tan\alpha \tan\beta}$ and rearranging the exponential terms in the expression, one can finally put the two–solitons solution in the following form:

$$\omega_{2\text{–sol}}^{(\mathrm{SG})}(u,v) = 4 \arctan \left\{ \frac{(\beta_2 - \beta_1) \sinh \left[\frac{u}{2\rho}(\beta_2 - \beta_1) + \frac{v}{2\rho}(\beta_2^{-1} - \beta_1^{-1}) + c \right]}{(\beta_2 - \beta_1) \cosh \left[\frac{u}{2}\rho(\beta_2 + \beta_1) + \frac{v}{2\rho}(\beta_2^{-1} + \beta_1^{-1}) + c' \right]} \right\}, \tag{4.5}$$

where $c \equiv c_2 - c_1$ and $c' \equiv c_2 + c_1$. Note that solutions (4.2) and (4.5) represent the angles between two asymptotic lines of different pseudospherical surfaces, whose expressions may be obtained after substituting the equations (4.2), (4.5) in the first fundamental quadratic form (now associated to a surface of revolution) and integrating the expression obtained [RoS002]. Then one finds that the surface corresponding to the one–soliton solution (4.2) is given by the *Beltrami's pseudosphere* (see Figure 7) and the one associated to the two–solitons solution (4.5) is the so–called *two–soliton surface* depicted in the image on the left of Figure 8. It should be also emphasized that other interesting solutions may be obtained by starting from different seed solutions ω_0 and by searching for periodic solutions, known as **breathers**. For example, the so–called *two–soliton breathers* may be obtained from solution (4.5) using complex Bäcklund parameters $\beta_{1,2} = z \pm \imath w$ and introducing a proper coordinate systems[36]; the surface associate to the two–solitons breather solution is depicted in Figure 8.

4.2 BT in other setting and the connection with IST. So far we've considered the BT for the SG equation; we wish now to generalize the argument to other nonlinear systems. Before starting, however, it should be stressed that since there is *no unique* definition of BT, there might be different approaches to the general problem; in this section we'll mainly follow the argumentations proposed by Ablowitz and Segur in the third chapter of their monograph[37] [AbS981]. Let's start with the following

Definition 4.1 - *Let* $\mathcal{D}[u] = 0$ *and* $\mathcal{E}[v] = 0$ *be two PDEs. A (set of) relation(s) of the form*

$$\mathbb{L}(u, v, u_x, v_x, u_t, v_t, \ldots; x, t) = 0$$

is said to map $\mathcal{E}$ into $\mathcal{D}$ *if every (local) solution of* $\mathcal{E}[u] = 0$ *uniquely defines a (local) solution of* $\mathcal{D}[v] = 0$.

Note that in this paragraph u and v are smooth functions of the independent variables $(x, t) \in \mathbb{R}^2$.

We have already encounter such relations: indeed the *Miura transform* $u = -v_x - v^2$ (see §2.1) is a map of mKdV into KdV, i.e. $\mathcal{K}[u] = -(\partial_x + 2v)\mathcal{M}[v]$ and the same is true for the *Hopf–Cole transform* $u \equiv -2\nu \frac{v_x}{v}$ (see §1.2), which maps solutions of the heat equation into solutions of the Burgers' one.

[36] See e.g. C. Rogers, W. K. Schief – *Bäcklund and Darboux transformations*, §1.4, pgg. 31–41.

[37] Another approach, based on a different BT definition, can be found in the work of R. L. Anderson, N. J. Ibragimov (1979) for the *Lie–Backlund transformations* [AnI979] and in the successive papers of N. J. Ibragimov, A. B. Shabat [IbS979] and A. S. Fokas [Fok980]; another one, based on *local jet–bundles*, was given by F. A. E. Pirani [Pir979].

Let's therefore give the Ablowitz–Segur definition for the BT (AS–BT) [AbS981].

DEFINITION 4.2 - *A set of relations involving* $\{x,t;u(x,t)\}$, $\{\chi,t;v(\chi,t)\}$ *and the derivatives of* u *and* v *is a* ***Bäcklund transform*** *between* $\mathcal{D}[u;x,t]=0$ *and* $\mathcal{E}[v;\chi,t]=0$ *if:*

- ◇ *the BT is integrable*[38] *for* v *(respectively for* u*) iff* $\mathcal{D}[u]=0$ $(\mathcal{E}[v]=0)$*;*
- ◇ *given* u *(respectively* v*) s.t.* $\mathcal{D}[u]=0$ $(\mathcal{E}[v]=0)$*, the BT defines* v (u) *to within a finite set of constants and* $\mathcal{E}[v]=0$ $(\mathcal{D}[v]=0)$*.*

Besides the BT for the SG, that we've already discussed, there exist other well known transformations in the literature that satisfy the AS–BT: e.g. in the theory of complex variables, the *Cauchy–Riemann conditions* (giving a necessary and sufficient condition for a function to be *complex differentiable*)

$$u_x = v_y, \qquad u_y = -v_x,$$

identify an auto–BT for the Laplace equation, in fact one has respectively that $u_{xx}+u_{yy}=0$ and $v_{xx}+v_{yy}=0$. Then, given v satisfying the Laplace equation, one has that the CR–conditions defines a new solution u of the same equation, to within a constant $u_0=u(x_0,y_0)$. Another interesting example is the *scattering problem for the KdV*, given by the set of equations (2.12) and (2.13), namely

$$\psi_{xx}+(\zeta^2+u)\psi=0, \qquad \psi_t=u_x\psi+2(2\zeta^2-u)\psi.$$

One immediately sees that these relations are also a BT between the KdV equation and

$$\psi_t+\psi_{xxx}-6\zeta^2\psi_x-3\frac{\psi_x\psi_{xx}}{\psi}=0, \tag{4.6}$$

obtained by solving the Schrödinger equation for u and substituting into the associated time–evolution equation; recall that the KdV follows from the compatibility condition $\psi_{xxt}=\psi_{txx}$. Note that u is uniquely determined from ψ (for $\psi\neq 0$), whereas ψ is only determined by u to within two arbitrary constants corresponding to ψ and ψ_x at (x_0,t_0) [AbS981].

REMARK 4.3 – There is a distinction between a BT and a mapping. Once v is given, a mapping uniquely defines u without specifying anything about *either* $\mathcal{D}[u]=0$ or $\mathcal{E}[v]=0$, whilst a BT need not define u uniquely but does specify *both* $\mathcal{D}[u]=0$ and $\mathcal{E}[v]=0$. Also, a BT can be constructed from a mapping by associating an appropriate evolution equation. For example, the Schrödinger equation is a mapping from ψ to u (in a domain where $\psi\neq 0$), but alone does not identifies a BT; another example is the Miura transform (2.2), which is a map of mKdV into KdV, but not a BT; however, the couple

$$v_x=-(u-v^2), \qquad v_t=6v^2v_x-v_{xxx}, \tag{4.7}$$

is indeed a BT between the KdV and the mKdV equations.

We come now to the main point of the section: what do the BT has to do with IST? We have already seen that the KdV scattering problem is also a BT between KdV and equation (4.6). Now let's consider the following problem: once the BT for a certain NLEE is given, can we recover the associated scattering problem? As a test case, consider the BT for the SG: following a procedure proposed originally by M. WADATI, H. SANUKI & K. KONNO [WSK975] and later independently discovered by R. K. DODD & R. K. BULLOUGH [DoB976], set (for notation convenience) $\omega\equiv\omega_0$ and $\widetilde{\omega}\equiv\omega_1$ and introduce the new variable $\gamma:=\tan(\frac{\widetilde{\omega}+\omega}{4})$. Under this transformation, $\mathbb{B}_\beta$ becomes

$$\mathbb{B}_\beta \;:\; \begin{cases} \gamma_x &= \dfrac{1}{2}(1+\gamma^2)\omega_x+\dfrac{\beta}{2\rho}(1+\gamma^2)\sin\left(\dfrac{\widetilde{\omega}+\omega}{2}\right)=\dfrac{1}{2}(1+\gamma^2)\omega_x+\dfrac{\beta\gamma}{\rho}, \\ \gamma_t &= \dfrac{1}{2\beta\rho}(1+\gamma^2)\sin\left(\dfrac{\widetilde{\omega}-\omega}{2}\right)=\dfrac{\gamma}{\beta\rho}\cos\omega-\dfrac{1}{2\beta\rho}(1-\gamma^2)\sin\omega, \end{cases} \tag{4.8}$$

[38] Here the sentence "$v_x=f(x,t)$ and $v_t=g(x,t)$ are *integrable*" must be read as "they're compatible".

where we have used the identities $(1+\gamma^2)\sin\left(\frac{\widetilde{\omega}-\omega}{2}\right) = 2\gamma$, $(1+\gamma^2)\sin\left(\frac{\widetilde{\omega}-\omega}{2}\right) = \gamma\cos\omega - (1-\gamma^2)\sin\omega$, which follows from the duplication formulas of the sine and cosine functions. The system (4.8) is a set of Riccati–type equations which can be linearized by applying the transformation $\gamma = \varphi_2/\varphi_1$, defined in a domain of $\mathbb{R}^2$ where $\varphi_1 \neq 0$ and $\varphi_2 \neq 0$ (the last condition holds since $\gamma = 0$ isn't a solution of system (4.8)). With this change of variable, the system (4.8) gets the form

$$\begin{cases} \varphi_1\left(\varphi_{2x} - \frac{1}{2}\omega_x\varphi_1 - \frac{\beta}{2\rho}\varphi_2\right) = \varphi_2\left(\varphi_{1x} + \frac{1}{2}\omega_x\varphi_2 + \frac{\beta}{2\rho}\varphi_1\right), \\ \varphi_1\left(\varphi_{2t} - \frac{1}{2\beta\rho}\varphi_2\cos\omega + \frac{1}{2\beta\rho}\varphi_1\sin\omega\right) = \varphi_2\left(\varphi_{1t} + \frac{1}{2\beta\rho}\varphi_1\cos\omega + \frac{1}{2\beta\rho}\varphi_2\sin\omega\right). \end{cases}$$

Invoking now the conditions $\varphi_1, \varphi_2 \neq 0$ and setting to zero each coefficient in the above system, gives

$$\begin{pmatrix}\varphi_{1x}\\ \varphi_{2x}\end{pmatrix} = \frac{1}{2}\begin{pmatrix}-\frac{\beta}{\rho} & -\omega_x\\ \omega_x & \frac{\beta}{\rho}\end{pmatrix}\begin{pmatrix}\varphi_1\\ \varphi_2\end{pmatrix}, \qquad \begin{pmatrix}\varphi_{1t}\\ \varphi_{2t}\end{pmatrix} = \frac{1}{2\beta\rho}\begin{pmatrix}-\cos\omega & -\sin\omega\\ -\sin\omega & \cos\omega\end{pmatrix}\begin{pmatrix}\varphi_1\\ \varphi_2\end{pmatrix}. \tag{4.9}$$

which correspond exactly to the *ZS spectral problem* for the SG equation (see relations (3.18), §3.2) with $\rho = 1$ and $\beta = 2\imath\zeta$. Therefore we can conclude that the *linearized aBT* for the SG equation (with curvature $\mathcal{K} = -1$) is equivalent to the ZS spectral problem for the SG; note that the Bäcklund parameter of the BT is related to the spectral parameter of the IST via the relation $\beta = 2\imath\zeta$.

Recalling now the role that the ZS spectral problem had in the contest of the IST method, the results of the following theorem should not came as a great surprise [AbS981].

THEOREM 4.1 - *Consider the ZS spectral problem given by the set of equations (3.6), (3.7) and let* $\mathbf{u} \equiv (u, v)$ *satisfy the evolution equations* $\mathcal{D}[\mathbf{u}] = \mathbf{0}$*, each one compatible with (3.6), (3.7). Then for every* $\zeta \in \mathbb{R}^*$ *the ZS spectral problem is a BT between* $\mathcal{D}[\mathbf{u}] = \mathbf{0}$ *and some* $\mathcal{E}[\boldsymbol{\varphi}; \zeta] = \mathbf{0}$*, where* $\mathcal{E}[\boldsymbol{\varphi}; \zeta] = \mathbf{0}$ *is a pair of PDEs for* $\boldsymbol{\varphi} \equiv (\varphi_1, \varphi_2)$ *involving the spectral parameter* ζ*, but not* $\mathbf{u}$*.*

Proof. See M. J. ABLOWITZ, H. SEGUR – *Solitons and the Inverse ...*, Ch. 3, pgg. 156–157. □

We won't prove the theorem above, since we wish to discuss a general procedure to derive BTs for IST–solvable NLEEs starting from the associated inverse scattering problems. The approach we'll follow was originally proposed by H.–H. CHEN in 1974 [Che974]. Consider the ZS spectral problem

$$\begin{cases} \varphi_{1x} = -\imath\zeta\varphi_1 + u(x,t)\varphi_2, \\ \varphi_{2x} = v(x,t)\varphi_1 + \imath\zeta\varphi_2, \end{cases} \qquad \begin{cases} \varphi_{1t} = \mathcal{A}\varphi_1 + \mathcal{B}\varphi_2, \\ \varphi_{2t} = \mathcal{C}\varphi_1 - \mathcal{A}\varphi_2; \end{cases}$$

recall that the system is integrable (in the sense of AKNS) iff $\mathcal{A}, \mathcal{B}$ and $\mathcal{C}$ satisfy the equations (3.10), obtained by forcing the compatibility condition $\boldsymbol{\varphi}_{xt} = \boldsymbol{\varphi}_{tx}$. We have seen that finite expansions of $\mathcal{A}, \mathcal{B}, \mathcal{C}$ in terms of ζ and appropriate choices of the potentials u, v reduce the ZS spectral problem to NLEEs of interest like the KdV, mKdV, SG, ShG or NLS (see §3.2). Now manipulate the ZS systems in the following way: multiply both members of the first equations for both systems by φ_2/φ_2^2 and both member of the second equations by φ_1/φ_2^2, to get the systems of Riccati equations

$$\begin{cases} \frac{\varphi_{1x}\varphi_2}{\varphi_2^2} = -\imath\zeta\frac{\varphi_1}{\varphi_2} + u(x,t), \\ \frac{\varphi_{2x}\varphi_1}{\varphi_2^2} = v(x,t)\frac{\varphi_1^2}{\varphi_2^2} + \imath\zeta\frac{\varphi_1}{\varphi_2}, \end{cases} \qquad \begin{cases} \frac{\varphi_{1t}\varphi_2}{\varphi_2^2} = \mathcal{A}\frac{\varphi_1}{\varphi_2} + \mathcal{B}, \\ \frac{\varphi_{2t}\varphi_1}{\varphi_2^2} = \mathcal{C}\frac{\varphi_1^2}{\varphi_2^2} - \mathcal{A}\frac{\varphi_1}{\varphi_2}, \end{cases}$$

respectively. Subtracting now the second equations from the first ones, we get the **Chen system**

$$\begin{cases} \gamma_x = -2\imath\zeta\gamma + u - v\gamma^2, \\ \gamma_t = 2\mathcal{A}\gamma + \mathcal{B} - \mathcal{C}\gamma^2, \end{cases} \tag{4.10}$$

having defined the linearizing variable $\gamma := \varphi_1/\varphi_2$ in a proper domain of $\mathbb{R}^2$ where $\varphi_1, \varphi_2 \neq 0$.

Therefore we have to specify the class to which the NLEE of interest belongs. Let's start with the one related to the KdV, for which $v=-1$ (see §3.2). Setting $\beta \equiv \imath\zeta$, the Chen system becomes

$$\mathbb{I} \;:\; \begin{cases} \gamma_x &= -2\beta\gamma + u + \gamma^2, \\ \gamma_t &= 2\mathcal{A}\gamma + \mathcal{B} - \mathcal{C}\gamma^2, \end{cases}$$

where the expressions for $\mathcal{A}, \mathcal{B}$ and $\mathcal{C}$ follow from equations (3.16):

$$\mathcal{A}_{\mathrm{KdV}} = 4\beta^3 + 2\beta u - u_x, \quad \mathcal{B}_{\mathrm{KdV}} = -4\beta^2 u + 2\beta u_x - 2u^2 - u_{xx}, \quad \mathcal{C}_{\mathrm{KdV}} = 4\beta^2 + 2u.$$

Now, solving the first equation of the Chen system for u gives $u = \gamma_x + 2\beta\gamma - \gamma^2$; inserting this into the previous expressions for $\mathcal{A}_{\mathrm{KdV}}, \mathcal{B}_{\mathrm{KdV}}$ and $\mathcal{C}_{\mathrm{KdV}}$ yields

$$\begin{aligned} \mathcal{A}_{\mathrm{KdV}} &= 4\beta^3 + 4\beta^2\gamma - 2\beta\gamma^2 - \gamma_{xx} + 2\gamma\gamma_x, \\ \mathcal{B}_{\mathrm{KdV}} &= -8\beta^2\gamma - 4\beta^2\gamma^2 - 8\beta\gamma\gamma_x + 4\gamma^2\gamma_x + 8\beta\gamma^3 - 2\gamma^4 - \gamma_{xxx} + 2\gamma\gamma_{xx}, \\ \mathcal{C}_{\mathrm{KdV}} &= 4\beta^2 + 2\gamma_x + 4\beta\gamma - 2\gamma^2. \end{aligned}$$

Finally, substituting these expressions into the second equation of the Chen system (4.10) we get

$$\gamma_t - 6\gamma^2\gamma_x + 12\beta\gamma\gamma_x + \gamma_{xxx} = 0, \tag{4.11}$$

which represents a mixed KdV–mKdV equation and is indeed equivalent (up to a change of variables) to equation (4.6). In particular, if $\beta = 0$ then equation (4.11) reduces to the pure mKdV and the Chen system (4.10) reproduces the BT between KdV and mKdV given in (4.7).

Now it comes the main aspect of the Chen's approach. Note that equation (4.11) is invariant under the *parity–transformations* $(\gamma, \beta) \mapsto (-\gamma, -\beta)$, which means that if u is a solution of the KdV, then another KdV–solution $\widetilde{u}$ exists, s.t. $-\gamma_x = -2\beta\gamma + \widetilde{u} + \gamma^2$, $-\gamma_t = -2\widetilde{\mathcal{A}}\gamma + \widetilde{\mathcal{B}} - \widetilde{\mathcal{C}}\gamma^2$, where $\widetilde{\mathcal{A}} \equiv \mathcal{A}(u', -\beta)$, $\widetilde{\mathcal{B}} \equiv \mathcal{B}(u', -\beta)$ and $\widetilde{\mathcal{C}} \equiv \mathcal{C}(u', -\beta)$ [Che974]. Therefore, one is leaved with the two solutions $(u, \widetilde{u})$ of the KdV equation belonging to the two Chen systems

$$u \;:\; \begin{cases} \gamma_x &= -2\beta\gamma + u + \gamma^2, \\ \gamma_t &= 2\mathcal{A}\gamma + \mathcal{B} - \mathcal{C}\gamma^2, \end{cases} \qquad \widetilde{u} \;:\; \begin{cases} \gamma_x &= 2\beta\gamma - \widetilde{u} - \gamma^2, \\ \gamma_t &= 2\widetilde{\mathcal{A}}\gamma - \widetilde{\mathcal{B}} + \widetilde{\mathcal{C}}\gamma^2. \end{cases} \tag{4.12}$$

Adding and subtracting the equations belonging to the same derivatives, yields finally

$$\begin{cases} u + \widetilde{u} &= 2(2\beta - \gamma)\gamma, \\ u - \widetilde{u} &= 2\gamma_x, \end{cases} \qquad \begin{cases} \gamma^2(\mathcal{C} + \widetilde{\mathcal{C}}) - 2\gamma(\mathcal{A} - \widetilde{\mathcal{A}}) - (\mathcal{B} - \widetilde{\mathcal{B}}) = 0, \\ 2\gamma_t + \gamma^2(\mathcal{C} - \widetilde{\mathcal{C}}) - 2\gamma(\mathcal{A} + \widetilde{\mathcal{A}}) - (\mathcal{B} - \widetilde{\mathcal{B}}) = 0. \end{cases}$$

To complete the task, let's make the substitution $u = \omega_x$: therefore one has $\gamma = \frac{1}{2}(\omega - \widetilde{\omega}) + \beta$ (having fixed the constant of integration with β for historical reasons, clarified below) and then

$$\frac{u + \widetilde{u}}{2} = 2\beta\gamma - \gamma^2 \qquad \xrightarrow{\gamma = \frac{1}{2}(\omega - \widetilde{\omega}) + \beta} \qquad \left(\frac{\omega + \widetilde{\omega}}{2}\right)_x = \beta^2 - \left(\frac{\omega - \widetilde{\omega}}{2}\right)^2, \tag{4.13}$$

which is the x–component of the aBT of the KdV, originally found by H. D. Wahlquist & F. B. Estabrook in 1973 [WaE973]. For the t–component one can insert γ in (4.11) to get [WSK975]

$$\left(\frac{\omega - \widetilde{\omega}}{2}\right)_t = \left(\omega_x + \widetilde{\omega}_x\right)^2 - 2\left(\frac{\omega - \widetilde{\omega}}{2}\right)\left(\frac{\omega - \widetilde{\omega}}{2}\right)_{xx}, \tag{4.14}$$

or equivalently into the second equation for the Chen system[39] [Che974, AbS981]. The set of equations (4.13), (4.14) is an aBT for the KdV equation with Bäcklund parameter $\beta = \imath\zeta$.

[39] Note that other forms for the KdV–aBT might be obtained by following different ways during the calculations [Che974]. We want to stress here that this is not wrong, in fact *the BT for a given NLEE is not necessary unique.*

Consider now a different class, namely the one associated to the SG and the mKdV equations; in §3.2 we've seen that these NLEEs are recovered from the ZS spectral problem by choosing $v = -u$, having supposed u real. Setting $\beta \equiv 2\imath\zeta$, the Chen system (4.10) becomes

$$\mathbb{III} \ : \ \begin{cases} \gamma_x \ = -\beta\gamma + u(1+\gamma^2), \\ \gamma_t \ = 2\mathcal{A}\gamma + \mathcal{B} - \mathcal{C}\gamma^2. \end{cases}$$

In order to work with the system above, introduce the convenient re–parametrization $\gamma = \tan\frac{w}{2}$, being $w = w(x,t)$ an $\mathrm{S}(\mathbb{R}^2)$ function. The following relations follow immediately:

$$\frac{\gamma_{x,t}}{1+\gamma^2} = \partial_{x,t}\arctan(\gamma) = \frac{1}{2}w_{x,t}, \quad \frac{\gamma}{1+\gamma^2} = \frac{1}{2}\sin w, \quad \frac{1}{1+\gamma^2} = \cos^2\left(\tfrac{1}{2}w\right), \quad \frac{\gamma^2}{1+\gamma 2} = \sin^2\left(\tfrac{1}{2}w\right).$$

Therefore, dividing both equations of the Chen system $\mathbb{III}$ by $1+\gamma^2$, one finds

$$\mathbb{III} \ : \ \begin{cases} u \ = \frac{1}{2}\left(w_x + \beta\sin w\right), \\ \frac{1}{2}w_t \ = \mathcal{A}\sin w + \mathcal{B}\cos^2\left(\frac{1}{2}w\right) - \mathcal{C}\sin^2\left(\frac{1}{2}w\right). \end{cases} \tag{4.15}$$

Let's specialize to particular cases. We start from the mKdV: recalling formulas (3.17), (3.18), one has

$$\begin{aligned} \mathcal{A}_{\mathrm{mKdV}} &= \tfrac{1}{2}\beta^3 + \tfrac{1}{4}\beta^3\sin^2 w + \tfrac{1}{2}\beta^2 w_x \sin w + \tfrac{1}{2}\beta w_x^2, \\ \mathcal{B}_{\mathrm{mKdV}} &= -\tfrac{1}{2}\beta^3\sin w\left(1+\tfrac{1}{2}\sin^2 w\right) - \tfrac{1}{2}\beta^2 w_x\left(\tfrac{3}{2}\sin^2 w - \cos w + 1\right) + \\ &\quad - \tfrac{1}{4}\beta w_x^2\sin w + \tfrac{1}{2}\beta w_x\left(1-\cos w\right) - \tfrac{1}{4}w_x^3 - \tfrac{1}{2}w_{xxx}, \\ \mathcal{C}_{\mathrm{mKdV}} &= \tfrac{1}{2}\beta^3\sin w\left(1+\tfrac{1}{2}\sin^2 w\right) + \tfrac{1}{2}\beta^2 w_x\left(\tfrac{3}{2}\sin^2 w + \cos w + 1\right) + \\ &\quad + \tfrac{1}{4}\beta w_x^2\sin w + \tfrac{1}{2}\beta w_x\left(1+\cos w\right) + \tfrac{1}{4}w_x^3 + \tfrac{1}{2}w_{xxx}; \end{aligned}$$

substitution of the previous expressions in the second equation for the Chen system $\mathbb{III}$ (4.15) gives

$$w_t + \tfrac{1}{2}w_x^3 + w_{xxx} + \tfrac{3}{2}\beta^2 w_x\sin^2 w = 0. \tag{4.16}$$

Equation (4.16) is invariant under the transform $(\gamma,\beta) \mapsto (\pm\gamma, \pm\beta)$. Thus considering, e.g. the invariance $(\gamma,\beta) \mapsto (-\gamma,-\beta)$, one has two mKdV–solutions $u, \widetilde{u}$, belonging to two Chen systems of the form (4.15): adding and subtracting relative expressions yields $u + \widetilde{u} = \beta\sin w$ and $u - \widetilde{u} = w_x$, respectively. Setting $u = \frac{\omega_x}{2}$ we obtain $w = \frac{1}{2}(\omega - \widetilde{\omega})$, having fixed (for simplicity) to zero che integration constant. Substituting the previous relations in $u + \widetilde{u} = \beta\sin w$ gives immediately the x–component of the MkDv–aBT, that is $\frac{1}{2}(\omega-\widetilde{\omega})_x = \beta\sin\left[\frac{1}{2}(\omega-\widetilde{\omega})\right]$. The t–component is obtained by inserting $w = \frac{1}{2}(\omega-\widetilde{\omega})$ in equation (4.16). Ultimately, one finds the following expressions for the **mKdV–aBT** [Che974]

$$\begin{aligned} \left(\frac{\omega \pm \widetilde{\omega}}{2}\right)_x &= \beta\sin\left(\frac{\omega \mp \widetilde{\omega}}{2}\right), \\ \left(\frac{\omega \pm \widetilde{\omega}}{2}\right)_t &= -2\beta\omega_x^2\sin\left(\frac{\omega \mp \widetilde{\omega}}{2}\right) \mp 2\beta\,\omega_{xx}\cos\left(\frac{\omega-\widetilde{\omega}}{2}\right) \mp \beta^3\sin\left(\frac{\omega \pm \widetilde{\omega}}{2}\right) \pm 2\beta^2\omega_x. \end{aligned} \tag{4.17}$$

Consider now the SG equation: in this case we have $\mathcal{A}_{\mathrm{SG}} = -\frac{1}{2\beta}\cos\omega$, $\mathcal{B}_{\mathrm{SG}} = \mathcal{C}_{\mathrm{SG}} = \beta^{-1}u_t$. It turns to be convenient to set $u = -\frac{\omega_x}{2}$ at this stage s.t. $\mathcal{A}_{\mathrm{SG}} = -\frac{1}{2\beta}(1-4u_t^2)^{1/2}$, being $\cos\omega = (1-\sin^2\omega)^{1/2}$ $= (1-\omega_{xt}^2)^{1/2}$, where we have assumed that ω satisfies the SG equation); note that the expression for u is given by the first equation in (4.15), i.e. $u_t = \frac{1}{2}(w_{xt} + \beta w_t\cos w)$. Now express functions $\mathcal{A}_{\mathrm{SG}}$ and $\mathcal{B}_{\mathrm{SG}}$ in terms of $w = 2\arctan(\gamma)$ and insert the results in the second equation of the Chen's system (4.15): using well–known trigonometric identities, we finally obtain

$$w_{xt} = \sin w\sqrt{1-\beta^2 w_t^2}. \tag{4.18}$$

Equation (4.18) is $(\gamma,\beta)\mapsto(\pm\gamma,\pm\beta)$–invariant; thus, as for mKdV, there exist two SG–solutions $(u,\widetilde{u})$, with $\widetilde{u}$ associated (say) to the $(\gamma,\beta)\mapsto(-\gamma,-\beta)$ invariance, s.t. $u+\widetilde{u}=\beta\sin w$ and $u-\widetilde{u}=w_x$. Having chosen $u\equiv-\frac{\omega_x}{2}$, one has $w=\frac{1}{2}(\widetilde{\omega}-\omega)$ and then $\frac{1}{2}(\omega+\widetilde{\omega})_x=\beta\sin\left[\frac{1}{2}(\omega-\widetilde{\omega})\right]$. Insert now $w=\frac{1}{2}(\widetilde{\omega}-\omega)$ in the second equation of (4.15): since $\mathcal{B}_{\rm SG}=\mathcal{C}_{\rm SG}$ one has

$$\tfrac{1}{2}w_t=\mathcal{A}_{\rm SG}\sin w+\mathcal{B}_{\rm SG}\cos w,$$

where we have used the identity $\cos^2(w/2)-\sin^2(w/2)=\cos w$. Invoking now the original expressions $\mathcal{A}_{\rm SG}=-\frac{1}{2\beta}\cos\omega$ and $\mathcal{B}_{\rm SG}=-\frac{1}{2\beta}\sin\omega$, one finds $\frac{1}{2}(\omega-\widetilde{\omega})_t=\frac{1}{\beta}\sin(w+\omega)=\frac{1}{\beta}\sin\left[\frac{1}{2}(\widetilde{\omega}+\omega)\right]$. We have therefore recovered the original expressions for the SG–aBT, namely

$$\left(\frac{\omega\pm\widetilde{\omega}}{2}\right)_x=\beta\sin\left(\frac{\omega\mp\widetilde{\omega}}{2}\right),\qquad\left(\frac{\omega\mp\widetilde{\omega}}{2}\right)_t=\frac{1}{\beta}\sin\left(\frac{\omega\pm\widetilde{\omega}}{2}\right),\tag{4.19}$$

where the equations with different signs follow from the $(\gamma,\beta)\mapsto(\pm\gamma,\mp\beta)$ invariances [Che974].

Proceeding in the same fashion as before, one can treats other classes of NLEEs and deduce the associated aBTs: e.g. the ShG–aBT might be obtained by studying the class $v=u$ and the NLS–aBT (or, more generally, aBTs associated to NLS–type equations) from the class $v=-u^*$.

4.2.1 The connection with multi–solitons solutions. The auto–Bäcklund transforms for the KdV, mKdV and SG equations can be deduced by following a different approach, suggested by K. Konno & M. Wadati [KoW975]. The main idea is to construct a transform $(\gamma,u)\mapsto(\widetilde{\gamma},\widetilde{u})$ that leaves invariant the system (4.10). For the class $\mathbb{I}$ (i.e. $v=-1$) the couple $(\widetilde{\gamma},\widetilde{u})$ can be chosen s.t.

$$\gamma'\equiv-\gamma+2\beta,\qquad\widetilde{u}\equiv u+2\widetilde{\gamma}_x,\tag{4.20}$$

where $\beta\equiv-\imath\zeta$. It is easy to verify that $(\widetilde{\gamma},\widetilde{u})$ satisfy the Chen system $\mathbb{I}$. If we chose $u\equiv\omega_x$, $\widetilde{u}\equiv\widetilde{\omega}_x$ in the second transformation $\widetilde{\gamma}_x=\frac{1}{2}(\widetilde{u}-u)$, we get the relation $\widetilde{\gamma}\equiv\frac{1}{2}(\widetilde{\omega}-\omega)-\beta$, where the integration constant has been chosen s.t. $\gamma=\frac{1}{2}(\widetilde{\omega}-\omega)+\beta$. Inserting this relation into $\mathbb{I}$, one finds

$$\begin{cases}\gamma_x=-2\beta\gamma+u+\gamma^2,\\ \gamma_t=2\mathcal{A}\gamma+\mathcal{B}-\mathcal{C}\gamma^2,\end{cases}\xrightarrow{\gamma=\frac{1}{2}(\widetilde{\omega}-\omega)+\beta}\begin{cases}\left(\omega+\omega'\right)_x=2\beta^2-\frac{1}{2}\left(\omega-\omega'\right)^2,\\ \left(\omega-\omega'\right)_t=2\mathcal{A}\left[(\omega-\omega')+2\beta\right]+2\mathcal{B}-\mathcal{C}\left[(\omega-\omega')+2\beta\right]^2,\end{cases}$$

which are to the KdV–aBT. For the class $\mathbb{III}$ (i.e. $v=-u$), one has the transforms [KoW975]

$$\widetilde{\gamma}=\frac{1}{\gamma},\qquad\widetilde{u}=u-2\partial_x\arctan(\gamma).\tag{4.21}$$

Now distinguish between the mKdV and the SG equations: choose $(\widetilde{u},u)\equiv(\frac{\widetilde{\omega}}{2},\frac{\omega}{2})$ for the former and $(\widetilde{u},u)\equiv(-\frac{\widetilde{\omega}}{2},-\frac{\omega}{2})$ for the latter. From the second transformation in (4.21), one therefore obtains

$$\text{mKdV}\ :\ \gamma=\tan\left(\frac{\omega-\widetilde{\omega}}{4}\right),\qquad\text{SG}\ :\ \gamma=\tan\left(\frac{\widetilde{\omega}-\omega}{4}\right),$$

respectively. Inserting these expressions in $\mathbb{III}$, yields ultimately the aBTs for the mKdV and the SG.

Indeed it might be thought that the transformations (4.20), (4.21) have been chosen on purpose, just to recover the correct expressions for the aBTs. Actually, those transforms emerge from a deep argumentation, developed independently by three groups, namely M. Wadati, H. Sanuki & K Konno [WSK975], P. Deift & E. Trubowitz [DeT979] and F. Calogero [Cal978], which is at the heart of the connection between the aBTs of NLEEs and *multi–solitons solutions*. The common aspect here is that the main effect of a BT on a given solution of a NLEE is to **add** or (regarding the inverse BT) **subtract one soliton**. We have already encountered a similar property in the beginning of §4.1, when we have derived, starting with the seed solution $\omega_0\equiv0$, the one–soliton (4.2) and two–solitons (4.5)

solutions of the SG equation from the associated aBT. However, a general theorem can be proved stating that if the original solution $u_0 = u_0(x,t)$ of a given NLEE satisfies the condition

$$\int_{\mathbb{R}} |u_0(x,t)|\big(1+|x|\big)\,\mathrm{d}\,x < \infty, \tag{4.22}$$

then the new solution $u_1 \equiv \mathbb{B}_\beta[u_0]$ satisfies condition (4.22) and its spectrum (thinking of u_1 as a potential in the associated spectral problem) differs from that of u_0 by exactly one discrete eigenvalue [AbS981]. We're not going to prove this result in general; however, following the work of M. Wadati *et al.* [WSK975], we are going to show how this happens for the KdV equation. We will be then in the position to justify the relations chosen at the beginning for the transforms (4.20), (4.21).

Let's therefore start from the KdV scattering problem, that is the Schrödinger equation[40] $\psi_{xx} + [\zeta^2 + u(x)]\psi = 0$; let $\psi_0 = \psi_0(x)$ an eigenfunction with eigenvalue ζ_0 and $\psi(x,\zeta)$ an arbitrary solution. Consider now the so–called **Crum Transform** (CT) here defined:

$$\psi(x,\zeta) \xrightarrow{\mathrm{C}_{\zeta_0}} \widetilde{\psi}(x,\zeta) \equiv \frac{\mathfrak{W}\big\{\psi(x,\zeta);\psi_0(x)\big\}}{(\zeta^2-\zeta_0^2)\psi_0(x)}, \tag{4.23}$$

where $\mathfrak{W}\{\psi,\psi_0\} := \psi_x\psi_0 - \psi\psi_{0x}$ is the *Wronskian* between functions ψ, ψ_0. A transformation similar to the one in (4.23) was introduced in 1955 by M. M. Crum to change a Sturm–Liouville operator defined over a finite interval into an operator having one less eigenvalue than the original one [Cru955]. Recall now two useful properties of the Schrödinger equation, summarized in the following

Theorem 4.2 - *The function* (4.23) *is a solution of the Schrödinger equation with potential*

$$\widetilde{u}(x) \equiv u(x) + \mathcal{F}[u(x)], \qquad \mathcal{F}[u(x)] \equiv 2\big[\ln\big(\psi_0(x)\big)\big]_{xx}, \tag{4.24}$$

$\forall\,\zeta \neq \zeta_0$. *Also the transformation inverse to* (4.23) *is given by*

$$\psi(x,\zeta) = \frac{\mathfrak{W}\big\{\widetilde{\psi}(x,\zeta);\widetilde{\psi_{01}}(x)\big\}}{\widetilde{\psi}_{01}(x)}, \qquad \widetilde{\psi}_{01}(x) \equiv \psi_0^{-1}(x). \tag{4.25}$$

Proof. See M. Wadati *et al.*, *Inverse Method, Bäcklund Transform and ...*, pg. 426 [WSK975]. □

It turns out that the CT is closely related to the KdV–aBT. To show this, let $u_{(N)}$ and $u_{(N+1)}$ be the N and $(N+1)$–solitons solution, respectively. We know from §2.2 that KdV–soliton solutions are associated to *bound states* of the Schrödinger equation. Consider therefore the system

$$\begin{cases} \big\{\psi_i \in L_2(\mathbb{R}) : \partial_x^2\psi_i + [\zeta_i^2 + u_{(N)}]\psi_i = 0,\ \psi_i(x) \xrightarrow{|x|\to+\infty} 0\big\}_{i=1,2,\dots,N}, \\ \big\{\phi_i \in L_2(\mathbb{R}) : \partial_x^2\phi_i + [\zeta_i^2 + u_{(N+1)}]\phi_i = 0,\ \phi_i(x) \xrightarrow{|x|\to+\infty} 0\big\}_{i=1,2,\dots,N,N+1}. \end{cases} \tag{4.26}$$

Without loss of generality, reorder the eigenvalues $\{\zeta_0,\zeta_1,\dots,\zeta_N,\zeta_{N+1}\}$ s.t. $\zeta_{N+1}^2 > \zeta_N^2 > \cdots > \zeta_1^2 > \zeta_0^2 > 0$. Now add to the system (4.26) a new eigenfunction ψ_{N+1} satisfying the equation

$$\partial_{xx}\psi_{N+1} + \big[\zeta_{N+1}^2 + u_{(N)}\big]\psi_{N+1} = 0; \tag{4.27}$$

note that this is not a bounded solution. As shown in Th. 4.2, the CT exists between two Schrödinger equations with same eigenvalues and potentials given by relation (4.24). Thus, identifying $\zeta_0^2 \equiv \zeta_{N+1}^2$, $\widetilde{u} \equiv u_{(N)}$, $u \equiv u_{(N+1)}$ and $\psi_0(x) \equiv \phi_{N+1}(x) = 1/\psi_{N+1}(x)$, we find from the relations (4.24), (4.25)

$$u_{(N+1)}(x) = u_{(N)}(x) + 2\big[\ln\psi_{N+1}(x)\big]_{xx}. \tag{4.28}$$

Equations (4.27) and (4.28) are the same that Wahlquist and Estabrook had derived from the KdV–BT [WaE973] and gives exactly the transforms (4.20) and (4.21) for the classes $\mathbb{I}$ and $\mathbb{III}$ [WSK975].

[40] Note that the time dependence is irrelevant in this argument and therefore it is suppressed.

5 The gauge–Bäcklund Transform (gBT)

In this section we are going to discuss an alternative procedure to obtain BTs for NLEEs, starting from the associate spectral problem, by means of the so–called *gauge transforms*. We will also show that within this approach, once the BTs for the given NLEEs are have been found, it's possible to straightforwardly prove the corresponding permutability theorem, just as the Bianchi's one for the SG equation. The presentation given here is based on the works of M. Boiti & G. Z. Tu [BoT982] and of M. Boiti, F. Pempinelli & G. Z. Tu [BPT983]; see also the other references given therein.

To introduce the argument, consider the following generalization for the ZS–spectral problem

$$\begin{cases} \boldsymbol{\varphi}_x = \mathcal{X}\boldsymbol{\varphi}, \\ \boldsymbol{\varphi}_t = \mathcal{T}\boldsymbol{\varphi}, \end{cases} \qquad \mathcal{X} \equiv \mathcal{X}\big[\mathcal{Q}(x,t);\zeta\big], \quad \mathcal{T} \equiv \mathcal{T}\big[\mathcal{Q}(x,t);\zeta\big], \tag{5.1}$$

where $\boldsymbol{\varphi}, \mathcal{X}$ and $\mathcal{T}$ are complex–valued $N \times N$ matrix functions of x, t, ζ and $\mathcal{X}$, $\mathcal{T}$ are rational in the spectral parameter ζ; assume also that the components of the matrix $\mathcal{Q} = \mathcal{Q}(x,t)$ are $\mathrm{S}(\mathbb{R})$ functions. Note that if $\mathcal{X} = \mathcal{Q} - \imath\zeta\sigma_3$ (see equation (3.9)) then the ZS–spectral problem is recovered. Also recall that the compatibility condition $\boldsymbol{\varphi}_{xt} = \boldsymbol{\varphi}_{tx}$ furnishes the so–called Lax representation in the AKNS formalism for the NLEE on is looking for (as shown in §3.2 for the ZS–system).

Let's now follow a different approach. Let $\mathcal{G} = \mathcal{G}[\mathcal{Q}(x,t);\zeta]$ be the matrix representation (over the space where the components of $\boldsymbol{\varphi}$ are defined, usually $L_2(\mathbb{R})$) of an *infinitesimal gauge transformation* $\mathfrak{g}$ of $\boldsymbol{\varphi}$ such that $\boldsymbol{\varphi} \mapsto \boldsymbol{\varphi}' := \mathcal{G}\boldsymbol{\varphi}$. This formulation must be consistent with the AKNS' one, thus $\mathcal{G}$ must be a complex–valued $N \times N$ matrix function of x, t, ζ rational in ζ, defined as

$$\mathcal{G}\big[\mathcal{Q}(x,t);\zeta\big] = \mathbb{1}_N + \mathcal{T}\big[\mathcal{Q}(x,t);\zeta\big]\,\mathrm{d}\,t. \tag{5.2}$$

Note that equation (5.2) and the time–evolution equation in (5.1) are equivalent, since $\boldsymbol{\varphi} + \boldsymbol{\varphi}_t\,\mathrm{d}\,t = \boldsymbol{\varphi}' = \big[\mathbb{1}_N + \mathcal{T}\,\mathrm{d}\,t\big]\boldsymbol{\varphi}$ which implies $\boldsymbol{\varphi}_t = \mathcal{T}\boldsymbol{\varphi}$. Under the action of the gauge transform, the principal spectral problem in (5.1) transforms simultaneously as $\boldsymbol{\varphi}'_x = \mathcal{X}'\boldsymbol{\varphi}'$, where $\mathcal{X}'$ satisfies the equation

$$\mathcal{G}_x = \mathcal{X}'\mathcal{G} - \mathcal{G}\mathcal{X}, \tag{5.3}$$

which follows by noting that the identity $(\mathcal{G}_x + \mathcal{G}\mathcal{X})\boldsymbol{\varphi} = (\mathcal{G}\boldsymbol{\varphi})_x = \boldsymbol{\varphi}'_x = \mathcal{X}'\boldsymbol{\varphi}' = (\mathcal{X}'\mathcal{G})\boldsymbol{\varphi}$ holds $\forall\,\boldsymbol{\varphi} \neq \mathbf{0}$. Now assume that $\mathcal{X}'$ is the infinitesimal time translation of $\mathcal{X}$ under the action of the gauge, i.e.

$$\mathcal{X}' = \mathcal{X}\big[\mathcal{Q}'(x,t),\zeta\big] \equiv \mathcal{X} + \mathcal{X}_t\,\mathrm{d}\,t, \qquad \mathcal{Q}'(x,t) \equiv \mathcal{Q}(x,t+\mathrm{d}\,t). \tag{5.4}$$

Inserting the relation (5.4) in equation (5.3), by means of (5.2) one has $\mathcal{T}_x\,\mathrm{d}\,t = \big(\mathcal{X}+\mathcal{X}_t\,\mathrm{d}\,t\big)\big(\mathbb{1}_N+\mathcal{T}\,\mathrm{d}\,t\big) - \big(\mathbb{1}_N + \mathcal{T}\,\mathrm{d}\,t\big)\mathcal{X} = \big(\mathcal{X}_t + [\mathcal{X},\mathcal{T}]\big)\,\mathrm{d}\,t + \mathcal{O}(\mathrm{d}\,t^2)$; neglecting higher order infinitesimals, we find

$$\mathcal{X}_t - \mathcal{T}_x + [\mathcal{X},\mathcal{T}] = \mathbb{O}.$$

Thus we have recovered equation (3.10) by means of an infinitesimal gauge–transform of $\boldsymbol{\varphi}$. Note that, since the Lax equation (3.2) involves isospectral flows (see §3.1), the equations in the system (3.8) are often called in the literature as *isospectral deformation equations*[41] for the matrix potential $\mathcal{Q}$.

5.1 Gauge–Bäcklund transforms and the permutability theorem. Having in mind the foregoing arguments, consider another infinitesimal gauge $\mathfrak{b}$ of $\boldsymbol{\varphi}$ s.t. $\boldsymbol{\varphi} \mapsto \widetilde{\boldsymbol{\varphi}} := \mathcal{B}\boldsymbol{\varphi}$, being $\mathcal{B}$ the $N \times N$ matrix representative of $\mathfrak{b}$ over the space where $\boldsymbol{\varphi}$ is defined. Assume then $\mathfrak{b}$ to be a more general transform than $\mathfrak{g}$, in the sense that $\mathcal{B} = \mathcal{B}[\mathcal{Q}(x,t), \widetilde{\mathcal{Q}}(x,t);\zeta]$, where the matrix potentials $\mathcal{Q}$ and $\widetilde{\mathcal{Q}}$ are solutions of (possibly different) isospectral deformation equations. Under the action of $\mathfrak{b}$ the principal spectral problem transforms as $\widetilde{\boldsymbol{\varphi}}_x = \widetilde{\mathcal{X}}\widetilde{\boldsymbol{\varphi}}$, where $\widetilde{\mathcal{X}} = \mathcal{X}\big(\widetilde{\mathcal{Q}}(x,t);\zeta\big)$ satisfies the equation

$$\mathcal{B}_x = \widetilde{\mathcal{X}}\mathcal{B} - \mathcal{B}\mathcal{X}, \tag{5.5}$$

[41] Therefore, we'll say that in the AKNS representation for the Lax formalism, the NLEEs can be obtain as isospectral deformation equations of the linear spectral problem $\boldsymbol{\varphi}_x = \mathcal{X}\boldsymbol{\varphi}$ [BoT982].

formally equivalent to equation (5.3). Since we wish to deduce the BTs from the infinitesimal gauge transforms, let's consider the double transform $\widetilde{\boldsymbol{\varphi}}' = \mathcal{G}_{\mathcal{B}}\widetilde{\boldsymbol{\varphi}}$, obtained by applying first $\mathfrak{b}$ and then $\mathfrak{g}$ on $\boldsymbol{\varphi}$ (note that the index in $\mathcal{G}_{\mathcal{B}}$ reminds the order of the transforms). The global effect of $\mathfrak{g}$ is an infinitesimal time–translation of $\widetilde{\mathcal{Q}} = \widetilde{\mathcal{Q}}(x,t)$; therefore, in accordance with the condition (5.2), we must have

$$\mathcal{G}_{\mathcal{B}} = \mathbb{1}_N + \mathcal{T}_{\mathcal{B}}\,\mathrm{d}\,t, \tag{5.6}$$

where $\mathcal{T}_{\mathcal{B}}$ is a complex–valued $N \times N$ matrix representative, related to the $\mathfrak{b}$–transformed auxiliary spectral problem, that is $\widetilde{\boldsymbol{\varphi}}_t = \mathcal{T}_{\mathcal{B}}\widetilde{\boldsymbol{\varphi}}$. Therefore $\widetilde{\boldsymbol{\varphi}}' = \widetilde{\boldsymbol{\varphi}} + \widetilde{\boldsymbol{\varphi}}_t\,\mathrm{d}\,t$ and so $(\mathcal{B} + \mathcal{T}_{\mathcal{B}}\mathcal{B}\,\mathrm{d}\,t)\boldsymbol{\varphi} = \mathcal{G}_{\mathcal{B}}\mathcal{B}\boldsymbol{\varphi} = \widetilde{\boldsymbol{\varphi}}' = [\mathcal{B} + (\mathcal{B}_t + \mathcal{B}\mathcal{T})\,\mathrm{d}\,t]\boldsymbol{\varphi}$, which holds $\forall\boldsymbol{\varphi} \neq \mathbf{0}$; then, we find the equation

$$\mathcal{B}_t = \mathcal{T}_{\mathcal{B}}\mathcal{B} - \mathcal{B}\mathcal{T}. \tag{5.7}$$

Remark 5.1 – The infinitesimal gauges transforms $\mathfrak{b}$ and $\mathfrak{g}$ *commutes*, in the sense that

$$\big(\mathcal{G}_{\mathcal{B}}\mathcal{B} - \mathcal{B}'\mathcal{G}\big)\boldsymbol{\varphi} = \mathbb{O}, \qquad \forall\,\boldsymbol{\varphi} \neq \mathbf{0}. \tag{5.8}$$

In fact, having in mind that $\mathcal{B}' = \mathcal{B} + \mathcal{B}_t\,\mathrm{d}\,t$ and recalling equations (5.2), (5.6), one finds

$$\mathcal{G}_{\mathcal{B}}\mathcal{B} - \mathcal{B}'\mathcal{G} = \big(\mathbb{1}_N + \mathcal{T}_{\mathcal{B}}\,\mathrm{d}\,t\big)\mathcal{B} - \big(\mathcal{B} + \mathcal{B}_t\,\mathrm{d}\,t\big)\big(\mathbb{1}_N + \mathcal{T}\,\mathrm{d}\,t\big) = \big(\mathcal{T}_{\mathcal{B}}\mathcal{B} - \mathcal{B}\mathcal{T} - \mathcal{B}_t\big)\,\mathrm{d}\,t + \mathcal{O}(\mathrm{d}\,t^2);$$

invoking equation (5.7) and neglecting higher order infinitesimals, we get the desired result.

We are therefore in the position to introduce the following

Definition 5.1 - *If the matrix $\mathcal{B} = \mathcal{B}[\mathcal{Q}, \widetilde{\mathcal{Q}}; \zeta]$ satisfies the system of equations (5.6), (5.7), i.e.*

$$\begin{cases} \mathcal{B}_x &= \widetilde{\mathcal{X}}\mathcal{B} - \mathcal{B}\mathcal{X}, \\ \mathcal{B}_t &= \mathcal{T}_{\mathcal{B}}\mathcal{B} - \mathcal{B}\mathcal{T}, \end{cases} \tag{5.9}$$

then the transform $\boldsymbol{\varphi} \mapsto \widetilde{\boldsymbol{\varphi}}' := \mathcal{G}_{\mathcal{B}}\mathcal{B}\boldsymbol{\varphi}$ is called a ***gauge–Bäcklund transform*** (**gBT**).

Now we desire to determine the equation satisfied by the matrix potential $\mathcal{Q} = \mathcal{Q}(x,t)$. In light of the $\mathfrak{b}$–transformed principal spectral problem, differentiate with respect to x the gBT to get $\widetilde{\boldsymbol{\varphi}}'_x = (\mathcal{G}_{\mathcal{B}x} + \mathcal{G}_{\mathcal{B}}\widetilde{\mathcal{X}})\widetilde{\boldsymbol{\varphi}}$. Then, note that $\mathfrak{g}$ and $\mathfrak{b}$ are both gauge transforms, thus $\mathfrak{g}\circ\mathfrak{b}$ is also a gauge transform and under its action the principal spectral problem stays formally unchanged, i.e. $\widetilde{\boldsymbol{\varphi}}'_x = \widetilde{\mathcal{X}}'\widetilde{\boldsymbol{\varphi}}'$. Substituting this in the previous expression for $\widetilde{\boldsymbol{\varphi}}'_x$, we find that $\widetilde{\mathcal{X}}'$ satisfies the relation

$$\mathcal{G}_{\mathcal{B}x} = \widetilde{\mathcal{X}}'\mathcal{G}_{\mathcal{B}} - \mathcal{G}_{\mathcal{B}}\widetilde{\mathcal{X}}, \tag{5.10}$$

which is formally equivalent to equation (5.3). Therefore, in analogy with our previous argumentations, we can assume $\widetilde{\mathcal{X}}' = \mathcal{X}[\widetilde{\mathcal{Q}}'(x,t); \zeta]$ to be the infinitesimal time–translation of $\widetilde{\mathcal{X}}$ under the action of $\mathfrak{g}$, s.t. $\widetilde{\mathcal{X}}' = \widetilde{\mathcal{X}} + \widetilde{\mathcal{X}}_t\,\mathrm{d}\,t$. Inserting this relation and the consistency condition (5.6) in equation (5.10), one finds $\mathcal{T}_{\mathcal{B}x}\,\mathrm{d}\,t = (\widetilde{\mathcal{X}} + \widetilde{\mathcal{X}}_t\,\mathrm{d}\,t)(\mathbb{1}_N + \mathcal{T}_{\mathcal{B}}\,\mathrm{d}\,t) - (\mathbb{1}_N + \mathcal{T}_{\mathcal{B}}\,\mathrm{d}\,t)\widetilde{\mathcal{X}} = (\widetilde{\mathcal{X}}_t + [\mathcal{X}, \mathcal{T}_{\mathcal{B}}])\,\mathrm{d}\,t + \mathcal{O}(\mathrm{d}\,t^2)$, which ultimately tells us that the matrix potential $\widetilde{\mathcal{Q}} = \widetilde{\mathcal{Q}}(x,t)$ satisfies the Lax equation

$$\widetilde{\mathcal{X}}_t - \mathcal{T}_{\mathcal{B}x} + [\widetilde{\mathcal{X}}, \mathcal{T}_{\mathcal{B}}] = \mathbb{O}.$$

As a particular case, assume $\mathcal{T}_{\mathcal{B}} = \widetilde{\mathcal{T}} \equiv \mathcal{T}[\widetilde{\mathcal{Q}}; \zeta]$: then equation (5.7) is replaced with $\mathcal{B}_t = \widetilde{\mathcal{T}}\mathcal{B} - \mathcal{B}\mathcal{T}$ and the matrix potential $\widetilde{\mathcal{Q}}$ satisfies the same Lax equation having $\mathcal{Q}$ as a solution, i.e. $\widetilde{\mathcal{X}}_t - \widetilde{\mathcal{T}}_x + [\widetilde{\mathcal{X}}, \widetilde{\mathcal{T}}] = \mathbb{O}$. Since this will be the working case in what follows, let's remark it with the definition below.

Definition 5.2 - *A gBT $\boldsymbol{\varphi} \mapsto \widetilde{\boldsymbol{\varphi}}' := \mathcal{G}_{\mathcal{B}}\mathcal{B}\boldsymbol{\varphi}$ is called an* ***auto–gauge–Bäcklund Transform*** (*hereafter abbreviated as* **agBT**) *if the matrix $\mathcal{B} = \mathcal{B}[\mathcal{Q}, \widetilde{\mathcal{Q}}; \zeta]$ satisfies the system*

$$\begin{cases} \mathcal{B}_x &= \widetilde{\mathcal{X}}\mathcal{B} - \mathcal{B}\mathcal{X}, \\ \mathcal{B}_t &= \widetilde{\mathcal{T}}\mathcal{B} - \mathcal{B}\mathcal{T}. \end{cases} \tag{5.11}$$

REMARK 5.2 – Forcing the compatibility condition $\mathcal{B}_{xt} = \mathcal{B}_{tx}$ in the system (5.11) yields

$$\begin{cases} \mathcal{B}_{xt} = (\widetilde{\mathcal{X}}_t + \widetilde{\mathcal{X}}\widetilde{\mathcal{T}})\mathcal{B} - \widetilde{\mathcal{X}}\mathcal{B}\mathcal{T} - \widetilde{\mathcal{T}}\mathcal{B}\mathcal{X} + \mathcal{B}(\mathcal{T}\mathcal{X} - \mathcal{X}_t), \\ \mathcal{B}_{tx} = (\widetilde{\mathcal{T}}_x - \widetilde{\mathcal{T}}\widetilde{\mathcal{X}})\mathcal{B} - \widetilde{\mathcal{T}}\mathcal{B}\mathcal{X} - \widetilde{\mathcal{X}}\mathcal{B}\mathcal{T} + \mathcal{B}(\mathcal{X}\mathcal{T} - \mathcal{T}_x), \end{cases} \xrightarrow{\mathcal{B}_{xt}=\mathcal{B}_{tx}} \Big(\widetilde{\mathcal{X}}_t - \widetilde{\mathcal{T}}_x + \big[\widetilde{\mathcal{X}}, \widetilde{\mathcal{T}}\big]\Big)\mathcal{B} = \mathcal{B}\Big(\mathcal{X}_t - \mathcal{T}_x + \big[\mathcal{X}, \mathcal{T}\big]\Big).$$

Therefore, if $\mathcal{B} = \mathcal{B}[\mathcal{Q}, \widetilde{\mathcal{Q}}; \zeta]$ satisfies the system (5.11) and if the compatibility condition holds, then the matrix potentials $\mathcal{Q}, \widetilde{\mathcal{Q}}$ satisfy the same Lax equation.

The system (5.11) gives us an expression for the BT of the NLEE under scrutiny. Actually, it should be noted that the x–component in (5.11) depends just on $\mathcal{X}$ and therefore it holds for all the NLEEs belonging the same hierarchy associated to the spectral problem (5.1), whilst the expression for the t–component varies case by case. We had already encountered this property in §4.2, where we saw that the x–component of the aBT was the same for each NLEE belonging to the same class[42]. However, comparing with the Chen's approach or the Wadati–Sanuki–Konno's one, a substantial simplification occurs within this framework: the expression for the t–component in system (5.11) is fixed by its *asymptotics*, e.g. as $x \to -\infty$. To show this property, consider the spectral problem

$$\boldsymbol{\phi}_x := \widetilde{\mathcal{X}}'\boldsymbol{\phi}. \tag{5.12}$$

Recalling relations (5.3), (5.10) it is easy to verify that $\boldsymbol{\phi} \equiv (\widetilde{\mathcal{G}}\mathcal{B} - \mathcal{B}'\mathcal{G})\boldsymbol{\varphi}$ is a solution of (5.12):

$$\boldsymbol{\phi}_x = \big[\widetilde{\mathcal{G}}_x\mathcal{B} + \widetilde{\mathcal{G}}\mathcal{B}_x - \mathcal{B}'_x\mathcal{G} - \mathcal{B}'\mathcal{G}_x + (\widetilde{\mathcal{G}}\mathcal{B} - \mathcal{B}'\mathcal{G})\mathcal{X}\big]\,\boldsymbol{\varphi} = \widetilde{\mathcal{X}}'\big(\widetilde{\mathcal{G}}\mathcal{B} - \mathcal{B}'\mathcal{G}\big)\boldsymbol{\varphi} = \widetilde{\mathcal{X}}'\boldsymbol{\phi}.$$

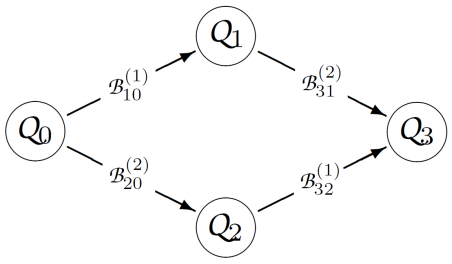
Then it follows that the gauges $\mathfrak{g}$ and $\mathfrak{b}$ *commute iff their asymptotics commute* (in fact, if $\boldsymbol{\phi}$ is zero at (say) $x \to -\infty$, then it is identically zero), which means that

$$\big[\widetilde{\mathcal{G}}^{(-)}\mathcal{B}^{(-)} - \mathcal{B}'^{(-)}\mathcal{G}^{(-)}\big]\boldsymbol{\varphi} = \mathbb{0}, \qquad \forall\, \boldsymbol{\varphi} \neq \mathbf{0}. \tag{5.13}$$

Recalling that the equations (5.7) and (5.8) are equivalent, we can replace the (5.13) with

$$\mathcal{B}_t^{(-)} = \widetilde{\mathcal{T}}^{(-)}\mathcal{B}^{(-)} - \mathcal{B}^{(-)}\mathcal{T}^{(-)}, \tag{5.14}$$

and we can infer that the behavior of the t–component in the system (5.11) is determined by the asymptotics (5.14). Therefore, in order to obtain the expression for the BT related to the desired NLEE, one need to solve for $\mathcal{B}$ the first equation in the system (5.11) and then to fix the time–depending integration constants by requiring that the asymptotics (5.14) is satisfied.



Lastly we have to highlight a remarkable advantage of the agBT method, related to the **permutability theorem**. Since the composition of two gauge transforms is also a gauge transform, the permutability theorem holds trivially within this framework and can be summarized with the identity

$$\mathcal{B}_{31}^{(2)}\mathcal{B}_{10}^{(1)} = \mathcal{B}_{32}^{(1)}\mathcal{B}_{20}^{(2)}, \tag{5.15}$$

where $\mathcal{B}_{ij}^{(k)} \equiv \mathcal{B}[\mathcal{Q}_i, \mathcal{Q}_j; \zeta; \{\beta_k\}]$, being $\{\beta_k\}$ the set of gauge–Bäcklund parameters characterizing the transform. The situation is schematically represented with the *commutative gauge–Bianchi diagram* above. Therefore, if the potential matrices $\mathcal{Q}_i$ with $i = 0, 1, 2, 3$ go to zero as $|x| \to +\infty$ (which is the case of work, having assumed $\mathcal{Q}_{nm} \in \mathrm{S}(\mathbb{R})\ \forall\, n, m = 1, 2, \ldots, N$ where $\mathcal{Q}_{nm}$ is the (nm)–th component of $\mathcal{Q}$) it is easy to verify equation (5.15) as $x \to -\infty$ and this proves the theorem, recalling that the behavior of each $\mathcal{B}$ is univocally determined by its asymptotics. At last, note that equation (5.15) yields the non–linear superposition formula, once the NLEE of interest has been chosen.

[42] E.g. the x–components of the mKdV and SG–aBTs given in equations (4.17) and (4.19), respectively.

5.2 Applications to the ZS spectral problem. We desire now to apply the scheme discussed so far to the 2×2 ZS spectral problem (3.9). Let the 2×2 matrix $\mathcal{T}=\mathcal{T}[\mathcal{Q};\zeta]$ in the auxiliary spectral problem to have the form of a polynomial of degree n in the spectral parameter ζ, that is

$$\mathcal{T}[\mathcal{Q}(x,t);\zeta]:=\sum_{j=0}^{n}\zeta^{n-j}\mathcal{T}_j[\mathcal{Q}(x,t)], \tag{5.16}$$

where the $\mathcal{T}_j$'s are 2×2 complex–valued matrices to be determined. Inserting the relation (5.16) in the AKNS representation for the Lax equation gives ultimately the following equation

$$\zeta^0\big(\mathcal{Q}_t-\mathcal{T}_{nx}+[\mathcal{Q},\mathcal{T}_n]\big)-\sum_{j=0}^{n-1}\zeta^{n-j}\left(\mathcal{T}_{j\,x}-[\mathcal{Q},\mathcal{T}_j]+\imath[\sigma_3,\mathcal{T}_{j+1}]\right)-\zeta^{n+1}[\sigma_3,\mathcal{T}_n]=\mathbb{O},$$

begin $\mathcal{X}=\mathcal{Q}-\imath\zeta\sigma_3$ and $\zeta_t=0$. Setting to zero the coefficients of the powers of ζ, yields the system

$$\text{ZS}\;:\;\begin{cases}[\sigma_3,\mathcal{T}_0]=\mathbb{O}, & (5.17a)\\ \mathcal{T}_{j\,x}-[\mathcal{Q},\mathcal{T}_j]+\imath[\sigma_3,\mathcal{T}_{j+1}]=\mathbb{O},\quad j=0,1,\dots,n-1, & (5.17b)\\ \mathcal{Q}_t-\mathcal{T}_{nx}+[\mathcal{Q},\mathcal{T}_n]=\mathbb{O}, & (5.17c)\end{cases}$$

where the last equation gives the desired NLEE; note that the second equation in the system (5.17) is a recursion relation for the set $\{\mathcal{T}_j\}_{j=0,1,\dots,n-1}$. Now recall that $\mathcal{Q}$ is a 2×2 *off–diagonal* matric, whilst σ_3 is a *diagonal* one; thus, if we separate $\mathcal{T}={}_d\mathcal{T}+{}_f\mathcal{T}$, being ${}_d\mathcal{T}\equiv\operatorname{diag}(\mathcal{T})$ and ${}_f\mathcal{T}\equiv\mathcal{T}-{}_d\mathcal{T}$, it's possible to rewrite more conveniently the system (5.17) in the form[43]

$$\begin{cases}{}_f\mathcal{T}_0=\mathbb{O}, & & (5.18a)\\ {}_d\mathcal{T}_{j\,x}=[\mathcal{Q},{}_f\mathcal{T}_j], & j=0,1,\dots,n, & (5.18b)\\ {}_f\mathcal{T}_{j\,x}=[\mathcal{Q},{}_d\mathcal{T}_j]-2\imath\sigma_3{}_f\mathcal{T}_{j+1}, & j=0,1,\dots,n-1, & (5.18c)\\ \mathcal{Q}_t-{}_f\mathcal{T}_{n\,x}+[\mathcal{Q},{}_d\mathcal{T}_n]=\mathbb{O}. & & (5.18d)\end{cases}$$

Integrating equation (5.18b) with respect to x gives

$${}_d\mathcal{T}_j=\mathcal{I}\big[[\mathcal{Q},\mathcal{T}_j]\big]+\eta_j,\qquad j=0,1,\dots,n, \tag{5.19}$$

where $\mathcal{I}[\,\cdot\,]:=\int_{-\infty}^{x}[\,\cdot\,]\,\mathrm{d}\,x'$ is the x–integral operator over $[-\infty,x]$ and $\{\nu_j\}_{j=0,1,\dots,n}$ is a set of 2×2 arbitrary diagonal matrices independent on the variable x. Inserting equation (5.19) in (5.18c) yields

$${}_f\mathcal{T}_{j\,x}=\Big[\mathcal{Q},\mathcal{I}\big[[\mathcal{Q},{}_f\mathcal{T}_j]\big]\Big]+\big[\mathcal{Q},\eta_j\big]-2\imath\sigma_3{}_d\mathcal{T}_{j+1}.$$

Taking into account that $\sigma_3^2=\mathbb{1}_2$, let's multiply both members of the previous equation by $-\frac{\imath}{2}\sigma_3$: isolating then ${}_f\mathcal{T}_{j+1}$ we obtain the following *recursive relation* for the ZS spectral problem

$${}_f\mathcal{T}_{j+1}=\mathcal{L}[{}_f\mathcal{T}_j]+\mathcal{Q}_j,\qquad {}_f\mathcal{T}_0=\mathbb{O}, \tag{5.20}$$

where $\mathcal{Q}_j(x,t)\equiv\frac{\imath}{2}\sigma_3[\eta_j(t),\mathcal{Q}(x,t)]$ with $j=0,1,2\dots,n-1$ and $\mathcal{L}$ is an *integro–differential operator* acting on the space of (at least) C^1 2×2 off–diagonal complex–valued matrices, defined as

$$\mathcal{L}[\,\cdot\,]:\mathcal{A}\in\mathrm{F}(2,\mathbb{C})\mapsto\mathcal{L}[\mathcal{A}]:=\frac{\imath}{2}\sigma_3\Big(\mathcal{A}_x-\big[\mathcal{Q},\mathcal{I}\big[[\mathcal{Q},\mathcal{A}]\big]\big]\Big)\in\mathrm{F}(2,\mathbb{C}).$$

[43] Let $\mathrm{D}(N,\mathbb{C})$ and $\mathrm{F}(N,\mathbb{C})$ be the sets of $N\times N$ diagonal and off–diagonal matrices, respectively; if $N=2$ then an off–diagonal matrix is also anti–diagonal. This property leads trivially to the following algebra: $d\circ d'\in\mathrm{D}(2,\mathbb{C})$, $f\circ d, d\circ f\in\mathrm{F}(2,\mathbb{C})$ and $f\circ f'\in\mathrm{D}(2,\mathbb{C})$, being $d,d'\in\mathrm{D}(2,\mathbb{C})$, $f,f'\in\mathrm{F}(2,\mathbb{C})$. Moreover, one finds that $[d,d']=\mathbb{O}$, $[d,f],[f,d]\in\mathrm{F}(2,\mathbb{C})$ and $[f,f']=\alpha\sigma_3$ where $\alpha\in\mathbb{C}$ is a constant. Also, note that $[\sigma_3,f]=2\sigma_3 f$ which is true $\forall\, f\in\mathrm{F}(2,\mathbb{C})$.

We'll call $\mathcal{L}$ the **ZS–spectral recursion operator** (not to be confused with the recursion operator $\mathfrak{R}$ defined in §2.4.1). Proceeding in the same fashion for the NLEE (5.18d), we obtain

$$\frac{\imath}{2}\sigma_3 \mathcal{Q}_t = \mathcal{L}[{}_f\mathcal{T}_n] + \mathcal{Q}_n. \tag{5.21}$$

Equations (5.19), (5.20) and (5.21) can be rewritten in a more elegant form, with observing that

$$\begin{aligned} {}_f\mathcal{T}_j &= \mathcal{L}[{}_f\mathcal{T}_{j-1}] + \mathcal{Q}_j = \mathcal{L}^2[{}_f\mathcal{T}_{j-2}] + \mathcal{L}[\mathcal{Q}_{j-1}] + \mathcal{Q}_j = \cdots \\ &= \mathcal{L}^j[{}_f\mathcal{T}_0] + \mathcal{L}^{j-1}[\mathcal{Q}_0] + \mathcal{L}^{j-1}[\mathcal{Q}_1] + \cdots + \mathcal{L}[\mathcal{Q}_{j-1}] + \mathcal{Q}_j. \end{aligned}$$

Recalling that ${}_f\mathcal{T}_0 = \mathbb{O}$ (see equation (5.18a)), we finally find

$$\begin{cases} {}_f\mathcal{T}_j = \displaystyle\sum_{k=0}^{j-1} \mathcal{L}^{j-1-k}[\mathcal{Q}_k], \\ {}_d\mathcal{T}_j = \mathcal{I}\big[[\mathcal{Q}, {}_f\mathcal{T}_j]\big] + \eta_j, \end{cases} \qquad j = 0, 1, \ldots, n. \tag{5.22}$$

We have thus univocally determined, up to the set of matrix integration constants $\{\eta_j\}_{j=0,1,\ldots,n}$, the operator $\mathcal{T}$ of the ZS–auxiliary spectral problem and deduced simultaneously an expression for the associated NLEEs, collected in different *hierarchies* according to the value of n, that is

$$\frac{\imath}{2}\sigma_3 \mathcal{Q}_t(x,t) = \sum_{k=0}^{n} \mathcal{L}^{n-k}[\mathcal{Q}_k(x,t)], \qquad \mathcal{Q}_k(x,t) \equiv \frac{\imath}{2}\sigma_3[\eta_k(t), \mathcal{Q}(x,t)]. \tag{5.23}$$

Remark 5.3 – Although the presence of the integro–differential operator $\mathcal{L}$, the NLEEs in (5.23) are pure PDEs. In fact, it can be proved that the ZS–system is an infinite–dimensional *integrable Hamiltonian system* [Kar998, AbS981], which means that all the NLEEs derivable from equation (5.23) can be casted in symplectic–like form (see equation (2.37)), namely $\mathcal{Q}_t = \mathfrak{J}\delta_{\mathcal{Q}}\mathcal{H}_n[\mathcal{Q}]$ where $\delta_{\mathcal{Q}}\mathcal{H}_n[\mathcal{Q}] := \mathrm{diag}(\delta_u\mathcal{H}[u,v], \delta_v\mathcal{H}[u,v])$, for some (possibly singular) skew–adjoint linear operator $\mathfrak{J}$.

To give some concrete examples of the above considerations, let's consider the first three equations of the ZS–hierarchy. It is convenient to deduce first the NLEE from (5.23) and then the associated auxiliary spectral operator $\mathcal{T}$ from (5.22). If $n = 0$, one finds immediately that

$$\frac{\imath}{2}\sigma_3 \mathcal{Q}_t = \mathcal{Q}_0 = \frac{\imath}{2}\sigma_3\big[\eta_0, \mathcal{Q}\big] \qquad \Longrightarrow \qquad \mathcal{Q}_t(x,t) = \big[\eta_0(t), \mathcal{Q}(x,t)\big],$$

which is a LEE, then uninteresting. Consider therefore the $n = 1$ case: equation (5.23) becomes $\frac{\imath}{2}\sigma_3\mathcal{Q}_t = \mathcal{L}[\mathcal{Q}_0] + \mathcal{Q}_1$, which can be rewritten explicitly by invoking the definitions of $\mathcal{L}$, $\mathcal{Q}_1$ and $\mathcal{Q}_0$ as

$$\mathcal{Q}_t = \big[\eta_1, \mathcal{Q}\big] + \frac{\imath}{2}\sigma_3\big[\eta_0, \mathcal{Q}_x\big] - \frac{\imath}{2}\Big[\mathcal{Q}, \mathcal{I}\big[[\mathcal{Q}, \sigma_3[\eta_0, \mathcal{Q}]]\big]\Big],$$

having noted that $\eta_{0x} = 0$ by definition. Observe now that the matrices $\mathcal{Q}$ and σ_3 satisfy the the following identities: $\mathcal{Q}\sigma_3\mathcal{Q} = -uv\sigma_3$, $\mathcal{Q}^2 = uv\mathbb{1}_2$ and $\mathcal{Q}\sigma_3 = -\sigma_3\mathcal{Q}$; thus the commutator involved in the x–integral operator on the r.h.s. of the above NLEE is zero, that is

$$\big[\mathcal{Q}, \sigma_3[\eta_0, \mathcal{Q}]\big] = \mathcal{Q}\sigma_3\mathcal{Q}\eta_0 + \sigma_3\mathcal{Q}\eta_0\mathcal{Q} - \mathcal{Q}\sigma_3\mathcal{Q}\eta_0 - \sigma_3\eta_0\mathcal{Q}^2 = \mathbb{O}.$$

Therefore, we find the (still uninteresting) LEE $\mathcal{Q}_t = [\eta_1, \mathcal{Q}] + \frac{\imath}{2}\sigma_3[\eta_0, \mathcal{Q}_x]$. Let's try with $n = 2$: in this case we have $\frac{\imath}{2}\sigma_3\mathcal{Q}_t = \mathcal{L}^2[\mathcal{Q}_0] + \mathcal{L}[\mathcal{Q}_1] + \mathcal{Q}_2$ and it can immediately be verified that

$$\mathcal{L}^2[\mathcal{Q}_0] = -\frac{\imath}{8}\sigma_3\Big([\eta_0, \mathcal{Q}_{xx}] - \Big[\mathcal{Q}, \mathcal{I}\big[[\mathcal{Q}, [\eta_0, \mathcal{Q}_x]]\big]\Big]\Big), \qquad \mathcal{L}[\mathcal{Q}_1] = -\frac{1}{4}[\eta_1, \mathcal{Q}_x],$$

where we have used again the fact that $[\mathcal{Q}, \sigma_3[\eta_1, \mathcal{Q}]] = \mathbb{O}$. The NLEE becomes

$$\mathcal{Q}_t = -\frac{1}{4}[\eta_0, \mathcal{Q}_{xx}] + \frac{\imath}{2}\sigma_3[\eta_1, \mathcal{Q}_x] + \frac{1}{4}\Big[\mathcal{Q}, \mathcal{I}\big[[\mathcal{Q}, [\eta_0, \mathcal{Q}_x]]\big]\Big] + [\eta_2, \mathcal{Q}].$$

This equation can be further simplified by noting that each $\eta_i = \eta_i(t)$ is by definition a 2×2 diagonal matrix, thus we may set $\eta_i = \mathrm{diag}(\eta_i{}^{(1)}, \eta_i{}^{(2)})$ with $i = 1, 2, \dots, n$. This yields immediately the identity $[\eta_i, f] = (\eta_i{}^{(1)} - \eta_i{}^{(2)})\sigma_3 f$, which is true $\forall\, f \in \mathrm{F}(2, \mathbb{C})$; therefore one finds

$$\Big[\mathcal{Q}, \mathcal{I}\big[[\mathcal{Q}, [\eta_0, \mathcal{Q}_x]]\big]\Big] = \Big(\eta_0^{(1)} - \eta_0^{(2)}\Big)\Big[\mathcal{Q}, \mathcal{I}\big[[\mathcal{Q}, \sigma_3\mathcal{Q}_x]\big]\Big] = 2\Big(\eta_0^{(1)} - \eta_0^{(2)}\Big)\, uv\sigma_3\mathcal{Q},$$

having noted that $[\mathcal{Q}, \sigma_3\mathcal{Q}_x] = -(uv)_x\sigma_3$ so that $\mathcal{I}\big[[\mathcal{Q}, \sigma_3\mathcal{Q}_x]\big] = -\mathcal{I}\big[(uv)_x\big]\sigma_3 = -uv\sigma_3$ and used then the property $[\sigma_3, f] = 2\sigma_3 f$ with $f \in \mathrm{F}(2, \mathbb{C})$. In the light of the previous observations, we find

$$\mathcal{Q}_t = -\frac{1}{4}\Big(\eta_0^{(1)} - \eta_0^{(2)}\Big)\sigma_3\mathcal{Q}_{xx} + \frac{\imath}{2}\Big(\eta_1^{(1)} - \eta_1^{(2)}\Big)\mathcal{Q}_x + \frac{1}{2}\Big(\eta_0^{(1)} - \eta_0^{(2)}\Big)uv\sigma_3\mathcal{Q} + \Big(\eta_2^{(1)} - \eta_2^{(2)}\Big)\sigma_3\mathcal{Q}, \tag{5.24}$$

which can be equivalently rewritten in terms of its components as

$$\begin{cases} u_t = \dfrac{1}{4}\Big(\eta_0^{(1)} - \eta_0^{(2)}\Big)\big(2u^2v - u_{xx}\big) + \dfrac{\imath}{2}\Big(\eta_1^{(1)} - \eta_1^{(2)}\Big)u_x + \Big(\eta_2^{(1)} - \eta_2^{(2)}\Big)u, & \text{(5.25a)} \\ v_t = \dfrac{1}{4}\Big(\eta_0^{(1)} - \eta_0^{(2)}\Big)\big(v_{xx} - 2uv^2\big) + \dfrac{\imath}{2}\Big(\eta_1^{(1)} - \eta_1^{(2)}\Big)v_x - \Big(\eta_2^{(1)} - \eta_2^{(2)}\Big)v. & \text{(5.25b)} \end{cases}$$

As a particular case, let $\eta_2(t) = a_0(t)\sigma_3$, $\eta_1(t) = a_1(t)\sigma_3$ and $\eta_0(t) = a_2(t)\sigma_3$, where a_0, a_1 and a_2 are arbitrary complex–valued functions independent on x; then the system of NLEEs (5.25) reduces to the system (3.12), obtained by following the AKNS method (see §3.2). Recall that the system (3.12) contains classes of NLEEs, corresponding to appropriate choices of a_0, a_1, a_2, u and v; in particular, setting $a_0 = 0$, $a_1 = 0$, $a_2 = \imath\alpha$ and $v = \varepsilon u^*$, where $\alpha \in \mathbb{R}$ and $\varepsilon = \pm 1$, one recovers respectively the attractive and repulsive NLS equations (see relations (3.14), (3.15)).

Lastly we have to determine the expression for the auxiliary spectral operator $\mathcal{T}$ associated to the family of NLEEs given in (5.24). Setting $n = 2$ in (5.16) leads to $\mathcal{T} = \zeta^2\mathcal{T}_0 + \zeta\mathcal{T}_1 + \mathcal{T}_2$. Invoking therefore the equations (5.22) and repeating the same arguments considered above, we find

$$\mathcal{T}_0 = \eta_0, \quad \mathcal{T}_1 = \frac{\imath}{2}\sigma_3\big[\eta_0, \mathcal{Q}\big] + \eta_1, \quad \mathcal{T}_2 = -\frac{1}{4}\mathcal{I}\big[[\mathcal{Q}, [\eta_0, \mathcal{Q}_x]]\big] - \frac{1}{4}\big[\eta_0, \mathcal{Q}_x\big] + \frac{\imath}{2}\sigma_3\big[\eta_1, \mathcal{Q}\big] + \eta_2.$$

Recalling the notation introduced before for the diagonal matrices η_i, one finally obtains

$$\mathcal{T} = \zeta^2\eta_0 + \zeta\eta_1 + \frac{1}{4}\Big(\eta_0^{(1)} - \eta_0^{(2)}\Big)\big(2\imath\zeta\mathcal{Q} + uv\sigma_3 - \sigma_3\mathcal{Q}_x\big) + \frac{\imath}{2}\Big(\eta_1^{(1)} - \eta_1^{(2)}\Big)\mathcal{Q} + \eta_2. \tag{5.26}$$

5.2.1 Gauge–Bäcklund Transforms for the ZS spectral problem. Let's now determine the aBTs for the ZS system by means of the associated agBTs. Recall that the matrix $\mathcal{B} = \mathcal{B}[\mathcal{Q}, \widetilde{\mathcal{Q}}; \zeta]$ must satisfy the system (5.11), where now $\mathcal{X} = \mathcal{Q} - \imath\zeta\sigma_3$ and $\widetilde{\mathcal{X}} = \widetilde{\mathcal{Q}} - \imath\zeta\sigma_3$, being $\widetilde{\mathcal{Q}}(x,t) := \mathcal{Q}|_{(u,v)\equiv(\widetilde{u},\widetilde{v})}(x,t)$; note also that $\widetilde{\mathcal{T}}$ is obtained from $\mathcal{T}$ via the relation $\widetilde{\mathcal{T}} = \mathcal{T}[\mathcal{Q}; \zeta]_{\mathcal{Q}=\widetilde{\mathcal{Q}}}$. By analogy with the procedure followed in §5.2, assume $\mathcal{B}$ to be a matrix polynomial of degree m in the spectral parameter ζ, namely

$$\mathcal{B}[\mathcal{Q}, \widetilde{\mathcal{Q}}; \zeta] = \sum_{j=0}^{m} \zeta^{m-j}\mathcal{B}_j[\mathcal{Q}, \widetilde{\mathcal{Q}}], \tag{5.27}$$

where the $\mathcal{B}_j$'s are 2×2 complex valued matrices to be determined. Inserting the expression (5.27) into the first equation of the system (5.11), that is $\mathcal{B}_x = \widetilde{\mathcal{Q}}\mathcal{B} - \mathcal{B}\mathcal{Q} + \imath\zeta[\mathcal{B}, \sigma_3]$, we obtain

$$\sum_{j=0}^{m} \zeta^{m-j}\Big(\mathcal{B}_j{}_x - \widetilde{\mathcal{Q}}\mathcal{B}_j + \mathcal{B}_j\mathcal{Q}\Big) = \imath\sum_{j=0}^{m} \zeta^{m-j+1}\big[\mathcal{B}_j, \sigma_3\big].$$

Setting to zero the coefficients belonging to the same powers of ζ, yields finally

$$\text{ZS} \;:\; \begin{cases} [\sigma_3, \mathcal{T}_0] = \mathbb{O}, & (5.28a) \\ \mathcal{B}_{j\,x} = \widetilde{\mathcal{Q}}\mathcal{B}_j - \mathcal{B}_j\mathcal{Q} - \imath[\sigma_3, \mathcal{B}_{j+1}], \quad j = 0, 1, \ldots, m-1, & (5.28b) \\ \mathcal{B}_{m\,x} - \widetilde{\mathcal{Q}}\mathcal{B}_m + \mathcal{B}_m\mathcal{Q} = \mathbb{O}. & (5.28c) \end{cases}$$

Note the strong analogy with the system (5.17); in particular, observe that the equation (5.28c) gives the expression for the x–component of the ZS–aBT between the matrix potentials $\mathcal{Q}$ and $\widetilde{\mathcal{Q}}$.

The system (5.28) can be further simplified with separating $\mathcal{B}$ into a sum of diagonal and off–diagonal matrices, namely $\mathcal{B} = {}_d\mathcal{B} + {}_f\mathcal{B}$. Inserting this into the system (5.28) and recalling the identities $[d, d'] = \mathbb{O}$ $\forall\, d \in \mathrm{D}(2, \mathbb{C})$ and $[\sigma_3, f] = 2\sigma_3 f$ $\forall\, f \in \mathrm{F}(2, \mathbb{C})$, we ultimately obtain

$$\begin{cases} {}_f\mathcal{B}_0 = \mathbb{O}, & (5.29a) \\ {}_d\mathcal{B}_{j\,x} = \widetilde{\mathcal{Q}}_f\mathcal{B}_j - {}_f\mathcal{B}_j\mathcal{Q}, \qquad j = 0, 1, \ldots, m, & (5.29b) \\ {}_f\mathcal{B}_{j\,x} = \widetilde{\mathcal{Q}}_d\mathcal{B}_j - {}_d\mathcal{B}_j\mathcal{Q} - 2\imath\sigma_3 {}_f\mathcal{B}_{j+1}, \quad j = 0, 1, \ldots, m-1, & (5.29c) \\ {}_f\mathcal{B}_{m\,x} = \widetilde{\mathcal{Q}}_d\mathcal{B}_m - {}_d\mathcal{B}_m\mathcal{Q}. & (5.29d) \end{cases}$$

Integration with respect to x of equation (5.29b) leads to

$${}_d\mathcal{B}_j = \mathcal{I}\big[\widetilde{\mathcal{Q}}_f\mathcal{B}_j - {}_f\mathcal{B}_j\mathcal{Q}\big] + \gamma_i, \qquad j = 0, 1, \ldots, m, \tag{5.30}$$

where the γ_i's are 2×2 matrix integration constants independent of x. By analogy with §5.2 we can rewrite the recursive relations (5.29c), (5.29a) in a more elegant form as

$${}_f\mathcal{B}_{j+1} = \Lambda[{}_f\mathcal{B}_j] + \Pi_j, \qquad \Pi_j \equiv -\frac{\imath}{2}\sigma_3\big(\widetilde{\mathcal{Q}}\gamma_j - \gamma_j\mathcal{Q}\big), \qquad j = 0, 1, \ldots, m-1, \tag{5.31}$$

having introduced the integro–differential operator Λ, which acts over the space of (at least) C^1 off–diagonal 2×2 complex–valued matrices and is defined as

$$\Lambda[\,\cdot\,] : f \in \mathrm{F}(2, \mathbb{C}) \mapsto \Lambda[f] := \frac{\imath}{2}\sigma_3\Big(f_x + \mathcal{I}\big[\widetilde{\mathcal{Q}}f - f\mathcal{Q}\big]\mathcal{Q} - \widetilde{\mathcal{Q}}\mathcal{I}\big[\widetilde{\mathcal{Q}}f - f\mathcal{Q}\big]\Big) \in \mathrm{F}(2, \mathbb{C}).$$

Note that the ZS–recursive spectral operator $\mathcal{L}$ is a particular case of the **ZS–recursive gauge–Bäcklund operator** Λ since $\Lambda|_{\widetilde{\mathcal{Q}}=\mathcal{Q}} = \mathcal{L}$. Therefore the gauge–Bäcklund matrix $\mathcal{B} = \mathcal{B}[\mathcal{Q}, \widetilde{\mathcal{Q}}; \zeta]$ is univocally determined, up to the set of matrices $\{\gamma_j\}_{j=0,1,\ldots,m}$, by the following expressions

$$\begin{cases} {}_f\mathcal{B}_j = \displaystyle\sum_{k=0}^{j-1} \Lambda^{j-1-k}[\Pi_j], \\ {}_d\mathcal{B}_j = \mathcal{I}\big[\widetilde{\mathcal{Q}}_f\mathcal{B}_j - {}_f\mathcal{B}_j\mathcal{Q}\big] + \gamma_j, \end{cases} \qquad j = 0, 1, \ldots, m, \tag{5.32}$$

being ${}_f\mathcal{B}_0 = \mathbb{O}$. Within this new formalism, the agBT given by equation (5.29d) becomes

$$\mathbb{O} = {}_f\mathcal{B}_{m\,x} - \widetilde{\mathcal{Q}}_d\mathcal{B}_m + {}_d\mathcal{B}_m\mathcal{Q} = \Lambda[{}_f\mathcal{B}_m] \qquad \Longrightarrow \qquad \sum_{k=0}^{m-1} \Lambda^{m-1-k}[\Pi_k] = \mathbb{O}, \tag{5.33}$$

where we have used again the property ${}_f\mathcal{B}_0 = \mathbb{O}$. The equation (5.33) represents the x component of the aBT for the ZS system and holds for all the NLEEs belonging to the hierarchy. To obtain the t–component of ZS–aBT (which depend on the expression for the auxiliary spectral operator $\mathcal{T}$ related to the particular NLEE under scrutiny) one need to study the asymptotic behavior (say at $x \to -\infty$) of the second equation in the system (5.11). Practically speaking, it is sufficient to insert the expressions (5.32) in equation (5.14) and determining in this way the set of matrices $\{\gamma_j\}_{j=0,1,\ldots,m}$.

As a concrete example, consider the $m = 1$ case: equation (5.27) yields $\mathcal{B} = \zeta\mathcal{B}_0 + \mathcal{B}_1$ and from the equations (5.32) we get $\mathcal{B}_0 = \gamma_0$ and $\mathcal{B}_1 = \Pi_0 + \mathcal{I}\big[\widetilde{\mathcal{Q}}\Pi_0 - \Pi_0\mathcal{Q}\big] + \gamma_1$. Therefore, one has

$$\mathcal{B} = \zeta\gamma_0 + \gamma_1 + \Pi_0 + \mathcal{I}\big[\widetilde{\mathcal{Q}}\Pi_0 - \Pi_0\mathcal{Q}\big].$$

For future reference, let's calculate the integral operator involved in the relation above: recalling that $\Pi_0 = -\frac{\imath}{2}\sigma_3\big(\widetilde{\mathcal{Q}}\gamma_0 - \gamma_0\mathcal{Q}\big)$ and that $\widetilde{\mathcal{Q}}\sigma_3\widetilde{\mathcal{Q}} = -\widetilde{u}\widetilde{v}\sigma_3$, $\mathcal{Q}^2 = uv\mathbb{1}_2$, one finally obtains

$$\mathcal{I}\big[\widetilde{\mathcal{Q}}\Pi_0 - \Pi_0\mathcal{Q}\big] = \frac{\imath}{2}\mathcal{I}\big[\widetilde{u}\widetilde{v} - uv\big]\sigma_3\gamma_0. \tag{5.34}$$

Inserting the equation (5.34) in the expression for the gauge–Bäcklund matrix $\mathcal{B}$ leads to

$$\mathcal{B} = \zeta\gamma_0 + \gamma_1 + \frac{\imath}{2}\sigma_3\Big[\mathcal{I}[\widetilde{u}\widetilde{v} - uv]\gamma_0 - \big(\widetilde{\mathcal{Q}}\gamma_0 - \gamma_0\mathcal{Q}\big)\Big], \tag{5.35}$$

which can be rewritten in terms of its components as

$$\mathcal{B} = \begin{pmatrix} \zeta\gamma_0^{(1)} + \gamma_1^{(1)} + \frac{\imath}{2}\mathcal{I}[\widetilde{u}\widetilde{v} - uv]\gamma_0^{(1)} & \frac{\imath}{2}\Big(\gamma_0^{(1)}u - \gamma_0^{(2)}\widetilde{u}\Big) \\ \frac{\imath}{2}\Big(\gamma_0^{(1)}\widetilde{v} - \gamma_0^{(2)}v\Big) & \zeta\gamma_0^{(2)} + \gamma_1^{(2)} - \frac{\imath}{2}\mathcal{I}[\widetilde{u}\widetilde{v} - uv]\gamma_0^{(2)} \end{pmatrix}.$$

Finally, let's calculate the associated aBTs. Setting $m = 1$ in (5.33), yields $\Lambda[\Pi_0] + \Pi_1 = \mathbb{O}$; thus we have to calculate the expression for the recursive gauge–Bäcklund operator. By definition, one finds

$$\begin{aligned} \Lambda[\Pi_0] &= \frac{\imath}{2}\sigma_3\Big(\Pi_{0x} + \mathcal{I}\big[\widetilde{\mathcal{Q}}\Pi_0 - \Pi_0\mathcal{Q}\big]\mathcal{Q} - \widetilde{\mathcal{Q}}\mathcal{I}\big[\widetilde{\mathcal{Q}}\Pi_0 - \Pi_0\mathcal{Q}\big]\Big) \\ &= \Big(\frac{\imath}{2}\sigma_3\Big)^2\Big\{\mathcal{I}[\widetilde{u}\widetilde{v} - uv]\big(\widetilde{\mathcal{Q}}\gamma_0 + \gamma_0\mathcal{Q}\big) - \big(\widetilde{\mathcal{Q}}_x\gamma_0 - \gamma_0\mathcal{Q}\big)\Big\}. \end{aligned}$$

Having in mind that $\Pi_1 = -\frac{\imath}{2}\sigma_3(\widetilde{\mathcal{Q}}\gamma_1 - \gamma_1\mathcal{Q})$, we finally obtain

$$\big(\widetilde{\mathcal{Q}}_x\gamma_0 - \gamma_0\mathcal{Q}_x\big) - \mathcal{I}[\widetilde{u}\widetilde{v} - uv]\big(\widetilde{\mathcal{Q}}\gamma_0 + \gamma_0\mathcal{Q}\big) - 2\imath\sigma_3\big(\widetilde{\mathcal{Q}}\gamma_1 - \gamma_1\mathcal{Q}\big) = \mathbb{O}, \tag{5.36}$$

which can be written in terms of its components by defining $\gamma_i \equiv \mathrm{diag}(\gamma_i{}^{(1)}, \gamma_i{}^{(2)})$, so that

$$\begin{cases} \Big(\gamma_0^{(2)}\widetilde{u}_x - \gamma_0^{(1)}u_x\Big) - \mathcal{I}[\widetilde{u}\widetilde{v} - uv]\Big(\gamma_0^{(2)}\widetilde{u} + \gamma_0^{(1)}u\Big) - 2\imath\Big(\gamma_1^{(2)}\widetilde{u} - \gamma_1^{(1)}u\Big) = 0, & \text{(5.37a)} \\ \Big(\gamma_0^{(1)}\widetilde{v}_x - \gamma_0^{(2)}v_x\Big) - \mathcal{I}[\widetilde{u}\widetilde{v} - uv]\Big(\gamma_0^{(1)}\widetilde{v} + \gamma_0^{(2)}v\Big) + 2\imath\Big(\gamma_1^{(1)}\widetilde{v} - \gamma_1^{(2)}v\Big) = 0. & \text{(5.37b)} \end{cases}$$

The system (5.37) contains, as particular examples, all the expressions for the x–components of the aBT we've found in §4.2 by following the Chen's method and the Wadati–Konno's approach.

Note that the integral $\mathcal{I}[\widetilde{u}\widetilde{v} - uv]$ can be put in local form by means of the so–called *elemental BTs* [Pem995], which are obtained when the matrices γ_0 and γ_1 are singular; in this case the matrix $\mathcal{B}$ is given by the matrix representative of the agBTs generators. However, for some particular systems different procedures may be used; we present here an approach, suggested in private conversation by B. G. Konopelchenko, for the mKdV equation $u_t - 6\varepsilon u^2u_x + u_{xxx} = 0$ with $\varepsilon = \pm 1$ and $u = \varepsilon u$.

Remark 5.4 – Before entering the discussion, note that the matrix potentials' components are $\mathrm{S}(\mathbb{R})$ functions, so $\mathcal{Q}, \widetilde{\mathcal{Q}} \to 0$ as (say) $x \to -\infty$; the systems (5.22), (5.32) behave then asymptotically as

$$\mathcal{T}^{(-)} \;:\; \begin{cases} {}_f\mathcal{T}_j^{(-)} = \mathbb{O}, \\ {}_d\mathcal{T}_j^{(-)} = \eta_j, \end{cases} \quad j = 0, 1, \dots, n, \qquad \mathcal{B}^{(-)} \;:\; \begin{cases} {}_f\mathcal{B}_i^{(-)} = \mathbb{O}, \\ {}_d\mathcal{B}_i^{(-)} = \gamma_i, \end{cases} \quad i = 0, 1, \dots, m. \tag{5.38}$$

Therefore, in order to fix the expressions for the matrices $\{\eta_j\}_{j=0,1,\dots,n}$ and $\{\gamma_i\}_{i=1,2,\dots,m}$ we need to insert the above expressions into equation (5.14); doing this, we get

$$\mathcal{B}_{it}^{(-)} = \sum_{j=0}^{n}\zeta^{n-j}\Big(\widetilde{\mathcal{T}}_j^{(-)}\mathcal{B}_i^{(-)} - \mathcal{B}_i^{(-)}\mathcal{T}_j^{(-)}\Big) \qquad \Longrightarrow \qquad \gamma_{it} = \gamma_i\sum_{j=0}^{n}\zeta^{n-j}\big(\widetilde{\eta}_j - \eta_j\big) = \mathbb{O},$$

being $\gamma_i, \eta_j \in \mathrm{D}(2,\mathbb{C})$ and having used the fact that $\widetilde{\eta}_j = \eta_j$ for all $j = 0, 1, \dots, n$ since $\widetilde{\mathcal{T}} = \mathcal{T}[\mathcal{Q};\zeta]_{\mathcal{Q}=\widetilde{\mathcal{Q}}}$. It must be emphasized that the property $\gamma_{it} = \mathbb{O}\ \forall\, i = 1, 2, \dots, n$ holds in the case under scrutiny, that is the (2×2) ZS system, but it could not be true for more general system.

From the above consideration it follows that $\gamma_{0t} = \mathbb{O} = \gamma_{1t}$. Now insert the reduction $u = \varepsilon v$ into the system (5.37): equations (5.37a) and (5.37b) are consistent iff $\gamma_0 = -2\imath\mathbb{1}_2$ and $\gamma_1 = \gamma\sigma_3$, being $\gamma \in \mathbb{C}$ an arbitrary constant. Thus the system of equations (5.37) reduces to the expression

$$\big(\widetilde{u} - u\big)_x = \big(\widetilde{u} - u\big)\big(\gamma + \varepsilon\mathcal{I}\big[\widetilde{u}^2 - u^2\big]\big)\,. \tag{5.39}$$

Let's express $\mathcal{I}[\widetilde{u}^2 - u^2]$ in local form by means of the Konopelchenko's approach. Consider the mKdV–BTs with $\gamma_0 = -2\imath\mathbb{1}_2$ and $\gamma_1 = \gamma\sigma_3$ but without the reduction $v = \varepsilon u$. The system (5.37) becomes

$$\begin{cases} \big(\widetilde{u} - u\big)_x = \big(\widetilde{u} + u\big)\Big(\mathcal{I}\big[\widetilde{u}\widetilde{v} - uv\big] + \gamma\Big), & \text{(5.40a)} \\ \big(\widetilde{v} - v\big)_x = \big(\widetilde{v} + v\big)\Big(\mathcal{I}\big[\widetilde{u}\widetilde{v} - uv\big] + \gamma\Big). & \text{(5.40b)} \end{cases}$$

Now multiply the equations (5.40a) by $\widetilde{v}$ and (5.40b) by u and subtract the latter from the former; repeat the procedure, this time with multiplying (5.40a) by v and (5.40b) by $\widetilde{u}$. Ultimately, one finds

$$\begin{cases} \widetilde{u}_x\widetilde{v} + v_x u - \big(u\widetilde{v}\big)_x = \big(\widetilde{u}\widetilde{v} - uv\big)\Big(\mathcal{I}\big[\widetilde{u}\widetilde{v} - uv\big] + \gamma\Big), \\ \big(\widetilde{u}v\big)_x - u_x v - \widetilde{v}_x\widetilde{u} = -\big(\widetilde{u}\widetilde{v} - uv\big)\Big(\mathcal{I}\big[\widetilde{u}\widetilde{v} - uv\big] + \gamma\Big). \end{cases}$$

Subtracting the latter expression from the former, yields finally

$$\big[(\widetilde{u} - u)(\widetilde{v} - v)\big]_x = 2\big(\widetilde{u}\widetilde{v} - uv\big)\Big(\mathcal{I}\big[\widetilde{u}\widetilde{v} - uv\big] + \gamma\Big).$$

Integrate this equation once with respect to x' over the interval $(-\infty, x)$: recalling that $\partial_x\{\big(\mathcal{I}[f(x)]\big)^2\} = 2\mathcal{I}[f(x)]\partial_x\{\mathcal{I}[f(x)]\} = 2f(x)\mathcal{I}[f(x)]$, being f an arbitrary $\mathrm{C}^0(\mathbb{R})$ function, we find

$$\Big(\mathcal{I}\big[\widetilde{u}\widetilde{v} - uv\big]\Big)^2 + 2\gamma\mathcal{I}\big[\widetilde{u}\widetilde{v} - uv\big] - \big(\widetilde{u} - u\big)\big(\widetilde{v} - v\big) = 0, \tag{5.41}$$

where the integration constant has been set to zero since $\mathcal{I}[\widetilde{u}\widetilde{v} - uv] = 0$ for $(\widetilde{u}, \widetilde{v}) = (u, v)$. Equation (5.41) is an algebraic equation of the second order in the unknown $\mathcal{I}[\widetilde{u}\widetilde{v} - uv]$ with solution

$$\mathcal{I}\big[\widetilde{u}\widetilde{v} - uv\big] = \sqrt{\gamma^2 + \big(\widetilde{u} - u\big)\big(\widetilde{v} - v\big)} - \gamma, \tag{5.42}$$

where we have excluded the solution with the minus sign, since $\mathcal{I}[\widetilde{u}\widetilde{v} - uv]$ is zero for $(\widetilde{u}, \widetilde{v}) = (u, v)$. Forcing the reduction $v = \varepsilon u$ in (5.42) and inserting this expression into (5.39), we find

$$\big(\widetilde{u} - u\big)_x = \big(\widetilde{u} + u\big)\sqrt{\gamma^2 + \varepsilon\big(\widetilde{u} - u\big)^2}, \tag{5.43}$$

which is equivalent to the formula (4.18) for the x components of the mKdV–aBTs.

Lastly, let's deduce the *mKdV–solitary–wave solution* from equation (5.43). In order to do this, set $\varepsilon = -1$ in (5.43) and choose $u(x,t) = 0$ as seed–solution; one then obtains

$$\widetilde{u}_x = \widetilde{u}\sqrt{\gamma^2 - \widetilde{u}^2} \qquad \xrightarrow{g \equiv \widetilde{u}/\gamma} \qquad g_x = \gamma g\sqrt{1 - g^2}.$$

The integral $\int[g(1-g^2)^{1/2}]^{-1}\,\mathrm{d}\,g$ can be casted in a fundamental form by means of the substitution $g \mapsto y = (1-g^2)^{1/2}$, which leads to $\gamma x + f(t) = \int(y^2-1)^{-1}\,\mathrm{d}\,y = -\mathrm{arctanh}(y)$. Returning then to the variable $g \equiv \widetilde{u}/\gamma$ and invoking the well–known identity $\cosh^2 x - \sinh^2 x = 1$, we ultimately obtain

$$\widetilde{u}(x,t) = \gamma\mathrm{sech}\big[\gamma x + f(t)\big],$$

where the integration constant $f(t)$ is fixed by assuming $\widetilde{u}$ to be a solution of the mKdV, so that $f(t) = -\gamma^3 t + x_0$. Therefore, the mKdV–solitary–wave solution has the form

$$\widetilde{u}(x,t) = \gamma \operatorname{sech}\big[\gamma(x - \gamma^2 t) + x_0\big], \tag{5.44}$$

being γ the height of the solitary–wave and γ^2 its velocity.

5.3 Permutability theorem and nonlinear superposition formula. In §5.1 we have discussed the role that the permutability theorem plays within the gauge–Bäcklund transform theory and summarized it with equation (5.15). Here we wish first to verify the validity of the condition (5.15) for the agBTs of the ZS spectral problem and then to obtain the associated nonlinear superposition formula.

The proof of the gauge–Bianchi permutability theorem for the ZS spectral problem follows almost trivially; in fact, we've already seen that a sufficient condition to prove the commutativity relation (5.15) is to verify it asymptotically (say) as $x \to -\infty$. But from equation (5.38) we know that the $\mathcal{B}$'s behave asymptotically like pure diagonal matrices, which always commute and this completes the proof.

Let's now focus on the ZS–nonlinear superposition formula. Recalling the arguments discussed in §4.1 for the classical aBT of the SG–equation, we want here to develop an equivalent analysis that would allow us to determine a closed formula expressing *without quadratures* the matrix potential $\mathcal{Q}_3$ in terms of $\mathcal{Q}_0$, $\mathcal{Q}_1$ and $\mathcal{Q}_2$, where $\mathcal{Q}_0$ is a given *seed matrix potential* (see the commutative gauge–Bianchi diagram in §5.1), usually taken as $\mathcal{Q}(x,t) = \mathbb{O}$. Therefore, consider the commutative condition (5.15) and assume $m = 1$ for simplicity: having in mind the shorthand definition for the $\mathcal{B}_{ij}{}^{(k)}$'s given in §5.1 and the fact that (for $m = 1$) the $\mathcal{B}$'s are univocally determined by equations (5.32) up to the matrix integration constants γ_0, γ_1, we ultimately have to calculate explicitly the expression

$$\mathcal{B}\big[\mathcal{Q}_3, \mathcal{Q}_1; \zeta; \alpha_2, \beta_2\big]\mathcal{B}\big[\mathcal{Q}_1, \mathcal{Q}_0; \zeta; \alpha_1, \beta_1\big] = \mathcal{B}^{(2)}_{31}\mathcal{B}^{(1)}_{10} = \mathcal{B}^{(1)}_{32}\mathcal{B}^{(2)}_{20} = \mathcal{B}\big[\mathcal{Q}_3, \mathcal{Q}_2; \zeta; \alpha_1, \beta_1\big]\mathcal{B}\big[\mathcal{Q}_2, \mathcal{Q}_0; \zeta; \alpha_2, \beta_2\big],$$

having defined $(\alpha \equiv \beta) \equiv (\gamma_0, \gamma_1)$, for notation simplicity. Inserting in the above relation the formula (5.35) for the $\mathcal{B}$'s matrices in the $m = 1$ case, leads us the relation

$$\begin{aligned}
&\Big\{\zeta\alpha_2 + \beta_2 + \frac{\imath}{2}\sigma_3\big[(\mathcal{I}_3 - \mathcal{I}_1)\alpha_2 - (\mathcal{Q}_3\alpha_2 - \alpha_2\mathcal{Q}_1)\big]\Big\}\Big\{\zeta\alpha_1 + \beta_1 + \frac{\imath}{2}\sigma_3\big[(\mathcal{I}_1 - \mathcal{I}_0)\alpha_1 - (\mathcal{Q}_1\alpha_1 - \alpha_1\mathcal{Q}_0)\big]\Big\} \\
&= \Big\{\zeta\alpha_1 + \beta_1 + \frac{\imath}{2}\sigma_3\big[(\mathcal{I}_3 - \mathcal{I}_2)\alpha_1 - (\mathcal{Q}_3\alpha_1 - \alpha_1\mathcal{Q}_2)\big]\Big\}\Big\{\zeta\alpha_2 + \beta_2 + \frac{\imath}{2}\sigma_3\big[(\mathcal{I}_2 - \mathcal{I}_0)\alpha_2 - (\mathcal{Q}_2\alpha_2 - \alpha_2\mathcal{Q}_0)\big]\Big\},
\end{aligned}$$

where $\mathcal{I}_i \equiv \frac{1}{2}\int_{-\infty}^{x} u_i v_i \,\mathrm{d}\,x'$ for $i = 0, 1, 2, 3$. Equate to zero the coefficients belonging to same powers of ζ: it follows that ζ^1 and ζ^2 return trivial identities, whilst ζ^0 gives a non–trivial equation that can be conveniently solved by separating again the diagonal terms from the off–diagonal ones. The final results turns out to be [BoT982, Pem995]

$$\mathcal{Q}_3 = 2\sigma_3\big(\mathcal{F}\mathcal{P} - \mathcal{M}\mathcal{N}\big)\big(\sigma_1\mathcal{M}\sigma_1\mathcal{M} - \mathcal{P}^2\big)^{-1}, \tag{5.45}$$

where we operators $\mathcal{F}$, $\mathcal{M}$, $\mathcal{M}$ and $\mathcal{B}$ are defined as

$$\mathcal{P} \equiv \frac{1}{2}\sigma_3\big(\alpha_2\mathcal{Q}_1\alpha_1 - \alpha_1\mathcal{Q}_2\alpha_2\big), \tag{5.46a}$$

$$\mathcal{M} \equiv \big(\mathcal{I}_2 - I - 1\big)\sigma_3\alpha_2\alpha_1 + \imath\big(\alpha_2\beta_1 - \alpha_1\beta_2\big), \tag{5.46b}$$

$$\mathcal{F} \equiv \big(I_1 + I_2 - I_0\big)\sigma_3\mathcal{M} - \frac{1}{2}\mathcal{P}\sigma_3\mathcal{Q}_0 - \frac{1}{4}\alpha_2\big(\mathcal{Q}_2^2 - \mathcal{Q}_1^2\big)\alpha_1, \tag{5.46c}$$

$$\mathcal{N} \equiv \frac{1}{2}\mathcal{M}\sigma_3\mathcal{Q}_0 + \frac{\imath}{2}\sigma_3\big(\beta_2\mathcal{Q}_1\alpha_1 - \alpha_2\mathcal{Q}_1\beta_1 - \beta_1\mathcal{Q}_2\alpha_2 + \alpha_1\mathcal{Q}_2\beta_2\big) - \mathcal{I}_0\mathcal{P}\sigma_3. \tag{5.46d}$$

Note that all the integral operators in (5.46) can be put in local form by means of the elemental gBTs.

6 Painlevé Transcendents

In previous sections we have discussed some properties that IST–solvable NLEEs have in common. In particular, we've understood that this class of PDEs possess a remarkably rich structure: Lax and AKNS pairs, soliton solutions, an infinite sets of conserved quantities, Hamiltonian and bi–Hamiltonian structures, Bäcklund and gauge–Bäcklund transformations. However, as we've pointed out elsewhere through the exposition, we still don't know whether any condition is both necessary and sufficient for a NLEE to be IST–solvable. The theory of the *Painlevé transcendents* places itself in this contest as a possibly conclusive approach to the problem. In fact, with a series of papers published in the early 80's, the team of M. Ablowitz, A. Ramani and H. Segur introduced a conjecture, known today as the **Painlevé conjecture**, which tries to give a unified view of the above scenario [ARS978, ARS80a, ARS80b]. However, despite the lack of counterexamples and the existence of many arguments that points toward the validity of the conjecture [ARS80a], we still lack a complete proof. This section provides a brief introduction to the theory of Painlevé trascendents and its connection with IST–solvable NLEEs. Note that the exposition mainly follows the one given by M. Ablowitz & H. Segur in their monograph [AbS981]; for further readings, we remind to the works of J. Weiss *et al.* [WTC983, Wei983].

We start by recalling some properties of ODEs in complex domain. Consider the n–th order ODE

$$\frac{\mathrm{d}^n w}{\mathrm{d}\, z^n} + \mathcal{P}_1(z)\frac{\mathrm{d}^{n-1} w}{\mathrm{d}\, z^{n-1}} + \cdots + \mathcal{P}_{n-1}(z)\frac{\mathrm{d}\, w}{\mathrm{d}\, z} + \mathcal{P}_n(z) w = 0, \tag{6.1}$$

where $w = w(z)$. If the functions $\mathcal{P}_1, \mathcal{P}_2, \ldots, \mathcal{P}_n$ are analytic at $z = z_0$, then $z_0 \in \mathbb{C}$ is said to be a *regular point* for the ODE (6.1) and for a given initial condition there exists a unique analytic solution in the form of a Laurent series, i.e. $w(z) = \sum_{k\in\mathbb{Z}} a_k (z - z_0)^k$. Thus any singular point of the solutions of the ODE (6.1) must be located only at the singularities of the $\{\mathcal{P}_i\}_{i=1,2,\ldots,n}$, which means that all these singularities are *fixed*, since their location doesn't depend on the initial conditions. This is a common property of ODEs and is lost in the case of NLODEs. A simple example is given by

$$w_z + w^2 = 0 \qquad \Longrightarrow \qquad w\big(z; z_0\big) = \frac{1}{z - z_0},$$

where the location of the singularity depends on the constant of integration. For this reason, these points are called *movable singularities*; note that NLODEs may exhibit both movable and fixed singularities.

Since singular points will play a relevant role in the arguments that will follow, it is convenient to recall some basics about their classification, summarized in the definitions below [AbF003].

Definition 6.1 - *Let z_0 be a point of the compactified complex plane $\mathbb{C}^* := \mathbb{C} \cup \{\infty\}$ s.t. $f \notin \mathrm{C}^\omega(z_0)$. We say that z_0 is an **isolated singularity** of f if $\exists\, r \in \mathbb{R}^+$ and a punctured neighborhood $\Omega_r(z_0) \equiv \{z \in \mathbb{C}^* : |z - z_0| \in \mathbb{R}^+\}$ s.t. $f(z) \in \mathrm{C}^\omega(\Omega_r(z_0))$, $\forall\, z \in \Omega_r(z_0)$. An isolated singularity z_0 of f is called*

- ◇ ***removable*** *iff $a_k \equiv 0\ \forall\, k \in \mathbb{Z}^-$, i.e. iff $f(z) = \sum_{k\in\mathbb{N}_0} a_k (z - z_0)^k$, $\forall\, z \in \Omega_r(z_0)$;*
- ◇ *a **pole of order m** iff $\exists\, m \in \mathbb{N}$ finite s.t. $f(z) = \sum_{k=-m}^{+\infty} a_k (z - z_0)^k$, $\forall\, z \in \Omega_r(z_0)$;*
- ◇ ***essential*** *iff it's a pole of infinite order, i.e. iff $f(z) = \sum_{k\in\mathbb{Z}} a_k (z - z_0)^k$, $\forall\, z \in \Omega_r(z_0)$.*

Definition 6.2 - *Let $f \in \mathrm{C}^\omega(\Omega)$, Ω an open set of $\mathbb{C}^*$. A point $z_0 \in \Omega$ is called a **branch point** of f if $f(z)$ is discontinuous upon traversing a small circuit around z_0*[44].

Hereafter we will call **critical points** the family of all singular points that are not poles (of any order).

The need for a complete classification of singularities has its origin in a problem arose in the late 19th century, that is the classification of ODEs on the basis of the singularities they admit [Inc956]. In particular, a special interest was reserved for ODEs satisfying the following

[44] Branch points are divided into three subclasses: *algebraic*, *transcendental* and *logarithmic*. Note that all these classes of singularities aren't independent, e.g. transcendental and logarithmic branch points are also essential singularities.

PROPERTY 6.1 - *Let $\mathcal{F} = \mathcal{F}(\mathrm{d}_z^{n-1} w, \mathrm{d}_z^{n-2} w, \ldots, w; z)$ be a rational function in w and its derivatives, which is also locally analytic in z. We say that the n–th order ODE*

$$\frac{\mathrm{d}^n w}{\mathrm{d} z^n} = \mathcal{F}\left(\frac{\mathrm{d}^{n-1} w}{\mathrm{d} z^{n-1}}, \frac{\mathrm{d}^{n-2} w}{\mathrm{d} z^{n-2}}, \ldots, w; z\right), \tag{6.2}$$

satisfies the ***Painlevé property*** *(equiv. is a $\mathscr{P}$–**type**) if its movable singularities are* at worst *poles.*

First attempts in the classification of ODEs of the form (6.2) where made by J. H. POINCARÉ (1854–1912) and J. I. FUCHS (1833–1902), who showed that out of the class of first order equations, the only ones of $\mathscr{P}$–type can be transformed into generalized Riccati equations, i.e. $w' = \mathcal{P}_0 + \mathcal{P}_1 w + \mathcal{P}_2 w^2$. After these results, C. E. PICARD (1856–1941) pointed out that for orders greater than one, movable critical points can occur, but failed in trying to find new examples. Around 1900 S. V. KOVALEVSKAYA (1850—1891), P. PAINLEVÉ (1863–1933) and B. GAMBIER studied second order ODEs of the form (6.2) satisfying the Painlevé property and found that, up to a *Möebius transform*

$$\mathcal{W}(\widetilde{z}) := \frac{a(z) w + b(z)}{c(z) w + d(z)}, \qquad \widetilde{z} \equiv \phi(w; z), \tag{6.3}$$

where a, b, c, d and ϕ are locally analytic functions, every $\mathscr{P}$–type 2nd order ODE can be put into one of fifty *canonical forms* [Kov889, Pai900, Pai902, Gam910]. Out of these, fortyfour types could be reduced to already known ODEs (solvable in terms of trigonometric or elliptic functions), whilst the remaining six defined new NLODEs that cannot be reduced to any simpler known ODE. These are the so–called **Painlevé transcendents** (or *Kovalevskaya–Painlevé–Gambier transcendents*), listed below:

$$\mathscr{P}_{\mathrm{I}} \quad : \quad \frac{\mathrm{d}^2 w}{\mathrm{d} z^2} = 6w^2 + z, \tag{6.4}$$

$$\mathscr{P}_{\mathrm{II}} \quad : \quad \frac{\mathrm{d}^2 w}{\mathrm{d} z^2} = 2w^3 + wz + \alpha, \tag{6.5}$$

$$\mathscr{P}_{\mathrm{III}} \quad : \quad \frac{\mathrm{d}^2 w}{\mathrm{d} z^2} = \frac{1}{w}\left(\frac{\mathrm{d} w}{\mathrm{d} z}\right)^2 - \frac{1}{z}\frac{\mathrm{d} w}{\mathrm{d} z} + \frac{\alpha w^2 + \beta}{z} + \gamma w^3 + \frac{\delta}{w}, \tag{6.6}$$

$$\mathscr{P}_{\mathrm{IV}} \quad : \quad \frac{\mathrm{d}^2 w}{\mathrm{d} z^2} = \frac{1}{2w}\left(\frac{\mathrm{d} w}{\mathrm{d} z}\right)^2 \frac{3}{2} w^3 + \frac{3w^3}{2} + 4zw^2 + 2(z^2 - \alpha)w + \frac{\beta}{w}, \tag{6.7}$$

$$\mathscr{P}_{\mathrm{V}} \quad : \quad \frac{\mathrm{d}^2 w}{\mathrm{d} z^2} = \left(\frac{1}{2w} + \frac{1}{w-1}\right)\left(\frac{\mathrm{d} w}{\mathrm{d} z}\right)^2 - \frac{1}{z}\left(\frac{\mathrm{d} w}{\mathrm{d} z}\right) + \frac{(w-1)^2}{z^2}\left(\alpha w + \frac{\beta}{w}\right), \tag{6.8}$$

$$\begin{aligned} \mathscr{P}_{\mathrm{VI}} \quad : \quad \frac{\mathrm{d}^2 w}{\mathrm{d} z^2} &= \frac{1}{2}\left(\frac{1}{w} + \frac{1}{w-1} + \frac{1}{w-z}\right)\left(\frac{\mathrm{d} w}{\mathrm{d} z}\right)^2 - \left(\frac{1}{z} + \frac{1}{z-1} + \frac{1}{w-z}\right)\frac{\mathrm{d} w}{\mathrm{d} z} + \\ &+ \frac{w(w-1)(w-z)}{z^2(z-1)^2}\left(\alpha + \beta\frac{z}{w^2} + \gamma\frac{z-1}{(w-1)^2} + \delta\frac{z(z-1)}{(w-z)^2}\right), \end{aligned} \tag{6.9}$$

being α, β, γ and δ are arbitrary constants[45]. Although we will focus on $\mathscr{P}_{\mathrm{I}}$–$\mathscr{P}_{\mathrm{VI}}$, it has to be noted that the classification of higher order $\mathscr{P}$–type ODEs is still an open problem. Actually we know just few examples of $\mathscr{P}$–type ODEs of 3rd order [Cha911] and even fewer of 4th and 5th order[46] [Kud997].

There are at least two tests to check if an ODE is of $\mathscr{P}$–type. The first one follows from the fact that if a 2nd order ODE possesses the Painlevé property, then it is either linearizable or can be put into one of the six Painlevé transcendents by appropriate coordinate transforms. Otherwise, especially

[45] More correctly, one should call the solutions of $\mathscr{P}_{\mathrm{I}}$–$\mathscr{P}_{\mathrm{VI}}$ as Painlevé transcendents, since for arbitrary values of the parameters α, β, γ and δ the general solutions of $\mathscr{P}_{\mathrm{I}}$–$\mathscr{P}_{\mathrm{VI}}$ are *transcendental*, that is they cannot be expressed in closed–form by means of elementary functions. However, the six equations are often referred in the literature as the six Kovalevskaya–Painlevé–Gambier transcendents and for this reason we'll adopt this choice hereafter.

[46] For further readings see U. MUČAN & F. JRAD, *Painlevé test and higher order ODEs* [MuJ002] and references therein.

if one suspects that the ODE is not of $\mathscr{P}$–type, a *singular point analysis* may be performed [Dun010]. In fact, if an n–th order ODE is of $\mathscr{P}$–type, then its general solution admits a Laurent expansion with a finite number of negative powers. Assume therefore that the leading term in the expansion goes like $w(z) \sim a(z-z_0)^p$ as $z \to z_0$, with $a, p \in \mathbb{C}$ and $a \neq 0$. Inserting this into the ODE and requiring the *maximal balance condition* (i.e. two (or more) terms must be of equal maximally small order as $z \to z_0$), one determines a and p and finally the form of a solution around z_0. Therefore, if z_0 is a singularity we can determine if it's movable and to what class it belongs. As an example, consider the ODE

$$w' = w^3 + z;$$

the maximal balance condition gives $ap(z-z_0)^{p-1} \sim a^3(z-z_0)^{3p}$ so $p = -1/2$, $a = \pm\imath/\sqrt{2}$ and then

$$w(z) \overset{z\to z_0}{\sim} \pm\frac{\imath}{\sqrt{2}}\big(z-z_0\big)^{-1/2}.$$

Therefore $w = w(z)$ possesses a movable branch point at $z = z_0$ and the ODE is not of $\mathscr{P}$–type.

REMARK 6.1 – The Painlevé property guarantees that the solutions of equations $\mathscr{P}_{\rm I}$–$\mathscr{P}_{\rm VI}$ are single valued, thus giving rise to proper functions. The importance of the Painlevé transcendents is that they properly define *new transcendental functions*, just as e.g. the exponential function $w(z) = e^z$ can be defined as the general solution of the ODE $w' = w$ with initial condition $w(0) = 1$ [Dun010]. A subtle issue related to these new transcendental functions is their **irreducibility**, in the sense of the *Umemura definition* [Ume987]. Roughly speaking, this means that within the field of smooth functions, the solutions of $\mathscr{P}_{\rm I}$–$\mathscr{P}_{\rm VI}$ form a set that is disjoint from the one of *classical functions*, which is defined by starting off with the rational functions $\mathfrak{Q}$ and adjoining those functions which arise as solutions of algebraic and linear differential equations with coefficients in $\mathfrak{Q}$. Painlevé himself anticipated that the solutions of $\mathscr{P}_{\rm I}$–$\mathscr{P}_{\rm VI}$ define "irreducible functions" (i.e. new transcendentals) but some rigorous proofs appeared only recently for $\mathscr{P}_{\rm I}$ [Ume987, Ume990], $\mathscr{P}_{\rm II}$ and $\mathscr{P}_{\rm IV}$ [NoO997]. What's interesting is that to show this property the authors use a far reaching extension of the Galois Theory from the number fields to differential fields of functions [Dun010]; in this way the irreducibility problem for the Painlevé equations becomes noting but than the analogous problem to the existence of non–algebraic numbers[47].

Finally note that the Painlevé transcendents satisfy the so–called *coalescence cascade* relations, which can be schematically expressed by means of the so–called *coalescence diagram* [Oka986]

$$\begin{array}{ccccc} \mathscr{P}_{\rm VI} & \longrightarrow & \mathscr{P}_{\rm V} & \longrightarrow & \mathscr{P}_{\rm IV} \\ & & \downarrow & & \downarrow \\ & & \mathscr{P}_{\rm III} & \longrightarrow & \mathscr{P}_{\rm II} \longrightarrow \mathscr{P}_{\rm I} \end{array}$$

For example, setting $\alpha \equiv 4\varepsilon^{-15}$ and $w(z;\alpha) = \varepsilon\omega(\zeta) + \varepsilon^{-5}$ with $z = \varepsilon^2\zeta - 6\varepsilon^{-10}$ in $\mathscr{P}_{\rm II}$, yields

$$\omega_{\zeta\zeta} = 6\omega^2 + \zeta + \varepsilon^6\big(2\omega^3 + \zeta\omega\big) \qquad \overset{\varepsilon\to 0}{\longrightarrow} \qquad \big(\mathscr{P}_{\rm I}\big) \;:\; \omega'' = 6\omega^2 + \zeta.$$

Similar transformations allows one to visit the remaining directions in the diagram [OLBC10].

6.1 Lax pairs and Hamiltonian structure. In this section we'll show that the Painlevé equations are actually *isospectral deformation equations*, which can be obtained from a Lax–type compatibility condition. We stress that here the role of the independent variables x, t does not necessarily parameterize space and time; indeed, the whole argument is extrapolated from a much more general contest that won't be discussed [FlN980]. Also, we'll see how $\mathscr{P}_{\rm I}$–$\mathscr{P}_{\rm VI}$ can be interpreted as Hamiltonian systems.

[47] For further readings see V. I. GOMAK, I. LAINE, S. SMIMOMURA – *Painlevé differential equations in the complex plane* [GLS002] or the recent monograph of R. CONTE, M. MUSETTE entitled *The Painlevé Handbook* [CoM008].

Let's start from the equation $\mathscr{P}_{\mathrm{I}}$. Introduce the linear operators $\mathcal{X}_{\mathscr{P}_{\mathrm{I}}} = \mathcal{X}_{\mathscr{P}_{\mathrm{I}}}[w(x,t);x,t]$ and $\mathcal{T}_{\mathscr{P}_{\mathrm{I}}} = \mathcal{T}_{\mathscr{P}_{\mathrm{I}}}[w(x,t);x,t]$ associated, say, to the principal and auxiliary spectral problems, respectively

$$\begin{aligned}\mathcal{X}_{\mathscr{P}_{\mathrm{I}}}[w;x,t] &\equiv \big(4x^4+2w^2+t\big)\sigma_3 - \imath\big(4x^2w+2w^2+t\big)\sigma_2 - \Big(2xw'+\frac{1}{2x}\Big)\sigma_1,\\ \mathcal{T}_{\mathscr{P}_{\mathrm{I}}}[w;x,t] &\equiv \Big(x+\frac{1}{x}\Big)\sigma_3 - \imath\frac{w}{x}\sigma_2,\end{aligned} \tag{6.10}$$

where σ_1,σ_2 and σ_3 are the Pauli–matrices. The compatibility condition force then $\mathcal{X}_{\mathscr{P}_{\mathrm{I}}}$ and $\mathcal{T}_{\mathscr{P}_{\mathrm{I}}}$ to satisfy the Lax–equation in AKNS formalism (3.8), conveniently rewritten here as

$$\mathcal{X}_{\mathscr{P}_{\mathrm{I}}t} - \mathcal{T}_{\mathscr{P}_{\mathrm{I}}x} = \big[\mathcal{T}_{\mathscr{P}_{\mathrm{I}}},\, \mathcal{X}_{\mathscr{P}_{\mathrm{I}}}\big].$$

Recalling the commutativity relation $[\sigma_i,\sigma_j] = \imath\varepsilon_{ijk}\sigma_k$, being ε_{ijk} the Levi–Civita symbol, one finds

$$\begin{aligned}\mathcal{X}_{\mathscr{P}_{\mathrm{I}}t} - \mathcal{T}_{\mathscr{P}_{\mathrm{I}}x} &= \Big(4ww' - \frac{w'}{x} + \frac{w}{x^2}\Big)\sigma_3 + \imath\Big(\frac{w'}{x} - \frac{w}{x^2} - 4ww' - 4x^2w' - 1\Big)\sigma_2 - 2xw''\sigma_1,\\ \big[\mathcal{T}_{\mathscr{P}_{\mathrm{I}}},\mathcal{X}_{\mathscr{P}_{\mathrm{I}}}\big] &= \Big(2ww' + \frac{w}{2x^2}\Big)\sigma_3 - \imath\Big(2x^2w' + 2ww' + \frac{1}{2} + \frac{w}{2x^2}\Big)\sigma_2 - \big(6xw^2+tx\big)\sigma_1.\end{aligned}$$

Equating the coefficients of σ_1 (being the only ones to contain a non–trivial NLEE), one finds

$$w'' = 3w^2 + \tfrac{1}{2}t;$$

thus $w = w(x,t)$ satisfies the first Painlevé transcendental equation with $z = t$.

In analogous fashion, consider the following two linear operators for the equation $\mathscr{P}_{\mathrm{II}}$:

$$\begin{aligned}\mathcal{X}_{\mathscr{P}_{\mathrm{II}}}[w;x,t] &\equiv -\imath\big(4x^2+2w^2+t\big)\sigma_3 - 2w'\sigma_2 + \big(4xw - \alpha x^{-1}\big)\sigma_1,\\ \mathcal{T}_{\mathscr{P}_{\mathrm{II}}}[w;x,t] &\equiv w\sigma_1 + \imath x\sigma_3,\end{aligned} \tag{6.11}$$

where $\alpha\in\mathbb{C}$ is an arbitrary constant. Following the same procedure as for $\mathscr{P}_{\mathrm{I}}$, we obtain

$$\begin{aligned}\mathcal{X}_{\mathscr{P}_{\mathrm{II}}t} - \mathcal{T}_{\mathscr{P}_{\mathrm{II}}} &= -2\imath ww'\sigma_3 - \big(2w^3+tw+\alpha\big)\sigma_2 + 2xw'\sigma_1,\\ \big[\mathcal{T}_{\mathscr{P}_{\mathrm{II}}},\mathcal{X}_{\mathscr{P}_{\mathrm{II}}}\big] &= -4\imath ww'\sigma_3 - 2w''\sigma_2 + (4x-1)w'\sigma_1.\end{aligned}$$

Equating the σ_2's coefficients, we find the equation (6.5). Analogous, but more involved, isospectral deformation equations can be obtained for the remaining transcendentals [JiM981, FlN980].

Finally, let's briefly discuss the connection with the Hamiltonian formulation. The main observation is that the equations $\mathscr{P}_{\mathrm{I}}$–$\mathscr{P}_{\mathrm{VI}}$ can be written in the form of Hamilton–Liouville equations

$$\frac{\mathrm{d}\,q}{z} = \frac{\partial\mathcal{H}[q,p;z]}{\mathrm{d}\,\partial p}, \qquad \frac{\mathrm{d}\,q}{z} = -\frac{\partial\mathcal{H}[q,p;z]}{\mathrm{d}\,\partial q}$$

for *non-autonomous Hamiltonian systems*, i.e. systems whose Hamiltonian depends explicitly on the independent variables. For instance, consider the following non–autonomous Hamilton function

$$\mathcal{H}_{\mathscr{P}_{\mathrm{I}}}[q,p;z] = \tfrac{1}{2}p^2 - 2q^3 - zq \qquad\Longrightarrow\qquad \begin{cases} q' = p,\\ p' = 6q^2 + z,\end{cases}$$

thus $q'' = 6q^2 + z$ and $q = q(z)$ satisfies $\mathscr{P}_{\mathrm{I}}$. Similarly it's straightforward to verify that

$$\mathcal{H}_{\mathscr{P}_{\mathrm{II}}}[q,p;z] = \tfrac{1}{2}p^2 - \big(q^2+\tfrac{1}{2}z\big)p - \big(\alpha+\tfrac{1}{2}\big)q,$$

is the Hamilton function for $\mathscr{P}_{\mathrm{II}}$. In fact, one finds $q' = p - q^2 - \frac{z}{2}$ and $p' = 2qp + \alpha + \frac{1}{2}$ so that

$$q'' = q^3 - \tfrac{1}{2}qz + \alpha + 1,$$

which means that $q = q(z)$ satisfies $\mathscr{P}_{\rm II}$ with $\alpha \equiv \alpha - 1$. Finally the Hamiltonian for $\mathscr{P}_{\rm III}$ is

$$\mathcal{H}_{\mathscr{P}_{\rm III}}[q,p;z,\theta_0,\theta_\infty,\kappa_0,\kappa_\infty] = \frac{1}{z}q^2p^2 - \frac{1}{z}\big[\kappa_\infty z q^2 + \big(2\theta_0+1\big)q - \kappa_0 z\big]p + \kappa_\infty\big(\theta_0+\theta_\infty\big)q,$$

where $\theta_0, \theta_\infty, \kappa_0$ and κ_∞ are arbitrary constants. In fact, it can be proved that $q = q(z)$ satisfies $\mathscr{P}_{\rm III}$ with $(\alpha,\beta,\gamma,\delta) \equiv \big(-2\kappa_\infty\theta_\infty, 2\kappa_0(\theta_0+1), \kappa_\infty^2, -\kappa_0^2\big)$ [OLBC10]. Analogous (but rather involved) Hamiltonian structures can be introduced for the remaining transcendentals $\mathscr{P}_{\rm IV}$–$\mathscr{P}_{\rm VI}$ [FoW004, Oka987].

6.2 Connection with IST–solvable NLEEs. In the last paragraph we've noted that the Painlevé transcendents have many features in common with the NLEEs studied so far. Those common structures suggest the presence of a direct connection with (say) KdV, SG or NLS, but don't give a clear answer to what this connection could be. This paragraph is devoted to give a clue on how the Painlevé transcendents are related to IST–solvable NLEEs. As an introductory example, consider the KdV equation $\mathcal{K}[u] = 0$: note that it admits traveling–wave solutions, corresponding to the scaling relations

$$\xi = x - ct, \qquad u(x,t) = \mathcal{U}(\xi), \tag{6.12}$$

which yield the relations $u_t = \mathcal{U}_\xi \xi_t = -c\mathcal{U}_\xi$ and $\mathcal{U}_x = \mathcal{U}_\xi$. Therefore the function $\mathcal{U} = \mathcal{U}(z)$ satisfies the 2nd order ODE $\mathcal{U}_{\xi\xi\xi} + 6\mathcal{U}\mathcal{U}_\xi - c\mathcal{U}_\xi = 0$ which can be integrated once with respect to z in order to give

$$\mathcal{U}_{\xi\xi} + 3\mathcal{U}^2 - c\mathcal{U} = \alpha' \qquad \xrightarrow{\mathcal{U}\mapsto w + c/6} \qquad w'' + 3w^2 + \alpha = 0,$$

where $\alpha \equiv \alpha' + \frac{c^2}{16}$ is an arbitrary constant. The solution of the rescaled ODE above is given by the Weierstrass elliptic function whose only singularities are poles, thus satisfying the Painlevé property. Now consider the re–scaling associated to the so–called *similarity* KdV–solution, given by

$$z = x + 3\lambda t^2, \qquad u(x,t) = w(z) - \lambda t,$$

being $\lambda \in \mathbb{C}$; in this case $w = w(z)$ satisfies the ODE $w''' + 6ww' - \lambda = 0$, whose integration gives

$$w'' + 3w^2 - \lambda z = \alpha \qquad \xrightarrow{w\mapsto -w} \qquad w'' = 3w^2 - \lambda z - \alpha.$$

Confronting the scaled ODE above with the first Painlevé transcendent (6.4), one deduces that $w = w(z)$ is expressible in terms of solutions of $\mathscr{P}_{\rm I}$. Equivalently, it can be verified for the scaling reduction

$$z = x(3t)^{-1/3}, \qquad u(x,t) = \big(3t\big)^{-2/3}\big[w'(z) + w^2(z)\big], \tag{6.13}$$

that the function $w = w(z)$ satisfies the second Painlevé transcendent (see appendix A).

DEFINITION 6.3 - *Each ODE obtained from a PDE by a proper scaling reduction (i.e. by suitably restricting the set of its solutions) is called an* ***exact reduction*** *of the PDE.*

Another example is given by the mKdV: an exact reduction corresponds to the scaling relation

$$z = x(3t)^{-1/3}, \qquad u(x,t) = \big(3t\big)^{-1/3}w(z), \tag{6.14}$$

which leads to $u_t = -(3t)^{-4/3}(zw)'$ and $u_x = (3t)^{-2/3}w'$, so that one finds

$$w''' - 6w'w^2 = (zw)' \qquad \longrightarrow \qquad w'' = 2w^3 + zw + \alpha.$$

Thus $w = w(z)$ satisfies $\mathscr{P}_{\rm II}$ with α an integration constant. Similar considerations follow also for the Boussinesq equation $u_{tt} = u_{xx} + \frac{1}{2}(u^2)_{xx} + \frac{1}{4}u_{x^4}$ (see §1.2). An exact reduction was obtained by Zakharov by looking for traveling–wave solution $u(x,t) = w(z)$ with $z = x - ct$, where $w = w(z)$ satisfies the ODE $w'''' + 2(w^2)'' + 2(1-c^2)w'' = 0$, which can be integrated twice to give

$$w'' + 2w^2 + 4(1-c)w + \alpha z + \beta = 0 \qquad \xrightarrow{w\mapsto w-1+c^2} \qquad w'' + 2w^2 + \alpha z + \beta' = 0,$$

where $\beta' \equiv \beta + 4c^2 - 2c^4 - 2$ is an arbitrary constant. Depending weather $\alpha = 0$ or $\alpha \neq 0$, the function $w = w(z)$ is expressible in terms of the Weierstrass elliptic function or solutions of $\mathscr{P}_{\rm I}$, respectively. Finally, consider the (light–cone) *SG* equation $u_{xt} = \sin u$ and assume the scaling reduction

$$z = xt, \qquad u(x,t) = w(z), \tag{6.15}$$

where the function $w = w(z)$ satisfies the ODE

$$zw'' + w' = \sin w \qquad \xrightarrow{w \mapsto w = \imath \ln f} \qquad f'' = \frac{(f')^2}{f} - \frac{f'}{z} + \frac{f^2}{2z} - \frac{1}{2z},$$

which is equivalent to $\mathscr{P}_{\rm III}$ with $\alpha = -\beta = 1/2$ and $\gamma = \delta = 0$. The previous equations represent only few examples of many other known IST–solvable NLEEs that have $\mathscr{P}$–type exact reductions, e.g. it has been proved that the derivative–NLS (see equation (1.24)) eventually reduces to $\mathscr{P}_{\rm IV}$ [ARS80b].

What is remarkable in the above scenario is that all those NLEEs are IST–solvable, thus suggesting an apparent correspondence between $\mathscr{P}$–type exact reductions and IST–solvability. Having in mind the connection between IST–solvable NLEE and *integrable* infinite dimensional Hamiltonian systems (see §2.4.1), Ablowitz, Ramani and Segur made in 1980 the following, still unproved, conjecture:

ARS Conjecture - *A NLEE is integrable only if all its nonlinear exact reductions are of $\mathscr{P}$–type.*

Let us assume the validity of the conjecture in order to discuss how it can be used to test PDEs' integrability. First note that the scaling reductions (6.12) –(6.15) are actually group–invariant solutions of the equation under the action of some particular *Lie point symmetry*. As an example, consider the SG equation $u_{xt}(x,t) = \sin[u(x,t)]$ and the symmetry [Olv993] $(x,t) \mapsto (\widetilde{x},\widetilde{t}) \equiv (cx, c^{-1}t)$ with $c \neq 0$; the group invariant solutions are of the form $u(x,t) = \mathcal{F}(z)$ where $z = xt$ is an invariant of the symmetry. Similar arguments follow for the mKdV equation $\mathcal{M}[u] = 0$, for which one may introduce the Lie–point symmetry $(u;x,t) \mapsto (\widetilde{u};\widetilde{x},\widetilde{t}) \equiv (c^\alpha u; c^\beta x, c^\gamma t)$ with $c \neq 0$. The symmetry condition holds if all the terms in the triple have equal weights, i.e. $\alpha - \gamma = 3\alpha - \beta = \alpha - 3\beta$ which yields $\beta = -\alpha$, $\gamma = -3\alpha$ and α arbitrary. The corresponding symmetry group, depending on the parameter c^α, is generated by $\mathcal{V} = v\partial_v - x\partial_x - 3t\partial_t$ and admits the invariants $z = (3t)^{-1/3}x$ and $w = (3t)^{1/3}v$ (having introduced the constant factor $3^{1/3}$ for convenience). Thus the group invariant solutions are of the form $v(x,t) = (3t)^{-1/3}w(z)$ with $z = (3t)^{-1/3}x$, which is the scaling reduction given in equation (6.14).
We can therefore introduce an integrability test for PDEs, known as **Painlevé test**, summarized with the help of the following algorithm [Dun010]: given a PDE

$\wp_1$) find all the associated Lie–point symmetries;

$\wp_2$) construct exact reductions from the group invariant solutions;

$\wp_3$) check if all the non–trivial ODEs obtained satisfy the Painlevé property.

If all the reductions are of $\mathscr{P}$–type then one may start looking for a BT or a scattering problem with some confidence. Note that the Painlevé test just gives a *necessary condition* for integrability, which means that if all the reductions of a given PDE are of $\mathscr{P}$–type, then the PDE does not have to be integrable in general. On the other hand, if an exact reduction is not of $\mathscr{P}$–type, then one can be quite confident (in the sense of the conjecture above) that the PDE under scrutiny is not IST–solvable in its present form. In this case, a transformation may be available to make the not–$\mathscr{P}$–type ODE of $\mathscr{P}$–type; if such a transformation exists, then the transformed PDE is a candidate for IST[48].

[48] These transformations are often suggested by the details of the singular point analysis [AbS981].

7 Asymptotic perturbative methods and universality

In the previous sections we've discussed the structure of IST–solvable NLEEs. Now we wish to study a different property of certain NLEEs, namely their **universality**. Recall that in §1.2 we said that, out of the family of the so–called derivative NLEs (1.24), the NLS (1.23) plays an important role, since it is possible to put a generic dispersive NLEE into the NLS form after an appropriate *multiple scale limit*. We will refer hereafter to this property as the universality for the NLS equation. To discuss this argument, we'll need to introduce some basics of *asymptotic analysis* and *perturbative methods*; however, we refer the interested reader to the monograph of A. Jeffrey & T. Kawahara entitled "*Asymptotic methods in Nonlinear wave theory*", from which the above argumentations are taken.

7.1 Asymptotic series and perturbations. First recall some basic notions about the so–called *Landau notation*. Let $f = f(x)$ and $g = g(x)$ be two functions defined on some open subset of $\mathbb{R}$, with $g(x) \neq 0$ for values of x sufficiently close to zero: then $f(x) \in o\big(g(x)\big)$ as $x \to 0$ iff $\limsup_{x\to 0} |f(x)/g(x)| < \infty$ and $f(x) \in \mathcal{O}\big(g(x)\big)$ iff $\lim_{x\to 0} f(x)/g(x) \to 0$. Thus we can introduce the following

Definition 7.1 - *Consider the sequence of functions* $\{\varphi_n(\varepsilon)\}_{n\in\mathbb{N}_0}$ *depending on the parameter* ε*; if*

$$\varphi_{n+1}(\varepsilon) \in o\big(\varphi_n(\varepsilon)\big) \quad \text{as } \varepsilon \to 0, \qquad \forall\, n \in \mathbb{N}_0, \tag{7.1}$$

then $\{\varphi_n(\varepsilon)\}_{n\in\mathbb{N}_0}$ *is said to be an* ***asymptotic sequence*** *as* $\varepsilon \to 0$*.*

A trivial example is the power sequence, i.e. $\varphi_n(\varepsilon) = \varepsilon^n$, and it'll be used frequently in what follows.

Consider now a function $f : \mathcal{I} \subset \mathbb{R} \to \mathbb{R}$ depending also on the parameter ε and assume that, as $\varepsilon \to 0$, there exists a positive integer N s.t. f can be approximated on $\mathcal{I}$ by the sum

$$f(x;\varepsilon) \xrightarrow{\varepsilon\to 0} \sum_{n=0}^{N} \varphi_n(\varepsilon) f_n(x) + o\big(\varphi_N(\varepsilon)\big). \tag{7.2}$$

Then, this expression is called an asymptotic approximation to $N+1$ terms of $f = f(x;\varepsilon)$ as $\varepsilon \to 0$ with respect to the asymptotic sequence of functions $\{\varphi_n(\varepsilon)\}_{n\in\mathbb{N}_0}$.

Definition 7.2 - *If equation (7.2) holds* $\forall\, N \in \mathbb{N}_0$*, then it's an* ***asymptotic expansion*** *of* $f = f(x;\varepsilon)$ *as* $\varepsilon \to 0$*. In particular, if it (doesn't) holds uniformly*[49] *on* $\mathcal{I}$*, it's said to be (non)uniformly valid.*

Note that convergent series expansions are asymptotic, but the converse is not true in general; asymptotic series are, in fact, usually divergent. Remark also that termwise differentiation and multiplication of asymptotic expansions is not always possible; however, this is not a big deal in our case, since we'll work with *asymptotic power sequences*, i.e. with asymptotic expansions of the form

$$f(x,\varepsilon) \xrightarrow{\varepsilon\to 0} \sum_{n=0}^{N} \varepsilon^n f_n(x) + \mathcal{O}(\varepsilon^{N+1}), \qquad \forall\, x \in \mathcal{I}. \tag{7.3}$$

In this particular case it can be proved that the above operations are admitted [JeK982].

Remark 7.1 – Asymptotic expansions occur frequently in Physics, where problems contain often one or more small parameters. For instance, consider a boundary–value problem depending on small parameter ε. If the equations are friendly workable for $\varepsilon = 0$ (which is called the *reduced problem*), then a solution $f = f(x;\varepsilon)$ might be obtained by means of an asymptotic expansion like the one in (A.3), where the first term $f_0 = f_0(x)$ is the solution of the reduced problem. When the result of this **perturbation analysis** approximates the solution of the original equation and it's uniformly valid in the range of the independent variable considered, then the approximation is called *asymptotic*.

[49] Recall that $\{f_n : \mathcal{S} \subseteq \mathbb{R} \to \mathbb{R}\}_{n\in\mathbb{N}}$ *converges uniformly* to $f : \mathcal{S} \to \mathbb{R}$ if $\forall\, \varepsilon \in \mathbb{R}^+$, $\exists\, N \in \mathbb{N}_0$ s.t. $\forall\, x \in \mathcal{S}$ and $\forall\, n \geq N$ it is $|f_n(x) - f(x)| < \varepsilon$.

DEFINITION 7.3 - *The perturbation problem is said to be* ***regular*** *iff the asymptotic expansion* (7.3) *converges uniformly on* $\mathcal{I}$ *as* $\varepsilon \to 0$ *or* ***singular*** *iff* $f = f(x;\varepsilon)$ *doesn't have a uniform limit in some regions of* $\mathcal{I}$*; these regions in which regular perturbations break down are called regions of non–uniformity.*

Remark that singular perturbation problems are very common in asymptotic analysis; indeed, it is true to say that they are the role rather than the exception in the theory.

Typical examples of perturbation problems are classified as follows [Nay973]:

- ▷ sources of non–uniformity appear in relation to an infinite domain, e.g. **secular terms**[50] of the form $x^n \cos x$, $x^n \sin x$ in nonlinear oscillations, which make $f_n(x)/f_{n-1}(x)$ unbounded as $x \to \infty$;
- ▷ a small parameter multiplies the highest–order derivative terms in a differential equation;
- ▷ there is a change of type of a PDE;
- ▷ the presence of singularities that doesn't manifest themselves in the exact solution but appear at a given order in the perturbative expansion and intensify when going to higher orders.

In what follows we'll deal only with *secular–type problems* and our main purpose will be to develop some basic techniques which eliminate non–uniformities (or singularities) from the perturbation problem, leading to uniformly valid approximate expressions. These perturbation techniques constitute what we shall call as **singular perturbation methods**. In particular, we'll use as a working laboratory the so–called *Duffing equation* (a well–known nonlinear oscillation equation) to show how singular perturbation techniques (i.e. the Poincaré and multiple–scale methods) do apply.

Before doing this, however, let's illustrate the regular perturbation procedure; in order to do this, assume ε to be a small real parameter and consider the following Cauchy problem

$$f''(t) + a(t)f'(t) + \big[b(t) + \varepsilon g(t)\big]f(t) = 0, \qquad f(0) = c_0, \quad f'(t)|_{t=0} = c_1. \tag{7.4}$$

Following the so–called *Poincaré approach*, introduce the power–series expansion

$$f(t;\varepsilon) = \sum_{n\in\mathbb{N}_0} \varepsilon^n f_n(t), \qquad \forall\, t \in \mathcal{I}, \tag{7.5}$$

where the f_n's are C^1 functions to be determined. Inserting equation (7.5) in (7.4) and setting to zero the coefficients belonging to the same powers of ε, we give rise to the infinite system of ODEs

$$\begin{cases} f_0''(t) + a(t)f_0'(t) + b(t)f_0(t) = 0, \\ f_1''(t) + a(t)f_1'(t) + b(t)f_1(t) = -g(t)f_0(t), \\ f_2''(t) + a(t)f_2'(t) + b(t)f_2(t) = -g(t)f_1(t), \\ \qquad\qquad \vdots \\ f_n''(t) + a(t)f_n'(t) + b(t)f_n(t) = -g(t)f_{n-1}(t), \\ \qquad\qquad \vdots \end{cases} \qquad \begin{aligned} f_0(0) &= c_0, \qquad f_0'(t)|_{t=0} = c_1, \\ f_n(0) &= 0 = f_n'(t)|_{t=0}, \ \ \forall\, n \in \mathbb{N}. \end{aligned} \tag{7.6}$$

Note that the series (7.5) converges uniformly on $\mathcal{I}$, thus allowing the termwise differentiation. The system (7.6) can now be solved iteratively, until one arrives to the desired number of terms.

[50] The term "*secular*" dates back to the early days of celestial mechanics, when these problems first appeared.

7.2 Singular perturbation methods: the Duffing equation. Now let's discuss how secular singularities arise in perturbative methods. Consider the following family of oscillation equation with a weak nonlinearity $F = F[f(t), f'(t)]$, whose intensity is determined by the parameter $\varepsilon \in \mathbb{R}_0^+ : \varepsilon \ll 1$:

$$f''(t) + f(t) = \varepsilon F[f(t), f'(t)]. \tag{7.7}$$

Assume also the frequency of the reduced problem (i.e. for $\varepsilon = 0$) to be normalized to one. A pedagogical example is given by the choice $F = -f^3$, corresponding to the so–called **Duffing equation**

$$f''(t) + f(t) = -\varepsilon f^3(t). \tag{7.8}$$

Following the Poincaré procedure above, introduce the power–series expansion (7.5) into (7.8) and equate the coefficients belonging to $\varepsilon^0, \varepsilon^1$ and ε^2, in order to get the following system

$$\begin{aligned} \mathcal{O}(\varepsilon^0) \quad &: \quad f_0''(t) + f_0(t) = 0, \\ \mathcal{O}(\varepsilon^1) \quad &: \quad f_1''(t) + f_1(t) = -f_0^3(t), \\ \mathcal{O}(\varepsilon^2) \quad &: \quad f_2''(t) + f_2(t) = -3f_0^2(t) f_1(t). \end{aligned} \tag{7.9}$$

The equation for the reduced problem is the usual one for the linear oscillator, whose general solution is given by $f_0(t) = a\cos(t+\phi)$, being[51] $\omega = 1$. Inserting f_0 into the problem of order $\mathcal{O}(\varepsilon)$ leads to $f_1''(t) + f_1(t) = -a^3\cos^3(t+\phi) = -\frac{1}{4}a^3\big[\cos\big(3(t+\phi)\big) + 3\cos(t+\phi)\big]$, whose general solution is[52]

$$f_1(t) = b\cos(t+\psi) - \tfrac{3}{8}a^3 t\sin(t+\phi) + \tfrac{1}{32}a^3\cos\big[3(t+\phi)\big],$$

having absorbed the term $\frac{1}{32}a^3\cos(t+\phi)$ into the characteristic solution. This quantity can in turn be eventually included into the zeroth–order solution, when b and ψ should the be determined from the initial conditions. The general perturbation solution up to orders higher than $\mathcal{O}(\varepsilon)$ is

$$f(t;\varepsilon) = a\cos(t+\phi) + \tfrac{1}{8}\varepsilon a^3\Big[\tfrac{1}{4}\cos\big[3(t+\phi)\big] - 3t\sin(t+\phi)\Big] + \mathcal{O}(\varepsilon^2). \tag{7.10}$$

Therefore, it is evident from equation (7.10) that f_1/f_0 diverges as $t \to \infty$, due to the emergence of the secular term $t\sin(t+\phi)$; also, if we proceed to a higher–order approximation, secular terms like $t^m\sin(t+\phi)$ or $t^m\cos(t+\phi)$ will appear, thus intensifying the divergence of the expansion.

The usefulness of the Duffing equation lies in the fact that it can be solved exactly. Indeed, multiplying equation (7.8) by f' and integrating once with respect to t, yields immediately $f'^2 + f^2 + \frac{1}{2}\varepsilon f^4 = 2E$ with E an integration constant. We obtain the implicit solution

$$t = \int \frac{\mathrm{d}\, f}{\big(2E - f^2 - \frac{1}{2}\varepsilon f^4\big)^{1/2}} \overset{f\equiv a\sin\theta}{=\!=\!=} \int \frac{\mathrm{d}\,\theta}{\big[1 + \varepsilon a^2\big(1 - \frac{1}{2}\cos^2\theta\big)\big]^{1/2}}, \qquad a^2 \equiv \frac{1}{\varepsilon}\big[(1+4\varepsilon E)^{1/2} - 1\big].$$

[51] Note that a and ϕ are constants to be determined after an appropriate choice of the initial conditions. However in this section we shall not impose any initial condition since we're concerned with general properties of solutions.

[52] This differential equation can be solved in various ways (by means of e.g. the similarity technique or the Lagrange method). Here we present an alternative powerful approach proposed by F. G. Tricomi (1897–1978). Introduce the linear operator $\mathscr{L} : \mathrm{C}^n(\mathcal{I}) \to \mathrm{C}^0(\mathcal{I})$ and consider then the non–homogeneous n–th order linear ODE

$$[\mathscr{L}y](x) \equiv \Big[\tfrac{\mathrm{d}^n}{\mathrm{d}\,x^n} + a_1(x)\tfrac{\mathrm{d}^{n-1}}{\mathrm{d}\,x^{n-1}} + \cdots + a_{n-1}(x)\tfrac{\mathrm{d}}{\mathrm{d}\,x} + a_n(x)\Big]y(x) = g(x).$$

Let $y_c = y_c(x)$ be its complementary solution (i.e. the solution of $[\mathscr{L}y](x) = 0$); thus the general solution is [Tri961]

$$y(x) = y_c(x) + \int_0^x g(\xi)\mathcal{Y}(x-\xi)\,\mathrm{d}\,\xi, \qquad \mathcal{Y}(x) := \begin{cases} y_c(0) = y_c'(x)|_{t=0} = y_c''(x)|_{x=0} = \cdots = y_c^{(n-2)}(x)|_{x=0} = 0, \\ y_c^{(n-1)}(x)|_{x=0} = 1. \end{cases}$$

Let's apply this method to the case under scrutiny, i.e. $f_1'' + f_1 = -a^3\cos^3 x$ with $x \equiv t + \phi$. The characteristic solution is given by $f_{1\,c}(x) = c_1 e^{-\imath x} + c_2 e^{\imath x}$ so that $\mathcal{Y}_1(x) = \frac{1}{2\imath}(e^{\imath x} - e^{-\imath x}) = \sin x$ and thus the particular solution is

$$f_{1\,p}(x) = \int_0^x a^3\cos^3\xi\sin(\xi - x)\,\mathrm{d}\,\xi = \tfrac{1}{32}a^3\big(\cos 3x - 12x\sin x - \cos x\big).$$

The integral above is an elliptic one, thus f is periodic in t and the period of oscillation is given by

$$\mathcal{T} = 4\int_0^{\pi/2} \frac{\mathrm{d}\,\theta}{\left[1+\varepsilon a^2\left(1-\frac{1}{2}\cos^2\theta\right)\right]^{1/2}} = 4\int_0^{a} \frac{\mathrm{d}\,f}{\left(2E - f^2 - \frac{1}{2}\varepsilon f^4\right)^{1/2}}. \tag{7.11}$$

Since εa^2 is small, we can expand in Taylor series the function in (7.11), yielding

$$\mathcal{T} = 4\int_0^{\pi/2} \left[1 - \tfrac{1}{2}\varepsilon a^2(1-\tfrac{1}{2}\cos^2\theta) + \mathcal{O}(\varepsilon^2 a^4)\right]\mathrm{d}\,\theta = 2\pi\left[1 - \tfrac{3}{8}\varepsilon a^2 + \mathcal{O}(\varepsilon^2 a^4)\right]. \tag{7.12}$$

Therefore, defining with $\Omega = 2\pi\mathcal{T}$ the angular frequency of the nonlinear oscillator, one has

$$\Omega = 1 + \tfrac{3}{8}\varepsilon a^2 + \mathcal{O}\big(\varepsilon^2 a^4\big).$$

The above approach suggests that, providing $\varepsilon a^2 \ll 1$, the period $\mathcal{T}$ is substantially independent of the amplitude and therefore it seems reasonable to construct a perturbation solution in terms of periodic functions, each with linearized angular frequency of 1. However, for large t, the phase–difference between the exact phase Ωt and the one of the linearized oscillator, i.e. $1t$, becomes very large even when Ω is very close to unity. This is the reason why the perturbation expansion based on the periodic solution of the reduced problem fails, giving rise to the secular term $t\sin(t+\phi)$ in (7.10). One may therefore think to construct a secular–free solution by means of a perturbation approach based on the exact period (7.11) or on the improved one given by the equation (7.12). These observations point out the role that independent coordinates have in the search for a uniform expansion; indeed, the secularity may be eliminated by using an asymptotic expansion in terms of a new time–coordinate $\tau = \tau(t;\varepsilon)$ so that

$$f(t;\varepsilon) = \sum_{n\in\mathbb{N}_0} \varepsilon^n f_n[\tau(t;\varepsilon)], \qquad \tau = \big(1 + \tfrac{3}{8}\varepsilon a^2\big)t + \phi. \tag{7.13}$$

In terms of the scaled time–variable τ, the equation (A.10) can be written in the secular–free form

$$f(t;\varepsilon) = a\cos\tau + \tfrac{1}{32}\varepsilon a^3\cos 3\tau + \mathcal{O}(\varepsilon^2), \tag{7.14}$$

since $\cos\tau = \cos(t+\phi+\frac{3}{8}\varepsilon a^2 t) = \cos(t+\phi) + \frac{3}{8}\varepsilon a^2 t\sin(t+\phi) + \mathcal{O}(\varepsilon^2)$ and $\cos 3\tau = \varepsilon\cos[3(t+\phi)] + \mathcal{O}(\varepsilon^2)$, begin $0 < \varepsilon a^2 \ll 1$. The new coordinate $\tau = \tau(t;\varepsilon)$ is thus called the *optimal coordinate*. This result indicates that an appropriate transformation of independent variables is useful for rendering perturbation expansions uniformly valid. The following two sections are devoted in the description of two basic techniques, based on scaling transforms of the independent variables, that are often used in singular perturbation theory, i.e. the *Poincaré* and the *multiple–scale methods*.

7.3 The Poincaré method. In this method the solution is expanded in the form of power series in ε in order to remove perturbation's secularities. Recalling that the angular frequency of the linearized equation is normalized to unity, assume the following expansion

$$f(t;\varepsilon) = \sum_{n\in\mathbb{N}_0} \varepsilon^n f_n[\tau(t;\varepsilon)], \qquad \tau(t;\varepsilon) \equiv \omega(\varepsilon)t, \quad \omega(\varepsilon) := 1 + \varepsilon\omega_1 + \varepsilon^2\omega_2 + \cdots. \tag{7.15}$$

The dependence of τ on ε is thus determined by forcing the *non–secularity conditions*, i.e. by setting to zero the coefficients belonging to the secular terms of the form $t^m\cos t$ and $t^m\sin t$.

To show how this method works, insert the expressions (7.15) into (7.7) in order to obtain

$$\sum_{n\in\mathbb{N}_0} \varepsilon^n f_n''(\tau)\Big(\sum_{m\in\mathbb{N}_0}\varepsilon^m\omega_m\Big)^2 + \sum_{n\in\mathbb{N}_0}\varepsilon^n f_n(\tau) = \varepsilon F\Big[\sum_{n\in\mathbb{N}_0}\varepsilon^n f_n(\tau), \sum_{n,m\in\mathbb{N}_0}\varepsilon^{n+m}\omega_m f_n'(\tau)\Big],$$

with $\omega_0 = 1$. Equating the coefficients belonging to same powers of ε, one finds the system

$$\begin{aligned} \mathcal{O}(\varepsilon^0) \quad &: \quad f_0''(\tau) + f_0(\tau) = 0, \\ \mathcal{O}(\varepsilon^1) \quad &: \quad f_1''(\tau) + f_1(\tau) = -2\omega_1 f_0''(\tau) + F\big[f_0(\tau), f_0'(\tau)\big]. \end{aligned}$$

Inserting the general solution of the reduced problem $f_0 = a\cos(\tau+\phi)$ into the 1st order equation yields

$$f_1''(\tau) + f_1(\tau) = -2a\omega_1\cos(\theta) + F\big[a\cos\theta, -a\sin\theta\big], \tag{7.16}$$

being $\theta \equiv \tau + \phi$. Therefore, if the r.h.s. contains terms proportional to $\cos\theta$ or $\sin\theta$, then singularities must occur and no periodic solutions exist. However, those singularities can now be eliminated by choosing suitable values for a and ω_1. To do so, let's perform a Fourier series expansion of F so that

$$F\big[a\cos\theta, -a\sin\theta\big] = \frac{\mathcal{A}_0}{2} + \sum_{n\in\mathbb{N}}\big[\mathcal{A}_n(a)\cos n\theta + \mathcal{B}_n(a)\sin n\theta\big],$$

$$\mathcal{A}_n(a) := \frac{1}{\pi}\int_0^{2\pi} F[a\cos\theta, -a\sin\theta]\cos n\theta\,\mathrm{d}\,\theta, \quad \mathcal{B}_n(a) := \frac{1}{\pi}\int_0^{2\pi} F[a\cos\theta, -a\sin\theta]\sin n\theta\,\mathrm{d}\,\theta.$$

Inserting the above expansion into (7.16), it follows that the secularities can be removed by forcing the conditions $2\omega_1 a + \mathcal{A}_1(a) = 0$ and $\mathcal{B}_1(a) = 0$, which gives us the relations for ω_1 and a. In fact, having in mind that $\tau = \omega t = \big[1 - \frac{\varepsilon}{2a}\mathcal{A}_1(a) + \mathcal{O}(\varepsilon^2)\big]t$, we obtain the lowest–order approximate solution

$$f(t;\varepsilon) = a\cos\left[\left(1 - \tfrac{\varepsilon}{2a}\mathcal{A}_1(a)\right)t + \phi\right], \tag{7.17}$$

which is correct up to $\mathcal{O}(\varepsilon)$. In the case of the Duffing equation, one has $F[f, f'] = -f^3$ so that $F[a\cos\theta, -a\sin\theta] = -a^3\cos^3\theta$ and the $\mathcal{O}(\varepsilon)$ equation (7.16) takes the explicit form

$$f_1'' + f_1 = -\tfrac{1}{4}a^3\cos 3\theta + \left(2\omega_1 - \tfrac{3}{4}a^2\right)a\cos\theta. \tag{7.18}$$

The non–secularity condition is $\omega_1 = \frac{3}{8}a^2$ so that $f_1'' + f_1 = -\frac{1}{4}a^3\cos 3\theta$, whose particular solution is $f_{1_p}(\theta) = \frac{1}{32}a^3(\cos 3\theta - \cos\theta)$. Since the second term in the r.h.s. can be absorbed into the general solution, we have $f_{1_p}(\theta) = \frac{1}{32}a^3\cos 3\theta$ and the perturbative solution valid up to $\mathcal{O}(\varepsilon^2)$ terms is

$$f = a\cos\big(\omega(\varepsilon)t + \phi\big) + \tfrac{1}{32}\varepsilon a^3\cos\big[3(\omega(\varepsilon)t + \phi)\big] + \mathcal{O}(\varepsilon^2), \qquad \omega(\varepsilon) = 1 + \tfrac{3}{8}\varepsilon a^2 + \mathcal{O}(\varepsilon^2). \tag{7.19}$$

We have then reproduced our previous result given in equations (7.13) and (7.14). Note that the analysis presented here can be straightforwardly extended to an arbitrary (but still finite) order $\mathcal{O}(\varepsilon^n)$, arriving this way at a higher–order perturbative analysis of the solution.

7.4 Multiple–scale method. In the literature there is a variety of versions of the so–called multiple–scale method; here we discuss a procedure known as the *derivative expansion approach*. First, let's expand the independent variable t to many variables, each with a different scale, by introducing the set $\{t_n \equiv \varepsilon^n t\}_{n=0,1,\ldots,N}$; now expand the dependent variable f into an asymptotic series of the form

$$f(t;\varepsilon) = \sum_{n=0}^{N}\varepsilon^n f_n(t_0, t_1, \ldots, t_N) + \mathcal{O}\big(\varepsilon^{N+1}\big), \tag{7.20}$$

having assumed that each function f_n depends on the multiple–scales $\{t_n\}_{n=0,1,\ldots,N}$. Note that

$$\frac{\mathrm{d}}{\mathrm{d}\,t} = \sum_{n=0}^{N}\varepsilon^n\frac{\partial}{\partial t_n},$$

which is the reason why this approach has the name it has. Inserting (7.20) into (7.8), yields

$$\left(\sum_{m=0}^{N}\varepsilon^n\partial_{t^n}\right)^2\sum_{n=0}^{N}\varepsilon^n f_n(t_0, t_1, \ldots, t_N) + \sum_{n=0}^{N}\varepsilon^n f_n(t_0, t_1, \ldots, t_N) = -\varepsilon\left(\sum_{n=0}^{N}\varepsilon^n f_n(t_0, t_1, \ldots, t_N)\right)^3.$$

Equating to zero the coefficients belonging to the same powers of ε, gives the first two equations

$$\begin{aligned} \mathcal{O}(\varepsilon^0) \quad &: \quad \partial^2_{t_0} f_0 + f_0 = 0, \\ \mathcal{O}(\varepsilon^1) \quad &: \quad \partial^2_{t_0} f_1 + f_1 = -2\partial_{t_0}\partial_{t_1} f_0 - f_0^3. \end{aligned} \tag{7.21}$$

The general solution of the first equation is the usual one of the linear oscillator, namely

$$f_0(t_0, t_1, \ldots, t_N) = a(t_1, t_2, \ldots, t_N) \cos\big[t_0 + \Phi(t_1, t_2, \ldots, t_N)\big], \tag{7.22}$$

where $a = a(t_1, \ldots, t_N)$ and $\Phi = \Phi(t_1, \ldots, t_N)$ are functions that might depend on the "*slow*" variables $t_1, t_2, \ldots, t_N$ and have to be determined. Inserting (7.22) into the second equation in (7.21) gives

$$\partial^2_{t_0} f_1 + f_1 = 2a_{t_1} \sin(t_0 + \Phi) + \Big(2a\Phi_{t_1} - \tfrac{3}{4}a^2\Big) a \cos(t_0 + \Phi) - \tfrac{1}{4}a^3 \cos\big[3(t_0 + \Phi)\big],$$

and the non–secularity conditions for f_1 become now $\partial_{t_1} a = 0$ and $\partial_{t_1}\Phi = \frac{3}{8}a^2$, that is

$$\Phi(t_1, t_2, \ldots, t_N) = \tfrac{3}{8} t_1 a(t_2, \ldots, t_N) + \phi(t_2, \ldots, t_N), \tag{7.23}$$

having noted that $a = a(t_2, \ldots, t_N)$. Therefore, if we are interested into solutions that are valid up to $\mathcal{O}(\varepsilon)$ orders, one can assume a and ϕ as constants. The secular–free equation is then $\partial^2_{t_0} f_1 + f_1 = -\frac{1}{4}\cos 3\theta$ with $\theta \equiv t_0 + \Phi$, which means that the uniformly valid solution up to $\mathcal{O}(\varepsilon)$ is given by

$$f = a\cos\theta + \frac{1}{32}\varepsilon a^3 \cos 3\theta + \mathcal{O}(\varepsilon^2), \qquad \theta \equiv \Big(1 + \tfrac{3}{8}\varepsilon a^2\Big) t + \phi. \tag{7.24}$$

Thus the same result obtained in (7.14) and (7.19) is again reproduced. Note that the main idea behind this method is to extend the independent variable to many variables, including the slowly varying ones; such an increase in the number of degrees of freedom is then used to make the perturbation uniformly valid. Since this method is systematic, it is a straightforward matter to proceed to higher order perturbations and this is why we'll use it in the following applications in Nonlinear Theory.

7.5 Multiple–scale approach to NLEEs. The previous subsections were concerned with the study of perturbative solutions of nonlinear oscillation problems and so they involved only NLODEs. It turns out that the essential ideas we've presented are equally useful in the study of PDEs and, in particular, of *nonlinear dispersive evolution equations* (NLDEEs). Indeed, some additional ideas will be needed for such problems. The aim of this paragraph it to show, by means of the perturbative approach, the emergence of a "universal" property of some NLEEs within the class of the NLDEEs. In particular, we'll see how the differential expansion method introduced in §7.4 leads to the NLS equation, in presence of wave modulation of quasi–monochromatic waves [Kaw973], and to the KdV within the long–wave approximation of the Boussinesq equation, which is a representative model for the family of NLDEEs.

Before entering the discussion, we'll need to recall some properties of NLDEEs. Let's therefore consider a simple but paradigmatic model described by the so–called *Boussinesq–type equation*

$$u_{tt} - c^2 u_{xx} - \mu u_{ttxx} = \tfrac{1}{2}\big(u^2\big)_{xx}, \tag{7.25}$$

where c is the phase velocity in the limit of long–wavelength and μ is a dispersive coefficient.

We search for harmonic wave–type solutions of (7.25) having the form $u(x,t) = Ae^{\imath(kx-\omega t)} + \text{c.c.}$, where k is the wave number, ω is the frequency, A is the complex–valued amplitude and c.c. denotes the complex conjugate of the preceding term. The latter solves the linearized Boussinesq–type equation $u_{tt} - c^2 u_{xx} - \mu u_{ttxx} = 0$ iff the following *dispersive relation* holds:

$$\mathscr{D}(k;\omega) = c^2k^2 - \omega^2 - \mu k^2\omega^2 \equiv 0, \qquad k \in \mathbb{R}\setminus\{0\}. \tag{7.26}$$

REMARK 7.1 – Recall that, for an arbitrary linear dispersive equation, the relation $\mathscr{D}(k;\omega) \equiv 0$ may have several roots for ω, when we may write $\omega = \mathscr{W}_l(k)$ with $l = 1, 2, \ldots, m$, where the functions $\mathscr{W}$ so

defined can be complex–valued functions of k. For real roots we'll say that the relation $\omega = \mathscr{W}(k)$ is *purely dispersive*, where the term "dispersive" is used iff the condition $\mathscr{W}''(k) \not\equiv 0$ is satisfied. E.g., the dispersive relation (7.26) for the linearized Boussinesq–type equation has two distinct roots

$$\omega = \pm \frac{ck}{(1+\mu k^2)^{1/2}}, \qquad k \in \mathbb{R} \setminus \{0\},$$

so that it is purely dispersive, since $\mathscr{W}''(k) \not\equiv 0 \; \forall\, k \in \mathbb{R} \setminus \{0\}$ and its roots are both real.

Recall that a solution of the linearized equation can be obtained by means of a superposition of Fourier integrals, whose number is determined by the number of the roots of $\mathscr{D}(k;\omega) \equiv 0$. In particular, for the linearized Boussinesq–type equation one has $\omega = \pm\mathscr{W}(k)$, with $\mathscr{W}(k) = ck(1+\mu k^2)^{-1/2}$ so that

$$u(x,t) = \int_{\mathbb{R}} f_1(k) e^{-\imath \chi_1(k) t}\, \mathrm{d}\, k + \int_{\mathbb{R}} f_2(k) e^{-\imath \chi_2(k) t}\, \mathrm{d}\, k, \qquad \chi_{1,2}(k) \equiv \pm \mathscr{W}(k) - k\frac{x}{t}.$$

In what follows, we'll be interested in the *far field limit*, i.e. $|t| \to +\infty$ with x/t kept fixed, representing the case of progressive waves moving with fixed velocity x/t. Within this limit, the integral above can be solved by means of the *stationary phase* method, thus obtaining the asymptotic behavior of linear dispersive waves [JeK982].

7.5.1 Nonlinear wave modulation and the NLS equation. In order to apply the multiple–scale method to the equation (7.25), let's rewrite it in the following form

$$\mathscr{L}(\partial_x, \partial_t)[u(x,t)] = \mathscr{N}(\partial_x, \partial_t)\big[u^2(x,t)\big], \qquad \begin{aligned} \mathscr{L}(\partial_x, \partial_t) &\equiv \partial_t^2 - c^2\partial_x^2 - \mu \partial_x^2 \partial_t^2, \\ \mathscr{N}(\partial_x, \partial_t) &\equiv \tfrac{1}{2}\partial_x^2. \end{aligned} \tag{7.27}$$

In the derivative expansion method, the number of independent variables is extended to include also slowly–varying variables. Applying the procedure to x and t, we introduce the set of multiple–scales

$$\big\{(x_n, t_n) \equiv (\varepsilon^n x, \varepsilon^n t)\big\}_{n=0,1,\ldots,M},$$

being ε a parameter characterizing the smallness of the associated terms. Accordingly, the dependent variable must be a function of the new variables, that is $u = u(\{x_n\}_{n=0,1,\ldots,M}, \{t_n\}_{n=0,1,\ldots,M})$. Furthermore, we assume that the dependent variable has an asymptotic representation of the form[53]

$$u(x_0, x_1, \ldots, x_M; t_0, t_1, \ldots, t_M; \varepsilon) = \sum_{m=1}^{M} \varepsilon^m u_m(x_0, x_1, \ldots, x_M; t_0, t_1, \ldots, t_M) + \mathcal{O}(\varepsilon^{M+1}). \tag{7.28}$$

It's clear that, after the multiple–scaling, the operators $\mathscr{L}$ and $\mathscr{N}$ transform simultaneously as

$$\mathscr{L}(\partial_x, \partial_t) = \sum_{n=0}^{M} \varepsilon^n \mathscr{L}_n\big(\partial_{x_0}, \ldots, \partial_{x_M}; \partial_{t_0}, \ldots, \partial_{t_M}\big) + \mathcal{O}\big(\varepsilon^{M+1}\big), \tag{7.29a}$$

$$\mathscr{N}(\partial_x, \partial_t) = \sum_{n=0}^{M} \varepsilon^n \mathscr{N}_n\big(\partial_{x_0}, \ldots, \partial_{x_M}; \partial_{t_0}, \ldots, \partial_{t_M}\big) + \mathcal{O}\big(\varepsilon^{M+1}\big). \tag{7.29b}$$

On the other hand we know that $\partial_x = \sum_{n=0}^{M} \varepsilon^n \partial_{x_n}$ and $\partial_t = \sum_{n=0}^{M} \varepsilon^n \partial_{t_n}$ so that, inserting the last

[53] In general $u = u(x,t)$ can be expanded in terms of another small parameter which measures the weakness of the nonlinearity of the wave; for the sake of simplicity, however, here we expand u as an asymptotic power series of ε.

transformation into the original definitions of $\mathscr{L}$ and $\mathscr{N}$ given in (7.27), one finds

$$\begin{aligned}
\mathcal{O}(\varepsilon^0) \quad &: \quad \begin{cases} \mathscr{L}_0 = \partial_{t_0}^2 - c^2\partial_{x_0}^2 - \mu\partial_{x_0}^2\partial_{t_0}^2, \\ \mathscr{N}_0 = \frac{1}{2}\partial_{x_0}^2, \end{cases} \\
\mathcal{O}(\varepsilon^1) \quad &: \quad \begin{cases} \mathscr{L}_1 = 2\partial_{t_0}\partial_{t_1} - 2c^2\partial_{x_0}\partial_{x_1} - 2\mu\big(\partial_{x_0}\partial_{x_1}\partial_{t_0}^2 + \partial_{t_0}\partial_{t_1}\partial_{x_0}^2\big), \\ \mathscr{N}_1 = \partial_{x_0}\partial_{x_1}, \end{cases} \\
\mathcal{O}(\varepsilon^2) \quad &: \quad \begin{cases} \mathscr{L}_2 = \partial_{t_1}^2 + 2\partial_{t_0}\partial_{t_2} - c^2\big(\partial_{x_1}^2 + 2\partial_{x_0}\partial_{x_2}\big) + \\ \qquad - \mu\big(\partial_{t_0}^2\partial_{x_1}^2 + \partial_{t_1}^2\partial_{x_0}^2 + 4\partial_{t_0}\partial_{t_1}\partial_{x_0}\partial_{x_1} + 2\partial_{t_0}^2\partial_{x_0}\partial_{x_2} + 2\partial_{t_0}\partial_{t_2}\partial_{x_0}^2\big), \\ \mathscr{N}_2 = \frac{1}{2}\partial_{x_1}^2 + \partial_{x_0}\partial_{x_2}, \end{cases}
\end{aligned}$$

Using the operators above, we obtain the first four perturbation equation in the following form

$$\mathcal{O}(\varepsilon^1) \quad : \quad \mathscr{L}_0[u_1] = 0, \tag{7.30a}$$

$$\mathcal{O}(\varepsilon^2) \quad : \quad \mathscr{L}_0[u_2] + \mathscr{L}_1[u_1] = \mathscr{N}_0[u_1^2], \tag{7.30b}$$

$$\mathcal{O}(\varepsilon^3) \quad : \quad \mathscr{L}_0[u_3] + \mathscr{L}_1[u_2] + \mathscr{L}_2[u_1] = \mathscr{N}_0[2u_1u_2] + \mathscr{N}_1[u_1^2], \tag{7.30c}$$

$$\mathcal{O}(\varepsilon^4) \quad : \quad \mathscr{L}_0[u_4] + \mathscr{L}_1[u_3] + \mathscr{L}_2[u_2] + \mathscr{L}_3[u_1] = \mathscr{N}_0[u_2^2 + 2u_1u_3] + \mathscr{N}_1[2u_1u_2] + \mathscr{N}_2[u_1^2]. \tag{7.30d}$$

Being interested in nonlinear wave modulation, assume that (7.30a) has a solution of the form

$$u_1(x_0, x_1, \ldots, x_M; t_0, t_1, \ldots, t_M) = \mathscr{A}(x_1, \ldots, x_M; t_1, \ldots, t_M)e^{\imath\theta} + \text{c.c.}, \qquad \theta \equiv kx_0 - \omega t_0. \tag{7.31}$$

which is nothing more than the harmonic wave–type solution of the linearized Boussinesq–type equation (7.25). Therefore, the wave number k and the frequency ω are not arbitrary, but must satisfy the dispersion relation (7.26). Since we wish to insert the solution (7.31) into (7.30b), let's first note that

$$\mathscr{L}_1[u_1] = -\imath\big(\mathscr{A}_{t_1}\mathscr{D}_\omega - \mathscr{A}_{x_1}\mathscr{D}_k\big)e^{\imath\theta} + \text{c.c.}, \qquad \mathscr{N}_0\big[u_1^2\big] = -2k^2A^2e^{2\imath\theta} + \text{c.c.},$$

where we have used the following facts: $\mathscr{D}_\omega(k;\omega) = -2\omega(1+\mu k^2)$, $\mathscr{D}_k(k;\omega) = 2k(c^2 - \mu\omega^2)$ and $u_1^2 = \mathscr{A}^2e^{2\imath\theta} + \mathscr{A}^{*2}e^{-2\imath\theta} + 2|\mathscr{A}|^2$. Thus, the second–order perturbation equation (7.30b) becomes

$$\mathscr{L}_0[u_2] = \imath\big(\mathscr{A}_{t_1}\mathscr{D}_\omega - \mathscr{A}_{x_1}\mathscr{D}_k\big)e^{\imath\theta} - 2k^2\mathscr{A}^2e^{2\imath\theta} + \text{c.c.}. \tag{7.32}$$

Imposing now the non–secularity condition, i.e. the coefficient of $e^{\imath\theta}$ vanishes, gives us the equation

$$\mathscr{A}_{t_1}\mathscr{D}_\omega = \mathscr{A}_{x_1}\mathscr{D}_k \qquad \Longleftrightarrow \qquad \frac{\mathscr{A}_{t_1}}{\mathscr{A}_{x_1}} = \frac{\mathscr{D}_k}{\mathscr{D}_\omega} = -v_g, \tag{7.33}$$

where v_g is the group–velocity[54] of the wave–train (7.31). If (7.33) and its complex conjugate are satisfied, then (7.32) gets the form $\mathscr{L}_0[u_2] = -2k^2\mathscr{A}^2e^{2\imath\theta} + \text{c.c.}$ and admits the uniformly valid solution

$$\begin{aligned}
u_2(x_0, x_1, \ldots, x_M; t_0, t_1, \ldots, t_M) = -\frac{2k}{\mathscr{D}(2k;2\omega)}\mathscr{A}^2(x_1, \ldots, x_M; t_1, \ldots, t_M)e^{2\imath\theta} + \\
+ \mathscr{E}(x_1, \ldots, x_M; t_1, \ldots, t_M)e^{\imath\theta} + \text{c.c.} + \mathscr{F}(x_1, \ldots, x_M; t_1, \ldots, t_M),
\end{aligned} \tag{7.34}$$

where $\mathscr{E}$ and $\mathscr{F}$ are respectively complex–valued and real–valued functions of higher–order scaled to be determined in higher–order perturbations[55]. Note that $\mathscr{E}$ can be absorbed into u_1 by redefining $\mathscr{A}$.

[54] Equation (7.33) follows immediately by differentiating both sides of equation (7.26) squared, that is

$$2\omega\,\mathrm{d}\,\omega = 2c^2\left(\frac{k}{1+\mu k^2} - \frac{2\mu k^3}{(1+\mu k^2)^2}\right)\mathrm{d}\,k = \frac{2k(c^2-\mu\omega^2)}{1+\mu k^2}\,\mathrm{d}\,k \qquad \Longrightarrow \qquad v_g := \frac{\mathrm{d}\,\omega}{\mathrm{d}\,k} = \frac{k(c^2-\mu\omega^2)}{\omega(1+\mu k^2)}.$$

[55] Note that the solution (7.34) has been obtained by searching for a particular solution of the non–secular second–order perturbative equation (7.32) of the form $u_2 = \alpha e^{2\imath\theta} + \mathscr{E}e^{\imath\theta} + \text{c.c.} + \mathscr{F}$, where $\mathscr{E}$ is complex–valued and $\mathscr{F}$ real.

To study the next–order equation, we need to insert the relations (7.34) and (7.31) into (7.30c). With this aim, let's calculate $\mathscr{L}_1[u_2]$, $\mathscr{L}_2[u_1]$, $\mathscr{N}_0[2u_1u_2]$ and $\mathscr{N}_1[u_1^2]$: after some calculation, one has

$$\begin{aligned}
\mathscr{L}_1\,[u_2] &= \frac{\imath k^2}{\mathscr{D}_2}\big(\mathscr{D}_{2k}\mathscr{A}_{x_1} - \mathscr{D}_{2\omega}\mathscr{A}_{t_1}\big)e^{2\imath\theta} - \imath\big(\mathscr{D}_k\mathscr{E}_{x_1} - \mathscr{D}_\omega\mathscr{E}_{t_1}\big)e^{\imath\theta} + \text{c.c.},\\
\mathscr{L}_2\,[u_1] &= -\frac{1}{2}\big(\mathscr{D}_{\omega\omega}\mathscr{A}_{t_1t_1} + \mathscr{D}_{kk}\mathscr{A}_{x_1x_1} - 2\mathscr{D}_{k\omega}\mathscr{A}_{t_1x_1}\big)e^{\imath\theta} - \imath\big(\mathscr{D}_k\mathscr{A}_{x_2} - \mathscr{D}_\omega\mathscr{A}_{t_2}\big)e^{\imath\theta} + \text{c.c.},\\
\mathscr{N}_0\,[2u_1u_2] &= \frac{18k^4}{\mathscr{D}_2}\mathscr{A}^3e^{3\imath\theta} - 4k^2\mathscr{A}\mathscr{E}e^{2\imath\theta} - k^2\left(\mathscr{A}\mathscr{F} - \frac{2k^2}{\mathscr{D}_2}\mathscr{A}^2\mathscr{A}\right)e^{\imath\theta} + \text{c.c.},\\
\mathscr{N}_1\,[u_1^2] &= 4\imath k\mathscr{A}\mathscr{A}_{x_1}e^{2\imath\theta} + \text{c.c.},
\end{aligned}$$

where we have introduced the new function $\mathscr{D}_2(k;\omega) \equiv \mathscr{D}(2k;2\omega)$ in order to simplify the notation and noted that the product of the two functions u_1 and u_2 takes the form

$$u_1u_2 = -\frac{2k^2}{\mathscr{D}_2}\mathscr{A}^3e^{3\imath\theta} + \mathscr{A}\mathscr{E}e^{2\imath\theta} + \left(\mathscr{A}\mathscr{F} - \frac{2k^2}{\mathscr{D}_2}\mathscr{A}^2\mathscr{A}^*\right)e^{\imath\theta} + \mathscr{A}^*\mathscr{E} + \text{c.c.},$$

in the third relation. The non–secularity condition is more involved than before and is given by

$$\begin{aligned}
\imath\Big(\mathscr{D}_k\mathscr{A}_{x_2} - \mathscr{D}_\omega\mathscr{A}_{t_2}\Big) + \tfrac{1}{2}\Big(\mathscr{D}_{\omega\omega}\mathscr{A}_{t_1t_1} - 2\mathscr{D}_{\omega k}\mathscr{A}_{t_1x_1} + \mathscr{D}_{kk}\mathscr{A}_{x_1x_1}\Big) +\\
+ k^2\left(\frac{2k^2}{\mathscr{D}_2}\mathscr{A}^2\mathscr{A}^* - \mathscr{A}\mathscr{F}\right) + \imath\big(\mathscr{D}_k\mathscr{E}_{x_1} - \mathscr{D}_\omega\mathscr{E}_{t_1}\big) = 0,
\end{aligned} \tag{7.35}$$

and its complex conjugate relation. The function $\mathscr{F}$ can be determined by means of the non–secularity condition for constant terms in the $\mathcal{O}(\varepsilon^4)$ problem. The calculations show that the terms which contribute are $\mathscr{L}_2[u_2]$ with $(\partial_{t_1}^2 - c^2\partial_{x_1}^2)\mathscr{F}$ and $\mathscr{N}_2[u_1^2]$ with $\partial_{x_1}^2|\mathscr{A}|^2$; therefore, $\mathscr{F}$ satisfies the equation

$$\big(\partial_{t_1}^2 - c^2\partial_{x_1}^2\big)\mathscr{F} = \partial_{x_1}^2|\mathscr{A}|^2. \tag{7.36}$$

If we assume that $\mathscr{A}$ and $\mathscr{F}$ depend on x_1, t_1 only through $\xi \equiv x_1 - v_gt_1$, i.e. if they are considered in a coordinate system moving with v_g, one can integrate twice $(v_g^2 - c^2)\partial_\xi^2\mathscr{F} = \partial_\xi^2|\mathscr{A}|^2$, so that

$$\mathscr{F}(\xi, x_2, \ldots, x_M; t_2, \ldots, t_M) = \frac{1}{v_g^2 - c^2}\big|\mathscr{A}\big|^2(x_1, \ldots, x_M; t_1, \ldots, t_M) + \beta(x_2, \ldots, x_M; t_2, \ldots, t_M), \tag{7.37}$$

where the term $\alpha(x_2, \ldots, x_M; t_2, \ldots, t_M)\xi$ has been absorbed in the expression[56] for $\mathscr{F}$. Note that the function $\beta = \beta(x_2, \ldots, x_M; t_2, \ldots, t_M)$ is determined by choosing suitable boundary or initial conditions.

In order to obtain a closed expression for $\mathscr{A}$, we need to determine $\mathscr{E}$ in equation (7.34). It can be demonstrated that this quantity can be transferred to the lowest–order solution (7.31) and the transferred quantity can then be regarded as a new expression for $\mathscr{A}$, which still satisfies equation (7.34). Thus we may drop the last term in (7.34) and we can rewrite it in the form

$$\imath\mathscr{D}_\omega\Big(\mathscr{A}_{t_2} - v_g\mathscr{A}_{x_2}\Big) - \tfrac{1}{2}\Big(\mathscr{D}_{kk} + 2v_g\mathscr{D}_{k\omega} + v_g^2\mathscr{D}_{\omega\omega}\Big)\mathscr{A}_{\xi\xi} + k^2\left[\left(\frac{1}{v_g^2 - c^2} - \frac{2k^2}{\mathscr{D}_2}\right)\big|\mathscr{A}\big|^2\mathscr{A} + \beta\mathscr{A}\right] = 0. \tag{7.38}$$

Equation (7.38) can be further simplified by means of the following observations:

$$\begin{aligned}
\frac{\mathrm{d}\,v_g}{\mathrm{d}\,k} = \frac{(c^2 - \mu\omega^2) - 2\mu k\omega v_g}{\omega(1 + \mu k^2)} - \frac{k(c^2 - \mu\omega^2)}{\omega^2(1 + \mu k^2)}\big[v_g(1 + \mu k^2) + 2\mu k\omega\big] = -\frac{1}{\mathscr{D}_\omega}\Big(\mathscr{D}_{kk} + 2v_g\mathscr{D}_{\omega k} + v_g^2\mathscr{D}_{\omega\omega}\Big),\\
\frac{1}{2k^2}\mathscr{D}_2(\omega; k) = \frac{2}{k^2}(c^2k^2 - \omega^2 - \mu k^2\omega^2) - 6\mu\omega^2 = \mathscr{D}(\omega; k) - 6\mu\omega^2 = -6\mu\omega^2,
\end{aligned}$$

[56] When $v_g \to c$ the resonant interaction between short waves and long waves becomes important and has to be treated separately. We refer the interested reader to the monograph of A. Jeffrey & T. Kawahara [JeK982].

having invoked the dispersive relation in the last identity. Thus, equation (7.38) becomes

$$\imath\Big(\mathscr{A}_{t_2} - v_g\mathscr{A}_{x_2}\Big) + \frac{1}{2}\mathscr{A}_{\xi\xi}\frac{\mathrm{d}\, v_g}{\mathrm{d}\, k} + \frac{k^2}{\mathscr{D}_\omega}\left[\left(\frac{1}{v_g^2 - c^2} + \frac{1}{6\mu\omega^2}\right)|\mathscr{A}|^2\mathscr{A} + \beta\mathscr{A}\right] = 0, \tag{7.39}$$

where $\mathscr{A} = \mathscr{A}(\xi, x_2, \ldots, x_M; t_2, \ldots, t_M)$ (since we have absorbed in it the term $\mathscr{E}\xi$, otherwise present in equation (7.37)). Note that if we're interested in second order perturbations, we can neglect the dependence of $\beta = \beta(x_2, \ldots, x_M; t_2, \ldots, t_M)$ from slowly–varying variables and considerer it as a constant within the approximation. Finally, introducing the scaling transforms

$$\xi \equiv \frac{1}{\varepsilon}(x_2 - v_g t_2) = x_1 - v_g t_1 = \varepsilon(x - v_g t), \qquad \tau \equiv t_2 = \varepsilon t_1 = \varepsilon^2 t,$$

we can write equation (7.39) in the standard form of the **nonlinear Schrödinger equation**

$$\imath\mathscr{A}_\tau + \frac{1}{2}\frac{\mathrm{d}\, v_g}{\mathrm{d}\, k}\mathscr{A}_{\xi\xi} + \frac{k^2}{\mathscr{D}_2}\left[\left(\frac{1}{v_g^2 - c^2} + \frac{1}{6\mu\omega^2}\right)|\mathscr{A}|^2 + \beta\right]\mathscr{A} = 0, \tag{7.40}$$

having used the transformations $\partial_{x_2} = \frac{1}{\varepsilon}\partial_\xi$ and $\partial_{t_2} = -\frac{v_g}{\varepsilon}\partial_\xi + \partial_\tau$.

REMARK 7.2 – As mentioned in the note preceding formula (7.28), one can expand u in terms of another small parameter δ measuring the strength of the nonlinearity, by assuming the asymptotic formula

$$u(x_0, x_1, \ldots, x_M; t_0, t_1, \ldots, t_M; \varepsilon, \delta) = \sum_{n=1}^{N} \delta^n u_n(x_0, x_1, \ldots, x_M; t_0, t_1, \ldots, t_M) + \mathcal{O}(\delta^{N+1}). \tag{7.41}$$

Several cases may occur depending on the relative importance of the parameters δ and ε, i.e. between the nonlinear and the dispersive properties of the wave. It is possible to show that [JeK982]:

- ◇ the case $\delta = \varepsilon$ corresponds to the one we've already studied;
- ◇ the case $\delta = \varepsilon^2$ describes a wave system *dominated by dispersive interaction*, thus characterized by a linear propagation of the envelopes. Proceeding in an analogous way to the case $\delta = \varepsilon$, we obtain a linear equation that is like (7.39), but without the nonlinear term $|\mathscr{A}|^2\mathscr{A}$;
- ◇ the case $\delta = \varepsilon^{1/2}$ describes a wave system *dominated by the nonlinear interaction*. The non–linearity appears already in the first non–secularity condition and the dispersive term is absent (i.e. the coefficient $\frac{\mathrm{d}\, v_g}{\mathrm{d}\, k} = \frac{\mathrm{d}\,\omega}{\mathrm{d}\, k} \propto \mathscr{W}''(k)$ vanishes). In an analogous way, one obtains

$$\imath\Big(\mathscr{A}_{t_1} + v_g\mathscr{A}_{x_1}\Big) + \frac{k^2}{\mathscr{D}_\omega}\left(\frac{1}{6\mu\omega^2}|\mathscr{A}|^2 + \beta\right)\mathscr{A} = 0,$$

  where $\mathscr{A}$ and β are functions of the slow scales $x_1, \ldots, x_M$ and $t_1, \ldots, t_M$.

In summary, in this section we've shown how, starting from a paradigmatic model of a NLDEE (the Boussinesq–type equation (7.25)), it's possible to recover the NLS equation under an appropriate multiple–scale limit. In this sense, one can somehow interpret the NLS as an "*attractor*" in the space of all possible NLDEEs, thus explaining its, already mentioned, universal character.

7.5.2 Long–wave approximation and the KdV equation. Here we derive the KdV equation from the multiple–scale approach to the Boussinesq–type equation (7.25) in the so–called *long–wave approximation*. First note that, within the Boussinesq wave–model, the limit of long waves implies that $\omega \to 0$ as $k \to 0$, owning to the dispersion relation (7.26). Therefore in the expansions (7.29) for the operators

$\mathscr{L}$ and $\mathscr{N}$, we can naturally remove both ∂_{x_0} and ∂_{t_0}, since they don't contribute to the perturbative equations (7.30). Thus $\mathscr{L}_0 = \mathscr{L}_1 = \mathscr{N}_0 = \mathscr{N}_1 = 0$ and the first few operators are reduced to

$$\mathcal{O}(\varepsilon^2) \quad : \quad \mathscr{L}_2 \equiv \partial^2_{t_1} - c^2\partial^2_{x_1}, \qquad \mathscr{N}_2 \equiv \tfrac{1}{2}\partial^2_{x_1}, \tag{7.42a}$$

$$\mathcal{O}(\varepsilon^3) \quad : \quad \mathscr{L}_3 \equiv \partial_{t_1}\partial_{t_2} - 2c^2\partial_{x_1}\partial_{x_2}, \qquad \mathscr{N}_3 \equiv \partial_{x_1}\partial_{x_2}, \tag{7.42b}$$

$$\mathcal{O}(\varepsilon^4) \quad : \quad \mathscr{L}_4 \equiv \partial^2_{t_2} + 2\partial_{t_1}\partial_{t_3} - c^2\big(\partial^2_{x_2} + 2\partial_{x_1}\partial_{x_3}\big) - \mu\partial^2_{t_1}\partial^2_{x_1}, \tag{7.42c}$$

The dependent variable u is then expanded completely into the asymptotic series

$$u(x_0, x_1, \ldots, x_M; t_0, t_1, \ldots, t_M; \varepsilon, \delta) = \sum_{n=1}^{N} \delta^n u_n(x_0, x_1, \ldots, x_M; t_0, t_1, \ldots, t_M) + \mathcal{O}(\delta^{N+1}). \tag{7.43}$$

We now have several cases, depending on the relative importance of the dispersion and nonlinearity effects into the wave–problem under scrutiny. Let's consider the case $\delta = \varepsilon^2$; we have

$$\mathcal{O}(\varepsilon^4) \quad : \quad \mathscr{L}_2[u_1] = 0, \tag{7.44a}$$

$$\mathcal{O}(\varepsilon^5) \quad : \quad \mathscr{L}_3[u_1] = 0, \tag{7.44b}$$

$$\mathcal{O}(\varepsilon^6) \quad : \quad \mathscr{L}_4[u_1] + \mathscr{L}_2[u_2] = \mathscr{N}_2[u_1^2]. \tag{7.44c}$$

Equation (7.44a) is a d'Alembert equation in the scaled variables (x_1, t_1), that is $(\partial^2_{t_1} - c^2\partial^2_{x_1})u_1 = 0$ and it is satisfied if u_1 depends on x_1 and t_1 only thought $\xi_1 \equiv x_1 - ct_1$ (equivalently, $\xi_1 \equiv x_1 + ct_1$ for waves propagating in the negative direction). Inserting what found into (7.44b), we obtain

$$\big(\partial_{t_2} + c\partial_{x_2}\big)\partial_{\xi_1} u_1(\xi, x_2, \ldots, x_M; t_2, \ldots, t_M) = 0, \tag{7.45}$$

having noted that $\partial_{x_1} = \partial_{\xi_1}$ and $\partial_{t_1} = -c\partial_{\xi_1}$. In turn, equation (7.45) is satisfied if u_1 depends on x_2 and t_2 through $\xi_2 \equiv x_2 - ct_2$; accordingly, the operators $\mathscr{N}_2$ and $\mathscr{L}_4$ transforms as $\mathscr{N}_2 = \frac{1}{2}\partial^2_{\xi_1}$ and $\mathscr{L}_4 = -c(2\partial_{t_3}\partial_{x_1} + 2c\partial_{x_3}\partial_{\xi_1} + \mu c\partial^4_{\xi_1})$ respectively, and equation (7.44c) reduces to

$$\Big(\partial^2_{t_1} - c^2\partial^2_{x_1}\Big)u_2 - c\Big(2\partial_{t_3} + 2c\partial_{x_3} + \mu c\partial^3_{\xi_1}\Big)\partial_{\xi_1} u_1 = \partial_{\xi_1}\big(u_1\partial_{\xi_1} u_1\big). \tag{7.46}$$

Note that the term $\mathscr{L}_2[u_2]$ in (7.46) has not been transformed since we don't know (in general) how u_2 depends from the scaled variables x_1, t_1. *If* we further assume that also u_2 depends on x_1 and t_1 only through $\xi_1 = x_1 - ct_1$, then the first term in equation (7.46) vanishes[57] and we are left with

$$\partial_{t_3} u_1 + c\partial_{x_3} u_1 + \tfrac{1}{2}\mu c\partial^3_{\xi_1} u_1 + \tfrac{1}{2c} u_1\partial_{\xi_1} u_1 = 0, \tag{7.47}$$

where we have integrated once with respect to ξ_1 and set to zero the integration constant. Finally, transforming to a coordinate system moving with the phase velocity c of the long–wave, i.e. introducing the new variables $\xi_3 \equiv x_3 - ct_3$ and $\tau \equiv t_3$, and replacing ξ_1 by ξ, we arrive to the equation

$$\partial_\tau u_1 + \frac{1}{2c} u_1\partial_\xi u_1 + \frac{1}{2}\mu c\partial^3_\xi u_1 = 0, \tag{7.48}$$

which is the KdV equation for the function u_1 with respect to the independent variables ξ and τ. Note that, differently from the asymptotic expansion discussed in §7.5.2, here it turns necessary to move the coordinate system with the *phase velocity* (and not with the group velocity) of the wave.

REMARK 7.3 – Other expansions can be analyzed, e.g. it can be shown that the choice $\delta = \varepsilon^3$ leads to a linear equation, whilst $\delta = \varepsilon$ leads to a nonlinear evolution equation with a dispersive term of the form

$$\partial_{t_2} u_1 + c\partial_{x_2} u_1 + \frac{1}{2c} u_1\partial_\xi u_1 = 0,$$

where the independent variables are the lower–order scales x_2 and t_2. Applications to more general systems can be found in the monograph of A. JEFFREY & T. KAWAHARA, *Asymptotic methods in Nonlinear Wave Theory*, §3.5.3, pgg. 98–105 [JeK982] and in references quoted therein.

[57] Being "on–shell", that is the function u_2 satisfies the d'Alembert equation $(\partial^2_{t_1} - c^2\partial^2_{x_1})u_2 = 0$.

A Bäcklund transform for the second Painlevé equation

In §6.2 we've discussed the connections between IST–solvable NLEEs and the Painlevé transcendents. Having in mind those properties, one may have an intuition about the reasons why the Painlevé transcendents have many common features (Lax pairs and Hamiltonian formalism, see §6.1) with IST–solvable NLEEs studies so far. Here we wish to use this connection to gain the BTs of the Painlevé transcendents from the BTs of the corresponding NLEEs. For instance, we shall treat the case of the KdV equation and its reduction towards $\mathscr{P}_{\rm II}$ given by the scaling (6.13).

With this aim, consider the so–called *potential–KdV* $w_t - 3w_x^2 + w_{xxx} = 0$ where $w_x(x,t) = u(x,t)$ and $\mathcal{K}[u] = 0$; assume then the scaling reduction [FoA982]

$$\eta \equiv \frac{x}{(6t)^{1/3}}, \qquad w(x,t) = -\frac{1}{(6t)^{1/3}}\left[\tfrac{1}{2}\eta^2 + 2\varphi(\eta)\right],$$

where $\varphi = \varphi(\eta)$ satisfies the similarity equation $\varphi_{\eta\eta\eta} + 6\varphi_\eta^2 + 4\eta\varphi_\eta - 2\varphi = 0$. Multiplying both sides by $2\varphi_{\eta\eta}$, adding and subtracting $4\varphi_\eta^2$ one can easily integrate the above equation to get

$$\varphi_{\eta\eta}^2 + 4\varphi_\eta^3 + 4\eta\varphi_\eta^2 - 4\varphi\varphi_\eta = \mu^2, \tag{A.1}$$

with μ^2 an integration constant. Now multiply the similarity equation by $2\varphi_\eta$ and subtract it from (A.1), in order to obtain the equation

$$2\varphi_\eta\varphi_{\eta\eta\eta} + 8\varphi_\eta^3 + 4\eta\varphi_\eta^2 - \varphi_{\eta\eta}^2 + \mu^2 = 0,$$

which corresponds, after a proper transformation of η and φ_η, to one of the fifty 2nd order ODEs classified by Painlevé & Gambier. Its general solution is therefore given by

$$\varphi_\eta(\eta;\mu^2) = -2^{-1/3}\left[\mathcal{V}_z(z;\mu) + \mathcal{V}^2(z;\mu) + \tfrac{1}{2}z\right], \qquad z \equiv 2^{1/3}\eta \tag{A.2}$$

where $\mathcal{V} = \mathcal{V}(z;\mu)$ satisfies $\mathscr{P}_{\rm II}$, i.e. $\mathcal{V}_{zz}(z;\eta) = 2\mathcal{V}^3(z;\eta) + z\mathcal{V}(z;\eta) + \mu - 1/2$. Having in mind this equation, we find $2\varphi_\eta\mathcal{V} = -2^{-1/3}\big(2\mathcal{V}\mathcal{V}_z + 2\mathcal{V}^3 + z\mathcal{V}\big) = -2^{-1/3}(2\mathcal{V}\mathcal{V}_z + \mathcal{V}_{zz} - \mu + 1/2)$; on the other hand, the equation (A.2) yields $\varphi_{\eta\eta} = -\big(\mathcal{V}_{zz} - 2\mathcal{V}\mathcal{V}_z + 1/2\big)$, so that

$$\mathcal{V}(z;\mu) = 2^{-1/3}\frac{\varphi_{\eta\eta}(\eta;\mu^2) + \mu}{2\varphi_\eta(\eta;\mu^2)}, \tag{A.3}$$

which can be inverted in order to express $\varphi = \varphi(\eta;\mu^2)$ in terms of $\mathcal{V} = \mathcal{V}(z;\mu)$, thus giving

$$\varphi(\eta;\mu^2) = \frac{1}{2^{2/3}}\Big\{\mathcal{V}_z^2(z;\mu) - \left[\mathcal{V}^2(z;\mu) + \tfrac{1}{2}z\right]^2 - 2\mu\mathcal{V}(z;\mu)\Big\}, \tag{A.4}$$

as it can be easily verified[58]. Inserting equation (A.4) into equation (A.1) and turning back to the original KdV equation, we finally obtain the general similarity KdV–solution, given by the scaling reduction

$$z \equiv \frac{x}{(3t)^{1/3}}, \qquad u(x,t) = \frac{1}{(3t)^{2/3}}\left[\mathcal{V}_z(z;\mu) + \mathcal{V}^2(z;\mu)\right]. \tag{A.5}$$

We derive now an expression for the $\mathscr{P}_{\rm II}$–aBTs by starting from the KdV's ones. Recall therefore the equations (4.13) and (4.14): having assumed similarity solutions (which are not invariant under Galilei's transformations), we have to conveniently replace them with the following ones

$$\begin{cases} \widetilde{w}_x + w_x &= \tfrac{1}{2}(\widetilde{w} - w)^2, \\ \widetilde{w}_t + w_t &= 2w_x^2 + w_x\big(\widetilde{w} - w\big)^2 + 2w_{xx}\big(\widetilde{w} - w\big). \end{cases} \tag{A.6}$$

[58] One has $\varphi_\eta = (2)^{-1/3}\left[2\mathcal{V}_z\mathcal{V}_{zz} - 2\big(\mathcal{V}^2 + \tfrac{1}{2}z\big)\big(2\mathcal{V}\mathcal{V}_z + \tfrac{1}{2}\big) - 2\mu\mathcal{V}_z\right] = -2^{-1/3}(\mathcal{V}^2 + \mathcal{V}_z + \tfrac{1}{2})$, as in the equation (A.2).

Let's rewrite the equations (A.6) in terms of $\varphi = \varphi(\eta;\mu^2)$. In order to do this, note that

$$\begin{aligned}
\widetilde{w}_x + w_x &= -2(6t)^{-2/3}(\widetilde{\varphi}_\eta + \varphi_\eta + \eta),\\
\widetilde{w}_t + w_x &= 2(6t)^{-4/3}\big[3\eta^2 + 2(\widetilde{\varphi}+\varphi) - 2\eta^2 - 2\eta(\widetilde{\varphi}-\varphi)^2\big],\\
2w_x^2 + w_x(\widetilde{w} - w)^2 + 2w_{xx}(\widetilde{w} - w) &= 2(6t)^{-2/3}\big[\eta^2 + 4\varphi_\eta^2 + 4\eta\varphi_\eta +\\
&\quad - 2(\eta + 2\varphi_\eta)(\widetilde{\varphi}-\varphi)^2 + 2(1+2\varphi_{\eta\eta})(\widetilde{\varphi}-\varphi)\big].
\end{aligned}$$

Inserting these expressions into the equations (A.6), we finally obtain

$$\begin{cases} \widetilde{\varphi}_\eta &= -\eta - \varphi_\eta - (\widetilde{\varphi}-\varphi)^2,\\ \varphi &= \varphi_\eta^2 + \eta\varphi_\eta - \varphi_\eta(\widetilde{\varphi}-\varphi)^2 + \varphi_{\eta\eta}(\widetilde{\varphi}-\varphi). \end{cases} \tag{A.7}$$

Note that if $\varphi = 0$ in (A.7), then we're left just with the first expression; otherwise, if $\varphi \neq 0$ then the second equation is algebraic of the 2nd order in $(\widetilde{\varphi}-\varphi)$ and admits the solution

$$\widetilde{\varphi}(\eta;\widetilde{\mu}^2) = \varphi(\eta;\mu^2) + \frac{\varphi_{\eta\eta}(\eta;\mu^2) \pm \mu}{2\varphi_\eta(\eta;\mu^2)},$$

having used equation (A.1); note that in (A.1) the integration constant appears as μ^2, thus we can ignore the minus sign in the above expression and absorb the $\pm$ as an indetermination for $\mathrm{sgn}(\mu)$. Also, observe that inserting the expression for $\widetilde{\varphi}$ into the analogous one for φ, one finds $\widetilde{\mu}^2 = (1-\mu)^2$ which implies $\widetilde{\varphi}(\eta;\widetilde{\mu}^2) = \varphi(\eta;(1-\mu)^2)$. Therefore, recalling the relation (A.3) for $\mathcal{V} = \mathcal{V}(z;\mu)$, we obtain an expression for the aBT of the exact KdV–reduction associated to the scaling transforms (A.5), namely

$$\varphi\big(\eta;(1-\mu)^2\big) = \varphi(\eta;\mu^2) + 2^{1/3}\mathcal{V}(z;\mu). \tag{A.8}$$

Therefore the effect of two successive aBTs of the form (A.8) is

$$\varphi\big(\eta;(2-\mu)^2\big) = \widetilde{\varphi}(\eta;\widetilde{\mu}^2) + \frac{\widetilde{\varphi}_{\eta\eta}(\eta;\widetilde{\mu}^2) + \widetilde{\mu}}{2\widetilde{\varphi}_\eta(\eta;\widetilde{\mu}^2)} = \varphi(\eta;\mu^2) + 2^{1/3}\big[\mathcal{V}(z;\mu) + \mathcal{V}(z;\mu-1)\big], \tag{A.9}$$

having noted that the choice $\widetilde{\mu} = 1-\mu$ returns the original solution, since

$$\widetilde{\widetilde{\mu}}^{\,2} = \big(1-\widetilde{\mu}\big)^2 = \begin{cases} \mu^2 & \text{if } \ \widetilde{\mu} = 1-\mu,\\ (2-\mu)^2 & \text{if } \ \widetilde{\mu} = \mu - 1. \end{cases}$$

The generalization of the equation (A.9) to an arbitrary number n of aBTs follows immediately

$$\varphi\left(\eta;(n-\mu)^2\right) = \varphi\big(\eta;\mu^2\big) + 2^{1/3}\sum_{k=0}^{n-1}\mathcal{V}\big(z;\mu-k\big). \tag{A.10}$$

The above structure reflects itself quite simply onto the solution $\mathcal{V} = \mathcal{V}(z;\mu)$ of the second Painlevé transcendent. In fact, from the equation (A.3) it follows that

$$\mathcal{V}(z;-\mu) = \mathcal{V}(z;\mu) + \frac{\mu}{\mathcal{V}_z(z;\mu) + \mathcal{V}^2(z;\mu) + \frac{1}{2}z},$$

and, recalling that for $\widetilde{\mu} = 1-\mu$ we get back the original solution, one finally finds

$$\mathcal{V}(z;\mu) + \mathcal{V}(z;1-\mu) = 0. \tag{A.11}$$

References

[AKNS73] M. J. Ablowitz, D. J. Kaup, A. C. Newell, H. Segur – *Method for solving the sine–Gordon equation*, Phys. Rev. Lett. **30**: 1262–1264 (1973).

[AKNS74] M. J. Ablowitz, D. J. Kaup, A. C. Newell, H. Segur – *The inverse–scattering transform – Fourier analysis for nonlinear problems*, Stud. Appl. Math. **53**: 249–315 (1974).

[ARS978] M. J. Ablowitz, A. Ramani, H. Segur – *Nonlinear evolution equations and ordinary differential equations of Painlevé type*, Lett. Nuovo Cim. **23**: 333–338 (1978).

[ARS80a] M. J. Ablowitz, A. Ramani, H. Segur – *A connection between nonlinear evolution equations and ordinary differential equations of* $\mathbb{P}$*-type I*, J. Math. Phys. **21**: 715–721 (1980).

[ARS80b] M. J. Ablowitz, A. Ramani, H. Segur – *A connection between nonlinear evolution equations and ordinary differential equations of* $\mathbb{P}$*-type II*, J. Math. Phys. **21**: 1006–15 (1980).

[AbS981] M. J. Ablowitz, H. Segur – *Solitons and the Inverse Scattering Transform*, Soc. for Indus. and Appl. Math., Philadelphia (1981).

[AbC991] M. J. Ablowitz, P. A. Clarkson – *Solitons, Nonlinear evolution equations and Inverse Scattering*, London, Math. Soc. Lecture Notes Series Vol.**149**, Cambridge (1991).

[AbF003] M. J. Ablowitz, A. S. Fokas – *Complex Variables: introduction and applications*, Cambridge texts in Appl. Math. (2nd ed.), Cambridge Uni. Press (2003).

[AbS972] M. Abramowitz, I. A. Stegun – *Handbook of mathematical functions with formulas, graphs and mathematical tables*, New York, Dover, Ch. 16: 567–581 (1972).

[Akt009] T. Aktosun – *Inverse Scattering Transform and the theory of solitons*, In: R. A. Meyers (eds.) – *Encyclopedia of Complexity and System Science*, Springer, New York: 4960–4971 (2009).

[AnI979] R. L. Anderson, N. J. Ibragimov – *Lie–Bäcklund transformations in Applications*, SIAM studies 1, Society for Industrial and Applied Mathematics, Philadelphia (1979).

[Arn989] V. I. Arnol'd – *Mathematical methods of classical mechanics, II ed.*, Graduate Texts in Mathematics, Vol. **60**, Springer (1989).

[BeR969] D. J. Benney, G. J. Roskes – *Wave instabilities*, Stud. Appl. Math. **48**: 377–358 (1969).

[Bia879] L. Bianchi – *Ricerche sulle superfici a curvatura costante e sulle elicoidi*, Tesi di Abilitazione, Ann. Scuola Norm. Sup. Pisa(1) **2**, 285–304 (1879).

[Bia885] L. Bianchi – *Sopra i sistemi tripli ortogonali di Weingarten*, Ann. Matem. **13**, 177–234 (1885).

[Bia923] L. Bianchi – *Lezioni di geometria differentiale* **1**–**4**, Zanichelli, Bologna (1923–1927).

[BoT982] M. Boiti, G. Z. Tu – *Bäcklund Transformations via Gauge Transformations*, Il Nuovo Cimento **71**(2): 253–264 (1982).

[BPT983] M. Boiti, F. Pempinelli, G. Z. Tu – *The Nonlinear Evolution Equation related to the Wadati–Konno–Ichikawa spectral problem*, Prog. of Th. Phys. **69**(1): 48–64 (1983).

[BLMP88] M. Boiti, J. Léon, L. Martina, F. Pempinelli – *Scattering of localized solitons in the plane*, Phys. Lett. A**132**(8–9): 432–439 (1988).

[BMP993] M. Boiti, L. Martina, F. Pempinelli – *Multidimensional localized solitons*, Chaos, Solitons & Fractals **5**(12): 2377–2417 (1993).

[Bou871] J. Boussinesq – *Théorie de l'intumescence liquide, applelée onde solitaire ou de translation, se propageant dans un canal rectangulaire*, Comp. Rend. de l'Ac. des Sciences **72**: 755–759 (1871).

[Cal978] F. Calogero – *Nonlinear evolution equations solvable by the inverse spectral transform*, in F. Calogero (eds.) – *Mathematical Problems in Theoretical Physics*, Springer–Verlag, New York (1978).

[Cha911] J. Chazy – *Sur les équations différentielles du troisieme et d'ordre supérieur dont l'intégrale générale a ses points critiques fixés*, Acta Math. **34**: 317–383 (1911).

[Che974] H.–H. Chen – *General derivation of Bäcklund Transformations from Inverse Scattering Problems*, Phys. Rev. Lett. **33**(15): 925–929 (1974).

[CoM008] R. Conte, M. Musette – *The Painlevé Handbook*, Springer–Verlag (2008).

[Cru955] M. M. Crum – *Associated Sturm–Liouville systems*, Quart. J. Math. Oxford **6**(2): 121–127 (1955).

[ChI965] B. V. Chirikov, F. M. Izrailev – *Statistical Properties of a Nonlinear String*, Institute of Nuclear Physics, Novosibirsk, USSR (1965).

[DPR005] T. Dauxois, M. Peyrard, S. Ruffo – *The Fermi–Pasta–Ulam ""numerical experiment": history and pedagogical perspectives*, Eur. Jour. Phys. **26**: S3–S11 (2005).

[Dau008] T. Dauxois – *Fermi, Pasta, Ulam and a misterious lady*, Physics Today **61**(1), 55–57 (2008).

[DaR008] T. Dauxois, S. Ruffo – *Fermi–Pasta–Ulam nonlinear lattice oscillations*, Scholarpedia **3**(8), 5538 (2008).

[DeS974] A. Davey, K. Stewartson – *On three dimensional packets of surface waves*, Proc. R. Soc. A **338**(1613): 101–110 (1974).

[DeT979] P. Deift, E. Trubowitz – *Inverse Scattering on the line*, Camm. Pure Appl. Math. **32**: 121–251 (1979).

[OLBC10] F. W. Olver, D. W. Lozier, R. F. Boisvert, C. W. Clark – *NIST Handbook of Mathematical Functions*, Cambridge Univ. Press, New York, NY (2010); print companion to NILF Digital Library of Mathematical Functions, http://dlmf.nist.gov/, Release 1.0.5 of 2012–10–01 (2012).

[DoB976] R. K. Dodd, R. K. Bullough – *Bäcklund transformations for the sine–Gordon equations*, Proc. R. Soc. Lond., A **351**: 499–523 (1976).

[DrJ993] P. G. Drazin, R. S. Johnson – *Solitons, an introduction*, Cambridge Texts in App. Math., Cambridge Uni. Press (1993).

[Dun010] M. Dunajski – *Solitons, Instantons and Twistors*, Oxford Uni. Press Inc., New York (2010); see also *Integrable systems*, Lecture notes, Cambridge University (2012).

[Eis947] L. P. Eisenhart – *An introduction to Differential Geometry*, Princeton Uni. Press (1947).

[FaT007] L. D. Faddeev, L. Takhatajan – *Hamiltonian methods in the Theory of Solitons*, Classics in Mathematics, Springer (2007).

[FPU955] E. Fermi, J. Pasta, S. Ulam – *Studies of Nonlinear Problems, I*, Los Alamos Report LA-1940 (1955), later publisched in *Collected Papers of Enrico Fermi, Vol. II*, Ed. E. Segre, University of Chicago Press: 978–989 (1965).

[FlN980] H. Flaschka, A. C. Newell – *Monodromy – and spectrum – preserving deformations*, Comm. Math. Phys. **75**: 65–116 (1980).

[Fok980] A. S. Fokas – *A symmetry approach to exactly solvable evolution equations*, J. Math. Phys. **21**: 1318–1326 (1980).

[FoA982] A. S. Fokas, M. J. Ablowitz – *On a unified approach to transformations and elementary solutions of Painlevé equations*, J. Math. Phys. **23**(11): 2033–42 (1982).

[For992] J. Ford – *The Fermi–Pasta–Ulam problem: paradox turns discovery*, Phys. Rep.**213**, 271–310 (1992).

[FoW004] P. J. Forrester, N. S. Witte – *Application of the τ–function theory of Painlevé equations to random matrices: PII, PIV and the GUE*, Comm. Math. Phys. **219**(2): 357–398 (2001); *Application of the τ–function theory of Painlevé equations to random matrices: PV, PIII and the LUE, JUE and CUE*, Comm. Pure Appl. Math. **55**(6): 679–727 (2001); *Application of the τ–function theory of Painlevé equations to random matrices: PVI and the JUE, CyUE, cJUE and scaled limits*, Nagoya Math. J. **174**: 29–114 (2004).

[FOR996] A. S. Fokas, P. J. Olver, P. Rosenau – *A plethora of integrable bi–Hamiltonian equations*, in *Algebraic aspects of itnegrable systems*, Progress in Nonlin. Diff. Equations and their appl. **26**: 93–101 (1996).

[FlN975] H. Flaschka, A. C. Newell – *Integrable systems of nonlinear evolution equations*, in J. Moser (ed.) – *Dynamical Systems: Theory and Applications*, Lecture Notes in Physics, **38**, New York (1975).

[Gam910] B. Gambier – *Sur les équations différentielles du second ordre et du premier degré dont l'intégrale générale est à points critiques fixes*, Act. Math. **33**: 1–55 (1910).

[GGKM67] C. S. Gardner, J. M. Greene, M. D. Kruskal, R. M. Miura – *Method for solving the Korteweg–de Vries equation*, Phys. Rev. Lett. **19**: 1095–7 (1967).

[GGKM74] C. S. Gardner, J. M. Greene, M. D. Kruskal, R. M. Miura – *Korteweg–de Vries equation and generalizations VI: methods for exact solutions*, Comm. Pure Appl. Math. **27**: 97–133 (1974).

[GeF963] I. M. Gel'fand, S. V. Fomin – *Calculus of Variations*, Prentice Hall Inc. (1963); Dover (2000).

[GeL951] I. M. Gel'fand, B. M. Levitan – *On the determination of a diffferential equation from its spectral function*, Izv. Akad. Nauk SSSR Ser. Mat. **15**:309–366 (1951).

[GPS001] H. Goldstein, C. P. Poole Jr., J. L. Safko – *Classical Mechanics*, Addison–Wesley, 3rd edition (2001).

[GLS002] V. I. Gomak, I. Laine, S. Smimomura – *Painlevé differential equations in the complex plane*, de Gruyter Studies in Mathematics **28**, Berlin (2002).

[IbS979] N. J. Ibragimov, A. B. Shabat – *Korteweg–de Vries equation from the group standard–point*, Doklady Akad. Nauk., USSR **244**: 57–61 (1979).

[Inc956] E. L. Ince – *Ordinary differential equations*, Dover, New York, Ch.12–14 (1956).

[JeK982] A. Jeffrey, T. Kawahara – *Asymptotic methods in Nonlinear Wave Theory*, Pitman Adv. Pub. Porgram (1982).

[JiM981] M. Jimbo, T. Miwa – *Monodromy preserving deformation of linear ordinary differential equations with rational coefficients II*, Phys. D. **2**(3): 407–448 (1981).

[Kar998] S. Karigiannis – *The Inverse Scattering Transform and integrability of nonlinear evolution equations*, Minor Thesis, Harvard Uni. (1998).

[Kaw973] T. Kawahara – *The derivative–expansion method and nonlinear dispersive waves*, J. Phys. Soc. Japan **35**: 1537–44 (1973).

[KoW975] K. Konno, M. Wadati – *Simple derivation of Bäcklund Transformation from Riccati form of Inverse Method*, Prog. of Th. Phys. **53**(6): 1652–56 (1975).

[Kov889] S. V. Kowalewskaya – *Sur le probléme de la rotation d'un corps solide autour d'un point*, Act. Math. **12**: 177–232 (1889).

[Kud997] N. A. Kudryashov – *The first and second Painlevé equations of higher order ans some relations between them*, Phys. Lett. A**224**: 353–360 (1997).

[Lam976] G. L. Lamb – *Bäcklund transformation at the turn of the century*, in R. M. Miura (eds.) – *Bäcklund transformations*, Lecture Notes in Mathematics **515**, 69–7, Springer–Verlag, New York (1976).

[Lax968] P. D. Lax – *Integrals of nonlinear equations of evolution and solitary waves*, Comm. Pure Appl. Math. **21**: 467–490 (1968).

[Lax978] P. D. Lax – *A Hamiltonian approach to the KdV and other equations*, in: M. G. Crandall (eds.) – *Nonlinear Evolution Equations*, Academic Press, New York, 207–244 (1978).

[Mag978] F. Magri – *A simple model of the integrable Hamiltonian equation*, J. Math. Phys. **19**, 1156 (1978).

[Mar955] V. A. Marchenko – *On the reconstruction of the potential energy from phase of the scattered waves*, Dokt. Akad. Nauk SSSR **104**: 695–698 (1955).

[McL994] R. McLachlan – *A gallery of constant–negative–curvature surfaces*, The Math. Intell. **14**(4) 31–37, Springer–Verlag (1994).

[Miu968] R. M. Miura – *Korteweg–de Vries equation and generalizations I. A remarkable explicit nonlinear transformation*, J. Math. Phys. **9**: 1202–4 (1968).

[MGK968] R. M. Miura, C. S. Gardner, M. D. Kruskal – *Korteweg–de Vries equation and generalizations II. Existence of conservation laws and constants of motion*, J. Math. Phys. **9**: 1204–9 (1968).

[MuJ002] U. Muğan, F. Jrad – *Painlevé Test and higher order differential equations*, J. Nonlinear Math. Phys. **9**(3): 282–310 (2002).

[Nay973] A. H. Nayfeh – *Perturbation methods*, Wiley–Interscience Pub., New York (1973).

[New985] A. C. Newell – *Solitons in Mathematics and Physics*, Soc. for Inf. and App. Math., Philadelphia (1985).

[New983] R. G. Newton – *The Marchenko and Gel'fand–Levitan methods in the inverse scattering problem in one and three dymensions*, in: J. B. Bednar, R. Redner, E. Robinson, A. Weglein (eds.) – *Conference on inverse scattering: theory and application*, SIAM, Philadelphia: 1–74 (1983).

[NoO997] M. Noumi, K. Okamoto – *Irreducibility of the Second and the Fourth Painlevé equations*, Fun. Ekv. **40**: 139–163 (1997).

[NMPZ84] S. Novikov, S. V. Manakov, L. P. Pitaevskii, V. E. Zakharov – *Theory of Solitons: the Inverse Scattering Method*, Plenyum Publ., New York (1984).

[Oka986] K. Okamoto – *Isomonodromic deformation and Painlevé equations, and the Garnier system*, J. Fac. Sci. Univ. Tokyo, Sect IA **33**: 575–618 (1986).

[Oka987] K. Okamoto – *Studies on the Painlevé equations. I. Sixt Painlevé equation, P_{VI}*, Ann. Mat. Pura Appl. **146**(4): 337–381 (1987); *Studies on the Painlevé equations. II. Fifth Painlevé equation P_V*, Japan J. Math. (N.S.) **13**(1): 47–76 (1987); *Studies on the Painlevé equations. III. Second and fourth Painlevé equations, P_{II} and $P_{\rm IV}$*, Math. Ann. **275**(2): 221–255 (1986).

[Olv993] P. J. Olver – *Applications of Lie groups to differential equations*, Graduate text in Mathematics, **107**, Springer–Verlag, New york (1993).

[Pai900] P. Painlevé – *Memoire sur les équations différentialles dont l'intégrale générale est uniforme*, Bull. Soc. Math. Phys. France **28**: 1–85 (1902).

[Pai902] P. Painlevé – *Sul les équations différentielles du second ordre et d'ordre supérieur dont l'intégrale générale est uniforme*, Act. Math. **25**: 201–261 (1900).

[Pal000] R. S. Palais – *An introduction to Wave Equations and Solitons*, lecture notes, Chinese Acad. of Science, Beijing (2000).

[Pem995] F. Pempinelli – *Gauge–Bäcklund Transformations*, Lecture Notes for the Nonlinear Physics class at Uni. of Lecce (1995).

[Pir979] F. A. E. Pirani – *Local jet–bundle formulation of Bäcklund–Transformations*, Reidel, Boston (1979).

[RoS002] C. Rogers, W. K. Schief – *Bäcklund and Darboux transformations, Geometry and modern applications in Soliton Theory*, Cambridge Text in Appl. Math., Cambridge Uni. Press (2002).

[Tri961] F. G. Tricomi – *Equazioni Differenziali*, Boringhieri (1961).

[TuM972] J. L. Tuck, M. T. Menzel – *The superperiod of the nonlinear weighted string (FPU) problem*, Adv. Math. **9**: 399–407 (1972).

[Ume987] H. Umemura – *On the irreducibility of the first differential equation of Painlevé*, Alg. Geom. and Com. Alg. **2**: 771–789, Tokyo (1987).

[Ume990] H. Umemura – *Second proof of the irreducibility of the first differential equation of Painlevé*, Nagoya Math. J. **117**: 125–171 (1990).

[Wad972] M. Wadati – *The modified Korteweg-de Vries equation*, J. Phys. Soc. Japan, **32**: 1681ff (1972).

[WSK975] M. Wadati, H. Sanuki, K. Konno – *Relationships among inverse method, Bäcklund transformation and infinite number of conservation laws*, Prog. of Th. Physics **53**(2): 419–436 (1975).

[WaE973] H. D. Wahlquist, F. B. Estabrook – *Bäcklund Transformation for solutions of the Kortweg–de Vries equation*, Phys. Rev. Lett. **31**, 1386–90 (1973).

[WTC983] J. Weiss, M. Tabor, G. Carnevale – *The Painlevé property for partial differential equations*, J. Math. Phys. **24**(3): 522–527 (1982).

[Wei983] J. Weiss – *The Painlevé property for partial differential equations. II: Bäcklund transformation, Lax pairs and the Schwarzian derivative*, J. Math. Phys. **24**(6): 1405–14 (1983).

[ZaK965] N. J. Zabusky, M. D. Kruskal – *Interaction of ""solitons" in a collisionless plasma and the recurrence of initial states*, Phys. Rev. Lett. **15**: 240–243 (1965).

[ZaF971] V. E. Zakharov, L. D. Faddeev – *Kortweg–de Vries equation, a completely integrable Hamiltonian system*, Funct. Anal. Appl., **5**: 280–287 (1971).

[ZaM974] V. E. Zakharov, S. V. Manakov – *On the complete integrability of a nonlinear Schrödinger equation*, Jour. of The. and Math. Physics, Vol.**19**, No. 3: 551–559 (1974).

[ZaS972] V. E. Zakharov, A. B. Shabat – *Exact Theory of two-dimensional self-focusing and one–dimensional self–modulation of waves in nonlinear media*, Soviet Phys. JETP **34**: 62–69 (1972).