Physics AUC, vol. 15 (part I), 31-37 (2005)

PHYSICS AUC

Hirota's bilinear method and soliton solutions

Jarmo Hietarinta*

Department of Physics, University of Turku, Turku, Finland

Abstract

In this lecture we will first discuss integrability in general, its meaning and significance, and then make some general observations about solitons. We will then introduce Hirota's bilinear method, which is particularly useful in constructing multisoliton solutions for integrable nonlinear evolution equations.

1 Why is integrability important?

In very general terms integrability means *regularity in time evolution.* This is due to the existence of many conserved quantities. Most dynamical systems have some conserved quantities, such as energy or total momentum, but integrable systems have more of them. How many is enough depends on the system. For example, a Hamiltonian system of N coordinates and N momenta is said to be *Liouville integrable* if it has N conserved quantities, which furthermore must be sufficiently regular (analytic) and mutually commuting under the Poisson bracket. In principle the system can then be solved by quadratures.

The existence of such a large quantity of conserved quantities is a special requirement for the system, and therefore integrable systems are rare. Already in 1880's Poincaré pointed out that the measure of integrable systems is 0 in the space of all dynamical systems. Such a comparison of chaotic vs. integrable system is similar to comparing real number vs. prime numbers, and even though prime numbers are rare nobody questions their importance.

Since integrable systems are so rare, why should we be interested in them?

1. **Regularity is a good property.** The existence of conserved quantities implies long time predictability, i.e., there is something to observe. In the case of solitons this means that we have conserved waves in various media, including optical fibers, with possible commercial application. Planetary motion in the solar system is not exactly integrable, but is it close enough when we consider dynamics in human timescales, and therefore it is possible to construct accurate calendars and tables for star-navigation.
2. **Integrability is a dominating property.** Exact, mathematical integrability is rare, but an integrable system dominates the dynamics of nearby systems. Proof: Solitons have been observed experimentally. Natural systems are never exactly integrable, since an exactly integrable system is always the result of various idealizations. Nevertheless soliton-like behavior has been observed in many natural systems such as surface waves in shallow channels, internal waves in oceans, light pulses in optical fibers, etc. Thus integrable systems provide new starting points for perturbative expansions. But one needs many integrable systems so that the starting point is not too far. For example, the free motion in space is of course integrable but it is not a good starting point for a description of the dynamics in the solar system, for this purpose one uses Kepler's two-body systems, which are also integrable and much closer to the problem in question.
3. **Why are certain nonlinear PDE's both widely applicable and integrable?** This question was asked by F. Calogero in [1]. He started with a fairly generic PDE and derived the behavior of a weakly dispersive wave for this system, and it turned out that

*e-mail: Jarmo.Hietarinta@utu.fi

the behavior was usually governed by the integrable nonlinear Schrödinger equation. This is another manifestation of the dominating property discussed above.

4. **Integrable systems contain interesting mathematics.** First of all, since integrability implies some level of solvability, one can study details of the system analytically. In these studies many interesting mathematical structures have been discovered, e.g., Kac-Moody algebras and quantum groups. Integrable systems also appear naturally in various other theories: the Sieberg-Witten theory is related to the integrable Calogero-Moser system, correlation functions often solve some integrable Painlevé equation, etc. The fundamental theory of solitons is based on the mathematical theory of Grassmannian manifolds, as described by the Kyoto school in the Sato theory [2], and from it one obtains practically all integrable hierarchies of PDE's.

2 Solitons observations

In this sections we will mention some experimental manifestations of solitons. But since most of the material has been discussed in detail elsewhere, we will just give some annotated web-links.

The first observation and written description of surface wave solitons in shallow water was given by John Scott Russell (observation 1834, article 1845). For details see
`http://www.ma.hw.ac.uk/~chris/scott_russell.html`.
This initial observation was recreated in 1995
`http://www.ma.hw.ac.uk/solitons/press.html`.
Shallow water solitons are in fact quite easy to create and some science parks contain such displays, see e.g.
`http://www.ma.hw.ac.uk/solitons/Snibston/fig2.jpg`

Several beautiful natural solitons are illustrated in
`http://www.amath.washington.edu/~bernard/kp/waterwaves.html`

In addition to surface waves, solitons can appear as internal waves under the surface of the sea. Such waves describe the motion of the interface between cold and warm, or salty and sweet, water. Internal waves have been observed by sonar, but perhaps the most interesting observations have been made by satellites. One of the first observations of internal wave solitons was by the Apollo-Soyuz satellite [3]
`http://atlas.cms.udel.edu/database/images/andaman_sea/Apollo1976.jpg`
In these pictures the solitons are seen through the small turbulent ripples on the otherwise flat sea-surface, the ripple waves are caused by water motion under the surface. Many more pictures and information about internal waves can be found in
`http://www.internalwaveatlas.com/Atlas2_index.html`,
`http://www.ifm.uni-hamburg.de/ers-sar/Sdata/oceanic/intwaves/index.html`
These internal waves can be very large (amplitude 30-120 meters) and there have been reports of huge oil drilling rigs being displaced by them.

The internal water waves discussed above are perhaps the biggest solitons observed. At the opposite end we have solitons as light pulses in optical fibers, described by the nonlinear Schrödinger equation. Such pulses can be used in telecommunication, or for making ultrashort pulses using oscillating breather solitons [4].

3 Soliton solutions

3.1 Soliton equations

The modern history of solitons starts with the numerical observations of the *elastic scattering* of solitary waves in the Korteweg–de Vries equation

$$u_t + 6uu_x + u_{xxx} = 0 \tag{1}$$

by Zabusky and Kruskal [5] (Here and in the following subscripts indicate partial derivatives.) We observe that this equation has two parts with different tendencies: the linear part $u_t+u_{xxx} = 0$ would alone describe dissipative, spreading waves, while the shock part $u_t + 6uu_x = 0$ would describe sharpening waves. Together, however, they can describe stable travelling waves: a single travelling wave solution can be easily obtained with the ansatz $u = f(x - ct)$ and yields

$$u = 2\,\partial_x^2 \log(1 + e^{p(x-p^2t)+c}) = \frac{p^2/2}{\cosh^2(\frac{1}{2}[p(x-p^2t)+c])}.$$

Since the travelling waves collided elastically, just like particles, they were called *solitons*.

After this seminal work of Zabusky and Kruskal the theory developed rapidly. The properties of the KdV equation were studied intensively: infinite number of conserved quantities were found, along with methods of solving the initial value problem. Subsequently the theory was extended to many other integrable soliton equations, most of which were $1+1$-dimensional, but some $2+1$-dimensional integrable equations have also been found. The set of integrable PDE's include the following:

The Kadomtsev-Petviashvili (KP) equation

$$(u_t + 6uu_x + u_{xxx})_x \pm u_{yy} = 0, \tag{2}$$

Lax' fifth-order equation (Lax5)

$$u_{xxxxx} + 10uu_{xxx} + 20u_xu_{xx} + 30u^2u_x + u_t = 0, \tag{3}$$

The modified KdV-equation (mKdV)

$$v_{xxx} + 6v^2v_x + v_t = 0, \tag{4}$$

The sine-Gordon (sG) equation

$$u_{xt} = \sin(u) \tag{5}$$

The nonlinear Schrödinger equation (NLS)

$$i\phi_t + \phi_{xx} + 2|\phi|^2\phi = 0. \tag{6}$$

Further details can be found from soliton textbooks, e.g. [6].

3.2 Hirota's method

Single soliton solutions can be found easily by using the travelling wave ansatz $f = f(x - vt)$, but what about multisoliton solutions? First of all it must be stated that explicit N-soliton solutions can only be found for integrable equations. If one is only interested in finding multisoliton solutions the best tool is Hirota's bilinear method [7], although many other methods can also be used.

It is important to realize that PDE's appearing in a given physical problem are not usually in the best form for the subsequent mathematical analysis. Hirota noticed that the best dependent variables for constructing soliton solutions are those in which the solution appears as a finite sum of exponentials. This was based on the observation that the solutions of (1) as obtained from the inverse scattering method were in the form

$$u = 2\partial_x^2 \log \det M,$$

where M was a matrix in which the x, t dependency entered through $e^{a_ix+b_it}$, i.e., $\det M$ was a polynomial of exponentials. Hirota then made the logical step of introducing a new dependent variable F by

$$u = 2\partial_x^2 \log F, \tag{7}$$

If this is substituted in (1) it can be written as

$$\partial_x \left[F_{xxxx}F - 4F_{xxx}F_x + 3{F_{xx}}^2 + FF_{xt} - F_xF_t\right] = 0. \tag{8}$$

This equation is fifth order in derivatives, but one overall derivative can be extracted.

In order to apply Hirota's method it is necessary that the equation is quadratic and that the derivatives only appear in combinations that can be expressed using Hirota's D-operator defined by:

$$D_x^n\, f\cdot g = (\partial_{x_1}-\partial_{x_2})^n f(x_1)g(x_2)\big|_{x_2=x_1=x}. \tag{9}$$

For example,

$$\begin{aligned} D_x\, f\cdot g &= f_x g - f g_x,\\ D_x D_t\, f\cdot g &= f_{xt} g - f_x g_t - f_t g_x + f g_{xt}. \end{aligned}$$

Thus D operates on a product of two functions like the Leibniz rule, except that it is antisymmetric,

$$D_x^n f\cdot g = (-1)^n D_x^n g\cdot f.$$

Equation (8) can be written in terms of the D-operator, the part in the square brackets is

$$(D_x^4 + D_x D_t)\, F\cdot F = 0. \tag{10}$$

Below we will show the usefulness of this formulation.

Unfortunately the process of bilinearization is far from being algorithmic. It is even difficult to know beforehand how many new independent and/or dependent variables are needed for bilinearization.

For the previously mentioned equations one finds the following bilinear forms, usually with the substitution (7):

KP:

$$[D_x^4 + D_x D_t \pm D_y^2] F\cdot F = 0,$$

Lax5:

$$\begin{cases} [D_x^6 + D_x D_t - \frac{5}{3}(D_x^4 + D_s^2)] F\cdot F &= 0,\\ (D_x^4 + D_x D_s) F\cdot F &= 0. \end{cases} \tag{11}$$

What is surprising here is that we have two equations for one function F. However, note that there is also a new independent variable s. By suitable cross differentiations one can eliminate the s-dependence and the result is eqn. (3).

For mKdV one uses a slightly different substitution

$$u = \partial_x\, 2\arctan(G/F), \text{ i.e., } u = 2\frac{D_x G\cdot F}{F^2+G^2}, \tag{12}$$

after which the once integrated (or potential) form of mKdV reads

$$(F^2+G^2)[(D_x^3 + D_t)G\cdot F] + 3(D_x F\cdot G)[D_x^2(F\cdot F + G\cdot G)] = 0. \tag{13}$$

It can be split as

$$\begin{cases} [D_x^3 + D_t] G\cdot F &= 0,\\ D_x^2(F\cdot F + G\cdot G) &= 0. \end{cases} \tag{14}$$

Thus in this case we need two dependent functions, F, G, and therefore two equations.

The bilinear form of SG is similar, with substitution

$$u = 4\arctan(G/F) \tag{15}$$

we get an equation that splits into

$$\begin{cases} [D_x D_t - 1] G\cdot F &= 0,\\ (D_x^2[F\cdot F - G\cdot G] &= 0. \end{cases} \tag{16}$$

Finally for NLS we use substitution

$$\phi = G/F, \quad G \text{ complex, } F \text{ real}, \tag{17}$$

which yields

$$F\,[(iD_t + D_x^2) G\cdot F] - G\,[D_x^2 F\cdot F - 2|G|^2] = 0. \tag{18}$$

For normal (bright) solitons we split this as

$$\begin{cases} (iD_t + D_x^2)G \cdot F &= 0, \\ D_x^2 F \cdot F &= 2|G|^2. \end{cases} \tag{19}$$

For a further discussion of bilinearization, see e.g., [8, 9].

Here are some useful properties of the bilinear derivative [10] (P is a polynomial):

$$\begin{aligned}
P(D)f \cdot g &= P(-D)g \cdot f, && (20)\\
P(D)1 \cdot f = P(-\partial)f, \quad & P(D)f \cdot 1 = P(\partial)f, && (21)\\
P(D)e^{px} \cdot e^{qx} &= P(p-q)e^{(p+q)x}, && (22)\\
P(D)e^{px}F \cdot e^{px}G &= e^{2px}P(D)F \cdot G && (23)\\
\partial_x^2 \log f &= (D_x^2 f \cdot f)/(2f^2), && (24)\\
\partial_x^4 \log f &= (D_x^4 f \cdot f)/(2f^2) - 3(D_x^2 f \cdot f)^2/(2f^4). && (25)
\end{aligned}$$

4 Constructing multisoliton solutions

4.1 The vacuum and the one-soliton solution

Here we will only consider KdV-type equations, by which we mean equations whose bilinear form is

$$P(D_x, D_y, ...)F \cdot F = 0, \tag{26}$$

where P is some polynomial in the Hirota partial derivatives D. It is quite easy to deal with this whole class at the same time. We may assume that P is even, because the odd terms cancel due to the antisymmetry of the D-operator (20).

Let us start with the zero-soliton solution or the "vacuum". We know that the KdV equation has the solution $u \equiv 0$ and now we want to find the corresponding F. From (7) we see that $F = e^{2\phi(t)x+\beta(t)}$ yields a u that solves (10), and in view of the gauge freedom (23) we can choose $F = 1$ as our vacuum solution. It solves (26) provided that

$$P(0, 0, \dots) = 0, \tag{27}$$

which is a condition that we have to impose on the polynomial P.

The multisoliton solutions are obtained by finite perturbation expansions around the vacuum $F = 1$:

$$F = 1 + \epsilon\, f_1 + \epsilon^2\, f_2 + \epsilon^3\, f_3 + \cdots \tag{28}$$

Here ϵ is a formal expansion parameter, and for an N-soliton solution of (26) the expansion stops at ϵ^N. If the bilinear form contains several functions similar expansions should be written for all of them. However, symmetry consideration may be useful, e.g., for (19) F is even and G odd in ϵ.

If we substitute

$$F = 1 + \epsilon\, f_1 \tag{29}$$

into (26) we obtain

$$P(D_x, \dots)\{1 \cdot 1 + \epsilon\, 1 \cdot f_1 + \epsilon\, f_1 \cdot 1 + \epsilon^2\, f_1 \cdot f_1\} = 0.$$

The term of order ϵ^0 vanishes because of (27). For the terms of order ϵ^1 we use property (21) so that, since now P is even, we get

$$P(\partial_x, \partial_y, \dots)f_1 = 0. \tag{30}$$

The soliton solutions correspond to the exponential solutions of (30). For a 1SS we take an f_1 with just one exponential

$$f_1 = e^{\eta}, \quad \eta = px + qy + \cdots + \text{ const}, \tag{31}$$

and then (30) becomes the *dispersion relation* on the parameters $p, q, \dots$

$$P(p, q, \dots) = 0. \tag{32}$$

Finally, with (31) the order ϵ^2 term vanishes because $P(\vec{D})e^{\eta} \cdot e^{\eta} = e^{2\eta} P(\vec{p} - \vec{p}) = 0$, by (27).

In summary, the 1SS is given by (29,31) where the parameters are constrained by (32). For KdV the dispersion relation is $q^3 = p$.

4.2 The two-soliton solution

The 2SS is built from two 1SS's, and one important principle is that for integrable systems one must be able to combine *any* pair of 1SS's built on top of the same vacuum. Thus if we have two 1SS's, $F_1 = 1 + e^{\eta_1}$ and $F_2 = 1 + e^{\eta_2}$, we should be able to combine them into a form $F = 1 + f_1 + f_2$, where $f_1 = e^{\eta_1} + e^{\eta_1}$. Gauge invariance suggests that we should try the combination

$$F = 1 + e^{\eta_1} + e^{\eta_2} + A_{12} e^{\eta_1 + \eta_2} \tag{33}$$

where there is just one arbitrary constant A_{12}. Substituting this into (26) yields

$$\begin{array}{rcccccccl}
P(D)\{ \; 1 \cdot 1 & + & 1 \cdot e^{\eta_1} & + & 1 \cdot e^{\eta_2} & + & \underline{A_{12}\, 1 \cdot e^{\eta_1+\eta_2}} & + & \\
e^{\eta_1} \cdot 1 & + & e^{\eta_1} \cdot e^{\eta_1} & + & \underline{e^{\eta_1} \cdot e^{\eta_2}} & + & A_{12}\, e^{\eta_1} \cdot e^{\eta_1+\eta_2} & + & \\
e^{\eta_2} \cdot 1 & + & \underline{e^{\eta_2} \cdot e^{\eta_1}} & + & e^{\eta_2} \cdot e^{\eta_2} & + & A_{12}\, e^{\eta_2} \cdot e^{\eta_1+\eta_2} & + & \\
\underline{A_{12} e^{\eta_1+\eta_2} \cdot 1} & + & A_{12} e^{\eta_1+\eta_2} \cdot e^{\eta_1} & + & A_{12} e^{\eta_1+\eta_2} \cdot e^{\eta_2} & + & A_{12}^2 e^{\eta_1+\eta_2} \cdot e^{\eta_1+\eta_2} & & \} = 0.
\end{array}$$

In this equation all non-underlined terms vanish due to (27,32). Since P is even, the underlined terms combine as $2A_{12} P(\vec{p}_1 + \vec{p}_2) + 2P(\vec{p}_1 - \vec{p}_2) = 0$, from which A_{12} can be solved as

$$A_{12} = -\frac{P(\vec{p}_1 - \vec{p}_2)}{P(\vec{p}_1 + \vec{p}_2)}. \tag{34}$$

The important thing about this result is that we were able to construct a two-soliton solution for a huge class of equations, namely all those whose bilinear form is of type (26). In particular this includes many non-integrable systems.

4.3 Multi-soliton solutions

The above shows that for the KdV class (26) the existence of a 2SS is not strongly related to integrability. However, it turns out that the existence on 3SS is very restrictive.

Following the above ideas we conclude that a 3SS should start with $f_1 = e^{\eta_1} + e^{\eta_2} + e^{\eta_3}$ and contain terms up to f_3. Furthermore, we will use the requirement that the solution $u = 2\partial_x^2 \log F$ should reduce to a 2SS when the third soliton is far away (which corresponds to $\eta_k \to \pm\infty$). From these one finds that F must have the form

$$\begin{aligned} F &= 1 + e^{\eta_1} + e^{\eta_2} + e^{\eta_3} \\ &\quad + A_{12} e^{\eta_1+\eta_2} + A_{13} e^{\eta_1+\eta_3} + A_{23} e^{\eta_2+\eta_3} + A_{12}A_{13}A_{23} e^{\eta_1+\eta_2+\eta_3}. \end{aligned} \tag{35}$$

Note in particular that this expression contains *no additional freedom*. The parameters p_i are only required to satisfy the dispersion relation (32) and the phase factors A were already determined (34). This form extends to NSS [11]:

$$F = \sum_{\substack{\mu_i = 0,1 \\ 1 \le i \le N}} \exp\left(\sum_{1 \le i < j \le N} \varphi(i,j) \mu_i \mu_j + \sum_{i=1}^{N} \mu_i \eta_i \right), \tag{36}$$

where ($A_{ij} = e^{\varphi(i,j)}$). Thus the ansatz for a NSS is completely fixed and the requirement that it be a solution of (26) implies conditions on the equation itself. Only for integrable equations can we combine solitons in this simple way. This leads us to consider the following:

Definition: *A set of equations written in the Hirota bilinear form is* **Hirota integrable**, *if one can combine any number N of one-soliton solutions into an NSS, and the combination is a finite polynomial in the e^{η}'s involved.*

In all cases studied so far, Hirota integrability has turned out to be equivalent to more conventional definitions of integrability.

4.4 Searching for integrable PDE's

Since the existence of a 3SS is very restrictive, one can use it as a method for searching for new integrable equations. All search methods depend on some initial assumptions about the structure of the class of equations considered. In this case it is necessary to assume that the nonlinear PDE can be put into a bilinear form, but no assumptions need to be made for example on the number of independent variables.

In papers [12, 13, 14] such a search was conducted for the bilinear classes of KdV, MKdV, SG and NLS type. This resulted in some interesting equations, among them the following:

$$(D_x^3 D_t + aD_x^2 + D_t D_y) F \cdot F = 0, \tag{37}$$

$$(D_x^4 - D_x D_t^3 + aD_x^2 + bD_x D_t + cD_t^2) F \cdot F = 0, \tag{38}$$

$$\begin{cases} [D_x D_y D_t + aD_x + bD_y] G \cdot F = 0, \\ D_x D_y [F \cdot F + G \cdot G] = 0. \end{cases} \tag{39}$$

$$\begin{cases} [aD_x^3 + bD_y^3 + D_t] G \cdot F = 0, \\ D_x D_y [F \cdot F + G \cdot G] = 0. \end{cases} \tag{40}$$

$$\begin{cases} (i\alpha D_x^3 + 3D_x D_y - 2iD_t + c)\, G \cdot F = 0, \\ (a(\alpha^2 D_x^4 - 3D_y^2 + 4\alpha D_x D_t) + bD_x^2)\, F \cdot F = |G|^2. \end{cases} \tag{41}$$

Perhaps the most interesting new equation above is the combination in (41) of the two most important $(2+1)$-dimensional equations, Davey-Stewartson and Kadomtsev-Petviashvili equations, see [15], page 315.

References

[1] F. Calogero, in " What is Integrability?", V.E. Zakharov (ed.), Springer.

[2] T. Miwa, M. Jimbo and E. Date, " Mathematics of Solitons", Cambridge University Press (2000). Y. Ohta, J. Satsuma, D. Takahashi and T. Tokihiro, *Prog. Theor. Phys. Suppl.*, 94 (1988) 210–41.

[3] A.R. Osborne and T.L. Burch, *Science*, 208 (1980) 451–460.

[4] L. F. Mollenauer, R. H. Stolen and J. P. Gordon, *Phys. Rev.Lett.* 45 (1980) 1095

[5] N.J. Zabusky and M.D. Kruskal, *Phys. Rev. Lett.* 15 (1965) 240-243.

[6] M. J. Ablowitz and P. A. Clarkson 1991 *Solitons, Nonlinear Evolution Equations and Inverse Scattering*, Cambridge, U.K.: Cambridge University Press.

[7] R. Hirota, *Phys. Rev. Lett.* **27**, 1192 (1971).

[8] R. Hirota in "Solitons", R.K. Bullough and P.J. Caudrey (eds.), Springer (1980), p. 157.

[9] J. Hietarinta, in "Partially Integrable Evolution Equations in Physics", R. Conte and N. Boccara (eds.), Kluwer Academic (1990), p. 459.

[10] R. Hirota, *Progr. Theor. Phys.* **52**, 1498 (1974).

[11] R. Hirota, *J. Math. Phys.* **14**, 810 (1973).

[12] J. Hietarinta, *J. Math. Phys.* **28**, 1732 (1987).

[13] J. Hietarinta, *J. Math. Phys.* **28**, 2094, 2586 (1987).

[14] J. Hietarinta, *J. Math. Phys.* **29**, 628 (1988).

[15] J. Hietarinta, in "Nonlinear evolution equations: integrability and spectral methods", A. Degasperis, A.P. Fordy and M. Lakshmanan (eds.), Manchester University Press (1990), p. 307.