GENERAL ARTICLE

The Idea of a Lax Pair–Part II*

Continuum Wave Equations

Govind S. Krishnaswami and T R Vishnu

In Part I [1], we introduced the idea of a Lax pair and explained how it could be used to obtain conserved quantities for systems of particles. Here, we extend these ideas to continuum mechanical systems of fields such as the linear wave equation for vibrations of a stretched string and the Korteweg-de Vries (KdV) equation for water waves. Unlike the Lax matrices for systems of particles, here Lax pairs are differential operators. A key idea is to view the Lax equation as a compatibility condition between a pair of linear equations. This is used to obtain a geometric reformulation of the Lax equation as the condition for a certain curvature to vanish. This 'zero curvature representation' then leads to a recipe for finding (typically an infinite sequence of) conserved quantities.

Govind Krishnaswami is on the faculty of the Chennai Mathematical Institute. He works on various problems in theoretical and mathematical physics.

T R Vishnu is a PhD student at the Chennai Mathematical Institute. He has been working on integrable systems.

Keywords
Continuum mechanics, fields, linear wave equation, Lax pair, KdV equation, conserved quantities, zero curvature representation, monodromy matrix.

1. Introduction

In the first part of this article [1], we introduced the idea of a dynamical system: one whose variables evolve in time, typically via differential equations. It was pointed out that conserved quantities, which are dynamical variables that are constant along trajectories help in simplifying the dynamics and solving the equations of motion (EOM) both in the classical and quantum settings. We then introduced the concept of a Lax pair of $p \times p$ square matrices (L, A) which, when available, is a useful tool to find conserved quantities. A system admits a Lax pair if its EOM are equivalent to the Lax equation (which is a system of p^2 equations)

$$L_t = \dot{L} \equiv \frac{dL}{dt} = [L, A] = LA - AL. \quad (1)$$

*Vol.26, No.2, DOI: https://doi.org/10.1007/s12045-021-1124-1

Here, the matrix elements of L and A depend on the dynamical variables. As a consequence of (1), the Lax matrix L evolves isospectrally: its eigenvalues and traces of its powers (tr $L^k, k = 1, 2, \ldots$) are conserved. This was illustrated using the harmonic oscillator, Toda chain, and Euler top which are classical mechanical systems with finitely many degrees of freedom (one for the oscillator, three for the top and N for a Toda chain of N atoms).

Here, we move from systems of particles with finitely many degrees of freedom to continuum mechanical systems of fields with infinitely many degrees of freedom. A simple example of a field is the height $u(x, t)$ of a vibrating stretched string. Classical fields evolve via partial differential equations (PDEs) such as the wave equation $u_{tt} = c^2 u_{xx}$, as opposed to Newton's ordinary differential equations (e.g. $m\ddot{q} = -kq$) of particle mechanics. Here, subscripts on u denote its partial derivatives: $u_t = \partial_t u = \partial u/\partial t$, etc. To specify an instantaneous configuration of the string, one needs to specify the height field $u(x, t)$ at each point x along the string. It is in this sense that fields possess infinitely many degrees of freedom. The height of a surface water wave, the pressure in a fluid or the magnetic field around the Earth are other examples of dynamical fields.

Continuum mechanical systems of classical fields evolve via partial differential equations such as the wave equation $u_{tt} = c^2 u_{xx}$, as opposed to Newton's ordinary differential equations (e.g. $m\ddot{q} = -kq$) of particle mechanics.

Our first example of a continuum mechanical system with a Lax pair is perhaps the simplest example of a field equation: the linear wave equation $u_t + cu_x = 0$. The latter describes waves that maintain their shape and move rightwards at a constant speed $c > 0$. Unlike the matrices encountered in Part I [1], the Lax pair L and A are now a pair of differential operators. As one might suspect from the existence of a Lax pair, the wave equation admits an infinite number of conserved quantities, since the differential operator L may be viewed as an infinite-dimensional matrix whose spectrum (eigenvalues) is conserved. We then generalize this Lax pair to a richer physical system, the nonlinear Korteweg-de Vries (KdV) equation for water waves: $u_t - 6uu_x + u_{xxx} = 0$. The Lax pair for KdV (consisting of second and third-order differential operators L and A) can be used to find an infinite number of unexpected and nontrivial conserved quantities. In fact, it turns

Unlike the finite-dimensional Lax matrices for systems of particles, in continuum mechanics (systems of fields), L and A are differential operators.

out that the wave and KdV equations are the first two members of an infinite 'KdV hierarchy' of equations that share the same set of conserved quantities!

We then use the KdV example to pass to a more symmetrical reformulation of the Lax pair idea where the Lax equation is viewed as a compatibility condition for a pair of linear equations involving only first derivatives with respect to space and time. This compatibility condition has a geometric meaning: it says that a certain curvature or fictitious electromagnetic field strength vanishes. More generally, we will say that a (nonlinear) system of field equations admits a zero curvature representation if the equations are equivalent to the condition for a certain curvature to vanish. Remarkably, a number of interesting nonlinear field equations, especially in one spatial dimension (such as the mKdV, nonlinear Schrödinger, sine-Gordon, Heisenberg magnetic chain and principal chiral), admit zero curvature representations. What is more, one can use the vanishing of this curvature to obtain (infinitely many) conserved quantities for these systems. This proceeds via an object called the monodromy or parallel transport matrix which governs how vectors change when one goes around a closed spatial loop. Our exploration of the Lax pair idea will also come around a full circle when we show that the monodromy matrix itself satisfies a Lax-like equation so that the traces of its powers furnish a set of conserved quantities! The presence of these conserved quantities gives these systems very remarkable features: aside from admitting exact solutions, these systems typically admit special types of spatially localized solitary waves called solitons which can scatter in complicated ways and yet reemerge while retaining their shapes.

The Lax equation admits a geometric reformulation known as the zero curvature representation which is obtained by viewing it as a compatibility condition between two first-order linear equations. This compatibility condition is equivalent to the vanishing of a certain curvature (which means that a certain space is flat).

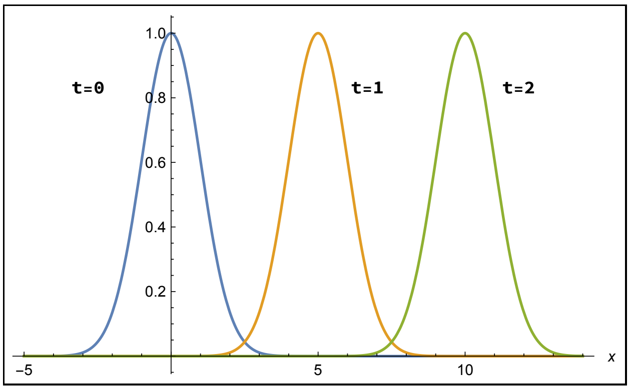
The 1st order wave equation $u_t + cu_x = 0$ is related to the 2nd order wave equation $(\partial_t^2 - c^2\partial_x^2)\phi = 0$. Indeed, the d'Alembert wave operator $\partial_t^2 - c^2\partial_x^2$ may be factorized as $(\partial_t + c\partial_x)(\partial_t - c\partial_x)$. The first-order equations $u_t + cu_x = 0$ and $v_t - cv_x = 0$ describe right/left-moving waves $u = f(x - ct)$ and $v = g(x + ct)$ while the 2nd order wave equation describes bi-directional propagation: $\phi(x, t) = f(x - ct) + g(x + ct)$.

2. Wave Equation

One of the simplest field equations in one dimension (1D) is the wave equation:

$$u_t + cu_x = 0 \quad \text{for constant} \quad c. \tag{2}$$

Figure 1. Right-moving solitary wave $u(x,t) = f(x - ct)$ at times $t = 0, 1, 2$ with profile $f = \exp(-x^2/2)$ and speed $c = 5$.



Here, $u(x,t)$ could represent the amplitude/height of the wave (sound/water, etc.) at position x at time t. For $c > 0$, this partial differential equation describes right-moving waves that travel at speed c while maintaining their shape (see *Figure* 1).

In the time-independent Schrödinger equation $L\psi = \lambda\psi$, $L = -\partial_x^2 + u(x,t)$ plays the role of the Hamiltonian (energy operator), u that of potential energy and λ the energy eigenvalue in units where $\hbar^2/2m = 1$. In one-dimensional quantum systems, bound state energy eigenvalues are nondegenerate (see p. 99 of [2]). Thus, the corresponding energy eigenspaces are one-dimensional, as assumed in §3 of [1].

Indeed, one checks that $u(x,t) = f(x - ct)$ is a solution of (2) for any differentiable function f. We seek a Lax pair of differential operators L and A (depending on u) such that $L_t = [L, A]$ is equivalent to (2). It is convenient to take L to be the Schrödinger operator $L = -\partial^2 + u(x,t)$, where $\partial = \partial_x = \partial/\partial x$. L is familiar from Sturm-Liouville theory as well as from quantum mechanics as the Hamiltonian of a particle moving in the potential $u(x,t)$. Since L is symmetric (hermitian), $L_t = u_t$ is also symmetric, so for the Lax equation to make sense $[L, A]$ must also be symmetric. As in §2 of [1], choosing A to be anti-symmetric (up to the addition of an operator that commutes with L) guarantees this. It turns out that $A = c\partial$ does the job (see *Box* 1). Indeed, using the commutator $[\partial, f] = f'$ for any function f, we see that the Lax equation is equivalent to the wave equation:

$$L_t = u_t = [L, A] = [-\partial^2 + u(x,t), c\partial] = [u, c\partial] = -cu_x. \qquad (3)$$

We will use this Lax pair as a stepping stone to find a Lax pair for the KdV equation, which is a nonlinear wave equation with widespread applications. As discussed in Part I [1], the existence of a Lax pair is usually associated with the presence of conserved quantities. For example, integrating (2) in x, we get

$$\frac{d}{dt}\int_{-\infty}^{\infty} u\,dx = -c\int_{-\infty}^{\infty} u_x dx = -c(u(\infty) - u(-\infty)) = 0 \qquad (4)$$

assuming $u \to 0$ as $x \to \pm\infty$. Thus, $C_1 = \int_{-\infty}^{\infty} u\,dx$ is conserved. The reason this worked is that (2) takes the form of a local conservation law: $\partial_t\rho + \partial_x j = 0$ with $\rho = u$ and $j = cu$.

Box 1. Lax pair for the wave equation.

The choice $A = c\partial$ to partner the Schrödinger operator L in the Lax pair (3) for the wave equation (2) can be arrived at by starting from the simplest of differential operators, a first order differential operator $\alpha(x,t)\partial + \beta(x,t)$ and imposing some consistency conditions. We shall see in *Box* 4, that this approach generalizes to other equations. To make A anti-symmetric, we subtract its adjoint and consider

$$A_1 = (\alpha\partial + \beta - \partial^\dagger\alpha - \beta) = (\alpha\partial + \partial\alpha) = [\alpha,\partial]_+ = (\alpha' + 2\alpha\partial) \quad \text{where} \quad \alpha' = \frac{\partial\alpha}{\partial x}. \qquad (5)$$

Here, we used (i) $\partial^\dagger = -\partial$, which is familiar from the hermiticity of momentum $p = -i\hbar\partial$ in quantum mechanics, (ii) $g^\dagger = g$ for any real function g and (iii) $(\partial\alpha)(f) = \alpha' f + \alpha f'$ so that $\partial\alpha = \alpha' + \alpha\partial$. The commutator with the Schrödinger operator L is then

$$[L, A_1] = [-\partial^2 + u, \alpha' + 2\alpha\partial] = -\alpha''' - 4\alpha''\partial - 4\alpha'\partial^2 - 2\alpha u'. \qquad (6)$$

Here, we used $[\partial,\alpha] = \alpha'$, the Leibnitz product rule, linearity and anti-symmetry of commutators to obtain

$$\begin{aligned} [u,\alpha'] &= 0, \quad [u, 2\alpha\partial] = -2\alpha u', \quad [-\partial^2,\alpha'] = -\partial[\partial,\alpha'] - [\partial,\alpha']\partial = -\alpha''' - 2\alpha''\partial \quad \text{and} \\ [-\partial^2, 2\alpha\partial] &= -\partial[\partial, 2\alpha\partial] - [\partial, 2\alpha\partial]\partial = -\partial(2\alpha'\partial) - (2\alpha'\partial)\partial = -4\alpha'\partial^2 - 2\alpha''\partial. \end{aligned} \qquad (7)$$

In the Lax equation $L_t = [L, A_1]$, $L_t = u_t$ is multiplication by $u_t(x,t)$. For $[L, A_1]$ in (6) to also be a multiplication operator, the coefficients of ∂ and ∂^2 must vanish which implies $\alpha' = \alpha'' \equiv 0$ for all x. This implies $\alpha = \alpha(t)$ is a function of time alone. Thus, $L_t = [L, A_1]$ becomes $u_t = -2\alpha(t)u_x$. For this to be equivalent to the wave equation $u_t + cu_x = 0$, we must pick $\alpha(t) = c/2$, so that A_1 reduces to $A = c\partial$.

Integrating an equation in local conservation form implies the conservation of $\int_{-\infty}^{\infty}\rho dx$, provided the 'flux' of j across the 'boundary' vanishes: $j(\infty) - j(-\infty) = 0$. Similarly, multiplying (2) by u leads to an equation that is again in local conservation form: $\partial_t(u^2/2) + c\partial_x(u^2/2) = 0$. Thus, $C_2 = \int_{-\infty}^{\infty} u^2\,dx$ is *also* conserved. In a similar manner, we find that $\partial_t u^n + c\partial_x u^n = 0$, so that

$C_n = \int_{-\infty}^{\infty} u(x,t)^n \, dx$ is conserved for any $n = 1, 2, 3, \ldots$. Thus, the wave equation admits infinitely many constants of motion.

The infinite sequence of conserved quantities Q_n are special: they are conserved for both the linear wave and KdV equations. In fact, they are conserved for *all* the equations in the KdV 'hierarchy'.

However, unlike in Part I [1], C_n have not been obtained from the Lax operator L. As we will see in *Box* 2, the wave equation also admits *another* infinite sequence of conserved quantities Q_n that may be obtained from L. Unlike C_n, the Q_n turn out to be very special: they are conserved quantities both for the wave equation and its upcoming nonlinear generalization, the KdV equation.

Box 2. Infinitely many conserved quantities for the linear wave equation.

Since $L = -\partial^2 + u(x,t)$ and $A = c\partial$ are unbounded differential operators, we do not try to make sense of tr L^n to find conserved quantities by the method of §4 of [1]. Nevertheless, conserved quantities *can* be obtained from the pair of equations $L\psi = \lambda\psi$ and $\psi_t = -A\psi$ (see §3 of [1]). Indeed, suppose we put $\lambda = k^2$ and change variables from the wavefunction ψ to a new function ρ defined via the transformation

$$\psi(x,t;k) = \exp\left[-ikx + \int_{-\infty}^{x} \rho(y,k,t) \, dy\right]. \tag{8}$$

Then, by studying the quantum mechanical scattering problem for a plane wave with one dimensional wavevector k in the potential u (assumed to vanish at $\pm\infty$), it can be shown that $\int_{-\infty}^{\infty} \rho(x,k,t) \, dx$ (which is the reciprocal of the transmission amplitude) is conserved in time for any k. We will use this to find an infinite sequence of integrals of motion (in terms of u). Putting (8) in $L\psi = k^2\psi$, we get a Riccati-like equation relating ρ to u: $\rho_x + \rho^2 - 2ik\rho = u(x,t)$. Since ρ is a conserved density, so are the coefficients ρ_n in an asymptotic series in inverse powers of k: $\rho = \sum_{n=1}^{\infty} \rho_n(x,t)/(2ik)^n$ which is a bit like a 'semiclassical' expansion. Comparing coefficients of different powers of k, one finds

$$\text{at} \quad O(k^0) : \rho_1 = -u, \quad \text{at} \quad O(1/k) : \rho_2 = \partial\rho_1 \quad \text{and} \quad \text{at} \quad O(1/k^n) : \rho_{n+1} = \partial\rho_n + \sum_{m=1}^{n-1} \rho_m \rho_{n-m}. \tag{9}$$

Using this recursion relation we may express ρ_n in terms of u and its derivatives:

$$\begin{aligned} \rho_1 &= -u, \quad \rho_2 = -u_x, \quad \rho_3 = u^2 - u_{xx}, \quad \rho_4 = (2u^2 - u_{xx})_x, \\ \rho_5 &= -u_{4x} + 2(u^2)_{xx} + u_x^2 + 2uu_{xx} - 2u^3, \quad \rho_6 = \left(-u_{4x} + 18uu_{2x} - \frac{16}{3}u^3\right)_x \quad \text{etc.} \end{aligned} \tag{10}$$

Contd.

Box 2. ***Contd.***

The even coefficients integrate to zero while ρ_{2n+1} lead to nontrivial conserved quantities defined as

$$Q_n = \frac{(-1)^{n+1}}{2} \int_{-\infty}^{\infty} \rho_{2n+1}\, dx \quad \text{for} \quad n = 0, 1, 2 \ldots . \tag{11}$$

The first few of these conserved quantities for the wave equation are:

$$Q_0 = \int \frac{u}{2} dx, \quad Q_1 = \int \frac{u^2}{2} dx, \quad Q_2 = \int \left(\frac{u_x^2}{2} + u^3 \right) dx \quad \text{and} \quad Q_3 = \frac{1}{2} \int \left[5u^4 + 10uu_x^2 + u_{2x}^2 \right] dx. \tag{12}$$

3. Korteweg-de Vries (KdV) Equation

The KdV equation, with subscripts denoting partial derivatives,

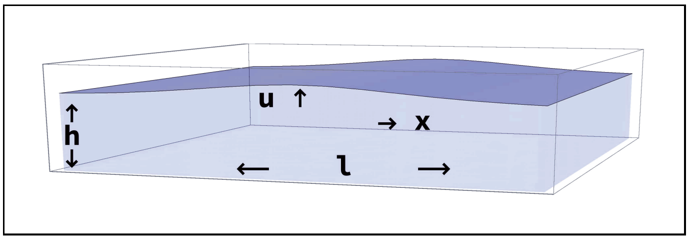
$$u_t - 6uu_x + u_{xxx} = 0, \tag{13}$$

describes long wavelength ($l \gg h$, 'shallow-water') surface waves of elevation $u(x, t) \ll h$ in water flowing in a narrow canal of depth h (see *Figure* 2).

The most famous solution of the KdV equation is the *soliton* $u = -\frac{c}{2}\mathrm{sech}^2\left[\frac{\sqrt{c}(x-ct)}{2}\right]$. It describes a localized solitary wave of depression that travels at velocity c while retaining its shape. Observation of such a wave was reported in 1834 by Scott Russell while riding a horse along the Edinburgh–Glasgow canal.

The KdV equation for the field u describes the evolution of infinitely many degrees of freedom labelled by points x lengthwise along the canal. While the nonlinear advection term uu_x can steepen the slope of a wave profile, the 'dispersive' u_{xxx} term tends to spread the wave out (see *Box* 3). A balance between the two effects can lead to localized solitary waves or 'solitons' that can propagate while maintaining their shape. What is more, two such solitons can collide and reemerge while retaining their shapes. These phenomena, which were discovered via laboratory and numerical experiments, suggested that the KdV equation may possess several constants of motion.

Figure 2. Surface wave profile in a canal.



Box 3. Nondispersive versus dispersive propagation.

A linear evolutionary partial differential equation (such as the wave equation) is nondispersive if the phase velocity $v_p(k) = \omega(k)/k$ of a plane wave solution $e^{i(kx-\omega(k)t)}$ is independent of the wavevector k. This happens if the angular frequency-wavevector dispersion relation $\omega = \omega(k)$ is linear. For the wave equation $u_t + cu_x = 0$, we have $\omega = ck$ so that $v_p = c$ is a constant implying nondispersive propagation. For a nondispersive equation, all Fourier components (labelled by k) travel at the same speed so that a wave packet does *not* spread out. For example, light waves in vacuum are nondispersive. On the other hand, free particle matter waves described by the Schrödinger equation $i\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\partial_x^2\psi$ are dispersive: $\omega = \hbar k^2/2m$ or $v_p(k) = \hbar k/2m$. Here, higher k modes travel faster and a wave packet broadens out with time.

In fact, the KdV equation admits some elementary conserved quantities [3, 4]. For instance, integrating (13) gives

$$\frac{d}{dt}\int_{-\infty}^{\infty} u\,dx = \int_{-\infty}^{\infty} \left(3u^2 - u_{xx}\right)_x\,dx = 0, \tag{14}$$

assuming $u \to 0$ as $x \to \pm\infty$. This leads to the conservation of the mean height $2Q_0 = \int_{-\infty}^{\infty} u\,dx$. Furthermore, one may check by differentiating in time and using (13) that

The conserved quantities P and E are related to symmetries of the KdV equation under space-time translations via a theorem of Emmy Noether. See Chapt. 1 of [5] for more on symmetries of the KdV equation.

$$2Q_1 = P = \int_{-\infty}^{\infty} u^2 dx \quad \text{and} \quad Q_2 = E = \int_{-\infty}^{\infty} \left(u^3 + \frac{u_x^2}{2}\right) dx \tag{15}$$

are also conserved. P and E can be interpreted as the momentum and energy of the wave. While these conservation laws could perhaps be guessed, in what came as a major surprise, in 1967–68, Whitham and then Kruskal and Zabusky discovered a fourth (Q_3 from *Box* 2) and fifth conserved quantity. Miura discovered yet more and the list grew to eleven conserved quantities. In fact, it was shown by Gardner, Greene, Kruskal and Miura that the KdV equation admits an infinite sequence of independent conserved quantities. These turn out to be the same as the Q_n of *Box* 2 for reasons that will soon become apparent.

At around the same time, in 1968, Peter Lax [6] proposed an (L, A) pair for the KdV equation. As for the wave equation in §2., L is the Schrödinger operator, but A is a third order operator (see *Box* 4 for an indication of how one arrives at A):

Interestingly, the work of Gardner, Greene, Kruskal and Miura was motivated by Kruskal and Zabusky's 1965 observation of 'recurrent behavior' and 'soliton scattering' in numerical solutions of the KdV equation.

$$L = -\partial^2 + u(x,t) \quad \text{and} \quad A = 4\partial^3 - 6u\partial - 3u_x. \tag{16}$$

As before, $L_t = u_t$. The commutator $[L, A]$ receives two contributions. With u' denoting u_x, the 3$^{\text{rd}}$ order term in A gives

$$[-\partial^2 + u, 4\partial^3] = -4(u''' + 3u''\partial + 3u'\partial^2). \tag{17}$$

As for the first order part of A, the calculation is essentially the same as in (6), with $\alpha = -3u$:

$$[-\partial^2 + u, -3(2u\partial + u_x)] = -3(u''' - 4u''\partial - 4u'\partial^2 - 2uu'). \tag{18}$$

Adding these, the differential operator terms in $L_t = [L, A]$ cancel, leaving us with the KdV equation (13):

The existence of a Lax representation for an evolutionary partial differential equation is usually associated with the presence of infinitely many independent conserved quantities.

$$u_t = [L, 4\partial^3 - 6u\partial - 3u_x] = -u''' + 6uu'. \tag{19}$$

The Lax representation helps us understand roughly why KdV admits infinitely many conserved quantities. Indeed, $L = -\partial^2 + u$ may be viewed as an infinite dimensional matrix, all of whose eigenvalues are conserved. In fact, the method of *Box* 2 for

Box 4. Lax pair for the KdV equation.

Here we adapt the method of *Box* 1 to explain the choice of the 3rd order differential operator $A = 4\partial^3 - 6u\partial - 3u_x$ in the KdV Lax pair (16). From §2., we know that $A = c\partial$ and the Schrödinger operator $L = -\partial^2 + u$ furnish a Lax pair for the linear wave equation. To find a Lax pair for the 3rd order KdV equation, we will retain $L = -\partial^2 + u$ with L_t being the multiplication operator u_t, while allowing for A to be of order higher than one. The simplest possibility is a 2nd order operator, but this does not work. Indeed, anti-symmetrization reduces it to a 1st order operator which is no different from (5) with $\alpha = -(e' + gf')$:

$$A_2 = e\partial^2 + f\partial g\partial - (e\partial^2 + f\partial g\partial)^\dagger = e\partial^2 + f\partial g\partial - \partial^2 e - \partial g\partial f = -(e' + gf')' - 2(e' + gf')\partial. \qquad (20)$$

The next possibility is a 3rd order operator. For simplicity, we try the operator $b\partial^3$ where b is a constant. Upon anti-symmetrizing,

$$A_3 = b\partial^3 - (b\partial^3)^\dagger = b\partial^3 + \partial^3 b = 2b\partial^3. \qquad (21)$$

As in *Box* 1, using the product rule and $[\partial, h] = h'$ we find that

$$[L, A_3] = [-\partial^2 + u, 2b\partial^3] = -2b(u''' + 3u''\partial + 3u'\partial^2). \qquad (22)$$

While this includes a u''' term, it lacks the uu' term in (13) and is not purely a multiplication operator. Here, A_1 from *Box* 1 comes to the rescue. Thus, let us consider $A = A_3 + A_1 = 2b\partial^3 + 2\alpha\partial + \alpha'$ so that

$$[L, A] = [-\partial^2 + u, 2b\partial^3 + 2\alpha\partial + \alpha'] = -\alpha''' - 2\alpha u' - 2bu''' - (6bu'' + 4\alpha'')\partial - (6bu' + 4\alpha')\partial^2. \qquad (23)$$

For $[L, A]$ to be a multiplication operator, the coefficients of ∂ and ∂^2 must vanish. Thus, $\alpha' = -(3/2)bu'$ which implies $\alpha = -(3/2)bu + \alpha_0$ for an integration constant α_0. Eliminating α, the Lax equation becomes

$$L_t = u_t = [L, A] = -\frac{b}{2}u''' + (3bu - 2\alpha_0)u'. \qquad (24)$$

Comparing with $u_t = 6uu_x - u_{3x}$ fixes $b = 2$ and $\alpha_0 = 0$ so that $A = 4\partial^3 - 6u\partial - 3u'$ as claimed. Note that we may add to A an arbitrary function of time (which would commute with L) without affecting the Lax equation.

finding conserved quantities for the wave equation from its (L, A) pair also applies to the KdV equation. What is more, since the two equations share the same Lax operator L, it turns out that they also possess the same set of conserved quantities Q_n. Moreover, treating Q_n as a sequence of 'energies' or Hamiltonians, one obtains the 'KdV' hierarchy of field equations. The linear wave and KdV equations are the first two in this hierarchy, while $u_t = u_{5x} - 10uu_{3x} - 20u_x u_{2x} + 30u^2 u_x$ is the third. The Schrödinger operator $L = -\partial^2 + u$ serves as a common Lax operator for all of them though the operator A (which enters through $\psi_t = -A\psi$) differs for the various members of this hierarchy. Remarkably, it turns out that the Q_n of *Box* 2 are integrals of motion for each of the equations in this hierarchy.

The Schrödinger operator $L = -\partial^2 + u$ serves as a common Lax operator for all the equations in the KdV hierarchy.

4. From Lax Pair to Zero Curvature Representation

The zero curvature representation generalizes the idea of a Lax pair to a wider class of nonlinear evolution equations for systems especially in one spatial dimension. To understand how this works, we change our viewpoint and regard the nonlinear Lax equation $L_t = [L, A]$ as a compatibility condition for the following pair of linear equations to admit simultaneous solutions:

The zero curvature representation generalizes the idea of a Lax pair to a wider class of nonlinear evolution equations in one spatial dimension. The reason for the name zero curvature is explained in *Box* 5.

$$L\psi = \lambda\psi \quad \text{and} \quad \psi_t = -A\psi \quad \text{with} \quad \lambda \quad \text{a constant.} \qquad (25)$$

Indeed, by differentiating $L\psi = \lambda\psi$ in time and using the second equation, it is verified that for the eigenvalue λ of L to be time-independent, L and A must satisfy the Lax equation $L_t = [L, A]$. Unlike in §3 of [1], here there is no need for λ to be a nondegenerate eigenvalue of L.

The transformation from the KdV Lax pair (L, A) to the matrices (U, V) is somewhat analogous to the one from Newton's second order equation to Hamilton's first order equations which treat position and momentum on a more equal footing.

In the case of the KdV equation (13), $L = -\partial_x^2 + u$ involves 2nd order space derivatives, so that the two equations in (25) are somewhat asymmetrical. There is a way of replacing (25) with a more symmetric pair of linear equations involving only 1st order derivatives:

$$\partial_x F = UF \quad \text{and} \quad \partial_t F = VF. \qquad (26)$$

The price to be paid is that U and V are now square matrices and F a column vector (of size equal to the order of the differential operator L) whose components depend on location. The matrix elements of U and V depend on the dynamical variables (such as u for KdV) as well as on the eigenvalue λ which is now called the spectral parameter.

$\partial_x - U$ and $\partial_t - V$ may be viewed as the space and time components of a 'covariant derivative'. Thus, the auxiliary linear equations (26) require that every vector field $F(x, t)$ is 'covariantly' constant.

Eqn. (26) is called the auxiliary linear system of equations. It is overdetermined (more equations than unknowns) in the sense that U and V must satisfy a compatibility (consistency) condition for solutions F to exist. Indeed, equating mixed partials $\partial_x\partial_t F = \partial_t\partial_x F$, we get the consistency condition

$$\partial_t U - \partial_x V + [U, V] = 0. \tag{27}$$

The original nonlinear evolution equations are said to have a zero curvature representation if they are equivalent to (27) for some pair of matrices U and V. Before explaining how this scheme may be used to find conserved quantities, let us use the KdV equation to provide an example.

For a Lax operator which is an n^{th}-order spatial differential operator, we may express the Lax equation as a system of n first order equations for the column vector $(\psi, \psi_x, \psi_{xx}, \cdots, \psi_{(n-1)x})^T$ comprising the first $(n-1)$ derivatives of the eigenfunction ψ. [Here G^T denotes the transpose of the vector G]. U and V then become $n \times n$ matrices. For the KdV equation, $n = 2$.

To find U for KdV, we write the eigenvalue problem for the Lax operator $(-\partial_x^2 + u)\psi = \lambda\psi$ as a pair of first order equations by introducing the column vector $F = (f_0, f_1)^T = (\psi, \psi_x)^T$:

$$\partial_x \begin{pmatrix} f_0 \\ f_1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ u-\lambda & 0 \end{pmatrix} \begin{pmatrix} f_0 \\ f_1 \end{pmatrix} \quad \Rightarrow \quad U = \begin{pmatrix} 0 & 1 \\ u-\lambda & 0 \end{pmatrix}, \tag{28}$$

upon comparing with (26). Next, we use $\psi_t = -A\psi$ with $A\psi = 4\psi_{xxx} - 6u\psi_x - 3u_x\psi$ to find V such that $\partial_t F = VF$. We may express $\psi_t = -A\psi$ as a system of two first order ODEs. First, we differentiate $L\psi = \lambda\psi$ in x to express ψ_{xxx} as $u_x\psi + u\psi_x - \lambda\psi_x$. Thus, $A\psi$ can be written in terms of ψ and ψ_x:

$$A\psi = -2(u + 2\lambda)\psi_x + u_x\psi. \tag{29}$$

The parameter λ is known as the *spectral parameter* because in the KdV case, it arose as an eigenvalue (part of the spectrum) of the Lax operator L.

Next, using $F = (\psi, \psi_x)^T$, $\psi_t = -A\psi$ takes the form

$$\partial_t \begin{pmatrix} f_0 \\ f_1 \end{pmatrix} = V \begin{pmatrix} f_0 \\ f_1 \end{pmatrix} \quad \text{with} \quad V = \begin{pmatrix} -u_x & 2(u+2\lambda) \\ 2u^2 - u_{xx} + 2u\lambda - 4\lambda^2 & u_x \end{pmatrix}. \tag{30}$$

Here, the second row of the matrix V is obtained by taking the x derivative of the first equation in (30) and using $L\psi = \lambda\psi$.

The parameter λ that appears in U and V originally arose as the eigenvalue of the Lax operator L. This explains the name *spectral parameter*. More generally, a zero curvature representation need not arise from a Lax pair and the corresponding spectral parameter λ may not admit an interpretation as an eigenvalue.

Box 5. Why the name 'zero curvature'?

Einstein's theory of gravity teaches us that a gravitational field is associated to space-time curvature. It turns out that an electromagnetic field is *also* associated to curvature, though not of space-time but of an 'internal' space. Now, the electric and magnetic fields **E** and **B** may be packaged in the components of the field strength: $F_{0i} = E_i/c$ and $F_{ij} = \sum_k \epsilon_{ijk} B_k$ for $1 \leq i, j, k \leq 3$, where c denotes the speed of light. Thus, the field strength is a measure of curvature. What is more, specializing to one spatial dimension ($x^0 = t, x^1 = x$) and introducing the scalar and vector potentials A_0 and A_1, we have $F_{01} = \partial_t A_1 - \partial_x A_0$. More generally, in the non-abelian version of electromagnetism relevant to the strong and weak interactions, A_0 and A_1 become matrices and the field strength acquires an extra commutator term: $F_{01} = \partial_t A_1 - \partial_x A_0 + [A_1, A_0]$. Now making the substitutions $A_1 \to U$ and $A_0 \to V$, we see that the consistency condition (27) states that the field strength or curvature of this generalized electromagnetic field vanishes. Hence the name zero curvature condition is used.

5. Conserved Quantities from the Zero Curvature Condition

Here, we will learn how the zero curvature representation may be used to construct conserved quantities. Let us consider the first of the auxiliary linear equations in (26) for the column vector F: $\partial_x F = U(x)F(x)$. Let us imagine solving this equation for F from an initial location x to a final point y. If $y = x + \delta x$ for small δx, then

$$F(x + \delta x) \approx [\mathbf{1} + \delta x\, U(x)]\; F(x). \tag{31}$$

More generally, linearity suggests that the solution may be written as $F(y) = T(y, x)F(x)$. Here $T(y, x)$ may be viewed as transforming $F(x)$ into $F(y)$ and is called the *transition matrix*. For this to work, $T(y, x)$ must satisfy the equation and 'boundary' condition

The transition matrix $T(y, x)$ transforms the vector field F at x to its value at y. Such a matrix is sometimes called a parallel transport operator.

$$\partial_y T(y, x; \lambda) = U(y; \lambda)T(y, x; \lambda) \quad \text{and} \quad T(x, x; \lambda) = \mathbf{1}. \tag{32}$$

Box 6. Time evolution operator and the ordered exponential.

In *Box* 2 of Part I [1], we encountered an equation for the time evolution operator $S(t)$

$$\dot{S} = -A(t)S, \quad \text{with the initial condition} \quad S(0) = \mathbf{1}, \quad \text{the identity matrix.} \tag{33}$$

The same equation also arises as the second of the auxiliary linear equations in (26) and as the Schrödinger equation in quantum mechanics for the time dependent 'Hamiltonian' $-i\hbar A(t)$. Here, we explain how this equation may be solved. When A is independent of time the solution is the matrix exponential $S = \exp(-At)$. However, for time-dependent A, this formula does not satisfy (33) if $A(t)$ at distinct times do not commute. To solve (33), we first integrate it in time form 0 to t to get an integral equation that automatically encodes the initial condition:

$$S(t) - \mathbf{1} = -\int_0^t A(t_1)S(t_1)\, dt_1. \tag{34}$$

S appears on both sides, so this is not an explicit solution. Iterating once, we get

$$S(t) = \mathbf{1} - \int_0^t dt_1\, A(t_1)\left(\mathbf{1} - \int_0^{t_1} dt_2\, A(t_2)S(t_2)\right). \tag{35}$$

Repeating this process, we get an infinite sum of multiple integrals,

$$S(t) = \mathbf{1} - \int_0^t dt_1\, A(t_1) + \int_0^t \int_0^{t_1} dt_1 dt_2\, A(t_1)A(t_2) - \cdots = \sum_{n=0}^{\infty} (-1)^n \int \cdots \int_{0<t_n<\cdots<t_1<t} dt_1 \cdots dt_n\, A(t_1)A(t_2)\cdots A(t_n). \tag{36}$$

Now, if we define time ordering denoted by the symbol T via

$$\mathrm{T}(A(t_1)A(t_2)) = \begin{cases} A(t_1)A(t_2) & \text{if} \quad t_1 \geq t_2 \\ A(t_2)A(t_1) & \text{if} \quad t_2 \geq t_1, \end{cases} \tag{37}$$

and use the identity $\int_{t_1>t_2} dt_1 dt_2 A(t_1)A(t_2) = \int_{t_2>t_1} dt_1 dt_2 A(t_2)A(t_1)$, we obtain

$$\int_0^t dt_1 \int_0^{t_1} dt_2 A(t_1)A(t_2) = \frac{1}{2}\int_0^t dt_1 \int_0^t dt_2\, \mathrm{T}(A(t_1)A(t_2)). \tag{38}$$

Thus, we have expressed an integral over a triangle in the t_1-t_2 plane as half the integral over a square. Similarly, for $n = 3$ we may express the integral over a pyramid as one/sixth of that over a cube. Proceeding this way, we get

$$\int_0^t dt_1 \int_0^{t_1} dt_2 \cdots \int_0^{t_{n-1}} dt_n\, A(t_1)A(t_2)\cdots A(t_n) = \frac{1}{n!}\int_0^t \cdots \int_0^t dt_1 dt_2 \cdots dt_n \mathrm{T}(A(t_1)A(t_2)\cdots A(t_n))$$
$$\text{so that} \quad S(t) = \sum_0^{\infty} \frac{(-1)^n}{n!} \int_0^t \cdots \int_0^t dt_1 dt_2 \cdots dt_n \mathrm{T}(A(t_1)A(t_2)\cdots A(t_n)) =: \mathrm{T}\exp\left[-\int_0^t A(t')dt'\right]. \tag{39}$$

This series is called a time-ordered exponential and denoted T exp. If time is replaced with a spatial coordinate, then it is called a path-ordered exponential and abbreviated 'P exp'.

This is obtained by inserting $F(y) = T(y, x)F(x)$ in the auxiliary linear equation $\partial_y F(y) = U(y)F(y)$ and requiring it to hold for any $F(x)$. In *Box* 6, we learn that the transition matrix $T(y, x)$ may be expressed (essentially by iterating (31)) as an *ordered exponential series* which we abbreviate as

$$T(y, x; \lambda) = \mathrm{P}\exp \int_x^y U(z; \lambda)\, dz. \qquad (40)$$

The transition matrix $T(y, x; \lambda)$ is a sort of exponential along a path from x to y. If the path is a closed loop that goes around a spatial circle once with $y = x$, then T is called the monodromy matrix.

For simplicity, we henceforth suppose that our one-dimensional system is defined on the spatial interval $-a \leq x \leq a$ with *periodic boundary conditions*, so that $U(-a) = U(a)$ and $V(-a) = V(a)$. Thus, we may view our spatial coordinate x as parametrizing a circle of circumference $2a$. So far, we have been working at one instant of time. It turns out that the transition matrix around the full circle ($x = -a$ to $y = a$), which is also called the monodromy matrix,

$$T_a(t, \lambda) = \mathrm{P}\exp \int_{-a}^{a} U(z; t, \lambda)\, dz \qquad (41)$$

has remarkably simple time evolution. In fact, using the derivative of the transition matrix (32) and the zero curvature condition (27), one may show (see *Box* 7 or §3 of Chapter 1 of [7]) that the transition matrix evolves according to:

$$\partial_t T(y, x; t) = V(y; t)T(y, x; t) - T(y, x; t)V(x; t). \qquad (42)$$

The time derivative of the transition matrix $T(y, x)$ is not quite a commutator (see Eqn. (42)). On the other hand, just like the Lax matrix L, the monodromy matrix $T(a, -a)$ evolves via a commutator (43) when periodic boundary conditions are imposed.

Now, specializing to $x = -a$ and $y = a$ and using periodic boundary conditions, we obtain an evolution equation for the monodromy matrix $T_a(t) = T(a, -a; t)$:

$$\partial_t T_a(t, \lambda) = [V(a; t, \lambda), T_a(t, \lambda)]. \qquad (43)$$

The trace of the monodromy matrix $\mathrm{tr}\ T_a(t, \lambda)$ is independent of time for any value of the spectral parameter λ and can be used to generate (infinitely many) conserved quantities.

We are now in familiar territory: this equation has the same structure as the Lax equation (1) upon making the replacements $T_a \mapsto L$ and $V \mapsto -A$. As explained in §3 of Part I [1], the spectrum of the Lax matrix L is independent of time. This immediately implies that the trace of the monodromy $\mathrm{tr}\ T_a(t, \lambda)$ is independent of time. Moreover, this is true for any value of the spectral parameter λ. Thus, if we expand $\mathrm{tr}\ T_a(\lambda)$ in a series in (positive and negative) powers of λ, then each of the coefficients is a conserved quantity. In many interesting cases such as the KdV and nonlinear Schrödinger equations, one obtains infinitely many conserved quantities in this way.

Box 7. Time evolution of the transition matrix $T(y, x; t)$

Recall that the transition matrix $T(y, x; t)$ 'propagates' vectors in the auxiliary linear space from x to y: $F(y; t) = T(y, x; t)F(x; t)$ and may be expressed as a path ordered exponential as in (40). To obtain Eqn. (42) for its time evolution, we first differentiate Eqn. (32) [$\partial_y T(y, x; t) = U(y; t)T(y, x; t)$] in time:

$$\partial_t \partial_y \, T(y, x; t) = \partial_t U(y; t) \, T(y, x; t) + U(y; t) \, \partial_t T(y, x; t). \tag{44}$$

Then we use the zero curvature condition $\partial_t U(y) - \partial_y V(y) + [U(y), V(y)] = 0$ and Eqn. (32) again to get:

$$\partial_t \partial_y \, T(y, x; t) = (\partial_y V)T + VUT - UVT + U(\partial_t T) = \partial_y(VT) + U(\partial_t T - V(y)T). \tag{45}$$

Thus, we have

$$\partial_y W(y, x; t) = U(y)W(y, x; t) \quad \text{where} \quad W(y, x; t) = \partial_t T - V(y)T. \tag{46}$$

Thus both $W(y, x; t)$ and $T(y, x; t)$ satisfy the *same* differential equation (32) though they obey different 'boundary conditions' $W(x, x; t) = -V(x)$ while $T(x, x; t) = I$. We now use this to check that $\tilde{W}(y, x; t) = -T(y, x; t)V(x)$ also satisfies the same differential equation with the desired boundary condition $\tilde{W}(x, x; t) = -V(x)$. Exploiting the uniqueness of solutions to (32) for a given boundary condition, we conclude that $W(y, x; t) = \tilde{W} = -T(y, x; t)V(x)$. Substituting this in the definition of $W(y, x; t)$ (46), we obtain the evolution equation (42) for the transition matrix: $\partial_t T(y, x; t) = V(y; t)T(y, x; t) - T(y, x; t)V(x; t)$.

6. Epilogue

Though it is not always possible or easy to find a Lax pair for a given system, it is possible to generate lots of Lax pairs and thereby discover systems with numerous conserved quantities. Some of these turn out to be interesting 'exactly solvable' or 'integrable' systems.

As one may infer from these examples, there is no step-by-step procedure to find a Lax pair for a given system or even to know whether it admits a Lax pair. One first needs to determine some properties of the system (say numerically, analytically or experimentally as happened with KdV) to develop a feeling for whether a Lax pair might exist. As a rule of thumb, equations whose trajectories are 'regular' or for which (some) analytic solutions can be obtained often do admit a Lax pair, while those that display irregular/chaotic behavior do not. Even if one suspects the presence of a Lax pair, finding one may not be easy and requires playing around with the equations as we have done for the harmonic oscillator, Euler top, wave equation and the KdV equation. However, if one *does* find a Lax pair, it opens up a whole new window to the problem and brings to bear new tools that can be applied to its understanding. Indeed, Lax pairs are the tip of an iceberg in the study of (Hamiltonian) dynamical systems. While it helps to have conserved quantities, one can do more if they are sufficiently numerous and generate 'commuting' flows on the state space (i.e., if their Poisson brackets vanish). In such cases, there is (at least in principle) a way of changing variables to so-called action-angle variables in which the solutions to the EOM may be written down by inspection! Moreover, continuum systems in *one spatial dimension* (such as the KdV, nonlinear Schrödinger and sine-Gordon equations) which have a Lax pair and an infinite tower of conserved quantities typically admit solitary wave solutions called solitons. Two such solitons can collide with each other and interact in a complicated way but emerge after the collision retaining their original shapes and speeds, thus mimicking the elastic scattering of particles. This *soliton scattering* behavior can be regarded as a generalization to nonlinear systems of the superposition principle for linear equations. These nonlinear field equations also admit a remarkable generalization of the Fourier transform technique of solving linear PDEs such as the heat or wave equations. This technique is called the 'inverse scattering transform' and can be used to solve the initial value problem of determining the fields at time t given their values at $t = 0$.

Conserved quantities can be particularly helpful in solving the equations of motion if they are sufficiently numerous and generate 'commuting' flows on the state space (i.e., if their Poisson brackets vanish).

When solitary waves undergo 'soliton scattering', they emerge from the collision region retaining their shapes and speeds (as in the elastic scattering of particles) despite interacting in a complicated manner.

Acknowledgements

We thank an anonymous referee for useful comments and references. This work was supported in part by the Infosys Foundation, J N Tata Trust and grants (MTR/2018/000734, CRG/2018/002040) from the Science and Engineering Research Board, Govt. of India.

Suggested Reading

[1] G S Krishnaswami and T R Vishnu, The idea of a Lax pair - Part I: Conserved quantities for a dynamical system, *Resonance*, Vol.25, No.12, pp.1705–1720, 2020.

[2] D J Griffiths, *Introduction to Quantum Mechanics*, Second Edition, Pearson Education, Dorling Kindersley Indian Ed., New Delhi (2005).

[3] P G Drazin and R S Johson, *Solitons: An Introduction*, Cambridge University Press, Cambridge, 1989.

[4] A Das, *Integrable Models*, World Scientific Publishing, Singapore (1989).

[5] T Miwa, M Jimbo and E Date, *Solitions: Differential Equations, Symmetries and Infinite Dimensional Algebras*, Cambridge University Press, Cambridge, 2011.

[6] P D Lax, Integrals of nonlinear equations of evolution and solitary waves, *Comm. Pure Appl. Math.*, 21, 467, 1968.

[7] L D Faddeev and L A Takhtajan, *Hamiltonian Methods in the Theory of Solitons*, Springer-Verlag, Berlin, 1987.

Address for Correspondence
Govind Krishnaswami[1]
T R Vishnu[2]
Chennai Mathematical Institute
H1 SIPCOT IT Park
Siruseri, Kelambakkam 603 103
Tamil Nadu.
Email: [1]govind@cmi.ac.in
[2]vishnu@cmi.ac.in