

Learning Lax Pairs: Revisiting the Classical Paradigm

Jimmie Adriaola^{*1,2}, Gino Biondini³, Wei Zhu⁴, and Panayotis G. Kevrekidis^{5,6,7}

¹*School of Mathematical and Statistical Sciences, Arizona State University, Tempe, AZ*

²*Simon A. Levin Mathematical, Computational and Modeling Sciences Center, Arizona State University, AZ*

³*Department of Mathematics, State University of New York at Buffalo, Buffalo, NY*

⁴*School of Mathematics, Georgia Institute of Technology, Atlanta, GA*

⁵*Department of Mathematics and Statistics, University of Massachusetts Amherst, 01003 Amherst, MA*

⁶*Department of Physics, University of Massachusetts Amherst, 01003 Amherst, MA*

⁷*Department of Mechanical Engineering, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul 08826, South Korea*

August 6, 2026

This work is dedicated to the memory of Francesco Calogero.

Abstract

A Lax pair (L, P) is sometimes thought of as a structural certificate, in that the spatial operator L carries the spectral data of an integrable system, and its isospectral evolution under $\partial_t L = [L, P]$ encodes the nonlinear dynamics. Yet, experience shows that the correspondence between equations and Lax pairs is much more nuanced than this picture suggests. Equations can admit Lax pairs that fail to encode the expected integrable structure. This paper probes that anomalous corner of the Lax pair landscape through five case studies (the Euler top, the free Schrödinger equation, the inviscid Burgers equation, the shallow water system, and the Korteweg–de Vries equation), each illustrating a different way the link to integrability can be distorted. The approach combines analytical calculations with the Sparse Identification of Lax Operators (SILO) framework, which proved useful throughout, in some cases confirming the textbook pair and in others surfacing alternatives worth understanding on their own terms. The recurring lesson across the five cases is that compatibility underdetermines the Lax representation, so that anomalous pairs are regular features of the landscape rather than pathologies. Notably, we show that a spectrally degenerate Korteweg–de Vries Lax pair, classified as fake by standard criteria, still generates the full conservation hierarchy through its operator algebra, which shows that a blunt dichotomy between true and fake Lax pairs can be too reductive.

1 Introduction

The standard picture of a Lax pair comes in several forms [1, 4, 20, 22, 44]. In the original operator version [44], one seeks operators L and P acting on a fixed Hilbert space and depending on a field u , such that

$$\partial_t L = [L, P] \tag{1.1}$$

^{*}Corresponding author: jimmie.adriazola@asu.edu

is equivalent to the evolution equation for u . The isospectral flow generated by (1.1) produces an infinite family of conserved quantities through the spectral data of L , and, under favorable circumstances, opens the door to the inverse scattering transform [2, 27]. The same compatibility logic applies when L and P are finite-dimensional matrices (as in the Manakov pair for the Euler top [51]), symbols on phase space (with the commutator replaced by a Poisson bracket, as in dispersionless hierarchies [66]), or connection forms (with the commutator replaced by a curvature, as in zero-curvature representations [67]). The correspondence between integrable equations and Lax pairs, in any of these forms, has been a productive organizing principle in integrable systems theory [1, 19, 21, 22], though one with subtleties [13].

What the standard picture leaves implicit is that the same equation can admit different Lax pairs, and the same Lax pair can be compatible with different equations. For example, a single integrable equation typically admits many inequivalent Lax pairs, related by gauge transformations [22], spectral shifts [51, 52], or less structured redundancies [13]. Some of these encode the same spectral content in different coordinates. Others do not, as was already recognized by Calogero and Nucci’s “Lax pairs galore” [13]. Calogero and Nucci showed that a pair satisfying (1.1) need not certify integrability at all, and Sakovich [60], and later Butler and Hay [12], gave systematic methods to detect such “fake” Lax pairs. Following these previous works, we call a pair fake if it satisfies a Lax compatibility relation, such as (1.1), on solutions of the equation but its spectral data certify no integrable structure and support no inverse scattering transform. Additionally and throughout, we use the word anomalous as an umbrella term for any pair that satisfies a Lax compatibility relation yet departs from the classical spectral picture, whether by spectral collapse or by hiding an entire parametric family behind a single base pair. We use it informally and indicate the individual departures as they arise throughout the paper.

The phenomenon has practical consequences. Anomalous pairs have been identified in discrete integrable equations [32], in modified Korteweg–de Vries (KdV) families [11], in the 2D and 3D Euler equations of incompressible fluids [26, 47, 48], and in the broader Calogero–Nucci construction that produces a pair from any conservation law [13]. Any method that tries to discover a Lax pair [3, 18, 40, 49] as a structural object for scientific applications, whether for the nonlinear Fourier transform developed for fiber-optic communication [36, 61, 63], the nonlinear-spectrum analysis of measured ocean records, which classifies rogue waves [45, 57] and has lately resolved the first soliton-gas sea states observed in the deep open ocean [46], reduced-order modeling along the lines of [29], neural surrogates that exploit Lax structure [58], symbolic-computation pipelines that produce Lax pairs algorithmically [8], or in a principled data-driven fashion [3] the practitioner has to reckon with the possibility that the discovered pair is useless.

In parallel, the existence of Lax-like structures that do not conform to the classical spectral picture is, in places, already well developed in the literature. For KdV alone, Miura, Gardner, and Kruskal [53, 54] showed that the conservation hierarchy follows from a generating-function argument that does not require the Schrödinger spectral problem. The Lambert–Musette Bell-polynomial formalism [23, 30, 42, 43] recasts bilinearizations, Bäcklund transformations, and Lax pairs of KdV-type equations through Bell polynomials, exposing algebraic content that the operator-theoretic presentation does not make visible. In the pseudopotential and prolongation-structure tradition [64], first-order scalar auxiliary equations generate conservation laws for a range of integrable PDEs without passing through a second-order operator.

Our work sits alongside these developments and revisits the landscape of Lax representations via a computational approach using the Sparse Identification of Lax Operators (SILO) framework, developed in [3] and revisited here. Concentrating our attention on five selected case studies, we investigate the algebraic and structural features that distinguish one representation from another,

and on what a computational search turns up when it is not constrained to land on a pair that fits a pre-existing template. The SILO framework formulates the search for Lax pairs as a structured sparse regression problem: a library of candidate operators is combined against the compatibility residual of (1.1), in any of its forms, evaluated pointwise in phase or parameter space through the chain rule, with sparsity regularization selecting parsimonious solutions and problem-specific penalties preventing degenerate ones. At the same time, we emphasize that SILO is not the only tool of this paper; it is an instrument that paves numerous further directions. For instance, in what follows, SILO may lead to the straightforward verification of a classical pair, it may spearhead the surfacing of genuinely anomalous Lax pair structure or it may suggest some Lax pair modification that is subsequently worked out by hand.

The paper is organized as follows. Section 2 treats the Euler top, where an unconstrained sparse search recovers a dynamically correct but spectrally trivial pair whose invariants fail to generate the full Liouville structure, a finite-dimensional instance of the fake Lax pair phenomenon in its cleanest form. A Manakov-type spectral shift [51] rescues the situation, and the shift itself is computed by linear algebra rather than further numerics. Section 3 studies the free Schrödinger equation, where two different nondegeneracy penalties in the SILO loss return two structurally different pairs, the textbook pair [25] and another, spectrally degenerate pair in a precise algebraic sense. Section 4 turns to the inviscid Burgers equation, where a symbolic regression on the Clairaut compatibility identity led us to an infinite-dimensional gauge freedom parametrized by an arbitrary function of one variable, and across spatial orders to a continuum $S_x^m = au + \lambda$ of auxiliary systems valid for every $m \in \mathbb{R}_{>0}$. Section 5 treats the shallow water equations on two registers. Within the Laurent layer of dispersionless Lax functions, the classical Brunelli–Das representation [10] is the only compatible pair our sparse search returned. Outside the Laurent layer, the Calogero–Nucci hodograph mechanism applied to an Euler–Poisson–Darboux equation in Riemann coordinates produces an infinite-dimensional family of compatible auxiliary systems, encompassing the polynomial Whitham hierarchy of conserved densities, an algebraic spectral-discriminant continuum, and a dressing transform that links them.

Section 6 focuses on the first-order scalar operator $L = u + \partial_x$ that has recently been revealed for the KdV equation via SILO in [3]. Although, to the best of our knowledge, this pair does not appear in the literature, the existence of such a pair for any scalar equation in conservation form ($u_t = q_x$) is, in fact, classical. This traces to separate discussions with Manakov and Ablowitz in the 1980s [34], transmitted through the Calogero–Nucci construction [13], where the prevailing verdict was that these Lax pairs were useless. Here we show that, while the spectrum of L is trivial by every classical measure (empty L^2 point spectrum, periodic eigenvalues encoding only the mean of u , monodromy reducing to the mass), powers of L nevertheless generate a ring of differential polynomials, the complete exponential Bell polynomials in the derivatives of u , and the classical KdV conservation laws appear as the kernel of the KdV time derivative on this ring at successive subscript-sum levels. The Bell polynomial connection to KdV-type equations is well developed in the Lambert–Musette–Springael formalism [43], where these polynomials enter through logarithmic linearization in the variable $v = \ln \phi$ with ϕ a Schrödinger eigenfunction, so that v_x plays the role of a Riccati pseudopotential. Our framing is different, however. Taking $v = \int u dx$, the gauge identity $L^s = e^{-v} \partial_x^s e^v$ identifies the Bell polynomials with the zeroth-order parts of L^s acting on the constant function. Through this identification, the Lax equation $\partial_t L = [L, P]$ descends to a $\mathbb{Q}$ -linear derivation on the ring of integrated ρ -monomials modulo total x -derivatives. Its kernel at each subscript-sum level is the KdV conservation hierarchy, extracted by finite-dimensional linear algebra over $\mathbb{Q}$.

We suggest keeping in mind the following while reading this work. First, this is not a paper about discovering integrability. Every equation considered here is classically known to be integrable

(or, in the Burgers case, linearizable). The object of study is the Lax representation itself: what forms it can take, what it can and cannot encode, and how algebraic and spectral content trade off across different Lax gauges. Second, SILO appears throughout the paper but the method itself is not the central point. Readers interested in SILO as a methodology are directed to [3], where the framework is carefully developed from the ground up. The broader argument this paper makes is that anomalous Lax pairs are regular features of the integrable landscape that any data-driven discovery or machine learning pipeline will inevitably encounter, and whose structure deserves to be understood on its own terms if future searching is to land on pairs that carry real content for scientific applications.

2 The Euler Top: Fake Lax Pairs in Finite Dimensions

The cleanest finite-dimensional illustration of an anomalous Lax pair lives within the Euler top. The system is simple, the classical Lax pair is well known, and yet an unconstrained search for Lax representations returns a pair that satisfies the Lax equation exactly while carrying strictly less information than the classical one. The phenomenon is a genuine feature of the Lax correspondence in the finite-dimensional Lie–Poisson setting, and it admits a classical remedy through a Manakov-type spectral shift [51]. We present both the degenerate pair and the remedy here, and in doing so fix the vocabulary of fake, degenerate, and spectrally incomplete that will recur throughout the paper.

2.1 The system and its classical Lax representation

The angular velocities Ω_i of the Euler top satisfy

$$I_i \dot{\Omega}_i = (I_j - I_k) \Omega_j \Omega_k, \quad (i, j, k) \text{ cyclic in } (1, 2, 3), \quad (2.1)$$

where $I_i > 0$ are the principal moments of inertia. As a Hamiltonian system evolving on the dual Lie algebra $\mathfrak{so}(3)^*$, the Euler top admits conserved quantities tied directly to its underlying geometric structure. In addition to the kinetic energy, the flow preserves the quadratic Casimir $\|M\|^2 := \sum_{i=1}^3 M_i^2$. The preservation of $\|M\|^2$ reflects the degeneracy of the Lie–Poisson structure on $\mathfrak{so}(3)^*$, which foliates phase space into symplectic leaves given by the level sets of the Casimir. Together, these invariants constrain the motion to the intersection of two quadratic surfaces in $\mathbb{R}^3$, providing a geometric manifestation of integrability and sharply restricting the admissible dynamics [7, 37].

The Euler top admits a Lax pair [51]. Indeed, Equation (2.1) is compatible with

$$L = \begin{pmatrix} 0 & -M_3 & M_2 \\ M_3 & 0 & -M_1 \\ -M_2 & M_1 & 0 \end{pmatrix}, \quad P = \begin{pmatrix} 0 & -\Omega_3 & \Omega_2 \\ \Omega_3 & 0 & -\Omega_1 \\ -\Omega_2 & \Omega_1 & 0 \end{pmatrix}, \quad (2.2)$$

with $M_i = I_i \Omega_i$, so that the Lax equation $\dot{L} = [L, P]$ reproduces the dynamics. Although this formulation correctly encodes the equations of motion, it does not capture the full integrable structure of the Euler top. The Euler top has two functionally independent first integrals: the Hamiltonian $H = \frac{1}{2} \sum_{i=1}^3 I_i \Omega_i^2 = \frac{1}{2} \sum_{i=1}^3 M_i^2 / I_i$ and the Casimir $\|M\|^2 = \sum_{i=1}^3 M_i^2$. On $\mathfrak{so}(3)^* \cong \mathbb{R}^3$ the symplectic leaves are the spheres $\|M\|^2 = \text{const}$, each of dimension two, so a single integral in involution suffices for Liouville integrability on a leaf, and H plays that role.

The pair (2.2) reproduces the equations of motion, but its spectral data carries less information, that is, it recovers the Casimir $\|M\|^2$ but not the Hamiltonian H . Observe that the matrix L is skew-symmetric with eigenvalues 0 and $\pm i\|M\|$, so $\text{tr}(L^2) = -2\|M\|^2$ and every higher trace invariant is determined by $\|M\|^2$. The Hamiltonian, which weighs the M_i^2 by $1/I_i$, is not among them. The

pair (2.2) is therefore spectrally incomplete. Its Lax equation $\dot{L} = [L, P]$ reproduces the equations of motion, so the dynamics is correct, but the invariants of L recover only $\|M\|^2$, which fixes the sphere $\|M\|^2 = \text{constant}$ on which the motion lies, and never H , the integral that selects the trajectory within it. Recovering the Casimir but not the Hamiltonian does not certify Liouville integrability [7, 59].

The classical remedy, due to Manakov [51], introduces a diagonal matrix

$$J = \frac{1}{2} \text{diag}(I_2 + I_3 - I_1, I_1 + I_3 - I_2, I_1 + I_2 - I_3), \quad (2.3)$$

under which the shifted Lax equation

$$\frac{d}{dt}(L + \lambda J^2) = [L + \lambda J^2, P + \lambda J]$$

holds identically in $\lambda \in \mathbb{C}$. The matrix $\hat{L} = L + \lambda J^2$ is then isospectral, so every $\text{tr } \hat{L}^n$ is conserved, and the two lowest traces already carry both integrals: $\text{tr } \hat{L}^2$ returns the Casimir $\|M\|^2$ and $\text{tr } \hat{L}^3$ returns the Hamiltonian H (the computation is collected at the end of this subsection). Higher traces add nothing functionally independent on the three-dimensional top. The spectral parameter is what lifts the pair (2.2) from a dynamical certificate to a Liouville certificate. This is the first structural lesson we encounter: a Lax pair and its spectral completion are different objects. We also learn that nothing here demands that every pair admit a spectral completion.

2.2 Naive searches yield rank-deficient Lax pairs

For what follows, it is useful to express $\dot{L}$ through its Hamiltonian chain rule. The Hamiltonian $H = \frac{1}{2} \sum_{i=1}^3 I_i \Omega_i^2$ generates the equations of motion via the Lie–Poisson bracket, written here in the convention $\dot{f} = \{H, f\}$, so that $\dot{\Omega}_i = \{H, \Omega_i\}$, where $\{\cdot, \cdot\}$ denotes the bracket on $\mathfrak{so}(3)^*$ pulled back through $M_i = I_i \Omega_i$. Applying the chain rule and noting that the moments of inertia I_i are constant,

$$\frac{dL}{dt} = \sum_{i=1}^3 \frac{\partial L}{\partial \Omega_i} \dot{\Omega}_i = \sum_{i=1}^3 \frac{\partial L}{\partial \Omega_i} \{H, \Omega_i\} = [L, P]. \quad (2.4)$$

This valuable identity expresses the Lax equation as a pointwise constraint in phase space. It avoids time discretization entirely and converts the search for Lax pairs into an operator-identification problem evaluated on samples of Ω . Every sparse regression in this paper uses some version of this identity, either directly or in spirit.

Now, suppose we assume nothing about the structure of the Lax pair beyond a linear dependence on the state variables,

$$\tilde{L}_{i,j} = \sum_k \xi_{i,j,k} \Omega_k, \quad \tilde{P}_{i,j} = \sum_k \zeta_{i,j,k} \Omega_k, \quad (2.5)$$

and minimize the natural relative residual

$$\mathcal{J}_{\text{loss}}(\tilde{L}, \tilde{P}) = \mathbb{E}_{\Omega \sim \mathcal{P}} \frac{\sum_{j,k} \left(\sum_i \frac{\partial \tilde{L}}{\partial \Omega_i} \{H, \Omega_i\} - [\tilde{L}, \tilde{P}] \right)_{j,k}^2}{\sum_{j,k} \left(\sum_i \frac{\partial \tilde{L}}{\partial \Omega_i} \{H, \Omega_i\} \right)_{j,k}^2}$$

with a sparsity penalty,

$$\min_{\eta \in \mathbb{R}^N} \mathcal{J}_{\text{natural}}[\eta] := \min_{\eta \in \mathbb{R}^N} \mathcal{J}_{\text{loss}}[\eta] + s \mathcal{S}[\eta], \quad (2.6)$$

where $\mathcal{P}$ denotes the uniform distribution on the unit cube centered at the origin, η collects the 54 parameters $\xi_{i,j,k}, \zeta_{i,j,k}$, and $\mathcal{S}$ can be any viable sparsity functional, e.g., $\|\eta\|_{l^1(\mathbb{R}^N)}$. The denominator normalization suppresses the trivial solution $\tilde{L} \equiv 0$.

What comes back is not the classical pair (2.2). Instead, repeatedly and across random initializations, SILO returns rank deficient Lax pairs. For purposes of this discussion, one such pair is

$$L_{\text{deficient}} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \Omega_1 & \Omega_2 \\ 0 & \frac{I_2(I_3 - I_2)}{I_1(I_3 - I_1)}\Omega_2 & -\Omega_1 \end{pmatrix}, \quad P_{\text{deficient}} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \frac{I_3 - I_1}{2I_2}\Omega_3 \\ 0 & \frac{I_2 - I_3}{2I_1}\Omega_3 & 0 \end{pmatrix}. \quad (2.7)$$

A direct calculation confirms that this pair satisfies $\dot{L}_{\text{deficient}} = [L_{\text{deficient}}, P_{\text{deficient}}]$ exactly, with only 6 of 54 parameters active. From the perspective of $\mathcal{J}_{\text{natural}}$, it is the best possible solution: perfect residual with maximal sparsity. From the perspective of integrability, it is useless. The matrix $L_{\text{deficient}}$ is rank-two, Ω_3 does not appear in it at all, and its spectral data do not recover either the Casimir or the Hamiltonian.

This is an even worse dilemma than being a fake Lax pair in the strict sense of [12, 32]: the compatibility condition is satisfied, yet the integrable structure is not certified and the dynamics are not even reproduced. We believe that this phenomenon is really a consequence of a bad search geometry, i.e., a generic necessary-but-not-sufficient objective that is given a generic hypothesis. These generic searches consistently trip over rank-deficient pairs that populate the Lax landscape. Clearly, the route to discovery of more meaningful Lax pairs requires more principled searching.

2.3 Recovering the classical structure

Two structural constraints suffice to rule out (2.7) and restore the classical pair. First, we enforce skew-symmetry in (2.5), i.e., $\tilde{L} = -\tilde{L}^\top$ and $\tilde{P} = -\tilde{P}^\top$, the superscript $\top$ denoting matrix transpose. In addition to fixing the geometry of the hypothesis, we have the added benefit of reducing the parameter count from 54 to 18 while eliminating the degenerate pair (2.7) from the hypothesis class. This does not rule out every possible rank deficient Lax pair, therefore, in addition, we add a penalty that forces every variable Ω_k to remain represented in both operators,

$$\mathcal{J}_{\text{full}} = \prod_{k=1}^3 \left\| \frac{\partial \tilde{L}}{\partial \Omega_k} \right\|_F^2 \left\| \frac{\partial \tilde{P}}{\partial \Omega_k} \right\|_F^2, \quad (2.8)$$

where $\|A\|_F^2 := \text{tr}(A^\top A)$. The product structure is essential in that if any Ω_k is dropped from either operator, the corresponding factor vanishes and the reciprocal $\mathcal{J}_{\text{full}}^{-1}$ blows up.

The full optimization problem becomes

$$\min_{\eta \in \mathbb{R}^N} \mathcal{J}_{\text{Euler}}[\eta] = \min_{\eta \in \mathbb{R}^N} \delta \mathcal{J}_{\text{full}}^{-1}[\eta] + (1 - \delta) \mathcal{J}_{\text{loss}}[\eta] + s \mathcal{S}[\eta], \quad (2.9)$$

with $\delta \in (0, 1)$ balancing nondegeneracy against compatibility, $\mathcal{S}[\eta] = \|\eta\|_{\ell^1}$ the sparsity penalty, and $s > 0$ its strength. Empirically, both ingredients are necessary. We emphasize again that without the Frobenius product penalty, numerical results consistently collapse onto fake pairs with the same structural character as (2.7), different in the specific variable dropped and coefficients present but identical in their uselessness. Meanwhile, without the skew-symmetry constraint layered on top of the Frobenius penalty, superfluous pairs begin to appear that are not genuinely skew and do not carry the $\mathfrak{so}(3)^*$ structure. Only when both are imposed does SILO reliably recover a pair with the same structure as the classical pair (2.2).

2.4 From the discovered pair to the Manakov completion

The Frobenius-regularized SILO search of the previous subsection recovers a skew-symmetric base pair (L_*, P_*) with the structure of (2.2). This pair reproduces the Euler top but, as noted above,

fails to generate the full Liouville hierarchy. The point of this subsection is that, given (L_*, P_*) , that are skew-symmetric, the spectral completion does not need to be accomplished by further numerics. The completion is the solution of a linear system posed inside the Lie algebra, and once the ansatz is fixed there is no residual numerical search. We make this explicit to mark the division of labor: SILO supplies the full-rank skew symmetric skeleton, and the spectral completion is pure linear algebra in the spirit of Manakov [51] and its semisimple generalization by Mishchenko and Fomenko [52].

We review the method here for completeness. Posit the polynomial family linear in λ ,

$$L(\lambda) = L_* + \lambda L_1, \quad P(\lambda) = P_* + \lambda P_1, \quad (2.10)$$

with L_1, P_1 constant matrices to be determined. Requiring $\dot{L}(\lambda) = [L(\lambda), P(\lambda)]$ for all λ and matching orders gives

$$\lambda^0: \quad \dot{L}_* = [L_*, P_*], \quad (2.11)$$

$$\lambda^1: \quad [L_*, P_1] + [L_1, P_*] = 0, \quad (2.12)$$

$$\lambda^2: \quad [L_1, P_1] = 0. \quad (2.13)$$

Equation (2.11) holds by construction of (L_*, P_*) . The remaining two are a linear system in (L_1, P_1) , decoupled from Ω sampling. To extract the Ω -independent content of (2.12), expand L_* and P_* in the standard basis E_k of $\mathfrak{so}(3)$:

$$L_*(\Omega) = \sum_{k=1}^3 I_k \Omega_k E_k, \quad P_*(\Omega) = \sum_{k=1}^3 \Omega_k E_k.$$

Substituting into (2.12) and using linear independence of the Ω_k , the coupled condition decouples into three matrix equations:

$$I_k [E_k, P_1] = [E_k, L_1], \quad k = 1, 2, 3. \quad (2.14)$$

Now, restrict (2.14) to the diagonal commuting ansatz $L_1 = \text{diag}(l_1, l_2, l_3)$, $P_1 = \text{diag}(p_1, p_2, p_3)$, which automatically satisfies (2.13). A direct computation (using that $[E_k, \text{diag}(d_1, d_2, d_3)]$ has a single off-diagonal entry $d_j - d_l$, where (j, l) are the indices other than k) reduces (2.14) to three scalar relations,

$$I_k(p_j - p_l) = l_j - l_l, \quad (k, j, l) \text{ cyclic}.$$

The order- λ relations $I_k(p_j - p_l) = l_j - l_l$ are the only constraints. Summing the three cyclic copies cancels the l 's and leaves a single relation on P_1 , solved by $p_k = \text{const}$ and by $p_k = I_k$, which span its solution space; hence $p_k = a + b I_k$. Back-substituting gives $l_j - l_l = b I_k(I_j - I_l)$, so $l_k = d - b I_j I_l = d - b I_1 I_2 I_3 / I_k$. The completion is therefore

$$P_1 = \text{diag}(a + b I_k), \quad L_1 = \text{diag}\left(d - b \frac{I_1 I_2 I_3}{I_k}\right),$$

with a and d the inert shifts $P_1 \rightarrow P_1 + aI$, $L_1 \rightarrow L_1 + dI$ and b an overall scale. Requiring the Hamiltonian in the spectrum fixes only $b \neq 0$: the order- λ part of $\text{tr}(L^3)$ is $3 \sum_k l_k (M_k^2 - \|M\|^2) = -6b I_1 I_2 I_3 H$ modulo the Casimir. The constants a and d are inert identity shifts and b is an overall scale, so on the three-dimensional top the completion is fixed only up to this gauge, and any nonzero b already certifies integrability.

The simplest choice makes the point. Take $a = d = 0$ and $b = 1$, so that

$$P_1 = \text{diag}(I_1, I_2, I_3), \quad L_1 = -I_1 I_2 I_3 \text{diag}(I_1^{-1}, I_2^{-1}, I_3^{-1}).$$

As discussed, $\text{tr}(L^2)$ returns the Casimir at order λ^0 , and $\text{tr}(L^3)$ returns $-6I_1I_2I_3H$ at order λ^1 , so both functionally independent integrals are present and the flow is Liouville-certified. Note that this completion is not the Manakov J given by Equation (2.3) resulting from the parameters

$$a = \frac{1}{2}(I_1 + I_2 + I_3), \quad b = -1, \quad d = \frac{1}{4}(I_1 + I_2 + I_3)^2 - (I_1I_2 + I_2I_3 + I_3I_1).$$

Nothing internal to the Euler top singles out $L_1 = J^2$, $P_1 = J$ over it: on $\mathfrak{so}(3)$ that form is one representative of the gauge family (a, b, d) , distinguished only as the $n = 3$ restriction of the n -dimensional rigid-body pair [51, 52], where J is the physical mass matrix and the Manakov shift/scale becomes forced.

Remark 2.1. SILO's role is limited to the first step of finding meaningful Lax pairs, discovering (L_*, P_*) from the compatibility residual; the λ -graded completion is then linear algebra over $\mathfrak{so}(3)$. Extensions of SILO to higher-dimensional Lie-Poisson systems should be explored in future work, particularly within the rigid-body context. Two natural targets are the n -dimensional rigid body on $\mathfrak{so}(n)$ [51, 52], whose spectral completion via the shift-of-argument method on semisimple Lie algebras tests the algebraic step of Section 2 at scale, and the Kowalewski top [6], whose Lax representation was found only in 1989 and is a natural stress test for any sparse-regression discovery procedure.

3 The Free Schrödinger Equation

We now move from the matrix setting to a scalar auxiliary formulation for PDEs, using the free Schrödinger equation,

$$q_t = i q_{xx}, \tag{3.1}$$

as a test case. The system is linear and classically solved via Fourier transforms. The associated Lax pair, in the Unified Transform Method (UTM) formalism [24, 25], consists of two first-order equations for an auxiliary field depending on a spectral parameter k . The example is simple enough that every step can be done by hand, and yet it already exhibits a phenomenon that will recur: two different nondegeneracy penalties, applied to the same compatibility condition, return two structurally different Lax pairs, only one of which is classical.

3.1 The Lax pair as a compatibility condition, learning ansatz and algebraic closure

A Lax pair for (3.1) consists of two first-order equations for an auxiliary field $\mu(x, t; k)$:

$$\mu_x = f(\mu, q; k), \quad \mu_t = g(\mu, q, q_x; k). \tag{3.2}$$

This may be viewed as defining the Pfaffian one-form

$$\omega = d\mu - f dx - g dt.$$

Compatibility, $\mu_{xt} = \mu_{tx}$, is equivalent to Frobenius integrability of $\omega = 0$. Indeed, along a graph $\mu = \mu(x, t; k)$ we have $d\mu = \mu_x dx + \mu_t dt$, and inserting this into $\omega = 0$ reproduces the pair. The one-form language is, of course, a compact way to encode the overdetermined first-order system and its compatibility, yet for our search purposes, the key relation is

$$\partial_t f(\mu, q; k) = \partial_x g(\mu, q, q_x; k).$$

This type of Clairaut compatibility also applies for integrable nonlinear PDEs [2, 55], and will be further exploited in subsequent sections.

Now, the classical Lax pair corresponds to the choice

$$f(\mu, q; k) = ik\mu + q, \quad g(\mu, q, q_x; k) = -ik^2\mu + iq_x - kq. \quad (3.3)$$

To pose a search problem for (3.2), we hypothesize an affine form for f and g :

$$\tilde{f} = a_1(k)q + a_2(k)\mu, \quad \tilde{g} = a_3(k)q + a_4(k)q_x + a_5(k)\mu,$$

with coefficients at most quadratic in the spectral parameter k ,

$$a_j(k) = \sum_{l=0}^2 z_{lj} k^l, \quad z_{lj} \in \mathbb{C}, \quad k \in \mathbb{R}.$$

A direct computation of $\tilde{f}_t$ and $\tilde{g}_x$ under Schrödinger dynamics gives

$$\tilde{f}_t = ia_1 q_{xx} + a_2(a_3 q + a_4 q_x + a_5 \mu), \quad \tilde{g}_x = a_3 q_x + a_4 q_{xx} + a_5(a_1 q + a_2 \mu).$$

The Clairaut compatibility $\tilde{f}_t - \tilde{g}_x = 0$ reads

$$\tilde{f}_t - \tilde{g}_x = (ia_1 - a_4)q_{xx} + (a_2 a_3 - a_5 a_1)q + (a_2 a_4 - a_3)q_x + (a_2 a_5 - a_5 a_2)\mu = 0,$$

where the last term vanishes identically. The remaining coefficients yield three nonlinear algebraic constraints:

$$ia_1 - a_4 = 0, \quad a_2 a_3 - a_5 a_1 = 0, \quad a_2 a_4 - a_3 = 0, \quad (3.4)$$

or more compactly, $F(\mathbf{a}) = \mathbf{0}$. It is clear that the k -dependence can be encoded in a matrix,

$$Z\mathbf{k} = \mathbf{a}, \quad Z = \begin{pmatrix} z_{01} & z_{02} & z_{03} & z_{04} & z_{05} \\ z_{11} & z_{12} & z_{13} & z_{14} & z_{15} \\ z_{21} & z_{22} & z_{23} & z_{24} & z_{25} \end{pmatrix}^\top, \quad \mathbf{k} = \begin{pmatrix} 1 \\ k \\ k^2 \end{pmatrix}.$$

Rather than solve $F(\mathbf{a}) = 0$ symbolically, we relax the root-finding problem into a sparse regression over the coefficient vector $z = \text{vec}(Z)$:

$$\min_{z \in \mathbb{C}^{15}} \frac{1}{2} \mathbb{E}_{k \sim \rho} \left[\sum_{j=1}^3 \|F_j(\mathbf{a})\|^2 \right] + s \|z\|_{\ell^1(\mathbb{C}^{15})}, \quad \mathbf{a} = Z\mathbf{k},$$

with $Z \in \mathbb{C}^{5 \times 3}$ and ρ uniform on $[-1, 1]$. The numerics recover the sparsest nontrivial Lax pair

$$\mu_x = iq, \quad \mu_t = -q_x,$$

which is irrelevant for UTM purposes: it contains no spectral parameter, and therefore it cannot be used to generate a meaningful transform in the UTM sense. This is the first nondegeneracy obstacle, and it is resolved by steering away from rank-deficient coefficient matrices.

Before doing so, we note that this type of degeneracy is general for linear PDE. The pair $\mu_x = iq$, $\mu_t = -q_x$ is the $k = 0$ reduction of the Fokas unified-transform Lax pair [24, 25], and every scalar constant-coefficient linear evolution equation has one. For $q_t + \omega(-i\partial_x)q = 0$ with polynomial dispersion ω obeying $\omega(0) = 0$, that pair is

$$\mu_x - ik\mu = q, \quad \mu_t + \omega(k)\mu = i \frac{\omega(k) - \omega(-i\partial_x)}{k + i\partial_x} q,$$

whose compatibility $\mu_{xt} = \mu_{tx}$ returns the equation for every k . At $k = 0$ it collapses to

$$\mu_x = q, \quad \mu_t = \partial_x^{-1} q_t = -\partial_x^{-1} \omega(-i\partial_x) q,$$

which carries no spectral parameter and reproduces the equation as the single conservation law $\mu_{xt} = \mu_{tx}$. The free Schrödinger case is $\omega(k) = ik^2$, for which $\mu_t = \partial_x^{-1}(iq_{xx}) = iq_x$, and the rescaling $\mu \mapsto i\mu$ returns the degenerate SILO pair $\mu_x = iq$, $\mu_t = -q_x$.

3.2 Two nondegeneracy penalties and the resulting Lax pairs

The classical Schrödinger Lax pair has coefficient matrix

$$Z_{\text{classical}} = \begin{pmatrix} 1 & 0 & 0 & i & 0 \\ 0 & i & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -i \end{pmatrix}^{\top},$$

which has full column rank. A direct way to penalize spectral collapse is thus to add a barrier against near-singular $Z^{\dagger}Z$. Two natural choices present themselves.

The determinant barrier penalizes volume collapse of the parallelepiped spanned by the columns of Z ,

$$\min_{\mathbf{a} \in \mathcal{A}} \frac{1-r}{2} \mathbb{E}_{k \sim \rho} \left[\sum_{j=1}^3 \|F_j(\mathbf{a})\|^2 \right] + \frac{r}{2} |\mathcal{D}|^{-2} + s \|z\|_{\ell^1(\mathbb{C}^{15})}, \quad \mathcal{D} = \det(Z^{\dagger}Z).$$

The Frobenius pseudoinverse surrogate penalizes the worst-conditioned singular direction,

$$\min_{\mathbf{a} \in \mathcal{A}} \frac{1-r}{2} \mathbb{E}_{k \sim \rho} \left[\sum_{j=1}^3 \|F_j(\mathbf{a})\|^2 \right] + \frac{r}{2} \tilde{\mathcal{D}} + s \|z\|_{\ell^1(\mathbb{C}^{15})}, \quad \tilde{\mathcal{D}} = \|Z^+\|_F^2 = \text{tr}((Z^{\dagger}Z)^{-1}).$$

Geometrically, both penalties act on the Gram matrix $G = Z^{\dagger}Z$. Full column rank of Z is equivalent to positive definiteness of G , or equivalently to all singular values $\sigma_1, \sigma_2, \sigma_3$ of Z being nonzero. The determinant

$$\det(Z^{\dagger}Z) = \sigma_1^2 \sigma_2^2 \sigma_3^2$$

measures the squared volume of the parallelepiped spanned by the columns of Z , so a determinant barrier prevents collapse of this volume and hence penalizes loss of rank. By contrast,

$$\|Z^+\|_F^2 = \text{tr}((Z^{\dagger}Z)^{-1}) = \sigma_1^{-2} + \sigma_2^{-2} + \sigma_3^{-2}$$

measures inverse conditioning: it blows up whenever any singular direction of Z becomes small. The determinant barrier is naturally interpreted as a volume-collapse penalty, whereas the pseudoinverse penalty is a conditioning penalty that more strongly discourages thin or nearly collapsed directions. We note that both barriers can be weakened by rescaling z to enlarge the singular values of Z ; however, the ℓ^1 term counteracts this by penalizing the overall size of the coefficients.

The two penalties select different pairs. The determinant penalty regularly returns the classical five-term Schrödinger Lax pair. The Frobenius surrogate, however, returns a different sparse solution, with coefficients of the form

$$a_2(k) \equiv z_4, \quad a_5(k) \equiv z_{11},$$

and

$$a_1(k) = z_1 + z_2 k + z_3 k^2, \quad a_3(k) = z_5 + z_6 k + z_7 k^2, \quad a_4(k) = z_8 + z_9 k + z_{10} k^2,$$

where the eleven complex parameters z_j satisfy (3.4). This second pair is sparse, compatible, and not equivalent to the classical one by any straightforward gauge. It is, however, spectrally degenerate in a precise algebraic sense, as a brief manual calculation shows.

For demonstration purposes, set $z_4 = 1$ and $z_{11} = i$. With this choice, Equations (3.4) are satisfied for every k by any function $M = M(k)$ through the relations $a_1(k) = M(k)$ and $a_3(k) = a_4(k) = iM(k)$. The resulting pair is

$$\mu_x = M(k)q + \mu, \quad \mu_t = iM(k)(q + q_x) + i\mu, \quad (3.5)$$

which admits the Schrödinger equation $q_t = i q_{xx}$ as its Clairaut compatibility condition for every choice of the function $M(k)$. The arbitrariness of $M(k)$ is structurally revealing: the pair carries no spectral information at all. Different $M(k)$ produce different auxiliary systems, all compatible with the same underlying PDE, with no way to select among them by the compatibility condition alone. In the language of the UTM, this family cannot produce a nontrivial transform kernel; the spectral parameter is present but acts as a free function rather than as the encoding of a scattering problem.

In fact, it is easy to see that pair given by Equation (3.5) is fake using the diagnostics of Appendix A.2. The pair is *g*-fake since the gauge $\mu \mapsto M(k)^{-1}\mu$ above sends it to the $M(k) \equiv 1$ pair, removing k . It is *u*-fake: imposing (3.4) on the generalized coefficients with $a_2 = z_4$, $a_5 = z_{11}$ gives

$$a_4 = i a_1, \quad a_3 = i z_4 a_1, \quad z_{11} = i z_4^2,$$

so $a_1 = M(k)$ is unconstrained and the compatibility system is underdetermined. Additionally, the Krichever test of Appendix A.3 agrees: linearizing at $q = q_0$, the auxiliary system supports the single mode $\mu \propto e^{x+it}$, independent of k , so ω_{Lax} is constant while $\omega_{\text{Schr}}(\xi) = \xi^2$.

The broader lesson of this section is that this is another example where compatibility does not select among Lax representations. Even after enforcing rank-fullness in coefficient space, the Clairaut condition leaves room for analytically unhelpful alternatives. Structural selection, i.e., determining which compatible pair actually encodes scattering, conservation, or inverse-problem data, is a separate question from compatibility itself, and must be addressed through additional constraints that reflect the intended use of the pair.

4 The Inviscid Burgers Equation and Infinite Parametric Families of Lax Pairs

The Schrödinger example exhibited nonuniqueness in the form of a single spectrally degenerate alternative to the classical pair. In this section, we will show that the Hopf equation exhibits a different yet also striking form of nonuniqueness: an infinite parametric family of compatible auxiliary systems, indexed by a positive integer, all of which reduce to a base pair under Galilean transformations. The family emerged through a SILO search that was deliberately open about the exponent structure of the compatibility, and once one member was identified numerically, the full family could be verified by hand for every integer $m \geq 1$.

4.1 A pen-and-paper auxiliary pair and SILO Formulation

The inviscid Burgers (or Hopf) equation,

$$u_t + u u_x = 0, \tag{4.1}$$

admits a WKB-type auxiliary formulation in which a scalar phase field $S(x, t)$ plays the role of the WKB phase. A pen-and-paper calculation of the dispersionless limit of the Kadomtsev-Petviashvili equation [66] produces, as a special case, the pair

$$S_x^2 =: f(u; \lambda), \quad S_t =: g(u, S_x), \tag{4.2}$$

with

$$f(u; \lambda) := -\frac{1}{6}u - \lambda, \quad g(u, S_x) := -4S_x^3 - u S_x, \tag{4.3}$$

and with spectral parameter $\lambda \in \mathbb{C}$. This is a dispersionless Lax-type pair in the sense that compatibility of (4.2) reproduces the dynamics of u through a Clairaut-type condition.

Before posing the search problem, it is useful to derive the compatibility conditions in a form that makes clear what we are enforcing. Write $s := S_x$ and treat $f = f(u)$ and $g = g(u, s)$. Differentiating $f(u) = s^2$ in x gives $2s s_x = f_u(u) u_x$, implying

$$s_x = \frac{f_u(u)}{2s} u_x.$$

Differentiating $f(u) = s^2$ in t and using (4.1) yields $2s s_t = f_u(u) u_t = -f_u(u) u u_x$, implying

$$s_t = -\frac{f_u(u)}{2s} u u_x.$$

On the other hand, differentiating the second of (4.2) in x gives

$$s_t = S_{xt} = (S_t)_x = g_u(u, s) u_x + g_s(u, s) s_x.$$

Substituting the expression for s_x and imposing $S_{xt} = S_{tx}$ yields, after factoring out u_x ,

$$g_u(u, s) + \frac{f_u(u)}{2s} g_s(u, s) = -\frac{f_u(u)}{2s} u,$$

or, equivalently the compatibility identity

$$2s g_u + f_u(g_s + u) = 0. \quad (4.4)$$

A direct substitution verifies that the pen-and-paper pair (4.2)–(4.3) satisfies (4.4) identically. This is the constraint that the regression enforces.

Because (4.4) involves only u and s , we hypothesize f and g as polynomials in those variables,

$$\tilde{f}(u) = \sum_{j=0}^{N_1} \xi_j u^j, \quad \tilde{g}(s, u) = \sum_{j=0}^{N_2} \sum_{k=0}^{N_3} \zeta_{j,k} s^j u^k, \quad (4.5)$$

and collect the unknowns into $\eta = (\Xi, \mathcal{Z}) \in \mathbb{R}^N$ with $N = (N_1 + 1) + (N_2 + 1)(N_3 + 1)$. Substituting (4.5) into (4.4) produces a residual that is polynomial in (u, s) and quadratic in η . SILO in this setting thus takes the form

$$\min_{\eta \in \mathbb{R}^N} \mathbb{E}_{(u,s) \sim \nu} \left[\frac{(2s \partial_u \tilde{g} + \partial_u \tilde{f} (\partial_s \tilde{g} + u))^2}{\|\Xi\|_\infty \|\mathcal{Z}\|_\infty} \right] + s \|\eta\|_{\ell^1}, \quad (4.6)$$

with ν uniform on $[-1, 1]^2$. The ℓ^∞ product normalization in the denominator plays the role of the Frobenius product penalty used in Section 2: it diverges whenever either Ξ or $\mathcal{Z}$ collapses to zero, and so suppresses the trivial solutions $\tilde{f} \equiv 0$ and $\tilde{g} \equiv 0$. The ℓ^1 term once again promotes parsimony.

4.2 Three numerically discovered families

Three sparse families emerge from the numerics. The first is the pen-and-paper pair. The second, discovered by SILO, may be written (after rescaling) as

$$S_x^2 = au + \lambda, \quad S_t = bu - uS_x - \frac{b}{a} S_x^2 + \frac{2}{3a} S_x^3, \quad (4.7)$$

with $a \neq 0$ and free constants (a, b, λ) . A third, higher-degree family also appears:

$$S_x^2 = au + \lambda, \quad S_t = c_1 S_x^4 + c_2 S_x^3 + c_3 S_x^2 + c_4 u S_x^2 + c_5 u S_x + c_6 u + c_7 u^2, \quad (4.8)$$

with the c_j satisfying linear relations derived below.

It is clear that even higher-degree families would emerge from the numerics if we include more terms in the polynomial library. Rather than verify each family independently or proceed by induction, we seek to solve the compatibility PDE (4.4) in closed form. Doing so reveals that the pattern is more general than the two polynomial samples above suggest: there is an infinite-dimensional family of admissible temporal parts S_t , indexed by an arbitrary function of one variable, every element of which collapses to the base pair on the constraint surface. The second and third numerical families are the linear and quadratic polynomial samples, respectively.

The compatibility PDE and its general solution. Fix the spatial ansatz $S_x^2 = au + \lambda$ with $a \neq 0$, and, again, write $s := S_x$ for brevity. With $f(u) = au + \lambda$, the Clairaut identity (4.4) becomes the linear first-order PDE

$$a g_s + 2s g_u = -au \quad (4.9)$$

for the unknown $g(s, u) = S_t$. Solving (4.9) is thus a routine application of the method of characteristics with characteristic curves that satisfy $ds/a = du/(2s)$, or equivalently, $d(s^2 - au) = 0$, and therefore coincide with the level sets of the spatial constraint: the characteristic variable is $\xi := s^2 - au$, and the surface $\xi = \lambda$ is the one on which the auxiliary pair is evaluated.

Proposition 4.1. *The general C^1 solution of (4.9) is*

$$g(s, u) = \frac{2}{3a} s^3 - us + \Phi(s^2 - au), \quad (4.10)$$

where $\Phi: \mathbb{R} \rightarrow \mathbb{R}$ is arbitrary. Every Lax pair for the inviscid Burgers equation with spatial ansatz $S_x^2 = au + \lambda$ has temporal part of this form. Moreover, on the constraint surface $s^2 = au + \lambda$, every element of the family (4.10) reduces to

$$S_t = -\frac{S_x^3}{3a} + \frac{\lambda}{a} S_x + \Phi(\lambda), \quad (4.11)$$

a Galilean translate of the base pair with additive constant $\mu := \Phi(\lambda)$.

Proof. The function $g_0(s, u) := 2s^3/(3a) - us$ is a particular solution of (4.9), since $(g_0)_s = 2s^2/a - u$ and $(g_0)_u = -s$ imply

$$a(g_0)_s + 2s(g_0)_u = 2s^2 - au - 2s^2 = -au.$$

The associated homogeneous equation $ah_s + 2sh_u = 0$ has general solution $h = \Phi(\xi)$ with $\xi = s^2 - au$, since ξ is constant along characteristics. Summing gives (4.10). Equation (4.11) then follows by simply substituting $u = (s^2 - \lambda)/a$ into (4.10), which yields

$$g = \frac{2}{3a} s^3 - \frac{s^2 - \lambda}{a} s + \Phi(\lambda) = -\frac{s^3}{3a} + \frac{\lambda s}{a} + \Phi(\lambda). \quad \square$$

The two numerical families as polynomial samples. The second and third families correspond to linear and quadratic choices of Φ , respectively. Taking $\Phi(\xi) = -b\xi/a$ in (4.10) gives

$$g(s, u) = \frac{2}{3a} s^3 - us - \frac{b}{a} (s^2 - au) = \frac{2}{3a} S_x^3 - u S_x - \frac{b}{a} S_x^2 + bu,$$

the second family. Taking $\Phi(\xi) = \alpha_0 + \alpha_1 \xi + \alpha_2 \xi^2$ and expanding yields

$$g(s, u) = \frac{2}{3a} s^3 - us + \alpha_0 + \alpha_1 (s^2 - au) + \alpha_2 (s^2 - au)^2.$$

Matching monomial coefficients against the third family's basis $\{s^4, s^3, s^2, us^2, us, u, u^2, 1\}$ gives

$$c_1 = \alpha_2, \quad c_2 = \frac{2}{3a}, \quad c_3 = \alpha_1, \quad c_4 = -2a\alpha_2, \quad c_5 = -1, \quad c_6 = -a\alpha_1, \quad c_7 = a^2\alpha_2,$$

with additive constant α_0 . The seven c_j are thus determined by three free parameters $(\alpha_0, \alpha_1, \alpha_2)$, with c_2 and c_5 pinned by compatibility and (c_1, c_4, c_7) , (c_3, c_6) related through α_2, α_1 respectively. The relations among the c_j reported by the numerics are exactly these.

The infinite family. Proposition 4.1 says that the temporal part S_t is determined only up to an additive homogeneous solution $\Phi(\xi)$, $\xi = s^2 - au$, a function constant along the characteristics of the compatibility PDE. As solutions of that PDE on the full (s, u) plane these are genuinely distinct, and the solution space is infinite-dimensional. As Lax pairs they are not, that is, the auxiliary system lives only on the constraint surface $\xi = \lambda$, where $\Phi(\xi)$ collapses to the single constant $\Phi(\lambda)$, and that constant is the potential gauge $S \mapsto S + \Phi(\lambda)t$, which fixes S_x and the recovered equation. Every member therefore reduces on the surface to the base pair up to this gauge, as the Corollary records. Polynomial choices of Φ of degree N give (s, u) -polynomial S_t of bidegree up to $(2N, N)$, and the sequence of “new” families SILO returns for $N = 2, 3, \dots$ is the artifact of evaluating Φ before restricting to $\xi = \lambda$. There is no ceiling to N , and no new Lax pair past the base one.

4.3 An infinite family of pairs

In light of the above results, we applied the same approach assuming $S_x^3 = f(u; \lambda)$. Once again SILO returned valid Lax pairs, but this time with a noticeably different algebraic structure. Pushing further, while investigating hypotheses of the form $S_x^m = f(u; \lambda)$, SILO discovered the following pair,

$$S_x^m = au + \lambda, \quad S_t = -\frac{1}{a(m+1)}S_x^{m+1} + \frac{\lambda}{a}S_x + \mu, \quad (4.12)$$

which is valid for every integer $m > 0$, with $a \neq 0$ and arbitrary constants (λ, μ) .

Showing this by hand for general m is not difficult. Set $s := S_x$ and assume $s^m = au + \lambda$, so that $u = (s^m - \lambda)/a$. Differentiating in x yields $ms^{m-1}s_x = au_x$, i.e., $u_x = \frac{m}{a}s^{m-1}s_x$. Differentiating in t and using inviscid Burgers yields instead

$$ms^{m-1}s_t = au_t = -auu_x = -a\left(\frac{s^m - \lambda}{a}\right)\left(\frac{m}{a}s^{m-1}s_x\right).$$

Canceling the common factor ms^{m-1} ,

$$s_t = -\frac{1}{a}(s^m - \lambda)s_x = \frac{1}{a}(\lambda - s^m)s_x.$$

If we now seek $S_t = P(s)$, then $s_t = S_{xt} = (S_t)_x = P'(s)s_x$, so that $P'(s) = (\lambda - s^m)/a$ implies

$$P(s) = \frac{1}{a}\left(\lambda s - \frac{s^{m+1}}{m+1}\right) + \mu,$$

which is exactly (4.12). The $m = 2$ case coincides with the pair reported earlier in this section.

To interpret this result it is useful to recall Calogero and Nucci's key observation [13] that any PDE with a local conservation law admits a Lax pair built by hodograph dressing of a constant-coefficient linear pair. We can therefore ask whether the inviscid Burgers family (4.12) could also be derived from a conservation-law hodograph in the dispersionless WKB setting. The answer is that indeed it can, as we now show.

The starting observation is that inviscid Burgers admits a fractional-power conservation hierarchy. For any real $m > 0$,

$$\partial_t u^{1/m} = \frac{1}{m} u^{1/m-1} u_t = -\frac{1}{m} u^{1/m} u_x = -\frac{1}{m+1} \partial_x u^{(m+1)/m},$$

so $\partial_t u^{1/m} + \frac{1}{m+1} \partial_x u^{(m+1)/m} = 0$ is a conservation law of inviscid Burgers for every $m > 0$. Allowing a Galilean shift, set $v := au + \lambda$ for constants $a \neq 0$ and λ . The inviscid Burgers in u becomes $v_t = -(v - \lambda)v_x/a$ in v . A direct computation then gives

$$\partial_t v^{1/m} = \frac{1}{m} v^{1/m-1} v_t = -\frac{(v - \lambda)v_x}{ma} v^{1/m-1} = \frac{\lambda v^{1/m-1} v_x}{ma} - \frac{v^{1/m} v_x}{ma},$$

and each term on the right is a total x -derivative:

$$-\frac{v^{1/m} v_x}{ma} = -\frac{1}{a(m+1)} \partial_x v^{(m+1)/m}, \quad \frac{\lambda v^{1/m-1} v_x}{ma} = \frac{\lambda}{a} \partial_x v^{1/m}.$$

Combining,

$$\partial_t v^{1/m} = \partial_x \left[\frac{\lambda}{a} v^{1/m} - \frac{1}{a(m+1)} v^{(m+1)/m} \right],$$

which is a one-parameter family of conservation laws indexed by $m > 0$, with densities f_m and fluxes g_m given by

$$f_m := (au + \lambda)^{1/m}, \quad g_m := \frac{\lambda}{a} (au + \lambda)^{1/m} - \frac{1}{a(m+1)} (au + \lambda)^{(m+1)/m}.$$

With these conservation laws in hand, we now apply the hodograph construction. Recall that a local conservation law $\partial_t f = \partial_x g$ is the closedness condition for the one-form $\omega = f dx + g dt$, so on a simply connected solution domain there exists a scalar potential $S(x, t)$ with $\omega = dS$, giving $S_x = f$ and $S_t = g$. The conservation law of the underlying PDE is thereby promoted to a first-order auxiliary system in S . Apply this with $f = f_m$ and $g = g_m$ from above by setting

$$S_x = f_m = (au + \lambda)^{1/m}, \quad S_t = g_m + \mu,$$

where the free parameter μ is the additive potential-gauge freedom $S \mapsto S + \mu t$, which leaves S_x invariant and shifts S_t . The mixed-partial consistency $S_{xt} = S_{tx}$ reads $\partial_t f_m = \partial_x g_m$, exactly the conservation law derived in the previous paragraph, so S exists along solutions of inviscid Burgers.

Lastly, to recover (4.12), simply raise S_x to the m -th power, and we have immediately the spatial part $S_x^m = au + \lambda$. For the temporal part, note that $S_x^{m+1} = (au + \lambda)^{(m+1)/m}$, so each term in g_m rewrites cleanly as a power of S_x :

$$g_m = \frac{\lambda}{a} S_x - \frac{1}{a(m+1)} S_x^{m+1}.$$

Substituting into $S_t = g_m + \mu$ gives the temporal equation of (4.12).

The family that SILO discovered is therefore a dispersionless first-order incarnation of the Calogero–Nucci mechanism, with the integer $m > 0$ indexing the fractional-power root of the underlying conservation law and the constants (a, λ, μ) playing the roles of scaling, Galilean shift, and gauge. The integer indexing is itself an artifact of our polynomial library. Indeed, the derivation above goes through for any real $m > 0$, since the conservation law $\partial_t v^{1/m} = \partial_x [\dots]$ requires nothing of m beyond $m \neq 0$, and the spatial relation $S_x^m = au + \lambda$ is a well-defined algebraic constraint on any branch where $au + \lambda > 0$. Inviscid Burgers therefore admits a genuine continuum of compatible dispersionless auxiliary systems indexed by $m \in \mathbb{R}_{>0}$. This is the inviscid Burgers incarnation of “Lax pairs galore” in its strongest form, with uncountably many compatibility certificates.

5 Shallow Water Equations and Dispersionless Lax Functions

Shallow water extends the inviscid Burgers analysis of section 4 to two field components, and the Lax landscape splits into two registers. Within the Laurent layer of dispersionless Lax functions in a spectral parameter p , the classical Brunelli–Das pair [10] is the only object our sparse search returns across the Laurent and polynomial degrees we explored. Outside the Laurent layer, the Calogero–Nucci hodograph mechanism produces an infinite-dimensional family of compatible auxiliary systems, parametrized by solutions of an Euler–Poisson–Darboux equation in the Riemann invariants. The Brunelli–Das pair sits inside this continuum as the algebraic family generated by the spectral discriminant.

5.1 The Brunelli–Das dispersionless Lax representation

The shallow water system (also known as the Saint-Venant system, which is equivalent to the polytropic gas equations with a pressure term $P = \pm \frac{1}{2} a \rho^\gamma$ with $\gamma = 2$) reads

$$\rho_t + (u\rho)_x = 0, \quad u_t + uu_x + a\rho_x = 0, \quad a = \pm 1, \quad (5.1)$$

and arises as the dispersionless (semiclassical) limit of focusing and defocusing NLS equations in its Madelung (i.e., density and phase gradient/velocity) representation [33]. Brunelli and Das [10] gave a Lax description of the polytropic gas hierarchy; here we limit our considerations to the case $\gamma = 2$. Their result is the dispersionless Lax function

$$L(p; u, \rho) = p + u + \frac{a\rho}{p}, \quad (5.2)$$

together with the polynomial generator

$$B(p; u) = \frac{1}{2} p^2 + up, \quad (5.3)$$

related by the standard hierarchy projection

$$B = \left(\frac{1}{2} L^2 \right)_{>0}, \quad (5.4)$$

where $(\cdot)_{>0}$ keeps strictly positive powers of p . The dispersionless Lax equation

$$L_t = \{L, B\}, \quad \{L, B\} := L_p B_x - L_x B_p, \quad (5.5)$$

in which p is an independent spectral variable, so L_t, L_x, B_x act through $u(x, t), \rho(x, t)$ at fixed p and L_p, B_p are the explicit p -derivatives, reproduces (5.1) after matching coefficients of p^0 and p^{-1} . We refer the reader to [9, 10, 16] for the algebraic structure (recursion relations, biHamiltonian content, and generating functions for the conserved charges), and to [35] for a treatment of dispersionless coupled-KdV hodographs in which the Euler–Poisson–Darboux equation that we encounter below appears as the compatibility datum for hierarchy critical points.

Similarly as in Section 4, the Lax function (5.2) admits an equivalent phase-function form: introducing a phase $S(x, t; \lambda)$ such that $p := S_x$, the spectral curve $L = \lambda$ becomes the quadratic relation

$$p^2 + (u - \lambda)p + a\rho = 0, \quad (5.6)$$

and (5.5) is equivalent, on the level set $L = \lambda$, to the Clairaut compatibility $S_{xt} = S_{tx}$ for $S_x = p$ and $S_t = -B$. We will return to this form in Section 5.3 with more detail when we generate alternative pairs by the hodograph mechanism.

5.2 SILO formulation and recovery of the known pair

The hierarchy projection (5.4) is the structural feature that makes a SILO formulation fairly straightforward: the time generator B is not an independent unknown, and the inverse problem reduces to learning a single object L . We therefore posit for L the truncated Laurent expansion

$$\tilde{L}(p; u, \rho) = \sum_{m=-M_-}^{M_+} c_m(u, \rho) p^m, \quad c_m(u, \rho) = \sum_{j=0}^{N_u} \sum_{k=0}^{N_\rho} \theta_{j,k}^{(m)} u^j \rho^k, \quad (5.7)$$

with normalization $c_1 \equiv 1$ to fix the spectral scaling. To derive the regression residuals, we expand $L_t = L_u u_t + L_\rho \rho_t$ and substitute the dynamical equations (5.1), obtaining

$$L_t = (-u L_u - \rho L_\rho) u_x + (-a L_u - u L_\rho) \rho_x.$$

For the Poisson bracket $\{L, B\} = L_p B_x - L_x B_p$, the spatial derivatives L_x and B_x at fixed p expand via the chain rule as

$$L_x = L_u u_x + L_\rho \rho_x, \quad B_x = B_u u_x + B_\rho \rho_x.$$

Substituting and grouping by u_x and ρ_x ,

$$\{L, B\} = L_p B_x - L_x B_p = L_p (B_u u_x + B_\rho \rho_x) - (L_u u_x + L_\rho \rho_x) B_p = (L_p B_u - L_u B_p) u_x + (L_p B_\rho - L_\rho B_p) \rho_x.$$

Equating coefficients gives two pointwise residuals

$$\mathcal{R}_1(\tilde{L}) = -u \tilde{L}_u - \rho \tilde{L}_\rho - (\tilde{L}_p \tilde{B}_u - \tilde{L}_u \tilde{B}_p), \quad (5.8a)$$

$$\mathcal{R}_2(\tilde{L}) = -a \tilde{L}_u - u \tilde{L}_\rho - (\tilde{L}_p \tilde{B}_\rho - \tilde{L}_\rho \tilde{B}_p), \quad (5.8b)$$

with $\tilde{B} = (\frac{1}{2} \tilde{L}^2)_{>0}$, leading to the SILO problem

$$\min_{\theta \in \Theta} (1-r) \mathbb{E}_{(u, \rho, p) \sim \nu} \left[|\mathcal{R}_1|^2 + |\mathcal{R}_2|^2 \right] + r \mathcal{N}(\tilde{L}) + 2s \|\theta\|_{\ell^1}, \quad (5.9)$$

where ν is supported on a bounded set with $\rho > 0$ and $|p| \geq p_{\min} > 0$. The penalty

$$\mathcal{N}(\tilde{L}) = \frac{1}{\mathbb{E}_{(u, \rho, p) \sim \nu} [\tilde{L}_u^2 \tilde{L}_\rho^2]} \quad (5.10)$$

is the shallow-water analog of the Frobenius product penalty (2.8) used for the Euler top: it diverges whenever $\tilde{L}_u$ or $\tilde{L}_\rho$ vanishes identically on the support of ν , forcing both fields to remain represented in $\tilde{L}$. Without it, the optimizer collapses onto $\tilde{L} = p$, which satisfies (5.8a)–(5.8b) identically but carries no field content.

In the experiment we take $M_- = 2$, $M_+ = 1$, and $N_u = N_\rho = 2$, so the classical three-term ansatz is not built in a priori. Across random initializations the optimizer returns the Brunelli–Das pair (5.2) and its constant-shift gauges $L \mapsto L + c$, $B \mapsto B + c p$, and reports no other compatible solution within the ansatz. Within the Laurent class, the known Lax function is the only object SILO could find, even after enlarging the search window to $M_-, M_+ \leq 5$ and the polynomial bidegrees to $N_u, N_\rho \leq 3$. Outside the Laurent class, however, the picture is entirely different, as we show next.

5.3 A Calogero–Nucci cornucopia of pairs via the Euler–Poisson–Darboux equation

Recall that section 4 produced an infinite family of Lax representations for inviscid Burgers via the Calogero–Nucci hodograph mechanism. In the present case, a SILO sweep on a hodograph ansatz would, by same token, surface a handful of low-degree members of the family one at a time and leave the structural pattern to be recognized afterward by hand. Here, we skip the relevant numerics. The mechanism applies verbatim to the shallow water equations, but with a resulting family that is substantially richer than for the Hopf equation, because the shallow water equations have two Riemann invariants instead of just one. Note that in this subsection we set $a = +1$ for definiteness. The case $a = -1$ only requires a sign tracking throughout.

Riemann coordinates and the compatibility equation. The Riemann invariants of (5.1) are

$$r_+ = u + 2\sqrt{\rho}, \quad r_- = u - 2\sqrt{\rho}, \quad (5.11)$$

and a direct calculation gives the Riemann form

$$(r_{\pm})_t + c_{\pm} (r_{\pm})_x = 0, \quad c_{\pm} = u \pm \sqrt{\rho} = \frac{3r_{\pm} + r_{\mp}}{4}, \quad (5.12)$$

two coupled Hopf-like equations whose characteristic speeds depend on both invariants. Unlike scalar Hopf, therefore, no single-variable function $f(r_+)$ alone is conserved. Thus, the analog of the fractional-power trick of Section 4 requires a joint density $F(r_+, r_-)$.

A Calogero–Nucci hodograph pair takes the form $S_x = F(u, \rho)$, $S_t = G(u, \rho)$, with compatibility $S_{xt} = S_{tx}$ equivalent on solutions of (5.1) to the conservation law

$$\partial_t F = \partial_x G. \quad (5.13)$$

Working in Riemann coordinates and writing $F = F(r_+, r_-)$, $G = G(r_+, r_-)$, the chain rule together with the Riemann form $(r_{\pm})_t = -c_{\pm}(r_{\pm})_x$ gives

$$\begin{aligned} \partial_t F &= F_{r_+}(r_+)_t + F_{r_-}(r_-)_t = -c_+ F_{r_+}(r_+)_x - c_- F_{r_-}(r_-)_x, \\ \partial_x G &= G_{r_+}(r_+)_x + G_{r_-}(r_-)_x. \end{aligned}$$

Matching coefficients of $(r_+)_x$ and $(r_-)_x$ in $\partial_t F = \partial_x G$ forces the flux to satisfy

$$G_{r_+} = -c_+ F_{r_+}, \quad G_{r_-} = -c_- F_{r_-}. \quad (5.14)$$

Differentiating these, and using $(c_+)_{r_-} = (c_-)_{r_+} = 1/4$ from $c_{\pm} = (3r_{\pm} + r_{\mp})/4$,

$$G_{r_+ r_-} = -\frac{1}{4} F_{r_+} - c_+ F_{r_+ r_-}, \quad G_{r_- r_+} = -\frac{1}{4} F_{r_-} - c_- F_{r_+ r_-}.$$

Setting $G_{r_+ r_-} = G_{r_- r_+}$ and using $c_+ - c_- = (r_+ - r_-)/2$ reduces to the linear second-order PDE

$$\frac{1}{2}(r_+ - r_-) F_{r_+ r_-} + \frac{1}{4}(F_{r_+} - F_{r_-}) = 0. \quad (5.15)$$

This is a classical Euler–Poisson–Darboux (EPD) equation [17, 65], the same equation that organizes the Carrier–Greenspan transformation of long-wave runup [14] and the hydrodynamic-type integrability theory of Tsarev [62], and that arises in [35] as the equation governing hodograph critical points of the dispersionless coupled-KdV hierarchies. The solution space of (5.15) is infinite-dimensional, and every solution gives a Lax pair for shallow water through (5.14) and the hodograph identification $S_x = F$, $S_t = G$. Here we exhibit three families explicitly.

The polynomial Whitham hierarchy. Restricting (5.15) to polynomial $F(r_+, r_-)$ and solving by total degree, the basis grows by exactly one new direction per degree. The reason is dimensional since the space of homogeneous polynomials of degree d in (r_+, r_-) is $(d+1)$ -dimensional, while EPD applied to a homogeneous degree- d polynomial returns a homogeneous expression of degree $d-1$ whose vanishing imposes d linear constraints, one for each monomial $r_+^i r_-^{d-1-i}$ with $i = 0, \dots, d-1$. This leaves one free direction at each degree.

Degree ≤ 1 is trivial: $F = a + b(r_+ + r_-)$ has $F_{r_+ r_-} = 0$ and $F_{r_+} - F_{r_-} = 0$, so EPD reduces to $0 = 0$ for any constants a, b . At degree 2, write $F = \alpha r_+^2 + \beta r_+ r_- + \gamma r_-^2$. Then $F_{r_+ r_-} = \beta$ and $F_{r_+} - F_{r_-} = (2\alpha - \beta)r_+ + (\beta - 2\gamma)r_-$. Substituting into EPD,

$$\frac{1}{2}\beta(r_+ - r_-) + \frac{1}{4}[(2\alpha - \beta)r_+ + (\beta - 2\gamma)r_-] = 0,$$

and matching coefficients of r_+ and r_- gives $\beta = -2\alpha$ and $\beta = -2\gamma$, so $\alpha = \gamma$. Setting $\alpha = 1$, the new direction is $F_2 = r_+^2 - 2r_+ r_- + r_-^2 = (r_+ - r_-)^2$. The degree-3 calculation is analogous and forces the homogeneous cubic part to be proportional to $(r_+ + r_-)(r_+ - r_-)^2$. The pattern continues, giving a basis $\{F_0, F_1, F_2, \dots\}$ of polynomial EPD solutions,

$$F_0 = 1, \quad F_1 = r_+ + r_-, \quad F_2 = (r_+ - r_-)^2, \quad F_3 = (r_+ + r_-)(r_+ - r_-)^2, \quad \dots,$$

where the subspace of solutions of total degree at most d has dimension $d+1$.

Translating back to the original physical variables (u, ρ) via $r_+ + r_- = 2u$ and $(r_+ - r_-)^2 = 16\rho$, the first six generators are, up to overall normalization,

$$1, \quad u, \quad \rho, \quad u\rho, \quad \rho^2 + u^2\rho, \quad u\rho(3\rho + u^2), \quad \dots \quad (5.16)$$

We recognize this pattern as the well-known Kupershmidt–Manin / Whitham hierarchy of conserved densities for shallow water [41, 65].

For each density F , the matching flux G is determined directly by the conservation law $\partial_t F = \partial_x G$ on shallow-water solutions. Two examples make this concrete. The mass density $F = \rho$ uses $\rho_t = -(u\rho)_x$ from (5.1):

$$\partial_t \rho = -(u\rho)_x = \partial_x(-u\rho), \quad G = -u\rho,$$

giving the hodograph pair $(S_x, S_t) = (\rho, -u\rho)$. The momentum-energy density $F = u\rho$ uses both PDEs and the product rule:

$$\partial_t(u\rho) = u_t \rho + u \rho_t = (-uu_x - \rho_x)\rho - u(u\rho)_x = -2u\rho u_x - \rho\rho_x - u^2\rho_x = -\partial_x(u^2\rho + \frac{1}{2}\rho^2),$$

giving the hodograph pair $(S_x, S_t) = (u\rho, -u^2\rho - \frac{1}{2}\rho^2)$.

None of these is a Laurent function in any spectral parameter λ ; they are polynomial in the field variables, and they are invisible to the Brunelli–Das presentation because that presentation parameterizes the spectral curve $L = \lambda$, not the conserved densities directly.

The algebraic continuum via the spectral discriminant. A second family of EPD solutions comes from the algebraic ansatz

$$F(r_+, r_-; \lambda) = (r_+ - \lambda)^a (r_- - \lambda)^b$$

with constants $a, b \in \mathbb{R}$ and spectral parameter $\lambda \in \mathbb{C}$. The derivatives are

$$F_{r_+ r_-} = ab(r_+ - \lambda)^{a-1}(r_- - \lambda)^{b-1}, \quad F_{r_+} - F_{r_-} = a(r_+ - \lambda)^{a-1}(r_- - \lambda)^b - b(r_+ - \lambda)^a(r_- - \lambda)^{b-1}.$$

Substituting into EPD and factoring out $(r_+ - \lambda)^{a-1}(r_- - \lambda)^{b-1}$ leaves

$$\frac{1}{2}ab(r_+ - r_-) + \frac{1}{4}[a(r_- - \lambda) - b(r_+ - \lambda)] = 0,$$

which expanded becomes $b(2a-1)r_+ + a(1-2b)r_- + (b-a)\lambda = 0$. Matching coefficients of r_+, r_-, λ separately forces $a = b = 1/2$ as the unique nontrivial product solution. Hence

$$F_\lambda(r_+, r_-) = \sqrt{(\lambda - r_+)(\lambda - r_-)}. \quad (5.17)$$

To find the matching flux G_λ , use the relations (5.14):

$$G_{r_+} = -c_+ F_{\lambda, r_+} = \frac{(3r_+ + r_-)(\lambda - r_-)}{8F_\lambda}.$$

Take the ansatz $G_\lambda = h(r_+, r_-, \lambda) F_\lambda$. Then $G_{r_+} = h_{r_+} F_\lambda + h F_{\lambda, r_+}$. Multiplying both sides of the constraint by F_λ , using $F_\lambda^2 = (\lambda - r_+)(\lambda - r_-)$, and dividing through by $(\lambda - r_-)$ gives the reduced equation

$$h_{r_+}(\lambda - r_+) = \frac{1}{8}(3r_+ + r_-) + \frac{1}{2}h,$$

with the symmetric companion $h_{r_-}(\lambda - r_-) = \frac{1}{8}(r_+ + 3r_-) + \frac{1}{2}h$ from the G_{r_-} constraint. Trying $h = \alpha r_+ + \beta r_- + \gamma \lambda$ and matching coefficients of r_+, r_-, λ on both sides forces $\alpha = \beta = -1/4$ and $\gamma = -1/2$, giving

$$G_\lambda(r_+, r_-) = -\frac{1}{4}(r_+ + r_- + 2\lambda) \sqrt{(\lambda - r_+)(\lambda - r_-)}. \quad (5.18)$$

In (u, ρ) coordinates, $(\lambda - r_+)(\lambda - r_-) = (\lambda - u)^2 - 4\rho$ and $r_+ + r_- + 2\lambda = 2(u + \lambda)$, so

$$F_\lambda(u, \rho) = \sqrt{(\lambda - u)^2 - 4\rho}, \quad G_\lambda(u, \rho) = -\frac{1}{2}(\lambda + u) \sqrt{(\lambda - u)^2 - 4\rho}. \quad (5.19)$$

The expression $F_\lambda^2 = (\lambda - u)^2 - 4\rho$ is the discriminant of the Brunelli–Das spectral curve (5.6). Solving $L = \lambda$ for p ,

$$p_\pm(u, \rho; \lambda) = \frac{1}{2}[(\lambda - u) \pm \sqrt{(\lambda - u)^2 - 4\rho}] = \frac{1}{2}[(\lambda - u) \pm F_\lambda], \quad (5.20)$$

so $p_+ - p_- = F_\lambda$. The algebraic continuum (5.19) is therefore the Brunelli–Das pair recast in hodograph coordinates rather than a genuinely new Lax representation. What is gained is that F_λ is irrational in p and lives outside any Laurent layer in p . This is the shallow-water analogue of the observation in Section 4 that the $S_x^m = au + \lambda$ family of inviscid Burgers is generated by fractional-power conservation laws and is invisible to a polynomial-in- S_x search.

The dressing transform. The two families above are connected by a third construction. The EPD equation is linear, so any superposition of solutions is again a solution. For any contour $C \subset \mathbb{C}$ and any measure $m(\lambda) d\lambda$ for which the integrals below converge,

$$F_m(u, \rho) = \int_C m(\lambda) F_\lambda(u, \rho) d\lambda, \quad G_m(u, \rho) = \int_C m(\lambda) G_\lambda(u, \rho) d\lambda \quad (5.21)$$

solve the EPD equation and the conservation law $\partial_t F_m = \partial_x G_m$ respectively, by (formally) passing the (x, t) -derivatives through the λ -integral. The hodograph pair $S_x = F_m$, $S_t = G_m$ is therefore a valid Lax representation for every choice of (C, m) . Two specializations make this concrete.

If m is a finite sum of point masses, $m(\lambda) = \sum_{j=1}^N \alpha_j \delta(\lambda - \lambda_j)$, then $F_m = \sum_{j=1}^N \alpha_j F_{\lambda_j}$ is a linear combination of the algebraic continuum at N chosen spectral points. For $N = 2$ with $\alpha_1 = -\alpha_2 = 1$ and distinct λ_1, λ_2 ,

$$F(u, \rho) = \sqrt{(\lambda_1 - u)^2 - 4\rho} - \sqrt{(\lambda_2 - u)^2 - 4\rho}$$

is irrational in (u, ρ) and inherits its content from two distinct points on the spectral curve. It is a perfectly valid hodograph Lax representation that no Laurent-in- p ansatz could capture.

The more interesting specialization extracts Laurent coefficients of F_λ at $\lambda = \infty$. Choose C to be a counterclockwise circle of radius $R \rightarrow \infty$ enclosing the branch cut of F_λ , and take $m(\lambda) = \lambda^{-n-2}/(2\pi i)$ for $n \geq 0$. For $|\lambda| > |u| + 2\sqrt{\rho}$ the radicand is positive, and

$$F_\lambda(u, \rho) = \lambda \sqrt{1 - \frac{2u}{\lambda} + \frac{u^2 - 4\rho}{\lambda^2}}.$$

Applying the binomial series $\sqrt{1+w} = 1 + w/2 - w^2/8 + w^3/16 - \dots$ with $w = -2u/\lambda + (u^2 - 4\rho)/\lambda^2$ and collecting powers of λ produces

$$F_\lambda(u, \rho) = \sum_{n \geq 0} \mathcal{F}_n(u, \rho) \lambda^{1-n}. \quad (5.22)$$

The first three coefficients come straight from bookkeeping: at order λ^1 the leading 1 inside the root gives $\mathcal{F}_0 = 1$; at order λ^0 only $w/2$ contributes, giving $\mathcal{F}_1 = -u$; at order λ^{-1} the term $(u^2 - 4\rho)/(2\lambda^2)$ from $w/2$ combines with $-u^2/(2\lambda^2)$ from $-w^2/8$ to give $\mathcal{F}_2 = -2\rho$. Higher orders continue mechanically. Similarly,

$$G_\lambda(u, \rho) = -\frac{1}{2}(\lambda + u) F_\lambda(u, \rho) = \sum_{n \geq 0} \mathcal{G}_n(u, \rho) \lambda^{2-n}. \quad (5.23)$$

The Cauchy residue theorem identifies these Laurent coefficients as contour integrals,

$$\mathcal{F}_n = \frac{1}{2\pi i} \oint_{|\lambda|=R} \lambda^{n-2} F_\lambda d\lambda, \quad \mathcal{G}_n = \frac{1}{2\pi i} \oint_{|\lambda|=R} \lambda^{n-3} G_\lambda d\lambda,$$

so each $\mathcal{F}_n, \mathcal{G}_n$ is the dressing transform (5.21) of the corresponding power-law measure. The first six pairs are

n	0	1	2	3	4	5	
$\mathcal{F}_n$	1	$-u$	-2ρ	$-2u\rho$	$-2\rho(\rho + u^2)$	$-2u\rho(3\rho + u^2)$,	(5.24)
$\mathcal{G}_n$	$-\frac{1}{2}$	0	$\rho + \frac{1}{2}u^2$	$2u\rho$	$\rho(\rho + 2u^2)$	$2u\rho(2\rho + u^2)$.	

Substituting (5.22) and (5.23) into $\partial_t F_\lambda = \partial_x G_\lambda$ and matching coefficients of λ^k yields a tower of conservation laws on the shallow-water flow. The factor of λ in $G_\lambda = -\frac{1}{2}(\lambda + u)F_\lambda$ offsets the Laurent indexing by one: the term in F_λ at power λ^k is $\mathcal{F}_{1-k}$, the term in G_λ at the same power is $\mathcal{G}_{2-k}$, and density $\mathcal{F}_n$ pairs with flux $\mathcal{G}_{n+1}$:

$$\partial_t \mathcal{F}_n = \partial_x \mathcal{G}_{n+1}, \quad n \geq 0. \quad (5.25)$$

The first three nontrivial cases, read off from (5.24), are

$n = 1:$	$\partial_t(-u) = \partial_x(\rho + \frac{1}{2}u^2)$	(momentum),
$n = 2:$	$\partial_t(-2\rho) = \partial_x(2u\rho)$	(mass),
$n = 3:$	$\partial_t(-2u\rho) = \partial_x(\rho(\rho + 2u^2))$	(energy).

These are the classical momentum, mass, and energy conservation laws of shallow water, three consecutive rungs in an infinite tower generated by the single algebraic function $F_\lambda = \sqrt{(\lambda - u)^2 - 4\rho}$. The remainder of the tower is the Kupershmidt–Manin / Whitham hierarchy (5.16), up to overall sign and normalization inherited from the F_λ ansatz. The hodograph Lax pair at level n is $S_x = \mathcal{F}_n$, $S_t = \mathcal{G}_{n+1}$.

This section shows that each measure on $\mathbb{C} \cup \{\infty\}$ defines a hodograph Lax pair through (5.21), so shallow water has an infinite-dimensional family of Lax representations. Power-law measures at ∞ produce the Whitham hierarchy and the classical conservation laws via Cauchy residues. Point masses at finite λ produce the algebraic continuum F_λ and combinations $\sum_j \alpha_j F_{\lambda_j}$. Smooth densities on a contour interpolate between these extremes. Thus, the “Lax pairs galore” phenomenon of Calogero and Nucci [13] appears here as a function space of representations parameterized by measures on $\mathbb{C} \cup \{\infty\}$.

6 A First-Order Lax Operator for the KdV equation

The four case studies so far have exhibited Lax pairs that are anomalous in various limited senses: pairs that are spectrally collapsed, infinite-parametric, or arranged into a function space of representations parameterized by a measure. This section presents a case where SILO returned something substantially farther from the classical picture: a first-order scalar Lax operator for the KdV equation, living alongside the classical second-order Schrödinger operator in the Lax landscape of KdV. This first-order pair was first reported in [3], and numerical details of that study can be found there. Here, we study some of its properties.

The first-order pair is interesting for two distinct reasons. Diagnostically, it is by every classical measure too simple to carry integrable structure, since, as we will show, its L^2 point spectrum is empty, its periodic eigenvalues are uniformly spaced and encode only the mean of u , and its monodromy reduces to the mass. Despite this, the full KdV conservation hierarchy is still present in the operator algebra of its powers, through a mechanism that connects the operator to the classical Bell-polynomial formalism for soliton equations [23, 42, 43] once the connection is made explicit. As mentioned in the introduction, the verdict that such conservation-law pairs are spectrally inert is itself longstanding, going back to the 1980s. What the present section shows is that this inertness coexists with a full operator-algebraic recovery of the conservation hierarchy, so the standard verdict, while correct about the spectrum, is incomplete about the algebra.

6.1 The first-order Lax pair and its fakeness

The KdV equation in the comoving frame reads

$$u_t = 6uu_x + u_{xxx}. \quad (6.1)$$

Its classical Lax pair, due to Lax [44], is built from the self-adjoint Schrödinger operator $L_{\text{Schr}} = -u - \partial_x^2$ and the skew-adjoint $P_{\text{Schr}} = -4\partial_x^3 - 6u\partial_x - 3u_x$, with $\partial_t L = [L, P]$ equivalent to (6.1). The spectral theory of L_{Schr} provides the machinery for exact solution and, through the trace formula $d[\text{tr}(L_{\text{Schr}}^s)]/dt = 0$, generates the infinite hierarchy of KdV conservation laws via the Gel’fand–Dikii polynomials [28].

Applied to KdV with a wide operator hypothesis, SILO recovers this classical pair and, alongside it, a second pair of entirely different character, written here in an algebraically convenient form:

$$L = u + \partial_x, \quad P = \kappa u + 3u^2 + u_{xx} + \kappa \partial_x = 3u^2 + u_{xx} + \kappa L, \quad (6.2)$$

where $\kappa \in \mathbb{R}$ is a free parameter.

Theorem 6.1 (First-order Lax pair for KdV). *The pair (6.2) satisfies $\partial_t L = [L, P]$ on solutions of (6.1), acting on $C^3(\mathbb{R})$.*

Proof. Write $L = u + D$ with $D = \partial_x$, and $P = (3u^2 + u_{xx}) + \kappa L$. Since $[L, \kappa L] = 0$,

$$[L, P]\varphi = [L, 3u^2 + u_{xx}]\varphi = (6uu_x + u_{xxx})\varphi,$$

and equating with $(\partial_t L)\varphi = u_t \varphi$ gives (6.1). $\square$

Two different diagnostics show that the pair (6.2) is fake. The Krichever integrability test [39] compares two dispersion laws obtained by linearizing at a constant background $u = u_0$: the law $\omega_{\text{KdV}}(\xi)$ coming from the linearized KdV equation, and the law $\omega_{\text{Lax}}(\xi)$ coming from the auxiliary linear system $L\varphi = \lambda\varphi$, $\varphi_t = -P\varphi$. For a Lax pair carrying algebro-geometric content, the two agree (up to overall normalization) and are polynomial in the wavenumber ξ . For our first-order pair, $\omega_{\text{KdV}}(\xi) = \xi^3 - 6u_0\xi$ is cubic, while a short calculation (Appendix A.3) gives $\omega_{\text{Lax}}(\xi) = -3iu_0^2$, constant in ξ . The pair satisfies Krichever's polynomiality requirement only in the degenerate sense that a constant is a polynomial of degree zero, and encodes none of the cubic ξ -structure that the KdV dispersion carries.

Additionally, the first-order pair is considered fake according to the Butler and Hay diagnostic [12]. Setting $S = e^U$ with $U = \int u dx$, the gauge transformation $\varphi \mapsto S\varphi$ conjugates the Lax operator to

$$SLS^{-1} = e^U(u + \partial_x)e^{-U} = \partial_x,$$

a constant-coefficient operator in which every trace of u has been removed. Butler and Hay take precisely this reducibility as a sufficient condition for fake-ness: a Lax pair gauge-equivalent to one independent of the dependent variable cannot have been encoding the nonlinear dynamics, however neatly the compatibility identity is satisfied. The same verdict follows from their second test, an excess-freedom count on the auxiliary system, which Appendix A.2 works out alongside the gauge construction.

6.2 Spectral theory of the first-order operator

Spectrum on $\mathbb{R}$. Before turning to the algebraic structure that carries the KdV content, we record the spectral picture for the operator L defined in (6.2). To this end, let

$$U(x) := \int_0^x u(s) ds. \quad (6.3)$$

The eigenvalue problem $L\varphi = \lambda\varphi$ reads $\varphi'(x) = (\lambda - u(x))\varphi(x)$, with explicit solution

$$\varphi(x; \lambda) = C \exp(\lambda x - U(x)), \quad (6.4)$$

with $U(x)$ given in (6.3) and C an arbitrary constant. The following proposition might seem obvious, but it is nevertheless useful to write down the result.

Proposition 6.2. *For localized $u \in L^1(\mathbb{R})$, the operator $L = u + \partial_x$ has no $L^2(\mathbb{R})$ eigenvalues for any $\lambda \in \mathbb{C}$.*

Proof. As $|x| \rightarrow \infty$, $U(x) \rightarrow U_{\pm}$, which is finite for $u \in L^1(\mathbb{R})$, so $|\varphi(x; \lambda)| \sim |C_{\pm}| e^{(\text{Re } \lambda)x}$. Membership in $L^2(\mathbb{R})$ would require $\text{Re } \lambda < 0$ (to decay at $+\infty$) and $\text{Re } \lambda > 0$ (to decay at $-\infty$), which is impossible to combine. Hence no $\lambda \in \mathbb{C}$ admits an L^2 eigenfunction. $\square$

Another direct computation reinforces this degenerate spectral picture. The eigenfunction (6.4) has the normalized transfer multiplier

$$\lim_{x \rightarrow +\infty} e^{-\lambda x} \varphi(x; \lambda) / \lim_{x \rightarrow -\infty} e^{-\lambda x} \varphi(x; \lambda) = \exp\left(-\int_{-\infty}^{\infty} u dx\right),$$

which is independent of λ and encodes only the mass. Thus even the natural first-order analogue of scattering data contains no spectral dependence. In particular, there is no counterpart of the nontrivial reflection/transmission data that drive the Schrödinger inverse-scattering theory [2, 22].

Periodic spectrum. Without loss of generality (thanks to the scaling invariance of the KdV equation), we choose the spatial period to be 2π . The periodic spectrum of L is characterized by the following:

Theorem 6.3. *On the interval $[0, 2\pi]$ with periodic boundary conditions, the eigenvalues of L are $\lambda_n = \bar{u} + in$ for $n \in \mathbb{Z}$, with eigenfunctions $\varphi_n(x) = w(x) e^{inx}$ where $w(x) = \exp(-\int_0^x (u - \bar{u}) ds)$.*

Proof. Recall that the general solution of the eigenvalue problem $L\varphi = \lambda\varphi$ is given by (6.4) with $C := \varphi(0)$. To impose periodicity, separate the mean of u from its fluctuation. Set

$$\bar{u} := \frac{1}{2\pi} \int_0^{2\pi} u(s) ds, \quad U_0(x) := \int_0^x (u(s) - \bar{u}) ds.$$

Then $\int_0^x u(s) ds = \bar{u}x + U_0(x)$, and (6.4) factors as

$$\varphi(x; \lambda) = C \exp((\lambda - \bar{u})x) w(x), \quad w(x) := e^{-U_0(x)}. \quad (6.5)$$

The function w is 2π -periodic, since

$$U_0(x + 2\pi) - U_0(x) = \int_x^{x+2\pi} (u(s) - \bar{u}) ds = \int_0^{2\pi} (u(s) - \bar{u}) ds = 0$$

by translation invariance of the integral over a full period and the definition of $\bar{u}$. In particular, $w(0) = w(2\pi) = 1$.

The periodic boundary condition $\varphi(2\pi; \lambda) = \varphi(0; \lambda)$ applied to (6.5) yields $C \exp(2\pi(\lambda - \bar{u})) \cdot w(2\pi) = C \cdot w(0)$, which, after cancelling C and using $w(2\pi) = w(0) = 1$, reduces to $\exp(2\pi(\lambda - \bar{u})) = 1$. The complex exponential equals 1 precisely when its argument is an integer multiple of $2\pi i$, so $2\pi(\lambda - \bar{u}) = 2\pi i n$ for $n \in \mathbb{Z}$, giving $\lambda_n = \bar{u} + in$. Substituting $\lambda = \lambda_n$ into (6.5) and absorbing C into the normalization, $\varphi_n(x) = w(x) e^{inx}$, which is the claimed eigenfunction. $\square$

Thus, the eigenvalues $\lambda_n = \bar{u} + in$ lie on a single vertical line in $\mathbb{C}$ at uniform spacing, with real part fixed by the mean $\bar{u}$ and imaginary part labelled by $n \in \mathbb{Z}$. Consequently, the eigenfunctions $\varphi_n(x) = w(x) e^{inx}$ are just weighted Fourier modes with weight $w(x) = e^{-U_0(x)}$. There are no bands, no gaps, no nontrivial Floquet discriminant [50]. The Schrödinger-Gel'fand-Dikii route [28], which extracts the KdV hierarchy from the resolvent expansion of L_{Schr} , has no analogue here.

6.3 From the first-order pair to the KdV conservation hierarchy

The diagnostics in Section 6.1 declare the first-order pair (6.2) fake, yet the KdV conservation hierarchy is encoded in the operator algebra generated by powers of L in the ungauged frame. We set $\kappa = 0$ in (6.2) to drop the κL tail, which lies in the centralizer of L and contributes nothing, leaving

$$(L, P) = (u + \partial_x, 3u^2 + u_{xx}). \quad (6.6)$$

Here $P = 3u^2 + u_{xx}$ is the KdV flux, and the Lax equation collapses to the conservation form of (6.1),

$$[L, P] = [\partial_x, 3u^2 + u_{xx}] = 6uu_x + u_{xxx} = u_t, \quad u_t = \partial_x P.$$

Remark 6.4. This is a special case of a general freedom in the Lax representation, noted to us by B. Konopelchenko [34]. The same equation is obtained from a modified pair in which $P \mapsto P + F(L)$ for any function F , since $[L, F(L)] = 0$, so the time operator is determined only modulo the centralizer of L . The κL tail is the linear instance $F(L) = \kappa L$. This centralizer freedom is the operator-algebra counterpart of the recurring lesson of this paper, that compatibility underdetermines the Lax representation, joining the gauge, spectral-shift, and hodograph freedoms encountered in the previous sections.

The machinery that extracts the hierarchy from this pair is classical, and we use it as such. The densities produced by L are Bell polynomials [5, 15], conserved densities are read off by the direct method of Göktaş and Hereman [31], which rests on the variational fact that a density is conserved exactly when its time derivative is a total spatial derivative [56], and the Bell-polynomial form of this calculation is standard for soliton equations [23, 30, 43]. What we add is the operator that feeds the machine. The densities are manufactured by the spectrally fake pair (6.6), and the construction asks nothing of the equation beyond its conservation form $u_t = \partial_x P$, so the same device runs on any scalar evolution equation written that way. We keep the exposition self-contained so the procedure can be run without chasing references, but none of the underlying identities below are new.

Bell polynomials. The operator $L = u + \partial_x$ has the gauge presentation

$$L = e^{-U} \partial_x e^U, \quad U := \int u dx, \quad (6.7)$$

checked directly: $e^{-U} \partial_x (e^U f) = u f + f_x = L f$. Induction gives $L^s = e^{-U} \partial_x^s e^U$, and acting on the constant function 1,

$$\rho_s := L^s \cdot 1 = e^{-U} \partial_x^s (e^U) = Y_s(u, u_x, \dots, u^{(s-1)}), \quad (6.8)$$

the s -th complete exponential Bell polynomial [5, 15], with first values

$$\begin{aligned} \rho_1 &= u, & \rho_2 &= u^2 + u_x, & \rho_3 &= u^3 + 3u u_x + u_{xx}, \\ \rho_4 &= u^4 + 6u^2 u_x + 3u_x^2 + 4u u_{xx} + u_{xxx}, \\ \rho_5 &= u^5 + 10u^3 u_x + 15u u_x^2 + 10u^2 u_{xx} + 10u_x u_{xx} + 5u u_{xxx} + u_{xxxx}. \end{aligned}$$

The complete Bell polynomials and their cumulant inverse are mutually inverse polynomial changes of variable [15], so

$$u^{(j)} = \widehat{Y}_{j+1}(\rho_1, \dots, \rho_{j+1}), \quad j \geq 0, \quad (6.9)$$

with first instances $u = \rho_1$, $u_x = \rho_2 - \rho_1^2$, and $u_{xx} = \rho_3 - 3\rho_1 \rho_2 + 2\rho_1^3$. The generating sets $\{u, u_x, u_{xx}, \dots\}$ and $\{\rho_1, \rho_2, \dots\}$ describe the same polynomials in u and its derivatives, and we move freely between them.

The two recursions. Both ∂_t and ∂_x act on the ρ_s by explicit polynomial rules. For the time derivative, the Lax equation $\partial_t L = [L, P]$ telescopes to $\partial_t L^s = [L^s, P]$. Applying this to 1 gives $\partial_t \rho_s = L^s P - P \rho_s$, and expanding $L^s P = e^{-U} \partial_x^s (e^U P)$ by Leibniz with $\partial_x^k e^U = \rho_k e^U$ leaves, after the $j = 0$ term cancels $P \rho_s$,

$$\partial_t \rho_s = \sum_{j=1}^s \binom{s}{j} \rho_{s-j} P^{(j)}, \quad P^{(j)} := D_x^j P, \quad s \geq 1. \quad (6.10)$$

The flux derivatives are explicit, $P^{(0)} = 3u^2 + u_{xx}$, $P^{(1)} = 6u u_x + u_{xxx}$, $P^{(2)} = 6u_x^2 + 6u u_{xx} + u_{xxxx}$, and so on, and with Leibniz (6.10) propagates ∂_t to every polynomial in the ρ_s . For the space derivative, $L \rho_s = \rho_{s+1}$ and $L = u + \partial_x$ give

$$(\rho_s)_x = \rho_{s+1} - \rho_1 \rho_s, \quad (6.11)$$

which evaluates D_x inside $R_\rho := \mathbb{Q}[\rho_1, \rho_2, \dots]$. This D_x is the ordinary total x -derivative written in the ρ variables, as one sees by differentiating the generating function $A(t) = e^{U(x+t) - U(x)}$ in x , so reducing modulo total x -derivatives means the same thing in either generating set.

Reading off the hierarchy. A local functional $\int F dx$ is conserved exactly when $\partial_t F$ is a total x -derivative [56], so the conserved densities are the kernel of ∂_t acting on R_ρ modulo $D_x R_\rho$, and this kernel is a finite computation level by level. Grade R_ρ by subscript sum, with ρ_s in degree s and $\rho_{i_1} \cdots \rho_{i_k}$ in degree $i_1 + \cdots + i_k$; the monomials of subscript sum at most N are finite in number. Take the ansatz

$$F = \sum_{|\alpha| \leq N} c_\alpha m_\alpha, \quad m_\alpha = \rho_{i_1} \cdots \rho_{i_k}, \quad |\alpha| = i_1 + \cdots + i_k,$$

compute $\partial_t F$ from (6.10) with Leibniz, and reduce modulo $D_x R_\rho$ by (6.11), which is integration by parts. Requiring the surviving residue monomials to cancel, $\partial_t F \equiv 0$, is a homogeneous linear system $Mc = 0$ over $\mathbb{Q}$, and $\ker M$ is the space of KdV conservation densities of subscript sum at most N . This is the Göktaş–Hereman direct method [31] carried out in the Bell variables natural to L , in exact rational arithmetic, with no spectral information about L entering at any step. Appendix A.1 states the procedure as an algorithm and runs it to recover the first five KdV conserved quantities.

The hierarchy that comes out is the standard one, and the route to it through the Schrödinger spectrum is unnecessary, as Miura, Gardner, and Kruskal showed by other means [54]. The point we draw is the provenance. The densities are manufactured by the pair (6.6), which Section 6.1 certifies as fake by the Krichever and Butler–Hay tests, and the conservation content rides on the operator algebra of L in a fixed gauge, the structure those gauge-invariant tests are built to discard. A pair with no spectral content still delivers the full hierarchy.

7 Concluding Remarks

The five case studies presented in this work share one common lesson: compatibility underdetermines the Lax representation. A summary of the five case studies pursued in this paper and the SILO variant under which each pair was surfaced is given in Table 1. Every equation studied in this work is classically integrable or linearizable. Thus, the object of study here was not integrability, but rather the shape of the representations that compatibility allows, and the present study demonstrated that many possible scenarios can arise.

The Euler top returned rank-deficient pairs that satisfy $\dot{L} = [L, P]$ while encoding less than the dynamics, with skew-symmetry and a Frobenius product penalty needed to surface a Manakov skeleton whose spectral completion is then linear algebra on $\mathfrak{so}(3)$. The free Schrödinger equation showed the nondegeneracy penalty itself choosing the pair: a determinant barrier returns the textbook UTM pair, a Frobenius surrogate returns a spectrally empty pair carrying a free function $M(k)$. Inviscid Burgers carried an infinite gauge freedom in its Clairaut compatibility and an exponent continuum $S_x^m = au + \lambda$, both instances of the Calogero–Nucci hodograph mechanism on the fractional-power conservation laws. Shallow water widened that continuum into a function space of hodograph pairs parametrized by measures on $\mathbb{C} \cup \{\infty\}$, with the Brunelli–Das pair the unique Laurent representative and the Whitham hierarchy and spectral-discriminant continuum lying outside the Laurent layer entirely. None of these degeneracies is visible to the compatibility condition, which is satisfied by all of them at once.

The first-order KdV pair $L = u + \partial_x$ is the sharpest case, and it is worth ending on. By every gauge-invariant diagnostic it is fake. The Krichever dispersion collapses to a constant (Appendix A.3), and both Butler–Hay tests apply (Appendix A.2): the gauge $S = e^U$ sends L to the bare ∂_x , and the excess-freedom count leaves the flux q unconstrained. After that gauge $\partial_x^s \cdot 1 = 0$ and the whole construction collapses, exactly as the classification predicts. Yet in the fixed gauge $v = \int u$ the same pair drives the classical direct method to the full KdV hierarchy, I_1 through I_5 at subscript-sum levels 1, 2, 4, 6, 8 and onward. The two verdicts do not contradict each other. They measure different things. The gauge-invariant tests read the spectral content, which is empty, while the conservation

System	Ansatz	Compatibility	Degeneracy control
$I\dot{\Omega} + \Omega \times (I\Omega) = 0$	Linear in Ω , skew-symmetric	Lie–Poisson chain rule $\{H, L\} = [L, P]$	Frobenius product penalty $\prod_k \ \partial_{\Omega_k} \tilde{L}\ _F^2 \ \partial_{\Omega_k} \tilde{P}\ _F^2$
$i q_t + q_{xx} = 0$	Affine $(\tilde{f}, \tilde{g})$, polynomial in k	Clairaut $\mu_{xt} = \tilde{f}_t = \tilde{g}_x = \mu_{tx}$	$ \det(Z^\dagger Z) ^{-2}$ barrier, or Frobenius penalty $\ Z^+\ _F^2$
$u_t + uu_x = 0$	Polynomial in (u, s) , $s = S_x$	Clairaut $2s g_u + f_u(g_s + u) = 0$	ℓ^∞ product penalty $\ \Xi\ _\infty^{-1} \ \mathcal{Z}\ _\infty^{-1}$
$\rho_t + (\rho u)_x = 0$ $u_t + uu_x + \rho_x = 0$	L Laurent in p , $B = (\frac{1}{2}L^2)_{>0}$	Chain rule on (u, ρ) , $L_t = \{L, B\}$	Inverse-mean product $1/\mathbb{E}_V[\tilde{L}_u^2 \tilde{L}_p^2]$
$u_t - 6uu_x - u_{xxx} = 0$	$\tilde{L} = \xi_1 u + \xi_2 \partial_x + \xi_3 \partial_x^2$, $\tilde{P}$ general tensor	Hamiltonian chain rule $\partial_t \tilde{L} = (\partial_u \tilde{L}) Q \frac{\delta H}{\delta u} = [\tilde{L}, \tilde{P}]$	Inverse-norm penalty $\ (\partial_u \tilde{L}) Q(\delta H / \delta u)\ _2^{-2}$

Table 1: A summary of the five case studies pursued throughout this paper (top to bottom: Euler top, free Schrödinger, inviscid Burgers, shallow water, first-order KdV) and the SILO variant under which each pair was surfaced. Degeneracy control takes a different form in every case, and each variant reflects the structural geometry of the underlying problem. The first-order KdV formulation and numerical details are from [3] and not rerun here.

content lives in the operator algebra of L in a fixed gauge, the structure those tests are built to quotient away.

We do not propose a refined taxonomy. The Butler–Hay test is correct for what it tests, the conservation content here is gauge-dependent, and the two simply do not see each other. Butler and Hay anticipated the gap, writing that “the extent to which a fake Lax pair can be used to gain information about its associated system is also open to debate.” Section 6 is one answer for the conservation-form class: a pair declared fake on gauge-invariant grounds still carries the full conservation ring of an arbitrary $u_t = q_x$ when read in the gauge those grounds discard. Whether other fake pairs in the literature hide algebraic content of this kind, and whether the Butler–Hay criteria admit a gauge-fixed refinement separating the algebraically active fakes from the empty ones, are questions we leave open.

The machinery doing the work is classical, as Section 6.3 points out. The hierarchy is manufactured by a spectrally fake pair, and the construction needs nothing of the equation past its conservation form, so the same device runs on any scalar $u_t = q_x$. The universality is a special case of Calogero–Nucci, and the combination we use, the substrate $(u + \partial_x, q)$ in the gauge $v = \int u$, is elementary enough that we make no claim of priority.

The broader point returns to where the paper started. Anomalous Lax pairs are not pathologies at the edge of the theory. They are regular inhabitants of the Lax landscape that any discovery procedure, sparse regression or otherwise, will keep returning, because compatibility alone cannot tell a spectrally informative pair from a degenerate or fake one. The burden of telling them apart falls on the search, on the ansatz, the gauge, and the penalty geometry, not on the compatibility condition, which certifies all of them equally. As data-driven methods providing such certification continue to develop, it will be important to continue expanding the relevant detailed pair characterization and whether the latter can be used to extract the infinite hierarchy of conservation laws (of the integrable system) or/and whether it can be leveraged to develop systematically the inverse scattering analysis of the system at hand. A complementary question is whether these fake pairs, beyond carrying the conservation hierarchy, can be used to constructively recover the genuine spectral pair of the underlying equation. For scalar conservation laws of Gelfand–Dikii type,

this turns out to be possible. The first-order operator supplies the differential skeleton through its antisymmetric part $\partial_x = \frac{1}{2}(L - L^\dagger)$, its conserved-quantity gradients supply the field content, and the Krichever test certifies the result. We will develop this constructive converse, which is the closest scalar analogue of the Manakov completion of Section 2, in forthcoming work.

Acknowledgements

J.A. gratefully acknowledges helpful discussions during the early stages of this work with Roy H. Goodman and Luis García-Naranjo about integrable rigid body dynamics and Manakov completions. J.A. also thanks Ildar Gabitov for pointing out the Krichever test and Oleksandr Minakov and Ibrahim Fatkullin for separate discussions concerning the spectral theory of the first order KdV Lax pair. The authors thank Boris Konopelchenko for a careful reading of an earlier draft and for the historical and structural comments that sharpened Sections 6 and the discussion of representation freedom. J.A. acknowledges support from NSF grant PHY-2316622. W.Z. acknowledges support from the National Science Foundation under awards DMS-2502900 and DMS-2540370, and from the Air Force Office of Scientific Research under Grant No. FA9550-25-1-0079. G.B. acknowledges support from the Simons foundation under grant number SFI-MPS-TSM-00013369. This material is based upon work supported by the National Science Foundation under Grant No. PHY-2408988 (PGK). This research was partly conducted while P.G.K. was visiting the Okinawa Institute of Science and Technology (OIST) through the Theoretical Sciences Visiting Program (TSVP), the University of Sydney through the visitor program of the Sydney Mathematical Research Institute (SMRI) and the Department of Mechanical Engineering at Seoul National University through a Fulbright Fellowship. Their support is gratefully acknowledged. This work was also supported by a grant from the Simons Foundation [SFI-MPS-SFM-00011048, P.G.K]. The authors collectively thank the Banff International Research Station (BIRS) at the Instituto de Matemáticas de la Universidad de Granada and the Institute for Computational and Experimental Mathematics (ICERM) where portions of this work were carried out during week long research workshops.

Conflict of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Use of AI Tools

The authors used Claude Opus 4.8 to verify algebraic and symbolic computations and to assist with literature review and positioning. All derivations, results, and text were produced and checked by the authors, who take full responsibility for the content.

References

- [1] M. J. Ablowitz, D. J. Kaup, A. C. Newell, and H. Segur. The inverse scattering transform–Fourier analysis for nonlinear problems. *Studies in Applied Mathematics*, 53(4):249–315, 1974.
- [2] M. J. Ablowitz and H. Segur. *Solitons and the Inverse Scattering Transform*. SIAM, Philadelphia, 1981.
- [3] J. Adriaola, W. Zhu, P. G. Kevrekidis, and A. Aceves. Computer assisted discovery of integrability via SILO: sparse identification of Lax operators. *SIAM Journal on Applied Dynamical Systems*, 25(1):131–159, 2026.
- [4] O. Babelon, D. Bernard, and M. Talon. *Introduction to Classical Integrable Systems*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, Cambridge, 2003.
- [5] E. T. Bell. Exponential polynomials. *Annals of Mathematics*, 35(2):258–277, 1934.

- [6] A. I. Bobenko, A. G. Reyman, and M. A. Semenov-Tian-Shansky. The Kowalewski top 99 years later: A Lax pair, generalizations and explicit solutions. Communications in Mathematical Physics, 122(2):321–354, 1989.
- [7] A. V. Bolsinov and A. T. Fomenko. Integrable Hamiltonian Systems: Geometry, Topology, Classification. Chapman & Hall/CRC, Boca Raton, 2004.
- [8] T. Bridgman, W. Hereman, G. R. W. Quispel, and P. H. van der Kamp. Symbolic computation of Lax pairs of partial difference equations using consistency around the cube. Foundations of Computational Mathematics, 13(4):517–544, 2013.
- [9] J. C. Brunelli. Dispersionless limit of integrable models. Brazilian Journal of Physics, 30(2):455–468, 2000. Preprint [nlin/0207042](#).
- [10] J. C. Brunelli and A. Das. A Lax description for polytropic gas dynamics. Physics Letters A, 235(6):597–602, 1997. Preprint [solv-int/9706005](#).
- [11] G. I. Burde. Lax pairs for the modified kdv equation. Axioms, 13(2), 2024.
- [12] S. Butler and M. Hay. Simple identification of fake Lax pairs. arXiv preprint, 2013.
- [13] F. Calogero and M. C. Nucci. Lax pairs galore. Journal of Mathematical Physics, 32:72–74, 1991.
- [14] G. F. Carrier and H. P. Greenspan. Water waves of finite amplitude on a sloping beach. Journal of Fluid Mechanics, 4(1):97–109, 1958.
- [15] L. Comtet. Advanced Combinatorics: The Art of Finite and Infinite Expansions. D. Reidel, Dordrecht, 1974.
- [16] A. Constandache, A. Das, and F. Toppan. Lucas polynomials and a standard Lax representation for the polytropic gas dynamics. Letters in Mathematical Physics, 60:197–209, 2002.
- [17] R. Courant and D. Hilbert. Methods of Mathematical Physics, Vol. II: Partial Differential Equations. Wiley-Interscience, New York, 1962.
- [18] P. de Koster and S. Wahls. Data-driven identification of the spectral operator in AKNS Lax pairs using conserved quantities. Wave Motion, 127:103273, 2024.
- [19] L. A. Dickey. Soliton Equations and Hamiltonian Systems, volume 26 of Advanced Series in Mathematical Physics. World Scientific, Singapore, 2nd edition, 2003.
- [20] A. Doikou. Selected topics in classical integrability. International Journal of Modern Physics A, 27(5):1230003, 2012.
- [21] P. G. Drazin and R. S. Johnson. Solitons: An Introduction. Cambridge Texts in Applied Mathematics. Cambridge University Press, Cambridge, 1989.
- [22] L. D. Faddeev and L. A. Takhtajan. Hamiltonian Methods in the Theory of Solitons. Classics in Mathematics. Springer, Berlin, 1987.
- [23] E. Fan. The integrability of nonisospectral and variable-coefficient KdV equation with binary Bell polynomials. Physics Letters A, 375(3):493–497, 2011.
- [24] A. S. Fokas. A unified transform method for solving linear and certain nonlinear PDEs. Proceedings of the Royal Society A, 453(1962):1411–1443, 1997.
- [25] A. S. Fokas. A Unified Approach to Boundary Value Problems. SIAM, 2008.
- [26] S. Friedlander and M. M. Vishik. Lax pair formulation for the Euler equation. Physics Letters A, 148(6–7):313–319, 1990.
- [27] C. S. Gardner, J. M. Greene, M. D. Kruskal, and R. M. Miura. Method for solving the Korteweg–de Vries equation. Physical Review Letters, 19:1095–1097, 1967.
- [28] I. M. Gel’fand and L. A. Dikii. Asymptotic behaviour of the resolvent of Sturm–Liouville equations and the algebra of the Korteweg–de Vries equations. Russian Mathematical Surveys, 30:77–113, 1975.
- [29] J.-F. Gerbeau and D. Lombardi. Approximated Lax pairs for the reduced order integration of nonlinear evolution equations. Journal of Computational Physics, 265:246–269, 2014.

- [30] C. Gilson, F. Lambert, J. J. C. Nimmo, and R. Willox. On the combinatorics of the Hirota D-operators. Proceedings of the Royal Society of London A, 452(1945):223–234, 1996.
- [31] U. Göktaş and W. Hereman. Symbolic computation of conserved densities for systems of nonlinear evolution equations. Journal of Symbolic Computation, 24(5–6):591–621, 1997.
- [32] G. Gubbiotti, C. Scimiterna, and D. Levi. Linearizability and a fake lax pair for a nonlinear nonautonomous quad-graph equation consistent around the cube. Theoretical and Mathematical Physics, 189:1459–1471, 2016.
- [33] S. Jin, C. D. Levermore, and D. W. McLaughlin. The semiclassical limit of the defocusing NLS hierarchy. Communications on Pure and Applied Mathematics, 52(5):613–654, 1999.
- [34] B. Konopelchenko, 2026. private communication.
- [35] B. Konopelchenko, L. Martínez Alonso, and E. Medina. Hodograph solutions of the dispersionless coupled KdV hierarchies, critical points and the Euler–Poisson–Darboux equation. Journal of Physics A: Mathematical and Theoretical, 43(43):434020, 2010. Preprint [arXiv:1003.2892](https://arxiv.org/abs/1003.2892).
- [36] O. Kotlyar, M. Pankratova, M. Kamalian-Kopae, A. Vasylenkova, J. E. Prilepsky, and S. K. Turitsyn. Combining nonlinear Fourier transform and neural network-based processing in optical communications. Optics Letters, 45(13):3462–3465, 2020.
- [37] V. V. Kozlov. Symmetries, Topology and Resonances in Hamiltonian Mechanics. Ergebnisse der Mathematik und ihrer Grenzgebiete. Springer, Berlin, 1993.
- [38] I. M. Krichever. Integration of nonlinear equations by the methods of algebraic geometry. Functional Analysis and Its Applications, 11(1):12–26, 1977.
- [39] I. M. Krichever. Methods of algebraic geometry in the theory of non-linear equations. Russian Mathematical Surveys, 32(6):185–213, 1977.
- [40] S. Krippendorf, D. Lüst, and M. Syvaeri. Integrability ex machina. Fortschritte der Physik, 69(7):2100057, 2021.
- [41] B. A. Kupershmidt and Y. I. Manin. Long-wave equations with a free surface. I. Conservation laws and solutions. Functional Analysis and Its Applications, 11(3):188–197, 1977. English translation of Funkts. Anal. Prilozhen. **11** (1977), no. 3, 31–42.
- [42] F. Lambert, S. Leblé, and J. Springael. Binary Bell polynomials and Darboux covariant Lax pairs. Glasgow Mathematical Journal, 43A:53–63, 2001.
- [43] F. Lambert and J. Springael. Soliton equations and simple combinatorics. Acta Applicandae Mathematicae, 102(2–3):147–178, 2008.
- [44] P. D. Lax. Integrals of nonlinear equations of evolution and solitary waves. Communications on Pure and Applied Mathematics, 21(5):467–490, 1968.
- [45] Y.-C. Lee, M. Brühl, D.-J. Doong, and S. Wahls. Nonlinear Fourier classification of 663 rogue waves measured in the Philippine Sea. PLOS ONE, 19(5):e0301709, 2024.
- [46] Y.-C. Lee and S. Wahls. Field observation of soliton gases in the deep open ocean, 2025. arXiv preprint.
- [47] Y. C. Li. A Lax pair for the two-dimensional Euler equation. Journal of Mathematical Physics, 42(8):3552–3553, 2001.
- [48] Y. C. Li and A. V. Yurov. Lax pairs and Darboux transformations for Euler equations. Studies in Applied Mathematics, 111(1):101–113, 2003.
- [49] S. Lin and Y. Chen. Lax-Pair-FIND: Discovering Lax pair from scarce data via deep learning. Chaos: An Interdisciplinary Journal of Nonlinear Science, 35(11):113120, 2025.
- [50] W. Magnus and S. Winkler. Hill’s Equation. Interscience Publishers, New York, 1966. Reprinted by Dover Publications, 1979.
- [51] S. V. Manakov. Note on the integration of Euler’s equations of the dynamics of an n -dimensional rigid body. Functional Analysis and Its Applications, 10(4):328–329, 1976.

- [52] A. S. Mishchenko and A. T. Fomenko. Euler equations on finite-dimensional Lie groups. Mathematics of the USSR-Izvestiya, 12(2):371–389, 1978.
- [53] R. M. Miura. Korteweg–de Vries equation and generalizations. I. A remarkable explicit nonlinear transformation. Journal of Mathematical Physics, 9(8):1202–1204, 1968.
- [54] R. M. Miura, C. S. Gardner, and M. D. Kruskal. Korteweg–de Vries equation and generalizations. II. Existence of conservation laws and constants of motion. Journal of Mathematical Physics, 9(8):1204–1209, 1968.
- [55] S. P. Novikov, S. V. Manakov, L. P. Pitaevskii, and V. E. Zakharov. Theory of solitons: the inverse scattering transform. Plenum, 1984.
- [56] P. J. Olver. Applications of Lie Groups to Differential Equations, volume 107 of Graduate Texts in Mathematics. Springer, New York, 2 edition, 1993.
- [57] A. R. Osborne. Nonlinear Fourier analysis: rogue waves in numerical modeling and data analysis. Journal of Marine Science and Engineering, 8(12):1005, 2020.
- [58] J. Pu and Y. Chen. Lax pairs informed neural networks solving integrable systems. Journal of Computational Physics, 510:113090, 2024. Preprint [arXiv:2401.04982](https://arxiv.org/abs/2401.04982).
- [59] A. G. Reyman. Group-theoretical methods in the theory of finite-dimensional integrable systems. Russian Mathematical Surveys, 35(1):57–91, 1980.
- [60] S. Sakovich. True and fake Lax pairs: how to distinguish them. Nonlinear Phenomena in Complex Systems, 23(3):338–341, 2020. Preprint [arXiv:nlin/0112027](https://arxiv.org/abs/nlin/0112027) (2001).
- [61] E. Sedov, I. Chekhovskoy, M. Fedoruk, and S. Turitsyn. Numerical approaches in nonlinear fourier transform-based signal processing for telecommunications. Studies in Applied Mathematics, 154(1):e12795, 2025.
- [62] S. P. Tsarev. The geometry of Hamiltonian systems of hydrodynamic type. The generalized hodograph method. Mathematics of the USSR-Izvestiya, 37(2):397–419, 1991.
- [63] S. K. Turitsyn, J. E. Prilepsky, S. T. Le, S. Wahls, L. L. Frumin, M. Kamalian, and S. A. Derevyanko. Nonlinear Fourier transform for optical data processing and transmission: advances and perspectives. Optica, 4(3):307–322, 2017.
- [64] H. D. Wahlquist and F. B. Estabrook. Prolongation structures of nonlinear evolution equations. Journal of Mathematical Physics, 16(1):1–7, 1975.
- [65] G. B. Whitham. Linear and Nonlinear Waves. Wiley-Interscience, New York, 1974.
- [66] V. Zakharov. Dispersionless limit of integrable systems in $2 + 1$ dimensions. In N. M. Ercolani, I. R. Gabitov, C. D. Levermore, and D. Serre, editors, Singular Limits of Dispersive Waves, pages 165–174. Plenum Press, 1994.
- [67] V. E. Zakharov and A. B. Shabat. Exact theory of two-dimensional self-focusing and one-dimensional self-modulation of waves in nonlinear media. Soviet Physics JETP, 34(1):62–69, 1972.

Appendix

A.1 KdV CQ Algorithm

Algorithm 1 summarizes the approach of section 6 to the conservation laws of the KdV equation. For further clarity and sake of concreteness, we run the above algorithm on the KdV equation, with $P = 3u^2 + u_{xx}$. We use the equivalence symbol $\equiv$ to denote equality modulo total D_x derivatives, and we work either on a periodic domain or on $\mathbb{R}$ with rapid decay.

Level 1. The single class $[\rho_1]$ has $\partial_t \rho_1 = P^{(1)} = (3u^2 + u_{xx})_x$, a total D_x . Conserved direction:

$$I_1 = \int u dx.$$

Level 2. The new class is $[\rho_2]$, with $[\rho_1^2] = [\rho_2]$ via (6.11). Compute

$$\partial_t \rho_2 = 2\rho_1 P^{(1)} + P^{(2)}.$$

Both terms reduce: $P^{(2)} = D_x P^{(1)}$ is a total derivative, and $2\rho_1 P^{(1)} = (2uP)_x - 2u_x P$ with $-2u_x P = -2u_x(3u^2 + u_{xx}) = -2(u^3)_x - (u_x^2)_x$. New conserved direction:

$$I_2 = \int u^2 dx.$$

Level 3. The new class $[\rho_3]$ is the unique level-3 representative (since $[\rho_3] = [\rho_1 \rho_2] = [\rho_1^3]$). Direct computation of

$$\partial_t \rho_3 = 3\rho_2 P^{(1)} + 3\rho_1 P^{(2)} + P^{(3)}$$

and reduction by integration by parts give $\partial_t \rho_3 \equiv 3u_x^3$, which is not a total derivative. Level 3 contributes no new direction.

Level 4. New representatives mod D_x : $[\rho_1 \rho_3]$, $[\rho_2^2]$, $[\rho_1^4]$, with $[\rho_4] = [\rho_1 \rho_3]$ and $[\rho_1^2 \rho_2] = [\rho_1^4]$ from (6.11). Combined with the level-3 carry-over $[\rho_3]$, the candidate space is four-dimensional. The reduced time derivatives sit in a two-dimensional residue space spanned by u_x^3 and uu_x^3 :

$$\partial_t \rho_3 \equiv 3u_x^3, \quad \partial_t(\rho_1 \rho_3) \equiv 12uu_x^3 - 6u_x^3, \quad \partial_t \rho_1^4 \equiv 12uu_x^3, \quad \partial_t \rho_2^2 \equiv 12uu_x^3 + 6u_x^3.$$

Algorithm 1 Conservation hierarchy for $u_t = q_x$ via the first-order Lax pair $(u + \partial_x, q)$.

Require: Conservation-form PDE $u_t = q_x$. Target subscript-sum level $N \geq 1$.

Ensure: A basis $\{F_\beta\}$ for the space of conservation densities at level $\leq N$.

```

1: Set  $L \leftarrow u + \partial_x$ ,  $P \leftarrow q$ ,  $\rho_0 \leftarrow 1$ .
2: for  $s = 1, \dots, N$  do ▷ Bell polynomials
3:    $\rho_s \leftarrow u \rho_{s-1} + (\rho_{s-1})_x$ .
4: end for
5: for  $j = 0, \dots, N$  do ▷ Spatial derivatives of  $P$ 
6:    $P^{(j)} \leftarrow D_x^j P$ .
7: end for
8: for  $s = 1, \dots, N$  do ▷ Time derivatives via (6.10)
9:    $\partial_t \rho_s \leftarrow \sum_{j=1}^s \binom{s}{j} \rho_{s-j} P^{(j)}$ .
10: end for
11: Enumerate partitions  $\alpha$  with  $|\alpha| \leq N$ ; form  $\rho$ -monomials  $m_\alpha = \rho_{i_1} \cdots \rho_{i_k}$ .
12: Reduce mod  $D_x$  via (6.11); pick representatives  $\{m_{\alpha_1}, \dots, m_{\alpha_n}\}$ .
13: for  $\alpha \in \{\alpha_1, \dots, \alpha_n\}$  do
14:   Compute  $\partial_t m_\alpha$  by Leibniz from  $\partial_t \rho_s$ .
15:   Reduce mod  $D_x$  via integration by parts to  $\partial_t m_\alpha \equiv \sum_{k=1}^m M_{k\alpha} R_k$ .
16: end for
17: Form the  $m \times n$  matrix  $M = (M_{k\alpha})$  over  $\mathbb{Q}$ .
18: Compute  $\ker M = \{c \in \mathbb{Q}^n : Mc = 0\}$ .
19: for each basis vector  $c^{(\beta)} \in \ker M$  do
20:    $F_\beta \leftarrow \sum_\alpha c_\alpha^{(\beta)} m_\alpha$ .
21: end for
22: return  $\{F_\beta\}$ . Each  $\int F_\beta dx$  is conserved.
```

Only one linear combination kills both residues. The unique new kernel direction is

$$F_3 = \rho_1^3 + \rho_1^2 \rho_2 - \frac{1}{2} \rho_2^2 - \frac{1}{2} \rho_1^4 \equiv u^3 - \frac{1}{2} u_x^2,$$

giving

$$I_3 = \int (u^3 - \frac{1}{2} u_x^2) dx. \quad (\text{A.1})$$

Level 5. The output residues from new level-5 candidates are independent of every residue seen at lower levels. The kernel does not extend.

Level 6. The kernel computation, with new candidates at subscript sum 6 and carry-overs from level 4, returns one new direction:

$$I_4 = \int (u^4 - 2uu_x^2 + \frac{1}{5} u_{xx}^2) dx, \quad (\text{A.2})$$

or equivalently, after clearing denominators, $5I_4 = \int (5u^4 - 10uu_x^2 + u_{xx}^2) dx$. The translation to ρ -form via (6.8) and (6.9) is mechanical: $u^4 = \rho_1^4$, $uu_x^2 = \rho_1 \rho_2^2 - 2\rho_1^3 \rho_2 + \rho_1^5$, and $u_{xx}^2 = \rho_3^2 - 6\rho_1 \rho_2 \rho_3 + 4\rho_1^3 \rho_3 + 9\rho_1^2 \rho_2^2 - 12\rho_1^4 \rho_2 + 4\rho_1^6$. Substituting and collecting yields a finite combination of integrated ρ -monomials at subscript sum ≤ 6 .

Level 7. Same story as level 5. No new direction.

Level 8. The first new direction beyond $I_1, \dots, I_4$ appears here. The kernel computation, on the ansatz $u^5 + a_1 u^2 u_x^2 + a_2 uu_{xx}^2 + a_3 u_{xxx}^2$ with the requirement that its time derivative reduces to zero, returns the unique solution $(a_1, a_2, a_3) = (-5, 1, -1/14)$. After clearing denominators,

$$I_5 = \int (14u^5 - 70u^2 u_x^2 + 14uu_{xx}^2 - u_{xxx}^2) dx. \quad (\text{A.3})$$

The pattern. At each subscript-sum level, a finite linear system over $\mathbb{Q}$ either extends the kernel by one direction or returns nothing. The new directions, in order, are I_1, I_2, I_3, I_4, I_5 , appearing at levels 1, 2, 4, 6, 8. The cross-product structure (the coefficient $-\frac{1}{2}$ in I_3 , the coefficients -2 and $\frac{1}{5}$ in I_4) is the unique direction in the level- N candidate space whose time derivative vanishes modulo D_x , identified by row-reducing a matrix of $\mathbb{Q}$ -coefficients.

A.2 The Butler–Hay Test for the First-Order KdV Pair

The Butler–Hay test [12] gives two diagnostics for fake Lax pairs. The gauge test seeks a transformation $\varphi \mapsto S\varphi$ conjugating L to constant coefficients in u ; if one exists, the pair is fake, since the gauged spatial problem could belong to any PDE on the same substrate. The excess-freedom (u -fake) test replaces each u -dependent coefficient by an arbitrary function and imposes compatibility; leftover freedom signals a fake, whereas a genuine substrate fixes P rigidly. We run both on (6.6).

The gauge test is immediate. The presentation $L = e^{-U} \partial_x e^U$ of (6.7), read backwards with $S = e^U$, gives

$$SLS^{-1} = e^U (u + \partial_x) e^{-U} = \partial_x,$$

constant-coefficient and free of u . The gauged eigenvalue problem $\partial_x \tilde{\varphi} = \lambda \tilde{\varphi}$ has solutions $\tilde{\varphi} = e^{\lambda x}$ carrying no information about u , and would govern any PDE on the substrate $(u + \partial_x, q)$. The pair is fake.

The excess-freedom test agrees. Since $L = u + \partial_x$ gives $[L, P] = P_x$ for any multiplication operator P , compatibility reads $u_t = P_x$ and holds for every conservation-form PDE $u_t = q_x$ with $P = q$. The

flux is unconstrained, and KdV with $q = 3u^2 + u_{xx}$ is one member of an infinite family sharing the substrate. This is the conservation-law dressing of Calogero and Nucci [13], whose fake pair carries arbitrary f, g with compatibility $f_t = g_x$; the first-order pair reported in [3] is one realization, and Sakovich [60] recast the same fact as a gauge invariant via the cyclic basis.

By contrast, the classical pair $L_{\text{Schr}} = -u - \partial_x^2$, $P_{\text{Schr}} = -4\partial_x^3 - 6u\partial_x - 3u_x$ passes both tests substantively: conjugating a second-order operator by $S = e^\phi$ produces a first-order term $-2\phi_x\partial_x$, so removing it forces $\phi_x = 0$ and leaves u intact, and $\partial_t L_{\text{Schr}} = [L_{\text{Schr}}, P_{\text{Schr}}]$ fixes P_{Schr} up to gauge with no alternative flux. The first-order pair is fake by every gauge-invariant measure here, in agreement with the Krichever collapse of Appendix A.3 and the spectral findings of Section 6.2, yet Section 6.3 recovers the full hierarchy from it.

A.3 The Krichever Test for the First-Order KdV Pair

The Krichever test [38, 39] is a necessary spectral condition for algebro-geometric content. Linearizing $u_t = K[u]$ and its auxiliary system $L\varphi = \lambda\varphi$, $\varphi_t = -P\varphi$ at a constant background $u = u_0$ yields two dispersion laws, $\omega_{\text{PDE}}(\xi)$ from the plane wave $\delta u = e^{i(\xi x - \omega t)}$ and $\omega_{\text{Lax}}(\xi)$ from the temporal frequency of eigenfunctions of L ; a genuine pair requires ω_{Lax} polynomial on the spectral curve and equal to ω_{PDE} up to normalization.

For KdV, $\delta u_t = 6u_0\delta u_x + \delta u_{xxx}$ gives

$$\omega_{\text{KdV}}(\xi) = \xi^3 - 6u_0\xi. \quad (\text{A.4})$$

For the first-order pair, $L\varphi = \mu\varphi$ at u_0 gives $\varphi = e^{(\mu - u_0)x}$, so $\mu = u_0 + i\xi$ and $\varphi = e^{i\xi x}$ on the linear curve $\mu - u_0 = i\xi$. There $u_{xx} = 0$ and $P = 3u^2 + u_{xx}$ collapses to the constant $3u_0^2$, so $\varphi_t = -3u_0^2\varphi$ gives

$$\omega_{\text{Lax}}(\xi) = -3iu_0^2, \quad (\text{A.5})$$

constant in ξ (the imaginary value reflects the non-self-adjointness of L). Against the cubic (A.4) this is polynomial only degenerately, as a degree-zero constant, and recovers none of the dispersive ξ^3 structure of KdV for any u_0 .

The classical pair passes substantively. At u_0 the eigenfunctions of $L_{\text{Schr}} = -\partial_x^2 - u$ are $\varphi = e^{i\xi x}$ with $\lambda = \xi^2 - u_0$, and $P_{\text{Schr}}\varphi = i(4\xi^3 - 6u_0\xi)\varphi$ gives $\omega_{\text{Lax}}^{\text{Schr}}(\xi) = 4\xi^3 - 6u_0\xi$, cubic on the curve and matching (A.4) up to the cubic normalization. The first-order pair carries none of this dispersive content.