

This is a repository copy of *Classical Yang-Baxter equation, Lagrangian multiforms and ultralocal integrable hierarchies*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/204743/>

Version: Published Version

Article:

Caudrelier, Vincent, Stoppato, Matteo and Vicedo, Benoit orcid.org/0000-0003-3003-559X (2024) Classical Yang-Baxter equation, Lagrangian multiforms and ultralocal integrable hierarchies. *Communications in Mathematical Physics*. 12. ISSN 1432-0916

<https://doi.org/10.1007/s00220-023-04871-x>

Reuse

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:

<https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



Classical Yang–Baxter Equation, Lagrangian Multiforms and Ultralocal Integrable Hierarchies

Vincent Caudrelier¹ , Matteo Stoppato¹, Benoît Vicedo²

¹ School of Mathematics, University of Leeds, Leeds LS2 9JT, UK.

E-mail: v.caudrelier@leeds.ac.uk; stoppato.matteo@gmail.com

² Department of Mathematics, University of York, York YO10 5DD, UK.

E-mail: benoit.vicedo@gmail.com

Received: 9 February 2022 / Accepted: 29 October 2023

© The Author(s) 2024

Abstract: We cast the classical Yang–Baxter equation (CYBE) in a variational context for the first time, by relating it to the theory of Lagrangian multiforms, a framework designed to capture integrability in a variational fashion. This provides a significant connection between Lagrangian multiforms and the CYBE, one of the most fundamental concepts of integrable systems. This is achieved by introducing a *generating* Lagrangian multiform which depends on a skew-symmetric classical r -matrix with spectral parameters. The multiform Euler–Lagrange equations produce a generating Lax equation which yields a generating zero curvature equation. The CYBE plays a role at three levels: (1) it ensures the commutativity of the flows of the generating Lax equation; (2) it ensures that the generating zero curvature equation holds; (3) it implies the closure relation for the generating Lagrangian multiform. The specification of an integrable hierarchy is achieved by fixing certain data: a finite set $S \subset \mathbb{CP}^1$, a Lie algebra $\mathfrak{g}$, a $\mathfrak{g}$ -valued rational function with poles in S and an r -matrix. We show how our framework is able to generate a large class of *ultralocal* integrable hierarchies by providing several known and new examples pertaining to the rational or trigonometric class. These include the Ablowitz–Kaup–Newell–Segur hierarchy, the sine-Gordon (sG) hierarchy and various hierarchies related to Zakharov–Mikhailov type models which contain the Faddeev–Reshetikhin (FR) model and recently introduced deformed Gross–Neveu models as particular cases. The versatility of our method is illustrated by showing how to couple integrable hierarchies together to create new examples of integrable field theories and their hierarchies. We provide two examples: the coupling of the nonlinear Schrödinger system to the FR model and the coupling of sG with the anisotropic FR model.

Contents

1. Introduction	
1.1 Context and background	
1.1.1 Integrability in the Hamiltonian framework	

1.1.2	Integrability in the Lagrangian framework	...
1.1.3	Motivating example: Ablowitz–Kaup–Newell–Segur hierarchy	...
1.2	Motivation, main results and plan	...
2.	Lie Algebra of $\mathfrak{g}$ -valued Adèles	...
2.1	General setup	...
2.2	Rational r -matrix	...
2.3	Trigonometric r -matrix	...
3.	Generating Lagrangian Multiform and CYBE	...
3.1	Dynamical equations	...
3.1.1	Adjoint orbit	...
3.1.2	Generating Lax equation	...
3.1.3	Generating zero curvature equation	...
3.1.4	Connection to integrable field theory and FNR procedure	...
3.2	Generating Lagrangian multiform	...
3.2.1	Extracting Lagrangians and Lagrangian multiforms	...
3.2.2	Generating multiform Euler–Lagrange equations	...
3.2.3	Generating closure relation	...
4.	AKNS Hierarchy	...
5.	Sine-Gordon Hierarchy	...
6.	Hierarchies of Zakharov–Mikhailov Type	...
6.1	Rational Zakharov–Mikhailov models	...
6.1.1	Case of simple poles	...
6.1.2	Case of higher poles	...
6.1.3	Interplay between hierarchies associated to simple and higher order poles	...
6.2	Trigonometric Zakharov–Mikhailov models	...
6.3	Deformed Gross–Neveu models	...
7.	Coupling Integrable Hierarchies Together	...
7.1	AKNS-FR hierarchy	...
7.2	sG-aFR hierarchy	...
8.	Discussion and Conclusion	...
A.	Comparison of Trigonometric r -matrices for Sine-Gordon	...

1. Introduction

1.1. Context and background.

1.1.1. Integrability in the Hamiltonian framework A profound discovery in the modern theory of integrable systems was that the special partial differential equations originally treated in the seminal works [GGKM,ZS], using what is now known as the Inverse Scattering Method, were also infinite dimensional Hamiltonian systems [G] for which an analog of the Liouville theorem for finite dimensional Hamiltonian systems could be established [ZF,ZMan]. This allows one, in particular, to see such systems as Hamiltonian field theories. The developments based on these early examples led to the beautiful theory of the classical r -matrix which captures the special Hamiltonian features of these models [Dr1,STS]. An important condition usually required of the r -matrix is that it satisfies the classical Yang–Baxter equation (CYBE)

$$[r_{12}(\lambda, \mu), r_{13}(\lambda, \nu)] + [r_{12}(\lambda, \mu), r_{23}(\mu, \nu)] - [r_{13}(\lambda, \nu), r_{32}(\nu, \mu)] = 0. \quad (1.1)$$

It ensures that a certain Poisson bracket defined using r satisfies the Jacobi identity. Another important condition is to decide if r is skew-symmetric or not, i.e. whether or not it satisfies

$$r_{12}(\lambda, \mu) = -r_{21}(\mu, \lambda).$$

This has deep mathematical and physical implications. If the r -matrix is skew-symmetric, the associated field theories are called ultralocal while they are non-ultralocal otherwise. In the present work, we restrict our attention to the ultralocal case.

A characteristic feature of integrable field theories is that their equations of motion come in hierarchies. Specifically, any given integrable Hamiltonian field theory has infinitely many conserved charges which can, themselves, be used as Hamiltonians to define flows with respect to the Poisson bracket. Because all the conserved charges Poisson commute amongst themselves, it is possible to impose all these flows simultaneously on the fields of the theory and thus treat the latter as depending on infinitely many times. The collection of equations of motion thus obtained is referred to as an integrable hierarchy. Schematically, for a scalar field theory with field u , there would be a countable number of conserved charges H_j , labelled by integers $j \geq 1$ say, in involution with respect to a given Poisson bracket, namely

$$\{H_i, H_j\} = 0$$

for every $i, j \geq 1$. The hierarchy would then consist of all the equations

$$\partial_{t_j} u = \{H_j, u\}, \quad (1.2)$$

where we have introduced an infinite number of times t_j for $j \geq 1$. Among all the conserved charges H_j , one of them can be taken to be the Hamiltonian of the integrable field theory one started with. Studying the hierarchy as a whole can reveal much more structure and properties of the initial model. This is of course not a new idea but here we depart from the established point of view in that we want to exploit this idea in a *Lagrangian* setting.

1.1.2. Integrability in the Lagrangian framework When turning to the Lagrangian setting, one is immediately faced with the following question: how should integrable hierarchies be captured in the Lagrangian formalism? This question found an answer relatively recently in the theory of *Lagrangian multiforms* which was introduced in the seminal paper [LN] and has rapidly developed in various direction. More recently, several works cast the original idea into the context of continuous integrable field theories, see [SV, V, SNC, PV, SNC2, CS1, CS2, CS3] for examples of two-dimensional field theories (e.g. Korteweg–de Vries, sine-Gordon and nonlinear Schrödinger) and [SNC2, SNC3] for a three-dimensional example (Kadomtsev–Petviashvili). For a two-dimensional field theory, the central object is a differential two-form

$$\mathcal{L}[u] = \sum_{i,j} \mathcal{L}_{ij}[u] dt_i \wedge dt_j \quad (1.3)$$

on an infinite-dimensional space $\mathbb{R}^\infty$ parametrised by the infinite collection of times t_i of the hierarchy. The coefficients $\mathcal{L}_{ij}[u]$ are Lagrangians depending on the fields of the theory, which are collectively denoted by u here for simplicity (even though we are no longer restricting to the case of a single scalar field). For each Lagrangian

coefficient $\mathcal{L}_{ij}[u]$ we can consider the associated action $S_{ij}[u] = \int_{\mathbb{R}^2} \mathcal{L}_{ij}[u] dt_i \wedge dt_j$. Using the differential two-form (1.3) we can succinctly rewrite all these actions as $S_{ij}[u] = \int_{\sigma_{ij}} \mathcal{L}[u]$, where the integral here is over the two dimensional plane $\sigma_{ij} \simeq \mathbb{R}^2$ spanned by the coordinates t_i and t_j in $\mathbb{R}^\infty$. At this point, of course, there is no reason for the field theories described by the actions $S_{ij}[u]$ to belong to the same integrable hierarchy, let alone to produce equations of motion that are integrable! The key new ingredient is to impose a generalised variational principle on the more general action

$$S[u, \sigma] = \int_{\sigma} \mathcal{L}[u], \quad (1.4)$$

which now also depends on an arbitrary choice of two-dimensional smooth surface σ in $\mathbb{R}^\infty$. Note, in particular, that $S_{ij}[u] = S[u, \sigma_{ij}]$. The generalised variational principle which ties all these theories together is a least action principle for $S[u, \sigma]$ **simultaneously for all smooth surfaces** σ . This results in what are called the **multiform Euler–Lagrange (EL) equations**. These were first derived in [SV] for the two-form case that we consider in this paper. It can be shown [SV, SNC] that they can be written compactly as

$$\delta d\mathcal{L} = 0, \quad (1.5)$$

where d is the usual exterior derivative and δ denotes the variational derivative. In the Lagrangian multiform theory, the above generalised variational principle is complemented by another requirement: on critical points, one also requires that the action be stationary with respect to arbitrary local variations of σ . This gives us the important **closure relation** on the equations of motion, i.e. *on shell*

$$d\mathcal{L} = 0. \quad (1.6)$$

Intuitively, requiring criticality of the action for an arbitrary surface is the new feature that encodes variationally the commutativity of the flows known to be a signature of integrability in the Hamiltonian world. Roughly speaking, the connection with (1.2) is that the Lagrangian coefficients $\mathcal{L}_{1j}$ correspond by a Legendre transform to the Hamiltonians H_j , with the understanding that the time t_1 plays some preferred role (the “space” variable) and the t_j , $j \geq 2$ are all the higher times of the hierarchy. The interpretation of all the other Lagrangian coefficients $\mathcal{L}_{ij}$ is best obtained by casting the hierarchy as a collection of compatible zero curvature equations involving Lax matrices $V_j(\lambda)$, namely

$$\partial_{t_j} V_i(\lambda) - \partial_{t_i} V_j(\lambda) + [V_i(\lambda), V_j(\lambda)] = 0 \quad (1.7)$$

for $i, j \geq 1$. It is known that all these equations are in fact Hamiltonian, see e.g. [AC], and the case $i = 1$ corresponds to (1.2). One of the main points of the present work is that they are also variational with Lagrangian $\mathcal{L}_{ij}$. It is important to realise that the multiform EL equations are largely overdetermined equations for the coefficients $\mathcal{L}_{ij}$. Part of these equations impose restrictions on the allowed coefficients, the idea being that they enforce the integrability of the corresponding theories. The rest consist of standard EL equations associated to these coefficients and give the equations of motion of the integrable hierarchy.

1.1.3. Motivating example: Ablowitz–Kaup–Newell–Segur hierarchy In [CS3], on the example of the Ablowitz–Kaup–Newell–Segur (AKNS) hierarchy, the notion of Lagrangian multiform was successfully combined with the idea of “compounding hierarchies” introduced in the Lagrangian framework in [N1] (itself inspired from the use of the generating formalism for integrable hierarchies, see e.g. [N2]). This naturally leads to working with generating functions when dealing with hierarchies. The key object was what we can call a *generating Lagrangian multiform*. The simple idea is to organise the Lagrangian coefficients $\mathcal{L}_{ij}$ of the 2-form (1.3) into a generating series involving formal (spectral) parameters

$$\mathcal{L}(\lambda, \mu) = \sum_{i,j} \frac{\mathcal{L}_{ij}}{\lambda^{i+1} \mu^{j+1}}. \quad (1.8)$$

It is clear that there is a one-to-one corresponding between $\mathcal{L}[u]$ and $\mathcal{L}(\lambda, \mu)$ where from the latter, one can extract the coefficients by the formula

$$\mathcal{L}_{ij} = \text{res}_\lambda \text{res}_\mu \left(\lambda^i \mu^j \mathcal{L}(\lambda, \mu) \right),$$

where res_λ returns the coefficient of λ^{-1} in the series expansion, and similarly for res_μ . One advantage of working with generating series such as (1.8) stems from the usefulness of generating functions in general: properties of their coefficients are more easily studied and derived from those of the generating function. In our context, this means that we can manipulate an integrable hierarchy as a whole instead of studying each Lagrangian coefficient $\mathcal{L}_{ij}$ individually. Originally, the latter approach was used in the sense that only a given “starting” Lagrangian coefficient was known, say $\mathcal{L}_{12}$, and one would try to build the higher coefficients $\mathcal{L}_{ij}$ so as to obtain a consistent Lagrangian multiform. Methods to compute these coefficients were introduced for instance in [V,SNC2]. Although the recursive algorithm could be applied in principle, in practice this is hard to implement beyond the first few coefficients. Moreover, the Lagrangians $\mathcal{L}_{ij}$ obtained in this way usually contain derivatives with respect to t_1 or t_2 (the times associated with $\mathcal{L}_{12}$). These are not natural times from the point of $\mathcal{L}_{ij}$: this is the so-called problem of “alien derivatives” which was identified and explained in [V]. Having a generating Lagrangian multiform circumvents these issues. This will be elaborated upon in the examples.

For the AKNS hierarchy, the generating Lagrangian multiform can be written as [CS3]

$$\mathcal{L}(\lambda, \mu) = i \text{Tr} \left(\phi(\mu)^{-1} \mathcal{D}_\lambda \phi(\mu) \sigma_3 - \phi(\lambda)^{-1} \mathcal{D}_\mu \phi(\lambda) \sigma_3 \right) - \text{Tr} \frac{Q(\lambda) Q(\mu)}{\mu - \lambda}, \quad (1.9)$$

with $Q(\lambda) = -i \phi(\lambda) \sigma_3 \phi(\lambda)^{-1}$, $\phi(\lambda)$ being a group-valued formal series in $1/\lambda$ with constant term equal to the identity and whose coefficients contain the dynamical variables. The operator $\mathcal{D}_\lambda = \sum_{j \geq 0} \lambda^{-j-1} \partial_{t_j}$ is a formal collection of all the AKNS flows ∂_{t_j} ,

and similarly for $\mathcal{D}_\mu$. The generating Lagrangian multiform (1.9) generates all the coefficients $\mathcal{L}_{ij}$ systematically and without the problem of alien derivatives, reproducing the first few coefficients which had been constructed in [SNC,SNC2,PV], as it should. Its

multiform EL equations yield the defining equations of the AKNS hierarchy as discussed by Flaschka–Newell–Ratiu (FNR) in [FNR], namely¹

$$\partial_{t_i} Q(\lambda) = [Q^{(i)}(\lambda), Q(\lambda)], \quad i \geq 0, \quad (1.10)$$

where $Q(\lambda) = \sum_{j=0}^{\infty} Q_j \lambda^{-j}$ and $Q^{(i)}(\lambda) = \sum_{j=0}^i Q_j \lambda^{i-j}$ and $Q_0 = -i\sigma_3$. More precisely, the multiform EL equations for (1.9) produce the equations (1.10) in generating form

$$\mathcal{D}_\mu Q(\lambda) = \frac{[Q(\mu), Q(\lambda)]}{\mu - \lambda}, \quad (1.11)$$

where we used the formal series identity

$$\sum_{k=0}^{\infty} \frac{Q^{(k)}(\lambda)}{\mu^{k+1}} = \frac{Q(\mu)}{\mu - \lambda}. \quad (1.12)$$

1.2. Motivation, main results and plan. **Motivation** The present work is motivated by the following observations made on the generating Lagrangian multiform (1.9) and the generating FNR equations (1.11):

1. The potential term in $\mathcal{L}(\lambda, \mu)$ has a characteristic form which can be identified as the expression

$$\mathrm{Tr}_{12} (r_{12}(\lambda, \mu) Q_1(\lambda) Q_2(\mu))$$

where $r_{12}(\lambda, \mu) = \frac{P_{12}}{\mu - \lambda}$ is the rational r -matrix, known to describe the Hamiltonian structure of the AKNS hierarchy. One could then imagine replacing this particular r -matrix with another skew-symmetric r -matrix. This leads to the question of whether the nice properties of the generating Lagrangian multiform still hold. One of our main results is that this is the case by virtue of the CYBE. Correspondingly, the RHS of (1.11) can also be written as $[\mathrm{Tr}_2 r_{12}(\lambda, \mu) Q_2(\mu), Q_1(\lambda)]$ and the same generalisation can be contemplated.

2. The choice of expanding all the objects as formal series in $1/\lambda$ and $1/\mu$ is a sign that one is performing an expansion around the point at infinity. However, nothing would prevent us from considering other points in $\mathbb{CP}^1$.
3. The Pauli matrix σ_3 appearing in (1.9) is a special choice of constant element in the underlying loop algebra of $\mathfrak{sl}_2$ and the form of $Q(\lambda)$ indicates that one is building a phase space for the field theory as a (co)adjoint orbit around this particular element. One could consider other elements in the loop algebra to construct different phase spaces and hence different models. Moreover, one could also consider other Lie algebras than $\mathfrak{sl}_2$.

The careful implementation of these natural observations requires some machinery which is presented Sect. 2. In a first instance, the reader may choose to read the rest of this introduction containing a summary of the formalism and results, and go directly to Sect. 3.

¹ The flow t_0 is the trivial linear flow but is included in the construction for convenience. In practice, one is interested in the nonlinear flows t_j , $j \geq 1$.

The idea is to substitute the loop algebra of $\mathfrak{sl}_2$ with a much more versatile structure: the Lie algebra of $\mathfrak{g}$ -valued adèles associated with a Lie algebra $\mathfrak{g}$. This algebra is presented in [STS2] as the relevant structure to implement the second observation above. By doing so in our context, we build a “universal” generating Lagrangian multiform which is capable of describing a large class of ultralocal integrable hierarchies and we provide a large variety of examples.

In a nutshell, for a matrix Lie algebra $\mathfrak{g}$, the Lie algebra of $\mathfrak{g}$ -valued adèles is defined as

$$\mathcal{A}_\lambda(\mathfrak{g}) := \coprod_{a \in \mathbb{C}P^1} \mathfrak{g} \otimes \mathbb{C}((\lambda_a)),$$

where $\lambda_a = \lambda - a$ for $a \in \mathbb{C}$ and $\lambda_\infty = \frac{1}{\lambda}$ are the local series expansion parameters. An element $X(\lambda) = (X^a(\lambda_a))_{a \in \mathbb{C}P^1}$ of this algebra consist of tuples with all but finitely many of the formal Laurent series $X^a(\lambda_a) \in \mathfrak{g} \otimes \mathbb{C}((\lambda_a))$ being Taylor series in λ_a , i.e. there exists a finite subset $S \subset \mathbb{C}P^1$ such that $X^a(\lambda_a) \in \mathfrak{g} \otimes \mathbb{C}[[\lambda_a]]$ for every $a \in \mathbb{C} \setminus S$. Let $R_\lambda(\mathfrak{g})$ denote the Lie algebra of $\mathfrak{g}$ -valued rational functions in the formal variable λ and define the map

$$\iota_\lambda : R_\lambda(\mathfrak{g}) \longrightarrow \mathcal{A}_\lambda(\mathfrak{g}), \quad f \longmapsto (\iota_{\lambda_a} f)_{a \in \mathbb{C}P^1} \quad (1.13)$$

where $\iota_{\lambda_a} f \in \mathfrak{g} \otimes \mathbb{C}((\lambda_a))$ is the Laurent expansion of $f \in R_\lambda(\mathfrak{g})$ at $a \in \mathbb{C}P^1$. Using certain solutions of the CYBE, it is possible to obtain a direct sum decomposition of this Lie algebra into two maximally isotropic Lie subalgebras

$$\mathcal{A}_\lambda(\mathfrak{g}) = \mathcal{A}_\lambda^+(\mathfrak{g}) \dot{+} \iota_\lambda R_\lambda(\mathfrak{g}). \quad (1.14)$$

We can also define a group $\mathcal{A}_\lambda^+(G)$ associated to $\mathcal{A}_\lambda^+(\mathfrak{g})$. If μ is another formal variable, we can work with double formal series locally in λ_a and μ_b , $a, b \in \mathbb{C}P^1$ and consider tuples of the form $X(\lambda, \mu) = (X^{a,b}(\lambda_a, \mu_b))_{a,b \in \mathbb{C}P^1}$.

Thanks to this adèlic framework, we can retain the power of the algebraic formulation of formal power series while working locally around arbitrary points in $\mathbb{C}P^1$. We introduce the following generalisation of (1.9) which realises the above three observations

$$\mathcal{L}(\lambda, \mu) := K(\lambda, \mu) - U(\lambda, \mu) \quad (1.15)$$

where the kinetic and potential terms are given by

$$K(\lambda, \mu) := \text{Tr}(\phi(\lambda)^{-1} \mathcal{D}_\mu \phi(\lambda) (\iota_\lambda F(\lambda))_-) - \text{Tr}(\phi(\mu)^{-1} \mathcal{D}_\lambda \phi(\mu) (\iota_\mu F(\mu))_-), \quad (1.16a)$$

$$U(\lambda, \mu) := \frac{1}{2} \text{Tr}_{12}((\iota_\lambda \iota_\mu + \iota_\mu \iota_\lambda) r_{12}(\lambda, \mu) \mathbf{Q}_1(\lambda) \mathbf{Q}_2(\mu)). \quad (1.16b)$$

Here $\phi(\lambda)$ is an element of the group $\mathcal{A}_\lambda^+(G)$, $\mathcal{Q}(\lambda) = \phi(\lambda)(\iota_\lambda F(\lambda))_- \phi(\lambda)^{-1}$ is an element of $\mathcal{A}_\lambda(\mathfrak{g})$, where $(\iota_\lambda F(\lambda))_- = (F^a(\lambda_a)_-)_{a \in \mathbb{C}P^1}$ is the collection of principal parts of a $\mathfrak{g}$ -valued rational function $F(\lambda) \in R_\lambda(\mathfrak{g})$. In terms of components of the tuples, we have

$$\begin{aligned} \mathcal{L}^{a,b}(\lambda_a, \mu_b) &= \text{Tr}(\phi^a(\lambda_a)^{-1} \mathcal{D}_{\mu_b} \phi^a(\lambda_a) F^a(\lambda_a)_-) - \text{Tr}(\phi^b(\mu_b)^{-1} \mathcal{D}_{\lambda_a} \phi^b(\mu_b) F^b(\mu_b)_-) \\ &\quad - \frac{1}{2} \text{Tr}_{12}((\iota_{\lambda_a} \iota_{\mu_b} + \iota_{\mu_b} \iota_{\lambda_a}) r_{12}(\lambda, \mu) \mathcal{Q}_1^a(\lambda_a) \mathcal{Q}_2^b(\mu_b)), \end{aligned}$$

for every $a, b \in \mathbb{C}P^1$. The operator $\mathcal{D}_\lambda := (\mathcal{D}_{\lambda_a})_{a \in \mathbb{C}P^1}$ denotes the $\mathbb{C}P^1$ -tuple of formal operators $\mathcal{D}_{\lambda_a}$ which contain the partial differential operators $\partial_{t_n^a}$ (see (3.6)). The times t_n^a will be the times of the integrable hierarchies we describe. The rational function $r_{12}(\lambda, \mu)$ is the classical r -matrix defining the type of ultralocal hierarchies we consider (e.g. rational or trigonometric) and corresponds to the r -matrix yielding the decomposition (1.14).

Main results

1. We show that the *generating Lax equation*

$$\mathcal{D}_\mu \mathcal{Q}_1(\lambda) = [\text{Tr}_2 (\iota_\lambda \iota_\mu r_{12}(\lambda, \mu) \mathcal{Q}_2(\mu)), \mathcal{Q}_1(\lambda)] \quad (1.17)$$

is variational. It arises as the multiform EL equations associated to our generating Lagrangian multiform (1.15). This is the content of Theorem 3.12. This generalises the analogous result first obtained in [SNC] in the context of the Zakharov–Mikhailov models [ZM1]. The generating Lax equation plays here for field theories a role similar to the traditional Lax equation for finite dimensional systems. This is explained in Sect. 3.1. We relate it to a *generating zero curvature equation* which is shown to hold as a consequence of the CYBE for the r -matrix appearing in (1.15).

2. We relate for the first time the CYBE with the relatively recent notion of Lagrangian multiforms. The closure relation (1.6) in generating form, i.e. the closure relation for (1.15), is shown to be a consequence of the CYBE for the r -matrix appearing in (1.15), see Theorem 3.13. On the one hand, this provides a variational interpretation of the CYBE, a fundamental equation that has only been introduced and studied from a Hamiltonian point of view so far. On the other hand, given the importance of the CYBE as a criterion for classical integrability, this further establishes the Lagrangian multiform approach as a variational criterion for integrability.
3. Specialising the generating Lagrangian multiform (1.15), we recover known examples of integrable hierarchies and produce several new examples. We also introduce an easy method for coupling hierarchies together.

Plan of the paper

In Sect. 2, we introduce the Lie algebra of $\mathfrak{g}$ -valued adèles and establish its decomposition into two complementary maximal isotropic Lie subalgebras which allows us to introduce the classical r -matrix of interest via the corresponding projectors onto the Lie subalgebras. This generalises to the adèles case the well-known interpretation of a classical r -matrix as a difference of projectors. This is done explicitly for the rational and trigonometric cases. Section 3 introduces the main elements of our framework: we state the generalisation of the generating FNR equations (1.11), which we call the generating Lax equation, taking into account the above observations. Its properties are directly connected to the CYBE. Then we introduce the generating Lagrangian multiform that produces the generating Lax equation as its multiform EL equations. Again, its properties, in particular the closure relation, are shown to be a direct consequence of the CYBE. The subsequent Sects. 4–6 are devoted to examples. Several were known previously, and these are used to show how our framework contains them naturally, e.g. the AKNS hierarchy and the sine-Gordon hierarchy. For the latter example, we explain in detail how the first few known Lagrangian coefficients are recovered but without the problem of alien derivatives. Other examples, such as the trigonometric Zakharov–Mikhailov hierarchy, are new. For the recently introduced deformed Gross-Neveu models, the new feature brought in by our construction is that they are naturally embedded into an integrable hierarchy. Various conclusions and discussions are presented in Sect. 8. Finally, we recall

in an “Appendix” the relationship between the trigonometric r -matrix used in this paper and the more familiar r -matrix of the sine-Gordon model.

2. Lie Algebra of $\mathfrak{g}$ -valued Adèles

2.1. General setup. Let $N \in \mathbb{Z}_{\geq 1}$ and consider either the Lie algebra $\mathfrak{gl}_N$ of all $N \times N$ matrices with complex entries or its Lie subalgebra $\mathfrak{sl}_N$ of traceless matrices. We will treat both of these cases in parallel, using the common notation $\mathfrak{g}$ throughout. The generalisation to other matrix Lie algebras is straightforward but for simplicity we shall restrict to these two cases. We also denote by G the associated Lie group which corresponds either to the general linear group GL_N of invertible $N \times N$ matrices or to its Lie subgroup SL_N of matrices with determinant 1.

We use the trace $\text{Tr} : \mathfrak{gl}_N \rightarrow \mathbb{C}$ to endow the Lie algebra $\mathfrak{g}$ with the non-degenerate invariant symmetric bilinear form $\mathfrak{g} \times \mathfrak{g} \rightarrow \mathbb{C}$ given by $(X, Y) \mapsto \text{Tr}(XY)$. Let P_{12} be the tensor Casimir of $\mathfrak{g}$ with the property that $\text{Tr}_2(P_{12}X_2) = X$ for any $X \in \mathfrak{g}$. Explicitly, for $\mathfrak{gl}_N$ it is given by $P_{12} = \sum_{i,j=1}^N E_{ij} \otimes E_{ji}$ where E_{ij} for $i, j = 1, \dots, N$ is the standard basis of $\mathfrak{gl}_N$. Similarly, for $\mathfrak{sl}_N$ we can write $P_{12} = \sum_a I_a \otimes I^a$ where $\{I_a\}$ and $\{I^a\}$ are dual bases of $\mathfrak{sl}_N$ with respect to the above bilinear form. For clarity, let us also recall that the notation X_2 means $\mathbf{1} \otimes X$ and the notation $\text{Tr}_2(\dots)$ means that we apply the trace only in the second tensor factor.

Let λ be a formal variable. For any $a \in \mathbb{C}$ we define the formal local coordinate around a as $\lambda_a := \lambda - a$ and to the point at infinity we associate the formal local coordinate $\lambda_\infty := \lambda^{-1}$. We consider the Lie algebra of $\mathfrak{g}$ -valued adèles defined as

$$\mathcal{A}_\lambda(\mathfrak{g}) := \coprod_{a \in \mathbb{C}P^1} \mathfrak{g} \otimes \mathbb{C}((\lambda_a)).$$

Its elements consist of tuples $X(\lambda) = (X^a(\lambda_a))_{a \in \mathbb{C}P^1}$ with all but finitely many of the formal Laurent series $X^a(\lambda_a) \in \mathfrak{g} \otimes \mathbb{C}((\lambda_a))$ being Taylor series, i.e. there exists a finite subset $S \subset \mathbb{C}P^1$ such that $X^a(\lambda_a) \in \mathfrak{g} \otimes \mathbb{C}[[\lambda_a]]$ for every $a \in \mathbb{C} \setminus S$. The Lie bracket of two elements $X(\lambda) = (X^a(\lambda_a))_{a \in \mathbb{C}P^1}$ and $Y(\lambda) = (Y^a(\lambda_a))_{a \in \mathbb{C}P^1}$ in $\mathcal{A}_\lambda(\mathfrak{g})$ is defined component-wise, as

$$[X(\lambda), Y(\lambda)] = ([X^a(\lambda_a), Y^a(\lambda_a)])_{a \in \mathbb{C}P^1}.$$

Let R_λ denote the algebra of rational functions in the formal variable λ . The Laurent expansion of a rational function $f \in R_\lambda$ at any $a \in \mathbb{C}P^1$ defines a homomorphism

$$\iota_{\lambda_a} : R_\lambda \longrightarrow \mathbb{C}((\lambda_a)), \quad f \longmapsto \iota_{\lambda_a} f. \quad (2.1)$$

We will consider two possible non-degenerate invariant bilinear forms on the Lie algebra $\mathcal{A}_\lambda(\mathfrak{g})$, namely

$$\langle\langle \cdot, \cdot \rangle\rangle_k : \mathcal{A}_\lambda(\mathfrak{g}) \times \mathcal{A}_\lambda(\mathfrak{g}) \longrightarrow \mathbb{C} \quad (2.2a)$$

for $k = 0$ and $k = -1$, defined as

$$\langle\langle X(\lambda), Y(\lambda) \rangle\rangle_k := \sum_{a \in \mathbb{C}P^1} \text{res}_a^\lambda \text{Tr} (X^a(\lambda_a) Y^a(\lambda_a)) \lambda^k d\lambda, \quad (2.2b)$$

for any $X(\lambda) = (X^a(\lambda_a))_{a \in \mathbb{C}P^1}$ and $Y(\lambda) = (Y^a(\lambda_a))_{a \in \mathbb{C}P^1}$. Strictly speaking, the rational function λ^k on the right hand side of (2.2b) should be expanded at $a \in \mathbb{C}P^1$,

namely we should write $\iota_{\lambda_a} \lambda^k$ instead of λ^k . In order to simplify the notation, such expansions will always be implicit when taking residues. Here, for any $a \in \mathbb{C}P^1$, the residue $\text{res}_a^\lambda : \mathbb{C}((\lambda_a)) d\lambda_a \rightarrow \mathbb{C}$ returns the coefficient of $\lambda_a^{-1} d\lambda_a$. For computing the residue at infinity we note that $d\lambda = -\lambda_\infty^{-2} d\lambda_\infty$. Note that only finitely many terms in the sum in (2.2b) are non-zero by definition of $\mathcal{A}_\lambda(\mathfrak{g})$.

Let $R_\lambda(\mathfrak{g}) := \mathfrak{g} \otimes R_\lambda$ denote the Lie algebra of $\mathfrak{g}$ -valued rational functions in the formal variable λ . We have an embedding of Lie algebras

$$\iota_\lambda : R_\lambda(\mathfrak{g}) \longrightarrow \mathcal{A}_\lambda(\mathfrak{g}), \quad f \longmapsto (\iota_{\lambda_a} f)_{a \in \mathbb{C}P^1} \quad (2.3)$$

where $\iota_{\lambda_a} f \in \mathfrak{g} \otimes \mathbb{C}((\lambda_a))$ is the Laurent expansion of $f \in R_\lambda(\mathfrak{g})$ at $a \in \mathbb{C}P^1$ in the second tensor factor, as in (2.1). The Lie subalgebra $\iota_\lambda R_\lambda(\mathfrak{g}) \subset \mathcal{A}_\lambda(\mathfrak{g})$ is maximally isotropic with respect to $\langle \cdot, \cdot \rangle_k$, for any $k \in \mathbb{Z}$, by the strong residue theorem; see for instance [Ta, Corollary 1].

In the remainder of this section we will describe two possible complementary Lie algebras to $\iota_\lambda R_\lambda(\mathfrak{g})$ in $\mathcal{A}_\lambda(\mathfrak{g})$, which are maximally isotropic with respect to $\langle \cdot, \cdot \rangle_0$ and $\langle \cdot, \cdot \rangle_{-1}$, respectively. These two main examples, which can be found for instance in [Dr2, Example 4], correspond to the rational r -matrix and the trigonometric r -matrix, respectively.

Notation We will generally use boldface to denote $\mathbb{C}P^1$ -tuples. For instance, given any $n \in \mathbb{Z}$ we will write $\lambda^n X(\lambda)$ for the element $(\lambda_a^n X^a(\lambda_a))_{a \in \mathbb{C}P^1} \in \mathcal{A}_\lambda(\mathfrak{g})$ of the Lie algebra of $\mathfrak{g}$ -valued adèles. More generally, we would write $\lambda^n X(\lambda) d\lambda$ as a shorthand for the $\mathbb{C}P^1$ -tuple $(\lambda_a^n X^a(\lambda_a) d\lambda_a)_{a \in \mathbb{C}P^1}$. Note, crucially, that although $d\lambda_a = d\lambda$ for all $a \in \mathbb{C}$, we have $d\lambda_\infty = -\lambda_\infty^{-2} d\lambda$ for the point at infinity. Therefore the two expressions $\lambda^n X(\lambda) d\lambda$ and $\lambda^n X(\lambda) d\lambda$ subtly differ only in the component at infinity. If μ is another formal variable then μ will denote a separate $\mathbb{C}P^1$ -tuple carrying an independent label $b \in \mathbb{C}P^1$. For instance, we would have

$$[X(\lambda), Y(\mu)] = (\delta_{ab} [X^a(\lambda_a), Y^b(\mu_b)])_{a,b \in \mathbb{C}P^1}$$

for any $X(\lambda) \in \mathcal{A}_\lambda(\mathfrak{g})$ and $Y(\mu) \in \mathcal{A}_\mu(\mathfrak{g})$. We will make use of such notation with multiple formal variables extensively from Sect. 3 onwards.

2.2. Rational r -matrix. Throughout this section we fix the choice $k = 0$ in the bilinear form (2.2). Consider the Lie subalgebra of $\mathfrak{g}$ -valued integral adèles

$$\mathcal{A}_\lambda^{\text{rat}}(\mathfrak{g}) := \mathfrak{g} \otimes \lambda_\infty \mathbb{C}[[\lambda_\infty]] \times \coprod_{a \in \mathbb{C}} \mathfrak{g} \otimes \mathbb{C}[[\lambda_a]]. \quad (2.4)$$

Note that we have excluded the constant term from the Taylor series at infinity. We shall also need the corresponding group

$$\mathcal{A}_\lambda^{\text{rat}}(G) := \widehat{G}_\infty \times \coprod_{a \in \mathbb{C}} \widehat{G}_a, \quad (2.5)$$

where in the GL_N case $\widehat{G}_a$ consists of all invertible $N \times N$ matrices with entries in $\mathbb{C}[[\lambda_a]]$ while $\widehat{G}_\infty$ consists of all invertible $N \times N$ matrices with off-diagonal entries in $\lambda_\infty \mathbb{C}[[\lambda_\infty]]$ and diagonal entries in $1 + \lambda_\infty \mathbb{C}[[\lambda_\infty]]$. In the SL_N case the groups $\widehat{G}_a$ for all $a \in \mathbb{C}P^1$ are defined in the same way but with the added condition that the matrices are of determinant 1.

For later practical purposes, it is convenient to collect the following notations in a definition.

Definition 2.1. Let $a \in \mathbb{C}$ and $X^a(\lambda_a) \in \mathfrak{g} \otimes \mathbb{C}((\lambda_a))$ be a Laurent series in λ_a with coefficients in $\mathfrak{g}$. We shall use the notation

$$X^a(\lambda_a)^{\text{rat}} \in \mathfrak{g} \otimes \lambda_a^{-1} \mathbb{C}[\lambda_a^{-1}] \quad (2.6a)$$

to represent the pole part of $X^a(\lambda_a)$. Similarly, for $X^\infty(\lambda_\infty) \in \mathfrak{g} \otimes \mathbb{C}((\lambda_\infty)) = \mathfrak{g} \otimes \mathbb{C}((\lambda^{-1}))$, we denote by

$$X^\infty(\lambda_\infty)^{\text{rat}} \in \mathfrak{g} \otimes \mathbb{C}[\lambda_\infty^{-1}] = \mathfrak{g} \otimes \mathbb{C}[\lambda] \quad (2.6b)$$

the pole part of $X^\infty(\lambda_\infty)$. Note that the constant term in λ_∞ is included around infinity.

The Lie subalgebra $\mathcal{A}_\lambda^{\text{rat}}(\mathfrak{g}) \subset \mathcal{A}_\lambda(\mathfrak{g})$ is clearly maximally isotropic with respect to the bilinear form $\langle\langle \cdot, \cdot \rangle\rangle_0$ defined in (2.2). Here we made use of the fact that the constant term was excluded from the Taylor series at infinity in the definition (2.4). It follows that the Lie algebra $\mathcal{A}_\lambda(\mathfrak{g})$ decomposes as a direct sum of vector spaces

$$\mathcal{A}_\lambda(\mathfrak{g}) = \mathcal{A}_\lambda^{\text{rat}}(\mathfrak{g}) \dot{+} \iota_\lambda R_\lambda(\mathfrak{g}) \quad (2.7)$$

into complementary Lagrangian (i.e. maximal isotropic) Lie subalgebras. Let $\pi_\pm^{\text{rat}}$ denote the projections onto $\mathcal{A}_\lambda^{\text{rat}}(\mathfrak{g})$ and $\iota_\lambda R_\lambda(\mathfrak{g})$, respectively, relative to (2.7).

Definition 2.2. (*Rational r -matrix*). Recall the notation P_{12} for the tensor Casimir of $\mathfrak{g}$ from Sect. 2.1. The rational r -matrix is defined as the following rational function of the formal variables λ and μ :

$$r_{12}^{\text{rat}}(\lambda, \mu) = \frac{P_{12}}{\mu - \lambda}. \quad (2.8)$$

As is well-known, it is skew-symmetric: $r_{12}^{\text{rat}}(\lambda, \mu) = -r_{21}^{\text{rat}}(\mu, \lambda)$. The following result shows that its known connection to projectors associated to the decomposition of a Lie algebra into isotropic Lie subalgebras extends to the present adèles setting.

Proposition 2.3. *For any $X(\lambda) \in \mathcal{A}_\lambda(\mathfrak{g})$, its projections onto the complementary subalgebras $\mathcal{A}_\lambda^{\text{rat}}(\mathfrak{g})$ and $\iota_\lambda R_\lambda(\mathfrak{g})$ relative to the direct sum decomposition (2.7) are given respectively by $\pi_+^{\text{rat}} X(\lambda) = ((\pi_+^{\text{rat}} X)^a(\lambda_a))_{a \in \mathbb{C}P^1}$ and $\pi_-^{\text{rat}} X(\lambda) = ((\pi_-^{\text{rat}} X)^a(\lambda_a))_{a \in \mathbb{C}P^1}$ where*

$$(\pi_+^{\text{rat}} X)^a(\lambda_a) = \sum_{b \in \mathbb{C}P^1} \text{res}_b^\mu \text{Tr}_2 \left(\iota_{\mu_b} \iota_{\lambda_a} r_{12}^{\text{rat}}(\lambda, \mu) X^b(\mu_b)_2 \right) d\mu, \quad (2.9a)$$

$$(\pi_-^{\text{rat}} X)^a(\lambda_a) = - \sum_{b \in \mathbb{C}P^1} \text{res}_b^\mu \text{Tr}_2 \left(\iota_{\lambda_a} \iota_{\mu_b} r_{12}^{\text{rat}}(\lambda, \mu) X^b(\mu_b)_2 \right) d\mu. \quad (2.9b)$$

Proof. Let $X(\lambda) \in \mathcal{A}_\lambda(\mathfrak{g})$. We consider, to begin with, its projection onto $\iota_\lambda R_\lambda(\mathfrak{g})$. The $\mathfrak{g}$ -valued rational function in $R_\lambda(\mathfrak{g})$ constructed out of the pole parts of the collection of Laurent series in $X(\lambda)$ is given by

$$\begin{aligned} \sum_{b \in \mathbb{C}P^1} \text{res}_b^\mu \text{Tr}_2 \left(\iota_{\mu_b} \frac{P_{12}}{\lambda - \mu} X^b(\mu_b)_2 \right) d\mu &= \sum_{b \in \mathbb{C}} \sum_{n=0}^{\infty} \text{res}_b^\mu \frac{\mu_b^n}{\lambda^{n+1}} X^b(\mu_b) d\mu \\ &- \sum_{n=0}^{\infty} \text{res}_\infty^\mu \frac{\lambda^n}{\mu^{n+1}} X^\infty(\mu_\infty) d\mu = \sum_{b \in \mathbb{C}P^1} X^b(\lambda_b)_-^{\text{rat}} \end{aligned}$$

where in the first equality we took the trace and split the term at $b = \infty$ from the rest of the sum over $b \in \mathbb{C}P^1$. The expression (2.9b) is then obtained by taking the Laurent series expansion of this rational function at each $a \in \mathbb{C}P^1$, corresponding to applying the map (2.3).

Consider now the projection of $X(\lambda) \in \mathcal{A}_\lambda(\mathfrak{g})$ onto $\mathcal{A}_\lambda^{\text{rat}}(\mathfrak{g})$. Note that for any $a \in \mathbb{C}$ we have

$$\sum_{b \in \mathbb{C}P^1} \text{res}_b^\mu \text{Tr}_2 \left(\iota_{\mu_b} \iota_{\lambda_a} \frac{P_{12}}{\mu - \lambda} X^b(\mu_b)_2 \right) d\mu = \sum_{n=0}^{\infty} \lambda_a^n \sum_{b \in \mathbb{C}P^1} \text{res}_b^\mu \iota_{\mu_b} \mu_a^{-n-1} X^b(\mu_b) d\mu.$$

If $X(\lambda) \in \iota_\lambda R_\lambda(\mathfrak{g})$, say $X(\lambda) = \iota_\lambda f(\lambda)$ for some $f(\lambda) \in R_\lambda(\mathfrak{g})$, then the above vanishes at each order in the λ_a -expansion by the residue theorem. Indeed, the coefficient of λ_a^n is given by the sum of all the residues of $(\mu - a)^{-n-1} f(\mu) d\mu$. On the other hand, if $X(\lambda) \in \mathcal{A}_\lambda^{\text{rat}}(\mathfrak{g})$ then the only term contributing to the sum over $b \in \mathbb{C}P^1$ is the term for $b = a$ which is equal to $X^a(\lambda_a)$. The same statements hold for $a = \infty$ and hence the result follows. $\square$

Define the linear operator $r^{\text{rat}} := \pi_+^{\text{rat}} - \pi_-^{\text{rat}}$. It follows from Proposition 2.3 that its kernel is given by

$$\left((\iota_{\mu_b} \iota_{\lambda_a} + \iota_{\lambda_a} \iota_{\mu_b}) \frac{P_{12}}{\mu - \lambda} \right)_{a,b \in \mathbb{C}P^1}. \quad (2.10)$$

The kernel of the identity operator $\text{id} = \pi_+^{\text{rat}} + \pi_-^{\text{rat}}$ is similarly given by an expansion of zero (see e.g. [LL, Chap. 2]) since

$$\left((\iota_{\mu_b} \iota_{\lambda_a} - \iota_{\lambda_a} \iota_{\mu_b}) \frac{P_{12}}{\mu - \lambda} d\mu \right)_{a,b \in \mathbb{C}P^1} = (P_{12} \delta_{ab} \delta(\lambda_a, \mu_a) d\mu_a)_{a,b \in \mathbb{C}P^1} \quad (2.11)$$

where we defined

$$\delta(\lambda_a, \mu_a) := \sum_{n \in \mathbb{Z}} \lambda_a^n \mu_a^{-n-1}. \quad (2.12)$$

Lemma 2.4. *Let $X(\mu) = (X^a(\mu_a))_{a \in \mathbb{C}P^1} \in \mathcal{A}_\mu(\mathfrak{g})$ with $X^a(\mu_a) = \sum_{n=-N_a}^{\infty} X_n^a \mu_a^n$ for some $N_a \in \mathbb{Z}$, where $N_a > 0$ for finitely many $a \in \mathbb{C}P^1$. For any $a \in \mathbb{C}$ we have*

$$\iota_{\mu_a} \frac{X^a(\mu_a)}{\mu - \lambda} = - \sum_{r=-N_a}^{\infty} \mu_a^r (\lambda_a^{-r-1} X^a(\lambda_a))_-^{\text{rat}}$$

and at infinity we have

$$\iota_{\mu_\infty} \frac{X^\infty(\mu_\infty)}{\mu - \lambda} = \sum_{r=-N_\infty}^{\infty} \mu_\infty^{r+1} (\lambda_\infty^{-r} X^\infty(\lambda_\infty))_-^{\text{rat}}.$$

Proof. First, let $a \in \mathbb{C}$. Then we have

$$\begin{aligned} \iota_{\mu_a} \frac{X^a(\mu_a)}{\mu - \lambda} &= - \sum_{n=-N_a}^{\infty} X_n^a \mu_a^n \sum_{s=0}^{\infty} \mu_a^s \lambda_a^{-s-1} = - \sum_{n=-N_a}^{\infty} \sum_{r=n}^{\infty} X_n^a \mu_a^r \lambda_a^{n-r-1} \\ &= - \sum_{r=-N_a}^{\infty} \mu_a^r \sum_{n=-N_a}^r X_n^a \lambda_a^{n-r-1} = - \sum_{r=-N_a}^{\infty} \mu_a^r (\lambda_a^{-r-1} X^a(\lambda_a))_{-}^{\text{rat}} \end{aligned}$$

where in the second equality we changed variables from s to $r = s + n$ in the second sum and in the second line we changed the order of the sums.

Consider now the point at infinity. We have

$$\begin{aligned} \iota_{\mu_{\infty}} \frac{X^{\infty}(\mu_{\infty})}{\mu - \lambda} &= \sum_{n=-N_{\infty}}^{\infty} X_n^{\infty} \mu_{\infty}^n \sum_{s=0}^{\infty} \mu_{\infty}^{s+1} \lambda_{\infty}^{-s} = \sum_{n=-N_{\infty}}^{\infty} \sum_{r=n}^{\infty} X_n^{\infty} \mu_{\infty}^{r+1} \lambda_{\infty}^{n-r} \\ &= \sum_{r=-N_{\infty}}^{\infty} \mu_{\infty}^{r+1} \sum_{n=-N_{\infty}}^r X_n^{\infty} \lambda_{\infty}^{n-r} = \sum_{r=-N_{\infty}}^{\infty} \mu_{\infty}^{r+1} (\lambda_{\infty}^{-r} X^{\infty}(\lambda_{\infty}))_{-}^{\text{rat}} \end{aligned}$$

where in the second equality we changed variables $s = r - n$ as before and in the second line we changed the order of the sums. $\square$

2.3. Trigonometric r -matrix. Throughout this section we will choose $k = -1$ in the bilinear form (2.2). We shall also make use of the standard nilpotent subalgebras $\mathfrak{n}_{\pm}$ and Borel subalgebras $\mathfrak{b}_{\pm}$ of $\mathfrak{g}$. Explicitly, $\mathfrak{n}_{+}$ (resp. $\mathfrak{n}_{-}$) is spanned by E_{ij} for $i < j$ (resp. $i > j$). In the $\mathfrak{gl}_N$ case $\mathfrak{b}_{+}$ (resp. $\mathfrak{b}_{-}$) is spanned by E_{ij} for $i \leq j$ (resp. $i \geq j$) while in the $\mathfrak{sl}_N$ case $\mathfrak{b}_{+}$ (resp. $\mathfrak{b}_{-}$) is spanned by E_{ij} for $i < j$ (resp. $i > j$) together with $E_{ii} - E_{jj}$ for $i < j$. The Cartan subalgebra $\mathfrak{h}$ is spanned by E_{ii} for $i = 1, \dots, N$ in the $\mathfrak{gl}_N$ case and by $E_{ii} - E_{jj}$ for $i < j$ in the $\mathfrak{sl}_N$ case. We have the direct sum decompositions $\mathfrak{b}_{\pm} = \mathfrak{h} \oplus \mathfrak{n}_{\pm}$. We shall also make use of the corresponding subgroups $N_{\pm}$, $B_{\pm}$ and H in G . For GL_N these are the groups of unipotent upper/lower-triangular $N \times N$ matrices, invertible upper/lower-triangular $N \times N$ matrices and invertible diagonal $N \times N$ matrices, respectively. For SL_N we add the condition that the matrices are of determinant 1.

Recall the notation P_{12} for the tensor Casimir of $\mathfrak{g}$ from Sect. 2.1. We can split this into three parts as $P_{12} = P_{12}^{-} + P_{12}^0 + P_{12}^{+}$ where $P_{12}^{\pm} \in \mathfrak{n}_{\pm} \otimes \mathfrak{n}_{\mp}$ and $P_{12}^0 \in \mathfrak{h} \otimes \mathfrak{h}$. Explicitly, in the $\mathfrak{gl}_N$ case these read

$$P_{12}^{+} = \sum_{\substack{i,j=1 \\ i < j}}^N E_{ij} \otimes E_{ji}, \quad P_{12}^0 = \sum_{i=1}^N E_{ii} \otimes E_{ii}, \quad P_{12}^{-} = \sum_{\substack{i,j=1 \\ i < j}}^N E_{ji} \otimes E_{ij}.$$

For $\mathfrak{sl}_N$ the expression for P_{12}^0 is given in terms of dual bases $\{u^i\}$ and $\{u_i\}$ of the Cartan subalgebra $\mathfrak{h}$ with respect to the trace bilinear form as $P_{12}^0 = \sum_{i=1}^{N-1} u^i \otimes u_i$. We note that $P_{21}^{+} = P_{12}^{-}$, $P_{21}^0 = P_{12}^0$ and $P_{21}^{-} = P_{12}^{+}$. We also define the corresponding projectors $P^{\pm} : \mathfrak{g} \rightarrow \mathfrak{n}_{\pm}$ and $P^0 : \mathfrak{g} \rightarrow \mathfrak{h}$ onto the nilpotent Lie subalgebras $\mathfrak{n}_{\pm}$ and the Cartan subalgebra $\mathfrak{h}$, respectively, given for any $X \in \mathfrak{g}$ as

$$P^{\pm} X := \text{Tr}_2(P_{12}^{\pm} X_2), \quad P^0 X := \text{Tr}_2(P_{12}^0 X_2),$$

so that $\text{id}_{\mathfrak{g}} = P^- + P^0 + P^+$.

In the trigonometric setting, the role of the Lie subalgebra $\mathcal{A}_\lambda^{\text{rat}}(\mathfrak{g}) \subset \mathcal{A}_\lambda(\mathfrak{g})$ in (2.4) will be played by the following alternative Lie subalgebra

$$\mathcal{A}_\lambda^{\text{trig}}(\mathfrak{g}) := \mathcal{B}_\lambda^{0,\infty}(\mathfrak{g}) \times \coprod_{a \in \mathbb{C}^\times} \mathfrak{g} \otimes \mathbb{C}[\lambda_a] \quad (2.13)$$

where $\mathbb{C}^\times := \mathbb{C} \setminus \{0\}$ and

$$\mathcal{B}_\lambda^{0,\infty}(\mathfrak{g}) \subset (\mathfrak{b}_+ \oplus \mathfrak{g} \otimes \lambda \mathbb{C}[\lambda]) \times (\mathfrak{b}_- \oplus \mathfrak{g} \otimes \lambda_\infty \mathbb{C}[\lambda_\infty])$$

is the Lie subalgebra consisting of pairs of Taylor series $X^0(\lambda) = \sum_{n=0}^\infty X_n^0 \lambda^n$ and $X^\infty(\lambda_\infty) = \sum_{n=0}^\infty X_n^\infty \lambda_\infty^n$ with $X_n^0, X_n^\infty \in \mathfrak{g}$ for all $n \geq 1$ but with $X_0^0 \in \mathfrak{b}_+$ and $X_0^\infty \in \mathfrak{b}_-$ subject to the constraint $P^0 X_0^0 = -P^0 X_0^\infty$. We shall also need the corresponding group $\mathcal{A}_\lambda^{\text{trig}}(G)$ defined as follows.

In the GL_N case we let $\widehat{B}_+$ denote the group of all invertible $N \times N$ matrices with entries below the diagonal in $\lambda \mathbb{C}[\lambda]$ and entries on or above the diagonal in $\mathbb{C}[\lambda]$. Likewise, we let $\widehat{B}_-$ be the group of all invertible $N \times N$ matrices with entries on or below the diagonal in $\mathbb{C}[\lambda_\infty]$ and entries above the diagonal in $\lambda_\infty \mathbb{C}[\lambda_\infty]$. Concretely, an element of $\widehat{B}_+$ can be expanded as a Taylor series $g(\lambda) = \sum_{n=0}^\infty g_n \lambda^n$ with g_0 upper triangular and $g_n \in \mathfrak{gl}_N$ for $n \geq 1$, while an element of $\widehat{B}_-$ is a Taylor series $h(\lambda_\infty) = \sum_{n=0}^\infty h_n \lambda_\infty^n$ with h_0 lower triangular and $h_n \in \mathfrak{gl}_N$ for $n \geq 1$. As usual, in the SL_N case we define the subgroups $\widehat{B}_\pm$ as in the GL_N case but with the added condition that the matrices are of determinant 1. We then set

$$\mathcal{A}_\lambda^{\text{trig}}(G) := \mathcal{B}_\lambda^{0,\infty}(G) \times \coprod_{a \in \mathbb{C}} \widehat{G}_a \quad (2.14)$$

where the first factor is the subgroup $\mathcal{B}_\lambda^{0,\infty}(GL_N) \subset \widehat{B}_+ \times \widehat{B}_-$ consisting of pairs of Taylor series $g^0(\lambda) = \sum_{n=0}^\infty g_n^0 \lambda^n$ and $g^\infty(\lambda_\infty) = \sum_{n=0}^\infty g_n^\infty \lambda_\infty^n$ with $g_n^0, g_n^\infty \in \mathfrak{gl}_N$ for all $n \geq 1$ but where the upper triangular matrix g_0^0 and the lower triangular matrix g_0^∞ are subject to the constraint $P^0 g_0^0 = (P^0 g_0^\infty)^{-1}$.

Note that for consistency we should really keep denoting the local coordinate at the origin as λ_0 , following the general notation introduced in Sect. 2.1. However, since λ_0 is nothing but λ , we will most often prefer to write the local coordinate at the origin simply as λ , rather than use the more cumbersome notation λ_0 .

It will be convenient in what follows to introduce slightly different notions of pole parts of Laurent series at the origin and infinity in the trigonometric case. As they are important in practical calculations, we gather them in the following definition.

Definition 2.5. Given any $X^0(\lambda) = \sum_{n=-N_0}^\infty X_n^0 \lambda^n \in \mathfrak{g} \otimes \mathbb{C}((\lambda))$ we define

$$X^0(\lambda)_-^{\text{trig}} := (P^- + \frac{1}{2}P^0)X_0^0 + X^0(\lambda)_-^{\text{rat}} \in \mathfrak{b}_- \oplus \mathfrak{g} \otimes \lambda^{-1}\mathbb{C}[\lambda^{-1}]. \quad (2.15a)$$

Similarly, for any $X^\infty(\lambda^{-1}) = \sum_{n=-N_\infty}^\infty X_n^\infty \lambda^{-n} \in \mathfrak{g} \otimes \mathbb{C}((\lambda^{-1}))$ we define

$$X^\infty(\lambda_\infty)_-^{\text{trig}} := (P^+ + \frac{1}{2}P^0)X_0^\infty + \sum_{n=-N_\infty}^{-1} X_n^\infty \lambda_\infty^n \in \mathfrak{b}_+ \oplus \mathfrak{g} \otimes \lambda_\infty^{-1}\mathbb{C}[\lambda_\infty^{-1}]. \quad (2.15b)$$

Furthermore, for a Laurent series $X^b(\lambda_b) = \sum_{n=-N_b}^{\infty} X_n^b \lambda_b^n \in \mathfrak{g} \otimes \mathbb{C}((\lambda_b))$ at any other point $b \in \mathbb{C}^\times$ we set

$$X^b(\lambda_b)_-^{\text{trig}} := -\left(P^- + \frac{1}{2}P^0\right)X^b(-b)_-^{\text{rat}} + X^b(\lambda_b)_-^{\text{rat}} \in \mathfrak{b}_- \oplus \mathfrak{g} \otimes \lambda_b^{-1}\mathbb{C}[\lambda_b^{-1}], \quad (2.15c)$$

where in the first term $X^b(-b)_-^{\text{rat}}$ is the pole part $X^b(\lambda_b)_-^{\text{rat}}$ at b evaluated at $\lambda = 0$. In particular, as compared to the pole part $X^b(\lambda_b)_-^{\text{rat}} \in \mathfrak{g} \otimes \lambda_b^{-1}\mathbb{C}[\lambda_b^{-1}]$ introduced in (2.6a), we note that the pole part $X^b(\lambda_b)_-^{\text{trig}}$ includes a constant term (provided that $(P^- + \frac{1}{2}P^0)X^b(-b)_-^{\text{rat}} \neq 0$) which, moreover, is valued in $\mathfrak{b}_-$.

Proposition 2.6. *The Lie subalgebra $\mathcal{A}_\lambda^{\text{trig}}(\mathfrak{g}) \subset \mathcal{A}_\lambda(\mathfrak{g})$ is maximally isotropic with respect to $\langle\langle \cdot, \cdot \rangle\rangle_{-1}$. Moreover, we have a direct sum of vector spaces*

$$\mathcal{A}_\lambda(\mathfrak{g}) = \mathcal{A}_\lambda^{\text{trig}}(\mathfrak{g}) \dot{+} \iota_\lambda R_\lambda(\mathfrak{g}) \quad (2.16)$$

into complementary Lagrangian (i.e. maximal isotropic) Lie subalgebras.

Proof. To see that $\mathcal{A}_\lambda^{\text{trig}}(\mathfrak{g})$ is isotropic with respect to the bilinear form $\langle\langle \cdot, \cdot \rangle\rangle_{-1}$, let $X(\lambda), Y(\lambda) \in \mathcal{A}_\lambda^{\text{trig}}(\mathfrak{g})$ be arbitrary and consider the pairing $\langle\langle X(\lambda), Y(\lambda) \rangle\rangle_{-1}$ as given in (2.2b). There are no contributions from any $a \in \mathbb{C}^\times$. The only contributions come from 0 and ∞ , which read

$$\begin{aligned} & \text{res}_0^\lambda \text{Tr} \left(X^0(\lambda) Y^0(\lambda) \right) \lambda^{-1} d\lambda + \text{res}_\infty^\lambda \text{Tr} \left(X^\infty(\lambda_\infty) Y^\infty(\lambda_\infty) \right) \lambda^{-1} d\lambda \\ &= \text{Tr}(X_0^0 Y_0^0) - \text{Tr}(X_0^\infty Y_0^\infty) \\ &= \text{Tr}(P^0(X_0^0)P^0(Y_0^0)) - \text{Tr}(P^0(X_0^\infty)P^0(Y_0^\infty)) = 0. \end{aligned}$$

In the first equality we wrote $X^0(\lambda) = \sum_{n=0}^\infty X_n^0 \lambda^n$, $X^\infty(\lambda_\infty) = \sum_{n=0}^\infty X_n^\infty \lambda_\infty^n$ and similarly for $Y^0(\lambda)$ and $Y^\infty(\lambda_\infty)$. The second equality above follows from the fact that $X_0^0, Y_0^0 \in \mathfrak{b}_+$ and $X_0^\infty, Y_0^\infty \in \mathfrak{b}_-$ and the last step makes use of the conditions in the definition of $\mathcal{B}_\lambda^{0,\infty}(\mathfrak{g})$ that $P^0 X_0^0 = -P^0 X_0^\infty$ and $P^0 Y_0^0 = -P^0 Y_0^\infty$. In order to show that $\mathcal{A}_\lambda^{\text{trig}}(\mathfrak{g})$ is maximally isotropic it suffices to prove the second statement, namely that we have the direct sum decomposition of vector spaces as in (2.16).

To any $X(\lambda) \in \mathcal{A}_\lambda(\mathfrak{g})$ we associate the rational function

$$f_X(\lambda) = X^0(\lambda)_-^{\text{trig}} + \sum_{b \in \mathbb{C}^\times} X^b(\lambda_b)_-^{\text{trig}} + X^\infty(\lambda_\infty)_-^{\text{trig}} \quad (2.17)$$

in $R_\lambda(\mathfrak{g})$. Consider the element $\tilde{X}(\lambda) = (\tilde{X}^a(\lambda_a))_{a \in \mathbb{C}P^1}$ defined by

$$\tilde{X}^a(\lambda_a) := X^a(\lambda_a) - \iota_{\lambda_a} f_X(\lambda)$$

for every $a \in \mathbb{C}P^1$. We have $\tilde{X}^a(\lambda_a) \in \mathfrak{g} \otimes \mathbb{C}[[\lambda_a]]$ for every $a \in \mathbb{C}P^1$. But more precisely, noting that

$$X^b(\lambda_b)_-^{\text{trig}}|_{\lambda=0} = (P^+ + \frac{1}{2}P^0)X^b(-b)_-^{\text{rat}}, \quad X^\infty(\lambda_\infty)_-^{\text{trig}}|_{\lambda=0} = (P^+ + \frac{1}{2}P^0)X_0^\infty$$

for every $b \in \mathbb{C}^\times$, we have, in fact, $\tilde{X}^0(\lambda_0) \in \mathfrak{b}_+ \oplus \mathfrak{g} \otimes \lambda \mathbb{C}[[\lambda]]$ whose leading term in $\mathfrak{b}_+$ is given by

$$(P^+ + \tfrac{1}{2}P^0)(X_0^0 - X_0^\infty - X^b(-b)_-^{\text{rat}}) \in \mathfrak{b}_+. \quad (2.18)$$

Likewise, we have

$$\begin{aligned} X^b(\lambda_b)_-^{\text{trig}}|_{\lambda=\infty} &= -(P^- + \tfrac{1}{2}P^0)X^b(-b)_-^{\text{rat}}, \\ X^0(\lambda)_-^{\text{trig}}|_{\lambda=\infty} &= (P^- + \tfrac{1}{2}P^0)X_0^0 \end{aligned}$$

from which it follows that $\tilde{X}^\infty(\lambda_\infty) \in \mathfrak{b}_- \oplus \mathfrak{g} \otimes \lambda_\infty \mathbb{C}[[\lambda_\infty]]$ with leading coefficient in $\mathfrak{b}_-$ given by

$$(P^- + \tfrac{1}{2}P^0)(-X_0^0 + X_0^\infty + X^b(-b)_-^{\text{rat}}) \in \mathfrak{b}_-. \quad (2.19)$$

Moreover, comparing the Cartan components of (2.18) and (2.19) we see that these are opposite. Hence we conclude that $\tilde{X}(\lambda) \in \mathcal{A}_\lambda^{\text{trig}}(\mathfrak{g})$. In other words,

$$X(\lambda) = \tilde{X}(\lambda) + \iota_\lambda f_X(\lambda)$$

gives the desired decomposition of a general element $X(\lambda) \in \mathcal{A}_\lambda(\mathfrak{g})$ as in (2.16).

This decomposition is unique since any element which belongs to both $\mathcal{A}_\lambda^{\text{trig}}(\mathfrak{g})$ and $\iota_\lambda R_\lambda(\mathfrak{g})$ must vanish. Indeed, suppose $f(\lambda) \in R_\lambda(\mathfrak{g})$ is such that $\iota_\lambda f(\lambda) \in \mathcal{A}_\lambda^{\text{trig}}(\mathfrak{g})$. Then it is clear from the definition of $\mathcal{A}_\lambda^{\text{trig}}(\mathfrak{g})$ in (2.13) that $f(\lambda)$ cannot be singular at any point in $\mathbb{C}P^1$ and so must be constant. But then it follows from the definition of $\mathcal{B}_\lambda^{0,\infty}(\mathfrak{g})$ that this constant must in fact be zero. $\square$

Definition 2.7. (*Trigonometric r -matrix*). The trigonometric r -matrix is defined as the following function of two formal variables λ and μ :

$$r_{12}^{\text{trig}}(\lambda, \mu) = \frac{1}{2} \left(P_{12}^+ - P_{12}^- + \frac{\mu + \lambda}{\mu - \lambda} P_{12} \right) = \frac{\mu P_{12}}{\mu - \lambda} - P_{12}^- - \frac{1}{2} P_{12}^0. \quad (2.20)$$

It is skew-symmetric, namely we have $r_{21}^{\text{trig}}(\mu, \lambda) = -r_{12}^{\text{trig}}(\lambda, \mu)$. It provides the trigonometric counterpart of the kernel (2.10) for the choice of complement (2.13). Indeed, we have the following analogue of Proposition 2.3 in the trigonometric case.

Proposition 2.8. *For any $X(\lambda) \in \mathcal{A}_\lambda(\mathfrak{g})$, its projections onto the complementary subalgebras $\mathcal{A}_\lambda^{\text{trig}}(\mathfrak{g})$ and $\iota_\lambda R_\lambda(\mathfrak{g})$ relative to the direct sum decomposition (2.16) are given respectively by $\pi_+^{\text{trig}} X(\lambda) = ((\pi_+^{\text{trig}} X)^a(\lambda_a))_{a \in \mathbb{C}P^1}$ and $\pi_-^{\text{trig}} X(\lambda) = ((\pi_-^{\text{trig}} X)^a(\lambda_a))_{a \in \mathbb{C}P^1}$ where*

$$(\pi_+^{\text{trig}} X)^a(\lambda_a) = \sum_{b \in \mathbb{C}P^1} \text{res}_b^\mu \text{Tr}_2 (\iota_{\mu_b} \iota_{\lambda_a} r_{12}^{\text{trig}}(\lambda, \mu) X^b(\mu_b)_2) \mu^{-1} d\mu, \quad (2.21a)$$

$$(\pi_-^{\text{trig}} X)^a(\lambda_a) = - \sum_{b \in \mathbb{C}P^1} \text{res}_b^\mu \text{Tr}_2 (\iota_{\lambda_a} \iota_{\mu_b} r_{12}^{\text{trig}}(\lambda, \mu) X^b(\mu_b)_2) \mu^{-1} d\mu. \quad (2.21b)$$

Proof. We first describe the image of π_-^{trig} explicitly. Then we show that π_-^{trig} sends $\mathcal{A}_\lambda^{\text{trig}}(\mathfrak{g})$ to zero and that it acts as the identity on $\iota_\lambda R_\lambda(\mathfrak{g})$. Hence, π_-^{trig} is the projection onto $\iota_\lambda R_\lambda(\mathfrak{g})$ along $\mathcal{A}_\lambda^{\text{trig}}(\mathfrak{g})$. Similarly, we prove that π_+^{trig} is the projection onto $\mathcal{A}_\lambda^{\text{trig}}(\mathfrak{g})$ along $\iota_\lambda R_\lambda(\mathfrak{g})$.

Given any $X(\lambda) \in \mathcal{A}_\lambda(\mathfrak{g})$, consider the $\mathfrak{g}$ -valued rational function

$$\begin{aligned} f_X(\lambda) &= - \sum_{b \in \mathbb{C}P^1} \text{res}_b^\mu \text{Tr}_2 (\iota_{\mu_b} r_{12}^{\text{trig}}(\lambda, \mu) X^b(\mu_b)_2) \mu^{-1} d\mu \\ &= \sum_{b \in \mathbb{C}P^1} \text{res}_b^\mu \left(\iota_{\mu_b} \mu^{-1} (P^- + \tfrac{1}{2} P^0) (X^b(\mu_b)) + \iota_{\mu_b} \frac{1}{\lambda - \mu} X^b(\mu_b) \right) d\mu. \end{aligned}$$

We compute the residues at each $b \in \mathbb{C}^\times$ and then at the origin and infinity. Firstly, for the residue at $b \in \mathbb{C}^\times$ we find $X^b(\lambda_b)_-^{\text{trig}}$. For the residue at the origin we find $X^0(\lambda)_-^{\text{trig}}$ and, likewise, for the residue at infinity we find

$$-(P^- + \tfrac{1}{2} P^0) X_0^\infty + X^\infty(\lambda_\infty)_-^{\text{rat}} = X^\infty(\lambda_\infty)_-^{\text{trig}},$$

where in the first expression we are using the pole part $X^\infty(\lambda_\infty)_-^{\text{rat}} \in \mathfrak{g} \otimes \mathbb{C}[\lambda_\infty^{-1}]$ defined in (2.6b) of a Laurent series at infinity, and in the second expression we are using the other notion of pole part $X^\infty(\lambda_\infty)_-^{\text{trig}}$ introduced above in (2.15b). Putting the above together we conclude that $f_X(\lambda)$ is the rational function (2.17) used in the proof of Proposition 2.6. By construction we have $(\pi_-^{\text{trig}} X)^a(\lambda_a) = \iota_{\lambda_a} f_X(\lambda)$ for every $a \in \mathbb{C}P^1$.

Now suppose $X(\lambda) \in \mathcal{A}_\lambda^{\text{trig}}(\mathfrak{g})$. Clearly $X^b(\lambda_b)_-^{\text{rat}} = 0$, hence also $X^b(\lambda_b)_-^{\text{trig}} = 0$ using the definition (2.15c), so the sum over $b \in \mathbb{C}^\times$ on the right hand side of (2.17) vanishes. On the other hand, $X^0(\lambda)_-^{\text{trig}} = \tfrac{1}{2} P^0 X_0^0$ and $X^\infty(\lambda_\infty)_-^{\text{trig}} = \tfrac{1}{2} P^0 X_0^\infty$. But since $P^0 X_0^0 = -P^0 X_0^\infty$ by definition of $X(\lambda)$ belonging to $\mathcal{A}_\lambda^{\text{trig}}(\mathfrak{g})$, it follows that the remaining two terms in (2.17) cancel. So we have shown that $\pi_-^{\text{trig}} X(\lambda) = 0$ for any $X(\lambda) \in \mathcal{A}_\lambda^{\text{trig}}(\mathfrak{g})$.

On the other hand, suppose now that $X(\lambda) = \iota_\lambda f(\lambda)$ for some $f(\lambda) \in R_\lambda(\mathfrak{g})$. If the latter has a pole at some $a \in \mathbb{C}^\times$ then its pole part there is given by $X^a(\lambda_a)_-^{\text{rat}}$. If it has a pole at the origin then its pole part there is equal to

$$X^0(\lambda)_-^{\text{trig}} - \sum_{b \in \mathbb{C}^\times} (P^- + \tfrac{1}{2} P^0) X^b(-b)_-^{\text{rat}} - (P^- + \tfrac{1}{2} P^0) X_0^\infty, \quad (2.22)$$

where X_0^∞ is the constant term in the expansion of $f(\lambda)$ at infinity. Indeed, recall from (2.15a) that $X^0(\lambda)_-^{\text{trig}}$ is given by the pole part $X^0(\lambda)_-^{\text{rat}}$ at the origin plus $(P^- + \tfrac{1}{2} P^0) X_0^0$ where X_0^0 is given here by the value at the origin of all the other pole parts of $f(\lambda)$. This is why we must subtract the latter from $X^0(\lambda)_-^{\text{trig}}$ in (2.22) to be left only with the desired pole part at the origin. Finally, the pole part of $f(\lambda)$ at infinity is given by

$$X^\infty(\lambda_\infty)_-^{\text{trig}} + (P^- + \tfrac{1}{2} P^0) X_0^\infty. \quad (2.23)$$

Indeed, the pole part at infinity should contain the constant term but $X^\infty(\lambda_\infty)_-^{\text{trig}}$ only contains part of it. The remaining part is precisely the piece added in (2.23). It now follows that the expression on the right hand side of (2.17) built from $X(\lambda) = \iota_\lambda f(\lambda)$

coincides exactly with the partial fraction decomposition of $f(\lambda)$. This establishes that $\pi_-^{\text{trig}} X(\lambda) = X(\lambda)$ for any $X(\lambda) \in \iota_\lambda f(\lambda)$. In other words, we have therefore shown that π_-^{trig} is indeed the projection onto $\iota_\lambda R_\lambda(\mathfrak{g})$ along $\mathcal{A}_\lambda^{\text{trig}}(\mathfrak{g})$.

It remains to consider π_+^{trig} . For any $X(\lambda) \in \mathcal{A}_\lambda(\mathfrak{g})$ and $a \in \mathbb{C}$ we have

$$\begin{aligned} (\pi_+^{\text{trig}} X)^a(\lambda_a) &= - \sum_{b \in \mathbb{C}P^1} \text{res}_b^\mu \left(\iota_{\mu_b} \mu^{-1} (P^- + \tfrac{1}{2} P^0) (X^b(\mu_b)) \right) d\mu \\ &\quad + \sum_{n=0}^{\infty} \lambda_a^n \sum_{b \in \mathbb{C}P^1} \text{res}_b^\mu \iota_{\mu_b} \mu_a^{-n-1} X^b(\mu_b) d\mu. \end{aligned} \quad (2.24)$$

If $X(\lambda) \in \iota_\lambda R_\lambda(\mathfrak{g})$ then the first term on the right hand side vanishes by the residue theorem and the second term likewise at each order in the λ_a -expansion. If instead we consider $a = \infty$ then

$$\begin{aligned} (\pi_+^{\text{trig}} X)^\infty(\lambda_\infty) &= - \sum_{b \in \mathbb{C}P^1} \text{res}_b^\mu \left(\iota_{\mu_b} \mu^{-1} (P^- + \tfrac{1}{2} P^0) (X^b(\mu_b)) \right) d\mu \\ &\quad - \sum_{n=0}^{\infty} \lambda^{-n-1} \sum_{b \in \mathbb{C}P^1} \text{res}_b^\mu \iota_{\mu_b} \mu^n X^b(\mu_b) d\mu, \end{aligned} \quad (2.25)$$

but both terms vanish once again by the residue theorem if $X(\lambda) \in \iota_\lambda R_\lambda(\mathfrak{g})$. So we deduce that $\pi_+^{\text{trig}} X(\lambda) = 0$ for every $X(\lambda) \in \iota_\lambda R_\lambda(\mathfrak{g})$.

Suppose now that $X(\lambda) \in \mathcal{A}_\lambda^{\text{trig}}(\mathfrak{g})$. The first term on the right hand side of (2.24) gets a contribution only from the terms $b = 0$ and $b = \infty$, which read

$$\begin{aligned} & - \text{res}_0^\mu \left(\mu^{-1} (P^- + \tfrac{1}{2} P^0) (X^0(\mu)) \right) d\mu - \text{res}_\infty^\mu \left(\mu^{-1} (P^- + \tfrac{1}{2} P^0) (X^\infty(\mu_\infty)) \right) d\mu \\ &= - (P^- + \tfrac{1}{2} P^0) X_0^0 + (P^- + \tfrac{1}{2} P^0) X_0^\infty = (P^- + P^0) X_0^\infty = X_0^\infty \end{aligned}$$

where we wrote $X^0(\mu) = \sum_{n=0}^{\infty} X_n^0 \mu^n$ and $X^\infty(\mu_\infty) = \sum_{n=0}^{\infty} X_n^\infty \mu_\infty^n$. The second equality above follows since by assumption we have $X_0^0 \in \mathfrak{b}_+$ so that $P^- X_0^0 = 0$ and also $P^0 X_0^0 = -P^0 X_0^\infty$. The third equality also follows since by assumption $X_0^\infty \in \mathfrak{b}_-$. The second sum in (2.24) is just as in the rational case, however since the series at infinity now contains a constant term we get a contribution to the sum over $b \in \mathbb{C}P^1$ from both $b = a$ and $b = \infty$, yielding $X^a(\lambda_a) - X_0^\infty$. So in total, we deduce that $(\pi_+^{\text{trig}} X)^a(\lambda_a) = X^a(\lambda_a)$ for every $a \in \mathbb{C}$.

Consider now the case $a = \infty$. The first term on the right hand side of (2.25) is again equal to X_0^∞ while the second term gives $\sum_{n=1}^{\infty} X_n^\infty \lambda_\infty^n = X^\infty(\lambda_\infty) - X_0^\infty$. Putting these

together we deduce that $(\pi_+^{\text{trig}} X)^\infty(\lambda_\infty) = X^\infty(\lambda_\infty)$. In conclusion, we have shown that $\pi_+^{\text{trig}} X(\lambda) = X(\lambda)$ for all $X(\lambda) \in \mathcal{A}_\lambda^{\text{trig}}(\mathfrak{g})$ so that π_+^{trig} is the projection onto $\mathcal{A}_\lambda^{\text{trig}}(\mathfrak{g})$ along $\iota_\lambda R_\lambda(\mathfrak{g})$, as claimed. $\square$

We can now define the linear operator $r^{\text{trig}} := \pi_+^{\text{trig}} - \pi_-^{\text{trig}}$. It follows from Proposition 2.8 that its kernel reads

$$\left((\iota_{\mu_b} \iota_{\lambda_a} + \iota_{\lambda_a} \iota_{\mu_b}) r_{12}^{\text{trig}}(\lambda, \mu) \right)_{a,b \in \mathbb{C}P^1}. \quad (2.26)$$

Moreover, the kernel of the identity operator $\text{id} = \pi_+^{\text{trig}} + \pi_-^{\text{trig}}$ is similarly given by an expansion of zero since

$$((\iota_{\mu_b} \iota_{\lambda_a} - \iota_{\lambda_a} \iota_{\mu_b}) r_{12}^{\text{trig}}(\lambda, \mu) \mu^{-1} d\mu)_{a,b \in \mathbb{C}P^1} = (P_{12} \delta_{ab} \delta(\lambda_a, \mu_a) d\mu_a)_{a,b \in \mathbb{C}P^1} \quad (2.27)$$

using the same notation δ_{ab} and $\delta(\lambda, \mu)$ as in the rational case.

The following is the analogue of Lemma 2.4 in the trigonometric case.

Lemma 2.9. *Let $X(\mu) = (X^a(\mu_a))_{a \in \mathbb{C}P^1} \in \mathcal{A}_\mu(\mathfrak{g})$ with $X^a(\mu_a) = \sum_{n=-N_a}^\infty X_n^a \mu_a^n$ for some $N_a \in \mathbb{Z}$, where $N_a > 0$ for finitely many $a \in \mathbb{C}P^1$. For any $a \in \mathbb{C}$ we have*

$$\iota_{\mu_a} \text{Tr}_2(r_{12}^{\text{trig}}(\lambda, \mu) X^a(\mu_a)_2) = - \sum_{r=-N_a}^\infty \mu_a^r ((\lambda_a^{-r} + a \lambda_a^{-r-1}) X^a(\lambda_a))_-^{\text{trig}},$$

while at infinity we have

$$\iota_{\mu_\infty} \text{Tr}_2(r_{12}^{\text{trig}}(\lambda, \mu) X^\infty(\mu_\infty)_2) = \sum_{r=-N_\infty}^\infty \mu_\infty^r (\lambda_\infty^{-r} X^\infty(\lambda_\infty))_-^{\text{trig}}.$$

Proof. First, let $a \in \mathbb{C}^\times$. We have

$$\begin{aligned} \iota_{\mu_a} \text{Tr}_2(r_{12}^{\text{trig}}(\lambda, \mu) X^a(\mu_a)_2) &= -(P^- + \tfrac{1}{2} P^0)(X^a(\mu_a)) - \iota_{\mu_a} \frac{\mu}{\lambda - \mu} X^a(\mu_a) \\ &= - \sum_{r=-N_a}^\infty \mu_a^r (P^- + \tfrac{1}{2} P^0) X_r^a - \sum_{n=-N_a}^\infty \sum_{s=0}^\infty (\mu_a + a) \frac{\mu_a^{n+s}}{\lambda_a^{s+1}} X_n^a \\ &= - \sum_{r=-N_a}^\infty \mu_a^r \left((P^- + \tfrac{1}{2} P^0) X_r^a + \sum_{n=-N_a}^{r-1} \lambda_a^{n-r} X_n^a + a \sum_{n=-N_a}^r \lambda_a^{n-r-1} X_n^a \right). \end{aligned}$$

In the third equality we split the double sum into two terms, containing μ_a and a respectively from the first factor. We changed variable from s to $r = s + n + 1$ in the first and from s to $r = s + n$ in the second, and then changed the order of the two double sums. It remains to note that

$$\sum_{n=-N_a}^{r-1} \lambda_a^{n-r} X_n^a + a \sum_{n=-N_a}^r \lambda_a^{n-r-1} X_n^a = ((\lambda_a^{-r} + a \lambda_a^{-r-1}) X^a(\lambda_a))_-^{\text{rat}}$$

and that this is equal to $-X_r^a$ when evaluated at $\lambda = 0$. The result at $a \in \mathbb{C}^\times$ now follows by definition (2.15c) of the pole part at a .

At the origin we have

$$\begin{aligned} \iota_\mu \text{Tr}_2(r_{12}^{\text{trig}}(\lambda, \mu) X^0(\mu)_2) &= - \sum_{r=-N_0}^\infty \mu^r (P^- + \tfrac{1}{2} P^0) X_r^0 - \sum_{n=-N_0}^\infty \sum_{s=1}^\infty \mu^{n+s} \lambda^{-s} X_n^0 \\ &= - \sum_{r=-N_0}^\infty \mu^r \left((P^- + \tfrac{1}{2} P^0) X_r^0 + \sum_{n=-N_0}^{r-1} \lambda^{n-r} X_n^0 \right) = - \sum_{r=-N_0}^\infty \mu^r (\lambda^{-r} X^0(\lambda))_-^{\text{trig}}. \end{aligned}$$

In the second equality we changed variable in the double sum from s to $r = n + s$ and then changed the order of the two sums. We have also added the term $r = -N_0$ in this double sum since this term vanishes due to the range in the sum over n being empty. The last equality uses the definition (2.15a). Note that the result at the origin coincides with the result obtained above for $a \in \mathbb{C}^\times$ but taken at $a = 0$. This is not completely obvious since the definitions of the pole parts (2.15a) and (2.15c) at 0 and a generic point $a \in \mathbb{C}^\times$ are different. Likewise, at infinity we have

$$\begin{aligned} \iota_{\mu_\infty} \text{Tr}_2(r_{12}^{\text{trig}}(\lambda, \mu) X^\infty(\mu_\infty)_2) &= - \sum_{r=-N_\infty}^{\infty} \mu_\infty^r (P^- + \tfrac{1}{2} P^0) X_r^\infty + \sum_{n=-N_\infty}^{\infty} \sum_{s=0}^{\infty} \mu_\infty^{n+s} \lambda_\infty^{-s} X_n^\infty \\ &= \sum_{r=-N_\infty}^{\infty} \mu_\infty^r \left(- (P^- + \tfrac{1}{2} P^0) X_r^\infty + \sum_{n=-N_\infty}^r \lambda_\infty^{n-r} X_n^\infty \right) \\ &= \sum_{r=-N_\infty}^{\infty} \mu_\infty^r (\lambda_\infty^{-r} X^\infty(\lambda_\infty))_-^{\text{trig}}. \end{aligned}$$

In the second equality we changed variable in the double sum from s to $r = n + s$ and then changed the order of the two sums. The last equality uses (2.15b). $\square$

3. Generating Lagrangian Multiform and CYBE

In this section we will treat uniformly both the rational and trigonometric cases discussed in Sect. 2.2 and Sect. 2.3, respectively. More precisely, we shall work with the Lie algebra of $\mathfrak{g}$ -valued adèles $\mathcal{A}_\lambda(\mathfrak{g})$ equipped with the bilinear form (2.2) with either $k = 0$ or $k = -1$. The corresponding vector space direct sum decompositions (2.7) and (2.16) will be denoted by

$$\mathcal{A}_\lambda(\mathfrak{g}) = \mathcal{A}_\lambda^+(\mathfrak{g}) \dot{+} \iota_\lambda R_\lambda(\mathfrak{g})$$

where $\mathcal{A}_\lambda^+(\mathfrak{g})$ stands for the rational Lie subalgebra $\mathcal{A}_\lambda^{\text{rat}}(\mathfrak{g})$ when $k = 0$ and the trigonometric Lie subalgebra $\mathcal{A}_\lambda^{\text{trig}}(\mathfrak{g})$ when $k = -1$. Correspondingly, we shall use the common notation $\mathcal{A}_\lambda^+(G)$ for the groups $\mathcal{A}_\lambda^{\text{rat}}(G)$ in the rational case and $\mathcal{A}_\lambda^{\text{trig}}(G)$ in the trigonometric case.

Given a general element $X(\lambda) = (X^a(\lambda_a))_{a \in \mathbb{C}P^1} \in \mathcal{A}_\lambda(\mathfrak{g})$ of the Lie algebra of $\mathfrak{g}$ -valued adèles, we will also denote by $X^a(\lambda_a)_-$ the principal part of the formal Laurent series $X^a(\lambda_a)$, which stands for $X^a(\lambda_a)^{\text{rat}}_-$ in the rational case, see (2.6), or for $X^a(\lambda_a)^{\text{trig}}_-$ in the trigonometric case, see (2.15).

3.1. Dynamical equations. We will describe integrable field theories by taking the point of view that the spatial coordinate x , on which all the Hamiltonian fields are usually taken to depend, should be treated on an equal footing to all the other times in the hierarchy. To explain this new perspective on integrable hierarchies it is useful to begin by recalling the traditional point of view.

The dynamical equations of different integrable field theories in the same hierarchy are usually described as zero curvature equations

$$\partial_x V_n^a(\lambda) - \partial_{t_n}^a U(\lambda) + [V_n^a(\lambda), U(\lambda)] = 0 \quad (3.1)$$

where the Lax matrix $U(\lambda) \in R_\lambda(\mathfrak{g})$ is a coadjoint orbit of $\mathcal{A}_\lambda^+(G)$ in $R_\lambda(\mathfrak{g})$ which encodes the finite collection of *fields* of the hierarchy. The $V_n^a(\lambda) \in R_\lambda(\mathfrak{g})$, associated to the times t_n^a for some labels $a \in \mathbb{C}$ and $n \in \mathbb{Z}$ to be specified below and which we also refer to as Lax matrices, are coadjoint orbits of $\mathcal{A}_\lambda^+(G)$ in $R_\lambda(\mathfrak{g})$ built out of differential polynomials in the fields. From this traditional point of view, (3.1) represents a set of equations which is seen as a natural extension of the Lax equations $\partial_{t_n^a} L(\lambda) = [M_n^a(\lambda), L(\lambda)]$, used to describe finite-dimensional systems, to the field theory case where every degree of freedom now depends on x . In particular, $U(\lambda)$ is usually treated as the fundamental object since the $V_n^a(\lambda)$ can all be built out of it and as such it is seen as the natural analogue of the Lax matrix $L(\lambda)$ in the field theory case.

The crucial point is that the particular flow ∂_x can, and from our point of view *should*, be thought of as a linear combination of some of the *elementary* time flows $\partial_{t_n^a}$. But if we are to treat the coadjoint orbit $U(\lambda) \in R_\lambda(\mathfrak{g})$ on an equal footing to all the other coadjoint orbits $V_n^a(\lambda) \in R_\lambda(\mathfrak{g})$ then we should also abandon the idea that each $V_n^a(\lambda)$ is parametrised by differential polynomials with respect to x of the finite collection of fields contained in $U(\lambda)$. Instead, we should treat all the coadjoint orbits $V_n^a(\lambda) \in R_\lambda(\mathfrak{g})$ as truly independent. We shall see, in a sense which is much closer in spirit to the Lax formalism for finite-dimensional systems, that all the Lax matrices $V_n^a(\lambda)$ can be derived from a single object $\mathcal{Q}(\lambda) \in \mathcal{A}_\lambda(\mathfrak{g})$, a certain *adjoint* orbit of $\mathcal{A}_\lambda^+(G)$ in the full space of adèles $\mathcal{A}_\lambda(\mathfrak{g})$. In particular, the latter will satisfy a Lax equation (see (3.16) below)

$$\partial_{t_n^a} \mathcal{Q}(\lambda) = [\iota_\lambda V_n^a(\lambda), \mathcal{Q}(\lambda)]$$

with respect to all the times t_n^a . As such, in our approach to hierarchies of integrable field theories, $\mathcal{Q}(\lambda)$ will play a very similar role to that of the usual Lax matrix $L(\lambda)$ for finite-dimensional systems. For us, the fundamental object will therefore be $\mathcal{Q}(\lambda)$ rather than $U(\lambda)$. The relationship between these two objects, and in particular the connection between our approach to hierarchies of integrable field theories and the usual one recalled above, comes from fixing a particular linear combination of time flows as our choice of spatial derivative ∂_x . We discuss this in detail in Sect. 3.1.4, together with what we call the FNR procedure.

Since there is a close parallel between our treatment of integrable field theories and various familiar constructions in the theory of finite-dimensional integrable systems, we will draw the comparison throughout this section in a series of remarks.

3.1.1. Adjoint orbit Let $\phi(\lambda) = (\phi^a(\lambda_a))_{a \in \mathbb{C}P^1} \in \mathcal{A}_\lambda^+(G)$. We regard the entries of the matrix coefficients in the expansions

$$\phi^a(\lambda_a) = \sum_{n=0}^{\infty} \phi_n^a \lambda_a^n$$

for all $a \in \mathbb{C}P^1$ as an infinite collection of dynamical variables. In general, these are not all independent. For instance, $\phi^a(\lambda_a)$ should be invertible in the GL_N case, which means that the first term ϕ_0^a should be invertible, or $\phi^a(\lambda_a)$ should have determinant 1 in the SL_N case which will impose non-trivial relations between the coefficients at each order in λ_a . The infinitely many degrees of freedom contained in $\phi(\lambda)$, or equivalently in $\mathcal{Q}(\lambda)$ defined in (3.3) below, will be used to describe infinitely many different integrable hierarchies of integrable field theories. We will refer to these as *group* or *algebra coordinates* (respectively): they represent the dependent variables and are the fields satisfying the equations of motion of a hierarchies.

A particular integrable hierarchy will be determined by a choice of non-dynamical rational function, with poles in a finite subset $S \subset \mathbb{C}P^1$, which we can write using a partial fraction decomposition as

$$F(\lambda) = \sum_{a \in S} F^a(\lambda_a)_- \in R_\lambda(\mathfrak{g})$$

where $F^a(\lambda_a)_- \in \mathfrak{g} \otimes \mathbb{C}[\lambda_a^{-1}]$ are (rational or trigonometric, depending on the case) principal parts at each $a \in S$. In particular, $F^a(\lambda_a)_- d\lambda$ has a pole of order $N_a > 0$ at any $a \in S \cap \mathbb{C}$ and a pole of order $N_\infty + 2 \geq 2$ at infinity if $\infty \in S$. Its expansion at all of the points $a \in \mathbb{C}P^1$ defines an element $\iota_\lambda F(\lambda) = (\iota_{\lambda_a} F(\lambda))_{a \in \mathbb{C}P^1} \in \mathcal{A}_\lambda(\mathfrak{g})$ of the $\mathfrak{g}$ -valued adèles, via the embedding (2.3). By design, we have $(\iota_{\lambda_a} F(\lambda))_- = F^a(\lambda_a)_-$ for each $a \in S$ and $(\iota_{\lambda_a} F(\lambda))_- = 0$ for every other points $a \in \mathbb{C}P^1 \setminus S$. The element of the $\mathfrak{g}$ -valued adèles with these components, which we can denote by

$$(\iota_\lambda F(\lambda))_- := (F^a(\lambda_a)_-)_{a \in \mathbb{C}P^1} \in \mathcal{A}_\lambda(\mathfrak{g}), \quad (3.2)$$

is just a finite collection of principal parts. We consider its adjoint orbit under the group element $\phi(\lambda) \in \mathcal{A}_\lambda^+(G)$ introduced above, namely

$$\mathcal{Q}(\lambda) := \phi(\lambda)(\iota_\lambda F(\lambda))_- \phi(\lambda)^{-1} \in \mathcal{A}_\lambda(\mathfrak{g}). \quad (3.3)$$

Explicitly, its component at any pole $a \in S$ is $\mathcal{Q}^a(\lambda_a) = \phi^a(\lambda_a) F^a(\lambda_a)_- \phi^a(\lambda_a)^{-1}$ while the component at any other $a \in \mathbb{C}P^1 \setminus S$ vanishes. We can further expand the latter as a Laurent series in λ_a , namely

$$\mathcal{Q}^a(\lambda_a) = \sum_{n=-N_a}^{\infty} \mathcal{Q}_n^a \lambda_a^n, \quad (3.4)$$

for some $\mathcal{Q}_n^a \in \mathfrak{g}$, where $N_a > 0$ is the order of the pole of $F(\lambda)$ at $a \in S \cap \mathbb{C}$. For the point at infinity we can have $N_\infty \geq 0$.

Remark 3.1. The adjoint orbit (3.3) within the full Lie algebra of $\mathfrak{g}$ -valued adèles $\mathcal{A}_\lambda(\mathfrak{g})$ will play the role of the Lax matrix in the present infinite-dimensional setting. For comparison, it is useful to recall that in the finite-dimensional setting the Lax matrix is given by a coadjoint orbit

$$L(\lambda) = \Pi_- \left(\phi(\lambda)(\iota_\lambda F(\lambda))_- \phi(\lambda)^{-1} \right) \in R_\lambda(\mathfrak{g}) \quad (3.5)$$

where Π_- denotes either π_-^{rat} or π_-^{trig} , depending on whether we are in the rational or trigonometric setting, but without applying the expansion ι_λ so that we obtain an element of $R_\lambda(\mathfrak{g})$ rather than $\iota_\lambda R_\lambda(\mathfrak{g})$. In particular, the rational function $L(\lambda)$ depends only on finitely many dynamical variables in $\phi(\lambda) \in \mathcal{A}_\lambda(\mathfrak{g})$.

3.1.2. Generating Lax equation As our aim is to work with hierarchies of equations of motion, to each point $a \in \mathbb{C}P^1$ we attach an infinite family of time coordinates t_n^a for $n \in \mathbb{Z}$. Related to each time is the usual partial derivative $\partial_{t_n^a}$ (meant as a total derivative when acting on functions of the fields). For our purposes, let us define the following generating operators

$$\mathcal{D}_{\lambda_a} := \sum_{n \in \mathbb{Z}} \lambda_a^n \partial_{t_n^a}, \quad a \in \mathbb{C}, \quad \mathcal{D}_{\lambda_\infty} := \sum_{n \in \mathbb{Z}} \lambda_\infty^{n+k+1} \partial_{t_n^\infty} \quad (3.6)$$

with $k = 0$ in the rational case and $k = -1$ in the trigonometric case. We let $\mathcal{D}_\lambda := (\mathcal{D}_{\lambda_a})_{a \in \mathbb{C}P^1}$ denote the $\mathbb{C}P^1$ -tuple of these differential operators. Then, if μ is another formal variable we will use the notation

$$\mathcal{D}_\lambda \mathcal{Q}(\mu) = \left(\mathcal{D}_{\lambda_a} \mathcal{Q}^b(\mu_b) \right)_{a,b \in \mathbb{C}P^1} \quad (3.7)$$

which encodes the flows $\partial_{t_m^a} \mathcal{Q}_n^b$ of all the dynamical variables $\mathcal{Q}_n^b$ with respect to all the times t_m^a for each pair of points $a, b \in \mathbb{C}P^1$.

Following the first observation in Sect. 1.2 of the introduction, we want to declare the evolution of $\mathcal{Q}(\lambda) \in \mathcal{A}_\lambda(\mathfrak{g})$ with respect to the above infinite family of times t_n^a to be governed by the following general Lax equation in r -matrix and generating form

$$\mathcal{D}_\mu \mathcal{Q}_1(\lambda) = [\text{Tr}_2 (\iota_\lambda \iota_\mu r_{12}(\lambda, \mu) \mathcal{Q}_2(\mu)), \mathcal{Q}_1(\lambda)]. \quad (3.8)$$

However, a few comments and precautions are necessary. First, writing such an equation with the understanding that $\mathcal{D}_\mu$ is the $\mathbb{C}P^1$ -tuple of commuting differential operators defined in (3.6) assumes that the vector fields on the right-hand side commute, if we want to be able to interpret the times t_n^a as coordinates on a manifold. In other words, defining the generating vector $\mathcal{X}_\mu$ acting on $\mathcal{A}_\lambda(\mathfrak{g})$ by

$$\mathcal{X}_\mu \mathcal{Q}_1(\lambda) = [\text{Tr}_2 (\iota_\lambda \iota_\mu r_{12}(\lambda, \mu) \mathcal{Q}_2(\mu)), \mathcal{Q}_1(\lambda)], \quad (3.9)$$

we must first prove that $[\mathcal{X}_\mu, \mathcal{X}_\nu] = 0$. Only then can we set $\mathcal{X}_\mu = \mathcal{D}_\mu$ and view the generating Lax equation (3.8) as describing compatible time flows on $\mathcal{A}_\lambda(\mathfrak{g})$. This is shown below in Proposition 3.4 and is a beautiful consequence of the CYBE for r .

Second, note that the right-hand side of (3.8) lives in $\coprod_{a,b \in \mathbb{C}P^1} \mathfrak{g} \otimes \lambda^{-N_a} \mu^{-N_b} \mathbb{C}[[\lambda, \mu]]$. Indeed, at $b \in \mathbb{C}P^1$ the power of μ_b is bounded below by $-N_b$ since $\iota_{\lambda_a} \iota_{\mu_b} r_{12}(\lambda, \mu)$ is a Taylor series in μ_b while $\mathcal{Q}_2^b(\mu_b)$ is a Laurent series with leading term of order $\mu_b^{-N_b}$ by definition (3.4). By the following lemma we then also deduce that at $a \in \mathbb{C}P^1$ the power of λ_a on the right hand side of (3.8) is bounded below by $-N_a$. For the left hand side, this means that the flow with respect to the times t_m^b , with $m < -N_b$ are trivial: $\mathcal{Q}(\lambda)$ does not depend on those times and for all practical purposes related to a hierarchy of field theories, they can be ignored.

Lemma 3.2. *We have*

$$[\text{Tr}_2 (\iota_\lambda \iota_\mu r_{12}(\lambda, \mu) \mathcal{Q}_2(\mu)), \mathcal{Q}_1(\lambda)] = [\text{Tr}_2 (\iota_\mu \iota_\lambda r_{12}(\lambda, \mu) \mathcal{Q}_2(\mu)), \mathcal{Q}_1(\lambda)].$$

Proof. Using the identity (2.11) (or (2.27) in the trigonometric case) we deduce that for any $a, b \in \mathbb{C}P^1$ we have

$$[\text{Tr}_2 ((\iota_{\lambda_a} \iota_{\mu_b} - \iota_{\mu_b} \iota_{\lambda_a}) r_{12}(\lambda, \mu) \mathcal{Q}_2^b(\mu_b)), \mathcal{Q}_1^a(\lambda_a)] \propto \delta(\lambda_a, \mu_a) [\mathcal{Q}^a(\lambda_a), \mathcal{Q}^a(\mu_a)].$$

Since $[\mathcal{Q}^a(\lambda_a), \mathcal{Q}^a(\mu_a)]$ vanishes when $\lambda_a = \mu_a$ it is proportional to $\lambda_a - \mu_a$ and so it follows that the right hand side above vanishes, as required. $\square$

Remark 3.3. The Lax equation (3.8) is to be compared with the Lax equation in the usual finite-dimensional setting for the evolution of the Lax matrix with respect to the times associated with the coefficients in the partial fraction decomposition of the quadratic Hamiltonian

$$H(\mu) = \frac{1}{2} \operatorname{Tr} (L(\mu)^2) = \sum_{a \in S} \sum_{n=0}^{n_a-1} \frac{H_n^a}{(\mu - a)^{n+1}}.$$

If we gather together the flows $\partial_{t_n^a} = \{H_n^a, \cdot\}$ associated with the Hamiltonians H_n^a by defining the differential operator valued rational function

$$\mathcal{D}_\mu = \sum_{a \in S} \sum_{n=0}^{n_a-1} \frac{\partial_{t_n^a}}{(\mu - a)^{n+1}},$$

which is to be compared with the adèlic object (3.6) in the present infinite-dimensional setting, then the Lax equations in the finite-dimensional setting take the form

$$\mathcal{D}_\mu L_1(\lambda) = [\operatorname{Tr}_2 (r_{12}(\lambda, \mu) L_2(\mu)), L_1(\lambda)]. \quad (3.10)$$

Both sides of this equation are $\mathfrak{g}$ -valued rational functions in both λ and μ with poles in λ and μ at each $a \in S$ of order at most N_a .

Proposition 3.4. *The flows (3.8) are compatible as a consequence of the commutativity of the corresponding vector fields, i.e. for any three formal variables λ , μ and ν we have*

$$\mathcal{X}_\nu \mathcal{X}_\mu \mathcal{Q}(\lambda) = \mathcal{X}_\mu \mathcal{X}_\nu \mathcal{Q}(\lambda). \quad (3.11)$$

Proof. We have

$$\begin{aligned} \mathcal{X}_\nu \mathcal{X}_\mu \mathcal{Q}_1(\lambda) &= [\operatorname{Tr}_2 (\iota_\lambda \iota_\mu r_{12}(\lambda, \mu) \mathcal{X}_\nu \mathcal{Q}_2(\mu)), \mathcal{Q}_1(\lambda)] \\ &\quad + [\operatorname{Tr}_2 (\iota_\lambda \iota_\mu r_{12}(\lambda, \mu) \mathcal{Q}_2(\mu)), \mathcal{X}_\nu \mathcal{Q}_1(\lambda)] \\ &= \operatorname{Tr}_{23} [\iota_\lambda \iota_\mu r_{12}(\lambda, \mu) [\iota_\mu \iota_\nu r_{23}(\mu, \nu) \mathcal{Q}_3(\nu), \mathcal{Q}_2(\mu)], \mathcal{Q}_1(\lambda)] \\ &\quad + \operatorname{Tr}_{23} [\iota_\lambda \iota_\mu r_{12}(\lambda, \mu) \mathcal{Q}_2(\mu), [\iota_\lambda \iota_\nu r_{13}(\lambda, \nu) \mathcal{Q}_3(\nu), \mathcal{Q}_1(\lambda)]]. \end{aligned} \quad (3.12)$$

By using the cyclicity of the trace over space 2 in the first term on the right hand side and the Jacobi identity on the last term, this can be rewritten as

$$\begin{aligned} \mathcal{X}_\nu \mathcal{X}_\mu \mathcal{Q}_1(\lambda) &= \operatorname{Tr}_{23} [\iota_\lambda \iota_\mu \iota_\nu [r_{12}(\lambda, \mu), r_{23}(\mu, \nu)] \mathcal{Q}_2(\mu) \mathcal{Q}_3(\nu), \mathcal{Q}_1(\lambda)] \\ &\quad + \operatorname{Tr}_{23} [\iota_\lambda \iota_\mu \iota_\nu [r_{12}(\lambda, \mu), r_{13}(\lambda, \nu)] \mathcal{Q}_2(\mu) \mathcal{Q}_3(\nu), \mathcal{Q}_1(\lambda)] \\ &\quad + \operatorname{Tr}_{23} [\iota_\lambda \iota_\nu r_{13}(\lambda, \nu) \mathcal{Q}_3(\nu), [\iota_\lambda \iota_\mu r_{12}(\lambda, \mu) \mathcal{Q}_2(\mu), \mathcal{Q}_1(\lambda)]]. \end{aligned}$$

Likewise, exchanging $\mu \leftrightarrow \nu$ in (3.12) we obtain

$$\begin{aligned} \mathcal{X}_\mu \mathcal{X}_\nu \mathcal{Q}_1(\lambda) &= \operatorname{Tr}_{23} [\iota_\lambda \iota_\nu r_{13}(\lambda, \nu) [\iota_\nu \iota_\mu r_{32}(\nu, \mu) \mathcal{Q}_2(\mu), \mathcal{Q}_3(\nu)], \mathcal{Q}_1(\lambda)] \\ &\quad + \operatorname{Tr}_{23} [\iota_\lambda \iota_\nu r_{13}(\lambda, \nu) \mathcal{Q}_3(\nu), [\iota_\lambda \iota_\mu r_{12}(\lambda, \mu) \mathcal{Q}_2(\mu), \mathcal{Q}_1(\lambda)]] \\ &= \operatorname{Tr}_{23} [\iota_\lambda \iota_\mu \iota_\nu [r_{13}(\lambda, \nu), r_{32}(\nu, \mu)] \mathcal{Q}_2(\mu) \mathcal{Q}_3(\nu), \mathcal{Q}_1(\lambda)] \\ &\quad + \operatorname{Tr}_{23} [\iota_\lambda \iota_\nu r_{13}(\lambda, \nu) \mathcal{Q}_3(\nu), [\iota_\lambda \iota_\mu r_{12}(\lambda, \mu) \mathcal{Q}_2(\mu), \mathcal{Q}_1(\lambda)]], \end{aligned}$$

where in the second equality we used Lemma 3.2 to swap the order of ι_v and ι_μ in the first term, along with the cyclicity of the trace over space 3. Thus $[\mathcal{X}_v, \mathcal{X}_\mu] \mathcal{Q}_1(\lambda)$ equals

$$\begin{aligned} \text{Tr}_{23} \Big[& \iota_\lambda \iota_\mu \iota_v \Big([r_{12}(\lambda, \mu), r_{13}(\lambda, v)] + [r_{12}(\lambda, \mu), r_{23}(\mu, v)] \\ & - [r_{13}(\lambda, v), r_{32}(v, \mu)] \Big) \mathcal{Q}_2(\mu) \mathcal{Q}_3(v), \mathcal{Q}_1(\lambda) \Big] \end{aligned}$$

which vanishes as a consequence of the CYBE (1.1). $\square$

3.1.3. Generating zero curvature equation In the context of integrable field theories the role of the Lax equation, cf. (3.10), is replaced by the zero curvature equation for a Lax connection. Therefore, as a first step towards relating the present formalism to integrable field theories, we now associate with each time t_n^a , for any $a \in S$ and $n \geq -N_a$, a rational Lax matrix $V_n^a(\lambda) \in R_\lambda(\mathfrak{g})$ such that any pair of these satisfies a zero curvature equation.

The equations of motion (3.8) can be written succinctly as

$$\mathcal{D}_\mu \mathcal{Q}(\lambda) = [\iota_\lambda V(\lambda; \mu), \mathcal{Q}(\lambda)] \quad (3.13)$$

where we have introduced

$$V(\lambda; \mu) := \text{Tr}_2 (\iota_\mu r_{12}(\lambda, \mu) \mathcal{Q}_2(\mu)). \quad (3.14)$$

Note that in (3.14) we do *not* expand the right hand side in powers of λ_a for $a \in \mathbb{C}P^1$, i.e. we do not apply the homomorphism ι_λ . Instead, this expansion is taken explicitly in (3.13). In particular, the semi-colon in the notation $V(\lambda; \mu)$ is used to emphasise that λ is just a formal variable whereas μ is the usual boldface notation used as a shorthand for a collection $(V^b(\lambda; \mu_b))_{b \in \mathbb{C}P^1}$ where

$$V^b(\lambda; \mu_b) = \sum_{n=-N_b}^{\infty} V_n^b(\lambda) \mu_b^n, \quad b \in \mathbb{C}, \quad (3.15a)$$

$$V^\infty(\lambda; \mu_\infty) = \sum_{n=-N_\infty}^{\infty} V_n^\infty(\lambda) \mu_\infty^{n+k+1}. \quad (3.15b)$$

As usual, we take $k = 0$ in the rational case and $k = -1$ in the trigonometric case. Here $V_n^b(\lambda) \in R_\lambda(\mathfrak{g})$ are $\mathfrak{g}$ -valued rational functions in λ with a pole at $\lambda = b$. Unpacking the notation in (3.13) slightly, recalling the definition of the operators $\mathcal{D}_\mu$ and (3.6), we see that the flow of $\mathcal{Q}(\lambda)$ with respect to the time t_n^a is controlled by $V_n^a(\lambda)$, namely we have the Lax equation

$$\partial_{t_n^a} \mathcal{Q}(\lambda) = [\iota_\lambda V_n^a(\lambda), \mathcal{Q}(\lambda)]. \quad (3.16)$$

Moreover, by the following proposition $V^b(\lambda; \mu_b)$ can be seen as a generating series in μ_b of a hierarchy of Lax matrices $V_n^b(\lambda)$ associated with the times t_n^b .

Proposition 3.5. *We have the zero curvature equation in generating form*

$$\mathcal{D}_v V(\lambda; \mu) - \mathcal{D}_\mu V(\lambda; v) + [V(\lambda; \mu), V(\lambda; v)] = 0. \quad (3.17)$$

Equivalently, in components we have the zero curvature equation

$$\partial_{t_n^b} V_m^a(\lambda) - \partial_{t_m^a} V_n^b(\lambda) + [V_m^a(\lambda), V_n^b(\lambda)] = 0$$

for every $a, b \in \mathbb{C}P^1$ and $m \geq -N_a$ and $n \geq -N_b$.

Proof. Using the Lax equation (3.8) we find

$$\begin{aligned}\mathcal{D}_v V(\lambda; \mu) &= \text{Tr}_2 (\iota_\mu r_{12}(\lambda, \mu) \mathcal{D}_v \mathcal{Q}_2(\mu)) \\ &= \text{Tr}_{23} (\iota_\mu r_{12}(\lambda, \mu) [\iota_\mu \iota_v r_{23}(\mu, v) \mathcal{Q}_3(v), \mathcal{Q}_2(\mu)]) \\ &= \text{Tr}_{23} (\iota_\mu \iota_v [r_{12}(\lambda, \mu), r_{23}(\mu, v)] \mathcal{Q}_2(\mu) \mathcal{Q}_3(v)),\end{aligned}$$

where in the last equality we used the cyclicity of the trace in space 2. Likewise, we also have

$$\begin{aligned}\mathcal{D}_\mu V(\lambda; v) &= \text{Tr}_3 (\iota_v r_{13}(\lambda, v) \mathcal{D}_\mu \mathcal{Q}_3(v)) \\ &= \text{Tr}_{23} (\iota_v r_{13}(\lambda, v) [\iota_v \iota_\mu r_{32}(v, \mu) \mathcal{Q}_2(\mu), \mathcal{Q}_3(v)]) \\ &= \text{Tr}_{23} (\iota_\mu \iota_v [r_{13}(\lambda, v), r_{32}(v, \mu)] \mathcal{Q}_2(\mu) \mathcal{Q}_3(v))\end{aligned}$$

where in the final step we used Lemma 3.2 to swap the order of ι_v and ι_μ , before using the cyclicity of the trace in space 3. Finally, we have

$$\begin{aligned}[V(\lambda; \mu), V(\lambda; v)] &= \text{Tr}_{23} [\iota_\mu r_{12}(\lambda, \mu) \mathcal{Q}_2(\mu), \iota_\lambda \iota_v r_{13}(\lambda, v) \mathcal{Q}_3(v)] \\ &= \text{Tr}_{23} (\iota_\mu \iota_v [r_{12}(\lambda, \mu), r_{13}(\lambda, v)] \mathcal{Q}_2(\mu) \mathcal{Q}_3(v)).\end{aligned}$$

The result now follows by the classical Yang–Baxter equation (1.1). $\square$

Remark 3.6. Note the clear resemblance between the generating series (3.14) for the hierarchy of Lax matrices $V_n^a(\lambda)$ and the usual generating rational function

$$M(\lambda; \mu) = \text{Tr}_2 (r_{12}(\lambda, \mu) L_2(\mu))$$

in the finite-dimensional case. The coefficients in the partial fraction decomposition of the latter with respect to μ are $\mathfrak{g}$ -valued rational matrices $M_n^a(\lambda)$ which control the flow of the Lax matrix $L(\lambda)$ with respect to the associated time t_n^a via the Lax equation $\partial_{t_n^a} L(\lambda) = [M_n^a(\lambda), L(\lambda)]$, which is to be compared with (3.16).

In the finite-dimensional case, however, one may also need to consider the more general generating rational function

$$M^{(n)}(\lambda; \mu) = \text{Tr}_2 (r_{12}(\lambda, \mu) L_2(\mu)^{n-1})$$

for integers $n \geq 2$. Indeed, the Lax equation in (3.10) involves $M(\lambda; \mu) = M^{(2)}(\lambda; \mu)$ which is only associated with the quadratic Hamiltonians $H(\mu) = \frac{1}{2} \text{Tr}(L(\mu)^2)$. But in the finite-dimensional setting one should equally consider the Lax equations where $M^{(n)}(\lambda; \mu)$ replaces $M^{(2)}(\lambda; \mu)$ since these describe the flows of the Lax matrix $L(\lambda)$ with respect to the higher order Hamiltonians built from $\frac{1}{n} \text{Tr}(L(\mu)^n)$.

In the present infinite-dimensional context, we observe that the generating series (3.14) is sufficient to produce the infinite number of Lax matrices $V_n^a(\lambda)$ associated with the infinite number of times t_n^a in the hierarchy that one expects from the traditional examples of the AKNS or the sine-Gordon hierarchies (see below). It is not clear to us what an appropriate analog of taking higher powers of $L(\lambda)$ is in terms of $\mathcal{Q}(\lambda)$ and whether the resulting Lax matrices and commuting flows would be independent of those obtained already.

Remark 3.7. It is instructive to compare the generating series (3.14) for the hierarchy of Lax matrices $V_n^a(\lambda)$ with formulas for similar generating series of Lax matrices obtained in the more traditional approach to integrable field theories which involves the monodromy matrix associated to a given auxiliary equation $\partial_x \Psi = U \Psi$. For example, in [FT, pp. 203–204], it is shown that the object

$$V(x, \lambda, \mu) = \frac{1}{2(\lambda - \mu)} (\mathbf{1} + W(x, \mu)) (-i\sigma_3) (\mathbf{1} + W(x, \mu))^{-1} \quad (3.18)$$

“is the generating series of the Lax matrices $V_n(x, \lambda)$ appearing in the zero curvature equation representation of the higher NS equations”. The expansion is to be understood as

$$V(x, \lambda, \mu) = \sum_{n=1}^{\infty} V_n(x, \lambda) \mu^{-n}. \quad (3.19)$$

The point is that (3.18) can be rewritten as

$$V(x, \lambda, \mu) = -\frac{1}{2} \text{Tr}_2 \left((\iota_{\mu_\infty} r_{12}(\lambda, \mu))_2 (\mathbf{1} + W(x, \mu))_2 (-i\sigma_3)_2 (\mathbf{1} + W(x, \mu))_2^{-1} \right). \quad (3.20)$$

We note the explicit dependence of the preferred variable x , indicative of the fact that this object has been built from a particular, preferred time x associated to the Lax matrix denoted $U(x, \lambda)$, which is nothing but $V_1(x, \lambda)$, as it should be. Other than this dependence, formula (3.18) has exactly the same structure as our formula (3.14) when specialised to the AKNS hierarchy, see Sect. 4. Indeed, in that case the only pole to consider is at infinity and the function $F(\lambda)$ is taken to be $-i\sigma_3$. Hence the only non zero element in the tuple (3.14) is

$$V^\infty(\lambda; \mu_\infty) = \text{Tr}_2 \left((\iota_{\mu_\infty} r_{12}(\lambda, \mu))_2 \phi_2^\infty(\mu_\infty) (-i\sigma_3)_2 \phi_2^\infty(\mu_\infty)^{-1} \right). \quad (3.21)$$

To complete the comparison, note that the term $(\mathbf{1} + W(x, \mu))$ in (3.20) comes from writing the monodromy matrix $T(x, y, \lambda)$ on the finite interval $[y, x]$, associated to $U(x, \lambda)$, as

$$T(x, y, \mu) = (\mathbf{1} + W(x, \mu)) e^{Z(x, y, \mu)} (\mathbf{1} + W(y, \mu))^{-1} \quad (3.22)$$

where Z is a diagonal matrix and both Z and W are Taylor series in $1/\mu$ (with no constant term for W). We refer the curious reader to [FT] for more details about Z and W which are not of importance for our discussion here. Considering for instance the case of fast decaying fields as $|x| \rightarrow \infty$, we can work with the monodromy matrix on $(-\infty, x)$

$$T^-(x, \mu) = (\mathbf{1} + W(x, \mu)) e^{Z^-(x, \mu)}. \quad (3.23)$$

This is the object that plays the role of our group element $\phi^\infty(\mu_\infty)$. Indeed, formally plugging $T^-(x, \mu)$ into (3.21) in place of $\phi^\infty(\mu_\infty)$, and remembering that $e^{Z^-(x, \mu)}$ commutes with σ_3 , we see that we get (3.20) (up to an irrelevant factor $-1/2$ which comes from a different choice of normalisation between us and [FT]).

In [ACDK, AC], the argument from [FT] was generalised to obtain the analog of formula (3.20) but where one now builds it from the monodromy matrix associated to the time t_k and Lax matrix $V_k(t_k, \lambda)$, for an arbitrary but fixed $k \geq 1$. This represented

the first step towards providing a generating function of Lax matrices that treats all times in the AKNS hierarchy equally. Our formula (3.14) achieves this fully in that it makes no reference to a preferred time and an associated monodromy matrix as a starting point. It is also valid well beyond the realm of AKNS only, as our various examples below demonstrate.

It will be useful, in view of applying our general framework to construct explicit examples in the next few sections, to be more explicit about the form of the Lax matrices $V_n^a(\lambda)$. This can be done using Lemma 2.4 in the rational case or Lemma 2.9 in the trigonometric case.

Proposition 3.8. *In the rational case, for every $a \in \mathbb{C}$ and $n \geq -N_a$, we have*

$$V_n^a(\lambda) = -(\lambda_a^{-n-1} Q^a(\lambda_a))_-^{\text{rat}},$$

while at infinity, for any $n \geq -N_\infty$ we have

$$V_n^\infty(\lambda) = (\lambda_\infty^{-n} Q^\infty(\lambda_\infty))_-^{\text{rat}}.$$

In the trigonometric case, for every $a \in \mathbb{C}$ and $n \geq -N_a$ we have

$$V_n^a(\lambda) = -((\lambda_a^{-n} + a\lambda_a^{-n-1}) Q^a(\lambda_a))_-^{\text{trig}},$$

which at the origin simply reads $V_n^0(\lambda) = -(\lambda^{-n} Q^0(\lambda))_-^{\text{trig}}$, while at infinity we have, for every $n \geq -N_\infty$,

$$V_n^\infty(\lambda) = (\lambda^n Q^\infty(\lambda^{-1}))_-^{\text{trig}}.$$

Proof. In the rational (resp. trigonometric) case this is a direct consequence of the definition (3.15) together with Lemma 2.4 (resp. Lemma 2.9). $\square$

Recall that $Q^a(\lambda_a) = 0$ if $a \in \mathbb{C}P^1 \setminus S$ so that, in fact, $V_n^a(\lambda) = 0$ unless $a \in S$. By construction each Lax matrix $V_n^a(\lambda) \in R_\lambda(\mathfrak{g})$ for any $a \in S$ and $n \geq -N_a$, or rather their embedding in $\mathcal{A}_\lambda(\mathfrak{g})$ via (2.3), is a coadjoint orbit in $\mathfrak{t}_\lambda R_\lambda(\mathfrak{g})$. For instance, in the rational case for $a \in \mathbb{C} \cap S$ we have

$$V_n^a(\lambda) = -\left(\phi^a(\lambda_a) \lambda_a^{-n-1} F^a(\lambda_a) - \phi^a(\lambda_a)^{-1}\right)_-^{\text{rat}} \in \mathfrak{g} \otimes \mathbb{C}[\lambda_a^{-1}] \subset R_\lambda(\mathfrak{g}).$$

3.1.4. Connection to integrable field theory and FNR procedure Up to this point, the framework we have been discussing is very similar to the one used to describe finite-dimensional integrable systems, as emphasised in Remarks 3.1, 3.3 and 3.6. However, as we will see explicitly in all the examples discussed in later sections, our formalism encodes entire hierarchies of integrable field theories!

The FNR procedure

One way to make explicit contact with the traditional approach to integrable field theory is to choose a preferred coordinate, denote it by x and set it as a particular combination of the fundamental times t_n^a for $a \in S$ and $n \geq -N_a$. Quite generally, we can choose some finite subsets $T_a \subset \mathbb{Z}_{\geq -N_a}$ for each $a \in S$ and define $\partial_x := \sum_{a \in S} \sum_{n \in T_a} r_n^a \partial_{t_n^a}$ for some $r_n^a \in \mathbb{C}^*$. The Lax matrix associated with the coordinate x is then given by

$$U(\lambda) := \sum_{a \in S} \sum_{n \in T_a} r_n^a V_n^a(\lambda) \in R_\lambda(\mathfrak{g}). \quad (3.24)$$

As explained above, $\iota_\lambda U(\lambda)$ is then a coadjoint orbit in the dual space $\iota_\lambda R_\lambda(\mathfrak{g})$ of $\mathcal{A}_\lambda^+(\mathfrak{g})$. This is the coadjoint orbit alluded to at the very start of this section which encodes the finite collection of fields of our integrable hierarchy. It follows from (3.13), or even more directly from (3.16), that the spatial dependence of $\mathcal{Q}(\lambda)$ is governed by the Lax equation

$$\partial_x \mathcal{Q}(\lambda) = [\iota_\lambda U(\lambda), \mathcal{Q}(\lambda)]. \quad (3.25)$$

As we will see on examples, the equation (3.25) can be solved recursively to express the coefficients Q_n^a of $Q^a(\lambda_a)$, cf. (3.4), as differential polynomials in the fields, i.e. the variables contained in the Lax matrix $U(\lambda)$. All other Lax matrices $V_n^a(\lambda)$ associated to the fundamental times t_n^a will then have components expressed as differential polynomials of the fields.

We will outline below how (3.25) can, in principle, be solved recursively for each Q_n^a . Since certain details of the recursive procedure depend on the model considered, we will only illustrate here the part of the construction which applies universally to all models in Lemma 3.9 below. We will see later on examples how to apply this construction to specific models.

To state the lemma, we first need to make a few observations and definitions. Since the Laurent expansions $\iota_{\lambda_a} V_n^a(\lambda)$ each have a non-zero principal part, it follows from the definition (3.24) that we can write

$$\iota_{\lambda_a} U(\lambda) = \sum_{p=-n_a}^{\infty} U_p^a \lambda_a^p \quad (3.26)$$

for some $n_a \geq 1$ and non-zero leading coefficient $U_{-n_a}^a \in \mathfrak{g}$. By definition (3.3) we have that $Q^a(\lambda_a) = \phi^a(\lambda_a) F^a(\lambda_a) \phi^a(\lambda_a)^{-1}$. It thus follows from the relationship between each $\iota_{\lambda_a} V_n^a(\lambda)$ and $Q^a(\lambda_a)$, as described explicitly in Proposition 3.8, that the coefficients of the most singular terms in the formal Laurent series $Q^a(\lambda_a)$ and $\iota_{\lambda_a} U(\lambda)$, given in (3.4) and (3.26) respectively, are proportional. In other words, we have $Q_{-N_a}^a = c U_{-n_a}^a$ for some $c \in \mathbb{C}^*$. Explicitly, it can be seen from Proposition 3.8 that c is given up to a sign by the coefficient r_n^a in (3.24) with $n = \max T_a$. We can thus write

$$\lambda_a^{N_a} Q^a(\lambda_a) = c U_{-n_a}^a + \sum_{r=1}^{\infty} Q_{-N_a+r}^a \lambda_a^r. \quad (3.27)$$

Now let $\mathfrak{k} := \ker(\text{ad } U_{-n_a}^a)$ and $\mathfrak{i} := \text{im}(\text{ad } U_{-n_a}^a)$. We fix any complements $\mathfrak{k}'$ of $\mathfrak{k}$ and $\mathfrak{i}'$ of $\mathfrak{i}$ in $\mathfrak{g}$ so that we have the direct sum decompositions

$$\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{k}' = \mathfrak{i} \oplus \mathfrak{i}'. \quad (3.28)$$

Let $\pi_{\mathfrak{k}} : \mathfrak{g} \rightarrow \mathfrak{k}$ and $\pi_{\mathfrak{k}'} : \mathfrak{g} \rightarrow \mathfrak{k}'$ denote the projections onto $\mathfrak{k}$ and $\mathfrak{k}'$ relative to the first decomposition. Likewise, let $\pi_{\mathfrak{i}} : \mathfrak{g} \rightarrow \mathfrak{i}$ and $\pi_{\mathfrak{i}'} : \mathfrak{g} \rightarrow \mathfrak{i}'$ denote the projections onto $\mathfrak{i}$ and $\mathfrak{i}'$ relative to the second decomposition in (3.28).

Lemma 3.9. *For any $r \geq 1$, $\pi_{\mathfrak{k}'}(Q_{-N_a+r}^a)$ is expressible as a differential polynomial in x of the elements $Q_{-N_a+s}^a$ for $s < r$.*

Proof. Using the explicit forms (3.4) and (3.26) for the Laurent series of $Q^a(\lambda_a)$ and $\iota_{\lambda_a} U(\lambda)$, we may rewrite the component of (3.25) at $a \in S$ more explicitly as

$$\sum_{n=-N_a}^{\infty} \lambda_a^n \partial_x Q_n^a = \sum_{m=-N_a}^{\infty} \sum_{p=-n_a}^{\infty} \lambda_a^{m+p} [U_p^a, Q_m^a] = \sum_{n=-N_a-n_a}^{\infty} \lambda_a^n \sum_{p=-n_a}^{\infty} [U_p^a, Q_{n-p}^a].$$

In the second equality we have changed variables in the double sum from $m \geq -N_a$ to $n := m + p \geq -N_a - n_a$. Comparing the coefficients of $\lambda_a^{-N_a-n_a+r}$ on both sides of the above equation for all $r \geq 0$ we find the following. For every $0 \leq r \leq n_a - 1$,

$$[U_{-n_a}^a, Q_{-N_a+r}^a] = - \sum_{q=1}^r [U_{-n_a+q}^a, Q_{-N_a+r-q}^a] \quad (3.29a)$$

where we changed variables in the sum from p to $q := p + n_a$. Notice that for $r = 0$ this gives $[U_{-n_a}^a, Q_{-N_a}^a] = 0$ which is consistent with the observation in (3.27) that $Q_{-N_a}^a$ is proportional to $U_{-n_a}^a$. On the other hand, for $r \geq n_a$ we have

$$[U_{-n_a}^a, Q_{-N_a+r}^a] = \partial_x Q_{-N_a-n_a+r}^a - \sum_{q=1}^r [U_{-n_a+q}^a, Q_{-N_a+r-q}^a]. \quad (3.29b)$$

Denoting the right hand side of the equations (3.29) by B_r , for each $r \geq 0$ we can rewrite all of them more uniformly as

$$[U_{-n_a}^a, Q_{-N_a+r}^a] = B_r \quad (3.30)$$

for $r \geq 0$. Since the left hand side of (3.30) lies in $\mathfrak{i}$ we have, for every $r \geq 0$,

$$[U_{-n_a}^a, \pi_{\mathfrak{k}'}(Q_{-N_a+r}^a)] = \pi_{\mathfrak{i}}(B_r), \quad 0 = \pi_{\mathfrak{i}'}(B_r), \quad (3.31)$$

where in the first equation we have also decomposed $Q_{-N_a+r}^a$ relative to the first decomposition in (3.28) and used the fact that $\pi_{\mathfrak{k}}(Q_{-N_a+r}^a)$ commutes with $U_{-n_a}^a$.

Now the linear map $\text{ad } U_{-n_a}^a : \mathfrak{k}' \rightarrow \mathfrak{i}$ is a bijection. Indeed, it is clearly surjective by definition of $\mathfrak{i}$. To see that it is injective, note that if $[U_{-n_a}^a, X] = [U_{-n_a}^a, Y]$ for any $X, Y \in \mathfrak{k}'$ then $X - Y \in \mathfrak{k}$ and hence $X - Y = 0$, as required. It follows that $\pi_{\mathfrak{k}'}(Q_{-N_a+r}^a)$ is uniquely determined in terms of $\pi_{\mathfrak{i}}(B_r)$ for every $r \geq 0$ by the first equation in (3.31). The result now follows. $\square$

In order to completely determine the coefficients $Q_{-N_a+r}^a$ for $r \geq 0$, it remains to show that the $\pi_{\mathfrak{k}}(Q_{-N_a+r}^a)$ for every $r \geq 0$ can also be determined recursively. This is the part which will typically depend on the model considered. Here we will show, generalising an argument for the ZS-AKNS $n \times n$ hierarchy given in [TU, Theorem 2.2], see also [Sa], how this can be done under the assumption that there is a polynomial P_a with coefficients in $\mathbb{C}[\lambda_a]$ such that $P_a(\lambda_a^{N_a} F^a(\lambda_a)_{-}) = 0$ and $P'_a(c U_{-n_a}^a) \in \mathbb{C}[\lambda_a]$ is invertible in $\mathbb{C}[[\lambda_a]]$. Recalling (3.27), we have the identity

$$P_a \left(c U_{-n_a}^a + \sum_{r=1}^{\infty} Q_{-N_a+r}^a \lambda_a^r \right) = 0.$$

And using the fact that each $\pi_{\mathfrak{k}}(Q_{-N_a+r}^a)$ commutes with $U_{-n_a}^a$, by definition of $\mathfrak{k}$, we can then rewrite the above in the form

$$P_a(c U_{-n_a}^a) + P'_a(c U_{-n_a}^a) \sum_{r=1}^{\infty} \pi_{\mathfrak{k}}(Q_{-N_a+r}^a) \lambda_a^r = \mathcal{R}\left(\{Q_{-N_a+r}^a\}_{r=1}^{\infty}\right) \quad (3.32)$$

where the right hand side is a sum of terms, each of which contains either higher powers of $\sum_{r=1}^{\infty} \pi_{\mathfrak{k}}(Q_{-N_a+r}^a) \lambda_a^r$ or at least one factor of $\sum_{r=1}^{\infty} \pi_{\mathfrak{k}'}(Q_{-N_a+r}^a) \lambda_a^r$. Since we are assuming that $P'_a(c U_{-n_a}^a) \in \mathbb{C}[\lambda_a]$ is invertible, it follows by comparing powers of λ_a^r on both sides of (3.32) that $\pi_{\mathfrak{k}}(Q_{-N_a+r}^a)$ can be expressed as a finite sum of terms involving only $\pi_{\mathfrak{k}}(Q_{-N_a+s}^a)$ for $s < r$ or $\pi_{\mathfrak{k}'}(Q_{-N_a+s}^a)$ for $s \leq r$.

In conjunction with Lemma 3.9, this shows that each $\pi_{\mathfrak{k}}(Q_{-N_a+r}^a)$ and $\pi_{\mathfrak{k}'}(Q_{-N_a+r}^a)$, and therefore $Q_{-N_a+r}^a$ itself, can be determined recursively for each $r \geq 0$. In particular, all the coefficients Q_n^a , $n \geq -N_a$ of the Laurent series $Q^a(\lambda_a)$ in (3.4) can be expressed as differential polynomials in x of the coefficients of the rational function $U(\lambda)$. The same conclusion still holds even when there is no polynomial P_a with the above properties, as will be shown on the example of the sine-Gordon hierarchy in Sect. 5.

It is important to observe that our choice of ‘spatial’ coordinate x defined by the linear combination $\partial_x = \sum_{a \in S} \sum_{n \in T_a} r_n^a \partial_{t_n^a}$ and its associated Lax matrix in (3.24) was completely arbitrary. Indeed, one of the main advantages of working with the adjoint orbit $\mathcal{Q}(\lambda)$ in $\mathcal{A}_{\lambda}(\mathfrak{g})$ rather than the coadjoint orbit $U(\lambda)$ in $R_{\lambda}(\mathfrak{g})$ is that it keeps all the times on an equal footing by not singling out a particular (linear combination of) time as ‘space’.

On the redundancy of the FNR procedure The previous discussion casts in the present framework the original idea of [FNR] whereby one should first solve for the coordinates in $\mathcal{Q}(\lambda)$ in terms of the finite collection of fields contained in a given Lax matrix $U(\lambda)$, now interpreted as fields depending on a preferred space variable x . The other times in the hierarchies are viewed as (compatible) time flows imposed on this finite collection of fields and define a preferred field theory alongside its higher symmetries.

Here we want to elaborate on a point of view originally advocated in [CS3] whereby the above “traditional” approach is not needed at all and, in fact, represents a conceptual obstruction to the formalism we want to put forward in this work: we treat all the *times* in a hierarchy **as well as** all the (*algebra or group*) *coordinates* (i.e. the dependent variables contained in $\mathcal{Q}(\lambda)$ or $\phi(\lambda)$ respectively) on the same footing. From this point of view, one should consider the entirety of the Lax equations contained in the generating Lax equation (3.8), or equivalently, the collection of zero curvature equations (3.5). The point is that the latter implement the FNR procedure anyway but they present the advantage of being amenable to a covariant Hamiltonian formulation, which was one of the main results of [CS1, CS3]. This aspect is beyond the scope of the present work but remains one motivation for it. The fact that the zero curvature equations contain the equations of the FNR procedure was already observed and used in the particular example of the AKNS hierarchy in [AC]. For convenience, let us sketch the argument here in the simplest case of a single pole $a \in \mathbb{C}$, with a collection of times t_n^a , $n \geq -N_a$. Suppose we fix $n \geq -N_a$ and we want to solve

$$\partial_{t_n^a} Q^a(\lambda_a) = [\iota_{\lambda_a} V_n^a(\lambda), Q^a(\lambda_a)], \quad (3.33)$$

given $Q_{-N_a}^a$, along the lines of Lemma 3.9 and the discussion after it. Without loss of generality, shifting the power of λ_a by N_a , we can always assume for simplicity that

$N_a = 0$. Then, (3.33) amounts to the collection of equations

$$\partial_{t_n^a} Q_j^a = \sum_{p=0}^n [Q_{j+n-p+1}^a, Q_p^a], \quad j \geq 0. \quad (3.34)$$

As discussed above, in certain cases (which include the AKNS hierarchy and the sG hierarchy as we show explicitly in Sect. 5), this allows one to express all the algebra coordinates in Q_j , $j \geq n$ as differential polynomials with respect to t_n^a in the coordinates contained in Q_k , $k = 0, \dots, n$. Now consider the zero curvature equations, for $m \geq n+1$,

$$\partial_{t_n^a} V_m^a(\lambda) - \partial_{t_m^a} V_n^a(\lambda) + [V_n^a(\lambda), V_m^a(\lambda)] = 0. \quad (3.35)$$

Looking at the coefficient of $1/\lambda^j$, for $j = n+2, \dots, m+1$, we find that they contain the equations

$$\partial_{t_n^a} Q_{m+1-k}^a = \sum_{p=0}^n [Q_{m+n+2-k-p}^a, Q_p^a], \quad k = n+2, \dots, m+1. \quad (3.36)$$

If we set $j = m+1-k$, these become

$$\partial_{t_n^a} Q_j^a = \sum_{p=0}^n [Q_{j+n+1-p}^a, Q_p^a], \quad j = 0, \dots, m-n-1. \quad (3.37)$$

So the collection of zero curvature equations (3.35) for $m \geq n+1$ produces exactly the set of FNR equations (3.34). Hence, there is no point in implementing the FNR procedure a priori to determine the “fields” and then impose the zero curvature equations to determine their equations of motion. The latter suffices. With this in mind, we will come back to this point in certain examples below to illustrate our position and show how abandoning the FNR procedure allows us to eliminate the problem of alien derivatives mentioned in the introduction.

3.2. Generating Lagrangian multiform. In this section, we introduce the main object of this paper, the generating Lagrangian multiform (1.15)–(3.40), and we show that the Lax equation (3.8) as it derives from $\mathcal{L}(\lambda, \mu)$. Although the equations of motion (3.8) can be written in terms of $Q(\lambda) \in \mathcal{A}_\lambda(\mathfrak{g})$ alone, in order to write $\mathcal{L}(\lambda, \mu)$ we need the group-valued element $\phi(\lambda) \in \mathcal{A}_\lambda^+(G)$. This is very reminiscent of the fact that writing down the Zakharov–Mikhailov action describing the Zakharov–Shabat equations of motion requires introducing a group valued field [ZM1]. Recall the definition of $Q(\lambda) \in \mathcal{A}_\lambda(\mathfrak{g})$ in (3.3) as an adjoint orbit of the element $(\iota_\lambda F(\lambda))_- \in \mathcal{A}_\lambda(\mathfrak{g})$, defined in (3.2), under the action of $\phi(\lambda) \in \mathcal{A}_\lambda^+(G)$.

We consider the following generating Lagrangian multiform

$$\mathcal{L}(\lambda, \mu) := K(\lambda, \mu) - U(\lambda, \mu) \quad (3.38)$$

where the kinetic and potential terms are given by

$$K(\lambda, \mu) := \text{Tr}(\phi(\lambda)^{-1} \mathcal{D}_\mu \phi(\lambda) (\iota_\lambda F(\lambda))_-) - \text{Tr}(\phi(\mu)^{-1} \mathcal{D}_\lambda \phi(\mu) (\iota_\mu F(\mu))_-), \quad (3.39a)$$

$$U(\lambda, \mu) := \frac{1}{2} \text{Tr}_{12}((\iota_\lambda \iota_\mu + \iota_\mu \iota_\lambda) r_{12}(\lambda, \mu) Q_1(\lambda) Q_2(\mu)). \quad (3.39b)$$

As mentioned at the end of Sect. 2.1, the boldface notation (1.15) is used as a shorthand for an equality of components

$$\mathcal{L}^{a,b}(\lambda_a, \mu_b) = K^{a,b}(\lambda_a, \mu_b) - U^{a,b}(\lambda_a, \mu_b)$$

for every $a, b \in \mathbb{C}P^1$, and the kinetic and potential terms (1.16) in components are given explicitly by

$$\begin{aligned} K^{a,b}(\lambda_a, \mu_b) &= \text{Tr} \left(\phi^a(\lambda_a)^{-1} \mathcal{D}_{\mu_b} \phi^a(\lambda_a) F^a(\lambda_a) - \right. \\ &\quad \left. - \text{Tr} \left(\phi^b(\mu_b)^{-1} \mathcal{D}_{\lambda_a} \phi^b(\mu_b) F^b(\mu_b) - \right), \right. \end{aligned} \quad (3.40a)$$

$$U^{a,b}(\lambda_a, \mu_b) = \frac{1}{2} \text{Tr}_{12} \left((\iota_{\lambda_a} \iota_{\mu_b} + \iota_{\mu_b} \iota_{\lambda_a}) r_{12}(\lambda, \mu) Q_1^a(\lambda_a) Q_2^b(\mu_b) \right). \quad (3.40b)$$

The kinetic term (3.39a) is clearly skew-symmetric under the exchange $\lambda \leftrightarrow \mu$, so the skew-symmetry of $\mathcal{L}(\lambda, \mu)$ is equivalent to the skew-symmetry of the potential term (1.16b), namely

$$\begin{aligned} &\text{Tr}_{12} \left((\iota_{\lambda} \iota_{\mu} + \iota_{\mu} \iota_{\lambda}) r_{12}(\lambda, \mu) Q_1(\lambda) Q_2(\mu) \right) \\ &= - \text{Tr}_{12} \left((\iota_{\lambda} \iota_{\mu} + \iota_{\mu} \iota_{\lambda}) r_{21}(\mu, \lambda) Q_1(\lambda) Q_2(\mu) \right). \end{aligned}$$

This holds since r is skew-symmetric.

3.2.1. Extracting Lagrangians and Lagrangian multiforms We have been using the generating formalism efficiently so far. Here, we spend some time discussing the connection of our generating Lagrangian multiform with Lagrangians and Lagrangian multiforms. This will be useful to reformulate the multiform EL equations and the closure relation in generating form, allowing to continue to take advantage of this for general computations.

From the definition of the generating Lagrangian multiform (1.15), we see that the kinetic term $K^{a,b}(\lambda_a, \mu_b)$ given by (3.40a) is a Laurent series in both λ_a and μ_b , with powers bounded below by $-N_a$ and $-N_b$, respectively. In particular, for any $m, n \in \mathbb{Z}$ the coefficient of $\lambda_a^m \mu_b^n$ is well defined. The same is true for the potential term (3.40b) by the following lemma.

Lemma 3.10. *For any $m, n \in \mathbb{Z}$ and any $a, b \in \mathbb{C}P^1$, the coefficient of $\lambda_a^m \mu_b^n$ in the potential term $U^{a,b}(\lambda_a, \mu_b)$ given by (3.40b) is a well defined expression which is quadratic in the coefficients of $Q^a(\lambda_a)$ and $Q^b(\mu_b)$.*

Proof. If $b \neq a$ then $(\iota_{\lambda_a} \iota_{\mu_b} + \iota_{\mu_b} \iota_{\lambda_a}) r_{12}(\lambda, \mu)$ is valued in $(\mathfrak{g} \otimes \mathfrak{g}) \otimes \mathbb{C}[[\lambda_a, \mu_b]]$. Since by definition (3.4) we have $Q^a(\lambda_a) \in \mathfrak{g} \otimes \lambda_a^{-N_a} \mathbb{C}[[\lambda_a]]$ and $Q^b(\mu_b) \in \mathfrak{g} \otimes \mu_b^{-N_b} \mathbb{C}[[\mu_b]]$, it follows that $U^{a,b}(\lambda_a, \mu_b)$ is a Laurent series in both λ_a and μ_b , with powers bounded below by $-N_a$ and $-N_b$, respectively.

If $b = a$ then $(\iota_{\lambda_a} \iota_{\mu_a} + \iota_{\mu_a} \iota_{\lambda_a}) r_{12}(\lambda, \mu)$ contains a doubly infinite Laurent series in $\lambda_a \mu_a^{-1}$ coming from the expansion of $1/(\lambda - \mu)$, possibly also multiplied by some polynomial in λ_a and μ_a depending on the precise form of the r -matrix. Multiplying this by the Laurent series $Q^a(\lambda_a) \in \mathfrak{g} \otimes \lambda_a^{-N_a} \mathbb{C}[[\lambda_a]]$ and $Q^a(\mu_a) \in \mathfrak{g} \otimes \mu_a^{-N_a} \mathbb{C}[[\mu_a]]$, we produce terms of the form $\lambda_a^{r+j+p} \mu_a^{s-j+q}$ with $r, s \geq -N_a$, $j \in \mathbb{Z}$ and p, q ranging over finitely many possible values. In order to form a term proportional to $\lambda_a^m \mu_a^n$ we need $m = r + j + p$ and $n = s - j + q$. But then $m - j - p = r \geq -N_a$ so that $j \leq m + N_a - p$ and also $n + j - q = s \geq -N_a$ so that $j \geq -n - N_a + q$. In other words, $j \in \mathbb{Z}$ must be bounded from above and below so that it ranges only over finitely many values. Hence, there are only finitely many terms contributing to the coefficient of $\lambda_a^m \mu_a^n$ and the result follows. $\square$

As a consequence, for any $a, b \in \mathbb{C}P^1$ and $m, n \in \mathbb{Z}$ with $m \geq -N_a$ and $n \geq -N_b$, we may now extract the following Lagrangian coefficients associated to the times t_n^a and t_m^b :

Definition 3.11 (*Elementary Lagrangians*).

$$\mathcal{L}_{m,n}^{a,b} := \text{res}_a^\lambda \text{res}_b^\mu \mathcal{L}^{a,b}(\lambda_a, \mu_b) \lambda^{-m-1} d\lambda \mu^{-n-1} d\mu. \quad (3.41)$$

Recall the notational convention explained after (2.2b), in particular for residues computed at infinity. In short, Definition (3.41) means that $\mathcal{L}_{m,n}^{a,b}$ is the coefficient of $\lambda_a^m \mu_b^n$ in the expansion of $\mathcal{L}^{a,b}(\lambda_a, \mu_b)$, as one would want. This is what we use to compute elementary Lagrangians in all our examples.

As explained below, when building a hierarchy, one chooses a finite set $S \in \mathbb{C}P^1$ and all but a finite number of the elementary Lagrangians $\mathcal{L}_{m,n}^{a,b}$ vanish (those for which a and/or b is in $\mathbb{C}P^1 \setminus S$). The Lagrangian multiform of the hierarchy is then given by

$$\mathcal{L}^S := \frac{1}{2} \sum_{a,b \in S} \sum_{m,n} \mathcal{L}_{m,n}^{a,b} dt_m^a \wedge dt_n^b = \sum_{(m,a) < (n,b)} \mathcal{L}_{m,n}^{a,b} dt_m^a \wedge dt_n^b. \quad (3.42)$$

Note that we introduced an order on the pairs $(m, a) \in \mathbb{Z} \times S$ in the last equality (recall that $\mathcal{L}_{m,n}^{a,b} = -\mathcal{L}_{n,m}^{b,a}$). With $S = \{a_1, \dots, a_n\}$, it is defined by

$$(m, a_i) < (n, a_j) \Leftrightarrow i < j \text{ or } (i = j \text{ and } m < n).$$

These definitions generalise the correspondence explained in the introductory section 1.1.3 between $\mathcal{L}[u]$ and $\mathcal{L}(\lambda, \mu)$ for the AKNS hierarchy. As we will see in detail in Sect. 4, the latter indeed corresponds to the case where $S = \{\infty\}$. In practice, one calculates the elementary Lagrangians (3.41) directly by computing the appropriate Laurent series expansion of $\mathcal{L}^{a,b}(\lambda_a, \mu_b)$. The corresponding Lagrangian multiform is easily obtained as in (3.42).

The essential point of the present discussion is to identify the generating form of the two main equations of the theory of Lagrangian multiforms: the multiform EL equations $\delta d\mathcal{L}^S = 0$ and the closure relation $d\mathcal{L}^S = 0$ which should hold on solutions of the multiform EL equations. We see that the key object to translate in generating form is therefore $d\mathcal{L}^S$. In view of (3.42), $d\mathcal{L}^S$ has the form

$$d\mathcal{L}^S = \sum_{(k,c) < (m,a) < (n,b)} \left(\partial_{t_k^c} \mathcal{L}_{m,n}^{a,b} + \partial_{t_n^b} \mathcal{L}_{k,m}^{c,a} + \partial_{t_m^a} \mathcal{L}_{n,k}^{b,c} \right) dt_k^c \wedge dt_m^a \wedge dt_n^b.$$

The generating function corresponding to the coefficient $\partial_{t_k^c} \mathcal{L}_{m,n}^{a,b} + \partial_{t_n^b} \mathcal{L}_{k,m}^{c,a} + \partial_{t_m^a} \mathcal{L}_{n,k}^{b,c}$ is

$$\mathcal{D}_{v_c} \mathcal{L}^{a,b}(\lambda_a, \mu_b) + \mathcal{D}_{\mu_b} \mathcal{L}^{c,a}(v_c, \lambda_a) + \mathcal{D}_{\lambda_a} \mathcal{L}^{b,c}(\mu_b, v_c).$$

Summarizing our discussion, the set S was fixed but arbitrary, so going back to the adélic setting, we will be working compactly with

$$\delta \mathcal{D}_v \mathcal{L}(\lambda, \mu) + \delta \mathcal{D}_\mu \mathcal{L}(v, \lambda) + \delta \mathcal{D}_\lambda \mathcal{L}(\mu, v)$$

when deriving the multiform EL equations in generating form, and with

$$\mathcal{D}_v \mathcal{L}(\lambda, \mu) + \mathcal{D}_\mu \mathcal{L}(v, \lambda) + \mathcal{D}_\lambda \mathcal{L}(\mu, v)$$

when studying the closure relation.

3.2.2. Generating multiform Euler–Lagrange equations Having introduced the main object of our framework, we proceed to derive the associated multiform EL equations (in generating form) and show that they give the generating Lax equation (3.8).

Theorem 3.12. *The generating Lax equation (3.8) is variational: the multiform EL equations deriving from the generating Lagrangian multiform $\mathcal{L}(\lambda, \mu)$ take the form*

$$\mathcal{D}_\mu \mathcal{Q}_1(\lambda) = [\text{Tr}_2 (\iota_\lambda \iota_\mu r_{12}(\lambda, \mu) \mathcal{Q}_2(\mu)), \mathcal{Q}_1(\lambda)].$$

Proof. We derive the equations induced by the requirement $\delta d\mathcal{L} = 0$ in generating form. This means that we compute $\delta \mathcal{D}_\nu \mathcal{L}(\lambda, \mu) + \circlearrowleft = 0$, where $\circlearrowleft$ means cyclic permutations of λ, μ, ν , and set the independent coefficients to zero. We start with the kinetic terms.

$$\begin{aligned} \mathcal{D}_\nu \mathcal{K}(\lambda, \mu) = & \text{Tr} \left(-\phi^{-1}(\lambda) \mathcal{D}_\nu \phi(\lambda) \phi^{-1}(\lambda) \mathcal{D}_\mu \phi(\lambda) (\iota_\lambda F(\lambda))_- + \phi^{-1}(\lambda) \mathcal{D}_\nu \mathcal{D}_\mu \phi(\lambda) (\iota_\lambda F(\lambda))_- \right. \\ & \left. + \phi^{-1}(\mu) \mathcal{D}_\nu \phi(\mu) \phi^{-1}(\mu) \mathcal{D}_\lambda \phi(\mu) (\iota_\mu F(\mu))_- - \phi^{-1}(\mu) \mathcal{D}_\nu \mathcal{D}_\lambda \phi(\mu) (\iota_\mu F(\mu))_- \right) \end{aligned}$$

so that $\mathcal{D}_\nu \mathcal{K}(\lambda, \mu) + \circlearrowleft$ is equal to

$$\text{Tr} \left([\phi^{-1}(\lambda) \mathcal{D}_\mu \phi(\lambda), \phi^{-1}(\lambda) \mathcal{D}_\nu \phi(\lambda)] (\iota_\lambda F(\lambda))_- \right) + \circlearrowleft. \quad (3.43)$$

After we apply the δ differential we get

$$\begin{aligned} \delta \mathcal{D}_\nu \mathcal{K}(\lambda, \mu) + \circlearrowleft = & \text{Tr} \left(\mathcal{D}_\nu \phi(\lambda) \phi^{-1}(\lambda) \mathcal{D}_\mu \mathcal{Q}(\lambda) - \mathcal{D}_\mu \phi(\lambda) \phi^{-1}(\lambda) \mathcal{D}_\nu \mathcal{Q}(\lambda) \right) \delta \phi(\lambda) \phi^{-1}(\lambda) \\ & + \text{Tr} \left(\phi^{-1}(\lambda) \mathcal{D}_\nu \mathcal{Q}(\lambda) \delta \mathcal{D}_\mu \phi(\lambda) - \phi^{-1}(\lambda) \mathcal{D}_\mu \mathcal{Q}(\lambda) \delta \mathcal{D}_\nu \phi(\lambda) \right) + \circlearrowleft. \end{aligned}$$

We now turn to the the potential term

$$U(\lambda, \mu) = \frac{1}{2} \text{Tr}_{12} ((\iota_\lambda \iota_\mu + \iota_\mu \iota_\lambda) r_{12}(\lambda, \mu) \mathcal{Q}_1(\lambda) \mathcal{Q}_2(\mu)). \quad (3.44)$$

We drop λ and μ in ϕ and $\mathcal{Q}$ for conciseness since they follow the spaces 1 and 2 consistently. Let us also denote $(\iota_\lambda \iota_\mu + \iota_\mu \iota_\lambda) r_{12}(\lambda, \mu)$ by r_{12} . We compute

$$\mathcal{D}_\nu U(\lambda, \mu) = \frac{1}{2} \text{Tr}_{12} (r_{12} (\mathcal{D}_\nu \mathcal{Q}_1 \mathcal{Q}_2 + \mathcal{Q}_1 \mathcal{D}_\nu \mathcal{Q}_2)) \quad (3.45)$$

and after applying the δ -differential we get

$$\delta \mathcal{D}_\nu U(\lambda, \mu) = \frac{1}{2} \text{Tr}_{12} (r_{12} (\delta \mathcal{D}_\nu \mathcal{Q}_1 \mathcal{Q}_2 + \mathcal{D}_\nu \mathcal{Q}_1 \delta \mathcal{Q}_2 + \delta \mathcal{Q}_1 \mathcal{D}_\nu \mathcal{Q}_2 + \mathcal{Q}_1 \delta \mathcal{D}_\nu \mathcal{Q}_2)) \quad (3.46)$$

and similarly for the cyclic permutations. We use the following identities

$$\begin{aligned} \text{Tr}_{12} r_{12} \delta \mathcal{D}_\nu \mathcal{Q}_1 \mathcal{Q}_2 &= \text{Tr}_{12} (-\mathcal{Q}_2 r_{12} \mathcal{D}_\nu \mathcal{Q}_1 - \mathcal{Q}_1 \mathcal{D}_\nu \phi_1 \phi_1^{-1} \mathcal{Q}_2 r_{12} \\ &\quad + \mathcal{D}_\nu \phi_1 \phi_1^{-1} \mathcal{Q}_2 r_{12} \mathcal{Q}_1 + \phi_1 \mathcal{D}_\nu X_1 \phi_1^{-1} \mathcal{Q}_2 r_{12}) \delta \phi_1 \phi_1^{-1} \\ &\quad + \text{Tr}_{12} [\mathcal{Q}_1, r_{12} \mathcal{Q}_2] \delta \mathcal{D}_\nu \phi_1 \phi_1^{-1} \\ &= \text{Tr}_{12} ([\mathcal{D}_\nu \mathcal{Q}_1, r_{12} \mathcal{Q}_2] - \mathcal{D}_\nu \phi_1 \phi_1^{-1} [\mathcal{Q}_1, r_{12} \mathcal{Q}_2]) \delta \phi_1 \phi_1^{-1} \\ &\quad + \text{Tr}_{12} [\mathcal{Q}_1, r_{12} \mathcal{Q}_2] \delta \mathcal{D}_\nu \phi_1 \phi_1^{-1}, \\ \text{Tr}_{12} r_{12} \mathcal{Q}_1 \delta \mathcal{D}_\nu \mathcal{Q}_2 &= \text{Tr}_{12} (-r_{12} \mathcal{Q}_1 \mathcal{D}_\nu \mathcal{Q}_2 - \mathcal{Q}_2 \mathcal{D}_\nu \phi_2 \phi_2^{-1} r_{12} \mathcal{Q}_1 \\ &\quad + \mathcal{D}_\nu \phi_2 \phi_2^{-1} r_{12} \mathcal{Q}_1 \mathcal{Q}_2 + \phi_2 \mathcal{D}_\nu X_2 \phi_2^{-1} r_{12} \mathcal{Q}_1) \delta \phi_2 \phi_2^{-1} \\ &\quad + \text{Tr}_{12} [\mathcal{Q}_2, r_{12} \mathcal{Q}_1] \delta \mathcal{D}_\nu \phi_2 \phi_2^{-1} \\ &= \text{Tr}_{12} ([\mathcal{D}_\nu \mathcal{Q}_2, r_{12} \mathcal{Q}_1] - \mathcal{D}_\nu \phi_2 \phi_2^{-1} [\mathcal{Q}_2, r_{12} \mathcal{Q}_1]) \delta \phi_2 \phi_2^{-1} \\ &\quad + \text{Tr}_{12} [\mathcal{Q}_2, r_{12} \mathcal{Q}_1] \delta \mathcal{D}_\nu \phi_2 \phi_2^{-1}, \end{aligned}$$

and

$$\begin{aligned}\mathrm{Tr}_{12} \, r_{12} \delta \mathcal{Q}_1 \mathcal{D}_v \mathcal{Q}_2 &= \mathrm{Tr}_{12} [\mathcal{Q}_1, r_{12} \mathcal{D}_v \mathcal{Q}_2] \delta \phi_1 \phi_1^{-1}, \\ \mathrm{Tr}_{12} \, r_{12} \mathcal{D}_v \mathcal{Q}_1 \delta \mathcal{Q}_2 &= \mathrm{Tr}_{12} [\mathcal{Q}_2, r_{12} \mathcal{D}_v \mathcal{Q}_1] \delta \phi_2 \phi_2^{-1},\end{aligned}$$

to express $\delta \mathcal{D}_v U(\lambda, \mu)$ on the basis of $\delta \phi_1$, $\delta \mathcal{D}_\mu \phi_1$ and $\delta \mathcal{D}_v \phi_1$ (and similarly on the space 2). Then, we collect the coefficients of $\delta \phi_1$, $\delta \mathcal{D}_\mu \phi_1$ and $\delta \mathcal{D}_v \phi_1$ which provide the independent equations. From $\delta \mathcal{D}_v K(\lambda, \mu) + \circlearrowright$ we have

$$\begin{aligned}\mathrm{Tr}_1 (-\mathcal{D}_\mu \phi_1 \phi_1^{-1} \mathcal{D}_v \mathcal{Q}_1 + \mathcal{D}_v \phi_1 \phi_1^{-1} \mathcal{D}_\mu \mathcal{Q}_1) \delta \phi_1 \phi_1^{-1} + \mathrm{Tr}_1 (-\mathcal{D}_\mu \delta \mathcal{D}_v \phi_1 \phi_1^{-1} \\ + \mathcal{D}_v \mathcal{Q}_1 \delta \mathcal{D}_\mu \phi_1 \phi_1^{-1}) + \circlearrowright\end{aligned}\quad (3.47)$$

and from $\delta \mathcal{D}_v U(\lambda, \mu) + \circlearrowright$, using the skew-symmetry of r , we obtain

$$\begin{aligned}\mathrm{Tr}_{12} [\mathcal{Q}_1, r_{12} \mathcal{Q}_2] \delta \mathcal{D}_v \phi_1 \phi_1^{-1} - \mathrm{Tr}_{13} [\mathcal{Q}_1, r_{13} \mathcal{Q}_3] \delta \mathcal{D}_\mu \phi_1 \phi_1^{-1} \\ + \mathrm{Tr}_{12} ([\mathcal{D}_v \mathcal{Q}_1, r_{12} \mathcal{Q}_2] - \mathcal{D}_v \phi_1 \phi_1^{-1} [\mathcal{Q}_1, r_{12} \mathcal{Q}_2] + [\mathcal{Q}_1, r_{12} \mathcal{D}_v \mathcal{Q}_2]) \delta \phi_1 \phi_1^{-1} \\ + \mathrm{Tr}_{13} (-[\mathcal{D}_\mu \mathcal{Q}_1, r_{13} \mathcal{Q}_3] + \mathcal{D}_\mu \phi_1 \phi_1^{-1} [\mathcal{Q}_1, r_{13} \mathcal{Q}_3] - [\mathcal{Q}_1, r_{13} \mathcal{D}_\mu \mathcal{Q}_3]) \delta \phi_1 \phi_1^{-1} + \circlearrowright.\end{aligned}\quad (3.48)$$

The coefficients of $\delta \mathcal{D}_\mu \phi_1$ and $\delta \mathcal{D}_v \phi_1$ in $\delta \mathcal{D}_v \mathcal{L}(\lambda, \mu) + \circlearrowright = 0$ give

$$\mathcal{D}_\mu \mathcal{Q}_1 = \frac{1}{2} [\mathrm{Tr}_2 \, r_{12} \mathcal{Q}_2, \mathcal{Q}_1], \quad \mathcal{D}_v \mathcal{Q}_1 = \frac{1}{2} [\mathrm{Tr}_3 \, r_{13} \mathcal{Q}_3, \mathcal{Q}_1],$$

i.e. two equivalent copies of the same equation under the irrelevant change $2 \leftrightarrow 3$ and $\mu \leftrightarrow v$. Explicitly, it reads

$$\mathcal{D}_\mu \mathcal{Q}_1(\lambda) = \frac{1}{2} \left[\mathrm{Tr}_2 \left((\iota_\lambda \iota_\mu + \iota_\mu \iota_\lambda) r_{12}(\lambda, \mu) \mathcal{Q}_2(\mu) \right), \mathcal{Q}_1(\lambda) \right], \quad (3.49)$$

which gives the desired result (3.8) upon recalling Lemma 3.2. The coefficient of $\delta \phi_1$ is just a consequence of this equation and of the commutativity of the flows: $[\mathcal{D}_\mu, \mathcal{D}_v] = 0$. The coefficients of $\delta \phi_2$, $\delta \phi_3$ etc. contained in the cyclic permutations $\circlearrowright$ give equivalent equations under the corresponding cyclic permutations of the spectral parameters and auxiliary spaces.

□

3.2.3. Generating closure relation

Theorem 3.13. *The generating closure relation*

$$\mathcal{D}_v \mathcal{L}(\lambda, \mu) + \mathcal{D}_\lambda \mathcal{L}(\mu, v) + \mathcal{D}_\mu \mathcal{L}(v, \lambda) = 0. \quad (3.50)$$

holds when (3.8) is satisfied. It is a consequence of the CYBE for r .

Proof. First consider the kinetic term (3.39a). We have

$$\begin{aligned}\mathcal{D}_v K(\lambda, \mu) &= \mathrm{Tr} \left(\phi(\lambda)^{-1} \mathcal{D}_v \mathcal{D}_\mu \phi(\lambda) (\iota_\lambda F(\lambda))_- \right) \\ &\quad - \mathrm{Tr} \left(\phi(\mu)^{-1} \mathcal{D}_v \mathcal{D}_\lambda \phi(\mu) (\iota_\mu F(\mu))_- \right) \\ &\quad - \mathrm{Tr} \left(\phi(\lambda)^{-1} \mathcal{D}_v \phi(\lambda) \phi(\lambda)^{-1} \mathcal{D}_\mu \phi(\lambda) (\iota_\lambda F(\lambda))_- \right) \\ &\quad + \mathrm{Tr} \left(\phi(\mu)^{-1} \mathcal{D}_v \phi(\mu) \phi(\mu)^{-1} \mathcal{D}_\lambda \phi(\mu) (\iota_\mu F(\mu))_- \right).\end{aligned}$$

It follows by adding the cyclic permutations of this expression in the variables λ , μ and ν that

$$\mathcal{D}_\nu K(\lambda, \mu) + \mathcal{D}_\lambda K(\mu, \nu) + \mathcal{D}_\mu K(\nu, \lambda) = 0. \quad (3.51)$$

Consider now the potential term (1.16b). Using Theorem 3.12 we find

$$\begin{aligned} \mathcal{D}_\nu U(\lambda, \mu) &= \frac{1}{2} \text{Tr}_{12} \left((\iota_\lambda \iota_\mu + \iota_\mu \iota_\lambda) r_{12}(\lambda, \mu) \mathcal{D}_\nu Q_1(\lambda) Q_2(\mu) \right) \\ &\quad + \frac{1}{2} \text{Tr}_{12} \left((\iota_\lambda \iota_\mu + \iota_\mu \iota_\lambda) r_{12}(\lambda, \mu) Q_1(\lambda) \mathcal{D}_\nu Q_2(\mu) \right) \\ &= \frac{1}{2} \text{Tr}_{123} \left((\iota_\lambda \iota_\mu + \iota_\mu \iota_\lambda) r_{12}(\lambda, \mu) [\iota_\lambda \iota_\nu r_{13}(\lambda, \nu) Q_3(\nu), Q_1(\lambda)] Q_2(\mu) \right) \\ &\quad + \frac{1}{2} \text{Tr}_{123} \left((\iota_\lambda \iota_\mu + \iota_\mu \iota_\lambda) r_{12}(\lambda, \mu) Q_1(\lambda) [\iota_\mu \iota_\nu r_{23}(\mu, \nu) Q_3(\nu), Q_2(\mu)] \right). \end{aligned}$$

By using the cyclicity of the trace in space 1 and 2 in the first and second terms, respectively, we may write this as

$$\begin{aligned} \mathcal{D}_\nu U(\lambda, \mu) &= -\frac{1}{2} \text{Tr}_{123} \left([(\iota_\lambda \iota_\mu + \iota_\mu \iota_\lambda) r_{12}(\lambda, \mu) Q_2(\mu), Q_1(\lambda)] \iota_\lambda \iota_\nu r_{13}(\lambda, \nu) Q_3(\nu) \right) \\ &\quad - \frac{1}{2} \text{Tr}_{123} \left([(\iota_\lambda \iota_\mu + \iota_\mu \iota_\lambda) r_{12}(\lambda, \mu) Q_1(\lambda), Q_2(\mu)] \iota_\mu \iota_\nu r_{23}(\mu, \nu) Q_3(\nu) \right) \\ &= -\text{Tr}_{123} \left([\iota_\lambda \iota_\mu r_{12}(\lambda, \mu) Q_2(\mu), Q_1(\lambda)] \iota_\lambda \iota_\nu r_{13}(\lambda, \nu) Q_3(\nu) \right) \\ &\quad - \text{Tr}_{123} \left([\iota_\lambda \iota_\mu r_{12}(\lambda, \mu) Q_1(\lambda), Q_2(\mu)] \iota_\mu \iota_\nu r_{23}(\mu, \nu) Q_3(\nu) \right) \end{aligned}$$

where in the second equality we used Lemma 3.2 in both terms. By using once again the cyclicity of the trace in space 1 and 2 in the first and second terms, respectively, we arrive at the expression

$$\begin{aligned} \mathcal{D}_\nu U(\lambda, \mu) &= \text{Tr}_{123} \left(\iota_\lambda \iota_\mu \iota_\nu ([r_{12}(\lambda, \mu), r_{13}(\lambda, \nu)] \right. \\ &\quad \left. + [r_{12}(\lambda, \mu), r_{23}(\mu, \nu)]) Q_1(\lambda) Q_2(\mu) Q_3(\nu) \right). \end{aligned} \quad (3.52a)$$

Likewise, using the skew-symmetry of the r -matrix we find

$$\begin{aligned} \mathcal{D}_\lambda U(\mu, \nu) &= -\frac{1}{2} \text{Tr}_{123} \left((\iota_\mu \iota_\nu + \iota_\nu \iota_\mu) r_{23}(\mu, \nu) [\iota_\mu \iota_\lambda r_{12}(\lambda, \mu) Q_1(\lambda), Q_2(\mu)] Q_3(\nu) \right) \\ &\quad - \frac{1}{2} \text{Tr}_{123} \left((\iota_\mu \iota_\nu + \iota_\nu \iota_\mu) r_{23}(\mu, \nu) Q_2(\mu) [\iota_\nu \iota_\lambda r_{13}(\lambda, \nu) Q_1(\lambda), Q_3(\nu)] \right). \end{aligned}$$

Then by following the same steps as above for $\mathcal{D}_\nu U(\lambda, \mu)$ we deduce that

$$\begin{aligned} \mathcal{D}_\lambda U(\mu, \nu) &= \text{Tr}_{123} \left(\iota_\lambda \iota_\mu \iota_\nu ([r_{12}(\lambda, \mu), r_{23}(\mu, \nu)] \right. \\ &\quad \left. + [r_{13}(\lambda, \nu), r_{23}(\mu, \nu)]) Q_1(\lambda) Q_2(\mu) Q_3(\nu) \right). \end{aligned} \quad (3.52b)$$

Similarly, we also find using the skew-symmetry of the r -matrix that

$$\begin{aligned} \mathcal{D}_\mu U(\nu, \lambda) &= \text{Tr}_{123} \left(\iota_\lambda \iota_\mu \iota_\nu ([r_{13}(\lambda, \nu), r_{23}(\mu, \nu)] \right. \\ &\quad \left. + [r_{12}(\lambda, \mu), r_{13}(\lambda, \nu)]) Q_1(\lambda) Q_2(\mu) Q_3(\nu) \right). \end{aligned} \quad (3.52c)$$

It now follows from combining the three equations in (3.52) and using the classical Yang–Baxter equation for the skew-symmetry r -matrix that

$$\mathcal{D}_\nu U(\lambda, \mu) + \mathcal{D}_\lambda U(\mu, \nu) + \mathcal{D}_\mu U(\nu, \lambda) = 0. \quad (3.53)$$

The result now follows from (3.51) and (3.53) but together. $\square$

The rest of the paper is devoted to examples. To specify an example, the following ingredients need to be fixed:

- (i) a skew-symmetric r -matrix as in Sect. 2 (rational or trigonometric in this work),
- (ii) an effective divisor $\mathcal{D} := \sum_{a \in S} N_a a$, in particular with support given by a finite subset $S \subset \mathbb{CP}^1$ and with $N_a \in \mathbb{Z}_{\geq 1}$ for each $a \in S$, ($N_\infty \in \mathbb{Z}_{\geq 0}$ if $\infty \in S$),
- (iii) a Lie algebra $\mathfrak{g}$ which for simplicity we take to be either $\mathfrak{gl}_N$ or $\mathfrak{sl}_N$,
- (iv) a $\mathfrak{g}$ -valued rational function $F(\lambda) \in R_\lambda(\mathfrak{g})$ with pole divisor $(F)_\infty = \mathcal{D}$, i.e. with a pole of order $N_a \in \mathbb{Z}_{\geq 1}$ at each point $a \in S$, ($N_\infty \in \mathbb{Z}_{\geq 0}$ if $\infty \in S$).

Each section contains an example of a hierarchy for which the above formalism produces Lagrangian multiforms, Lax matrices and zero curvature equations. Some sections consist of known examples that we recover or cast in a new light, e.g. AKNS and sine-Gordon. Other examples are new to the best of our knowledge and show the power of the formalism, e.g. the trigonometric Zakharov–Mikhailov class of models or the examples where we couple different integrable field theories together.

4. AKNS Hierarchy

We keep this section short as it is a matter of “closing the loop”: we reproduce the motivating example of Sect. 1.1.3 which was dealt with in detail in [CS3] and the starting point of this whole project. The main objective is to illustrate how to use our machinery on the simplest and most well known example. We choose the rational r -matrix and we fix the required data as follows:

$$S = \{\infty\}, \quad N_\infty = 0, \quad \mathfrak{g} = \mathfrak{sl}_2, \quad F(\lambda) = -i\sigma_3. \quad (4.1)$$

The adjoint orbit description of Sect. 3.1 is implemented with

$$\phi^\infty(\lambda_\infty) = \mathbf{1} + \sum_{n=1}^{\infty} \phi_n^\infty \lambda_\infty^n, \quad (4.2)$$

and gives

$$Q^\infty(\lambda_\infty) = \sum_{n=0}^{\infty} Q_n^\infty \lambda_\infty^n, \quad (4.3)$$

with $Q_0^\infty = -i\sigma_3$ and $Q_1^\infty = i[\sigma_3, \phi_1^\infty]$, the familiar first two elements in the AKNS hierarchy. Since there is only one pole in this example, let us drop the subscripts and superscripts and simply write the fundamental objects in (4.2) and (4.3) as

$$\phi(\lambda) = \mathbf{1} + \sum_{n=1}^{\infty} \phi_n \lambda^{-n}, \quad Q(\lambda) = \sum_{n=0}^{\infty} Q_n \lambda^{-n}. \quad (4.4)$$

Similarly, we will just write t_n instead of t_n^∞ for the times of the hierarchy. The generating Lax equation (3.8) gives us, using the definitions (3.6), (3.14) and (3.15),

$$\partial_{t_n} Q(\lambda) = [V_n(\lambda), Q(\lambda)] \quad (4.5)$$

where

$$V_n(\lambda) = \sum_{r=0}^n Q_r \lambda^{n-r} \quad (4.6)$$

are the Lax matrices of the hierarchy. Eqs (4.5) are the central equations of [FNR] where only Hamiltonian aspects of the theory were developed. The associated zero curvature equations read

$$\partial_{t_k} V_n(\lambda) - \partial_{t_n} V_k(\lambda) + [V_n(\lambda), V_k(\lambda)] = 0, \quad n, k \geq 0, \quad (4.7)$$

and produce the equations of motion of the hierarchy. The famous (unreduced) NLS system corresponds to $n = 1$ and $k = 2$. From our generating Lagrangian (1.15), we can of course reproduce the generating Lagrangian of [CS3] and all the Lagrangians forming the Lagrangian multiform that gives these equations as its (multiform) EL equations. Since $S = \{\infty\}$ we only have $\mathcal{L}^{\infty, \infty}(\lambda_\infty, \mu_\infty)$ to consider. As above, let us simply denote it as $\mathcal{L}(\lambda, \mu)$. The coefficient $\mathcal{L}_{mn}$ of $\lambda^{-m-1} \mu^{-n-1}$ in its expansion reads

$$\mathcal{L}_{mn} = \sum_{i=1}^m \text{Tr} \tilde{\phi}_i \partial_{t_n} \phi_{m-i+1} X_0 - \sum_{i=1}^n \text{Tr} \tilde{\phi}_i \partial_{t_m} \phi_{n-i+1} X_0 - U_{mn} \quad (4.8)$$

where we wrote $\phi^{-1}(\lambda) = \mathbf{1} + \sum_{n=1}^{\infty} \tilde{\phi}_n \lambda^{-n}$ for convenience and where U_{mn} is given by

$$U_{mn} = -\text{Tr} \sum_{j=0}^m Q_{m+n+1-j} Q_j. \quad (4.9)$$

These are the coefficients of the AKNS Lagrangian multiform found in [CS3] (up to an overall minus sign) to which we refer for more details. It was explained in [CS3] that

there exists a parametrization of $\phi(\lambda)$ in terms of very nice coordinates $e(\lambda) = \sum_{i=1}^{\infty} e_i \lambda^{-i}$,

$$f(\lambda) = \sum_{i=1}^{\infty} f_i \lambda^{-i} \text{ as}$$

$$\phi(\lambda) = \frac{1}{\sqrt{2i}} \begin{pmatrix} \sqrt{2i - e(\lambda)f(\lambda)} & e(\lambda) \\ -f(\lambda) & \sqrt{2i - e(\lambda)f(\lambda)} \end{pmatrix}. \quad (4.10)$$

For the reader's convenience, let us give for instance

$$\mathcal{L}_{12} = \frac{1}{2}(f_1 \partial_{t_2} e_1 - e_1 \partial_{t_2} f_1) - \frac{1}{2} \sum_{j=1}^2 (f_j \partial_{t_1} e_{2-j+1} - e_j \partial_{t_1} f_{2-j+1}) - 2ie_2 f_2 - e_1^2 f_1^2 \quad (4.11)$$

and

$$\begin{aligned} \mathcal{L}_{13} = & \frac{1}{2}(f_1 \partial_{t_3} e_1 - e_1 \partial_{t_3} f_1) - \frac{1}{2} \sum_{j=1}^3 (f_j \partial_{t_1} e_{3-j+1} - e_j \partial_{t_1} f_{3-j+1}) \\ & - 2i(e_2 f_3 + e_3 f_2) - \frac{3}{2} e_1 f_1 (f_1 e_2 + f_2 e_1) \end{aligned} \quad (4.12)$$

Of course, one can check that the equations of motion for these Lagrangians give precisely the zero curvature equations (4.7) for $(k, n) = (1, 2)$ and $(k, n) = (1, 3)$ respectively. For instance, varying $\mathcal{L}_{12}$ with respect to $e_j, f_j, j = 1, 2$, we have

$$\partial_{t_1} e_1 + 2i e_2 = 0, \quad \partial_{t_1} f_1 - 2i f_2 = 0, \quad (4.13)$$

$$\partial_{t_2} e_1 - \partial_{t_1} e_2 - 2e_1^2 f_1 = 0, \quad \partial_{t_2} f_1 - \partial_{t_1} f_2 + 2f_1^2 e_1 = 0. \quad (4.14)$$

This is equivalent to (4.7) for $(k, n) = (1, 2)$, upon recalling that

$$Q_1 = \begin{pmatrix} 0 & \sqrt{2i} e_1 \\ \sqrt{2i} f_1 & 0 \end{pmatrix}, \quad Q_2 = \begin{pmatrix} e_1 f_1 & \sqrt{2i} e_2 \\ \sqrt{2i} f_2 & -e_1 f_1 \end{pmatrix}.$$

The top two equations can be used to eliminate e_2, f_2 in the bottom two equations. With $t_2 = t, t_1 = x, e_1 = \frac{1}{\sqrt{2i}} q, f_1 = \frac{1}{\sqrt{2i}} r$ we get

$$i \partial_t q + \frac{1}{2} \partial_x^2 q - q^2 r = 0, \quad -i \partial_t r + \frac{1}{2} \partial_x^2 r - r^2 q = 0, \quad (4.15)$$

and the reduction $r = \mp r^*$ yields the well-known (de)focusing NLS equation

$$i \partial_t q + \frac{1}{2} \partial_x^2 q \pm |q|^2 q = 0$$

for the complex field q . Similarly, $\mathcal{L}_{13}$ gives the complex modified KdV equation.

5. Sine-Gordon Hierarchy

For the example of the sine-Gordon equation

$$u_{xy} + \sin u = 0, \quad (5.1)$$

we choose the trigonometric r -matrix (2.20). The required data is fixed as follows

$$S = \{0, \infty\}, \quad N_0 = 1 = N_\infty, \quad \mathfrak{g} = \mathfrak{sl}_2, \quad F(\lambda) = \frac{i}{2} \left(\frac{1}{\lambda} \sigma_+ + \sigma_- - \sigma_+ - \lambda \sigma_- \right), \quad (5.2)$$

and we work with the basis $\sigma_3, \sigma_+, \sigma_-$. The adjoint orbit description of Sect. 3.1 is implemented with

$$\phi^0(\lambda) = \sum_{n=0}^{\infty} \phi_n^0 \lambda^n, \quad \phi_0^0 = e^{i \frac{\pi}{4} \sigma_3}, \quad (5.3)$$

$$\phi^\infty(\lambda_\infty) = \sum_{n=0}^{\infty} \phi_n^\infty \lambda_\infty^n, \quad \phi_0^\infty = e^{-i \frac{\pi}{4} \sigma_3}. \quad (5.4)$$

The phase space coordinate u will be the sine-Gordon field as will become clear soon. This gives, with $(\iota_{\lambda_0} F(\lambda))_-^{\text{trig}} = \frac{i}{2} \left(\frac{1}{\lambda} \sigma_+ + \sigma_- \right)$ and $(\iota_{\lambda_\infty} F(\lambda))_-^{\text{trig}} = -\frac{i}{2} (\lambda \sigma_- + \sigma_+)$,

$$Q^0(\lambda_0) = \frac{i}{2} \phi^0(\lambda) \left(\frac{1}{\lambda} \sigma_+ + \sigma_- \right) \phi^0(\lambda)^{-1} = \sum_{n=-1}^{\infty} Q_n^0 \lambda^n, \quad (5.5)$$

$$Q^\infty(\lambda_\infty) = -\frac{i}{2} \phi^\infty(\lambda_\infty) \left(\frac{1}{\lambda_\infty} \sigma_- + \sigma_+ \right) \phi^\infty(\lambda_\infty)^{-1} = \sum_{n=-1}^{\infty} Q_n^\infty \lambda_\infty^n, \quad (5.6)$$

with $Q_{-1}^0 = \frac{i}{2}e^{\frac{iu}{2}}\sigma_+$ and $Q_{-1}^\infty = -\frac{i}{2}e^{\frac{iu}{2}}\sigma_-$. We now show how to use our formalism to recover the sine-Gordon equation (in light cone coordinates) as well as its first higher compatible flow which is nothing but the modified KdV equation, as presented in [Su]. We take advantage of this example to illustrate how our formalism also produces the Lagrangian multiform corresponding to these 3 times. In this context, our motivation is to show that the so-called “alien derivatives” problem that was discussed in [V] does not appear with our approach. The problem only arises if one insists on using the variational equations we obtain to eliminate some of the phase space coordinates in favour of the sine-Gordon field u and its derivatives with respect to a given time. In other words, we show in detail how our general discussion about the FNR procedure, when applied at the variational level, leads to this alien derivative problem. This is yet another reason in our opinion why it is preferable to work with the natural phase space coordinates that are provided by $\phi(\lambda)$.

It is convenient to parametrise (5.3)–(5.4) as

$$\phi^0(\lambda_0) = e^{i\frac{u}{4}\sigma_3}(\mathbf{1} + \psi^0(\lambda)), \quad \psi^0(\lambda) = \sum_{n=1}^{\infty} \psi_n^0 \lambda^n, \quad (5.7)$$

$$\phi^\infty(\lambda_\infty) = e^{-i\frac{u}{4}\sigma_3}(\mathbf{1} + \psi^\infty(\lambda_\infty)), \quad \psi^\infty(\lambda_\infty) = \sum_{n=1}^{\infty} \psi_n^\infty \lambda_\infty^n, \quad (5.8)$$

where we recall that $\det \phi^0 = 1 = \phi^\infty$ should hold. Using the gauge freedom of multiplying $\phi^0(\lambda_0)$ (resp. $\phi^\infty(\lambda_\infty)$) on the right by a matrix which commutes with $(\iota_{\lambda_0} F(\lambda))_-^{\text{trig}}$ (resp. $(\iota_{\lambda_\infty} F(\lambda))_-^{\text{trig}}$), we can work with

$$\psi^0(\lambda) = \sum_{n=1}^{\infty} \psi_n^0 \lambda^n, \quad \psi_n^0 = \begin{pmatrix} A_n^0 & 0 \\ C_n^0 & D_n^0 \end{pmatrix}, \quad (5.9)$$

$$\psi^\infty(\lambda_\infty) = \sum_{n=1}^{\infty} \psi_n^\infty \lambda_\infty^n, \quad \psi_n^\infty = \begin{pmatrix} A_n^\infty & B_n^\infty \\ 0 & D_n^\infty \end{pmatrix}. \quad (5.10)$$

Note that one can show that there is a bijection between the group coordinates A_n^0, C_n^0, D_n^0 and $A_n^\infty, C_n^\infty, D_n^\infty$, and the algebra coordinates a_n^0, b_n^0 and c_n^0 , which we would introduce via $Q_n^0 = a_n^0 \sigma_3 + b_n^0 \sigma_+ + c_n^0 \sigma_-$ (and similarly at ∞). The reader familiar with the FNR construction or only interested in zero curvature equations would tend to use the algebra coordinates. However, since our Lagrangians are naturally expressed with group coordinates, we use the latter both for the zero curvature equations and the Lagrangians. It also facilitates comparison between the two ways of obtaining the equations of motion.

By our general results in Sects. 3.1.2 and 3.1.3, all the time flows commute and all the corresponding zero curvature equations of Proposition 3.5 hold, with the Lax matrices reading for $n \geq -1$, (see Proposition 3.8)

$$V_n^0(\lambda) = -(P^- + \frac{1}{2}P^0)Q_n^0 - \frac{1}{\lambda}Q_{n-1}^0 - \cdots - \frac{1}{\lambda^n}Q_0^0 - \frac{1}{\lambda^{n+1}}Q_{-1}^0, \quad (5.11)$$

$$V_n^\infty(\lambda) = (P^+ + \frac{1}{2}P^0)Q_n^\infty + \lambda Q_{n-1}^\infty + \cdots + \lambda^n Q_0^\infty + \lambda^{n+1} Q_{-1}^\infty. \quad (5.12)$$

The sine-Gordon equation is recovered by taking the pair of Lax matrices $(V_0^0(\lambda), V_0^\infty(\lambda))$ and the compatible higher flow attached to the pair $(V_0^\infty(\lambda), V_1^\infty(\lambda))$ gives the mKdV

equation in potential form. The third possible Lax pair is $(V_0^0(\lambda), V_1^\infty(\lambda))$ and will be called the mixed equation. For convenience, let us label the corresponding times as follows $t_0^0 = y, t_0^\infty = x, t_1^\infty = z$. Therefore, we focus on the following three zero curvature equations

1. $\partial_x V_0^0(\lambda) - \partial_y V_0^\infty(\lambda) + [V_0^0(\lambda), V_0^\infty(\lambda)] = 0$ (sG);
2. $\partial_z V_0^\infty(\lambda) - \partial_x V_1^\infty(\lambda) + [V_0^\infty(\lambda), V_1^\infty(\lambda)] = 0$ (mKdV);
3. $\partial_z V_0^0(\lambda) - \partial_y V_1^\infty(\lambda) + [V_0^0(\lambda), V_1^\infty(\lambda)] = 0$ (mixed).

A direct calculation gives

$$Q_{-1}^0 = \frac{i}{2} e^{\frac{i u}{2}} \sigma_+, \quad Q_0^0 = \frac{i}{2} \begin{pmatrix} -C_1^0 & 2A_1^0 e^{i \frac{u}{2}} \\ e^{-i \frac{u}{2}} & C_1^0 \end{pmatrix}, \quad (5.13)$$

$$Q_1^0 = \frac{i}{2} \begin{pmatrix} -C_2^0 - A_1^0 C_1^0 & (2A_2^0 + (A_1^0)^2) e^{i \frac{u}{2}} \\ (2D_1^0 - (C_1^0)^2) e^{-i \frac{u}{2}} & C_2^0 + A_1^0 C_1^0 \end{pmatrix}, \quad (5.14)$$

$$Q_{-1}^\infty = -\frac{i}{2} e^{\frac{i u}{2}} \sigma_-, \quad Q_0^\infty = -\frac{i}{2} \begin{pmatrix} B_1^\infty & e^{-i \frac{u}{2}} \\ 2D_1^\infty e^{i \frac{u}{2}} & -B_1^\infty \end{pmatrix}, \quad (5.15)$$

$$Q_1^\infty = -\frac{i}{2} \begin{pmatrix} B_2^\infty + B_1^\infty D_1^\infty & (2A_1^\infty - (B_1^\infty)^2) e^{-i \frac{u}{2}} \\ (2D_2^\infty + (D_1^\infty)^2) e^{i \frac{u}{2}} & -B_2^\infty - B_1^\infty D_1^\infty \end{pmatrix}. \quad (5.16)$$

Hence,

$$V_0^0(\lambda) = -(P^- + \frac{1}{2} P^0) Q_0^0 - \frac{1}{\lambda} Q_{-1}^0 = -\frac{i}{4} \begin{pmatrix} -C_1^0 & 2e^{i \frac{u}{2}} / \lambda \\ 2e^{-i \frac{u}{2}} & C_1^0 \end{pmatrix}, \quad (5.17)$$

$$V_0^\infty(\lambda) = (P^+ + \frac{1}{2} P^0) Q_0^\infty + \lambda Q_{-1}^\infty = -\frac{i}{4} \begin{pmatrix} B_1^\infty & 2e^{-i \frac{u}{2}} \\ 2\lambda e^{i \frac{u}{2}} & -B_1^\infty \end{pmatrix}, \quad (5.18)$$

$$\begin{aligned} V_1^\infty(\lambda) &= (P^+ + \frac{1}{2} P^0) Q_1^\infty + \lambda Q_0^\infty + \lambda^2 Q_{-1}^\infty \\ &= -\frac{i}{4} \begin{pmatrix} 2\lambda B_1^\infty + B_2^\infty - A_1^\infty B_1^\infty & 2(\lambda + 2A_1^\infty - B_1^\infty) e^{-i \frac{u}{2}} \\ 2\lambda(\lambda - 2A_1^\infty) e^{i \frac{u}{2}} & -2\lambda B_1^\infty - B_2^\infty + A_1^\infty B_1^\infty \end{pmatrix}. \end{aligned} \quad (5.19)$$

Therefore, we obtain the following equations of motion from the zero curvature equations:

$$(sG) \quad \begin{cases} C_1^0 = -u_y, \\ B_1^\infty = -u_x, \\ \partial_x C_1^0 + \partial_y B_1^\infty - 2 \sin u = 0. \end{cases} \quad (5.20)$$

The first two equations show that the group coordinates C_1^0, B_1^∞ can be thought of as auxiliary fields and can be eliminated from the dynamics to get (5.1), as desired.

$$(mKdV) \quad \begin{cases} B_1^\infty = -u_x, \\ A_1^\infty = -\frac{i}{2} \partial_x B_1^\infty + \frac{1}{4} (B_1^\infty)^2, \\ u_z - u_x (2A_1^\infty - (B_1^\infty)^2) - 2i \partial_x (2A_1^\infty - (B_1^\infty)^2) - (B_1^\infty)^3 + 3A_1^\infty B_1^\infty - B_2^\infty = 0 \\ u_z + 2u_x A_1^\infty - 4i \partial_x A_1^\infty - A_1^\infty B_1^\infty - B_2^\infty = 0, \\ \partial_z B_1^\infty + \partial_x (A_1^\infty B_1^\infty - B_2^\infty) = 0 \end{cases} \quad (5.21)$$

We see that both (5.20) and (5.21) contain the same equation for B_1^∞ in terms of u , as it should be. A comment is in order. Under the first two equations, the third and fourth equation consistently give the same expression for B_2^∞ . In turn, replacing all the auxiliary fields into the last equation yields mKdV in potential form (i.e. mKdV for $v = u_x$)

$$u_{xz} + u_{xxx} + \frac{3}{2}u_{xx}u_x^2 = 0. \quad (5.22)$$

$$(mixed) \quad \begin{cases} C_1^0 = -u_y, \\ A_1^\infty B_1^\infty - B_2^\infty = u_z, \\ \partial_y B_1^\infty = \sin u, \\ i\partial_y (e^{-iu/2}(2A_1^\infty - (B_1^\infty)^2)) + \frac{1}{2}C_1^0(2A_1^\infty - (B_1^\infty)^2)e^{-iu/2} + B_1^\infty e^{iu/2} = 0, \\ 2i\partial_y (A_1^\infty e^{iu}) + B_1^\infty e^{-iu/2} - A_1^\infty C_1^0 e^{iu/2} = 0, \\ i\partial_z C_1^0 - i\partial_y (A_1^\infty B_1^\infty - B_2^\infty) + 2A_1^\infty (e^{iu} + e^{-iu}) - (B_1^\infty)^2 e^{-iu} = 0. \end{cases} \quad (5.23)$$

Using the first two equations to eliminate the auxiliary fields and noting that the fourth and fifth equations are equivalent (modulo the third equation), we obtain after simplification the following system of equations for the three fields u , A_1^∞ and B_1^∞ ,

$$(mixed) \quad \begin{cases} \partial_y B_1^\infty = \sin u, \\ \partial_y A_1^\infty = \frac{i}{2}B_1^\infty e^{-iu}, \\ -2iu_{yz} + 2A_1^\infty (e^{iu} + e^{-iu}) - (B_1^\infty)^2 e^{-iu} = 0. \end{cases} \quad (5.24)$$

Note that this system of equations in (y, z) can be perfectly studied on its own and is integrable. However, from our point of view, it should be included together with (5.20) and (5.21) into the sG hierarchy. This leads to interesting observations which are related to the Lagrangian multiform description we present below. First of all, using $B_1^\infty = -u_x$ and the sine-Gordon equation, we see that the first equation in (5.24) is trivially satisfied. Similarly, the second equation in (5.24) is a consequence of $B_1^\infty = -u_x$, the sine-Gordon equation and the second equation in (5.21). Perhaps more interesting is the fact that combining the first three equations in (5.21) with the second equation in (5.23) yields

$$u_z + u_{xxx} + \frac{1}{2}u_x^3 = 0, \quad (5.25)$$

of which (5.22) is simply a differential consequence.

We now turn to the extraction of the coefficients of the Lagrangian multiform for the corresponding time flows. We need $\mathcal{L}_{00}^{0,\infty}$ for (sG), $\mathcal{L}_{01}^{0,\infty}$ for (mixed) and $\mathcal{L}_{01}^{\infty,\infty}$ for (mKdV). We have

$$\begin{aligned} K^{0,\infty}(\lambda_0, \mu_\infty) = & \operatorname{Tr} \left[\frac{i}{2} \left(\frac{1}{\lambda} \sigma_+ + \sigma_- \right) \left(\phi_0^{(0)} + \lambda \phi_1^{(0)} + \lambda^2 \phi_2^{(0)} + O(\lambda^3) \right)^{-1} \right. \\ & \times \left(\frac{1}{\mu_\infty} \partial_{t_{-1}}^\infty + \partial_{t_0}^\infty + \mu_\infty \partial_{t_1}^\infty + O(\mu_\infty^2) \right) \left(\phi_0^{(0)} + \lambda \phi_1^{(0)} + \lambda^2 \phi_2^{(0)} \cdots + O(\lambda^3) \right) \Big] \\ & + \operatorname{Tr} \left[\frac{i}{2} \left(\frac{1}{\mu_\infty} \sigma_- + \sigma_+ \right) \left(\phi_0^\infty + \mu_\infty \phi_1^\infty + \mu_\infty^2 \phi_2^\infty + O(\lambda^3) \right)^{-1} \right. \\ & \times \left(\frac{1}{\lambda} \partial_{t_{-1}}^0 + \partial_{t_0}^0 + \lambda \partial_{t_1}^0 + O(\lambda^2) \right) \left(\phi_0^\infty + \mu_\infty \phi_1^\infty + \mu_\infty^2 \phi_2^\infty \cdots + O(\mu_\infty^3) \right) \Big]. \end{aligned}$$

Hence, dropping irrelevant total derivative terms and using again $t_0^\infty = y$, $t_0^\infty = x$, $t_1^\infty = z$ for convenience, we find

$$K_{00}^{0,\infty} = \frac{1}{4}C_1^0 u_x + \frac{1}{4}B_1^\infty u_y, \quad K_{01}^{0,\infty} = \frac{1}{4}C_1^0 u_z - \frac{1}{4}(A_1^\infty B_1^\infty - B_2^\infty)u_y \\ - A_1^\infty \partial_y B_1^\infty + B_1^\infty \partial_y A_1^\infty.$$

To compute the potential terms, observe that for the trigonometric r -matrix, we have

$$(\iota_{\lambda_0} \iota_{\mu_\infty} + \iota_{\mu_\infty} \iota_{\lambda_0})r_{12}(\lambda, \mu) = -2P_{12}^- - P_{12}^0 + 2P_{12} \sum_{n=0}^{\infty} \lambda^n \mu_\infty^n.$$

Hence,

$$U_{00}^{0,\infty} = \text{Tr} \left(Q_0^0 (P^+ + \frac{1}{2}P^0) Q_0^\infty + Q_{-1}^0 Q_{-1}^\infty \right) = \frac{1}{4} (e^{iu} + e^{-iu} - C_1^0 B_1^\infty), \\ U_{01}^{0,\infty} = \text{Tr} \left(Q_0^0 (P^+ + \frac{1}{2}P^0) Q_1^\infty + Q_{-1}^0 Q_0^\infty \right) \\ = \frac{1}{4} (2A_1^\infty - (B_1^\infty)^2) e^{-iu} + \frac{1}{4} C_1^0 (A_1^\infty B_1^\infty - B_2^\infty) - \frac{1}{2} A_1^\infty e^{iu}.$$

This gives us the desired Lagrangian densities for (sG) and (mixed) as

$$\mathcal{L}_{\text{sG}} \equiv \mathcal{L}_{00}^{0,\infty} = K_{00}^{0,\infty} - U_{00}^{0,\infty}, \quad (5.26)$$

and

$$\mathcal{L}_{\text{mixed}} \equiv \mathcal{L}_{01}^{0,\infty} = K_{01}^{0,\infty} - U_{01}^{0,\infty}. \quad (5.27)$$

Similarly, we find

$$K_{01}^{\infty,\infty} = -\frac{1}{4}B_1^\infty u_z - \frac{1}{4}(A_1^\infty B_1^\infty - B_2^\infty)u_x - \frac{i}{2}A_1^\infty \partial_x B_1^\infty + \frac{i}{2}B_1^\infty \partial_x A_1^\infty$$

and, with

$$(\iota_{\lambda_\infty} \iota_{\mu_\infty} + \iota_{\mu_\infty} \iota_{\lambda_\infty})r_{12}(\lambda, \mu) = -2P_{12}^- - P_{12}^0 + P_{12} \sum_{n=0}^{\infty} \frac{\mu_\infty^n}{\lambda_\infty^n} - P_{12} \sum_{n=0}^{\infty} \frac{\lambda_\infty^{n+1}}{\mu_\infty^{n+1}},$$

we get

$$U_{01}^{\infty,\infty} = \text{Tr} \left(Q_0^\infty (P^+ + \frac{1}{2}P^0) Q_1^\infty \right) = \frac{1}{4}B_1^\infty (A_1^\infty B_1^\infty - B_2^\infty) + \frac{1}{2}A_1^\infty (2A_1^\infty - (B_1^\infty)^2).$$

Thus, the Lagrangian density for (mKdV) is given by

$$\mathcal{L}_{\text{mKdV}} \equiv \mathcal{L}_{01}^{\infty,\infty} = K_{01}^{\infty,\infty} - U_{01}^{\infty,\infty}. \quad (5.28)$$

It remains to derive the EL equations associated to each Lagrangian. For instance, by varying B_1^∞ , C_1^0 and u in $\mathcal{L}_{\text{sG}}$ we find exactly the three equations in (5.20). Similarly, it can be checked that the E-L equations for $\mathcal{L}_{\text{mKdV}}$ and $\mathcal{L}_{\text{mixed}}$ reproduce (5.21) and (5.23) respectively.

In particular, all the equations that determine the group coordinates in terms of u and its (relevant) derivatives are reproduced variationally. This is an interesting feature that the FNR procedure is also obtained variationally with our construction. An important by-product is that the so-called problem of “alien-derivatives” is eliminated systematically. In the present context, the manifestation of this problem would be for instance that the Lagrangian $\mathcal{L}_{\text{mixed}}$ contains terms with derivatives of u with respect to x , while this Lagrangian is supposed to produce equations of motion with respect to the variables y and z only. Clearly, our Lagrangians do not suffer from this problem since by construction, they always only involve the two times they are supposed to produce equations of motion for. The problem is an artefact of using some of the equations of motion to solve for some of the fields in terms of u and its derivatives. In other words, it is an artefact of implementing the FNR procedure *a priori* to eliminate some of the group coordinates. If we do implement this procedure of elimination, we obtain Lagrangians which form a Lagrangian multiform equivalent to the one given originally in [Su] and which suffers from this problem. Eliminating the auxiliary fields in favour of u and its derivatives, we obtain

$$\mathcal{L}_{\text{sG}} = -\frac{1}{4}u_x u_y - \frac{1}{2}\cos u$$

which is a well-known Lagrangian for (5.1), as well as

$$\mathcal{L}_{\text{mKdV}} = \frac{1}{4}u_x u_z + \frac{1}{16}u_x^4 - \frac{1}{4}u_{xx}^2 - \frac{i}{4}\partial_x \left(\frac{1}{6}u_x^3 + iu_x u_{xx} \right)$$

and

$$\mathcal{L}_{\text{mixed}} = -\frac{1}{4}u_y u_z - \frac{1}{2}u_{xx}(u_{xy} + \sin u) + \frac{1}{4}u_x^2 \cos u - \frac{i}{4}\partial_y \left(\frac{1}{6}u_x^3 + iu_x u_{xx} \right).$$

Changing x to $-x$, multiplying all our Lagrangian by 2 and dropping the irrelevant total derivatives in x and y , we recover exactly the three Lagrangian coefficients, eqs (31)–(33), in [Su]. Our Lagrangian multiform expressed with the group coordinates (and restricted to the three times x, y, z) is thus equivalent to that in [Su] but, as noted before, it does not suffer from the alien derivative problem.

The poles at 0 and ∞ play a symmetric role in the construction so it would be natural to consider also the time t_1^0 and the associated Lax matrix $V_1^0(\lambda)$. This naturally leads to two additional zero curvature equations (denote $t_1^0 = t$ and the other times as above) that can be combined with (sG)

1. $\partial_t V_0^0(\lambda) - \partial_y V_1^0(\lambda) + [V_0^0(\lambda), V_1^0(\lambda)] = 0$ (mKdV2);
2. $\partial_t V_0^\infty(\lambda) - \partial_x V_1^\infty(\lambda) + [V_0^\infty(\lambda), V_1^\infty(\lambda)] = 0$ (mixed 2);

The first one is called (mKdV2) as it is another copy of the mKdV equation but in (y, t) instead of (x, t) . It is a compatible flow with (sG) where we can think of the roles of x and y being swapped. Then, naturally (mixed 2) is the remaining compatible flow between the variables x and t . All the expressions for the Lax matrices, the zero curvature equations and the corresponding Lagrangians are similar to the above ones with the appropriate changes and we omit them. To complete the picture related to the four times we have focussed on, it would remain to consider the zero curvature equation

$$\partial_z V_1^0(\lambda) - \partial_t V_1^\infty(\lambda) + [V_1^0(\lambda), V_1^\infty(\lambda)] = 0.$$

The set of equations of motion is not particularly enlightening. When embedded in the hierarchy of the five zero curvature equations already discussed, this system is a consequence of them, as it should be. Our construction gives us the means to derive the corresponding Lagrangian density $\mathcal{L}_{11}^{(0,\infty)}$ if required but again we omit its lengthy expression here.

FNR procedure for the sine-Gordon hierarchy

We have discussed the FNR procedure at the level of the EL equations above, using some of the equations to eliminate certain auxiliary coordinates/fields. Here, we discuss it at the level of the algebra coordinates using the Lax equation. This is more in line with the original work [FNR] and with the explanation around Lemma 3.9 for which it provides an illustration in the sG case. We recall that our point of view is that the procedure is unnecessary. We show it in the sG case to make contact with a more traditional approach but also because to our knowledge, this is the first time that the FNR construction is obtained for a hierarchy other than AKNS. In the present sG case, it is based on the Lax equations (5.29)–(5.30) below.

The generating Lax equation (3.8) gives the following equations, for $n \geq -1$,

$$\begin{aligned} \partial_{t_0} Q^0(\lambda) &= \left[V_n^0(\lambda), Q^0(\lambda) \right] = \left[-(\lambda^{-n} Q^0(\lambda))_-^{\text{trig}}, Q^0(\lambda) \right], \\ \partial_{t_n^\infty} Q^\infty(\lambda_\infty) &= \left[V_n^\infty(\lambda), Q^\infty(\lambda_\infty) \right] = \left[(\lambda_\infty^{-n} Q^\infty(\lambda_\infty))_-^{\text{trig}}, Q^\infty(\lambda_\infty) \right]. \end{aligned} \quad (5.29)$$

$$(5.30)$$

We could use the Lax equations (5.29)–(5.30) to derive the coefficients of Q_n^0 and Q_n^∞ as differential polynomials in the coordinate u . Given the form of $F(\lambda)$ here, we do not fall into the area of applicability of the argument given after Lemma 3.9. Nevertheless, it is still possible to proceed. We illustrate this with (5.30), the other case being similar.

Our choices (5.2) and (5.3) give $c_{-1} = -\frac{i}{2}e^{iu/2}$, $a_{-1} = 0 = b_{-1}$. Then, consider (5.30) for $n = 0$ with $Q^\infty(\lambda) = \begin{pmatrix} a(\lambda) & b(\lambda) \\ c(\lambda) & -a(\lambda) \end{pmatrix}$ (we drop the superscript for conciseness). Writing $t_0^\infty = x$ for convenience and projecting onto σ_3 , σ_+ and σ_- , we obtain

$$\begin{cases} \partial_x a(\lambda) = b_0 c(\lambda) - \frac{1}{\lambda} c_{-1} b(\lambda), \\ \partial_x b(\lambda) = a_0 b(\lambda) - 2b_0 a(\lambda), \\ \partial_x c(\lambda) = -a_0 c(\lambda) + \frac{2}{\lambda} c_{-1} a(\lambda). \end{cases} \quad (5.31)$$

Looking at the λ^j coefficient, this yields the following system

$$\begin{cases} \partial_x a_j = b_0 c_j - c_{-1} b_{j+1}, \\ \partial_x b_j = a_0 b_j - 2b_0 a_j, \\ \partial_x c_j = -a_0 c_j + 2c_{-1} a_{j+1}, \end{cases} \quad (5.32)$$

which we should use to determine the coefficients recursively. Suppose, we have determined a_k, b_k, c_k for $k = 1, \dots, n-1$ then the first equation gives us b_n and hence the second equation yields a_n . However, we cannot deduce c_n from the third equation since it would require the knowledge of a_{n+1} . It is possible to replace (5.31) by the following equivalent system

$$\begin{cases} \partial_x a(\lambda) = b_0 c(\lambda) - \frac{1}{\lambda} c_{-1} b(\lambda), \\ \partial_x b(\lambda) = a_0 b(\lambda) - 2b_0 a(\lambda), \\ a^2(\lambda) + b(\lambda)c(\lambda) = -\frac{\lambda}{4}. \end{cases} \quad (5.33)$$

To see this, note that (5.31) implies $\partial_x \operatorname{Tr} Q^\infty(\lambda_\infty)^2 = 0$ so that

$$\mathcal{L}_{-1-1}^{a_i, b_j} = -\operatorname{Tr} \left((\phi_0^{a_i})^{-1} \partial_{t_{-1}^{b_j}} \phi_0^{a_i} A_{i0} - (\phi_0^{b_j})^{-1} \partial_{t_{-1}^{a_i}} \phi_0^{b_j} B_{j0} - \frac{\phi_0^{a_i} A_{i0} (\phi_0^{a_i})^{-1} \phi_0^{b_j} B_{j0} (\phi_0^{b_j})^{-1}}{a_i - b_j} \right)$$

as it should by construction. Conversely, assume (5.33) holds. The third equation implies $2\partial_x a(\lambda)a(\lambda) + \partial_x b(\lambda)c(\lambda) + b(\lambda)\partial_x c(\lambda) = 0$. Using the first two equations to eliminate $\partial_x a(\lambda)$ and $\partial_x b(\lambda)$ yields $b(\lambda) \left(\partial_x c(\lambda) + a_0 c(\lambda) - \frac{2}{\lambda} c_{-1} a(\lambda) \right) = 0$, and the claim follows. Now the advantage of system (5.33) is that the j -th term of the third equation gives the following relation:

$$\sum_{i=0}^{j+1} (a_i a_{j-i} + b_i c_{j-i}) = -\frac{1}{4} \delta_{j,-1}. \quad (5.34)$$

Spelling it out, it can be seen that it can be used to determine c_n from $a_k, b_k, c_k, k = 1, \dots, n-1$ and b_n, a_n obtained from the first two equations as explained before. Thus, (5.33) allow us to determine all $a_j, b_j, c_j, j \geq 0$ recursively. We find the first few as

$$a_0 = \frac{i}{2} u_x, \quad b_0 = -\frac{i}{2} e^{-iu/2}, \quad c_0 = \frac{i}{2} e^{iu/2} \left(1 + i u_{xx} + \frac{1}{2} u_x^2 \right), \quad (5.35)$$

$$a_1 = -\frac{i}{2} \left(u_x + u_{xxx} + \frac{1}{2} u_x^3 \right), \quad b_1 = \frac{i}{2} e^{-iu/2} \left(1 - i u_{xx} + \frac{1}{2} u_x^2 \right), \quad (5.36)$$

$$c_1 = -\frac{i}{2} e^{iu/2} \left(\frac{1}{2} u_x^2 + \frac{3}{8} u_x^4 + i u_{xx} + u_x u_{xxx} - \frac{1}{2} u_{xx}^2 + i u_{xxx} + \frac{3i}{8} u_{xx} u_x^2 \right). \quad (5.37)$$

Now, for instance, the expression we find for a_0 is consistent with the fact that $a_0 = -\frac{i}{2} B_1^\infty$ from (5.15) and with the second equation in (5.20). This is what we mean when we say that the FNR procedure is automatically implemented with our Lagrangian approach. We reiterate that the advantage of not applying it is that the problem of alien derivatives disappears and that dependent variables are also treated on an equal footing, like the independent variables.

6. Hierarchies of Zakharov–Mikhailov Type

In this section, we introduce a rather large class of models and their hierarchies by using the following data

$$S = \{a_1, \dots, a_P\} \subset \mathbb{C}, \quad P > 0, \quad \mathfrak{g} = \mathfrak{gl}_N, \quad (6.1)$$

$$F(\lambda) = - \sum_{i=1}^P \sum_{r=0}^{n_i} \frac{A_{ir}}{(\lambda - a_i)^{r+1}}. \quad (6.2)$$

Each $A_{ir} \in \mathfrak{gl}_N$ is a non-dynamical constant matrix and we have chosen to write the order N_{a_i} of the pole $a_i, i = 1, \dots, P$ as $N_{a_i} = n_i + 1$ for convenience. All the poles in S are distinct. The r -matrix can be the rational or trigonometric one at this stage.

The motivation behind such choices is that in the simplest setting (rational r -matrix and simple poles), our construction reproduces the Zakharov–Shabat Lax pair with simple poles whose equations of motion were cast in variational form in [ZM1]. In fact,

our construction automatically embeds this single Lax pair, its zero curvature equation and its Lagrangian into an integrable hierarchy. This point of view was first introduced in [SNC] where the class of Zakharov–Mikhailov (ZM) models was cast into the formalism of Lagrangian multiforms. Allowing for higher order poles gives us the generalisation discussed in [Di, Chap. 20]. When we switch to the trigonometric r -matrix, we produce for the first time the trigonometric version of the large class of ZM models and their hierarchies. Finally, when specialising the construction via an appropriate reduction and choice of matrices A_{ir} , we obtain as a special case the class of models studied in [ABW]. Their integrability is guaranteed by construction and they are naturally embedded in an integrable hierarchy, a new feature for these models that were originally obtained as standalone models by a different method related to the $4d$ Chern–Simons construction (see conclusions for details and references). These examples are detailed in the next three subsections.

6.1. Rational Zakharov–Mikhailov models. We first describe in detail how to reproduce the class of Lax pairs and Lagrangians originally discussed in the pioneering paper [ZM1]. The generalisation to higher order poles presented in [Di] will be straightforward. The r -matrix is fixed to be the rational one in this subsection. We split the data (6.1)–(6.2) in the following way: $P = P_1 + P_2$, $P_1, P_2 > 0$, and

$$S = \{a_1, \dots, a_{P_1}, b_1, \dots, b_{P_2}\} \subset \mathbb{C}, \quad \mathfrak{g} = \mathfrak{gl}_N, \quad (6.3)$$

$$F(\lambda) = - \sum_{i=1}^{P_1} \sum_{r=0}^{n_i} \frac{A_{ir}}{(\lambda - a_i)^{r+1}} - \sum_{j=1}^{P_2} \sum_{r=0}^{m_j} \frac{B_{jr}}{(\lambda - b_j)^{r+1}}. \quad (6.4)$$

For notational convenience, we simply denoted $A_{j+P_1, r} = B_{jr}$ and $n_{j+P_1} = m_j$ for $j = 1, \dots, P_2$.

6.1.1. Case of simple poles Following [ZM1], let us consider a Lax pair of the form²

$$U(\lambda) = \sum_{i=1}^{P_1} \frac{U_i}{\lambda - a_i}, \quad V(\lambda) = \sum_{j=1}^{P_2} \frac{V_j}{\lambda - b_j}. \quad (6.5)$$

A prominent example of an integrable field theory that falls into this class is the Faddeev–Reshetikhin model [FR] which was proposed as an ultralocal variant of the principal chiral model. The main result of [ZM1] is that the equations of motions encoded in the zero curvature equation $\partial_\eta U(\lambda) - \partial_\xi V(\lambda) + [U(\lambda), V(\lambda)] = 0$ associated to the auxiliary problem

$$\Psi_\xi = U \Psi, \quad \Psi_\eta = V \Psi, \quad (6.6)$$

are variational and are obtained as the EL equations of the following Lagrangian density

$$\mathcal{L}_{ZM} = \text{Tr} \left(\sum_{i=1}^{P_1} \phi_i^{-1} \partial_\eta \phi_i U_i^{(0)} - \sum_{j=1}^{P_2} \psi_j^{-1} \partial_\xi \psi_j V_j^{(0)} - \sum_{i=1}^{P_1} \sum_{j=1}^{P_2} \frac{\phi_i U_i^{(0)} \phi_i^{-1} \psi_j V_j^{(0)} \psi_j^{-1}}{a_i - b_j} \right). \quad (6.7)$$

² In [ZM1], the authors include an additional term in U and V corresponding to a pole at ∞ but it can be gauged away.

The key insight to obtain this result is to parametrise U_i as $\varphi_i U_i^{(0)} \varphi_i^{-1}$ and V_j as $\psi_j V_j^{(0)} \psi_j^{-1}$. The matrices $U_i^{(0)}$ and $V_j^{(0)}$ are constant and all the dynamical variables are contained in the fields φ_i and ψ_j .

We can reproduce (6.5) by choosing $n_i = 0$ and $m_j = 0$ in our data (6.4). Since

$$\left(\iota_{\lambda_{a_i}} F(\lambda) \right)_{-}^{\text{rat}} = -\frac{A_{i0}}{\lambda - a_i}, \quad \left(\iota_{\lambda_{b_j}} F(\lambda) \right)_{-}^{\text{rat}} = -\frac{B_{j0}}{\lambda - b_j},$$

a direct calculation using Proposition 3.8 gives

$$V_{-1}^{a_i}(\lambda) = \frac{\phi_0^{a_i} A_{i0} (\phi_0^{a_i})^{-1}}{\lambda - a_i}, \quad V_{-1}^{b_j}(\lambda) = \frac{\phi_0^{b_j} B_{j0} (\phi_0^{b_j})^{-1}}{\lambda - b_j}. \quad (6.8)$$

Therefore, it remains to make the identifications $\phi_0^{a_i} = \varphi_i$ and $A_{i0} = U_i^{(0)}$, and $\phi_0^{b_j} = \psi_j$ and $B_{j0} = V_j^{(0)}$ and take linear combinations $\partial_{\xi} = \sum_{i=1}^{P_1} \partial_{t_{-1}^{a_i}}$, $\partial_{\eta} = \sum_{j=1}^{P_2} \partial_{t_{-1}^{b_j}}$ of the elementary time flows $\partial_{t_{-1}^{a_i}}$ and $\partial_{t_{-1}^{b_j}}$. The corresponding Lax matrices are simply the sum of the elementary Lax matrices (6.8) which gives precisely (6.5).

To understand how to recover the Lagrangian (6.7) with our method, note that the zero curvature equation associated to the elementary times $t_{-1}^{a_i}$ and $t_{-1}^{b_j}$ reads

$$\partial_{t_{-1}^{a_i}} V_{-1}^{b_j}(\lambda) - \partial_{t_{-1}^{b_j}} V_{-1}^{a_i}(\lambda) + [V_{-1}^{b_j}(\lambda), V_{-1}^{a_i}(\lambda)] = 0. \quad (6.9)$$

Summing these elementary zero curvature equations over $i = 1, \dots, P_1$ and $j = 1, \dots, P_2$ yields the desired $\partial_{\eta} U(\lambda) - \partial_{\xi} V(\lambda) + [U(\lambda), V(\lambda)] = 0$. Therefore, to find the Lagrangian $\mathcal{L}_{ZM}$ it suffices to sum the elementary Lagrangians $\mathcal{L}_{-1-1}^{a_i, b_j}$ (the coefficient of $\lambda_{a_i}^{-1} \mu_{b_j}^{-1}$ in $\mathcal{L}^{a_i, b_j}(\lambda_{a_i}, \mu_{b_j})$ which yields the equations of motion in (6.9)). A direct calculation gives

$$\mathcal{L}_{-1-1}^{a_i, b_j} = -\text{Tr} \left((\phi_0^{a_i})^{-1} \partial_{t_{-1}^{b_j}} \phi_0^{a_i} A_{i0} - (\phi_0^{b_j})^{-1} \partial_{t_{-1}^{a_i}} \phi_0^{b_j} B_{j0} - \frac{\phi_0^{a_i} A_{i0} (\phi_0^{a_i})^{-1} \phi_0^{b_j} B_{j0} (\phi_0^{b_j})^{-1}}{a_i - b_j} \right) \quad (6.10)$$

and the claim follows, i.e., with identifications made above, we derive $\mathcal{L}_{ZM}$ (up to an irrelevant minus sign) as in (6.7) by taking the double sum $\sum_{i=1}^{P_1} \sum_{j=1}^{P_2} \mathcal{L}_{-1-1}^{a_i, b_j}$.

It was shown for the first time in [SNC] that the ZM Lagrangian can be incorporated into a Lagrangian multiform where each coefficient is a copy of the original ZM Lagrangian associated to the corresponding times. The explicit case of 3 times was considered. We now explain how to recover this multiform from our data. Instead of splitting the data (6.1)–(6.2) into two types of poles as in (6.3)–(6.4), we split it into three types

of poles by setting $P = P_1 + P_2 + P_3$ and restrict our attention to simple poles, i.e. we set

$$S = \{a_1, \dots, a_{P_1}, b_1, \dots, b_{P_2}, c_1, \dots, c_{P_3}\} \subset \mathbb{C}, \quad P_1, P_2, P_3 > 0, \quad \mathfrak{g} = \mathfrak{gl}_N, \quad (6.11)$$

$$F(\lambda) = - \sum_{i=1}^{P_1} \frac{A_i}{\lambda - a_i} - \sum_{j=1}^{P_2} \frac{B_j}{\lambda - b_j} - \sum_{k=1}^{P_3} \frac{C_k}{\lambda - c_k}. \quad (6.12)$$

As before, we take the linear combinations $\partial_\xi = \sum_{i=1}^{P_1} \partial_{t_{-1}^{a_i}}, \partial_\eta = \sum_{j=1}^{P_2} \partial_{t_{-1}^{b_j}}$ of the elementary time flows, as well as the new combinations $\partial_v = \sum_{k=1}^{P_3} \partial_{t_{-1}^{c_k}}$. The original ZM Lagrangian is now denoted by $\mathcal{L}_{\xi\eta}$ and is accompanied by two new copies

$$\mathcal{L}_{\eta v} = \sum_{j=1}^{P_2} \sum_{k=1}^{P_3} \mathcal{L}_{-1-1}^{b_j, c_k}, \quad \mathcal{L}_{v\xi} = \sum_{k=1}^{P_3} \sum_{i=1}^{P_1} \mathcal{L}_{-1-1}^{c_k, a_i}. \quad (6.13)$$

The Lagrangian multiform in [SNC, Section 2.4] is precisely

$$\mathcal{L} = \mathcal{L}_{\xi\eta} d\xi \wedge d\eta + \mathcal{L}_{\eta v} d\eta \wedge dv + \mathcal{L}_{v\xi} dv \wedge d\xi. \quad (6.14)$$

The associated Lax matrices and zero curvature equations also reproduce those of [SNC].

6.1.2. Case of higher poles The generalisation of the ZM result to Lax matrices with higher order poles of the form

$$U(\lambda) = \sum_{i=1}^{P_1} U_i(\lambda), \quad V(\lambda) = \sum_{j=1}^{P_2} V_j(\lambda), \quad (6.15)$$

where

$$U_i = \sum_{r=0}^{n_i} \frac{U_{ir}}{(\lambda - a_i)^{r+1}}, \quad V_j = \sum_{r=0}^{m_j} \frac{V_{jr}}{(\lambda - b_j)^{r+1}} \quad (6.16)$$

was presented in [Di]. We can reproduce it by simply allowing n_i and m_j in the data (6.4) to be arbitrary positive integers and by following the same steps as for simple poles. In that case we find

$$V_{-1}^{a_i}(\lambda) = - \sum_{r=0}^{n_i} \frac{Q_{-r-1}^{a_i}}{(\lambda - a_i)^{r+1}} \equiv U_i(\lambda), \quad V_{-1}^{b_j}(\lambda) = - \sum_{r=0}^{m_j} \frac{Q_{-r-1}^{b_j}}{(\lambda - b_j)^{r+1}} \equiv V_j(\lambda) \quad (6.17)$$

where the coefficients are identified as

$$Q_{-r-1}^{a_i} =: -U_{ir}, \quad i = 1, \dots, P_1, \quad r = 0, \dots, n_i, \quad (6.18)$$

$$Q_{-r-1}^{b_j} =: -V_{jr}, \quad j = 1, \dots, P_2, \quad r = 0, \dots, m_j, \quad (6.19)$$

and calculated from the group coordinates using the following expansions

$$\begin{aligned} Q^{a_i}(\lambda_{a_i}) &= -\phi^{a_i}(\lambda_{a_i}) \sum_{r=0}^{n_i} \frac{A_{ir}}{(\lambda - a_i)^{r+1}} \phi^{a_i}(\lambda_{a_i})^{-1} = \sum_{k=-n_i-1}^{\infty} Q_k^{a_i}(\lambda - a_i)^k, \quad (6.20) \\ Q^{b_j}(\lambda_{b_j}) &= -\phi^{b_j}(\lambda_{b_j}) \sum_{r=0}^{m_j} \frac{B_{jr}}{(\lambda - b_j)^{r+1}} \phi^{b_j}(\lambda_{b_j})^{-1} = \sum_{k=-m_j-1}^{\infty} Q_k^{b_j}(\lambda - b_j)^k. \end{aligned} \quad (6.21)$$

As before, we simply assemble the elementary time flows into $\partial_\xi = \sum_{i=1}^{P_1} \partial_{t_{-1}^{a_i}}$ and $\partial_\eta = \sum_{j=1}^{P_2} \partial_{t_{-1}^{b_j}}$ which have the desired Lax pair (6.15). This gives the corresponding equations of motion in zero curvature form $\partial_\eta U(\lambda) - \partial_\xi V(\lambda) + [U(\lambda), V(\lambda)] = 0$. The Lagrangian producing these equations of motion is obtained by adding the elementary Lagrangians $\mathcal{L}_{-1-1}^{a_i, b_j}$. We give some details to show that we recover exactly [Di, Formula 20.2.12] (in the case of non coinciding poles which we consider here).

The kinetic part of $\mathcal{L}_{-1-1}^{a_i, b_j}$ reads

$$\begin{aligned} K_{-1-1}^{a_i, b_j} &= \operatorname{res}_{\lambda=a_i} \operatorname{res}_{\mu=b_j} \operatorname{Tr}(-\phi^{a_i}(\lambda_{a_i})^{-1} \mathcal{D}_{\mu_{b_j}} \phi^{a_i}(\lambda_{a_i}) \sum_{r=0}^{n_i} \frac{A_{ir}}{(\lambda - a_i)^{r+1}} \\ &\quad + \phi^{b_j}(\lambda_{b_j})^{-1} \mathcal{D}_{\lambda_{a_i}} \phi^{b_j}(\lambda_{b_j}) \sum_{r=0}^{m_j} \frac{B_{jr}}{(\lambda - b_j)^{r+1}}) \\ &= \operatorname{Tr}(-\operatorname{res}_{\lambda=a_i} \phi^{a_i}(\lambda_{a_i})^{-1} \partial_{t_{-1}^{b_j}} \phi^{a_i}(\lambda_{a_i}) \sum_{r=0}^{n_i} \frac{A_{ir}}{(\lambda - a_i)^{r+1}} \\ &\quad + \operatorname{res}_{\mu=b_j} \phi^{b_j}(\lambda_{b_j})^{-1} \partial_{t_{-1}^{a_i}} \phi^{b_j}(\lambda_{b_j}) \sum_{r=0}^{m_j} \frac{B_{jr}}{(\lambda - b_j)^{r+1}}) \\ &= \operatorname{Tr}(-\operatorname{res}_{\lambda=a_i} g_i^{-1} \partial_{t_{-1}^{b_j}} g_i A_i + \operatorname{res}_{\mu=b_j} h_j^{-1} \partial_{t_{-1}^{a_i}} h_j B_j) \end{aligned}$$

where in the last equality, we introduced g_i (resp. h_j) to denote the truncation of $\phi^{a_i}(\lambda_{a_i})$ (resp. $\phi^{b_j}(\lambda_{b_j})$) up to the order n_i (resp. m_j), in order to help make the comparison with Dickey's formula. The equality holds since the truncation is possible under the residue.

We also denoted $A_i := \sum_{r=0}^{n_i} \frac{A_{ir}}{(\lambda - a_i)^{r+1}}$ and $B_j := \sum_{r=0}^{m_j} \frac{B_{jr}}{(\lambda - b_j)^{r+1}}$ for conciseness.

The potential term reads, noting that $\iota_{\lambda_{a_i}} \iota_{\mu_{b_j}} = \iota_{\mu_{b_j}} \iota_{\lambda_{a_i}}$ when $a_i \neq b_j$,

$$\begin{aligned} U_{-1-1}^{a_i, b_j} &= \text{Tr}_{12} \left(\text{res}_{\lambda=a_i} \text{res}_{\mu=b_j} \iota_{\lambda_{a_i}} \iota_{\mu_{b_j}} \frac{P_{12}}{\mu - \lambda} (Q_{a_i}(\lambda_{a_i}))_1 (Q_{b_j}(\mu_{b_j}))_2 \right) \\ &= - \text{res}_{\lambda=a_i} Q^{a_i}(\lambda_{a_i}) \left(Q^{b_j}(\lambda_{b_j}) \right)_-^{\text{rat}} \\ &= - \text{res}_{\lambda=a_i} (Q^{a_i}(\lambda_{a_i}))_-^{\text{rat}} (Q^{b_j}(\lambda_{b_j}))_-^{\text{rat}} \\ &= - \text{res}_{\lambda=a_i} (g_i A_i g_i^{-1})_-^{\text{rat}} (h_j B_j h_j^{-1})_-^{\text{rat}}. \end{aligned}$$

We obtain Dickey's Lagrangian, up to an overall sign and a relative sign due to a different convention in the zero-curvature equation, by taking the following sums

$$L_D = - \sum_{i=1}^{P_1} \sum_{j=1}^{P_2} \left(K_{-1-1}^{a_i b_j} - U_{-1-1}^{a_i b_j} \right). \quad (6.22)$$

6.1.3. Interplay between hierarchies associated to simple and higher order poles Following Proposition 3.8, the Lax matrices read, for each $n \geq -n_i - 1$ and $i = 1, \dots, P_1$, and for each $m \geq -m_j - 1$ and $j = 1, \dots, P_2$:

$$V_n^{a_i}(\lambda) = - \left(\frac{Q^{a_i}(\lambda_{a_i})}{(\lambda - a_i)^{n+1}} \right)_-^{\text{rat}} = - \sum_{r=0}^{n+n_i+1} \frac{Q_{n-r}^{a_i}}{(\lambda - a_i)^{r+1}}, \quad (6.23)$$

$$V_m^{b_j}(\lambda) = - \left(\frac{Q^{b_j}(\lambda_{b_j})}{(\lambda - b_j)^{m+1}} \right)_-^{\text{rat}} = - \sum_{r=0}^{m+m_j+1} \frac{Q_{m-r}^{b_j}}{(\lambda - b_j)^{r+1}}. \quad (6.24)$$

At first glance, it is tempting to suggest that a Dickey hierarchy with certain fixed order n_i and m_j simply sits higher or lower in another Dickey hierarchy with different fixed n_i and m_j . The situation is much more complicated in general. To illustrate what we mean and show that this is too naive, let us focus on the field content of a Lax matrix around a pole a and compare the ZM case (where a is a simple pole) with the Dickey case (where a has order $n_1 + 1 > 1$). The corresponding Lax matrices are

$$V_n^{ZM,a}(\lambda) = - \left(\frac{Q^{ZM,a}(\lambda_a)}{(\lambda - a)^{n+1}} \right)_-^{\text{rat}} = - \sum_{r=0}^{n+1} \frac{Q_{n-r}^{ZM,a}}{(\lambda - a)^{r+1}}, \quad n \geq -1, \quad (6.25)$$

and

$$V_n^{D,a}(\lambda) = - \left(\frac{Q^{D,a}(\lambda_a)}{(\lambda - a)^{n+1}} \right)_-^{\text{rat}} = - \sum_{r=0}^{n+n_1+1} \frac{Q_{n-r}^{D,a}}{(\lambda - a)^{r+1}}, \quad n \geq -n_1 - 1. \quad (6.26)$$

In general, it is always the case that the Dickey hierarchy contains the ZM case as its lowest level. Indeed,

$$V_{-n_1-1}^{D,a}(\lambda) = \left((\lambda - a)^{n_1} \phi^a(\lambda_a) \sum_{r=0}^{n_1} \frac{A_r^D}{(\lambda - a)^{r+1}} (\phi^a(\lambda_a))^{-1} \right)_-^{\text{rat}} = \frac{\phi_0^a A_{n_1}^D (\phi_0^a)^{-1}}{\lambda - a} \quad (6.27)$$

and it suffices to choose $A_{n_1}^D = A_0^{ZM}$ to see that this is equal to

$$V_{-1}^{ZM,a}(\lambda) = \left(\phi^a(\lambda_a) \frac{A_0^{ZM}}{\lambda - a} (\phi^a(\lambda_a))^{-1} \right)_{-}^{\text{rat}} = \frac{\phi_0^a A_0^{ZM} (\phi_0^a)^{-1}}{\lambda - a}. \quad (6.28)$$

However, the crucial point is that $Q^{ZM,a}(\lambda_a)$ and $Q^{D,a}(\lambda_a)$ are constructed as orbits around different elements in general so the phase space is different in general. This means that the previous identification only gives *some* of the fields of the Dickey case which happen to be identifiable with the full phase space for ZM. The “converse” is not true in general. The Dickey case can only be seen as a higher flow in the ZM hierarchy if we construct it around a special element of the form $\sum_{r=0}^{n_1} \frac{A_r^D}{(\lambda - a)^{r+1}}$ with $A_r^D = 0$ for $r = 1, \dots, n_1$ and $A_{n_1}^D = A_0^{ZM}$. In that case, we see that

$$V_n^{D,a}(\lambda) = V_{n+n_1}^{ZM,a}(\lambda) \quad (6.29)$$

so that the two hierarchies simply correspond to shifting the starting point in the elementary times t_j^a . This discussion was local in the sense that we looked at a typical pole a . Of course, similar conclusions hold around the other poles. If one assembles them to obtain compound times, then the situation is similar but technically more complicated. The summary is that in general, the Dickey case is a genuine generalisation of the ZM case unless it is constructed as an orbit around a specific element dictated by the ZM element. Of course, this comparison extends to the corresponding Lagrangians since the building blocks are the same as for the Lax matrices.

6.2. Trigonometric Zakharov–Mikhailov models. We can repeat the construction of the previous subsection but with the rational r -matrix replaced by the trigonometric one. To the best of our knowledge, this produces for the first time a new class of models which we call *trigonometric Zakharov–Mikhailov models*.

For conciseness, we simply illustrate this on the simplest example of simple poles in the data (6.4). To derive the elementary Lax matrices, we need to use the trigonometric formula in Proposition 3.8 which brings interesting differences compared to the rational case, already for the lowest times $t_{-1}^{a_i}$ and $t_{-1}^{b_j}$. With $Q_{-1}^{a_i} = \phi_0^{a_i} A_{i0} (\phi_0^{a_i})^{-1}$ and $Q_{-1}^{b_j} = \phi_0^{b_j} B_{j0} (\phi_0^{b_j})^{-1}$, the corresponding elementary Lax matrices read

$$V_{-1}^{a_i}(\lambda) = -\frac{a_i Q_{-1}^{a_i}}{\lambda - a_i} - \left(P^- + \frac{1}{2} P^0 \right) Q_{-1}^{a_i}, \quad (6.30)$$

$$V_{-1}^{b_j}(\lambda) = -\frac{b_j Q_{-1}^{b_j}}{\lambda - b_j} - \left(P^- + \frac{1}{2} P^0 \right) Q_{-1}^{b_j}. \quad (6.31)$$

It will be convenient to introduce the following notations, for $M \in \mathfrak{gl}_N$:

$$\left(P^+ + \frac{1}{2} P^0 \right) M = M^>, \quad \left(P^- + \frac{1}{2} P^0 \right) M = M^<. \quad (6.32)$$

In particular $M = M^> + M^<$. We derive from our general formula the following elementary Lagrangian:

$$\begin{aligned} \mathcal{L}_{-1-1}^{a_i, b_j} = & \text{Tr} \left((\phi_0^{a_i})^{-1} \partial_{t_{-1}^{b_j}} \phi_0^{a_i} A_{i0} - (\phi_0^{b_j})^{-1} \partial_{t_{-1}^{a_i}} \phi_0^{b_j} B_{j0} \right) \\ & - \frac{b_j}{b_j - a_i} \text{Tr} \left(\phi_0^{a_i} A_{i0} (\phi_0^{a_i})^{-1} \phi_0^{b_j} B_{j0} (\phi_0^{b_j})^{-1} \right) + \text{Tr} \left[\phi_0^{a_i} A_{i0} (\phi_0^{a_i})^{-1} \left(\phi_0^{b_j} B_{j0} (\phi_0^{b_j})^{-1} \right)^< \right]. \end{aligned} \quad (6.33)$$

The last term represents the main difference with the rational case, see (6.10).

We now show that the so-called anisotropic chiral model presented in Section 6 of [FR] can be obtained as a particular case of our trigonometric ZM Lagrangians and ZS Lax matrices. We will refer to it as *anisotropic Faddeev-Reshetikhin model* to avoid the confusion with the “anisotropic chiral model” terminology used in [FR] which would assume that we parametrise the currents differently from our coadjoint parametrization, see (6.34).

We proceed in two steps. First, we specialise our data as follows: in (6.3), we take $P_1 = P_2 = 1$ and write $a_1 = a$ and $b_1 = b$; in (6.4), we simply write

$$F(\lambda) = \frac{A}{\lambda - a} + \frac{B}{\lambda - b}.$$

We also restrict $\mathfrak{g}$ to $\mathfrak{sl}_2$.³ Second, we apply the automorphism discussed in “Appendix A” to make the connection with [FR] easier. Let us denote for convenience $t_{-1}^a = \xi$, $t_{-1}^b = \eta$,

$$Q_{-1}^a = \phi_0^a A (\phi_0^a)^{-1} \equiv J_0, \quad Q_{-1}^b = \phi_0^b B (\phi_0^b)^{-1} \equiv J_1, \quad (6.34)$$

and the Lax pair (6.30),

$$V_{-1}^a(\lambda) \equiv U(\lambda) = -\frac{aJ_0}{\lambda - a} - J_0^<, \quad (6.35)$$

$$V_{-1}^b(\lambda) \equiv V(\lambda) = -\frac{bJ_1}{\lambda - b} - J_1^<. \quad (6.36)$$

The Lagrangian (6.33) becomes

$$\mathcal{L}_{\text{aFR}} = \text{Tr} \left((\phi_0^a)^{-1} \partial_\eta \phi_0^a A - (\phi_0^b)^{-1} \partial_\xi \phi_0^b B - \frac{b}{b-a} J_0 J_1 + J_0 J_1^< \right). \quad (6.37)$$

Varying with respect to ϕ_0^a and ϕ_0^b , the EL equations read⁴

$$\partial_\eta J_0 = \left[-\frac{bJ_1}{a-b} - J_1^<, J_0 \right], \quad \partial_\xi J_1 = \left[-\frac{aJ_0}{b-a} - J_0^<, J_1 \right]. \quad (6.38)$$

Projecting on the basis $J_{0,1} = J_{0,1}^+ \sigma_+ + J_{0,1}^- \sigma_- + J_{0,1}^3 \sigma_3$, we get

$$\begin{cases} \partial_\eta J_0^+ = \frac{2b}{a-b} J_1^+ J_0^3 - \frac{a+b}{a-b} J_1^3 J_0^+, \\ \partial_\eta J_0^- = \frac{2a}{b-a} J_1^- J_0^3 + \frac{a+b}{a-b} J_1^3 J_0^-, \\ \partial_\eta J_0^3 = \frac{b}{b-a} J_1^+ J_0^- + \frac{a}{a-b} J_1^- J_0^+, \end{cases} \quad \begin{cases} \partial_\xi J_1^+ = \frac{2a}{b-a} J_0^+ J_1^3 + \frac{a+b}{a-b} J_0^3 J_1^+, \\ \partial_\xi J_1^- = \frac{2b}{a-b} J_0^- J_1^3 - \frac{a+b}{a-b} J_0^3 J_1^-, \\ \partial_\xi J_1^3 = \frac{a}{a-b} J_0^+ J_1^- + \frac{b}{b-a} J_0^- J_1^+. \end{cases} \quad (6.39)$$

³ Thus, it would be more accurate to say that we derive the $\mathfrak{sl}_2$ anisotropic FR model, as opposed to the $su(2)$ version of [FR]. This is not important for our considerations here.

⁴ The property $\text{Tr}(J_0 J_1^<) = \text{Tr}(J_0^> J_1)$ is useful in deriving the EL equations.

The residue at infinity of the zero curvature equation for the Lax pair (6.35)–(6.36) yields the equation $-\partial_\eta J_0^< + \partial_\xi J_1^< + [J_0^<, J_1^<] = 0$ in addition to (6.38). However, when projecting, one can see that this is a consequence of the system (6.39).

To make the comparison with the equations for the fields $S_{1,2,3}$ and $T_{1,2,3}$ used in [FR], we use the automorphism mentioned above and express the final answer using the Pauli matrices $\sigma_{1,2,3}$. We also implement the changes $\lambda \rightarrow e^{2\lambda}$, $a \rightarrow e^{2a}$, $b \rightarrow e^{-2a}$ to go from rational to hyperbolic parametrisation. We find

$$e^{\lambda/2\sigma_3} U(e^{2\lambda}) e^{-\lambda/2\sigma_3} = -\frac{1}{2} \left[w_1(\lambda - a) \frac{1}{2} (e^a J_0^+ + e^{-a} J_0^-) \sigma_1 \right. \\ \left. + w_2(\lambda - a) \frac{i}{2} (e^a J_0^+ - e^{-a} J_0^-) \sigma_2 + w_3(\lambda - a) J_0^3 \sigma_3 \right],$$

and

$$e^{\lambda/2\sigma_3} V(e^{2\lambda}) e^{-\lambda/2\sigma_3} = -\frac{1}{2} \left[w_1(\lambda + a) \frac{1}{2} (e^{-a} J_1^+ + e^a J_1^-) \sigma_1 \right. \\ \left. + w_2(\lambda + a) \frac{i}{2} (e^{-a} J_1^+ - e^a J_1^-) \sigma_2 + w_3(\lambda + a) J_1^3 \sigma_3 \right],$$

where $w_1(\lambda) = w_2(\lambda) = \frac{1}{\sinh \lambda}$, $w_3(\lambda) = \coth \lambda$. It remains to compare with the Lax operator (6.22) in [FR] and remember that they work with x and t instead of the light-cone coordinates ξ and η . This leads to the identifications

$$\begin{cases} S_1 = -\frac{1}{4} (e^a J_0^+ + e^{-a} J_0^-), \\ S_2 = -\frac{i}{4} (e^a J_0^+ - e^{-a} J_0^-), \\ S_3 = -\frac{1}{2} J_0^3, \end{cases} \quad \begin{cases} T_1 = -\frac{1}{4} (e^{-a} J_1^+ + e^a J_1^-), \\ T_2 = -\frac{i}{4} (e^{-a} J_1^+ - e^a J_1^-), \\ T_3 = -\frac{1}{2} J_1^3. \end{cases} \quad (6.40)$$

Using (6.40), eqs (6.39) become

$$\partial_\eta S_a = 2i \sum_{b,c} \epsilon^{abc} w_b(2a) T_b S_c, \quad (6.41)$$

$$\partial_\xi T_a = -2i \sum_{b,c} \epsilon^{abc} w_b(2a) S_b T_c, \quad (6.42)$$

which are of the same form as (6.26)–(6.27) in [FR] when moving from the light-cone coordinates ξ, η to the coordinates x, t .

6.3. Deformed Gross–Neveu models. Here, we show how to produce the Lax pair and Lagrangian for the deformed Gross–Neveu model discussed in [ABW, Section 16.2] (see also [By] and references therein for the particular case of rank $M = 1$) as a particular case of our construction. The deformation is controlled by the r -matrix in the potential term which appears naturally in our construction. In fact, more than just the single Lagrangian and its Lax pair, we can in principle generate all the elementary Lagrangians in the whole Lagrangian multiform and all the elementary Lax pairs for the hierarchy containing this model as its main representative. This explains the origin of the integrability of such a class of models observed in [ABW, By] and is seen to be a particular case of our construction.

The idea is to apply a *reduction*, in the spirit of [Mik], to a Zakharov–Mikhailov model. The r -matrix could in principle be any skew-symmetric solution of the CYBE as we have already mentioned. Of course, if we want to resort to our explicit formulas for Lagrangians or Lax matrices, then it will be either the rational or trigonometric one since we have given an explicit construction only in those cases. Nevertheless, we will write most results without specifying the r -matrix to emphasize this observation.

Choose the data in (6.3)–(6.4) as follows

$$S = \{a, a^*\}, \quad a \notin \mathbb{R}, \quad \mathfrak{g} = \mathfrak{gl}_N, \quad (6.43)$$

$$F(\lambda) = \frac{A}{\lambda - a} - \frac{A^\dagger}{\lambda - a^*}. \quad (6.44)$$

In particular, we chose $N_a = 1 = N_{a^*}$. As mentioned, we want to use the idea of reduction which we implement as a reality condition on the objects of the theory. Writing

$$Q^a(\lambda_a) = \sum_{k=-1}^{\infty} Q_k^a(\lambda - a)^k \quad (6.45)$$

and

$$Q^{a^*}(\lambda_{a^*}) = \sum_{k=-1}^{\infty} Q_k^{a^*}(\lambda - a^*)^k \quad (6.46)$$

we require $Q_k^{a^*} = -(Q_k^a)^\dagger$ for all $k \geq -1$. Accordingly, at the group level, we require that when writing

$$\varphi_a^{-1}(\lambda_a) = \sum_{k=0}^{\infty} \tilde{\varphi}_k^a(\lambda - a)^k \quad (6.47)$$

and

$$\varphi_{a^*}(\lambda_{a^*}) = \sum_{k=0}^{\infty} \varphi_k^{a^*}(\lambda - a^*)^k \quad (6.48)$$

we must have $\tilde{\varphi}_k^a = (\varphi_k^{a^*})^\dagger$ for all $k \geq 0$. Then, for any skew-symmetric r -matrix which is well-defined at $\lambda = a$ and $\mu = a^*$, a direct computation gives

$$\begin{aligned} \mathcal{L}_{-1-1}^{a,a^*} = & \operatorname{Tr} \left(\left((\varphi_0^{a^*})^\dagger \partial_{t_{-1}^a} \varphi_0^a A + (\varphi_0^a)^\dagger \partial_{t_{-1}^{a^*}} \varphi_0^{a^*} A^\dagger \right) \right. \\ & \left. + \operatorname{Tr}_{12} \left(r_{12}(a, a^*) \left(\varphi_0^a A \left((\varphi_0^{a^*})^\dagger \right) \right)_1 \left(\varphi_0^a A \left((\varphi_0^{a^*})^\dagger \right) \right)_2^\dagger \right) \right). \end{aligned} \quad (6.49)$$

It remains to choose A as a rank M matrix and parametrize it as $A = (uv)^\dagger$ where u is a constant $N \times M$ matrix and v is a constant $M \times N$ matrix ($M \leq N$). Then, setting $U = \varphi_0^{a^*} u$, $V = v (\varphi_0^a)^\dagger$, $t_{-1}^a = \bar{z}$ and $t_{-1}^{a^*} = z$, we get

$$\mathcal{L}_{-1-1}^{a,a^*} = \operatorname{Tr} \left(V \partial_{\bar{z}} U + U^\dagger \partial_z V^\dagger \right) + \operatorname{Tr}_{12} \left(r_{12}(a, a^*) (UV)_1 (UV)_2^\dagger \right). \quad (6.50)$$

This is the Lagrangian given in [ABW] (without the covariant derivative), with the relation to their notation being $r_a(A)_1 = \text{Tr}_2(r_{12}(a, a^*)A_2)$ so that the potential term reads

$$\text{Tr}_{12} \left(r_{12}(a, a^*) (UV)_1^\dagger (UV)_2 \right) = \text{Tr} \left(r_a(UV) (UV)^\dagger \right).$$

The interpretation of the parameter appearing in the r -matrix (a here, s in [ABW]) is clear in our context: it corresponds to the pole structure of the constant matrix in our data (6.4).

The corresponding Lax pair is derived from (3.14) and reads, with $K = UV$,

$$V_{-1}^a(\lambda) = \text{Tr}_2(r_{12}(\lambda, a)K_2^\dagger), \quad V_{-1}^{a^*}(\lambda) = -\text{Tr}_2(r_{12}(\lambda, a^*)K_2), \quad (6.51)$$

and coincides with the Lax connection (16.7) in [ABW]. Hence, the zero curvature equation yields

$$\begin{aligned} \partial_z \text{Tr}_2 \left(\text{res}_\lambda^a r_{12}(\lambda, a)K_2^\dagger \right) &= \left[\text{Tr}_2 \left(\text{res}_\lambda^a r_{12}(\lambda, a)K_2^\dagger \right), \text{Tr}_2 \left(r_{12}(a, a^*)K_2 \right) \right], \\ \partial_{\bar{z}} \text{Tr}_2 \left(\text{res}_\lambda^{a^*} r_{12}(\lambda, a^*)K_2 \right) &= \left[\text{Tr}_2 \left(r_{12}(\lambda, a)K_2^\dagger \right), \text{Tr}_2 \left(\text{res}_\lambda^{a^*} r_{12}(\lambda, a^*)K_2 \right) \right], \end{aligned}$$

which reduces to⁵

$$\begin{aligned} \partial_z K^\dagger &= \left[K^\dagger, \text{Tr}_2 \left(r_{12}(a, a^*)K_2 \right) \right], \\ \partial_{\bar{z}} K_2 &= \left[\text{Tr}_2 \left(r_{12}(\lambda, a)K_2^\dagger \right), K \right]. \end{aligned}$$

In our opinion, it is rather beautiful that our generating Lagrangian multiform produces this class of models which was originally obtained via a completely different method, related to $4d$ Chern–Simons theory (see the conclusion for details and references). Unlike the latter method which necessarily focuses on a single Lagrangian at a time, we can also obtain all the Lagrangians corresponding to the higher commuting flows of the hierarchy, if desired.

7. Coupling Integrable Hierarchies Together

To show the flexibility of the construction, we explain by way of two examples how we can couple integrable field theories together in a simple way. The reader familiar with integrable hierarchies will recognize the procedure of assembling elementary time flows and the corresponding Lax matrices into linear combinations. What we gain here is the possibility to derive the corresponding Lagrangian (multiform) systematically for the new model as well. The procedure is an analog in the ultralocal case of the construction presented in [DLMV1] for a class of non ultralocal field theories. Unlike the latter, the coupling here is at the level of an entire hierarchy. We give an example in the rational class and one in the trigonometric class of models. In the rational class, we couple together the AKNS hierarchy with the hierarchy of the Faddeev–Reshetikhin model (the simplest instance of a ZM model). In the trigonometric class, we couple the sine-Gordon hierarchy as discussed in Sect. 5 with the hierarchy of the anisotropic Faddeev–Reshetikhin model as presented in Sect. 6.2. In each case, for conciseness, we present all the details for the lowest levels of the hierarchy but it should be clear by now that one can extract higher levels (Lagrangians and Lax matrices) systematically if desired.

⁵ This is true for r -matrices whose singular part at $\lambda = \mu$ is of the form $\frac{f(\mu)}{\mu - \lambda} P_{12}$, which is the case for the rational and trigonometric matrices we work with here.

7.1. AKNS-FR hierarchy. To couple models in the AKNS hierarchy with models in the simplest ZM hierarchy (with two poles), we assemble the corresponding data as

$$S = \{a, -a, \infty\}, \quad a \in \mathbb{C}^\times, \quad N_a = N_b = 1, \quad N_\infty = 0, \quad \mathfrak{g} = \mathfrak{sl}_2, \quad (7.1)$$

and we choose

$$F(\lambda) = -i\alpha\sigma_3 + \frac{A}{\lambda - a} + \frac{B}{\lambda + a} \equiv \alpha F^{AKNS}(\lambda) + F^{FR}(\lambda), \quad (7.2)$$

where A, B are constant $\mathfrak{sl}_2$ matrices. The parameter α is the coupling between the two theories: $\alpha = 0$ gives a pure FR theory while sending α to infinity produces a pure AKNS hierarchy. The effect of multiplying $F^{AKNS}(\lambda) = -i\sigma_3$ by α is to yield $Q^\infty(\lambda_\infty) = \alpha Q(\lambda)$ where $Q(\lambda)$ is the AKNS series (4.4). Hence the Lax matrix $V_n^\infty(\lambda)$ is equal to the AKNS Lax matrix $V_n(\lambda)$ multiplied by α . With this in mind, we have for instance $V_1^\infty(\lambda) = -i\alpha\lambda\sigma_3 + \alpha Q_1$.

For simplicity, we illustrate the coupling by looking at the two main models in each hierarchy (NLS in AKNS and FR in ZM), i.e. by considering the Lax pair

$$V_{-1}^a(\lambda) + V_1^\infty(\lambda) \equiv U(\lambda), \quad V_{-1}^b(\lambda) + V_2^\infty(\lambda) \equiv V(\lambda), \quad (7.3)$$

with associated times ξ and η respectively. This choice of Lax pair corresponds to assembling the four flows t_1^∞, t_2^∞ (AKNS) and t_{-1}^a, t_{-1}^{-a} (FR) such that $\partial_\xi = \partial_{t_{-1}^a} + \partial_{t_1^\infty}$ and $\partial_\eta = \partial_{t_{-1}^{-a}} + \partial_{t_2^\infty}$.

Denoting $Q_{-1}^a = J_0$ and $Q_{-1}^{-a} = J_1$ and recalling the above comments on the effect of multiplying by α , we have

$$U(\lambda) = \frac{J_0}{\lambda - a} - i\alpha\lambda\sigma_3 + \alpha Q_1 \equiv U_{FR}(\lambda) + \alpha U_{NLS}(\lambda), \quad (7.4)$$

$$V(\lambda) = \frac{J_1}{\lambda - b} - i\lambda^2\alpha\sigma_3 + \lambda\alpha Q_1 + \alpha Q_2 \equiv V_{FR}(\lambda) + \alpha V_{NLS}(\lambda). \quad (7.5)$$

The zero curvature equation $\partial_\eta U(\lambda) - \partial_\xi V(\lambda) + [U(\lambda), V(\lambda)] = 0$ yields the following four (matrix) equations by looking at the residue at $\lambda = a, \lambda = -a, \lambda = \infty$ and at the constant term in the $1/\lambda$ expansion respectively,

$$\partial_\eta J_0 + \frac{1}{2a} [J_0, J_1] + \alpha [J_0, V_{NLS}(a)] = 0, \quad (7.6)$$

$$\partial_\xi J_1 + \frac{1}{2a} [J_0, J_1] - \alpha [U_{NLS}(-a), J_1] = 0, \quad (7.7)$$

$$\alpha \partial_\xi Q_1 + i\alpha^2 [\sigma_3, Q_2] + i\alpha [J_0, \sigma_3] = 0, \quad (7.8)$$

$$\alpha \partial_\eta Q_1 - \alpha \partial_\xi Q_2 + \alpha^2 [Q_1, Q_2] - i\alpha\alpha [J_0, \sigma_3] - i\alpha [\sigma_3, J_1] + \alpha [J_0, Q_1] = 0. \quad (7.9)$$

Setting $\alpha = 0$, (7.6) and (7.7) gives the FR version of the principal chiral model [ZM2, FR] which is usually written as

$$\partial_\eta J_0 + \partial_\xi J_1 + \frac{1}{a} [J_0, J_1] = 0, \quad \partial_\eta J_0 - \partial_\xi J_1 = 0. \quad (7.10)$$

In the limit $\alpha \rightarrow \infty$ (recall that ∂_ξ scales like $\alpha \partial_{t_1}$ and ∂_η scales like $\alpha \partial_{t_2}$, with t_1, t_2 the NLS times), we see that (7.8)–(7.9) yield the NLS system (4.13)–(4.14)

$$\partial_{t_1} Q_1 + i[\sigma_3, Q_2] = 0, \quad \partial_{t_2} Q_1 - \partial_{t_1} Q_2 + [Q_1, Q_2] = 0. \quad (7.11)$$

The Lagrangian of this coupled model is obtained by adding the NLS Lagrangian $\mathcal{L}_{12}^{\infty\infty}$ (which is $\mathcal{L}_{12}$ in (4.11) properly rescaled)

$$\begin{aligned} \mathcal{L}_{12}^{\infty\infty} &= \frac{\alpha}{2} (f_1 \partial_{t_2}^\infty e_1 - e_1 \partial_{t_2}^\infty f_1) - \frac{\alpha}{2} \sum_{j=1}^2 (f_j \partial_{t_1}^\infty e_{2-j+1} - e_j \partial_{t_1}^\infty f_{2-j+1}) \\ &\quad - \alpha^2 \left(2i e_2 f_2 + e_1^2 f_1^2 \right), \end{aligned} \quad (7.12)$$

the FR Lagrangian $\mathcal{L}_{-1-1}^{a,-a}$

$$\mathcal{L}_{-1-1}^{a,-a} = \text{Tr} \left[(\phi_0^a)^{-1} \partial_{t_1}^{-a} \phi_0^a A - (\phi_0^{-a})^{-1} \partial_{t_1}^{-a} \phi_0^{-a} B - \frac{J_0 J_1}{2a} \right], \quad (7.13)$$

and the following two mixed elementary Lagrangians (discarding some irrelevant total derivatives),

$$\begin{aligned} \mathcal{L}_{-12}^{a,\infty} &= \text{Tr} \left[(\phi_0^a)^{-1} \partial_{t_2}^\infty \phi_0^a A \right] - \frac{\alpha}{2} \sum_{j=1}^2 (f_j \partial_{t_1}^a e_{2-j+1} - e_j \partial_{t_1}^a f_{2-j+1}) \\ &\quad - \alpha \text{Tr} [J_0 V_{NLS}(a)], \end{aligned} \quad (7.14)$$

$$\mathcal{L}_{1-1}^{\infty,-a} = \frac{\alpha}{2} (f_1 \partial_{t_1}^{-a} e_1 - e_1 \partial_{t_1}^{-a} f_1) - \text{Tr} \left[(\phi_0^{-a})^{-1} \partial_{t_1}^\infty \phi_0^{-a} B + \alpha J_1 U_{NLS}(-a) \right]. \quad (7.15)$$

Summing we get our Lagrangian for the coupled model

$$\begin{aligned} \mathcal{L}_{\text{NLS-FR}} &= \text{Tr} \left[(\phi_0^a)^{-1} \partial_\eta \phi_0^a A - (\phi_0^{-a})^{-1} \partial_\xi \phi_0^{-a} B \right] \\ &\quad + \frac{\alpha}{2} (f_1 \partial_\eta e_1 - e_1 \partial_\eta f_1) - \frac{\alpha}{2} \sum_{j=1}^2 (f_j \partial_\xi e_{2-j+1} - e_j \partial_\xi f_{2-j+1}) \\ &\quad - \alpha^2 \left(2i e_2 f_2 + e_1^2 f_1^2 \right) - \text{Tr} \left[\frac{J_0 J_1}{2a} + \alpha J_0 V_{NLS}(a) - \alpha J_1 U_{NLS}(-a) \right]. \end{aligned} \quad (7.16)$$

It can be checked directly that the variations with respect to $\phi_0^a, \phi_0^{-a}, e_2, f_2$ and e_1, f_1 gives (7.6), (7.7), (7.8) and (7.9) respectively.

7.2. sG-aFR hierarchy. The same strategy can of course be applied in the trigonometric case and we illustrate this by assembling the data of the sine-Gordon (sG) hierarchy as in Sect. 5 with that of the anisotropic Faddeev-Reshetikhin (aFR) model as in Sect. 6.2, in the following way

$$S = \{0, a, b, \infty\}, \quad a, b \in \mathbb{C}^\times, \quad N_0 = N_a = N_b = N_\infty = 1, \quad \mathfrak{g} = \mathfrak{sl}_2, \quad (7.17)$$

and we choose

$$F(\lambda) = \frac{i\beta}{2} \left(\frac{1}{\lambda} \sigma_+ + \sigma_- - \sigma_+ - \lambda \sigma_- \right) + \frac{A}{\lambda - a} + \frac{B}{\lambda - b}, \quad (7.18)$$

where A, B are constant $\mathfrak{sl}_2$ matrices and it is understood that $b = 1/a$. We keep b instead of $1/a$ as it makes notations lighter but all calculations are done with $b = 1/a$. The parameter β is the coupling between the two theories: $\beta = 0$ gives a pure aFR theory while sending β to infinity produces a pure sG model.

To illustrate the procedure on the easiest case, we choose the main representative of each hierarchy, i.e. we consider the Lax pair (recall from Sect. 6.2 that we set $J_0 = Q_{-1}^a = \phi_0^a A (\phi_0^a)^{-1}$ and $J_1 = Q_{-1}^b = \phi_0^b B (\phi_0^b)^{-1}$)

$$\begin{aligned} U(\lambda) &= V_{-1}^a(\lambda) + V_0^0(\lambda) = \frac{-aJ_0^> - \lambda J_0^<}{\lambda - a} - \frac{i\beta}{4} \begin{pmatrix} -C_1^0 & 2e^{i\frac{\eta}{2}}/\lambda \\ 2e^{-i\frac{\eta}{2}} & C_1^0 \end{pmatrix} \\ &\equiv U_{\text{aFR}}(\lambda) + \beta U_{\text{sG}}(\lambda), \end{aligned} \quad (7.19)$$

$$\begin{aligned} V(\lambda) &= V_{-1}^b(\lambda) + V_0^\infty(\lambda) = \frac{-bJ_1^> - \lambda J_1^<}{\lambda - b} - \frac{i\beta}{4} \begin{pmatrix} B_1^\infty 2e^{-i\frac{\eta}{2}} \\ 2\lambda e^{i\frac{\eta}{2}} & -B_1^\infty \end{pmatrix} \\ &\equiv V_{\text{aFR}}(\lambda) + \beta V_{\text{sG}}(\lambda), \end{aligned} \quad (7.20)$$

with associated times ξ and η respectively. This corresponds to assembling the two sG times t_0^0, t_0^∞ with the two aFR times t_{-1}^a, t_{-1}^b such that $\partial_\xi = \partial_{t_{-1}^a} + \partial_{t_0^0}$ and $\partial_\eta = \partial_{t_{-1}^b} + \partial_{t_0^\infty}$. The zero curvature equation $\partial_\eta U(\lambda) - \partial_\xi V(\lambda) + [U(\lambda), V(\lambda)] = 0$ yields the following four equations by looking at the residue at $\lambda = 0, \lambda = \infty, \lambda = a, \lambda = b$ respectively,

$$u_\eta + \beta B_1^\infty = -2iJ_1^3, \quad (7.21)$$

$$u_\xi + \beta C_1^0 = -2iJ_0^3, \quad (7.22)$$

$$\partial_\eta J_0 = [V_{\text{aFR}}(a), J_0] - \beta [J_0, V_{\text{sG}}(a)], \quad (7.23)$$

$$\partial_\xi J_1 = [U_{\text{aFR}}(b), J_1] + \beta [U_{\text{sG}}(b), J_1]. \quad (7.24)$$

Equations (7.21)–(7.22) should be compared with the first two equations in (5.20) and (7.23)–(7.24) should be compared with (6.38). The last independent equation contained in the zero curvature can be obtained for instance by setting $\lambda = 1$. It can be shown

that only the component σ_3 gives an equation that is not a consequence of those already written. It takes the form

$$\begin{aligned} & \frac{i\beta}{4} \left(\partial_\eta C_1^0 + \partial_\xi B_1^\infty \right) - \frac{i\beta^2}{2} \sin u \\ & + \frac{1}{2} \frac{a+1}{a-1} \left(\partial_\eta J_0^3 + \partial_\xi J_1^3 \right) + \frac{1}{(a-1)(b-1)} (aJ_0^+ J_1^- - bJ_0^- J_1^+) \\ & + \frac{i\beta}{2(a-1)} \left(J_0^- e^{-i\frac{u}{2}} - aJ_0^+ e^{i\frac{u}{2}} - J_1^+ e^{-i\frac{u}{2}} + aJ_1^- e^{i\frac{u}{2}} \right) = 0. \end{aligned} \quad (7.25)$$

We can use (7.21)–(7.24) to cast this equation in the following more suggestive form which shows the coupling between sG and the aFR currents

$$i \left(\partial_\eta J_0^3 + \partial_\xi J_1^3 \right) + u_{\eta\xi} + \beta^2 \sin u + \frac{\beta}{2} \left((J_0^- - J_1^+) e^{-i\frac{u}{2}} + a(J_1^- - J_0^+) e^{i\frac{u}{2}} \right) = 0. \quad (7.26)$$

We can derive the Lagrangian producing (7.21)–(7.24) and (7.26) by adding the sine-Gordon Lagrangian (5.26) (with appropriate inclusion of β)

$$\mathcal{L}_{\text{sG}} = \frac{\beta}{4} C_1^0 \partial_{t_0^\infty} u + \frac{\beta}{4} B_1^\infty \partial_{t_0} u - \frac{\beta^2}{4} \left(e^{iu} + e^{-iu} - C_1^0 B_1^\infty \right), \quad (7.27)$$

the anisotropic FR Lagrangian (6.37)

$$\mathcal{L}_{\text{aFR}} = \text{Tr} \left((\phi_0^a)^{-1} \partial_{t_{-1}^b} \phi_0^a A - (\phi_0^b)^{-1} \partial_{t_{-1}^a} \phi_0^b B - \frac{b}{b-a} J_0 J_1 + J_0 J_1^< \right), \quad (7.28)$$

and the following two mixed elementary Lagrangians

$$\mathcal{L}_{-10}^{a,\infty} = \text{Tr} \left((\phi_0^a)^{-1} \partial_{t_0^\infty} \phi_0^a A + \frac{\beta}{4} C_1^0 \partial_{t_{-1}^a} u - \beta J_0 V_{\text{sG}}(a) \right), \quad (7.29)$$

$$\mathcal{L}_{0-1}^{0,b} = \text{Tr} \left(\frac{\beta}{4} C_1^0 \partial_{t_{-1}^b} u - (\phi_0^b)^{-1} \partial_{t_0} \phi_0^b B + \beta J_1 U_{\text{sG}}(b) \right). \quad (7.30)$$

We obtain

$$\begin{aligned} \mathcal{L}_{\text{sG-aFR}} &= \frac{\beta}{4} C_1^0 \partial_\eta u + \frac{\beta}{4} B_1^\infty \partial_\xi u + \text{Tr} \left((\phi_0^a)^{-1} \partial_\eta \phi_0^a A - (\phi_0^b)^{-1} \partial_\xi \phi_0^b B \right) \\ &- \frac{\beta^2}{4} \left(e^{iu} + e^{-iu} - C_1^0 B_1^\infty \right) - \text{Tr} \left(\frac{b}{b-a} J_0 J_1 + J_0 J_1^< + \beta J_0 V_{\text{sG}}(a) - \beta J_1 U_{\text{sG}}(b) \right). \end{aligned} \quad (7.31)$$

The variation with respect to C_1^0 , B_1^∞ , ϕ_0^a , ϕ_0^b , and u gives (7.21), (7.22), (7.23), (7.24), and (7.25) respectively.

8. Discussion and Conclusion

By introducing a certain generating Lagrangian multiform, we were able to relate two important but so far separate aspects of integrable systems: the well established theory of the classical r -matrix and the comparatively much newer framework of Lagrangian multiforms. In doing so, we bring closer together the vast amount of results in the Hamiltonian approach to integrable systems and the Lagrangian approach in the form advocated in the seminal paper [LN]. A rich byproduct of this effort is that the generating Lagrangian multiform and its accompanying generating Lax equation and zero curvature equation provide a systematic framework to construct integrable hierarchies of field theories, both in terms of Lagrangians and of Lax matrices. This was illustrated at length over many examples, both known and new. As already emphasised in the introduction, this versatility to accommodate a very large class of examples stems from the fact that we work in the adèlic framework.

The most immediate open question that comes to mind relates to the restrictions imposed on the classical r -matrix appearing in the generating Lagrangian multiform. Certain aspects of our construction appear to remain true under only the assumption that r is a solution of the CYBE (1.1). In particular, the restriction to the rational or trigonometric case that we studied in detail only played a role in the explicit construction of the projectors associated to the decomposition of the Lie algebra of $\mathfrak{g}$ -valued adèles. It is easy to imagine that one could use a more general skew-symmetric r -matrix provided similar technicalities can be dealt with. Specifically, given a solution r of the CYBE, one would like to establish results along the following lines:

- Define a pair of linear operators on the Lie algebra of $\mathfrak{g}$ -valued adèles $\mathcal{A}_\lambda(\mathfrak{g})$ as

$$\pi_\pm : \mathcal{A}_\lambda(\mathfrak{g}) \longrightarrow \mathcal{A}_\lambda(\mathfrak{g}), \quad X(\lambda) \longmapsto ((\pi_\pm X)_a(\lambda_a))_{a \in \mathbb{C}P^1} \quad (8.1)$$

with formulas similar to e.g. (2.21).

- Show that the linear maps $\pi_\pm$ so defined are projection operators onto complementary subspaces of $\mathcal{A}_\lambda(\mathfrak{g})$, i.e. $(\pi_\pm)^2 = \pi_\pm$ and $\pi_+ + \pi_- = \text{id}$.

- Show that the images $\pi_\pm \mathcal{A}_\lambda(\mathfrak{g})$ of the projection operators $\pi_\pm$ are both Lie subalgebras of $\mathcal{A}_\lambda(\mathfrak{g})$ and are isotropic with respect to the bilinear form analogous to that defined in (2.2).

If one could accomplish this then it would follow that one would have a direct sum decomposition of $\mathcal{A}_\lambda(\mathfrak{g})$ into complementary Lagrangian Lie subalgebras

$$\mathcal{A}_\lambda(\mathfrak{g}) = \pi_+ \mathcal{A}_\lambda(\mathfrak{g}) \dot{+} \pi_- \mathcal{A}_\lambda(\mathfrak{g}).$$

The corresponding r -matrix would be defined as $r := \pi_+ - \pi_- \in \text{End } \mathcal{A}_\lambda(\mathfrak{g})$ and would presumably have a kernel of the form $((\iota_{\mu_b} \iota_{\lambda_a} + \iota_{\lambda_a} \iota_{\mu_b}) r_{12}(\lambda, \mu))_{a,b \in \mathbb{C}P^1}$. We could then use this kernel into our generating Lagrangian multiform and construct integrable hierarchies by the same method as we have done. One candidate to see if such a programme can be realised is the elliptic r -matrix [Sk, Be].

The other obvious restriction of the present work is the condition that r be skew-symmetric. In fact, we wrote the CYBE (1.1) in its non-skew-symmetric form on purpose. Once again, some of our results appear to hold without this assumption. This is the case for the commutativity of the vector fields (3.9) as can be seen from the proof of Proposition 3.4. The extension of our construction to the non-skew-symmetric case, hence to non-ultralocal field theories, appears rather challenging as the current form of our generating Lagrangian multiform simply does not allow for such an extension. We are currently investigating this exciting issue which promises to have connection with the

framework of classical affine Gaudin models, developed in [V1,DLMV2], that provides a unifying formalism for constructing and studying a very broad class of non-ultralocal classical integrable field theories. A first step in that direction was achieved recently [CDS] where it was shown how to incorporate the non-skew symmetric case naturally in the context of finite-dimensional integrable hierarchies.

It was shown in [V2] that classical affine Gaudin models are closely related to $4d$ mixed topological-holomorphic Chern–Simons theory introduced and studied in [Cos1, Cos2,CWY1,CWY2,CY], see also [DLMV3,BSV,LV]. In fact, $4d$ Chern–Simons theory also naturally provides a framework for constructing a very broad class of ultralocal integrable field theories (see also [Zo] for a description of ultralocal integrable field theories as affine Gaudin models). In this context, it was shown in [CSV], see also [FSY], that the rational Zakharov–Mikhailov models, one of the main classes of examples that we reproduced here, could be obtained from $4d$ Chern–Simons theory with certain line defects. However, the construction of [CSV] is, by design, able to produce only the action of a single Zakharov–Mikhailov model, as opposed to its entire hierarchy, starting from that of $4d$ Chern–Simons theory. It seems natural to wonder if such a construction, and in fact the whole $4d$ Chern–Simons approach, could be adapted to our generating Lagrangian multiform framework in order to derive entire integrable hierarchies and not just single models from this point of view.

In the simplest case of the AKNS hierarchy, the concept of Hamiltonian multiform, initially introduced in [CS2], was illustrated in [CS3]. The main idea is that it is possible to apply a version of the covariant Legendre transformation to an entire Lagrangian multiform to obtain the Hamiltonian analog of a multiform. Each coefficient of the resulting Hamiltonian multiform can be seen as a covariant Hamiltonian for the field theory described by the associated Lagrangian coefficient in the Lagrangian multiform. Important accompanying objects are the symplectic multiform and the multitime Poisson bracket which generalise to an entire hierarchy the concepts of multisymplectic form and of covariant Poisson bracket respectively. The latter are essential ingredients of the framework generally called covariant Hamiltonian field theory, see e.g. [Gi] and references therein for a very useful recent review of the many facets of this rich topic. We believe it is important to try and obtain the generating Hamiltonian multiform and related structures corresponding to our generating Lagrangian multiform. Indeed, historically, one of the driving motivations of the above mentioned covariant Hamiltonian approach to field theory has been to allow for a (canonical) quantization of field theories that removes from the start the breaking of covariance associated to the standard Hamiltonian approach. The idea of covariant Hamiltonian field theory is to use a Poisson bracket that does not suffer from the lack of covariance of the traditional Poisson bracket: a *covariant* Poisson bracket. The results of [CS2,CS3] show that one can extend this idea to a whole integrable hierarchy and that the classical r -matrix plays a key role in this “covariant” context, see also [CS1,CSV]. The hope is that this could allow one to use the nice features of integrability encoded in the passage from the classical r -matrix to the quantum R -matrix, to fully implement the idea of covariant canonical quantization for such field theories.

Finally, our work also opens the possibility for quantization using another route: combining Feynman’s path integral ideas with a Lagrangian multiform, thus taking advantage again of integrability features now encoded in a Lagrangian object entering the path integral. This tantalising idea was first put forward and explored in [KN] but is still very much in its infancy.

Acknowledgements V.C and M.S. would like to acknowledge the vibrant atmosphere of the Leeds research group on Lagrangian multiforms involving F. Nijhoff, J. Richardson, D. Sleight and M. Vermeeren.

Data Availability Data sharing not applicable to this article as no datasets were generated or analysed in this work.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

A. Comparison of Trigonometric r -matrices for Sine-Gordon

For the reader's convenience, we make the connection between the trigonometric r -matrix we used in this paper and the perhaps more familiar one usually used for treating the sine-Gordon model. The former reads

$$r_{12}(\lambda, \mu) = \sigma_+ \otimes \sigma_- - \sigma_- \otimes \sigma_+ + \frac{\mu + \lambda}{\mu - \lambda} P_{12}, \quad (\text{A.1})$$

with

$$P_{12} = \frac{1}{2} (\mathbf{1} \otimes \mathbf{1} + \sigma_1 \otimes \sigma_1 + \sigma_2 \otimes \sigma_2 + \sigma_3 \otimes \sigma_3), \quad (\text{A.2})$$

while the latter, which can be found for instance in [FT, pp. 432–433], reads

$$\tilde{r}_{12}(\lambda, \mu) = \frac{\gamma}{2} \frac{\mu^2 + \lambda^2}{\lambda^2 - \mu^2} (\mathbf{1} \otimes \mathbf{1} - \sigma_3 \otimes \sigma_3) - \frac{\gamma \mu \lambda}{\lambda^2 - \mu^2} (\sigma_1 \otimes \sigma_1 + \sigma_2 \otimes \sigma_2), \quad (\text{A.3})$$

where γ is related to the coupling constant of the sine-Gordon model and is set to 1 as it is not relevant here. We relate the two matrices by showing that they both give rise to the same matrix in trigonometric form. Set $\lambda = e^{2i\alpha}$, $\mu = e^{2i\beta}$ and define

$$M(\alpha) = e^{\frac{i\alpha}{2} \sigma_3} \quad (\text{A.4})$$

with property $M(\alpha) \sigma_{\pm} M^{-1}(\alpha) = e^{\pm i\alpha} \sigma_{\pm}$. Then we have

$$\begin{aligned} & M(\alpha) \otimes M(\beta) r_{12}(\alpha, \beta) M^{-1}(\alpha) \otimes M^{-1}(\beta) \\ &= -\frac{1}{2i \sin(\alpha - \beta)} [2 (\sigma_+ \otimes \sigma_- + \sigma_- \otimes \sigma_+) + \cos(\alpha - \beta) (\sigma_3 \otimes \sigma_3 + \mathbf{1} \otimes \mathbf{1})]. \end{aligned} \quad (\text{A.5})$$

Now set instead $\lambda = e^{i\alpha}$, $\mu = e^{i\beta}$, and recall the relation $\sigma_1 \otimes \sigma_1 + \sigma_2 \otimes \sigma_2 = 2 (\sigma_+ \otimes \sigma_- + \sigma_- \otimes \sigma_+)$, to get

$$\tilde{r}_{12}(\alpha, \beta) = -\frac{1}{2i \sin(\alpha - \beta)} [2 (\sigma_+ \otimes \sigma_- + \sigma_- \otimes \sigma_+) + \cos(\alpha - \beta) (\sigma_3 \otimes \sigma_3 - \mathbf{1} \otimes \mathbf{1})]. \quad (\text{A.6})$$

The term proportional to $\mathbf{1} \otimes \mathbf{1}$ is irrelevant as it plays no role in the Sklyanin bracket or in the CYBE.

There is a deeper reason for this connection which has to do with the fact that the twist of a loop algebra by an inner automorphism is isomorphic to the loop algebra. In simple terms here, the version of sine-Gordon considered in [FT] is built on the twisted loop algebra $\mathcal{L}^\theta(\mathfrak{g})$ where θ is an automorphism of order 2 of $\mathfrak{g}$, defined by $\theta X = \sigma_3 X \sigma_3$ for all $X \in \mathfrak{g}$, and extended to the algebra

$$\mathcal{L}(\mathfrak{g}) = \bigoplus_{n \in \mathbb{Z}} \mathfrak{g} \otimes \lambda^n \quad (\text{A.7})$$

by setting $\theta(X\lambda^n) = (-1)^n \theta(X) \lambda^n$ for all $X \in \mathfrak{g}$. With $\mathfrak{g} = \mathfrak{n}_- \oplus \mathfrak{h} \oplus \mathfrak{n}_+$, $\mathcal{L}^\theta(\mathfrak{g})$ can be decomposed as

$$\mathcal{L}^\theta(\mathfrak{g}) = \left(\bigoplus_{n \in \mathbb{Z}} \mathfrak{n}_- \otimes \lambda^{2n+1} \right) \oplus \left(\bigoplus_{n \in \mathbb{Z}} \mathfrak{h} \otimes \lambda^{2n} \right) \oplus \left(\bigoplus_{n \in \mathbb{Z}} \mathfrak{n}_+ \otimes \lambda^{2n+1} \right). \quad (\text{A.8})$$

Now we apply the map

$$\sigma_+ \lambda^{2n+1} \mapsto \sigma_+ \lambda^{2n}, \quad \sigma_- \lambda^{2n+1} \mapsto \sigma_- \lambda^{2n+2}, \quad \sigma_3 \lambda^{2n} \mapsto \sigma_3 \lambda^{2n} \quad (\text{A.9})$$

which amounts to the transformation $X(e^{2i\alpha}) \mapsto M(-\alpha)X(e^{2i\alpha})M^{-1}(-\alpha)$ to obtain that $\mathcal{L}^\theta(\mathfrak{g})$ is isomorphic to

$$\left(\bigoplus_{n \in \mathbb{Z}} \mathfrak{n}_- \otimes \lambda^{2n+2} \right) \oplus \left(\bigoplus_{n \in \mathbb{Z}} \mathfrak{h} \otimes \lambda^{2n} \right) \oplus \left(\bigoplus_{n \in \mathbb{Z}} \mathfrak{n}_+ \otimes \lambda^{2n} \right). \quad (\text{A.10})$$

We now apply a second map

$$\sigma_i \lambda^{2n} \mapsto \sigma_i \lambda^n, \quad i = 3, \pm, \quad (\text{A.11})$$

to obtain that $\mathcal{L}^\theta(\mathfrak{g})$ is isomorphic to

$$\left(\bigoplus_{n \in \mathbb{Z}} \mathfrak{n}_- \otimes \lambda^n \right) \oplus \left(\bigoplus_{n \in \mathbb{Z}} \mathfrak{h} \otimes \lambda^n \right) \oplus \left(\bigoplus_{n \in \mathbb{Z}} \mathfrak{n}_+ \otimes \lambda^n \right) = \mathcal{L}(\mathfrak{g}). \quad (\text{A.12})$$

References

- [ABW] Affleck, I., Bykov, D., Wamer, K.: Flag manifold sigma models: spin chains and integrable theories. [arXiv:2101.11638](https://arxiv.org/abs/2101.11638)
- [AC] Avan, J., Caudrelier, V.: On the origin of dual Lax pairs and their r -matrix structure. *J. Geom. Phys.* **120**, 106–128 (2017)
- [ACDK] Avan, J., Caudrelier, V., Doikou, A., Kundu, A.: Lagrangian and Hamiltonian structures in an integrable hierarchy and space-time duality. *Nucl. Phys. B* **902**, 415 (2016)
- [BBT] Babelon, O., Bernard, D., Talon, M.: *Introduction to Classical Integrable Systems*. Cambridge University Press (2003)
- [Be] Belavin, A.A.: Discrete groups and the integrability of quantum systems. *Funktsional. Anal. i Prilozhen.* **14**(4), 18–26 (1980)
- [BD1] Belavin, A.A., Drinfeld, V.G.: Triangle equation for simple Lie algebras. In: Novikov, R., et al. (eds.) *Mathematical Physics Reviews*, p. 93. Harwood, New York (1984)
- [BD2] Belavin, A.A., Drinfeld, V.G.: Classical Yang–Baxter equation for simple Lie algebra. *Funct. Anal. Appl.* **17**, 220 (1983)

- [BSV] Benini, M., Schenkel, A., Vicedo, B.: Homotopical analysis of 4D Chern–Simons theory and integrable field theories. [arXiv:2008.01829](#)
- [By] Bykov, D.: The CP^{n-1} -model with fermions: a new look. [arXiv:2009.04608](#)
- [CDS] Caudrelier, V., Dell’Atti, M., Singh, A.A.: Lagrangian multiforms on coadjoint orbits for finite-dimensional integrable systems. [arXiv:2307.07339](#)
- [CS1] Caudrelier, V., Stoppato, M.: A connection between the classical r-matrix formalism and covariant Hamiltonian field theory. *J. Geom. Phys.* **148**, 103546 (2020)
- [CS2] Caudrelier, V., Stoppato, M.: Hamiltonian multiform description of an integrable hierarchy. *J. Math. Phys.* **61**, 123506 (2020)
- [CS3] Caudrelier, V., Stoppato, M.: Multiform description of the AKNS hierarchy and classical r-matrix. *J. Phys. A* **54**, 235204 (2021)
- [CSV] Caudrelier, V., Stoppato, M., Vicedo, B.: On the Zakharov–Mikhailov action: 4d Chern–Simons origin and covariant Poisson algebra of the Lax connection. *Lett. Math. Phys.* **111**, 82 (2021)
- [Cos1] Costello, K.: Supersymmetric gauge theory and the Yangian. [arXiv:1303.2632](#)
- [Cos2] Costello, K.: Integrable lattice models from four-dimensional field theories. *Proc. Symp. Pure Math.* **88**, 3 (2014)
- [CWY1] Costello, K., Witten, E., Yamazaki, M.: Gauge theory and integrability, I. *ICCM Not.* **6**, 46–119 (2018)
- [CWY2] Costello, K., Witten, E., Yamazaki, M.: Gauge theory and integrability, II. *ICCM Not.* **6**, 120–149 (2018)
- [CY] Costello, K., Yamazaki, M.: Gauge Theory and Integrability, III. [arXiv:1908.02289](#)
- [DLMV1] Delduc, F., Lacroix, S., Magro, M., Vicedo, B.: Integrable coupled σ models. *Phys. Rev. Lett.* **122**(4), 041601 (2019)
- [DLMV2] Delduc, F., Lacroix, S., Magro, M., Vicedo, B.: Assembling integrable σ -models as affine Gaudin models. *JHEP* **06**, 017 (2019)
- [DLMV3] Delduc, F., Lacroix, S., Magro, M., Vicedo, B.: A unifying 2D action for integrable σ -models from 4D Chern–Simons theory. *Lett. Math. Phys.* **110**, 1645–1687 (2020)
- [Di] Dickey, L.A.: *Soliton Equations and Hamiltonian Systems*. World Scientific (2003)
- [Dr1] Drinfel’d, V.G.: Hamiltonian structures on Lie groups, Lie bialgebras, and the geometric meaning of the classical Yang–Baxter equations. *Dokl. Akad. Nauk SSSR* **268**, 285 (1983)
- [Dr2] Drinfel’d, V.G.: Quantum groups. *J. Math. Sci.* **41**, 898 (1988)
- [FR] Faddeev, L.D., Reshetikhin, N.Y.: Integrability of the principal chiral field model in (1+1)-dimension. *Ann. Phys.* **167**, 227 (1986)
- [FT] Faddeev, L.D., Takhtajan, L.A.: *Hamiltonian Methods in the Theory of Solitons*. Springer (2007)
- [FNR] Flaschka, H., Newell, A.C., Ratiu, T.: Kac–Moody Lie algebras and soliton equations: II. Lax equations associated with $A^{(1)}$. *Physica D* **9**, 300–323 (1983)
- [FSY] Fukushima, O., Sakamoto, J.I., Yoshida, K.: Faddeev–Reshetikhin model from a 4D Chern–Simons theory. *JHEP* **02**, 115 (2021)
- [G] Gardner, C.S.: Korteweg–de Vries equation and generalizations. IV. The Korteweg–de Vries equation as a Hamiltonian system. *J. Math. Phys.* **12**, 1548 (1971)
- [GGKM] Gardner, C.S., Greene, J.M., Kruskal, M.D., Miura, R.M.: Method for solving the Korteweg–de Vries equation. *Phys. Rev. Lett.* **19**, 1095 (1967)
- [Gi] Gieres, F.: Covariant canonical formulations of classical field theories. [arXiv:2109.07330](#)
- [KN] King, S.D., Nijhoff, F.W.: Quantum variational principle and quantum multiform structure: the case of quadratic Lagrangians. *Nucl. Phys. B* **947**, 114686 (2019)
- [LV] Lacroix, S., Vicedo, B.: Integrable $\mathcal{E}$ -models, 4D Chern–Simons theory and affine Gaudin models. I. Lagrangian aspects. *SIGMA* **17**, 058 (2021)
- [LL] Lepowsky, J., Li, H.: *Introduction to Vertex Operator Algebras and Their Representations*, vol. 227. Springer, New York (2004)
- [LN] Lobb, S., Nijhoff, F.W.: Lagrangian multiforms and multidimensional consistency. *J. Phys. A* **42**, 454013 (2009)
- [Mik] Mikhailov, A.V.: The reduction problem and the inverse scattering method. *Physica D* **3**(1–2), 73 (1981)
- [N1] Nijhoff, F.W.: Integrable hierarchies, Lagrangian structures and non-commuting flows. In: Ablowitz, M.J., Fuchssteiner, B., Kruskal, M. (eds.) *Topics in Soliton Theory and Exactly Solvable Nonlinear Equations*, pp. 150–181. World Scientific, Singapore (1987)
- [N2] Nijhoff, F.W.: Linear integral transformations and hierarchies of integrable nonlinear evolution equations. *Physica D* **31**, 339 (1988)
- [PV] Petrer, M., Vermeeren, M.: Variational symmetries and pluri-Lagrangian structures for integrable hierarchies of PDEs. *Eur. J. Math.* **7**, 741 (2021)

- [Sa] Sattinger, D.H.: Hamiltonian hierarchies on semisimple Lie algebras. *Stud. Appl. Math.* **72**(1), 65–86 (1985)
- [STS] Semenov-Tian-Shansky, M.A.: What is a classical r -matrix? *Funct. Anal. Its Appl.* **17**, 259 (1983)
- [STS2] Semenov-Tian-Shansky, M.A.: Integrable systems: an r -matrix approach, Kyoto preprint RIMS-1650 (2008). www.kurims.kyoto-u.ac.jp/preprint/file/RIMS1650.pdf
- [SNC] Sleigh, D.G., Nijhoff, F.W., Caudrelier, V.: A variational approach to Lax representations. *J. Geom. Phys.* **142**, 66 (2019)
- [SNC2] Sleigh, D.G., Nijhoff, F.W., Caudrelier, V.: Variational symmetries and Lagrangian multiforms. *Lett. Math. Phys.* **110**, 805 (2020)
- [SNC3] Sleigh, D.G., Nijhoff, F.W., Caudrelier, V.: Lagrangian multiforms for Kadomtsev–Petviashvili (KP) and the Gelfand–Dickey hierarchy. *Int. Math. Res. Not.* rna288 (2021)
- [Sk] Sklyanin, E.K.: On complete integrability of the Landau–Lifshitz equation, LOMI-79-3 (1979)
- [Su] Suris, Y.B.: Variational symmetries and pluri-Lagrangian systems. In: *Dynamical Systems, Number Theory and Applications: A Festschrift in Honor of Armin Leutbecher's 80th Birthday*, pp. 255–266 (2016)
- [SV] Suris, Y., Vermeeren, M.: On the Lagrangian structure of integrable hierarchies. In: Bobenko, A. (ed.) *Advances in Discrete Differential Geometry*, pp. 347–378. Springer, Berlin (2016)
- [Ta] Takhtajan, L.A.: Quantum field theories on algebraic curves. I. Additive bosons, 2013 Russian Academy of Sciences, (DoM) and London Mathematical Society, Turpion Ltd—Izvestiya: Mathematics, 77(2)
- [TU] Terng, C.-L., Uhlenbeck, K.: Bäcklund transformations and loop group actions. *Commun. Pure Appl. Math.* **53**(1), 1–75 (2000)
- [V] Vermeeren, M.: Continuum limits of pluri-Lagrangian systems. *J. Integr. Syst.* **4**, xyy020 (2019)
- [V1] Vicedo, B.: On integrable field theories as dihedral affine Gaudin models. *Int. Math. Res. Not.* **2020**(15), 4513–4601 (2020)
- [V2] Vicedo, B.: Holomorphic Chern–Simons theory and affine Gaudin models. [arXiv:1908.07511](https://arxiv.org/abs/1908.07511)
- [ZF] Zakharov, V.E., Faddeev, L.D.: Korteweg–de Vries equation, a complete integrable Hamiltonian system. *Funct. Anal. Appl.* **5**, 280 (1971)
- [ZMan] Zakharov, V.E., Manakov, S.V.: On the complete integrability of a nonlinear Schrödinger equation. *Theor. Math. Phys.* **19**, 551 (1975)
- [ZM1] Zakharov, V.E., Mikhailov, A.V.: Variational principle for equations integrable by the inverse problem method. *Funct. Anal. Its Appl.* **14**, 43–44 (1980)
- [ZM2] Zakharov, V.E., Mikhailov, A.V.: Relativistically invariant two-dimensional models of field theory which are integrable by means of the inverse scattering problem method. *Sov. Phys. JETP* **47**, 1017 (1978)
- [ZS] Zakharov, V.E., Shabat, A.B.: Exact theory of two-dimensional self-focusing and one-dimensional self-modulation of waves in non-linear media. *Sov. Phys. JETP* **34**, 62 (1972)
- [Zo] Zotov, A.V.: 1+1 Gaudin model. *SIGMA* **7**, 067 (2011)