GLM Equations, Tau Function and Scattering Data

Max R. Niedermaier

Max-Planck-Institut für Physik, Werner Heisenberg Institut, Föhringer Ring 6, D-80805 Munich, Germany

Abstract: The direct and the inverse scattering problem for affine Toda/mKdV systems is addressed and is found to develop non-standard features within the framework of the inverse scattering method. A solution scheme based on the tau function formalism is described. The inverse problem is shown to be equivalent to a set of decoupled, scalar Gelfand-Levitan-Marchenko-type equations. The Fredholm-Grothendieck determinants of the latter are shown to define tau-functions in the sense of the Kyoto School. In particular, a simple monodromy formula allows the derivation of trace identities.

1. Introduction

For many integrable field theories the inverse scattering method (ISM) provides the most complete and physically compelling insight into the structure of the classical phase space. The basic discovery, dating back to Gardner, Greene, Kruskal and Miura [9], is that a generic solution can be parametrized through the scattering data of some auxiliary linear system (generalised Schrödinger equation). Although initially designed for the specific example of the KdV equation, the principle turned out to be systematically applicable to a wide range of systems (see e.g. the book [21] for an exposition). Moreover the scattering data were found to be related to action-angle variables, turning these models into infinite dimensional completely integrable Hamiltonian systems [8].

An independent development was initiated by the observation of Hirota that many of these nonlinear equations could be bilinearized by a suitable change of variables, so that a direct construction of solutions became possible [12]. In addition, these variables ("τ-functions") were discovered by the Kyoto School to describe the orbits of affine Lie groups in a particular realization [4, 14]. This lead to a systematic construction and classification scheme for integrable systems in terms of the data associated with some affine Lie algebra. An overlapping, presumably yet broader algebraic scheme was developed by Drinfeld and Sokolov [6] (see also [5, 13]). We

will collectively refer to these systems as being *algebraically integrable*. This means that a certain algebraic scenario (linear system, infinite set of conserved charges, bihamiltonian structure, etc.), which usually goes along with complete integrability (suitably defined), can be constructed for these models.

There is, however, a mismatch between the large number of algebraically integrable systems and the lack of further knowledge on the structure of their phase space. In particular, given the preferred role of the scattering data in a Hamiltonian formulation, it seems to be desirable to extend the parametrization of the phase space through scattering data beyond the rank 1 systems [associated with the Lie algebra $\widehat{sl}(2)$] typically considered in the literature on the ISM.[1] In this paper we provide such an extension for a typical series of algebraically integrable models in the above sense. We find that the principle aims of the ISM can still be achieved, although on the methodological level considerable deviations from the standard scheme are necessary. Before describing these deviations, it might be useful to recall the aims of the ISM.

Aims of the ISM: The ISM aims 1) to achieve a parametrization of the phase space through the scattering data of an auxiliary linear problem. The classical phase space is identified with the space of solutions to the field equations with specified boundary conditions. On the scattering data the dynamics becomes trivial and the mapping $\mathscr{F}$: *solution* → *scattering data* may be viewed as a nonlinear analogue of the Fourier transform. The solution here is supposed to be given and one just seeks to identify the parameter space relevant to the class of solutions aimed at (e.g. $\mathscr{C}^\infty$ with rapidly decreasing b.c.). In the second step one tries to invert $\mathscr{F}$, i.e. 2) to construct the classical solution from given scattering data. This is called the *inverse problem* and can usually be shown to be equivalent (for generic data) to the solution of certain matrix integral equations (of Wiener-Hopf type for the Riemann-Hilbert problem or of Fredholm type for the matrix GLM equations). In a final step 3) one aims to derive *trace identities*, i.e. to find explicit expressions for the infinite set of conserved charges in terms of the scattering data. Symbolically, find $I^{(n)}(\cdot)$ in

$$I^{(n)}[solution] = I^{(n)}(scattering\ data).$$

For the Hamiltonian ($n = 1$, say), in particular, this gives the classical masses/energy levels of the various excitations. Generally, the significance of trace identities lies in the fact that within a Hamiltonian formulation the quantities $I^{(n)}(scattering\ data)$ identify a (complete) set of action variables on the phase space.

In the following we will consider the program 1)–3) for the case of affine Toda theories. These are 2-dim. relativistic field theories generalizing the Sinh/Sine-Gordon model, which are defined in terms of the data of some affine Lie aglebra. Because of the relevance to perturbed conformal field theories and the extension of the bootstrap approach, the aim of constructing the corresponding QFTs has received considerable attention in the last years (see e.g. [19] for references). In lightcone dynamics, the affine Toda theories appear as extensions of the (generalized) mKdV systems and to some extent both types of systems can (and should) be studied in parallel. Here we consider the affine Toda theories of the A series in the principal graduation $AT(r, 1)$. For the notation and a collection of basic results we refer to Appendix A. The generalization to other Lie algebras should provide no prinicple problems.[2]

---

[1] Results on higher rank systems of this type are surveyed in Chap III.3 of [21]. The analogue of the matrix Ω below, however, is assumed to be real there, so that the complications we find are absent

[2] The τ-functions of $AT(\hat{g}, s)$ will only partially be characterized by bilinear identities, namely those of the mKdV systems, which then have to be supplemented by other relations/requirements

Non-standard features. The attempt to realize the program 1)–3) by application of the standard scheme for the ISM meets certain obstructions. For the mKdV systems of rank $r > 1$ the inverse problem is not equivalent to a standard Riemann Hilbert problem and also the derivation of Gelfand-Levitan-Marchenko (GLM) equations is problematic.[3] The reason is that the matrix Jost solutions (and hence the monodromy matrix) lack simple analyticity properties in the spectral parameter. The analyticity properties of the Jost solutions in the parameter μ turn out to be governed by the falloff properties of exponentials of the form $e^{-i\mu x \Omega}$, as $x \to \pm \infty$, where $\Omega = (1, \omega, \ldots, \omega^r)$ and $\omega = e^{\frac{2\pi i}{r+1}}$. The crucial difference between the rank 1 systems and their higher rank generalizations lies in the fact that only for $r = 1$ is the matrix Ω real, so that suitable rows/columns of the Jost solutions admit an analytic extension to either the upper or the lower half μ-plane. For $r > 1$ these simple analyticity properties fail and most of the techniques based on them fail likewise. Imposing, however, a certain technical restriction on the scattering data (condition B in Sect. 2.6), which presumably renders them generic only for the affine Toda but not for the mKdV theories, a system of GLM equations can be derived. For the $AT(r, 1)$ models we show in Sect. 3:

• The inverse problem is equivalent to a set of scalar, decoupled Gelfand-Levitan-Marchenko-type equations.

Each of the equations constitutes a family of Fredholm equations indexed by $x \in \mathbb{R}$ on a non-compact interval. The kernels of the associated integral operators $\hat{F}_{j,x}^{(\pm)}$, $j = 0, \ldots, r$, $x \in \mathbb{R}$ are specified in terms of the scattering data (whence of "GLM-type"). The operators $\hat{F}_{j,x}^{(\pm)}$ are *not* trace-class[4] but a version of Grothendieck's generalized Fredholm theory on Banach spaces [11] can be applied and leads to the result

• A ("generic") solution of the affine Toda equations can be constructed via

$$\phi^a = -\frac{1}{\varepsilon\beta} \sum_{j=0}^{r} \alpha_j^a \ln D_j^{(\pm)},$$

where $D_j^{(\pm)}(x)$ are the Fredholm-Grothendieck determinants of $\hat{F}_{j,x}^{(\pm)}$. Moreover the determinants $D_j^{(\pm)}$ define generalized τ-functions in the sense of the Kyoto School, in a parametrization through scattering data. The matching of the solutions "from the right" $(\phi^a)^{(+)}$ and "from the left" $(\phi^a)^{(-)}$ is guaranteed by construction.

The last point refers to the consistency condition for the GLM approach to the inverse problem. It is also the basis for interpreting the GLM equations as Zakharov-Shabat dressing problem [30]. The relation to the τ-functions further gives a new way to derive trace identities [19, 20]

---

[3] An approach could be via reduction of a non-local RH-problem [17] For special data a relation to τ-functions has already been observed in [22, 25]. The problem consists in formulating the inverse problem for generic data

[4] For this reason the solution of the decoupled GLM equations is not quite equivalent to a standard Bruhat decomposition in the affine group. Recovering such a formulation would require to generalize the notion of a dressing group to include determinants of nuclear operators on Banach spaces (cf Sect. 4)

- The derivation of trace identities is possible from the formulae

$$I_{\pm}^{(n)} = \mp \frac{r+1}{n} (\varepsilon\beta)^{-(n+1)} \partial_{x_n^{\pm}} \ln \tau_j \Big|_{x^- = -\infty}^{\infty} ,$$

where $I_{\pm}^{(n)}$ are the conserved charges in lightcone dynamics and the right-hand sides are independent of the flow variables $x_n^{\pm}$ as well as j-independent.

To keep the paper of reasonable size we defer the derivation of trace identities to a separate publication. In Sect. 3 we show that the inverse problem is equivalent to the set of decoupled GLM-type equations described above. The solution of these equations is given in Sect. 4 in terms of the Fredholm-Grothendieck determinants $D_j^{(\pm)}$, which are shown to define τ-functions in a parametrization through scattering data. To prepare the ground we have to extend the notion of scattering data to the higher rank models in question.

2. The Scattering Data

We assume some familiarity with affine Toda theories and the content of references [6, 23, 24]. Our conventions and a number of basic results have been summarized in Appendix A. In particular, we recall from there that the equations of motion can be obtained from the integrability condition of a linear system in the affine Lie algebra. This can be reduced to a linear problem in the loop algebra/group. For $g = A_r$ and the principle graduation, the latter takes the form

$$\mathscr{L}_{\pm} W = 0, \qquad \mathscr{L}_{\pm} = \partial_{\pm} - A_{\pm} ,$$

$$A_+ = -\frac{m}{2} e^{-\varepsilon\beta\phi \cdot H} (-i\mu\Lambda)^{-1} e^{\varepsilon\beta\phi \cdot H} , \qquad (2.1)$$

$$A_- = -\varepsilon\beta H \cdot \partial_-\phi - \frac{m}{2} i\mu\Lambda .$$

Here ϕ^a are a set of real scalar fields considered as the components of a cartan subalgebra-valued field w.r.t. a basis H^a, $a = 1, \ldots, r$ and Λ is the matrix $(\delta_{a,b-1})_{0 \le a, b \le r}$ with the indices taken modulo $r + 1$. Further $\varepsilon = 1$, i for real and imaginary coupling models, respectively. For convenience we have also extracted a factor $-i$ from the parameter of the loop group. The conventions are then s.t. the operator defining the genuine eigenvalue problem is formally selfadjoint on a suitable space of wave functions for $\varepsilon = 1$ (see Appendix A). This means that for real coupling affine Toda theories μ is real.

2.1 Wave Functions

To address problem 1 of the introduction, suppose now a classical solution $\phi \cdot H$ to be given. The integrability condition $[\mathscr{L}_+, \mathscr{L}_-] = 0$ is then satisfied and it suffices to consider the linear equation $\mathscr{L}_- W = 0$. To simplify the notation set $D = \beta H \cdot \partial_-\phi = \text{diag}(d_0, \ldots, d_r)$, where $d_j = \beta \hat{h}_j \cdot \partial_-\phi$ and $\hat{h}_j$ is the jth weight of the defining representation of A_r. Rescaling finally $x^{\pm} \to x^{\pm} \frac{2}{m}$ the equation now reads

$$\mathscr{L} W = 0, \qquad \mathscr{L} := \mathscr{L}_- = \partial_- + \varepsilon D + i\mu\Lambda =: \partial_- - A_- \qquad (2.2)$$

It is convenient to diagonalize Λ via $U^{-1}\Lambda U = \mathrm{diag}(1, \omega, \ldots, \omega^r) =: \Omega$, where
$U = \dfrac{1}{\sqrt{r+1}}\, (\omega^{ab})_{0 \le a, b \le r}$, $\omega = e^{\frac{2\pi i}{r+1}}$. Set

$$\mathscr{L}^U := U^{-1}\mathscr{L}U = \partial_- + \varepsilon D^U + i\mu\Omega\,, \tag{2.3}$$

where $D^U = \sum\limits_{a=1}^{r} D_a \Lambda^{-a}$ is a ω^0-circulant with diagonal entries $D_a = \dfrac{1}{r+1} \times$
$\sum\limits_{j=0}^{r} \omega^{-aj} d_j$. The associated wave function is $W^U = U^{-1}WU$. Now observe that
$\mathscr{L}^U = \mathscr{L}^U(\mu)$ enjoys the involution $\Lambda\mathscr{L}^U(\mu)\Lambda^{-1} = \mathscr{L}^U(\omega\mu)$. For a given potential
D^U and fixed initial (or asymptotic) value the solution of $\mathscr{L}^U W^U = 0$ is unique so
that W^U has to obey the same involution

$$\Lambda W^U(\mu)\Lambda^{-1} = W^U(\omega\mu)\,. \tag{2.4}$$

The general solution of (2.4) is given by an ω-circulant in the arguments, i.e.

$$W^U(\mu) = \sum_{a=0}^{r} w_a(\Omega\mu)\Lambda^{-a}\,, \tag{2.5}$$

where
$$w_a(\Omega\mu) := \mathrm{diag}(w_a(\mu),\, w_a(\omega\mu),\, \ldots,\, w_a(\omega^r\mu))\,.$$

In particular, any wave function is parametrized by $r + 1$ independent functions
$w_a(x^+, x^-; \mu)$, $0 \le a \le r$. The original wave function is then given by $W = (W_{kl})_{0 \le k, l \le r}$,

$$W_{kl}(\mu) = \frac{1}{r+1} \sum_{a=0}^{r} \sum_{j=0}^{r} \omega^{ak+j(k-l)} w_a(\omega^{a+j}\mu)\,, \tag{2.6}$$

and solves $\Omega^{-1}W(\mu)\Omega = W(\omega\mu)$. For a domain in the complex μ-plane, where the
$w_a(\mu)$ are analytic an equivalent expression is

$$W(\mu) = \sum_{a=0}^{r} w_a(\Lambda\mu)\Omega^a\,, \tag{2.7}$$

where $w_a(\Lambda\mu)$ is defined through the power series expansion of $w_a(\mu)$ with μ replaced
by $\Lambda\mu$.

A similar, but less stringent involution exists under complex conjugation. Writing
momentarily $\mathscr{L}^U(\varepsilon, \mu)$ for the Lax operator (2.3) one verifies $(\mathscr{L}^U(\varepsilon, \mu))^* = J\mathscr{L}^U(\varepsilon^*, -\mu^*)J$, where $J = U^2$. This implies

$$W^U(\varepsilon, \mu)^* = JW^U(\varepsilon^*, -\mu^*)J\,, \tag{2.8}$$

in the same notation. In terms of components this becomes

$$w_a(\varepsilon, \mu)^* = w_{\bar{a}}(\varepsilon^*, -\mu^*)\,. \tag{2.9}$$

Because the ε-dependence of w_a will not be known explicitly in general, the relation
(2.9) will be useful only for $\varepsilon = 1$.

2.2. *Existence of Jost Solutions*

For any interval $I \subset \mathbb{R}$ let $L_1(I, A_r)$ denote the space of A_r-valued functions (in the defining representation), which are absolutely integrable on I w.r.t. some matrix-norm $\|\cdot\|$, i.e.

$$\int_I dx^- \|F(x^-)\| < \infty. \tag{2.10}$$

In particular for the potentials D in (2.2) we will assume *rapidly decreasing boundary conditions*, i.e. require $D \in L_1(\mathbb{R}, A_r)$. To a given classical solution of this type one can uniquely associate a pair of *Jost solutions* of $\mathcal{L}W = 0$ with the asymptotics (w.r.t. the L_1-norm)

$$W^{(\pm)}(x^+, x^-; \mu) \to e^{-i\mu\Lambda x^-} + o(1), \tag{2.11}$$

for $x^- \to \pm\infty$, respectively.

To do this, introduce the *transition matrix* w.r.t. (x^-, y^-) by the conditions

$$\mathcal{L}T(x^-, y^-; \mu) = 0, \qquad T(x^-, x^-; \mu) = I, \tag{2.12}$$

where to simplify the notation we have suppressed the dependence on x^+, y^+. A straightforward extension of the reasoning in [8] (p. 30ff., p. 39ff.) yields the

Lemma 1. a) *The limits*

$$W^{(\pm)}(x^-; \mu) := \lim_{y^- \to \pm\infty} T(x^-, y^-; \mu) e^{-i\mu y^- \Lambda} \tag{2.13}$$

exist for $\mu \in S := \bigcup_{0 \le a \le r} \omega^a \mathbb{R}$.

b) *There exist kernels* $\Gamma^{(\pm)}(x^-, \cdot) \in L_1(I^\pm_{x^-}, A_r)$ *s.t.*

$$W^{(\pm)}(x^-; \mu) = e^{-ix^- \mu\Lambda} + \int_{I^\pm_{x^-}} dz\, \Gamma^{(\pm)}(x^-, z^-) e^{-iz^- \mu\Lambda}, \tag{2.14}$$

where $I^+_{x^-} = [x^-, -\infty[, I^+_{x^-} = [\infty, x^-[$ *In particular* (2.11) *holds.*

Remark Similar formulae hold for the transformed quantities $A^U = U^{-1}AU$ with $e^{-ix^- \mu\Lambda}$ replaced by $e^{-ix^- \mu\Omega}$. In particular set $W^{(U,\pm)} := U^{-1}W^{(\pm)}U$ and $\Gamma^{(U,\pm)} := U^{-1}\Gamma^{(\pm)}U$. In the A_1 case one has $\Omega = \text{diag}(1, -1)$ so that (2.14) can be used to deduce that the columns of $W^{(U,\pm)}(\mu)$ admit an analytic extension to either the upper or the lower half μ-plane. For rank $r > 1$ the phase factors are complex and the rows/columns of $W^{(U,\pm)}(\mu)$ do not admit analytic extension off the star S in general.

We note some basic properties of the Jost solutions. A standard argument shows $\det W^{(\pm)} = 1$. From (2.5) it follows further that the Jost solutions $W^{(\pm)}$ are parametrized by $r + 1$ functions $w_a^{(\pm)}(x^+, x^-; \mu)$ subject to the asymptotics

$$\begin{aligned} w_0^{(\pm)} &\to e^{-i\mu x^-} + o(1) \\ w_a^{(\pm)} &\to 0, \qquad a = 1, \ldots, r \end{aligned} \qquad \text{for} \quad x^- \to \pm\infty, \ \mu \in S, \tag{2.15}$$

where the limits are taken w.r.t. the (scalar) L_1-norm. In particular,

$$W_{kl}^{(\pm)}(x^+, x^-; \mu) \to \frac{1}{r+1} \sum_{j=0}^{r} \omega^{j(k-l)} e^{-i\mu\omega^j x^-}, \qquad x^- \to \pm\infty, \qquad (2.16)$$

consistent with (2.11).

As outlined in Appendix A the pair (x^+, x^-) is only the lowest member of an infinite sequence of higher order flow variables $x_n^{\pm}$ associated with the equations of the affine Toda/mKdV hierarchy. In particular $\mathscr{L} = \mathscr{L}_- = \mathscr{L}_-^1$ is the lowest member of the sequence of matrix differential operators defining the mKdV hierarchy. There asymptotics is $\lim_{x^- \to \pm\infty} \mathscr{L}_-^n = \frac{\partial}{\partial x_n^-} - (-i\mu\Lambda)^n$. Correspondingly the Jost solutions for the hierarchy are defined through the condition

$$W^{(\pm)}(\underline{x}^+, \underline{x}^-; \mu) = e^{\xi(x^-, -i\mu\Lambda)} + o(1), \qquad x^- \to \pm\infty, \qquad (2.17)$$

where $\xi(x^-, \mu) = \sum_{n \in E} x_n^- \mu^n$ and $\underline{x}^{\pm} = (x_1^{\pm} = x^{\pm}, x_2^{\pm}, \ldots, x_n^{\pm}, \ldots)$. To have the functions $\xi(x; \mu)$ or $\xi(x; \Lambda\mu)$ etc. analytic in all its variables, we will always assume that the hierarchy variables satisfy the condition

$$\lim_{n \to \infty} \sup |x_n^{\pm}|^{1/n} = 0. \qquad (2.18)$$

In particular (2.18) holds if all but a finite number of $x_n^{\pm}$'s vanish.

2.3. The Monodromy Matrix

For a given potential D (and hence $\mathscr{L}_-^n$, $n \in E$) the Jost solutions provide a pair of fundamental solutions to a set of linear matrix differential equations. Hence they can not be independent but must be linearly related via

$$W^{(-)}(\underline{x}^+, \underline{x}^-; \mu) = W^{(+)}(\underline{x}^+, \underline{x}^-; \mu) T(\underline{x}^+; \mu), \qquad \mu \in S. \qquad (2.19)$$

The matrix $T(\underline{x}^+; \mu)$ depends on the "time" variables only and is called *reduced monodromy matrix*. Initially it is defined on S and then possibly through analytic continuation. Clearly $\det T = 1$. Once more, a uniqueness argument implies that

$$\Lambda T^U(\underline{x}^+; \mu) \Lambda^{-1} = T^U(\underline{x}^+, \omega\mu), \qquad (2.20)$$

for $T^U = U^{-1}TU$. This means that also T^U is parametrized by $r + 1$ independent functions $t_a(\underline{x}^+; \mu)$, $a = 0, \ldots, r$ via

$$T^U(\underline{x}^+; \mu) = \sum_{a=0}^{r} t_a(\underline{x}^+; \Omega\mu) \Lambda^{-a}, \qquad (2.21)$$

$$t_a(\underline{x}^+, \Omega\mu) := \text{diag}(t_a(\underline{x}^+; \mu), t_a(\underline{x}^+; \omega\mu), \ldots, t_a(\underline{x}^+; \omega^r \mu)).$$

For T one has an expression analogous to (2.6). From (2.20) it follows that the functions $t_a(\mu)$, $0 \le a \le r$ are defined on the region $S = \bigcup_{0 \le a \le r} \omega^a \mathbb{R}$.

Proposition 1. *The "time" dependence of t_a, $0 \le a \le r$ is given by*

$$t_a(\underline{x}^+; \mu) = t_a(\underline{0}; \mu) \exp\left(\sum_{n \in E} (1 - \omega^{an}) (i\mu)^{-n} x_n^+\right). \qquad (2.22)$$

In particular $t_0(\underline{x}^+; \mu) = t_0(\underline{0}; \mu) =: t_0(\mu)$ is time independent.

Proof. First note that (2.22) is equivalent to

$$\frac{\partial}{\partial x_n^+} TU = [(i\mu\Omega)^{-n}, T^U], \tag{2.23}$$

using $[\Omega^n, \Lambda^a] = (1 - \omega^{an})\Omega^n\Lambda^a$. To show (2.23) express the transition matrix $T(x^-, y^-; \mu)$ in terms of the Jost solutions

$$\begin{aligned}
T(x^-, y^-; \mu) &= W^{(+)}(x^-; \mu) W^{(+)}(y^-; \mu)^{-1} \\
&= W^{(-)}(x^-; \mu) W^{(-)}(y^-; \mu)^{-1} \\
&= W^{(-)}(x^-; \mu) T^{-1}(\mu) W^{(+)}(y^-; \mu)^{-1} \\
&= W^{(+)}(x^-; \mu) T(\mu) W^{(-)}(y^-; \mu)^{-1}.
\end{aligned} \tag{2.24}$$

The first two equalities follow from the fact that the r.h.s. solves the differential equation and has the correct normalization. The rest follows upon insertion of (2.19). In particular,

$$\begin{aligned}
T(\mu) &= \lim_{x^- \to \infty, y^- \to -\infty} W^{(+)}(x^-; \mu)^{-1} T(x^-, y^-; \mu) W^{(-)}(y^-; \mu) \\
&= \lim_{x^- \to \infty, y^- \to -\infty} e^{i\mu\Lambda x^-} T(x^-, y^-; \mu) e^{-i\mu\Lambda y^-},
\end{aligned} \tag{2.25}$$

$$\begin{aligned}
W^{(\pm)}(x^-; \mu) &= T(x^-, y^-; \mu) W^{(\pm)}(y^-; \mu) \\
&= \lim_{y^- \to \pm\infty} T(x^-, y^-; \mu) e^{-i\mu\Lambda y^-},
\end{aligned} \tag{2.26}$$

$$T(\mu) = \lim_{x^- \to \pm\infty} e^{i\mu x^- \Lambda} W^{(\pm)}(x^-; \mu).$$

To proceed, derive the evolution equation for the transition matrix parallel to [8] (Eqn. (I.3.21))

$$\frac{\partial T(x^-, y^-; \mu)}{\partial x_n^+} = A_+^n(x^-, y^-; \mu) T(x^-, y^-; \mu) - T(x^-, y^-; \mu) A_+^n(y^-; \mu), \tag{2.27}$$

where $\mathscr{L}_+^n = \dfrac{\partial}{\partial x_n^+} - A_+^n$ are the higher order Lax operators associated with x_n^+, $n \in E$. Now diagonalize Λ and rewrite all equations in terms of the transformed quantities $A^U = U^{-1}AU$. Using (2.25) on the l.h.s. of (2.27) and (2.26) on the r.h.s. one arrives at (2.23), if $\lim_{x^- \to \infty} U^{-1}A_+^nU = (i\mu\Omega)^{-n}$ is taken into account $\square$

In particular Proposition 1 implies that the functions t_a, $a = 0, \ldots, r$, uniquely associated to a solution of the field equations, can serve as a set of scattering data.

2.4 Scalar Formulation

In Sect. 2.2 a Lax formulation in terms of first order matrix differential operators was used. It is well-known that the mKdV systems admit an equivalent formulation in terms of scalar differential operators. For technical reasons it is often useful to exploit the scalar formulation. In this section we relate the scalar and the matrix formulations.

First note that from each matrix solution of $\mathscr{L}W = 0$ one can construct $r + 1$ vector solutions $\mathscr{L}\underline{W}_a = 0$ via

$$\underline{W}_a = W\underline{v}_a, \qquad \underline{v}_a = (1, \omega^a, \ldots, \omega^{ra})^T, \tag{2.28}$$

where $\underline{v}_a$, $a = 0, \ldots, r$ are the eigenvectors of $\Lambda\colon \Lambda\underline{v}_a = \omega^a\underline{v}_a$. Set $\underline{W}_a = (V_{0,a}, \ldots, V_{r,a})^T$. Applying (2.7) (within some domain of analyticity in μ) gives

$$V_{j,a}(\mu) = \omega^{aj} \sum_{k=0}^{r} W_{jk}(\omega^a \mu) \,. \tag{2.29}$$

Alternatively one can directly evaluate

$$V_{j,a}(\mu) = (W\underline{v}_a)_j = \sum_{k=0}^{r} W_{jk}(\mu)\omega^{ak} \,. \tag{2.30}$$

Inserting the explicit form of W_{kl} in (2.6) one obtains in both cases

$$V_{j,a}(\mu) = \sum_{b=0}^{r} \omega^{(a+b)j} w_b(\omega^{a+b}\mu) \,. \tag{2.31}$$

In particular

$$V_{j,a}\mu) = \omega^{aj} V_{j,0}(\omega^a \mu) \,, \tag{2.32}$$

so that only the j-index labels independent functions. Inverting (2.31) yields

$$w_a(\mu) = \frac{1}{r+1} \sum_{k=0}^{r} \omega^{-k(a+b)} V_{k,b}(\omega^{-a-b}\mu) \,, \tag{2.33}$$

where (2.32) can be used to check that the r.h.s. is b-independent.

Consider now the scalar eigenvalue problem

$$\begin{aligned} L_j V &= \mu^{r+1} V \,, \\ L_j &= (i\partial_- + i\varepsilon d_{r+j})\ldots(i\partial_- + i\varepsilon d_j) \,, \end{aligned} \tag{2.34}$$

(the indices taken modulo $r+1$). For fixed j the functions $V_{j,a}$, $a = 0, \ldots, r$ then provide a fundamental system of solutions to (2.34). To see this, write out $\mathscr{L}\underline{W}_a = 0$ to find

$$(i\partial_- + i\varepsilon d_j)V_{j,a} = \mu V_{j+1,a} \,, \qquad 0 \le j \le r \,.$$

Acting successively with $(i\partial_- + i\varepsilon d_j)$, $k > j$ on the jth equation gives $L_j V_{j,a} = \mu^{r+1} V_{j,a}$ in the rth step. In particular, starting from the matrix Jost solutions $W^{(\pm)}$ in (2.28) one obtains a system of Jost solutions $V_{j,a}^{(\pm)}$ for the operators L_j. From (2.15), (2.31) there asymptotics is given by

$$V_{j,a}^{(\pm)} \to \omega^{aj} e^{-i\mu\omega^a x^-} + o(1) \,, \qquad x^- \to \pm\infty \,. \tag{2.35}$$

Forming two fundamental systems of solutions of (2.34), the Jost solutions $V_{j,a}^{(\pm)}$ must be linearly related. In fact

$$V_{j,a}^{(-)} = \sum_{b=0}^{r} t_{b-a}(\omega^b \mu) V_{j,b}^{(+)} \,, \qquad j = 0, \ldots, r \,. \tag{2.36}$$

To see this, rewrite (2.33) as

$$(W^U)_{j,a}(\mu) = \frac{1}{r+1} \sum_{k=0}^{r} \omega^{-jk} V_{k,a}(\mu) \,. \tag{2.37}$$

From (2.19) one finds

$$(W^{(U,-)})_{j,a} = \sum_{b=0}^{r} t_{b-a}(\omega^b \mu) (W^{(U,+)})_{j,b}, \tag{2.38}$$

which results in (2.36).

Remark. From the viewpoint of the scalar eigenvalue problem Eq. (2.36) is nontrivial. The definition of the Jost solutions for (2.34) implies only that

$$V_{j,a}^{(-)} = \sum_{b=0}^{r} (S_j)_{ab} (V_{j,b}^{(+)}),$$

for some matrices S_j. The relation (2.36) tells that S_j is j-independent and coincides with the transpose of T^U.

2.5. Integral Representation for Scalar Jost Solutions

Recall from Lemma 1 the integral representations

$$W^{(U,\pm)}(x;\mu) = e^{-ix\mu\Omega} + \int_{I_x^{\pm}} dz\, \Gamma^{(U,\pm)}(x,z) e^{-iz\mu\Omega}, \qquad \mu \in S, \tag{2.39}$$

where for simplicity we set $x := x^-$, etc. From $\Lambda\Omega\Lambda^{-1} = \omega\Omega$ and (2.4) it follows that the kernels $\Gamma^{(\pm)}$, being μ-independent, must satisfy the involution $\Lambda\Gamma^{(\pm)}(x,z)\Lambda^{-1} = \Gamma^{(\pm)}(x,z)$. This means that the matrix-kernels $\Gamma^{(\pm)}$ are parametrized by only $r+1$ independent scalar kernels $K_c^{(\pm)}(x,z)$, $c = 0, \ldots, r$ via

$$\Gamma^{(U,\pm)}(x,z) = \sum_{c=0}^{r} K_c^{(\pm)}(x,z) \Lambda^{-c}. \tag{2.40}$$

In components one has $(W^{(U,\pm)}(x;\mu))_{ab} = w_{a-b}^{(\pm)}(x;\omega^a\mu)$ and $(\Gamma^{(U,\pm)}(x,z))_{ab} = K_{a-b}^{(\pm)}(x,z)$, so that (2.39) becomes

$$w_c^{(\pm)}(x;\omega^{b+c}\mu) = \delta_{c,0} e^{-ix\omega^b\mu} + \int_{I_x^{\pm}} dz\, K_c^{(\pm)}(x,z) e^{-iz\omega^b\mu}. \tag{2.41}$$

Operating with $\sum_{c=0}^{r} \omega^{j(b+c)}$ on these equations and making use of (2.31) one finds

$$V_{j,a}^{(\pm)}(\mu) = \omega^{ja} \left[e^{-ix\omega^a\mu} + \int_{I_x^{\pm}} dz\, K_j^{(\pm)}(x,z) e^{-iz\omega^a\mu} \right], \tag{2.42}$$

where

$$K_j^{(\pm)}(x,z) = \sum_{c=0}^{r} \omega^{jc} K_c^{(\pm)}(x,z). \tag{2.43}$$

From the integral representations (2.39), (2.42) one can also determine the analyticity properties of the Jost solutions. The exponential factors $e^{-iz\omega^a\mu}$ in (2.42) will decay

as $z \to \pm\infty$ if $\mathrm{Im}(\omega^a \mu) \lessgtr 0$, respectively. This means that $V_{j,a}^{(\pm)}(\mu)$, initially defined on S, admit an analytic continuation to the half-planes $\mathrm{Im}(\omega^a \mu) \lessgtr 0$, respectively. For fixed x^- this also implies that

$$e^{ix^- \omega^a \mu} V_{j,a}^{(\pm)}(x^-; \mu) \to \omega^{ja} + o(1/|\mu|), \qquad (2.44)$$

as $|\mu| \to \infty$ and $\mathrm{Im}(\omega^a \mu) \lessgtr 0$, respectively. Similar statements hold for the columns of the matrix Jost solutions. The ath column of U^{-1} (2.14) U reads

$$\underline{U}_a^{(+)}(x; \mu) = \underline{e}_a e^{-ix\omega^a \mu} + \int_{I_x^{\pm}} dz\, \Gamma^{(U,+)}(x, z) \underline{e}\, e^{-iz\omega^a \mu}, \qquad (2.45)$$

where $\underline{e}_a$ is the ath vector of the canonical basis of $\mathbb{R}^{r+1}$ in column notation. Thus, the ath volumn of $W^{(U,\pm)}$ admits an analytic continuation to the half-planes $\mathrm{Im}(\omega^a \mu) \lessgtr 0$, respectively. For fixed x^- one has the fall-off properties

$$e^{ix^- \omega^a \mu} \underline{U}_a^{(\pm)}(x^-; \mu) \to \underline{e}_a + o(1/|\mu|), \qquad (2.46)$$

as $|\mu| \to \infty$ and $\mathrm{Im}(\omega^a \mu) \lessgtr 0$, respectively.

2.6. Characterization of the Discrete Spectrum

Here we show that the discrete spectrum of the various scattering problems encountered above can be characterized in terms of the zeros of a function $\Delta_+(\mu)$ constructed from the minors of the monodromy matrix. It is convenient to start with the scalar eigenvalue problems $L_j V = \mu^{r+1} V$. Recall the asymptotic form (2.35) of the Jost solutions. Clearly, for given μ, $V_{j,a}^{(-)}(\mu)$ decays as $x \to -\infty$ iff $a \in N(\mu)$, and $V_{j,a}^{(+)}(\mu)$ decays as $x \to \infty$ iff $a \in \bar{N}(\mu) := \{0, \dots, r\} \backslash N(\mu)$, where $N(\mu) = \{0 \le a \le r \mid \mathrm{Im}(\omega^a \mu) > 0\}$. Eigenfunctions of the point spectrum of L_j should be normalizable w.r.t. $(f, g) = \int_{-\infty}^{\infty} dx^- f^*(x^-) g(x^-)$, i.e. proper eigenfunctions. As the Jost solutions $V_{j,a}^{(\pm)}(\mu)$ from a fundamental system, any solution of (2.34) can be expressed as a linear combination of either $V_{j,a}^{(+)}(\mu)$ or $V_{j,a}^{(-)}(\mu)$. In particular, for a proper eigenfunction corresponding to $\mu_k \in \sigma_p(L_j)$ in the point spectrum, the linear combination has to decay as $|x^-| \to \infty$. Thus, $\mu_k \in \sigma_p(L_j)$ if and only if

$$\sum_{a \in N(\mu_k)} c_{a,k}^- V_{j,a}^{(-)}(\mu_k) = \sum_{a \in \bar{N}(\mu_k)} c_{a,k}^+ V_{j,a}^{(+)}(\mu_k). \qquad (2.47)$$

We next show that (2.47) holds iff μ_k is a zero of $\Delta(\mu) := \det(T_{ab}^U(\mu))_{a,b \in N(\mu)}$. For functions $V_0, \dots, V_r$ define the Wronskian $W[V_0, \dots, V_r] = \det(\partial_-^a V_b)_{0 \le a, b \le r}$. If the functions V_a are solutions of (2.34) the Wronskian is x^--independent and vanishes if and only if the solutions are linearly related. For the Jost solutions one has

$$W[V_{j,0}^{(\pm)}, \dots, V_{j,r}^{(\pm)}] = (-i\mu\omega^j)^{\frac{r(r+1)}{2}} \prod_{a<b} (\omega^a - \omega^b) \ne 0. \qquad (2.48)$$

Consider then $W[V_{j,a_1}^{(-)}, \ldots, V_{j,a_l}^{(-)}, V_{j,b_1}^{(+)}, \ldots, V_{j,b_{r+1-l}}^{(+)}]$, where $N(\mu) := \{a_1, \ldots,$ $a_l\}$, $a_i < a_j$, $i < j$ and $\bar{N}(\mu) = \{b_1, \ldots, b_{r+1-l}\}$, $b_i < b_j$, $i < j$ are enumerations of the index sets in (2.47). Inserting the relation $V_{j,a}^{(-)}(\mu) = \sum\limits_{b=0}^{r} T_{ba}(\mu) V_{j,b}^{(+)}$, one obtains

$$
\begin{aligned}
& W[V_{j,a_1}^{(-)}, \ldots, V_{j,a_l}^{(-)}, V_{j,b_1}^{(+)}, \ldots, V_{j,b_{r+1-l}}^{(+)}] \\
& = \sum_{c_1, \ldots, c_l \in N(\mu)} T_{c_1 a_1}(\mu) \ldots T_{c_l a_l}(\mu) \\
& \quad \times W[V_{j,c_1}^{(+)}, \ldots, V_{j,c_l}^{(+)}, V_{j,b_1}^{(+)}, \ldots, V_{j,b_{r+1-l}}^{(+)}].
\end{aligned}
\tag{2.49}
$$

Because of (2.48) the r.h.s. is seen to be proportional to $\det(T_{ab}^U)_{a,b \in N(\mu)} =: \Delta(\mu)$. We have proved the

Proposition 2. *The operators* L_j, $j = 0, \ldots, r$ *have a common point spectrum, which is given by*

$$
\mu^{r+1} \in \sigma_p(L_j) \quad \text{iff} \quad \Delta(\mu) = 0.
\tag{2.50}
$$

However, since the function $\Delta(\mu)$ is defined through different minors of $T(\mu)$ for different values of μ, the lemma can serve only to test independently given candidates $\mu_k \in \mathbb{C}$: For given $\mu_k \in \mathbb{C}$ determine $N(\mu_k)$, calculate $\Delta(\mu_k)$ and check whether it vanishes. In order to use Proposition 2 as a device to calculate the discrete spectrum we have to lift the $\mu \to \omega^a \mu$ ambiguities arising from the fact that the Jost solutions are initially defined on S. From (2.47) one sees that this can be achieved by requiring that $0 \in N(\mu)$ for all candidates $\mu^{r+1} \in \sigma_p(L_j)$. W.r.t. this framing the discrete spectrum will be characterized by a subset of the upper half plane. (Similarly the choice $a_0 \in N(\mu)$ for some fixed $0 \leq a_0 \leq r$ would characterize the discrete spectrum as a subset of $\{\mu \in \mathbb{C} \mid \mathrm{Im}(\omega^{a_0}\mu) > 0\}$.)

Remark. i. The necessity for such a framing is not specific for the higher rank models. Even for $r = 1$ one can use either $W[V_{1,0}^{(-)}, V_{1,1}^{(+)}] = t_0(\mu)$ or $W[V_{1,1}^{(-)}, V_{1,0}^{(+)}] = t_0(-\mu)$ as a starting point, which corresponds to the functions $\Delta(\mu) = t_0(\mu)$ and $\Delta(\mu) = t_0(-\mu)$, respectively. In the first case the discrete spectrum of L_j is characterized by a subset of the upper half plane, in the second case by a subset of the lower half plane.
ii. For $\varepsilon = 1$ the operators L_j are formally selfadjoint on a suitable space of functionals $a[d_0, \ldots, d_r]$ (cf. Appendix A). This implies that $\mu^{r+1} \in \mathbb{R}$, i.e. $\mu \in \bigcup\limits_{0 \leq a \leq r} e^{\frac{i\pi a}{r+1}}\mathbb{R}$. Usually one expects $\mu_k \in S$ not to give rise to normalizable eigenfunctions, in which case the discrete spectrum of L_j (w.r.t. the above framing to the upper half plane) is given by

$$
\sigma_p(L_j) \subset \bigcup_{0 \leq k \leq [(r-1)/2]} e^{\frac{i\pi(2k+1)}{r+1}}\mathbb{R}^+, \quad \varepsilon = 1.
\tag{2.51}
$$

This generalizes the well known result that for a Schrödinger operator with real, rapidly decreasing potential all eigenvalues lie on the positive imaginary axis.

The reasoning leading to Proposition 2 can be repeated for the matrix differential operator $\tilde{\mathscr{L}}$ defining the eigenvalue problem (A.13) equivalent to $\mathscr{L}W = 0$.

Equation (2.47) gets replaced by: $\mu_k \in \sigma_p(\tilde{\mathscr{L}})$ iff

$$\sum_{a \in N(\mu_k)} c_{a,k}^- \underline{U}_a^{(-)}(\mu_k) = \sum_{a \in \tilde{N}(\mu_k)} c_{a,k}^+ \underline{U}_a^{(+)}(\mu_k). \tag{2.52}$$

(One can show that the constants $c_{a,k}^\pm$ in (2.47) and (2.52) coincide; cf. below.) Equation (2.49) gets replaced by

$$W[\underline{U}_{a_1}^{(-)}, \ldots, \underline{U}_{a_l}^{(-)}, \underline{U}_{b_1}^{(+)}, \ldots, \underline{U}_{b_{r+1-l}}^{(+)}] = \Delta(\mu), \tag{2.53}$$

where $[\underline{U}_0, \ldots, \underline{U}_r]$ is the matrix with columns $\underline{U}_0, \ldots, \underline{U}_r$. Again, since the solutions $\underline{U}_a^{(\pm)}$ are defined on S we lift the degeneracy under $\mu \rightarrow \omega^a \mu$ by requiring $0 \in N(\mu)$ for all $\mu \in \sigma_p(\tilde{\mathscr{L}})$. Let thus $\mathbb{C}^+ := \{\mu \in \mathbb{C} \,|\, \mathrm{Im} \geq 0\}$ and set $\Delta_+(\mu) = t_0(\mu) \prod_{a \in N_+(\mu)} t_0(\omega^a \mu)$, where $N_+(\mu) \subset \{1, \ldots, r\}$ is defined s.t. Δ_+ coincides with Δ on $S \cap \mathbb{C}^+$.

Proposition 2′. *The function $\Delta_+(\mu)$ admits an analytic continuation to the upper half μ-plane. The point spectrum $\sigma_p(\tilde{\mathscr{L}})$ of $\tilde{\mathscr{L}}$ is given by the zeros of Δ_+ in the upper μ-half plane.*

As a corollary we note that the fall-off properties (2.44) imply that

$$\Delta_+(\mu) \rightarrow 1 + o(1), \qquad |\mu| \rightarrow \infty, \, \mathrm{Im}(\mu) > 0. \tag{2.54}$$

Together with the analyticity it then follows that the zeros of $\Delta_+(\mu)$ are located in a bounded region of the upper half plane and may only accumulate towards the real line. If one assumes in addition that $t_a(\mu)$, $a > 1$ are of Schwartz type on $\mathbb{R}$, this implies $t_0(\mu) \rightarrow 1 + 0(1)$, $|\mu| \rightarrow \infty$, $\mathrm{Im}\, \mu > 0$. Unless $r = 1$, $t_0(\mu)$ will however in general not admit an analytic continuation off the region S. For technical reasons it is convenient (cf. [8] p. 49) to assume the following

Condition A. The zeros of $\Delta_+(\mu)$ are simple and there are no zeros of $\Delta_+(\mu)$ on $S = \bigcup_{0 \leq a \leq r} \omega^a \mathbb{R}$.

For $r = 1$ it is known that this condition is mild in that the scattering data satisfying A describe a dense subspace of the phase space.

Remark. For $\varepsilon = 1$ also the matrix differential operator $\tilde{\mathscr{L}}$ is formally selfadjoint on a suitable space of matrix functionals. By construction (cf. Appendix A) every normalizable matrix wave function gives rise to a normalizable scalar wave function, but not necessarily vice versa. This implies that for $\varepsilon = 1$,

$$\sigma_p(\tilde{\mathscr{L}}) \subset \sigma_p(L_j). \tag{2.55}$$

In the $r = 1$ case one can in addition conclude from $\det T(\mu) = 1$ that $\sigma_p(\tilde{\mathscr{L}})$ is empty, but this argument fails for $r > 1$.

So far we have considered only the spectral analysis of $\mathscr{L} = \mathscr{L}_-$. The additional dependence of the functions $d_j(x^+, x^-)$ on x^+ was treated as purely parametric. In particular, it is consistent for the eigenvalues of $\mathscr{L}$ to depend on these external parameters. If, however, one is interested in the spectral analysis of the (compatible) pair $\mathscr{L}_-$, $\mathscr{L}_+$ a "time" dependence of the simultaneous eigenvalues is no longer permitted. One would expect that the space of simultaneous wavefunctions $\mathscr{L}_\pm W = 0$

becomes "smaller" and that suitably constrained scattering data are sufficient to construct a dense subspace. We impose the following

Condition B. The functions $t_a(\mu)$, $a = 1, \ldots, r$ are of Schwartz type on S and satisfy $t_a(\omega^a \mu) = 0$ for all $0 \neq a \in N(\mu)$, $\mu \in S \cap \mathbb{C}^+$.

Explicitly, the condition reads (for $a \neq 0$)

$$
\begin{aligned}
t_a(\omega^{a+b}\mathbb{R}^+) &= 0, \qquad 1 \leq a + b \leq [r/2], \\
t_a(\omega^{a+b}\mathbb{R}^-) &= 0, \qquad [(r+1)/2] + 1 \leq a + b \leq r.
\end{aligned}
\tag{2.56}
$$

Notice that (for $\varepsilon = 1$) this is compatible with $t_a(\mu)^* = t_{\bar{a}}(-\mu^*)$, which follows from the involution (2.8). The condition (2.56) guarantees that the minors defining $\Delta(\mu)$ in the different half-lines of S have trigonal form, so that

$$
\Delta(\mu) = \prod_{a \in N(\mu)} t_0(\omega^a \mu),
\tag{2.57}
$$

which, in particular, is independent of $x^{\pm}$. We shall refer to data satisfying condition B as AT-*reflective*. If all the $t_a(\mu)$, $a \geq 1$ vanish identically the spectrum can be seen to be purely discrete and in extension to the $r = 1$ case we shall refer to such data as *reflectionless*. In the AT-reflective case a continuous spectrum is present and the quantities $r_a(\mu) := \dfrac{t_a(\omega^a \mu)}{t_0(\mu)}$, $a \geq 1$ turn out to play the role of generalized *reflection coefficients*. Functions $t_a(\mu)$, $a \geq 1$ violating (2.56) may still serve as scattering data for the mKdV systems, although the reasoning of Sect. 3 will not go through. For the AT systems we expect the data satisfying condition B to be generic in the sense that they cover a subset of the phase space which is dense w.r.t. a suitable topology; although we have no proof (or counter-example) to offer.

3. Decoupled GLM-Equations

In this section we show that for the class of models considered, the inverse problem for AT-*reflective* data is equivalent to a set of *decoupled, scalar* GLM-type equations, which are solved in terms of τ-functions. To do this, a number of preparatory results are needed. Assume AT-reflective data and let μ_k, $k \in K \subset \mathbb{N}$ be an enumeration of the zeros of $\Delta(\mu)_+$ defined from (2.57). Clearly, K can be partitioned into (by condition A non-intersecting) subsets K_a for which $t_0(\omega^a \mu_k)$ vanishes. For these values of the spectral parameter the (components of the) Jost solutions have linear dependencies in addition to (2.47), (2.52).

3.1. Linear Dependencies at Zeros of $\Delta_+(\mu)$

Lemma 2. *For* $\mu = \mu_k$, $k \in K_a$ *the functions* $V_{j,0}^{(-)}(\omega^a \mu_k)$, $V_{j,1}^{(+)}(\omega^a \mu_k)$, $\ldots$, $V_{j,r}^{(+)}(\omega^a \mu_k)$ *are linearly related via*

$$
V_{j,0}^{(-)}(\omega^a \mu_k) = \sum_{b=1}^{r} c_{b,k} V_{j,b}^{(+)}(\omega^a \mu_k),
\tag{3.1}
$$

for complex constants $c_{b,k}$ *and* $k \in K_a$.

Proof Recall the definition of the Wronskian in Sect. 2.6 and consider $W[V_{j,0}^{(-)},$
$V_{j,1}^{(+)}, \ldots, V_{j,r}^{(+)}]$, where the argument is $(x^-; \omega^a\mu)$ in all entries. Then use (2.36) to
express $V_{j,0}^{(-)}(\omega^a\mu)$ in terms of $V_{j,b}^{(+)}(\omega^a\mu)$, $1 \le b \le r$. Rearranging the columns given

$$W[V_{j,0}^{(-)}, V_{j,1}^{(+)}, \ldots, V_{j,r}^{(+)}] = t_0(\omega^a\mu) W[V_{j,0}^{(+)}, \ldots, V_{j,r}^{(+)}]. \quad \Box \qquad (3.2)$$

Remark. A similar calculation shows

$$W[V_{j,a}^{(-)}, V_{j,1}^{(+)}, \ldots, V_{j,r}^{(+)}] = t_{\bar{a}}(\mu), \qquad\qquad (3.3)$$

with arguments $(x^-; \mu)$ in all entries. For AT-reflective data the r.h.s. vanishes for
$\bar{a} \in N(\omega^a\mu)\backslash\{0\}$, i.e. for $a \ne 0$ and $\operatorname{Im}\mu > 0$. Thus,

$$V_{j,a}^{(-)}(\mu) = \sum_{b=1}^{r} c_{a,b} V_{j,b}^{(+)}(\mu), \qquad\qquad (3.4)$$

valid for all $\mu \in \mathbb{C}^+$. In particular, for $\mu = \mu_k \in \mathbb{C}^+$, $k \in K_a$ the subset $N(\mu_k)\backslash\{0\}$
of a's also satisfies $\operatorname{Im}(\omega^a\mu_k) > 0$ and can be used to eliminate $V_{j,a}^{(-)}(\mu_k)$ from the
l.h.s. of the general relation (2.47); which reproduces (3.1).

   For later use we note also that (3.1) can be reformulated in terms of the columns
of $W^{(U,\pm)}$. Let $\underline{e}_b = (\delta_{ab})_{0 \le a \le r}$, $b = 0, \ldots, r$ be the standard basis vectors of $\mathbb{R}^{r+1}$
in column notation. Set

$$\underline{U}_b^{(\pm)}(\mu) := (U_{0,b}^{(\pm)}(\mu), \ldots, U_{r,b}^{(\pm)}(\mu))^T$$
$$:= W^{(U,\pm)}(\mu)\underline{e}_b = (w_{a-b}^{(\pm)}(x, \omega^a\mu))_{0 \le a \le r} \qquad (3.5)$$

for the bth column of the matrix-Jost solutions $W^{(U,\pm)}$. Inserting (2.31) into (3.1)
gives

$$w_c^{(-)}(\omega^{a+c}\mu_k) = \sum_{b=1}^{r} c_{b,k} w_{c-b}^{(+)}(\omega^{a+c}\mu_k), \qquad k \in K_a, \qquad (3.6)$$

which is equivalent to

$$\underline{U}_0^{(-)}(x; \omega^a\mu_k) = \sum_{b=1}^{r} c_{b,k}\underline{U}_b^{(+)}(x; \omega^a\mu_k). \qquad\qquad (3.7)$$

A relation of this form can of course also directly be obtained from the matrix formula-
tion, where one finds from (2.38) and $\det W^{(U,\pm)} = 1$ that $\det[\underline{U}_0^{(-)}, \underline{U}_1^{(+)}, \ldots, \underline{U}_r^{(+)}] = $
$t_0(\omega^a\mu)$ [again with arguments $(x; \omega^a\mu)$ on the r.h.s.]. The above derivation shows in
addition that the constants $c_{b,k}$ in (3.1) and (3.7) coincide. The same argument has
been used to show that the constants $c_{a,k}^{\pm}$ in (2.47) and (2.52) coincide.

3 2. Matrix GLM Equations

Recall the integral representation (2.39) for the columns of $W^{(U,+)}$. Further consider
the first column of the defining relation (2.38) for T^U,

$$\underline{U}_0^{(-)}(x; \mu) = \sum_{b=0}^{r} t_b(\omega^b\mu)\underline{U}_0^{(+)}(x; \mu) \qquad\qquad (3.8)$$

(again suppressing the "time" variables). For real μ rewrite this as

$$\frac{1}{t_0(\omega^a \mu)} \underline{U}_0^{(-)}(x; \omega^a \mu) - \underline{e}_0 e^{-ix\omega^a \mu}$$

$$= \underline{U}_0^{(+)}(x; \omega^a \mu) - \underline{e}_0 e^{-ix\omega^a \mu} + \sum_{b=1}^{r} \frac{t_b(\omega^{a+b} \mu)}{t_0(\omega^a \mu)} \underline{U}_b^{(+)}(x; \omega^a \mu), \qquad (3.9)$$

where for $\mu \in \mathbb{C}$, $a = a(\mu)$ is supposed to obey $\mathrm{Im}(\omega^a \mu) \geq 0$. In particular, on the real line this means explicitly that $0 \leq a(\mu) = a(+) \leq [(r+1)/2]$ for $\mu \in \mathbb{R}^+$ and $0 \leq a(\mu) = a(-) \leq [(r+1)/2]$ for $\mu \in \mathbb{R}^-$. To simplify the notation we will suppress the dependence of a on μ in the following. Take now the Fourier transform $\int_{-\infty}^{\infty} d\mu\, e^{iy\omega^a \mu}$ of (3.9). For $x < y$ the l.h.s. can be evaluated by complex contour deformation. The fall-off properties (2.46) imply

$$\underline{U}_0^{(\pm)}(x; \omega^a \mu) \to \underline{e}_0 e^{-ix\omega^a \mu} + o(1)$$

and $t_0(\omega^a \mu) \to 1$ as $|\mu| \to \infty$, $\mathrm{Im}\,\omega^a \mu > 0$, so that the integrand vanishes in this limit. Thus, for $x < y$ the contour can be closed in the upper half μ-plane and the integral receives contributions only from the simple poles of $1/t_0(\omega^a \mu)$,

$$\text{l.h.s.} = \sum_{k \in K_a} \operatorname*{res}_{\mu = \mu_k} \left[\frac{1}{t_0(\omega^a \mu)} \underline{U}_0^{(-)}(x; \omega^a \mu)\, e^{iy\omega^a \mu} \right]. \qquad (3.10)$$

Using (3.7) and the integral representation for $\underline{U}_b^{(+)}(x; \mu)$ this becomes

$$\text{l.h.s.} = \sum_{k \in K_a} \sum_{b=1}^{r} \frac{\omega^{-a} c_{b,k}}{\dot{t}_0(\omega^a \mu_k)} \underline{e}_b\, e^{i\omega^a \mu_k(y - \omega^b x)}$$

$$+ \sum_{k \in K_a} \sum_{b=1}^{r} \frac{\omega^{-a} c_{b,k}}{\dot{t}_0(\omega^a \mu_k)} \int_x^{\infty} dz\, \Gamma^{(+)}(x, z) \underline{e}_b\, e^{i\omega^a \mu_k(y - \omega^b z)}, \qquad (3.11)$$

where $\dot{t}_0 = \dfrac{d}{d\mu} t_0(\mu)$. For the r.h.s. one finds upon insertion of the integral representation (2.45),

$$\text{r.h.s.} = \Gamma^{(+)}(x, y) \underline{e}_0 + \sum_{b=1}^{r} \underline{e}_b \int_{-\infty}^{\infty} d\mu\, e^{i\omega^a \mu(y - \omega^b x)} \frac{t_b(\omega^{a+b} \mu)}{t_0(\omega^a \mu)}$$

$$+ \int_x^{\infty} dz\, \Gamma^{(+)}(x, z) \left(\sum_{b=1}^{r} \underline{e}_b \int_{-\infty}^{\infty} d\mu\, e^{i\omega^a \mu(y - \omega^b z)} \frac{t_b(\omega^{a+b} \mu)}{t_0(\omega^a \mu)} \right). \qquad (3.12)$$

Equating both sides results in

$$\Gamma^{(+)}(x, y) \underline{e}_0$$

$$+ \sum_{b=0}^{r} \left\{ F_b^{(+)}(x, y) \underline{e}_b + \int_x^{\infty} dz\, \Gamma^{(+)}(x, z) \underline{e}_b F_b^{(+)}(z, y) \right\} = 0, \qquad x < y, \quad (3.13)$$

where

$$F_b^{(+)}(x,y) = - \sum_{k \in K_a} \omega^{-a} C_b(\omega^a \mu_k) e^{i\omega^a \mu_k(y - \omega^b x)}$$

$$+ \int_{-\infty}^{\infty} d\mu \, e^{i\omega^a \mu(y - \omega^b x)} r_b(\omega^a \mu) \,, \tag{3.14}$$

introducing the notation $C_b(\mu_k) := \dfrac{c_{b,k}}{t_0(\mu_k)}$ and $r_b(\mu) := \dfrac{t_b(\omega^b \mu)}{t_0(\mu)}$.

Remark. We have suppressed the a-dependence in (3.13) and (3.14) because it just labels the decomposition of the discrete spectrum into the subsets K_a relative to the real integration variable μ of the continuous spectrum contributions; the only subtlety being the μ-dependence of a.

It is convenient to rewrite the result (3.13) in matrix form. To do this, multiply (3.13) by Λ^{-c} from the left and sum up the resulting equations for $c = 0, \ldots, r$. Using $\Lambda^{-c} \underline{e}_b = \underline{e}_{b+c}$ the first term in (3.13) just generates $\Gamma^{(+)}(x,y)$. The column $\sum_{b=1}^{r} F_b^{(+)} \underline{e}_b = (0, F_1^{(+)}, \ldots, F_r^{(+)})^T$ generates $F^{(+)} := \sum_{c=1}^{r} F_c^{(+)} \Lambda^{-c}$. Together

$$\Gamma^{(+)}(x,y) + F^{(+)}(x,y) + \int_{x}^{\infty} dz \, \Gamma^{(+)}(x,z) F^{(+)}(z,y) = 0 \,, \qquad x < y \,. \tag{3.15}$$

This is the *matrix*-GLM *equation from the right*. It generalizes the corresponding equations for various rank 1 systems such as the (m)KdV system [21], the SG model in lightcone dynamics [8, Sects. II.5, 7] and (essentially) the NLS model [8, Sect. II.4].

By a similar chain of arguments one arrives at the *matrix*-GLM *equation from the left*

$$\Gamma^{(-)}(x,y) + F^{(-)}(x,y) + \int_{-\infty}^{x} dz \, \Gamma^{(-)}(x,z) F^{(-)}(z,y) = 0 \,, \qquad x > y \,, \tag{3.16}$$

where $F^{(-)} = \sum_{b=1}^{r} F_b^{(-)} \Lambda^b$,

$$F_b^{(-)}(x,y) = - \sum_{k \in K_a} \omega^a C_b(\omega^a \mu_k)^* e^{-i\omega^{-a} \mu_k^*(y - \omega^{-b} x)}$$

$$+ \int_{-\infty}^{\infty} d\mu \, e^{-i\omega^{-a} \mu(y - \omega^{-b} x)} r_b(\omega^a \mu)^* = F^{(+)}(x,y)^* \,. \tag{3.17}$$

We just give some comments on the derivation. The starting point is the complex conjugate of Eq. (3.9). The condition $x > y$ arises because now $\int_{-\infty}^{x} dz \, \Gamma^{(-)}(x,z) \times \delta(y-z)\underline{e}_0$ should result in $\Gamma^{(-)}(x,y)\underline{e}_0$. Since $x > y$ one has to use the complex conjugate of the integral representation (2.45) and the Fourier transform $\int_{-\infty}^{\infty} d\mu \, e^{-iy\omega^a \mu}$

to obtain the required fall-off properties of the integrand on the l.h.s. This leads to

$$\Gamma^{(-)}(x,y)^* \underline{e}_0$$
$$+ \sum_{b=0}^{r} \left\{ F_b^{(-)}(x,y)\underline{e}_b + \int_{-\infty}^{x} dz\, \Gamma^{(-)}(x,z)^* \underline{e}_b F_b^{(-)}(z,y) \right\} = 0\,, \qquad x > y\,. \quad (3.18)$$

When rewritten in matrix form one can apply the involution $\Gamma^{(-)}(x,y)^* = J\Gamma^{(-)}(x,y)\,J(*)$. For $\varepsilon = 1$ the relation $(*)$ is a consequence of (2.8) and the integral representation (2.14). For $\varepsilon = i$ one can compensate the change of sign in $\varepsilon \to \varepsilon^*$ by changing the sign of β. The latter is not an automorphism of the equations of motion for $\phi \cdot H$ but it is one for the bilinear identities (A.9). In Sect. 4 it will turn out that $\Gamma^{(-)}(x,y)$ is expressible in terms of the τ-functions, so that $(*)$ holds for both $\varepsilon = 1$ and $\varepsilon = i$. From $J\Lambda J = \Lambda^{-1}$ one finally obtains (3.16).

3.3. Scalar GLM Equations

With these preparations we can formulate

Theorem 1. *For the affine Toda systems the inverse problem for AT-reflective scattering data is equivalent to a set of scalar, decoupled GLM-type equations*

$$K_j^{(\pm)}(x,y) + F_j^{(\pm)}(x,y) + \int_{I_x^\pm} dz\, K_j^{(\pm)}(x,z)\, F_j^{(\pm)}(x,z) = 0\,, \qquad x \lessgtr y\,, \quad (3.19)$$

for $j = 0, \ldots, r$.

$$F_j^{(\pm)} = \sum_{b=1}^{r} \omega^{jb} F_b^{(\pm)}\,,$$

$$F_b^{(+)}(x,y) = -\sum_{k \in K} \omega^{-a} C_b(\omega^a \mu_k)\, e^{\xi(y, i\omega^a \mu_k) - \xi(x, i\omega^{a+b} \mu_k)}$$
$$\qquad\qquad + \int_{-\infty}^{\infty} d\mu\, r_b(\omega^a \mu)\, e^{\xi(y, i\omega^a \mu) - \xi(x, i\omega^{a+b} \mu)} = F_b^{(-)}(x,y)^*\,. \quad (3.20)$$

The flow variables $x_n^\pm$, $n \in E$ enter via

$$\xi(x,\mu) = \sum_{n \in E} (x_n^+ \mu^{-n} + x_n^- (-\mu)^n)\,, \quad (3.21)$$

where in (3.19) $(x,y) := (x_1^-, y_1^-)$ etc., the other variables being suppressed.

Proof. Consider the bth row of the vector equation (3.13) for the "+" case, say

$$K_b^{(+)}(x,y) + F_b^{(+)}(x,y) + \sum_{c=0}^{r} \int_x^{\infty} dz\, K_{b-c}^{(+)}(x,z)\, F_c^{(+)}(z,y) = 0\,, \qquad x < y\,, \quad (3.22)$$

Operate with $\sum\limits_{b=1}^{r} \omega^{jb}$ on this equation to find

$$K_j^{(+)}(x,y) + F_j^{(+)}(x,y) + \int\limits_x^\infty dz\, K_j^{(+)}(x,z)\, F_j^{(+)}(x,y) = 0\,, \qquad x < y\,, \qquad (3.23)$$

with the notation (2.43) and $F_j^{(+)} = \sum\limits_{b=1}^{r} \omega^{jb} F_b^{(+)}$. The dependence on the flow variables follows from (2.22) and Apendix A. Similarly one obtains the equations from the left. $\square$

Equations (3.19) are the main result of this section. They constitute a set of *scalar, decoupled* GLM-type equations, where the kernels $F_j^{(\pm)}(x,y)$ are specified in terms of the scattering data through (3.20).

Remark. Notice that Eqs. (3.19) do *not* coincide with the equations for the resolvent kernels of the Volterra integral operators $\hat{F}_j^{(\pm)}(x,y)$ (where the role of F_j and K_j would be interchanged). Note further that also at the level of Eqs. (3.19) the $r = 1$ case is special in that the restriction (2.26) is empty and the kernel $F_{b=1}^{(\pm)}$ is symmetric $F_1^{(\pm)}(x,y) = F_1^{(\pm)}(y,x)$. Only in this case are the GLM equations equivalent to a local RH-problem [17].

An alternative derivation of Eqs. (3.19) starts directly from the scalar linear problem (2.34). For the scalar Jost solutions $V_{j,a}^{(\pm)}$ one has the relation (2.36). Rewrite the $a = 0$ equation as

$$\frac{1}{t_0(\omega^a\mu)}\, V_{j,0}^{(-)}(x;\omega^a\mu) - e^{-ix\omega^a\mu}$$

$$= V_{j,0}^{(+)}(x;\omega^a\mu) - e^{-ix\omega^a\mu} + \sum_{b=1}^{r} r_b(\omega^a\mu)\, V_{j,b}^{(+)}(x;\omega^a\mu)\,, \qquad (3.24)$$

and take the Fourier transform $\int\limits_{-\infty}^{\infty} d\mu\, e^{iy\omega^a\mu}$. From the falloff properties (2.44) the l.h.s. can again be evaluated by complex contour deformation. Using (3.1) and (2.42) one finds for $x < y$

$$\text{l.h.s.} = - \sum_{K_a} \sum_{b=1}^{r} C_b(\omega^a\mu_k)\, e^{i\omega^a\mu_k y}$$

$$+ \sum_{k\in K_a} \sum_{b=1}^{r} \int\limits_x^\infty dz\, K_j^{(+)}(x,z) \left(\int\limits_{-\infty}^{\infty} d\mu\, e^{i\omega^a\mu(y-\omega^b z)} r_b(\omega^a\mu) \right). \qquad (3.25)$$

On the r.h.s. insertion of the integral representation (2.42) gives for $x < y$,

$$\text{r.h.s.} = K_j^{(+)}(x,y) + \sum_{b=1}^{r} \omega^{jb} \int\limits_{-\infty}^{\infty} d\mu\, e^{i\omega^a\mu(y-\omega^b x)} r_b(\omega^a\mu)$$

$$+ \int\limits_x^\infty dz\, K_j^{(+)}(x,z) \left(\sum_{b=1}^{r} \omega^{jb} \int\limits_{-\infty}^{\infty} d\mu\, e^{i\omega^a\mu(y-\omega^b z)} r_b(\omega^a\mu) \right). \qquad (3.26)$$

Equating both sides one recovers (3.19). Similarly a scalar derivation for the equations from the left is possible.

3.4. Interpretation as ZS-Dressing Problem

The GLM-type equations (3.19) can be interpreted as defining a dressing problem in the sense of Zakharov-Shabat [30] for the linear problem (2.34). Rewrite (2.42) as

$$V_{j,a}^{(\pm)}(\mu) = \omega^{ja}[\mathbb{1} + \hat{K}_j^{(\pm)}]e^{-ix\omega^a\mu}, \qquad (3.27)$$

for upper/lower Volterra integral operators $\hat{K}_j^{(\pm)}$. Since the Jost solutions $V_{j,a}^{(\pm)}$ form a fundamental system, both of these operators solve the *dressing equation*

$$L_j(\mathbb{1} + \hat{K}_j^{(\pm)}) = (\mathbb{1} + \hat{K}_j^{(\pm)})(i\partial_-)^{r+1}. \qquad (3.28)$$

Supposing that $\mathbb{1} + \hat{K}_j^{(\pm)}$ have an inverse (cf. the remark following Theorem 2) the operators

$$\begin{aligned}
\mathbb{1} + \hat{F}_j^{(+)} &= (\mathbb{1} + \hat{K}_j^{(+)})^{-1}(\mathbb{1} + \hat{K}_j^{(-)}), \\
\mathbb{1} + \hat{F}_j^{(-)} &= (\mathbb{1} + \hat{K}_j^{(-)})^{-1}(\mathbb{1} + \hat{K}_j^{(+)}),
\end{aligned} \qquad (3.29)$$

commute with $(i\partial_-)^{r+1}$ if and only if (3.28) holds. This means that $\hat{F}_j^{(\pm)}$ solves a linear equation and the mapping $L_j \to \hat{F}_j^{(\pm)}$ can be interpreted a direct ("scattering") transformation linearising the dynamics. The GLM-type equations (3.19) in addition yield a parametrization of $\hat{F}_j^{(\pm)}$ through scattering data. Conversely, for given $\hat{F}_j^{(\pm)}$ the inverse problem amounts to a factorization of $\hat{F}_j^{(\pm)}$ into upper/lower Volterra integral operators. Rewriting $(\mathbb{1} + \hat{K}_j^{(\pm)})(\mathbb{1} + \hat{F}_j^{(\pm)}) = (\mathbb{1} + \hat{K}_j^{(\mp)})$ in terms of the kernels one recovers the GLM-type equations (3.19).

4. Fredholm Determinant and Tau Function

4.1. Solution of the GLM Equations

Each of Eqs. (3.19) constitutes a family of Fredholm equations indexed by $x \in \mathbb{R}$ on a non-compact interval. These equations are no longer amenable to the classical fredholm theory. Moreover, even for continuous kernels the operators

$$\hat{F}_j^{(\pm)} : f(x) \to \int_{I_x^{(\pm)}} dz\, F_j^{(\pm)}(x, z)\, f(z)$$

are *not* trace-class. However these operators act as *nuclear* operators on a suitable Banach space of continuous functions on I_x, so that Grothendieck's generalized Fredholm theory can be applied. Some of the basic results and the application to the case at hand have been summarized in Appendix B. If the reflection coefficients $r_b(\mu)$ are of Schwartz type one readily verifies that the kernels $F_j^{(\pm)}$ are of type $\mathscr{C}(I_x^2)$ and the results of Theorems B.1 and B.2 apply. For $x \in \mathbb{R}$ define the family of

resolvent kernels by $(I + \hat{F}_j^{(\pm)})^{-1} f(x) = f(x) + \int_{I_x} K_j^{(\pm)}(x,y) f(y)$. Theorem B.2.b) then gives

$$K_j^{(\pm)}(x,y) = \frac{D_j^{(\pm)}(F \mid x, y)}{D_j^{(\pm)}(F \mid x)}, \tag{4.1}$$

where $D_j^{(\pm)}(F \mid x, y)$ and $D_j^{(\pm)}(F \mid x)$, $x \in \mathbb{R}$ are the family of Fredholm-Grothendieck minors and -determinants, respectively. The expansion coefficients in $D_j^{(\pm)}(zF \mid x, y) = \sum_{n \geq 0} D_{j,n}^{(\pm)}(F \mid x, y) z^n$ and $D_j^{(\pm)}(zF \mid x) = \sum_{n \geq 0} D_{j,n}^{(\pm)}(F \mid x) z^n$ are given by (B.4). To calculate the determinant explicitly we make use of the following

Lemma 3. *Let $d_n^{(+)}(x)$ denote the nth expansion coefficient of the determinant $D^{(+)}(x)$ with kernel*

$$F(x,y) = \int_{C^2} d\mu(p,q) \, e^{\xi(x,p) - \xi(y,q)}, \tag{4.2}$$

where $d\mu(p,q)$ is some measure on $\mathbb{C}^2$ having support only for $\operatorname{Re} p < \operatorname{Re} q$ and which is otherwise s.t. $F \in \mathscr{C}(I_x^2)$. Let $\psi(p)$ be a free Fermi field on semi-infinite wedge space ([15], p. 315) satisfying $\langle 0| \, \psi(p) \psi^*(q) \, |0\rangle = \dfrac{q}{p - q}$. Then:

$$D_n^{(+)}(\underline{x}) = \langle 0| \, A(\underline{x})^n \, |0\rangle,$$
$$A(\underline{x}) = \int \frac{d\mu(p,q)}{q} \, \psi(p) \psi^*(q) \, e^{\xi(x,p) - \xi(x,q)}, \tag{4.3}$$

where the powers of A are understood to be normal ordered.

Proof. Write $\underline{x}(s)$ for $s, x_2, \ldots, x_n, \ldots)$ and recall from (B.4)

$$D_n(x) = \int_{(I_x^+)^n} ds_1 \ldots ds_n F \begin{bmatrix} s_1 \ldots s_n \\ s_1 \ldots s_n \end{bmatrix}. \tag{4.4}$$

Explicitly,

$$F \begin{bmatrix} s_1 \ldots s_n \\ s_1 \ldots s_n \end{bmatrix} = \det(F(s_i, s_j))_{1 \leq i,j \leq n}$$

$$= \sum_{\pi \in S_n} (-)^\pi \prod_{l=1}^n F(s_l, s_{\pi(l)})$$

$$= \int_{C^{2n}} d\mu(p_1, q_1) \ldots d\mu(p_n, q_n)$$

$$\times \prod_{l=1}^n \sum_{\pi \in S_n} (-)^\pi \, e^{\xi(x(s_l), p_l) - \xi(x(s_l), q_{\pi^{-1}(l)})}. \tag{4.5}$$

The integrand is invariant under permutations of the p's and q's separately, so that the integration variables can be ordered according to $\operatorname{Re} p_1 < \ldots < \operatorname{Re} p_n < \operatorname{Re} q_n <$

$\ldots < \operatorname{Re} q_1$ at the expense of an overall factor $1/n!)^2$. In particular $\operatorname{Re} p_l < \operatorname{Re} q_m$, $1 \leq l, m \leq n$. The s-integration in (4.4) can now be done, the basic integral being

$$\int_x^\infty ds \, e^{\xi(\underline{x}(s),p)-\xi(\underline{x}(s),q)} = \frac{1}{p-q} \, e^{\xi(x,p)-\xi(x,q)} \,. \tag{4.6}$$

The ordering can now be lifted again so that

$$D_n(\underline{x}) = \int_{C^{2n}} d\mu(p_1,q_1) \ldots d\mu(p_n,q_n)$$

$$\times \prod_{l=1}^n \sum_{\pi \in S_n} \frac{(-)^\pi}{p_l - q_{\pi^{-1}(l)}} \, e^{\xi(x_l,p_l)-\xi(x_l,q_{\pi^{-1}(l)})}$$

$$= \int_{C^{2n}} d\mu(p_1,q_1) \ldots d\mu(p_n,q_n)$$

$$\times \det\left(\frac{1}{p_i - q_j}\right)_{1 \leq i,j \leq n} \prod_{l=1}^n e^{\xi(x_l,p_l)-\xi(x_l,q_l)} \,. \tag{4.7}$$

The integrand has an almost factorized by w.r.t. the $d\mu(p,q)$-integrations. A complete factorization can be achieved by inserting

$$\langle 0| \, \psi(p_1)\psi^*(q_1) \ldots \psi(p_n)\psi^*(q_n) \, |0\rangle = \det\left(\frac{q_j}{p_i - q_j}\right)_{1 \leq i,j \leq n} \,. \quad \square \tag{4.8}$$

We can now prove

Theorem 2. a) *The determinants $D_j^{(\pm)}(F \mid \underline{x})$ can be written as matrix elements in the wedge space realization of (a version of) the affine group $\widehat{SL}_{r+1}$ and define generalized τ-functions in the sense of the Kyoto School,*

$$D_j^{(\pm)}(F \mid \underline{x}) = \langle \Lambda_j| \, e^{\hat{E}(x)} g(A^{(\pm)}) e^{-\hat{E}(x)} \, |\Lambda_j\rangle = \tau_j(A^{(\pm)} \mid \underline{x}) \,. \tag{4.9}$$

In particular, they solve the bilinear equations (A.8). All of the quantities in (4.9) refer to the wedge space realization. The elements $g(A^{(\pm)})$ of $\widehat{SL}_{r+1}$ are defined in terms of Lie algebra elements $A^{(\pm)}$, which in turn are functions of the scattering data (given by Eqs. (4.17), (4.18) below). $|\Lambda_j\rangle$ is the h.w.v. of the jth fundamental representation of $\widehat{sl}_{r+1}$. The evolution operators $e^{\pm\hat{E}(x)}$ are defined in terms of the principle Heisenberg subalgebra $[\hat{E}_n, \hat{E}_m] = m\delta_{m+n,0}K$ of $\widehat{sl}_{r+1}$ via $\hat{E}(x) = \sum_{n \in E} [x_n^+ \hat{E}_{-n} + (-)^n x_n^- \hat{E}_n]$.

b) *In terms of the τ-functions the solution of the affine Toda equations is given by*

$$\Phi = -\frac{1}{\varepsilon\beta} \sum_{j=0}^r \check{\alpha}_j \ln \tau_j(A^{(\pm)}) + \zeta(A^{(\pm)}) K \,, \tag{4.10}$$

where the "+" and "−" parts of the solution match whenever the $A^{(\pm)}$ initial data are consistent.

Remark The formulation in a) explicitly refers to the wedge space realization and avoids the notion of the dressing group (see [2] for a discussion) for the following reason. The expectation values in $\widehat{SL}_{r+1}$, i.e. the τ-functions are *defined* here in terms of the Fredholm-Grothendieck determinant. The latter are *not* determinants of trace-class operators, so that the resulting τ-functions are "generalized" τ-functions. The normal ordered exponentials $e^{A^{(\pm)}}$, although by construction well-defined in the wedge space realization, need not correspond to a (realization-independent) element $g(A^{(\pm)}) = g_-^{-1}g_+$ of the standard dressing group. The Bruhat decomposition for (centrally extended) loop groups, which usually guarantees this gives τ-functions which are determinants of trace-class operators. To ensure the existence of a dressing group the determinant bundles of the Grassmannian formulation [28] would have to be generalized to nuclear operators on Banach spaces (cf. Appendix B).

Proof. a) We wish to apply Lemma 3 to the kernels $F_j^{(\pm)}$ in (3.20). Clearly, choosing for the measure $d\mu(p, q)$ a suitable sum of δ-function contributions, one can produce the kernels $F_j^{(\pm)}$, which have already been seen to be of type $\mathscr{C}(I_x^2)$. Consider first the "+" case. In the analogue of Eq. (4.5) there will appear two types of exponentials of the form $e^{\xi(s_l, p_l) - \xi(s_l, q_{\pi^{-1}(l)})}$, with the substitutions,

$$p_l = i\omega^{a(\mu_{k_l})}\mu_{k_l}, \quad q_l = i\omega^{a(\mu_{k_l})+b_l}\mu_{k_l}, \quad \text{for the discrete part,}$$

$$p_l = i\omega^{a(\mu_l)}\mu_l, \quad q_l = i\omega^{a(\mu_l)+b_l}\mu_l, \quad \text{for the continuous part.}$$

To perform the s-integrations we have to show that $\operatorname{Re} p_l < \operatorname{Re} q_l$ for all l. For the continuous part this is guaranteed by condition B on the reflection coefficients. If $\mu \in \mathbb{R}^+$ one needs

$$\operatorname{Re}[i\mu\omega^{a(+)}(1 - \omega^b)] = \mu(\sin(\theta_{a(+)} + \theta_b) - \sin\theta_{a(+)}) < 0,$$

[where $\theta_a = 2\pi a/(r + 1)$] at least for those $b \in \{1, \ldots, r\}$ for which $t_b(\omega^b)$ is non-vanishing. By condition B the latter is the case only if $\sin(\theta_{a(+)} + \theta_b) < 0$, for given $a(+) \in \{0, \ldots, [(r+1)/2]\}$. But then also $\sin\theta_{a(+)}$ is non-negative and $\operatorname{Re} p < \operatorname{Re} q$ holds. Similarly one checks

$$\operatorname{Re}[i\mu\omega^{a(-)}(1 - \omega^b)] = \mu(\sin(\theta_{a(-)} + \theta_b) - \sin\theta_{a(-)}) < 0,$$

for $\mu \in \mathbb{R}^-$. This means that for the continuous spectrum contributions the boundary terms at $x^- = \pm\infty$ vanish. For the discrete spectrum contributions the analogous condition is $\operatorname{Re}[i\mu_k\omega^{a(\mu_k)}(1 - \omega^b)] < 0$. By construction $\mu_k\omega^{a(\mu_k)}$ is of the form $p_k e^{i\alpha_k}$, where $p_k > 0$ and $0 < \alpha_k < \pi$. The condition then becomes $p(\sin(\alpha_k + \theta_{b_k}) - \sin\alpha_k) < 0$. In principle this gives a constraint on the possible b-sectors which contribute to $t_0(\mu\omega^a)$. In practice also μ_k will depend on the type b so that the discussion is specific for each class of excitations in the discrete spectrum. We defer the details to a later publication but anticipate that usually no restriction arises. Thus, the boundary terms vanish also for the discrete spectrum contributions

and the result of the s-integration is

$$D_{j,n}^{(+)}(\underline{x}) = \int\limits_{(I_x^+)^n} ds_1 \dots ds_n \det(F_j^{(+)}(s_i, s_j))_{1 \le i,j \le n}$$

$$= \sum_{b_1 \dots b_n} \omega^{j(b_1 + \dots + b_n)} D_{b_1 \dots b_n}^{(+)};$$ \hfill (4.11)

$$D_{b_1 \dots b_n}^{(+)} = (-)^n \sum_{k_1 \dots k_n \in K_a} \left[C_{b_1}(-ip_1) \dots C_{b_n}(-ip_n) \right.$$

$$\left. \times \det \left(\frac{1}{p_i - q_j} \right) \prod_{l=1}^n e^{\xi(x_l, p_l) - \xi(x_l, q_l)} \right]_{\substack{p_l = i\omega^{a(\mu_{k_l})} \mu_{k_l} \\ q_l = i\omega^{a(\mu_{k_l}) + b_l} \mu_{k_l}}}$$

$$+ \int\limits_{-\infty}^{\infty} d\mu_1 \dots d\mu_n \left[r_{b_1}(-ip_1) \dots r_{b_n}(-ip_n) \right.$$

$$\left. \times \det \left(\frac{1}{p_i - q_j} \right) \prod_{l=1}^n e^{\xi(x_l, p_l) - \xi(x_l, q_l)} \right]_{\substack{p_l = i\omega^{a(\mu_l)} \mu_l \\ q_l = i\omega^{a(\mu_l) + b_l} \mu_l}},$$ \hfill (4.12)

so that

$$D_{j,n}^{(+)} = \langle 0 | (A_j^{(\pm)}(\underline{x}))^n | 0 \rangle,$$ \hfill (4.13)

where $A_j^{(+)}(\underline{x}) = \sum\limits_{b=1}^r \omega^{jb} A_b^{(+)}(\underline{x})$ and

$$A_b^{(+)}(\underline{x}) = - \sum_{k \in K_a} \left[\frac{C_b(-ip)}{\omega^{a(\mu_k)} q_k} \psi(p) \psi^*(q) e^{\xi(x,p) - \xi(x,q)} \right]_{\substack{p_k = i\omega^{a(\mu_k)} \mu_k \\ q_k = i\omega^{a(\mu_k) + b} \mu_k}}$$

$$+ \int\limits_{-\infty}^{\infty} d\mu \left[\frac{r_b(-ip)}{q} \psi(p) \psi^*(q) e^{\xi(x,p) - \xi(x,q)} \right]_{\substack{p = i\omega^{a(\mu)} \mu \\ q = i\omega^{a(\mu) + b} \mu}}.$$ \hfill (4.14)

Recall now that the fundamental representations ϱ_j, $0 \le j \le r$ of $\hat{A}_r^{(1)}$ admit a fermionic realization on semi-infinite wedge space $\wedge C^\infty$ [14, 15]. In particular, recall that finite or infinite (if well-defined) superpositions of bilinears $\psi(p) \psi^*(q)$ realize elements of $\widehat{sl}_{r+1}$ iff $p/q = \omega^b$ for $b \in \{0, \dots, r\}$. The latter is the case for all the terms in (4.14), so that (for $r_b(\mu)$ of Schwartz-type) (4.14) defines an element of $\widehat{sl}_{r+1}$. Further let $\hat{E}_n$, $n \in \mathbb{Z}$ be the generators of the principle Heisenberg subalgebra in this realization and set $\hat{E}(x) = \sum\limits_{n \in E} (x_n^+ \hat{E}_{-n} + (-)^n x_n^- \hat{E}_n)$. Then

$$e^{\hat{E}(x)} \psi(p) e^{-\hat{E}(x)} = e^{\xi(x,p)} \psi(p),$$
$$e^{\hat{E}(x)} \psi^*(p) e^{-\hat{E}(x)} = e^{-\xi(x,p)} \psi^*(p).$$ \hfill (4.15)

If finally $|\Lambda_j\rangle$ denotes the h.w.v. in the wedge space realization one has

$$\langle \Lambda_j | \psi(p) \psi(q)^* | \Lambda_j \rangle = \frac{(p/q)^j}{1 - (p/q)} = \omega^{-bj} \langle 0 | \psi(p) \psi(q)^* | 0 \rangle.$$

for all the terms in (4.14). Combining (4.11)–(4.15) one ends up with

$$D_{j,n}^{(+)}(\underline{x}) = \langle \Lambda_j | \, e^{\hat{E}(x)} (A^{(+)})^n \, e^{-\hat{E}(x)} \, | \Lambda_j \rangle \,, \tag{4.16}$$

where $A^{(+)} = \sum\limits_{b=1}^{r} A_b^{(+)}$ and (notice that the sign of b flips)

$$
\begin{aligned}
A_b^{(+)} = -\sum_{k \in K_a} & \left[\frac{C_b(-ip_k)}{\omega^{a(\mu_k)} q_k} \, \psi(p_k) \psi^*(q_k) \right]_{\substack{p_k = i\omega^{a(\mu_k)} \mu_k \\ q_k = i\omega^{a(\mu_k)-b} \mu_k}} \\
& + \int_{-\infty}^{\infty} d\mu \left[\frac{r_b(-ip)}{q} \, \psi(p) \psi^*(q) \right]_{\substack{p = i\omega^{a(\mu)} \mu \\ q = i\omega^{a(\mu)-b} \mu}} \,.
\end{aligned}
\tag{4.17}
$$

Finally summing up the Fredholm series yields the claim. The result for the "$-$" sector is obtained similarly. The conditions for the vanishing of the boundary terms coincide with that of the "$+$" case (since Re $z = $ Re z^*) and the s integrations can be performed. The element $A^{(-)} = \sum\limits_{b=1}^{r} A_b^{(-)}$ takes the form

$$
\begin{aligned}
A_b^{(-)} = -\sum_{k \in K_a} & \left[\frac{C_b(-ip_k)^*}{\omega^{-a(\mu_k)} q_k} \, \psi(p_k) \psi^*(q_k) \right]_{\substack{p_k = -i\omega^{-a(\mu_k)} \mu_k \\ q_k = -i\omega^{-a(\mu_k)+b} \mu_k}} \\
& + \int_{-\infty}^{\infty} d\mu \left[\frac{r_b(-ip)^*}{q} \, \psi(p) \psi^*(q) \right]_{\substack{p = -i\omega^{-a(\mu)} \mu \\ q = -i\omega^{-a(\mu)+b} \mu}} \,.
\end{aligned}
\tag{4.18}
$$

b) From Eq. (4.9) it is clear that the dynamics of $D_j^{(\pm)}(x)$ is the same for both, the "$+$" and the "$-$" sectors. But this means that the matching of the "$+$" and the "$-$" parts of the solution is automatically guaranteed whenever the initial data are consistent, i.e. whenever the branches of the logarithms can be choosen s.t. $\sum\limits_{j=0}^{r} \check{\alpha}_j \ln \tau_j(A^{(+)} \,|\, 0) = \sum\limits_{j=0}^{r} \check{\alpha}_j \ln \tau_j(A^{(-)} \,|\, 0)$ and $\zeta(A^{(+)} \,|\, 0) = \zeta(A^{(-)} \,|\, 0)$. $\square$

Remark. i. Since $A^{(-)} = (A^{(+)})^*$ the τ-functions $\tau_j(A^{(\pm)} \,|\, x)$ will just be complex conjugates of each other. The branches of the logarithms can therefore always be adjusted to achieve $\Phi(A^{(+)} \,|\, 0) = \Phi(A^{(-)} \,|\, 0)$ for the initial data. Theorem 2 in particular then guarantees that both parts of the solution match consistently under the evolution of all the $x_n^\pm$ flows and the "$\pm$" parts of the solution will be construction always admit an extension to the vicinity of $x^- \to \mp\infty$. Thus, it suffices to consider only one part of the solution. In particular, we may simplify the notation and work with the "$+$" sector alone, dropping the "$+$" superscripts in all quantities.
ii. Theorem 2 holds with minor changes also for the mKdV hierarchies. From the discussion in Appendix A it follows that the mKdV lows can be identified with the $\partial_{x_n^-}$-flows of the AT hierarchy. Thus, dropping x_n^+ variables in (4.10) yields τ-functions for the mKdV system. The mKdV variables are recovered from

$$q_j = -\frac{1}{\varepsilon\beta} \, \partial_x \ln \left(\frac{\tau_j^{(\pm)}}{\tau_{j-1}^{(\pm)}} \right) = -\frac{1}{\varepsilon\beta} \left(K_j^{(\pm)}(x,x) - K_{j-1}^{(\pm)}(x,x) \right), \tag{4.19}$$

using $K_j^{(\pm)}(x,x) = \partial_{x_1} \ln D_j^{(\pm)}(x)$ for the second equality. As remarked before the AT-reflective data are however presumably not generic for the mKdV systems.

The preceding remark applies whenever the initial value problem is well-defined. Here we consider affine Toda theories in ∂_+-lightcone dynamics. As in the SG model [16, 8] this implies that the initial data $\Phi(0, x^-)$ are subject to an infinite number of constraints. One can argue that these constraints are satisfied if $t_a(\mu)$, $a > 0$ vanish together with all its derivatives at $\mu = 0$. For the first few constraint equations and in the linearized limit this can be checked explicitly. In particular, in the linearized limit it is a well-known property of lightcone dynamics to enforce the Fourier transform of the initial data to vanish at the origin together with all its derivatives. It is then convenient to reverse the viewpoint ([8] p. 451) and define the constraints (for the class of solutions considered) by the condition on $t_a(\mu)$, $a > 0$. Thus, for a discussion of the initial value problem in lightcone dynamics, we impose finally

Condition C. $\left(\dfrac{d}{d\mu}\right)^n t_a(\mu)\Big|_{\mu=0} = 0,\ n \geq 0,\ a = 1, \ldots, r.$

The identification of the τ-functions with Fredholm determinants also gives the

Corollary 1. a) *For* AT-*reflective data (in particular $r_a(\mu)$, $a > 0$ of Schwartz type on S) and* $\lim\limits_{n\to\infty} \sup |x_n^\pm|^{1/n} = 0$ *the τ-functions (4.9) are $\mathscr{C}^\infty$ in all its arguments.*
b) *Scaling the Lie algebra element by $z \in \mathbb{C}^*$, the functions $\tau_j(zA^{(\pm)}) \,|\, x)$ are entire analytic functions of z. They admit a product representation of the form* B.1.b) *with $x^\pm$-dependent spectral values $\lambda_k(x^+, x^-)$.*

Proof. a) The condition on the reflection coefficients guarantees that the kernels are of type $\mathscr{C}(I_x^2)$ (cf. Appendix B). The condition on the flow variables ensures that $\xi(x; \mu)$ is analytic in its arguments. The proof then rests on a standard argument based on Hadamards inequality, which is also used to prove the convergence of the Fredholm expansions (B.1), (B.2). We omit the details. Part b) is a consequence of Theorem B.1. $\square$

Remark The corollary does not imply that the solutions Φ of the field equations are $\mathscr{C}^\infty$ because the evolutions in $x_n^\pm$ may cross the branch cuts of the logarithms. This can be detected from the product representation B.1.b). At the zeros of this product the operators $I + \hat{K}_j^{(\pm)}$ have a nontrivial kernel and the solution Φ crosses a branch cut. In principle, the product representation for $\tau_j^{(\pm)}$ would allow also to deduce restriction on the eigenvalues $\lambda_k(x^+, x^-)$ and to discuss their dependence on the scattering data, but this is beyond the scope of the present paper.

As a further application of Theorem 2 we can also express the Jost solutions in terms of the τ-functions.

Corollary 2. *The Jost solutions $V_{j,a}^{(\pm)}$ of the scalar linear problem (2.34) are given by*

$$V_{j,0}^{(\pm)} = \frac{\tau_j(x^+, \pm x^- - \varepsilon((\mp i\mu)^{-1}))}{\tau_j(x^+, \pm x^-)}\, e^{\xi(x^+, \pm x^- ; \mp i\mu)}, \qquad (4.20)$$

and $V_{j,a}^{(\pm)}(\mu) = \omega^{ja} V_{j,0}(\omega^a \mu)$. The Jost solutions $W^{(\pm)}$ of the matrix linear problem (2.2) are given by (2.6) with

$$w_a^{(\pm)}(\mu) = \frac{1}{r+1} \sum_{j=0}^{r} \omega^{-ja}\, \frac{\tau_j(x^+, \pm x^- - \varepsilon((\mp i\mu)^{-1}))}{\tau_j(x^+, \pm x^-)}\, e^{\xi(x^+, \pm x^- ; \mp i\mu)}. \qquad (4.21)$$

Proof. We recall from [29] that there exists a Hirota-Hierarchy of the form,

$$H_n^{(+)}[D_1, \ldots, D_n]\tau_j v_j = G_n^{(+)}[D_1, \ldots, D_n]\tau_{j+1}v_{j-1},$$

$$H_n^{(-)}[D_1, \ldots, D_n]\tau_{j+1}v_j = G_n^{(-)}[D_1, \ldots, D_n]\tau_j v_{j+1}, \quad n \geq 1,$$

where $H_n^{(\pm)}$, $G_n^{(\pm)}$ are Hirota polynomials of degree n in the Hirota operators $D_n = D_n^{\pm}$ for $x_n^{\pm}$. Further $v_j(x^+, x^-; \mu) = \tau_j(x^+, x^- - \varepsilon((-i\mu)^{-1})e^{\xi(x^-;-i\mu)}$ and $\varepsilon(\mu) = (\mu, \mu^2/2\,\mu^2, \ldots, \mu^n/n, \mu^n, \ldots)$. The lowest non-trivial equations are

$$D_1^+ \tau_h v_j = \tau_{j+1}v_{j-1},$$
$$D_1^- \tau_{j+1}v_j = \tau_j v_{j+1}. \tag{4.22}$$

Set $W_j(x^+, x^-; \mu) = \dfrac{v_j(x^+, x^-; \mu)}{\tau_j(x^+, x^-)}$ and recall $\phi^a = -\dfrac{1}{\varepsilon\beta}\sum_{j=0}^{r}\alpha_j^a \ln \tau_j$ w.r.t. the

basis H^a of h. After a rescaling $\partial_+ \to -i\mu\dfrac{2}{m}\partial_+$, $\partial_- \to -(i\mu)^{-1}\dfrac{2}{m}\partial_-$, Eqs. (4.22) become

$$\partial_- W_j = -\varepsilon\beta\hat{h}_j \cdot \partial_-\phi W_j - i\mu\frac{m}{2}W_{j+1},$$
$$\partial_+ W_j = -(i\mu)^{-1}\frac{m}{2}e^{\varepsilon\beta\alpha_j}\,{}^\phi W_{j-1}, \tag{4.23}$$

which is the explicit form of (2.1). In particular $(W_0, \ldots, W_r)^T$ forms a vector solution of $\mathscr{L}_- W = 0$. In general the asymptotic behaviour of the τ-functions would not be controllable. For AT-reflective data the expression (4.9), (4.11) in terms of the Fredholm determinant ensures that

$$W_j \to e^{\xi(x^-;-i\mu)}, \quad x^- \to \infty.$$

Comparing this with (2.35) one deduces $V_{j,0}^{(+)} = W_j$ (from the uniqueness of the Jost solution with given asymptotics) so that we have gained an expression for the Jost solution $V_{j,0}^{(+)}$ in terms of the τ-functions. Since $\mathscr{L}_-$ is invariant under the simultaneous sign change $x^- \to -x^-$ and $\mu \to -\mu$, also the other Jost solution can be obtained, which results in (4.20). The expression for the functions $w_0, \ldots, w_r$ parametrizing the matrix Jost solutions is then obtained form (2.33). $\square$

Given the fact that the Jost solutions are expressible in terms of the τ-functions one expects that also the monodromy matrix can be reconstructed from the τ-functions. This is of particular interest because the solution of the GLM equations in terms of the affine group orbits shifts emphasis from the construction of solutions to the reconstruction of their (preferred) parameters. In fact, recall that $\Delta_+(\mu)$ defined in Sect. 2.6 admits an analytic continuation to the upper half μ-plane and has at most a finite number of zeros contained in some bounded region $|\mu| \leq R$, $\text{Im}\,\mu > 0$ of the upper half plane.

Proposition 3. *For AT-reflective data*

$$\Delta_+(\mu) = \lim_{x^- \to \infty} \prod_{a \in N(\mu)} \frac{\tau_j(x^+, -x^- - \varepsilon((i\omega^a\mu)^{-1}))}{\tau_j(x^+, -x^-)}, \quad \text{Im}\,\mu > 0, |\mu| > R, \tag{4.24}$$

where the r.h.s. is independent of the flow variables $x_n^{\pm}$ as well as j-independent.

Equation (4.24) is the basis of the approach [19, 20] to use the τ-function formalism for the derivation of trace identities. It allows one to recover the relevant part of the scattering data from a given set of τ-functions. The evaluation for the τ-functions (4.9) is deferred to a separate publication. The proof of Proposition 3 and the contributions to the trace identities from the solitons and breathers can be found in [20].

5. Conclusion

We have given a survey of the results already in the introduction. Perhaps here is the place to list some of the problems that were not addressed. Most importantly, we ignored questions of completeness, i.e. whether or not the solutions constructed from scattering data subject to conditions A, B, C cover some dense subspace of the phase space. For the AT theories (but not for the mKdV) systems we believe this to be the case. Further, the construction has been on the level of the τ-functions and we did not discuss the conditions under which the dynamics of $\ln \tau_j$ is confined to a given branch of the logarithm (cf. the remark after Corollary 1). A mathematical desideratum is to generalize the Grassmannian formulation of Segal and Wilson [28] and the notion of a dressing group from (the determinants of) trace-class operators to (the determinants of) nuclear operators. In particular, this should yield a classification of linear integral equations whose solutions are described in terms of affine group orbits. From a physical viewpoint it should be interesting to repeat the analysis in finite volume and seek contact to the quasi-classical limit of the form factor equations [27].

Appendix

A. The Models

Here we summarize some basic definitions and results for affine Toda theories. In lightcone dynamics these models are closely related to the mKdV systems. Each of the models is specified by the following

Data.
- An affine Lie algebra $\hat{g}$, where the finite dimensional Lie algebra g is simply laced of rank r.

- A graduation $\mathbf{s} = (s_0, \ldots, s_r) \in \mathbb{N}_0^{r+1}$ by means of which $\hat{g}$ can be realized as centrally extended loop algebra

$$\hat{g} \overset{s}{\cong} (g \otimes C[\mu, \mu^{-1}]) \oplus C(K + d_s)\,,$$

where K is the central extension and d_s is the scaling element.

- A field $\Phi : \mathbb{R}^{1,1} \to \hat{h}$ on 2-dim. Minkowski space, which takes values in the Cartan subalgebra $\hat{h}$ of $\hat{g}$,

$$\Phi = \tfrac{1}{2}\phi \cdot H + \eta d_s + \zeta K\,,$$

where $\phi \cdot H = \sum_{a=1}^{r} \phi^a H^a$ and H^a, $a = 1, \ldots, r$ is a basis of the cartan subalgebra h of g, normalized s.t. $(H^a, H^b) = 2\delta_{ab}$. The coefficient field $\eta : \mathbb{R}^{1,1} \to \mathbb{C}$ of the scaling

element will be set identically to zero later, while $\zeta : \mathbb{R}^{1,1} \to$ becomes a functional of $\phi \cdot H$ through the equations of motion.

Lagrangian. The conformally extended affine Toda theory CAT$(\hat{g}, s)$ [3] associated with these data is defined by the Lagrangian

$$\mathscr{L} = \frac{1}{2} \partial_\mu \phi \cdot \partial^\mu \phi + 4 \partial_\mu \eta \partial^\mu \zeta - \frac{m^2}{(\varepsilon\beta)^2} \sum_{j=0}^{r} a_j \, e^{\varepsilon\beta \left(\alpha_j \cdot \phi + \frac{2}{N} s_j \eta \right)} . \qquad (A.1)$$

Here $\beta \in \mathbb{R}$ is the modulus of the coupling constant, m is the mass scale and $\varepsilon = 1$, i for real and imaginary coupling models, respectively. Further $\alpha_1, \ldots, \alpha_r$ are the simple roots of g and $\alpha_0 = -\theta$ is minus the highest root. $N = \sum_{i=0}^{r} a_i s_i$ and $\check{a}_i = a_i$ are the labels of the (dual) Dynkin diagram.

Equations of motion. The equations of motion are

$$\partial^\mu \partial_\mu \phi^a + \frac{m^2}{\varepsilon\beta} \sum_{j=0}^{r} a_j \alpha_j^a \, e^{\varepsilon\beta \left(\alpha_j \cdot \phi + \frac{2}{N} s_j \eta \right)} = 0 \,,$$

$$\partial^\mu \partial_\mu \zeta + \frac{m^2}{2\varepsilon\beta N} \sum_{j=0}^{r} a_j s_j \, e^{\varepsilon\beta \left(\alpha_j \cdot \phi + \frac{2}{N} s_j \eta \right)} = 0 \,, \qquad (A.2)$$

$$\partial^\mu \partial_\mu \eta = 0 \,.$$

Introducing $\varrho(s) = \sum_{i=1}^{r} c_i \lambda_i$, $c_i = \dfrac{a_i s_i}{\check{a}_i N}$, $0 \le i \le r$, the second equation becomes

$$\partial^\mu \partial_\mu (\zeta - \varrho(s) \cdot \phi) = -\frac{m^2}{2\varepsilon\beta}, \text{ so that one is left with the first equation. The equations}$$

of motion of the affine Toda theories AT$(\hat{g}, s)$ are obtained by setting η to zero. These admit the following reformulation. Let $h_i(s) = \check{a}_i + c_i K$, $0 \le i \le r$ denote the Chevelley generators of $\hat{h}$ in the graduation s, where $\check{a}_i = \frac{1}{2} \alpha_i \cdot H$ are the simple coroots ($\alpha_i \cdot \alpha_i = 2$). Let $\Lambda_i(s) \in \hat{h}^*$ denote the basis of fundamental weights dual to them $\langle \Lambda_i(s), h_j(s) \rangle = \delta_{ij}$. Explicitly

$$\Lambda_j(s) = \lambda_j - \check{a}_j \varrho(s) + \check{a}_j \hat{\lambda}_0 \,, \qquad j = 0, \ldots, r \,, \qquad (A.3)$$

where $\lambda_1, \ldots, \lambda_r$ are the fundamental weights of g, $\lambda_0 = 0$ and $\hat{\lambda}_0$ is dual to K, i.e. $\langle K, \hat{\lambda}_0 \rangle = 1$. Define τ-variables $\tau_j : \mathbb{R}^{1,1} \to \mathbb{C}$ by ($\eta \equiv 0$),

$$-\frac{1}{\varepsilon\beta} \ln \tau_j(s) := \langle \Lambda_j(s), \Phi \rangle + \check{a}_j \frac{m^2 x^2}{4\varepsilon\beta}$$

$$= \lambda_j \cdot \phi - \check{a}_j \varrho(s) \cdot \phi + \check{a}_j \left(\zeta + \frac{m^2 x^2}{8\varepsilon\beta} \right) , \qquad (A.4)$$

where one is still free to choose the branch of the logarithm. In particular,

$$\zeta(s) = -\frac{1}{\varepsilon\beta} \ln \left(\prod_{j=0}^{r} \tau_j(s)^{c_j} \right) - \frac{m^2 x^2}{8\varepsilon\beta} . \qquad (A.5)$$

Conversely,

$$\Phi = -\frac{1}{\varepsilon\beta} \sum_{j=0}^{r} h_j(s) \ln(\tau_j(s)\, e^{\frac{1}{8}\, \check{a}_j\, m^2 x^2})$$

$$= -\frac{1}{\varepsilon\beta} \sum_{j=0}^{r} \check{a}_j \ln \tau_j(s) + \zeta(s)\, K\,. \tag{A.6}$$

The equations of motion for $AT(\hat{g}, s)$ are then equivalent to

$$\partial\mu\partial_\mu \ln \tau_j(s) - m^2 \check{a}_j \left(\prod_{k=1}^{r} \tau_k(s)^{-a_{kj}} - 1 \right) = 0\,, \tag{A.7}$$

$$\partial\mu\partial_\mu(\zeta(s) - \varrho(s)\cdot\phi) = -\frac{m^2}{2\varepsilon\beta}\,.$$

The normalizations are choosen s.t. the vacuum solution ($\alpha_j \cdot \phi \equiv 0$ for $\varepsilon = 1$; $\alpha_j \cdot \phi \equiv 2\pi n/\beta$, $n \in \mathbb{Z}$ for $\varepsilon = i$) $\Phi_{\text{vac}} = -\dfrac{m^2}{8\varepsilon\beta}\, x^2 K$ corresponds to $\tau_j(s) = 1$, $j = 0, \ldots, r$. In particular, for $g = A_r$ the equation for the τ-variables is

$$\tau_j\partial_+\partial_-\tau_j - \partial_+\tau_j\partial_-\tau_j = \frac{m^2}{4}\, (\tau_{j+1}\tau_{j-1} - \tau_j^2)\,. \tag{A.8}$$

Linear system. The equations of motion (A.3) can be recovered from the compatibility condition $[\hat{\mathscr{L}}_+, \hat{\mathscr{L}}_-] = 0$ of the following linear system:

$$\hat{\mathscr{L}}_\pm = \partial_\pm - \hat{A}_\pm\,,$$

$$\hat{A}_- = -2\varepsilon\beta\partial_-\Phi + \frac{m}{2}\, \hat{E}\,, \tag{A.9}$$

$$\hat{A}_+ = -\frac{m}{2}\, e^{-2\varepsilon\beta\Phi}\, \hat{F}\, e^{2\varepsilon\beta\Phi}\,,$$

where $\hat{E} = \sum_{i=0}^{r} e_i$, $\hat{F} = \sum_{i=0}^{r} \check{a}_i f_i$ are the standard regular elements of $\hat{g}$, and $[d_s, e_j] = \dfrac{s_j}{N}\, e_j$ and $[d_s, f_j] = -\dfrac{s_j}{N}\, f_j$ has been used.

To a given solution Φ of the field equations one can associate a wave function solving

$$\hat{\mathscr{L}}_\pm\hat{W} = 0 \tag{A.10}$$

and vice versa, where $\hat{W}$ takes values in the affine group. Because of the rapidly decreasing boundary conditions (cf. Sect. 2.2) one has

$$\hat{A}_- \to \frac{m}{2}\, \hat{E} + \frac{m^2}{4}\, x^+ K\,,$$

$$\hat{A}_+ \to -\frac{m}{2}\, \hat{F}\,,$$

as $|x^-| \to \infty$. Let $E(x^+, x^-)$ denote the asymptotic wave function solving

$$\left(\lim_{|x^-|\to\infty} \hat{\mathscr{L}}_\pm \right) E = 0\,.$$

Using $[\hat{E}, \hat{F}] = K$ one checks

$$E(x^+, x^-) = e^{\frac{m^2}{4} x^+ x^- K} e^{\frac{m}{2} x^- \hat{E}} e^{-\frac{m}{2} x^+ \hat{F}} = e^{-\frac{m}{2} x^+ \hat{F}} e^{\frac{m}{2} x^- \hat{E}}.$$

Because of the simple K-dependence, the problem of solving $\hat{\mathscr{L}}_\pm \hat{W} = 0$ can be reduced to an equivalent linear problem $\mathscr{L}_\pm W = 0$ in the loop group. Set

$$A_- = -\varepsilon\beta\partial_-\phi \cdot H + \frac{m}{2}\,\tilde{E}(s)\,,$$

$$A_+ = -\frac{m}{2}\,e^{-\varepsilon\beta\phi\ H}\tilde{F}(s)e^{\varepsilon\beta\phi\ H}\,, \tag{A.11}$$

where $\tilde{E}(s)$, $\tilde{F}(s)$ are the projections of $\hat{E}$, $\hat{F}$ into the loop algebra. Set $\mathscr{L}_\pm = \partial_\pm - A_\pm$ and consider $\mathscr{L}_\pm W = 0$.

Lemma 2. [2] *Any solution W of $\mathscr{L}_\pm W = 0$ can be lifted to a solution $\hat{W}$ of* (A.10).

From now on we specialize to $g = A_r = sl_{r+1}$ and the principle graduation $s = (1, \ldots, 1)$ and set $\mathrm{AT}(r, 1) := \mathrm{AT}(\widehat{sl}_{r+1}, 1)$.

Spectral problem. For a given solution of the field equations the integrability condition $[\mathscr{L}_-, \mathscr{L}_+] = 0$ is satisfied and it suffices to consider the linear equation $\mathscr{L}_- W = 0$. This can be rewritten as a genuine eigenvalue problem, which for the $\mathrm{AT}(r, 1)$ models takes the form

$$\tilde{\mathscr{L}}\tilde{W} = \mu\tilde{W}\,, \qquad \tilde{\mathscr{L}} := i\Lambda^r\partial_- + i\varepsilon D\Lambda^r\,, \qquad \tilde{W} := \Lambda W\,, \tag{A.12}$$

where the parameter of the loop group was redefined by a factor of $-i$, so that

$$\tilde{E}(1) = -i\mu\Lambda\,, \qquad \Lambda_{ab} = \delta_{a,b-1}\,,$$

with the indices taken modulo $r + 1$. Introduce the matrix $U_{ab} = \dfrac{1}{\sqrt{r+1}}\,\omega^{ab}$, satisfying $U^2 = J$, $J_{ab} = \delta_{a,r+1-b}$, $U^4 = J^2 = I$. On the space of matrix-valued functionals $A = A[d_0, \ldots, d_r]$ introduce an inner product by [7],

$$(A, B) = \int\limits_{-\infty}^{\infty} dx^-\,\mathrm{Tr}(A^\dagger J B)\,. \tag{A.13}$$

The conventions in (A.10), (A.11) are then chosen s.t. for the real coupling models ($\varepsilon = 1$) the operator $\tilde{\mathscr{L}}$ is formally selfadjoint w.r.t. to the inner product (A.12) on a suitable subspace of functionals A. To see this introduce the automorphism $\kappa A[d_0, \ldots, d_r] = A[\kappa d_0, \ldots, \kappa d_r]$, where $\kappa d_i = -d_{r+2-i}$, $i = 0, \ldots, r$. It satisfies $(\kappa A, \kappa B) = (A, B)$. From the analysis of the scalar linear problem it follows that one can restrict attention to the subspace of κ-invariant functionals A (cf. the remark below and [18] for a discussion of this fact in the context of W-algebras). From $J\Lambda J^{-1} = \Lambda^{-1}$, $J\Lambda D J\Lambda = -\kappa D$ one checks that $\tilde{\mathscr{L}}^\dagger = J(\kappa\tilde{\mathscr{L}})J$, for $\varepsilon = 1$. This implies

$$(A, \tilde{\mathscr{L}}B) = (\tilde{\mathscr{L}}\kappa A, \kappa B)\,, \tag{A.14}$$

so that $\tilde{\mathscr{L}}$ is formally selfadjoint for $\varepsilon = 1$ on the space of κ-invariant functionals A. In particular this means that for the real coupling models the spectral values μ are real.

Remark. For the scalar operators L_j the restriction to κ-invariant wave functions is possible because also the operators κL_j have the key property of lying in the commutant of r screening operators [18]. Alternatively one could work directly with the symmetrized scalar differential operators $L_j^{\text{sym}} = \frac{1}{2}(L_j + \kappa L_j)$. We preferred here the first viewpoint (restriction to invariant wave functions) because there is no simple matrix analogue of L_j^{sym} (in particular $\frac{1}{2}(\mathscr{L}_- + \kappa\mathscr{L}_-)$ would not do). Notice also that by means of the relation (2.37) the inner product (A.12) induces an inner product on the space of scalar wave functions in terms of the standard inner product $(a, b) = \int dx^- a^* b$. W.r.t. the latter the operators L_j are formally selfadjoint on κ-invariant scalar functions a. By construction every normalizable matrix wave function gives rise to a normalizable scalar wave function, but not necessarily vice versa. This implies that for $\varepsilon = 1$

$$\sigma_p(\tilde{\mathscr{L}}) \subset \sigma_p(L_j) .$$

Affine Toda vs. mKdV hierarchy. A partial differential equation arising from the integrability condition of a linear system $\mathscr{L}_+$, $\mathscr{L}_-$ can be systematically extended to an infinite hierarchy of partial differential equations by studying the kernels of ad $\mathscr{L}_\pm$.[5]

Proposition A. *Any element of* Ker(ad $\mathscr{L}_\pm$) *can be written in the form* $G_\pm S_\pm G_\pm^{-1}$, *where*

$$G_- = \mathbb{1} + \sum_{n=1}^{\infty} g_{-,n}(i\mu\Lambda)^{-n} ,$$

$$G_+ = e^{-\varepsilon\beta\phi \ H} \left(\mathbb{1} + \sum_{n=1}^{\infty} g_{+,n}(-i\mu\Lambda)^n \right) \quad\quad (A.15)$$

and $g_{\pm,n}$ *are diagonal matrices. Further,* $S_- = \sum_{n \le m_-} s_{-,n}(i\mu\Lambda)^n$ *and* $S_+ = \sum_{n \ge m_+} s_{+,n}(-i\mu\Lambda)^n$, *where* $s_{\pm,n}$ *are constant diagonal matrices and* $m_\pm \in \mathbb{Z}$.

Proof (Sketch). For the "$-$" component this is Proposition 1.2, Lemma 1.3 of [6] (see also [23, 24]). For the "$+$" case observe that

$$e^{\varepsilon\beta\phi \ H}\mathscr{L}_+ e^{-\varepsilon\beta\phi \ H} = \partial_+ - \varepsilon\beta\partial_+\phi \cdot H - (i\mu\Lambda)^{-1} ,$$

which is form-identical to $\mathscr{L}_-$ upon replacing ∂_+ with ∂_- and $\mu\Lambda$ with $(\mu\Lambda)^{-1}$ and changing the sign of β. Thus

$$V_- \in \text{Ker(ad } \mathscr{L}_-) \text{ iff } V_+ := e^{\varepsilon\beta\phi \ H}V_- e^{-\varepsilon\beta\phi \ H}\Big|_{\substack{\partial_- \to \partial_+, \beta \to -\beta \\ \mu\Lambda \to (\mu\Lambda)^{-1}}} \in \text{Ker(ad } \mathscr{L}_+). \quad \square$$

Each element of Ker(ad $\mathscr{L}_\pm$) can now serve to define an evolution equation consistent with the original one. The coefficients of the positive/negative powers of $i\mu$ in $S_\pm$ do not enter the construction (cf. below) so that if suffices to consider

$$V_\pm := G_\pm(\pm i\mu\Lambda)^{\mp 1}G_\pm^{-1} , \quad\quad (A.16)$$

---

[5] Depending on the context such kernels are also known as the kernel of the Adler map or the kernel of the Poisson pencil

and powers thereof. Observe that the matrices $V_\pm$ define a resolution of the identity, $V_\pm^{r+1} = (\pm i\mu)^{r+1} \mathbb{1}$. For any Laurent series V in $i\mu$ use the notation $(V)_{\leq k}$ and $(V)_{\geq l}$ for the parts containing powers $\leq k$ and $\geq l$ respectively. Similarly let $(V)_k$ denote the coefficient of $(i\mu)^k$. In this notation the defining property $V_\pm \in \mathrm{Ker}(\mathrm{ad}\,\mathscr{L}_\pm)$ implies that

$$[(V_-^n)_{\geq 0}, \mathscr{L}_-] = [(V_-^n)_{-1}\Lambda^{-1}, \Lambda] = -[(V_-^n)_{\leq -1}, \mathscr{L}_-],$$

are μ-independent and

$$[(V_+^n)_{<0}, \mathscr{L}_+] = [(V_+^n)_0, e^{-\varepsilon\beta\phi\,^H}\Lambda^{-1}e^{-\varepsilon\beta\phi\,^H}] = -[(V_-^n)_{\geq 0}, \mathscr{L}_+],$$

are of order $(i\mu)^{-1}$. This means that for $V_\pm \in \mathrm{Ker}(\mathrm{ad}\,\mathscr{L}_\pm)$ one has a consistent set of evolution equations

$$\frac{\partial}{\partial x_n^\pm}\mathscr{L}_\pm = [A_\pm^n, \mathscr{L}_\pm], \tag{A.17}$$

where $A_-^n := (V_-^n)_{\geq 0}$, $A_+^n := (V_+^n)_{<0}$. In particular one shows [6]

$$A_\pm^1 = A_\pm, \tag{A.18}$$

so that one can identify $x_1^\pm$ with $x^\pm$ and the $n = 1$ equations become trivial.

So far the "+" and the "−" sectors have been treated as independent. But because of $[\mathscr{L}_-, \mathscr{L}_+] = 0$ the kernels $\mathrm{ad}\,\mathscr{L}_+$ and $\mathrm{ad}\,\mathscr{L}_-$ intersect. To see this, note that $[\mathscr{L}_-, \mathscr{L}_+] = 0$ implies $[\partial_- - A_-^G, G_-\mathscr{L}_+ G_-^{-1}] = 0$, so that $G_-\mathscr{L}_+ G_-^{-1}$ is of the form $G_-\mathscr{L}_+ G_-^{-1} = \partial_+ + C$, where C takes values in $\mathrm{Ker}(\mathrm{ad}\,\Lambda)$. Thus,

$$[\mathscr{L}_+, V_-^n] = 0, \quad n \geq 1. \tag{A.19}$$

Similarly it follows $[\mathscr{L}_-, V_+^n] = 0$, $n \geq 1$. Correspondingly there is a consistent set of "crossed" evolution equations

$$\frac{\partial}{\partial x_n^\pm}\mathscr{L}_\mp = [A_\pm^n, \mathscr{L}_\mp]. \tag{A.20}$$

For $n = 1$ one recovers $[\mathscr{L}_+, \mathscr{L}_-] = 0$, i.e. the original affine Toda equation.

Remark. i. Notice that the elements of $\mathrm{Ker}(\mathrm{ad}\,\mathscr{L}_\pm)$ are uniquely determined by their constant part. In particular $V_\pm^n = (\pm i\mu\Lambda)^{\mp n} + \ldots$, which also implies for rapidly decreasing boundary conditions

$$\lim_{|x^-|\to\infty} A_\pm^n = (\pm i\mu\Lambda)^{\mp n}. \tag{A.21}$$

ii. Since $\mathscr{L}_- = \partial_- + \varepsilon D - i\mu\Lambda$ coincides with the differential operator defining the mKdV hierarchy in the DS scheme, the "−" flow equations (A.17) coincide with that of the mKdV hierarchy in the variables $d_0, \ldots, d_r$. In particular, the flow variables x_-^2 and x_-^1 correspond to the usual time and space coordinate in the mKdV equation. Since $\mathscr{L}_+$ can be made form-identical to $\mathscr{L}_-$ by a gauge transformation (and $\mu\Lambda \to (\mu\Lambda)^{-1}$, $\beta \to -\beta$) the "+" flow equations in (A.17) essentially also define a mKdV-hierarchy in disguised form. The higher order generalization of the affine Toda equations arise from the crossed flow equations (A.20).

---

[6] This holds only in the principle graduation

Abelization Maps. As a corollary of the previous construction one obtains the Lax representation and the zero curvature condition for the hierarchies. To simplify the notation set $t_n = x_n^-$, $t_{-n} = x_n^+$, $n \in E$ and similarly $A^n = A_-^n$, $A^{-n} = A_+^n$, $n \in E$ and $V^n = V_-^n$, $V^{-n} = V_+^n$, $n \in E$. Then

$$\frac{\partial}{\partial t_m} V^n - [A^n, V^n] = 0, \tag{A.22}$$

$$\frac{\partial}{\partial t_m} A^n - \frac{\partial}{\partial t_n} A^m - [A^n, A^m] = 0. \tag{A.23}$$

The second equation is just the integrability condition for the Eqs. (A.17) and (A.20). The first equation can be obtained by differentiating $[\mathscr{L}_\pm, V^n] = 0$ w.r.t. $\frac{\partial}{\partial t_m}$, which results in

$$\left[\frac{\partial}{\partial T_m} V^n - [A^m, V^n], \mathscr{L}_\pm \right] = 0.$$

The left entry of the commutator is again an element of $\mathrm{Ker}(\mathrm{ad}\,\mathscr{L}_\pm)$. Since elements of $\mathrm{Ker}(\mathrm{ad}\,\mathscr{L}_\pm)$ are uniquely determined by there constant parts (in t_n) and the latter vanishes for the quantity in question one arrives at the Lax representation (A.22).

Inserting now $V_\pm = G_\pm^{-1}(\pm i\mu\Lambda)^{\mp 1} G_\pm$ into the Lax representation one finds that

$$G_\pm^{-1} \left(\frac{\partial}{\partial t_m} - A_\pm^m \right) G_\pm = \frac{\partial}{\partial t_m} - A^{G,m}, \tag{A.24}$$

for some $\mathrm{Ker}(\mathrm{ad}\,\Lambda)$-valued $A^{G,m}$. Since $\mathrm{Ker}(\mathrm{ad}\,\Lambda)$ is abelian this means that $\mathrm{Ad}\,G_\pm$ (considered as a function of all the flow variables) serves as an abelization map for all of the higher order linear operators $\mathscr{L}^n = \frac{\partial}{\partial t_n} - A^m$. The case $m = 1$ leads back to the proof of Proposition A. Since the transformed connections $A^{G,m}$ are abelian the zero curvature conditions (A.23) imply

$$\frac{\partial}{\partial t_m} A^{G,n}(\mu) - \frac{\partial}{\partial t_n} A^{G,m}(\mu) = 0, \tag{A.25}$$

so that $A^{G,n}(\mu)$ are generating functions for the conserved charges. In particular for $(m, n) = (2, 1)$ one obtains the conservation equation for the lowest KdV flow while $(m, n) = (-1, 1)$ gives the conservation equation $\partial_+ A^{G,1} = \partial_- A^{G,-1}$ for the lowest affine Toda flow. In particular, we note explicitly the

Corollary A. *The functional $A^{G,1}(\mu) = A^{G,1}[d_0, \ldots, d_r](\mu)$ is the generating functional for the homogeneous, polynomial conserved currents of both, the $\partial_{x_2^-}$-KdV flow and the $\partial_{x_1^+}$-affine Toda flow in lightcone dynamics*

Here, homogeneous and polynomial means that the expansion coefficients in powers of $i\mu\Lambda$ are differential polynomials in $d_0, \ldots, d_r$ and are homogeneous of degree $n + 1$ w.r.t. the grading $\deg d_j = \deg \partial_- = 1$.

Equations (A.25) can also be used to relate the abelization maps $\mathrm{Ad}\,G_\pm$ to wave functions $W_\pm$ solving $\mathscr{L}^n W_\pm = 0$, $n \in \pm E$, where $\mathscr{L}^n = \frac{\partial}{\partial t_n} - A^n$. To find the relation, rewrite this as

$$W_\pm^{-1} \mathscr{L}^n W_\pm = \frac{\partial}{\partial t_n}, \tag{A.26}$$

which means that $\operatorname{Ad} W_\pm$ can be viewed as an abelization map for which the transformed connection vanishes identically.[7] Given $\operatorname{Ad} G_\pm$ satisfying (A.24) and Eqs. (A.25) for the transformed connections an abelization map $\operatorname{Ad} W_\pm$ implementing (A.26) can be obtained as follows. Define

$$\omega(\mu) := \sum_{n \in \pm E} A^{G,n} dt_n , \qquad (A.27)$$

which by (A.25) is a closed 1-form on some contractable region of the phase space (where t_n, $n \in E$ are good local coordinates). Hence it is exact and has a potential $\hat{\gamma} = \hat{\gamma}(t_n, n \in \pm E; \mu)$ satisfying $\partial_{t_n} \hat{\gamma} = A^{G,n}$. Explicitly,

$$\hat{\gamma}(t; i\mu \Lambda) = \int_0^1 ds \left(\sum_{n \in \pm E} A^{G,n}(st; \mu) t_n \right) + \hat{\gamma}_0 , \qquad (A.28)$$

where $st = (st_{\pm 1}, st_{\pm 2}, \ldots)$ and the integration constant $\hat{\gamma}_0$ is chosen s.t. $\lim_{x^- \to \infty} \hat{\gamma} = 0$. Identifying

$$W_\pm = G_\pm e^{\hat{\gamma}} , \qquad (A.29)$$

one checks

$$\mathscr{L}^n = G_\pm \left(\frac{\partial}{\partial t_n} - A^{G,n} \right) G_\pm^{-1} = (G_\pm e^{\hat{\gamma}}) \frac{\partial}{\partial t_n} (G_\pm e^{\hat{\gamma}})^{-1} .$$

Asymptotics. The asymptotics of the function $\hat{\gamma}$ is related to that of the τ-functions. Consider the "$-$" part of the relation (A.29) and return to the notation $U^{-1} W_-^{(\pm)} U =: W^{(U,\pm)}$ for the Jost solutions of $\mathscr{L}$. Since the matrices $U^{-1} g_{-,n} U$ are differential polynomials in $d_0, \ldots, d_r$, they vanish for $|x^-| \to \infty$. Thus,

$$\lim_{x^- \to -\infty} e^{-\xi(x^+, x^-; i\mu\Omega)} W^{(U,+)}$$
$$= \lim_{x^- \to -\infty} e^{-\xi(x^+, x^-; i\mu\Omega) + \hat{\gamma}(x^+, x^-; i\mu\Omega)} . \qquad (A.30)$$

Observe from (A.21) that as $x^- \to -\infty$, $\hat{\gamma}$ contains a field independent part $\hat{\gamma} = \xi(x, i\mu\Omega) + \gamma$, which cancels against the factor ξ on the r.h.s. of (A.30). In terms of components (A.30) reads

$$\lim_{x^- \to -\infty} e^{-\xi(x^+, x^-; i\mu\Omega)} w_{a-b}^{(+)}(\omega^a \mu)$$
$$= \begin{cases} 0, & a \neq b, \\ \lim_{x^- \to -\infty} e^{\gamma(x^+, x^-; i\mu\omega^a)}, & a = b. \end{cases} \qquad (A.31)$$

Inserting the expression (4.20) for the Jost solutions this becomes

$$\lim_{x^- \to -\infty} e^{\gamma(x^+, x^-; i\mu)} = \lim_{x^- \to -\infty} \frac{\tau_j(x^+, x^- - \varepsilon((-i\mu)^{-1}))}{\tau_j(x^+, x^-)} . \qquad (A.32)$$

---

[7] The advantage of using an abelization map of the form (A 24) with nontrivial image in $\operatorname{Ker}(\operatorname{ad} \Lambda)$ is that the recursion relations encountered in the proof of proposition A can be solved locally, i.e. no integrations are necessary

From here one can deduce the relations $I_\pm^{(n)} = \mp \dfrac{r+1}{n} (\varepsilon\beta)^{-(n+1)} \partial_{x_n^\pm} \ln \tau_j \Big|_{x^- = -\infty}^\infty$ for the conserved charges quoted in the introduction. See also [20].

B. Fredholm Theory on Banach Spaces

The decoupled GLM equations (3.19) constitute a family of Fredholm equations indexed by $x \in \mathbb{R}$ on a non-compact interval. These equations are not directly amenable to the classical Fredholm theory. In particular, Eqs. (3.19) do *not* coincide with the equations for the resolvent kernel of the Volterra integral operators $\hat{F}_j^{(\pm)}$. Nevertheless, up to some technical modifications the major results of the classical Fredholm theory carry over. One way to show this would be to go back to the original discretization procedure and to check whether for suitable continuous kernels $F_j^{(\pm)}$ the required bounds can be established. For several reasons it is however more convenient to embed the set-up into the broader context of Grothendieck's generalized Fredholm theory on Banach spaces [11]. For orientation we remark that the classical Fredholm theory is recovered for the Banach space $\mathscr{C}(I)$ of continuous functions on some compact interval $I = [a, b] \subset \mathbb{R}$ with the usual norm $\|f\| = \sup\limits_{x \in I} |f(x)|$. A continuous kernel $A \in \mathscr{C}(I \times I)$ then defines a nuclear operator and its trace coincides with the "naive" trace $\int\limits_a^b dx\, A(x, x)$.

Fredholm theory on Banach spaces. Let E be a Banach space satisfying the (Banach) approximation property. We refrain from giving the definition because all the standard Banach spaces (continuous functions, Hilbert spaces, L^p-spaces, Sobolev spaces) satisfy the approximation property. Let E' be the dual of E and use $\langle v, w \rangle$, $v \in E'$, $w \in E$ to denote the duality pairing. The relevant class of operators (replacing the trace-class operators on Hilbert spaces) are the *nuclear operators* in E. A nuclear operator A in E is one which can (non-uniquely) be represented as a series $A = \sum\limits_n a_n$, where each a_n is of rank 1 and $\sum\limits_n \|a_n\|$ is finite (so that the series A converges in operator norm). Explicitly one may write $a_n = \lambda_n |v_n\rangle \langle v'_n|$, for $v_c \in E$, $v'_n \in E'$ s.t. $\|v_n\| = \|v'_n\| = 1$ and $\lambda_n \in \mathbb{C}$. The *nuclear norm* $\|a\|_1$ is the infimum of the set of numbers $\sum\limits_n \|a_n\| = \sum\limits_n |\lambda_n|$ for all decompositions $A = \sum\limits_n a_n$. The space $L^1(E)$ of nuclear operators then is a Banach space for the nuclear norm. The crucial properties of nuclear operators are

N1. Let E be a Banach space satisfying the approximation property. There exists a linear form Tr on $L^1(E)$ s.t. $|\operatorname{Tr}(A)| \leq \|A\|_1$ and taking $|v\rangle\langle w| (c \in E, w \in E')$ to $\langle v, w \rangle$.

N2. For any $A \in L^1(E)$ and $\varepsilon > 0$ there exists a Hilbert space H and bounded operators $\alpha : H \to E$, $\alpha' : E \to H$ s.t. $A = \alpha\alpha'$ and $\alpha'\alpha$ is a Hilbert Schmidt operator in H with $\|\alpha'\alpha\|_2 \leq \|A\|_1 + \varepsilon$ (where $\|\cdot\|_2$ is the Hilbert Schmidt norm on $L^2(H)$.)

The modified determinant of $A \in L^1(E)$ can then be defined by $\det_2(\mathbb{1} + A) := \det_2(\mathbb{1} + \alpha'\alpha)$, where the r.h.s. is the determinant for Hilbert Schmidt operators (see e.g. [26, 10]). Finally, define the ("Grothendieck") determinant of $A \in L^1(E)$ by

$$\det(\mathbb{1} + A) := e^{\operatorname{Tr}(A)} \det_2(\mathbb{1} + A). \tag{B.1}$$

Using the analogous properties for the determinant on $L^2(H)$ one deduces

Theorem B1. *For $A \in L^1(E)$, a) $|\det(\mathbb{1}+A)| \leq \exp(\|A\|_1)$. The map $z \to \det(\mathbb{1}+zA)$ is an entire analytic function.*
b) $\det(\mathbb{1} + zA) = e^{z\,\mathrm{Tr}(A)} \prod_{n\geq 1} (1 - \lambda_n z) e^{\lambda_n z}$, *where* $\|A\|_1 = \sum_n |\lambda_n|$.
c) $\det(\mathbb{1} + A)\det(\mathbb{1} + B) = \det(\mathbb{1} + A + B + AB)$.
d) *If $-\lambda^{-1}$ is an eigenvalue of A then $\det(\mathbb{1}+zA)$ has a zero of order n at $z = -\lambda^{-1}$, where n is the algebraic multiplicity of $-\lambda^{-1}$.*
e) $\det(I + zA) = \sum_{n\geq 0} z^n d_n(A)$, *where for a decomposition $A = \sum_n |v_n\rangle \langle v_n'|$ one has*
$d_n(A) = \sum_I \det(\langle v_i, v_j'\rangle)_{i,j\in I}$, *with $I = \{(i_1, \ldots, i_n) \in \mathbb{N}^n \mid i_1 < \ldots < i_n\}$.*
f) $\det(\mathbb{1} + zA) = \exp\left(\sum_{n\geq 1} \frac{1}{n} \mathrm{Tr}(A^n) (-z)^n\right)$, *for $|z| < 1/\|A\|_1$.*

In particular, it follows that for $A \in L^1(E)$ the resolvent set is given by $\varrho(A) = \{-z^{-1} \in \mathbb{C} \mid \det(\mathbb{1}+zA) \neq 0\}$. To describe the resolvent, define the modified first Fredholm minor by $D_z^2(A) := D_z^2(\alpha'\alpha)$, where the r.h.s. denotes the Fredholm minor for Hilbert Schmidt operators, i.e. $D_z^2(A) := \det_2(\mathbb{1} + zA)(\mathbb{1} + zA)^{-1}A$. Then define

$$D_z(A) := e^{z\,\mathrm{Tr}(A)} D_z^2(A) \qquad\qquad (\text{B.2})$$

to be the Fredholm minor for $A \in L^1(E)$. Using the analogous properties for the Fredholm minor on $L^2(H)$ (e.g. [26, 10]) one deduces

Theorem B2. a) $\|D_z(A)\|_1 \leq C_1 \exp(z\gamma_1\|A\|_1)$ *for constants C_1, γ_1.*
b) *For $-z^{-1} \in \varrho(A)$ the resolvent is given by*

$$(\mathbb{1} + zA)^{-1} = \mathbb{1} - z \frac{D_z(A)}{\det(\mathbb{1} + zA)}.$$

c) $D_z(A) = \sum_{n\geq 0} z^n D_n(A)$, *where $D_n(A)$ is given by a Plemelj-Smithies formula (in terms of the nuclear trace).*

Fredholm theory for GLM equations Set $I_x = [x, \infty[$ and consider the Banach space of continuous functions $\mathscr{C}(I_x) = \{f : I_x \to \mathbb{C} \text{ continuous} \mid \|f\| := \|f\|_{x,\nu} < \infty\}$ with norm $\|f\|_{x,\nu} = \sup_{s\in I_x} |f(s)| (s-x+1)^\nu$, $\frac{1}{2} < \nu \leq 1$. Similarly define $\mathscr{C}(I_x^2)$ with norm $\|A\| = \sup_{(s,t)\in I_x^2} |A(s,t)| (s - x + 1)^\nu (t - x + 1)^\nu$. Then, if $A \in \mathscr{C}(I_x^2)$ is a continuous kernel, the function $\hat{A}f$ defined by $\hat{A}f = \int_{I_x} dy\, A(x,y)\, f(y)$ belongs to $\mathscr{C}(I_x)$ if f does. The linear operator $\hat{A}$ is bounded with norm $\|\hat{A}\| \leq \frac{1}{2\nu}\|A\|$. Since $A \in \mathscr{C}(I_x^2)$ is also square integrable $\hat{A}$ can also be considered as the restriction $\hat{A} = \hat{A}_{\mathrm{HS}}|_{\mathscr{C}(I_x)}$ of a Hilbert Schmidt operator acting on $L^2(I_x)$. One can prove the

Lemma. *The operator $\hat{A}$ acts on $\mathscr{C}(I_x)$ as a nuclear operator. The Hilbert Schmidt operator $\hat{A}_{\mathrm{HS}}$ approximates $\hat{A}$ in the sense of N2. Moreover, the trace of $\hat{A}$ as a nuclear operator coincides with the "naive" trace $\mathrm{Tr}\,\hat{A} = \int_{I_x} ds\, A(s,s)$.*

Remark. We emphasize that $\hat{A}$ does *not* act on $L^2(I_x)$ as a trace-class operator and the sum $\sum_n \lambda_n$ does not converge, generally speaking. A criterion for $\hat{A}$ to be trace-class is that the kernel is hermitian positive, i.e. satisfies $\int_{I_x^2} ds\,dt\,A(s,t)f(s)f(t)^* \geq 0$ for all $f \in \mathscr{C}(I_x)$ (e.g. [10], p. 114). For the class of kernels relevant to Eqs. (3.19), however, this fails. For the same reason the Grassmannian formulation of Segal and Wilson [28] (where only determinants of trace-class operators are considered) does *not* apply here.

From (B.1), (B.2) it follows that

$$
\det(\mathbb{1} + z\hat{A}) = \exp\left(z \int_{I_x} ds\, A(s,s) \right) \det_2(\mathbb{1} + z\hat{A}_{\mathrm{HS}}) \,,
$$

$$
D_z(A) = \exp\left(z \int_{I_x} ds\, A(s,s) \right) D_z^2(A) \,.
$$

(B.3)

In particular one obtains from (B.3) expressions for the expansion coefficients $d_n(A)$ and $D_n(A)$. These turn out to be of the same type as in the classical Fredholm theory, because the $e^{z\,\mathrm{Tr}(\hat{A})}$ factor essentially re-inserts the diagonal terms into $A(s_i, s_j)_{1 \leq i,j \leq n}$, which are absent in $d_n^2(A)$. Thus, $d_0(A\,|\,x) = 1$, $D_0(A\,|\,x,y) = A(x,y)$ and

$$
d_n(A\,|\,x) = \int_{I_x^n} ds_1 \ldots ds_n\, A \begin{bmatrix} s_1 \ldots s_n \\ s_1 \ldots s_n \end{bmatrix}, \qquad n \geq 1 \,,
$$

$$
D_n(A\,|\,x,y) = \int_{I_x^n} ds_1 \ldots ds_n\, A \begin{bmatrix} x & s_1 \ldots s_n \\ y & s_1 \ldots s_n \end{bmatrix}, \qquad n \geq 1 \,,
$$

(B.4)

where

$$
A \begin{bmatrix} s_1 \ldots s_n \\ t_1 \ldots t_n \end{bmatrix} = \det(A(s_i, t_j)_{1 \leq i,j \leq n}) \,.
$$

References

1. Ablowitz, M.J., Kaup, D.J., Newell, A.C., Segur, H.: Method for solving the sine-Gordon equation. Phys. Rev. Lett. **30**, 1262–1264 (1973)
2. Babelon, O., Bernard, D.: Affine Solitons. A relation between τ-function, dressing and Bäcklund transformations. Int. J. Mod. Phys. A **8**, 507–544 (1993)
3. Babelon, O., Bonora, L.: Conformal affine Toda field theory. Phys. Lett. B **244**, 220 (1990)
4. Date, E., Jimbo, M., Kashiwara, M, Miwa, T.: In: Nonlinear integrable systems. RIMS Symposium, Jimbo, M. et al. (eds.) Singapore: World Scientific 1983
5. DeGroot, M., Hollowood, T.J, Miramontes, J.L.: Generalized Drinfel'd-Sokolov hierarchies Commun. Math. Phys. **145**, 57–84 (1992)

6 Drinfel'd, V.G., Sokolov, V.V.: Lie algebras and equations of KdV type J. Sov. Math. **30**, 1975 (1984)

7. Evans, J.M : Complex Toda theories and twisted reality conditions. Nucl. Phys B **390**, 225–250 (1993)

8 Faddeev, L.D., Takhtajan, L.A.: Hamiltonian methods in the theory of solitons. Berlin, Heidelberg, New York: Springer 1987

9. Gardner, C.S., Greene, J.M., Kruskal, M.D., Miura, R.M.: Method for Solving the Korteweg-de Vries equation. Phys. Rev. Lett. **19**, 1095–1097 (1967)

10. Gohberg, I.C., Krein, M.G : Introduction into the theory of linear non-selfadjoint operators. Providence, RI: AMS 1969

11. Grothendieck, A : La théorie de Fredholm. Bull Soc. Math. France **84**, 319–384 (1956)

12. Hirota, R.: Direct Methods in soliton theory. In: Soliton Bullough, R.K., Chaudry, P S. (eds.) 1980

13. Hollowood, T., Miramontes, J.L.: τ-functions and generalized integrable hierarchies Commun. Math. Phys **157**, 99–117 (1993)

14. Jimbo, M., Miwa, T : Solitons and infinite dimensional Lie algebras. Publ. RIMS, Kyoto Univ. **19**, 943 (1983)

15 Kac, V : Infinite dimensional Lie algebras Third edition Cambridge: Cambridge University Press 1990

16 Kaup, D.J., Newell, A.C : The Coursat and Cauchy problems for the sine-Gordon equation. SIAM J. Appl. Math., 37–54 (1978)

17. Manakov, S.V., Zakharov, V.E.: Construction of higher-dimensional nonlinear integrable systems and of their solutions. Funct. Anal Appl. **19**, 89–101 (1985)

18. Niedermaier, M.R.: Irrational free field resolutions for $W(sl(n))$ and extended Sugawara construction. Commun. Math. Phys. **148**, 249–281 (1992)

19. Niedermaier, M R.: The spectrum of the conserved charges in affine Toda theories. Preprint DESY-92-105

20. Niedermaier, M.R.: Trace identities from τ-functions Preprint MPI-Ph/93-16

21. Novikov, S P., Manakov, S V., Pitaievski, L.P., Zakharov, V.E.: Theory of solitons. The Inverse Method. New York: Plenum 1984

22. Oishi, S.: Relationship between Hirota's method and the inverse spectral method – the KdV case. J. Phys. Soc. Jpn. **47**, 1037–1038 (1979)

23. Olive, D , Turok, N.: Local conserved densities and zero curvature condition for Toda lattice field theories. Nucl. Phys. B **257**, 277 (1985)

24. Olive, D., Turok, N.: The Toda lattice field hierarchies and zero curvature conditions for Kac-Moody algebras. Nucl Phys. B **265**, 469 (1986)

25. Pöppe, C., Sattinger, D.H.: Fredholm determinants and the τ-function for the Kadomtsev-Petviashvili hierarchy Publ. RIMS **24**, 505–538 (1988)

26. Reed, M., Simon, B.: Methods of modern mathematical physics Vol. 4. New York: Academic Press 1978

27 Smirnov, F.A.: Form factors, deformed KZ equations and finite gap integration. Preprint Newton Institute (Sept. 1992)

28. Segal, G., Wilson, G.: Publ. I.H.E.S. **61**, 1 (1985)

29 Ueno, K., Takasaki, K : Toda lattice hierarchy. In: Group representations and systems of differential equations. Adv. Sud. Pure Math. Vol 4 Tokyo: Kinokuniya 1984

30 Zakharov, V.E., Shabat, A.B.: A scheme for integrating the nonlinear equations of mathematical physics by the method of the inverse scattering problem Funct. Anal. Appl. **8**, 226–235 (1974)

31. Beals, R , Deift, P., Tomei, C.: Direct and inverse scattering on the line. Math. Monographs **28**, AMS 1988

Communicated by R. H Dijkgraaf