Integrable Systems and Operator Equations

Habilitationsschrift

vorgelegt am

der Fakultät für Mathematik und Informatik
der Friedrich-Schiller-Universität Jena

von
Dr. rer. nat. Cornelia Schiebold
aus Frankfurt am Main

Gutachter

1.
2.
3.

Erteilung der Lehrbefähigung am

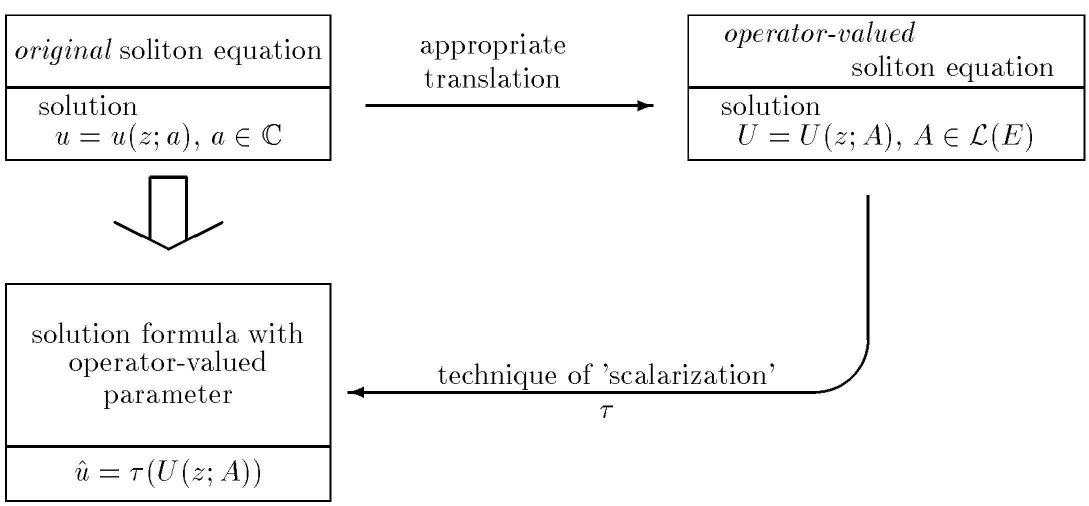
To Egmont

Introduction

The topic of the present monograph is to study integrable systems from an operator theoretic point of view. Except for a digression on the KP equations, it is devoted to a uniform treatment of the AKNS system. Its main intention is to show that not only the construction of explicit solution formulas, but also a great deal of the structural analysis of solution classes can be pursued on this general level.

In the beginning soliton equations arose as distinguished nonlinear evolution equations with striking similarities and irritating differences. A great step towards understanding the connection between the equations was achieved in the landmark paper of M. J. Ablowitz, D. J. Kaup, A. C. Newell and H. Segur [3]. Building on important work of V. E. Zakharov and A. B. Shabat [101], they discovered that the most prominent integrable systems in one space variable can be derived via reduction from a more general integro-differential system in two unknown functions. This was a crucial progress both in the effort to unify the inverse scattering method and to explain the different characteristic properties of the equations. For example the Nonlinear Schrödinger equation, displaying the complex nature of quantum mechanics, results from another type of reduction than the 'essentially real' sine-Gordon equation. Moreover it is remarkable that the AKNS system is in general non-integrable in the sense that there are solutions with instantaneous singularities, and the inverse scattering method applies to a certain extent only formally. One needs appropriate reductions (we call them $\mathbb{C}$- and $\mathbb{R}$-reduction) to arrive at soliton equations.

In the present work we will approach the AKNS system by an operator method. Our guiding principle can be described as follows: One translates a soliton equation and a special solution simultaneously to an operator equation and a corresponding operator-valued solution. Then one tries to regain solutions of the original equation by the use of an appropriate functional.



Typically, the solution we start with depends on a scalar parameter $a \in \mathbb{C}$. The goal of our strategy is to construct solution formulas depending on an operator-valued parameter $A \in \mathcal{L}(E)$, E some Banach space, which can be viewed as a blow-up of a. As indicated in the diagram above, this can be achieved via a detour through the operator-level.

The original idea of this strategy is due to V. A. Marchenko, who pursued it in his pioneering work [55] for differential algebras. In his applications these are always realized by operators on Hilbert spaces. Then B. Carl had the idea to place it into the frame of Banach operator ideals (in the sense of A. Pietsch [72]) with the intention to establish a link between soliton theory and the geometry of Banach spaces. In a joint paper with H. Aden [8] this was done for the Korteweg-de Vries equation, and the role of the trace as an appropriate choice for the scalarization functional was clarified. In the sequel, the method proved to be very flexible. As shown by Carl and the author, it works for the most prominent soliton equations, even for discrete ones as the Toda lattice [18], [19], [88], [89], [92]. But it turned out that there is no universal algorithm to produce the right translation to the operator level. In another way, the efficiency of the method was confirmed by the work of Aden and H. Blohm [7], [13], [14], who showed that all solutions covered by the inverse scattering method can be realized in this frame. Here semigroup techniques play an important role [17].

Among the various other operator-theoretic approaches to soliton theory we can only mention some which are closely related to our own work. In the work of C. Pöppe [75], [76] determinants of Fredholm integral operators are used for the construction of solutions. His joint article [10] with W. Bauhardt was one of the starting points of our work on the AKNS system. In [83], [84], A. Sakhnovich uses the method of operator identities to face non-commutativity in the study of matrix differential equations. We refer also to [40], [41], [42], [85] for various aspects of reduction. A combination of soliton theory, infinite-dimensional analysis, and Hamiltonian techniques [34], [36] was invented by B. Fuchssteiner. It is remarkable that A. R. Chowdhury and Fuchssteiner [35] obtain the operator version of the KdV from a completely different point of view.

Let us now explain the main aspects of the present work.

Throughout the whole text our intention was to treat the AKNS system in a uniform way. As a rule this means that the constructions on the operator-level are carried out for the general AKNS system, and everything which concerns applications to soliton equations is done for its $\mathbb{C}$-reduction (and, with ameliorations, for the subordinated $\mathbb{R}$-reduction). In particular, we can cover a large part of soliton theory without ad hoc choices once the general techniques are established.

Our operator-theoretic treatment of the AKNS system starts from ideas of Bauhardt and Pöppe. In [10] they wrote down the right operator version of the AKNS system. But for the applications we have in mind we need a substantial generalization of their operator solution. More precisely we construct in Theorem 1.2.1 an operator solution depending on *two* parameters $A \in \mathcal{L}(E), B \in \mathcal{L}(F)$, where E, F are possibly different Banach spaces, in contrast to [10] where only the case $A = B$ is considered. The basis of all later applications are the determinant formulas in Theorem 2.4.4.

We focus on two major applications. The first is a detailed study of negatons (or multiple pole solutions) for the complete $\mathbb{C}$-reduced AKNS system. From the earlier literature on, discussions of these solutions had been appearing occasionally [34], [71], [96], [98], [99], mostly in very particular cases. But the question of a complete and rigorous asymptotic characterization was, to the best of our knowledge, first asked by Matveev [59] in the related context of positons (see also [79] for corresponding material on negatons). For several individual equations, a complete answer was obtained by the author in [88], [90], [91]. We point out that the topic is particularly attractive in the context of the AKNS system, because its negatons are always regular as proved in Proposition 4.3.7. In fact they are accessible by the inverse scattering method.

The main result is formulated in Theorem 5.1.2. In the proof we will first observe that negatons consist of groups of solitons which are weakly bounded. These groups interact in an analogous way as the particles of N-solitons. Then we will split them up and analyse the interior structure of each group separately.

For the $\mathbb{R}$-reduction, the breather is a further distinguished solitary solution. It can also be considered as the simplest example of so-called formations [54], [93]. In our result for the $\mathbb{R}$-reduction in Theorem 5.2.2 we also describe negatons consisting of breathers. It is interesting that the collision of breathers affects not only the trajectories but also the oscillations.

Our second main application is the construction of countable superpositions, a topic which was initiated by F. Gesztesy, W. Karwowski, and Z. Zhao [39] and intensely studied, for solitons, by Gezstesy, W. Renger and collaborators [39], [40], [81]. For related results see also [8], [18], [88], [89]. In the present work we establish for the first time countable superpositions of negatons.

The integral terms of the AKNS system force us to include assumptions on the width of the superposed waves. However, for individual equations, for which the integral terms cancel, this restriction can be completely dropped. Exemplarily we will give sharpened results in Theorem 7.5.5 and Theorem 7.6.3 for the Nonlinear Schrödinger and the modified Korteweg-de Vries equations. In the proof the amelioration relies on sophisticated Banach space techniques.

As a digression to equations in two space variables, we will provide the operator-theoretic basis for the treatment of the KP-I and KP-II equations. The main result is the construction of an operator-valued solution in Theorem 3.2.1. Already in [18], we obtained a solution formula with two commuting operator parameters in joint work with B. Carl. The novelty in Theorem 3.2.1 is that we allow operator-parameters operating on different spaces. This increases the complexity of the situation enormously, and for some time we even had doubts about the validity of the result. But then new evidence was given in discussions with A. Sakhnovich. He had discovered a different solution of the matrix-KP with non-commuting parameters and produced a computer-supported proof of the solution property. Strongly encouraged by this, we finally succeeded in proving Theorem 3.2.1. Our argument cuts down the calculations by a consequent use of recursion relations. Later on we found an alternative, more elegant proof relying on the properties of the Miura transformation. Finally we carry out the familiar scalarization procedure. A generalization of the process also leads to explicit solution formulas for the matrix KP.

In the remainder we outline the organization of the text and survey its most important results. In Chapters 1–3 we lay the operator-theoretic ground of our work and derive explicit solution formulas for the general AKNS system and the KP equations. Chapters 4–7 are devoted to applications. Here we discuss in particular negatons and countable superpositions. For a more details, references, and further informations we refer to the introductions preceding the respective chapters.

Chapter 1 Here we treat the general AKNS system on the operator level. We formulate an operator version and give in Theorem 1.2.1 a solution depending on two completely independent operator-valued parameters. The main part of the chapter is then concerned with the proof of the solution property. The development of the general theory is continuously accompanied by a discussion of the prototypical equations (the Nonlinear Schrödinger, the sine-Gordon, and the modified Korteweg-de Vries equations).

Chapter 2 Next we explain how to extract scalar solution formulas from our operator soliton. The basic idea is to apply a functional with convenient multiplicativity properties. It will turn out that this can only be done for operator solutions with specified range (see Theorem 2.2.1), and we subsequently discuss a systematic way to arrange this. Next we use the theory of traces and determinants on Banach ideals to derive determinant formulas which are crucial for our applications. Our main result, after a further useful amelioration, is recorded in Theorem 2.4.4.

Chapter 3 In this excursus we turn to the KP equation. In Theorem 3.2.1 we construct a solution of the operator-valued KP. The hard part of the proof is to treat non-commuting parameters. Then we extend the scalarization process such that it gives access not only to the original scalar KP, but also to the more complicated matrix KP (see Theorem 3.3.4. and Theorem 3.4.9). Finally we give an alternative appproach via the bilinear KP and Miura transformations. The chapter concludes with a series of examples and computer graphics, mainly for resonance phenomena in the context of the KP-II, indicating how to exploit our solution formula in future research.

Chapter 4 In the first part we study soliton-like solutions of the general AKNS system. These are the solutions which reduce to N-solitons if appropriate constraints are imposed. In Theorem 4.1.4 we obtain explicit expressions of these solutions. The second part may be considered as a preparation for the deeper study of negatons. We arrange the needed formulas, clarify the notion of negatons, and provide a motivating discussion of the particular case of N-solitons. The most substantial results of the second part are Proposition 4.3.7, Proposition 4.4.2, and Proposition 4.4.7, where conditions for global regularity and reality are proved.

Chapter 5 The contents of this chapter is the complete asymptotic description of negatons for the $\mathbb{C}$-reduced AKNS system. This is done in Theorem 5.1.2. In Theorem 5.2.2 we provide a corresponding result for the $\mathbb{R}$-reduced AKNS system. It is worth mentioning that in the latter case we integrated also the simplest type of formation of solitons into the analysis. Formations are solutions forming bound states in some sense and cannot be separated in asymptotic terms. Our result means that negatons can also consist of breathers, not only of solitons. The largest part of the chapter is concerned with the geometric part of the proof of Theorem 5.1.2. To illustrate our result we finally gather computer graphics of negatons for the sine-Gordon and the Nonlinear Schrödinger equations.

Chapter 6 Here we supply the calculations needed to determine the phase-shifts in Theorem 5.1.2. The argument reduces to the evaluation of very complicated determinants. The main result in Theorem 6.1.1 is a substantial generalization of a classical identity of Cauchy.

Chapter 7 In Theorem 7.5.5 we construct countable superpositions of negatons for the $\mathbb{C}$-reduced AKNS system, and in Theorem 7.6.3 for the $\mathbb{R}$-reduced AKNS system even superpositions where beside solitons also breathers can be admitted. Because of the integral terms appearing in the AKNS system, it seems unavoidable to assume that the appearing waves are of controlled width. But for equations for which the integral terms cancel, we can go further and drop this restriction (see Theorem 7.5.11 for the Nonlinear Schrödinger and

Theorem 7.6.4 for the modified Korteweg-de Vries equations). The idea is to compensate the non-solvability of elementary equations by a gain of summability obtained by appropriate factorizations through intermediate spaces, in the spirit of the Grothendieck theorem.

Appendices In Appendix A we establish the (non-obvious) link to the standard way of constructing negatons by Wronskian determinants. These formulas can be derived via Darboux transformations [56]. Appendix B contains a concise introduction to the theory of traces and determinants on quasi-Banach operator ideals. In Appendix C we give, for the sake of comparison with the general case, a straightforward proof of Theorem 3.2.1 in the particular case of commuting parameters. This result was already stated, but not proved, in a former joint article [18] with B. Carl.

Acknowledgment

I am grateful to many colleagues and friends. First of all I have to mention my academical teachers Bertram Huppert, who led my first steps in science, and especially Bernd Carl. I owe him the introduction to my area of research, many hints and ideas, and support without reserve in the critical moments of my work. For scientific exchange and inspirations I would like to thank the members of our working group, Ernst Dietzel, Aicke Hinrichs, Bernd Jende, Rainer Oloff, Christian Richter and Irmtraud Stephani.

I had the good luck to work in the mathematical institute of the university Jena in optimal conditions and a friendly atmosphere. Exemplarily I would like to thank the professors Hans Triebel as the speaker of the Graduiertenkolleg ‘Analytische und Stochastische Strukturen und Systeme’ where I started and Erich Novak as the actual director of the institute.

My research has been generously supported by two research projects ”‘Methoden der Geometrischen Analysis und der Funktionalanalysis bei der Behandlung von nichtlinearen Gleichungen der Solitonentheorie”’ and ”‘Integrable Systeme in der Theorie zweidimensionaler Wasserwellen, der Quantenoptik und der Allgemeinen Relativitätstheorie”’ of the Deutsche Forschungsgemeinschaft as well as a Habilitationsstelle of the Land Thüringen. Without these positions the present work would not have been possible.

To Alexander Sakhnovich I am grateful for important hints concerning the KP equation. I thank Heinz Steudel for giving advice on the appearance of muliple pole solutions in the literature and for providing unpublished work of his. For fruitfull discussions and hints I would like to thank Benno Fuchssteiner, Gerd Leuchs, Reinhard Meinel, and Gernot Neugebauer.

Finally I would like to thank the professors Bernd Carl, Benno Fuchssteiner and Alexander Sakhnovich for kindly agreeing to write the reports.

Contents

Chapter 1

Operator-theoretic treatment of the AKNS system

One of the main mysteries in soliton theory is whether there is a precise meaning of integrability or, at least, whether the familiar soliton equations of mathematical physics can be understood as members of a higher structure. Building on important work of Zhakharov and Shabat [101], it was discovered by Ablowitz, Kaup, Newell, and Segur [3] (see also [5]) that a good deal of the most prominent soliton equations (the Nonlinear Schrödinger equation, the Korteweg-de Vries and modified Korteweg-de Vries equations, the sine-Gordon and sinh-Gordon equations) can in fact be obtained as reductions of a general system of two equations. Moreover, the structural difference between the Nonlinear Schrödinger equation with its typical complex nature and the remaining above mentioned equations is reflected by a different level of reduction.

The operator method used in the present work was first developed for particular equations ([8], [18], [88], [89], [92], see [19] for a comparison and further references). It turned out that there is no canonical way to guess the appropriate operator equation, and it was not clear why certain equations required much more complicated twists than others.

This was one of our main motivations to obtain a complete operator treatment for the general AKNS system, which would yield the comprised individual equations via reduction. As for the operator equation we can build on previous work by Bauhardt and Pöppe [10]. But we need several extensions: Most importantly, their operator solution is not a generalization of the full scalar one-soliton (1.3), (1.4) with *two* parameters a, b but only of the one-soliton subject to the constraint $a = b$. As a consequence it seems hard (if possible) to derive merely the full family of N-solitons, whereas our formula with the *two* operator-valued parameters yield them via the most obvious choices. Furthermore the verification of the solution property in [10] is only formal to a certain extent. These gaps are supplemented in Theorem 1.2.1 and its proof. With respect to scalarization, [10] does not furnish handy determinant formulas, which are badly needed in our applications. This will be the content of Chapter 2. It should be mentioned that there was a different operator approach by Blohm [13], [14]. Based on operator algebra of Marchenko [55], he described a computational algorithm how to construct solutions for a hierarchy closely connected with the ZS system. But he did not consider uniform formulas for the whole system.

We proceed as follows. First we briefly discuss the AKNS system together with a number of prototypical soliton equations, which are contained in it.

Section 1.2 contains the essence of the chapter. Here we formulate the operator-valued AKNS system and construct in Theorem 1.2.1 a solution class depending on two independent operator-valued parameters. This operator solution is the basis of all later applications.

The next task is to take reductions into account. On the operator level, we hit on certain difficulties, which we will explain at the examples of the prototypical equations. It will turn out that reductions have to be treated in connection with scalarization, which will be the topic of Chapter 2.

1.1 The AKNS system

For the sake of motivation we will, following [3], [5], recall the basic facts on the scalar AKNS system and its reductions. In particular we will record how to obtain the Nonlinear Schrödinger, the modified Korteweg-de Vries, and the sine-Gordon equations as special cases.

For given non-trivial polynomials f, $g \in \mathbb{C}[x]$ the AKNS system reads

$$g\left(T_{r,q} \right) \begin{pmatrix} r_t \\ q_t \end{pmatrix} = f\left(T_{r,q} \right) \begin{pmatrix} r \\ -q \end{pmatrix}. \tag{1.1}$$

It is an integro-differential system in two unknown functions $r(x,t)$, $q(x,t)$. By $T_{r,q}$ we denote the (r,q)-dependent operator

$$\begin{pmatrix} u \\ v \end{pmatrix} \longmapsto T_{r,q} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} u_x - 2r\left(\int_{-\infty}^x qu\, d\xi + \int_{-\infty}^x rv\, d\xi \right) \\ -v_x + 2q\left(\int_{-\infty}^x qu\, d\xi + \int_{-\infty}^x rv\, d\xi \right) \end{pmatrix}. \tag{1.2}$$

Interpreting $\left(T_{r,q} \right)^n$ as n-fold iteration, we obtain operators $f\left(T_{r,q} \right)$, $g\left(T_{r,q} \right)$ acting on pairs of functions and arrive at (1.1) by inserting $\begin{pmatrix} r \\ -q \end{pmatrix}$, $\begin{pmatrix} r_t \\ q_t \end{pmatrix}$.

Often it will be more conceptual to formulate conditions in terms of the rational function $f_0 = f/g$. Some calculations even get formally simpler for

$$\begin{pmatrix} r_t \\ q_t \end{pmatrix} = f_0\left(T_{r,q} \right) \begin{pmatrix} r \\ -q \end{pmatrix}.$$

But we will avoid this because manipulation with $f_0(T_{r,q})$ leads to serious justification problems concerning invertibility.

To settle existence in (1.1), one may restrict to functions r, q which are sufficiently smooth on an open set $\mathbb{R}_x \times (t_1, t_2)$ and decay sufficiently rapidly for $x \to -\infty$.

In practice, this means that explicit formulas for $f\left(T_{r,q} \right)$, $g\left(T_{r,q} \right)$ can be computed, where we are allowed to simplify expressions by (i) interchanging differentiation and integration and (ii) applying partial integration on intervals $(-\infty, x)$ with zero boundary values in $x = -\infty$ (see Section 1.1.1 for an example).

It is beyond the scope of these introductory remarks to give an adequate impression of the by now classical theory of (1.1) as developed by means of the inverse scattering method. However we mention that without further restrictions (1.1) is not an integrable system and the inverse scattering method applies only to a certain extent. In particular the corresponding Levitan-Gelfand-Marchenko equation need not always be solvable.

Geometrically this is reflected by the existence of solutions with instantaneous singularities. As remarked in [3], V.A, one may for example look at the one-soliton, depending on four complex parameters a, b, φ, ψ. Setting

$$\ell(x,t) = \exp(ax + f_0(a)t + \varphi), \qquad m(x,t) = \exp(bx - f_0(-b)t + \psi),$$

the one-soliton is given by

$$q = (a+b)m/(1 - \ell m), \tag{1.3}$$
$$r = (a+b)\,\ell/(1 - \ell m). \tag{1.4}$$

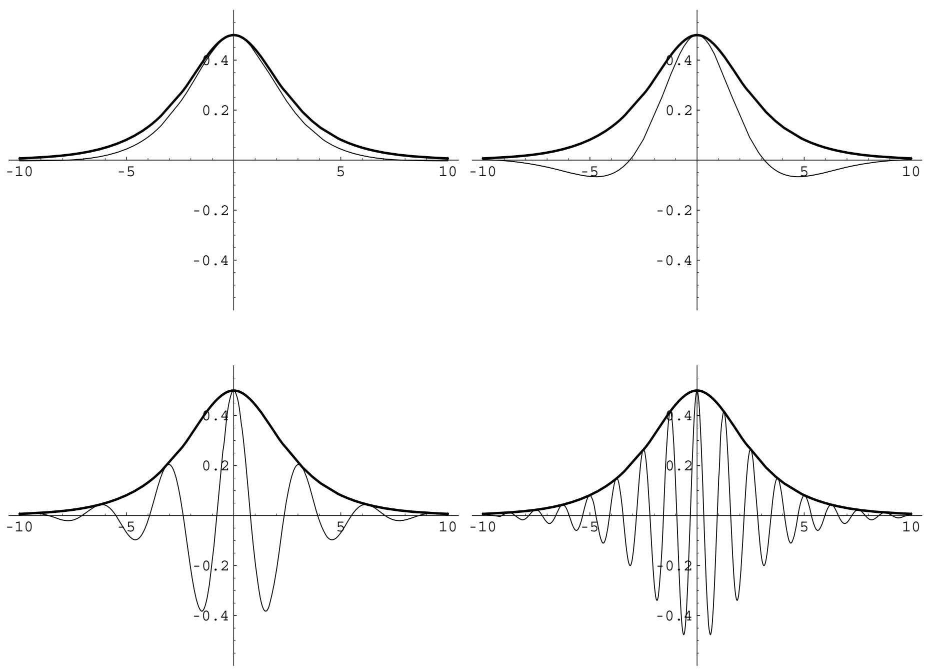
Now it can easily be arranged that the denominator does not vanish, say, for $t = 0$ and all $x \in \mathbb{R}$, but that there are singularities after a finite time t_0. Hence the AKNS system is not specific enough to prevent solutions from exploding!

The key idea is to enforce integrability by reduction. A first reduction, which we will call $\mathbb{C}$-reduction, consists in requiring $r = -\overline{q}$ leading to the condition $\overline{f_0(z)} = f_0(-\overline{z})$ for the rational function $f_0 = f/g$. For the one-soliton (1.3), $\mathbb{C}$-reduction implies $a = \overline{b}$, $\varphi = \overline{\psi} + \mathrm{i}\pi$ (see [3]), and a straightforward calculation yields

$$q(x,t) \quad = \quad -\mathrm{Re}(a)\, e^{-\mathrm{i}\,\mathrm{Im}\left(\Gamma(x,t)\right)}\, \cosh^{-1}\left(\mathrm{Re}\left(\Gamma(x,t)\right)\right)$$
$$\text{for } \Gamma(x,t) = ax + f_0(a)t + \varphi.$$

In particular, q is regular everywhere, has amplitude $\mathrm{Re}(a)$, and moves with velocity $v = -\mathrm{Re}\left(f_0(a)\right)/\mathrm{Re}(a)$.

For illustration the reader finds figures for $\varphi = \mathrm{i}\pi$ and varying a below. The figures show the modulus (thick line) and the real part (thin line) of $q(x,0)$ for the fixed time $t = 0$.



In all diagrams $\mathrm{Re}(a) = \frac{1}{2}$, and successively $\mathrm{Im}(a) = \frac{1}{5}$, $\frac{1}{2}$, 2, and 5.

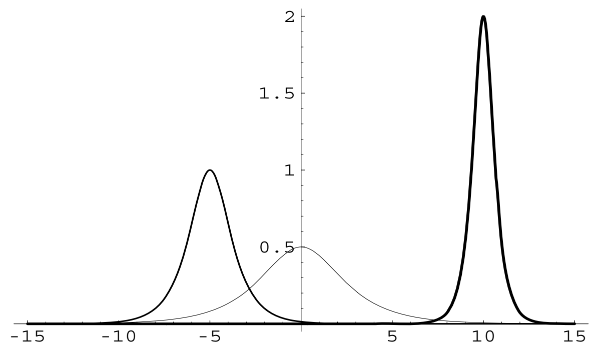
It is an important result that stability of one-solitons is typical. In [3] (cf. [26]) it is shown that the inverse scattering method applies to full extent to the $\mathbb{C}$-reduced AKNS system.

There is a second reduction where one assumes in addition that the functions r, q, and f_0 are real. Then a in the one-soliton of the $\mathbb{C}$-reduction is real, and $\varphi = \varphi_0 + \mathrm{i}\pi k$ with $\varphi_0 \in \mathbb{R}$, $k \in \mathbb{Z}$. We obtain

$$q(x,t) = -a\,\epsilon\,\cosh^{-1}\left(\Gamma(x,t)\right) \qquad \text{for } \Gamma(x,t) = ax + f_0(a)t + \varphi_0,$$

where $\epsilon = -1$ for k odd and $\epsilon = 1$ for k even. Note that there is no oscillating term. We call the second reduction the $\mathbb{R}$-reduced AKNS system.

The figures below show $q(x,0)$ for $\epsilon=-1$ and varying values of a. To plot clear figures, here we added position shifts φ_0.



Successively $a=1$, $\frac{1}{2}$, and 2.

1.1.1 Prototypical equations in the AKNS system: Nonlinear Schrödinger, sine-Gordon, and modified Korteweg-de Vries equations

There are three prototypical equations contained in the AKNS system: The Nonlinear Schrödinger equation (NLS), the modified Korteweg-de Vries equation (mKdV), and the sine-Gordon equation (sG). In the sequel we explain how they can be obtained and why they are prototypical. In the first example we will perform the ensuing calculations in detail.

The Nonlinear Schrödinger equation

To obtain the Nonlinear Schrödinger equation, we consider $f_0(z)=-\mathrm{i}z^2$. In other words, we have $f(z)=-\mathrm{i}z^2$, $g(z)=1$. Because

$$\begin{aligned}
T_{r,q}\begin{pmatrix} r \\ -q \end{pmatrix} &= \begin{pmatrix} r_x - 2r\Big(\int_{-\infty}^x rq\,d\xi - \int_{-\infty}^x rq\,d\xi\Big) \\ q_x + 2q\Big(\int_{-\infty}^x rq\,d\xi - \int_{-\infty}^x rq\,d\xi\Big) \end{pmatrix} = \begin{pmatrix} r_x \\ q_x \end{pmatrix}, \\
\big(\,T_{r,q}\,\big)^2\begin{pmatrix} r \\ -q \end{pmatrix} &= T_{r,q}\begin{pmatrix} r_x \\ q_x \end{pmatrix} = \begin{pmatrix} r_{xx} - 2r\Big(\int_{-\infty}^x r_x q\,d\xi + \int_{-\infty}^x rq_x\,d\xi\Big) \\ -q_{xx} + 2q\Big(\int_{-\infty}^x r_x q\,d\xi + \int_{-\infty}^x rq_x\,d\xi\Big) \end{pmatrix} \\
&= \begin{pmatrix} r_{xx} - 2r\int_{-\infty}^x (rq)_x\,d\xi \\ -q_{xx} + 2q\int_{-\infty}^x (rq)_x\,d\xi \end{pmatrix} = \begin{pmatrix} r_{xx} - 2r^2q \\ -q_{xx} + 2q^2r \end{pmatrix},
\end{aligned}$$

the system (1.1) associated with this choice of f_0 becomes

$$\begin{aligned}
-\mathrm{i}r_t + r_{xx} - 2r^2q &= 0, \\
\mathrm{i}q_t + q_{xx} - 2q^2r &= 0.
\end{aligned}$$

Actually this choice is quite natural, since we aim at an equation which is first order in the time vaiable t and second order in the space variable x. Thus it is immediately clear that f_0 should be a polynomial of degree two.

In particular, after the reduction $r=-\overline{q}$, we arrive at the NLS

$$\mathrm{i}q_t + q_{xx} + 2q^2\overline{q} = 0. \tag{1.5}$$

If we omit all possible position shifts, the one-soliton of the NLS is

$$q(x,t) = -\mathrm{Re}(a)\, e^{-\mathrm{i}\left(\mathrm{Im}(a)x + \left[\mathrm{Im}(a)^2 - \mathrm{Re}(a)^2\right]t\right)} \cosh^{-1}\Big(\mathrm{Re}(a)\, \left[\, x + 2\mathrm{Im}(a)t \,\right]\Big).$$

Its amplitude is determined by $\mathrm{Re}(a)$, whereas its velocity equals $v = -2\mathrm{Im}(a)$. In particular, both can be chosen independently.

The modified Korteweg-de Vries equation

For the modified Korteweg-de Vries equation, the right choice again is a polynomial. If we take $f_0(z) = -z^3$, then (1.1) reads

$$\begin{aligned} r_t + r_{xxx} - 6rr_xq &= 0, \\ q_t + q_{xxx} - 6qq_xr &= 0. \end{aligned}$$

In contrast to the first example, here we not only assume $r = -\overline{q}$ but also that q is real. Consequently $r = -q$, and we end up with the mKdV

$$q_t + q_{xxx} + 6q^2 q_x = 0. \tag{1.6}$$

As for the one-soliton, we get, neglecting again position shifts,

$$q(x,t) \;=\; -a\cosh^{-1}\left(a[x - a^2 t]\right),$$

where $a \in \mathbb{R}$. Here the amplitude $-a$ and the velocity $v = a^2$ are coupled.

The sine-Gordon equation

Whereas in both other examples f_0 was a polynomial, the sine-Gordon equation is obtained for $f_0(z) = 1/z$. Using the right interpretation of the system (1.1) with $g(z) = z$, $f(z) = 1$, we get

$$\begin{aligned} r_{tx} - 2r\int_{-\infty}^{x} (qr)_t \, d\xi - r &= 0, \\ q_{tx} - 2q\int_{-\infty}^{x} (qr)_t \, d\xi - q &= 0. \end{aligned}$$

In particular, the same restriction as in the second example, $r = -\overline{q}$, q real, yields a derivative version of the sG,

$$q_{tx} + 2q\int_{-\infty}^{x} (q^2)_t \, d\xi = q. \tag{1.7}$$

Its one-soliton is

$$q(x,t) = -a\cosh^{-1}\left(ax + t/a\right)$$

for $a \in \mathbb{R}$. Again amplitude $-a$ and velocity $v = -1/a^2$ are coupled.

The relation of (1.7) to the usual form of the sine-Gordon equation $u_{xt} = \sin(u)$ is given by the transformation $u = -2\int_{-\infty}^{x} q\, d\xi$ (see [3]). More precisely, if u is a solution of the sine-Gordon equation decaying sufficiently fast for $x \to -\infty$, then we have for $q = -u_x/2$

$$\begin{aligned} q_{tx} + 2q\int_{-\infty}^{x} (q^2)_t \, d\xi &= -\frac{1}{2}\left(u_{txx} + u_x\int_{-\infty}^{x} u_{xt}u_x \, d\xi\right) \\ &= -\frac{1}{2}\left((\sin u)_x + u_x\int_{-\infty}^{x} \sin(u)u_x \, d\xi\right) = -\frac{1}{2}\left(\cos(u)u_x - u_x\int_{-\infty}^{x} (\cos u)_x \, d\xi\right) \\ &= -\frac{1}{2}u_x \cdot \lim_{x\to-\infty} \cos(u) \;=\; q. \end{aligned}$$

Thus $q = -u_x/2$ solves (1.7).

Nevertheless, the sine-Gordon itself can also be treated directly by the operator-method presented in this work. The reader may find the corresponding results (including the rigorous asymptotic analysis of negatons) in [88], [90].

Maybe it is worth remarking that $u = -2\int_{-\infty}^{x} q d\xi$ is just the kink or antikink solution of the sine-Gordon equation (depending on the sign of a). Transferred to laboratory coordinates, one obtains the usual pictures.

1.2 The non-abelian AKNS system

Next we present an operator-valued version of the AKNS system. Our approach is a generalization of the work of Bauhardt and Pöppe. In [10] they introduced and studied the system (1.8), (1.9). The most important difference to our treatment is that we deduce solution formulas with *two* operator parameters. This will, already in the case of solitons, be essential to obtain complete families of solutions (see Remark 1.2.12 b)).

For given polynomials $f, g \in \mathbb{C}[x]$, the non-abelian AKNS system reads

$$g\left(\mathcal{T}_{R,Q}\right)\begin{pmatrix} R_t \\ Q_t \end{pmatrix} = f\left(\mathcal{T}_{R,Q}\right)\begin{pmatrix} R \\ -Q \end{pmatrix}, \tag{1.8}$$

where the unknown operator-valued functions $R(x,t)$, $Q(x,t)$ take values in $\mathcal{L}(F,E)$, $\mathcal{L}(E,F)$, respectively, and $\mathcal{T}_{R,Q}$ denotes the operator

$$\begin{pmatrix} U \\ V \end{pmatrix} \longmapsto \mathcal{T}_{R,Q}\begin{pmatrix} U \\ V \end{pmatrix} = \begin{pmatrix} U_x - \Big(R\int_{-\infty}^{x}(QU+VR)d\xi + \int_{-\infty}^{x}(UQ+RV)d\xi\, R\Big) \\ -V_x + \Big(Q\int_{-\infty}^{x}(UQ+RV)d\xi + \int_{-\infty}^{x}(QU+VR)d\xi\, Q\Big) \end{pmatrix}, \tag{1.9}$$

for $U = U(x,t) \in \mathcal{L}(F,E)$, $V = V(x,t) \in \mathcal{L}(E,F)$. The operators $f\left(\mathcal{T}_{R,Q}\right)$, $g\left(\mathcal{T}_{R,Q}\right)$ are defined as in the preceding section. For Banach spaces E, F, we denote by $\mathcal{L}(E,F)$ as usual the Banach space of bounded linear operators equipped with the standard operator norm.

Existence of the expressions appearing in (1.8) may be ensured similarly as in the scalar case. One requires that R, Q are sufficiently smooth and behave sufficiently well for $x \to -\infty$. More precisely the latter means for the operator-function $R = R(x,t)$: For some sufficiently large $n_0 \in \mathbb{N}$ (depending on the degrees of f, g), one requires that the t-dependent expressions

$$\| R(\cdot,t) \|_{\kappa,\lambda} = \sup_{x\in\mathbb{R}} (1+|x|)^{\lambda} \left\| \frac{\partial^{\kappa}}{\partial x^{\kappa}} R(x,t) \right\|$$

are finite for κ, $\lambda \le n_0$. Moreover, the map $t \mapsto R(\cdot,t)$ has to be continuous with respect to $\|\cdot\|_{\kappa,\lambda}$, $\kappa, \lambda \le n_0$. Finally, the t-derivative $R_t(x,t)$ is required to have the same properties.

The main point is that then the integrals appearing in the definition of $\mathcal{T}_{R,Q}$ can be evaluated as Bochner integrals. The continuity hypothesis in t ensures that in (1.8) both sides are at least continuous.

However, these aspects are somewhat peripheral to our purposes, because existence will always be evident in our main results, and the solutions will even decay faster than any polynomial together with their derivatives. Therefore we do not make an effort to state optimal assumptions on smoothness and decay. Roughly speaking, almost all later difficulties will be situated in the target spaces $\mathcal{L}(F,E)$, $\mathcal{L}(E,F)$, and not in spacetime.

The following theorem is our fundamental tool for the study of the AKNS system.

Theorem 1.2.1. *Let E, F be Banach spaces, $A \in \mathcal{L}(E)$, $B \in \mathcal{L}(F)$ 0 constant operators such that* $\operatorname{spec}(A) \cup \operatorname{spec}(-B)$ *is contained in the domain where f_0 is holomorphic.*

Assume that $L = L(x,t) \in \mathcal{L}(F,E)$, $M = M(x,t) \in \mathcal{L}(E,F)$ are operator-valued functions which, on a strip $\{(x,y)|x \in \mathbb{R}, y \in (t_1,t_2)\}$, are sufficiently smooth and behave sufficiently well for $x \to -\infty$, and solve the base equations

$$\begin{aligned} L_x &= AL, & L_t &= f_0(A)L, \\ M_x &= BM, & M_t &= -f_0(-B)M. \end{aligned}$$

Assume furthermore that $(I-LM)$, $(I-ML)$ are invertible on $\mathbb{R} \times \big(t_1,t_2\big)$.

Then

$$Q = (I-ML)^{-1}(BM+MA), \tag{1.10}$$

$$R = (I-LM)^{-1}(AL+LB). \tag{1.11}$$

is a solution of (1.8) on $\mathbb{R} \times \big(t_1,t_2\big)$.

Note that the expressions $f_0(A)$, $f_0(-B)$ do not cause complications because the arguments A, B are bounded operators and the holomorphic function calculus works in the usual way (see [82], [97]).

Remark 1.2.2. *a) In [10] several additional technical assumptions are made. But the main difference is that [10] assumes $A = B$.*

b) The solution (1.10), (1.11) is formally analogous to the one-soliton. Indeed, we recover (1.3), (1.4) if we take all functions scalar-valued. In the applications we shall see that the operator-valued parameters A, B give access to a huge variety of solutions.

c) A comparison with the inverse scattering method supports the point of view that the underlying spectral data are encoded in the operators A, B. The fact that A, B map between different Banach spaces reflects that, in the scalar case of the general AKNS system, the contributions of $r(x,0)$, $q(x,0)$ to the spectral data are completely unrelated.

Before entering the proof, we collect some preparational material. We start with the following easy observation.

Lemma 1.2.3. *Let E, F be Banach spaces and $S \in \mathcal{L}(F,E)$, $T \in \mathcal{L}(E,F)$ arbitrary operators such that the inverses $(I-ST)^{-1}$, $(I-TS)^{-1}$ exist. Then the following identities hold:*

$$(I-TS)^{-1}T = T(I-ST)^{-1}, \tag{1.12}$$

$$S(I-TS)^{-1}T = (I-ST)^{-1} - I, \tag{1.13}$$

Proof To verify (1.12), we use $T(I-ST) = (I-TS)T$ and multiply it by $(I-ST)^{-1}$ from the left and by $(I-TS)^{-1}$ from the right. As a consequence,

$$\begin{aligned} S(I-TS)^{-1}T &= ST(I-ST)^{-1} \\ &= \Big(I - (I-ST)\Big)(I-ST)^{-1} \\ &= (I-ST)^{-1} - I \end{aligned}$$

which is (1.13). □

Next we recall the non-abelian differentiation rule for inverse operators.

Lemma 1.2.4. *Let $T = T(s) \in \mathcal{L}(E)$ be a family of operators depending on a real variable s which is differentiable with respect to s and invertible for all $s \in \mathbb{R}$. Then $T^{-1}(s)$ is differentiable, and we have*

$$T_s^{-1} = -T^{-1}T_sT^{-1}.$$

In the proof of Theorem 1.2.1 we will introduce certain operator-valued functions. The subsequent lemmata contain rules for the manipulations with these functions.

Lemma 1.2.5. *Let E, F be a Banach spaces and $A \in \mathcal{L}(E)$, $B \in \mathcal{L}(F)$ constant operators. Let $L = L(s) \in \mathcal{L}(F,E)$, $M = M(s) \in \mathcal{L}(E,F)$ be operator-valued functions which are differentiable with respect to the real variable s and satisfy the base equations*

$$L_s = AL, \qquad M_s = BM,$$

and assume that $(I - LM)$, $(I - ML)$ are always invertible.

Furthermore, let $A_n \in \mathcal{L}(E)$, $B_n \in \mathcal{L}(F)$, $n \in \mathbb{N}_0$, be constant operators with $[A, A_n] = [B, B_n] = 0$ for all n.

Define the following operator-valued functions

$$T = (I - LM)^{-1}(AL + LB), \qquad \widehat{T} = (I - LM)^{-1}(A + LBM),$$
$$S = (I - ML)^{-1}(BM + MA), \qquad \widehat{S} = (I - ML)^{-1}(B + MAL),$$

and, for $n \in \mathbb{N}$,

$$T_n = (I - LM)^{-1}(A_nL + LB_n), \qquad \widehat{T}_n = (I - LM)^{-1}(A_n + LB_nM),$$
$$S_n = (I - ML)^{-1}(B_nM + MA_n), \qquad \widehat{S}_n = (I - ML)^{-1}(B_n + MA_nL).$$

Then the following derivation rules hold for all $n \in \mathbb{N}_0$:

$$T_{n,s} = \widehat{T}T_n, \tag{1.14}$$
$$\widehat{T}_{n,s} = TS_n, \tag{1.15}$$
$$S_{n,s} = \widehat{S}S_n, \tag{1.16}$$
$$\widehat{S}_{n,s} = ST_n. \tag{1.17}$$

Proof We start with (1.14). From Lemma 1.2.4, the base equations, and the fact that $[A, A_n] = 0$, we observe

$$\begin{aligned}
T_{n,s} &= -(I - LM)^{-1}(-LM)_s(I - LM)^{-1}\Big(A_nL + LB_n\Big) \\
&\qquad\qquad\qquad\qquad + (I - LM)^{-1}\Big(A_nL + LB_n\Big)_s \\
&= (I - LM)^{-1}\Big((AL + LB)M\Big)(I - LM)^{-1}\Big(A_nL + LB_n\Big) \\
&\qquad\qquad\qquad\qquad + (I - LM)^{-1}A\Big(A_nL + LB_n\Big) \\
&= (I - LM)^{-1}\Big((AL + LB)M + A(I - LM)\Big)\, T_n \\
&= (I - LM)^{-1}\Big(A + LBM\Big)\, T_n \\
&= \widehat{T}\, T_n.
\end{aligned}$$

Analogously we find

$$\begin{aligned}
\widehat{T}_{n,s} &= -(I - LM)^{-1}(-LM)_s(I - LM)^{-1}\Big(A_n + LB_nM\Big) \\
&\qquad\qquad\qquad\qquad + (I - LM)^{-1}\Big(A_n + LB_nM\Big)_s \\
&= (I - LM)^{-1}\Big((AL + LB)M\Big)(I - LM)^{-1}\Big(A_n + LB_nM\Big) \\
&\qquad\qquad\qquad\qquad + (I - LM)^{-1}\Big(\big(AL + LB\big)B_nM\Big)
\end{aligned}$$

$$
\begin{aligned}
&= T\Big(M(I-LM)^{-1}\Big(A_n+LB_nM\Big)+B_nM\Big)\\
&= T\,(I-ML)^{-1}\Big(M\Big(A_n+LB_nM\Big)+(I-ML)B_nM\Big)\\
&= T\,(I-ML)^{-1}\Big(MA_n+B_nM\Big)\\
&= T\,S_n,
\end{aligned}
$$

where we have in addition used Lemma 1.2.3 for the fourth identity.

This is (1.15). The argument for (1.16), (1.17) is completely symmetric. □

The next lemma only concerns constant operators.

Lemma 1.2.6. *Let E, F be a Banach spaces and $L \in \mathcal{L}(F,E)$, $M \in \mathcal{L}(E,F)$ be operators such that $(I-LM)$, $(I-ML)$ are invertible. Furthermore, let $A, A_n \in \mathcal{L}(E)$, $B, B_n \in \mathcal{L}(F)$, $n \in \mathbb{N}_0$, satisfy (i) $[A,A_n]=[B,B_n]=0$ $\forall n$ and (ii) $A_{n+1}=A_nA$, $B_{n+1}=-B_nB$ $\forall n$. Define*

$$
\begin{aligned}
T &= (I-LM)^{-1}(AL+LB), & \widehat{T} &= (I-LM)^{-1}(A+LBM),\\
S &= (I-ML)^{-1}(BM+MA), & \widehat{S} &= (I-ML)^{-1}(B+MAL),
\end{aligned}
$$

and, for $n \in \mathbb{N}_0$,

$$
\begin{aligned}
T_n &= (I-LM)^{-1}(A_nL+LB_n), & \widehat{T}_n &= (I-LM)^{-1}(A_n+LB_nM),\\
S_n &= (I-ML)^{-1}(B_nM+MA_n), & \widehat{S}_n &= (I-ML)^{-1}(B_n+MA_nL).
\end{aligned}
$$

Then the following identities hold for all $n \in \mathbb{N}_0$:

$$
\begin{aligned}
\widehat{T}_n\widehat{T}-T_nS &= \widehat{T}_{n+1}, && (1.18)\\
\widehat{T}\widehat{T}_n-TS_n &= \widehat{T}_{n+1}, && (1.19)\\
\widehat{T}_nT-T_n\widehat{S} &= T_{n+1}, && (1.20)\\
\widehat{T}T_n-T\widehat{S}_n &= T_{n+1}, && (1.21)\\
\widehat{S}_n\widehat{S}-S_nT &= -\widehat{S}_{n+1}, && (1.22)\\
\widehat{S}\widehat{S}_n-ST_n &= -\widehat{S}_{n+1}, && (1.23)\\
\widehat{S}_nS-S_n\widehat{T} &= -S_{n+1}, && (1.24)\\
\widehat{S}S_n-S\widehat{T}_n &= -S_{n+1}. && (1.25)
\end{aligned}
$$

Proof Let us start with (1.18). We calculate

$$
\begin{aligned}
&(I-LM)\,T_nS = \Big(A_nL+LB_n\Big)(I-ML)^{-1}\Big(MA+BM\Big)\\
&\quad= A_n\Big(L(I-ML)^{-1}M\Big)A+LB_n\Big((I-ML)^{-1}\Big)BM\\
&\qquad\qquad +LB_n\Big((I-ML)^{-1}M\Big)A+A_n\Big(L(I-ML)^{-1}\Big)BM.
\end{aligned}
$$

At this point we apply Lemma 1.2.3 to the large brackets in order to change $(I-ML)^{-1}$ into $(I-LM)^{-1}$ wherever it appears. We get

$$
\begin{aligned}
&(I-LM)\,T_nS\\
&\quad= A_n\Big((I-LM)^{-1}-I\Big)A+LB_n\Big(I+M(I-LM)^{-1}L\Big)BM\\
&\qquad\qquad +LB_n\Big(M(I-LM)^{-1}\Big)A+A_n\Big((I-LM)^{-1}L\Big)BM\\
&\quad= -\Big(AA_n-LBB_nM\Big)+\Big(A_n+LB_nM\Big)(I-LM)^{-1}\Big(A+LBM\Big)
\end{aligned}
$$

$$
\begin{aligned}
&= -\Big(A_{n+1} + LB_{n+1}M\Big) + \Big(A_n + LB_nM\Big)(I-LM)^{-1}\Big(A + LBM\Big) \\
&= (I-LM)\Big(-\widehat{T}_{n+1} + \widehat{T}_n\widehat{T}\Big),
\end{aligned}
$$

which proves (1.18). As for (1.19), we can use completely the same line of arguments just exchanging the roles of A and A_n, B and B_n.

As for (1.20), we calculate

$$
\begin{aligned}
(I-LM)\,\widehat{T}_nT &= \Big(A_n + LB_nM\Big)(I-LM)^{-1}\Big(AL + LB\Big) \\
&= LB_n\Big(M(I-LM)^{-1}L\Big)B + A_n\Big((I-LM)^{-1}\Big)AL \\
&\qquad +LB_n\Big(M(I-LM)^{-1}\Big)AL + A_n\Big((I-LM)^{-1}L\Big)B \\
&= LB_n\Big((I-ML)^{-1} - I\Big)B + A_n\Big(I + L(I-ML)^{-1}M\Big)AL \\
&\qquad +LB_n\Big((I-ML)^{-1}M\Big)AL + A_n\Big(L(I-ML)^{-1}\Big)B \\
&= \Big(AA_nL - LBB_n\Big) + \Big(A_nL + LB_n\Big)(I-ML)^{-1}\Big(B + MAL\Big) \\
&= \Big(A_{n+1}L + LB_{n+1}\Big) + \Big(A_nL + LB_n\Big)(I-ML)^{-1}\Big(B + MAL\Big) \\
&= (I-LM)\Big(T_{n+1} + T_n\widehat{S}\Big),
\end{aligned}
$$

and again (1.21) can be verified by the same calculation, exchanging the roles of A and A_n, B and B_n.

Now we turn to (1.22). Here the argument is symmetric to the verification of (1.18), but there is an additional sign in front of the term $\widehat{S}_{n+1}$, which is due to the difference in the recursion relation for B_n. We observe

$$
\begin{aligned}
(I-ML)\,S_nT &= \Big(MA_n + B_nM\Big)(I-LM)^{-1}\Big(AL + LB\Big) \\
&= B_n\Big(M(I-LM)^{-1}L\Big)B + MA_n\Big((I-LM)^{-1}\Big)AL \\
&\qquad +B_n\Big(M(I-LM)^{-1}\Big)AL + MA_n\Big((I-LM)^{-1}L\Big)B \\
&= B_n\Big((I-ML)^{-1} - I\Big)B + MA_n\Big(I + L(I-ML)^{-1}M\Big)AL \\
&\qquad +B_n\Big((I-ML)^{-1}M\Big)AL + MA_n\Big(L(I-ML)^{-1}\Big)B \\
&= \Big(-BB_n + MAA_nL\Big) + \Big(B_n + MA_nL\Big)(I-ML)^{-1}\Big(B + MAL\Big) \\
&= \Big(B_{n+1} + MA_{n+1}L\Big) + \Big(B_n + MA_nL\Big)(I-ML)^{-1}\Big(B + MAL\Big) \\
&= (I-ML)\Big(\widehat{S}_{n+1} + \widehat{S}_n\widehat{S}\Big),
\end{aligned}
$$

and (1.23) follows by exchanging A and A_n, B and B_n.

Finally, we check (1.24). Here the arguments are close to the calculations for (1.20) except for the different recursion relation.

$$
\begin{aligned}
(I-LM)\,\widehat{S}_nS &= \Big(B_n + MA_nL\Big)(I-ML)^{-1}\Big(BM + MA\Big) \\
&= MA_n\Big(L(I-ML)^{-1}M\Big)A + B_n\Big((I-ML)^{-1}\Big)BM \\
&\qquad +MA_n\Big(L(I-ML)^{-1}\Big)BM + B_n\Big((I-ML)^{-1}M\Big)A
\end{aligned}
$$

$$
\begin{aligned}
&= MA_n\Big((I-LM)^{-1}-I\Big)A + B_n\Big(I+M(I-LM)^{-1}L\Big)BM \\
&\qquad\qquad +MA_n\Big((I-LM)^{-1}L\Big)BM + B_n\Big(M(I-LM)^{-1}\Big)A \\
&= \Big(BB_nM - MAA_n\Big) + \Big(B_nM+MA_n\Big)(I-LM)^{-1}\Big(A+LBM\Big) \\
&= -\Big(B_{n+1}M + MA_{n+1}\Big) + \Big(B_nM+MA_n\Big)(I-LM)^{-1}\Big(A+LBM\Big) \\
&= (I-ML)\Big(-S_{n+1}+S_n\widehat{T}\Big),
\end{aligned}
$$

Exchanging the roles of A and A_n, B and B_n, this argument also shows (1.25).
This completes the proof. □

The last preparation concerns the existence of the integrals in Theorem 1.2.1.

Lemma 1.2.7. *The derivative of the operator-valued functions $Q = Q(x,t)$, $R = R(x,t)$ given in Theorem 1.2.1 with respect to t is:*

$$
\begin{aligned}
Q_t &= (I-ML)^{-1}\Big(-f_0(-B)+Mf_0(A)L\Big)\,Q \\
R_t &= (I-LM)^{-1}\Big(\ \ f_0(A)\ -\ Lf_0(-B)M\Big)\,R.
\end{aligned}
$$

Proof The calculations are quite similar as the ones in Lemma 1.2.5. The only difference is that here we use the base equations $L_t = f_0(A)L$, $M_t = -f_0(-B)M$. Let us start with

$$
\begin{aligned}
Q_t &= -(I-ML)^{-1}(-ML)_t(I-ML)^{-1}\Big(BM+MA\Big) \\
&\qquad\qquad\qquad +(I-ML)^{-1}\Big(BM+MA\Big)_t \\
&= (I-ML)^{-1}\Big(\big(-f_0(-B)M+Mf_0(A)\big)L\Big)(I-ML)^{-1}\Big(BM+MA\Big) \\
&\qquad\qquad\qquad +(I-ML)^{-1}(-f_0(-B))\Big(BM+MA\Big) \\
&= (I-ML)^{-1}\Big(\big(-f_0(-B)M+Mf_0(A)\big)L - f_0(-B)(I-ML)\Big)Q \\
&= (I-ML)^{-1}\Big(-f_0(-B)+Mf_0(A)L\Big)\,Q,
\end{aligned}
$$

and, analogously,

$$
\begin{aligned}
R_t &= -(I-LM)^{-1}(-LM)_t(I-LM)^{-1}\Big(AL+LB\Big) \\
&\qquad\qquad\qquad +(I-LM)^{-1}\Big(AL+LB\Big)_t \\
&= (I-LM)^{-1}\Big(\big(f_0(A)L-Lf_0(-B)\big)M\Big)(I-LM)^{-1}\Big(AL+LB\Big) \\
&\qquad\qquad\qquad +(I-LM)^{-1}f_0(A)\Big(AL+LB\Big) \\
&= (I-LM)^{-1}\Big(\big(f_0(A)L-Lf_0(-B)\big)M + f_0(A)(I-LM)\Big)R \\
&= (I-LM)^{-1}\Big(f_0(A)-Lf_0(-B)M\Big)\,R,
\end{aligned}
$$

which completes the proof. □

Let us recall a specific form of the fundamental theorem of calculus for operator-valued functions depending on a real variable (see for example [24]).

Lemma 1.2.8. *Let E, F be Banach spaces and $T = T(s) \in \mathcal{L}(E,F)$ a family of operators which is continuously differentiable with respect to a real variable s and satisfies the boundary condition $T(s) \to 0$ for $s \to -\infty$. Then, for all s,*

$$T(s) = \int_{-\infty}^{s} T_s(\sigma) d\sigma.$$

Lemma 1.2.9. *Let $L = L(x,t)$, $M = M(x,t)$ be as given in Theorem 1.2.1. Then, for all constant operators $\widehat{A} \in \mathcal{L}(E)$, $\widehat{B} \in \mathcal{L}(F)$, the operator-functions*

$$\begin{aligned} \widehat{T} &= (I - LM)^{-1} \, (\widehat{A}L + L\widehat{B}), \\ \widehat{S} &= (I - ML)^{-1}(M\widehat{A} + \widehat{B}M), \end{aligned}$$

decay as $x \to -\infty$.

Proof It suffices to show the assertion for $\widehat{S} = \widehat{S}(x,t)$. To this end, fix t and choose x sufficiently large such that $||L(x,t)||$, $||M(x,t)|| < 1$. Then, by the Neumann series argument,

$$||(1 - ML)^{-1}|| = ||\sum_{k=0}^{\infty}(ML)^k|| \le \sum_{k=0}^{\infty} ||M||^k ||L||^k = \frac{1}{1 - ||M|| \, ||L||}.$$

Since $\widehat{S} = (1 - ML)^{-1}(\widehat{B}M + M\widehat{A})$, the boundary condition for $\widehat{S}$ follows easily from

$$||\widehat{S}|| \quad \le \quad (||\widehat{A}|| + ||\widehat{B}||)\frac{||M||}{1 - ||M|| \, ||L||} \to 0$$

as $x \to -\infty$ by assumption. □

Now we are in position to give the proof of Theorem 1.2.1.

Proof (of Theorem 1.2.1) The prove is divided into three steps. The aim of the first and the second step is to derive an explicit expression for the n-fold iteration $(\ \mathcal{T}_{R,Q}\ )^n$ applied to

$$\begin{pmatrix} R \\ -Q \end{pmatrix}, \qquad \begin{pmatrix} R_t \\ Q_t \end{pmatrix},$$

respectively. The last step combines these expressions to derive (1.8).

Step 1: To start with, we define the following hierarchy of auxiliary operator-valued functions for $n \in \mathbb{N}_0$:

$$\begin{aligned} R_n &= (I - LM)^{-1}\Big(A^n L - L(-B)^n\Big), \\ \widehat{R}_n &= (I - LM)^{-1}\Big(A^n - L(-B)^n M\Big), \\ Q_n &= (I - ML)^{-1}\Big(MA^n - (-B)^n M\Big), \\ \widehat{Q}_n &= (I - ML)^{-1}\Big(MA^n L - (-B)^n\Big). \end{aligned}$$

Note $R_0 = 0$, $R_1 = R$, $\widehat{R}_0 = I$, and, correspondingly, $Q_0 = 0$, $Q_1 = Q$, $\widehat{Q}_0 = -I$. In addition we define

$$U_n = \widehat{R}_n R_1, \qquad \text{and} \qquad V_n = \widehat{Q}_n Q_1.$$

Then we claim that, for $n \in \mathbb{N}_0$,

$$(\mathcal{T}_{R,Q})^n \begin{pmatrix} R \\ -Q \end{pmatrix} = \begin{pmatrix} U_n \\ V_n \end{pmatrix}. \tag{1.26}$$

Proof of Step 1: By Lemma 1.2.5 (with $A_n = A^n$, $B_n = -(-B)^n$), the differentiation rules for the operator-valued functions R_n, $\widehat{R}_n$, Q_n, and $\widehat{Q}_n$ are

$$\begin{aligned} R_{n,x} &= \widehat{R}_1 R_n, & (1.27)\\ \widehat{R}_{n,x} &= R_1 Q_n, & (1.28)\\ Q_{n,x} &= \widehat{Q}_1 Q_n, & (1.29)\\ \widehat{Q}_{n,x} &= Q_1 R_n. & (1.30) \end{aligned}$$

Next we state rules for the evaluation of certain products of the R_n, $\widehat{R}_n$, Q_n, and $\widehat{Q}_n$, following from Lemma 1.2.6 (with the same choices for A_n, B_n as above),

$$\begin{aligned} \widehat{R}_n \widehat{R}_1 - R_n Q_1 &= \widehat{R}_{n+1}, & (1.31)\\ \widehat{R}_n R_1 - R_n \widehat{Q}_1 &= R_{n+1}, & (1.32)\\ \widehat{R}_1 R_n - R_1 \widehat{Q}_n &= R_{n+1}, & (1.33)\\ \widehat{Q}_n \widehat{Q}_1 - Q_n R_1 &= -\widehat{Q}_{n+1}, & (1.34)\\ \widehat{Q}_n Q_1 - Q_n \widehat{R}_1 &= -Q_{n+1}, & (1.35)\\ \widehat{Q}_1 Q_n - Q_1 \widehat{R}_n &= -Q_{n+1}, & (1.36) \end{aligned}$$

Now, for $n \in \mathbb{N}_0$, we in addition define the operator-valued functions

$$W_n = R_n Q_1, \qquad \text{and} \qquad Z_n = Q_n R_1,$$

and claim that the following recursion relations hold:

$$\begin{aligned} -U_{n+1} + U_{n,x} &= W_n R_1 + R_1 Z_n, & (1.37)\\ V_{n+1} + V_{n,x} &= Z_n Q_1 + Q_1 W_n, & (1.38)\\ W_{n,x} &= U_n Q_1 + R_1 V_n, & (1.39)\\ Z_{n,x} &= Q_1 U_n + V_n R_1. & (1.40) \end{aligned}$$

Indeed (1.37)–(1.40) can be seen as follows. For (1.37), we first use the derivation rules (1.27), (1.28), and then (1.31) to find

$$\begin{aligned} U_{n,x} &= \widehat{R}_n R_{1,x} + \widehat{R}_{n,x} R_1 \ = \ (\widehat{R}_n \widehat{R}_1 + R_1 Q_n) R_1 \\ &= \Big(\widehat{R}_{n+1} + (R_n Q_1 + R_1 Q_n) \Big) R_1 \\ &= U_{n+1} + (W_n R_1 + R_1 Z_n). \end{aligned}$$

Analogously, (1.38) follows from first applying (1.29), (1.30), and then (1.34):

$$\begin{aligned} V_{n,x} &= \widehat{Q}_n Q_{1,x} + \widehat{Q}_{n,x} Q_1 \ = \ (\widehat{Q}_n \widehat{Q}_1 + Q_1 R_n) Q_1 \\ &= \Big(-\widehat{Q}_{n+1} + (Q_1 R_n + Q_n R_1) \Big) Q_1 \\ &= -V_{n+1} + (Q_1 W_n + Z_n Q_1). \end{aligned}$$

To prove (1.39) we first apply the derivation rules (1.27), (1.29), and then (1.32) and (1.33). This yields

$$\begin{aligned} W_{n,x} &= R_n Q_{1,x} + R_{n,x} Q_1 \ = \ (R_n \widehat{Q}_1 + \widehat{R}_1 R_n) Q_1 \\ &= \Big((-R_{n+1} + \widehat{R}_n R_1) + (R_{n+1} + R_1 \widehat{Q}_n) \Big) Q_1 \\ &= (\widehat{R}_n R_1 + R_1 \widehat{Q}_n) Q_1 \\ &= U_n Q_1 + R_1 V_n. \end{aligned}$$

Finally we check (1.40). Here we again need the derivation rules (1.27), (1.29), then we use (1.35) and (1.36),

$$\begin{aligned}
Z_{n,x} &= Q_{n,x}R_1 + Q_nR_{1,x} \;=\; (\widehat{Q}_1Q_n + Q_n\widehat{R}_1)R_1 \\
&= \Big((-Q_{n+1} + Q_1\widehat{R}_n) + (Q_{n+1} + \widehat{Q}_nQ_1)\Big)R_1 \\
&= (Q_1\widehat{R}_n + \widehat{Q}_nQ_1)R_1 \\
&= Q_1U_n + V_nR_1.
\end{aligned}$$

Thus we have shown the recursion relations.

Inserting (1.39), (1.40) into the recursion relation (1.37), and using Lemma 1.2.8, we obtain

$$\begin{aligned}
U_{n+1} &= U_{n,x} - \Big(\int_{-\infty}^{x} (U_nQ_1 + R_1V_n)d\xi \; R_1 + R_1 \int_{-\infty}^{x} (Q_1U_n + V_nR_1)d\xi \Big) \\
&= U_{n,x} - \Big(\int_{-\infty}^{x} (U_nQ + RV_n)d\xi \; R + R \int_{-\infty}^{x} (QU_n + V_nR)d\xi \Big),
\end{aligned}$$

where the integrals exist because W_n, Z_n decay for $x \to -\infty$ (see Lemma 1.2.9).

Analogously, inserting (1.39), (1.40) into the recursion relation (1.38), we get

$$V_{n+1} \;=\; -V_{n,x} + \Big(\int_{-\infty}^{x} (QU_n + V_nR)d\xi \; Q + Q \int_{-\infty}^{x} (U_nQ + RV_n)d\xi \Big) .$$

In summary, we have shown $\begin{pmatrix} U_{n+1} \\ V_{n+1} \end{pmatrix} = \mathcal{T}_{R,Q} \begin{pmatrix} U_n \\ V_n \end{pmatrix}$, and thus, by induction,

$$\begin{pmatrix} U_n \\ V_n \end{pmatrix} = \big(\mathcal{T}_{R,Q} \big)^n \begin{pmatrix} U_0 \\ V_0 \end{pmatrix} = \big(\mathcal{T}_{R,Q} \big)^n \begin{pmatrix} R \\ -Q \end{pmatrix}.$$

This completes the first step. □

Step 2: The strategy of the second step is essentially the same as in the first step. Now we need a second hierarchy of auxiliary operator-valued functions. For $n \in \mathbb{N}_0$, we define:

$$\begin{aligned}
T_n &= (I - LM)^{-1}\Big(\big(A^nf_0(A) \big) L - L \big((-B)^nf_0(-B) \big) \Big), \\
\widehat{T}_n &= (I - LM)^{-1}\Big(\big(A^nf_0(A) \big) - L \big((-B)^nf_0(-B) \big) M \Big), \\
S_n &= (I - ML)^{-1}\Big(M \big(A^nf_0(A) \big) - \big((-B)^nf_0(-B) \big) M \Big), \\
\widehat{S}_n &= (I - ML)^{-1}\Big(M \big(A^nf_0(A) \big) L - \big((-B)^nf_0(-B) \big) \Big).
\end{aligned}$$

In addition we define

$$\widetilde{U}_n = \widehat{T}_nR_1, \qquad \text{and} \qquad \widetilde{V}_n = \widehat{S}_nQ_1.$$

Then, for $n \in \mathbb{N}_0$, we claim that

$$\big(\mathcal{T}_{R,Q} \big)^n \begin{pmatrix} R_t \\ Q_t \end{pmatrix} = \begin{pmatrix} \widetilde{U}_n \\ \widetilde{V}_n \end{pmatrix}. \tag{1.41}$$

Proof of Step 2: Again the differentiation rules for the operator-valued functions T_n, $\widehat{T}_n$, S_n, and $\widehat{S}_n$ can be taken from Lemma 1.2.5 (here with $A_n = A^nf_0(A)$, $B_n = -(-B)^nf_0(-B)$). As a result, it holds

$$T_{n,x} \;=\; \widehat{R}_1T_n, \tag{1.42}$$

$$\widehat{T}_{n,x} \;=\; R_1S_n, \tag{1.43}$$

$$
\begin{aligned}
S_{n,x} &= \widehat{Q}_1 S_n, & (1.44)\\
\widehat{S}_{n,x} &= Q_1 T_n. & (1.45)
\end{aligned}
$$

The necessary rules for the evaluation of certain products of the T_n, $\widehat{T}_n$, S_n, and $\widehat{S}_n$ are taken from Lemma 1.2.6 (with the same choice for A_n, B_n as before), yielding

$$
\begin{aligned}
\widehat{T}_n\widehat{R}_1 - T_nQ_1 &= \widehat{T}_{n+1}, & (1.46)\\
\widehat{T}_nR_1 - T_n\widehat{Q}_1 &= T_{n+1}, & (1.47)\\
\widehat{R}_1T_n - R_1\widehat{S}_n &= T_{n+1}, & (1.48)\\
\widehat{S}_n\widehat{Q}_1 - S_nR_1 &= -\widehat{S}_{n+1}, & (1.49)\\
\widehat{S}_nQ_1 - S_n\widehat{R}_1 &= -S_{n+1}, & (1.50)\\
\widehat{Q}_1S_n - Q_1\widehat{T}_n &= -S_{n+1}. & (1.51)
\end{aligned}
$$

Next we introduce, for $n \in \mathbb{N}_0$, the operator-valued functions

$$
\widetilde{W}_n = T_nQ_1, \qquad \text{and} \qquad \widetilde{Z}_n = S_nR_1.
$$

and derive the following recursion relations:

$$
\begin{aligned}
-\widetilde{U}_{n+1} + \widetilde{U}_{n,x} &= \widetilde{W}_nR_1 + R_1\widetilde{Z}_n, & (1.52)\\
\widetilde{V}_{n+1} + \widetilde{V}_{n,x} &= \widetilde{Z}_nQ_1 + Q_1\widetilde{W}_n, & (1.53)\\
\widetilde{W}_{n,x} &= \widetilde{U}_nQ_1 + R_1\widetilde{V}_n & (1.54)\\
\widetilde{Z}_{n,x} &= Q_1\widetilde{U}_n + \widetilde{V}_nR_1, & (1.55)
\end{aligned}
$$

To prove these recursion relations we proceed as follows. As for (1.52), using the derivation rules (1.27), (1.43), and then (1.46), we see

$$
\begin{aligned}
\widetilde{U}_{n,x} &= \widehat{T}_nR_{1,x} + \widehat{T}_{n,x}R_1 \;=\; (\widehat{T}_n\widehat{R}_1 + R_1S_n)R_1\\
&= \Big(\widehat{T}_{n+1} + (T_nQ_1 + R_1S_n)\Big)R_1\\
&= \widetilde{U}_{n+1} + (\widetilde{W}_nR_1 + R_1\widetilde{Z}_n).
\end{aligned}
$$

Accordingly, (1.53) follows from the derivation rules (1.29), (1.45), and then the rule (1.49),

$$
\begin{aligned}
\widetilde{V}_{n,x} &= \widehat{S}_nQ_{1,x} + \widehat{S}_{n,x}Q_1 \;=\; (\widehat{S}_n\widehat{Q}_1 + Q_1T_n)Q_1\\
&= \Big(-\widehat{S}_{n+1} + (S_nR_1 + Q_1T_n)\Big)Q_1\\
&= -\widetilde{V}_{n+1} + (Q_1\widetilde{W}_n + \widetilde{Z}_nQ_1).
\end{aligned}
$$

To see (1.54), we apply the derivation rules (1.29), (1.42), and subsequently use (1.47) and (1.48). This yields

$$
\begin{aligned}
\widetilde{W}_{n,x} &= T_nQ_{1,x} + T_{n,x}Q_1 \;=\; (T_n\widehat{Q}_1 + \widehat{R}_1T_n)Q_1\\
&= \Big((-T_{n+1} + \widehat{T}_nR_1) + (T_{n+1} + R_1\widehat{S}_n)\Big)Q_1\\
&= (\widehat{T}_nR_1 + R_1\widehat{S}_n)Q_1\\
&= R_1\widetilde{V}_n + \widetilde{U}_nQ_1.
\end{aligned}
$$

Finally, (1.55) follows from the derivation rules (1.27), (1.44), and then the rules (1.50) and (1.51),

$$
\begin{aligned}
\widetilde{Z}_{n,x} &= S_{n,x}R_1 + S_nR_{1,x} \;=\; (\widehat{Q}_1S_n + S_n\widehat{R}_1)R_1\\
&= \Big((-S_{n+1} + Q_1\widehat{T}_n) + (S_{n+1} + \widehat{S}_nQ_1)\Big)R_1\\
&= (Q_1\widehat{T}_n + \widehat{S}_nQ_1)R_1\\
&= Q_1\widetilde{U}_n + \widetilde{V}_nR_1.
\end{aligned}
$$

Thus the recursion relations are shown.

Inserting (1.54), (1.55) into the recursion relations (1.52), (1.53), and using Lemma 1.2.8, the same argument as in the first step yields

$$\begin{aligned}
\widetilde{U}_{n+1} &= \widetilde{U}_{n,x} - \Big(\int_{-\infty}^{x} (\widetilde{U}_n Q + R\widetilde{V}_n) d\xi \, R + R \int_{-\infty}^{x} (Q\widetilde{U}_n + \widetilde{V}_n R) d\xi \Big) \\
\widetilde{V}_{n+1} &= -\widetilde{V}_{n,x} + \Big(\int_{-\infty}^{x} (Q\widetilde{U}_n + \widetilde{V}_n R) d\xi \, Q + Q \int_{-\infty}^{x} (\widetilde{U}_n Q + R\widetilde{V}_n) d\xi \Big)
\end{aligned}$$

(note that $\widetilde{W}_n$ $\widetilde{Z}_n$ decay for $x \to -\infty$ by Lemma 1.2.9, which guarantees the existence of the integrals).

In summary, $\begin{pmatrix} \widetilde{U}_{n+1} \\ \widetilde{V}_{n+1} \end{pmatrix} = \mathcal{T}_{R,Q} \begin{pmatrix} \widetilde{U}_n \\ \widetilde{V}_n \end{pmatrix}$, which by induction implies

$$\begin{pmatrix} \widetilde{U}_n \\ \widetilde{V}_n \end{pmatrix} = \big(\mathcal{T}_{R,Q} \big)^n \begin{pmatrix} \widetilde{U}_0 \\ \widetilde{V}_0 \end{pmatrix} = \big(\mathcal{T}_{R,Q} \big)^n \begin{pmatrix} R_t \\ Q_t \end{pmatrix},$$

where we have used $\widetilde{U}_0 = \widehat{T}_0 R_1 = R_t$ and $\widetilde{V}_0 = \widehat{S}_0 Q_1 = Q_t$ (confer Lemma 1.2.7).

This completes the second step. □

Step 3: To conclude the proof, we assume that the polynomials f, g are concretely given by

$$f(z) = \sum_{n=0}^{N} a_n z^n, \qquad g(z) = \sum_{n=0}^{N} b_n z^n.$$

To have the same order for both polynomials we allow leading coefficients to vanish. Then the connection between U_n and $\widetilde{U}_n$ is given by

$$\begin{aligned}
\sum_{n=0}^{N} b_n \widetilde{U}_n &= \Big(\sum_{n=0}^{N} b_n \widehat{T}_n \Big) R_1 \\
&= (I - LM)^{-1} \Big(\sum_{n=0}^{N} b_n \Big(\big(A^n f_0(A) \big) - L \big((-B)^n f_0(-B) \big) M \Big) \Big) R_1 \\
&= (I - LM)^{-1} \Big(\big(g(A) f_0(A) \big) - L \big(g(-B) f_0(-B) M \big) \Big) R_1 \\
&= (I - LM)^{-1} \Big(f(A) - L f(-B) M \Big) R_1 \\
&= (I - LM)^{-1} \Big(\sum_{n=0}^{N} a_n (A^n - L(-B)^n M) \Big) R_1 \\
&= \Big(\sum_{n=0}^{N} a_n \widehat{R}_n \Big) R_1 \\
&= \sum_{n=0}^{N} a_n U_n, \qquad (1.56)
\end{aligned}$$

and, analogously,

$$\sum_{n=0}^{N} b_n \widetilde{V}_n = \sum_{n=0}^{N} a_n V_n. \qquad (1.57)$$

As a result,

$$\begin{aligned}
g\,(\,\mathcal{T}_{R,Q}\,)\begin{pmatrix} R_t \\ Q_t \end{pmatrix} &= \sum_{n=0}^{N} b_n\,(\,\mathcal{T}_{R,Q}\,)^n \begin{pmatrix} R_t \\ Q_t \end{pmatrix} \overset{(1.41)}{=} \sum_{n=0}^{N} b_n \begin{pmatrix} \widetilde{U}_n \\ \widetilde{V}_n \end{pmatrix} \\
&\overset{(1.56),(1.57)}{=} \sum_{n=0}^{N} a_n \begin{pmatrix} U_n \\ V_n \end{pmatrix} \overset{(1.26)}{=} \sum_{n=0}^{N} a_n\,(\,\mathcal{T}_{R,Q}\,)^n \begin{pmatrix} R \\ -Q \end{pmatrix} \\
&= f\,(\,\mathcal{T}_{R,Q}\,)\begin{pmatrix} R \\ -Q \end{pmatrix}.
\end{aligned}$$

This completes the proof. □

Remark 1.2.10. *One major difference between our treatment and [10] is that the latter authors work directly with a meromorphic function $f_0 = f/g$. This leads to serious difficulties concerning invertibility (or more general concerning the holomorphic function calculus on non Banach spaces) for $\mathcal{T}_{R,Q}$. In fact, a major part of the arguments in [10] is only formal.*

That is also the reason why we restricted to polynomials f, g. The point is that we only need iterations of $\mathcal{T}_{R,Q}$ up to a certain finite degree. In all known applications polynomials are sufficient.

1.2.1 Non-abelian versions of the prototypical equations

Transition from Theorem 1.2.1 to soliton equations is accomplished in two steps. First one needs an appropriate choice of f_0 determining the coupled system under consideration. Secondly, one reduces to a single equation by adding a linear relation between R and Q.

Here we will perform the first step. The second step will be postponed because it requires (at least for the $\mathbb{C}$-reduction) further choices and becomes only transparent when considered in the context of scalarization. But we will briefly discuss the difficulties which have to be overcome (see Remarks 1.2.12, 1.2.14).

The non-abelian Nonlinear Schrödinger equation

As in the scalar case we choose $f_0(z) = -\mathrm{i}z^2$ (confer Section 1.1.1). Then

$$\mathcal{T}_{R,Q}\begin{pmatrix} R \\ -Q \end{pmatrix} = \begin{pmatrix} R_x - \Big(R\int_{-\infty}^{x}(QR-QR)\,d\xi + \int_{-\infty}^{x}(RQ-RQ)\,d\xi\,R\Big) \\ Q_x + \Big(Q\int_{-\infty}^{x}(RQ-RQ)\,d\xi + \int_{-\infty}^{x}(QR-QR)\,d\xi\,Q\Big) \end{pmatrix} = \begin{pmatrix} R_x \\ Q_x \end{pmatrix},$$

and

$$\begin{aligned}
(\,\mathcal{T}_{R,Q}\,)^2\begin{pmatrix} R \\ -Q \end{pmatrix} &= \mathcal{T}_{R,Q}\begin{pmatrix} R_x \\ Q_x \end{pmatrix} \\
&= \begin{pmatrix} R_{xx} - \Big(R\int_{-\infty}^{x}(QR_x+Q_xR)\,d\xi + \int_{-\infty}^{x}(R_xQ+RQ_x)\,d\xi\,R\Big) \\ -Q_{xx} + \Big(Q\int_{-\infty}^{x}(R_xQ+RQ_x)\,d\xi + \int_{-\infty}^{x}(QR_x+Q_xR)\,d\xi\,Q\Big) \end{pmatrix} \\
&= \begin{pmatrix} R_{xx} - \Big(R\int_{-\infty}^{x}(QR)_x\,d\xi + \int_{-\infty}^{x}(RQ)_x\,d\xi\,R\Big) \\ -Q_{xx} + \Big(Q\int_{-\infty}^{x}(RQ)_x\,d\xi + \int_{-\infty}^{x}(QR)_x\,d\xi\,Q\Big) \end{pmatrix} \\
&= \begin{pmatrix} R_{xx} - 2RQR \\ -Q_{xx} + 2QRQ \end{pmatrix},
\end{aligned}$$

and the associated system (1.8) becomes

$$-\mathrm{i}R_t + R_{xx} - 2RQR = 0, \tag{1.58}$$
$$\mathrm{i}Q_t + Q_{xx} - 2QRQ = 0. \tag{1.59}$$

By Theorem 1.2.1, we obtain an explicit solution of (1.58), (1.59) as follows.

Proposition 1.2.11. *Let E be a Banach space and $A,\ B \in \mathcal{L}(E)$.*

Assume that $L = L(x,t)$, $M = M(x,t) \in \mathcal{L}(E)$ are operator-valued functions which, on a strip $\mathbb{R}\times(t_1,t_2)$, are sufficiently smooth and behave sufficiently well for $x \to -\infty$, and solve the base equations

$$L_x = AL, \quad L_t = -\mathrm{i}A^2L, \qquad \text{and} \qquad M_x = BM, \quad M_t = \mathrm{i}B^2M.$$

Assume that $(I+LM)$, $(I+ML)$ are invertible on $\mathbb{R}\times(t_1,t_2)$.

Then $R = (I+LM)^{-1}(AL+LB)$, $Q = -(I+ML)^{-1}(BM+MA)$, solve the operator-valued NLS system (1.58), (1.59) on $\mathbb{R}\times(t_1,t_2)$.

Proof The proof is a straightforward application of Theorem 1.2.1 with an additional sign added to the operator-function M. □

Remark 1.2.12. *a) It is instructive to restrict Proposition 1.2.11 to the scalar setting. Then the solution reduces to the one-soliton for $b = \overline{a}$ (see Section 1.1.1 for the precise formula). In contrast, the condition $b = a$ would lead to*

$$q(x,t) = -a\, e^{\mathrm{i}a^2 t}\cosh^{-1}(ax) \qquad \text{with } a \text{ real},$$

which means that one parameter of the one-soliton, namely its velocity, is lost. Thus, already from this simple example it becomes clear that it is essential to have two operator parameters A, B in Theorem 1.2.1, and not one as in [10].

b) The scalar NLS belongs to the $\mathbb{C}$-reduced AKNS system, which means that the linear relation $r = -\overline{q}$ is imposed. Note that there is no straightforward way to express $R = -\overline{Q}$ on a general Banach space. We will fix this later by working on sequence spaces where complex conjugation has a natural interpretation.

The non-abelian modified Korteweg-de Vries equation

For the modified Korteweg-de Vries equation, $f_0(z) = -z^3$ (confer Section 1.1.1). With this choice (1.8) reads

$$R_t + R_{xxx} - 3\,(\,RQR_x + R_xQR\,) = 0,$$
$$Q_t + Q_{xxx} - 3\,(\,QRQ_x + Q_xRQ\,) = 0.$$

Imposing the linear relation $R = -Q$ yields the operator-valued mKdV

$$Q_t + Q_{xxx} + 3\,(\,Q^2Q_x + Q_xQ^2\,) = 0. \tag{1.60}$$

An explicit solution of (1.60) is given by the next proposition.

Proposition 1.2.13. *Let E be a Banach space and $A \in \mathcal{L}(E)$.*

Assume that $L = L(x,t) \in \mathcal{L}(E)$ is an operator-valued function which, on a strip $\mathbb{R}\times(t_1,t_2)$, is sufficiently smooth and behaves sufficiently well for $x \to -\infty$, and solves the base equations

$$L_x = AL \qquad \text{and} \qquad L_t = -A^3L.$$

Assume that $(I+L^2)$ is invertible on $\mathbb{R}\times(t_1,t_2)$.

Then $Q = -(I+L^2)^{-1}(AL+LA)$ solves the operator-valued mKdV equation (1.60) on $\mathbb{R}\times(t_1,t_2)$.

Proof To arrange $R = -Q$ in Theorem 1.2.1, we set $B = A$ and $M = -L$. Then the assertion is an immediate consequence of Theorem 1.2.1. $\square$

Together with the operator-valued mKdV (1.60), the operator-valued Korteweg-de Vries equation (KdV)

$$U_t + U_{xxx} + 3\,(\,UU_x + U_xU\,) = 0$$

has been treated already in the author's thesis [88]. We want to mention that in this context also the Miura transformation, linking mKdV and KdV, and their discretizations (Langmuir and Wadati lattices together with the coninuum approximation) have been generalized to the operator-level in [88].

Remark 1.2.14. *The scalar mKdV (and also the derivative sG in the next example) belong to the $\mathbb{R}$-reduced AKNS system, where the linear relation $r = -q$ is imposed. The condition $R = -Q$ generalizes naturally to the operator-level. But now the requirement that the solution be real admits no canonical interpretation in general Banach spaces.*

The non-abelian sine-Gordon equation

As in Section 1.1.1, we have $f_0(z) = 1/z$. Here (1.8) becomes

$$\begin{aligned} R_{tx} - \left(R \int_{-\infty}^{x} (QR)_t \, d\xi + \int_{-\infty}^{x} (RQ)_t \, d\xi \, R \right) - R &= 0, \\ Q_{tx} - \left(Q \int_{-\infty}^{x} (RQ)_t \, d\xi + \int_{-\infty}^{x} (QR)_t \, d\xi \, Q \right) - Q &= 0. \end{aligned}$$

With the relation $R = -Q$, we obtain the following non-abelian version of the derivative sine-Gordon equation

$$Q_{tx} + \left(Q \int_{-\infty}^{x} (Q^2)_t \, d\xi + \int_{-\infty}^{x} (Q^2)_t \, d\xi \, Q \right) - Q = 0. \tag{1.61}$$

Again an explicit solution to (1.61) is given by Theorem 1.2.1.

Proposition 1.2.15. *Let E be a Banach space and $A \in \mathcal{L}(E)$ invertible.*

Assume that $L = L(x,t) \in \mathcal{L}(E)$ is an operator-valued function which, on a strip $\mathbb{R} \times (t_1, t_2)$, is sufficiently smooth and behaves sufficiently well for $x \to -\infty$, and solves the base equations

$$L_x = AL \qquad \text{and} \qquad L_t = A^{-1}L.$$

Assume that $(I + L^2)$ is invertible on $\mathbb{R} \times (t_1, t_2)$.

Then $Q = -(I + L^2)^{-1}(AL + LA)$ solves the operator-valued derivative sG (1.61) on $\mathbb{R} \times (t_1, t_2)$.

Proof To achieve $R = -Q$ we set $B = A$, $M = -L$ in Theorem 1.2.1. The assertion then follows by Theorem 1.2.1. $\square$

Also the sine-Gordon equation in its usual form, $u_{xt} = \sin(u)$, has been successfully treated by the operator-method in [88] (see also [90]).

Chapter 2

From the non-abelian to the scalar AKNS system: Extracting solution formulas

In this chapter we will go back the way from the operator solutions constructed in the previous chapter to the scalar setting. Scalarization is easier to explain for an individual equation: Then the operator solution is a family of endomorphisms, and one can try to descent by applying an appropriate functional τ, typically the trace. As the solution property is to be preserved, one quickly sees that τ has to enjoy certain multiplicativity properties, which cannot be valid for arbitrary operators. But it can be guaranteed for operator-functions taking values in $\mathcal{S}_a$, the space of one-dimensional operators with a fixed kernel (prescribed by a functional a). In Section 2.1.2 it is explained how scalarization looks like for an individual equation. We mention that the construction of $\mathcal{S}_a$-valued solutions is closely related to the use of projectors to one-dimensional subspaces. In the Hilbert space setting the two approaches are more or less equivalent (see Marchenko [55]). In general Banach space projections to subspaces do not always come with nice expressions and our dual approach seems preferable.

For the general AKNS system scalarization becomes harder because one starts from a pair of operator solutions $Q \in \mathcal{L}(E,F)$, $R \in \mathcal{L}(F,E)$, mapping between different Banach spaces E, F. In particular we cannot simply apply traces. But we can still choose Q, R with values in spaces $\mathcal{S}_a$, $\mathcal{S}_b$, $a \in E'$, $b \in F'$, and scalarize by cross evaluation (see Theorem 2.2.1). This was partly inspired by related techniques appearing in [13].

The formulas determined by cross evaluation still contain inverse operators, which are extremely embarassing in applications. Our next step is to eliminate inverses by transition to determinants. During the calculations one needs to apply traces also to expressions which are no longer of finite rank. Hence we need extensions of the elementary trace on the finite rank operators, a topic which is exhaustively treated in the theorey of traces and determinants on quasi-Banach operator ideals (see [41], [73]). The final result, after some additional refinements, is recorded in Theorem 2.4.4. It will be the basis of all later applications.

But before we have to clarify how the one-dimensionality conditions can be met. Looking at the explicit expressions of the operator solutions in Theorem 1.2.1, we see that we have to make the elementary expression $AX + XB$ one-dimensional for given operators $A \in \mathcal{L}(E)$, $B \in \mathcal{L}(F)$. Under the condition $0 \notin \operatorname{spec}(A) + \operatorname{spec}(B)$, one-dimensionality can be settled by means of the fundamental theorem of Eschmeier [28] and Dash/Schechter [22] (and a refinement by Aden [7]).

From the viewpoint of our applications in Chapter 7, the condition $0 \notin \operatorname{spec}(A)+\operatorname{spec}(B)$

appears to be relatively massive. We will see in Chapter 7 that there is a subtler way to produce one-dimensionality by replacing the Eschmeier and Dash/Schechter theorem by factorization techniques in the spirit of the Grothendieck theorem.

Finally we obtain an amelioration for the $\mathbb{R}$-reduced AKNS system (containing in particular the modified Korteweg-de Vries and the sine-Gordon equations) where one can reduce the size of the determinants by half. Furthermore, we provide, for the Nonlinear Schrödinger and the modified Korteweg-de Vries equations, solution formulas without any growth condition. Note that growth conditions are always necessary for the treatment of the general AKNS system to ensure existence of the integral operator $\mathcal{T}_{R,Q}$.

2.1 Background for scalarization

We will first introduce some terminology concerning duality and one-dimensional operators in Banach spaces. Then we recall, for the sake of motivation, known scalarization techniques in the simplified setting $E = F$.

2.1.1 Algebraic terminology

For bounded linear functionals $a \in E'$ (E' the dual Banach space of E) we write

$$a(x) = \langle x, a \rangle, \qquad x \in E.$$

A one-dimensional operator is an operator $T \in \mathcal{L}(E,F)$ with one-dimensional range. Every such operator can be written as $a \otimes y$ with appropriate $a \in E'$, $y \in F$, where the map $a \otimes y$ is defined by

$$a \otimes y(x) = \langle x, a \rangle y.$$

For convenience, we note some frequently used calculation rules for one-dimensional operators. Let $a \in E'$, $y \in F$. Then

$$T(a \otimes y)S = (S'a) \otimes (Ty) \qquad \forall S \in \mathcal{L}(E_1, E), T \in \mathcal{L}(F, F_1).$$

In particular,

$$(a \otimes y)(b \otimes x) \quad = \quad \langle x, a \rangle b \otimes y$$

for $a \in E'$, $b \in E_1'$, $x \in E$, and $y \in F$.

A finite-rank operator is an operator $T \in \mathcal{L}(E,F)$ with finite-dimensional range. We set $\operatorname{rank}(T) = \dim(\operatorname{ran}(T))$. Thus $\mathcal{F}(E,F)$, the set of all finite-rank operators, is a (not necessarily closed) subspace of $\mathcal{L}(E,F)$. It is easily verified that any operator of rank N can be written as $T = \sum_{j=1}^{N} a_j \otimes y_j$ with appropriate $a_j \in E'$, $y_j \in F$.

Definition 2.1.1. *For a nonzero $a \in E'$, we define the vector space*

$$\mathcal{S}_a(E,F) = \Big\{ a \otimes y \;\Big|\; y \in F \Big\},$$

and we write $\mathcal{S}_a(E)$ instead of $\mathcal{S}_a(E,E)$.

This means that $\mathcal{S}_a(E,F)$ consists of all one-dimensional operators between E and F whose behaviour is governed by the functional $a \in E'$. Equivalently, one could define $\mathcal{S}_a(E,F)$ as the space of all $T \in \mathcal{L}(E,F)$ with $\ker(a) \subset \ker(T)$. In particular we have $\mathcal{S}_a(E) = \mathcal{S}_{\widehat{a}}(E)$ if and only if a and $\widehat{a}$ are linearly dependent.

Lemma 2.1.2. *$\mathcal{S}_a(E,F)$ is closed under multiplication from the left with operators in $\mathcal{L}(F)$. In particular $\mathcal{S}_a(E)$ is a left ideal in $\mathcal{L}(E)$.*

Moreover, $\mathcal{S}_a(E)$ is a Banach algebra.

There is a canonical continuous linear functional on $\mathcal{S}_a(E)$, the evaluation functional ev_a, defined by

$$\mathrm{ev}_a(a \otimes y) = \langle y, a \rangle. \tag{2.1}$$

It is easily verified that ev_a remains unchanged, if we replace a by $\widehat{a}$ with $\mathcal{S}_a(E) = \mathcal{S}_{\widehat{a}}(E)$.

The following lemma shows that ev_a is even a continuous algebra homomorphism.

Proposition 2.1.3. *On $\mathcal{S}_a(E)$, the functional ev_a is multiplicative.*

Proof Let T, $S \in \mathcal{S}_a(E)$ be operators governed by the same functional $a \in E'$, say $T = a \otimes y$, $S = a \otimes z$ with y, $z \in E$. Since $TS = \langle z, a \rangle a \otimes y$,

$$\begin{aligned} \mathrm{ev}_a(TS) &= \mathrm{ev}_a\left(\langle z, a \rangle a \otimes y\right) = \langle z, a \rangle \, \mathrm{ev}_a(a \otimes y) = \langle z, a \rangle \langle y, a \rangle \\ &= \mathrm{ev}_a(a \otimes z) \, \mathrm{ev}_a(a \otimes y) = \mathrm{ev}_a(S) \, \mathrm{ev}_a(T). \end{aligned}$$

□

2.1.2 Scalarization for endomorphisms

For the sake of motivation we explain the idea of the scalarization process in the simplified setting of endomorphisms. As model equation we take the modified Korteweg-de Vries equation (1.6).

Assume that $Q = Q(x,t) \in \mathcal{L}(E)$ is an operator-solution of the non-abelian mKdV (1.60). A natural ansatz to derive scalar solutions $q = \tau(Q)$ for (1.6) is to apply a continuous linear functional τ to the operator solution Q.

By linearity of τ,

$$\begin{aligned} 0 &= \tau\Big(Q_t + Q_{xxx} + 3(Q^2 Q_x + Q_x Q^2)\Big) \\ &= \tau(Q_t) + \tau(Q_{xxx}) + 3\,\left(\,\tau(Q^2 Q_x) + \tau(Q_x Q^2)\,\right). \end{aligned}$$

At this point we would like to continue by

$$\begin{aligned} &= \tau(Q_t) + \tau(Q_{xxx}) + 6\tau(Q)^2 \tau(Q_x) \\ &= \tau(Q)_t + \tau(Q)_{xxx} + 6\tau(Q)^2 \tau(Q)_x, \end{aligned}$$

the latter by continuity of τ. Then $q = \tau(Q)$ would indeed be a solution of (1.6).

The above calculation shows that, in order to maintain the solution property under scalarization, the nonlinearity of the mKdV enforces multiplicativity of the functional τ in some sense.

This requirement can be satisfied in the following way:

1. We assume that the operator solution Q belongs to $\mathcal{S}_a(E)$ for a constant $a \in E'$. In other words, Q is one-dimensional with fixed kernel.

2. For scalarization, we use the functional ev_a on $\mathcal{S}_a(E)$ defined by $\mathrm{ev}_a(a \otimes y) = \langle y, a \rangle$. Then multiplicativity of ev_a on $\mathcal{S}_a(E)$ is guaranteed by Proposition 2.1.3.

Observe that ev_a is the restriction of the trace, defined for endomorphisms of finite rank, to the $\mathcal{S}_a(E)$. Actually, earlier work of Aden, Carl, and the author (see for example [8], [18]) focused on the particular role of the trace on quasi-Banach operator ideals. But the treatment becomes more transparent if one uses this theory only for the later improvements of the solution formulas.

For the AKNS system we encounter the difficulty that the operator solution in general acts between different Banach spaces. In particular, the definition of evaluation functionals does not even make sense in this case. In the next section we explain how to overcome this problem.

2.2 The scalarization process for operators between different Banach spaces

In this section we explain how to use the non-abelian AKNS system to construct solutions of the scalar AKNS system. We start with an operator solution $Q = Q(x,t) \in \mathcal{L}(E,F)$, $R = R(x,t) \in \mathcal{L}(F,E)$ where E, F are possibly different Banach spaces.

We adhere to the idea to choose the operator functions Q, R one-dimensional with fixed kernel, say

$$Q = a \otimes d(x,t) \in \mathcal{S}_a(E,F), \qquad R = b \otimes c(x,t) \in \mathcal{S}_b(F,E),$$

for fixed, constant and non-vanishing functionals $a \in E'$, $b \in F'$. It is clear that we cannot apply $a \in E'$ to the vector function $d = d(x,t) \in F$ since they do not match. In contrast, it is a natural ansatz to try cross-evaluation,

$$q(x,t) := \langle d(x,t), b \rangle, \qquad r(x,t) := \langle c(x,t), a \rangle.$$

Note that r is given in terms of $c = c(x,t)$, which encodes the information from R.

Theorem 2.2.1. *Let E, F be Banach spaces and $Q = Q(x,t) \in \mathcal{L}(E,F)$, $R = R(x,t) \in \mathcal{L}(F,E)$ operator-valued functions which solve the non-abelian AKNS system (1.8).*

If, in addition, there exist constant functionals $0 \neq a \in E'$, $0 \neq b \in F'$, and vector-functions $c = c(x,t) \in E$, $d = d(x,t) \in F$ such that

$$Q(x,t) = a \otimes d(x,t), \qquad R(x,t) = b \otimes c(x,t),$$

then a solution of the scalar AKNS system (1.1) is given by

$$q(x,t) = \langle d(x,t), b \rangle, \qquad r(x,t) = \langle c(x,t), a \rangle.$$

When supposing that R, Q solve the AKNS system, it is understood that R, Q are sufficiently smooth and behave sufficiently well for $x \to -\infty$.

Proof First we compute the effect of n-fold iteration of $\mathcal{T}_{R,Q}$ when evaluated on operator functions $U \in \mathcal{S}_b(F,E)$, $V \in \mathcal{S}_a(E,F)$. Let $U = b \otimes \widehat{c}$, $V = a \otimes \widehat{d}$, with vector functions $\widehat{c} = \widehat{c}(x,t) \in E$, $\widehat{d} = \widehat{d}(x,t) \in F$. As usual we assume U, V to be sufficiently smooth and sufficiently well behaved for $x \to -\infty$.

Claim 1: For $n \in \mathbb{N}_0$, set

$$\begin{pmatrix} U_n \\ V_n \end{pmatrix} = (\ \mathcal{T}_{R,Q}\ )^n \begin{pmatrix} U \\ V \end{pmatrix}.$$

a) Then $U_n \in \mathcal{S}_b(F,E)$, $V_n \in \mathcal{S}_a(E,F)$ for all $n \in \mathbb{N}_0$.

In particular, there exist vector functions $\widehat{c}_n = \widehat{c}_n(x,t) \in E$, $\widehat{d}_n = \widehat{d}_n(x,t) \in F$ such that $U_n = b \otimes \widehat{c}_n$, $V_n = a \otimes \widehat{d}_n$.

b) Assigning to U_n, V_n the scalar functions $u_n = \langle \widehat{c}_n, a\rangle$, $v_n = \langle \widehat{d}_n, b\rangle$, the corresponding scalar relation holds, i.e.,

$$\begin{pmatrix} u_n \\ v_n \end{pmatrix} = (\ T_{r,q}\ )^n \begin{pmatrix} u \\ v \end{pmatrix},$$

where $u := \langle \widehat{c}, a\rangle$, $v := \langle \widehat{d}, b\rangle$ are the scalar functions assigned to U, V.

Proof of Claim 1: For $n = 0$ the assertion is trivial. Assume now the assertion for a certain $n \in \mathbb{N}_0$.

Then there exist vector functions $\widehat{c}_n = \widehat{c}_n(x,t) \in E$, $\widehat{d}_n = \widehat{d}_n(x,t) \in F$ such that $U_n = b \otimes \widehat{c}_n$ and $V_n = a \otimes \widehat{d}_n$. Since the functional b is constant,

$$\begin{aligned} QU_n + V_nR &= (a\otimes d)(b\otimes \widehat{c}_n) + (a\otimes \widehat{d}_n)(b\otimes c) \ = \ b\otimes \Big(\langle \widehat{c}_n, a\rangle d + \langle c, a\rangle \widehat{d}_n\Big) \\ \Longrightarrow \qquad & \int_{-\infty}^{x} (QU_n + V_nR)d\xi \ = \ b\otimes \int_{-\infty}^{x} \Big(\langle \widehat{c}_n, a\rangle d + \langle c, a\rangle \widehat{d}_n\Big) d\xi \ =: \ b\otimes f_n. \end{aligned}$$

Therefore, we obtain

$$\begin{aligned} R\int_{-\infty}^{x} (QU_n + V_nR)d\xi &= (b\otimes c)(b\otimes f_n) = b\otimes \Big(\langle f_n, b\rangle c\Big), \\ \int_{-\infty}^{x} (QU_n + V_nR)d\xi\, Q &= (b\otimes f_n)(a\otimes d) = a\otimes \Big(\langle d, b\rangle f_n\Big). \end{aligned}$$

Analogously, the constancy of the functional a yields

$$\begin{aligned} U_nQ + RV_n &= (b\otimes \widehat{c}_n)(a\otimes d) + (b\otimes c)(a\otimes \widehat{d}_n) \ = \ a\otimes \Big(\langle d, b\rangle \widehat{c}_n + \langle \widehat{d}_n, b\rangle c\Big) \\ \Longrightarrow \qquad & \int_{-\infty}^{x} (U_nQ + RV_n)d\xi \ = \ a\otimes \int_{-\infty}^{x} \Big(\langle d, b\rangle \widehat{c}_n + \langle \widehat{d}_n, b\rangle c\Big) d\xi \ =: \ a\otimes g_n, \end{aligned}$$

which implies

$$\begin{aligned} Q\int_{-\infty}^{x} (U_nQ + RV_n)d\xi &= (a\otimes d)(a\otimes g_n) = a\otimes \Big(\langle g_n, a\rangle d\Big), \\ \int_{-\infty}^{x} (U_nQ + RV_n)d\xi\, R &= (a\otimes g_n)(b\otimes c) = b\otimes \Big(\langle c, a\rangle g_n\Big). \end{aligned}$$

As a consequence,

$$\begin{pmatrix} U_{n+1} \\ V_{n+1} \end{pmatrix} = \mathcal{T}_{R,Q}\begin{pmatrix} U_n \\ V_n \end{pmatrix} = \begin{pmatrix} b\otimes \Big[\quad \widehat{c}_{n,x} - \Big(\langle f_n, b\rangle c + \langle c, a\rangle g_n\Big)\Big] \\ a\otimes \Big[-\widehat{d}_{n,x} + \Big(\langle g_n, a\rangle d + \langle d, b\rangle f_n\Big)\Big] \end{pmatrix}.$$

Hence $U_{n+1} = b\otimes \widehat{c}_{n+1} \in \mathcal{S}_b(F,E)$, $V_{n+1} = a\otimes \widehat{d}_{n+1} \in \mathcal{S}_a(E,F)$, with the vector functions $\widehat{c}_{n+1} = \widehat{c}_{n,x} - \big(\ \langle f_n, b\rangle c + \langle c, a\rangle g_n\ \big)$, $\widehat{d}_{n+1} = -\widehat{d}_{n,x} + \big(\ \langle g_n, a\rangle d + \langle d, b\rangle f_n\ \big)$. Part a) of the assertion is proved by induction.

Next observe

$$\begin{aligned} &\langle f_n, b\rangle + \langle g_n, a\rangle = \\ &= \ \Big\langle \int_{-\infty}^{x} \Big[\langle \widehat{c}_n, a\rangle d + \langle c, a\rangle \widehat{d}_n\Big] d\xi, b\Big\rangle + \Big\langle \int_{-\infty}^{x} \Big[\langle d, b\rangle \widehat{c}_n + \langle \widehat{d}_n, b\rangle c\Big] d\xi, a\Big\rangle \\ &= \ 2\int_{-\infty}^{x} \Big(\langle \widehat{c}_n, a\rangle\langle d, b\rangle + \langle c, a\rangle\langle \widehat{d}_n, b\rangle\Big) d\xi \\ &= \ 2\int_{-\infty}^{x} \big(\ u_nq + rv_n\ \big)\, d\xi. \end{aligned}$$

Inserting this identity, we find

$$
\begin{aligned}
u_{n+1} &= \langle \widehat{c}_{n+1}, a \rangle \;=\; \Big\langle \widehat{c}_{n,x} - \big(\, \langle f_n, b \rangle c + \langle c, a \rangle g_n \,\big), a \Big\rangle \\
&= \langle \widehat{c}_n, a \rangle_x - \langle c, a \rangle \, \big(\, \langle f_n, b \rangle + \langle g_n, a \rangle \,\big) = \; u_{n,x} - 2r \int_{-\infty}^{x} \big(\, u_n q + r v_n \,\big) \, d\xi, \\
v_{n+1} &= \langle \widehat{d}_{n+1}, b \rangle \;=\; \Big\langle -\, \widehat{d}_{n,x} + \big(\, \langle g_n, a \rangle d + \langle d, b \rangle f_n \,\big), b \Big\rangle \\
&= -\langle \widehat{d}_n, b \rangle_x + \langle d, b \rangle \, \big(\, \langle g_n, a \rangle + \langle f_n, b \rangle \,\big) = \; -v_{n,x} + 2q \int_{-\infty}^{x} \big(\, u_n q + r v_n \,\big) \, d\xi.
\end{aligned}
$$

In other words,

$$
\begin{pmatrix} u_{n+1} \\ v_{n+1} \end{pmatrix} = T_{r,q} \begin{pmatrix} u_n \\ v_n \end{pmatrix},
$$

and part b) of the assertion follows again by induction. The proof of Claim 1 is complete.

Claim 2: Choose $c_0 \in E$, $d_0 \in F$ such thay $\langle c_0, a \rangle = \langle d_0, b \rangle = 1$. Then, for all $n \in \mathbb{N}_0$, the following identities hold:

$$
u_n = \langle U_n d_0, a \rangle, \qquad v_n = \langle V_n c_0, b \rangle.
$$

Proof of Claim 2: Applying the operator $U_n = b \otimes \widehat{c}_n \in \mathcal{S}_b(F, E)$ to the vector $d_0 \in F$, we obtain $U_n d_0 = \langle d_0, b \rangle \widehat{c}_n = \widehat{c}_n$. Thus $\langle U_n d_0, a \rangle = \langle \widehat{c}_n, a \rangle = u_n$. Analogously,

$$
\langle V_n c_0, b \rangle = \langle (a \otimes \widehat{d}_n) c_0, b \rangle = \langle \langle c_0, a \rangle \widehat{d}_n, b \rangle = \langle \widehat{d}_n, b \rangle,
$$

and Claim 2 is shown.

To conclude the proof, we assume that the polynomials f, g are concretely given by

$$
f(z) = \sum_{n=0}^{N} a_n z^n, \qquad g(z) = \sum_{n=0}^{N} b_n z^n,
$$

where we allow leading terms to vanish in order to have the same N for both polynomials. Let $Q = Q(x,t) \in \mathcal{S}_a(E, F)$, $R = R(x,t) \in \mathcal{S}_b(F, E)$ be a solution of the non-abelian AKNS system (1.8), and set

$$
\begin{pmatrix} R_n \\ Q_n \end{pmatrix} = \big(\, \mathcal{T}_{R,Q} \,\big)^n \begin{pmatrix} R \\ -Q \end{pmatrix}, \qquad \begin{pmatrix} \widehat{R}_n \\ \widehat{Q}_n \end{pmatrix} = \big(\, \mathcal{T}_{R,Q} \,\big)^n \begin{pmatrix} R_t \\ Q_t \end{pmatrix},
$$

for $n \in \mathbb{N}_0$. Then the non-abelian AKNS system (1.8) reads

$$
\sum_{n=0}^{N} a_n R_n = \sum_{n=0}^{N} b_n \widehat{R}_n, \qquad \sum_{n=0}^{N} a_n Q_n = \sum_{n=0}^{N} b_n \widehat{Q}_n. \tag{2.2}
$$

By Claim 1, R_n, $\widehat{R}_n \in \mathcal{S}_b(F, E)$, and Q_n, $\widehat{Q}_n \in \mathcal{S}_a(E, F)$ for all n.

Moreover, by the Claims 1 and 2, the scalar functions $r_n = \langle R_n d_0, a \rangle$, $\widehat{r}_n = \langle \widehat{R}_n d_0, a \rangle$, $q_n = \langle Q_n c_0, b \rangle$, and $\widehat{q}_n = \langle \widehat{Q}_n c_0, b \rangle$ assigned to these operators (where $c_0 \in E$, $d_0 \in F$ are the constant functionals from Claim 2) satisfy

$$
\begin{pmatrix} r_n \\ q_n \end{pmatrix} = \big(\, T_{r,q} \,\big)^n \begin{pmatrix} r \\ -q \end{pmatrix}, \qquad \begin{pmatrix} \widehat{r}_n \\ \widehat{q}_n \end{pmatrix} = \big(\, T_{r,q} \,\big)^n \begin{pmatrix} r_t \\ q_t \end{pmatrix}, \tag{2.3}
$$

By (2.2),

$$\begin{aligned}
\sum_{n=0}^{N} a_n r_n &= \sum_{n=0}^{N} a_n \langle R_n d_0, a\rangle = \Big\langle \Big(\sum_{n=0}^{N} a_n R_n \Big) d_0, a\Big\rangle \\
&= \Big\langle \Big(\sum_{n=0}^{N} b_n \widehat{R}_n \Big) d_0, a\Big\rangle = \sum_{m=0}^{N} b_n \langle \widehat{R}_n d_0, a\rangle = \sum_{m=0}^{N} b_n \widehat{r}_n.
\end{aligned}$$

Similarily, we see $\sum_{n=0}^{N} a_n q_n = \sum_{n=0}^{N} b_n \widehat{q}_n$.

As a consequence,

$$\begin{aligned}
f\,(\,T_{r,q}\,)\begin{pmatrix} r \\ -q \end{pmatrix} &= \sum_{n=0}^{N} a_n\,(\,T_{r,q}\,)^n \begin{pmatrix} r \\ -q \end{pmatrix} \overset{(2.3)}{=} \sum_{n=0}^{N} a_n \begin{pmatrix} r_n \\ q_n \end{pmatrix} \\
&= \sum_{n=0}^{N} b_n \begin{pmatrix} \widehat{r}_n \\ \widehat{q}_n \end{pmatrix} \overset{(2.3)}{=} \sum_{n=0}^{N} b_n\,(\,T_{r,q}\,)^n \begin{pmatrix} r_t \\ q_t \end{pmatrix} = g\,(\,T_{r,q}\,)\begin{pmatrix} r_t \\ q_t \end{pmatrix}.
\end{aligned}$$

This completes the proof. □

2.3 Derivation of solution formulas

In this section we carry out the scalarization process in Section 2.2 for the concrete operator-valued solution, the operator soliton, derived in Theorem 1.2.1 and present a first solution formula for the AKNS system. The central condition in this context is that the operator soliton be one-dimensional. We show that there is a systematic way to fulfill this condition using the theory of so-called elementary operators. Finally, we achieve an improvement of the solution formula particularly adequate for applications. The main tool are determinants on quasi-Banach operator ideals.

2.3.1 Carrying out the scalarization process for the operator soliton

Let E, F be Banach spaces, and let $A \in \mathcal{L}(E)$, $B \in \mathcal{L}(F)$ be constant operators such that $\operatorname{spec}(A) \cup \operatorname{spec}(-B)$ is contained in the domain where f_0 is holomorphic. Then the base equations in Theorem 1.2.1 are obviously fulfilled for

$$L(x,t) = \widehat{L}(x,t)C, \qquad M(x,t) = \widehat{M}(x,t)D,$$

where $\widehat{L}(x,t) = \exp\big(Ax + f_0(A)t\big)$, $\widehat{M}(x,t) = \exp\big(Bx - f_0(-B)t\big)$, and constant operators $C \in \mathcal{L}(F,E)$, $D \in \mathcal{L}(E,F)$.

This means that the operator-valued solution in Theorem 1.2.1 reads

$$Q = (I - ML)^{-1}\widehat{M}(BD + DA), \qquad R = (I - LM)^{-1}\widehat{L}(AC + CB).$$

Thus it is clear that the operator functions Q, R will only become one-dimensional (with fixed kernels) if the expressions

$$BD + DA, \qquad AC + CB,$$

can be chosen one-dimensional.

Proposition 2.3.1. *Let E, F be Banach spaces and $A \in \mathcal{L}(E)$, $B \in \mathcal{L}(F)$ such that* $\operatorname{spec}(A) \cup \operatorname{spec}(-B)$ *is contained in the domain where f_0 is holomorphic and* $\exp(Ax)$, $\exp(Bx)$ *behave sufficiently well for $x \to -\infty$.*

Assume that $C \in \mathcal{L}(F,E)$, $D \in \mathcal{L}(E,F)$ are constant operators satisfying the one-dimensionality conditions

$$AC + CB = b \otimes c, \qquad BD + DA = a \otimes d, \tag{2.4}$$

for $0 \neq a \in E'$, $c \in E$, $0 \neq b \in F'$, and $d \in F$. Define the operator-functions

$$\begin{aligned} L(x,t) &= \widehat{L}(x,t)C, & \widehat{L}(x,t) &= \exp\big(Ax + f_0(A)t\big), \\ M(x,t) &= \widehat{M}(x,t)D, & \widehat{M}(x,t) &= \exp\big(Bx - f_0(-B)t\big). \end{aligned}$$

Then, on every domain $\mathbb{R} \times (t_1, t_2)$ on which $(I - LM)^{-1}$, $(I - ML)^{-1}$ are defined, a solution of the scalar AKNS-system (1.1) is given by

$$q = \operatorname{tr}\Big(\big(I - ML\big)^{-1}\widehat{M}(b \otimes d)\Big), \tag{2.5}$$

$$r = \operatorname{tr}\Big(\big(I - LM\big)^{-1}\widehat{L}(a \otimes c)\Big). \tag{2.6}$$

Recall the notation tr for the trace on the finite-rank endomorphisms (see Appendix B).

Proof According to the explanations preceding the theorem, the coupling conditions guarantee $Q = Q(x,t) \in \mathcal{S}_a(E,F)$, $R = R(x,t) \in \mathcal{S}_b(F,E)$ for the operator solutions Q, R in Theorem 1.2.1. Namely,

$$\begin{aligned} Q &= (I - ML)^{-1}\widehat{M}\,(a \otimes d) \;=\; a \otimes \big(\,(I - ML)^{-1}\widehat{M}d\,\big), \\ R &= (I - LM)^{-1}\widehat{L}\,(b \otimes c) \;=\; b \otimes \big(\,(I - LM)^{-1}\widehat{L}c\,\big)\,. \end{aligned}$$

Application of Theorem 2.2.1 then shows that a solution of the scalar AKNS-system (1.1) is given by

$$\begin{aligned} q &= \Big\langle (I - ML)^{-1}\widehat{M}\,d, b\Big\rangle \;=\; \operatorname{tr}\Big(\big(I - ML\big)^{-1}\widehat{M}(b \otimes d)\Big), \\ r &= \Big\langle (I - LM)^{-1}\widehat{L}\,c, a\Big\rangle \;=\; \operatorname{tr}\Big(\big(I - LM\big)^{-1}\widehat{L}(a \otimes c)\Big), \end{aligned}$$

the latter by definition of tr, see Appendix B.

This completes the proof. □

2.3.2 On elementary operators: How and why to solve the operator equation $AX + XB = C$

In order to apply Proposition 2.3.1, we have to choose C, D such that the one-dimensionality conditions (2.4) are satisfied. In other words, we ask whether the expression $\Phi_{A,B}(X) = AX + XB$ can be made one-dimensional by an appropriate choice of X. Fortunately, the theory of the map $\Phi_{A,B}$ is fairly well understood and provides powerful results on existence and properties of solutions.

For a detailed discussion of the operator equation $AX + XB = C$ and its applications we refer to the survey article of Bhatia and Rosenthal [12].

Definition 2.3.2. *Let $\mathcal{A}$ be a p-Banach operator ideal $(0 < p \leq 1)$ and $A \in \mathcal{L}(E)$, $B \in \mathcal{L}(E)$. Then the operator $\Phi_{A,B} : \mathcal{A}(F,E) \longrightarrow \mathcal{A}(F,E)$ is defined by*

$$\Phi_{A,B}(X) = AX + XB.$$

Note that $\Phi_{A,B}$ is an endomorphism of $\mathcal{A}(F,E)$ for every p-Banach operator ideal $\mathcal{A}$.

The above defined operator is a particular case of a so-called elementary operator, i.e., an operator $\Phi : \mathcal{A}(F,E) \longrightarrow \mathcal{A}(F,E)$ of the form

$$\Phi = p(R_B, L_A) \qquad \text{for a polynomial } p,$$

where R_B, L_A denote the multiplication with $A \in \mathcal{L}(E)$ from the left and $B \in \mathcal{L}(F)$ from the right, respectively.

In fundamental papers Dash/Schechter [22] and Eschmeier [28] proved that, provided that $\mathcal{A}$ is a Banach operator ideal, $\operatorname{spec}(\Phi)$ is independent of the choice of $\mathcal{A}$ and given by the catching formula

$$\operatorname{spec}(\Phi) = p(\operatorname{spec}(B), \operatorname{spec}(A)).$$

Next we cite a result of Aden extending, for the operator $\Phi_{A,B}$, this result to p-Banach operator ideals $\mathcal{A}$, $0 < p \le 1$.

Proposition 2.3.3. *([7], Theorem III.2.7) Let $\mathcal{A}$ be a p-Banach operator ideal ($0 < p \le 1$). Then the spectrum of the operator $\Phi_{A,B}$ is given by*

$$\operatorname{spec}(\Phi_{A,B}) = \operatorname{spec}(A) + \operatorname{spec}(B).$$

In particular, $\operatorname{spec}(\Phi_{A,B})$ *does not depended on the p-Banach operator ideal $\mathcal{A}$ under consideration.*

The next proposition is an immediate consequence. It stresses that the operator equation $AX + XB = C$ with given $C \in \mathcal{A}(F,E)$ can always be solved if $0 \notin \operatorname{spec}(A) + \operatorname{spec}(B)$. Moreover, the solution $X = \Phi_{A,B}^{-1}(C)$ is as good as C is, more precisely, $X \in \mathcal{A}(F,E)$.

Proposition 2.3.4. *Let $\mathcal{A}$ be a p-Banach operator ideal ($0 < p \le 1$) and $0 \notin \operatorname{spec}(A) + \operatorname{spec}(B)$. Then, for every operator $C \in \mathcal{A}(F,E)$, the equation $AX + XB = C$ has a unique solution $X \in \mathcal{A}(F,E)$, namely*

$$X = \Phi_{A,B}^{-1}(C).$$

Remark 2.3.5. *In the above results p-Banach operator ideals could be replaced by the seemingly more general but in fact equivalent quasi-Banach operator ideals (see Pietsch [72]).*

To conclude the section, we state the impact of the above theory to our solution formula.

Proposition 2.3.6. *Let E, F be Banach spaces and $A \in \mathcal{L}(E)$, $B \in \mathcal{L}(F)$ such that $0 \notin \operatorname{spec}(A) + \operatorname{spec}(B)$. Then, for all $a \in E'$, $c \in E$, $b \in F'$, $d \in F$, the operators*

$$C = \Phi_{A,B}^{-1}(b \otimes c), \qquad D = \Phi_{B,A}^{-1}(a \otimes d),$$

satisfy the one-dimensionality conditions (2.4) in Proposition 2.3.1.

2.3.3 Solution formulas in terms of determinants

The solution formulas (2.5), (2.6) in Proposition 2.3.1 have the disadvantage that one has to compute the inverse operators $(I - LM)^{-1}$, $(I - ML)^{-1}$. In this section we improve these formulas in order to make them accessible for explicit calculations. This will be crucial in the applications in Chapters 4 – 7.

Starting point is the observation that the functional ev_a, considered as a map on $\mathcal{S}_a(E)$, coincides with the well-known trace. Thus it is tempting to use the calculus of traces and determinants to improve the solution formula. The problem is that the intermediate calculations pass through traces of non-finite operators, whose existence is not obvious at all. The adequate framework to overcome this problem is theory of traces and determinants on quasi-Banach operator ideals ([41], [73], see also Appendix B).

The following theorem is the crucial result of this chapter.

Theorem 2.3.7. *Let E, F be Banach spaces, and let $A \in \mathcal{L}(E)$, $B \in \mathcal{L}(F)$ be such that* $\operatorname{spec}(A) \cup \operatorname{spec}(-B)$ *is contained in the domain where f_0 is holomorphic and* $\exp(Ax)$, $\exp(Bx)$ *behave sufficiently well as $x \to -\infty$.*

a) Assume that the one-dimensionality conditions (2.4) are satisfied with $C \in \mathcal{A}(F,E)$, $D \in \mathcal{A}(E,F)$, and $0 \neq a \in E'$, $0 \neq b \in F'$, where $\mathcal{A}$ is an arbitrary quasi-Banach operator ideal admitting a continuous determinant δ. Define

$$\begin{aligned} L(x,t) &= \widehat{L}(x,t)C \qquad \textit{with } \widehat{L}(x,t) = \exp\big(Ax + f_0(A)t\big), \\ M(x,t) &= \widehat{M}(x,t)D \qquad \textit{with } \widehat{M}(x,t) = \exp\big(Bx - f_0(-B)t\big). \end{aligned}$$

Then, on strips $\mathbb{R} \times (t_1, t_2)$ on which $\delta\,\big(\, I - LM \,\big)$ does not vanish, a solution of the scalar AKNS-system (1.1) is given by

$$q = 1 - \frac{\delta\,\big(\, I - ML - \widehat{M}\,(b \otimes d)\,\big)}{\delta\,\big(\, I - ML\,\big)}, \tag{2.7}$$

$$r = 1 - \frac{\delta\,\big(\, I - LM - \widehat{L}\,(a \otimes c)\,\big)}{\delta\,\big(\, I - LM\,\big)}. \tag{2.8}$$

b) If $0 \notin \operatorname{spec}(A) + \operatorname{spec}(B)$, then a) holds with $C = \Phi_{A,B}^{-1}(b \otimes c)$, $D = \Phi_{B,A}^{-1}(a \otimes d)$, $a, b \neq 0$, and any p-Banach operator ideal $\mathcal{A}$ admitting a continuous determinant δ.

Proof a) We show that the solution q, r given by (2.5), (2.6) in Proposition 2.3.1 can be rewritten in the form (2.7), (2.8). To this end, we first use the well-known identity $\det(I+T) = 1 + \operatorname{tr}(T)$, valid for any one-dimensional operator T. This yields

$$\begin{aligned} q &= \operatorname{tr}\Big((I - ML)^{-1}\,\widehat{M}(b \otimes d)\Big) \\ &= 1 - \det\Big(I - (I - ML)^{-1}\,\widehat{M}(b \otimes d)\Big) \\ &= 1 - \det\Big((I - ML)^{-1}\Big(I - ML - \widehat{M}(b \otimes d)\Big)\Big). \end{aligned}$$

Here det denotes the canonical determinant on the finite-rank operators.

Note that we cannot use the multiplicativity of the determinant det, because the operator ML is not of finite rank. Thus we switch to the determinant δ on the quasi-Banach operator ideal $\mathcal{A}$. By the definition of general determinants (see Appendix B), we have $\delta|_{\mathcal{F}} = \det$. Thus

$$\begin{aligned} q &= 1 - \delta\Big((I - ML)^{-1}\Big(I - ML - \widehat{M}(b \otimes d)\Big)\Big) \\ &= 1 - \frac{\delta\Big(I - ML - \widehat{M}(b \otimes d)\Big)}{\delta\Big(I - ML\Big)}. \end{aligned}$$

The reformulation of the formula for r follows analogously.

b) This is an immediate consequence of Proposition 2.3.4. $\square$

2.4 The solution formula revisited

For the applications in particular in context with the asymptotic behaviour of negatons (see Chapter 5), a reformulation of the solution formula which provides an even better access for calculations will be decisive. To this end we need some tools for the calculation with general determinants on quasi-Banach operator ideals.

2.4.1 Key relations for general determinants

As a rule, the following identities will be elementary for finite matrices. In the general case we have to check the arguments carefully to guarantee existence. Note also that we suppose continuity of δ but not that the finite-rank operators $\mathcal{F}$ are $\|\cdot|\mathcal{A}\|$-dense in $\mathcal{A}$.

Proposition 2.4.1. *Let E, F be Banach spaces and $\mathcal{A}$ a quasi-Banach operator ideal admitting a continuous determinant δ. Then*

a) $\begin{pmatrix} 0 & U \\ V & 0 \end{pmatrix} \in \mathcal{A}(E \oplus F)$ *whenever* $U \in \mathcal{A}(F,E)$, $V \in \mathcal{A}(E,F)$.

b) If, in addition, $U = U(x)$, $V = V(x)$ depend smoothly on a real variable x, if $||U(x)||$, $||V(x)|| \to 0$ for $x \to -\infty$, and if $I_E - UV$ is always invertible, then

$$\delta \begin{pmatrix} I_E & U \\ V & I_F \end{pmatrix} = \delta \left(I_E - UV \right) .$$

Proof a) Let $P_E : E \oplus F \to E$, $P_F : E \oplus F \to F$ be the canonical projections from $E \oplus F$ to E, F, respectively, and $J_E : E \to E \oplus F$, $J_F : F \to E \oplus F$ the canonical embeddings of E, F, respectively, in $E \oplus F$. Then a) follows immediately from the factorization

$$\begin{pmatrix} 0 & U \\ V & 0 \end{pmatrix} = J_E U P_F + J_F V P_E$$

and the ideal properties of $\mathcal{A}$.

b) Abbreviate $W = \begin{pmatrix} 0 & U \\ V & 0 \end{pmatrix}$, $I = I_{E\oplus F}$, and set $J = \begin{pmatrix} I_E & 0 \\ 0 & -I_F \end{pmatrix}$. First we observe

$$\delta(I - W) \;=\; \delta(I + JWJ) \;=\; \delta(I + J^2 W) \;=\; \delta(I + W).$$

As a consequence,

$$\begin{aligned} \Big(\delta(I+W)\Big)^2 \;&=\; \delta(I+W)\,\delta(I-W) \;=\; \delta\Big((I+W)(I-W)\Big) \;=\; \delta(I-W^2) \\ &=\; \delta \begin{pmatrix} I_E - UV & 0 \\ 0 & I_F - VU \end{pmatrix} = \delta \begin{pmatrix} I_E - UV & 0 \\ 0 & I_F \end{pmatrix} \delta \begin{pmatrix} I_E & 0 \\ 0 & I_F - VU \end{pmatrix} \\ &=\; \delta\big(I_E - UV\big)\,\delta\big(I_F - VU\big) \qquad\qquad (2.9) \\ &=\; \Big(\delta(I_E - UV)\Big)^2. \end{aligned}$$

Since we deal with general determinants, the identity (2.9) requires a comment. In fact, with the same notations as in a), (2.9) follows from

$$\delta\Big(\begin{pmatrix} I_E + T & 0 \\ 0 & I_F \end{pmatrix}\Big) \;=\; \delta(I + J_E T P_E) \;=\; \delta(I_E + P_E J_E T) \;=\; \delta(I_E + T)$$

for all $T \in \mathcal{L}(E)$. The second identity holds by Definition B.1.8 (iii).

To conclude the proof, consider the function

$$f(x) \;=\; \delta\big(I+W(x)\big)\Big/\delta\big(I_E-U(x)V(x)\big),$$

which is well-defined by assumption. Obviously f is a continuous function, and we have shown that it takes its values in $\{-1,1\}$. Thus f is constant. Finally, from $U(x)$, $V(x)\to 0$ for $x\to-\infty$ it is clear that $f(x)\to 1$ for $x\to-\infty$. Thus $f\equiv 1$, and the proof is complete. □

Corollary 2.4.2. *Under the assumptions of Proposition 2.4.1, the following relation holds in the perturbed case,*

$$\delta\begin{pmatrix} I_E-a\otimes c & U \\ V & I_F\end{pmatrix}=\delta\left(I_E-(UV+a\otimes c)\right).$$

for any $a=a(x)\in F'$, $c=c(x)\in F$.

Proof Note that $\begin{pmatrix} I_E & U \\ V & I_F\end{pmatrix}$ is invertible by Proposition 2.4.1, and we check

$$\begin{pmatrix} I_E & U \\ V & I_F\end{pmatrix}^{-1}=\begin{pmatrix} (I_E-UV)^{-1} & -(I_E-UV)^{-1}U \\ -(I_F-VU)^{-1}V & (I_F-VU)^{-1}\end{pmatrix}.$$

Therefore,

$$\begin{aligned}\begin{pmatrix} I_E-a\otimes c & U \\ V & I_F\end{pmatrix}&=\begin{pmatrix} I_E & U \\ V & I_F\end{pmatrix}\left(\begin{pmatrix} I_E & U \\ V & I_F\end{pmatrix}^{-1}\begin{pmatrix} I_E-a\otimes c & U \\ V & I_F\end{pmatrix}\right)\\ &=\begin{pmatrix} I_E & U \\ V & I_F\end{pmatrix}\begin{pmatrix} I_E\;-\;(I_E-UV)^{-1}\,(a\otimes c) & 0 \\ (I_F-VU)^{-1}V(a\otimes c) & I_F\end{pmatrix}\\ &=\begin{pmatrix} I_E & U \\ V & I_F\end{pmatrix}\big(I+f\otimes g\big),\end{aligned}$$

where we have abbreviated $I=I_{E\oplus F}$, $f=(a,0)$, $g=\big(-(I_E-UV)^{-1}c,(I_F-VU)^{-1}Vc\big)$. Next we use that the value of determinants is prescribed on one-dimensional operators. We thus get

$$\begin{aligned}\delta(I+f\otimes g)&=1+\langle\, g,f\,\rangle=1-\langle\,(I_E-UV)^{-1}c,a\,\rangle\\ &=\delta\Big(I_E-a\otimes\big((I_E-UV)^{-1}c\big)\Big)\\ &=\delta\big(I_E-(I_E-UV)^{-1}(a\otimes c)\big).\end{aligned}$$

Now multiplicativity of the determinant, the above identity, and Proposition 2.4.1 yield

$$\begin{aligned}\delta\begin{pmatrix} I_E-a\otimes c & U \\ V & I_F\end{pmatrix}&=\delta\begin{pmatrix} I_E & U \\ V & I_F\end{pmatrix}\delta(I+f\otimes g)\\ &=\delta\big(I_E-UV\big)\;\delta\big(I_E-(I_E-UV)^{-1}(a\otimes c)\big)\\ &=\delta\big(I_E-(UV+a\otimes c)\big),\end{aligned}$$

which is the assertion. □

For later use we also note a symmetry relation for the expression of our solution formulas.

Proposition 2.4.3. *Let E, F be Banach spaces, and $\mathcal{A}$ a quasi-Banach operator ideal with a continuous determinant δ. Then the following relation holds*

$$1-\frac{\delta\begin{pmatrix} I_E - a\otimes c & U \\ V & I_F \end{pmatrix}}{\delta\begin{pmatrix} I_E & U \\ V & I_F \end{pmatrix}} = \frac{\delta\begin{pmatrix} I_E + a\otimes c & U \\ V & I_F \end{pmatrix}}{\delta\begin{pmatrix} I_E & U \\ V & I_F \end{pmatrix}} - 1$$

for $U \in \mathcal{A}(F,E)$, $V \in \mathcal{A}(E,F)$, and $a \in E'$, $c \in E$.

Proof To simplify writing, set $W = \begin{pmatrix} 0 & U \\ V & 0 \end{pmatrix}$, $I = I_{E\oplus F}$, $f = (a,0)$, $g = (c,0)$. Then

$$\begin{aligned}
1-\frac{\delta(I+W-f\otimes g)}{\delta(1+W)} &= 1-\delta\Big((I+W)^{-1}(I+W-f\otimes g)\Big) \\
= \quad 1-\delta\Big(I-(I+W)^{-1}(f\otimes g)\Big) &= \tau\Big((I+W)^{-1}(f\otimes g)\Big) \\
= \quad -1+\delta\Big(I+(I+W)^{-1}(f\otimes g)\Big) &= -1+\delta\Big((I+W)^{-1}(1+W+f\otimes g)\Big) \\
= \quad -1+\frac{\delta(I+W+f\otimes g)}{\delta(I+W)}, &
\end{aligned}$$

where we have used the identity $\delta(1+T) = 1+\tau(T)$ which is valid for one-dimensional operators $T \in \mathcal{A}(E)$. $\square$

2.4.2 Statement of the main theorem

It remains to formulate the main theorem of the present chapter. It is an immediate consequence of Theorem 2.3.7, Proposition 2.4.1, and Corollary 2.4.2.

Theorem 2.4.4. *Let E, F be Banach spaces, and let $A \in \mathcal{L}(E)$, $B \in \mathcal{L}(F)$ be such that $\operatorname{spec}(A) \cup \operatorname{spec}(-B)$ is contained in the domain where f_0 is holomorphic and $\exp(Ax)$, $\exp(Bx)$ behave sufficiently well as $x \to -\infty$.*

a) Assume that the one-dimensionality conditions (2.4) are satisfied with $C \in \mathcal{A}(F,E)$, $D \in \mathcal{A}(E,F)$, $0 \neq a \in E'$, $0 \neq b \in F'$, where $\mathcal{A}$ is an arbitrary quasi-Banach operator ideal admitting a continuous determinant δ. Define

$$\begin{aligned}
L(x,t) &= \widehat{L}(x,t)C \qquad \text{with } \widehat{L}(x,t) = \exp\big(Ax+f_0(A)t\big), \\
M(x,t) &= \widehat{M}(x,t)D \qquad \text{with } \widehat{M}(x,t) = \exp\big(Bx-f_0(-B)t\big).
\end{aligned}$$

Then, on strips $\mathbb{R}\times(t_1,t_2)$ on which p does not vanish, a solution of the scalar AKNS-system (1.1) is given by

$$q = 1-P/p, \qquad r = 1-\widehat{P}/p, \tag{2.10}$$

where

$$P = \delta\begin{pmatrix} I_E & L \\ M & I_F-\widehat{M}(b\otimes d) \end{pmatrix}, \quad \widehat{P} = \delta\begin{pmatrix} I_E-\widehat{L}(a\otimes c) & L \\ M & I_F \end{pmatrix}, \quad p = \delta\begin{pmatrix} I_E & L \\ M & I_F \end{pmatrix}.$$

b) If $0 \notin \operatorname{spec}(A)+\operatorname{spec}(B)$, then a) holds with $C = \Phi_{A,B}^{-1}(b\otimes c)$, $D = \Phi_{B,A}^{-1}(a\otimes d)$, $a, b \neq 0$, and any p-Banach operator ideal $\mathcal{A}$ admitting a continuous determinant δ.

2.5 Improvements for the $\mathbb{R}$-reduction

In this section we aim at an improvement of the solution formula in Theorem 2.3.7 for the $\mathbb{R}$-reduced AKNS system. Namely, we reduce the involved determinants to half of their size. This is of course a considerable advantage in explicit calculations.

Referring to Section 4.2 for a systematic introduction of $\mathbb{R}$-reduction, we just suppose that $f_0(z) = -f_0(-z)$ holds at every z in the domain where f_0 is holomorphic. Then the $\mathbb{R}$-reduced equation is obtained by setting $r = -q$. In particular, it is natural to restrict to

$$F = E, \qquad B = A, \qquad b = -a, \quad d = c. \tag{2.11}$$

Then we obtain a nice solution formula, which is a logarithmic derivative.

Theorem 2.5.1. *Let E be a Banach space and $A \in \mathcal{L}(E)$ such that* $\operatorname{spec}(A)$ *is contained in the domain where f_0 is holomorphic and* $\exp(Ax)$ *behaves sufficiently well as $x \to -\infty$.*

a) Assume that the one-dimensionality condition

$$AC + CA = a \otimes c \tag{2.12}$$

for $0 \neq a \in E'$, $c \in E$, is satisfied with $C \in \mathcal{A}(E)$, where $\mathcal{A}$ is an arbitrary quasi-Banach operator ideal admitting a continuous determinant δ. Define

$$L(x,t) = \widehat{L}(x,t)C, \qquad \text{with } \widehat{L} = \exp\big(\, Ax + f_0(A)t \,\big)\,.$$

Then, on strips $\mathbb{R} \times (t_1, t_2)$ on which $\delta(I \pm \mathrm{i}L)$ do not vanish, a solution of the $\mathbb{R}$-reduced AKNS-system (4.6) is given by

$$q = \mathrm{i}\,\frac{\partial}{\partial x}\,\log\frac{\delta(I + \mathrm{i}L)}{\delta(I - \mathrm{i}L)}. \tag{2.13}$$

b) If $0 \notin \operatorname{spec}(A) + \operatorname{spec}(A)$, then a) holds with $C = \Phi_{A,A}^{-1}(a \otimes c)$, $a \neq 0$, and any p-Banach operator ideal $\mathcal{A}$ admitting a continuous determinant δ.

Proof If we apply Proposition 2.3.1 with the particular choice (2.11), then $r = -q$ follows from $\widehat{M} = \widehat{L}$ and the fact that we can take $D = -C$. It remains to show that the formula (2.5) for q can be rewritten in the form (2.13).

To this end we denote by τ the trace corresponding to δ via to the Trace-determinant theorem (see Proposition B.2.11). Since all traces coincide on $\mathcal{F}$, the operator ideal of finite-rank operators, we have $\tau|_{\mathcal{F}} = \mathrm{tr}$. Therefore, and by the coupling condition (2.12),

$$\begin{aligned} q &= -\mathrm{tr}\Big(\, (\, I + L^2\,)^{-1}\; \widehat{L}(a \otimes c)\Big) \\ &= -\tau\Big(\, (\, I + L^2\,)^{-1}\; (AL + LA)\Big). \end{aligned}$$

Next we use the linearity of the trace τ (note that this is allowed because $L \in \mathcal{A}(E)$) to observe

$$\begin{aligned} q &= -\tau\Big(\, (\, I + L^2\,)^{-1}\; AL\Big) - \tau\Big(\, (\, I + L^2\,)^{-1}\; LA\Big) \\ &= -2\tau\Big(\, (\, I + L^2\,)^{-1}\; L_x\Big), \end{aligned}$$

the latter by the property of traces ($\tau(ST) = \tau(TS)$) and the base equation for L. Using the identity $2(I + L^2)^{-1} = (I + \mathrm{i}L)^{-1} + (I - \mathrm{i}L)^{-1}$, we finally get

$$\begin{aligned} q &= \mathrm{i}\,\tau\Big((I + \mathrm{i}L)^{-1}(\mathrm{i}L)_x\Big) - \mathrm{i}\,\tau\Big((I - \mathrm{i}L)^{-1}(-\mathrm{i}L)_x\Big) \\ &= \mathrm{i}\,\frac{\partial}{\partial x}\,\log\frac{\delta(I + \mathrm{i}L)}{\delta(I - \mathrm{i}L)}, \end{aligned}$$

the latter by the differentiation rule for determinants, see Proposition B.2.12. $\square$

2.6 Ameliorated formulas for NLS and mKdV

Theorem 1.2.1 give at one stroke solutions for *all* equations of the AKNS system. Note that the appearance of the integral term $\mathcal{T}_{R,Q}$ enforced strong assumptions on the behaviour for the functions $L = L(x,t)$, $M = M(x,t)$ for $x \to -\infty$. As already seen, $\mathcal{T}_{R,Q}$ may cancel during the calculations. This may allow us to generalize the resulting solution formulas by direct verifications avoiding $\mathcal{T}_{R,Q}$. We illustrate this for the Nonlinear Schrödinger and the modified Korteweg-de Vries equations.

For the Nonlinear Schrödinger equation, the following result holds.

Proposition 2.6.1. *Let E, F be Banach spaces and $A \in \mathcal{L}(E)$, $B \in \mathcal{L}(F)$.*

a) Assume that the one-dimensionality conditions (2.4) are satisfied with $C \in \mathcal{A}(F,E)$, $D \in \mathcal{A}(E,F)$, $0 \neq a \in E'$, $0 \neq b \in F'$, where $\mathcal{A}$ is an arbitrary quasi-Banach operator ideal admitting a continuous determinant δ. Define

$$\begin{aligned} L(x,t) &= \widehat{L}(x,t)C \qquad \text{with } \widehat{L}(x,t) = \exp\big(Ax - \mathrm{i}A^2t\big), \\ M(x,t) &= \widehat{M}(x,t)D \qquad \text{with } \widehat{M}(x,t) = \exp\big(Bx + \mathrm{i}B^2t\big). \end{aligned}$$

Then, on $\Omega = \{(x,t) \mid p(x,t) \neq 0\}$, a solution of the scalar NLS system

$$\begin{aligned} -\mathrm{i}r_t + r_{xx} - 2r^2q &= 0, \\ \mathrm{i}q_t + q_{xx} - 2q^2r &= 0. \end{aligned}$$

is given by

$$q = 1 - P/p, \qquad r = 1 - \widehat{P}/p,$$

where

$$P = \delta\begin{pmatrix} I & L \\ M & I - \widehat{M}b \otimes d \end{pmatrix}, \quad \widehat{P} = \delta\begin{pmatrix} I - \widehat{L}a \otimes c & L \\ M & I \end{pmatrix}, \quad p = \delta\begin{pmatrix} I & L \\ M & I \end{pmatrix}.$$

b) If $0 \notin \mathrm{spec}(A) + \mathrm{spec}(B)$, then a) holds with $C = \Phi_{A,B}^{-1}(b \otimes c)$, $D = \Phi_{B,A}^{-1}(a \otimes d)$, $a, b \neq 0$, and any p-Banach operator ideal $\mathcal{A}$ admitting a continuous determinant δ.

We want to stress again that the freedom to choose the two operators A, $B \in \mathcal{L}(E)$ independently is essential. Otherwise even solitons could not be derived in full generality (see Remark 1.2.12).

Proof To avoid the growth conditions in the proposition, we have to verify that the operator-functions Q, R given by (1.10), (1.11) are a solution of the non-abelian NLS system (1.58), (1.59) without using Theorem 1.2.1. To this end, observe the rules

$$\begin{aligned} R_t &= -\mathrm{i}\widehat{R}_2R, \qquad R_x = \widehat{R}_1R, \\ Q_t &= \ \mathrm{i}\widehat{Q}_2Q, \qquad Q_x = \widehat{Q}_1Q, \end{aligned}$$

for $\widehat{R}_j = (I - LM)^{-1}\big(A^j + (-1)^{j-1}LB^jM\big)$, $\widehat{Q}_j = (I - ML)^{-1}\big(B^j + (-1)^{j-1}MA^jL\big)$, $j = 1, 2$, which follow from Lemma 1.2.7, Lemma 1.2.5.
Using Lemma 1.2.5 once more, we observe

$$\widehat{R}_{1,x} = RQ, \qquad \widehat{Q}_{1,x} = QR.$$

This easily yields

$$R_{xx} = RQR + \widehat{R}_1^2R, \qquad Q_{xx} = QRQ + \widehat{Q}_1^2Q.$$

Moreover, by Lemma 1.2.6, we have

$$\widehat{R}_1^2 = RQ + \widehat{R}_2, \qquad \widehat{Q}_1^2 = QR + \widehat{Q}_2.$$

Inserting these relations, we get

$$\begin{aligned}
-\mathrm{i}R_t + R_{xx} - 2RQR &= -\widehat{R}_2R + (RQR + \widehat{R}_1^2R) - 2RQR \\
&= -\widehat{R}_2R + (\widehat{R}_1^2 - RQ)R \;=\; 0, \\
\mathrm{i}Q_t + Q_{xx} - 2QRQ &= -\widehat{Q}_2Q + (QRQ + \widehat{Q}_1^2Q) - 2QRQ \\
&= -\widehat{Q}_2Q + (\widehat{Q}_1^2 - QR)Q \;=\; 0.
\end{aligned}$$

Thus Q, R solve the non-abelian NLS system (1.58), (1.59). Now the scalarization process explained in this chapter applies and yields the assertion. □

In analogy with Section 2.5 we get a simpler solution formula for the modified Korteweg-de Vries equation.

Proposition 2.6.2. *Let E be a Banach space and $A \in \mathcal{L}(E)$.*

a) Assume that the one-dimensionality condition (2.12) is satisfied with $C \in \mathcal{A}(E)$ and $0 \neq a \in E'$, where $\mathcal{A}$ is an arbitrary quasi-Banach operator ideal admitting a continuous determinant δ. Define

$$L(x,t) = \widehat{L}(x,t)C, \qquad \text{with } \widehat{L} = \exp\left(\, Ax - A^3t\,\right).$$

Then, on $\Omega = \{(x,t) \mid \delta\big(I \pm \mathrm{i}L(x,t)\big) \neq 0\}$, a solution of the mKdV (1.6) is given by

$$q = \mathrm{i}\,\frac{\partial}{\partial x}\,\log\frac{\delta(I+\mathrm{i}L)}{\delta(I-\mathrm{i}L)}.$$

b) If $0 \notin \operatorname{spec}(A) + \operatorname{spec}(A)$, then a) holds with $C = \Phi_{A,A}^{-1}(a \otimes c)$, $a \neq 0$, and any p-Banach operator ideal $\mathcal{A}$ admitting a continuous determinant δ.

Proof The main point is to show directly that the non-abelian mKdV (1.60) is satisfied for $Q = -(I+L^2)^{-1}(AL + LA)$. To this end, we first collect the necessary manipulation rules. Define

$$\begin{aligned}
Q_j &= (-1)^j\,(I+L^2)^{-1}\Big(A^jL + (-1)^{j-1}LA^j\Big), \\
\widehat{Q}_j &= (-1)^{j-1}\,(I+L^2)^{-1}\Big(A^j + (-1)^jLA^jL\Big),
\end{aligned}$$

for $j = 1,2,3$. In particular, $Q_1 = Q$. Then, by Lemma 1.2.7, Lemma 1.2.5, and Lemma 1.2.6 (with $B = A$, $M = -L$), we observe

$$Q_{1,t} = -\widehat{Q}_3Q_1, \qquad Q_{j,x} = \widehat{Q}_1Q_j, \qquad \widehat{Q}_{j,x} = (-1)^jQ_1Q_j,$$

and

$$\widehat{Q}_1\widehat{Q}_j - (-1)^jQ_1Q_j = -\widehat{Q}_{j+1}, \qquad \widehat{Q}_1Q_j + (-1)^jQ_1\widehat{Q}_j = -Q_{j+1}.$$

Thus we find

$$\begin{aligned}
Q_{1,xx} &= -Q_1^3 + \widehat{Q}_1^2Q_1, \\
Q_{1,xxx} &= \widehat{Q}_1^3Q_1 - 2\widehat{Q}_1Q_1^3 - Q_1\widehat{Q}_1Q_1^2 - 2Q_1^2\widehat{Q}_1Q_1.
\end{aligned}$$

Finally, we insert the above relations to check

$$\begin{aligned}
Q_{1,xxx} + 3(Q_{1,x}Q_1^2 + Q_1^2 Q_{1,x}) \;&=\; \Big(\widehat{Q}_1^3 + \widehat{Q}_1 Q_1^2 - Q_1\widehat{Q}_1 Q_1 + Q_1^2\widehat{Q}_1\Big)Q_1 \\
=\;& \Big(\widehat{Q}_1(\widehat{Q}_1^2 + Q_1^2) + Q_1(Q_1\widehat{Q}_1 - \widehat{Q}_1 Q_1)\Big)Q_1 \;=\; \big(-\widehat{Q}_1\widehat{Q}_2 + Q_1 Q_2\big)Q_1 \;=\; \widehat{Q}_3 Q_1 \\
=\;& -Q_{1,t},
\end{aligned}$$

which shows that $Q_1 = Q$ solves (1.60).

Here even the scalarization process in the simplified setting for endomorphisms (see Section 2.1.2) applies, and the reformulation in terms of determinants can be copied from the proof of Theorem 2.5.1. □

Let us summarize the advantages of Proposition 2.6.1 and Proposition 2.6.2 compared to Theorem 2.2.1.

a) No growth condition for $x \to -\infty$ is assumed.

b) In Theorem 2.2.1 we had to assume the solution to be regular on strips of the form $\mathbb{R} \times (t_1, t_2)$. Proposition 2.6.1 also comprises solutions with poles along real curves $\{p(x,t) = 0\}$. From the mathematical point of view, solutions with *controlled* singularities can be interesting (see [19], [91]).

Chapter 3

Steps towards two-dimensional soliton equations

In this chapter we lay the ground for an operator theoretic investigation of the KP-I and KP-II equations. The KP equations are the most prominent soliton equations in two space dimensions. For analytic and geometric reasons they are harder than one-dimensional equations. Many basic questions are not completely understood and form a topic of recent research (see [2], [6], [15] [61], [86], and references therein).

Here we can obtain the corresponding operator equation as a straightforward generalization of the matrix KP equation, which is itself a topic of independent interest (see [21], [53], [86]). As for a solution corresponding to the one-soliton, there is a natural choice (3.12). From the very beginning it is quite clear that it should contain *two* operator-valued parameters A, B, corresponding to the fact that for line-solitons the velocities and the angles with the x-axis are determined by independent sets of parameters.

In our personal work, the main difficulty was to prove the solution property without constraints on A, B. In a joint article with B. Carl, we obtained the result under the additional assumption $[A, B] = 0$ which in particular means that A, B map between the *same* space. Then the verification of the solution property is a lengthy but straightforward calculation (confer Appendix C). Without $[A, B] = 0$, the complexity explodes and seems to prevent further understanding. Nevertheless we got convinced of the truth of the general formula in discussions with A. Sakhnovich, who had in [86] discovered a related but different matrix solution with non-commuting parameters and had produced a proof by computer algebra.

Finally we succeeded in proving Theorem 3.2.1 without using computers. Our method to cut down the number of appearing terms dramatically consists in a systematic use of recursive relations, very close in spirit to those used for the AKNS system. Thereafter we follow the scalarization techniques familiar from Chapter 2 to obtain solution formulas in determinant form.

But then we will see that also solutions to the matrix KP can be extracted from the same operator solutions. To this end we consider solutions with values in the rank n operators with fixed kernel (the matrix KP being viewed as a system in the $n \times n$-matrices) and descent via a natural multiplicative evaluation map.

Hirota's method is one of the most important direct approaches in soliton theory. It relies on the examination of bilinear versions of the soliton equations. In Proposition 3.5.1 we will see that the determinant appearing in our solution formula (3.33) even satisfies the bilinear KP equation. The proof will give us the occasion to study the Miura transformation on the operator-level. Since our argument is independent from what we have done before, we also obtain an alternative elegant proof of Theorem 3.3.4.

We will conclude the chapter with a preliminary discussion of first examples, focusing on features which underline the difference to the one-dimensional case like line-solitons with parallel wavefront or identical shape. Perhaps the most interesting examples are resonant structures (so called Miles structures) which we obtain in a strikingly simple way. We plan to treat this in greater detail in a forthcoming publication.

3.1 The Kadomtsev-Petviashvili equations

The Kadomtsev-Petviashvili (KP) equation is one of the few soliton equations, which describe physical phenomena in two-dimensional space. It was introduced by Kadomtsev and Petviashvili [50] to discuss stability of one-dimensional solitons in a nonlinear media with weak dispersion.

There are several ways to write down the KP. Following [69], we shall use the system

$$u_t + 6uu_x + u_{xxx} = -3\alpha^2 w_y, \tag{3.1}$$

$$u_y = w_x, \tag{3.2}$$

for $\alpha \in \{1, \mathrm{i}\}$. In the case $\alpha = \mathrm{i}$, the system (3.1), (3.2) is called KP-I, and KP-II for $\alpha = 1$. In both cases the KP describes the propagation of shallow water waves. It depends on whether surface tension or gravitation dominates whether one arrives at KP-I or KP-II.

This is well reflected by certain peculiar solution classes. For the KP-II there are ordinary nonlinear superpositions of line-solitons (see [87], [102]), but also resonance phenomena are possible leading to waves with a tree shaped profile, the so-called Miles structures (see [61], [64], [70]). For the KP-I one can construct rational structures with particle character, the so-called lumps (see [4], [54]). Generically superpositions of lumps interact without phase-shift. But there is also weakly bound superposition, roughly comparable to negatons, which has recently aroused a lot of scientific interest. For results and further references the reader may consult [1], [2], [6], [36], [66], [76], [77].

For the sake of illustration, we mention the soliton solution. For both the KP-I and KP-II equations, it formally reads

$$\begin{aligned} u(x,y,t) &= 2\frac{\partial^2}{\partial x^2}\log\big(1+\ell(x,y,t)\big) \\ &\quad \text{with } \ell(x,y,t) = \exp\Big((a+b)x + \frac{1}{\alpha}(a^2-b^2)y - 4(a^3+b^3)t\Big), \end{aligned} \tag{3.3}$$

where $a, b \in \mathbb{C}$ are complex parameters. In the case of the KP-II equation the soliton becomes real for $a, b \in \mathbb{R}$, and is called line-soliton. Rewriting (3.3),

$$u(x,y,t) = \frac{(a+b)^2}{2}\cosh^{-2}\Big(\frac{1}{2}\Big((a+b)x + (a^2-b^2)y - 4(a^3+b^3)t\Big)\Big),$$

one observes that the line-soliton is a bell shaped wave front of infinite length. The parameters a, b characterize the height $(a+b)^2/2$ of the wave front and its angle φ, $\tan\varphi = -(a-b)^{-1}$ to the x-axis.

3.2 Solution of the operator-valued KP

In this section we consider the non-abelian Kadomtsev-Petviashvili equation,

$$U_t + 3\{U, U_x\} + U_{xxx} = -3\alpha^2 W_y - 3\alpha[U, W], \tag{3.4}$$

$$U_y = W_x, \tag{3.5}$$

where the two unknown functions $U = U(x,y,t)$, $W = W(x,y,t)$ take values in the bounded operators $\mathcal{L}(F)$ on some Banach space F. As usual $\{\cdot,\cdot\}$ denotes the anticommutator, $[\cdot,\cdot]$ the commutator.

Restricted to finite square matrices, the above system is known as matrix KP and is a topic of independent research (see [21], [53], [86]). Our intention is to study (3.4), (3.5) on the operator level both for its own sake and as a tool for the investigation of the scalar case.

The additional term $-3\alpha[U, W]$ in (3.4) can be motivated as follows. From [25], [102], we know that a Lax pair for the KP is

$$L = \frac{\partial^2}{\partial x^2} + \alpha\frac{\partial}{\partial y} + u, \tag{3.6}$$

$$B = \frac{\partial}{\partial t} + 4\frac{\partial^3}{\partial x^3} + 6u\frac{\partial}{\partial x} + 3u_x - 3\alpha w, \tag{3.7}$$

In other words, the scalar KP is obtained as the integrability condition $[L, B] = 0$. If we calculate $[L, B]$ without the assumption that the terms u, w in (3.6), (3.7) commute, we arrive at (3.4), (3.5). Note also that $-3\alpha[U, W]$ is harmless with respect to scalarization because it is annihilated by multiplicative functionals.

The following theorem is the main result of this section. It states an explicit operator-valued solution for the non-abelian KP (3.4), (3.5).

Theorem 3.2.1. *Let E, F be a Banach spaces, and $A \in \mathcal{L}(E)$, $B \in \mathcal{L}(F)$ arbitrary constant operators.*

Assume that $L = L(x, y, t) \in \mathcal{L}(F, E)$, $M = M(x, y, t) \in \mathcal{L}(E, F)$ are operator-valued functions which are C^4-smooth and solve the base equations

$$L_x = AL, \qquad L_y = \frac{1}{\alpha}A^2L, \qquad L_t = -4A^3L, \tag{3.8}$$

$$M_x = BM, \qquad M_y = -\frac{1}{\alpha}B^2M, \qquad M_t = -4B^3M. \tag{3.9}$$

Then, on $\Omega = \{(x, y, t) \in \mathbb{R}^3 \mid (I + LM) \text{ is invertible}\}$, a solution of the non-abelian KP (3.4), (3.5) is given by

$$U = 2V_x, \tag{3.10}$$

$$W = 2V_y, \tag{3.11}$$

where the operator-valued function $V = V(x, y, t) \in \mathcal{L}(F)$ is defined by

$$V = M(I + LM)^{-1}(AL + LB). \tag{3.12}$$

For the sake of comparison, we provide a direct proof of Theorem 3.2.1 for the case that the parameter operators A and B commute in Appendix C. This gives us an opportunity to show a result which was stated without proof in an earlier publication (Proposition 5.1 in [18]). In the commutative case, Theorem 3.2.1 follows from a lengthy but straightforward calculation.

In contrast a direct approach seems to be out of discussion for the non-commutative case. The reader should compare this to [86], where a related result was proved by computer algebra. One of our motivations was to find structural reasons behind such formulas which seem to be hopelessly complicated at first sight.

The idea of our argument is motivated by the techniques we used in the treatment of the AKNS system, to get hands on iterations of operators. Roughly speaking, we try to condense the calculation by reduction to terms with accessible 'reproduction properties'.

At first we observe that we can reduce the complexity of the proof of Theorem 3.2.1 by transition to an integrated version of the operator KP equation (3.4), (3.5).

Lemma 3.2.2. *If the operator-function $V = V(x,y,t) \in \mathcal{L}(F)$ is a solution of the integrated non-abelian KP,*

$$\Big(V_t + 6(V_x)^2 + V_{xxx}\Big)_x = -3\alpha^2 V_{yy} - 6\alpha[V_x, V_y], \tag{3.13}$$

then $U = 2V_x$, $W = 2V_y$ solve (3.4), (3.5).

Proof The proof is straightforward. Namely, by definition of U, W, we get

$$\begin{aligned} U_t + 3\{U, U_x\} + U_{xxx} &= 2\Big(V_t + 6(V_x)^2 + V_{xxx}\Big)_x \overset{(3.13)}{=} -6\alpha^2 V_{yy} - 12\alpha[V_x, V_y] \\ &= -3\alpha^2 W_y - 3\alpha[U, W], \end{aligned}$$

which is (3.4). (3.5) is obvious. □

As a second preparation, we provide some tools for the manipulation with some operator-valued functions.

Lemma 3.2.3. *Let E, F be Banach spaces and $L \in \mathcal{L}(F,E)$, $M \in \mathcal{L}(E,F)$ arbitrary operators such that the inverses $(I+LM)^{-1}$, $(I+ML)^{-1}$ exist. Then the following identities hold:*

$$(I+LM)^{-1}L = L(I+ML)^{-1}, \tag{3.14}$$

$$M(I+LM)^{-1}L = I - (I+ML)^{-1}. \tag{3.15}$$

The correponding identities with the roles of L, M exchanged hold, too.

Proof To verify (3.14), we use $L(I+ML) = (I+LM)L$ and multiply it by $(I+ML)^{-1}$ from the left and $(I+LM)^{-1}$ from the right. As a consequence,

$$\begin{aligned} M(I+LM)^{-1}L &= ML(I+ML)^{-1} \\ &= \Big((I+ML) - I\Big)(I+ML)^{-1} \\ &= I - (I+ML)^{-1} \end{aligned}$$

which is (3.15). □

The next lemma furnishes the recursive identities which allow to cut down the proof of Theorem 3.2.1 to a reasonable size.

Lemma 3.2.4. *Let E, F be Banach spaces and $A \in \mathcal{L}(E)$, $B \in \mathcal{L}(F)$ constant operators.*

Let $L = L(x_1, x_2, \dots) \in \mathcal{L}(F,E)$ and $M = M(x_1, x_2, \dots) \in \mathcal{L}(E,F)$ be operator-valued functions, depending on infinitely many real variables x_j, $j \in \mathbb{N}$, which are differentiable with respect to x_j for all j and satisfy the base equations

$$L_{x_j} = A^j L \qquad \text{and} \qquad M_{x_j} = -(-B)^j M \qquad \text{for all } j \in \mathbb{N},$$

and assume that $(I+LM)$, $(I+ML)$ are always invertible.

Define, for $j \in \mathbb{N}$, the following operator-valued functions

$$\begin{aligned} V_j &= (I+LM)^{-1}\Big(A^j L - L(-B)^j\Big) & \widehat{V}_j &= (I+LM)^{-1}\Big(A^j + L(-B)^j M\Big) \\ W_j &= (I+ML)^{-1}\Big(-(-B)^j M + MA^j\Big) & \widehat{W}_j &= (I+ML)^{-1}\Big(-(-B)^j - MA^j L\Big) \end{aligned}$$

Then the following derivation rules hold for all $i, j \in \mathbb{N}$:

$$V_{j,x_i} = \widehat{V}_i V_j, \tag{3.16}$$

$$\widehat{V}_{j,x_i} = -V_i W_j, \tag{3.17}$$

$$W_{j,x_i} = \widehat{W}_i W_j, \tag{3.18}$$
$$\widehat{W}_{j,x_i} = -W_i V_j. \tag{3.19}$$

Moreover, the following identities hold for all $i, j \in \mathbb{N}$:

$$\widehat{V}_i V_j - V_i \widehat{W}_j = V_{i+j}, \tag{3.20}$$
$$\widehat{V}_i \widehat{V}_j + V_i W_j = \widehat{V}_{i+j}, \tag{3.21}$$
$$\widehat{W}_i W_j - W_i \widehat{V}_j = -W_{i+j}, \tag{3.22}$$
$$\widehat{W}_i \widehat{W}_j + W_i V_j = -\widehat{W}_{i+j}. \tag{3.23}$$

Of course, a version in three variables would be sufficient for our applications. But the use of infinitely many variables is notationally more convenient and shows better the recursive structure. Note also the structural similarity of Lemma 3.2.4 to the Lemmas 1.2.5, 1.2.6.

Proof We start with (3.16). Using Lemma 1.2.4 and the base equations, we obtain

$$\begin{aligned}
V_{j,x_i} &= -(I+LM)^{-1}(LM)_{x_i}(I+LM)^{-1}\Big(A^j L - L(-B)^j\Big) \\
&\qquad\qquad +(I+LM)^{-1}\Big(A^j L - L(-B)^j\Big)_{x_i} \\
&= -(I+LM)^{-1}\Big((A^i L - L(-B)^i)M\Big)(I+LM)^{-1}\Big(A^j L - L(-B)^j\Big) \\
&\qquad\qquad +(I+LM)^{-1}A^i\Big(A^j L - L(-B)^j\Big) \\
&= (I+LM)^{-1}\Big(-(A^i L - L(-B)^i)M + A^i(I+LM)\Big)\, V_j \\
&= (I+LM)^{-1}\Big(A^i + L(-B)^i M\Big)\, V_j \\
&= \widehat{V}_i\, V_j
\end{aligned}$$

and, analogously,

$$\begin{aligned}
\widehat{V}_{j,x_i} &= -(I+LM)^{-1}(LM)_{x_i}(I+LM)^{-1}\Big(A^j + L(-B)^j M\Big) \\
&\qquad\qquad +(I+LM)^{-1}\Big(A^j + L(-B)^j M\Big)_{x_i} \\
&= -(I+LM)^{-1}\Big((A^i L - L(-B)^i)M\Big)(I+LM)^{-1}\Big(A^j + L(-B)^j M\Big) \\
&\qquad\qquad +(I+LM)^{-1}\Big((A^i L - L(-B)^i)(-B)^j M\Big) \\
&= V_i\left(-M(I+LM)^{-1}\Big(A^j + L(-B)^j M\Big) + (-B)^j M\right) \\
&= V_i\,(I+ML)^{-1}\left(-M\Big(A^j + L(-B)^j M\Big) + (I+ML)(-B)^j M\right) \\
&= V_i\,(I+ML)^{-1}\Big(-\big(MA^j - (-B)^j M\big)\Big) \\
&= -V_i\, W_j,
\end{aligned}$$

where we have used Lemma 3.2.3 for the fourth identity.

This shows (3.17). The identities (3.18), (3.19) follow by exchanging the roles of L and M, A^j and $-(-B)^j$.

Now we turn to (3.20). First we calculate

$$(I+LM)\,\widehat{V}_i V_j = \Big(A^i + L(-B)^i M\Big)(I+LM)^{-1}\Big(A^j L - L(-B)^j\Big)$$

$$
\begin{aligned}
= \quad & -L(-B)^i\Big(M(I+LM)^{-1}L\Big)(-B)^j + A^i\Big((I+LM)^{-1}\Big)A^jL \\
& \qquad -A^i\Big((I+LM)^{-1}L\Big)(-B)^j + L(-B)^i\Big(M(I+LM)^{-1}\Big)A^jL.
\end{aligned}
$$

Next we apply Lemma 3.2.3 to the large brackets in order to change $(I+LM)^{-1}$ into $(I+ML)^{-1}$ wherever it appears. We thus get

$$
\begin{aligned}
& (I+LM)\,\widehat{V}_iV_j \\
= \quad & -L(-B)^i\Big(I-(I+ML)^{-1}\Big)(-B)^j + A^i\Big(I-L(I+ML)^{-1}M\Big)A^jL \\
& \qquad -A^i\Big(L(I+ML)^{-1}\Big)(-B)^j + L(-B)^i\Big((I+ML)^{-1}M\Big)A^jL \\
= \quad & \Big(A^{i+j}L - L(-B)^{i+j}\Big) - \Big(A^iL - L(-B)^i\Big)(I+ML)^{-1}\Big((-B)^j + MA^jL\Big) \\
= \quad & (I+LM)V_{i+j} + (I+LM)V_i\widehat{W}_j,
\end{aligned}
$$

which shows (3.20). Analogously, we get

$$
\begin{aligned}
& (I+LM)\,\widehat{V}_i\widehat{V}_j = \Big(A^i + L(-B)^iM\Big)(I+LM)^{-1}\Big(A^j + L(-B)^jM\Big) \\
= \quad & L(-B)^i\Big(M(I+LM)^{-1}L\Big)(-B)^jM + A^i\Big((I+LM)^{-1}\Big)A^j \\
& \qquad +L(-B)^i\Big(M(I+LM)^{-1}\Big)A^j + A^i\Big((I+LM)^{-1}L\Big)(-B)^jM \\
= \quad & L(-B)^i\Big(I-(I+ML)^{-1}\Big)(-B)^jM + A^i\Big(I-L(I+ML)^{-1}M\Big)A^j \\
& \qquad +L(-B)^i\Big((I+ML)^{-1}M\Big)A^j + A^i\Big(L(I+ML)^{-1}\Big)(-B)^jM \\
= \quad & \Big(A^{i+j} + L(-B)^{i+j}M\Big) - \Big(A^iL - L(-B)^i\Big)(I+ML)^{-1}\Big(MA^j - (-B)^jM\Big) \\
= \quad & (I+LM)\widehat{V}_{i+j} - (I+LM)V_iW_j
\end{aligned}
$$

and thus (3.21) holds.

As for (3.22), (3.23), we repeat the arguments above. Note that in this case we cannot exchange the roles of A^j and $-(-B)^j$ because of the signs. We observe

$$
\begin{aligned}
& (I+ML)\,\widehat{W}_iW_j = -\Big((-B)^i + MA^iL\Big)(I+ML)^{-1}\Big(MA^j - (-B)^jM\Big) \\
= \quad & (-B)^i\Big((I+ML)^{-1}\Big)(-B)^jM - MA^i\Big(L(I+ML)^{-1}M\Big)A^j \\
& \qquad -(-B)^i\Big((I+ML)^{-1}M\Big)A^j + MA^i\Big(L(I+ML)^{-1}\Big)(-B)^jM \\
= \quad & (-B)^i\Big(I-M(I+LM)^{-1}L\Big)(-B)^jM - MA^i\Big(I-(I+LM)^{-1}\Big)A^j \\
& \qquad -(-B)^i\Big(M(I+LM)^{-1}\Big)A^j + MA^i\Big((I+LM)^{-1}L\Big)(-B)^jM \\
= \quad & \Big((-B)^{i+j}M - MA^{i+j}\Big) + \Big(MA^i - (-B)^iM\Big)(I+LM)^{-1}\Big(A^j + L(-B)^jM\Big) \\
= \quad & -(I+ML)W_{i+j} + (I+ML)W_i\widehat{V}_j,
\end{aligned}
$$

yielding (3.22), and (3.23) follows from

$$
\begin{aligned}
& (I+ML)\,\widehat{W}_i\widehat{W}_j = \Big((-B)^i + MA^iL\Big)(I+ML)^{-1}\Big((-B)^j + MA^jL\Big) \\
= \quad & (-B)^i\Big((I+ML)^{-1}\Big)(-B)^j + MA^i\Big(L(I+ML)^{-1}M\Big)A^jL \\
& \qquad +(-B)^i\Big((I+ML)^{-1}M\Big)A^jL + MA^i\Big(L(I+ML)^{-1}\Big)(-B)^j
\end{aligned}
$$

$$
\begin{aligned}
&= (-B)^i\Big(I - M(I+LM)^{-1}L\Big)(-B)^j + MA^i\Big(I-(I+LM)^{-1}\Big)A^jL \\
&\qquad\qquad +(-B)^i\Big(M(I+LM)^{-1}\Big)A^jL + MA^i\Big((I+LM)^{-1}L\Big)(-B)^j \\
&= \Big((-B)^{i+j} + MA^{i+j}L\Big) - \Big(MA^i - (-B)^iM\Big)(I+LM)^{-1}\Big(A^jL - L(-B)^j\Big) \\
&= -(I+ML)\widehat{W}_{i+j} - (I+ML)W_iV_j.
\end{aligned}
$$

This completes the proof. □

Now we are in position to give the proof of the main result of this section.

Proof (of Theorem 3.2.1) By Lemma 3.2.2 it suffices to show that

$$V = M(I+LM)^{-1}(AL+LB)$$

solves (3.13). To unify the calculations, we use the coordinate transformation

$$x_1 = x, \qquad x_2 = \frac{1}{\alpha}y, \qquad x_3 = -4t, \tag{3.24}$$

under which the base equations (3.8), (3.9) become $L_{x_i} = A^iL$ and $M_{x_i} = -(-B)^iM$ for $i = 1,2,3$. Then it remains to show that $V = V(x_1,x_2,x_3)$ solves

$$\Big(V_{x_1x_1x_1} + 6(V_{x_1})^2 - 4V_{x_3}\Big)_{x_1} = -3V_{x_2x_2} - 6[V_{x_1}, V_{x_2}]. \tag{3.25}$$

To start with the proof, we remark the connection between V and the operator-functions used in Lemma 3.2.4. Namely, by Lemma 3.2.3,

$$
\begin{aligned}
V &= M(I+LM)^{-1}(AL+LB) \\
&= (I+ML)^{-1}M(AL+LB) \\
&= (I+ML)^{-1}\ \big(\ (MAL-B) + (I+ML)B\ \big) \\
&= -\widehat{W}_1 + B
\end{aligned}
$$

Note also that V, V_1 do not denote the same operator-function.

First we use Lemma 3.2.4 to calculate the derivatives of V. By (3.19), we get, for $j = 1,2,3$,

$$V_{x_j} = -\widehat{W}_{1,x_j} = W_jV_1. \tag{3.26}$$

Using again Lemma 3.2.4, namely (3.16), (3.18), and then (3.22), we find, for all $j \in \mathbb{N}$ and $i = 1,2,3$,

$$
\begin{aligned}
(W_jV_1)_{x_i} &= W_{j,x_i}V_1 + W_jV_{1,x_i} \\
&= (\widehat{W}_iW_j + W_j\,\widehat{V}_i)V_1 \\
&= -W_{i+j}V_1 + (W_i\widehat{V}_j + W_j\widehat{V}_i)V_1.
\end{aligned} \tag{3.27}
$$

In particular, for $j = 1,2,3$,

$$V_{x_jx_j} = -W_{2j}V_1 + 2\,W_j\widehat{V}_jV_1. \tag{3.28}$$

Next we calculate the derivatives of the operators in (3.27). Since we do only need higher derivatives with respect to the variable x_1, we restrict to this case. Starting from (3.16), (3.17), (3.18), we observe for $i,j \in \mathbb{N}$,

$$
\begin{aligned}
&(W_i\widehat{V}_jV_1)_{x_1} = W_{i,x_1}\widehat{V}_jV_1 + W_i\widehat{V}_{j,x_1}V_1 + W_i\widehat{V}_jV_{1,x_1} \\
&= (\widehat{W}_1W_i)\widehat{V}_jV_1 - W_iV_1W_jV_1 + W_i(\widehat{V}_j\widehat{V}_1)V_1
\end{aligned}
$$

$$
\begin{aligned}
&= (W_1\widehat{V}_i - W_{i+1})\widehat{V}_j V_1 - W_i V_1 W_j V_1 + W_i(\widehat{V}_{j+1} - V_j W_1)V_1 \\
&= (W_i\widehat{V}_{j+1} - W_{i+1}\widehat{V}_j)V_1 + \Big(W_1(\widehat{V}_i\widehat{V}_j) - W_iV_1W_j - W_iV_jW_1\Big)V_1 \\
&= \Big(W_i\widehat{V}_{j+1} - W_{i+1}\widehat{V}_j + W_1\widehat{V}_{i+j}\Big)V_1 - \Big(W_1V_iW_j + W_iV_1W_j + W_iV_jW_1\Big)V_1
\end{aligned}
$$

where we have applied (3.21), (3.22) to the terms in the brackets of the second, and (3.21) to the term in the brackets of the fourth identity. In particular,

$$
\begin{aligned}
(W_1\widehat{V}_1V_1)_{x_1} &= -W_2\widehat{V}_1V_1 + 2W_1\widehat{V}_2V_1 - 3W_1V_1W_1V_1, \\
(W_1\widehat{V}_2V_1 - W_2\widehat{V}_1V_1)_{x_1} &= (W_1\widehat{V}_3 + W_3\widehat{V}_1)V_1 - 2W_2\widehat{V}_2V_1 \\
&\qquad +2(W_2V_1W_1V_1 - W_1V_1W_2V_1).
\end{aligned}
$$

Using (3.26), (3.28), we can rewrite those equations as

$$
(W_1\widehat{V}_1V_1)_{x_1} = -W_2\widehat{V}_1V_1 + 2W_1\widehat{V}_2V_1 - 3(V_{x_1})^2, \tag{3.29}
$$

$$
(W_1\widehat{V}_2V_1 - W_2\widehat{V}_1V_1)_{x_1} = -W_4V_1 + (W_1\widehat{V}_3 + W_3\widehat{V}_1)V_1 - V_{x_2x_2} - 2[V_{x_1}, V_{x_2}]. \tag{3.30}
$$

Now, starting from (3.28), and then applying (3.27), (3.29), we calculate

$$
\begin{aligned}
V_{x_1x_1x_1} &= -(W_2V_1)_{x_1} + 2(W_1\widehat{V}_1V_1)_{x_1} \\
&= \Big(W_3V_1 - (W_1\widehat{V}_2 + W_2\widehat{V}_1)V_1\Big) - 2\Big(W_2\widehat{V}_1V_1 - 2W_1\widehat{V}_2V_1 + 3(V_{x_1})^2\Big) \\
&= W_3V_1 + 3(W_1\widehat{V}_2 - W_2\widehat{V}_1)V_1 - 6(V_{x_1})^2.
\end{aligned}
$$

Therefore, taking also (3.26) into account, we obtain

$$
V_{x_1x_1x_1} + 6(V_{x_1})^2 - 4V_{x_3} = -3W_3V_1 + 3(W_1\widehat{V}_2 - W_2\widehat{V}_1)V_1,
$$

which finally, by (3.27) and (3.30), yields

$$
\begin{aligned}
&\Big(V_{x_1x_1x_1} + 6(V_{x_1})^2 - 4V_{x_3}\Big)_{x_1} = -3(W_3V_1)_{x_1} + 3(W_1\widehat{V}_2V_1 - W_2\widehat{V}_1V_1)_{x_1} \\
&= -3\Big(-W_4V_1 + (W_1\widehat{V}_3 + W_3\widehat{V}_1)V_1\Big) \\
&\qquad\qquad +3\Big(-W_4V_1 + (W_1\widehat{V}_3 + W_3\widehat{V}_1)V_1 - V_{x_2x_2} - 2[V_{x_1}, V_{x_2}]\Big) \\
&= -3V_{x_2x_2} - 6[V_{x_1}, V_{x_2}].
\end{aligned}
$$

Thus (3.25) is proved. □

3.3 Derivation of solution formulas for the scalar KP

In this section the explicit solution of the operator-valued KP is used to construct solution formulas for the scalar KP.

3.3.1 The scalarization process

Let us briefly recall the idea for the scalarization process.

Starting from an operator-valued solution $U = U(x, y, t)$, $W = W(x, y, t) \in \mathcal{L}(F)$ of the non-abelian KP equation (3.4), (3.5), application of a functional τ yields scalar functions $u = \tau(U)$, $w = \tau(W)$. It is our purpose to obtain solutions of the scalar KP equation (3.1), (3.2) this way. Thus τ has to maintain the solution property. Since the KP equation is nonlinear, τ has to be multiplicative in a suitable sense.

The discussion in Chapter 2.1.2 motivates the following ansatz. Fix $a \in F'$ and choose (i) $U, W \in \mathcal{S}_a(F)$, (ii) ev_a as the canonical functional on $\mathcal{S}_a(F)$. Then multiplicativity of ev_a on $\mathcal{S}_a(F)$ is assured by Proposition 2.1.3.

Pursuing this strategy, we end up with the following result.

Theorem 3.3.1. *Let F be a Banach space and $a \in F'$ a constant functional.*
If $V = V(x, y, t) \in \mathcal{S}_a(F)$ is a family of bounded operators such that $U = V_x$, $W = V_y$ solve the non-abelian KP equation (3.4), (3.5), then

$$\begin{aligned} u &= v_x, \\ w &= v_y, \end{aligned}$$

where the function $v = v(x, y, t)$ is defined by

$$v(x, y, t) = \mathrm{ev}_a\big(V(x, y, t)\big),$$

solve the scalar KP equation (3.1), (3.2).

Proof First observe that $V \in \mathcal{S}_a(F)$ implies $U, W \in \mathcal{S}_a(F)$. Thus, the non-abelian KP (3.4), (3.5) is an operator equation in the Banach algebra $\mathcal{S}_a(F)$. By Proposition 2.1.3, on this Banach algebra ev_a defines a multiplicative functional.

Application of ev_a to the non-abelian KP (3.4), (3.5) yields successively, by i) linearity, ii) multiplicativity, and iii) continuity,

$$\begin{aligned} 0 &= \mathrm{ev}_a\Big(U_t + 3\{U, U_x\} + U_{xxx} + 3\alpha^2 W_y + 3\alpha[U, W]\Big) \\ &\overset{\text{i)}}{=} \mathrm{ev}_a(U_t) + 3\,\mathrm{ev}_a\,\big(\ \{U, U_x\}\ \big) + \mathrm{ev}_a(U_{xxx}) + 3\alpha^2\,\mathrm{ev}_a(W_y) + 3\alpha\,\mathrm{ev}_a\,\big(\ [U, W]\ \big) \\ &\overset{\text{ii)}}{=} \mathrm{ev}_a(U_t) + 3\,\big\{\ \mathrm{ev}_a(U), \mathrm{ev}_a(U_x)\ \big\} + \mathrm{ev}_a(U_{xxx}) + 3\alpha^2\,\mathrm{ev}_a(W_y) + 3\alpha\,\big[\ \mathrm{ev}_a(U), \mathrm{ev}_a(W)\ \big] \\ &= \mathrm{ev}_a(U_t) + 6\,\mathrm{ev}_a(U)\,\mathrm{ev}_a(U_x) + \mathrm{ev}_a(U_{xxx}) + 3\alpha^2\,\mathrm{ev}_a(W_y) \\ &\overset{\text{iii)}}{=} \mathrm{ev}_a(U)_t + 6\,\mathrm{ev}_a(U)\,\mathrm{ev}_a(U)_x + \mathrm{ev}_a(U)_{xxx} + 3\alpha^2\,\mathrm{ev}_a(W)_y \\ &= u_t + 6uu_x + u_{xxx} + 3\alpha^2 w_y \end{aligned}$$

with $u = \mathrm{ev}_a(U) = \mathrm{ev}_a(V_x) = \mathrm{ev}_a(V)_x = v_x$, $w = \mathrm{ev}_a(W) = \mathrm{ev}_a(V_y) = \mathrm{ev}_a(V)_y = v_y$. This shows (3.1), and (3.2) follows analogously. $\square$

There is a slightly different way to see Theorem 3.3.1 by using duality directly. Although it is less economic, we will discuss it because it contains a clue for the scalarization leading to the matrix KP equation.

Since $V = V(x, y, t) \in \mathcal{S}_a(F)$, there is a vector function $c = c(x, y, t)$ such that $V = a \otimes c$. Consequently,

$$U = a \otimes d, \qquad W = a \otimes f$$

with $d = c_x$ and $f = c_y$. By the calculation rules for one-dimensional operators (see Section 2.1.1) this means that the operator-valued KP equation reads

$$\begin{aligned} &a \otimes d_t + 3(\langle d_x, a\rangle a \otimes d + \langle d, a\rangle a \otimes d_x) + a \otimes d_{xxx} \\ &\qquad\qquad = -3\alpha^2 a \otimes f_y - 3\alpha(\langle f, a\rangle a \otimes d - \langle d, a\rangle a \otimes f), \\ &a \otimes f_x = a \otimes d_y. \end{aligned}$$

Take $c_0 \in F$ such that $\langle c_0, a\rangle = 1$. Applying the operator equation to c_0, we get the vector equation

$$
\begin{aligned}
&d_t + 3(\langle d_x, a\rangle d + \langle d, a\rangle d_x) + d_{xxx} = -3\alpha^2 f_y - 3\alpha(\langle f, a\rangle d - \langle d, a\rangle f),\\
&f_x = d_y.
\end{aligned}
$$

Applying the functional a again, end up with the scalar equation

$$
\begin{aligned}
\langle d_t, a\rangle + 6\langle d, a\rangle\langle d_x, a\rangle + \langle d_{xxx}, a\rangle &= -3\alpha^2\langle f_y, a\rangle,\\
\langle f_x, a\rangle &= \langle d_y, a\rangle.
\end{aligned}
$$

Since the functional a is continuous, we can interchange the order of derivation and the application of a. Therefore, the functions $u = \langle d, a\rangle$, $w = \langle f, a\rangle$ solve the scalar KP equation (3.1), (3.2).

To complete the argument, observe $u = \langle d, a\rangle = \langle c_x, a\rangle = \langle c, a\rangle_x = v_x$, $w = \langle f, a\rangle = \langle c_y, a\rangle = \langle c, a\rangle_y = v_y$, with $v = \langle c, a\rangle$.

3.3.2 Solution formulas in terms of determinants

Next we carry out the scalarization process for the concrete operator soliton derived in Theorem 3.2.1. Moreover, we follow the line of arguments of Chapter 2 to derive compact solution formulas in terms of determinants.

We start with the following result which we obtain by imposing appropriate conditions such that the operator soliton in Theorem 3.2.1 becomes one-dimensional.

Proposition 3.3.2. *Let E, F be Banach spaces and $A \in \mathcal{L}(E)$, $B \in \mathcal{L}(F)$, $D \in \mathcal{L}(E, F)$ arbitrary constant operators. Assume in addition that the constant operator $C \in \mathcal{L}(F, E)$ satisfies the one-dimensionality condition*

$$AC + CB = a \otimes c \tag{3.31}$$

for $0 \neq a \in F'$, $c \in E$. Define the operator-functions

$$
\begin{aligned}
L(x,y,t) &= \widehat{L}(x,y,t)C, & \widehat{L}(x,y,t) &= \exp\Big(\, Ax + \frac{1}{\alpha}A^2 y - 4A^3 t \,\Big),\\
M(x,y,t) &= \widehat{M}(x,y,t)D, & \widehat{M}(x,y,t) &= \exp\Big(\, Bx - \frac{1}{\alpha}B^2 y - 4B^3 t \,\Big)\,.
\end{aligned}
$$

Then, on $\Omega = \{(x,y,t)\,|\,(I + LM)$ is invertible$\}$, a solution of the scalar KP equation (3.1), (3.2) is given by

$$
\begin{aligned}
u &= v_x,\\
w &= v_y,
\end{aligned}
$$

where

$$v = 2\,\mathrm{tr}\Big(M\ \big(\ I + LM\ \big)^{-1}\ \widehat{L}(a \otimes c)\Big). \tag{3.32}$$

Remark 3.3.3. *We want to emphasize that the one-dimensionality condition does not involve the operator D. Thus, the choice of D is completely free!*

Proof Obviously the operator functions $L = L(x,y,t) \in \mathcal{L}(F,E)$, $M = M(x,y,t) \in \mathcal{L}(E,F)$ solve the base equations in Theorem 3.3.1. Hence Theorem 3.2.1 provides us with the solution $U = V_x$, $W = V_y$ of the non-abelian KP equation (3.4), (3.5) on Ω, where

$$V \;=\; 2M(I + LM)^{-1}(AL + LB).$$

By (3.31),

$$\begin{aligned} V &= 2M(I+LM)^{-1}\widehat{L}(AC+CB) \\ &= 2M(I+LM)^{-1}\widehat{L}(a\otimes c) \\ &= a\otimes\big(\,2M(I+LM)^{-1}\widehat{L}c\,\big)\,. \end{aligned}$$

Thus $V(x,y,t) = a\otimes\big(\,\widehat{c}(x,y,t)\,\big)$ with the vector function $\widehat{c} = 2M(I+LM)^{-1}\widehat{L}\,c \in F$. In particular, $V \in \mathcal{S}_a(F)$.

Now we are in the position to apply Theorem 3.3.1. This shows that on Ω a solution of the scalar KP equation (3.1), (3.2) is given by $u = v_x$, $w = v_y$ with

$$\begin{aligned} v &= \mathrm{ev}_a\big(a\otimes\widehat{c}\big) \;=\; \langle\widehat{c},a\rangle \;=\; \mathrm{tr}\big(a\otimes\widehat{c}\big) \\ &= \mathrm{tr}\Big(a\otimes\big(2M(I+LM)^{-1}\widehat{L}\,c\big)\Big) \\ &= 2\,\mathrm{tr}\Big(M(I+LM)^{-1}\widehat{L}(a\otimes c)\Big). \end{aligned}$$

This is the desired solution formula. $\square$

Our next aim is to improve the solution formula (3.32). Exploiting the theory of traces and determinants on quasi-Banach ideals, we can get rid of expressions containing inverse operators like $(I+LM)^{-1}$. The main result of the section is the following.

Theorem 3.3.4. *Let E, F be Banach spaces, and let $A \in \mathcal{L}(E)$, $B \in \mathcal{L}(F)$, $D \in \mathcal{L}(E,F)$ arbitrary constant operators.*

a) Assume that the one-dimensionality condition (3.31) is satisfied with $C \in \mathcal{A}(F,E)$ and $0 \neq a \in F'$, where $\mathcal{A}$ is an arbitrary quasi-Banach ideal admitting a continuous determinant δ. Define

$$\begin{aligned} L(x,y,t) &= \widehat{L}(x,y,t)\,C, \qquad \widehat{L}(x,y,t) = \exp(Ax + \frac{1}{\alpha}A^2y - 4A^3t), \\ M(x,y,t) &= \widehat{M}(x,y,t)D, \qquad \widehat{M}(x,y,t) = \exp(Bx - \frac{1}{\alpha}B^2y - 4B^3t). \end{aligned}$$

Then, on $\Omega = \{(x,y,t)\,|\,\delta(I+LM)\neq 0\}$, a solution of the scalar KP (3.1), (3.2) is given by

$$\begin{aligned} u &= v_x, \\ w &= v_y, \end{aligned}$$

where

$$v = 2\frac{\partial}{\partial x}\log\delta(I+LM). \tag{3.33}$$

b) If $0 \notin \mathrm{spec}(A) + \mathrm{spec}(B)$, then a) holds with $C = \Phi_{A,B}^{-1}(a\otimes c)$, $a \neq 0$, and any p-Banach operator ideal $\mathcal{A}$ admitting a continuous determinant δ.

Proof Recall that the solution u, w of the KP in Proposition 3.3.2 is determined by the function v given in (3.32). We show that v can be written in the form (3.33).

To this end we denote by τ the trace on $\mathcal{A}$ which corresponds to the determinant δ according to the Trace-determinant theorem (confer Proposition B.2.11). Since all traces coincide on the ideal $\mathcal{F}$ of finite-rank operators, we have $\tau_{\mathcal{F}} = \mathrm{tr}$, and thus

$$\begin{aligned} v &= 2\,\mathrm{tr}\Big(M(I+LM)^{-1}\widehat{L}(a\otimes c)\Big) \\ &= 2\,\tau\Big(M(I+LM)^{-1}\widehat{L}(a\otimes c)\Big) \\ &= 2\,\tau\Big(M(I+LM)^{-1}(AL+LB)\Big) \\ &= 2\,\tau\Big((I+LM)^{-1}(AL+LB)M\Big), \end{aligned}$$

the latter by the trace property. By the base equations $L_x = AL$, $M_x = BM$, this shows

$$v \;=\; 2\,\tau\Big((I+LM)^{-1}(LM)_x\Big).$$

Since C, and thus LM, belongs to the quasi-Banach ideal $\mathcal{A}$ by assumption, we can exploit the differentiation rule for determinants on quasi-Banach ideals, see Proposition B.2.12, which yields

$$v \;=\; 2\,\frac{\partial}{\partial x}\log\delta(I+LM).$$

Therefore, part a) is shown. Part b) is an immediate consequence of Proposition 2.3.4. $\square$

3.4 From the operator-valued to the matrix KP

As its scalar analogue, the matrix KP is a soliton equation, and one is interested in integrability techniques, Lax pairs, explicit solutions and so on. For results in this direction, we refer to [21], [53], [86].

In the present section we show that the operator-valued KP equation can be also used to construct solutions of the matrix KP.

We proceed as follows. First we provide the necessary algebraic tools. Then we extend the scalarization process to the matrix KP equation and carry it out for the operator soliton derived in Theorem 3.2.1. Finally we give some ameliorations of the solution formula.

3.4.1 Algebraic tools for the scalarization process to matrix-valued soliton equations

To start with, we introduce the vector space $\mathcal{S}_{a_1,\dots,a_n}(E,F)$ consisting of all operators with range at most n between the Banach spaces E and F whose behaviour is governed by the functionals $a_1,\dots,a_n \in E'$.

Definition 3.4.1. *For linearly independent $a_1,\dots,a_n \in E'$, we define the vector space*

$$\mathcal{S}_{a_1,\dots,a_n}(E,F) = \Big\{\sum_{j=1}^{n} a_j \otimes y_j \;\Big|\; y_1,\dots,y_n \in F\Big\}.$$

and we write $\mathcal{S}_{a_1,\dots,a_n}(E)$ instead of $\mathcal{S}_{a_1,\dots,a_n}(E,E)$.

One could also introduce for some subspace $K \subset E$ of finite codimension $\mathcal{S}_K$ as the space of all $T \in \mathcal{L}(E,F)$ with $\ker(T) \supset K$. We have $\mathcal{S}_K(E,F) = \mathcal{S}_{a_1,\dots,a_m}(E,F)$ iff $K = \bigcap_{j=1}^{m}\ker(a_j)$ and $\operatorname{codim}(K) = m$. We prefer the less invariant notation $\mathcal{S}_{a_1,\dots,a_m}(E,F)$ for notational convenience.

Lemma 3.4.2. *Any $T \in \mathcal{S}_{a_1,\dots,a_n}(E,F)$ has a unique representation $T = \sum_{j=1}^{n} a_j \otimes y_j$ with $y_1,\dots,y_n \in F$.*

Lemma 3.4.3. *$\mathcal{S}_{a_1,\dots,a_n}(E,F)$ is closed under multiplication from the left with operators in $\mathcal{L}(F)$. In particular, $\mathcal{S}_{a_1,\dots,a_n}(E)$ is a left ideal in $\mathcal{L}(E)$.*

Moreover, $\mathcal{S}_{a_1,\dots,a_n}(E)$ is a Banach algebra.

Once $a_1,\dots,a_n$ are given, there is a continuous linear map which naturally relates n-dimensional operators on a Banach space E to $n\times n$-matrices with complex entries. Namely, we define the map $\sigma : \mathcal{S}_{a_1,\dots,a_n}(E) \to \mathcal{M}_{n,n}(\mathbb{C})$ by

$$\sigma\Big(\sum_{j=1}^n a_j \otimes y_j\Big) = \Big(\operatorname{tr}(a_i \otimes y_j)\Big)_{i,j=1}^n. \tag{3.34}$$

Of course $\sigma = \sigma_{a_1,\dots,a_n}$. In the sequel we do not mention the dependence on the functionals $a_1, \dots, a_n$ if it is clear from the context.

Proposition 3.4.4. *On $\mathcal{S}_{a_1,\dots,a_n}(E)$, the map σ is an algebra homomorphism.*

Proof It suffices to show multiplicativity. Let T, $S \in \mathcal{S}_{a_1,\dots,a_n}(E)$ be operators governed by the same functionals $a_1, \dots, a_n \in E'$, where $T = \sum_{j=1}^n a_j \otimes y_j$ and $S = \sum_{j=1}^n a_j \otimes z_j$ with $y_1, \dots, y_n$, and $z_1, \dots, z_n \in E$. By the multiplication rule for one-dimensional operators,

$$\begin{aligned}
\sigma(TS) &= \sigma\Big(\sum_{k=1}^n a_k \otimes y_k \cdot \sum_{j=1}^n a_j \otimes z_j\Big) = \sigma\Bigg(\sum_{j=1}^n a_j \otimes \Big(\sum_{k=1}^n \langle z_j, a_k\rangle y_k\Big)\Bigg) \\
&= \Bigg(\operatorname{tr}\Big(a_i \otimes \big(\sum_{k=1}^n \langle z_j, a_k\rangle y_k\big)\Big)\Bigg)_{i,j=1}^n = \Bigg(\Big\langle \sum_{k=1}^n \langle z_j, a_k\rangle y_k, a_i\Big\rangle\Bigg)_{i,j=1}^n \\
&= \Bigg(\sum_{k=1}^n \langle y_k, a_i\rangle \langle z_j, a_k\rangle\Bigg)_{i,j=1}^n = \Big(\langle y_j, a_i\rangle\Big)_{i,j=1}^n \Big(\langle z_j, a_i\rangle\Big)_{i,j=1}^n \\
&= \Big(\operatorname{tr}(a_i \otimes y_j)\Big)_{i,j=1}^n \Big(\operatorname{tr}(a_i \otimes z_j)\Big)_{i,j=1}^n = \sigma\Big(\sum_{j=1}^n a_j \otimes y_j\Big)\sigma\Big(\sum_{j=1}^n a_j \otimes z_j\Big) \\
&= \sigma(T)\sigma(S).
\end{aligned}$$

$\square$

Lemma 3.4.5. *Let $a_1, \dots, a_n \in E'$ and $\widehat{a}_1, \dots, \widehat{a}_n \in E'$ be two different sets of linearly independent functionals with $\mathcal{S}_{a_1,\dots,a_n}(E) = \mathcal{S}_{\widehat{a}_1,\dots,\widehat{a}_n}(E)$. Then the corresponding maps σ and $\widehat{\sigma}$ are gauge equivalent.*

Two maps ϕ, $\psi : V \to \mathcal{M}_{n,n}(\mathbb{C})$ are called gauge equivalent iff there is an invertible $A \in \mathcal{M}_{n,n}(\mathbb{C})$ such that $\phi(v) = A^{-1}\psi(v)A$ for all $v \in V$.

Proof Since $\mathcal{S}_{a_1,\dots,a_n}(E) = \mathcal{S}_{\widehat{a}_1,\dots,\widehat{a}_n}(E)$, there are $\lambda_{jk} \in \mathbb{C}$ such that $\widehat{a}_j = \sum_{k=1}^n \lambda_{jk} a_k$, and we set $\Lambda = \big(\lambda_{jk}\big)_{j,k=1}^n$. Of course Λ is invertible.

Now let $T \in \mathcal{S}_{a_1,\dots,a_n}(E) = \mathcal{S}_{\widehat{a}_1,\dots,\widehat{a}_n}(E)$ be given arbitrarily with its corresponding representations $T = \sum_{j=1}^n a_j \otimes y_j = \sum_{j=1}^n \widehat{a}_j \otimes \widehat{y}_j$. Then

$$\sum_{j=1}^n \widehat{a}_j \otimes \widehat{y}_j = \sum_{j=1}^n \Big(\sum_{k=1}^n \lambda_{jk} a_k\Big) \otimes \widehat{y}_j = \sum_{k=1}^n a_k \otimes \Big(\sum_{j=1}^n \lambda_{jk} \widehat{y}_j\Big),$$

and since the representation of T in $\mathcal{S}_{a_1,\dots,a_n}(E)$ is unique, $y_k = \sum_{j=1}^n \lambda_{jk}\widehat{y}_j$.

Therefore,

$$\begin{aligned}
\widehat{\sigma}(T)\Lambda &= \Bigg(\sum_{k=1}^n \operatorname{tr}(\widehat{a}_i \otimes \widehat{y}_k) \cdot \lambda_{kj}\Bigg)_{i,j=1}^n = \Bigg(\operatorname{tr}\Big(\widehat{a}_i \otimes \big(\sum_{k=1}^n \lambda_{kj}\widehat{y}_k\big)\Big)\Bigg)_{i,j=1}^n \\
&= \Big(\operatorname{tr}(\widehat{a}_i \otimes y_j)\Big)_{i,j=1}^n = \Bigg(\operatorname{tr}\Big(\big(\sum_{k=1}^n \lambda_{ik} a_k\big) \otimes y_j\Big)\Bigg)_{i,j=1}^n \\
&= \Bigg(\sum_{k=1}^n \lambda_{ik} \cdot \operatorname{tr}(a_k \otimes y_j)\Bigg)_{i,j=1}^n = \Lambda\sigma(T)
\end{aligned}$$

which shows the assertion. $\square$

3.4.2 Scalarization from a different point of view

Next we use the extended scalarization process to obtain solutions of the matrix KP equation. For fixed dimension $n \in \mathbb{N}$, this is the system

$$u_t + 3\{u, u_x\} + u_{xxx} = -3\alpha^2 w_y - 3\alpha[u, w], \tag{3.35}$$

$$w_x = u_y, \tag{3.36}$$

taking its values in $\mathcal{M}_{n,n}(\mathbb{C})$, the $n \times n$-matrices with complex entries.

In principle the scalarization process works as before. The only difference is that the solutions to be constructed are not scalar anymore but matrix-valued. This is taken into account if we use a map σ taking its values in $\mathcal{M}_{n,n}(\mathbb{C})$ instead of a functional.

Our ansatz is analogous to the one for the scalar KP. Fix $a_1, \dots, a_n \in F'$ linearly independent and choose (i) U, $W \in \mathcal{S}_{a_1,\dots,a_n}(F)$, (ii) σ as the map from $\mathcal{S}_{a_1,\dots,a_n}(F)$ to $\mathcal{M}_{n,n}(\mathbb{C})$. Since σ is an algebra homomorphism by Proposition 3.4.4, multiplicativity of σ on $\mathcal{S}_{a_1,\dots,a_n}(F)$ is assured.

With this strategy we arrive at the following result.

Theorem 3.4.6. *Let F be a Banach space and $a_1, \dots, a_n \in F'$ constant, linearly independent functionals.*

If $V = V(x, y, t) \in \mathcal{S}_{a_1,\dots,a_n}(F)$ is a family of bounded operators such that $U = V_x$, $W = V_y$ solve the non-abelian KP equation (3.4), (3.5), then

$$\begin{aligned} u &= v_x, \\ w &= v_y, \end{aligned}$$

where the function $v = v(x, yt)$ is defined by

$$v(x, y, t) \quad = \quad \sigma\left(V(x, y, t) \right),$$

solve the matrix KP equation (3.35), (3.36).

Proof Observe U, $W \in \mathcal{S}_{a_1,\dots,a_n}(F)$, because the same holds for V. Therefore we can read the non-abelian KP (3.4), (3.5) as an operator equation in the Banach algebra $\mathcal{S}_{a_1,\dots,a_n}(F)$. Since σ is an algebra homomorphism by Proposition 3.4.4, the proof of Theorem 3.3.1 can be carried over literally. □

As in Section 3.3.1 we give a slightly different proof of Theorem 3.4.6. Its argument is quite instructive, since it in particular explains the concrete choice of the map σ used for the extended scalarization process.

Since $V = V(x, y, t) \in \mathcal{S}_{a_1,\dots,a_n}(F)$, there are vector-functions $c_i = c_i(x, y, t)$, $i = 1, \dots, n$, such that $V = \sum_{i=1}^n a_i \otimes c_i$. Hence,

$$U = \sum_{i=1}^n a_i \otimes d_i, \qquad W = \sum_{i=1}^n a_i \otimes f_i,$$

where $d_i = c_{i,x}$, $f_i = c_{i,y}$. Using the calculation rules for one-dimensional operators, see Section 2.1.1, the operator-valued KP equation reads

$$\begin{aligned} \sum_{i=1}^n a_i \otimes d_{i,t} + 3 \sum_{i=1}^n a_i \otimes \left(\sum_{k=1}^n \Big(\langle d_{i,x}, a_k \rangle d_k + \langle d_i, a_k \rangle d_{k,x} \Big) \right) + \sum_{i=1}^n a_i \otimes d_{i,xxx} \\ = -3\alpha^2 \sum_{i=1}^n a_i \otimes f_{i,y} - 3\alpha \sum_{i=1}^n a_i \otimes \left(\sum_{k=1}^n \Big(\langle f_i, a_k \rangle d_k - \langle d_i, a_k \rangle f_k \Big) \right), \end{aligned}$$

$$\sum_{i=1}^n a_i \otimes f_{i,x} = \sum_{i=1}^n a_i \otimes d_{i,y}.$$

There exist vectors $c_{j,0} \in E$ such that $\langle c_{j,0}, a_i\rangle = \delta_{ij}$, $i, j = 1, \ldots, n$. If we now apply the operator KP to $c_{j,0}$, $j = 1, \ldots, n$, we obtain the vector equations

$$\begin{aligned} d_{j,t} + 3\sum_{k=1}^n \Big(\langle d_{j,x}, a_k\rangle d_k + \langle d_j, a_k\rangle d_{k,x} \Big) + d_{j,xxx} \\ = -3\alpha^2 f_{j,y} - 3\alpha \sum_{k=1}^n \Big(\langle f_j, a_k\rangle d_k - \langle d_j, a_k\rangle f_k \Big), \\ f_{j,x} = d_{j,y}, \end{aligned}$$

$j = 1, \ldots, n$, and application of the functional a_i, $i = 1, \ldots, n$, to these vector equations yields the scalar system

$$\begin{aligned} \langle d_{j,t}, a_i\rangle + 3\sum_{k=1}^n \Big(\langle d_{j,x}, a_k\rangle \langle d_k, a_i\rangle + \langle d_j, a_k\rangle \langle d_{k,x}, a_i\rangle \Big) + \langle d_{j,xxx}, a_i\rangle \\ = -3\alpha^2 \langle f_{j,y}, a_i\rangle - 3\alpha \sum_{k=1}^n \Big(\langle f_j, a_k\rangle \langle d_k, a_i\rangle - \langle d_j, a_k\rangle \langle f_k, a_i\rangle \Big), \\ \langle f_{j,x}, a_i\rangle = \langle d_{j,y}, a_i\rangle, \end{aligned}$$

$i, j = 1 \ldots, n$.

Since the functional a_i is continuous, we can interchange the order of derivation with the application of a_i. As a consequence, we have shown the following system for the functions $u_{ij} = \langle d_j, a_i\rangle$, $w_{ij} = \langle f_j, a_i\rangle$, $i, j = 1, \ldots, n$,

$$\begin{aligned} (u_{ij})_t + 3\sum_{k=1}^n \Big((u_{kj})_x u_{ik} + u_{kj}(u_{ik})_x \Big) + (u_{ij})_{xxx} = \\ = -3\alpha^2 (w_{ij})_y - 3\alpha \sum_{k=1}^n \Big(w_{kj} u_{ik} - u_{kj} w_{ik} \Big), \\ (w_{ij})_x = (u_{ij})_y, \end{aligned}$$

which is the matrix KP equation in its component form. In other words, $u = (u_{ij})_{i,j=1}^n$, $w = (w_{ij})_{i,j=1}^n$ solve the matrix KP equation (3.35), (3.36).

Finally note that

$$\begin{aligned} u &= \Big(\langle d_j, a_i\rangle\Big)_{i,j=1}^n = \Big(\langle c_{j,x}, a_i\rangle\Big)_{i,j=1}^n = \Big((\langle c_j, a_i\rangle)_{i,j=1}^n\Big)_x = v_x, \\ w &= \Big(\langle f_j, a_i\rangle\Big)_{i,j=1}^n = (\langle c_{j,y}, a_i\rangle\Big)_{i,j=1}^n = \Big((\langle c_j, a_i\rangle)_{i,j=1}^n\Big)_y = v_y, \end{aligned}$$

for

$$v = \Big(\langle c_j, a_i\rangle\Big)_{i,j=1}^n = \Big(\mathrm{tr}(a_i \otimes c_j)\Big)_{i,j=1}^n = \sigma\Big(\sum_{i=1}^n a_i \otimes c_i\Big) = \sigma(V).$$

This completes the alternative proof of Theorem 3.4.6.

3.4.3 Resulting solution formulas

In the sequel we carry out the scalarization process for the concrete operator soliton in Theorem 3.2.1. Moreover, we derive an ameliorated solution formula which is formulated in terms of determinants.

As result a result of scalarization we obtain the following proposition.

Proposition 3.4.7. *Let E, F be Banach spaces and $A \in \mathcal{L}(E)$, $B \in \mathcal{L}(F)$, $D \in \mathcal{L}(E,F)$ arbitrary constant operators. Moreover, assume that the constant operator $C \in \mathcal{L}(F,E)$ satisfies the n-dimensionality condition*

$$AC + CB = \sum_{j=1}^{n} a_j \otimes c_j \tag{3.37}$$

for $a_1, \dots, a_n \in F'$ linearly independent and $c_1, \dots, c_n \in E$. Define the operator-functions

$$\begin{aligned} L(x,y,t) &= \widehat{L}(x,y,t)C, & \widehat{L}(x,y,t) &= \exp\Big(Ax + \frac{1}{\alpha}A^2 y - 4A^3 t \Big), \\ M(x,y,t) &= \widehat{M}(x,y,t)D, & \widehat{M}(x,y,t) &= \exp\Big(Bx - \frac{1}{\alpha}B^2 y - 4B^3 t \Big). \end{aligned}$$

Then, on $\Omega = \{(x,y,t) \mid (I+LM) \text{ is invertible}\}$, a solution of the matrix KP equation (3.35), (3.36) is given by

$$\begin{aligned} u &= v_x, \\ w &= v_y, \end{aligned}$$

where

$$v = 2\Big(\mathrm{tr}\Big(M\ (\ I + LM\ )^{-1}\ \widehat{L}(a_i \otimes c_j)\Big)\Big)_{i,j=1}^{n}. \tag{3.38}$$

Remark 3.4.8. *Again we stress that the operator D can be chosen without any restriction.*

Proof Obviously the operator functions $L = L(x,y,t) \in \mathcal{L}(F,E)$, $M = M(x,y,t) \in \mathcal{L}(E,F)$ solve the base equations in Theorem 3.2.1. Hence Theorem 3.2.1 provides us with the solution $U = V_x$, $W = V_y$ of the non-abelian KP equation (3.4), (3.5) on Ω, where

$$V = 2M(I+LM)^{-1}(AL + LB).$$

By (3.37),

$$\begin{aligned} V &= 2M(I+LM)^{-1}\widehat{L}(AC+CB) \\ &= 2M(I+LM)^{-1}\widehat{L}\Big(\sum_{j=1}^{n} a_j \otimes c_j\Big) \\ &= \sum_{j=1}^{n} a_j \otimes \big(\ 2M(I+LM)^{-1}\widehat{L}c_j\ \big). \end{aligned}$$

Thus $V(x,y,t) = \sum_{j=1}^{n} a_j \otimes \big(\ \widehat{c}_j(x,y,t)\ \big)$ with vector functions $\widehat{c}_j = 2M(I+LM)^{-1}\widehat{L}\, c_j \in F$. In particular, $V \in \mathcal{S}_{a_1,\dots,a_n}(F)$.

Now we are in the position to apply Theorem 3.4.6. This shows that on Ω a solution of the matrix KP equation (3.35), (3.36) is given by $u = v_x$, $w = v_y$ with

$$\begin{aligned} v &= \sigma\Big(\sum_{j=1}^{n} a_j \otimes \widehat{c}_j\Big) = \Big(\mathrm{tr}\Big(a_i \otimes \widehat{c}_j\Big)\Big)_{i,j=1}^{n} \\ &= \Big(\mathrm{tr}\Big(a_i \otimes \big(2M(I+LM)^{-1}\widehat{L}\, c_j\big)\Big)\Big)_{i,j=1}^{n} \\ &= 2\Big(\mathrm{tr}\Big(M(I+LM)^{-1}\widehat{L}(a_i \otimes c_j)\Big)\Big)_{i,j=1}^{n}. \end{aligned}$$

This is the desired solution formula. □

Again the next step is to gain a solution formula which avoids the calculation of the inverse operator $(I+LM)^{-1}$. The resulting formulas, which are stated in the next theorem, have a similar structure as those for the AKNS system in Theorem 2.3.7.

Theorem 3.4.9. *Let E, F be Banach spaces, and let $A \in \mathcal{L}(E)$, $B \in \mathcal{L}(F)$, $D \in \mathcal{L}(F,E)$ arbitrary constant operators.*

a) Assume that the n-dimensionality condition (3.37) is satisfied with $C \in \mathcal{A}(E,F)$ and linearly independent $a_1, \dots, a_n \in F'$, where $\mathcal{A}$ is an arbitrary quasi-Banach ideal admitting a continuous determinant δ. Define

$$\begin{aligned} L(x,y,t) &= \widehat{L}(x,y,t)\, C, \qquad \widehat{L}(x,y,t) = \exp(Ax + \frac{1}{\alpha}A^2 y - 4A^3 t), \\ M(x,y,t) &= \widehat{M}(x,y,t) D, \qquad \widehat{M}(x,y,t) = \exp(Bx - \frac{1}{\alpha}B^2 y - 4B^3 t). \end{aligned}$$

Then, on $\Omega = \{(x,y,t) \,|\, (I+LM) \text{ is invertible}\}$, a solution of the matrix KP (3.35), (3.36) is given by

$$\begin{aligned} u &= v_x, \\ w &= v_y, \end{aligned}$$

where

$$v = 2 \left(1 - \frac{\delta\big(I + LM - \widehat{L}(a_i \otimes c_j)M\big)}{\delta\big(I+LM\big)} \right)_{i,j=1}^{n}. \tag{3.39}$$

b) If $0 \notin \operatorname{spec}(A) + \operatorname{spec}(B)$, then a) holds with $C = \Phi_{A,B}^{-1}(\sum_{j=1}^{n} a_j \otimes c_j)$, linearly independent $a_1, \dots, a_n \in F'$, and any p-Banach operator ideal $\mathcal{A}$ admitting a continuous determinant δ.

Proof Recall that the solution u, w of the matrix KP in Proposition 3.4.7 are determined in terms of the matrix-valued function v given in (3.38). We show that v can be rewritten in the form (3.39).

To this end we denote by τ the trace on $\mathcal{A}$ which corresponds to the determinant δ according to the Trace-determinant theorem (see Proposition B.2.11). Since all traces coincide on the ideal $\mathcal{F}$ of finite-rank operators, we have $\tau_{\mathcal{F}} = \operatorname{tr}$ and hence

$$\begin{aligned} v &= 2 \left(\operatorname{tr}\Big(M(I+LM)^{-1}\widehat{L}(a_i \otimes c_j)\Big) \right)_{i,j=1}^{n} \\ &= 2 \left(\tau\Big(M(I+LM)^{-1}\widehat{L}(a_i \otimes c_j)\Big) \right)_{i,j=1}^{n} \\ &= 2 \left(1 - \delta\Big(I - M(I+LM)^{-1}\widehat{L}(a_i \otimes c_j)\Big) \right)_{i,j=1}^{n}, \end{aligned}$$

where the latter identity follows by the common relation $\delta(I+T) = 1 + \tau(T)$ for one-dimensional operators T. Using the refomulation

$$\begin{aligned} \delta\Big(I - M(I+LM)^{-1}\widehat{L}(a_i \otimes c_j)\Big) &= \delta\Big(I - (I+LM)^{-1}\widehat{L}(a_i \otimes c_j)M\Big) \\ &= \delta\Big(\ (I+LM)^{-1}\big(I + LM - \widehat{L}(a_i \otimes c_j)M\ \big)\Big) \\ &= \frac{\delta\big(I + LM - \widehat{L}(a_i \otimes c_j)M\big)}{\delta\big(I+LM\big)} \end{aligned}$$

part a) is shown. Part b) is an immediate consequence of Proposition 2.3.4. □

It is remarkable that for the (usual) trace of the solution of the matrix KP a comparable smooth formula holds as for the solutions of the scalar KP.

Proposition 3.4.10. *If the requirements of Theorem 3.4.9 a) are met, then for the (usual) trace of the solution given in (3.39), it holds*

$$\mathrm{tr}(v) \;=\; 2\,\frac{\partial}{\partial x}\log\delta(I+LM).$$

Proof We start from the representation of v given in (3.38),

$$v \;=\; 2\,\Big(\,\mathrm{tr}\Big(M(I+LM)^{-1}\widehat{L}(a_i\otimes c_j)\Big)\,\Big)_{i,j=1}^{n}.$$

For the trace of the solution v of the matrix KP, calculated in its usual way, applying successively the n-dimensionality condition (3.37), the trace property, and the base equations $L_x = AL$, $M_x = BM$, we obtain

$$\begin{aligned}
\mathrm{tr}(v) &= 2\sum_{j=1}^{n}\mathrm{tr}\Big(M(I+LM)^{-1}\widehat{L}(a_j\otimes c_j)\Big)\\
&= 2\,\mathrm{tr}\Big(M(I+LM)^{-1}\widehat{L}\sum_{j=1}^{n}a_j\otimes c_j\Big)\\
&= 2\,\mathrm{tr}\Big(M(I+LM)^{-1}\widehat{L}(AC+CB)\Big)\\
&= 2\,\mathrm{tr}\Big(M(I+LM)^{-1}(AL+LB)\Big)\\
&= 2\,\mathrm{tr}\Big((I+LM)^{-1}(AL+LB)M\Big)\\
&= 2\,\mathrm{tr}\Big((I+LM)^{-1}(LM)_x\Big).
\end{aligned}$$

Let us again denote by τ the trace on $\mathcal{A}$ corresponding to the determinant δ. Since on the finite-rank operators τ coincides with tr, we have

$$\begin{aligned}
\mathrm{tr}(v) &= 2\,\tau\Big((1+LM)^{-1}(LM)_x\Big)\\
&= 2\,\frac{\partial}{\partial x}\log\delta(1+LM),
\end{aligned}$$

the latter following from Proposition B.2.12 (note that LM belongs to $\mathcal{A}$ because C does). This is the assertion. □

3.5 Some remarks concerning Hirota's bilinear equation and Miura transformations

In the present section we prove that the solution formula in Theorem 3.3.4 even provides a solution of the bilinear KP equation. For convenience, we briefly recall Hirota's formalism for the KP. Then we state our result. The basic idea of the proof is to avoid a direct operator-valued treatment of Hirota's equation, but to reformulate it first in terms of the Miura transformation. Then we observe that the Miura transformation translates nicely to the operator level together with a good deal of its crucial properties. This will rather smoothly lead to the desired result.

In soliton theory, Hirota's bilinear method has proved to be a powerful technique for finding explicit solutions (see the classical references [46], [47] and also [68]). Starting from a soliton equation the main idea is, roughly speaking, to develop it into a bilinear equation, i.e., a differential equation in twice as many variables as before, and to look at the restriction to the diagonal. The latter can be investigated by perturbation methods.

For the KP equation, the bilinear form was derived in [87]. It reads

$$(D_1^4 - 4D_{1,3}^2 + 3D_2^2)(p \cdot p) = 0. \tag{3.40}$$

For a multiindex $\alpha = (\alpha_1, \ldots, \alpha_k)$, $\alpha_j \in \{1,2,3\}$, and two functions p, q we define the Hirota derivative

$$D_\alpha^k \, (\, p \cdot q \,) \, (x) = \left[\left(\frac{\partial}{\partial x_{\alpha_1}} - \frac{\partial}{\partial \xi_{\alpha_1}} \right) \cdots \left(\frac{\partial}{\partial x_{\alpha_k}} - \frac{\partial}{\partial \xi_{\alpha_k}} \right) p(x) \, q(\xi) \right] \Bigg|_{\, x = \xi}$$

for $x = (x_1, x_2, x_3)$, $\xi = (\xi_1, \xi_2, \xi_3)$. The right side is defined on the diagonal in $\mathbb{R}_x^3 \times \mathbb{R}_\xi^3$ which is in the obvious way parametrized by $x \in \mathbb{R}_x^3$. If $\alpha = (j, \ldots, j)$, we simply write D_j^k.

Symmetry implies many useful identities for the Hirota derivatives (see [48]). For example one readily sees that $D_j^k(p \cdot p)$ vanishes identically if k is odd, whereas $D_j^k(p \cdot p)$ is in general nontrivial for even k.

The dependent variable transformation $u = \partial_{x_1}^2 \log p$ leads to the KP equation in a slightly different form as the one used up to now, namely

$$u_{x_3} - \frac{1}{4} u_{x_1 x_1 x_1} - 3 u u_{x_1} = \frac{3}{4} \int_{-\infty}^{x_1} u_{x_2 x_2} d\xi_1, \tag{3.41}$$

where appropriate boundary conditions have to be imposed for $x \to -\infty$.

By the coordinate transformation $x = x_1$, $y = \alpha x_2$, and $t = -x_3/4$, rescaling $v = 2u$ and differentiating once, we arrive at

$$\big(v_t + v_{xxx} + 6 v v_x \big)_x = -3\alpha^2 v_{yy},$$

the scalar version of (3.13), which obviously provides solutions of (3.1), (3.2). In the sequel we focus on (3.41), which is the common form with respect to the bilinear formalism.

Proposition 3.5.1. *Let E, F be Banach spaces and $A \in \mathcal{L}(E)$, $B \in \mathcal{L}(F)$ arbitrary constant operators.*

If $L = L(x_1, x_2, x_3) \in \mathcal{A}(F, E)$, $M = M(x_1, x_2, x_3) \in \mathcal{A}(E, F)$, for $\mathcal{A}$ a quasi-Banach operator ideal admitting a continuous determinant δ, are operator-valued C^4-smooth functions, satisfy the base equations

$$\begin{array}{lll} L_{x_1} = AL, & L_{x_2} = A^2 L, & L_{x_3} = A^3 L, \\ M_{x_1} = BM, & M_{x_2} = -B^2 M, & M_{x_3} = B^3 M, \end{array}$$

and the condition

$$(AL + LB)P = (AL + LB) \tag{3.42}$$

holds for a constant, one-dimensional projector $P \in \mathcal{L}(F)$, then a solution of the bilinear form (3.40) of the KP equation is given by

$$p = \delta(I + LM).$$

Before we give the proof, we show that the bilinear KP equation (3.40) is intimately linked with Miura type transformations. Writing out (3.40) explicitly, we obtain becomes

$$(p_{x_1x_1x_1x_1} p - 4 p_{x_1x_1x_1} p_{x_1} + 3 p_{x_1x_1}^2) + 3(p_{x_2x_2} p - p_{x_2}^2) - 4(p_{x_1x_3} p - p_{x_1} p_{x_3}) = 0,$$

or, at points where $p(x_1,x_2,x_3) \neq 0$,

$$\Big(\frac{p_{x_1x_1x_1x_1}}{p} - 4\frac{p_{x_1x_3}}{p} + 3\frac{p_{x_2x_2}}{p}\Big) + 3\Big(\Big(\frac{p_{x_1x_1}}{p}\Big)^2 - \Big(\frac{p_{x_2}}{p}\Big)^2\Big) - 4\Big(\frac{p_{x_1x_1x_1}}{p} - \frac{p_{x_3}}{p}\Big)\frac{p_{x_1}}{p} = 0. \tag{3.43}$$

Define $w_j = p_{x_j}/p$. There is a systematic way to express (3.43) in terms of the w_j's and Miura-type transformations of the w_j's. Namely, if we denote by M_j the Miura-type transformation given by

$$M_j(f) = f_{x_j} + w_j f,$$

then, taking logarithmic derivatives, it is easy to verify the following properties:

$$\begin{aligned} \frac{p_{x_jx_i}}{p} &= M_j(w_i) = M_i(w_j), \qquad i,j = 1,2,3, \\ \frac{\partial_{x_1}^{n+1} p}{p} &= M_1^n(w_1), \qquad n \in \mathbb{N}_0. \end{aligned}$$

Using these properties, (3.43) becomes

$$\Big(M_1^3(w_1) - 4M_3(w_1) + 3M_2(w_2)\Big) + 3\Big(\big(M_1(w_1)\big)^2 - (w_2)^2\Big) - 4\Big(M_1^2(w_1) - w_3\Big)w_1 = 0. \tag{3.44}$$

We want to point out that (3.44) can also be viewed as a reformulation of (3.43) in terms of the w_j's and their derivatives. Indeed, evaluating the Miura-type transforms in (3.44) we end up with

$$w_{1,x_1x_1x_1} + 6w_{1,x_1}^2 - 4w_{3,x_1} + 3w_{2,x_2} = 0. \tag{3.45}$$

In the sequel we extend this reasoning to the operator-level. To this end, we first define operator-valued analogues of the Miura-type transformations, discuss their properties, and solve the operator pendant (3.46) of (3.44).

Proposition 3.5.2. *Let E, F be Banach spaces, and $A \in \mathcal{L}(E)$, $B \in \mathcal{L}(F)$ arbitrary constant operators. Assume that $L = L(x_1,x_2,x_3) \in \mathcal{L}(F,E)$, $M = M(x_1,x_2,x_3) \in \mathcal{L}(E,F)$, are operator-valued functions satisfying the base equations*

$$\begin{aligned} L_{x_1} &= AL, & L_{x_2} &= A^2L, & L_{x_3} &= A^3L, \\ M_{x_1} &= BM, & M_{x_2} &= -B^2M, & M_{x_3} &= B^3M, \end{aligned}$$

and assume that $(I+LM)$ is always invertible.

For $n \in \mathbb{N}$, define the operator-valued functions

$$Z_n = (I+ML)^{-1}(MA^nL + (-B)^n),$$

and, for $j = 1,2,3$, the operator-valued Miura-type transform M_j as the map from $\mathcal{L}(F)$ to $\mathcal{L}(F)$ given by

$$M_j(W) = W_{x_j} + Z_jW \qquad \text{for} \qquad W \in \mathcal{L}(F).$$

Then,

$$M_j(Z_n) = Z_{n+j}$$

for all $j = 1,2,3$ and all $n \in \mathbb{N}$.

Proof In the notation of Lemma 3.2.4 we have $Z_n = -\widehat{W}_n$. Thus, by the definition of the Miura-type transformation, a successive application of the lemma yields

$$M_j(Z_n) = Z_{n,x_j} + Z_j Z_n = -\widehat{W}_{n,x_j} + \widehat{W}_j \widehat{W}_n = W_j V_n + \widehat{W}_j \widehat{W}_n = -\widehat{W}_{n+j} = Z_{n+j},$$

what had to be shown. □

Corollary 3.5.3. *Let the requirements of Proposition 3.5.2 be met. Then,*

$$Z_{i,x_j} - Z_{j,x_i} = [Z_i, Z_j]$$

for $i, j = 1, 2, 3$.

Proof By the definition of the Miura-type transformation, we have

$$Z_{i,x_j} - Z_{j,x_i} = \big(\, M_j(Z_i) - Z_j Z_i \,\big) - \big(\, M_i(Z_j) - Z_i Z_j \,\big) = [Z_i, Z_j],$$

the latter by Lemma 3.5.2. □

Next we solve the operator-valued pendant of (3.44).

Proposition 3.5.4. *Let the requirements of Proposition 3.5.2 be met. Then the operator-functions* Z_1, Z_2, Z_3 *solve the operator equation*

$$\Big(M_1^3(Z_1) - 4M_3(Z_1) + 3M_2(Z_2)\Big) + 3\Big(\big(M_1(Z_1)\big)^2 - (Z_2)^2\Big) - 4\Big(M_1^2(Z_1) - Z_3\Big)Z_1 = 0. \tag{3.46}$$

Proof Successive application of Lemma 3.5.2 to each of the terms in the large brackets shows that all respective terms vanish. The assertion follows. □

Evaluating the Miura-type transformations in (3.46) yields an operator-valued pendant of (3.45).

Lemma 3.5.5. *The operator equation (3.46) in Proposition 3.5.4 is equivalent to*

$$Z_{1,x_1x_1x_1} + 6(Z_{1,x_1})^2 - 4Z_{1,x_3} + 3Z_{2,x_2} = -3[Z_1, Z_{1,x_2}]. \tag{3.47}$$

In the proof of Theorem 3.2.1 we have shown that the operator function $V = -\widehat{W}_1 + B$ solves (3.25), where $-\widehat{W}_1 = Z_1$ in the notation of Lemma 3.5.5. Thus the above lemma states that even an integrated version of (3.25) holds on the operator-level.

Proof Recall $M_2(Z_2) - (Z_2)^2 = Z_{2,x_2}$, $M_3(Z_1) - Z_3 Z_1 = Z_{1,x_3}$, by the definition of the respective Miura-type transformations. Moreover, we observe for iterated applications of the Miura-type transformation M_1 to Z_1,

$$\begin{aligned}
M_1(Z_1) &= Z_{1,x_1} + Z_1^2, \\
M_1^2(Z_1) &= Z_{1,x_1x_1} + Z_{1,x_1}Z_1 + 2Z_1Z_{1,x_1} + Z_1^3, \\
M_1^3(Z_1) &= Z_{1,x_1x_1x_1}Z_1 + Z_{1,x_1x_1} + 31Z_{1,x_1x_1} + 3Z_{1,x_1}^2 \\
&\qquad + Z_{1,x_1}Z_1^2 + 2Z_1Z_{1,x_1}Z_1 + 3Z_1^2 Z_{1,x_1} + Z_1^4.
\end{aligned}$$

By a straightforward calculation, these identities yield

$$Z_{1,x_1x_1x_1} + 6(Z_{1,x_1})^2 - 4Z_{1,x_3} + 3Z_{2,x_2} = -3[Z_1, Z_{1,x_1x_1} + 2Z_1Z_{1,x_1}].$$

Next, expressing reversely derivatives by Miura-type transformations and applying Lemma 3.5.2 again, we get

$$\begin{aligned} Z_{1,x_1x_1} + 2Z_1Z_{1,x_1} &= M_1^2(Z_1) - Z_{1,x_1}Z_1 - Z_1^3 \;=\; M_1^2(Z_1) - M_1(Z_1)Z_1 \\ &= Z_3 - Z_2Z_1 \;=\; M_2(Z_1) - Z_2Z_1 \;=\; Z_{1,x_2}, \end{aligned}$$

which shows the desired reformulation. □

In the next lemma we establish the link from solutions of operator equation (3.47) to (scalar) solutions of the bilinear form (3.40).

Lemma 3.5.6. *Let the assumptions of Proposition 3.5.1 be met and denote by τ the trace which corresponds to the determinant δ on $\mathcal{A}$. Then, for $j = 1, 2, 3$,*

$$\frac{\partial}{\partial x_j} \log p = \tau\Big(Z_j - (-B)^j\Big).$$

Proof Using successively Proposition B.2.12, the base equations in Proposition 3.5.1, the trace property, and Lemma 3.2.3, we get

$$\begin{aligned} \frac{\partial}{\partial x_j} p &= \frac{\big(\ \delta(I+LM)\ \big)_{x_j}}{\delta(I+LM)} \\ &= \tau\Big((I+LM)^{-1}(LM)_{x_j}\Big) \\ &= \tau\Big((I+LM)^{-1}\ \big(\ (A^jL - L(-B)^j)M\ \big)\Big) \\ &= \tau\Big(M(I+LM)^{-1}\ \big(\ A^jL - L(-B)^j\ \big)\Big) \\ &= \tau\Big((I+ML)^{-1}M\ \big(\ A^jL - L(-B)^j\ \big)\Big) \\ &= \tau\Big((I+ML)^{-1}\ \big(\ MA^jL - (1+ML)(-B)^j + (-B)^j\ \big)\Big) \\ &= \tau\big(Z_j - (-B)^j\big), \end{aligned}$$

which is the assertion. □

Now we are prepared for the proof of Proposition 3.5.1.

Proof (of Proposition 3.5.1) We have to show that $p = \delta(I + LM)$ is a solution of the bilinear KP equation (3.40). To this end set

$$w_j = \frac{\partial}{\partial x_j} \log p.$$

By Lemma 3.5.6, $w_j = \tau(W_j)$, where $W_j = Z_j - (-B)^j$. Moreover, since the trace τ is continuous, we have

$$w_{j,x_i} \;=\; \tau\big(W_j\big)_{x_i} = \tau(W_{j,x_i}) = \tau(Z_{j,x_i}),$$

and corresponding identities hold for higher order derivatives of w_j. Next we note that, because

$$W_1 = (I+ML)^{-1}M\ \big(\ AL + LB\ \big),$$

(3.42) guarantees $W_1P = W_1$. Therefore $Z_{1,x_1}P = W_{1,x_1}P = W_{1,x_1} = Z_{1,x_1}$ as well, and, as a consequence,

$$\begin{aligned} \tau\ \big(\ (Z_{1,x_1})^2\ \big) &= \tau\ \big(\ (Z_{1,x_1}P)^2\ \big) = \tau\ \big(\ (Z_{1,x_1}P^2)^2\ \big) \\ &= \tau\ \big(\ Z_{1,x_1}P^2Z_{1,x_1}P^2\ \big) = \tau\ \big(\ (PZ_{1,x_1}P)(PZ_{1,x_1}P)\ \big) \end{aligned}$$

$$= \tau\big(PZ_{1,x_1}P\big)\,\tau\big(PZ_{1,x_1}P\big)$$
$$= \Big(\tau\big(Z_{1,x_1}P^2\big)\Big)^2 = \Big(\tau\big(Z_{1,x_1}P\big)\Big)^2 = \Big(\tau\big(Z_{1,x_1}\big)\Big)^2.$$

This means that the trace τ is multiplicative on Z_{1,x_1}. Furthermore, we obviously have $\tau([Z_1, Z_{1,x_2}]) = 0$.

With these preparations, we now proceed as follows. By Lemma 3.5.5, Z_1 solves (3.47). Now we apply the trace τ and use that it is linear, continuous, and multiplicative on Z_{1,x_1} to observe that w_j satisfies (3.45).

It remains to check that the line of argument which led from (3.43) to (3.45) can be reversed with $w_j = p_{x_j}/p$. This shows that p solves the bilinear KP equation (3.40) on $\Omega = \{(x_1, x_2, x_3) | p = 0\}$. Since p is real-analytic, the theorem follows $\square$

3.6 Line-solitons and Miles structures

In this section we discuss a first application of our solution formula. Taking the operator-valued parameters A, B simply as diagonal matrices of the same size, and $D = I$, we obtain the N-line-solitons of the KP-II equation. We discuss regularity, include some special features often neglected in the literature, and provide computer graphics for illustration. Furthermore, we explain how Miles structures fit into the picture.

We start with the proof that, for finite-dimensional parameters $A \in \mathcal{M}_{n,n}(\mathbb{C})$ and $B \in \mathcal{M}_{m,m}(\mathbb{C})$, the essential data of the generated solution is encoded in the Jordan canonical form of A, B.

Lemma 3.6.1. *Let $A \in \mathcal{M}_{n,n}(\mathbb{C})$, $B \in \mathcal{M}_{m,m}(\mathbb{C})$, and let U, V be matrices transforming A, B into Jordan canonical form J_A, J_B, respectively, i.e., $A = U^{-1}J_AU$, $B = V^{-1}J_BV$.*

Then the solution in Theorem 3.3.4 b) is not altered if we replace simultaneously A, B by J_A, J_B, the matrix D by VDU^{-1}, and the vectors a, c by $(V^{-1})'a$, Uc.

Proof Recall that we are in the finite-dimensional setting ($E = \mathbb{C}^n$, $F = \mathbb{C}^m$). It is sufficient to verify that the asserted replacements do not change the determinant $p = \det(I + LM)$ in Theorem 3.3.4.

From the very definition of the exponential function, it is easy to check that, for the functions $\widehat{L}$, $\widehat{M}$ as defined in Theorem 3.3.4, it holds

$$\widehat{L} = U^{-1}J_{\widehat{L}}U \qquad \text{for } J_{\widehat{L}} = \exp\big(J_A\,x + \frac{1}{\alpha}J_A^2\,y - 4J_A^3\,t\big),$$
$$\widehat{M} = V^{-1}J_{\widehat{M}}V \qquad \text{for } J_{\widehat{M}} = \exp\big(J_B\,x - \frac{1}{\alpha}J_A^2\,y - 4J_A^3\,t\big).$$

Abbreviating $C = \Phi_{A,B}^{-1}(a\otimes c)$, from $a\otimes c = AC + CB = U^{-1}\big(J_A(UCV^{-1}) + (UCV^{-1})J_B\big)V$ we next observe

$$UCV^{-1} = \Phi_{J_A,J_B}^{-1}\ \big(\ U(a\otimes c)V^{-1}\ \big) = \Phi_{J_A,J_B}^{-1}\ \big(\ ((V^{-1})'a)\otimes(Uc)\ \big)\ . \tag{3.48}$$

As a consequence,

$$p = \det\Big(I + \widehat{L}C\,\widehat{M}D\Big) = \det\Big(I + (U^{-1}J_{\widehat{L}}U)C(V^{-1}J_{\widehat{M}}V)D\Big)$$
$$= \det\Big(I + J_{\widehat{L}}(UCV^{-1})J_{\widehat{M}}(VDU^{-1})\Big)$$

which, by (3.48), shows the assertion. $\square$

Next we derive the formal N-soliton solution for both the KP (i.e., both the KP-I and KP-II) equation.

Proposition 3.6.2. *For complex numbers p_j, q_j, $j = 1, \ldots, N$, satisfying $p_i + q_j \neq 0\ \forall i, j$, a solution of the KP equation (3.1), (3.2) on $\Omega = \{(x, y, t) \mid p(x, y, t) \neq 0\}$ is given by*

$$\begin{aligned} u &= v_x, \\ w &= v_y, \end{aligned}$$

with $v = 2\dfrac{\partial}{\partial x} \log p$, where

$$p(x, y, t) = 1 + \sum_{n=1}^{N} \sum_{i_1 < \ldots < i_n} \prod_{j=1}^{n} f_{i_j}(x, y, t) \prod_{\substack{j,j'=1 \\ j<j'}}^{n} \frac{(p_{i_j} - p_{i_{j'}})(q_{i_j} - q_{i_{j'}})}{(p_{i_j} + q_{i_{j'}})(q_{i_j} + p_{i_{j'}})} \tag{3.49}$$

and $f_j(x, y, t) = \exp\Big((p_j + q_j)x + \frac{1}{\alpha}(p_j^2 - q_j^2)y - 4(p_j^3 + q_j^3)t + \varphi_j\Big)$ with arbitrary (complex) constants φ_j.

Proof Consider the diagonal matrices A, $B \in \mathcal{M}_{N,N}(\mathbb{C})$ given by

$$A = \begin{pmatrix} p_1 & & 0 \\ & \ddots & \\ 0 & & p_N \end{pmatrix}, \qquad B = \begin{pmatrix} q_1 & & 0 \\ & \ddots & \\ 0 & & q_N \end{pmatrix},$$

where p_j, $q_j \in \mathbb{C}$ with $p_i + q_j \neq 0$ for all $i, j = 1, \ldots, N$.

Our aim is to apply Theorem 3.3.4 b). To this end, we need the following preparations:

(i) Given a, $c \in \mathbb{C}^N$, a solution of the coupling condition $AC + CB = a \otimes c$ is given by the matrix

$$C = \Big(\frac{a_j c_i}{p_i + q_j}\Big)_{i,j=1}^{N}.$$

(ii) The matrix exponential function

$$\widehat{L}(x, y, t) = \exp\big(Ax + \frac{1}{\alpha}A^2 y - 4A^3 t\big)$$

is again a diagonal matrix with the entries $\ell_j = \ell_j(x, y, t) = \exp\big(p_j x + \frac{1}{\alpha}p_j^2 y - 4p_j^3 t\big)$ on its diagonal, and an analogous statement holds for

$$\widehat{M}(x, y, t) = \exp\big(Bx - \frac{1}{\alpha}B^2 y - 4B^3 t\big).$$

The coupling condition in (i) is easily verified by

$$\begin{aligned} &< (AC + BC)e_j, e_i > = < ACe_j, e_i > + < CBe_j, e_i > \\ &\quad = \ < Ce_j, A'e_i > + < CBe_j, e_i > = < Ce_j, p_i e_i > + < Cq_j e_j, e_i > \\ &\quad = \ (p_i + q_j) < Ce_j, e_i > = \ a_j c_i \\ &\quad = \ < (a \otimes c)e_j, e_i > . \end{aligned}$$

As for (ii) we observe that for any diagonal matrix $T = \operatorname{diag}\{t_j \mid j = 1, \ldots, N\}$ the n-th power can be taken entry by entry, i.e., $T^n = \operatorname{diag}\{t_j^n \mid j = 1, \ldots, N\}$. Thus (ii) follows directly from the definition of exponential function as a power series.

Now we are in position to apply Theorem 3.3.4 b). With $D = I$, we thus obtain a solution of the KP equation (3.1), (3.2) by

$$u = v_x, \qquad w = v_y$$

with $v = 2\dfrac{\partial}{\partial x} \log p$, where $p = \det\big(I + \widehat{L}C\widehat{M}\big)$.

Using the well known expansion rule for determinants (confer Lemma 4.1.6), we can evaluate p explicitly. Namely, by (i), (ii), we find

$$\begin{aligned} p &= \det\left(\delta_{ij} + \frac{a_j c_i}{p_i + q_j}\ell_i m_j\right)_{i,j=1}^{N} \\ &= 1 + \sum_{n=1}^{N} \sum_{i_1 < \ldots < i_n} \det\left(\frac{a_{i_{j'}} c_{i_j}}{p_{i_j} + q_{i_{j'}}}\ell_{i_j} m_{i_{j'}}\right)_{j,j'=1}^{n} \\ &= 1 + \sum_{n=1}^{N} \sum_{i_1 < \ldots < i_n} \prod_{j=1}^{n} a_{i_j} c_{i_j} \cdot m_{i_j}\ell_{i_j} \cdot \det\left(\frac{1}{p_{i_j} + q_{i_{j'}}}\right)_{j,j'=1}^{n} \\ &= 1 + \sum_{n=1}^{N} \sum_{i_1 < \ldots < i_n} \prod_{j=1}^{n} \frac{a_{i_j} c_{i_j}}{p_{i_j} + q_{i_j}} m_{i_j}\ell_{i_j} \prod_{\substack{j,j'=1 \\ j<j'}}^{n} \frac{(p_{i_j} - p_{i_{j'}})(q_{i_j} - q_{i_{j'}})}{(p_{i_j} + q_{i_{j'}})(q_{i_j} + p_{i_{j'}})}, \end{aligned}$$

where the last identity goes back to a remark of Cauchy (see [20], p. 151-159, [78], VII, §1, Nr. 3 and Lemma 6.1.2).

To complete the proof, it remains to rewrite the expressions $\dfrac{a_j c_j}{p_j + q_j}\ell_j m_j = f_j$. □

For convenience we show that the solution derived in Proposition 3.6.2 coincides with the representation of the N-solitons as given by [87]. Since there probably are some misprints in this version, we here refer to the representation in [47]. It reads

$$u = \partial_x^2 \log f_N,$$

$$f_N = \sum_{\mu=0,1} \exp\Big(\sum_{i=1}^{N} \mu_i\eta_i + \sum_{\substack{i,j=1 \\ i<j}}^{N} \mu_i\mu_j A_{ij}\Big) \quad \text{the sum runs over all } \mu = (\mu_1, \ldots, \mu_N) \text{ with } \mu_i \in \{0,1\},$$

$$\eta_i = \kappa_i x + \lambda_i y - \omega_i t + \eta_i^{(0)},$$

$$\exp(A_{ij}) = -\frac{(\kappa_i - \kappa_j)^4 + 3\alpha^2(\lambda_i - \lambda_j)^2 - (\kappa_i - \kappa_j)(\omega_i - \omega_j)}{(\kappa_i + \kappa_j)^4 + 3\alpha^2(\lambda_i + \lambda_j)^2 - (\kappa_i + \kappa_j)(\omega_i + \omega_j)},$$

and the dispersion relation $\kappa_i^4 + 3\alpha^2\lambda_i^2 - \kappa_i\omega_i = 0$ holds.

To see that this equals our formula, take $\kappa_i = p_i + q_i$ and $\lambda_i = (p_i^2 - q_i^2)/\alpha$. Then the dispersion relation shows $\omega_i = 4(p_i^3 + q_i^3)$, hence $m_i\ell_i = \exp(\eta_i)$. After some calculations,

$$\exp(A_{ij}) = \frac{(\kappa_i - \kappa_j)^2 - \alpha^2(\lambda_i/\kappa_i - \lambda_j/\kappa_j)^2}{(\kappa_i + \kappa_j)^2 - \alpha^2(\lambda_i/\kappa_i - \lambda_j/\kappa_j)^2} = \frac{(p_i - p_j)(q_i - q_j)}{(p_i + q_j)(p_j + q_i)}.$$

Thus Proposition 3.6.2 indeed describes the N-soliton solution for the KP equation.

We now come to line-solitons which are regular, real-valued solutions of the KP-II equation (i.e., in the case $\alpha = 1$).

Proposition 3.6.3. *Let p_j, $q_j \in \mathbb{R}$ for $j = 1, \ldots, N$, and assume that for all $i \neq j$ one of the following conditions is satisfied:*
(i) $-q_i < -q_j < p_j < p_i$,
(ii) $-q_i < p_i < -q_j < p_j$.
Then a real-valued solution of the KP-II equation (3.1), (3.2) with $\alpha = 1$ is given by

$$\begin{aligned} u &= v_x, \\ w &= v_y, \end{aligned}$$

with $v = 2\dfrac{\partial}{\partial x}\log p$, *where*

$$p(x,y,t) = 1 + \sum_{n=1}^{N} \sum_{i_1<\ldots<i_n} \prod_{j=1}^{n} f_{i_j}(x,y,t) \prod_{\substack{j,j'=1\\ j<j'}}^{n} \frac{(p_{i_j}-p_{i_{j'}})(q_{i_j}-q_{i_{j'}})}{(p_{i_j}+q_{i_{j'}})(q_{i_j}+p_{i_{j'}})} \tag{3.50}$$

and $f_j(x,y,t) = \exp\Big((p_j+q_j)x + (p_j^2-q_j^2)y - 4(p_j^3+q_j^3)t + \varphi_j\Big)$ *with* $\varphi_j \in \mathbb{R}$ *arbitrarily.*

Moreover, this solution is regular on the whole of $\mathbb{R}^3$.

Proof The solution property is an immediate consequence of Proposition 3.6.2, and reality of the solution is obvious. Thus it remains to show regularity. But since either of the conditions (i), (ii) implies that the phase-shift term

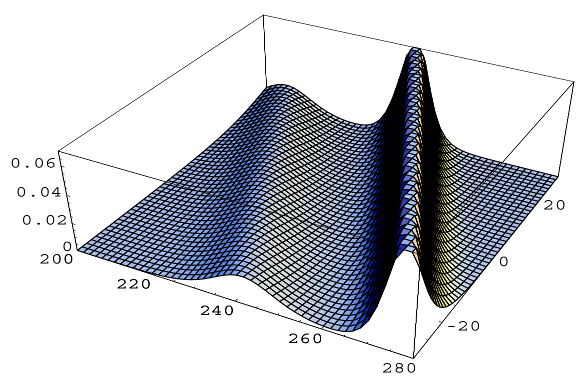
$$\frac{(p_i-p_j)(q_i-q_j)}{(p_i+q_j)(q_i+p_j)}$$

is positive, we have $p(x,y,t) > 1$ for all (x,y,t). This shows the assertion. $\square$

To illustrate our result, we turn to $u = u(x,y,t)$, which are the N-line-solitons in the closer sense. Recall that a single line-soliton corresponding to the parameters p_j, q_j is characterized by its angle to the x-axis and its shape, which are encoded in $(p_j - q_j)$, $(p_j + q_j)$, respectively (confer (3.3) and the subsequent discussion).

Now there are two particularly interesting cases.

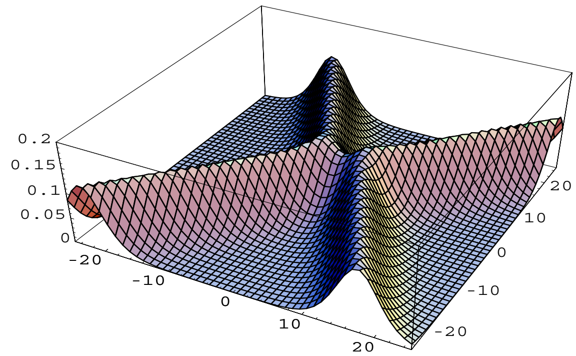
Existence of line-solitons with the same angle to the x-axis Using the condition (i) in Proposition 3.6.3 one easily arranges parameters $-q_i < -q_j < p_j < p_i$ such that $p_i - q_i = p_j - q_j$. In the N-line-soliton this yields two solitons of different shape but with the same angle to the x-axis.



The 2-line soliton for $p_1 = 0.5$, $q_1 = -0.1$, $p_2 = 0.4$, $q_2 = -0.2$, plotted at the time $t = 200$.

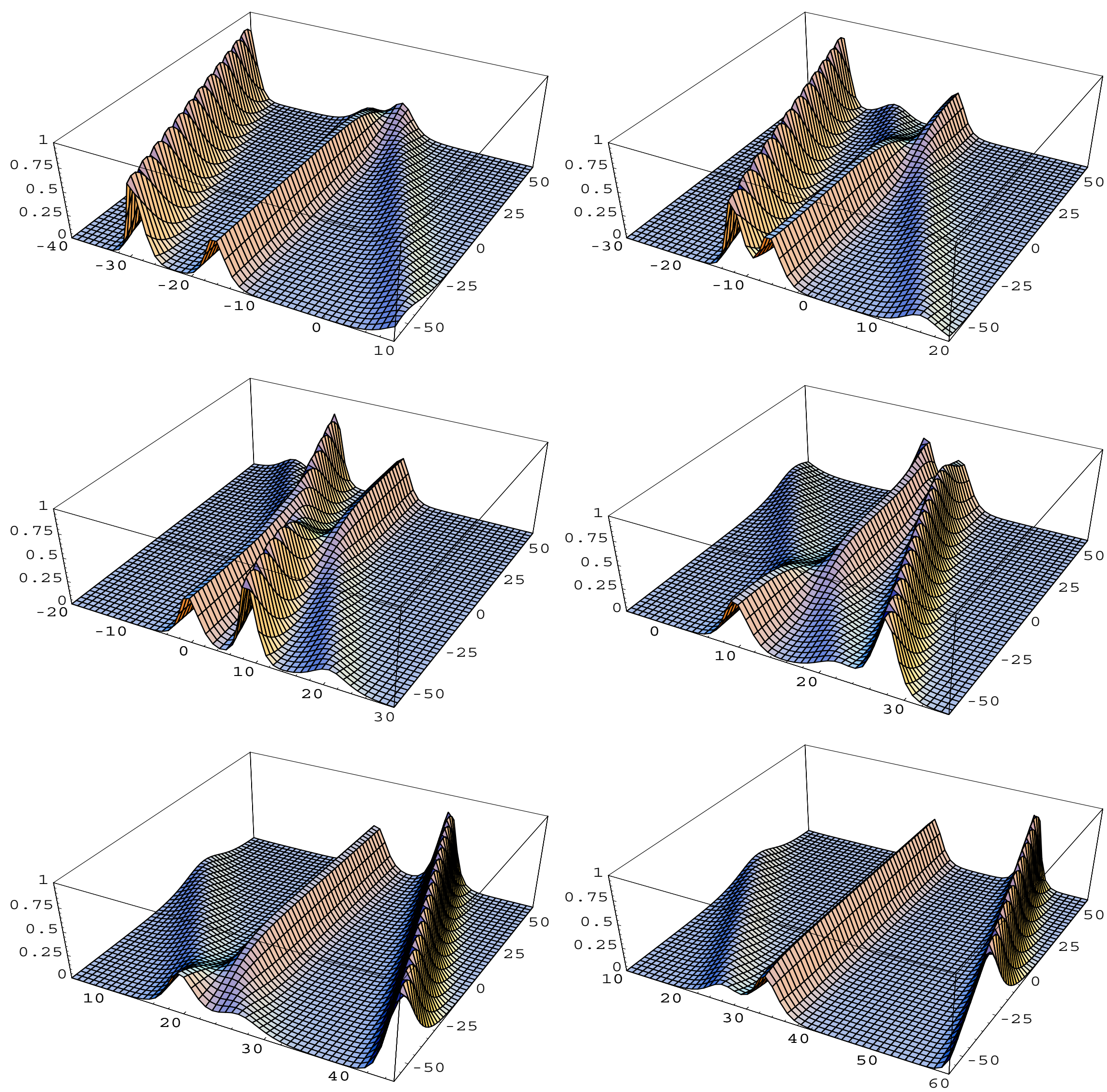
Existence of line-solitons with identical shape In contrast, the condition (ii) in Proposition 3.6.3 allows to construct even line-solitons with identical shape. One has only to arrange the parameters $-q_i < p_i < -q_j < p_j$ such that $p_i + q_i = p_j + q_j$.

Note that solutions of this particular type cannot be achieved by condition (i), which is the standard regularity condition in the literature (see [37], [81], [55]).



The 2-line soliton for $p_1 = 0.5$, $q_1 = 0$, $p_2 = -0.2$, $q_2 = 0.7$, plotted at the time $t = 0$.

Snapshots of the three-soliton during interaction



Successively the solution is plotted for $t = -20$,-10, 0, 10, 20, and 30. The parameters of the three participating line-solitons are $p_1 = 0.7$, $q_1 = 0.6$, $p_2 = 0.5$, $q_2 = 0.5$, $p_3 = 0.4$, $q_3 = 0.2$.

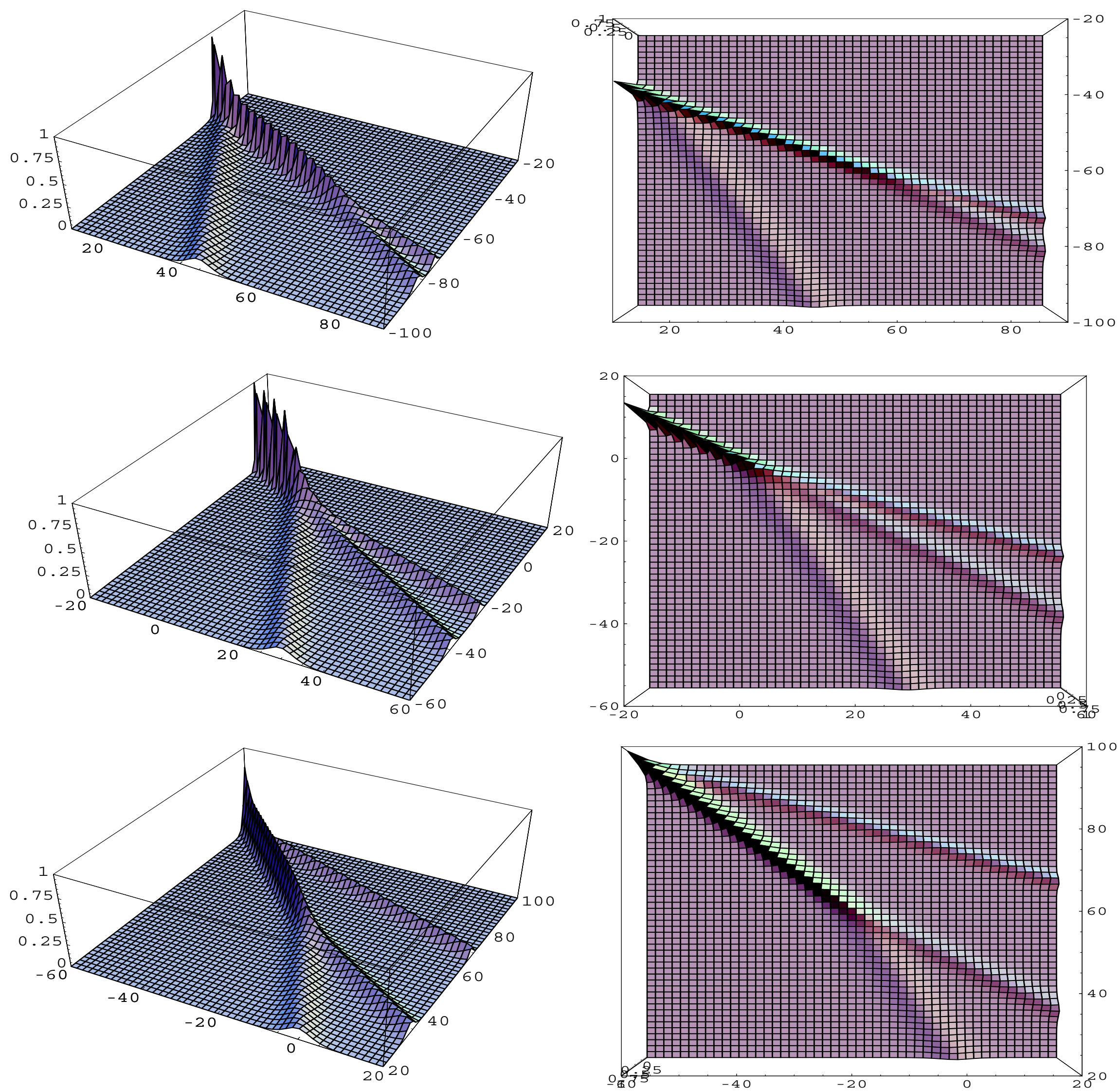
Resonance phenomena Finally, we will briefly discuss resonance phenomena. These arise in the same way as the line-solitons in Proposition 3.6.3, but without the condition that the parameters p_j, q_j are pairwise different. For example, if we set

$$A = pI, \qquad B = \begin{pmatrix} q_1 & & 0 \\ & \ddots & \\ 0 & & q_N \end{pmatrix},$$

with $-q_N < \ldots < -q_1 < p$, the resulting solution is a typical Miles structure. It reads precisely as in Proposition 3.6.3, but with

$$p(x, y, t) = 1 + \sum_{j=1}^{N} f_j(x, y, t)$$

instead of (3.50). In particular it is real-valued and regular. By a more detailed investigation it can be verified that it looks like a tree with a trunk and N branches. The trunk corresponds to the soliton with parameters p, q_N, the first branch to p, q_1, and the other $N-1$ branches to $-q_j$, q_{j+1} for $j = 1, \ldots, N-1$.

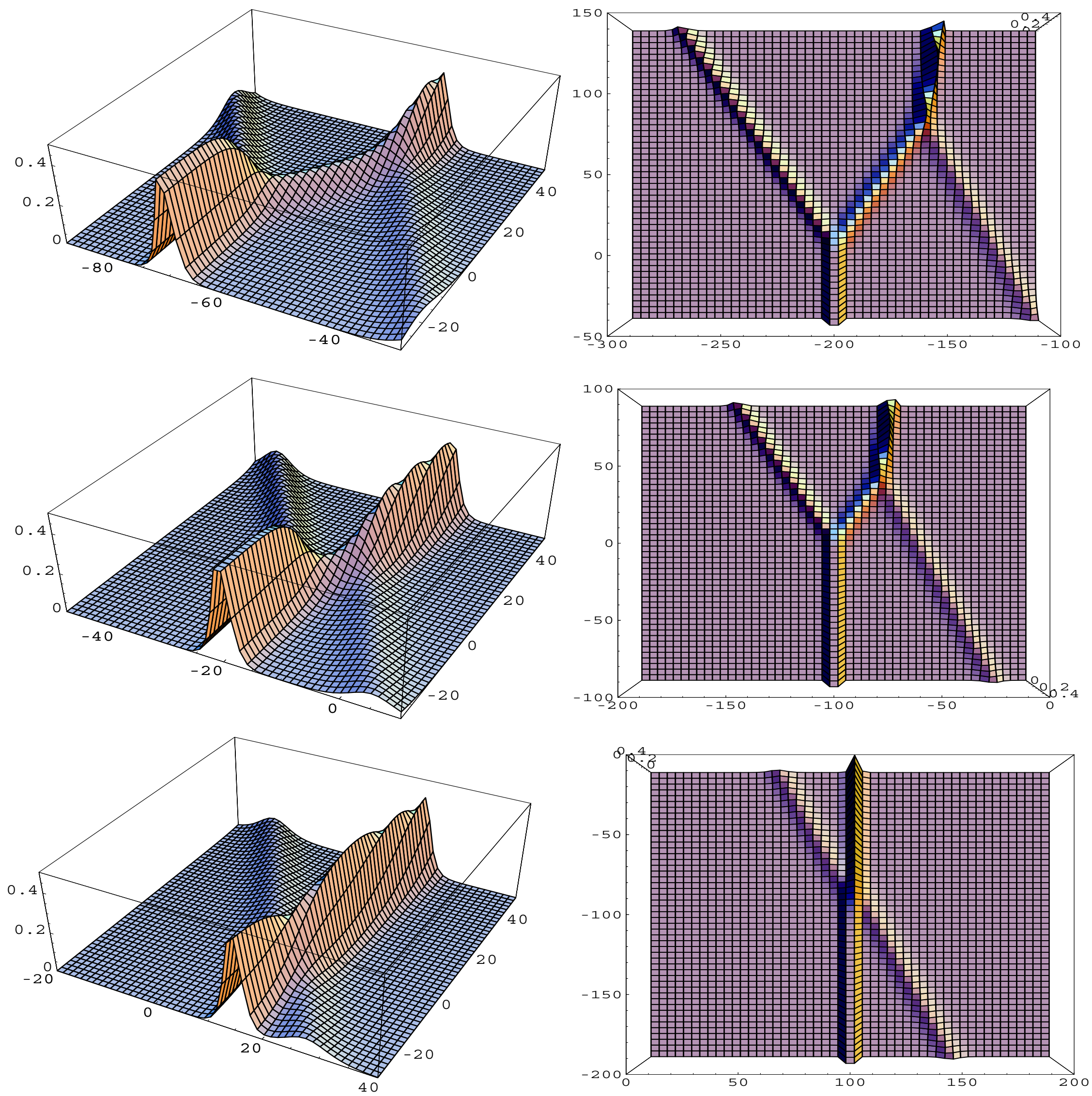


Miles structure ($p_1 = p_2 = p_3 = 0$, $q_1 = 0.5$, $q_2 = 1$, $q_3 = 1.5$)
plotted successively at the times $t = -5, 0, 10$

The plots above show such a Miles structure. Beside the usual pictures, we here found it instructive to plot also corresponding pictures from above. Note also that we have plotted $u(x, -y, t)$ to make the diagrams clearer.

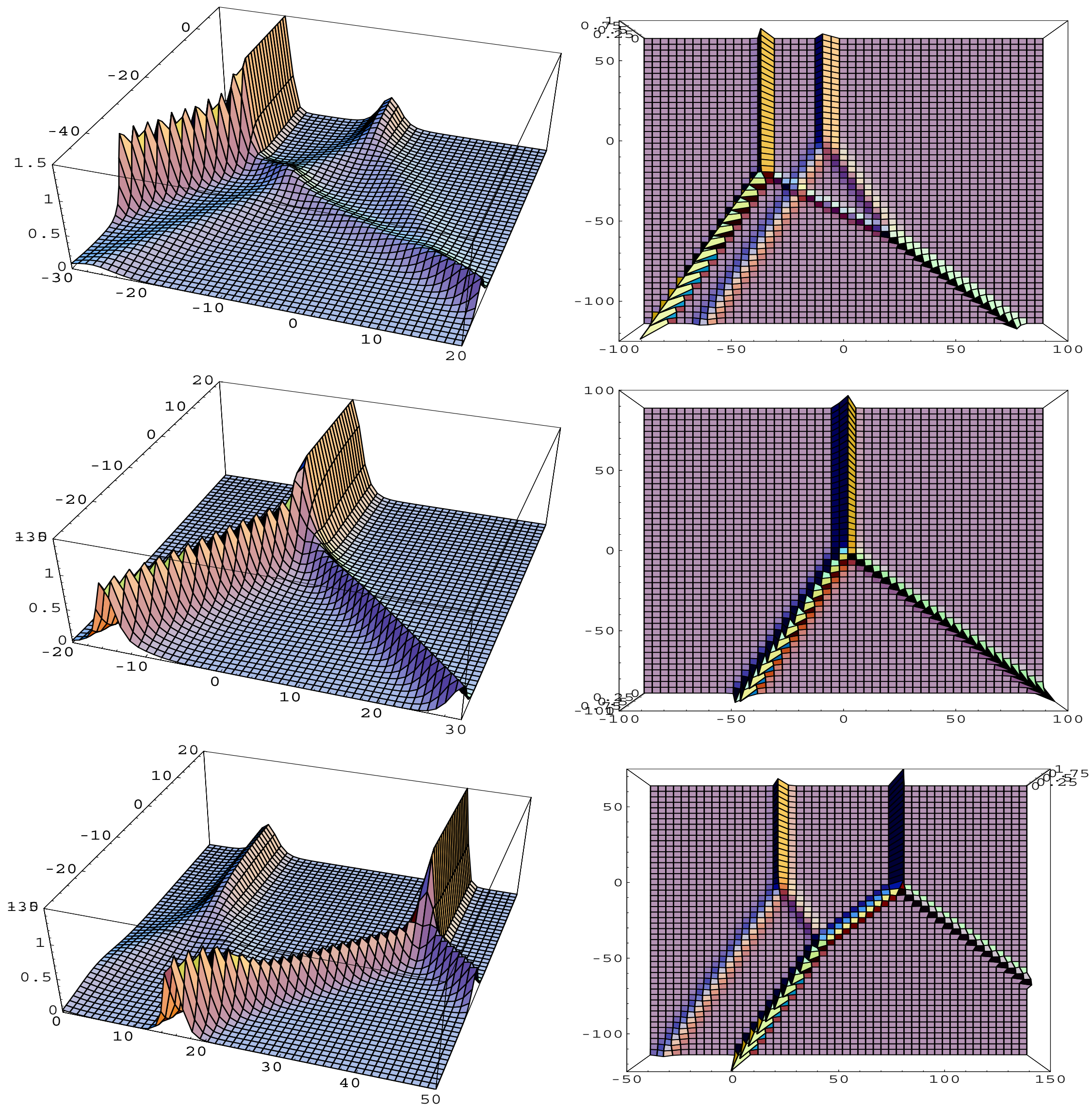
Some other impressing examples of what can happen in degenerated cases which can be similarly obtained as the one above are gathered below. Among the examples where the colliding solitons change their shape in the interaction, and a structure developping an inner closed cell.

Example 1



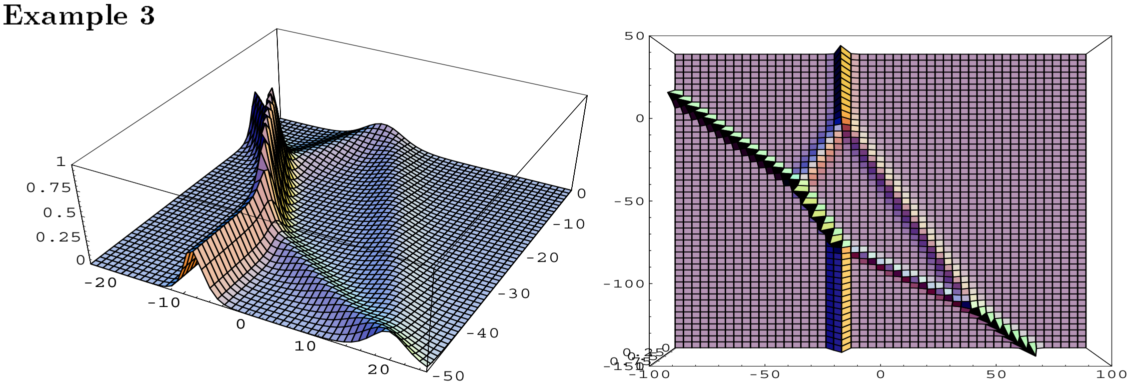
The above plots show solitons which change their shape during interaction. The parameters are $p_1 = 0.5$, $q_1 = 0.5$, $p_2 = 0.5$, $q_2 = 0$, $p_3 = 0.4$, $q_3 = 0$. The solution is plotted successively at the times $t = -70, -20, 15$ and $t = -200, -100, 100$, respectively for the left and right column.

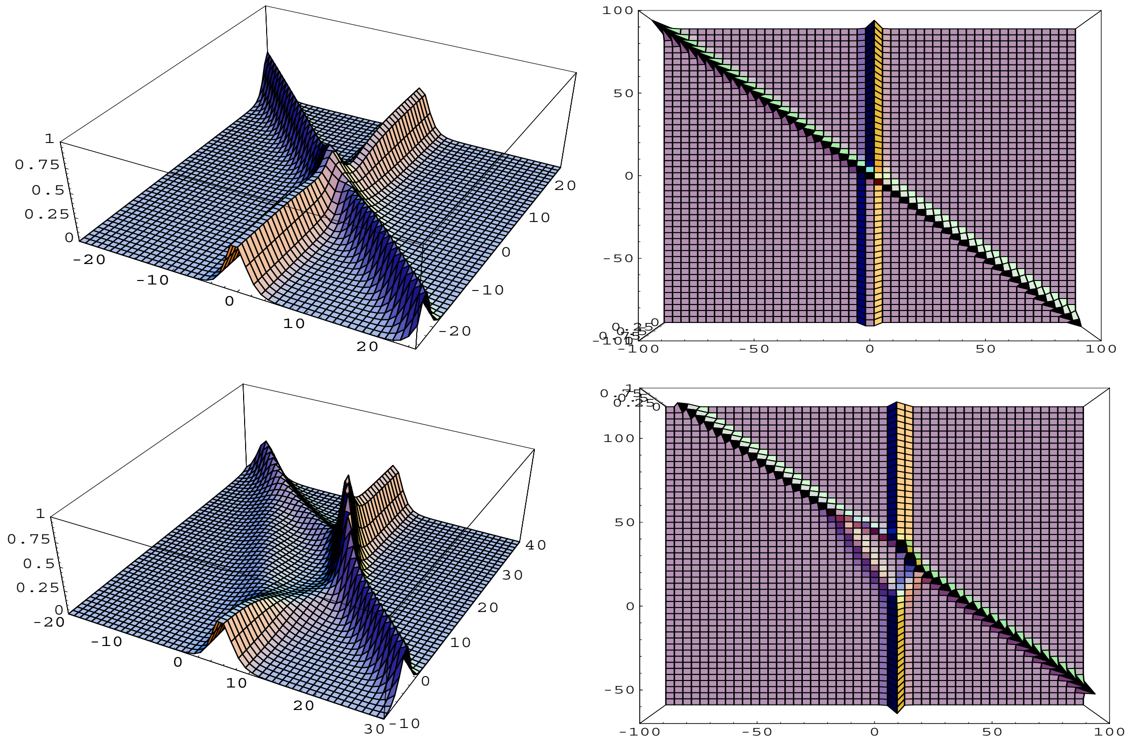
Example 2



Above the parameters are $p_1 = 0$, $q_1 = 0.5$, $p_2 = 0.5$, $q_2 = 0.5$, $p_3 = 0.5$, $q_3 = 1$, $p_4 = 1$, $q_4 = 1$, and the solution is plotted successively at the times $t = -5, 0, 10$ and $t = -10, 0, 20$, respectively for the left and right column.

Example 3





In these final plots, an inner cell develops. The parameters are $p_1 = 0$, $q_1 = 0.5$, $p_2 = 0.5$, $q_2 = 0.5$, $p_3 = 0.5$, $q_3 = 1$, $p_4 = 0$, $q_4 = 1$. Note that the solution $u(x, -y, t)$ is plotted to make the pictures clearer. The successive times are $t = -8, 0, 6$ and $t = -20, 0, 10$, respectively for the left and right column.

Chapter 4

Soliton-like solutions for the AKNS system and properties of its reductions

The ultimate purpose of Chapters 4, 5, and 6 is to achieve a reasonably complete asymptotic understanding of the solutions one can obtain by inserting finite matrices into the formulas of Chapter 2. In the present chapter we present the necessary solution formulas and deduce structural properties of more general character. Furthermore, we will focus on soliton-like solutions, for which some of the results are more explicit.

The first section is concerned with the general AKNS system. Here we have independent generating matrices $A \in \mathcal{M}_{n,n}(\mathbb{C})$, $B \in \mathcal{M}_{m,m}(\mathbb{C})$, which may even be of different dimension. Then we observe that our constructions behave naturally with respect to similarity between matrices, which yields reduction to matrices in Jordan canonical form. This simple fact indicates that one may hope for a relation between the dynamics of the solutions and the algebraic properties of the matrices. For the generic case that A and B are diagonalizable, we are able to evaluate the solution formula in closed form (including explicit evaluation of all appearing determinants). To the best of our knowledge, for the general AKNS system this is done for the first time. For the $\mathbb{C}$-reduced AKNS system, corresponding formulas are given in [56], [63] in terms of Wronskian determinants, but full evaluation of the determinants is only achieved in the $\mathbb{R}$-reduced case.

We call these solutions soliton-like because of their formal analogy with solitons. As already mentioned in [3], these solutions may present instantaneous singularities. This is no surprise and illustrates the known fact that the unrestricted AKNS system does not yield soliton equations in general. In the framework of [3], [5], this is reflected by the circumstance that the inverse scattering method applies only formally. In particular, the Gelfand-Levitan-Marchenko equation need not be solvable.

But we do obtain honest soliton equations, if we restrict to convenient reductions of the AKNS system. Following [3], we consider successively the $\mathbb{C}$-reduced and $\mathbb{R}$-reduced AKNS system. In our approach the two cases are formally independent. But it is helpful to keep in mind that the $\mathbb{R}$-reduced system becomes a further reduction of the $\mathbb{C}$-reduced one under the physically plausible assumption that its solutions be real. As a matter of fact, the most prominent equations in the AKNS system (the Nonlinear Schrödinger, the sine-Gordon, and the modified Korteweg-de Vries equation) are already contained in these two reductions.

After reduction we are left with a single generating matrix in the solution formula. This allows us to introduce negatons as the solutions corresponding to matrices where each eigenvalue appears in exactly one Jordan block (otherwise there would be cancellation phenomena). Historically, negatons (and the corresponding concept of positions) were systematically derived via Wronskian determinants by Matveev et al. ([11], [57], [58], [59], [60], [95], see also [71], [79], [96], [98], [99]). Our motivation to consider negatons was mainly to prove conjectures of Matveev about the asymptotic behaviour of negatons in the general case. This will be done in Chapters 5 and 6.

In the remainder of Chapter 4 we clarify structural questions, mainly regularity and, for the $\mathbb{R}$-reduction, also reality. To prepare the transition to a finer analysis, we study in some detail the case of solitons. Already in this simplified setting, the explicit formulas we obtain for the $\mathbb{C}$-reduction are new. Then we explain how asymptotic analysis gives a precise meaning to what is usually called the particle behaviour of solitons. A further point of independent interest is that we include breathers into the description for the $\mathbb{R}$-reduction. These are the easiest case of formations of solitons (see [69]), i.e., solitons moving with equal velocity, which cannot be separated in asymptotic terms.

Finally, we emphasize that our formulas are quite different from the expressions which are derived by Wronskian techniques. In Appendix A we indicate how to translate formulas from [62], [63], [96] into our formalism.

4.1 Explicit formulas for soliton-like solutions of the AKNS system

In this section we explain what can be said for the AKNS system as a whole. We derive a basic solution formula and establish reduction to Jordan form. Finally we give an expression in closed form of soliton-like solutions, the class which specialises to solitons under the reductions to be discussed later.

4.1.1 Transition to Jordan data

For convenience we briefly recall some standard terminology. Let $\mathbb{C}^n$ be the Hilbert space of all complex n-tuples $\xi = (\xi_i)_{i=1}^n$, equipped with its usual inner product $(\xi, \eta) = \sum_{j=1}^n \xi_j \overline{\eta}_j$ for $\xi,\ \eta \in \mathbb{C}^n$. Its standard basis is denoted by $\{e_j | j = 1, \ldots, n\}$, where e_j is the vector with 1 in the j-th entry and zero elsewhere.

By the Riesz lemma, any functional $f_a \in (\mathbb{C}^n)'$ can be identified with a vector $a = (a_j)_{j=1}^n \in \mathbb{C}^n$ by the assignment $< \xi, f_a > = \sum_{j=1}^n \xi_j a_j = (\xi, \overline{a})$ for $\xi = (\xi_j)_{j=1}^n \in \mathbb{C}^n$. We do not distinguish between f_a and a in notation.

Consider now the one-dimensional operator $a \otimes c$ for $a,\ c \in \mathbb{C}^n$. From its definition,

$$(a \otimes c)(\xi) =< \xi, a > c =\Big(\sum_{j=1}^n \xi_j a_j c_i \Big)_{i=1}^n,$$

it is clear that it corresponds to the matrix $\big(a_j c_i \big)_{ij=1}^n$.

Let us now reformulate the solution formula of Theorem 2.4.4 b) for matrices.

Theorem 4.1.1. *Let $A \in \mathcal{M}_{n,n}(\mathbb{C})$ and $B \in \mathcal{M}_{m,m}(\mathbb{C})$ be matrices satisfying* $\mathrm{spec}(A) \cup \mathrm{spec}(B) \subseteq \{z \in \mathbb{C} |\, \mathrm{Re}(z) > 0\}$ *and* $\mathrm{spec}(A) \cup \mathrm{spec}(-B) \subseteq \{z \in \mathbb{C} |\, f_0(z)$ *finite* $\}$. *Let $0 \neq a$, $c \in \mathbb{C}^n$, $0 \neq b$, $d \in \mathbb{C}^m$ be arbitrary. We define the operator-functions*

$$\widehat{L}(x,t) = \exp\big(Ax + f_0(A)t\big), \qquad \widehat{M}(x,t) = \exp\big(Bx - f_0(-B)t\big),$$

and abbreviate

$$L = \widehat{L}\Phi_{A,B}^{-1}(b \otimes c), \qquad M = \widehat{M}\Phi_{B,A}^{-1}(a \otimes d),$$
$$L_0 = \widehat{L}(a \otimes c), \qquad M_0 = \widehat{M}(b \otimes d).$$

Then a solution of the AKNS system (1.1) is given by the pair

$$q = 1 - P/p, \qquad r = 1 - \widehat{P}/p,$$

where

$$P = \det \begin{pmatrix} I & L \\ M & I - M_0 \end{pmatrix}, \quad \widehat{P} = \det \begin{pmatrix} I - L_0 & L \\ M & I \end{pmatrix}, \quad p = \det \begin{pmatrix} I & L \\ M & I \end{pmatrix}.$$

This solution is defined on any strip $\mathbb{R} \times (t_1, t_2)$ on which p does not vanish.

Proof Note that $\mathrm{spec}(A) \cup \mathrm{spec}(B) \subseteq \{z \in \mathbb{C} \mid \mathrm{Re}(z) > 0\}$ implies $0 \notin \mathrm{spec}(A) + \mathrm{spec}(B)$, and hence the existence of $\Phi^{-1}_{A,B}$, $\Phi^{-1}_{B,A}$. Thus the only thing that requires a little justification is that $\exp(Ax)$, $\exp(Bx)$ behave sufficiently well for $x \to -\infty$.

To this end let U be a matrix transforming A into Jordan form J, say $A = U^{-1}JU$, where $J = \mathrm{diag}\{J_j \mid j = 1, \dots, N\}$ with Jordan block J_j of dimensions n_j corresponding to eigenvalues α_j. Then $\exp(Ax) = U^{-1}\exp(Jx)U = U^{-1}\mathrm{diag}\{\exp(J_j) \mid j = 1, \dots, N\}U$ and

$$\exp(J_j x) = \begin{pmatrix} \gamma_j^{(0)} & & \gamma_j^{(n_j-1)} \\ & \ddots & \\ & & \gamma_j^{(0)} \end{pmatrix} \exp(\alpha_j x),$$

with constants $\gamma_j^{(\mu)} = \alpha_j^{\mu}/\mu!$ for $\mu = 0, \dots, n_j - 1$. From this it is clear the $\exp(Ax)$ behaves sufficiently well for $x \to -\infty$. It remains to apply Theorem 2.4.4 b). $\square$

The following lemma shows that all relevant properties of the generating matrices A, B are encoded in their Jordan canonical form.

Lemma 4.1.2. *Let $A \in \mathcal{M}_{n,n}(\mathbb{C})$, $B \in \mathcal{M}_{m,m}(\mathbb{C})$ be as in Theorem 4.1.1, and let U, V be matrices transforming A, B into Jordan form J_A, J_B, respectively, namely $A = U^{-1}J_AU$, $B = V^{-1}J_BV$.*

Then the solution in Theorem 4.1.1 is not altered if we replace simultaneously A, B by J_A, J_B and the vectors a, b, c, d by $(U^{-1})'a$, $(V^{-1})'b$, Uc, Vd.

Proof Let us start with the first member $q = 1 - P/p$ of the solution in Theorem 4.1.1 and show that the asserted replacements do not change the determinants P, p.

From the very definition of the exponential function, it is easy to check

$$\begin{aligned} \widehat{L} &= U^{-1}J_{\widehat{L}}U \qquad \text{for } J_{\widehat{L}} := \exp\big(J_Ax + f_0(J_A)t\big), \\ \widehat{M} &= V^{-1}J_{\widehat{M}}V \qquad \text{for } J_{\widehat{M}} := \exp\big(J_Bx - f_0(-J_B)t\big). \end{aligned}$$

Using the abbreviation $C := \Phi^{-1}_{A,B}(b \otimes c)$, $D := \Phi^{-1}_{B,A}(a \otimes d)$, we next observe

$$b \otimes c = AC + CB = U^{-1}\ \big(\ J_A(UCV^{-1}) + (UCV^{-1})J_B\ \big)\ V.$$

It immediately follows

$$UCV^{-1} = \Phi^{-1}_{J_A,J_B}\ \big(\ U(b \otimes c)V^{-1}\ \big) = \Phi^{-1}_{J_A,J_B}\ \big(\ ((V^{-1})'b) \otimes (Uc)\ \big),$$

and, for the same reason, $VDU^{-1} = \Phi^{-1}_{J_B,J_A}\ \big(\ ((U^{-1})'a) \otimes (Vd)\ \big)$.

Therefore,

$$\begin{aligned} P &= \det\left(I + \begin{pmatrix} U^{-1} & 0 \\ 0 & V^{-1} \end{pmatrix} \begin{pmatrix} 0 & J_{\widehat{L}}\, UCV^{-1} \\ J_{\widehat{M}}\, VDU^{-1} & -J_{\widehat{M}}\, V(b \otimes d)V^{-1} \end{pmatrix} \begin{pmatrix} U & 0 \\ 0 & V \end{pmatrix} \right) \\ &= \det\left(I + \begin{pmatrix} 0 & J_{\widehat{L}}(UCV^{-1}) \\ J_{\widehat{M}}(VDU^{-1}) & -J_{\widehat{M}}\ \big(\ ((V^{-1})'b) \otimes (Vd)\ \big) \end{pmatrix} \right), \end{aligned}$$

which means that we can replace A, B by their Jordan forms if we use the vectors $(U^{-1})'a$, $(V^{-1})'b$, Uc, Vd instead of a, b, c, d. A similar calculation applies to p.

Finally, an analogous argument shows that $\widetilde{P}$ transforms properly. □

For later use, we also state the following representation of $\Phi_{A,B}^{-1}(a \otimes c)$.

Proposition 4.1.3. *Let $A \in \mathcal{M}_{n,n}(\mathbb{C})$, $B \in \mathcal{M}_{m,m}(\mathbb{C})$ be in Jordan canonical form: $A = \operatorname{diag}\{A_i \mid i = 1, \ldots, N\}$, $B = \operatorname{diag}\{B_j \mid j = 1, \ldots, M\}$ with Jordan blocks A_i, B_j of dimensions n_i, m_j corresponding to eigenvalues α_i, β_j, respectively. Assume $\alpha_i + \beta_j \neq 0$ for all $i = 1, \ldots, N$, $j = 1, \ldots, M$.*

Then

$$\Phi_{A,B}^{-1}(a \otimes c) = \big(\, \Gamma_l(c_i)\, T_{ij}\, \Gamma_r(a_j)\, \big)_{\substack{i=1,\ldots,N \\ j=1,\ldots,M}}$$

with the upper left and right band matrices $\Gamma_l(c_i)$, $i = 1, \ldots, N$ and $\Gamma_r(a_j)$, $j = 1, \ldots, M$, given by

$$\Gamma_l(c_i) = \begin{pmatrix} c_i^{(1)} & & c_i^{(n_i)} \\ & \cdot^{\cdot^{\cdot}} & \\ c_i^{(n_i)} & & 0 \end{pmatrix}, \qquad \Gamma_r(a_j) = \begin{pmatrix} a_j^{(1)} & & a_j^{(m_j)} \\ & \ddots & \\ 0 & & a_j^{(1)} \end{pmatrix},$$

where the vectors $a \in \mathbb{C}^m$, $c \in \mathbb{C}^n$ are decomposed according to the relevant Jordan decomposition, namely

$$\begin{aligned} c &= (c_1, \ldots, c_N) \quad \textit{with } c_i = (c_i^{(1)}, \ldots, c_i^{(n_i)}), \\ a &= (a_1, \ldots, a_M) \quad \textit{with } a_j = (a_j^{(1)}, \ldots, a_j^{(m_j)}), \end{aligned}$$

and

$$T_{ij} = \left((-1)^{\mu+\nu} \binom{\mu+\nu-2}{\mu-1} \left(\frac{1}{\alpha_i + \beta_j} \right)^{\mu+\nu-1} \right)_{\substack{\nu=1,\ldots,n_i \\ \mu=1,\ldots,m_j}}.$$

Proof Let us first consider the case that A, B are Jordan blocks of dimensions n, m with eigenvalues α, β, respectively. We claim $A'T + TB = e_m^{(1)} \otimes e_n^{(1)}$, where $e_k^{(1)}$ is the first standard basis vector in $\mathbb{C}^k$ and

$$T = (t_{\nu\mu})_{\substack{\nu=1,\ldots,n \\ \mu=1,\ldots,m}} \qquad \text{with } t_{\nu\mu} = (-1)^{\mu+\nu} \binom{\mu+\nu-2}{\mu-1} \left(\frac{1}{\alpha+\beta} \right)^{\mu+\nu-1}.$$

Since

$$(A'T)_{\nu\mu} = \begin{cases} \alpha t_{1\mu}, & \nu = 1, \\ \alpha t_{\nu\mu} + t_{(\nu-1)\mu}, & \nu > 1, \end{cases} \qquad (TB)_{\nu\mu} = \begin{cases} \beta t_{\nu 1}, & \mu = 1, \\ \beta t_{\nu\mu} + t_{\nu(\mu-1)}, & \mu > 1, \end{cases}$$

we check immediately $(A'T+TB)_{11} = 1$, $(A'T+TB)_{1\mu} = (A'T+TB)_{\nu 1} = 0$ for $\mu, \nu > 1$. Finally, for $\mu, \nu > 1$,

$$\begin{aligned} (A'T+TB)_{\nu\mu} &= (\alpha+\beta)t_{\nu\mu} + t_{(\nu-1)\mu} + t_{\nu(\mu-1)} \\ &= (-1)^{\mu+\nu} \left[\binom{\mu+\nu-2}{\mu-1} - \binom{\mu+\nu-3}{\mu-1} - \binom{\mu+\nu-3}{\mu-2} \right] \left(\frac{1}{\alpha+\beta} \right)^{\mu+\nu-2} \\ &= 0. \end{aligned}$$

This proves the claim.

Observe $[\Gamma_r(a), B] = 0$ and $A\Gamma_l(c) = \Gamma_l(c)A'$. Thus,

$$\begin{aligned} A\Big(\Gamma_l(c)T\Gamma_r(a)\Big) + \Big(\Gamma_l(c)T\Gamma_r(a)\Big)B \;&=\; \Gamma_l(c)\Big(A'\,T + TB\Big)\Gamma_r(a) \\ =\;\; \Gamma_l(c)\Big(e_m^{(1)} \otimes e_n^{(1)}\Big)\Gamma_r(a) \;&=\; \Big(\Gamma_r(a)'e_m^{(1)}\Big) \otimes \Big(\Gamma_l(c)e_n^{(1)}\Big) \;=\; a \otimes c. \end{aligned}$$

Consequently, $\Gamma_l(c)T\Gamma_r(a) = \Phi_{A,B}^{-1}(a \otimes c)$. This completes the proof for single Jordan blocks A, B of possibly different dimensions.

In the general case the assertion follows from

$$\Phi_{A,B}^{-1}(a \otimes c) = \Big(\Phi_{A_i,B_j}^{-1}(a_j \otimes c_i)\Big)_{\substack{i=1,\dots,N\\ j=1,\dots,M}}.$$

□

4.1.2 Soliton-like solutions for the AKNS system

As already mentioned in the introduction, the unreduced AKNS system seems to be too general to hope for a significant solution theory. Nevertheless it is remarkable that the formula in Theorem 4.1.1 leads to an explicit description of a solution class which lifts the N-solitons to the general level.

Theorem 4.1.4. *For complex numbers α_i, $i = 1, \dots n$, and β_j, $j = 1, \dots m$, such that $\mathrm{Re}(\alpha_i) > 0$, $\mathrm{Re}(\beta_j) > 0$ and $f_0(\alpha_i)$, $f_0(\beta_j)$ is finite for all i, j, a solution of the AKNS system (1.1) is given by*

$$q \;=\; \frac{1}{p}\Big[\sum_{j=1}^{m} g_j + \sum_{\kappa=1}^{\min(m-1,n)} (-1)^\kappa \sum_{\substack{i_1,\dots,i_\kappa=1\\ i_1<\dots<i_\kappa}}^{n} \; \sum_{\substack{j_1,\dots,j_{\kappa+1}=1\\ j_1<\dots<j_{\kappa+1}}}^{m} \widetilde{p}\begin{pmatrix} i_1, \dots, i_\kappa \\ j_1, \dots, j_{\kappa+1}\end{pmatrix}\Big] \tag{4.1}$$

$$r \;=\; \frac{1}{p}\Big[\sum_{i=1}^{n} f_i + \sum_{\kappa=1}^{\min(m,n-1)} (-1)^\kappa \sum_{\substack{i_1,\dots,i_{\kappa+1}=1\\ i_1<\dots<i_{\kappa+1}}}^{n} \; \sum_{\substack{j_1,\dots,j_\kappa=1\\ j_1<\dots<j_\kappa}}^{m} \widetilde{p}\begin{pmatrix} i_1, \dots, i_{\kappa+1} \\ j_1, \dots, j_\kappa\end{pmatrix}\Big] \tag{4.2}$$

where $p = 1 + \displaystyle\sum_{\kappa=1}^{\min(m,n)} (-1)^\kappa \sum_{\substack{i_1,\dots,i_\kappa=1\\ i_1<\dots<i_\kappa}}^{n} \; \sum_{\substack{j_1,\dots,j_\kappa=1\\ j_1<\dots<j_\kappa}}^{m} \widetilde{p}\begin{pmatrix} i_1, \dots, i_\kappa \\ j_1, \dots, j_\kappa\end{pmatrix}$ *and where we use the abbreviation*

$$\widetilde{p}\begin{pmatrix} i_1, \dots, i_\kappa \\ j_1, \dots, j_\lambda\end{pmatrix} = \prod_{\mu=1}^{\kappa} f_{i_\mu} \prod_{\nu=1}^{\lambda} g_{j_\nu} \prod_{\substack{\mu,\nu=1\\ \mu<\nu}}^{\kappa} (\alpha_{i_\mu} - \alpha_{i_\nu})^2 \prod_{\substack{\mu,\nu=1\\ \mu<\nu}}^{\lambda} (\beta_{j_\mu} - \beta_{j_\nu})^2 \Big/ \prod_{\mu=1}^{\kappa}\prod_{\nu=1}^{\lambda} (\alpha_{i_\mu} + \beta_{j_\nu})^2.$$

with $f_i(x,t) = f_i^{(0)} \exp\big(\, \alpha_i x + f_0(\alpha_i)t \,\big)$, $g_j(x,t) = g_j^{(0)} \exp\big(\, \beta_j x - f_0(-\beta_j)t \,\big)$, and arbitrary constants $f_i^{(0)}$, $g_j^{(0)} \in \mathbb{C}\backslash\{0\}$, $i = 1, \dots, n$, $j = 1, \dots, m$.

More precisely, q, r is a solution on all strips $\mathbb{R} \times (t_1, t_2)$ on which p does not vanish.

In particular, for $n = m = 1$, we recover (1.3), (1.4). Note also that the final requirement on the domain of the solution is essential for the general AKNS system. In fact, instantaneous singularities may be caused by vanishing denominators.

Remark 4.1.5. *To the best of the author's knowledge this is the first time that explicit soliton formulas for the AKNS system as a whole are deduced. For the $\mathbb{C}$-reduced AKNS* *system, soliton formulas were derived by Marchenko [55], Matveev [56], Meinel, Neugebauer, Steudel [63] et al.. In the mentioned articles, solutions are expressed by Wronskian determinants. However, these are only evaluated for the $\mathbb{R}$-reduced AKNS system (for example the sine-Gordon equation, but not the Nonlinear Schrödinger equation), where the determinants in the formulas are much simpler (compare Section 4.4).*

For a detailed comparison with existing results for the $\mathbb{C}$-reduced AKNS system, in particular with respect to the representation of solitons in terms of Wronskian determinants, we refer to Appendix A.

Proof Choose $a, c \in \mathbb{C}^n$, $b, d \in \mathbb{C}^m$ all different from zero, and define the generating matrices $A \in \mathcal{M}_{n,n}(\mathbb{C})$, $B \in \mathcal{M}_{m,m}(\mathbb{C})$ in the solution formula of Theorem 4.1.1 as

$$A = \begin{pmatrix} \alpha_1 & & 0 \\ & \ddots & \\ 0 & & \alpha_n \end{pmatrix}, \qquad B = \begin{pmatrix} \beta_1 & & 0 \\ & \ddots & \\ 0 & & \beta_m \end{pmatrix}.$$

Since $\mathrm{Re}(\alpha_i), \mathrm{Re}(\beta_j) > 0 \ \forall i, j$ by assumption, we in particular have $0 \notin \mathrm{spec}(A) + \mathrm{spec}(B)$.

Thus we are in position to apply Theorem 4.1.1. As a result, we obtain a solution q, r of the AKNS system. To fix ideas, we focus on the formula for q which is given by $q = 1 - P/p$, where

$$\begin{aligned} P &= \det\Big(I + \begin{pmatrix} 0 & \widehat{L}\Phi_{A,B}^{-1}(b \otimes c) \\ \widehat{M}\Phi_{B,A}^{-1}(a \otimes d) & -\widehat{M}(b \otimes d) \end{pmatrix}\Big), \\ p &= \det\Big(I + \begin{pmatrix} 0 & \widehat{L}\Phi_{A,B}^{-1}(b \otimes c) \\ \widehat{M}\Phi_{B,A}^{-1}(a \otimes d) & 0 \end{pmatrix}\Big), \end{aligned}$$

with the exponential functions $\widehat{L}(x,t) = \exp\big(Ax + f_0(A)t\big)$, $\widehat{M}(x,t) = \exp\big(Bx - f_0(-B)t\big)$, and the constant operators $\Phi_{A,B}^{-1}(b \otimes c)$, $\Phi_{B,A}^{-1}(a \otimes d)$. In the remainder of the proof we calculate P, p explicitly. We claim:

(i) For $b \in \mathbb{C}^m$, $c \in \mathbb{C}^n$, the unique solution $C := \Phi_{A,B}^{-1}(b \otimes c)$ of the matrix equation $AX + XB = b \otimes c$ is given by

$$C = \left(\frac{b_j c_i}{\alpha_i + \beta_j}\right)_{\substack{i=1,\dots,n \\ j=1,\dots,m}}.$$

(ii) As for the exponential functions, $\widehat{L} \in \mathcal{M}_{n,n}(\mathbb{C})$, $\widehat{M} \in \mathcal{M}_{m,m}(\mathbb{C})$ again are diagonal matrices with the entries $\ell_i = \ell_i(x,t) = \exp\big(\ \alpha_i x + f_0(\alpha_i)t\ \big)$ and $m_j = m_j(x,t) = \exp\big(\ \beta_j x - f_0(-\beta_j)t\ \big)$ on the diagonals, respectively.

Both claims are almost obvious: (i) follows from $b \otimes c = \big(b_j c_i\big)_{\substack{i=1,\dots,n \\ j=1,\dots,m}}$, and (ii) from the definition of the exponential function as a power series.

This provides us with the factorizations $\Phi_{A,B}^{-1}(b \otimes c) = D_c C_0 D_b$, $\Phi_{B,A}^{-1}(a \otimes d) = D_d C_0' D_a$, and $b \otimes d = D_d(e_0 \otimes e_0)D_b$, where

$$C_0 = \left(\frac{1}{\alpha_i + \beta_j}\right)_{\substack{i=1,\dots,n \\ j=1,\dots,m}} \in \mathcal{M}_{n,m}(\mathbb{C}), \qquad e_0 = (1, \dots, 1) \in \mathbb{C}^m,$$

and D_a, D_b, D_c, D_d the diagonal matrices with diagonal given by a, b, c, d, respectively. We infer

$$P = \det\Big(\ I + \begin{pmatrix} 0 & \widehat{L}D_c C_0 D_b \\ \widehat{M}D_d C_0' D_a & -\widehat{M}D_d(e_0 \otimes e_0)D_b \end{pmatrix}\Big)$$

$$= \det\left(I + \begin{pmatrix} \widehat{L}D_c & 0 \\ 0 & \widehat{M}D_d \end{pmatrix} \begin{pmatrix} 0 & C_0 \\ C_0' & -e_0\otimes e_0 \end{pmatrix} \begin{pmatrix} D_a & 0 \\ 0 & D_b \end{pmatrix} \right)$$

$$= \det\left(I + \begin{pmatrix} D_aD_c\widehat{L} & 0 \\ 0 & D_bD_d\widehat{M} \end{pmatrix} \begin{pmatrix} 0 & C_0 \\ C_0' & -e_0\otimes e_0 \end{pmatrix} \right)$$

$$= \det\left(I + \begin{pmatrix} 0 & FC_0 \\ GC_0' & -G(e_0\otimes e_0) \end{pmatrix} \right), \tag{4.3}$$

with $F = D_aD_c\widehat{L}$, $G = D_bD_d\widehat{M}$, and, analogously,

$$p = \det\left(1 + \begin{pmatrix} 0 & FC_0 \\ GC_0' & 0 \end{pmatrix} \right).$$

To calculate P, p, we use the the expansion rule for determinants, see Lemma 4.1.6, together with Lemma 4.1.7. This yields

$$p = 1 + \sum_{k=1}^{n+m} \sum_{\substack{I_1,\dots,I_k=1 \\ I_1<\dots<I_k}}^{n+m} \det\left(\begin{pmatrix} 0 & FC_0 \\ GC_0' & 0 \end{pmatrix}^{(I_1,\dots,I_k)} \right)$$

$$= 1 + \sum_{\kappa=1}^{\min(m,n)} \sum_{\substack{i_1,\dots,i_\kappa=1 \\ i_1<\dots<i_\kappa}}^{n} \sum_{\substack{j_1,\dots,j_\kappa=1 \\ j_1<\dots<j_\kappa}}^{m} \det \begin{pmatrix} 0 & (FC_0)^{\binom{i_1,\dots,i_\kappa}{j_1,\dots,j_\kappa}} \\ (GC_0')^{\binom{j_1,\dots,j_\kappa}{i_1,\dots,i_\kappa}} & 0 \end{pmatrix}.$$

In the latter expression, the notation $T^{\binom{i_1,\dots,i_\kappa}{j_1,\dots,j_\kappa}}$ indicates that we only keep the rows number $i_1,\dots,i_\kappa$ and the columns number $j_1,\dots,j_\kappa$ in T.

We only calculate this determinant for the situation that $\kappa = n \le m$ and $(j_1,\dots,j_n) = (1,\dots,n)$. The remaining cases can be obtained by an obvious renumbering. For notational simplicity we can even assume $m = n$. Then, by Proposition 4.1.10 a),

$$\det\begin{pmatrix} 0 & FC_0 \\ GC_0' & 0 \end{pmatrix} = \det(F)\det(G)\,\det\begin{pmatrix} 0 & C_0 \\ C_0' & 0 \end{pmatrix}$$

$$= (-1)^n \prod_{i=1}^n f_i(x,t)g_i(x,t) \prod_{\substack{i,j=1 \\ i<j}}^n \big((\beta_i-\beta_j)(\alpha_i-\alpha_j)\big)^2 \Big/ \prod_{i,j=1}^n (\alpha_i+\beta_j)^2,$$

where $f_i(x,t) = a_ic_i\ell_i(x,t)$, $g_i(x,t) = b_id_im_i(x,t)$.

In summary, we have shown

$$p = 1 + \sum_{\kappa=1}^{\min(m,n)} (-1)^\kappa \sum_{\substack{i_1,\dots,i_\kappa=1 \\ i_1<\dots<i_\kappa}}^{n} \sum_{\substack{j_1,\dots,j_\kappa=1 \\ j_1<\dots<j_\kappa}}^{m} \widetilde{p}\begin{pmatrix} i_1,\dots,i_\kappa \\ j_1,\dots,j_\kappa \end{pmatrix}.$$

To calculate P, we follow the same path. Because of the one-dimensional perturbation, it is slightly more involved to determine the principal minors in the expansion. Here they contribute whenever, for the dimensions κ of the block in the upper left, λ in the lower right corner, it holds $\lambda = \kappa, \kappa+1$, see Proposition 4.1.9. In the case $\kappa = 0$, the corresponding principal minor reduces to a number. In the case $\kappa > 0$, we again confine to the calculation of the prototypical principal minor, which by Proposition 4.1.10 b) yields

$$\det\begin{pmatrix} 0 & FC_0 \\ GC_0' & -G(e_0\otimes e_0) \end{pmatrix} = \det(F)\det(G)\,\det\begin{pmatrix} 0 & C_0 \\ C_0' & -e_0\otimes e_0 \end{pmatrix} \tag{4.4}$$

$$= (-1)^m \prod_{i=1}^n f_i(x,t) \prod_{j=1}^m g_j(x,t) \prod_{\substack{i,j=1 \\ i<j}}^m (\beta_i-\beta_j)^2 \prod_{\substack{i,j=1 \\ i<j}}^n (\alpha_i-\alpha_j)^2 \Big/ \prod_{i=1}^n \prod_{j=1}^m (\alpha_i+\beta_j)^2$$

for $m = n, n+1$.

As a result,

$$P = p - \sum_{\kappa=1}^{m} g_\kappa + \sum_{\kappa=1}^{\min(m-1,n)} (-1)^{\kappa+1} \sum_{\substack{i_1,\dots,i_\kappa=1\\ i_1<\dots<i_\kappa}}^{n} \sum_{\substack{j_1,\dots,j_{\kappa+1}=1\\ j_1<\dots<j_{\kappa+1}}}^{m} \widetilde{p}\begin{pmatrix} i_1,\dots,i_\kappa \\ j_1,\dots,j_{\kappa+1}\end{pmatrix}.$$

Pasting all formulas together completes the proof for q, and r can be treated in a similar manner. $\square$

To fill in the remaining gaps, we supply the tools for the calculation of determinants which we have used the proof of Theorem 4.1.4. Although some of these techniques will be extended in Chapter 6, we give a complete treatment for the sake of motivation. First we recall the expansion rule for determinants.

Lemma 4.1.6. *Let $T \in \mathcal{M}_{n,n}(\mathbb{C})$ and denote by $T^{(i_1,\dots,i_k)}$ the minor obtained from T by keeping only the rows and columns number $i_1,\dots,i_k$. Then the following expansion of T holds,*

$$\det(I+T) = 1 + \sum_{k=1}^{n} \sum_{\substack{i_1,\dots,i_k=1\\ i_1<\dots<i_k}}^{n} \det\left(T^{(i_1,\dots,i_k)}\right),$$

In particular, the summand corresponding to $k=1$ is the trace of T, the summand corresponding to $k=n$ the determinant of T.

We also need the following observations concerning the evaluation of determinants with a certain block structure.

Lemma 4.1.7. *If $T \in \mathcal{M}_{n,n}(\mathbb{C})$ is a matrix with a k-dimensional block $\widehat{T} = 0$ on the diagonal and $k > n-k$, then $\det(T) = 0$.*

Proof This is shown by a careful expansion of the determinant. $\square$

Lemma 4.1.8. *If $S \in \mathcal{M}_{n,m}(\mathbb{C})$ and $T \in \mathcal{M}_{m,n}(\mathbb{C})$, then*

$$\det\begin{pmatrix} 0 & S \\ T & 0 \end{pmatrix} = \begin{cases} (-1)^n \det(S)\det(T), & n = m, \\ 0, & n \neq m. \end{cases}$$

Proof By Lemma 4.1.7 it remains to calculate the value of the determinant for $m = n$. Since in this case

$$\det\begin{pmatrix} 0 & S \\ T & 0 \end{pmatrix} = \det(S)\det(T) \det\begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix} = (-1)^{n^2} \det(S)\det(T)$$

and $(-1)^{n^2} = (-1)^n$, the asserted formula holds. $\square$

Proposition 4.1.9. *Let $S \in \mathcal{M}_{n,m}(\mathbb{C})$, $T \in \mathcal{M}_{m,n}(\mathbb{C})$ be arbitrary, and b, $d \in \mathbb{C}^m$. Then, for $m < n$ and $n+1 < m$,*

$$\det\begin{pmatrix} 0 & S \\ T & b \otimes d \end{pmatrix} = 0.$$

Proof By Lemma 4.1.7 we have only to consider the case $n+1 < m$. Without loss of generality, b, $d \neq 0$, since otherwise $b \otimes d = 0$, and again Lemma 4.1.7 can be applied. Consequently, there are U, $V \in \mathcal{M}_{m,m}(\mathbb{C})$, both invertible, such that $b = Ue_m^{(1)}$ and $d =$

$Ve_m^{(1)}$, where $e_m^{(1)}$ is the first standard basis vector in $\mathbb{C}^m$. Thus $b \otimes d = (Ue_m^{(1)}) \otimes (Ve_m^{(1)}) = V(e_m^{(1)} \otimes e_m^{(1)})U'$, and it follows

$$\begin{aligned}
\det \begin{pmatrix} 0 & S \\ T & b \otimes d \end{pmatrix} &= \det \left(\begin{pmatrix} I & 0 \\ 0 & V \end{pmatrix} \begin{pmatrix} 0 & S(U')^{-1} \\ V^{-1}T & e_m^{(1)} \otimes e_m^{(1)} \end{pmatrix} \begin{pmatrix} I & 0 \\ 0 & U' \end{pmatrix} \right) \\
&= \det(U)\det(V) \det \begin{pmatrix} 0 & S(U')^{-1} \\ V^{-1}T & e_m^{(1)} \otimes e_m^{(1)} \end{pmatrix}.
\end{aligned}$$

Therefore we can assume $b = d = e_m^{(1)}$. Note that $e_m^{(1)} \otimes e_m^{(1)}$ is an $m \times m$-matrix with the $(1,1)$-entry being the only non-vanishing entry. Set

$$R = \begin{pmatrix} 0 & S \\ T & e_m^{(1)} \otimes e_m^{(1)} \end{pmatrix}.$$

Expanding $\det(R)$ with respect to the $(n+1)$-th row, we get

$$\det(R) = \det(R_{n+1}) + \sum_{\mu=1}^{n} (-1)^{n+1+\mu}\, t_{1\mu} \, \det(R_\mu),$$

where R_μ arises from R by deleting the $(n+1)$-th row and the μ-th column, and $t_{1\mu}$ are the entries in the first row of T. Taking a closer look at the structure of the matrices R_μ, we find that all of them have a zero block of dimension $m-1$ in the lower right corner. Since $m-1 > n$, we can apply Lemma 4.1.7, which shows $\det(R_\mu) = 0$ for $\mu = 1, \ldots, n+1$. This completes the proof. $\square$

Finally we need to evaluate some particular determinants.

Proposition 4.1.10. *Let α_i, $i = 1, \ldots, n$, and β_j, $j = 1, \ldots, m$, be complex numbers such that $\alpha_i + \beta_j \neq 0$ for all i, j. Define*

$$C_0 = \left(\frac{1}{\alpha_i + \beta_j} \right)_{\substack{i=1,\ldots,n \\ j=1,\ldots,m}} \in \mathcal{M}_{n,m}(\mathbb{C}), \qquad e_0 = (1, \ldots, 1) \in \mathbb{C}^m.$$

Moreover, we define the constants

$$\gamma_{nm} = (-1)^m \prod_{\substack{i,j=1 \\ i<j}}^{m} (\beta_i - \beta_j)^2 \prod_{\substack{i,j=1 \\ i<j}}^{n} (\alpha_i - \alpha_j)^2 \Big/ \prod_{i=1}^{n} \prod_{j=1}^{m} (\alpha_i + \beta_j)^2.$$

Then the following statements hold:

a) If $m = n$, then

$$\det \begin{pmatrix} 0 & C_0 \\ C_0' & 0 \end{pmatrix} = \gamma_{nn}.$$

b) If $m \in \{n, n+1\}$, then

$$\det \begin{pmatrix} 0 & C_0 \\ C_0' & -e_0 \otimes e_0 \end{pmatrix} = \gamma_{nm}.$$

Proposition 4.1.10 extends a well-known remark of Cauchy (see [20], p. 151-159, [78], VII, §1, Nr. 3, and Lemma 6.1.2). In fact, Lemma 6.1.2 yields the value of $\det(C_0)$ for $m = n$.

A further generalization of Proposition 4.1.10 will be given in Theorem 6.2.1. Nevertheless we provide the argument because it is short and motivates the much more complicated later extensions.

Proof a) The assertion is an immediate consequence of Lemma 4.1.8 and Lemma 6.1.2.
b) Without loss of generality the α_i are pairwise different. Otherwise C_0 would contain linearly dependent rows, and the assertion would be clear. Analogously we can assume the β_j to be pairwise different.

First we treat the case $m = n$. Then C_0 is a square matrix, and invertible by Lemma 6.1.2. Hence,

$$\det \begin{pmatrix} 0 & C_0' \\ C_0 & -e_0 \otimes e_0 \end{pmatrix} = \det \begin{pmatrix} I & 0 \\ -(e_0 \otimes e_0)(C_0')^{-1} & I \end{pmatrix} \det \begin{pmatrix} 0 & C_0' \\ C_0 & 0 \end{pmatrix},$$

where the first determinant is obviously equal to 1. Thus a) can be applied.

It remains to consider the case $m = n+1$. Here the basic idea of the proof is to imitate the strategy of [20]:

(i) (Manipulations with respect to the last m columns) Subtract the $(n+m)$-th column from the $(n+j)$-th column for $j = 1, \ldots, m-1$. Then the entries of the $(n+j)$-th column, $j < m$, become

$$\frac{\beta_m - \beta_j}{\alpha_i + \beta_m} \frac{1}{\alpha_i + \beta_j} \quad \text{for } i \le n \qquad \text{and} \qquad 0 \quad \text{for } i > n,$$

and extract common factors $(\beta_m - \beta_j)$ in the $(n+j)$-th column, $j = 1, \ldots, m-1$, and $1/(\alpha_i + \beta_m)$ in the i-th row, $i = 1, \ldots, n$.

(ii) (Manipulations with respect to the last m rows) Subtract the $(n+m)$-th row from the $(n+i)$-th row for $i = 1, \ldots, m-1$. Then the entries of the $(n+i)$-th row, $i < m$, become

$$\frac{\beta_m - \beta_i}{\alpha_j + \beta_m} \frac{1}{\alpha_j + \beta_i} \quad \text{for } j = 1 \le n \qquad \text{and} \qquad 0 \quad \text{for } j > n,$$

and extract common factors $(\beta_m - \beta_i)$ in the $(n+i)$-th row, $i = 1, \ldots, m-1$, and $1/(\alpha_j + \beta_m)$ in the j-th column, $j = 1, \ldots, n$.

As a result,

$$\det \begin{pmatrix} 0 & C_0 \\ C_0' & -e_0 \otimes e_0 \end{pmatrix} = \prod_{i=1}^{n} \left[\frac{\beta_{n+1} - \beta_i}{\alpha_i + \beta_{n+1}} \right]^2 \det \begin{pmatrix} 0 & \widehat{C}_0 & \widehat{e}_0 \\ \widehat{C}_0{}' & 0 & 0 \\ \widehat{e}_0{}' & 0 & -1 \end{pmatrix},$$

where $\widehat{C}_0$ is obtained from C_0 by deleting the last column and $\widehat{e}_0 \in \mathbb{C}^n$ denotes the vector with entries all equal to 1. Note also that we have used $m = n+1$.

Next we add the $(n+m)$-th column to the i-th column for $i = 1, \ldots, n$, and the $(n+m)$-th row to the j-th row for $j = 1, \ldots, n$. This yields

$$\det \begin{pmatrix} 0 & \widehat{C}_0 & \widehat{e}_0 \\ \widehat{C}_0{}' & 0 & 0 \\ \widehat{e}_0{}' & 0 & -1 \end{pmatrix} = \det \begin{pmatrix} \widehat{e}_0 \otimes \widehat{e}_0 & \widehat{C}_0 & 0 \\ \widehat{C}_0{}' & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} = -\det \begin{pmatrix} \widehat{e}_0 \otimes \widehat{e}_0 & \widehat{C}_0 \\ \widehat{C}_0{}' & 0 \end{pmatrix},$$

the latter by expanding. Since $m = n+1$, all blocks of the latter determinant are square matrices of dimension n. Thus a similar argument as in the case $m = n$ applies, and, pasting together the formulas, we can conclude the proof. □

4.2 Reductions of the AKNS system

The AKNS system (1.1) does not have good stability properties in general (see [3], p.278, for an example of a solution with smooth initial data which becomes unbounded after a finite time). To arrive at soliton equations, the system has to be reduced. In practice this means that we impose a relation on the two unknown functions q, r, and require a symmetry condition for f_0.

The reductions we will consider already appear in [3]. In this paper, the reduction allows to carry out the inverse scattering method completely. In our method the treatment of the unreduced and the reduced AKNS system does not differ formally. But it should be observed that reduction is the right way to ensure global regularity.

More precisely, we consider two types of reductions. In the first, which we call the complex reduction of the AKNS system, or $\mathbb{C}$-reduced AKNS system for short, we suppose that, for two given polynomials f, g, the rational function $f_0 = f/g$ satisfies

$$\overline{f_0(z)} = -f_0(-\overline{z}) \quad \text{for all } z \in \mathbb{C} \text{ where } f_0 \text{ is holomorphic.}$$

Then the $\mathbb{C}$-reduced AKNS system reads

$$\begin{cases} g(T_{r,q}) \begin{pmatrix} r_t \\ q_t \end{pmatrix} = f(T_{r,q}) \begin{pmatrix} r \\ -q \end{pmatrix} \\ r = -\overline{q}. \end{cases} \tag{4.5}$$

For the real reduction of the AKNS system, briefly $\mathbb{R}$-reduced AKNS, we suppose

$$f_0(z) = -f_0(-z) \quad \text{for all } z \in \mathbb{C} \text{ where } f_0 \text{ is holomorphic.}$$

Then the $\mathbb{R}$-reduced AKNS reads

$$\begin{cases} g(T_{r,q}) \begin{pmatrix} r_t \\ q_t \end{pmatrix} = f(T_{r,q}) \begin{pmatrix} r \\ -q \end{pmatrix} \\ r = -q. \end{cases} \tag{4.6}$$

In this context one is mainly interested in real solutions. Then one has to assume in addition that f_0 is real in the sense that $f_0(x) \in \mathbb{R}$ for all $x \in \mathbb{R}$ where f_0 is finite. If one assumes reality of f_0 (what we will not do) one can regard $\mathbb{R}$-reduction as a further specialization of $\mathbb{C}$-reduction.

In what follows we shall consider (4.5), (4.6), respectively, as an equation for the single unknown function q.

We want to stress that these reductions comprise all important soliton equations of the AKNS system. For example, the Nonlinear Schrödinger equation is contained in the $\mathbb{C}$-reduced AKNS, the sine-Gordon and modified Korteweg-de Vries equation are contained in the $\mathbb{R}$-reduced AKNS.

4.3 The $\mathbb{C}$-reduced AKNS system, negatons, and their regularity properties

This section will contain the first step torwards to the study of negatons. We start by providing the $\mathbb{C}$-reduction of the solution formula given in Theorem 4.1.1. This will lead to appropriate conditions on the eigenvalues and thereby to the notion of negatons.

Afterwards we discuss structural properties of this solution class, mainly the regularity question. At the end we explain how the familiar N-solitons integrate into this picture.

4.3.1 Negatons

Starting point of this section is the solution formula in Theorem 4.1.1.

Recall that, for the $\mathbb{C}$-reduced AKNS system, $\overline{f_0(z)} = -f_0(-\overline{z})$ holds for all $z \in \mathbb{C}$ where f_0 is finite. To realize the linear relation $r = -\overline{q}$ for the $\mathbb{C}$-reduced AKNS system, we use the following simple fact:

$$\overline{\det(1+T)} = \det(1+\overline{T}) \qquad \text{for } T \in \mathcal{M}_{n,n}(\mathbb{C}),$$

where $\overline{T}$, which we will call the complex conjugate of T, is the matrix with the complex conjugate entries of T.

Proposition 4.3.1. *Assume that $A \in \mathcal{M}_{n,n}(\mathbb{C})$ satisfies $\operatorname{spec}(A) \subseteq \{z \in \mathbb{C} \mid \operatorname{Re}(z) > 0$ and $f_0(z)$ is finite$\}$ and let $0 \neq a,\ c \in \mathbb{C}^n$ be arbitrary. Define the operator-functions*

$$\begin{aligned} L(x,t) &= \exp\big(\, Ax + f_0(A)t \,\big)\, \Phi^{-1}_{A,\overline{A}}(\overline{a} \otimes c), \\ L_0(x,t) &= \exp\big(\, Ax + f_0(A)t \,\big)\, (a \otimes c). \end{aligned}$$

Then a solution of the $\mathbb{C}$-reduced AKNS system (4.5) is given by

$$q = 1 - P/p, \text{ where}$$

$$P = \det\begin{pmatrix} I & -L \\ \overline{L} & I + \overline{L_0} \end{pmatrix}, \qquad p = \det\begin{pmatrix} I & -L \\ \overline{L} & I \end{pmatrix},$$

on every strip $\mathbb{R} \times (t_1, t_2)$ on which the denominator p does not vanish.

Proof Apply Theorem 4.1.1 with (i) $B = \overline{A}$ and (ii) $b = -\overline{a}$, $d = \overline{c}$. By Lemma 7.1.6, we have $\operatorname{spec}(-\overline{A}) \subseteq \{z \in \mathbb{C} \mid f_0(z) \text{ is finite}\}$, and $-f_0(-\overline{A}) = \overline{f_0(A)}$. Hence (i) implies that the operator functions $\widehat{M}$ and $\widehat{L}$ are complex conjugate to each other.

Set $C = \Phi^{-1}_{A,\overline{A}}(\overline{a} \otimes c)$. By definition, C is the unique solution of the matrix equation $AX + X\overline{A} = \overline{a} \otimes c$. Thus, from

$$\overline{AC} + \overline{C}A = \overline{AC + C\overline{A}} = \overline{\overline{a} \otimes c} = a \otimes \overline{c},$$

we infer $\Phi^{-1}_{\overline{A},A}(a \otimes \overline{c}) = \overline{C}$.

By (i), (ii), this yields $\Phi^{-1}_{A,B}(b \otimes c) = -C$, $\Phi^{-1}_{B,A}(a \otimes d) = \overline{C}$, and the solution formula in Theorem 4.1.1 reads $q = 1 - P/p$, $r = 1 - \widehat{P}/p$, where

$$P = \det\begin{pmatrix} I & -\widehat{L}C \\ \overline{\widehat{L}C} & I + \overline{\widehat{L}}(\overline{a} \otimes \overline{c}) \end{pmatrix}, \quad \widehat{P} = \det\begin{pmatrix} I - \widehat{L}(a \otimes c) & -\widehat{L}C \\ \overline{\widehat{L}C} & I \end{pmatrix},$$

$$\text{and} \qquad p = \det\begin{pmatrix} I & -\widehat{L}C \\ \overline{\widehat{L}C} & I \end{pmatrix}.$$

It remains to show the linear relation $r = -\overline{q}$ for the $\mathbb{C}$-reduced AKNS system. To this end, we calculate

$$\begin{aligned} \overline{P} &= \det\Bigg(I + \overline{\begin{pmatrix} 0 & -\widehat{L}C \\ \overline{\widehat{L}C} & \overline{\widehat{L}}(\overline{a} \otimes \overline{c}) \end{pmatrix}} \Bigg) \\ &= \det\Bigg(I + \begin{pmatrix} 0 & -\overline{\widehat{L}C} \\ \widehat{L}C & \widehat{L}(a \otimes c) \end{pmatrix} \Bigg) \\ &= \det\Bigg(I + \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \begin{pmatrix} \widehat{L}(a \otimes c) & -\widehat{L}C \\ \overline{\widehat{L}C} & 0 \end{pmatrix} \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix} \Bigg) \\ &= \det\Bigg(I + \begin{pmatrix} \widehat{L}(a \otimes c) & -\widehat{L}C \\ \overline{\widehat{L}C} & 0 \end{pmatrix} \Bigg). \end{aligned}$$

Analogously, $\overline{p} = p$.

Finally, using Proposition 2.4.3, we observe

$$
\begin{aligned}
-\overline{q} \;&=\; -1+\frac{\overline{P}}{p} \;=\; -1+\frac{\det\left(I+\begin{matrix}\widehat{L}(a\otimes c) & -\widehat{L}C\\ \overline{\widehat{L}C} & 0\end{matrix}\right)}{\det\left(I+\begin{matrix}0 & -\widehat{L}C\\ \overline{\widehat{L}C} & 0\end{matrix}\right)}\\
&=\; 1-\frac{\det\left(I+\begin{matrix}-\widehat{L}(a\otimes c) & -\widehat{L}C\\ \overline{\widehat{L}C} & 0\end{matrix}\right)}{\det\left(I+\begin{matrix}0 & -\widehat{L}C\\ \overline{\widehat{L}C} & 0\end{matrix}\right)} \;=\; 1-\frac{\widehat{P}}{p} \;=\; r.
\end{aligned}
$$

This completes the proof. $\square$

Lemma 4.3.2. *Let $A\in\mathcal{M}_{n,n}(\mathbb{C})$ be as in Proposition 4.3.1 and U a matrix transforming A into Jordan form J_A, namely $A=U^{-1}J_AU$.*

Then the solution formula of Proposition 4.3.1 is not altered if we replace simultaneously A by J_A and the vectors a, c by $(U^{-1})'a$, Uc.

Proof In fact, Lemma 4.3.2 is a corollary of Lemma 4.1.2. To see this, we only have to check that the replacements in Lemma 4.1.2 are compatible with the special choices (i), (ii) made in the proof of Proposition 4.3.1. But by (i) it is possible to take $V:=\overline{U}$, and then it is evident that the replacement of the vectors in Lemma 4.1.2 fits to (ii). $\square$

By Lemma 4.3.2 any solution coming from a finite matrix can also be generated by its Jordan form. From elementary linear algebra we recall that the Jordan form is essentially unique. We call negaton any solution generated by a Jordan matrix A with the following property:

The eigenvalues $\alpha_i\in\operatorname{spec}(A)$ satisfy $\operatorname{Re}(\alpha_i)>0$ for all i.

Our notion of negaton is very general. For example, it comprises for the sine-Gordon equation usual solitons (kinks), breathers, as well as groups of solitons drifting apart with logarithmic velocity. The latter type may be called negatons in the closer sense (compare [58], [79]). Sometimes we will use the term negaton also in the more specific sense.

4.3.2 Regularity conditions

Next we prove that the solutions of the $\mathbb{C}$-reduced AKNS system in Proposition 4.3.1 are globally regular. The main ingredient is Proposition 4.3.6. As a first step we prove the following factorization result.

Proposition 4.3.3. *Let A, $B\in\mathcal{M}_{n,n}(\mathbb{C})$ be in Jordan form: $A=\operatorname{diag}\{A_j\,|\,j=1,\dots,N\}$, $B=\operatorname{diag}\{B_j\mid j=1,\dots,N\}$, with Jordan blocks A_j, B_j of dimension n_j corresponding to the eigenvalues α_j, β_j, respectively. Assume $\operatorname{Re}(\alpha_j)$, $\operatorname{Re}(\beta_j)>0$ $\forall j$.*

Let a, $c\in\mathbb{C}^n$ (when decomposed according to the Jordan form, i.e., $a=(a_j)_{j=1}^N$ with $a_j=(a_j^{(\mu)})_{\mu=1}^{n_j}$) satisfy $a_j^{(1)}c_j^{(n_j)}\neq 0$ $\forall j$.

Then the following factorization holds:

$$\Phi_{A,B}^{-1}(a\otimes c)=ST,$$

where

$$T:\bigoplus_{i=1}^{N}\mathbb{C}^{n_i}\longrightarrow L_2(-\infty,0],\qquad S:L_2(-\infty,0]\longrightarrow\bigoplus_{i=1}^{N}\mathbb{C}^{n_i},$$

are defined by

$$\begin{aligned}(Te_{ii'})(\sigma) &= \frac{1}{(i'-1)!}\frac{\partial^{i'-1}}{\partial\beta_i^{i'-1}}\Big(e^{\beta_i\sigma+\psi(\beta_i)}\Big),\\ Sf &= \sum_{i=1}^{N}\sum_{i'=1}^{n_i}\Big(\int_{-\infty}^{0} f(s)\frac{1}{(n_i-i')!}\frac{\partial^{n_i-i'}}{\partial\alpha_i^{n_i-i'}}e^{\alpha_i s+\varphi(\alpha_i)}ds\Big)e_{ii'},\end{aligned}$$

and φ, ψ are functions of one complex variable z which fulfill the set of conditions

$$\frac{1}{(j'-1)!}\frac{\partial^{j'-1}}{\partial z^{j'-1}}e^{\varphi(z)}\Big|_{z=\alpha_j}=c_j^{(n_j-j'+1)},\qquad \frac{1}{(j'-1)!}\frac{\partial^{j'-1}}{\partial z^{j'-1}}e^{\psi(z)}\Big|_{z=\beta_j}=a_j^{(j')},\quad (4.7)$$

for $j=1,\ldots,N$, $j'=1,\ldots,n_j$.

Moreover, TS is the Fredholm integral operator $R_h\in\mathcal{L}\left(L_2(-\infty,0]\right)$ defined by

$$(R_hf)(s)=\int_{-\infty}^{0} f(\sigma)h(s,\sigma)d\sigma$$

$$\text{with kernel } h(s,\sigma)=\sum_{i=1}^{N}\frac{1}{(n_i-1)!}\left(\frac{\partial}{\partial\alpha_i}+\frac{\partial}{\partial\beta_i}\right)^{n_i-1}e^{\alpha_i s+\beta_i\sigma+\chi(\alpha_i,\beta_i)},$$

where $\chi(z_1,z_2)=\varphi(z_1)+\psi(z_2)$.

According to the decomposition $\mathbb{C}^n=\bigoplus_{\lambda=1}^{N}\mathbb{C}^{n_\lambda}$, $e_{ii'}\in\mathbb{C}^n$ denotes the vector which is the i'-th standard basis vector on the component $\mathbb{C}^{n_i}$ and the zero-vector on all other components $\mathbb{C}^{n_\lambda}$, $\lambda\neq i$.

Proof First note that (4.7) is always solvable, since we have assumed $a_j^{(1)}c_j^{(n_j)}\neq 0$. Next, calculating the entries of $ST\in\mathcal{M}_{n,n}(\mathbb{C})$, we observe

$$\begin{aligned}\langle STe_{jj'},e_{ii'}\rangle &= \int_{-\infty}^{0}\frac{1}{(j'-1)!}\frac{\partial^{j'-1}}{\partial\beta_j^{j'-1}}e^{\beta_j s+\psi(\beta_j)}\frac{1}{(n_i-i')!}\frac{\partial^{n_i-i'}}{\partial\alpha_i^{n_i-i'}}e^{\alpha_i s+\varphi(\alpha_i)}ds\\ &= \sum_{\mu=0}^{j'-1}\sum_{\nu=0}^{n_i-i'}\frac{1}{(j'-1-\mu)!}\frac{\partial^{j'-1-\mu}}{\partial\beta_j^{j'-1-\mu}}e^{\psi(\beta_j)}\frac{1}{(n_i-i'-\nu)!}\frac{\partial^{n_i-i'-\nu}}{\partial\alpha_i^{n_i-i'-\nu}}e^{\varphi(\alpha_i)}\\ &\qquad\cdot\int_{-\infty}^{0}\frac{1}{\mu!}\frac{\partial^\mu}{\partial\beta_j^\mu}e^{\beta_j s}\frac{1}{\nu!}\frac{\partial^\nu}{\partial\alpha_i^\nu}e^{\alpha_i s}ds\\ &= \sum_{\mu=0}^{j'-1}\sum_{\nu=0}^{n_i-i'}a_j^{(j'-\mu)}c_i^{(i'+\nu)}(-1)^{\mu+\nu}\binom{\mu+\nu}{\mu}\left(\frac{1}{\alpha_i+\beta_j}\right)^{\mu+\nu+1}.\end{aligned}$$

For the last identity we used, beside the defining relations (4.7) for the functions φ, ψ, the elementary fact that

$$\int_{-\infty}^{0}\frac{\partial^\mu}{\partial\beta^\mu}\frac{\partial^\nu}{\partial\alpha^\nu}e^{(\alpha+\beta)s}ds=\int_{-\infty}^{0}s^{\mu+\nu}e^{(\alpha+\beta)s}ds=(-1)^{\mu+\nu}(\mu+\nu)!\left(\frac{1}{\alpha+\beta}\right)^{\mu+\nu+1}.$$

Comparison with Proposition 4.1.3, where $\Phi_{A,B}^{-1}(a \otimes c) = (C_{ij})_{i,j=1}^{N}$ was determined explicitly, shows that $\langle STe_{jj'}, e_{ii'} \rangle$ coincides with the $i'j'$-th entry of C_{ij}. Thus $\Phi_{A,B}^{-1}(a \otimes c) = ST$, and we have obtained a factorization through the Hilbert space $L_2(-\infty, 0]$. Vice versa,

$$\begin{aligned}
&(TSf)(\sigma) = \\
&= \sum_{i=1}^{N} \sum_{i'=1}^{n_i} \int_{-\infty}^{0} f(s) \frac{1}{(n_i - i')!} \frac{\partial^{n_i - i'}}{\partial \alpha_i^{n_i - i'}} e^{\alpha_i s + \varphi(\alpha_i)} ds \cdot \frac{1}{(i'-1)!} \frac{\partial^{i'-1}}{\partial \beta_i^{i'-1}} e^{\beta_i \sigma + \psi(\beta_i)} \\
&= \int_{-\infty}^{0} f(s) \left[\sum_{i=1}^{N} \frac{1}{(n_i - 1)!} \sum_{i'=1}^{n_i} \binom{n_i - 1}{i' - 1} \frac{\partial^{n_i - i'}}{\partial \alpha_i^{n_i - i'}} e^{\alpha_i s + \varphi(\alpha_i)} \frac{\partial^{i'-1}}{\partial \beta_i^{i'-1}} e^{\beta_i \sigma + \psi(\beta_i)} \right] ds \\
&= \int_{-\infty}^{0} f(s) \left[\sum_{i=1}^{N} \frac{1}{(n_i - 1)!} \left(\frac{\partial}{\partial \alpha_i} + \frac{\partial}{\partial \beta_i} \right)^{n_i - 1} e^{\alpha_i s + \beta_i \sigma + \varphi(\alpha_i) + \psi(\beta_i)} \right] ds \\
&= (T_h f)(\sigma).
\end{aligned}$$

This completes the proof. $\square$

Remark 4.3.4. *Note that Proposition 4.3.3 and its proof also yields the following two statements:*

a) If $\alpha_j = \beta_j$ for all j, then the kernel of the Fredholm integral operator R_h is even given by

$$h(s, \sigma) = \sum_{j=1}^{N} \frac{\partial^{n_j - 1}}{\partial \alpha_j^{n_j - 1}} \exp\Big(\alpha_j (s + \sigma) + \chi(\alpha_j) \Big)$$

with $\chi(z) = \varphi(z) + \psi(z)$.

b) Let us assume that φ, ψ are already constructed as in Proposition 4.3.3. If we start anew from the complex conjugate data, namely the eigenvalues $\overline{\alpha}_j$, $\overline{\beta}_j$, $j = 1, \ldots, N$, and the vectors $\overline{a}$, $\overline{c}$, then (4.7) shows that we can choose the corresponding functions $\widehat{\varphi}$, $\widehat{\psi}$ such that $\widehat{\varphi}(\overline{z}) = \overline{\varphi(z)}$, $\widehat{\psi}(\overline{z}) = \overline{\psi(z)}$ for all z. In particular, we have $\widehat{h} = \overline{h}$ for the respective integral kernels.

Remark 4.3.5. *As a byproduct, Proposition 4.3.3 establishes a nice relation to the methods of Pöppe [10], [75], who used Fredholm integral operators for the construction of solutions to soliton equations. For the sine-Gordon equation, [75] constructed multiple pole solutions. Our result shows in particular how these fit in the general concept of negatons (see [88]).*

Proposition 4.3.6. *Let the requirements of Proposition 4.3.1 be met with $A \in \mathcal{M}_{n,n}(\mathbb{C})$ in Jordan form: $A = diag\{A_j \,|\, j = 1, \ldots, N\}$ with Jordan blocks A_j of dimension n_j corresponding to the eigenvalue α_j. Then $L\overline{L}$ is related to the integral operator R on $L_2(-\infty, 0]$ given by*

$$(Rf)(s) = \int_{-\infty}^{0} f(\sigma) K(s, \sigma) d\sigma \quad \text{with kernel } K(s, \sigma) = \int_{\infty}^{0} k(s + \rho)\, \overline{k(\sigma + \rho)}\, d\rho,$$

where

$$k(\rho) = \sum_{j=1}^{N} \frac{1}{(n_j - 1)!} \frac{\partial^{n_j - 1}}{\partial \alpha_j^{n_j - 1}} \exp\big(\alpha_j \rho + \chi(\alpha_j) \big),$$

and $\chi = \chi(\alpha; x, t)$ is an appropriate $\mathcal{C}^\infty$-function whose derivatives with respect to α at α_j satisfy the property $\overline{\chi^{(\mu)}(\alpha_j)} = \chi^{(\mu)}(\overline{\alpha}_j)$ for $\mu = 0, \ldots, n_j - 1$.

In particular, R is self-adjoint and positive.

Proof Set $\widehat{L}(x,t) = \exp\big(Ax + f_0(A)t\big)$ and $C = \Phi^{-1}_{A,\overline{A}}(\overline{a}\otimes c)$. Thus $L = \widehat{L}C$. Since $\widehat{L}$ commutes with A,

$$AL + L\overline{A} = \widehat{L}(AC + C\overline{A}) = \widehat{L}(\overline{a}\otimes c) = \overline{a}\otimes(\widehat{L}c),$$

which means that $L = \Phi^{-1}_{A,\overline{A}}\big(\overline{a}\otimes(\widehat{L}c)\big)$. Now we are in position to apply Proposition 4.3.3, which yields the factorization

$$L\overline{L} = (S\overline{T})(\overline{S}T),$$

where $T,\, \overline{T} : \bigoplus_{\mu=1}^{N}\mathbb{C}^{n_\mu} \to L_2(-\infty,0]$ are given by

$$Te_{jj'} = f_j^{(j')}, \qquad \overline{T}e_{jj'} = \overline{f_j^{(j')}},$$
$$\text{with } f_j^{(j')}(s) = \frac{1}{(j'-1)!}\frac{\partial^{j'-1}}{\partial\alpha_j^{j'-1}}\exp\Big(\alpha_j s + \psi(\alpha_j)\Big),$$

and $S,\, \overline{S} : L_2(-\infty,0] \to \bigoplus_{\mu=1}^{N}\mathbb{C}^{n_\mu}$ are given by

$$\langle Sf, e_{jj'}\rangle = \int_{-\infty}^{0} f(s)g_j^{(j')}(s)ds, \qquad \langle\overline{S}f, e_{jj'}\rangle = \int_{-\infty}^{0} f(s)\overline{g_j^{(j')}(s)}ds,$$
$$\text{with } g_j^{(j')}(s) = \frac{1}{(n_j-j')!}\frac{\partial^{n_j-j'}}{\partial\alpha_j^{n_j-j'}}\exp\Big(\alpha_j s + \phi(\alpha_j)\Big).$$

The functions $\psi = \psi(\alpha;x,t)$, $\phi = \phi(\alpha;x,t)$ were appropriately constructed. They can be chosen such that its derivatives with respect to α at α_j up to order n_j-1 satisfy $\overline{\phi^{(\mu)}(\alpha_j)} = \phi^{(\mu)}(\overline{\alpha}_j)$, $\overline{\psi^{(\mu)}(\alpha_j)} = \psi^{(\mu)}(\overline{\alpha}_j)$ (for details see Proposition 4.3.3).

Here $e_{jj'} \in \mathbb{C}^n = \bigoplus_{\mu=1}^{N}\mathbb{C}^{n_\mu}$ is the vector which is the j'-th standard basis vector $e_{n_j}^{(j')}$ on the component $\mathbb{C}^{n_j}$ and zero else.

As a consequence, $L\overline{L}$ is related to the operator R on $L_2(-\infty,0]$ given by

$$R = (TS)(\overline{TS}),$$

where by Proposition 4.3.3 (see also Remark 4.3.4) shows that TS and $\overline{TS}$ are integral operators with kernels that are complex conjugated, namely

$$(TSf)(s) = \int_{-\infty}^{0} f(\sigma)k(s+\sigma)d\sigma, \qquad (\overline{TS}f)(s) = \int_{-\infty}^{0} f(\sigma)\overline{k(s+\sigma)}d\sigma,$$

where $k(\rho) = \displaystyle\sum_{j=1}^{N}\frac{1}{(n_j-1)!}\frac{\partial^{n_j-1}}{\partial\alpha_j^{n_j-1}}\exp\Big(\alpha_j\rho + \chi(\alpha_j)\Big)$,
and $\chi = \psi + \phi$. In particular, the derivatives of χ with respect to α at α_j up to order $n_j - 1$ again satisfy $\overline{\chi^{(\mu)}(\alpha_j)} = \chi^{(\mu)}(\overline{\alpha}_j)$.

The composition of those two integral operators can be easily calculated to be again an integral operator, namely

$$(Rf)(s) = \int_{-\infty}^{0} f(\sigma)K(s,\sigma)d\sigma$$

with kernel $K(s,\sigma) = \displaystyle\int_{-\infty}^{0} k(s+\rho)\overline{k(\sigma+\rho)}d\rho.$

Observe $\overline{K(s,\sigma)} = K(\sigma, s)$. Therefore, the usual argument $(\ Rg, f\ )=(\ g, Rf\ )$ for all $f,\ g \in L_2(-\infty, 0]$, together with the fact that R is a bounded operator, implies that R is self-adjoint.

It remains to prove that R is positive. To this end, we calculate

$$\begin{aligned}
(\ Rf, f\ ) &= \int_{-\infty}^{0} \int_{-\infty}^{0} f(\sigma) K(s,\sigma) d\sigma \quad \overline{f(s)} ds \\
&= \int_{-\infty}^{0} \left(\int_{-\infty}^{0} k(s+\rho)\overline{f(s)} ds \right) \left(\int_{-\infty}^{0} \overline{k(\sigma+\rho)} f(\sigma) d\sigma \right) d\rho \\
&= \int_{-\infty}^{0} \left| \int_{-\infty}^{0} k(s+\rho)\overline{f(s)} ds \right|^2 \geq \ 0
\end{aligned}$$

for all $f \in L_2(-\infty, 0]$. This completes the proof. $\square$

As an immediate consequence, we state that negatons are globally regular.

Proposition 4.3.7. *The solutions given in Proposition 4.3.1 are defined and regular on all of* $\mathbb{R}^2$.

Proof Without loss of generality A is in Jordan canonical form (see Lemma 4.3.2). To prove regularity of the solutions in Proposition 4.3.1 it is sufficient, thanks to Lemma 4.3.8, to show $\det\left(I + L\overline{L}\right) > 0$. Actually, we will even show

$$\det\left(I + L\overline{L}\right) \geq 1. \tag{4.8}$$

This is done exploiting the fact that $L\overline{L}$ is related to the integral operator R on $L_2(-\infty, 0]$ in Proposition 4.3.6. Since R is positive, its eigenvalues are non-negative. By the principle of related operators, see Proposition B.2.3, also $L\overline{L}$ has only non-negative eigenvalues. Hence $\det(I + L\overline{L}) \geq 1$. $\square$

We supply the relation for determinants used in the preceding proof.

Lemma 4.3.8. *For all $S,\ T \in \mathcal{M}_{n,n}(\mathbb{C})$, the following relation holds*

$$\det \begin{pmatrix} I & -S \\ T & I \end{pmatrix} = \det\left(I + ST\right).$$

In comparison to the analogous statement on the operator-level in Proposition 2.4.1, no assumption on S, T is needed here.

Proof In the finite-dimensional situation, this proof simply relies on the following elementary manipulations of the determinant on the left-hand side: Multiply the blocks in the second row by S and add them to the blocks in the first row. The result is

$$\det \begin{pmatrix} I & -S \\ T & I \end{pmatrix} = \det \begin{pmatrix} I+ST & 0 \\ T & I \end{pmatrix},$$

and the assertion follows. $\square$

4.3.3 Solitons and their characteristic properties

In this section, we construct N-solitons for the $\mathbb{C}$-reduced AKNS system and report on their characteristic property: They collide elastically without change of velocity and shape, the only effect is the so-called phase-shift. In mathematical language this behaviour can be understood in terms of their asymptotic behaviour.

Explicit formula for N-solitons

In the sequel we show that generating matrices A in diagonal form in Proposition 4.3.1 lead to superpositions of solitons.

Assume $A = \mathrm{diag}\{\alpha_i \mid i = 1, \dots, N\}$ such that we have $\mathrm{Re}(\alpha_i) > 0$ and $f_0(\alpha_i)$ is finite $\forall i$. Then the operators L, L_0 in Proposition 4.3.1 read

$$\begin{aligned} L(x,t) &= \left(\frac{\overline{a}_j c_i}{\alpha_i + \overline{\alpha}_j}\, \widehat{\ell}_i(x,t) \right)_{i,j=1}^{N}, \\ L_0(x,t) &= \left(a_j c_i\, \widehat{\ell}_i(x,t) \right)_{i,j=1}^{N}, \end{aligned}$$

where $\widehat{\ell}_i(x,t) = \exp\big(\alpha_i x + f_0(\alpha_i) t\big)$ (confer the arguments in the proof of Theorem 4.1.4).

The explicit formula for the resulting N-solitons is the following.

Proposition 4.3.9. *For complex numbers α_j, $j = 1, \dots, N$, which are pairwise different, satisfy $\mathrm{Re}(\alpha_j) > 0$ $\forall j$, and are contained in the set where f_0 is holomorphic, a solution of the $\mathbb{C}$-reduced AKNS system (4.5) is given by*

$$q = -\frac{1}{p} \Big[\sum_{j=1}^{N} \overline{\ell}_j + \sum_{\kappa=1}^{N-1} \sum_{\substack{i_1,\dots,i_\kappa=1 \\ i_1<\dots<i_\kappa}}^{N} \sum_{\substack{j_1,\dots,j_{\kappa+1}=1 \\ j_1<\dots<j_{\kappa+1}}}^{N} \widetilde{p}\begin{pmatrix} i_1, \dots, i_\kappa \\ j_1, \dots, j_{\kappa+1} \end{pmatrix} \Big] \tag{4.9}$$

where $p = 1 + \displaystyle\sum_{\kappa=1}^{N} \sum_{\substack{i_1,\dots,i_\kappa=1 \\ i_1<\dots<i_\kappa}}^{N} \sum_{\substack{j_1,\dots,j_\kappa=1 \\ j_1<\dots<j_\kappa}}^{N} \widetilde{p}\begin{pmatrix} i_1, \dots, i_\kappa \\ j_1, \dots, j_\kappa \end{pmatrix}$ *and we have abbreviated*

$$\widetilde{p}\begin{pmatrix} i_1, \dots, i_\kappa \\ j_1, \dots, j_\lambda \end{pmatrix} = \prod_{\mu=1}^{\kappa} \ell_{i_\mu} \prod_{\nu=1}^{\lambda} \overline{\ell}_{j_\nu} \prod_{\substack{\mu,\nu=1 \\ \mu<\nu}}^{\kappa} (\alpha_{i_\mu} - \alpha_{i_\nu})^2 \prod_{\substack{\mu,\nu=1 \\ \mu<\nu}}^{\lambda} (\overline{\alpha}_{j_\mu} - \overline{\alpha}_{j_\nu})^2 \Big/ \prod_{\mu=1}^{\kappa} \prod_{\nu=1}^{\lambda} (\alpha_{i_\mu} + \overline{\alpha}_{j_\nu})^2$$

with $\ell_j(x,t) = \ell_j^{(0)} \exp\big(\ \alpha_j x + f_0(\alpha_j) t\ \big)$, and arbitrary constants $\ell_j^{(0)} \in \mathbb{C}\backslash\{0\}$, $j = 1, \dots, N$.

Moreover, then solution q is regular on all of $\mathbb{R}^2$.

We assumed the α_j to be pairwise different in order to avoid cancellation phenomena. In fact it can be shown that, if an eigenvalue appears with higher multiplicity, there is nevertheless only one soliton with corresponding velocity.

Proof Proposition 4.3.9 follows from Proposition 4.3.1 in completely the same way as Theorem 4.1.4 from Theorem 4.1.1. Since this amounts to setting (i) $B = \overline{A}$ and (ii) $b = -\overline{a}$, $d = \overline{c}$ in the proof of Theorem 4.1.4 (compare the proof of Proposition 4.3.1), the resulting formulas simply can be carried over by replacing

a) m, n by N, b) f_j, g_j by $-\ell_j$, $\overline{\ell_j}$, and c) q by $-q$.

The reason for c) is the following: The sign in (ii) implies that there is no sign in front of the lower right block in (4.3). As a consequence there is an additional sign in the evaluation of the corresponding principal minor (4.4) if $m = n + 1$.

Regularity follows from Proposition 4.3.7. □

In particular, for $N = 1$ we recover the one-soliton

$$\begin{aligned} q(x,t) &= -\Big(1 + \ell(x,t)\overline{\ell(x,t)}/(\alpha + \overline{\alpha})^2\Big)^{-1} \overline{\ell(x,t)} \\ &= -\mathrm{Re}(\alpha)\, e^{-\mathrm{i}\,\mathrm{Im}\left(\Gamma(x,t)\right)} \cosh^{-1}\Big(\mathrm{Re}\big(\Gamma(x,t)\big)\Big), \end{aligned} \tag{4.10}$$

where $\Gamma(x,t) = \alpha x + f_0(\alpha) t + \varphi$ and $\exp(\varphi) = \ell^{(0)}/(\alpha + \overline{\alpha})$.

Asymptotic behaviour

First of all we have to define when two functions coincide asymptotically.

Definition 4.3.10. *Two functions $u = u(x,t)$, $v = v(x,t)$ are of the same asymptotic behaviour for $t \to \infty$ ($t \to -\infty$), if for every $\epsilon > 0$ there is a t_ϵ such that for $t > t_\epsilon$ ($t < t_\epsilon$) we have $|u(x,t) - v(x,t)| < \epsilon$ uniformly in x.*

In this case, we write $u(x,t) \approx v(x,t)$ for $t \approx \infty$ ($t \approx -\infty$).

The following result describes the characteristic behaviour of N-solitons. It is well-known for particular equations of the $\mathbb{C}$-reduced AKNS system (see for example [69] for the sine-Gordon and the Nonlinear Schrödinger equation). We will not give a proof here because the result is a special case of Theorem 5.1.2.

Theorem 4.3.11. *Let $\alpha_j \in \mathbb{C}$, $j = 1, \ldots, N$, are as assumed in Proposition 4.3.9, and choose $\ell_j^{(0)} \in \mathbb{C}\backslash\{0\}$, $j = 1, \ldots, N$, arbitrarily. Moreover, assume that the v_j defined by*

$$v_j = -\mathrm{Re}\big(f_0(\alpha_j)\big)/\mathrm{Re}(\alpha_j) \tag{4.11}$$

are pairwise different.

To these data we associate the single solitons

$$\begin{aligned} q_j^{\pm}(x,t) \;&=\; -\mathrm{Re}(\alpha_j)\, e^{-\mathrm{i}\,\mathrm{Im}\left(\Gamma_j^{\pm}(x,t)\right)} \cosh^{-1}\Big(\mathrm{Re}\big(\Gamma_j^{\pm}(x,t)\big)\Big) \\ &\qquad\qquad \text{with } \Gamma_j^{\pm}(x,t) = \alpha_j x + f_0(\alpha_j)t + \varphi_j + \varphi_j^{\pm}, \end{aligned} \tag{4.12}$$

where the quantities φ_j are given by $\exp(\varphi_j) = \ell_j^{(0)}/(\alpha_j + \overline{\alpha}_j)$ and $\varphi_j^{\pm}$ by the explicit formula (modulo $2\pi\mathrm{i}\mathbb{Z}$)

$$\exp\left(\varphi_j^{\pm}\right) = \prod_{k \in \Lambda_j^{\pm}} \left[\frac{\alpha_k - \alpha_j}{\overline{\alpha}_k + \alpha_j}\right]^2 \tag{4.13}$$

with the index sets $\Lambda_j^{\pm} = \{k = 1, \ldots, N \mid v_j \gtrless v_k\}$.

Then the asymptotic behaviour of the solution in Proposition 4.3.9 is described by

$$q(x,t) \;\approx\; \sum_{j=1}^{N} q_j^{\pm}(x,t) \qquad \text{for } t \approx \pm\infty. \tag{4.14}$$

The v_j are the velocities of the solitons $q_j^{\pm}$. Observe that the assumption that the v_j are pairwise different implies the same for the α_j. The quantities $\varphi_j^{\pm}$, which indicate a position shift in the asymptotic forms, are called phase-shifts.

Geometrically, the above result can be visualized as follows:

> *For large negative times, all solitons are well separated, each travelling with constant velocity. As time goes by, faster solitons will overtake slower ones. The resulting collisions do not change shape and velocity of the solitons, the only effect being a phase-shift. For large positive times the picture is the same as in the beginning except for the fact that the solitons travel in reversed order.*

Note that the assumption $\mathrm{Re}(\alpha_j) > 0$ assures regularity of the solitons. The assumption that the v_j are pairwise different means that the solitons move with different velocities. Groups of solitons moving with identical velocity exist and are called formations of solitons in [69]. It is tempting to exclude them in general asymptotic results because formations cannot be separated in asymptotic terms. But they can be treated as a whole, giving rise to more and more complicated 'atomic' building blocks. We will not formalize this, but we emphasize that formations, at least in examples of lower complexity, can be understood in asymptotic terms by our methods. As an example we give in Theorem 5.2.2 a complete result, including breathers, for the $\mathbb{R}$-reduced AKNS system.

An immediate consequence is the following conservation law.

Corollary 4.3.12. *The sum of all phase-shifts vanishes,* $\sum_{j=1}^{N}(\varphi_j^+ - \varphi_j^-) = 0 \ (\mathrm{mod}\, 2\pi \mathrm{i})$.

4.4 The $\mathbb{R}$-reduced AKNS system and reality questions

For the $\mathbb{R}$-reduced AKNS system we get more accessible solution formulas, where the structure of the determinant, always the main difficulty in applications, becomes considerably simpler. Then we discuss reality conditions.

Regarding the $\mathbb{R}$-reduced AKNS as a further reduction of the $\mathbb{C}$-reduced AKNS system is just an alternative approach to our topic, which enables us to carry over certain structural properties proved in the previous section.

The $\mathbb{R}$-reduced case presents a distinguished kind of formations of solitons, the so-called breathers. In preparation of the full negaton case we will focus on a solution class which combines solitons and breathers.

4.4.1 The improved solution formula and reality conditions

For the $\mathbb{R}$-reduced AKNS system we obtained in Theorem 2.5.1 a solution formula which looks more tractible than the form for the $\mathbb{C}$-reduced AKNS system. In the sequel we will look for appropriate conditions sufficient for the reality of the solutions. In the applications, this will amount to choosing the eigenvalues of the generating matrix A as a symmetric set with respect to the real axis.

The following proposition is a reformulation of Theorem 2.5.1 b) for matrices.

Proposition 4.4.1. *Let* $A \in \mathcal{M}_{n,n}(\mathbb{C})$ *satisfy* $\mathrm{spec}(A) \subseteq \{z \in \mathbb{C} \mid \mathrm{Re}(z) > 0$ *and* $f_0(z)$ *is finite*$\}$*, and let* $a,\ c \in \mathbb{C}^n$ *be non-zero. Define the operator-function*

$$L(x,t) = \exp\big(\,Ax + f_0(A)t\,\big)\,\Phi_{A,A}^{-1}(a \otimes c).$$

Then a solution of the $\mathbb{R}$*-reduced AKNS system (4.6) is given by*

$$q \;=\; \mathrm{i}\,\frac{\partial}{\partial x}\,\log\frac{\det\big(I + \mathrm{i}L\big)}{\det\big(I - \mathrm{i}L\big)}, \tag{4.15}$$

on strips $\mathbb{R} \times (t_1, t_2)$ *where both determinants do not vanish.*

For the $\mathbb{R}$-reduced AKNS system we are mainly interested in real-valued solutions. This can be achieved by imposing a certain relation between the data A, a, c and their complex conjugates. Recall that a rational function f_0 is called real if $f_0(x) \in \mathbb{R}$ for all $x \in \mathbb{R}$ where f_0 is finite.

Proposition 4.4.2. *Let the requirement of Proposition 4.4.1 be met and* f_0 *be real. Assume, in addition, that there is an invertible matrix* $\Pi \in \mathcal{M}_{n,n}(\mathbb{C})$ *such that*

$$\overline{A} = \Pi A \Pi^{-1}, \qquad \Pi'\overline{a} = a, \qquad \Pi^{-1}\overline{c} = c. \tag{4.16}$$

Then the solution q *in Proposition 4.4.1 is real-valued on the strips* $\mathbb{R} \times (t_1, t_2)$ *where it is defined.*

Proof By Lemma 4.4.4, reality of f_0 shows that $f_0 = \hat{f}/\hat{g}$ for two relatively prime polynomials $\hat{f}$, $\hat{g}$ with real coefficients. Since $\operatorname{spec}(A) \cap \{z \in \mathbb{C} \,|\, \hat{g}(z) = 0\} = \emptyset$, the holomorphic function calculus yields $(\hat{f}/\hat{g})(A) = \big(\hat{g}(A)\big)^{-1}\hat{f}(A)$ (see [82]). For the polynomial $\hat{g}$ (and also for $\hat{f}$) we have $\overline{\hat{g}(A)} = \hat{g}(\overline{A})$, and therefore $\overline{f_0(A)} = f_0(\overline{A})$.

As a consequence, $\overline{f_0(A)} = \Pi f_0(A)\Pi^{-1}$, and the definition of the exponential function $\widehat{L}(x,t) = \exp\big(Ax + f_0(A)t\big)$ as a power series yields

$$\begin{aligned}\overline{\widehat{L}(x,t)} &= \overline{\exp\big(Ax + f_0(A)t\big)} \;=\; \exp\big(\overline{A}x + f_0(\overline{A})t\big)\\ &= \Pi \exp\big(\; Ax + f_0(A)t \;\big)\,\Pi^{-1} \;=\; \Pi\widehat{L}(x,t)\Pi^{-1}.\end{aligned}$$

Secondly, from the unique solvability of the equation $AX + XA = a \otimes c$ with solution $C := \Phi_{A,A}^{-1}(a \otimes c)$ and (4.16) we infer from

$$\overline{a} \otimes \overline{c} \;=\; \overline{a \otimes c} = \overline{AC + CA} = \Pi\,\Big(\; A(\Pi^{-1}\overline{C}\Pi) + (\Pi^{-1}\overline{C}\Pi)A \;\Big)\,\Pi^{-1}$$

that $\Pi^{-1}\overline{C}\Pi = \Phi_{A,A}^{-1}\big(\Pi^{-1}(\overline{a}\otimes\overline{c})\Pi\big) = \Phi_{A,A}^{-1}\Big((\Pi'\overline{a}) \otimes (\Pi^{-1}\overline{c})\Big) = \Phi_{A,A}^{-1}(a \otimes c)$. In other words, $\overline{C} = \Pi C \Pi^{-1}$. This yields the following identity

$$\begin{aligned}\overline{\det\big(I - \mathrm{i}L\big)} &= \overline{\det\big(I - \mathrm{i}\widehat{L}C\big)} \;=\; \det\big(I + \mathrm{i}\overline{\widehat{L}C}\big) \;=\; \det\big(I + \mathrm{i}\,\Pi(\widehat{L}C)\Pi^{-1}\big)\\ &= \det\big(I + \mathrm{i}\widehat{L}C\big) \;=\; \det\big(I + \mathrm{i}L\big).\end{aligned}$$

To conclude,

$$\overline{q} = -\mathrm{i}\,\frac{\partial}{\partial x}\,\log\overline{\left(\frac{\det\big(I + \mathrm{i}L\big)}{\det\big(I - \mathrm{i}L\big)}\right)} = -\mathrm{i}\,\frac{\partial}{\partial x}\,\log\frac{\det\big(I - \mathrm{i}L\big)}{\det\big(I + \mathrm{i}L\big)} = \mathrm{i}\,\frac{\partial}{\partial x}\,\log\frac{\det\big(I + \mathrm{i}L\big)}{\det\big(I - \mathrm{i}L\big)} = q,$$

and q is real. □

Remark 4.4.3. *The above proof shows also that, under the assumption of Proposition 4.4.2, the solution q in Proposition 4.4.1 can be rewritten as*

$$q(x,t) = -\mathrm{i}\frac{\partial}{\partial x}\log\frac{\overline{p(x,t)}}{p(x,t)} \qquad \textit{with } p(x,t) = \det\big(I + \mathrm{i}L(x,t)\big).$$

Lemma 4.4.4. *Let f be a rational function with $f(x) \in \mathbb{R}$ for all $x \in \mathbb{R}$ where f is finite. Then f is the ratio of two relatively prime polynomials with real coefficients.*

Proof Consider the unique reduced factorization

$$f(z) \;=\; c\,\frac{\prod_{j=1}^{m}(z - a_j)}{\prod_{k=1}^{n}(z - b_k)}. \tag{4.17}$$

Then, for all $x \in \mathbb{R}$ where f is finite,

$$c\prod_{j=1}^{m}(x - a_j)\prod_{k=1}^{n}(x - \overline{b}_k) \;=\; c\,\frac{\prod_{j=1}^{m}(x - a_j)\prod_{k=1}^{n}|x - b_k|^2}{\prod_{k=1}^{n}(x - b_k)} \in \mathbb{R} \tag{4.18}$$

The polynomial $g(z) = c\prod_{j=1}^{m}(z - a_j)\prod_{k=1}^{n}(z - \overline{b}_k)$ coincides with (4.18) for $z \in \mathbb{R}$. Expand $g|_{\mathbb{R}}$ in its Taylor series,

$$g(x) \;=\; \sum_{j=0}^{n+m} c_j x^j, \qquad c_j \in \mathbb{R}.$$

This implies $g(z) = \sum_{j=0}^{n+m} c_j z^j$ and thus $g \in \mathbb{R}[x]$. As a consequence, the zeros of g are either real or come in complex conjugated pairs. Since (4.17) is reduced, no b_k can be conjugated to some a_j. Therefore, $\prod_{j=1}^{m}(z - a_j)$, $\prod_{k=1}^{n}(z - b_k) \in \mathbb{R}[z]$ and $c \in \mathbb{R}$. □

4.4.2 An alternative deduction of solutions from the $\mathbb{C}$-reduced AKNS by symmetry constraints

Subsequently we derive the solution formula for the $\mathbb{R}$-reduced AKNS system in the following manner: We start from the solution formula for the $\mathbb{C}$-reduced AKNS system in Proposition 4.3.1 and impose the condition (4.16) on the involved data. The main advantage of this link between $\mathbb{R}$-reduction and $\mathbb{C}$-reduction of the AKNS system is that it becomes possible to derive some properties of the $\mathbb{R}$-reduction, for example regularity, as immediate corollaries.

Proposition 4.4.5. *Let $A \in \mathcal{M}_{n,n}(\mathbb{C})$ satisfy $\operatorname{spec}(A) \subseteq \{z \in \mathbb{C} \mid \operatorname{Re}(z) > 0$ and $f_0(z)$ is finite$\}$, and let $a,\ c \in \mathbb{C}^n \backslash \{0\}$. Assume that f_0 is real, and that there is an invertible matrix $\Pi \in \mathcal{M}_{n,n}(\mathbb{C})$ such that (4.16) holds.*

Then any solution in Proposition 4.4.2 belongs to the solutions constructed in Proposition 4.3.1.

Proof First observe that, for f_0 real, the conditions $\overline{f_0(z)} = -f_0(-\overline{z})$ and $f_0(z) = -f_0(-z)$ for the $\mathbb{C}$-reduced and the $\mathbb{R}$-reduced AKNS systems, respectively, are equivalent.

Consequently, $\overline{\widehat{L}(x,t)} = \Pi \widehat{L}(x,t)\Pi^{-1}$ for $\widehat{L}(x,t) = \exp\big(Ax + f_0(A)t\big)$, because

$$\begin{aligned} \overline{\widehat{L}(x,t)} &= \overline{\exp\big(Ax + f_0(A)t\big)} \;=\; \exp\big(\overline{A}x - f_0(-\overline{A})t\big) \\ &= \Pi \exp\big(\ Ax - f_0(-A)t\ \big)\,\Pi^{-1} \;=\; \Pi \widehat{L}(x,t)\Pi^{-1}. \end{aligned}$$

Next, set $C = \Phi^{-1}_{A,A}(a \otimes c)$, and observe the identities $\Phi^{-1}_{A,\overline{A}}(\overline{a}\otimes c) = C\Pi^{-1}$, $\Phi^{-1}_{\overline{A},A}(a\otimes\overline{c}) = \Pi C$, which follow from

$$\begin{aligned} \overline{a} \otimes c &= \big((\Pi')^{-1}a\big) \otimes c = (a \otimes c)\Pi^{-1} = (AC + CA)\Pi^{-1} = A(C\Pi^{-1}) + (C\Pi^{-1})\overline{A}, \\ a \otimes \overline{c} &= a \otimes (\Pi c) = \Pi(a \otimes c) = \Pi(AC + CA) = \overline{A}(\Pi C) + (\Pi C)A, \end{aligned}$$

and $\overline{a} \otimes \overline{c} = \Pi(a \otimes c)\Pi^{-1}$.

Thus we get

$$\widehat{L}\Phi^{-1}_{A,\overline{A}}(\overline{a} \otimes c) = (\widehat{L}C)\Pi^{-1}, \quad \overline{\widehat{L}}\Phi^{-1}_{\overline{A},A}(a \otimes \overline{c}) = \Pi(\widehat{L}C), \quad \overline{\widehat{L}}(\overline{a} \otimes \overline{c}) = \Pi\big(\widehat{L}(a \otimes c)\big)\Pi^{-1}.$$

As a consequence, the solution q in Proposition 4.3.1 reads

$$\begin{aligned} q &= 1 - P/p, \text{ where} \\ & P = \det\left(I + \begin{pmatrix} 0 & -(\widehat{L}C)\Pi^{-1} \\ \Pi(\widehat{L}C) & \Pi(\widehat{L}(a \otimes c))\Pi^{-1} \end{pmatrix}\right), \\ & p = \det\left(I + \begin{pmatrix} 0 & -(\widehat{L}C)\Pi^{-1} \\ \Pi(\widehat{L}C) & 0 \end{pmatrix}\right). \end{aligned}$$

The following manipulation shows that the matrix Π cancels in the above expressions,

$$\begin{aligned} P &= \det\left(I + \begin{pmatrix} I & 0 \\ 0 & \Pi \end{pmatrix}\begin{pmatrix} 0 & -\widehat{L}C \\ \widehat{L}C & \widehat{L}(a \otimes c) \end{pmatrix}\begin{pmatrix} I & 0 \\ 0 & \Pi^{-1} \end{pmatrix}\right) \\ &= \det\left(I + \begin{pmatrix} 0 & -\widehat{L}C \\ \widehat{L}C & \widehat{L}(a \otimes c) \end{pmatrix}\right). \end{aligned}$$

Hence the coincidence of the solution formula for q in Proposition 4.3.1 with (4.15) can be verified by the subsequent arguments.

$$q = 1 - \frac{\det\left(I + \begin{pmatrix} 0 & -\widehat{L}C \\ \widehat{L}C & \widehat{L}(a \otimes c) \end{pmatrix}\right)}{\det\left(I + \begin{pmatrix} 0 & -\widehat{L}C \\ \widehat{L}C & 0 \end{pmatrix}\right)}$$

$$
\begin{aligned}
&= 1-\det\Big(I+\Big(I+\begin{pmatrix}0 & -\widehat{L}C\\ \widehat{L}C & 0\end{pmatrix}\Big)^{-1}\begin{pmatrix}0 & 0\\ 0 & \widehat{L}(a\otimes c)\end{pmatrix}\Big)\\
&= -\mathrm{tr}\Big(\Big(I+\begin{pmatrix}0 & -\widehat{L}C\\ \widehat{L}C & 0\end{pmatrix}\Big)^{-1}\begin{pmatrix}0 & 0\\ 0 & \widehat{L}(a\otimes c)\end{pmatrix}\Big),
\end{aligned}
$$

because $\det(I+b\otimes d)=1+\langle d,b\rangle = 1+\mathrm{tr}(b\otimes d)$ for one-dimensional matrices $b\otimes d$. Now we invert explicitly

$$
\Big(I+\begin{pmatrix}0 & -\widehat{L}C\\ \widehat{L}C & 0\end{pmatrix}\Big)^{-1} = \begin{pmatrix}\big(I+(\widehat{L}C)^2\big)^{-1} & 0\\ 0 & \big(I+(\widehat{L}C)^2\big)^{-1}\end{pmatrix}\begin{pmatrix}I & \widehat{L}C\\ -\widehat{L}C & I\end{pmatrix}
$$

and multiply the appearing matrices with each other. For the trace, this yields

$$
\begin{aligned}
q &= -\mathrm{tr}\Big(\big(I+(\widehat{L}C)^2\big)^{-1}\widehat{L}(a\otimes c)\Big)\\
&= -\mathrm{tr}\Big(\big(I+(\widehat{L}C)^2\big)^{-1}\widehat{L}(AC+CA)\Big)\\
&= -2\,\mathrm{tr}\Big(\big(I+(\widehat{L}C)^2\big)^{-1}A(\widehat{L}C)\Big) \qquad (4.19)\\
&= -\mathrm{tr}\Big(\Big[(I+\mathrm{i}L)^{-1}+(I-\mathrm{i}L)^{-1}\Big]L_x\Big)\\
&= \mathrm{i}\,\mathrm{tr}\Big((I+\mathrm{i}L)^{-1}(\mathrm{i}L)_x\Big) - \mathrm{i}\,\mathrm{tr}\Big((I-\mathrm{i}L)^{-1}(-\mathrm{i}L)_x\Big)\\
&= \mathrm{i}\frac{\partial}{\partial x}\log\frac{\det(I+\mathrm{i}L)}{\det(I-\mathrm{i}L)},
\end{aligned}
$$

where we used the trace property in (4.19), $L=\widehat{L}C$, and the derivation rule for determinants (see Proposition B.2.12) for the last equality. □

Exploiting the link between $\mathbb{C}$- and $\mathbb{R}$-reduction in Proposition 4.4.5, the following statements are immediate corollaries of Lemma 4.3.2, Poposition 4.3.7.

Lemma 4.4.6. *Let $A\in\mathcal{M}_{n,n}(\mathbb{C})$ be as in Proposition 4.4.1 and U a matrix transforming A into Jordan form J_A, namely $A=U^{-1}J_AU$. Suppose that the assumptions of Proposition 4.4.2 are met.*

Then the solution is not altered upon replacing simultaneously A by J_A and the vectors a, c by $(U^{-1})'a$, Uc.

Proposition 4.4.7. *The solutions in Proposition 4.4.2 are regular on all of $\mathbb{R}^2$.*

Remark 4.4.8. *By Lemma 4.4.6, we can always assume that A is in Jordan form. For a Jordan block A_j and a vector $v\in\mathbb{C}^n$, we call the entries of v appearing in the same lines as A_j the part of v corresponding to A_j. There is a natural choice to fulfill (4.16).*

1. *For the real eigenvalues the parts of a, c corresponding to the associated Jordan blocks have real entries.*

2. *Nonreal eigenvalues always appear in complex conjugate pairs α_j, $\alpha_{\overline{j}}=\overline{\alpha}_j$. There are unique associated Jordan blocks A_j, $A_{\overline{j}}$ of equal size, and the entries of a, c corresponing to A_j, $A_{\overline{j}}$ are complex conjugate, respectively.*

4.4.3 Solitons, bound states of solitons, and their particle behaviour

Remarkably simpler soliton formulas can be obtained for the $\mathbb{R}$-reduced AKNS system. For the sine-Gordon equation, a comparable result was already obtained in [90].

Proposition 4.4.9. *For pairwise different complex numbers α_j, $j = 1, \ldots N$, stisfying $\mathrm{Re}(\alpha_j) > 0$ and contained in the domain where f_0 is holomorphic, a solution of the $\mathbb{R}$-reduced AKNS system (4.6) is given by*

$$q = -\mathrm{i}\frac{\partial}{\partial x}\log\frac{\overline{p}}{p},$$
$$\textit{where} \qquad p = 1 + \sum_{\kappa=1}^{N} \mathrm{i}^{\kappa} \sum_{\substack{j_1,\ldots,j_\kappa=1 \\ j_1<\cdots<j_\kappa}}^{N} \prod_{\mu=1}^{\kappa} \ell_{j_\mu} \prod_{\substack{\mu,\nu=1 \\ \mu<\nu}}^{\kappa} \left[\frac{\alpha_{j_\mu}-\alpha_{j_\nu}}{\alpha_{j_\mu}+\alpha_{j_\nu}}\right]^2 \tag{4.20}$$

with $\ell_j(x,t) = \ell_j^{(0)} \exp\big(\,\alpha_j x + f_0(\alpha_j)t\,\big)$ and arbitrary constants $\ell_j^{(0)} \in \mathbb{C}\backslash\{0\}$, $j = 1, \ldots, N$.

Assume that for each j, α_j and $\ell_j^{(0)}$ are either real or there is a unique index $\overline{\jmath}$ with $\alpha_{\overline{\jmath}} = \overline{\alpha}_j$ and $\ell_{\overline{\jmath}}^{(0)} = \overline{\ell}_j^{(0)}$. Then q is real-valued and regular on all of $\mathbb{R}^2$.

Proof Let us define the generating matrix $A \in \mathcal{M}_{N,N}(\mathbb{C})$ in the solution formula of Proposition 4.4.1 as $A = \mathrm{diag}\,\{\,\alpha_j \,|\, j = 1, \ldots, N\,\}$. Since $\mathrm{Re}(\alpha_j) > 0\ \forall j$ by assumption, we in particular we have $0 \notin \mathrm{spec}(A) + \mathrm{spec}(A)$.

Thus we are in position to apply Proposition 4.4.1 and Remark 4.4.3 and obtain a solution q of the $\mathbb{C}$-reduced AKNS system by

$$q = -\mathrm{i}\,\frac{\partial}{\partial x}\log\frac{\overline{p}}{p}, \qquad \text{where} \qquad p_\pm = \det\big(I + \mathrm{i}\widehat{L}\Phi_{A,A}^{-1}(a\otimes c)\big),$$

where $\widehat{L}(x,t) = \exp\big(Ax + f_0(A)t\big)$.

One directly verifies

(i) $\quad \widehat{L} = \mathrm{diag}\,\{\,\widehat{\ell}_j \,|\, j = 1, \ldots, N\,\}$ with $\widehat{\ell}_j(x,t) = \big(\alpha_j x + f_0(\alpha_j)t\big)$,

(ii) $\quad \Phi_{A,A}^{-1}(a\otimes c) = \left(\dfrac{a_{j'}c_j}{\alpha_j+\alpha_{j'}}\right)_{j,j'=1}^{N}.$

Inserting this data into the solution formula, and using the expansion rule for determinants (confer Lemma 4.1.6), we derive

$$\begin{aligned}
p &= \det\left(\,\delta_{jj'}\ \mathrm{i}\frac{a_{j'}c_j}{\alpha_j+\alpha_{j'}}\widehat{\ell}_j\,\right)_{j,j'=1}^{N} \\
&= 1 + \sum_{\kappa=1}^{N}\mathrm{i}^{\kappa}\sum_{\substack{j_1,\ldots,j_\kappa=1 \\ j_1<\cdots<j_\kappa}}^{N} \det\left(\frac{a_{j_\nu}c_{j_\mu}}{\alpha_{j_\mu}+\alpha_{j_\nu}}\widehat{\ell}_{j_\mu}\right)_{\mu,\nu=1}^{\kappa} \\
&= 1 + \sum_{\kappa=1}^{N}\mathrm{i}^{\kappa}\sum_{\substack{j_1,\ldots,j_\kappa=1 \\ j_1<\cdots<j_\kappa}}^{N} \prod_{\mu=1}^{\kappa} a_{j_\mu}c_{j_\mu}\widehat{\ell}_{j_\mu} \quad \det\left(\frac{1}{\alpha_{j_\mu}+\alpha_{j_\nu}}\right)_{\mu,\nu=1}^{\kappa} \\
&= 1 + \sum_{\kappa=1}^{N}\mathrm{i}^{\kappa}\sum_{\substack{j_1,\ldots,j_\kappa=1 \\ j_1<\cdots<j_\kappa}}^{N} \prod_{\mu=1}^{\kappa} \frac{a_{j_\mu}c_{j_\mu}}{2\alpha_{j_\mu}}\widehat{\ell}_{j_\mu} \quad \prod_{\substack{\mu,\nu=1 \\ \mu<\nu}}^{\kappa}\left[\frac{\alpha_{j_\mu}-\alpha_{j_\nu}}{\alpha_{j_\mu}+\alpha_{j_\nu}}\right]^2,
\end{aligned}$$

where the last equality relies on Lemma 6.1.2. Reality of the solutions follows from Proposition 4.4.2, regularity from Proposition 4.4.7. □

In particular, for $N=1$ and $\alpha=\alpha_1$, $\ell^{(0)}=\ell_1^{(0)}$ real, we get

$$\begin{aligned} q(x,t) &= -\mathrm{i}\,\frac{\partial}{\partial x}\log\frac{\overline{p(x,t)}}{p(x,t)}, && \text{where } p(x,t)=1+\mathrm{i}\,\ell(x,t), \\ &= -\alpha\epsilon\,\cosh^{-1}\big(\Gamma(x,t)\big) \end{aligned}$$

with the real-valued function $\Gamma(x,t)=\alpha x+f_0(\alpha)t+\varphi$ and $\ell^{(0)}=\epsilon\exp(\varphi)$ with $\varphi\in\mathbb{R}$, $\epsilon=\pm1$. This solution is called soliton for $\operatorname{sgn}(\alpha\epsilon)<0$ and antisoliton for $\operatorname{sgn}(\alpha\epsilon)>0$.

The next interesting special case is $N=2$, and $\alpha=\alpha_1=\overline{\alpha}_2$, $\ell^{(0)}=\ell_1^{(0)}=\overline{\ell}_2^{(0)}$. In this case we obtain the so-called breather, or pulsating soliton,

$$q(x,t)=-\mathrm{i}\,\frac{\partial}{\partial x}\log\frac{\overline{p(x,t)}}{p(x,t)},$$

where

$$\begin{aligned} p(x,t) &= 1+\mathrm{i}\,\big(\ell(x,t)+\overline{\ell(x,t)}\big)-\left(\frac{\alpha-\overline{\alpha}}{\alpha+\overline{\alpha}}\right)^2\ell(x,t)\overline{\ell(x,t)} \\ &= 1+\mathrm{i}\gamma^{-1}\Big(\exp\big(\Gamma(x,t)\big)+\exp\big(\overline{\Gamma(x,t)}\big)\Big)+\exp\Big(\Gamma(x,t)+\overline{\Gamma(x,t)}\Big) \end{aligned}$$

for $\Gamma(x,t)=\alpha x+f_0(\alpha)t+\varphi+\log\gamma$ with $\gamma=\big|\mathrm{Im}(\alpha)/\mathrm{Re}(\alpha)\big|$, and $\exp(\varphi)=\ell^{(0)}$.

Then

$$\begin{aligned} q(x,t) &= -\mathrm{i}\,\frac{\partial}{\partial x}\log\frac{1-\mathrm{i}f(x,t)}{1+\mathrm{i}f(x,t)} \;=\; -2\big(1+f(x,t)^2\big)^{-1}\frac{\partial}{\partial x}f(x,t), \\ &\qquad \text{where } f(x,t)=\gamma^{-1}\cos\Big(\mathrm{Im}\big(\Gamma(x,t)\big)\Big)\cosh^{-1}\Big(\mathrm{Re}\big(\Gamma(x,t)\big)\Big). \end{aligned}$$

This solution represents a bound state of two solitons, see [69].

The solutions in Proposition 4.4.9 allow a similar asymptotic analysis as those in Proposition 4.3.9. Qualitatively the result is the following:

Any solution given in Proposition 4.4.9 asymptotically (for t large in modulus) behaves as a sum of n_1 solitons and n_2 breathers, where $n_1=|\{j\,|\,\mathrm{Im}(\alpha_j)=0\}|$, $n_2=|\{j\,|\,\mathrm{Im}(\alpha_j)>0\}|$.

The interesting point is that breathers, which should be viewed as formations of two solitons moving with identical speeds, can be treated as atomic building blocks. We will not give the precise formulas here. They can be extracted from Theorem 5.2.2 upon setting $n_j=1$.

Chapter 5

Asymptotic analysis of negatons

Negatons are solutions where solitons and antisolitons appear in weakly bounded states. In terms of the inverse scattering method these correspond to multiple poles of the reflection coefficient. In the literature these solutions (often called multiple pole solutions) were discovered in the eighties and discussed in several papers ([9], [34], [49], [71], [75], [96], [98], [99]). The interest in a complete asymptotic description was aroused by Matveev ([58], [59], see also [60] and references therein), who treated the related class of positons to a certain extent and formulated expectations for the general case. In this spirit particular cases of negatons were examined in [79]. In [88], [90], [91], the author gave a complete and rigorous description of the negatons of the Korteweg-de Vries equation, the sine-Gordon equation, and the Toda lattice, the sine-Gordon equation being particularly interesting because its negatons are smooth.

In the present chapter we give a reasonable complete analysis of the negatons of the $\mathbb{C}$-reduced AKNS system. In our formalism, negatons are obtained by inserting admissible square matrices into the soution formulas of Proposition 4.3.1. The essence of the main result in Theorem 5.1.2 is that for an admissible matrix the dynamical properties of the ensuing solution can be read off from the Jordan normal form. To every Jordan block there corresponds a single negaton, which is a group of weakly bound solitons, their number equals the size of the block. Negatons are weakly localized meaning that they shrink and extend, but with sublinear velocity. The striking feature is now that negatons as a whole interact with a phase-shift, in the same way as solitons. In particular we obtain Theorem 4.3.11 if all blocks are of size 1×1.

It is also possible to analyse the interior structure of a single negaton. Its member solitons travel on logarithmic trajectories in spacetime organized around a geometric center moving itself with constant velocity. There are internal collisions which do not effect the path of the geometric center. Moreover internal and external collisions are coherent in so far as the collision of two negatons can also be understood as the sum of all pairwise collisions of respective member solitons. We stress that an isolated member soliton (if the block is larger than 1×1) cannot be traced back to a solution of the equation (like the members of an N-soliton).

The preceding qualitative description can be rigourosly reformulated in terms of asymptotic analysis. Actually the proof follows roughly the order of our explanations. Whereas we show asymptotic coincidence only for very negative/positive times, the computer graphics in Section 5.3 actually show that the convergence is very fast.

As expected the asymptotics of negatons simplifies in the $\mathbb{R}$-reduced case. This allows us to go one step further and include also negatons of breathers. Breathers are the most important example of so-called formations, i.e., groups of solitons moving with identical speed. The purpose of Section 5.2 is to show that, at least in cases of reasonable complexity, formations are accessible by our methods.

It is illustrative to explain the difference of our results to the work of Schuur [93]. He considered those solutions of the reduced AKNS systems which can be obtained from rapidly decreasing potentials by the inverse scattering method, and shows that, asymptotically, only the contribution corresponding to the discrete part of the spectrum survives. Note that, for the reduced AKNS systems, negatons belong to this reflectionless part. But he does not determine the asymptotics of solutions with discrete spectrum.

5.1 Negatons of the $\mathbb{C}$-reduction

5.1.1 Statement of the main result

In Section 4.3.1 we have introduced the notion of negatons. For the $\mathbb{C}$-reduced AKNS-system, an N-negaton is a solution as given in Proposition 4.3.1 with the generating matrix A chosen according to the following assumption.

Assumption 5.1.1. *Let $A \in \mathcal{M}_{n,n}(\mathbb{C})$ be given in Jordan form with N Jordan blocks A_j of dimension n_j and with eigenvalues α_j, i.e.,*

$$A = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & A_N \end{pmatrix}, \qquad A_j = \begin{pmatrix} \alpha_j & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \alpha_j \end{pmatrix} \in \mathcal{M}_{n_j,n_j}(\mathbb{C}).$$

Assume $\operatorname{spec}(A) \subseteq \{z \in \mathbb{C} \mid \operatorname{Re}(z) > 0$ *and* $f_0(z)$ *finite*$\}$.

It is useful to adapt our notation to the given Jordan structure of A. For a vector $v \in \mathbb{C}^n$, its decompositions according to the Jordan structure of A reads

$$v = (v_j)_{j=1}^N \qquad \text{with} \qquad v_j = (v_j^{(\mu)})_{\mu=1}^{n_j} \in \mathbb{C}^{n_j},$$

and, analogously, for a matrix $T \in \mathcal{M}_{n,n}(\mathbb{C})$, we write

$$T = (T_{ij})_{ij=1}^N \qquad \text{with} \qquad T_{ij} = (t_{ij}^{(\nu\mu)})_{\substack{\nu=1,\dots,n_i \\ \mu=1,\dots,n_j}} \in \mathcal{M}_{n_i,n_j}(\mathbb{C}).$$

Now we are in position to state the main result of this chapter.

Theorem 5.1.2. *Let Assumption 5.1.1 be fulfilled. Assume that a, $c \in \mathbb{C}^n$ (when decomposed according to the Jordan form of A) satisfy $a_j^{(1)} c_j^{(n_j)} \neq 0$ for $j = 1, \dots, N$. Define*

$$v_j = -\operatorname{Re}\big(\, f_0(\alpha_j)\,\big) / \operatorname{Re}(\alpha_j). \tag{5.1}$$

Assume, in addition,

(i) v_j *are pairwise different,*

(ii) $v_j + f_0'(\alpha_j) \neq 0 \ \forall j$.

To these data we associate for $j' = 0, \dots, n_j - 1$, the single solitons

$$\begin{aligned} q_{jj'}^{\pm}(x,t) &= (-1)^{j'+1} \operatorname{Re}(\alpha_j) e^{-\mathrm{i}\operatorname{Im}\left(\Gamma_{jj'}^{\pm}(x,t)\right)} \cosh^{-1}\Big(\operatorname{Re}\big(\Gamma_{jj'}^{\pm}(x,t)\big)\Big) \\ &\quad \textit{with } \Gamma_{jj'}^{\pm}(x,t) = \alpha_j x + f_0(\alpha_j) t \mp J' \log|t| + \varphi_j + \varphi_j^{\pm} + \varphi_{jj'}^{\pm}, \end{aligned} \tag{5.2}$$

where we have set $J' = -(n_j - 1) + 2j'$.

Modulo $2\pi\mathrm{i}\mathbb{Z}$, the quantities φ_j are determined by $\exp(\varphi_j) = a_j^{(1)} c_j^{(n_j)} / (\alpha_j + \overline{\alpha}_j)^{n_j}$ *and $\varphi_j^{\pm}$, $\varphi_{jj'}^{\pm}$ by the explicit formulas*

$$\exp(\varphi_j^{\pm}) = \prod_{k \in \Lambda_j^{\pm}} \left[\frac{\alpha_j - \alpha_k}{\alpha_j + \overline{\alpha}_k}\right]^{2n_k}, \tag{5.3}$$

$$\exp\left(\pm \varphi^{\pm}_{jj'} \right) \;=\; \frac{j'!}{(j'-J')!} d_j^{-J'}, \tag{5.4}$$

with the index sets $\Lambda_j^{\pm} = \{k \mid v_k \lessgtr v_j\}$, *and* $d_j = (\alpha_j + \overline{\alpha}_j)\big(v_j + f_0'(\alpha_j)\big)$.

Then the asymptotic behaviour of the solution in Proposition 4.3.1 is described by

$$q(x,t) \;\approx\; \sum_{j=1}^{N} \sum_{j'=0}^{n_j-1} q^{\pm}_{jj'}(x,t) \qquad \textit{for } t \approx \pm\infty \tag{5.5}$$

Hence, the solitons $q^{\pm}_{jj'}$ (or more precisely their maxima) travel asymptotically along the logarithmic curves

$$\begin{aligned} &\Big\{(x,t)\in\mathbb{R}^2 \,\Big|\, \mathrm{Re}\big(\Gamma^{\pm}_{jj'}(x,t)\big) = 0\Big\} \\ &= \;\Big\{(x,t)\in\mathbb{R}^2 \,\Big|\, \mathrm{Re}(\alpha_j)\big[x - v_j t\big] \mp J' \log|t| \; + \mathrm{Re}\,\left(\, \varphi_j + \varphi_j^{\pm} + \varphi^{\pm}_{jj'} \,\right) = 0\Big\}, \end{aligned}$$

and the term

$$\mathrm{Im}\,\left(\, \Gamma^{\pm}_{jj'}(x,t) \,\right) = \;\mathrm{Im}\,\Big(\, \alpha_j x + f_0(\alpha_j)t + \varphi_j + \varphi_j^{\pm} + \varphi^{\pm}_{jj'} \,\Big)$$

encodes the oscillation of the soliton. In particular, v_j is the velocity of $q^{\pm}_{jj'}$ up to a logarithmic error, and φ_j, the initial phase, determines the position of $q^{\pm}_{jj'}$ at a prescribed time. It should be stressed that φ_j is a parameter of the solution, but not a shift.

Furthermore, the quantities $\varphi_j^{\pm}$, $\varphi^{\pm}_{jj'}$ indicate position shifts in the asymptotic form. The former are due to external collisions of negatons with different velocities. Note that the index sets Λ_j^{+} (resp. Λ_j^{-}) stand for those negatons which move slower (respectively faster) than the j-th negaton. The latter come from internal collisions between the solitons belonging to the same negaton.

Both, $\varphi_j^{\pm}$ and $\varphi^{\pm}_{jj'}$, are called (external and internal, resp.) phase-shifts.

Qualitatively, our main result can be visualized as follows:

Interpretation a) *First consider a single eigenvalue* α *of multiplicity* n. *Then the solution is a wave packet consisting of* n *solitons. We call such a solution a (single) negaton of order* n. *The main observation is that the geometric center of the wave packet propagates with constant velocity* $v = -\mathrm{Re}\big(f_0(\alpha)\big)/\mathrm{Re}(\alpha)$, *but its members drift away from each other at most logarithmically.*

Hence, for large negative times we can imagine each soliton to be located on one side of the center, approaching the center logarithmically. At some moment it changes sides, and for large positive times it is located on the other side of the center, moving away from the center again logarithmically. Hence the solitons appear in reversed order for $\pm\infty$. *Finally we stress that the path of the geometric center is not affected by the internal collisions.*

b) In the general case of N *eigenvalues* $\alpha_1, \dots, \alpha_N$ *of algebraic multiplicities* $n_1, \dots, n_N$, *the solution is a superposition of* N *wave packets as in a). Their behaviour under collision is a natural generalization of what is known for* N*-solitons. In particular, every wave packet as a whole suffers a phase-shift.*

Note that not only the curves on which the solitons move but also their oscillations experience phase-shifts. However, the logarithmic deviation only has an effect on the path of the solitons.

Now we can justify the assumptions. First note that Assumption 5.1.1 guarantees regularity of the N-negaton as shown in Section 4.3.2. Assumption (i) means that the N single negatons all move with different velocities and thus can be distinguished by the asymptotic analysis (confer Section 4.3.3). in particular, it also shows that the α_j are pairwise different, which avoid cancellation phenomena. Finally, (ii) is supposed to avoid degenerations.

The following conservation law for the total complex phase-shifts holds modulo $2\pi\mathrm{i}\mathbb{Z}$. Of course there is no amibiguity for its real part responsible for translation in space.

Corollary 5.1.3. *The sum of all phase-shifts vanishes:*

$$\sum_{j=1}^{N}\sum_{j'=0}^{n_j-1}\Big(\,[\varphi_j^+ + \varphi_{jj'}^+] - [\varphi_j^- + \varphi_{jj'}^-]\,\Big) = 0 \qquad (\text{mod } 2\pi\mathrm{i}).$$

At first glance the assumption on the vectors a, c (namely $a_j^{(1)} c_j^{(n_j)} \neq 0\ \forall j$) looks artificial. However, the following lemma shows that it can indeed be supposed without loosing anything.

Lemma 5.1.4. *Let Assumption 5.1.1 be met for $A \in \mathcal{M}_{n,n}(\mathbb{C})$, and choose a, $c \in \mathbb{C}^n$ with $a \otimes c \neq 0$. If $c_k^{(n_k)} = 0$, then the solution given in Proposition 4.3.1 is not altered if we replace simultaneously A by $\widetilde{A}$, and a, c by $\widetilde{a}$, $\widetilde{c}$, where*

(i) $\widetilde{A}$ is the matrix obtained from A by reducing the block A_k in dimension by one, i.e., by omitting its last row and column,

(ii) $\widetilde{a}$, $\widetilde{c}$ are the vectors obtained from a, c by deleting the entries $a_k^{(n_k)}$, $c_k^{(n_k)}$.

The case $a_k^{(1)} = 0$ is analogous. Then we have to delete the entries $a_k^{(1)}$, $c_k^{(1)}$.

Proof By Proposition 4.3.1, the solution q is given by $q = 1 - P/p$ with

$$P = \det\Big(I + \begin{pmatrix} 0 & -L \\ \overline{L} & \overline{L}_0 \end{pmatrix}\Big), \qquad p = \det\Big(I + \begin{pmatrix} 0 & -L \\ \overline{L} & 0 \end{pmatrix}\Big), \tag{5.6}$$

where $L = \widehat{L}\Phi^{-1}_{A,\overline{A}}(\overline{a} \otimes c)$, $L_0 = \widehat{L}(a \otimes c)$, and $\widehat{L} = \widehat{L}(x,t) = \exp\big(Ax + f_0(A)t\big)$. From the fact that $\widehat{L}$ commutes with A, we deduce

$$L = \Phi^{-1}_{A,\overline{A}}(\overline{a} \otimes d), \quad L_0 = a \otimes d, \qquad \text{for } d = d(x,t) = \widehat{L}(x,t)c.$$

Obviously, $\widehat{L} = \operatorname{diag}\{\widehat{L}_j \,|\, j = 1, \dots, N\}$ with $\widehat{L}_j(x,t) = \exp\big(A_j x + f_0(A_j)t\big)$. Thus $d = (d_j)_j$, where $d_j(x,t) = \widehat{L}_j(x,t)c_j$. Since $c_k^{(n_k)} = 0$ and $\widehat{L}_k$ is an upper triangular matrix, $d_k^{(n_k)} = 0$. Moreover,

$$\widehat{L}_k c_k \;= \big(\,\widetilde{L}_k \widetilde{c}_k,\; 0\,\big)', \qquad \text{where } \widetilde{L}_k(x,t) = \exp\big(\widetilde{A}_k x + f_0(\widetilde{A}_k)t\big),$$

and $\widetilde{A}_k$ the Jordanblock with respect to the eigenvalue α_k but of dimension $n_k - 1$.

The very definition of one-dimensional matrices implies that $a \otimes d = \big(a_j \otimes d_i\big)_{i,j=1}^{n}$ has a zero-row, namely the n_k-th row of all blocks $a_j \otimes d_k$, $j = 1, \dots, N$. Moreover, the same holds true for $\Phi^{-1}_{A,\overline{A}}(\overline{a} \otimes d)$, $\Phi^{-1}_{\overline{A},A}(a \otimes \overline{d})$ by Proposition 4.1.3. Therefore the assertion follows from expanding the determinants in (5.6). $\square$

5.1.2 Proof of Theorem 5.1.2

The proof is divided into two parts. The first step contains some necessary technical reductions and the second step is devoted to the asymptotic analysis. It will rely on the evaluation of complicated determinants which is postponed to Chapter 6.

Part 1. Technical reductions

First we reduce the $2n$ a priori independent parameters $a_j^{(j')}$ and $c_j^{(j')}$, $j' = 1, \dots, n_j$, $j = 1, \dots, N$ $(n = \sum_{j=1}^N n_j)$, by half.

Proposition 5.1.5. *The formula of the solution given in Proposition 4.3.1 can be reformulated as follows:* $q(x,t) = 1 - P(x,t)/p(x,t)$ *with*

$$\begin{aligned} p(x,t) &= \det\left(I + \begin{pmatrix} M & 0 \\ 0 & \overline{M} \end{pmatrix} \begin{pmatrix} 0 & -T \\ T' & 0 \end{pmatrix}\right), \\ P(x,t) &= \det\left(I + \begin{pmatrix} M & 0 \\ 0 & \overline{M} \end{pmatrix} \begin{pmatrix} 0 & -T \\ T' & f \otimes f \end{pmatrix}\right), \end{aligned}$$

where
(i) $T = (T_{ij})_{i,j=1}^N$ *with*

$$T_{ij} = \left((-1)^{\nu+\mu} \binom{\nu+\mu-2}{\nu-1} \left(\frac{1}{\alpha_i + \overline{\alpha}_j}\right)^{\nu+\mu-1} \right)_{\substack{\nu=1,\dots,n_i \\ \mu=1,\dots,n_j}},$$

(ii) $f \in \mathbb{C}^n$ *denotes the vector* $f = (e_{n_1}^{(1)}, \dots, e_{n_N}^{(1)})$ *consisting of the first standard basis vectors* $e_{n_j}^{(1)} \in \mathbb{C}^{n_j}$ *for* $j = 1, \dots, N$, *and*
(iii) $M = diag\{M_j \mid j = 1, \dots, N\}$ *with*

$$M_j = \begin{pmatrix} m_j^{(1)} & & m_j^{(n_j)} \\ & \begin{smallmatrix} & \cdot \\ \cdot & \end{smallmatrix} & \\ m_j^{(n_j)} & & 0 \end{pmatrix} \in \mathcal{M}_{n_j,n_j}(\mathbb{C}), \tag{5.7}$$

the entries given by

$$m_j^{(\mu)} = m_j^{(\mu)}(x,t) = \sum_{\kappa=1}^{n_j-(\mu-1)} b_j^{(\mu-1+\kappa)} \frac{1}{(\kappa-1)!} \frac{\partial^{\kappa-1}}{\partial \alpha_j^{\kappa-1}} \exp\big(\alpha_j x + f_0(\alpha_j) t\big) \tag{5.8}$$

with constants $b_j^{(\mu)} = \sum_{\kappa=1}^{n_j-(\mu-1)} a_j^{(\kappa)} c_j^{(\mu-1+\kappa)}$ *for* $\mu = 1, \dots, n_j$.

Proof By Proposition 4.3.1 the solution q is given by $q = 1 - P/p$ with

$$P = \det\left(I + \begin{pmatrix} 0 & -L \\ \overline{L} & \overline{L_0} \end{pmatrix}\right), \qquad p = \det\left(I + \begin{pmatrix} 0 & -L \\ \overline{L} & 0 \end{pmatrix}\right),$$

where $L = \widehat{L}\Phi_{A,\overline{A}}^{-1}(\overline{a} \otimes c)$, $L_0 = \widehat{L}(a \otimes c)$, and $\widehat{L} = \widehat{L}(x,t) = \exp\big(Ax + f_0(A)t\big)$. Our aim is to show that they can be rewritten as stated in the assertion. We focus on P, which is the more involved case. The arguments for p are similar but easier.

Observe $\widehat{L} = \mathrm{diag}\{\widehat{L}_j \mid j = 1, \dots, N\}$, where $\widehat{L}_j = \widehat{L}_j(x,t) = \exp\big(A_j x + f_0(A_j)t\big)$. Set $\ell_j(x,t) = \exp\big(\alpha_j x + f_0(\alpha_j)t\big)$. For the Jordan block A_j, it is straightforward to calculate

$$\widehat{L}_j = \begin{pmatrix} \ell_j^{(1)} & & \ell_j^{(n_j)} \\ & \ddots & \\ 0 & & \ell_j^{(1)} \end{pmatrix} \quad \text{with } \ell_j^{(\mu)} = \frac{1}{(\mu-1)!} \frac{\partial^{\mu-1}}{\partial \alpha_j^{\mu-1}} \ell_j \text{ for } \mu = 1, \dots, n_j. \tag{5.9}$$

Next, by Proposition 4.1.3,

$$\Phi^{-1}_{A,\overline{A}}(\overline{a}\otimes c)=\Gamma_c T\overline{\Gamma}_a \quad \text{ and } \quad \Phi^{-1}_{\overline{A},A}(a\otimes \overline{c})=\overline{\Gamma}_c T'\Gamma_a$$

(Note $T'\neq \overline{T}$!) for $\Gamma_c=\operatorname{diag}\{\Gamma_l(c_j)\mid j=1,\dots,N\}$ and $\Gamma_a=\operatorname{diag}\{\Gamma_r(a_j)\mid j=1,\dots,N\}$ with the diagonal consisting of upper left and upper right band matrices

$$\Gamma_l(c_j)=\begin{pmatrix} c_j^{(1)} & & c_j^{(n_j)}\\ & \iddots & \\ c_j^{(n_j)} & & 0\end{pmatrix} \quad \text{ and } \quad \Gamma_r(a_j)=\begin{pmatrix} a_j^{(1)} & & a_j^{(n_j)}\\ & \ddots & \\ 0 & & a_j^{(1)}\end{pmatrix},$$

which are defined in terms of the constant vectors $c_j=(c_j^{(\mu)})_{\mu=1}^{n_j}$, $a_j=(a_j^{(\mu)})_{\mu=1}^{n_j}\in\mathbb{C}^{n_j}$. Moreover, we check $\overline{a}\otimes\overline{c}=\overline{\Gamma}_c(f\otimes f)\overline{\Gamma}_a$.

Therefore,

$$\begin{aligned} P &= \det\left(I+\begin{pmatrix}\widehat{L} & 0\\ 0 & \overline{\widehat{L}}\end{pmatrix}\begin{pmatrix}0 & -\Gamma_c T\overline{\Gamma}_a\\ \overline{\Gamma}_c T'\Gamma_a & \overline{\Gamma}_c(f\otimes f)\overline{\Gamma}_a\end{pmatrix}\right)\\ &= \det\left(I+\begin{pmatrix}\widehat{L}\Gamma_c & 0\\ 0 & \overline{\widehat{L}}\overline{\Gamma}_c\end{pmatrix}\begin{pmatrix}0 & -T\\ T' & f\otimes f\end{pmatrix}\begin{pmatrix}\Gamma_a & 0\\ 0 & \overline{\Gamma}_a\end{pmatrix}\right)\\ &= \det\left(I+\begin{pmatrix}\Gamma_a\widehat{L}\Gamma_c & 0\\ 0 & \overline{\Gamma_a\widehat{L}\Gamma_c}\end{pmatrix}\begin{pmatrix}0 & -T\\ T' & f\otimes f\end{pmatrix}\right).\end{aligned}$$

Obviously, $M=\Gamma_a\widehat{L}\Gamma_c$ again is a diagonal matrix with the blocks $M_j=\Gamma_r(a_j)\widehat{L}_j\Gamma_l(c_j)$, $j=1,\dots,N$, on the diagonal. As for those blocks, we use the fact that upper right band matrices commute to see

$$M_j=\Gamma_r(a_j)\widehat{L}_j\Gamma_l(c_j)=\widehat{L}_j(\Gamma_r(a_j)\Gamma_l(c_j)).$$

It is straightforward to verify

$$\Gamma_r(a_j)\Gamma_l(c_j)=\Gamma_l(b_j)=\begin{pmatrix} b_j^{(1)} & & b_j^{(n_j)}\\ & \iddots & \\ b_j^{(n_j)} & & 0\end{pmatrix},$$

and then it follows easily that M_j is of the form (5.7) with entries given by (5.8). □

Part 2. Asymptotic estimates

Again we proceed in two steps. In the first step we show, that the N-negaton asymptotically is a superposition of N single negatons. Then, in the second step, we investigate how a single negaton behaves.

We only consider $t\to-\infty$, since the case $t\to+\infty$ is completely symmetric.

In particular we can always assume $t<0$.

Step 1. To distinguish the single negatons, we associate the velocity v_{j_0} given by (5.1) to the j_0-th negaton (i.e., the negaton corresponding to the Jordan block A_{j_0}).

Then the index set $\Lambda_{j_0}^-=\{j\mid v_j>v_{j_0}\}$ corresponds to the negatons which move faster and will hence overtake the j_0-th negaton, $\Lambda_{j_0}^+=\{j\mid v_j<v_{j_0}\}$ to the slower negatons, which will be overtaken by the j_0-th negaton.

Proposition 5.1.6. *$q(x,t) \approx \sum_{j_0=1}^{N} q^{(j_0)}(x,t)$ for $t \approx -\infty$,*

where $q^{(j_0)} = 1 - P^{(j_0)}/p^{(j_0)}$ with

$$\begin{aligned} P^{(j_0)} &= \det\left(\begin{pmatrix} I^{(j_0)} & 0 \\ 0 & I^{(j_0)} \end{pmatrix} + \begin{pmatrix} M^{(j_0)} & 0 \\ 0 & \overline{M}^{(j_0)} \end{pmatrix}\begin{pmatrix} 0 & -T^{(j_0)} \\ (T^{(j_0)})' & f^{(j_0)} \otimes f^{(j_0)} \end{pmatrix}\right) \\ p^{(j_0)} &= \det\left(\begin{pmatrix} I^{(j_0)} & 0 \\ 0 & I^{(j_0)} \end{pmatrix} + \begin{pmatrix} M^{(j_0)} & 0 \\ 0 & \overline{M}^{(j_0)} \end{pmatrix}\begin{pmatrix} 0 & -T^{(j_0)} \\ (T^{(j_0)})' & 0 \end{pmatrix}\right), \end{aligned}$$

and the entries are defined by

$$\begin{aligned} T^{(j_0)} &= \big(\,T_{ij}\,\big)_{i,j\in\Lambda^-_{j_0}\cup\{j_0\}}, \qquad f^{(j_0)} = \big(\,e^{(1)}_{n_j}\,\big)_{j\in\Lambda^-_{j_0}\cup\{j_0\}}, \\ I^{(j_0)} &= \operatorname{diag}\big\{I^{(j_0)}_j \mid j\in\Lambda^-_{j_0}\cup\{j_0\}\big\} \ \text{ with the blocks } I^{(j_0)}_j = \begin{cases} I_{n_{j_0}}, & j=j_0, \\ 0, & j\in\Lambda^-_{j_0}, \end{cases} \\ M^{(j_0)} &= \operatorname{diag}\big\{M^{(j_0)}_j \mid j\in\Lambda^-_{j_0}\cup\{j_0\}\big\} \ \text{ with the blocks } M^{(j_0)}_j = \begin{cases} M_{j_0}, & j=j_0, \\ I_{n_{j_0}}, & j\in\Lambda^-_{j_0}. \end{cases} \end{aligned}$$

Note that all the matrices appearing in the above formula belong to $\mathcal{M}_{n^{(j_0)},n^{(j_0)}}(\mathbb{C})$ for $n^{(j_0)} = \sum_{j\in\Lambda^-_{j_0}\cup\{j_0\}} n_j$.

For the proof we need the following elementary perturbation lemma. If S is a matrix, we denote by $|S|$ the maximum of the moduli of its entries.

Lemma 5.1.7. *Let δ, $\delta_0 > 0$ and $t_0 < 0$. Furthermore, let $S(t)$, $S_0(t) \in \mathcal{M}_{m,m}(\mathbb{C})$ be defined for $t \le t_0$ and satisfy $|S_0(t)| \le \exp(-\delta_0 t)$, $|S(t) - S_0(t)| \le \exp(\delta t)$. Then there is a constant $\gamma > 0$, only depending on m, such that $|\det\big(S(t)\big) - \det\big(S_0(t)\big)| \le \gamma\exp\big((\delta - m\,\delta_0)t\big)$ for all $t \le t_0$.*

Proof Set $\widehat{S} = S - S_0$. With $S_0 = \big(s_{0,ij}\big)_{i,j=1}^m$, $\widehat{S} = \big(\widehat{s}_{ij}\big)_{i,j=1}^m$, the very definition of the determinant yields

$$\det(S) = \det\big(\widehat{S} + S_0\big) = \sum_{\pi\in\mathrm{Perm}(m)} \operatorname{sgn}(\pi) \prod_{j=1}^{m} (\widehat{s}_{j\pi(j)} + s_{0,j\pi(j)}),$$

where $\mathrm{Perm}(m)$ is the set of all permutations of $\{1,\ldots,m\}$. Multiplying out, and collecting all terms belonging only to S_0, we get

$$\det(S) - \det(S_0) = \sum_{\pi\in\mathrm{Perm}(m)} \sum_{\mu=1}^{m} \sum_{i_1<\ldots<i_\mu} \left[\operatorname{sgn}(\pi) \prod_{\kappa=1}^{\mu} \widehat{s}_{i_\kappa\pi(i_\kappa)} \prod_{j\notin\{i_1,\ldots,i_\mu\}} s_{0,j\pi(j)}\right]. \tag{5.10}$$

By assumption,

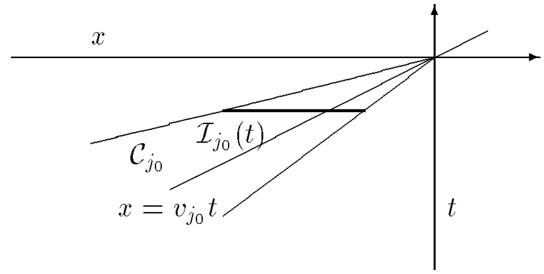
$$\begin{aligned} \Big|\prod_{\kappa=1}^{\mu} \widehat{s}_{i_\kappa\pi(i_\kappa)} \prod_{j\ne i_1,\ldots,i_\mu} s_{0,j\pi(j)}\Big| &\le \exp\big((\mu\delta - (m-\mu)\delta_0)t\big) \\ &\le \exp\big((\delta - m\,\delta_0)t\big), \end{aligned}$$

the latter because $t < 0$. Inserting this into (5.10), we see that for $\gamma = \gamma(m) = (2^m - 1)m!$ we have

$$|\det(S) - \det(S_0)| \le \gamma\exp\big((\delta - m\,\delta_0)t\big).$$

for all $t \le t_0$. □

Proof (of Proposition 5.1.6) Fix j_0. Let $x \in \mathcal{I}_{j_0}(t) = \big((v_{j_0}+\delta_{j_0})t, (v_{j_0}-\delta_{j_0})t\big)$. This interval has the center $v_{j_0}t$, and its diameter grows linearly with t. Furthermore, $\mathcal{C}_{j_0} = \bigcup_{t\le 0} \mathcal{I}_{j_0}(t)$ is a cone with vertex in the origin.
We choose $\delta_{j_0} < \min_{j\neq j_0} |v_j - v_{j_0}|$ in order not to cross the path of another negaton.



We subdivide our task in three steps:

Step a: First we show that, in the asymptotic sense, the only contribution to the N-negaton q in $\mathcal{C}_{j_0}$ is due to the j_0-th negaton $q^{(j_0)}$.

Step b: Next we show that outside $\mathcal{C}_{j_0}$ the j_0-th negaton $q^{(j_0)}$ asymptotically vanishes.

Setp c: Finally we show that the N-negaton q vanishes asymptotically outside $\bigcup_{j_0=1}^{N} \mathcal{C}_{j_0}$.

In fact, any other cone with parallel edges could be used likewise (modulo obvious modifications in the estimates). The reason is that the width of $\mathcal{I}_{j_0}(t)$ grows linearly in t whereas that of the negaton grows only logarithmically. Hence $\mathcal{I}_{j_0}(t)$ will contain an arbitrarily large portion of the negaton, if t is large enough.

Note that $\mathcal{I}_{j_0}(t)$ is a proper interval since we have assumed $t < 0$.

Step a: Let us start with some elementary considerations, where we aim at estimates for the entries of M_j which hold uniformly for $x \in \mathcal{I}_{j_0}(t)$.

For $j \in \Lambda^+_{j_0}$, $\gamma_{j_0 j} := (v_{j_0} - v_j) - \delta_{j_0}$ is positive by the choice of δ_{j_0}, and we have $x - v_j t \le \gamma_{j_0 j} t$ for all $x \in \mathcal{I}_{j_0}(t)$. Recall that, by assumption, $\mathrm{Re}(\alpha_j) > 0\ \forall j$. Thus, from (5.7), (5.8), we immediately get:

$$\exists \gamma_{j_0} > 0,\ t_{j_0} < 0 : \quad |M_j| < \exp(\gamma_{j_0} t) \quad \forall\, t \le t_{j_0},\ x \in \mathcal{I}_{j_0}(t).$$

An analogous estimate can be derived for M_j^{-1} if $j \in \Lambda^-_{j_0}$. To this end note that M_j^{-1} is of the same structure as derived for M_j in (5.7), (5.8), of course with different constants, but the exponential function in (5.8) just has to be replaced by its inverse.

In summary, we get

$$\exists \gamma_{j_0} > 0,\ t_{j_0} < 0 : \quad \left\{ \begin{array}{rl} |M_j| < \exp(\gamma_{j_0} t), & \text{if } j \in \Lambda^+_{j_0}, \\ |M_j^{-1}| < \exp(\gamma_{j_0} t), & \text{if } j \in \Lambda^-_{j_0}, \end{array} \right. \tag{5.11}$$

for all $t \le t_{j_0}$ and $x \in \mathcal{I}_{j_0}(t)$.

This motivates to replace P, p by determinants $\widehat{P}, \widehat{p}$ with the property that all parameter-dependent entries, except of the one corresponding to the j_0-th negaton, are of the form (5.11) and thus decrease exponentially. This is done as follows:

$$\widehat{P} = \det \begin{pmatrix} D & 0 \\ 0 & \overline{D} \end{pmatrix} P \quad \text{and} \quad \widehat{p} = \det \begin{pmatrix} D & 0 \\ 0 & \overline{D} \end{pmatrix} p, \tag{5.12}$$

where D is a matrix eliminating the exploding blocks of M:

$$D = \mathrm{diag}\{D_j \mid j = 1, \dots, N\} \text{ with } D_j = \left\{ \begin{array}{ll} M_j^{-1}, & j \in \Lambda^-_{j_0}, \\ I_{n_j}, & j \in \Lambda^+_{j_0} \cup \{j_0\}. \end{array} \right.$$

As a consequence,

$$DM = \mathrm{diag}\{D_j M_j \mid j = 1, \dots, N\} \text{ with } D_j M_j = \begin{cases} I_{n_j}, & j \in \Lambda^-_{j_0}, \\ M_j, & j \in \Lambda^+_{j_0} \cup \{j_0\}. \end{cases}$$

The above manipulation of P, p does not alter the solution, since $q = 1 - P/p = 1 - \widehat{P}/\widehat{p}$.

Let us now focus on $\widehat{P}$. The arguments for $\widehat{p}$ are completely the same. By the product rule for determinants,

$$\begin{aligned} \widehat{P} &= \det\left(\begin{pmatrix} D & 0 \\ 0 & \overline{D} \end{pmatrix} + \begin{pmatrix} DM & 0 \\ 0 & \overline{DM} \end{pmatrix} \begin{pmatrix} 0 & -T \\ T' & f \otimes f \end{pmatrix} \right) \\ &= \det \begin{pmatrix} D & -DMT \\ \overline{DM}T' & \overline{D} + \overline{DM}(f \otimes f) \end{pmatrix} = \det \begin{pmatrix} D & -S \\ \widehat{S} & \overline{D} + R \end{pmatrix} \end{aligned}$$

with

$$\begin{aligned} S_{ij} &= \begin{cases} T_{ij}, & i \in \Lambda^-_{j_0}, \\ M_i T_{ij}, & i \in \Lambda^+_{j_0} \cup \{j_0\}, \end{cases} \qquad \widehat{S}_{ij} = \begin{cases} T'_{ji}, & i \in \Lambda^-_{j_0}, \\ \overline{M}_i T'_{ji}, & i \in \Lambda^+_{j_0} \cup \{j_0\}, \end{cases} \\ \text{and } R_{ij} &= \begin{cases} (f_j \otimes f_i), & i \in \Lambda^-_{j_0}, \\ \overline{M}_i (f_j \otimes f_i), & i \in \Lambda^+_{j_0} \cup \{j_0\}. \end{cases} \end{aligned}$$

Next we will interpret $\widehat{P}$ as a perturbation of a determinant $\widehat{P}^{(j_0)}$ of a matrix whose rows are t-independent except the two rows corresponding to the index j_0. To carry out this argument, we define

$$\widehat{P}^{(j_0)} = \det \begin{pmatrix} D^{(j_0)} & -S^{(j_0)} \\ \widehat{S}^{(j_0)} & D^{(j_0)} + R^{(j_0)} \end{pmatrix}$$

with the blocks

$$\begin{aligned} S^{(j_0)}_{ij} &= \begin{cases} T_{ij}, & i \in \Lambda^-_{j_0}, \\ M_{j_0} T_{j_0 j}, & i = j_0, \\ 0, & i \in \Lambda^+_{j_0}, \end{cases} \qquad \widehat{S}^{(j_0)}_{ij} = \begin{cases} T'_{ji}, & i \in \Lambda^-_{j_0}, \\ \overline{M}_{j_0} T'_{j j_0}, & i = j_0, \\ 0, & i \in \Lambda^+_{j_0}, \end{cases} \\ \text{and } R^{(j_0)}_{ij} &= \begin{cases} (f_j \otimes f_i), & i \in \Lambda^-_{j_0}, \\ \overline{M}_{j_0} (f_j \otimes f_{j_0}), & i = j_0, \\ 0, & i \in \Lambda^+_{j_0}, \end{cases} \end{aligned}$$

and $D^{(j_0)} = \mathrm{diag}\{D^{(j_0)}_j \mid j = 1, \dots, N\}$ with $D^{(j_0)}_j = \begin{cases} 0, & j \in \Lambda^-_{j_0}, \\ I_{n_j}, & j \in \Lambda^+_{j_0} \cup \{j_0\}, \end{cases}$

By Lemma 5.1.7, we obtain

$$\sup_{x \in \mathcal{I}_{j_0}(t)} \left| \widehat{P}(x,t) - \widehat{P}^{(j_0)}(x,t) \right| \lesssim \exp\left(\widehat{\gamma}_{j_0} t\right) \quad \text{for all } t \le t_{j_0},$$

where $\widehat{\gamma}_{j_0} = \gamma_{j_0} - 2n\delta_{j_0}/\mathrm{Re}(\alpha_{j_0})$. Note that possibly γ_{j_0} has to be reduced in size to include the constants from the blocks $T_{j_0 j}$, $T'_{j j_0}$ $(\forall j)$ in the estimates. Choosing δ_{j_0} small enough, we can achieve $\widehat{\gamma}_{j_0} > 0$.

Let us take a closer look at the determinant $\widehat{P}^{(j_0)}$ defined above. For $i \in \Lambda^+_{j_0}$ all entries in the i-th row of $S^{(j_0)}$, $\widehat{S}^{(j_0)}$ vanish and the only non-vanishing entry in the i-th row of both $D^{(j_0)}$ and $D^{(j_0)} + R^{(j_0)}$ is I_{n_i} in the i-th column. Therefore, straightforward expansion of the determinant $\widehat{P}^{(j_0)}$ shows that $\widehat{P}^{(j_0)}$ actually coincides with the determinant $P^{(j_0)}$ defined in the assertion.

Therefore,

$$\sup_{x\in\mathcal{I}_{j_0}(t)} \left|\widehat{P}(x,t) - P^{(j_0)}(x,t)\right| \longrightarrow 0 \quad \text{as } t\to-\infty$$

and, by the same arguments, also $\sup_{x\in\mathcal{I}_{j_0}(t)} \left|\widehat{p}(x,t) - p^{(j_0)}(x,t)\right|$ converges to zero.

Consider now $q = 1 - P/p = 1 - \widehat{P}/\widehat{p}$. Since q is a regular solution by assumption, $\widehat{p}(x,t) \neq 0$ for all x, t. The above convergence implies that also $p^{(j_0)}(x,t) \neq 0$ for all x, t. Thus we can transfer the convergence from $\widehat{P}$, $\widehat{p}$ to q.

In summary, we have shown that q asymptotically behaves like $q^{(j_0)} = 1 - P^{(j_0)}/p^{(j_0)}$ on $\mathcal{C}_{j_0}$ as t tends to $-\infty$. More precisely, we have $1_{\mathcal{C}_{j_0}}\, q \approx 1_{\mathcal{C}_{j_0}} q^{(j_0)}$ for $t\approx-\infty$, where $1_{\mathcal{C}_{j_0}}$ is the characteristic function of $\mathcal{C}_{j_0}$. Step a is complete.

Step b: Next we discuss the behaviour of the j_0-th negaton $q^{(j_0)}$ outside of $\mathcal{I}_{j_0}(t)$. We distinguish two cases:

(i) Let $x \in \mathcal{I}^-_{j_0}(t) = \big(-\infty, (v_{j_0}+\delta_{j_0})t\,\big]$. Here the entries of $M^{(j_0)}_{j_0}$ decay exponentially. Similar arguments as before show

$$\sup_{x\in\mathcal{I}^-_{j_0}(t)} \left|P^{(j_0)}(x,t) - C\right| \longrightarrow 0, \qquad \sup_{x\in\mathcal{I}^-_{j_0}(t)} \left|p^{(j_0)}(x,t) - c\right| \longrightarrow 0, \qquad \text{as } t\to-\infty,$$

where

$$C = \det\begin{pmatrix} 0 & -T^- \\ (T^-)' & f^-\otimes f^- \end{pmatrix}, \qquad c = \det\begin{pmatrix} 0 & -T^- \\ (T^-)' & 0 \end{pmatrix},$$

with $T^- = \big(\, T_{ij}\,\big)_{i,j\in\Lambda^-_{j_0}}$ and $f^- = \big(\, f_j\,\big)_{j\in\Lambda^-_{j_0}}$. Moreover, by Lemma 5.1.8, the values c, C of the two determinants coincide, and are nonzero by Theorem 6.1.1.

(ii) Let $x \in \mathcal{I}^+_{j_0}(t) = \big[\,(v_{j_0}-\delta_{j_0})t, \infty\big)$. In this case the entries $M^{(j_0)}_{j_0}$ of the determinants $P^{(j_0)}$, $p^{(j_0)}$ explode. Thus we again replace $P^{(j_0)}$, $p^{(j_0)}$ by determinants $\widehat{P}^{(j_0)}$, $\widehat{p}^{(j_0)}$ with

$$\begin{aligned}
\widehat{P}^{(j_0)} &= \det\left(\begin{pmatrix} M^{(j_0)} & 0 \\ 0 & \overline{M}^{(j_0)} \end{pmatrix}^{-1}\right) P^{(j_0)} \\
&= \det\left(\begin{pmatrix} (M^{(j_0)})^{-1} I^{(j_0)} & 0 \\ 0 & (\overline{M}^{(j_0)})^{-1} I^{(j_0)} \end{pmatrix} + \begin{pmatrix} 0 & -T^{(j_0)} \\ (T^{(j_0)})' & f^{(j_0)}\otimes f^{(j_0)} \end{pmatrix}\right),
\end{aligned}$$

where

$$(M^{(j_0)})^{-1} I^{(j_0)} = \operatorname{diag}\big\{ (M^{(j_0)}_j)^{-1} I^{(j_0)}_j \mid j\in\Lambda^-_{j_0}\cup\{j_0\}\big\}$$
$$\text{with } (M^{(j_0)}_j)^{-1} I^{(j_0)}_j = \begin{cases} M^{-1}_{j_0}, & j = j_0, \\ 0, & j\in\Lambda^-_{j_0}, \end{cases}$$

and $\widehat{p}^{(j_0)}$ is modified analogously.

Then

$$\sup_{x\in\mathcal{I}^+_{j_0}(t)} \left|\widehat{P}^{(j_0)}(x,t) - \widehat{C}\right| \longrightarrow 0, \qquad \sup_{x\in\mathcal{I}^+_{j_0}(t)} \left|\widehat{p}^{(j_0)}(x,t) - \widehat{c}\right| \longrightarrow 0, \qquad \text{as } t\to-\infty,$$

where

$$\widehat{C} = \det\begin{pmatrix} 0 & -T^{(j_0)} \\ (T^{(j_0)})' & f^{(j_0)} \otimes f^{(j_0)} \end{pmatrix}, \qquad \widehat{c} = \det\begin{pmatrix} 0 & -T^{(j_0)} \\ (T^{(j_0)})' & 0 \end{pmatrix}.$$

Again $\widehat{C} = \widehat{c} \neq 0$ by Lemma 5.1.8 and Theorem 6.1.1.

Consequently, $q^{(j_0)}$ asymptotically vanishes on $\mathbb{R}^2 \backslash \mathcal{C}_{j_0}$ as t tends to $-\infty$. More precisely, $(1 - 1_{\mathcal{C}_{j_0}})q^{(j_0)} \approx 0$ for $t \approx -\infty$. Step b is complete.

Step c: It remains to show that q asymptotically vanishes outside $\cup_{j_0=1}^{N} \mathcal{C}_{j_0}$ as $t \to -\infty$. This is done in the remainder of the proof with analogous arguments as before.

If necessary, we shrink δ_{j_0} once more such that $\delta_{j_0} < \frac{1}{2} \min_{j \neq j_0} |v_{j_0} - v_j|$ for all j_0. Consider $v_{j_2} < v_{j_1}$ such that no other v_j lies between them, and let

$$x \in \mathcal{I}_{j_1,j_2}(t) = \big[(v_{j_1} - \delta_{j_1})t, (v_{j_2} + \delta_{j_2})t\big].$$

Those are just the intervals covering the gaps between the two neighbouring strips $\mathcal{I}_{j_1}(t)$, $\mathcal{I}_{j_2}(t)$. They are proper because δ_{j_1}, δ_{j_2} are chosen small enough.

Now it is straightforward to check a similar estimate as before, namely

$$\exists \gamma_{j_1,j_2} > 0,\ t_{j_0} < 0 : \quad \begin{cases} |M_j| < \exp(\gamma_{j_1,j_2} t), & \text{if } j \in \Lambda^+_{j_2}, \\ |M_j^{-1}| < \exp(\gamma_{j_1,j_2} t), & \text{if } j \in \Lambda^-_{j_1}, \end{cases}$$

for all $t \leq t_{j_0}$ and $x \in \mathcal{I}_{j_1,j_2}(t)$. Note that $\Lambda^-_{j_1} = \{j \mid v_j > v_{j_1}\}$, $\Lambda^+_{j_2} = \{j \mid v_{j_2} > v_j\}$ cover all possible indices j.

Following the same line of reasoning as before, we next

(i) replace P, p by $\widehat{P}$, $\widehat{p}$ in the same way as in (5.12) with the only difference that we use here the index sets $\Lambda^-_{j_1}$ instead of $\Lambda^-_{j_0}$ and $\Lambda^+_{j_2}$ instead of $\Lambda^+_{j_0} \cup \{j_0\}$,

(ii) apply Lemma 5.1.7, and expand the resulting determinant.

As a result,

$$\sup_{x \in \mathcal{I}_{j_1,j_2}(t)} \Big|\widehat{P}(x,t) - \widehat{C}\Big| \longrightarrow 0, \qquad \sup_{x \in \mathcal{I}_{j_1,j_2}(t)} \Big|\widehat{p}(x,t) - \widehat{c}\Big| \longrightarrow 0, \qquad \text{as } t \to -\infty,$$

where

$$\widehat{C} = \det\begin{pmatrix} 0 & -\widehat{T} \\ \widehat{T}' & \widehat{f} \otimes \widehat{f} \end{pmatrix}, \qquad \widehat{c} = \det\begin{pmatrix} 0 & -\widehat{T} \\ \widehat{T}' & 0 \end{pmatrix},$$

with $\widehat{T} = (\ T_{ij}\ )_{i,j \in \Lambda^-_{j_1}}$ and $\widehat{f} = (\ f_j\ )_{j \in \Lambda^-_{j_1}}$. Again Lemma 5.1.8, Theorem 6.1.1 yield $\widehat{C} = \widehat{c} \neq 0$.

We still have to consider (i) $x \in \mathcal{I}_{\min}(t) = \big(-\infty, (v_{j_{\min}} + \delta_{j_{\min}})t\ \big]$ for $v_{j_{\min}} = \min_j(v_j)$, and (ii) $x \in \mathcal{I}_{\max}(t) = \big[\ (v_{j_{\max}} - \delta_{j_{\max}})t, \infty\big)$ for $v_{j_{\max}} = \max_j(v_j)$. In the case (ii), all M_j decay exponentially. Thus

$$\sup_{x \in \mathcal{I}_{\max}(t)} \Big|P(x,t) - 1\Big| \longrightarrow 0, \qquad \sup_{x \in \mathcal{I}_{\max}(t)} \Big|p(x,t) - 1\Big| \longrightarrow 0, \qquad \text{as } t \to -\infty,$$

and we are done.

In the case (i), all M_j explode. Here we replace P, p by $\widehat{P}, \widehat{p}$ with

$$\widehat{P} = \det\begin{pmatrix} M^{-1} & -T \\ T' & \overline{M}^{-1} + f \otimes f \end{pmatrix}, \quad \widehat{p} = \det\begin{pmatrix} M^{-1} & -T \\ T' & \overline{M}^{-1} \end{pmatrix}.$$

Then

$$\sup_{x \in \mathcal{I}_{\min}(t)} \left| \widehat{P}(x,t) - C \right| \longrightarrow 0, \qquad \sup_{x \in \mathcal{I}_{\min}(t)} \left| \widehat{p}(x,t) - c \right| \longrightarrow 0, \qquad \text{as } t \to -\infty,$$

$$C = \det\begin{pmatrix} 0 & -T \\ T' & f \otimes f \end{pmatrix}, \qquad c = \det\begin{pmatrix} 0 & -T \\ T' & 0 \end{pmatrix}$$

and, by Lemma 5.1.8, Theorem 6.1.1, we are done again.

Thus q asymptotically vanishes on $\mathbb{R}^2 \setminus \cup_{j_0=1}^{N} \mathcal{C}_{j_0}$ as t tends to $-\infty$. More precisely, $\big(1 - \sum_{j_0=1}^{N} 1_{\mathcal{C}_{j_0}}\big) q \approx 0$ for $t \approx -\infty$. Step c is complete.

In summary, Proposition 5.1.6 is proved. $\square$

Finally we supply the simple fact on determinants used throughout the proof.

Lemma 5.1.8. *Let U, V, $W \in \mathcal{M}_{n,n}(\mathbb{C})$ be arbitrary square matrices. Then the following identity holds,*

$$\det\begin{pmatrix} 0 & -U \\ V & W \end{pmatrix} = \det(U)\det(V)$$

Proof The assertion follows easily by the subsequent calculation

$$\begin{aligned} \det\begin{pmatrix} 0 & -U \\ V & W \end{pmatrix} &= \det\left(\begin{pmatrix} U & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} 0 & -I \\ I & W \end{pmatrix} \begin{pmatrix} V & 0 \\ 0 & I \end{pmatrix} \right) \\ &= \det(U)\det(V)\ \det\begin{pmatrix} 0 & -I \\ I & W \end{pmatrix} \\ &= (-1)^n \det(U)\det(V)\ \det\begin{pmatrix} -I & 0 \\ W & I \end{pmatrix} \\ &= \det(U)\det(V), \end{aligned}$$

where for the third identity we have used the obvious permutation argument with respect to the columns. $\square$

Step 2. As a preparation for the inner analysis of the j_0-th negaton, we first give the basic estimates. To this end we study what happens if we deviate from its path logarithmically, that is $x - v_{j_0} t \sim \log|t|$.

Proposition 5.1.9. *On the curve $\big(\gamma_\rho(t), t\big)$, where $\gamma_\rho(t) = v_{j_0} t + \big(\rho \log|t|\big)/\mathrm{Re}(\alpha_{j_0})$ for $\rho \in \mathbb{R}$, the determinants $P^{(j_0)}$, $p^{(j_0)}$ behave according to*

$$p^{(j_0)}\big(\gamma_\rho(t), t\big) \;=\; C \left(\sum_{\kappa=0}^{n_{j_0}} C_{\kappa\kappa}\, F_\kappa(t)\overline{F_\kappa(t)}\, |t|^{2\kappa\rho}\, t^{2(n_{j_0}-\kappa)\kappa} \right) \Big[1 + \mathcal{O}\Big(\frac{\log|t|}{t}\Big) \Big]$$

and

$$P^{(j_0)}(\gamma_\rho(t),t) \quad = \quad C\,\Bigg(\sum_{\kappa=0}^{n_{j_0}} C_{\kappa\kappa}\; F_\kappa(t)\overline{F_\kappa(t)}\;|t|^{2\kappa\rho}\;t^{2(n_{j_0}-\kappa)\kappa}\;+$$

$$+\sum_{\kappa=0}^{n_{j_0}-1}(-1)^\kappa C_{\kappa(\kappa+1)}\;F_\kappa(t)\overline{F_{\kappa+1}(t)}\;|t|^{(2\kappa+1)\rho}\;t^{2(n_{j_0}-\kappa)\kappa+(n_{j_0}-2\kappa-1)}\Bigg)\Big[1+\mathcal{O}\Big(\frac{\log|t|}{t}\Big)\Big]$$

with $C=\Big(\det\big(T_{ij}\big)_{ij\in\Lambda^-_{j_0}}\Big)^2$,

$$C_{\kappa\lambda} \quad = \quad \left[\frac{1}{\alpha_{j_0}+\overline{\alpha}_{j_0}}\right]^{2\kappa\lambda}\prod_{j\in\Lambda^-_{j_0}}\left[\frac{\alpha_j-\alpha_{j_0}}{\overline{\alpha}_j+\alpha_{j_0}}\right]^{2\kappa n_j}\left[\frac{\overline{\alpha}_j-\overline{\alpha}_{j_0}}{\alpha_j+\overline{\alpha}_{j_0}}\right]^{2\lambda n_j},$$

$$\textit{and } F_\kappa(t) \quad = \quad \frac{\prod_{k=1}^{\kappa-1}k!}{\prod_{k=1}^{\kappa}(n_{j_0}-k)!}\Big(b_{j_0}^{(n_{j_0})}c_{j_0}^{n_{j_0}-\kappa}\exp\big(\,\mathrm{i}\Upsilon_{j_0}(\gamma_\rho(t),t)\,\big)\Big)^\kappa,$$

where $c_{j_0}=v_{j_0}+f_0'(\alpha_{j_0})$ *and* $\Upsilon_{j_0}(x,t)=\mathrm{Im}\,\big(\,\alpha_{j_0}x+f_0(\alpha_{j_0})t\,\big)$.

The parameter ρ controls the distance from the geometric center. Note that it enters in the asymptotic via the powers of $|t|$ and the term $\Upsilon_{j_0}(\gamma_\rho(t),t)$.

The main part of the proof is to expand the determinants $P^{(j_0)}$, $p^{(j_0)}$. We proceed in a similar manner as for solitons where the well-known expansion rule for determinants was used (see Lemma 4.1.6). Here we use a form of this rule which is particularly adapted to our problem. We need some more notation.

Consider the matrix

$$S=\begin{pmatrix} U & V\\ W & Z\end{pmatrix},$$

where each of the blocks U,V,W,Z itself has a block structure which is given by

$$U=\big(\,U_{ij}\,\big)_{ij\in\Lambda^-_{j_0}\cup\{j_0\}}\qquad\text{with}\qquad U_{ij}=\big(\,U_{ij}^{(\nu\mu)}\,\big)_{\substack{\nu=1,\dots,n_i\\ \mu=1,\dots,n_j}}$$

and for V,W,Z analogously.

For index tuples $J=(\sigma_1,\dots,\sigma_\kappa)$, $K=(\tau_1,\dots,\tau_\lambda)$ with $\sigma_k,\tau_l\in\{1,\dots,n_{j_0}\}$ for $1\le k\le\kappa$, $1\le l\le\lambda$, we define the matrix $S[J\times K]$ by

$$S[J\times K]=\begin{pmatrix} U[J,J] & V[J,K]\\ W[K,J] & Z[K,K]\end{pmatrix}.$$

Here $U[J,K]$ denotes the matrix with the blocks $U[J,K]_{ij}$, where

1. $U[J,K]_{ij}=U_{ij}$ if $i\neq j_0$, $j\neq j_0$,
2. $U[J,K]_{j_0j}$ is obtained from U_{j_0j} by maintaining only the rows indexed by J for $j\neq j_0$,
3. $U[J,K]_{ij_0}$ is obtained from U_{ij_0} by maintaining only the columns indexed by K for $i\neq j_0$,
4. $U[J,K]_{j_0j_0}$ is obtained from $U_{j_0j_0}$ by maintaining only the rows indexed by J and the columns indexed by K,

and, furthermore, all maintained columns and rows in the blocks $U[J,K]_{ij}$ appear precisely in the order which is indicated by the tuples J,K.

By $|J|$ we denote the length of the index tuple J, in the case at hand $|J| = \kappa$. Note that also the trivial case of empty index tuples is admitted, where no rows or columns are maintained in the respective blocks. Then the length of the index tuple is zero.

Finally, we need a substitute $\widehat{I}$ of the identity matrix. With respect to the expansion we have in mind (confer Proposition 5.1.6) we define

$$\widehat{I} = \begin{pmatrix} I^{(j_0)} & 0 \\ 0 & I^{(j_0)} \end{pmatrix}$$

for the diagonal matrix $I^{(j_0)}$ with the blocks $I_j^{(j_0)} = 0$, $j \in \Lambda_{j_0}^-$, and $I_{j_0}^{(j_0)} = I_{n_{j_0}}$ on the diagonal.

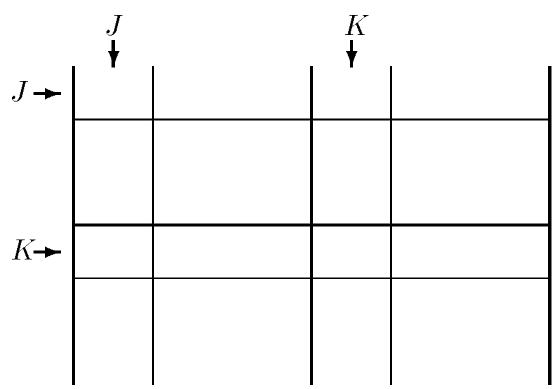
Routine arguments yield the following expansion rule.

Lemma 5.1.10. *In the situation described above, the following expansion rule holds:*

$$\det\left(\,\widehat{I} + S\,\right) = \sum_{\kappa=0}^{n_{j_0}} \sum_{\lambda=0}^{n_{j_0}} \quad {\sum_{|J|=\kappa}}' {\sum_{|K|=\lambda}}' \det\left(\,S[J \times K]\,\right),$$

where the inner sums are taken over all index tuples J, K from $\{1, \dots, n_{j_0}\}$ and the prime means that only index sets with strictly increasing entries are admitted.

Note that the cases $\kappa = 0$, $\lambda = 0$ correspond to the appearance of empty index sets.



To illustrate the expansion, we indicate the position of the blocks where only the indices listed in J, K, respectively, are maintained in the case $j_0 = 1$.

We also need another, technical result for the evaluation of a certain determinant which has already been proved in [88]. Let us start with the following preparation.

Lemma 5.1.11. *For all $\gamma \in \mathbb{C}$,*

$$\det\left(\prod_{\kappa=1}^{\mu-1} \left[\,\big(\gamma - (\nu - 1)\big) - \kappa\,\right] \right)_{\nu,\mu=1}^{k} = (-1)^{\frac{k(k+3)}{2}} \prod_{\kappa=1}^{k-1} \kappa! \quad ,$$

where it is understood that empty products equal 1.

Proof With special regard to the order induced by the numbering of the indices, we pursue the following strategy: Multiply the μ-th column by $\left[\big(\gamma - (n-1)\big) - \mu\right]$ and subtract it from the $(\mu + 1)$-th column for $\mu = k-1, \dots, 1$. This yields

$$\Delta \;=\; \det\left(\prod_{\kappa=1}^{\mu-1} \left[\big(\gamma - (\nu - 1)\big) - \kappa\right] \right)_{\nu,\mu=1}^{k}$$

$$= \ \det\left(1, \quad (k-\nu)\prod_{\kappa=1}^{\mu-2}\left[\big(\gamma-(\nu-1)-\kappa\big)\right]\right)_{\substack{\nu\geq 1\\ \mu>1}}$$

Now we expand the determinant with respect to the k-th row, which is zero except of the first entry. Then we extract the factor $(k-\nu)$, which is common to the ν-th row $(\nu = 1,\ldots,k-1)$. As a result,

$$\Delta \ = \ (-1)^{k+1}(k-1)! \quad \det\left(\prod_{\kappa=1}^{\mu-1}\left[\big(\gamma-(\nu-1)\big)-\kappa\right]\right)_{\nu,\mu=1}^{k-1},$$

and the assertion follows by induction. □

Corollary 5.1.12. *Let $n \geq k$ be natural numbers. Then the determinant of the matrix $F = \big(\mathrm{fac}_{\nu\mu}\big)_{\nu,\mu=1}^{k} \in \mathcal{M}_{k,k}(\mathbb{C})$ with the entries*

$$\mathrm{fac}_{\nu\mu} = \begin{cases} \dfrac{1}{[n-(\nu+\mu-1)]!}, & \nu+\mu \leq n+1, \\ 0, & \nu+\mu > n+1, \end{cases}$$

has the value

$$\det(F) \ = \ (-1)^{\frac{k(k+3)}{2}} \ \frac{\prod_{\kappa=1}^{k-1}\kappa!}{\prod_{\kappa=1}^{k}(n-\kappa)!},$$

where it is understood that $0! = 1$.

Proof For $\nu+\mu \leq n+1$, we have

$$(n-\nu)!\,\mathrm{fac}_{\nu\mu} = \frac{(n-\nu)!}{\big(n-(\nu+\mu-1)\big)!} = \prod_{\kappa=1}^{\mu-1}\left[\big(n-(\nu-1)\big)-\kappa\right]. \tag{5.13}$$

If, on the other hand, $\nu+\mu > n+1$ the identity (5.13) holds all the more since the product on the right then contains the factor corresponding to $\kappa = n-(\nu-1) < \mu$ and hence vanishes. Therefore,

$$\det(F) = \det\left(\frac{1}{(n-\nu)!}\cdot\prod_{\kappa=1}^{\mu-1}\left[\big(n-(\nu-1)\big)-\kappa\right]\right)_{\nu,\mu=1}^{k},$$

and, extracting the factors $1/(n-\nu)!$ common to the ν-th row for all ν, the assertion follows from Lemma 5.1.11. □

Proof (of Proposition 5.1.9) For the j_0-th negaton $q^{(j_0)}$ the only dependence on the variables x,t is due to the block M_{j_0}. Therefore, we take a closer look at the behaviour of its entries $m_{j_0}^{(\mu)}$ along the curve $\big(\gamma_\rho(t),t\big)$ for $\rho \in \mathbb{R}$ fixed.

To begin with, for $\ell(x,t) = \exp\big(\alpha_{j_0}x + f_0(\alpha_{j_0})t\big) = \exp\big(\,\mathrm{Re}(\alpha_{j_0})[x-v_{j_0}t] + \mathrm{i}\Upsilon_{j_0}(x,t)\,\big)$ we have

$$\ell\big(\gamma_\rho(t),t\big) = |t|^\rho \exp\Big(\mathrm{i}\Upsilon_{j_0}\big(\gamma_\rho(t),t\big)\Big).$$

Set $Q_\mu(x,t) = \big(\partial^\mu\ell(x,t)/\partial\alpha_{j_0}^\mu\big)/\ell(x,t)$. Then Q_μ are polynomials satisfying the recursion relation $Q_{\mu+1} = Q_\mu Q_1 + Q_\mu'$ with $Q_0 = 1$, $Q_1 = x + f_0'(\alpha_{j_0})t$ (the prime denoting the derivative with respect to α_{j_0}). Thus an induction argument shows

$$Q_\mu\big(\gamma_\rho(t),t\big) = (c_{j_0}t)^\mu\ \left[1+\mathcal{O}\Big(\frac{\log|t|}{t}\Big)\right],$$

and from (5.8) we get

$$m_{j_0}^{(\mu)}(\gamma_\rho(t),t) = b_{j_0}^{(n_{j_0})} \exp\Big(\mathrm{i}\Upsilon_{j_0}(\gamma_\rho(t),t)\Big) \frac{(c_{j_0}t)^{n_{j_0}-\mu}}{(n_{j_0}-\mu)!}|t|^\rho \ \Big[1+\mathcal{O}\Big(\frac{\log|t|}{t}\Big)\Big]. \tag{5.14}$$

Now we consider the determinants $P^{(j_0)} = \det\big(\widehat{I}+S\big)$, $p^{(j_0)} = \det\big(\widehat{I}+S_0\big)$. Application of Lemma 5.1.10 combined with Proposition 4.1.9 yields the expansions

$$\begin{aligned}
p^{(j_0)} &= \sum_{\kappa=0}^{n_{j_0}} \ {\sum_{|J|=\kappa}}' \ {\sum_{|K|=\kappa}}' \det\big(\, S_0[J\times K]\,\big), \\
P^{(j_0)} &= \sum_{\kappa=0}^{n_{j_0}} \ {\sum_{|J|=\kappa}}' \ {\sum_{|K|=\kappa}}' \det\big(\, S[J\times J]\,\big) + \\
&\qquad\qquad + \sum_{\kappa=0}^{n_{j_0}-1} \ {\sum_{|J|=\kappa}}' \ {\sum_{|K|=\kappa+1}}' \det\big(\, S[J\times K]\,\big)\,.
\end{aligned} \tag{5.15}$$

In the sequel we restrict the arguments to the treatment of $P^{(j_0)}$, which is the more involved case. Indeed the result for $p^{(j_0)}$ follows by completely the same line of arguments but only the sum concerning $|K|=\kappa$ has to be considered.

Our next aim is to calculate the principal minors $\det\big(S[J\times K]\big)$ for $|J|=\kappa$, $|K|=\lambda$ with $\lambda\in\{\kappa,\kappa+1\}$, where, by Proposition 5.1.6,

$$S = \begin{pmatrix} 0 & -M^{(j_0)}T^{(j_0)} \\ \overline{M}^{(j_0)}(T^{(j_0)})' & \overline{M}^{(j_0)}(f^{(j_0)}\otimes f^{(j_0)}) \end{pmatrix},$$

To reduce the expression to the calculation of determinants with known values, we have to expand once more.

For the moment abbreviate $U = -M^{(j_0)}T^{(j_0)}$. Then $U = \big(U_{ij}\big)_{i,j\in\Lambda_{j_0}^-\cup\{j_0\}}$, and its blocks are given by $U_{ij} = -T_{ij}$ for $i\in\Lambda_{j_0}^-$ and $U_{j_0 j} = -M_{j_0}T_{j_0 j}$, where

$$M_{j_0}T_{j_0 j} = \Big(\sum_{\kappa=1}^{n_{j_0}-(\nu-1)} t_{j_0 j}^{(\kappa\mu)} m_{j_0}^{(\nu+\kappa-1)}\Big)_{\substack{\nu=1,\dots,n_{j_0}\\ \mu=1,\dots,n_j}}.$$

Observe that the matrices in the lower left and right corners of S have a similar structure.

By the usual rules for the calculation of determinants (linearity with respect to rows), we observe for $J=(\sigma_1,\dots,\sigma_\kappa)$, $K=(\tau_1,\dots,\tau_\lambda)$ with strictly increasing indices

$$\begin{aligned}
\det\big(S[J\times K]\big) \ &= \sum_{\widehat{\sigma}_1=1}^{n_{j_0}-\sigma_1+1}\Big(m_{j_0}^{(\sigma_1+\widehat{\sigma}_1-1)}\cdots\sum_{\widehat{\sigma}_\kappa=1}^{n_{j_0}-\sigma_\kappa+1}\Big(m_{j_0}^{(\sigma_\kappa+\widehat{\sigma}_\kappa-1)} \\
&\sum_{\widehat{\tau}_1=1}^{n_{j_0}-\tau_1+1}\Big(\overline{m}_{j_0}^{(\tau_1+\widehat{\tau}_1-1)}\cdots\sum_{\widehat{\tau}_\lambda=1}^{n_{j_0}-\tau_\lambda+1}\Big(\overline{m}_{j_0}^{(\tau_\lambda+\widehat{\tau}_\lambda-1)} \quad \det\big(R[\widehat{J}\times\widehat{K}]\big)\dots\Big)\Big)\dots\Big)
\end{aligned} \tag{5.16}$$

with $\widehat{J} := (\widehat{\sigma}_1,\dots,\widehat{\sigma}_\kappa)$, $\widehat{K} := (\widehat{\tau}_1,\dots,\widehat{\tau}_\lambda)$, and

$$R = \begin{pmatrix} 0 & -T^{(j_0)} \\ (T^{(j_0)})' & f^{(j_0)}\otimes f^{(j_0)} \end{pmatrix}.$$

Recall that J,K contain strictly increasing indices. Moreover, $\det\big(R[\widehat{J}\times\widehat{K}]\big)=0$ whenever $\widehat{J}$ (or $\widehat{K}$) contains two coiniciding indices. Thus

$$\begin{array}{ll} \widehat{\sigma}_1<\ldots<\widehat{\sigma}_\kappa & \text{and } \widehat{\sigma}_k \text{ pairwise different for } k=1,\dots,\kappa, \\ \widehat{\tau}_1<\ldots<\widehat{\tau}_\lambda & \text{and } \widehat{\tau}_l \text{ pairwise different for } l=1,\dots,\lambda. \end{array} \tag{5.17}$$

Searching for the leading term in t of (5.16), we consult (5.14). This shows that in (5.16) the following powers of t occur:

$$
\begin{aligned}
&\sum_{k=1}^{\kappa}\big(\rho+n_{j_0}-(\sigma_k+\widehat{\sigma}_k-1)\big)+\sum_{l=1}^{\lambda}\big(\rho+n_{j_0}-(\tau_l+\widehat{\tau}_l-1)\big)\\
&\qquad = (\kappa+\lambda)\big(\rho+n_{j_0}+1\big)-\Big(\sum_{k=1}^{\kappa}(\sigma_k+\widehat{\sigma}_k)+\sum_{l=1}^{\lambda}(\tau_l+\widehat{\tau}_l)\Big),
\end{aligned}
$$

with the constraint $\sigma_\kappa+\widehat{\sigma}_\kappa,\ \tau_\lambda+\widehat{\tau}_\lambda\le n_{j_0}+1$ for $1\le\kappa\le k$, $1\le\lambda\le l$ due to the limitations of the summation in (5.16). Taking into account (5.17), this expression is maximized precisely by the choices

(i) $\sigma_k=k$ and $\widehat{\sigma}_k=\pi(k)$ for $\pi\in\mathrm{Perm}'(\kappa)$,

(ii) $\tau_l=l$ and $\widehat{\tau}_l=\chi(l)$ for $\chi\in\mathrm{Perm}'(\lambda)$.

Here $\mathrm{Perm}(\kappa)$ denotes the group of permutations of $\{1,\dots,\kappa\}$ and $\mathrm{Perm}'(\kappa)$ is the subset of those of the permutations $\pi\in\mathrm{Perm}(\kappa)$ which satisfy $k+\pi(k)\le n_{j_0}+1$ for all $k=1,\dots,\kappa$. Therefore, the leading term in t of (5.16) is given by

$$
\det\big(\ S[J\times K]\ \big)=H_{\kappa\lambda}(t)\Big[1+\mathcal{O}\Big(\frac{1}{t}\Big)\Big] \tag{5.18}
$$

with

$$
\begin{aligned}
H_{\kappa\lambda}(t) &= \sum_{\pi\in\mathrm{Perm}'(\kappa)}\ \sum_{\chi\in\mathrm{Perm}'(\lambda)}\Bigg(\prod_{k=1}^{\kappa} m_{j_0}^{(k+\pi(k)-1)}\prod_{l=1}^{\lambda}\overline{m}_{j_0}^{(l+\chi(l)-1)}\\
&\qquad\cdot\det\Big(R\big[\big(\pi(1),\dots,\pi(\kappa)\big)\times\big(\chi(1),\dots,\chi(\lambda)\big)\big]\Big)\Bigg).
\end{aligned} \tag{5.19}
$$

As for the calculation of $H_{\kappa\lambda}$, we first exploit (5.14) and use $\sum_{k=1}^{\kappa}(n_{j_0}-\big(k+\pi(k)-1\big))=(n_{j_0}+1)\kappa-2\sum_{k=1}^{\kappa}k=(n_{j_0}-\kappa)\kappa$, which yields

$$
\prod_{k=1}^{\kappa} m_{j_0}^{(k+\pi(k)-1)}=\frac{G_\kappa(t)\ t^{(n_{j_0}-\kappa)\kappa}|t|^{\kappa\rho}}{\prod_{k=1}^{\kappa}\Big(n_{j_0}+1-\big(k+\pi(k)\big)\Big)!}\quad\Big[1+\mathcal{O}\Big(\frac{\log|t|}{t}\Big)\Big]
$$

$$
\text{with } G_\kappa(t)=\Big(b_{j_0}^{(n_{j_0})}(c_{j_0})^{n_{j_0}-\kappa}\exp\big(\ \mathrm{i}\Upsilon_{j_0}(\gamma_\rho(t),t)\ \big)\ \Big)^{\kappa}
$$

(note that $G_\kappa(t)$ does not contribute to the growth in t since $|G_\kappa(t)|$ is constant). Inserting this into (5.19), we get

$$
H_{\kappa\lambda}(t)\ =\ D_{\kappa\lambda}\ G_\kappa(t)\overline{G_\lambda(t)}\quad |t|^{(\kappa+\lambda)\rho}\ t^{\kappa(n_{j_0}-\kappa)+\lambda(n_{j_0}-\lambda)}\quad\Big[1+\mathcal{O}\Big(\frac{\log|t|}{t}\Big)\Big], \tag{5.20}
$$

where

$$
\begin{aligned}
D_{\kappa\lambda} &= \sum_{\substack{\pi\in\mathrm{Perm}'(\kappa)\\ \chi\in\mathrm{Perm}'(\lambda)}}\prod_{k=1}^{\kappa}\frac{1}{\big(n_{j_0}+1-\big(k+\pi(k)\big)\big)!}\prod_{l=1}^{\lambda}\frac{1}{\big(n_{j_0}+1-\big(l+\chi(l)\big)\big)!}\\
&\qquad\cdot\det\Big(R\big[\big(\pi(1),\dots,\pi(\kappa)\big)\times\big(\chi(1),\dots,\chi(\lambda)\big)\big]\Big).
\end{aligned}
$$

We can drop the restriction on the permutations $\pi\in\mathrm{Perm}'(\kappa)$, $\chi\in\mathrm{Perm}'(\lambda)$ if we pass from the faculties to the quantities

$$
\mathrm{fac}_{\nu\mu}=\begin{cases}\dfrac{1}{(n_{j_0}+1-(\mu+\nu))!}, & \mu+\nu\le n_{j_0}+1,\\ 0, & \mu+\nu>n_{j_0}+1.\end{cases}.
$$

Thus, by Corollary 5.1.12,

$$\begin{aligned}
D_{\kappa\lambda} &= \sum_{\substack{\pi\in\mathrm{Perm}(\kappa)\\ \chi\in\mathrm{Perm}(\lambda)}} \prod_{k=1}^{\kappa} \mathrm{fac}_{k\pi(k)} \prod_{l=1}^{\lambda} \mathrm{fac}_{l\chi(l)} \ \det\Big(R\big[(\pi(1),\ldots,\pi(\kappa))\times(\chi(1),\ldots,\chi(\lambda))\big]\Big)\\
&= \det\Big(R\big[(1,\ldots,\kappa)\times(1,\ldots,\lambda)\big]\Big)\\
&\qquad\cdot \sum_{\pi\in\mathrm{Perm}(\kappa)} \Big(\mathrm{sgn}(\pi)\prod_{k=1}^{\kappa}\mathrm{fac}_{k\pi(k)}\Big) \sum_{\chi\in\mathrm{Perm}(\lambda)} \Big(\mathrm{sgn}(\chi)\prod_{l=1}^{\lambda}\mathrm{fac}_{l\chi(l)}\Big)\\
&= \det\big(\,\mathrm{fac}_{\mu\nu}\,\big)_{\mu,\nu=1}^{\kappa}\ \det\big(\,\mathrm{fac}_{\mu\nu}\,\big)_{\mu,\nu=1}^{\lambda}\ \det\Big(R\big[(1,\ldots,\kappa)\times(1,\ldots,\lambda)\big]\Big)\\
&= (-1)^{\kappa(\kappa+3)/2}\frac{\prod_{k=1}^{\kappa-1}k!}{\prod_{k=1}^{\kappa}(n_{j_0}-k)!}\quad (-1)^{\lambda(\lambda+3)/2}\frac{\prod_{l=1}^{\lambda-1}l!}{\prod_{l=1}^{\lambda}(n_{j_0}-l)!}\\
&\qquad\cdot\det\Big(R\big[(1,\ldots,\kappa)\times(1,\ldots,\lambda)\big]\Big).
\end{aligned}$$

Therefore,

$$\begin{aligned}
&D_{\kappa\lambda}G_\kappa(t)\overline{G_\lambda(t)} =\\
&= (-1)^{\frac{\kappa(\kappa+3)}{2}}(-1)^{\frac{\lambda(\lambda+3)}{2}}F_\kappa(t)\overline{F_\lambda(t)}\ \det\Big(R\big[(1,\ldots,\kappa)\times(1,\ldots,\lambda)\big]\Big).
\end{aligned}\tag{5.21}$$

As a consequence of (5.18), (5.20), and (5.21), we infer

$$\begin{aligned}
\det\big(S[J\times K]\,\big) = {}& (-1)^{\frac{\kappa(\kappa+3)}{2}}(-1)^{\frac{\lambda(\lambda+3)}{2}}\det\Big(R\big[(1,\ldots,\kappa)\times(1,\ldots,\lambda)\big]\Big)\\
&\cdot F_\kappa(t)\overline{F_\lambda(t)}\ \ |t|^{(\kappa+\lambda)\rho}\ t^{\kappa(n_{j_0}-\kappa)+\lambda(n_{j_0}-\lambda)}\quad \Big[1+\mathcal{O}\Big(\frac{\log|t|}{t}\Big)\Big].
\end{aligned}\tag{5.22}$$

Finally, we state

$$\det\Big(R\big[(1,\ldots,\kappa)\times(1,\ldots,\lambda)\big]\Big) = CC_{\kappa\lambda}\qquad \text{for } \lambda=\kappa,\kappa+1.\tag{5.23}$$

Actually the proof of (5.23) is very involved. We refer to Theorem 6.2.1 in Chapter 6.

Inserting (5.22), (5.23) into (5.15), and using $(-1)^{\kappa^2}=(-1)^{\kappa}$ yields Proposition 5.1.9. □

Now we are in position to give the asymptotic behaviour of the j_0-th negaton.

Proposition 5.1.13. $q^{(j_0)}(x,t) \approx \sum_{j_0{}'=0}^{n_{j_0}-1} q^{(j_0j_0')}(x,t)$ *for* $t\approx-\infty$

with

$$q^{(j_0j_0')}(x,t) = (\alpha_{j_0}+\overline{\alpha}_{j_0})\,\frac{(-1)^{j_0'+1}\,\epsilon^{J_0'}\ \exp\Big(\overline{\Gamma^{-}_{j_0j_0'}(x,t)}\Big)}{1+\exp\Big(\Gamma^{-}_{j_0j_0'}(x,t)+\overline{\Gamma^{-}_{j_0j_0'}(x,t)}\Big)}\tag{5.24}$$

where $\epsilon=\mathrm{sgn}(t)$ *and the function* $\Gamma^{-}_{j_0j_0'}=\Gamma^{-}_{j_0j_0'}(x,t)$ *is defined by*

$$\Gamma^{-}_{j_0j_0'}(x,t) = \alpha_{j_0}x + f_0(\alpha_{j_0})t - J_0'\log|t| + \varphi_{j_0} + \varphi^{-}_{j_0} - \varphi^{-}_{j_0j_0'}$$

for the index $J_0' = -(n_{j_0}-1)+2j_0'$ *associated to* j_0'*, and the phase* φ_{j_0}*, and the phase-shifts* $\varphi^{-}_{j_0}$*,* $\varphi^{-}_{j_0j_0'}$ *as defined in Theorem 5.1.2.*

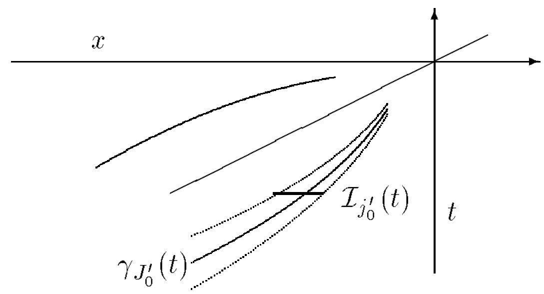
Remark 5.1.14. *In (5.24) we did not include the sign $\epsilon^{J_0'}$ in the power of (-1) in order to stress the symmetry between the cases $t \approx -\infty$, $t \approx +\infty$.*

Proof Without loss of generality, $t < -1$. (This assumption is only needed to guarantee that the intervals defined below are proper).

Recall from Proposition 5.1.9 that the curve $\big(t, \gamma_\rho(t)\big)$ for $\rho \in \mathbb{R}$ is defined by

$$\gamma_\rho(t) = v_{j_0} t + \rho \, \log|t|/\mathrm{Re}(\alpha_{j_0}).$$

Fix j_0' and set $J_0' = -(n_{j_0} - 1) + 2j_0'$.
Let $x \in \mathcal{I}_{j_0'}(t) = \big(\gamma_{J_0'-\frac{1}{2}}(t), \gamma_{J_0'+\frac{1}{2}}(t)\big)$.
This interval $\mathcal{I}_{j_0'}(t)$ has the center $\gamma_{J_0'}(t) = v_{j_0} t + J_0' \log|t|/\mathrm{Re}(\alpha_{j_0})$, and its diameter grows logarithmically with t.
Moreover, we define $\mathcal{C}_{j_0'} = \bigcup_{t \le -1} \mathcal{I}_{j_0'}(t)$.



We subdivide our task in three steps:

Step a: First we show that, in the asymptotic sense, the only contribution to the j_0-th negaton $q^{(j_0)}$ in $\mathcal{C}_{j_0'}$ is due to the soliton $q^{(j_0 j_0')}$.

Step b: Next we show that outside $\mathcal{C}_{j_0'}$ the soliton $q^{(j_0 j_0')}$ asymptotically vanishes.

Step c: At last we show that the j_0-th negaton $q^{(j_0)}$ vanishes asymptotically outside $\bigcup_{j_0'=0}^{n_{j_0}-1} \mathcal{C}_{j_0'}$.

To avoid confusion we point out that we are not going to show that the maximum of the j_0-th soliton moves along the center $\gamma_{J_0'}(t)$ of the interval $\mathcal{I}_{j_0'}(t)$, but on a curve parallel to $\big(\gamma_{J_0'}(t), t\big)$ where the distance between the curves is determined by the term $\varphi_{j_0} + \varphi_{j_0}^- - \varphi_{j_0 j_0'}^-$. Note that there is a certain flexibility in the choice of the $\mathcal{I}_{j_0'}(t)$. For example, they could be replaced by their translates $\mathcal{I}_{j_0'}(t) + t_0$ by a fixed but arbitrary $t_0 \in \mathbb{R}$.

Step a: Parametrize $x = \gamma_{J_0'+\rho}(t)$ by $\rho \in (-\frac{1}{2}, \frac{1}{2})$.

By Proposition 5.1.9, the order of $P^{(j_0)}\big(\gamma_{J_0'+\rho}(t), t\big)$, $p^{(j_0)}\big(\gamma_{J_0'+\rho}(t), t\big)$ in t is determined by the exponents

$$\begin{aligned} f_\rho(\kappa) &= -2\kappa^2 + 2\big(n_{j_0} + (J_0' + \rho)\big)\kappa \\ &= -2\kappa^2 + 2\big((2j_0' + 1) + \rho\big)\kappa \qquad \text{for} \quad \kappa \in \{0, \ldots, n_{j_0}\} \end{aligned} \tag{5.25}$$

and

$$\begin{aligned} g_\rho(\kappa) &= -2\kappa^2 + 2\big(n_{j_0} - 1 + (J_0' + \rho)\big)\kappa + \big(n_{j_0} - 1 + (J_0' + \rho)\big) \\ &= -2\kappa^2 + 2(2j_0' + \rho)\kappa + (2j_0' + \rho) \qquad \text{for} \quad \kappa \in \{0, \ldots, n_{j_0} - 1\}. \end{aligned} \tag{5.26}$$

Again we search for the exponents which maximize the order of t.

Regarding κ as a continuous variable for the moment, the function f_ρ is a concave parabola. It attains its global maximum in $\kappa_0(\rho) = (j_0' + \frac{1}{2}) + \rho/2$. Since $|\rho| < \frac{1}{2}$, candidates for the dominating exponents correspond to the indices $\kappa = j_0', j_0' + 1$. Moreover, the other exponents do not disturb the order estimation, since the minimal distance to them is

$$\begin{aligned} &\min\Big\{f_\rho(j_0'), f_\rho(j_0' + 1)\Big\} - \max\Big\{f_\rho(j_0' - 1), f_\rho(j_0' + 2)\Big\} = \\ &= \left\{ \begin{array}{rcll} f_\rho(j_0' + 1) - f_\rho(j_0' - 1) &=& 4(1 + \rho), & \text{if } \rho \le 0 \\ f_\rho(j_0') - f_\rho(j_0' + 2) &=& 4(1 - \rho), & \text{if } \rho \ge 0 \end{array} \right. \\ &> 2. \end{aligned}$$

Analogously, g_ρ attains its maximum for $\kappa_0(\rho) = j_0' + \rho/2$ what shows that $\kappa = j_0'$ maximizes the concerning exponents. Here the minimal distance to the other exponents can be estimated from below by $g_\rho(j_0) = \max\{g_\rho(j_0'-1), g_\rho(j_0'+1)\} > 1$.

Keeping only those terms in Proposition 5.1.9 (note that $\gamma_{\rho+J_0'}(t)$ has to be inserted), we get for $q^{(j_0)} = 1 - P^{(j_0)}/p^{(j_0)} = \big(p^{(j_0)} - P^{(j_0)}\big)/p^{(j_0)}$

$$\begin{aligned}
&q^{(j_0)}\big(\gamma_{J_0'+\rho}(t), t\big) \;= \\
&= \frac{\Big(-(-1)^{j_0'}\, C_{j_0'(j_0'+1)}\; F_{j_0'}(t)\overline{F_{j_0'+1}(t)}\;\; |t|^{\rho+J_0'} t^{-J_0'}\Big)\;\Big[1+\mathcal{O}\Big(\tfrac{\log|t|}{t}\Big)\Big]}{\Big(C_{j_0'j_0'}\; F_{j_0'}(t)\overline{F_{j_0'}(t)} + C_{(j_0'+1)(j_0'+1)}\; F_{j_0'+1}(t)\overline{F_{j_0'+1}(t)}\;\; |t|^{2(\rho+J_0')} t^{-2J_0'}\Big)\;\Big[1+\mathcal{O}\Big(\tfrac{\log|t|}{t}\Big)\Big]} \\
&= \frac{(-1)^{j_0'+1}\epsilon^{J_0'}\; C_{j_0'(j_0'+1)}\; F_{j_0'}(t)\overline{F_{j_0'+1}(t)}\;\; |t|^{\rho}}{C_{j_0'j_0'}\; F_{j_0'}(t)\overline{F_{j_0'}(t)} + C_{(j_0'+1)(j_0'+1)}\; F_{j_0'+1}(t)\overline{F_{j_0'+1}(t)}\;\; |t|^{2\rho}} \;+\; \mathcal{O}\Big(\frac{\log|t|}{t}\Big), \qquad (5.27)
\end{aligned}$$

where the sign ϵ is defined by $t = \epsilon|t|$.

Recall the definitions of $C_{\kappa\lambda}$, $F_\kappa(t)$ in Proposition 5.1.9. With regard to

$$d_{j_0} = (\alpha_{j_0} + \overline{\alpha}_{j_0})c_{j_0}, \qquad \exp(\varphi_{j_0}) = b_{j_0}^{(n_{j_0})}/(\alpha_{j_0} + \overline{\alpha}_{j_0})^{n_{j_0}}$$

and the definitions of the phase-shifts $\varphi_{j_0}^-$, $\varphi_{j_0j_0'}^-$ in Theorem 5.1.2, these read

$$C_{\kappa\lambda} = \left(\frac{1}{\alpha_{j_0} + \overline{\alpha}_{j_0}}\right)^{2\kappa\lambda} \Big(\exp(\varphi_{j_0}^-)\Big)^{\kappa}\Big(\overline{\exp(\varphi_{j_0}^-)}\Big)^{\lambda}$$

and

$$\begin{aligned}
F_{\kappa+1}(t)/F_\kappa(t) \;&=\; \frac{\kappa!}{(n_{j_0}-\kappa-1)!} c_{j_0}^{n_{j_0}-2\kappa-1}\;\; b_{j_0}^{(n_{j_0})} \exp\big(\, \mathrm{i}\Upsilon^{(j_0)}(\gamma_{J_0'+\rho}(t), t)\,\big) \\
&=\; \frac{\kappa!}{(\kappa-K)!} c_{j_0}^{-K}\;\; (\alpha_{j_0}+\overline{\alpha}_{j_0})^{n_{j_0}}\;\; \exp\big(\, \mathrm{i}\Upsilon^{(j_0)}(\gamma_{J_0'+\rho}(t), t) + \varphi_{j_0}\,\big) \\
&=\; (\alpha_{j_0}+\overline{\alpha}_{j_0})^{2\kappa+1}\;\; \frac{\kappa!}{(\kappa-K)!} d_{j_0}^{-K}\;\; \exp\big(\, \mathrm{i}\Upsilon^{(j_0)}(\gamma_{J_0'+\rho}(t), t) + \varphi_{j_0}\,\big) \\
&=\; (\alpha_{j_0}+\overline{\alpha}_{j_0})^{2\kappa+1}\;\; \exp\big(\, \mathrm{i}\Upsilon^{(j_0)}(\gamma_{J_0'+\rho}(t), t) + \varphi_{j_0} - \varphi_{j_0\kappa}^-\,\big)\,.
\end{aligned}$$

for $K = -(n_{j_0}-1) + 2\kappa$, the index associated to κ.

Inserting these into (5.27) and reordering a bit, we end up with

$$q^{(j_0)}\big(\gamma_{J_0'+\rho}(t), t\big) \;=\; (\alpha_{j_0}+\overline{\alpha}_{j_0}) \frac{\overline{\mathcal{P}^{(j_0j_0')}(t)}\;\; |t|^{\rho}}{1 + \mathcal{P}^{(j_0j_0')}(t)\overline{\mathcal{P}^{(j_0j_0')}(t)}\;\; |t|^{2\rho}} \;+\; \mathcal{O}\Big(\frac{\log|t|}{t}\Big), \qquad (5.28)$$

$$\text{for } \mathcal{P}^{(j_0j_0')}(t) = (-1)^{j_0'+1}\epsilon^{J_0'} \exp\Big(\; \mathrm{i}\Upsilon^{(j_0)}(\gamma_{J_0'+\rho}(t), t) + \varphi_{j_0} + \varphi_{j_0}^- - \varphi_{j_0j_0'}^- \;\Big)\,. \quad (5.29)$$

Our aim is to compare $q^{(j_0)}$ with $q^{(j_0j_0')}$ on $\mathcal{I}_{j_0'}(t)$. To this end, we also have to calculate $q^{(j_0j_0')}\big(\gamma_{J_0'+\rho}(t), t\big)$. From

$$\begin{aligned}
&\Gamma_{j_0j_0'}^-\big(\gamma_{J_0'+\rho}(t), t\big) \;= \\
&=\; \mathrm{Re}(\alpha_{j_0})\big[\gamma_{J_0'+\rho}(t) - v_{j_0}t\big] - J_0'\log|t| + \mathrm{i}\Upsilon\big(\gamma_{J_0'+\rho}(t), t\big) + \varphi_{j_0} + \varphi_{j_0}^- - \varphi_{j_0j_0'}^- \\
&=\; \rho\log|t| + \mathrm{i}\Upsilon\big(\gamma_{J_0'+\rho}(t), t\big) + \varphi_{j_0} + \varphi_{j_0}^- - \varphi_{j_0j_0'}^-
\end{aligned}$$

it is straightforward to check that

$$q^{(j_0j_0')}\big(\gamma_{J_0'+\rho}(t), t\big) = (\alpha_{j_0}+\overline{\alpha}_{j_0}) \frac{\overline{\mathcal{P}^{(j_0j_0')}(t)}\;\; |t|^{\rho}}{1 + \mathcal{P}^{(j_0j_0')}(t)\overline{\mathcal{P}^{(j_0j_0')}(t)}\;\; |t|^{2\rho}}. \qquad (5.30)$$

As a consequence of (5.28) and (5.30), we may conclude

$$\sup_{x\in\mathcal{I}_{j_0'}(t)}\left|q^{(j_0)}(x,t)-q^{(j_0j_0')}(x,t)\right|=\mathcal{O}\Big(\frac{\log|t|}{t}\Big),$$

and hence the supremum converges to zero as $t\to-\infty$.

In summary, we have shown that $q^{(j_0)}$ asymptotically behaves like the soliton $q^{(j_0j_0')}$ on $\mathcal{C}_{j_0'}$ as t tends to $-\infty$. More precisely, we have $1_{\mathcal{C}_{j_0'}}q^{(j_0)}\approx 1_{\mathcal{C}_{j_0'}}q^{(j_0j_0')}$ for $t\approx-\infty$, where $1_{\mathcal{C}_{j_0'}}$ is the characteristic function of $\mathcal{C}_{j_0'}$. Step a is complete.

Step b: Next, we discuss the behaviour of the soliton $q^{(j_0j_0')}$ outside of $\mathcal{I}_{j_0'}(t)$. We distinguish two cases:

(i) Let $x\in\mathcal{I}^-_{j_0'}(t)=\big(-\infty,\gamma_{J_0'-\frac{1}{2}}(t)\,\big]$, say $x=\gamma_{J_0'+\rho}(t)$ with $\rho\le-\frac{1}{2}$.

First recall that $q^{(j_0j_0')}\big(\gamma_{J_0'+\rho}(t),t\big)$ has been determined in (5.30) with $\mathcal{P}^{(j_0j_0')}(t)$ as defined in (5.29). Note that the modulus of $\mathcal{P}^{(j_0j_0')}(t)$ does not depend on t. Set $c=|\mathcal{P}^{(j_0j_0')}(t)|$. Then, by (5.30), for $|t|>c^2$,

$$\begin{aligned}\left|q^{(j_0j_0')}\big(\gamma_{J_0'+\rho}(t),t\big)\right| &\le |\alpha_{j_0}+\overline{\alpha}_{j_0}|\frac{c|t|^{\rho}}{1-c^2|t|^{2\rho}}\\ &\le |\alpha_{j_0}+\overline{\alpha}_{j_0}|\frac{c|t|^{-\frac{1}{2}}}{1-c^2|t|^{-1}}\;=\;\mathcal{O}\big(t^{-\frac{1}{2}}\big)\qquad\forall\rho\le-\frac{1}{2}.\end{aligned}$$

(ii) Let $x\in\mathcal{I}^+_{j_0'}(t)=\big[\,\gamma_{J_0'+\frac{1}{2}}(t),\infty\,\big)$, say $x=\gamma_{J_0'+\rho}(t)$ with $\rho\ge\frac{1}{2}$.

Here we rewrite (5.30) in the following way:

$$q^{(j_0j_0')}\big(\gamma_{J_0'+\rho}(t),t\big)=-(\alpha_{j_0}+\overline{\alpha}_{j_0})\frac{\big(\,\mathcal{P}^{(j_0j_0')}(t)\,\big)^{-1}\,|t|^{-\rho}}{1+\big(\,\mathcal{P}^{(j_0j_0')}(t)\overline{\mathcal{P}^{(j_0j_0')}(t)}\,\big)^{-1}\,|t|^{-2\rho}}.$$

Now the same argument as in (i) yields $\left|q^{(j_0j_0')}\big(\gamma_{J_0'+\rho}(t),t\big)\right|=\mathcal{O}\big(t^{-\frac{1}{2}}\big)$ for all $\rho\ge\frac{1}{2}$.

Consequently,

$$\sup_{x\in\mathcal{I}^-_{j_0'}(t)\cup\mathcal{I}^+_{j_0'}(t)}\left|q^{(j_0j_0')}(x,t)\right|\to 0\qquad\text{as } t\to-\infty.$$

Thus $q^{(j_0j_0')}$ vanishes on $\mathbb{R}^2\backslash\mathcal{C}_{j_0'}$ as t tends to $-\infty$. More precisely, $(1-1_{\mathcal{C}_{j_0'}})q^{(j_0j_0')}\approx 0$ for $t\approx-\infty$. Step b is complete.

Step c: It remains to check that $q^{(j_0)}$ asymptotically vanishes outside $\cup_{j_0'=0}^{n_{j_0}-1}\mathcal{C}_{j_0'}$ as $t\to-\infty$. This is done in the remainder of the proof.

Consider $x\in\widetilde{\mathcal{I}}_{j_0'}(t)=\big[\gamma_{J_0'+\frac{1}{2}}(t),\gamma_{J_0'+\frac{3}{2}}(t)\big]$. It is clear that, if J_0' is the index associated to j_0' according to Theorem 5.1.2, then $J_0'+2$ is the index associated to $j_0'+1$. Thus, $\widetilde{\mathcal{I}}_{j_0'}(t)$ is just the interval covering the gap between the two neighbouring strips $\mathcal{I}_{j_0'}(t)$ and $\mathcal{I}_{j_0'+1}(t)$. Parametrize $x=\gamma_{J_0'+\rho}(t)$ with $\rho\in\big[\frac{1}{2},\frac{3}{2}\big]$.

As before we investigate the behaviour of $P^{(j_0)}\big(\gamma_\rho(t),t\big)$, $p^{(j_0)}\big(\gamma_\rho(t),t\big)$ with the aim to find the maximal contribution to the order of t. The latter is again determined by the exponents $f_\rho(\kappa)$, $g_\rho(\kappa)$ defined in (5.25), (5.26).

Recall that f_ρ attains its maximum for $\kappa_0(\rho) = (j_0 + \frac{1}{2}) + \rho/2$. Since $\frac{1}{2} \le \rho \le \frac{3}{2}$, the only candidate for the dominating exponent is $\kappa = j_0' + 1$. As for the minimal distance to the other exponents, we observe

$$
\begin{aligned}
& f_\rho(j_0'+1) - \max\Big\{ f_\rho(j_0'), f_\rho(j_0'+2) \Big\} = \\
= & \left\{ \begin{array}{rcll} f_\rho(j_0'+1) - f_\rho(j_0') & = & 2\rho, & \text{if } \rho \le 1 \\ f_\rho(j_0'+1) - f_\rho(j_0'+2) & = & 2(2-\rho), & \text{if } \rho \ge 1 \end{array} \right. \\
\ge & \; 1.
\end{aligned}
$$

The maximum of g_ρ is attained for $\kappa_0(\rho) = j_0' + \rho/2$ and thus $\kappa = j_0', j_0'+1$ are the candidates to maximize the concerning exponents. Here the minimal distance to the other exponents can be estimated from below by

$$
\begin{aligned}
& \min\Big\{ g_\rho(j_0'), g_\rho(j_0'+1) \Big\} - \max\Big\{ g_\rho(j_0'-1), g_\rho(j_0'+2) \Big\} = \\
= & \left\{ \begin{array}{rcll} g_\rho(j_0'+1) - g_\rho(j_0'-1) & = & 4\rho, & \text{if } \rho \le 1 \\ g_\rho(j_0') - g_\rho(j_0'+2) & = & 4(2-\rho), & \text{if } \rho \ge 1 \end{array} \right. \\
\ge & \; 2.
\end{aligned}
$$

Keeping only those terms in Proposition 5.1.9, we get

$$
q^{(j_0)}\big(\gamma_{J_0'+\rho}(t), t\big) = \frac{\Big(G_{j_0'}(t) + G_{j_0'+1}(t)\; |t|^{2(J_0'+\rho)} t^{-2(J_0'+1)} \Big) \; \Big[1 + \mathcal{O}\Big(\frac{\log|t|}{t} \Big) \Big]}{\Big(\widehat{G}_{j_0'+1}(t)\; |t|^{(J_0'+\rho)} t^{-J_0'} \Big) \; \Big[1 + \mathcal{O}\Big(\frac{\log|t|}{t} \Big) \Big]}
$$

for $G_{j_0'}(t) = (-1)^{j_0'+1} C_{j_0'(j_0'+1)} F_{j_0'}(t) \overline{F_{j_0'+1}(t)}$ and $\widehat{G}_{j_0'}(t) = C_{j_0'j_0'} F_{j_0'}(t) \overline{F_{j_0'}(t)}$.

Recall that the moduli of $G_{j_0'}(t)$ and $\widehat{G}_{j_0'}(t)$ do not depend on t, say $|G_{j_0'}(t)| = c_{j_0'}$, $|\widehat{G}_{j_0'}(t)| = \widehat{c}_{j_0'}$. Then

$$
\Big| q^{(j_0)}\big(\gamma_\rho(t), t\big) \Big| \le \frac{1}{\widehat{c}_{j_0'+1}} \Big(c_{j_0'}\; |t|^{-\rho} + c_{j_0'+1}\; |t|^{\rho-2} \Big) \Big[1 + \mathcal{O}\Big(\frac{\log|t|}{t} \Big) \Big] = \mathcal{O}\Big(t^{-\frac{1}{2}} \Big)
$$

since $\frac{1}{2} \le \rho \le \frac{3}{2}$. This shows

$$
\sup_{x \in \widetilde{\mathcal{I}}_{j_0'}(t)} \Big| q^{(j_0)}(x,t) \Big| \longrightarrow 0 \qquad \text{as } t \to -\infty
$$

for all $j_0' = 0, \ldots, n_j - 1$.

We still have to consider the boundary regions. To this end we keep the notations $G_{j_0}(t)$, $\widehat{G}_{j_0}(t)$ from above and proceed as before.

(i) Let $x \in \widetilde{\mathcal{I}}_{\min}(t) = \big(-\infty, \gamma_{(-(n_{j_0}-1)-\frac{1}{2})}(t) \big]$. Note that, for $j_0' = 0$, $J_0' = -(n_{j_0} - 1)$, the associated interval $\mathcal{I}_{j_0'}(t)$ is the one which is the farthest to the left. Thus $\widetilde{\mathcal{I}}_{\min}(t)$ closes the gap between $-\infty$ and $\mathcal{I}_{j_0'}(t)$.

Parametrize $x = \gamma_{(-(n_{j_0}-1)+\rho)}(t)$ with $\rho \in \big(-\infty, -\frac{1}{2} \big]$.

Recall that the functions f_ρ, g_ρ responsible for the maximal contributions to the order of t are given by (5.25), (5.26). Candidates for the dominating exponents are 0 in either case. The minimal distance to the other exponents can be estimated from below by

$$
\begin{array}{rcrcl} f_\rho(0) - f_\rho(1) & = & -2\rho & \ge & 1, \\ g_\rho(0) - g_\rho(1) & = & -2(\rho-1) & \ge & 3. \end{array}
$$

If we keep only these terms in Proposition 5.1.9, we get

$$q^{(j_0)}\big(\gamma_{(-(n_{j_0}-1)+\rho)}(t),t\big) = \frac{G_0(t)\ |t|^{(-(n_{j_0}-1)+\rho)} t^{n_{j_0}-1}\ \Big[1+\mathcal{O}\Big(\frac{\log|t|}{t}\Big)\Big]}{\widehat{G}_0(t)\ \Big[1+\mathcal{O}\Big(\frac{\log|t|}{t}\Big)\Big]},$$

showing $|q^{(j_0)}\big(\gamma_{(-(n_{j_0}-1)+\rho)}(t),t\big)| = \mathcal{O}\big(t^{-\frac{1}{2}}\big)$.

(ii) Let $x \in \widetilde{\mathcal{I}}_{\max}(t) = \big[\gamma_{((n_{j_0}-1)+\frac{1}{2})}(t),\infty\big)$. Since the interval $\mathcal{I}_{j_0'}(t)$ corresponding to $j_0' = (n_{j_0}-1)$, $J_0' = (n_{j_0}-1)$, is the farthest to the right, this covers the gap between $\mathcal{I}_{j_0'}(t)$ and $+\infty$.

Parametrize $x = \gamma_{((n_{j_0}-1)+\rho)}(t)$ with $\rho \in \big[\frac{1}{2},\infty\big)$.

Again the functions f_ρ, g_ρ which are responsible for the maximal contributions to the order of t are given by (5.25), (5.26). With regard to the different sets of admitted indices for f_ρ and g_ρ, here the candidates for the dominating exponents are n_{j_0} for f_ρ and $n_{j_0}-1$ for g_ρ. The minimal distance to the other exponents is

$$\begin{array}{rcccl} f_\rho(n_{j_0}) - f_\rho(n_{j_0}-1) & = & 2\rho & \geq & 1, \\ g_\rho(n_{j_0}-1) - g_\rho(n_{j_0}-2) & = & 2(\rho+1) & \geq & 3. \end{array}$$

Thus

$$q^{(j_0)}\big(\gamma_{((n_{j_0}-1)+\rho)}(t),t\big) = \frac{G_{n_{j_0}-1}(t)\ |t|^{(2n_{j_0}-1)(\rho+(n+j_0-1))} t^{n_{j_0}-1}\ \Big[1+\mathcal{O}\Big(\frac{\log|t|}{t}\Big)\Big]}{\widehat{G}_{n_{j_0}}(t)\ \ |t|^{2n_{j_0}(\rho+(n+j_0-1))}\ \Big[1+\mathcal{O}\Big(\frac{\log|t|}{t}\Big)\Big]},$$

showing again $|q^{(j_0)}\big(\gamma_{((n_{j_0}-1)+\rho)}(t),t\big)| = \mathcal{O}\big(t^{-\frac{1}{2}}\big)$.

Thus $q^{(j_0)}$ vanishes on $\mathbb{R}^2 \backslash \cup_{j_0=0}^{n_{j_0}-1} \mathcal{C}_{j_0}$ as t tends to $-\infty$. More precisely, $(1-\sum_{j_0'=0}^{n_{j_0}-1} 1_{\mathcal{C}_{j_0'}})q^{(j_0)} \approx 0$ for $t \approx -\infty$. Step c is complete.

In summary, Proposition 5.1.13 is proved. □

Proof (of Theorem 5.1.2) For the proof of Theorem 5.1.2, we collect what we have achieved so far. By Proposition 5.1.6 and Proposition 5.1.13,

$$q(x,t) \approx \sum_{j=1}^{N} \sum_{j'=0}^{n_j-1} q^{(jj')}(x,t) \qquad \text{for } t \approx -\infty$$

with

$$\begin{aligned} q^{(jj')}(x,t) \ &= \ (\alpha_j + \overline{\alpha}_j) \frac{(-1)^{j'+1}\, \epsilon^{J'}\ \exp\big(\overline{\Gamma^-_{jj'}(x,t)}\big)}{1+\exp\big(\Gamma^-_{jj'}(x,t)\big)\exp\big(\overline{\Gamma^-_{jj'}(x,t)}\big)} \\ &= \ \ (-1)^{j'+1}\, \epsilon^{J'}\ \ \mathrm{Re}(\alpha_j)\exp\Big(-\mathrm{i}\,\mathrm{Im}\big(\Gamma^-_{jj'}(x,t)\big)\Big)\cosh^{-1}\Big(\mathrm{Re}\big(\Gamma^-_{jj'}(x,t)\big)\Big) \end{aligned}$$

and

$$\Gamma^-_{jj'}(x,t) = \alpha_j x + f_0(\alpha_j)t - J'\log|t| + \varphi_j + \varphi_j^- - \varphi_{jj'}^-.$$

To obtain the assertion for $t \approx -\infty$, we have to integrate the sign ϵ. This is done by reordering the solitons in the following manner.

Replace j' by $k' = (n_j - 1) - j'$, and let $K' = -(n_j - 1) + 2k'$ be the index associated to k' according to Theorem 5.1.2. Thus,

$$J' = -K', \qquad j' = k' - K'.$$

Note that this replacement just reverses the order of the solitons $q^{(jj')}$ within the j-th negaton. In particular, $0 \leq k' \leq n_j - 1$. Moreover, we check

$$\exp\big(- \varphi_{jj'}^{-}\big) = \frac{j'!}{(j'-J')!}\, d_j^{-J'} = \frac{(k'-K')!}{k'!}\, d_j^{K'} = \exp\big(+ \varphi_{jk'}^{-}\big).$$

Thus, the reordering results in

$$q^{(jk')}(x,t) = (-1)^{k'+1}\ \mathrm{Re}(\alpha_j) \exp\Big(- \mathrm{i}\, \mathrm{Im}\big(\Gamma_{jk'}^{-}(x,t)\big)\Big) \cosh^{-1}\Big(\mathrm{Re}\big(\Gamma_{jk'}^{-}(x,t)\big)\Big)$$

with

$$\Gamma_{jk'}^{-}(x,t) = \alpha_j x + f_0(\alpha_j)t + K' \log|t| + \varphi_j + \varphi_j^{-} + \varphi_{jk'}^{-}.$$

This completes the proof in the case $t \approx -\infty$.

As for the asymptotic result in the case $t \approx +\infty$, the proof is completely the same as in the case $t \approx -\infty$ with the following two modifications:

(i) In Proposition 5.1.6 the index set $\Lambda_{j_0}^{-}$ is to be replaced by $\Lambda_{j_0}^{+}$.

(ii) The reordering procedure above can be skipped. Here it is not necessary because the sign ϵ produces no effect. Since $\exp\big(- \varphi_{jj'}^{-}\big) = \exp\big(+ \varphi_{jj'}^{+}\big)$, the assertion in this case follows immediately.

This completes the proof of Theorem 5.1.2. □

Remark 5.1.15. *Imposing assumption (ii) in Theorem 5.1.2 we excluded the degenerate case $d_j = 0$ for some j. In this case the estimates (5.14) in the proof of Proposition 5.1.9 are not valid anymore. Instead, contributions to the leading order of t now are due to higher derivatives of f_0. However, for all relevant equations (ii) is automatically satisfied.*

5.2 Negatons of the $\mathbb{R}$-reduction – How to include breathers

In this section we will see that the $\mathbb{R}$-reduction of the AKNS system allows to include a new element in the context of negatons. In this case there are two natural choices leading to real solutions: a) real eigenvalues and b) pairs of complex conjugated eigenvalues. The former correspond to the usual bell-shaped solitons, the latter to the famous breathers, a particularly interesting class of solutions.

In [69], breathers are interpreted as bound states of two solitons. Thus they are the lowest-dimensional example of a formation of solitons. In general, a formation is a group of solitons moving with exactly the same velocity, which hence cannot be separated in terms of the asymptotic analysis. Therefore, breathers were not included in our treatment so far. The aim of this section is to explain how the asymptotic analysis can be extended to negatons consisting of breathers.

We are only interested in real solutions of the $\mathbb{R}$-reduced AKNS-system. Thus, we call N-negaton any solution given in Proposition 4.4.1 generated by a matrix A chosen according to the following assumption.

Assumption 5.2.1. *Let $A \in \mathcal{M}_{n,n}(\mathbb{C})$ be given in Jordan form with N Jordan blocks A_j of dimension n_j and with eigenvalues α_j, i.e.,*

$$A = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & A_N \end{pmatrix}, \qquad A_j = \begin{pmatrix} \alpha_j & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \alpha_j \end{pmatrix} \in \mathcal{M}_{n_j,n_j}(\mathbb{C}).$$

Assume $\operatorname{spec}(A) \subseteq \{z \in \mathbb{C} \mid \operatorname{Re}(z) > 0$ *and* $f_0(z)$ *is finite*$\}$.

Moreover, let f_0 be a real function in sense of the explanations before Proposition 4.4.2, and assume either $\alpha_j \in \mathbb{R}$ or there exists a unique index $\overline{\jmath} \neq j$ such that $\alpha_{\overline{\jmath}} = \overline{\alpha}_j$ and $n_{\overline{\jmath}} = n_j$.

In the case of diagonal matrices A, real eigenvalues give rise to solitons and antisolitons, whereas pairs of eigenvalues which are complex conjugate lead to breathers. For details, see Section 4.4.3. Recall Definition 4.3.10 for the notion of asymptotic behaviour.

Theorem 5.2.2. *Let Assumption 5.2.1 be fulfilled. Assume that $a,\, c \in \mathbb{C}^n$ (when decomposed according to the Jordan form of A) satisfy $a_j^{(1)} c_j^{(n_j)} \neq 0$ for $j = 1, \ldots, N$, and, that $a_j,\, c_j \in \mathbb{R}^{n_j}$ if $\alpha_j \in \mathbb{R}$ whereas $a_{\overline{\jmath}} = \overline{a}_j$, $c_{\overline{\jmath}} = \overline{c}_j$ if $\alpha_j \notin \mathbb{R}$ and $\overline{\jmath}$ is the unique index with $\alpha_{\overline{\jmath}} = \overline{\alpha}_j$. Define*

$$v_j = -\operatorname{Re}\big(f_0(\alpha_j)\big)/\operatorname{Re}(\alpha_j).$$

Assume, in addition,

(i) v_j are pairwise different for $j \in \{k \mid \operatorname{Im}(\alpha_k) \geq 0\}$,

(ii) $v_j + f_0'(\alpha_j) \neq 0 \;\forall j$.

To these data we associate, for $\operatorname{Im}(\alpha_j) = 0$, the solitons

$$q_{jj'}^{\pm} \;=\; -\mathrm{i}\frac{\partial}{\partial x} \log \frac{\overline{p}_{jj'}^{\,\pm}}{p_{jj'}^{\,\pm}}, \qquad \textit{where} \;\; p_{jj'}^{\pm}(x,t) = 1 + \mathrm{i}(-1)^{j'} \epsilon_j \;\; \exp\big(\Gamma_{jj'}^{\pm}(x,t)\big),$$

and, for $\operatorname{Im}(\alpha_j) > 0$, the breathers

$$Q_{jj'}^{\pm} \;=\; -\mathrm{i}\frac{\partial}{\partial x} \log \frac{\overline{P}_{jj'}^{\,\pm}}{P_{jj'}^{\,\pm}},$$

where $P_{jj'}^{\pm}(x,t) = 1 + \dfrac{\mathrm{i}}{\gamma_j}\Big(\exp\big(\Gamma_{jj'}^{\pm}(x,t)\big) + \exp\big(\overline{\Gamma_{jj'}^{\pm}(x,t)}\big)\Big) + \exp\Big(\Gamma_{jj'}^{\pm}(x,t) + \overline{\Gamma_{jj'}^{\pm}(x,t)}\Big),$

with

$$\Gamma_{jj'}^{\pm}(x,t) = \alpha_j x + f_0(\alpha_j) t \;\mp\; J' \log|t| \;+ \varphi_j + \varphi_j^{\pm} + \varphi_{jj'}^{\pm} \;+\; (n_j \pm J') \log \gamma_j,$$

where we have set $J' = -(n_j - 1) + 2j'$.

Modulo $2\pi\mathrm{i}$, the quantities $\varphi_j^{\pm}$, and $\varphi_{jj'}^{\pm}$ are determined by

$$\exp(\varphi_j^{\pm}) \;=\; \prod_{k \in \Lambda_j^{\pm}} \left[\frac{\alpha_j - \alpha_k}{\alpha_j + \alpha_k}\right]^{2n_k}, \tag{5.31}$$

$$\exp\big(\pm \varphi_{jj'}^{\pm}\big) \;=\; \frac{j'!}{(j'-J')!} d_j^{-J'},$$

with the index sets $\Lambda_j^{\pm} = \{k \mid v_k \lessgtr v_j\}$, and $d_j = 2\alpha_j\big(v_j + f_0'(\alpha_j)\big)$. Furthermore,

$$\gamma_j = \begin{cases} 1, & \textit{if } \operatorname{Im}(\alpha_j) = 0, \\ |\operatorname{Im}(\alpha_j)/\operatorname{Re}(\alpha_j)|, & \textit{if } \operatorname{Im}(\alpha_j) > 0. \end{cases}$$

Finally, the φ_j are determined by $a_j^{(1)} c_j^{(n_j)}/(2\alpha_j)^{n_j} = \epsilon_j \exp(\varphi_j)$, $\varphi_j \in \mathbb{R}$, $\epsilon_j = \pm 1$, if $\operatorname{Im}(\alpha_j) = 0$ and (modulo $2\pi\mathrm{i}$) by $a_j^{(1)} c_j^{(n_j)}/(2\alpha_j)^{n_j} = \exp(\varphi_j)$, $\varphi_j \in \mathbb{C}$, if $\operatorname{Im}(\alpha_j) > 0$.

Then the asymptotic behaviour of the solution in Proposition 4.4.1 is described by

$$q(x,t) \approx \sum_{\substack{j=1 \\ \mathrm{Im}(\alpha_j)=0}}^{N} \sum_{j'=0}^{n_j-1} q^{\pm}_{jj'}(x,t) + \sum_{\substack{j=1 \\ \mathrm{Im}(\alpha_j)>0}}^{N} \sum_{j'=0}^{n_j-1} Q^{\pm}_{jj'}(x,t) \qquad \text{for } t \approx \pm\infty.$$

Since the geometric content of the theorem is quite similar to that of Theorem 5.1.2, we concentrate on the differences. First recall that the assumptions on the eigenvalues (and on the corresponding parts of the vectors a, c) either to be real or to appear in complex conjugate pairs guarantees reality of the solution (see Proposition 4.4.2).

There are negatons consisting of solitons and antisolitons ($\mathrm{Im}(\alpha_j) = 0$) and negatons consisting of breathers ($\mathrm{Im}(\alpha_j) > 0$). For a description of solitons, antisolitons, and breathers we refer to Section 4.4.3, in particular the comments after the proof of Proposition 4.4.9. Note that in the case $\mathrm{Im}(\alpha_j) = 0$ all phase-shifts $\varphi^{\pm}_j$, $\varphi^{\pm}_{jj'}$, and thus also the functions $\Gamma^{\pm}_{jj'}$ determining the path of the solitons in the j-th negaton, are real.

In contrast, if $\mathrm{Im}(\alpha_j) > 0$, the functions $\Gamma^{\pm}_{jj'}$ are no longer real. Their real parts give the trajectories of the breathers, the imaginary parts their oscillations. Accordingly, the phase-shifts $\varphi^{\pm}_j$, $\varphi^{\pm}_{jj'}$ effect both the trajectories and the oscillation. Note also that the logrithmic term is absorbed in the real part. Moreover, the additional term γ_j is only present in this case.

A remarkable difference between Theorems 5.1.2 and 5.2.2 concerns the expressions for the external phase-shifts (5.3) and (5.31). For illustration, let $k \in \Lambda^+_j$ with $\mathrm{Im}(\alpha_k) \neq 0$. In the $\mathbb{C}$-reduction there need not be another index $\overline{k}$ with $\alpha_{\overline{k}} = \overline{\alpha}_k$. Thus here the contribution to (5.3) is

$$\left(\frac{\alpha_j - \alpha_k}{\alpha_j + \overline{\alpha}_k}\right)^{2n_k}.$$

In the $\mathbb{R}$-reduced case, we have k, $\overline{k} \in \Lambda^+_j$, and therefore the contribution to (5.31) is

$$\left(\left[\frac{\alpha_j - \alpha_k}{\alpha_j + \alpha_k}\right]\left[\frac{\alpha_j - \overline{\alpha}_k}{\alpha_j + \overline{\alpha}_k}\right]\right)^{2n_k}.$$

However, again the sum over all phase-shifts is a conserved quantity.

Corollary 5.2.3. *The sum over all phase-shifts vanishes:*

$$\sum_{j=1}^{N}\sum_{j'=0}^{n_j} \left([\varphi^+_j + \varphi^+_{jj'}] - [\varphi^-_j + \varphi^-_{jj'}] \right) = 0 \qquad (\mathrm{mod}\ 2\pi \mathrm{i}).$$

Another conserved quantity is the 'topological charge' Q,

$$Q = \frac{1}{2\pi}\int_{-\infty}^{\infty} u(x,t)dx,$$

for u a soliton, an antisoliton, or a breather, see [69]. For solitons, we have $Q = 1$, for antisolitons, $Q = -1$, and breathers - as bound states of a soliton and an antisoliton - yield $Q = 0$.

This generalizes naturally to negatons.

Corollary 5.2.4. *The topological charge Q of the N-negaton is given by*

$$Q = -\sum\nolimits_j \epsilon_j,$$

where the sum ranges over all j with (i) $\mathrm{Im}(\alpha_j) = 0$ and (ii) n_j is odd.

Already the N-solitons show that every integer is attained as topological charge. In a single negaton solitons and antisolitons alternate. Hence (ii) means that the charge of a negaton with an even number of members always vanishes.

For completeness we give an interpretation of our result.

Interpretation 5.2.5. *a) To start with, consider a single eigenvalue $\alpha \in \mathbb{R}$ of multiplicity n. Then the solution is a cluster containing n waves which are either solitons or antisolitons. Such a solution is called a (single) negaton. The main observation is that the geometric center of the cluster propagates with constant velocity $v = -f_0(\alpha)/\alpha$, whereas its members drift away on logarithmic curves.*

Thus we may visualize, for large negative times, each soliton and antisoliton on one definite side of the center approaching it logarithmically. As time goes by, they get closer, collide, and separate again. For large positive times they can be found on the opposite side of the center, moving away from it again logarithmically. In particular, solitons and antisolitons appear exactly in reversed order in the asymptotic forms for $-\infty$ and $+\infty$.

Moreover, solitons and antisolitons always alternate. In particular, for each n there are only two different types of asymptotic forms. Namely, it suffices to know whether the wave which is the farthest to the left in the asymptotic form for $-\infty$ is a soliton or an antisoliton. The latter is determined by the sign ϵ of the parameter responsible for the initial position of the cluster.

It is remarkable that the cluster itself, or, more precisely, the path of its geometric center, is not disturbed by the internal collisions which the solitons and antisolitons experience on their way from one to the other side of the center.

b) Next consider a pair of complex conjugate eigenvalues α, $\overline{\alpha}$, of the same multiplicity n. Again the solution is a cluster, but now containing n breathers. Except of the fact that the velocity of the cluster as a whole is given by $v = \mathrm{Re}\big(f_0(\alpha)\big)/\mathrm{Re}(\alpha)$, it behaves as described in a).

In the asymptotic forms for $\pm\infty$, now only breathers appear. However, the statement that solitons and antisolitons always alternate remains true if we interprete breathers as a bound state of a soliton and an antisoliton (confer [69]). Let us in addition point out that the oscillation of the breathers is synchronized in the asymptotic forms for $\pm\infty$.

c) In general, the solution consists of $M = M_1 + M_2$ single negatons,

$$M_1 = \#|\{j = 1, \dots, N \mid \mathrm{Im}(\alpha_j) = 0\}| \text{ clusters of solitons/antisolitons as in a)}$$

$$M_2 = \#|\{j = 1, \dots, N \mid \mathrm{Im}(\alpha_j) > 0\}| \text{ clusters of breathers as in b).}$$

Hence, it is called an M-negaton (in general $M \neq N$).

These negatons rather behave like solitons. They collide elastically and the only effect of the collision is a phase-shift (of the whole negaton). Moreover, the resulting formulas for phase-shifts (5.31) are a very natural extension of the known formulas for the collision of solitons, from which they differ only by the exponents n_j.

The proof of Theorem 5.2.2 follows precisely the line of arguments pursued for Theorem 5.1.2. The main advantage in this case is, that the relevant determinant p in the solution formula of Proposition 4.4.1 is considerably easier to handle, because its dimension is only half as large.

On the other hand we have to overcome the difficulty to include also breathers in our asymptotic analysis. As bound states of two solitons, these are the lowest-dimensional examples of a formation of solitons. Solitons in such a formation cannot be separated in asymptotic terms but have to be treated as a joint entity.

In the sequel we will not repeat the full argument but carefully indicate all changes necessary to adapt the proof of Theorem 5.1.2 to the case at hand.

Sketch of the proof of Theorem 5.2.2

The organization of the proof will be the same as that of Theorem 5.1.2. For every step we will explain only the necessary modifications. As the arguments for p and p_x are very similar, we concentrate on the treatment of p.

Part 1. Technical reductions

As for the preparational reduction, the main difference to the proof of Proposition 5.1.6 is that, instead of the operators $\Phi^{-1}_{A,\overline{A}}(\overline{a}\otimes c)$, $\Phi^{-1}_{\overline{A},A}(a\otimes\overline{c})$, now only one operator (but a different one), namely $\Phi^{-1}_{A,A}(a\otimes c)$, is involved. The result is:

Proposition 5.2.6. *The formula of the solution given in Proposition 4.4.1 can be reformulated as follows:*

$$q(x,t) = -\mathrm{i}\frac{\partial}{\partial x}\log\frac{\overline{p(x,t)}}{p(x,t)} \qquad \textit{with } p(x,t) = \det\big(1+\mathrm{i}MT\big),$$

where

(i) $T=\big(T_{ij}\big)_{i,j=1}^{N}$ *with*

$$T_{ij} = \left((-1)^{\nu+\mu}\binom{\nu+\mu-2}{\nu-1}\left(\frac{1}{\alpha_i+\alpha_j}\right)^{\nu+\mu-1}\right)_{\substack{\nu=1,\dots,n_i\\ \mu=1,\dots,n_j}},$$

(ii) M *as defined in Proposition 5.1.6.*

Note that the definition of the matrix T is different matrix from that in Proposition 5.1.6. Subsequently this will lead to minor changes whenever the determinant of T is evaluated. To be precise, this changes

a) the constants $C_{\kappa\lambda}$ in Proposition 5.2.9,

b) the initial phases φ_j, the phase-shifts $\varphi_j^{\pm}$, and the constant d_j.

Part 2. Asymptotic estimates

Recall that, for $j\in\{1,\dots,N\}$, the index $\overline{\jmath}\in\{1,\dots,N\}$ is the unique index such that $\alpha_{\overline{\jmath}}=\overline{\alpha}_j$. There are two cases: If $\alpha_j\in\mathbb{R}$, then $\overline{\jmath}=j$, else $\overline{\jmath}\neq j$.

Step 1. The essential change concerns the index sets, because two indices $j,\overline{\jmath}$ with $\overline{\jmath}\neq j$ correspond to solitons with the same velocities $v_{\overline{\jmath}}=v_j$. Since they cannot be separated in asymptotic terms, they have to be treated on equal footing.

The same reasoning as in the proof of Proposition 5.1.9 yields the following result, which shows how q decomposes into single negatons.

Proposition 5.2.7. $q(x,t)\approx\sum\limits_{j_0=1}^{N} q^{(j_0)}(x,t)$ *for* $t\approx-\infty$,

where

$$q^{(j_0)}(x,t) = -\mathrm{i}\frac{\partial}{\partial x}\log\frac{\overline{p^{(j_0)}(x,t)}}{p^{(j_0)}(x,t)} \qquad \textit{with } p^{(j_0)} = \det\Big(1^{(j_0)}+\mathrm{i}M^{(j_0)}T^{(j_0)}\Big)$$

and the entries are defined by

$$
\begin{aligned}
T^{(j_0)} &= (\ T_{ij}\ )_{i,j\in\Lambda^-_{j_0}\cup\{j_0,\overline{\jmath}_0\}},\\
I^{(j_0)} &= \operatorname{diag}\{I^{(j_0)}_j \mid j\in\Lambda^-_{j_0}\cup\{j_0,\overline{\jmath}_0\}\}\ \textit{with blocks}\ I^{(j_0)}_j = \begin{cases} I_{n_{j_0}}, & j\in\{j_0,\overline{\jmath}_0\},\\ 0, & j\in\Lambda^-_{j_0},\end{cases}\\
M^{(j_0)} &= \operatorname{diag}\{M^{(j_0)}_j \mid j\in\Lambda^-_{j_0}\cup\{j_0,\overline{\jmath}_0\}\}\ \textit{with blocks}\ M^{(j_0)}_j = \begin{cases} M_{j_0}, & j\in\{j_0,\overline{\jmath}_0\},\\ I_{n_{j_0}}, & j\in\Lambda^-_{j_0}.\end{cases}
\end{aligned}
$$

Step 2. Let us first adapt the expansion rule, which has been the main tool of the proof of Proposition 5.1.13, to the present situation. To this end, let us consider a matrix S with the particular block structure

$$
S = (\ S_{ij}\ )_{i,j\in\Lambda^-_{j_0}\cup\{\overline{\jmath}_0,j_0\}} \qquad \text{with} \qquad S_{ij} = (\ S^{(\nu\mu)}_{ij}\ )_{\substack{\nu=1,\dots,n_i\\ \mu=1,\dots,n_j}}\ .
$$

Let us assume $\overline{\jmath}_0 \neq j_0$ for the moment.

For index tuples $J = (\sigma_1,\dots,\sigma_\kappa)$ and $K = (\tau_1,\dots,\tau_\lambda)$ with $\sigma_k,\tau_l \in \{1,\dots,n_{j_0}\}$, $1\le k\le\kappa$, $1\le l\le\lambda$, we now define $S[J\widehat{\times}K]$ as the matrix

$$
S[J\widehat{\times}K] = (\ S[J\widehat{\times}K]_{ij}\ )_{i,j}
$$

where the block $S[J\widehat{\times}K]_{ij}$ is obtained from S_{ij} by maintaining only

1. the rows indexed by $\begin{cases} J, & \text{if } i=j_0,\\ K, & \text{if } i=\overline{\jmath}_0,\end{cases}$
2. the columns indexed by $\begin{cases} J, & \text{if } j=j_0,\\ K, & \text{if } j=\overline{\jmath}_0,\end{cases}$

and, moreover, all maintained columns and rows in the blocks $S[J\widehat{\times}K]_{ij}$ appear in the order which is indicated by the index sets J, K.

For $\overline{\jmath}_0 = j_0$, we can use the usual expansion rule.

Lemma 5.2.8. *In the situation described above, the following expansion rules holds:*

a) For $\overline{\jmath}_0 = j_0$,

$$
\det\ (\ I^{(j_0)} + S\ ) = \sum_{\kappa=0}^{n_{j_0}}\ \sum_{|J|=\kappa}{}' \det\ (\ S[J,J]\ ),
$$

where the inner sum is taken over all index tuples J from $\{1,\dots,n_{j_0}\}$.

b) For $\overline{\jmath}_0 \neq j_0$,

$$
\det\ (\ I^{(j_0)} + S\ ) = \sum_{\kappa=0}^{n_{j_0}}\sum_{\lambda=0}^{n_{j_0}}\ \sum_{|J|=\kappa}{}' \sum_{|K|=\lambda}{}' \det\ (\ S[J\widehat{\times}K]\ ),
$$

where the inner sums are taken over all index tuples J, K from $\{1,\dots,n_{j_0}\}$.

In both cases the prime means that only index sets with strictly increasing entries are admitted.

Following the arguments of the proof of Proposition 5.1.9 (note that for the $\mathbb{R}$-reduction Theorem 6.1.1 suffices for the calculation of $\det(T)$) we get as result:

Proposition 5.2.9. *On the curve $(\gamma_\rho(t),t)$, where $\gamma_\rho(t) = v_{j_0}t + (\rho\log|t|)/\mathrm{Re}(\alpha_{j_0})$ for $\rho\in\mathbb{R}$, the determinant $p^{(j_0)}$ behaves according to*

a) If $\overline{\jmath}_0 = j_0$, then

$$p^{(j_0)}(\gamma_\rho(t),t) \;=\; C\left(\sum_{\kappa=0}^{n_{j_0}} \mathrm{i}^\kappa(-1)^{\frac{\kappa(\kappa+3)}{2}}\; D_\kappa\, G_\kappa |t|^{\rho\kappa} t^{(n_{j_0}-\kappa)\kappa}\right)\Big[1+\mathcal{O}\Big(\frac{\log|t|}{t}\Big)\Big],$$

with $C = \det\big(T_{ij}\big)_{i,j\in\Lambda_{j_0}^-}$,

$$D_\kappa \;=\; \left[\frac{1}{2\alpha_{j_0}}\right]^{\kappa^2}\prod_{j\in\Lambda_{j_0}^-}\left[\frac{\alpha_j-\alpha_{j_0}}{\alpha_j+\alpha_{j_0}}\right]^{2\kappa n_j},$$

$$G_\kappa \;=\; \frac{\prod_{k=1}^{\kappa-1}k!}{\prod_{k=1}^{\kappa}(n_{j_0}-k)!}\Big(b_{j_0}^{(n_{j_0})}\big(v_{j_0}+f_0'(\alpha_{j_0})\big)^{n_{j_0}-\kappa}\Big)^\kappa .$$

b) If $\overline{\jmath}_0 \neq j_0$, then

$$\begin{aligned} p^{(j_0)}(\gamma_\rho(t),t) \;=\; C\Bigg(\sum_{\kappa=0}^{n_{j_0}}\sum_{\lambda=0}^{n_{j_0}} &\mathrm{i}^{\kappa+\lambda}(-1)^{\frac{\kappa(\kappa+3)}{2}}(-1)^{\frac{\lambda(\lambda+3)}{2}}\; C_{\kappa\lambda}\, F_\kappa(t)\overline{F_\lambda(t)}\,\cdot \\ &\cdot|t|^{\rho(\kappa+\lambda)}t^{(n_{j_0}-\kappa)\kappa+(n_{j_0}-\lambda)\lambda}\Bigg)\Big[1+\mathcal{O}\Big(\frac{\log|t|}{t}\Big)\Big],\end{aligned}$$

with C as in a),

$$C_{\kappa\lambda} \;=\; \left[\frac{1}{2\alpha_{j_0}}\right]^{\kappa^2}\left[\frac{1}{2\overline{\alpha}_{j_0}}\right]^{\lambda^2}\left[\frac{\alpha_{j_0}-\overline{\alpha}_{j_0}}{\alpha_{j_0}+\overline{\alpha}_{j_0}}\right]^{2\kappa\lambda}\prod_{j\in\Lambda_{j_0}^-}\left[\frac{\alpha_j-\alpha_{j_0}}{\alpha_j+\alpha_{j_0}}\right]^{2\kappa n_j}\left[\frac{\alpha_j-\overline{\alpha}_{j_0}}{\alpha_j+\overline{\alpha}_{j_0}}\right]^{2\lambda n_j},$$

$$F_\kappa(t) \;=\; \frac{\prod_{k=1}^{\kappa-1}k!}{\prod_{k=1}^{\kappa}(n_{j_0}-k)!}\bigg(b_{j_0}^{(n_{j_0})}\big(v_{j_0}+f_0'(\alpha_{j_0})\big)^{n_{j_0}-\kappa}\exp\Big(\mathrm{i}\,\mathrm{Im}\big(\alpha_j x+f_0(\alpha_j)t\big)\Big)\bigg)^\kappa .$$

In Lemma 5.2.8, Proposition 5.2.9, the statement a) is a special case of b), namely $\lambda = 0$ (and α_j real).

Now the proof of Proposition 5.1.13 carries over in a straightforward manner, and we obtain the following description of the interior structure of a single negaton.

Proposition 5.2.10. *$q^{(j_0)}(x,t)\approx\sum_{j_0'=0}^{n_{j_0}-1}q^{(j_0j_0')}(x,t)$ for $t\approx-\infty$ with*

$$q^{(j_0j_0')}(x,t) = -\mathrm{i}\frac{\partial}{\partial x}\log\frac{\overline{p^{(j_0j_0')}(x,t)}}{p^{(j_0j_0')}(x,t)},$$

where

$$p^{(j_0j_0')}(x,t) = \begin{cases} 1+\mathrm{i}(-1)^{j_0'}\epsilon_{j_0}\epsilon^{J_0'}\exp\Big(\Gamma^-_{j_0j_0'}(x,t)\Big), & \text{if } \overline{\jmath}_0 = j_0,\\ 1+\mathrm{i}\epsilon^{J_0'}\gamma_{j_0}^{-1}\Big(\exp\big(\Gamma^-_{j_0j_0'}(x,t)\big)+\exp\big(\overline{\Gamma^-_{j_0j_0'}(x,t)}\big)\Big) & \\ \qquad\qquad +\exp\Big(\Gamma^-_{j_0j_0'}(x,t)+\overline{\Gamma^-_{j_0j_0'}(x,t)}\Big), & \text{if } \overline{\jmath}_0 \neq j_0,\end{cases}$$

and where $\epsilon = sgn(t)$, the function $\Gamma^-_{j_0j_0'}(x,t)$ is defined by

$$\Gamma^-_{j_0j_0'}(x,t) = \alpha_{j_0}x + f_0(\alpha_{j_0})t - J_0'\log|t| + \varphi_{j_0} + \varphi^-_{j_0} - \varphi^-_{j_0j_0'} + (n_{j_0}+J_0')\log\gamma_{j_0}$$

for the index J_0' associated to j_0', and the phase φ_{j_0} (together with ϵ_{j_0} if $\overline{\jmath}_0 = j_0$), the phase-shifts $\varphi^-_{j_0}$, $\varphi^-_{j_0j_0'}$, and γ_{j_0} as defined in Theorem 5.2.2.

By the same reordering argument as in the proof of Theorem 5.1.2, the assertion of Theorem 5.2.2 follows. The sketch of the proof is complete. $\square$

5.3 Illustration of the result in the lowest-dimensional cases

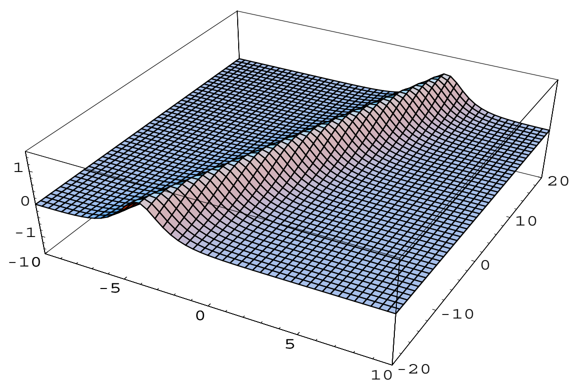
In this section, we assemble some computer graphics, which we found instructive during the preparation. The main point is of course the difference between straight lines of solitary waves and logarithmic rayes of the members of clusters. Furthermore the diagrams confirm that the convergence we have established by asymptotic formulas is in fact very rapid.

5.3.1 Computer graphics for the derivative sine-Gordon equation in laboratory coordinates

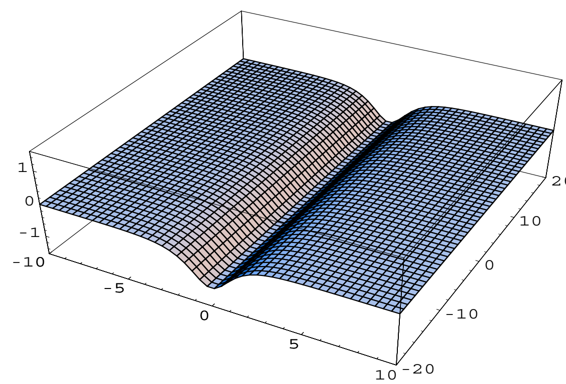
First we provide material for the derivative sine-Gordon equation, which serves as prototypical example of the $\mathbb{R}$-reduced AKNS system. For a better geometric description, we have turned to coordinates $\xi = x + t$, $\tau = -x + t$.

First we present pictures of solitons and breathers, which are the building blocks for the negatons. Note that the plots for breather confirms nicely its interpretation as a bound state of a soliton and an antisoliton (see Novikov [69]).

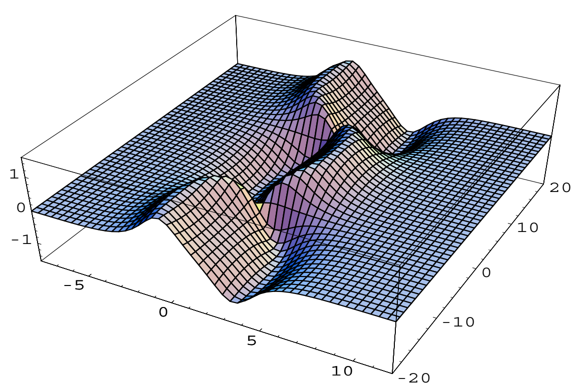
In the pictures, the variables ξ and τ are depicted in canonical fashion.



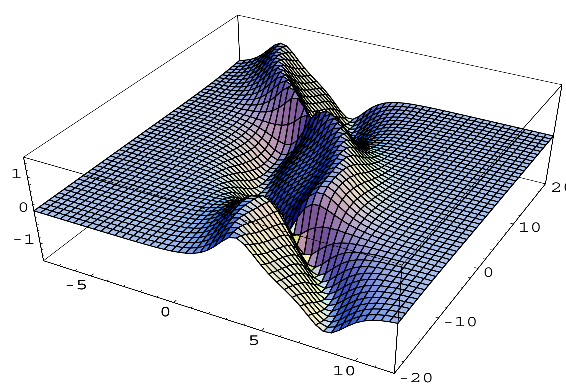
moving soliton ($a = 1.2$)



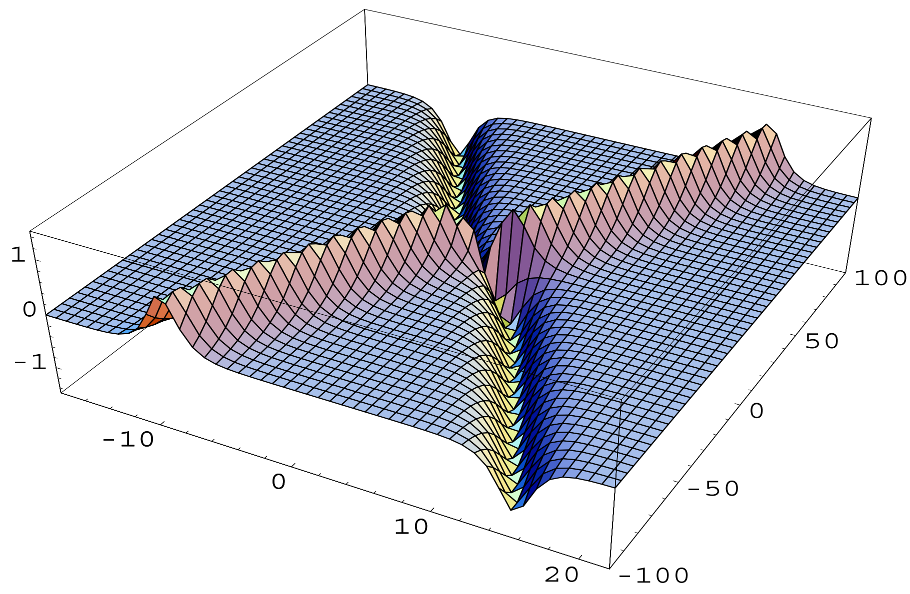
stationary antisoliton ($a = 1$)



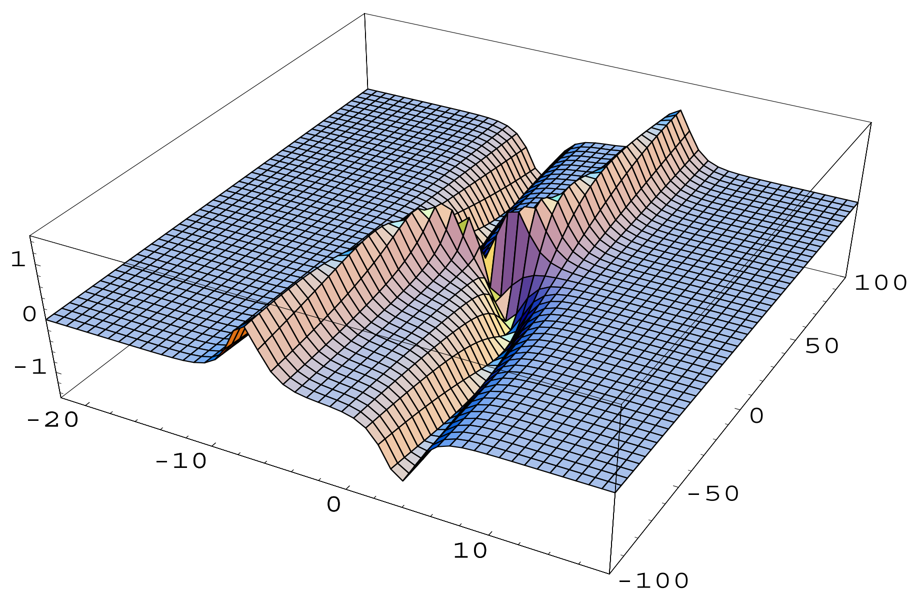
stationary breather ($a = \sqrt{1 - 0.2^2} + 0.2\mathrm{i}$)



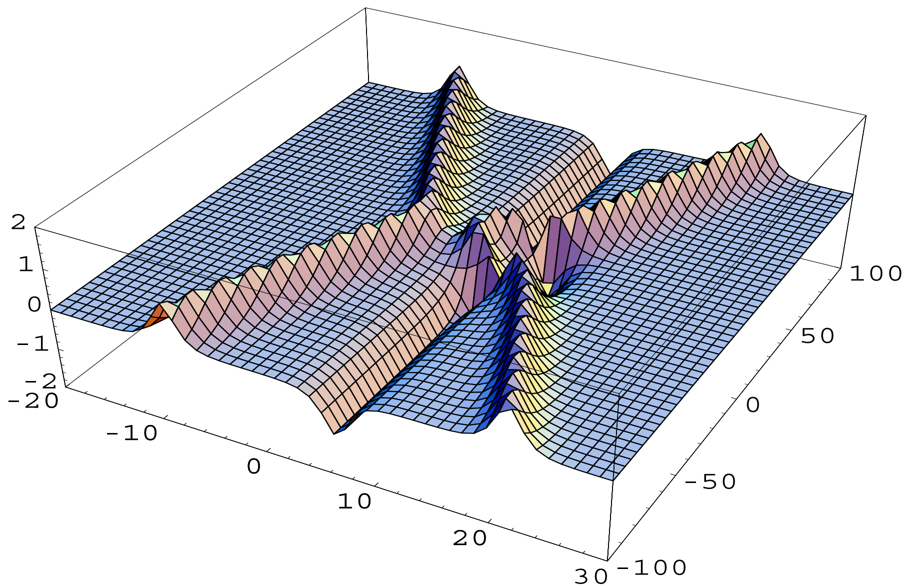
moving breather ($a = 0.8\big(\sqrt{1 - 0.18^2} + 0.18\mathrm{i}\big)$)



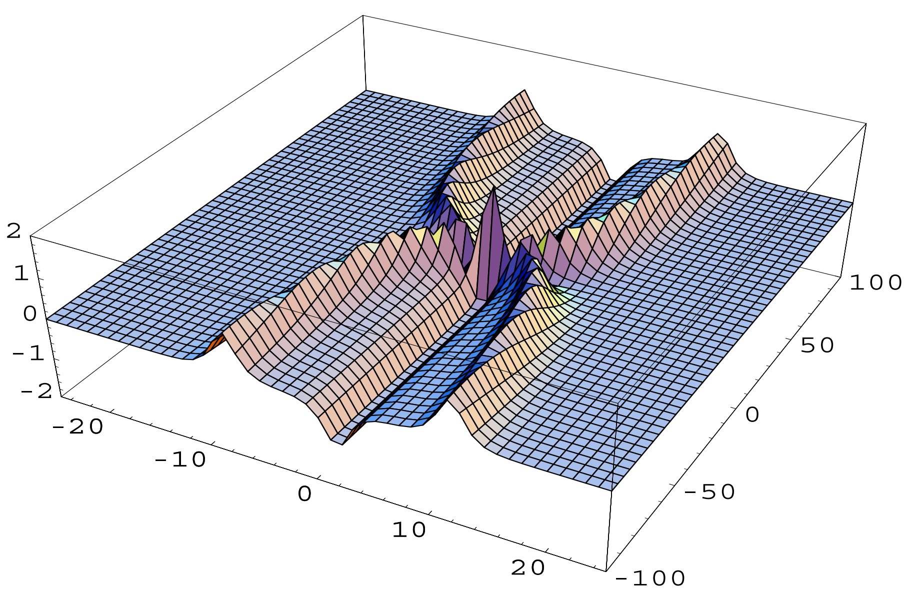
soliton $(a_1 = 1.1)$ meets antisoliton $(a_2 = 0.9)$



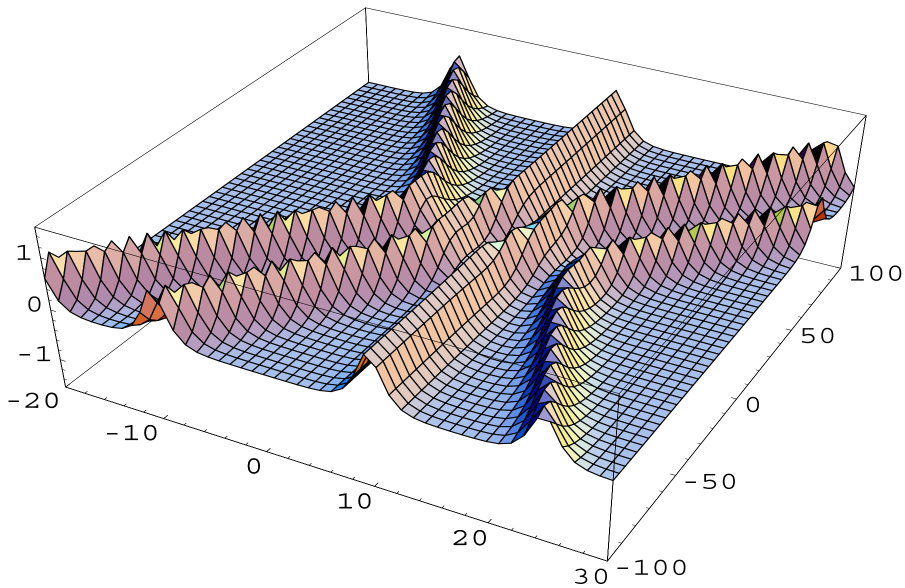
negaton $(a = 1)$ consisting of a soliton and an antisoliton



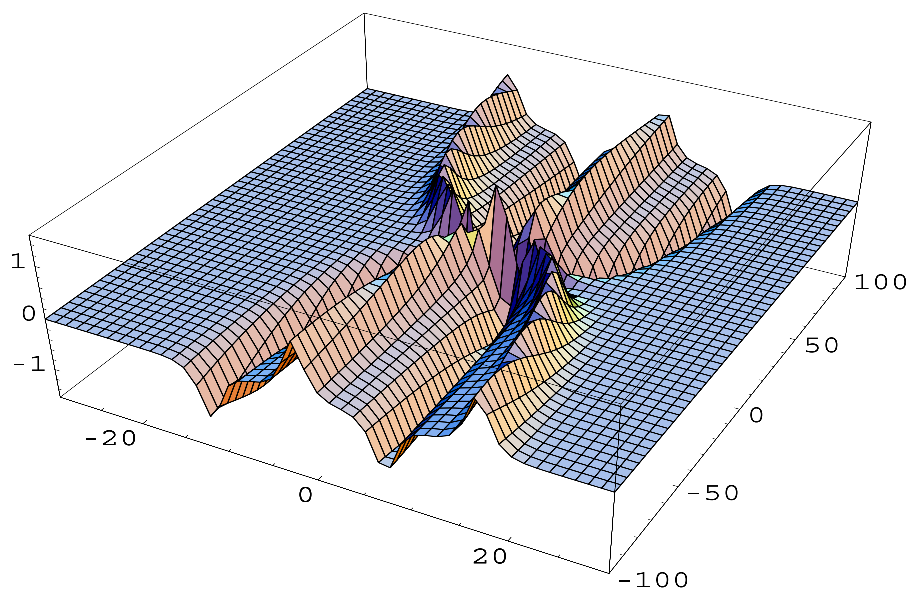
two solitons ($a_1 = 0.9$, $a_2 = 1.1$) and an antisoliton ($a_3 = 1$) meet



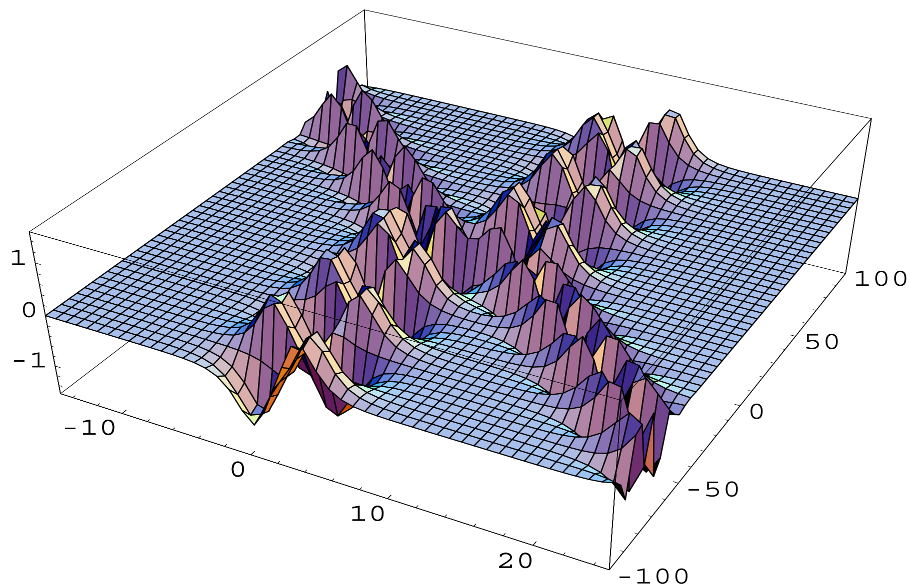
negaton ($a = 1$) consisting of two solitons and an antisoliton



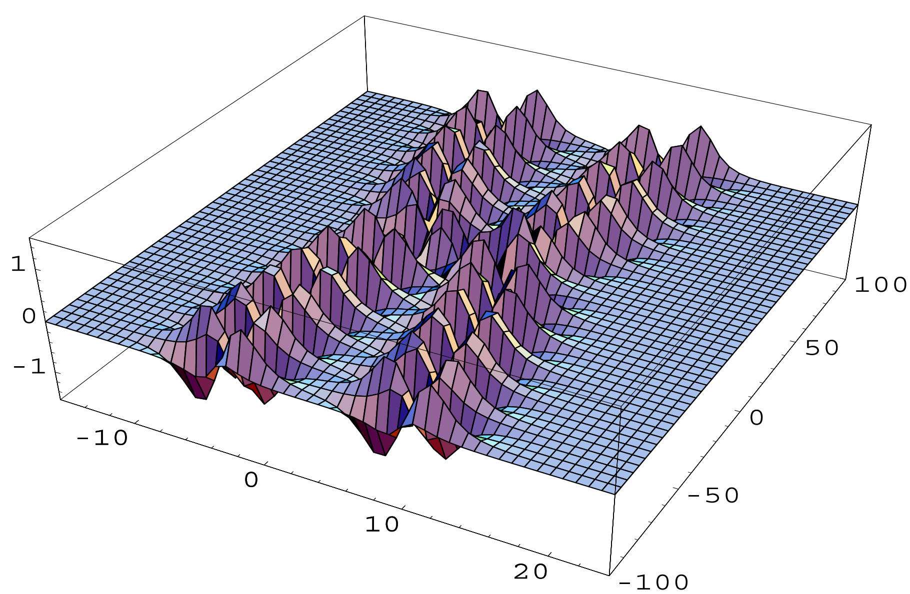
four solitons ($a_1 = 0.9$, $a_2 = 1$, $a_3 = 1.2$, $a_4 = 1.25$)



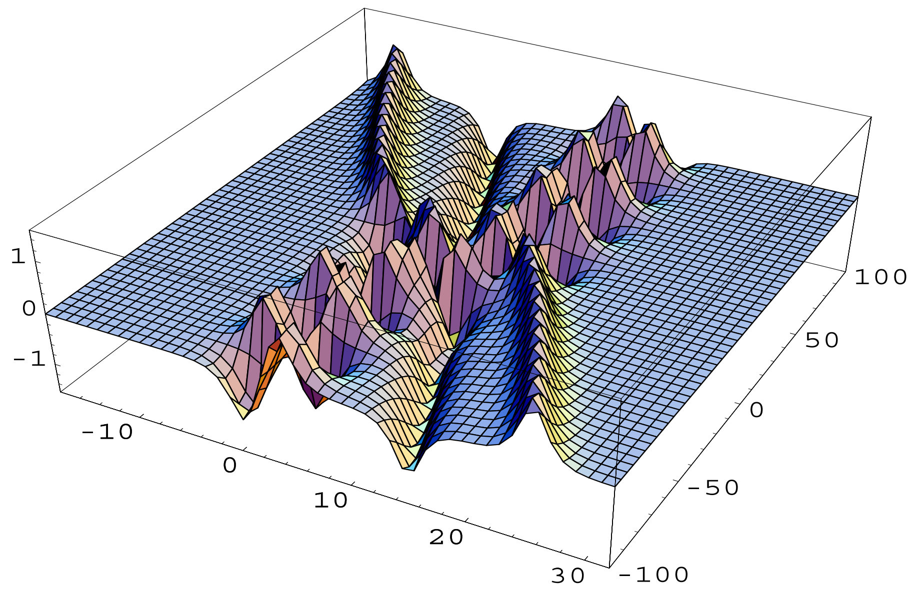
negaton ($a = 1$) consisting of two solitons and two antisolitons



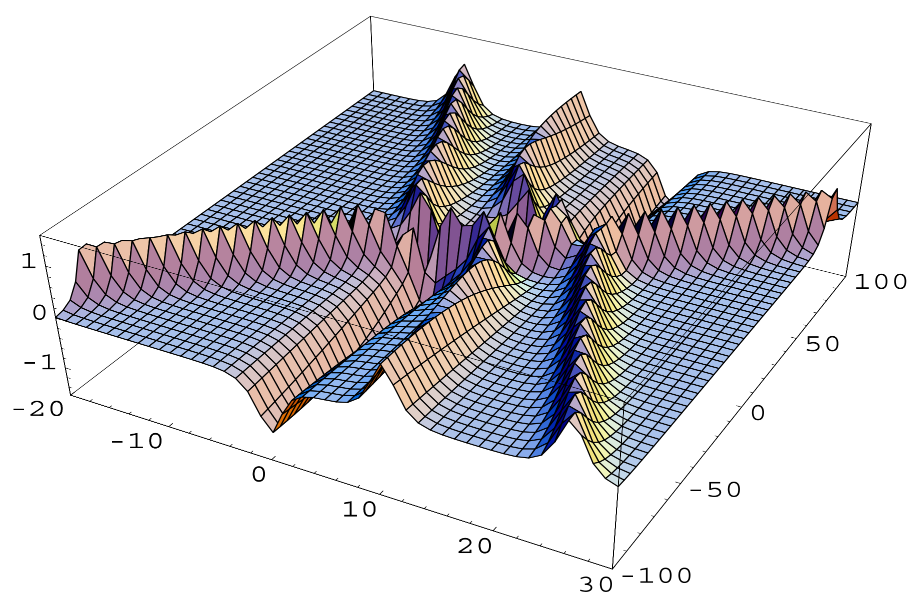
two breathers ($a_1 = 0.8\left(\sqrt{1-0.4^2}+0.4\mathrm{i}\right)$, $a_2 = \sqrt{1-0.2^2}+0.2\mathrm{i}$) meet



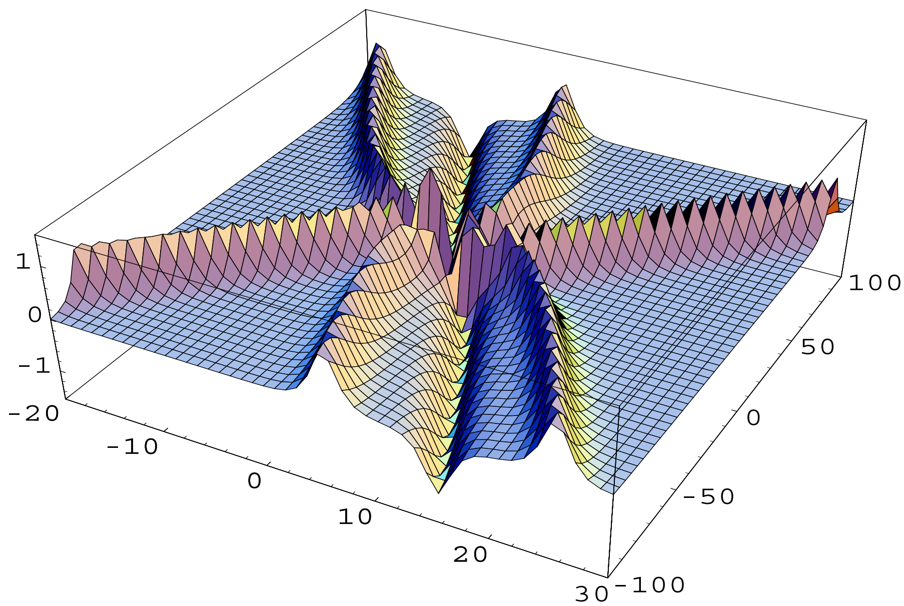
negaton ($a = \sqrt{1-0.4^2}+0.4\mathrm{i}$) consisting of two breathers



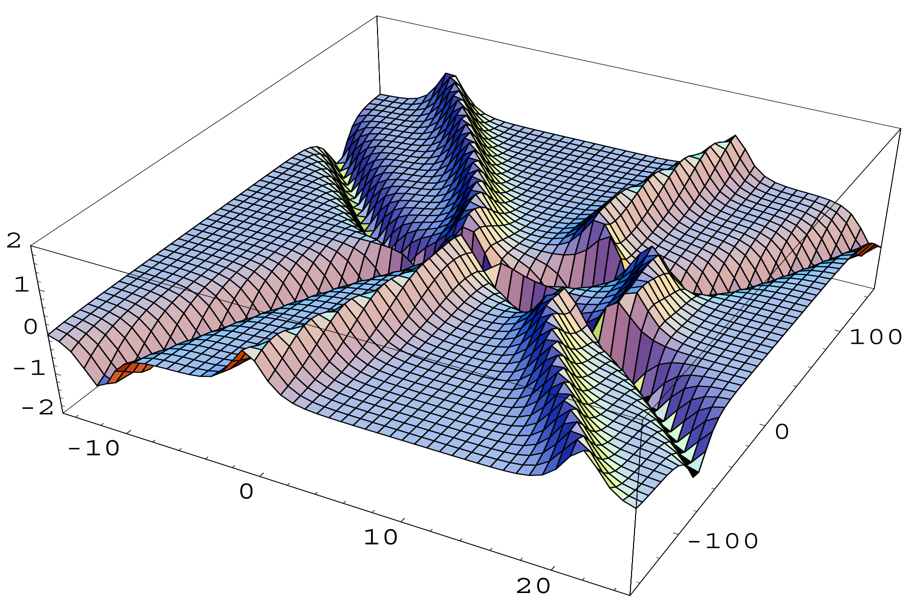
negaton ($a_1 = 0.9$) consisting of a soliton and an antisoliton meets breather ($a_2 = \sqrt{1-0.2^2} + 0.2\mathrm{i}$)



negaton ($a_1 = 1$) consisting of a soliton and an antisoliton meets two solitons ($a_2 = 0.9$, $a_3 = 1.3$)



negaton ($a_1 = 0.9$) consisting of two solitons and an antisoliton meets a soliton ($a_2 = 1$)

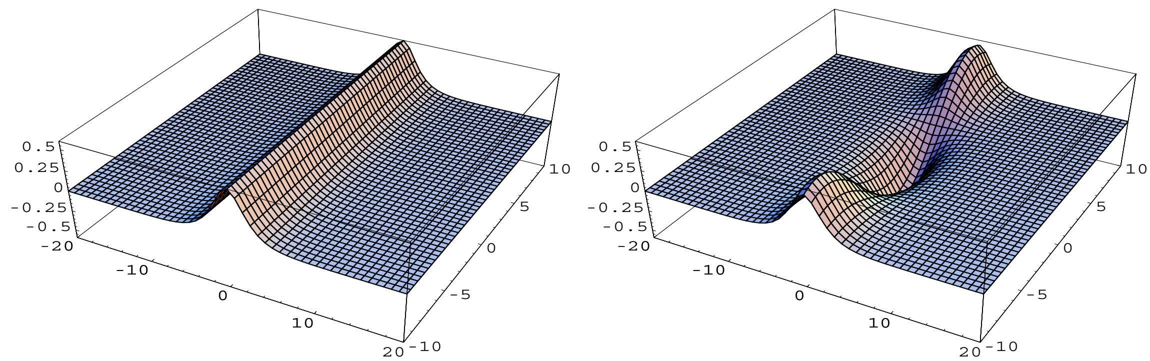


two negatons ($a_1 = 0.9$, $a_2 = 1.05$), each consisting of a soliton and an antisoliton, meet

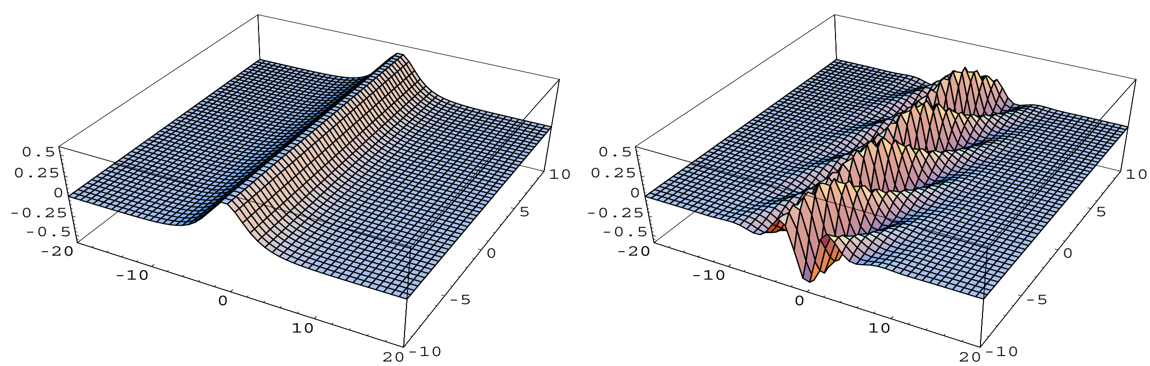
5.3.2 Computer graphics for the Nonlinear Schrödinger equation

Secondly, we turn to the Nonlinear Schrödinger equation, which is the most prominent member of the $\mathbb{C}$-reduced AKNS system. Since its solutions genuinely are complex, here we plot real part and modulus.

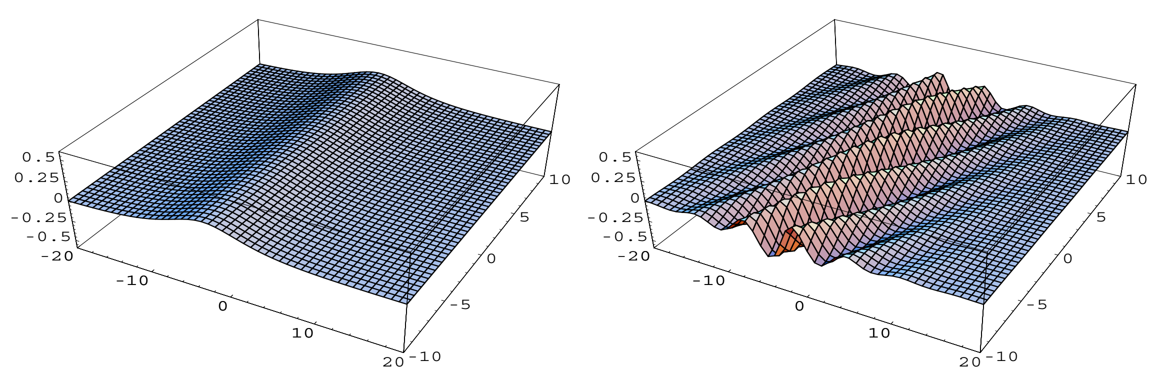
Again we start with a series of plots of one-solitons. Note that these are plotted in the coordinates $(x - vt, t)$, where v is the velocity of the soliton.



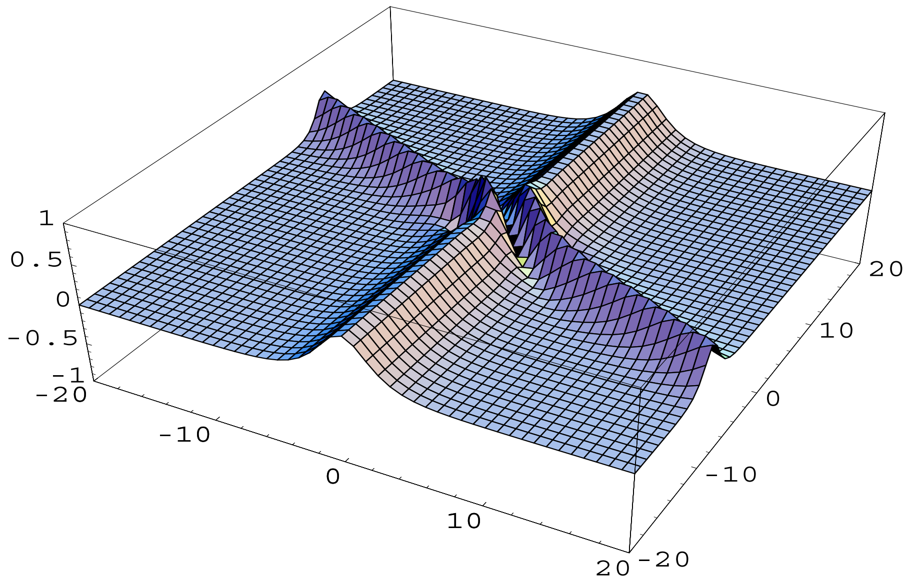
modulus and real part of the one-soliton with $a = 0.6$

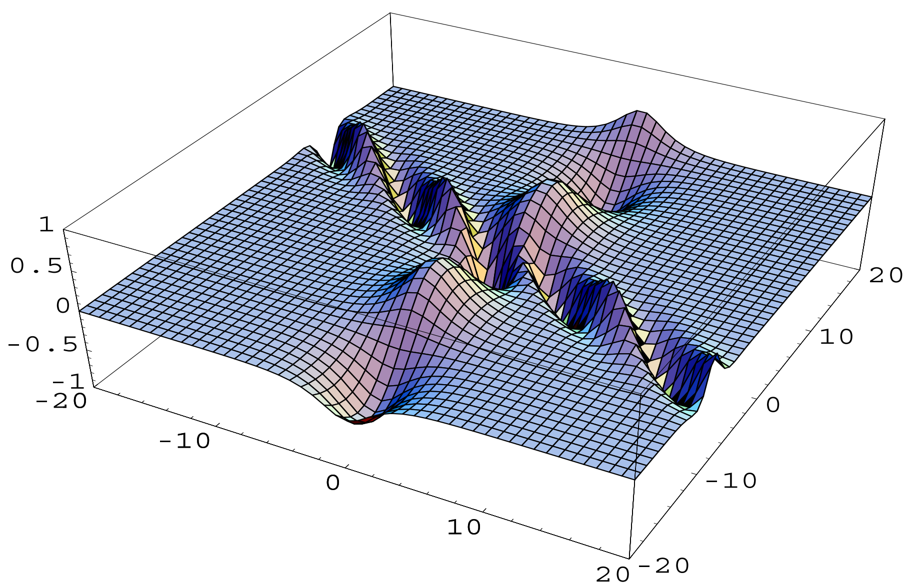


modulus and real part of the one-soliton with $a = 0.5 + \mathrm{i}$



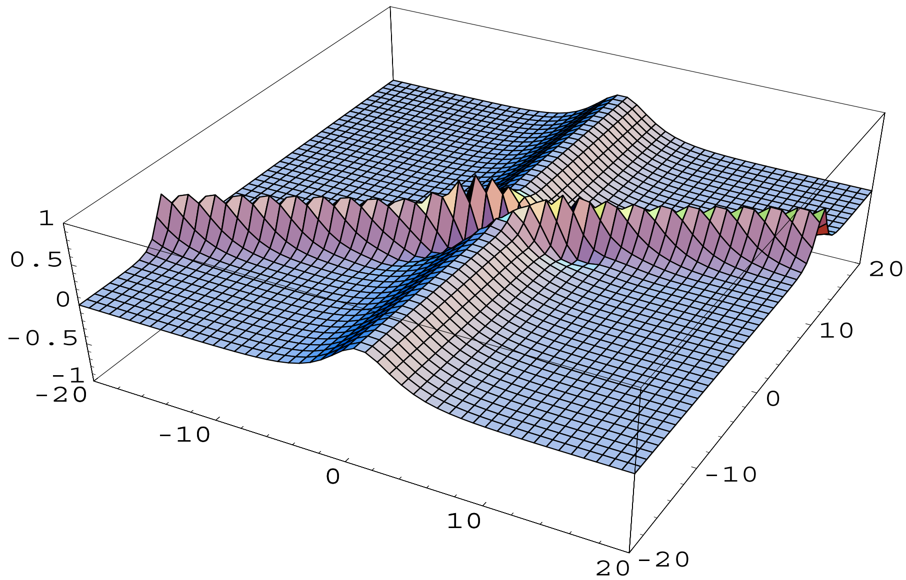
modulus and real part of the one-soliton with $a = 0.25 + \mathrm{i}$

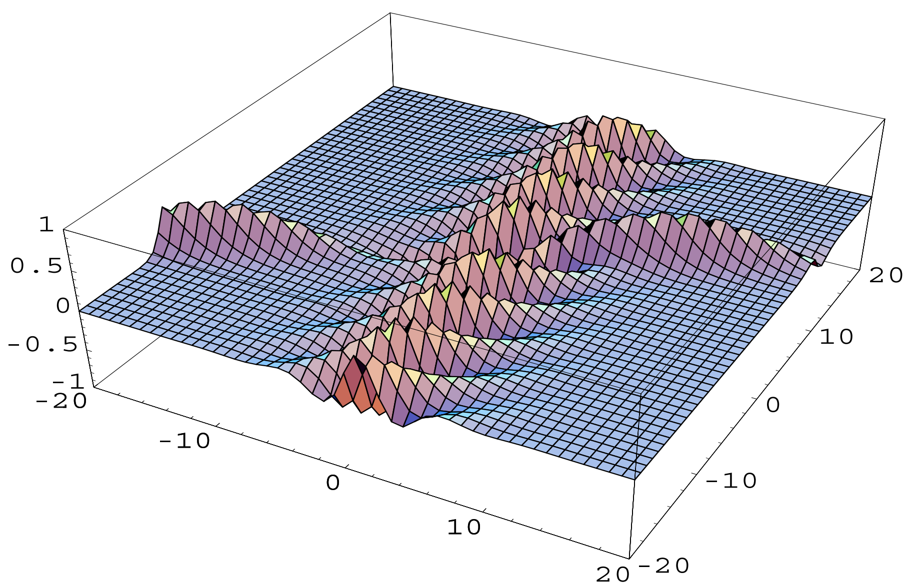




two solitons ($a_1 = 0.5 + \mathrm{i}$, $a_2 = 0.6$) meet

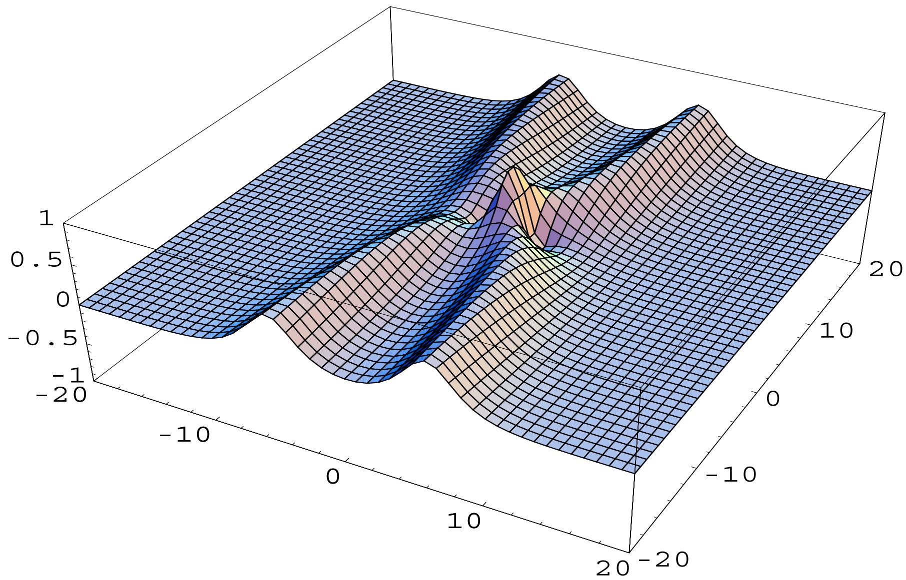
Note that the solution is drawn as usual in the coordinates (x, t). The plot above shows the modulus, the plot below the real part of the solution.

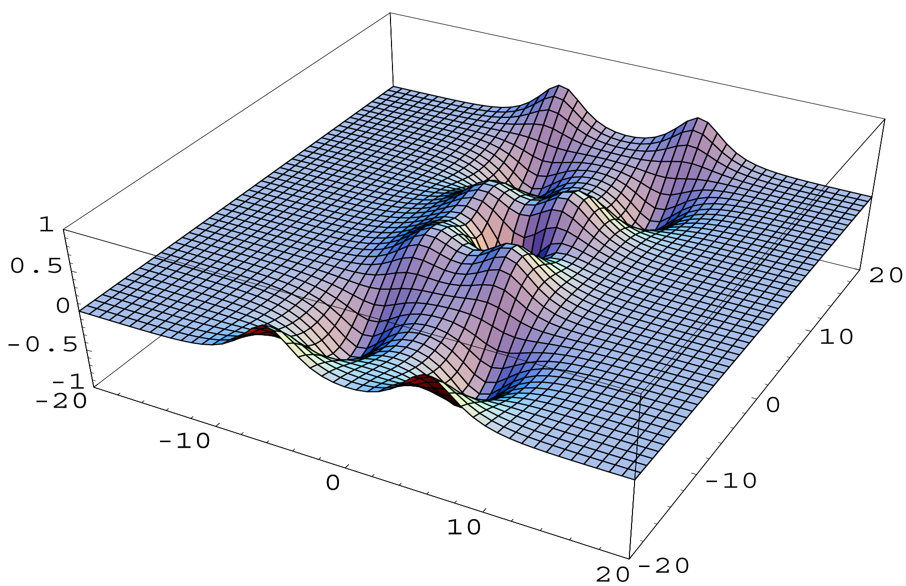




two solitons ($a_1 = 0.5 + \mathrm{i}$, $a_2 = 0.6$) meet

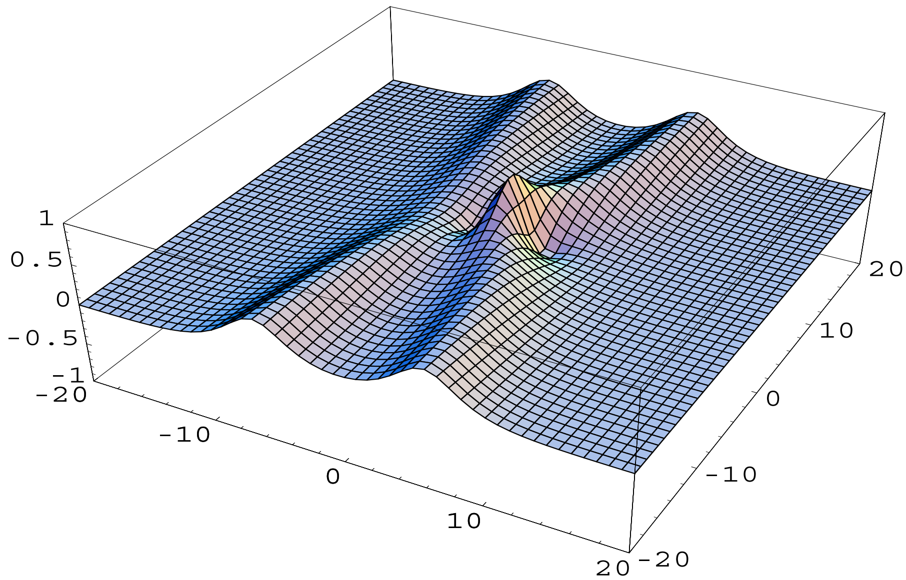
Here the same solution is drawn as before, but in the coordinates $(x - v_1 t, t)$, $v_1 = -2$ the velocity of the soliton corresponding to a_1, Again the plot above show the modulus, the plot below the real part of the solution.

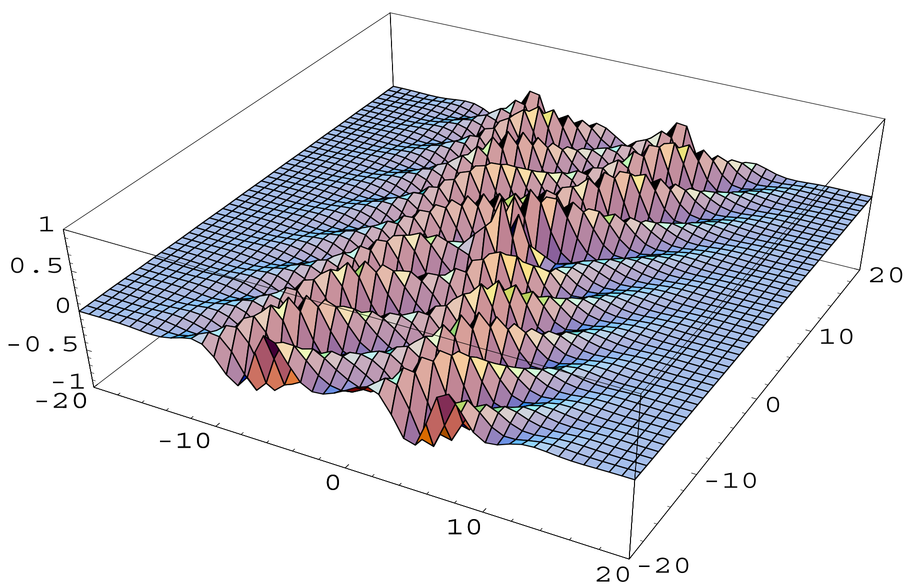




negaton ($a = 0.6$) consisting of two solitons

This solution is a stationary negaton, which is drawn in the coordinates (x, t). The plot above shows its modulus, the plot below its real part.





negaton ($a = 0.5 + \mathrm{i}$) consisting of two solitons

This solution moves with velocity $v = -2$ (corresponding to a), and is drawn in the coordinates $(x - vt, t)$. Once more the plot above shows its modulus, the plot below its real part.

Chapter 6

Determination of the phase-shifts

In the proofs of Theorem 5.1.2 and Theorem5.2.2 we used for the determination of the phase-shift explicit knowledge of several complicated determinants. The content of this chapter is to complete the asymptotic analysis by proving the Theorems 6.1.1, 6.2.1.

The results of this chapter may be also of independent operator-theoretic interest, because they compute the determinants of solutions of the operator equation $AX + XB = C$ (see Proposition 4.1.3).

As the following chapter will be very technical, it may be helpful to survey the main points. We consider matrices of the form $T = \left(T_{ij}\right)_{\substack{i=1,\dots,N \\ j=1,\dots,M}}$ with the blocks

$$T_{ij} = \left(\left(\frac{1}{\alpha_i + \beta_j} \right)^{\nu+\mu-1} \binom{(\nu-1)+(\mu-1)}{(\nu-1)} \right)_{\substack{\nu=1,\dots,n_i \\ \mu=1,\dots,m_j}} \in \mathcal{M}_{n_i,m_j}(\mathbb{C}),$$

where $\alpha_i + \beta_j \neq 0$ $(\forall i,j)$. Note $T \in \mathcal{M}_{n,m}(\mathbb{C})$ with $n = \sum_{i=1}^N n_i$, $m = \sum_{j=1}^M m_j$. In particular T need not be square.

To prove Theorem 5.1.2, first we have to evaluate

$$\det \begin{pmatrix} 0 & -T \\ T' & 0 \end{pmatrix}.$$

One readily sees (confer Lemma 4.1.8) that this determinant is only non-zero for square matrices T, in which case our task can be reduced to the evaluation of $\det(T)$.

In the very special case that all $m_j = n_j = 1$ (corresponding to N-solitons), the result is classical (see [20], p. 151-159, [78], VII, §1, Nr. 3). The general case is substantially more involved and will be treated in Theorem 6.1.1. The proof contains the main ideas of the present chapter.

For $m_j = n_j$ the result was already proved in the author's thesis [88]. Actually this case would be sufficient for the proof of the asymptotics in the $\mathbb{R}$-reduced case in Theorem 5.2.2 (see also [90], [91]), but not for the $\mathbb{C}$-reduction in Theorem 5.1.2.

Next we turn to the evaluation of determinants of the form

$$\det \begin{pmatrix} 0 & -T \\ T' & f \otimes f \end{pmatrix},$$

where $f \otimes f$ is a matrix representing a one-dimensional operator. Now there are two cases where the determinant can be non-trivial (see Proposition 4.1.9). If T is a square matrix, general arguments reduce our task to Theorem 6.1.1, but we have also examine the case where T is an $n \times (n+1)$-matrix. This is done in Theorem 6.2.1.

6.1 Extension of a result of Cauchy

In this section we prove Theorem 6.1.1. In the simplest case ($n_j = m_j = 1$) the result was already known by Cauchy ([20], p. 151-159), see also Lemma 6.1.2. For $n_j = m_j$ the result was shown in the author's thesis ([88], Theorem 4.2.1). The proof presented below is an extension of the argument in [88].

To start with, we introduce the following notation. Calculating determinants, the first row/column often has to be treated separately. In this case we write

$$\det\left(T_{ij}\right)_{\substack{i=1,\dots,N\\ j=1,\dots,M}} = \det\begin{pmatrix} T_{11} & T_{1j} \\ T_{i1} & T_{ij}\end{pmatrix}_{\substack{i>1\\ j>1}}.$$

Now we state the first main result of this chapter.

Theorem 6.1.1. *Assume that α_i, $i = 1, \dots, N$, and β_j, $j = 1, \dots, M$, are complex numbers satisfying $\alpha_i + \beta_j \neq 0$ for all i, j. Let n_i, $i = 1, \dots, N$, and m_j, $j = 1, \dots, M$, be natural numbers such that $\sum_{i=1}^N n_i = \sum_{j=1}^M m_j = n$.*

Then the determinant of the matrix $T = (T_{ij})_{\substack{i=1,\dots,N\\ j=1,\dots,M}} \in \mathcal{M}_{n,n}(\mathbb{C})$ with the blocks

$$T_{ij} = \left(\left(\frac{1}{\alpha_i+\beta_j}\right)^{\nu+\mu-1}\binom{\nu+\mu-2}{\nu-1}\right)_{\substack{\nu=1,\dots,n_i\\ \mu=1,\dots,m_j}} \in \mathcal{M}_{n_i,m_j}(\mathbb{C})$$

has the following value:

$$\det(T) = \prod_{\substack{i<j\\ i,j=1}}^{N}(\alpha_i-\alpha_j)^{n_in_j}\prod_{\substack{i<j\\ i,j=1}}^{M}(\beta_i-\beta_j)^{m_im_j}\Big/\prod_{i=1}^{N}\prod_{j=1}^{M}(\alpha_i+\beta_j)^{n_im_j}.$$

Before we enter the proof, we discuss two special cases, each of them requiring a strategy of its own. In the proof of Theorem 6.1.1 we will combine both strategies skilfully.

First we treat the case of one-dimensional blocks, i.e., $n_j = m_j = 1$ for $j = 1, \dots, N$. Here the result is classical, see [20]. We recall the proof in [78], but modify the arguments slightly to adapt it to the proof of Theorem 6.1.1.

Lemma 6.1.2. *Let α_i, $\beta_j \in \mathbb{C}$, $i, j = 1, \dots, N$, with $\alpha_i + \beta_j \neq 0$ $\forall i, j$. Then*

$$\det\left(\frac{1}{\alpha_i+\beta_j}\right)_{i,j=1}^{N} = \prod_{i=1}^{N}\frac{1}{(\alpha_i+\beta_i)}\prod_{\substack{i,j=1\\ i<j}}^{N}\frac{(\alpha_i-\alpha_j)(\beta_i-\beta_j)}{(\alpha_i+\beta_j)(\alpha_j+\beta_i)}.$$

Proof We pursue the following strategy:

(i) (Manipulations with respect to rows) Subtract the first row from the i-th row for $i = 2, \dots, N$.

(ii) (Manipulations with respect to columns) Multiply the first column by $(\alpha_1+\beta_1)/(\alpha_1+\beta_j)$ and then subtract it from the j-th column for $j = 2, \dots, N$.

This yields

$$\Delta \;:=\; \det\left(\frac{1}{\alpha_i+\beta_j}\right)_{i,j=1}^{N}$$

$$\stackrel{(i)}{=} \det \begin{pmatrix} \dfrac{1}{\alpha_1+\beta_j} \\ \left[\dfrac{1}{\alpha_i+\beta_j} - \dfrac{1}{\alpha_1+\beta_j}\right] \end{pmatrix}_{\substack{i>1\\ j\geq 1}}$$

$$= \det \begin{pmatrix} \dfrac{1}{\alpha_1+\beta_j} \\ \dfrac{1}{\alpha_i+\beta_j}\,\dfrac{\alpha_1-\alpha_i}{\alpha_1+\beta_j} \end{pmatrix}_{\substack{i>1\\ j\geq 1}}$$

$$\stackrel{(ii)}{=} \det \begin{pmatrix} \dfrac{1}{\alpha_1+\beta_1} & 0 \\ \dfrac{1}{\alpha_1+\beta_1}\,\dfrac{\alpha_1-\alpha_i}{\alpha_i+\beta_1} & \left[\dfrac{1}{\alpha_i+\beta_j} - \dfrac{1}{\alpha_i+\beta_1}\right]\dfrac{\alpha_1-\alpha_i}{\alpha_1+\beta_j} \end{pmatrix}_{\substack{i>1\\ j>1}}$$

$$= \det \begin{pmatrix} \dfrac{1}{\alpha_1+\beta_1} & 0 \\ \dfrac{1}{\alpha_1+\beta_1}\left[\dfrac{\alpha_1-\alpha_i}{\alpha_i+\beta_1}\right] & \dfrac{1}{\alpha_i+\beta_j}\left[\dfrac{\beta_1-\beta_j}{\alpha_1+\beta_j}\right]\left[\dfrac{\alpha_1-\alpha_i}{\alpha_i+\beta_1}\right] \end{pmatrix}_{\substack{i>1\\ j>1}}.$$

Next we extract the common factors $(\beta_1-\beta_j)/(\alpha_1+\beta_j)$ from the j-th column ($j=2,\ldots,N$) and $(\alpha_1-\alpha_i)/(\alpha_i+\beta_1)$ from the i-th row ($i=2,\ldots,N$). Finally, expanding the determinant with respect to the first column, we obtain

$$\Delta = \frac{1}{\alpha_1+\beta_1} \prod_{1<j} \frac{(\alpha_1-\alpha_j)(\beta_1-\beta_j)}{(\alpha_j+\beta_1)(\alpha_1+\beta_j)} \det\left(\frac{1}{\alpha_i+\beta_j}\right)_{i,j=2}^{N},$$

and the assertion follows by induction. □

Next we consider the case that T only consists of a single block, i.e., $M=N=1$.

Lemma 6.1.3. *For $\gamma \in \mathbb{C}$, it holds*

$$\det\left(\gamma^{\nu+\mu-1} \quad \binom{\nu+\mu-2}{\nu-1}\right)_{\nu,\mu=1}^{n} = \gamma^{n^2}.$$

Proof With special regard to the order induced by the numbering of the indices, we pursue the following strategy:

(i) (Manipulations with respect to colums) Multiply the $(\mu-1)$-th column by γ and subtract it from the μ-th column for $\mu=n,\ldots,2$.

(ii) (Manipulations with respect to rows) Multiply the $(\nu-1)$-th row by γ and subtract it from the ν-th row for $\nu=n,\ldots,2$.

This results in

$$\Delta := \det\left(\gamma^{\nu+\mu-1}\binom{\nu+\mu-2}{\nu-1}\right)_{\nu,\mu=1}^{n}$$

$$\stackrel{(i)}{=} \det\left(\gamma^{\nu} \qquad \gamma^{\nu+\mu-1}\left[\binom{\nu+\mu-2}{\nu-1} - \binom{\nu+\mu-3}{\nu-1}\right]\right)_{\substack{\nu\geq 1\\ \mu>1}}$$

$$= \det\begin{pmatrix} \gamma & 0 \\ \gamma^{\nu} & \gamma^{\nu+\mu-1}\dbinom{\nu+\mu-3}{\nu-2} \end{pmatrix}_{\substack{\nu>1\\ \mu>1}}$$

$$\stackrel{(ii)}{=} \det \begin{pmatrix} \gamma & 0 & \\ 0 & \gamma^{\mu+1}\left[\,1-0\,\right] & \\ 0 & \gamma^{\nu+\mu-1}\left[\binom{\nu+\mu-3}{\nu-2} - \binom{\nu+\mu-4}{\nu-3}\right] & \end{pmatrix}_{\substack{\nu>2\\ \mu\geq 2}}$$

$$= \det \begin{pmatrix} \gamma & 0 \\ 0 & \gamma^2\cdot\gamma^{\nu+\mu-3}\binom{\nu+\mu-4}{\nu-2} \end{pmatrix}_{\substack{\nu>1\\ \mu>1}} .$$

Next we expand the determinant, and then extract the factor γ^2, which all the remaining rows and columns have in common. We end up with

$$\Delta \;=\; \gamma\cdot\gamma^{2(n-1)} \quad \det\left(\gamma^{\nu+\mu-1} \quad \binom{\nu+\mu-2}{\nu-1}\right)_{\nu,\mu=1}^{n-1},$$

and the assertion again follows by induction. $\square$

Keeping these strategies in mind, we now enter the proof of Theorem 6.1.1.

Proof (of Theorem 6.1.1) It suffices to consider the situation where $\alpha_i \neq \alpha_j$ for all $i,j = 1,\ldots,N$, $i\neq j$, and $\beta_i \neq \beta_j$ for all $i,j=1,\ldots,M$, $i\neq j$, since otherwise the matrix T would contain linearly dependent columns or rows.

Our aim is to argue by induction. To keep the manipulations as clear as possible, we replace the usual operations of columns/rows by the multiplication with corresponding matrices.

We use the following notations. By T_{ij} we denote the ij-th block of T, and for its entries we write

$$T_{ij} \;=\; \left(\left(\frac{1}{\alpha_i+\beta_j}\right)^{\nu+\mu-1} t[i,j]_{\nu\mu}\right)_{\substack{\nu=1,\ldots,n_i\\ \mu=1,\ldots,m_j}} . \tag{6.1}$$

In the proof we will construct matrices $T^{\langle k\rangle}$ with blocks $T_{ij}^{\langle k\rangle}$. The numbers $t[i,j]_{\nu\mu}^{\langle k\rangle}$ will be related to the entries of $T_{ij}^{\langle k\rangle}$ exactly as in (6.1).

Moreover, we define

$$\Phi_i = \frac{\alpha_1-\alpha_i}{\alpha_i+\beta_1}, \quad \phi_{ij} = \frac{\alpha_1-\alpha_i}{\alpha_1+\beta_j}, \quad \Psi_j = \frac{\beta_1-\beta_j}{\alpha_1+\beta_j}, \quad \text{and} \quad \psi_{ji} = \frac{\beta_1-\beta_j}{\alpha_i+\beta_1} \tag{6.2}$$

for $i=1,\ldots,N$, $j=1,\ldots,M$. Note in particular, $\phi_{1j}=0$ and $\psi_{1i}=0$ $(\forall i,j)$.

Claim 1: $\det(T)=\det(T^{\langle 1\rangle})$, where the blocks $T_{ij}^{\langle 1\rangle}$, $i=1,\ldots,N$, $j=1,\ldots,M$, of $T^{\langle 1\rangle}$ are given by (modulo (6.1))

$$\left(t[i,j]_{\nu\mu}^{\langle 1\rangle}\right)_{\substack{\nu=1,\ldots,n_i\\ \mu=1,\ldots,m_j}} = \begin{pmatrix} 1 & \phi_{ij} \\ \psi_{ji} & \binom{\nu+\mu-4}{\nu-1}\phi_{ij} + \binom{\nu+\mu-4}{\mu-1}\psi_{ji} + \binom{\nu+\mu-4}{\nu-2}\left[1+\phi_{ij}\psi_{ji}\right] \end{pmatrix}_{\substack{\nu>1\\ \mu>1}} .$$

In particular,

$$T_{1j}^{\langle 1\rangle} = \begin{pmatrix} \dfrac{1}{\alpha_1+\beta_j} & 0 \\ * & * \end{pmatrix}_{\substack{\nu>1\\ \mu>1}} \qquad \text{and} \qquad T_{i1}^{\langle 1\rangle} = \begin{pmatrix} \dfrac{1}{\alpha_i+\beta_1} & * \\ 0 & * \end{pmatrix}_{\substack{\nu>1\\ \mu>1}} .$$

Proof of Claim 1: (Arguments within the single blocks) Here we use the strategy developed in Lemma 6.1.3 to create zero entries in the first column of the blocks T_{i1}, $i=1,\dots,N$, and in the first row of the blocks T_{1j}, $j=1,\dots,M$.

To this end, define

$$X_i = \begin{pmatrix} 1 & & & 0 \\ x_i & \cdot & & \\ & \cdot & \cdot & \\ & & \cdot & \cdot \\ 0 & & x_i & 1 \end{pmatrix} \in \mathcal{M}_{n_i,n_i}(\mathbb{C}), \quad Y_j = \begin{pmatrix} 1 & & & 0 \\ y_j & \cdot & & \\ & \cdot & \cdot & \\ & & \cdot & \cdot \\ 0 & & y_j & 1 \end{pmatrix} \in \mathcal{M}_{m_j,m_j}(\mathbb{C}),$$

with $x_i = -\dfrac{1}{\alpha_i+\beta_1}$, $y_j = -\dfrac{1}{\alpha_1+\beta_j}$ for $i=1,\dots,N$ and $j=1,\dots,M$.

Since the $\det(X_i) = \det(Y_j) = 1$ $\forall i,j$, the following manipulation does not change the value of the determinant of T:

$$\begin{aligned} \det(T) &= \det\left(\begin{pmatrix} X_1 & 0 & \cdots & 0 \\ 0 & X_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & X_N \end{pmatrix} T \begin{pmatrix} Y_1' & 0 & \cdots & 0 \\ 0 & Y_2' & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & Y_M' \end{pmatrix} \right) \\ &= \det\left(X_i T_{ij} Y_j' \right)_{\substack{i=1,\dots,N \\ j=1,\dots,M}} . \end{aligned}$$

Let us now explicitly calculate $T_{ij}^{\langle 1\rangle} = X_i T_{ij} Y_j' \in \mathcal{M}_{n_i,m_j}(\mathbb{C})$, where we use the notation introduced in the beginning of the proof for the entries of $T_{ij}^{\langle 1\rangle}$ (confer (6.1)). We compute

$$\begin{aligned} \left(\frac{1}{\alpha_i+\beta_j}\right)^{\nu+\mu-1} t[i,j]^{\langle 1\rangle}_{\nu\mu} &= \left(\frac{1}{\alpha_i+\beta_j}\right)^{\nu+\mu-1} t[i,j]_{\nu\mu} \\ &\quad +(1-\delta_{1\mu})\, y_j \left(\frac{1}{\alpha_i+\beta_j}\right)^{\nu+\mu-2} t[i,j]_{\nu(\mu-1)} \\ &\quad +(1-\delta_{1\nu})\, x_i \left(\frac{1}{\alpha_i+\beta_j}\right)^{\nu+\mu-2} t[i,j]_{(\nu-1)\mu} \\ &\quad +(1-\delta_{1\mu})(1-\delta_{1\nu})\, y_j x_i \left(\frac{1}{\alpha_i+\beta_j}\right)^{\nu+\mu-3} t[i,j]_{(\nu-1)(\mu-1)}. \end{aligned}$$

Therefore,

$$\begin{aligned} t[i,j]^{\langle 1\rangle}_{\nu\mu} &= t[i,j]_{\nu\mu} - (1-\delta_{1\mu})\frac{\alpha_i+\beta_j}{\alpha_1+\beta_j}\, t[i,j]_{\nu(\mu-1)} - (1-\delta_{1\nu})\frac{\alpha_i+\beta_j}{\alpha_i+\beta_1}\, t[i,j]_{(\nu-1)\mu} \\ &\quad +(1-\delta_{1\mu})(1-\delta_{1\nu})\frac{\alpha_i+\beta_j}{\alpha_1+\beta_j}\frac{\alpha_i+\beta_j}{\alpha_i+\beta_1}\, t[i,j]_{(\nu-1)(\mu-1)} \\ &\overset{(6.2)}{=} t[i,j]_{\nu\mu} - (1-\delta_{1\mu})\left[\,1-\phi_{ij}\,\right] t[i,j]_{\nu(\mu-1)} - (1-\delta_{1\nu})\left[\,1-\psi_{ji}\,\right] t[i,j]_{(\nu-1)\mu} \\ &\quad +(1-\delta_{1\mu})(1-\delta_{1\nu})\left[\,1-\phi_{ij}\,\right]\left[\,1-\psi_{ji}\,\right] t[i,j]_{(\nu-1)(\mu-1)}. \end{aligned}$$

Since $t[i,j]_{\nu\mu} = \binom{\nu+\mu-2}{\nu-1}$, we immediately find

$$t[i,j]^{\langle 1\rangle}_{\nu\mu} = \begin{cases} 1, & \nu=1, \mu=1, \\ \phi_{ij}, & \nu=1, \mu>1, \\ \psi_{ji}, & \nu>1, \mu=1, \end{cases}$$

and as for $\nu>1$, $\mu>1$, we finally calculate

$$\begin{aligned} t[i,j]^{\langle 1\rangle}_{\nu,\mu} &= \binom{\nu+\mu-2}{\nu-1} - (1-\phi_{ij})\binom{\nu+\mu-3}{\nu-1} - (1-\psi_{ji})\binom{\nu+\mu-3}{\nu-2} \\ &\quad +(1-\phi_{ij})(1-\psi_{ji})\binom{\nu+\mu-4}{\nu-2} \end{aligned}$$

$$
\begin{aligned}
&= \phi_{ij}\binom{\nu+\mu-3}{\nu-1} + \psi_{ji}\binom{\nu+\mu-3}{\nu-2} + (1-\phi_{ij})(1-\psi_{ji})\binom{\nu+\mu-4}{\nu-2} \\
&= \phi_{ij}\Big(\binom{\nu+\mu-3}{\nu-1} - \binom{\nu+\mu-4}{\nu-2}\Big) + \psi_{ji}\Big(\binom{\nu+\mu-3}{\nu-2} - \binom{\nu+\mu-4}{\nu-2}\Big) \\
&\qquad\qquad\qquad\qquad\qquad\qquad + (1+\phi_{ij}\psi_{ji})\binom{\nu+\mu-4}{\nu-2} \\
&= \phi_{ij}\binom{\nu+\mu-4}{\nu-1} + \psi_{ji}\Big(\binom{\nu+\mu-3}{\mu-1} - \binom{\nu+\mu-4}{\mu-2}\Big) \\
&\qquad\qquad\qquad\qquad\qquad\qquad + (1+\phi_{ij}\psi_{ji})\binom{\nu+\mu-4}{\nu-2} \\
&= \phi_{ij}\binom{\nu+\mu-4}{\nu-1} + \psi_{ji}\binom{\nu+\mu-4}{\mu-1} + (1+\phi_{ij}\psi_{ji})\binom{\nu+\mu-4}{\nu-2}.
\end{aligned}
$$

This completes the proff of Claim 1.

Claim 2: $\det(T^{\langle 1\rangle}) = \det(T^{\langle 2\rangle})$, where the blocks $T_{ij}^{\langle 2\rangle}$, $i=1,\ldots,N$, $j=1,\ldots,M$, of $T^{\langle 2\rangle}$ are given by

$$
t[1,j]_{11}^{\langle 2\rangle} = \begin{cases} 1, & j=1, \\ 0, & j>1, \end{cases} \qquad t[i,j]_{11}^{\langle 2\rangle} = \begin{cases} \phi_{i1}, & j=1, \\ \phi_{ij}\psi_{ji}, & j>1, \end{cases} \quad \text{for } i>1,
$$

and $t[i,j]_{\nu\mu}^{\langle 2\rangle} = t[i,j]_{\nu\mu}^{\langle 1\rangle}$ whenever $(\nu,\mu)\neq(1,1)$. In other words, in each block $T_{ij}^{\langle 1\rangle}$ we do only change the $(1,1)$-entry.

In particular,

$$
T_{11}^{\langle 2\rangle} = \begin{pmatrix} \dfrac{1}{\alpha_1+\beta_1} & * \\ 0 & * \end{pmatrix}_{\substack{\nu>1\\ \mu>1}} \qquad \text{and} \qquad T_{1j}^{\langle 2\rangle} = \begin{pmatrix} 0 & 0 \\ * & * \end{pmatrix}_{\substack{\nu>1\\ \mu>1}} \quad \text{for } j>1.
$$

Proof of Claim 2: (Arguments between the single blocks) Now we apply the strategy explained in Lemma 6.1.2 with respect to the $(1,1)$-entries of all blocks.

Recall that we denote by $e_k^{(1)}$ the first standard basis vector in $\mathbb{C}^k$. Define the matrices

$$
X_i = -e_{n_1}^{(1)} \otimes e_{n_i}^{(1)} \in \mathcal{M}_{n_i,n_1}(\mathbb{C}), \quad Y_j = -y_j\, e_{m_1}^{(1)} \otimes e_{m_j}^{(1)} \in \mathcal{M}_{m_j,m_1}(\mathbb{C}),
$$

with $y_j = \dfrac{\alpha_1+\beta_1}{\alpha_1+\beta_j}$ for $i=2,\ldots,N$, $j=2,\ldots,M$, and consider the manipulation

$$
\begin{aligned}
\det(T^{\langle 1\rangle}) &= \det\left(\begin{pmatrix} I & 0 & \cdots & 0 \\ X_2 & I & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ X_N & 0 & \cdots & I \end{pmatrix} T^{\langle 1\rangle} \begin{pmatrix} I & Y_2' & \cdots & Y_M' \\ 0 & I & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & I \end{pmatrix} \right) \\
&= \det(T^{\langle 2\rangle}),
\end{aligned}
$$

where

$$
T_{ij}^{\langle 2\rangle} = T_{ij}^{\langle 1\rangle} + (1-\delta_{1i})X_i T_{1j}^{\langle 1\rangle} + (1-\delta_{1j})T_{i1}^{\langle 1\rangle}Y_j' + (1-\delta_{1i})(1-\delta_{1j})X_i T_{11}^{\langle 1\rangle} Y_j'
$$

for $i=1,\ldots,N$, $j=1,\ldots,M$.

Concretely this means that we subtract the first row from the first rows of the horizontal strips $(T_{i1}^{\langle 1\rangle},\ldots,T_{iM}^{\langle 1\rangle})$, $i=2,\ldots,N$, and the (modified) first column from the first columns of the vertical strips $(T_{1j}^{\langle 1\rangle},\ldots,T_{Nj}^{\langle 1\rangle})'$, $2=1,\ldots,M$.

Exploiting the concrete form of $T_{1j}^{\langle 1\rangle}$ $(\forall j)$ and $T_{i1}^{\langle 1\rangle}$ $(\forall i)$ as given by Claim 1, we directly verify

$$X_i T_{1j}^{\langle 1\rangle} \;=\; -\Big((T_{1j}^{\langle 1\rangle})' e_{n_1}^{(1)} \Big) \otimes e_{n_i}^{(1)} \;=\; -\frac{1}{\alpha_1+\beta_j} e_{m_j}^{(1)} \otimes e_{n_i}^{(1)},$$

and, analogously,

$$T_{i1}^{\langle 1\rangle} Y_j' \;=\; -y_j \frac{1}{\alpha_i+\beta_1} e_{m_j}^{(1)} \otimes e_{n_i}^{(1)}, \qquad X_i T_{11}^{\langle 1\rangle} Y_j' \;=\; y_j \frac{1}{\alpha_1+\beta_1} e_{m_j}^{(1)} \otimes e_{n_i}^{(1)}.$$

From the fact that the matrices $e_{m_j}^{(1)} \otimes e_{n_i}^{(1)} \in \mathcal{M}_{n_i,m_j}(\mathbb{C})$ are zero except of the $(1,1)$-entry, it is clear that the performed manipulation does only change the $(1,1)$-entries of $T_{ij}^{\langle 1\rangle}$, in other words $t[i,j]_{\nu\mu}^{\langle 2\rangle} = t[i,j]_{\nu\mu}^{\langle 1\rangle}$ whenever $(\nu,\mu) \neq (1,1)$.

As for the $(1,1)$-entries we get

$$\begin{aligned}\frac{1}{\alpha_i+\beta_j} t[i,j]_{11}^{\langle 2\rangle} \;&=\; \frac{1}{\alpha_i+\beta_j} - (1-\delta_{1i})\frac{1}{\alpha_1+\beta_j} - (1-\delta_{1j}) y_j \frac{1}{\alpha_i+\beta_1} \\ &\qquad + (1-\delta_{1i})(1-\delta_{1j}) y_j \frac{1}{\alpha_1+\beta_1},\end{aligned}$$

yielding

$$\begin{aligned} t[i,j]_{11}^{\langle 2\rangle} \;&=\; 1 - (1-\delta_{1i})\frac{\alpha_i+\beta_j}{\alpha_1+\beta_j} - (1-\delta_{1j})\frac{\alpha_i+\beta_j}{\alpha_i+\beta_1}\frac{\alpha_1+\beta_1}{\alpha_1+\beta_j} \\ &\qquad + (1-\delta_{1i})(1-\delta_{1j})\frac{\alpha_i+\beta_j}{\alpha_1+\beta_j} \\ &=\; 1 - (1-\delta_{1j})\frac{\alpha_1+\beta_1}{\alpha_i+\beta_1}\frac{\alpha_i+\beta_j}{\alpha_1+\beta_j} - \delta_{1j}(1-\delta_{1i})\frac{\alpha_i+\beta_1}{\alpha_1+\beta_1} \\ &\qquad \text{(we may replace } \beta_j \text{ by } \beta_1 \text{ because of the Kronecker symbol)} \\ &\overset{(6.2)}{=}\; 1 - (1-\delta_{1j}) \left[\, 1 - \phi_{ij}\psi_{ji} \,\right] - \delta_{1j}(1-\delta_{1i}) \left[\, 1 - \phi_{i1} \,\right]\end{aligned}$$

Thus it is straightforward to check

$$t[i,1]_{11}^{\langle 2\rangle} = \begin{cases} 1, & i=1, \\ \phi_{i1}, & i>1, \end{cases} \qquad \text{and} \qquad t[i,j]_{11}^{\langle 2\rangle} = \phi_{ij}\psi_{ji} \quad \text{for } j>1.$$

Since $\psi_{1i} = 0$ for all i, this completes the proof of Claim 2.

Claim 3: $\det(T^{\langle 2\rangle}) = \left(\dfrac{1}{\alpha_1+\beta_1}\right)^{n_1+m_1-1} \det(T^{\langle 3\rangle})$, where the blocks $T_{ij}^{\langle 3\rangle}$, $i=1,\dots,N$, $j=1,\dots,M$, of $T^{\langle 3\rangle}$ are given by

$$\begin{aligned}
\Big(t[1,1]_{\nu\mu}^{\langle 3\rangle}\Big)_{\substack{\nu=1,\dots,\hat{n}_1\\ \mu=1,\dots,\hat{m}_1}} \;&=\; \left(\binom{\nu+\mu-2}{\nu-1}\right)_{\substack{\nu\geq 1\\ \mu\geq 1}}, \\
\Big(t[i,1]_{\nu\mu}^{\langle 3\rangle}\Big)_{\substack{\nu=1,\dots,\hat{n}_i\\ \mu=1,\dots,\hat{m}_1}} \;&=\; \begin{pmatrix} \Phi_i \\ \binom{\nu+\mu-3}{\nu-2} + \binom{\nu+\mu-2}{\nu-1}\Phi_i \end{pmatrix}_{\substack{\nu>1\\ \mu\geq 1}}, \qquad \text{for } i>1, \\
\Big(t[1,j]_{\nu\mu}^{\langle 3\rangle}\Big)_{\substack{\nu=1,\dots,\hat{n}_1\\ \mu=1,\dots,\hat{m}_j}} \;&=\; \begin{pmatrix} \Psi_j & \binom{\nu+\mu-3}{\nu-1} + \binom{\nu+\mu-2}{\mu-1}\Psi_j \end{pmatrix}_{\substack{\nu\geq 1\\ \mu>1}}, \qquad \text{for } j>1, \\
\Big(t[i,j]_{\nu\mu}^{\langle 3\rangle}\Big)_{\substack{\nu=1,\dots,\hat{n}_i\\ \mu=1,\dots,\hat{m}_j}} \;&=\; \begin{pmatrix} \phi_{ij}\psi_{ji} & \phi_{ij} \\ \psi_{ji} & \binom{\nu+\mu-4}{\nu-1}\phi_{ij} + \binom{\nu+\mu-4}{\mu-1}\psi_{ji} + \binom{\nu+\mu-4}{\nu-2}\left[1+\phi_{ij}\psi_{ji}\right] \end{pmatrix}_{\substack{\nu>1\\ \mu>1}}, \\
&\qquad \text{for } i>1,\ j>1.
\end{aligned}$$

The new dimensions are defined by $\widehat{n}_1 = n_1 - 1$ and $\widehat{n}_i = n_i$ for $i = 2, \ldots, N$, as well as $\widehat{m}_1 = m_1 - 1$ and $\widehat{m}_j = m_j$ for $j = 2, \ldots, M$. For simplicity, we consider matrices of the types $0 \times k$, $k \times 0$ as non-existent.

In particular, $T^{\langle 3 \rangle} \in \mathcal{M}_{n-1,n-1}(\mathbb{C})$.

Proof of Claim 3: (Expansion of the determinant) As a preparation, we simplify the results observed so far.

To this end, abbreviate $\gamma = (\alpha_1 + \beta_1)^{-1}$. The entries of $T_{11}^{\langle 2 \rangle}$ only experienced the strategy of the first step. This was described in Lemma 6.1.3 and we can copy the results from there (they are recorded below).

As for the entries of $T_{1j}^{\langle 2 \rangle}$, $j > 1$, we rewrite $\psi_{j1} = (\alpha_1 + \beta_j)\gamma\Psi_j$ (recall $\phi_{1j} = 0$), to see

$$\begin{aligned}
t[1,j]_{\nu 1}^{\langle 2 \rangle} &= (\alpha_1 + \beta_j)\gamma\Psi_j && (\forall \nu > 1), \\
t[1,j]_{\nu\mu}^{\langle 2 \rangle} &= \tbinom{\nu+\mu-4}{\nu-2} + \tbinom{\nu+\mu-4}{\mu-1}(\alpha_1 + \beta_j)\gamma\Psi_j \\
&\overset{(6.3)}{=} (\alpha_1 + \beta_j)\gamma\left(\tbinom{\nu+\mu-4}{\nu-2}(1 + \Psi_j) + \tbinom{\nu+\mu-4}{\mu-1}\Psi_j\right) \\
&= (\alpha_1 + \beta_j)\gamma\left(\tbinom{\nu+\mu-4}{\nu-2} + \tbinom{\nu+\mu-3}{\mu-1}\Psi_j\right) && (\forall \mu > 1, \forall \nu > 1).
\end{aligned}$$

As for the entries of $T_{i1}^{\langle 2 \rangle}$, $i > 1$, inserting $\phi_{i1} = (\alpha_i + \beta_1)\gamma\Phi_i$ (recall also $\psi_{1i} = 0$) we analogously get

$$\begin{aligned}
t[i,1]_{1\mu}^{\langle 2 \rangle} &= (\alpha_i + \beta_1)\gamma\Phi_i && (\forall \mu), \\
t[i,1]_{\nu\mu}^{\langle 2 \rangle} &= \tbinom{\nu+\mu-4}{\nu-2} + \tbinom{\nu+\mu-4}{\nu-1}(\alpha_i + \beta_1)\gamma\Phi_i \\
&\overset{(6.3)}{=} (\alpha_i + \beta_1)\gamma\left(\tbinom{\nu+\mu-4}{\nu-2}(1 + \Phi_i) + \tbinom{\nu+\mu-4}{\nu-1}\Phi_i\right) \\
&= (\alpha_i + \beta_1)\gamma\left(\tbinom{\nu+\mu-4}{\nu-2} + \tbinom{\nu+\mu-3}{\nu-1}\Phi_i\right) && (\forall \mu > 1, \forall \nu > 1).
\end{aligned}$$

Above we have used the identities

$$\frac{1}{\alpha_1 + \beta_j}\gamma^{-1} = 1 + \Psi_j, \quad \frac{1}{\alpha_i + \beta_1}\gamma^{-1} = 1 + \Phi_i. \tag{6.3}$$

There is no need to consider $T_{ij}^{\langle 2 \rangle}$ for $i > 1$, $j > 1$, since these blocks are not altered by the expansion below.

Let us sum up what we have achieved so far.

$$\begin{aligned}
T_{11}^{\langle 2 \rangle} &= \begin{pmatrix} \gamma & 0 \\ 0 & \gamma^2 \left(\frac{1}{\alpha_1+\beta_1}\right)^{\nu+\mu-3} \binom{\nu+\mu-4}{\nu-2} \end{pmatrix}_{\substack{\nu>1 \\ \mu>1}}, \\
T_{1j}^{\langle 2 \rangle} &= \begin{pmatrix} 0 & 0 \\ \gamma\left(\frac{1}{\alpha_1+\beta_j}\right)^{\nu-1}\Psi_j & \gamma\left(\frac{1}{\alpha_1+\beta_j}\right)^{\nu+\mu-2}\left(\binom{\nu+\mu-4}{\nu-2} + \binom{\nu+\mu-3}{\mu-1}\Psi_j\right) \end{pmatrix}_{\substack{\nu>1 \\ \mu>1}}, \quad j > 1, \\
T_{i1}^{\langle 2 \rangle} &= \begin{pmatrix} \gamma\Phi_i & \gamma\left(\frac{1}{\alpha_i+\beta_1}\right)^{\mu-1}\Phi_i \\ 0 & \gamma\left(\frac{1}{\alpha_i+\beta_1}\right)^{\nu+\mu-2}\left(\binom{\nu+\mu-4}{\nu-2} + \binom{\nu+\mu-3}{\nu-1}\Phi_i\right) \end{pmatrix}_{\substack{\nu>1 \\ \mu>1}}, \quad i > 1.
\end{aligned}$$

Note that, in the first row of $T^{\langle 2 \rangle}$, only the first entry is non-zero. Hence expanding reduces the dimension by one, and Claim 3 then follows by extracting the factor γ which is common to

(i) the ν-th row, $\nu = 2, \ldots, n_1$, of the blocks $T_{1j}^{\langle 2 \rangle}$ $(\forall j)$,

(ii) the μ-th columns, $\mu = 2, \ldots, m_1$, of the blocks $T_{i1}^{\langle 2 \rangle}$ $(\forall i)$.

Claim 4: $\det(T^{\langle 3\rangle}) = \Big(\prod_{i=2}^{N}\Phi_i^{n_i}\prod_{j=2}^{M}\Psi_j^{m_j}\Big)\det(\widehat{T})$, where the blocks $\widehat{T}_{ij}$ for $i=1,\ldots,N$, $j=1,\ldots,M$ of $\widehat{T}$ are given by

$$\widehat{T}_{ij} = \left(\left(\frac{1}{\alpha_i+\beta_j}\right)^{\nu+\mu-1}\binom{\nu+\mu-2}{\nu-1}\right)_{\substack{\nu=1,\ldots,\widehat{n}_i\\ \mu=1,\ldots,\widehat{m}_j}} \in \mathcal{M}_{\widehat{n}_i,\widehat{m}_j}(\mathbb{C})$$

and $\widehat{n}_i$, $\widehat{m}_j$ are defined as in Claim 3.

Therefore, $\widehat{T}$ is a matrix of the same form as T but of lower dimension.

Proof of Claim 4: (Reestablishing the original structure) For $i=2,\ldots,N$, $j=2,\ldots,M$ we define the matrices

$$X_i = \Big(x[i]_{\nu\mu}\Big)_{\nu,\mu=1}^{\widehat{n}_i} \in \mathcal{M}_{\widehat{n}_i,\widehat{n}_i}(\mathbb{C}) \quad \text{with} \quad x[i]_{\nu\mu} = \begin{cases} 0, & \mu>\nu,\\ x_i^{\nu-\mu}, & \mu\le\nu,\end{cases}$$

$$Y_j = \Big(y[j]_{\nu\mu}\Big)_{\nu,\mu=1}^{\widehat{m}_j} \in \mathcal{M}_{\widehat{m}_j,\widehat{m}_j}(\mathbb{C}) \quad \text{with} \quad y[j]_{\nu\mu} = \begin{cases} 0, & \mu>\nu,\\ y_j^{\nu-\mu}, & \mu\le\nu,\end{cases}$$

where $x_i = -\dfrac{1}{\alpha_i+\beta_1}\Phi_i^{-1}$, $y_j = -\dfrac{1}{\alpha_1+\beta_j}\Psi_j^{-1}$. Again $\det(X_i)=\det(Y_j)=1$, and thus the following manipulations are allowed,

$$\begin{aligned}
\det(T^{\langle 3\rangle}) &= \det\left(\begin{pmatrix} I & 0 & \cdots & 0\\ 0 & X_2 & \cdots & 0\\ \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & \cdots & X_N\end{pmatrix} T^{\langle 3\rangle} \begin{pmatrix} I & 0 & \cdots & 0\\ 0 & Y_2' & \cdots & 0\\ \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & \cdots & Y_M'\end{pmatrix}\right)\\
&= \det\begin{pmatrix} T_{11}^{\langle 3\rangle} & T_{1j}^{\langle 3\rangle}Y_j'\\ X_iT_{i1}^{\langle 3\rangle} & X_iT_{ij}^{\langle 3\rangle}Y_j'\end{pmatrix}_{\substack{i>1\\ j>1}}\\
&= \det(T^{\langle 4\rangle}).
\end{aligned}$$

Here $T^{\langle 4\rangle}$ has the blocks $T_{ij}^{\langle 4\rangle} = X_iT_{ij}^{\langle 3\rangle}Y_j'$, where, for the sake of convenience, we adopt the convention $X_1 = I \in \mathcal{M}_{\widehat{n}_1,\widehat{n}_1}(\mathbb{C})$ and $Y_1 = I \in \mathcal{M}_{\widehat{m}_1,\widehat{m}_1}(\mathbb{C})$. With regard to $0^0=1$ the latter can also be stated by $x_1=y_1=0$.

Calculating the entries of $T_{ij}^{\langle 4\rangle}$ we get

$$\begin{aligned}
\left(\frac{1}{\alpha_i+\beta_j}\right)^{\nu+\mu-1} t[i,j]_{\nu\mu}^{\langle 4\rangle} &= \sum_{\lambda=1}^{\nu}\sum_{\kappa=1}^{\mu}\left(\frac{1}{\alpha_i+\beta_j}\right)^{\lambda+\kappa-1} x_i^{\nu-\lambda}\ t[i,j]_{\lambda\kappa}^{\langle 3\rangle}\ y_j^{\mu-\kappa}\\
= \left(\frac{1}{\alpha_i+\beta_j}\right)^{\nu+\mu-1} &\sum_{\lambda=1}^{\nu}\sum_{\kappa=1}^{\mu}\Big((\alpha_i+\beta_j)x_i\Big)^{\nu-\lambda}\Big((\alpha_i+\beta_j)y_j\Big)^{\mu-\kappa} t[i,j]_{\lambda\kappa}^{\langle 3\rangle}.
\end{aligned}$$

With $p_{ij} := (\alpha_i+\beta_j)x_i$ and $q_{ij} := (\alpha_i+\beta_j)y_j$, the above identity rewrites as

$$t[i,j]_{\nu\mu}^{\langle 4\rangle} = \sum_{\lambda=1}^{\nu}\sum_{\kappa=1}^{\mu} p_{ij}^{\nu-\lambda}q_{ij}^{\mu-\kappa}t[i,j]_{\lambda\kappa}^{\langle 3\rangle}, \tag{6.4}$$

where, for later use, we note

$$p_{ij} = \begin{cases} 0, & i=1,\ j\ge 1,\\ -\Phi_i^{-1}, & i>1,\ j=1,\\ 1-\phi_{ij}^{-1}, & i>1,\ j>1,\end{cases} \qquad q_{ij} = \begin{cases} 0, & j=1,\ i\ge 1,\\ -\Psi_j^{-1}, & j>1,\ i=1,\\ 1-\psi_{ji}^{-1}, & j>1,\ i>1.\end{cases} \tag{6.5}$$

To evalulate (6.4), the following simple identity, which can be shown by an obvious argument with telescope sums, is helpful. For $\gamma \in \mathbb{C}$ and $S,\ R \in \mathbb{N}$,

$$\sum_{r=2}^{R} \gamma^{R-r} \left[\binom{S+r-3}{S-1} - \gamma^{-1}\binom{S+r-2}{S-1}\right] = \gamma^{R-2} - \gamma^{-1}\binom{S+R-2}{S-1}. \tag{6.6}$$

We start with the calculation of (6.4) in the case $i = 1,\ j > 1$. Since $p_{1j} = 0,\ \Psi_j = -q_{1j}^{-1}$ for $j > 1$ by (6.5), we observe

$$\begin{aligned}
t[1,j]^{\langle 4\rangle}_{\nu\mu} &= \sum_{\kappa=1}^{\mu} q_{1j}^{\mu-\kappa} t[1,j]^{\langle 3\rangle}_{\nu\kappa} \\
&= q_{1j}^{\mu-1}\Psi_j + \sum_{\kappa=2}^{\mu} q_{1j}^{\mu-\kappa}\left[\binom{\nu+\kappa-3}{\nu-1} + \binom{\nu+\kappa-2}{\mu-1}\Psi_j\right] \\
&= -q_{1j}^{\mu-2} + \sum_{\kappa=2}^{\mu} q_{1j}^{\mu-\kappa}\left[\binom{\nu+\kappa-3}{\nu-1} - q_{1j}^{-1}\binom{\nu+\kappa-2}{\nu-1}\right] \\
&\overset{(6.6)}{=} -q_{1j}^{-1}\binom{\nu+\mu-2}{\nu-1} \\
&= \Psi_j\binom{\nu+\mu-2}{\nu-1}.
\end{aligned}$$

Analogously, in the case $i > 1,\ j = 1$ we get $t[i,1]^{\langle 4\rangle}_{\nu\mu} = \Phi_i\binom{\nu+\mu-2}{\nu-1}$.

The case $i > 1,\ j > 1$ is slightly more involved. Here $\phi_{ij} = (1-p_{ij})^{-1},\ \psi_{ji} = (1-q_{ij})^{-1}$, again by (6.5). Thus the entries $t[i,j]^{\langle 3\rangle}_{\lambda\kappa}$ we start from are represented as

$$\begin{aligned}
t[i,j]^{\langle 3\rangle}_{11} &= \phi_{ij}\psi_{ji} = \big(\,(1-p_{ij})(1-q_{ij})\,\big)^{-1}, \\
t[i,j]^{\langle 3\rangle}_{\lambda 1} &= \psi_{ji} = \big(\,(1-p_{ij})(1-q_{ij})\,\big)^{-1}\cdot(1-p_{ij}), \qquad \lambda > 1, \\
t[i,j]^{\langle 3\rangle}_{1\kappa} &= \phi_{ij} = \big(\,(1-p_{ij})(1-q_{ij})\,\big)^{-1}\cdot(1-q_{ij}), \qquad \kappa > 1, \\
t[i,j]^{\langle 3\rangle}_{\lambda\kappa} &= \binom{\lambda+\kappa-4}{\lambda-1}\phi_{ij} + \binom{\lambda+\kappa-4}{\kappa-1}\psi_{ji} + \binom{\lambda+\kappa-4}{\lambda-2}\left[1+\phi_{ij}\psi_{ji}\right] \\
&= \big((1-p_{ij})(1-q_{ij})\big)^{-1}\Big[\binom{\lambda+\kappa-2}{\lambda-1} - q_{ij}\binom{\lambda+\kappa-3}{\lambda-1} - p_{ij}\binom{\lambda+\kappa-3}{\lambda-2} + p_{ij}q_{ij}\binom{\lambda+\kappa-4}{\lambda-2}\Big], \\
&\qquad\qquad \lambda > 1, \kappa > 1,
\end{aligned}$$

the latter by the usual properties of binomial coefficients. Let $\mu > 1,\ \nu > 1$. Exploiting these representations, we obtain

$$\begin{aligned}
&(1-p_{ij})(1-q_{ij})\sum_{\kappa=2}^{\mu} q_{ij}^{\mu-\kappa}\, t[i,j]^{\langle 3\rangle}_{\lambda\kappa} \\
&= -q_{ij}\sum_{\kappa=2}^{\mu} q_{ij}^{\mu-\kappa}\Big[\binom{\lambda+\kappa-3}{\lambda-1} - q_{ij}^{-1}\binom{\lambda+\kappa-2}{\lambda-1}\Big] + p_{ij}q_{ij}\sum_{\kappa=2}^{\mu} q_{ij}^{\mu-\kappa}\Big[\binom{\lambda+\kappa-4}{\lambda-2} - q_{ij}^{-1}\binom{\lambda+\kappa-3}{\lambda-2}\Big] \\
&\overset{(6.6)}{=} -q_{ij}\Big(q_{ij}^{\mu-2} - q_{ij}^{-1}\binom{\lambda+\mu-2}{\lambda-1}\Big) + p_{ij}q_{ij}\Big(q_{ij}^{\mu-2} - q_{ij}^{-1}\binom{\lambda+\mu-3}{\lambda-2}\Big) \\
&= \Big[\binom{\lambda+\mu-2}{\lambda-1} - p_{ij}\binom{\lambda+\mu-3}{\lambda-2}\Big] - q_{ij}^{\mu-1}(1-p_{ij}) \\
&= \Big[\binom{\lambda+\mu-2}{\mu-1} - p_{ij}\binom{\lambda+\mu-3}{\mu-1}\Big] - q_{ij}^{\mu-1}(1-p_{ij}), \qquad \text{if } \lambda > 1.
\end{aligned}$$

Hence,

$$(1-p_{ij})(1-q_{ij})\sum_{\lambda=2}^{\nu}\sum_{\kappa=2}^{\mu} p_{ij}^{\nu-\lambda} q_{ij}^{\mu-\kappa}\, t[i,j]^{\langle 3\rangle}_{\lambda\kappa}$$

$$
\begin{aligned}
&= \quad -p_{ij}\sum_{\lambda=2}^{\nu} p_{ij}^{\nu-\lambda}\Big[\tbinom{\lambda+\mu-3}{\mu-1} - p_{ij}^{-1}\tbinom{\lambda+\mu-2}{\mu-1}\Big] + p_{ij}q_{ij}^{\mu-1}\sum_{\lambda=2}^{\nu} p_{ij}^{\nu-\lambda}(1-p_{ij}^{-1}) \\
&\overset{(6.6)}{=} \quad -p_{ij}\Big(p_{ij}^{\nu-2} - p_{ij}^{-1}\tbinom{\nu+\mu-2}{\mu-1}\Big) + p_{ij}q_{ij}^{\mu-1}\Big(p_{ij}^{\nu-2} - p_{ij}^{-1}\Big) \\
&= \quad \tbinom{\nu+\mu-2}{\mu-1} + \big(\, (1-p_{ij}^{\nu-1})(1-q_{ij}^{\mu-1})\,\big) - 1. \qquad (6.7)
\end{aligned}
$$

In addition, we immediately verify

$$
\begin{aligned}
&(1-p_{ij})(1-q_{ij})\Big(p_{ij}^{\nu-1}q_{ij}^{\mu-1}t[i,j]_{11}^{\langle 3\rangle} + q_{ij}^{\mu-1}\sum_{\lambda=2}^{\nu} p_{ij}^{\nu-\lambda}\, t[i,j]_{\lambda 1}^{\langle 3\rangle} + p_{ij}^{\nu-1}\sum_{\kappa=2}^{\mu} q_{ij}^{\mu-\kappa}\, t[i,j]_{1\kappa}^{\langle 3\rangle}\Big) \\
&= \quad p_{ij}^{\nu-1}q_{ij}^{\mu-1} + q_{ij}^{\mu-1}\sum_{\lambda=2}^{\nu} p_{ij}^{\nu-\lambda}(1-p_{ij}) + p_{ij}^{\nu-1}\sum_{\kappa=2}^{\mu} q_{ij}^{\mu-\kappa}(1-q_{ij}) \\
&= \quad -(1-p_{ij}^{\nu-1})(1-q_{ij}^{\mu-1}) + 1. \qquad (6.8)
\end{aligned}
$$

Consequently, inserting (6.7), (6.8) into (6.4) yields $(1-p_{ij})(1-q_{ij})\, t[i,j]_{\lambda\kappa}^{\langle 4\rangle} = \binom{\nu+\mu-2}{\nu-1}$, which by (6.2), (6.5) finally shows

$$
t[i,j]_{\lambda\kappa}^{\langle 4\rangle} \;=\; \phi_{ij}\psi_{ji}\tbinom{\nu+\mu-2}{\nu-1} \;=\; \Phi_i\Psi_j\tbinom{\nu+\mu-2}{\nu-1}.
$$

To sum up,

$$
T_{11}^{\langle 4\rangle} = \widehat{T}_{11}, \qquad \begin{array}{ll} T_{i1}^{\langle 4\rangle} = \Phi_i\widehat{T}_{i1}, & i>1, \\ T_{1j}^{\langle 4\rangle} = \Psi_j\widehat{T}_{1j}, & j>1, \end{array} \qquad T_{ij}^{\langle 4\rangle} = \Phi_i\Psi_j\widehat{T}_{ij}, \quad i>1, j>1,
$$

and Claim 4 follows by extracting common factors.

Induction with respect to the dimension n: We conclude by carrying out the induction argument. To this end, assume that the assertion holds for all dimensions less then n. By Claim 1 to Claim 4,

$$
\det(T) \;=\; \left[\frac{1}{\alpha_1+\beta_1}\right]^{n_1+m_1-1} \prod_{i=2}^{N}\left[\frac{\alpha_1-\alpha_i}{\alpha_i+\beta_1}\right]^{n_i} \prod_{j=2}^{M}\left[\frac{\beta_1-\beta_j}{\alpha_1+\beta_j}\right]^{m_j} \quad \det(\widehat{T}), \qquad (6.9)
$$

where $\widehat{T}\in\mathcal{M}_{n-1,n-1}(\mathbb{C})$ is of the same structure as T. Thus, by assumption,

$$
\begin{aligned}
\det(\widehat{T}) &= \prod_{\substack{i<j\\ i,j=1}}^{N}(\alpha_i-\alpha_j)^{\widehat{n}_i\widehat{n}_j} \prod_{\substack{i<j\\ i,j=1}}^{M}(\beta_i-\beta_j)^{\widehat{m}_i\widehat{m}_j} \Big/ \prod_{i=1}^{N}\prod_{j=1}^{M}(\alpha_i+\beta_j)^{\widehat{n}_i\widehat{m}_j} \\
&= \quad \Delta\left[\frac{1}{\alpha_1+\beta_1}\right]^{(n_1-1)(m_1-1)} \prod_{i=2}^{N}\left[\frac{(\alpha_1-\alpha_i)^{n_1-1}}{(\alpha_i+\beta_1)^{m_1-1}}\right]^{n_i} \prod_{j=2}^{M}\left[\frac{(\beta_1-\beta_j)^{m_1-1}}{(\alpha_1+\beta_j)^{n_1-1}}\right]^{m_j} \qquad (6.10) \\
&\qquad \text{for } \Delta = \prod_{\substack{i,j=2\\ i<j}}^{N}(\alpha_i-\alpha_j)^{n_in_j} \prod_{\substack{i,j=2\\ i<j}}^{M}(\beta_i-\beta_j)^{m_im_j} \Big/ \prod_{i=2}^{N}\prod_{j=2}^{M}(\alpha_i+\beta_j)^{n_im_j}.
\end{aligned}
$$

Inserting (6.10) into (6.9) immediately yields the desired formula for $\det(T)$.

Therefore, Theorem 6.1.1 is shown. $\square$

6.2 Determinants of double size with a one-dimensional perturbation

Now we are in position to fill in the last remaining gap in Theorem 5.1.2. The following theorem contains the determination of the phase-shifts.

Theorem 6.2.1. *Assume that α_i, $i = 1, \ldots, N$, and β_j, $j = 1, \ldots, M$, are complex numbers with the property $\alpha_i + \beta_j \neq 0$ for all i, j. Let n_i, $i = 1, \ldots, N$, and m_j, $j = 1, \ldots, M$, be natural numbers, and set $n = \sum_{i=1}^{N} n_i$, $m = \sum_{j=1}^{M} m_j$.*

Define the matrix $T = (T_{ij})_{\substack{i=1,\ldots,N \\ j=1,\ldots,M}} \in \mathcal{M}_{n,m}(\mathbb{C})$ with the blocks

$$T_{ij} = \left(\left(\frac{1}{\alpha_i + \beta_j} \right)^{\nu+\mu-1} \binom{\nu+\mu-2}{\nu-1} \right)_{\substack{\nu=1,\ldots,n_i \\ \mu=1,\ldots,m_j}} \in \mathcal{M}_{n_i,m_j}(\mathbb{C})$$

and the vector $f = (e_{m_1}^{(1)}, \ldots, e_{m_M}^{(1)})' \in \mathbb{C}^m$ consisting of the first standard basis vectors $e_{m_j}^{(1)} \in \mathbb{C}^{m_j}$ for $j = 1, \ldots, M$.

If $m \in \{n, n+1\}$, then

$$\det \begin{pmatrix} 0 & -T \\ T' & f \otimes f \end{pmatrix} = \prod_{\substack{i<j \\ i,j=1}}^{N} (\alpha_i - \alpha_j)^{2n_in_j} \prod_{\substack{i<j \\ i,j=1}}^{M} (\beta_i - \beta_j)^{2m_im_j} \Big/ \prod_{i=1}^{N} \prod_{j=1}^{M} (\alpha_i + \beta_j)^{2n_im_j}.$$

Recall that T' denotes the transposed of T.

Proof As in the proof of Theorem 6.1.1, it suffices to consider the situation where $\alpha_i \neq \alpha_j$ for all $i, j = 1, \ldots, N$, $i \neq j$, and $\beta_i \neq \beta_j$ for all $i, j = 1, \ldots, M$, $i \neq j$.

To start with, we consider the case $m = n$. Then T is a square matrix, and invertible by Theorem 6.1.1. By Lemma 4.1.8

$$\det \begin{pmatrix} 0 & -T \\ T' & f \otimes f \end{pmatrix} = \det \begin{pmatrix} I & 0 \\ -(f \otimes f)T^{-1} & I \end{pmatrix} \det \begin{pmatrix} 0 & -T \\ T' & 0 \end{pmatrix} = \det (T)^2.$$

Thus our task is reduced to Theorem 6.1.1 and the assertion follows. It remains to treat the case $m = n + 1$.

In the case $m = n + 1$, the proof follows to some extent the arguments of the proof of Theorem 6.1.1. For convenience we premise an outline of the main steps.

First, to simplify the argument, we write

$$\det \begin{pmatrix} 0 & -T \\ T' & f \otimes f \end{pmatrix} = (-1)^n \det(S), \quad \text{where } S = \begin{pmatrix} 0 & T \\ T' & f \otimes f \end{pmatrix}.$$

As for the calculation of $\det(S)$, we proceed as follows:

Step 1: (Preparational manipulations) First we apply the strategy developed in Claim 1 and Claim 2 of the proof of Theorem 6.1.1 to the block T in the upper right corner of S, where we only have to pay attention to the fact that T is no longer a square matrix. Secondly we perform the transposed strategy with respect to the block T' in the lower left corner of S.

If the manipulations applied to T are written as matrix multiplication XTY' with $X \in \mathcal{M}_{n,n}(\mathbb{C})$, $Y \in \mathcal{M}_{m,m}(\mathbb{C})$, then the transposed manipulations amount to $YT'X'$. As a consequence,

$$\begin{aligned} \det(S) &= \det\left(\begin{pmatrix} X & 0 \\ 0 & Y \end{pmatrix} \begin{pmatrix} 0 & T \\ T' & f \otimes f \end{pmatrix} \begin{pmatrix} X' & 0 \\ 0 & Y' \end{pmatrix} \right) \\ &= \det \begin{pmatrix} 0 & XTY' \\ (XTY')' & (Yf) \otimes (Yf) \end{pmatrix}. \end{aligned} \tag{6.11}$$

Obviously these manipulations do not change the zero block in the upper left corner of S, but we have to check their effect on the one-dimensional perturbation $f \otimes f$ in the lower right corner of S.

Step 2: (Expansion of the determinant) By the proof of Theorem 6.1.1, the block XTY' has zero entries in the first row with the only exception of the $(1,1)$-entry. In analogy, the block $(XTY')'$ has zero entries in the first column, again except of the $(1,1)$-entry.

Thus we can expand $\det(S)$, with respect to (i) the first column and (ii) the first row, reducing both dimensions n, m of the problem by one. As a result we obtain

$$\det(S) = -\lambda \det(\widehat{S}) \qquad \text{with } \lambda \in \mathbb{C} \text{ and } \widehat{S} \in \mathcal{M}_{n+m-2,n+m-2}(\mathbb{C}).$$

Step 3: (Reestablishing the original structure) In the last step we prove that

$$\det(\widehat{S}) = \widehat{\lambda} \det \begin{pmatrix} 0 & \widehat{T} \\ \widehat{T}' & \widehat{f} \otimes \widehat{f} \end{pmatrix},$$

where $\widehat{T} \in \mathcal{M}_{n-1,m-1}(\mathbb{C})$, $\widehat{f} \in \mathbb{C}^{m-1}$ are of the same structure as in Theorem 6.2.1, and $\widehat{\lambda} \in \mathbb{C}$. To this end, we use the strategy developed in Claim 4 of the proof of Theorem 6.1.1 in the same manner as in Step 1.

Summing up the content of Step 1 to Step 3, we observe

$$\begin{aligned} \det \begin{pmatrix} 0 & -T \\ T' & f \otimes f \end{pmatrix} &= (-1)^n \det(S) \;=\; (-1)^{n+1} \lambda \widehat{\lambda} \det(\widehat{S}) \\ &= \lambda \widehat{\lambda} \det \begin{pmatrix} 0 & -\widehat{T} \\ \widehat{T}' & \widehat{f} \otimes \widehat{f} \end{pmatrix}. \end{aligned}$$

The result then follows by induction.

Recall that for $T = \left(T_{ij} \right)_{\substack{i=1,\dots,N \\ j=1,\dots,M}}$ (and analogously for $f = \left(f_j \right)_{j=1}^{M}$) we often use the notation

$$\begin{pmatrix} T_{11} & T_{1j} \\ T_{i1} & T_{ij} \end{pmatrix}_{\substack{i>1 \\ j>1}} \qquad \text{or} \qquad \begin{pmatrix} f_1, & f_j \end{pmatrix}_{j>1} \qquad \text{respectively,}$$

if the blocks $T_{ij} \in \mathcal{M}_{n_i m_j}(\mathbb{C})$ with $i > 1$, $j > 1$ (or the vector $f_j \in \mathbb{C}^{m_j}$ with $j > 1$) are treated separately from the others.

Let us now enter the proof. In the sequel we will use the notation $X_i^{\langle k \rangle}$, $Y_j^{\langle k \rangle}$ for the matrices X_i, Y_j used in the proof of the k-th claim of Theorem 6.1.1. Note that their dimension depends on k, for example $Y_j^{\langle 1 \rangle} \in \mathcal{M}_{m_j,m_j}(\mathbb{C})$ for $j = 1, \dots, M$ and $Y_j^{\langle 2 \rangle} \in \mathcal{M}_{m_j,m_1}(\mathbb{C})$ for $j = 2, \dots, M$.

To keep the presentation as clear as possible, we gather the $Y_j^{\langle k\rangle}$ in one common matrix $Y^{\langle k\rangle}$. Analogously, the matrix $X^{\langle k\rangle}$ collects the $X_j^{\langle k\rangle}$. For example,

$$Y^{\langle 1\rangle} = \begin{pmatrix} Y_1^{\langle 1\rangle} & & 0 \\ & \ddots & \\ 0 & & Y_M^{\langle 1\rangle} \end{pmatrix}, \qquad Y^{\langle 2\rangle} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ Y_2^{\langle 2\rangle} & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ Y_M^{\langle 2\rangle} & 0 & \cdots & 1 \end{pmatrix}. \tag{6.12}$$

Recall $\det(X^{\langle k\rangle}) = \det(Y^{\langle k\rangle}) = 1$ for all k.

Note that the manipulations with respect to rows and columns used in the proof of Theorem 6.1.1 apply for n, m arbitrary. The fact that T is a square matrix was only needed for the existence of $\det(T)$. Hence we can be brief in the following arguments.

Step 1: With $X = X^{\langle 2\rangle}X^{\langle 1\rangle}$, $Y = Y^{\langle 2\rangle}Y^{\langle 1\rangle}$, we can apply (6.11), since $\det(X) = \det(Y) = 1$, and it follows

$$\det(S) = \det\begin{pmatrix} 0 & T^{\langle 2\rangle} \\ (T^{\langle 2\rangle})' & f^{\langle 2\rangle} \otimes f^{\langle 2\rangle} \end{pmatrix},$$

where $T^{\langle 2\rangle}$ is the matrix obtained in Claim 1 and Claim 2 of the proof of Theorem 6.1.1, and $f^{\langle 2\rangle} = Yf$. With $f = (f_j)_{j=1}^M$, we infer by (6.12)

$$f^{\langle 2\rangle} = \left(Y_1^{\langle 1\rangle}f_1, \quad Y_j^{\langle 1\rangle}f_j \;+\; Y_j^{\langle 2\rangle}Y_1^{\langle 1\rangle}f_1\right)_{j>1}.$$

Inserting the concrete forms for $Y_j^{\langle 1\rangle}$, $Y_j^{\langle 2\rangle}$, as given in the proofs of Claim 1 and Claim 2, Theorem 6.1.1, and $f_j = e_{m_j}^{(1)}$, we get

$$\begin{aligned} Y_j^{\langle 1\rangle}f_j &= e_{m_j}^{(1)} - \frac{1}{\alpha_1+\beta_j}e_{m_j}^{(2)}, & j &= 1,\dots,M, \\ Y_j^{\langle 2\rangle}Y_1^{\langle 1\rangle}f_1 &= -\frac{\alpha_1+\beta_1}{\alpha_1+\beta_j}e_{m_j}^{(1)}, & j &= 2,\dots,M, \end{aligned}$$

where, for simplicity, we consider the κ-th standard basis vector $e_k^{(\kappa)} \in \mathbb{C}^k$ as non-existent for $k < \kappa$. In summary,

$$f^{\langle 2\rangle} = \left(e_{m_1}^{(1)} - \frac{1}{\alpha_1+\beta_1}e_{m_1}^{(2)}, \qquad -\frac{\beta_1-\beta_j}{\alpha_1+\beta_j}e_{m_j}^{(1)} - \frac{1}{\alpha_1+\beta_j}e_{m_j}^{(2)}\right)_{j>1}. \tag{6.13}$$

The key point is that, except of the $(1,1)$-entry, $T^{\langle 2\rangle}$ has only zero entries in the first row. Similarly, $(T^{\langle 2\rangle})'$ has only zero entries in the first column except of the $(1,1)$-entry. The value of these $(1,1)$-entries is $(\alpha_1+\beta_1)^{-1}$.

Step 2: Set $\gamma = (\alpha_1+\beta_1)^{-1}$. In the proof of Claim 3, Theorem 6.1.1, we have shown that the matrix $T^{\langle 2\rangle}$ is of the form

$$T^{\langle 2\rangle} = \begin{pmatrix} \gamma & 0 & 0 \\ * & \gamma^2 T_{11}^{\langle 3\rangle} & \gamma T_{1j}^{\langle 3\rangle} \\ * & \gamma T_{i1}^{\langle 3\rangle} & T_{ij}^{\langle 3\rangle} \end{pmatrix}_{\substack{i>1 \\ j>1}}, \qquad \begin{array}{l} T_{1j}^{\langle 3\rangle} \text{ with } n_1 - 1 \text{ rows } (\forall j), \\ T_{i1}^{\langle 3\rangle} \text{ with } m_1 - 1 \text{ columns } (\forall i), \end{array}$$

with $T_{ij}^{\langle 3\rangle}$ as defined in Claim 3 of the proof of Theorem 6.1.1.

Moreover, by (6.13) we have obtained $f^{\langle 2\rangle} = (\, f_j^{\langle 2\rangle} \,)_{j=1}^{M}$, where $f_1^{\langle 2\rangle} = (1, -\gamma, 0, \dots, 0) = (1, -\gamma e^{(1)}_{m_1-1})$. Let us define

$$f_j^{\langle 3\rangle} = \begin{cases} e^{(1)}_{m_1-1}, & j = 1, \\ -f_j^{\langle 2\rangle} \;=\; \dfrac{\beta_1 - \beta_j}{\alpha_1 + \beta_j} e^{(1)}_{m_j} + \dfrac{1}{\alpha_1 + \beta_j} e^{(2)}_{m_j}, & j > 1. \end{cases} \tag{6.14}$$

Then $f^{\langle 2\rangle} = -\,(\, -1, \quad \gamma f_1^{\langle 3\rangle}, \quad f_j^{\langle 3\rangle} \,)_{j>1}$.

Since $f^{\langle 2\rangle} \otimes f^{\langle 2\rangle} = (-f^{\langle 2\rangle}) \otimes (-f^{\langle 2\rangle})$, the one-dimensional block becomes

$$f^{\langle 2\rangle} \otimes f^{\langle 2\rangle} = \begin{pmatrix} 1 & * & * \\ * & \gamma^2\, f_1^{\langle 3\rangle} \otimes f_1^{\langle 3\rangle} & \gamma\, f_j^{\langle 3\rangle} \otimes f_1^{\langle 3\rangle} \\ * & \gamma\, f_1^{\langle 3\rangle} \otimes f_i^{\langle 3\rangle} & f_j^{\langle 3\rangle} \otimes f_i^{\langle 3\rangle} \end{pmatrix}_{\substack{i>1 \\ j>1}} .$$

Therefore[1], the expansion results in

$$\det(S) = -\left(\frac{1}{\alpha_1 + \beta_1}\right)^{2(n_1 + m_1 - 1)} \det(\widehat{S}) \qquad \text{with } \widehat{S} = \begin{pmatrix} 0 & T^{\langle 3\rangle} \\ (T^{\langle 3\rangle})' & f^{\langle 3\rangle} \otimes f^{\langle 3\rangle} \end{pmatrix},$$

where, as usual, $f^{\langle 3\rangle} = (\, f_j^{\langle 3\rangle} \,)_{j=1}^{M}$.

Step 3: To reestablish the original structure of the determinant, we apply the manipulations used in Claim 4 of the proof of Theorem 6.1.1. Arguing analogously to (6.11), we infer

$$\det(\widehat{S}) = \det(S^{\langle 4\rangle}) \qquad \text{with} \qquad S^{\langle 4\rangle} = \begin{pmatrix} 0 & T^{\langle 4\rangle} \\ (T^{\langle 4\rangle})' & f^{\langle 4\rangle} \otimes f^{\langle 4\rangle} \end{pmatrix}$$

where, by Claim 4 in the proof of Theorem 6.1.1, $T^{\langle 4\rangle}$ differs from a matrix $\widehat{T}$ of the same structure as in the assertion only by certain factors. To be precise, $T^{\langle 4\rangle}$ has the blocks

$$T_{11}^{\langle 4\rangle} = \widehat{T}_{11}, \qquad \begin{array}{ll} T_{i1}^{\langle 4\rangle} = \Phi_i \widehat{T}_{i1}, & i > 1, \\ T_{1j}^{\langle 4\rangle} = \Psi_j \widehat{T}_{1j}, & j > 1, \end{array} \qquad T_{ij}^{\langle 4\rangle} = \Phi_i \Psi_j \widehat{T}_{ij}, \quad i > 1, j > 1,$$

with Φ_i, Ψ_j defined as in (6.2) and $\widehat{T} = (\, \widehat{T}_{ij} \,)_{\substack{i=1,\dots,N \\ j=1,\dots,M}}$ as in the assertion of Claim 4.

Moreover,

$$f^{\langle 4\rangle} = Y^{\langle 4\rangle} f^{\langle 3\rangle}.$$

[1] For illustration, the determinant to be expanded has the form

$$\left|\begin{array}{ccc|ccc} 0 & 0 & 0 & \gamma & 0 & 0 \\ 0 & 0 & 0 & * & \gamma^2\, T_{11}^{\langle 3\rangle} & \gamma\, T_{1j}^{\langle 3\rangle} \\ 0 & 0 & 0 & * & \gamma\, T_{i1}^{\langle 3\rangle} & T_{ij}^{\langle 3\rangle} \\ \hline \gamma & * & * & 1 & * & * \\ 0 & \gamma^2\, (T_{11}^{\langle 3\rangle})' & \gamma\, (T_{j1}^{\langle 3\rangle})' & * & \gamma^2\, f_1^{\langle 3\rangle} \otimes f_1^{\langle 3\rangle} & \gamma\, f_j^{\langle 3\rangle} \otimes f_1^{\langle 3\rangle} \\ 0 & \gamma\, (T_{1i}^{\langle 3\rangle})' & (T_{ji}^{\langle 3\rangle})' & * & \gamma\, f_1^{\langle 3\rangle} \otimes f_i^{\langle 3\rangle} & f_j^{\langle 3\rangle} \otimes f_i^{\langle 3\rangle} \end{array}\right|$$

Inserting $Y^{\langle 4\rangle} = \mathrm{diag}\{Y_j^{\langle 4\rangle} | j = 1, \ldots, M\}$ with $Y_1^{\langle 4\rangle} = 1$, and using (6.2), (6.14), we observe

$$f^{\langle 4\rangle} = \left(e^{(1)}_{m_1-1}, \quad Y_j^{\langle 4\rangle}\left[\Psi_j e^{(1)}_{m_j} + \frac{1}{\alpha_1+\beta_j} e^{(2)}_{m_j}\right]\right)_{j>1}.$$

As in Claim 4, set $y_j = -\left((\alpha_1+\beta_j)\Psi_j \right)^{-1}$. Then, from the concrete form of $Y_j^{\langle 4\rangle}$, $j = 2, \ldots, M$, in Claim 4, we immediately find

$$\begin{aligned} Y_j^{\langle 4\rangle}\left[\Psi_j e^{(1)}_{m_j} + \frac{1}{\alpha_1+\beta_j} e^{(2)}_{m_j}\right] &= \Psi_j e^{(1)}_{m_j} + \sum_{\mu=2}^{m_j}\left(\Psi_j y_j^{m_j-1} + \frac{1}{\alpha_1+\beta_j} y_j^{m_j-2}\right) e^{(\mu)}_{m_j} \\ &= \Psi_j e^{(1)}_{m_j} + \frac{1}{\alpha_1+\beta_j}\sum_{\mu=2}^{m_j}\left(-\frac{1}{y_j} y_j^{m_j-1} + y_j^{m_j-2}\right) e^{(\mu)}_{m_j} \\ &= \Psi_j e^{(1)}_{m_j}. \end{aligned}$$

As a consequence,

$$f^{\langle 4\rangle} = \left(e^{(1)}_{m_1-1}, \quad \Psi_j e^{(1)}_{m_j}\right)_{j>1} = \left(\widehat{f}_1, \quad \Psi_j \widehat{f}_j \right)_{j>1},$$

if we define $\widehat{f}_1 = e^{(1)}_{m_1-1}$ and $\widehat{f}_j = e^{(1)}_{m_j}$. In addition, set $\widehat{f} = (\widehat{f}_j)_{j=1}^M$.

Therefore[2], we can extract the factor Φ_i from the n_i rows, $i = 2, \ldots, N$, of the blocks in the upper right corner of $S^{\langle 4\rangle}$ and the factor Ψ_i from the m_i rows, $i = 2, \ldots, M$, of the blocks in the lower left corner of $S^{\langle 4\rangle}$. Similarly, we extract the factor Ψ_j from the m_j columns, $j = 2, \ldots, M$, of the blocks in the upper right corner of $S^{\langle 4\rangle}$ and the factor Φ_j from the n_j columns, $j = 2, \ldots, N$, of the blocks in the lower left corner of $S^{\langle 4\rangle}$.

As a result,

$$\det(\widehat{S}) = \prod_{i=2}^{N}\Phi_i^{2n_i}\prod_{j=2}^{M}\Psi_j^{2m_j} \quad \det\begin{pmatrix} 0 & \widehat{T} \\ \widehat{T}' & \widehat{f}\otimes\widehat{f}\end{pmatrix}.$$

Result of Step1 to Step 3: Let us sum up what we have achieved so far. Namely, starting from the original matrix, we get

$$\begin{aligned} \det\begin{pmatrix} 0 & -T \\ T' & f\otimes f\end{pmatrix} &= (-1)^n \det(S) \\ &\overset{\text{Step 2}}{=} (-1)^{n+1}\left[\frac{1}{\alpha_1+\beta_1}\right]^{2(n_1+m_1-1)} \det(\widehat{S}) \\ &\overset{\text{Step 3}}{=} (-1)^{n+1}\left[\frac{1}{\alpha_1+\beta_1}\right]^{2(n_1+m_1-1)} \prod_{i=2}^{N}\Phi_i^{2n_i}\prod_{j=2}^{M}\Psi_j^{2m_j} \det\begin{pmatrix} 0 & \widehat{T} \\ \widehat{T}' & \widehat{f}\otimes\widehat{f}\end{pmatrix} \\ &= \left[\frac{1}{\alpha_1+\beta_1}\right]^{2(n_1+m_1-1)} \prod_{i=2}^{N}\Phi_i^{2n_i}\prod_{j=2}^{M}\Psi_j^{2m_j} \quad \det\begin{pmatrix} 0 & -\widehat{T} \\ \widehat{T}' & \widehat{f}\otimes\widehat{f}\end{pmatrix}. \end{aligned}$$

---

[2] For illustration, in summary we now have observed

$$|S^{\langle 4\rangle}| = \begin{vmatrix} 0 & T^{\langle 4\rangle} \\ (T^{\langle 4\rangle})' & f^{\langle 4\rangle}\otimes f^{\langle 4\rangle}\end{vmatrix} = \left|\begin{array}{cc|cc} 0 & 0 & \widehat{T}_{11} & \Psi_j\widehat{T}_{1j} \\ 0 & 0 & \Phi_i\widehat{T}_{i1} & \Phi_i\Psi_j\widehat{T}_{ij} \\ \hline \widehat{T}'_{11} & \Phi_j\widehat{T}'_{j1} & \widehat{f}_1\otimes\widehat{f}_1 & \Psi_j\widehat{f}_j\otimes\widehat{f}_1 \\ \Psi_i\widehat{T}'_{1i} & \Psi_i\Phi_j\widehat{T}'_{ji} & \Psi_i\widehat{f}_1\otimes\widehat{f}_i & \Psi_i\Psi_j\widehat{f}_j\otimes\widehat{f}_i \end{array}\right|$$

Inserting (6.2), we in summary have proved

$$\det \begin{pmatrix} 0 & -T \\ T' & f \otimes f \end{pmatrix} = \tag{6.15}$$
$$= \left[\frac{1}{\alpha_1+\beta_1}\right]^{2(n_1+m_1-1)} \prod_{i=2}^{N} \left[\frac{\alpha_1-\alpha_i}{\alpha_i+\beta_1}\right]^{2n_i} \prod_{j=2}^{M} \left[\frac{\beta_1-\beta_j}{\alpha_1+\beta_j}\right]^{2m_j} \det \begin{pmatrix} 0 & -\widehat{T} \\ \widehat{T}' & \widehat{f} \otimes \widehat{f} \end{pmatrix},$$

where the latter matrix is of the same structure as in the assertion, but of lower dimension.

Induction: Comparison to the proof of Theorem 6.1.1 shows that the induction step with respect to (6.15) can be carried over almost literally. The only difference is an additional square appearing in the factors.

This completes the proof. □

Chapter 7

Countable superpositions of solitons and negatons for the reduced AKNS systems

It is a characteristic property for soliton equations that solutions with particle-like behaviour can be combined to solutions representing the interaction of finitely many particles by 'nonlinear superposition'. Hence it is quite naturally to ask whether it is possible to superpose also countably many particles.

For solitons, the problem is the following: Given a sequence

$$\alpha_1, \quad \alpha_2, \quad \ldots, \quad \alpha_N, \Big| \, \alpha_{N+1}, \quad \ldots$$

the first N parameters correspond to an N-soliton $q^{(N)}$ (neglecting initial positions for the moment). Now one would like to find conditions on this sequence such that (i) the limit $q = \lim_{N\to\infty} q^{(N)}$ exists and (ii) q still solves the reduced AKNS system.

The study of the question was initiated by Gesztesy, Karwowski, and Zhao. In [38], [39] they were able to find for the KdV equation sufficient conditions for the convergence of countable superpositions of solitons. The passage to the limit is achieved by hard analysis going through the complete inverse scattering method. Related results were established for the Toda lattice in [40] and the KP and mKP equation in [81].

The main novelties of this chapter are the following: First we prove the existence of countable superpositions for the $\mathbb{C}$- and $\mathbb{R}$-reduced AKNS systems. This is done in full generality, i.e., for any admissible choice of f_0. Secondly this is not only done for solitons (as in all preceding references) but also for negatons and breathers. Since negatons are regular for both the $\mathbb{C}$- and $\mathbb{R}$-reduced AKNS systems (indeed even treatable by the inverse scattering method), this result is particularly interesting.

Furthermore we observe that, working on the general AKNS system, one cannot hope for optimal conditions, simply because the integral terms require strong decay conditions for $x \to -\infty$. We will show that for individual equations for which the integral terms cancel the superposition results can be considerably sharpened. In Sections 7.5.4, 7.6.2 this will be done for the Nonlinear Schrödinger and the modified Korteweg-de Vries equations. In the former case countable superpositions appear for the first time in the literature. On the level of individual equations we have established results about countable superpositions of solitons before ([8] for the Korteweg-de Vries equation, [88] for the sine-Gordon equation, [18] for the Kadomtsev-Petviashvili equation, and [89] for the Toda lattice). It can be shown (see [19], [89]) that the corresponding theorems in [38], [40], [81] are covered by our results.

We give the full program for the $\mathbb{C}$-reduced AKNS system. The first essential step is to derive a solution formula for generating operators A on sequence spaces. Sequence spaces are not only the natural framework to treat countable superpositions, but also have the advantage that a bounded operator T has a canonical complex conjugate $\overline{T}$, which is necessary to achieve the reduction. At first glance one could be tempted to do this with adjoint operators on Hilbert spaces, but this would contradict the essence of our approach which is, very roughly speaking, that ℓ_1 possesses the 'best' determinant.

The use of generalized diagonal operators A to realize countable superpositions of negatons is clearly motivated by Chapter 5. Then the decisive condition on shape and velocity of the appearing negatons is encoded in $\operatorname{spec}(A)$ and the multiplicity of the eigenvalues. Our main result in this context is Theorem 7.5.5 which proves, for the full $\mathbb{C}$-reduced AKNS system, the existence of countable superpositions of negatons, if the real parts of the eigenvalues are positive, bounded away from zero, and the size of the negatons (the number of their solitons, which coincides with the multiplicity of the corresponding eigenvalue) is bounded. In this case we have $0 \notin \operatorname{spec}(A) + \operatorname{spec}(\overline{A})$ and the theorem of Eschmeier and Dash/Schechter [22], [28] yields a systematic and satisfying method to extract scalar solutions.

On the other hand it is known from the results about individual equations, that countable superposition is already possible under weaker assumptions than $0 \notin \operatorname{spec}(A)+\operatorname{spec}(\overline{A})$. In Theorem 7.5.10 we present the strengthened result for the Nonlinear Schrödinger equation. Geometrically, we can drop the assumption that the 'width' of the waves is bounded. In the proof we can no longer use the Eschmeier, Dash/Schechter theorem, but have to produce one-dimensionality by using factorization techniques in the spirit of the Grothendieck theorem. Roughly speaking the idea is to exploit convergence not in the original, but in appropriate intermediate spaces.

As usual we obtain nicer formulas for the $\mathbb{R}$-reductions. These are gathered in Section 7.6. In addition we will explain how to integrate breathers into the picture. Finally we state sharpened results for the modified Korteweg-de Vries equation.

On the various levels outlined above we will provide a construction called universal realization. It tells that all our solutions can a posteriori reformulated in terms of the nuclear determinant on ℓ_1. This eventually proves that our superpositions are limits in the usual sense. In particular, we see that the negatons encoded in A really appear in the solution and are not lost by some hidden cancellation. As corollaries we will obtain results on global regularity and reality (in the $\mathbb{R}$-reduced case).

7.1 Solution formulas for the $\mathbb{C}$-reduced AKNS system based on sequence spaces

Throughout this chapter we will work on sequence spaces, the appropriate framework for countable superposition. Note that, for the $\mathbb{C}$-reduced AKNS system, it is furthermore important to work on spaces where a canonical complex conjugate of operators is defined.

In this section we deduce the crucial solution formulas from the general formulas of Chapter 2. Together with those for the $\mathbb{C}$-reduced AKNS system, we also provide a solution formula for the Nonlinear Schrödinger equation which has the advantage to come without any growth conditions.

But first we gather some background material on traces and determinants on quasi-Banach operator ideals, which will play a decisive role in the sequel. In Appendix B the reader finds a concise introduction to this topic. For thorough information see [41], [73].

7.1.1 Preliminaries

To start with, we introduce the quasi-Banach ideal of operators factorizing over an L_1-space, a Hilbert space and an L_∞-space. Let E, F be arbitrary Banach spaces. An operator

$T \in \mathcal{L}(E,F)$ belongs to the ideal $\mathcal{L}_\infty \circ \mathcal{H} \circ \mathcal{L}_1$ if there exists a Hilbert space H, an L_1-space, and an L_∞-space such that T factors in the following manner

$$T = YSRX \quad \text{with}$$
$$X \in \mathcal{L}(E, L_1),\ R \in \mathcal{L}(L_1, H),\ S \in \mathcal{L}(H, L_\infty),\ \text{and } Y \in \mathcal{L}(L_\infty, F).$$

With respect to the quasi-norm

$$\| T \mid \mathcal{L}_\infty \circ \mathcal{H} \circ \mathcal{L}_1 \| = \inf \left\{ \| Y \| \cdot \| S \| \cdot \| R \| \cdot \| X \| \,\middle|\, T = YSRX \right\},$$

where the infimum is taken over all factorizations of T, $\mathcal{L}_\infty \circ \mathcal{H} \circ \mathcal{L}_1$ becomes a quasi-Banach operator ideal.

Proposition 7.1.1. *The quasi-Banach operator ideal $\mathcal{L}_\infty \circ \mathcal{H} \circ \mathcal{L}_1$ possesses a spectral determinant* $\det_\lambda$, *which is even continuous.*

The proof can be found in Aden/Carl [8]. It mainly bases on Grothendieck's theorem (see Pisier [74]) and an application of the deep result of White [100] (confer also Proposition B.2.5).

Let us next recall a well-known property of nuclear operators on the sequence space ℓ_1 (see for example Pietsch [72]). Any operator $T \in \mathcal{N}(\ell_1)$ can be expressed by an infinite matrix, say $T = (\tau_{ij})_{i,j=1}^\infty$, such that $\| T \mid \mathcal{N} \| = \sum_i \sup_j \mid \tau_{ij} \mid$. For the determinant $\det_{\mathcal{N}}(I+T)$ the following expansion holds

$$\det_{\mathcal{N}}(I+T) = 1 + \sum_{n=1}^{\infty} \alpha_n(T) \tag{7.1}$$
$$\text{where } \alpha_n(T) = \frac{1}{n!} \sum_{i_1=1}^{\infty} \cdots \sum_{i_n=1}^{\infty} \det \begin{pmatrix} \tau_{i_1 i_1} & \cdots & \tau_{i_1 i_n} \\ \vdots & & \vdots \\ \tau_{i_n i_1} & \cdots & \tau_{i_n i_n} \end{pmatrix}.$$

In the following chapter we shall mainly use three determinants: the above mentioned determinants on $\mathcal{N}(\ell_1)$ and $\mathcal{L}_\infty \circ \mathcal{H} \circ \mathcal{L}_1$, and furthermore the spectral determinant on $\mathcal{N}_{\frac{2}{3}}$.

Next we turn to more specific properties of these determinants for operators on sequence spaces. To this end, let E, F be classical sequence spaces, i.e., c_0 or ℓ_p $(1 \le p < \infty)$. First note that any operator $T \in \mathcal{L}(E,F)$ is represented by an infinite matrix $(\tau_{ij})_{i,j=1}^\infty$ in the sense that

$$T\xi = \Big(\sum_j \tau_{ij} \xi_j \Big)_i \quad \text{for } \xi = (\xi_i)_i.$$

Then the notion of the complex conjugate $\overline{T}$ of T can be canonically defined as the operator $\overline{T} \in \mathcal{L}(E,F)$ generated by the infinite matrix $(\overline{\tau}_{ij})_{i,j=1}^\infty$.

We need a lemma about compatibility of determinants with complex conjugation.

Lemma 7.1.2. *Assume that one of the following situations holds for a r-Banach operator ideal $\mathcal{A}$ $(0 < r \le 1)$ with determinant δ and a Banach space E:*

1. *$\mathcal{A}$ is $\mathcal{N}_{\frac{2}{3}}$ $(r = \frac{2}{3})$, δ the spectral determinant* $\det_\lambda$ *on $\mathcal{N}_{\frac{2}{3}}$, and E a classical sequence space.*
2. *$\mathcal{A}$ is $\mathcal{N}$ $(r = 1)$, δ the nuclear determinant* $\det_{\mathcal{N}}$ *on $\mathcal{N}$ restricted to the class of Banach spaces with approximation property (a.p.), and $E = \ell_1$.*

Then, for any $T \in \mathcal{A}(E)$, we have $\overline{T} \in \mathcal{A}(E)$, and it holds

$$\delta(I + \overline{T}) = \overline{\delta(I+T)}. \tag{7.2}$$

By the classical sequence spaces we mean one of the spaces c_0, ℓ_p, $1 \le p < \infty$.

Proof In both cases $\overline{T} \in \mathcal{A}(E)$ is clear from the definition of $\mathcal{A}$. Identity (7.2) follows in the first case from the fact that δ is spectral and the eigenvalues of $\overline{T}$ are conjugate to those of T. In the second case it is an immediate consequence of (7.1). $\square$

The argument for the first case gives also the following lemma.

Lemma 7.1.3. *If T, $\overline{T} \in \mathcal{L}_\infty \circ \mathcal{H} \circ \mathcal{L}_1(E)$, E a classical sequence space, then (7.2) holds for the spectral determinant* $\det_\lambda$ *on* $\mathcal{L}_\infty \circ \mathcal{H} \circ \mathcal{L}_1$.

As for the universal realization, we have to replace some given determinants by a universal choice of a determinant. To this end we need the following lemma, which is mainly a consequence of the closed graph theorem (see Aden/Carl [8]).

Lemma 7.1.4. *Let $\mathcal{A}$ and $\mathcal{B}$ be quasi-Banach operator ideals. If $\mathcal{A}(E) \subset \mathcal{B}(E)$ (E Banach space) and the finite rank operators $\mathcal{F}(E)$ are $\| \cdot \mid \mathcal{A} \|$-dense in $\mathcal{A}(E)$, then there is a unique continuous trace* $\mathrm{tr}_\mathcal{A}$ *on $\mathcal{A}(E)$ and, for any continuous trace τ on $\mathcal{B}(E)$, we have*

$$\tau(T) = tr_\mathcal{A}(T) \quad \textit{for all} \quad T \in \mathcal{A}(E).$$

Of course, the same assertion applies to determinants.

In particular, we note the following consequence.

Corollary 7.1.5. *Let T belong to $\mathcal{N}_{\frac{2}{3}}(\ell_1)$ or to $\mathcal{L}_\infty \circ \mathcal{H} \circ \mathcal{L}_1(\ell_1)$. Then $T \in \mathcal{N}(\ell_1)$, and the identity*

$$\det{}_\lambda(I+T) = \det{}_\mathcal{N}(I+T) \tag{7.3}$$

holds with $\det_\mathcal{N}$ *denoting the nuclear determinant on $\mathcal{N}$ restricted to the class of Banach spaces with a.p., and* $\det_\lambda$ *the spectral determinant on $\mathcal{N}_{\frac{2}{3}}$ or $\mathcal{L}_\infty \circ \mathcal{H} \circ \mathcal{L}_1$, respectively.*

For $\mathcal{N}_{\frac{2}{3}}$ the assumptions of Lemma 7.1.4 follow immediately from the definitions. For $\mathcal{L}_\infty \circ \mathcal{H} \circ \mathcal{L}_1$ this is quite non-trivial. To sketch the argument, $\mathcal{L}_\infty \circ \mathcal{H} \circ \mathcal{L}_1(\ell_1) \subset \mathcal{N}(\ell_1)$ follows from Grothendieck's theorem, which states that operators in $\mathcal{L}(L_\infty, L_1)$ and in $\mathcal{L}(L_1, L_2)$ are 2-summing (confer for example [52], [74]), and the fact that the product of 2-summing operators is nuclear (see [72]). To observe that the finite rank operators $\mathcal{F}(\ell_1)$ on ℓ_1 are $\| \cdot \mid \mathcal{L}_\infty \circ \mathcal{H} \circ \mathcal{L}_1 \|$-dense in $\mathcal{L}_\infty \circ \mathcal{H} \circ \mathcal{L}_1(\ell_1)$, the idea is (i) to factorize the 2-summing operator in $\mathcal{L}(L_\infty, L_1)$ through a Hilbert space (see [73], 1.5.1), (ii) to pick out the operator between the two Hilbert spaces, say H, K, which is Hilbert Schmidt, and (iii) to show that it factors as $H \to \ell_\infty \to \ell_2 \to K$ with the middle part being a diagonal operator (see [73], 1.7.10, and use well-known coincidences of ideals on Hilbert spaces). The latter can be obviously approximated by finite operators. Putting this together, we get a new factorization, consistent with the definition of the ideal, a part of which is approximable by finite operators.

In the following statement E can be any of the spaces $\mathbb{C}^N$ with $N \in \mathbb{N}$, or the sequence spaces $c_0, \ell_p, 1 \leq p < \infty$.

Lemma 7.1.6. *Let f_0 be a rational function satisfying $\overline{f_0(z)} = -f_0(-\overline{z})$ at all $z \in \mathbb{C}$ where f_0 is finite. Assume for a given $T \in \mathcal{L}(E)$ that* $\mathrm{spec}(T)$ *does not intersect the set P of poles of f_0. Then $f_0(-\overline{T})$ is defined and we have*

$$\overline{f_0(T)} = -f_0(-\overline{T}). \tag{7.4}$$

Proof If $\epsilon > 0$ is sufficiently small and $R > 0$ sufficiently large, the contour

$$\Gamma = \{\zeta \mid |\zeta| = R\} \cup \bigcup_{z\in P} \{\zeta \mid |\zeta - z| = \epsilon\} \tag{7.5}$$

surrounds $\mathrm{spec}(T)$, and we have

$$f_0(T) = \frac{1}{2\pi i}\int_{|\zeta|=R} f_0(\zeta)(\zeta I - T)^{-1}\, d\zeta - \frac{1}{2\pi i}\sum_{z\in P}\int_{|\zeta-z|=\epsilon} f_0(\zeta)(\zeta I - T)^{-1}\, d\zeta, \tag{7.6}$$

where all circles are oriented counter-clockwise. We compute

$$\begin{aligned} \overline{f_0(T)} &= -\frac{1}{2\pi i}\int_{|\zeta|=R} \overline{f_0(\zeta)}(\overline{\zeta} I - \overline{T})^{-1}\, d\overline{\zeta} + \frac{1}{2\pi i}\sum_{z\in P}\int_{|\zeta-z|=\epsilon} \overline{f_0(\zeta)}(\overline{\zeta} I - \overline{T})^{-1}\, d\overline{\zeta} \\ &= \frac{1}{2\pi i}\int_{|\zeta|=R} f_0(-\overline{\zeta})(\overline{\zeta} I - \overline{T})^{-1}\, d\overline{\zeta} + \frac{1}{2\pi i}\sum_{z\in P}\int_{|\zeta-z|=\epsilon} f_0(-\overline{\zeta})(\overline{\zeta} I - \overline{T})^{-1}\, d\overline{\zeta}. \end{aligned}$$

The assumption $\overline{f_0(z)} = -f_0(-\overline{z})$ implies that P is symmetric with respect to the imaginary axis. Hence the transformation $\eta = -\overline{\zeta}$ maps Γ to itself while changing the orientations of the circles. Hence we can continue

$$\overline{f_0(T)} = -\frac{1}{2\pi i}\int_{|\zeta|=R} f_0(\eta)(\eta I - \overline{T})^{-1}\, d\eta + \frac{1}{2\pi i}\sum_{z\in P}\int_{|\zeta-z|=\epsilon} f_0(\eta)(\eta I - \overline{T})^{-1}\, d\eta. \tag{7.7}$$

Since $\mathrm{spec}(-\overline{T}) = -\overline{\mathrm{spec}(T)}$ and P is symmetric with respect to the imaginary axis, Γ surrounds $\mathrm{spec}(-\overline{T})$. Hence the last expression equals $-f_0(-\overline{T})$, and the proof is complete. $\square$

Finally we shall often use the following fact without further mentioning. If $T_{ij} \in \mathcal{A}(E)$ for some quasi-Banach operator ideal $\mathcal{A}$, $i, j = 1, 2$, then

$$T = \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} \in \mathcal{A}(E \oplus E).$$

Of course this is easily proved by choosing appropriate projections and embeddings and using the ideal properties (see Proposition 2.4.1).

Furthermore, we shall always identify $a \in E'$ with its standard representation as a sequence $(a_i)_i$.

7.1.2 The basic solution formula

Here we provide the basic solution formula for the $\mathbb{C}$-reduced AKNS system. Note that it is valid for generating operators A on sequence spaces.

Theorem 7.1.7. *Let E be a classical sequence space and $\mathcal{A}$ a p-Banach operator ideal $(0 < p \leq 1)$ with a continuous determinant δ which is compatible with respect to conjugation in the sense that*

$$\delta(I + \overline{T}) = \overline{\delta(I + T)}$$

for all $T \in \mathcal{A}(E)$ with $\overline{T} \in \mathcal{A}(E)$.

Let $A \in \mathcal{L}(E)$ with $0 \notin \mathrm{spec}(A) + \mathrm{spec}(\overline{A})$, and let $\mathrm{spec}(A)$ be contained in the domain where f_0 is holomorphic. Choose $0 \neq a \in E'$, $c \in E$ arbitrarily, and define the operator-functions

$$\begin{aligned} L(x,t) &= \exp\big(\, Ax + f_0(A)t \,\big)\, \Phi^{-1}_{A,\overline{A}}(\overline{a} \otimes c), \\ L_0(x,t) &= \exp\big(\, Ax + f_0(A)t \,\big)\, (a \otimes c). \end{aligned}$$

Then L, $\overline{L}$ belong to $\mathcal{A}$ and L_0, $\overline{L}_0$ are even one-dimensional.

Assume in addition that $\exp(Ax)$ *behaves sufficiently well for* $x \to -\infty$. *Then,* $q = 1 - P/p$, *where*

$$P = \delta\begin{pmatrix} I & -L \\ \overline{L} & I + \overline{L}_0 \end{pmatrix}, \qquad p = \delta\begin{pmatrix} I & -L \\ \overline{L} & I \end{pmatrix},$$

solves the $\mathbb{C}$*-reduced AKNS system (4.5) on every strip* $\mathbb{R} \times (t_1, t_2)$ *on which the denominator* p *does not vanish.*

Proof We want to apply Theorem 2.4.4 b) with the particular choice

(i) $B = \overline{A}$,

(ii) $b = -\overline{a}$, and $d = \overline{c}$.

Recall that, for the $\mathbb{C}$-reduced AKNS system, $\overline{f_0(z)} = -f_0(-\overline{z})$ for all $z \in \mathbb{C}$ where f_0 is finite. Let the operators $\widehat{L}$, $\widehat{M}$ be as defined in Theorem 2.4.4. By Lemma 7.1.6 and (i), $\widehat{M}$ is well-defined and $\widehat{L}$, $\widehat{M}$ are complex conjugate to each other.

Next, for $C = \Phi_{A,B}^{-1}(b \otimes c)$, $D = \Phi_{B,A}^{-1}(a \otimes d)$, the conditions (i), (ii) immediately show $D = -\overline{C}$. In particular, $L = -\widehat{L}C$, $\overline{L} = \widehat{M}D$ belong to $\mathcal{A}$, and one-dimensionality of L_0, $\overline{L}_0$ is obvious. Thus Theorem 2.4.4 b) provides us with the solution formula $q = 1 - P/p$, $r = 1 = \widehat{P}/p$, where

$$P = \delta\begin{pmatrix} I & -L \\ \overline{L} & I + \overline{L}_0 \end{pmatrix}, \quad \widehat{P} = \delta\begin{pmatrix} I - L_0 & -L \\ \overline{L} & I \end{pmatrix}, \quad p = \delta\begin{pmatrix} I & -L \\ \overline{L} & I \end{pmatrix}.$$

It remains to establish the linear relation $r = -\overline{q}$ for the $\mathbb{C}$-reduced AKNS system.

To this end, we use Proposition 2.4.3 to rewrite r as $r = \widetilde{P}/p - 1$, where

$$\widetilde{P} = \delta\begin{pmatrix} I + L_0 & -L \\ \overline{L} & I \end{pmatrix}.$$

By the compatibility condition and the usual property of determinants, we infer

$$\begin{aligned}
\overline{P} &= \overline{\delta\Big(I + \begin{pmatrix} 0 & -L \\ \overline{L} & \overline{L}_0 \end{pmatrix}\Big)} \\
&= \delta\Big(I + \begin{pmatrix} 0 & -\overline{L} \\ L & L_0 \end{pmatrix}\Big) \\
&= \delta\Big(I + \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}\begin{pmatrix} 0 & -\overline{L} \\ L & L_0 \end{pmatrix}\begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}\Big) \\
&= \delta\Big(I + \begin{pmatrix} L_0 & -L \\ \overline{L} & 0 \end{pmatrix}\Big) \\
&= \widetilde{P},
\end{aligned}$$

and obviously $\overline{p} = p$. Consequently, $\overline{q} = 1 - \overline{P}/\overline{p} = 1 - \widetilde{P}/p = -r$.

This completes the proof. □

7.1.3 Ameliorations for the Nonlinear Schrödinger equation

For the Nonlinear Schrödinger equation, we have the ameliorated solution formulas of Section 2.6 at our disposal. The crucial point is that these avoid the condition that $\exp(Ax)$ behaves sufficiently well for $x \to -\infty$. Moreover, since $f_0(z) = -\mathrm{i}z^2$ for the Nonlinear Schrödinger equation, the condition that $\mathrm{spec}(A)$ is contained in the domain where f_0 is holomorphic becomes superfluous.

The ameliorated solution formula for the Nonlinear Schrödinger equation based on sequence spaces reads as follows.

Theorem 7.1.8. *Let E be a classical sequence space and $\mathcal{A}$ a quasi-Banach operator ideal with a continuous determinant δ which is compatible with respect to conjugation in the sense that*

$$\delta(I+\overline{T}) = \overline{\delta(I+T)}$$

for all $T \in \mathcal{A}(E)$ with $\overline{T} \in \mathcal{A}(E)$.

Let $A \in \mathcal{L}(E)$, and let $a \in E'$, $c \in E$ be given. Assume that there exist $C \in \mathcal{A}(E)$, $0 \neq a \in E'$ such that $AC + C\overline{A} = \overline{a} \otimes c$. We define the operator-functions

$$\begin{aligned} L(x,t) &= \exp\left(Ax - \mathrm{i}A^2 t \right) C, \\ L_0(x,t) &= \exp\left(Ax - \mathrm{i}A^2 t \right) (a \otimes c). \end{aligned}$$

Then L, $\overline{L}$ belong to $\mathcal{A}$ and L_0, $\overline{L}_0$ are even one-dimensional.

Moreover, $q = 1 - P/p$, where

$$P = \delta \begin{pmatrix} I & -L \\ \overline{L} & I + \overline{L}_0 \end{pmatrix}, \qquad p = \delta \begin{pmatrix} I & -L \\ \overline{L} & I \end{pmatrix},$$

is a solution of the Nonlinear Schrödinger equation (1.5) wherever the denominator p does not vanish.

The proof can be taken over from the proof of Proposition 7.1.7. One just has to start from Proposition 2.6.1 instead of Theorem 2.4.4. Furthermore there are some simplifications because of the simple form of f_0.

7.2 Countable superposition of solitons for the $\mathbb{C}$-reduction

To place ourselves first into a setting where the operator theory is transparent, we start with countable superpositions of solitons. These can be realized by diagonal operators.

In the sequel, E will always be one of the classical sequence spaces c_0, ℓ_p $(1 \leq p < \infty)$ and $A \in \mathcal{L}(E)$ the diagonal operator generated by a bounded sequence $\alpha = (\alpha_i)_i \in \ell_\infty$,

$$A : E \longrightarrow E \quad \text{with} \quad A(\xi_i)_i = (\alpha_i \xi_i)_i.$$

Observe $\operatorname{spec}(A) = \overline{\{\alpha_i | i \in \mathbb{N}\}}$.

It should be remarked that all results in this section can be transferred to weighted sequence spaces.

In order to apply Theorem 7.1.7, we need to guarantee that $\exp(Ax)$ behave sufficiently well for $x \to -\infty$. A neat way to arrange this is to suppose $\inf_i \operatorname{Re}(\alpha_i) > 0$. Geometrically this means that the solitons are localized to a certain extent. More precisely, one can see that with $\operatorname{Re}(\alpha_i) \to 0$ for $i \to \infty$ the solitons become lower but broader.

Theorem 7.2.1. *Let $\alpha = (\alpha_i)_i$ be a bounded sequence such that $\inf_i \operatorname{Re}(\alpha_i) > 0$ and $\sup_i |f_0(\alpha_i)| < \infty$. Then the operators defined by the infinite matrices*

$$\begin{aligned} L(x,t) &= \left(\frac{\overline{a}_j c_i}{\alpha_i + \overline{\alpha}_j} \exp\left(\alpha_i x + f_0(\alpha_i) t \right) \right)_{i,j=1}^{\infty} \\ \textit{and } L_0(x,t) &= \left(a_j c_i \exp\left(\alpha_i x + f_0(\alpha_i) t \right) \right)_{i,j=1}^{\infty} \end{aligned}$$

belong to the component $\mathcal{N}_{\frac{2}{3}}(E)$ of the $\frac{2}{3}$-Banach operator ideal $\mathcal{N}_{\frac{2}{3}}$ for all sequences $c = (c_i)_i \in E$, $0 \neq a = (a_i)_i \in E'$, and the same holds true for the complex conjugate operators.

Moreover, $q = 1 - P/p$, where

$$P \;=\; \det_\lambda \begin{pmatrix} I & -L \\ L & I + L_0 \end{pmatrix}, \qquad p \;=\; \det_\lambda \begin{pmatrix} I & -L \\ L & I \end{pmatrix},$$

is a solution of the $\mathbb{C}$-reduced AKNS system (4.5) on strips $\mathbb{R} \times (t_1, t_2)$ where the denominator p does not vanish.

In the above statement $\det_\lambda$ denotes the unique, continuous, spectral determinant on the $\frac{2}{3}$-Banach operator ideal $\mathcal{N}_{\frac{2}{3}}$ of $\frac{2}{3}$-nuclear operators. In Theorem 7.3.2 we shall see that the solutions in Theorem 7.2.1 are globally regular.

Proof Recall that A is a diagonal operator generated by the sequence $(\alpha_i)_i$. Thus, also $\exp\big(Ax + f_0(A)t\big)$ is a diagonal operator generated by the sequence $\big(\exp(\alpha_i x + f_0(\alpha_i)t)\big)_i$. Since $\inf_i \mathrm{Re}(\alpha_i) > 0$, $\exp(Ax)$ behaves sufficiently well as $x \to -\infty$. Furthermore the condition that $|f_0(\alpha_i)|$ be uniformly bounded for all i shows that $\mathrm{spec}(A) = \overline{\{\alpha_i \mid i\}}$ is contained in the domain where f_0 is holomorphic.

Next, it is straightforward to check that the operator $C : E \longrightarrow E$ generated by the infinite matrix

$$\left(\frac{\overline{a}_j c_i}{\alpha_i + \overline{\alpha}_j}\right)_{i,j=1}^{\infty}$$

is bounded and satisfies $AC + C\overline{A} = \overline{a} \otimes c$. Since $0 \notin \mathrm{spec}(A) + \mathrm{spec}(\overline{A})$ by assumption, we can apply Proposition 2.3.4, which shows $C = \Phi_{A,\overline{A}}^{-1}(\overline{a} \otimes c) \in \mathcal{N}_{\frac{2}{3}}(E)$. Moreover, $a \otimes c$ is obviously generated by the infinite matrix $\big(a_j c_i\big)_{i,j=1}^{\infty}$.

Thus the assertion immediately follows by applying Theorem 7.1.7 with respect to the $\frac{2}{3}$-Banach operator ideal $\mathcal{N}_{\frac{2}{3}}$ of $\frac{2}{3}$-nuclear operators (see Lemma 7.1.2). $\square$

Remark 7.2.2. *Note that in Theorem 7.2.1 any continuous determinant δ on a p-Banach operator ideal $\mathcal{A}$, $0 < p \leq 1$, can be used which has the property that $\overline{\delta(I+T)} = \delta(I+\overline{T})$ for all $T \in \mathcal{A}(E)$ satisfying $\overline{T} \in \mathcal{A}(E)$.*

7.3 Universal realization and regularity

In this section we will see that the previous solutions can be reinterpreted uniformly in terms of the nuclear determinant $\det_{\mathcal{N}}$ on ℓ_1. The practical use of this is that the determinant can be evaluated by the concrete formula (7.1). This can be exploited to prove that the solutions considered in the previous section are globally regular. Note that regularity is not obvious at all in our general setting. Furthermore we will obtain a better understanding of the parameters a, c.

Theorem 7.3.1. *Let $\alpha = (\alpha_i)_i \in \ell_\infty$ be a bounded sequence such that* $\inf_i \mathrm{Re}(\alpha_i) > 0$ *and* $\sup_i |f_0(\alpha_i)| < \infty$*. Then the operators defined by the infinite matrices*

$$\begin{aligned} L(x,t) \;&=\; \left(\frac{d_i}{\alpha_i + \overline{\alpha}_j} \exp\,\big(\,\alpha_i x + f_0(\alpha_i)t\,\big)\right)_{i,j=1}^{\infty} \\ \text{and } L_0(x,t) \;&=\; \Big(\,d_i\,\exp\,\big(\,\alpha_i x + f_0(\alpha_i)t\,\big)\,\Big)_{i,j=1}^{\infty} \end{aligned}$$

belong to the nuclear component $\mathcal{N}(\ell_1)$ for all sequences $0 \neq d = (d_i)_i \in \ell_1$, and the same is true for the complex conjugate operators.

Moreover $q = 1 - P/p$, *where*

$$P = \det_{\mathcal{N}} \begin{pmatrix} I & -L \\ \overline{L} & I + \overline{L}_0 \end{pmatrix}, \quad p = \det_{\mathcal{N}} \begin{pmatrix} I & -L \\ \overline{L} & I \end{pmatrix},$$

is a solution of the $\mathbb{C}$*-reduced AKNS system (4.5) on strips* $\mathbb{R} \times (t_1, t_2)$ *where the denominator* p *does not vanish.*

In addition, each solution of Theorem 7.2.1 can be expressed explicitly in this way.

In the above statement $\det_{\mathcal{N}}$ denotes the unique continuous determinant on the Banach operator ideal $\mathcal{N}$ of nuclear operators restricted to the class of Banach spaces with a.p..

Proof (of Theorem 7.3.1) As for the proof that q is a solution of the $\mathbb{C}$-reduced AKNS system, we refer to Theorem 7.2.1 with the following modifications:

- The determinant $\det_{\mathcal{N}}$ on the nuclear component $\mathcal{N}(\ell_1)$ is used, see Remark 7.2.2 and Lemma 7.1.2.
- The particular choice $d \in \ell_1$, $e_0 := (1, 1, \dots) \in \ell_\infty$ for the sequences is made.

Therefore, the only thing left to show is that each solution of Theorem 7.2.1 can be expressed explicitly in this way. For concreteness we fix such a solution.

Our first objective is to provide an appropriate factorization of the operators L, L_0 in Theorem 7.2.1.

To this end, consider the diagonal operators $A \in \mathcal{L}(\ell_\infty)$, $\widehat{A} \in \mathcal{L}(\ell_1)$ generated by the sequences α, $\overline{\alpha}$, respectively. Note that $\widehat{A}$ is not the complex conjugate operator of A because the underlying spaces differ. Since $0 \notin \operatorname{spec}(A) + \operatorname{spec}(\widehat{A})$, Proposition 2.3.4 implies that the operator equation $AX + X\widehat{A} = e_0 \otimes e_0$ on $\mathcal{N}_{\frac{2}{3}}(\ell_1, \ell_\infty)$ has the unique solution $C_0 := \Phi_{A,\widehat{A}}^{-1}(e_0 \otimes e_0)$. We verify $C_0 = \big(1/(\alpha_i + \overline{\alpha}_j)\big)_{i,j=1}^{\infty}$.

In addition we define $\widehat{L}(x,t) \in \mathcal{L}(E)$ as the diagonal operator generated by the sequence $\big(\exp(\alpha_i x + f_0(\alpha_i)t)\big)_i$. Moreover, let $D_a \in \mathcal{L}(E, \ell_1)$ and $D_c \in \mathcal{L}(\ell_\infty, E)$ be the diagonal operators generated by the sequences a, c, respectively.

Then the operators L, $L_0 \in \mathcal{N}_{\frac{2}{3}}(E)$ in Theorem 7.2.1 factorize as $L = \widehat{L} D_c C_0 \overline{D}_a$, $L_0 = \widehat{L} D_c (e_0 \otimes e_0) D_a$.

Consequently the following operators are related,

$$\begin{pmatrix} 0 & -L \\ \overline{L} & \overline{L}_0 \end{pmatrix}, \quad \begin{pmatrix} 0 & -M \\ \overline{M} & \overline{M}_0 \end{pmatrix}$$

with $M_0 := D_a \widehat{L} D_c (e_0 \otimes e_0)$, $M := D_a \widehat{L} D_c C_0 \in \mathcal{N}_{\frac{2}{3}}(\ell_1)$. Namely, one immediately sees

$$\begin{aligned}
\begin{pmatrix} 0 & -L \\ \overline{L} & \overline{L}_0 \end{pmatrix} &= \begin{pmatrix} 0 & -\widehat{L} D_c C_0 \overline{D}_a \\ \overline{\widehat{L}} \overline{D}_c \overline{C}_0 D_a & \overline{\widehat{L}} \overline{D}_c (e_0 \otimes e_0) \overline{D}_a \end{pmatrix} \\
&= \begin{pmatrix} \widehat{L} D_c & 0 \\ 0 & \overline{\widehat{L}} \overline{D}_c \end{pmatrix} \begin{pmatrix} 0 & -C_0 \\ \overline{C}_0 & e_0 \otimes e_0 \end{pmatrix} \begin{pmatrix} D_a & 0 \\ 0 & \overline{D}_a \end{pmatrix}
\end{aligned}$$

and

$$\begin{aligned}
\begin{pmatrix} 0 & -M \\ \overline{M} & \overline{M}_0 \end{pmatrix} &= \begin{pmatrix} 0 & -D_a \widehat{L} D_c C_0 \\ \overline{D}_a \overline{\widehat{L}} \overline{D}_c \overline{C}_0 & \overline{D}_a \overline{\widehat{L}} \overline{D}_c (e_0 \otimes e_0) \end{pmatrix} \\
&= \begin{pmatrix} D_a & 0 \\ 0 & \overline{D}_a \end{pmatrix} \begin{pmatrix} \widehat{L} D_c & 0 \\ 0 & \overline{\widehat{L}} \overline{D}_c \end{pmatrix} \begin{pmatrix} 0 & -C_0 \\ \overline{C}_0 & e_0 \otimes e_0 \end{pmatrix}.
\end{aligned}$$

As a result,

$$\det_\lambda\Big(I + \begin{pmatrix} 0 & -L \\ \overline{L} & \overline{L_0} \end{pmatrix}\Big) = \det_\lambda\Big(I + \begin{pmatrix} 0 & -M \\ \overline{M} & \overline{M_0} \end{pmatrix}\Big),$$

$\det_\lambda$ denoting the spectral determinant on $\mathcal{N}_{\frac{2}{3}}$. In the latter expression $\det_\lambda$ can be replaced by the nuclear determinant $\det_{\mathcal{N}}$ on the component $\mathcal{N}(\ell_1)$, see Corollary 7.1.5.

Finally we observe that M, M_0 are of the form as required in the statement, namely

$$\begin{aligned} M &= \left(\frac{d_i}{\alpha_i + \overline{\alpha}_j}\exp\Big(\alpha_i x + f_0(\alpha_i)t\Big)\right)_{i,j=1}^{\infty}, \\ M_0 &= \Big(d_i \exp\Big(\alpha_i x + f_0(\alpha_i)t\Big)\Big)_{i,j=1}^{\infty} \end{aligned}$$

for $d_i := a_i c_i$. In particular, $d = (d_i)_i \in \ell_1$.

Carrying out the same manipulations for the denominator p of the solution, we complete the proof. □

Finally, we prove global regularity of the solutions constructed so far.

Theorem 7.3.2. *The solutions in Theorem 7.3.1 (and hence in Theorem 7.2.1) are defined and regular on all of* $\mathbb{R}^2$.

Proof We have only to show that the denominator p does not vanish. By (7.1), p is the limit of p_N for $N \to \infty$, where p_N is the denominator appearing in the formula of the N-soliton corresponding to the first N members of the countable superposition. But in the proof of Proposition 4.3.7 we have seen that $p_N \geq 1$. Hence the same holds true for p, and the proof is complete. □

7.4 Sharper results for the Nonlinear Schrödinger equation

Next we turn to degenerate cases. As explained in the introduction, this should be done in the context of concrete equations. Here we will study the Nonlinear Schrödinger equation in detail and show that under suitable conditions we can handle the case $\inf_i \mathrm{Re}(\alpha_i) = 0$. Geometrically this means that we can superpose solitons whose widths diverge.

In the case $\inf_i \mathrm{Re}(\alpha_i) > 0$ we obtained an operator $C = \Phi^{-1}_{A,\overline{A}}(\overline{a} \otimes c)$ with $AC + C\overline{A} = \overline{a} \otimes c$ by the Eschmeier-Dash/Schechter theorem. Furthermore for any choice $(a,\, c)$ it was clear that C belonged to any p-Banach operator ideal $\mathcal{A}$. In the present situation the underlying mapping $\Phi^{-1}_{A,\overline{A}}$ is no longer defined. Nevertheless the matrix expression $\big((\overline{a}_j c_i)/(\alpha_i + \overline{\alpha}_j)\big)_{i,j=1}^{\infty}$ still yields a formal candidate for C, which may a priori even be unbounded. Our task is now to ensure by a clever choice of $(a,\, c)$ that the candidate is contained in an appropriate quasi-Banach operator ideal.

The following lemma is the crucial step.

Lemma 7.4.1. *Let $\alpha = (\alpha_i)_i \in \ell_\infty$ be a bounded sequence with $\mathrm{Re}(\alpha_i) > 0$ $\forall i$. Then the operator $\widetilde{C}_0 : \ell_1 \longrightarrow \ell_\infty$ defined by the infinite matrix*

$$\widetilde{C}_0 = \left(\frac{\sqrt{\mathrm{Re}(\alpha_i)}\sqrt{\mathrm{Re}(\alpha_j)}}{\alpha_i + \overline{\alpha}_j}\right)_{i,j=1}^{\infty}$$

belongs to the quasi-Banach operator ideal $\mathcal{L}_\infty \circ \mathcal{H} \circ \mathcal{L}_1$.

The same statement holds true for the complex conjugate operator $\overline{\widetilde{C}_0}$ of $\widetilde{C}_0$.

Proof We show that $\widetilde{C}_0$ factors through the Hilbert space $L_2[0,\infty)$. To this end, consider the operator $S : \ell_1 \mapsto L_2[0,\infty)$ which is defined on the standard basis $\{e_i \mid i\}$ of ℓ_1 by $Se_i := \overline{f}_i$ with $f_i(s) = \sqrt{\mathrm{Re}(\alpha_i)} \exp\big(-\alpha_i s\big)$. Because of

$$\begin{aligned}
\langle S'Se_j, e_i\rangle &= \langle Se_j, Se_i\rangle = \int_0^\infty \overline{f_j(s)} f_i(s)\, ds \\
&= \sqrt{\mathrm{Re}(\alpha_i)}\sqrt{\mathrm{Re}(\alpha_j)} \int_0^\infty \exp\big(-(\alpha_i + \overline{\alpha}_j)s\big)\; ds \\
&= \frac{\sqrt{\mathrm{Re}(\alpha_i)}\sqrt{\mathrm{Re}(\alpha_j)}}{\alpha_i + \overline{\alpha}_j} = \langle \widetilde{C}_0 e_j, e_i\rangle
\end{aligned}$$

actually $\widetilde{C}_0 = S'S$. Similarly as above we check $\| f_j \|_{L_2[0,\infty)} = \sqrt{\frac{1}{2}}$, which in turn proves that S is bounded.

As for the complex conjugate operator $\overline{\widetilde{C}_0}$ of $\widetilde{C}_0$, a factorization $\overline{\widetilde{C}_0} = R'R$ through $L_2[0,\infty)$ is obtained via $R : \ell_1 \to L_2[0,\infty)$ given by $Re_i := f_i$ with f_i as above. □

Theorem 7.4.2. *Let $\alpha = (\alpha_j)_j \in \ell_\infty$ be a bounded sequence with $\mathrm{Re}(\alpha_j) > 0\ \forall j$. Then the operators defined by the infinite matrices*

$$\begin{aligned}
L(x,t) &= \left(\frac{\overline{a}_{j'}c_j}{\alpha_j + \overline{\alpha}_{j'}} \exp\big(\,\alpha_j x - \mathrm{i}\alpha_j^2 t\,\big)\right)_{j,j'=1}^\infty \\
\textit{and } L_0(x,t) &= \Big(a_{j'}c_j \exp\big(\,\alpha_j x - \mathrm{i}\alpha_j^2 t\,\big)\Big)_{j,j'=1}^\infty
\end{aligned}$$

belong to the component $\mathcal{L}_\infty \circ \mathcal{H} \circ \mathcal{L}_1(E)$ of the quasi-Banach operator ideal $\mathcal{L}_\infty \circ \mathcal{H} \circ \mathcal{L}_1$ for all sequences $c = (c_j)_j$, $0 \neq a = (a_j)_j$ with $(c_j/\sqrt{\mathrm{Re}(\alpha_j)})_j \in E$, $a = (a_j/\sqrt{\mathrm{Re}(\alpha_j)})_j \in E'$, and the same is true for the complex conjugate operators.

Moreover, $q = 1 - P/p$, where

$$P = \det{}_\lambda \begin{pmatrix} I & -L \\ \overline{L} & I + \overline{L}_0 \end{pmatrix}, \quad p = \det{}_\lambda \begin{pmatrix} I & -L \\ \overline{L} & I \end{pmatrix},$$

is a solution of the Nonlinear Schrödinger equation (1.5) wherever the denominator p does not vanish.

In the above statement $\det_\lambda$ denotes the continuous and spectral determinant on the quasi-Banach operator ideal $\mathcal{L}_\infty \circ \mathcal{H} \circ \mathcal{L}_1$.

Proof Recall that A is a diagonal operator generated by the sequence $(\alpha_j)_j$. Thus, also $\exp\big(Ax - \mathrm{i}A^2 t\big)$ is a diagonal operator, and it is generated by the sequence $\big(\exp(\alpha_j x - \mathrm{i}\alpha_j^2 t)\big)_j$.

It is also clear that $a \otimes c$ is generated by the infinite matrix $\big(a_{j'}c_j\big)_{j,j'=1}^\infty$.

Next we introduce the diagonal operators $D_{a/\sqrt{\mathrm{Re}(\alpha)}} : E \to \ell_1$, $D_{c/\sqrt{\mathrm{Re}(\alpha)}} : \ell_\infty \to E$ generated by the sequences $(a_j/\sqrt{\mathrm{Re}(\alpha_j)})_j \in E'$, $(c_j/\sqrt{\mathrm{Re}(\alpha_j)})_j \in E$, respectively. By Hölders inequality, both are bounded. Thus Lemma 7.4.1 implies that the operator

$$C := D_{c/\sqrt{\mathrm{Re}(\alpha)}}\, \widetilde{C}_0\, \overline{D}_{a/\sqrt{\mathrm{Re}(\alpha)}} \;:\; E \longrightarrow E$$

belongs to the quasi-Banach operator ideal $\mathcal{L}_\infty \circ \mathcal{H} \circ \mathcal{L}_1$. Obviously, C is represented by the infinite matrix $\big((\overline{a}_{j'}c_j)/(\alpha_j + \overline{\alpha}_{j'})\big)_{j,j'=1}^\infty$, and the equation $AC + C\overline{A} = \overline{a} \otimes c$ can be immediately verified.

Similarly, $\overline{C} \in \mathcal{L}_\infty \circ \mathcal{H} \circ \mathcal{L}_1(E)$ solving $\overline{AC} + \overline{C}A = a \otimes \overline{c}$.

Now the assertion follows by applying Theorem 7.1.8 with respect to the quasi-Banach operator ideal $\mathcal{L}_\infty \circ \mathcal{H} \circ \mathcal{L}_1$ of operators factorizing through an L_1-space, a Hilbert space, and an L_∞-space. $\square$

Next we show that, even for the solutions of the Nonlinear Schrödinger equation which only satisfy the weaker condition $\mathrm{Re}(\alpha_j) > 0$ for all j, the counterpart of the universal realization in Theorem 7.3.1 holds.

Theorem 7.4.3. *Let $\alpha = (\alpha_j)_j \in \ell_\infty$ be a bounded sequence with $\mathrm{Re}(\alpha_j) > 0$ for all j. Then the operators defined by the infinite matrices*

$$\begin{aligned} L(x,t) &= \left(\frac{d_j}{\alpha_j + \overline{\alpha}_{j'}} \exp\left(\alpha_j x - \mathrm{i}\alpha_j^2 t \right) \right)_{j,j'=1}^{\infty} \\ \text{and } L_0(x,t) &= \left(d_j \exp\left(\alpha_j x - \mathrm{i}\alpha_j^2 t \right) \right)_{j,j'=1}^{\infty} \end{aligned}$$

belong to the nuclear component $\mathcal{N}(\ell_1)$ for all sequences $0 \neq d = (d_j)_j$ with $(d_j/\mathrm{Re}(\alpha_j))_j \in \ell_1$, and the same holds true for the complex conjugate operators.

Moreover $q = 1 - P/p$, where

$$P = \det{}_{\mathcal{N}} \begin{pmatrix} I & -L \\ \overline{L} & I + \overline{L}_0 \end{pmatrix}, \quad p = \det{}_{\mathcal{N}} \begin{pmatrix} I & -L \\ \overline{L} & I \end{pmatrix},$$

is a solution of the Nonlinear Schrödinger equation (1.5) wherever the denominator p does not vanish.

In addition, each solution of Theorem 7.4.2 can be expressed explicitly in this way.

In the above statement $\det_{\mathcal{N}}$ denotes the unique continuous determinant on the Banach operator ideal $\mathcal{N}$ of nuclear operators restricted to the class of Banach spaces with approximation property.

Proof First we show that q solves the Nonlinear Schrödinger equation. Observe that the operator $C : \ell_1 \to \ell_1$ generated by the infinite matrix

$$C = \left(\frac{d_j}{\alpha_j + \overline{\alpha}_{j'}} \right)_{j,j'=1}^{\infty}$$

is nuclear. Indeed $|\alpha_j + \overline{\alpha}_{j'}| \geq \mathrm{Re}(\alpha_j)$ implies

$$\| C \mid \mathcal{N} \| = \sum_j \sup_{j'} \left| \frac{d_j}{\alpha_j + \overline{\alpha}_{j'}} \right| \leq \sum_j \left| \frac{d_j}{\mathrm{Re}(\alpha_j)} \right| = \| \left(d_j/\mathrm{Re}(\alpha_j) \right)_j \|_1 < \infty.$$

Obviously, $AC + C\overline{A} = e_0 \otimes d$ for $e_0 = (1,1,\ldots) \in \ell_\infty$, $d \in \ell_1$ (here we have used that $(\alpha_j)_j \in \ell_\infty$, $(d_j/\mathrm{Re}(\alpha_j))_j \in \ell_1$ yields $(d_j)_j \in \ell_1$).

Thus we can apply Theorem 7.1.8 with respect to the nuclear determinant $\det_{\mathcal{N}}$ on ℓ_1 (see Lemma 7.1.2).

It remains to show that each solution of Theorem 7.4.2 can be expressed explicitly in this form. For concreteness fix such a solution. We use the following factorization for the operators L, L_0 in Theorem 7.4.2:

$$\begin{aligned} L &= \widehat{L} D_{c/\sqrt{\mathrm{Re}(\alpha)}} \widetilde{C}_0 \overline{D}_{a/\sqrt{\mathrm{Re}(\alpha)}}, \\ L_0 &= \widehat{L} D_{c/\sqrt{\mathrm{Re}(\alpha)}} \big(\sqrt{\mathrm{Re}(\alpha)} \otimes \sqrt{\mathrm{Re}(\alpha)} \big) D_{a/\sqrt{\mathrm{Re}(\alpha)}}, \end{aligned}$$

where $D_{a/\sqrt{\mathrm{Re}(\alpha)}} \in \mathcal{L}(E, \ell_1)$, $D_{c/\sqrt{\mathrm{Re}(\alpha)}} \in \mathcal{L}(\ell_\infty, E)$, and $\widetilde{C}_0 \in \mathcal{L}_\infty \circ \mathcal{H} \circ \mathcal{L}_1(\ell_1, \ell_\infty)$ have been introduced in the proof of Theorem 7.4.2, and the diagonal operator $\widehat{L}(x,t) \in \mathcal{L}(E)$ is generated by the sequence $\big(\exp(\alpha_j x - \mathrm{i}\alpha_j^2 t)\big)_j$. By the same calculation as in the proof of Theorem 7.3.1 we check that the following operators are related,

$$\begin{pmatrix} 0 & -L \\ L & L_0 \end{pmatrix}, \quad \begin{pmatrix} 0 & -M \\ M & M_0 \end{pmatrix}$$

with $M_0 := D_{a/\sqrt{\mathrm{Re}(\alpha)}} \widehat{L} D_{c/\sqrt{\mathrm{Re}(\alpha)}} \big(\sqrt{\mathrm{Re}(\alpha)} \otimes \sqrt{\mathrm{Re}(\alpha)}\big)$, $M := D_{a/\sqrt{\mathrm{Re}(\alpha)}} \widehat{L} D_{c/\sqrt{\mathrm{Re}(\alpha)}} \widetilde{C}_0 \in \mathcal{L}_\infty \circ \mathcal{H} \circ \mathcal{L}_1(\ell_1)$.

Hence, by the principle of related operators, we obtain

$$\det\nolimits_\lambda \left(I + \begin{pmatrix} 0 & -L \\ L & L_0 \end{pmatrix}\right) = \det\nolimits_\lambda \left(I + \begin{pmatrix} 0 & -M \\ M & M_0 \end{pmatrix}\right),$$

for the spectral determinant $\det_\lambda$ on $\mathcal{L}_\infty \circ \mathcal{H} \circ \mathcal{L}_1$. In the latter expression $\det_\lambda$ can be replaced by the nuclear determinant $\det_{\mathcal{N}}$ on the component $\mathcal{N}(\ell_1)$, see Corollary 7.1.5. Moreover, we observe the following matrix representations for M, M_0,

$$\begin{aligned} M &= \left(\sqrt{\frac{\mathrm{Re}(\alpha_{j'})}{\mathrm{Re}(\alpha_j)}}\, \frac{d_j}{\alpha_j + \overline{\alpha}_{j'}} \exp\left(\alpha_j x - \mathrm{i}\alpha_j^2 t\right)\right)_{j,j'}^\infty \\ M_0 &= \left(\sqrt{\frac{\mathrm{Re}(\alpha_{j'})}{\mathrm{Re}(\alpha_j)}}\, d_j \exp\left(\alpha_j x - \mathrm{i}\alpha_j^2 t\right)\right)_{j,j'=1}^\infty \end{aligned}$$

for $d_j := a_j c_j$. In particular, $\big(d_j / \mathrm{Re}(\alpha_j)\big)_j \in \ell_1$.

Let us assume for a moment that the sequence $\big(\mathrm{Re}(\alpha_i)\big)_j$ is bounded away from zero. Then we can conclude the proof by

$$\begin{aligned} &\det\nolimits_{\mathcal{N}} \left(I + \begin{pmatrix} 0 & -M \\ M & M_0 \end{pmatrix}\right) = \\ &= \det\nolimits_{\mathcal{N}} \left(I + \begin{pmatrix} D_{\sqrt{\mathrm{Re}(\alpha)}} & 0 \\ 0 & D_{\sqrt{\mathrm{Re}(\alpha)}} \end{pmatrix} \begin{pmatrix} 0 & -M \\ M & M_0 \end{pmatrix} \begin{pmatrix} D_{1/\sqrt{\mathrm{Re}(\alpha)}} & 0 \\ 0 & D_{1/\sqrt{\mathrm{Re}(\alpha)}} \end{pmatrix}\right) \\ &= \det\nolimits_{\mathcal{N}} \left(I + \begin{pmatrix} 0 & -N \\ N & N_0 \end{pmatrix}\right), \end{aligned}$$

where N, N_0 are just the operators defined in the assertion.

For arbitrary sequences $(\alpha_j)_j \in \ell_\infty$ we first note $N, N_0 \in \mathcal{N}(\ell_1)$, which has been shown in the first part of the proof. Thus we can use formula (7.1) for the calculation of the nuclear determinant $\det_{\mathcal{N}}(I+T)$, and the assertion follows by the same argument as before, but now applied to the finite-dimensional principal minors.

Carrying out the same manipulations for the denominator p of the solution, we complete the proof. $\square$

Now we can achieve regularity also for the solutions of the Nonlinear Schrödinger equation constructed in this section. The proof is literally the same as for Theorem 7.3.2.

Theorem 7.4.4. *The solutions in Theorem 7.4.3 (and hence in Theorem 7.4.2) are defined and regular all of* $\mathbb{R}^2$.

7.5 Countable superposition of negatons

Now we come to our main topic, the superposition of countably many negatons. Here we will use so-called generalized diagonal operators A, i.e., an operator with Jordan blocks on the diagonal. To this end, we shall first introduce some additional terminology.

7.5.1 Vector-valued sequence spaces

Let $(n_j)_j \in \ell_\infty$ with $n_j \in \mathbb{N}$ be a bounded sequence of natural numbers, and E one of the classical sequence spaces ℓ_p $(1 \le p < \infty)$ or c_0.

For $1 \le q \le \infty$, we define the vector-valued sequence space

$$E\big(\ell_q(n_j)\big) = \Big\{\xi = (\xi_j)_j \;\Big|\; \xi_j \in \ell_q(n_j) \ \forall j \text{ such that } \big(\| \xi_j \|_q\big)_j \in E\Big\}.$$

Here $\ell_q(n_j)$ denotes $\mathbb{C}^{n_j}$ equipped with the q-norm. $E\big(\ell_q(n_j)\big)$ becomes a Banach space with respect to the norm

$$\| \xi \|_{E(\ell_q(n_j))} = \| \big(\| \xi_j \|_q\big)_j \|_E \, .$$

We write elements ξ in vector-valued sequence spaces $E(\ell_q(n_j))$ as $\xi = (\xi_j)_j$ with vector entries $\xi_j \in \ell_q(n_j)$.

The following duality statement can be shown in the same manner as for ordinary sequence spaces.

Proposition 7.5.1. *For $1 \le q \le \infty$, the pairing*

$$\langle x, a\rangle = \sum\nolimits_j \langle x_j, a_j\rangle, \quad a = (a_j)_j \in E'\big(\ell_{q'}(n_j)\big),\ x = (x_j)_j \in E\big(\ell_q(n_j)\big),$$

yields a metrical isomorphism between $E\big(\ell_q(n_j)\big)'$ and $E'\big(\ell_{q'}(n_j)\big)$ $(\frac{1}{q} + \frac{1}{q'} = 1)$.

We say that an operator $T \in \mathcal{L}\big(E(\ell_q(n_j)), F(\ell_r(n_j))\big)$ is generated by a generalized infinite matrix $(T_{ij})_{i,j=1}^\infty$ if it is given by the rule

$$T\xi = \Big(\sum_k T_{jk}\xi_k \Big)_j, \qquad \text{for } \xi = (\xi_j)_j \in E(\ell_q(n_j)),$$

with matrices $T_{ij} \in \mathcal{M}_{n_i,n_j}(\mathbb{C})$. If not stated otherwise, T_{ij} is always viewed as an operator from $\ell_q(n_j)$ to $\ell_r(n_i)$.

If, in particular, only the matrices T_{jj} on the diagonal are non-zero, then we call T a generalized diagonal operator generated by the sequence $(T_{jj})_j$.

It is not difficult to find conditions such that generalized infinite matrices define bounded operators. Exemplarily, we state the following lemmata.

Lemma 7.5.2. *Any generalized infinite matrix $(T_{ij})_{i,j=1}^\infty$ with $\sup_{i,j} \| T_{ij} \| < \infty$ defines a bounded operator T from $\ell_1(\ell_q(n_j))$ to $\ell_\infty(\ell_r(n_j))$ (where $1 \le q, r \le \infty$).*

Note that $\| T_{ij} \|$ abbreviates $\| T_{ij} \|_{\mathcal{L}(\ell_q(n_j),\ell_r(n_j))}$, as said before.

Proof We observe the elementary estimate

$$\begin{aligned} \| \sum_{j=1}^\infty T_{ij}\xi_j \|_r &\le \sum_{j=1}^\infty \| T_{ij} \| \, \| \xi_j \|_q \le \sup_j \| T_{ij} \| \sum_{j=1}^\infty \| \xi_j \|_q \\ &= \Big(\sup_j \| T_{ij} \| \Big) \, \| \xi \|_{\ell_1(\ell_q(n_j))} \, . \end{aligned}$$

This yields

$$\| T\xi \|_{\ell_\infty(\ell_r(n_j))} = \sup_i \| \sum_{j=1}^{\infty} T_{ij}\xi_j \|_r \leq \left(\sup_{ij} \| T_{ij} \| \right) \| \xi \|_{\ell_1(\ell_q(n_j))}$$

for $\xi = (\xi_j)_j \in \ell_1(\ell_q(n_j))$. Thus $T \in \mathcal{L}\big(\ell_1(\ell_q(n_j)), \ell_\infty(\ell_r(n_j))\big)$. □

Lemma 7.5.3. *Any sequence* $(T_j)_j$, $T_j \in \mathcal{M}_{n_j,n_j}(\mathbb{C})$, *with* $\sup_j \| T_j \| < \infty$ *defines an operator* $T \in \mathcal{L}\big(E(\ell_q(n_j))\big)$ *(where* $1 \leq q \leq \infty$*).*

Proof Set $t = \sup_j \| T_j \|$. Then from

$$\| T\xi \|_{E(\ell_q(n_j))} = \| \ \big(\| T_j\xi_j \|_q \big)_j \ \|_E \ \ \leq t \ \| \ \big(\| \xi_j \|_q \big)_j \ \|_E = t \ \| \xi \|_{E(\ell_q(n_j))}$$

the assertion follows immediately. □

For operators T which are generated by generalized infinite matrices, again the notion of the complex conjugate operator $\overline{T}$ of T can be canonically defined as the operator generated by $(\overline{T}_{ij})_{i,j=1}^{\infty}$.

7.5.2 The main theorem for the $\mathbb{C}$-reduction

Throughout this section we will assume that $(n_j)_j$ is a bounded sequence of natural numbers and E one of the classical sequence spaces ℓ_p $(1 \leq p < \infty)$ or c_0.

Let $\alpha = (\alpha_j)_j \in \ell_\infty$ be a bounded sequence.

To these data we associate the sequence $(A_j)_j$, where $A_j \in \mathcal{M}_{n_j,n_j}(\mathbb{C})$ is the Jordan block of dimension n_j corresponding to the eigenvalue α_j. Then we can define the generalized diagonal operator $A \in \mathcal{L}(E\big(\ell_q(n_j)\big))$, $1 \leq q \leq \infty$, generated by the sequence $(A_j)_j$, i.e.,

$$A(\xi_j)_j = (A_j\xi_j)_j \quad \text{for } (\xi_j)_j \in E\big(\ell_q(n_j)\big),$$

see Lemma 7.5.3.

The following lemma determines $\operatorname{spec}(A)$. Here we will use that the size of the Jordan blocks is bounded.

Lemma 7.5.4. *For the operator A defined above we have* $\operatorname{spec}(A) = \overline{\{\alpha_j \mid j \in \mathbb{N}\}}$.

Proof Obviously $\overline{\{\alpha_j \mid j \in \mathbb{N}\}} \subseteq \operatorname{spec}(A)$. Consider now $\lambda \notin \overline{\{\alpha_j \mid j \in \mathbb{N}\}}$. Then there exists $\epsilon > 0$ such that $|\alpha_j - \lambda| > \epsilon$ for all j. Without loss of generality, $\epsilon < 1$.

Define the generalized diagonal operator T generated by the sequence $(T_j)_j$ with $T_j = (A_j - \lambda I_{n_j})^{-1}$. Then $T_j = \sum_{k=1}^{n_j} (\alpha_j - \lambda)^{-k} N_j^{k-1}$ with N_j the nilpotent matrix of dimension n_j with the entries 1 on the off-diagonal and zero else.

Consequently,

$$\| T_j \| \leq \sum_{k=1}^{n_j} \frac{1}{|\alpha_j - \lambda|^k} \ \leq \ n_j \epsilon^{-n_j} \ \leq \ n_0 \epsilon^{-n_0}$$

for $n_0 := \sup_j n_j$. By Lemma 7.5.3, this shows $T \in \mathcal{L}\big(E(\ell_q(n_j))\big)$, and it is clear that $T = (A - \lambda I)^{-1}$. Thus $\lambda \notin \operatorname{spec}(A)$. □

Review of negatons

For convenience we recall some material on negatons.

To simplify the formulas, we will use the notation $\Gamma_l(\xi), \Gamma_r(\xi) \in \mathcal{M}_{k,k}(\mathbb{C})$ for the upper left and upper right band matrices given in terms of the vector $\xi = (\xi^{(\kappa)})_{\kappa=1}^k \in \mathbb{C}^k$ by the following assignment

$$\Gamma_l(\xi) = \begin{pmatrix} \xi^{(1)} & & \xi^{(k)} \\ & \iddots & \\ \xi^{(k)} & & 0 \end{pmatrix}, \qquad \Gamma_r(\xi) = \begin{pmatrix} \xi^{(1)} & & \xi^{(k)} \\ & \ddots & \\ 0 & & \xi^{(1)} \end{pmatrix}.$$

We write $e_k^{(\kappa)}$ for the κ-th standard basis vector of $\mathbb{C}^k$.

For the construction of finite superpositions of negatons (see Proposition 4.3.1) we used the following two types of matrices:

1. the exponential function $\widehat{L}_i(x,t) := \exp\big(A_i x + f_0(A_i)t\big)$,
2. the solution $\Phi^{-1}_{A_i,\overline{A}_j}(\overline{a}_j \otimes c_i) \in \mathcal{M}_{n_i,n_j}(\mathbb{C})$ of the matrix equation $A_i X + X\overline{A}_j = \overline{a_j} \otimes c_i$ for $c_i \in \mathbb{C}^{n_i}$, $\overline{a}_j \in \mathbb{C}^{n_j}$,

where $A_i \in \mathcal{M}_{n_i,n_i}(\mathbb{C})$ are the Jordan block of dimension n_i corresponding to the eigenvalue α_i. Recall that these matrices have the concrete form given by

$$\widehat{L}_i(x,t) := \Gamma_r\big(\ell_i(x,t)\big) \tag{7.8}$$
$$\text{for the vector } \ell_i(x,t) = \Big(\frac{1}{(\nu-1)!}\frac{\partial^{\nu-1}}{\partial \alpha_i^{\nu-1}} \exp\big(\alpha_i x + f_0(\alpha_i)t\big) \Big)_{\nu=1}^{n_i}$$

and $\Phi^{-1}_{A_i,\overline{A}_j}(\overline{a}_j \otimes c_i) = \Gamma_l(c_i) C_{0,ij} \Gamma_r(\overline{a}_j)$ with

$$C_{0,ij} := \left((-1)^{\nu+\mu} \binom{\nu+\mu-2}{\mu-1} \left(\frac{1}{\alpha_i + \overline{\alpha}_j}\right)^{\nu+\mu-1} \right)_{\substack{\nu=1,\dots,n_i \\ \mu=1,\dots,n_j}}. \tag{7.9}$$

by Proposition 4.3.3.

Observe in particular $C_{0,ij} = \Phi^{-1}_{A_i,\overline{A}_j}(e^{(1)}_{n_j} \otimes e^{(1)}_{n_i})$.

Superpositions of negatons

Now we are in position to extend Theorem 7.2.1 from solitons to negatons. Recall that E always denotes a classical sequence space.

Theorem 7.5.5. *Let n_j, $j \in \mathbb{N}$, be natural numbers with $\sup_j n_j < \infty$ and $\alpha = (\alpha_j)_j \in \ell_\infty$ a bounded sequence with $\inf_j \operatorname{Re}(\alpha_j) > 0$ and $\sup_j |f_0(\alpha_j)| < \infty$.*

For $\widehat{L}_i(x,t)$, $C_{0,ij}$ as given in (7.8), (7.9), respectively, define operators generated by the following generalized infinite matrices

$$\begin{aligned} L(x,t) &= \Big(\widehat{L}_i(x,t)\, \Gamma_l(c_i) C_{0,ij} \Gamma_r(\overline{a}_j) \Big)_{i,j=1}^\infty, \\ L_0(x,t) &= \Big(\widehat{L}_i(x,t)\, \Gamma_l(c_i)(e^{(1)}_{n_j} \otimes e^{(1)}_{n_i}) \Gamma_r(a_j) \Big)_{i,j=1}^\infty. \end{aligned}$$

Then $L, L_0 \in \mathcal{N}_{\frac{2}{3}}\big(E(\ell_q(n_j))\big)$ for all $c = (c_j)_j \in E(\ell_q(n_j))$, $0 \neq a = (a_j)_j \in E(\ell_q(n_j))'$ $(1 \leq q \leq \infty)$, and the same holds true for the complex conjugate operators.

Moreover, $q = 1 - P/p$, where

$$P = \det_\lambda \begin{pmatrix} I & -L \\ \overline{L} & I + \overline{L}_0 \end{pmatrix}, \quad p = \det_\lambda \begin{pmatrix} I & -L \\ \overline{L} & I \end{pmatrix},$$

is a solution of the $\mathbb{C}$-reduced AKNS system system (4.5) on strips $\mathbb{R} \times (t_1, t_2)$ where the denominator p does not vanish.

In the above statement $\det_\lambda$ denotes the unique, continuous, spectral determinant on the $\frac{2}{3}$-Banach operator ideal $\mathcal{N}_{\frac{2}{3}}$ of $\frac{2}{3}$-nuclear operators. Again we will see later that the assumption on strips is superfluous.

Proof Consider the generalized diagonal operator $A \in \mathcal{L}\big(E(\ell_q(n_j))\big)$ generated by the sequence $(A_j)_j$ of Jordan blocks of dimension n_j corresponding to the eigenvalue α_j. It follows immediately that $\exp\big(Ax + f_0(A)t\big)$ is a generalized diagonal operator, too, which is generated by the sequence $\big(\exp\big(A_j x + f_0(A_j)t\big)\big)_j = \big(\widehat{L}_j(x,t)\big)_j$.

We also have to verify that $\exp(Ax)$ behaves sufficiently well for $x \to -\infty$. Since $(n_j)_j$ is bounded, and $\inf_j \operatorname{Re}(\alpha_j) > 0$, it sufficient to show this for a single Jordan block A_j. The latter follows from the fact that

$$\exp(A_j x) = \exp(\alpha_j x) \begin{pmatrix} p_j^{(0)} & & p_j^{(n_j-1)} \\ & \ddots & \\ 0 & & p_j^{(0)} \end{pmatrix}$$

with polynomials $p_j^{(0)} = 1, \ldots, p_j^{(n_j-1)}$ (compare (5.9) in the proof of Proposition 5.1.5).

Furthermore, the condition that $|f_0(\alpha_j)|$ be uniformly bounded implies that f_0 is holomorphic near $\operatorname{spec}(A)$.

By assumption, $0 \notin \operatorname{spec}(A) + \operatorname{spec}(\overline{A})$. Hence Proposition 2.3.4 guarantees the existence of $\Phi_{A,\overline{A}}^{-1}(\overline{a} \otimes c) \in \mathcal{N}_{\frac{2}{3}}\big(E(\ell_q(n_j))\big)$. We show that this operator is generated by the generalized infinite matrix $\big(\Gamma_l(c_i) C_{0,ij} \Gamma_r(\overline{a}_j)\big)_{i,j=1}^{\infty}$.

To this end we define the generalized diagonal operators

$$D_1 : \quad E(\ell_q(n_j)) \longrightarrow \ell_1(\ell_1(n_j)) \quad \text{generated by } \big(\Gamma_r(a_j)\big)_j,$$
$$D_2 : \quad \ell_\infty(\ell_\infty(n_j)) \longrightarrow E(\ell_q(n_j)) \quad \text{generated by } \big(\Gamma_l(c_j)\big)_j.$$

Both operators are bounded, which is verified as follows:

First note that for $a_j = (a_j^{(\mu)})_{\mu=1}^{n_j} \in \ell_{q'}(n_j)$ we have $\Gamma_r(a_j) = \sum_{\mu=1}^{n_j} a_j^{(\mu)} N_j^{\mu-1}$, where N_j is the nilpotent operator with 1 on the first off-diagonal as only non-vanishing entries. Thus $\| \Gamma_r(a_j) \| \leq \sum_{\mu=1}^{n_j} |a_j^{(\mu)}| = \| a_j \|_1 \leq n_j^{1/q} \| a_j \|_{q'}$ ($\frac{1}{q} + \frac{1}{q'} = 1$), and we infer

$$\begin{aligned} \| D_1 \xi \|_{\ell_1(\ell_1((n))} &= \sum_j \| \Gamma_r(a_j) \xi_j \|_1 \quad \leq \quad \sum_j \| \Gamma_r(a_j) \| \, \| \xi_j \|_q \\ &\leq \big(\sup_j n_j\big)^{1/q} \sum_j \| a_j \|_{q'} \| \xi_j \|_q \\ &\leq \big(\sup_j n_j\big)^{1/q} \ \| a \|_{E'(\ell_{q'}(n_j))} \ \| \xi \|_{E(\ell_q(n_j))} \end{aligned}$$

by Hölder's inequality. Similarly, $\| \Gamma_l(c_j) \| \leq n_j^{1/q'} \| c_j \|_q$ and

$$\begin{aligned} \| D_2 \xi \|_{E(\ell_q(n_j))} &= \| \big(\| \Gamma_l(c_j) \xi_j \|_q \big)_j \|_E \ \leq \ \big(\sup_j n_j\big)^{1/q'} \ \| \big(\| c_j \|_q \| \xi_j \|_\infty \big)_j \|_E \\ &\leq \big(\sup_j n_j\big)^{1/q'} \Big(\sup_j \| \xi_j \|_\infty \Big) \ \| \big(\| c_j \|_q \big)_j \|_E \end{aligned}$$

$$= \quad \big(\sup_j n_j\big)^{1/q'} \ \| \, c \, \|_{E(\ell_q(n_j))} \, \| \, \xi \, \|_{\ell_\infty(\ell_\infty(n_j))} \, .$$

Next, define C_0 as the operator generated by the generalized infinite matrix $(C_{0,ij})_{i,j}$. Then $C_0 \in \mathcal{L}\big(\ell_1(\ell_1(n_j)), \ell_\infty(\ell_\infty(n_j))\big)$ by Lemma 7.5.2.

Set

$$C := D_2 C_0 \overline{D}_1 \in \mathcal{L}\big(E(\ell_q(n_j))\big).$$

Observe that C is generated by the generalized infinite matrix $\big(\ \Gamma_l(c_i) C_{0,ij} \Gamma_r(\overline{a}_j)\ \big)_{i,j=1}^{\infty}$. Moreover, from $\Gamma_l(c_i) C_{0,ij} \Gamma_r(\overline{a}_j) = \Phi^{-1}_{A_i,\overline{A}_j}(\overline{a}_j \otimes c_i)$ it is straightforward to check that C solves $AC + C\overline{A} = \overline{a} \otimes c$. By uniqueness of this solution, $C = \Phi^{-1}_{A,\overline{A}}(\overline{a} \otimes c)$.

Finally, the one-dimensional operator $a \otimes c$ is generated by the generalized infinite matrix

$$\Big(\Gamma_l(c_i)\ \big(\ e^{(1)}_{n_j} \otimes e^{(1)}_{n_i}\ \big)\ \Gamma_r(a_j)\Big)_{i,j=1}^{\infty}$$

because $\Gamma_l(c_i)\ \big(\ e^{(1)}_{n_j} \otimes e^{(1)}_{n_i}\ \big)\ \Gamma_r(a_j) = \big(\Gamma_r(a_j)' e^{(1)}_{n_j}\big) \otimes \big(\Gamma_l(c_i) e^{(1)}_{n_i}\big) = a_j \otimes c_i$.

For later use we state that, in particular, $a \otimes c = D_2(e_0 \otimes e_0) D_1$, where $e_0 := (e^{(1)}_{n_j})_j$.

Now the assertion follows immediately by applying Theorem 7.1.7 with respect to the $\frac{2}{3}$-Banach operator ideal $\mathcal{N}_{\frac{2}{3}}$ of $\frac{2}{3}$-nuclear operators (see Lemma 7.1.2). $\square$

Remark 7.5.6. *Note that again Theorem 7.5.5 remains valid for any continuous determinant δ on a p-Banach operator ideal $\mathcal{A}$, $0 < p \le 1$, with the property $\overline{\delta(I+T)} = \delta(I + \overline{T})$ for all $T \in \mathcal{A}(E)$ satisfying $\overline{T} \in \mathcal{A}(E)$.*

7.5.3 Universal realization and regularity revisited

We round off the picture by providing also the universal realization of superposition of negatons and giving regularity conditions.

Theorem 7.5.7. *Let n_j, $j \in \mathbb{N}$, be natural numbers with $\sup_j n_j < \infty$ and $\alpha = (\alpha_j)_j \in \ell_\infty$ be a bounded sequence with $\inf_j \operatorname{Re}(\alpha_j) > 0$ and $\sup_j |f_0(\alpha_j)| < \infty$.*

For $\widehat{L}_i(x,t)$, $C_{0,ij}$ as given in (7.8), (7.9), respectively, define operators generated by the following generalized infinite matrices

$$\begin{aligned} L(x,t) &= \Big(\ \widehat{L}_i(x,t)\ \Gamma_l(d_i) C_{0,ij}\ \Big)_{i,j=1}^{\infty}, \\ L_0(x,t) &= \Big(\widehat{L}_i(x,t)\ \Gamma_l(d_i)(e^{(1)}_{n_j} \otimes e^{(1)}_{n_i})\Big)_{i,j=1}^{\infty}. \end{aligned}$$

Then L, $L_0 \in \mathcal{N}\big(\ell_1(\ell_1(n_j))\big)$ for all $0 \neq d = (d_j)_j \in \ell_1(\ell_1(n_j))$, and the same is true for the complex conjugate operators.

Moreover, $= 1 - P/p$, where

$$P = \det\nolimits_{\mathcal{N}} \begin{pmatrix} I & -L \\ \overline{L} & I + \overline{L}_0 \end{pmatrix}, \quad p = \det\nolimits_{\mathcal{N}} \begin{pmatrix} I & -L \\ \overline{L} & I \end{pmatrix},$$

is a solution of the $\mathbb{C}$-reduced AKNS system (4.5) on strips $\mathbb{R} \times (t_1, t_2)$ where the denominator p does not vanish.

In addition, each solution of Theorem 7.5.5 can be expressed explicitly in this way.

In the above statement $\det_{\mathcal{N}}$ denotes the unique continuous determinant on the Banach operator ideal $\mathcal{N}$ of nuclear operators restricted to the class of Banach spaces with approximation property.

Proof As for the proof that q is a solution of the $\mathbb{C}$-reduced AKNS system, we refer to Theorem 7.5.5 with the following modifications:

- The determinant $\det_{\mathcal{N}}$ on the nuclear component $\mathcal{N}\big(\ell_1(\ell_1(n_j))\big)$ is used, see Remark 7.5.6 and Lemma 7.1.2.
- The particular choice $d \in \ell_1(\ell_1(n_j))$, $e_0 := (e^{(1)}_{n_j})_j \in \ell_\infty(\ell_\infty(n_j))$ for the sequences is made.

Therefore, it only remains show that each solution of Theorem 7.5.5 can be expressed explicitly in this way. For concreteness we fix such a solution.

We start from the factorization of the operators L, $L_0 \in \mathcal{N}(\ell_q(n_j))$ in Theorem 7.5.5 as provided in the proof. Indeed, we have $L = \widehat{L}D_2C_0\overline{D}_1$, $L_0 = \widehat{L}D_2(e_0 \otimes e_0)D_1$, where $\widehat{L} = \operatorname{diag}\{\widehat{L}_j \mid j\}$ and all other ingredients are as defined in the proof of Theorem 7.5.5. Note that

$$C_0 \in \mathcal{N}_{\frac{2}{3}}\big(\ell_1(\ell_1(n_j)), \ell_\infty(\ell_\infty(n_j))\big)$$

as the unique solution of the equation $AX + X\widehat{A} = e_0 \otimes e_0$, where $A \in \mathcal{L}\big(\ell_\infty(\ell_\infty(n_j))\big)$, $\widehat{A} \in \mathcal{L}\big(\ell_1(\ell_1(n_j))\big)$ are the generalized diagonal operators generated by the sequences $(A_j)_j$, $(\overline{A}_j)_j$, respectively (see Proposition 2.3.4). Consequently the following operators are related

$$\begin{pmatrix} 0 & -L \\ \overline{L} & \overline{L_0} \end{pmatrix}, \qquad \begin{pmatrix} 0 & -M \\ \overline{M} & \overline{M_0} \end{pmatrix}$$

with $M_0 = D_1\widehat{L}D_2(e_0 \otimes e_0)$, $M = D_1\widehat{L}D_2C_0 \in \mathcal{N}_{\frac{2}{3}}\big(\ell_1(\ell_1((n_j)))\big)$ (This is verified as in the proof of Theorem 7.2.1). As a result,

$$\det{}_\lambda\left(I + \begin{pmatrix} 0 & -L \\ \overline{L} & \overline{L_0} \end{pmatrix}\right) = \det{}_\lambda\left(I + \begin{pmatrix} 0 & -M \\ \overline{M} & \overline{M_0} \end{pmatrix}\right),$$

$\det_\lambda$ denoting the spectral determinant on $\mathcal{N}_{\frac{2}{3}}$.

In the latter expression $\det_\lambda$ can be replaced by the nuclear determinant $\det_{\mathcal{N}}$ on the component $\mathcal{N}\big(\ell_1(\ell_1(n_j))\big)$, see Corollary 7.1.5.

Finally we observe that M, M_0 are of the form as required in the statement because $\Gamma_r(a_j)\Gamma_l(c_j) = \Gamma_l(d_j)$, where

$$d_j = \Big(\sum_{\kappa=1}^{n_j-\mu+1} a_j^{(\kappa)} c_j^{(\mu-1+\kappa)}\Big)_{\mu=1}^{n_j}.$$

Since we have $\| d_j \|_1 \le \sum_{\mu,\kappa=1}^{n_j} |a_j^{(\mu)}||c_j^{(\kappa)}| = \big(\sum_{\mu=1}^{n_j} |a_j^{(\mu)}|\big)\,\big(\sum_{\kappa=1}^{n_j} |c_j^{(\kappa)}|\big) = \| a_j \|_1 \| c_j \|_1$ $\le n_j \| a_j \|_{q'} \| c_j \|_q$, we infer $d \in \ell_1(\ell_1(n_j))$, again by Hölder's inequality.

Carrying out the same manipulations for the denominator p of the solution, we complete the proof. $\square$

Finally, we prove global regularity also for countable superpositons of negatons.

Theorem 7.5.8. *The solutions in Theorem 7.5.7 (and hence in Theorem 7.5.5) are defined and regular on all of* $\mathbb{R}^2$.

Proof The proof is virtually the same as that of Theorem 7.3.2. The only difference is that we now approximate by finite superposition of negatons. $\square$

7.5.4 Again sharper results for the Nonlinear Schrödinger equation

In Section 7.4 we have seen that, for the Nonlinear Schrödinger equation, it is possible to construct countable superpositions even in the case that 0 is contained in $\mathrm{spec}(A)+\mathrm{spec}(\overline{A})$. The main ingredient in the proof was an appropriate factorization of the formal solution C of the equation $AX+X\overline{A}=\overline{a}\otimes c$ to the effect that C belongs to the quasi-Banach operator ideal $\mathcal{L}_\infty \circ \mathcal{H} \circ \mathcal{L}_1$.

To extend this result to negatons, we need the following counterpart of Lemma 7.4.1.

Lemma 7.5.9. *Let n_j, $j \in \mathbb{N}$, be natural numbers with $\sup_j n_j < \infty$ and $\alpha = (\alpha_j)_j \in \ell_\infty$ a bounded sequence with $\mathrm{Re}(\alpha_j) > 0\ \forall j$.*

Then the operator $\widetilde{C}_0 : \ell_1(\ell_1(n_j)) \longrightarrow \ell_\infty(\ell_\infty(n_j))$ generated by the generalized infinite matrix

$$\widetilde{C}_0 = \Big(\mathrm{Re}(\alpha_j)^{n_j-\frac{1}{2}} \mathrm{Re}(\alpha_{j'})^{n_{j'}-\frac{1}{2}} C_{0,jj'} \Big)_{j,j'=1}^{\infty},$$

with $C_{0,jj'}$ as defined in (7.9), belongs to the quasi-Banach operator ideal $\mathcal{L}_\infty \circ \mathcal{H} \circ \mathcal{L}_1$.

The same statement holds for the complex conjugate operator $\overline{\widetilde{C}_0}$ of $\widetilde{C}_0$.

Proof We have to show that $\widetilde{C}_0$ factors through the Hilbert space $L_2[0,\infty)$. To this end, first observe the reformulation

$$C_{0,jj'} = \left(\frac{1}{(\nu-1)!} \frac{\partial^{\nu-1}}{\partial \alpha_j^{\nu-1}} \frac{1}{(\mu-1)!} \frac{\partial^{\mu-1}}{\partial \overline{\alpha}_{j'}^{\mu-1}} \frac{1}{\alpha_j + \overline{\alpha}_{j'}} \right)_{\substack{\nu=1,\dots,n_j \\ \mu=1,\dots,n_{j'}}}.$$

From this reformulation we can more or less read off the factorization of $\widetilde{C}_0$.

Consider the operator $S : \ell_1(\ell_1(n_j)) \longrightarrow L_2[0,\infty)$ which is defined on the standard basis $\{e_{n_j}^{(\mu)} \mid \mu = 1,\dots,n_j; j \in \mathbb{N}\}$ of $\ell_1(\ell_1(n_j))$ by

$$S e_{n_j}^{(\mu)} = \overline{f_j^{(\mu)}} \quad \text{for } f_j^{(\mu)}(s) = \mathrm{Re}(\alpha_j)^{n_j-\frac{1}{2}} \frac{1}{(\mu-1)!} \frac{\partial^{\mu-1}}{\partial \alpha_j^{\mu-1}} \exp\big(-\alpha_j s\big).$$

Then

$$\begin{aligned}
\langle S'S e_{n_{j'}}^{(\mu)}, e_{n_j}^{(\nu)} \rangle &= \langle S e_{n_{j'}}^{(\mu)}, S e_{n_j}^{(\nu)} \rangle \\
&= \int_0^\infty \overline{f_{j'}^{(\mu)}(s)} f_j^{(\nu)}(s) ds \\
&= \mathrm{Re}(\alpha_j)^{n_j-\frac{1}{2}} \mathrm{Re}(\alpha_{j'})^{n_{j'}-\frac{1}{2}} \int_0^\infty \frac{1}{(\nu-1)!} \frac{\partial^{\nu-1}}{\partial \alpha_j^{\nu-1}} \frac{1}{(\mu-1)!} \frac{\partial^{\mu-1}}{\partial \overline{\alpha}_{j'}^{\mu-1}} \exp\big(-(\alpha_j + \overline{\alpha}_{j'})s\big) ds \\
&= \mathrm{Re}(\alpha_j)^{n_j-\frac{1}{2}} \mathrm{Re}(\alpha_{j'})^{n_{j'}-\frac{1}{2}} \frac{1}{(\nu-1)!} \frac{\partial^{\nu-1}}{\partial \alpha_j^{\nu-1}} \frac{1}{(\mu-1)!} \frac{\partial^{\mu-1}}{\partial \overline{\alpha}_{j'}^{\mu-1}} \frac{1}{\alpha_j + \overline{\alpha}_{j'}} \\
&= \langle \widetilde{C}_0 e_{n_{j'}}^{(\mu)}, e_{n_j}^{(\nu)} \rangle.
\end{aligned}$$

Thus we have shown $\widetilde{C}_0 = S'S$.

It remains to prove that S is a bounded operator. Using the definition of $\widetilde{C}_0$ together with (7.9), we get

$$\begin{aligned}
\| f_j^{(\mu)} \|_{L_2[0,\infty)}^2 &= \langle \widetilde{C}_0 e_{n_j}^{(\mu)}, e_{n_j}^{(\mu)} \rangle \\
&= \big(\mathrm{Re}(\alpha_j)\big)^{2n_j-1} \binom{2\mu-2}{\mu-1} \left(\frac{1}{\alpha_j + \overline{\alpha}_j} \right)^{2\mu-1}
\end{aligned}$$

$$
\begin{aligned}
&= \left(\mathrm{Re}(\alpha_j)\right)^{2(n_j-\mu)} \binom{2\mu-2}{\mu-1} \left(\frac{1}{2}\right)^{2\mu-1} \\
&\le \frac{1}{2}(2n_j-2)! \max(1,|\alpha_j|)^{2n_j-2}
\end{aligned}
$$

Because $\sup_j n_j < \infty$ and $(\alpha_j)_j$ is a bounded sequence, there is a constant $c > 0$ such that $\| f_j^{(\mu)} \|_{L_2[0,\infty)} < c$ for all j. Thus we observe

$$
\begin{aligned}
\| S\xi \|_{L_2[0,\infty)} &= \| \sum_{j=1}^{\infty}\sum_{\mu=1}^{n_j} \xi_j^{(\mu)} f_j^{(\mu)} \|_{L_2[0,\infty)} \le \sum_{j=1}^{\infty}\sum_{\mu=1}^{n_j} |\xi_j^{(\mu)}| \, \| f_j^{(\mu)} \|_{L_2[0,\infty)} \\
&\le c \sum_{j=1}^{\infty}\sum_{\mu=1}^{n_j} |\xi_j^{(\mu)}| = c \, \| \xi \|_{\ell_1(\ell_1(n_j))},
\end{aligned}
$$

and S is bounded.

As for the complex conjugate operator $\overline{\widetilde{C}_0}$ of $\widetilde{C}_0$, a factorization $\overline{\widetilde{C}_0} = R'R$ through $L_2[0,\infty)$ is obtained via $R : \ell_1(\ell_1(n_j)) \to L_2[0,\infty)$ given by $Re_{n_j}^{(\mu)} = f_j^{(\mu)}$ with $f_j^{(\mu)}$ as above. □

The extension of Theorem 7.4.2 for negatons reads as follows.

Theorem 7.5.10. *Let n_j, $j \in \mathbb{N}$, be natural numbers with $\sup_j n_j < \infty$ and $\alpha = (\alpha_j)_j \in \ell_\infty$ be a bounded sequence with $\mathrm{Re}(\alpha_j) > 0 \ \forall j$.*

For $\widehat{L}_j(x,t)$, $C_{0,jj'}$ as given in (7.8) with $f_0(z) = -\mathrm{i}z^2$, (7.9), respectively, define operators generated by the following generalized infinite matrices

$$
\begin{aligned}
L(x,t) &= \left(\widehat{L}_j(x,t)\, \Gamma_l(c_j) C_{0,jj'} \Gamma_r(\overline{a}_{j'}) \right)_{j,j'=1}^{\infty}, \\
L_0(x,t) &= \left(\widehat{L}_j(x,t)\, \Gamma_l(c_j) (e_{n_{j'}}^{(1)} \otimes e_{n_j}^{(1)}) \Gamma_r(a_{j'}) \right)_{j,j'=1}^{\infty}.
\end{aligned}
$$

Then L, $L_0 \in \mathcal{L}_\infty \circ \mathcal{H} \circ \mathcal{L}_1\big(E(\ell_q(n_j))\big)$ for all $c = (c_j)_j$, $0 \neq a = (a_j)_j$ with $\left(c_j/\sqrt{\mathrm{Re}(\alpha_j)}^{2n_j-1}\right)_j \in E(\ell_q(n_j))$, $\left(a_j/\sqrt{\mathrm{Re}(\alpha_j)}^{2n_j-1}\right)_j \in E(\ell_q(n_j))'$, and the same is true for the complex conjugate operators.

Moreover, $q = 1 - P/p$, where

$$
P = \det\nolimits_\lambda \begin{pmatrix} I & -L \\ \overline{L} & I + \overline{L}_0 \end{pmatrix}, \quad p = \det\nolimits_\lambda \begin{pmatrix} I & -L \\ \overline{L} & I \end{pmatrix},
$$

is a solution of the Nonlinear Schrödinger equation (1.5) wherever the denominator p does not vanish.

In the above statement $\det_\lambda$ denotes the continuous and spectral determinant on the quasi-Banach operator ideal $\mathcal{L}_\infty \circ \mathcal{H} \circ \mathcal{L}_1$.

Proof Let us introduce the generalized diagonal operators

$$
\begin{aligned}
D_1 :\quad & E(\ell_q(n_j)) \to \ell_1(\ell_1(n_j)) \quad \text{generated by} \quad \left(\Gamma_r\left(a_j/\sqrt{\mathrm{Re}(\alpha_j)}^{2n_j-1}\right) \right)_j, \\
D_2 :\quad & \ell_\infty(\ell_\infty(n_j)) \to E(\ell_q(n_j)) \quad \text{generated by} \quad \left(\Gamma_l\left(c_j/\sqrt{\mathrm{Re}(\alpha_j)}^{2n_j-1}\right) \right)_j.
\end{aligned}
$$

Both operators are bounded which is shown precisely as in the proof of Theorem 7.5.5 with regard to the fact that $\left(a_j/\sqrt{\mathrm{Re}(\alpha_j)}^{2n_j-1}\right)_j \in E(\ell_q(n_j))'$, $\left(c_j/\sqrt{\mathrm{Re}(\alpha_j)}^{2n_j-1}\right)_j \in E(\ell_q(n_j))$. By Lemma 7.5.9, we observe that

$$
C := D_2 \, \widetilde{C}_0 \, \overline{D_1} \in \mathcal{L}_\infty \circ \mathcal{H} \circ \mathcal{L}_1\big(E(\ell_q(n_j))\big).
$$

Moreover, because $\Gamma_l\Big(c_j/\sqrt{\mathrm{Re}(\alpha_j)}^{2n_j-1}\Big) = \mathrm{Re}(\alpha_j)^{n_j-\frac{1}{2}}\Gamma_l(c_j)$, we easily see that C is generated by the generalized infinite matrix

$$C = \Big(\ \Gamma_l(c_j)C_{0,jj'}\Gamma_r(a_{j'})\ \Big)_{j,j'=1}^{\infty}.$$

This in turn shows that C satisfies the equation $AX + X\overline{A} = \overline{a} \otimes c$.

Similarly, $\overline{C} \in \mathcal{L}_\infty \circ \mathcal{H} \circ \mathcal{L}_1\big(E(\ell_q(n_j))\big)$ solves the equation $\overline{A}X + XA = a \otimes \overline{c}$.

For later use let us also remark the following factorization for the one-dimensional operator $a \otimes c$. It holds

$$a \otimes c = D_2\big(\widetilde{e}_0 \otimes \widetilde{e}_0\big)D_1,$$

with the vector $\widetilde{e}_0 = \big(\sqrt{\mathrm{Re}(\alpha_j)}^{2n_j-1} e_{n_j}^{(1)}\big)_j$. This can be checked blockwise for the generating infinite matrices by

$$\begin{aligned}
&\Gamma_l\big(c_j/\sqrt{\mathrm{Re}(\alpha_j)}^{2n_j-1}\big)\Big(\sqrt{\mathrm{Re}(\alpha_{j'})}^{2n_{j'}-1} e_{n_{j'}}^{(1)} \otimes \sqrt{\mathrm{Re}(\alpha_j)}^{2n_j-1} e_{n_j}^{(1)}\Big)\Gamma_r\big(a_{j'}/\sqrt{\mathrm{Re}(\alpha_{j'})}^{2n_{j'}-1}\big) = \\
&= \ \Gamma_l(c_j)\ \big(\ e_{n_{j'}}^{(1)} \otimes e_{n_j}^{(1)}\ \big)\ \Gamma_r(a_{j'}) \ = \ \big(\Gamma_r(a_{j'})' e_{n_{j'}}^{(1)}\big) \otimes \big(\Gamma_l(c_j) e_{n_j}^{(1)}\big) \ = \ a_{j'} \otimes c_j.
\end{aligned}$$

Now the assertion follows by applying Theorem 7.1.8 with respect to the quasi-Banach operator ideal $\mathcal{L}_\infty \circ \mathcal{H} \circ \mathcal{L}_1$ of operators factorizing through an L_1-space, a Hilbert space, and an L_∞-space (see Lemma 7.1.2). □

Next we give the counterpart of the universal realization in Theorem 7.5.7 for superpositions of negatons with the the weaker growth condition.

Theorem 7.5.11. *Let n_j, $j \in \mathbb{N}$, be natural numbers with $\sup_j n_j < \infty$ and $\alpha = (\alpha_j)_j \in \ell_\infty$ be a bounded sequence with $\mathrm{Re}(\alpha_j) > 0$ for all j.*

For $\widehat{L}_j(x,t)$, $C_{0,jj'}$ as given in (7.8) with $f_0(z) = -\mathrm{i}z^2$, (7.9), respectively, define operators generated by the following generalized infinite matrices

$$\begin{aligned}
L(x,t) \quad &= \quad \Big(\ \gamma_{jj'}\widehat{L}_j(x,t)\ \Gamma_l(d_j)C_{0,jj'}\ \Big)_{j,j'=1}^{\infty}, \\
L_0(x,t) \quad &= \quad \Big(\gamma_{jj'}\widehat{L}_j(x,t)\ \Gamma_l(d_j)(e_{n_{j'}}^{(1)} \otimes e_{n_j}^{(1)})\Big)_{j,j'=1}^{\infty}, \qquad \gamma_{jj'} := \frac{\mathrm{Re}(\alpha_{j'})^{n_{j'}-\frac{1}{2}}}{\mathrm{Re}(\alpha_j)^{n_j-\frac{1}{2}}}.
\end{aligned}$$

Then L, $L_0 \in \mathcal{N}\big(\ell_1(\ell_1(n_j))\big)$ for all $0 \neq d = (d_j)_j$ with $(d_j/\mathrm{Re}(\alpha_j)^{2n_j-1})_j \in \ell_1(\ell_1(n_j))$, and the same is true for the complex conjugate operators.

Moreover, $q = 1 - P/p$, where

$$P = \det\nolimits_{\mathcal{N}}\begin{pmatrix} I & -L \\ \overline{L} & I + \overline{L}_0 \end{pmatrix}, \quad p = \det\nolimits_{\mathcal{N}}\begin{pmatrix} I & -L \\ \overline{L} & I \end{pmatrix},$$

is a solution of the Nonlinear Schrödinger equation (1.5) wherever the denominator p does not vanish.

In addition, each solution of Theorem 7.5.10 can be expressed explicitly in this way.

In the above statement $\det_{\mathcal{N}}$ denotes the unique continuous determinant on the Banach operator ideal $\mathcal{N}$ of nuclear operators restricted to the class of Banach spaces with approximation property.

Remark 7.5.12. *a) It should be stressed that the role of the factors $\gamma_{jj'}$ is to ensure nuclearity of the operator*

$$\begin{pmatrix} 0 & -L \\ \overline{L} & \overline{L_0} \end{pmatrix}.$$

Once this is valid, they actually cancel in the evaluation of the nuclear determinant, because this can be done by taking the limit of finite-dimensional determinants, see (7.1).

b) If we want to get rid of the factors $\gamma_{jj'}$ in the solution formula, we need a sharper condition on the sequence d, confer Proposition 7.5.13.

Proof First we show that q solves the Nonlinear Schrödinger equation. Application of Theorem 7.5.10 with $(c_j)_j = \big(d_j/\mathrm{Re}(\alpha_j)^{n_j-\frac{1}{2}}\big)_j \in \ell_1(\ell_1(n_j))$, $(a_j)_j = \big(e_{n_j}^{(1)} \cdot \mathrm{Re}(\alpha_j)^{n_j-\frac{1}{2}}\big)_j \in \ell_\infty(\ell_\infty(n_j))$ yields the solution property of q expressed in terms of the determinant $\det_\lambda$ on the component $\mathcal{L}_\infty \circ \mathcal{H} \circ \mathcal{L}_1\big(\ell_1(\ell_1(n_j))\big)$. By Lemma 7.1.5, we can use the nuclear determinant $\det_{\mathcal{N}}$ on the component $\mathcal{N}\big(\ell_1(\ell_1(n_j)\big)$ instead.

Thus it remains to show that each solution of Theorem 7.5.10 can be expressed explicitly in this way. For concreteness fix such a solution. We start from the factorization of the operators L, $L_0 \in \mathcal{L}_\infty \circ \mathcal{H} \circ \mathcal{L}_1\big(E(\ell_q(n_j))\big)$ as provided in the proof. Indeed we have $L = \widehat{L} D_2 \widetilde{C}_0 \overline{D}_1$, $L_0 = \widehat{L} D_1 \big(\widetilde{e}_0 \otimes \widetilde{e}_0\big) D_1$, where $\widehat{L} = \mathrm{diag}\{\widehat{L}_j \mid j\}$ and all other ingredients are defined in the proof of Theorem 7.5.10. Recall in particular

$$\widetilde{C}_0 \in \mathcal{L}_\infty \circ \mathcal{H} \circ \mathcal{L}_1\big(\ell_1(\ell_1(n_j)), \ell_\infty(\ell_\infty(n_j))\big).$$

Thus, the following operators are related

$$\begin{pmatrix} 0 & -L \\ \overline{L} & \overline{L_0} \end{pmatrix}, \qquad \begin{pmatrix} 0 & -M \\ \overline{M} & \overline{M_0} \end{pmatrix}$$

with $M_0 = D_1 \widehat{L} D_2 (\widetilde{e}_0 \otimes \widetilde{e}_0)$, $M = D_1 \widehat{L} D_2 \widetilde{C}_0 \in \mathcal{L}_\infty \circ \mathcal{H} \circ \mathcal{L}_1\big(\ell_1(\ell_1(n_j))\big)$. Hence, by the principle of related operators

$$\det\nolimits_\lambda \left(I + \begin{pmatrix} 0 & -L \\ \overline{L} & \overline{L_0} \end{pmatrix}\right) = \det\nolimits_\lambda \left(I + \begin{pmatrix} 0 & -M \\ \overline{M} & \overline{M_0} \end{pmatrix}\right),$$

$\det_\lambda$ denoting the spectral determinant on $\mathcal{L}_\infty \circ \mathcal{H} \circ \mathcal{L}_1$.

In the latter expression $\det_\lambda$ can be replaced by the nuclear determinant $\det_{\mathcal{N}}$ on the component $\mathcal{N}\big(\ell_1(\ell_1(n_j))\big)$, see Corollary 7.1.5.

Finally we observe that M, M_0 are of the form as required in the statement because $\Gamma_r(a_j)\Gamma_l(c_j) = \Gamma_l(d_j)$, where

$$d_j = \Big(\sum_{\kappa=1}^{n_j-\mu+1} a_j^{(\kappa)} c_j^{(\mu-1+\kappa)} \Big)_{\mu=1}^{n_j}.$$

To check $\big(d_j/\mathrm{Re}(\alpha_j)^{2n_j-1}\big)_i \in \ell_1(\ell_1(n_j)n)$, we need the simple estimate

$$\begin{aligned}
\| d_j/\mathrm{Re}(\alpha_j)^{2n_j-1} \|_1 &\le \Big(\sum_{\mu,\kappa=1}^{n_i} |a_i^{(\mu)}||c_i^{(\kappa)}| \Big) \mathrm{Re}(\alpha_i)^{2n_i-1} \\
&= \| a_j/\sqrt{\mathrm{Re}(\alpha_j)}^{2n_j-1} \|_1 \, \| c_j/\sqrt{\mathrm{Re}(\alpha_j)}^{2n_j-1} \|_1 \\
&\le n_j \, \| a_i/\sqrt{\mathrm{Re}(\alpha_i)}^{2n_i-1} \|_{q'} \, \| c_i/\sqrt{\mathrm{Re}(\alpha_i)}^{2n_i-1} \|_q,
\end{aligned}$$

and then apply Hölder's inequality again.

Carrying out the same manipulations for the denominator p of the solution, we complete the proof. $\square$

Proposition 7.5.13. *Let n_j, $j \in \mathbb{N}$, be natural numbers with $n_0 := \sup_j n_j < \infty$ and $\alpha = (\alpha_j)_j \in \ell_\infty$ be a bounded sequence with $\mathrm{Re}(\alpha_j) > 0$ for all j.*

For $\widehat{L}_j(x,t)$, $C_{0,jj'}$ as given in (7.8) with $f_0(z) = -\mathrm{i}z^2$, (7.9), respectively, we define operators generated by the following generalized infinite matrices

$$\begin{aligned} L(x,t) &= \Big(\widehat{L}_j(x,t)\, \Gamma_l(d_j) C_{0,jj'} \Big)_{j,j'=1}^{\infty}, \\ L_0(x,t) &= \Big(\widehat{L}_j(x,t)\, \Gamma_l(d_j) (e_{n_{j'}}^{(1)} \otimes e_{n_j}^{(1)}) \Big)_{j,j'=1}^{\infty}, \end{aligned}$$

Then L, $L_0 \in \mathcal{N}\big(\ell_1(\ell_1(n_j))\big)$ for all $0 \neq d = (d_i)_i$ with $(d_j/\mathrm{Re}(\alpha_j)^{2n_0-1})_j \in \ell_1(\ell_1(n_j))$, and the same is true for the complex conjugate operators.

Moreover, $q = 1 - P/p$, where

$$P = \det_{\mathcal{N}} \begin{pmatrix} I & -L \\ \overline{L} & I + \overline{L}_0 \end{pmatrix}, \quad p = \det_{\mathcal{N}} \begin{pmatrix} I & -L \\ \overline{L} & I \end{pmatrix},$$

is a solution of the Nonlinear Schrödinger equation (1.5) wherever the denominator p does not vanish.

Proof To show that q solves the Nonlinear Schrödinger equation, consider the operator $C : \ell_1(\ell_1(n_j)) \longrightarrow \ell_1(\ell_1(n_j))$ generated by the generalized infinite matrix

$$C = \Big(\Gamma_l(d_j) C_{0,jj'} \Big)_{j,j'=1}^{\infty}.$$

Our first aim is to show $C \in \mathcal{N}\big(\ell_1(\ell_1(n_j))\big)$.

Observe

$$\Gamma_l(d_j) C_{0,jj'} = \Big(\sum_{\kappa=1}^{n_j-\nu+1} d_j^{(\kappa+\nu-1)} c_{0,jj'}^{(\kappa\mu)} \Big)_{\substack{\nu=1,\dots,n_j \\ \mu=1,\dots,n_{j'}}},$$

where $c_{0,jj'}^{(\nu\mu)}$ abbreviates the $(\nu\mu)$-th entry of $C_{0,jj'}$ as given in (7.9). Thus, by (7.9) and the obvious estimate $|\alpha_j + \overline{\alpha}_{j'}| \geq \mathrm{Re}(\alpha_j)$ $\forall j$ (recall $\mathrm{Re}(\alpha_j) > 0$ $\forall j$ by assumption), we infer

$$\begin{aligned} \mathrm{Re}(\alpha_j)^{2n_0-1}\, |c_{0,jj'}^{(\nu\mu)}| &\leq \binom{\nu+\mu-2}{\nu-1} \mathrm{Re}(\alpha_j)^{(n_0-\nu)+(n_0-\mu)} \\ &\leq (n_j + n_{j'} - 2)!\ \max\big(1, |\alpha_j|\big)^{2n_0-2} \\ &\leq (2n_0 - 2)!\ \max\big(1, \| \alpha \|_\infty\big)^{2n_0-2} \ =: c \end{aligned}$$

for all $\nu = 1, \dots, n_j$, $\mu = 1, \dots, n_{j'}$.
Therefore,

$$\begin{aligned} \Big| \sum_{\kappa=1}^{n_j-\nu+1} d_j^{(\kappa+\nu-1)} c_{0,jj'}^{(\kappa\mu)} \Big| &\leq c \sum_{\kappa=1}^{n_j-\nu+1} |d_j^{(\kappa+\nu-1)}| / \mathrm{Re}(\alpha_j)^{2n_0-1} \\ &\leq c \, \| d_j \|_1 / \mathrm{Re}(\alpha_j)^{2n_0-1}. \end{aligned}$$

Therefore, abbreviating the $(\nu\mu)$-th entry of the (jj')-th block of C by $c_{jj'}^{(\nu\mu)}$ as above, the nuclear norm of C can by calculated as

$$\| C \,|\, \mathcal{N} \| = \sum_{j=1}^{\infty} \sum_{\nu=1}^{n_j} \sup_{j' \in \mathbb{N}} \sup_{\mu=1,\dots,n_{j'}} |c_{jj'}^{(\nu\mu)}|$$

$$
\begin{aligned}
&= \sum_{j=1}^{\infty} \sum_{\nu=1}^{n_j} \sup_{j' \in \mathbb{N}} \sup_{\mu=1,\dots,n_{j'}} \; | \sum_{\kappa=1}^{n_j-\nu+1} d_j^{(\kappa+\nu-1)} c_{0,jj'}^{(\kappa\mu)} | \\
&\leq c \sum_{j=1}^{\infty} n_j \parallel d_j \parallel_1 / \mathrm{Re}(\alpha_j)^{2n_0-1} \\
&\leq c\, n_0 \parallel (d_j/\mathrm{Re}(\alpha_j)^{2n_0-1})_j \parallel_{\ell_1(\ell_1(n_j))} < \infty.
\end{aligned}
$$

Thus C is indeed nuclear.

It is straightforward to verify that C solves the equation $AX + X\overline{A} = e_0 \otimes d$ for $e_0 = (e_{n_j}^{(1)})_j \in \ell_\infty(\ell_\infty(n_j))$, $d = (d_j)_j \in \ell_1(\ell_1(n_j))$ (here we have used $(\alpha_j)_j \in \ell_\infty$, $\big(d_j/\mathrm{Re}(\alpha_j)^{2n_j-1}\big)_j \in \ell_1(\ell_1(n_j))$, and $n_j \leq n_0$ for all j).

Thus the assertion follows by applying Theorem 7.1.8 with respect to the nuclear determinant $\det_{\mathcal{N}}$ on $\ell_1(\ell_1(n_j))$ (see Lemma 7.1.2). □

Again we have global regularity of the solutions constructed so far.

Theorem 7.5.14. *The solutions in Theorem 7.5.11 (and hence in Theorem 7.5.10) and Proposition 7.5.13 are defined and regular on all of* $\mathbb{R}^2$.

7.6 Countable superpositions for the $\mathbb{R}$-reduced AKNS system, and the modified Korteweg-de Vries equation

In the last section we turn to the $\mathbb{R}$-reduced AKNS system. First we prove the basic solution formulas for the $\mathbb{R}$-reduced AKNS system, as well as an ameliorated version for the modified Korteweg-de Vries equation. Note that these formulas are considerably simpler than the ones used for the $\mathbb{C}$-reductions. Moreover, due to the fact that the relation $r = -q$ can already be realized on the abstract operator level, the concept of the complex conjugate operator is superfluous for the $\mathbb{R}$-reduced AKNS system.

Finally, we briefly outline the results on countable superpositions, including regularity and reality conditions. It should be emphasized that these conditions are general enough to include not only the usual solitons but also breathers into the countable superpositions.

7.6.1 Basic solution formulas

We start with the solution formula for the $\mathbb{R}$-reduced AKNS system in general, which is valid for generating operators A on sequence spaces. It is an immediate consequence of Theorem 2.5.1 and reads as follows.

Theorem 7.6.1. *Let E be a classical sequence space and $\mathcal{A}$ a p-Banach operator ideal $(0 < p \leq 1)$ with a continuous determinant δ.*

Let $A \in \mathcal{L}(E)$ with $0 \notin \mathrm{spec}(A) + \mathrm{spec}(A)$, let $\mathrm{spec}(A)$ be contained in the domain where f_0 is holomorphic, and assume that $\exp(Ax)$ behaves sufficiently well for $x \to -\infty$. Then the operator-function

$$L(x,t) = \exp\big(\, Ax + f_0(A)t \,\big)\, \Phi_{A,A}^{-1}(a \otimes c).$$

belongs to $\mathcal{A}$ for arbitrary $0 \neq a \in E'$, $c \in E$.

Moreover,

$$q \;=\; \mathrm{i}\frac{\partial}{\partial x} \log \frac{\delta\,(\, I - \mathrm{i}L \,)}{\delta\,(\, I + \mathrm{i}L \,)}$$

solves the $\mathbb{R}$-reduced AKNS system (4.6) on every strip $\mathbb{R} \times (t_1, t_2)$ on which both determinants do not vanish.

As for the modified Korteweg-de Vries equation, we will build on the following ameliorated solution formula which is an immediate consequence of Proposition 2.6.2.

Theorem 7.6.2. *Let E be a classical sequence space and $\mathcal{A}$ a quasi-Banach operator ideal with a continuous determinant δ.*

Let $A \in \mathcal{L}(E)$, and let $0 \neq a \in E'$, $c \in E$ be given. Assume that there exists $C \in \mathcal{A}(E)$ such that $AC + CA = a \otimes c$. Then the operator-function

$$L(x,t) = \exp\left(Ax - A^3 t \right) C,$$

belongs to $\mathcal{A}$, and

$$q \;=\; \mathrm{i}\frac{\partial}{\partial x} \log \frac{\delta\left(I - \mathrm{i}L \right)}{\delta\left(I + \mathrm{i}L \right)}$$

solves the modified Korteweg-de Vries equation (1.6) wherever both determinants do not vanish.

7.6.2 Review of the results

Let us start with the construction of countable superpositions of negatons for the $\mathbb{R}$-reduced AKNS system.

To this end we adapt the material on negatons according to the situation at hand. Here the essential building blocks are

$$\widehat{L}_i(x,t) := \Gamma_r\big(\ell_i(x,t)\big) \tag{7.10}$$
$$\text{for the vector } \ell_i(x,t) = \left(\frac{1}{(\nu-1)!}\frac{\partial^{\nu-1}}{\partial \alpha_i^{\nu-1}} \exp\big(\alpha_i x + f_0(\alpha_i)t\big) \right)_{\nu=1}^{n_i}$$

as before, and

$$C_{0,ij} := \left((-1)^{\nu+\mu}\binom{\nu+\mu-2}{\mu-1}\left(\frac{1}{\alpha_i+\alpha_j}\right)^{\nu+\mu-1} \right)_{\substack{\nu=1,\dots,n_i\\ \mu=1,\dots,n_j}}. \tag{7.11}$$

Note that in the latter expression complex conjugate terms are absent.

The following theorem sums up the result on countable superpositions of negatons for the $\mathbb{R}$-reduction as a whole. It also gives the corresponding universal realization.

Theorem 7.6.3. *Let n_j, $j \in \mathbb{N}$, be natural numbers with $\sup_j n_j < \infty$ and $\alpha = (\alpha_j)_j \in \ell_\infty$ be a bounded sequence with $\inf_j \mathrm{Re}(\alpha_j) > 0$ and $\sup_j |f_0(\alpha_j)| < \infty$.*

a) Then, for $\widehat{L}_i(x,t)$, $C_{0,ij}$ as given in (7.10), (7.11), respectively, the operator generated by the generalized infinite matrix

$$L(x,t) = \left(\widehat{L}_i(x,t)\, \Gamma_l(c_i) C_{0,ij} \Gamma_r(a_j) \right)_{i,j=1}^{\infty} \tag{7.12}$$

belongs to the component $\mathcal{N}_{\frac{2}{3}}\big(E(\ell_q(n_j))\big)$ of the $\frac{2}{3}$-Banach operator ideal $\mathcal{N}_{\frac{2}{3}}$ for all sequences $c = (c_j)_j \in E(\ell_q(n_j))$, $a = (a_j)_j \in E(\ell_q(n_j))'$. Moreover

$$q \;=\; \mathrm{i}\frac{\partial}{\partial x} \log \frac{\det_\lambda\left(I - \mathrm{i}L \right)}{\det_\lambda\left(I + \mathrm{i}L \right)} \tag{7.13}$$

is a solution of the $\mathbb{R}$-reduced AKNS system (4.6) on strips $\mathbb{R} \times (t_1,t_2)$ where both determinants $\det_\lambda(I \pm \mathrm{i}L)$ do not vanish.

b) Each of the solutions in a) can be expressed explicitly in the form

$$q \;=\; \mathrm{i}\frac{\partial}{\partial x}\log\frac{\det_{\mathcal{N}}\,\left(\,I-\mathrm{i}M\,\right)}{\det_{\mathcal{N}}\,\left(\,I+\mathrm{i}M\,\right)},$$

where the operator $M = M(x,t) \in \mathcal{N}\big(\ell_1(\ell_1(n_j))\big)$ *is generated by the special generalized infinite matrix*

$$M(x,t) = \Big(\widehat{L}_i(x,t)\;\Gamma_l(d_i)C_{0,ij}\Big)_{i,j=1}^{\infty},$$

and we have $0 \neq d = (d_j)_j \in \ell_1(\ell_1(n_j))$.

In the above statement $\det_\lambda$ denotes the spectral determinant on the $\frac{2}{3}$-Banach operator ideal $\mathcal{N}_{\frac{2}{3}}$, and $\det_{\mathcal{N}}$ is the nuclear determinant on the component $\mathcal{N}\big(\ell_1(\ell_1(n_j))\big)$.

Countable superpositions of solitons are comprised in Theorem 7.6.3 in the particular case $n_j = 1$ for all j. In this case, the operator (7.12) becomes

$$L(x,t) = \left(\frac{a_j c_i}{\alpha_i+\alpha_j}\exp\;\left(\;\alpha_i x + f_0(\alpha_i)t\;\right)\right)_{i,j=1}^{\infty},$$

where $a = (a_j)_j$, $c = (c_j)_j$ now are ordinary sequences.

We round off the picture by providing the sharpened result for the modified Korteweg-de Vries equation, and give again its universal realization.

Theorem 7.6.4. *Let* n_j, $j \in \mathbb{N}$, *be natural numbers with* $\sup_j n_j < \infty$ *and* $\alpha = (\alpha_j)_j \in \ell_\infty$ *be a bounded sequence with* $\mathrm{Re}(\alpha_j) > 0\ \forall j$.

a) Then, for $\widehat{L}_i(x,t)$, $C_{0,ij}$ *as given in (7.10) with* $f_0(z) = -z^3$, *(7.11), respectively, the operator generated by the generalized infinite matrix*

$$L(x,t) = \Big(\;\widehat{L}_i(x,t)\;\Gamma_l(c_i)C_{0,ij}\Gamma_r(a_j)\;\Big)_{i,j=1}^{\infty}$$

belongs to the component $\mathcal{L}_\infty \circ \mathcal{H} \circ \mathcal{L}_1\big(E(\ell_q(n_j))\big)$ *of the quasi-Banach operator ideal* $\mathcal{L}_\infty \circ \mathcal{H} \circ \mathcal{L}_1$ *for all sequences* $c = (c_j)_j$, $0 \neq a = (a_j)_j$ *with*

$$\Big(c_j/\sqrt{\mathrm{Re}(\alpha_j)}^{\,2n_j-1}\Big)_j \in E(\ell_q(n_j)), \quad \Big(a_j/\sqrt{\mathrm{Re}(\alpha_j)}^{\,2n_j-1}\Big)_j \in E(\ell_q(n_j))'.$$

Moreover

$$q \;=\; \mathrm{i}\frac{\partial}{\partial x}\log\frac{\det_\lambda\,\left(\,I-\mathrm{i}L\,\right)}{\det_\lambda\,\left(\,I+\mathrm{i}L\,\right)}$$

is a solution of the modified Korteweg-de Vries equation (1.6) wherever both determinants $\det_\lambda(I \pm \mathrm{i}L)$ *do not vanish.*

b) Each of the solutions in a) can be expressed explicitly in the form

$$q \;=\; \mathrm{i}\frac{\partial}{\partial x}\log\frac{\det_{\mathcal{N}}\,\left(\,I-\mathrm{i}M\,\right)}{\det_{\mathcal{N}}\,\left(\,I+\mathrm{i}M\,\right)},$$

where the operator $M = M(x,t) \in \mathcal{N}\big(\ell_1(\ell_1(n_j))\big)$ *is generated by the special generalized infinite matrix*

$$M(x,t) = \Big(\gamma_{ij}\;\widehat{L}_i(x,t)\;\Gamma_l(d_i)C_{0,ij}\Big)_{i,j=1}^{\infty},$$

with a vector $0 \neq d = (d_j)_j$ *satisfying* $\big(d_j/\mathrm{Re}(\alpha_j)^{2n_j-1}\big)_i \in \ell_1(\ell_1(n_j))$, *and we have* $\gamma_{ij} = \mathrm{Re}(\alpha_j)^{n_j-\frac{1}{2}}/\mathrm{Re}(\alpha_i)^{n_i-\frac{1}{2}}$.

In the above statement $\det_\lambda$ denotes the spectral determinant on $\mathcal{L}_\infty \circ \mathcal{H} \circ \mathcal{L}_1$, and again $\det_{\mathcal{N}}$ is the nuclear determinant on the component $\mathcal{N}\big(\ell_1(\ell_1(n_j))\big)$.

The parts b) of Theorem 7.6.3 and Theorem 7.6.4 give us access to reality conditions global and regularity for countable superpositions of negatons as well for the $\mathbb{R}$-reduced AKNS system as for the sharpener result in the context of the modified Korteweg-de Vries equation.

Theorem 7.6.5. *Let the α_j be as assumed in Theorem 7.6.3 or Theorem 7.6.4, respectively, with the following additional properties:*

(i) If $\alpha_j \in \mathbb{R}$, then $d_j \in \ell_q(n_j, \mathbb{R})$.

(ii) There is a permutation π of $J = \{j \in \mathbb{N} \mid \alpha_j \notin \mathbb{R}\}$ such that for each $j \in J$ we have $\alpha_{\pi(j)} = \overline{\alpha}_j$ and $d_{\pi(j)} = \overline{d_j}$.

Then the solutions in Theorems 7.6.3 or Theorem 7.6.4, respectively, are defined on all of $\mathbb{R}^2$, real-valued, and regular.

Note that real eigenvalues, see (i), lead to solitons/antisolitons, whereas pairs of complex conjugate eigenvalues as in (ii) give rise to breathers (see Section 4.4.3). Thus breathers are indeed contained in our treatment of countable superpositions.

Recall finally that the factors γ_{ij} in Theorem 7.6.4 are necessary to ensure $M \in \mathcal{N}$, but they cancel in the actual evaluation of the nuclear determinant. Note, however, that in the soliton case these factors always equal 1. In the negaton case, a slightly stronger condition ensures nuclearity of M even for $\gamma_{ij} = 1$.

Proposition 7.6.6. *If the sequence $d = (d_j)_j$ in Theorem 7.6.4 b) even satisfies*

$$(d_j/\mathrm{Re}(\alpha_j)^{2n_0-1})_j \in \ell_1(\ell_1(n_j)),$$

where $n_0 := \sup_j n_j$, then the factors γ_{ij} can be chosen equal to 1.

Appendix A

Connection to Wronskian representations

Up to now our treatment of solitons and negatons was exclusively based on our operator-theoretical approach. Although our work was inspired by the method of Marchenko [55], the structure of the resulting solution formulas is quite different. Even in the simplest case of N-solitons, the coincidence of the corresponding formulas is not obvious at all.

In Marchenko's work, N-solitons are represented in terms of Wronskian determinants. The latter appear also in constructions based on Darboux transforms (confer [56]). For our purposes the construction of positions by Matveev and ensuing developements are very important ([57], [58], [59], [60], see also [11], [79], [95]). For related material the interested reader may also consult [30], [31], [62], [63], [67].

The purpose of the present appendix is to establish the link between our solution formulas and those constructed by Wronskian techniques.

A.1 The pure soliton case

For the $\mathbb{C}$-reduced AKNS system, a representation of n-solitons in terms of Wronskian-type determinants is given by

$$q = 2\mathrm{i}(-1)^n \Delta_1/\Delta_2, \tag{A.1}$$

with

$$\Delta_1 = \det \begin{pmatrix} \lambda_1^n & \cdots & \lambda_1 & 1 & -\lambda_1^{n-2}\varphi_1 & \cdots & -\lambda_1\varphi_1 & -\varphi_1 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots \\ \lambda_n^n & \cdots & \lambda_n & 1 & -\lambda_n^{n-2}\varphi_n & \cdots & -\lambda_n\varphi_n & -\varphi_n \\ \overline{\lambda}_1^n & \cdots & \overline{\lambda}_1 & 1 & \overline{\lambda}_1^{n-2}/\overline{\varphi}_1 & \cdots & \overline{\lambda}_1/\overline{\varphi}_1 & 1/\overline{\varphi}_1 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots \\ \overline{\lambda}_n^n & \cdots & \overline{\lambda}_n & 1 & \overline{\lambda}_n^{n-2}/\overline{\varphi}_n & \cdots & \overline{\lambda}_n/\overline{\varphi}_n & 1/\overline{\varphi}_n \end{pmatrix}$$

and

$$\Delta_2 = \det \begin{pmatrix} \lambda_1^{n-1} & \cdots & \lambda_1 & 1 & -\lambda_1^{n-1}\varphi_1 & \cdots & -\lambda_1\varphi_1 & -\varphi_1 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots \\ \lambda_n^{n-1} & \cdots & \lambda_n & 1 & -\lambda_n^{n-1}\varphi_n & \cdots & -\lambda_n\varphi_n & -\varphi_n \\ \overline{\lambda}_1^{n-1} & \cdots & \overline{\lambda}_1 & 1 & \overline{\lambda}_1^{n-1}/\overline{\varphi}_1 & \cdots & \overline{\lambda}_1/\overline{\varphi}_1 & 1/\overline{\varphi}_1 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots \\ \overline{\lambda}_n^{n-1} & \cdots & \overline{\lambda}_n & 1 & \overline{\lambda}_n^{n-1}/\overline{\varphi}_n & \cdots & \overline{\lambda}_n/\overline{\varphi}_n & 1/\overline{\varphi}_n \end{pmatrix},$$

where the functions φ_j are defined by $\varphi_j(x,t) = \exp\big(2\mathrm{i}\lambda_j x + f_0(2\mathrm{i}\lambda_j)t + \varphi_j^{(0)}\big)$ and the constants λ_j are assumed to be pairwise different.

Note that the particular case $n = 1$ yields $q = -2\mathrm{i}(\lambda - \overline{\lambda})/(\varphi + 1/\overline{\varphi})$, which coincides with the one-soliton solution (4.10) of the $\mathbb{C}$-reduced AKNS system if we set $\alpha := 2\mathrm{i}\lambda$.

The representation (A.1) is taken from [62] combined with [63]. Similar formulas were already derived by Matveev [56] in the context of Darboux transformations for the zero curvature equations.

In the following lemma we show how to translate (A.1) into our framework.

Proposition A.1.1. *Every soliton solution as in (A.1) can be realized as one of the solutions constructed in Proposition 4.3.1 with A being a diagonal matrix.*

Proof Step 1: To simplify arguments, rename $\alpha_j = 2\mathrm{i}\lambda_j$, $\overline{\alpha}_j = -2\mathrm{i}\overline{\lambda}_j$. Then q in (A.1) is given by almost the same formula with λ_j, $\overline{\lambda}_j$ replaced by α_j, $-\overline{\alpha}_j$ and the factor 2i in front of the whole formula cancelled.

Step 2: Next we take a closer look on the numerator Δ_1. Interchanging columns, we can put the first into $(n+1)$-th column without changing the order of the other columns. Now we can compare Δ_1 with the denominator Δ_2, from which it only differs in the $(n+1)$-th column. Thus $\Delta_1 = (-1)^n\big(\Delta_2 - \Delta\big)$ with

$$\Delta = \det\begin{pmatrix} \alpha_1^{n-1} & \cdots & 1 & -\alpha_1^{n-1}\varphi_1 - \alpha_1^n & -\alpha_1^{n-2}\varphi_1 & \cdots & -\varphi_1 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \alpha_n^{n-1} & \cdots & 1 & -\alpha_n^{n-1}\varphi_n - \alpha_n^n & -\alpha_n^{n-2}\varphi_n & \cdots & -\varphi_n \\ (-\overline{\alpha}_1)^{n-1} & \cdots & 1 & (-\overline{\alpha}_1)^{n-1}/\overline{\varphi}_1 - (-\overline{\alpha}_1)^n & (-\overline{\alpha}_1)^{n-2}/\overline{\varphi}_1 & \cdots & 1/\overline{\varphi}_1 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ (-\overline{\alpha}_n)^{n-1} & \cdots & 1 & (-\overline{\alpha}_n)^{n-1}/\overline{\varphi}_n - (-\overline{\alpha}_n)^n & (-\overline{\alpha}_n)^{n-2}/\overline{\varphi}_n & \cdots & 1/\overline{\varphi}_n \end{pmatrix} \tag{A.2}$$

Next we introduce $A = \operatorname{diag}\{\alpha_j \mid j = 1, \ldots, n\}$, $L = \operatorname{diag}\{\varphi_j \mid j = 1, \ldots, n\}$ and the two Vandermonde matrices $V = \big(\alpha_i^{n-j}\big)_{i,j=1}^n$, $W = \big((-\overline{\alpha}_i)^{n-j}\big)_{i,j=1}^n$. Then, as a result of the manipulations of this step, we get

$$\begin{aligned} q &= 1 - \Delta/\Delta_2 \\ &\qquad \text{with } \Delta_2 = \det\begin{pmatrix} V & -LV \\ W & \overline{L}^{-1}W \end{pmatrix} \quad \text{and} \quad \Delta = \det\begin{pmatrix} V & -LV - AV \\ W & \overline{L}^{-1}W + \overline{A}W \end{pmatrix}. \end{aligned}$$

Note that we subtracted in (A.2) the j-th from the $(n+j+1)$-th column (for $j = 1, \ldots, n-1$) to obtain the expression of Δ.

Step 3: We proceed by the subsequent simultaneous manipulations for numerator Δ and denominator Δ_2.

$$\begin{aligned} \Delta_2 &= \det\left(\begin{pmatrix} I & 0 \\ 0 & \overline{L}^{-1} \end{pmatrix}\begin{pmatrix} I & -LVW^{-1} \\ \overline{L}WV^{-1} & I \end{pmatrix}\begin{pmatrix} V & 0 \\ 0 & W \end{pmatrix}\right) \\ &= \det(\overline{L}^{-1}VW)\ \det\begin{pmatrix} I & -LVW^{-1} \\ \overline{L}WV^{-1} & I \end{pmatrix} \end{aligned}$$

and

$$\Delta = \det\left(\begin{pmatrix} I & 0 \\ 0 & \overline{L}^{-1} \end{pmatrix}\begin{pmatrix} I & -LVW^{-1} - AVW^{-1} \\ \overline{L}WV^{-1} & I + \overline{LA} \end{pmatrix}\begin{pmatrix} V & 0 \\ 0 & W \end{pmatrix}\right)$$

$$= \det(\overline{L}^{-1}VW)\,\det\begin{pmatrix} I & -LVW^{-1}-AVW^{-1} \\ \overline{L}WV^{-1} & I+\overline{LA} \end{pmatrix}$$

$$= \det(\overline{L}^{-1}VW)\,\det\begin{pmatrix} I & -LVW^{-1} \\ \overline{L}WV^{-1} & I+\overline{L}WV^{-1}\left(VW^{-1}\overline{A}+AVW^{-1}\right) \end{pmatrix},$$

where we have used Lemma A.1.2 for the latter reformulation.

Set $C = VW^{-1}$. Then, by definition of V, W, it is clear that $W = \epsilon\overline{V}\Sigma$ for $\epsilon = (-1)^{n-1}$ and $\Sigma = \operatorname{diag}\{(-1)^{j-1} \mid j = 1, \ldots, n\}$, and the following calculation shows $WV^{-1} = \overline{C}$,

$$WV^{-1} = \epsilon\overline{V}\Sigma\, V^{-1} = \epsilon\overline{V}\,\overline{\left(\Sigma\overline{V^{-1}}\right)} = \overline{V\left(\epsilon\Sigma^{-1}\overline{V}^{-1}\right)} = \overline{VW^{-1}}.$$

To sum up,

$$q = 1 - \frac{\det\begin{pmatrix} I & -LC \\ \overline{LC} & I+\overline{LC}(AC+C\overline{A}) \end{pmatrix}}{\det\begin{pmatrix} I & -LC \\ \overline{LC} & I \end{pmatrix}}.$$

Step 4: Let us define $C_0 = \epsilon\, C = V\Sigma\overline{V}^{-1}$. Then we have $C_0^{-1} = \overline{C}_0$. Moreover, by Proposition A.3.3 a) there exist $\overline{b}$, $d \in \mathbb{C}^n$, both non-zero, such that $C_0 = \Phi^{-1}_{A,\overline{A}}(\overline{b}\otimes d)$. In other words,

$$AC_0 + C_0\overline{A} \;=\; \overline{b}\otimes d, \tag{A.3}$$

and, by conjugation,

$$\overline{AC_0} + \overline{C}_0 A \;=\; b\otimes\overline{d}. \tag{A.4}$$

In particular, $\overline{C}_0 = \Phi^{-1}_{\overline{A},A}(b\otimes\overline{d})$.

Now, multiplying (A.3) with $\overline{C}_0$ from the left and (A.4) with C_0 from the right yields

$$\overline{b}\otimes(\overline{C}_0 d) = \overline{C}_0(\overline{b}\otimes d) = \overline{A} + \overline{C}_0 A C_0 = (b\otimes\overline{d})C_0 = (C_0' b)\otimes\overline{d}.$$

Since the range of $\overline{b}\otimes(\overline{C}_0 d)$ is $\langle\overline{C}_0 d\rangle$ and the range of $(C_0'b)\otimes\overline{d}$ is $\langle\overline{d}\rangle$, there exists $\overline{\lambda}\in\mathbb{C}$ such that $\overline{C}_0 d = \overline{\lambda}\overline{d}$, and, because the operators are the same, $C_0'b = \overline{\lambda}\overline{b}$. Next, from $d = C_0\overline{C}_0 d = \overline{\lambda}C_0\overline{d} = \overline{\lambda}\,\overline{\overline{C}_0 d} = |\lambda|^2 d$ we infer $|\lambda|^2 = 1$.

Step 5: As a last preparation, we use $\epsilon^2 = 1$, $|\lambda|^2 = 1$, and $C = \epsilon C_0$ to rewrite

$$\begin{aligned} q &= 1 - \frac{\det\left(\begin{pmatrix} I & 0 \\ 0 & \epsilon\overline{\lambda} \end{pmatrix}\begin{pmatrix} I & -LC \\ \overline{LC} & I+\overline{LC}(AC+C\overline{A}) \end{pmatrix}\begin{pmatrix} I & 0 \\ 0 & \epsilon\lambda \end{pmatrix}\right)}{\det\left(\begin{pmatrix} I & 0 \\ 0 & \epsilon\overline{\lambda} \end{pmatrix}\begin{pmatrix} I & -LC \\ \overline{LC} & I \end{pmatrix}\begin{pmatrix} I & 0 \\ 0 & \epsilon\lambda \end{pmatrix}\right)} \\ &= 1 - \frac{\det\begin{pmatrix} I & -\epsilon\lambda LC \\ \epsilon\overline{\lambda LC} & 1+\overline{LC}(AC+C\overline{A}) \end{pmatrix}}{\det\begin{pmatrix} I & -\epsilon\lambda LC \\ \epsilon\overline{\lambda LC} & I \end{pmatrix}} \\ &= 1 - \frac{\det\begin{pmatrix} I & -\lambda LC_0 \\ \overline{\lambda LC_0} & I+\overline{LC}_0(AC_0+C_0\overline{A}) \end{pmatrix}}{\det\begin{pmatrix} I & -\lambda LC_0 \\ \overline{\lambda LC_0} & I \end{pmatrix}}. \end{aligned} \tag{A.5}$$

Step 6: In the last step we show that there are vectors a, $c \in \mathbb{C}^n$ such that (A.5) coincides with a solution as given in Proposition 4.3.1. Namely, using the vectors b, $d \in \mathbb{C}^n$ constructed in the fourth step, we define

$$a = b, \quad c = \lambda F d,$$

where $F = \operatorname{diag}\{ \exp(\varphi_j^{(0)}) \mid j = 1, \dots, n\}$ is the diagonal matrix containing the initial position shifts.

It is straightforward to observe $L = \widehat{L}F$ with $\widehat{L}(x,t) = \exp\big(Ax + f_0(A)t\big)$, and thus

$$\lambda L C_0 = \widehat{L} \ \big(\ \lambda F \Phi_{A,\overline{A}}^{-1}(\overline{b} \otimes d) \ \big) = \widehat{L}\Phi_{A,\overline{A}}^{-1} \ \big(\ \overline{b} \otimes (\lambda F d) \ \big) = \widehat{L}\Phi_{A,\overline{A}}^{-1}(\overline{a} \otimes c).$$

It remains to show $\overline{LC}_0\big(AC_0 + C_0\overline{A}\big) = \overline{\widehat{L}}(\overline{a} \otimes \overline{c})$. To this end, we apply the results of the fourth step. Consequently,

$$\begin{aligned} \overline{LC}_0\big(AC_0 + C_0\overline{A}\big) &= \overline{LC}_0\big(\overline{b} \otimes d\big) \ = \ \overline{L}\big(\overline{b} \otimes (\overline{C}_0 d)\big) \\ &= \overline{L}\big(\overline{b} \otimes (\overline{\lambda d})\big) \ = \ \overline{\widehat{L}}\big(\overline{b} \otimes (\overline{\lambda F d})\big) \\ &= \overline{\widehat{L}}(\overline{a} \otimes \overline{c}). \end{aligned}$$

As a result we have proved that the solution (A.1) can be realized in the form of Proposition 4.3.1 by a particular choice of a, $c \in \mathbb{C}^n$, and the generating matrix A being diagonal. This is shows the assertion. □

Lemma A.1.2. *Let R, S, T, and $R_0 \in \mathcal{M}_{n,n}(\mathbb{C})$ be arbitrary square matrices. Then the following identity holds:*

$$\det \begin{pmatrix} I & R + R_0 \\ S & T \end{pmatrix} = \det \begin{pmatrix} I & R \\ S & T - SR_0 \end{pmatrix}.$$

Proof Since the block in the upper left corner is the identity matrix $I \in \mathcal{M}_{n,n}(\mathbb{C})$ and the dimensions of the blocks are equal, we can eliminate R_0 by ordinary operations with columns. It is elementary to check the effect of these operations on the block in the lower right corner. □

A.2 The case of a single negaton

Literature on positons and negatons of the AKNS-system is very limited. Discussions of low-dimensional cases can be found for example in [11], [79] for the modified Korteweg-de Vries and in [95] for the sine-Gordon equation. Their techniques rely on the Darboux transformation techniques initiated by Matveev and do not yield solution formulas in closed form.

Another approach is to explicitly solve the Gelfand-Levitan-Marchenko equations for scattering data with multiple poles in the reflection coefficient. (Single poles correspond to solitons). This ansatz has been pursued in [98] for the sine-Gordon and in [71] for the Nonlinear Schrödinger equation, where also asymptotics are discussed.

In the sequel we shall focus on unpublished work of Steudel [96], where the single negaton of order n of the AKNS-system is given in terms of Wronskians. After a slight adaption to the formula (A.1), the solution formula of Steudel reads

$$q = 2\mathrm{i}(-1)^n \Delta_1 / \Delta_2, \tag{A.6}$$

with

$$\Delta_1 = \det \begin{pmatrix} \left(\frac{\partial^{i-1}}{\partial\lambda^{i-1}} \left(\lambda^{n-j}\right) \right)_{\substack{i=1,\dots,n \\ j=0,\dots,n}} & \left(-\frac{\partial^{i-1}}{\partial\lambda^{i-1}} \left(\lambda^{n-j}\varphi(\lambda)\right) \right)_{\substack{i=1,\dots,n \\ j=2,\dots,n}} \\ \left(\frac{\partial^{i-1}}{\partial\overline{\lambda}^{i-1}} \left(\overline{\lambda}^{n-j}\right) \right)_{\substack{i=1,\dots,n \\ j=0,\dots,n}} & \left(\frac{\partial^{i-1}}{\partial\overline{\lambda}^{i-1}} \left(\overline{\lambda}^{n-j}/\varphi(\overline{\lambda})\right) \right)_{\substack{i=1,\dots,n \\ j=2,\dots,n}} \end{pmatrix}$$

and

$$\Delta_2 = \det \begin{pmatrix} \left(\frac{\partial^{i-1}}{\partial\lambda^{i-1}} \left(\lambda^{n-j}\right) \right)_{i,j=1}^{n} & \left(-\frac{\partial^{i-1}}{\partial\lambda^{i-1}} \left(\lambda^{n-j}\varphi(\lambda)\right) \right)_{i,j=1}^{n} \\ \left(\frac{\partial^{i-1}}{\partial\overline{\lambda}^{i-1}} \left(\overline{\lambda}^{n-j}\right) \right)_{i,j=1}^{n} & \left(\frac{\partial^{i-1}}{\partial\overline{\lambda}^{i-1}} \left(\overline{\lambda}^{n-j}/\varphi(\overline{\lambda})\right) \right)_{i,j=1}^{n} \end{pmatrix},$$

where $\varphi(\lambda) = \exp\big(2\mathrm{i}\lambda x + f_0(2\mathrm{i}\lambda)t + \varphi^{(0)}(2\mathrm{i}\lambda)\big)$ and $\varphi^{(0)}$ shares the property of f_0 for the $\mathbb{C}$-reduced AKNS system, i.e., $\overline{\varphi^{(0)}(t)} = -\varphi^{(0)}(-\overline{t})$. In particular,

$$\varphi(\overline{\lambda}) = 1/\overline{\varphi(\lambda)}.$$

Note the formal analogy to the determinants appearing in the Wronskian formula (A.1) for the n-solitons. In fact, the first and the n-th rows of Δ_1 in (A.1) and of the actual Δ_1 coincide exactly. Whereas in the previous case the matrix is filled by independent rows of the same type, in the case at hand we obtain the remaining rows by successive differentiation with respect to λ, $\overline{\lambda}$.

In the following lemma we translate (A.6) into our framework.

Proposition A.2.1. *Every single negaton as in (A.6) can be realized as one of the solutions constructed in Proposition 4.3.1 with A being to a single Jordan block.*

Proof We closely follow the proof of Proposition A.1.1. Therefore, we only indicate the necessary modifications in each step.

Step 1: Again we set $\alpha = 2\mathrm{i}\lambda$, $\overline{\alpha} = -2\mathrm{i}\overline{\lambda}$. Consequently, the derivatives are replaced by the rule $\partial/\partial\alpha = \frac{1}{2\mathrm{i}}\partial/\partial\lambda$, $\partial/\partial\overline{\alpha} = -\frac{1}{2\mathrm{i}}\partial/\partial\overline{\lambda}$. Carrying out this replacement carefully, we end up with the same formulas as in (A.6) with

- λ, $\overline{\lambda}$ replaced by α, $-\overline{\alpha}$,
- $\dfrac{\partial}{\partial\lambda}$, $\dfrac{\partial}{\partial\overline{\lambda}}$ replaced by $\dfrac{\partial}{\partial\alpha}$, $\dfrac{\partial}{\partial\overline{\alpha}}$,
- $\varphi(\lambda)$, $\varphi(\overline{\lambda})$ replaced by $\psi(\alpha)$, $1/\widehat{\psi}(\overline{\alpha})$ with $\psi(\alpha) := \exp\big(\alpha x + f_0(\alpha)t + \varphi^{(0)}(\alpha)\big)$, $\widehat{\psi}(\overline{\alpha}) := \exp\big(-\overline{\alpha}x + f_0(-\overline{\alpha})t + \varphi^{(0)}(-\overline{\alpha})\big)$ (note that the latter equals $1/\overline{\psi(\alpha)}$),

and the factor $2\mathrm{i}$ in front of the whole formula cancelled.

Step 2: By precisely the same argument as in the proof of Proposition A.1.1, we obtain the reformulation $q = 1 - \Delta/\Delta_2$ with

$$\Delta = \det \begin{pmatrix} \left(\frac{\partial^{i-1}}{\partial\alpha^{i-1}} \left(\alpha^{n-j}\right) \right)_{i,j=1}^{n} & \left(-\frac{\partial^{i-1}}{\partial\alpha^{i-1}} \Big(\alpha^{n-j}\big(\psi(\alpha)+\alpha\big)\Big) \right)_{i,j=1}^{n} \\ \left(\frac{\partial^{i-1}}{\partial\overline{\alpha}^{i-1}} \left((-\overline{\alpha})^{n-j}\right) \right)_{i,j=1}^{n} & \left(\frac{\partial^{i-1}}{\partial\overline{\alpha}^{i-1}} \Big((-\overline{\alpha})^{n-j}\big(1/\widehat{\psi}(\overline{\alpha})+\overline{\alpha}\big)\Big) \right)_{i,j=1}^{n} \end{pmatrix},$$

As for a more convenient representation of the inner matrices, define

$$V_0 = \Big(\frac{\partial^{i-1}}{\partial \alpha^{i-1}} \left(\alpha^{n-j}\right) \Big)_{i,j=1}^{n} \qquad \text{and} \qquad W_0 = \Big(\frac{\partial^{i-1}}{\partial \overline{\alpha}^{i-1}} \left((-\overline{\alpha})^{n-j}\right) \Big)_{i,j=1}^{n}.$$

In addition, we use the following factorization, valid for an arbitrary function f, which can be easily verified by the product rule,

$$\Big(\frac{\partial^{i-1}}{\partial \alpha^{i-1}} \left(\alpha^{n-j} f(\alpha)\right) \Big)_{i,j=1}^{n} = \Gamma \widehat{D}_f \Gamma^{-1} V_0 \qquad \text{for } \Gamma := \begin{pmatrix} 0 & & 0! \\ & \iddots & \\ (n-1)! & & 0 \end{pmatrix}$$

$$\text{and } \widehat{D}_f \text{ with the entries } (\widehat{D}_f)_{ij} := \begin{cases} \dfrac{1}{(j-i)!} \dfrac{\partial^{j-i}}{\partial \alpha^{j-i}} f(\alpha), & j \geq i, \\ 0, & j < i. \end{cases}$$

In particular, $\widehat{D}_{\text{id}} = A$, where A is the Jordan block corresponding to the eigenvalue α with dimension n. Moreover, we define $L := \widehat{D}_\psi$.

With these ingredients, we check

$$\begin{aligned} \Delta &= \det \begin{pmatrix} V_0 & -\Gamma(L+A)\Gamma^{-1} V_0 \\ W_0 & \Gamma(\overline{L}^{-1} + \overline{A})\Gamma^{-1} W_0 \end{pmatrix} = \det \begin{pmatrix} \Gamma^{-1} V_0 \Gamma & -(L+A)\, \Gamma^{-1} V_0 \Gamma \\ \Gamma^{-1} W_0 \Gamma & (\overline{L}^{-1} + \overline{A})\, \Gamma^{-1} W_0 \Gamma \end{pmatrix} \\ &= \det \begin{pmatrix} V & -(L+A) V \\ W & (\overline{L}^{-1} + \overline{A}) W \end{pmatrix}, \end{aligned}$$

where we have set $V = \Gamma^{-1} V_0 \Gamma$, $W = \Gamma^{-1} W_0 \Gamma$. Of course, Δ_2 can be rewritten analogously.

To sum up,

$$\begin{aligned} q &= 1 - \Delta/\Delta_2 \text{ with} \\ & \qquad \Delta_2 = \det \begin{pmatrix} V & -LV \\ W & \overline{L}^{-1} W \end{pmatrix} \qquad \text{and} \qquad \Delta = \det \begin{pmatrix} V & -(L+A)V \\ W & (\overline{L}^{-1} + \overline{A}) W \end{pmatrix}. \end{aligned}$$

Step 3: This step can be carried over literally. For $C := VW^{-1}$, the result is

$$q = 1 - \frac{\det \begin{pmatrix} I & -LC \\ \overline{LC} & I + \overline{LC}(AC + C\overline{A}) \end{pmatrix}}{\det \begin{pmatrix} I & -LC \\ \overline{LC} & I \end{pmatrix}}.$$

To perform the reformulation in this step we have used $WV^{-1} = \overline{C}$, which can be seen by the following argument. By definition of V_0, W_0, we have $W_0 = \epsilon \overline{V}_0 \Sigma$ with $\Sigma = \operatorname{diag}\{(-1)^{j-1} \mid j = 1, \ldots, n\}$ and $\epsilon = (-1)^{n-1}$. Therefore,

$$C = VW^{-1} = \epsilon\, \Gamma^{-1} V_0 \Sigma \overline{V}_0^{-1} \Gamma$$

and, analogously, $WV^{-1} = \epsilon\, \Gamma^{-1} \overline{V}_0 \Sigma V_0^{-1} \Gamma = \overline{C}$. In particular, $C^{-1} = \overline{C}$.

Step 4: Now we apply the factorization result for matrices X such that $AX + X\overline{A}$ is one-dimensional. Recall the definition of V_A,

$$V_A = \Big(\frac{1}{(n-i)!} \frac{\partial^{n-i}}{\partial \alpha^{n-i}} (\alpha^{j-1}) \Big)_{i,j=1}^{n},$$

and define $C_0 = V_A \Sigma V_{\overline{A}}^{-1}$. Then, by Proposition A.3.6 a), there exist b, $d \in \mathbb{C}^n$, both nonzero, such that $C_0 = \Phi_{A,\overline{A}}^{-1}(\overline{b} \otimes d)$, and, by conjugation, $\overline{C}_0 = \Phi_{\overline{A},A}^{-1}(b \otimes \overline{d})$. Moreover, the same argument as in the proof of Proposition A.1.1 shows that there is a complex number λ, $|\lambda| = 1$, with $C_0 \overline{d} = \lambda d$ and $C_0' b = \overline{\lambda b}$.

To finish the step, we establish the relation between C_0 and C. Observe

$$V_A = \Gamma^{-1} V_0 J \qquad \text{with } J = \begin{pmatrix} 0 & & 1 \\ & \cdot^{\cdot^{\cdot}} & \\ 1 & & 0 \end{pmatrix}.$$

Thus $C = \epsilon\,(\Gamma^{-1} V_0) \Sigma (\overline{V}_0^{-1} \Gamma) = V_A (\epsilon\, J \Sigma J) V_{\overline{A}}^{-1} = V_A \Sigma V_{\overline{A}}^{-1} = C_0$.

Step 5: Using $|\lambda| = 1$, $C = C_0$, and following precisely the argument in the proof of Proposition A.1.1, we obtain

$$q = 1 - \frac{\det \begin{pmatrix} I & -\lambda L C_0 \\ \overline{\lambda L C_0} & I + \overline{L C}_0 (A C_0 + C_0 \overline{A}) \end{pmatrix}}{\det \begin{pmatrix} I & -\lambda L C_0 \\ \overline{\lambda L C_0} & I \end{pmatrix}}. \tag{A.7}$$

Step 6: Finally we construct vectors a, $c \in \mathbb{C}^n$ such that (A.7) coincides with a solution as given in Proposition 4.3.1. Starting with the vectors b, $d \in \mathbb{C}^n$ of the fourth step, we set

$$a = b, \quad c = \lambda F d.$$

where the matrix $F = \widehat{D}_{\exp(\varphi(0))}$ contains the initial position shifts and their derivatives. Then $L = \widehat{L} F$ with $\widehat{L}(x,t) = \exp\big(Ax + f_0(A)t\big)$, which is simply the product rule, and $\lambda F \Phi_{A,\overline{A}}^{-1}(\overline{b} \otimes d) = \Phi_{A,\overline{A}}^{-1}\big(\overline{b} \otimes (\lambda F d)\big)$ because A and F commute. This immediately shows that $\lambda L C_0 = \widehat{L}\, \lambda F \Phi_{A,\overline{A}}^{-1}(\overline{b} \otimes d) = \widehat{L} \Phi_{A,\overline{A}}^{-1}(\overline{a} \otimes c)$, $\overline{\lambda L C_0} = \overline{\widehat{L}} \Phi_{\overline{A},A}^{-1}(a \otimes \overline{c})$. The identity $\overline{L C}_0 \big(A C_0 + C_0 \overline{A}\big) = \overline{\widehat{L}}(\overline{a} \otimes \overline{c})$ can be verified as in the proof of Proposition A.1.1. Therefore, (A.6) can be realized in the form of Proposition 4.3.1 by a particular choice of a, $c \in \mathbb{C}^n$, and the generating matrix A being a Jordan block. The proof is complete. □

A.3 Factorization of solutions X of the operator equation $AX + XB$

In the preceding sections we have established the link to solutions of the $\mathbb{C}$-reduced AKNS system which are constructed using Wronskian techniques. The aim of this section is to supplement the factorization results Proposition A.3.3 and Proposition A.3.6 needed in the proofs of Proposition A.1.1 and Proposition A.2.1.

A.3.1 Vandermonde-type factorization for A, B diagonal

To the diagonal matrix $A = \operatorname{diag}\{\alpha_j \,|\, j = 1, \dots, n\} \in \mathcal{M}_{n,n}(\mathbb{C})$, we assign the Vandermonde matrix

$$V_A = \Big(\alpha_i^{n-j} \Big)_{i,j=1}^{n} = \begin{pmatrix} \alpha_1^{n-1} & \dots & \alpha_1 & 1 \\ \vdots & & \vdots & \vdots \\ \alpha_n^{n-1} & \dots & \alpha_n & 1 \end{pmatrix} \in \mathcal{M}_{n,n}(\mathbb{C})$$

and

$$W_A = \Big(\sum_{\substack{\lambda_1 < \ldots < \lambda_{i-1} \\ j \notin \{\lambda_1, \ldots, \lambda_{i-1}\}}} \alpha_{\lambda_1} \ldots \alpha_{\lambda_{i-1}} \Big)_{i,j=1}^{n} = \begin{pmatrix} 1 & \ldots & 1 \\ \sum_{j \neq 1} \alpha_j & \ldots & \sum_{j \neq n} \alpha_j \\ \sum_{\substack{j<j' \\ j,j' \neq 1}} \alpha_j \alpha_{j'} & \ldots & \sum_{\substack{j<j' \\ j,j' \neq n}} \alpha_j \alpha_{j'} \\ \vdots & & \vdots \\ \prod_{j \neq 1} \alpha_j & \ldots & \prod_{j \neq n} \alpha_j \end{pmatrix}.$$

As the next lemma shows, the latter matrix is the main part of V_A^{-1}.

Lemma A.3.1. *The inverse of the Vandermonde matrix $V_A = \big(\alpha_i^{n-j}\big)_{i,j=1}^n \in \mathcal{M}_{n,n}(\mathbb{C})$, with α_j pairwise different, is*

$$V_A^{-1} = \Sigma W_A D_A \qquad \textit{with } D_A = \operatorname{diag}\ \Big\{ 1 \Big/ \prod_{\substack{\lambda=1 \\ \lambda \neq j}}^{n} (\alpha_j - \alpha_\lambda) \ \Big| \ j = 1, \ldots, n \ \Big\}$$
$$\textit{and } \Sigma = \operatorname{diag}\ \big\{ \ (-1)^{j-1} \ \big| \ j = 1, \ldots, n \ \big\} \ .$$

Proof The proof can be done by direct verification. Namely, denoting the ij-th entry of a matrix T as usual by T_{ij}, matrix calculation shows

$$\begin{aligned} \big(\ V_A \Sigma W_A \ \big)_{ij} &= \alpha_i^{n-1} - \alpha_i^{n-2} \sum_{\lambda \neq j} \alpha_\lambda + \alpha_i^{n-3} \sum_{\lambda, \lambda' \neq j} \alpha_\lambda \alpha_{\lambda'} + \ldots + (-1)^{n-1} \prod_{\lambda \neq j} \alpha_\lambda \\ &= \prod_{\lambda \neq j} (\alpha_i - \alpha_\lambda) \\ &= \begin{cases} \prod_{\lambda \neq j} (\alpha_j - \alpha_\lambda), & i = j \\ 0 & i \neq j \end{cases} . \end{aligned}$$

Therefore, $V_A \Sigma W_A = D_A^{-1}$, which completes the proof. □

Corollary A.3.2. *Let $A = \operatorname{diag}\{\alpha_j | j = 1, \ldots, n\}$, $B = \operatorname{diag}\{\beta_j | j = 1, \ldots, n\} \in \mathcal{M}_{n,n}(\mathbb{C})$. Then*

a) $V_A W_B = \Big(\prod_{\substack{\lambda=1 \\ \lambda \neq j}}^{n} (\alpha_i + \beta_\lambda) \Big)_{i,j=1}^n$.

b) $A(V_A W_B) + (V_A W_B) B = \Big(\prod_{\lambda=1}^{n} (\alpha_i + \beta_\lambda) \Big)_{i,j=1}^n = e_0 \otimes c$, *where $e_0 \in \mathbb{C}^n$ is the vector with all entries equal to 1 and* $c = \big(\prod_{\lambda=1}^n (\alpha_i + \beta_\lambda) \big)_{i=1}^n$.

Proof As for a), the proof can be almost literally copied from the proof of Lemma A.3.1, and b) can be immediately checked by matrix multiplication. □

In summary, we have shown the following factorization result.

Proposition A.3.3. *Let $A = \operatorname{diag}\{\alpha_j \mid j = 1, \ldots, n\}$, $B = \operatorname{diag}\{\beta_j \mid j = 1, \ldots, n\} \in \mathcal{M}_{n,n}(\mathbb{C})$ satisfy (i) β_j pairwise different and (ii) $\alpha_i + \beta_j \neq 0$ for all i, j. Then the following factorization holds:*

a) There exist vectors $a, c \in \mathbb{C}^n$ such that

$$V_A \Sigma V_B^{-1} \ = \ \Phi_{A,B}^{-1}(a \otimes c),$$

namely $a = \big(\prod_{\lambda=1, \lambda \neq i}^n (\beta_j - \beta_\lambda)^{-1} \big)_{i=1}^n$, $c = \big(\prod_{\lambda=1}^n (\alpha_i + \beta_\lambda) \big)_{i=1}^n$.

b) For any $b, d \in \mathbb{C}^n$,

$$\Phi^{-1}_{A,B}(b \otimes d) = D_c D_d^{-1}\, V_A \Sigma V_B^{-1}\, D_a D_b^{-1},$$

where $D_f \in \mathcal{M}_{n,n}(\mathbb{C})$, for $f = (f_j)_{j=1}^n \in \mathbb{C}^n$, denotes the diagonal matrix with the entries f_j on the diagonal, and $a, c \in \mathbb{C}^n$ are as defined in a).

Proof Part a) is just a reformulation of Lemma A.3.1, Corollary A.3.2 b). For part b) we use that the matrices $D_1 = D_d D_c^{-1}$, $D_2 = D_b D_a^{-1}$ are diagonal. Hence, from the fact that $C = V_A \Sigma V_B^{-1}$ solves $AC + CB = a \otimes c$ with a, c as defined in a), it follows that $\widehat{C} = D_1 C D_2$ solves $A\widehat{C} + \widehat{C}B = D_1(AC + CB)D_2 = D_1(a \otimes c)D_2 = (D_2' a) \otimes (D_1 c) = b \otimes d$. $\square$

A.3.2 Factorization for Jordan blocks A, B

To the Jordan block $A \in \mathcal{M}_{n,n}(\mathbb{C})$ with eigenvalue α, we assign

$$V_A = \Big(\frac{1}{(n-i)!} \frac{\partial^{n-i}}{\partial \alpha^{n-i}} \alpha^{j-1} \Big)_{i,j=1}^n \mathcal{M}_{n,n}(\mathbb{C}),$$

$$\text{in other words } (V_A)_{ij} = \begin{cases} \binom{j-1}{n-i} \alpha^{(j-1)-(n-i)} & n \le i+j-1, \\ 0 & n > i+j-1. \end{cases}$$

for the ij-th entry $(V_A)_{ij}$ of V_A.

By the subsequent lemma, the main part of V_A^{-1} is an upper left triangle matrix of a quite similar structure, namely

$$W_A = \Big((-1)^{n-j} \frac{1}{(i-1)!} \frac{\partial^{i-1}}{\partial \alpha^{i-1}} \alpha^{n-j} \Big)_{i,j=1}^n .$$

Lemma A.3.4. *The inverse of the matrix $V_A \in \mathcal{M}_{n,n}(\mathbb{C})$ is*

$$V_A^{-1} = \Sigma W_A \qquad \text{with } \Sigma = \operatorname{diag}\big\{ (-1)^{j-1} \,\big|\, j = 1, \dots, n \big\} .$$

Proof Since $(V_A)_{i\kappa} = 0$ if $\kappa < n+1-i$, and $(W_A)_{\kappa j} = 0$ if $\kappa > n+1-j$, $\big(V_A \Sigma \cdot W_A \big)_{ij} = 0$ for $i < j$. For $i \ge j$, we directly calculate

$$\begin{aligned}
\big(V_A \Sigma \cdot W_A \big)_{ij} &= \sum_{\kappa=n+1-i}^{n+1-j} (V_A)_{i\kappa} (-1)^{\kappa-1} (W_A)_{\kappa j} \\
&= \sum_{\mu=0}^{i-j} (V_A)_{i(\mu+n+1-i)} \cdot (-1)^{\mu+n-i} \cdot (W_A)_{(\mu+n+1-i)j} \\
&= \sum_{\mu=0}^{i-j} \binom{n-i+\mu}{n-i} \alpha^{\mu} \cdot (-1)^{n-i+\mu} \cdot (-1)^{n-j} \binom{n-j}{n-i+\mu} \alpha^{i-j-\mu} \\
&= (-\alpha)^{i-j} \binom{n-j}{i-j} \sum_{\mu=0}^{i-j} (-1)^{\mu} \binom{i-j}{\mu} \\
&= \delta_{ij}.
\end{aligned}$$

Thus $V_A \Sigma W_A = 1$, yielding the assertion. $\square$

Corollary A.3.5. *Let $A, B \in \mathcal{M}_{n,n}(\mathbb{C})$ be Jordan blocks with eigenvalues α, β, respectively. Then*

a) $V_A W_B = \left((-1)^{n-j} \frac{1}{(n-i)!} \frac{\partial^{n-i}}{\partial(\alpha+\beta)^{n-i}} (\alpha+\beta)^{n-j} \right)_{i,j=1}^{n}.$

b) $A(V_A W_B) + (V_A W_B)B = e_n^{(1)} \otimes c$*, where* $e_n^{(1)} \in \mathbb{C}^n$ *is the first standard basis vector and*

$$c = \left((-1)^{n-1} \frac{1}{(n-i)!} \frac{\partial^{n-i}}{\partial(\alpha+\beta)^{n-i}} (\alpha+\beta)^{n} \right)_{i=1}^{n} \in \mathbb{C}^n.$$

Proof The proof of part a) is the same as the one of Lemma A.3.4. As for b), we start with the calculation of $A(V_A W_B)$. Matrix multiplication yields that, for $i < n$,

$$\begin{aligned} \left(A(V_A W_B) \right)_{ij} &= \alpha\, (-1)^{n-j} \frac{1}{(n-i)!} \frac{\partial^{n-i}}{\partial(\alpha+\beta)^{n-i}} (\alpha+\beta)^{n-j} \\ &\quad + (-1)^{n-j} \frac{1}{(n-i-1)!} \frac{\partial^{n-i-1}}{\partial(\alpha+\beta)^{n-i-1}} (\alpha+\beta)^{n-j}. \end{aligned}$$

It is clear that the first summand does not vanish if and only if $i - j \geq 0$, the second if and only if $i - j + 1 \geq 0$. Thus the following cases are relevant:

a) $i - j < -1$. Then $\left(A(V_A W_B) \right)_{ij} = 0$ since both summands are zero.

b) $i - j = -1$. Here only the second summand is non-zero, and $\left(A(V_A W_B) \right)_{ij} = (-1)^{n-j}$.

$\mathbf{c_1}$) $i - j \geq 0$. Now the both summands are non-trivial and we calculate

$$\begin{aligned} &\left(A(V_A W_B) \right)_{ij} \\ &= \alpha\, (-1)^{n-j} \binom{n-j}{n-i} (\alpha+\beta)^{i-j} + (-1)^{n-j} \binom{n-j}{n-i-1} (\alpha+\beta)^{i-j+1} \\ &= (-1)^{n-j} \left[\binom{n-j+1}{n-i} \alpha + \binom{n-j}{n-i-1} \beta \right] (\alpha+\beta)^{i-j}. \end{aligned}$$

$\mathbf{d_1}$) In the case $i = n$, which has to be treated separately, we get $\left(A(V_A W_B) \right)_{nj} = \alpha\, (-1)^{n-j} (\alpha+\beta)^{n-j}$.

Next we calculate $(V_A W_B)B$. In an analogous manner as above, for $j > 1$,

$$\begin{aligned} \left((V_A W_B)B \right)_{ij} &= \beta\, (-1)^{n-j} \frac{1}{(n-i)!} \frac{\partial^{n-i}}{\partial(\alpha+\beta)^{n-i}} (\alpha+\beta)^{n-j} \\ &\quad + (-1)^{n-j+1} \frac{1}{(n-i)!} \frac{\partial^{n-i}}{\partial(\alpha+\beta)^{n-i}} (\alpha+\beta)^{n-j+1}. \end{aligned}$$

Again the first summand does not vanish if and only if $i - j \geq 0$, the second if and only if $i - j + 1 \geq 0$, and we have to distinguish the cases:

a) $i - j < -1$. Again $\left((V_A W_B)B \right)_{ij} = 0$ because both summands vanish.

b) $i - j = -1$. Here only the second summand is non-zero, yielding $\left((V_A W_B)B \right)_{ij} = (-1)^{n-j+1}$.

$\mathbf{c_2}$) $i - j \geq 0$. Then both summands are non-zero, and we check

$$\begin{aligned} &\left((V_A W_B)B \right)_{ij} \\ &= \beta\, (-1)^{n-j} \binom{n-j}{n-i} (\alpha+\beta)^{i-j} + (-1)^{n-j+1} \binom{n-j+1}{n-i} (\alpha+\beta)^{i-j+1} \end{aligned}$$

$$= \begin{cases} (-1)^{n-j}\left[\binom{n-j}{n-i-1}\beta - \binom{n-j+1}{n-i}\alpha\right](\alpha+\beta)^{i-j}, & i < n, \\ (-1)^{n-j+1}\alpha(\alpha+\beta)^{n-j}, & i = n. \end{cases}$$

d₂) The case $j = 1$ has to be treated separately. Here we get

$$\left(\,(V_A W_B)B\,\right)_{i1} \;=\; \beta(-1)^{n-1}\binom{n-1}{n-i}(\alpha+\beta)^{i-1}.$$

It is obvious[1] that in the cases **a)** and **b)** the entries $\left(\,A(V_A W_B)+(V_A W_B)B\,\right)_{ij}$ always vanish. Therefore, $A(V_A W_B) + (V_A W_B)B$ is a lower left triangular matrix. To check the remaining entries $i - j \leq 0$, we proceed along the different cases discussed above.

$\mathbf{c_1}) \wedge \mathbf{c_2})$ $i < n$, $j > 1$. Then we obviously have $\left(\,A(V_A W_B) + (V_A W_B)B\,\right)_{ij} = 0$.

$\mathbf{d_1}) \wedge \mathbf{c_2})$ $i = n$, $j > 1$. Also in this case we directly see $\left(\,A(V_A W_B) + (V_A W_B)B\,\right)_{nj} = 0$.

$\mathbf{c_1}) \wedge \mathbf{d_2})$ $i < n$, $j = 1$. Then

$$\begin{aligned} &\left(\,A(V_A W_B) + (V_A W_B)B\,\right)_{i1} \\ &= (-1)^{n-1}\left[\binom{n}{n-i}\alpha + \binom{n-1}{n-i-1}\beta\right](\alpha+\beta)^{i-1} \\ &\qquad\qquad + \beta(-1)^{n-1}\binom{n-1}{n-i}(\alpha+\beta)^{i-1} \\ &= (-1)^{n-1}\binom{n}{n-i}(\alpha+\beta)^{i}. \end{aligned}$$

$\mathbf{d_1}) \wedge \mathbf{d_2})$ $i = n$, $j = 1$. Then

$$\begin{aligned} &\left(\,A(V_A W_B) + (V_A W_B)B\,\right)_{n1} \\ &= \alpha(-1)^{n-1}(\alpha+\beta)^{n-1} + \beta(-1)^{n-1}\binom{n-1}{0}(\alpha+\beta)^{n-1} \\ &= (-1)^{n-1}(\alpha+\beta)^{n}. \end{aligned}$$

In summary, we have shown that the matrix $A(V_A W_B) + (V_A W_B)B$ has only non-vanishing entries in the first column, namely

$$\begin{aligned} \left(\,A(V_A W_B) + (V_A W_B)B\,\right)_{i1} &= (-1)^{n-1}\binom{n}{n-i}(\alpha+\beta)^{i} \\ &= (-1)^{n-1}\frac{1}{(n-i)!}\frac{\partial^{n-i}}{\partial(\alpha+\beta)^{n-i}}(\alpha+\beta)^{n}. \end{aligned}$$

which is equivalent to the assertion. □

[1] Illustration of the cases which had to be treated separately in the proof of Lemma A.3.5. The first matrix is a scheme for the calculation of $A\,(\,V_A W_B\,)$, the second for $(\,V_A W_B\,)\,B$.

$$\begin{pmatrix} & \times & & a) & \\ & & \ddots & & \\ & c_1) & & \times & \\ & & & & \times \\ \bullet & \bullet & \ldots & \bullet & \bullet \end{pmatrix} \qquad \begin{pmatrix} \circ & \times & & a) & \\ \vdots & & \ddots & & \\ \circ & & & \times & \\ \circ & c_2) & & & \times \\ \circ & & & & \end{pmatrix} \qquad \times : b) \qquad \begin{matrix} \bullet : d_1) \\ \circ : d_2) \end{matrix}$$

To state the factorization result, we recall the notation $\Gamma_r(\xi) \in \mathcal{M}_{n,n}(\mathbb{C})$ for the upper right band matrix given in terms of a vector $\xi = (\xi_j)_{j=1}^n \in \mathbb{C}^n$ by

$$\Gamma_r(\xi) = \begin{pmatrix} \xi_1 & & \xi_n \\ & \ddots & \\ 0 & & \xi_1 \end{pmatrix}$$

Note that $\Gamma_r(\xi)$ is invertible if and only if $\xi_1 \neq 0$. In this case also $\Gamma(\xi)^{-1}$ is an upper right band matrix. Furthermore, for $\xi \in \mathbb{C}^n$ we set $\widehat{\xi} = (\xi_{n-i+1})_{i=1}^n$. Then we have $\Gamma_r(\widehat{\xi})e_n^{(n)} = \xi$ and $\Gamma_r(\xi)'e_n^{(1)} = \xi$, where $e_n^{(1)}$, $e_n^{(n)}$ are the first and the n-th standard basis vectors.

Proposition A.3.6. *For single Jordan blocks $A, B \in \mathcal{M}_{n,n}(\mathbb{C})$ with eigenvalues α, β, respectively, which satisfy $\alpha + \beta \neq 0$, the following factorization holds:*

a) There exist vectors $a, c \in \mathbb{C}^n$ such that

$$V_A \Sigma V_B^{-1} \ = \ \Phi_{A,B}^{-1}(a \otimes c),$$

namely $a = e_n^{(1)}$, $c = \left((-1)^{n-1} \frac{1}{(n-i)!} \frac{\partial^{n-i}}{\partial(\alpha+\beta)^{n-i}} (\alpha+\beta)^n \right)_{i=1}^n.$

b) For any $b, d \in \mathbb{C}^n$,

$$\Phi_{A,B}^{-1}(b \otimes d) = \Gamma_r(\widehat{d})\Gamma_r(\widehat{c})^{-1} \cdot V_A \Sigma V_B^{-1} \cdot \Gamma_r(b),$$

where c is defined as in a).

Proof Part a) is just a reformulation of Lemma A.3.4, Corollary A.3.5 b). As for part b), we use that upper band matrices commute. By a) we infer that $C = V_A \Sigma V_B^{-1}$ solves $AC + CB = e_n^{(1)} \otimes c = e_n^{(1)} \otimes \left(\Gamma_r(\widehat{c})e_n^{(n)}\right) = \Gamma_r(\widehat{c})(e_n^{(1)} \otimes e_n^{(n)})$. Thus $\widehat{C} = \Gamma_r(\widehat{d})\Gamma_r(\widehat{c})^{-1}C\Gamma_r(b)$ solves $A\widehat{C} + \widehat{C}B = \Gamma_r(\widehat{d})(e_n^{(1)} \otimes e_n^{(n)})\Gamma_r(b) = (\Gamma_r(b)'e_n^{(1)}) \otimes (\Gamma_r(\widehat{d})e_n^{(n)}) = b \otimes d$. Note that the invertibility of $\Gamma_r(\widehat{c})$ is guaranteed by the assumption that $\alpha + \beta \neq 0$. $\square$

Appendix B

A brief survey on traces and determinants on quasi-Banach operator ideals

In this appendix we give a concise and selfcontained introduction to the theory of traces and determinants on quasi-Banach operator ideals including all the material needed in this text. We shall also encompass several recent results. As a general reference, we cite the monographs of König [52], Pietsch [73], Reed/Simon [80], and Simon [94].

On the following aspect we lay particular emphasis.

> Basing on the well-known fact that there exists a unique trace on the smallest operator ideal $\mathcal{F}$ of finite rank operators which is, indeed, spectral, two natural procedures are considered how to extend the trace to (larger) quasi-Banach operator ideals.
>
> In this context the following operator ideals, which, from the modern point of view, play a crucial role for developing a trace theory, come up, namely the ideal $\mathcal{N}_r$ of r-nuclear operators ($0 < r \leq 1$) in the sense of Grothendieck on the one hand and ideals $\mathcal{A} \subset \mathcal{S}_1^{eig}$ of so-called eigenvalue type 1 on the other hand.
>
> Both of those possibilities to extend the trace are discussed in detail leading to some background information as well as a motivation for the construction of countable superpositions of solitons.

The outline presented here is essentially taken from [18]. If not stated otherwise, the cited results can be found in [73].

B.1 Basic notions

For the general notion of a quasi-Banach operator ideal we refer to the monographs of Defant/Floret [23] and Pietsch [72].

Definition B.1.1. (Operator ideal) *The class* $\mathcal{A} := \bigcup_{E,F} \mathcal{A}(E,F)$ *with given subsets* $\mathcal{A}(E,F) \subset \mathcal{L}(E,F)$ *for each pair of Banach spaces* E *and* F *is called an* operator ideal *if the following conditions are satisfied:*

(i) $a \otimes y \in \mathcal{A}(E,F)$ *for any* $a \in E'$ *and* $y \in F$,

(ii) *If* $S, T \in \mathcal{A}(E,F)$, *then* $S + T \in \mathcal{A}(E,F)$,

(iii) *If* $X \in \mathcal{L}(E_0,E)$, $T \in \mathcal{A}(E,F)$, *and* $Y \in \mathcal{L}(F,F_0)$, *then* $YTX \in \mathcal{A}(E_0,F_0)$.

Remark B.1.2. *An operator $T \in \mathcal{L}(E,F)$ is a finite rank operator if there exists a finite representation $T = \sum_{i=1}^{n} a_i \otimes y_i$ with $a_i \in E'$, $y_i \in F$, and the collection of all finite rank operators from E into F is denoted by $\mathcal{F}(E,F)$.*

The class $\mathcal{L}$ of all (bounded linear) operators is the largest and the class $\mathcal{F}$ of finite rank operators the smallest operator ideal.

Definition B.1.3. *a) A function $\| \cdot \mid \mathcal{A} \|$ assigning a non-negative number $\| T \mid \mathcal{A} \|$ to every operator $T \in \mathcal{A}$ is called a* quasi-norm *on the operator ideal $\mathcal{A}$ if it has the following properties:*

(i) $\quad \| a \otimes y \mid \mathcal{A} \| = \| a \| \cdot \| y \|$ *for* $a \in E'$, $y \in F$,

(ii) $\quad \| S + T \mid \mathcal{A} \| \leq c_{\mathcal{A}} \Big(\| S \mid \mathcal{A} \| + \| T \mid \mathcal{A} \| \Big)$ *for* $S, T \in \mathcal{A}(E,F)$,

(iii) $\quad \| YTX \mid \mathcal{A} \| \leq \| Y \| \cdot \| T \mid \mathcal{A} \| \cdot \| X \|$ *for* $X \in \mathcal{L}(E_0, E)$, $T \in \mathcal{A}(E,F)$, *and* $Y \in \mathcal{L}(F, F_0)$.

In the case $c_{\mathcal{A}} = 1$ we simply speak of a norm.

b) An operator ideal $\mathcal{A}$ is called a quasi-Banach operator ideal *if all components $\mathcal{A}(E,F)$ are complete with respect to the quasi-norm $\| \cdot \mid \mathcal{A} \|$ given on $\mathcal{A}$. If $\| \cdot \mid \mathcal{A} \|$ is an r-norm $(0 < r < 1)$, $\mathcal{A}$ is called an* r-Banach operator ideal*, and if $\| \cdot \mid \mathcal{A} \|$ is even a norm, then we call $\mathcal{A}$ a Banach operator ideal.*

Remark B.1.4. $\| T \| \leq \| T \mid \mathcal{A} \|$ *for all* $T \in \mathcal{A}$.

The concept of an operator ideal $\mathcal{A} = \bigcup_{E,F} \mathcal{A}(E,F)$ also makes sense if E, F range only over a subclass of Banach spaces. Of particular interest is the subclass of Hilbert spaces.

Definition B.1.5. *Let $\mathcal{A}$, $\mathcal{B}$ be quasi-Banach operator ideals.*

a) The product *$\mathcal{B} \circ \mathcal{A}$ consisting of operators $T \in \mathcal{L}(E,F)$ which can be written in the form $T = YX$ with $X \in \mathcal{A}(E,G)$, $Y \in \mathcal{B}(G,F)$ becomes a quasi-Banach operator ideal with respect to $\| T \mid \mathcal{B} \circ \mathcal{A} \| = \inf\{\| Y \mid \mathcal{B} \| \| X \mid \mathcal{A} \|\}$, where the infimum is taken over all possible factorizations.*

b) The sum *$\mathcal{A} + \mathcal{B}$ consisting of operators $T \in \mathcal{L}(E,F)$ which can be written in the form $T = X + Y$ with $X \in \mathcal{A}(E,F)$, $Y \in \mathcal{B}(E,F)$ becomes a quasi-Banach operator ideal with respect to $\| T \mid \mathcal{A} + \mathcal{B} \| = \inf\{\| X \mid \mathcal{A} \| + \| Y \mid \mathcal{B} \|\}$, where the infimum is taken over all possible decompositions.*

The following axiomatic approach to abstract traces on arbitrary operator ideals is due to Pietsch [73].

Definition B.1.6. (Trace)

a) Let $\mathcal{A}$ be an operator ideal. A complex valued function $\tau : \bigcup_E \mathcal{A}(E) \longrightarrow \mathbb{C}$ is called a trace *on $\mathcal{A}$ if the following properties are satisfied:*

(i) $\quad \tau$ *is linear on each component* $\mathcal{A}(E)$,

(ii) $\quad \tau(a \otimes y) = \langle y, a \rangle$ *for all* $a \in E'$, $y \in E$,

(iii) $\quad \tau(XT) = \tau(TX)$ *for any* $T \in \mathcal{A}(E,F)$, $X \in \mathcal{L}(F,E)$.

b) A trace τ defined on a quasi-Banach operator ideal $\mathcal{A}$ is said to be continuous *if the function $T \mapsto \tau(T)$ is continuous on every component $\mathcal{A}(E)$.*

c) A trace τ is called spectral *if it is given as the sum of the eigenvalues, i.e., $\tau(T) = \sum_i \lambda_i(T)$ for all T. Obviously an operator ideal admits at most one spectral trace, which we will always denote by* tr_λ.

Lemma B.1.7. *A trace τ defined on a quasi-Banach operator ideal $\mathcal{A}$ is continuous if and only if there exists a universal constant $c \geq 1$ such that $\mid \tau(T) \mid \leq c \parallel T \mid \mathcal{A} \parallel$ for all $T \in \mathcal{A}(E)$ (and all Banach spaces E).*

In analogy to the concept of traces, determinants on operator ideals can be defined.

Definition B.1.8. (Determinant)
a) Let $\mathcal{A}$ be an operator ideal. A complex valued function $\delta : \bigcup_E \mathcal{A}(E) \longrightarrow \mathbb{C}$ is called a determinant *on $\mathcal{A}$ if the following properties are satisfied:*

(i) $\delta((I+S)(I+T)) = \delta(I+S)\,\delta(I+T)$ *for all* $S, T \in \mathcal{A}(E)$,

(ii) $\delta(I + a \otimes y) = 1 + \langle y, a \rangle$ *for all* $a \in E'$, $y \in E$,

(iii) $\delta(I_E + XT) = \delta(I_F + TX)$ *for any* $T \in \mathcal{A}(E,F)$, $X \in \mathcal{L}(F,E)$,

(iv) For every $T \in \mathcal{A}(E)$ (and every arbitrary Banach space E), $\delta(I + zT)$ is an entire function in z,

where I (or I_E to emphasize the underlying Banach space E) always denotes the identity operator on E.

b) A determinant δ defined on a quasi-Banach operator ideal $\mathcal{A}$ is said to be continuous *if the function $T \mapsto \delta(I+T)$ is continuous on every component $\mathcal{A}(E)$.*

c) A determinant δ is called spectral *if it is given by $\delta(I+T) = \prod_i \big(1 + \lambda_i(T)\big)$ for all T. Obviously an operator ideal admits at most one spectral determinant, which we will always denote by* $\det_\lambda$.

Lemma B.1.9. *The operator $I + T$ is invertible if and only if $\delta(I+T) \neq 0$.*

Lemma B.1.10. *Let $\mathcal{A}$ be an operator ideal, τ a trace, and δ a determinant on $\mathcal{A}$. Then, for all $T \in \mathcal{A}$ with* $\operatorname{rank}(T) = 1$,

$$\delta(1+T) = 1 + \tau(T).$$

Remark B.1.11. *The concept of a trace/determinant is also used for operator ideals defined on a subclass of Banach spaces.*

B.2 Outline of results

The first aim of this section is to explain how, basing on the trace formula for finite rank operators, the trace can be extended to (larger) quasi-Banach operator ideals. Afterwards, the link between traces and determinants is discussed.

As the starting point, we consider the operator ideal $\mathcal{F}$ of finite rank operators. It is a well-known fact that there is a unique trace tr on $\mathcal{F}$ which is given by

$$(1) \qquad \operatorname{tr}(T) = \sum_{i=1}^{n} \langle y_i, a_i \rangle \ \text{ for an arbitrary (finite) representation } \ T = \sum_{i=1}^{n} a_i \otimes y_i,$$

and moreover, since the trace tr on the finite rank operators $\mathcal{F}$ is spectral,

$$(2) \qquad \operatorname{tr}(T) = \sum_{i=1}^{N} \lambda_i(T) \qquad (\lambda_i(T) \text{ the eigenvalues of } T).$$

Each of those two possibilities to express the trace tr on the finite rank operators $\mathcal{F}$ leads in a natural manner to an ansatz for the extension of the trace to (larger) quasi-Banach operator ideals, namely, for operators

(1) T with an infinite representation $T = \sum_{i=1}^{\infty} a_i \otimes y_i$

where $\sum_{i=1}^{\infty} \mid \langle y_i, a_i \rangle \mid \leq \sum_{i=1}^{\infty} \parallel a_i \parallel \parallel y_i \parallel < \infty$ is assumend,

(2) T possessing absolutely summing eigenvalues, i.e. $\sum_i \mid \lambda_i(T) \mid < \infty$.

Unfortunately, neither of them goes through automatically and we therefore have to take a closer look to both of them.

First we turn to the extension-procedure indicated by (2).

Let $T \in \mathcal{L}(E)$ be a Riesz operator. Because all non-zero elements $\lambda \neq 0$ of the spectrum $\mathrm{spec}(T)$ of a Riesz operator are isolated eigenvalues with finite algebraic multiplicity, we may assign an eigenvalue sequence $(\lambda_i(T))_i$ to every Riesz operator $T \in \mathcal{L}(E)$ by the following rule:

Counting every eigenvalue according to its algebraic multiplicity, the eigenvalues are arranged in order of non-increasing absolute values. In order to deal always with an infinite sequence of eigenvalues, in the case that T possesses exactly N eigenvalues, we define $\lambda_i(T) = 0$ for $i > N$.

Definition B.2.1. *Let $0 < p < \infty$. By $\mathcal{S}_p^{eig}$ we denote the class $\mathcal{S}_p^{eig} = \bigcup_{E,F} \mathcal{S}_p^{eig}(E,F)$ consisting of the sets*

$$\mathcal{S}_p^{eig}(E,F) = \left\{ T \in \mathcal{L}(E,F) \,\middle|\, ST \text{ is Riesz with } p\text{-summing eigenvalues } \forall S \in \mathcal{L}(F,E) \right\}.$$

Remark B.2.2.

a) The operator ideal $\mathcal{F}$ of finite rank operators is contained in the class $\mathcal{S}_p^{eig}$.

b) The class $\mathcal{S}_p^{eig}$ satisfies the multiplicativity property $\mathcal{L} \circ \mathcal{S}_p^{eig} \circ \mathcal{L} \subset \mathcal{S}_p^{eig}$ (that is ideal property (iii) in the Definition of an operator ideal).

The latter property is an immediate consequence of the principle of related operators (cf. Pietsch [73]).

Proposition B.2.3. (Principle of related operators) *The operators $S \in \mathcal{L}(E)$ and $T \in \mathcal{L}(F)$ are called* related *if there exist $A \in \mathcal{L}(F,E)$, $B \in \mathcal{L}(E,F)$ such that $S = AB$ and $T = BA$. Then S is a Riesz operator if and only if T is Riesz, and both operators have the same non-zero eigenvalues with the same multiplicities.*

Nevertheless, it is not true that $\mathcal{S}_p^{eig}$ is an operator ideal considered over the class of all Banach spaces. More precisely, it can be shown that for all $0 < p < \infty$ there exists a Banach space E such that $\mathcal{S}_p^{eig}(E,E)$ is no vector space.

In contrast, we mention that $\mathcal{S}_p^{eig}(H)$ is an operator ideal over the class of the separable infinite-dimensional Hilbert spaces H. Indeed, for $0 < p < \infty$ the class $\mathcal{S}_p^{eig}(H)$ coincides with the Schatten ideal $\mathcal{S}_p(H)$ of type l_p (confer Pietsch [72] for those statements).

Considering quasi-Banach operator ideals $\mathcal{A}$ contained in $\mathcal{S}_p^{eig}$, the p-norm of the eigenvalue sequence can be estimated.

Proposition B.2.4. (Principle of boundedness) *Let $\mathcal{A}$ be a quasi-Banach operator ideal with $\mathcal{A} \subset \mathcal{S}_p^{eig}$. Then there exists a universal constant $c \geq 1$ such that*

$$\Big(\sum_{i=1}^{\infty} |\, \lambda_i(T) \,|^p \Big)^{\frac{1}{p}} \leq c \, \| \, T \mid \mathcal{A} \, \| \qquad \text{for all } T \in \mathcal{A}(E) \text{ and all Banach spaces } E.$$

If, in addition, we confine ourselves to the class $\mathcal{S}_1^{eig}$, the following deep result was obtained by White [100].

Proposition B.2.5. (Spectral trace)
Let $\mathcal{A}$ be a quasi-Banach operator ideal such that $\mathcal{A} \subset \mathcal{S}_1^{eig}$. For arbitrary Banach spaces E and every $T \in \mathcal{A}(E)$ we define

$$\mathrm{tr}_\lambda(T) = \sum_{i=1}^{\infty} \lambda_i(T).$$

Then the function tr_λ is a continuous spectral trace on $\mathcal{A}$.

In general the spectral trace tr_λ is not unique as an observation of Kalton [51] shows. He has proved the existence of a quasi-Banach operator ideal $\mathcal{A} \subset \mathcal{S}_1^{eig}$ admitting different continuous traces.

To guarantee the uniqueness of the trace tr_λ, we use the above result of White combined with the trace extension theorem (cf. Pietsch [73]).

Proposition B.2.6. (Trace extension theorem)
Let $\mathcal{A}$ be a quasi-Banach operator ideal such that for all Banach spaces E and F the components $\mathcal{F}(E, F)$ of finite rank operators are $\| \cdot \mid \mathcal{A} \|$-dense in $\mathcal{A}(E, F)$. If there exists a constant $c \geq 1$ with

$$|\, \mathrm{tr}(T) \,| \leq c \, \| \, T \mid \mathcal{A} \, \| \qquad \text{for all } T \in \mathcal{F}(E) \text{ and all Banach spaces } E,$$

then $\mathcal{A}$ admits a unique continuous trace denoted by $\mathrm{tr}_{\mathcal{A}}$.

The unique trace $\mathrm{tr}_{\mathcal{A}} : \mathcal{A}(E) \to \mathbb{C}$ can be defined for all $T \in \mathcal{A}(E)$ by the $\| \cdot \mid \mathcal{A} \|$-continuous extension of $\mathrm{tr} : \mathcal{F}(E) \to \mathbb{C}$.

Remark B.2.7. *Let us further mention (confer Pietsch [73]) that there exist quasi-Banach operator ideals $\mathcal{A}$, $\mathcal{B} \subset \mathcal{S}_1^{eig}$ such that $\mathcal{A} + \mathcal{B} \not\subset \mathcal{S}_1^{eig}$. Thus, if tr_λ denotes the spectral trace on $\mathcal{A}$ and $\mathcal{B}$, respectively, then by*

$$\tau(A + B) = \mathrm{tr}_\lambda(A) + \mathrm{tr}_\lambda(B) \qquad \text{for all } A \in \mathcal{A}(E) \text{ and } B \in \mathcal{B}(E)$$

we obtain a continuous trace on $\mathcal{A} + \mathcal{B}$ which is not spectral.

In particular, the class $\mathcal{S}_1^{eig}$ does not contain a quasi-Banach operator ideal containing all other ideals in $\mathcal{S}_1^{eig}$.

Next we turn to the approach to extensions of the trace motivated by (1).

To this end, we introduce the quasi-Banach operator ideal $\mathcal{N}_r$ of r-nuclear operators in the sense of Grothendieck.

Definition B.2.8. *Let $0 < r \leq 1$. An operator $T \in \mathcal{L}(E, F)$ is called* r-nuclear *if it admits a so-called r-nuclear representation*

$$T = \sum_{i=1}^{\infty} a_i \otimes y_i \qquad \text{with } \sum_{i=1}^{\infty} \| \, a_i \, \|^r \cdot \| \, y_i \, \|^r < \infty \qquad (a_i \in E', y_i \in F).$$

By $\mathcal{N}_r$ we denote the class of all r-nuclear operators. $\mathcal{N}_r$ becomes an r-Banach operator ideal with respect to the quasi-norm

$$\| T \mid \mathcal{N}_r \| = \inf \left\{ \Big(\sum_{i=1}^{\infty} \| a_i \|^r \cdot \| y_i \|^r \Big)^{\frac{1}{r}} \right\} \qquad \textit{for } T \in \mathcal{N}_r(E,F)$$

where the infimum is taken over all possible representations.

If $r = 1$, $\mathcal{N}_r = \mathcal{N}_1 =: \mathcal{N}$ is a Banach operator ideal and we simply speak of nuclear operators in this situation.

Remark B.2.9. *$\mathcal{N}_r$ is the smallest r-Banach operator ideal.*

Now, given $T \in \mathcal{N}(E)$, E a Banach space, the natural idea to extend the trace tr on $\mathcal{F}$ as indicated by (1) would be to define

$$\mathrm{tr}_{\mathcal{N}}(T) = \sum_{i=1}^{\infty} \langle y_i, a_i \rangle \quad \text{for } T = \sum_{i=1}^{\infty} a_i \otimes y_i, \quad (a_i \in E', y_i \in E).$$

Unfortunately this expression depends on the underlying representation of the operator T. This was a long outstanding problem finally solved by Enflo [27] who constructed a Banach space without approximation property.

> *A Banach space E has the* approximation property (a.p.) *if, given any precompact subset M of E and any $\epsilon > 0$, there exists a finite rank operator $L \in \mathcal{F}(E)$ such that $\| z - Lz \| \leq \epsilon$ for all $z \in M$.*

Since the approximation property of a Banach space E is equivalent to the estimate $| \mathrm{tr}_{\mathcal{N}}(T) | \leq \| T \mid \mathcal{N} \|$ for $T \in \mathcal{F}(E)$, by the trace extension theorem we observe

Proposition B.2.10. *Considering the Banach operator ideal $\mathcal{N}$ restricted to the class of Banach spaces with a.p., the expression $\mathrm{tr}_{\mathcal{N}}$ defines a unique continuous trace.*

However, this trace may behave rather strangely. Since the ideal components $\mathcal{N}(l_p)$, $1 \leq p \leq \infty$, belong to $\mathcal{S}_q^{eig}(l_p)$ with the optimal eigenvalue type $\frac{1}{q} = 1 - \mid \frac{1}{2} - \frac{1}{p} \mid$ (see König [52]), the following situations occur.

$p = 2$ (best eigenvalue type l_1). By the spectral-trace theorem we obtain Lidskij's well-known spectral trace formula $\mathrm{tr}_{\mathcal{N}}(T) = \mathrm{tr}_{\lambda}(T)$ for $T \in \mathcal{N}(l_2)$ (see Reed/Simon [80]). (As a remark, the Hilbert space components $\mathcal{N}(H)$, H a Hilbert space, are the well-known trace classes of Schatten.)

$p = 1$ (worst eigenvalue type l_2). By Enflo [27] there exists an operator $S \in \mathcal{N}(l_1)$ with $\mathrm{tr}_{\mathcal{N}}(S) = 1$ and $S^2 = 0$. Because the nilpotent operator S does not possess any eigenvalue $\lambda_0 \neq 0$, it is impossible to compute the trace $\mathrm{tr}_{\mathcal{N}}(S)$ from the trivial eigenvalue sequence $(0, 0, \ldots)$. Thus, the trace $\mathrm{tr}_{\mathcal{N}}$ is not spectral.

We have noticed that even on the smallest Banach operator ideal $\mathcal{N}$ of nuclear operators over the class of (all) Banach spaces a well defined trace $\mathrm{tr}_{\mathcal{N}}$ does not exist. From this point of view, we now are interested in the smaller r-Banach operator ideals $\mathcal{N}_r$ $(0 < r \leq 1)$.

Due to the fact that the r-Banach operator ideal $\mathcal{N}_r$ belongs to $\mathcal{S}_q^{eig}$ with $\frac{1}{q} = \frac{1}{r} - \frac{1}{2}$, we obtain $\mathcal{N}_r \subset \mathcal{S}_1^{eig}$ for $0 < r \leq \frac{2}{3}$, and hence there exists the spectral trace tr_{λ} on $\mathcal{N}_r$. Moreover, in this case it is known that

$$\mathrm{tr}_{\mathcal{N}_r}(T) = \sum_{i=1}^{\infty} \langle y_i, a_i \rangle \quad \text{for} \quad T = \sum_{i=1}^{\infty} a_i \otimes y_i \in \mathcal{N}_r(E)$$

defines a trace on $\mathcal{N}_r(E)$ over the class of (all) Banach spaces E. This trace turns out to be unique, and, therefore it coincides with the spectral trace on $\mathcal{N}_r$.

Summarizing the previous discussion, we point out that the smallest Banach operator ideal $\mathcal{N}$ of nuclear operators considered over the class of all Banach spaces does not possess a continuous trace. Even if we restrict the considerations to the class of Banach spaces with approximation property, the obtained trace fails to be spectral. Therefore, searching for spectral traces, we are necessarily led to the context of quasi-Banach operator ideals.

Let us now explain the connection of the concept of traces on quasi-Banach operator ideals with that one of determinants. In general, the link between traces and determinants is governed by the trace-determinant theorem (confer Grobler/Raubenheim/Eldik [45], Pietsch [73]).

Proposition B.2.11. (Trace-determinant theorem). *There exists a one-to-one correspondence between continuous traces and continuous determinants on every quasi-Banach operator ideal.*

Furthermore, the following result concerning the differentiation of determinants holds.

Proposition B.2.12. *Let $\mathcal{A}$ be a quasi-Banach operator ideal admitting a continuous determinant δ.*

Suppose that the $\mathcal{A}(E)$-valued function $T(z)$ is defined on a domain of the complex plane. If $T(z)$ is complex-differentiable at a point z_0 with respect to the quasi-norm $\| \cdot \mid \mathcal{A} \|$, then the complex-valued function $\delta(I+T(z))$ is differentiable at z_0 as well.

In the particular case that $I+T(z_0)$ is invertible, the derivative is given by

$$\frac{\partial}{\partial z}\Big(\delta(I+T(z_0))\Big) = \tau\Big((I+T(z_0))^{-1}\frac{\partial}{\partial z}T(z_0)\Big)\ \delta\Big(I+T(z_0)\Big)$$

where τ is the corresponding trace defined by $\tau(S) := \lim\limits_{z\to 0}\frac{1}{z}\Big(\delta(I+zS)-1\Big)$ for $S\in\mathcal{A}(E)$.

As counterparts to the trace formulae, at last we state some results about the extensions of the determinant det on $\mathcal{F}$.

On the Banach operator ideal $\mathcal{N}$ of nuclear operators restricted to the class of Banach spaces with a.p. the unique determinant $\det_{\mathcal{N}}$ is described by

$$\det{}_{\mathcal{N}}(I+T) \;=\; 1+\sum_{n=1}^{\infty}\alpha_n(T)$$

$$\text{with}\quad \alpha_n(T) = \frac{1}{n!}\sum_{i_1=1}^{\infty}\cdots\sum_{i_n=1}^{\infty} det\begin{pmatrix} \langle x_{i_1}, a_{i_1}\rangle & \cdots & \langle x_{i_1}, a_{i_n}\rangle \\ \vdots & & \vdots \\ \langle x_{i_n}, a_{i_1}\rangle & \cdots & \langle x_{i_n}, a_{i_n}\rangle \end{pmatrix}$$

for $T=\sum\limits_{i=1}^{\infty} a_i\otimes x_i\in\mathcal{N}(E)$.

For the r-Banach operator ideal $\mathcal{N}_r$ of r-nuclear operators over the class of all Banach spaces $(0<r\le\frac{2}{3})$ the above formula holds as well.

And again, on quasi-Banach operator ideals $\mathcal{A}\subset\mathcal{S}_1^{eig}$ we obtain the spectral determinant $\det_\lambda$ given by $\det_\lambda(I+T)=\prod_i\big(1+\lambda_i(T)\big)$.

Appendix C

Supplement to the KP equation: The case of commuting parameters

For the sake of comparability, in this appendix we provide the proof of Theorem 3.2.1 in the particular case that the operator-valued parameters A, B commute. Proposition C.1.13 has already been stated in [18], Proposition 5.1, but without proof. We use the opportunity to supplement this here.

However, note that in this simplified setting the proof can be obtained by an although lengthy but completely straightforward calculation.

Proposition C.1.13. *Let E be a Banach space and A, $B \in \mathcal{L}(E)$ with $[A, B] = 0$.*

Assume that $L = L(x, y, t) \in \mathcal{L}(E)$ is an operator-valued function which solves the base equations

$$L_x = (A + B)\, L, \quad L_y = \frac{1}{\alpha}(A^2 - B^2)\, L, \quad L_t = -4(A^3 + B^3)\, L.$$

Then, on $\Omega = \{(x, y, t) \in \mathbb{R}^3 \,|\, (I + L) \text{ is invertible}\}$, a solution of the non-abelian KP (3.4), (3.5) is given by

$$\begin{aligned} U &= 2V_x, \\ W &= 2V_y, \end{aligned}$$

where the operator-valued function $V = V(x, y, t) \in \mathcal{L}(E)$ is defined by

$$V = (I + L)^{-1}(AL + LB).$$

Note that in the case of commuting parameter operators A, B, we also lose the freedom to choose the underlying Banach spaces independently.

Proof First recall that, by Lemma 3.2.2, it suffices to show that the operator-valued function $V = (I + L)^{-1}(AL + LB)$ solves the integrated version (3.13) of the non-abelian KP equation.

Next we introduce the abbreviations

$$a = A + B, \quad b = \frac{1}{\alpha}(A^2 - B^2), \quad c = -\,4(A^3 + B^3).$$

Note that c actually is a function of a, b,

$$c = -(a^4 + 3\alpha^2 b^2)/a. \tag{C.1}$$

With these abbrviations, the base equations simply read

$$L_x = aL, \qquad L_y = bL, \qquad L_t = cL.$$

Setting $M = (AL + LB)$, it is clear that M satisfies the same base equations as L.

Next we state two simple auxiliary identities, where we use Lemma 1.2.4 and the base equations for the second, namely

$$\begin{aligned}
(I+L)^{-1}L &= (I+L)^{-1}\big((I+L)-I\big) = I-(I+L)^{-1} = L(I+L)^{-1},\\
\big((I+L)^{-1}\big)_x &= -(I+L)^{-1}L_x(I+L)^{-1}\\
&= -(I+L)^{-1}aL(I+L)^{-1}\\
&= -(I+L)^{-1}a\big(I-(I+L)^{-1}\big).
\end{aligned}$$

With $X = (I+L)^{-1}$, these auxiliary identities read

$$XL = I - X = LX, \tag{C.2}$$

$$X_x = -Xa + XaX, \tag{C.3}$$

and analogously $X_y = -Xb + XbX$, $X_t = -Xc + XcX$.

After these preliminarities, we start with the calculation of the derivatives of $V = XM$. By the usual product rule, the base equations for M, and (C.3), we get

$$\begin{aligned}
V_x &= X_xM + XM_x\\
&= (-Xa + XaX)M + XaM\\
&= XaXM\\
&= XaV
\end{aligned}$$

and

$$\begin{aligned}
V_{xx} &= X_xaV + XaV_x\\
&= (-Xa + XaX)aV + (Xa)^2V\\
&= -Xa^2V + 2(Xa)^2V.
\end{aligned}$$

Analogously $V_y = XbV$, $V_{yy} = -Xb^2V + 2(Xb)^2V$ and $V_t = XcV$. Finally we have to differentiate V_{xx} once more with respect to x. Since

$$\begin{aligned}
\big((Xa)^2V\big)_x &= X_xaXaV + XaX_xaV + (Xa)^2V_x\\
&= (-Xa + XaX)aXaV + Xa(-Xa + XaX)aV + (Xa)^3V\\
&= -Xa^2XaV - XaXa^2V + 3(Xa)^3V
\end{aligned}$$

and

$$\begin{aligned}
(Xa^2V)_x &= X_xa^2V + Xa^2V_x\\
&= (-Xa + XaX)a^2V + Xa^2XaV\\
&= -Xa^3V + XaXa^2V + Xa^2XaV,
\end{aligned}$$

we obtain

$$V_{xxx} = Xa^3V - 3XaXa^2V - 3Xa^2XaV + 6(Xa)^3V.$$

To keep calculations as clear as possible, from now on we use the parameters A, B and a, b, c simultaneously. The following identity is the tool to control the nonlinear terms in (3.13). By the definitions of V, M, and (C.2),

$$\begin{aligned}
VX &= X(AL + LB)X\\
&= XA(LX) + (XL)BX\\
&= XA + BX - XaX.
\end{aligned} \tag{C.4}$$

It immediately yields

$$\begin{aligned} V_x^2 &= XaVXaV \\ &= Xa(XA+BX-XaX)aV \\ &= XaXaAV + XaBXaV - (Xa)^3V. \end{aligned}$$

Therefore,

$$\begin{aligned} & V_t + 6(V_x)^2 + V_{xxx} \\ &\quad = X(c+a^3)V + 3XaX(2aA-a^2)V + 3X(2aB-a^2)XaV \\ &\quad = X(c+a^3)V + 3\big(XaX(\alpha b) - X(\alpha b)Xa\big)V, \end{aligned}$$

where we have used

$$\begin{aligned} (2A-a)a &= (2A-(A+B))(A+B) \;=\; A^2-B^2 \;=\; \alpha b, \\ (2B-a)a &= (2B-(A+B))(A+B) \;=\; -(A^2-B^2) \;=\; -\alpha b. \end{aligned}$$

As a consequence,

$$\begin{aligned} & \big(V_t + 6(V_x)^2 + V_{xxx}\big)_x \\ &\quad = X_x(c+a^3)V + X(c+a^3)V_x \\ &\qquad\qquad +3\big(X_xaX(\alpha b)V + XaX_x(\alpha b)V + XaX(\alpha b)V_x\big) \\ &\qquad\qquad -3\big(X_x(\alpha b)XaV + X(\alpha b)X_xaV + X(\alpha b)XaV_x\big) \\ &\quad = (-Xa+XaX)(c+a^3)V + X(c+a^3)XaV \\ &\qquad +3\Big(\big(-Xa+XaX\big)aX(\alpha b) + Xa\big(-Xa+XaX\big)(\alpha b) + XaX(\alpha b)Xa\Big)V \\ &\qquad -3\Big(\big(-Xa+XaX\big)(\alpha b)Xa + X(\alpha b)\big(-Xa+XaX\big)a + X(\alpha b)XaXa\Big)V \\ &\quad = -X(c+a^3)aV + XaX(c+a^3)V + X(c+a^3)XaV \\ &\qquad\qquad -3Xa^2X(\alpha b)V - 3XaXa(\alpha b)V + 6XaXaX(\alpha b)V + 3XaX(\alpha b)XaV \\ &\qquad\qquad +3Xa(\alpha b)XaV + 3X(\alpha b)Xa^2V - 3XaX(\alpha b)XaV - 6X(\alpha b)XaXaV \\ &\quad = -X(c+a^3)aV \\ &\qquad\qquad +XaX(c+a^3-3a(\alpha b))V + X(c+a^3+3a(\alpha b))XaV \\ &\qquad +3(X(\alpha b)Xa^2 - Xa^2X(\alpha b))V \\ &\qquad +6\Big((Xa)^2\big(X(\alpha b)\big) - \big(X(\alpha b)\big)(Xa)^2\Big)V. \end{aligned}$$

Using (C.1) and $a^{-1}(\alpha b) + a = (A+B)^{-1}(A^2-B^2) + (A+B) = 2A$ or, with the other sign, $a^{-1}(\alpha b) - a = (A+B)^{-1}(A^2-B^2) - (A+B) = -2B$, we infer

$$\begin{aligned} (c+a^3)a &= -3(\alpha b)^2, \\ c+a^3+3a(\alpha b) &= -3(\alpha b)\big(a^{-1}(\alpha b)-a\big) = 6(\alpha b)B, \\ c+a^3-3a(\alpha b) &= -3(\alpha b)\big(a^{-1}(\alpha b)+a\big) = -6(\alpha b)A. \end{aligned}$$

Inserting these yields

$$\begin{aligned} \big(V_t + 6(V_x)^2 + V_{xxx}\big)_x &= 3X(\alpha b)^2V \\ &\quad +6\Big(X(\alpha b)BXa - XaX(\alpha b)A\Big)V + 3\big(X(\alpha b)Xa^2 - Xa^2X(\alpha b)\big)V \\ &\quad +6\Big((Xa)^2\big(X(\alpha b)\big) - \big(X(\alpha b)\big)(Xa)^2\Big)V. \end{aligned}$$

By (C.4) we get for the remaining nonlinear term

$$\begin{aligned}
[V_x, V_y] &= XaV\,XbV - XbV\,XaV \\
&= Xa\big(XA + BX - XaX\big)bV - Xb\big(XA + BX - XaX\big)aV \\
&= \Big(XaXbA - XbBXa\Big)V + \big(XaBXb - XbXAa\big)V \\
&\qquad +\Big(Xb(Xa)^2 - (Xa)^2Xb\Big)V,
\end{aligned}$$

and thus

$$\begin{aligned}
&\big(V_t + 6(V_x)^2 + V_{xxx}\big)_x + 6\alpha[V_x, V_y] = \\
&\quad = 3X(\alpha b)^2V + 3\big(X(\alpha b)Xa^2 - Xa^2X(\alpha b)\big)V \\
&\qquad\qquad\qquad\qquad +6\big(XaBX(\alpha b) - X(\alpha b)XAa\big)V \\
&\quad = \Big(-3\alpha^2V_{yy} + 6\big(X(\alpha b)\big)^2V\Big) + 3\big(X(\alpha b)Xa^2 - Xa^2X(\alpha b)\big)V \\
&\qquad\qquad\qquad\qquad +6\big(XaBX(\alpha b) - X(\alpha b)XAa\big)V \\
&\quad = -3\alpha^2V_{yy} + 6X(\alpha b)X(\alpha b)V \\
&\qquad\quad +3\Big(X(\alpha b)X(a^2 - 2aA) - X(a^2 - 2aB)X(\alpha b)\Big)V \\
&\quad = -3\alpha^2V_{yy} \\
&\qquad\quad +3\Big(X(\alpha b)X\big(a^2 - 2aA + (\alpha b)\big) - X\big(a^2 - 2aB - (\alpha b)\big)X(\alpha b)\Big)V.
\end{aligned}$$

To see that the term in the brackets vanishes, we just have to use $\alpha b + a^2 = 2aA$, $\alpha b - a^2 = -2aB$ once again. This completes the proof. $\square$

Bibliography

[1] M. J. Ablowitz and P. A. Clarkson. *Solitons, Nonlinear Evolution Equations and Inverse Scattering.* London Mathematical Society Lecture Notes Series **149**, Cambridge University Press 1991.

[2] M. J. Ablowitz, S. Chakravarty, A. D. Trubatch and J. Villaroel. *A novel class of solutions of the non-stationary Schrödinger and the Kadomtsev-Petviashvili I equations.* Phys. Lett. A **267**, 132-146 (2000).

[3] M. J. Ablowitz, D. J. Kaup, A. C. Newell, and H. Segur. *The inverse scattering transform – Fourier analysis for nonlinear problems.* Stud. Appl. Math. **53**, 249-315 (1974).

[4] M. J. Ablowitz and J. Satsuma. *Solitons and rational solutions of nonlinear equations.* J. Math. Phys. **19**, 2180-2186 (1978).

[5] M. J. Ablowitz and H. Segur. *Solitons and the Inverse Scattering Transform.* SIAM, Philadelphia 1981.

[6] M. J. Ablowitz and J. Villarroel. *On the discrete spectrum of the nonstationary Schrödinger equation and multipole lumps of the Kadomtsev-Petviashvili I equation.* Comm. Math. Phys. **207**, 1-42 (1999).

[7] H. Aden. *Elementary Operators and Solutions of the Korteweg-de Vries Equation.* Thesis, Jena 1996.

[8] H. Aden and B. Carl. *On realizations of solutions of the KdV equation by determinants on operator ideals.* J. Math. Phys. **37** (4), 1833-1857 (1996).

[9] S. Barran and M. Kovalyov. *A note on slowly decaying solutions of the defocusing nonlinear Schrödinger equation.* J. Phys. A, Math. Gen. **32**, 6121-6125 (1999).

[10] W. Bauhardt and C. Pöppe. *The Zakharov-Shabat inverse spectral problem for operators.* J. Math. Phys. **34**, 3073-3086 (1993).

[11] R. Beutler. *Positon solutions of the sine-Gordon equation.* J. Math. Phys. **34** (7), 3098-3109 (1993).

[12] R. Bhatia and P. Rosenthal. *How and why to solve the operator equation $AX-XB=Y$.* Bull. London Math. Soc. **29**, 1-21 (1997).

[13] H. Blohm. *On the Solution of Nonlinear Evolution Equations and Corresponding Cauchy Problems by Trace Methods.* Thesis, Jena 1999.

[14] H. Blohm. *Solution of nonlinear equations by trace methods.* Nonlinearity **13**, 1925-1964 (2000).

[15] A. Boutet de Monvel, E. Khruslov, and V. Kotlyarov. *Soliton asymptotics of the rear part of non-localized solutions of the Kadomtsev-Petviashvili equation.* J. Math. Phys. **9**, 58-76 (2002).

[16] R. K. Bullough and P. J. Caudrey (Ed.). *Solitons.* Topics in Current Physics **17**, Springer 1980.

[17] B. Carl and S.-Z. Huang. *On realizations of solutions of the KdV equation by the C_0-semigroup method.* Amer. J. Math. **122**, 403-438 (2000).

[18] B. Carl and C. Schiebold. *Nonlinear equations in soliton physics and operator ideals.* Nonlinearity **12**, 333-364 (1999).

[19] B. Carl and C. Schiebold. *Ein direkter Ansatz zur Untersuchung von Solitonengleichungen.* Jber. d. Dt. Math.-Verein. **102**, 102-148 (2000).

[20] A. Cauchy. *Exercises d'analyse et de physique mathématique, Tome 2.* Bachelier, Paris 1841.

[21] D. V. Chudnovsky. *The generalized Riemann-Hilbert problem and the spectral interpretation.* Nonlinear Evolution Equations and Dynamical Systems. Proc. Meeting Univ. Lecce, Lecce 1979, Lect. Notes Phys. **120**, Springer 1980, 103-149.

[22] A. T. Dash and M. Schechter. *Tensor products and joint spectra.* Israel J. Math. **8**, 191-193 (1970).

[23] A. Defant and K. Floret. *Tensor Norms and Operator Ideals.* North-Holland, Amsterdam (1993).

[24] J. Dieudonne. *Foundations in Modern Analysis.* 7th ed., Academic Press 1968.

[25] V. S. Drjuma. *On the analytical solution of the two-dimensional Korteweg-de Vries equation.* Sov. Phys. JETP Lett. **19**, 735-737 (1973).

[26] W. Eckhaus and A. van Harten. *The Inverse Scattering Transformation and the Theory of Solitons.* Mathematics Studies **50**, North-Holland Publishing Company 1981.

[27] P. Enflo. *A counterexample to the approximation problem in Banach spaces.* Acta Math. **130**, 309-317 (1973).

[28] A. Eschmeier. *Tensor products and elementary operators.* J. reine und angew. Mathematik **390**, 47-66 (1988).

[29] A. P. Fordy (Ed.). *Soliton Theory: A Survey of Result.* Nonlinear Science: Theory and Application. Manchester University Press 1990.

[30] N. C. Freeman. *Soliton solutions of non-linear evolution equations.* IMA J. Appl. Math. **32**, 125-145 (1984).

[31] N. C. Freeman and J. J. C. Nimmo. *Soliton solutions of the Korteweg-de Vries and Kadomtsev-Petviashvili equations: the Wronskian technique.* Proc. R. Soc. London A **389**, 319-329 (1983).

[32] B. Fuchssteiner. *Integrable nonlinear evolution equations with time-dependent coefficients.* J. Math. Phys. **34**, 5140-5158 (1993).

[33] B. Fuchssteiner. *Some tricks from the symmetry-toolbox for nonlinear equations: Generalizations of the Camassa-Holm equation.* Physica D **95**, 229-243 (1996).

[34] B. Fuchssteiner and R. N. Aiyer. *Multisolitons, or the discrete eigenfunctions of the recursion operator on non-linear evolution equations: II. Background.* J. Phys. A, Math. Gen. **20**, 375-388 (1987).

[35] B. Fuchssteiner and A. R. Chowdhury. *A new approach to the Quantum KdV.* Chaos, Solitons & Fractals **5**, 2345-2355 (1995).

[36] B. Fuchssteiner and W. Oevel. *Explicit formulas for symmetries and conservation laws of the Kadomtsev-Petviashvili equation.* Physics Lett. A **88**, 323-345 (1981).

[37] F. Gesztesy, H. Holden, E. Saab, and B. Simon. *Explicit construction of solutions of the modified Kadomtsev-Petviashvili equation.* J. Funct. Anal. **98**, 211-228 (1991).

[38] F. Gesztesy, W. Karwowski, and Z. Zhao. *Limits of soliton solutions.* Duke Math. J. **68**, 101-150 (1992).

[39] F. Gesztesy, W. Karwowski, and Z. Zhao. *New types of soliton solutions.* Bull. Am. Math. Soc. (New Series) **27**, 266-272 (1992).

[40] F. Gesztesy and W. Renger. *New classes of Toda soliton solutions.* Commun. Math. Phys. **184**, 27-50 (1997).

[41] I. C. Gohberg. *Traces and Determinants of Linear Operators.* Operator Theory: Advances and Applications **116**, Birkhäuser 2000.

[42] I. Gohberg, M. A. Kaashoek, and A. L. Sakhnovich. *Canonical systems with rational spectral densities: Explicit formulas and applications.* Math. Nachr. **194**, 93-125 (1998).

[43] I. Gohberg, M. A. Kaashoek, and A. L. Sakhnovich. *Pseudo-canonical systems with rational Weyl functions: Explicit formulas and applications.* J. Diff. Equations **146**, 375-398 (1998).

[44] I. Gohberg, M. A. Kaashoek, and A. L. Sakhnovich. *Scattering problems for a canonical system with a pseudo-exponential potential.* Asymptotic Analysis **29**, 1-38 (2002).

[45] J. Grobler, H. Raubenheimer, and P. van Eldik. *Fredholm theory for operators in an operator ideal with a trace.* Integral Equations Operator Theory **5**, 774-790 (1982).

[46] R. Hirota. *Direct methods in soliton theory.* In [16], 157-176.

[47] R. Hirota. *Direct methods of finding exact solutions of nonlinear evolution equations.* In [65], 40-69.

[48] R. Hirota and J. Satsuma. *A variety of nonlinear network equations generated from the Bäcklund transformation for the Toda lattice.* Progr. Theoret. Phys. Suppl. **59**, 64-100 (1976).

[49] M. Jaworski and J. Zagrodzinski. *Positons and positon-like solutions of the Korteweg-de Vries equation and the sine-Gordon equaiton.* Chaos, Solitons, and Fractals **5**, 229-2233 (1995).

[50] B. B. Kadomtsev and V. L. Petviashvili. *On the stability of solitary waves in weakly dispersing media.* Sov. Phys. Dokl. **15**, 539-541 (1970).

[51] N. Kalton. *Unusual traces on operator ideals.* Math. Nachrichten **134**, 119-130 (1987).

[52] H. König. *Eigenvalue Distribution of Compact Operators.* Birkhäuser, Basel 1986.

[53] B. G. Konopelchenko. *On the structure of nonlinear evolution equations and their Bäcklund transformations connected with the matrix nonstationary Schrödinger spectral problem.* J. Phys. A **15**, 3425-3437 (1982).

[54] S. V. Manakov, V. E. Zakharov, L. A. Bordag, .A. R. Its, and V. B. Matveev. *Two-dimensional solitons of the Kadomtsev-Petviashvili equation and their interaction.* Phys. Lett. A **63**, 205-206 (1977).

[55] V. A. Marchenko. *Nonlinear Equations and Operator Algebras.* Reidel, Dordrecht 1988.

[56] V. B. Matveev and M. A. Salle. *Darboux Transformations and Solitons.* Springer, Berlin 1991.

[57] V. B. Matveev. *Theory of Positons I.* Preprint of the Max Planck Institut für Metallforschung Stuttgart (1992).

[58] V. B. Matveev. *Generalized Wronskian formula for solutions of the KdV equations: first applications.* Phys. Lett. A **166**, 205-208 (1992).

[59] V. B. Matveev. *Asymptotics for the multipositon-soliton τ function of the Korteweg-de Vries equation and the supertransparency.* J. Math. Phys. **35**, 2955-2970 (1994).

[60] V. B. Matveev. *Positons: slowly decreasing analogues of solitons.* Theor. Math. Phys. **131**, 483-497 (2002).

[61] E. Medina. *An N soliton resonance solution for the KP equation: interaction with change of form and velocity.* Lett. Math. Phys. **62**, 91-99 (2002).

[62] R. Meinel, G. Neugebauer, and H. Steudel. *Vandermonde-like determinants and N-fold Darboux/Bäcklund transformations.* J. Math. Phys. **38**, 4692-4695 (1984).

[63] R. Meinel, G. Neugebauer, and H. Steudel. *Solitonen: Nichtlineare Strukturen.* Akademie-Verlag 1991.

[64] J. W. Miles. *Diffraction of solitary waves.* J. Fluid. Mech **79**, 171-179 (1977).

[65] R. Miura (Ed.). *Bäcklund Transformations, the Inverse Scattering Method, Solitons, and Their Applications.* NSF Research Workshop on Contact Transformations. Lecture Notes in Mathematics 515, Springer 1976.

[66] E. Date, M. Jimbo, and T. Miwa. *Solitons: Differential Equations, Symmetries and Infinite Dimensional Algebras.* Cambridge University Press 2000.

[67] J. J. C. Nimmo. *Wronskian determinants, the KP hierarchy and supersymmetric polynomials.* J. Phys. A **22**, 3213-3221 (1989).

[68] J. J. C. Nimmo. *Hirota's method.* In [29], 75-96.

[69] S. Novikov, S. V. Manakov, L. P. Pitaevskii, and V. E. Zakharov. *Theory of Solitons. The Inverse Scattering Method.* Consultants Bureau, New York 1984.

[70] K. Ohkuma and M. Wadati. *The Kadomtsev-Petviashvili equation: the trace method and soliton resonances.* J. Phys. Soc. Jpn. **52**, 749-760 (1983).

[71] E. Olmedilla. *Multiple pole solutions of the non-linear Schrödinger equation.* Physica D **25**, 330-346 (1986).

[72] A. Pietsch. *Operator Ideals.* VEB Deutscher Verlag der Wissenschaften, Berlin 1978.

[73] A. Pietsch. *Eigenvalues and s-Numbers.* Cambridge University Press, Cambridge 1987.

[74] G. Pisier. *Factorization of Linear Operators and Geometry of Banach spaces.* CBMS Regional Conference Series **60** AMS (1986).

[75] C. Pöppe. *Construction of solutions of the sine-Gordon equation by means of Fredholm determinants.* Physica **9D**, 103-139 (1983).

[76] C. Pöppe. *General determinants and the τ function for the Kadomtsev-Petviashvili hierarchy.* Inverse Problems **5**, 613-630 (1989).

[77] C. Pöppe and D. H. Sattinger. *Fredholm determinants and the τ function for the Kadomtsev-Petviashvili hierarchy.* Publ. RIMS, Kyoto Univ. **24**, 505-538 (1988).

[78] G. Pólya and G. Szegö. *Aufgaben und Lehrsätze zur Analysis II.* Springer (1976).

[79] C. Rasinariu, U. Sukhatme, and A. Khare. *Negaton and positon solutions of the KdV and mKdV hierarchy.* J. Phys. A: Math. Gen. **29**, 1803-1823 (1996).

[80] M. Reed and B. Simon. *Methods of Modern Mathematical Physics*, Vol. IV. Academic Press, New York 1979.

[81] W. Renger. *A new class of soliton solutions for the (modified) Kadomtsev-Petviashvili equation.* Math. Nachr. **202**, 125-140 (1999).

[82] W. Rudin. *Functional Analysis.* Tata Mc Graw-Hill 1988.

[83] A. L. Sakhnovich. *Exact solutions of nonlinear equations and the method of operator identities.* Linear Algebra Appl. **182**, 109-126 (1993).

[84] A. L. Sakhnovich. *Dressing procedure for solutions of non-linear equations and the method of operator identities.* Inverse Problems **10**, 699-710 (1994).

[85] A. L. Sakhnovich. *Dirac type system on the axis: Explicit formulae for matrix potentials with singularities and soliton-positon interactions.* Inverse Problems **19**, 845-854 (2003).

[86] A. L. Sakhnovich. *Matrix Kadomtsev-Petviashvili eqaution: matrix identities and explicit non-singular solutions.* J. Phys. A: Math. Gen. **36**, 5023-5033 (2003).

[87] J. Satsuma. *N-soliton solution of the two-dimensional Korteweg-de Vries equation.* J. Phys. Soc. Japan **40**, 286-290 (1976).

[88] C. Schiebold. *Funktionalanalytische Methoden bei der Behandlung von Solitonengleichungen.* Thesis, Jena 1996.

[89] C. Schiebold. *An operator theoretic approach to the Toda lattice equation.* Physica D **122**, 37-61 (1998).

[90] C. Schiebold. *Solitons of the sine-Gordon equation coming in clusters.* Revista Matemática Complutense **15**, 265-325 (2002).

[91] C. Schiebold. *On negatons of the Toda lattice.* J. Nonlin. Math. Phys. **10**, Suppl. 2, 180-192 (2003).

[92] C. Schiebold. *From the non-abelian to the scalar two-dimensional Toda lattice.* Glasgow Math J. (to appear)

[93] P. C. Schuur. *Asymptotic Analysis of Soliton Problems. An Inverse Scattering Approach.* Lecture Notes in Mathematics 1232, Springer 1986.

[94] B. Simon. *Trace Ideals and Their Applications.* London Math. Lecture Notes 35, Cambridge University Press, Cambridge 1979.

[95] A. A. Stahlhofen. *Positons of the modified Korteweg-de Vries equation.* Ann. Physik **1**,554-569 (1992).

[96] H. Steudel. *Cloning of Solitons.* Manuscript 1994.

[97] E. L. Stout. *Theory of Uniform Algebras.* Bogden and Quigley 1971.

[98] H. Tsuru and M. Wadati. *The multiple pole solutions of the sine-Gordon equation.* J. Phys. Soc. Japan **9**, 2908-2921 (1984).

[99] M. Wadati and K. Ohkuma. *Multiple pole solutions of the modified Korteweg-de Vries equation.* J. Phys. Soc. Jpn. **51**, 2029-2035 (1981).

[100] M. C. White. *Analytic multivalued functions and spectral trace.* Math. Ann. **304**, 669-683 (1996).

[101] V. E. Zakharov and A. B. Shabat. *Exact theory of two-dimensional self focusing and one-dimensional self modulation of waves in non-linear media.* Sov. Phys. JETP **34**, 62-69 (1972).

[102] V. E. Zakharov and A. B. Shabat. *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).