Nonlinear Fourier transform—towards the construction of nonlinear Fourier modes

To cite this article: Pavle Saksida 2018 J. Phys. A: Math. Theor. 51 015205

View the article online for updates and enhancements.
Nonlinear Fourier transform—towards the construction of nonlinear Fourier modes

Pavle Saksida

Institute of Mathematics, Physics and Mechanics, Jadranska 19, 1000 Ljubljana, Slovenia
Faculty of Mathematics and Physics, University of Ljubljana, Jadranska 21, 1000 Ljubljana, Slovenia
E-mail: Pavle.Saksida@fmf.uni-lj.si

Received 12 April 2017, revised 15 October 2017
Accepted for publication 18 October 2017
Published 5 December 2017

Abstract
We study a version of the nonlinear Fourier transform associated with ZS-AKNS systems. This version is suitable for the construction of nonlinear analogues of Fourier modes, and for the perturbation-theoretic study of their superposition. We provide an iterative scheme for computing the inverse of our transform. The relevant formulae are expressed in terms of Bell polynomials and functions related to them. In order to prove the validity of our iterative scheme, we show that our transform has the necessary analytic properties. We show that up to order three of the perturbation parameter, the nonlinear Fourier mode is a complex sinusoid modulated by the second Bernoulli polynomial. We describe an application of the nonlinear superposition of two modes to a problem of transmission through a nonlinear medium.

Keywords: nonlinear Fourier transform, inverse scattering, integrable systems, nonlinear Fourier modes, Bell polynomials, Bernoulli polynomials, nonlinear Schrödinger equation

1. Introduction

The theory of the inverse scattering transform is the most important tool for solving nonlinear integrable partial differential equations. It can be thought of as a nonlinear analogue of the Fourier approach to solving certain linear partial differential equations. The Fourier method of solving the initial value problems can be summarized by the well-known diagram

\[ q(x, 0) \xrightarrow{F} F[q](0; z) \xrightarrow{\text{time evolution}} F[q](t; z) \xrightarrow{F^{-1}} q(x, t). \]

First, we map the initial condition \( q(x, 0) \) by the Fourier transform \( F \), then we apply the time evolution operator to the transformed data, and finally we map the time-evolved Fourier data
by means of the inverse Fourier transform in order to obtain the state of our system at any desired future time \( t \).

The method of the inverse scattering transform can be illustrated by the same diagram, with \( F \) replaced by a nonlinear operator \( \mathcal{F} \) called the scattering transform, or the nonlinear Fourier transform. The inverse scattering method works very well for the open-ended initial value problems in which the solution paths \( t \mapsto q(x, t) \) lie in, for instance, the Schwartz class of functions of the space variable \( x \). If we impose the periodic or quasi-periodic boundary conditions, the problem becomes more difficult and the original version of the inverse scattering method alone does not suffice; an additional effort is required. A rich class of the so-called algebraic solutions of problems with periodic or quasi-periodic boundary conditions is provided by the finite-gap integration theory. This theory was developed by many authors over a long period of time, beginning in the mid 1970s. The volume of the relevant literature is large, and we only cite \([3, 11]\) and the historical overview in \([14]\). This beautiful and very successful combination of inverse scattering and algebraic geometry produced a host of extremely important results and insights. Nevertheless, it provides only the solutions with finitely many degrees of freedom and therefore cannot solve the general initial value problems. The work of Knörrer, Trubowitz, Feldman, Müller, Schmidt, and others on the theory of curves with infinite genus fills this gap by providing solutions with infinitely many degrees of freedom whose initial values are arbitrary elements of suitable \( L^2 \) spaces. These results are described in \([7, 15, 25]\), and in many other important and interesting papers.

The essential building blocks of the linear Fourier analysis are Fourier modes. The forward Fourier transform decomposes the initial condition into a set of Fourier modes. Their time evolution is easily found and gives rise to one degree of freedom solutions. In other words, a Fourier mode remains a Fourier mode throughout its time evolution. The final solution is then obtained by the superposition of these elementary solutions by means of the inverse Fourier transform. One way to try to solve the nonlinear initial value problem with periodic boundary conditions could be to try to mimic the linear theory closely.

In this paper, we study a version of the nonlinear Fourier transform \( \mathcal{F} \), defined on the space \( L^2[0, 2\pi] \), by means of which one will be able to construct nonlinear Fourier modes and also find their nonlinear superposition. This will facilitate the study of certain deformations of the ZS-AKNS integrable equations and, presumably, also the construction of some new integrable systems, not necessarily closely linked to the existing ZS-AKNS equations. For all these tasks, one needs a good grip on the inverse of \( \mathcal{F} \). We provide an iterative scheme which yields explicit approximations to the inverse image with respect to \( \mathcal{F} \) of arbitrary argument and to any desired accuracy. The construction of the inverse \( \mathcal{F}^{-1} \) is the central result of this paper.

Besides the motivation described above, there are other reasons for studying various aspects of nonlinear Fourier transforms. Different versions appear in many contexts. In \([26]\), the authors study a certain discrete version and its many applications. A better understanding of the nonlinear Fourier transform, and, in particular, of nonlinear superposition, is important for researchers who study the propagation of signals in optical fibres. These problems are treated, for example, in \([30, 31]\), in \([27, 29]\) and in \([28]\). A somewhat different version of nonlinear Fourier transformation was studied by Fokas and collaborators in \([9, 10]\), and in other papers. A kind of initial boundary problem was treated by means of this approach in, for example, \([8, 13]\). Important and profound results concerning analytical problems, relating to various versions of nonlinear Fourier transform, were obtained by Christ, Kiselev, Muscalu, Thiele, (see \([5, 4, 16, 17]\)), to name but a few.
1.1. Presentation of results

Our version of the nonlinear Fourier transform will be associated with the integrable systems of the ZS-AKNS type. The equations of the ZS-AKNS type are the integrable partial differential equations whose Lax pairs \((L, A)\) have the \(L\)-matrix of the form

\[
L_{q,r}(x, t; z) = \begin{pmatrix}
\frac{q(x)}{r(x, t)} & q(x, t) \\
-r(x, t) & -\frac{r(x)}{q(x)}
\end{pmatrix},
\]

where \(q\) and \(r\) are complex valued functions and \(z\) is the spectral parameter. In many cases of interest, we have \(r(x) = \pm q(x)\). These systems were first studied by Ablowitz, Kaup, Newell, and Segur ([1, 2]), and by Zakharov and Shabat [32].

By \(\text{LSL}(2; \mathbb{C})\) we shall denote the space of bi-infinite sequences in the special linear group \(\text{SL}(2; \mathbb{C})\). These sequences converge to the identity element \(I\) at both infinities. More precisely,

\[
\text{LSL}(2; \mathbb{C}) = \{g(n) = \begin{pmatrix} 1 + a(n) & b(n) \\ c(n) & 1 + d(n) \end{pmatrix} \}_{n \in \mathbb{Z}}
\]

and

\[
\{a(n)\}_{n \in \mathbb{Z}}, \{b(n)\}_{n \in \mathbb{Z}}, \{c(n)\}_{n \in \mathbb{Z}} \quad \text{and} \quad \{d(n)\}_{n \in \mathbb{Z}} \in \ell^2_\mathbb{Z},
\]

where \(\ell^2_\mathbb{Z}\) is the space of the square-summable bi-infinite complex sequences. We can think of the space \(\text{LSL}(2; \mathbb{C})\) as of a discrete loop group (consisting of loops based at the identity) in the Lie group \(\text{SL}(2; \mathbb{C})\).

Recall that the holonomy \(\text{Hol}[q, r](z)\) of the matrix \(L_{q,r}(x; z)\) is given by

\[
\text{Hol}[q, r](z) = \Phi(2\pi; z), \quad \text{where} \quad \Phi(x; z) = L_{q,r}(x; z) \cdot \Phi(x; z), \quad \Phi(0; z) = I.
\]

**Definition 1.** Let \((q, r) \in L^2[0, 2\pi] \times L^2[0, 2\pi]\) and let \(\{\sigma_n\}_{n \in \mathbb{Z}} \in \ell^2_\mathbb{Z}\). The map

\[
\mathcal{F} : L^2[0, 2\pi] \times L^2[0, 2\pi] \times \ell^2_\mathbb{Z} \longrightarrow \text{LSL}(2; \mathbb{C})
\]

defined by

\[
\mathcal{F}[q, r, \sigma](n) = (-1)^n \text{Hol}[q, r](n + \sigma_n)
\]

is called the nonlinear Fourier transform associated with ZS-AKNS systems.

We shall see later in the text that the target space of \(\mathcal{F}\) is indeed the discrete loop group \(\text{LSL}(2; \mathbb{C})\) described above. The matrix \(L_{q,r}(z)\) is traceless and therefore \(\mathcal{F}[q, r, \sigma](n)\) is unimodular for every \(q, r, \sigma\) and for every \(n \in \mathbb{Z}\). The value

\[
\mathcal{F}[q, r, \sigma] = \begin{pmatrix} 1 + \{a(n)\}_{n \in \mathbb{Z}} & \{b(n)\}_{n \in \mathbb{Z}} \\ \{c(n)\}_{n \in \mathbb{Z}} & 1 + \{d(n)\}_{n \in \mathbb{Z}} \end{pmatrix}
\]

is therefore given by the three \(\ell^2_\mathbb{Z}\) sequences \(\{a(n)\}_{n \in \mathbb{Z}}, \{b(n)\}_{n \in \mathbb{Z}}, \{c(n)\}_{n \in \mathbb{Z}}\), provided all \(\alpha_n > -1\). It will be convenient to use the auxiliary map

\[
\mathcal{H} : L^2[0, 2\pi] \times L^2[0, 2\pi] \times \ell^2_\mathbb{Z} \longrightarrow (\ell^2_\mathbb{Z})^3
\]

given by

\[
\mathcal{H}[q, r, \sigma](n) = (\alpha(n), \beta(n), \gamma(n)) = \left(\alpha[q, r, \sigma](n), \beta[q, r, \sigma](n), \gamma[q, r, \sigma](n)\right).
\]

(4)
The unimodularity of $\mathcal{F}[q,r,\sigma](n)$ for every $n$ ensures that the maps $\mathcal{F}$ and $\mathcal{H}$ are equivalent. The advantage of $\mathcal{H}$ over $\mathcal{F}$ lies in the fact that the image of $\mathcal{H}$ is a linear Hilbert space, while the image of $\mathcal{F}$ is a nonlinear Hilbert manifold. The map $\mathcal{H}$ will also be called the nonlinear Fourier transform.

In the paper, we prove two theorems.

**Theorem 1.** The nonlinear Fourier transform

$$\mathcal{H} : L^2[0,2\pi] \times L^2[0,2\pi] \times \mathbb{F}_2^3 \longrightarrow (\mathbb{F}_2^3)^3$$

is an analytic map. It is also invertible in the vicinity of the origin. There exist neighbourhoods $U \subset (L^2[0,2\pi])^2 \times \mathbb{F}_2$ and $V \subset (\mathbb{F}_2^3)^3$ of the respective origins such that the restriction $\mathcal{H} : U \longrightarrow V$ is invertible, and the inverse

$$\mathcal{H}^{-1} : V \longrightarrow U$$

is also analytic.

The next theorem describes an iterative scheme for the calculation of the inverse $\mathcal{H}^{-1}$ for an arbitrary argument and to any desired degree of accuracy. Theorem 1 ensures that our scheme is convergent.

**Theorem 2.** Let

$$s \longmapsto \{A_n(s)\}_{n \in \mathbb{Z}}, \{B_n(s)\}_{n \in \mathbb{Z}}, \{C_n(s)\}_{n \in \mathbb{Z}} = \left\{ \sum_{m=1}^{\infty} \frac{s^m}{m!} (A_n^{(m)}, B_n^{(m)}, C_n^{(m)}) \right\}_{n \in \mathbb{Z}}$$

be an analytic curve in $(\mathbb{F}_2^3)^3$ starting at the origin. The inverse image with respect to the nonlinear Fourier transform of this curve is the analytic curve

$$\mathcal{H}^{-1}\left( \left\{ \sum_{m=1}^{\infty} \frac{s^m}{m!} (A_n^{(m)}, B_n^{(m)}, C_n^{(m)}) \right\}_{n \in \mathbb{Z}} \right) = \left\{ \sum_{m=1}^{\infty} \frac{s^m}{m!} (q^{(m)}(x), r^{(m)}(x)), \left\{ \sum_{m=1}^{\infty} \frac{1}{m!} \sigma_n^{(m)} \right\}_{n \in \mathbb{Z}} \right\}$$

in $(L^2[0,2\pi])^2 \times \mathbb{F}_2$. The coefficients $q^{(m)}(x)$, $r^{(m)}(x)$ and $\{\sigma_n^{(m)}\}_{n \in \mathbb{Z}}$ can be computed iteratively by means of the system

$$\sigma_n^{(m)} = -\frac{i}{\pi} (A_n^{(m)} + B_{1,1}^{(m)} [q,r,\sigma](n))$$

$$q^{(m)}(x) = F^{-1}\left( \left\{ B_n^{(m)} \right\}_{n \in \mathbb{Z}} + \left\{ B_{1,2}^{(m)} [q,r,\sigma](n) \right\}_{n \in \mathbb{Z}} \right)$$

$$r^{(m)}(x) = F^{-1}\left( \left\{ C_{-n}^{(m)} \right\}_{n \in \mathbb{Z}} + \left\{ B_{2,1}^{(m)} [q,r,\sigma](-n) \right\}_{n \in \mathbb{Z}} \right),$$

where

$$B_{ij}^{(m)} [q,r,\sigma] = B_{ij}^{(m)} (q^{(1)}, \ldots, q^{(m-1)}, r^{(1)}, \ldots, r^{(m-1)}, \sigma^{(1)}, \ldots, \sigma^{(m-1)}),$$

and $F^{-1}$ denotes the linear inverse Fourier transform

$$F^{-1}\left( \left\{ K_n \right\}_{n \in \mathbb{Z}} \right) = \sum_{n=-\infty}^{\infty} K_n e^{inx}.$$
The values $B_{ij}^{(m)}[q,r,\sigma]$ can be expressed in terms of the Bell polynomials. The explicit expressions of $B_{ij}^{(m)}[q,r,\sigma]$ are given by formulae (41)–(43) on page 22.

As an immediate trivial corollary of theorem 2, we obtain the scheme for calculating preimages of points (as opposed to curves) in $(U_2^2)^3$ with respect to $\mathcal{H}$. Let \( \{A_n\}_{n \in \mathbb{Z}}, \{B_n\}_{n \in \mathbb{Z}}, \{C_n\}_{n \in \mathbb{Z}} \in (U_2^2)^3 \).\footnote{We use $U_2$ for the unitary 2x2 matrices.} We have

\[
\mathcal{H}^{-1}\left(\{A_n\}_{n \in \mathbb{Z}}, \{B_n\}_{n \in \mathbb{Z}}, \{C_n\}_{n \in \mathbb{Z}}\right) = \mathcal{H}^{-1}\left(s\{A_n\}_{n \in \mathbb{Z}}, s\{B_n\}_{n \in \mathbb{Z}}, s\{C_n\}_{n \in \mathbb{Z}}\right)_{i=1}.
\]

In this case, all coefficients $A_{(m)}^l$, $B_{(m)}^l$ and $C_{(m)}^l$ for $m > 1$ in theorem 2 are equal to zero.

The nonlinear Fourier transform defined above is suitable for the construction of nonlinear Fourier modes. Let $M_d$ be a matrix in $SL(2; \mathbb{C})$, and let the elements $\{M(n)\}_{n \in \mathbb{Z}}$ of $LSL(2; \mathbb{C})$ be given by

\[
M(n) = I + M_d \cdot \delta_{n,0},
\]

where $\delta_{n,0}$ is the Kronecker delta. Let

\[
\mathcal{F}^{-1}\{(M(n))_{n \in \mathbb{Z}}\} = (q_{M_d}(x), r_{M_d}(x), \{\sigma_n\}_{n \in \mathbb{Z}}).
\]

Then, the pair of functions $(q_{M_d}(x), r_{M_d}(x))$ can serve as a nonlinear analogue of a Fourier mode. It has the following property. There exists an $I$-sequence $\sigma = \{\sigma_n\}_{n \in \mathbb{Z}}$ such that

\[
\mathcal{F}[q_{M_d}, r_{M_d}, \sigma] = (-1)^n \text{Hol}[q, r](n + \sigma_n) = I + M_d \cdot \delta_{n,0}.
\]

Let $(q_{M_d}(x), r_{M_d}(x))$ be an initial state of an integrable ZS-AKNS system. Its time evolution is then given by

\[
\mathcal{F}\{(q_{M_d}(x), r_{M_d}(x), \sigma)\} = \{\text{Ad}_{\Psi(t,\sigma)} \mathcal{F}[q_{M_d}(x), r_{M_d}(x), \sigma](n)\}_{n \in \mathbb{Z}} = I + \text{Ad}_{\Psi(t,\sigma)} M_d \cdot \delta_{n,0}.
\]

for a suitable matrix function $\Psi(t; z)$. This is a nonlinear analogue of the time evolution of a linear Fourier mode, where the time evolution is confined to one degree of freedom. Moreover, the linearization of $(q_{M_d}, r_{M_d})$ is precisely the linear Fourier mode of frequency $d$. We will show on page 16 that the derivative of $\mathcal{F}$ at the origin is given by

\[
D_0 F[u,v,\tau](n) = \begin{pmatrix} i\pi n & F[u](n) \\ F[v](n) & -i\pi n \end{pmatrix} = \begin{pmatrix} \int_0^{2\pi} e^{i\pi n} q(x) \; dx \\ \int_0^{2\pi} e^{i\pi n} r(x) \; dx \end{pmatrix}.
\]

where $F$ is the linear Fourier transform. Thus, for every curve $s \mapsto (q_d(x;s), r_d(x;s))$ of the nonlinear Fourier modes of frequency $d$, the derivative $(d/ds)|_{s=0}(q_d(x,s); r_d(x,s))$ is a linear Fourier mode of the same frequency. Indeed, the nonlinear Fourier modes can be viewed as self-modulated linear modes. They have the form of an amplitude modulated complex or real sinusoids. Preliminary calculations show that the modulation envelope can be given as a perturbation series whose terms are expressed in terms of the Bernoulli periodic functions $B_n\left(\frac{x}{2\pi}\right)$, given by

\[
B_n\left(\frac{x}{2\pi}\right) = \frac{n!}{(2\pi)^n} \left(\text{Li}_n(e^{ix}) + (-1)^n \text{Li}_n(e^{-ix})\right),
\]

where $\text{Li}_n(z)$ is the polylogarithm. In this paper, we prove this claim up to the order three, where the first nontrivial term appears. We defer a more general and thorough discussion of this topic to another paper.
The iterative scheme, given in theorem 1, also enables us to calculate the nonlinear superposition of the nonlinear Fourier modes. The superposition of modes \((q_{M_d}, r_{M_d})\) is given by

\[
(q_{M_d}, r_{M_d}) = \text{proj} \left( F^{-1}\{M_d, g(n)\}_{n \in \mathbb{Z}} \right),
\]

where

\[
M_d, g(n) = I + M_d \cdot \delta_{n,d} + \bar{M}_g \cdot \delta_{n,g},
\]

and ‘\text{proj}’ is the projection on the functional part of the triple \((q_{M_d}, r_{M_d}, \sigma)\). Clearly, this procedure works for any finite or infinite number of Fourier modes. The author is convinced that this method of constructing a nonlinear superposition in which we have a rather direct control of the relevant parameters, will turn out to be useful in many different contexts. In this paper, we only propose a very simplified scheme of communication through a nonlinear medium (e.g. an optical fibre), described by the nonlinear Schrödinger equation. We intend to present a more thorough description of various applications in another paper.

The results of theorems 1 and 2 can be interpreted in the vein of Sturm–Liouville theory. Consider the operator

\[
L_{q,r} = \begin{pmatrix}
\frac{d}{dx} & 0 \\
0 & -\frac{d}{dx}
\end{pmatrix} + \begin{pmatrix} 0 & -q(x) \\
r(x) & 0
\end{pmatrix}
\]

acting on \(C^2\)-valued functions, defined on \([0, 2\pi]\). Let the matrix \(\Phi(x; z)\) be the solution of the initial problem

\[
\Phi_\chi = L_{q,r}(x; z) \cdot \Phi, \quad \Phi(0; z) = I,
\]

and let \(M(z) = \Phi(2\pi; z)\). Let

\[
V_2 = \{ \phi(x) = (\phi_1(x), \phi_2(x))^T; \phi(2\pi) = M(z) \cdot \bar{\phi}(0) \}
\]

be the subspace of \((L^2[0, 2\pi])^2\), defined by the quasi-periodicity condition in the bracket. Then, \(iz/2\) is the double eigenvalue of the restriction of \(L_{q,r}\) to \(V_2\), and the columns of \(\Phi(x; z)\) are a basis of the corresponding two-dimensional eigenspace. The nonlinear Fourier transform assigns to the operator \(L_{q,r}\) and to a sequence \(\{\sigma_n\}_{n \in \mathbb{Z}} \in \ell^2_\mathbb{Z}\) the sequence of (modified) monodromies \((-1)^n M_n = F[q, r, \sigma](n)\) such that for every integer \(n\) the value \(i(n + \sigma_n)/2\) is the double eigenvalue of \(L_{q,r}\) restricted to \(V_{n+\sigma_n}\). By theorem 1 this association is analytic in all the variables \(q, r\), \(\sigma\). More interestingly, theorem 1 ensures that the inverse problem is also solvable and that the solution is unique. Namely, given the element \(\{M_n\}_{n \in \mathbb{Z}} \in LSL(2; \mathbb{C})\), there exists precisely one operator \(L_{q,r}\) and one element \(\sigma \in \ell^2_\mathbb{Z}\) such that \(i(n + \sigma_n)/2\) are double eigenvalues of the restrictions of \(L_{q,r}\) to the subspaces \(V_{n+\sigma_n}\) which contain \(M_n\)-quasiperiodic \(L^2\) vector functions \((q, r)\). Theorem 2 provides an iterative scheme for the construction of \(L_{q,r}\) and \(\sigma\). If \((q, r)\) is a nonlinear Fourier mode of frequency \(d\), then for all \(n \neq d\) we have \(M_n = (-1)^n I\), and the values \(i(n + \sigma_n)/2\) are double eigenvalues of \(L_{q,r}\) restricted to the subspace of \(2\pi\)-periodic or anti-periodic functions. More generally, let \(D \subset \mathbb{Z}\) be a finite set. Functions \((q, r)\), such that \(i(n + \sigma_n)/2\) for \(n \in \mathbb{Z}\setminus D\) is the periodic–anti periodic double spectrum of the operator \(L_{q,r}\), are the initial conditions of the finite degrees of freedom solutions of ZS-AKNS type equations.

In the theory of the ZS-AKNS equations, the relevant elements \(\sigma \in \ell^2_\mathbb{Z}\) are the periodic and Dirichlet spectra of the operator \(L_{q,r}\). In this paper we allow \(\sigma\) to vary as a parameter. This additional amount of freedom will enable us to study certain deformations of the ZS-AKNS equations and, possibly, to construct some new integrable systems in a future work.
The approach to the nonlinear Fourier transform, presented in this paper, is motivated by some observations set forth in [21], [22], where the close relationship between the method of the inverse scattering transform and the linear Fourier analysis was considered. The nonlinear Fourier transform \( F \) studied here is related to the one described in my paper [23]. The essential difference between the two is the fact that by means of \( F \), one can construct the nonlinear Fourier modes and their superposition. The nonlinear Fourier transform from [23] does not allow that and this is its most important drawback. In my conference paper [24], two results of the present paper are presented in a rudimentary form. The general principle of the construction of \( F^{-1} \) is introduced and a nonlinear mode is calculated to the order three. No essential proofs are given. They are only announced to appear in the present paper. Here, we first of all prove the analyticity of \( F \) and of \( F^{-1} \), which guarantees the convergence of our iterative scheme. We also provide the newly constructed explicit formulae for the calculation of \( F^{-1} \).

The well-known Bernoulli polynomials play an important role here. In section 2 we prove theorem 1, and in section 3 we prove theorem 2. In section 4, we describe the application of our construction in a problem of communication through a nonlinear medium. We end the paper with some remarks collected in the final section.

2. Analytic properties

Our aim in this section is to prove theorem 1. We have to show that \( F \), or equivalently \( \mathcal{H} \), is analytic. By a well-known theorem of nonlinear functional analysis, an operator between two complex Banach spaces is analytic if it is locally bounded and weakly holomorphic. For the proof, see, for example, [12] or [18]. After some preparation, we shall prove the local boundedness of \( \mathcal{H} \) and then its weak holomorphicity. At the end, we shall prove the local invertibility of \( \mathcal{H} \) in the vicinity of the origin. We will do this by proving that the derivative \( D_0 \mathcal{H} \) of \( \mathcal{H} \) at the origin is a bounded invertible linear map. By the holomorphic version of the inverse mapping theorem for Banach spaces (see [12]), the map \( \mathcal{H} \) is invertible and the inverse \( \mathcal{H}^{-1} \) is also an analytic map.

2.1. A more explicit expression of \( \mathcal{H}(q, r, \sigma) \)

The definition 1 is slightly awkward and implicit. Therefore, we will begin with a more explicit expression of \( \mathcal{H} \).

First, let us change the gauge of the linear initial problem with the coefficient matrix \( L_{q, r} \) and let the gauge transformation matrix be given by

\[
G(x; z) = \begin{pmatrix} e^{-\frac{i}{2}z} & 0 \\ 0 & e^{\frac{i}{2}z} \end{pmatrix}.
\]

The matrix \( \Phi_{q, r}(x; z) \) is the solution of the \( L_{q, r} \)-initial problem if and only if the matrix

\[
\Phi_{q, r}^G(x; z) = G(x; z) \cdot \Phi_{q, r}(x; z)
\]

is the solution of

\[
(\Phi_{q, r}^G)_x = L_{q, r}^G \cdot \Phi_{q, r}^G, \quad \Phi_{q, r}^G(0; z) = I, \tag{6}
\]

where

\[
L_{q, r}^G(x; z) = \left( G_x \cdot G^{-1} + G \cdot L_{q, r} \cdot G^{-1} \right)(x; z) = \begin{pmatrix} 0 & e^{iz} q(x) \\ e^{iz} r(x) & 0 \end{pmatrix}.
\]
In terms of Dyson’s expansion, the solution of (6) is given by the convergent series
\[
\Phi^G(x; z) = I + \sum_{m=1}^{\infty} \int_{\Delta_m(x)} L^G_{q,r}(\xi_1; z) \cdot L^G_{r,q}(\xi_2; z) \cdots L^G_{q,r}(\xi_m; z) \, d\vec{\xi},
\]
where
\[
\Delta_m(x) = \{ (\xi_1, \xi_2, \ldots, \xi_m) : x \geq \xi_1 \geq \xi_2 \geq \ldots \geq \xi_m \geq 0 \}
\]
is the ordered simplex of dimension \( m \). Evaluation at \( x = 2\pi \) gives
\[
\text{Hol}^G_{q,r}(z) = \Phi^G_{q,r}(x = 2\pi; z) = \begin{pmatrix}
1 + \tilde{a}[q, r](z) & \tilde{b}[q, r](z) \\
\tilde{c}[q, r](z) & 1 + \tilde{d}[q, r](z)
\end{pmatrix},
\]
where
\[
\tilde{a}[q, r](z) = \sum_{m=1}^{\infty} \int_{\Delta_{2m}} e^{-i\pi(\xi_1 - \xi_2 + \cdots + \xi_{2m})} q(\xi_1) r(\xi_2) q(\xi_3) \cdots r(\xi_{2m}) \, d\vec{\xi}
\]
(7)
and \( \triangle_k = \Delta_k(x = 2\pi) \). The matrices \( \text{Hol}_{q,r}(z) \) and \( \text{Hol}^G_{q,r}(z) \) are related by the equation
\[
\text{Hol}_{q,r}(z) = G^{-1}(2\pi; z) \cdot \text{Hol}^G_{q,r}(z).
\]
If we denote
\[
\text{Hol}_{q,r}(z) = \begin{pmatrix}
\text{Hol}_{q,r}(z) & \tilde{b}[q, r](z) \\
\tilde{c}[q, r](z) & \text{Hol}^G_{q,r}(z)
\end{pmatrix},
\]
we get
\[
\tilde{a}[q, r](z) = \sum_{m=1}^{\infty} \int_{\Delta_{2m}} e^{-i\pi(\xi_1 - \xi_2 + \cdots + \xi_{2m})} q(\xi_1) r(\xi_2) q(\xi_3) \cdots r(\xi_{2m}) \, d\vec{\xi}
\]
(11)
and
\[
\tilde{b}[q, r](z) = \sum_{m=1}^{\infty} \int_{\Delta_{2m-1}} e^{-i\pi(\xi_1 - \xi_2 + \cdots + \xi_{2m-1})} q(\xi_1) r(\xi_2) q(\xi_3) \cdots q(\xi_{2m-1}) \, d\vec{\xi}
\]
(12)
and
\[
\tilde{c}[q, r](z) = \sum_{m=1}^{\infty} \int_{\Delta_{2m-1}} e^{i\pi(\xi_1 - \xi_2 + \cdots + \xi_{2m-1})} r(\xi_1) q(\xi_2) r(\xi_3) \cdots r(\xi_{2m-1}) \, d\vec{\xi}
\]
(13)
Evaluation of (10) at \( z = z_n = n + \sigma_n \) and multiplication by \((-1)^n\) finally give
\[
F[q, r, \sigma](n) = \begin{pmatrix}
1 + \alpha[q, r, \sigma](n) & \beta[q, r, \sigma](n) \\
\gamma[q, r, \sigma](n) & 1 + \delta[q, r, \sigma](n)
\end{pmatrix}
\]
with
\[
\alpha[g, r, \sigma](n) = (e^{i\pi \sigma} - 1) + (-1)^n \tilde{a}[g, r](n + \sigma_n),
\]
(14)
\[
\beta[g, r, \sigma](n) = (-1)^n \tilde{b}[g, r](n + \sigma_n),
\]
(15)
\[
\gamma[g, r, \sigma](n) = (-1)^n \tilde{c}[g, r](n + \sigma_n).
\]
(16)

We shall now separate the variables \(q\) and \(r\) on the one hand, and \(\sigma\) on the other, as much as possible. Let us expand the functions \(\exp(-i\xi_\alpha)\) into the Fourier series (in reverse order) on the interval \([-\pi, \pi]\). We have
\[
e^{-i\xi_\alpha x} = \sum_{k=-\infty}^{\infty} \tilde{P}_{n,k}[\sigma] e^{-ikx},
\]
(17)

where
\[
\tilde{P}_{n,k}[\sigma] = \frac{1}{2\pi} \int_{\pi} \pi e^{-i\xi_\alpha x} e^{ikx} dx = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-i(n-k)x} e^{-i\sigma x} dx.
\]

Let us denote
\[
P_{n,k}[\sigma] = (-1)^{n+k} \tilde{P}_{n,k}[\sigma].
\]
(18)

**Proposition 1.** _The three components of the nonlinear Fourier transform
\[
\mathcal{H}[q, r, \sigma] = \left(\alpha[q, r, \sigma], \beta[q, r, \sigma], \gamma[q, r, \sigma]\right)
\]
can be given by the formulæ
\[
\alpha[q, r, \sigma](n) = (e^{i\pi \sigma} - 1) + \sum_{k=-\infty}^{\infty} P_{n,k}[\sigma] \tilde{a}[g, r](k)
\]
\[
\beta[q, r, \sigma](n) = \sum_{k=-\infty}^{\infty} P_{n,k}[\sigma] \tilde{b}[g, r](k)
\]
\[
\gamma[q, r, \sigma](n) = \sum_{k=-\infty}^{\infty} P_{-\sigma, -n-k}[\sigma] \tilde{c}[g, r](k),
\]

where the coefficients \(P_{n,k}[\sigma]\) are given by (18), and \(\tilde{a}, \tilde{b}, \tilde{c}\) by (7)–(9), respectively.

**Proof.** We shall prove only the first equation. The other two are obviously proved in the same way. Let us denote the terms in the infinite sum (7) for \(\tilde{a}[g, r](z)\) by
\[
\tilde{a}_{2m}[g, r](z) = \int_{\Delta_{2m}} e^{-i\xi(z_1 + \ldots + z_{2m})} q(\xi_1)q(\xi_2)q(\xi_3) \ldots q(\xi_{2m}) d\xi.
\]

We introduce the new variable \(u = \xi_1 + \xi_2 + \ldots + \xi_{2m}\) in the above expression. This enables us to express \(\tilde{a}_{2m}[g, r](z)\) in the form
\[
\tilde{a}_{2m}[g, r](z) = \int_{0}^{2\pi} e^{-i\alpha u} T_{2m}(u) du,
\]

...
where
\[ T_{2m}(u) = \int_{D_{2m}(u)} q(\xi_1) r(\xi_2) q(\xi_3) \cdots r(\xi_{2m}) \, d\xi \]
and
\[ D_{2m}(u) = \{(\xi_1, \xi_2, \ldots, \xi_{2m}) \in \triangle_{2m}; \xi_1 - \xi_2 + \xi_3 - \ldots - \xi_{2m} = u\}. \]

If we denote
\[ T(u) = \sum_{m=1}^{\infty} T_{2m}(u), \]
we get
\[ \tilde{a}[q, r](z) = \int_0^{2\pi} e^{-i(\pi - u)} T(u) \, du. \]

Upon introduction of the new variable \( v = u - \pi \), the expression (11) for \( \tilde{a}[q, r](z) \) can now be rewritten in the form
\[ \tilde{a}[q, r](z) = \int_0^{2\pi} e^{-i(\pi - u)} T(u) \, du = \int_{-\pi}^{\pi} e^{-i\pi v} T(v + \pi) \, dv. \]

We now evaluate \( \tilde{a}[q, r](z) \) at \( z = zn \), and use (14) and (17) to obtain the desired result:
\[ \alpha[q, r, \sigma](n) = (e^{i\pi \sigma n} - 1) + (-1)^n \int_{-\pi}^{\pi} e^{-ikv} T(v + \pi) \, dv \]
\[ = (e^{i\pi \sigma n} - 1) + (-1)^n \sum_{k=-\infty}^{\infty} \tilde{P}[\sigma]_{n,k} e^{-ikv} T(v + \pi) \, dv \]
\[ = (e^{i\pi \sigma n} - 1) + (-1)^n \sum_{k=-\infty}^{\infty} \tilde{P}[\sigma]_{n,k} \left( -1 \right)^k \int_{-\pi}^{\pi} e^{-iku} T(u) \, du \]
\[ = (e^{i\pi \sigma n} - 1) + \sum_{k=-\infty}^{\infty} P[\sigma]_{n,k} \tilde{a}[q, r](k). \]

\[ \square \]

2.2. Local boundedness of \( \mathcal{H} \)

Our next goal is to prove that the target space of the operator \( \mathcal{H} \) is indeed the Hilbert space \( (l^2_Z)^3 \) and that \( \mathcal{H} \) is locally bounded.

Proposition 2. The nonlinear operator
\[ \mathcal{H} : L^2[0, 2\pi] \times L^2[0, 2\pi] \times l^2_Z \longrightarrow (l^2_Z)^3 \]
is locally bounded with respect to the usual \( L^2 \)-norms in the domain and in the target spaces.
Proof. From the definition given by (4) we see that it is enough to prove the local boundedness of the three components $\alpha[q, r, \sigma]$, $\beta[q, r, \sigma]$ and $\gamma[q, r, \sigma]$ of $\mathcal{H}$. Again, the proofs for the three components are virtually the same. Therefore, we shall only give the proof for $\alpha[q, r, \sigma]$. We shall construct a continuous function $B$ of three variables such that
\[
\|\alpha[q, r, \sigma]\|_P < B(\|q\|_{L^2}, \|r\|_{L^2}, \|\sigma\|_P).
\] (19)

First, we note that $(q, r, \sigma) \mapsto \alpha[q, r, \sigma]$ is a sum of two operators. The first summand
\[
(q, r, \sigma) \mapsto \{e^{i\pi \sigma n} - 1\}_{n \in \mathbb{Z}}
\]
depends only on $\sigma$ and it is clearly bounded with respect to the $L^2$ norms. It is not difficult to see that
\[
\|\{e^{i\pi \sigma n} - 1\}_{n \in \mathbb{Z}}\|^2 < (e^{\pi \|\sigma\|} - 1)^2.
\] (20)

It remains to prove that the second summand
\[
(q, r, \sigma) \mapsto \{ \sum_{k=-\infty}^{\infty} P[\sigma]_{\mu \lambda} \tilde{a}[q, r](k) \}_{n \in \mathbb{Z}}
\]
of $\alpha$ is also bounded. Let us denote
\[
\tilde{a}[q, r, \sigma](n) = \sum_{k=-\infty}^{\infty} P[\sigma]_{\mu \lambda} \tilde{a}[q, r](k).
\]
The two sequences $\{\tilde{a}[q, r, \sigma](n)\}_{n \in \mathbb{Z}}$ and $\{\tilde{a}[q, r](k)\}_{k \in \mathbb{Z}}$ can be thought of as two vectors with infinitely many components and $P[\sigma]_{\mu \lambda}$ is an infinite matrix. The map $\tilde{a}[q, r, \sigma]$ is the composition
\[
(q, r, \sigma) \mapsto (A[q, r], \sigma) \mapsto P[\sigma] \circ A[q, r] = \{ \sum_{k=-\infty}^{\infty} P[\sigma]_{\mu \lambda} \tilde{a}[q, r](k) \}_{n \in \mathbb{Z}},
\]
where we denoted
\[
A[q, r] = \{\tilde{a}[q, r](k)\}_{k \in \mathbb{Z}}.
\]
We shall prove that the operators
\[
A : L^2[0, 2\pi] \times L^2[0, 2\pi] \rightarrow l^2 \quad \text{and} \quad P : l^2 \times l^2 \rightarrow l^2
\]
are both locally bounded.

The boundedness of $A$ was essentially already proved in [23]. Nevertheless, we shall briefly repeat the argument for the sake of clarity and completeness. Above, we have seen that $\tilde{a}$ can be expressed in the form
\[
\{\tilde{a}[q, r](k)\}_{k \in \mathbb{Z}} = \int_0^{2\pi} e^{-iku} T(u) \, du \}_{k \in \mathbb{Z}} = F(T(u))
\]
where $F$ is the linear Fourier transform (Fourier series) of the function $T(u)$. Recall that one can consider $T$ to be an operator, defined on $L^2[0, 2\pi] \times L^2[0, 2\pi]$:
\[ T[q, r](u) = \sum_{m=1}^{\infty} T_{2m}[q, r](u) = \sum_{m=1}^{\infty} \int_{D_{2m}(u)} q(\xi_1) r(\xi_2) q(\xi_3) \cdots r(\xi_{2m}) \, d\xi. \]

Let us now introduce the functions
\[ S_{2m}(\xi_1, \ldots, \xi_{2m}) = \prod_{l=1}^{2m} (|q(\xi_l)| + |r(\xi_l)|). \]

For every \( m \in \mathbb{N} \) and every \( u \in [0, 2\pi] \), we have
\[ |T_{2m}[q, r](u)| \leq \int_{\Delta_{2m}} |q(\xi_1)| |r(\xi_2)| \cdots |r(\xi_{2m})| \, d\xi \leq \int_{\Delta_{2m}} S_{2m}(\xi_1, \ldots, \xi_{2m}) \, d\xi. \]

The advantage of the functions \( S_{2m} \) over \( T_{2m} \) is their invariance with respect to the permutations of the arguments \( \xi \). For every permutation \( \tau \in S_{2m} \), we have
\[ S_{2m}(\xi_{\tau(1)}, \xi_{\tau(2)}, \ldots, \xi_{\tau(2m)}) = S_{2m}(\xi_1, \xi_2, \ldots, \xi_{2m}). \]

Therefore,
\[ \int_{\Delta_{2m}} S_{2m}(\xi_1, \xi_2, \ldots, \xi_{2m}) \, d\xi = \frac{1}{(2m)!} \int_{\triangle_{2m}} S_{2m}(\xi_1, \xi_2, \ldots, \xi_{2m}) \, d\xi, \]
where \( \triangle_{2m} = [0, 2\pi]^{2m} \) is the hypercube. This gives
\[ \int_{\Delta_{2m}} S(\xi_1, \xi_2, \ldots, \xi_{2m}) \, d\xi = \frac{(||q||_1 + ||r||_1)^{2m}}{(2m)!}. \]

Summation over \( m \) gives
\[ |T[q, r](u)| \leq \cosh (||q||_1 + ||r||_1), \quad \text{for every } u \in [0, 2\pi]. \]

The functions \( q \) and \( r \) are elements of \( L^2[0, 2\pi] \), therefore they are also elements of \( L^1[0, 2\pi] \) and by the Cauchy–Schwarz inequality we have
\[ ||q||_1 \leq \sqrt{2\pi} \, ||q||_2, \quad ||r||_1 \leq \sqrt{2\pi} \, ||r||_2, \]
and finally
\[ ||T[q, r]||_2 \leq \sqrt{2\pi} \cosh (\sqrt{2\pi} (||q||_2 + ||r||_2)). \]

By Parseval’s theorem, the Fourier transform \( F \) is an \( L^2 \)-isometry, so we finally get
\[ ||A(q, r)||_2 = ||F(T[q, r])||_2 \leq \sqrt{2\pi} \cosh (\sqrt{2\pi} (||q||_2 + ||r||_2)). \]

The norm of \( A(q, r) \) is controlled by the norms of \( q \) and \( r \). In other words, the operator \( A \) is bounded.

In order to prove the local boundedness of \( \tilde{\alpha}[q, r, \sigma] \) with respect to \( (q, r) \) and with respect to \( \sigma \), we shall construct a continuous function \( M(x) \) of a real variable such that for every \( \sigma \in \mathbb{L}^2 \) and for every \( v \in \mathbb{L}^2 \), we shall have
\[ |\tilde{\alpha}[q, r, \sigma](x) - \tilde{\alpha}[q, r, \sigma](y)| \leq M(|x - y|). \]
\[ \|P[\sigma](v)\| \leq M(\|\sigma\|)\|v\|. \]  

(24)

For arbitrary \(v \in L^2_\mathbb{R}\), the \(n\)th term of the sequence \(P[\sigma](v)\) is given by

\[
\sum_{k=-\infty}^{\infty} P[\sigma]_{n,k} v(k) = \left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ikx} e^{-i\sigma x} \, dx \right\}_{k \in \mathbb{Z}}, \{v(k)\}_{k \in \mathbb{Z}}_2.
\]

Let the function \(f(x) \in L^2[-\pi, \pi]\) be given by \(f(x) = \sum_{k=-\infty}^{\infty} v(k) e^{-ikx}\). Then, the Parseval equation gives

\[
\left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ikx} e^{-i\sigma x} \, dx \right\}_{k \in \mathbb{Z}}, \{v(k)\}_{k \in \mathbb{Z}}_2 = \langle e^{-i(n+\sigma)x}, f(x) \rangle_{L^2} = \langle e^{-in\sigma x}, e^{i\sigma x} f(x) \rangle_{L^2} = \langle e^{-in\sigma x} - 1, e^{i\sigma x} f(x) \rangle_{L^2} + \langle 1, e^{i\sigma x} f(x) \rangle_{L^2}.
\]

Now, the Cauchy–Bunyakovsky–Schwarz inequality gives

\[
\| \sum_{k=-\infty}^{\infty} P[\sigma]_{n,k} v(k) \| \leq \|e^{-in\sigma x} - 1\| \cdot \|e^{i\sigma x} f(x)\| + |v(n)|
\]

\[
= \|e^{-in\sigma x} - 1\| \cdot \|f(x)\| + |v(n)|.
\]

(25)

For the \(L^2[-\pi, \pi]\)-norm of the function \(e^{-in\sigma x} - 1\), we have

\[
\|e^{-in\sigma x} - 1\| = \left\| \sum_{k=1}^{\infty} \frac{1}{k!} (-i\sigma x)^k \right\| \leq \sum_{k=1}^{\infty} \frac{1}{k!} \|\sigma x\|^k
\]

\[
= \sum_{k=1}^{\infty} \frac{\pi^k |\sigma|^k}{k!} \frac{1}{\sqrt{2k+1}} \leq \sum_{k=1}^{\infty} \frac{\pi^k |\sigma|^k}{k!} = e^{\pi |\sigma|} - 1.
\]

The triangle inequality, formula (25), Cauchy–Bunyakovsky–Schwartz, and the above calculation give

\[
\|P[\sigma](v)\|^2 = \sum_{n=-\infty}^{\infty} \| \sum_{k=-\infty}^{\infty} P[\sigma]_{n,k} v(k) \|^2
\]

\[
\leq \|f(x)\|^2 \sum_{n=-\infty}^{\infty} \|e^{-in\sigma x} - 1\|^2 + \|v\|^2
\]

\[
+ 2 \|v\| \| \{ e^{i|\sigma|} - 1 \}_{n \in \mathbb{Z}}, \{v(n)\}_{n \in \mathbb{Z}} \|_2
\]

\[
\leq \|v\|^2 \| \{ e^{i|\sigma|} - 1 \}_{n \in \mathbb{Z}} \|^2 + \|v\|^2
\]

\[
+ 2 \|v\|^2 \| \{ e^{i|\sigma|} - 1 \}_{n \in \mathbb{Z}} \|^2
\]

\[
= \|v\|^2 \left( \| \{ e^{i|\sigma|} - 1 \}_{n \in \mathbb{Z}} \|^2 + 1 \right)^2.
\]

Again, we have taken into account the fact that \(\|f(x)\| = \|v\|\). The positivity of the entries of the sequences \(\{|\sigma|^k\}_{n \in \mathbb{Z}}\) gives the estimate.
\[ \sum_{n=-\infty}^{\infty} |\sigma_n|^k < \left( \sum_{n=-\infty}^{\infty} |\sigma_n|^2 \right)^{\frac{k}{2}} = \|\sigma\|^k \]

for every \( k = 2, 3, \ldots \). So, expanding into the Taylor series and reversing the order of summations with respect to \( n \) and \( k \) yields

\[ \sum_{n=-\infty}^{\infty} \left( e^{n|\sigma_n|} - 1 \right)^2 = \sum_{n=-\infty}^{\infty} \left( \sum_{k=2}^{\infty} \frac{(2\pi|\sigma_n|)^k}{k!} - 2 \sum_{k=2}^{\infty} \frac{(\pi|\sigma_n|)^k}{k!} \right) < \sum_{k=2}^{\infty} \left( \frac{(2\pi\|\sigma\|)^k}{k!} + 2 \frac{(\pi\|\sigma\|)^k}{k!} \right) < e^{2\pi\|\sigma\|} + 2e^{\pi\|\sigma\|}. \]

The above estimates finally give us the \( l^2 \)-inequality

\[ \|P[\sigma](v)\|^2 < \|v\|^2 (\sqrt{e^{2\pi\|\sigma\|} + 2e^{\pi\|\sigma\|}} + 1)^2 < \|v\|^2 (e^{\pi\|\sigma\|} + 2)^2. \]  

(26)

Thus, the continuous function \( M(\|\sigma\|) \), needed in equation (24), exists. We can take

\[ M(\|\sigma\|) = e^{\pi\|\sigma\|} + 2. \]

Collecting the estimates (23) and (26) together, finally gives us

\[ \tilde{\alpha}[q, r, \sigma] = P[\sigma] \circ A[q, r] < \sqrt{2\pi}(e^{\pi\|\sigma\|} + 2) \cosh (\sqrt{2\pi}(\|q\| + \|r\|)). \]  

(27)

We now have the estimates for both summands of the map \( \alpha[q, r, \sigma] \). For the sought-for continuous function figuring in (19) we can take the sum of the right-hand sides of (20) and of (27). We get the estimate

\[ \|\alpha[q, r, \sigma]\| < (e^{\|\sigma\|} - 1) + \sqrt{2\pi}(e^{\pi\|\sigma\|} + 2) \cosh (\sqrt{2\pi}(\|q\| + \|r\|)). \]

This proves that the first component of the nonlinear Fourier transform \( H \) is indeed locally bounded with respect to all three variables \( q, r \) and \( \sigma \). Clearly, the same is true for the other two components, so our proposition is proved.

\[ \square \]

2.3. Analyticity and invertibility of \( H \)

To conclude the proof of the analyticity of \( H \), we have to prove its weak holomorphicity. Let the domain and the target spaces be equipped with their natural complex structures.

**Proposition 3.** The map

\[ H : L^2[0, 2\pi] \times L^2[0, 2\pi] \times l^2_k \longrightarrow (l^2_k)^3 \]

is weakly holomorphic at every point of the domain space.

**Proof.** Again, it suffices to prove the proposition for each component of \( H \) separately. The proofs for all the components are essentially the same, therefore we shall again focus only on the component \( \alpha \).
To prove the weak holomorphicity of \( \alpha \), we have to show that the complex function

\[
\zeta \mapsto \mathcal{L}\left(\alpha[q + \zeta t, r + \zeta s, \sigma + \zeta \tau]\right)
\]

is a holomorphic function of \( \zeta \) on some disc \( \Delta_\rho = \{ \zeta; |\zeta| < \rho \} \subseteq \mathbb{C} \), for arbitrary choices of \( (q, r), (t, s) \) in \( L^2[0, 2\pi] \times L^2[0, 2\pi] \) and \( \sigma, \tau \) in \( \mathbb{Z} \). It is a consequence of the Riesz representation theorem that every continuous linear functional on \( l^2 \) is a linear combination of the evaluation functionals \( \{ \alpha_n \}_{n \in \mathbb{Z}} \mapsto \alpha_n \), and that the coefficients of this linear combination must form a square-summable sequence. Therefore, it is enough to prove the holomorphicity of the map

\[
(\zeta, \eta) \mapsto f_{n_0}(\zeta, \eta) = \alpha[q + \zeta t, r + \zeta s, \sigma + \eta \tau](n_0)
\]

on \( \Delta_\rho \times \Delta_\rho \), for every \( n_0 \in \mathbb{Z} \). As a consequence, \( f_{n_0}(\zeta, \zeta) \) will also be holomorphic and this is what has to be proved.

From (10), (11) and (14) we get

\[
f_{n_0}(\zeta, \eta) = (e^{i\pi (\alpha_0 + \eta n_0)} - 1) + (-1)^n_0 \int_0^{2\pi} e^{-i(\alpha_0 + \eta n_0)}(x - \pi) \sum_{m=-\infty}^{\infty} T_{2m}[q + \zeta t, r + \zeta s](x) \, dx,
\]

where \( z_{n_0} = n_0 + \sigma n_0 \). The dependence of \( f_{n_0} \) on \( \eta \) is clearly holomorphic. Every \( T_{2m} \) is a polynomial of degree \( 2m \) of \( \zeta \). Thus, the series \( \sum_{m=-\infty}^{\infty} T_{2m} \) depends holomorphically on \( \zeta \), if it converges uniformly on every closed disc \( \Delta_\rho \) contained in \( \Delta_\rho \). From (21) and (22) we get

\[
|T_{2m}[q + \zeta t, r + \zeta s](x)| \leq \left( \|q\|_1 + \|r\|_1 + R(\|s\|_1 + \|\tau\|_1) \right) \frac{2m}{2m!},
\]

for every \( \zeta \in \Delta_\rho \), and so

\[
\sum_{m=1}^{\infty} |T_{2m}[q + \zeta t, r + \zeta s](x)| \leq \cosh (\|q\|_1 + \|r\|_1 + R(\|s\|_1 + \|\tau\|_1)).
\]

From this the uniform convergence of the series \( \sum_{m=1}^{\infty} T_{2m} \) on \( \Delta_\rho \) follows immediately (see [23]).

**Proposition 4.** The derivative \( D_0 H \) is invertible and the inverse is continuous.

**Proof.** We have proved that the map \( H \) is analytic at every point of the domain space. So, in particular, it is differentiable in the Fréchet sense and its strong derivative is equal to its directional derivative. We note that \( P[0] = I \) on \( l^2 \). From expressions (7)-(9) we get

\[
\frac{d}{dt}\bigg|_{t=0} \tilde{a}(eu, ev)(n) = 0, \quad \frac{d}{dt}\bigg|_{t=0} \tilde{b}(eu, ev)(n) = F[u](n), \quad \frac{d}{dt}\bigg|_{t=0} \tilde{c}(eu, ev)(n) = F[v'](-n),
\]

where \( F[u] = \{ F[u](n) \}_{n \in \mathbb{Z}} \) is the linear Fourier transform (series in our case) of the function \( u(x) \in L^2[0, 2\pi] \). A short calculation in which we use the above facts, definition (4) of \( H \), proposition 1, the Leibnitz rule, and the fact that \( P[0] = I \) shows that the expression for the derivative of \( H \) at the origin of \( L^2[0, 2\pi] \times l^2 \) is given by
\[ D_0 \mathcal{H}[u,v,\tau](n) = \left( \text{i}\pi\tau_n, F[u](n), F[v](-n) \right) \]

From Parseval’s equation it follows that the Fréchet derivative
\[ D_0 \mathcal{H} : L^2[0, 2\pi] \times L^2[0, 2\pi] \times \ell^2 \mathbb{Z} \longrightarrow (\ell^2 \mathbb{Z})^3 \]
is a linear isomorphism. \(\square\)

As explained at the beginning of this section, propositions 2–4 constitute the proof of theorem 1.

### 3. The inverse of the nonlinear Fourier transform

In this section, we will prove theorem 2. We shall provide a perturbation formula for the inverse
\[ \mathcal{H}^{-1} : V \subset (\ell^2 \mathbb{Z})^3 \longrightarrow L^2[0, 2\pi] \times L^2[0, 2\pi] \times \ell^2 \mathbb{Z}, \]
where \(V\) is a small enough neighbourhood of the origin in \((\ell^2 \mathbb{Z})^3\).

#### 3.1. The ubiquity of the Bell polynomials

We have seen above that Dyson’s series of the matrix \(L_{q,r}\) is a very important tool in the analysis of the nonlinear Fourier transform. Dyson’s series can be thought of as a non-commutative version of the exponential. Indeed, it is often called the path-ordered exponential. The generating function of the Bell polynomials is given by the exponentiation of the suitable power series. It is therefore not surprising that the Bell polynomials play an important role in the perturbative construction of the inverse of the nonlinear Fourier transform.

The exponential Bell polynomials \(B_{n,k}, n \in \mathbb{N}, k = 1, \ldots, n\) are defined as
\[
B_{n,k}(a_1, a_2, \ldots, a_{n-k+1}) = \frac{1}{k!} \sum_{r_1 + r_2 + \cdots + r_k = n} \left( \begin{array}{c} n \\ r_1, r_2, \ldots, r_k \end{array} \right) a_{r_1}a_{r_2}\cdots a_{r_k}. \tag{28}
\]
The above summation is taken over all ordered \(k\)-tuples \((r_1, r_2, \ldots, r_k)\), and the bracketed expression in the sum is the usual notation for the multinomial symbol. More information on Bell polynomials can be found, for example, in [19] or in [20]. (An alternative, equivalent definition can be found in [19], but for us, formula (28) is more useful.) The generating function for the Bell polynomials is
\[
G(x,s) = \exp \left( \sum_{j=1}^{\infty} \frac{s^j}{j!} a_j \right) = 1 + \sum_{m=1}^{\infty} \frac{s^m}{m!} \left( \sum_{k=1}^{m} x^k B_{m,k}(a_1, a_2, \ldots, a_{m-k+1}) \right). \tag{29}
\]

Returning to the nonlinear Fourier transform \(\mathcal{H}\), let \(s \mapsto \{\sigma_n(s)\}_{n \in \mathbb{Z}}\) be an analytic curve in \(\ell^2 \mathbb{Z}\) such that \(\{\sigma_n(0)\}_{n \in \mathbb{Z}} = 0\). We shall denote...
are related by the formula
\[ \sin \sigma_n(s) = \tilde{\sigma}_n(s) = \sum_{m=1}^{\infty} \frac{d^m}{m!} \tilde{\sigma}_n^{(m)}. \]

The following proposition follows from (29).

**Proposition 5.** For every integer \(n\) and for every positive integer \(m\), the terms \(S_n^{(m)}[\sigma] = (d^m/dx^m)|_{x=0} \exp (i \pi \sigma_n(s))\) and \(\tilde{\sigma}_n^{(m)} = i \pi \sigma_n^{(m)}\) are related by the formula
\[ S_n^{(m)}[\sigma] = \frac{d^m}{dx^m}|_{x=0} \exp (i \pi \sigma_n(s)) = \sum_{k=1}^{m} B_{m,k}(\tilde{\sigma}_n^{(1)}, \tilde{\sigma}_n^{(2)}, \ldots, \tilde{\sigma}_n^{(m)}). \] (30)

**Remark 1.** The well-known Arbogast—Faà di Bruno formula (see [19]), applied to the functions \(f(x) = \log(x)\) and \(g(x) = 1 + \sum_{m=1}^{\infty} \frac{e^x}{m!} S_n^{(m)}\), yields the ‘inverse’ of (30), namely
\[ \tilde{\sigma}_n^{(m)}[S] = \sum_{k=1}^{m} (-1)^k (k-1)! B_{m,k}(S_n^{(1)}, S_n^{(2)}, \ldots, S_n^{(m-k+1)}). \]

For later use, we can write equation (30) slightly more explicitly:
\[ S_n^{(m)} = \tilde{\sigma}_n^{(m)} + \sum_{k=2}^{m} B_{m,k}(\tilde{\sigma}_n^{(1)}, \tilde{\sigma}_n^{(2)}, \ldots, \tilde{\sigma}_n^{(m-k+1)}). \] (31)

A modification of the above reasoning provides us with formulae for the derivatives
\[ \frac{d^m}{dx^m}|_{x=0} P[\sigma(s)]_{n,k} = P[\sigma_n^{(m)}]_{n,k} \]
of the infinite matrices \(P[\sigma(s)]_{n,k}\). These derivatives will also be needed below.

**Proposition 6.** Let the Toeplitz matrices \(M_{n,k}^{(j)}\) be given by
\[ M_{n,k}^{(j)} = \frac{(-1)^{n+k}}{2\pi} \int_{-\pi}^{\pi} x^j e^{-i(n-k)x} \, dx. \]
The derivatives of the matrix \(P[\sigma(s)]\) are given by
\[ P[\sigma_n^{(m)}] = \sum_{j=1}^{m} B_{m,j}(-i\sigma_n^{(1)}, -i\sigma_n^{(2)}, \ldots, -i\sigma_n^{(m-j+1)}) M_{n,k}^{(j)}. \] (32)

**Proof.** By (29) we have
\[ e^{-i\sigma_n(s)x} = 1 + \sum_{m=1}^{\infty} \frac{s^m}{m!} \left( \sum_{j=1}^{m} x^j B_{m,j}(-i\sigma_n^{(1)}, -i\sigma_n^{(2)}, \ldots, -i\sigma_n^{(m-j+1)}) \right). \]

From now on, we shall use the abbreviated notation
\[ B_{m,j}[\sigma](n) = B_{m,j}(i\sigma_n^{(1)}, i\sigma_n^{(2)}, \ldots, i\sigma_n^{(m-j+1)}). \]
The \((n, k)\)th entry of the infinite matrix \(P[\sigma(s)]\) is the \((n - k)\)th Fourier coefficient of the function \(\exp(-i\sigma(s)x)\). This gives
These three sequences are the entries of
\[ i - \sum \] in terms of the
shall now describe. Let


\[ P[\sigma(s)]_{n,k} = (\pm 1)^{n+k} \tilde{P}[\sigma(s)]_{n,k} = \frac{(-1)^{n+k}}{2\pi} \int_{-\pi}^{\pi} e^{-i\sigma(s)x} e^{-i(n-k)x} \, dx \]

\[ = (\pm 1)^{n+k} \delta_{n,k} + \sum_{m=1}^{\infty} \frac{s^m}{m!} \sum_{j=1}^{m} B_{m,j}[\sigma]_{n}(j) \left( -\frac{1}{2\pi} \right)^{n+k} \int_{-\pi}^{\pi} x^j e^{-i(n-k)x} \, dx \]

\[ = \delta_{n,k} + \sum_{m=1}^{\infty} \frac{s^m}{m!} \sum_{j=1}^{m} B_{m,j}[\sigma]_{n} M^{(j)}_{n,k}. \]

The \( m \)th derivation of the above expression gives (32).

Let now \( s \mapsto q(x; s) \) and \( s \mapsto r(x; s) \) be two analytic curves in \( L^2[0, 2\pi] \), defined on \( \Delta_p \), both starting at the origin. Let

\[ q(x; s) = \sum_{m=1}^{\infty} \frac{s^m}{m!} q^{(m)}(x), \quad r(x; s) = \sum_{m=1}^{\infty} \frac{s^m}{m!} r^{(m)}(x) \]

be their Taylor expansions. We shall derive formulae for the terms in the Taylor expansions of \( \hat{a}(q(s), r(s))[n] \), \( \hat{b}(q(s), r(s))[n] \) and \( \hat{c}(q(s), r(s))[n] \). These three sequences are the entries of the Taylor expansion of the holonomy of

\[ L_{q(s),r(s)}^{G}(x; n) = \sum_{m=1}^{\infty} \frac{s^m}{m!} L_{n}^{(m)}(x), \]

where

\[ L_{n}^{(m)}(x) = \begin{pmatrix} 0 & e^{-i\sigma x^{(m)}(x)} \\ e^{i\sigma x^{(m)}(x)} & 0 \end{pmatrix}. \]

The holonomy

\[ \text{Hol}^{G}(q(s), r(s))[n] = \sum_{m=1}^{\infty} \frac{s^m}{m!} H_{n}^{(m)}[q, r] \]

is obtained by Dyson’s series of \( L_{q(s),r(s)}^{G}(x; n) \). Therefore, the expressions for \( H_{n}^{(m)} \) in terms of \( L_{n}^{(m)}(x) \) will be provided by a non-commutative version of the Bell polynomials which we shall now describe. Let \( \hat{a} = (a_1, a_2, a_3, \ldots) \) be an infinite vector of indeterminates. The defining equation (29) for the Bell polynomials \( B_{n,k}(\hat{a}) = B_{n,k}(a_1, a_2, \ldots, a_{n-k+1}) \) is equivalent to the sequence of matrix equations

\[ \text{Exp}(tA_N) = I + B_N, \quad N \in \mathbb{N} \]

where \( A_N \) and \( B_N \) are strictly upper triangular Toeplitz matrices whose nonzero entries are given by

\[ (A_N)_{ij} = \frac{s^{j-i}}{(j-i)!} a_{j-i}, \quad (B_N)_{ij} = \frac{s^{j-i}}{(j-i)!} \sum_{k=1}^{j-i} \frac{j-i}{k} B_{j-i-k}(\hat{a}). \]

A straightforward check shows that the terms \( H_{n}^{(m)} \) of the \( N \)-jet \( I + \sum_{m=1}^{\infty} \frac{s^m}{m!} H_{n}^{(m)} \) of \( \text{Hol}^{G}(q(s), r(s))[n] \) are the entries of \( \text{Hol}^{G}_{N}[q, r](n) = \Phi_{N}(n)(2\pi) \), where \( \Phi_{N}(n)(x) \) is the solution of the initial problem
\((\Phi_N)(n)_j(x) = \mathcal{L}_N[q,r](n) \cdot \Phi_N(n)(x), \quad \Phi_N(n)(0) = I,\)

with

\[
\mathcal{L}_N[q,r](n) = \begin{pmatrix}
0 & \frac{1}{\pi} L_n^{(1)} & \frac{1}{\pi} L_n^{(2)} & \cdots & \frac{1}{\pi} L_n^{(N)} \\
0 & 0 & \frac{1}{\pi} L_n^{(1)} & \cdots & \frac{1}{\pi} L_n^{(N-1)} \\
0 & 0 & 0 & \cdots & \frac{1}{\pi} L_n^{(N-2)} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & \frac{1}{\pi} L_n^{(1)} \\
0 & 0 & 0 & \cdots & 0
\end{pmatrix}
\]

and

\[
\Phi_N(n) = \begin{pmatrix}
1 & \frac{1}{\pi} \Phi_n^{(1)} & \frac{1}{\pi} \Phi_n^{(2)} & \cdots & \frac{1}{\pi} \Phi_n^{(N)} \\
0 & 1 & \frac{1}{\pi} \Phi_n^{(1)} & \cdots & \frac{1}{\pi} \Phi_n^{(N-1)} \\
0 & 0 & 1 & \cdots & \frac{1}{\pi} \Phi_n^{(N-2)} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & \frac{1}{\pi} \Phi_n^{(1)} \\
0 & 0 & 0 & \cdots & 1
\end{pmatrix}
\]

The matrix \(\mathcal{L}_N(n) = \mathcal{L}_N[q,r](n)\) is nilpotent. Therefore, its path-ordered exponential \(\text{Hol}_N[q,r](n)\) has only \(N\) terms and is very similar to the triangular Toeplitz matrix \(I + B_N\) which defines the Bell polynomials. We have

\[
H_n^{(m)} = \sum_{k=1}^{m} \hat{B}_{m,k}[q,r](n),
\]

where

\[
\hat{B}_{m,k}[q,r](n) = \sum_{j_1+\cdots+j_k=m} \begin{pmatrix} j_1, j_2, \ldots, j_k \end{pmatrix} \int_{\Delta_k} L_n^{(j_1)}(\xi_1) \cdot L_n^{(j_2)}(\xi_2) \cdots L_n^{(j_k)}(\xi_k) \, d\xi.
\]

(34)

We observe that for every \(m\) we have

\[
\hat{B}_{m,1}[q,r](n) = \int_0^{2\pi} L_n^{(m)}(\xi) \, d\xi = \begin{pmatrix} 0 & 2\pi e^{-i\xi} q^{(m)}(\xi) \\ 0 & 2\pi \end{pmatrix}
\]

\[
= \begin{pmatrix} 0 & F[q^{(m)} (n)] \\ F[q^{(m)} (-n)] & 0 \end{pmatrix}
\]

(35)

Formula (34) is very similar to formula (28) for the Bell polynomial. The essential differences are the presence of the integration in (34), and, even more importantly, the non-commutativity of the variables \(L_n^{(j)}(\xi)\). The occurrence of the factor \(1/k!\) in (28) is mirrored by the integration over the ordered simplex \(\Delta_k\) whose volume is \((2\pi)^k/k!\). The order of the variables \(L_n^{(j)}(\xi)\) is essential. In other words, the functions \(\hat{B}_{m,k}\) are not symmetric in their variables \(L_n^{(j)}(\xi)\).
Remark 2. In the case where all the matrices \( L_n^{(b)}(\xi_j) \) (for varying \( j \) and \( \xi_j \)) commute, we have
\[
\tilde{B}_{m,k}(L_n(n)) = B_{m,k} \left( \int_0^{2\pi} L_n^{(1)}(\xi) \, d\xi \right) \int_0^{2\pi} L_n^{(2)}(\xi) \, d\xi \ldots \int_0^{2\pi} L_n^{(m-1)}(\xi) \, d\xi,
\]
where the polynomial \( B_{m,k} \) is defined by (28).

3.2. The iterative scheme

In terms of the matrix-valued nonlinear Fourier transform \( F \), the result of proposition 1 in section 2 can be rewritten in the form
\[
F[q, r, \sigma](n) = \begin{pmatrix} e^{i\sigma_n} & 0 \\ 0 & e^{-i\sigma_n} \end{pmatrix} + \sum_{k=-\infty}^{\infty} \begin{pmatrix} P[\sigma]_{n,k} & 0 \\ 0 & P[-\sigma]_{-n,-k} \end{pmatrix} \begin{pmatrix} \tilde{a}[q, r](k) & \tilde{b}[q, r](k) \\ \tilde{c}[q, r](k) & \tilde{d}[q, r](k) \end{pmatrix}.
\]

We shall abbreviate the above notation and write
\[
F[q, r, \sigma] = \text{Exp}[\tilde{P}[\sigma]] \cdot \mathcal{F}_n[q, r],
\]
where
\[
\tilde{P}[\sigma]_{n,k} = \begin{pmatrix} P[\sigma]_{n,k} & 0 \\ 0 & P[-\sigma]_{-n,-k} \end{pmatrix}, \quad \text{and} \quad \mathcal{F}_n[q, r](k) = \begin{pmatrix} \tilde{a}[q, r](k) & \tilde{b}[q, r](k) \\ \tilde{c}[q, r](k) & \tilde{d}[q, r](k) \end{pmatrix}.
\]

Let now
\[
s \mapsto \{ \sigma_n(s) \}_{n \in \mathbb{Z}} = \sum_{m=1}^{\infty} \frac{s^m}{m!} (\sigma_n^{(m)})_{n \in \mathbb{Z}}, \quad s \mapsto q(s) = \sum_{m=1}^{\infty} \frac{s^m}{m!} q^{(m)}, \quad s \mapsto r(s) = \sum_{m=1}^{\infty} \frac{s^m}{m!} r^{(m)}
\]
be analytic curves in \( \mathbb{C}^2 \) and in \( L^2[0, 2\pi] \) which all start at respective origins. We have proved that \( F[q, r, \sigma] \) is an analytic map with respect to all three variables \( q, r, \) and \( \sigma \). Therefore, the path
\[
s \mapsto F[q(s), r(s), \sigma(s)] - I = \begin{pmatrix} \alpha[q(s), r(s), \sigma(s)] & \beta[q(s), r(s), \sigma(s)] \\ \gamma[q(s), r(s), \sigma(s)] & \delta[q(s), r(s), \sigma(s)] \end{pmatrix}
\]
is also an analytic map into \( (\mathbb{C}^2)^4 \). In particular, for every \( n \in \mathbb{Z} \), the path
\[
s \mapsto F[q(s), r(s), \sigma(s)](n) = (-1)^n \text{Hol}[q(s), r(s)](n + \sigma_n(s))
\]
is an analytic map from a suitably chosen interval around \( 0 \in \mathbb{R} \) into the space of the unimodular \( 2 \times 2 \)-matrices.

Let now
\[
s \mapsto \{ A_n \}_{n \in \mathbb{Z}}(s) = \left\{ \begin{pmatrix} 1 + A_n(s) & B_n(s) \\ C_n(s) & 1 + D_n(s) \end{pmatrix} \right\}_{n \in \mathbb{Z}}
\]
\[
= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \sum_{m=1}^{\infty} \frac{s^m}{m!} \begin{pmatrix} A_n^{(m)} & B_n^{(m)} \\ C_n^{(m)} & D_n^{(m)} \end{pmatrix}_{n \in \mathbb{Z}}
\]

be an analytic path in the discrete loop group \( LSL(2, \mathbb{C}) \). By this we mean that all the curves
\( s \mapsto \{ A_s \}_{n \in \mathbb{Z}}(s), \quad s \mapsto \{ B_s \}_{n \in \mathbb{Z}}(s), \quad s \mapsto \{ C_s \}_{n \in \mathbb{Z}}(s), \quad s \mapsto \{ D_s \}_{n \in \mathbb{Z}}(s) \)
are analytic curves in \( \ell^2_{\mathbb{Z}} \), and that they all start at the origin of \( \ell^2_{\mathbb{Z}} \). The entire \( \{ A_s \}_{n \in \mathbb{Z}}(s) \) is determined by the curves \( \{ A_s \}_{n \in \mathbb{Z}}(s), \{ B_s \}_{n \in \mathbb{Z}}(s), \) and \( \{ C_s \}_{n \in \mathbb{Z}}(s) \) because of unimodularity.

Our aim is to provide the inverse image
\[
F^{-1}(\{ A \}_{n \in \mathbb{Z}}(s)) = \left( q(s), r(s), \sigma(s) \right).
\]
We have proved that the map
\[
F^{-1} : LSL(2, \mathbb{C}) \longrightarrow L^2[0, 2\pi] \times L^2[0, 2\pi] \times \ell^2_{\mathbb{Z}}
\]
is analytic. Therefore, there exist unique convergent series
\[
q(x; s) = \sum_{m=1}^{\infty} \frac{x^m}{m!} q^{(m)}(x), \quad r(x; s) = \sum_{m=1}^{\infty} \frac{x^m}{m!} r^{(m)}(x)
\]
(36)
\[
\{ \sigma_n \}_{n \in \mathbb{Z}}(s) = \sum_{m=1}^{\infty} \frac{s^m}{m!} \{ \sigma_n^{(m)} \}_{n \in \mathbb{Z}}
\]
(37)
in \( L^2[0, 2\pi] \times L^2[0, 2\pi] \) and in \( \ell^2_{\mathbb{Z}} \) such that
\[
F[q(x; s), r(x; s), \{ \sigma_n \}_{n \in \mathbb{Z}}(s)] = A(s).
\]
(38)

Our iterative scheme for the calculation of the terms \( q^{(m)}(x), r^{(m)}(x), \) and \( \sigma_n^{(m)}(x) \), for every integer \( m \), is constructed as follows. Deriving equation (38) \( m \) times with respect to \( s \) and evaluating at \( s = 0 \) gives
\[
\tilde{S}(m)[\sigma] + \sum_{j=1}^{m} \binom{m}{j} \tilde{P}[\sigma]^{(m-j)} \cdot F_{\mathbb{Z}}[q, r]^{(j)} = A^{(m)} = \frac{d^m}{ds^m} |_{s=0} A(s),
\]
where \( \tilde{S}(m)[\sigma] = \{ \tilde{S}_n^{(m)} \}_{n \in \mathbb{Z}} \) and
\[
\tilde{S}_n^{(m)} = \begin{pmatrix} S_n^{(m)}[\sigma] & 0 \\ 0 & S_n^{(m)}[-\sigma] \end{pmatrix} = \begin{pmatrix} \tilde{\sigma}_n^{(m)} + \sum_{d=2}^{m} B_{m,d}[\pi \sigma] & 0 \\ 0 & -\tilde{\sigma}_n^{(m)} + \sum_{d=2}^{m} B_{m,d}[-\pi \sigma] \end{pmatrix}
\]
as noted in (31). Similarly, \( \tilde{P}[\sigma]^{(k)} \) denotes the \( k \)th derivative of \( \tilde{P}[\sigma(s)] \) at \( s = 0 \) and \( F_{\mathbb{Z}}[q, r]^{(k)} \) the \( k \)th derivative of \( F_{\mathbb{Z}}[q(s), r(s)] \) at \( s = 0 \). We can separate the terms which contain the highest order derivatives in the unknowns \( q, r, \sigma \) and obtain the equation
\[
\tilde{S}(m)[\sigma] + F_{\mathbb{Z}}^{(m)}[q, r] = A^{(m)} - \sum_{j=1}^{m} \binom{m}{j} \tilde{P}[\sigma]^{(m-j)} \cdot F_{\mathbb{Z}}[q, r]^{(j)}.
\]
(39)
Equations (33) and (34) give
\[
F_{\mathbb{Z}}[q, r]^{(m)}(n) = H_n^{(m)} = \sum_{l=1}^{m} B_{m,l}[q, r](n), \quad \text{for every } n \in \mathbb{N}.
\]
In particular, formula (35) for $\widehat{B}[q, r]_{m, 1}(n)$ gives
\[
\mathcal{F}_{Z}[q, r]^{(m)}(n) = \begin{pmatrix} 0 & F[q^{(m)}](n) \\ F[r^{(m)}](-n) & 0 \end{pmatrix} + \sum_{j=2}^{m} \widehat{B}_{m,j}[q, r](n).
\]

From formula (32) in proposition 6, we get
\[
\widehat{P}[^{(l)}\sigma]_{\sigma, n, k} = \sum_{l=1}^{j} \left( \begin{array}{cc} (B_{l}-\sigma)(n) \cdot M_{n,k}^{(l)} & 0 \\ 0 & (B_{l})(\sigma)(n) \cdot M_{n,-n-k}^{(l)} \end{array} \right).
\]

Plugging the above expressions into equation (39) gives
\[
\iota\pi\sigma^{(m)} + \widehat{B}_{m,1}[q, r] = A^{(m)} - \sum_{l=2}^{m} \widehat{B}_{m,l}[\sigma] + \sum_{l=2}^{m} \widehat{B}_{m,l}[q, r] - \sum_{j=1}^{m-1} \sum_{l=1}^{j} \widehat{P}[\sigma]^{m-j} \cdot \sum_{l=1}^{j} \widehat{B}_{l}[q, r],
\]

where
\[
\iota\pi\sigma^{(m)} = \begin{pmatrix} \iota\pi\sigma^{(m)}_{n, \in Z} & 0 \\ 0 & -\iota\pi\sigma^{(m)}_{n, \in Z} \end{pmatrix}, \quad \widehat{B}_{m,l}[\sigma] = \begin{pmatrix} B_{m,l}[\sigma] & 0 \\ 0 & B_{m,l}[-\sigma] \end{pmatrix}.
\]

Now we shall divide the $2 \times 2$-matrix equation (40) into its diagonal part and its anti-diagonal part. In the diagonal part, both entries yield essentially the same result. Therefore, we can keep only the top-left entry which will give us the expression for $\sigma^{(m)}$. The anti-diagonal part consists of two different equations which yield the expressions for the functions $q^{(m)}$ and $r^{(m)}$. If we denote the right-hand side (without $A^{(m)}$) of equation (40) by $B^{(m)}[q, r, \sigma]$, then its $(i,j)$-entries $B_{i,j}^{(m)}[q, r, \sigma]$ are given by
\[
B_{1,1}^{(m)}[q, r, \sigma] = -\sum_{d=2}^{m} B_{m,d}[\sigma] - \sum_{d=2}^{m} (\widehat{B}_{m,d}[q, r])_{1,1} - \sum_{j=1}^{m-1} \sum_{l=1}^{j} B_{m-j,l}[-\sigma] \left( \sum_{d=1}^{j} M^{(l)} \cdot (\widehat{B}_{d,l}[q, r])_{1,1} \right)
\]

(41)

\[
B_{1,2}^{(m)}[q, r, \sigma] = -\sum_{d=2}^{m} (\widehat{B}_{m,d}[q, r])_{1,2} - \sum_{j=1}^{m-1} \sum_{l=1}^{j} B_{m-j,l}[-\sigma] \left( \sum_{d=1}^{j} M^{(l)} \cdot (\widehat{B}_{d,l}[q, r])_{1,2} \right)
\]

(42)

\[
B_{2,1}^{(m)}[q, r, \sigma] = -\sum_{d=2}^{m} (\widehat{B}_{m,d}[q, r])_{2,1} - \sum_{j=1}^{m-1} \sum_{l=1}^{j} B_{m-j,l}[\sigma] \left( \sum_{d=1}^{j} M^{(l)} \cdot (\widehat{B}_{d,l}[q, r])_{2,1} \right)
\]

(43)

and $\widehat{M}^{(l)} = M^{(l)}_{n,-n-k}$. Above, $M^{(l)} \cdot (\widehat{B}_{d,l}[q, r])_{n,s}$ denotes the product of the infinite matrix $M^{(l)}$ and the element of $P^{(l)}$. As before, let $F^{-1}$ denote the inverse of the linear Fourier transform
\[
F^{-1}(\{a_{n}\}_{n \in \mathbb{Z}}) = \sum_{n=-\infty}^{\infty} a_{n} e^{i\pi n}.
\]
Thus, by the analytic continuation the map 

\[ \text{iterative scheme cannot be expected to converge.} \]

Therefore, system (44) with the right-hand sides given by (41) and of the mapping theorem. The operator

\[ \text{discussion, therefore we shall only indicate the issues involved.} \]

The fundamental tool which ensures the convergence of (45) is the holomorphic inverse

\[ \text{matrix equation (40) is equivalent to the system of scalar equations} \]

\[
\begin{align*}
\sigma_n^{(m)} &= -\frac{i}{π} \left( A_n^{(m)} + B_n^{(m)}[q, r, σ](n) \right) \\
q^{(m)}(x) &= F^{-1}\left( \left\{ B_n^{(m)}[q, r, σ](n) \right\}_{n ∈ ℤ} \right) \\
r^{(m)}(x) &= F^{-1}\left( \left\{ C^{(m)}_{-n}[q, r, σ](-n) \right\}_{n ∈ ℤ} \right).
\end{align*}
\]

(44)

The definition of the Bell polynomial implies that

\[ \text{Therefore, system (44) with the right-hand sides given by (41)–(43) provides the iterative expressions of the terms} \]

\[ \text{and of the Taylor series (36) and (37).} \]

This concludes the proof of theorem 2.

An important and natural question arises: what are the limitations on the data \( \{ M(n) \}_{n ∈ ℤ} \in \text{LSL}(2; ℂ) \) under which the constructions described above make sense. If we restrict ourselves to considering lines instead of arbitrary analytic curves, we can rephrase the question more concretely. Let

\[ \{ M(n) \}_{n ∈ ℤ} = \left\{ \left( \begin{array}{cc} 1 + α_n & β_n \\ γ_n & 1 + δ_n \end{array} \right) \right\}_{n ∈ ℤ} \in \text{LSL}(2; ℂ) \]

and let us again replace \( \text{LSL}(2; ℂ) \) by \((L^2_δ)^3\). We shall concentrate on a single direction in \((L^2_δ)^3\).

By the holomorphic version of the inverse mapping theorem for Hilbert spaces, there exists \( \{ (α_n, β_n, γ_n) \}_{n ∈ ℤ} \in (L^2_δ)^3 \) with \( \| (α_n, β_n, γ_n) \|_{n ∈ ℤ} > 0 \) such that the series

\[ \mathcal{H}^{-1}\left\{ \{ s α_n, s β_n, s γ_n \} \right\}_{n ∈ ℤ} = Q(s) = \sum_{m=1}^{∞} \frac{s^m}{m!} \left( q^{(m)}(x), r^{(m)}(x), \{ σ_n^{(m)} \}_{n ∈ ℤ} \right) \]

converges for every \( s ∈ Δ_δ \), where \( Δ_δ \) is the unit disc in \( ℂ \). What is the actual convergence radius of (45)? An exhaustive and thorough answer to this question would demand a long discussion, therefore we shall only indicate the issues involved.

The fundamental tool which ensures the convergence of (45) is the holomorphic inverse mapping theorem. The operator \( \mathcal{H} \) satisfies the hypotheses of this theorem in the vicinity of the origin. Therefore \( \mathcal{H}^{-1} \) is analytic on some neighbourhood \( V \) of \( 0 \in (L^2_δ)^3 \), and the composition of an analytic curve in \( V \) and of the mapping \( \mathcal{H}^{-1} \) is an analytic curve in \((L^2_δ[0, 2π])^3 \times L^2_δ\). The key hypothesis is the invertibility of the derivative \( \mathcal{D}_H \) of \( H \) at the origin. More generally, \( \mathcal{H}^{-1} \) is invertible in a vicinity of every point \( (q_0, r_0, σ_0) \in (L^2_δ)^3 \) such that \( \mathcal{D}_{(q_0, r_0, σ_0)} \mathcal{H} \) is invertible.

Thus, by the analytic continuation the map \( s ↦ (q^{(m)}(s), r^{(m)}(s), σ^{(m)}(s)) \) on \( Δ_δ \), if for every \( s_0 ∈ Δ_δ \) the derivative \( \mathcal{D}_{(q, r, σ)} \mathcal{H} \) at \( (q_0, r_0, σ_0) = \mathcal{Q}(s_0) \) is invertible. Consider now the curve \( s ↦ \mathcal{Q}(s) = (q(s), r(s), σ(s)) \). Let the derivative \( \mathcal{D}_{(q, r, σ)} \mathcal{H} \) be singular at \( (q_ε, r_ε, σ_ε) = \mathcal{Q}(s_c) \) and let \( s_c \) have the smallest modulus among all such \( s \). Then \( \rho_c = |s_c| \) is the convergence radius of \( \mathcal{H}^{-1} \{ \{ s α_n, s β_n, s γ_n \} \}_{n ∈ ℤ} \).

Beyond this radius our iterative scheme cannot be expected to converge.

Let

\[ \{ T_n \}_{n ∈ ℤ} = \left\{ \begin{array}{c} \frac{t_2(x)}{t_1(x)} \\
\frac{t_1(x)}{t_2(x)} \end{array} \right\}_{n ∈ ℤ} \]
and let us denote the derivative of $\mathcal{F}$ at $(q, r, \sigma)$ in the direction $(t_1, t_2, \tau)$ by
\[
(D_{q,r,\sigma})\mathcal{F}(t_1, t_2, \tau)(n) = \left. \frac{d}{du} \right|_{u=0} \mathcal{F}[q + ut_1, r + ut_2, \sigma + ut\tau](n).
\]
A simple calculation shows
\[
(D_{q,r,\sigma})\mathcal{F}(t_1, t_2, \tau)(n) = \mathcal{F}[q, r, \sigma](n) \cdot \int_0^{2\pi} \Phi^{-1}(x;n) \cdot T_n(x) \cdot \Phi(x;n) \, dx,
\]  
(46)

where
\[
\Phi_e(x;n) = L_{q,r}(x; n + \sigma_n) \cdot \Phi(x;n), \quad \Phi(0;n) = I.
\]

For details of a similar case, see page 12 of [23]. Let us denote
\[
\mathcal{F}[q + ut_1, r + ut_2, \sigma + ut\tau](n) = \begin{pmatrix}
1 + \alpha_e(u) & \beta_e(u) \\
\gamma_e(u) & 1 + \delta_e(u)
\end{pmatrix} = \begin{pmatrix}
a_e(u) & b_e(u) \\
c_e(u) & d_e(u)
\end{pmatrix}.
\]

The derivation of the identity
\[
\det \left( \mathcal{F}[q + ut_1, r + ut_2, \sigma + ut\tau](n) \right) = a_e(u)d_e(u) - b_e(u)c_e(u) = 1
\]
with respect to $u$ and the evaluation at $u = 0$ tells us that $(\alpha'_e(0), \beta'_e(0), \gamma'_e(0)) = (0, 0, 0)$ if and only if $(a'_e(0), b'_e(0), c'_e(0), d'_e(0)) = (0, 0, 0, 0)$, provided $a_e(0) \neq 0$. This means that the derivatives of $\mathcal{F}$ and of $\mathcal{H}$ are singular simultaneously. Since $\mathcal{F}[q, r, \sigma](n)$ are always invertible, we see from (46) that
\[
(D_{q,r,\sigma})\mathcal{F}(t_1, t_2, \tau)(n) = 0, \quad \text{if and only if} \quad \int_0^{2\pi} \Phi^{-1}(x;n) \cdot T(x) \cdot \Phi(x;n) \, dx = 0.
\]

Let us denote
\[
\Phi_e(x;n) = \begin{pmatrix}
\phi_{11}(x;n) & \phi_{12}(x;n) \\
\phi_{21}(x;n) & \phi_{22}(x;n)
\end{pmatrix}.
\]

A direct calculation now shows that the triple $(q, r, \sigma)$ is a critical triple $(q_e, r_e, \sigma_e)$, if for the solutions $\Phi_e(x;n)$ of the initial problems
\[
(\Phi_e)_e(n) = L_{q_e,r_e}(x; n + (\sigma_e)_n) \cdot \Phi_e(x;n), \quad \Phi_e(0;n) = I, \quad n \in \mathbb{Z}
\]
there exists a triple $(t_1(x), t_2(x), \tau) \in (L^2[0, 2\pi])^2 \times L^2$ such that
\[
\int_0^{2\pi} \begin{pmatrix}
\phi_{11}(x) & \phi_{12}(x) \\
2\phi_{21}(x) & -\phi_{22}(x)
\end{pmatrix} \begin{pmatrix}
\phi_{11}(x) & \phi_{12}(x) \\
2\phi_{21}(x) & -\phi_{22}(x)
\end{pmatrix}^T (x;n) \cdot \begin{pmatrix}
t_1(x) \\
t_2(x)
\end{pmatrix} \, dx = \begin{pmatrix}
0 \\
0
\end{pmatrix}
\]

holds for every $n \in \mathbb{Z}$.

4. Application

In this section, we shall briefly indicate the possible use of the nonlinear Fourier modes. In section 4.1 we give the explicit calculation of a nonlinear Fourier mode up to the third order
in the perturbation parameter. We shall also construct a nonlinear superposition of two Fourier modes up to the third order. The third order term is the first nontrivial one and it allows for the illustration of some of the issues involved. In section 4.2 we shall apply our expressions for the inverse nonlinear Fourier transform to a problem of communication through a nonlinear medium. We intend to address the problem in greater generality and in more detail in a separate paper.

4.1. Appearance of the Bernoulli polynomials

Let the nonlinear Fourier spectrum be given by

\[
\{A_n(s)\}_{n \in \mathbb{Z}} = \sum_{m=1}^{\infty} \frac{s^m}{m!} A_d^{(m)} \delta_{n,d},
\]

\[
\{B_n(s)\}_{n \in \mathbb{Z}} = \sum_{m=1}^{\infty} \frac{s^m}{m!} B_d^{(m)} \delta_{n,d}, \quad \{C_n(s)\}_{n \in \mathbb{Z}} = \sum_{m=1}^{\infty} \frac{s^m}{m!} C_d^{(m)} \delta_{n,d},
\]

where \(\delta_{n,d}\) is the Kronecker delta. Formulae (41)–(44) are trivial for \(s=1\). They give

\[
\sigma_n^{(1)} = -\frac{i}{\pi} A_d^{(1)} \delta_{n,d}, \quad q^{(1)}(x) = B_d^{(1)} e^{idx}, \quad r^{(1)}(x) = C_d^{(1)} e^{-idx}.
\]

The \(s^2\) coefficients are

\[
\sigma_n^{(2)} = -\frac{2B_d^{(1)} C_d^{(1)}}{n - d}, \quad \text{for } n \neq d, \quad \text{and } \sigma_d^{(2)} = -\frac{i}{\pi} (A_d^{(2)} - (A_d^{(1)})^2 - 2\pi i (B_d^{(1)} C_d^{(1)}))
\]

and again

\[
q^{(2)}(x) = B_d^{(2)} e^{idx}, \quad r^{(2)}(x) = C_d^{(2)} e^{-idx}.
\]

Now we can use (41)–(44) to calculate \(q^{(3)}\) and \(r^{(3)}\). For \(n \neq d\), we get

\[
q^{(3)}(x) = Q_3(A_d, B_d, C_d) e^{idx} + Mq \sum_{n \neq d} \frac{e^{inx}}{(n - d)^2},
\]

where

\[
Mq = (4\pi^3 + 2\pi)(B_d^{(1)})^2 C_d^{(1)}
\]

\[
Q_3(A_d, B_d, C_d) = B_d^{(3)} - B_d^{(1)} (A_d^{(2)} + (A_d^{(1)})^2 + \frac{(2\pi)^3}{6} B_d^{(1)} C_d^{(1)}).
\]

The infinite sum in the expression for \(q^{(3)}\) is explicitly computable. From the definition of the polylogarithm function

\[
Li_k(x) = \sum_{n=1}^{\infty} \frac{x^n}{n^k},
\]

we get

\[
\sum_{n \neq d} \frac{e^{inx}}{(n - d)^2} = e^{idx} \left( Li_2(e^{idx}) + Li_2(e^{-idx}) \right).
\]
Recall the well-known relation between the polylogarithms and the Bernoulli polynomials

$$\text{Li}_k(e^{ix}) + (-1)^k \text{Li}_k(e^{-ix}) = -\frac{(2\pi i)^k}{k!} B_k(\frac{x}{2\pi}).$$  \hfill (48)

This gives us the expression

$$q^{(3)}(x) = Q_3 e^{idx} + M^g B_d(\frac{x}{2\pi}) e^{idx}.$$  \hfill (49)

Collecting the terms $q^{(1)}, q^{(2)}, q^{(3)}$ and $r^{(1)}, r^{(2)}, r^{(3)}$, where also the calculation of $r^{(3)}$ is now obvious, gives

$$q(x, s) = C^d(s) e^{idx} + s^3 M^d B(\frac{x}{2\pi}) e^{idx} + \mathcal{O}(4),$$  \hfill (49)

$$r(x, s) = C^g(s) e^{-idx} + s^3 M^g B_d(\frac{x}{2\pi}) e^{-idx} + \mathcal{O}(4).$$  \hfill (50)

The coefficients above are given by

$$M^d = (4\pi^3 + 2\pi) (B_d^{(1)})^2 C_d^{(1)}, \quad M^g = (4\pi^3 + 2\pi) (C_g^{(1)})^2 B_d^{(1)}$$

and

$$C^d(s) = s B_d^{(1)} + \frac{s^2}{2} B_d^{(2)} + \frac{s^3}{6} Q_3(A_d, B_d, C_d),$$

$$C^g(s) = s C_g^{(1)} + \frac{s^2}{2} C_g^{(2)} + \frac{s^3}{6} Q_3(A_d, C_d, B_d).$$

One can think of $q(x)$ as a self modulated waveform with the carrier oscillation of frequency $d$ and the modulation envelope $s^3 M^d B_d(x/2\pi)$.

In the same way, with longer but still straightforward calculations, one can calculate the nonlinear superposition of two complex nonlinear Fourier modes of respective frequencies $d$ and $g$. Up to the third order, the final result for the function $q(x)$ is

$$q(x) = C^d_d(s) e^{idx} + C^g_g(s) e^{ixg}$$

$$+ s^3 \left( M^d_d B_d(\frac{x}{2\pi}) e^{idx} + M^g_g B_d(\frac{x}{2\pi}) e^{ixg} \right)$$

$$+ s^3 \left( I^d_d B_d(\frac{x}{2\pi}) (e^{idx} - e^{ixg}) \right).$$  \hfill (51)

The modulation terms are given by

$$M^d_d = (4\pi^3 + 2\pi) (B_d^{(1)})^2 C_d^{(1)}, \quad M^g_g = (4\pi^3 + 2\pi) (B_g^{(1)})^2 C_g^{(1)},$$

and the interaction term is

$$I^d_d = (-2\pi i) \frac{B_d^{(1)} B_g^{(1)} (C_d^{(1)} - C_g^{(1)})}{d - g}.$$  \hfill (51)

Since they are lengthy and will not be needed in the sequel, we will omit writing down the terms $C^d_d$ and $C^g_g$. The expression for $r(x)$ is obtained from the expression for $q(x)$ simply by interchanging the variables $B_d^{(1)}, B_g^{(1)}$ and $C_d^{(1)}, C_g^{(1)}$, wherever applicable. Since the Bernoulli functions $B_k(\frac{x}{2\pi})$, defined by (48), are periodic for all integers $n > 1$, the function $q(x)$, given by (51), is $2\pi$-periodic.
The function \( q(x) \), described above, can be chosen as the initial conditions for the complex ZS-AKNS systems, for instance, for the vector nonlinear Schrödinger system

\[
iq_t = q_{xx} - 2rq^2, \quad -ir_t = r_{xx} - 2qr^2.
\]

In most physical applications, various reality conditions are imposed. The standard reality conditions for the above system are \( r = \pm \bar{q} \). They yield the two versions of the nonlinear Schrödinger equation

\[
iq_t = q_{xx} \pm 2|q|^2q.
\]

(52)

In the real cases, the sensible elementary nonlinear Fourier modes are the superposition of two complex Fourier modes of the opposite frequencies, \( g = -d \). We can choose the coefficients \( B_d^0 \ldots \) in (51) in such a way that we get

\[
q(x,s) = C_d(s)e^{idx} + \bar{C}_d(s)e^{-idx} + s^3 \left( M_d e^{idx} + \bar{M}_d e^{-idx} \right) B_2 \frac{x}{2\pi} + O(4).
\]

(53)

Above, we have also chosen \( C_g = C_d \), which caused the vanishing of the interaction term containing the function \( B_1(\frac{x}{2\pi}) \).

4.2. Communication through a nonlinear medium

Recall that the evolution of the complex envelope \( q(\tau, l) \) of a narrow band signal in an optical fibre from the transmitter, located at \( l = 0 \), to the receiver at \( l = L \), is described by the nonlinear Schrodinger equation

\[
iq(\tau, l) = q_{\tau\tau}(\tau, l) + |q(\tau, l)|^2q(\tau, l)
\]

(54)

on the domain \( \{ (\tau, l) \in [0, \infty) \times [0, L] \} \). Here \( \tau = t - \beta l \) denotes the retarded time, where \( t \) is the ordinary time and \( \beta \) is a suitable parameter. The description of this and of many other physical phenomena which can be modeled by integrable equations can be found, for example, in [6]. In (54), all the physically relevant parameters are set to 1 for simplicity. Note that here the roles of time and space are reversed, compared with their more usual roles, as given in (52).

Suppose we want to transmit a quadruplet of real numbers \( (I_1, I_2, I_3, I_4) \). We shall describe a simple possible scheme of such communication. Let our slowly varying envelope be given by a perturbation series

\[
q(\tau, l) = s \ q^{(1)}(\tau, l) + \frac{s^2}{2!} q^{(2)}(\tau, l) + \frac{s^3}{3!} q^{(3)}(\tau, l) + \ldots
\]

We shall assume that our receiver detects the signal up to the order three. The quadruplet will be transmitted by means of a nonlinear superposition of two complex nonlinear Fourier modes. In our scheme, the numerical value of the perturbation parameter \( s \) is known in advance to the transmitter as well as to the receiver.

4.2.1. Coding and transmission. First we choose any quadruplet of complex numbers \( (B_d^{(1)}, B_d^{(2)}, B_g^{(1)}, B_g^{(2)}) \) such that the equations

\[
I_1 = |B_d^{(1)}|^2, \quad I_2 = \text{Re} \left( B_d^{(1)} \overline{B_d^{(2)}} + B_d^{(1)} \overline{B_d^{(2)}} - B_g^{(1)} \overline{B_g^{(2)}} \right),
\]

\[
I_3 = |B_g^{(1)}|^2, \quad I_4 = \text{Re} \left( B_g^{(1)} \overline{B_d^{(2)}} + B_g^{(1)} \overline{B_d^{(2)}} - B_d^{(1)} \overline{B_d^{(2)}} \right).
\]

(55)

(56)
are satisfied. Now, we encode the data \((B_d^{(1)}, B_d^{(2)}, B_g^{(1)}, B_g^{(2)})\) at the transmitter end as an approximate element of \(LSU(2) \subset LSL(2, \mathbb{C})\):

\[
\{K_{d,g}(n)\}_{n \in \mathbb{Z}} = I + s\left(\frac{A_d^{(1)} \delta_{n,d} + A_g^{(1)} \delta_{n,g}}{-B_d^{(1)} \delta_{n,d} - B_g^{(1)} \delta_{n,g}} \frac{B_d^{(1)} \delta_{n,d} + B_g^{(1)} \delta_{n,g}}{A_d^{(1)} \delta_{n,d} + A_g^{(1)} \delta_{n,g}}\right)
+ \frac{s^2}{2} \left(\frac{A_d^{(2)} \delta_{n,d} + A_g^{(2)} \delta_{n,g}}{A_d^{(2)} \delta_{n,d} + A_g^{(2)} \delta_{n,g}} 0\right)
+ \frac{s^3}{6} \left(\frac{A_d^{(3)} \delta_{n,d} + A_g^{(3)} \delta_{n,g}}{-B_d^{(3)} \delta_{n,d} - B_g^{(3)} \delta_{n,g}} \frac{B_d^{(2)} \delta_{n,d} + B_g^{(2)} \delta_{n,g}}{A_d^{(3)} \delta_{n,d} + A_g^{(3)} \delta_{n,g}}\right).
\]

The terms \(A_d^{(k)}, A_g^{(k)}, k = 1, 2, 3\) are not arbitrary. They are determined by two stipulations. Denote

\[
F^{-1}[(\{K_{d,g}(n)\}_{n \in \mathbb{Z}})] = (q_{d,g}(\tau, 0), \{\sigma_n\}_{n \in \mathbb{Z}}).
\]

We shall choose \(A_d^{(k)}, A_g^{(k)}, k = 1, 2, 3\) in such a way that, up to the third order in \(s\), we will have \(\sigma_d = 0\) and \(\sigma_g = 0\). In other words, up to the third order, we will have

\[
F[q_{d,g}(\tau, 0), \{\sigma_n\}_{n \in \mathbb{Z}}](d) = (-1)^d \text{Hol}[q_{d,g}(\tau, 0), \sigma](d) = K_{d,d}(d)
F[q_{d,g}(\tau, 0), \{\sigma_n\}_{n \in \mathbb{Z}}](g) = (-1)^d \text{Hol}[q_{d,g}(\tau, 0), \sigma](g) = K_{d,g}(g).
\]

From (41)–(44) we get the following expressions:

\[
\sigma_d^{(1)} = -\frac{i}{\pi} A_d^{(1)} ,
\]

\[
\sigma_d^{(2)} = \frac{i}{\pi} A_d^{(2)} + \frac{2}{d - g} \left(2 \text{Re}(B_d^{(1)} \overline{B_g^{(1)}}) + |B_g^{(1)}|^2\right) + i \frac{2 \pi}{d} |B_g^{(1)}|^2 ,
\]

\[
\sigma_d^{(3)} = \frac{i}{\pi} A_d^{(3)} - 4 \pi^2 \text{Re}(B_d^{(1)} \overline{B_g^{(2)}}) - \frac{4 \pi i}{d - g} \text{Re}(B_d^{(1)} B_d^{(2)} B_g^{(1)} B_g^{(2)} - B_g^{(1)} B_d^{(2)}).
\]

We get the expression for \(\sigma_g^{(k)}, k = 1, 2, 3\) simply by interchanging the roles of \(d\) and \(g\) everywhere in the formulae for \(\sigma_d^{(k)}\). Setting \(\sigma_d\) and \(\sigma_g\) to zero then gives

\[
A_d^{(1)} = 0 ,
\]

\[
A_d^{(2)} = -2 \pi^2 |B_d^{(1)}|^2 - \frac{2 \pi i}{d - g} \left(2 \text{Re}(B_d^{(1)} \overline{B_g^{(1)}}) + |B_g^{(1)}|^2\right) ,
\]

\[
A_d^{(3)} = -\frac{4 \pi^2}{d - g} \text{Re}(B_d^{(1)} B_d^{(2)} B_g^{(1)} B_g^{(2)} - B_g^{(1)} B_d^{(2)}) + i \frac{2 \pi}{d} |B_g^{(1)}|^2 .
\]

We now transmit the waveform \(q_{d,g}(\tau, 0)\) through the waveguide and obtain the function \(q_{d,g}(\tau, L), 0 \leq \tau \leq 2\pi\) at the receiver end.
4.2.2. Decoding. The function \( q_{d,g}(\tau, l) \) is the solution of the integrable nonlinear Schrödinger equation (54) and \( q_{d,g}(\tau, L) \) is its evaluation at \( l = L \). We will use the integrability of (54) to extract \( (I_1, I_2, I_3, I_4) \) from \( q_{d,g}(\tau, L) \).

As is well known, (54) can be rewritten in terms of a zero-curvature condition. The function \( q_{d,g}(\tau, l) \) is \( 2\pi \)-periodic in \( \tau \), therefore the \( l \)-evolution of the holonomy \( \text{Hol}_{d,g}(l, z) = \text{Hol}[q_{d,g}(\tau, l)](z) \) of the associated \( L \)-matrix is given by

\[
\text{Hol}_{d,g}(l, z) = \Psi(l, z) \cdot \text{Hol}_{d,g}(0, z) \cdot \Psi^{-1}(l, z),
\]

for every value of the spectral parameter \( z \). The matrix \( \Psi \) is obtained from the \( M \)-matrix of the Lax pair. From this we see that for every value \( z \) the expressions \( \text{tr}(\text{Hol}_{d,g}(l, z)^k) \) for \( k = 1, 2 \) are conserved quantities, that is, they are independent of \( l \).

Suppose now that we detected the transmitted waveform \( q_{d,g}(\tau, L) \) over the interval \( 0 \leq \tau \leq 2\pi \). Let us calculate its forward nonlinear Fourier transform

\[
\mathcal{F}[q_{d,g}(\tau, L), \{ \sigma_n \}_{n \in \mathbb{Z}}](n) = (-1)^n \text{Hol}(L, n + \sigma_n).
\]

In general, we do not know the values \( \sigma_n \) at which to evaluate the above expression in order to retrieve the transmitted information. The spectrum \( \{ \sigma_n \}_{n \in \mathbb{Z}} \) was not transmitted. But in our case, the transmitter and the receiver have agreed that the transmitter will ensure that \( \sigma_d = 0 \) and \( \sigma_g = 0 \). Therefore, we get

\[
\text{tr}\left( (\mathcal{F}[q_{d,g}(\tau, L), \sigma](d))^k \right) = (-1)^k \text{tr}(\text{Hol}(L, d)^k) + \mathcal{O}(4)
\]

\[
\text{tr}\left( (\mathcal{F}[q_{d,g}(\tau, L), \sigma](g))^k \right) = (-1)^k \text{tr}(\text{Hol}(L, g)^k) + \mathcal{O}(4)
\]

for \( k = 1, 2 \). The calculation shows that up to the order 3 of \( s \), we have

\[
\text{tr}(\mathcal{K}_{d,g}(d)) = 2 + s^2 \text{ Re}(A_{d}^{(2)}) + \frac{s^3}{3} \text{ Re}(A_{d}^{(3)}) + \mathcal{O}(4)
\]

\[
\text{tr}(\mathcal{K}_{d,g}(d)^2) = 2 + s^2 \left( 2 \text{ Re}(A_{d}^{(2)}) - 2 |B_{d}^{(1)}|^2 \right) + \frac{2s^3}{3} \text{ Re}(A_{d}^{(3)}) + \mathcal{O}(4)
\]

\[
\text{tr}(\mathcal{K}_{d,g}(g)) = 2 + s^2 \text{ Re}(A_{g}^{(2)}) + \frac{s^3}{3} \text{ Re}(A_{g}^{(3)}) + \mathcal{O}(4)
\]

\[
\text{tr}(\mathcal{K}_{d,g}(g)^2) = 2 + s^2 \left( 2 \text{ Re}(A_{g}^{(2)}) - 2 |B_{g}^{(1)}|^2 \right) + \frac{2s^3}{3} \text{ Re}(A_{g}^{(3)}) + \mathcal{O}(4).
\]

From the above equations and from (55)–(58) we get the linear systems

\[
2\pi^2 s^2 I_1 + \frac{4\pi^2 s^3}{3(d-g)} I_2 = \mathcal{I}_{1,d}
\]

\[
(4\pi^2 + 2)s^2 I_1 + \frac{8\pi^2 s^3}{3(d-g)} I_2 = \mathcal{I}_{2,d}, \quad (59)
\]

\[
2\pi^2 s^2 I_3 + \frac{4\pi^2 s^3}{3(g-d)} I_4 = \mathcal{I}_{1,g}
\]

\[
(4\pi^2 + 2)s^2 I_3 + \frac{8\pi^2 s^3}{3(g-d)} I_4 = \mathcal{I}_{2,g}, \quad (60)
\]

where \( \mathcal{I}_{k,d} = -\text{tr}(\mathcal{K}_{d,g}(d)^k) + 2, k = 1, 2 \), and similarly for \( \mathcal{I}_{k,g} \). Solving the linear systems (59) and (60) gives the approximate value of the desired information \( (I_1, I_2, I_3, I_4) \).

One of the advantages of using the superposition of nonlinear Fourier modes is the fact that both, on the transmitter end and on the receiver end, a relatively small amount of calculation has to be performed.
5. Conclusion

In section 4, we have computed a perturbation expression for a nonlinear Fourier mode up to the third order and we have encountered the Bernoulli polynomials. As a rule, each iteration in a recursive computation is more complex than the previous one and inputs can soon become too vast to handle. In our case, the appearance of the Bernoulli polynomials greatly simplifies the calculations. As already mentioned in the introduction, preliminary results indicate that the higher order terms of nonlinear Fourier modes are also expressible in terms of the Bernoulli polynomials. The families of the Bernoulli polynomials and their relatives (like polylogarithms) have many good properties and are richly structured. In the last two decades, they have appeared in rather different contexts, from analytic number theory to different areas of mathematical physics. Therefore one can expect that further investigation in this direction could yield interesting results and insights.

Studying the nonlinear superposition could also lead to interesting observations. One degree of freedom solutions (Fourier modes) of integrable partial differential equations with periodic boundary conditions are usually expressed in terms of various elliptic functions, while the $N$-degrees of freedom solutions for $N > 1$ are expressed in terms of theta functions, (see, for example, [3] and references therein). The application of the construction of the superposition proposed above could hopefully shed some new light on various relationships among these algebro-geometric functions. This problem is treated by the authors in [3], but their approach works only for certain special cases.

Acknowledgment

The research for this paper was supported in part by the research programme Analysis and Geometry, P1-0291 from ARRS, Republic of Slovenia.

ORCID iDs

Pavle Saksida © https://orcid.org/0000-0003-3093-9863

References


P Saksida