**Topic 7: The Inverse-Scattering Transform for the Focusing Nonlinear Schrödinger Equation**

Recall that the focusing nonlinear Schrödinger equation

\[ i\psi_t + \frac{1}{2} \psi_{xx} + |\psi|^2 \psi = 0 \]

for a complex-valued field \( \psi(x, t) \) is another equation that can be viewed as a compatibility condition for the simultaneous linear equations of a Lax pair:

\[ \frac{\partial w}{\partial x} = Uw := \begin{bmatrix} -i\lambda & \psi \\ \psi^* & i\lambda \end{bmatrix} w \]

and

\[ \frac{\partial w}{\partial t} = Vw := \begin{bmatrix} -i\lambda^2 + i\lambda |\psi|^2 & \lambda \psi + i\frac{1}{2} \psi_x \\ -\lambda \psi^* + i\frac{1}{2} \psi_x^* & \lambda^2 - i\frac{1}{2} |\psi|^2 \end{bmatrix} w. \]

In other words, the focusing NLS equation is equivalent to the zero-curvature condition

\[ \frac{\partial U}{\partial t} - \frac{\partial V}{\partial x} + [U, V] = 0 \]

(the left-hand side is independent of \( \lambda \) and vanishes when the focusing nonlinear Schrödinger equation (1) is satisfied by \( \psi \)).

Here we indicate how the inverse-scattering transform for this equation differs from that appropriate for the defocusing version of the nonlinear Schrödinger equation.

**Basic scattering theory.** Suppose that \( t \) is fixed and \( \psi(x) = \psi(x, t) \) is an absolutely integrable function of \( x \in \mathbb{R} \). For \( \lambda \in \mathbb{R} \) we may again define Jost matrix solutions \( J^\pm(x; \lambda) \) now of the equation (2), and again normalized by the boundary conditions \( J^\pm(x; \lambda)e^{i\lambda \sigma_3} \to \mathbb{I} \) as \( x \to \pm \infty \). There is again a scattering matrix \( S(\lambda) \) defined by \( J^+(x; \lambda) = J^-(x; \lambda)S(\lambda) \) for \( \lambda \in \mathbb{R} \). The Jost matrices and the scattering matrix are all unimodular (unit determinant) for exactly the same reason as in the defocusing case.

The first difference from the defocusing case is that the coefficient matrix \( U(x; \lambda) \) now has a different symmetry for \( \lambda \in \mathbb{R} \), namely,

\[ U(x; \lambda)^* = (i\sigma_2)^{-1}U(x; \lambda)(i\sigma_2), \quad \sigma_2 := \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}. \]

This implies that whenever \( w(x; \lambda) \) is a solution of (2) for some \( \lambda \in \mathbb{R} \), then so is \( (i\sigma_2)w(x; \lambda)^* \). Recall that we are using asterisks on matrices and vectors to denote elementwise complex conjugation (no transpose).

Since the Jost matrices \( J^\pm(x; \lambda) \) are uniquely determined for \( \lambda \in \mathbb{R} \) by the fact that they satisfy (2) and by their asymptotic behavior as \( x \to \pm \infty \), it follows that

\[ J^\pm(x; \lambda)^* = (i\sigma_2)^{-1}J^\pm(x; \lambda)(i\sigma_2). \]

From this it in turn follows that \( S(\lambda)^* = (i\sigma_2)^{-1}S(\lambda)(i\sigma_2) \), which implies that \( S(\lambda) \) can be written in the form

\[ S(\lambda) = \begin{bmatrix} a(\lambda)^* & b(\lambda)^* \\ -b(\lambda) & a(\lambda) \end{bmatrix}, \quad \lambda \in \mathbb{R} \]

for some complex-valued functions \( a(\lambda) \) and \( b(\lambda) \). The condition that \( \det(S(\lambda)) = 1 \) then implies that

\[ |a(\lambda)|^2 + |b(\lambda)|^2 = 1. \]

As in the defocusing case, the quantity \( T(\lambda) := 1/a(\lambda) \) is the transmission coefficient and the ratio \( R(\lambda) := b(\lambda)/a(\lambda) \) is the reflection coefficient, as they have exactly the same interpretation in terms of scattering of a unit-amplitude incident wave from an obstacle. However, in the focusing case we deduce from (5) that

\[ |T(\lambda)|^2 - |R(\lambda)|^2 = 1, \]

which is a rather non-physical relation for a conservative scattering process.
Time dependence of scattering matrix for $\lambda \in \mathbb{R}$. Provided that $\psi = \psi(x, t)$ is a solution of the focusing nonlinear Schrödinger equation (1) that is in $L^1(\mathbb{R})$ as a function of $x$ for each $t$, we may again ask about the time dynamics of the elements of the scattering matrix. By exactly the same arguments as in the defocusing case (involving writing differential equations in $t$ satisfied by the Jost matrices with the help of (3)), we may draw the same conclusions, namely that

$$a(t; t) = a(\lambda; 0) \quad \text{and} \quad R(t; t) = R(\lambda; 0)e^{2i\lambda x t}, \quad \lambda \in \mathbb{R},$$

provided that also $\psi$ and $\psi_x$ both vanish as $|x| \to \infty$ and that $\psi_t \in L^1(\mathbb{R})$ for each $t$. It is therefore tempting to define a scattering transform of $\psi$ as before by $\psi \mapsto R$, as such a mapping obviously trivializes the dynamics of the focusing nonlinear Schrödinger equation (1). However, unlike in the defocusing case, the mapping $\psi \mapsto R$ fails to be injective$^1$ in the focusing case. Indeed, it can be shown that the two $L^1(\mathbb{R})$ functions $\psi(x) \equiv 0$ and $\psi(x) = 2be^{-2iax}\text{sech}(2bx)$ (with parameters $a, b \in \mathbb{R}$) both yield $R(\lambda) \equiv 0$. This means that we will need to augment the scattering transform with additional data if we wish the transform to be invertible in the focusing case.

Additional properties of Jost solutions. The iteration argument based on Volterra integral equations that we used to study the Jost matrices in the defocusing case is rather soft, in the sense that it does not change substantially by merely altering the sign of the entry $U_{21}$ in the coefficient matrix $U(x; \lambda)$. We can therefore deduce that, provided $\psi \in L^1(\mathbb{R})$,

- $j^{-1}(x; \lambda)$ is defined for $\text{Im}(\lambda) \geq 0$ as a solution of (2) that for every $x \in \mathbb{R}$ is continuous for $\text{Im}(\lambda) > 0$ and analytic for $\text{Im}(\lambda) > 0$, and that satisfies
  
  $$j^{-1}(x; \lambda) = O(e^{\text{Im}(\lambda)x}), \quad x \to -\infty$$

  (exponential decay as $x \to -\infty$ for $\text{Im}(\lambda) > 0$) and

  $$\lim_{\lambda \to \infty} e^{i\lambda x} j^{-1}(x; \lambda) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$ 

- $j^{-2}(x; \lambda)$ is defined for $\text{Im}(\lambda) \leq 0$ as a solution of (2) that for every $x \in \mathbb{R}$ is continuous for $\text{Im}(\lambda) \leq 0$ and analytic for $\text{Im}(\lambda) < 0$, and that satisfies
  
  $$j^{-2}(x; \lambda) = O(e^{-\text{Im}(\lambda)x}), \quad x \to -\infty$$

  (exponential decay as $x \to -\infty$) and

  $$\lim_{\lambda \to \infty} e^{-i\lambda x} j^{-2}(x; \lambda) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$ 

- $j^{+1}(x; \lambda)$ is defined for $\text{Im}(\lambda) \leq 0$ as a solution of (2) that for every $x \in \mathbb{R}$ is continuous for $\text{Im}(\lambda) \leq 0$ and analytic for $\text{Im}(\lambda) < 0$, and that satisfies
  
  $$j^{+1}(x; \lambda) = O(e^{\text{Im}(\lambda)x}), \quad x \to +\infty$$

  (exponential decay as $x \to +\infty$) and

  $$\lim_{\lambda \to \infty} e^{i\lambda x} j^{+1}(x; \lambda) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$ 

- $j^{+2}(x; \lambda)$ is defined for $\text{Im}(\lambda) \geq 0$ as a solution of (2) that for every $x \in \mathbb{R}$ is continuous for $\text{Im}(\lambda) \geq 0$ and analytic for $\text{Im}(\lambda) > 0$, and that satisfies
  
  $$j^{+2}(x; \lambda) = O(e^{-\text{Im}(\lambda)x}), \quad x \to +\infty$$

  (exponential decay as $x \to +\infty$) and

  $$\lim_{\lambda \to \infty} e^{-i\lambda x} j^{+2}(x; \lambda) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$ 

$^1$We have not yet proven that the mapping $\psi \mapsto R$ is injective in the defocusing case, but it is true, and we will do this eventually, by developing the theory of Riemann-Hilbert problems.
Exactly as in the defocusing case, the spectral functions $a(\lambda)$ and $b(\lambda)$ can be expressed as Wronskians of Jost solutions by the same formulae:

$$a(\lambda) = \det[j^{-1}(x; \lambda), j^{+2}(x; \lambda)] \quad \text{and} \quad b(\lambda) = \det[j^{+1}(x; \lambda), j^{-1}(x; \lambda)], \quad \lambda \in \mathbb{R},$$

so again we deduce that $a(\lambda)$ is the boundary value taken on the real axis of a function continuous for $\Im(\lambda) \geq 0$ and analytic for $\Im(\lambda) > 0$ and that satisfies $a(\lambda) \to 1$ as $\lambda \to \infty$ in the closed upper half-plane. The function $b(\lambda)$ is continuous for $\lambda \in \mathbb{R}$ and decays to zero as $\lambda \to \infty$ along the real line in either direction, but it is in general not the boundary value of any analytic function in either the upper or lower half-planes.

**Inverse scattering. Riemann-Hilbert problem.** Exactly as in the defocusing case, we consider the $2 \times 2$ matrix function $M(\lambda; x)$ defined by the formula

$$M(\lambda; x) := \begin{cases}
\begin{bmatrix}
\frac{a(\lambda)}{i\lambda} e^{i\lambda x} j^{+1}(x; \lambda), & e^{-i\lambda x} j^{-1}(x; \lambda) \\
\frac{1}{a(\lambda)} e^{-i\lambda x} j^{+1}(x; \lambda), & e^{i\lambda x} j^{-1}(x; \lambda)
\end{bmatrix}, & \Im(\lambda) > 0, \\
\begin{bmatrix}
\frac{1}{a(\lambda)} e^{i\lambda x} j^{+2}(x; \lambda), & e^{-i\lambda x} j^{-2}(x; \lambda) \\
\frac{1}{a(\lambda)} e^{-i\lambda x} j^{+2}(x; \lambda), & e^{i\lambda x} j^{-2}(x; \lambda)
\end{bmatrix}, & \Im(\lambda) < 0.
\end{cases}$$

As in the defocusing case, $\det(M(\lambda; x)) = 1$ holds as an identity by the Wronskian formula for $a(\lambda)$. Also like in the defocusing case, $M(\lambda; x) \to \mathbb{I}$ as $\lambda \to \infty$ in every direction of the complex plane, including along the top and bottom “edges” of $\mathbb{R}$. The Schwarz reflection symmetry of $M(\lambda; x)$ is slightly different in the focusing case compared with the defocusing case; from the relations $j^{\pm 2}(x; \lambda) = (i\sigma_2)^{-1} j^{\pm 1}(x; \lambda)$ that hold for all $\lambda \in \mathbb{R}$ we get by analytic continuation that $j^{-2}(x; \lambda^*) = (i\sigma_2)^{-1} j^{-1}(x; \lambda)$ holds whenever $\Im(\lambda) \geq 0$, while $j^{-2}(x; \lambda^*) = (i\sigma_2)^{-1} j^{+1}(x; \lambda)$ holds whenever $\Im(\lambda) \leq 0$. Therefore, we obtain

$$M(\lambda; x) = (i\sigma_2)M(\lambda^*; x)^*(i\sigma_2)^{-1}, \quad \lambda \in \mathbb{C} \setminus \mathbb{R}. \quad (6)$$

At this juncture we have to point out another major difference between the focusing and defocusing theory: unlike in the defocusing case, the matrix $M(\lambda; x)$ might not be analytic at every point of $\mathbb{C} \setminus \mathbb{R}$, and it might fail to take continuous boundary values on $\mathbb{R}$ from the upper/lower half-planes. This is because in the focusing case the function $a(\lambda)$ might have zeros for $\Im(\lambda) \geq 0$.

Indeed, for $\lambda \in \mathbb{R}$ the identity $\det(S(\lambda)) = 1$ implies, in the focusing case, that $|a(\lambda)| \leq 1$ (whereas in the defocusing case we had $|a(\lambda)| > 1$). As in the defocusing case, the existence of a complex zero $\lambda_0$ of $a(\lambda)$ with $\Im(\lambda_0) > 0$ implies the existence of a nontrivial solution of (2) for $\lambda = \lambda_0$ that decays exponentially as $|x| \to \infty$. In the focusing case the equation (2) can be re-written in the form

$$Lw = \lambda w, \quad L := i\sigma_3 \frac{d}{dx} - i \begin{bmatrix} 0 & \psi(x) \\ \psi(x)^* & 0 \end{bmatrix},$$

so $\lambda_0$ is necessarily an $L^2$ eigenvalue of the operator $L$. Whereas in the defocusing case the analogue of the operator $L$ was self-adjoint, a fact that precluded the possibility of complex zeros of $a(\lambda)$, in the focusing case $L$ is not self-adjoint with respect to any inner product, so complex zeros cannot be ruled out in general.

While zeros of $a(\lambda)$ for $\Im(\lambda) \geq 0$ cannot be avoided in general, certain simplifications can be assumed for “generic” potentials $\psi \in L^1(\mathbb{R})$. Indeed, Beals and Coifman have shown [1, Theorem A] that there is an open dense subset $P_0$ of $L^1(\mathbb{R})$ consisting of potentials $\psi$ for which

- $a(\lambda)$ does not vanish for $\lambda \in \mathbb{R}$,
- $a(\lambda)$ has only finitely many zeros for $\Im(\lambda) > 0$, and each zero is simple.

The good news is that if $\psi(\cdot, 0)$ lies in $P_0$, then since the spectral function $a(\cdot)$ is independent of time $t$ when $\psi$ evolves under (1), then the corresponding properties of the zeros of $a$ persist for all time. We shall develop the Riemann-Hilbert problem of inverse scattering assuming that for some time $t \in \mathbb{R}$, $\psi(\cdot, t)$ lies in $P_0$ so that $a$ has only finitely many zeros, all simple, lying only in the open upper half-plane. However, one should be aware that there exist very nice potentials $\psi$ that have so-called “spectral singularities”, i.e., real zeros of $a(\cdot)$. For a simple class of examples, suppose that $\psi$ is real-valued and strictly positive, and consider $\lambda = 0$ in (2):

$$\frac{1}{\psi(x)} \frac{\partial \psi}{\partial x} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} w.$$
which becomes a constant-coefficient problem upon introducing a new independent variable by \( y'(x) = \psi(x) \):

\[
\frac{\partial w}{\partial y} = \begin{bmatrix} 0 & 0 \\ -1 & 0 \end{bmatrix} w \implies w = \alpha \begin{bmatrix} \sin(y) \\ \cos(y) \end{bmatrix} + \beta \begin{bmatrix} \cos(y) \\ -\sin(y) \end{bmatrix}.
\]

Here the general solution is parametrized by two integration constants, \( \alpha \) and \( \beta \). From this we can write down the Jost solutions \( j^{-1}(x; 0) \) and \( j^{+2}(x; 0) \) explicitly:

\[
j^{-1}(x; 0) = \begin{bmatrix} \cos \left( \int_{-\infty}^{x} \psi(x') \, dx' \right) \\ -\sin \left( \int_{-\infty}^{x} \psi(x') \, dx' \right) \end{bmatrix} \quad \text{and} \quad j^{+2}(x; 0) = \begin{bmatrix} \sin \left( -\int_{x}^{\infty} \psi(x') \, dx' \right) \\ \cos \left( -\int_{x}^{\infty} \psi(x') \, dx' \right) \end{bmatrix},
\]

from which it follows by trig identities that

\[
a(0) = \det[j^{-1}(x, 0), j^{+2}(x, 0)] = \cos \left( \int_{-\infty}^{\infty} \psi(x) \, dx \right) = \cos(||\psi||_1)
\]

where the final equality holds because \( \psi \) is positive. Therefore we see that \( \lambda = 0 \) is a spectral singularity for such potentials whenever \( ||\psi||_1 \) is of the form \( (n + \frac{1}{2})\pi, \, n \in \mathbb{Z} \); observe that this condition has nothing to do with how well-behaved the function \( \psi \) is because given any such function, \( a(0) = 0 \) can be arranged simply by replacing the potential \( \psi \) by a multiple of itself. It can also happen that for certain potentials \( \psi \) so nice as to lie in the Schwartz class \( \mathcal{S}(\mathbb{R}) \), there exist infinitely many zeros of \( a \) in the open upper half-plane that necessarily accumulate at one or more points of the real axis (because \( a(\lambda) \to 1 \) as \( \lambda \to \infty \) and \( a \) can have only isolated zeros in its domain of analyticity), forming a particularly exotic type of spectral singularity. See [2, Example 3.3.16] for details.

It is obvious from the analyticity properties of the function \( a(\lambda) \) and of the Jost solutions that \( M(\lambda; x) \) is analytic for \( \lambda \in \mathbb{C} \setminus \mathbb{R} \) with the sole exception of the complex zeros of \( a \) for \( \text{Im}(\lambda) > 0 \) and their complex conjugates. Under the condition that the zeros are all assumed to be simple, each is a simple pole of \( M(\lambda; x) \) (albeit in the first column only), as is its complex conjugate (albeit in the second column only).

Let \( \lambda_n, \text{Im}(\lambda_n) > 0, \, n = 1, \ldots, N \), denote the simple zeros of \( a(\lambda) \). By the Wronskian formula for \( a(\lambda) \), at each \( \lambda = \lambda_n \), the vector solutions \( j^{-1}(x; \lambda_n) \) and \( j^{+2}(x; \lambda_n) \) of (2) are linearly dependent; however neither of them can be the zero solution of (2) due to their normalization as \( x \to -\infty \) and \( x \to +\infty \) respectively. Therefore, there must exist a nonzero proportionality constant \( \gamma_n \in \mathbb{C} \), \( \gamma_n \neq 0 \), such that

\[
j^{-1}(x; \lambda_n) = \gamma_n j^{+2}(x; \lambda_n), \quad x \in \mathbb{R}.
\]

As the singularity of \( M(\lambda; x) \) at \( \lambda = \lambda_n \) is a simple pole, we may compute its residue explicitly:

\[
\text{Res}_{\lambda = \lambda_n} M(\lambda; x) = \begin{bmatrix} e^{i\lambda_n x} a'(\lambda_n) \\ a(\lambda_n) \end{bmatrix}^{-1} j^{-1}(x; \lambda_n), 0
\]

\[
= \frac{\gamma_n e^{i\lambda_n x}}{a'(\lambda_n)} j^{+2}(x; \lambda_n), 0
\]

\[
= \frac{\gamma_n e^{2i\lambda_n x}}{a'(\lambda_n)} e^{-i\lambda_n x} j^{+2}(x; \lambda_n), 0
\]

\[
= \lim_{\lambda \to \lambda_n} M(\lambda; x) \begin{bmatrix} 0 \\ c_n e^{2i\lambda_n x} \end{bmatrix},
\]

where

\[
c_n := \frac{\gamma_n}{a'(\lambda_n)}.
\]

Using the Schwarz symmetry formula (6), we can obtain the analogous relation that holds at the simple pole \( \lambda = \lambda_n^* \) in the lower half-plane:

\[
\text{Res}_{\lambda = \lambda_n^*} M(\lambda; x) = \lim_{\lambda \to \lambda_n^*} M(\lambda; x) \begin{bmatrix} 0 \\ -c_n^* e^{-2i\lambda_n^* x} \end{bmatrix}.
\]

Thus, in both cases the residue at a point of the column of \( M(\lambda; x) \) experiencing the singularity is proportional to the value of the other column at the same point. These residue conditions are very similar to those that occurred in the Riemann-Hilbert problem formulation of the inverse problem for the Toda lattice as well as to
those that we imposed in the dressing method to produce dark soliton solutions of the defocusing nonlinear Schrödinger equation. The points $\psi_1, \ldots, \psi_N$ in the upper half-plane and the corresponding residue constants $c_1, \ldots, c_N$ turn out to be the additional information that we need to keep track of for the scattering map to be injective. Therefore we have the following definition.

**Definition 1** (Scattering map for the focusing nonlinear Schrödinger equation). Let $\psi \in P_0 \subset L^1(\mathbb{R})$. The scattering map assigns to $\psi$ the following data:

- its reflection coefficient $R(\lambda)$ defined for $\lambda \in \mathbb{R}$, a continuous function decaying to zero for large real $\lambda$,
- its finite set (for $\psi \in P_0$) of eigenvalues of $L$ in $\mathbb{C}_+$, i.e., simple (for $\psi \in P_0$) zeros of $a(\lambda)$, $\lambda_1, \lambda_2, \ldots, \lambda_N$, and
- its corresponding complex residue constants $c_1, c_2, \ldots, c_N$.

It will turn out that with the reflection coefficient augmented with the discrete data $\{(\lambda_n, \gamma_n)\}_{n=1}^N$ in this way, the scattering map is injective on $P_0$. However, for this to be a suitable mapping for studying the focusing nonlinear Schrödinger equation (1), we need to deduce simple dynamics for the discrete data analogous to the simple dynamics of the reflection coefficient.

Since the function $a(\lambda; t) = a(\lambda; 0)$ for $\lambda \in \mathbb{R}$ does not change in time, its analytic continuation into the upper half-plane is also unchanged in time. In particular, the zeros $\{(\lambda_n)\}_{n=1}^N$ are fixed; the eigenvalues of $L$ are therefore constants of the motion (as for KdV). The residue constants $\{c_n\}_{n=1}^N$ are, however, not constant in time. Since $a'(\lambda_n)$ will be constant in time, it suffices to describe the time evolution of the proportionality constants $\gamma_n = \gamma_n(t)$ in $\mathbb{R}$. To derive a formula for them, we can use the defining relation

$$(7) \quad j^{-1}(x, t; \lambda_n) = \gamma_n(t) j^{+2}(x, t; \lambda_n)$$

and differentiate with respect to $t$:

$$(8) \quad \frac{\partial j^{-1}}{\partial t} = \gamma_n \frac{\partial j^{+2}}{\partial t} + d\gamma_n \frac{\partial j^{+2}}{\partial t}, \quad \lambda = \lambda_n.$$ 

As in the defocusing case, when $\psi$ evolves in time according to (1), the Jost matrices evolve according to

$$\frac{\partial J^\pm}{\partial t} = i\lambda^2 J^\pm \sigma_3 + V J^\pm, \quad \lambda \in \mathbb{R},$$

where now $V$ is the coefficient matrix from (3). Analytically continuing the first column of this equation for $J^-$ and the second column of this equation for $J^+$ into the upper half-plane yields

$$\frac{\partial j^{-1}}{\partial t} = (V + i\lambda^2 \mathbb{I}) j^{-1} \quad \text{and} \quad \frac{\partial j^{+2}}{\partial t} = (V - i\lambda^2 \mathbb{I}) j^{+2},$$

so evaluating at $\lambda = \lambda_n$ and substituting into (8) gives

$$V j^{-1} + i\lambda^2 j^{-1} = \gamma_n V j^{+2} - i\lambda^2 \gamma_j^{+2} + \frac{d\gamma_n}{dt} j^{+2}, \quad \lambda = \lambda_n.$$ 

Finally, eliminating $j^{-1}(x, t; \lambda_n)$ using (7) we arrive at

$$\frac{d\gamma_n}{dt} j^{+2}(x, t; \lambda_n) = 2i\lambda^2 \gamma_n j^{+2}(x, t; \lambda_n).$$

Since $j^{+2}(x, t; \lambda_n)$ is not the zero vector, we see that

$$\frac{d\gamma_n}{dt} = 2i\lambda^2 \gamma_n,$$

so $\gamma_n(t) = \gamma_n(0)e^{2i\lambda^2 t}$. Restoring the time-independent factor $1/a'(\lambda_n)$, we find at last that

$$c_n(t) = c_n(0)e^{2i\lambda^2 t}, \quad n = 1, \ldots, N.$$ 

This proves that, like the reflection coefficient, the discrete data $\{(\lambda_n, c_n)\}_{n=1}^N$ also evolves simply in time when the potential $\psi$ evolves according to the focusing nonlinear Schrödinger equation (1).

The last difference with the defocusing case arises in the details of the calculation of the jump matrix relating the boundary values $M_+ (\lambda; x)$ taken by $M(\lambda; x)$ on the real axis from the upper and lower half-planes. The basic calculation proceeds in exactly the same way as for the defocusing case, but now the form
(4) of the scattering matrix is different by a sign in one of the elements, and this sign propagates through the formal calculation. The result is as follows:

\[ M_+(\lambda; x) = M_-(\lambda; x)D(\lambda; x), \quad \lambda \in \mathbb{R} \]

where the jump matrix is defined in terms of the reflection coefficient as

\[ D(\lambda; x) := \begin{bmatrix} 1 + |R(\lambda)|^2 & e^{-2i\lambda x}R(\lambda)^* \\ e^{2i\lambda x}R(\lambda) & 1 \end{bmatrix}, \quad \lambda \in \mathbb{R}. \]

As in the defocusing case, the matrix \( M(\lambda; x) \) can be recovered from its scattering data by solving a matrix Riemann-Hilbert problem, and then the potential \( \psi(x) \) can be obtained from \( M(\lambda; x) \) via the formula

\[ \psi(x) = 2i \lim_{\lambda \to \infty} \lambda M_{12}(\lambda; x). \]

Including in the scattering data its explicit time evolution as implied by the focusing nonlinear Schrödinger equation (1) for \( \psi(x, t) \), we arrive at an algorithm for solving the initial-value problem when \( \psi(\cdot, 0) \in P_0 \subset L^1(\mathbb{R}) \).

**Summary:** the inverse-scattering transform solution algorithm for the initial-value problem for the focusing nonlinear Schrödinger equation. We wish to find the solution \( \psi(x, t) \) to the nonlinear equation (1) that decays sufficiently rapidly as \(|x| \to \infty\), and that satisfies the initial condition \( \psi(x, 0) = \psi_0(x) \) with \( \psi_0 \in P_0 \subset L^1(\mathbb{R}) \). The algorithm for solving this problem is as follows:

1. By computing Jost solutions associated with the linear system (2) of differential equations determine the reflection coefficient \( R(\lambda) \), the eigenvalues \( \lambda_1, \ldots, \lambda_N \) in the upper half-plane, and the residue constants \( c_1, \ldots, c_N \), where \( c_n := \gamma_n/a'(\lambda_n) \) (Note that \( a'(\lambda_n) \neq 0 \) whenever \( a(\lambda_n) = 0 \) for \( \psi_0 \in P_0 \)).
2. Solve the following Riemann-Hilbert problem: seek a \( 2 \times 2 \) matrix \( M(\lambda; x, t) \) satisfying the following properties:
   - **Analyticity:** \( M(\lambda; x, t) \) is an analytic function of \( \lambda \) for \( \lambda \in \mathbb{C} \setminus (\mathbb{R} \cup \{\lambda_1, \ldots, \lambda_N, \lambda_1^*, \ldots, \lambda_N^*\}) \).
   - **Residues:** At \( \lambda = \lambda_n \) and \( \lambda = \lambda_n^* \), \( M(\lambda; x, t) \) has simple poles and the residues satisfy the conditions
     \[ \text{Res}_{\lambda=\lambda_n} M(\lambda; x, t) = \lim_{\lambda \to \lambda_n} M(\lambda; x, t) \begin{bmatrix} 0 & 0 \\ c_n(x, t) & 0 \end{bmatrix}, \]
     and
     \[ \text{Res}_{\lambda=\lambda_n^*} M(\lambda; x, t) = \lim_{\lambda \to \lambda_n^*} M(\lambda; x, t) \begin{bmatrix} 0 & -c_n(x, t)^* \\ 0 & 0 \end{bmatrix}, \]
     where \( c_n(x, t) := c_n e^{2i(\lambda_n x + \lambda_n^2 t)} \).
   - **Jump Condition:** The matrix \( M(\lambda; x, t) \) takes continuous boundary values on the real axis
     \[ M_\pm(\lambda; x, t) := \lim_{\epsilon \downarrow 0} M(\lambda \pm i\epsilon; x, t), \]
     that are related by the condition
     \[ M_+(\lambda; x, t) = M_-(\lambda; x, t)D(\lambda; x, t), \]
     where
     \[ D(\lambda; x, t) := \begin{bmatrix} 1 + |R(\lambda)|^2 & e^{-2i(\lambda x + \lambda^2 t)}R(\lambda)^* \\ e^{2i(\lambda x + \lambda^2 t)}R(\lambda) & 1 \end{bmatrix}. \]
   - **Normalization:** As \( \lambda \to \infty \), \( M(\lambda; x, t) \to I \).
3. From the solution of this Riemann-Hilbert problem given the scattering data, one reconstructs the solution of the initial-value problem \( \psi(x, t) \) from the limit
   \[ \psi(x, t) = 2i \lim_{\lambda \to \infty} \lambda M_{12}(\lambda; x, t). \]
Reflectionless potentials. Multisoliton solutions. As remarked earlier, for some \( \psi_0(x) \) other than the zero potential \( \psi_0(x) \equiv 0 \), it can turn out that the reflection coefficient \( R(\lambda) \) vanishes identically for all \( \lambda \in \mathbb{R} \). Such potentials \( \psi_0(x) \), called reflectionless potentials, for obvious reasons, are then completely characterized by a number \( N \) of eigenvalues \( \lambda_1, \ldots, \lambda_N \) in the upper half-plane, and a corresponding number of nonzero complex numbers \( c_1, \ldots, c_N \). Since \( R(\lambda) \equiv 0 \), the jump matrix \( D(x; t) \) becomes the identity matrix, so there is no disagreement between the boundary values \( M_\pm(x; t) \) taken on the real axis in the \( \lambda \)-plane.

This means that (if \( N \) is finite) \( M(\lambda; x, t) \) is a meromorphic function of \( \lambda \): it is analytic\(^2\) except for simple poles at the points \( \lambda_1, \ldots, \lambda_N \) and their complex conjugates. We should observe that in such a situation, the Riemann-Hilbert problem for \( M(\lambda; x, t) \) begins to look like that which arose in the solution of the Toda lattice. That latter problem was solved by observing that the residue conditions were equivalent to a finite-dimensional linear system of algebraic equations that encoded the Gram-Schmidt process. Something similar takes place for reflectionless potentials for the focusing nonlinear Schrödinger equation.

Any meromorphic function can be represented by an expansion in partial fractions. Thus, when \( R(\lambda) \equiv 0 \), we may represent \( M(\lambda; x, t) \) in the form

\[
M(\lambda; x, t) = \mathbb{I} + \sum_{n=1}^{N} \frac{A_n(x, t)}{\lambda - \lambda_n} + \sum_{n=1}^{N} \frac{B_n(x, t)}{\lambda - \lambda_n^*},
\]

for some undetermined matrices \( A_1(x, t), \ldots, A_N(x, t) \) and \( B_1(x, t), \ldots, B_N(x, t) \). Note that this formula builds in the normalization condition that \( M(\lambda; x, t) \to \mathbb{I} \) as \( \lambda \to \infty \). To determine the matrix coefficients, we need to impose the conditions that constrain the residues. Since according to our formula for \( M(\lambda; x, t) \) we have

\[
\text{Res}_{\lambda=\lambda_n} M(\lambda; x, t) = A_n(x, t), \quad \text{and} \quad \text{Res}_{\lambda=\lambda_n^*} M(\lambda; x, t) = B_n(x, t),
\]

the residue conditions require that the matrices \( A_n(x, t) \) have zeros in the second column:

\[
A_n(x, t) = \begin{bmatrix} r_n(x, t) & 0 \\ s_n(x, t) & 0 \end{bmatrix},
\]

while the matrices \( B_n(x, t) \) have zeros in the first column:

\[
B_n(x, t) = \begin{bmatrix} 0 & u_n(x, t) \\ 0 & v_n(x, t) \end{bmatrix},
\]

and then the functions \( r_n(x, t), s_n(x, t), u_n(x, t), \) and \( v_n(x, t) \) are constrained by the relations

\[
\begin{bmatrix} r_n(x, t) \\ s_n(x, t) \end{bmatrix} = c_n e^{2i(\lambda_n x + \lambda_n^* t)} \left( \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \sum_{k=1}^{N} \frac{1}{\lambda_n - \lambda_k} \begin{bmatrix} u_k(x, t) \\ v_k(x, t) \end{bmatrix} \right), \quad n = 1, \ldots, N,
\]

(first the column of the residue at \( \lambda = \lambda_n \)), and

\[
\begin{bmatrix} u_n(x, t) \\ v_n(x, t) \end{bmatrix} = -c_n^* e^{-2i(\lambda_n^* x + \lambda_n^* t)} \left( \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \sum_{k=1}^{N} \frac{\lambda_n - \lambda_k}{\lambda_n^* - \lambda_k} \begin{bmatrix} r_k(x, t) \\ s_k(x, t) \end{bmatrix} \right), \quad n = 1, \ldots, N,
\]

(second the column of the residue at \( \lambda = \lambda_n^* \)). This constitutes a square system of linear equations for unknowns \( r_n, s_n, u_n, \) and \( v_n \), for \( n = 1, \ldots, N \). Solving these equations (possible if a determinant is nonzero) gives us \( M(\lambda; x, t) \). From \( M(\lambda; x, t) \) one recovers \( \psi(x, t) \) by the usual formula, which in this case amounts to a relation between \( \psi(x, t) \) and the quantities \( \{u_n(x, t)\}_{n=1}^{N} \):

\[
(9) \quad \psi(x, t) = 2i \lim_{\lambda \to \infty} \lambda M_{12}(\lambda; x, t) = 2i \lim_{\lambda \to \infty} \lambda \sum_{n=1}^{N} \frac{u_n(x, t)}{\lambda - \lambda_n^*} = 2i \sum_{n=1}^{N} u_n(x, t).
\]

\(^2\)The fact that the agreement of the boundary values means that \( M(\lambda; x, t) \) can be defined by continuity for real \( \lambda \) so as to become a matrix function analytic in a neighborhood of the real axis follows from Morera’s Theorem.
We can therefore find $\psi(x,t)$ by solving for $u_1(x,t), \ldots, u_N(x,t)$, and we can get a closed system of linear equations for the quantities $\{u_n(x,t)\}$ alone by explicitly eliminating the quantities $\{r_n(x,t)\}$:

$$ u_n(x,t) = -c_n^* e^{-2i(\lambda_n^* x + \lambda_n^* t)} \left( 1 + \sum_{k=1}^N \frac{1}{\lambda_n^* - \lambda_k} c_k e^{2i(\lambda_k x + \lambda_k^* t)} \sum_{j=1}^N \frac{1}{\lambda_k - \lambda_j} u_j(x,t) \right), \quad n = 1, \ldots, N. $$

The simplest example. Let’s try this procedure out in the simplest case: $N = 1$. Then the equation governing $u_1$ is just

$$ u_1(x,t) = -c_1^* e^{-2i(\lambda_1^* x + \lambda_1^* t)} \left( 1 + \frac{c_1 e^{2i(\lambda_1 x + \lambda_1^* t)}}{4\text{Im}(\lambda_1)^2} u_1(x,t) \right). $$

Writing $\lambda_1 = a + ib$, note that there are some real constants $\phi$ and $\Delta$ so that

$$ c_1 e^{2i(\lambda_1 x + \lambda_1^* t)} = -e^{2i(ax+(a^2-b^2)t+\phi)} e^{-2b(x+2at)+\Delta}. $$

Therefore,

$$ u_1(x,t) = \frac{e^{-2i(ax+(a^2-b^2)t+\phi)} e^{-2b(x+2at)+\Delta}}{1 + \frac{1}{4\phi^2} e^{-2b(x+2at)+2\Delta}}. $$

Finally, multiplying numerator and denominator by $2be^{2b(x+2at)-\Delta}$ and combining $\Delta$ and $\log(2b)$ into another constant $x_0$, we find

$$ u_1(x,t) = be^{-2i(ax+(a^2-b^2)t+\phi) \text{sech}(2b((x-x_0)+2at))}. $$

Therefore, for some other arbitrary phase constant $\theta$ we have a formula for a solution of the focusing nonlinear Schrödinger equation:

$$ \psi(x,t) = 2iu_1(x,t) = 2be^{-2i(ax+(a^2-b^2)t+\phi) \text{sech}(2b((x-x_0)+2at))}. $$

This formula is a soliton solution of the focusing nonlinear Schrödinger equation. It consists of a complex exponential “carrier wave” multiplied by a pulse-shaped envelope of amplitude $2b$ and width $(2b)^{-1}$, propagating to the right at a constant velocity $-2a$. Note that for the soliton solution of the focusing nonlinear Schrödinger equation, the amplitude and velocity of the soliton envelope are independent real parameters. This is quite different from KdV, where the speed and height of the solitary wave are always related.

The case $N > 1$. Reflectionless potentials for which $N > 1$ correspond to nonlinear interactions of solitons. We will later be able to give a proof, without calculating any determinant, that the linear system (10) can always be solved for any $N$. We will do this by proving that the Riemann-Hilbert problem is always solvable. However, it is also possible to analyze the determinant directly.

First we rewrite the system. Let $s_n(x,t) \neq 0$ denote any number such that $s_n(x,t)^2 = c_n e^{2i(\lambda_n x + \lambda_n^* t)} \neq 0$. Introducing new variables $\hat{u}_1, \ldots, \hat{u}_N$ by the invertible transformation $u_n(x,t) = s_n(x,t)^* \hat{u}_n(x,t)$, we write the system of equations (10) in the form $(I + AA^*)n = f$, where $A = A(x,t)$ is the $N \times N$ matrix with elements

$$ A_{nk}(x,t) = -s_n(x,t)^* s_k(x,t) \frac{\lambda_n^* - \lambda_k}{\lambda_n^* - \lambda_k} $$

and $f = f(x,t)$ is a certain well-defined $N$-component vector. Here $A^*$ refers, as usual, to the component-wise complex conjugate matrix (no transpose). Indeed, in terms of the quantities $s_n(x,t)^2$, the system (10) takes the form

$$ u_n(x,t) = -s_n(x,t)^2 \left( 1 + \sum_{k=1}^N \sum_{j=1}^N \frac{s_k(x,t)^2}{(\lambda_n^* - \lambda_k)(\lambda_k - \lambda_j^*)} u_j(x,t) \right), \quad n = 1, \ldots, N. $$

Making the indicated substitution $u_n(x,t) = s_n(x,t)^* \hat{u}_n(x,t)$ and dividing through by $s_n(x,t)^*$ gives

$$ \hat{u}_n(x,t) = -s_n(x,t)^* - \sum_{k=1}^N \sum_{j=1}^N \frac{s_n(x,t)^* s_k(x,t)^2 s_j(x,t)^*}{(\lambda_n^* - \lambda_k)(\lambda_k - \lambda_j^*)} \hat{u}_j(x,t), \quad n = 1, \ldots, N. $$

Inserting $1 = (-i)i = (-i)(-i)^*$, this system clearly has the desired form provided we take the vector $f$ to have components $f_n(x,t) := -s_n(x,t)^*$. 

8
The matrix $A$ is Hermitian. Indeed, we calculate:

$$(A^d)_{jk} := A^d_{kj} = \left[ -i \frac{s_k(x,t)^* s_j(x,t)}{\lambda_k - \lambda_j} \right]^* = i \frac{s_k(x,t) s_j(x,t)^*}{\lambda_k - \lambda_j} = -i \frac{s_j(x,t)^* s_k(x,t)}{\lambda_j^* - \lambda_k} = A_{jk}.$$  

Therefore, $A^d = A$ as claimed. But more is true: $A$ is also positive-definite. To see this, for fixed $(x,t) \in \mathbb{R}^2$, consider the functions $h_n(y) := s_n(x,t)e^{iy \lambda} \text{ for } n = 1, \ldots, N$. Note that

$$\int_0^\infty |h_n(y)|^2 dy = |s_n(x,t)|^2 \int_0^\infty e^{iy \lambda}|^2 dy = |s_n(x,t)|^2 \int_0^\infty e^{-2b_n y} dy = \frac{|s_n(x,t)|^2}{2b_n} < \infty,$$

because $b_n := \text{Im}(\lambda_n) > 0$. Hence $h_n \in L^2(0, \infty)$. Also, since $\lambda_1, \ldots, \lambda_N$ are distinct, we can show that the functions $h_1, \ldots, h_N$ are linearly independent for $0 < y < \infty$. Indeed, suppose we have constants $c_1, \ldots, c_N$ such that $c_1 h_1(y) + \cdots + c_N h_N(y) = 0$ holds identically on $(0, \infty)$. Then by differentiating this identity $N - 1$ times we can assemble the system of equations

$$\begin{bmatrix}
    h_1(y) & \cdots & h_N(y) \\
    h_1'(y) & \cdots & h_N'(y) \\
    \vdots & \ddots & \vdots \\
    h_1^{(N-1)}(y) & \cdots & h_N^{(N-1)}(y)
\end{bmatrix}
\begin{bmatrix}
    c_1 \\
    c_2 \\
    \vdots \\
    c_N
\end{bmatrix} = \begin{bmatrix}
    0 \\
    0 \\
    \vdots \\
    0
\end{bmatrix}.$$

Taking the derivatives, which only affect the factor $e^{i \lambda_n y}$, we can write this system in the form

$$\begin{bmatrix}
    1 & \cdots & 1 & \cdots & 1 \\
    i \lambda_1 & \cdots & i \lambda_N \\
    \vdots & \ddots & \vdots & \ddots & \vdots \\
    (i \lambda_1)^{N-1} & \cdots & (i \lambda_N)^{N-1}
\end{bmatrix}
\begin{bmatrix}
    s_1(x,t) c_1 \\
    s_2(x,t) c_2 \\
    \vdots \\
    s_N(x,t) c_N
\end{bmatrix} = \begin{bmatrix}
    0 \\
    0 \\
    \vdots \\
    0
\end{bmatrix}.$$

The coefficient matrix of this system is the famous Vandermonde matrix, and it is well-known to be invertible provided the parameters $i \lambda_1, \ldots, i \lambda_N$ are distinct numbers. Therefore, multiplying through by the inverse, we deduce that $s_n(x,t) c_n = 0$ for all $n = 1, \ldots, N$. But since $s_n(x,t) \neq 0$, this implies that $c_1 = c_2 = \cdots = c_N = 0$, which proves the linear independence. The Gram matrix $G$ of inner products $G_{nk} := \langle h_n, h_k \rangle$ of linearly independent elements of an inner product space is well-known to be a Hermitian positive-definite matrix. Indeed, if $v \in \mathbb{C}^N$ is any vector,

$$v^\dagger G v = \sum_{n=1}^N \sum_{k=1}^N v_n^* G_{nk} v_k = \sum_{n=1}^N \sum_{k=1}^N v_n^* \langle h_n, h_k \rangle v_k = \langle v_1 h_1 + \cdots + v_N h_N, v_1 h_1 + \cdots + v_N h_N \rangle = \|v_1 h_1 + \cdots + v_N h_N\|^2$$

where the norm is induced in the usual way from the inner product. This is obviously nonnegative, and since $h_1, \ldots, h_N$ are linearly independent, $v^\dagger G v$ can only be zero if $v$ is the zero vector. Calculating the inner products gives

$$G_{nk} := \langle h_n, h_k \rangle := \int_0^\infty h_n(y)^* h_k(y) dy = s_n(x,t)^* s_k(x,t) \int_0^\infty e^{i(\lambda_k - \lambda_n) y} dy = \frac{s_n(x,t)^* s_k(x,t)}{i(\lambda_k - \lambda_n^*)} = A_{nk}.$$  

This proves that $A$ is in fact a Gram matrix $G$ of inner products of linearly independent functions, and hence is positive-definite.

Because $A$ is Hermitian positive-definite, it has a positive-definite square root $A^{1/2}$, that we can compute by the spectral theorem. Indeed, $A = A^d$ can be diagonalized as $A = U D U^\dagger$, where $U$ is a unitary matrix of eigenvectors of $A$, and where $D = \text{diag}(\alpha_1, \ldots, \alpha_N)$ is the corresponding diagonal matrix of (positive real, because $A$ is positive-definite) eigenvalues of $A$. Let $\sqrt{\alpha_n}$ denote the positive square root of $\alpha_n > 0$, for $n = 1, \ldots, N$. Then a positive-definite square root of $A$ is given by the formula

$$A^{1/2} := U \text{ diag}(\sqrt{\alpha_1}, \ldots, \sqrt{\alpha_N}) U^\dagger.$$
This is obviously Hermitian positive-definite, and also it is obvious (since $U$ is unitary) that $A^{1/2}A^{1/2} = A$. Note that positive definite square roots of matrices are not unique, because the eigenvector matrix $U$ involves making some choices that do not cancel out from the formula for $A^{1/2}$.

From these results, we can show easily that the matrix $A^{1/2}A^*A^{1/2}$ is Hermitian positive-definite too. Indeed, first, note that $A^{1/2}A^*A^{1/2}$ is Hermitian because $A$ and $A^{1/2}$ are:

$$(A^{1/2}A^*A^{1/2})^\dagger = (A^{1/2})^\dagger A^T(A^{1/2})^\dagger = A^{1/2}A^*A^{1/2}.$$ 

Now let $v \in \mathbb{C}^N$, and consider $v^\dagger A^{1/2}A^*A^{1/2}v$. Since $A^{1/2}$ is Hermitian, we can write this as $u^\dagger A^*u$, where $u := A^{1/2}v$; note also that since $A^{1/2}$ is positive-definite it is invertible, and therefore $u = 0$ if and only if $v = 0$. Furthermore, it is easy to see that if $A$ is Hermitian positive-definite, then so is $A^* = A^T$. Therefore $u^\dagger A^*u \geq 0$ and it vanishes only if $u = 0$.

With these preliminaries, we now show that $I + AA^*$ is invertible for every $(x, t) \in \mathbb{R}^2$, hence proving the existence of the $N$-soliton solution. We first conjugate $I + AA^*$ by the invertible matrix $A^{1/2}$ to obtain a matrix $B$ that necessarily has the same determinant as does $I + AA^*$:

$$B := (A^{1/2})^{-1}(I + AA^*)A^{1/2} = I + A^{1/2}A^*A^{1/2}$$

where we used the identity $A = A^{1/2}A^{1/2}$. But the eigenvalues of $B$ are $1$ plus those of the Hermitian positive-definite matrix $A^{1/2}A^*A^{1/2}$, which are necessarily positive by part (e). Therefore the determinant of $B$ is a product of positive factors that all exceed $1$, and therefore $\det(I + AA^*) = \det(B) > 1$. It follows that $I + AA^*$ is invertible.

References
