

Finite Genus Solutions to the Ablowitz-Ladik Equations

PETER D. MILLER¹

Program in Applied Mathematics, University of Arizona

AND

NICHOLAS M. ERCOLANI

Department of Mathematics, University of Arizona

AND

IGOR M. KRICHEVER

The Landau Institute for Theoretical Physics

AND

C. DAVID LEVERMORE

Department of Mathematics, University of Arizona

Abstract

Generic wavetrain solutions to the complex Ablowitz-Ladik equations are developed using methods of algebraic geometry. The inverse spectral transform is used to realize these solutions as potentials in a spatially discrete linear operator. The manifold of wavetrains is infinite dimensional, but is stratified by finite dimensional submanifolds indexed by nonnegative integers g . Each of these strata is a foliation whose leaves are parametrized by the moduli space of (possibly singular) hyperelliptic Riemann surfaces of genus g . The generic leaf is a g dimensional complex torus. Thus, each wavetrain is constructed from a finite number of complex numbers comprising a set of spectral data, indicating that the wavetrain has a finite number of degrees of freedom. Our construction uses a new Lax pair differing from that originally given by Ablowitz and Ladik. This new Lax pair allows a simplified construction that avoids some of the degeneracies encountered in previous analyses making use of the original discretized AKNS Lax pair. Generic wavetrains are built from Baker-Akhiezer functions on nonsingular Riemann surfaces having distinct branch points, and the construction is extended to handle singular Riemann surfaces that are pinched off at a coinciding pair of branch points. The corresponding solutions in the pinched case may also be found from wavetrains belonging to nonsingular surfaces using Bäcklund transformations. The problem of the reduction of the complex Ablowitz-Ladik equations to the focusing and defocusing versions of the discrete nonlinear Schrödinger equation is solved by specifying which spectral data correspond to focusing or defocusing potentials. Within the class of finite genus complex potentials, spatially periodic potentials are isolated, resulting in a formula for the solution to the spatially periodic initial

¹CURRENT AFFILIATION: OPTICAL SCIENCES CENTRE, AUSTRALIAN NATIONAL UNIVERSITY

value problem. Formal modulation equations governing slow evolution of $g+1$ phase wavetrains are developed, and a gauge invariance is used to simplify the equations in the focusing and defocusing cases. In both of these cases, the modulation equations can be either hyperbolic (suggesting modulational stability) or elliptic (suggesting modulational instability) depending upon the local initial data. As has been shown to be the case with modulation equations for other integrable systems, hyperbolic data will remain hyperbolic under the evolution at least until infinite derivatives develop.

1. Introduction.

The program of understanding the macroscopic behavior of oscillations in nonlinear dispersive wave systems began in the 1960's with the work of Whitham [50], who showed how to describe slowly varying single phase wavetrains both through averaging a variational principle describing the microscopic waves and also through the averaging of local conservation laws. Whitham's analysis yielded a set of modulation equations that described the evolution (in appropriately scaled slow time and space variables) of the amplitude and wavenumber of the microscopic waves. The first steps in generalizing these methods to handle slowly varying multiphase waves were taken by Ablowitz and Benney [2] and Ablowitz [1], who explained how to use the method of multiple scales to obtain modulation equations for multiphase waves, given that one could isolate a family of exact multiphase waves. These methods had limited applicability because families of multiphase waves were difficult to obtain, and because the Cauchy problem for the resulting modulation equations was often ill-posed, which was interpreted as indicating the modulational instability of the multiphase wavetrain. These elliptic modulation equations could only be successfully integrated for initial data that were real analytic in the macroscopic spatial variable.

The discovery that many physical nonlinear dispersive wave equations are integrable by an inverse scattering transform² was the breakthrough that allowed the modulation theory of multiphase waves to continue. The thirteen years between the discovery of the inverse scattering transform for the Korteweg-de Vries equation by Gardner, Greene, Kruskal, and Miura [22] and the derivation of the modulation equations for g -phase wavetrains in the Korteweg-de Vries equation by Flaschka, Forest, and McLaughlin [20] saw the development of algebro-geometric methods for solving the Korteweg-de Vries equation (and other integrable systems) in the class of multiphase waves. The spectral theory of the quasiperiodic Schrödinger operator associated with the Korteweg-de Vries equation was studied (see, for example Dubrovin, Matveev, and Novikov [13]) and soon thereafter a direct algebro-geometrical construction of multiphase wave solutions was advanced. This construction was based on the complex structure of Riemann surfaces on which the quasiperiodic Schrödinger eigenfunction is single valued (see Krichever [29]).

²See the book by Faddeev and Takhtajan [18] for an exposition about the inverse scattering transform using the nonlinear Schrödinger equation as a central example.

With the practical difficulties of obtaining formal modulation equations in integrable nonlinear dispersive wave equations under control, it remained to determine the ubiquity of these modulation equations. There was little doubt that as long as the modulation equations were hyperbolic, they would correctly describe the slow evolution of a system that was appropriately prepared to have a slowly varying wavetrain as an initial condition. It was also shown that the zero-dispersion limits of the Korteweg-de Vries equation [34] and the nonlinear Schrödinger equation [27, 16] are weak limits described by multiphase modulation equations³. However what was not known was whether modulated wavetrains could appear spontaneously in nonlinear dispersive wave equations with arbitrary initial data in the limit of long time, as is known to be the case in linear dispersive wave theory. Happily, this question was answered in the affirmative by Deift, Venakides, and Zhou [8] in the case of the Korteweg-de Vries equation. Their method involves the interpretation of oscillatory Riemann-Hilbert problems and has been extended to the AKNS hierarchy in a study of the modified Korteweg-de Vries equation [9], again resulting in the conclusion that the long time limit is locally given by a multiphase wavetrain whose long scale behavior is governed by modulation equations.

Most of the progress in multiphase modulation theory has occurred in the study of dispersive wave systems that are spatially continuous. A notable exception is the Toda lattice system, the modulation theory of which has been studied by Bloch and Kodama [5]. Since most familiar spatially discrete integrable systems have natural continuum limits that are also integrable, it is easy to incorrectly assume that the modulational description of the discrete system will parallel that of the associated continuum system. The modulational description of the discrete system is in fact richer than that of the associated continuum system because the two differ significantly for finite lattice spacings h , and yet the discrete system becomes the continuous system as h goes to zero.

These ideas are stressed in the recent work of Levermore and Liu [38, 37] and Hays, Levermore, and Miller [25] who undertake the macroscopic description of oscillations in general conservative lattices, governed on a microscopic level by systems of (infinitely many) coupled ordinary differential equations in time. The microscopic equations describe the evolution of a quantity depending on an integer position index, n , and a real time variable, t ; the goal is to identify and describe behavior on spatial and temporal scales much longer than those characteristic of n and t . The main tool that is used here is the averaging of local conservation laws about oscillatory solutions to obtain formal modulation equations for the macroscopic variables. This procedure consists in general of several steps:

1. Identify a class of oscillatory solutions of the microscopic dynamical system at hand. This class is parametrized by a finite number of real constants of motion.

³A recent review of zero-dispersion limits of physical systems was written by Lax, Levermore, and Venakides [35].

2. Allow these parameters to depend on the slow modulational variables $X = hn$ and $T = ht$, where h is the small lattice spacing. Substitute this modulated solution into a set of local conservation laws. There should be as many independent real local conservation laws as there are real parameters.
3. Formally pass to the limit $h \downarrow 0$ in the local conservation laws to obtain partial differential equations for the motion parameters as functions of X and T . In general, this step involves averaging the conserved densities and fluxes over oscillations in the fast variables n and t .
4. Study the resulting system of conservation laws to determine the domains of hyperbolicity and ellipticity.

In the paper [25], this procedure was carried out for harmonic plane wave solutions of a family of lattice equations given by

$$(1.1) \quad i\partial_t A_n + f(|A_n|^2)(A_{n+1} + A_{n-1}) + F(|A_n|^2)A_n = 0,$$

where n is an integer index, A_n is a complex function of time, and f and F are real valued functions. Since the family of harmonic plane waves is described by two parameters, an amplitude ρ and a wavenumber k , the result of the averaging procedure is a set of two coupled first order partial differential equations in X and T describing the evolution of $\rho(X, T)$ and $k(X, T)$. Since in this case there are only two equations, they may in principle be cast into Riemann invariant form whenever they are hyperbolic. This fact was exploited in [25] in the study of data for which the modulation equations are hyperbolic at $T = 0$ to determine conditions on the data sufficient to prevent the equations from dynamically changing type from hyperbolic to elliptic.

In studying the macroscopic behavior of solutions more complicated than harmonic plane waves, serious problems, some merely technical but some analytically substantial, arise in the first, second, and fourth steps. More complicated exact solution families are very difficult to find⁴, and the system (1.1) is known to have in general only two local conservation laws, causing the procedure to fail in the second step for any solution families described by more than two real parameters. Even if modulation equations could be obtained by other means⁵,

⁴In seeking multiphase wavetrains in a spatially discrete system such as (1.1), one first makes the Ansatz $A_n(t) = \exp(i\theta_0)B(\theta_1, \dots, \theta_N)$ where $\theta_i = k_i n - \omega_i t$ and B is taken to be periodic with period 2π in each of its arguments. This leads to the problem of describing periodic solutions of a nonlinear functional (partial) differential equation for B with k_i and ω_i as parameters. The theory of functional differential equations is not as complete as that of ordinary and partial differential equations, even at the level of existence of solutions. The book by Hale and Lunel [23] gives a very good introduction to the difficulties that can arise in the analysis of functional (ordinary) differential equations. Chapter 11 in particular deals with periodic solutions to autonomous equations. In wave equations that are spatially continuous, the situation is somewhat better, since the equation for B is a partial differential equation, and not a functional equation. However, even in this case, strategies for dealing with finding multiphase waves in the spatially continuous case have been primarily perturbative and numerical, as discussed by Ablowitz [1].

⁵It is also possible to obtain modulation equations for parametrized families of waves in systems that can be described by a variational principle involving an action functional and its

the analytical advantage of putting the modulation equations into Riemann invariant form can only be gained when the family of microstates is described by no more than two parameters (in the absence of any additional structure). Thus, the kind of change-of-type analysis carried out in [25] cannot be carried out for modulation equations in the general case. But even more importantly, we expect that the complete nonequilibrium thermodynamical description of a typical system of the form (1.1) does not consist of microscopic behavior that is controlled by local constants of motion at all, but rather consists of microscopic motion that is irregular and may need to be described statistically. Understanding the long range correlations in a spatially and temporally chaotic dynamical system is a different program entirely, requiring other methods.

However, it is possible to learn a great deal by concentrating on special cases of the general system of lattice equations (1.1) for which the methods outlined in the four steps above work very well. For example, when $f(\rho) = 1 \pm \rho$ and $F(\rho) = -2$, the family of lattice equations becomes the Ablowitz-Ladik equations

$$(1.2) \quad i\partial_t A_n + (1 \pm |A_n|^2)(A_{n+1} + A_{n-1}) - 2A_n = 0.$$

This system of equations is a discrete analog of the nonlinear Schrödinger partial differential equation, and like its continuum limit partner, it is known to be integrable by means of an inverse spectral transform, and has been dealt with in the whole line [4] and periodic [6] cases. In principle, the methods of integrable systems may be used to construct very broad classes of multiphase wavetrain solutions of arbitrary complexity, which adhere to a macroscopic modulational description also available within the framework of the integrability. As we will show, this modulational description is a set of first order partial differential equations for quantities such as wavenumbers, frequencies, and amplitudes, that vary on the slow scales of X and T . In integrable systems similar to the Ablowitz-Ladik system, these modulation equations have been shown to correctly describe emergent macroscopic phenomena. For example, the long scale description of the nonequilibrium region behind the shock front in the Toda shock problem, as described by Venakides, Deift, and Oba [49] and constructed as a long time limit of local averages in the exact shock solution, is equivalently given by the formal modulation equations provided by Bloch and Kodama [5]. These modulation equations provide a more generic description of macroscopic behavior than do local averages constructed from a particular exact solution because the modulation equations are derived in the case of an arbitrary local solution in the class of multiphase wavetrains. Furthermore, the modulation equations have the advantage that they may always be written in Riemann invariant form. Thus,

associated Lagrangian function. The Lagrangian is averaged holding the parameters fixed, and an variational principle is postulated for the parameters using the averaged Lagrangian integrated over X and T as the action functional. This approach can give formal results even when there are too few conservation laws to permit one to proceed as described in the four steps above. The averaging of Lagrangian functionals to derive modulation equations was first done by Whitham; a good description appears in his book [50]. Furthermore, formal modulation equations may be developed by directly employing the method of multiple scales and an appropriate solution Ansatz, as discussed by Ablowitz and Benney [2].

it is desirable to understand the multiphase wavetrains and their corresponding modulational descriptions in the special case of the integrable Ablowitz-Ladik system as a benchmark from which perturbative methods may be used to extend the theory as much as possible to generic systems of the form (1.1).

Henceforth, we will be concerned with the construction of some solutions to the system of nonlinear ordinary differential equations given by

$$\begin{aligned}
 (1.3) \quad & -i\partial_t Q(n,t) - \left[Q(n+1,t) - 2Q(n,t) + Q(n-1,t) \right] \\
 & + Q(n,t)R(n,t) \left[Q(n+1,t) + Q(n-1,t) \right] = 0, \\
 & -i\partial_t R(n,t) + \left[R(n+1,t) - 2R(n,t) + R(n-1,t) \right] \\
 & - R(n,t)Q(n,t) \left[R(n+1,t) + R(n-1,t) \right] = 0,
 \end{aligned}$$

where n is an integer index specifying location in the lattice, and t is a continuous time variable. This system of equations is a complexified version of (1.2). The latter is recovered by taking $R(n,t) = \pm Q(n,t)$ for real t . It is easier to work with the complexified system (1.3) since the construction that follows produces complex potentials R and Q that are related only by being a solution pair to (1.3); the symmetries that guarantee that $R(n,t) = \pm Q(n,t)$ are imposed only after the relationship between the data that we will use to generate a solution and the solution (Q, R) itself is made transparent.

Our objective is not to solve the initial value problem, only to produce a class of multiphase wave solutions to (1.3) using methods of algebraic geometry. However, we will show that this class of solutions includes almost all spatially periodic solutions⁶ of (1.3), and we will provide a closed form formula for the solution to the initial value problem for a dense class of initial data in the case of periodic boundary conditions. Although an algorithm for computing the solution to the periodic initial value problem has been given [6], a formula for the solution has not appeared in the literature.

⁶This is an expected result, placing the periodic initial value problem for the Ablowitz-Ladik equations in natural context with other integrable systems. The initial value problem for integrable wave equations with periodic boundary conditions is solved in terms of potentials constructed using Riemann surfaces (possibly having infinite genus) because in each case, a solution of such an equation is a potential in a linear operator (depending upon a complex spectral parameter, say, λ) that is second order (or higher order) in the independent spatial variable. The fundamental spectral quantity in the periodic case is the Floquet multiplier, which is naturally interpreted as a function on a multiple cover of the λ plane, where the multiplicity is equal to the order of the corresponding linear operator. For examples, see Kac and van Moerbeke [28] regarding the Toda lattice, Ablowitz and Ma [39] regarding the nonlinear Schrödinger equation, and McKean and Trubowitz [40] regarding the Korteweg-de Vries equation. There is only one special case in which we do not provide the solution to the periodic initial value problem in terms of finite genus solutions — the case when there is some lattice point n such that $Q(n,0)R(n,0) = 1$. Basically, that point decouples from the rest of the lattice, as can be seen from (1.3). Reasons for considering this as a special case are provided in Appendix A.

Let us begin to describe the scheme we will employ to construct this class of solutions. Ablowitz and Ladik [4] observed that the system (1.3) is the simultaneous solvability condition for the spatial linear problem

$$(1.4) \quad \mathbf{v}(n+1, t, \lambda) = \mathbf{L}(n, t, \lambda)\mathbf{v}(n, t, \lambda),$$

where

$$(1.5) \quad \mathbf{L}(n, t, \lambda) = \begin{bmatrix} \lambda & Q(n, t) \\ R(n, t) & \lambda^{-1} \end{bmatrix},$$

and the temporal linear problem

$$(1.6) \quad -i\partial_t \mathbf{v}(n, t, \lambda) = \mathbf{B}(n, t, \lambda)\mathbf{v}(n, t, \lambda),$$

where

$$(1.7) \quad \mathbf{B}(n, t, \lambda) = \begin{bmatrix} \lambda^2 - 1 - Q(n, t)R(n-1, t) & Q(n, t)\lambda - Q(n-1, t)\lambda^{-1} \\ R(n-1, t)\lambda - R(n, t)\lambda^{-1} & 1 + R(n, t)Q(n-1, t) - \lambda^{-2} \end{bmatrix}.$$

The solutions are two-component vectors $\mathbf{v} = (v_1, v_2)^T$, and λ is a complex parameter on which the solutions implicitly depend. For historical reasons, the functions $Q(n, t)$ and $R(n, t)$ appearing in these two linear problems are called *potentials*⁷. This representation⁸ is purely local in n and t , so it works regardless of any particular boundary conditions imposed on the potentials $Q(n, t)$ and $R(n, t)$.

There is a redundancy in these two linear problems, based upon the substitution of $-\lambda$ for λ , that has been observed ever since the introduction of (1.4) as a scattering problem in [4]. When Ablowitz and Ladik developed the inverse scattering scheme for (1.4) in order to reconstruct solutions (Q, R) of (1.3) that decay rapidly as $|n| \rightarrow \infty$, they observed that the reflection coefficients were

⁷The first problem treated by the inverse scattering method, the Korteweg-de Vries equation [22], was associated (in analogy with the association between the nonlinear system (1.3) and the spatial linear problem (1.4)) with a stationary Schrödinger equation for a wavefunction ψ having energy λ . The solution $u(x, t)$ of the Korteweg-de Vries equation, interpreted as a function of x with a parameter t , appears as the potential energy function in this Schrödinger equation.

⁸A representation of a nonlinear equation for potential functions as a compatibility condition for a spatial linear problem with a temporal linear problem is called a *zero-curvature* representation [18]. This nomenclature comes from the interpretation of the linear problems as representing infinitesimal parallel translations of a vector \mathbf{V} in space and time with respect to an affine connection parametrized by the potentials. The curvature tensor of this connection vanishes if the two linear problems are compatible.

odd functions of λ ; this had the consequence of causing the discrete spectrum of (1.4) to come in positive-negative pairs and the continuous spectrum of (1.4) to have the same symmetry so that the reconstruction involved a contour integral over only that part of the unit circle lying in the right half λ plane [4]. Similarly, Bogolyubov and Prikarpat-skii observed that the elements of the monodromy matrix derived from (1.4) in the case of spatially periodic potentials had either even or odd parity in λ ; this led them to construct solutions from Riemann surfaces whose branch points, like the discrete eigenvalues of the whole line problem, came in positive-negative pairs [6].

Our experience in using the linear problems (1.4) and (1.6) to construct general solutions to (1.3) using Baker-Akhiezer functions led us to analogous results. We had wanted to proceed by directly constructing a simultaneous solution $\mathbf{v} = (v_1, v_2)^T$ of (1.4) and (1.6) for all values of the spectral parameter λ from which the consistent potentials $Q(n, t)$ and $R(n, t)$ could be deduced; the general approach was to parallel those of Dubrovin [11], Krichever [29, 30], and Previato [43] who treated similar problems in other nonlinear systems. For example, to find solutions to the Toda lattice equations⁹,

$$(1.8) \quad \partial_t^2 q(n, t) = \exp(q(n+1, t) - q(n, t)) - \exp(q(n, t) - q(n-1, t)),$$

one considers the two linear problems (whose consistency condition is (1.8))

$$(1.9) \quad \begin{aligned} v_1(n+1, t) &= (\partial_t q(n, t) + \lambda)v_1(n, t) + \exp(q(n, t))v_2(n, t), \\ v_2(n+1, t) &= -\exp(-q(n, t))v_1(n, t), \end{aligned}$$

and

$$(1.10) \quad \begin{aligned} \partial_t v_1(n, t) &= -\exp(q(n, t)), \\ \partial_t v_2(n, t) &= \exp(-q(n-1, t))v_1(n, t) + \lambda v_2(n, t), \end{aligned}$$

where we have used the notation of Faddeev and Takhtajan [18]. These linear problems are singular when λ is large; asymptotic analysis gives two possible dominant balances for λ near ∞ . If λ is reinterpreted as a sheet projection function of a hyperelliptic Riemann surface with two points over ∞ given by ∞^+ and ∞^- , one can ask that the functions v_1 and v_2 be functions on the

⁹The periodic problem for the Toda lattice was first solved in terms of Abelian integrals by Kac and van Moerbeke [28], and a development of solutions using Baker-Akhiezer functions is given by Krichever [30].

Riemann surface with behavior given (in terms of a point P on the surface) by

$$(1.11) \quad \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \lambda^n \exp(\lambda^{-1}t) \begin{pmatrix} \mathcal{O}(1) \\ \mathcal{O}(\lambda^{-1}) \end{pmatrix}, \quad P \rightarrow \infty^+,$$

$$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \lambda^{-n} \exp(\lambda t) \begin{pmatrix} \mathcal{O}(\lambda^{-1}) \\ \mathcal{O}(1) \end{pmatrix}, \quad P \rightarrow \infty^-.$$

It turns out that by insisting further that the functions v_1 and v_2 have certain poles on the Riemann surface, and by appropriately normalizing them, one can conclude that the two functions are unique, and that they solve the two linear problems, as long as $q(n, t)$ is related to a certain leading coefficient near ∞^+ . There are expressions for v_1 and v_2 in terms of Riemann theta functions, and from these, one derives an expression for a solution of the Toda problem, $q(n, t)$. This is in fact the formula that Bloch and Kodama inserted into the local conservation laws prior to averaging and passing to the continuum limit in the parameters of the potential $q(n, t)$ (the parameters characterizing the Riemann surface employed in the construction), eventually obtaining formal modulation equations for the moduli of a hyperelliptic Riemann surface [5].

In contrast, when we studied the linear problems (1.4) and (1.6) for the Ablowitz-Ladik equations, the construction was not so simple. The asymptotic analysis led us to consider Riemann surfaces with four punctures (∞^+ , ∞^- , 0^+ , and 0^-), and by specifying the poles of v_1 and v_2 and normalizing appropriately, we were able to show that the two functions are unique. Where the construction deviated essentially from that for the Toda lattice was at the point where one concludes that v_1 and v_2 solve (1.4) and (1.6). In fact, we were unable to solve these linear problems with the functions v_1 and v_2 without further imposing the symmetry

$$(1.12) \quad \begin{aligned} v_1(n, t, \sigma(P)) &= (-1)^n v_1(n, t, P), \\ v_2(n, t, \sigma(P)) &= (-1)^{n+1} v_2(n, t, P), \end{aligned}$$

where P denotes a point on the Riemann surface, and σ denotes the unique involution of the surface that covers $\lambda \mapsto -\lambda$ and preserves the sign of the square root function that defines the surface. We found that the involution σ only exists on Riemann surfaces of odd genus whose branch points enjoy the symmetry of $\lambda \mapsto -\lambda$. In this case, we were able to obtain solutions to (1.3) as in the Toda problem, however due to the symmetry, the formulas were reducible to simpler formulas involving what appeared at the time to be an auxiliary Riemann surface

of smaller genus. The details of the construction on Riemann surfaces of odd genera and the corresponding reduction can be found in the dissertation of Miller [41].

These clues led us to believe that the Lax pair (1.4) and (1.6) could be replaced by another Lax pair that takes into account the $\lambda \mapsto -\lambda$ symmetry. The construction of solutions using this new Lax pair would then proceed in the same manner as that for the Toda lattice, without the need to impose any additional symmetries. Such a Lax pair is easy to find by introducing a simple transformation of \mathbf{v} ; this induces a corresponding transformation in the linear problems (1.4) and (1.6) yielding the new Lax pair. Consider a new vector function \mathbf{u} defined by

$$(1.13) \quad \begin{aligned} u_1(n, t) &= \lambda^n v_1(n, t), \\ u_2(n, t) &= \lambda^{n+1} v_2(n, t). \end{aligned}$$

The linear problems (1.4) and (1.6) are then correspondingly transformed into equations for \mathbf{u} :

$$(1.14) \quad \mathbf{u}(n+1, t, z) = \tilde{\mathbf{L}}(n, t, z)\mathbf{u}(n, t, z),$$

where

$$(1.15) \quad \tilde{\mathbf{L}}(n, t, z) = \begin{bmatrix} z & Q(n, t) \\ zR(n, t) & 1 \end{bmatrix},$$

and

$$(1.16) \quad -i\partial_t \mathbf{u}(n, t, z) = \tilde{\mathbf{B}}(n, t, z)\mathbf{u}(n, t, z),$$

where

$$(1.17) \quad \begin{aligned} &\tilde{\mathbf{B}}(n, t, z) \\ &= \begin{bmatrix} z-1-Q(n, t)R(n-1, t) & Q(n, t)-z^{-1}Q(n-1, t) \\ zR(n-1, t)-R(n, t) & 1-z^{-1}+R(n, t)Q(n-1, t) \end{bmatrix}. \end{aligned}$$

In these equations, the spectral parameter z has been substituted for λ^2 , thereby eliminating any redundancy associated with the symmetry $\lambda \mapsto -\lambda$. The linear equations (1.14) and (1.16) make up the transformed Lax pair that we will use almost exclusively in the remainder of this paper¹⁰; thus, we will henceforth

¹⁰In Appendix B we will use the original Lax pair of Ablowitz and Ladik to describe allowable branch point configurations for periodic solutions of the focusing Ablowitz-Ladik equations. We will also use a unimodular version of this original Lax pair to discuss branch point configurations in the defocusing problem.

drop the tildes on the transformed matrices. These linear problems may appear to be less symmetrical than (1.4) and (1.6) in their dependence on the spectral parameter, but they will allow a streamlined construction of solutions to (1.3).

This paper shows how to use this new Lax pair to find a class of algebro-geometric solutions to (1.3) that are associated with finite genus Riemann surfaces, first assuming distinct branch points, and subsequently relaxing that constraint. Using these solutions, we will solve the spatially periodic (and later twist-periodic¹¹ initial value problem. We will also show how to select the solutions that satisfy reality conditions and thus solve the integrable discretization of the cubic nonlinear Schrödinger equation (1.2). Finally, we will develop the modulation equations for multiphase waves in this spatially discrete setting. Throughout, we will emphasize the similarities and differences between the Ablowitz-Ladik equations and the (continuum) nonlinear Schrödinger equation. The main differences will be:

- The spatially twist-periodic potentials of the Ablowitz-Ladik equations always correspond to (possibly singular) Riemann surfaces of finite genus, whereas most twist-periodic potentials of the nonlinear Schrödinger equation correspond to surfaces of infinite genus,
- In the defocusing case of the Ablowitz-Ladik equations, there are nonempty classes of both stable and unstable multiphase waves of each genus, whereas the defocusing nonlinear Schrödinger equation has only stable multiphase waves, and
- In the complex Ablowitz-Ladik equations, each hyperelliptic Riemann surface and nonspecial divisor gives rise to *two* distinct algebro-geometric solutions, whereas in the complexified nonlinear Schrödinger equation there is only one solution for each Riemann surface and divisor.

The structure of the remainder of this paper is as follows. First, Section 2 contains a detailed analysis of the asymptotic behavior of solutions to the streamlined Lax pair (1.14) and (1.16) in the singular limits of $z \rightarrow 0$ and $z \rightarrow \infty$, and a development of the theory of functions on hyperelliptic Riemann surfaces that behave asymptotically for large and small z like solutions of these two linear problems. At the end of this section we will have found a pair of functions u_1 and u_2 that are likely candidates for a simultaneous solution of (1.14) and (1.16). Then, in Section 3, the function theory will be used to *prove* that the candidate functions u_1 and u_2 actually solve both linear problems — not just asymptotically, but globally — as long as the potentials Q and R are related to certain leading coefficients in the expansions of u_1 and u_2 near singularities. Since, for these potentials, the problems (1.14) and (1.16) are obviously consistent, Q and R are solutions, constructed from a finite set of complex spectral data, to the Ablowitz-Ladik equations (1.3). The final sections of the paper are devoted to the

¹¹Twist-periodic boundary conditions on the potentials Q and R are defined by choosing an angle θ and an integer N and insisting that $Q(n+N) = \exp(i\theta)Q(n)$ and $R(n+N) = \exp(-i\theta)R(n)$ for all n .

analysis of the constructed solutions. Section 4 is dedicated to the reduction of the complex Ablowitz-Ladik equations (1.3) to the focusing and defocusing cases of the integrable discrete nonlinear Schrödinger equation (1.2). Realizing this reduction at the level of the finite genus potentials requires specifying conditions on the spectral data equivalent to the reality conditions on the potentials themselves: $R(n, t) = \pm \overline{Q(n, t)}$. The reality conditions are most easily expressed in terms of spectral data using polynomial (in z) squared eigenfunctions which are introduced in Section 4 as a tool. Section 5 deals with the solutions to Ablowitz-Ladik equations (1.3) restricted to be periodic in n . These periodic solutions are characterized in terms of their spectral data, and the initial value problem for periodic boundary conditions is coincidentally solved in terms of formulas for $Q(n, t)$ and $R(n, t)$ first presented at the end of Section 3. Then, the periodic theory will be extended to handle cases that are only twist-periodic in n . With the elementary description of the finite genus exact solutions to (1.3) complete, Section 7 undertakes a description of approximate solutions to (1.3) that appear locally in n and t to be exact finite genus solutions, but that vary slowly on the slow scales $X = hn$ and $T = ht$ for a small lattice spacing h . Formal modulation equations are derived, and a gauge symmetry group is introduced that leaves the modulation equations invariant, but alters their appearance. The gauge group is used to simplify the modulation equations in the focusing and defocusing cases. The resulting simplified modulation equations can locally be either hyperbolic or elliptic, depending upon the local branch point configuration, but if initially hyperbolic for all X , the equations cannot evolve to become elliptic as long as the data are smooth in X . In Section 8, we summarize the results and discuss open problems. Some technical matters are delegated to the appendices. Appendix A contains the description of solutions obtained from Riemann surfaces in the singular limit of coinciding branch points; this material is more technical, but is used in the complete solution of the initial value problem for periodic boundary conditions given in Section 5. Finally, in Appendix B we prove two theorems concerning allowable branch point configurations in the periodic focusing and defocusing cases. The first theorem states that in the focusing case the periodic and antiperiodic Floquet eigenvalues may only lie on the unit circle if they have multiplicity greater than one. The second states that for defocusing potentials satisfying $|Q(n, t)| < 1$ for all n , the periodic and antiperiodic Floquet eigenvalues must lie on the unit circle. In both cases, we suggest how the results might be extended from the strictly periodic to the general finite genus cases. The paper is intended to be somewhat self-contained; however, any technical details not presented here can be found in Dubrovin [11].

2. Singular asymptotics of the linear problems. The Baker-Akhiezer function.

In order to build a solution $\mathbf{u} = (u_1, u_2)^T$ for every value of the complex parameter z , the behavior of the two linear problems (1.14) and (1.16) near the singular values $z = 0$ and $z = \infty$ must, in particular, be examined. First,

consider the spatial problem (1.14) in the case of z small. There are two dominant balances possible:

$$(2.1) \quad \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = z^n \begin{pmatrix} \mathcal{O}(1) \\ \mathcal{O}(z) \end{pmatrix}, \quad \text{and} \quad \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} \mathcal{O}(1) \\ \mathcal{O}(1) \end{pmatrix},$$

Similarly, in the case of z large, there are two dominant balances possible:

$$(2.2) \quad \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = z^n \begin{pmatrix} \mathcal{O}(1) \\ \mathcal{O}(1) \end{pmatrix}, \quad \text{and} \quad \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} \mathcal{O}(1) \\ \mathcal{O}(z) \end{pmatrix}.$$

These balances will be modified by overall factors depending on z and t obtained by substituting them into the temporal linear problem (1.16) and seeking a dominant balance. The final result of this procedure is the pair of dominant balances valid for z near 0,

$$(2.3) \quad \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = z^n \begin{pmatrix} \mathcal{O}(1) \\ \mathcal{O}(z) \end{pmatrix}, \quad \text{and} \quad \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \exp(i(1 - z^{-1})t) \begin{pmatrix} \mathcal{O}(1) \\ \mathcal{O}(1) \end{pmatrix},$$

and the pair of dominant balances valid for z near ∞ ,

$$(2.4) \quad \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = z^n \exp(i(z - 1)t) \begin{pmatrix} \mathcal{O}(1) \\ \mathcal{O}(1) \end{pmatrix}, \quad \text{and} \quad \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} \mathcal{O}(1) \\ \mathcal{O}(z) \end{pmatrix}.$$

If the function \mathbf{u} depending on the spectral parameter z is to include all four of the above kinds of behavior, it is necessary to enlarge the scope of the problem. Notice that reinterpreting the spectral parameter z as a sheet projection function of a two-sheeted (hyperelliptic) Riemann surface Γ and demanding that the solution \mathbf{u} live on Γ yields the possibility of two different kinds of asymptotic behavior near each of $z = 0$ and $z = \infty$. Thus, to accommodate the function \mathbf{u} , we introduce the smooth Riemann surface Γ as the algebraic curve of genus $g \geq 0$ associated with the relation

$$(2.5) \quad y^2 = \prod_{i=1}^{2g+2} (z - z_i).$$

If the z_i are all finite and nonzero then $z = 0$ and $z = \infty$ will each have two preimages. Labeling one of the preimages on Γ of $z = 0$ as $P = 0^+$ and the other preimage as $P = 0^-$, and similarly labeling preimages of $z = \infty$ as $P = \infty^+$ and $P = \infty^-$, we insist that our solution be a function of n , t , and $P \in \Gamma$ having the

asymptotic behavior¹²:

$$\begin{aligned}
 \begin{pmatrix} u_1(n, t, P) \\ u_2(n, t, P) \end{pmatrix} &= z^n \begin{pmatrix} \mathcal{O}(1) \\ \mathcal{O}(z) \end{pmatrix}, & P \rightarrow 0^-, \\
 \begin{pmatrix} u_1(n, t, P) \\ u_2(n, t, P) \end{pmatrix} &= \exp(i(1 - z^{-1})t) \begin{pmatrix} \mathcal{O}(1) \\ \mathcal{O}(1) \end{pmatrix}, & P \rightarrow 0^+, \\
 \begin{pmatrix} u_1(n, t, P) \\ u_2(n, t, P) \end{pmatrix} &= z^n \exp(i(z - 1)t) \begin{pmatrix} \mathcal{O}(1) \\ \mathcal{O}(1) \end{pmatrix}, & P \rightarrow \infty^+, \\
 \begin{pmatrix} u_1(n, t, P) \\ u_2(n, t, P) \end{pmatrix} &= \begin{pmatrix} \mathcal{O}(1) \\ \mathcal{O}(z) \end{pmatrix}, & P \rightarrow \infty^-.
 \end{aligned}
 \tag{2.6}$$

It is important to observe that there are four possible ways to label the two points over $z = 0$ as 0^\pm and the two points over $z = \infty$ as ∞^\pm , and that these four different labeling choices lead to four different kinds of asymptotic behavior on Γ as described by (2.6). However, only two of these four labelings are essential, since the simultaneous permutation of the two zeros and of the two infinities leads to a function $\tilde{\mathbf{u}}$ that is just the hyperelliptic involute of the original function \mathbf{u} . So, without loss of generality, the distant point on Γ that corresponds to large positive y for large positive z will be denoted ∞^+ ; the other point on Γ over $z = \infty$ will be denoted ∞^- . Having made this choice, however, we *cannot* choose an arbitrary labeling of the two points over $z = 0$ without loss of generality. The two possible labelings for the points over $z = 0$ lead to two genuinely different kinds of asymptotic behavior; these in turn will lead to two different classes of solutions to (1.3). Such a partitioning of solutions will occur in every nonlinear system that arises as a zero-curvature condition of 2-by-2 linear problems that are singular for more than one value of the spectral parameter. Thus, while the solutions to the the Ablowitz-Ladik equations come in pairs due to this labeling ambiguity, the solutions to the continuum limit nonlinear Schrödinger equation do not, since as described by Previato [43], the 2-by-2 linear problems giving rise to the nonlinear Schrödinger equation are singular only when the spectral parameter is large; a solution of the linear problems on the Riemann surface Γ would only be singular near the points ∞^\pm and permutation of these two points corresponds to hyperelliptic involution of the solution $\mathbf{u}(P)$. This is one point in which the description of finite genus solutions for the Ablowitz-Ladik system (1.3) differs significantly from that for the nonlinear Schrödinger equation.

¹²To the reader familiar with Baker-Akhiezer function solutions of Lax pairs it may seem incorrect to refer to the point ∞^- as a singular point or a puncture on the surface Γ since there is no essential singularity (or the discrete version thereof, z^n) there. The relative magnitudes of the two components u_1 and u_2 are, however, determined by the two linear problems at this point independently of the potentials, and thus we must specify this behavior of the common solution \mathbf{u} . Furthermore, the corresponding solution \mathbf{v} of the original Lax pair does indeed have a singularity of the form λ^{-n} at the point ∞^- .

In the context of the solutions of the complex Ablowitz-Ladik equations (1.3), it is not difficult to unify the two different components of the solution manifold. In fact they may be analytically continued into each other through the following device. Rather than thinking of specifying $2g + 2$ distinct branch points z_i (that is, an element of the set¹³ $(\mathbb{C}^{*(2g+2)} \setminus \Delta)/\mathcal{S}_{2g+2}$) in order to specify the Riemann surface Γ , consider specifying an element of the *double cover* of $(\mathbb{C}^{*(2g+2)} \setminus \Delta)/\mathcal{S}_{2g+2}$ entering through the quantity

$$(2.7) \quad \eta = \sqrt{\prod_{i=1}^{2g+2} z_i}.$$

We denote this set of genus g Riemann surfaces with choice of η as \mathcal{M}_g , and from now on, by the symbol Γ , we will mean an element of \mathcal{M}_g . Appendix A includes a discussion of how to include the diagonals Δ in the set of branch points, resulting in a slightly bigger set \mathcal{M}'_g ; an element Γ of this enlarged set consists of a Riemann surface of genus $\tilde{g} \leq g$ with choice of η , and a divisor \mathcal{E} on \mathbb{C}^* . We will occasionally abuse notation by referring to the Riemann surface determined by $\Gamma \in \mathcal{M}'_g$ by the same name, Γ . Concretely specifying an element of \mathcal{M}_g involves the following steps. First, choose a point (ξ, η) with $\xi \in \mathbb{C}^*$ from the Riemann surface of the algebraic relation

$$(2.8) \quad \eta^2 = \xi.$$

Then select $2g + 2$ distinct points z_1, \dots, z_{2g+2} from \mathbb{C}^* that satisfy

$$(2.9) \quad \prod_{i=1}^{2g+2} z_i = \xi.$$

These points determine the Riemann surface. Finally, the labeling of the two points over $z = 0$ is determined in this way: 0^+ is the point over $z = 0$ such that $y(0^+) = \eta$, and correspondingly 0^- is the point over $z = 0$ such that $y(0^-) = -\eta$. Thus, given a surface Γ with one particular labeling of the points over $z = 0$, one may analytically deform the surface in such a way that the branch points return to their original configuration, but that the points 0^\pm have been exchanged. An elementary deformation that achieves this result consists of holding all but one branch point fixed, and moving the remaining branch point through a closed loop encircling $z = 0$, the point of monodromy. In the course of this deformation, the branch points have all returned to their original configuration, but the quantity η has changed sign. A similar kind of double covering occurs in the study of periodic solutions to the sine-Gordon equation, because the simultaneous solution to the two linear problems giving rise to the (complex) sine-Gordon equation has singularities when the spectral parameter

¹³ \mathbb{C}^* denotes the set of all finite nonzero complex numbers, \mathcal{S}_m denotes the symmetric group of permutations of m symbols, and Δ denotes the *diagonal*, consisting of sets of branch points where $z_i = z_j$ for some i and j .

is both small and large. As explained by Ercolani, Forest and McLaughlin [14], the solution to the sine-Gordon equation involves the square root of the product of the finite nonzero branch points. Thus, each Riemann surface supports two classes of complex solutions, one for each sign of the square root, and these two classes may be continued into each other by the same device described above. However, when reality conditions are imposed on the solution, these conditions correspondingly introduce constraints on the allowable Riemann surfaces; these constraints sever the analytic connection between the two components. We will show at the end of Section 4 that the same is true when the complex Ablowitz-Ladik equations (1.3) are reduced to the focusing and defocusing cases (1.2).

It is time to develop the theory of functions on Γ with asymptotic behavior given by (2.6). To begin with, we say that a function $\psi(n, t, P)$ of $P \in \Gamma$ and depending parametrically on an integer n and a real number t has *singularities of type 1* if it has the following behavior:

$$\begin{aligned}
 (2.10) \quad & \psi(n, t, P) = a_1(n, t)z^n(1 + \mathcal{O}(z^{-1})) \exp(i(z - 1)t), & P \rightarrow \infty^+, \\
 & \psi(n, t, P) = b_1(n, t)(1 + \mathcal{O}(z^{-1})), & P \rightarrow \infty^-, \\
 & \psi(n, t, P) = c_1(n, t)z^n(1 + \mathcal{O}(z)), & P \rightarrow 0^-, \\
 & \psi(n, t, P) = d_1(n, t)(1 + \mathcal{O}(z)) \exp(i(1 - z^{-1})t), & P \rightarrow 0^+,
 \end{aligned}$$

where $a_1, b_1, c_1,$ and d_1 are arbitrary functions $n, t,$ not identically zero. Similarly, a function $\phi(n, t, P)$ has *singularities of type 2* if it has the following behavior:

$$\begin{aligned}
 (2.11) \quad & \phi(n, t, P) = a_2(n, t)z^n(1 + \mathcal{O}(z^{-1})) \exp(i(z - 1)t), & P \rightarrow \infty^+, \\
 & \phi(n, t, P) = b_2(n, t)z(1 + \mathcal{O}(z^{-1})), & P \rightarrow \infty^-, \\
 & \phi(n, t, P) = c_2(n, t)z^{n+1}(1 + \mathcal{O}(z)), & P \rightarrow 0^-, \\
 & \phi(n, t, P) = d_2(n, t)(1 + \mathcal{O}(z)) \exp(i(1 - z^{-1})t), & P \rightarrow 0^+,
 \end{aligned}$$

where again a_2, b_2, c_2, d_2 are arbitrary nonzero functions. Let $\mathcal{D} = P_1 + \dots + P_g$ be a nonspecial integral divisor¹⁴ on Γ of degree g .

¹⁴The points $\mathcal{D} = P_1 + \dots + P_g$ must make up what is called a *nonspecial* divisor on Γ . This concretely means that if $P^+(\mu)$ is one of the g points P_j , then $P^-(\mu)$ cannot be, and vice-versa, where μ is not one of the branch points and $P^+(\mu)$ and $P^-(\mu)$ refer to the two distinct points over $z = \mu$. Such a condition of nonspeciality ensures that the function $\Theta(A(P) - A(\mathcal{D}) - \mathcal{K})$ does not vanish identically, but rather has exactly g zeros on Γ at the points P_1, \dots, P_g . Most sets of g points make up nonspecial divisors; accordingly we will sometimes refer to the g points of such a nonspecial divisor as being *in general position*.

If $\mathcal{D} = P_1 + \dots + P_g$ is the divisor of the poles (zeros) of a function u , then the poles (zeros) of u are confined to the points P_1, \dots, P_g , with allowable multiplicities at the point P_j determined by the number of times P_j appears in the formal sum \mathcal{D} . Divisor notation is reviewed in [11].

DEFINITION 2.1. $\Lambda_1(\Gamma, \mathcal{D})$ is the linear space of all functions on Γ with singularities of type 1 and otherwise meromorphic, having \mathcal{D} as the divisor of the poles.

DEFINITION 2.2. $\Lambda_2(\Gamma, \mathcal{D})$ is the linear space of all functions on Γ with singularities of type 2 and otherwise meromorphic, having \mathcal{D} as the divisor of the poles.

Functions in $\Lambda_1(\Gamma, \mathcal{D})$ and $\Lambda_2(\Gamma, \mathcal{D})$ are called *Baker-Akhiezer functions*; a vector function whose first component is in Λ_1 and whose second component is in Λ_2 is called a *vector Baker-Akhiezer function*. It is a consequence of the complex structure of the Riemann surface Γ that the spaces $\Lambda_1(\Gamma, \mathcal{D})$ and $\Lambda_2(\Gamma, \mathcal{D})$ are one-dimensional in the sense that any two elements of either of these two spaces are proportional via a function of n and t only. We make this idea concrete in the following two lemmas.

LEMMA 2.3. $\dim \Lambda_1(\Gamma, \mathcal{D}) = 1$.

Proof: Begin by supposing the existence of a nonzero function $\psi^{(0)}$ in the space $\Lambda_1(\Gamma, \mathcal{D})$ that has minimal order at the points of the divisor \mathcal{D} . This means that if the point P appears in the divisor \mathcal{D} with multiplicity m , then $\psi^{(0)}$ has a pole of order m at P . With this assumption it is possible to show that $\dim \Lambda_1(\Gamma, \mathcal{D}) = 1$. Let ψ be an arbitrary function in Λ_1 . Then, in view of the expansions (2.10), the quotient $\xi = \psi/\psi^{(0)}$ is holomorphic near the four singular points 0^\pm and ∞^\pm . The only poles of ξ on Γ lie in the nonspecial divisor of zeros of $\psi^{(0)}$ (since the poles of ψ and $\psi^{(0)}$ cancel) which has degree g . From the Riemann-Roch theorem, the dimension of the space of all such functions is exactly 1, and thus ξ is a constant function on Γ (that is, dependent on n and t only). Thus, $\psi \propto \psi^{(0)}$ for all ψ in Λ_1 . So if $\Lambda_1(\Gamma, \mathcal{D})$ contains a function like $\psi^{(0)}$, then $\Lambda_1(\Gamma, \mathcal{D})$ is exactly one dimensional.

Thus, to deduce that $\dim \Lambda_1(\Gamma, \mathcal{D}) = 1$, it must be shown that there exists a nonzero element of this space having minimal order at the points of the divisor \mathcal{D} . Define the function:

$$(2.12) \quad \begin{aligned} \psi^{(0)}(n, t, P) &= \frac{\Theta(A(P) - A(\mathcal{D}) - \mathcal{K} + \mathbf{U}n + \mathbf{V}t)}{\Theta(A(P) - A(\mathcal{D}) - \mathcal{K})} \\ &\times \exp\left(n \int_{P_0}^P \omega_{(3)} + t \int_{P_0}^P \omega_{(2)}\right). \end{aligned}$$

In this formula, A is the Abel mapping¹⁵, \mathcal{K} is the Riemann constant vector¹⁶ of Γ , and Θ is the Riemann theta function¹⁷ of the surface Γ . The path on Γ in the integrals in the exponent is chosen to be the same as the path in the Abel mapping. The differentials $\omega_{(3)}$ and $\omega_{(2)}$ on Γ are normalized to have vanishing integrals over all a -cycles of a canonical homology basis¹⁸. The differential $\omega_{(3)}$

¹⁵The Abel mapping is defined as $A : \Gamma \rightarrow \text{Jac}(\Gamma)$ by

$$A_i(P) \equiv \int_{P_0}^P \nu_i,$$

where $\{\nu_1 \dots \nu_g\}$ is the basis of holomorphic differentials on Γ dual to a canonical homology basis. In this paper, the base point P_0 is fixed to be one of the branch points of the surface Γ . The Abel map is extended by linearity to all divisors on Γ so that $A(D) = A(P_1) + \dots + A(P_g)$. In order for the Abel map to be single valued, the integrals must make sense regardless of the chosen path on Γ from P_0 to P . The value of the integral is thus considered modulo integrals over noncontractable closed loops, and this is what we mean by using the symbol “ \equiv ” in the definition. The corresponding range when the map is appropriately extended to all divisors is a multitorus in \mathbb{C}^g denoted by $\text{Jac}(\Gamma)$, the Jacobian variety of Γ . This nonlinear mapping is often used to linearize nonlinear differential equations. See [11] for details.

¹⁶The Riemann constant vector of a hyperelliptic Riemann surface gives the location of the zeros of the corresponding theta function in the Jacobian of Γ . Farkas and Kra [19] give a useful formula for the Riemann constant vector in terms of the Abel map and the branch points. Let the base point of the integral in the Abel map be the branch point z_1 . Then the Riemann constant vector is defined on $\text{Jac}(\Gamma)$ as

$$\mathcal{K} \equiv \sum_{i=1}^g A(z_{2i+1}),$$

where the g branch points z_{2i+1} are chosen (uniquely, up to permutation) so that $A(z_{2i+1})$ is an *odd half period*. See [19] for details.

¹⁷The Riemann theta function is a function of $\mathbf{w} \in \mathbb{C}^g$ depending parametrically on a g -by- g matrix \mathbf{M} having a negative definite real part. The function is defined by the Fourier series

$$\Theta(\mathbf{w}) = \sum_{\mathbf{n} \in \mathbb{Z}^g} \exp\left(\frac{1}{2} \mathbf{n}^T \mathbf{M} \mathbf{n} + \mathbf{n}^T \mathbf{w}\right).$$

The Riemann theta function of the Riemann surface Γ corresponds to choosing the matrix \mathbf{M} so that

$$M_{ij} = \oint_{b_i} \nu_j,$$

where the b_i refer to the g b -cycles of a canonical homology basis, and the ν_j refer to the elements of the basis of holomorphic differentials on Γ dual to the chosen homology basis. See [11] for details.

¹⁸A canonical homology basis consists of $2g$ noncontractable oriented loops on the surface Γ labeled as $a_1, \dots, a_g, b_1, \dots, b_g$, modulo smooth deformations of the loops. The loops must be independent, and must satisfy the intersection conditions

$$a_i \circ a_j = b_i \circ b_j = 0,$$

and

$$a_i \circ b_j = \delta_{ij},$$

where δ here denotes the Kronecker delta function. The symbol “ \circ ” denotes the canonical intersection number of the two loops; $x \circ y$ is computed as the number of times the cycle y intersects the cycle x , with the contribution being positive (negative) if y intersects x from the right (left). Each cycle x can be lifted by the Abel integration map to \mathbb{C}^g where we denote the corresponding vector by \mathbf{x} . The basis of holomorphic differentials dual to a canonical homology

is an Abelian differential of the third kind having simple poles at 0^- and ∞^+ with residues 1 and -1 respectively and no other singularities. The differential $\omega_{(2)}$ is an Abelian differential of the second kind having singularities only at the points ∞^+ and 0^+ where it has the expansions:

$$(2.13) \quad \begin{aligned} \omega_{(2)} &= (i + \mathcal{O}(z^{-2})) dz, & P \rightarrow \infty^+, \\ \omega_{(2)} &= (iz^{-2} + \mathcal{O}(1)) dz, & P \rightarrow 0^+. \end{aligned}$$

The function $\psi^{(0)}$ must first be shown to be a well defined function on Γ , since some ambiguity enters the expression for $\psi^{(0)}$ through the path in the Abel map, which is only determined modulo cycles on Γ . However, if the choice of path in the Abel map is changed by some cycle $\sum l_i a_i + m_i b_i$ (and correspondingly adjust the path in the exponent), the function transforms as¹⁹

$$(2.14) \quad \begin{aligned} \psi^{(0)} \rightarrow \exp\left(\sum_{i=1}^g \left[l_i \left(n \oint_{a_i} \omega_{(3)} + t \oint_{a_i} \omega_{(2)} \right) \right. \right. \\ \left. \left. + m_i \left(n \oint_{b_i} \omega_{(3)} + t \oint_{b_i} \omega_{(2)} - U_i n - V_i t \right) \right] \right) \psi^{(0)}. \end{aligned}$$

So, choosing the vectors \mathbf{U} and \mathbf{V} so that

$$(2.15) \quad U_i = \oint_{b_i} \omega_{(3)}, \quad V_i = \oint_{b_i} \omega_{(2)},$$

and using the fact that the two differentials have vanishing integrals over a -cycles, the exponential factor becomes unity and the function $\psi^{(0)}$ is well defined²⁰. It remains only to show that $\psi^{(0)}$ is in Λ_1 and has minimal order at the points of \mathcal{D} . The exponential factor gives $\psi^{(0)}$ singularities of type 1, and the solution

basis is the set of differentials of the first kind ν_k such that

$$\oint_{a_j} \nu_k = 2\pi i \delta_{jk}.$$

Thus, we have $\mathbf{a}_k = 2\pi i \mathbf{e}_k$, where the \mathbf{e}_k are the usual unit vectors in \mathbb{C}^g . Since we will be integrating differentials of the third kind with residues at ∞^+ , ∞^- , and 0^- , we must be aware that a smooth deformation of a cycle on Γ may introduce a residue contribution. The correct way to handle this ambiguity is to consider a homology basis on the surface Γ with punctures at these three points. Around a puncture at a point P , we place a small closed oriented loop c_P . Thus, the homology basis we use is $\{a_1, \dots, a_g, b_1, \dots, b_g, c_{\infty^+}, c_{\infty^-}, c_{0^-}\}$, where a cycle may not be deformed in such a way that it crosses a puncture. With this basis, loop integrals of differentials of the third kind will be well defined.

¹⁹This calculation requires the transformation law for Riemann theta functions as found, for example, in [11].

²⁰Strictly speaking, we must also consider changing the path by one of the cycles c_{∞^+} , c_{∞^-} , or c_{0^-} . These cycles, being contractable on Γ do not contribute to the Abel mapping, and the exponential factor contributes nothing because the third kind differential $\omega_{(3)}$ has integer residues.

of the Jacobi inversion problem in terms of Riemann theta functions gives the denominator *exactly* g zeros in the divisor \mathcal{D} , in accordance with Riemann's theorem [11]. This means that $\psi^{(0)}$ is a well defined function on Γ with the required behavior. Thus $\dim \Lambda_1(\Gamma, \mathcal{D}) = 1$.

LEMMA 2.4. $\dim \Lambda_2(\Gamma, \mathcal{D}) = 1$.

Proof: The proof that Λ_2 containing an appropriate element $\phi^{(0)}$ implies $\dim \Lambda_2 = 1$ is identical to the corresponding part of the above proof. Whereas the proof of Lemma 2.3 employs the function $\psi^{(0)}$, the proof of this lemma requires the existence of a nonzero function in $\Lambda_2(\Gamma, \mathcal{D})$ of minimal order at the points of \mathcal{D} :

$$(2.16) \quad \begin{aligned} \phi^{(0)}(n, t, P) &= \frac{\Theta(A(P) - A(\mathcal{D}) - \mathcal{K} + \mathbf{U}n + \mathbf{V}t + \mathbf{W})}{\Theta(A(P) - A(\mathcal{D}) - \mathcal{K})} \\ &\times \exp \left(n \int_{P_0}^P \omega_{(3)} + t \int_{P_0}^P \omega_{(2)} + \int_{P_0}^P \omega \right). \end{aligned}$$

Here, ω is the Abelian differential of the third kind on Γ having simple poles at 0^- and ∞^- with residues 1 and -1 respectively, and having vanishing integrals over all a -cycles in the homology basis of Γ . Choosing the vector \mathbf{W} so that

$$(2.17) \quad W_i = \oint_{b_i} \omega = A_i(0^-) - A_i(\infty^-),$$

makes $\phi^{(0)}$ a well defined function. By virtue of the integrals in the exponent it has singularities of type 2, and again the solution of the Jacobi inversion problem by Riemann's theorem [11] gives the denominator exactly g zeros in the divisor \mathcal{D} . It follows that $\dim \Lambda_2(\Gamma, \mathcal{D}) = 1$.

This all means that if it is further stipulated that the functions u_1 and u_2 are to be meromorphic on Γ away from the four singular points, and are to have at most g simple poles at the fixed points P_1, \dots, P_g making up the nonspecial divisor \mathcal{D} , then the functions u_1 and u_2 are unique up to factors $\alpha_1(n, t)$ and $\alpha_2(n, t)$ constant on the surface Γ . By choosing these two factors, we fix the

functions u_1 and u_2 and obtain their expansions near the four singular points:

$$\begin{aligned}
 & u_1(n, t, P) \\
 (2.18) \quad & = \begin{cases} z^n \left(c_1^{(1)}(n, t) + c_1^{(2)}(n, t)z + \mathcal{O}(z^2) \right), & P \rightarrow 0^-, \\ \exp(i(1 - z^{-1})t) \left(d_1^{(1)}(n, t) + d_1^{(2)}(n, t)z + \mathcal{O}(z^2) \right), & P \rightarrow 0^+, \\ z^n \exp(i(z - 1)t) \left(1 + a_1^{(2)}(n, t)z^{-1} + \mathcal{O}(z^{-2}) \right), & P \rightarrow \infty^+, \\ b_1^{(1)}(n, t) + b_1^{(2)}(n, t)z^{-1} + \mathcal{O}(z^{-2}), & P \rightarrow \infty^-, \end{cases} \\
 (2.19) \quad &
 \end{aligned}$$

$$u_2(n, t, P)$$

$$= \begin{cases} z^{n+1} \left(c_2^{(1)}(n, t) + c_2^{(2)}(n, t)z + \mathcal{O}(z^2) \right), & P \rightarrow 0^-, \\ \exp(i(1 - z^{-1})t) \left(1 + d_2^{(2)}(n, t)z + \mathcal{O}(z^2) \right), & P \rightarrow 0^+, \\ z^n \exp(i(z - 1)t) \left(a_2^{(1)}(n, t) + a_2^{(2)}(n, t)z^{-1} + \mathcal{O}(z^{-2}) \right), & P \rightarrow \infty^+, \\ z \left(b_2^{(1)}(n, t) + b_2^{(2)}(n, t)z^{-1} + \mathcal{O}(z^{-2}) \right), & P \rightarrow \infty^-. \end{cases}$$

Thus, u_1 has been normalized to have a leading coefficient of 1 near ∞^+ and u_2 has been normalized to have a leading coefficient of 1 near 0^+ . The expansion coefficients depending on n and t are working variables determined completely by the choice of Γ and \mathcal{D} that will remain in use until the end of Section 3. They do not appear in any final formulas.

Exact formulas for the functions u_1 and u_2 are obtained by normalizing the functions $\psi^{(0)}$ and $\phi^{(0)}$ that appeared in the proofs of the lemmas. Any concrete representation of u_1 and u_2 requires corresponding concrete representations of the three differentials $\omega_{(3)}$, $\omega_{(2)}$ and ω . These representations are

$$(2.20) \quad \omega_{(3)} = \left[\frac{1}{2} \left(\frac{1}{z} - \frac{\eta}{zy} \right) + \frac{1}{2} \left(\frac{1}{z} + \frac{z^{g+1}}{zy} \right) - \frac{1}{2z} + \frac{P_{g-1}^{(3)}(z, \mathbf{z})}{y} \right] dz,$$

$$\begin{aligned}
 (2.21) \quad \omega_{(2)} = & i \left[\frac{1}{2} \left(1 + \frac{z^{g+1}}{y} \right) + \frac{1}{2} \left(\frac{1}{z^2} + \frac{\eta}{z^2 y} \right) - \frac{\eta}{4zy} \sum_{j=1}^{2g+2} \frac{1}{z_j} \right. \\
 & \left. - \frac{z^g}{4y} \sum_{j=1}^{2g+2} z_j + \frac{P_{g-1}^{(2)}(z, \mathbf{z})}{y} \right] dz,
 \end{aligned}$$

$$(2.22) \quad \omega = \left[\frac{1}{2} \left(\frac{1}{z} - \frac{\eta}{zy} \right) + \frac{1}{2} \left(\frac{1}{z} - \frac{z^{g+1}}{zy} \right) - \frac{1}{2z} + \frac{P_{g-1}(z, \mathbf{z})}{y} \right] dz.$$

At this point, to obtain concrete formulas, we must choose a homology basis $\{a_i, b_i, c_{\infty+}, c_{\infty-}, c_{0-}\}$ for $i = 1 \dots g$ on Γ . Then, $P_{g-1}^{(3)}$, $P_{g-1}^{(2)}$, and P_{g-1} are polynomials of degree $g - 1$ in z whose coefficients are symmetrical functions of the all the branch points contained in the vector \mathbf{z} and are determined by the g conditions on each differential that the integrals over a -cycles should vanish. The point here is that, while the functions u_1 and u_2 do not depend on any choice of homology, the formulas we will write down *do* depend on this choice as a result of normalizing differentials with the a -cycles. Thus, while making a specific choice of homology is nonphysical, like choosing gauge in a field theory, specific choice of gauge can be used to improve cosmetic features of equations and formulas. This homology gauge choice will be useful in simplifying modulation equations in Section 7.

We will also need the expansions of integrals of the three differentials $\omega_{(3)}$, $\omega_{(2)}$ and ω near the points 0^+ (where u_2 is normalized) and ∞^+ (where u_1 is normalized). The required asymptotic expansions are

$$(2.23) \quad \begin{aligned} \int_{P_0}^P \omega_{(3)} &= g_3 + \mathcal{O}(z), & P \rightarrow 0^+, \\ \int_{P_0}^P \omega_{(3)} &= \log z + G_3 + \mathcal{O}(z^{-1}), & P \rightarrow \infty^+, \end{aligned}$$

$$(2.24) \quad \begin{aligned} \int_{P_0}^P \omega_{(2)} &= -\frac{i}{z} + g_2 + \mathcal{O}(z), & P \rightarrow 0^+, \\ \int_{P_0}^P \omega_{(2)} &= iz + G_2 + \mathcal{O}(z^{-1}), & P \rightarrow \infty^+, \end{aligned}$$

$$(2.25) \quad \begin{aligned} \int_{P_0}^P \omega &= f + \mathcal{O}(z), & P \rightarrow 0^+, \\ \int_{P_0}^P \omega &= F + \mathcal{O}(z^{-1}), & P \rightarrow \infty^+. \end{aligned}$$

There are corresponding expressions of the order 1 terms in the above expansions in terms of (sometimes singular) integrals:

$$(2.26) \quad g_3 = \int_{P_0}^{0^+} \omega_{(3)},$$

$$(2.27) \quad G_3 = -\log(z(P_0)) + \int_{P_0}^{\infty^+} \left[\omega_{(3)} - \frac{dz}{z} \right],$$

$$(2.28) \quad g_2 = \frac{i}{z(P_0)} + \int_{P_0}^{0^+} \left[\omega_{(2)} - i \frac{dz}{z^2} \right],$$

$$(2.29) \quad G_2 = -iz(P_0) + \int_{P_0}^{\infty^+} [\omega_{(2)} - i dz],$$

$$(2.30) \quad f = \int_{P_0}^{0^+} \omega,$$

$$(2.31) \quad F = \int_{P_0}^{\infty^+} \omega.$$

Choosing g points in general position on Γ as the nonspecial divisor $\mathcal{D} = P_1 + \dots + P_g$ gives the formulas for u_1 and u_2 :

$$(2.32) \quad u_1(n, t, P) = \frac{\Theta(A(\infty^+) - \mathbf{Z})\Theta(A(P) - \mathbf{Z} + \mathbf{U}n + \mathbf{V}t)}{\Theta(A(\infty^+) - \mathbf{Z} + \mathbf{U}n + \mathbf{V}t)\Theta(A(P) - \mathbf{Z})} \\ \times \exp\left(- (G_2 + i)t - G_3n + \int_{P_0}^P n\omega_{(3)} + t\omega_{(2)}\right),$$

(2.33)

$$u_2(n, t, P) = \frac{\Theta(A(0^+) - \mathbf{Z})\Theta(A(P) - \mathbf{Z} + \mathbf{U}n + \mathbf{V}t + \mathbf{W})}{\Theta(A(0^+) - \mathbf{Z} + \mathbf{U}n + \mathbf{V}t + \mathbf{W})\Theta(A(P) - \mathbf{Z})} \\ \times \exp\left(- (g_2 - i)t - g_3n - f + \int_{P_0}^P n\omega_{(3)} + t\omega_{(2)} + \omega\right),$$

where $\mathbf{Z} = A(\mathcal{D}) + \mathcal{K}$, \mathcal{K} is the Riemann constant of Γ and Θ is the Riemann theta function of Γ .

3. Solving the linear problems with the Baker-Akhiezer function.

In this section, we will show that the vector function \mathbf{u} on Γ , specially constructed to have the correct asymptotic behavior near each of the four singular points, and having poles in the divisor \mathcal{D} , actually solves (1.14) and (1.16) globally on Γ , as long as the potentials $Q(n, t)$ and $R(n, t)$ are taken to be certain expansion coefficients of u_1 and u_2 . First, we deal with the spatial problem (1.14). Consider the two functions on Γ given by

$$(3.1) \quad \begin{aligned} \psi(n, P) &= z^{-1}u_2(n+1, P) - R(n)u_1(n, P) - z^{-1}u_2(n, P), \\ \phi(n, P) &= u_1(n+1, P) - zu_1(n, P) - Q(n)u_2(n, P), \end{aligned}$$

where we have suppressed the t dependence. These functions must be identically zero on Γ in order for \mathbf{u} to satisfy (1.4). Near singular points, they have the asymptotic expansions (suppressing the dependence of the expansion coefficients and potentials Q and R on the time t)

$$(3.2) \quad \psi = \begin{cases} z^n \left(\left[-R(n)c_1^{(1)}(n) - c_2^{(1)}(n) \right] + \mathcal{O}(z) \right), & P \rightarrow 0^-, \\ \exp(i(1-z^{-1})t) \\ \times \left(\left[d_2^{(2)}(n+1) - R(n)d_1^{(1)}(n) - d_2^{(2)}(n) \right] + \mathcal{O}(z) \right), & P \rightarrow 0^+, \\ z^n \exp(i(z-1)t) \left(\left[a_2^{(1)}(n+1) - R(n) \right] + \mathcal{O}(z^{-1}) \right), & P \rightarrow \infty^+, \\ b_2^{(1)}(n+1) - R(n)b_1^{(1)}(n) - b_2^{(1)}(n) + \mathcal{O}(z^{-1}), & P \rightarrow \infty^-, \end{cases}$$

$$(3.3) \quad \phi = \begin{cases} z^{n+1} \left(\left[c_1^{(1)}(n+1) - c_1^{(1)}(n) - Q(n)c_2^{(1)}(n) \right] + \mathcal{O}(z) \right), & P \rightarrow 0^-, \\ \exp(i(1-z^{-1})t) \left(\left[d_1^{(1)}(n+1) - Q(n) \right] + \mathcal{O}(z) \right), & P \rightarrow 0^+, \\ z^n \exp(i(z-1)t) \\ \times \left(\left[a_1^{(2)}(n+1) - Q(n)a_2^{(1)}(n) - a_1^{(2)}(n) \right] + \mathcal{O}(z^{-1}) \right), & P \rightarrow \infty^+, \\ z \left(\left[-b_1^{(1)}(n) - Q(n)b_2^{(1)}(n) \right] + \mathcal{O}(z^{-1}) \right), & P \rightarrow \infty^-. \end{cases}$$

From these expansions, it is evident that ψ has singularities of type 1, while ϕ has singularities of type 2. Furthermore, away from the four singular points, both of these functions have their poles confined to \mathcal{D} because u_1 and u_2 do. Thus, $\psi \in \Lambda_1(\Gamma, \mathcal{D})$ and $\phi \in \Lambda_2(\Gamma, \mathcal{D})$, so that $\psi \propto u_1$ and $\phi \propto u_2$.

At this point, if one chooses the potentials to satisfy

$$(3.4) \quad \begin{aligned} Q(n, t) &= d_1^{(1)}(n+1, t), \\ R(n, t) &= a_2^{(1)}(n+1, t), \end{aligned}$$

then the leading coefficient of ψ near ∞^+ vanishes identically in n and t and the leading coefficient of ϕ near 0^+ vanishes identically in n and t . Thus, the constants of proportionality are identically zero in n and t , so that $\psi \equiv 0$ and $\phi \equiv 0$ globally as functions on Γ . The function \mathbf{u} solves (1.14) where the potentials are given by (3.4). We have established the following lemma.

LEMMA 3.1. *The vector Baker-Akhiezer function \mathbf{u} solves the spatial linear problem (1.14) globally in n , t , and $P \in \Gamma$.*

With the spatial linear problem (1.14) solved, it is now possible to solve the temporal linear problem (1.16) as well. Consider the functions

$$(3.5) \quad \begin{aligned} \psi &= -i\partial_t u_1(n) - [z - 1 - Q(n)R(n-1)] u_1(n) \\ &\quad - [Q(n) - Q(n-1)z^{-1}] u_2(n), \end{aligned}$$

$$\begin{aligned} \phi &= -i\partial_t u_2(n) - [R(n-1)z - R(n)] u_1(n) \\ &\quad - [1 + R(n)Q(n-1) - z^{-1}] u_2(n). \end{aligned}$$

These functions should vanish identically on Γ in order for \mathbf{u} to satisfy (1.16). Near the four singular points, they have the expansions (once again suppressing the dependence of the expansion coefficients and the potentials Q and R on t)

$$(3.6) \quad \psi = \begin{cases} z^n \left(\left[-i\partial_t c_1^{(1)}(n) + (1 + Q(n)R(n-1)) c_1^{(1)}(n) \right. \right. \\ \quad \left. \left. + Q(n-1)c_2^{(1)}(n) \right] + \mathcal{O}(z) \right), & P \rightarrow 0^-, \\ \exp(i(1-z^{-1})t) \left(\left[-d_1^{(1)}(n) + Q(n,t) \right] z^{-1} + \mathcal{O}(1) \right), & P \rightarrow 0^+, \\ z^n \exp(i(z-1)t) \\ \quad \times \left(\left[Q(n) \left(R(n-1) - a_2^{(1)}(n) \right) \right] + \mathcal{O}(z^{-1}) \right), & P \rightarrow \infty^+, \\ \left[-b_1^{(1)}(n) - Q(n)b_2^{(1)}(n) \right] z + \mathcal{O}(1), & P \rightarrow \infty^-, \end{cases}$$

$$(3.7) \quad \phi = \begin{cases} z^{n+1} \left(\left[R(n)c_1^{(1)}(n) + c_2^{(1)}(n) \right] z^{-1} + \mathcal{O}(1) \right), & P \rightarrow 0^-, \\ \exp(i(1-z^{-1})t) \\ \quad \times \left(\left[R(n) \left(d_1^{(1)}(n) - Q(n-1) \right) \right] + \mathcal{O}(z) \right), & P \rightarrow 0^+, \\ z^n \exp(i(z-1)t) \left(\left[a_2^{(1)}(n) - R(n-1) \right] z + \mathcal{O}(1) \right), & P \rightarrow \infty^+, \\ \left[-i\partial_t b_2^{(1)}(n) - R(n-1)b_1^{(1)}(n) \right. \\ \quad \left. - (1 + R(n)Q(n-1)) b_2^{(1)}(n) \right] z + \mathcal{O}(1), & P \rightarrow \infty^-. \end{cases}$$

Upon applying (3.4) and using the facts (following from the vanishing of the expansions (3.3) and (3.2))

$$(3.8) \quad \begin{aligned} b_1^{(1)}(n, t) + Q(n, t)b_2^{(1)}(n, t) &= 0, \\ c_2^{(1)}(n, t) + R(n, t)c_1^{(1)}(n, t) &= 0, \end{aligned}$$

the expansions (3.6) and (3.7) may be reduced to

$$(3.9) \quad \psi = \begin{cases} z^n \mathcal{O}(1), & P \rightarrow 0^-, \\ \exp(i(1 - z^{-1})t) \mathcal{O}(1), & P \rightarrow 0^+, \\ z^n \exp(i(z - 1)t) \mathcal{O}(z^{-1}), & P \rightarrow \infty^+, \\ \mathcal{O}(1), & P \rightarrow \infty^-, \end{cases}$$

$$(3.10) \quad \phi = \begin{cases} z^{n+1} \mathcal{O}(1), & P \rightarrow 0^-, \\ \exp(i(1 - z^{-1})t) \mathcal{O}(z), & P \rightarrow 0^+, \\ z^n \exp(i(z - 1)t) \mathcal{O}(1), & P \rightarrow \infty^+, \\ \mathcal{O}(z), & P \rightarrow \infty^-. \end{cases}$$

Now, it is clear from these expressions that ψ has singularities of type 1 and ϕ has singularities of type 2. Away from the four singular points on Γ , these two functions have their poles confined to the divisor \mathcal{D} because u_1 and u_2 do. Thus, $\psi \in \Lambda_1(\Gamma, \mathcal{D})$ and $\phi \in \Lambda_2(\Gamma, \mathcal{D})$, so that $\psi \propto u_1$ and $\phi \propto u_2$. But, from the leading behavior of ψ near ∞^+ and ϕ near 0^+ , it appears that the constants of proportionality vanish as functions of n and t , and thus $\psi \equiv 0$ and $\phi \equiv 0$ globally on Γ . The function \mathbf{u} solves (1.16) where the potentials are given by (3.4). We have established the following lemma.

LEMMA 3.2. *The vector Baker-Akhiezer function \mathbf{u} solves the temporal linear problem (1.16) globally in n , t , and $P \in \Gamma$.*

We have constructed a simultaneous solution of (1.14) and (1.16) valid for all values of the spectral parameter $P \in \Gamma$, subject to the choice for the potentials $Q(n, t)$ and $R(n, t)$ given by (3.4). It follows that the two linear problems (1.14) and (1.16) are consistent for these potentials. Since the consistency conditions for these two linear problems,

$$(3.11) \quad \partial_t \mathbf{L}(n, t, z) + \mathbf{L}(n, t, z) \mathbf{B}(n, t, z) - \mathbf{B}(n + 1, t, z) \mathbf{L}(n, t, z) = 0,$$

are equivalent to the Ablowitz-Ladik equations (1.3), the potentials $Q(n, t)$ and $R(n, t)$ comprise a solution to this nonlinear problem.

THEOREM 3.3. *The pair of functions $Q(n, t) = d_1^{(1)}(n+1, t)$ and $R(n, t) = a_2^{(1)}(n+1, t)$ solves the Ablowitz-Ladik equations (1.3) for all integer n and complex t .*

Using the expressions for u_1 and u_2 in terms of the Riemann theta function of Γ , it is easy to extract the coefficients $d_1^{(1)}(n, t)$ and $a_2^{(1)}(n, t)$, resulting in the following expressions for the solution to (1.3):

$$(3.12) \quad Q(n-1, t) = \frac{\Theta(A(\infty^+) - \mathbf{Z})\Theta(A(0^+) - \mathbf{Z} + \mathbf{U}n + \mathbf{V}t)}{\Theta(A(\infty^+) - \mathbf{Z} + \mathbf{U}n + \mathbf{V}t)\Theta(A(0^+) - \mathbf{Z})} \\ \times \exp((g_3 - G_3)n + (g_2 - G_2 - 2i)t),$$

$$(3.13) \quad R(n-1, t) = \frac{\Theta(A(0^+) - \mathbf{Z})\Theta(A(\infty^+) - \mathbf{Z} + \mathbf{U}n + \mathbf{V}t + \mathbf{W})}{\Theta(A(0^+) - \mathbf{Z} + \mathbf{U}n + \mathbf{V}t + \mathbf{W})\Theta(A(\infty^+) - \mathbf{Z})} \\ \times \exp((F - f) - (g_3 - G_3)n - (g_2 - G_2 - 2i)t).$$

It is a consequence of the above construction that the function $u_1(0, 0, P)$ is the constant function 1 on the surface Γ . It follows that $d_1^{(1)}(0, 0) = 1$ and hence $Q(-1, 0) = 1$ for all constructed solutions to (1.3). It is easy to remove this constraint by making use of the following simple symmetry of (1.3): if the pair of functions $(Q(n, t), R(n, t))$ are a solution to (1.3), then $(\xi Q(n, t), \xi^{-1}R(n, t))$ will also be a solution for an arbitrary nonzero complex number ξ . In contrast with the construction of Previato [43] for the cubic nonlinear Schrödinger equation, this \mathbb{C}^* action cannot be absorbed into the unscaled formulas (3.12) and (3.13) through a change of the divisor \mathcal{D} to a linearly equivalent²¹ divisor \mathcal{D}' , without realizing the flow of (1.3) on a manifold that Previato refers to as a generalized Jacobian.

A family of solutions to (1.3) has been constructed from a set of data that is in analogy to the scattering data used to reconstruct the potential function in whole-line inverse scattering problems. In our case, the data making up a given solution pair $(Q(n, t), R(n, t))$ are

- the $2g+2$ distinct finite complex branch points z_i and the number η making up the Riemann surface with labeling from the set \mathcal{M}_g ,

²¹A divisor \mathcal{D} is linearly equivalent to another divisor \mathcal{D}' if $\mathcal{D} = \mathcal{D}' + P_1 + \dots + P_m - Q_1 - \dots - Q_m$, where the P_i are the zeros and the Q_i are the poles of some meromorphic function on Γ . It follows from Abel's theorem that $A(\mathcal{D})$ and $A(\mathcal{D}')$ differ by a lattice vector of the form

$$\sum_{i=1}^g m_i \mathbf{a}_i + n_i \mathbf{b}_i,$$

where the m_i and n_i are integers.

- the g points $P_j = (z_j, y_j)$ on Γ making up the nonspecial divisor \mathcal{D} , and
- the complex scaling parameter ξ .

We remark that the formulas simplify in the special case of $g = 0$. In this case, the theta quotients disappear, leaving only the complex exponentials. In Appendix A, an enlarged class of solutions will be developed which will include the above formulas as the generic case, but will allow the constraint that $z_i \neq z_j$ for $i \neq j$ to be effectively relaxed.

The solutions given by formulas (3.12) and (3.13) (or alternatively by the generalized formulas (A.13) and (A.14)) show explicitly that the potentials Q and R are functions of $g + 1$ phase variables θ_j of the form $k_j n - \omega_j t$. For $1 \leq j \leq g$, the wavenumbers k_j and frequencies ω_j are calculated from integrals over b -cycles of the differentials $\omega_{(2)}$ and $\omega_{(3)}$ according to (2.15); one finds $k_j = -iU_j$ and $\omega_j = iV_j$. For $j = 0$, however, we have the harmonic wavenumber

$$(3.14) \quad k_0 = -i(g_3 - G_3) = i \int_{0+}^{\infty+} \left[\omega_{(3)} - \frac{dz}{z+1} \right],$$

and the harmonic frequency

$$(3.15) \quad \omega_0 = 2 + i(g_2 - G_2) = 2 - i \int_{0+}^{\infty+} \left[\omega_{(2)} - i dz - i \frac{dz}{z^2} \right].$$

Note that these expressions are independent of the base point P_0 although the values of $g_2, G_2, g_3,$ and G_3 are base point dependent. If we can further establish that the k_j and ω_j are real, then the functions Q and R will in fact be periodic in each θ_j independently (for real t and integer n), and $Q(n, t)$ and $R(n, t)$ will represent $g + 1$ phase wavetrains. In fact, whenever we consider the reduction of the Ablowitz-Ladik equations (1.3) to the integrable discrete nonlinear Schrödinger equation (1.2), the required reality will be present, and the finite genus potentials will represent multiphase wavetrains. This reduction is considered in the next section, and in Section 7 we will show that in the focusing and defocusing cases it is always possible to choose the homology gauge so that the wavenumbers and frequencies are real.

4. Quasiperiodic, focusing, and defocusing solutions.

Numerical reconstructions of the potentials $Q(n, t)$ and $R(n, t)$ done using the formulas given in this paper ((3.12) and (3.13)) show that, for general choices of the data (the labeled surface from the set \mathcal{M}_g , the pole divisor \mathcal{D} , and the scaling parameter ξ), $Q(n, t)$ and $R(n, t)$ grow exponentially in both n and t . Even if one considers purely real t , which is of eventual interest in the case where the Ablowitz-Ladik equations go over into the integrable discrete nonlinear Schrödinger equation (1.2), this exponential growth is generally observed. It seems that in the general case, the finite genus solution cannot be correctly called a multiphase wavetrain. However, the quantity $Q(n, t)R(n, t)$ can be shown to

always be bounded, since the growth rates in $Q(n, t)$ are exactly cancelled by decay rates in $R(n, t)$. So, if it is known that $R(n, t) = \pm \overline{Q(n, t)}$ for real t , then in particular, the two potentials will be of the same magnitude, and neither will be able to undergo growth at the expense of the other. Following this lead, we consider in this section selecting the branch points, pole divisor, and scaling parameter so that $R(n, t) = \overline{Q(n, t)}$ for real t , the so-called defocusing case, or $R(n, t) = -\overline{Q(n, t)}$ for real t , the so-called focusing case. In Section 7 we will prove that by careful choice of homology gauge, real (focusing or defocusing) potentials are always quasiperiodic functions of both n and t .

In order to study this problem, it is useful to pass over to a more algebraic formulation analogous to that used by Previato [43] to study the solutions of the focusing and defocusing NLS equation. Previato mentions that similar constructions appeared in the work of Jacobi. Suppose that for a given complex z , one has *any* two solutions to the linear problems (1.14) and (1.16), \mathbf{u}^+ and \mathbf{u}^- . Consider the *squared eigenfunctions* given in terms of these two solutions by

$$\begin{aligned}
 \varphi(n, t, z) &= u_1^+(n, t, z)u_1^-(n, t, z), \\
 (4.1) \quad \chi(n, t, z) &= u_2^+(n, t, z)u_2^-(n, t, z), \\
 f(n, t, z) &= \frac{1}{2} \left(u_1^+(n, t, z)u_2^-(n, t, z) + u_1^-(n, t, z)u_2^+(n, t, z) \right).
 \end{aligned}$$

It follows that these quantities solve the system of equations (suppressing the dependence on the parameter z and the dependence of the matrix elements on n and t)

$$\begin{aligned}
 \varphi(n+1, t) &= L_{11}^2 \varphi(n, t) + L_{12}^2 \chi(n, t) + 2L_{11}L_{12}f(n, t), \\
 (4.2) \quad \chi(n+1, t) &= L_{21}^2 \varphi(n, t) + L_{22}^2 \chi(n, t) + 2L_{21}L_{22}f(n, t), \\
 f(n+1, t) &= L_{11}L_{21}\varphi(n, t) + L_{12}L_{22}\chi(n, t) \\
 &\quad + (L_{11}L_{22} + L_{21}L_{12})f(n, t),
 \end{aligned}$$

and the system

$$\begin{aligned}
 -i\partial_t \varphi(n, t) &= 2B_{11}\varphi(n, t) + 2B_{12}f(n, t), \\
 (4.3) \quad -i\partial_t \chi(n, t) &= 2B_{22}\chi(n, t) + 2B_{21}f(n, t), \\
 -i\partial_t f(n, t) &= B_{21}\varphi(n, t) + B_{12}\chi(n, t) + (B_{11} + B_{22})f(n, t).
 \end{aligned}$$

We introduce the quantity

$$(4.4) \quad J(n, t, z) = f^2(n, t, z) - \varphi(n, t, z)\chi(n, t, z).$$

From the evolution equations for φ , χ , and f , one derives the evolution equations for J :

$$(4.5) \quad \begin{aligned} J(n+1, t, z) &= [\det(\mathbf{L}(n, t, z))]^2 J(n, t, z) \\ &= z^2(1 - Q(n, t)R(n, t))^2 J(n, t, z), \end{aligned}$$

and

$$(4.6) \quad \begin{aligned} -i\partial_t J(n, t, z) &= 2\text{tr}(\mathbf{B}(n, t, z))J(n, t, z) \\ &= 2\left[z - z^{-1} + R(n, t)Q(n-1, t) \right. \\ &\quad \left. - Q(n, t)R(n-1, t)\right]J(n, t, z). \end{aligned}$$

Thus, it is possible to write

$$(4.7) \quad J(n, t, z) = z^{2n} \exp(2i(z - z^{-1})t) \beta^2(n, t)p(z),$$

where β depending on only n and t solves the equation

$$(4.8) \quad \beta(n+1, t) = (1 - Q(n, t)R(n, t))\beta(n, t),$$

and the equation

$$(4.9) \quad -i\partial_t \beta(n, t) = (R(n, t)Q(n-1, t) - Q(n, t)R(n-1, t))\beta(n, t).$$

The quantity $J(n, t, z)$ is useful because all of the nontrivial dependence on the spectral parameter z is contained in the factor $p(z)$, which does *not* depend upon n or t . The factor $p(z)$ is a generating function for constants of motion of the dynamical system (1.3).

We now examine the squared eigenfunctions φ , χ , and f constructed from the vector Baker-Akhiezer function \mathbf{u} through the relations (4.1). We set²²

$$(4.10) \quad \mathbf{u}^\pm(n, t, z) = \mathbf{u}(n, t, P^\pm(z)).$$

Using the description of \mathbf{u} in terms of its behavior near the four singular points of Γ and the locations of the poles in \mathcal{D} , we obtain the expressions

$$(4.11) \quad \varphi(n, t, z) = z^n \exp(i(z - z^{-1})t) \varphi_g(n, t) \frac{\prod_{k=1}^g (z - \mu_k(n, t))}{\prod_{j=1}^g (z - \mu_j(0, 0))},$$

²²We use the notation $P^\pm(w)$ for the pair of points on Γ that are preimages of the sheet projection function $z(P)$ at the point w . That is, $z(P^\pm(w)) = w$.

$$(4.12) \quad \chi(n, t, z) = z^{n+1} \exp(i(z - z^{-1})t) \chi_g(n, t) \frac{\prod_{k=1}^g (z - \zeta_k(n, t))}{\prod_{j=1}^g (z - \mu_j(0, 0))},$$

$$(4.13) \quad f(n, t, z) = z^n \exp(i(z - z^{-1})t) f_{g+1}(n, t) \frac{\prod_{k=0}^g (z - \gamma_k(n, t))}{\prod_{j=1}^g (z - \mu_j(0, 0))}.$$

In these expressions, the complex numbers $\mu_j(0, 0)$ are the z projections of the points P_j on Γ in the pole divisor \mathcal{D} . The numbers $\mu_j(n, t)$ are the z projections of the moving zeros of the component u_1 ; the numbers $\zeta_j(n, t)$ are the z projections of the moving zeros of the component u_2 . The numbers $\gamma_k(n, t)$ can be expressed in terms of u_1 and u_2 , but not as easily as $\mu_k(n, t)$ or $\zeta_k(n, t)$. Multiplication by a function of z only does not change the fact that these expressions solve (4.2) and (4.3), so we renormalize them by multiplying by the common denominator to be polynomials in z multiplied by a common factor $z^n \exp(i(z - z^{-1})t)$:

$$(4.14) \quad \begin{aligned} \varphi(n, t, z) &= z^n \exp(i(z - z^{-1})t) \left[\sum_{k=0}^g z^k \varphi_k(n, t) \right] \\ &= z^n \exp(i(z - z^{-1})t) \left[\varphi_g(n, t) \prod_{j=1}^g (z - \mu_j(n, t)) \right], \\ \chi(n, t, z) &= z^n \exp(i(z - z^{-1})t) \left[z \sum_{k=0}^g z^k \chi_k(n, t) \right] \\ &= z^n \exp(i(z - z^{-1})t) \left[z \chi_g(n, t) \prod_{j=1}^g (z - \zeta_j(n, t)) \right], \\ f(n, t, z) &= z^n \exp(i(z - z^{-1})t) \left[\sum_{k=0}^{g+1} z^k f_k(n, t) \right]. \end{aligned}$$

So, when n and t are equal to zero, f , φ , and χ are polynomials in z . It then follows from the systems of equations (4.2) and (4.3) that the coefficients $f_{g+1}(n, t)$ and $f_0(n, t)$ *both* solve the same linear equations as the quantity $\beta(n, t)$; thus the two coefficients are proportional via a constant (independent of n , t , and z). We further normalize the squared eigenfunctions φ , χ , and f so that

$$(4.15) \quad f_{g+1}(0, 0) = 1.$$

It then follows that

$$(4.16) \quad f_0(0, 0) = -\eta.$$

This normalization reveals the nature of the quantity $J(n, t, z)$. In particular,

$$(4.17) \quad J(0, 0, z) = \prod_{k=1}^{2g+2} (z - z_k).$$

The evolution of $J(n, t, z)$ in n and t is through the function $\beta(n, t)$; the zeros z_k in z remain fixed as n and t vary.

Since these polynomials were constructed from a vector Baker-Akhiezer function satisfying $u_1(0, 0, P) = 1$, there is a constraint on the coefficients, namely that $\varphi_0(0, 0) = 2f_0(0, 0)$. By dropping this constraint, we will be able to describe potentials that arise from the formulas (3.12) and (3.13) through the scaling parameter ξ . We will now show that the set of (almost) all polynomials of this general type, for $n = 0$ and $t = 0$, provides a coordinatization of the phase space of (1.3) in the class of finite genus potentials. These coordinates will be particularly useful because the reality conditions will find a very easy expression therein. To begin with, we need a definition of finite genus potentials. Let \mathcal{M}'_g denote the set of labeled surfaces \mathcal{M}_g with the diagonals $z_i = z_j$ included; this enlarged set is described in detail in Appendix A. An element Γ of \mathcal{M}'_g consists of a Riemann surface of genus $\tilde{g} \leq g$ having distinct branch points, and an integral divisor \mathcal{E} of degree $g - \tilde{g}$ on \mathbb{C}^* . If $\tilde{g} = g$, then Γ is in \mathcal{M}_g . Section 2 contains the description of the fibers of Baker-Akhiezer functions associated with elements Γ of \mathcal{M}_g ; the prescription for associating a fiber of generalized Baker-Akhiezer functions to arbitrary elements of \mathcal{M}'_g is given in Appendix A.

DEFINITION 4.1. *The class of genus g potentials is the set of all pairs of functions of n , $(Q(n), R(n))$, such that for some element Γ of \mathcal{M}'_g , some set of g points $\mathcal{D} = P_1 + \dots + P_g$ in general position on the Riemann surface associated with Γ , and a given complex number ξ , $Q(n) = \xi Q(n, 0; \Gamma, \mathcal{D})$ and $R(n) = \xi^{-1} R(n, 0; \Gamma, \mathcal{D})$ where $Q(n, t; \cdot)$ and $R(n, t; \cdot)$ are constructed from the generalized Baker-Akhiezer function \mathbf{u} .*

Most genus g potentials correspond to taking $\Gamma \in \mathcal{M}_g$ so that the branch points are distinct. Such a genus g potential is determined by the branch points z_1, \dots, z_{2g+2} together with the sign of η , the g points of the divisor \mathcal{D} , and the scaling parameter ξ . Up to the sheet indices (± 1) of the points of the divisor \mathcal{D} , and also the sign of η , these data make up $3g + 3$ independent complex parameters. Similarly, if $J(0, 0, z)$ has distinct roots, then with $f_{g+1}(0, 0) = 1$, there remain in polynomials $f(0, 0, z)$, $\varphi(0, 0, z)$, and $\chi(0, 0, z)$ $3g + 3$ independent complex coefficients. In fact, more is true.

LEMMA 4.2. *The class of genus g potentials $(Q(n), R(n))$ is in one-to-one correspondence with the class of (almost) all triples of polynomials $(f(z), \varphi(z), \chi(z))$ of the form (4.14) with $n = 0$, $t = 0$, and $f_{g+1} = 1$. The meaning of “almost” here is that a triple of polynomials must lead to a nonspecial divisor \mathcal{D} and*

nonzero branch points. Such polynomials with otherwise arbitrary coefficients provide global coordinates for the phase space of genus g potentials.

Proof: We prove here only the association between genus g potentials corresponding to $\Gamma \in \mathcal{M}_g$ and polynomials for which $J(z)$ has distinct roots, leaving the discussion of the diagonal cases for Appendix A. Let $(Q(n), R(n))$ be a genus g potential constructed from the data $(z_1, \dots, z_{2g+2}, \eta, \mathcal{D}, \xi)$ with the z_i distinct. Construct the vector Baker-Akhiezer function \mathbf{u} from the labeled curve Γ and the pole divisor \mathcal{D} exactly as described in Section 2, and let \mathbf{u}^\pm be the corresponding pair of vector functions of z obtained as described above. Set n and t equal to zero, and construct the squared eigenfunctions (scaled as in (4.14), with $f_{g+1} = 1$), calling them $f^{(0)}(z)$, $\varphi^{(0)}(z)$, and $\chi^{(0)}(z)$. We have yet to make use of the scaling parameter ξ ; we take this into account by defining the polynomials $f(z) = f^{(0)}(z)$, $\varphi(z) = \xi\varphi^{(0)}(z)$, and $\chi(z) = \xi^{-1}\chi^{(0)}(z)$. On the other hand, let arbitrary polynomials $f(z)$, $\varphi(z)$, and $\chi(z)$ of the form (4.14) with $n = 0$, $t = 0$, and $f_{g+1} = 1$ be given, such that $J(z)$ has distinct nonzero roots. The branch points z_i are obtained as the zeros of the polynomial $f^2(z) - \varphi(z)\chi(z)$; the labeling parameter of the surface Γ is given by $\eta = -f_0$. The sheet projections of the points of the pole divisor \mathcal{D} are given by the zeros of the polynomial $\varphi(z)$; the sheet indices of the points of \mathcal{D} are determined from the relations²³ $y(P_j) = -f(z(P_j))$. The only remaining piece of information to determine is the scaling parameter ξ , which is given by $\varphi_0/2f_0$. Given these associations, we can now very precisely describe the class of polynomials that give rise to genus g potentials:

- If μ_i and μ_j are roots of the polynomial φ such that $\mu_i = \mu_j$, then the corresponding points of the divisor \mathcal{D} must lie on the same sheet of Γ in order for \mathcal{D} to be nonspecial; that is it must be true that $f(\mu_i) = f(\mu_j)$.
- In order for the branch points to be nonzero, it must be true that $f_0 \neq 0$.

The polynomials we must exclude therefore make up a very low dimensional subset of \mathbb{C}^{3g+3} , justifying our use of the phrase “almost all triples of polynomials”.

So, the phase space of genus g potentials is isomorphic to the set of almost all complex polynomials of the form (4.14) with $f_{g+1} = 1$. We will use this representation of the phase space to obtain reality conditions on the problem data sufficient to have $R(n, t) = \pm \overline{Q(n, t)}$ for real t . We introduce here two involutions on the phase space of all potentials $(Q(n), R(n))$ (these will preserve the classes of genus g potentials):

²³The correct signs of η and $y(P_j)$ come from the representation of the Baker-Akhiezer function at $n = t = 0$ in terms of the squared eigenfunction polynomials. We have $u_1(0, 0, P) = 1$, and

$$u_2(0, 0, P) = \frac{f(z(P)) - y(P)}{\varphi(z(P))}.$$

- the defocusing involution, given by

$$(4.18) \quad D : \begin{pmatrix} Q(n) \\ R(n) \end{pmatrix} \mapsto \begin{pmatrix} \overline{R(n)} \\ \overline{Q(n)} \end{pmatrix},$$

- and the focusing involution, given by

$$(4.19) \quad F : \begin{pmatrix} Q(n) \\ R(n) \end{pmatrix} \mapsto \begin{pmatrix} -\overline{R(n)} \\ -\overline{Q(n)} \end{pmatrix}.$$

The set fixed points of D , the defocusing potentials, is invariant under the flow of (1.3) for real t . Likewise the set of fixed points of F , the focusing potentials, is invariant under the flow of (1.3) for real t . According to the lemma, there is an induced action of D (respectively F) on the set of normalized polynomials of the form (4.14) called D_* (respectively F_*). The induced action is given by

$$(4.20) \quad D_* : \begin{pmatrix} f(z) \\ \varphi(z) \\ \chi(z) \end{pmatrix} \mapsto \frac{z^{2n+g+1}}{f_0} \begin{pmatrix} \overline{f(1/\bar{z})} \\ \overline{\chi(1/\bar{z})} \\ \overline{\varphi(1/\bar{z})} \end{pmatrix},$$

and

$$(4.21) \quad F_* : \begin{pmatrix} f(z) \\ \varphi(z) \\ \chi(z) \end{pmatrix} \mapsto \frac{z^{2n+g+1}}{f_0} \begin{pmatrix} \overline{f(1/\bar{z})} \\ -\overline{\chi(1/\bar{z})} \\ -\overline{\varphi(1/\bar{z})} \end{pmatrix}.$$

Defocusing potentials are constructed from triples of polynomials that are fixed by D_* . Likewise, focusing potentials are constructed from triples of polynomials that are fixed by F_* . The fixed points are easily characterized; this, along with the construction given in the proof of Lemma 4.2, provides a prescription for generating the generalized Baker-Akhiezer function data $(\Gamma, \mathcal{D}, \xi)$ necessary to build both focusing and defocusing solutions to (1.3).

THEOREM 4.3. *For almost every set of data consisting of*

- some $\theta \in [0, 2\pi)$,
- g complex numbers $(f_1 \dots f_g)$ satisfying

$$\overline{f_k} \exp(i\theta) = f_{g+1-k},$$

- $g + 1$ arbitrary complex numbers $(\varphi_0 \dots \varphi_g)$

there is a focusing potential constructed via the polynomials

$$\begin{aligned} f(z) &= z^n [z^{g+1} + f_g z^g + \dots + f_1 z + \exp(i\theta)], \\ \varphi(z) &= z^n \left[\sum_{k=0}^g \varphi_k z^k \right], \\ \chi(z) &= -z^{2n+g+1} \overline{\varphi(1/\bar{z})} \exp(i\theta), \end{aligned}$$

and a defocusing potential constructed via the polynomials

$$\begin{aligned} f(z) &= z^n [z^{g+1} + f_g z^g + \dots + f_1 z + \exp(i\theta)] , \\ \varphi(z) &= z^n \left[\sum_{k=0}^g \varphi_k z^{2k} \right] , \\ \chi(z) &= z^{2n+g+1} \overline{\varphi(1/\bar{z})} \exp(i\theta) , \end{aligned}$$

using the construction of Lemma 4.2. The meaning of the word “almost” is that the resulting polynomials must satisfy the mild conditions mentioned at the end of the proof of Lemma 4.2 in order to avoid special divisors and branching at $z = 0$.

It should be remarked that the conditions of reality sufficient to guarantee focusing and defocusing solutions can be imposed on three levels:

- the level of the potentials, where reality means $R = \pm \bar{Q}$;
- the level of the squared eigenfunction polynomials, where the reality conditions are given by Theorem 4.3;
- the level of the Baker-Akhiezer function data (the element of \mathcal{M}'_g along with divisor \mathcal{D} and scaling parameter).

In the past, reality conditions have been transcendently expressed at the level of the Baker-Akhiezer function data via the Abel map (see for example [31]); here we have chosen to characterize the spectral data corresponding to real potentials at the level of the squared eigenfunction polynomials because the reality conditions are simple in form and because the real Baker-Akhiezer function data $(\Gamma, \mathcal{D}, \xi)$ can be recovered from the squared eigenfunctions using only algebraic operations, as described in the proof of Lemma 4.2.

In particular, it is possible to conclude from Theorem 4.3 that Riemann surfaces for both focusing and defocusing solutions must have branch points that are, as a set, symmetrical in reflection through the unit circle. Finding the set of all divisors giving focusing (defocusing) potentials is an algebraic problem to be solved after fixing a Riemann surface whose branch points enjoy reflection symmetry through the unit circle. The monic polynomial $p(z)$ whose roots are the branch points must be decomposed into polynomials f , φ , and χ having focusing (defocusing) symmetry as described in Theorem 4.3 such that $f^2 - \varphi\chi = p$. Each possible decomposition gives a focusing (defocusing) divisor \mathcal{D} that is recovered from φ and f . That there exist such decompositions is suggested by a calculation in Section 7 that demonstrates the existence of a real part of the Jacobian of any Riemann surface whose branch points have the unit circle symmetry. In practical calculations, it is most efficient to proceed according to the following steps (as outlined by Tracy, Chen, and Lee [46] in the context of the nonlinear Schrödinger equation):

1. Choose branch points z_k having reflection symmetry through the unit circle and choose the sign of η . Build the monic polynomial $P(z)$ whose roots are the branch points.
2. Choose a monic polynomial f with $f(0) = -\eta$, whose coefficients have the symmetry described in Theorem 4.3.
3. Factor the polynomial $P(z) - f^2(z)$. In order for there to be a solution, none of the roots may lie on the unit circle. From each pair of roots symmetric with respect to unit circle reflection, select one root. These will be the pole projections $z(P_j)$.
4. Calculate $|\xi|^2$ as

$$|\xi|^2 = \pm \frac{(-1)^g}{4\eta} \prod_{j=1}^g \overline{z(P_j)} \lim_{z \downarrow 0} \left[\frac{1}{z} (P(z) - f^2(z)) \right],$$

where $+$ ($-$) indicates defocusing (focusing). In order for there to be a solution, this value must be real and positive.

This procedure involves only root finding operations, and produces a solution for each appropriate choice of the polynomial f as long as none of the roots of $P - f^2$ lie on the unit circle and as long as the value of $|\xi|^2$ is real and positive. The possibility of finding a solution for a given set of branch points and value of η indicates an allowable branch point configuration. Not all symmetric branch point configurations are allowable.

As a demonstration of the allowable branch point configurations, it is useful to consider the simplest case, $g = 0$. A simple application of the quadratic formula shows that

- in the focusing case the two branch points do *not* lie on the unit circle, and,
- in the defocusing case, the two branch points lie on the unit circle if $|\varphi_0|^2 < 4$ and are split off the unit circle if $|\varphi_0|^2 > 4$. In terms of the potentials, since $\varphi_0 = -2\eta Q(-1, 0)$, the separatrix between the two cases occurs whenever there exists an n for which $|Q(n, t)|^2 = 1$.

Similar results hold for $g > 0$. More specifically, we will show in Appendix B that in the focusing case, the branch points must always lie off of the unit circle unless they are double. The defocusing case is less straightforward, due to the presence of a separatrix in the phase space. The set of phase points satisfying $|Q(n, t)|^2 = 1$ for some n make up a dynamical barrier for orbits of (1.3) in the defocusing case. That is, if for some n and t one has $|Q(n, t)|^2 < 1$, then this condition persists for all real t , and similarly if one has $|Q(n, t)|^2 > 1$ for some n and t , then this condition is also permanent. The defocusing phase space is thus divided into dynamically disjoint sectors Σ_κ indexed by sets of lattice points $\kappa = \{n_1, n_2, \dots\}$ indicating that $|Q(n, t)|^2 < 1$ exclusively for those n in κ . In Appendix B, we will show that the branch points for all potentials in

the component $\Sigma_{\mathbb{Z}}$ (where $|Q(n, t)|^2 < 1$ for all n) are *all* constrained to lie on the unit circle. The constraints on the branch points corresponding to other defocusing components Σ_{κ} may not be so strict. These statements are proved in Appendix B using periodic boundary conditions on n , but by density arguments, the branch point symmetries should carry over to general finite genus potentials.

With this information, it is possible to make a connection between the spectral properties of the linear problem (1.14) and those of its continuum limit, the AKNS scattering problem for the focusing and defocusing nonlinear Schrödinger equations [3]. The fixed point set of the antiholomorphic spectral involution in the continuous case is the real axis rather than the unit circle. In the defocusing case, the AKNS spectral problem is self-adjoint, and thus the branch points (periodic and antiperiodic Floquet eigenvalues) are constrained to lie on the real axis, within the fixed point set. On the other hand, in the focusing case, spectral convexity considerations [39] show that the branch points cannot lie on the real axis unless they are double. Now, in the continuum limit, the cutoff condition for the defocusing component $\Sigma_{\mathbb{Z}}$ scales as

$$(4.22) \quad |Q(n, t)|^2 < \frac{1}{h^2},$$

where h is the vanishing lattice spacing. Thus, the only component of the defocusing phase space that contributes in the continuum limit is $\Sigma_{\mathbb{Z}}$. In this component, the branch points must lie on the unit circle. Furthermore, the branch points for all focusing potentials must lie off of the unit circle. These conjectures agree with the formal connection between the discrete spectral parameter z and the AKNS spectral parameter ζ

$$(4.23) \quad z = \exp(-ih\zeta).$$

This transformation takes the real axis to the unit circle. The remaining components of the defocusing phase space Σ_{κ} where $\kappa \neq \mathbb{Z}$ contain behavior that has absolutely no analog in the continuous defocusing nonlinear Schrödinger equation. It is likely that these components of the defocusing phase space are noncompact, as it has been observed that having a lattice point at which the modulus of the solution exceeds the value 1 can lead to blowup in finite time [26, 24].

Another interesting structural feature we deduce in both of the real cases is that the focusing and defocusing submanifolds of the set \mathcal{M}_g each consist of two disconnected components. This is a consequence of the arrangement of the branch points about the unit circle; they must either come in pairs reflected through the circle, or be on the circle itself in which case they must also come in pairs, since the total number of branch points is even. It is easy to see that any deformation in this class of Riemann surfaces that returns the branch points to their initial configuration must either involve an even number of branch points traveling around the origin, so that the points 0^{\pm} *cannot* be exchanged one for the other, or one branch point passing through another on the unit circle. Thus,

the elements of \mathcal{M}_g given by $(\{z_i\}, \eta)$ and $(\{z_i\}, -\eta)$ cannot be connected by a smooth path of real elements of \mathcal{M}_g . In the defocusing case, when a pair of multiplicity 1 branch points may lie on the unit circle, it is possible to make the connection, but only by passing through an element of the more general set \mathcal{M}'_g when the two branch points on the circle collide.

5. Spatially periodic solutions of the Ablowitz-Ladik equations.

In this section we identify the Baker-Akhiezer function data $(\Gamma, \mathcal{D}, \xi)$ that correspond to finite genus complex solutions of (1.3) that are periodic in n , and correspondingly describe the solution of the spatially periodic initial value problem in terms of explicit formulas. In doing so, we will make contact with the work of Bogolyubov and Prikarpat-skii [6], who considered the solution of the initial value problem in the focusing and defocusing cases described by (1.2).

The analysis in the case when the potentials $Q(n, t)$ and $R(n, t)$ are periodic in the integer index n with period N begins with the *monodromy matrix* of the spatial linear problem (1.14). This matrix is given by

$$(5.1) \quad \mathbf{S}(n, t, z) = \mathbf{L}(n + N - 1, t, z)\mathbf{L}(n + N - 2, t, z) \cdots \mathbf{L}(n + 1, t, z)\mathbf{L}(n, t, z).$$

The monodromy matrix is the linear mapping that takes a solution \mathbf{u} of the spatial linear problem (1.14) at position n to the same solution at position $n + N$. It is easy to see that the eigenvalues of $\mathbf{S}(n, t, z)$ are independent of n and t . The n independence follows from the similarity of $\mathbf{S}(n, t, z)$ and $\mathbf{S}(n + 1, t, z)$:

$$(5.2) \quad \begin{aligned} \mathbf{S}(n + 1, t, z) &= \mathbf{L}(n + N, t, z)\mathbf{S}(n, t, z)\mathbf{L}^{-1}(n, t, z) \\ &= \mathbf{L}(n, t, z)\mathbf{S}(n, t, z)\mathbf{L}^{-1}(n, t, z). \end{aligned}$$

The t independence comes from the fact that the monodromy matrix satisfies a Lax equation in time; since

$$(5.3) \quad \begin{aligned} -i\partial_t \mathbf{u}(n + N, t, z) &= -i\partial_t \left(\mathbf{S}(n, t, z)\mathbf{u}(n, t, z) \right) \\ &= \left[-i\partial_t \mathbf{S}(n, t, z) \right] \mathbf{u}(n, t, z) \\ &\quad + \mathbf{S}(n, t, z) \left[-i\partial_t \mathbf{u}(n, t, z) \right] \\ &= \left[-i\partial_t \mathbf{S}(n, t, z) + \mathbf{S}(n, t, z)\mathbf{B}(n, t, z) \right] \mathbf{u}(n, t, z), \end{aligned}$$

and also

$$(5.4) \quad \begin{aligned} -i\partial_t \mathbf{u}(n + N, t, z) &= \mathbf{B}(n + N, t, z)\mathbf{u}(n + N, t, z) \\ &= \mathbf{B}(n, t, z)\mathbf{S}(n, t, z)\mathbf{u}(n, t, z), \end{aligned}$$

the monodromy matrix solves

$$(5.5) \quad -i\partial_t \mathbf{S}(n, t, z) = [\mathbf{B}(n, t, z), \mathbf{S}(n, t, z)].$$

Thus, if $\mathbf{A}(n, t, z)$ is a matrix such that

$$(5.6) \quad -i\partial_t \mathbf{A}(n, t, z) = \mathbf{B}(n, t, z)\mathbf{A}(n, t, z),$$

then it will be true that

$$(5.7) \quad \mathbf{S}(n, t, z) = \mathbf{A}(n, t, z)\mathbf{S}(n, 0, z)\mathbf{A}^{-1}(n, t, z),$$

hence, the eigenvalues of $\mathbf{S}(n, t, z)$ will be independent of t as well. The trace and determinant of $\mathbf{S}(n, t, z)$ therefore depend on z only; accordingly, we write

$$(5.8) \quad T_s(z) = \text{tr}(\mathbf{S}(n, t, z)),$$

$$(5.9) \quad D_s(z) = \det(\mathbf{S}(n, t, z)).$$

The invariant eigenvalues of $\mathbf{S}(n, t, z)$ are the *Floquet multipliers* of the linear problem (1.14). They have the form

$$(5.10) \quad \rho^\pm(z) = \frac{1}{2} \left[T_s(z) \pm \sqrt{T_s^2(z) - 4D_s(z)} \right],$$

where we use $T_s^2(z)$ to denote $(T_s(z))^2$ here and throughout the paper. The two multipliers may be analytically continued into each other in z ; it is useful to consider them as a single meromorphic function $\rho(P)$ on a hyperelliptic Riemann surface which we now proceed to describe. The polynomial under the radical is written in factored form:

$$(5.11) \quad T_s^2(z) - 4D_s(z) = s^2(z)p(z),$$

where $s(z)$ and $p(z)$ are monic polynomials, possibly with some roots in common, but such that $p(z)$ has distinct roots. Then, the surface used in the analysis of the periodic problem is the surface of the algebraic relation

$$(5.12) \quad y^2 = p(z),$$

In the case where $s^2(z) = 1$, the expression on the right hand side is a monic polynomial of degree $2N$ with $p(0) = 1$; thus, the genus is $g = N - 1$, and the initial value problem will be solved in terms of the formulas for $Q(n, t)$ (3.12) and $R(n, t)$ (3.13). However, in general, the genus will be less, $s^2(0)p(0) = 1$, and the generalized formulas for $Q(n, t)$ (A.13) and $R(n, t)$ (A.14) from Appendix A will be needed to solve the initial value problem. In any case, the Floquet multiplier is then the function

$$(5.13) \quad \rho(P) = \frac{1}{2} \left(T_s(z(P)) + s(z(P))y(P) \right).$$

We will now prove the following lemma.

LEMMA 5.1. *Let $Q(n, t)$ and $R(n, t)$ be generalized finite genus potentials, given by (A.13) and (A.14), that are periodic in n with period N . When $1 \leq g \leq N - 1$, the vector $N\mathbf{U}$ is in the period lattice of the hyperelliptic Riemann surface characterized by the branch points z_i satisfying $p(z_i) = 0$ and $\eta = 1/s(0)$.*

Proof: Using Riemann’s bilinear identity²⁴ and the definition of the differential $\omega_{(3)}$, one obtains

$$(5.14) \quad N U_j = N \oint_{b_j} \omega_{(3)} = N A_j(0^-) - N A_j(\infty^+) = A_j(N 0^- - N \infty^+).$$

But, by using Abel’s theorem [11], we see that $N\mathbf{U}$ is in the period lattice if and only if the divisor $N 0^- - N \infty^+$ is the divisor of a meromorphic function on the Riemann surface of the relation (5.12). In such a case, when in addition the surface is labeled over $z = 0$ by $\eta = 1/s(0)$, we have such a function available; it is the Floquet multiplier function $\rho(P)$.

Thus in the formulas for periodic generalized potentials with $g \geq 1$, the vector $N\mathbf{U}$ has an expression of the form

$$(5.15) \quad N \mathbf{U} = \sum_{j=1}^g m_j \mathbf{b}_j + 2\pi i n_j \mathbf{e}_j.$$

Once it is known that $N\mathbf{U}$ may be written in this way, it is possible to use the transformation law for theta functions in the formulas (A.13) and (A.14) to insist on periodicity and hence deduce that the harmonic wavenumber $g_3 - G_3$ obeys the relation

$$(5.16) \quad \exp \left(\sum_{j=1}^g m_j A_j(\infty^+ - 0^+) + N(g_3 - G_3) \right) = 1.$$

Using these relations and examining the generalized vector Baker-Akhiezer function \mathbf{u} , we find that \mathbf{u} is a Floquet eigenfunction (an eigenfunction of \mathbf{S}). The link between the Floquet multiplier and the differential $\omega_{(3)}$ is given by

$$(5.17) \quad - \sum_{j=1}^g m_j \nu_j + N \omega_{(3)} = d \log \rho,$$

²⁴Riemann’s bilinear identity provides a link between differentials of the third kind and differentials of the first kind. Let τ_{PQ} be the unique differential of the third kind with residue 1 (-1) at the point P (Q), holomorphic otherwise, and having vanishing integrals over all a -cycle representatives of a chosen generalized homology basis $\{a_1, \dots, a_g, b_1, \dots, b_g, c_{\infty^+}, c_{\infty^-}, c_{0^-}\}$. Then

$$\oint_{b_j} \tau_{PQ} = A_j(P) - A_j(Q).$$

See [11] for details.

where ν_j are the elements in the basis of normalized first kind Abelian differentials on Γ . Actually, these expressions can be simplified by choosing different a -cycles (with respect to which $\omega_{(3)}$ is normalized) in the canonical homology basis. Since the differential $d \log \rho$ has integer residues at 0^- and ∞^+ , it is possible to add to any a -cycle an integer sum of the cycles c_{∞^+} and c_{0^-} so that for the new a -cycles:

$$(5.18) \quad N\omega_{(3)} = d \log \rho,$$

and therefore

$$(5.19) \quad N\mathbf{U} = \sum_{j=1}^g 2\pi i n_j \mathbf{e}_j,$$

$$(5.20) \quad N(g_3 - G_3) = 2\pi i n_0,$$

for some integer n_0 .

We can now characterize precisely the elements Γ of the set \mathcal{M}'_{N-1} corresponding to periodic potentials of period N arising from the generalized formulas (A.13) and (A.14). To state the theorem, we need to use the fact that an element Γ in the set \mathcal{M}'_{N-1} consists of a Riemann surface of genus $g \leq N-1$ with choice of sign of η and having distinct branch points z_1, \dots, z_{2g+2} , and an integral divisor of degree $N-1-g$ on \mathbb{C}^* , $\mathcal{E} = m_1 w_1 + \dots + m_j w_j$.

THEOREM 5.2. *For the potentials $Q(n, t)$ and $R(n, t)$ constructed from the data $(\Gamma, \mathcal{D}, \xi)$ to be periodic functions of n of period N , it is necessary and sufficient for $\Gamma \in \mathcal{M}'_{N-1}$ to be such that the branch points z_k and the divisor \mathcal{E} satisfy*

$$(5.21) \quad \left[\prod_{l=1}^j (z - w_l)^{m_l} \right]^2 \left[\prod_{k=1}^{2g+2} (z - z_k) \right] = s^2(z)p(z) = [\tau(z)]^2 + \delta z^N,$$

for some constant $\delta \neq 0$ and some monic polynomial $\tau(z)$ of degree N satisfying $\tau(0) = 1$, and for the labeling of the points 0^\pm on the Riemann surface to correspond to

$$(5.22) \quad \eta = s(0)^{-1} = \left[\prod_{l=1}^j (-w_l)^{m_l} \right]^{-1}.$$

Proof: We will give the proof for the generic case of $\Gamma \in \mathcal{M}_{N-1}$ and thus $j = 0$ and $g = N-1$. The necessity follows from the fact that the trace of the monodromy matrix has the form

$$(5.23) \quad T_{\mathbf{s}}(z) = z^N + \tau_{N-1} z^{N-1} + \dots + \tau_1 z + 1,$$

and the determinant of the monodromy matrix has the form

$$(5.24) \quad D_{\mathbf{s}}(z) = z^N \prod_{k=n}^{n+N-1} (1 - Q(k, t)R(k, t)).$$

Thus, $\tau(z) = T_{\mathbf{s}}(z)$, and $\delta = -z^{-N}D_{\mathbf{s}}(z)/4$. The sufficiency follows from two facts. First, as a consequence of the above lemma, of the Baker-Akhiezer function data $(\Gamma, \mathcal{D}, \xi)$, only the branch points z_i and the labeling parameter η determine whether the potentials as given by the formulas (3.12) and (3.13) represent periodic functions of n . Thus, it is sufficient to appropriately specify the Riemann surface Γ with choice of sign of η from the set \mathcal{M}_{N-1} . Second, as we consider varying the potentials within the class of periodic potentials of period N , the coefficients in the trace and determinant of the monodromy matrix $\mathbf{S}(n, t, z)$ take on all possible complex values (subject to the restriction that the trace is monic with constant term equal to 1). Thus, any branch points satisfying the conditions given in the theorem must generate only solutions that are periodic in n .

The extension of the proof to cases in which Γ is in \mathcal{M}'_{N-1} but is not generic (so that it is *not* in \mathcal{M}_{N-1}) involves associating Γ to τ and δ where $[\tau(z)]^2 + \delta z^N$ has multiple roots.

It is not possible to associate an element of \mathcal{M}'_{N-1} to the case of $\delta = 0$ because, as explained in Appendix A, elements of \mathcal{M}'_{N-1} describe hyperelliptic Riemann surfaces that are branched in at least two places, and when $\delta = 0$, all roots of $[\tau(z)]^2 + \delta z^N$ are double and there are no branch points at all. It is possible for the corresponding potentials to be periodic with period N ($\delta = 0$ in the periodic initial value problem if for some n , $Q(n, 0)R(n, 0) = 1$), but they are not strictly given in terms of generalized Baker-Akhiezer functions. The degenerate nature of this case is also discussed in Appendix A.

In any case, this theorem gives a prescription for generating periodic solutions to the Ablowitz-Ladik equations (1.3) having period N . Choose an arbitrary monic polynomial $\tau(z)$ of degree N with $\tau(0) = 1$, and an arbitrary nonzero constant δ . Uniquely factor $[\tau(z)]^2 + \delta z^{g+1}$ as $s^2(z)p(z)$. The $2g+2$ distinct roots of the polynomial $p(z)$ are the branch points of the curve associated with $\Gamma \in \mathcal{M}'_{N-1}$. The divisor \mathcal{E} associated with $\Gamma \in \mathcal{M}'_{N-1}$ is equal to $m_1 w_1 + \dots + m_j w_j$ where the w_l are the roots of the polynomial $s(z)$ of multiplicity m_l . The point 0^+ must be chosen so that

$$(5.25) \quad y(0^+) = s(0)^{-1} = \left[\prod_{l=1}^j (-w_l)^{m_l} \right]^{-1}.$$

If $g = N - 1$, the solution is given by the formulas (3.12) and (3.13). If the genus is less, the solution is generated from a generalized Baker-Akhiezer function as described in Appendix A.

In particular, it follows from the condition $\tau(0) = 1$ that for periodic potentials, the product of the roots of $[\tau(z)]^2 + \delta z^N$ is always equal to 1. If the roots are distinct, then the corresponding Riemann surface will be constructed from branch points whose product is equal to 1. Since generic finite genus potentials do not have such a constraint on their branch points, we see that in order to approximate an arbitrary genus g potential (say with the product of the branch points equal to ζ) by a sequence of periodic potentials of arbitrarily large periods, N , we must use periodic potentials derived from polynomials $\tau_N(z)$ and constants δ_N such that the $2N$ roots of $[\tau_N(z)]^2 + \delta_N z^N$ come together in pairs for large N , *except* for $2g + 2$ roots making up the branch points of the limiting curve, such that the product of these ultimately distinct roots converges to ζ as N increases. Furthermore, the z -projections of the points of the divisors \mathcal{D}_N must align themselves so that in the limit all but g of them are positioned at double points of $[\tau_N(z)]^2 + \delta_N z^N$. This ensures that the limiting aperiodic wavetrain has exactly g degrees of freedom.

Now, we turn our attention toward the initial value problem for spatially periodic initial data. An algorithm for solving this initial value problem was given in [6]; here, we endeavor to provide an explicit closed form formula for the solution. The elements of the monodromy matrix are closely related to the squared eigenfunctions used in the previous section. The connection is given by the relations

$$\begin{aligned}
 \varphi(n, t, z) &= -2S_{12}(n, t, z)\beta(n, t)z^n \exp(i(z - z^{-1})t) , \\
 (5.26) \quad \chi(n, t, z) &= 2S_{21}(n, t, z)\beta(n, t)z^n \exp(i(z - z^{-1})t) , \\
 f(n, t, z) &= [S_{11}(n, t, z) - S_{22}(n, t, z)]\beta(n, t)z^n \exp(i(z - z^{-1})t) ,
 \end{aligned}$$

which follow from comparing the equations for \mathbf{S} to those for the squared eigenfunctions, and using the normalization $f_{g+1}(0, 0) = 1$. The monodromy matrix $\mathbf{S}(0, 0, z)$ is easy to compute from the initial data; using the above relations along with the construction in Lemma 4.2 gives an explicit construction of the solution to the initial value problem:

1. Calculate from the initial data the monodromy matrix $\mathbf{S}(0, 0, z)$.
2. Calculate the appropriate element of \mathcal{M}'_{N-1} . Factor the polynomial $\text{tr}^2 \mathbf{S} - 4 \det \mathbf{S}$ as $s^2(z)p(z)$. The branch points $\{z_k\}$ of the Riemann surface are the distinct roots of $p(z)$. The roots w_l of $s(z)$ having multiplicity m_l make up the divisor $\mathcal{E} = m_1 w_1 + \dots + m_j w_j$. Label the points 0^\pm of the Riemann surface according to $y(0^+) = 1/s(0)$, where

$$s(z) = \prod_{l=1}^j (z - w_l)^{m_l} .$$

3. Calculate the divisor \mathcal{D} that determines the vector Baker-Akhiezer function \mathbf{u} . The z projections of the points of \mathcal{D} are the $N-1$ zeros of the polynomial

$S_{12}(0, 0, z)$ and the sheet indices are determined from the relations

$$\frac{S_{11}(n, t, z(P_j)) - S_{22}(n, t, z(P_j))}{s(z(P_j))} = y(P_j).$$

4. Determine the scaling constant: $\xi = Q(-1, 0)$.
5. The solution to the initial value problem is then given by the explicit (generalized) formulas (A.13) and (A.14), appropriately scaled by the complex parameter ξ .

This solution of the initial value problem is valid whenever there are at least two roots of $\text{tr}^2\mathbf{S} - 4\det\mathbf{S}$ with odd multiplicity, although when there are repeated roots the solution must be built from generalized Baker-Akhiezer functions as described in Appendix A. It is interesting to observe that it is not possible for finite initial data to give rise to a special divisor \mathcal{D} , since a special divisor would cause the formulas (A.13) and (A.14) to be undefined (in particular at $t = 0$). The solution to the initial value problem given here is complete except in the special case of $\det\mathbf{S} = 0$, which is discussed in Appendix A.

6. The extension to twist-periodicity.

In this section, we consider extending the periodic theory to include the case of twist-periodic boundary conditions in n . Choose some $\theta \in [0, 2\pi)$ and some integer N . Then twist-periodic boundary conditions are defined by

$$(6.1) \quad \begin{aligned} Q(n + N, t) &= \exp(i\theta)Q(n, t), \\ R(n + N, t) &= \exp(-i\theta)R(n, t). \end{aligned}$$

These boundary conditions include periodic boundary conditions as a special case when $\theta = 0$.

It is possible to treat those cases of twist-periodic boundary conditions of twist-period N for which θ is a rational multiple of π within the context of the periodic theory, and we will use this fact to demonstrate a predictable degeneracy in the solution procedure that will lead us to the streamlined approach to be given below. Suppose that $\theta = 2\pi p/q$ for relatively prime integers p and q . Then, the potentials are periodic with period $M = qN$. The monodromy matrix for this periodic problem is

$$(6.2) \quad \mathbf{T}(z) = \mathbf{S}^q(z),$$

where the *twist-periodic monodromy matrix* \mathbf{S} is defined by

$$(6.3) \quad \begin{aligned} \mathbf{S}(z) &= \mathbf{Q}^{-1}(\theta)\mathbf{L}(N-1, 0, z)\mathbf{L}(N-2, 0, z)\dots\mathbf{L}(0, 0, z) \\ &= \begin{bmatrix} A(z) & B(z) \\ C(z) & D(z) \end{bmatrix}, \end{aligned}$$

and the twist matrix is given by

$$(6.4) \quad \mathbf{Q}(\theta) = \begin{bmatrix} \exp(i\theta/2) & 0 \\ 0 & \exp(-i\theta/2) \end{bmatrix}.$$

As a consequence of the factorization of the monodromy matrix \mathbf{T} , the polynomials take the factorized forms

$$(6.5) \quad \begin{aligned} \varphi(z) &= -B(z)F_q(z), \\ \chi(z) &= C(z)F_q(z), \\ f(z) &= (A(z) - D(z))F_q(z). \end{aligned}$$

The first few polynomials $F_q(z)$ are

$$(6.6) \quad \begin{aligned} F_1(z) &= 1, \\ F_2(z) &= A(z) + D(z), \\ F_3(z) &= A^2(z) + A(z)D(z) + B(z)C(z) + D^2(z), \\ F_4(z) &= A^3(z) + A^2(z)D(z) + 2A(z)B(z)C(z) \\ &\quad + 2B(z)C(z)D(z) + A(z)D^2(z) + D^3(z). \end{aligned}$$

Thus, the polynomial $f^2(z) - \varphi(z)\chi(z)$ takes the form

$$(6.7) \quad f^2(z) - \varphi(z)\chi(z) = \text{tr}^2(\mathbf{T}) - 4\det(\mathbf{T}) = (\text{tr}^2(\mathbf{S}) - 4\det(\mathbf{S}))F_q^2(z),$$

and it will have roots of even multiplicity at the zeros of the polynomial $F_q(z)$. The totality of the roots of $f^2(z) - \varphi(z)\chi(z)$ are constants of the motion for (1.3). In particular, the roots of $F_q(z)$ are constants of the motion. But then, those roots of $\varphi(z)$ that are in the factor $F_q(z)$ represent zeros of u_1 that are fixed in n and t . These fixed zeros never leave the pole divisor \mathcal{D} , and thus at these points, neither poles nor zeros may be detected for any n or t in the component u_1 . This

result suggests that no more than $N - 1$ poles are really needed in the divisor \mathcal{D} , and correspondingly that the degree of the divisor \mathcal{E} could be (in most cases²⁵) taken to be zero.

Now, we will see how to deflate the problem data to include only that which is essential; we will replace the (true) monodromy matrix \mathbf{T} with the twist-periodic monodromy matrix $\mathbf{S}(z)$. It will then be clear how to extend the theory to cover cases in which $\theta/2\pi$ is irrational.

The twist-periodic boundary conditions on the potentials induce the following relations on the matrices $\mathbf{L}(n, t, z)$ and $\mathbf{B}(n, t, z)$:

$$(6.8) \quad \mathbf{L}(n + N, t, z) = \mathbf{Q}(\theta)\mathbf{L}(n, t, z)\mathbf{Q}^{-1}(\theta),$$

$$(6.9) \quad \mathbf{B}(n + N, t, z) = \mathbf{Q}(\theta)\mathbf{B}(n, t, z)\mathbf{Q}^{-1}(\theta).$$

Using the equations (1.14) and (1.16) along with the above relations shows that the twist-periodic monodromy matrix $\mathbf{S}(z)$ obeys the relations:

$$(6.10) \quad \mathbf{S}(n + 1, t, z) = \mathbf{L}(n, t, z)\mathbf{S}(n, t, z)\mathbf{L}(n, t, z)^{-1},$$

$$(6.11) \quad -i\partial_t\mathbf{S}(n, t, z) = \left[\mathbf{B}(n, t, z), \mathbf{S}(n, t, z) \right],$$

describing the evolution in n and t . It follows that the eigenvalues of $\mathbf{S}(n, t, z)$ are independent of n and t . These are in fact the same differential equations obeyed by the periodic monodromy matrix $\mathbf{T}(n, t, z)$. This shows that the essential dynamical information is contained in the twist-periodic monodromy matrix $\mathbf{S}(n, t, z)$. One builds the Baker-Akhiezer functions u_1 and u_2 from the polynomials

$$(6.12) \quad \begin{aligned} \varphi(z) &= -S_{12}(z), \\ \chi(z) &= S_{21}(z), \\ f(z) &= S_{11}(z) - S_{22}(z). \end{aligned}$$

Prior to using these polynomials to construct the Baker-Akhiezer function \mathbf{u} , all three must be multiplied by a common normalization factor to make $f(z)$ be a monic polynomial. One then constructs the potentials from these polynomials exactly as described in the proof of Lemma 4.2. The corresponding potentials $Q(n, t)$ and $R(n, t)$ solve the initial value problem where the polynomials are constructed from $\mathbf{S}(0, 0, z)$ and the scaling parameter $\xi = Q(-1, 0)$ is used.

Although this approach was suggested by considering those values of θ that were rational multiples of 2π , and consequently embedding the twist-periodic problem into a larger periodic problem, θ can now be taken to be an arbitrary real number in $[0, 2\pi)$.

²⁵It is possible that even after taking into account the degeneracy in \mathbf{T} arising from the twist-periodicity there still remain multiple roots in the polynomial $\text{tr}^2(\mathbf{S}) - 4\det(\mathbf{S})$, however the roots are generically distinct.

To construct the solution with the least redundancy in its expression, it is important to find the smallest N for which the initial data is twist-periodic for some θ . Even if one has found the smallest such N , the polynomial $f^2(z) - \varphi(z)\chi(z)$ may still have repeated roots in some special cases. In these cases, the Baker-Akhiezer functions u_1 and u_2 will have to be constructed from data having $r > 0$, where r is the number of homogeneous linear conditions imposed on the generalized Baker-Akhiezer functions u_1 and u_2 as discussed in Appendix A. There is evidence that this degeneracy vanishes if one considers focusing or defocusing potentials. In these cases, choosing the smallest possible N may guarantee that $r = 0$.

7. Modulation equations.

In this section, we begin to describe the modulational behavior of the solutions to (1.3) constructed above. The procedure we follow was first used by Flaschka, Forest, and McLaughlin [20] with respect to the Korteweg-de Vries equation, and was recently used by Kodama and Bloch [5] to study the Toda lattice. The modulation equations we will derive are based on an averaging principle that is merely postulated as a correct description of a modulated wavetrain. Thus, the technique is more formal than analytical. However it has been shown by Krichever [32] that for general integrable systems expressible as 2-by-2 zero-curvature conditions²⁶, these formal equations must be satisfied by any slowly varying wavetrain in order for the true evolution to be uniformly approximated to leading order (in the asymptotic limit of long modulational scales) by exact finite genus solutions.

Modulation theory begins with local conservation laws that control the microscopic motion. In the integrable setting, it is possible to write down a generator that gives rise to an infinite number of local conservation laws. Let us construct this generator. It is a consequence of the commutativity of the unit shift operator in n , Δ (defined by $\Delta a(n) = a(n+1)$), and differential operators in t that the first component of the Baker-Akhiezer function satisfies

$$(7.1) \quad \partial_t \log \left(\frac{\Delta u_1(n, t, P)}{u_1(n, t, P)} \right) = (\Delta - 1) \left(\frac{1}{u_1(n, t, P)} \partial_t u_1(n, t, P) \right).$$

This is an equation for functions on the Riemann surface; with an application of the operator d of differentiation on the surface, it becomes an equation for differentials

$$(7.2) \quad \partial_t F(n, t, P; (\Gamma, \mathcal{D}, \xi)) = (\Delta - 1) G(n, t, P; (\Gamma, \mathcal{D}, \xi)),$$

where F is a generating function for local conserved densities:

$$(7.3) \quad F(n, t, P; (\Gamma, \mathcal{D}, \xi)) = d \log \left(\frac{\Delta u_1(n, t, P)}{u_1(n, t, P)} \right),$$

²⁶The τ -function of the universal Whitham hierarchy giving the formal modulational description of these systems was given by Krichever in [33]

and G is a generating function for the corresponding fluxes:

$$(7.4) \quad G(n, t, P; (\Gamma, \mathcal{D}, \xi)) = d\partial_t \log u_1(n, t, P).$$

The parametric dependence on P allows an infinite number of local conservation laws to be derived, for example by expanding in a formal series near a singular point such as $P = \infty^+$. Note that although the density generator F is manifestly in the range of $\Delta - 1$ and thus seems to be a trivial conserved density, the individual expansion coefficients in a formal series are not trivial, since the operator $\Delta - 1$ is acting on a nonlocal sequence in n . Indeed, the coefficients turn out to be the nontrivial local conserved densities for (1.3) [41]. Here, however, we retain the dependence on P in order to average over rapid oscillations in all local conservation laws simultaneously. Consider the equation (7.2) in the case when the reconstructed potentials $R(n, t)$ and $Q(n, t)$ are quasiperiodic functions of n and t , for real t . Imagine the data $(\Gamma, \mathcal{D}, \xi)$ to be no longer fixed, but to depend upon the slow variables $X = hn$ and $T = ht$ where h is the small lattice spacing (to denote this slow dependence through the data, we will write $F = F(n, t, X, T)$ and $G = G(n, t, X, T)$). For a function w depending on both the fast and slow scales, make the usual multiple scale replacements

$$(7.5) \quad \Delta w \rightarrow \Delta w + h\partial_X \Delta w,$$

and

$$(7.6) \quad \partial_t w \rightarrow \partial_t w + h\partial_T w.$$

The conservation laws then take the form

$$(7.7) \quad \begin{aligned} \partial_t F(n, t, X, T) + h\partial_T F(n, t, X, T) \\ = (\Delta - 1)G(n, t, X, T) + h\partial_X \Delta G(n, t, X, T). \end{aligned}$$

The averaging operation for *fixed* X and T

$$(7.8) \quad \langle w(n, t, X, T) \rangle = \lim_{N \uparrow \infty} \frac{1}{2N+1} \sum_{n=-N}^N w(n, t, X, T),$$

where the resulting object depends on X and T only because we assume that the fast variables n and t are tied together in $g+1$ independent phases²⁷, is used to eliminate the order 1 terms²⁸, yielding

$$(7.9) \quad \partial_T \langle F \rangle = \partial_X \langle \Delta G \rangle.$$

²⁷If the phases are not independent, so that the evolution in n and t does not cover the Jacobian ergodically, it is necessary to proceed differently, introducing additional modulation equations for the locked phases. Modulation theory in the presence of phase locking is discussed in [42].

²⁸The averaging operation is applied directly to the equation (7.7). The order 1 terms vanish because they vanish for fixed X and T , and the average of 0 is 0. The averaging operation commutes with the slow differential operators ∂_X and ∂_T .

The equation (7.9) contains all the information about the modulational behavior of a slowly varying $g + 1$ phase wavetrain. The first task at hand is to realize this equation in terms of the Baker-Akhiezer function data. Consider the averaged conserved density generator $\langle F \rangle$:

$$(7.10) \quad \langle F \rangle = \left\langle d \log \frac{\Delta u_1(n, t, P)}{u_1(n, t, P)} \right\rangle.$$

Replacing u_1 using the formula (2.32) and keeping only those terms depending on $P \in \Gamma$ due to the presence of the operator d transforms this into

$$(7.11) \quad \begin{aligned} \langle F \rangle &= \left\langle \omega_{(3)} + d \log \frac{\Theta(A(P) - \mathbf{Z} + \mathbf{U}(n+1) + \mathbf{V}t)}{\Theta(A(P) - \mathbf{Z} + \mathbf{U}n + \mathbf{V}t)} \right\rangle \\ &= \omega_{(3)} + \left\langle d \log \frac{\Theta(A(P) - \mathbf{Z} + \mathbf{U}(n+1) + \mathbf{V}t)}{\Theta(A(P) - \mathbf{Z} + \mathbf{U}n + \mathbf{V}t)} \right\rangle. \end{aligned}$$

Similarly, the averaged flux generator $\langle \Delta G \rangle$ takes the concrete form

$$(7.12) \quad \begin{aligned} \langle \Delta G \rangle &= \left\langle \omega_{(2)} + d \frac{\partial_t \Theta(A(P) - \mathbf{Z} + \mathbf{U}(n+1) + \mathbf{V}t)}{\Theta(A(P) - \mathbf{Z} + \mathbf{U}(n+1) + \mathbf{V}t)} \right\rangle \\ &= \omega_{(2)} + \left\langle d \frac{\partial_t \Theta(A(P) - \mathbf{Z} + \mathbf{U}(n+1) + \mathbf{V}t)}{\Theta(A(P) - \mathbf{Z} + \mathbf{U}(n+1) + \mathbf{V}t)} \right\rangle. \end{aligned}$$

So, the modulation equations take the form

$$(7.13) \quad \begin{aligned} \partial_T \omega_{(3)} &- \partial_X \omega_{(2)} \\ &= \partial_X \left\langle d \frac{\partial_t \Theta(A(P) - \mathbf{Z} + \mathbf{U}(n+1) + \mathbf{V}t)}{\Theta(A(P) - \mathbf{Z} + \mathbf{U}(n+1) + \mathbf{V}t)} \right\rangle \\ &- \partial_T \left\langle d \log \frac{\Theta(A(P) - \mathbf{Z} + \mathbf{U}(n+1) + \mathbf{V}t)}{\Theta(A(P) - \mathbf{Z} + \mathbf{U}n + \mathbf{V}t)} \right\rangle. \end{aligned}$$

Modulation equations for multiphase waves in integrable systems have appeared in the literature, but *without* the terms involving the theta averages. We call these terms *gauge terms* because their form may be altered by making a change of specific choice of homology basis cycles²⁹ on the surface $\Gamma(X, T)$. The gauge group of the modulation equations is the group that takes cycles to

²⁹We want to emphasize that the modulational equations themselves are gauge invariant, since they derive from the Baker-Akhiezer function which does not depend on any choice of cycles. Only the *form* of the modulational equations is affected by a change of gauge.

cycles, preserving the intersection numbers; this group is $Sp(g, \mathbb{Z})$, the group of symplectic $2g$ by $2g$ matrices that have integer entries. In the cases worked out in the literature (the Korteweg-de Vries equation [20], the cubic nonlinear Schrödinger equation [21], and the Toda chain [5]), an implicit choice of gauge was made that causes the gauge terms to vanish entirely. If we could make such a choice of gauge in the modulation equations (7.13), the differentials $\omega_{(2)}$ and $\omega_{(3)}$ would become normalized with respect to a distinguished basis of cycles; we could write these as $\Omega_{(2)}$ and $\Omega_{(3)}$. Then, the modulation equations would take on the familiar form

$$(7.14) \quad \partial_T \Omega_{(3)} = \partial_X \Omega_{(2)}.$$

Modulation equations of the form (7.14) can be cast into a revealing form by integration of them around b -cycles on the Riemann surface

$$(7.15) \quad \partial_T \mathbf{U} = \partial_X \mathbf{V},$$

and by expanding the path integral of (7.14) near 0^+ and ∞^+ on the Riemann surface to obtain

$$(7.16) \quad \partial_T (G_3 - g_3) = \partial_X (G_2 - g_2).$$

Together, equations (7.15) and (7.16) make up $g + 1$ equations representing wave conservation in the $g + 1$ phases. They are $g + 1$ equations on $2g + 2$ unknowns (in the generic case of $\Gamma(X, T) \in \mathcal{M}_g$, $\Gamma(X, T)$ is parametrized by the $2g + 2$ branch points and the sheet index of η). The remaining half of the modulational degrees of freedom contribute to defining the wave size and shape. Their modulational description would be contained in (7.14) as well.

In the case that the branch points z_i are invariant as a set under reflection through the unit circle in the z plane, it is possible to remove the gauge terms. In particular, the modulation equations for the focusing and defocusing versions of the Ablowitz-Ladik equations (1.2) can be put into the form (7.14). First, observe that in order for the gauge terms to vanish, it is sufficient for the wavenumber and frequency vectors, U and V , to have pure imaginary components. These components are the integrals of the differentials $\omega_{(2)}$ and $\omega_{(3)}$ around b -cycles; this fact is the link between the choice of cycles and the vanishing of the gauge terms. Just as we refer to the conditions necessary to cause the integrals of $\omega_{(2)}$ and $\omega_{(3)}$ over a -cycles to vanish as normalization conditions, we refer to the conditions necessary to cause the integrals of $\omega_{(2)}$ and $\omega_{(3)}$ over b -cycles to be imaginary as *gauge conditions*. Whenever the unit circle symmetry is present, there exists on the surface Γ an anti-holomorphic involution given by

$$(7.17) \quad I : (z, y) \mapsto (1/\bar{z}, \eta \bar{y} / \bar{z}^{g+1}).$$

This involution takes a Riemann surface whose branch points are symmetrical through the unit circle to itself, changing the complex structure everywhere by a sign. It permutes 0^+ with ∞^+ and 0^- with ∞^- . The differentials $\omega_{(2)}$ and $\omega_{(3)}$

transform by the involution I in such a way that it is easy to choose cycles with respect to which they should be normalized. We use an unnormalized variant of the Abel map to identify the homology group of cycles on Γ with the period lattice Λ in \mathbb{C}^g :

$$(7.18) \quad \begin{array}{ccc} \mathbf{h} : H_1(\Gamma, \mathbb{Z}) & \rightarrow & \Lambda \\ C & \mapsto & \int_C (dz/y, zdz/y, \dots, z^{g-1}dz/y)^T. \end{array}$$

Through this map, the involution I acts on \mathbb{C}^g :

$$(7.19) \quad \begin{aligned} h_k(I(C)) &= \int_{I(C)} \frac{z^{k-1}dz}{y} \\ &= \int_C I^* \left(\frac{z^{k-1}dz}{y} \right) \\ &= -\bar{\eta} \int_C \frac{\bar{z}^{g-k}d\bar{z}}{\bar{y}} \\ &= -\bar{\eta} \overline{h_{g+1-k}(C)}. \end{aligned}$$

It is clear that the involution I maps $H_1(\Gamma, \mathbb{Z})$ onto itself; thus, the action on the lattice Λ is also onto. The action of I on \mathbb{C}^g has a g real dimensional subspace of fixed points \mathbb{A} , made up of complex vectors \mathbf{a} whose components obey the symmetry

$$(7.20) \quad a_j = -\bar{\eta} \overline{a_{g+1-j}}.$$

This subspace contains g lattice points, independent over \mathbb{Z} , as a consequence of the fact that whenever $\mathbf{h}(C)$ is in Λ , $\mathbf{h}(I(C))$ is also in Λ , and thus the sum $\mathbf{h}(C) + \mathbf{h}(I(C))$ is a lattice vector in the subspace \mathbb{A} . These lattice points correspond to cycles C invariant under the involution, so that $I(C) = C$. Likewise, the orthogonal complement³⁰ of this subspace, \mathbb{A}^\perp contains a g dimensional real sublattice corresponding to cycles C such that $I(C) = -C$. Every lattice point $\mathbf{h}(C)$ in Λ can be uniquely written as the sum of a lattice point in \mathbb{A} and a lattice point in \mathbb{A}^\perp . The real sublattice contained within the subspace \mathbb{A} is generated by g independent cycles in $H_1(\Gamma, \mathbb{Z})$. For any two such cycles, C_1 and C_2 , the canonical intersection form can be written as

$$(7.21) \quad C_1 \circ C_2 = I(C_1) \circ I(C_2),$$

because the involution does not affect cycles in \mathbb{A} . Since the involution I acts on the canonical intersection number (for general cycles C_1 and C_2) as:

$$(7.22) \quad I(C_1) \circ I(C_2) = C_2 \circ C_1,$$

³⁰The orthogonal complement is taken with respect to the Euclidean inner product on $\mathbb{R}^{2g} \cong \mathbb{C}^g$.

it follows from the skew symmetry of the intersection form that for any pair of cycles both in \mathbb{A} , the canonical intersection number vanishes. We choose g independent lattice points in \mathbb{A} and call them $\mathbf{a}_1, \dots, \mathbf{a}_g$ (the corresponding cycles on Γ are denoted a_1, \dots, a_g); the canonical basis of cycles is completed by selecting g nonintersecting loops b_1, \dots, b_g so that $a_i \circ b_j = \delta_{ij}$.

Now, we consider the differentials $\omega_{(2)}$ and $\omega_{(3)}$. Begin by writing the differentials in the form

$$(7.23) \quad \omega_{(2,3)} = f^{(2,3)}(z) dz + \frac{g^{(2,3)}(z)}{y} dz + \frac{h^{(2,3)}(z)}{y} dz,$$

where

$$(7.24) \quad g^{(2)}(z) = i \left[\frac{1}{2} z^{g+1} + \frac{1}{2} \frac{\eta}{z^2} - \frac{1}{4} \frac{\eta}{z} \sum_{j=1}^{2g+2} \frac{1}{z_j} - \frac{1}{4} z^g \sum_{j=1}^{2g+2} z_j \right] dz,$$

$$(7.25) \quad g^{(3)}(z) = \frac{1}{2} \left[-\frac{\eta}{z} + z^g \right] dz,$$

and the freedom of normalization has been absorbed into the $g - 1$ degree polynomials $h^{(2,3)}(z)$. Since $z = 0$ and $z = \infty$ cannot be branch points, we may choose concrete contours on Γ representing any homology class such that these contours have z projections that do not encircle $z = 0$ or $z = \infty$. Using these representatives, the integral of the y independent terms $f^{(2,3)}(z) dz$ over any cycle vanishes, so they cannot contribute to the normalization conditions or the gauge conditions. The symmetry induced by the involution I is given in these terms by

$$(7.26) \quad I^* \frac{g^{(2,3)}(z)}{y} dz = \overline{\left[\frac{g^{(2,3)}(z)}{y} dz \right]},$$

where the asterisk denotes the pull-back. There is a g dimensional *real* subspace of holomorphic differentials obeying

$$(7.27) \quad I^* \alpha = \bar{\alpha}.$$

Notice that any differential obeying $I^* \alpha = \bar{\alpha}$ will have real valued integrals over all a-cycles, since

$$(7.28) \quad \oint_{a_j} \alpha = \oint_{I(a_j)} \alpha = \oint_{a_j} I^* \alpha = \oint_{a_j} \bar{\alpha}.$$

Thus, it is possible to normalize the differentials $\omega_{(2)}$ and $\omega_{(3)}$ to have vanishing integrals over a-cycles by adding a differential in the class described by (7.27). Ignoring those terms that do not contribute to integrals over closed cycles, the two normalized differentials, which we may now refer to as $\Omega_{(2)}$ and $\Omega_{(3)}$, also

have the symmetry of (7.27), since the two contributing terms in the expression (7.23) do. Now consider the integral

$$(7.29) \quad \oint_{b_j} \Omega,$$

where Ω is either $\Omega_{(2)}$ or $\Omega_{(3)}$. The cycle b_j can be decomposed as

$$(7.30) \quad b_j = \sum_{k=1}^g m_{jk} a_k + \tilde{a},$$

where the m_{jk} are integers and \tilde{a} corresponds to a lattice point $\tilde{\mathbf{a}}$ in the subspace \mathbb{A}^\perp . As a result of the normalization,

$$(7.31) \quad \oint_{b_j} \Omega = \oint_{\tilde{a}} \Omega.$$

This integral is pure imaginary, since

$$(7.32) \quad \oint_{\tilde{a}} \Omega = - \oint_{I(\tilde{a})} \Omega = - \oint_{\tilde{a}} I^* \Omega = - \oint_{\tilde{a}} \bar{\Omega}.$$

Thus, it is possible in the focusing and defocusing cases to choose a homology basis so that \mathbf{U} and \mathbf{V} have imaginary components. Similar arguments show that the same basis makes $G_3 - g_3$ and $G_2 - g_2$ imaginary, so that there is no growth from the exponential factor in the formula for Q either. Not only does this make the gauge terms in the modulation equations vanish, but it also proves that the focusing or defocusing potential $Q(n, t)$ is a multiphase wavetrain that is quasiperiodic in n (for fixed t) and in t (for fixed n).

Incidentally, the preceding arguments establish for any Riemann surface whose set of branch points is symmetrical about the unit circle in the z -plane the existence of a real subtorus of the Jacobian. This subtorus contains the images under the Abel map (7.18) of divisors giving rise to focusing or defocusing potentials. Given that, the reality conditions presented in Section 4 might be expressible entirely in terms of the involution I , using an approach similar to that employed by Krichever [31] in his study of the Kadomtsev-Petviashvili equation. Whether this method works depends crucially on the nature of the fixed point set of I ; there needs to be a large enough set of fixed points to separate the Riemann surface into two disjoint pieces. The fixed points of I must lie on one of the two unit circles on the Riemann surface, however the exact structure of the fixed point set depends on the number of branch points on the unit circle and the sign of η .

In the focusing and defocusing cases then, we consider a specific choice of cycles that removes the gauge terms and results in the system

$$(7.33) \quad \partial_T \Omega_{(3)} = \partial_X \Omega_{(2)}.$$

We continue by writing these modulation equations in Riemann invariant form, in which the individual motions of the branch points z_i are coupled only through $2g + 2$ characteristic velocities. This form of the modulation equations will help clarify the behavior of modulated wavetrains by revealing conditions under which the equations are hyperbolic and well-posed or elliptic and ill-posed. Expand the equations about the branch point $z = z_k$. Differentiation with respect to X or T of the local parameter $\sqrt{z - z_k}$ results in a second order pole in each term of (7.33). The equality of the coefficients of these poles is the expression

$$(7.34) \quad \partial_T z_k + c(z_k, \mathbf{z}) \partial_X z_k = 0,$$

a system of partial differential equations for the branch points of the labeled surface $\Gamma(X, T) \in \mathcal{M}_g$ with symmetrical characteristic speeds given by the function

$$(7.35) \quad c(z, \mathbf{z}) = \frac{z^{g+1} + \frac{\eta}{z^2} - \frac{\eta}{2z} \left(\sum_{j=1}^{2g+2} \frac{1}{z_j} \right) - \frac{z^g}{2} \left(\sum_{j=1}^{2g+2} z_j \right) + 2P_{g-1}^{(2)}(z, \mathbf{z})}{-\frac{\eta}{z} + z^g + 2P_{g-1}^{(3)}(z, \mathbf{z})}.$$

The characteristic speeds depend transcendently on the modulational variables $\{z_1 \dots z_{2g+2}\}$ through the coefficients of the polynomials $P_{g-1}^{(2)}$ and $P_{g-1}^{(3)}$ arising from the normalization of $\Omega_{(2)}$ and $\Omega_{(3)}$ with respect to the distinguished set of cycles $a_1, \dots, a_g, b_1, \dots, b_g$. In the symmetric case we have been describing, it is easy to see that the speed function must have the symmetry

$$(7.36) \quad \overline{c(\bar{z}^{-1}, \mathbf{z})} = c(z, \mathbf{z}).$$

Thus, if a branch point z_k lies on the unit circle, then its characteristic velocity will be real; its corresponding equation of motion becomes

$$(7.37) \quad \partial_T \theta_k + c(\exp(i\theta_k), \mathbf{z}) \partial_X \theta_k = 0,$$

where $z_k = \exp(i\theta_k)$. Thus the motion confines the point $z_k(X, T)$ to the unit circle. Furthermore, if all branch points are distinct and lie on the unit circle, then all the characteristic velocities are real and distinct; the modulation equations are strictly hyperbolic³¹. There are only two ways that this situation might be dynamically destroyed:

- The dynamics in X and T might cause two branch points to collide and then bifurcate off of the circle, leading to modulation equations that are locally elliptic, or

³¹This is a sufficient condition for hyperbolicity, but not necessary. It is possible for the characteristic speeds to be real although some branch points are not on the unit circle. See [42] for an example of a class of focusing potentials that are modulationally stable, in spite of the fact that none of the branch points lie on the unit circle.

- The dynamics in X and T might cause one of the quantities $\partial_X \theta_k$ to become infinite for finite X and T signaling the appearance of a hyperbolic shock.

In fact, the collision of branch points on the unit circle is impossible as long as the modulation equations are classical, since

$$(7.38) \quad |c(\exp(i\theta + i\epsilon), \mathbf{z}) - c(\exp(i\theta), \mathbf{z})| = \mathcal{O}(\epsilon),$$

where $\epsilon \ll 1$ measures the distance between two nearby branch points. Thus, the two branch points begin to move with the same velocity as they approach each other on the circle, and (7.38) shows that their velocities align sufficiently fast that the two branch points cannot catch up with each other in finite time. Of course the interpretation of $c(z, \mathbf{z})$ as a velocity is only valid when $\partial_X \theta_k$ exists; thus it is possible for branch points to collide on the unit circle, but shocks must then form also where the branch points coalesce. This result is the converse of the result of Levermore [36] for the Korteweg-de Vries modulation equations which says that shocks may only form only if branch points coalesce. The fact that the modulation equations remain hyperbolic as long as they remain classical is a familiar feature of the modulation equations of known integrable systems, and here we see that it follows from the fact that in integrable systems, the characteristic velocities are expressed as a single function evaluated at the various values of the dynamical variables. The analyticity of the function $c(\cdot, \mathbf{z})$ in its first argument leads directly to (7.38).

It is significant that in the defocusing case, there are two possibilities for the modulational behavior. Some branch point configurations have all branch points on the unit circle so that the modulation equations are hyperbolic, suggesting stable modulational behavior. However, there are also branch point configurations having some branch points split off of the unit circle in pairs; in this case the modulation equations may be elliptic, suggesting unstable behavior. This result should be contrasted with the modulational behavior of the continuum nonlinear Schrödinger equation, where in the defocusing case *all* admissible branch point configurations lead to hyperbolic modulation equations.

8. Discussion of results and future work.

We have presented in this paper a class of exact solutions to the Ablowitz-Ladik equations (1.3), along with criteria for selecting from this class of solutions those that are spatially periodic with a given period N , and those that are solutions to the focusing or defocusing version of the discrete nonlinear Schrödinger equation (1.2). The corresponding formal modulational equations have also been presented; these are the first ingredients in a program to understand the complete picture of macroscopic behavior in the discrete nonlinear Schrödinger equation.

We have left many questions unanswered for our future work to address. For instance, we would like to understand the Hamiltonian structure of the complex Ablowitz-Ladik equations (1.3) restricted to the genus g component of the phase space. For the strictly periodic case there is really no problem here, since in

the nonsingular cases there are $g + 1 = N$ independent branch points (action variables) and $g + 1$ corresponding angle variables. For finite genus potentials that are not periodic functions of n , however, finding the Hamiltonian structure will require identifying which of the $3g + 3$ coordinates are Casimir functions of the Poisson bracket

$$(8.1) \quad \begin{aligned} & \{f, g\} \\ &= \lim_{N \uparrow \infty} \frac{i}{2N + 1} \sum_{n=-N}^N (1 - Q(n)R(n)) \left[\frac{\partial f}{\partial Q(n)} \frac{\partial g}{\partial R(n)} - \frac{\partial g}{\partial Q(n)} \frac{\partial f}{\partial R(n)} \right]. \end{aligned}$$

The remaining coordinates can then be partitioned into canonical (complex) action and angle variables. In the focusing and defocusing cases, the actions should be real valued, and the potentials should be periodic functions of all the angle variables.

In the focusing and defocusing cases, the modulation equations found in Section 7 can also be given a Poisson structure. A general treatment of the relation of modulation equations to flat metrics and Poisson brackets of hydrodynamic type can be found in the paper [33] by Krichever. In some cases, this structure can be exploited to explicitly solve the modulation equations by a generalized hodograph transform first worked out by Tsarev and summarized by Dubrovin [12].

Although we have identified coordinates for the genus g part of the phase space of (1.3) in the general complex case, we have not yet addressed the topological properties of the genus g component of the phase space. It is likely that the manifold of complex genus g solutions is connected, but numerical constructions of exponentially growing potentials using the formulas (3.12) and (3.13) show that this manifold is not compact. Imposing reality on the complex solution manifold to isolate the focusing and defocusing submanifolds has the effect of destroying the connectedness. In particular, both the focusing and defocusing submanifolds of the phase space consist of at least two disconnected components, corresponding to the fact that the analytic double covering present in the set \mathcal{M}_g through the parameter η degenerates into two distinct copies of the unit circle on which η must live. Assessing the compactness of these components in the focusing and defocusing cases requires understanding where the special divisors lie in the Jacobian, as in the approach of Previato [43]. It is known that there is a noncompact component in the defocusing case, because defocusing initial conditions of sufficiently large amplitude experience blowup in finite time [26]. It should be possible to characterize exactly which spectral data lead to this catastrophic phenomenon.

We are interested in singular limits of the class of solutions to (1.3) that we have described. One kind of singular limit is the limit of infinite genus. In some cases, this limit can correspond to the continuum limit of smooth potentials. But also, we expect to find in this limit at least in the focusing and defocusing

cases, all discrete solutions corresponding to initial conditions that are bounded as functions of n .

As shown in Appendix A, a more tractable kind of singular limit is the limit of colliding branch points. The class of generalized potentials given there trivially includes the potentials satisfying twist-periodic boundary conditions and described in Section 6. The particular case of solutions having antiperiod 2 is included in this class; these solutions correspond to genus 1, and have the interesting property (observed numerically, and described in the context of finite genus modulation theory in [42]) that they do not evolve into more general conditionally periodic genus 1 waves under the flow of the modulation equations. But potentials in the limit of colliding branch points can behave in other ways also. In some cases, pinching the Riemann surface together in this way is known to lead to homoclinic orbits and (dark) solitons. We would like to find the solitons in formulas like (A.13) and (A.14) and describe more explicitly the Bäcklund transformations that generate these solitons from quasiperiodic potentials. The first steps in this program are already being carried out [17]. Yet another possible result of colliding branch points is that the limiting potential may achieve even or odd symmetry in n . For example, this is the situation in any neighborhood near the origin $n = 0$ of the Toda shock problem in the long time limit [49], where the pinched genus 2 potential that appears after a long time is constrained by the symmetry of the initial shock at the origin to be odd as a function of n .

Such long time limits as those explicitly calculated by Venakides, Deift and Oba [49] in the context of the Toda shock problem (and similarly the zero-dispersion limits first obtained for the Korteweg-de Vries equation by Lax and Levermore [34] and later extended by Venakides [48]) are the real proof of the validity of modulation equations of the type found in Section 7. These calculations always begin with a specific (but to some degree arbitrary) exact solution, and the long time asymptotics are calculated explicitly using the inverse spectral transform. The long time limit is given locally by a finite genus oscillation, with macroscopic behavior given by the same modulation equations that can be (formally, but more easily) obtained by other means. We would like to consider verifying the modulation equations given in Section 7 with similar calculations. The most effective method would be to recast the solution of (1.3) (considered on the whole line) as the solution of a matrix valued Riemann-Hilbert problem. Then, the very effective methods of Deift, Its, and Zhou [10] can be employed to examine the long time behavior of the Riemann-Hilbert problem, resulting in a description of the long time dynamics that is a natural nonlinear generalization of the steepest descent expansions familiar from the asymptotic theory of linear dispersive wave systems.

Regarding our original impetus for understanding the finite genus solutions to (1.2), we need to undertake a more detailed study of the hyperbolic structure of the modulation equations, and their corresponding evolutionary behavior. First, we need to consider the modulation equations in the general complex case, in order to determine how the modulation equations can be interpreted in the case that the spectrum has no particular symmetries. If the modulation equations

turn out to have some meaning in the complex case, we would like to understand what kind of simplifications can be introduced by careful choice of gauge. Even in the focusing and defocusing cases, there are questions to be answered. The most important question is that of the qualitative difference between potentials whose branch points lie on the unit circle having stable modulational behavior and potentials whose branch points have split off the unit circle in pairs. In Appendix B we have proved that in the focusing case *all* branch points lie off of the unit circle, and that in the defocusing case restricted to the component $\Sigma_{\mathbb{Z}}$ (from which the continuum limit emerges) the branch points are constrained all to lie on the unit circle. However, there is very little known about solutions belonging to the other components Σ_{κ} of the defocusing phase space. Finally, in the defocusing hyperbolic cases, we need to understand the consequences of wave breaking and shock formation. We expect that, as has been shown with other integrable systems, steepening fields in the modulation equations can be regularized locally by describing the local microscopic behavior with a solution of higher genus. Determining whether something similar is true in the case of the Ablowitz-Ladik equations requires comparing the formal structure of the genus g modulation equations with that of the genus $g + 1$ modulation equations.

Finally, we want to use the machinery presented in this paper to extend the results to the infinite hierarchy of flows that commute with the Ablowitz-Ladik equations (1.3). A systematic algebraic investigation of this hierarchy of equations was carried out by Schilling [44], who found a generator for a doubly infinite sequence of temporal linear problems that commute with the spatial linear problem (1.14) under the condition that the potentials $Q(n, t)$ and $R(n, t)$ satisfy certain nonlinear equations. In the context of the current paper, these “higher” Ablowitz-Ladik equations arise by specifying a new differential indexed by an integer $k = 2, 3, \dots, \omega_{(2)}^{(k)}$, in the place of $\omega_{(2)}$ (which we would write in this scheme as $\omega_{(2)}^{(1)}$) having singularities of order k at 0^+ and ∞^+ . For example, in the $k = 2$ case, if we specify that the new Baker-Akhiezer function \mathbf{u} behave as

$$(8.2) \quad \mathbf{u} = \begin{cases} \tilde{\mathbf{u}}, & P \rightarrow 0^-, \\ \exp(i(1 - z^{-2})t)\tilde{\mathbf{u}}, & P \rightarrow 0^+, \\ \tilde{\mathbf{u}}, & P \rightarrow \infty^-, \\ \exp(i(z^2 - 1)t)\tilde{\mathbf{u}}, & P \rightarrow \infty^+, \end{cases}$$

where the function $\tilde{\mathbf{u}}$ has the same n dependence as the vector Baker-Akhiezer function we have used in this paper, then this new function solves the usual spatial linear problem (1.14), and it is possible to deduce a “higher” temporal linear problem satisfied as well. The formulas for the new Baker-Akhiezer function are identical to (2.32) and (2.33) with the differential $\omega_{(2)}$ replaced by a new differential $\omega_{(2)}^{(2)}$. The consistency relation between the two linear problems is a

new nonlinear system, the second Ablowitz-Ladik equation. It takes the form

$$\begin{aligned}
 (8.3) \quad -i\partial_t Q(n) &+ \sigma(n) \left[(2 + \kappa(n))Q(n) + R(n) \left(Q(n+1)^2 + Q(n-1)^2 \right) \right. \\
 &\quad \left. - \left(\sigma(n-1)Q(n-2) + \sigma(n+1)Q(n+2) \right) \right] = 0, \\
 -i\partial_t R(n) &- \sigma(n) \left[(2 + \kappa(n))R(n) + Q(n) \left(R(n+1)^2 + R(n-1)^2 \right) \right. \\
 &\quad \left. - \left(\sigma(n-1)R(n-2) + \sigma(n+1)R(n+2) \right) \right] = 0,
 \end{aligned}$$

where

$$(8.4) \quad \sigma(n) = 1 - Q(n)R(n),$$

$$(8.5) \quad \kappa(n) = Q(n-1)R(n+1) + Q(n+1)R(n-1).$$

We remark that this system is Hamiltonian as well³², with the Poisson bracket (8.1) and the Hamiltonian

$$\begin{aligned}
 (8.6) \quad H^{(2)} &= \lim_{N \uparrow \infty} \frac{1}{2N+1} \sum_{n=-N}^N \left\{ Q(n)R(n)\kappa(n) \right. \\
 &\quad \left. + \frac{1}{2} R(n)^2 \left(Q(n+1)^2 + Q(n-1)^2 \right) \right. \\
 &\quad \left. - R(n) \left(Q(n+2) + Q(n-2) \right) + 2 \log \sigma(n) \right\}.
 \end{aligned}$$

Formally, at least, all statements we have made about the system (1.3) also hold for the higher system (8.3) by the formal replacement of $\omega_{(2)}$ by $\omega_{(2)}^{(2)}$. This procedure can be continued, resulting in a sequence of differentials $\omega_{(2)}^{(k)}$ controlling the time dependence of the Baker-Akhiezer function. The behavior of modulated solutions to the k -th flow will depend upon symmetry properties of these differentials. In particular, we expect the flows corresponding to odd k to have stability properties similar to those of modulated solutions to the system (1.3) we have considered at length in this paper.

Acknowledgement. The authors wish to thank Annalisa Calini for useful discussions, and Mark Hays for his careful notes on the paper [6].

The authors gratefully acknowledge the support provided to them: N.M.E. from the NSF under grant DMS-8942390 and from the Air Force under grant

³²The Hamiltonian of the Ablowitz-Ladik equations (1.3) with respect to the Poisson bracket (8.1) is

$$H^{(1)} = \lim_{N \uparrow \infty} \frac{1}{2N+1} \sum_{n=-N}^N Q(n) \left(R(n+1) + R(n-1) \right) + 2 \log \sigma(n).$$

AFOSR-FQ8671-9000589; C.D.L. and P.D.M. from the Air Force under grant AFOSR-90-021 and from the NSF under grant DMS-9404570; I.M.K. was partly supported by RFFI grant 93-011-16087 and ISF grant MD 8000. In addition, some of this work was carried out while N.M.E., C.D.L., and P.D.M. were visiting the Mathematical Sciences Research Institute (MSRI) in Berkeley, which is supported in part by the NSF under grant DMS-9022140.

Appendix A: Solutions derived from generalized Baker-Akhiezer functions.

The class of algebro-geometric solutions constructed in this paper is easily extended to include singular data giving rise to Riemann surfaces with nondistinct branch points. The approach taken here is to describe the analytical behavior of the vector Baker-Akhiezer function on a surface with choice of sign of η from the set \mathcal{M}_g having distinct branch points in the limit as two or more branch points are brought together. The limiting object is a function on a Riemann surface of lower genus, but having as many poles as the original Baker-Akhiezer function and satisfying a set of linear homogeneous relations holding where the surface has been pinched to yield the smaller surface. The pinched surfaces with labeling are elements of the more general set \mathcal{M}'_g . This appendix provides a characterization of Baker-Akhiezer functions associated with general elements of \mathcal{M}'_g and describes their associated solutions to the Ablowitz-Ladik equations (1.3), relating them to Bäcklund transformations.

Choose an integer $0 \leq r \leq g$, and let $\tilde{g} = g - r$. Consider as before the Riemann surface of the algebraic relation

$$(A.7) \quad y^2 = \prod_{k=1}^{2\tilde{g}+2} (z - z_k),$$

where the branch points z_k are finite, distinct, and nonzero, and a specific labeling of the points 0^\pm has been chosen. Introduce the *singularity locus* \mathcal{E} , an integral divisor on \mathbb{C}^* of degree r consisting of points w_i with multiplicities m_i . \mathcal{E} is called the singularity locus because it is at the points of \mathcal{E} where we consider a genus g Riemann surface to have been pinched off to yield the genus \tilde{g} surface given by (A.7). These data make up an element Γ of the set \mathcal{M}'_g . If $r = 0$, Γ is a member of the set \mathcal{M}_g introduced in Section 2. The case of $r > 0$ corresponds to allowing r pairs of branch points to come together over the points of \mathcal{E} . In Section 2 we described the fiber of Baker-Akhiezer functions, coordinatized by the nonspecial divisor \mathcal{D} , associated to each element of \mathcal{M}_g . We now introduce the generalized fibers of Baker-Akhiezer functions associated to elements of the closure \mathcal{M}'_g . Let \mathcal{D} be a nonspecial integral divisor on this surface of degree g .

DEFINITION A.1. $\Lambda_1(\Gamma, \mathcal{D})$ ($\Lambda_2(\Gamma, \mathcal{D})$) is the linear space of all functions $f(P)$ on the Riemann surface given by the relation (A.7) with singularities of type 1 (type 2) and otherwise meromorphic, having \mathcal{D} as the divisor of the poles,

and satisfying the following linear conditions at each point w of the singularity locus \mathcal{E} :

1. If w is not a branch point of the Riemann surface, and the multiplicity of w in \mathcal{E} is m , then

$$\partial_z^l f(P^+(z)) \Big|_{z=w} = \partial_z^l f(P^-(z)) \Big|_{z=w},$$

for $l = 0, \dots, m-1$.

2. If $w = z_k$ is a branch point of the Riemann surface, and the multiplicity of w in \mathcal{E} is m , then

$$\partial_\tau^l f(P(\tau)) \Big|_{\tau=0} = 0,$$

for $l = 1, \dots, m$, where $\tau = \sqrt{z-w}$ is the local parameter at $P = w$.

In the case $r = 0$, this definition agrees with the ones given in Section 2. A function in $\Lambda_1(\Gamma, \mathcal{D})$ (resp. $\Lambda_2(\Gamma, \mathcal{D})$) for $r > 0$ can be thought of as a limit of a sequence of functions in the space $\Lambda_1(\Gamma_s, \mathcal{D})$ (resp. $\Lambda_2(\Gamma_s, \mathcal{D})$) for $s = 1, 2, 3, \dots$, where Γ_s are elements of \mathcal{M}_g (so that $r = 0$) and the corresponding Riemann surfaces are derived from algebraic relations given in terms of small parameters $\varepsilon_{i,l,s}$ that control the pinching of the sequence of surfaces. These algebraic relations are of the form

$$(A.8) \quad y_s^2 = \prod_{k=1}^{2\tilde{g}+2} (z - z_k) \prod_{i=1}^j \prod_{l=1}^{2m_i} (z - w_i + \varepsilon_{i,l,s}),$$

where the complex quantities $\varepsilon_{i,l,s}$ approach zero for large s . Segal and Wilson [45] offer an analog of \mathcal{M}_g' for the Korteweg-de Vries equation, allowing arbitrary nonzero proportionality between the left and right hand sides of the constraint imposed through \mathcal{E} in the case when w is not a branch point. However, only when this constant of proportionality is taken to be unity does the constraint correspond to a handle of the Riemann surface being pinched off exactly at the point w . Date [7] and Krichever [31] have offered analogous constructions for Lax pairs associated with other nonlinear equations.

In contrast with the generic case $r = 0$, the spaces $\Lambda_1(\Gamma, \mathcal{D})$ and $\Lambda_2(\Gamma, \mathcal{D})$ for $r > 0$ are constrained by r conditions at the points of the divisor \mathcal{E} ; however there are also r additional degrees of freedom contained in the divisor \mathcal{D} of degree $\tilde{g} + r$. In fact, the spaces are one-dimensional for all integer r such that $0 \leq r \leq g$.

LEMMA A.2. $\dim \Lambda_1(\Gamma, \mathcal{D}) = 1$ and $\dim \Lambda_2(\Gamma, \mathcal{D}) = 1$.

Proof: We sketch the proof for the space Λ_1 . The extension to the case of the space Λ_2 is exactly as it was in Section 2, through the adjoining of the differential ω . Consider the enlarged linear space Υ_1 consisting of all functions on

the Riemann surface described by (A.7) having singularities of type 1 and poles in \mathcal{D} . $\Lambda_1(\Gamma, \mathcal{D})$ is a subspace of Υ_1 obtained by imposing the linear constraints embodied in the divisor \mathcal{E} . First, we will show that the dimension of Υ_1 is $r+1$ and provide a basis for this enlarged space. Then, we will impose the r linear conditions entering through the divisor \mathcal{E} to obtain the one-dimensional subspace $\Lambda_1(\Gamma, \mathcal{D})$.

Let $\psi^{(0)}$ be a fixed nonzero element of Υ_1 having the least possible order at each point of \mathcal{D} . This means that if P is in \mathcal{D} with multiplicity m then the function $\psi^{(0)}$ has a pole of order m at the point P . Later, by providing a basis, we will demonstrate that such a function exists. Consider the linear mapping taking each element ψ of Υ_1 to a meromorphic function on the algebraic curve,

$$(A.9) \quad \xi = \psi/\psi^{(0)}.$$

Consider the range of this transformation. It is the space of all meromorphic functions on the curve having poles in the divisor of zeros of the function $\psi^{(0)}$. This divisor is nonspecial and has degree g . Thus, by the Riemann-Roch theorem applied to the Riemann surface of genus $\tilde{g} = g - r$, the range has dimension $r+1$. But the linear mapping $\psi \mapsto \xi$ is invertible, so the dimension of Υ_1 is $r+1$ (given that an appropriate element $\psi^{(0)}$ exists).

Let us now provide a basis for Υ_1 . Choose $r+1$ linearly independent nonspecial integral divisors \mathcal{D}_l of degree \tilde{g} such that $\mathcal{D}_l \leq \mathcal{D}$. The basis elements are then $\psi^{(l)}$, given by

$$(A.10) \quad \psi^{(l)}(n, t, P) = \frac{\Theta(A(P) - A(\mathcal{D}_l) - \mathcal{K} + \mathbf{U}n + \mathbf{V}t)}{\Theta(A(P) - A(\mathcal{D}_l) - \mathcal{K})} \\ \times \exp\left(n \int_{P_0}^P \omega_{(3)} + t \int_{P_0}^P \omega_{(2)}\right).$$

A generic linear combination of these $r+1$ functions will have minimal order at all points of the divisor \mathcal{D} ; such a function can serve as $\psi^{(0)}$.

Now it is time to impose the linear constraints described by the singularity locus \mathcal{E} to obtain the subspace $\Lambda_1(\Gamma, \mathcal{D})$. There are r linear equations that must be satisfied by an element of the subspace $\Lambda_1(\Gamma, \mathcal{D})$. Whenever these equations have full rank, the dimension of $\Lambda_1(\Gamma, \mathcal{D})$ will be exactly 1. In fact, the equations generically have full rank; an argument will not be given here, except to say that one may argue inductively that by Gaussian elimination it is possible to produce a triangular system with nonzero pivots. At each step of the induction, one must demonstrate the existence of a function in Υ_1 that does *not* satisfy the homogeneous linear conditions entering through \mathcal{E} . The essential condition required is that of the *regularity* of the divisor \mathcal{D} : \mathcal{D} must be linearly equivalent to $\tilde{\mathcal{D}} + q_1 + \dots + q_r$ where $\tilde{\mathcal{D}}$ is a nonspecial integral divisor of degree \tilde{g} . The complete argument for the full rank condition in the case when none of the points of \mathcal{E} are branch points appears in [17]. Also, the full rank condition has been verified in the case when \mathcal{E} consists of a single branch point for $\tilde{g} = 0$.

Normalizing an element of $\Lambda_1(\Gamma, \mathcal{D})$ at ∞^+ and an element of $\Lambda_2(\Gamma, \mathcal{D})$ at 0^+ exactly as in Section 2, we obtain the *generalized Baker-Akhiezer functions* $u_1(n, t, P)$ and $u_2(n, t, P)$. These functions reduce to those given in Section 2 when $r = 0$ and thus Γ is in the set \mathcal{M}_g . It is easy to generalize the arguments of Section 3 to show that the generalized Baker-Akhiezer functions solve the spatial linear problem (1.14) and the temporal linear problem (1.16). The only required adjustment to the arguments of Section 3 is to check that the functions ψ and ϕ employed there satisfy the linear conditions imposed through the divisor \mathcal{E} . But, since u_1 and u_2 satisfy these conditions, they extend by linearity to the functions ψ and ϕ . The remainder of the arguments go through unchanged, leading to the following lemma.

LEMMA A.3. *The generalized Baker-Akhiezer functions u_1 and u_2 satisfy the spatial linear problem (1.14) and the temporal linear problem (1.16) globally in n , t , and for all P on the algebraic curve defined by (A.7), as long as $Q(n-1, t) = d_1^{(1)}(n, t)$ and $R(n-1, t) = a_2^{(1)}(n, t)$.*

Since the consistency conditions for (1.14) and (1.16) are the Ablowitz-Ladik equations (1.3), each generalized vector Baker-Akhiezer function \mathbf{u} leads to a solution $(Q(n, t), R(n, t))$ of (1.3). We state this as a theorem.

THEOREM A.4. *Let \mathbf{u} be the generalized vector Baker-Akhiezer function constructed from an arbitrary element of \mathcal{M}'_g and the divisor \mathcal{D} . Then, the coefficients $Q(n, t) = d_1^{(1)}(n+1, t)$ and $R(n, t) = a_2^{(1)}(n+1, t)$ solve the Ablowitz-Ladik equations (1.3).*

In order to give formulas for the solution it is necessary to calculate the appropriate linear combination of the basis elements of the inflated linear spaces Υ_1 and Υ_2 to satisfy the conditions imposed through the divisor \mathcal{E} . To calculate $Q(n, t)$ and $R(n, t)$, one first finds u_1 and u_2 :

$$(A.11) \quad \begin{aligned} & u_1(n, t, P) \\ &= \left\{ \sum_{l=1}^{r+1} \alpha_l(n, t) \frac{\Theta(A(\infty^+) - \mathbf{Z}_l) \Theta(A(P) - \mathbf{Z}_l + \mathbf{U}n + \mathbf{V}t)}{\Theta(A(\infty^+) - \mathbf{Z}_l + \mathbf{U}n + \mathbf{V}t) \Theta(A(P) - \mathbf{Z}_l)} \right\} \\ & \quad \times \exp \left(-(G_2 + i)t - G_3 n + \int_{P_0}^P n \omega_{(3)} + t \omega_{(2)} \right), \end{aligned}$$

$$\begin{aligned}
& u_2(n, t, P) \\
\text{(A.12)} \quad &= \left\{ \sum_{l=1}^{r+1} \beta_l(n, t) \frac{\Theta(A(\infty^+) - \mathbf{Z}_l) \Theta(A(P) - \mathbf{Z}_l + \mathbf{U}n + \mathbf{V}t + \mathbf{W})}{\Theta(A(\infty^+) - \mathbf{Z}_l + \mathbf{U}n + \mathbf{V}t + \mathbf{W}) \Theta(A(P) - \mathbf{Z}_l)} \right\} \\
& \quad \times \exp \left(-(g_2 - i)t - g_3 n - f + \int_{P_0}^P n \omega_{(3)} + t \omega_{(2)} + \omega \right),
\end{aligned}$$

where $\mathbf{Z}_l = A(\mathcal{D}_l) - \mathcal{K}$ and $\alpha_l(n, t)$ satisfying $\alpha_1(n, t) + \dots + \alpha_{r+1}(n, t) = 1$ and $\beta_l(n, t)$ satisfying $\beta_1(n, t) + \dots + \beta_{r+1}(n, t) = 1$ are coefficients used to achieve the r linear constraints on u_1 and u_2 entering through the singularity locus \mathcal{E} . The $\alpha_l(n, t)$ and $\beta_l(n, t)$ may be directly and easily calculated from (A.11) and (A.12) in terms of theta functions by imposing the r conditions at the points w_i of \mathcal{E} , although it is not easy to give a general formula for them. The corresponding potentials $Q(n, t)$ and $R(n, t)$ are given by

$$\begin{aligned}
& Q(n-1, t) \\
\text{(A.13)} \quad &= \left\{ \sum_{l=1}^{r+1} \alpha_l(n, t) \frac{\Theta(A(\infty^+) - \mathbf{Z}_l) \Theta(A(0^+) - \mathbf{Z}_l + \mathbf{U}n + \mathbf{V}t)}{\Theta(A(\infty^+) - \mathbf{Z}_l + \mathbf{U}n + \mathbf{V}t) \Theta(A(\infty^+) - \mathbf{Z}_l)} \right\} \\
& \quad \times \exp((g_3 - G_3)n + (g_2 - G_2 - 2i)t),
\end{aligned}$$

$$\begin{aligned}
& R(n-1, t) \\
\text{(A.14)} \quad &= \left\{ \sum_{l=1}^{r+1} \beta_l(n, t) \frac{\Theta(A(\infty^+) - \mathbf{Z}_l) \Theta(A(0^+) - \mathbf{Z}_l + \mathbf{U}n + \mathbf{V}t + \mathbf{W})}{\Theta(A(\infty^+) - \mathbf{Z}_l + \mathbf{U}n + \mathbf{V}t + \mathbf{W}) \Theta(A(\infty^+) - \mathbf{Z}_l)} \right\} \\
& \quad \times \exp((F - f) - (g_3 - G_3)n - (g_2 - G_2 - 2i)t).
\end{aligned}$$

In the case of $\tilde{g} = 0$, the formulas will be simpler, since the functions in Υ_1 and Υ_2 can be written as explicit rational functions of z and y multiplied by the exponential factor. The corresponding solutions $Q(n, t)$ and $R(n, t)$ will have rational exponential expressions.

These singular solutions given by (A.13) and (A.14) can be viewed as the result of applying Bäcklund transformations to nonsingular solutions of (1.3). Bäcklund transformations are nonlinear analogs of the superposition principle in linear systems; they allow solutions of a nonlinear problem to be combined (non-linearly) to yield a genuinely new solution to the same problem [15]. One constructs such a Bäcklund transformation for the Ablowitz-Ladik equations in the following way. Choose an integral divisor \mathcal{E} of degree r on \mathbb{C}^* that will characterize the transformation. Begin with $r+1$ nonsingular solutions $(Q_l(n, t), R_l(n, t))$ to (1.3) that correspond to the same Riemann surface (A.7) and to the non-special divisors \mathcal{D}_l for $l = 1, 2, \dots, r+1$. These solutions are all merely translates

of each other on the Jacobian of the Riemann surface. For each of these potentials, there is a (nonsingular) vector Baker-Akhiezer function \mathbf{u}_l . Assemble an arbitrary linear combination (with coefficients depending on n and t) of the first components of the Baker-Akhiezer functions and then fix the coefficients (maintaining the normalization) by imposing homogeneous linear conditions on the linear combination through \mathcal{E} as we have done above. Repeat the same for the second components of the Baker-Akhiezer functions. Again expanding these linear combinations near ∞^+ and 0^+ , one obtains new potentials $(Q(n, t), R(n, t))$ satisfying (1.3). It is a pleasant feature that each nonlinear Bäcklund transformation of potentials $Q_l(n, t)$ and $R_l(n, t)$ is actually a genuine linear superposition of the corresponding vector Baker-Akhiezer functions \mathbf{u}_l .

Let us now explain the role played by these generalized solutions to (1.3) in the complete solution of the initial value problem with periodic boundary conditions in n of period N . The procedure described in Section 5 began with polynomials $f(z)$, $\varphi(z)$ and $\chi(z)$ calculated from the monodromy matrix $\mathbf{S}(0, 0, z)$. The Riemann surface was given by the algebraic relation

$$(A.15) \quad y^2 = f^2(z) - \varphi(z)\chi(z),$$

whenever the roots of the polynomial $f^2 - \varphi\chi$ were distinct. In this way, an element of \mathcal{M}_{N-1} is associated with the initial data, and the polynomials φ and f are used to define the divisor \mathcal{D} that selects a vector Baker-Akhiezer function from the associated fiber. It is now possible to remove the constraint of distinct roots, as long as at least one pair of distinct branch points remains. Write the polynomial $f^2 - \varphi\chi$ in the form

$$(A.16) \quad f^2(z) - \varphi(z)\chi(z) = \left[\prod_{i=1}^j (z - w_i)^{m_i} \right]^2 \left[\prod_{k=1}^{2\tilde{g}+2} (z - z_k) \right],$$

where $\tilde{g} + m_1 + \dots + m_j = N - 1$, and the z_k are all distinct, as are the w_i . The Riemann surface is given by the familiar algebraic relation

$$(A.17) \quad y^2 = \prod_{k=1}^{2\tilde{g}+2} (z - z_k),$$

the labeling of the Riemann surface is determined from

$$(A.18) \quad y(0^+) = \left[\prod_{i=1}^j (-w_i)^{m_i} \right]^{-1}.$$

The divisor \mathcal{E} is given by

$$(A.19) \quad \mathcal{E} = \sum_{i=1}^j m_i w_i.$$

These two objects specify an element of the set \mathcal{M}'_{N-1} . Then, the divisor \mathcal{D} of degree $N - 1$ is obtained from the polynomials φ and f exactly as follows. The z -projections of the points of \mathcal{D} are the $N - 1$ roots of φ . The y -projections of the points of \mathcal{D} are given by:

$$(A.20) \quad y(P) = \frac{f(z(P))}{\prod_{i=1}^j (z(P) - w_i)^{m_i}},$$

thus selecting a function \mathbf{u} from the associated fiber. Finally, the scaling parameter ξ is given by $Q(-1, 0)$. The corresponding solution $(\xi Q(n, t), \xi^{-1} R(n, t))$ obtained from the generalized vector Baker-Akhiezer function with the data $\Gamma \in \mathcal{M}'_{N-1}$ and \mathcal{D} then solves the initial value problem, for all finite initial data except that giving rise to an invariant polynomial having *no* distinct roots.

The case in which *all* the roots of $f^2 - \varphi\chi$ coincide in pairs requires special treatment because in the limit as the branch points coincide, the Riemann surface splits into two disjoint copies of the Riemann sphere. If the point 0^+ remains in the same connected component of the split surface as ∞^+ as the gaps between neighboring branch points close, the limiting Baker-Akhiezer function (for concreteness, consider the component u_1) is not well defined, the limit depending upon the order in which the tubes connecting the two spheres are pinched off. This is because there are singularities of the form z^n at the points 0^- and ∞^+ , each of which ends up in a different sphere when the final pair of branch points come together; thus, in the limit, n poles and n zeros must appear in some configuration where the gaps have been pinched (the situation becomes increasingly complicated as more pairs of branch points are brought together in unison, rather than one pair at a time). This configuration will depend upon which pair(s) of branch points merged last. On the other hand, if the point 0^- remains in the same connected component of the split surface as ∞^+ as the gaps close, the limiting Baker-Akhiezer function *is* well defined, and it is possible to give a unique construction of the solution in this case.

Unfortunately, it is the former case that applies to the solution of the initial value problem for periodic initial data. This is true because, from the description of the periodic problem given in Section 5, it can be seen that the case of all branch points being double corresponds to $\delta = 0$ where

$$(A.21) \quad \delta = 4 \prod_{i=k}^N (1 - Q(k, t)R(k, t)),$$

so that the function y becomes simply $\pm \text{tr} \mathbf{S}$. The “branch” of this function containing the point ∞^+ is $y = \text{tr} \mathbf{S}$, whose value at $z = 0$ is 1, so that the point connected to ∞^+ is 0^+ . It is unlikely that a Baker-Akhiezer function can be used to describe the dynamics in this case, since there is only one singular point in each sphere for the spatial dynamics, and thus, the evolution in n cannot be described by a simple differential of the third kind. To be more concrete, the two Floquet multipliers become $\rho = 0$ and $\rho = \text{tr} \mathbf{S}$ in this case; finding a third

kind differential $\omega_{(3)}$ having integer residues and satisfying

$$(A.22) \quad N\omega_{(3)} = d \log \text{tr} \mathbf{S}$$

on the Riemann sphere is impossible unless $\text{tr} \mathbf{S} = (z + \exp(2\pi ik/N))^N$ for some integer k , since these are the only degree N monic polynomials having the constant term equal to 1 where all N zeros coincide, as do all N poles.

Although our solution to the initial value problem with periodic boundary conditions cannot be strictly considered complete without treating the case $\delta = 0$, the solution for $\delta = 0$ will not be presented here, since it does not fall into the class of finite genus solutions to (1.3). But beyond the fact that the construction for $\delta = 0$ does not involve Baker-Akhiezer functions, the problem itself is a bizarre special case in which even the notion of periodic boundary conditions appears to break down. Let us explain this statement. In the case of $\delta = 0$, it can be seen from (A.21) that for some k , $Q(k, 0)R(k, 0) = 1$. As is easily verified from the equations of motion (1.3), the motion at the site k is trivial; indeed

$$(A.23) \quad Q(k, t) = Q(k, 0) \exp(-2it) \quad \text{and} \quad R(k, t) = R(k, 0) \exp(2it).$$

Furthermore, the section of the lattice to the left of site k is dynamically decoupled from the section of the lattice to the right of site k . Thus, the periodic problem becomes an initial value problem with exponentially evolving boundary conditions at the lattice points k and $k + N$, repeated periodically along the lattice. This initial value problem with boundary conditions at the endpoints given by (A.23) was treated in the defocusing case by Vekslerchik [47] who gave an algorithm for the solution of the initial value problem in terms of Toeplitz determinants. The dynamics are described by rational trigonometric functions; evidently the harmonic driving at the endpoints of each decoupled segment of the chain excites only a finite number of modes in this case.

Appendix B: Spectral symmetries of the Ablowitz-Ladik equations.

It was shown in Section 4 that in both the focusing and defocusing cases of the Ablowitz-Ladik equations, the finite genus solutions are built from Riemann surfaces that have their branch points invariant as a set with respect to reflection through the unit circle. We claimed that more was true; there were additional symmetries that applied to the focusing case and to the defocusing $|Q| < 1$ case. In this Appendix, we describe these additional symmetries. First, we deal with the focusing case.

THEOREM B.1. *In the focusing case, the branch points come in pairs reflected through the unit circle, and can only lie on the circle if the multiplicity is greater than one.*

Proof: We prove the theorem for periodic potentials $Q(n, t)$, and we will later appeal to density arguments to extend the result to general finite genus

potentials. We will employ a spectral convexity argument essentially identical to that used by Ma and Ablowitz [39] to achieve an analogous result for the periodic focusing nonlinear Schrödinger equation. It is useful to deal with the original spectral problem (1.4), based on the spectral parameter λ . In the focusing case, the spectral problem takes the form:

$$(B.24) \quad \mathbf{v}(n+1, t) = \begin{bmatrix} \lambda & Q(n) \\ -\overline{Q(n)} & \lambda^{-1} \end{bmatrix} \mathbf{v}(n, t).$$

In a periodic problem of period N , the Floquet multipliers are the eigenvalues of the monodromy matrix which is the product of the shift matrices. In the focusing case, it has the form:

$$(B.25) \quad \mathbf{S} = \begin{bmatrix} a(\lambda) & b(\lambda) \\ -\overline{b(1/\bar{\lambda})} & \overline{a(1/\bar{\lambda})} \end{bmatrix},$$

where $a(\lambda)$ and $b(\lambda)$ are Laurent polynomials in λ . The zeros in \mathbb{C}^* of the Floquet discriminant

$$(B.26) \quad \Delta(\lambda) = \text{tr}^2 \mathbf{S} - 4 \det \mathbf{S},$$

determine the branch points of the Floquet multiplier curve. Thus, the branch points (in the λ spectral variable) are the roots of

$$(B.27) \quad \left[a(\lambda) + \overline{a(1/\bar{\lambda})} \right]^2 - 4D = 0,$$

where the determinant of \mathbf{S} is

$$(B.28) \quad D = a(\lambda)\overline{a(1/\bar{\lambda})} + b(\lambda)\overline{b(1/\bar{\lambda})} = \prod_{k=1}^N (1 + |Q(k)|^2) \geq 1.$$

It is important that D is independent of λ . Define

$$(B.29) \quad a_R(\lambda) = \frac{a(\lambda) + \overline{a(1/\bar{\lambda})}}{2},$$

$$(B.30) \quad a_I(\lambda) = \frac{a(\lambda) - \overline{a(1/\bar{\lambda})}}{2i},$$

$$(B.31) \quad b_R(\lambda) = \frac{b(\lambda) + \overline{b(1/\bar{\lambda})}}{2},$$

$$(B.32) \quad b_I(\lambda) = \frac{b(\lambda) - \overline{b(1/\bar{\lambda})}}{2i}.$$

In terms of these four meromorphic functions of λ , the branch points solve the equation

$$(B.33) \quad a_R^2(\lambda) = D,$$

where the λ independent determinant has the expression

$$(B.34) \quad D = a_R^2(\lambda) + a_I^2(\lambda) + b_R^2(\lambda) + b_I^2(\lambda).$$

If $|\lambda| = 1$, then all four of the quantities a_R , a_I , b_R , and b_I are real. Thus, if λ' is a branch point on the unit circle, one has $a_I(\lambda') = b_R(\lambda') = b_I(\lambda') = 0$. The convex shape of the graph of (B.34) allows us to deduce this; it is the reason for the term ‘‘spectral convexity’’. Differentiating (B.34) with respect to λ , and evaluating at the branch point λ' on the unit circle gives

$$(B.35) \quad a_R(\lambda') \frac{d}{d\lambda} a_R(\lambda) \Big|_{\lambda=\lambda'} = 0,$$

which, since $a_R^2(\lambda') = D \geq 1$, implies

$$(B.36) \quad \frac{d}{d\lambda} a_R(\lambda) \Big|_{\lambda=\lambda'} = 0.$$

This proves that the multiplicity of any unimodular root λ' of $a_R^2(\lambda) - D = 0$ is at least two. Since the branch points z_i in this paper are just the squares of the branch points $\pm\lambda_i$ (they come in positive negative pairs, as discussed in Section 1), the branch points in a focusing periodic problem that lie on the unit circle are multiple.

Now we move on to the defocusing case. Our results are restricted to the case where the potential satisfies $|Q(n, t)|^2 < 1$ for all n , the component of the phase space that in Section 4 we have called $\Sigma_{\mathbb{Z}}$.

THEOREM B.2. *In the defocusing case, finite genus potentials in the continuum limit component $\Sigma_{\mathbb{Z}}$ of the phase space are built from Riemann surfaces that have all branch points on the unit circle.*

Proof: We will now show that, subject to periodic boundary conditions, the branch points of Riemann surfaces corresponding to defocusing potentials satisfying the constraint $|Q(n, t)| < 1$ must lie on the unit circle in the z -plane. Later, we will appeal to density arguments to extend the result to general finite genus potentials. Here, it is useful to consider the unimodular version of the Lax pair of Ablowitz and Ladik, obtained through the transformation of the Baker-Akhiezer function:

$$(B.37) \quad \mathbf{v}(n, t, \lambda) = \mathbf{w}(n, t, \lambda) \prod_{k=-\infty}^{n-1} \sqrt{1 - Q(n, t)R(n, t)}.$$

The defocusing spatial linear problem for the vector \mathbf{w} then takes the form

$$(B.38) \quad \mathbf{w}(n+1, t) = \frac{1}{\sqrt{1-|Q(n, t)|^2}} \begin{bmatrix} \lambda & Q(n) \\ \overline{Q(n)} & \lambda^{-1} \end{bmatrix} \mathbf{w}(n, t).$$

Observe that this linear problem can be written as a genuine eigenvalue problem:

$$(B.39) \quad \begin{aligned} \sqrt{1-|Q(n, t)|^2} w_1(n+1, t) - Q(n, t) w_2(n, t) &= \lambda w_1(n, t), \\ \overline{Q(n-1, t)} w_1(n, t) + \sqrt{1-|Q(n-1, t)|^2} w_2(n-1, t) &= \lambda w_2(n, t), \end{aligned}$$

which has the form

$$(B.40) \quad \Lambda \mathbf{w} = \lambda \mathbf{w},$$

where the operator Λ involves the unit shift operator Δ , and can be written as

$$(B.41) \quad \Lambda = \begin{bmatrix} \sqrt{1-|Q(n, t)|^2} \Delta & -Q(n, t) \\ \overline{Q(n-1, t)} & \sqrt{1-|Q(n-1, t)|^2} \Delta^{-1} \end{bmatrix}.$$

The operator Λ is invertible. The inverse is easily calculated:

$$(B.42) \quad \Lambda^{-1} = \begin{bmatrix} \sqrt{1-|Q(n-1, t)|^2} \Delta^{-1} & Q(n-1, t) \\ -\overline{Q(n, t)} & \sqrt{1-|Q(n, t)|^2} \Delta \end{bmatrix}.$$

Furthermore, the adjoint (for boundary boundary conditions) with respect to the inner product on \mathbf{w} given by

$$(B.43) \quad \langle \mathbf{w}, \mathbf{w}' \rangle = \sum_n \overline{\mathbf{w}}^T \mathbf{w}',$$

can be calculated as

$$(B.44) \quad \Lambda^\dagger = \begin{bmatrix} \sqrt{1-|Q(n-1, t)|^2} \Delta^{-1} & Q(n-1, t) \\ -\overline{Q(n, t)} & \sqrt{1-|Q(n, t)|^2} \Delta \end{bmatrix}.$$

Thus, periodic boundary conditions on $Q(n, t)$, and the constraint that $|Q(n, t)|$ is less than 1 for all n clearly imply that

$$(B.45) \quad \Lambda^{-1} = \Lambda^\dagger.$$

That is, the operator Λ is unitary. This is the discrete analog of the self-adjointness for the AKNS spectral problem corresponding to the defocusing nonlinear Schrödinger equation. Since Λ is unitary, its Floquet spectrum in the λ -plane lies on the unit circle. The corresponding z -values then also lie on the unit circle, since $z = \lambda^2$.

By the discussion at the end of Section 5 concerning approximating arbitrary finite genus potentials by periodic potentials, these results should carry over to the general finite genus case as well, thus verifying Theorems B.1 and B.2 in their entirety.

Bibliography

- [1] M.J. Ablowitz. Approximate methods for obtaining multi-phase modes in nonlinear dispersive wave problems. *Stud. Appl. Math.*, 51(1):17–55, 1972.
- [2] M.J. Ablowitz and D.J. Benney. The evolution of multiphase modes for nonlinear dispersive waves. *Stud. Appl. Math.*, 42(3):225–238, 1970.
- [3] M.J. Ablowitz, D.J. Kaup, A.C. Newell, and H. Segur. The inverse scattering transform – Fourier analysis for nonlinear problems. *Stud. Appl. Math.*, 53(4):249–315, 1974.
- [4] M.J. Ablowitz and J.F. Ladik. Nonlinear differential-difference equations and Fourier analysis. *J. Math. Phys.*, 17:1011–1018, 1976.
- [5] A.M. Bloch and Y. Kodama. Dispersive regularization of the Whitham equation for the Toda lattice. *SIAM J. Appl. Math.*, 52(4):909–928, 1992.
- [6] N.N. Bogolyubov, Jr. and A.K. Prikarpat-skii. The inverse periodic problem for a discrete approximation of a nonlinear Schrödinger equation. *Sov. Phys. Dokl.*, 27(2):113–116, 1982.
- [7] E. Date. Multi-soliton solutions and quasi-periodic solutions of nonlinear equations of sine-Gordon type. *Osaka J. Math.*, 19:125–158, 1982.
- [8] P. Deift, S. Venakides, and X. Zhou. The collisionless shock region for the long-time behavior of solutions of the KdV equation. *Comm. Pure Appl. Math.*, 47(2):199–206, 1994.
- [9] P. Deift and X. Zhou. A steepest descent method for oscillatory Riemann-Hilbert problems. asymptotics for the MKdV equation. *Annals of Math.*, 137(2):295–368, 1993.
- [10] P.A. Deift, A.R. Its, and X. Zhou. Long-time asymptotics for integrable nonlinear wave equations. In A.S. Fokas and V.E. Zakharov, editors, *Important Developments in Soliton Theory*, pages 181–204. Springer Verlag, Berlin, 1993.
- [11] B.A. Dubrovin. Theta functions and non-linear equations. *Uspekhi Mat. Nauk*, 36(2):11–80, 1981.
- [12] B.A. Dubrovin. *Geometry of Hamiltonian Evolutionary Systems*. Bibliopolis, Napoli, 1991.
- [13] B.A. Dubrovin, V.B. Matveev, and S.P. Novikov. Non-linear equations of Korteweg-deVries type, finite-zone linear operators, and Abelian varieties. *Russ. Math. Surveys*, 31(1):59–146, 1976.
- [14] N.M. Ercolani, M.G. Forest, and D.W. McLaughlin. The geometry of the modulational instability, part I: Local theory. to appear in *Memoirs of the AMS*.
- [15] N.M. Ercolani, M.G. Forest, and D.W. McLaughlin. The geometry of the modulational instability, part III: Homoclinic orbits for the periodic sine-Gordon equation. *Physica D*, 43:349–384, 1990.
- [16] N.M. Ercolani, S. Jin, C.D. Levermore, and W. MacEvoy. The zero dispersion limit of the NLS/mKdV hierarchy for the nonselfadjoint ZS operator. University of Arizona preprint, 1992.

- [17] N.M. Ercolani and P.D. Miller. Singular limits of multiphase solutions of the Ablowitz-Ladik equations. in preparation.
- [18] L.D. Faddeev and L.A. Takhtajan. *Hamiltonian Methods in the Theory of Solitons*. Springer Verlag, Berlin, 1987.
- [19] H.M. Farkas and I. Kra. *Riemann Surfaces*. Springer Verlag, Berlin, 2 edition, 1991.
- [20] H. Flaschka, M.G. Forest, and D.W. McLaughlin. Multiphase averaging and the inverse spectral solution of the Korteweg-de Vries equation. *Comm. Pure Appl. Math.*, 33:739–784, 1980.
- [21] M.G. Forest and J.E. Lee. Geometry and modulation theory for the periodic nonlinear Schrödinger equation. In C. Dafermos, J.L. Erickson, D. Kinderlehrer, and M. Slemrod, editors, *Oscillation Theory, Computation, and Methods of Compensated Compactness*, volume 3 of *IMA Volumes on Mathematics and its Applications*, pages 35–69. Springer Verlag, New York, 1986.
- [22] C.S. Gardner, J.M. Greene, M.D. Kruskal, and R.M. Miura. Method for solveing the Korteweg-de Vries equation. *Phys. Rev. Lett.*, 19:1095–1097, 1967.
- [23] J.K. Hale and S.M.V. Lunel. *Introduction to Functional Differential Equations*, volume 89 of *Applied Mathematical Sciences*. Springer Verlag, Berlin, 1993.
- [24] M.H. Hays. *Classical and Quantum Theories of Nonlinear Lattices*. PhD thesis, University of Arizona, 1995.
- [25] M.H. Hays, C.D. Levermore, and P.D. Miller. Macroscopic lattice dynamics. *Physica D*, 79(1):1–15, 1994.
- [26] M.H. Hays and Yixia Lu. Private Communication.
- [27] Shan Jin, C. David Levermore, and David W. McLaughlin. The behavior of solutions of the NLS equation in the semiclassical limit. In N. Ercolani, I. Gabitov, D. Levermore, and D. Serre, editors, *Singular Limits of Dispersive Waves*, pages 235–255. Plenum NATO ARW series, New York, 1994.
- [28] M. Kac and P. van Moerbeke. A complete solution of the periodic Toda problem. *Proc. Nat. Acad. Sci. USA*, 72(8):2879–2880, 1975.
- [29] I.M. Krichever. Methods of algebraic geometry in the theory of nonlinear equations. *Russian Math. Surveys*, 32(6):185–213, 1977.
- [30] I.M. Krichever. Algebraic curves and non-linear difference equations. *Russian Math. Surveys*, 33(4):255–256, 1978.
- [31] I.M. Krichever. Spectral theory of nonstationary Schrödinger operators and nonstationary Peierls model. *Funk. Anal. Pril.*, 20(3):42–51, 1986.
- [32] I.M. Krichever. Averaging method for two-dimensional integrable equations. *Funk. Anal. Pril.*, 22(3):37–52, 1988.
- [33] I.M. Krichever. Tau-function of the universal Whitham hierarchy. *Comm. Pure Appl. Math.*, 47:1–40, 1994.
- [34] P.D. Lax and C.D. Levermore. The small dispersion limit of the Korteweg-de Vries equation I, II, III. *Comm. Pure Appl. Math.*, 36:253–290, 571–593, 809–829, 1983.
- [35] P.D. Lax, C.D. Levermore, and S. Venakides. The generation and propagation of oscillations in dispersive initial value problems and their limiting behavior. In A.S. Fokas and V.E. Zakharov, editors, *Important Developments in Soliton Theory*, pages 205–241. Springer Verlag, Berlin, 1993.
- [36] C.D. Levermore. The hyperbolic nature of the zero dispersion KdV limit. *Comm. Partial Diff. Eq.*, 13(4):495–514, 1988.
- [37] C.D. Levermore and J.G. Liu. Large oscillations arising in a dispersive numerical scheme. Submitted to *Physica D*, 1995.
- [38] C.D. Levermore and J.G. Liu. Oscillations arising in numerical experiments. In N. Ercolani, I. Gabitov, D. Levermore, and D. Serre, editors, *Singular Limits of Dispersive Waves*, pages 329–346. Plenum NATO ARW series, New York, 1994.
- [39] Y.C. Ma and M.J. Ablowitz. The periodic cubic Schrödinger equation. *Stud. Appl. Math.*, 65(2):113–158, 1981.
- [40] H.P. McKean and E. Trubowitz. Hill’s operator and hyperelliptic function theory in the

- presence of infinitely many branch points. *Comm. Pure Appl. Math.*, 29(2):143–226, 1976.
- [41] P.D. Miller. *Macroscopic Lattice Dynamics*. PhD thesis, University of Arizona, 1994.
 - [42] P.D. Miller, N.M. Ercolani, and C.D. Levermore. Modulation theory in the presence of phase locking. Submitted to *Physica D*, 1994.
 - [43] E. Previato. Hyperelliptic quasi-periodic and soliton solutions of the nonlinear Schrödinger equation. *Duke Mathematical Journal*, 52(2):329–377, 1985.
 - [44] R.J. Schilling. A systematic approach to the soliton equations of a discrete eigenvalue problem. *J. Math. Phys.*, 30(7):1487–1501, 1989.
 - [45] G. Segal and G. Wilson. Loop groups and equations of KdV type. *Publications Mathématiques, Institut des Hautes Études Scientifiques*, 61:5–65, 1985.
 - [46] E.R. Tracy, H.H. Chen, and Y.C. Lee. Study of quasiperiodic solutions of the nonlinear Schrödinger equation and the nonlinear modulational instability. *Phys. Rev. Lett.*, 53(3):218–220, 1984.
 - [47] V.E. Vekslerchik. Finite nonlinear Schrödinger chain. *Phys. Lett. A*, 174(4):285–288, 1993.
 - [48] S. Venakides. Higher order Lax-Levermore theory. *Comm. Pure Appl. Math.*, 43:335–362, 1990.
 - [49] S. Venakides, P. Deift, and R. Oba. The Toda shock problem. *Comm. Pure Appl. Math.*, 44:1171–1242, 1991.
 - [50] G. Whitham. *Linear and Nonlinear Waves*. Wiley, New York, 1974.

Received (date of receipt).

Revised (date of revision).