An offprint from the

DUKE MATHEMATICAL JOURNAL

Duke University Press
Durham, NC 27708
U.S.A.

Vol. 55, No. 4 DUKE MATHEMATICAL JOURNAL © December 1987

HAMILTONIAN STRUCTURE FOR THE MODULATION EQUATIONS OF A SINE-GORDON WAVETRAIN

N. ERCOLANI, M. G. FOREST, D. W. MCLAUGHLIN, AND R. MONTGOMERY

I. Introduction. We study the sine-Gordon equation,

$$\varepsilon^2\left(\partial_t^2 - \partial_x^2\right)u + \sin u = 0, \tag{I.1}$$

where ε is a small parameter.[1] This equation is an example of a conservative, nonlinear, dispersive wave equation which enjoys the additional special property that it is integrable as an (infinite-dimensional) Hamiltonian system. It has exact solutions in the form

$$u(x,t) = W_N\left(\frac{\vec{\theta}}{\varepsilon}(x,t); \vec{\kappa}, \vec{\omega}\right). \tag{I.2}$$

These solutions (I.2) depend upon $2N$ real parameters $\vec{\kappa} = (\kappa_1, \dots, \kappa_N)$ and $\vec{\omega} = (\omega_1, \dots, \omega_N)$ and N "phases,"

$$\theta_i(x,t) = \kappa_i x + \omega_i t + \theta_i^{(0)}.$$

The x and t dependence enters the waveform (I.2) only *linearly* through these phases. For each $\vec{\kappa}$ and $\vec{\omega}$, W_N is a real function on the N-torus T^N (2π-periodic in each θ_i/ε), which has an explicit representation in terms of the Riemann theta function. Because of this 2π periodicity, the parameters $\vec{\kappa}$ and $\vec{\omega}$ are interpreted as spatial wave numbers and temporal frequencies.

Thus, (I.2) represents a (real) $3N$-dimensional family of exact solutions of the sine-Gordon equation, each member of which is quasiperiodic in both space and time. We call this family of solutions the "*N-phase, quasiperiodic waves*." When $N = 1$, this family reduces to the well-known "periodic traveling waves" for the sine-Gordon equation. Because ε is small, these traveling waves are rapidly

Received June 20, 1986. Revision received February 24, 1987. First two authors supported in part by the National Science Foundation, Grant DMS-8411002. Third author supported in part by the National Science Foundation Grant DMS-8703397 and ONR Engineering Grant N00014-85-K-0412. Last author supported in part by University of California Regents Grant and a Sloan Foundation Grant.

[1]For small ε, this problem is equivalent to the sine-Gordon equation $U_{TT} - U_{XX} + \sin U = 0$ on asymptotically long spatial and temporal scales, $X = x/\varepsilon$ and $T = t/\varepsilon$. We prefer the scaling (I.1) for fixed x and t as $\varepsilon \to 0$.

oscillating, with spatial and temporal periods $0(\varepsilon)$ on the x and t scales. For $N > 1$, solutions (I.2) should be thought of as N traveling waves *in interaction*.

As a solution of the sine-Gordon equation (I.1) evolves in time from initial data which is localized in space, it organizes (through a dispersive mechanism) into rapidly oscillating wavetrains [Whitham (1974)]. Although these wavetrains are not exactly given by N-phase, quasiperiodic solutions of (I.1), they very closely resemble members of family (I.2). The main difference between these emerging wavetrains and solutions (I.2) is that the wavetrains have physical characteristics such as wave numbers and frequencies which change slowly over large distances in space and time. If these emerging wavetrains were to be represented in the notation of (I.2), $\vec{\kappa}$ and $\vec{\omega}$ could not be constant, but would have to be slowly varying functions of x and t, $\vec{\kappa} = \vec{\kappa}(x, t)$, $\vec{\omega} = \vec{\omega}(x, t)$. Formula (I.2), with parameters which vary with x and t, is no longer an exact solution of the sine-Gordon equation (I.1).

In linear theory this situation is well understood. Localized initial data flows into nearly monochromatic wave packets. These wave packets are constructed mathematically with asymptotic methods such as geometrical optics, WKB theory, and stationary phase. These methods yield transport equations for the evolution of the amplitude and phase of the packet.

The mathematical problem is to construct an asymptotic solution of the sine-Gordon equation (I.1) in the form of a slowly varying N-phase wavetrain. This construction will produce evolution (transport) equations for the slowly varying wave numbers $\vec{\kappa}(x, t)$ and frequencies $\vec{\omega}(x, t)$. We will call these equations for $\vec{\kappa}(x, t)$ and $\vec{\omega}(x, t)$ the "*modulation equations*."

I.A. Historical background. In the single ($N = 1$) phase case, this mathematical problem is essentially solved. For a general class of nonlinear dispersive wave equations, [Whitham (1974)] derived the modulation equations for slowly varying traveling waves. He used two beautiful methods: first, he averaged conservation laws [Whitham (1965)]; later, he averaged Langrangians [Whitham (1974)]. The formal construction of a solution in the form of a slowly modulated traveling wave was given later by [Luke (1966)]. This construction was extended to the multiphase case by [Ablowitz and Benney (1970)].

[Whitham (1974)] describes the properties of the single phase modulation equations. These are first-order, nonlinear partial differential equations. When they are strictly hyperbolic, the slowly varying traveling wave is modulationally stable; long wavelength instabilities (related to "modulational" and "side-band" instabilities) arise when this strict hyperbolicity is lost. When the modulation equations are strictly hyperbolic, the distinct characteristic speeds are interpreted as the nonlinear generalization of group velocity. For the general nonlinear dispersive wave equations, most of the additional analysis of the modulation equations is restricted to the small amplitude, nearly linear regime.

However, when the underlying nonlinear wave equation is integrable (e.g., a soliton equation), its modulation equations enjoy special properties. For example, Whitham studied the modulation equations for the traveling waves of the Korteweg–de Vries (KdV) equation. Since the KdV equation is third-order in spatial derivatives, its traveling waves are a four-parameter family of solutions; in addition to κ, ω, and θ_0, the mean of the waves $\langle u \rangle$ must be included in the parameter list. The modulation equations are a strictly hyperbolic system for three unknowns—$(\kappa, \omega, \langle u \rangle)$. Even though there are three state variables, [Whitham (1974)], and later [Miura and Kruskal (1974)], explicitly found Riemann invariants which diagonalized these modulation equations for KdV. This Riemann invariant form of the KdV modulation equations was used in [Gurevich and Pitaevskii (1974), Fornberg and Whitham (1978)] to study the *conservative* smoothing of shocks.

We now understand that the Riemann invariant form of the KdV modulation equations is a direct consequence of the integrability of KdV with the inverse spectral transformation. [Flaschka, Forest, and McLaughlin (1980, 1981)] derived an "invariant form" of the modulation equations for N-phase KdV wavetrains in terms of Abelian differentials. Other equivalent forms of the modulation equations quickly follow from this invariant form. In particular, the Riemann invariant structure is immediate. In later work [Forest and McLaughlin (1983), and Forest and Lee (1986)], this invariant representation of the modulation equations was extended to the sine-Gordon equation (I.1) and the nonlinear Schrödinger equation. Clearly, these methods apply to any of the soliton equations. To date, only these soliton equations have N-phase modulation wavetrains. For these soliton equations, inverse spectral theory provides nice representations of the solution of the Cauchy initial-value problem. In the KdV case, [Lax and Levermore (1979, 1982)] used one of these representations to show that the modulation equations govern the (weak) limit of zero dispersion. [Venakides (1985)] used another of these representations to capture the oscillatory structure itself. The status of modulation equations for integrable waves is surveyed in [Ercolani, Forest, and McLaughlin (1984) and McLaughlin (1980)].

I.B. Summary of wave results. In this paper we focus on one property of the modulation equations—their structure as a Hamiltonian system. In the single-phase case [Hayes (1973)], and also [Whitham (1974)], placed the single-phase modulation equations in a Hamiltonian form which had a particularly interesting symplectic structure. However, these authors did not relate the full Hamiltonian system to a reduced one for the modulation equations; rather, they placed the reduced equations in Hamiltonian form. [Forest and McLaughlin (1979)] developed a formal and very heuristic description of the reduction process. In fact, it was with this final reduction process that they first obtained the N-phase modulation equations for the sine-Gordon equation. (The derivation presented in [Forest and McLaughlin (1979)] was done after this heuristic reduction.) In this present work we return to that early reduction and improve and complete it.

There are two main results in this paper. The first appears in section I, where the modulational Poisson structure is constructed by a natural reduction of the

Poisson structure for the original sine-Gordon system (Prop. V.3). Though we only study the sine-Gordon system here, this reduction process is rather general in that it can be applied to any Hamiltonian system having a family of N-phase solutions. The only examples known to date are soliton equations.

However, in the special case of an integrable PDE, such as sine-Gordon, the Poisson structure is very rich. For example, canonical variables are built out of conformal ingredients such as differentials on Riemann surfaces and theta functions. Analytic constructions appear at many other levels of the theory. Of course, on the one hand, this is merely a consequence of the spectral theory used to integrate these systems. On the other hand, from a strictly Hamiltonian viewpoint, these analytic objects seem alien. We feel, therefore, that it is reasonable to inquire whether these conformal constructs have a natural meaning in the purely Hamiltonian context of a phase space. Our second main result is a step in this direction. We obtain a useful spectral representation of the vector fields $\delta w_n/\delta\theta_i$ associated to the angle variables θ_i on a Liouville torus. Precisely,

$$\frac{\delta w_n}{\delta \theta_i} = \frac{i}{2\pi}\int_{a_i}\frac{d\mu}{\mu}\cdot\frac{f(\mu)}{R(\mu)},$$

where $f(\mu)/R(\mu)$ is the usual squared-eigenfunction representation of a tangent vector. The loop integral $\int_{a_i}$ is a certain continuous superposition of these eigenfunctions. This result is the key step for a long computation in section VI that derives the modulational Poisson structure in terms of conformal ingredients *directly* from the structure obtained by reduction in section V.

We conclude this introduction with a comment about the origin of the slowly varying modulations of a conservative dispersive wave. These modulations are somewhat subtle. With no "external" perturbation such as dissipation present, the wavetrain itself is an exact solution of the full equation. Why does it modulate? The answer is that dispersion causes initial data to filter into wavepackets; these wavepackets in turn share and couple excitations among neighbors in parameter space; this coupling generates modulations and dispersive spreading. Although this coupling and transport process is clear in linear wavepacket theory, it is often misunderstood in nonlinear theory. Here we emphasize that this coupling in parameter space is captured by the Hamiltonian form of the modulation equations; in particular, the coupling is the origin of the modulational symplectic structure (as discussed in section V).

II. *N*-Phase quasiperiodic waves. In this section we summarize the representation of "N-phase, quasiperiodic waves" for the sine-Gordon equation (I.1). For more details, we refer the reader to [Forest and McLaughlin (1982, 1983)], [Ercolani and Forest (1985)], and [McKean (1981)].

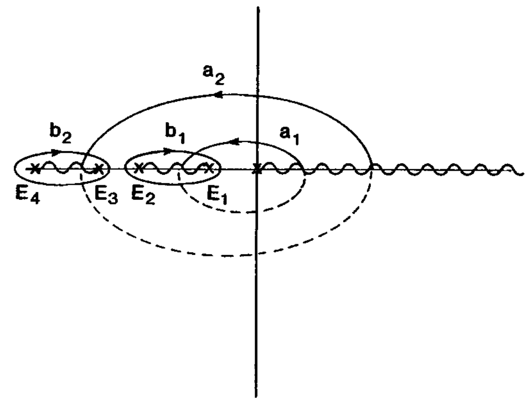

FIGURE 1

The sine-Gordon equation has exact solutions in the form

$$u(x,t) = W_N\left(\frac{\vec{\theta}(x,t)}{\varepsilon}; \vec{E}\right). \tag{II.1a}$$

Here $\vec{E} = (E_1, E_2, \dots, E_{2N}) \in \mathbf{C}^{2N}$ are parameters which are constant in x and t and which satisfy the reality constraints that E_j are either negative-real ($E_j \in (-\infty, 0)$) or occur in conjugate pairs E_j, $E_{j+1} = E_j^*$. The phases $\vec{\theta}$ and the waveform W_N are constructed from the Riemann surface $\mathscr{R} = (E, R(E))$, where

$$R^2(E) = R^2(E, \vec{E}) = E\prod_{j=1}^{2N}(E - E_j). \tag{II.1b}$$

This construction is as follows: Let (a_i, b_i), $i = 1, 2, \dots, N$ denote a particular basis of homology cycles on $\mathscr{R}$ (see Fig. 1) [see Ercolani and Forest (1984)].

On this Riemann surface we define two (unique) Abelian differentials $\Omega^{(x)}$ and $\Omega^{(t)}$:

$$\Omega^{(x)} = \frac{1}{2}\left(1 + (-1)^M\frac{\Gamma^{(x)}}{16E}\right)\prod_{k=1}^{N}\left(E - c_k^{(x)}\right)\frac{dE}{R(E)}$$

$$\Omega^{(t)} = \frac{1}{2}\left(1 + (-1)^{M+1}\frac{\Gamma^{(t)}}{16E}\right)\prod_{k=1}^{N}\left(E - c_k^{(t)}\right)\frac{dE}{R(E)},$$

where the constants $\Gamma^{(x)}, \Gamma^{(t)}$ are given by

$$\Gamma^{(x)} = \sqrt{\prod_{j=1}^{2N} E_j} \Big/ \prod_{j=1}^{N} c_j^{(x)}$$

$$\Gamma^{(t)} = \sqrt{\prod_{j=1}^{N} E_j} \Big/ \prod_{j=1}^{N} c_j^{(t)},$$

and $c_j^{(x)}, c_j^{(t)}$ are defined by the normalization conditions $\int_{a_i}\Omega^{(x)} = 0$, $\int_{a_i}\Omega^{(t)} = 0$. M is the charge of u.

In terms of these differentials, the wave numbers $\vec{\kappa} = \vec{\kappa}(\vec{E})$, the frequencies $\vec{\omega} = \vec{\omega}(\vec{E})$, and the phases $\vec{\theta}$ are given by

$$\kappa_i(\vec{E}) = \int_{b_i}\Omega^{(x)}$$

$$\omega_i(\vec{E}) = \int_{b_i}\Omega^{(t)}$$

$$\theta_i(x, t) = \kappa_i x + \omega_i t + \theta_i^{(0)}, \tag{II.1c}$$

from which we see that the $2N$ parameters $\vec{E} = (E_1, E_2, \ldots, E_{2N})$ are equivalent to the more physical parameters $\vec{\kappa}(\vec{E}) = (\kappa_1, \ldots, \kappa_N)$ and $\vec{\omega}(\vec{E}) = (\omega_1, \ldots, \omega_N)$. The only way x and t enter the waveform W_N is (linearly) through the phases θ_i. For each $\vec{E}$ the waveform W_N is a real function on the N-torus T^N, which is given explicitly in terms of Riemann theta functions (see [Forest and McLaughlin (1983)]);

$$W_N\colon T^N \to \mathbf{R} \quad \text{by } W_N = W_N(\vec{\theta}; \vec{E}). \tag{II.1d}$$

Thus, the family of real, N-phase, quasiperiodic waves (II.1a) is $3N$-real-dimensional, parameterized by $(\vec{E}, \vec{\theta}^{(0)})$, or equivalently by $(\vec{\kappa}, \vec{\omega}, \vec{\theta}^{(0)})$.

III. The modulation framework. To modulate the wavetrain, we allow $\vec{E}$, while satisfying the reality constraints, to depend upon x and t,

$$\vec{E} = \vec{E}(x, t); \tag{III.1a}$$

we construct $\vec{\kappa}$ and $\vec{\omega}$ from $\vec{E}(x, t)$ through (II.1c),

$$\kappa_i(x, t) = \kappa_i\big(\vec{E}(x, t)\big) = \int_{b_i}\Omega^{(x)}$$

$$\omega_i(x, t) = \omega_i\big(\vec{E}(x, t)\big) = \int_{b_i}\Omega^{(t)}. \tag{III.1b}$$

Then we make an ansatz for the wave u in the form

$$u^{\varepsilon}(x,t) = W_N\left(\frac{\vec{\theta}(x,t)}{\varepsilon}; \vec{E}(x,t)\right) + 0(\varepsilon), \tag{III.1c}$$

where

$$\vec{\theta}_x = \vec{\kappa}(x,t)$$
$$\vec{\theta}_t = \vec{\omega}(x,t). \tag{III.1d}$$

Clearly, for consistency of (III.1d), the wave numbers $\vec{\kappa}(x,t)$ and the frequencies $\vec{\omega}(x,t)$ must satisfy

$$\vec{\kappa}_t = \vec{\omega}_x. \tag{III.2}$$

Equation (III.2) provides N evolution equations among the $2N$ variables $\vec{E} = (E_1, E_2, \dots, E_{2N})$. We need N additional evolution equations in order to close this system. These can be deduced from the demand [McLaughlin (1981), Forest and McLaughlin (1983)] that

$$u^{(0)}(x,t) = W_N\left(\frac{\vec{\theta}(x,t)}{\varepsilon}; \vec{E}(x,t)\right)$$

satisfy the sine-Gordon equation (I.1) with an $0(\varepsilon)$ error, *valid for times* $t = 0(1)$.[2] In this manner (III.2) is closed by adding N equations, which can be placed in the form [Forest and McLaughlin (1983)]

$$\vec{J}_t = \partial_x \mathrm{grad}_{\vec{\kappa}} h, \tag{III.3}$$

where

$$h = h(\vec{\kappa}, \vec{J}) = H\left(\vec{W}_N(\cdot, \vec{E})\right)\Big|_{\vec{E}=\vec{E}(\vec{\kappa},\vec{J})} \tag{III.4}$$

and $H(\vec{u})$ is the sine-Gordon Hamiltonian

$$H(\vec{u}) = \lim_{L\to\infty} \frac{1}{2L}\int_{-L}^{L}\left[\frac{u_t^2}{2} + \frac{u_x^2}{2} + (1-\cos u)\right]dx. \tag{III.5}$$

[2] In the equivalent scaling,

$$U_{TT} - U_{XX} + \sin u = 0, \qquad W_N\left(\frac{\vec{\theta}(\varepsilon X, \varepsilon T)}{\varepsilon}; \vec{E}(\varepsilon X, \varepsilon T)\right)$$

satisfies the equation with an $0(\varepsilon)$ error, *valid for times* $T = 0(\varepsilon^{-1})$.

In these modulation equations (III.2) and (III.3), $(\vec{\kappa}, \vec{\omega}, \vec{J})$ are defined in terms of the basic parameter $\vec{E} = (E_1, \ldots, E_{2N})$ by

$$\kappa_n = \int_{b_n} \Omega^{(x)}$$

$$\omega_n = \int_{b_n} \Omega^{(t)}$$

$$J_n = \frac{i}{\pi} \int_{a_n} \ln \lambda \Omega^{(x)}(\lambda). \qquad \text{(III.6a, b, c)}$$

Equations (III.2) and (III.3) provide $2N$ equations for the $2N$ parameters $\vec{E} = \vec{E}(x, t)$. The map $\vec{E} \to (\vec{\kappa}, \vec{J})$ is invertible, so we can write $\vec{E} = \vec{E}(\vec{\kappa}, \vec{J})$, as in (III.4).

Equations (III.2) and (III.3) provide $2N$ equations for the $2N$ parameters $\vec{E} = \vec{E}(x, t)$. By manipulation [Forest and McLaughlin (1983)], these modulation equations[3] (III.2) and (III.3) can be placed in Hamiltonian form

$$\vec{\kappa}_t = \partial_x \mathrm{grad}_{\vec{J}} h$$

$$\vec{J}_t = \partial_x \mathrm{grad}_{\vec{\kappa}} h. \qquad \text{(III.7)}$$

Thus, the modulation equations for a slowly varying, N-phase wavetrain are a Hamiltonian system with canonical variables the wave numbers $\vec{\kappa}$ and the "actions" $\vec{J}$, together with a Poisson structure generated by the skew symmetric differential operator $J = I\partial/\partial x$. The wave numbers and the actions are defined in terms of $\vec{E}$ through the basic differentials $\Omega^{(x)}$ and $\Omega^{(t)}$, by (III.6). The Hamiltonian h (Equations (III.4) and (III.5)) is the sine-Gordon Hamiltonian evaluated on the N-phase quasiperiodic waves and viewed as a function of $\vec{\kappa}, \vec{J}$. The actions $\vec{J}$ and the frequencies $\vec{\omega}$ are related [Forest and McLaughlin (1983)] by

$$\vec{\omega} = \mathrm{grad}_{\vec{J}} h. \qquad \text{(III.8)}$$

(III.8) shows the equivalence of (III.2) and (III.6a).

In [Forest and McLaughlin (1983)], the Hamiltonian form (III.7) of the modulation equations was obtained by a manipulation of the invariant representation $\partial_t \Omega^{(x)} = \partial_x \Omega^{(t)}$ of the modulation equations. In the next section we deduce this Hamiltonian form of the modulation equations, together with the "modulational Poisson structures" generated by $I\partial_x$, directly from the standard

[3]Actually, the most useful form of the modulation equations is the "invariant representation" [Forest and McLaughlin (1983)] $\partial_t \Omega^{(x)} = \partial_x \Omega^{(t)}$, from which the Hamiltonian form (III.7) follows.

Hamiltonian form of the sine-Gordon equation. Although formal, this derivation represents a substantial improvement from the heuristic arguments of [Forest and McLaughlin (1979), Appendix A].

IV. Hamiltonian structure. We can write the sine-Gordon equation (I.1) as a Hamiltonian system:

$$\varepsilon \vec{u}_t = J \operatorname{grad} H^\varepsilon, \tag{IV.1}$$

where $\vec{u} = (u, v)$ represents the canonically conjugate fields, $J \equiv \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, and the Hamiltonian

$$H^\varepsilon \equiv H^\varepsilon(\vec{u}) = \int dx \left[\frac{v^2}{2} + \frac{\varepsilon^2 u_x^2}{2} + (1 - \cos u) \right]. \tag{IV.2}$$

Throughout this section we assume that the field $\vec{u}$ lives in a phase space $\mathscr{F}$ of functions of x, which die off as $x \to \infty$ sufficiently fast that integration by parts is justified.

In this framework, we consider the modulational ansatz as defining a manifold $\mathscr{M}^\varepsilon$ which is embedded in the phase space $\mathscr{F}$ in an ε-dependent way. This manifold will be coordinatized by $3N$ functions of x, $(\vec{\theta}(x), \vec{E}(x))$, with $\vec{E}$ subject to the reality constraints and with $\vec{\theta}$ subject to the constraint $\partial_x \vec{\theta} = \vec{\kappa}(\vec{E}(x))$. The modulation equations are related to a flow on $\mathscr{M}^\varepsilon$.

The modulational ansatz now takes the form

$$u^\varepsilon = W_N\left(\frac{\vec{\theta}(x)}{\varepsilon}; \vec{E}(x) \right)$$

$$v^\varepsilon = \vec{\omega}\left(\vec{E}(x)\right) \cdot \vec{\nabla} W_N, \tag{IV.3}$$

where $\vec{\nabla} \equiv \partial / \partial \vec{\theta}$ and where the $\vec{E}(x)$ are fixed (nice) functions which satisfy the reality constraints. To obtain the Hamiltonian structure of the modulation equations, we use this ε-dependent embedding (IV.3) to pull back the standard Hamiltonian structure on $\mathscr{F}$ to $\mathscr{M}^\varepsilon$, and then evaluate the result asymptotically as $\varepsilon \to 0$.

The canonical one-form $p\,dq$ on (u, v) space is written in (u, v) coordinates as $v\,du$. Its value on a variation (tangent vector) $(\delta u, \delta v)$ is

$$\int v(x) \delta u(x)\, dx. \tag{IV.4}$$

Pulling this back by our ε-dependent embedding yields the one-form $v_\varepsilon\, du_\varepsilon$ on

$\mathcal{M}^\varepsilon$. Its value on a variation $(\delta\vec{\theta}, \delta\vec{E})$ is found by noting that

$$\delta u_\varepsilon = \frac{1}{\varepsilon}\sum_{i=1}^{N}\frac{\partial W}{\partial\theta^i}\delta\theta^i + \sum_{i=1}^{2N}\frac{\partial W}{\partial E^j}\delta E^j.$$

Then

$$\int v_\varepsilon \delta u_\varepsilon = \frac{1}{\varepsilon}\int(\vec{\omega}\cdot\nabla)W_N\cdot\left(\sum_{i=1}^{N}\frac{\partial W}{\partial\theta^i}\delta\theta^i\right)dx + 0(1),$$

$$= \frac{1}{\varepsilon}\int J_i^\varepsilon(x)\delta\theta^i(x)\,dx + 0(1),$$

where

$$J_i^\varepsilon(x) = \sum_{j=1}^{N}\omega^j(\vec{E}(x))\frac{\partial W_n}{\partial\theta^j}\left(\frac{1}{\varepsilon}\vec{\theta}(x), \vec{E}(x)\right)\frac{\partial W}{\partial\theta^i}\left(\frac{1}{\varepsilon}\vec{\theta}(x), \vec{E}(x)\right).$$

Note that the 0(1) term does not contain any variations of $\vec{\theta}(x)$. We evaluate the leading order term asymptotically as $\varepsilon \to 0$ by making the *ergodic hypothesis*

$$\int J_i^\varepsilon(x)\,\delta\theta^i(x)\,dx \sim \int\langle J_i^\varepsilon(x)\rangle_{E(x)}\delta\theta^i(x)\,dx + 0(\varepsilon),$$

where the brackets "$\langle\ \rangle$" denote the phase average at constant parameter value $\vec{E} = \vec{E}(x)$:

$$\langle J_i^\varepsilon(x)\rangle_{E(x)} = \frac{1}{(2\pi)^N}\iiint_{T^N}\cdots\int\sum_{j=1}^{N}\omega^j(\vec{E})\frac{\partial W_N}{\partial\theta^j}(\vec{\theta};\vec{E})\frac{\partial W_n}{\partial\theta^i}(\vec{\theta};\vec{E})\,d^N\theta. \tag{IV.5}$$

Remarks.

1. The functions

$$J_i(x) = \langle J_i^\varepsilon(x)\rangle_{\vec{E}(x)}$$

constitute a vector $\vec{J}(x)$ called the wave-action density. $\vec{J}(x)$ seems the natural action from the point of view of canonical coordinates on $\mathcal{M}^\varepsilon$. However, its representation (IV.5) is not very useful. [Forest and McLaughlin (1979), (1983)] found a more useful expression (III.5) for the wave-action density. In sections IV–VII we show that these two expressions agree.

2. For $N = 1$, the validity of the ergodic hypothesis can be proved, provided $\partial\theta/\partial x$ is never zero (or has a finite number of finite-order zeros),

by a stationary-phase argument. For $N > 1$, one can attempt to extend the stationary-phase argument, but the condition $\partial\theta/\partial x \neq 0$ must be replaced by the infinite number of nonresonance conditions $\vec{\nu} \cdot \vec{\kappa}(x) \neq 0$, where $\vec{\nu} \in \mathbf{Z}^N - \{0\}$ and $\vec{\kappa} = \vec{\nabla}\theta$. A *K-A-M* type argument can be attempted in order to take care of these resonances, and might prove the validity of this ergodic hypothesis if we put some restrictions on $\vec{\kappa}(x)$, such as

$$\frac{\partial}{\partial x}\left(\frac{\vec{\kappa}(x)}{|\kappa(x)|}\right) \neq 0.$$

We believe the ergodic hypothesis holds for $N = 2$.

Using the wave-action density, we can summarize our calculation by

$$v_\varepsilon \, du_\varepsilon \sim \frac{1}{\varepsilon}\vec{J} \cdot d\vec{\theta} \tag{IV.6}$$

where we have dropped the term of order 1. Note that $\vec{J}$ depends on x only through $\vec{E} = \vec{E}(x)$. It is known that the map $\vec{E} \to (\vec{\kappa}(\vec{E}), \vec{J}(\vec{E}))$ is invertible,[4] so that the exact N-phase waves can also be parameterized by $\vec{\kappa}$, $\vec{J}$, and $\vec{\theta}_0$. Thus the space $\mathscr{M}^\varepsilon$ of modulated waves can be reparameterized by functions $(\vec{\theta}(x), \vec{J}(x))$. Now there *are no constraints on* $\vec{\theta}$, and the only constraints on $\vec{J}$ are those induced by the reality constraints on $\vec{E}$. Note that the reparameterization

$$\left(\vec{\theta}(x), \vec{E}(x)\right) \to \left(\vec{\theta}(x), \vec{J}(x)\right)$$

is only invertible when we constrain $\vec{\theta}$ of $(\vec{\theta}, \vec{E})$ by $\vec{\theta}_x = \vec{\kappa}(\vec{E}(x))$; otherwise, it is a many-to-one map.

Differentiating (IV.6), we obtain

$$du_\varepsilon \wedge dv_\varepsilon \sim \frac{1}{\varepsilon}\sum_{i=1}^{N} d\theta^i \wedge dJ_i. \tag{IV.7}$$

This is the highest-order term in the asymptotic expansion for the canonical two-form on $\mathscr{M}^\varepsilon$. Evaluated for a pair of variations $(\delta\vec{\theta}_1, \delta\vec{J}_1), (\delta\vec{\theta}_2, \delta\vec{J}_2)$ in $\mathscr{M}^\varepsilon$, this two-form gives the number

$$\frac{1}{\varepsilon}\int\left\{\delta\vec{\theta}_1(x) \cdot \delta\vec{J}_2(x) - \delta\vec{\theta}_2(x) \cdot \delta\vec{J}_1(x)\right\} dx + 0(1).$$

V. Canonical properties of $\mathscr{M}^\varepsilon$. In the last section we showed that the modulation manifold $\mathscr{M}^\varepsilon$, which is embedded in $\mathscr{F}$ by the ε-dependent embed-

[4] Forest and McLaughlin's expression for $\vec{J}$ must be used to see this.

ding (IV.3), inherits a canonical structure from that of $\mathscr{F}$. In particular, we found that $\mathscr{M}^{\varepsilon}$ could be coordinatized by $(\vec{\theta}(x), \vec{J}(x))$, with one- and two-forms given by

$$v_\varepsilon\, du_\varepsilon \sim \frac{1}{\varepsilon}\vec{J}\cdot d\vec{\theta}$$

$$du_\varepsilon \wedge dv_\varepsilon \sim \frac{1}{\varepsilon}\sum_{i=1}^{N} d\theta_i \wedge dJ_i. \qquad \text{(V.1a, b)}$$

In this section we analyze some consequences of this canonical structure on $\mathscr{M}^{\varepsilon}$.

Now consider the act of shifting phases on $\mathscr{M}^{\varepsilon}$:

$$\vec{\theta}(x) \to \vec{\theta}(x) + \vec{\theta}_0, \qquad \vec{\theta}_0 \in T^N$$

$$\vec{J}(x) \quad \text{fixed.} \qquad \text{(V.2)}$$

This is a torus action which leaves the one-form $\vec{J}\cdot d\vec{\theta}$ invariant, and hence is a canonical torus action on $\mathscr{M}^{\varepsilon}$, at least to leading order in ε. (If we simply drop the higher-order terms in the expansion of the two-form, or assume the ergodic hypothesis applies to these higher-order terms, then this torus action is, respectively, exactly canonical or canonical to first order.) By the standard theory of momentum maps [Abraham and Marsden (1978), Ch. 4.2] the momentum map, or Noether-conserved quantity, for this action is

$$\frac{1}{\varepsilon}\int \vec{J}(x)\, dx \in \mathbf{R}^N.$$

We can also (Poisson-) reduce by this torus action to obtain the brackets on $\vec{\kappa}(x)$, $\vec{J}(x)$ space. The (Poisson) reduction of a phase space by a canonical group action is the quotient space. In our case, the quotient space $\mathscr{M}^{\varepsilon}/T^N$ can be identified with the pairs $(\vec{\theta}_x = \vec{\kappa}, \vec{J})$; taking the x-derivative of $\vec{\theta}(x) + \vec{\theta}_0$ annihilates the constant phase shift $\vec{\theta}_0$.

Poisson brackets on the reduced phase space are computed as follows: Take two functionals F and G of $\vec{\kappa}$ and $\vec{J}$. Consider them as functions of $\vec{\theta}$ and $\vec{J}$ through the relation $\vec{\theta}_x = \vec{\kappa}$. Take their $(\vec{\theta}, \vec{J})$ Poisson brackets and then re-express this bracket in terms of $\vec{\kappa}$ and $\vec{J}$. Specifically,

$$\tilde{F}\big(\vec{\theta}(x), \vec{J}(x)\big) = F\big(\vec{\theta}_x(x), \vec{J}(x)\big)$$

or

$$\tilde{F} = \tilde{F}\circ(\partial_x, \text{identity}).$$

Then

$$\frac{\delta\tilde{F}}{\delta\vec{J}} = \frac{\delta F}{\delta\vec{J}}.$$

But

$$D_{\vec{\theta}}\tilde{F} = D_{\vec{\kappa}}F \circ \partial_x,$$

so that

$$\int \frac{\delta\tilde{F}}{\delta\vec{\theta}}\delta\vec{\theta} = \int \frac{\delta F}{\delta\vec{\kappa}}\partial_x \delta\vec{\theta}$$

$$= -\int \partial_x \frac{\delta F}{\delta\vec{\kappa}} \cdot \delta\vec{\theta}.$$

Therefore,

$$\frac{\delta\tilde{F}}{\delta\vec{\theta}} = -\partial_x \frac{\delta F}{\delta\vec{\kappa}},$$

and similarly for $\tilde{G}(\vec{\theta}(x), \vec{J}(x)) = G(\vec{\theta}_x(x), \vec{J}(x))$. Formula (V.1) says that $\vec{\theta}$ and $\vec{J}$ are canonically conjugate (with a factor of ε in front), so

$$\{F, G\}(\vec{\kappa}, \vec{J}) = \{\tilde{F}, \tilde{G}\}(\vec{\theta}, \vec{J}) \quad \text{(by definition of the reduced bracket)}$$

$$= \varepsilon \int \frac{\delta\tilde{F}}{\delta\vec{\theta}} \cdot \frac{\delta\tilde{G}}{\delta\vec{J}} - \frac{\delta\tilde{G}}{\delta\vec{\theta}} \cdot \frac{\delta\tilde{F}}{\delta\vec{J}}\, dx$$

$$= \varepsilon \int -\partial_x \frac{\delta F}{\delta\vec{\kappa}} \cdot \frac{\delta G}{\delta\vec{J}} + \partial_x \frac{\delta G}{\delta\vec{\kappa}} \cdot \frac{\delta F}{\delta\vec{J}}\, dx$$

$$= \varepsilon \int \frac{\delta F}{\delta\vec{\kappa}} \cdot \partial_x \frac{\delta G}{\delta\vec{J}} - \frac{\delta G}{\delta\vec{\kappa}} \cdot \partial_x \frac{\partial F}{\delta\vec{J}}\, dx.$$

This Poisson bracket was found earlier by [Forest and McLaughlin (1979)]. From the general theory of reduction, the momentum map $1/\varepsilon \int \vec{J}(x)\, dx$ should be a Casimir for this bracket. This can be seen by hand: Set

$$I^i = \frac{1}{\varepsilon}\int J^i(x)\, dx.$$

Then

$$\frac{\delta I^i}{\delta \vec{\kappa}} = 0, \qquad \frac{\delta I^i}{\delta J^j} = \frac{1}{\varepsilon}\delta^i_j.$$

So

$$\{F, I^i\} = -\int \partial_x \frac{\delta F}{\delta \kappa^i}\, dx = 0$$

(assuming fields die off sufficiently at infinity).

Finally, for all of this to be of any use, the Hamiltonian must be phase-shift-invariant so that it can be considered a function of $\vec{\kappa}$ and $\vec{J}$ and not $\vec{\theta}_0$. This will not be exactly true, but it will be to highest order if we make the ergodic assumption for H. We write this result as

$$H(\vec{\kappa}, \vec{J}) = \int \langle h \rangle (\vec{\kappa}(x), \vec{J}(x))\, dx.$$

We summarize the discussion of sections IV and V in a

PROPOSITION. *Assume that the ergodic hypothesis holds for the canonical one-form* $v_\varepsilon\, du_\varepsilon$ *and for the Hamiltonian* H *on* $\mathscr{M}^\varepsilon$. *Then*

(A) $$v_\varepsilon\, du_\varepsilon = \frac{1}{\varepsilon} \sum_{i=1}^{N} J_i\, d\theta^i + 0(1), \tag{V.3a}$$

where J_i *is the wave action given by phase averaging* (IV.5).

Drop the 0(1) *term in* $v_\varepsilon\, du_\varepsilon$. *Then the torus action* (V.2) *is a canonical action on* $\mathscr{M}^\varepsilon$. *The corresponding Poisson-reduced space* $\mathscr{M}^\varepsilon / T^N$ *is identifiable with pairs of functions* $(\vec{\kappa}(x), \vec{J}(x))$. *The Poisson bracket on this phase space is given by*

$$\{F, G\}(\kappa, J) = \varepsilon \int \left(\frac{\delta F}{\delta \vec{\kappa}} \cdot \partial_x \frac{\delta G}{\delta \vec{J}} - \frac{\delta G}{\delta \vec{\kappa}} \cdot \partial_x \frac{\delta F}{\delta \vec{J}} \right) dx. \tag{V.3b}$$

The N *functionals*

$$\vec{I} = \frac{1}{\varepsilon} \int \vec{J}(x)\, dx \tag{V.3c}$$

compose the momentum map for the torus action and are also the Casimirs for the

Poisson bracket. The modulation equations can be written in the Hamiltonian form

$$F_t = \{F, H\}$$

or

$$\vec{\kappa}_t = \varepsilon \partial_x \frac{\delta H}{\delta \vec{J}}$$

$$\vec{J}_t = \varepsilon \partial_x \frac{\delta H}{\delta \vec{\kappa}}. \tag{V.3d}$$

We close this section with a comment about the nature of the modulational Poisson structure as generated by the skew-symmetric operator $I\partial_x$. We focus on the Hamiltonian form of the modulation equations (V.3d). While the Hamiltonian H does not depend on $\vec{\theta}$, it does depend upon $\vec{\kappa}$; hence, it depends upon $\vec{\theta}_x$. This dependence couples a wavetrain located at (x) with its neighbor located at $(x + dx)$. This coupling induces the modulations through (V.3d), and eventually leads to phenomena such as "dispersive spreading."

VI. Spectral representation of Grad$_{\theta}W_N$. In section IV we derived the action $\langle \vec{J} \rangle$, Equation (IV.5),

$$\langle J_j \rangle \equiv \frac{1}{(2\pi)^N} \int_{T^N} \left\{ \left[\left(\vec{\omega}(\vec{E}) \cdot \vec{\nabla}_\theta \right) W_N(\vec{\theta}; \vec{E}) \right] \frac{\partial}{\partial \theta_j} W_N(\vec{\theta}; \vec{E}) \right\} d^N\theta, \tag{VI.1a}$$

by pulling back the canonical structure onto the modulation manifold $\mathscr{M}^\varepsilon$. Although its derivation is canonical, Formula (IV.5) for the action $\langle \vec{J} \rangle$ is far less explicit and less useful than the representation of the action $\vec{J}$ in terms of Abelian integrals, Equation (III.6c),

$$J_j = -\frac{i}{\pi} \int_{a_j} \ln \lambda \Omega^{(x)}(\lambda). \tag{VI.1b}$$

In the next two sections we use the theory of the spectral transform to express the action $\langle \vec{J} \rangle$, as defined by (VI.1a), in terms of spectral variables; thus, we show that $\langle \vec{J} \rangle$ has a compact expression in terms of Abelian integrals. In fact, up to sign, $\langle \vec{J}_i \rangle = \vec{J}_i$, which we show in section VII. First, in this section we derive a new spectral formula for $\partial/\partial\theta_i W_N$.

VI.A. The inverse spectral representation. The space of N-phase waves W_N (§1), on which the modulation theory is built, can be coordinatized by the data of a spectral problem. This spectral problem [Ablowitz, Kaup and Newell (1974), Fadeev and Takhtajian (1974)] has the form

$$\mathscr{L}(\vec{u}, E)\vec{\psi} = 0,$$

where $\vec{u}$ is the N-phase wave, E is the eigenvalue parameter, $\vec{\psi}(x, E) = (\psi_1, \psi_2)^T \in \mathbf{C}^2$, and $\mathscr{L}$ is a 2×2-matrix differential operator in x. It is a fact [Forest and McLaughlin (1983), Daté (1980)] that $\vec{u}$ is N-phase if and only if there exist $\vec{\psi}(E), \vec{\phi}(E)$ in the E-eigenspace of $\mathscr{L}(\vec{u}, E)$ for all E such that

$$g(E) = \phi_1\psi_1, \qquad h(E) = -\phi_2\psi_2$$

are polynomials of degree N in E. Thus we have, possibly after a normalization,

$$g(E) = -\prod_{l=1}^{N}(E - \mu_l) = -\left(E^N + G_1E^{N-1} + \cdots + G_N\right)$$

$$h(E) = \prod_{l=1}^{N}(E - \nu_l) = \left(E^N + H_1E^{N-1} + \cdots + H_N\right). \tag{VI.2}$$

The coefficients $\{G_i, H_i\}$ depend on x and t through the wave $\vec{u}$. They are equivalent to $\{\mu_l, \nu_l\}$, the former being symmetric polynomials of the latter. It is also true that

$$f(E) = \frac{i}{2}(\phi_1\psi_2 + \psi_1\phi_2) = \sqrt{E}\left(F_1E^{N-1} + \cdots + F_N\right), \tag{VI.3}$$

and

$$f^2 - gh = \prod_{k=1}^{2N}(E - E_k) = s^2(E) \tag{VI.4a}$$

is a polynomial whose roots $\{E_1, \ldots, E_{2N}\}$ are invariant under the sine-Gordon flows.

These eigenvalues $\{E_1, \ldots, E_{2N}\}$ were already mentioned (§1) in our parameterization of N-phase waves. The main point of the spectral coordinates is to replace $(\theta_1, \ldots, \theta_N)$ in our earlier parameterization by $(\mu_1, \ldots, \mu_N)$.

This might appear to be a needless complication of affairs. On the contrary, the spectral coordinates are fundamental to the integrable structure of sine-Gordon and even to the definition of the angles $\vec{\theta}$.

These $\vec{\mu}$-variables have analogues in the classical work of Jacobi on integrability; this classical viewpoint has been revived by many people recently, with a special emphasis on its connections to algebraic geometry. It is these connections which enable one to carry out computations. We will briefly review them below.

The remainder of this section proceeds in three parts:
VI.B Changing coordinates to spectral variables
VI.C Translation invariant vector fields
VI.D A fundamental formula.

VI.B. Spectral coordinates. Because $u = W_N$ is real and quasiperiodic, it follows that the roots of $f^2 - gh$, $\{E_1, \dots, E_{2N}\}$, are all distinct and satisfy the "reality constraints": either E_i is real and negative or, if $E_i \notin \mathbf{R}^{(-)}$, then E_i is nonreal and $E_i^* = E_j$ for some $j \neq i$. Since the E_i are invariant under the sine-Gordon flows, the level sets $\{E_i = \text{const}\}_{i=1}^{2N}$ in function space are invariant sets for these flows.

Comparing coefficients in (IV.3),

$$f^2 - gh = s^2,$$

we desire $2N$ polynomial equations for the $3N$ complex variables $(G_1, \dots, G_N, H_1, \dots, H_N, F_1, \dots, F_N)$ which define the N-dimensional level sets of constant E_i, denoted $\mathcal{M}_{\vec{E}}$. When $\vec{u}$ is real-valued the variables satisfy the constraints

$$G_i = H_i^*, \qquad F_i \text{ real.}$$

We will call such points in C^{3N} "real points" and will use them at several times. However, to describe our constructions it is easier if we allow these variables to be complex. The level set $\mathcal{M}_{\vec{E}}$ is locally coordinatized by

$$\{(\mu_1, s(\mu_1)), \dots, (\mu_N, s(\mu_N))\}: \qquad \text{(VI.4b)}$$

$\{(\mu_l, s(\mu_l))\}$ determine g and f; h is then determined from $f^2 - gh = s^2$. (For a more detailed discussion of this see [Mumford (1983)].)

Replace $s(\mu)$ by

$$R(\mu) = \sqrt{\mu}\, s(\mu).$$

Let $\mathcal{R}$ be the Riemann surface of $R(\mu)$. The level set $\mathcal{M}_{\vec{E}}$ is isomorphic to

$$(\mathcal{R} - \{\infty\})^{(N)} \qquad \text{(VI.5)}$$

(($)^{(N)}$ denotes the N-fold symmetric product), which is locally coordinatized by (VI.4). To pass from the coordinates (VI.4) to the complex angles θ_i, we introduce the Abel map:

$$\Phi: \mathcal{R}^{(N)} \to \operatorname{Jac}(\mathcal{R}) = \mathbf{C}^N/\Lambda:$$

$$(P_l = (\mu_l, R(\mu_l))) \qquad P_1 + \cdots + P_N \to \begin{bmatrix} \sum_{i=1}^{N} \int_\infty^{P_i} v_1 \\ \ddots \\ \sum_{i=1}^{N} \int_\infty^{P_i} v_N \end{bmatrix}.$$

Definition of terms and remarks:

(i) A canonical basis $a_1, \ldots, a_N, b_1, \ldots, b_N$ is chosen for the homology group of one-cycles on $\mathscr{R}$, $a_1, \ldots, a_N$ are uniquely determined, up to ordering and orientation, by the condition that any closed loop of real points of $\mathscr{M}_{\vec{E}}$ decomposes in μ_l-coordinates into a sum of N loops which are homologous to an integral combination of a_i. *Canonical* means that under the intersection pairing, $(a_i, a_j) = (b_i, b_j) = 0$, $(a_i, b_j) = \delta_{ij}$. (See [Ercolani and Forest (1985)] for more details.)

(ii) $v_1, \ldots, v_N$ is the basis of the space of holomorphic differentials on normalized, so that

$$\int_{a_i} v_j = \delta_{ij}.$$

(iii) Λ is the lattice of periods of Φ. It is generated by integer sums of $(\int_{a_i}\vec{v}, \int_{b_i}\vec{v})$.

(iv) ϕ is a biholomorphism between $\mathscr{R}^{(N)}$ and Jac($\mathscr{R}$). It takes $(\mathscr{R} - \infty)^{(N)}$ onto Jac($\mathscr{R}$) $- \{\phi\}$.

(v) The real angles θ_i parameterize a translation of the real span of the a_i periods:

$$\{(0, \ldots, \theta_i/2\pi, \ldots, 0) : \theta_i \bmod 2\pi\} \simeq \left(\vec{\Lambda}_0 + \text{span of } \mathbf{R} \cdot \int_{a_i} \vec{v}\right) \Big/ \Lambda,$$

where $\vec{\Lambda}_0$ is a constant translation. This translate, denoted by T, is a real N-torus lying entirely in Jac $\mathbf{R} - \{0\}$. Under ϕ^{-1} this torus is isomorphic to the real points of $\mathscr{M}_{\vec{E}}$ [Ercolani and Forest (1985)].

VI.C. Translation invariant vector fields. Our goal in section VI is to find a useful spectral representation of $\partial W_n/\partial\theta_i$; hence, we must study the vector fields $\partial/\partial\theta_i$. Clearly, $\{\partial W_N/\partial\theta_1, \ldots, \partial W_N/\partial\theta_N\}$ is a complex basis for the translation-invariant vector fields. We use the Abel map ϕ to pull these back to $(\mathscr{R} - \{\infty\})^{(N)}$, the spectral coordinates, i.e., we identify

$$(\Phi)^* \frac{\partial}{\partial\theta_i}.$$

We get at this in a somewhat roundabout fashion by constructing a family of invariant vector fields parameterized by $\mathscr{R} - \{\infty\}$.[5]

[5] For those familiar with spectral theory in the periodic case, these vector fields are the analog of the gradients

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \frac{\delta \cos^{-1} \frac{\Delta}{2}(E)}{\delta \vec{u}},$$

where Δ is the Floquet discriminant.

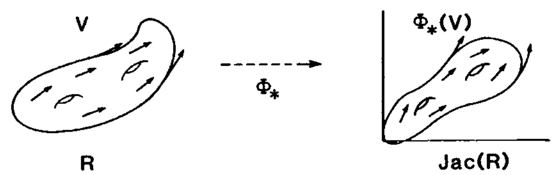

FIGURE 2

Let V be a nonvanishing holomorphic vector field on $\mathscr{R} - \{\infty\}$. Consider the restricted Abel map

$$\Phi(P + (N-1)\infty) = \begin{bmatrix} \int_\infty^P v_1 \\ \vdots \\ \int_\infty^P v_N \end{bmatrix}.$$

The closure of the image of Φ is a copy of $\mathscr{R}$ embedded in Jac($\mathscr{R}$) with $\Phi(\infty) = \vec{0}$ [Ercolani and Forest (1985)]. Φ_* embeds the vector field V as a vector field along $\Phi(\mathscr{R} - \{\infty\}) \subseteq \text{Jac}(\mathscr{R})$. See Figure 2.

Now $\text{Jac}(\mathscr{R}) = C^N/\Lambda$ is a flat torus, so we can extend the vector $\phi_* V(p)$ at $\Phi(p)$ to a translation invariant vector field on all of Jac($\mathscr{R}$). Let $D_p(V)$ denote this extension in terms of the translation-invariant basis $\{\partial/\partial\theta_i\}$; we have the formula [Mumford (1983)]

$$D_p(V) = -2\pi \sum_{i=1}^{N} \langle v_i(p), V(p)\rangle \frac{\partial}{\partial\theta^i},$$

where the v_i are the basis of holomorphic differentials on $\mathscr{R}$ and where the pairing $\langle , \rangle$ is the natural pairing between covectors and vectors. A point-independent way of representing this construction is

$$D = -2\pi \sum_{i=1}^{N} v_i \otimes \frac{\partial}{\partial\theta^i}. \tag{VI.6}$$

From a differential geometric point of view, D is a (holomorphic) one-form on $\mathscr{R}$ with values in the space of invariant vector fields on Jac($\mathscr{R}$).

The pullback of D by the *full* Abel map,

$$\phi^* D = -2\pi \sum_i v_i \otimes \phi^* \frac{\partial}{\partial\theta^i},$$

can also be computed.

The result [Ercolani (1986), Mumford (1983)] is best given in a point-independent manner by evaluating this derivation, which by abuse of notation we will also call D, on the coordinate polynomials g, h, and f on $\mathscr{R}^{(N)}$. For our calculations we only need the result for g/s:

$$D\left(\frac{g(E)}{s(E)}\right) = \frac{d\mu}{s(\mu)} \cdot \frac{g(E)f(\mu) - f(E)g(\mu)}{(E-\mu)s(E)}, \tag{VI.7}$$

where $(\mu, \sqrt{\mu}\, s(\mu))$ is the point P on $\mathscr{R}$ and where E is the parameter of the functions f, g, and h. A remark on interpreting this formula may be helpful. Fix $E \in \mathscr{R}$. Then $f(E)$, $g(E)$, $h(E)$, and $s(E)$ are functions of $(G_1, \dots, G_N, H_1, \dots, H_N, F_1, \dots, F_N)$, and hence functions on $\mathscr{R}^{(N)}$. For example, $g(0)/s(0)$ is one such function. Now D has the form $\Sigma v_i \otimes X^i$, where the v_i are holomorphic differentials on $\mathscr{R}$ and the X^i are vector fields on $\mathscr{R}^{(N)}$. Hence

$$D\left(\frac{g(E)}{s(E)}\right) = \Sigma v_i X^i \left[\frac{g(E)}{s(E)}\right]$$

should be a one-form on $\mathscr{R}$, as is the right-hand side of (VI.7).

VI.D. A fundamental formula. We will now use (III.6) and (III.7) to derive a formula which is fundamental for our final computation.

At $E = 0$,

$$\frac{g(0)}{s(0)} = \frac{-G_N}{\sqrt{\Pi E_i}} = \frac{-\Pi \mu_l}{\sqrt{\Pi E_i}} = -e^{-iu}, \tag{VI.8}$$

where $u = W_N$ is an N-phase wave on $\mathscr{M}_{\vec{E}}$. The final equality here is derived by [Forest and McLaughlin (1983)]. Then

$$D\left(\frac{g(0)}{s(0)}\right) = i\,Du \cdot e^{-iu}.$$

Also,

$$D\left(\frac{g(0)}{s(0)}\right) = \frac{g(0)f(\mu)\,d\mu}{s(0)(-\mu)s(\mu)} \qquad \text{(by (VI.7))}$$

$$= \frac{e^{-iu}f(\mu)}{\mu s(\mu)}\,d\mu.$$

Therefore,

$$Du = -i\frac{f(\mu)}{s(\mu)} \cdot \frac{d\mu}{\mu}. \tag{VI.9}$$

Taking a continuous superposition of D_P along a_j, we have

$$\int_{a_j} Du = -2\pi \int_{a_j} \sum_{i=1}^{N} u_{\theta_i} v_i \qquad \text{(by (VI.6))}$$

$$= -2\pi u_{\theta_j}. \qquad \text{(by normalization of } v_i\text{)}$$

On the other hand,

$$\int_{a_j} Du = -i \int_{a_j} \frac{f(\mu)}{s(\mu)} \cdot \frac{d\mu}{\mu}. \qquad \text{(by (VI.9))}$$

Therefore, we have a fundamental formula

$$\boxed{\frac{\partial W_N}{\partial \theta_j} = \frac{i}{2\pi} \int_{a_j} \frac{d\mu}{\mu} \cdot \frac{f(\mu)}{\sqrt{\prod_{k=1}^{2N} (\mu - E_k)}}}, \qquad \text{(VI.10)}$$

which is the main result of section VI. Note that the denominator in (VI.10) may be expressed as $\sqrt{\mu}\, R(\mu)$.

VII. A spectral representation of the action $\langle J \rangle$. Armed with the fundamental formula (VI.10), we can now derive a spectral representation of $\langle J \rangle$. We begin from (VI.1a) ($u = W_N$ below),

$$\langle J_i \rangle \equiv \langle u_t, u_{\theta_i} \rangle = \langle \vec{\omega} \cdot \vec{\nabla}_\theta u, u_{\theta_i} \rangle,$$

and continue to compute:

(i)
$$\langle J_i \rangle = \langle \vec{\omega} \cdot \vec{\nabla} u, u_{\theta_i} \rangle$$

$$= \frac{i}{2\pi} \int_{a_i} \frac{d\mu}{\sqrt{\mu}\, R(\mu)} \vec{\omega} \cdot \langle \vec{\nabla} u, f(\mu) \rangle \qquad \text{(by (VI.10))}$$

$$= \frac{-i}{2\pi} \int_{a_i} \frac{d\mu}{\sqrt{\mu}\, R(\mu)} \vec{\omega} \cdot \langle u, \vec{\nabla} f(\mu) \rangle \quad \text{(integration by parts)}$$

$$= \frac{-i}{2\pi} \int_{a_i} \frac{d\mu}{\sqrt{\mu}\, R(\mu)} \langle u, f_t(\mu)) \rangle. \qquad \text{(VII.1)}$$

(ii) In [Forest and McLaughlin (1983)], Formula (II.4′), we find that

$$f_t = i\sqrt{\mu}\left[\left(1 + \frac{e^{iu}}{16\mu}\right)g + \left(1 + \frac{e^{-iu}}{16\mu}\right)h\right].$$

Inserting this into (VII.1), we have

$$\langle J_i\rangle = \frac{1}{2\pi}\left[\int_{a_i}\frac{d\mu}{R(\mu)}\left\langle\left(1 + \frac{e^{iu}}{16\mu}\right)gu\right\rangle + \int_{a_i}\frac{d\mu}{R(\mu)}\left\langle\left(1 + \frac{e^{-iu}}{16\mu}\right)hu\right\rangle\right]. \quad \text{(VII.2)}$$

(iii) It follows from the reality of u that

$$(e^{iu})^* = e^{-iu}$$

(( )* denotes complex conjugate) and

$$(h(\mu))^* = g(\mu^*).$$

$(1 + e^{-iu}/16\mu)h(\mu)u$ is of the form $(1/\mu)F(\mu)$, where $F(\mu)$ is a polynomial in μ. The coefficients of F are functions of $\{(\mu_1, R(\mu_1)), \ldots, (\mu_N, (R(\mu_N))\}$. $(F(\mu^*))^*$ replaces the coefficients of F by their conjugates. Since these coefficients are just symmetric polynomials in $\{\mu_1, \ldots, \mu_N\}$, this conjugation amounts to replacing $\{\mu_i\}$ by $\{\mu_i^*\}$. However, if $(\mu_1, \ldots, \mu_N)$ are coordinates of a point on T, $(\mu_1^*, \ldots, \mu_N^*)$ are coordinates of another point on T. This follows from the structure of the reality constraints. Also, this conjugation preserves volumes on T. Thus, the average over all of T satisfies

$$\left\langle\frac{1}{\mu}F(\mu)\right\rangle = \frac{1}{\mu}\langle F(\mu)\rangle = \frac{1}{\mu}\langle(F(\mu^*))^*\rangle = \left\langle\frac{1}{\mu}(F(\mu^*))^*\right\rangle.$$

Hence,

$$\left\langle\left(1 + \frac{e^{-iu}}{16\mu}\right)h(\mu)u\right\rangle = \left\langle\left(1 + \frac{e^{iu}}{16\mu}\right)(h(\mu^*))^*u\right\rangle$$

$$= \left\langle\left(1 + \frac{e^{iu}}{16\mu}\right)g(\mu)u\right\rangle.$$

Therefore, (VII.2) reduces to

$$\langle J_i \rangle = \frac{1}{\pi} \int_{a_i} \frac{d\mu}{R(\mu)} \left\langle \left(1 + \frac{e^{iu}}{16\mu}\right) gu \right\rangle. \tag{VII.3}$$

(iv) We now rewrite the phase average in (VII.3) entirely in terms of the spectral variables $\{\mu_i\}$. This is done by multiphase averaging as described in [Forest and McLaughlin (1983)]:

$$\left\langle \left(1 + \frac{e^{iu}}{16\mu}\right) g(\mu) u \right\rangle$$

$$= \int_{a_1} \cdots \int_{a_N} \left(1 + \frac{1}{16} \frac{\sqrt{\Pi E_k}}{\mu\mu_1 \cdots \mu_N}\right) \prod_{i=1}^{N} (\mu - \mu_i) \cdot$$

$$\left[\left(\sum_{j=1}^{N} i(\ln \mu_j)\right) - i \ln \sqrt{\Pi E_k}\right] \frac{\prod_{i<j} (\mu_i - \mu_j)}{\prod_{l=1}^{N} R(\mu_l)} d\mu_1 \cdots d\mu_N$$

and

$$\frac{\left\langle \left(1 + \frac{e^{iu}}{16\mu}\right) gu \right\rangle d\mu}{R(\mu)} = -i \ln \sqrt{\Pi E_k}\, \Omega^{(x)}$$

$$+ \left[\int_{a_i} \cdots \int_{a_N} \left(1 + \frac{1}{16} \frac{\sqrt{\Pi E_k}}{\mu\mu_1 \cdots \mu_N}\right) \prod_{i=1}^{N} (\mu - \mu_i) \right.$$

$$\left. \cdot \left(\sum_{j=1}^{N} i(\ln \mu_j)\right) \frac{\prod_{1 \leqslant i<j} (\mu_i - \mu_j)}{\prod_{l=1}^{N} R(\mu_l)} d\mu_1 \cdots d\mu_N \right] \frac{d\mu}{R(\mu)}.$$

The first term is established by section III.D of [Forest and McLaughlin (1983)].

Since, $\int_{a_i}\Omega^{(x)} = 0$, this term contributes nothing to (VII.3), and so

$$\langle J_p\rangle = \frac{1}{\pi}\int_{a_p}\int_{a_1}\cdots\int_{a_N}\left(1 + \frac{1}{16}\frac{\sqrt{\prod E_k}}{\mu_0\cdots\mu_N}\right)\prod_{i=1}^{N}(\mu_0-\mu_i)$$

$$\times\left(\sum_{j=1}^{N} i(\ln\mu_j)\right)\frac{\prod_{1<i<j}(\mu_i-\mu_j)}{\prod_{l=0}^{N}R(\mu_l)}\,d\mu_0\cdots d\mu_N, \tag{VII.4}$$

where we have replaced μ by μ_0, the variable of integration for the outer integral $\int_{a_p}$.

$$\langle J_p\rangle = \frac{i}{\pi}\sum_{j=1}^{N}\int_{a_p}\int_{a_1}\cdots\int_{a_N}\ln(\mu_j)\left(1+\frac{1}{16}\frac{\sqrt{\prod E_i}}{\mu_0\cdots\mu_N}\right)$$

$$\times\frac{\prod_{0<i<j<N}(\mu_i-\mu_j)}{\prod_{l=0}^{N}R(\mu_l)}\,d\mu_0\cdots d\mu_N. \tag{VII.5}$$

Formula (VII.5) evinces the symmetry of the μ_0 variable with the other μ_i (except in $\ln(\mu_j)$).

(v) In this part we will show that if $j \neq p$, the jth term in (VII.5) vanishes. The jth term is

$$\frac{i}{\pi}\int_{a_p}\int_{a_1}\cdots\int_{a_N}\ln(\mu_j)\left(1+\frac{1}{16}\frac{\sqrt{\prod E_i}}{\mu_0\cdots\mu_N}\right)\frac{\prod_{0\leqslant k<m<N}(\mu_k-\mu_m)}{\prod_{l=0}^{N}R(\mu_l)}\,d\mu_0\cdots d\mu_N.$$

By Fubini's theorem one can interchange the 0th and the jth place in the order of integration:

$$j\text{th term} = \frac{1}{\pi}\int_{a_j}\int_{a_1}\cdots\int_{a_{j-1}}\int_{a_p}\int\cdots\int_{a_N}\ln(\mu_j)\left(1+\frac{1}{16}\frac{\sqrt{\prod_k E_k}}{\mu_0\cdots\mu_N}\right)$$

$$\times\left\{\frac{\prod_{0<k<m<N}(\mu_k-\mu_m)}{\prod_{l=0}^{N}R(\mu_l)}\right\}d\mu_j\,d\mu_1\cdots d\mu_{j-1}\,d\mu_p\,d\mu_{j+1}\cdots d\mu_N.$$

The term in the brackets can be rewritten

$$\{\ \} = (-1)^j \frac{\prod_k'(\mu_j - \mu_k)\prod_{k<m}'(\mu_k - \mu_m)}{\prod_l' R(\mu_l)} \times \frac{1}{R(\mu_j)},$$

where the prime in the product symbol $\prod'$ means multiply over all indices *not* equal to j. The sign $(-1)^j$ comes out, since j is the number of indices k with $k < j$. Thus

$$j\text{th term} = \frac{(-1)^j i}{\pi} \int_{a_j} \ln(\mu_j)[\ ]_j \frac{d\mu_j}{R(\mu_j)} \qquad \text{(VII.6a)}_j$$

where

$$[\ ]_j = \int_{a_1} \cdots \int_{a_{j-1}} \int_{a_p} \int_{a_j} \cdots \int_{a_N} \left(1 + \frac{1}{16} \frac{\sqrt{\prod_k E_k}}{\prod_k \mu_k}\right)$$

$$\times \prod'(\mu_j - \mu_k) \frac{\prod_{k<m}'(\mu_k - \mu_m)}{\prod_k' R(\mu_k)} d\mu_1 \cdots d\mu_0 \cdots d\mu_N. \qquad \text{(VII.6b)}_j$$

If $p \neq j$, then the loop integral $\int_{a_p}$ is repeated *twice* in $[\ ]_j$. We apply Fubini to the integral $[\ ]_j$ and interchange the pth and the jth (now occupied by μ_0) variables of integration:

$$[\ ]_j = \int_{a_1} \cdots \int_{a_i} \int_{a_{i+1}} \cdots \int_{a_{j-1}} \int_{a_p} \int_{a_{j+1}} \cdots \int_{a_N} \{\ \}$$

$$= \int_{a_1} \cdots \int_{a_p} \cdots \int_{a_p} \cdots \int_{a_N} \left(1 + \frac{\sqrt{\prod_k E_k}}{16 \prod_k \mu_k}\right) \frac{1}{\prod_k' R(\mu_k)}$$

$$+ \sigma_{p0}\left[\prod_k'(\mu_j - \mu_k)\prod_{k<m}'(\mu_k - \mu_m)\right] d\mu_1 \cdots d\mu_0 \cdots d\mu_N, \qquad \text{(VII.7)}_j$$

where σ_{p0} just permutes μ_p and μ_0 in the product $[\prod_k'(\mu_j - \mu_k)\prod_{k<m}'(\mu_k - \mu_m)]$. This leaves $\prod_k'(\mu_j - \mu_k)$ unaffected since $j \neq 0$ or p (by assumption).

$\Pi'_{k<m}(\mu_k - \mu_m)$ is just the van der Monde determinant of

$$\begin{bmatrix} 1 & \mu_0 & \mu_0^2 & \cdots & \mu_0^N \\ 1 & \mu_1 & \mu_1^2 & \cdots & \mu_1^N \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \mu_{j-1} & \mu_{j-1}^2 & \cdots & \mu_{j-1}^N \\ 1 & \mu_{j+1} & \mu_{j+1}^2 & \cdots & \mu_{j+1}^N \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \mu_N & \mu_N^2 & \cdots & \mu_N^N \end{bmatrix}.$$

σ_{p0} corresponds to the interchange of the 0th and the pth row. Hence, its effect on the determinant is to change the sign. Therefore, (VII.7)$_j$ becomes

$$[\ ]_j = -[\ ]_j,$$

and so $[\ ]_j = 0$ and the jth term (VII.6)$_j$ for $p \neq j$ is zero. (Note that if $p = j$, then σ_{p0} also changes the sign of $\Pi'_k(\mu_j - \mu_k)$, and so the pth term needn't vanish.)

(vi) From the previous discussion and the expression (VII.6a, b)$_j$ with $j = p$, we have

$$\langle J_p \rangle = (-1)^p \frac{\sqrt{-1}}{\pi} \int_{a_p} \ln \mu_p [\ ]_p \frac{d\mu_p}{R(\mu_p)},$$

where

$$\begin{aligned} [\ ]_p &= \int_{a_1} \cdots \int_{a_N} \left(1 + \frac{1}{16} \frac{\sqrt{\Pi_k E_k}}{\Pi_k \mu_k}\right) \\ &\quad \times \Pi'_k(\mu_p - \mu_k) \frac{\Pi'_{k<m}(\mu_k - \mu_m)}{\Pi'_k R(\mu_k)} \, d\mu_1 \cdots d\mu_0 \cdots d\mu_N \\ &= \Omega^{(x)}(\mu_p); \end{aligned}$$

the last equality comes from §III.D of [Forest and McLaughlin (1983)]. Thus

$$\begin{aligned} \langle J_p \rangle &= (-1)^p \frac{\sqrt{-1}}{\pi} \int_{a_p} \ln(\mu) \Omega^{(x)}(\mu) \\ &= J_p \cdot (-1)^{p+1}, \end{aligned}$$

where J_p is as defined by (VI.1b).

Thus we have shown that, up to sign, the component $\langle J_p \rangle$ and J_p of the two wave actions agree. This sign difference is unimportant, since both components act as action variables. Moreover, the sign difference can be attributed to the orientation of the a_p cycles and hence of the angle θ_p. We summarize the results of this section in a

PROPOSITION. *The two wave actions J_p and $\langle J_p \rangle$ are equal up to sign. Hence the canonical structure of* [*Forest and McLaughlin* (1979)] *agrees with the structure obtained in section* V *via pull-back.*

VIII. Conclusion

1. We have shown that the full sine-Gordon Hamiltonian structure, when restricted to slowly modulating N-phase wavetrains and reduced by the action of phase shifting, induces the Hamiltonian structure for the modulation equations. In this reduced Hamiltonian system, the canonical variables are the local wave numbers $\vec{\kappa}(x)$ and the local actions $\vec{J}(x)$; the reduced Hamiltonian is a phase average of the original Hamiltonian. Notice that the local action variables depend upon time. This time dependence of $\vec{J}(x)$ arises because the Hamiltonian $\hat{H}$ which governs the reduced system is not the (completely integrable) Hamiltonian in the class of N-phase waves with fixed spatial wave numbers. Rather, it is an average over such local Hamiltonians (see the appendix for more on this point of view); the wave numbers $\vec{\kappa}(x)$ and $\hat{H}$ depend upon the slow spatial variable x. This change in spatial wave numbers induces the time dependence of the action variables $\vec{J}(x)$.

2. The Hamiltonian form of the modulation equations for the sine-Gordon system was presented in a compact and useful form in [Forest, McLaughlin (1983)]:

$$\vec{\kappa}_t = \partial_x \frac{\partial H}{\partial \vec{J}}$$

$$\vec{J}_t = \partial_x \frac{\partial H}{\partial \vec{\kappa}}, \tag{VIII.1}$$

where

$$\kappa_i = \int_{b_i} \Omega^{(x)} \qquad J_i = \int_{a_i} \frac{1}{\mu} \int^{\mu} \Omega^{(x)}.$$

This is a consequence of the more fundamental relation

$$\Omega_t^{(x)} - \Omega_x^{(t)} = 0. \tag{VIII.2}$$

$\Omega^{(x)}$ and $\Omega^{(t)}$ are meromorphic differentials on the Riemann surface $\mathscr{R}$ which enter the theory through the spectral problem that integrates sine-Gordon.

variables: (J_i, θ_i), $i = 1, \ldots, N$. In these variables, (A.2) is completely integrable:

$$\dot{J}_i = -\frac{\partial H_N}{\partial \theta_i} = 0,$$

$$\dot{\theta}_i = \frac{\partial H_N}{\partial J_i} \equiv \omega_i \quad (\text{the } i\text{th temporal frequency}). \tag{A.3a, b}$$

In addition, the angles θ_i satisfy

$$(\theta_i)_x = \kappa_i \quad (\text{the } i\text{th spatial frequency}). \tag{A.3c}$$

A.2. Modulations of the parameters. The wave numbers κ_i and the actions J_i are constants which parameterize the exact N-phase solutions u_N. We aim to parameterize a solution $u(x, t)$ of the full system (A.1) locally by u_N. For this approximation to remain uniformly valid over long distances, the parameters (κ_i, J_i) must vary on slow scales.

To describe this situation, we use the scaling parameter ε ($0 < \varepsilon \ll 1$), as it appears in the sine-Gordon equation (I.1), to define multiple scales:

$$x, X = \frac{x}{\varepsilon} \quad \text{and} \quad t, T = \frac{t}{\varepsilon} \tag{A.1a$^\varepsilon$}$$

$$H^\varepsilon(u) = \int_{-\infty}^{\infty} h^\varepsilon(u(x))\, dx \tag{A.1b$^\varepsilon$}$$

$$\varepsilon \vec{u}_t = \mathscr{J}^\varepsilon \frac{\delta H^\varepsilon}{\delta \vec{u}}. \tag{A.1c$^\varepsilon$}$$

In addition to the slow phases θ_j, we introduce fast phases

$$\Theta_j \equiv \theta_j(x, t)/\varepsilon, \qquad j = 1, 2, \ldots, N.$$

In terms of these variables, the wave u is described locally by the N-phase wave form evaluated on these phases:

$$u \simeq u_N\left(\frac{\theta_i(x,t)}{\varepsilon}, \ldots, \frac{\theta_N(x,t)}{\varepsilon}; \vec{\kappa}(x,t), \vec{J}(x,t)\right) \equiv u^\varepsilon.$$

Here, as in the body of the text, u^ε is the modulation ansatz. For sine-Gordon u^ε is $(u^\varepsilon, v^\varepsilon)$ as in the text. This N-phase waveform will be a local solution to (A.1b)$^\varepsilon$, provided the phases $\vec{\theta}(x, t; X, T)$ are tied to the parameters $(\vec{\kappa}, \vec{J})$ as

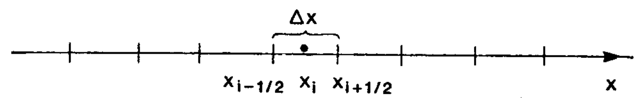

FIGURE 3

follows:

$$\frac{\partial}{\partial x}\theta_j = \kappa_j$$

$$\frac{\partial}{\partial t}\theta_j = \omega_j = \frac{\delta H_N^{(\varepsilon)}}{\delta J_j}. \qquad \text{(A.4a, b)}$$

This means that locally the wave is a highly oscillatory N-phase wave with wave numbers and frequencies which change moderately over long distances. Because of these modulations, we do not know the x, t dependence of the phases as yet.

To describe this situation, we construct a "subspace" $\hat{\mathscr{M}}^\varepsilon$ of the full phase space $\mathscr{F}$ by pasting together local phase spaces $\mathscr{F}_N(\vec{\kappa}(x))$, each of which is indexed by the long scale x. For this construction of $\hat{\mathscr{M}}^\varepsilon$, consider the x-axis partitioned into blocks of width Δx (Figure 3).

In a block, the slow x dependence of the wave u on the parameters $\vec{\kappa}(x, t)$ and $\vec{J}(x, t)$ is essentially constant. However, as one moves from block to block, these parameters change. The subspace $\hat{\mathscr{M}}^\varepsilon$ is the two-timing version of the space $\mathscr{M}^\varepsilon$ in the body of the text.

Thus, we construct a phase space $\hat{\mathscr{M}}^\varepsilon$ by pasting together the spaces of N-phases waves $\mathscr{F}_N$ for each block:

$$\hat{\mathscr{M}}^\varepsilon = \prod_{x \in \mathbf{R}} \mathscr{F}_N[\vec{\kappa}(x)].$$

We treat this "space of slowly varying N-phase waves" as a subspace of the full phase space $\mathscr{F}$. The coordinates of a slowly modulating wavetrain in $\hat{\mathscr{M}}^\varepsilon$ are

$$\vec{J}(x)\text{—local actions}$$

$$\vec{\Theta}(x)\text{—local phases,}$$

one for each slow variable x, just as we found for $\mathscr{M}^\varepsilon$ in the body of the text. The phases satisfy

$$\vec{\Theta}_x = \frac{\vec{\kappa}}{\varepsilon}(x).$$

The local wavetrain in the ith box depends upon the local wavetrain in adjacent boxes. This coupling of the local wavetrains results mathematically from the restriction of the full Hamiltonian to slowly varying N-phase waves in $\hat{\mathcal{M}}^{\varepsilon}$, $H\colon \hat{\mathcal{M}}^{\varepsilon} \to \mathbf{R}$. We compute this restriction next.

A.3. *A reduced Hamiltonian*. The idea now is to use the full Hamiltonian H^{ε}, (A.1b)$^{\varepsilon}$, but restricted to the subspace $\hat{\mathcal{M}}^{\varepsilon}$, to characterize the dynamics of the slow modulations. Upon approximating this restriction, we obtain an *approximate Hamiltonian*:

$$\hat{H}^{\varepsilon}\colon \hat{\mathcal{M}}^{\varepsilon} \to \mathbf{R} \text{ by } \hat{H}^{\varepsilon}[u^{\varepsilon}] = \int_{-\infty}^{\infty} H_N^{\varepsilon}[u^{\varepsilon}(\cdot, x)]\, dx. \tag{A.5a}$$

We emphasize that the Hamiltonian *d*ensity for the approximation Hamiltonian $\hat{H}^{\varepsilon}$ is just the local N-phase Hamiltonian $H_N^{(\varepsilon)}$ indexed by the slow scale x. Thus, the approximate Hamiltonian $\hat{H}^{(\varepsilon)}$ is an "average over the slow scale x" of the N-phase local Hamiltonian. In the remainder of this subsection A.3, we will carry out the reduction of the full Hamiltonian $H^{(\varepsilon)}$ to this approximate Hamiltonian $\hat{H}^{(\varepsilon)}$.

To begin, we evaluate the full Hamiltonian H on a slowly modulating N-phase wavetrain u^{ε}. Then we approximate the Hamiltonian by partitioning the x-axis into blocks of length Δx; these blocks are moderate on the x scale, but long on the x/ε scale (Figure 3):

$$H^{\varepsilon}(u^{\varepsilon}) = \sum_{i=-\infty}^{\infty} \int_{x_{i-1/2}}^{x_{i+1/2}} h^{\varepsilon}\left[u_N\left(\frac{\vec{\theta}(x)}{\varepsilon}; \vec{\kappa}(x), \vec{J}(x)\right)\right] dx \quad (\text{definition of } H^{\varepsilon})$$

$$\simeq \sum_{i=-\infty}^{+\infty} \int_{x_i-1/2}^{x_{i+1/2}} h^{\varepsilon}\left[u_N\left(\frac{\vec{\theta}(x)}{\varepsilon}; \vec{\kappa}(x_i), \vec{J}(x_i)\right)\right] dx$$

(freezing the slow dependence of $\vec{J}$, $\vec{\kappa}$ at the midpoint of each block)

$$\simeq \sum_{i=-\infty}^{+\infty} \langle h^{\varepsilon}\rangle \Delta x_i \qquad (\text{defines the spatial average } \langle h\rangle)$$

$$\simeq \int_{-\infty}^{\infty} \langle h\rangle\, dx, \qquad (\text{Riemann sum})$$

where the average $\langle h^{\varepsilon}\rangle$ is defined by

$$\langle h^{\varepsilon}\rangle \equiv \frac{1}{\Delta x_i} \int_{x_{i-1/2}}^{x_{i+1/2}} h^{\varepsilon}\left[u_N\left(\frac{\vec{\theta}(x)}{\varepsilon}; \vec{\kappa}(x_i), \vec{J}(x_i)\right)\right] dx$$

and where

$$\langle h \rangle \equiv \lim_{\varepsilon \to 0} \langle h^\varepsilon \rangle = H_N\big(u_N\big(\cdot, \vec{\kappa}(x_i), \vec{J}(x_i)\big)\big),$$

which yields (A.5a):

$$\hat{H}^\varepsilon(u^\varepsilon) \equiv \int_{-\infty}^{\infty} H_N^\varepsilon(u^\varepsilon(\cdot, x))\, dx$$

$$= H(\vec{\kappa}, \vec{J}), \quad \text{of §V.}$$

This last equality is clear from the ergodicity of the translational flow when all the κ_i are irrationally related. When some of the κ_i are rationally related, it follows from continuity.

A.4. Modulational equations as a Hamiltonian system. In this section we show that the modulation equations themselves are a Hamiltonian system with Hamiltonian $\hat{H}^\varepsilon$:

$$\frac{d}{dt}\begin{pmatrix} \vec{\kappa} \\ \vec{J} \end{pmatrix} = \mathscr{J}\begin{pmatrix} \dfrac{\delta \hat{H}^\varepsilon}{\delta \vec{\kappa}} \\ \dfrac{\delta \hat{H}^\varepsilon}{\delta \vec{J}} \end{pmatrix}. \tag{A.6}$$

Here the phase point has coordinates $(\vec{\kappa}, \vec{J})$, and $\mathscr{J}$ denotes the antisymmetric differential operator

$$\mathscr{J} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \frac{\partial}{\partial x}.$$

To deduce this Hamiltonian form of the modulation equations, we assume that the reduced Hamiltonian $\hat{H}^\varepsilon$ generates the correct flow on $\mathscr{M}^\varepsilon$ by

$$\varepsilon \frac{d}{dt}\vec{\Theta} = \frac{\delta \hat{H}^\varepsilon}{\delta \vec{J}}$$

$$\varepsilon \frac{d}{dt}\vec{J} = -\frac{\delta \hat{H}^\varepsilon}{\delta \vec{\Theta}}.$$

We differentiate the first equation with respect to x, and use $\vec{\Theta}_x = \vec{\kappa}/\varepsilon$ to replace it by

$$\frac{d}{dt}\vec{\kappa} = \frac{\partial}{\partial x}\frac{\delta \hat{H}^\varepsilon}{\delta \hat{J}}.$$

Now consider the second equation. Notice that the approximate Hamiltonian $\hat{H}^\varepsilon$ depends on $\vec{\Theta}$ only through $\vec{\Theta}_x = \vec{\kappa}/\varepsilon$; therefore,

$$\frac{\delta \hat{H}^\varepsilon}{\delta \vec{\Theta}} = -\frac{\partial}{\partial x}\frac{\partial \hat{H}^\varepsilon}{\partial \vec{\Theta}_x} = -\varepsilon \frac{\partial}{\partial x}\frac{\delta \hat{H}^\varepsilon}{\delta \vec{\kappa}}.$$

With these observations, the action-angle formulas lead to a *Hamiltonian form of the modulation equations*:

$$\frac{d}{dt}\vec{\kappa}(x) = \frac{\partial}{\partial x}\left[\frac{\partial H_N^\varepsilon(u_N(\cdot, x))}{\partial \vec{J}(x)}\right],$$

$$\frac{d}{dt}\vec{J}(x) = \frac{\partial}{\partial x}\left[\frac{\partial H_N^\varepsilon(u_N(\cdot, x))}{\partial \vec{\kappa}(x)}\right];$$

that is, to (A.6).

Note that this calculation of $\mathscr{J}$ is precisely the calculation of the reduced brackets done in §V.

We have again shown that the full Hamiltonian, when restricted to the slowly modulating N-phase wavetrains, induces a Hamiltonian structure for the modulation equations, and have arrived at exactly the same modulational Poisson structure as in the text.

REFERENCES

1. M. J. ABLOWITZ AND D. J. BENNEY, *The evolution of multi-phase modes for nonlinear dispersive waves*, Stud. Appl. Math. **49** (1970), 225–238.
2. M. J. ABLOWITZ, D. KAUP, A. NEWELL, AND H. SEGUR, Phys. Rev. Lett. **30** (1973), 1262.
3. R. ABRAHAM AND J. E. MARSDEN, *Foundation of Mechanics*, Benjamin/Cummings, Reading, MA, 1978.
4. E. DATÉ, *Multisoliton solutions and quasi-periodic solutions of nonlinear equations of sine-Gordon type*, Osaka J. Math. **19** (1982), 125–128.
5. N. M. ERCOLANI AND M. G. FOREST, *The geometry of real sine-Gordon wavetrains*, Comm. Math. Phys. **99** (1985), 1–49.
6. N. ERCOLANI, M. G. FOREST, AND D. W. MCLAUGHLIN, *Modulational stability of two phase sine-Gordon wavetrains*, Stud. Appl. Math. **71** (1984), 91–101.
7. ______, "Oscillations and instabilities in near integrable PDE's," in *Proc. of AMS-SIAM Conference on Nonlinear Evolution Equations*, Santa Fe, ed. D. Holm, J. M. Hyman, and B. Nicolaenko, 1985.
8. J. FAY, *Theta Functions on Riemann Surfaces*, Lecture Notes in Mathematics **352**, Springer-Verlag, New York, 1984.
9. H. FLASCHKA, M. G. FOREST, AND D. W. MCLAUGHLIN, *Multiphase averaging and the inverse spectral solution of the Korteweg–deVries equation*, Comm. Pure Appl. Math. **33** (1980), 739–784.
10. H. FLASCHKA AND D. W. MCLAUGHLIN, *Canonically conjugate variables for the KdV...*, Prog. Theor. Phys. **55** (1976), 438–456.
11. H. FLASCHKA, A. C. NEWELL, AND T. RATIU, *Kac-Moody Lie algebras and soliton equations, II*, Physica **9D** (1983), 300–323.

12. M. G. Forest and D. W. McLaughlin, *Canonical variables for the periodic sine-Gordon equation and a method of averaging*, Report of the Los Alamos National Laboratory, LA-UR 78-3318, (1978).

13. ______, *Modulations and sinh-Gordon and sine-Gordon wavetrains*, Stud. Appl. Math. **68** (1983), 11–59.

14. ______, *Modulations of perturbed KdV wavetrains*, SIAM J. Appl. Math. **44** (1984), 287–300.

15. ______, *Spectral theory for the periodic sine-Gordon equation: A concrete viewpoint*, J. Math. Phys. **23** (1982), 1248–1277.

16. B. Fornberg and G. B. Whitham, *A numerical and theoretical study of certain nonlinear wave phenomena*, Phil. Trans. Roy. Soc. Lond. **289** (1978), 373–404.

17. P. Griffiths and J. Harris, *Principles of Algebraic Geometry*, Wiley-Interscience, New York, 1978.

18. A. V. Gurevich and L. P. Pitaevskii, *Nonstationary structure of a collisionless shock wave*, Soviet Phys. JETP **38**(2) (1974); and B. Fornberg and G. B. Whitham, *A numerical and theoretical study of certain nonlinear wave phenomena*, Phil. Trans. Roy. Soc. Lond. **289** (1978), 373–404.

19. W. D. Hayes, *Group velocity and nonlinear dispersive wave propagation*, Proc. Royal Soc. London **A332** (1973), 199–221.

20. P. D. Lax and C. D. Levermore, *The small dispersion limit for the KdV equation, I*, Comm. Pure Appl. Math. **36** (1983), 253–290; *II*, Comm. Pure Appl. Math. **36** (1983), 571–594; *III*, Comm. Pure Appl. Math. **36** (1983), 809–829.

21. ______, *The zero dispersion limit for the KdV equation*, Proc. Nat. Acad. Sci. **76** (1979), 3602–3606.

22. J. C. Luke, *A perturbation method for nonlinear dispersive wave problems*, Proc. Roy. Soc. **A292** (1966), 403–412.

23. H. P. McKean, *The sine-Gordon and sinh-Gordon equations on the circle*, C.P.A.M. **34** (1981), 197.

24. D. W. McLaughlin, *Modulations of KdV wavetrains*, Physica **D3** (1981), 335–343.

25. ______, "On the construction of a modulating, multiphase wavetrain for a perturbed KdV equation," to appear in *Proc. of Conf. on Rapid Oscillation Theory*, Minnesota, ed. D. Kenderlehrer and M. Slemrod, 1985.

26. R. Miura and M. Kruskal, *Application of a nonlinear WKB method to the Korteweg–deVries equation*, SIAM J. Appl. Math. **26** (1974), 376–395.

27. D. Mumford, *Tata Lectures on Theta II*, Birkhäuser, Boston, 1983.

28. M. E. Schonbek, *Convergence of solutions to nonlinear dispersive equations*, preprint, Univ. Rhode Island, 1981.

29. L. A. Takhtajian and L. D. Faddeev, *Essentially nonlinear one-dimensional model of a classical field theory*, Theor. Math. Phys. **21** (1974), 1046.

30. S. Venakides, *The generation of modulated wavetrains in the solutions of the Korteweg-deVries equation*, preprint, 1985.

31. ______, *The zero dispersion limit of the Korteweg–deVries equation for initial potentials with non-trivial reflection coefficient*, dissertation, New York University, 1982.

32. ______, *The zero dispersion limit of the Korteweg–deVries equation for initial potentials with non-trivial reflection coefficient*, Comm. Pure Appl. Math. **38** (1985), 125–155.

33. G. B. Whitham, *Linear and Nonlinear Dispersive Waves*, Wiley-Interscience, New York, 1974.

34. ______, *Nonlinear dispersive waves*, Proc. Royal Soc. London **A283** (1965), 238–261.

Ercolani and Forest: Department of Mathematics, The Ohio State University, Columbus, Ohio 43210

McLaughlin: Department of Mathematics, University of Arizona, Tucson, Arizona 85721

Montgomery: Department of Mathematics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

DUKE MATHEMATICAL JOURNAL

Current subscriptions: For libraries and institutions, $192 per year for two volumes of six issues total. For individuals, $96 per year for two volumes of six issues total. For individuals whose institution already subscribes (please provide name and address of institutional affiliation), $48 per year for two volumes of six issues total. *For subscriptions outside the U.S.*, please add $9 per year for postage and handling. All subscription orders must be on a calendar-year basis. Send orders to Journals Dept., Duke University Press, 6697 College Station, Durham, NC 27708, USA.

Back orders and single issues: Back copies from Volume 35 (1967) to the present are available from Duke University Press at the address above. The price for back volumes is $128; for single issues, $32. For information on volumes prior to 1968, write to Johnson Reprint Corp., 111 Fifth Ave., New York, NY 10003.

Permissions: Please address all requests for permission to use copyrighted material (beyond "fair use") to the Permissions Dept., Duke University Press, 6697 College Station, Durham, NC 27708, USA.

Library exchanges: Proposals for exchange of library materials in return for subscriptions to the *Duke Mathematical Journal* will be considered. The proposals should be sent to the Gift and Exchange Dept., Perkins Library, Duke University, Durham, NC 27706, USA.

The *Duke Mathematical Journal* (ISSN 0012-7094) is published bimonthly at the prices listed above, by Duke University Press, 6697 College Station, Durham, NC 27708. Second-class postage paid at Durham, N.C. POSTMASTER: Send address changes to Journals Dept., Duke University Press, 6697 College Station, Durham, NC 27708.