Math. Proc. Camb. Phil. Soc. (1997), **121**, 147
Printed in Great Britain

Multi-symplectic structures and wave propagation

By THOMAS J. BRIDGES

Department of Mathematical and Computing Sciences, University of Surrey, Guildford, GU2 5XH

(*Received* 2 *June* 1994; *revised* 3 *April* 1995)

Abstract

A Hamiltonian structure is presented, which generalizes classical Hamiltonian structure, by assigning a distinct symplectic operator for each unbounded space direction and time, of a Hamiltonian evolution equation on one or more space dimensions. This generalization, called multi-symplectic structures, is shown to be natural for dispersive wave propagation problems. Application of the abstract properties of the multi-symplectic structures framework leads to a new variational principle for space-time periodic states reminiscent of the variational principle for invariant tori, a geometric reformulation of the concepts of action and action flux, a rigorous proof of the instability criterion predicted by the Whitham modulation equations, a new symplectic decomposition of the Noether theory, generalization of the concept of reversibility to space-time and a proof of Lighthill's geometric criterion for instability of periodic waves travelling in one space dimension. The nonlinear Schrödinger equation and the water-wave problem are characterized as Hamiltonian systems on a multi-symplectic structure for example. Further ramifications of the generalized symplectic structure of theoretical and practical interest are also discussed.

1. *Introduction*

An understanding of the existence, propagation, stability, bifurcation, dynamics, breakup and other properties of wave motion are of fundamental importance in physical problems such as oceanic waves, atmospheric dynamics, wave guides, optics, the nervous system, shear flows, acoustics, gas dynamics and many other areas. In many cases of wave propagation, particularly the equations governing ocean waves and atmospheric flow, a conservative model is accurate and when studying conservative partial differential equations it is natural to appeal to the powerful geometric methods of Lagrangian and Hamiltonian mechanics.

The Lagrangian formulation and the Hamiltonian formulation for a conservative system are usually considered to be dually related through the Legendre transform and, in finite dimensions, when the Legendre transform is non-degenerate, the duality is exact. However in infinite dimensions, particularly systems governing wave propagation where one or more spatial directions is infinite, the Lagrange–Hamiltonian duality is no longer uniquely defined. To clarify this point consider the following elementary example of a nonlinear Klein–Gordon equation

$$u_{tt} - u_{xx} = V'(u) \quad x \in \mathbb{R}, \quad t > 0 \tag{1·1}$$

https://doi.org/10.1017/S0305004196001429 Published online by Cambridge University Press

where $V(u)$ is some smooth nonlinear function of u. The Lagrangian formulation for (1·1) is well understood; the Lagrangian functional for (1·1) is

$$\mathscr{L} = \int_{t_1}^{t_2}\int_{x_1}^{x_2} L(u, u_t, u_x)\,dx\,dt \quad \text{with} \quad L(u, u_t, u_x) = \tfrac{1}{2}u_t^2 - \tfrac{1}{2}u_x^2 + V(u) \tag{1·2}$$

and the equation (1·1) is formally recovered from the Euler–Lagrange equation:

$$0 = \frac{\partial}{\partial t}\left(\frac{\partial L}{\partial u_t}\right) + \frac{\partial}{\partial x}\left(\frac{\partial L}{\partial u_x}\right) - \frac{\partial L}{\partial u} = u_{tt} - u_{xx} - V'(u).$$

The Lagrangian formulation of the equations governing wave propagation problems forms the basis for the Whitham modulation theory for periodic waves (Whitham [**33**, **34**, **35**]) which we will discuss in more detail shortly.

Formally taking the Legendre transform of the Lagrange density L results in a Hamiltonian formulation; let $v = \partial L/\partial u_t = u_t$. Then the system (1·1) has 'the' Hamiltonian formulation

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\frac{\partial}{\partial t}\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} \delta H/\delta u \\ \delta H/\delta v \end{pmatrix}, \tag{1·3a}$$

where

$$H(u, v) = \int_{x_1}^{x_2} [\tfrac{1}{2}v^2 + \tfrac{1}{2}u_x^2 - V(u)]\,dx. \tag{1·3b}$$

An advantage of the Hamiltonian formulation is that the system is in the form of an evolution equation in time and therefore one can appeal to results in the literature for existence and other properties of the initial value problem. Another advantage is the organising structure provided by the symplectic operator. On the other hand, to be precise, it is necessary to consider a space of functions, including integration over x, for which the integral of H is well defined. Therefore such a Hamiltonian formulation is most useful when the spatial domain is finite or when one is interested in a wave with a particular spatial variation – for example, a spatially periodic wave or a class of waves that decays exponentially as $x \to \pm\infty$.

However, to say that (1·3) is the Hamiltonian formulation corresponding to (1·1) or to say that the Hamiltonian formulation (1·3) is dual to the Lagrangian formulation (1·2) is not completely correct. In effect the Legendre transform leading to (1·3) is only a partial Legendre transform. A complete Legendre transform would also eliminate the x-derivatives in the Lagrangian density $L(u, u_t, u_x)$. In other words let $w = \partial L/\partial u_x = u_x$; then the transformation of equation (1·1), after the complete Legendre transform, takes the form

$$\mathbf{M}Z_t + \mathbf{K}Z_x = \nabla S(Z) \quad \text{where} \quad Z = \begin{pmatrix} u \\ v \\ w \end{pmatrix} \in \mathscr{M} \subset \mathbb{R}^3, \tag{1·4}$$

$$\mathbf{M} = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \mathbf{K} = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

and

$$S(Z) = \tfrac{1}{2}(w^2 - v^2) + V(u).$$

The representation (1·4) organizes neatly each facet of the equation. All time

https://doi.org/10.1017/S0305004196001429 Published online by Cambridge University Press

derivatives appear in the term $\mathbf{M}Z_t$, all space derivatives appear in the term $\mathbf{K}Z_x$ and the gradient of $S(Z)$ is defined with respect to an inner product on $\mathscr{M}$ which in this case is $\mathbb{R}^3$.

The system (1·4) is a Hamiltonian formulation of (1·1) on a multi-symplectic structure, since, instead of a single symplectic form, as in (1·3), there are two pre-symplectic operators $\mathbf{M}$ and $\mathbf{K}$ – that is; skew-symmetric and can be identified with closed two forms – that define a generalized symplectic structure for the system.

To appreciate the distinction between the classical Hamiltonian formulation (1·3), with a single symplectic operator, and the Hamiltonian formulation on a multi-symplectic structure in (1·4) it is useful to reconsider the role of the Legendre transform. In the first, partial, Legendre transform (a) a new set of variables $(u, v) \stackrel{\text{def}}{=} (u, u_t)$ was introduced, (b) a Hamiltonian functional $H(u, v)$ was generated and (c) an action density was created. The action density in this case is vu_t and the gradient of action (with respect to an inner product that includes integration over t) results in the left-hand side of (1·3) and hence is responsible for generating the single symplectic operator in the system.

In addition to new variables the complete Legendre transform contributes a family of action densities and so a family of symplectic operators and in addition a new Hamiltonian functional is created which is stripped of any explicit derivatives.

In other words a partial Legendre transform and a complete Legendre transform will lead to non-trivially different symplectic structures and Hamiltonian systems. It is the main argument, and central organizing feature, of this paper that the Hamiltonian formulation on a multi-symplectic structure, of the abstract form (1·4) and its generalization to higher space dimension, is a natural framework for analysing, and proving particular properties of, dispersive wave propagation in conservative systems. One of the main results of setting the equations on a multi-symplectic structure is a framework leading to a rigorous proof of the instability of periodic travelling waves that is predicted by the Whitham modulation equations.

The Whitham theory for modulation of periodic travelling waves, which was developed within a Lagrangian framework (Whitham [**33**, **34**, **35**]) is summarized as follows, using (1·1) as an example. Let $\theta = kx - \omega t$; then the organizing centre for the Whitham theory is the averaged Lagrangian

$$\mathscr{L} = \int_0^{2\pi} L(u, -\omega u_\theta, k u_\theta)\, d\theta. \tag{1·5}$$

The existence theory of such waves proceeds by studying critical points of (1·5) resulting in a family of periodic travelling waves $(u(\theta; \omega, k), \omega, k)$. Modulated travelling waves are treated by allowing ω and k to be slowly varying functions of x and t. Defining $\omega = -\theta_t$ and $k = \theta_x$ leads to the kinematic relation

$$k_t + \omega_x = 0. \tag{1·6}$$

Whitham then defines an action and action flux by

$$\mathscr{A} = \frac{\partial \mathscr{L}}{\partial \omega} \quad \text{and} \quad \mathscr{B} = -\frac{\partial \mathscr{L}}{\partial k} \tag{1·7}$$

https://doi.org/10.1017/S0305004196001429 Published online by Cambridge University Press

and introduces the *conservation of wave action*

$$\frac{\partial \mathscr{A}}{\partial t}+\frac{\partial \mathscr{B}}{\partial x}=0. \tag{1·8}$$

Combining (1·6) and (1·8) then leads to the coupled Whitham modulation equations

$$\begin{pmatrix}\mathscr{A}_\omega & \mathscr{A}_k\\ 0 & 1\end{pmatrix}\begin{bmatrix}\omega\\ k\end{bmatrix}_t+\begin{pmatrix}\mathscr{B}_\omega & \mathscr{B}_k\\ 1 & 0\end{pmatrix}\begin{bmatrix}\omega\\ k\end{bmatrix}_x=0. \tag{1·9a}$$

The principal use of the above modulation equations is for predicting the linear instability of periodic travelling waves of arbitrary amplitude. The linear stability problem is formulated by taking

$$\begin{pmatrix}\omega\\ k\end{pmatrix}=\begin{pmatrix}\omega_0\\ k_0\end{pmatrix}+\begin{pmatrix}\hat{\omega}\\ \hat{k}\end{pmatrix}e^{i(\alpha x-\lambda t)},$$

where (ω_0, k_0) represents the basic periodic state. Linearizing (1·9a) about (ω_0, k_0) results in the eigenvalue problem

$$\det\left[-i\lambda\begin{pmatrix}\mathscr{A}_\omega & \mathscr{A}_k\\ 0 & 1\end{pmatrix}+i\alpha\begin{pmatrix}\mathscr{B}_\omega & \mathscr{B}_k\\ 1 & 0\end{pmatrix}\right]=0$$

and the expression

$$\lambda=\frac{\alpha}{\mathscr{A}_\omega}[\tfrac{1}{2}(\mathscr{A}_k-\mathscr{B}_\omega)\pm\sqrt{-\Delta_{\mathscr{L}}}]\quad \mathscr{A}_\omega\neq 0 \tag{1·9b}$$

for the linear stability exponent, where

$$\Delta_{\mathscr{L}}=-\tfrac{1}{4}(\mathscr{A}_k-\mathscr{B}_\omega)^2-\mathscr{A}_\omega\mathscr{B}_k. \tag{1·10a}$$

In other words if there exists a root $\lambda\in\mathbb{C}$ with $\Im(\lambda)\neq 0$, equivalently $\Delta_{\mathscr{L}}>0$, there is an exponentially growing solution of the system (1·9a) linearized about a periodic travelling wave and hence predicts linear instability. Lighthill [**21**, **22**] noted that

$$\Delta_{\mathscr{L}}=\det\begin{pmatrix}\mathscr{L}_{\omega\omega} & \mathscr{L}_{\omega k}\\ \mathscr{L}_{k\omega} & \mathscr{L}_{kk}\end{pmatrix} \tag{1·10b}$$

and the following geometric interpretation: if $\mathscr{L}$ (*or* $-\mathscr{L}$) *in* (1·5) *is a convex function of* ω *and* k *then the basic periodic wave is linearly unstable.*

A proof of validity of the Whitham modulation equations would require showing that solutions of the modulation equation remain near solutions of the original equations, with respect to a suitably defined norm, on a sufficiently long time scale. Formal results on the validity have been given by Luke and Whitham using the method of multiple scales (cf. Whitham [**35**, §14.4]).

On the other hand the most significant application of the Whitham modulation equations is to the prediction of linear instability of periodic travelling waves of arbitrary amplitude. In this paper we formulate rigorously and give a complete proof of, for periodic travelling waves in one or two space dimensions, the instability result predicted by the Whitham theory and the geometric instability criterion of Lighthill. Moreover the reformulation of the problem leads not only to a precise list of the

https://doi.org/10.1017/S0305004196001429 Published online by Cambridge University Press

hypotheses under which the theory is valid, and a complete proof, but also to further results including dual geometric criteria for instability. It is precisely the multi-symplectic structures framework and a geometric reformulation of the concepts of action and action flux that make a rigorous theory possible.

The concepts of action and action flux and the conservation of wave action are recognized as fundamental concepts in wave propagation and, since they were introduced by Whitham, there has been significant generalizations of action and its relatives (Hayes[**18**, **19**], Andrews & McIntyre[**1**], Grimshaw[**17**] and references therein). However in the definition of action and action flux in (1·7), and all related definitions in the literature, the action and action flux are defined in terms of averaged quantities. The lack of structure in the definition of action and action flux is an obstacle to a rigorous theory. One of the main results of this paper is to define primitive functionals for action and action flux; they are one-forms whose exterior derivative results in closed two forms and it is precisely the structure of the closed two forms that forms a basis for the rigorous theory of wave instability.

Although many of the properties of wave dynamics can be appreciated within a Lagrangian framework, a Hamiltonian framework is a useful one within which to prove results. A prototype is the water wave problem. The Lagrangian formulation for water waves was discovered before the Hamiltonian formulation and many properties of waves were illuminated in the Lagrangian framework (cf. Whitham[**34**, **35**]). However the later discovery of a Hamiltonian formulation (cf. Zakharov[**36**], Broer[**15**]) for water waves lead to proofs, and in particular geometric proofs, of many facets of the water-wave problem. For example Benjamin & Olver[**5**] were able to enumerate rigorously, and prove finiteness of, the set of conservation laws for water waves; Saffman[**31**] gave a geometric proof of the superharmonic instability of water waves that used only the Hamiltonian structure and MacKay & Saffman[**23**] defined a signature invariant of the eigenspace corresponding to the linear stability problem for water waves, using the symplectic operator, that proved useful for classifying instabilities of large-amplitude travelling water waves. However in the results above on water waves a partial Hamiltonian structure was used; a classical Hamiltonian structure with a single symplectic operator. By considering the water wave problem as a Hamiltonian system on a multi-symplectic structure further results are possible. In Bridges & Mielke[**13**] a proof of the Benjamin–Feir instability is given which relies crucially on the interplay between the spatial and temporal symplectic structures. In Bridges[**11**] the multi-symplectic structures framework is applied to formulate the instability problem for arbitrary periodic patterns, interacting with a mean flow in finite depth, on the ocean surface.

Abstractly, we call the formulation represented by (1·4), and its generalizations, a Hamiltonian system on a multi-symplectic structure and as far as we are aware the formulation and analysis of a such a structure is new. A symplectic structure $(\mathscr{M}, \omega)$ consists of a manifold $\mathscr{M}$, the phase space, on which there is a non-degenerate closed two-form. A Hamiltonian system is then a triplet $(\mathscr{M}, \omega, H)$ where $H: \mathscr{M} \to \mathbb{R}$ is a functional. The generalisation of this, a multi-symplectic structure $(\mathscr{M}, \omega^{(1)}, \ldots, \omega^{(n)})$, consists of a manifold $\mathscr{M}$, the phase space, and a family of closed, in general non-commuting, two-forms. In the multi-symplectic case we relax the requirement of non-degeneracy; a closed but possibly degenerate two-form is usually called pre-symplectic (cf. Marsden[**25**, p. 29]) but the distinction is not important in the present

https://doi.org/10.1017/S0305004196001429 Published online by Cambridge University Press

theory. A Hamiltonian system on a multi-symplectic structure is then represented symbolically by $(\mathscr{M}, \omega^{(1)}, \ldots, \omega^{(n)}, S)$ with governing equation

$$\omega^{(1)}\left(\frac{\partial Z}{\partial t}, \mathscr{V}\right) + \ldots + \omega^{(n)}\left(\frac{\partial Z}{\partial x_m}, \mathscr{V}\right) = \langle \nabla S(Z), \mathscr{V} \rangle_m \tag{1·11}$$

for all $\mathscr{V} \in T_m \mathscr{M}$ where $\langle \cdot, \cdot \rangle_m$ is an inner product on $T_m \mathscr{M}$ at $m \in \mathscr{M}$. For definiteness we will suppose that the system is tri-symplectic (two space dimensions and time) with

$$\omega^{(1)}(Z_t, \mathscr{V}) \stackrel{\text{def}}{=} \langle \mathbf{M}(Z) Z_t, \mathscr{V} \rangle_m,$$

$$\omega^{(2)}(Z_x, \mathscr{V}) \stackrel{\text{def}}{=} \langle \mathbf{K}(Z) Z_x, \mathscr{V} \rangle_m,$$

$$\omega^{(3)}(Z_y, \mathscr{V}) \stackrel{\text{def}}{=} \langle \mathbf{L}(Z) Z_y, \mathscr{V} \rangle_m.$$

When the inner product $\langle \cdot, \cdot \rangle_m$ is independent of $m \in \mathscr{M}$ the system (1·11) may be written

$$\mathbf{M}(Z) Z_t + \mathbf{K}(Z) Z_x + \mathbf{L}(Z) Z_y = \nabla S(Z) \quad Z \in \mathscr{M}. \tag{1·12}$$

In general the phase space $\mathscr{M}$ is a manifold modelled on $\mathbb{R}^n$, or in infinite dimensions a Hilbert or Banach space. To avoid unnecessary technicalities we will take $\mathscr{M}$ to be a linear space unless noted otherwise.

The formulation (1·11)–(1·12) has properties that make it natural for studying wave propagation in open systems. In fact the components of the formulation (1·11), the Hamiltonian functional S and the skew-symmetric operators $\mathbf{M}(Z)$, $\mathbf{K}(Z)$ and $\mathbf{L}(Z)$, can be identified with quantities that are known to be important for wave propagation in conservative systems. In one space dimension the Hamiltonian functional $S(Z)$ is related to the flow force, or in some cases the momentum flux, for the system (see Appendix B). The definition of the flow force and its implications for wave propagation are discussed in Benjamin [**3**] and references therein. In two space dimensions the connection between $S(Z)$ and a physically relevant functional is not so clear. The operators $\mathbf{M}(Z)$, $\mathbf{K}(Z)$ and $\mathbf{L}(Z)$ are structure operators for the action and action-flux vector.

The hypothesis of closedness for the symplectic operators $\mathbf{M}(Z)$, $\mathbf{K}(Z)$ and $\mathbf{L}(Z)$ is necessary to ensure that, at least locally, the terms $\mathbf{M}(Z) Z_t$, $\mathbf{K}(Z) Z_x$ and $\mathbf{L}(Z) Z_y$ are gradients of certain functionals. In order to avoid unnecessary technicalities we will also assume that the symplectic operators are associated with two-forms that are exact, which is often the case in applications, in which case the terms $\mathbf{M}(Z) Z_t, \ldots,$ are globally defined as gradient operators. In particular there exist, formally, functionals $A(Z)$, $B_1(Z)$ and $B_2(Z)$ such that

$$\nabla A = \mathbf{M}(Z) Z_t$$

$$\nabla B_1 = \mathbf{K}(Z) Z_x$$

$$\nabla B_2 = \mathbf{L}(Z) Z_y$$

where the gradients are defined with respect to an inner product that includes integration over x, y or t. The densities A, B_1 and B_2 are independent action functionals for the t, x and y directions respectively and can be identified with one-

https://doi.org/10.1017/S0305004196001429 Published online by Cambridge University Press

forms; that is, there exist, formally, vector-valued functions $\alpha(Z)$, $\beta_1(Z)$ and $\beta_2(Z)$ such that

$$\begin{aligned} A(Z) &= \langle \alpha(Z), Z_t \rangle_m \\ B_1(Z) &= \langle \beta_1(Z), Z_x \rangle_m \\ B_2(Z) &= \langle \beta_2(Z), Z_y \rangle_m. \end{aligned} \tag{1·13}$$

When the inner product does not depend explicitly on $m \in \mathscr{M}$ there is a simple relation between the forms α, β_1 and β_2 and the symplectic operators, namely

$$\begin{aligned} \mathbf{M}(Z) &= D\alpha(Z)^* - D\alpha(Z) \\ \mathbf{K}(Z) &= D\beta_1(Z)^* - D\beta_1(Z) \\ \mathbf{L}(Z) &= D\beta_2(Z)^* - D\beta_2(Z) \end{aligned} \tag{1·14}$$

where D represents differentiation on the phase space (when $\mathscr{M}$ is finite dimensional it is the Jacobian operation) and $*$ indicates formal adjoint. For example for the multi-symplectic structure presented in equation (1·4) for the wave equation (1·1) the action one-forms are

$$A(Z) = \langle \alpha(Z), Z_t \rangle \quad \text{with} \quad \alpha(Z) = \begin{pmatrix} v \\ 0 \\ 0 \end{pmatrix},$$

$$B(Z) = \langle \beta(Z), Z_x \rangle \quad \text{with} \quad \beta(Z) = \begin{pmatrix} w \\ 0 \\ 0 \end{pmatrix},$$

where, here, $\langle \cdot, \cdot \rangle$ is the standard inner product on $\mathbb{R}^3$. In this case the Jacobian operation is the usual derivative and the adjoint operation is matrix transposition.

When ensemble averaged or averaged over a phase, the values of the functionals $\mathscr{A}$, $\mathscr{B}_1$ and $\mathscr{B}_2$, which represent the averaged values of A, B_1 and B_2, recover the Whitham–Hayes definition of action and (two-component) action flux respectively. Although, in the present formulation, a conservation law for action is neither implied nor required but for historical reasons we refer to the primitive definitions in (1·13), even without averaging, as the action and action-flux densities.

A remarkable consequence of the formulation (1·11)–(1·12) is that a number of general results about waves can be derived and in some cases rigorously proved using only the abstract properties of the mult-symplectic structure. In Section 2 a new variational principle for space and time periodic states is introduced. In one-space dimension the idea can be seen using (1·12) restricted to one space dimension. With $\tau = \omega t$ and $\xi = kx$ (1·12), restricted to one space dimension and 2π-periodic, in (ξ, τ), functions is transformed to

$$\omega \mathbf{M}(Z) \frac{\partial Z}{\partial \tau} + k \mathbf{K}(Z) \frac{\partial Z}{\partial \xi} = \nabla S(Z) \quad \text{with} \quad (\tau, \xi) \in \mathbb{T}^2. \tag{1·15}$$

However since $\mathbf{M}(Z)\, Z_\tau$ and $\mathbf{K}(Z)\, Z_\xi$ are gradients of functionals A and B_1 respectively (averaged over $\mathbb{T}^2$) the equation (1·15) is formally the Lagrange necessary condition for the constrained variational principle

$$\text{crit}\,(\mathscr{S})\,|_{\mathscr{A} = I_1,\, \mathscr{B}_1 = I_2} \tag{1·16}$$

https://doi.org/10.1017/S0305004196001429 Published online by Cambridge University Press

in which case ω and k are Lagrange multipliers. The generalization of this variational principle to two-space dimensions is

$$\text{crit}\,(\mathscr{S})\,|_{\mathscr{A}=I_1,\,\mathscr{B}_1=I_2,\,\mathscr{B}_2=I_3} \tag{1·17}$$

with ω, k and l as Lagrange multipliers, where l is the wavenumber in the y-direction. The variational principles given by (1·15)–(1·17) are analogous to the variational principles for invariant toroids of finite-dimensional Hamiltonian systems (cf. Percival[**30**]).

It is an elementary consequence of the Lagrange multiplier theory (cf. Section 2) that

$$\omega = \frac{\partial S}{\partial I_1}, \quad k = \frac{\partial S}{\partial I_2}, \quad \text{and} \quad l = \frac{\partial S}{\partial I_3};$$

that is, the frequency and wavenumbers are given by certain slopes in S versus action space. The above identities also lead to

$$\text{Hess}_I(\mathscr{S}) = \begin{pmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{pmatrix} = \begin{pmatrix} \dfrac{\partial\omega}{\partial I_1} & \dfrac{\partial\omega}{\partial I_2} & \dfrac{\partial\omega}{\partial I_3} \\ \dfrac{\partial k}{\partial I_1} & \dfrac{\partial k}{\partial I_2} & \dfrac{\partial k}{\partial I_3} \\ \dfrac{\partial l}{\partial I_1} & \dfrac{\partial l}{\partial I_2} & \dfrac{\partial l}{\partial I_3} \end{pmatrix} \tag{1·18}$$

where $S_{ij} = \partial^2 S/\partial I_i\,\partial I_j$ for $i,j = 1,2,3$. It is shown in Section 2 that a necessary condition for the variational principle for waves to be non-degenerate is that

$$\Delta_S \overset{\text{def}}{=} \det\,(\text{Hess}_I(\mathscr{S})) \neq 0.$$

It is interesting that this non-degeneracy condition is reminiscent of the non-degeneracy condition for invariant tori in the KAM theory (cf. Percival[**30**]).

In Section 3 the properties of the Whitham modulation equations for periodic travelling waves in two space dimensions are recalled and they are cast into a form for comparison with the variational principle of Section 2 and the rigorous theory of Section 4. The important point here is to indicate precisely what one can prove using the methods of this paper about the Whitham modulation equations. In particular a proof is given of the validity of the linear instability results predicted by the Whitham theory. By checking the hypotheses of Theorem 4·1 in Section 4 it is clear also that a wave state predicted by the Whitham modulation equations to be stable is not necessarily stable in the full system from which the equations were derived.

In Section 4 a rigorous theory for the instability of periodic waves of arbitrary amplitude travelling in two space dimensions is presented. In the stability analysis the existence of a basic state is taken as a hypothesis and it is of the form

$$\{\hat{Z}(\theta;I),\,\omega(I),\,k(I),\,l(I);\,\theta = kx+ly-\omega t,\,I\in\mathscr{U}\subset\mathbb{R}^3\} \tag{1·19}$$

and satisfies (1·12). The dependence on the parameters I_1, I_2 and I_3 (level sets of the action constraint sets) is assured by the variational principle of Section 2. The main result of Section 4 is a proof of the relation between the stability exponent of the system (1·12) linearised about the basic state (1·19) and the fundamental matrix

https://doi.org/10.1017/S0305004196001429 Published online by Cambridge University Press

$\mathrm{Hess}_I(\mathscr{S})$ of the basic state. The result of Theorem 4·1 is precisely a proof of the *instability* result predicted by the Whitham modulation equation. The proof however does not appeal to the conservation of wave action nor the Whitham modulation equations but works directly with the governing equations appealing instead to the geometry presented by the multi-sympletic structures framework. Theorem 4·1 also connects explicitly the linear stability exponents and the fundamental matrix $\mathrm{Hess}_I(\mathscr{S})$ in (1·18). In terms of $\mathrm{Hess}_I(\mathscr{S})$ the instability result is

- if $\mathrm{Hess}_I(\mathscr{S})$ is positive definite the wave is linearly unstable;
- if $S_{22}S_{33}-S_{23}S_{32}<0$ the wave is linearly unstable;
- if $S_{22}S_{33}-S_{23}S_{32}>0$ but $S_{33}\Delta_K>0$ the wave is linearly unstable.

The proof of instability for waves travelling in one space dimension follows similar lines – indeed is a special case – but has some interesting features that are treated separately in Section 5. In particular a proof of Lighthill's geometric criterion is given and moreover the proof is obtained by first showing that there is a dual criterion for instability in terms of the geometry of the flow force on action space.

In Section 6 two examples are treated, the Nonlinear Schrödinger (NLS) equation and the water-wave problem on an infinite-depth fluid both for the case of two space dimensions. The purpose of Section 6 is to illustrate how the theory of the present paper applies to wave instability problems that are generally well understood. For the NLS equation the hypotheses of Theorem 4·1 in Section 4 can be explicitly verified and application of the theory to NLS leads to interesting geometric properties of the instability. For the water-wave problem the application of the theory in Section 6 is formal; this example is a precursor to more complete results that we presented elsewhere. A complete proof of instability for water waves travelling in one space dimension has been given in Bridges & Mielke [**13**]. Also in a companion paper (Bridges [**11**]) complete details of the application of the present theory to the instability problem for periodic patterns on the ocean surface affected by mean-flow dynamics is given.

Some auxiliary results are presented in the Appendices. A fundamental concept in the theory of classical Hamiltonian systems is the interplay between reversibility and symplecticity. The generalization to the multi-symplectic framework is presented in Appendix A.

An advantage of the multi-symplectic structures framework is that it decomposes neatly the different facets of the governing equation. Appendix B shows how the impulse and action are simply related in the framework leading to a useful decomposition of the impulse and energy conservation laws.

When a conservative system has a Hamiltonian structure the symplectic operator gives a natural correspondence between symmetries and invariants (Olver [**28**, **29**]). But the classical theory does not establish a relation between symplecticity and fluxes. Such a connection is only possible by introducing additional sympletic operators. In Appendix C a new decomposition of the Noether theory is given using the multi-sympletic structures theory. Such a decomposition is of fundamental importance when analysing the stability of waves in the presence of symmetry or mean-flow effects (cf. Bridges [**11**]). Finally, in Appendix D some special properties of the Hamiltonian functional on a multi-symplectic structure are presented.

https://doi.org/10.1017/S0305004196001429 Published online by Cambridge University Press

2. *Toral geometry and variational principles for space-time periodic states*

The restriction of the generalized Hamiltonian structure (1·11) or (1·12) to $\mathbb{T}^n$; that is, functions that are periodic in n-directions leads to a new constrained variational principle for space and time periodic waves. Here we consider the wave states that are periodic in 2-space dimensions and time.

Scale x, y and t so that the system (1·12) is restricted to 2π-periodic functions

$$\tau = \omega t, \quad \xi = kx, \quad \zeta = ly \quad \text{with} \quad (\xi, \zeta, \tau) \in \mathbb{T}^3.$$

Then (1·12) becomes

$$\omega \mathbf{M}(Z) Z_\tau + k\mathbf{K}(Z) Z_\xi + l\mathbf{L}(Z) Z_\zeta = \nabla S(Z) \quad \text{with} \quad Z\colon \mathbb{T}^3 \to \mathscr{M}. \tag{2·1}$$

Because of the hypothesis that $\mathbf{M}(Z)$, $\mathbf{K}(Z)$ and $\mathbf{L}(Z)$ are associated with closed two forms, the terms $\mathbf{M}(Z) Z_\tau$, $\mathbf{K}(Z) Z_\xi$ and $\mathbf{L}(Z) Z_\zeta$ can be identified with the gradients of certain functionals,

$$\begin{aligned} \nabla A(Z) &= \mathbf{M}(Z) Z_\tau \\ \nabla B_1(Z) &= \mathbf{K}(Z) Z_\xi \\ \nabla B_2(Z) &= \mathbf{L}(Z) Z_\zeta, \end{aligned}$$

where here the gradient is defined with respect to an inner product that includes integration over $\mathbb{T}^3$. Therefore (2·1) can be written

$$\nabla S(Z) = \omega \nabla A(Z) + k \nabla B_1(Z) + l \nabla B_2(Z) \tag{2·2}$$

which can equivalently be characterized in terms of a constrained variational principle.

Variational principle for space-time periodic states. Space-time periodic states of the Hamiltonian system (1·12) on the multi-symplectic structure $(\mathscr{M}, \omega_m^{(1)}, \omega_m^{(2)}, \omega_m^{(3)}, S)$ correspond to the critical points: $\hat{Z}\colon \mathbb{T}^3 \to \mathscr{M}$ of S restricted to level sets of the action functionals $\mathscr{A}$, $\mathscr{B}_1$ and $\mathscr{B}_2$ or

$$\operatorname{crit}\{S\colon \mathscr{A} = I_1, \mathscr{B}_1 = I_2 \text{ and } \mathscr{B}_2 = I_3\}, \tag{2·3}$$

where $I \in \mathscr{U} \subset \mathbb{R}^3$ are values of the action level sets.

The symbols $\mathscr{S}$, $\mathscr{A}$, $\mathscr{B}_1$ and $\mathscr{B}_2$ correspond to S, A, B_1 and B_2, respectively, averaged over $\mathbb{T}^3$. The verification of the above variational principle follows from the Lagrange multiplier theory. The statement is formal; application to particular examples may encounter technical difficulties such as small divisors. On the other hand it leads to interesting structural properties of the solutions including linear stability information (cf. Section 4).

An interesting property of the above variational principle is its relation with the variational principle for invariant toroids of finite-dimensional Hamiltonian systems. This correspondence has been discussed in Bridges [**9**] for plane waves on one-space dimension in the context of the nonlinear Schrödinger equation. However, the connection with invariant tori is more general and follows from the multi-symplectic structure.

For example, consider the finite-dimensional Hamiltonian system

$$\mathbf{J} U_t = \nabla H(U) \quad \text{with} \quad U \in \mathscr{M} \subset \mathbb{R}^6 \tag{2·4}$$

https://doi.org/10.1017/S0305004196001429 Published online by Cambridge University Press

and $\mathbf{J}$ the standard symplectic operator on $\mathbb{R}^6$. Restrict (2·4) to an invariant torus by taking coordinates $\xi = (\xi_1, \xi_2, \xi_3) \in \mathbb{T}^3$ with $\xi_j = \omega_j t$ for $j = 1, 2, 3$. Then (2·4) restricted to $\mathbb{T}^3$ becomes

$$\omega_1 \mathbf{J}\frac{\partial U}{\partial \xi_1} + \omega_2 \mathbf{J}\frac{\partial U}{\partial \xi_2} + \omega_3 \mathbf{J}\frac{\partial U}{\partial \xi_3} = \nabla H(U). \tag{2·5}$$

The most significant difference between (2·1) and (2·5) is the symplectic operators. In (2·5) the symplectic operators commute, indeed are the same, whereas in (2·1) the symplectic operators are different and in general non-commutative. Introduce actions for the tori satisfying (2·5),

$$I_j = \oint_{\mathbb{T}^3} \tfrac{1}{2}\langle U, \mathbf{J}\partial_{\xi_j} U\rangle_{\mathbb{R}^6}\, d\xi \quad j = 1, 2, 3.$$

It is evident that, formally, (2·5) is the Lagrange necessary condition for the variational principle of finding critical points of the Hamiltonian on level sets of the actions with ω_j, $j = 1, 2, 3$, treated as Lagrange multipliers; this is in fact Percival's first variational principle for invariant tori of fixed action (cf. Percival [**30**]). The relation of the variational principle (2·3) with invariant toroids is further emphasized below where it is shown that the variational principle (2·3) is non-degenerate if the frequency map is non-degenerate which is reminiscent of the non-degeneracy condition for invariant toroids of finite-dimensional Hamiltonian systems.

The Lagrange necessary condition for the variational principle (2·3) is that $Z \in \mathscr{M}$ should satisfy

$$\nabla \mathscr{F} = 0 \quad \text{with} \quad \mathscr{F}(Z; \omega, k, l) = \mathscr{S}(Z) - \omega \mathscr{A}(Z) - k\mathscr{B}_1(Z) - l\mathscr{B}_2(Z) \tag{2·6}$$

as well as the three constraints. A basic solution of the variational principle depends on $I \in \mathbb{R}^3$ in the following way

$$\{\hat{Z}(\xi, \zeta, \tau; I),\, \omega(I),\, k(I),\, l(I);\, I = (I_1, I_2, I_3) \in \mathscr{U} \subset \mathbb{R}^3\}. \tag{2·7}$$

This is verified as follows. A solution of (2·6) has the form $Z(\xi, \zeta, \tau; \omega, k, l)$; but substitution of this expression into the constraint set results in

$$\left.\begin{aligned} \mathscr{A}(Z(\xi, \zeta, \tau; \omega, k, l)) &= I_1 \\ \mathscr{B}_1(Z(\xi, \zeta, \tau; \omega, k, l)) &= I_2 \\ \mathscr{B}_2(Z(\xi, \zeta, \tau; \omega, k, l)) &= I_3 \end{aligned}\right\} \tag{2·8}$$

which when inverted leads to $\omega(I)$, $k(I)$ and $l(I)$. Back substitution into $Z(\xi, \zeta, \tau; \omega, k, l)$ leads to the dependence in (2·7). Note however that (2·8) is solvable, by an implicit function theorem argument, if and only if $\det(\mathbf{D}) \neq 0$ where

$$\mathbf{D} = \begin{pmatrix} \dfrac{\partial \omega}{\partial I_1} & \dfrac{\partial \omega}{\partial I_2} & \dfrac{\partial \omega}{\partial I_3} \\ \dfrac{\partial k}{\partial I_1} & \dfrac{\partial k}{\partial I_2} & \dfrac{\partial k}{\partial I_3} \\ \dfrac{\partial l}{\partial I_1} & \dfrac{\partial l}{\partial I_2} & \dfrac{\partial l}{\partial I_3} \end{pmatrix}. \tag{2·9}$$

In the analysis of Section 4 it will be shown that the matrix (2·9) contains linear stability information. It is also reminiscent of the frequeny map on action space in the KAM theory. The variational principle (2·3) is non-degenerate if $\det(\mathbf{D}) \neq 0$. The

https://doi.org/10.1017/S0305004196001429 Published online by Cambridge University Press

matrix $\mathbf{D}$ is symmetric. The symmetry of $\mathbf{D}$ can be seen by using the following identities

$$\omega(I) = \frac{\partial \mathscr{S}}{\partial I_1}, \quad k(I) = \frac{\partial \mathscr{S}}{\partial I_2} \quad \text{and} \quad l(I) = \frac{\partial \mathscr{S}}{\partial I_3}. \tag{2·10}$$

These identities follow from the classical Lagrange multiplier theory and are verified as follows: for $j = 1, 2, 3$ we have that

$$\begin{aligned}
\frac{\partial \mathscr{S}}{\partial I_j} &= \left[\nabla S(Z), \frac{\partial Z}{\partial I_j}\right] \\
&= \left[\omega \nabla A(Z) + k\nabla B_1(Z) + l\nabla B_2(Z), \frac{\partial Z}{\partial I_j}\right] \\
&= \omega \frac{\partial \mathscr{A}}{\partial I_j} + k\frac{\partial \mathscr{B}_1}{\partial I_j} + l\frac{\partial \mathscr{B}_2}{\partial I_j} \\
&= \omega\delta_{1j} + k\delta_{2j} + l\delta_{3j}.
\end{aligned}$$

Using the identities (2·10) it follows that the matrix $\mathbf{D}$ can also be represented as

$$\mathbf{D} = \mathrm{Hess}_I(\mathscr{S}) = \begin{pmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{pmatrix} \quad \text{where} \quad S_{ij} \overset{\text{def}}{=} \frac{\partial^2 \mathscr{S}}{\partial I_i \partial I_j}. \tag{2·11}$$

A sensitivity matrix of the form (2·9) carries information about critical point type for constrained variational principles with a well-defined Morse index (cf. Maddocks [**24**]). However critical point type for solutions of the variational principle (2·3) is not well-defined. The second variation of $\mathscr{F}$ in (2·6) defined by

$$\mathscr{L} = D^2\mathscr{F}(Z; \omega, k, l) = D^2S(Z) - \omega D^2A(Z) - kD^2B_1(Z) - lD^2B_2(Z)$$

has a countable number of negative as well as positive eigenvalues. This can be easily seen already in the case when space variation is absent (i.e. $\mathscr{F} = \mathscr{S}(Z) - \omega\mathscr{A}(Z)$) which is just Hamilton's principle for periodic orbits of classical Hamiltonian systems. Technically the critical points are saddle points of infinite order. However we will be able to extract useful information about linear stability from the constrained variational principle and sensitivity matrix by working directly with the linear stability problem (Section 4).

Finally we remark that the variational principle (2·3) is for nonlinear states. It has no meaning for the linear problem – it is trivially degenerate. In the linear problem the frequency and wavenumbers are related by the dispersion relation and therefore the columns of the matrix $\mathbf{D}$ in (2·9) will be linearly dependent resulting in degeneracy.

3. *Structure of Whitham's modulation equations*

The Whitham theory for modulated travelling waves in two space dimensions, in the absence of mean flow effects, is based on the equations

$$\begin{aligned}
k_y - l_x &= 0 \\
k_t + \omega_x &= 0 \\
l_t + \omega_y &= 0 \\
(\mathscr{A})_t - (\mathscr{B}_1)_x - (\mathscr{B}_2)_y &= 0.
\end{aligned} \tag{3·1}$$

https://doi.org/10.1017/S0305004196001429 Published online by Cambridge University Press

The first three of these are kinematic conditions on the phase θ and

$$\omega = -\theta_t, \quad k = \theta_x \quad \text{and} \quad l = \theta_y \tag{3·2}$$

are interpreted as slowly varying functions of x, y and t. The fourth equation of (3·1) is the hypothesis of conservation of wave action where $\mathscr{A}$, $\mathscr{B}_1$ and $\mathscr{B}_2$ are averages of the action and action fluxes respectively that are considered as slowly varying functions of x, y and t (cf. Whitham [**34**, **35**] and Hayes [**18**, **19**]).

In this section we will first make some observations about the Hamiltonian structure of the equations (3·1). Then the equations will be cast into a form where the instability of waves predicted by (3·1) can be compared with the information generated by the variational principle of Section 2 and the rigorous theory to be presented in Section 4.

First we remark that the first three equations of (3·1) correspond to a Hamiltonian system; they can be rewritten in the form

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} \omega \\ k \\ l \end{pmatrix} = \begin{pmatrix} k_y - l_x \\ -\omega_y \\ \omega_x \end{pmatrix} \tag{3·3}$$

or $$\mathbf{J}_1 \Omega_t = \nabla H(\Omega), \quad \text{where} \quad \Omega = \begin{pmatrix} \omega \\ k \\ l \end{pmatrix}, \quad \mathbf{J}_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}$$

and $$H(\Omega) = \int_{y_1}^{y_2} \int_{x_1}^{x_2} \omega(k_y - l_x)\, dx\, dy.$$

With suitable fixed endpoint conditions for Ω on the boundary of the rectangle $[x_1, x_2] \times [y_1, y_2]$ the gradient of $H(\Omega)$ is

$$\nabla H(\Omega) \overset{\text{def}}{=} \begin{pmatrix} \delta H/\delta\omega \\ \delta H/\delta k \\ \delta H/\delta l \end{pmatrix} = \begin{pmatrix} k_y - l_x \\ -\omega_y \\ \omega_x \end{pmatrix}.$$

The governing equation for Ω has another interpretation as a Hamiltonian system on a multi-symplectic structure. First note that

$$\nabla H(\Omega) = \begin{pmatrix} k_y - l_x \\ -\omega_y \\ \omega_x \end{pmatrix} = \begin{pmatrix} -l_x \\ 0 \\ \omega_x \end{pmatrix} + \begin{pmatrix} k_y \\ -\omega_y \end{pmatrix} = -\mathbf{J}_2 \Omega_x - \mathbf{J}_3 \Omega_y,$$

where $$\mathbf{J}_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} \quad \text{and} \quad \mathbf{J}_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Therefore the set (3·1) has the equivalent representation

$$\left.\begin{aligned} \mathbf{J}_1 \Omega_t + \mathbf{J}_2 \Omega_x + \mathbf{J}_3 \Omega_y &= 0 \\ (\mathscr{A})_t - (\mathscr{B}_1)_x - (\mathscr{B}_2)_y &= 0. \end{aligned}\right\} \tag{3·4}$$

https://doi.org/10.1017/S0305004196001429 Published online by Cambridge University Press

The governing equation for Ω is a Hamiltonian system on a multi-symplectic structure: $(\mathbb{R}^3, \omega^{(1)}, \omega^{(2)}, \omega^{(3)})$ with

$$\omega^{(1)} = dk \wedge dl, \quad \omega^{(2)} = d\omega \wedge dl \quad \text{and} \quad \omega^{(3)} = dk \wedge d\omega$$

with trivial Hamiltonian function. Moreover it is interesting to note that the symplectic forms $\omega^{(1)}$, $\omega^{(2)}$ and $\omega^{(3)}$ form a basis for the Lie algebra $so(3)$.

In light of the variational principle of Section 2, a reasonable hypothesis is to suppose that ω, k and l depend only on the values of the actions $\mathscr{A}$, $\mathscr{B}_1$ and $\mathscr{B}_2$ and do not otherwise depend explicitly on x, y and t. Therefore define

$$\mathbf{D} = \begin{pmatrix} \nabla\omega \\ \nabla k \\ \nabla l \end{pmatrix}$$

where ∇ is a gradient on action space. Then

$$\Omega_t = \mathbf{D}\mathscr{I}_t \quad \text{where} \quad \mathscr{I} = \begin{pmatrix} \mathscr{A} \\ \mathscr{B}_1 \\ \mathscr{B}_2 \end{pmatrix}$$

with similar expressions for Ω_x and Ω_y. On action space the coupled equations (3·4) then become

$$\begin{aligned} \mathbf{J}_1\mathbf{D}\mathscr{I}_t + \mathbf{J}_2\mathbf{D}\mathscr{I}_x + \mathbf{J}_3\mathbf{D}\mathscr{I}_y = 0 \\ \langle e_1, \mathscr{I}_t\rangle - \langle e_2, \mathscr{I}_x\rangle - \langle e_3, \mathscr{I}_y\rangle = 0 \end{aligned} \tag{3·5}$$

where e_1, e_2 and e_3 are unit vectors on $\mathbb{R}^3$ and here $\langle\cdot,\cdot\rangle$ represents the standard inner product on $\mathbb{R}^3$ (compare with Hayes[**19**, equations (3·1)–(3·2)]).

Let $\mathscr{I}_0$ be some fixed constant value on action space representative of a periodic travelling wave in two space dimensions and linearize (3·5) about this state with

$$\mathscr{I} = \mathscr{I}_0 + \hat{\mathscr{I}} e^{i(\alpha x + \beta y - \lambda t)}$$

where $(\alpha, \beta) \in \mathbb{R}^2$ and $\lambda \in \mathbb{C}$. Then the linearization of (3·5) about $\mathscr{I}_0$ becomes

$$\left.\begin{aligned} [-\lambda\mathbf{J}_1 + \alpha\mathbf{J}_2 + \beta\mathbf{J}_3]\mathbf{D}\hat{\mathscr{I}} = 0 \\ \langle \Gamma, \hat{\mathscr{I}}\rangle = 0 \end{aligned}\right\} \tag{3·6}$$

where $\Gamma = \lambda e_1 + \alpha e_e + \beta e_3$ and $\mathbf{D}$ is now evaluated at $\mathscr{I}_0$. Supposing $|\lambda|^2 + \alpha^2 + \beta^2 \neq 0$,

$$\text{Kernel}\,[-\lambda\mathbf{J}_1 + \alpha\mathbf{J}_2 + \beta\mathbf{J}_3] = \text{span}\,\{\Gamma\}.$$

Therefore the general solution of the first equation of (3·6) is

$$\mathbf{D}\hat{\mathscr{I}} \in \text{Kernel}\,[-\lambda\mathbf{J}_1 + \alpha\mathbf{J}_2 + \beta\mathbf{J}_3] \Rightarrow \hat{\mathscr{I}} = c\mathbf{D}^{-1}\Gamma$$

for some non-zero scalar c. Substitution into the second equation of (3·6) results in

$$c\langle \Gamma, \mathbf{D}^{-1}\Gamma\rangle = 0.$$

Therefore define

$$\Lambda_2(\lambda, \alpha, \beta) = \langle \Gamma, \mathbf{D}^{-1}\Gamma\rangle = \begin{pmatrix} \lambda \\ \alpha \\ \beta \end{pmatrix}, \begin{pmatrix} \omega_1 & \omega_2 & \omega_3 \\ k_1 & k_2 & k_3 \\ l_1 & l_2 & l_3 \end{pmatrix}^{-1} \begin{pmatrix} \lambda \\ \alpha \\ \beta \end{pmatrix}. \tag{3·7}$$

https://doi.org/10.1017/S0305004196001429 Published online by Cambridge University Press

The interpretation of (3·7) in terms of (3·5) is as follows. If for some $(\alpha, \beta) \in \mathbb{R}^2$, there exists a $\lambda \in \mathbb{C}$ with $\Im(\lambda) \neq 0$ and $\Lambda(\lambda, \alpha, \beta) = 0$ there exists an unstable solution near $\mathscr{I}_0$ on action space.

This completes the reappraisal of the Whitham modulation equations. In Section 4 we will give a proof of the above instability criterion and the proof is given without direct recourse to the modulation equations or the conservation of wave action.

4. *A rigorous geometric theory for plane wave instability*

In this section the linear instability problem for periodic travelling waves of the system

$$\mathbf{M}(Z)\, Z_t + \mathbf{K}(Z)\, Z_x + \mathbf{L}(Z)\, Z_y = \nabla S(Z), \tag{4·1}$$

where $\mathbf{M}(Z)$, $\mathbf{K}(Z)$, $\mathbf{L}(Z)$ and S are identified with a multi-symplectic structure and the skew-symmetric operators are in general Z-dependent, is formulated rigorously. The basic state is a travelling plane wave of the form $\hat{Z}(\theta, \cdot)$ with $\theta = kx + ly - \omega t$. A connection is established between the stability exponent in the time-dependent problem and the 'KAM matrix' $\mathrm{Hess}_I(S)$ that arises in the variational principle of Section 2. The instability criterion is established independent of the particular equation and relies only on the geometry of the multi-symplectic structure.

Introduce new coordinates

$$\theta = kx + ly - \omega t, \quad y' = y \quad \text{and} \quad t' = t;$$

in terms of which equation (4·1) becomes

$$\mathbf{M}(Z)\, Z_{t'} + \mathbf{L}(Z)\, Z_{y'} + \mathbf{J}(Z)\, Z_\theta = \nabla S(Z) \tag{4·2}$$

where, for simplicity the primes on y, t will henceforth be dropped, and

$$\mathbf{J}(Z) = k\mathbf{K}(Z) + l\mathbf{L}(Z) - \omega \mathbf{M}(Z). \tag{4·3}$$

It is taken as a hypothesis that there exists a basic travelling wave state of the form

(**H1**) $$\{\hat{Z}(\theta; I),\, \omega(I),\, k(I),\, l(I);\, I \in \mathscr{U} \subset \mathbb{R}^3\}$$

and that ω, k and l are sufficiently smooth so that the matrix $\mathbf{D}$ in (2·9), evaluated at the basic state, is well defined and non-degenerate and $\hat{Z}$ is sufficiently differentiable. It follows from (4·2) that $\hat{Z}$ satisfies

$$\mathbf{J}(\hat{Z})\, \hat{Z}_\theta = \nabla S(\hat{Z}). \tag{4·4}$$

To study the linear stability of the state (H1) let $Z = \hat{Z}(\theta; I) + \tilde{Z}(\theta, y, t)$ and linearize (4·2) about the state $\hat{Z}$ resulting in

$$\mathscr{L}\tilde{Z} = \mathbf{M}(\hat{Z})\, \tilde{Z}_t + \mathbf{L}(\hat{Z})\, \tilde{Z}_y, \tag{4·5}$$

where $$\mathscr{L} \overset{\text{def}}{=} D^2 S(\hat{Z}) - \mathbf{J}(\hat{Z}) \frac{\partial}{\partial \theta} - \{d\mathbf{J}(\hat{Z}), \cdot\}\, \hat{Z}_\theta.$$

The coefficients of the linear partial differential equation (4·5) are independent of y and t. Therefore let $\tilde{Z} = \Re(\tilde{U} e^{i(\beta y + \lambda t)})$ with $\tilde{U}$ satisfying

$$\mathscr{L}\tilde{U} = i\lambda \mathbf{M}(\hat{Z})\, \tilde{U} + i\beta \mathbf{L}(\hat{Z})\, \tilde{U}. \tag{4·6}$$

https://doi.org/10.1017/S0305004196001429 Published online by Cambridge University Press

In terms of (4·6) the linear instability problem can be stated as follows.

Linear instability. Suppose that for some $\beta \in \mathbb{R}$ and $\lambda \in \mathbb{C}$ such that $\Im(\lambda) \neq 0$ there exists a bounded solution $\tilde{U}$ of (4·6). Then the basic state $\hat{Z}$ is unstable.

In order to give a rigorous formulation of the instability problem it is necessary to characterize the bounded states of (4·6). This is straightforward if Floquet's theorem is applicable. Therefore we take the following hypothesis.

(**H2**) $\begin{cases} \text{there exists a bounded solution of (4·6) of the form } \tilde{U} = Ue^{i\alpha\theta} \\ \text{with } \alpha \in \mathbb{R} \text{ and } U \text{ } 2\pi\text{-periodic in } \theta. \end{cases}$

First we note that when the system (4·6) is an ODE Floquet's theorem is immediate and the hypothesis is unnecessary. The hypothesis (H2) is necessary for the case when $\mathscr{L}$ is an elliptic operator because there exists counterexamples (albeit pathological) to Floquet's theorem for elliptic operators (cf. Kuchment [**20**]). In the sequel we are not interested in studying all solutions of (4·5); the idea is to demonstrate, under the given hypotheses, the existence of an unstable state. The governing equation for the periodic function U is

$$\mathscr{L}U = i\lambda \mathbf{M}(\hat{Z})\, U + i\alpha \mathbf{J}(\hat{Z})\, U + i\beta \mathbf{L}(\hat{Z})\, U. \tag{4·7}$$

In terms of (4·7) the instability condition is: suppose for some $(\alpha, \beta) \in \mathbb{R}^2$ and $\lambda \in \mathbb{C}$ such that $\Im(\lambda) \neq 0$ there exists a periodic function U satisfying (4·7). Then the basic state (H1) is linearly unstable.

To give a rigorous formulation the eigenvalue problem (4·7) is first written as an operator equation on a suitable function space. Let $\mathscr{Y}$ and $\mathscr{Z}$ be Hilbert spaces such that $\mathbf{M}(\hat{Z})$, $\mathbf{K}(\hat{Z})$, $\mathbf{L}(\hat{Z})$ and D^2S are bounded linear operators from $\mathscr{Y}$ into $\mathscr{Z}$ and define

$$\mathbb{X}_1 = \mathscr{C}^1(\mathbb{S}^1, \mathscr{Y})$$
$$\mathbb{X}_0 = \mathscr{C}^0(\mathbb{S}^1, \mathscr{Z}).$$

Define the linear operator $\Psi\colon \mathbb{X}_1 \times \mathbb{C} \times \mathbb{R} \times \mathbb{R} \to \mathbb{X}_0$ by

$$\Psi(U, \lambda, \alpha, \beta) = \mathscr{L}U - i\lambda \mathbf{M}(\hat{Z})\, U - i\alpha \mathbf{J}(\hat{Z})\, U - i\beta \mathbf{L}(\hat{Z})\, U. \tag{4·8}$$

For functions $a, b \in \mathbb{X}_0$ define the bilinear pairing

$$[a, b] = \frac{1}{2\pi} \int_0^{2\pi} \langle \cdot, \cdot \rangle_m \, d\theta.$$

It is clear that the tangent vector to the basic state, defined by

$$\phi_0 = \frac{\partial}{\partial \theta} \hat{Z}(\theta; I),$$

is in the kernel of $\mathscr{L}$. This is verified by differentiating (4·4) with respect to θ.

(**H3**) $$\operatorname{Ker}(\mathscr{L})|_{\mathbb{X}_1} = \{\phi_0\}.$$

The purpose of (H3) is to ensure that the kernel of $\mathscr{L}$ is not larger. When a continuous symmetry is present, the hypothesis (H3) is violated and this case is treated by including the tangent vectors to the group orbit in the kernel of $\mathscr{L}$. This occurs when

https://doi.org/10.1017/S0305004196001429 Published online by Cambridge University Press

studying instabilities in the water-wave problem when mean-flow effects are important for example (cf. Bridges[**11**]). The hypothesis (H3) is also a genericity hypothesis; at distinguished values of the parameters the kernel can be larger.

The idea is to apply the Lyapunov–Schmidt method to the operator equation (4·8) (cf. Bridges & Rowlands[**14**] for a similar analysis for the instability of spatially quasiperiodic states). Decompose the spaces $\mathbb{X}_0$ and $\mathbb{X}_1$ as

$$\begin{aligned}\mathbb{X}_1 &= \mathrm{Ker}\,(\mathscr{L}) \oplus M \quad \text{where} \quad M = \mathrm{range}\,(\mathscr{L}) \cap \mathbb{X}_1\\ \mathbb{X}_0 &= \mathrm{range}\,(\mathscr{L}) \oplus \ker(\mathscr{L})\end{aligned}$$

and introduce the projection operator $\mathbf{P}\colon \mathbb{X}_0 \to \ker(\mathscr{L})$,

$$\mathbf{P}f \stackrel{\mathrm{def}}{=} \frac{[\phi_0, f]\,\phi_0}{[\phi_0, \phi_0]}. \tag{4·9}$$

Then any element in $\mathbb{X}_1$ has the representation

$$U = \mathbb{C}(\phi_0 + W(\theta;\lambda,\alpha,\beta)) \quad \text{with} \quad W(\theta;0,0,0) = 0. \tag{4·10}$$

Without loss of generality the complex multiplier in (4·10) can be taken to be unity. The operator equation (4·8) can be decomposed as

$$\begin{aligned}\mathbf{P}\cdot\Psi(\phi_0 + W(\theta;\lambda,\alpha,\beta),\lambda,\alpha,\beta) &= 0\\ (\mathbf{I}-\mathbf{P})\cdot\Psi(\phi_0 + W(\theta;\lambda,\alpha,\beta),\lambda,\alpha,\beta) &= 0.\end{aligned}$$

Theorem 4·1. *Under the hypotheses* (*H*1)–(*H*3) *every solution of* (4·6) *for* $|\lambda|$, $|\alpha|$ *and* $|\beta|$ *sufficiently small is in one-to-one correspondence with roots of the 'dispersion relation'*

$$\Lambda(\lambda,\alpha,\beta) = [\phi_0, \Psi(\phi_0 + W(\theta;\lambda,\alpha,\beta),\lambda,\alpha,\beta)] = 0.$$

Moreover

$$\Lambda(\lambda,\alpha,\beta) = \left\langle \begin{pmatrix} -\lambda+\alpha\omega \\ \alpha k \\ \beta+\alpha l \end{pmatrix}, \begin{bmatrix} \dfrac{\partial\omega}{\partial I_1} & \dfrac{\partial\omega}{\partial I_2} & \dfrac{\partial\omega}{\partial I_3} \\ \dfrac{\partial k}{\partial I_1} & \dfrac{\partial k}{\partial I_2} & \dfrac{\partial k}{\partial I_3} \\ \dfrac{\partial l}{\partial I_1} & \dfrac{\partial l}{\partial I_2} & \dfrac{\partial l}{\partial I_3} \end{bmatrix}^{-1} \begin{pmatrix} -\lambda+\alpha\omega \\ \alpha k \\ \beta+\alpha l \end{pmatrix} \right\rangle + r_3(\lambda,\alpha,\beta) \tag{4·11}$$

where $|r_3| = o(|\lambda|,|\alpha|,|\beta|)^2$ *as* $(|\lambda|,|\alpha|,|\beta|) \to 0$.

Before beginning the proof of Theorem 4·1 we establish its corollaries which connect the linear instability to the properties of the matrix $\mathrm{Hess}_I(S)$.

Corollary 4·2. *Let* (*H*1) *be a periodic travelling wave solution of* (4·1) *and suppose moreover that*

$$\Delta_S \neq 0 \quad \textit{and} \quad \det\begin{pmatrix} S_{22} & S_{23} \\ S_{32} & S_{33} \end{pmatrix} \neq 0.$$

(*a*) *If* $\det\begin{pmatrix} S_{22} & S_{23} \\ S_{32} & S_{33} \end{pmatrix} < 0$ (*H*1) *is linearly unstable.*

(*b*) *If* $\det\begin{pmatrix} S_{22} & S_{23} \\ S_{32} & S_{33} \end{pmatrix} > 0$ *and* $S_{33}\Delta_S > 0$ (*H*1) *is linearly unstable.*

https://doi.org/10.1017/S0305004196001429 Published online by Cambridge University Press

Proof. By Theorem 4·1 the stability exponents, for $|\lambda|$, $|\alpha|$ and $|\beta|$ sufficiently small are given by $\Lambda(\lambda,\alpha,\beta)=0$. First use the identity

$$\begin{bmatrix} \frac{\partial\omega}{\partial I_1} & \frac{\partial\omega}{\partial I_2} & \frac{\partial\omega}{\partial I_3} \\ \frac{\partial k}{\partial I_1} & \frac{\partial k}{\partial I_2} & \frac{\partial k}{\partial I_3} \\ \frac{\partial l}{\partial I_1} & \frac{\partial l}{\partial I_2} & \frac{\partial l}{\partial I_3} \end{bmatrix}^{-1} = \begin{bmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{bmatrix}^{-1} \stackrel{\text{def}}{=} \begin{pmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \end{pmatrix} \tag{4·12}$$

and let $$\Omega=-\lambda+\alpha\omega, \quad p=\alpha k, \quad \text{and} \quad q=\beta+\alpha l.$$

Then $\Lambda(\lambda,\alpha,\beta)$ in (4·11) can be written as

$$\begin{aligned}\Lambda(\lambda,\alpha,\beta) &= d_{11}\Omega^2+2d_{12}\Omega p+2d_{13}\Omega q+d_{22}p^2+2d_{23}pq+d_{33}q^2+r_3\\ &= d_{11}\left(\Omega+\frac{d_{12}p+d_{13}q}{d_{11}}\right)^2+\left(\frac{d_{11}d_{22}-d_{12}^2}{d_{11}}\right)p^2\\ &\quad+2\left(\frac{d_{11}d_{23}-d_{12}d_{13}}{d_{11}}\right)pq+\left(\frac{d_{11}d_{33}-d_{13}^2}{d_{11}}\right)q^2+r_3.\end{aligned}$$

Now, from (4·12) we have the following identities

$$d_{11}=\frac{1}{\Delta_S}\det\begin{pmatrix} S_{22} & S_{23} \\ S_{32} & S_{33}\end{pmatrix}$$

$$d_{11}d_{22}-d_{12}^2=\frac{S_{33}}{\Delta_S}$$

$$d_{11}d_{23}-d_{12}d_{13}=-\frac{S_{32}}{\Delta_S}$$

$$d_{11}d_{33}-d_{13}^2=\frac{S_{22}}{\Delta_S}.$$

By hypothesis $\Delta_S\neq 0$ and $d_{11}\neq 0$. Therefore the dispersion relation can be written

$$\Lambda(\lambda,\alpha,\beta)=\frac{1}{d_{11}}\left[d_{11}^2\left(\Omega+\frac{d_{12}p+d_{13}q}{d_{11}}\right)^2+Q(q,p)+d_{11}r_3\right]$$

where $$Q(q,p)=\frac{1}{\Delta_S}\begin{pmatrix}p\\ q\end{pmatrix}\begin{pmatrix} S_{33} & -S_{32} \\ -S_{23} & S_{22}\end{pmatrix}\begin{pmatrix}p\\ q\end{pmatrix}.$$

Then for $|q|$ and $|p|$ sufficiently small $\Lambda=0$ results in

$$d_{11}\left(\Omega+\frac{d_{12}p+d_{13}q}{d_{11}}\right)=\pm\sqrt{[-Q(q,p)]}+\dots.$$

Therefore an unstable solution results (that is, $\Im(\lambda)\neq 0$ which is implied if $\Im(\Omega)\neq 0$) if, for some $(q,p)\in\mathbb{R}^2$ but $q^2+p^2\neq 0$ and sufficiently small $Q(p,q)>0$. An analysis of $Q(p,q)$ results in the conditions (a) and (b) stated.

COROLLARY 4·3. *Suppose* $\mathrm{Hess}_I(S)$ *is positive definite for the basic state* $(H1)$; *or equivalently, suppose* S *is a convex function on action space at* $(H1)$. *Then it is linearly unstable.*

https://doi.org/10.1017/S0305004196001429 Published online by Cambridge University Press

Proof. By hypothesis S is smooth enough so that $\mathrm{Hess}_I(S)$ is well-defined in which case convexity of S implies positivity of $\mathrm{Hess}_I(S)$. A necessary and sufficient condition for $\mathrm{Hess}_I(S)$ to be positive definite is

$$S_{33} > 0, \quad \begin{vmatrix} S_{22} & S_{23} \\ S_{32} & S_{33} \end{vmatrix} > 0 \quad \text{and} \quad \Delta_S > 0.$$

But by part (*b*) of Corollary 4·2 such a wave is unstable.

Proof of Theorem 4·1. The proof proceeds as follows. The technical points follow from the Lyapunov–Schmidt theory. The general solution of (4·7) is of the form (4·10) with

$$W = i\lambda W_1 + i\alpha W_2 + i\beta W_3 + r_2(\lambda, \alpha, \beta).$$

The remainder term $r_2 = o(|\lambda|, |\alpha|, |\beta|)$ and

$$W_1 = -i\frac{\partial W}{\partial \lambda}\bigg|_{\lambda=\alpha=\beta=0},$$

$$W_2 = -i\frac{\partial W}{\partial \alpha}\bigg|_{\lambda=\alpha=\beta=0},$$

$$W_3 = -i\frac{\partial W}{\partial \beta}\bigg|_{\lambda=\alpha=\beta=0}.$$

The implicit function theorem is applied to $(\mathbf{I}-\mathbf{P})\Psi = 0$ to prove the existence and smoothness (in λ, α and β) of W. The function W has a convergent power series in terms of (λ, α, β), for $(|\lambda|, |\alpha|, |\beta|)$ sufficiently small, with leading order expression as given above. A sufficient condition for linear instability will be obtained using only the leading terms W_1, W_2 and W_3. Substituting (4·10) into (4·7) results in the following inhomogeneous problems for W_j, $j = 1, 2, 3$:

$$\begin{aligned} \mathscr{L} W_1 &= \mathbf{M}(\hat{Z})\,\phi_0 \\ \mathscr{L} W_2 &= \mathbf{J}(\hat{Z})\,\phi_0 \\ \mathscr{L} W_3 &= \mathbf{L}(\hat{Z})\,\phi_0. \end{aligned} \tag{4·13}$$

A remarkable fact is that these equations can be solved exactly. To see this we first note that the basic state depends on I. Therefore define

$$\phi_j = \frac{\partial \hat{Z}}{\partial I_j} \quad j = 1, 2, 3$$

and note that these functions satisfy

$$\mathscr{L}\phi_j = \frac{\partial k}{\partial I_j}\mathbf{K}(\hat{Z})\,\phi_0 + \frac{\partial l}{\partial I_j}\mathbf{L}(\hat{Z})\,\phi_0 - \frac{\partial \omega}{\partial I_j}\mathbf{M}(\hat{Z})\,\phi_0 \quad \text{for} \quad j = 1, 2, 3 \tag{4·14}$$

(obtained by simply differentiating the governing equation (4·4) for $\hat{Z}$). The idea is then to represent solutions of (4·13) in terms of linear combinations of the complementary functions $\phi_j\, j = 1, 2, 3$. Let

$$\left.\begin{aligned} W_1 &= a_1\phi_1 + a_2\phi_2 + a_3\phi_3 \\ W_2 &= b_1\phi_1 + b_2\phi_2 + b_3\phi_3 \\ W_3 &= c_1\phi_1 + c_2\phi_2 + c_3\phi_3 \end{aligned}\right\}. \tag{4·15}$$

https://doi.org/10.1017/S0305004196001429 Published online by Cambridge University Press

Then substituting (4·15) into (4·13) and use of (4·14) results in

$$\begin{pmatrix} \dfrac{\partial\omega}{\partial I_1} & \dfrac{\partial\omega}{\partial I_2} & \dfrac{\partial\omega}{\partial I_3} \\ \dfrac{\partial k}{\partial I_1} & \dfrac{\partial k}{\partial I_2} & \dfrac{\partial k}{\partial I_3} \\ \dfrac{\partial l}{\partial I_1} & \dfrac{\partial l}{\partial I_2} & \dfrac{\partial l}{\partial I_3} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix} \quad \text{or} \quad \mathbf{Da} = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}. \tag{4·16}$$

It is interesting to note that the $\mathbf{D}$ matrix, generated by the variational principle (6·3), appears in a natural way in the linear stability analysis. The linear equation (4·16) is solvable if the variational principle (6·3) is non-degenerate; that is, $\det(\mathbf{D}) \neq 0$. Similar equations are found for the coefficients $\mathbf{b} \in \mathbb{R}^3$ and $\mathbf{c} \in \mathbb{R}^3$:

$$\mathbf{Db} = \begin{pmatrix} \omega \\ k \\ l \end{pmatrix} \quad \text{and} \quad \mathbf{Dc} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

The above equations can be combined as

$$\begin{pmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{pmatrix} \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} = \begin{pmatrix} -1 & \omega & 0 \\ 0 & k & 0 \\ 0 & l & 1 \end{pmatrix}. \tag{4·17}$$

This completes the construction of the solution (4·10) up to terms linear in the parameters (λ, α, β).

The operator $\mathscr{L}$ is formally symmetric and therefore the equation $\mathbf{P}\Psi = 0$ reduces to

$$\Lambda(\lambda, \alpha, \beta) = -[\phi_0, i\lambda \mathbf{M}(\hat{Z})\, U + i\alpha \mathbf{J}(\hat{Z})\, U + i\beta \mathbf{L}(\hat{Z})\, U] = 0. \tag{4·18}$$

The expression (4·18) is the dispersion relation for the linear stability problem evaluated at the nonlinear wave (H1). To obtain the form (4·11) we take successive derivatives. It is evident that $\Lambda(0,0,0) = 0$. For the first derivatives we find

$$\left.\frac{\partial\Lambda}{\partial\lambda}\right|_{\lambda=\alpha=\beta=0} = i[\phi_0, \mathbf{M}(\hat{Z})\, \phi_0] = 0,$$

$$\left.\frac{\partial\Lambda}{\partial\alpha}\right|_{\lambda=\alpha=\beta=0} = i[\phi_0, \mathbf{J}(\hat{Z})\, \phi_0] = 0,$$

$$\left.\frac{\partial\Lambda}{\partial\beta}\right|_{\lambda=\alpha=\beta=0} = i[\phi_0, \mathbf{L}(\hat{Z})\, \phi_0] = 0.$$

The vanishing of the above three expressions follows from the skew-symmetry of the linear operators $\mathbf{M}$, $\mathbf{K}$ and $\mathbf{L}$. Therefore the dispersion relation is quadratic at leading order and takes the form

$$\begin{aligned} \Lambda(\lambda, \alpha, \beta) = {} & \lambda^2[\phi_0, \mathbf{M}(\hat{Z})\, W_1] + \lambda\alpha[\phi_0, \mathbf{M}(\hat{Z})\, W_2] + \lambda\beta[\phi_0, \mathbf{M}(\hat{Z})\, W_3] \\ & + \alpha\lambda[\phi_0, \mathbf{J}(\hat{Z})\, W_1] + \alpha^2[\phi_0, \mathbf{J}(\hat{Z})\, W_2] + \alpha\beta[\phi_0, \mathbf{J}(\hat{Z})\, W_3] \\ & + \beta\lambda[\phi_0, \mathbf{L}(\hat{Z})\, W_1] + \beta\alpha[\phi_0, \mathbf{L}(\hat{Z})\, W_2] + \beta^2[\phi_0, \mathbf{L}(\hat{Z})\, W_3] + r_3 \end{aligned} \tag{4·19}$$

https://doi.org/10.1017/S0305004196001429 Published online by Cambridge University Press

where r_3 contains terms of degree 3 and higher in $|\lambda|$, $|\alpha|$ and $|\beta|$. The integrals in (4·19) are all of the form

$$[\phi_0, \mathbf{M}(\hat{Z})\,\phi_j], \quad [\phi_0, \mathbf{J}(\hat{Z})\,\phi_j] \quad \text{and} \quad [\phi_0, \mathbf{L}(\hat{Z})\,\phi_j]$$

for $j = 1, 2, 3$ which are in general non-zero. It is a remarkable fact, following from the variational principle of Section 2, that these integrals can be evaluated exactly:

$$\left.\begin{aligned} [\phi_0, \mathbf{M}(\hat{Z})\,\phi_j] &= \delta_{1j} \\ [\phi_0, \mathbf{L}(\hat{Z})\,\phi_j] &= -\delta_{3j} \\ [\phi_0, \mathbf{J}(\hat{Z})\,\phi_j] &= \omega\delta_{1j} - k\delta_{2j} - l\delta_{3j} \end{aligned}\right\} \quad \text{for} \quad j = 1, 2, 3. \tag{4·20}$$

These identities are verified as follows. The first identity will be verified as the other two follow the same line. Using the fact that $\mathscr{A}(Z) = I_1$ we have that

$$\begin{aligned} \delta_{1j} = \frac{\partial A}{\partial I_j} &= \left[\nabla A(\hat{Z}), \frac{\partial \hat{Z}}{\partial I_j}\right] \\ &= -\left[\mathbf{M}(\hat{Z})\frac{\partial \hat{Z}}{\partial \theta}, \phi_j\right] \\ &= [\phi_0, \mathbf{M}(\hat{Z})\,\phi_j] \quad \text{for} \quad j = 1, 2, 3. \end{aligned}$$

The result follows from the fact that I_1, I_2 and I_3 are level sets of the three action functionals and the fact that the gradients of the action functionals generate skew-symmetric operators.

Substitution of the identities (4·20) into (4·19) reduces the dispersion relation to

$$\begin{aligned} \Lambda(\lambda, \alpha, \beta) &= \begin{pmatrix} -\lambda + \alpha\omega \\ \alpha k \\ \beta + \alpha l \end{pmatrix} \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix} \begin{pmatrix} \lambda \\ \alpha \\ \beta \end{pmatrix} + r_3(\lambda, \alpha, \beta) \\ &= \begin{pmatrix} -\lambda + \alpha\omega \\ \alpha k \\ \beta + \alpha l \end{pmatrix} \begin{bmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{bmatrix}^{-1} \begin{pmatrix} -1 & \omega & 0 \\ 0 & k & 0 \\ 0 & l & 1 \end{pmatrix} \begin{pmatrix} \lambda \\ \alpha \\ \beta \end{pmatrix} + r_3(\lambda, \alpha, \beta) \\ &= \begin{pmatrix} -\lambda + \alpha\omega \\ \alpha k \\ \beta + \alpha l \end{pmatrix} \begin{bmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{bmatrix}^{-1} \begin{pmatrix} -\lambda + \alpha\omega \\ \alpha k \\ \beta + \alpha l \end{pmatrix} + r_3(\lambda, \alpha, \beta) \end{aligned}$$

completing the proof. The elimination of the three by three matrix containing the coefficients a_i, b_i and c_i for $i = 1, 2, 3$ follows using (4·17).

5. *Action, flow force and a proof of Lighthill's instability criterion*

The proof of instability for plane waves in one space dimension follows precisely as in Section 4 but there are some special properties of this case worth separate attention. In particular a proof of Lighthill's instability criterion is given. The proof is obtained by first showing that there is an interesting dual criterion for instability in terms of the flow force.

Consider a dispersive wave system in one space dimension formulated as a Hamiltonian system on a bi-symplectic structure with governing equation

$$\mathbf{M}(Z)\,Z_t + \mathbf{K}(Z)\,Z_x = \nabla S(Z) \quad Z \in \mathscr{M} \tag{5·1}$$

https://doi.org/10.1017/S0305004196001429 Published online by Cambridge University Press

where $\mathscr{M}$ is the phase space and $\mathbf{M}(Z)$ and $\mathbf{K}(Z)$ are skew-symmetric operators associated with exact two-forms. The Hamiltonian functional $S(Z)$ can be identified with the static part of the flow force for the system. The identification of $S(Z)$ with the flow force is a special property of Hamiltonian systems on a multi-symplectic structure in one space dimension (see Appendix B).

Suppose there exists a two-parameter family of periodic travelling waves parametrized by the actions as in hypothesis (H1) of Section 4 restricted to one space dimension. Linearizing (5·1) about the basic (H1) travelling wave results in a linear stability problem of the form

$$\mathscr{L}U = i\lambda\mathbf{M}(\hat{Z})\,U + i\alpha(k\mathbf{K}(\hat{Z}) - \omega\mathbf{M}(\hat{Z}))\,U \tag{5·2}$$

with terms defined as in Section 4. The analogue of Theorem 4·1 is

THEOREM 5·1. *Under the hypotheses* (*H*1)–(*H*3), *restricted to the case of waves in one space dimension, every solution of the linear stability problem* (5·2) *for* $|\lambda|$ *and* $|\alpha|$ *sufficiently small is in one to one correspondence with roots of the following dispersion relation*

$$\Lambda(\lambda,\alpha) = \left\langle \begin{pmatrix} -\lambda+\alpha\omega \\ \alpha k \end{pmatrix}, \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}^{-1} \begin{pmatrix} -\lambda+\alpha\omega \\ \alpha k \end{pmatrix} \right\rangle + r_3(\lambda,\alpha) \tag{5·3}$$

where the inner product $\langle\cdot,\cdot\rangle$ *is the usual real inner product on* $\mathbb{R}^2$ *and* $|r_3| = o(|\lambda|,|\alpha|)^2$.

The proof follows exactly the proof of Theorem 4·1 and leads immediately to

COROLLARY 5·2. *Suppose* $\mathscr{S}$ (*or* $-\mathscr{S}$) *evaluated on a travelling wave of* (5·1) *under the hypotheses* (*H*1)–(*H*3) *is a convex function on action space. Then the wave is linearly unstable.*

Proof. By Theorem 5·1 the analysis for $|\alpha|$ and $|\lambda|$ sufficiently small reduces to the analysis of $\Lambda(\lambda,\alpha)$ in (5·3) and, with $\Omega = -\lambda+\alpha\omega$,

$$\begin{aligned}\Lambda(\lambda,\alpha) &= \frac{1}{\Delta_S}(\Omega^2 S_{22} - 2S_{12}\,\Omega k\alpha + k^2\alpha^2 S_{11}) + r_3(\lambda,\alpha)\\ &= \frac{S_{22}}{\Delta_S}\left(\Omega - \frac{S_{12}}{S_{22}}\alpha k\right)^2 + \frac{(S_{11}S_{22}-S_{12}^2)}{S_{22}}k^2\alpha^2 + r_3(\lambda,\alpha)\\ &= \frac{S_{22}}{\Delta_S}\left(\Omega - \frac{S_{12}}{S_{22}}\alpha k\right)^2 + \frac{\Delta_S}{S_{22}}k^2\alpha^2 + r_3(\lambda,\alpha).\end{aligned}$$

When $\mathscr{S}$ (or $-\mathscr{S}$) is a convex function of the actions then $\Delta_S \overset{\text{def}}{=} S_{11}S_{22} - S_{12}^2$ is positive in which case $S_{22} \neq 0$. Therefore when $|\lambda|$ and $|\alpha|$ are sufficiently small and $\Delta_S > 0$ it follows that $\Lambda = 0$ has a root with $\Im(\lambda) \neq 0$ implying linear instability. ∎

The result in Corollary 5·2, that convexity of S on action space implies instability, was first stated for the special case of NLS in Bridges[**9**]. Corollary 5·2 generalizes this result to arbitrary systems of the form (5·1) and shows that the result is not equation dependent but is a property of systems with a multi-symplectic structure.

The constrained variational principle, in particular the fundamental properties of the Lagrange multiplier theory, lead to the precise dependence of the flow force on

https://doi.org/10.1017/S0305004196001429 Published online by Cambridge University Press

the actions (cf. Section 2). Although the importance of flow force for wave motion is well-established (cf. Benjamin [**3**] and references therein) the connection between flow force – and its dependence on action – and wave instability is new.

The Lagrange necessary condition for the constrained variational principle, with functional defined in equation (2·6), leads, when restricted to one space dimension, to the functional

$$\mathscr{F}(Z;\omega,k) = \mathscr{S}(Z) - \omega\mathscr{A}(Z) - k\mathscr{B}(Z) \tag{5·4}$$

where $\mathscr{S}$, $\mathscr{A}$ and $\mathscr{B}$ are S, A and B averaged over a wavelength. It follows that

$$\frac{\partial \mathscr{F}}{\partial \omega} = -\mathscr{A} \quad \text{and} \quad \frac{\partial \mathscr{F}}{\partial k} = -\mathscr{B}.$$

The image of the functional $\mathscr{F}$ is equivalent to (minus) the image of the averaged Lagrangian used in Lighthill [**21**, **22**]. The main distinction is that the structure of each term in (5·4) is characterized precisely; that is, the functional $S(Z)$ is separated from the rest of the Lagrange functional and, not only the value of each action functional is known, the structure – the set of skew-symmetric operators that define the gradients of action and action flux density – is also known and it is precisely the structure of the functionals that leads to a rigorous proof.

Once we identify the Lagrange functional $\mathscr{F}(Z;\omega,k)$ with the averaged Lagrangian of Whitham and Lighthill we obtain easily the following identities

$$\begin{aligned}\operatorname{sign}\begin{vmatrix} \mathscr{F}_{\omega\omega} & \mathscr{F}_{\omega k} \\ \mathscr{F}_{k\omega} & \mathscr{F}_{kk} \end{vmatrix} &= \operatorname{sign}\begin{vmatrix} \mathscr{A}_{\omega} & \mathscr{A}_{k} \\ \mathscr{B}_{\omega} & \mathscr{B}_{k} \end{vmatrix} \\ &= \operatorname{sign}\begin{vmatrix} \dfrac{\partial \omega}{\partial I_1} & \dfrac{\partial \omega}{\partial I_2} \\ \dfrac{\partial k}{\partial I_1} & \dfrac{\partial k}{\partial I_2} \end{vmatrix} = \operatorname{sign}\det\left[\operatorname{Hess}_I(\mathscr{S})\right] \end{aligned} \tag{5·5}$$

where I_1 and I_2 are values of the level sets of $\mathscr{A}$ and $\mathscr{B}$. Combining the above result with Corollary 5·2 proves the following.

Corollary 5·3 (Lighthill's instability criterion). *Suppose that the averaged Lagrangian* (*or minus the averaged Lagrangian*), *defined for the constrained variational principle* (5·4), *evaluated on a periodic travelling wave in one space dimension, is a convex function of the frequency and wavenumber. Then the wave is linearly unstable.*

It is important to note that the converse, when the sign of each of the determinants in (5·5) is negative, is a necessary but not sufficient condition for linear stability.

6. *Examples: NLS and the water-wave equations*

In this section two examples are presented. The examples, the Nonlinear Schrödinger equation (NLS) and the water-wave equations on infinite depth in two space dimensions, correspond to systems where the wave instability is well understood and are included here to show, in a familiar setting, how the theory of multi-symplectic structures applies to such systems. Two further examples, a nonlinear wave equation and a system of Boussinesq equations, are considered in

https://doi.org/10.1017/S0305004196001429 Published online by Cambridge University Press

Appendix A with particular attention to their reversible structure. The multi-symplectic structures theory also provides a framework for obtaining new results about waves. In a companion paper (Bridges[**11**]) results on the instability of periodic patterns, interacting with a mean flow, on the ocean surface are obtained using the multi-symplectic structures framework.

(a) *Nonlinear Schrödinger equation*. The form of the NLS equation considered here is

$$i\psi_t+\psi_{xx}+\epsilon_1\psi_{yy}+V'(|\psi|^2)\psi=0 \tag{6·1}$$

where $\psi(x,y,t)$ is complex-valued, i is the imaginary unit, $\epsilon_1=\pm1$ and $V(\cdot)$ is a smooth potential function. The cubic nonlinear Schrödinger equation is obtained by taking $V'(|\psi|^2)=\epsilon_2|\psi|^2$ with $\epsilon_2=\pm1$. The cubic NLS in two space dimensions appears as a model equation for the modulation of 3D travelling water waves on deep water. For example the model equation (6·1), with cubic nonlinearity, has been derived for gravity waves by Zakharov[**36**] (where $\epsilon_1=-1$ and $\epsilon_2=+1$) and for modulated capillary-gravity waves on deep water by Chen & Saffman [**16**, equation (2.20)] (where $\epsilon_1=+1$ and $\epsilon_2=+1$).

The system (6·1) has a well-known classical Hamiltonian structure. Let $\psi=u_1+iu_2$ for real-valued functions u_1 and u_2. Then (6·1) has the Hamiltonian formulation

$$u_t=\mathbf{J}\nabla H(u)\quad u=(u_1,u_2)$$

with $\mathbf{J}=\begin{pmatrix}0&1\\-1&0\end{pmatrix}$ and

$$H(u)=\int_{x_1}^{x_2}\int_{y_1}^{y_2}[\tfrac{1}{2}(u_{1x}^2+u_{2x}^2)+\tfrac{1}{2}\epsilon_1(u_{1y}^2+u_{2y}^2)-\tfrac{1}{2}V(u_1^2+u_2^2)]\,dy\,dx.$$

However, analysis of this system requires specification of the function space, including integration over x and y, on which $H(u)$ is well defined.

When $\epsilon_1\neq0$ the system (6·1) has a formulation as a Hamiltonian system on a tri-symplectic structure. Let

$$u_1+iu_2=\psi,\quad v_1+iv_2=\psi_x\quad\text{and}\quad w_1+iw_2=\psi_y,$$

where the six functions $u=(u_1,u_2)$, $v=(v_1,v_2)$ and $w=(w_1,w_2)$ are real-valued. In terms of the vector-valued set of variables $Z=(u,v,w)\in\mathbb{R}^6$ the tri-symplectic structure is given abstractly as $(\mathbb{R}^6,\omega^{(1)},\omega^{(2)},\omega^{(3)},S)$. The symplectic operators are defined by

$$\left.\begin{aligned}\omega^{(1)}(U,V)&=\langle U,\mathbf{M}V\rangle_{\mathbb{R}^6}\\ \omega^{(2)}(U,V)&=\langle U,\mathbf{K}V\rangle_{\mathbb{R}^6}\\ \omega^{(3)}(U,V)&=\langle U,\mathbf{L}V\rangle_{\mathbb{R}^6}\end{aligned}\right\}\quad\text{for any}\quad U,V\in\mathbb{R}^6$$

with
$$\mathbf{M}=\begin{pmatrix}\mathbf{J}&\mathbf{0}&\mathbf{0}\\\mathbf{0}&\mathbf{0}&\mathbf{0}\\\mathbf{0}&\mathbf{0}&\mathbf{0}\end{pmatrix}\quad\text{where}\quad\mathbf{J}=\begin{pmatrix}0&1\\-1&0\end{pmatrix},$$

$$\mathbf{K}=\begin{pmatrix}\mathbf{0}&-\mathbf{I}&\mathbf{0}\\\mathbf{I}&\mathbf{0}&\mathbf{0}\\\mathbf{0}&\mathbf{0}&\mathbf{0}\end{pmatrix}\quad\text{and}\quad\mathbf{L}=\begin{pmatrix}\mathbf{0}&\mathbf{0}&-\epsilon_1\mathbf{I}\\\mathbf{0}&\mathbf{0}&\mathbf{0}\\\epsilon_1\mathbf{I}&\mathbf{0}&\mathbf{0}\end{pmatrix}$$

https://doi.org/10.1017/S0305004196001429 Published online by Cambridge University Press

and $$S(Z) = \tfrac{1}{2}(v_1^2+v_2^2)+\tfrac{1}{2}\epsilon_1(w_1^2+w_2^2)+\tfrac{1}{2}V(u_1^2+u_2^2).$$

The governing equations, equivalent to NLS in (6·1), are then

$$\mathbf{M}Z_t+\mathbf{K}Z_x+\mathbf{L}Z_y = \nabla S(Z). \tag{6·2}$$

The plane wave solutions of NLS provide an elementary example of the rigorous theory of Section 4 where all the details can be worked out explicitly. In particular the hypotheses (H1)–(H3) of Theorem 4·1 can be verified explicitly. The plane wave solutions of NLS in (6·1) have the form

$$\psi(x,y,t) = R\,e^{i(kx+ly-\omega t)} \quad R\in\mathbb{C} \tag{6·3}$$

with (ω,k,l) satisfying

$$\omega = k^2+\varepsilon_1 l^2 - V'(|R|^2). \tag{6·4}$$

An analysis of the linear stability of this family of waves can be given exactly (cf. Newton & Keller [**27**]) and we will sketch the theory here for purposes of comparison with the rigorous characterisation of the instability given by Theorem 4·1.

For the linear stability analysis let

$$\psi(x,y,t) = [R+B(x,y,t)]\,e^{i(kx+ly-\omega t)}$$

where R is as above and B is complex-valued. Substitution into (6·1) leads to the following equation for B,

$$iB_t+B_{xx}+2ikB_x+\varepsilon_1 B_{yy}+2il\varepsilon_1 B_y+V''(|R|^2)\,R^2\bar{B}+V''(|R|^2)\,|R|^2 B = 0. \tag{6·5}$$

The equation governing B has constant coefficients and the general solution has the form

$$B(x,y,t) = U\,e^{i(\lambda t+\alpha x+\beta y)}+V\,e^{i(\bar{\lambda}t+\alpha x+\beta y)} \tag{6·6}$$

where U,V are complex-valued scalars, $\lambda\in\mathbb{C}$ and $(\alpha,\beta)\in\mathbb{R}^2$. Substitution of (6·6) into (6·5) leads to the algebraic equation

$$\begin{bmatrix} -\Lambda+|R|^2\,V''-\alpha^2-\varepsilon_1\beta^2 & V''(|R|^2)\,R^2 \\ V''(|R|^2)\,\bar{R}^2 & \Lambda+|R|^2\,V''-\alpha^2-\varepsilon_1\beta^2 \end{bmatrix}\begin{pmatrix} U \\ \bar{V}\end{pmatrix} = \begin{pmatrix} 0\\0\end{pmatrix}$$

where $\Lambda = \lambda+2k\alpha+2\varepsilon_1 l\beta$. Setting the determinant to zero results in the 'dispersion relation'

$$\tilde{\Delta}(\lambda,\alpha,\beta) = \Lambda^2+(\alpha^2+\varepsilon_1\beta^2)\,[2\,|R|^2\,V''(|R|^2)-(\alpha^2+\varepsilon_1\beta^2)] = 0. \tag{6·7}$$

If for some $(\alpha,\beta)\in\mathbb{R}^2$ there exists a root $\lambda\in\mathbb{C}$ of $\tilde{\Delta}(\lambda,\alpha,\beta)=0$ with $\Im(\lambda)\neq 0$ the plane wave is unstable. It is evident from (6·7) that

$$\begin{aligned} &(a)\ \text{if } \varepsilon_1 = -1 \text{ the wave is unstable,}\\ &(b)\ \text{if } \varepsilon_1 = +1 \text{ and } V''>0 \text{ the wave is unstable.}\end{aligned} \tag{6·8}$$

The special case of the cubic NLS is recovered by taking $V'(|\psi|^2) = \varepsilon_2|\psi|^2$ with $\varepsilon_2 = \pm 1$. The classification (6·8) recovers the classical instability results for NLS (Newton & Keller [**27**]).

https://doi.org/10.1017/S0305004196001429 Published online by Cambridge University Press

Before proceeding to analyse the instability using the multi-symplectic structure we first rewrite (6·7). Divide (6·7) by $2V''(|R|^2)$ supposing that $V'' \neq 0$ and define

$$\Delta(\lambda,\alpha,\beta) = \frac{\tilde{\Delta}(\lambda,\alpha,\beta)}{2V''(|R|^2)}.$$

Then (6·7) has the equivalent representation

$$\Delta(\lambda,\alpha,\beta) = \frac{1}{2V''}\begin{pmatrix} -\lambda \\ \alpha \\ \beta \end{pmatrix}\begin{bmatrix} 1 & -2k & -2\varepsilon_1 l \\ -2k & 4k^2+2|R|^2V'' & 4\varepsilon_1 kl \\ -2\varepsilon_1 l & 4\varepsilon_1 kl & 4l^2+2\varepsilon_1|R|^2V'' \end{bmatrix}\begin{pmatrix} -\lambda \\ \alpha \\ \beta \end{pmatrix} - \frac{(\alpha^2+\varepsilon_1\beta^2)^2}{2V''}. \tag{6·9}$$

We now give an analysis of the instability in terms of the multi-symplectic structure and the rigorous theory of Section 4. Consider NLS in the form (6·2). The plane wave (6·3) in terms of the $Z = (u,v,w)^T$ coordinates, has the form

$$\begin{aligned} u &= \mathscr{R}_\theta \hat{u}, \quad \theta = x+z-t, \quad \hat{u}\in\mathbb{R}^2 \\ v &= -kJ\mathscr{R}_\theta \hat{u} \\ w &= -lJ\mathscr{R}_\theta \hat{u} \quad \text{with} \quad \mathscr{R}_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}. \end{aligned} \tag{6·10}$$

The frequency and wavenumbers again satisfy (6·4) with $|R|^2$ identified with $|\hat{u}|^2$. However the form (6·4) is not complete with respect to the variational principle of Section 2; (ω,k,l) are in fact functions of the values of the level sets of the actions.

The three action functionals for NLS averaged over $\mathbb{T}^3$ are

$$\begin{aligned} \mathscr{A} &= \oint_{\mathbb{T}^3} \tfrac{1}{2}(u, Ju_t)\,dx\,dy\,dt, \\ \mathscr{B}_1 &= \oint_{\mathbb{T}^3} (v,u_x)\,dx\,dy\,dt \quad \text{and} \quad \mathscr{B}_2 = \oint_{\mathbb{T}^3} (w,u_y)\,dx\,dy\,dt, \end{aligned} \tag{6·11}$$

where $(\cdot,\cdot)$ is an inner product on $\mathbb{R}^2$. Substitution of the plane wave (6·10) into the actions results in

$$I_1 = \mathscr{A}(\omega,k,l) = -\tfrac{1}{2}|\hat{u}|^2, \quad I_2 = \mathscr{B}_1(\omega,k,l) = k|\hat{u}|^2 \quad \text{and} \quad I_3 = \mathscr{B}_2(\omega,k,l) = \epsilon_1 l|\hat{u}|^2. \tag{6·12}$$

In (6·12), the dependency of A, B_1 and B_2 on (ω,k,l) is given, implicitly, by inverting

$$V'(|\hat{u}|^2) = k^2+\varepsilon_1 l^2-\omega. \tag{6·13}$$

Therefore

$$\nabla_{(\omega,k,l)} A = \frac{1}{2V''}(1,-2k,-2\varepsilon_1 l)$$

$$\nabla_{(\omega,k,l)} B_1 = \frac{1}{2V''}(-2k, 4k^2+2|\hat{u}|^2V'', 4\varepsilon_1 kl)$$

$$\nabla_{(\omega,k,l)} B_2 = \frac{1}{2V''}(-2\varepsilon_1 l, 4\varepsilon_1 kl, 4l^2+2\varepsilon_1|\hat{u}|^2V'').$$

https://doi.org/10.1017/S0305004196001429 Published online by Cambridge University Press

Using the fact that the Jacobian of (A, B_1, B_2) with respect to (ω, k, l) is the inverse of $\mathrm{Hess}_I(S)$ we obtain the following expression for the $\mathrm{Hess}_I(S)$ evaluated on an NLS travelling wave,

$$\mathrm{Hess}_I(S)^{-1} = \frac{1}{2V''}\begin{pmatrix} 1 & -2k & -2\varepsilon_1 l \\ -2k & 4k^2+2|\hat{u}|^2 V'' & 4\varepsilon_1 kl \\ -2\varepsilon_1 l & 4\varepsilon_1 kl & 4l^2+2\epsilon_1|\hat{u}|^2 V'' \end{pmatrix}. \tag{6·14}$$

From which it follows that

$$\Delta_S \overset{\text{def}}{=} \det\left(\mathrm{Hess}_I(S)\right) = \frac{2\varepsilon_1 V''}{|\hat{u}|^4}, \quad \det\begin{pmatrix} S_{22} & S_{23} \\ S_{32} & S_{33} \end{pmatrix} = \frac{\varepsilon_1}{|\hat{u}|^2} \quad \text{and} \quad S_{33} = \frac{\Delta_S}{2V''}$$

and therefore

$$\operatorname{sign}\begin{vmatrix} S_{22} & S_{23} \\ S_{32} & S_{33} \end{vmatrix} = \varepsilon_1 \quad \text{and} \quad \operatorname{sign}\left(S_{33}\Delta_S\right) = \operatorname{sign}\left(V''\right).$$

According to Corollary 4·2, if $\varepsilon_1 = -1$ the wave is unstable and if $\varepsilon_1 = +1$ but $V'' > 0$ the wave is unstable; in agreement with the classical theory in (6·8).

Comparison of (6·14) with (6·9) shows that

$$\Delta(\lambda, \alpha, \beta) = \begin{pmatrix} -\lambda \\ \alpha \\ \beta \end{pmatrix}\begin{bmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{bmatrix}^{-1}\begin{pmatrix} -\lambda \\ \alpha \\ \beta \end{pmatrix} - \frac{(\alpha^2+\varepsilon_1\beta^2)^2}{2V''} \tag{6·15}$$

which is exact for the NLS equation. It is interesting to note that the terms of degree four in (α, β) in (6·15) provide the bandwidth of the unstable wavenumbers; whereas the theory of Section 4 predicts only a band of unstable wavenumbers but not the precise width.

(b) *Formal application to the water-wave problem*. A second example is the water-wave problem in three space dimensions (two, unbounded, evolution directions (x, y) and a vertical dimension z). The problem of gravity waves at the surface of an inviscid irrotational fluid of constant density is considered. The analysis of this example is formal. No attempt will be made to give a rigorous proof of the instability criterion here, although a rigorous proof is possible for water waves in some cases; for example, Bridges & Mielke [**13**] give a rigorous proof of sideband instability for two-dimensional water waves travelling in finite depth – the Benjamin–Feir instability, the necessary technicalities would take us too far afield here.

The object is to transform the governing equations so that the water-wave problem can be cast as a Hamiltonian system on a multi-symplectic structure.

The governing equations for the water-wave problem are as follows (cf. Whitham [**35**, §13]). In the interior of the fluid the governing equation is Laplace's equation for the velocity potential:

$$\Delta\phi = \frac{\partial^2\phi}{\partial x^2}+\frac{\partial^2\phi}{\partial y^2}+\frac{\partial^2\phi}{\partial z^2} = 0 \quad -h < z < \eta(x, y, t) \tag{6·16}$$

for all $(x, y) \in \mathbb{R}^2$. At the bottom we have

$$\phi_z = 0 \quad \text{at} \quad z = -h \quad \text{for all} \quad (x, y) \in \mathbb{R}^2. \tag{6·17}$$

https://doi.org/10.1017/S0305004196001429 Published online by Cambridge University Press

And at the interface the boundary conditions are

$$\left.\begin{aligned}\eta_t+\phi_x\eta_x+\phi_y\eta_y-\phi_z&=0\\ \phi_t+\tfrac{1}{2}(\phi_x^2+\phi_y^2+\phi_z^2)+g\eta&=0\end{aligned}\right\}\quad\text{at}\quad z=\eta(x,z,t).\tag{6·18}$$

This system has a characterization as a Hamiltonian evolution equation of the form (Zakharov [**36**], Broer [**15**], Benjamin & Olver [**5**])

$$\Phi_t=\frac{\delta H}{\delta\eta}\quad\text{and}\quad\eta_t=-\frac{\delta H}{\delta\Phi}$$

where $\Phi(x,y,t)\stackrel{\text{def}}{=}\phi(x,y,z,t)\,|_{z=\eta(x,y,t)}$ and H is the total energy

$$H=\int_{x_1}^{x_2}\int_{y_1}^{y_2}\left[\int_{-h}^{\eta}\tfrac{1}{2}|\nabla\phi|^2\,dz\right]+\tfrac{1}{2}g\eta^2\,dx\,dy.$$

We now show that the above-defined water-wave problem has a characterization as a Hamiltonian system on the following multi-symplectic structure:

$$(\mathscr{M},\omega_m^{(1)},\omega_m^{(2)},\omega_m^{(3)},S).$$

To verify this, we first define the following skew-symmetric operators:

$$\mathbf{M}(Z)=\begin{cases}-1 & \text{if } i=1 \text{ and } j=2\\ +1 & \text{if } i=2 \text{ and } j=1\\ 0 & \text{otherwise}\end{cases}\quad\text{with}\quad i,j=1,\dots,5$$

$$\mathbf{K}(Z)=\begin{pmatrix}0&-\mathbf{u}&0&0&0\\ \mathbf{u}&0&0&0&0\\ 0&0&0&-1&0\\ 0&0&1&0&0\\ 0&0&0&0&0\end{pmatrix}\quad\text{and}\quad\mathbf{L}(Z)=\begin{pmatrix}0&-\mathbf{v}&0&0&0\\ \mathbf{v}&0&0&0&0\\ 0&0&0&0&-1\\ 0&0&0&0&0\\ 0&0&1&0&0\end{pmatrix}$$

where $\mathbf{u}=\phi_x\,|_{z=\eta}$ and $\mathbf{v}=\phi_y\,|_{y=\eta}$. Define the five-component vector

$$Z=\begin{pmatrix}\Phi\\ \eta\\ \phi\\ u\\ v\end{pmatrix}\quad\text{with}\quad\Phi\stackrel{\text{def}}{=}\phi\,|_{z=\eta},\quad u=\phi_x,\quad v=\phi_y\tag{6·19}$$

where $(\Phi,\eta)\in\mathbb{R}^2$ with $h+\eta>0$ and (ϕ,u,v) are defined on the cross-section $z\in(-h,\eta)$. For vector-valued functions of the form (6·19) a suitable inner product is

$$\langle U,V\rangle_m=U_1V_1+U_2V_2+\int_{-h}^{\eta}(U_3V_3+U_4V_4+U_5V_5)\,dz$$

where the m indicates the dependence of the inner product on position through η appearing in the integral. The Hamiltonian function S is defined to be

$$S(Z)=\frac{1}{2}\int_{-h}^{\eta}(u^2+v^2-\phi_z^2)\,dz-\tfrac{1}{2}g\eta^2\tag{6·20a}$$

https://doi.org/10.1017/S0305004196001429 Published online by Cambridge University Press

with gradient (defined using the above inner product $\langle \cdot, \cdot \rangle_m$)

$$\nabla S(Z) \stackrel{\text{def}}{=} \begin{pmatrix} \delta S/\delta \Phi \\ \delta S/\delta \eta \\ \delta S/\delta \phi \\ \delta S/\delta u \\ \delta S/\delta v \end{pmatrix} = \begin{pmatrix} -\phi_z|_{z=\eta} \\ \frac{1}{2}(\mathbf{u}^2+\mathbf{v}^2+\phi_z^2)-g\eta \\ \phi_{zz} \\ u \\ v \end{pmatrix}. \tag{6·20b}$$

Now, using the above definitions we find

$$\mathbf{M}(Z)\,Z_t+\mathbf{K}(Z)\,Z_x+\mathbf{L}(Z)\,Z_y = \begin{pmatrix} -\eta_t \\ \Phi_t \\ 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} -\mathbf{u}\eta_x \\ \mathbf{u}\Phi_x \\ -u_x \\ \phi_x \\ 0 \end{pmatrix} + \begin{pmatrix} -\mathbf{v}\eta_y \\ \mathbf{v}\Phi_y \\ -v_y \\ 0 \\ \phi_y \end{pmatrix}$$

and equating this to $\nabla S(Z)$ in (6·20b) results in

$$\begin{aligned} -\eta_t-\mathbf{u}\eta_x-\mathbf{v}\eta_y &= -\phi_z|_{z=\eta} \\ \Phi_t+\mathbf{u}\Phi_x+\mathbf{v}\Phi_y &= \tfrac{1}{2}(\mathbf{u}^2+\mathbf{v}^2+\phi_z^2)|_{z=\eta}-g\eta \\ -u_x-v_y &= \phi_{zz} \\ \phi_x &= u \\ \phi_y &= v. \end{aligned}$$

The last two equations are identities and the third to last equation recovers Laplace's equation. The first equation is the kinematic free surface condition and the second equation recovers the Bernoulli condition (using the identities $\Phi_{(j)} = \phi_j|_{z=\eta}+\phi_y|_{z=\eta}\,\eta_j$ for $j = t$, x or y).

We have verified that the representation

$$\mathbf{M}(Z)\,Z_t+\mathbf{K}(Z)\,Z_x+\mathbf{L}(Z)\,Z_y = \nabla S(Z), \tag{6·21}$$

along with the definition of Z and the boundary conditions $\Phi = \phi|_{z=\eta}$ and $\phi_z|_{z=-h} = 0$, formally recovers the governing equations for the specified water-wave problem.

For the case of water waves where the free surface position is an unknown the representation of the governing equations (6·21) in the abstract form (1·11) is preferable. This is because the phase space for the system includes functions defined on the interval $z \in (-h, \eta)$ which clearly depends on η (note also that the inner product depends on η).

It remains to verify that the three symplectic forms $\omega^{(j)}, j = 1, 2, 3$ are closed. However, introduce the following one-forms,

$$\alpha^{(1)} = -\Phi\,d\eta, \quad \alpha^{(2)} = \int_{-h}^{\eta} u\,d\phi\,dz \quad \text{and} \quad \alpha^{(3)} = \int_{-h}^{\eta} v\,d\phi\,dz.$$

Taking the exterior derivative of the above one-forms results in

$$\begin{aligned} d\alpha^{(1)} &= d\eta \wedge d\Phi \\ d\alpha^{(2)} &= \int_{-h}^{\eta} du \wedge d\phi\,dz + \mathbf{u}\,d\eta \wedge d\Phi \\ d\alpha^{(3)} &= \int_{-h}^{\eta} dv \wedge d\phi\,dz + \mathbf{v}\,d\eta \wedge d\Phi. \end{aligned}$$

https://doi.org/10.1017/S0305004196001429 Published online by Cambridge University Press

Evaluation on tangent vectors leads immediately to the identities $\omega^{(j)} = -d\alpha^{(j)}$ for $j = 1, 2, 3$; the two-forms are exact and therefore, formally, closed.

In the case of one evolution direction ($\partial_y = 0$) the Hamiltonian functional in (6·21) can be identified precisely with the flow force for the water-wave problem (see Appendix B). For the case of two (evolution) space dimensions the flow force is tensor valued. For the water wave problem (6·21) the impulse has two components, $-\Phi\eta_x$ and $-\Phi\eta_y$, and the first component satisfies the conservation law

$$(-\Phi\eta_x)_t + (S_{11})_x + (S_{13})_y = 0$$

where
$$S_{11} = S - A - \int_{-h}^{\eta} v^2\, dz \quad \text{and} \quad S_{13} = \int_{-h}^{\eta} uv\, dz$$

(cf. Bridges [**10**, Section 2]). The functional S is the Hamiltonian function for the system (6·21) defined in (6·20b) and $A = -\Phi\eta_t$ is the action density for the system. When $\partial_y = 0$ the term S_{11} reduces to $S - A$ and therefore the Hamiltonian function is precisely the static part (the term A vanishing on time-independent states) of the flow force. However when $v \not\equiv 0$ the precise identification of S with a physically relevant quantity is not immediately evident.

The theory of Sections 2 and 4 is formally applied to the problem of instability of travelling waves on fluid of infinite depth in two space dimensions (the third, non-evolutionary, space dimension associated with the vertical direction). In one space dimension the instability of gravity waves was first demonstrated theoretically and experimentally by Benjamin & Feir [**4**], Lighthill [**22**] and Whitham [**34**]. A rigorous proof of instability for travelling waves in finite depth has recently been given by Bridges & Mielke [**13**]. The instability of gravity waves in two space dimensions on water of infinite depth was first demonstrated by Zakharov [**36**] and Benney & Roskes [**6**] by deriving a modulation equation for the wave amplitude. The purpose here is to sketch how the theory of Section 4 is formally applied to the simplest water-wave instability in two space dimensions: gravity waves of infinitesimal amplitude on deep water.

With $Z = (\phi, \eta, \phi, u, v)$ the basic functionals for the water wave problem averaged over $\mathbb{T}^3$ with $h \to \infty$ are (cf. Section 4)

$$\begin{aligned}
\mathscr{S} &= \oint_{\mathbb{T}^3} \left[\int_{-\infty}^{\eta} \tfrac{1}{2}(u^2 + v^2 - \phi_z^2)\, dz - \tfrac{1}{2} g\eta^2 \right] dT \\
\mathscr{A} &= \oint_{\mathbb{T}^3} -\Phi\eta_t\, dT \\
\mathscr{B}_1 &= \oint_{\mathbb{T}^3} \int_{-\infty}^{\eta} u\phi_x\, dz\, dT \\
\mathscr{B}_2 &= \oint_{\mathbb{T}^3} \int_{-\infty}^{\eta} v\phi_y\, dz\, dT
\end{aligned} \tag{6·22}$$

where
$$\oint_{\mathbb{T}^3} (\cdot)\, dT = \frac{1}{(2\pi)^3} \int_0^{2\pi} \int_0^{2\pi} \int_0^{2\pi} (\cdot)\, d\xi\, d\zeta\, d\tau$$

with $\xi = kx$, $\zeta = ly$ and $\tau = \omega t$. For a weakly nonlinear wave travelling in two space dimensions we take the ansatz

https://doi.org/10.1017/S0305004196001429 Published online by Cambridge University Press

$$\eta = a_1 \cos(x+z-t) + a_2 \cos 2(x+z-t) + \dots$$
$$\phi = b_1 e^{pz} \sin(x+z-t) + b_2 e^{2pz} \sin 2(x+z-t) + \dots \tag{6·23}$$

where $p = \sqrt{(k^2+l^2)}$.

By the variational principle of Section 2, the necessary condition for existence is that $\nabla \mathscr{F} = 0$ where

$$\mathscr{F}(a_1, \dots, b_1, \dots; \omega, k, l) = \mathscr{S} - \omega \mathscr{A} - k\mathscr{B}_1 - l\mathscr{B}_2.$$

Substituting (6·23) into (6·22) and solving $\nabla \mathscr{F} = 0$ results in

$$b_1 = \frac{\omega}{p} a_1 (1 - \tfrac{3}{2} p a_2 + \tfrac{1}{8} p^2 a_1^2) + \dots$$

$$b_2 = \frac{\omega}{p} a_2 - \tfrac{1}{2} \omega a_1^2 + \dots$$

$$a_2 = \tfrac{1}{2} p a_1^2 + \dots$$

with dispersion relation

$$\hat{\mathscr{F}}(a_1^2; \omega, k, l) = \omega^2 - gp - gp^3 a_1^2 + \dots = 0. \tag{6·24}$$

Note that, with respect to the variational principle of Section 2, the problem is not complete in the form (6·24); it is completed by substituting (6·24) into the constraint sets $\mathscr{A} = I_1$, $\mathscr{B}_1 = I_2$ and $\mathscr{B}_2 = I_3$ resulting in ω, k and l as functions of $I \in \mathscr{U} \subset \mathbb{R}^3$.

For the instability theory the matrix $\mathrm{Hess}_I(S)$ is necessary. However, using the fact that

$$\mathrm{Hess}_I(S)^{-1} = \begin{pmatrix} \nabla A \\ \nabla B_1 \\ \nabla B_2 \end{pmatrix}, \tag{6·25}$$

where $\nabla = (\partial_\omega, \partial_k, \partial_l)$, the inverse is constructed. For the actions we find

$$\mathscr{A} = -\frac{1}{2}\frac{\omega}{p} a_1^2 + \dots, \quad \mathscr{B}_1 = \frac{1}{2}\frac{\omega}{p} c_g^{(1)} a_1^2 + \dots \quad \text{and} \quad \mathscr{B}_2 = \frac{1}{2}\frac{\omega}{p} c_g^{(2)} a_1^2 + \dots \tag{6·26}$$

where $(c_g^{(1)}, c_g^{(2)}) = (\omega_k, \omega_l)$ are the components of the group velocity. Therefore (6·26) can be expressed as

$$(\mathscr{A}, \mathscr{B}_1, \mathscr{B}_2) = -\mathscr{A}(-1, \omega_k, \omega_l) + \dots$$

and therefore

$$\mathrm{Hess}_I(\mathscr{S})^{-1} = \begin{pmatrix} \mathscr{A}_\omega & \mathscr{A}_k & \mathscr{A}_l \\ -\omega_k \mathscr{A}_\omega & -\omega_k \mathscr{A}_k - \omega_{kk} \mathscr{A} & -\omega_k \mathscr{A}_l - \omega_{kl} \mathscr{A} \\ -\omega_l \mathscr{A}_\omega & -\omega_l \mathscr{A}_k - \omega_{lk} \mathscr{A} & -\omega_l \mathscr{A}_l - \omega_{ll} \mathscr{A} \end{pmatrix} + \dots.$$

For a_1^2 sufficiently small

$$\det(\mathrm{Hess}_I(\mathscr{S})^{-1}) = \mathscr{A}^2 \mathscr{A}_\omega \begin{vmatrix} \omega_{kk} & \omega_{kl} \\ \omega_{lk} & \omega_{ll} \end{vmatrix} + \dots.$$

However

$$\mathrm{sign} \begin{vmatrix} S_{22} & S_{23} \\ S_{32} & S_{33} \end{vmatrix} = \mathrm{sign}(\mathscr{A}_\omega)\, \mathrm{sign}(\Delta_S)$$

and so

$$\mathrm{sign} \begin{vmatrix} S_{22} & S_{23} \\ S_{32} & S_{33} \end{vmatrix} = \mathrm{sign} \begin{vmatrix} \omega_{kk} & \omega_{kl} \\ \omega_{lk} & \omega_{ll} \end{vmatrix} \quad \text{for } a_1^2 \text{ sufficiently small.} \tag{6·27}$$

https://doi.org/10.1017/S0305004196001429 Published online by Cambridge University Press

By formal application of Corollary 4·2 the wave is unstable if $\omega_{kk}\,\omega_{ll}-\omega_{kl}^2<0$. But $\omega^2=gp$ and so

$$\begin{pmatrix}\omega_{kk} & \omega_{ll}\\ \omega_{lk} & \omega_{ll}\end{pmatrix}=\frac{g}{4\omega p^3}\begin{pmatrix}2l^2-k^2 & -3kl\\ -2kl & 2k^2-l^2\end{pmatrix}$$

and
$$\det\begin{vmatrix}\omega_{kk} & \omega_{ll}\\ \omega_{lk} & \omega_{ll}\end{vmatrix}=\left(\frac{g}{4\omega p^3}\right)^2[(2l^2-k^2)(2k^2-l^2)-9k^2l^2]=-\frac{1}{8}\frac{g}{p^3}<0$$

recovering the classic instability of deep water gravity waves in two space dimensions. In fact (6·27) also recovers the well-known result, for infinitesimal waves travelling in two-space dimensions, that if the dispersion relation is indefinite, $\omega_{kk}\,\omega_{ll}-\omega_{kl}<0$, the weakly nonlinear wave is unstable.

7. *Concluding remarks*

The properties and applications of a Hamiltonian structure, which generalizes classical Hamiltonian structure for evolution equations by assigning a distinct symplectic operator for each unbounded direction and time, has been presented. It was shown that the multi-symplectic structure has general properties that make it natural for the analysis of wave propagation problems. Applications of the abstract properties of the multi-symplectic structure were presented: a variational principle for space-time periodic states reminiscent of the variational principle for invariant tori in finite-dimensional Hamiltonian systems, a geometric reformulation of the concepts of action and action flux, a rigorous proof of the instability criterion predicted by the Whitham modulation equations, a new symplectic decomposition of the Noether theory, generalization of the concept of the reversibility to space-time and a proof of Lighthill's geometric criterion for instability of periodic waves travelling in one space dimension.

It is noteworthy that the variational principle presented in Section 2 is not restricted to periodic travelling waves. It applies to all space-time periodic waves and a notable example is standing waves. However standing waves are truly space and time periodic (whereas periodic travelling waves are time-independent in a suitable moving frame of reference) and are therefore prone to small divisor difficulties. The small divisor difficulty can already be seen by treating the wave equation on a fixed spatial interval (cf. Wayne[**32**]). However it would be of great interest to extend the instability theory of Section 4 to the case of standing waves. The difficulty with standing waves (outside of the functional analytic difficulties and small divisors) is that the tangent space to the basic state includes $\{Z_t, Z_x\}$. Therefore the kernel of the linear operator $\mathscr{L}$ in (4·5), that appears in the linear stability problem, is larger. Some results towards this problem are in Bridges[**11**]. In the case of periodic travelling waves, treated here, the tangent vectors Z_t and Z_x are not independent restricting the kernel to the single element Z_θ as in (H3) of Section 4.

The concept of signature is important in the theory of Hamiltonian systems for classifying instabilities. For example, Mackay & Saffman[**23**] use the theory of signature to classify the instabilities of large amplitude travelling water waves. In the multi-symplectic structures framework new results on signature are possible by generating a set of signature functions using each symplectic operator. In particular, in Bridges[**12**], the properties of the signature family for linearized Hamiltonian

https://doi.org/10.1017/S0305004196001429 Published online by Cambridge University Press

systems on a multi-symplectic are introduced and it is shown that, whereas the classical signature, associated with the temporal symplectic form, gives information about eigenvalue collision leading to instability, the spatial symplectic forms generate signature invariants that further classify the spatial spreading of instabilities, and in some cases distinguish between absolute and convective instabilities.

An interesting ramification of the multisymplectic structures framework is that mean flow effects, which are important in the dynamics of water waves in shallow depth fluid, can be characterized as drift along a group orbit. The connection between symmetry and mean flow is treated in detail in Bridges[**11**] where it is also incorporated into a theory for linear instability of periodic patterns on the ocean surface interacting with a mean flow.

Finally we remark that a perturbation theory that uses the geometry of the multi-symplectic structure can be constructed for systems with non-Hamiltonian perturbations (damping, inhomogeneities, external forcing) of conservative systems. The idea is to write a perturbed wave problem in the form (restricting to one space dimension for example)

$$\mathbf{M}Z_t+\mathbf{K}Z_x-\nabla S(Z)=\epsilon F(Z,x,t;\epsilon) \tag{7·1}$$

where $|\epsilon|\ll 1$ is a small parameter. A familiar example that can be cast into this form is the perturbed Sine–Gordon equation

$$\frac{\partial^2 u}{\partial t^2}-\varepsilon_1\frac{\partial^2 u}{\partial x^2}+\sin u=-\epsilon\left(a\frac{\partial u}{\partial t}+b\cos\omega t+c\cos kx\right). \tag{7·2}$$

The coefficients (a,b,c) are real numbers and $\varepsilon_1=\pm 1$. When $\varepsilon_1=+1$ the unperturbed equation is the Sine–Gordon equation whereas when $\varepsilon_1=-1$ the unperturbed equation is a semi-linear elliptic equation. Letting

$$v=u_t+\epsilon\frac{b}{\omega}\sin\omega t \quad \text{and} \quad w=u_x-\epsilon\varepsilon_1\frac{c}{k}\sin kx,$$

the equation (7·2) can be expressed in the form (7·1) by taking $S(Z)=\frac{1}{2}(v^2-\varepsilon_1 w^2)-\cos u$,

$$Z=\begin{pmatrix} u\\ v\\ w\end{pmatrix},\quad \mathbf{M}=\begin{pmatrix} 0 & -1 & 0\\ 1 & 0 & 0\\ 0 & 0 & 0\end{pmatrix},\quad \mathbf{K}=\begin{pmatrix} 0 & 0 & \varepsilon_1\\ 0 & 0 & 0\\ -\varepsilon_1 & 0 & 0\end{pmatrix}$$

and

$$F(Z,x,t)=\begin{pmatrix} av-\epsilon\dfrac{ab}{\omega}\sin\omega t\\ -\dfrac{b}{\omega}\sin\omega t\\ -\dfrac{c}{k}\sin kx\end{pmatrix}.$$

Both cases $\varepsilon_1=\pm 1$ can be characterized as perturbed multi-symplectic Hamiltonian systems. The advantage of the multi-symplectic structure in (7·1) is that it provides structural information about the unperturbed problem which contributes to the analysis of the perturbed problem.

https://doi.org/10.1017/S0305004196001429 Published online by Cambridge University Press

Appendix A. Symplectic and anti-symplectic involutions and reversibility

In this appendix the concept of reversibility of a Hamiltonian system is generalized to the case of more than one reversible direction. Recall the definition of reversibility for a classical Hamiltonian system written in the form

$$\mathbf{J}U_t = \nabla H(U) \quad U \in \mathscr{M} \tag{A 1}$$

where $\mathscr{M}$ is the phase space, $\mathbf{J}$ is a skew-symmetric operator associated with a closed two-form and $\nabla H(U)$ is the gradient of the Hamiltonian functional. The system (A 1) is called a reversible-Hamiltonian system if there exists an anti-symplectic involution $\mathbf{R}$; that is, $\mathbf{R}$ satisfies

$$\mathbf{R} = \mathbf{R}^{-1} \quad \text{and} \quad \mathbf{R}^*\mathbf{J}\mathbf{R} = -\mathbf{J},$$

where $\mathbf{R}^*$ is the formal adjoint of $\mathbf{R}$, such that H is $\mathbf{R}$-invariant: $H(\mathbf{R}\cdot U) = H(U)$. For such systems $\mathbf{R}U(-t)$ is a solution whenever $U(t)$ is a solution since, acting an (A 1) with $\mathbf{R}^*$ results in

$$\mathbf{R}^*\mathbf{J}U_t = \mathbf{R}^*\nabla H(U). \tag{A 2}$$

But $\mathbf{R}^*\mathbf{J} = -\mathbf{J}\mathbf{R}$, since $\mathbf{R}$ is anti-symplectic, and $\mathbf{R}$-invariance of H implies that $\mathbf{R}^*\nabla H(U) = \nabla H(\mathbf{R}.U)$ and therefore

$$-\mathbf{J}(\mathbf{R}U)_t = \nabla H(\mathbf{R}U), \tag{A 3}$$

showing that $\mathbf{R}U(-t)$ is also a solution.

The concept of reversibility is now generalized to higher dimension with particular attention to the role of the multi-symplectic structure. For simplicity consider a Hamiltonian system on a bi-symplectic structure with governing equation

$$\mathbf{M}(Z)\,Z_t + \mathbf{K}(Z)\,Z_x = \nabla S(Z) \quad Z \in \mathscr{M}, \tag{A 4}$$

where $\mathscr{M}$ is the phase space and the operators $\mathbf{M}(Z)$ and $\mathbf{K}(Z)$ are skew-symmetric and associated with closed two-forms. There are two independent directions, x and t, in which the system can be reversible. We introduce the following definitions.

The system (A 4) is called *t-reversible*, or time-reversible, if there exists an involution $\mathbf{R}_t$ acting on $\mathscr{M}, \mathbf{R}_t = \mathbf{R}_t^{-1}$, such that $S(\mathbf{R}_t\cdot Z) = S(Z)$,

$$\mathbf{R}_t^*\,\mathbf{M}(\mathbf{R}_t\cdot Z)\,\mathbf{R}_t = -\mathbf{M}(Z) \quad \text{and} \quad \mathbf{R}_t^*\,\mathbf{K}(\mathbf{R}_t\cdot Z)\,\mathbf{R}_t = \mathbf{K}(Z).$$

In other words the involution is *anti-symplectic* with respect to the operator $\mathbf{M}(Z)$ but *symplectic* with respect to the operator $\mathbf{K}(Z)$.

Similarly, the system (A 4) is called *x-reversible* if there exists an involution $\mathbf{R}_x$ acting on $\mathscr{M}, \mathbf{R}_x = \mathbf{R}_x^{-1}$, such that $S(\mathbf{R}_x.Z) = S(Z)$,

$$\mathbf{R}_x^*\,\mathbf{M}(\mathbf{R}_x\cdot Z)\,\mathbf{R}_x = \mathbf{M}(Z) \quad \text{and} \quad \mathbf{R}_x^*\,\mathbf{K}(\mathbf{R}_x.Z)\,\mathbf{R}_x = -\mathbf{K}(Z).$$

In this case the involution is required to be *symplectic* with respect to the operator $\mathbf{M}(Z)$ but *anti-symplectic* with respect to the operator $\mathbf{K}(Z)$.

The above definition of reversibility, in the context of multi-symplectic structures, has natural further generalization to systems on higher space dimension.

A simple example of the utility of the above definitions is given by a semilinear wave equation in one space dimension

$$u_{tt} - u_{xx} = V'(u) \tag{A 5}$$

https://doi.org/10.1017/S0305004196001429 Published online by Cambridge University Press

where $V(u)$ is some smooth nonlinear function of u. This system is clearly reversible in both x and t; that is, if $u(x,t)$ is a solution then so is $u(-x,t)$ and $u(x,-t)$. By letting $v = u_t$ and $w = u_x$ the system (A 5) can be reformulated as a Hamiltonian system on a multi-symplectic structure

$$\begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} u \\ v \\ w \end{pmatrix}_t + \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} u \\ v \\ w \end{pmatrix}_x = \begin{pmatrix} V'(u) \\ -v \\ w \end{pmatrix}$$

or with

$$\mathbf{M} = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \mathbf{K} = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \mathbf{Z} = \begin{pmatrix} u \\ v \\ w \end{pmatrix} \tag{A 6}$$

and $S(Z) = \frac{1}{2}(w^2 - v^2) + V(u)$ the system (A 5) has the form

$$\mathbf{M}Z_t + \mathbf{K}Z_x = \nabla S(Z) \quad Z \in \mathscr{M} \subset \mathbb{R}^3.$$

Introduce the following reversors

$$\mathbf{R}_t = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad \mathbf{R}_x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

It is clear that $\mathbf{R}_t = \mathbf{R}_t^{-1}$ and $\mathbf{R}_x = \mathbf{R}_x^{-1}$ and since $S(Z)$ is even in both v and w it follows that $S(\mathbf{R}_t \cdot Z) = S(Z) = S(\mathbf{R}_x \cdot Z)$. It remains to check the symplecticity of $\mathbf{M}$ and $\mathbf{K}$. Using the definitions of $\mathbf{M}$ and $\mathbf{K}$ in (A 6) we find

$$\begin{aligned} \mathbf{R}_t \mathbf{M} \mathbf{R}_t &= -\mathbf{M} \\ \mathbf{R}_t \mathbf{K} \mathbf{R}_t &= \mathbf{K} \end{aligned} \quad \text{and} \quad \begin{aligned} \mathbf{R}_x \mathbf{M} \mathbf{R}_x &= \mathbf{M} \\ \mathbf{R}_x \mathbf{K} \mathbf{R}_x &= -\mathbf{K}. \end{aligned}$$

Therefore $\mathbf{R}_t$ is symplectic (anti-symplectic) with respect to $\mathbf{K}$ (respectively $\mathbf{M}$) and $\mathbf{R}_x$ is symplectic (anti-symplectic) with respect to $\mathbf{M}$ (respectively $\mathbf{K}$); establishing a connection between the generalized reversibility and the generalized symplecticity of the wave equation (A 5).

As a second example of the above definitions the Boussinesq equations will be characterised as an x-reversible system in the above sense. In this example one of the symplectic operators is Z-dependent. The Boussinesq equations are a model for shallow water wave propagation and the system favoured by Boussinesq (cf. Whitham [**33**], [**35**, p. 462]) is

$$\begin{aligned} h_t + uh_x + hu_x &= 0 \\ u_t + uu_x + gh_x + \tfrac{1}{3}h_0 h_{xtt} &= 0 \end{aligned} \tag{A 7}$$

where g is the gravitational constant, h_0 is the stillwater level, $u(x,t)$ is the local velocity and $h(x,t)$ is a positive elevation function. Introducing a velocity potential ϕ, with defining equation $u = \phi_x$, allows integration of the second equation of (A 7) to

$$\phi_t + \tfrac{1}{2}u^2 + gh + \tfrac{1}{3}h_0 \eta_t = R(t) \tag{A 8}$$

where $R(t)$ is some function of time and $\eta = h_t$. The above system can be written as a Hamiltonian system on a multi-symplectic structure by introducing

$$Z = \begin{pmatrix} h \\ u \\ \eta \\ \phi \end{pmatrix} \in \mathbb{R}^4, \quad \mathbf{M} = \begin{pmatrix} 0 & 0 & \frac{1}{3}h_0 & 1 \\ 0 & 0 & 0 & 0 \\ -\frac{1}{3}h_0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}, \quad \mathbf{K}(Z) = \begin{pmatrix} 0 & 0 & 0 & u \\ 0 & 0 & 0 & h \\ 0 & 0 & 0 & 0 \\ -u & -h & 0 & 0 \end{pmatrix} \tag{A 9}$$

https://doi.org/10.1017/S0305004196001429 Published online by Cambridge University Press

and $S(Z) = Rh - \frac{1}{2}gh^2 + \frac{1}{2}hu^2 - \frac{1}{6}h_0\eta^2$ in which case the governing equations take the form

$$\mathbf{M}Z_t + \mathbf{K}(Z)Z_x = \nabla S(Z) \quad Z \in \mathscr{M} \subset \mathbb{R}^4. \tag{A 10}$$

It is easily verified that $\mathbf{M}$ and $\mathbf{K}(Z)$ are skew-symmetric (with respect to the standard inner product on $\mathbb{R}^4$) and correspond to exact two forms.

Introduce the following involution

$$\mathbf{R}_x = \mathrm{diag}\,[1, -1, 1, 1] \quad \text{with action} \quad \mathbf{R}_x \cdot Z = (h, -u, \eta, \phi).$$

Then clearly $\mathbf{R}_x = \mathbf{R}_x^{-1}$ and since $S(Z)$ is even in u it follows that $S(\mathbf{R}_x \cdot Z) = S(Z)$. It remains to check the action of $\mathbf{R}_x$ on the symplectic operators. Using the definitions of $\mathbf{M}$ and $\mathbf{K}(Z)$ in equations (A 9) we find

$$\mathbf{R}_x \mathbf{M} \mathbf{R}_x = \mathbf{M} \quad \text{and} \quad \mathbf{R}_x \mathbf{K}(\mathbf{R}_x \cdot Z)\mathbf{R}_x = -\mathbf{K}(\mathbf{R}_x \cdot Z); \tag{A 11}$$

that is, $\mathbf{R}_x$ is symplectic with respect to $\mathbf{M}$ but anti-symplectic with respect to $\mathbf{K}(Z)$. Therefore acting on (A 10) with $\mathbf{R}_x$ results in

$$\mathbf{R}_x \mathbf{M} Z_t + \mathbf{R}_x \mathbf{K}(Z) Z_x = \mathbf{R}_x \nabla S(Z) = \nabla S(\mathbf{R}_x Z).$$

Using the properties (A 11) then results in

$$\mathbf{M}(\mathbf{R}_x Z)_t - \mathbf{K}(\mathbf{R}_x Z)(\mathbf{R}_x Z)_x = \nabla S(\mathbf{R}_x Z).$$

In other words, $\mathbf{R}_x Z(-x, t)$ is a solution of (A 10) whenever $Z(x, t)$ is a solution. The x-reversibility is a symplectic abstraction of the property that the Boussinesq equations admit two-way propagation of waves.

Appendix B: *Decomposing impulse and energy conservation*

When a dispersive wave system is restricted to one space dimension there is an interesting connection between the action and the impulse and the similarity leads to a decomposition of the impulse conservation law.

The properties of impulse are defined in Benjamin [**3**] with respect to a classical Hamiltonian structure. Here the properties of impulse in the classical setting and the multi-symplectic structures setting will be compared. Consider a Hamiltonian system on one space dimension and time with governing equation

$$\mathbf{M}(Z)Z_t + \mathbf{K}(Z)Z_x = \nabla S(Z) \quad Z \in \mathscr{M} \tag{B 1}$$

or in classical form

$$\mathbf{M}(Z)Z_t = \nabla E(Z) \stackrel{\text{def}}{=} \nabla S(Z) - \mathbf{K}(Z)Z_x. \tag{B 2}$$

The gradient of E in (B 2) is defined with respect to an inner product on $\mathscr{M}$ that also includes integration over x whereas the gradient of S in (B 1) is with respect to an inner product on the phase space $\mathscr{M}$. The skew-symmetric operators $\mathbf{K}(Z)$ and $\mathbf{M}(Z)$ in (B 1) and (B 2) are taken to be associated with exact two forms. Therefore we have the following densities for the action and action flux (cf. Section 1)

$$\begin{aligned} A(Z) &= \langle \alpha(Z), Z_t \rangle \\ B(Z) &= \langle \beta(Z), Z_x \rangle \end{aligned} \tag{B 3}$$

https://doi.org/10.1017/S0305004196001429 Published online by Cambridge University Press

where $\langle \cdot, \cdot \rangle$ is an inner product associated with the phase space $\mathscr{M}$ (that is, without integration over x or t) and for simplicity is taken to be independent of $m \in \mathscr{M}$. The one forms (B 3) generate the symplectic operators as follows

$$\begin{aligned} \mathbf{M}(Z) &= D\alpha(Z)^* - D\alpha(Z) \\ \mathbf{K}(Z) &= D\beta(Z)^* - D\beta(Z) \end{aligned} \tag{B 4}$$

where $D\alpha(Z)$ is the Jacobian with respect to Z and * indicates formal adjoint.

The impulse is defined for a classical Hamiltonian system in one space dimension as the functional associated, through Noether's theorem, with the translation invariance of the system in the x-direction (cf. Benjamin [**3**]) and satisfies

$$\mathbf{M}(Z)\, Z_x = \nabla I(Z). \tag{B 5}$$

The gradient $\nabla I(Z)$ in (B 5) is defined with respect to an inner product on $\mathscr{M}$ that also includes integration over x. The definition of action in (B 3) and its relation with $\mathbf{M}(Z)$ in (B 4) lead immediately to the following definition for the impulse.

Given a Hamiltonian system in one space dimension (either classical or on a multi-symplectic structure), with action density $A(Z) = \langle \alpha(Z), Z_t \rangle$, the impulse density is

$$I(Z) = \langle \alpha(Z), Z_x \rangle. \tag{B 6}$$

This follows since

$$\begin{aligned} \frac{d}{d\varepsilon} I(Z + \varepsilon\xi)\,|_{\varepsilon=0} &= \langle D\alpha(Z)\,\xi, Z_x \rangle + \langle \alpha(Z), \xi_x \rangle \\ &= \langle D\alpha(Z)^* Z_x, \xi \rangle + \frac{\partial}{\partial x} \langle \alpha(Z), \xi \rangle - \langle D\alpha(Z)\, Z_x, \xi \rangle \\ &= \frac{\partial}{\partial x} \langle \alpha(Z), \xi \rangle + \langle \mathbf{M}(Z)\, Z_x, \xi \rangle \end{aligned}$$

where we have used the assumption that the inner product does not depend on position in the phase space (although such dependence is easily incorporated). Taking variations ξ on an interval $[x_1, x_2]$ with fixed endpoint conditions on ξ and integrating $I(Z)$ over the interval recovers the required expression (B 5). Comparison of (B 3) and (B 6) shows that the action and impulse density share the same one-form.

The above definition of action and impulse leads to a useful decomposition of the impulse conservation law. Here the derivation is easiest in terms of the multi-symplectic structure. Differentiate the impulse density with respect to t,

$$\begin{aligned} I_t &= \langle D\alpha(Z)\, Z_t, Z_x \rangle + \langle \alpha(Z), Z_{xt} \rangle \\ &= \langle D\alpha(Z)^* Z_x - D\alpha(Z)\, Z_x, Z_t \rangle + \frac{\partial}{\partial x} \langle \alpha(Z), Z_t \rangle \\ &= \langle \mathbf{M}(Z)\, Z_x, Z_t \rangle + A_x \\ &= A_x - \langle Z_x, \mathbf{M}(Z)\, Z_t \rangle \quad \text{(skew-symmetry of } \mathbf{M}(Z)\text{)}. \end{aligned}$$

But substitution of (B 1) and the skew-symmetry of $\mathbf{K}(Z)$ lead to

$$I_t = A_x - \langle \nabla S(Z) - \mathbf{K}(Z)\, Z_x, Z_x \rangle = A_x - \langle \nabla S(Z), Z_x \rangle = A_x - S_x$$

and so

$$I_t + (S - A)_x = 0. \tag{B 7}$$

https://doi.org/10.1017/S0305004196001429 Published online by Cambridge University Press

In other words the flow force has two components; a static component, here denoted by S, and a dynamic component which is precisely $-A$ where A is the action density. The decomposition into a static and dynamic part is interesting for the multi-symplectic structures framework because, in one space dimension, it is the *static* part of the flow force that appears as the generalized Hamiltonian functional. Note that in higher space dimension the impulse conservation law is tensor valued and the generalized Hamiltonian functional for a multi-symplectic structure theory is no longer related in an elementary way to the flow force.

Similar arguments can be applied to the one-form for action flux in (B 3) leading to a decomposition of the energy conservation law. If $B(Z) = \langle \beta(Z), Z_x \rangle$ is the action-flux density then

$$F(Z) = \langle \beta(Z), Z_t \rangle \tag{B 8}$$

is the energy-flux density. To verify, note that the enegy density is defined, using (B 2), as $E(Z) = S(Z) - B(Z)$ and so

$$\begin{aligned} E_t &= S_t - B_t \\ &= \langle \nabla S(Z), Z_t \rangle - \langle D\beta(Z)\, Z_t, Z_x \rangle - \langle \beta(Z), Z_{xt} \rangle \\ &= \langle \mathbf{K}(Z)\, Z_x + \mathbf{M}(Z)\, Z_t, Z_t \rangle - \frac{\partial}{\partial x} \langle \beta(Z), Z_t \rangle - \langle \mathbf{K}(Z)\, Z_x, Z_t \rangle \\ &= -\frac{\partial}{\partial x} \langle \beta(Z), Z_t \rangle \\ &= -F_x \end{aligned}$$

resulting in the following decomposition of the energy conservation law

$$(S-B)_t + F_x = 0 \tag{B 9}$$

where B and F are related by (B 3) and (B 8) and S is the static part of the flow force. Combining (B 7) and (B 9) with (B 1) results in the interesting identities

$$S_t = B_t - F_x = \langle \mathbf{K}(Z)\, Z_x, Z_t \rangle \quad \text{and} \quad S_x = A_x - I_t = \langle \mathbf{M}(Z)\, Z_t, Z_x \rangle.$$

Appendix C. Multi-symplectic structure and conservation laws

An important part of the theory of conservative systems is the Noether theory that relates symmetries and conservation laws. When a conservative system has a Hamiltonian structure the symplectic operator gives a natural correspondence between symmetries and invariants or conserved densities (Olver [**28**, **29**]). However in the classical setting there is only one symplectic operator and therefore there is no relation between symplecticity and the fluxes of a conservation law. In the multi-symplectic structures framework such a connection is possible and leads to a new and useful decomposition of the Noether theory.

In its simplest setting, finite-dimensional Hamiltonian systems, the connection between symmetry and conservation laws can be stated as follows. Suppose $U \in \mathbb{R}^{2n}$ and

$$\mathbf{J} U_t = \nabla H(U) \tag{C 1}$$

with $\mathbf{J}$ the usual unit symplectic operator on $\mathbb{R}^{2n}$. Suppose there exists a one-parameter Lie group $G(\epsilon)$ acting symplectically on $\mathbb{R}^{2n}$ which leaves the Hamiltonian functional invariant. Let

$$V = \frac{d}{d\epsilon} [G(\epsilon)\, U] \Big|_{\epsilon=0} \tag{C 2}$$

https://doi.org/10.1017/S0305004196001429 Published online by Cambridge University Press

and suppose

$$\mathbf{J}V = \nabla P(U) \tag{C 3}$$

for some functional $P(U)$. Then $\partial_t P = 0$. A proof of this result for Hamiltonian evolution equations is given in Olver [**28**]. Below we will generalise the result to include fluxes.

The above result connects the action of the symmetry group with the conserved densities but in the case of conservation laws of Hamiltonian evolution equations it does not establish a connection between the action of the symmetry and the fluxes. In the expression (C 3) it is the action of the symplectic operator, on the infinitesimal action of the Lie group, that generates the gradient of the conserved quantity. Therefore its generalisation to include fluxes is clear: act on V with each element in the family of skew-symmetric operators in the multi-symplectic structure to obtain all the components of the conservation law. This leads to a new decomposition of the Noether theory.

Let $Z \in \mathscr{M}$ and consider the Hamiltonian system

$$\mathbf{M}(Z)\, Z_t + \mathbf{K}(Z)\, Z_x + \mathbf{L}(Z)\, Z_y = \nabla S(Z) \tag{C 4}$$

on a multi-symplectic structure $(\mathscr{M}, \omega^{(1)}, \omega^{(2)}, \omega^{(3)}, S)$. We say that the system (C 4) is equivariant with respect to the action of a one parameter Lie group $G(\epsilon)$ if

$$\begin{aligned} S(G(\epsilon).Z) &= S(Z) \\ DG(\epsilon)^*\, \mathbf{M}(G(\epsilon)\cdot Z)\, DG(\epsilon) &= \mathbf{M}(Z) \\ DG(\epsilon)^*\, \mathbf{K}(G(\epsilon)\cdot Z)\, DG(\epsilon) &= \mathbf{K}(Z) \\ DG(\epsilon)^*\, \mathbf{L}(G(\epsilon)\cdot Z)\, DG(\epsilon) &= \mathbf{L}(Z). \end{aligned} \tag{C 5}$$

Proposition C1 (Symmetries and fluxes). *Let $(\mathscr{M}, \omega^{(1)}, \omega^{(2)}, \omega^{(3)}, S)$ be a Hamiltonian system on multi-symplectic structure equivariant with respect to the action of a one-parameter Lie group $G(\epsilon)$ with generator V. Suppose there exists a solution $Z(x, y, t)$ satisfying (C 4) and functionals $P(Z)$, $Q(Z)$ and $R(Z)$ such that*

$$\mathbf{M}(Z)\, V = \nabla P(Z), \quad \mathbf{K}(Z)\, V = \nabla Q(Z) \quad \text{and} \quad \mathbf{L}(Z)\, V = \nabla R(Z). \tag{C 6}$$

Then P, Q and R, evaluated at the solution Z of (C 4), satisfy the conservation law

$$\frac{\partial P}{\partial t} + \frac{\partial Q}{\partial x} + \frac{\partial R}{\partial y} = 0.$$

Proof. The proof of this result is straightforward. Differentiating the first equation of (C 5) with respect to ϵ and setting to zero results in

$$\begin{aligned} 0 = \frac{d}{d\epsilon} S(G(\epsilon)\cdot Z)\Big|_{\epsilon=0} &= \langle \nabla S(Z), V \rangle_m \\ &= \langle \mathbf{M}(Z)\, Z_t + \mathbf{K}(Z)\, Z_x + \mathbf{L}(Z)\, Z_y, V \rangle_m \quad \text{(using (C 4))} \\ &= -\langle Z_t, \mathbf{M}(Z)\, V \rangle_m - \langle Z_t, \mathbf{K}(Z)\, V \rangle_m - \langle Z_t, \mathbf{L}(Z)\, V \rangle_m \\ &= -\langle Z_t, \nabla P \rangle_m - \langle Z_x, \nabla Q \rangle_m - \langle Z_y, \nabla R \rangle_m \quad \text{(using (C 6))} \\ &= -\frac{\partial P}{\partial t} - \frac{\partial Q}{\partial x} - \frac{\partial R}{\partial y} \end{aligned}$$

proving the claim. ∎

https://doi.org/10.1017/S0305004196001429 Published online by Cambridge University Press

Note that it is essential that the inner product not include integration over x, y or t and only the abstract property that $\mathbf{M}$, $\mathbf{K}$ and $\mathbf{L}$ are skew-adjoint is used (they can be dependent on Z for example). The inner product is one associated with the phase space $\mathscr{M}$ and its tangent space. Note that the statement of the above result requires changes if P, Q or R depends explicitly on x, y or t.

The above result generalizes the usual Noether theorem to include the relation between the fluxes and the action of the symmetry group. This connection is of interest when proving instability results for systems with mean flow effects which can be characterized as space-time drift along the a group orbit (cf. Bridges[**11**]).

As an example of Proposition C1 it will be applied to the mass conservation law for water waves. The symmetry, that generates the mass conservation law for water waves is perturbation of the value of the velocity potential (Benjamin & Olver[**5**]). Using the definition of Z in (4·4) the action of this symmetry group on $\mathscr{M}$ is

$$G(\epsilon)\,Z = Z + \epsilon V \quad \text{with} \quad V = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \tag{C 7}$$

where V is the infinitesimal generator. Using the definition of $\mathbf{M}$, $\mathbf{K}$ and $\mathbf{L}$ for water waves given in Section 6 we find

$$\mathbf{M}(Z)\,V = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \nabla P \quad \text{with} \quad P = \eta \tag{C 8a}$$

$$\mathbf{K}(Z)\,V = \begin{pmatrix} 0 \\ \mathbf{u} \\ 0 \\ 1 \\ 0 \end{pmatrix} = \nabla Q \quad \text{with} \quad Q = \int_{-h}^{\eta} u\,dz \tag{C 8b}$$

$$\mathbf{L}(Z)\,V = \begin{pmatrix} 0 \\ \mathbf{v} \\ 0 \\ 0 \\ 1 \end{pmatrix} = \nabla R \quad \text{with} \quad R = \int_{-h}^{\eta} v\,dz. \tag{C 8c}$$

Therefore by Proposition C1 it follows that $P_t + Q_x + R_y = 0$ which is the well-known mass conservation law for the water-wave problem. The identity (C 8a) is well known and was first obtained in Benjamin & Olver[**5**] but the identities (C 8b, c) are new.

Appendix D. Hamiltonian functionals

A property of the Hamiltonian functional in the classical case is that, when the system is autonomous, the value of the Hamiltonian functional is invariant over time. For a Hamiltonian system on a multi-symplectic structure the generalized

https://doi.org/10.1017/S0305004196001429 Published online by Cambridge University Press

Hamiltonian functional is not an invariant over time or space, even when the system is autonomous; that is, $S, \mathbf{M}, \mathbf{K}, \mathbf{L}$ do not depend explicitly on x, y or t. For the case of one space dimension this can be seen as follows

$$S_t = \langle \nabla S(Z), Z_t \rangle_m = \langle \mathbf{M}Z_t + \mathbf{K}Z_x, Z_t \rangle_m = \langle \mathbf{K}Z_x, Z_t \rangle_m$$

and similarly $$S_x = \langle \mathbf{M}Z_t, Z_x \rangle_m.$$

However the identity $S_{xt} = S_{tx}$ generates the conservation law

$$\frac{\partial}{\partial t}\langle \mathbf{M}Z_t, Z_x \rangle_m + \frac{\partial}{\partial x}\langle \mathbf{K}Z_t, Z_x \rangle_m = 0.$$

For the case of two space dimensions the Hamiltonian function satisfies more complex identities. First consider the following general result.

Proposition D1. *Let $S(x, y, t)$ be a scalar-valued twice-continuously differentiable function of $(x, y, t) \in \mathscr{U} \subset \mathbb{R}^3$ and define*

$$a_1(x, y, t) = \frac{\partial S}{\partial t}, \quad a_2(x, y, t) = \frac{\partial S}{\partial x} \quad \text{and} \quad a_3(x, y, t) = \frac{\partial S}{\partial y}. \tag{D 1}$$

Then the equation governing $\mathbf{a} = (a_1, a_2, a_3)$ is a Hamiltonian system on a multi-symplectic structure $(\mathbb{R}^3, \omega^{(1)}, \omega^{(2)}, \omega^{(3)})$ with trivial Hamiltonian functional and where $\omega^{(1)}$, $\omega^{(2)}$ and $\omega^{(3)}$ span the Lie algebra so(3).

Proof. Since $S \in \mathscr{C}^2$ we obtain the following identities from (D 1) simply by cross-differentiation

$$\frac{\partial a_2}{\partial y} - \frac{\partial a_3}{\partial x} = 0, \quad \frac{\partial a_2}{\partial t} - \frac{\partial a_1}{\partial x} = 0 \quad \text{and} \quad \frac{\partial a_3}{\partial t} - \frac{\partial a_1}{\partial y} = 0$$

or $$\mathbf{J}_1 \mathbf{a}_t + \mathbf{J}_2 \mathbf{a}_x + \mathbf{J}_3 \mathbf{a}_y = 0 \tag{D 2}$$

with

$$\mathbf{J}_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, \quad \mathbf{J}_2 = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad \text{and} \quad \mathbf{J}_3 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{D 3}$$

Closure of the two-forms is immediate since they are constant. The Lie group $SO(3)$ is the group of orthogonal matrices on $\mathbb{R}^3$ with unit determinant. It is a three parameter group with Lie algebra spanned by $\{\mathbf{J}_1, \mathbf{J}_2, \mathbf{J}_3\}$ (Olver [**29**, p. 48]).

The vector $\mathbf{a}$ is in the normal space for level surfaces of $S(x, y, t)$ and (D 2) is the transport equation for $\mathbf{a}$. Proposition D1 expresses the vector identity, that the gradient operator is in the kernel of the curl operator, in terms of multi-symplectic structure. Application of Proposition D1 to the Hamiltonian functional S leads to

$$\begin{aligned} a_1 &= -S_t = -\langle \mathbf{K}Z_x, Z_t \rangle_m - \langle \mathbf{L}Z_y, Z_t \rangle_m \\ a_2 &= S_x = \langle \mathbf{M}Z_t, Z_x \rangle_m + \langle \mathbf{L}Z_y, Z_x \rangle_m \\ a_3 &= S_y = \langle \mathbf{M}Z_t, Z_y \rangle_m + \langle \mathbf{K}Z_x, Z_y \rangle_m \end{aligned}$$

with transport equation (D 2). It follows that the functional S is an invariant over space and time only under exceptional circumstances. A special case in which S is invariant is when Z_t, Z_x and Z_y are collinear (for example travelling waves).

Special cases of a Hamiltonian system on a multi-symplectic structure are

https://doi.org/10.1017/S0305004196001429 Published online by Cambridge University Press

obtained as follows. For simplicity of discussion consider the system restricted to one space dimension $(\mathscr{M}, \omega^{(1)}, \omega^{(2)}, S)$ with governing equation

$$\mathbf{M}(Z)\,Z_t + \mathbf{K}(Z)\,Z_x = \nabla S(Z). \tag{D 4}$$

The classic Hamiltonian structure for the system, denoted $(\mathscr{M}, \omega^{(1)}, \mathrm{H})$, is recovered by taking

$$\mathbf{M}Z_t = \nabla H(Z, Z_x) \quad \text{with} \quad H(Z, Z_x) = \int_{x_1}^{x_2} [S(Z) - B(Z)]\,dx.$$

This corresponds to a classic Hamiltonian structure for the time-evolution equation in terms of the larger set of corrdinates $Z \in \mathscr{M}$. The difficulty with analysing this formulation when $x \in \mathbb{R}$ is that one must specify the function space over x. On the other hand the roles of space and time can be reversed with the system $(\mathscr{M}, \omega^{(2)}, \tilde{H})$ with governing equation

$$\mathbf{K}Z_x = \nabla \tilde{H}(Z, Z_t) \quad \text{with} \quad \tilde{H}(Z, Z_t) = \int_{t_1}^{t_2} [S(Z) - A(Z)]\,dt. \tag{D 5}$$

In this case the function space over time (say periodic functions) is specified in advance and then one seeks all bounded solutions (in x) of the spatial 'evolution' equation. A special case is the analysis of time-independent states with evolution equation $\mathbf{K}(Z)\,Z_x = \nabla S(Z)$.

A third possibility is to consider solutions of (D 4) that are stationary relative to a moving frame. Let $\theta = x - ct$ then (D 4) becomes

$$(\mathbf{K} - c\mathbf{M})\,Z_\theta = \nabla S(Z). \tag{D 6}$$

The Hamiltonian structure of spatial 'evolution' equations of the form (D 5)–(D 6) are studied in Benjamin [**3**], Mielke [**26**], Baesens & MacKay [**2**] and Bridges [**7**, **11**]. For the system in (D 6) the functional S is an absolute spatial (θ) invariant,

$$\partial_\theta S = \langle \nabla S(Z), Z_\theta \rangle_m = \langle (\mathbf{K} - c\mathbf{M})\,Z_\theta, Z_\theta \rangle_m = 0$$

because $\mathbf{K} - c\mathbf{M}$ is skew-symmetric and S in this case can be identified with the flow force for the system.

The case of two-space dimensions (cf. (1·12)) contains further special cases with evolution in various spatial directions. Finally we note that the family of symplectic operators $\omega^{(1)}, \ldots, \omega^{(n)}$ is, to be precise, a basis for the multi-symplectic structure. For example one can introduce a non-degenerate linear transformation $\mathbf{T}$ taking $\omega^{(1)}, \ldots, \omega^{(n)}$ to a new basis $\Omega^{(1)}, \ldots, \Omega^{(n)}$.

Acknowledgements. The research and writing of this work was partially carried out while the author was supported by a Research Fellowship from the Alexander von Humboldt Foundation held at the University of Stuttgart during the 1993–4 academic year.

REFERENCES

[**1**] D. G. ANDREWS and M. E. MCINTYRE. On wave-action and its relatives. *J. Fluid Mech.* **89** (1978), 647–664.

[**2**] C. BAESENS and R. S. MACKAY. Uniformly travelling water waves from a dynamical systems viewpoint: some insights into bifurcations from Stokes' family. *J. Fluid Mech.* **241** (1992), 33–47.

https://doi.org/10.1017/S0305004196001429 Published online by Cambridge University Press

[3] T. B. Benjamin. Impulse, flow force and variational principles. *IMA J. Applied Math.* **32** (1984), 3–68.
[4] T. B. Benjamin and J. E. Feir. The disintegration of wavetrains on deep water. Part 1. Theory, *J. Fluid Mech.* **27** (1967), 417–430.
[5] T. B. Benjamin and P. J. Olver. Hamiltonian structure, symmetries and conservation laws for water waves. *J. Fluid Mech.* **125** (1982), 137–185.
[6] D. J. Benney and G. J. Roskes. Wave instabilities. *Stud. Appl. Math.* **48** (1969), 377–385.
[7] T. J. Bridges. Spatial Hamiltonian structure, energy flux and the water-wave problem. *Proc. Roy. Soc. London, Ser A* **439** (1992), 297–315.
[8] T. J. Bridges. Hamiltonian bifurcations of the spatial structure for coupled nonlinear Schrödinger equations. *Physica D* **57** (1992), 375–394.
[9] T. J. Bridges. Hamiltonian structure of plane wave instabilities, *Fields Inst. Comm.* **8** (1996), 19–33.
[10] T. J. Bridges. Hamiltonian spatial structure for three-dimensional water-waves in a moving frame of reference. *J. Nonlinear Science* **4** (1994), 221–251.
[11] T. J. Bridges. Periodic patterns, linear instability, symplectic structure and mean-flow dynamics for 3D surface waves. *Philos. Trans. Roy. Soc. London, Ser. A* **354** (1996), 533–574.
[12] T. J. Bridges. A geometric formulation of the conservation of wave action and its implications for signature and the classification of instabilities, preprint.
[13] T. J. Bridges and A. Mielke. A proof of the Benjamin–Feir instability. *Arch. Rational Mech. Anal.* **133** (1995), 145–198.
[14] T. J. Bridges and G. Rowlands. Instability of spatially quasiperiodic states of the Ginzburg–Landau equation. *Proc. Roy. Soc. London, Ser. A* **444** (1994), 347–362.
[15] L. J. F. Broer. On the Hamiltonian theory of surface waves. *Appl. Sci. Res.* **29** (1974), 430–446.
[16] B. Chen and P. G. Saffman. Three-dimensional stability and bifurcation of capillary and gravity waves on deep water. *Stud. Appl. Math.* **72** (1985), 125–147.
[17] R. Grimshaw. Wave action and wave-mean flow interaction with application to stratified shear flows. *Ann. Rev. Fluid Mech.* **16** (1984), 11–44.
[18] W. D. Hayes. Conservation of action and modal wave action. *Proc. Roy. Soc. London Ser. A* **320** (1970), 187–208.
[19] W. D. Hayes. Group velocity and nonlinear dispersive wave propagation. *Proc. Roy. Soc. London Ser. A* **332** (1973), 199–221.
[20] P. Kuchment. Floquet theory for partial differential equations (Birkhäuser, Verlag: Basel 1993).
[21] M. J. Lighthill. Contributions to the theory of waves in nonlinear dispersive systems. *J. Inst. Math. Appl.* **1** (1965), 269–306.
[22] M. J. Lighthill. Some special cases treated by the Whitham theory. *Proc. Roy. Soc. London Ser. A* **299** (1967), 28–53.
[23] R. S. MacKay and P. G. Saffman. Stability of water waves. *Proc. Roy. Soc. London Ser. A* **406** (1986), 115–125.
[24] J. H. Maddocks. On second-order conditions in constrained variational principles. *J. Optim. Theory Appl.*, to appear.
[25] J. E. Marsden. *Applications of global analysis in mathematical physics*, Math. Lect. Ser. **2** (Publish or Perish, 1974).
[26] A. Mielke. *Hamiltonian and Lagrangian flows on center manifolds with applications to elliptic variational problems.* Lect. Notes in Math. **1489** (Springer-Verlag, 1991).
[27] P. K. Newton and J. B. Keller. Stability of periodic plane waves. *SIAM J. Appl. Math.* **47** (1987), 959–964.
[28] P. J. Olver. On the Hamiltonian structure of evolution equations. *Math. Proc. Cambridge Philos. Soc.* **88** (1980), 71–88.
[29] P. J. Olver. *Applications of Lie groups to differential equations.* Graduate Texts in Math. **107** (Springer-Verlag, 1986).
[30] I. C. Percival. Variational principles for the invariant toroids of classical dynamics. *J. Phys. A: Math., Nucl. Gen.* **7** (1974), 794–802.
[31] P. G. Saffman. The superharmonic instability of finite-amplitude water waves. *J. Fluid Mech.* **159** (1985), 169–174.
[32] C. E. Wayne. *Periodic and quasiperiodic solutions of nonlinear wave equations via KAM theory. Comm. Math. Phys.* **127** (1990), 479–528.

https://doi.org/10.1017/S0305004196001429 Published online by Cambridge University Press

[**33**] G. B. WHITHAM. A general approach to linear and nonlinear dispersive waves using a Lagrangian. *J. Fluid Mech.* **22** (1965), 273–283.
[**34**] G. B. WHITHAM. Nonlinear dispersion of water-waves. *J. Fluid Mech.* **27** (1967), 399–412.
[**35**] G. B. WHITHAM. Linear and nonlinear waves (Wiley-Interscience, 1974).
[**36**] V. E. ZAKHAROV. Stability of periodic waves of finite amplitude on the surface of a deep fluid. *Zh. Prikl. Mekh. Fiz.* **9** (1968), 86–94 (English trans: *J. Appl. Mech. Tech. Phys.* **2** (1968), 190).

https://doi.org/10.1017/S0305004196001429 Published online by Cambridge University Press