NEW VARIATIONAL PRINCIPLES FOR SYSTEMS OF PARTIAL DIFFERENTIAL EQUATIONS

Thesis by

John R. Myers

In Partial Fulfillment of the Requirements

For the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California 91109

1972

(submitted May 22, 1972)

ACKNOWLEDGMENTS

The author wishes to thank Professor G. B. Whitham for suggesting the problem and also for many stimulating discussions. He has given generously of his time. His incredible physical insight has been invaluable.

The author wishes to thank Alrae Tingley for the painstaking work of typing this thesis.

He is grateful for a fellowship awarded under Title IV of the National Defense Education Act and for a National Science Foundation Traineeship. Also, the California Institute of Technology supported some of this work with various tuition scholarships and assistantships.

ABSTRACT

The question of finding variational principles for coupled systems of first order partial differential equations is considered. Using a potential representation for solutions of the first order system a higher order system is obtained. Existence of a variational principle follows if the original system can be transformed to a self-adjoint higher order system. Existence of variational principles for all linear wave equations with constant coefficients having real dispersion relations is established. The method of adjoining some of the equations of the original system to a suitable Lagrangian function by the method of Lagrange multipliers is used to construct new variational principles for a class of linear systems. The equations used as side conditions must satisfy highly-restrictive integrability conditions. In the more difficult nonlinear case the system of two equations in two independent variables can be analyzed completely. For systems determined by two conservation laws the side condition must be a conservation law in addition to satisfying the integrability conditions.

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION

Variational principles have proved of great value in the mathematical analysis of physical problems. The history of mechanics as an exact science is almost inseparable from the history of the two classical variational principles of Lagrange and Hamilton. The canonical form allows general discussion of physical systems. Furthermore, by providing an alternate statement of a problem the variational principle often allows one to find approximate solutions to problems which would be intractable by a straightforward application of Newton's laws.

In many classical fields, the appropriate variational principle is "well-known," but the question of how one finds a variational principle for a given system seems to have received very little discussion. This question has arisen frequently in recent years and is the topic of this thesis. A number of cases have arisen in fluid mechanics. There the issue revolves around the use of Eulerian coordinates. Hamilton's principle is easily formulated in the Lagrangian description. But the Eulerian description is usually preferable and for this reason variational principles have not been used extensively in fluid mechanics.

In 1945, John von Neumann [1] suggested that increased use of variational principles in hydrodynamics might prove as fruitful as their use had been in problems of Maxwellian electrodynamics and quantum mechanics noting that in contrast to the original equations

of motion "the great virtue of the variational treatment ... is that it permits efficient use, in the process of calculation, of any experimental or intuitive insight which one may possess concerning the problem which is to be solved by calculation."

Recent advances in the theory of non-linear dispersive wave equations rely heavily on variational principles through the method of averaged Lagrangians [2], [3], [4], [5]. The generality of the theory requires the existence of a variational principle for all systems of appropriate type and at first it was quite a challenge to find them. The question of existence is still very much open; despite intensive research, an automatic method to provide variational principles for given systems is not available.

The general mathematical problem is to decide when a given partial differential equation or system of partial differential equations is equivalent to a variational principle. The difficulty is that some change of variables or other transformation may be necessary before the system is in a form for which a variational principle may be found. The change of variables is usually not just algebraic but involves "potential representations" in which the original variables are expressed in terms of "potentials" and their derivatives.

For a single equation the question is relatively easy to investigate. For a linear equation with constant coefficients, for example, it is easily seen that the equation should contain only even derivatives or only odd derivatives. (The details are discussed in the first few pages of Chapter II.) Thus, for linear systems with constant coef-

ficients, one could eliminate variables and reduce the system to a single equation, and decide the question there. For non-linear systems, however, this is not always possible and, even when it is, usually gives a very awkward description. Moreover, the overwhelming evidence from known examples indicates one should work with a system but choose appropriate potentials. The potentials reduce the number of equations in the system by roughly a half, with a corresponding increase in the order of the derivatives, but there is by no means a full reduction to a single equation. Accordingly, even for linear systems, the question of potential representations and side conditions is the central theme. But the existence of a variational principle for the system may be settled by the reduction to a single equation and then tied in with a more direct attack on the system. The information learned in linear systems is used in the investigation of non-linear systems.

Of course a possibility for any given system would be to minimize the sum of the squares of the partial differential equations. That is, if the given system is

$$a[\underline{u}] = 0 \tag{1.1}$$

we may consider the variational principle

$$\delta J = \delta \int_R (a[\underline{u}])^2 dV = 0 .$$

Certainly J assumes its minimum value, zero, at solutions of (1.1). The trouble with this approach is that the corresponding Euler

equations are $a^*a[\underline{u}] = 0$, where a^* is the adjoint operator to a. (The adjoint operator appears in performing the integration by parts in obtaining the Euler equations.) The weakness of this seems to be that it imbeds the original system in the larger system $a^*a[\underline{u}] = 0$, which has solutions that are not solutions of $a[\underline{u}] = 0$. For many purposes this turns out to be unsatisfactory, although it has been much used in elasticity under the title of the "method of least squares."

It is well known that the Euler equations of a variational principle are self-adjoint. Conversely all self-adjoint systems of equations are equivalent to variational principles. The trouble stems from the fact that most non-linear problems are posed as a system of coupled first order equations which is usually not self-adjoint as it stands, and what is needed is a suitable transformation of the equations so that the transformed equations are self-adjoint.

As noted above, for problems of continuum mechanics a possible variational principle is Hamilton's Principle written in Lagrangian coordinates. With this description the continuum corresponds directly to a system of particles. But usually the Eulerian description is more useful. Moreover, in problems involving the interaction of the continuum with another field (plasma physics, for example) the Eulerian description is essential. In the Eulerian description the variational principles which have been found for problems in continuum mechanics have in common this feature: the Lagrangian function (written in Eulerian variables) is similar to the

one used in Hamilton's Principle but certain of the equations of the system under consideration have to be adjoined as side conditions with Lagrange multipliers. When the field quantities are referred to Eulerian coordinates the use of Hamilton's Principle by itself is inadequate; the side conditions are necessary to restrict the class of functions and lead to non-trivial flows.

Variational principles of this type have been found for problems of ideal fluid flow by Herivel [6]. Lin [7] proposed a modification of this variational principle to include rotational flows. More recently, Seliger and Whitham [8] have reconsidered Herivel's work noting that after an integration by parts Herivel's Lagrangian function is nothing but an expression for the pressure. Seliger and Whitham go on to find new variational principles for problems of rotating fluids, stratified fluids, and plasmas. Schutz [9] has extended these results to relativistic fluids. Extensive bibliographies may be found in the article by Seliger and Whitham and also in the review article by Serrin [10].

Upon analysis it is found that all variational principles of this type introduce a potential representation for solutions of the original system. The Lagrange multipliers become the potentials. It is often the case that the potential representation generated by the side condition technique has been used previously for some other special application but the full importance of the transformations to these potential coordinates has rarely been recognized. In terms of this potential representation the original system is self-adjoint. For

inviscid fluid mechanics, for example, the variational principle introduces the Clebsch representation. (For an account of the Clebsch representation see Lamb [11].)

Generally speaking there is no effective way to choose a suitable Lagrangian function nor is there an effective procedure for choosing which equations are to be used as side conditions, although some authors mention that it is easier to work with side conditions in conservation law form. Even the number of side conditions required is not known although Seliger and Whitham argue by analogy to Pfaff's theorem in the theory of differential forms that the number of side conditions for a system with n equations should be at most $\frac{n}{2}$ if n is even and $\frac{n+1}{2}$ if n is odd.

Maxwell's equations provide one of the simplest examples of the appropriate potential representation and its connection with side conditions. Usually the scalar and vector potentials of electricity and magnetism are introduced in order to satisfy certain of Maxwell's equations identically. However these potentials may also be obtained naturally by means of a variational principle with side conditions.

In order to keep the exposition as simple as possible we will take Maxwell's equations for free space. They may be written

$$\frac{\partial \underline{B}}{\partial t} + \nabla \times \underline{E} = \underline{0} \tag{1.2}$$

$$\frac{\partial \underline{E}}{\partial t} - \nabla \times \underline{B} = \underline{0} \tag{1.3}$$

$$\nabla \cdot \underline{B} = 0 \tag{1.4}$$

$$\nabla \cdot \underline{E} = 0 \, . \tag{1.5}$$

The usual form of the Lagrangian function for electricity and magnetism is the difference of the energy densities stored in the electric and in the magnetic fields

$$L = \tfrac{1}{2} \underline{B}^2 - \tfrac{1}{2} \underline{E}^2 \, .$$

Some restrictions on the variations are necessary, otherwise the stationary values are merely $\underline{B} = \underline{E} = \underline{0}$. If the variations are required to satisfy equations (1.3) and (1.5), these two equations may be adjoined as side conditions and we take

$$\mathfrak{L} = \tfrac{1}{2} \underline{B}^2 - \tfrac{1}{2} \underline{E}^2 + \underline{A} \cdot \left(\frac{\partial \underline{E}}{\partial t} - \nabla \times \underline{B} \right) + \varphi \nabla \cdot \underline{E}$$

where $\underline{A}$ and φ are Lagrange multipliers. Now consider the result of varying J where

$$J = \int_{T_1}^{T_2} \int_S \mathfrak{L} \, \underline{dx} \, dt \, .$$

Variations with respect to $\underline{E}$, $\underline{B}$, $\underline{A}$, and φ give, respectively

$$\begin{aligned} \partial \underline{E}: \quad & \underline{E} = - \frac{\partial \underline{A}}{\partial t} - \nabla \varphi \\ \partial \underline{B}: \quad & \underline{B} = \nabla \times \underline{A} \end{aligned} \tag{1.6}$$

$$\partial \underline{A}: \quad \frac{\partial \underline{E}}{\partial t} - \nabla \times \underline{B} = \underline{0}$$

$$\partial \varphi: \quad \nabla \cdot \underline{E} = 0 \, . \tag{1.7}$$

The field quantities $\underline{E}$ and $\underline{B}$ are obtained in terms of the usual potentials but they are introduced here as Lagrange multipliers. The expressions (1.6) satisfy identically the equations (1.2) and (1.4), which were not used as side conditions. For, we obtain

$$\frac{\partial \underline{B}}{\partial t} + \nabla \times \underline{E} = \frac{\partial}{\partial t}(\nabla \times \underline{A}) + \nabla \times (- \frac{\partial}{\partial t}\underline{A} - \nabla \varphi)$$

$$\equiv \underline{0}$$

and

$$\nabla \times \underline{B} = \nabla \cdot (\nabla \times \underline{A})$$

$$\equiv 0 \, .$$

The result of substituting the potential representation (1.6) into the equations (1.7) is a self-adjoint system of second order partial differential equations, i.e.

$$\underline{0} = - \frac{\partial^2}{\partial t^2} \underline{A} - \frac{\partial}{\partial t} \nabla \varphi - c^2 \nabla \times \nabla \times \underline{A}$$

and

$$0 = - \nabla \frac{\partial \underline{A}}{\partial t} - \nabla^2 \varphi \, .$$

We may now obtain the variational principle in terms of potentials. As we shall see, this is a consequence of the fact that

the potentials satisfy a self-adjoint system of equations. In terms of potentials, the Lagrangian function $\mathfrak{L}$ becomes

$$\mathfrak{L} = \tfrac{1}{2}\underline{B}^2 - \tfrac{1}{2}\underline{E}^2 + \underline{A}\cdot\left(\frac{\partial \underline{E}}{\partial t} - \nabla\times\underline{B}\right) + \varphi\nabla\cdot\underline{E}$$

$$= \tfrac{1}{2}(\nabla\times\underline{A})^2 - \tfrac{1}{2}\left(\frac{\partial}{\partial t}\underline{A} + \nabla\varphi\right)^2$$

$$+ \underline{A}\cdot\left(-\frac{\partial^2}{\partial t^2}\underline{A} - \nabla\frac{\partial}{\partial t}\varphi - \nabla\times\nabla\times\underline{A}\right)$$

$$- \varphi\left(\nabla\cdot\frac{\partial}{\partial t}\underline{A} + \nabla\cdot\nabla\varphi\right).$$

Integrating by parts this becomes

$$\mathfrak{L}^I = \tfrac{1}{2}(\nabla\times\underline{A})^2 - \tfrac{1}{2}\left(\frac{\partial \underline{A}}{\partial t} + \nabla\varphi\right)^2 + \left(\frac{\partial \underline{A}}{\partial t}\right)^2$$

$$+ 2\frac{\partial \underline{A}}{\partial t}\cdot\nabla\varphi - (\nabla\times\underline{A})^2 + (\nabla\varphi)^2$$

$$= -\tfrac{1}{2}(\nabla\times\underline{A})^2 + \tfrac{1}{2}\left(\frac{\partial \underline{A}}{\partial t} + \nabla\varphi\right)^2.$$

In terms of $\underline{E}$ and $\underline{B}$ the integrated Lagrangian $\mathfrak{L}^I$ is simply the negative of the original Lagrangian

$$\mathfrak{L}^I = -L.$$

This property is common to all variational principles for linear partial differential equations with side conditions but for non-linear problems the dual formulation gives an entirely different variational principle.

We may now see the general form of the method. Consider a system of conservation laws

$$G_i[\underline{u}] \equiv \frac{\partial}{\partial x_j} g_{ij}(\underline{u}) = 0 \qquad i = 1,2,\ldots,n$$

where $\underline{u}$ is an n-vector. Suppose a variational principle with some of the equations as side conditions exist and let the first m equations be the side conditions. We now vary

$$J = \int \{L(\underline{u}) + \lambda_\alpha G_\alpha[\underline{u}]\}\, \underline{dx} \tag{1.8}$$

where L is a suitable Lagrangian function, the λ_α's are Lagrange multipliers and α is understood to range from 1 to m. Variations with respect to $\underline{u}$ give the Euler equations

$$\delta u: \quad \frac{\partial L}{\partial u_j} = \frac{\partial \lambda_\alpha}{\partial x_i} \frac{\partial g_{\alpha i}}{\partial u_j} \tag{1.9}$$

and the system is completed by varying the side conditions

$$\delta\lambda_\alpha: \quad G_\alpha[\underline{u}] = 0 \qquad \alpha = 1,2,\ldots,m \tag{1.10}$$

to provide m + n equations for the m λ_α's and n u_i's. The n equations (1.9) should be solvable in general to express $\underline{u}$ in terms of $\partial\lambda_\alpha/\partial x_i$ giving a potential representation of $\underline{u}$; this would be substituted in (1.10) to give m equations for the λ_α's. The form of (1.9) is reminiscent of statements of Pfaff's theorem and this fact has led to speculation that the maximum number of potentials (or equivalently, side conditions) necessary is either n/2 or (n+1)/2.

If we integrate (1.8) by parts and eliminate $\underline{u}$ in favor of $\partial\underline{\lambda}/\partial\underline{x}$ we obtain a variational principle entirely in terms of $\underline{\lambda}$ and free of side conditions.

$$J = \int \left\{ L - g_{\alpha i} \frac{\partial \lambda_\alpha}{\partial x_i} \right\} \underline{dx} \tag{1.11}$$

where the potential representation

$$\underline{u} = \underline{U}\left(\frac{\partial \underline{\lambda}}{\partial \underline{x}}\right) \tag{1.12}$$

is determined from (1.10) and used in the arguments of L and $g_{\alpha i}$. Of course this dual form is usually of physical interest. The Euler equations arising from the variation of $\underline{\lambda}$ in this dual variational principle together with the potential representation (1.12) for $\underline{u}$ in terms of $\partial\underline{\lambda}/\partial\underline{x}$ form a system equivalent to the Euler equations (1.10) of the original variational principle. The Euler equations of the dual are

$$\delta\lambda_\alpha: \quad 0 = -\frac{\partial}{\partial x_j}\left[\left(\frac{\partial L}{\partial u_i} - \frac{\partial g_{\beta k}}{\partial u_i}\frac{\partial \lambda_\beta}{\partial x_k}\right)\frac{\partial U_i}{\partial \lambda_\alpha}\right] + \frac{\partial g_{\alpha i}}{\partial x_i} \quad .$$

The expression in parentheses vanishes when the potential representation (1.12) is used so we are left with

$$G_\alpha\left[\underline{U}\left(\frac{\partial \underline{\lambda}}{\partial \underline{x}}\right)\right] = 0 \qquad \alpha = 1, 2, \ldots, m \ .$$

This is a system of m second order partial differential equations for λ. That is, the result of varying the potentials in (1.11) is the same as varying the Lagrange multipliers in (1.8) and the two forms of the variational principle are entirely equivalent.

CHAPTER II

LINEAR SYSTEMS WITH CONSTANT COEFFICIENTS

II.1. Systems derived from a variational principle

We begin the study of systems of linear partial differential equations with constant coefficients by considering the special properties of equations which are derived from variational principles. For a linear system the Lagrangian is quadratic in the dependent variables and their derivatives.

We consider, as a typical case, a variational principle for functions $\underline{u}$ of two independent variables x_1 and x_2 and take

$$J = \int \mathfrak{L}[\underline{u}]\ \underline{dx}$$

where

$$\begin{aligned}\mathfrak{L}[\underline{u}] = {} & \underline{u}_{x_1}^T \psi_{11} \underline{u}_{x_1} + \underline{u}_{x_1}^T \psi_{12} \underline{u}_{x_2} \\ & + \underline{u}_{x_\psi}^T \psi_{22} \underline{u}_{x_2} + \underline{u}_{x_1}^T \psi_{13} \underline{u} \\ & + \underline{u}_{x_2}^T \psi_{23} \underline{u} + \underline{u}^T \psi_{33} \underline{u} \ . \end{aligned} \tag{1.1}$$

Here $\underline{u}$ is an n-vector and the ψ_{ij} are constant $n \times n$ matrices. Upon varying $\underline{u}$ we obtain the Euler equation

$$\begin{aligned}\delta\underline{u}: \quad & (\psi_{11} + \psi_{11}^T)\underline{u}_{x_1 x_1} + (\psi_{12} + \psi_{12}^T)\underline{u}_{x_1 x_2} \\ & + (\psi_{22} + \psi_{22}^T)\underline{u}_{x_2 x_2} + (\psi_{13} - \psi_{13}^T)\underline{u}_{x_1} \\ & + (\psi_{23} - \psi_{23}^T)\underline{u}_{x_2} - (\psi_{33} + \psi_{33}^T)\underline{u} = 0 \end{aligned} \tag{1.2}$$

where ψ_{ij}^T denotes the transpose of the matrix ψ_{ij}. It is immediately obvious that systems of second order equations go most naturally with variational principles involving a function and its first derivatives. This confirms the notion that a system of first order equations should be reduced somehow to roughly half the number of second order equations. It is true that by taking $\psi_{11} = \psi_{12} = \psi_{22} = 0$, i.e. omitting the quadratic terms in derivatives in $\mathcal{L}$, that a first order system results, but this restriction on $\mathcal{L}$ looks artificial. In any event, we see that the coefficients in (1.2) have quite special properties. The second derivatives must have symmetric coefficient matrices, the first derivatives have skew-symmetric coefficients and the undifferentiated terms have again symmetric coefficients. Indeed, one may note that there is no loss of generality in (1.1) if we take ψ_{11}, ψ_{12}, ψ_{22}, ψ_{33} to be symmetric and ψ_{13}, ψ_{23}, skew symmetric from the start.

Similar conditions follow for quadratic variational principles including higher derivatives and more independent variables in the Lagrangian. The result is that the Euler equations of such variational principles must have symmetric coefficient matrices for the even order derivative terms (including the undifferentiated term) while the coefficients of the odd order derivative terms must be skew-symmetric. Conversely, given any system satisfying these symmetry conditions a variational principle is immediately obvious.

For convenience in the algebraic manipulation let X_i represent the differential operator $\partial/\partial X_i$. For example,

$$X_1^2 X_2^4 \underline{u} \longleftrightarrow \frac{\partial^6 \underline{u}}{\partial X_1^2 X_2^4} \ .$$

Then, each term

$$(X_1^{n_1} X_2^{n_2} \ldots X_\ell^{n_\ell} \underline{u})^T \psi (X_1^{p_1} X_2^{p_2} \ldots X_\ell^{p_\ell} u) \ .$$

In a quadratic Lagrangian function corresponds to a term

$$[(-1)^{n_1+n_2+\ldots+n_\ell} \psi + (-1)^{p_1+p_2+\ldots+p_\ell} \psi^T] (X_1^{n_1+p_1} X_2^{n_2+p_2} \ldots X^{n_\ell+p_\ell} \underline{u})$$

in the Euler equations. The coefficient matrix (in brackets) is symmetric if

$$n_1 + p_1 + n_2 + p_2 + \ldots + n_\ell + p_\ell$$

is even, antisymmetric if it is odd. (We are considering only real variables; if complex variables are allowed the coefficients are Hermitian and skew-Hermitian respectively.) In the particular case of a scalar function $u(X_1, X_2)$, the variational principle

$$\delta \int\int \sum A_{ij} \frac{\partial^i u}{\partial X_1^i} \frac{\partial^j u}{\partial X_2^j}$$

has the Euler equation

$$\sum [(-1)^i + (-1)^j] A_{ij} \frac{\partial^{i+j} u}{\partial X_1^i \partial X_2^j} \ .$$

Therefore only even derivatives occur. Conversely, for a single equation with even derivatives one can immediately write down the variational principle. If only odd derivatives occur in an equation we

can take $u = \partial\varphi/\partial X_1$ to make it even and again find a variational principle.

These particular conditions for the existence of a variational principle for a linear system are extremely restrictive. Relatively simple systems do not meet them, yet appropriate variational principles can be found.

In time dependent problems, for example, one would naturally pose the system as

$$\underline{u}_t + b\underline{u}_x = \underline{0}$$

and the unit matrix for the t-derivative immediately means that these are not the Euler equations of a variational principle as they stand. Of course we then make a transformation

$$\underline{u} = r\underline{v}$$

and multiply on the left by ℓ to obtain

$$\ell r\underline{v}_t + \ell br\underline{v}_x = \underline{0} \ .$$

If the matrices ℓr and ℓbr are skew-symmetric the variational principle is obvious. Even in the case of only two equations such a transformation is only possible when b is a constant multiple of the unit matrix, and in that case the two equations uncouple. Such a trivial system is of no great interest. Later in this chapter we will construct variational principles free of this extreme limitation.

What is needed is some method of transforming the first order

system to some other system for which a variational principle is available. We shall first consider ways of transforming a system of n coupled first order equations to a single n^{th} order partial differential equation.

II. 2. Reduction to a single n^{th} order equation

For the case of a single n^{th} order partial differential equation the preceding result was that variational principles exist for equations which contain only even-order derivatives or only odd-order derivatives. Since we are interested in results for first order systems of equations, we consider ways to transform the first order system to a single n^{th} order equation. Then if the resulting equation contains only even derivative terms (or only odd derivative terms), a variational principle follows immediately.

Consider the linear first order partial differential equation

$$A[\underline{u}] = 0 \tag{2.1}$$

where $\underline{u}$ is a continuously differentiable n-vector $\underline{u}(x_1, \ldots, x_\ell)$, A is an operator of the form

$$A = a_1 X_1 + a_2 X_2 + \ldots + a_\ell X_\ell + b \tag{2.2}$$

and $a_1, \ldots, a_\ell$, b are constant $n \times n$ matrices. Operating on the left by A^*, the transpose of the cofactor matrix of A, we find that each component of $\underline{u}$ satisfies the n^{th} order partial differential equation

$$(\det A)[u_i] = 0 \qquad i = 1,\ldots,n\ . \tag{2.3}$$

Here, the operators $X_1,\ldots,X_\ell$ appear as entries in the determinant, but on expansion we have an n^{th} order partial differential operator.

For example, if we take the system

$$\underline{u}_t + \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix} \underline{u}_x + \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \underline{u} = \underline{0} \tag{2.4}$$

which may be written out as

$$\begin{aligned} u_{1_t} + u_{2_x} + u_3 &= 0 \\ u_{2_t} + u_{3_x} + u_4 &= 0 \\ u_{3_t} + u_{4_x} - u_1 &= 0 \\ u_{4_t} + u_{1_x} - u_2 &= 0 \end{aligned} \tag{2.5}$$

the transposed cofactor matrix has the first row

$$(T^3 - X^2 + T,\ -T^2X + X,\ X^2T - T^2 - 1,\ -X^3 + 2TX).$$

If we apply these operators to the successive equations in (2.5) we have

$$\{T^4 + 2T^2 + 1 - X^4\}u_1 = 0\ . \tag{2.6}$$

Of course, the operator in (2.6) is just the determinant of the system (2.5).

This reduction can also be carried out by a potential representation. If we take

$$\underline{u} = \underline{U}\,\lambda$$

where $\underline{U}$ is any column of the transposed cofactor matrix then (2.1) becomes

$$A\underline{U}\lambda = \underline{0}$$

and $A\underline{U}$ has only one non-zero component which is equal to det A. In the example (2.5) the first column of A^* is

$$\begin{pmatrix} T^3 - X^2 + T \\ -X^3 \\ TX^2 + T^2 + 1 \\ -T^2X - X \end{pmatrix} .$$

So we may substitute

$$u_1 = (T^3 - X^2 + T)\lambda$$

$$u_2 = -X^3\lambda$$

$$u_3 = (TX^2 + T^2 + 1)\lambda$$

$$u_4 = (-T^2X - X)\lambda$$

in (2.5). The first equation becomes

$$(T^4 + 2T^2 + 1 - X^4)\lambda = 0$$

and the others are satisfied identically.

More generally we may take

$$\underline{u} = A^{*}\underline{k}\lambda$$

where $\underline{k}$ is an arbitrary constant n-vector. Then we have

$$A[\underline{u}] = AA^{*}\underline{k}\lambda$$

$$= (\det A)\underline{k}\lambda$$

and the system is equivalent to

$$(\det A)[\lambda] = 0 .$$

In fact, an even more general potential representation for solutions of (2.1) is

$$\underline{u} = A^{*}[\underline{k}_1 \lambda_1 + \ldots + \underline{k}_p \lambda_p] \qquad \text{for } p \leq n \tag{2.7}$$

where the $\underline{k}_i$ are independent n-vectors and each of the λ_i satisfy the partial differential equation

$$(\det A)\lambda_i = 0 . \tag{2.8}$$

The proof of this is analogous to the proof immediately above for a single potential. Of course the $\underline{k}_i$ need not be independent for the purposes of the proof but in any event it is clear that the term

$$\underline{k}_1 \lambda_1 + \ldots + \underline{k}_p \lambda_p$$

may be represented in terms of an arbitrary set of orthonormal basis vectors without loss of generality.

By the above method the original system may be reduced to the single equation

$$D[\lambda] = 0$$

where D is the operator $\det A$. If D is even in the derivatives a variational principle exists for the potential equation. If D is odd, we introduce $\lambda = X_1 \mu$ say, to make it even. We will now note the conditions under which D is even.

The determinant for (2.2) may be written

$$|O + b|$$

where O is the operator matrix with elements $a_1 X_1 + \dots a_\ell X_\ell$. The requirement that this be even in the X's, when expanded, is that

$$|O + b| = 0$$

$$|O - b| = 0 .$$

This fits in nicely with the theory of dispersive waves which was one of the motivations for this work. For, suppose we have a wave system of the form

$$0 = a\underline{u}_t + b_1 \underline{u}_{x_1} + \dots + b_\ell \underline{u}_{x_\ell} + c\underline{u} .$$

The existence of dispersive waves is found from elementary solutions

$$\underline{u} = \underline{A} e^{i\omega t + ik_1 x_1 + \dots + ik_\ell x_\ell}$$

where $\underline{A}$, $\underline{k}$, ω are constants, $\underline{k}$ is the wave number and ω is the

frequency. On substitution we have

$$\{i(\omega a + b_1 k_1 + b_2 k_2 + \ldots + b_\ell k_\ell) + c\}\underline{A} = 0 .$$

For a non-trivial solution

$$D = |i(\omega a + b_1 k_1 + \ldots + b_\ell k_\ell) + c| = 0 .$$

This is the dispersion relation that determines the frequency ω as a function of the wave number $(k_1, \ldots, k_\ell)$. For undamped waves this must be an equation in $\omega, k_1, \ldots, k_\ell$ with real coefficients. The condition for this is that a change in the sign of i gives the same equation. This is exactly the condition noted above on

$$D(T, X_1, \ldots, X_\ell) = |(aT + b_1 X_1 + \ldots + b_\ell X_\ell) + c| .$$

For real dispersive waves, the general theory developed by Whitham requires the existence of a variational principle. Our argument here proves that such a variational principle exists for linear systems.

II.3. Systems and side conditions

While the reduction to a single equation is satisfactory in many ways for linear systems, we pursue the alternative possibility of working with potentials and side conditions which only partially reduce the number of equations. As explained in the introduction, this seems to be necessary for non-linear systems and we investigate the ideas on the simpler case of linear systems. In general we will be concerned with a system of first order partial differential equations

$$A[\underline{u}] = \underline{u}_{,1} + b_2\underline{u}_{,2} + \ldots + b_\ell\underline{u}_{,\ell} + c\underline{u} = \underline{0} \,. \qquad (3.1)$$

We will consider variational principles whose integrands are quadratic in the dependent variable and to which m side conditions taken from the system are adjoined by means of side conditions. That is,

$$\mathfrak{L}[\underline{u},\underline{\lambda}] = \tfrac{1}{2}\underline{u}^T a^{-1}\underline{u} + \underline{\lambda}^T p^T(\underline{u}_{,1} + b_2\underline{u}_{,2} + \ldots + b_\ell\underline{u}_{,\ell} + c\underline{u}) \qquad (3.2)$$

where $\underline{\lambda}(\underline{x})$ is an m-vector representing the potentials, a is a symmetric, non-singular $n \times n$ matrix, p is an $n \times m$ projection matrix which determines the choice of equations to be used as side conditions. If the analogy with Pfaff's theorem is correct m equal to $n/2$ or $(n+1)/2$ should always be a sufficient number of potentials. We will now explore conditions under which the variational principle whose Lagrangian function is (3.2) yields a consistent potential representation. Varying $\underline{u}$ we obtain the Euler equations

$$\underline{u} = ap\underline{\lambda}_{,1} + ab_2^T p\underline{\lambda}_{,2} + \ldots + ab_\ell^T p\underline{\lambda}_{,\ell} - ac^T p\underline{\lambda} \qquad (3.3)$$

and from the side conditions we have

$$p^T(\underline{u}_{,1} + b_2 u_{,2} + \ldots + b_\ell\underline{u}_{,\ell} + c\underline{u}) = \underline{0} \,.$$

The first set of Euler equations is required to be a potential representation for the solution $\underline{u}$ in terms of first derivatives $\underline{\lambda}_{,1},\ldots,\underline{\lambda}_{,m}$ of the m potentials. The result of substituting this potential representation for $\underline{u}$ into the original system ($A[\underline{u}] = 0$) must be a <u>consistent</u> set of n second-order equations for the m

potentials $\underline{\lambda}$. The question of consistency and possible restrictions on the choice of side conditions gets fairly involved and we will start with special cases.

II.4. Systems of two equations

We consider first the 2-system

$$\underline{u}_t + b\underline{u}_x = \underline{0}$$

where

$$u = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} , \qquad b = \begin{pmatrix} A & B \\ C & D \end{pmatrix} .$$

Here we need one potential λ and one linear combination of the equations as side conditions. In the case of two equations the proposed reduction to roughly half the number of equations coincides with the reduction to a single equation discussed in Section II.2. We may start, therefore, with the representations of that section and see how to tie them in with a variational principle plus side conditions. The transposed cofactor matrix for (4.1) is

$$\begin{pmatrix} T + DX & -BX \\ -CX & T + AX \end{pmatrix}$$

so that the representation (2.7) is

$$\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} k \\ \ell \end{pmatrix} \lambda_t + \begin{pmatrix} D & -B \\ -C & A \end{pmatrix} \begin{pmatrix} k \\ \ell \end{pmatrix} \lambda_x$$

i.e.

$$\underline{u} = \underline{k}\lambda_t + b^* \underline{k}\lambda_x \tag{4.2}$$

where $\underline{k}$ is an arbitrary vector and b^* is the transposed cofactor matrix of b. This representation may be used and the determinental equation (2.8) for λ is

$$\lambda_{tt} + (A + D)\lambda_{tx} + (AD - BC)\lambda_{xx} = 0$$

i.e.

$$\lambda_{tt} + (\text{tr } b)\lambda_{tx} + (\det b)\lambda_{xx} = 0 \ . \tag{4.3}$$

We can verify this directly in (4.1). For,

$$\underline{u}_t + b\underline{u}_x = \{\lambda_{tt} + (b + b^*)\lambda_{tx} + bb^*\lambda_{xx}\}\,\underline{k}$$

$$= \{\lambda_{tt} + (\text{tr } b)\lambda_{tx} + (\det b)\lambda_{xx}\}\,\underline{k}$$

and (4.3) follows. We now consider the relation of this representation with (3.2) and (3.3). That is we consider a variational principle with Lagrangian of the form

$$\mathfrak{L} = \tfrac{1}{2}\,\underline{u}^T a^{-1} \underline{u} + \lambda \underline{p}^T (\underline{u}_t + b\underline{u}_x) \tag{4.4}$$

where, in this case, a is a symmetric, 2×2 matrix and $\underline{p}$ is a 2-vector. What are the conditions on a and $\underline{p}$ in order that the Euler equations of this variational principle give a potential representation of the form (4.2) above? We obtain the Euler equations by varying $\underline{u}$; they are

$$\underline{u}^T = \underline{p}^T a(\lambda_t + b\lambda_x)$$

or, equivalently,

$$\underline{u} = a\underline{p}\lambda_t + ab^T\underline{p}\lambda_x \ .$$

This potential representation is of the form (4.2) for arbitrary $\underline{p}$ provided

$$a\underline{p} = \underline{k}$$

$$ab^T\underline{p} = b^*\underline{k} \ ,$$

Therefore we require

$$ab^T = b^*a \ .$$

Equating the elements in this matrix equation for a, it is found that the only solutions are given by

$$a = r\begin{pmatrix} B & \frac{1}{2}(D-A) \\ \frac{1}{2}(D-A) & -C \end{pmatrix}$$

where r is an arbitrary constant. We may take r to be 1 without loss of generality.

With this choice of a, the Euler equations for (4.4) are

$$\underline{u} = a(\lambda_t + b^T\lambda_x)\underline{p} \tag{4.5}$$

and the side conditions are

$$\underline{p}^T(\underline{u}_t + b\underline{u}_x) = 0 \ . \tag{4.6}$$

Substitution of (4.5) into (4.6) gives the basic determinental equation (4.3) provided

$$\underline{p}^T a \underline{p} \neq 0 \, , \tag{4.7}$$

It may be verified that (4.5) satisfies

$$\underline{n}^T(\underline{u}_t + b\underline{u}_x) = 0$$

identically where $\underline{n}$ is a vector orthogonal to $a\underline{p}$. The condition (4.7) above guarantees that this equation is independent of (4.6).

We next consider the effect of the inclusion of an undifferentiated term. Consider a system of two equations in the form

$$\begin{pmatrix} \alpha \\ \beta \end{pmatrix}_t + \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix}_x + \begin{pmatrix} E & F \\ G & H \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

or

$$\underline{u}_t + b\underline{u}_x + c\underline{u} = \underline{0} \ . \tag{4.8}$$

In section 2 of this chapter we found that a necessary and sufficient condition for the existence of a variational principle (by reduction to a single equation) is that

$$\det \begin{pmatrix} T + AX + E & BX + F \\ CX + G & T + DX + H \end{pmatrix}$$

$$= \det \begin{pmatrix} T + AX - E & BX - F \\ CX - G & T + DX - H \end{pmatrix} \ ;$$

that is,

$$A + D = 0$$

$$AH + DE - BG - CF = 0 . \tag{4.9}$$

If we take a Lagrangian with side condition of the form

$$\mathfrak{L} = \tfrac{1}{2}\underline{u}^T a^{-1}\underline{u} + \lambda \underline{p}^T(\underline{u}_t + b\underline{u}_x + c\underline{u})$$

then the potential representation obtained from the appropriate variational principle is

$$\underline{u} = a\underline{p}\lambda_t + ab^T\underline{p}\lambda_x - ac^T\underline{p}\lambda , \tag{4.10}$$

Upon substituting (4.10) into the original system (4.8) we obtain

$$[\, a\lambda_{tt} + (ba + ab^T)\lambda_{tx} + (bab^T)\lambda_{xx}$$

$$+ (ca - ac^T)\lambda_t + (cab^T - bac^T)\lambda_x - cac^T\lambda]\,\underline{p} = 0 .$$

Since we must require compatibility with the determinental equation for arbitrary $\underline{p}$, the expression in brackets must be

$$a[\det (T + bX + c)]\,\lambda .$$

This example contains the previous example as a special case so that a is already known to be the matrix

$$\begin{pmatrix} B & \tfrac{1}{2}(D - A) \\ \tfrac{1}{2}(D - A) & -C \end{pmatrix} .$$

Compability is then shown to follow when the conditions (4.9) are satisfied (i.e. only when odd derivative terms are absent from the

determinental equation.)

It is natural to try to extend these results to more space dimensions, that is to a system of two equations in the form

$$\underline{u}_t + b_1\underline{u}_{,1} + b_2\underline{u}_{,2} + \ldots + b_\ell\underline{u}_{,\ell} + c\underline{u} = 0 . \qquad (4.11)$$

The matrix a of the Lagrangian may be determined from b_1 and then $b_2,\ldots,b_\ell$ must not be incompatible with this choice of a. It can be shown that a variational principle with side condition exists only when $b_1,\ldots,b_\ell$ are linearly related by

$$b_i = \alpha_i b_i + \beta_i I \qquad 2 \le i \le \ell \qquad (4.12)$$

where α_i and β_i are scalars and the determinental equation of the system

$$\underline{u}_t + b_1\underline{u}_x + c\underline{u} = \underline{0}$$

contains only even derivative terms.

In summary, we have shown the existence of a variational principle using an equation of the system as a side condition for systems of two equations in the form

$$\underline{u}_t + (b_1\underline{u}_{x_1} + \alpha_2 b_1 + \beta_2 I)\underline{u}_{x_2} + \ldots$$

$$+ (\alpha_\ell b_1 + \beta_\ell I)\underline{u}_{x_\ell} + c\underline{u} = \underline{0}$$

for which the determinantal equation contains no odd derivative terms.

There is no side condition approach to the variational principle

for more general coefficient matrices b_i in (4.11), although we see from Section II.2 that representations and variational principles are still possible. This shows that the side condition technique need not work for systems for which variational principles are possible. It should be borne in mind, however, that physical systems do always have some special properties and it may be that these usually allow the side condition technique.

II.5. Potential representations for higher order systems in two independent variables

For systems of more than two equations, variational principles with side conditions are more difficult to analyze. With only two equations the form of the potential representation is uniquely determined (only the choice of $\underline{k}$ is arbitrary) and the potential satisfies a single second order equation. The coefficients of this equation are scalars which commute with the vectors involved. For a system of n equations, a variational principle with side conditions will lead to a potential representation in terms of a number of potentials. The second order potential equations will now be a system with matrix coefficients which will not commute with the projection matrix p. In this case it is surprising that we can make any progress at all!

We first consider in the next few sections, the question of potential representations of appropriate type irrespective of whether they can be associated with a variational principle and side conditions. As a starting point in this we will show that a system of 2m equations

of the form

$$\underline{u}_t + b\underline{u}_x = \underline{0}$$

always has a potential representation in terms of a potential m-vector $\underline{\lambda}$ which satisfies the second order potential equation

$$\underline{\lambda}_{tt} + 2\psi\underline{\lambda}_{tx} + \varphi\underline{\lambda}_{xx} = \underline{0}$$

where $\underline{\lambda}$ is an m-vector and ψ, φ are constant $m \times m$ matrices. In fact there is a large degree of arbitrariness in the choice of the potential representation just as there was with only two equations (for two independent variables).

Let k be an arbitrary $2m \times m$ matrix subject to the condition that the partitioned matrix

$$(bk;k)$$

be nonsingular. Then, if ψ and φ are defined by

$$\begin{pmatrix} 2\psi \\ -\varphi \end{pmatrix} = (bk;k)^{-1} b^2 k$$

the potential representation

$$\underline{u} = k\underline{\lambda}_t + (2k\psi - bk)\underline{\lambda}_x \qquad (5.1)$$

leads to the desired second order equation. To see the content of this we will give an example now and postpone the general proof of the assertion until the end of this section.

Example: Suppose

$$b = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

and k is chosen to be

$$k = \begin{pmatrix} 3 & 2 \\ 1 & -3 \\ 1 & 4 \\ -1 & 1 \end{pmatrix} .$$

Then

$$b^2k = \begin{pmatrix} 1 & 4 \\ -1 & 1 \\ 3 & 2 \\ 1 & -3 \end{pmatrix}$$

and

$$(bk;k) = \begin{pmatrix} 1 & -3 & 3 & 2 \\ 1 & 4 & 1 & -3 \\ -1 & 1 & 1 & 4 \\ 3 & 2 & -1 & 1 \end{pmatrix}$$

is nonsingular and has an inverse

$$(bk;k)^{-1} = \frac{1}{72} \begin{pmatrix} 9 & 0 & -9 & 18 \\ -6 & 10 & 10 & 2 \\ 15 & 14 & 5 & -8 \\ 0 & -6 & 12 & 6 \end{pmatrix} .$$

The expression

$$(bk;k)^{-1}b^2k$$

reduces to

$$\frac{1}{18}\begin{pmatrix} 0 & -9 \\ 4 & 0 \\ 2 & 27 \\ 12 & 0 \end{pmatrix}.$$

Hence,

$$2\psi = \frac{1}{18}\begin{pmatrix} 0 & -9 \\ 4 & 0 \end{pmatrix}$$

and

$$\varphi = -\frac{1}{18}\begin{pmatrix} 2 & 27 \\ 12 & 0 \end{pmatrix}.$$

According to (5.1) consider the potential representation

$$\underline{u} = k\underline{\lambda}_t + (2k\psi - bk)\underline{\lambda}_x$$

or

$$\underline{u} = \begin{pmatrix} 3 & 2 \\ 1 & -3 \\ 1 & 4 \\ -1 & 1 \end{pmatrix}\underline{\lambda}_t + \frac{1}{18}\begin{pmatrix} -10 & 27 \\ -30 & -81 \\ 34 & -27 \\ -50 & -27 \end{pmatrix}\underline{\lambda}_x . \qquad (5.2)$$

The result of substituting this potential representation into the original equation

$$\underline{u}_t + b\underline{u}_x = \underline{0}$$

is

$$\begin{pmatrix} 3 & 2 \\ 1 & -3 \\ 1 & 4 \\ -1 & 1 \end{pmatrix} \underline{\lambda}_{tt} + \frac{1}{18} \begin{pmatrix} 8 & -27 \\ -12 & -9 \\ 16 & -9 \\ 4 & 9 \end{pmatrix} \underline{\lambda}_{tx} + \frac{1}{18} \begin{pmatrix} -30 & -81 \\ -34 & -27 \\ -50 & -27 \\ -10 & 27 \end{pmatrix} \underline{\lambda}_{xx} = \underline{0}$$

which is in turn equal to

$$\begin{pmatrix} 3 & 2 \\ 1 & -3 \\ 1 & 4 \\ -1 & 1 \end{pmatrix} \left[\underline{\lambda}_{tt} + \frac{1}{18} \begin{pmatrix} 0 & -9 \\ 4 & 0 \end{pmatrix} \underline{\lambda}_{tx} - \frac{1}{18} \begin{pmatrix} 2 & 27 \\ 12 & 0 \end{pmatrix} \underline{\lambda}_{xx} \right] = \underline{0} \ .$$

For any potential vector $\underline{\lambda}$ satisfying the second order equation

$$\underline{\lambda}_{tt} + 2\psi \underline{\lambda}_{tx} + \varphi \underline{\lambda}_{xx} = \underline{0}$$

the expression (5.2) represents solutions of the first order system

$$\underline{u}_t + b\underline{u}_x = \underline{0} \ .$$

For more than two equations in two independent variables the choice of ψ and φ is dictated by the choice of k and this dependence has been shown explicitly.

Let us consider this method in general. Given the original first order system of $2m$ equations

$$\underline{u}_t + b\underline{u}_x = \underline{0} \tag{5.3}$$

we construct the $m \times m$ matrices ψ and φ defined by

$$\begin{pmatrix} 2\psi \\ -\varphi \end{pmatrix} = (bk;k)^{-1} b^2 k \tag{5.4}$$

for arbitrary k (where bk;k) is nonsingular). We claim that if the potentials $\underline{\lambda}$ satisfy the second order equation

$$\underline{\lambda}_{tt} + 2\psi\underline{\lambda}_{tx} + \varphi\underline{\lambda}_{xx} = 0 \tag{5.5}$$

then the potential representation

$$\underline{u} = k\underline{\lambda}_t + (2k\psi - bk)\underline{\lambda}_x \tag{2.6}$$

leads to a second order system which is compatible with (5.5).

To prove this substitute the potential representation (5.6) into the original equation (5.3)

$$\underline{u}_t + b\underline{u}_x = k\underline{\lambda}_{tt} + 2k\psi\underline{\lambda}_{tx} + (2bk\psi - b^2k)\underline{\lambda}_{xx}.$$

In order to establish compatibility with (5.5) we must show

$$k\varphi = 2bk\psi - b^2k$$

or

$$b^2k = (bk;k)\begin{pmatrix} 2\psi \\ -\varphi \end{pmatrix}.$$

Equation (5.7) must be satisfied because of the definition of ψ and φ in equation (5.4).*

---

*The methods of this section were first motivated by results for two equations in two independent variables. If b is a 2×2 matrix then

$$b^* = (\text{tr } b)I - b$$

so that (4.2) could be written

$$\underline{u} = \underline{k}\underline{\lambda}_t + [(\text{tr } b)I - b]\underline{k}\underline{\lambda}_x$$

which is close to (5.1). Krylov's theorem may be used to determine

tr b and det b given $\underline{k}$, $b\underline{k}$, $b^2\underline{k}$ where $\underline{k}$ is not an eigenvector of b. Not surprisingly, the condition that $\underline{k}$ is not an eigenvector of b is another way of stating (4.7).

II.6. Potential representations for general linear systems

Before discussing the existence of variational principles leading to potential representations like those above we will consider the effect of adding more space dimensions. (Adding an undifferentiated term is formally equivalent to adding an extra space dimension, as far as finding potential representation is concerned.)

Consider the potential representation

$$\underline{u} = k_1\underline{\lambda}_{,1} + k_2\underline{\lambda}_{,2} + \ldots + k_\ell\underline{\lambda}_{,\ell} \tag{6.1}$$

where the k_i's are constant $2m \times m$ matrices for $i = 1,\ldots,\ell$ and $\underline{\lambda}$ is an m-vector. It is of interest to determine conditions on the k_i's for which the potential representation (6.1) leads to a consistent set of second order partial differential equations for $\underline{\lambda}$ when it is substituted into the first order system

$$\underline{u}_{,1} + b_2\underline{u}_{,2} + \ldots + b_\ell\underline{u}_{,\ell} = \underline{0} . \tag{6.2}$$

On substitution we obtain $2m$ equations for the m λ_i's; since $2m > m$ the extra equations must be consistent. We will suppose that a basic second order system satisfied by $\underline{\lambda}$ is of the form

$$0 = \psi_{ij}\underline{\lambda}_{,ij} \qquad i,j = 1,\ldots,\ell \tag{6.3}$$

where the ψ_{ij} are constant $m \times m$ matrices and without loss of

generality we may assume $\psi_{ij} = \psi_{ji}$. For the purposes of this discussion let n be even. Here, motivated by the results of Section II-2, we assume that all components of $\underline{u}$ and $\underline{\lambda}$ satisfy the same $2m^{th}$ order partial differential equation.

By considering the highest x_1 derivative in the partial differential equation we may take without loss of generality

$$\psi_{11} = I_m \qquad \text{(the } m \times m \text{ identity matrix)}.$$

Then the remaining ψ_{ij} are uniquely specified. Our approach now is to substitute the potential representation (6.1) into the first order system (6.2) and to determine conditions under which the resulting second order system would be consistent with (4.3).

The result of substituting the potential representation into the first order system is

$$b_i k_j \underline{\lambda}_{,ij} = \underline{0}$$

where summation is $i,j = 1,\ldots,\ell$ and b_1 is the unit matrix. The consistency condition obtained by requiring the second order system to be some linear combination of the equations in (4.3) is

$$b_i k_j + b_j k_i = 2k\psi_{ij} \qquad i,j = 1,\ldots,\ell$$

where k is a $2m \times m$ constant matrix. For the case $i = 1$, $j = 1$ the consistency condition becomes

$$k_1 = k .$$

For $i = 1$, $j \neq 1$ the consistency condition defines the other k_j in terms of k and the ψ_{ij}'s

$$k_j + b_j k_1 = 2k\psi_{1j}$$

$$k_j = 2k\psi_{1j} - b_j k .$$

For $i \neq 1$, $j \neq 1$, $\frac{\ell(\ell-1)}{2}$ consistency conditions result

$$(b_i b_j + b_j b_i)k - 2b_i k\psi_{1j} - 2b_j k\psi_{1i} + 2k\psi_{ij} = 0 . \qquad (6.4)$$

These are equations for k and ψ_{ij}. If we require solutions for arbitrary k, we have $\frac{\ell(\ell-1)}{2}$ equations for $\frac{(\ell+1)\ell}{2} - 1$ ψ_{ij}'s (ψ_{11} is the identity matrix). Each of the ψ_{ij} has m^2 components while each equation of (6.4) has $2m^2$ components. We are left with $\ell(\ell-1)m^2$ equations for $\frac{(\ell+2)(\ell-1)}{2} m^2$ unknowns. The number of equations exceeds the number of unknowns for $\ell \geq 3$ so we would not expect solutions to exist except in certain special cases. (One of these special cases is the system of two equations analyzed by the method of Section II.2.) In Section II.5 we obtained a solution for arbitrary k and $\ell = 2$. In general we must expect that for $\ell \geq 3$ the freedom in the choice of k will be greatly reduced.

Special case of two equations

When there are only two equations ($n = 2$), the method used in Section II.2 may be used to generate the relevant potential representation. In order to develop a more flexible notation we will briefly review that work bearing in mind that for the present we are seeking

consistent potential representations without regard to the existence of a variational principle.

Consider the first order system of two equations

$$\underline{u}_{,1} + b_2 u_{,2} + \ldots + b_\ell \underline{u}_{,\ell} = 0 . \tag{6.5}$$

The second order partial differential operator of this system may be written

$$\det [IX_1 + b_2 X_2 + \ldots + b_\ell X_\ell]$$

$$= IX_1^2 + 2\psi_{12} X_1 X_2 + \ldots + 2\psi_{1\ell} X_1 X_\ell + \psi_{22} X_2^2 + \ldots + \psi_{\ell\ell} X_\ell^2 .$$

Here the ψ_{ij} are scalars and they must be the ψ's of equation (6.3). In particular

$$2\psi_{1i} = \operatorname{tr} b_i .$$

The relation of the ψ's to the b's may be seen by applying the Hamilton-Cayley theorem, i.e.

$$m^2 - (\operatorname{tr} m)m + (\det m)I = 0$$

to the matrix

$$IX_1 + b_2 X_2 + \ldots + b_\ell X_\ell .$$

We have

$$[IX_1 + b_2 X_2 + \ldots + b_\ell X_\ell]^2$$
$$- [2IX_1 + 2\psi_{12} X_2 + \ldots + 2\psi_{1\ell} X_\ell] (IX_1 + b_2 X_2 + \ldots + b_\ell X_\ell)$$
$$+ [IX_1^2 + 2\psi_{12} X_1 X_2 + \ldots + 2\psi_{1\ell} X_1 X_\ell + \psi_{33} X_2^2 + \ldots + \psi_{\ell\ell} X_\ell^2] I$$
$$= 0 .$$

Upon separating coefficients of like powers of $X_i X_j$ we obtain

$$0 = b_i b_j + b_j b_i - 2\psi_{1i} b_j - 2\psi_{1j} b_i + 2\psi_{ij}$$

for fixed i and j. These are precisely the consistency conditions obtained in (6.4) for the existence of a potential representation. Therefore a suitable potential representation for solutions of (6.5) is

$$\underline{u} = \underline{k}\lambda_{,1} + (2\psi_{12} - b_2)\underline{k}\lambda_{,2} + \ldots + (2\psi_{1\ell} - b_\ell)\underline{k}\lambda_{,\ell}$$

for arbitrary $\underline{k}$. (Note the extra conditions on the b_i in Section II.4 come from the further requirement that this representation comes from the side condition of a variational principle.) This proof hinges on the fact that the ψ_{ij} are scalars which commute with $\underline{k}$ and the b_i.

Special case of two independent variables

The case of an even number (2m) of equations in two independent variables was analyzed in Section II.5. There we found ψ_{12} and ψ_{22} could be determined from

$$\begin{pmatrix} 2\psi_{12} \\ -\psi_{22} \end{pmatrix} = (bk;k)^{-1} b^2 k \ .$$

II.7. Potential representations for more than two independent variables

The analysis of systems containing more than two coefficients is difficult because the coefficients of the potential equation (6.3) are functions of k. The preceding discussion shows that we should not expect potential representations for an arbitrary choice of k when the system is written in three or more independent variables. We will now consider the systems in three independent variables of the form

$$\underline{u}_{,1} + b_2\underline{u}_{,2} + b_3\underline{u}_{,3} = 0 . \tag{7.1}$$

Any consistent potential representation of solutions of (7.1) will be of the form

$$\underline{u} = k\underline{\lambda}_{,1} + (2k\psi_{12} - b_2k)\underline{\lambda}_{,2} + (2k\psi_{13} - b_3k)\underline{\lambda}_{,3} \tag{7.2}$$

where it is assumed that λ satisfies

$$\underline{\lambda}_{,11} + 2\psi_{12}\underline{\lambda}_{,12} + 2\psi_{13}\underline{\lambda}_{,13} + \psi_{22}\underline{\lambda}_{,22} + 2\psi_{23}\underline{\lambda}_{,23} + \psi_{33}\underline{\lambda}_{,33} = \underline{0} . \tag{7.3}$$

The consistency conditions (6.4) (obtained in this case by substituting (7.2) into (7.1) and comparing with (7.3)) are

$$k\psi_{22} = 2b_2k\psi_{12} - b_2^2k \tag{7.4}$$

$$2k\psi_{23} = 2b_2k\psi_{13} + 2b_3k\psi_{12} - b_2b_3k - b_3b_2k \tag{7.5}$$

$$k\psi_{33} = 2b_3k\psi_{13} - b_3^2k .$$

We may combine these into

$$(b_2k;b_3k;k)\begin{pmatrix} 2\psi_{12} & 2\psi_{13} & 0 \\ 0 & 2\psi_{12} & 2\psi_{13} \\ -\psi_{22} & -2\psi_{23} & -\psi_{33} \end{pmatrix} = (b_2^2k;(b_2b_3+b_3b_2)k;b_3^2k). \quad (7.7)$$

Now choose a $2m \times m$ matrix k such that

$$\text{rank}\,[\,b_2k;b_3k;k] = 2m$$

and, either

$$\text{rank}\,[\,b_2k;k] = m$$

or

$$\text{rank}\,[\,b_3k;k] = m\ .$$

Say the first alternative holds. We may solve (7.6) in the form

$$\begin{pmatrix} 2\psi_{13} \\ -\psi_{33} \end{pmatrix} = (b_3k;k)^{-1}b_3^2k\ . \quad (7.8)$$

Now, ψ_{12} and ψ_{13} may be determined from equation (7.5)

$$\begin{pmatrix} 2\psi_{12} \\ -2\psi_{13} \end{pmatrix} = (b_3k;k)^{-1}\Big((b_2b_3+b_3b_2)k - 2b_2k\psi_{13}\Big)\ . \quad (7.9)$$

The equation which remains to be solved is (7.4),

$$k\psi_{22} = 2b_2k\psi_{12} - b_2^2k\ . \quad (7.10)$$

ψ_{12} has already been determined so that we cannot solve this equation in the same way that we solved (7.5) and (7.6). However, we assumed that

$$\text{rank}\,[\,b_2k;k] = m$$

so that the solution to (7.10) may be found by multiplying both sides by k^T and solving the resultant equation

$$k^Tk\psi_{22} = 2k^Tb_2k\psi_{12} - k^Tb_2^2k\,.$$

The $m \times m$ matrix k has full rank since the matrix $(b_2k;b_3k;k)$ has rank $2m$. Therefore ψ_{22} may be written

$$\psi_{22} = (k^Tk)^{-1}(2k^Tb_2k\psi_{12} - k^Tb_2^2k) \tag{7.11}$$

where the right-hand side of (7.11) is known.

Example: Consider the model system introduced in section II.2,

$$\underline{u}_t + \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}\underline{u}_x + \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}\underline{u} = \underline{0} \tag{2.4'}$$

or

$$\underline{u}_t + b\underline{u}_x + c\underline{u} = 0\,.$$

First we have to determine k, a suitable choice is

$$k = \begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \\ 1 & -1 \end{pmatrix}$$

since

$$\text{rank (bk;ck;k)} = \text{rank}\begin{pmatrix} 1 & -1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 & 1 & -1 \end{pmatrix} = 4$$

while

$$\text{rank (bk;k)} = \text{rank}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \end{pmatrix} = 2$$

ψ_{13} and ψ_{33} are found from (7.8),

$$\begin{pmatrix} 2\psi_{13} \\ -\psi_{33} \end{pmatrix} = (\text{ck;k})^{-1} c^2 k$$

$$= \frac{1}{4}\begin{pmatrix} 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ -1 & 0 \\ 0 & -1 \end{pmatrix}.$$

That is,

$$\psi_{13} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

and

$$\psi_{33} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

We may now determine ψ_{12} and ψ_{23} from (7.9)

$$\begin{pmatrix} 2\psi_{12} \\ -2\psi_{23} \end{pmatrix} = (ck;k)^{-1}(bck + cbk - 2bk\psi_{13})$$

$$= \frac{1}{4}\begin{pmatrix} 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \end{pmatrix}\begin{pmatrix} 2 & -2 \\ 0 & 0 \\ -2 & 2 \\ 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & -1 \\ 1 & -1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

To find ψ_{22} we must solve (7.11)

$$\psi_{22} = (k^T k)^{-1}(2k^T bk\psi_{12} - k^T b^2 k)$$

$$= \frac{1}{4}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & -4 \\ -4 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix},$$

The potential representation determined by this particular choice of k is, from equation (7.2),

$$\underline{u} = k\underline{\lambda}_t + (2k\psi_{12} - bk)\underline{\lambda}_x + (2k\psi_{13} - ck)\lambda$$

$$= \begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \\ 1 & -1 \end{pmatrix}\underline{\lambda}_t + \begin{pmatrix} 1 & -1 \\ -1 & -1 \\ 1 & -1 \\ -1 & -1 \end{pmatrix}\underline{\lambda}_x + \begin{pmatrix} -1 & -1 \\ -1 & 1 \\ 1 & 1 \\ 1 & -1 \end{pmatrix}\lambda \qquad (7.12)$$

where $\underline{\lambda}$ satisfies (7.3) which may be written

$$\underline{\lambda}_{tt} + \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix} \underline{\lambda}_{tx} + \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \underline{\lambda}_{xx} + \underline{\lambda} = \underline{0} .$$

II.8. Construction of variational principles

At the beginning of Section II.3 we set out to examine linearized variational principles which use some of the equations of the original system as side conditions. Variational principles of this type seem to be correct for nonlinear systems. They work by introducing a potential representation for solutions to the original set of first order equations; in terms of potentials the original system reduces to a self-adjoint second order system for which a variational principle may be found.

Investigation of the linearized problem shows that even the question of the existence of a potential representation is not trivial. Even if a potential representation does exist, there is still no guarantee that a variational principle of the form

$$\delta J = \delta \int \{ \tfrac{1}{2} \underline{u}^T a^{-1} \underline{u} + \underline{\lambda}^T p^T (\underline{u}_{,1} + b_2 \underline{u}_{,2} + \ldots + b_\ell \underline{u}_{,\ell} + c\underline{u} \} \, \underline{dx}$$

may be used to generate it as we have seen from the detailed analysis of systems of two equations. A technique which has been used successfully to find a number of new variational principles for first order linear systems is described below. Basically we attempt to construct a variational principle which will lead to a given potential representation. Since we lack precise conditions for the existence of variational

principles (beyond the condition that the determinental equation has only even terms) it is pointless to try to maintain full generality. We take the case of two independent variables but include an undifferentiated term. (This corresponds to the case $\ell = 3$ as far as the algebra of possible potential representations is concerned.)

Suppose that

$$J = \int \{\tfrac{1}{2}\underline{u}^T a^{-1}\underline{u} + \underline{\lambda}^T p^T(\underline{u}_t + b\underline{u}_x + c\underline{u})\}\, \underline{dx} \tag{8.1}$$

is a variational principle for the system

$$\underline{u}_t + b\underline{u}_x + c\underline{u} = \underline{0}\,. \tag{8.2}$$

Here a, b and c are $2m \times 2m$ matrices, a is nonsingular and symmetric, p is a $2m \times m$ projection matrix with rank m, $\underline{u}$ is a 2m-vector and $\underline{\lambda}$ is an m-vector. The Euler equations of (8.1) give the potential representation

$$\underline{u} = ap\underline{\lambda}_t + ab^T p\underline{\lambda}_x - ac^T p\underline{\lambda} \tag{8.3}$$

and the side conditions give

$$p^T(\underline{u}_t + b\underline{u}_x + c\underline{u}) = 0\,. \tag{8.4}$$

The potential equation is obtained by substituting (8.3) into (8.4)

$$\begin{aligned} &p^T ap\underline{\lambda}_{tt} + p^T(ab^T + ba)p\underline{\lambda}_{tx} + p^T(bab^T)p\underline{\lambda}_{xx} \\ &\quad + p^T(ca - ac^T)p\underline{\lambda}_t + p^T(cab^T - bac^T)p\underline{\lambda}_x - p^T cac^T p\underline{\lambda} \\ &\quad = 0\,. \end{aligned} \tag{8.5}$$

The coefficient matrices of this equation satisfy the symmetry condition previously derived in Section II.2, of course.

We now have to consider whether a and p can be found so that (8.3) and (8.5) are a consistent potential representation and all the equations of

$$\underline{u}_t + b\underline{u}_x + c\underline{u} = 0$$

are satisfied, not just the subset (8.4). This will be so provided

$$ap\underline{\lambda}_{tt} + (ab^T + ba)p\underline{\lambda}_{tx} + bab^T p\underline{\lambda}_{xx}$$

$$+ (ca - ac^T)p\underline{\lambda}_t + (cab^T - bac^T)p\underline{\lambda}_x - cac^T p\underline{\lambda} = 0. \qquad (8.6)$$

These are 2m equations for the m components of $\underline{\lambda}$ and they must be consistent. (The system (8.5) is a subset of m of them.) For (8.6) to be consistent there must exist a $2m \times m$ matrix k and $m \times m$ matrices ψ_{ij} so that

$$\begin{aligned} ap &= k \\ (ab^T + ba)p &= 2k\psi_{12} \\ bab^T p &= k\psi_{22} \\ (ca - ac^T)p &= 2k\psi_{13} \\ (cab^T - bac^T)p &= 2k\psi_{23} \\ -cac^T p &= k\psi_{33}. \end{aligned} \qquad (8.7)$$

In Section II.7 we discussed the problem of finding possible potential representations

$$\underline{u} = k\underline{\lambda}_t + (2\psi_{12} - bk)\underline{\lambda}_x + (2\psi_{13} - ck)\lambda \qquad (8.8)$$

and the consistent reduction of the equations

$$\underline{u}_t + b\underline{u}_x + c\underline{u} = 0$$

to the system

$$\underline{\lambda}_{tt} + 2\psi_{12}\underline{\lambda}_{tx} + \psi_{22}\underline{\lambda}_{xx} + 2\psi_{13}\underline{\lambda}_t + 2\psi_{23}\underline{\lambda}_x + \psi_{33}\underline{\lambda} = 0. \qquad (8.9)$$

If possible k and ψ_{ij} have been found as described in Section II.7 then the representation (8.8) can also be obtained from (8.1), provided a and p satisfying (8.8) can also be found. So we now take k and ψ_{ij} to be known and consider solving (8.8) for a and p.

The system (8.7) may be used to determine ab^Tp and ac^Tp. For example, from the third equation of (8.7) we obtain

$$bab^Tp = k\psi_{22}$$

$$ab^Tp = b^{-1}k\psi_{22}.$$

Similarly, the sixth equation of (8.7) gives

$$ac^Tp = -c^{-1}k\psi_{33}$$

Note that the following matrices are symmetric and that the quantities enclosed in parentheses are known in terms of k and ψ_{ij} from (8.7):

$$
\begin{aligned}
&p^T(ap)\\
&p^T(bab^Tp)\\
&p^T(cac^Tp)\\
&p^T(bap + ab^Tp)\\
&p^T(cap + ac^Tp)\\
&p^T(cab^Tp + bac^Tp)\\
&p^T(cac^Tp)\,,
\end{aligned}
\tag{8.10}
$$

The following matrices are skew-symmetric:

$$
\begin{aligned}
&p^T(bap - ab^Tp)\\
&p^T(cap - ac^Tp)\\
&p^T(cab^Tp - bac^Tp).
\end{aligned}
\tag{8.11}
$$

Using (8.10) and (8.11) we determine p. In order that (8.5) be non-trivial, p^Tap must have full rank. Hence we require

$$\text{rank}\,(p^Tk) = m\,,$$

(For the case of two equations discussed in Section II.4 this condition may be written $\underline{p}^T a\underline{p} = \underline{p}^T\underline{k} \neq 0$.)

We may now determine a, if it exists, from

$$a(p;b^Tp;c^Tp) = (k;b^{-1}k\psi_{22};-c^{-1}k\psi_{33}) \tag{8.12}$$

using the newly-calculated p on the left and the previously-known k and ψ_{ij} on the right. The matrix a must be symmetric.

If any step of this construction fails then no variational principle with side conditions of type (8.1) can lead to the potential representation (8.6).

Example:

Let us consider the model system of four equations

$$\underline{u}_t + \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix} \underline{u}_x + \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \underline{u} = \underline{0} . \qquad (2.4')$$

We will try to construct a variational principle of the form (8.1) for

$$k = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} .$$

The technique developed in Section II.7 may be used to show that this choice of k does lead to a consistent potential representation with

$$\psi_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} , \qquad \psi_{22} = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} , \qquad \psi_{33} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} ,$$

$$\psi_{12} = \psi_{13} = \psi_{23} = 0 .$$

We find $ab^T p$ and $ac^T p$ from (8.7)

$$ab^T p = \begin{pmatrix} 0 & 0 \\ 0 & -1 \\ 0 & 0 \\ -1 & 0 \end{pmatrix} , \qquad ac^T p = \begin{pmatrix} 0 & 1 \\ 0 & 0 \\ -1 & 0 \\ 0 & 0 \end{pmatrix} .$$

The symmetry conditions (8.9) and (8.10) may now be used to show that p is of the form

$$p = \begin{pmatrix} 0 & \alpha \\ \beta & 0 \\ \alpha & 0 \\ 0 & -\beta \end{pmatrix}.$$

The condition for non-trivial solutions to (8.5) is satisfied provided $\alpha \neq 0$ since

$$\text{rank } (p^T k) = \text{rank} \begin{pmatrix} 0 & \alpha \\ \alpha & 0 \end{pmatrix}.$$

Finally we solve (8.12) to obtain a.

$$a \begin{pmatrix} 0 & \alpha & 0 & -\beta & -\alpha & 0 \\ \beta & 0 & 0 & \alpha & 0 & \beta \\ \alpha & 0 & \beta & 0 & 0 & \alpha \\ 0 & -\beta & \alpha & 0 & \beta & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \end{pmatrix}$$

$$a = \frac{1}{\beta^4 - \alpha^4} \begin{pmatrix} 0 & \beta^3 & -\alpha^3 & 0 \\ \beta^3 & \alpha^3 & 0 & 0 \\ -\alpha^3 & 0 & 0 & -\beta^3 \\ 0 & 0 & -\beta^3 & \alpha^3 \end{pmatrix}.$$

It may be verified that a and p determined above may be used in a variational principle of the form (8.1).

This example is atypical in that two arbitrary constants α and

β enter. Usually p is determined to within a constant.

There are potential representations which cannot come from a variational principle of the form (8.1). If, in the example above, we make a seemingly trivial change in the choice of k, then the construction may fail. Consider the choice

$$k = \begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

(This choice of k was used to derive a potential representation for (2.4) in Section II.7.) The matrices $ab^T p$ and $ac^T p$ are readily found to be

$$ab^T p = \begin{pmatrix} 1 & -1 \\ -1 & -1 \\ 1 & -1 \\ -1 & -1 \end{pmatrix} \qquad ac^T p = \begin{pmatrix} 1 & 1 \\ 1 & -1 \\ -1 & -1 \\ -1 & 1 \end{pmatrix}.$$

In this case there is no choice of p which will satisfy the symmetry conditions (8.9) and (8.10) and for which

$$\operatorname{rank} p^T k = 2 .$$

Hence no variational principle may be constructed to give the potential representation (7.12).

In summary, we have developed a constructive technique to find variational principles leading to a given potential representation, provided such a variational principle exists.

CHAPTER III

NON-LINEAR SYSTEMS

III.1. Introduction

When we turn to non-linear equations, not only is the analysis more complicated but we find even further restrictions on the choice of side conditions. For two equations with two independent variables we found in the linear case (Chapter II, section 4) that any linear combination $p^T(u_t + bu_x)$ of the equations could be used as a side condition, with the mild exception that $p^T ap \neq 0$. In a typical example for the non-linear case we find the opposite: only four choices are possible. We start with this example which is of two equations in conservation law form.

III.2. Two conservation laws

Consider a system of two non-linear first order partial differential equations in two independent variables determined by the two conservation laws

$$\begin{aligned} &\frac{\partial}{\partial t}\alpha + \frac{\partial}{\partial x} a(\alpha,\beta) = 0 \\ &\frac{\partial}{\partial t}\beta + \frac{\partial}{\partial x} b(\alpha,\beta) = 0 . \end{aligned} \tag{2.1}$$

Previous examples with known physical systems indicate that only equations of this system which are themselves conservation laws are used as side conditions. Accordingly, we consider the whole family of conservation laws that may be derived from (2.1).

Suppose that

$$\frac{\partial}{\partial t} u(\alpha, \beta) + \frac{\partial}{\partial x} f(\alpha, \beta) = 0$$

is a conservation law of the system (2.1). Then there must exist functions μ and ν such that

$$du = \mu\, d\alpha + \nu\, d\beta$$

$$df = \mu\, da + \nu\, db.$$

From the first of these we determine

$$\mu = u_\alpha$$

$$\nu = u_\beta$$

and from the second one we obtain

$$\begin{aligned} f_\alpha &= u_\alpha a_\alpha + u_\beta b_\beta \\ f_\beta &= u_\alpha a_\beta + u_\beta b_\beta . \end{aligned} \tag{2.2}$$

From the consistency condition for $f(\alpha, \beta)$ we find

$$\begin{aligned} 0 &= (a_\beta u_\alpha)_\alpha + (b_\beta u_\beta)_x - (a_\alpha u_\alpha)_\beta - (b_\alpha a_\beta)_\beta \\ &= a_\beta u_{\alpha\alpha} + (b_\beta - a_\alpha) u_{\alpha\beta} - b_\alpha u_{\beta\beta} . \end{aligned} \tag{2.3}$$

This second order equation must be satisfied by every conserved density of the system (2.1). It will be denoted by

$$0 = R u_{\alpha\alpha} + 2S u_{\alpha\beta} + T u_{\beta\beta} \tag{2.4}$$

where

$$R = a_\beta, \qquad S = \tfrac{1}{2}(b_\beta - a_\alpha), \qquad T = -b_\alpha .$$

Consider now a variational principle of the form

$$\delta J = \delta \int\int \left\{ H(\alpha,\beta) + \lambda \left[\frac{\partial}{\partial t} u(\alpha,\beta) + \frac{\partial}{\partial x} f(\alpha,\beta) \right] \right\} dx\, dt = 0. \quad (2.5)$$

Varying J with respect to α and β we obtain

$$\begin{aligned} \delta\alpha: &\qquad H_\alpha = u_\alpha \lambda_t + f_\alpha \lambda_x \\ \delta\beta: &\qquad H_\beta = u_\beta \lambda_t + f_\beta \lambda_x . \end{aligned} \quad (2.6)$$

These Euler equations may be used to determine λ_t and λ_x as functions of α and β. We have

$$\lambda_t = \frac{H_\alpha f_\beta - H_\beta f_\alpha}{u_\alpha f_\beta - u_\beta f_\alpha} = -\, h(\alpha,\beta)$$

$$\lambda_x = \frac{u_\alpha H_\beta - u_\beta H_\alpha}{u_\alpha f_\beta - u_\beta f_\alpha} = w(\alpha,\beta)$$

where w and h are defined in terms of u, f, and H. Thus, we find that the conservation law

$$\frac{\partial}{\partial t} w(\alpha,\beta) + \frac{\partial}{\partial x} h(\alpha,\beta) = 0$$

must hold by virtue of the consistency condition satisfied by λ_t and λ_x. Since $w(\alpha,\beta)$ is a conserved density of the system (2.1), it must also satisfy (2.4); that is

$$R w_{\alpha\alpha} + 2S w_{\alpha\beta} + T w_{\beta\beta} = 0. \quad (2.7)$$

The system (2.6) may be written in terms of w and h. It is

$$H_\alpha = - u_\alpha h + f_\alpha w$$

$$H_\beta = -u_\beta h + f_\beta w .$$

Now, apply the integrability condition to H_α and H_β determined just above. It is

$$0 = H_{\beta\alpha} - H_{\alpha\beta}$$

$$= - u_\beta h_\alpha + f_\beta w_\alpha + u_\alpha h_\beta - f_\alpha w_\beta .$$

Equation (2.2) may be used to eliminate derivatives of f in favor of derivatives of u, and similar expressions relate h to w. When f and h are eliminated from the equation just above we have

$$0 = - u_\beta(w_\alpha a_\alpha + w_\beta b_\alpha) + (u_\alpha a_\beta + u_\beta b_\beta)w_\alpha$$

$$+ w_\alpha(w_\alpha a_\beta + w_\beta b_\beta) - (u_\alpha a_\alpha + b_\beta b_\alpha)w_\beta$$

$$= a_\beta u_\alpha w_\alpha + (b_\beta - a_\alpha)(u_\alpha w_\beta + u_\beta w_\alpha) - b_\alpha u_\beta w_\beta .$$

The last equation is

$$Ru_\alpha w_\alpha + S(u_\alpha w_\beta + u_\beta w_\alpha) + Tu_\beta w_\beta \qquad (2.8)$$

where $R(\alpha,\beta)$, $S(\alpha,\beta)$, and $T(\alpha,\beta)$ are defined as in equation (2.4).

We have derived three partial differential equations (2.4), (2.7), (2.8) which must be satisfied by the conserved densities $u(\alpha,\beta)$ and $w(\alpha,\beta)$. We collect them here for ease of reference

$$
\begin{aligned}
&Ru_{\alpha\alpha} + 2Su_{\alpha\beta} + Tu_{\beta\beta} = 0 \\
&Rw_{\alpha\alpha} + 2Sw_{\alpha\beta} + Tw_{\beta\beta} = 0 \\
&Ru_{\alpha}w_{\alpha} + S(u_{\alpha}w_{\beta} + u_{\beta}w_{\alpha}) + Tu_{\beta}w_{\beta} = 0
\end{aligned}
\tag{2.9}
$$

where

$$R = a_{\beta}, \quad S = \tfrac{1}{2}(b_{\beta} - a_{\alpha}), \quad T = -b_{\alpha} .$$

Variational principles of the form (2.5) exist only when this overdetermined system can be solved. The side condition has density u and the companion equation, satisfied identically by the potential representation, has density w.

In the case of a linear system of two equations in the form

$$\begin{pmatrix} \alpha \\ \beta \end{pmatrix}_t + \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix}_x = \underline{0}$$

the first two equations of (2.9) are trivially satisfied; this just corresponds to the obvious fact that any linear combination of the equations may be put in conservation law form. The last equation of (2.9) reduces to

$$Bu_{\alpha}w_{\alpha} + \tfrac{1}{2}(D - A)(u_{\alpha}w_{\beta} + u_{\beta}w_{\alpha}) - Cu_{\beta}w_{\beta} = 0. \tag{2.10}$$

In section II.4 we had

$$(u_{\alpha}, u_{\beta}) = \underline{p}^{T}$$

and

$$(w_\alpha, w_\beta) = \underline{n}^T .$$

So (2.10) reduces to

$$\underline{p}^T a \underline{n} = 0$$

in the notation used there. This is just a statement of the fact that some equation of the system must be identically satisfied by the choice of the potential representation.

As might be expected, for nonlinear systems consistent solutions of (2.9) are very restricted and as a consequence only special choices out of the conservation laws for a given system may be used as side conditions in a variational principle. We return to the determination of u and w after a few comments on the corresponding variational principles.

In view of the symmetric conditions on u and w, it is clear that if a variational principle exists using the side condition

$$\frac{\partial}{\partial t} u(\alpha,\beta) + \frac{\partial}{\partial x} f(\alpha,\beta) = 0 ,$$

then there must be a related variational principle using

$$\frac{\partial}{\partial t} w(\alpha,\beta) + \frac{\partial}{\partial x} h(\alpha,\beta)$$

as a side condition. A suitable variational principle is

$$0 = \delta J' = \delta \int\int \left[H + uh - wf + \psi(w_t + h_x) \right] dx\, dt . \qquad (2.11)$$

This integrand (without side condition) may be obtained by integration by parts from the previous variational principle (2.5). For

$$\delta J = \delta \iint \left[H + \lambda(u_t + f_x) \right] \, dx \, dt = 0$$

is equivalent to

$$\delta \iint \left[H - u\lambda_t - f\lambda_x \right] \, dx \, dt = 0$$

and this is

$$\delta \iint \left[H + uh - wf \right] \, dx \, dt = 0 .$$

Similarly, upon integration by parts, J' becomes

$$J'' = \iint H(\alpha, \beta) \, dx \, dt .$$

The intimate relation between the existence of conservation laws satisfying the three equations (2.9) and the existence of a variational principle is confirmed by the equations satisfied by H and its dual

$$H' = H - uh + wf . \tag{2.12}$$

For H and H' are found by integrating differential expressions of the form (see (2.6))

$$dH = - h \, du + w \, df$$

and

$$dH' = - f \, dw + u \, dh .$$

III.3. Solution of equations (2.9)

Second order quasi-linear equations of the form appearing in (2.9) are most easily analyzed and solved when they are hyperbolic. When the equations (2.1) describe waves (which is the case motivating this investigation), both (2.1) and the equations in (2.9) are hyperbolic. We therefore consider this case in detail.

The characteristics $\varphi(\alpha,\beta)$ = constant of

$$Ru_{\alpha\alpha} + 2Su_{\alpha\beta} + Tu_{\beta\beta} = 0$$

are given by

$$R\varphi_\alpha^2 + 2S\varphi_\alpha\varphi_\beta + T\varphi_\beta^2 = 0 .$$

We let $\varphi(\alpha,\beta)$ and $\psi(\alpha,\beta)$ be two independent characteristic variables and now change the variables in (2.9) from α,β-coordinates to φ,ψ-coordinates. The equations become, respectively,

$$u_{\varphi\psi} + Pu_\varphi + Qu_\psi = 0 \tag{3.1}$$

$$w_{\varphi\psi} + Pw_\varphi + Qw_\psi = 0 \tag{3.2}$$

$$w_\psi u_\varphi + w_\varphi u_\psi = 0 \tag{3.3}$$

where $P(\varphi,\psi)$ and $Q(\varphi,\psi)$ are given by

$$P = \frac{1}{\Delta}\left(R\varphi_{\alpha\alpha} + 2S\varphi_{\alpha\beta} + T\varphi_{\beta\beta}\right)$$

$$Q = \frac{1}{\Delta}\left(R\psi_{\alpha\alpha} + 2S\psi_{\alpha\beta} + T\psi_{\beta\beta}\right)$$

and

$$\Delta = 2\left[R\varphi_\alpha\psi_\alpha + S(\varphi_\alpha\psi_\beta + \varphi_\beta\psi_\alpha) + T\varphi_\beta\psi_\beta \right] .$$

The system (3.1), (3.2), and (3.3) may be integrated partially. Multiply (3.1) by w_ψ, multiply (3.2) by u_ψ, add the resulting equations and then apply (3.3) to obtain

$$\frac{\partial}{\partial\varphi}(u_\psi w_\psi) + 2Qu_\psi w_\psi = 0 . \tag{3.4}$$

Similarly, multiplying (3.1) and (3.2) by w_φ and u_φ, respectively, we find that

$$\frac{\partial}{\partial\psi}(u_\varphi w_\varphi) + 2Pu_\varphi w_\varphi = 0 . \tag{3.5}$$

Equations (3.4) and (3.5) may be integrated in terms of arbitrary functions $m(\psi)$ and $n(\varphi)$ to give

$$u_\psi w_\psi = m(\psi)e^{-2\mathfrak{Q}}$$

$$u_\varphi w_\varphi = n(\varphi)e^{-2\mathfrak{P}}$$

where

$$\mathfrak{P}(\varphi,\psi) = \int_{\psi_o}^{\psi} P(\varphi,\eta)\, d\eta$$

$$\mathfrak{Q}(\varphi,\psi) = \int_{\varphi_o}^{\varphi} Q(\xi,\psi)\, d\xi .$$

Hence, we have

$$\frac{u_\varphi w_\varphi}{u_\psi w_\psi} = \frac{n(\varphi)}{m(\psi)}\, e^{\,2\mathfrak{Q} - 2\mathfrak{P}} . \tag{3.6}$$

Equation (3.2) may be written

$$\frac{w_{\varphi}}{w_{\psi}} = - \frac{u_{\varphi}}{u_{\psi}} \, . \tag{3.7}$$

We may use (3.7) to eliminate w from (3.6) giving a first order partial differential equation for u

$$u_{\varphi}^2 \pm \frac{n(\varphi)}{m(\psi)} e^{2\mathfrak{Q} - 2\mathfrak{P}} u_{\psi}^2 = 0 . \tag{3.8}$$

Solutions exist only for

$$\frac{n(\varphi)}{m(\psi)} < 0$$

in which case (3.8) can be factored and the two cases which arise may be analyzed separately. We have

$$u_{\varphi} + \mu u_{\psi} = 0 \tag{3.9}$$

or

$$u_{\varphi} - \mu u_{\psi} = 0 \tag{3.10}$$

where

$$\mu = \sqrt{\frac{n(\varphi)}{m(\psi)}} \; e^{\mathfrak{Q} - \mathfrak{P}} \, .$$

Of course, due to the symmetry of (3.1), (3.2) and (3.3), the equations (3.9) and (3.10) are not independent. Suppose $u(\varphi,\psi)$ satisfies (3.9). Then the corresponding $w(\varphi,\psi)$ defined by (3.7) will satisfy (3.10) and vice-versa. Suppose (3.9) holds. We may integrate this equation along characteristics to obtain

$$u(\varphi,\psi) = u(s(\varphi,\psi)) \tag{3.11}$$

where

$$s_{\varphi} = -\mu$$

$$s_{\psi} = 1 .$$

In order to obtain further conditions on $m(\psi)$ and $n(\varphi)$, which are still arbitrary, we substitue $u(s)$ into equation (3.1).

From equation (3.9), we see that

$$u_{\varphi} = -\mu \dot{u}(s)$$

$$u_{\psi} = \dot{u}(s)$$

$$u_{\psi\varphi} = -\mu \ddot{u}(s),$$

so that (3.1) may be written

$$\begin{aligned} 0 &= u_{\varphi\psi} + Pu_{\varphi} + Qu_{\psi} \\ &= -\mu \ddot{u}(s) - \mu P \dot{u}(s) + Q \dot{u}(s) . \end{aligned}$$

Therefore,

$$\frac{\ddot{u}(s)}{\dot{u}(s)} = \frac{Q - \mu P}{\mu} . \tag{3.12}$$

The left-hand side of (2.28) is a function of s. We define $F(s)$ by

$$\frac{Q - \mu P}{\mu} = F(s) . \tag{3.13}$$

The function F so determined must satisfy an integrability condition

of the form

$$s_{\psi}F_{\varphi} - s_{\varphi}F_{\psi} = 0. \tag{3.14}$$

This condition is used to obtain differential equations for the functions $m(\psi)$ and $n(\varphi)$ which were previously undetermined. The integrability condition (3.14) places severe restrictions on the choice of the functions $m(\psi)$ and $n(\varphi)$.

In the analysis presented above we integrated the system (3.1), (3.2), and (3.3) directly. An alternate approach may be used to obtain two second order partial differential equations for the single function $u(\varphi,\psi)$. This overdetermined system may be used to determine whether a given conservation law

$$\frac{\partial}{\partial t}u(\varphi,\psi) + \frac{\partial}{\partial x}f(\varphi,\psi) = 0$$

may be used as a side condition in a variational principle. We proceed by separating (3.3) in the form

$$\begin{aligned} w_{\varphi} &= u_{\varphi}e^{2\lambda(\varphi,\psi)} \\ w_{\psi} &= -u_{\psi}e^{2\lambda(\varphi,\psi)}. \end{aligned} \tag{3.15}$$

Consistency on $w(\varphi,\psi)$ gives

$$0 = 2u_{\varphi\psi} + u_{\psi}\lambda_{\varphi} + u_{\varphi}\lambda_{\psi}. \tag{3.16}$$

We may substitute w_{φ} and w_{ψ} determined by (3.15) into (3.2) to obtain

$$0 = u_{\varphi\psi} + (2\lambda_\psi + P)u_\varphi - Qu_\psi \ . \tag{3.17}$$

Now, if we use (3.1) to eliminate $u_{\varphi\psi}$ terms from (3.16) and (3.17), the resulting system may be solved for λ_φ and λ_ψ to give

$$\lambda_\varphi = Pu_\varphi u_\psi^{-1}$$

$$\lambda_\psi = Qu_\varphi^{-1} u_\psi \ .$$

Consistency on λ gives a second order equation for $u(\varphi,\psi)$ which is independent of (3.1),

$$(Pu_\varphi u_\psi^{-1})_\psi - (Qu_\varphi^{-1} u_\psi)_\varphi = 0 \ . \tag{3.18}$$

A variational principle of the form (2.5) exists for each function $u(\varphi,\psi)$ satisfying both (3.1) and (3.18).

In summary, we have obtained an overdetermined system, solutions of which lead to variational principles of the type (2.5). Hyperbolic systems may be integrated, after reduction to characteristic coordinates, and the existence question for any given system may be settled. In addition, we may quickly determine whether any proposed function $u(\varphi,\psi)$ may be used as the conserved density in the side condition of a variational principle.

III.4. Application to shallow water theory

The shallow water equations of hydrodynamics provide a good example of how this theory applies to a fully nonlinear problem. The shallow water theory is developed from equations of conservation of

mass and conservation of momentum of the form

$$\frac{\partial}{\partial t}\alpha + \frac{\partial}{\partial x}(\alpha\beta) = 0 \tag{4.1}$$

$$\frac{\partial}{\partial t}(\alpha\beta) + \frac{\partial}{\partial x}(\alpha\beta^2 + \tfrac{1}{2}\alpha) = 0 \tag{4.2}$$

where α and β are nondimensional height and velocity respectively. Given the conservation of mass equation, the momentum equation may be expanded and simplified to give the conservation law

$$\frac{\partial}{\partial t}\beta + \frac{\partial}{\partial x}(\beta^2 + \alpha) = 0. \tag{4.3}$$

We will work with equations (4.1) and (4.3). Referring to (2.4), we see that here

$$a(\alpha,\beta) = \alpha\beta$$

$$b(\alpha,\beta) = \beta^2 + \alpha$$

and

$$R = \alpha, \quad S = 0, \quad T = -1 .$$

Any conserved density of the shallow water equations must satisfy

$$\alpha u_{\alpha\alpha} - u_{\beta\beta} = 0 . \tag{4.4}$$

It is clear that this has an infinity of solutiions; hence there are an infinite number of conservation equations. From (4.4) we note that if $u(\alpha,\beta)$ is a conserved density then so is $u_\beta(\alpha,\beta)$. Moreover, given any conserved density, another conserved density may be obtained by integrating with respect to β and adding a suitable

function of α as a "constant of integration." An infinite sequence of polynomial conservation laws may be constructed in this fashion starting with equation (4.1). These conserved densities $u_k(\alpha,\beta)$ are given by

$$u_k(\alpha,\beta) = \sum_{m=0}^{[k/2]} \frac{k!}{(k-2m)!(m+1)!m!}\, \alpha^{m+1}\beta^{2k-m} \quad \text{for } k = 1,2,3,\ldots$$

where $[k/2]$ is defined to be the greatest integer less than or equal to $k/2$. It is easy to show that the $u_k(\alpha,\beta)$ so defined are conserved densities of the shallow water equations by direct substitution into equation (4.4).

Another infinite sequence of conservation laws may be obtained by repeated integration of the conservation law

$$\frac{\partial}{\partial t}\beta + \frac{\partial}{\partial x}(\tfrac{1}{2}\beta^2 + \alpha) = 0 .$$

The next conservation law in this sequence is

$$\frac{\partial}{\partial t}(\tfrac{1}{2}\beta^2 + \alpha \log \alpha - \alpha) + \frac{\partial}{\partial x}(\alpha\beta \log \alpha + \frac{1}{3}\beta^3) = 0 .$$

Still other solutions to (4.4) may be obtained by separation of variables assuming a conserved density of the form

$$u(\alpha,\beta) = A(\alpha)B(\beta) .$$

Despite the existence of these infinite families of conservation laws for the shallow water equations, the choice of a conservation law for use as the side condition of a variational principle is severly

restricted. In fact, by integrating the system (2.9) in characteristic coordinates we can show that only two pairs of conservation laws lead to variational principles.

In order to find all variational principles with side conditions we apply the theory of the previous section. The shallow water equations may be written in characteristic coordinates as the quasi-linear system

$$\begin{aligned} \varphi_t + (\tfrac{3}{4}\varphi + \tfrac{1}{4}\psi)\varphi_x &= 0 \\ \psi_t + (\tfrac{1}{4}\varphi + \tfrac{3}{4}\psi)\psi_x &= 0 \end{aligned} \tag{4.5}$$

where

$$\begin{aligned} \varphi &= \beta + 2\sqrt{\alpha} \\ \psi &= \beta - 2\sqrt{\alpha} \ . \end{aligned}$$

Using the values

$$R = \alpha, \quad S = 0, \quad T = -1$$

in the definition following equation (3.3), we determine that

$$\begin{aligned} P &= \tfrac{1}{2}(\varphi - \psi)^{-1} \\ Q &= -\tfrac{1}{2}(\varphi - \psi)^{-1}. \end{aligned}$$

The conditions for the existence of a variational principle with side condition from the beginning of Section III.3 may now be written

$$2(\varphi - \psi)u_{\varphi\psi} + u_{\varphi} - u_{\psi} = 0 \tag{2.6}$$

$$2(\varphi - \psi)w_{\varphi\psi} + w_{\varphi} - w_{\psi} = 0 \tag{4.7}$$

$$w_{\psi}u_{\varphi} + w_{\varphi}u_{\psi} = 0 . \tag{4.8}$$

From equations (3.4) and (3.5) we find that the system may be integrated to the form

$$u_{\psi}w_{\psi} = m(\psi)(\varphi - \psi)$$

$$u_{\varphi}\psi_{\varphi} = n(\varphi)(\varphi - \psi)$$

and equation (3.8) gives

$$m(\psi)u_{\varphi}^{2} + n(\varphi)u_{\psi}^{2} = 0 . \tag{4.9}$$

We see that solutions exist only for

$$m(\psi)n(\varphi) < 0$$

and (4.9) may be factored to

$$(\ell(\psi)u_{\varphi} + k(\varphi)u_{\psi})(\ell(\psi)u_{\varphi} - k(\varphi)u_{\psi}) = 0$$

where

$$k^{2}(\varphi) = \pm n(\varphi), \qquad \ell^{2}(\psi) = \mp m(\psi) .$$

We consider the case

$$\ell(\psi)u_{\varphi} + k(\varphi)u_{\psi} = 0 . \tag{4.10}$$

When we integrate this first order linear equation along characteristics

we find that

$$u = u(s(\varphi, \psi))$$

where

$$s_\varphi = k(\varphi), \qquad s_\psi = -\ell(\psi)$$

$$s = \int^{\varphi} k(\xi)\, d\xi \; - \int^{\psi} \ell(\eta)\, d\eta \; + A \; . \tag{4.11}$$

The function u must satisfy (3.3), that is

$$2(\varphi - \psi)u_{\varphi\psi} + u_\varphi - u_\psi = 0$$

$$-2(\varphi - \psi)k(\varphi)\ell(\psi)\, \ddot{u}(s) + [\, k(\varphi) + \ell(\psi)]\; \dot{u}(s) = 0$$

We find

$$\frac{\ddot{u}(s)}{\dot{u}(s)} = \tfrac{1}{2}(\varphi - \psi)^{-1}\ell^{-1} + \tfrac{1}{2}(\varphi - \psi)^{-1}k^{-1}. \tag{4.12}$$

The left-hand side of (4.12) is a function of s so that the function F defined by

$$\tfrac{1}{2}(\varphi - \psi)^{-1}\ell^{-1}(\psi) + \tfrac{1}{2}(\varphi - \psi)^{-1}k^{-1}(\varphi) = F \tag{4.13}$$

must satisfy the condition

$$0 = s_\psi F_\varphi - s_\varphi F_\psi$$

$$= -\ell[\, -\tfrac{1}{2}(\varphi - \psi)^{-1}\ell^{-1} - \tfrac{1}{2}(\varphi - \psi)^{-2}k^{-1} - \tfrac{1}{2}(\varphi - \psi)^{-1}k^{-1}\dot{k}\,]$$

$$-k[\tfrac{1}{2}(\varphi - \psi)^{-2}\ell^{-1} - \tfrac{1}{2}(\varphi - \psi)^{-1}\ell^{-2}\dot{\ell} + \tfrac{1}{2}(\varphi - \psi)^{-2}k^{-1}] \; .$$

Thus, we have

$$0 = \tfrac{1}{2}(\varphi - \psi)^{-2}k^{-1}\ell + \tfrac{1}{2}(\varphi - \psi)^{-1}k^{-2}\ell\, k'$$

$$- \tfrac{1}{2}(\varphi - \psi)^{-2}k\ell^{-1} + \tfrac{1}{2}(\varphi - \psi)^{-1}k\ell^{-2}\ell' . \qquad (4.14)$$

Upon multiplying both sides of (4.14) by $2(\varphi - \psi)^{2}k^{-1}\ell^{-1}$, we are left with

$$0 = k^{-2} + (\varphi - \psi)k^{-3}k' - \ell^{-2} + (\varphi - \psi)\ell^{-3}\ell' . \qquad (4.13)$$

Disregarding a constant multiplicative factor, the three solutions of (4.15) are given by

$$k_1 = 1 , \qquad \ell_1 = -1 ,$$

$$k_2 = \varphi^{-1} , \qquad \ell_2 = \psi^{-1} ,$$

$$k_3 = \varphi^{-1} , \qquad \ell_3 = -\psi^{-1} .$$

We consider the first solution. Equation (4.10) becomes

$$-u_{,\varphi} + u_{,\psi} = 0$$

which we integrate to

$$u_1 = \varphi + \psi .$$

The corresponding function $w_1(\varphi,\psi)$ is given by

$$w_1 = \tfrac{1}{2}(\varphi - \psi)^{2}.$$

This solution leads to the pair of variational principles which is already well-known for the shallow water equations. In α, β

coordinates the Lagrangian functions are given by

$$\mathfrak{L}_1 = \tfrac{1}{2}\alpha^2 + \lambda\left[\frac{\partial}{\partial t}u_1(\alpha,\beta) + \frac{\partial}{\partial x}f_1(\alpha,\beta)\right]$$

$$\mathfrak{L}_1^I = \tfrac{1}{2}\alpha\beta^2 - \tfrac{1}{2}\alpha^2 + \lambda\left[\frac{\partial}{\partial t}w_1(\alpha,\beta) + \frac{\partial}{\partial x}h_1(\alpha,\beta)\right]$$

where

$$u_1(\alpha,\beta) = \beta \qquad f_1(\alpha,\beta) = \alpha + \tfrac{1}{2}\beta^2$$

$$w_1(\alpha,\beta) = \alpha \qquad h_1(\alpha,\beta) = \alpha\beta\ ,$$

We now consider the second solution of (4.15). Equation (4.10) becomes

$$\psi^{-1}u_{2_\varphi} + \varphi^{-1}u_{2_\psi} = 0$$

from which we determine

$$u_2 = u_2(s)$$

where

$$s = \log\varphi - \log\psi + A\ .$$

In order to determine $u(s)$ we go to equation (4.12)

$$\frac{\ddot{u}_2(s)}{\dot{u}_2(s)} = \tfrac{1}{2}(\varphi - \psi)^{-1}(\varphi + \psi)$$

$$= \tfrac{1}{2}\,\frac{e^s + 1}{e^s - 1}\ .$$

This equation may be integrated to

$$u_2(s) = e^{\frac{1}{2}s} + e^{-\frac{1}{2}s}$$

$$u_2(\varphi,\psi) = \varphi^{\frac{1}{2}}\psi^{-\frac{1}{2}} + \varphi^{-\frac{1}{2}}\psi^{\frac{1}{2}} .$$

The corresponding f_2 is given by

$$f_2(\varphi,\psi) = \frac{1}{4}(\varphi^{3/2}\psi^{-\frac{1}{2}} - 2\varphi^{\frac{1}{2}}\psi^{\frac{1}{2}} + \varphi^{-\frac{1}{2}}\psi^{3/2}) .$$

Equation (4.8) is used to find

$$w_2(\varphi,\psi) = w(s)$$

where

$$s = \varphi\psi$$

and then, from (4.7), we have

$$w_2(\varphi,\psi) = \varphi^{\frac{1}{2}}\psi^{\frac{1}{2}}$$

$$h_2(\varphi,\psi) = \frac{1}{4}\varphi^{\frac{1}{2}}\psi^{\frac{1}{2}}(\varphi + \psi) .$$

Using these values of u, w, f, and h, we integrate the form

$$dH_1 = h_1\, du_1 - w_1\, d f_1$$

to obtain

$$H_1 = -\frac{1}{8}(\varphi - \psi)^2 .$$

Finally, (2.12) gives

$$H_2' = H_2 - u_2h_2 + w_2f_2$$
$$= -\frac{1}{8}\psi^2 - \frac{3}{4}\varphi\psi - \frac{1}{8}\varphi^2 .$$

The second pair of variational principles has Lagrangian functions of the form

$$\mathcal{L}_2 = H_2 + \lambda \left(\frac{\partial}{\partial t} u_2 + \frac{\partial}{\partial x} f_2 \right)$$

and

$$\mathcal{L}_2' = H_2' + \lambda \left(\frac{\partial}{\partial t} w_2 + \frac{\partial}{\partial x} h_2 \right) .$$

Returning to the α, β coordinates we have

$$H_2 = -\tfrac{1}{2}\alpha$$

$$H_2' = 2\alpha - \beta^2$$

$$u = \frac{2\beta}{\sqrt{|4\alpha - \beta^2|}}$$

$$f_2 = \frac{4\alpha}{\sqrt{|4\alpha - \beta^2|}}$$

$$w_2 = \sqrt{|4\alpha - \beta^2|}$$

$$h_2 = \tfrac{1}{2}\beta\sqrt{|4\alpha - \beta^2|} .$$

The third solution of (3.15) does not provide any additional variational principles since

$$u_3 = - w_2, \quad \text{etc.}$$

III. 5. Constraints on the form of a side condition

In the last two sections we assumed from the outset that side conditions were in conservation law form. Now consider the question of the exact form which a side condition must have for use in a variational principle. We shall find that a single side condition must be of the form

$$\frac{\partial}{\partial t} u(\alpha,\beta) + \frac{\partial}{\partial x} f(\alpha,\beta) + c(\alpha,\beta) = 0$$

if it is to lead to a potential representation for α and β.

Suppose we have a quasi-linear system

$$\frac{\partial}{\partial t}\begin{pmatrix}\alpha\\ \beta\end{pmatrix} + \begin{pmatrix}A & B\\ C & D\end{pmatrix}\frac{\partial}{\partial x}\begin{pmatrix}\alpha\\ \beta\end{pmatrix} + \begin{pmatrix}K\\ L\end{pmatrix} = 0 \qquad (5.1)$$

where A, B, C, D, K, L are, in general, functions of α and β. We seek functions $y(\alpha,\beta)$ and $z(\alpha,\beta)$ such that the equation

$$y(\alpha_t + A\alpha_x + B\beta_x + K) + z(\beta_t + C\alpha_x + D\beta_x + L) = 0 \qquad (5.2)$$

may be used as a side condition in a variational principle with Lagrangian

$$\mathfrak{L} = H(\alpha,\beta) + \lambda[\, y(\alpha_t + A\alpha_x + B\beta_x + K) + z(\beta_t + C\alpha_x + D\beta_x + L)] \qquad (5.3)$$

to give a consistent potential representation for α and β in terms of λ and its derivatives. Varying with respect to α and β we obtain

$\delta\alpha$:
$$0 = (y\lambda)_t + (Ay\lambda)_x + (Cz\lambda)_x - H_\alpha$$
$$- \lambda y_\alpha(\alpha_t + A\alpha_x + B\beta_x + K) - \lambda_y(A_\alpha\alpha_x + B_\alpha\beta_x + K_\alpha)$$
$$-\lambda z_\alpha(\beta_t + C\alpha_x + D\beta_x + L) - \lambda_z(C_\alpha\alpha_x + D_\alpha\beta_x + L_\alpha) \quad (5.4)$$

$\delta\beta$:
$$0 = (\lambda By)_x + (\lambda z)_t + (Dz\lambda)_x - H_\beta$$
$$-\lambda y_\beta(\alpha_t + A\alpha_x + B\beta_x + K) - \lambda_y(A_\beta\alpha_x + B_\beta\beta_x + K_\beta)$$
$$-\lambda z_\beta(\beta_t + C\alpha_x + D\beta_x + L) - \lambda_z(C_\beta\alpha_x + D_\beta\beta_x + L_\beta). \quad (5.5)$$

In order to obtain a representation for α and β we require the coefficients of α_t, β_t, α_x and β_x to vanish. From (5.4) we find

β_t:
$$0 = y_\beta - z_\alpha \quad (5.6)$$

β_x:
$$0 = (Ay)_\beta + (Cz)_\beta - (By)_\alpha - (Dz)_\alpha . \quad (5.7)$$

No further conditions may be obtained from (5.5).

From (5.6) we infer the existence of a function $u(\alpha, b)$ such that

$$y = u_\alpha$$
$$z = u_\beta .$$

Similarly, from (5.7) we infer the existence of a function $f(\alpha, \beta)$ such that

$$Ay + Cz = f_\alpha$$
$$By + Dz = f_\beta .$$

If we define a function $c(\alpha,\beta)$ by

$$c = yK + zL ,$$

then the original Lagrangian (5.3) may be written in terms of u, f and c. It is

$$\mathfrak{L} = H(\alpha,\beta) + \lambda\left[\frac{\partial}{\partial t}u(\alpha,\beta) + \frac{\partial}{\partial x}f(\alpha,\beta) + c(\alpha,\beta)\right].$$

We have shown that if any combination

$$y(\alpha_t + A\alpha_t + B\beta_x + K) + z(\beta_t + C\alpha_x + D\beta_x + L) = 0$$

of the equations in (5.1) is used as the side condition in a variational principle then y and z are quite special. The side condition may be written

$$u_t + f_x + c = 0 .$$

In particular, if the original system contained no undifferentiated terms then the side condition must be a conservation law.

III.6. Two equations with undifferentiated terms

The presence of an undifferentiated term complicates the analysis of the variational principle with

$$u_t + f_x + c = 0$$

as a side condition. Consider the variational principle

$$J = \int\int \left\{H + \lambda(u_t + f_x + c)\right\} dx\, dt$$

whose Euler equations are

$$\begin{aligned} H_\alpha + \lambda c_\alpha &= \lambda_t u_\alpha + \lambda_x f_\alpha \\ H_\beta + \lambda c_\beta &= \lambda_t u_\beta + \lambda_x f_\beta \,. \end{aligned} \tag{6.1}$$

In order to obtain a more compact notation we introduce the Poisson bracket (y,z) defined by

$$(y,z) = y_\alpha z_\beta - y_\beta z_\alpha \,.$$

The system (6.1) may be solved for λ_t and λ_x in terms of α, β, and λ in the form

$$\begin{aligned} \lambda_t &= -h - \lambda g \\ \lambda_x &= w + \lambda v \end{aligned} \tag{6.2}$$

where w, v, h, and g are given by

$$w = \frac{(u,H)}{(u,f)} \,, \qquad h = -\frac{(H,f)}{(u,f)} \,,$$

$$v = \frac{(u,c)}{(u,f)} \,, \qquad g = -\frac{(c,f)}{(u,f)} \,.$$

In the previously considered case of two conservation laws we were able to proceed directly to the equation $w_t + h_x = 0$. The presence of an undifferentiated term does not enable us to do this directly but the idea of obtaining further restrictions from the consistency condition for λ does seem right. From (6.2) we obtain

$$0 = \lambda_{xt} - \lambda_{tx}$$

$$= w_t + \lambda v_t + \lambda_t v + h_x + \lambda g_x + \lambda_x g$$

$$= w_t + h_x + (wg - vh) + \lambda(v_t + g_x) . \qquad (6.3)$$

We require (6.3) to hold independent of the choice of λ to obtain two additional equations

$$0 = \frac{\partial}{\partial t} v(\alpha,\beta) + \frac{\partial}{\partial x} g(\alpha,\beta) \qquad (6.4)$$

and

$$0 = \frac{\partial}{\partial t} w + \frac{\partial}{\partial x} h + wg - vh . \qquad (6.5)$$

The first of these (6.4) is actually a restriction on the choice of the side condition since it may be written

$$0 = \frac{\partial}{\partial t}\left[\frac{(u,c)}{(u,f)}\right] + \frac{\partial}{\partial x}\left[\frac{(f,c)}{(u,f)}\right] ,$$

a form which involves only u, f and c. For linear systems and non-linear systems containing no undifferentiated term this equation is trivially satisfied. In the second equation (6.5) may be simplified using the Poisson bracket notation. We have

$$wg - vh = \frac{(u,H)}{(u,f)}\frac{(f,c)}{(u,f)} - \frac{(u,c)}{(u,f)}\frac{(f,H)}{(u,f)}$$

$$= \frac{(H,c)}{(u,f)} .$$

The second equation may now be written

$$0 = \frac{\partial}{\partial t}\left[\frac{(u,H)}{(u,f)}\right] + \frac{\partial}{\partial x}\left[\frac{(f,H)}{(u,f)}\right] + \frac{(H,c)}{(u,f)} \tag{6.6}$$

or

$$0 = \frac{\partial}{\partial t} w(\alpha,\beta) + \frac{\partial}{\partial x} h(\alpha,\beta) + d(\alpha,\beta)$$

where

$$d = \frac{(H,c)}{(u,f)} .$$

In the case of two conservation laws considered in section III-2 we obtained further restrictions of the choice of the side condition by applying the consistency condition on H to the system (2.5) (analogous to (6.1) here). Equation (6.1) may be written in terms of v, w, g, h in the form

$$\begin{aligned} H_\alpha + \lambda c_\alpha &= - h u_\alpha + w f_\alpha + \lambda(-g u_\alpha + v f_\alpha) \\ H_\beta + \lambda c_\beta &= - h u_\beta + w f_\beta + \lambda(-g u_\beta + v f_\beta) \end{aligned} \tag{6.7}$$

It appears that the choice of the side condition will now be restricted by the consistency conditions for both H and c. It is possible to show that if

$$\frac{\partial}{\partial t} v(\alpha,\beta) + \frac{\partial}{\partial x} g(\alpha,\beta) = 0$$

is a conservation law of the system (5.1) then these two consistency conditions are equivalent. This result provides some confirmation of the method of separating (6.3) into (6.4) and (6.5).

Applied to $H(\alpha,\beta)$, the consistency condition may be written

$$0 = H_{\alpha\beta} - H_{\beta\alpha}$$

$$= (-hu_\alpha + wf_\alpha)_\beta - (-hu_\beta + wf_\beta)_\alpha$$

$$= (h,u) + (f,w) \qquad (6.8)$$

Before considering the consistency condition for c we will first observe that if

$$\frac{\partial}{\partial t} v(\alpha,\beta) + \frac{\partial}{\partial x} g(\alpha,\beta) = 0$$

belongs to the system determined by

$$u_t + f_x + c = 0$$

$$w_t + h_x + d = 0 ,$$

then an integrating function μ exists such that

$$dv = \mu d\, dw - \mu c\, du$$

and

$$dg = \mu d\, dh - \mu c\, df .$$

Now we apply the consistency condition to c (from (6.7)) to obtain

$$0 = (-gu_\alpha + vf_\alpha)_\beta - (-gu_\beta + vf_\beta)_\alpha$$

$$= g_\alpha u_\beta - g_\beta u_\alpha + f_\alpha v_\beta - f_\beta v_\alpha$$

$$= (\mu df_\alpha - \mu ch_\alpha)u_\beta - (\mu df_\beta - \mu ch_\beta)u_\alpha$$

$$+ f_\alpha(\mu du_\beta - \mu cw_\beta) - f_\beta(\mu du_\alpha - \mu cw_\alpha)$$

$$= \mu c[\,(u,h) + (w,f)] . \qquad (6.9)$$

Since μ and c are non-zero, (6.9) is equivalent to (6.8) above, the consistency condition for H.

In summary, if the undifferentiated term is included in (5.1), then the conditions for the existence of a variational principle become even more restrictive than they were for the case of two conservation laws. It is necessary to find u, v, w, f, g, h such that

$$\frac{\partial}{\partial t} u(\alpha,\beta) + \frac{\partial}{\partial x} f(\alpha,\beta) + c(\alpha,\beta) = 0 \tag{6.10}$$

$$\frac{\partial}{\partial t} v(\alpha,\beta) + \frac{\partial}{\partial x} g(\alpha,\beta) = 0 \tag{6.11}$$

$$\frac{\partial}{\partial t} w(\alpha,\beta) + \frac{\partial}{\partial x} h(\alpha,\beta) + d(\alpha,\beta) = 0 \tag{6.12}$$

all belong to the system (4.1). Furthermore we require

$$(u,h) + (w,f) = 0 . \tag{6.13}$$

Equations (6.10) and (6.12) correspond to (2.4) and (2.7) respectively. It is easy to show that (6.13) is equivalent to (2.8). The requirement that the system (5.1) must have a conservation law (6.11) certainly poses a severe restriction on the kinds of systems containing undifferentated terms for which variational principles may be found.

REFERENCES

1. John von Neumann, "Memorandum to O. Veblen," Collected Works, Vol. VI, pp. 357-9, A. H. Lamb, ed., Pergamon Press, New York, 1963.

2. G. B. Whitham, J. Fluid Mech. 22, 273 (1965).

3. G. B. Whitham, Proc. Roy. Soc. A299, 6 (1967).

4. G. B. Whitham, J. Fluid Mech. 27, 399 (1967).

5. G. B. Whitham, J. Fluid Mech. 44, 373 (1970).

6. J. W. Herivel, Proc. Camb. Phil. Soc. 51, 344 (1955).

7. C. C. Lin, Liquid helium, Proc. Int. School of Physics, Course XXI, p. 91, Academic Press, New York, 1963.

8. R. L. Seliger and G. B. Whitham, Proc. Roy. Soc. A305, 1 (1968).

9. B. F. Schutz, Relativistic velocity-potentials, hydrodynamics and stellar stability, Ph.D. Thesis, California Institute of Technology (1972).

10. J. Serrin, "Mathematical principles of classical fluid mechanics," in Handbuch der Physik, vol. 8/1, p. 125, S. Flugge, ed., Springer-Verlag, Berlin, 1959.

11. H. Lamb, Hydrodynamics, Sixth ed., p. 249, Dover, New York, 1932.