JOURNAL OF DIFFERENTIAL EQUATIONS 71, 10–33 (1988)

Darboux' Theorem for Hamiltonian Differential Operators

PETER J. OLVER[1]

School of Mathematics, University of Minnesota, Minneapolis, 55455 Minnesota

Received April 15, 1986

It is proved that any one-dimensional, first order Hamiltonian differential operator can be put into constant coefficient form by a suitable change of variables. Consequently, there exist canonical variables for any such Hamiltonian operator. In the course of the proof, a complete characterization of all first order Hamiltonian differential operators, as well as the general formula for the behavior of a Hamiltonian operator under a change of variables involving both the independent and the dependent variables are found. © 1988 Academic Press, Inc.

1. INTRODUCTION

In recent years there has been a great surge of interest in Hamiltonian systems of partial differential equations; see the papers in [10] for a representative sample of current research in this area. Although they are the natural infinite dimensional counterparts of the finite dimensional Hamiltonian systems of classical mechanics, the correct formulation of a Hamiltonian system of evolution equations has only been arrived at in the last decade. (See [15; Chap. 7] for a historical survey, as well as an introduction to the general theory.) Part of the difficulty in making the jump to infinite dimensions has been the excessive reliance on canonical coordinates, which, for finite dimensional Hamiltonian systems, are always guaranteed to exist by Darboux' Theorem, (cf. [18]). For evolution equations the Hamiltonian operators are usually differential operators, and it is a significant open problem as to whether some version of Darboux' Theorem allowing one to change to canonical variables is valid in this context. In this paper and the companion paper [16], we prove that in one special case—first order Hamiltonian differential operators for a single evolution equation—Darboux' Theorem is valid, and it is always possible to choose canonical coordinates.

Actually, there *is* an infinite-dimensional version of Darboux' Theorem due to Weinstein [19], but it does not appear to be applicable to the

[1] Research was supported in part by NSF Grant MCS 86–02004 and NATO Collaborative Research Grant RG 86/0055.

0022-0396/88 $3.00

Hamiltonian differential operators of interest. Weinstein requires some form of Banach manifold structure to effect his proof, but for differential operators which depend on the dependent variables it is not at all obvious how to impose such a structure. Even if one could mimic Weinstein's proof, the resulting changes of variable would be horribly nonlocal, and therefore be of limited use. Nevertheless, we can try to follow the broad outlines of Weinstein's proof, which consists of two distinct stages. First, the Hamiltonian operator is reduced to a constant coefficient differential operator—this is the harder part and is the one explicitly addressed in this paper. Second, it is shown how any constant coefficient Hamiltonian operator is equivalent to one in canonical form. This latter step turns out to be closely related to Dirac's theory of constraints for field theories [1, 11], and is dealt with in the companion paper [16].

A fundamental question is what type of changes of variable are allowed in the proof of Darboux' Theorem for differential operators? If x denotes the independent variables and u the dependent variables in the system of evolution equations, then it is *not* sufficient to just consider local changes in the dependent variable, $w = Q[u]$, where Q depends on x, u, and, possibly, derivatives of u; this will be shown in a particular example treated in the Appendix to the paper. However, if one also allows similar changes in the independent variables, so $w = Q[u]$, and $y = P[u]$, which are called *differential substitutions* by Ibragimov [6], then one *can* prove a version of Darboux' Theorem for first order operators. Thus, the class of changes of variable must be enlarged beyond what one might ordinary expect.

The proof itself is constructive. First, we use a slight generalization of a result of Gel'fand and Dorfman [4] to completely characterize all first order Hamiltonian operators. Then, given such a differential operator the requisite differential substitution that changes it into the constant coefficient form D_y is explicitly determined. This turns out to reduce to the Darboux Theorem for closed differential two-forms in $\mathbb{R}^3$! Although this constructive method of proof is fine from a computational point of view, it certainly does not admit easy generalization to either higher order operators or matrix operators. Thus the general truth of Darboux' Theorem for differential operators remains completely open, and of great significance. Time will tell whether it is true or false. (It should be noted that results of Dubrovin and Novikov [2, 3] indicate that Darboux' Theorem does not hold for matrix operators involving more than one independent variable. However, they do not consider the most general change of variables presented here, and it is not at all clear whether their result remains valid in this more general context.)

Finally, there is the issue of why one needs to find canonical coordinates for Hamiltonian systems. Among all the reasons, three stand out. First, most quantization procedures require that the Hamiltonian system be in canonical form before proceeding. (A significant exception is the theory of geometric quantization [20]. However, I am unaware of any attempts to apply this complicated theory to Hamiltonian systems of evolution equations.) Another reason is that Hamiltonian perturbation theories are much easier to develop in canonical coordinates (cf. [8, 14]). Finally, the fundamental theorem of Magri on biHamiltonian systems [9] can be significantly strengthened if one of the Hamiltonian operators is in constant coefficient form (cf. [17]).

2. Hamiltonian Operators

In this paper, we will be exclusively deal with Hamiltonian evolution equations involving a single spatial variable x and a single dependent variable u, although the present treatment of Hamiltonian systems of evolution equations easily generalizes to systems in several independent variables. The basic equation takes the form

$$u_t = K[u] = K(x, u, u_x, u_{xx}, \dots), \tag{1}$$

where the square brackets indicate that K is a *differential function*, meaning that it depends on x, u, and finitely many derivatives of u with respect to x; the solutions are then functions $u = u(x, t)$ depending on x and the temporal variable t. The evolution equation is said to be *Hamiltonian* if it can be written in the form

$$u_t = \mathcal{D} \cdot E_u(H),$$

where $\mathcal{D}$ is the Hamiltonian operator, $\mathcal{H}[u] = \int H[u]\, dx$ is the Hamiltonian functional, and E_u denotes the Euler operator or variational derivative with respect to u. The Hamiltonian operator $\mathcal{D}$ is a linear differential operator, which may depend on both x and u (but not t), and is required to be (formally) skew-adjoint relative to the L^2-inner product $\langle f, g \rangle = \int f \cdot g\, dx$ as well as satisfy a nonlinear "Jacobi condition" that the corresponding Poisson bracket

$$\{\mathcal{P}, \mathcal{Q}\} = \int E_u(P) \cdot \mathcal{D} E_u(Q)\, dx, \qquad \mathcal{P} = \int P[u]\, dx, \quad \mathcal{Q} = \int Q[u]\, dx,$$

satisfy the Jacobi identity.

The complicated Jacobi conditions for $\mathcal{D}$ can be considerably simplified by the "functional multi-vector" methods of [15; Chap. 7], which we now briefly review. (See also, [4, 13] for similar techniques.) Coordinates on the infinite jet space J_∞ are provided by the spatial variable x, the dependent variable u, and its derivatives u_x, u_{xx}, ..., $u_j = \partial^j u/\partial x^j$, ... A basis for the space of *vertical one-forms* to J_∞ is provided by the differentials du, du_x, du_{xx}, ..., du_j, ...; the dual basis for the space of vertical tangent vectors, or "uni-vectors," is denoted by θ, θ_x, θ_{xx}, ..., θ_j, ... (At this stage, the alternative notation $\partial/\partial u_j$ for θ_j would seem a bit more natural; however, this latter notation quickly becomes confusing, which is the reason for our choice of θ_j for these objects.) A *vertical multi-vector* is a finite sum of terms, each of which is the product of a differential function times a wedge product of the basic uni-vectors; for example,

$$\hat{\Theta} = xU_x\, \theta \wedge \theta_x$$

is a vertical bi-vector. The total derivative D_x acts as a Lie derivative on these vertical multi-vectors, with $D_x(u_j) = u_{j+1}$, and $D_x(\theta_j) = \theta_{j+1}$. For the above example,

$$D_x \hat{\Theta} = (xu_{xx} + u_x)\, \theta \wedge \theta_x + xu_x \theta \wedge \theta_{xx}.$$

(Since $\theta_x \wedge \theta_x = 0$.) The space of *functional multi-vectors* is the cokernel of D_x, so that two vertical multi-vectors determine the same functional multi-vector if and only if they differ by a total derivative. The functional multi-vector determined by $\hat{\Theta}$ is denoted, suggestively, by an integral sign: $\Theta = \int \hat{\Theta}\, dx$. In particular $\int \hat{\Theta}\, dx = 0$ if and only if $\hat{\Theta} = D_x \hat{\Psi}$ for some vertical multi-vector $\hat{\Psi}$. This implies that we can integrate functional multi-vectors by parts:

$$\int \hat{\Theta} \wedge (D_x \hat{\Psi})\, dx = -\int (D_x \hat{\Theta}) \wedge \hat{\Psi}\, dx. \qquad (2)$$

The principal example of a functional bi-vector is that determined by a Hamiltonian differential operator $\mathscr{D}$, which is

$$\Theta_{\mathscr{D}} = \int \theta \wedge \mathscr{D}(\theta)\, dx.$$

For instance, if $\mathscr{D} = 2uD_x + u_x$ (which is Hamiltonian), then

$$\Theta_{\mathscr{D}} = \int \theta \wedge (2uD_x\theta + u_x\theta)\, dx = \int 2u\theta \wedge \theta_x\, dx.$$

Finally, define the formal prolonged vector field

$$\text{pr}\, v_{\mathscr{D}\theta} = \mathscr{D}(\theta) \cdot \frac{\partial}{\partial u} + D_x \mathscr{D}(\theta) \frac{\partial}{\partial u_x} + D_x^2 \mathscr{D}(\theta) \cdot \frac{\partial}{\partial u_{xx}} + \cdots,$$

which acts on differential functions to produce uni-vectors; for example, if $\mathscr{D}$ is as above, and $P[u]=xuu_{xx}$, then

$$\begin{aligned}\operatorname{pr}\mathbf{v}_{\mathscr{D}\theta}(P) &= xu_{xx}\mathscr{D}(\theta)+xuD_x^2\mathscr{D}(\theta)\\ &= xu_{xx}(2u\theta_x+u_x\theta)+xuD_x^2(2u\theta_x+u_x\theta)\\ &= 2xu^2\theta_{xxx}+5xuu_x\theta_{xx}+6xuu_{xx}\theta_x+x(uu_{xxx}+u_xu_{xx})\,\theta.\end{aligned}$$

We further let $\operatorname{pr}\mathbf{v}_{\mathscr{D}\theta}$ act on vertical multi-vectors by wedging the result of its action on the coefficient differential functions with the product of the θ's, so, for example, $\operatorname{pr}\mathbf{v}_{\mathscr{D}\theta}(P\cdot\theta\wedge\theta_x)=\operatorname{pr}\mathbf{v}_{\mathscr{D}\theta}(P)\wedge\theta\wedge\theta_x$. Since $\operatorname{pr}\mathbf{v}_{\mathscr{D}\theta}$ commutes with the total derivative, there is also a well-defined action of $\operatorname{pr}\mathbf{v}_{\mathscr{D}\theta}$ on the space of functional multi-vectors, which essentially amounts to bringing it "under the integral sign."

THEOREM 1. *Let $\mathscr{D}$ be a skew-adjoint differential operator with corresponding bi-vector $\Theta_{\mathscr{D}}$ as above. Then $\mathscr{D}$ is a Hamiltonian operator if and only*

$$\operatorname{pr}\mathbf{v}_{\mathscr{D}\theta}(\Theta_{\mathscr{D}})=0. \tag{3}$$

The proof that (3) is equivalent to the Jacobi identity for the Poisson bracket determined by $\mathscr{D}$ can be found in [15; Chap. 7]. As a simple example, if $\mathscr{D}=2uD_x+u_x$, then (3) reads

$$\begin{aligned}\operatorname{pr}\mathbf{v}_{\mathscr{D}\theta}\left\{\int 2u\theta\wedge\theta_x\,dx\right\} &= \int \operatorname{pr}\mathbf{v}_{\mathscr{D}\theta}(2u)\wedge\theta\wedge\theta_x\,dx\\ &= \int 2(\mathscr{D}\theta)\wedge\theta\wedge\theta_x\,dx\\ &= \int 2(2u\theta_x+u_x\theta)\wedge\theta\wedge\theta_x\,dx=0,\end{aligned}$$

which trivially vanishes. This proves that $2uD_x+u_x$ is a Hamiltonian differential operator.

3. FIRST ORDER OPERATORS

In [4] Gel'fand and Dorfman claim to give a complete characterization of all first order Hamiltonian operators, but they only allow the coefficient functions of their operators to depend on u and the derivatives of u. Here we generalize their result to include operators which depend explicitly on x also.

THEOREM 2. *Let $\mathscr{D}$ be a first order Hamiltonian operator. Then $\mathscr{D}$ must be of the form*

$$\mathscr{D}=(E(A))^{-1}\cdot D_x\cdot(E(A))^{-1}, \tag{4}$$

where $A=A(x,u,u_x)$ depends only on first order derivatives of u, and $E(A)=A_u-D_xA_{u_x}$ is the usual Euler operator applied to A, with subscripts on A indicating partial derivatives. (The operator in (4) acts on a differential function by first dividing by $E(A)$, then differentiating, and then again dividing by $E(A)$.)

Proof. By an elementary lemma, every scalar first order skew-adjoint differential operator is of the form

$$\mathscr{D}=B[u]\cdot D_x\cdot B[u]$$

for some differential function B; the corresponding functional bi-vector is

$$\Theta_{\mathscr{D}}=\int B^2\cdot\theta\wedge\theta_x\,dx.$$

The tri-vector (3) whose vanishing proves the Jacobi identity is

$$\operatorname{pr}\mathbf{v}_{\mathscr{D}\theta}(\Theta_{\mathscr{D}})=\int 2B\cdot\sum\frac{\partial B}{\partial u_j}\cdot D_x^j(B^2\theta_x+B\cdot D_xB\cdot\theta)\wedge\theta\wedge\theta_x\,dx, \tag{5}$$

where $u_j=D_x^j(u)$, $\theta_j=D_x^j(\theta)$, and the sum is from $j=0$ to n, the highest order derivative of u appearing in B. Once the total derivatives are expanded, (5) can be written as a sum of terms of the form

$$\Psi_k=\int P_k[u]\cdot\theta_k\wedge\theta\wedge\theta_x\,dx,$$

for certain differential functions P_k, of which the highest order one is

$$\Psi_n=\int 2B^3\cdot\frac{\partial B}{\partial u_n}\cdot\theta_{n+1}\wedge\theta\wedge\theta_x\,dx.$$

The theorem rests on the following two results, the first of which provides a complete characterization of functional tri-vectors depending on one independent and one dependent variable, and is the key to any serious analysis of these objects.

LEMMA 3. *The tri-vectors $\theta_i \wedge \theta_j \wedge \theta_{j+1}$, $0 \leqslant i < j$ form a basis for the space of functional tri-vectors. In other words, every functional tri-vector can be written uniquely in the form*

$$\Psi = \int \sum_{i<j} A_{ij}[u] \cdot \theta_i \wedge \theta_j \wedge \theta_{j+1}\, dx \tag{6}$$

for certain uniquely determined differential functions A_{ij}. In particular, $\Psi = 0$ if and only if $A_{ij} = 0$ for all $i < j$.

Proof. The proof requires the methods from [15; Chap. 7], which we quote without proof. It is not difficult to use integration by parts to prove that every tri-vector depending on one independent and one dependent variable can be written in the form (6) for some finite set of coefficient functions A_{ij}; the only difficulty is with the uniqueness, and for this it suffices to prove that if $\Psi = 0$ then all the A_{ij} vanish. Suppose not, so for some natural number n, $A_{ij} = 0$ for all $j > n$, but $A_{in} \neq 0$ for at least one $i < n$. By Lemma 5.61 and the remark on page 431 of [15], the tri-vector (6) vanishes if and only if its evaluation

$$\langle \Psi; P, Q, R \rangle = \int \sum_{i<j\leqslant n} A_{ij} \cdot \det \begin{bmatrix} D_x^i P & D_x^i Q & D_x^i R \\ D_x^j P & D_x^j Q & D_x^j R \\ D_x^{j+1} P & D_x^{j+1} Q & D_x^{j+1} R \end{bmatrix} dx$$

vanishes for an arbitrary triple of differential functions P, Q, R. Expanding the determinant and integrating by parts, we arrive at a functional of the form $\int L \cdot R\, dx$, where

$$\begin{aligned} L = \sum \{ & (-D_x)^{j+1} A_{ij} (D_x^i P \cdot D_x^j Q - D_x^j P \cdot D_x^i Q) \\ & + (-D_x)^j A_{ij} (D_x^{j+1} P \cdot D_x^i Q - D_x^i P \cdot D_x^{j+1} Q) \\ & + (-D_x)^i A_{ij} (D_x^j P \cdot D_x^{j+1} Q - D_x^{j+1} P \cdot D_x^j Q) \}. \end{aligned}$$

According to Corollary 5.52 of [15], the functional $\int L \cdot R\, dx$ vanishes for all differential functions R if and only if L itself is zero, and this in turn must occur for all P and Q. Now the highest order derivative of Q appearing in L is $D_x^{2n+1} Q$, and its coefficient is easily seen to be

$$2(-1)^{n+1} \sum_i A_{in} \cdot D_x^i P,$$

which must necessary vanish for all P if L itself is to vanish for all differential functions Q. But this is impossible unless all the coefficient functions A_{in} vanish, which contradicts our initial assumption. This proves that all the A_{ij} in (6) vanish if Ψ is to vanish and completes the proof of uniqueness.

We now apply this basic result towards the vanishing of the Jacobi tri-vector (5) for our differential operator $\mathscr{D}$.

LEMMA 4. *The functional tri-vector*

$$\Psi = \int \sum_{k=2}^{n} P_k[u] \cdot \theta_k \wedge \theta \wedge \theta_x \, dx$$

vanishes if and only if

$$P_2 = D_x P_3 \qquad \text{and} \qquad P_k = 0 \qquad \text{for} \quad k \geqslant 4. \tag{7}$$

Proof. In order to apply Lemma 3 to the summand $\Psi_k = \int P_k[u] \cdot \theta_k \wedge \theta \wedge \theta_x \, dx$, we need to use the integration by parts formula (2) (with $\hat{\Psi} = \theta_k$ and $\hat{\Theta} = P_k \cdot \theta \wedge \theta_x$ in the initial application):

$$\begin{aligned}
\Psi_k &= -\int \{P_k \cdot \theta_{k-1} \wedge \theta \wedge \theta_{xx} + D_x P_k \cdot \theta_{k-1} \wedge \theta \wedge \theta_x\} \, dx \\
&= +\int \{P_k \cdot \theta_{k-2} \wedge \theta \wedge \theta_{xxx} + P_k \cdot \theta_{k-2} \wedge \theta_x \wedge \theta_{xx} + \cdots\} \, dx \\
&= -\int \{P_k \cdot \theta_{k-3} \wedge \theta \wedge \theta_4 + 2P_k \cdot \theta_{k-3} \wedge \theta_x \wedge \theta_{xxx} + \cdots\} \, dx \\
&= \cdots.
\end{aligned}$$

Continuing this process, the net result will be different depending on whether k is even or odd. If $k = 2l$ is even, then we end up with

$$\Psi_k = (-1)^{l-1} \int \{P_k \cdot \theta_{l+1} \wedge \theta \wedge \theta_l + \cdots\} \, dx,$$

where the omitted terms involve lower order basis tri-vectors, i.e., $\theta_i \wedge \theta_j \wedge \theta_{j+1}$, for $j < l$. On the other hand, if $k = 2l + 1$ is odd, the final version is

$$\Psi_k = (-1)^{l-1} \int \{(l-1) P_k \cdot \theta_{l+1} \wedge \theta_x \wedge \theta_l + \cdots\} \, dx,$$

the omitted terms again involving only lower order basis tri-vectors. Thus, after integrating by parts in the prescribed manner, if $n \geq 4$, the highest order basis tri-vector in $\Psi = \sum \Psi_k$ appears with a nonzero multiple of P_n, and hence the only way Ψ can vanish is if all $P_k = 0$ for $k \geq 4$. The $k = 3$ term is a bit different, with

$$\Psi_3 = \int P_3 \cdot \theta_{xxx} \wedge \theta \wedge \theta_x \, dx = -\int D_x P_3 \cdot \theta_{xx} \wedge \theta \wedge \theta_x \, dx.$$

Adding in Ψ_2 verifies the first condition in (7) and completes the proof of the lemma.

To prove the theorem, we apply Lemma 4 to check the vanishing of the tri-vector (5). The latter conditions in (7) imply immediately that B cannot depend on any derivatives of u of higher than second order, so $B = B(x, u, u_x, u_{xx})$. To implement the first condition in (7), we compute

$$P_3 = 2B^3 B_{u_{xx}} \qquad \text{and} \qquad P_2 = 2(5B^2 B_{u_{xx}} D_x B + B^3 B_{u_x}),$$

hence we require that

$$B(D_x B_{u_{xx}} - B_{u_x}) = 2B_{u_{xx}} D_x B \tag{8}$$

in order that $\mathscr{D}$ be Hamiltonian. (In (8) the subscripts on B denote partial derivatives.) To solve (8), we first note that the terms involving u_{xxx} are

$$B \cdot B_{u_{xx}u_{xx}} = 2(B_{u_{xx}})^2,$$

hence

$$B = 1/(F \cdot u_{xx} + G), \tag{9}$$

where $F(x, u, u_x)$ and $G(x, u, u_x)$ depend on at most first order derivatives of u. Substituting this expression into (8), we are left with the single condition

$$F_{u_x} \cdot u_{xx} + G_{u_x} = D_x F = F_x + F_u \cdot u_x + F_{u_x} \cdot u_{xx}.$$

Thus B, as given by (9), will determine a Hamiltonian differential operator $B \cdot D_x \cdot B$ if and only if F and G satisfy the first order partial differential equation

$$G_{u_x} = F_x + F_u \cdot u_x. \tag{10}$$

The general solution to (10) is

$$F = -\frac{\partial^2 A}{\partial u_x^2}, \qquad G = -\frac{\partial^2 A}{\partial x \partial u_x} - u_x \frac{\partial^2 A}{\partial u \partial u_x} + \frac{\partial A}{\partial u},$$

where $A(x, u, u_x)$ is arbitrary. Therefore

$$F \cdot u_{xx} + G = -D_x \frac{\partial A}{\partial u_x} + \frac{\partial A}{\partial u} = E(A),$$

which completes the proof of (4).

As mentioned above, Gel'fand and Dorfman have previously treated the case when B does not depend explicitly on x, in which case they determined that the coefficient B in $\mathscr{D} = B \cdot D_x \cdot B$ has the form

$$B(u, u_x, u_{xx}) = u_x / D_x C = u_x / (C_u \cdot u_x + C_{u_x} \cdot u_{xx}),$$

where $C(u, u_x)$ depends only on u and u_x. Their result, while seemingly dissimilar, is just a special case of the present result; indeed if we let

$$C(u, u_x) = u_x \frac{\partial A}{\partial u_x} - A,$$

where $A(u, u_x)$ does not depend on x, then, as the reader can check,

$$D_x C = u_x \cdot E(A),$$

and so the two formulas for B agree. Special cases are of interest. If C depends only on u (equivalently, A depends only on u) then we find that any operator of the form

$$\frac{f(u_x)}{u_{xx}} \cdot D_x \cdot \frac{f(u_x)}{u_{xx}},$$

where f is any function of u, is always Hamiltonian In particular, if $f(u) = \sqrt{2u}$ then we recover our earlier operator $2u \cdot D_x + u_x$. Alternatively, if C (or A) depend only on u_x, then we find that any operator of the form

$$\frac{f(u_x)}{u_{xx}} \cdot D_x \cdot \frac{f(u_x)}{u_{xx}},$$

where f again is an arbitrary function, is always Hamiltonian. The general case is like a combination of these two subclasses of Hamiltonian operators.

4. Third Order Operators

Although Gel'fand and Dorfman's claim to have completely characterized all first order Hamiltonian operators is simply misleading, in that they do not consider operators whose coefficients explicitly depend on the independent variable x, their claim in [4] to have completely characterized all third order Hamiltonian operators is completely wrong! Their general form of third order operator, which is equivalent to the expression

$$\mathscr{D} = (f'(u))^{-1} \cdot \{D_x^3 + \sqrt{f(u)} \cdot D_x \cdot \sqrt{f(u)}\} \cdot (f'(u))^{-1},$$

while certainly Hamiltonian, fails to include such elementary Hamiltonian operators as $D_x \cdot (u_x)^{-1} \cdot D_x \cdot (u_x)^{-1} \cdot D_x$ which occur in the case of nonlinear wave equations of the form $u_t = h(u)\, u_x$. In fact, it is not difficult to see that Gel'fand and Dorfman's operator is trivially equivalent to the second Hamiltonian operator of the Korteweg–deVries equation (cf. [15; Chap. 7])—just set $w = \sqrt{2} \cdot f(u)$, so that $\mathscr{D}$ becomes the operator

$$D_x^3 + \sqrt{2w} \cdot D_x \cdot \sqrt{2w} = D_x^3 + 2w \cdot D_x + w_x.$$

But there exist many more third order operators. (Incidentally, this operator *is* equivalent to the constant coefficient operator D_x under the Miura transformation $w = u_x - \frac{1}{2}u^2$. (See [16]).)

In point of fact, it *is* possible to generalize the methods of the preceding section to determine the general form of a third order Hamiltonian operator, but the intervening calculations are just too complex to warrant a detailed solution at this time. Suffice it to say that for a general third order operator $\mathscr{D} = P \cdot D_x^3 + D_x^3 \cdot P + Q \cdot D_x + D_x \cdot Q$, one can show that P can depend on at most fourth order derivatives and Q on at most sixth order derivatives of u. The resulting Jacobi conditions amount to eight complicated nonlinear differential equations for P and Q, which I have not attempted to solve. (However, see the note added in proof.)

5. Changes of Variable

Darboux' Theorem is concerned with the problem of simplifying Hamiltonian operators using an appropriate change of variables. In the present infinite-dimensional case, we can change not only the dependent variable u, but also the independent variable x. As we shall see, in order to prove a Darboux Theorem for Hamiltonian differential operators, we must be allowed to change both u and x in an arbitrary fashion, including hodograph-like transformations that mix up the roles of independent and dependent variables. The importance of such transformations for the theory of symmetry groups of differential equations was emphasized in [12].

Thus, along with Ibragimov [6], we are lead to consider general *differential substitutions*

$$y = P[u], \quad w = Q[u], \tag{11}$$

in which y is the new independent variable and w the new dependent variable, and P and Q are differential functions, which therefore are allowed to depend on x, u, *and* derivatives of u. The main technical complicated is that inversion of the change of variables (11) is a nonlocal operation, which involves the solution of the two ordinary differential equations determined by P and Q. However, given u as a function of x, Eq. (11) will (usually) determine w locally as a function of y. The goal now is to see how various operators change when subjected to (11). The easiest is the total derivative D_x, whose transformation rule is determined by the chain rule from elementary calculus:

$$D_x = D_xP \cdot D_y \qquad \text{or} \qquad D_y = (D_xP)^{-1} \cdot D_x, \tag{12}$$

where $(D_xP)^{-1}$ is just the reciprocal $1/D_xP$.

To determine more complicated changes of variable, we first need to recall the *Fréchet derivative* of a differential function, which is the differential operator D_P defined by the formula

$$D_P(v) = \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} P[u+\varepsilon v];$$

more explicitly,

$$D_P = \sum \frac{\partial P}{\partial u_k} \cdot D_x^k.$$

If $\mathscr{D}$ is any differential operator, we denote its (formal) L^2 adjoint by $\mathscr{D}^*$; in particular, the adjoint of the Fréchet derivative of a differential function plays an important role and has the explicit formula

$$D_P^* = \sum (-D_x)^k \cdot \frac{\partial P}{\partial u_k},$$

where the $\cdot$ means one first multiplies by $\partial P/\partial u_k$ and then differentiates. Note in particular that if $\psi(x)$ depends only on x, then

$$D_P^*(\psi) = E_u(\psi \cdot P). \tag{13}$$

With these definitions, the key operator associated with the change of variables (11) is

$$\mathbf{D} = D_xP \cdot D_Q - D_xQ \cdot D_P = \det\begin{pmatrix} D_xP & D_P \\ D_xQ & D_Q \end{pmatrix}. \tag{14}$$

Its adjoint is given by

$$\mathbf{D}^* = D_Q^* \cdot D_xP - D_P^* \cdot D_xQ.$$

LEMMA 5. *Let y, w be related to x, u by the differential substitution* (11), *with y the new independent and w the new dependent variable. Then*

$$w_t = (D_x P)^{-1} \cdot \mathbf{D}(u_t), \tag{15}$$

and the Euler operators are related by

$$E_u = \mathbf{D}^* \cdot E_w. \tag{16}$$

Proof. Formula (15) is proved in the same manner as the change of variables formulas from vector calculus. Using differential form notation, we have, by the definition of the Fréchet derivative,

$$dy = D_x P \cdot dx + D_P(u_t) \cdot dt, \qquad dw = D_x Q \cdot dx + D_Q(u_t) \cdot dt,$$

hence, solving the first for dx,

$$\begin{aligned} dw &= (D_x P)^{-1} \{ D_x Q \cdot dy + (D_x P \cdot D_Q(u_t) - D_x Q \cdot D_P(u_t)) \cdot dt \} \\ &= (D_x P)^{-1} \{ D_x Q \cdot dy + \mathbf{D}(u_t) \cdot dt \}, \end{aligned}$$

which simultaneously proves the two change of variables formulas (15) and (12) for the derivatives of w.

The formula (16) is a special case of a general formula for the behavior of the Euler operator under a change of variables given in [15; Exercise 5.39]. (N.B. there is a misprint in this exercise as it appears in the book: the change of variables should read $y = P(x, u^{(m)})$, $w = Q(x, u^{(m)})$, and in the subsequent two formulas w^α should be u^α, while u^β should be w^β!)

THEOREM 6. *Let $\mathscr{D}$ be a Hamiltonian operator depending on x and u, and let $y = P[u]$ and $w = Q[u]$ be related to x and u by a differential substitution. Then the corresponding Hamiltonian operator in the y, w-variables is*

$$\tilde{\mathscr{D}} = (D_x P)^{-1} \cdot \mathbf{D} \cdot \mathscr{D} \cdot \mathbf{D}^*, \tag{17}$$

where (12) *is used to related the total x and y derivatives.*

Proof. In other words, if

$$u_t = \mathscr{D} \cdot E_u(H)$$

is a Hamiltonian system in the x, u variables, then the corresponding evolution equation in the y, w variables will also be Hamiltonian

$$w_t = \tilde{\mathscr{D}} \cdot E_w(H), \tag{18}$$

with the Hamiltonian operator given by the formula (17). To prove this, we just compute using (15) and (16):

$$\begin{aligned} w_t &= (D_x P)^{-1} \cdot \mathbf{D}(u_t) \\ &= (D_x P)^{-1} \cdot \mathbf{D} \cdot \mathscr{D} \cdot E_u(H) \\ &= (D_x P)^{-1} \cdot \mathbf{D} \cdot \mathscr{D} \cdot \mathbf{D}^* \cdot E_w(H). \end{aligned}$$

Comparison with (18) completes the proof of (17).

Note in particular, if the differential substitution does not change the independent variable, so we just have $w = Q[u]$, then (17) reduces to the formula

$$\tilde{\mathscr{D}} = D_Q \cdot \mathscr{D} \cdot D_Q^*. \tag{19}$$

EXAMPLE. Consider the first order Hamiltonian operator

$$\mathscr{D} = (u_{xx})^{-1} \cdot D_x \cdot (u_{xx})^{-1} = (u_{xx})^{-3}\{u_{xx} D_x - u_{xxx}\}.$$

An example of an evolution equation which is Hamiltonian with respect to $\mathscr{D}$ is the potential form of the nonlinear wave equation $v_t = v v_x$:

$$u_t = u_x^2 = \mathscr{D} E_u(H), \qquad \text{where} \quad \mathscr{H} = \int H\, dx = \int -\frac{1}{60} u_x^5\, dx \tag{20}$$

is the Hamiltonian functional, and $v = u_x$. Suppose we are to use the change of variables

$$y = u_x, \qquad w = x u_x - u. \tag{21}$$

Then the operator $\mathbf{D}$ in (14) is

$$\mathbf{D} = u_{xx} \cdot (x D_x - 1) - x u_{xx} \cdot D_x = -u_{xx},$$

which is just a multiplication operator; in particular $\mathbf{D}$ is self-adjoint in this particular instance: $\mathbf{D}^* = \mathbf{D}$. Also, by (12),

$$D_y = (u_{xx})^{-1} \cdot D_x.$$

Therefore by formula (17), in the y, w variables, $\mathscr{D}$ takes on the simple form

$$\tilde{\mathscr{D}} = (u_{xx})^{-1} \cdot (-u_{xx}) \cdot \{(u_{xx})^{-1} \cdot D_x \cdot (u_{xx})^{-1}\} \cdot (-u_{xx}) = D_y.$$

Thus (21) is precisely the change of variables required to bring $\mathscr{D}$ into constant coefficient form! Indeed, note that by (15) $w_t = -u_t$, hence in the y, w variables, the wave equation (20) becomes

$$w_t = -y^2 = \tilde{\mathscr{D}} E_w(\tilde{H}), \qquad \text{since} \quad \mathscr{H} = \int \tilde{H}\, dy = \int -\frac{1}{60} y^5 w_{yy}\, dy$$

in the new variables. (To see this, note that $dx = (u_{xx})^{-1}\, dy$, and $w_y = x$, hence $w_{yy} = (u_{xx})^{-1}$, and $dx = w_{yy}\, dy$.)

The change of variables (21) looks a bit strange. Let us see how it can be built up from more familiar changes of variable. First of all, if we replace the potential u by $v = u_x$, then $\mathscr{D}$ becomes the third order operator

$$\mathscr{D}' = D_x \cdot (v_x)^{-1} \cdot D_x \cdot (v_x)^{-1} \cdot D_x$$

we encountered in Section 4. Second, we perform the "hodograph" transformation $y = v$, $z = x$, which interchanges the roles of independent and dependent variables, to change $\mathscr{D}'$ into

$$\mathscr{D}'' = D_y^3,$$

which is a constant coefficient operator, but of third order. Finally, the potential substitution $z = w_y$ changes $\mathscr{D}''$ into $\tilde{\mathscr{D}}$. The reader can check that the composition of these three changes of variable is the same as the original formula (21).

The reader may wonder whether it is possible to change this Hamiltonian operator to constant coefficient form by a restricted differential substitution in which the independent variables are not transformed, such as in (19). In fact this is *not* possible, but the proof is a bit technical and is relegated to the Appendix to this paper. Therefore, the Darboux Theorem for Hamiltonian differential operators is not true unless we admit the full range of differential substitutions (11).

7. Darboux' Theorem for First Order Operators

Armed with the change of variables formula (17), and the explicit characterization of first order Hamiltonian operators, we are in a position to prove Darboux' Theorem in this special case.

THEOREM 7. *Let $\mathscr{D}$ be a first order Hamiltonian differential operator. Then there exists a local change of variables $y = P[u]$, $w = Q[u]$ such that in the y, w variables $\mathscr{D}$ becomes the constant coefficient differential operator $\tilde{\mathscr{D}} = D_y$.*

Proof. It turns out that it suffices to consider a first order differential substitution

$$y = P(x, u, u_x), \qquad w = Q(x, u, u_x).$$

In this case, the operator $\mathbf{D}$ in (14) is at most a first order differential operator, with the explicit formula

$$\mathbf{D} = M \cdot D_x + N,$$

where

$$M = (P_x Q_{u_x} - P_{u_x} Q_x) + u_x(P_u Q_{u_x} - P_{u_x} Q_u),$$
$$N = (P_x Q_u - P_u Q_x) + u_{xx}(P_{u_x} Q_u - P_u Q_{u_x}).$$

Now, when $\mathscr{D}$ is given by (4), the only way for (17) to be a first order differential operator is for $\mathbf{D}$ to be a zeroth order differential operator, which means that M must vanish. In this case, (17) reads

$$\begin{aligned}\tilde{\mathscr{D}} &= (D_x P)^{-1} \cdot \mathbf{D} \cdot \mathscr{D} \cdot \mathbf{D}^* \\ &= (D_x P)^{-1} N \cdot (E(A))^{-1} \cdot D_x \cdot (E(A))^{-1} \cdot N \\ &= N \cdot (E(A))^{-1} \cdot D_y \cdot (E(A))^{-1} \cdot N.\end{aligned}$$

Therefore $\tilde{\mathscr{D}} = D_y$ if and only if $N = E(A) = F \cdot u_{xx} + G$, where F and G are as given in (9). Combining this with the vanishing of M, we are left with three equations relating P, Q, and A, which are easily seen to be equivalent to the equations

$$\begin{aligned} P_x Q_u - P_u Q_x &= G, \\ P_x Q_{u_x} - P_{u_x} Q_x &= u_x \cdot F, \\ P_{u_x} Q_u - P_u Q_{u_x} &= F. \end{aligned}$$

These in turn are equivalent to the two-form equation

$$dP \wedge dQ = G \cdot dx \wedge du + u_x F \cdot dx \wedge du_x + F \cdot du_x \wedge du \equiv \omega \tag{22}$$

involving the three "independent variables" x, u, u_x. It is easy to check that the two-form ω involving F and G is closed under the usual exterior derivative operator d if and only if F and G satisfy the key equation (10) that was required for $\mathscr{D}$ to be a genuine Hamiltonian operator! (Equivalently, in terms of the differential function A in (4),

$$\omega = d(A - u_x A_{u_x}) \wedge dx + d(A_{u_x}) \wedge du,$$

which is clearly closed.) Thus, by Darboux' Theorem for differential two-forms in $\mathbb{R}^3$ (cf. [18; Theorem 6.1]), unless both F and G vanish, which is not possible, there is a local change of coordinates $\xi = P(x, u, u_x)$, $\eta = Q(x, u, u_x)$, $\zeta = R(x, u, u_x)$, such that $\omega = d\xi \wedge d\eta$ ($= dP \wedge dQ$) in the ξ, η, ζ-coordinates. But this is precisely what we need to satisfy the condition (22)! Therefore, the change of variable needed to place a first order Hamiltonian operator in constant coefficient (Darboux) form D_y is determined by the change of coordinates needed to place the two-form ω into canonical (Darboux) form.

EXAMPLE. Consider the Hamiltonian operator $\mathscr{D} = (u_{xx})^{-1} \cdot D_x \cdot (u_{xx})^{-1}$ discussed earlier. In this case $F = 1$, $G = 0$, so the two-form ω in (22) is

$$\begin{aligned}\omega &= u_x dx \wedge du_x + du_x \wedge du \\ &= d(xu_x - u) \wedge du_x.\end{aligned}$$

Thus we recover the change of variables determined by $P = xu_x - u$, $Q = u_x$ deduced earlier. However, this is clearly not the only change of variables that satisfies (22), since we can, for instance, compose P, Q with any area-preserving (canonical) diffeomorphism of $\mathbb{R}^2$. Thus, for instance,

$$y = \tilde{P} = u_x(u_x - 1)^{-1}, \qquad w = \tilde{Q} = (xu_x - u)(u_x - 1)^2$$

is also a valid differential substitution which changes $\mathscr{D}$ into constant coefficient form. More complicated examples can, of course, be constructed, but the basic method should be clear.

APPENDIX

Here we prove that the first order Hamiltonian operator

$$\mathscr{D} = (u_{xx})^{-1} \cdot D_x \cdot (u_{xx})^{-1} = (u_{xx})^{-3}\{u_{xx} D_x - u_{xxx}\}$$

cannot be changed into a constant coefficient differential operator by a change of variables involving only the dependent variables. In other words, we consider differential substitutions either of the form

$$w = Q[u], \tag{A1}$$

or, alternatively, of the form

$$u = R[w], \tag{A2}$$

and where x is unchanged.

THEOREM. *Let* $\mathscr{D}=(u_{xx})^{-1}\cdot D_x\cdot(u_{xx})^{-1}$ *and let either* $w=Q[u]$ *or* $u=R[w]$ *be a differential substitution which does not change the independent variable* x. *Let* $\tilde{\mathscr{D}}$ *be the corresponding Hamiltonian operator in the* x, w*-coordinates. Then* $\tilde{\mathscr{D}}$ *must depend on* w *or its derivatives.*

Proof. To prove this, we assume the contrary, i.e., that

$$\tilde{\mathscr{D}}=\sum_{j=0}^{m} a_j(x)\cdot D_x^j, \tag{A3}$$

where the coefficients a_j depend only on x, and m denotes the order of $\tilde{\mathscr{D}}$.

Case (A2) is easy to dismiss. According to (19) $\mathscr{D}$ and $\tilde{\mathscr{D}}$ are related by the change of variables formula

$$\mathscr{D}=D_R\cdot\tilde{\mathscr{D}}\cdot D_R^*, \tag{A4}$$

where D_R is the Fréchet derivative of $R[w]$ with respect to w. Now $\mathscr{D}$ is a first order differential operator, so the only way that (A4) can hold is if $\tilde{\mathscr{D}}$ is also first order, and $R=R(x,w)$ does not depend on any derivatives of w (as otherwise D_R would be a differential operator of order at least one). In this case, $u_{xx}=D_x^2R$, so (A4) would read

$$(D_x^2R)^{-1}\cdot D_x\cdot(D_x^2R)^{-1}=(\partial R/\partial w)\cdot\tilde{\mathscr{D}}\cdot(\partial R/\partial w). \tag{A5}$$

The left hand side of (A5) depends on w_{xx}, but $\partial R/\partial w$ depends only on x and w, hence $\tilde{\mathscr{D}}$ must depend explicitly on w_{xx}, and cannot be of the form (A3). Thus $\mathscr{D}$ cannot be reduced to a differential operator depending only on x (in particular a constant coefficient operator) by a change of variables of the form (A2).

Now consider a change of variables of the form (A1). Again by (19) $\mathscr{D}$ and $\tilde{\mathscr{D}}$ are related by

$$\tilde{\mathscr{D}}=D_Q\cdot\mathscr{D}\cdot D_Q^*, \tag{A6}$$

where

$$D_Q=\sum_{j=0}^{n}\frac{\partial Q}{\partial u_j}\cdot D_x^j \tag{A7}$$

is the Fréchet derivative of $Q[u]$ with respect to u,

$$D_Q^*=\sum_{j=0}^{n}(-D_x)^j\cdot\frac{\partial Q}{\partial u_j} \tag{A8}$$

is its adjoint, and $u_j\equiv\partial^j u/\partial x^j$. We assume that u_n is the highest order derivative of u appearing in Q, hence $Q[u]=Q(x,u,u_x,\ldots,u_n)$, and $\partial Q/\partial u_n\neq 0$ does not vanish identically.

LEMMA A1. *Under the above assumptions,* $\tilde{\mathscr{D}}$ *is a* $(2n+1)^{st}$ *order differential operator, and*

$$\partial Q/\partial u_n = \alpha(x)\, u_{xx}, \tag{A9}$$

where $\alpha(x)^2 = (-1)^n\, a_{2n+1}(x)$ (*cf.* (A3)) *does not vanish identically.*

Proof. Substituting (A7) and (A8) into (A6), we find that the highest order term is

$$D_Q \cdot \mathscr{D} \cdot D_Q^* = (-1)^n\, u_{xx}^{-2} \left(\frac{\partial Q}{\partial u_n}\right)^2 D_x^{2n+1} + \cdots.$$

Equating this to the desired form (A3), we see that $m = 2n+1$, and hence (A9) holds.

LEMMA A2. *For* $Q[u]$ *as above,*

$$\ker D_Q \subset \{ax + b\colon a, b \in \mathbb{R}\}.$$

Proof. Suppose $S[u] = S(x, u, ..., u_m)$, depending on derivatives of u up to order m, lies in the kernel of D_Q, so

$$0 = D_Q(S) = \sum_{j=0}^{n} \frac{\partial Q}{\partial u_j} \cdot D_x^j\, S. \tag{A10}$$

If $m > 0$, then there is only one term in this equation involving u_{n+m}, which is

$$u_{n+m} \frac{\partial Q}{\partial u_n} \frac{\partial S}{\partial u_m}.$$

Since $\partial Q/\partial u_n$ does not vanish identically, the only way that (A10) could hold is if $\partial S/\partial u_m = 0$ everywhere, and we conclude that $S = S(x, u)$ depends only on x and u. Next, apply the operator $\partial/\partial u_n$ to (A10). Note that by (A9)

$$\frac{\partial^2 Q}{\partial u_n\, \partial u_j} = 0$$

for $j \neq 2$, hence the only terms which survive are

$$0 = \frac{\partial Q}{\partial u_n} \frac{\partial S}{\partial u} + \frac{\partial^2 Q}{\partial u_n \partial u_2} \cdot D_x^2 S$$
$$= \alpha(x) \left\{ u_{xx} \frac{\partial S}{\partial u} + D_x^2 S \right\}.$$

Since $\alpha(x) \neq 0$, it is easily seen that this requires S to be independent of u and at most linear in x: $S = ax + b$ for $a, b \in \mathbb{R}$.

Now, as $\tilde{\mathscr{D}}$ is a $(2n+1)^{\text{st}}$ order differential operator that depends only on x, it possesses a fundamental set of solutions $\psi_1(x), ..., \psi_{2n+1}(x)$ to the $(2n+1)^{\text{st}}$ order ordinary differential equation

$$\tilde{\mathscr{D}}(\psi) = 0. \tag{A11}$$

LEMMA A3. *For each* $1 \leqslant j \leqslant 2n+1$ *there exist constants* $a_j, b_j, c_j \in \mathbb{R}$, *and a differential function* $R_j[u]$ *such that*

$$\psi_j(x) \cdot Q[u] = D_x R_j + a_j \cdot (\tfrac{1}{3} x u_x^3 - u u_x^2) + b_j \cdot (\tfrac{1}{3} u_x^3) + c_j \cdot (\tfrac{1}{2} u_x^2). \tag{A12}$$

Proof. Let $\psi(x)$ be any solution to the ordinary differential equation (A11). Then, by (A6) and (13),

$$0 = \tilde{\mathscr{D}}(\psi) = D_Q \cdot \mathscr{D} \cdot D_Q^*(\psi) = D_Q \cdot \mathscr{D} \cdot E_u(\psi \cdot Q).$$

Therefore, $\mathscr{D} \cdot E_u(\psi \cdot Q)$ lies in the kernel of D_Q, and hence according to Lemma A2,

$$\mathscr{D} \cdot E_u(\psi \cdot Q) = ax + b$$

for constants $a, b \in \mathbb{R}$. Recalling the original definition of $\mathscr{D}$, we see that this is equivalent to the condition

$$D_x(u_{xx}^{-1} \cdot E_u(\psi \cdot Q)) = (ax+b)\, u_{xx},$$

or, equivalently,

$$E_u(\psi \cdot Q) = (ax+b)\, u_x u_{xx} - a u u_{xx} + c u_{xx} \tag{A13}$$

for some $c \in \mathbb{R}$. The right hand side of (A13) lies in the image of E_u, and hence we have the equivalent condition

$$E_u\{\psi \cdot Q - a \cdot (\tfrac{1}{3} x u_x^3 - u u_x^2) - b \cdot (\tfrac{1}{3} u_x^3) - c \cdot (\tfrac{1}{2} u_x^2)\} = 0.$$

The lemma now follows immediately from the characterization of the kernel of the Euler operator as the image of the total derivative D_x (cf. [15; Theorem 4.7]).

The goal now is to prove that the conditions (A12) are incompatible except when Q is independent of u, which is not an allowed change of variables (A1). To this end, we introduce ordinary differential operators Δ_k and functions $\varphi_k(x)$, $k = 1, ..., 2n+1$, by the recursive definition

$$\varphi_1 = 1/\psi_1, \qquad \Delta_1 = \varphi_1 \cdot D_x,$$

and

$$\varphi_{k+1} = 1/[\Delta_k^*(\psi_{k+1})], \qquad \Delta_{k+1} = \Delta_k \cdot \varphi_{k+1} \cdot D_x, \qquad 1 \leqslant k \leqslant 2n. \tag{A14}$$

Note in particular that

$$\Delta_k = \varphi_1 \cdot D_x \cdot \varphi_2 \cdot D_x \cdot \varphi_3 \cdot \cdots \cdot D_x \cdot \varphi_k \cdot D_x = \varphi_1 \varphi_2 \cdots \varphi_k \cdot D_x^k + \cdots. \tag{A15}$$

One important point that needs to be proved is that although the functions $\varphi_k(x)$ may have singularities, they are isolated; in particular there is an open interval $I \subset \mathbb{R}$ such that for every $1 \leqslant k \leqslant 2n+1$, $\varphi_k(x)$ is a smooth function for $x \in I$. In fact, suppose we let $\mathcal{R}$ denote the ring of smooth functions $f: \mathbb{R} \to \mathbb{R}$ which are either identically 0 or do not vanish on any open interval; note that the solutions $\psi_j(x)$ of (A11) belong to $\mathcal{R}$. Since $\mathcal{R}$ has no zero divisors, we can let $\mathcal{Q}$ denote the quotient field of $\mathcal{R}$, so that elements of $\mathcal{Q}$ are quotients of smooth functions $f(x)/g(x)$ where $f, g \in \mathcal{R}$, $g \neq 0$. Then we can prove the following:

LEMMA A4. *Let Δ_k and $\varphi_k(x)$, $k = 1, ..., 2n+1$, be defined by* (A14), *where $\psi_j(x)$, $j = 1, ..., 2n+1$, form a fundamental set of solutions to the $(2n+1)^{st}$ order ordinary differential equation* (A11). *Then $\varphi_k \in \mathcal{Q}$, $k = 1, ..., 2n+1$.*

Proof. It suffices to show that $(\Delta_k^*(\psi_{k+1})$ does not vanish identically, as both the derivative and the reciprocal of a function in $\mathcal{Q}$ remain in $\mathcal{Q}$. This will follow from the more basic lemma:

LEMMA A5. *If*

$$\Delta_k^*(\psi) = 0$$

vanishes identically, then there exists a constant $c \in \mathbb{R}$ such that

$$\Delta_{k-1}^*(\psi - c\psi_k) = 0$$

vanishes identically.

Proof. Taking the adjoint of the definition (A14) of Δ_k we have

$$0 = \Delta_k^*(\psi) = -D_x[\varphi_k \cdot \Delta_{k-1}^*(\psi)].$$

Therefore

$$\Delta_{k-1}^*(\psi) = c/\varphi_k = c\Delta_{k-1}^*(\psi_k)$$

for some $c \in \mathbb{R}$, which proves Lemma A5.

Returning to the proof of Lemma A4, suppose

$$\Delta_k^*(\psi_{k+1}) = 0$$

vanishes identically. Then by Lemma A5,

$$\Delta_{k-1}^*(\psi_{k+1} - c_k\psi_k) = 0$$

for some constant c_k. We can again apply Lemma A5, with k replaced by $k-1$, to get

$$\Delta_{k-2}^*(\psi_{k+1} - c_k\psi_k - c_{k-1}\psi_{k-1}) = 0$$

for some constant c_{k-1}. Proceeding by reverse induction on k, we are finally led to a relationship of the form

$$\psi_{k+1} - c_k\psi_k - c_{k-1}\psi_{k-1} - \cdots - c_1\psi_1 = 0,$$

which must hold identically in x. But this means $\psi_1, ..., \psi_{k+1}$ are linearly dependent, which contradicts the fact that they form a fundamental set of solutions to (A11). This completes the proof of Lemma A4.

Therefore, for each $k = 1, ..., 2n+1$, Δ_k is a kth order ordinary differential operator whose coefficients are smooth functions of x on an open interval $I \subset \mathbb{R}$. Define the differential polynomials

$$A = \tfrac{1}{3} x u_x^3 - u u_x^2, \qquad B = \tfrac{1}{3} u_x^3, \qquad C = \tfrac{1}{2} u_x^2, \tag{A16}$$

so that (A12) can be written as

$$\psi_j(x) \cdot Q[u] = D_x R_j + a_j A + b_j B + c_j C.$$

LEMMA A6. *For each $k = 1, ..., 2n+1$, there exist functions $f_i^k(x)$, $g_i^k(x)$, $h_i^k(x)$, $i = 1, ..., k-1$, and differential functions $S_k[u]$, all smoothly defined for $x \in I$, such that*

$$Q[u] = \Delta_k(S_k) + \sum_{i=0}^{k-1} \{f_i^k(x) \cdot D_x^i A + g_i^k(x) \cdot D_x^i B + h_i^k(x) \cdot D_x^i C\}. \tag{A17}$$

Proof. For $k = 1$ this reduces to (A12) when $j = 1$. Proceeding by induction on k, if (A17) holds for k, then

$$\begin{aligned}\psi_{k+1} \cdot Q[u] &= \psi_{k+1} \cdot \Delta_k(S_k) \\ &\quad + \sum_{i=0}^{k-1} \psi_{k+1} \cdot \{f_i^k(x) \cdot D_x^i A + g_i^k(x) \cdot D_x^i B + h_i^k(x) \cdot D_x^i C\}.\end{aligned}$$

We integrate each of the terms on the right hand side by parts, so

$$\begin{aligned}\psi_{k+1} \cdot Q[u] &= \Delta_k^*(\psi_{k+1}) \cdot S_k + f_*^k(x) \cdot A + g_*^k(x) \cdot B + h_*^k(x) \cdot C + D_x T_k \\ &= (1/\varphi_{k+1}) \cdot S_k + f_*^k(x) \cdot A + g_*^k(x) \cdot B + h_*^k(x) \cdot C + D_x T_k,\end{aligned} \tag{A18}$$

where

$$f_*^k(x) = \sum_{i=0}^{k-1} (-D_x)^i f_i^k(x), \qquad g_*^k(x) = \sum_{i=0}^{k-1} (-D_x)^i g_i^k(x),$$

$$h_*^k(x) = \sum_{i=0}^{k-1} (-D_x)^i h_i^k(x),$$

and $T_k[u]$ is some differential function whose precise expression is not required here.

Substituting (A18) into (A12), for $j = k+1$, we find that

$$S_k = \varphi_{k+1} \cdot (D_x(S_{k+1}) + \tilde{f}^k(x) \cdot A + \tilde{g}^k(x) \cdot B + \tilde{h}^k(x) \cdot C), \qquad \text{(A19)}$$

where

$$S_{k+1} = R_{k+1} - T_k,$$

and

$$\tilde{f}^k(x) = a_{k+1} - f_*^k(x), \qquad \tilde{g}^k(x) = b_{k+1} - g_*^k(x), \qquad \tilde{h}^k(x) = c_{k+1} - h_*^k(x).$$

Substituting (A19) into (A17) and recalling the definition of Δ_{k+1} completes the induction step from k to $k+1$.

We are now in a position to derive a contradiction and complete the proof of the theorem. Consider the case $k = n$ in (A17)

$$Q[u] = \Delta_n(S) + \sum_{i=0}^{n-1} \{f_i(x) \cdot D_x^i A + g_i(x) \cdot D_x^i B + h_i(x) \cdot D_x^i C\}, \qquad \text{(A20)}$$

where, for simplicity, we have denoted S_n by S, f_i^n by f_i, etc. Suppose S depends on at most mth order derivatives of u, so $S = S(x, u, ..., u_m)$. If $m \geqslant 1$, then apply the operator $\partial/\partial u_{n+m+1}$ to (A20); using (A15) we obtain

$$0 = \varphi_1 \cdot \cdots \cdot \varphi_n (\partial S/\partial u_m).$$

Since the $\varphi_j(x)$ have only isolated zeros and singularities, we deduce that $S = S(x, u)$ can depend on at most x and u. Finally, applying $\partial/\partial u_n$ to (A20), and using (A9) and the definition (A16) of A, B, and C, we find

$$\alpha(x)\, u_{xx} = \frac{\partial Q}{\partial u_n} = \varphi_1 \cdot \cdots \cdot \varphi_n \cdot \frac{\partial S}{\partial u} + f_n \cdot x u_x^2 + g_n \cdot u_x^2 + h_n \cdot u_x.$$

But this is clearly impossible unless both sides are identically zero, hence $\partial Q/\partial u_n = 0$ identically. This contradicts our initial assumption and completes the proof.

ACKNOWLEDGMENTS

It is a great preasure to thank Yavuz Nutku and John Verosky for valuable discussions and inspiration, without which this paper would not have been written.

Note added in proof. A recent paper of A. M. Astashov and A. M. Vinogradov, On the structure of Hamiltonian operators in field theory, *J. Geom. and Physics* **2** (1986), 263–287, gives another proof of Theorem 7, along with some extensions to higher order and higher dimensional operators.

REFERENCES

1. P. A. M. DIRAC, "Lectures on Quantum Mechanics," Belfer Graduate School Monograph Series No. 3, Yeshiva University, New York, 1964.
2. B. A. DUBROVIN, AND S. A. NOVIKOV, Hamiltonian formalism of one-dimensional systems of hydrodynamic type and the Bogolyubov–Whitham averaging method, *Soviet Math. Dokl.* **27** (1983), 665–669.
3. B. A. DUBROVIN, AND S. A. NOVIKOV, On Poisson brackets of hydrodynamic type, *Soviet Math. Dokl.* **30** (1984), 651–654.
4. I. M. GEL'FAND AND I. YA. DORFMAN, Hamiltonian operators and algebraic structures related to them, *Functional Anal. Appl.* **13** (1979), 248–262.
5. D. D. HOLM, J. E. MARSDEN, T. RATIU, AND A. WEINSTEIN, Nonlinear stability of fluid and plasma equilibria, *Phys. Rep.* **123** (1985), 1–116.
6. N. H. IBRAGIMOV, "Transformation Groups Applied to Mathematical Physics," Reidel, Boston, 1985.
7. Y. KODAMA, On integrable systems with higher order corrections, *Phys. Lett. A* **107** (1985), 245–249.
8. R. LITTLEJOHN, Hamiltonian perturbation theory in noncanonical coordinates, *J. Math. Phys.* **23** (1982), 742–747.
9. F. MAGRI, A simple model of the integrable Hamiltonian equation, *J. Math. Phys.* **19** (1978), 1156–1162.
10. J. E. MARSDEN, Fluids and plasmas: Geometry and dynamics, *Contemp. Math.* **28** (1984), 231–249.
11. Y. NUTKU, Hamiltonian formulation of the KdV equation, *J. Math. Phys.* **25** (1984), 2007–2008.
12. P. J. OLVER, Symmetry groups and group-invariant solutions of partial differential equations, *J. Diff. Geom.* **14** (1979), 497–542.
13. P. J. OLVER, On the Hamiltonian structure of evolution equations, *Math. Proc. Cambridge Philos. Soc.* **88** (1980), 71–88.
14. P. J. OLVER, Hamiltonian perturbation theory and water waves, *Contemp. Math.* **28** (1984), 231–249.
15. P. J. OLVER, "Applications of Lie Groups to Differential Equations," Graduate Texts in Mathematics, Vol. 107, Springer–Verlag, New York/Berlin, 1986.
16. P. J. OLVER, Dirac's theory of constraints for Hamiltonian differential operators, *J. Math. Phys.* **27** (1986), 2495–2501.
17. P. J. OLVER, BiHamiltonian systems, 9th Dundee Conference on the Theory of Ordinary and Partial Differential Equations, Proceedings, to appear.
18. S. STERNBERG, "Lectures on Differential Geometry," Prentice–Hall, Englewood Cliffs, NJ, 1964.
19. A. WEINSTEIN, Symplectic manifolds and their Lagrangian submanifolds, *Adv. Math.* **6** (1971), 329–346.
20. N. WOODHOUSE, "Geometric Quantization," Oxford Univ. Press, New York, 1980.