Math. Proc. Camb. Phil. Soc. (1983), **94**, 529
Printed in Great Britain

Conservation laws and null divergences

By PETER J. OLVER
School of Mathematics, University of Minnesota, Minneapolis, Minnesota 55455

(*Received* 14 *October* 1982; *revised* 8 *March* 1983)

1. *Conservation laws*

For a system of partial differential equations, the existence of appropriate conservation laws is often a key ingredient in the investigation of its solutions and their properties. Conservation laws can be used in proving existence of solutions, decay and scattering properties, investigation of singularities, analysis of integrability properties of the system and so on. Representative applications, and more complete bibliographies on conservation laws, can be found in references [7], [8], [**12**], [**19**]. The more conservation laws known for a given system, the more tools available for the above investigations. Thus a complete classification of all conservation laws of a given system is of great interest. Not many physical systems have been subjected to such a complete analysis, but two examples can be found in [**11**] and [**14**]. The present paper arose from investigations ([**15**], [**16**]) into the conservation laws of the equations of elasticity.

We begin by recalling the definition of a conservation law. Let $x = (x^1, \ldots, x^p)$ be the independent and $u = (u^1, \ldots, u^q)$ the dependent variables in the system. The notation $\partial^m u$ is an abbreviation for the collection of all mth order partial derivatives of the u's with respect to the x's, for which we use multi-index notation

$$u^\nu_J = \partial^m u^\nu / \partial x^{j_1} \ldots \partial x^{j_m}, \quad J = (j_1, \ldots, j_m),$$

for $1 \leqslant \nu \leqslant q$, $1 \leqslant j_\kappa \leqslant p$. For $1 \leqslant i \leqslant p$, the operator

$$D_i = \frac{\partial}{\partial x^i} + \sum_{\nu, J} u^\nu_{J,i} \frac{\partial}{\partial u^\nu_J}, \tag{1·1}$$

where

$$u^\nu_{J,i} = \partial u^\nu_J / \partial x^i,$$

when applied to functions of $x, u, \partial u, \ldots, \partial^n u$, is the *total derivative* with respect to x^i. It is defined so that, given $P(x, u, \ldots, \partial^n u)$,

$$(D_i P)(x, f(x), \ldots, \partial^{n+1} f(x)) = \frac{\partial}{\partial x^i}[P(x, f(x), \ldots, \partial^n f(x))],$$

for any smooth function $u = f(x)$. For example,

$$D_i\left[u \frac{\partial u}{\partial x^j}\right] = \frac{\partial u}{\partial x^i}\frac{\partial u}{\partial x^j} + u \frac{\partial^2 u}{\partial x^i \partial x^j}.$$

Higher order total derivatives are written using multi-index notation:

$$D^J = D_{j_1} D_{j_2} \ldots D_{j_k}, \quad J = (j_1, \ldots, j_k).$$

Consider a system of nth order differential equations

$$\Delta_\kappa(x, u, \partial u, \ldots, \partial^n u) = 0 \quad (\kappa = 1, \ldots, l), \tag{1·2}$$

where the Δ_κ are smooth functions of their arguments. A *conservation law* for this system is a divergence expression

$$\operatorname{Div} P = \sum_{i=1}^{p} D_i P_i = 0, \tag{1·3}$$

which vanishes for all solutions $u = f(x)$ of (1·2). The p-tuple $P = (P_1, \ldots, P_p)$ are functions of $x, u, \partial u, \ldots, \partial^k u$ for some k, and consist of the conserved density and corresponding fluxes. Under mild *nondegeneracy* assumptions on (1·2) (see the appendix), the fact that (1·3) vanishes on solutions of (1·2) can be replaced with the more explicit condition

$$\operatorname{Div} P = Q = \sum_{J,\kappa} Q_{J,\kappa} D^J \Delta_\kappa, \tag{1·4}$$

holding for all values of $x, u, \partial u, \ldots, \partial^m u$, for nonvanishing functions $Q_{J,\kappa}(x, u, \ldots, \partial^m u)$ to be determined. Thus the classification problem for conservation laws amounts to determining all functions P_i, $Q_{J,\kappa}$ depending on x, u and derivatives of u for which the identity (1·4) holds.

Example 1·1. For the wave equation

$$u_{tt} - u_{xx} = 0,$$

the divergence expression

$$D_t(\tfrac{1}{2}u_t^2 + \tfrac{1}{2}u_x^2) + D_x(-u_x u_t) = 0,$$

is a conservation law, since

$$D_t(\tfrac{1}{2}u_t^2 + \tfrac{1}{2}u_x^2) + D_x(-u_x u_t) = u_t(u_{tt} - u_{xx}).$$

The first component $\frac{1}{2}u_x^2 + \frac{1}{2}u_t^2$ is the conserved density, and, for solutions $u \to 0$ sufficiently rapidly as $|x| \to \infty$, we deduce the conservation of energy:

$$\int_{-\infty}^{\infty} \tfrac{1}{2}u_t^2 + \tfrac{1}{2}u_x^2 \, dx = \text{constant}.$$

In the general search for conservation laws, one usually begins by bounding the order m of partial derivatives on which the $Q_{J,\kappa}$ in (1·4) can depend. It is then extremely helpful to know an a priori bound on the order k of derivatives on which P can depend. Proving such a result is one of the main goals of this paper.

2. *Null divergences*

One must always distinguish between trivial and non-trivial conservation laws. There are two types of triviality which automatically imply the vanishing of (1·3). The first, which will not concern us here, is when the p-tuple P itself vanishes for all solutions of the system (1·2); for Δ nondegenerate, this is equivalent to

$$P_i = \sum P^i_{J,\kappa} D^J \Delta_\kappa \quad (i = 1, \ldots, p),$$

for certain $P^i_{J,\kappa}$.

The second type of triviality is when $\operatorname{Div} P$ vanishes identically, no matter what system of equations $u = f(x)$ satisfies. These P will be called null divergences:

Definition 2·1. A *null divergence* is a p-tuple of functions $P(x, u, \ldots, \partial^k u)$ such that

$$\operatorname{Div} P \equiv 0, \tag{2·1}$$

for all $x, u, \ldots, \partial^{k+1}u$. (It should be required that $u \in C^{k+1}$, but, as in [2], it really suffices to assume (2·1) holds distributionally.)

Example 2·2. A trivial linear null divergence is $P = (u_y, -u_x)$ $(x = x^1, y = x^2)$, since

$$\operatorname{Div} P = D_x u_y + D_y(-u_x) \equiv 0,$$

for $u \in C^2$.

Example 2·3. Less trivial is the three dimensional quadratic example $(x = x^1, y = x^2, z = x^3)$,

$$P = \left(\frac{\partial(u,v)}{\partial(y,z)}, \frac{\partial(u,v)}{\partial(z,x)}, \frac{\partial(u,v)}{\partial(x,y)}\right),$$

the entries being Jacobian determinants, e.g. $\partial(u,v)/\partial(x,y) = u_x v_y - u_y v_x$. It is easy to check that

$$\operatorname{Div} P = D_x \frac{\partial(u,v)}{\partial(y,z)} + D_y \frac{\partial(u,v)}{\partial(z,x)} + D_z \frac{\partial(u,v)}{\partial(x,y)} \equiv 0. \qquad (2\cdot2)$$

In Section 3, we describe higher order versions of this null divergence, and prove that these are essentially the only null divergences depending exclusively on first order derivatives of u.

The term 'null divergence' is in analogy with the concept of a *null Lagrangian*, investigated in detail in [2]. By definition, $L(x, u, \partial u, \ldots, \partial^k u)$ is a null Lagrangian if the Euler–Lagrange equations

$$\delta I/\delta u^\nu = 0, \quad (\nu = 1, \ldots, q), \qquad (2\cdot3)$$

for the variational problem

$$I[u] = \int_\Omega L(x, u, \ldots, \partial^k u)\, dx,$$

vanish identically. Clearly, if $L = \operatorname{Div} P$ is a divergence, by the divergence theorem I depends only on the boundary values of u, so $\delta I/\delta u^\nu \equiv 0$, and L is a null Lagrangian. The converse is also true:

THEOREM 2·4. *Let $L(x, u, \ldots, \partial^k u)$ be smooth. Then L is a null Lagrangian if and only if*

$$L = \operatorname{Div} P,$$

for some p-tuple $P(x, u, \ldots, \partial^m u)$.

There is an analogous result for null divergences.

THEOREM 2·5. *A p-tuple $P(x, u, \ldots, \partial^n u)$ is a null divergence if and only if there exist functions $Q_{ij}(x, u, \ldots, \partial^m u)$, $i, j = 1, \ldots, p$ satisfying*

$$\left.\begin{aligned} &\text{(i)}\ \ Q_{ij} = -Q_{ji}, \\ &\text{(ii)}\ \ P_i = \sum_{j=1}^{p} D_j Q_{ij}. \end{aligned}\right\} \qquad (2\cdot4)$$

Moreover, if P depends smoothly on parameters, $P_i(x, \ldots, \partial^n u, \lambda)$, $\lambda \in \mathbb{R}^r$, so do the corresponding $Q_{ij}(x, \ldots, \partial^m u, \lambda)$.

In fact, these two theorems are just the last two terms in a long exact sequence, similar to the de Rham sequence in differential geometry, of great importance in the calculus of variations, but only recently discovered. Avoiding the introduction of differential form notation, this result can be stated as follows.

Definition 2·6. Let $\mathscr{A}^l$ denote the vector space of $\binom{p}{l}$-tuples of functions: $\mathbf{P}(x, u, \ldots, \partial^n u) = (\ldots, P_{i_1 i_2 \ldots i_l}(x, u, \ldots, \partial^n u), \ldots)$ (n arbitrary) with $1 \leqslant i_\nu \leqslant p$, and where the functions $P_{i_1 \ldots i_l}$ are smooth and skew symmetric in their indices:

$$P_{i_1 \ldots i_\nu \ldots i_\mu \ldots i_l} = -P_{i_1 \ldots i_\mu \ldots i_\nu \ldots i_l}.$$

Define the *generalized total divergence* operator

$$\mathbf{Div}: \mathscr{A}^l \to \mathscr{A}^{l-1},$$

so that

$$\mathbf{Q} = \mathbf{Div}\,\mathbf{P}$$

means

$$Q_{i_1 \ldots i_{l-1}} = \sum_{j=1}^{p} D_j P_{i_1 \ldots i_{l-1} j}.$$

Thus, for $l = 1$,

$$\mathbf{Div}: \mathscr{A}^1 \to \mathscr{A}^0$$

coincides with the usual total divergence, while for $l = 2$,

$$\mathbf{Div}: \mathscr{A}^2 \to \mathscr{A}^1$$

agrees with the operator in (2·4) (with the roles of $\mathbf{P}$ and $\mathbf{Q}$ reversed). These total divergence operators form an exact complex:

THEOREM 2·7. *Given $l \neq 0, p$, and $\mathbf{P} \in \mathscr{A}^l$, then*

$$\mathbf{Div}\,\mathbf{P} = 0 \quad \textit{in } \mathscr{A}^{l-1} \tag{2·5}$$

if and only if

$$\mathbf{P} = \mathbf{Div}\,\mathbf{Q} \tag{2·6}$$

for some $\mathbf{Q} \in \mathscr{A}^{l+1}$. For $l = p$, the statement holds with (2·6) replaced by the condition that $\mathbf{P}$ be constant (independent of $x, u, \ldots, \partial^n u$). For $l = 0$, the conclusion holds with (2·5) replaced by the condition that $\mathbf{P}$ be a null Lagrangian. If $\mathbf{P} = \mathbf{P}(\lambda)$ depends smoothly on parameters $\lambda \in \mathbb{R}^r$, then $\mathbf{Q} = \mathbf{Q}(\lambda)$ does likewise.

Detailed developments of the theory and applications of Theorem 2·7, including proofs, can be found in [**1**], [**20**], [**21**], [**23**] and (in the polynomial case) [**18**].

3. *Homogeneous null divergences*

A function (or p-tuple) P is called *homogeneous* of order k if it depends exclusively on k-th order derivatives of u: $P = P(\partial^k u)$. We first consider the problem of explicitly classifying homogeneous null divergences; the more general case will be taken up in Section 4. (The term 'homogeneous' should not be confused with 'algebraically homogeneous' which refers to polynomials in all variables $u, \ldots, \partial^k u$).

The companion problem of classifying homogeneous null Lagrangians has already been solved by Anderson and Duchamp [**1**] and Ball, Currie and Olver [**2**]. (In [**1**], the result is not explicitly written down, but can be inferred from theorems 2·1 and 4·1 therein.) Another proof of the classification can be inferred from Vasilenko [**22**], using results in [**2**].

To state the basic classification, we use the notation of [**2**] for Jacobian determinants:

$$J_K^\alpha = \frac{\partial u^\alpha}{\partial x_K} = \frac{\partial(u_{I_1}^{\nu_1}, \ldots, u_{I_r}^{\nu_r})}{\partial(x^{k_1}, \ldots, x^{k_r})} = \det\left(\frac{\partial u_{I_i}^{\nu_i}}{\partial x^{k_j}}\right), \tag{3·1}$$

for $K = (k_1, \ldots, k_r)$, $\alpha = (\nu_1, I_1; \ldots; \nu_r, I_r)$. The Jacobian J^α_K is homogeneous of order k provided each multi-index I_j in α is of order $k-1$ (and, in any case, algebraically homogeneous of degree r.) It is easy to check that each Jacobian determinant is a null Lagrangian; the classification theorem states that these are essentially all the homogeneous null Lagrangians there are.

THEOREM 3·1. *Let L be a homogeneous function of $\partial^k u$. Then the following are equivalent.*

(1) *L is a null Lagrangian.*

(2) *$L = \operatorname{Div} P$ for some P.*

(3) *L is an affine combination of homogeneous Jacobian determinants*, i.e.

$$L = C^0_0 + \sum_{\alpha, k} C^\alpha_K J^\alpha_K, \tag{3·2}$$

for suitable constants C^α_K. (In particular, L is a polynomial function of $\partial^k u$.)

Now suppose P is a homogeneous null divergence. By Theorem 2·5, each component P_i is a divergence, and hence by Theorem 3·1 an affine combination of Jacobian determinants. However, the P_i are, of course, not independent, so more work is needed to completely classify all such P.

To accomplish this, we first write down some basic homogeneous null divergences generalizing the identities in Examples 2·2 and 2·3. Given α as above, and

$$L = (l_1, \ldots, l_{r+1}), \quad 1 \leqslant l_1 < l_2 < \ldots < l_{r+1} \leqslant p,$$

define N^α_L to be the p-tuple whose i-th entry is

$$N^\alpha_{L,i} = \begin{Bmatrix} 0 & i \notin L, \\ (-1)^{\kappa-1} J^\alpha_{L_{\hat\kappa}}, & i = l_\kappa. \end{Bmatrix} \tag{3·3}$$

Here $L_{\hat\kappa} = (l_1, \ldots, l_{\kappa-1}, l_{\kappa+1}, \ldots, l_{r+1})$. It is not too difficult to check directly that N^α_L is a null divergence, i.e.

$$\sum_{\kappa=1}^{r+1} (-1)^{\kappa-1} D_{l_\kappa} J^\alpha_{L_{\hat\kappa}} \equiv 0. \tag{3·4}$$

(However, the proof of our characterization theorem will also provide a computationally simpler proof of this identity.) The basic theorem to be proved is that the N^α_L provide a complete list of homogeneous null divergences:

THEOREM 3·2. *Let P be a homogeneous p-tuple of functions of $\partial^k u$. Then the following are equivalent:*

(1) *P is a null divergence*: $\operatorname{Div} P \equiv 0$.

(2) *$P_i = \Sigma D_j Q_{ij}$ with $Q_{ij} = -Q_{ji}$.*

(3) *P is an affine combination of the above 'Jacobian null divergences':*

$$P = C_0 + \sum_{\alpha, L} C^\alpha_L N^\alpha_L, \quad C_0 = (C^1_0, \ldots, C^p_0), \tag{3·5}$$

for suitable constants C^α_L.

The proof of this theorem relies on the transform theory developed in **[2]**, **[18]**. We therefore defer the details until Section 6.

It should, however, be remarked that the proof does not follow immediately from Theorems 3·1 and 2·5 owing to identities among Jacobian determinants stemming from the quadratic P-relations among (ordinary) determinants, cf. ([2]; p. 155). For instance, the null divergence

$$D_x\left(\frac{\partial(u_x,u_z)}{\partial(y,w)}-\frac{\partial(u_x,u_w)}{\partial(y,z)}\right)+D_z\frac{\partial(u_w,u_x)}{\partial(x,y)}+D_w\frac{\partial(u_x,u_y)}{\partial(x,z)}=0,$$

can be changed into an identity of the type discussed in Example 2·3 using the relations

$$\frac{\partial(u_x,u_z)}{\partial(y,w)}-\frac{\partial(u_x,u_w)}{\partial(y,z)}=\frac{\partial(u_x,u_y)}{\partial(z,w)},$$

$$\frac{\partial(u_w,u_x)}{\partial(x,y)}=\frac{\partial(u_x,u_y)}{\partial(w,x)}.$$

More complicated examples can easily be constructed.

4. *Low order divergence expressions*

More generally, we can ask the question as to how a null divergence (not necessarily homogeneous) depends on its highest order derivatives $\partial^k u$. For null Lagrangians, Anderson and Duchamp [1] generalized theorem 3·1 to show that if $L(x,u,\ldots,\partial^k u)$ is a null Lagrangian, and we fix $x,u,\ldots,\partial^{k-1}u$ to be constant, then the resulting function of $\partial^k u$ is also a null Lagrangian, and hence an affine combination of Jacobian determinants. An analogous statement holds for null divergences. We state this result in more general form.

PROPOSITION 4·1. *Suppose P depends on $x,u,\ldots,\partial^k u$, and* $\operatorname{Div} P$ *also depends on $x,u,\ldots,\partial^k u$ (but not $\partial^{k+1}u$). Then for any fixed $x=\tilde c$, $u=c_0,\ldots,\partial^{k-1}u=c_{k-1}$, the p-tuple $P(\tilde c,c_0,\ldots,c_{k-1},\partial^k u)$, as a function of $\partial^k u$, is a null divergence, hence of the form* (3·5). *In particular, $P(x,u,\ldots,\partial^k u)$ is a polynomial in $\partial^k u$.*

To prove this, we first break the total derivatives (1·1) up into homogeneous pieces:

$$D_i=\partial_i+\sum_{m=0}^{\infty}D_i^{(m)},\quad \partial_i=\partial/\partial x^i, \tag{4·1}$$

where $D_i^{(m)}$ denotes the sum of all terms in (1·1) with multi-indices J of order m. Similarly, let $\operatorname{Div}=\operatorname{div}+\Sigma\operatorname{Div}^{(m)}$.

Now if P depends on $x,u,\ldots,\partial^k u$, then the only terms in $\operatorname{Div}P$ which depend on $\partial^{k+1}u$ are those in $\operatorname{Div}^{(k)}P$. Thus if P satisfies the conditions of the proposition, then

$$\operatorname{Div}^{(k)}P=0. \tag{4·2}$$

For $Q(x,u,\ldots,\partial^k u)$ any function, let $Q|_c$ denote the homogeneous function

$$Q(\tilde c,c_0,\ldots,c_{k-1},\partial^k u)$$

for fixed $\tilde c,\ldots,c_{k-1}$. To complete the proof of the proposition, it suffices to note that for any $1\leqslant i\leqslant p$,

$$D_i(Q|_c)=(D_i^{(k)}Q)|_c. \tag{4·3}$$

Thus (4·2) implies that $\operatorname{Div}(P|_c)=0$, hence the proposition.

More generally, the same method of proof yields:

PROPOSITION 4·2. *Suppose $\mathbf{P}\in\mathscr{A}^l$ and* $\operatorname{Div}\mathbf{P}\in\mathscr{A}^{l-1}$ *both depend on $x,u,\ldots,\partial^k u$. Then, for each fixed $x=\tilde c$, $u=c_0,\ldots,\partial^{k-1}u=c_{k-1}$,*

$$\mathbf{P}|_c=\mathbf{P}(\tilde c,c_0,\ldots,c_{k-1},\partial^k u)$$

is a generalized null divergence:

$$\mathbf{Div}\,(\mathbf{P}|_c) = 0,$$

and is (*by Theorem* 3·1) *a polynomial in* $\partial^k u$.

(There is a generalization of the classification Theorem 3·2 to these generalized null divergences, but we will not require this result.)

5. *Order of derivatives in conservation laws*

Return to the problem of classifying conservation laws for a given system of differential equations. In the identity

$$\operatorname{Div} P = Q,$$

where Q is given by (1·4), it is often necessary to specify in advance the order of derivatives on which Q can depend, i.e. to assume $Q = Q(x, u, \ldots, \partial^k u)$. One would also like to assert that P can depend on at most k-th order derivatives, but it is conceivable that P depends on derivatives up to order $l > k$, but terms in $\operatorname{Div} P$ involving derivatives of order higher than k cancel out. A typical example would be a null divergence.

The aim of this section is to prove that, except for the trivial possibility of adding a null divergence, P can be assumed to depend on derivatives of order $\leqslant k$ if Q does. This result is fundamental in the systematic classification of conservation laws, especially those of bounded order. See also Anderson and Duchamp [1].

THEOREM 5·1. *Suppose*

$$\operatorname{Div} P = Q,$$

where $Q = Q(x, u, \ldots, \partial^{k+1} u)$, *and depends linearly on the highest order derivatives* $\partial^{k+1}u$. *Then there is a null divergence* N *such that* $\tilde{P} \equiv P - N$ *depends only on* $x, u, \ldots, \partial^k u$ *and also satisfies*

$$\operatorname{Div} \tilde{P} = Q.$$

If $Q(\lambda)$ *depends smoothly on parameters* $\lambda \in \mathbb{R}^r$, *so does* $\tilde{\mathbf{P}}(\lambda)$.

Proof. Actually, we will prove the more general result that if $\mathbf{P} \in \mathscr{A}^l$, $\mathbf{Q} \in \mathscr{A}^{l-1}$,

$$\mathbf{Div}\,\mathbf{P} = \mathbf{Q},$$

and $\mathbf{Q} = \mathbf{Q}(x, u, \ldots, \partial^k u; \lambda)$ is a polynomial of degree m in the highest order derivatives $\partial^k u$, then there exists $\tilde{\mathbf{P}}(x, u, \ldots, \partial^k u; \lambda) \in \mathscr{A}^l$ with $\tilde{\mathbf{P}}$ a polynomial of degree at most $m-1$ in the highest order derivatives, and

$$\mathbf{Div}\,\tilde{\mathbf{P}} = \mathbf{Q},$$

also. This will be proved by reverse induction on l.

For $l = p$ the result is easy to see. Indeed, $\mathbf{P}$ consists of a single function

$$P = P_{1,2\ldots p},$$

and $\mathbf{Div}$ can be identified with the 'total gradient',

$$Q_{1\ldots i-1, i+1\ldots p} = D_i P.$$

Its easy to see that if P depends on $\partial^n u$, then $D_i P$ must depend (linearly) on $\partial^{n+1} u$, so the result follows, with $\mathbf{Q}$ at most linear in its highest order derivatives.

Turning to the induction step, we assume the result is true for $l+1$ and prove it for l. Thus suppose $\mathbf{P} = \mathbf{P}(x, u, \ldots, \partial^n u; \lambda) \in \mathscr{A}^l$, and $\mathbf{Q} = \mathbf{Q}(x, u, \ldots, \partial^k u; \lambda) \in \mathscr{A}^{l-1}$ is as above, with $n \geqslant k$. By Proposition 4·2, $\mathbf{P}$ is a polynomial in $\partial^n u$, hence

$$\mathbf{P} = \mathbf{P}^0 + \ldots + \mathbf{P}^s, \quad \mathbf{P}^s \not\equiv 0, \quad s \geqslant 1,$$

where each $\mathbf{P}^j$ is an algebraically homogeneous polynomial in $\partial^n u$ of degree j. If $n = k$ and $s \leqslant m-1$ we are done. Otherwise, note that if we break up

$$\mathbf{Div} = \mathbf{div} + \Sigma\, \mathbf{Div}^{(m)},$$

as in (4·1), then

$$\mathbf{Div}^{(n)}\, \mathbf{P}^j = 0, \quad j = 1, \ldots, s. \tag{5·1}$$

Also, the only terms in $\mathbf{Div}\,\mathbf{P}$ which are polynomials of degree $s+1$ in $\partial^n u$ are

$$\mathbf{Div}^{(n-1)}\, \mathbf{P}^s = 0, \tag{5·2}$$

which holds since $\mathbf{Q}$ is either independent of $\partial^n u$ $(n > k)$ or of lower degree $(n = k)$.

Now, if $x = \tilde{c}$, $u = c_0, \ldots, \partial^{n-2} u = c_{n-2}$ are fixed, then (5·1) and (5·2) imply

$$\mathbf{Div}\, \mathbf{P}^s(\tilde{c}, c_0, \ldots, c_{n-2}, \partial^{n-1} u, \partial^n u; \lambda) = 0.$$

By the induction hypothesis, there exists $\mathbf{R} = \mathbf{R}(\tilde{c}, c_0, \ldots, c_{n-2}; x, u, \ldots, \partial^n u; \lambda) \in \mathscr{A}^{l+1}$, a polynomial of degree $\leqslant s-1$ in $\partial^n u$, depending smoothly on all its arguments, with

$$\mathbf{P}^s(\tilde{c}, \ldots, c_{n-2}, \partial^{n-1} u, \partial^n u; \lambda) = \mathbf{Div}\, \mathbf{R}(\tilde{c}, \ldots, c_{n-2}; x, \ldots, \partial^n u; \lambda). \tag{5·3}$$

Given $y \in \mathbb{R}^p$, $v \in \mathbb{R}^q$, consider the functions $\mathbf{R}(y, v, \ldots, \partial^{n-2} v; x, u, \ldots, \partial^n u; \lambda)$. Let $\mathbf{Div}^y$, $\mathbf{Div}^x$ denote the total divergences with respect to y and x respectively, treating v as a function only of y and u as a function only of x. By the chain rule, when $x = y$, $u = v$,

$$\begin{aligned}&\mathbf{Div}\, \mathbf{R}(x, u, \ldots, \partial^{n-2} u; x, u, \ldots, \partial^n u; \lambda)\\ &\quad = \mathbf{Div}^y\, \mathbf{R}(x, \ldots, \partial^{n-2} u; x, \ldots, \partial^n u; \lambda) + \mathbf{Div}^x\, \mathbf{R}(x, \ldots, \partial^{n-2} u; x, \ldots, \partial^n u; \lambda).\end{aligned}$$

Further note that (5·3) can be rewritten as

$$\mathbf{P}^s(x, u, \ldots, \partial^n u; \lambda) = \mathbf{Div}^x\, \mathbf{R}(x, \ldots, \partial^{n-2} u; x, \ldots, \partial^n u; \lambda).$$

Now set

$$\mathbf{N} = \mathbf{Div}\, \mathbf{R}(x, \ldots, \partial^{n-2} u; x, \ldots, \partial^n u; \lambda).$$

Then

$$\hat{\mathbf{P}}^s = \mathbf{P}^s - \mathbf{N} = \mathbf{Div}^y\, \mathbf{R}(x, \ldots, \partial^{n-2} u; x, \ldots, \partial^n u; \lambda)$$

is a polynomial in $\partial^n u$ of degree at most $s-1$. Thus

$$\hat{\mathbf{P}} = \mathbf{P} - \mathbf{N}$$

satisfies

$$\mathbf{Div}\, \hat{\mathbf{P}} = \mathbf{Div}\, \mathbf{P} = \mathbf{Q},$$

and is of degree at most $s-1$ in $\partial^n u$.

We now continue this process to reduce the degee s and the order n of the highest order derivatives entering into $\mathbf{P}$ until we reach $\tilde{\mathbf{P}}$ depending only on $x, u, \ldots, \partial^k u$, and of degree at most $m-1$ in $\partial^k u$. This completes the induction step, and hence proves the theorem.

In [16], this result will be of key importance in the classification of conservation laws for linear, isotropic elasticity. However, its potential range of applicability is much wider.

6. *Classification of homogeneous null divergences*

The proof of Theorem 3·2 uses the same transform techniques as developed in [2] to prove the classification Theorem 3·1 for homogeneous null Langrangians. We thus begin by briefly reviewing the transform, which maps differential polynomials to algebraic polynomials, thereby reducing questions about the former to problems in ordinary commutative algebra, to which the powerful methods of algebraic geometry can be applied. For the most part, the notation here is the same as in [2], although with the slight modifications introduced in [13].

Let $\mathscr{L}^r = \mathscr{L}^r(p,q,k)$ denote the space of all differential polynomials $L(u, \partial u, \ldots, \partial^k u)$ with complex coefficients, which are algebraically homogeneous polynomials of degree r in their arguments. Thus, for example,

$$uu_x \in \mathscr{L}^2, \quad u^2 u_x u_{xyy} \in \mathscr{L}^4, \quad \text{etc.}$$

(In [2], $\mathscr{L}^r$ denoted the subspace $\mathscr{L}^r_0$ of homogeneous polynomials $L(\partial^k u)$ depending only on k-th order derivatives). Let Z^r be the space of algebraic polynomials $\phi(a^1, b^1; a^2, b^2; \ldots; a^r, b^r)$, with $a^i \in \mathbb{R}^q$, $b^i \in \mathbb{R}^p$, which are linear in the a^i. Here Z^r is the direct sum of its homogeneous subspaces $Z^{r,k}$, consisting of those polynomials homogeneous of degree k in the b^i, as defined in [2]. The symmetric group of permutations of $\{1, \ldots, r\}$ acts on Z^r by

$$\hat{\pi}\phi(a^1, b^1; \ldots; a^r, b^r) = \phi(a^{\pi(1)}, b^{\pi(1)}; \ldots; a^{\pi(r)}, b^{\pi(r)}),$$

for π a permutation. Let Z^r_0 denote the subspace of symmetric polynomials in Z^r; in other words $\phi \in Z^r_0$ if and only if $\hat{\pi}\phi = \phi$ for all permutations π. Let

$$\sigma = \frac{1}{r!}\Sigma\hat{\pi}\colon Z^r \to Z^r_0$$

denote the natural projection, so $\sigma[\phi]$ is the 'symmetrized' version of ϕ.

Define the *transform*

$$\mathscr{F}\colon \mathscr{L}^r \to Z^r_0,$$

to be the linear map whose action on monomials is given by

$$\mathscr{F}(u^{\nu_1}_{I_1} \ldots u^{\nu_r}_{I_r}) = \sigma(a^1_{\nu_1}\, b^1_{I_1} \ldots a^r_{\nu_r} b^r_{I_r}). \tag{6·1}$$

Here for $1 \leqslant \nu \leqslant q$, $I = (i_1, \ldots, i_k)$, u^ν_I denotes the I-th partial derivative of u^ν, as defined in Section 2. Also, $b^j_I = b^j_{i_1} b^j_{i_2} \ldots b^j_{i_k}$, with b^j_i, a^j_ν denoting entries of b^j, a^j respectively. See [4], [18], [2], [13] for more details.

THEOREM 6·1 [2], [13]. *The transform $\mathscr{F}$ gives a linear isomorphism between $\mathscr{L}^r$ and Z^r_0.*

If $\Phi\colon \mathscr{L}^r \to \mathscr{L}^s$ is a linear map, then $\hat{\Phi}\colon Z^r_0 \to Z^s_0$ denotes its transform defined by $\mathscr{F}[\Phi(L)] = \hat{\Phi}(\mathscr{F}L)$.

LEMMA 6·2 [13], [18]. *The transform of the total derivative $D_i\colon \mathscr{L}^r \to \mathscr{L}^r$ is*

$$\hat{D}_i \phi(a^1, \ldots, b^r) = (b^1_i + \ldots + b^r_i)\,\phi(a^1, \ldots, b^r). \tag{6·2}$$

We also need the formula for the transform of a Jacobian determinant as given in [2]. For $K = (k_1, \ldots, k_r)$, and $\alpha = (\nu_1, I_1; \ldots; \nu_r, I_r)$ collections of indices and multi-indices as in (3·1), define B_K to be the $r \times r$ matrix with entries $b^i_{k_j}$, and $(A \otimes B)_\alpha$ to be the $r \times r$ matrix with entries $a^i_{\nu_j} b^i_{I_j}$.

LEMMA 6·3. [2]. *If J^{α}_{K} denotes the Jacobian determinant* (3·1) *then*

$$\mathscr{F}(J^{\alpha}_{K}) = \frac{1}{r!}\det(B_K)\det(A\otimes B)_{\alpha}. \tag{6·3}$$

Turning to the proof of Theorem 3·2, note first that by Theorems 3·1 and 2·5 each component P_i of the homogeneous null divergence P is a polynomial in $\partial^k u$. Clearly, each algebraically homogeneous summand of P is separately a null divergence, since Div preserves the algebraic homogeneity of each monomial. Thus, without loss of generality, assume that $P \in \mathscr{L}^{r,p}_0 \subset \mathscr{L}^{r,p}$, where $\mathscr{L}^{r,p}$ denotes the p-fold Cartesian product of $\mathscr{L}^r$. Let $\mathscr{F}^p : \mathscr{L}^{r,p} \to Z^{r,p}_0$ be the Cartesian product transform, $Z^{r,p}_0$ being again a p-fold Cartesian product of Z^r_0, and let $\phi = \mathscr{F}^p(P)$. Thus from Lemma 6·2 and the fact that $\mathscr{F}$ is an isomorphism, we conclude that P is a null divergence if and only if

$$\sum_{i=1}^{p}\left(\sum_{j=1}^{r} b^j_i\right)\phi_i(a^1,\ldots,b^r) = 0. \tag{6·4}$$

Let $y \in \mathbb{R}^p$, and define the polynomial

$$\psi(a^1,b^1;\ldots;a^r,b^r;y) = \sum_{i=1}^{p} y_i\phi_i(a^1,\ldots,b^r). \tag{6·5}$$

From (6·4) we see that $\psi = 0$ whenever $b^1+\ldots+b^r+y = 0$. Moreover, since P is homogeneous, ϕ_i is a homogeneous polynomial of degree k in $b^1,\ldots,b^r$; hence ψ is homogeneous of degree k in $b^1,\ldots,b^r$ and degree 1 in y. Thus for $\lambda_1,\ldots,\lambda_{r+1} \in \mathbb{C}$,

$$(\lambda_1\ldots\lambda_r)^k\lambda_{r+1}\psi(a^1,b^1;\ldots;a^r,b^r;y) = \psi(a^1,\lambda_1 b^1;\ldots;a^r,\lambda_r b^r;\lambda_{r+1}y) = 0,$$

whenever $\lambda_1 b^1+\ldots+\lambda_r b^r+\lambda_{r+1}y = 0$. By continuity, we conclude that $\psi = 0$ whenever $b^1,\ldots,b^r,y$ are linearly dependent.

Given a multi-index L with $1 \leqslant l_1 < l_2 < \ldots < l_{r+1} \leqslant p$, let Y_L denote the $(r+1)\times(r+1)$ matrix with (i,j)-th entry $b^i_{l_j}$ for $1 \leqslant i \leqslant r$, or y_{l_j} for $i = r+1$. Then $b^1,\ldots,b^r,y$ are linearly dependent if and only if $\det Y_L = 0$ for all such multi-indices L.

At this point we require some deep results from algebraic geometry. Let $\mathscr{I}$ denote the polynomial ideal generated by the determinants $\det Y_L$; $\mathscr{I}$ is known as a determinantal ideal. By the Hilbert Nullstellensatz (cf. theorem 4·6 in [2], or [6]; p. 254), since ψ vanishes whenever $\det Y_L = 0$, some power of ψ must be in the ideal $\mathscr{I}$. Moreover, by a theorem of Northcott [10] and Mount [9] (see also theorem 4·7 in [2]) the determinantal ideal $\mathscr{I}$ is prime, hence this power can be taken to be one. In other words, there exist polynomials ψ_L such that

$$\psi = \Sigma\psi_L \det Y_L. \tag{6·6}$$

(Compare the proof of lemma 4·8 in [2]). Expanding the determinants along the last row,

$$\det Y_L = \sum_{\kappa=1}^{r+1}(-1)^{\kappa+r+1}y_{l_\kappa}\det B_{L\hat{\kappa}}, \tag{6·7}$$

where we are using the same notation as in (3·3). From (6·5, 6, 7), we see that

$$\phi_i = \sum_{i=l_\kappa}(-1)^{\kappa+r+1}\psi_L\det B_{L\hat{\kappa}}. \tag{6·8}$$

Finally, we use the fact that each ϕ_i is symmetric. Thus applying the symmetrizing map σ to both sides of (6·8) we have

$$\phi_i = \Sigma(-1)^{\kappa+r+1}\tilde{\psi}_L \det B_{L\hat{\kappa}}, \tag{6·9}$$

where

$$\tilde{\psi}_L = \frac{1}{r!}\sum_{\pi}(\operatorname{sign}\pi)\,\hat{\pi}(\psi_L).$$

In particular, for any permutation π,

$$\hat{\pi}(\tilde{\psi}_L) = (\operatorname{sign}\pi)\,\tilde{\psi}_L.$$

Lemma 4·9 in [2] then implies that

$$\tilde{\psi}_L = \sum_{\alpha} C^{\alpha}_{L} \det(A \otimes B)_{\alpha}, \tag{6·10}$$

for suitable constants C^{α}_{L}. Inserting (6·10) into (6·9) and using the formula for the transform of a Jacobian determinant, we conclude that

$$\mathscr{F}^{-1}(\phi_i) = P_i = \Sigma\, C^{\alpha}_{L} N^{\alpha}_{L,i},$$

where $N^{\alpha}_{L,i}$ is given by (3·3). This completes the proof of Theorem 3·2.

Appendix. Nondegeneracy conditions

Let

$$\Delta_{\kappa}(x, u, \ldots, \partial^m u) = 0 \quad (\kappa = 1, \ldots, l),$$

be a system of partial differential equations, the Δ_κ assumed to be C^∞ functions. If $u = f(x)$ is a C^∞ solution, then u also satisfies all the 'prolonged' equations

$$D^J \Delta_\kappa = 0, \tag{A 1}$$

for all multi-indices J. The system is of *maximal rank* if for each n the Jacobian matrix of $D^J\Delta_\kappa$ for all $|J| \leqslant n$ with respect to all variables $x, u, \ldots, \partial^{n+m}u$ is of maximal rank whenever the equations are satisfied. For polynomial systems, this is equivalent to the statement that the Δ_κ generate a radical differential ideal, cf. [17]. (In [17] an example of a single prime ordinary differential polynomial whose differential ideal is nevertheless not radical is discussed!) For linear systems, or evolutionary systems, this condition is easy to verify.

A system is *locally solvable* if for any n and for each fixed $x_0, u_0, \ldots, \partial^{n+m}u_0$ satisfying the prolonged equations (A 1) for $|J| \leqslant n$, there is a C^∞ solution $u = f(x)$ defined in a neighbourhood of x_0 satisfying initial conditions $u_0 = f(x_0), \ldots, \partial^{n+m}u_0 = \partial^{n+m}f(x_0)$. (These initial conditions should be contrasted with the usual Cauchy problem where, except for ordinary differential equations, the initial data must be specified along an entire submanifold of $\mathbb{R}^p$.) For analytic systems, the Cauchy–Kowaleski theorem [3] ensures local solvability. However, counter-examples such as that constructed by Lewy (cf. [5]) show that this question is more delicate in general.

A system is *nondegenerate* if it is both of maximal rank and locally solvable. The importance of nondegeneracy for symmetry group theory is discussed in [12], where a proof of the following basic result is outlined. This provides the connection between (1·3) and (1·4).

THEOREM. *If $\Delta = 0$ is nondegenerate, and $Q(x, u, \ldots, \partial^k u) = 0$ whenever $u = f(x)$ is a solution, then*

$$Q = \Sigma Q_{J,\kappa} D^J \Delta_\kappa,$$

for suitable nonvanishing functions $Q_{J,\kappa}(x, u, \ldots, \partial^k u)$.

REFERENCES

[1] I. M. ANDERSON and T. DUCHAMP. On the existence of global variational principles. *Amer. J. Math.* **102** (1980), 781–868.

[2] J. M. BALL, J. C. CURRIE and P. J. OLVER. Null Lagrangians, weak continuity, and variational problems of arbitrary order. *J. Funct. Anal.* **41** (1981), 135–174.

[3] R. COURANT and D. HILBERT. *Methods of Mathematical Physics*, vol. II (Interscience, 1962).

[4] I. M. GEL'FAND and L. A. DIKII. Asymptotic behaviour of the resolvent of Sturm–Liouville equations and the algebra of the Korteweg–de Vries equation. *Russian Math. Surveys* **30** (1975), 77–113.

[5] L. HÖRMANDER. *Partial Differential Operators* (Springer-Verlag, 1969).

[6] N. JACOBSON. *Lectures in Abstract Algebra*, vol. 3 (Van Nostrand, 1964).

[7] J. K. KNOWLES and E. STERNBERG. On a class of conservation laws in linearized and finite elastostatics. *Arch. Rational Mech. Anal.* **44** (1972), 187–211.

[8] G. L. LAMB. *Elements of Soliton Theory* (Wiley-Interscience, 1980).

[9] K. R. MOUNT. A remark on determinantal loci. *J. London Math. Soc.* **42** (1967), 595–598.

[10] D. G. NORTHCOTT. Some remarks on the theory of ideals defined by matrices. *Quart. J. Math. Oxford* **14** (1963), 193–204.

[11] P. J. OLVER. Euler operators and conservation laws of the BBM equation. *Math. Proc. Cambridge Philos. Soc.* **85** (1979), 143–160.

[12] P. J. OLVER. *Applications of Lie Groups to Differential Equations* (Lecture Notes, University of Oxford, 1980).

[13] P. J. OLVER. Hyperjacobians, determinantal ideals and weak solutions to variational problems. *Proc. Roy. Soc. Edinburgh* (to appear).

[14] P. J. OLVER. Conservation laws of free boundary problems and the classification of conservation laws for water waves. *Trans. Amer. Math. Soc.*, **277**, (1983), 353–380..

[15] P. J. OLVER. Conservation laws in elasticity. I. General results. *Arch. Rational Mech. Anal.* (to appear).

[16] P. J. OLVER. Conservation laws in elasticity. II. Linear isotropic homogeneous elastostatics. *Arch. Rational Mech. Anal.* (to appear).

[17] J. F. RITT. *Differential Algebra* (Dover, 1966).

[18] C. SHAKIBAN. A resolution of the Euler operator. II. *Math. Proc. Cambridge Philos. Soc.* **89** (1981), 501–510.

[19] W. A. STRAUSS. Nonlinear invariant wave equations. In *Invariant Wave Equations*, Lecture Notes in Physics, vol. 73 (Springer-Verlag, 1978), pp. 197–249.

[20] F. TAKENS. A global version of the inverse problem of the calculus of variations. *J. Differential Geom.* **14** (1979), 543–562.

[21] W. M. TULCZYJEW. The Lagrange complex. *Bull. Soc. Math. France* **105** (1977), 419–431.

[22] G. N. VASILENKO. Weak continuity of Jacobians. *Siberian Math. J.* **22** (1981), 355–360.

[23] A. M. VINOGRADOV. A spectral sequence associated with a nonlinear differential equation and algebro-geometric foundations of Lagrangian field theory with constraints. *Soviet Math. Dokl.* **19** (1978), 144–148.