Proceedings of the Royal Society of Edinburgh, **95A**, 317–340, *1983*

Hyperjacobians, determinantal ideals and weak solutions to variational problems

Peter J. Olver†
School of Mathematics, University of Minnesota,
Minneapolis, MN 55455, U.S.A.

(MS received 8 February 1983. Revised MS received 19 April 1983)

Synopsis

The problem of classifying homogeneous null Lagrangians satisfying an nth order divergence identity is completely solved. All such differential polynomials are affine combinations of higher order Jacobian determinants, called hyperjacobians, which can be expressed as higher dimensional determinants of higher order Jacobian matrices. Special cases, called transvectants, are of importance in classical invariant theory. Transform techniques reduce this question to the characterization of the symbolic powers of certain determinantal ideals. Applications to the proof of existence of minimizers of certain quasi-convex variational problems with weakened growth conditions are discussed.

1. Introduction

In earlier joint work with J. M. Ball and J. C. Currie, [**5**], (hereafter referred to as BCO), on variational problems in non-linear elasticity, the concept of a null Lagrangian was the key ingredient in providing interesting classes of non-convex variational problems for which the existence of minimizers could be proved. If x, u are the independent and dependent variables in the problem, a *null Lagrangian* is a continuous function of x, u and the derivatives of u so that the Euler–Lagrange equations for the corresponding variational problem $\int L\,dx$ vanish identically. In essence, the variational problems treated in BCO are convex functions of null Lagrangians depending on the highest order derivatives of u, and satisfying certain growth conditions in these arguments.

A classical result states that L is a null Lagrangian if and only if it can be written as a divergence

$$L = \operatorname{Div} P, \tag{1.1}$$

for some P. The Jacobian determinant

$$\frac{\partial(u, v)}{\partial(x, y)} = u_x v_y - u_y v_x = D_x(uv_y) - D_y(uv_x) \tag{1.2}$$

is the simplest example. (Here subscripts denote partial derivatives, and D_x denotes the total derivative with respect to x.) It was noted in BCO that the divergence identity (1.1) can be used to "weakly" define L as a distribution over a wider class of functions, and thus enable one to weaken the growth conditions in

† Research supported in part by NSF grant MCS 81-00786 and a United Kingdom SRC research grant.

the relevant variational problems, yet still be able to prove existence of minimizers. (One slight difficulty is the as yet unsolved problem of whether this distributionally defined L agrees with L itself if the former is a function. See **[4]** for a discussion of this point.) It was noted in BCO that certain null Lagrangians had nicer identities, in other words could be written as higher order divergences, the archetypal example being

$$\begin{aligned} L = \frac{\partial(u_x, u_y)}{\partial(x, y)} &= u_{xx}u_{yy} - u_{xy}^2 \\ &= -D_x^2(\tfrac{1}{2}u_y^2) + D_x D_y(u_x u_y) - D_y^2(\tfrac{1}{2}u_x^2). \end{aligned} \tag{1.3}$$

Such identities can be similarly utilized to further weaken the growth conditions in the relevant types of variational problems.

One of the principal results of BCO was that any homogeneous null Lagrangian, meaning one depending exclusively on kth order partial derivatives of u for some fixed k, is necessarily an affine combination of Jacobian determinants whose arguments (e.g. the u, v in (1.2)) are $(k-1)$st order derivatives of u. (It should be mentioned that this result was independently proved by Anderson and Duchamp, **[3]**.) The problem was then raised of effecting a similar classification of homogeneous null Lagrangians which are nth *order divergences*, i.e. satisfy a "nice identity" of the form

$$L = \sum_{|I|=n} D^I Q_I \tag{1.4}$$

for certain Q_I. (See subsection 2.1 for the multi-index notation.)

The main result of this paper is that any homogeneous null Lagrangian which is an nth order divergence is necessarily an affine combination of an interesting new class of differential polynomials, which we name n-th order ***hyperjacobians***. These hyperjacobians bear an analogous relation to the ordinary Jacobian determinants (which are the same as first order hyperjacobians) to that which higher order derivatives bear to first order derivatives. Simple examples of second and third order hyperjacobians are

$$\begin{aligned} \frac{\partial^2(u, v)}{\partial(x, y)^2} &\equiv \frac{\partial(u_x, v_y)}{\partial(x, y)} - \frac{\partial(u_y, v_x)}{\partial(x, y)} \\ &= u_{xx}v_{yy} - 2u_{xy}v_{xy} + u_{yy}v_{xx} \\ &= -D_x^2(u_y v_y) + D_x D_y(u_x v_y + u_y v_x) - D_y^2(u_x v_x) \end{aligned} \tag{1.5}$$

and

$$\begin{aligned} \frac{\partial^3(u, v)}{\partial(x, y)^3} &\equiv \frac{\partial^2(u_x, v_y)}{\partial(x, y)^2} - \frac{\partial^2(u_y, v_x)}{\partial(x, y)^2} \\ &= u_{xxx}v_{yyy} - 3u_{xxy}v_{xyy} + 3u_{xyy}v_{xxy} - u_{yyy}v_{xxx} \\ &= -D_x^3(u_y v_{yy}) + D_x^2 D_y(2u_y v_{xy} + u_x v_{yy}) \\ &\quad - D_x D_y^2(u_y v_{xx} + 2u_x v_{xy}) + D_y^3(u_x v_{xx}). \end{aligned} \tag{1.6}$$

Note that the nice identity (1.3) for an ordinary Jacobian is a special case of the second order hyperjacobian identity (1.5) when $u = v$. The reader can easily imagine the general formula for a hyperjacobian, which can be expressed using Cayley's theory of higher dimensional determinants. They were apparently first written down by Escherich, [**7**], and Gegenbauer, [**9**]. Special types of hyperjacobians with polynomial functions as arguments are of great importance in classical invariant theory, where they are called transvectants (German: Überschiebung) [**10**], [**11**], [**12**]. Cayley originally investigated certain special cases in his theory of hyperdeterminants, cf. [**11**, p. 84]. All these earlier investigations were limited to polynomial functions and, moreover, as far as I can determine, none of these authors was aware of the key property that these differential polynomials are higher order divergences. See sections 2.4 and 3.3 for these connections.

In BCO, a transform similar to that discussed by Gel'fand and Dikii, [**8**], and Shakiban, [**21**], [**22**] was introduced. It changes questions about differential polynomials to problems in the theory of ordinary algebraic polynomials, to which powerful techniques in algebraic geometry and commutative algebra can be applied. (It is significant that, for polynomial functions, this transform is essentially equivalent to the powerful symbolic method of Aronhold in classical invariant theory.) The problem of classifying homogeneous null Lagrangians transforms into the problem of whether the *determinantal ideal* generated by the maximal minors of a matrix of independent variables is prime, a result proved by Northcott, [**17**], and Mount, [**16**], in the 1960's. The characterization of n-th order divergences as hyperjacobians rests on the deeper fact that the symbolic powers of this same determinantal ideal are the same as the actual powers of it. By a rather fortuitous coincidence, this theorem has been recently proved by Trung, [**24**], and, in further generality, by DeConcini, Eisenbud and Procesi, [**6**]. The application of this important result to our classification problem via transform theory is presented in section 3.

The final section returns to the original inspiration for the development of the preceding theory. An n-th order hyperjacobian of degree r can be defined for functions in Sobolev spaces with $[n/r]$ fewer derivatives. ([ ] denotes integer part.) Compact embeddings of Sobolev spaces over bounded domains then yield sequential weak continuity results for hyperjacobians of the type found in BCO for ordinary Jacobian determinants. This in turn, again by methods of BCO, yields existence results for minimizers of certain special types of quasi-convex variational problems with weakened growth conditions on the integrands. The application of this result, however, is somewhat limited until the problem on the agreement of weak and classical definition of these hyperjacobians mentioned above is fully resolved.

The one remaining problem in the development of hyperjacobians is the efficient computation of the polynomials Q_I in the identity (1.4). A quick perusal of the examples at the end of section 3 should convince the reader that for all but the simplest hyperjacobians this is a non-trivial algebraic computation. For ordinary Jacobian determinants, the standard divergence identities can be found from reading off the coefficient of $dx_{k_1} \wedge \ldots \wedge dx_{k_r}$ in the differential form identity

$$du^1 \wedge \ldots \wedge du^r = d(u^1\, du^2 \wedge \ldots \wedge du^r).$$

This indicates the possibility of developing a theory of higher order differential forms, so that the hyperjacobian identities can be found as coefficients in certain easier differential form identities. Thus, identities (1.5), (1.6) would be the coefficients of $(dx\wedge dy)^2$, $(dx\wedge dy)^3$ respectively in the "hyperform identities"

$$d^2u * d^2v = d^2(du * dv),$$

and

$$d^3u * d^3v = d^3(du * d^2v).$$

Such a theory has been developed in the companion paper **[20]** to this. The theory relies on the concept of a Schur functor, **[15]**, **[23]**, which has been recently developed for studying resolutions of determinantal ideals. The resulting differential hyperforms can be defined over smooth manifolds, and lead to interesting generalizations of the de Rham complex and "hypercohomology" theories based on higher order derivatives of the defining functions. Lack of space precludes any further discussion here of this theory, which is still under development.

2. Hyperjacobians

Suppose $u^1, \ldots, u^r$ are functions of $x^1, \ldots, x^r$. The Jacobian determinant

$$\frac{\partial(u^1, \ldots, u^r)}{\partial(x^1, \ldots, x^r)} = \det\left(\frac{\partial u^i}{\partial x^j}\right)$$

is well-known to be expressible in divergence form. In this section, higher order analogues of this Jacobian determinant, which we name "hyperjacobians", are introduced and their elementary properties derived. The resulting formulae can also be written using the theory of higher dimensional determinants, as discussed at the end of this section. For homogeneous polynomial functions, certain hyperjacobians appear in classical invariant theory under the same transvectant, a connection noted in section 3.3.

2.1. Multi-index notation

The spaces $X = \mathbb{R}^p$ and $U = \mathbb{R}^q$ representing the independent variables $x = (x_1, \ldots, x_p)$ and dependent variables $u = (u^1, \ldots, u^q)$, will be fixed throughout, the u's being thought of as functions of the x's. There are two types of multi-indices used in this paper. The first, denoted by I or J, are p-tuples $I = (i_1, \ldots, i_p)$ with $i_\nu \geqq 0$. Set $|I| = i_1 + \ldots + i_p$, $x^I = (x_1)^{i_1} \ldots (x_p)^{i_p}$, $\partial^I = (\partial_1)^{i_1} \ldots (\partial_p)^{i_p}$, where $\partial_j = \partial/\partial x_j$, etc. The Greek letters α, β will denote pairs $\alpha = (I, \nu)$ with I a multi-index and $1 \leqq \nu \leqq q$: these are in one-to-one correspondence with the partial derivatives $u_\alpha = u_I^\nu = \partial^I u^\nu$ of the u's.

The second type of multi-index, denoted by K or L, are r-tuples $K = (k_1, \ldots, k_r)$ with $1 \leqq k_\nu \leqq p$. Let $\#K = r$. If $1 \leqq \nu \leqq \#K$, let $K_{\hat{\nu}} = (k_1, \ldots, k_{\nu-1}, k_{\nu+1}, \ldots, k_r)$. Let x_K denote the r-tuple $(x_{k_1}, \ldots, x_{k_r})$. Similarly, if $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_r)$ where each α_i is a pair as above, $\mathbf{u}_{\boldsymbol{\alpha}}$ will denote the r-tuple of partial derivatives $(u_{\alpha_1}, \ldots, u_{\alpha_r})$.

2.2. Definition of hyperjacobians

Given an r-tuple of partial derivatives $\mathbf{u}_{\boldsymbol{\alpha}}$ and a multi-index $K = (k_1, \ldots, k_r)$, of the second type, set

$$u_{\boldsymbol{\alpha},K} = (\partial_{k_1} u_{\alpha_1}, \ldots, \partial_{k_r} u_{\alpha_r}).$$

Define the skew-symmetrized K-th derivative of $\mathbf{u}_{\boldsymbol{\alpha}}$ to be the formal linear sum of r-tuples

$$\frac{\partial \mathbf{u}_{\boldsymbol{\alpha}}}{\partial K} = \frac{\partial \mathbf{u}_{\boldsymbol{\alpha}}}{\partial x_K} = \sum \operatorname{sign}(\pi)\, \mathbf{u}_{\boldsymbol{\alpha}, \pi(K)}, \tag{2.1}$$

the sum being over all permutations π of $\{1, \ldots, r\}$, with $\pi(K) = (k_{\pi 1}, \ldots, k_{\pi r})$. For example,

$$\frac{\partial(u, v)}{\partial(x, y)} = (u_x, v_y) - (u_y, v_x).$$

Clearly, $\partial \mathbf{u}_{\boldsymbol{\alpha}}/\partial K$ is alternating in K;

$$\partial \mathbf{u}_{\boldsymbol{\alpha}}/\partial \pi(K) = \operatorname{sign} \pi \; \partial \mathbf{u}_{\boldsymbol{\alpha}}/\partial K$$

for any permutation π. The above definition extends linearly to $\mathbb{R}$-linear sums of r-tuples. An easy computation proves that these "derivatives" commute.

LEMMA 2.1. *For any* $\boldsymbol{\alpha}$, K, L,

$$\frac{\partial}{\partial K}\left[\frac{\partial \mathbf{u}_{\boldsymbol{\alpha}}}{\partial L}\right] = \frac{\partial}{\partial L}\left[\frac{\partial \mathbf{u}_{\boldsymbol{\alpha}}}{\partial K}\right] \equiv \frac{\partial^2 \mathbf{u}_{\boldsymbol{\alpha}}}{\partial K\, \partial L}.$$

This permits us to define unambiguously the n-th order derivative

$$\frac{\partial^n \mathbf{u}_{\boldsymbol{\alpha}}}{\partial \mathbf{K}} = \frac{\partial^n \mathbf{u}_{\boldsymbol{\alpha}}}{\partial K_1 \ldots \partial K_n},$$

which is symmetric in $\mathbf{K} = (K_1, \ldots, K_n)$. These lead directly to the definition of hyperjacobians.

DEFINITION 2.2. Given $\boldsymbol{\alpha}$, $\mathbf{K}$ as above, the n-th order *hyperjacobian* is the differential polynomial

$$J^{\boldsymbol{\alpha}}_{\mathbf{K}} = \frac{\partial^n \mathbf{u}_{\boldsymbol{\alpha}}}{\partial \mathbf{K}}, \tag{2.2}$$

where each r-tuple $\mathbf{u}_{\boldsymbol{\beta}}$ in the resulting sum is identified with the product of its entries.

EXAMPLE 2.3. For the first order case,

$$J^{\boldsymbol{\alpha}}_{K} = \frac{\partial(u_{\alpha_1}, \ldots, u_{\alpha_r})}{\partial(x_{k_1}, \ldots, x_{k_r})} = \sum_{\pi} (\operatorname{sign} \pi) \frac{\partial u_{\alpha_1}}{\partial x_{k_{\pi 1}}} \cdot \frac{\partial u_{\alpha_2}}{\partial x_{k_{\pi 2}}} \cdots \frac{\partial u_{\alpha_r}}{\partial x_{k_{\pi r}}}$$
$$= \det\left(\frac{\partial u_{\alpha_i}}{\partial x_{k_j}}\right).$$

Therefore, a first order hyperjacobian is nothing but an ordinary Jacobian determinant.

EXAMPLE 2.4. Now let $n = 2$. Second order hyperjacobians are symmetric functions of the dependent variables, and we can consider the special case when the arguments are all the same. Then,

$$\begin{aligned}\frac{\partial^2(u,\ldots,u)}{\partial K\,\partial L} &= \sum_\pi (\operatorname{sign}\pi)\frac{\partial(u_{k_{\pi 1}},\ldots,u_{k_{\pi r}})}{\partial L}\\ &= \sum_\pi\sum_\rho (\operatorname{sign}\pi)(\operatorname{sign}\rho)u_{k_{\pi 1}l_{\rho 1}}\ldots u_{k_{\pi r}l_{\rho r}}\\ &= r!\sum_\pi (\operatorname{sign}\pi)u_{k_1 l_{\pi 1}}\ldots u_{k_r l_{\pi r}}\\ &= r!\frac{\partial(u_{k_1},\ldots,u_{k_r})}{\partial(x_{l_1},\ldots,x_{l_r})} = r!\frac{\partial u_K}{\partial L}.\end{aligned}$$

(Here $u_i \equiv \partial u/\partial x_i$, etc.) Therefore, a Jacobian determinant whose arguments are all derivatives of a single dependent variable is actually a second order hyperjacobian in disguise. This will explain the observation in BCO that these Jacobians could be expressed as second order derivatives.

EXAMPLE 2.5. Let $p = r = 2$. The various hyperjacobians have the forms

$$\begin{aligned}\frac{\partial(u,v)}{\partial(x,y)} &= u_x v_y - u_y v_x,\\ \frac{\partial^2(u,v)}{\partial(x,y)^2} &= u_{xx}v_{yy} - 2u_{xy}v_{xy} + u_{yy}v_{xx},\\ \frac{\partial^3(u,v)}{\partial(x,y)^3} &= u_{xxx}v_{yyy} - 3u_{xxy}v_{xyy} + 3u_{xyy}v_{xxy} - u_{yyy}v_{xxx},\end{aligned}$$

and, by induction,

$$\frac{\partial^n(u,v)}{\partial(x,y)^n} = \sum_{i=0}^{n}(-1)^i\binom{n}{i}\frac{\partial^n u}{\partial x^{n-i}\,\partial y^i}\,\frac{\partial^n v}{\partial x^i\,\partial y^{n-i}}. \tag{2.3}$$

For polynomial functions u, v, the hyperjacobian (2.3) is known as the n-th *transvectant* of u and v, and can be found in Gordan, [**10**, p. 36], and Gurevich, [**12**, p. 227]. See section 3.3 for more details on this connection. Note that these are symmetric or alternating in u, v depending on whether n is even or odd. More generally, the following lemma can be easily proved by induction on n.

LEMMA 2.6. *The hyperjacobian $J^{\boldsymbol{\alpha}}_{\mathbf{K}}$ of order n is alternating or symmetric in* $\boldsymbol{\alpha}$ *depending on whether n is odd or even:*

$$J^{\pi(\boldsymbol{\alpha})}_{\mathbf{K}} = \begin{cases}(\operatorname{sign}\pi)J^{\boldsymbol{\alpha}}_{\mathbf{K}}, & n \text{ odd},\\ J^{\boldsymbol{\alpha}}_{\mathbf{K}}, & n \text{ even}.\end{cases}$$

2.3. Row expansion formulae

The above definition allows one to compute n-th order hyperjacobians of degree r recursively from the $(n-1)$-st order hyperjacobians of the same degree. An alternative procedure is to use the analogue of a row expansion formula for determinants expressing them in terms of n-th order hyperjacobians of degree $r-1$.

THEOREM 2.7. *Given* $\boldsymbol{\alpha}$, $\mathbf{K}$, *suppose* $1 \leqq \nu \leqq r$. *Then,*

$$\frac{\partial^n \mathbf{u}_{\boldsymbol{\alpha}}}{\partial \mathbf{K}} = (-1)^{n(\nu-1)} \sum_{J,\mathbf{L}} (\operatorname{sign} J)\, \partial_J u_{\alpha_\nu} \frac{\partial^n \mathbf{u}_{\boldsymbol{\alpha}_{\hat{\nu}}}}{\partial \mathbf{L}}, \tag{2.4}$$

the sum being over all $J = (j_1, \ldots, j_n)$, $1 \leqq j_1 \leqq \ldots \leqq j_n \leqq p$, *and all* $\mathbf{L} = (L_1, \ldots, L_n)$, $L_k = (l_1^k, \ldots, l_{r-1}^k)$, $1 \leqq l_1^k < \ldots < l_{r-1}^k \leqq p$, *such that* $K_k = \pi_k(j_k, l_1^k, \ldots, l_{r-1}^k)$ *for some permutation* π_k. *In formula* (2.4),

$$\operatorname{sign} J = \Pi \operatorname{sign}(\pi_k)$$

and

$$\partial_J = \partial_{j_1} \partial_{j_2} \ldots \partial_{j_n}.$$

Proof. This is done by induction on the order n. For $n = 1$, (2.4) is simply the expansion of a Jacobian determinant by its ν-th row.

Now suppose we have proved (2.4) for given n. Without loss of generality (see Lemma 2.6), set $\nu = 1$. Let $\boldsymbol{\beta} = (\alpha_2, \ldots, \alpha_r)$, $\alpha = \alpha_1$. Then, by (2.1), and the induction hypothesis,

$$\begin{aligned}\frac{\partial^{n+1} \mathbf{u}_{\boldsymbol{\alpha}}}{\partial K\, \partial \mathbf{K}} &= \sum_\rho \operatorname{sign} \rho \frac{\partial^n \mathbf{u}_{\boldsymbol{\alpha}, \rho(K)}}{\partial \mathbf{K}} \\ &= \sum_\rho \sum_{J,\mathbf{L}} \operatorname{sign} \rho \operatorname{sign} J\, \partial_i\, \partial_J u_\alpha \frac{\partial^n \mathbf{u}_{\boldsymbol{\beta}, L'_\rho}}{\partial \mathbf{L}},\end{aligned}$$

where $i = k_{\rho 1}$, $L'_\rho = (k_{\rho 2}, \ldots, k_{\rho r})$. Given K and ρ, let π be the permutation of $\{1, \ldots, r-1\}$ such that $\pi(L_\rho) = L'_\rho$, where $l_1 < l_2 < \ldots < l_{r-1}$. Then $\operatorname{sign} \rho = (\operatorname{sign} \pi)(\operatorname{sign} i)$, where $\operatorname{sign} i$ is the sign of the permutation π_0, with $\pi_0(i, l_1, ., l_{r-1}) = (k_1, \ldots, k_r)$. Thus,

$$\begin{aligned}\frac{\partial^{n+1} \mathbf{u}_{\boldsymbol{\alpha}}}{\partial K\, \partial \mathbf{K}} &= \sum_{i,J,\mathbf{L}} (\operatorname{sign} i)(\operatorname{sign} J)\, \partial_i\, \partial_J u_\alpha \sum_{\pi,L} \operatorname{sign} \pi \frac{\partial^n \mathbf{u}_{\boldsymbol{\beta},L}}{\partial \mathbf{L}} \\ &= \sum_{J',\mathbf{L}'} (\operatorname{sign} J')\, \partial_{J'} u_\alpha \cdot \frac{\partial^{n+1} \mathbf{u}_{\boldsymbol{\beta}}}{\partial \mathbf{L}'},\end{aligned}$$

where $J' = (i, j_1, \ldots, j_n)$, $\mathbf{L}' = (L, L_1, \ldots, L_n)$. This completes the induction.

EXAMPLE 2.8. Suppose $r = 2$. If $K = ((k_1^1, k_2^1), \ldots, (k_1^n, k_2^n))$, then (2.4) reads

$$\frac{\partial^n (u, v)}{\partial \mathbf{K}} = \sum_{\boldsymbol{\nu}} (-1)^{s(\boldsymbol{\nu})} \frac{\partial^n v}{\partial k_{\nu_1}^1 \ldots \partial k_{\nu_n}^n} \cdot \frac{\partial^n v}{\partial k_{\nu'_1}^1 \ldots \partial k_{\nu'_n}^n}, \tag{2.5}$$

where $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_n)$ ranges over all n-tuples with $\nu_j = 1$ or 2, and $\nu'_j = 3 - \nu_j$. Also, $s(\boldsymbol{\nu}) \equiv \#\{\nu_j = 2\} = \sum (\nu_j - 1)$. In particular, if $\mathbf{I} = ((1, 2), (1, 2), \ldots, (1, 2))$, we recover (2.3) by identifying $x_1 = x$, $x_2 = y$, since there are $\binom{n}{i}$ $\boldsymbol{\nu}$'s with i entries equal to 1 and $n - i$ entries equal to 2.

EXAMPLE 2.9. Suppose $r = 3 = p$. We compute the second and third order cases. By (2.4),

$$\frac{\partial^2(u, v, w)}{\partial(x, y, z)^2} = u_{xx} \frac{\partial^2(v, w)}{\partial(y, z)^2} - 2u_{xy} \frac{\partial^2(v, w)}{\partial(x, z)\, \partial(y, z)} + (\text{cyclic in } x, y, z),$$

the remaining terms being obtained by cyclically permuting the variables x, y, z. (However, any duplicate terms should be disregarded. Thus, we would write $u_{xx}v_{yy}-2u_{xy}v_{xy}+u_{yy}v_{xx}$ as $u_{xx}v_{yy}-2u_{xy}v_{xy}+(\text{cyclic in } x, y)$ in this notation.) These can be further expanded by (2.5), so

$$\frac{\partial^2(u, v, w)}{\partial(x, y, z)^2}=u_{xx}(v_{yy}w_{zz}+v_{zz}w_{yy}-2v_{yz}w_{yz})-2u_{xy}(v_{xy}w_{zz}+v_{zz}w_{xy}-v_{xz}w_{yz}-v_{yz}w_{xz})$$
$$+(\text{cyclic in } x, y, z).$$

Similar considerations give

$$\begin{aligned}\frac{\partial^3(u, v, w)}{\partial(x, y, z)^3}&=u_{xxx}\frac{\partial^3(v, w)}{\partial(y, z)^3}-3u_{xxy}\frac{\partial^3(v, w)}{\partial(y, z)^2\,\partial(x, z)}-6u_{xyz}\frac{\partial^3(v, w)}{\partial(x, y)\,\partial(x, z)\,\partial(y, z)}\\&\quad+(\text{cyclic in } x, y, z)\\&\quad=u_{xxx}(v_{yyy}w_{zzz}-3v_{yyz}w_{yzz}+3v_{yzz}w_{yyz}-v_{zzz}w_{yyy})\\&\quad-3u_{xxy}(v_{xxy}w_{zzz}-2v_{xyz}w_{yzz}-v_{yyz}w_{xzz}+v_{xzz}w_{yyz}+2v_{yzz}w_{xyz}-v_{zzz}w_{yyy})\\&\quad-6u_{xyz}(v_{xxy}w_{yzz}-v_{xxz}w_{yyz}+v_{xzz}w_{xyy}-v_{xyy}w_{xzz}+v_{yyz}w_{xxz}-v_{yzz}w_{xxy})\\&\quad+(\text{cyclic in } x, y, z)\\&=u_{xxx}v_{yyy}w_{zzz}-3u_{xxx}v_{yyz}w_{yzz}-6u_{xyz}v_{xxy}w_{yzz}+\dots,\end{aligned}\tag{2.6}$$

where in the last expression the remaining terms are obtained by cyclically permuting x, y, z *and* u, v, w, and multiplying by the sign of the permutation on u, v, w.

In fact, as can be seen in the above examples, a simple induction using the row expansion formula (2.4) yields a closed form expression for an n-th order hyperjacobian.

COROLLARY 2.10. *Let* $\boldsymbol{\alpha}$, $\mathbf{K}$ *be as above. Then,*

$$\frac{\partial^n \mathbf{u}_{\boldsymbol{\alpha}}}{\partial \mathbf{K}}=\sum_{\mathbf{J}}(\operatorname{sign}\mathbf{J})\,\partial_{J_1}u_{\alpha_1}\dots\partial_{J_r}u_{\alpha_r},\tag{2.7}$$

where the sum is over all $\mathbf{J}=(J_1,\dots,J_r)$, $J_\nu=(j_1^\nu,\dots,j_n^\nu)$, $1\leqq j_k^\nu\leqq p$, *such that*

$$K_k=\pi_k(j_k^1,\dots,j_k^r),$$

for some permutation π_k. *Finally,*

$$\operatorname{sign}\mathbf{J}=\Pi\operatorname{sign}\pi_k.$$

This, in (2.6), the coefficient of $u_{xyz}v_{xxy}w_{yzz}$ is -6, since there are exactly 6 ways of choosing the (unordered) triples (xyz), (xxy), (yzz) from the triples (xyz), (xyz), (xyz) with each of the former containing one symbol in each of the latter, and the permutations always have product of signs -1. For instance, in the case $J_1=(xyz)$, $J_2=(yxx)$, $J_3=(zzy)$, the permutations π_k have signs:

$$(xyz)=+(xyz),\quad (yxz)=-(xyz),\quad (zxy)=+(xyz).$$

2.4. Higher dimensional determinants

The above formulae for hyperjacobians can be rewritten by using Cayley's theory of higher dimensional determinants. Just as an ordinary Jacobian determinant is the determinant of a two-dimensional square matrix of first order partial derivatives of the arguments, so can an n-th order hyperjacobian be written as the "determinant" of an $(n+1)$-dimensional hypercubical array of n-th order partial derivatives. Our previous results on hyperjacobians thereby translate into known results in the theory of higher dimensional determinants. See Oldenburger, [**18**], and the references therein for a more complete exposition of this theory.

An m-*dimensional matrix* $\mathbf{A}$ of order (or size) r is a hypercubical array of r^m numbers a_I indexed by $I=(i_1,\ldots,i_m)$, $1\leqq i_\nu\leqq r$. The (*full signed*) *determinant* of $\mathbf{A}$ is the number

$$\det \mathbf{A} = \sum (\operatorname{sign} \mathbf{I}) a_{I_1} a_{I_2} \ldots a_{I_r}, \tag{2.8}$$

summed over all $\mathbf{I}=(I_1,\ldots,I_r)$, such that $I_j=(i_1^j,\ldots,i_m^j)$, $1\leqq i_\nu^j\leqq r$, $i_1^j=j$, and $i_\nu^j\neq i_\nu^k$ for $j\neq k$, $1\leqq\nu\leqq m$. For each such $\mathbf{I}$, sign $\mathbf{I}$ denotes the product of the signs of the permutations $(i_\nu^1,\ldots,i_\nu^r)$ of the integers $(1,\ldots,r)$.

The $(m-1)$-dimensional matrix obtained from $\mathbf{A}$ by fixing one of the indices $i_\nu=k$ is called the k-th ν-*layer* of $\mathbf{A}$, or layer for short. The layers just consist of those entries in $\mathbf{A}$ contained in a hyperplane orthogonal to one of the coordinate axes, generalizing the notion of row and column for an ordinary (two-dimensional) matrix. From (2.8), if a pair of ν-layers in $\mathbf{A}$ is interchanged, for m even, $\det \mathbf{A}$ always changes sign, whereas for m odd, $\det \mathbf{A}$ changes sign if $\nu\geqq 2$, but remains the same if $\nu=1$.

Given $\mathbf{J}=(J_1,\ldots,J_m)$, with $J_\nu=(j_\nu^1,\ldots,j_\nu^s)$, where $1\leqq s\leqq r$, $1\leqq j_\nu^i\leqq r$, and $j_\nu^i\neq j_\nu^k$ for $i\neq k$, define the $\mathbf{J}$-th *minor* of $\mathbf{A}$ to be the m-dimensional matrix $\mathbf{A}_\mathbf{J}$ of order s consisting of those entries a_I of $\mathbf{A}$ where, for $1\leqq\nu\leqq m$, $i_\nu=j_\nu^k$ for some $1\leqq k\leqq s$. In other words, $\mathbf{A}_\mathbf{J}$ consists of those entries of $\mathbf{A}$ lying on the ν-layers of $\mathbf{A}$ indexed by J_ν. The entries of $\mathbf{A}_\mathbf{J}$ are ordered so that the k-th ν-layer of $\mathbf{A}_\mathbf{J}$ consists of those entries of the j_ν^k-th ν-layer of $\mathbf{A}$ satisfying the above condition. (Note: if J_ν is not increasing, the ν-layers of $\mathbf{A}_\mathbf{J}$ will not be in the same order as those of $\mathbf{A}$.) More generally, we can consider an m-dimensional hyperrectangular matrix $\mathbf{B}$ of order $r_1\times\ldots\times r_m$ with entries b_I indexed by $I=(i_1,\ldots,i_m)$, $1\leqq i_\nu\leqq r_\nu$. The $\mathbf{J}$-th minor $\mathbf{B}_\mathbf{J}$ is hypercubical of order s, where $\mathbf{J}=(J_1,\ldots,J_m)$, $J_\nu=(j_\nu^1,\ldots,j_\nu^s)$, $1\leqq j_\nu^i\leqq r_\nu$, $j_\nu^i\neq j_\nu^k$ for $i\neq k$. Note that $s\leqq\min r_\nu$.

Suppose $u^1(x),\ldots,u^q(x)$ are smooth functions of $x_1,\ldots,x_p$. Given $n\geqq 1$, form the n-th order hyperjacobian matrix $\mathbf{U}^{(n)}$, which is defined to be the $(n+1)$-dimensional hyperrectangular matrix of order $q\times p\times\ldots\times p$ with entries $u_I^{(n)}=\partial_J u^i$ for $I=(i_1,\ldots,i_{n+1})$, $1\leqq i_1\leqq q$, $1\leqq i_\nu\leqq p$, $\nu\geqq 2$ and where $i=i_1$, $J=(i_2,\ldots,i_{n+1})$. The following result shows how n-th order hyperjacobians of the u^ν are given by determinants of suitable minors of the hyperjacobian matrix $U^{(n)}$.

THEOREM 2.11. *Given* $\mathbf{K}=(K_1,\ldots,K_p)$, $K_i=(k_1^i,\ldots,k_n^i)$, $1\leqq k_\nu^i\leqq p$, *and* $\boldsymbol{\alpha}=(\alpha_1,\ldots,\alpha_r)$, $\alpha_k=(0,i_k)$, $1\leqq i_k\leqq q$, *define* $\mathbf{J}=(J_1,\ldots,J_{n+1})$ *so that* $j_1^k=i_k$, $1\leqq k\leqq r$, $j_\nu^i=k_{\nu-1}^i$, $2\leqq\nu\leqq n+1$, $1\leqq i\leqq r$. *Then,*

$$\frac{\partial^n \mathbf{u}_{\boldsymbol{\alpha}}}{\partial \mathbf{K}} = \det \mathbf{U}_\mathbf{J}^{(n)}. \tag{2.9}$$

For more general $\boldsymbol{\alpha}$, (2.9) can clearly be generalized by forming the corresponding n-th order hyperjacobian matrix of the u_{α_i}. Note that $(J_2,\ldots,J_{n+1})$ is essentially the "transpose" of $(K_1,\ldots,K_r)$. For instance, if $\mathbf{K}=((1,2),(3,4),(5,6))$, then $(J_2,J_3)=((1,3,5),(2,4,6))$. The first multi-index J_1 tells which u^ν are the arguments in the hyperjacobian. The proof of (2.9) follows directly from a comparison of (2.7) and (2.8). The differential polynomials $\det \mathbf{U}_{\mathbf{J}}^{(n)}$ were first written down by Escherich, [**7**], and Gegenbauer, [**9**], where they proved that for homogeneous polynomial functions $u^1, \ldots, u^q$ these are covariants under the action of the general linear group $GL(p)$. Apparently they never noticed that $\det \mathbf{U}_{\mathbf{J}}^{(n)}$ could be expressed as an n-th order divergence.

In light of (2.9), our preceding results can be rederived from the elementary properties of higher dimensional determinants. In particular, the fact that an n-th order hyperjacobian is a symmetric or alternating function of its arguments depending on whether n is even or odd reflects the corresponding behaviour of an $(n+1)$-dimensional determinant under interchange of 1-layers. The remark on second order hyperjacobians in Example 2.4 arises from a theorem relating cubical determinants with identical 1-layers to an ordinary determinant of the given layer, [**18**, Theorem 9]. I have been unable to locate any further work on these determinants of hyperjacobian matrices after Escherich and Gegenbauer, although their work was largely subsumed by the invariant-theoretic notion of transvectant, discussed in section 3.3.

3. Transform theory and determinantal ideals

For the study of homogeneous null Lagrangians, BCO introduced a transform for differential polynomials inspired by work of Gel'fand and Dikii, [**8**], and Shakiban, [**22**], on the formal variational calculus. By analogy with the Fourier transform of classical analysis, this transform has the useful property of changing differential operations to algebraic operations, thereby enabling questions about differential polynomials to be attacked by the powerful techniques of algebraic geometry. Here we employ transform techniques to prove two fundamental results on hyperjacobians. First, each n-th order hyperjacobian is an n-th order divergence, i.e. satisfies an identity of the form (1.4). Secondly, any continuous function L of the k-th order derivatives of u for some fixed k which is an n-th order divergence is necessarily an affine combination of n-th order hyperjacobians with appropriate arguments. This provides a complete solution to the problem of classifying homogeneous null Lagrangians with nice identities. The connection between the transform and Aronhold's symbolic method in classical invariant theory is discussed in subsection 3.3.

3.1. The transform and derivatives

The first task is to investigate the transform introduced in BCO, whose notation we use without further comment, in more detail. Recall that the transform $\mathscr{F}$ provides a linear isomorphism between $\mathscr{L}^r$, the space of homogeneous differential polynomials of degree r, and Z_0^r, the space of symmetric algebraic polynomials $\varphi(a^1, b^1; \ldots; a^r, b^r)$, $a^\nu \in \mathbb{R}^q$, $b^\nu \in \mathbb{R}^p$, which are linear in the a^ν. Here, Z_0^r is simply the direct sum of the spaces $Z_0^{r,k}$ of polynomials linear in the a^ν and homogeneous of degree k in the b^ν considered in BCO. If we identify $\mathbf{u}_{\boldsymbol{\alpha}}$ with the product of its entries, as in section 2.2, then $\mathscr{F}$ has the explicit representation

$$\mathscr{F}(\mathbf{u}_{\boldsymbol{\alpha}}) = \sigma[(\mathbf{a}\otimes\mathbf{b})_{\boldsymbol{\alpha}}] = \frac{1}{r!}\sum (\mathbf{a}\otimes\mathbf{b})_{\pi(\boldsymbol{\alpha})},$$

where, for $\boldsymbol{\alpha}=((I_1,\nu_1),\ldots,(I_r,\nu_r))$, $(\mathbf{a}\otimes\mathbf{b})_{\boldsymbol{\alpha}}\equiv a^1_{\nu_1}b^1_{I_1}a^2_{\nu_2}\ldots b^{r-1}_{I_{r-1}}a^r_{\nu_r}b^r_{I_r}$; $b_I=b^I$ as in section 2, and the sum is over all permutations π of $\{1,\ldots,r\}$.

Given a linear map $G:\mathscr{L}^r\to\mathscr{L}^s$, there is an induced linear map $\hat{G}: Z^r_0\to Z^s_0$, called the *transform* of G, defined by

$$\hat{G}[\mathscr{F}(P)]=\mathscr{F}[G(P)],\quad P\in\mathscr{L}^r.$$

We need to find the transform of the total derivative operator; the proof is elementary, cf. [**22**].

LEMMA 3.1. *If* $1\leqq i\leqq p$, $\varphi\in Z^r_0$, *then*

$$\widehat{D_i}(\varphi)=(b^1_i+\ldots+b^r_i)\varphi. \tag{3.1}$$

More generally, for any multi-index I,

$$\widehat{D^I}(\varphi)=(b^1+\ldots+b^r)^I\varphi. \tag{3.1$'$}$$

We now consider the inverse problem of determining when a differential polynomial is an n-th order derivative. Our criterion will be in the transform space Z^r_0, but, as shall be indicated, this can easily be restated so as not to rely on the transform itself. Given $\alpha=(J,\nu)$, define the partial derivative $\partial_\alpha=\partial^{m+1}/\partial a^r_\nu(\partial b^r_1)^{i_1}\ldots(\partial b^r_p)^{i_p}$, acting on Z^r_0. Given $\varphi\in Z^r_0$, let $\varphi|_0\in Z^{r-1}_0$ denote the polynomial

$$\varphi|_0\equiv\varphi(a^1,b^1;\ldots;a^{r-1},b^{r-1};0,-b^1-b^2-\ldots-b^{r-1}).$$

LEMMA 3.2. *Let* $P\in\mathscr{L}^r$. *Then* P *is an* n*-th order divergence if and only if*

$$\partial_\alpha(\mathscr{F}P)|_0=0, \tag{3.2}$$

for all $\alpha=(J,\nu)$, *with* $|J|\leqq n-1$, $1\leqq\nu\leqq q$.

Proof. First recall a calculus lemma. Let $\varphi(y)$, $y=(y_1,\ldots,y_p)$, be a polynomial and let $c=(c_1,\ldots,c_p)$ be constant. Then φ can be written as a sum of terms, each of degree $\geqslant n$ in the monomials y_i-c_i if and only if

$$\left.\frac{\partial^J\varphi}{\partial y^J}\right|_{y=c}=0,$$

for all J with $|J|\leqq n-1$. By identifying y_i with b^r_i and c_i with $-b^1_i-\ldots-b^{r-1}_i$, and using (3.1$'$), we infer that P is an n-th order derivative if and only if

$$\frac{\partial^J}{(\partial b^r)^J}(\mathscr{F}P)|_{b^r=-b^1-\ldots-b^{r-1}}=0,\quad |J|\leqq n-1.$$

Finally, recall that $\mathscr{F}P$ is linear in a^r to complete the proof.

The counterpart of the operator $\partial_\alpha|_0$ on $\mathscr{L}^r$ is the higher order Euler operators introduced and studied independently in [**2**] and [**19**]. These are given by

$$E^\alpha=E^\nu_J\equiv\sum_{K\geqq J}\binom{K}{J}(-D)^{K-J}\frac{\partial}{\partial u^\nu_K},$$

and it can be shown, [21], that

$$\widehat{E^\alpha} = \frac{r}{J!}\partial_\alpha|_0.$$

Thus, we have proved the following:

LEMMA 3.3. *A differential polynomial P is an n-th order divergence if and only if $E^\alpha(P) = 0$ for all $\alpha = (J, \nu)$ with $|J| \leqq n-1$.*

We will not use this result here.

3.2. Transform of hyperjacobians

We now compute the transform of a hyperjacobian. Recall that if $C = (c^i_j)$ is an $r \times r$ matrix, the *permanant* of C is defined as

$$\operatorname{perm} C = \sum c^1_{\pi 1} c^2_{\pi 2} \ldots c^r_{\pi r},$$

the sum being over all permutations π of $\{1, \ldots, r\}$. Next, recall some notation from BCO. Given K, let B_K denote the $r \times r$ submatrix with entries $b^i_{k_j}$. Given $\boldsymbol{\alpha}$, $(A \otimes B)_{\boldsymbol{\alpha}}$ denotes the $r \times r$ matrix with entries $a^i_{\nu_j} b^i_{I_j}$. Note that $\operatorname{perm}(A \otimes B)_{\boldsymbol{\alpha}}$ is just $r!$ times the transform of the monomial $\mathbf{u}_{\boldsymbol{\alpha}}$.

LEMMA 3.4. *The transform of the n-th order hyperjacobian $J^{\boldsymbol{\alpha}}_{\mathbf{K}}$ is given by*

$$\mathcal{F}[J^{\boldsymbol{\alpha}}_{\mathbf{K}}] = \prod_{\nu=1}^{n} \det(B_{K_\nu}) \cdot \begin{Bmatrix} \det \\ \operatorname{perm} \end{Bmatrix} (A \otimes B)_{\boldsymbol{\alpha}}, \tag{3.3}$$

where det *is used for n odd, and* perm *for n even.*

Proof. Lemma 4.5 of BCO proves (3.3) for an ordinary Jacobian, and we use induction on n to prove it in general. Given $\mathbf{K}$, set $K = K_n$. By transforming the definition (2.2) of $J^{\boldsymbol{\alpha}}_{\mathbf{K}}$, and using (2.1) and the induction hypothesis, we find

$$\mathcal{F}[J^{\boldsymbol{\alpha}}_{\mathbf{K}}] = \sum_\pi (\operatorname{sign} \pi) \prod_{\nu=1}^{n-1} \det(B_{K_\nu}) \begin{Bmatrix} \operatorname{perm} \\ \det \end{Bmatrix} (A \otimes B)_{\boldsymbol{\alpha}, \pi(K)},$$

where throughout the proof the upper line in the brackets refers to n odd and the lower to n even. It thus suffices to prove the identity

$$\sum_\pi (\operatorname{sign} \pi) \begin{Bmatrix} \operatorname{perm} \\ \det \end{Bmatrix} (A \otimes B)_{\boldsymbol{\alpha}, \pi(K)} = \det(B_K) \begin{Bmatrix} \det \\ \operatorname{perm} \end{Bmatrix} (A \otimes B)_{\boldsymbol{\alpha}}.$$

Let $\boldsymbol{\alpha} = ((I_1, \nu_1), \ldots, (I_r, \nu_r))$; then, on expanding both sides of the previous formula,

$$\begin{aligned} &\sum_\pi (\operatorname{sign} \pi) \sum_\rho \begin{Bmatrix} 1 \\ \operatorname{sign} \rho \end{Bmatrix} \prod_{i=1}^{r} a^i_{\nu_{\rho i}} b^i_{I_{\rho i}} b^i_{k_{\rho \pi i}} \\ &\quad = \left\{ \sum_\pi (\operatorname{sign} \pi) \prod_{i=1}^{r} b^i_{k_{\pi i}} \right\} \times \left\{ \sum_\rho \begin{Bmatrix} \operatorname{sign} \rho \\ 1 \end{Bmatrix} \prod_{i=1}^{r} a^i_{\nu_{\rho i}} b^i_{I_{\rho i}} \right\}. \end{aligned}$$

These expressions are clearly equal.

3.3. Transvectants and the symbolic method

For homogeneous polynomial functions, the transform presented in BCO is equivalent to the powerful symbolic method of Aronhold in classical invariant theory. Here this connection will be briefly indicated, but lack of space precludes a more detailed exposition, and we refer the reader to **[10]**, **[11]**, **[12]**, for detailed developments of the symbolic method and its applications to computing covariants. Once the connection has been made, it is easy to see how transvectants arise as special types of hyperjacobians with polynomial arguments.

An m-th order homogeneous polynomial or *form* (in classical terminology)

$$u(x)=\sum b_I x^I, \quad x\in\mathbb{R}^p,$$

summed over $I=(i_1,\ldots,i_m)$, $1\leqq i_\nu\leqq p$, is represented by its coefficients

$$b_1=\frac{1}{m!}\partial^I u(x),$$

or, equivalently, all its m-th order partial derivatives. The *symbolic expression* for u is

$$u(x)\simeq(bx)^m,$$

where $bx=b_1x_1+\ldots+b_px_p$. The key to the symbolic method lies in the identification of the coefficient b_I with the product $b^I=b_1^{i_1}\ldots b_m^{i_m}$ in the m-th power of bx.

Given several forms of various degrees

$$u^\nu(x)=\sum b_I x^I=(b^\nu x)^{m_\nu}, \quad \nu=1,\ldots,r,$$

consider an algebraically homogeneous differential polynomial of the special form

$$P(u^1,\ldots,u^r)=\sum C_{\boldsymbol\alpha} u_{\boldsymbol\alpha}\in\mathscr{L}^r,$$

summed over $\boldsymbol\alpha=(\alpha_1,\ldots,\alpha_r)$, where $\alpha_j=(I_j,j)$. (In other words, each u^ν occurs precisely once in each monomial in P.) Clearly,

$$P\simeq\sum A_J x^J,$$

where each coefficient A_J is a polynomial in the coefficients b_I^ν of the forms. The *symbolic expression* for P is obtained by replacing b_I^ν by the product $(b^\nu)^I$, wherever it occurs in A_J. For example, the symbolic expression for a Jacobian determinant is

$$\frac{\partial(u^1,\ldots,u^r)}{\partial(x_1,\ldots,x_r)}\sim(\det B)(b^1x)^{m_1-1}\ldots(b^rx)^{m_r-1},$$

where $x=(x_1,\ldots,x_r)$ and B is the $r\times r$ matrix with entries b_i^ν, cf. [**12**, p. 192]. The essential equivalence between the symbolic expression for P and its transform can be easily seen.

THEOREM 3.5. *Let $P\in\mathscr{L}^r$ be as described above, and let $\mathscr{F}(P)$ be its transform. The symbolic expression for P is obtained by replacing each monomial*

$$a_{\nu_1}^1 b_{I_1}^1\ldots a_{\nu_r}^r b_{I_r}^r$$

in $\mathscr{F}(P)$ by the product

$$(b^{\nu_1}_{I_1} \dots b^{\nu_r}_{I_r})(b^{\nu_1}x)^{m_1 - i_1} \dots (b^{\nu_r}x)^{m_r - i_r},$$

where $i_j = |I_j|$.

More generally, if P has monomials in which some u^ν occurs more than once, a different symbolic expression for u^ν must be used for each occurrence. An analogous result holds, but is more complicated to formulate explicitly, and thus is left to the reader. In this case, the transform seems to offer slight advantages over the symbolic method. (Of course, our transform is more widely applicable than the symbolic method, being not restricted to just polynomial functions.)

Of particular importance for invariant theory are differential polynomials P left unchanged (up to a factor) under the action of the general linear group $GL(p)$; these are called *covariants*. A particular covariant is the n-th *transvectant*, which has symbolic expression

$$(u^1, \dots, u^r)^{(n)} = (\det B)^n (b^1 x)^{m_1 - n} \dots (b^r x)^{m_r - n},$$

using the same notation as in the above Jacobian determinant. As an immediate consequence of Theorem 3.5 we find:

COROLLARY 3.6. *The n-th transvectant of $u^1(x), \dots, u^r(x)$, $x = (x_1, \dots, x_r)$ is the special n-th order hyperjacobian*

$$(u^1, \dots, u^r)^{(n)} = \partial^n (u^1, \dots, u^r)/\partial(x_1, \dots, x_r)^n.$$

This explains the agreement between our hyperjacobian examples and the formulae for transvectants in [**10**], [**11**], [**12**]. Moreover, the general formula (2.7) for a hyperjacobian, or its equivalent higher order determinantal form (2.9), give an explicit non-symbolic representation of a general transvectant. I have been unable to find this general formula in the invariant theory literature, nor have I found any explicit connection between transvectants and multi-dimensional determinants. Moreover, the key result that an n-th order transvectant is an n-th order divergence does not appear in the invariant theory literature either. It would be interesting to pursue these links in more detail.

In the special case when all the arguments are the same, the n-th transvectant

$$(u, \dots, u)^{(n)} = \partial^n (u, \dots, u)/\partial(x_1, \dots, x_r)^n,$$

n necessarily even, was orginially considered by Cayley in his presymbolic theory of "hyperdeterminants", cf. [**11**, p. 84]. This justifies our introduction of the name "hyperjacobian" for our more general expression, rather than retaining the classical invariant theoretic term "transvectant", which is restricted to hyperjacobians of the special type in Corollary 3.6, with polynomial arguments.

3.4. Hyperjacobians as derivatives

The proof of the main result of this section, that n-th order divergences and n-th order hyperjacobians are essentially the same objects, parallels the proof of Theorem 4.1 in BCO identifying ordinary Jacobians and first order divergences. The relevant results from algebraic geometry, though, have only recently been proved.

THEOREM 3.7. *Let P be a function depending (smoothly) exclusively on the k-th order partial derivatives of u. Then P is an n-th order divergence if and only if P is an affine combination of n-th hyperjacobians whose arguments are $(k-n)$-th order derivatives of u.*

We first prove that each n-th order Jacobian is indeed an n-th order divergence. (This result is not as obvious as the fact that each Jacobian determinant is a divergence!). On expanding det B_K by its j-th row, we find

$$\det B_K = \sum_{\nu=1}^{r} b^j_{k_\nu} \det B_K^{jk_\nu},$$

where B_K^{jk} denotes the minor obtained from B_K by deleting the j-th row and the column corresponding to k in K. Also, for $i \neq j$, a similar expansion yields

$$0 = \sum_{\nu=1}^{r} b^i_{k_\nu} \det(B_K^{jk_\nu}).$$

By adding these formulae together, and using the expression (3.1) for the transform of the total derivative, we see that

$$\begin{aligned}\det(B_K) &= \sum_{\nu=1}^{r} \widehat{D_{k_\nu}} \det(B_K^{jk_\nu})\\ &= \sum_{i=1}^{p} \widehat{D_i} \det(B_K^{ji}),\end{aligned}$$

where, by convention, $B_K^{ji}=0$, if i does not appear in K. Thus, from Lemma 3.4 and the algebraic nature of the $\widehat{D_i}$,

$$\begin{aligned}\mathscr{F}(J_{\mathbf{K}}^{\boldsymbol{\alpha}}) &= \prod_{\nu=1}^{n}\left[\sum_{i=1}^{p} \widehat{D_i} \det B_{K_\nu}^{j_\nu i}\right] \times \begin{Bmatrix}\det\\ \text{perm}\end{Bmatrix}(A\otimes B)_{\boldsymbol{\alpha}},\\ &= \sum \widehat{D^I}\left[\frac{1}{I!}\sum_{\pi}\prod_{\nu=1}^{n} \det B_{K_\nu}^{j_\nu i_{\pi\nu}} \times \begin{Bmatrix}\det\\ \text{perm}\end{Bmatrix}(A\otimes B)_{\boldsymbol{\alpha}}\right]\\ &\equiv \sum \widehat{D^I}\widehat{Q_I},\end{aligned} \tag{3.4}$$

where $1 \leqq j_1, \ldots, j_n \leqq r$ can be chosen arbitrarily.

The fact that $\mathscr{F}$ is an isomorphism, and Lemma 3.1, imply that

$$J_{\mathbf{K}}^{\boldsymbol{\alpha}} = \sum D^I Q_I \tag{3.5}$$

for certain differential polynomials Q_I, which are defined by their transforms according to the previous formula, (3.4). Thus, each n-th order hyperjacobian is an n-th order derivative. The explicit form of Q_I is extremely complicated in general.

We now turn to a proof of the converse proposition, that every homogeneous n-th order divergence is an affine combination of hyperjacobians. Theorem 3.4, (vii) of BCO shows that all such Lagrangians are polynomials, and clearly we can restrict our attention to differential polynomials which are homogeneous of degree r and depend exclusively on k-th order derivatives. These all transform into the homogeneous space $Z_0^{r,k}$, and we first require a slight strengthing of

Lemma 3.2 for this space. Given $\beta=(i_1, j_1; \ldots; i_m, j_m)$ with $1 \leqq i_\nu \leqq p$, $1 \leqq j_\nu \leqq r$, let $\#\beta = m$, and $\partial_\beta = \partial^m/\partial b^{j_1}_{i_1} \ldots \partial b^{j_m}_{i_m}$.

LEMMA 3.8. *Let P be an r-th degree differential polynomial depending only on k-th order derivatives of u. Then P is an n-th order derivative if and only if, for all β with $\#\beta \leqq n-1$,*

$$\partial_\beta(\mathscr{F}P)[a^1, b^1; \ldots; a^r, b^r]=0, \tag{3.6}$$

whenever $b^1, \ldots, b^r$ are linearly dependent.

If $n=1$, Lemma 3.8 is just a restatement of condition iv) of Theorem 3.4 of BCO; thus the following proof specializes to a new proof of this condition, which, as discussed in BCO, originally arose in work of Murat and Tartar on compensated compactness.

Proof of Lemma 3.8. First note that conditions (3.2) of Lemma 3.2 can be replaced by equivalent conditions

$$\partial_\beta(\mathscr{F}P)=0, \tag{3.6$'$}$$

whenever $b^1+\ldots+b^r=0$ for all β with $\#\beta \leqq n-1$. Indeed, this includes the indices $(i_1, r; \ldots; i_m, r)$, so (3.6′) includes (3.2). It is easy to check that all the superfluous extra conditions in (3.6′) must also hold when (3.2) holds. Now, since $\mathscr{F}P$ is homogeneous of degree k in the b^ν,

$$\mathscr{F}P[a^1, \lambda_1 b^1; \ldots; a^r, \lambda_r b^r]=(\lambda_1\lambda_2\ldots\lambda_r)^k \cdot \mathscr{F}P[a^1, b^1; \ldots; a^r, b^r],$$

whenever $\lambda_1, \ldots, \lambda_r \in \mathbb{R}$. Differentiation proves that (3.6′) must also hold whenever $\lambda_1 b^1+\ldots+\lambda_r b^r=0$, thus proving the lemma.

Let $\mathscr{I}$ be the polynomial ideal introduced in BCO; $\mathscr{I}$ is generated by all the $r\times r$ minors of the $r\times p$ matrix $B=(b^i_j)$. A polynomial $\varphi(b^1, \ldots, b^r)$ is said to *vanish* on $\mathscr{I}$ if $\varphi(b^1, \ldots, b^r)=0$ whenever $b^1, \ldots, b^r$ are linearly dependent. (In algebraic geometry, φ vanishing on $\mathscr{I}$ means that $\varphi=0$ on the variety defined by the members of $\mathscr{I}$.) In BCO, we needed the result that $\mathscr{I}$ is a prime ideal, and hence φ vanishes on $\mathscr{I}$ if and only if φ belongs to $\mathscr{I}$, so $\varphi=\sum \psi_K \det B_K$ for certain polynomials ψ_K. Here we require a stronger result characterizing those polynomials which, together with all their partial derivatives of order less than n, vanish on $\mathscr{I}$. The ideal of such polynomials is known as the n-th *symbolic power* of $\mathscr{I}$, denoted by $\mathscr{I}^{(n)}$. The key result has fortunately been recently established, and states that the symbolic n-th power of $\mathscr{I}$ is the same as the algebraic n-th power of $\mathscr{I}$, which is the ideal generated by all n-fold products of elements of $\mathscr{I}$. Hochster, **[13]**, proved this in the case $p=r+1$; subsequently Trung, **[24]**, proved the result for general p, and DeConcini, Eisenbud and Procesi, **[6]**, have generalized this result to ideals generated by subdeterminants of arbitrary size. In this latter case, it is *not* true in general that the symbolic n-th power equals the actual n-th power.

THEOREM 3.9. (Hochster–Trung–DeConcini–Eisenbud–Procesi). *Let $\mathscr{I}$ be the ideal generated by the $r\times r$ minors* $\det B_K$ *of an $r\times p$ matrix of indeterminants B $(r\leqq p)$. For each positive integer n, the symbolic n-th power of $\mathscr{I}$ is equal to the* *algebraic n-th power of $\mathcal{I}$. In other words, for any polynomial φ,*

$$\partial_\beta \varphi(b^1, \ldots, b^r) = 0,$$

whenever $b^1, \ldots, b^r$ are linearly independent for all β with $\#\beta \leqq n-1$ if and only if

$$\varphi = \sum \psi_{\mathbf{K}} \cdot \prod_{\nu=1}^{n} \det(B_{K_\nu}), \tag{3.7}$$

suitable polynomials $\psi_{\mathbf{K}}$.

This theorem, in conjunction with Lemma 3.8, proves that P is an n-th order divergence if and only if its transform $\mathcal{F}P = \varphi$ is of the form (3.7) (where the $\psi_{\mathbf{K}}$ can now depend on the a^ν). Thus, to complete the proof of Theorem 3.9, we only need to show that the polynomials $\psi_{\mathbf{K}}$ can be taken to be linear combinations of determinants or permanents of the matrices $(A \otimes B)_{\boldsymbol{\alpha}}$ depending on the parity of n. This can be deduced from the fact that $\mathcal{F}P$ must be a symmetric function of its arguments, so Lemma 4.9 of BCO when n is odd and a similar lemma when n is even, coupled with the remaining arguments in the proof of Theorem 4.1 of BCO, can be utilized here to complete the proof of Theorem 3.9.

3.5. Quadratic p-relations

Any attempt to try to compute the number of linearly independent hyperjacobians is hampered, as for ordinary Jacobians, by the appearance of non-trivial linear relations amongst the hyperjacobians. These arise from the *quadratic p-relations* between products of pairs of determinants in the transform space.

THEOREM 3.10. *Suppose $\varphi(\ldots, y_K, \ldots)$, is a polynomial such that $\varphi(\ldots, \det B_K, \ldots)$ vanishes identically. Then φ is in the ideal generated by the quadratic p-relations, which are*

$$p_{KL} = \sum_{\nu=0}^{r} (-1)^\nu y_{K_{\hat\nu}} y_{k_\nu, L} = 0$$

for all $K = (k_0, \ldots, k_r)$, $L = (l_1, \ldots, l_{r-1})$.

For a proof see Hodge and Pedoe, [**14**]. Thus, all the relations amongst polynomials in the det B_K are derivable from the identities

$$\sum_{\nu=0}^{r} (-1)^\nu \det B_{K_{\hat\nu}} \det B_{k_\nu, L} = 0.$$

COROLLARY 3.11. *For each set of multi-indices $K = (k_0, \ldots, k_r)$, $L = (l_1, \ldots, l_{r-1})$, $\mathbf{K} = (k_3, \ldots, k_n)$ the relation*

$$\sum_{\nu=0}^{r} (-1)^\nu \frac{\partial^n \mathbf{u}_{\boldsymbol{\alpha}}}{\partial K_{\hat\nu}\, \partial(k_\nu, L)\, \partial \mathbf{K}} = 0$$

holds.

The proof is by transform. I suspect that these are all the relations satisfied by hyperjacobians, which means any other relation amongst n-th order hyperjacobians must either be a linear combination of the above relations, or derivable from a similar relation amongst m-th order hyperjacobians for some $m > n$. I do not have a proof of this conjecture, or know how to count the number of independent relations and so get a fix on the dimension of the space of hyperjacobians.

3.6. Divergence formulae and distributing derivatives

We now know that if $J^{\boldsymbol{\alpha}}_{\mathbf{K}}$ is an n-th order hyperjacobian, there is a nice identity of the form

$$J^{\boldsymbol{\alpha}}_{\mathbf{K}} = \sum_{|I|=n} D^I Q_I \tag{3.8}$$

for certain differential polynomials Q_I, which can, in principle, be constructed from (3.4). The explicit construction of the Q_I is an extremely difficult computational problem in general. It can, however, be implemented directly using a new theory of higher order differential forms based on the theory of Schur functors, which is developed in the paper [**20**]. Here we just present a couple of examples.

EXAMPLE 3.12. Let $r=2$, and consider the n-th order hyperjacobians $\partial^n(u, v)/\partial(x, y)^n$ given in Example 2.5. The n-th order divergence identities are

$$\frac{\partial^n(u, v)}{\partial(x, y)^n} = \sum_{i=0}^{n} D_x^{n-i} D_y^i \left\{ (-1)^{m+i+1} \sum_j \binom{m}{j}\binom{m+1}{i-j} \frac{\partial^m u}{\partial x^j\, \partial y^{m-j}} \frac{\partial^{m+1} v}{\partial x^{i-j}\, \partial y^{m+1+j-1}} \right\},$$

when $n = 2m+1$ is odd, and

$$\frac{\partial^n(u, v)}{\partial(x, y)^n} = \sum_{i=0}^{n} D_x^{n-i} D_y^i \left\{ (-1)^{m+i} \sum_j \binom{m}{j}\binom{m}{i-j} \frac{\partial^m u}{\partial x^j\, \partial y^{m-j}} \frac{\partial^m v}{\partial x^{i-j}\, \partial y^{m+j-1}} \right\},$$

when $n = 2m$ is even. (In both cases, the second sum is over all j such that the binomial coefficients are well defined.) These can be checked directly without too much difficulty.

EXAMPLE 3.13. Consider the third order hyperjacobian $\partial^3(u, v, w)/\partial(x, y, z)^3$ given in Example 2.9. To write down the identity, we use the notation

$$\frac{\partial(u, v, w)}{\partial(xx, xz, yz)} = \det\begin{pmatrix} u_{xx} & u_{xz} & u_{yz} \\ v_{xx} & v_{xz} & v_{yz} \\ w_{xx} & w_{xz} & w_{yz} \end{pmatrix},$$

(and similarly for other triples of pairs of independent variables) for Hessian-type determinants. Then the identity is

$$\begin{aligned} \frac{\partial^3(u, v, w)}{\partial(x, y, z)^3} = {}& -D_z^3 \frac{\partial(u, v, w)}{\partial(xx, xy, yy)} + D_z^2 D_y \left[2\frac{\partial(u, v, w)}{\partial(xx, xy, yz)} - \frac{\partial(u, v, w)}{\partial(xx, yy, xz)} \right] \\ & + D_x D_y D_z \left[\frac{\partial(u, v, w)}{\partial(xx, yy, zz)} + 4\frac{\partial(u, v, w)}{\partial(xy, xz, yz)} \right] + (\text{cyclic in } x, y, z). \end{aligned} \tag{3.9}$$

Attempts to derive, or even to verify, this rather pretty identity should be sufficient motivation for the reader to look at the theory of differential hyperforms, [**20**], from which it can be written down! Higher order cases can also be treated by this method with a minimum of computational distress.

The reader should note that in all the above identities, the various partial derivatives in the relevant Q_I's are distributed among the dependent variables as evenly as possible. For instance, in (3.9), all the u, v, w's appear with second order derivatives in the Hessian-type determinants therein. For the purposes of reducing growth conditions in the variational problems, we need to know this can always be done.

THEOREM 3.14. *Let $n = sr + t$, where s and $0 \leqq t < r$ are integers. Let $\boldsymbol{\alpha} = (i_1, I_1; \ldots; i_r, I_r)$ be homogeneous, i.e. $|I_\nu| = k$ for all ν. Then the hyperjacobian $J^{\boldsymbol{\alpha}}_{\mathbf{K}}$ has a nice identity of the form* (3.8) *with each Q_I being a linear combination of monomials each of which is a product of t $(k+n-s-1)$-st order derivatives and $r-t$ $(k+n-s)$-th order derivatives of the u^{i_ν}.*

To prove this result, it suffices to choose the indices $j_1, \ldots, j_n$ in (3.4) judiciously so that each $\hat{Q}_I$ is of degree $k+n-s-1$ in $b^1, \ldots, b^t$ and degree $k+n-s$ in $b^{t+1}, \ldots, b^r$. This can clearly be done.

4. Weak solutions to variational problems

The hyperjacobian identities allow one to define these special differential polynomials for functions in lower order Sobolev spaces than might ordinarily be expected. The method is a straightforward application of the Sobolev embedding theorem. Roughly speaking, an n-th order hyperjacobian of degree r can be defined for functions with n/r fewer derivatives than usual. (Of course, when r does not evenly divide n, the exact numerology is a little more complicated.) A slight strengthening of this condition allows one to use a compact Sobolev embedding and thereby prove weak continuity results for hyperjacobians of the type discussed in BCO. This in turn can be used to prove the existence of weak minimizers for certain special kinds of quasi-convex variational problems which could prove of interest in non-linear elasticity. One could attempt to write out a general theorem on the types of convexity conditions required to prove the existence of weak minimizers, but the rather special nature of the variational problems involved makes this kind of general result of minimal practical import. Rather, I have chosen to illustrate the kind of existence results obtainable with a couple of examples of interest.

4.1. Weak definition and continuity properties of hyperjacobians

Throughout this section, $\Omega \subset X = \mathbb{R}^p$ will be a bounded, connected open domain whose boundary $\partial\Omega$ is strongly Lipschitz, cf. BCO, although some of the results will hold under somewhat weaker hypotheses. Let $W^{l,\gamma} = W^{l,\gamma}(\Omega)$ denote the Sobolev space of (equivalence classes of) functions $u : \Omega \to U = \mathbb{R}^q$ with generalized l-th order derivatives in $L^\gamma(\Omega)$. We first determine on exactly which Sobolev spaces a hyperjacobian can be defined.

THEOREM 4.1. *Let $n = sr + t$, $0 \leqq t < r$. The n-th order hyperjacobian $J^{\boldsymbol{\alpha}}_{\mathbf{K}}(u)$ of degree r can be defined as a distribution provided $u \in W^{k+n-s,\gamma}$, where*

$$\gamma \geqq \max\left\{\frac{pr}{p+t}, r-t\right\}, \tag{4.1}$$

unless $r=p+t$ and $t>0$, in which case only the strict inequality in (4.1) *is allowed. When this hyperjacobian is defined distributionally rather than classically, it will be denoted as* $\widetilde{J^{\alpha}_{\mathbf{K}}(u)}$.

Note that in general to define an r-th degree differential polynomial which, like $J^{\alpha}_{\mathbf{K}}$, depnds on $(n+k)$-th order derivatives of u, u should be in the Sobolev space $W^{k+n,r}$. Thus, for an n-th order hyperjacobian, u needs s fewer derivatives, where s is the greatest integer $\leqq n/r$. When n is not evenly divisible by r, so $t>0$, further weakening of the requirements on u is reflected in the less stringent conditions on the exponent γ of the Sobolev space.

Proof of Theorem 4.1. To define $J^{\alpha}_{\mathbf{K}}$ as a distribution, it suffices to require that each of the polynomials Q_I in the identity (3.8) be in $L^1_{\mathrm{loc}}(\Omega)$. By theorem 3.14, each monomial in the Q_I is a product of t $(k+n-s-1)$-st order derivatives and $r-t$ $(k+n-s)$-th order derivatives of the components of u. The case $t=0$ is trivial, so assume $t>0$ for the remainder of the proof. The function u must then be an element of both $W^{k+n-s,\gamma}$ and $W^{k+n-s-1,\delta}$, where

$$\frac{r-t}{\gamma}+\frac{t}{\delta}\leqq 1. \tag{4.2}$$

To complete the argument, we require the Sobolev embedding theorem, cf. Adams [**1**], which gives an embedding

$$W^{l+1,\gamma}\hookrightarrow W^{l,\delta}, \tag{4.3}$$

where γ and δ satisfy one of the following conditions:

$$\begin{array}{llll} \text{(i)} & \gamma<p & \text{and} & \delta\leqq p\gamma/(p-\gamma); \\ \text{(ii)} & \gamma=p & \text{and} & \delta<\infty; \\ \text{(iii)} & \gamma>p & \text{and} & \delta\leqq\infty. \end{array} \tag{4.4}$$

It is easy to check that these conditions and (4.2) are both satisfied if and only if γ satisfies the conditions of the theorem, and hence $J^{\alpha}_{\mathbf{K}}$ is defined as a distribution precisely in these cases.

The reader might wonder whether the theorem could be improved by a suitably clever grouping of the various factors in the Q_I and further use of identities. For instance, row expansions of the Hessian determinants in the third order identity (3.9) would give terms involving second order derivatives of u multiplying Jacobian determinants of v and w. However, each n-th order identity only reduces the number of required derivatives by $[n/r]$, whereas to ensure that the term again lies in an L^{γ}-space requires the raising of the number of derivatives by n. Clearly, any such attempts are counter-productive, and so Theorem 4.1 cannot be improved.

An interesting consequence of this observation and of the classification of all null Lagrangians with nice identities is that we can obtain an upper bound on the order of derivatives which a differential polynomial can depend on in order to have any chance of being defined even distributionally over a given Sobolev space.

THEOREM 4.2. *Consider the Sobolev space $W^{l,r}$. Let P be an r-th degree differential polynomial depending exclusively on the n-th order derivatives of $u \in W^{l,r}$. Then P cannot be defined as a distribution over $W^{l,r}$ using divergence identities if*

$$n > \left[\frac{lr}{r-1}\right], \tag{4.5}$$

where [ ] denotes integer part.

Proof. Clearly, the greatest relaxation of requirements on derivatives is achieved for the hyperjacobians of the form $J_{\mathbf{K}}(u) = \partial^n(u^{i_1}, \ldots, u^{i_r})/\partial\mathbf{K}$, i.e. no derivatives of the u's appear in the arguments of $J_{\mathbf{K}}(u)$. If $u \in W^{l,r}$, then $J_{\mathbf{K}}(u)$ is well defined provided

$$l \geqq n - \left[\frac{n}{r}\right].$$

The upper bound of all such n is

$$n_0 = \left[\frac{lr}{r-1}\right].$$

(Note that if n is odd and q, the dimension of the range space, is $\leqq r$, this bound may be further reduced owing to the skew-symmetry of such hyperjacobians.)

Next, we look at the sequential weak continuity properties of hyperjacobians. The basic ingredient of the proof, which is then trivial, is that the embedding (4.3) is compact as long as γ and δ satisfy one of the conditions (4.4), the only proviso being that in condition (i) only the strict inequality is allowed.

THEOREM 4.3. *Let n, r, s, t, $J^{\boldsymbol{\alpha}}_{\mathbf{K}}$ be as in Theorem 4.1. The mapping $u \to J^{\boldsymbol{\alpha}}_{\mathbf{K}}(u)$ is sequentially weakly continuous from $W^{k+n-s,\gamma}$ to $\mathcal{D}'(\Omega)$ provided*

$$\gamma > \max\left\{\frac{pr}{r+t}, r-t\right\},$$

except in the special case $t=0$, when it is only sequentially weakly continuous from $W^{k+n-s+1,\gamma}$ to $\mathcal{D}'(\Omega)$ for any $\gamma \geqq 1$.

4.2. Examples of variational problems with weak minimizers

We now illustrate the application of the sequential weak continuity results of Theorem 4.3 with a couple of prototypical variational problems. The proofs follow exactly along the lines of BCO, to which we refer the reader for missing details.

In the first example, let u, v be functions of the real variables x, y. The second order hyperjacobians

$$\begin{aligned} J^2(u, v) &= \left(\frac{\partial^2(u, u)}{\partial(x, y)^2}, \frac{\partial^2(u, v)}{\partial(x, y)^2}, \frac{\partial^2(v, v)}{\partial(x, y)^2}\right) \\ &= (2u_{xx}u_{yy} - 2u_{xy}^2, u_{xx}v_{yy} + u_{yy}v_{xx} - 2u_{xy}v_{xy}, 2v_{xx}v_{yy} - 2v_{xy}^2), \end{aligned}$$

can be identified with the principal curvatures of the deformation of the plane described by the map (u, v). According to Theorem 4.3, these three functions are sequentially weakly continuous on $W^{2,\gamma}$ for any γ.

Consider the variational problem

$$I(u, v) = \iint_\Omega F(x, y; u, v; \nabla(u, v); \nabla^2(u, v))\, dx\, dy, \tag{4.6}$$

where $\nabla^k(u, v)$, $k = 1, 2$ denotes all the k-th order derivatives of u and v. We propose to minimize (4.6) subject to non-linear boundary conditions

$$G(x, y; u, v; \nabla(u, v)) = 0, \quad (x, y) \in \partial\Omega. \tag{4.7}$$

The hypotheses on F and G are as follows:

(H1) There is a function $\Phi(x, y; u, v; \nabla(u, v); \nabla^2(u, v); J^2(u, v))$ such that

$$F(x, y; \mathbf{c}_0; \mathbf{c}_1; H) = \Phi(x, y; \mathbf{c}_0; \mathbf{c}_1; H; J^2(H)),$$

for all $\mathbf{c}_0 \in \mathbb{R}^2$;, $\mathbf{c}_1 \in \mathbb{R}^4$, $H \in \mathbb{R}^6$ and almost all $(x, y) \in \Omega$.

(H2) $\Phi(\cdot; \mathbf{c}_0; \mathbf{c}_1; H; J^2): \Omega \to \bar{\mathbb{R}}$ is measurable for each $\mathbf{c}_0$, $\mathbf{c}_1$, H, J^2. ($\bar{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$.)

(H3) $\Phi(x, y; \cdot; \cdot; \cdot; \cdot): \mathbb{R}^{15} \to \bar{\mathbb{R}}$ is continuous for almost all $(x, y) \in \Omega$.

(H4) $\Phi(x, y; \mathbf{c}_0; \mathbf{c}_1; \cdot; \cdot): \mathbb{R}^9 \to \bar{\mathbb{R}}$ is convex for all $\mathbf{c}_0 \in \mathbb{R}^2$, $\mathbf{c}_1 \in \mathbb{R}^4$ and almost all $(x, y) \in \Omega$.

(H5) $F(x, y; \mathbf{c}_0; \mathbf{c}_1; H) \geqq \varphi(x, y) + C(|H| + \Psi(|J^2(H)|))$ (4.8)

for all $\mathbf{c}_0 \in \mathbb{R}^2$, $\mathbf{c}_1 \in \mathbb{R}^4$, $H \in \mathbb{R}^6$, almost all $(x, y) \in \Omega$ where $\varphi \in L^1(\Omega)$, $C > 0$, and $\Psi: \mathbb{R}^+ \to \mathbb{R}$ is convex and satisfies $\Psi(t)/t \to \infty$ as $t \to \infty$.

(C1) $G(\cdot; \mathbf{c}_0, \mathbf{c}_1): \partial\Omega \to \mathbb{R}^m$ is μ-measurable for all $\mathbf{c}_0 \in \mathbb{R}^2$, $\mathbf{c}_1 \in \mathbb{R}^4$, where μ denotes one-dimensional Hausdorff measure.

(C2) $G(x, y; \cdot; \cdot): \mathbf{R}^6 \to \mathbf{R}^m$ is continuous for μ-almost all $(x, y) \in \partial\Omega$.

(C3) There exist measurable subsets $\partial\Omega_i$ of $\partial\Omega$, $i = 1, 2$, with $\mu(\partial\Omega_i) > 0$, and a constant $K \geqq 0$ such that if $G(x, y; \mathbf{c}_0; \mathbf{c}_1) = 0$ for some $(x, y) \in \partial\Omega_i$ and $\mathbf{c}_0 = (c_0^1, c_0^2)$, then $|c_0^i| \leqq K$.

The integrand (4.6) as it stands cannot, strictly speaking, be defined on $W^{2,1}$, so we replace it by the weakly defined integral

$$\tilde{I}(u, v) = \iint_\Omega \Phi(x, y; u, v; \nabla(u, v); \nabla^2(u, v); \tilde{J}^2(u, v))\, dx\, dy,$$

where $\tilde{J}^2$ are the distributionally defined hyperjacobians as in Theorem 4.1. The set of admissible functions is

$$\mathscr{A} = \{(u, v) \in W^{2,1}: \tilde{I}(u, v) < \infty,\ G(x, y; u, v; \nabla(u, v)) = 0,$$
$$\mu\text{-almost everywhere in } \partial\Omega\}.$$

We assume $\mathscr{A}$ is non-empty.

THEOREM 4.4. *Under the above hypotheses, $\tilde{I}$ attains its minimum on $\mathscr{A}$.*

The key difference of this result to that of BCO is that the growth condition in (4.8) has been weakened from an exponent of 2 on $|H|$ needed in BCO. Note that $\tilde{I}$ is being minimized over the space of functions u, v such that the distributionally defined curvatures $\tilde{J}^2$ are functions. However, as mentioned in the introduction, it is not known whether in this case these agree in general with the actual curvatures, in case the latter can be defined, so the precise meaning of the above result is not entirely clear.

As a second example, we show how the higher order identities allow us to construct variational problems where the convexity conditions are placed mainly on lower order derivatives. Let x, y, u, v be as above, and consider the third order hyperjacobian

$$J^3(u, v) = \frac{\partial^3(u, v)}{\partial(x, y)^3} = u_{xxx}v_{yyy} - 3u_{xxy}v_{xyy} + 3u_{xyy}v_{xxy} - u_{yyy}v_{xxx}.$$

Theorem 4.3 shows that $J^3(u, v)$ is sequentially weakly continuous on $W^{2,\alpha}$ for $\alpha > 4/3$. Consider the variational problem

$$I = \iint_\Omega F(x, y; u, v; \nabla(u, v); \nabla^2(u, v); \nabla^3(u, v))\, dx\, dy \tag{4.9}$$

with boundary conditions (4.7). The hypotheses on G are the same as before; those on F are

(H1) There exists $\Phi(x, y; u, v; \nabla(u, v); \nabla^2(u, v); J^3(u, v))$,

such that

$$F(x, y; \mathbf{c}_0; \mathbf{c}_1; \mathbf{c}_2; H) = \Phi(x, y; \mathbf{c}_0; \mathbf{c}_1; \mathbf{c}_2; J^3(H)),$$

for all $\mathbf{c}_0 \in \mathbb{R}^2$, $\mathbf{c}_1 \in \mathbb{R}^4$, $\mathbf{c}_2 \in \mathbb{R}^6$, $H \in \mathbb{R}^8$ and almost all $(x, y) \in \Omega$.

(H5) $\Phi(x, y; \mathbf{c}_0; \mathbf{c}_1; H; J^3) \geqq \varphi(x, y) + C(|H|^\alpha + \Psi(|J^3|))$,

where φ, Ψ, C are as before and $\alpha > \frac{4}{3}$.

The remaining hypotheses (H2–4) are similar to these above, and are left to the reader to state explicitly. Again, the integral I must be replaced by

$$\tilde{I}(u, v) = \iint_\Omega \Phi(x, y; u, v; \nabla(u, v); \nabla^2(u, v); \tilde{J}^3(u, v)\, dx\, dy.$$

The set of admissible functions is

$$\mathcal{A} = \{(u, v) \in W^{2,\alpha} \colon \tilde{I}(u, v) < \infty,\ G(x, y; u, v; \nabla(u, v)) = 0$$
$$\text{for } \mu\text{-almost all } (x, y) \in \partial\Omega\},$$

and we assume $\mathcal{A}$ is non-empty.

THEOREM 4.5. *Under the above hypotheses, $\tilde{I}$ attains its minimum on $\mathcal{A}$.*

The point here is that the convexity conditions are placed mainly on the second derivatives of u, v, except for the convex dependence of the integrand on J^3.

Acknowledgments

The research in this paper was for the most part accomplished under a United Kingdom SRC research grant during 1980 at the University of Oxford. It is a pleasure to thank John Ball for helpful suggestions, comments and much needed encouragement.

References

1 R. A. Adams. *Sobolev Spaces* (New York: Academic Press, 1975).
2 S. J. Aldersley. Higher Euler operators and some of their applications. *J. Math. Phys.* **20** (1979), 522–531.
3 I. M. Anderson and T. Duchamp. On the existence of global variational principles. *Amer. J. Math.* **102** (1980), 781–868.
4 J. M. Ball. Convexity conditions and existence theorems in nonlinear elasticity. *Arch. Rational Mech. Anal.* **63** (1977), 337–403.
5 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.
6 C. DeConcini, D. Eisenbud and C. Procesi. Young diagrams and determinantal varieties. *Invent. Math.* **56** (1980), 129–165.
7 G. V. Escherich. Die Determinanten hoheren Ranges und ihre Verwendung zur Bildung von Invarianten. *Denkschr. Kais. Akad. Wiss. Wien* **43** (1882), 1–12.
8 I. M. Gel'fand and L. A. Dikii. Asymptotic behavior of the resolvent of Sturm-Liouville equations and the algebra of the Korteweg-DeVries equation. *Russian Math. Surveys* **30** (1975), 77–113.
9 L. Gegenbauer. Über Determinanten hoheren Ranges. *Denkschr. Kais. Akad. Wiss. Wien* **43** (1882), 17–32.
10 P. Gordan. *Vorlesung Über Invariantentheorie*, Vol. 2 (ed. G. Kerschensteiner) (Leipzig: Teubner, 1887).
11 J. H. Grace and A. Young. *The Algebra of Invariants* (Cambridge University Press, 1903).
12 G. B. Gurevich. *Foundations of the Theory of Algebraic Invariants* (Groningen: Noordhoff, 1964).
13 M. Hochster. Criteria for equality of ordinary and symbolic powers of primes. *Math. Z.* **133** (1973), 53–65.
14 W. V. D. Hodge and D. Pedoe. *Methods of Algebraic Geometry*, Vol. II (Cambridge University Press, 1954).
15 A. Lascoux. Syzygies des varietés determinantales. *Adv. in Math.* **30** (1978), 202–237.
16 K. R. Mount. A remark on determinantal loci. *J. London Math. Soc.* **42** (1967), 595–598.
17 D. G. Northcott. Some remarks on the theory of ideals defined by matrices. *Quart. J. Math. Oxford* **14** (1963), 193–204.
18 R. Oldenburger. Higher dimensional determinants. *Amer. Math. Monthly* **47** (1940), 25–33.
19 P. J. Olver. Euler operators and conservation laws of the BBM equation. *Math. Proc. Cambridge Philos. Soc.* **85** (1979), 143–160.
20 P. J. Olver. Differential hyperforms I. Univ. of Minnesota, Mathematics Report 82–101.
21 C. Shakiban. *The Euler operator in the formal calculus of variations* (Ph.D. Thesis, Brown University, Providence, R.I., 1979).
22 C. Shakiban. A resolution of the Euler operator II. *Math. Proc. Cambridge Philos. Soc.* **89** (1981), 501–510.
23 J. Towber. Two new functors from modules to algebras. *J. Algebra* **47** (1977), 80–104.
24 N. V. Trung. On the symbolic powers of determinantal ideals. *J. Algebra* **58** (1979), 361–369.

(*Issued* 8 *November* 1983)