THE PROLONGATION FORMULA FOR TENSOR FIELDS

SATYANAD KICHENASSAMY

Abstract. We derive a generalization of the prolongation formula for tensor-valued functions, by taking into account the natural action of diffeomorphisms on tensor fields. This differs from the usual procedure, where such fields are viewed as defining "multi-component scalars;" it replaces the ordinary derivative in the expression of the "characteristic" by a Lie derivative. The resulting version of Noether's theorem takes a form closer to what is familiar to theoretical physicists. A very simple proof of the conformal invariance of the Maxwell Lagrangian also follows from this procedure. The Eshelby tensor is also readily obtained, as a further illustration of the practical use of this new prolongation formula.

Contents

1. Introduction. **2**

2. Group actions on tensor bundles. **4**

3. Expression of the prolongation in local coordinates. **8**

3.1. Covectors

3.2. Vectors

3.3. General tensors

3.4. Prolongation

3.5. An alternative method

4. Applications. **13**

4.1. Variational Symmetries

4.2. Conservation laws

4.3. First-order Lagrangians

4.4. Example 1: Linear Elasticity

4.5. Example 2: Maxwell's equations

5. Concluding Remark **19**

References **19**

1. Introduction.

S. Lie (1881), in his study of symmetries of differential equations, introduced the *prolongation* of the infinitesimal generator of a symmetry, which expresses the infinitesimal variation of the unknown function and its derivatives, under the group action. E. Noether (1918) applied this procedure to the variation of integrals, and established in a very general framework the connection between the invariance group of an integral and the conservation laws of the associated Euler-Lagrange equations. The more general case when the first variation of the Lagrangian is a divergence (divergence symmetries) was studied by Bessel-Hagen (1921).

Only later was the prolongation method used in a systematic way to classify symmetries of large numbers of equations (Ovsiannikov (1962), Bluman and Cole (1974), Ibragimov (1975)); the procedure was streamlined in the process, by the introduction of a simple closed form expression for the nth prolongation, which is referred to as *the prolongation formula* (see Olver (1986) ch. 2, where one can also find other historical remarks). The modern presentation of the prolongation (Olver (1979)), is as follows:

For a differential equation in n independent variables and one scalar dependent variable u, we consider the space $X \times U \ni (x, u)$, where $X = \mathbf{R}^n$, and $U = \mathbf{R}$, or an open set thereof, on which a group generated by

$$v = \xi^i \partial_i + \varphi \frac{\partial}{\partial u}$$

acts locally. This group defines a natural transformation on functions $u = u(x)$ and their derivatives. The information is summarized into prolonged vector fields

$$\mathrm{pr}_v^{(k)} = \xi^i \partial_i + \varphi \frac{\partial}{\partial u} + \sum_{|J| \le k} \varphi_J \frac{\partial}{\partial u_{,J}},$$

$k = 1, 2, \ldots$, which can be viewed as vector fields on the jet spaces $J^k(X; U)$. For systems, one repeats the operation on each dependent variable, treated as a scalar. This version of the prolongation has been, since then, applied to the determination of symmetries of differential equations, the computation of finite-dimensional reductions (classical and "non-classical" (see Olver (1993)), and the formation of conservation laws, to name a few.

Conservation laws constructed in this manner are particularly useful in the proof of global bounds on solutions of nonlinear PDEs, and in the study of cracks in elastic materials. In particular, the procedure provides an explanation for the "multiplier," introduced empirically to derive conservation laws from Euler-Lagrange equations, in terms of the "characteristic" $\varphi - \xi^i u_{,i}$ of the symmetry.

With the development of field theories, many simplified accounts introduced weakened forms of Noether's results: see e.g. E. L. Hill (1951), I. M. Gel'fand and S. V. Fomin (1963), E. Schmutzer (1968), D. Edelen (1969), D. Anderson (1973), and J. D. Logan (1977). The last reference gives an elementary treatment of Noether's

theorems which does not rely on the fundamental variational formula, but rather uses a set of invariance identities obtained through the use of Euler's equations in the variation (see also D. Lovelock and H. Rund (1975)). One should also note that these simplified accounts limited to Noether's theorems did not lead to the important development of group-theoretic methods in PDE made possible by the prolongation method in the last twenty years. The case when the dependent variable is a *tensor* was handled essentially in three different ways:

(i) By treating each component as a scalar-valued function. This amounts to fixing once and for all a local trivialization of the tensor bundle under consideration. In this framework, we usually recover conservation laws with a simple physical interpretation only as a linear combination of those corresponding to groups acting on the independent variables alone and those acting on the dependent variables alone (see e.g. Ibragimov (1985)).

(ii) By breaking up the total variation of the dependent variables into the sum of "small variations" consecutive to: (a) variation of the point x; (b) variation of the dependent variable; (c) variation of "local frames" (Bogoliubov and Shirkov (1959), Schmutzer (1968), Edelen (1969), among others). Bogoliubov and Shirkov (1959) and Schmutzer (1968) also contain material on spinors. The variation of local frames was also used in some cases by Ibragimov (1985) and Fuschich *et al.,* (1987, 1993) in conjunction with method (i).

(iii) By equating (rather than adding) two expressions for the variation of the Lagrangian, one due to the integrand being a scalar density, and the other consecutive to the variation of the dependent variables by the group action (see J. N. Goldberg (1953), J. G. Fletcher (1960), and J. L. Anderson (1967)).

However, these developments call for the following commments:

(a) It is well-known that while the first order derivatives of a function define a covariant vector, there is no intrinsic meaning to higher derivatives. Thus, the Hessian has an invariant meaning only at critical points (which makes Morse theory possible). This is generally obviated by introducing jet spaces. But jet spaces are not designed to give an invariant meaning to individual derivatives—only to the notion of order of contact. In a vector bundle, the situation is even worse, since it is not clear how to differentiate sections even once, in the absence of a connection. Even with a connection, the commutation of derivatives is lost in general.

(b) If the part of the variation of a tensorial quantity due to the variation of "local frames" may be explained in concrete terms by the "dragging" of coordinates, one would like to have a systematic way of computing these terms, and of incorporating them in the prolongation formula, to recover the ease of use of the latter, especially for higher-order Lagrangians.

(c) One would like to explain why procedure (iii) is successful in yielding the correct conservation laws.

Now, we will see that in a tensor bundle, one can, despite point (a) above, define the first variation by pull-back and compute an intrinsic prolongation formula, provided one modifies the definition of the characteristic. This is due to the fact that any action on the base of a tensor bundle can be lifted to the full bundle in a definite way, as we show in §2. In terms of jet bundles, one must compute the k-jet extension of the transformation on sections of *tensor bundles* induced by the group, rather than on sections of a *product* bundle $X \times U$. The translation in local coordinates is given in §3, where the prolongation formula for tensors is stated and proved. To illustrate the practical application of this formalism, we discuss conservation laws for three-dimensional elasticity and for Maxwell's equations in §4, obtaining in particular a very simple proof of the conformal invariance of the latter. A more general classification of conservation laws for electromagnetism will be discussed elsewhere.

2. Group actions on tensor bundles.

We develop a few abstract results on group actions on tensor bundles. Their expression in local coordinates is given in this section and the next.

The material of the first two paragraphs of this section actually extends to a quite general vector bundle, while the second uses the fact that an invertible map on the base of a tensor bundle has a natural lift as bundle map.

2.1. Consider a smooth vector bundle (E, B, π), and a group G acting on E. Note that the elements of G are *not* bundle transformations: they do not map fibers to fibers in general. We may, for simplicity, assume that G is a one-parameter group. As usual, B is the base, and π is the projection from E onto it. Two examples are relevant for our applications:

Example 1: Product bundle. Let B be an open subset of $\mathbf{R}^n$, and let $E = B \times \mathbf{R}^p$. The projection is

$$\pi : (x^1, \dots, x^n; u^1 \dots, u^p) \mapsto (x^1, \dots, x^n).$$

A section of E defines a "multi-scalar" function $u = u(x)$:

$$s : (x^1, \dots, x^n) \mapsto (x^1, \dots, x^n; u^1(x) \dots, u^p(x))$$

Example 2: Tensor bundle. B is taken to be $\mathbf{R}^n$ (or any n-dimensional manifold), and E is a tensor bundle on it; local coordinates on B will have the form $(x^1, \dots, x^n)$, while those on E have the form

$$((x^i); (u^{j_1,\dots,j_q}_{i_1,\dots,i_p}))$$

where all indices take all possible values from 1 to n. Sections of this bundle define tensor fields on B, of type (p, q). It will often be convenient to abbreviate these coordinates to (x^i, u^I), where I stands for all tensor indices, covariant or contravariant.

The summation convention will be used throughout the paper. Note that a tensor field $u^I = u^I(x)$is identified with the section

$$s : (x^i) \mapsto (x^i; u^I(x)).$$

The projection has locally the form $\pi(x^i; u^I) = (x^i)$.

Example 1 leads to the usual prolongation formula, and Example 2 to the one derived in §3.

We keep the presentation at this more abstract level in this section since the procedure is also potentially applicable to principal bundles as well.

2.2. We wish to study the induced action of elements of G close to the identity on sections of E. In practice, it will be enough that the action be defined in the neighborhood of the fiber above some point of the base.

Let therefore $g \in G$ be close to the identity. Let us also fix a section $s : B \to E$ (so that, by definition, $\pi \circ s = \mathrm{id}$).[1] It defines an action f on the base, and an induced section s', which make the following diagram commutative:

$$\begin{array}{ccc} E & \xrightarrow{g} & E \\ {\scriptstyle s}\uparrow\downarrow{\scriptstyle \pi} & & {\scriptstyle s'}\uparrow\downarrow{\scriptstyle \pi} \\ B & \xrightarrow{f} & B \end{array}$$

More precisely,

$$f = \pi \circ g \circ s : B \to B; \quad x \mapsto x'(x, g). \tag{1}$$

Since g is close to the identity, this map is invertible. (Again, the inverse may only be locally defined in the applications). Define now the induced section $s'(x')$ so that $s' \circ f = g \circ s$:

$$s' := g \circ s \circ f^{-1}. \tag{2}$$

2.3. The passage from s to s' expresses the action of the group on E-valued fields on B; it has been described in an entirely coordinate-free manner.

Let us now turn to the problem of defining the prolongation of the group action. Since we have a map $s \mapsto s'$ from sections of E to sections of E, there is a natural k-jet extension for every k, which maps sections of $J^k(E)$ to sections of $J^k(E)$; the prolongation formula will describe this map at the infinitesimal level.

We therefore need to be able to compute "derivatives of s with respect to group parameters." It is however not possible to differentiate s with respect to group parameters in the obvious way because we cannot substract vectors attached to different points of the base. In Example 1, this difficulty disappears, because there is a canonical identification of all the fibers at different points. For Example 2, E is not a

[1] Olver (1979) considers also sections with 'vertical tangents,' but we limit ourselves to a more familiar set-up.

product bundle anymore: we may not identify all fibers with one another *a priori*, and another procedure is needed.

2.4. We now limit ourselves to Example 2 and show how to define intrinsically the infinitesimal group action. In that case, f is a diffeomorphism of B, and therefore induces a transformation on tensors of any type. It therefore induces in particular a bundle map ('push-forward') $F : E \to E$. If we compose g with F^{-1}, we obtain a map which acts as the identity on the base, but which reflects the group action on the fiber. Comparing it with the identity, we obtain a measure of the variation in the section:

$$\delta s = [F^{-1} \circ g - \mathrm{id}] \circ s = F^{-1} \circ s' \circ f - s.$$

Assume now that $g = g_\varepsilon$, where ε is the group parameter. To any section s, we associate as above a new section s'_ε, and maps f_ε and F_ε. The infinitesimal action of g_ε on E is of course represented by a vector field tangent to E, but it can also be represented more concretely, if s is given, by

$$\frac{d}{d\varepsilon} x'(x, u, \varepsilon)$$

and

$$\frac{d}{d\varepsilon} F_\varepsilon^{-1} \circ s'_\varepsilon \circ f_\varepsilon,$$

both evaluated at $\varepsilon = 0$. Indeed, the first is a vector field on B, while the second can be identified with a tensor field.

The role of F_ε^{-1} is to bring the point x' back to x so that we may always compare tensors attached to the same point, which should be independent of ε. The derivative of δs with respect to ε at $\varepsilon = 0$ is a well-defined section.

This is reminiscent of the procedure used in the construction of the Lie derivative.

The various mappings involved are now summarized in the following diagram.

$$\begin{array}{ccc} E & \xrightarrow{F_\varepsilon} & E \\ {\scriptstyle s}\uparrow\downarrow{\scriptstyle \pi} & & {\scriptstyle s'_\varepsilon}\uparrow\downarrow{\scriptstyle \pi} \\ B & \xrightarrow{f_\varepsilon} & B \end{array}$$

2.5. We now give the expression of g, f, F and δs in local coordinates.

Pick a coordinate system $\{x^i\}$ on B; it induces a local field of natural frames and coframes: $\{\partial/\partial x^i\}$ and $\{dx^i\}$, and more generally, by tensor product of these, fields $\{e_I\}$ which form a basis for (p, q)-tensors on B. The generic element of E is therefore represented by

$$(x^i; u) \quad \text{where} \quad u = u^I e_I.$$

The action of g is then represented by

$$(x^i; u^I) \mapsto (x'^i; u'^I).$$

Let us now choose a section s of E: $s : x \mapsto (x, u(x))$. It defines a mapping f from B to itself:

$$f : (x^1, \ldots, x^n) \mapsto (x'^1, \ldots, x'^n).$$

The $\{x'^i\}$ also induce new frames and coframes: $\{\partial/\partial x'^i\}$ and $\{dx'^i\}$. There is again a corresponding basis for tensors, which we call $\{e'_{I'}\}$.

We view f now as a *local diffeomorphism.*

The mapping F describes the transformation of tensor components under this diffeomorphism. To describe it explicitly, we introduce 1-forms $dx^{1'}$, ... ,$dx^{n'}$ by

$$dx^{1'} = \frac{\partial x'^1}{\partial x^j} dx^j, \ldots dx^{n'} = \frac{\partial x'^n}{\partial x^j} dx^j.$$

We then have

$$F(x, dx^{i'}) = (x', dx'^i).$$

Note that in this equation, one should think of $dx^{i'}$ as an element of $T^*_x B$ ("attached at x") while dx'^i lives in $T^*_{x'} B$ (and is therefore "attached at x'").

Similarly,

$$F(x, dx^i) = (x', \frac{\partial x^i}{\partial x'^j} dx'^j).$$

For vector fields, we define analogously $\partial_{1'}$, ... ,$\partial_{n'}$ by

$$\partial_{1'} = \frac{\partial x^j}{\partial x'^1} \frac{\partial}{\partial x^j}, \ldots, \partial_{n'} = \frac{\partial x^j}{\partial x'^n} \frac{\partial}{\partial x^j}.$$

We then have

$$F(x, \partial_{i'}) = (x', \frac{\partial}{\partial x'^i})$$

and

$$F(x, \partial_i) = (x', \frac{\partial x'^j}{\partial x^i} \frac{\partial}{\partial x'^j}).$$

Finally, general tensors can be expressed in either basis:

$$u = u^{I'} e'_{I'} = u^I e_I,$$

and

$$F(x, u^I e_I) = (x', u^{I'} e'_{I'}),$$

while

$$(x, u'^I e_I) = F^{-1}(x', u'^{I'} e'_{I'}).$$

More explicitly,

$$F(x^i; u^I) = (x'^1, \ldots, x'^n; u^{l_1 \ldots l_q}_{k_1 \ldots k_p} \frac{\partial x^{k_1}}{\partial x'^{i_1}} \cdots \frac{\partial x^{k_p}}{\partial x'^{i_p}} \frac{\partial x'^{j_1}}{\partial x^{l_1}} \cdots \frac{\partial x'^{j_q}}{\partial x'^{l_q}}). \tag{3}$$

We then find

$$\delta s(x) = (x, (u'^I - u^I) e_I).$$

REMARK: The operation whereby dx'^i, is brought back from x' to x (by F) has been interpreted in earlier works as *"dragging from one point to another."* It consists in the identification of dx'^i, which is viewed as belonging to the cotangent space above x', with $(\partial x'^i/\partial x^j)\partial x^j$, a combination of the basis elements of the cotangent space at x. Similarly, F "drags" $u'^i e_I$ back to x. This is made possible by the fact that the charts $\{x^i\}$ and $\{x'^i\}$ overlap, and therefore, the local fields of forms associated with $\{x^i\}$ are also defined in the vicinity of x', and can be used to represent tensors attached at x', and vice-versa. The usual pull-back is, in this framework, a combination of (a) the chain rule and (b) dragging. It is convenient to assume that F is the identity and write

$$dx^{i'} = dx'^i = \frac{\partial x'i}{\partial x^j} dx^j,$$

and we will often do so.

2.6. One may also argue in the following way. From the expression of the Lie derivative $\mathcal{L}_X s'$ of s' with respect to an arbitrary vector field X on the base B (unrelated to the group generators), one can find all the first-order derivatives of s', and therefore, the first prolongation. Higher derivatives can be computed from the action of several successive Lie derivatives on s'. Expressing these in local coordinates in two different ways will lead in the next section to a second proof of the prolongation formula; it is also noteworthy that there is a completely intrinsic way of computing the prolongation using Lie derivatives.

3. EXPRESSION OF THE PROLONGATION IN LOCAL COORDINATES.

We translate in local coordinates the construction of the prolongation in the previous section.

All calculations are carried out for a one-parameter group, and are always local. Dots in formulae (...) indicate terms of second or higher order in Taylor expansions with respect to the group parameter ε. The general prolongation formula will differ from the usual one only in the form of the characteristic.

Let I stand for an arbitrary set of indices (up, down, or both), corresponding to the type of tensors under consideration. We also write D_J for $D_{j_1} \dots D_{j_q}$, if $J = (j_1, \dots j_q)$. D_j stands for the total derivative with respect to x^j. We denote by $|J|$ the length of the multi-index J, i.e., the number of its components: $|(j_1, \dots, j_q)| = q$. The rest of the notation is as in §2. The group action has the form

$$x'^i = x^i + \varepsilon \xi^i(x, u) + \dots \tag{4}$$

$$u'^I = u^I + \varepsilon \phi^I(x, u) + \dots . \tag{5}$$

Since we are given a relation $u = u(x)$, the first line defines a point transformation. It will enable us to define new components $u'^{I'}$. The goal is to express the derivatives of

$u'^{I'}$ with respect to the x', and to compute the linear term in their Taylor expansion near $\varepsilon = 0$. The result is the prolongation formula.

This task is broken up in two parts: first, one must compute $u'^{I'}$; next, one must compute its successive derivatives.

We deal with the first task, on vectors and covectors, from which the case of general of general tensors is easy to describe. We then state and prove the prolongation formula by induction on the number of derivatives.

3.1. Covectors. We are interested here in the case $E = T^*\mathbf{R}^n$.

A section is now a one-form $u = u_a dx^a$. Since u is given as a function of x, the group induces a point transformation $x \mapsto x'$.

In local coordinates, we therefore write the group action infinitesimally:

$$x'^i = x^i + \varepsilon\xi^i(x,u) + \dots \tag{6}$$

$$u'_a = u_a + \varepsilon\phi_a(x,u) + \dots, \tag{7}$$

where i and a run from 1 to n. Substituting $u = u(x)$, we obtain the point transformation $x \mapsto x'$:

$$x'^i = x'^i(x, u(x), \varepsilon),$$

and the transformed function $u'(x')$ obtained by eliminating x between (6) and (7).

Now, using the notation of §2.5:

$$dx'^a = (\delta^a_b + \varepsilon\xi^a{}_{,b} + \dots)\, dx^b, \tag{8}$$

and

$$dx^a = (\delta^a_b - \varepsilon\xi^a{}_{,b} + \dots)\, dx'^b, \tag{9}$$

where the comma denotes a *total* derivative; as is commonly done, we will at times write $D_b\xi^a$ instead of $\xi^a{}_{,b}$, to emphasize this.

One can now express u' in two ways

$$u' = u'_a dx^a = u'_{a'} dx^{a'} (= u'_{a'} dx'^a).$$

Denoting by $\mathcal{L}_\xi$ the Lie derivative in the direction $\xi^i\partial_i$, we find

$$\begin{aligned} u'_{a'} &= (u_b + \varepsilon\phi_b + \dots)(\delta^b_a - \varepsilon\xi^b{}_{,a} + \dots) \\ &= u_a + \varepsilon[\phi_a - \xi^b{}_{,a} u_b] + \dots \\ &= u_a + \varepsilon[\phi_a - (\mathcal{L}_\xi u)_a + \xi^b u_{a,b}] + \dots. \end{aligned} \tag{10}$$

This is the desired formula.

3.2. Vectors. Let $E = T\mathbf{R}^n$.

A section is now a vector field $u = u^a\partial_a$. From considerations similar to those developed before, we see that the group action can be written

$$x'^i = x^i + \varepsilon\xi^i(x,u) + \dots \tag{11}$$

$$u'^a = u^a + \varepsilon\phi^a(x,u) + \dots . \tag{12}$$

Given u, this again defines a point transformation. Again, u' can be written in two different ways:

$$u' = u'^a\partial_a = u'^{a'}\partial_{a'},$$

where

$$\begin{aligned} u'^{a'} &= (u^b + \varepsilon\phi^b + \dots)(\delta^a_b + \varepsilon\xi^a{}_{,b} + \dots) \\ &= u^a + \varepsilon[\phi^a + \xi^a{}_{,b}u^b] + \dots \\ &= u^a + \varepsilon[\phi^a - (\mathcal{L}_\xi u)^a + \xi^b u^a{}_{,b}] + \dots, \end{aligned} \tag{13}$$

since $\partial_b = (\delta^a_b + \varepsilon\xi^a{}_{,b} + \dots)\partial_{a'}$

3.3. General tensors. The expression of u' in terms of x' now takes the form

$$u'^{I'} = u^I + \varepsilon[\phi^I - (\mathcal{L}_\xi u)^I + \xi^h u^I{}_{,h}] + \dots \tag{14}$$

by an argument similar to that of §§3.1 and 3.2.

Remark: One may recover from the factor of ε in (14) the total variation computed by method (ii) (see §1) in the treatments of conservation laws in [5, 7, 8, 9, 13, 23], and recognize its geometrical origin.

By analogy with the scalar case, we define the *characteristic* to be

$$C^I = \phi^I - (\mathcal{L}_\xi u)^I. \tag{15}$$

Note that C^I does transform like a tensor. For multi-component scalars, one should replace the Lie derivative by an ordinary derivative. This has the effect of making the two expressions of u' (namely $u'^{I'}$ and u'^I) identical, as it should be. The present formalism then reduces exactly to the usual one.

3.4. Prolongation. The prolongation formula gives the infinitesimal variation of the derivatives $D_{K'}u'^{I'}$ under the group action, where I' and K' are multi-indices. It will be convenient to use the convention that if the length of K is zero, then $D_K u^I = u^I$.

For the usual procedure, and many applications, see Olver [18, 20], Ibragimov [13].

The main result of this section is:

Theorem 1. *The prolongation formula for tensors reads:*

$$D_{I'}u^{J'} = D_I u^J + \varepsilon[D_I C^J + \xi^h u^J{}_{,Ih}] + \dots . \tag{16}$$

In other words, if

$$v = \xi^i \partial_i + \phi^J \frac{\partial}{\partial u^J},$$

we have

$$\mathrm{pr}_v^{(m)} = v + \sum_{|I| \leq m} \phi_I^J \frac{\partial}{\partial u^J{}_{,I}},$$

where

$$\phi_I^J = D_I C^J + \xi^h u^J{}_{,Ih}.$$

Proof: We already know from (14) that this formula is correct for derivatives of order 0.

For the sake of clarity, let us write out the case of one derivative, for a contravariant vector; the general case is treated next. Using (13) to compute $u'^{a'}$, we find

$$\begin{aligned} D_{k'} u'^{a'} &= (\delta_k^h - \varepsilon \xi^h{}_{,k} + \dots) D_h (u^a + \varepsilon[\phi^a + \xi^a{}_{,b} u^b] + \dots) \\ &= D_k u^a + \varepsilon[\phi^a{}_{,k} + D_k(\xi^a{}_{,b} u^b) - \xi^h{}_{,k} u^a_{,h}] + \dots \\ &= D_k u^a + \varepsilon[D_k(\phi^a + \xi^a{}_{,b} u^b - \xi^h D_h u^a) + \xi^h D_{hk} u^a] + \dots \\ &= D_k u^a + \varepsilon[D_k C^a + \xi^h D_h u^a{}_{,k}] + \dots, \end{aligned}$$

with $C^a = \phi^a - (\mathcal{L}_\xi u)^a$. We now prove the general formula by induction on the number of derivatives, i.e., on the length of the multi-index $|I|$. The case when $|I| = 0$ is given by (14).

Let us assume (16) for some I, and let u compute $D_{h'I'} u'^{J'}$:

$$\begin{aligned} D_{h'I'} u'^{J'} &= D_{h'}(D_{I'} u'^{J'}) \\ &= (\delta_h^j - \varepsilon \xi^j{}_{,h} + \dots) D_j (D_I u^J + \varepsilon(D_I C^J + \xi^k D_k u^J{}_{,I}) + \dots) \\ &= D_{hI} u^J + \varepsilon(D_{hI} C^J + D_h(\xi^k D_k u^J{}_{,I}) - \xi^j{}_{,h} D_j u^J{}_{,I}) + \dots \\ &= D_{hI} u^J + \varepsilon(D_{hI} C^J + \xi^k D_k u^J{}_{Ih}) + \dots, \end{aligned}$$

QED.

From a strictly computational view-point, this amounts to changing the definition of the characteristic given in [20], by adding a term to ϕ^I. One can compare the result of Th. 1 with procedure (i) of §1: if u is treated as a multi-scalar, and its vector nature is not taken into account, it will show in the non-invariance of the equations under natural transformations (such as rotations, for instance). One will be led to compensate for this by adding a term to ϕ^I. We have a systematic procedure for finding this term: it is the difference $\xi^k \partial_k u^I - (\mathcal{L}_\xi u)^I$. It is however difficult to explain this procedure and to make it systematic without the help of geometric ideas. At any rate, equations which failed to be invariant when their unknowns were viewed as scalars will now have their natural invariance properties.

It should also be pointed out that since there is no restriction on the choice of ϕ^I and ξ^h, the present procedure has the same level of generality as the usual one.

3.5. An alternative method. There is a more invariant, albeit more complicated, way to derive this prolongation formula. We present it only for the first derivative of a covariant vector, for the sake of simplicity.

Let X be a vector field, which is quite unrelated to the group generator. It can be expressed in two ways:

$$X = X^a \partial_a = X'^a \partial_{a'}.$$

Let us compute in particular the Lie derivative $\mathcal{L}_X u'$ in the primed variables in two different ways, and show how to deduce from this the expression for the first prolongation. For the rest of this section, we make all calculations modulo ε^2, and consequently omit the ....

METHOD 1: We start from the relation

$$\mathcal{L}_X u' = \mathcal{L}_X u + \varepsilon \mathcal{L}_X \phi.$$

It shows that

$$\begin{aligned}(\mathcal{L}_X u')_{a'} &= (\mathcal{L}_X u')_b (\delta^b_a - \varepsilon \xi^a{}_{,b}) \\ &= (\mathcal{L}_X u)_a + \varepsilon[(\mathcal{L}_X \phi)_a - \xi^a{}_{,b}(\mathcal{L}_X u)_b].\end{aligned}$$

The linear term in ε in the above reads

$$X^b \phi_{a,b} + X^b{}_{,a}\phi_b - \xi^b{}_a (X^c u_{b,c} + X^c{}_{,b} u_c). \tag{17}$$

METHOD 2: We may also compute $\mathcal{L}_X u'$ directly in the primed coordinates. This will involve the first prolongation ϕ_{ab} (without comma), which satisfies

$$u'_{a',b'} = u_{a,b} + \varepsilon \phi_{ab}.$$

Equating the results will enable us to compute ϕ_{ab}.

Let us first note that

$$X^{b'} = X^b + \varepsilon \xi^b{}_{,a} X^a,$$

and

$$\begin{aligned}X^{b'}{}_{,a'} &= (\delta^c_a - \varepsilon \xi^c{}_{,a})\partial_c (X^b + \varepsilon \xi^b{}_{,k} X^k) \\ &= X^b{}_{,a} + \varepsilon[\xi^b{}_{,ak} X^k - \xi^c{}_{,a} X^b{}_{,c} + \xi^b{}_{,k} X^k{}_{,a}].\end{aligned}$$

Now,

$$\begin{aligned}(\mathcal{L}_X u')_{a'} &= X^{b'} u'_{a',b'} + X^{b'}{}_{,a'} u'_{b'} \\ &= (X^b + \varepsilon \xi^b{}_{,m} X^m)(u_{a,b} + \varepsilon \phi_{ab}) \\ &\quad + (X^b{}_{,a} + \varepsilon[\xi^b{}_{,ak} X^k - \xi^c{}_{,a} X^b{}_{,c} + \xi^b{}_{,k} X^k{}_{,a}]) \\ &\quad \times (u_b + \varepsilon(\phi_b - \xi^c{}_{,b} u_c)).\end{aligned}$$

The linear term in ε is this expression is now

$$\begin{aligned} X^b\phi_{ab} + \xi^b{}_{,m}u_{a,b}X^m + X^b{}_{,a}(\phi_b - \xi^c{}_{,b}u_c) \\ + u_b[\xi^b{}_{,ak}X^k - \xi^c{}_{,a}X^b{}_{,c} + \xi^b{}_{,k}X^k{}_{,a}]. \end{aligned} \tag{18}$$

Equating (17) and (18) gives

$$X^b\left\{\phi_{ab} - D_b[\phi_a - \xi^c u_{a,c} - \xi^c{}_{,a}u_c] + \xi^c u_{a,bc}\right\} = 0.$$

Since X^b is arbitrary, we recover

$$\phi_{ab} = D_b C_a + \xi^c u_{a,bc},$$

which is, as it should be, identical with the result of Theorem 1.

4. Applications.

After defining the notion of variational symmetry in this set-up, we prove the corresponding version of Noether's theorem for Lagrangians of any order. We then specialize to the case of first-order Lagrangians where the dependent variable is a covector. We then treat two examples in detail: Maxwell's equations in Minkowski space, and the equations of three-dimensional elastostatics. We refrain from giving here a treatment of ordinary and generalized symmetries of equations without a variational structure. We also do not attempt any classification of variational symmetries since our aim is simply to illustrate the present procedure.

4.1. Variational Symmetries. We consider a Lagrangian of order m:

$$\int L(x, A^I, A^I{}_{,J}, \dots)\, dx,$$

where A^I is a tensor of arbitrary type (recall that I stands for a collection of indices which may by up, down, or in mixed positions); L is a scalar formed with the derivatives of A upto order m. We say that

$$v := \xi^i\partial_i + \phi^I\partial_{A^I}$$

generates a variational symmetry of this Lagrangian if

$$\mathrm{pr}_v^{(m)}(L) + LD_h\xi^h = 0,$$

where we have, from Theorem 1,

$$\mathrm{pr}_v^{(m)}(L) := \xi^h\partial L/\partial x^h + \sum_{I,J} L_I{}^J(C^I{}_{,J} + \xi^h A^I{}_{,Jh}),$$

$$L_I{}^J := \partial L/\partial A^I_{,J}$$

and C^I is the characteristic. This invariance condition can be written in a more compact form using the total derivative:

$$D_h(L\xi^h) + \sum_{I,J} L_I{}^J C^I{}_{,J} = 0. \tag{19}$$

The above sum includes the term $\sum_I L_I C^I$, corresponding to $|J| = 0$. Note also that the Euler-Lagrange equations take the form

$$E_I := \sum_H (-1)^{|H|} D_H L_I{}^H = 0.$$

This sum is always finite, since the Lagrangian depends on finitely many derivatives of A^I. The same will be true of the other sums below. We have explicitly written the summation operator at times, for added clarity, although all formulae are consistent with the summation convention. Also, capital letters (H, I, K) refer to multi-indices.

4.2. Conservation laws. We now prove a version of Noether's theorem for one-parameter groups for the present set-up.

Theorem 2. *The identity:*

$$C^I E_I + D_j P^j = 0, \tag{20}$$

where

$$P^j = L\xi^j + \sum_{H,K} (-1)^{|K|} C^I{}_{,H} L_I{}^{jHK}{}_{,K},$$

holds identically.

Remark: In particular, if the Euler equations hold, we have the conservation law

$$D_j P^j = 0.$$

The identity (20) is however still true even if A^I is not a solution. This has been very useful (cf. for instance the "Pohozhaev identity" for solutions of nonlinear elliptic equations).

Proof: The proof is a direct calculation (summation over I is understood in all formulae):

$$\begin{aligned}
&\sum_j D_j \sum_{H,K} (-1)^{|K|} C^I{}_{,H} L_I{}^{jHK}{}_{,K} \\
&= \sum_{H,K,j} (-1)^{|K|} [C^I{}_{,Hj} L_I{}^{jHK}{}_{,K} + C^I{}_{,H} L_I{}^{jHK}{}_{,jK}] \\
&= \sum_{|H|\geq 1, |K|\geq 0} (-1)^{|K|} C^I{}_{,H} L_I{}^{HK}{}_{,K} + \sum_{|H|\geq 0, |K|\geq 1} (-1)^{|K|+1} C^I{}_{,H} L_I{}^{HK}{}_{,K} \\
&= \sum_{|H|\geq 1} C^I{}_{,H} L_I{}^H - \sum_{|K|\geq 1} (-1)^{|K|} C^I L_I{}^K{}_{,K} \\
&= \sum_H C^I_{,H} L_I{}^H - C^I \sum_K (-1)^{|K|} D_K L_I{}^K \\
&= -D_h(L\xi^h) - C^I E_I,
\end{aligned}$$

QED. Note that we used the symmetry criterion in the last step.

4.3. First-order Lagrangians. We now write out the results for first-order Lagrangians of the form

$$L = L(A_{a,i}).$$

We let $L^{ak} = \partial L / \partial A_{a,k}$, and give the explicit form of the invariance criterion, and of the corresponding conservation laws.

The criterion for $\xi^i \partial_i + \phi_a \partial / \partial A_a$ to generate a variational symmetry reads:

$$L D_h \xi^h + L^{ak}(\phi_{a,k} - \xi^h{}_{,k} A_{a,h} - \xi^h{}_{,a} A_{h,k} - \xi^h{}_{,ak} A_h) = 0,$$

since

$$C_a = \phi_a - \xi^h A_{a,h} - \xi^h{}_{,a} A_h.$$

If this condition is satisfied, we obtain the conservation law:

$$\partial_i P^i = 0$$

with

$$P^i = L\xi^i + C_a L^{aj}.$$

where C_a is the characteristic given above. (This would also be the form of the conservation law for the most general first-order Lagrangian.) If $\phi_a = 0$, this reduces to

$$P^i = L\xi^i - (\mathcal{L}_\xi A_a) L^{ai}.$$

This last conservation law is sometimes justified without the use of prolongation (on scalars or vectors) as follows: The invariance of the Lagrangian is expressed by equating the variation consecutive to the change in the dependent variables induced

by the group action, with the variation of the Lagrangian scalar density (*viz.* with $D_i(L\xi^i)$). This gives

$$D_i(L\xi^i) = E^a(\mathcal{L}_\xi A_a) + D_j[(\mathcal{L}_\xi A_a)L^{aj}]. \tag{21}$$

Remarkably enough, taking into account the Euler equations $E^a = 0$, this is precisely the result obtained from the prolongation described here. This is procedure (iii) mentioned in §1.

4.4. Example 1: Linear Elasticity. The *Lagrangian* is:

$$L = \frac{1}{4}[\lambda({\sigma_k}^k)^2 + 2\mu\sigma^{ij}\sigma_{ij}] \tag{22}$$

where

$$\sigma_{ij} = \partial_i A_j + \partial_j A_i$$

and indices run from 1 to 3. We let $\sigma = {\sigma_k}^k$. Indices are raised and lowered by δ, and the A_a are the covariant components of the displacement.

$$L^{mi} = \lambda{\sigma_k}^k\delta^{mi} + 2\mu\sigma^{mi}.$$

The criterion for invariance simplifies due to the symmetry of L^{ak}: We only need

$$L{\xi^h}_{,h} + L^{ak}[\phi_{a,k} - {\xi^h}_{,k}(A_{h,a} + A_{a,h}) - {\xi^h}_{,ak}] = 0. \tag{23}$$

We now examine a few natural groups and list the corresponding conservation laws. In all these examples except dilatations, $\phi_a = 0$.

Translations: The generator is $a^m\partial_m$, where the a^m are constants, and

$$P^i = a^m T_{mi} \tag{24}$$

with

$$T^{mi} = \lambda\sigma(\frac{\sigma}{4}\delta^{mi} - \sigma^{mi} + A^{m,i}) + 2\mu(\frac{1}{4}\sigma^{jk}\sigma_{jk}\delta^{mi} - \sigma^{am}{\sigma_a}^i + \sigma^{hi}{A^m}_{,h}).$$

We recover the Eshelby tensor.

Rotations: Here, the generator is $\omega_{hk}(x^k\delta^{hi} - x^h\delta^{ki})\partial_i$, where ω_{hk} are constants satisfying $\omega_{hk} = -\omega_{kh}$, and $\partial_i P^i = 0$, with

$$P^i = \frac{1}{2}\omega_{hk}\{(x^h T^{ki} - x^k T^{hi}) + (A^h L^{ki} - A^k L^{hi})\}. \tag{25}$$

These conservation laws are the *sum* of two conservation laws obtained by Olver[19, 20], corresponding to the independent rotation of the space coordinates and the

displacement vector. In a physical rotation of an elastic body, the displacement does not remain constant in direction, but rather moves with the body—this is in fact a consequence of its vectorial nature. If it didn't, the body would undergo some additional stress. It is therefore normal that we should obtain both terms at once when considering rotations in x-space.

Dilatations: $\xi = x^i\partial_i + \alpha A_i \partial_{A_i}$. The criterion gives $\alpha = 1/2$. We now find

$$P^i = Lx^i + (\frac{1}{2}A_j + x^h A_{j,h})L^{ji}.$$

All the above conservation laws are in fact *linear combinations* of those obtained by Olver [19]; note that this paper also contains several *classification* results for the symmetries of the Euler equations.

4.5. Example 2: Maxwell's equations. The *Lagrangian* is now: $L = -F^2/4$, where

$$F^2 = F^{ij}F_{ij}; \quad F_{ij} = \partial_i A_j - \partial_j A_i.$$

Indices run from 0 to 3. Indices are raised an lowered using the Minkowski tensor η^{ij} with signature $(-,+,+,+)$. A short calculation gives $L^{ik} = F^{ik}$. The dependent variables are the A_j. If we consider the generator $\xi^i\partial_i$, the condition for this to generate a variational symmetry has a remarkable form: Since here $L^{ak} = -L^{ka}$, we find

$$L\xi^h{}_{,h} + L^{ak}[-\xi^h{}_{,k}(A_{a,h} - A_{h,a})] = 0,$$

or more simply

$$\theta^{kh}\xi_{h,k} = 0, \tag{26}$$

where

$$\theta^k{}_h = -\frac{1}{4}F^{ij}F_{ij}\delta^h_k + F^{ak}F_{ah}$$

is a symmetric,[2] traceless tensor known in the physics literature as the symmetric Maxwell energy-momentum tensor. Using symmetry, our condition further reduces to

$$\theta^{hk}(\xi_{h,k} + \xi_{k,h}) = 0. \tag{27}$$

It is therefore satisfied whenever

$$\mathcal{L}_\xi \eta = \lambda(x)\eta, \tag{28}$$

[2]that is, $\theta^{hk} = \theta^{kh}$.

where $\eta = \eta_{ab}\, dx^a \otimes dx^b$, that is, whenever ξ generates a group of conformal transformations. This very simple derivation is made possible by the present formalism. For earlier, related results via more complicated techniques, see [2, 3, 6].

Here again, the vectorial nature of the dependent variable has a direct physical meaning: under a Lorentz transformation, the electric and magnetic fields are scrambled and only the electromagnetic tensor has a geometric meaning, as a two-form (or a twice covariant tensor).

Let us now examine the conservation laws associated with the generators of the conformal group. Many of these can be found in the physics literature, as we point out, but their derivation has not always been integrated into the common framework of computations of symmetries. The present derivation seems to give the results in their simplest form. The reader is invited to compare the following with the following references: [2, 3, 6, 9]. Note also that P. J. Pohjanpelto (Dissertation, U. of Minnesota, 1989) has obtained a classification of symmetries, starting directly from Maxwell's equations.

Translations: The generator is $\xi = a^k \partial_k$, where the a^k are constants. The corresponding conservation law is $D_i P^i = 0$ with

$$P^i = a^k T^i{}_k,$$

where

$$T^i{}_k = -\frac{F^2}{4}\delta^i_k + F^{ia} A_{a,k} = \theta^i{}_k - F^{ih}\partial_h A_k.$$

These tensors have a physical interpretation: $T^i{}_k$ is known as the nonsymmetric canonical energy-momentum tensor, and $\theta^i{}_k$ is the symmetric Maxwell tensor encountered earlier. Their components are related to the Poynting vector, but, as we noted earlier, three dimensional language would be inadequate in this situation.

Rotations: The generator is $\xi = \frac{1}{2}\omega_{hk}(\eta^{hi}x^k - \eta^{ki}x^h)$, where, again, the ω_{hk} are constants. Note that $D_i\xi^i = 0$. The corresponding conservation law is $D_i P^i = 0$ with

$$\begin{aligned} P^i &= -\frac{1}{2}\omega_{hk}\{(T^{ki}x^h - T^{hi}x^k) + (A^h L^{ki} - A^k L^{hi})\} \\ &= \frac{1}{2}\omega_{hk} M^{hki}. \end{aligned}$$

$$\partial_i M^{hki} = 0; \quad M^{hki} = L^{hki} + S^{hki}.$$

Further,

$$M^{hki} = J^{hki} - \partial_l[F^{li}(x^h A^k - x^k A^h)].$$

Integrating the identity $\partial_i M^{hki}$ and using the divergence theorem gives the conservation of angular momentum. S^{hki} represents a spin, while L^{hki} represents an orbital angular momentum, and J^{hki} is a total angular momentum. M^{hki} is the total angular momentum (see [23] for the physical interpretation and applications).

Dilatations: The generator is $\xi = x^i \partial_i$.

$$P^i = Lx^i - (A_j + x^h A_{j,h}) F^{ji}.$$

Conformal transformations: The generator is $\xi^j \partial_j = a_i(2x^i x^j - x^k x_k \eta^{ij})\partial_j$, where the a^i are constants. The corresponding conservation law is $D_i P^i = 0$ with

$$P^i = L\xi^i - [\xi^h (A_{k,h} - A_{h,k}) - \partial_k(\xi^h A_h)] F^{ki}.$$

Recall that we proved in one blow at the beginning of this section that all of these generate variational symmetries.

5. Concluding Remark

The prolongation of group actions has been in the last twenty years a powerful tool for studying symmetries of differential equations and variational principles. Simplified accounts of Noether's theorems [12, 10, 24, 7, 15], although useful in obtaining familiar conservation laws of field theories, have neither the power nor the computational ease of use of the Lie-Noether procedure. We have extended here the latter to tensor-valued field variables, treating them not as multi-scalars but as tensor fields. Illustrative examples show the simplicity of use of this prolongation formula. It should provide the natural tool for studying symmetries of tensor field theories.

References

1. Anderson, D., Noether's theorem in generalized mechanics, *J. Phys., A* **6** (1973) 299–305.
2. Bateman, H., The transformations of electrodynamical equations, *Proc. London Math. Soc.*, **8** (1909) 223–264.
3. Bessel-Hagen, E., Über die Erhaltungssätze der Elektrodynamik, *Math. Ann.,* **84** (1921) 258–276.
4. Bluman, G. W., and Cole, J. D., *Similarity Methods for Differential Equations,* Appl. Math. Sci., **13**, Springer Verlag, New York, 1974.
5. Bogoliubov, N. N. and Shirkov, D. V., *Introduction to the Theory of Quantized Fields,* Wiley (Interscience), New York (1959).
6. Cunningham, E., The principle of relativity in electrodynamics and an extension thereof, *Proc. London Math. Soc.*, **8** (1909) 77–97.
7. Edelen, D. G. B., *Non-local Variations and Local Invariance of Fields,* American Elsevier Publ., New York, 1969.
8. Fushchich V. I. and Nikitin, A. G., *Symmetries of Maxwell's equations,* Reidel, Dordrecht, 1987.

9. Fushchich V. I., Shtelen, W. M. and Serov, N. I., *Symmetry Analysis and Exact Solutions of Equations of Nonlinear Mathematical Physics,* Kluwer, 1993.
10. Gel'fand, I. M. and Fomin, S. V., *Calculus of Variations,* Prentice-Hall, Englewood Cliffs, N.J., 1963.
11. Goldberg, J. N., Strong conservation laws and equations of motion in covariant field theories, *Phys. Rev.,* **89** (1953) 263. See also: Fletcher, J. G., Local conservation laws an generally covariant theories, *Rev. Mod. Phys.,* **32** (1960) 65–87, and Anderson, J. L., *Principles of Relativity Physics,* Academic Press, 1967.
12. Hill, E. L., Hamilton's principle and the conservation theorems of mathematical physics, *Rev. Mod. Phys.,* **23** (1951) 253–260.
13. Ibragimov, N. H. *Transformation Groups Applied to Mathematical Physics*, Mathematics and its Applications, Reidel, Dordrecht, 1985.
14. Lie, S., Über die Integration durch bestimmte Integrale von einer Klasse linear partieller Differentialgleichungen, *Arch. for Math,* **6** (1881) 328–268; see also *Gesammelte Abhandlungen* vol. 3, 492–523, B. G. Teubner, Leipzig, 1968.
15. Logan, J. D., *Invariant Variational Principles,* Academic Pres, New York, 1977.
16. Lovelock, D. and Rund, H., *Tensors, Differential Forms, and Variational Principles,* Wiley (Interscience), New York (1975).
17. Noether, E., Invariante Variationsprobleme, *Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl. II,* (1918) 235–237; translated by M. A. Tavel as 'Invariant variation problems,' in *Transport Theory and Stat. Phys.,* **1** (1971) 186–207.
18. Olver, P. J., Symmetry groups and group invariant solutions of partial differential equations, *J. Diff. Geo.*, **14** (1979) 497–542.
19. Olver, P. J., Conservation laws in Elasticity, I. General results, *Arch. Rat. Mech. Anal.,* **85**, 111–129 (1984); II. Linear homogeneous isotropic elastostatics, *ibid.*, **85**, 131–160 and **102** (1988) 385–387.
20. Olver, P. J., *Application of Lie Groups to Differential Equations,* Springer, New York, 1st edition, 1986; 2nd edition, 1993.
21. Olver, P. J., Direct reduction and differential constraints, preprint (1993).
22. Ovsiannikov, L. V., *Group Properties of Differential Equations,* Novosibirsk, Moscow, 1962 (in Russian), see also *Group Analysis of Differential Equations,* Academic Press, 1982.
23. Rohrlich, F., *Classical Charged Particles; foundation of their theory* Addison-Wesley, Reading, MA, 1965.
24. Schmutzer, E., *Relativistische Physik,* B. G. Teubner, Leipzig, 1968.
25. Trautman, A., Noether's equations and conservation laws, *Comm. Math. Phys.,* **6** (1967) 248–261.

School of Mathematics, University of Minnesota, Minneapolis, MN 55455-0487, USA