The integrability conditions in the inverse problem of the calculus of variations for second-order ordinary differential equations

W. Sarlet
Theoretical Mechanics Division, University of Gent
Krijgslaan 281, B–9000 Gent (Belgium)

M. Crampin
Department of Applied Mathematics, The Open University
Walton Hall, Milton Keynes MK7 6AA (U.K.)

E. Martínez
Departamento de Matemática Aplicada, Universidad de Zaragoza
María de Luna 3, E-50015 Zaragoza (Spain)

September 1, 1998

Abstract

A novel approach to a coordinate-free analysis of the multiplier question in the inverse problem of the calculus of variations, initiated in a previous publication [5], is completed in the following sense: under quite general circumstances, the complete set of passivity or integrability conditions is computed for systems with arbitrary dimension n. The results are applied to prove that the problem is always solvable in the case that the Jacobi endomorphism of the system is a multiple of the identity. This generalizes to arbitrary n a result derived by Douglas [6] for $n = 2$.

1 Introduction

In a fairly recent paper [5] we have presented a new geometrical approach to the solution of the hard part of the inverse problem of the calculus of variations, i.e. the problem of determining those systems of second-order ordinary differential equations, given initially in normal form, for which there is a non-singular multiplier matrix which will convert the given system to an equivalent one which is variational.

Our approach was very much tuned to the solution which Douglas gave to this problem [6] for the case of two degrees of freedom ($n = 2$). However, our methods have several marked advantages compared with the analytical ones used by Douglas. For example, whereas Douglas's calculations are very explicit and done in terms of the standard coordinate derivatives, our calculus is coordinate free: all computations are performed in terms of a local frame of vector fields which in general remains unspecified, but which in each specific case may be chosen optimally in relation to the natural geometrical structures which come with the given system. We thus reach a degree of generality which avoids the necessity of computing, for example, certain types of integrability conditions over and over again, either for the different subcases for each fixed number of degrees of freedom, or for different degrees of freedom.

The basic ingredients of our calculus are certain types of covariant derivative. There are two main sources which can be consulted for background, [19, 4], each giving a different perspective and general framework in which these derivations come to life in a natural way.

The geometrical setting for modelling time-dependent second-order ordinary differential equations is the first jet bundle $J^1\pi$ of a bundle $\pi : E \to \mathbb{R}$. The bundle E can be thought of as being diffeomorphic to a manifold of the form $\mathbb{R} \times M$, in which case $J^1\pi$ becomes diffeomorphic to $\mathbb{R} \times TM$, but it is not necessary to choose such a trivialization. A second-order system in normal form is a vector field Γ on $J^1\pi$ which satisfies $\langle \Gamma, dt \rangle = 1$ and which is annihilated by the contact forms $\theta^i = dx^i - v^i dt$. (Here (t, x^i, v^i) are local coordinates on $J^1\pi$ with t the coordinate on $\mathbb{R}$, x^i fibre coordinates on E, and v^i the jet coordinates.) It turns out that many other geometrical objects on $J^1\pi$ of interest are completely determined by tensor fields along the projection $\pi_1^0 : J^1\pi \to E$, and therefore have a smaller number of independent components than might have been expected. For example, the Poincaré-Cartan 2-form $d\theta_L$ determined by a Lagrangian L on $J^1\pi$, when expressed in terms of the local basis of 1-forms $\{dt, \theta^i, \eta^i\}$ adapted to the non-linear connection associated to the second-order differential equation field Γ corresponding to the Euler-Lagrange equations of L, turns out to have only $\frac{1}{2}n(n+1)$, as opposed to $n(2n+1)$, non-zero components; it is given by

$$d\theta_L = g_{ij}\, \eta^i \wedge \theta^j, \quad \text{with} \quad g_{ij}(t, x, v) = \frac{\partial^2 L}{\partial v^i \partial v^j}.$$

Here, the η^i are given by

$$\eta^i = dv^i - f^i dt + \Gamma^i_k \theta^k,$$

where the functions $f^i(t, x, v)$ are the right-hand sides of the second-order equations in normal form, and Γ^i_k denotes the connection coefficients of the non-linear connection determined by these equations, i.e.

$$\Gamma^i_k = -\tfrac{1}{2}\frac{\partial f^i}{\partial v^k}. \tag{1}$$

As such, $d\theta_L$ can be interpreted as coming from a kind of generalised metric tensor field

$$g = g_{ij}\,\theta^i \otimes \theta^j, \qquad g_{ij} = g_{ji}; \tag{2}$$

g is actually a tensor field along the projection π_1^0. The process which relates g to $d\theta_L$ is called the Kähler lift.

In fact, a second-order differential equation field Γ is itself determined by a vector field along π_1^0; it is simply the horizontal lift of the canonical vector field

$$\mathbf{T} = \frac{\partial}{\partial t} + v^i\,\frac{\partial}{\partial x^i} \tag{3}$$

along π_1^0. All such matters are explained in [19], which deals at length with the theory of derivations of forms along π_1^0.

A preliminary remark is in order here about the structure of $\mathcal{X}(\pi_1^0)$, the space of vector fields along π_1^0. The canonical vector field $\mathbf{T}$ along π_1^0 determines a natural splitting of $\mathcal{X}(\pi_1^0)$ in the form

$$\mathcal{X}(\pi_1^0) = \langle \mathbf{T} \rangle \oplus \overline{\mathcal{X}}(\pi_1^0);$$

in other words, every $X \in \mathcal{X}(\pi_1^0)$ has a representation in the form

$$X = X^0\mathbf{T} + \overline{X}, \qquad \overline{X} = X^i\frac{\partial}{\partial x^i}, \qquad X^0, X^i \in C^\infty(J^1\pi).$$

The vector fields $\{\mathbf{T}, \partial/\partial x^i\}$ constitute a basis for $\mathcal{X}(\pi_1^0)$ adapted to the splitting; the dual local basis of 1-forms along π_1^0 is given by $\{dt, \theta^i\}$.

The most interesting operations which emerge from the theory of derivations of forms along π_1^0 are vertical and horizontal covariant derivatives, denoted by D^V_X and D^H_X respectively, where X belongs to $\mathcal{X}(\pi_1^0)$. Being self-dual derivations of degree zero, their action extends to tensor fields along π_1^0 of arbitrary type. We have $\mathrm{D}^V_{\mathbf{T}} = 0$, whereas $\mathrm{D}^H_{\mathbf{T}}$ is an important derivation which has been given its own name and notation: we call it the dynamical covariant derivative and denote it by ∇. All further occurrences of vertical and horizontal covariant derivatives, therefore, will have for the argument X an element of $\overline{\mathcal{X}}(\pi_1^0)$. In fact, all tensor fields of interest in what follows will also have a non-zero action only on this submodule; moreover, if they are vector valued they will take values in $\overline{\mathcal{X}}(\pi_1^0)$ also. Therefore, in the rest of this paper all vector fields along π_1^0 will be elements of $\overline{\mathcal{X}}(\pi_1^0)$ unless it is stated explicitly otherwise; and to simplify notations, we will omit the overbar.

We are now ready to define the derivations under consideration by specifying their actions on the building blocks of all tensor fields along π_1^0, namely functions $F \in C^\infty(J^1\pi)$ and the bases of vector fields and 1-forms discussed earlier:

$$\nabla F = \Gamma(F), \quad \mathrm{D}^V_X F = X^V(F) = X^i\frac{\partial F}{\partial v^i}, \quad \mathrm{D}^H_X F = X^H(F) = X^i\left(\frac{\partial}{\partial x^i} - \Gamma^j_i\frac{\partial}{\partial v^j}\right)(F), \tag{4}$$

$$\nabla\left(\frac{\partial}{\partial x^j}\right) = \Gamma^i_j \frac{\partial}{\partial x^i}, \quad \mathrm{D}^V_X\left(\frac{\partial}{\partial x^j}\right) = 0, \quad \mathrm{D}^H_X\left(\frac{\partial}{\partial x^j}\right) = \left(X^k \frac{\partial \Gamma^i_j}{\partial v^k}\right)\frac{\partial}{\partial x^i}, \tag{5}$$

and by duality

$$\nabla\theta^i = -\Gamma^i_j\theta^j, \quad \mathrm{D}^V_X\theta^i = -X^i dt, \quad \mathrm{D}^H_X\theta^i = -\left(X^k\frac{\partial\Gamma^i_j}{\partial v^k}\right)\theta^j. \tag{6}$$

To this we add for completeness that

$$\nabla\mathbf{T} = 0, \quad \mathrm{D}^V_X\mathbf{T} = X, \quad \mathrm{D}^H_X\mathbf{T} = 0,$$

and dually

$$\nabla dt = 0, \quad \mathrm{D}^V_X dt = 0, \quad \mathrm{D}^H_X dt = 0.$$

That D^V_X, D^H_X and ∇ are indeed the operations of interest when analysing properties of second-order equations has become apparent in applications such as the study of complete separability in [11, 3], but is perhaps even more obvious from further theoretical foundations. The set $\mathcal{X}(\pi^0_1)$ is in fact the set of sections of the pull back bundle $\pi^{0*}_1(\tau_E)$ over $J^1\pi$, where τ_E is the tangent bundle projection $TE \to E$. It was shown in [4] that a second-order differential equation field Γ further determines a linear connection on $\pi^{0*}_1(\tau_E)$, and that various tensor fields encountered in the study of derivations of forms along π^0_1 acquire an elegant interpretation in this framework. To explain the direct link between [19] and [4], it suffices to point out that the covariant derivative operator D_ξ associated with the linear connection on $\pi^{0*}_1(\tau_E)$ can be written as

$$\mathrm{D}_\xi = \mathrm{D}^V_X + \mathrm{D}^H_Y + \langle\xi, dt\rangle\nabla. \tag{7}$$

Here, ξ is a vector field on $J^1\pi$, and X and Y are the elements of $\overline{\mathcal{X}}(\pi^0_1)$ uniquely determined by the decomposition $\xi = X^V + Y^H + \langle\xi, dt\rangle\Gamma$.

We now come back to the inverse problem. One virtue of our approach is that by using the kind of geometrical objects and operations referred to above, one arrives at a formulation of the Helmholtz conditions which reveals their geometrical content but at the same time stays very close to the analytical formulations which can be found e.g. in the work of Douglas [6] and in [18]. As we said at the beginning of this section, we made considerable progress in [5] towards understanding the geometry behind Douglas's solution of the inverse problem and its possible generalization to more than two degrees of freedom. However, at that time we did not complete the discussion of integrability conditions (the computation of 'alternants' in the terminology of the Riquier theory) at the same level of generality as we started it. It is precisely at this point that the new insights gained in [4] provide a breakthrough. Indeed, much of the work of computing further alternants can be clarified by the fact that we now have the so-called second Bianchi identities of the linear connection on $\pi^{0*}_1(\tau_E)$ at our disposal. There is another aspect of Riquier theory which we did not even touch in [5]. There were good reasons for that because, as our present contribution will show, it is a topic which cannot be discussed at the same level of coordinate free analysis and certainly not without entering the details of a classification into numerous subcases.

To be more precise, the Riquier theory [17], in the form we shall need it, depends on two basic notions, those of the orthonomicity and of the passivity of a system of partial differential equations. We take our definitions of these terms from the resumé of Riquier theory given by Douglas in his paper on the inverse problem, adjusting the wording as necessary to fit our exposition. Other useful descriptions, in the context of setting up computer algebra algorithms, can be found in [16, 22]. Further general references are of course the works of Janet [8, 9].

A system of partial differential equations is orthonomic if

1. it is solved for *distinct* derivatives of the unknown functions: these derivatives, each of which appears as the single term on the left-hand side of one of the equations, together with all their derivatives, are called the principal derivatives
2. a way of ordering the partial derivatives has been assigned so that in each equation the derivative appearing on the left-hand side has a higher order than any derivative which appears in the corresponding right-hand side
3. no right-hand side contains any principal derivative.

A derivative which is not principal is said to be parametric.

Repeated differentiation of the different equations of the given system will produce different expressions for the same principal derivative in terms of parametric derivatives, which will possibly lead to non-trivial relations between the parametric derivatives. Such relations are usually called integrability conditions, but Riquier (and Douglas) occasionally refer to them also as 'passivity conditions'. If in fact no such relations between parametric derivatives can be derived as the result of repeated differentiation the system is said to be passive.

The fundamental result of Riquier theory which we need is that every system of partial differential equations which is passive and orthonomic is formally integrable.

Needless to say, more precision is required when it comes to describing the details of the ordering process (for example, the axioms to be satisfied by what Riquier calls the 'cotes' assigned to all variables, the unknown functions and their derivatives). We will spell details out when we need them in the specific situation of the application discussed in Section 4. In fact, a very important feature of our approach to the inverse problem is the following. The standard Riquier theory, as sketched above, clearly is not geometrical: writing equations in orthonomic form and computing alternants to obtain integrability conditions are procedures that normally go hand in hand and depend, among other things, on the choice of principal derivatives. We will present a geometrical way of using these ideas in the context of the inverse problem. More precisely, we will carry out a large part of the integrability analysis of the Helmholtz conditions in a completely coordinate free way, which means that the process of repeated differentiation of equations and looking for new non-trivial relations is carried out by intrinsic operations and in a way which is independent of the selection of principal derivatives which will have to made at some later stage, when one is forced to enter a classification into different subcases.

The plan of the paper is as follows. In the next section, we recall the geometrical formulation of the Helmholtz conditions in the form of coordinate independent equations to be satisfied by the generalized metric tensor field g introduced above. To prepare the stage for the full integrability analysis, the first few steps in that process are explained up to the point where all purely algebraic restrictions on g have been obtained; the proof that there are no others is given in Appendix A. In Section 3 we show how the process of computing prolongations of the equations, with the possibility of obtaining new passivity requirements, terminates by virtue of the Bianchi identities referred to above, at least when there is no degeneracy (to be defined later on). The calculations required in the final stage are explained in Appendix B, though since they are quite tedious the explanation is brief. As an application and illustration of the generality of these results we prove in Section 4 that for the analogue of Douglas's Case I in arbitrary dimension, i.e. the case where the Jacobi endomorphism Φ is a multiple of the identity, the inverse problem always has solutions. This is the stage where the aspect of orthonomicity has to be addressed. Accordingly, we shall start this discussion from a coordinate representation of the equations, but we shall prove in fact that also this part of the Riquier analysis can be given a geometrical flavour. Indeed, the conclusion of our detailed investigation will be the following: the local frame of vector fields on $J^1\pi$ which is imposed upon us by the geometry of the problem (i.e. by the non-linear connection coming from the given second-order equation field) can replace the local basis of coordinate derivatives also in the process of checking orthonomicity. In Section 5 we set up a formula for the degree of arbitrariness in this solution, which again is the direct generalization of Douglas's claim for $n = 2$.

A few more words of introduction are in order. The first proof of the existence of a Lagrangian in Case I for arbitrary n that we are aware of is due to Anderson and Thompson [1]; their result for the number of arbitrary functions in the solution, however, differs from ours, and in fact is in conflict also with the claim of Douglas for $n = 2$. While the preparation of this paper was in progress, we have been informed of very similar results obtained by Grifone and Muzsnay [7, 13]. Their approach seems to have been worked out for autonomous differential equations only so far. It is very different from ours in that their analysis starts directly from the Euler-Lagrange operator, rather than from the Helmholtz conditions for the multiplier matrix g. Grifone and Muzsnay also prove the existence of a Lagrangian for Case I in arbitrary dimension, and their method seems to lead to the same number of arbitrary functions in the solution as in [1][1]. There are many conflicting statements in the literature about the arbitrariness of solutions of formally integrable systems, because such results depend very much on the representation of the solution (cf. [20]). We will try to clarify the origin of the different results obtained for the inverse problem. Finally, we mention a more fundamental issue: the term "integrability condition" does not seem to have a universal meaning in the different theories for partial differential equations which are now available. Some of the passivity requirements which we face in our Riquier-Janet approach would not appear as integrability conditions in the Spencer-Goldschmidt approach to formal integrability of partial differential equations. We dare to hope that the way we sketch the difference between integrability and passivity conditions here will also be of some interest outside the present context.

2 An efficient formulation of the Helmholtz conditions

In purely analytical terms, the problem we are investigating is the characterization of those systems of second-order ordinary differential equations of the form

$$\ddot{x}^j = f^j(t, x, \dot{x})$$

for which a non-singular symmetric matrix $g_{ij}(t, x, \dot{x})$ can be constructed, at least locally, such that the equivalent system $g_{ij}(\ddot{x}^j - f^j) = 0$ is a set of Euler-Lagrange equations. Necessary and sufficient conditions for the existence of such a multiplier matrix, known as the Helmholtz conditions, make their appearance in the literature in many different disguises. A good list of references for the state of the art prior to 1990 can be found in [12]. The better analytical formulations of the Helmholtz conditions (so we believe) are those which make reference to another matrix $\Phi^i_j(t, x, \dot{x})$ associated with the given system and defined by (the minus sign is a matter of convention):

$$\Phi^i_j = -\frac{\partial f^i}{\partial x^j} - \Gamma^i_k \Gamma^k_j - \Gamma(\Gamma^i_j). \tag{8}$$

This matrix is already present in the paper of Douglas, but seems subsequently to have been forgotten, until it was assigned an important role again in [18]. The minimal set of Helmholtz conditions reads:

$$\Gamma(g_{ij}) = g_{ik}\Gamma^k_j + g_{jk}\Gamma^k_i, \qquad \frac{\partial g_{ij}}{\partial v^k} = \frac{\partial g_{ik}}{\partial v^j}, \qquad g_{ij}\Phi^j_k = g_{kj}\Phi^j_i. \tag{9}$$

We next give a geometrical version of these equations. We know already that the object g we are looking for is a symmetric type (0,2) tensor field along π^0_1, of the form (2). As for

[1]Private communication.

the matrix Φ^i_j, it represents the components of a type (1,1) tensor field (a vector-valued 1-form) along π^0_1 of the form:

$$\Phi = \Phi^i_j \, \theta^j \otimes \frac{\partial}{\partial x^i}. \tag{10}$$

This tensor field plays an important part in all our recent contributions to the study of second-order equations. It is called the Jacobi endomorphism; this terminology was introduced, for autonomous equations, in [10]. The geometrical derivation of the Helmholtz conditions in [19] provides a direct coordinate-free transcription of the conditions (9), in the following form:

$$\begin{aligned} \nabla g &= 0 & (11)\\ \mathrm{D}^V g\,(Z,X,Y) &= \mathrm{D}^V g\,(Y,X,Z) & (12)\\ g\,(\Phi(X),Y) &= g\,(\Phi(Y),X). & (13) \end{aligned}$$

Here, we have introduced the vertical covariant differential D^V, which increases the covariant order of an arbitrary tensor field T by one, and is simply defined by:

$$\mathrm{D}^V T(Z,\ldots) = \mathrm{D}^V_Z T(\ldots).$$

There is, however, a technical point which needs to be clarified. Due to certain choices which were made in developing the theory of derivations in [19], $\mathrm{D}^V_X \theta^i$ is not zero and in fact introduces terms in dt into the picture (see (6)). Whenever we apply such operations in this paper, however, the resulting tensor field will be restricted to act on $\overline{\mathcal{X}}(\pi^0_1)$ only, which means that the terms in dt can be ignored. Let us agree therefore to incorporate this feature into the definition of the operator D^V, so that for all practical purposes we can take $\mathrm{D}^V\theta^i = 0$. For the same reason the properties of derivations that we need are essentially identical to those in the autonomous theory described in [10], which is what we exploited in [5].

The computation of integrability conditions for the equations for g is a fairly delicate business and cannot be decoupled completely from a classification into subcases derived from the algebraic restrictions on g. In his paper Douglas had to carry out a tedious integrability analysis for each of the many subcases in his classification. What we propose to do (in the next section) is to push the computation of integrability conditions as far as possible before even starting a classification into subcases, and to do this in a coordinate-free way. These general results will then be available for each case and will only require further integrability computations when there is a certain form of degeneracy. With respect to the modern theory of formal integrability of partial differential equations (see e.g. [15, 2]), one has to interpret what we are doing with some care. The term "integrability condition" does not seem to have the same meaning in the different approaches. What we are doing here, following Riquier, is to check "passivity" of the equations. New conditions have to be added on to the system, in this approach, whenever one encounters restrictions on what were previously parametric derivatives. Such restrictions may well be of the same order of differentiation as the prolongation under consideration. In the other approach, integrability conditions will turn up when a prolongation does not project onto the equations of the previous stage, so they will always be of lower order than the equations in the prolongation under consideration.

For the purpose of setting up some basic computational rules, we explain here the start of the integrability analysis. What we need in the first place are commutator properties of the derivations involved. They were already used in [5], but the sometimes tedious computations can be performed more efficiently if one learns to work directly with covariant differentials such as D^V, rather then with the degree zero derivation D^V_X for arbitrary X. Thus, the commutator relation

$$[\nabla, \mathrm{D}^V_X] = \mathrm{D}^V_{\nabla X} - \mathrm{D}^H_X \tag{14}$$

translates into

$$[\nabla, \mathrm{D}^V] = -\mathrm{D}^H, \tag{15}$$

D^H being defined in a way similar to D^V.

Comparing the D^V prolongation of (11) to the ∇ prolongation of (12) using (15) brings us to the integrability requirement

$$\mathrm{D}^H g\,(Z, X, Y) = \mathrm{D}^H g\,(Y, X, Z) \tag{16}$$

which has to be added to the three equations we start from. Note that these are true "integrability conditions" in any sense of the word, because they are new first-order equations resulting from prolonging the given system to second order (i.e. they are obstructions for the prolongation projecting onto the original system).

Compatibility of (11) with (13) leads to the new algebraic requirement

$$g\,(\nabla\Phi(X), Y) = g\,(\nabla\Phi(Y), X).$$

This process can obviously be continued to produce the infinite hierarchy of requirements

$$g\,(\nabla^k\Phi(X), Y) = g\,(\nabla^k\Phi(Y), X) \tag{17}$$

for all $k = 1, 2, 3, \ldots$, first discussed in [18]. How many of these conditions will have to be imposed will of course depend on the dimension n and on a case by case classification related to the structure of Φ.

A second hierarchy of algebraic conditions, first derived in [18], follows from compatibility of (12) and (13). This works as follows. The D^V prolongation of (13) is the equation

$$\mathrm{D}^V g\,(Z, \Phi(X), Y) - \mathrm{D}^V g\,(Z, \Phi(Y), X) = g\,(\mathrm{D}^V\Phi(Z, Y), X) - g\,(\mathrm{D}^V\Phi(Z, X), Y). \tag{18}$$

As before, the first objective is to see whether algebraic consequences of such prolongations can lead to new equations of lower order in the derivatives, which here means that one should look for possible combinations of these relations which eliminate the first-order terms in g. In view of the symmetry of g and the condition (12), the tensor field $\mathrm{D}^V g$ is symmetric in its three arguments. As a result, writing two more versions of this equation with cyclic permutations of X, Y, Z and adding them up, the left-hand side identically vanishes, yielding the new algebraic condition ($\sum$ meaning cyclic sum, here and below)

$$\sum g\,(\mathrm{D}^V\Phi(X,Y) - \mathrm{D}^V\Phi(Y,X), Z) = 0.$$

The geometrical meaning of this relation becomes more apparent if we recall the identity (see [19, 4])

$$\mathrm{D}^V\Phi(X,Y) - \mathrm{D}^V\Phi(Y,X) = 3\,R(X,Y), \tag{19}$$

where R is the curvature tensor (in our context a vector-valued 2-form along π_1^0) of the non-linear connection associated to the given second-order system Γ. (To be more precise, it is the restriction of the curvature tensor to $\overline{\mathcal{X}}(\pi_1^0)$, which was denoted by $\tilde{R}$ in the above cited references, but we will omit the tilde here.) In conclusion, the new algebraic restriction on g can be written as

$$\sum g\,(R(X,Y), Z) = 0. \tag{20}$$

Acting on this repeatedly with ∇ and taking account of (11) gives rise to the hierarchy of conditions

$$\sum g\,(\nabla^k R(X,Y), Z) = 0, \qquad \forall k. \tag{21}$$

Of course, having augmented the original Helmholtz conditions with the D^H condition (16), we should consider also its compatibility with the algebraic relations (13). However, this will not produce anything which has not already been mentioned, because we have the property

$$\mathrm{D}^H\Phi(X,Y) - \mathrm{D}^H\Phi(Y,X) = \nabla R(X,Y). \tag{22}$$

But we have now again augmented the original conditions with an extra algebraic restriction (20) (and its ∇ derivatives). Therefore, we have to consider the possibility that even more algebraic conditions might be created from suitable combinations of the D^V and D^H prolongations of these. The method of investigation is the same: one writes down the prolongation, for example

$$\begin{aligned}\mathrm{D}^V g\,(U, R(X,Y), Z) + \mathrm{D}^V g\,(U, R(Y,Z), X) + \mathrm{D}^V g\,(U, R(Z,X), Y) = \\ - g\,(\mathrm{D}^V R(U,X,Y), Z) - g\,(\mathrm{D}^V R(U,Y,Z), X) - g\,(\mathrm{D}^V R(U,Z,X), Y),\end{aligned} \tag{23}$$

and then makes the appropriate combination of cyclic permutations which, when added together, will eliminate all the derivative terms. The result is the algebraic relation

$$\begin{aligned} g\left(\sum \mathrm{D}^V R(U,X,Y), Z\right) - g\left(\sum \mathrm{D}^V R(X,Y,Z), U\right) \\ + g\left(\sum \mathrm{D}^V R(Y,Z,U), X\right) - g\left(\sum \mathrm{D}^V R(Z,U,X), Y\right) = 0,\end{aligned}$$

and similarly for D^{H}. Both of the resulting relations, however, are identically satisfied in view of the properties

$$\sum \mathrm{D}^{V} R(X,Y,Z) \;=\; 0 \tag{24}$$

$$\sum \mathrm{D}^{H} R(X,Y,Z) \;=\; 0. \tag{25}$$

It is worthwhile to observe that the properties (19), (22), (24) and (25) are among the first Bianchi identities satisfied by the curvature of the linear connection on $\pi_1^{0*}(\tau_E)$.

This is still not the end of the line in terms of the possibility of creating further independent algebraic conditions. Indeed, one can repeat the manipulations we have considered so far for the compatibility between the D^{V} (or D^{H}) equation and each of the algebraic relations in the lists (17) and (21) for $k > 0$. The commutator relation (15) shows that such an analysis will rely essentially on what happens when we interchange ∇ and D^{H}. To find the commutator of ∇ and D^{H} we take the formula for $[\nabla, \mathrm{D}^{H}_{X}]$. As we know from [10, 19], the i_* part of the action of a self-dual derivation of degree zero on 1-forms translates by duality into a corresponding algebraic derivation of vector fields. Therefore, in all generality we have

$$[\nabla, \mathrm{D}^{H}_{X}] = \mathrm{D}^{H}_{\nabla X} + \mathrm{D}^{V}_{\Phi(X)} + 2\mu_{i_X R} - \mu_{\mathrm{D}^{V}_{X}\Phi}, \tag{26}$$

where for any type (1,1) tensor field A, μ_A contains two terms: $\mu_A = a_A - i_A$. The action of these terms on a type $(1,p)$ tensor field T, or vector-valued covariant p-tensor, is given by

$$i_A T(X_1,\ldots,X_p) \;=\; \sum_{i=1}^{p} T(X_1,\ldots,A(X_i),\ldots,X_p), \tag{27}$$

$$a_A T(X_1,\ldots,X_p) \;=\; A(T(X_1,\ldots,X_p)). \tag{28}$$

When μ_A acts on a type $(0,p)$ tensor field T, however, the term $a_A T$ is absent. (In the final section of [5] a formula for $[\nabla, \mathrm{D}^{H}_{X}]$ was given which contained no a_A term; since the operators were used there to act only on covariant tensor fields such as g, this term was not needed.) Passing from the derivation D^{H}_{X} to the differential D^{H} and making use of (19), one obtains

$$[\nabla, \mathrm{D}^{H}] = \Phi \lrcorner\, \mathrm{D}^{V} - \mu_{\Psi}, \tag{29}$$

where Ψ is a type (1,2) tensor field along π_1^0, defined by

$$\Psi(X,Y) = R(X,Y) + \mathrm{D}^{V}\Phi(Y,X), \tag{30}$$

and the action of the operator μ_Ψ on vector-valued covariant tensor fields is given by

$$\mu_\Psi T(X,\ldots) = \mu_{X \lrcorner \Psi} T(\ldots). \tag{31}$$

By making use of this commutator relation one can show that no further algebraic restrictions can be obtained apart from those given in (17) and (21). To be precise, we state the following proposition, whose proof is given in Appendix A.

Proposition 1 *Assume g satisfies the differential conditions (11), (12), (16) and the algebraic conditions (17), (21) up to order $k = l$. Then the next level of integrability conditions between the differential and the algebraic equations generates the conditions (17), (21) for $k = l + 1$.*

3 The complete set of passivity conditions when there is no degeneracy

Summarizing the results of the previous section we have, so far, augmented the original Helmholtz conditions with the D^H equations (16), and we have found all generic algebraic conditions on g. The next stage will be to look for the possible integrability or passivity conditions for the extended system which may come from other prolongations to second-order equations for g. Some of these computations were done in [5] already, but will be repeated here with the more compact formulations that the use of covariant differentials makes possible.

Integrability conditions would occur if combinations of the prolongations to second-order equations turned out to produce new equations of lower order. Following the Riquier approach, on the other hand, other so-called passivity conditions may arise which are of the same order as the prolongation. In the process of bringing the system of partial differential equations into orthonomic or standard form (see the Introduction), the equations have to be written in such a way that a principal derivative appears in the left-hand side, while the right-hand side contains no such derivatives and no derivatives of higher order than the left-hand side. It may happen that all the terms of highest derivative order in an equation obtained by prolongation were previously designated parametric derivatives. One of them has then to be promoted to the rank of principal derivative, and the equation under consideration has to be added to the system.

Our purpose is to obtain the true integrability conditions, if any; and to obtain also, in advance of any choice of ordering, all the relations which can give rise to possible passivity conditions when some suitable choice of ordering has been made. In other words, our analysis is designed to be valid irrespective of the ordering which one may select later on in a case by case study.

The computations which follow are concerned with g and its covariant differentials only, which means that we will need commutator properties such as (29) only for the action on type $(0,p)$ tensor fields, so that the a term in μ has no effect.

Acting with ∇ on (16) and making use of the commutator (29), we obtain first:

$$(\Phi\lrcorner\, \mathrm{D}^V g + i_\Psi g)(Z,X,Y) = (\Phi\lrcorner\, \mathrm{D}^V g + i_\Psi g)(Y,X,Z).$$

Using the prolongation (18) to eliminate the terms in $\mathrm{D}^V g$, all terms become algebraic again and using the definition of Ψ, it turns out that we recover the curvature condition (20). Here we make use of the fact that

$$\Psi(X,Y) - \Psi(Y,X) = -R(X,Y). \tag{32}$$

The next stage will involve second-order covariant differentials. From the definition of D^V and D^H, it follows that for any covariant tensor field T,

$$\mathrm{D}^2\mathrm{D}^1 T(U,V,X,\ldots) = \mathrm{D}^2_U\mathrm{D}^1_V T(X,\ldots) - \mathrm{D}^1_{\mathrm{D}^2_U V} T(X,\ldots), \tag{33}$$

where D^2 and D^1 stand for either D^H or D^V. For the compatibility analysis of D^V and D^H prolongations of (12) and (16), we need the commutator properties of all possible combinations of covariant derivations D^V_X and D^H_Y. We used them extensively already in the final section of [5] and write them here directly in the form of properties of the corresponding covariant differentials. With T again representing any covariant tensor field, we have the identities:

$$\begin{aligned}
\mathrm{D}^V\mathrm{D}^V T(X,Y,Z,\ldots) - \mathrm{D}^V\mathrm{D}^V T(Y,X,Z,\ldots) &= 0 && (34)\\
\mathrm{D}^H\mathrm{D}^H T(X,Y,Z,\ldots) - \mathrm{D}^H\mathrm{D}^H T(Y,X,Z,\ldots) &= \\
\mathrm{D}^V T(R(X,Y),Z,\ldots) &- i_{\mathrm{Rie}(X,Y)} T(Z,\ldots) && (35)\\
\mathrm{D}^V\mathrm{D}^H T(X,Y,Z,\ldots) - \mathrm{D}^H\mathrm{D}^V T(Y,X,Z,\ldots) &= -i_{\theta(X,Y)} T(Z,\ldots). && (36)
\end{aligned}$$

In effect these relations define the curvature tensor of the linear connection on $\pi_1^{0*}(\tau_E)$ (cf. [4]), so it is best to regard the last two here as defining the tensor fields Rie and θ, which are type (1,3) tensor fields along π_1^0 (or better, covariant 2-tensors which take type (1,1) tensors as their values). As part of the first Bianchi identities satisfied by this curvature, one finds that θ is symmetric in all its vector arguments, i.e.

$$\theta(X,Y)Z = \theta(Y,X)Z = \theta(X,Z)Y; \tag{37}$$

and also that

$$\mathrm{Rie} = -\mathrm{D}^V R \quad \text{or} \quad \mathrm{Rie}(X,Y)Z = -\mathrm{D}^V_Z R(X,Y). \tag{38}$$

For completeness, we should add here again that this Rie is the restriction to $\overline{\mathcal{X}}(\pi_1^0)$ of the tensor with the same name in [19].

At this point, it is necessary to go through some of the calculations in detail, to show how the calculus works in practice. We take the prolongation to second-order horizontal derivatives as an example. Acting on an equation of the form (16) with another (arbitrary) horizontal covariant derivative yields a prolonged equation of the form

$$\mathrm{D}^H\mathrm{D}^H g\,(U,X,Y,Z) = \mathrm{D}^H\mathrm{D}^H g\,(U,Z,Y,X).$$

Interchanging the names of U and X, we could have written just as well:

$$\mathrm{D}^H\mathrm{D}^H g\,(X,U,Y,Z) = \mathrm{D}^H\mathrm{D}^H g\,(X,Z,Y,U).$$

The left-hand sides of these two representations are related through the identity (35). Replacing their difference by the lower order terms in (35) might give rise to new restrictions on parametric derivatives, i.e. to passivity conditions, or even integrability conditions if the second-order terms coming from the right-hand sides could also be eliminated in the same process.

Making use of (35) to replace both sides of the second equation and then substituting the term $\mathrm{D}^H\mathrm{D}^H g\,(U,X,Y,Z)$ on the left via the first equation, we obtain the intermediate relation

$$\begin{aligned}\mathrm{D}^H\mathrm{D}^H g\,(U,Z,Y,X) - \mathrm{D}^H\mathrm{D}^H g\,(Z,X,Y,U) = \\ -\mathrm{D}^V g\,(R(X,U),Y,Z) - g\,(\mathrm{D}^V R(Y,X,U),Z) - g\,(Y,\mathrm{D}^V R(Z,X,U)) \\ + \mathrm{D}^V g\,(R(X,Z),Y,U) + g\,(\mathrm{D}^V R(Y,X,Z),U) + g\,(Y,\mathrm{D}^V R(U,X,Z)).\end{aligned}$$

Now in all such computations we can freely add or subtract straight prolongations. The idea is that such terms will cancel out anyway when, for a given ordering, principal derivatives are substituted for in terms of parametric ones. Exactly which ordering is chosen is irrelevant for the present argument. Here, for example, we can interchange X and U in the second term on the left again, and then the two second-order derivative terms will be replaced by terms of lower order in view of (35). It is at this point in calculations of this type that we might find a new integrability or passivity condition. In the present case, however, it turns out that we get a combination of terms of the form $g\,(Y,\sum \mathrm{D}^V R(X,U,Z))$, which is identically zero in view of (24), whereas all the remaining terms precisely make up the D^V prolongation (23) of the curvature condition (20). We conclude that no new integrability or passivity requirements can arise here.

It is immediately clear that no conditions will arise for D^V prolongations of (12), in view of the lack of “vertical curvature” in the linear connection, as expressed by the zero in the right-hand side of (34).

The situation is different, however, for mixed horizontal and vertical derivatives. Consider the prolongation

$$\mathrm{D}^H\mathrm{D}^V g\,(U,X,Y,Z) = \mathrm{D}^H\mathrm{D}^V g\,(U,Z,Y,X)$$

of an equation of type (12) on the one hand, and the prolongation

$$\mathrm{D}^V\mathrm{D}^H g\,(X,U,Y,Z) = \mathrm{D}^V\mathrm{D}^H g\,(X,Z,Y,U)$$

of an equation of type (16) on the other. It is clear that if we use the defining relation (36) of the tensor θ for replacements in the second equation, we will create on the left-hand side a term which is identical to the left-hand side of the first equation. If we continue the same procedure explained above line by line, and make use of the symmetry properties (37) of θ, we can write the end result in the following form

$$\mathrm{D}^H\mathrm{D}^V g\,(U,Z,Y,X) - \mathrm{D}^H\mathrm{D}^V g\,(Z,U,Y,X) + g\,(\theta(Y,X)Z,U) - g\,(\theta(Y,X)U,Z) = 0. \quad (39)$$

The second-order derivative terms involve the same components of g (the third and fourth arguments are the same for both). Also, not all components of g can have all their D^H and D^V derivatives in the list of principal derivatives at the level of the first-order equations. Therefore, whatever ordering process one happens to select, there will be equations coming from (39) which establish a relation between parametric derivatives and therefore have to be added on as passivity conditions. An explicit example of this phenomenon is given in the next section.

The second-order operator in (39) may be expressed in a number of different equivalent ways by interchanging D^V and D^H derivatives through (36). We introduce $A(X,Y)$ as shorthand notation for it: thus

$$\begin{aligned}(A(X,Y)T)\,(Z,\ldots) &= \mathrm{D}^V\mathrm{D}^H T(X,Y,Z,\ldots) - \mathrm{D}^V\mathrm{D}^H T(Y,X,Z,\ldots) &(40)\\ &= \mathrm{D}^H\mathrm{D}^V T(Y,X,Z,\ldots) - \mathrm{D}^H\mathrm{D}^V T(X,Y,Z,\ldots) &(41)\\ &= \mathrm{D}^V\mathrm{D}^H T(X,Y,Z,\ldots) - \mathrm{D}^H\mathrm{D}^V T(X,Y,Z,\ldots) + i_{\theta(X,Y)}T(Z,\ldots). &(42)\end{aligned}$$

It is clear that $A(X,Y)$ is skew-symmetric in its two arguments. We further define an operator $\mathcal{A}(X,Y)$, specifically for the action on g. This is also skew-symmetric in X and Y, and is given by

$$(\mathcal{A}(X,Y)g)\,(U,V) = (A(X,Y)g)\,(U,V) + g\,(\theta(U,V)X,Y) - g\,(X,\theta(U,V)Y), \qquad (43)$$

so that the conditions (39) acquire the form $(\mathcal{A}(X,Y)g)\,(U,V) = 0$, for all X,Y,U,V. Observe that the right-hand side of (43) is also manifestly symmetric in U and V.

How many such passivity conditions will have to be imposed will depend on the choice of principal derivatives. If one were to write down an expression like (43) with leading terms which are principal derivatives, these would have to be replaced by a substitution from the first-order equations. Such an operation amounts to interchanging the two middle arguments in the leading terms. The self-consistency of the second-order conditions in relation to this operation is illustrated by the following result.

Proposition 2 *We have the property*

$$(\mathcal{A}(X,Y)g)\,(U,V) = (\mathcal{A}(X,U)g)\,(Y,V) \quad \text{modulo prolongations.} \qquad (44)$$

Proof In expressions like $\mathrm{D}^V g\,(X,Y,Z)$ or $\mathrm{D}^H g\,(X,Y,Z)$ and their prolongations, we can write the arguments X,Y,Z in any order we like in view of the symmetry of g and the conditions (12) and (16). Using the representation (42) of the A-operator, collecting all the algebraic terms and using the symmetry of g and θ, we can write

$$\begin{aligned}(\mathcal{A}(X,Y)g)\,(U,V) &= (\mathrm{D}^V\mathrm{D}^H g - \mathrm{D}^H\mathrm{D}^V g)(X,Y,U,V) + i_{\theta(X,Y)}g\,(U,V)\\ &\quad + i_{\theta(X,U)}g\,(Y,V) - i_{\theta(Y,U)}g\,(X,V).\end{aligned}$$

In the derivative terms, Y and U can be interchanged by using prolongations, whereas the algebraic part is now manifestly symmetric in Y and U. The result follows. □

It is now clear that, given a particular ordering of the first-order equations, the independent second-order conditions will come from those expressions (43) in which none of the second-order terms is a derivative of a term which figures in the left-hand side of first-order D^V or D^H equations.

At this point our set of equations has been augmented with a number of second-order equations, and the process of computing alternants has to be started all over again. The computations involved are extremely laborious. We give a sketch of what is involved in carrying them out in Appendix B, and content ourselves here with a statement of the results.

To compute the alternant of (39) (in any of its appearances) and (11), we apply ∇ to (39) and use the commutator properties (15) and (29) to bring ∇ inside, taking advantage in the end of the fact that $\nabla g = 0$. In the long process of checking whether this produces anything new, one has to make use of a large number of properties of tensors already introduced and prolongations of the equations we already have on our list. This will not be sufficient, however, to prove that no new conditions are born. What one needs in addition is the property

$$\nabla\theta(X,Y)Z = -\tfrac{1}{3}\sum \mathrm{D}^V\mathrm{D}^V\Phi(X,Y,Z). \tag{45}$$

This is one of the second Bianchi identities for the curvature of the linear connection on $\pi_1^{0*}(\tau_E)$, as proved in [4].

The computation of the alternant coming from the second-order conditions and the D^V equation (12) proceeds in much the same way: here, the essentially new property which has to be invoked is the full symmetry of $\mathrm{D}^V\theta$, following from

$$\mathrm{D}^V\theta(X,Y,Z)U = \mathrm{D}^V\theta(Y,X,Z)U, \tag{46}$$

which is another second Bianchi identity for the linear connection.

That the alternant with the D^H equation (16) does not give rise to new conditions either is due to the property

$$\mathrm{D}^V\mathrm{D}^V R(U,X,Y,Z) = \mathrm{D}^H\theta(Z,X,Y)U - \mathrm{D}^H\theta(Y,X,Z)U. \tag{47}$$

Again, this is a second Bianchi identity of the linear connection, although it has been shown in [4] that it is merely a consequence of the other two.

Finally, one must consider whether there could be a condition coming from the internal compatibility of the second-order conditions. If one starts with second prolongations of the conditions (39) and proceeds in much the same way as was explained in some detail for the D^H prolongations of (16), it is fairly easy to see that nothing else but direct prolongations can be produced.

Since the second-order conditions are not true integrability conditions (they would merely be part of the prolonged system in other approaches to formally integrable partial differential equations), there is no need to verify compatibility between them and the double hierarchy of algebraic restrictions: for example, it would not be possible to find suitable combinations of second-order prolongations of the algebraic conditions which would eliminate the derivative terms. So at last the process seems to have come to an end. This is only partially true, however, because the computations of Appendix B rely on the fact that the conditions (39) effectively express a dependency between second-order parametric derivatives. If, for example, the diagonal elements of g are given priority in the selection of principal derivatives, then the leading terms of the second-order passivity requirements (39) will involve only non-diagonal elements of g. But if, for example as a result of the algebraic requirements, some of these non-diagonal elements are zero, then the corresponding passivity conditions will "degenerate" into new first-order (or even algebraic) integrability conditions. One can see this happening in a large number of cases discussed by Douglas for $n = 2$. In such situations, the further computation of alternants will not follow the course described in Appendix B, and it seems unlikely that anything could be said about such cases at the same level of generality.

Let us summarize the whole scheme, in the form of a theorem.

Theorem 1 *The complete set of integrability and/or passivity conditions associated with the Helmholtz equations (11), (12) and (13)*

$$\begin{aligned} \nabla g &= 0 \\ \mathrm{D}^{V} g\,(Z, X, Y) &= \mathrm{D}^{V} g\,(Y, X, Z) \\ g\,(\Phi(X), Y) &= g\,(\Phi(Y), X), \end{aligned}$$

are equation (16)

$$\mathrm{D}^{H} g\,(Z, X, Y) = \mathrm{D}^{H} g\,(Y, X, Z),$$

the two sets of algebraic requirements (17)

$$g\,(\nabla^{k}\Phi(X), Y) = g\,(\nabla^{k}\Phi(Y), X) \qquad k = 1, 2, 3, \ldots$$

and (21)

$$\sum g\,(\nabla^{k} R(X, Y), Z) = 0 \qquad k = 0, 1, 2, \ldots,$$

and the second-order conditions (39)

$$\mathrm{D}^{H}\mathrm{D}^{V} g\,(U, Z, Y, X) - \mathrm{D}^{H}\mathrm{D}^{V} g\,(Z, U, Y, X) + g\,(\theta(Y, X)Z, U) - g\,(\theta(Y, X)U, Z) = 0.$$

The completeness of the scheme only applies when there is no degeneracy in the second-order passivity conditions.

4 The generalization of Douglas's Case I to arbitrary dimension

Recall that the classification of Douglas for $n = 2$, as explained in detail in [5], starts from the degree of linear independence of the matrices $I, \Phi, \nabla\Phi, \ldots$ in the hierarchy (17) of algebraic conditions. Observe, by the way, that Douglas did not have to worry about the other hierarchy (21), because these conditions are void in dimension 2. The case which is most easy to identify, which we will continue to call Case I, is the case where Φ is a multiple of the identity tensor:

$$\Phi = \mu I, \qquad \mu \in C^\infty(J^1\pi). \tag{48}$$

It is called (for autonomous equations) the (flat) isotropic case in [7].

Under these circumstances, it is obvious that all conditions (17) are satisfied. Also, we have from (19) that $3\,R(X,Y) = X^V(\mu)Y - Y^V(\mu)X$, so that for any symmetric g the cyclic sum (20) will be identically zero (and likewise for the other conditions in the hierarchy (21)). Thus the algebraic integrability tests place no restrictions on g.

In more explicit form, the resulting system of equations we have to deal with is obtained from equations (11), (12), (16) and (39) in the following way. We take a local basis $\{X_i\}_{1\le i\le n}$ of $\overline{\mathcal{X}}(\pi_1^0)$, so that $\{\Gamma, {X_i}^H, {X_i}^V\}$ will be a local frame of vector fields on $J^1\pi$; the system of equations referred to this basis, is

$$\begin{aligned}
\nabla g(X_i, X_j) &= 0 \\
\mathrm{D}^V g\,(X_k, X_i, X_j) &= \mathrm{D}^V g\,(X_j, X_i, X_k) \\
\mathrm{D}^H g\,(X_k, X_i, X_j) &= \mathrm{D}^H g\,(X_j, X_i, X_k) \\
(\mathcal{A}(X_k, X_l)g)\,(X_i, X_j) &= 0.
\end{aligned}$$

As i, j, k and l range over $1, 2, \ldots, n$ we obtain a system of partial differential equations for the components of g with respect to the X_i. We shall denote the components of g by g_{ij} (that is, $g_{ij} = g(X_i, X_j)$).

Since the terms appearing in the equations are tensorial, a different choice of local basis $\{X_i\}$ will produce an equivalent system of equations. We shall therefore take for $\{X_i\}$ a coordinate basis: $X_i = \partial/\partial x^i$. Notice that this does not imply that the corresponding local basis for $J^1\pi$ is a coordinate basis.

In fact it is not the full set of equations listed above that we shall be concerned with. In the first place, some of the equations are trivial for certain values of the suffices: for example, we do not have to consider the cases $j = k$ in equations (12) or (16). More significantly, only a certain subset of equations (39), which we will specify later, will be required.

The general theory developed in the previous section deals with the question of passivity almost in its entirety. Passivity conditions arise from the computation of alternants between different equations of the system. We know that equations (16) arise as passivity conditions for equations (11) and (12), and equations (39) contain all passivity conditions that might arise between the first three. Moreover, all conditions arising as alternants between the full set are satisfied automatically by virtue of the original equations, their prolongations, and the Bianchi identities for the curvature associated with the covariant derivative operators ∇, D^V and D^H. Provided that we ensure that the subset of equations (39) which we include in our system consists precisely of the passivity conditions which can arise from consideration of equations (11), (12) and (16), our system as a whole is bound to be passive. All that remains is to deal with orthonomicity.

The system consisting of equations (11), (12) and (16) and the passivity conditions to be found amongst equations (39) is clearly not orthonomic as it stands. What we shall do is show that it can be put in orthonomic form. For this purpose we must specify an ordering of the derivatives which appear in the equations, including the undifferentiated terms g_{ij}, which are treated as derivatives of degree zero. We must now spell out in more detail how this process can be carried out in full agreement with the general rules for orthonomicity mentioned in the Introduction.

As a preliminary step we specify an ordering of the components g_{ij} of g, the dependent variables in our system of differential equations. We may think of them as the elements of a symmetric matrix – and we need consider only those that come on and above the main diagonal. The diagonal elements come first, in their natural order; that is to say, g_{11} is the dependent variable of highest order, followed by g_{22}, and so on down to g_{nn}. The elements of the sub-diagonal immediately above the main diagonal come next, again in their natural order. We continue to work up through the sub-diagonals in the same way, the final element – the one of lowest order – being the one in the top right-hand corner of the matrix. To be specific: for two components of g, say g_{ij} and $g_{i'j'}$ (where we may assume without loss of generality that $i \leq j$ and $i' \leq j'$), we have $g_{ij} > g_{i'j'}$ if either $j - i < j' - i'$, or if $j - i = j' - i'$ and $i < i'$. This ordering is a natural extension of the one given by Douglas for $n = 2$.

In order to describe the ordering of the derivatives of the g_{ij} it is convenient to rename the independent variables temporarily, as follows: $t \mapsto x^0$; $v^i \mapsto x^{n+i}$. We can therefore deal with all the independent variables as x^a, say, with $a = 0, 1, 2, \ldots, n, n+1, \ldots, 2n$. The advantage of this manoeuvre is that it allows us to use the multi-index notation for derivatives: we write $\partial^{|A|} g_{ij}/\partial x^A$ for the partial derivative determined by the multi-index A, where $|A|$ is the sum of the entries of A. We remind the reader that there is a natural ordering of multi-indices, such that $A > A'$ just in case the first non-zero entry in $A - A'$ is positive.

We can now specify the ordering of the partial derivatives $\partial^{|A|} g_{ij}/\partial x^A$. We say that $\partial^{|A|} g_{ij}/\partial x^A > \partial^{|A'|} g_{i'j'}/\partial x^{A'}$ if

1. $|A| > |A'|$; or

2. $|A| = |A'|$ and $g_{ij} > g_{i'j'}$; or

3. $|A| = |A'|$, $i = i'$, $j = j'$ and $A > A'$.

That is to say, we order the partial derivatives first by degree of differentiation, then (for derivatives of the same degree) by the order of the dependent variable according to the ordering defined earlier, and finally (for derivatives of the same degree of the same dependent variable) according to the independent variables as determined by their multi-indices. In this last situation, t derivatives have higher order than x derivatives, which in turn have higher order than v derivatives. This system is again a direct generalization of the one used by Douglas in his analysis of Case I for $n = 2$. Douglas in fact uses Riquier's method of cotes to define his ordering; it is not necessary to introduce cotes into the discussion here, but it is easy to see how cotes could be defined to produce the ordering specified above for the general case.

We can now explain how to write the equations in orthonomic form. As a first step we shall say, in each case, which derivative is to be singled out to appear on the left-hand side of the rearranged equation, that is, to be the principal derivative. This must be one of the derivatives of highest degree which occur in the equation; in the case of equations (11), (12) and (16) it will be a derivative of degree 1, for equations (39) it will be a derivative of degree 2. All terms of lower degree may therefore be transferred immediately to the right-hand sides of the equations.

We shall always suppose that the suffices i, j on any term $g_{ij} = g_{ji}$ have been written in non-decreasing order.

The first derivative terms in equation (11) are

$$\frac{\partial g_{ij}}{\partial t} + v^k \frac{\partial g_{ij}}{\partial x^k} + f^k \frac{\partial g_{ij}}{\partial v^k}.$$

We must choose $\partial g_{ij}/\partial t$ as principal derivative, and transfer the remaining terms to the right-hand side of the equation. This clearly isolates the derivative of highest order on the left-hand side, as required.

The equations (12) take the form

$$\frac{\partial g_{ij}}{\partial v^k} = \frac{\partial g_{ik}}{\partial v^j} \text{ or } \frac{\partial g_{ki}}{\partial v^j}.$$

We consider only the non-trivial equations, for which i, j and k are not all the same. The dependent variables that appear are therefore different, and we must choose for the principal derivative in each equation that for which the dependent variable has the higher order according to our ordering scheme. In the case in which two of the indices are the same and the third is different, this simply requires that the derivative of the diagonal term be taken as principal. When all three indices are distinct, however, three equations can be written down, one of which is superfluous: for example, when i, j and k are 1, 2 and 3 in some order, the equations are

$$\frac{\partial g_{12}}{\partial v^3} = \frac{\partial g_{13}}{\partial v^2} = \frac{\partial g_{23}}{\partial v^1}.$$

Of the three distinct dependent variables which occur, one is of least order according to our ordering scheme, and we choose the two equations which involve this one, placing it on the right-hand side. Thus in the specific case just considered the chosen equations are

$$\frac{\partial g_{12}}{\partial v^3} = \frac{\partial g_{13}}{\partial v^2} \quad \text{and} \quad \frac{\partial g_{23}}{\partial v^1} = \frac{\partial g_{13}}{\partial v^2}.$$

It is immediate that the terms on the left-hand sides of our equations are all distinct, and that in each case the term on the right-hand side has lower order than the term on the left.

The equations (16) are (modulo the order of the suffices on g)

$$\frac{\partial g_{ij}}{\partial x^k} - \Gamma^l_k \frac{\partial g_{ij}}{\partial v^l} + \ldots = \frac{\partial g_{ik}}{\partial x^j} - \Gamma^l_j \frac{\partial g_{ik}}{\partial v^l} + \ldots$$

where the omitted terms do not involve derivatives of the dependent variables. We choose the principal derivative from between the two x derivatives using exactly the same procedure as we did for equations (12). On rearrangement the equation becomes (say)

$$\frac{\partial g_{ij}}{\partial x^k} = \Gamma^l_k \frac{\partial g_{ij}}{\partial v^l} + \frac{\partial g_{ik}}{\partial x^j} - \Gamma^l_j \frac{\partial g_{ik}}{\partial v^l} + \ldots.$$

The terms involving $\partial g_{ij}/\partial v^k$ are of lower order than the principal derivative because they are derivatives of the same dependent variable but with respect to v variables rather than x^k; the other derivatives on the right-hand side are derivatives of a dependent variable of lower order than g_{ij}.

With these provisions the first-order equations as rearranged are seen to satisfy the first two conditions for orthonomicity as described in the Introduction; but they are not yet in orthonomic form because some of the derivatives appearing on the right-hand sides are principal. However, these may clearly be eliminated by substitution in terms of parametric derivatives; each such substitution involves the replacement of a term by terms of lower order, so the second of the conditions for orthonomicity still obtains. So the first-order equations may be written in orthonomic form.

Notice that for each choice of values for the indices i, j and k, the derivatives $\partial g_{ij}/\partial x^k$ and $\partial g_{ij}/\partial v^k$ are either both principal or both parametric. The conditions on i, j and k that these derivatives are principal are that either $k < i$ or $j < k$. For these derivatives to be parametric, conversely, $i \leq k \leq j$. It is perhaps worth mentioning, what might easily be overlooked, that $\partial g_{ii}/\partial x^i$ and $\partial g_{ii}/\partial v^i$ are parametric, as indeed are all the undifferentiated dependent variables.

Notice also that the equations (12) by themselves are orthonomic. It is clear from our discussions of the passivity conditions in the previous section that these equations by themselves also comprise a passive system.

We turn now to consideration of the second-order equations (39). From their full expression, which is

$$\begin{aligned} &\mathrm{D}^H\mathrm{D}^V g\,(X_k, X_l, X_i, X_j) - g\,(\theta(X_i, X_j)X_k, X_l) \\ &\quad = \mathrm{D}^H\mathrm{D}^V g\,(X_l, X_k, X_i, X_j) - g\,(\theta(X_i, X_j)X_l, X_k), \end{aligned}$$

it can be seen that when $X_i = \partial/\partial x^i$ these equations take the form

$$\frac{\partial^2 g_{ij}}{\partial x^k \partial v^l} - \Gamma^m_k \frac{\partial^2 g_{ij}}{\partial v^m \partial v^l} + \ldots = \frac{\partial^2 g_{ij}}{\partial x^l \partial v^k} - \Gamma^m_l \frac{\partial^2 g_{ij}}{\partial v^m \partial v^k} + \ldots$$

where the omitted terms are, again, of lower order. We must first state which of these equations are to be included in our system. We wish to include only those which express relations between parametric derivatives for the first-order equations. A necessary condition for this is that the values of i, j, k and l must be restricted to those for which both $\partial g_{ij}/\partial x^k$ and $\partial g_{ij}/\partial v^l$ are parametric, and equally for which both $\partial g_{ij}/\partial x^l$ and $\partial g_{ij}/\partial v^k$ are parametric. There may be principal derivatives among terms like $\partial^2 g_{ij}/\partial v^m \partial v^l$, since all the derivatives of principal derivatives are themselves principal, by definition. However, it follows from the fact that equations (12) by themselves form a passive orthonomic system, and from the property of such systems noted earlier, that all such principal derivatives may be replaced by parametric derivatives by making appropriate substitutions from equations (12), and that the expression so obtained is uniquely determined. So this choice of values of i, j, k and l will give precisely the equations which arise as passivity conditions for the first-order equations. With these equations now added to the system we have to confirm that they can also be written in orthonomic form (which involves selecting some of what were parametric derivatives for the first-order system to be principal for the extended system). Clearly we must have $k \neq l$, and in accordance with our ordering principles we have to choose $\partial^2 g_{ij}/\partial x^k \partial v^l$ to be principal for $k < l$. When each equation containing such a term is written with it on the left-hand side, and all the principal derivatives on the right-hand side are eliminated by substitution in terms of parametric derivatives, the system becomes orthonomic. The derivative $\partial^2 g_{ij}/\partial x^k \partial v^l$ will be principal when $i \leq k < l \leq j$.

Thus the system we have defined is orthonomic and passive, and it follows that Case I is variational.

For a strict application of Riquier theory it was necessary to deal with the partial derivatives of the dependent variables in the way described above. But this approach appears very pedestrian, and foreign to the spirit of previous sections in which we made full use of the tensorial properties of our equations. A way of improving the situation is obtained by extending a procedure already used by Douglas: he elects to use (in our notation) Γ in place of $\partial/\partial t$ not just in calculating passivity conditions but also in testing for orthonomicity. As he puts it: 'The cote of x [our t] is thought of as associated with the operator d/dx [our Γ] rather than $\partial/\partial x$ [our $\partial/\partial t$].' The idea here is that for any ordering of the independent variables in which t has highest order, it makes no essential difference when considering orthonomicity whether we write the equation $\nabla g(X_i, X_j) = 0$ in the form $\partial g_{ij}/\partial t = \ldots$ or $\Gamma(g_{ij}) = \ldots$, since the difference between the left-hand sides of the two versions of the equation consists of terms which are necessarily of lower order than $\partial g_{ij}/\partial t$. Now the vector fields ${X_i}^H = \partial/\partial x^i - \Gamma_i^j \partial/\partial v^j$ share with Γ the property of containing a coordinate vector field with coefficient 1; it is clear that Douglas's idea may be extended to the ${X_i}^H$, provided that the independent variables v^i have lower order than the x^i. There is no need to extend the idea further since the ${X_i}^V$ are coordinate fields anyway.

The detailed arguments given above, in particular the reasoning which shows that in expressions like ${X_l}^V {X_k}^H(g_{ij})$ terms like $\partial^2 g_{ij}/\partial v^l \partial v^m$ will always become parametric by the end of the process, show that our system is passive and orthonomic in this generalized sense in which the operators $\{\partial/\partial t, \partial/\partial x^i, \partial/\partial v^i\}$ are replaced by $\{\Gamma, {X_i}^H, {X_i}^V\}$ (where $X_i = \partial/\partial x^i$), and is formally integrable. Since, however, there is no generalized version of Riquier theory available from which we could have started to argue immediately in these terms, we have chosen the safer way of showing the validity of such an approach directly, by relating it to the standard theory with coordinate derivatives.

5 Counting the degree of arbitrariness

As mentioned at the end of Section 1, the formula in [1] for the degree of arbitrariness of the solution is not in agreement with Douglas's result for $n = 2$. We wish to sketch here how Douglas's way of counting proceeds for general n. In addition, we wish to explain where the difference comes from.

Douglas's calculation of the number of arbitrary functions in the solution, and the number of variables occurring in each of these functions, uses the notion of a complete set of parametric derivatives. A finite set of parametric derivatives is termed complete if all parametric derivatives (of all degrees) can be derived from the set, in the following manner. First, with each parametric derivative from the complete set a certain collection of independent variables must be associated; these are called the multipliers for that derivative. For clarity we emphasise that different parametric derivatives may have different multipliers. Then for the set (with the associated assignment of multipliers to each element) to be complete, every parametric derivative must be expressible in one and only one way as a (multiple) derivative of one of the members of the set with respect to its multipliers. It is permissible (indeed in our case necessary) to include undifferentiated dependent variables among the elements of a complete set, regarding them as derivatives of degree zero in the usual way. It is a result of the Riquier theory that when a system of equations is orthonomic and passive, a complete set of parametric derivatives may be found for it; there is no implication that a complete set is uniquely determined, however. There is then an arbitrary function in the solution for each of the members of the complete set, and this is a function of the associated multipliers. If a complete set of parametric derivatives can be exhibited then the freedom or arbitrariness in the solution can be read off from it. Douglas shows how this can be done for Case I with $n = 2$. We shall explain how to generalize his result to arbitrary n.

Note first of all that since $\partial g_{ij}/\partial t$ is principal for all i and j, no derivative with respect to t can occur among the parametric derivatives. We may therefore simply ignore the variable t in the process of counting the freedom in the general solution.

It is instructive first to consider what happens for small values of n, beginning with $n = 2$, in order to recall Douglas's method.

Since parametric derivatives are defined in terms of principal derivatives, we shall need to consider all the principal derivatives (other than those involving $\partial g_{ij}/\partial t$) for small values of n. To do so it is enough to list the values of the indices i, j, k and (for the second-order derivatives) l for which the corresponding derivatives $\partial g_{ij}/\partial v^k$, $\partial g_{ij}/\partial x^k$ and $\partial^2 g_{ij}/\partial x^k \partial v^l$ are principal. Note that for the first-order derivatives each appropriate choice of values of the indices i, j and k identifies *two* terms, one a v derivative and the other an x derivative. We shall write the values in the form ij, k or ij, kl, with the comma to separate the suffices on the dependent variable (the first two) from the index (indices) on the independent variable(s) of differentiation. In order to be systematic we shall list these principal derivatives, in tabular form, in decreasing order reading down the columns, except that we shall give the first-order derivatives before the second-order ones.

Incidentally, the values of the indices we shall list will serve at the same time to identify the equations in the system, if the given values are substituted for i, j, k and l in

$$\begin{aligned} \mathrm{D}^V g\,(X_k, X_i, X_j) &= \mathrm{D}^V g\,(X_j, X_i, X_k) \\ \mathrm{D}^H g\,(X_k, X_i, X_j) &= \mathrm{D}^H g\,(X_j, X_i, X_k) \\ (\mathcal{A}(X_k, X_l) g)\,(X_i, X_j) &= 0 \end{aligned}$$

(where $X_i = \partial/\partial x^i$). Note again that for the first-order equations each entry in the list corresponds to two equations, one involving vertical derivatives only and the other involving horizontal derivatives.

For $n = 2$, with the notational conventions introduced above, the principal derivatives that occur in the equations of the system are represented by

11,2	12,12
22,1	

Of course, in addition to these all their derivatives are principal. The parametric derivatives can be described as follows: the undifferentiated dependent variables; any derivative of g_{11} with respect to the variables x^1 and/or v^1; any derivative of g_{22} with respect to the variables x^2 and/or v^2; any derivative of g_{12} with respect to any of x^1, x^2 and v^1; any derivative of $\partial g_{12}/\partial v^2$ with respect to any variable other than x^1. This description reveals a complete set of parametric derivatives: they are

Derivative	Multipliers
g_{11}	x^1, v^1
g_{22}	x^2, v^2
g_{12}	x^1, x^2, v^1
$\partial g_{12}/\partial v^2$	x^2, v^1, v^2

These parametric derivatives and multipliers are the ones given by Douglas, translated into our notation. It is easy to check that they do satisfy the condition for being a complete set. Consider, for example, a parametric derivative of g_{12}. If it includes any differentiations with respect to v^2 then it cannot also include any differentiations with respect to x^1; it can therefore be obtained by differentiating $\partial g_{12}/\partial v^2$ with respect to some or all of the variables x^2, v^1 and v^2. If it does not include any differentiations with respect to v^2 then it can be obtained by differentiating g_{12} with respect to any variable other than v^2.

The general solution in this case therefore contains two arbitrary functions of two variables and two arbitrary functions of three variables, which we shall indicate by saying that the freedom is

$$2\,f(2) + 2\,f(3).$$

Consider next the case $n = 3$. The principal derivatives are

11,2	12,3	12,12
11,3	23,1	23,23
22,1		
22,3		13,12
33,1		13,13
33,2		13,23

together with all their derivatives. The way in which the complete set of parametric derivatives is drawn up for $n = 2$ gives us plenty of guidance for the case $n = 3$. Clearly we must include in the complete set all the diagonal elements g_{ii}, with multipliers x^i and

v^i. Consider now the parametric derivatives of g_{12}. Since 12,3 is principal, neither x^3 nor v^3 can appear in a parametric derivative of g_{12}. Furthermore, 12,12 is principal, and is the only second-order principal derivative of g_{12}. Thus g_{12} for $n = 3$ behaves exactly like g_{12} for $n = 2$. The same is true for g_{23}, under the substitutions $2 \mapsto 1 \mapsto 3 \mapsto 2$. This shows what the elements of the complete set for these two dependent variables must be. Finally we must consider g_{13}. There are no first-order principal derivatives of this variable (as is the case for g_{12} when $n = 2$), but three second-order ones. A little thought shows how to extend the method for g_{12} from $n = 2$ to cover g_{13} for $n = 3$. We obtain the following complete set:

Derivative	Multipliers
g_{11}	x^1, v^1
g_{22}	x^2, v^2
g_{33}	x^3, v^3
g_{12}	x^1, x^2, v^1
$\partial g_{12}/\partial v^2$	x^2, v^1, v^2
g_{23}	x^2, x^3, v^2
$\partial g_{23}/\partial v^3$	x^3, v^2, v^3
g_{13}	x^1, x^2, x^3, v^1
$\partial g_{13}/\partial v^2$	x^2, x^3, v^1, v^2
$\partial g_{13}/\partial v^3$	x^3, v^1, v^2, v^3

The freedom is

$$3\, f(2) + 2 \times 2\, f(3) + 3\, f(4).$$

The induction process will become more apparent when we go one step further and look at the case $n = 4$. The principal derivatives are

11,2	12,3	12,12	14,12
11,3	12,4	23,23	14,13
11,4	23,1	34,34	14,14
22,1	23,4		14,23
22,3	34,1	13,12	14,24
22,4	34,2	13,13	14,34
33,1		13,23	
33,2	13,4		
33,4	24,1	24,23	
44,1		24,24	
44,2		24,34	
44,3			

The diagonal elements behave just as the diagonal elements did previously.

The elements in the first sub-diagonal (the one immediately above the main diagonal) behave just like g_{12} for $n = 2$. Indeed, the fact that 12,3 and 12,4 are principal means that only x^1, v^1, x^2 and v^2 can appear in parametric derivatives of g_{12}; then the fact that 12,12 is principal reduces the case to that of g_{12} for $n = 2$. Similar considerations apply to the other entries in this sub-diagonal, with appropriate substitutions of indices. There are three such terms in all.

The two elements of the next sub-diagonal behave just like g_{13} for $n = 3$; for g_{13} itself, the translation is direct, while for g_{24} one has to make the substitutions $2 \mapsto 1 \mapsto 4 \mapsto 3 \mapsto 2$.

Note that there are again no first-order principal derivatives of the top right-hand corner element g_{14}. Since the other terms have already been dealt with, it is enough to list the members of the complete set involving g_{14}. They are

Derivative	Multipliers
g_{14}	x^1, x^2, x^3, x^4, v^1
$\partial g_{14}/\partial v^2$	x^2, x^3, x^4, v^1, v^2
$\partial g_{14}/\partial v^3$	x^3, x^4, v^1, v^2, v^3
$\partial g_{14}/\partial v^4$	x^4, v^1, v^2, v^3, v^4

The freedom for $n = 4$ is

$$4\,f(2) + 3 \times 2\,f(3) + 2 \times 3\,f(4) + 4\,f(5).$$

The $f(2)$ terms come from the 4 diagonal elements, the $f(3)$ terms from the 3 elements in the first sub-diagonal, the $f(4)$ terms from the 2 elements in the next sub-diagonal, and the $f(5)$ terms from the top right-hand corner element.

It is now easy to see what the general formula for the freedom must be. If we index the sub-diagonals with m, so that the element g_{ij} lies in the mth sub-diagonal where $m = j - i$ (including the cases $m = 0$ – the main diagonal – and $m = n - 1$ – the top right-hand corner element), then there are $n - m$ elements in the mth sub-diagonal and each of them contributes $m + 1$ functions of $m + 2$ variables. Therefore, the formula

$$\sum_{m=0}^{n-1} (n - m)(m + 1) f(m + 2)$$

gives the total freedom in the solution for Case I in n dimensions.

A crucial part is played in this calculation by the top right-hand corner element g_{1n}. Let us denote it for convenience by ρ. There are no first-order principal derivatives of ρ, while all second-order derivatives of the form $\partial^2\rho/\partial x^k \partial v^l$ with $k < l$ are principal. The part of the table of a complete set of parametric derivatives for dimension n involving ρ, with the corresponding multipliers, will be

Derivative	Multipliers
ρ	$x^1,\ x^2,\ x^3,\ \ldots,\ x^{n-1},\ x^n,\ v^1$
$\partial\rho/\partial v^2$	$x^2,\ x^3,\ \ldots,\ x^{n-1},\ x^n,\ v^1,\ v^2$
$\partial\rho/\partial v^3$	$x^3,\ \ldots,\ x^{n-1},\ x^n,\ v^1,\ v^2,\ v^3$
$\vdots$	$\vdots$
$\partial\rho/\partial v^{n-1}$	$x^{n-1},\ x^n,\ v^1,\ v^2,\ \ldots,\ v^{n-1}$
$\partial\rho/\partial v^n$	$x^n,\ v^1,\ v^2,\ v^3,\ \ldots,\ v^{n-1},\ v^n$

A parametric derivative of ρ which contains a differentiation with respect to v^n cannot also contain a differentiation with respect to any x^k for $k < n$, because $\partial^2\rho/\partial x^k\partial v^n$ is principal when $k < n$. Thus any such derivative can be obtained by differentiating $\partial\rho/\partial v^n$ with respect to some of the variables x^n, v^1, v^2, v^3, $\ldots$, v^{n-1}, v^n. This accounts for the final row in the table. Once derivatives of ρ which contain a differentiation with respect to v^n have been dealt with, the situation reduces to that for the top right-hand corner element for dimension $n-1$, except that the list of multipliers for each parametric derivative must include an additional variable, namely x^n. The table has been built up inductively on this pattern from that for g_{12} when $n = 2$.

There is an heuristic method for finding the freedom in the solution which can be used to confirm the results given above. For any dependent variable, take all of its principal derivatives in the list for the dimension under consideration, and set them equal to zero. The resulting system of partial differential equations can be readily solved; the solution contains the same number of arbitrary functions of the same variables as the theory predicts will be associated with that dependent variable according to the method of complete sets of parametric derivatives. This approach is evidently consistent with (and indeed suggested by) the remarks above concerning g_{12} for $n = 3$, where it was pointed out that the fact that 12,3 is principal means in effect that no parametric derivative of g_{12} occurring in a complete set – including g_{12} itself – can have x^3 or v^3 among its multipliers. If the reader cares to check back (s)he will see that similar considerations apply to all the arguments given explicitly in the determination of the freedom for $n = 2$, 3 and 4.

Let us consider the case of the corner element $\rho = g_{1n}$ for arbitrary n from this point of view. There will be no first-order principal derivatives of ρ, and the second-order ones will be those for dimension $n - 1$ with in addition $\partial^2\rho/\partial x^k\partial v^n$ for $k = 1, 2, \ldots, n-1$. Consider the partial differential equations obtained by setting all these derivatives to zero; and first consider the set just explicitly identified:

$$\frac{\partial^2\rho}{\partial x^k \partial v^n} = 0, \qquad k = 1, 2, \ldots, n-1.$$

On solving these equations we obtain

$$\frac{\partial\rho}{\partial v^n} = F(x^n, v^1, v^2, \ldots, v^n),$$

and therefore

$$\rho = \hat{F}(x^n, v^1, v^2, \ldots, v^n) + G(x^1, x^2, \ldots, x^n, v^1, v^2, \ldots, v^{n-1}),$$

where F, $\hat{F}$ and G are functions of the indicated variables (and $F = \partial \hat{F}/\partial v^n$). This does not complete the determination of ρ of course: the remaining equations appropriate to dimension $n-1$ still have to be satisfied. Since each of these involves a derivative with respect to some x^k with $k < n$, the function $\hat{F}$ is not affected by these conditions; but G has in effect to satisfy the corresponding equations for the corner element in dimension $n-1$ (though of course it depends on one additional variable, namely x^n). Thus according to this reasoning, in dimension n there will be one more arbitrary function associated with ρ than there is for the corresponding element in dimension $n-1$, and each of the arbitrary functions will be a function of one more variable. This confirms inductively the freedom $n\,f(n+1)$ associated with ρ in dimension n. All the other contributions to the calculation of the overall freedom can be dealt with in similar ways.

We summarize the results of this and the previous section in the following statement.

Theorem 2 *For a system of second-order equations $\ddot{x}^j = f^j(t, x, \dot{x})$, satisfying the property $\Phi = \mu I$, it is always possible to construct a symmetric multiplier matrix g_{ij}, which solves the conditions (9) (or (11–13)). The freedom in choosing arbitrary functions in the general solution will be given by*

$$\sum_{m=0}^{n-1} (n-m)(m+1) f(m+2). \tag{49}$$

However, one should be a little sceptical about the formula for the freedom in the solution. It predicts, for example, far more arbitrary functions than the result given by Anderson and Thompson in [1]; their result is just the $m = n-1$ term in (49), namely $n\,f(n+1)$ (except that they made a slip of the tongue and said $(n+1)\,f(n)$ in their conclusion). Seiler [20] has discussed in detail the possible ambiguity in statements about arbitrariness in the solution of involutive systems; his concluding remarks suggest that one should not take any results in that area too seriously. The problem is that there are many ways of representing a solution and new arbitrary functions which arise in a step by step integration process can often be absorbed into previously obtained functions of more arguments. Seiler says that the best measure is probably provided by the coefficients of the Hilbert polynomial, but it would take us too far to enter into such a discussion here. What is important in Seiler's work is a theorem which proves that the only true invariant is the number of free functions of the maximal number of arguments. Other authors have made this observation as well. In the Spencer-Goldschmidt approach to formally integrable systems, the arbitrariness in the solution is measured in one way or another from the computation of the Cartan characters and it is known that only the last non-zero Cartan character (when the system is represented in 'δ-regular' coordinates) has an intrinsic meaning. From this point of view, therefore, the discrepancy between the formula for the freedom in the solution in [1] and the one obtained above (and previously by Douglas for $n = 2$) need not be a cause for concern.

Rather than going into more theoretical considerations, we will provide now a simplified model which exhibits a lot of the features of the complex set of passivity conditions of Case I (at least for $n = 2$), but is simple enough to let us understand exactly what is happening.

Consider the following problem with two unknowns ρ_i and four independent variables:

$$\begin{aligned} \frac{\partial \rho_1}{\partial x^2} &= \frac{\partial \rho_2}{\partial x^1} \\ \frac{\partial \rho_1}{\partial x^4} &= \frac{\partial \rho_2}{\partial x^3}. \end{aligned}$$

Equating $\partial^2 \rho_1 / \partial x^2 \partial x^4$ and $\partial^2 \rho_1 / \partial x^4 \partial x^2$, as computed from the above equations, one obtains the following relation between parametric derivatives:

$$\frac{\partial^2 \rho_2}{\partial x^1 \partial x^4} = \frac{\partial^2 \rho_2}{\partial x^2 \partial x^3}.$$

So this is a second-order passivity condition in the Riquier approach. Promoting the left-hand side to principal derivative, no further conditions can be obtained and we reach the conclusion that the system is formally integrable. A complete set of parametric derivatives and multipliers is given in the following table:

Derivative	Multipliers
ρ_1	x^1, x^3
ρ_2	x^1, x^2, x^3
$\partial \rho_2 / \partial x^4$	x^2, x^3, x^4

According to the Riquier method, therefore, the freedom is $f(2) + 2\,f(3)$. The first point we can illustrate explicitly here is that $2\,f(3)$ is the unambiguous part of this measurement of the freedom and that the $f(2)$, although we can see where it comes from, is somehow redundant. If we solve first the second-order equation involving ρ_2 only, the freedom in its solution clearly is $2\,f(3)$. Substituting this solution into the original equations and integrating for ρ_1 creates the additional freedom $f(2)$. But it is not hard to verify explicitly here that this arbitrary function of two variables can be absorbed into the two functions of three variables which are left free in ρ_2. Alternatively, one could proceed to solve the problem as follows. The first equation shows that $\rho_i = \partial \phi / \partial x^i$, $i = 1, 2$, for some function ϕ of all four independent variables; the necessary and sufficient condition for the second equation to be satisfied also is that

$$\frac{\partial^2 \phi}{\partial x^1 \partial x^4} = \frac{\partial^2 \phi}{\partial x^2 \partial x^3}.$$

By the same reasoning the freedom in ϕ is just $2\,f(3)$ and is now clearly the only freedom in the system.

The second point of interest about this example is that we can illustrate that the second-order passivity condition for ρ_2, which plays an essential role in the Riquier approach and manifestly provides information which is not explicitly contained in the given first-order equations, from yet another point of view is itself redundant. If we apply the coordinate transformation $y^1 = x^1 + x^4$, $y^2 = x^2$, $y^3 = x^3$, $y^4 = x^1 - x^4$ to the original equations, they become

$$\begin{aligned}\frac{\partial \rho_1}{\partial y^1} &= \frac{\partial \rho_1}{\partial y^4} + \frac{\partial \rho_2}{\partial y^3} \\ \frac{\partial \rho_2}{\partial y^1} &= \frac{\partial \rho_1}{\partial y^2} - \frac{\partial \rho_2}{\partial y^4}.\end{aligned}$$

The transformed system is of Cauchy-Kowalevski type, so there is no need to search for integrability conditions (and it is clear again that the solution will depend on $2\,f(3)$).

It is perhaps not entirely convincing to extrapolate the conclusions of this simple model problem to the complicated reality of the inverse problem in Case I. If the extrapolation were allowable we would conclude that only the functions depending on the largest number of variables, that is, the term corresponding to $m = n - 1$ in (49), which is the contribution of the top right-hand corner element, would represent the true freedom in the solution. The other functions of fewer variables might indeed turn up if one were able to solve the equations explicitly by the methods of Riquier theory; but even so it is quite plausible that they would not be functionally independent of the functions with the maximal number of independent variables.

It is worth saying, finally, that the overcounting which we have observed here, is not just a feature of Riquier-Janet theory versus more modern approaches. In Appendix C, we briefly discuss the matter of computing Cartan characters and provide another simple model example where different forms of measuring the freedom lead to different answers and where this time the Riquier method provides the more economical result.

6 Concluding remarks

Using the geometrical version (11–13) of the Helmholtz conditions and the calculus of covariant derivations which can be associated in a natural way to a given system of second-order differential equations, we have obtained all general passivity and integrability conditions which the Riquier theory can produce in the inverse problem of the calculus of variations. This is the maximal result one may hope to attain at this level of generality, i.e. without entering a case by case study based for example on the Jordan normal form of the matrix Φ (cf. [6, 5]). Further analysis is required when the second-order passivity conditions (39) degenerate into first- or zeroth-order conditions and thus become true integrability conditions. Our results remain useful for such cases as the conditions (39) are in any case the source of potentially new integrability requirements and a further investigation can then be taken up from this point.

We have applied our general theory to the solution of the inverse problem in Case I of Douglas for arbitrary dimension; this is the case where the nature (48) of Φ is such that the algebraic conditions have no effect on the generalized metric g we are looking for, so that no degeneracy can occur. We have shown that Case I is variational and have estimated the degree of arbitrariness in the solution which follows from the Riquier method. We finally have discussed to some extent the origin of conflicting statements in the literature concerning such estimates and have illustrated our remarks by some simple model problems.

In a forthcoming paper, we will work out an application on a case where there is degeneracy in the second-order passivity conditions, namely the generalization to n degrees of freedom of what Douglas identified as the separability Case IIa1.

Acknowledgements We are indebted to Geoff Prince and Gerard Thompson for many stimulating discussions. This research was partially supported by NATO Collaborative Research Grant No. CRG 940195. W.S. thanks the Fund for Scientific Research – Flanders (Belgium) for continuing support. W.S. and M.C. in particular thank La Trobe University, where part of this research was conducted, for its hospitality and financial support. We thank one of the referees for a number of useful comments.

A The hierarchy of algebraic conditions

The purpose of this appendix is to prove the proposition formulated at the end of Section 2. With this in mind, we first prove two lemmas whose function is to show that all further integrability conditions which might arise from acting with D^H on members of the lists (17) and (21) belong to the list (21). The inspiration comes directly from the commutator property (29).

The basic assumption for the lemmas below is that g satisfies the differential conditions (11), (12), (16), and the algebraic conditions (13), (20) (i.e. the $k = 0$ elements of both lists). The summation symbol $\sum$, when it appears without limits, indicates that a cyclic sum is to be taken.

Each of the lemmas involves an instance of the following general construction. If P is a vector-valued k-form along π_1^0, we define the type $(1, k+1)$ tensor field $P_{\Phi,R}$ by

$$P_{\Phi,R} = -\Phi \lrcorner \mathrm{D}^V P + \mu_\Psi P. \tag{50}$$

Lemma A1 *Let P be a vector-valued 1-form along π_1^0, satisfying*

$$g(P(X),Y) = g(P(Y),X);$$

then

$$\sum \Big[g(P_{\Phi,R}(X,Y),Z) - g(P_{\Phi,R}(Y,X),Z)\Big] = 0. \tag{51}$$

Proof We have (from (31) in combination with (27), (28) and (30))

$$\begin{aligned} P_{\Phi,R}(X,Y) \quad = \quad & -\mathrm{D}^V P(\Phi(X),Y) + \mathrm{D}^V\Phi(P(Y),X) + R(X,P(Y)) \\ & - P(\mathrm{D}^V\Phi(Y,X) + R(X,Y)). \end{aligned}$$

The first term, using the D^V-prolongation of the assumption on P (a relation of the form (18)), gives rise to

$$\begin{aligned} & -\sum \Big[g(\mathrm{D}^V P(\Phi(X),Y),Z) - g(\mathrm{D}^V P(\Phi(X),Z),Y)\Big] \\ & \quad = \sum \Big[\mathrm{D}^V g(\Phi(X),P(Y),Z) - \mathrm{D}^V g(\Phi(Z),P(Y),X)\Big], \end{aligned}$$

where on both sides we have used the freedom coming from the cyclic sum to rearrange arguments in the second term. Similar manipulations on the expressions generated by the second term of $P_{\Phi,R}(X,Y)$ give rise to exactly the same terms, with a minus sign.

The third term, again using the cyclic sum freedom, generates

$$\sum \Big[g(R(Y,P(Z)),X) - g(R(X,P(Z)),Y)\Big] = -\sum g(R(X,Y),P(Z)).$$

Here we have used (20) in the last transition. Finally, for the last combination of terms in $P_{\Phi,R}(X,Y)$, the $\mathrm{D}^V\Phi$ terms, due to the skew-symmetrization and (19), give rise to three curvature terms which precisely compensate the two other curvature terms in the same combination and the one generated above. The result follows. □

Lemma A2 *Let P be a vector-valued 2-form along π_1^0, satisfying*

$$\sum g(P(X,Y),Z) = 0;$$

then

$$\sum_{i=1}^{4} (-1)^{i-1} g\left(\sum P_{\Phi,R}(X_j,X_k,X_l),X_i\right) = 0. \tag{52}$$

It is to be understood here that (i,j,k,l) is a cyclic permutation of $(1,2,3,4)$, and that for each fixed value of i a cyclic sum is to be taken over the remaining arguments (indicated by the symbol $\sum$ without subscripts). The above equation is a concise way of writing an expression such as the one preceding equation (24).

PROOF This time we have

$$P_{\Phi,R}(X_j,X_k,X_l) = -\mathrm{D}^V P(\Phi(X_j),X_k,X_l) + \mathrm{D}^V\Phi(P(X_k,X_l),X_j) + R(X_j,P(X_k,X_l))$$
$$- P(\mathrm{D}^V\Phi(X_k,X_j),X_l) - P(X_k,\mathrm{D}^V\Phi(X_l,X_j)) - P(R(X_j,X_k),X_l) - P(X_k,R(X_j,X_l)).$$

One can easily verify that in a sum of type (52) the X_i argument can be interchanged with another argument, provided the expression is skew-symmetric in the remaining two arguments; the sign has to be changed in this process. Looking at the first term of $P_{\Phi,R}(X_j,X_k,X_l)$ with this in mind, and using also the D^V prolongation of the assumption on P (cf. (23)), we obtain:

$$\begin{aligned}
&-\sum_{i=1}^{4}(-1)^{i-1} g\left(\sum \mathrm{D}^V P(\Phi(X_j),X_k,X_l),X_i\right)\\
&\quad = +\sum_{i=1}^{4}(-1)^{i-1}\sum g\left(\mathrm{D}^V P(\Phi(X_i),X_j,X_k),X_l\right)\\
&\quad = -\sum_{i=1}^{4}(-1)^{i-1}\sum \mathrm{D}^V g(\Phi(X_i),P(X_j,X_k),X_l).
\end{aligned}$$

Likewise, from the second term, using the skew-symmetry of P and the prolongation (18) we can write

$$\begin{aligned}
&\sum_{i=1}^{4}(-1)^{i-1} g\left(\sum \mathrm{D}^V\Phi(P(X_j,X_k),X_l),X_i\right)\\
&\quad = -\sum_{i=1}^{4}(-1)^{i-1}\sum g\left(\mathrm{D}^V\Phi(P(X_j,X_k),X_i),X_l\right)\\
&\quad = -\tfrac{1}{2}\sum_{i=1}^{4}(-1)^{i-1}\sum\Big[g(\mathrm{D}^V\Phi(P(X_j,X_k),X_i),X_l) - g(\mathrm{D}^V\Phi(P(X_j,X_k),X_l),X_i)\Big]\\
&\quad = \tfrac{1}{2}\sum_{i=1}^{4}(-1)^{i-1}\sum\Big[\mathrm{D}^V g(P(X_j,X_k),\Phi(X_i),X_l) - \mathrm{D}^V g(P(X_j,X_k),\Phi(X_l),X_i)\Big]\\
&\quad = \sum_{i=1}^{4}(-1)^{i-1}\sum \mathrm{D}^V g(P(X_j,X_k),\Phi(X_i),X_l).
\end{aligned}$$

It is obvious that the two expressions so far computed cancel each other.

The four P terms in $P_{\Phi,R}(X_j,X_k,X_l)$ lead in a direct way to

$$\sum_{i=1}^{4}(-1)^{i-1} g\left(\sum P(R(X_j,X_k),X_l),X_i\right).$$

It remains to consider $\sum_{i=1}^{4}(-1)^{i-1} g\left(\sum R(X_j,P(X_k,X_l)),X_i\right)$. The calculation which follows shows that, using (20) and the similar assumption on P to bring the R inside

P, the result cancels the above term. The cyclic sum freedom and the skew-symmetry of both R and P is exploited in various places.

$$\begin{aligned}
&\sum_{i=1}^{4}(-1)^{i-1} g\left(\sum R(X_j, P(X_k, X_l), X_i\right) \\
&\quad = \tfrac{1}{2}\sum_{i=1}^{4}(-1)^{i-1}\sum\Big[g(R(X_j, P(X_k, X_l)), X_i) + g(R(P(X_k, X_l), X_i), X_j)\Big] \\
&\quad = -\tfrac{1}{2}\sum_{i=1}^{4}(-1)^{i-1}\sum g(R(X_i, X_j), P(X_k, X_l)) \\
&\quad = \tfrac{1}{2}\sum_{i=1}^{4}(-1)^{i-1}\sum\Big[g(X_k, P(X_l, R(X_i, X_j))) + g(X_k, P(R(X_i, X_l), X_j))\Big] \\
&\quad = -\tfrac{1}{2}\sum_{i=1}^{4}(-1)^{i-1}\sum\Big[g(X_i, P(X_l, R(X_k, X_j))) + g(X_i, P(R(X_k, X_l), X_j))\Big] \\
&\quad = -\sum_{i=1}^{4}(-1)^{i-1} g\left(X_i, \sum P(R(X_j, X_k), X_l)\right).
\end{aligned}$$

The final step which makes X_i the first argument of g relies on the skew-symmetry in X_l, X_j of the whole expression between square brackets. □

Proof of Proposition 1

If we assume that g satisfies the differential conditions of Section 2 and the algebraic requirements (17),(21) up to order l, then obviously, compatibility with $\nabla g = 0$ already generates the next algebraic conditions of order $l+1$. Compatibility with the D^V and D^H condition, following the procedure carried out in Section 2 for the first step, will at first create conditions of the following form (where D stands for either D^V or D^H):

$$\begin{aligned}
\sum\Big[g(\mathrm{D}\nabla^l\Phi(X,Y), Z) - g(\mathrm{D}\nabla^l\Phi(Y,X), Z)\Big] &= 0 \\
\sum_{i=1}^{4}(-1)^{i-1} g\left(\sum \mathrm{D}\nabla^l R(X_j, X_k, X_l), X_i\right) &= 0.
\end{aligned}$$

For $l = 0$ we have seen in Section 2 that $\mathrm{D} = \mathrm{D}^V$ creates the R condition which is already assumed in the present context, whereas $\mathrm{D} = \mathrm{D}^H$ creates the ∇R condition, which has been obtained at this level from the compatibility with $\nabla g = 0$. For $l > 0$ we can commute D with one ∇ operator. In the case of $\mathrm{D} = \mathrm{D}^H$, the result is, using also the fact that $\nabla g = 0$, an expression which is simply the ∇ prolongation of the condition obtained at the previous stage, plus an expression which is of the form of the left-hand sides of (51) or (52) respectively, with respectively $P = \nabla^{l-1}\Phi$, and $P = \nabla^{l-1}R$. The above lemmas just prove that these terms are zero as a result of the assumptions on P. In the case of $\mathrm{D} = \mathrm{D}^V$, the situation is even simpler as we are led back essentially to the

D^H analysis of the preceding case in view of the commutator (15). The result follows by induction. □

B Alternants with the second-order conditions

With the A-operator in the form (40), applying ∇ to the second-order conditions, using consecutively the commutators (15) and (29) and the identity (35) for the resulting double D^H derivatives, we obtain

$$\begin{aligned}&(\mathrm{D}^V(\Phi\lrcorner\mathrm{D}^V)g)\,(X,Y,U,Z) - (\mathrm{D}^V(\Phi\lrcorner\mathrm{D}^V)g)\,(Y,X,U,Z) =\\ &\quad (\mathrm{D}^V i_\Psi g)(Y,X,U,Z) - (\mathrm{D}^V i_\Psi g)(X,Y,U,Z) + \mathrm{D}^V g\,(R(X,Y),U,Z)\\ &\quad - i_{\mathrm{Rie}(X,Y)}g\,(U,Z) + g\,(X,\nabla\theta(U,Z)Y) - g\,(\nabla\theta(U,Z)X,Y).\end{aligned}$$

Let us sketch how the various parts of this equation can be manipulated further. The terms in the left-hand side can be brought to the form

$$\mathrm{D}^V\mathrm{D}^V g\,(X,\Phi Y,U,Z) - \mathrm{D}^V\mathrm{D}^V g\,(Y,\Phi X,U,Z) + 3\,\mathrm{D}^V g\,(R(X,Y),U,Z).$$

The terms involving Ψ on the right can be reduced in the first place to

$$i_{X\lrcorner\mathrm{D}^V_Y\Psi}g - i_{Y\lrcorner\mathrm{D}^V_X\Psi}g + i_{X\lrcorner\Psi}\mathrm{D}^V_Y g - i_{Y\lrcorner\Psi}\mathrm{D}^V_X g,$$

acting as a covariant 2-tensor on U and Z. The first two (algebraic) terms in the latter expression can be shown to equal $2\,i_{\mathrm{Rie}(X,Y)}g\,(U,Z)$. Thus an intermediate result for the complete equation reads

$$\begin{aligned}&\mathrm{D}^V\mathrm{D}^V g\,(X,\Phi(Y),U,Z) - \mathrm{D}^V\mathrm{D}^V g\,(Y,\Phi(X),U,Z) =\\ &\quad -2\,\mathrm{D}^V g\,(R(X,Y),U,Z) + i_{\mathrm{Rie}(X,Y)}g\,(U,Z) + i_{X\lrcorner\Psi}\mathrm{D}^V_Y g\,(U,Z)\\ &\quad - i_{Y\lrcorner\Psi}\mathrm{D}^V_X g\,(U,Z) + g\,(X,\nabla\theta(U,Z)Y) - g\,(\nabla\theta(U,Z)X,Y).\end{aligned}$$

The two remaining terms involving Ψ, using the defining relation (30), represent a sum of eight terms in $\mathrm{D}^V g$. For the term involving the Rie tensor, we use the property

$$\mathrm{Rie}(X,Y)Z = -\tfrac{1}{3}\mathrm{D}^V\mathrm{D}^V\Phi(Z,X,Y) + \tfrac{1}{3}\mathrm{D}^V\mathrm{D}^V\Phi(Z,Y,X) \qquad (53)$$

which follows from (38) and (19). The second derivative terms in the left-hand side are replaced by using a second D^V prolongation of the algebraic condition (13); it reads

$$\begin{aligned}&\mathrm{D}^V\mathrm{D}^V g\,(U,Z,\Phi(X),Y) - \mathrm{D}^V\mathrm{D}^V g\,(U,Z,\Phi(Y),X) =\\ &\quad \mathrm{D}^V g\,(Z,\mathrm{D}^V\Phi(U,Y),X) + \mathrm{D}^V g\,(U,\mathrm{D}^V\Phi(Z,Y),X) - \mathrm{D}^V g\,(Z,\mathrm{D}^V\Phi(U,X),Y)\\ &\quad - \mathrm{D}^V g\,(U,\mathrm{D}^V\Phi(Z,X),Y) + g\,(\mathrm{D}^V\mathrm{D}^V\Phi(U,Z,Y),X) - g\,(\mathrm{D}^V\mathrm{D}^V\Phi(U,Z,X),Y).\end{aligned}$$

When putting this all together one has to keep it in mind that in expressions involving $\mathrm{D}^V g$ one can write the three arguments in any desired order. Moreover, in view of the fact that there is no "vertical curvature", expressed by (34), the four arguments in $\mathrm{D}^V\mathrm{D}^V g$ may also be written in any preferred order. One finds in this way that a number of terms will cancel out as a result of the D^V prolongation (23) of the algebraic condition (20). The remaining terms are all algebraic in g and have either $\mathrm{D}^V\mathrm{D}^V\Phi$ or $\nabla\theta$ in one of their arguments. When we appeal to the Bianchi identity (45) these terms also finally cancel out.

Next, considering a D^V prolongation of the second-order conditions, we get relations of the form

$$\begin{aligned}\mathrm{D}^V\mathrm{D}^V\mathrm{D}^H g\,(W,X,Y,U,Z) &= \mathrm{D}^V\mathrm{D}^V\mathrm{D}^H g\,(W,Y,X,U,Z) + \mathrm{D}^V g\,(W,\theta(U,Z)Y,X)\\ &\quad - \mathrm{D}^V g\,(W,\theta(U,Z)X,Y) + g\,(\mathrm{D}^V\theta(W,U,Z)Y,X) - g\,(\mathrm{D}^V\theta(W,U,Z)X,Y).\end{aligned}$$

For potential compatibility problems, this should be compared to the following $\mathrm{D}^V\mathrm{D}^H$ prolongation of (12):

$$\mathrm{D}^V\mathrm{D}^H\mathrm{D}^V g\,(X,Y,W,U,Z) = \mathrm{D}^V\mathrm{D}^H\mathrm{D}^V g\,(X,Y,Z,U,W).$$

Indeed, one can use a D^V prolongation of the identity (36) applied to g to replace these third-order derivatives by derivatives of the form $\mathrm{D}^V\mathrm{D}^V\mathrm{D}^H g$ also, with an interchange of the second and third argument. The leading term on the left, in view of (34) again, becomes identical to the left-hand side of the first equation, and the compatibility consists of matching all the other terms. What we get, making use of (34) and the symmetry of $\mathrm{D}^V\theta$ (the Bianchi identity (46)), is the intermediate condition

$$\begin{aligned}&\mathrm{D}^V\mathrm{D}^V\mathrm{D}^H g\,(Y,W,X,U,Z) - \mathrm{D}^V\mathrm{D}^V\mathrm{D}^H g\,(Z,X,Y,U,W) =\\ &\quad \mathrm{D}^V g\,(Y,\theta(U,Z)X,W) - \mathrm{D}^V g\,(Z,\theta(W,Y)U,X) + g\,(\mathrm{D}^V\theta(Y,U,Z)X,W)\\ &\quad - g\,(\mathrm{D}^V\theta(Y,U,Z)W,X) + g\,(\mathrm{D}^V\theta(Z,U,W)X,Y) - g\,(\mathrm{D}^V\theta(X,W,Y)U,Z).\end{aligned}$$

Recalling that all such calculations can be freely changed modulo prolongations, we make three further substitutions for the highest order derivatives. The first one is taken from the D^V_Y prolongation of $(\mathcal{A}(W,X)g)\,(U,Z)=0$. The second one comes from the D^V_Z prolongation of $(\mathcal{A}(X,Y)g)\,(U,W)=0$. In the resulting expression we swap the first two arguments in the leading term by appealing to (34) again, and then make a third and final substitution from the D^V_Y prolongation of $(\mathcal{A}(Z,X)g)\,(U,W)=0$. In the course of these calculations, one repeatedly simplifies the expression by using the symmetry properties of θ and $\mathrm{D}^V\theta$. The result finally reads

$$\mathrm{D}^V\mathrm{D}^V\mathrm{D}^H g\,(Y,X,W,U,Z) = \mathrm{D}^V\mathrm{D}^V\mathrm{D}^H g\,(Y,X,Z,U,W),$$

and this is nothing but a second-order prolongation of the equation $\mathrm{D}^H g\,(W,U,Z) = \mathrm{D}^H g\,(Z,U,W)$. Hence, no integrability or passivity conditions are obtained.

We trust that a reader who has had the courage to get to this point will be able to see how to do the similar calculation for D^H. Very briefly, it is best to start from the expression (41) of the A operator (i.e. the second-order conditions in the form (39)); the manipulations on the leading terms are of the same form as above with repeated use of the identity (35) to swap arguments; that all algebraic terms cancel out in the end this time follows from the Bianchi identity (47), and the second D^V prolongation of the curvature condition (20).

C Elements of the theory of formally integrable systems

Our purpose in this appendix is to illustrate with a few simple examples how the comptutation of Cartan characters works and how this possibly relates to the counting procedure for the arbitrariness in the solutions, discussed in Section 5. We will show, in particular, that our counting procedure for the solution of the second-order equations on g_{1n} fits the process of checking involutivity of the symbol and relates to the last non-zero character. The reader may consult e.g. [2, 15], or the last chapter in [14] for an account of the Cartan-Kähler theory for partial differential equations. A concise introduction may be found also in [21].

Let us start with an equation for which we know the general solution, for example, the wave equation

$$\rho_{22} - \rho_{11} = 0,$$

a second-order equation for 1 unknown ρ in $N = 2$ independent variables x^1, x^2, where the subscripts indicate partial differentiations with respect to the corresponding variables. This equation defines a symbol G_2 of dimension 2: there is one linear relation among the three second-order derivative coordinates ρ_{11}, ρ_{12}, ρ_{22} on the appropriate jet space. According to [15], we should divide these coordinates, or 'components', into separate classes in the following way: ρ_{11}, ρ_{12} are the components of class 1, and ρ_{22} the component of class 2. If $(G_2)^i$ is the space formed from the elements of G_2 with zero components of class $1, \ldots, i$, with $(G_2)^0 = G_2$ and $(G_2)^N = 0$, then the Cartan characters α^i could be defined by

$$\alpha^i = \dim(G_2)^{i-1} - \dim(G_2)^i.$$

Alternatively, one can try to write as many equations as possible with class i coordinates as principal derivatives, starting with class N. The number of such equations in class i is denoted by β^i. There is a direct relation between α^i and β^i, which for a system of second-order equations in 1 unknown and N independent variables is

$$\alpha^i = (N - i + 1) - \beta^i,$$

and the sequences α^i and β^i are also related by the general property

$$\sum_{k=1}^{N} k\beta^k = \binom{N+2}{3} - \sum_{k=1}^{N} k\alpha^k.$$

The symbol is involutive if $\sum k\beta^k$ is equal to the number of independent equations in the prolongation of the defining relations (equivalently, $\sum k\alpha^k$ is the dimension of the symbol defined by the prolonged equations). One of the difficulties is that the computation of the numbers α^i and β^i is very much coordinate dependent, and the test of involutivity can only be positive with respect to so-called δ-regular (or quasi-regular) coordinates. Such coordinates correspond to a maximization of the sequence β^i, starting as before with β^N. An algorithmic way of constructing δ-regular coordinates is concisely described in the introduction of [21].

For the simple wave equation above, it is clear that $\beta^2 = 1$, $\beta^1 = 0$ and these are the maximal numbers one can obtain. Also, $\alpha^1 = 2$, $\alpha^2 = 0$. We have $\sum k\beta^k = 2$, and this is indeed the number of equations one obtains by prolonging the original one. As expected, we have involutivity, and since no integrability conditions will arise from the prolongation the equation is formally integrable. The Cartan-Kähler theorem then includes a statement about the number of arbitrary functions in the general solution and the number of variables they depend on, and these numbers are computed from the Cartan characters α^i. But there are different versions of this result in the literature which contradict each other (see [20]). For the present example, nothing can go wrong because there is only one non-zero α, and so the statement will be that the solution depends on two functions (the value of α^1) of one variable (the superscript of α^1), which is in agreement with our knowledge about the general solution of the wave equation. It is clear that the method explained in Section 5 gives the same answer.

Now, let us confuse the issue by considering again the wave equation, but in the form

$$\rho_{44} - \rho_{11} = 0,$$

where $N = 4$ this time, but the variables x_2, x_3 do not appear in the given equations explicitly. Computing the numbers α^i and β^i as before, one finds that $\beta^4 = 1$, $\beta^3 = \beta^2 = \beta^1 = 0$, and $\alpha^1 = 4$, $\alpha^2 = 3$, $\alpha^3 = 2$ and $\alpha^4 = 0$. We have: $\sum k\beta^k = 4$, which is indeed the number of (independent) equations in the prolongation, so the symbol is involutive. Pommaret's version of the Cartan-Kähler theorem (see [15], p.160) would now state that the freedom in the solution will be $4\,f(1) + 3\,f(2) + 2\,f(3)$. Clearly this is not the most economical measure, because we know the general solution: it is the same as before, but depending parametrically on the extra variables x_2, x_3, so the optimal answer should be $2\,f(3)$. This is again an illustration of the fact that really only the last non-zero character (in this case α^3) has an intrinsic meaning and should be taken into account in describing the arbitrariness in the solution (cf. [14]). It is rather odd that while [2] does mention this at some stage, it also contains a statement (on p.87) which seems to indicate that all non-zero Cartan characters contribute to the arbitrariness in the solution.

The counting procedure of Section 5 indicates that the solution to this problem will depend on $2\,f(3)$, the correct answer. In fact the equation $\rho_{44} - \rho_{11} = 0$ is equivalent, so far as this analysis goes, to the second-order equation for g_{12} for $n = 2$ discussed in Section 5: the two equations can be converted one into the other by a linear change of coordinates, modulo terms of degree less than 2.

To conclude this discussion we shall examine the general structure of the second-order conditions for $\rho = g_{1n}$ of Sections 4 and 5, for arbitrary n. The first point to be made is that the test of involutivity of the symbol, as it is described in [15], p.92, can be carried out at a fixed generic point of $J^1\pi$. It is easy to see that, given any point $p \in J^1\pi$, there is a coordinate transformation of jet coordinates which makes Γ^i_j zero at p: this is a consequence of the fact that Γ^i_j transforms affinely under a change of jet coordinates.

So without loss of generality we may assume in our calculations that $\Gamma^i_j = 0$. For convenience let us relabel the coordinates as in Section 4, so that $v^i \mapsto x^{n+i}$. Then for arbitrary n the whole system can be cast into the following form, modulo terms of lower degree:

$$\rho_{k,n+l} - \rho_{l,n+k} = 0, \tag{54}$$

with $k = 1, \ldots, n-1$, $l = 2, \ldots, n$ and $k < l$. The number of equations corresponds therefore to the number of independent elements in a skew-symmetric $n \times n$ matrix, which is $\binom{n}{2} = \frac{1}{2}n(n-1)$.

When we prolong these equations, we have to write down for a start $2n$ times that many equations, but there may be redundancies. As a matter of fact, the following identities hold amongst the 3-jets of ρ as a consequence of equations (54):

$$\sum(\rho_{k,n+l,m} - \rho_{k,n+m,l}) = 0$$

$$\sum(\rho_{k,n+l,n+m} - \rho_{k,n+m,n+l} = 0,$$

the sums being cyclic sums over k, l and m. Since k, l and m must be different, we can take $1 \le k < l < m \le n$. This gives the complete set of independent identities on the 3-jets of ρ; their number is twice the number of independent elements in an arbitrary completely skew 3-tensor in n dimensions, which is $2 \times \binom{n}{3} = \frac{1}{3}n(n-1)(n-2)$. It follows that the number of independent equations in the prolongation of the system (54) is

$$2n \times \tfrac{1}{2}n(n-1) - \tfrac{1}{3}n(n-1)(n-2) = \tfrac{2}{3}n(n^2-1). \tag{55}$$

The route to maximizing the numbers β^i goes through transforming terms like $\rho_{k,n+l}$ into $\rho_{aa} - \rho_{bb}$ via a change of coordinates of the form $y_a = x_k + v_l$, $y_b = x_k - v_l$. We can use such transformations to write the $\frac{1}{2}n(n-1)$ equations (54) in such a form that the $\frac{1}{2}n(n-1)$ derivatives

$$\rho_{2n-i,2n-j} \qquad \text{for} \quad i,j=0,\ldots,n-2 \quad \text{with} \quad i \geq j$$

appear as principal derivatives. In such coordinates we will have optimal values for the highest $n-1$ numbers β^i, namely

$$\begin{aligned} \beta^{2n-(k-1)} &= k \quad \text{for} \quad k=1,\ldots,n-1 \\ &= 0 \quad \text{for} \quad k \geq n. \end{aligned}$$

It follows that

$$\sum_{k=1}^{2n} k\,\beta^k = \sum_{k=1}^{n-1} (2n-(k-1))\,k,$$

which can easily be shown to equal $\frac{2}{3}n(n^2-1)$, the number of independent equations in the prolongation of the system as given in (55). We conclude that the symbol is involutive. Furthermore, it follows from the relation $\alpha^i = (2n-i+1) - \beta^i$ that the highest non-zero character is α^{n+1} and its value is n. The formula for the freedom in the solution for ρ found in Section 5 is $n\,f(n+1)$; for this system considered in its own right, this is consistent with the result given by the last non-zero Cartan character.

References

[1] I. Anderson and G. Thompson, The inverse problem of the calculus of variations for ordinary differential equations, *Memoirs Amer. Math. Soc.* **98** No. 473 (1992).

[2] R.L. Bryant, S.S. Chern, R.B. Gardner, H.L. Goldschmidt and P.A. Griffiths, *Exterior differential systems*, Math. Sciences Res. Inst. Public. **18**, (Springer-Verlag, New York) (1991).

[3] F. Cantrijn, W. Sarlet, A. Vandecasteele and E. Martínez, Complete separability of time-dependent second-order ordinary differential equations, *Acta Appl. Math.* **42** (1996) 309–334.

[4] M. Crampin, E. Martínez and W. Sarlet, Linear connections for systems of second-order ordinary differential equations, *Ann. Inst. H.Poincaré* **65** (1996) 223–249.

[5] M. Crampin, W. Sarlet, E. Martínez, G.B. Byrnes and G.E. Prince, Towards a geometrical understanding of Douglas's solution of the inverse problem of the calculus of variations, *Inverse problems* **10** (1994) 245–260.

[6] J. Douglas, Solution of the inverse problem of the calculus of variations, *Trans. Amer. Math. Soc.* **50** (1941) 71–128.

[7] J. Grifone and Z. Muzsnay, Sur le problème inverse du calcul des variations: existence de lagrangiens associées à un spray dans le cas isotrope, Laboratoire de Mathématiques, Université Toulouse III, Preprint Nr. 78 (1996).

[8] M. Janet, Sur les systèmes d'équations aux dérivées partielles, *J. Math. Pure Appl.* **3** (1920) 65–151.

[9] M. Janet, *Leçons sur les systèmes d'équations aux dérivées partielles*, Cahiers Scientifiques, Fascicule IV (Gauthier-Villars, Paris) (1929).

[10] E. Martínez, J.F. Cariñena and W. Sarlet, Derivations of differential forms along the tangent bundle projection. Part II, *Diff. Geometry and its Applications* **3** (1993) 1–29.

[11] E. Martínez, J.F. Cariñena and W. Sarlet, Geometric characterization of separable second-order differential equations, *Math. Proc. Camb. Phil. Soc.* **113** (1993) 205–224.

[12] G. Morandi, C. Ferrario, G. Lo Vecchio, G. Marmo and C. Rubano, The inverse problem in the calculus of variations and the geometry of the tangent bundle, *Phys. Rep.* **188** (1990) 147–284.

[13] Z. Muzsnay, *Sur le problème inverse du calcul des variations*, Thèse de Doctorat, Université Paul Sabatier (Toulouse III) (1997).

[14] P.J. Olver, *Equivalence, invariants, and symmetry*, (Cambridge University Press, Cambridge) (1995).

[15] J.F. Pommaret, *Systems of partial differential equations and Lie pseudogroups*, (Gordon and Breach, London) (1978).

[16] G.J. Reid, Algorithms for reducing a system of PDE's to standard form, determining the dimension of its solution space and calculating its Taylor series solution, *Eur. J. Appl. Math.* **2** (1991) 293–318.

[17] C. Riquier, *Les systèmes d'équations aux derivées partielles*, (Gauthier-Villars, Paris) (1910).

[18] W. Sarlet, The Helmholtz conditions revisited. A new approach to the inverse problem of Lagrangian dynamics, *J. Phys. A* **15** (1982) 1503–1517.

[19] W. Sarlet, A. Vandecasteele, F. Cantrijn and E. Martínez, Derivations of forms along a map: the framework for time-dependent second-order equations, *Diff. Geometry and its Applications* **5** (1995) 171–203.

[20] W.M. Seiler, On the arbitrariness of the general solution of an involutive partial differential equation, *J. Math. Phys.* **35** (1994) 486–498.

[21] W.M. Seiler and R.W. Tucker, Involution of constrained dynamics I: the Dirac approach, *J. Phys. A* **28** (1995) 4431–4451.

[22] V.L. Topunov, Reducing systems of linear differential equations to passive form, *Acta Appl. Math.* **16** (1989) 191–206.