Hydon, Peter E. (2023) *Partial Euler operators and the efficient inversion of Div.* European Journal of Applied Mathematics . ISSN 0956-7925.

PAPER

Partial Euler operators and the efficient inversion of Div

P. E. Hydon

School of Mathematics, Statistics and Actuarial Science, University of Kent, Canterbury CT2 7FS, UK
Email: P.E.Hydon@kent.ac.uk

Abstract
The problem of inverting the total divergence operator is central to finding components of a given conservation law. This might not be taxing for a low-order conservation law of a scalar partial differential equation, but integrable systems have conservation laws of arbitrarily high order that must be found with the aid of computer algebra. Even low-order conservation laws of complex systems can be hard to find and invert. This paper describes a new, efficient approach to the inversion problem. Two main tools are developed: partial Euler operators and partial scalings. These lead to a line integral formula for the inversion of a total derivative and a procedure for inverting a given total divergence concisely.

1. Introduction

Around 20 years ago, Stephen Anco and George Bluman [2, 3] introduced a comprehensive practical method for determining conservation laws of partial differential equations (PDEs) in Kovalevskaya form. The method is based on finding adjoint symmetries and applying Helmholtz conditions.[1] A key part of the calculation is the inversion of the total divergence operator Div to obtain the components of the conservation law. Usually, this can be done by using a homotopy operator, but the following three problems may occur with the standard homotopy formula (which is given by Olver in [8]).

1. The homotopy formula uses definite integrals. This works well if the divergence is a differential polynomial; by contrast, rational polynomials commonly have a singularity at one limit. Hickman [7] and Poole & Hereman [9] suggest avoiding these by working in terms of indefinite integrals, an approach that we use throughout this paper. Alternatively, one can move the singularity by modifying the dependent variable (see Anco & Bluman [3] and Poole & Hereman [9]).

2. Scaling is fundamental to the homotopy approach to inversion. For instance, varying the scaling parameter in the standard homotopy formula moves contributions from variables along a ray to the origin. However, scaling does not change rational polynomials that are homogeneous of degree zero, so the standard inversion process does not work for such terms. Deconinck & Nivala [5] discussed this problem in some detail (for one independent variable only) and suggested possible workarounds, but commented, 'We are unaware of a homotopy method that algorithmically avoids all problems like the ones demonstrated . . .'. Poole & Hereman [9] proposed an approach that works well for problems with one independent variable, but noted the difficulties of extending this to multiple independent variables (in a way that can be programmed).

---

[1] The deep theoretical foundation for this approach is discussed in Olver [8]. For a recent review of the method and its extension beyond equations in Kovalevskaya form, see Anco [1]. In particular, the method may be used for equations in extended Kovalevskaya form (see Popovych & Bihlo [10]).

3. The standard homotopy operator applies to a star-shaped domain and integrates along rays to the origin, changing all Cartesian coordinates at once. This gives an inefficient inversion, in the sense that the number of terms is generally very much greater than necessary; the homotopy formula creates superfluous elements of ker(Div). For polynomial divergences, Poole & Hereman [9] removed curls by parametrizing all terms given by the homotopy formula and optimizing the resulting linear system. This approach is very effective, because (except where there are cancellations) the homotopy formula tends to include every possible term that can appear in an inversion. However, it is unclear whether this approach can be generalized to non-polynomial divergences. Moreover, the removal of curl terms takes extra processing time and does not allow for the possibility that new terms might appear in the most concise inversion of the divergence. If inversion could be done with respect to one independent variable at a time, this might prevent the occurrence of superfluous curls from the outset.

Example 1.1. *To illustrate the inefficiency of homotopy operators on star-shaped domains, consider the following divergence in $\mathbb{R}^3$:*

$$\mathcal{C}(x, y, z) = 2xy \cos z.$$

The homotopy operator based on a star-shaped domain gives $\mathcal{C} = \mathrm{div}(x\phi, y\phi, z\phi)$, where

$$\phi = \int_0^1 \lambda^2 \mathcal{C}(\lambda x, \lambda y, \lambda z) \, d\lambda$$

$$= 2xy \left\{ \left(z^{-1} - 12z^{-3} + 24z^{-5}\right) \sin z + \left(4z^{-2} - 24z^{-4}\right) \cos z \right\}.$$

By comparison, for a given divergence $\mathcal{C}(x, y, z)$ that has no singularities on the coordinate axes, using a homotopy formula that integrates one variable at a time gives $\mathcal{C} = \mathrm{div}(F, G, H)$, where

$$F = \int_0^x \mathcal{C}(\lambda, y, z) \, d\lambda, \qquad G = \int_0^y \mathcal{C}(0, \lambda, z) \, d\lambda, \qquad H = \int_0^z \mathcal{C}(0, 0, \lambda) \, d\lambda.$$

This recovers the concise form $(F, G, H) = (x^2 y \cos z, 0, 0)$. However, the use of the lower limit makes the formula over-elaborate (and unsuitable for treating singularities). Indefinite integration is far more straightforward:

$$\mathcal{C} = \frac{\partial F}{\partial x}, \quad \text{where} \quad F = \int \mathcal{C}(x, y, z) \, dx = x^2 y \cos z.$$

So the homotopy formula for star-shaped domains gives 14 *more terms than the simple form above; the superfluous terms amount to*

$$\mathrm{curl} \left(4xy^2 \left\{ \left(3z^{-2} - 6z^{-4}\right) \sin z - \left(z^{-1} - 6z^{-3}\right) \cos z \right\}, \ x^2 y \sin z, \ xy\phi/2 \right).$$

The current paper extends the efficient one-variable-at-a-time approach to total derivatives. Indefinite integration is used, as advocated by Hickman [7] for standard homotopy operators; it avoids the complications resulting from singularities. From the computational viewpoint, the biggest advantage of integration with respect to just one independent variable is that the major computer algebra systems have efficient procedures for computing antiderivatives.

The keys to inverting a total divergence one variable at a time are 'partial Euler operators' (see Section 3). These enable the inversion of a total derivative D_x to be written as an indefinite line integral. Section 4 introduces a new iterative method for inverting a given total divergence; typically, this does not produce superfluous terms and very few iterations are needed. Furthermore, it can cope with components that are unchanged by the relevant scaling operator.

The methods in this paper are systematic, leading to procedures that are intended to be suitable for implementation in computer algebra systems.

2. Standard differential and homotopy operators

Here is a brief summary of the standard operators that are relevant to total divergences; for further details, see Olver [8]. The independent variables $\mathbf{x} = (x^1, \ldots, x^p)$ are local Cartesian coordinates, and the dependent variables $\mathbf{u} = (u^1, \ldots, u^q)$ may be real- or complex-valued. The Einstein summation convention is used to explain the main ideas and state general results. In examples, commonly used notation is used where this aids clarity. All functions are assumed to be locally smooth, to allow the key ideas to be presented simply.

Derivatives of each u^α are written as $u^\alpha_{\mathbf{J}}$, where $\mathbf{J} = (j^1, \ldots, j^p)$ is a multi-index; each j^i denotes the number of derivatives with respect to x^i, so $u^\alpha_{\mathbf{0}} = u^\alpha$. The variables x^i and $u^\alpha_{\mathbf{J}}$ can be regarded as jet space coordinates. The *total derivative* with respect to x^i,

$$D_i = \frac{\partial}{\partial x^i} + u^\alpha_{\mathbf{J}+\mathbf{1}_i} \frac{\partial}{\partial u^\alpha_{\mathbf{J}}}, \quad \text{where} \quad \mathbf{J} + \mathbf{1}_i = (j^1, \ldots, j^{i-1}, j^i + 1, j^{i+1}, \ldots, j^p),$$

treats each $u^\alpha_{\mathbf{J}}$ as a function of $\mathbf{x}$. To keep the notation concise, write

$$D_{\mathbf{J}} = D_1^{j^1} D_2^{j^2} \cdots D_p^{j^p};$$

note that $u^\alpha_{\mathbf{J}} = D_{\mathbf{J}}(u^\alpha)$. Let $[\mathbf{u}]$ represent $\mathbf{u}$ and finitely many of its derivatives; more generally, square brackets around an expression denote the expression and as many of its total derivatives as are needed.

A *total divergence* is an expression of the form

$$\mathcal{C} = \mathrm{Div}(\mathbf{F}) := D_i F^i(\mathbf{x}, [\mathbf{u}]).$$

(If all F^i depend on $\mathbf{x}$ only, $\mathcal{C}$ is an ordinary divergence.) A conservation law of a given system of partial differential equations (PDEs), $\mathcal{A}_\ell(\mathbf{x}, [\mathbf{u}]) = 0$, $\ell = 1, \ldots, L$, is a total divergence that is zero on all solutions of the system; each F^i is a finite sum of terms. By using elementary algebraic operations (in particular, expanding logarithms of products and products of sums), the number of linearly independent terms may be maximized. When the number of linearly independent terms is maximal for each i, we call the result the *fully expanded* form of $\mathbf{F}$.

When $p > 1$, the p-tuple of components, $\mathbf{F}$, is determined by $\mathcal{C}$ up to a transformation of the form

$$F^i \longmapsto F^i + D_j \left\{ f^{ij}(\mathbf{x}, [\mathbf{u}]) \right\}, \qquad \text{where } f^{ji} = -f^{ij}. \tag{2.1}$$

(If $p = 3$, such a transformation adds a total curl to $\mathbf{F}$.) The total number of terms in $\mathbf{F}$ is the sum of the number of terms in all of the fully expanded components F^i. If this cannot be lowered by any transformation (2.1), we call $\mathbf{F}$ *minimal*. Commonly, there is more than one minimal $\mathbf{F}$, any of which puts the inversion of Div in as concise a form as possible. If $p = 1$, the sole component F^1 (also denoted F) is determined up to an arbitrary constant, so the number of non-constant terms is fixed.

The formal adjoint of a differential operator (with total derivatives), $\mathcal{D}$, is the unique differential operator $\mathcal{D}^\dagger$ such that

$$f \mathcal{D}g - (\mathcal{D}^\dagger f)g$$

is a total divergence for all functions $f(\mathbf{x}, [\mathbf{u}])$ and $g(\mathbf{x}, [\mathbf{u}])$. In particular,

$$(D_{\mathbf{J}})^\dagger = (-D)_{\mathbf{J}} := (-1)^{|\mathbf{J}|} D_{\mathbf{J}}, \quad \text{where} \quad |\mathbf{J}| = j^1 + \cdots + j^p. \tag{2.2}$$

Thus, the (standard) Euler–Lagrange operator corresponding to variations in u^α is

$$\mathbf{E}_{u^\alpha} = (-D)_{\mathbf{J}} \frac{\partial}{\partial u^\alpha_{\mathbf{J}}}.$$

Total divergences satisfy a useful identity: a function $\mathcal{C}(\mathbf{x}, [\mathbf{u}])$ is a total divergence if and only if

$$\mathbf{E}_{u^\alpha}(\mathcal{C}) = 0, \qquad \alpha = 1, \ldots, q. \tag{2.3}$$

Given a Lagrangian function $L(\mathbf{x}, [\mathbf{u}])$, the Euler–Lagrange equations are $\mathbf{E}_{u^\alpha}(L) = 0$. Given a set of Euler–Lagrange equations that are polynomial in the variables $(\mathbf{x}, [\mathbf{u}])$, the function $\overline{L}$ given by the

homotopy formula

$$\overline{L}(\mathbf{x}, [\mathbf{u}]) = \int_0^1 u^\alpha \left\{ \mathbf{E}_{u^\alpha}(L) \right\} \Big|_{[\mathbf{u} \mapsto \lambda \mathbf{u}]} \, d\lambda \tag{2.4}$$

differs from L by a total divergence. (The same applies to many, but not all, non-polynomial Euler–Lagrange equations.)

When $p = 1$, the equation $P(x, [\mathbf{u}]) = D_x F$ is invertible (at least, for polynomial P) by the following standard homotopy formula:

$$F(x, [\mathbf{u}]) = \int_0^1 \sum_{i=1}^\infty D_x^{i-1} \left(u^\alpha \left\{ \sum_{k \geq i} \binom{k}{i} (-D_x)^{k-i} \frac{\partial P(x, [\mathbf{u}])}{\partial (D_x^k u^\alpha)} \right\} \Bigg|_{[\mathbf{u} \mapsto \lambda \mathbf{u}]} \right) d\lambda + \int_0^1 x P(\lambda x, [0]) \, d\lambda. \tag{2.5}$$

The operator acting on P in the braces above is the *higher Euler operator* of order i for $p = 1$. When $p \geq 2$, the standard homotopy formula is similar, but somewhat more complex (see Olver [8] for details); it is based on higher Euler operators and integration along rays in a totally star-shaped domain. The following example illustrates that even for quite simple divergences, this formula commonly yields inversions with many superfluous terms.

Example 2.1. *The Benjamin–Bona–Mahony (BBM) equation, $u_t - u u_x - u_{xxt} = 0$, has a conservation law*

$$\mathcal{C} = D_x F + D_t G = (u^2 + 2u_{xt})(u_t - u u_x - u_{xxt}). \tag{2.6}$$

The standard homotopy formula gives

$$F = -\frac{1}{3} u u_{xxtt} + \frac{2}{3} u_x u_{xtt} - \frac{1}{3} u_t u_{xxt} - \frac{1}{3} u_{xt}^2 - \frac{1}{2} u u_{tt} + \frac{1}{2} u_t^2 - \frac{2}{3} u^2 u_{xt} + \frac{2}{3} u u_x u_t - \frac{1}{4} u^4,$$

$$G = \frac{1}{3} u u_{xxxt} - \frac{1}{3} u_x u_{xxt} - \frac{2}{3} u_{xx} u_{xt} + \frac{1}{2} u u_{xt} + \frac{1}{2} u_x u_t - \frac{1}{3} u^2 u_{xx} - \frac{2}{3} u u_x^2 + \frac{1}{3} u^3,$$

a total of 17 terms. By contrast, careful integration by inspection yields

$$\mathcal{C} = D_x \left(u_t^2 - u_{xt}^2 - u^2 u_{xt} - \frac{1}{4} u^4 \right) + D_t \left(\frac{1}{3} u^3 \right), \tag{2.7}$$

which is minimal, having only five terms in the components.

The homotopy formulae above can be applied or adapted to some, but not all, classes of non-polynomial Lagrangians and divergences.

3. Partial Euler operators and partial scalings

This section introduces some ideas and results that underpin integration with respect to one independent variable at a time. The independent variable over which one integrates is distinguished; this is denoted by x. For instance, if $x = x^1$, replace the derivative index $\mathbf{J}$ by $(\mathbf{I}, j)$, where $j = j^1$ and $\mathbf{I} = (j^2, \ldots, j^p)$. So the dependent variables and their derivatives are denoted

$$u_{\mathbf{I},j}^\alpha = D_x^j u_{\mathbf{I}}^\alpha, \qquad \text{where } u_{\mathbf{I}}^\alpha = u_{\mathbf{I},0}^\alpha.$$

In examples, however, we write each $u_{\mathbf{I}}^\alpha$ more simply (as u, v_y, u_{yt}, and so on), using j for D_x^j.

3.1 Partial Euler operators

The *partial Euler operator* with respect to x and $u_{\mathbf{I}}^\alpha$ is obtained by varying each $u_{\mathbf{I}}^\alpha$ independently, treating x as the sole independent variable:

$$\mathbf{E}_{u_{\mathbf{I}}^\alpha}^x = (-D_x)^j \frac{\partial}{\partial u_{\mathbf{I},j}^\alpha}. \tag{3.1}$$

Consequently, the standard Euler operator with respect to u^α amounts to

$$\mathbf{E}_{u^\alpha} = (-D)_{\mathbf{I}} \mathbf{E}_{u_{\mathbf{I}}^\alpha}^x. \tag{3.2}$$

Similarly, the partial Euler operator with respect to x and $u_{\mathbf{I},k}^\alpha$ is

$$\mathbf{E}_{u_{\mathbf{I},k}^\alpha}^x = (-D_x)^j \frac{\partial}{\partial u_{\mathbf{I},j+k}^\alpha}. \tag{3.3}$$

Note that

$$\mathbf{E}_{u_{\mathbf{I},k}^\alpha}^x = \frac{\partial}{\partial u_{\mathbf{I},k}^\alpha} - D_x \mathbf{E}_{u_{\mathbf{I},k+1}^\alpha}^x. \tag{3.4}$$

The following identities are easily verified; here, $f(\mathbf{x}, [\mathbf{u}])$ is an arbitrary function.

$$\mathbf{E}_{u_{\mathbf{I}}^\alpha}^x (D_x f) = 0, \tag{3.5}$$

$$\mathbf{E}_{u_{\mathbf{I},k}^\alpha}^x (D_x f) = \frac{\partial f}{\partial u_{\mathbf{I},k-1}^\alpha}, \qquad k \geq 1, \tag{3.6}$$

$$\mathbf{E}_{u_{\mathbf{I},k}^\alpha}^x (D_i f) = D_i \left(\mathbf{E}_{u_{\mathbf{I},k}^\alpha}^x (f) \right) + \mathbf{E}_{u_{\mathbf{I}-1_i,k}^\alpha}^x (f), \qquad x^i \neq x, \tag{3.7}$$

where the last term in (3.7) is zero if $j^i = 0$.

3.2 Inversion of D_x

The identities (3.5), (3.6) and (3.7) are central to the inversion of total divergences, including the following inversion of $P = D_x F$ as an indefinite line integral.

Lemma 3.1. *If $P(\mathbf{x}, [\mathbf{u}]) = D_x F$, then, up to an (irrelevant) arbitrary function of all independent variables other than x,*

$$F(\mathbf{x}, [\mathbf{u}]) = \int \left(P - \sum_{k \geq 1} u_{\mathbf{I},k}^\alpha \, \mathbf{E}_{u_{\mathbf{I},k}^\alpha}^x (P) \right) dx + \sum_{k \geq 0} \mathbf{E}_{u_{\mathbf{I},k+1}^\alpha}^x (P) \, du_{\mathbf{I},k}^\alpha. \tag{3.8}$$

Proof. By the identity (3.6),

$$F(\mathbf{x}, [\mathbf{u}]) = \int \frac{\partial F}{\partial x} \, dx + \sum_{k \geq 0} \mathbf{E}_{u_{\mathbf{I},k+1}^\alpha}^x (D_x F) \, du_{\mathbf{I},k}^\alpha.$$

Moreover,

$$P = \frac{\partial F}{\partial x} + \sum_{k \geq 1} u_{\mathbf{I},k}^\alpha \frac{\partial F}{\partial u_{\mathbf{I},k-1}^\alpha} = \frac{\partial F}{\partial x} + \sum_{k \geq 1} u_{\mathbf{I},k}^\alpha \, \mathbf{E}_{u_{\mathbf{I},k}^\alpha}^x (D_x F).$$

Substituting P for $D_x F$ completes the proof. □

Example 3.1. *Locally, away from its singularities, the function*

$$P = \frac{u_{xx} v_y - u_x v_{xy}}{v_y^2} + \frac{u v_x - u_x v}{v(u+v)} + \frac{1}{x} \tag{3.9}$$

belongs to $\text{im}(D_x)$, *but cannot be inverted using the standard homotopy formula. Substituting*

$$\mathbf{E}^x_{u,1}(P) = \frac{\partial P}{\partial u_x} - D_x \frac{\partial P}{\partial u_{xx}} = -\frac{1}{u+v}, \qquad\qquad \mathbf{E}^x_{u,2}(P) = \frac{\partial P}{\partial u_{xx}} = \frac{1}{v_y},$$

$$\mathbf{E}^x_{v,1}(P) = \frac{\partial P}{\partial v_x} = \frac{u}{v(u+v)}, \qquad\qquad \mathbf{E}^x_{v_y,1}(P) = \frac{\partial P}{\partial v_{xy}} = -\frac{u_x}{v_y^2},$$

into (3.8) *yields the inversion:*

$$F = \int \frac{dx}{x} - \frac{du}{u+v} + \frac{du_x}{v_y} + \frac{u\,dv}{v(u+v)} - \frac{u_x\,dv_y}{v_y^2} = \ln\left|\frac{xv}{u+v}\right| + \frac{u_x}{v_y}. \tag{3.10}$$

3.3 Integration by parts

From here on, we will restrict attention to total divergences $\mathcal{C}$ whose fully expanded form has no terms that depend on $\mathbf{x}$ only. Such terms can be inverted easily by evaluating an indefinite integral, as explained in the Introduction. Henceforth, all indefinite integrals denote antiderivatives with the minimal number of terms in their fully expanded form. Any arbitrary constants and functions that would increase the number of terms are set to zero. This restriction facilitates the search for minimal inversions.

The indefinite line integral formula (3.8) is closely related to integration by parts. To see this, we introduce a *positive ranking* on the variables $u^\alpha_\mathbf{J}$; this is a total order $\preceq$ that is subject to two conditions:

$$(i)\ u^\alpha \prec u^\alpha_\mathbf{J}, \quad \mathbf{J} \neq \mathbf{0}, \qquad (ii)\ u^\beta_\mathbf{I} \prec u^\alpha_\mathbf{J} \Longrightarrow D_\mathbf{K} u^\beta_\mathbf{I} \prec D_\mathbf{K} u^\alpha_\mathbf{J}.$$

The *leading part* of a differential function is the sum of terms in the function that depend on the highest-ranked $u^\alpha_\mathbf{J}$, and the *rank* of the function is the rank of its leading part (see Rust *et al.* [12] for details and references). Let $f(\mathbf{x}, [\mathbf{u}])$ denote the leading part of the fully expanded form of F and let $U_{,k}$ denote the highest-ranked $u^\alpha_{\mathbf{I},k}$; then the highest-ranked part of $P = D_x F$ is $U_{,k+1}\partial f/\partial U_{,k}$. Then (3.8) includes the contribution

$$\int \mathbf{E}^x_{U_{,k+1}}(P)\,dU_{,k} = \int \frac{\partial f}{\partial U_{,k}}\,dU_{,k} = f + \text{lower-ranked terms}.$$

Integration by parts gives the same result. Subtracting f from F and iterating shows that evaluating the line integral (3.8) is equivalent to integrating by parts from the highest-ranked terms downwards.

Integration by parts is useful for splitting a differential expression $P(\mathbf{x}, [\mathbf{u}])$, with $P(\mathbf{x}, [0]) = 0$, into $D_x F$ and a remainder, R, whose x-derivatives are of the lowest-possible order. The splitting is achieved by the following procedure.

Procedure A. Integration by parts

Step 0. Choose a positive ranking in which $u^\alpha_{\mathbf{I},0} \prec D_x u^\beta$ for all $\alpha, \mathbf{I}$ and β. (We call such rankings *x-dominant*.) Initialize by setting $F := 0$ and $R := 0$.

Step 1. Identify the highest-ranked $u^\alpha_{\mathbf{I},k}$ in P; denote this by $U_{,k}$. If $k = 0$, add P to R and **stop**. Otherwise, determine the leading part, g, of P.

Step 2. Determine the sum $hU_{,k}$ of all terms in the fully expanded form of g that are of the form $\gamma U_{,k}$, where γ is ranked no higher than $U_{,k-1}$, and let

$$H = \int h\,dU_{,k-1}.$$

Step 3. Update F, R and P, as follows:

$$F := F + H, \qquad R := R + g - hU_{,k}, \qquad P := P - g + hU_{,k} - D_x H.$$

If $P \neq 0$, return to Step 1. Otherwise output F and R, then **stop**.

The reason for choosing an x-dominant ranking is to ensure that the derivative order with respect to x outweighs all other ranking criteria. Consequently, the minimally ranked remainder cannot contain x-derivatives of unnecessarily high order.

Example 3.2. *To produce a concise inversion of a conservation law of the Harry Dym equation (see Example* 4.1 *below), it is necessary to split*

$$P = -\frac{8}{3}u^2 u_{,4} - \frac{16}{3}uu_{,1}u_{,3} - 4uu_{,2}^2 + 4u_{,1}^2 u_{,2} - u^{-1}u_{,1}^4 .$$

Procedure A *gives the splitting*

$$P = D_x\left\{ -\frac{8}{3}u^2 u_{,3} + \frac{4}{3}u_{,1}^3 \right\} - 4uu_{,2}^2 - u^{-1}u_{,1}^4 .$$

Example 3.3. *The ranking criterion in Step* 2 *of Procedure* A *ensures that there are no infinite loops. It is not enough that terms are linear in the highest-ranked x-derivative, as shown by the following splitting of*

$$P = \frac{v_{,3}}{u_y} + \frac{u_{,2}}{v_y}$$

For the positive x-dominant ranking defined by $v \prec u \prec v_y$, Procedure A *yields*

$$P = D_x\left\{ \frac{v_{,2}}{u_y} + \frac{u_{,1}}{v_y} \right\} + \frac{v_{,2}u_{y,1}}{u_y^2} + \frac{u_{,1}v_{y,1}}{v_y^2} .$$

Both terms in the remainder are linear in their highest-ranked x-derivatives, which are $v_{,2}$ and $v_{y,1}$, respectively. However, further integration by parts would return P to a form with a higher-ranked remainder.

3.4 Partial scalings

To investigate partial Euler operators further, it is helpful to use a variant of the homotopy approach. The *partial scaling* (by a positive real parameter, λ) of a function $f(\mathbf{x}, [\mathbf{u}])$ with respect to x and $u_{\mathbf{I}}^\alpha$ is the mapping

$$\sigma_{u_{\mathbf{I}}^\alpha}^x : (f; \lambda) \mapsto f\big|_{\{u_{\mathbf{I},j}^\alpha \mapsto \lambda u_{\mathbf{I},j}^\alpha,\ j \geq 0\}}.$$

Again, each $u_{\mathbf{I}}^\alpha$ is treated as a distinct dependent variable. Note the identity

$$\sigma_{u_{\mathbf{I}}^\alpha}^x D_x = D_x \sigma_{u_{\mathbf{I}}^\alpha}^x . \tag{3.11}$$

Definition 3.1. *The partial scaling $\sigma_{u_{\mathbf{I}}^\alpha}^x$ is a good scaling for a given differential function $f(\mathbf{x}, [\mathbf{u}])$ if*

$$\sigma_{u_{\mathbf{I}}^\alpha}^x(f; \lambda) = \int \frac{\mathrm{d}}{\mathrm{d}\lambda}\left(\sigma_{u_{\mathbf{I}}^\alpha}^x(f; \lambda) \right) \, \mathrm{d}\lambda, \tag{3.12}$$

for all λ in some neighborhood of 1.

By definition, the partial scaling $\sigma_{u_{\mathbf{I}}^\alpha}^x$ fails to be a good scaling for f if and only if there are terms that are independent of λ in the fully expanded form of $\sigma_{u_{\mathbf{I}}^\alpha}^x(f; \lambda)$. The simplest cause of this is that the fully expanded form of f has terms that are independent of $u_{\mathbf{I}}^\alpha$ and its x-derivatives. However, this is not the only cause, as the following example illustrates.

Example 3.4. *The scalings σ_u^y and σ_v^y are not good scalings for*

$$\mathcal{C} = u_x(2u + v_y) - v_x(u_y + 2v_{yy}) + \frac{u_x}{u^2} + \frac{v_{yy}}{v_y} + \frac{2u_y}{u} \ln |u|, \tag{3.13}$$

because (in fully expanded form),

$$\sigma_u^y(\mathcal{C};\lambda) = \lambda(2uu_x - v_xu_y) + \frac{u_x}{\lambda^2 u^2} + \frac{2u_y}{u}\ln(\lambda) + \left\{u_xv_y - 2v_xv_{yy} + \frac{v_{yy}}{v_y} + \frac{2u_y}{u}\ln|u|\right\}, \tag{3.14}$$

$$\sigma_v^y(\mathcal{C};\lambda) = \lambda(u_xv_y - 2v_xv_{yy}) + \left\{2uu_x - v_xu_y + \frac{u_x}{u^2} + \frac{v_{yy}}{v_y} + \frac{2u_y}{u}\ln|u|\right\}. \tag{3.15}$$

The terms in braces are independent of λ; in (3.14) (resp. (3.15)), some of these depend on u (resp. v) and/or its y-derivatives. Part of the scaled logarithmic term is independent of λ, though part survives differentiation. Note that σ_u^y is a good scaling for the term u_x/u^2; the singularity at $u=0$ is not an obstacle.

Lemma 3.2. *The partial scaling $\sigma_{u_{\mathbf{I}}^\alpha}^x$ is a good scaling for $f(\mathbf{x},[\mathbf{u}])$ if and only if*

$$f = \lim_{\lambda \to 1} \int \frac{\mathrm{d}}{\mathrm{d}\lambda}\left(\sigma_{u_{\mathbf{I}}^\alpha}^x(f;\lambda)\right)\,\mathrm{d}\lambda. \tag{3.16}$$

Proof. If $\sigma_{u_{\mathbf{I}}^\alpha}^x$ is a good scaling, (3.16) is a consequence of $f = \sigma_{u_{\mathbf{I}}^\alpha}^x(f;1)$ and local smoothness. Conversely, suppose that (3.16) holds and let μ be a positive real parameter that is independent of λ. Then for μ sufficiently close to 1,

$$\sigma_{u_{\mathbf{I}}^\alpha}^x(f;\mu) = \lim_{\lambda \to 1} \int \frac{\mathrm{d}}{\mathrm{d}\lambda}\left(\sigma_{u_{\mathbf{I}}^\alpha}^x(f;\lambda\mu)\right)\,\mathrm{d}\lambda = \lim_{\lambda \to \mu}\int \frac{\mathrm{d}}{\mathrm{d}\lambda}\left(\sigma_{u_{\mathbf{I}}^\alpha}^x(f;\lambda)\right)\,\mathrm{d}\lambda = \int \frac{\mathrm{d}}{\mathrm{d}\mu}\left(\sigma_{u_{\mathbf{I}}^\alpha}^x(f;\mu)\right)\,\mathrm{d}\mu.$$

Therefore, $\sigma_{u_{\mathbf{I}}^\alpha}^x$ is a good scaling.

The use of the limit in (3.16) is needed to deal with any values of $u_{\mathbf{I},k}^\alpha$ for which the integral is an indeterminate form. For other values, simple substitution of $\lambda = 1$ gives the limit. $\square$

Given a partial scaling $\sigma_{u_{\mathbf{I}}^\alpha}^x$ and a differential function $f(\mathbf{x},[\mathbf{u}])$, let

$$\pi_{u_{\mathbf{I}}^\alpha}^x(f) = \lim_{\lambda \to 1}\int \frac{\mathrm{d}}{\mathrm{d}\lambda}\left(\sigma_{u_{\mathbf{I}}^\alpha}^x(f;\lambda)\right)\,\mathrm{d}\lambda. \tag{3.17}$$

For μ sufficiently close to 1 (using $\widehat{f}$ as shorthand for $\pi_{u_{\mathbf{I}}^\alpha}^x(f)$),

$$\sigma_{u_{\mathbf{I}}^\alpha}^x(\widehat{f};\mu) = \int \frac{\mathrm{d}}{\mathrm{d}\mu}\left(\sigma_{u_{\mathbf{I}}^\alpha}^x(f;\mu)\right)\mathrm{d}\mu = \int \frac{\mathrm{d}}{\mathrm{d}\mu}\left\{\int \frac{\mathrm{d}}{\mathrm{d}\mu}\left(\sigma_{u_{\mathbf{I}}^\alpha}^x(f;\mu)\right)\,\mathrm{d}\mu\right\}\mathrm{d}\mu = \int \frac{\mathrm{d}}{\mathrm{d}\mu}\left(\sigma_{u_{\mathbf{I}}^\alpha}^x(\widehat{f};\mu)\right)\mathrm{d}\mu;$$

the first equality comes from the proof of Lemma 3.2. Therefore, $\sigma_{u_{\mathbf{I}}^\alpha}^x$ is a good scaling for $\pi_{u_{\mathbf{I}}^\alpha}^x(f)$. Moreover, there are no terms in the fully expanded form of the remainder, $f - \pi_{u_{\mathbf{I}}^\alpha}^x(f)$, for which $\sigma_{u_{\mathbf{I}}^\alpha}^x$ is a good scaling, because

$$\frac{\mathrm{d}}{\mathrm{d}\mu}\left(\sigma_{u_{\mathbf{I}}^\alpha}^x(f;\mu)\right) - \frac{\mathrm{d}}{\mathrm{d}\mu}\left(\sigma_{u_{\mathbf{I}}^\alpha}^x(\widehat{f};\mu)\right) = 0.$$

So $\pi_{u_{\mathbf{I}}^\alpha}^x$ is the projection that maps a given function onto the component which has $\sigma_{u_{\mathbf{I}}^\alpha}^x$ as a good scaling.

Definition 3.2. *The partial scaling $\sigma_{u_{\mathbf{I}}^\alpha}^x$ is a poor scaling for a given differential function $f(\mathbf{x},[\mathbf{u}])$ if $f - \pi_{u_{\mathbf{I}}^\alpha}^x(f)$ depends on any $u_{\mathbf{I},k}^\alpha$.*

For instance, both σ_u^y and σ_v^y are poor scalings for (3.13), as explained in Example 3.4. Section 4.3 addresses the problem of inverting divergences such as (3.13) that have poor scalings. First, we develop the inversion process for general divergences. The following results are fundamental.

Theorem 3.1. *Let $f(\mathbf{x},[\mathbf{u}])$ be a differential function.*

1. If $f = D_xF$, then $f \in \ker(\mathbf{E}_{u_{\mathbf{I}}^\alpha}^x)$ for all α and $\mathbf{I}$; moreover,

$$\pi_{u_{\mathbf{I}}^\alpha}^x(F) = \lim_{\lambda \to 1}\int \sum_{j\geq 0} u_{\mathbf{I},j}^\alpha \,\sigma_{u_{\mathbf{I}}^\alpha}^x\left(\mathbf{E}_{u_{\mathbf{I},j+1}^\alpha}^x(f);\lambda\right)\,\mathrm{d}\lambda. \tag{3.18}$$

2. *If $f \in \ker(\mathbf{E}^x_{u^\alpha_{\mathbf{I}}})$, then $\pi^x_{u^\alpha_{\mathbf{I}}}(f) \in \mathrm{im}(D_x)$.*

3. *If $g = \mathbf{E}^x_{u^\alpha_{\mathbf{I}}}(f)$, then, up to terms in $\mathrm{im}(D_x)$,*

$$\pi^x_{u^\alpha_{\mathbf{I}}}(f) = \lim_{\lambda \to 1} \int u^\alpha_{\mathbf{I}} \, \sigma^x_{u^\alpha_{\mathbf{I}}}(g; \lambda) \; \mathrm{d}\lambda. \tag{3.19}$$

Proof. All three statements are proved by expanding $\pi^x_{u^\alpha_{\mathbf{I}}}(f)$:

$$\pi^x_{u^\alpha_{\mathbf{I}}}(f) = \lim_{\lambda \to 1} \int \sum_{j \geq 0} u^\alpha_{\mathbf{I},j} \, \sigma^x_{u^\alpha_{\mathbf{I}}} \left(\frac{\partial f}{\partial u^\alpha_{\mathbf{I},j}}; \lambda \right) \; \mathrm{d}\lambda \tag{3.20}$$

$$= \lim_{\lambda \to 1} \int u^\alpha_{\mathbf{I}} \, \sigma^x_{u^\alpha_{\mathbf{I}}} \left(\mathbf{E}^x_{u^\alpha_{\mathbf{I}}}(f); \lambda \right) \; \mathrm{d}\lambda + D_x h. \tag{3.21}$$

Here, $h(\mathbf{x}, [\mathbf{u}])$ is obtained by integrating by parts, using the identity (3.11).

If $f = D_x F$, the identity (3.5) amounts to $f \in \ker(\mathbf{E}^x_{u^\alpha_{\mathbf{I}}})$. Replace f by F in (3.20) and use the identity (3.6) to obtain (3.18). Statements 2 and 3 come directly from (3.21). □

Note that (3.18) is a homotopy formula for (at least partially) inverting $D_x F$, giving a third way to do this. The line integral formula (3.8) carries out the full inversion in one step, but may take longer to compute.

4. The inversion method for Div

This section introduces a procedure to invert Div, with a ranking heuristic (informed by experience) that is intended to keep the calculation short and efficient. To motivate the procedure, it is helpful to examine a simple example.

Example 4.1. *Wolf et al.* [14] *introduced computer algebra algorithms that can handle general (non-polynomial) conservation laws and used these to derive various rational conservation laws of the Harry Dym equation. In the (unique) x-dominant positive ranking, the equation is $\mathcal{A} = 0$, with*

$$\mathcal{A} = u_t - u^3 u_{,3}.$$

The highest-order conservation law derived in Wolf et al. [14] *is $\mathcal{C} = \mathcal{Q}\mathcal{A}$, where*

$$\mathcal{Q} = -8u u_{,4} - 16 u_{,1} u_{,3} - 12 u^2_{,2} + 12 u^{-1} u^2_{,1} u_{,2} - 3 u^{-2} u^4_{,1}.$$

Note that σ^x_u is a good scaling for $\mathcal{C}$. The first step in inverting $\mathcal{C} = D_x F + D_t G$ is to apply the partial Euler operator $\mathbf{E}^x_u$, to annihilate the term $D_x F$. There are only two independent variables, so the identity (3.2) shortens the calculation to

$$\mathbf{E}^x_u(\mathcal{C}) = D_t(\mathbf{E}^x_{u_t}(\mathcal{C})) = D_t(\mathcal{Q}).$$

Applying (3.19), then using Procedure A to integrate by parts (see Example 3.2) gives

$$\mathcal{Q} = \mathbf{E}^x_u \left\{ -\frac{8}{3} u^2 u_{,4} - \frac{16}{3} u u_{,1} u_{,3} - 4 u u^2_{,2} + 4 u^2_{,1} u_{,2} - u^{-1} u^4_{,1} \right\}$$

$$= \mathbf{E}^x_u \left\{ -4 u u^2_{,2} - u^{-1} u^4_{,1} \right\}.$$

Therefore,

$$\mathcal{C} = D_t \left(-4 u u^2_{,2} - u^{-1} u^4_{,1} \right) + \widetilde{\mathcal{C}},$$

where $\mathbf{E}^x_u(\widetilde{\mathcal{C}}) = 0$. As σ^x_u is a good scaling for

$$\widetilde{\mathcal{C}} = \mathcal{Q}\mathcal{A} + 8 u u_{,2} u_{t,2} + 4 u^{-1} u^3_{,1} u_{t,1} + \{4 u^2_{,2} - u^{-2} u^4_{,1}\} u_t,$$

the second part of Theorem 3.1 *states that* $\widetilde{\mathcal{C}} \in \operatorname{im}(D_x)$; *consequently,*

$$G = -4uu_{,2}^2 - u^{-1}u_{,1}^4 \,.$$

Either the line integral (3.8) *or Procedure* A *completes the inversion, giving* $\widetilde{\mathcal{C}} = D_x F$, *where*

$$F = 8uu_{,2}u_{t,1} - \{8uu_{,3} + 8u_{,1}u_{,2} - 4u^{-1}u_{,1}^3\}u_t + 4u^4u_{,3}^2 + 4u^3u_{,2}^3 - 6u^2u_{,1}^2u_{,2}^2 + 3uu_{,1}^4u_{,2} - \frac{1}{2}u_{,1}^6.$$

The fully expanded form of (F, G) *is minimal, having* 11 *terms rather than the* 12 *terms in Wolf et al.* [14]. *(Note: there is an equivalent conservation law, not in the form* $\mathcal{QA}$, *that has only* 10 *terms.)*

4.1 A single iteration

The basic method for inverting a given total divergence one independent variable at a time works similarly to the example above. Suppose that after n iterations the inversion process has yielded components F_n^i and that an expression of the form $\mathcal{C} = D_i f_n^i$ remains to be inverted. For the next iteration, let $\mathbf{E}_u^x$ be the partial Euler operator that is applied to $\mathcal{C}$. Here, u is one of the variables $u_{\mathbf{I}}^\alpha$, which is chosen to ensure that for each i such that $x^i \neq x$,

$$\mathbf{E}_u^x D_i f_n^i = D_i \mathbf{E}_u^x f_n^i. \tag{4.1}$$

This requires care, in view of the identity (3.7). However, it is achievable by using the variables $u_{\mathbf{I}}^\alpha$ in the order given by a ranking that is discussed in Section 4.2. This ranking is entirely determined by user-defined rankings of the variables x^j and u^α.

Taking (3.5) into account leads to the identity

$$\mathbf{E}_u^x(\mathcal{C}) = \sum_{x^i \neq x} D_i(\mathbf{E}_u^x(f_n^i)), \tag{4.2}$$

which, with together with Theorem 3.1, is the basis of the inversion method. The method works without modification provided that:

- there are no poor scalings for any terms in $\mathcal{C}$;
- the fully expanded form of $\mathcal{C}$ has no terms that are linear in $[\mathbf{u}]$.

We begin by restricting attention to divergences for which these conditions hold, so that

$$\mathbf{E}_u^x(\mathcal{C}) = \sum_{x^i \neq x} D_i(\mathbf{E}_u^x(\pi_u^x f_n^i)), \tag{4.3}$$

where every term in $\mathbf{E}_u^x(\pi_u^x f_n^i)$ depends on $[\mathbf{u}]$. The modifications needed if either condition does not hold are given in Sections 4.3 and 4.4.

The iteration of the inversion process runs as follows. Calculate $\mathbf{E}_u^x(\mathcal{C})$, which is a divergence $D_i P^i$ with no D_x term, by (4.3); it involves at most $p - 1$ (but commonly, very few) nonzero functions P^i. Invert this divergence, treating x as a parameter. *If it is possible to invert in more than one way, always invert into the P^i for which x^i is ranked as low as possible*; the reason for this is given in the next paragraph. If $\mathbf{E}_u^x(\mathcal{C})$ has nonlinear terms that involve derivatives with respect to more than one D_i (excluding D_x), this is accomplished by iterating the inversion process with as few independent variables as are needed. Otherwise, P^i can be determined more quickly by using integration by parts (Procedure A, with x replaced by the appropriate x^i), and/or the method for linear terms (see Procedure B in Section 4.4). Note that this shortcut can be used whenever there are only two independent variables.

At this stage, check that the fully expanded form of each P^i has no terms that are ranked lower than u. If any term is ranked lower than u, stop the calculation and try a different ranking of the variables x^j and/or u^α. This is essential because, to satisfy (4.1) and avoid infinite loops, the variables $u_{\mathbf{I}}^\alpha$ that are chosen to be u in successive iterations must progress upwards through the ranking. Where there is a

choice of inversion, the rank of each term in P^i is maximized by using the x^i of minimum order; this avoids unnecessary re-ranking.

Having found and checked P^i, use (3.21) to obtain

$$\pi_u^x(f_n^i) = \left\{ \lim_{\lambda \to 1} \int u\, \sigma_u^x(P^i; \lambda)\ d\lambda \right\} + D_x h^i, \tag{4.4}$$

for arbitrary functions $h^i(\mathbf{x}, [\mathbf{u}])$. Apply Procedure A to the function in braces and choose h^i to make the right-hand side of (4.4) equal the remainder from this procedure. This yields the representation of $\pi_u^x(f_n^i)$ that has the lowest-order derivatives (with respect to x) consistent with the inversion of $D_i P^i$; call this representation f^i. Commonly, such a lowest-order representation is needed to obtain a minimal inversion.

By Theorem 3.1, there exists ϕ such that

$$\pi_u^x \left(\mathcal{C} - \sum_{x^i \neq x} D_i f^i \right) = D_x \phi, \tag{4.5}$$

because (by construction) the expression in parentheses belongs to $\ker(\mathbf{E}_u^x)$. Use the line integral formula (3.8) or Procedure A to obtain ϕ, then set $f^i := \phi$ for $x^i = x$. Now update: set

$$\mathcal{C} := \mathcal{C} - D_i f^i, \qquad F_{n+1}^i := F_n^i + f^i.$$

4.2 Ranking and using the variables

Having described a single iteration, we now turn to the question of how to choose x and u effectively. The starting point is to construct a *derivative-dominant* ranking of the variables $u_{\mathbf{J}}^\alpha$. This is a positive ranking that is determined by:

- a ranking of the independent variables, $x^1 \prec x^2 \prec \cdots \prec x^p$;
- a ranking of the dependent variables, $u^1 \prec u^2 \prec \cdots \prec u^q$.

(Later in this section, we give a heuristic for ranking the dependent and independent variables effectively.) The derivative-dominant ranking (denoted $\mathbf{u}_p$) is constructed iteratively, as follows.

$$\mathbf{u}_0 = u^1 \prec \cdots \prec u^q,$$

$$\mathbf{u}_1 = \mathbf{u}_0 \prec D_1 \mathbf{u}_0 \prec D_1^2 \mathbf{u}_0 \prec \cdots,$$

$$\mathbf{u}_2 = \mathbf{u}_1 \prec D_2 \mathbf{u}_1 \prec D_2^2 \mathbf{u}_1 \prec \cdots,$$

$$\vdots \qquad \vdots$$

$$\mathbf{u}_p = \mathbf{u}_{p-1} \prec D_p \mathbf{u}_{p-1} \prec D_p^2 \mathbf{u}_{p-1} \prec \cdots.$$

In practice, very few $u_{\mathbf{J}}^\alpha$ are needed to carry out many inversions of interest, but it is essential that these are used in the order given by their ranking, subject to a constraint on $|\mathbf{I}|$ that is explained below.

Given an independent variable, x, we call $u_{\mathbf{I}}^\alpha$ *relevant* if the updated $\mathcal{C}$ depends on $u_{\mathbf{I},k}^\alpha$ for some $k \geq 0$. The first set of iterations uses $x = x^1$. For the initial iteration, u is the lowest-ranked relevant u^α. In the following iteration, u is the next-lowest-ranked relevant u^α and so on, up to and including u^q. (From (3.7), the condition (4.1) holds whenever $u = u^\alpha$, $\alpha = 1, \ldots, q$.) After these iterations, the updated $\mathcal{C}$ is independent of $\mathbf{u}$ and its unmixed x-derivatives.

If the updated $\mathcal{C}$ has any remaining x-derivatives, these are mixed. Thus, as $\mathcal{C}$ has no linear terms, a necessary condition for the inversion to be minimal is that every f_n^i is independent of $\mathbf{u}$ and its unmixed x-derivatives. Consequently, (4.1) holds for $u = u_{\mathbf{I}}^\alpha$ whenever $|\mathbf{I}| = 1$, because

$$\mathbf{E}_{u_{\mathbf{I}-\mathbf{1}_i}^\alpha}^x (f_n^i) = 0, \qquad x^i \neq x.$$

Therefore, the process can be continued using each relevant $u = u_{\mathbf{I}}^{\alpha}$ with $|\mathbf{I}| = 1$ in the ranked order. Iterating, the same argument is used with $|\mathbf{I}| = 2, 3, \ldots$, until $\mathcal{C}$ is independent of x-derivatives. Now set $x = x^2$ and iterate, treating x^1 as a parameter. In principle, this can be continued up to $x = x^p$; in practice, only a very few iterations are usually needed to complete the inversion. The best rankings invert many terms at each iteration. On the basis of some experience with conservation laws, the following heuristic for ranking the variables x^j and u^{α} is recommended.

Ranking heuristic. Apply criteria for ranking independent variables, using the following order of precedence.

1. Any independent variables that occur in the arguments of arbitrary functions of $\mathbf{x}$ should be ranked as high as possible, if they multiply terms that are nonlinear in $[\mathbf{u}]$. For instance, if nonlinear terms in a divergence depend on an arbitrary function, $f(t)$, set $x^p = t$.

2. If independent variables occur explicitly in non-arbitrary functions, they should be ranked as high as possible (subject to 1 above), with priority going to variables with the most complicated functional dependence. For instance, if $\mathcal{C}$ is linear in x^i and quadratic in x^j, then $x^i \prec x^j$ (so $i < j$ in our ordering).

3. If an unmixed derivative of any u^{α} with respect to x^i is the argument of a function other than a rational polynomial, rank x^i as low as possible.

4. Set $x^i \prec x^j$ if the highest-order *unmixed* derivative (of any u^{α}) with respect to x^i is of higher order than the highest-order unmixed derivative with respect to x^j.

5. Set $x^i \prec x^j$ if there are more occurrences of unmixed x^i-derivatives (in the fully expanded divergence) than there are of unmixed x^j-derivatives.

6. Apply criteria 3, 4, and 5 in order of precedence, replacing unmixed by 'minimally mixed' derivatives. Minimally mixed means that there are as few derivatives as possible with respect to any other variable(s).

The derivative indices $\mathbf{J}$ in a derivative-dominant ranking are ordered according to the ranking of the independent variables. This can be used to rank the dependent variables; if there is more than one dependent variable in $\mathcal{C}$, use the following criteria in order.

1. Let $u^{\alpha} \prec u^{\beta}$ if $\mathcal{C}$ is linear in $[u^{\alpha}]$ and nonlinear in $[u^{\beta}]$.

2. Let $u^{\alpha} \prec u^{\beta}$ if the lowest-ranked derivative of u^{α} that occurs in $\mathcal{C}$ is ranked lower than the lowest-ranked derivative of u^{β} in $\mathcal{C}$. (In conservation laws, the lowest-ranked derivative of u^{α} is commonly the undifferentiated u^{α}, which corresponds to $\mathbf{J} = 0$.)

3. Let $u^{\alpha} \prec u^{\beta}$ if the lowest-ranked derivative of u^{α} in the fully expanded form of $\mathcal{C}$ occurs in more terms than the corresponding derivative of u^{β} does.

These two sets of criteria are not exhaustive (allowing ties, which must be broken), but the aim that underlies them is to carry as few terms as possible into successive iterations of the procedure. Partial Euler operators with respect to unmixed derivatives are used in the earliest iterations; commonly, these are sufficient to complete the inversion.

4.3 *How to deal with poor scalings*

To remove an earlier restriction on the inversion process, we now address the problem of poor scalings, namely, that $u_{\mathbf{I}}^{\alpha}$ and its x-derivatives (denoted $[u_{\mathbf{I}}^{\alpha}]_x$) may occur in terms that belong to $\mathcal{C} - \pi_{u_{\mathbf{I}}^{\alpha}}^x(\mathcal{C})$. Such terms (when fully expanded) are products of homogeneous rational polynomials in $[u_{\mathbf{I}}^{\alpha}]_x$ of degree zero and logarithms of a single element of $[u_{\mathbf{I}}^{\alpha}]_x$. We refer to these collectively as *zero-degree* terms.

To overcome this difficulty, we modify the approach used by Anco & Bluman [3] to treat singularities. In our context, $[u_{\mathbf{I}}^{\alpha}]_x$ is replaced by $[u_{\mathbf{I}}^{\alpha} + U_{\mathbf{I}}^{\alpha}]_x$, where $U_{\mathbf{I}}^{\alpha}$ is regarded as a new dependent variable that

is ranked higher than $u_{\mathbf{I}}^{\alpha}$. This approach works equally well for logarithms, ensuring that $\pi_{u_{\mathbf{I}}^{\alpha}}^{x}$ is a good scaling for all terms that depend on $[u_{\mathbf{I}}^{\alpha}]_x$, so that its kernel consists only of terms that are independent of these variables. At the end of the calculation, all members of $[U_{\mathbf{I}}^{\alpha}]_x$ are set to zero. Note that there is no need to replace $u_{\mathbf{I}}^{\alpha}$ in terms that are not zero-degree in $[u_{\mathbf{I}}^{\alpha}]_x$, as total differentiation preserves the degree of homogeneity.[2]

Example 4.2. *To illustrate the inversion process for divergences that have zero-degree terms, we complete Example* 3.4 *by inverting*

$$\mathcal{C} = D_x F + D_y G = u_x(2u + v_y) - v_x(u_y + 2v_{yy}) + \frac{u_x}{u^2} + \frac{v_{yy}}{v_y} + \frac{2u_y}{u}\ln|u|;$$

this has no linear terms. The ranking heuristic gives $y \prec x$ *and* $u \prec v$. *The rest of the calculation goes as follows.*

1. $\mathcal{C} - \pi_u^y(\mathcal{C})$ *has just one zero-degree term (in* $[u]_y$*), namely* $(2u_y/u)\ln|u|$. *Replace this term by* $(2(u_y + U_y)/(u + U))\ln|u + U|$.

2. *Calculate* $\mathbf{E}_u^y(\mathcal{C}) = 2u_x + v_{xy} - 2u_x u^{-3} = D_x\{2u + v_y + u^{-2}\} = D_x\{\mathbf{E}_u^y(u^2 + uv_y - u^{-1})\}$. *No term in* $2u + v_y + u^{-2}$ *is ranked lower than* u, *as* u *is the lowest-ranked variable.*

3. *Then* $\pi_u^y(\mathcal{C} - D_x\{u^2 + uv_y - u^{-1}\}) = D_y\{-uv_x + (\ln|u + U|)^2\}$.

4. *Now set* $U = 0$ *to yield the remainder* $\mathcal{C}_1 = \mathcal{C} - D_x\{u^2 + uv_y - u^{-1}\} - D_y\{-uv_x + (\ln|u|)^2\}$ *at the close of the first iteration. This amounts to* $\mathcal{C}_1 = -2v_x v_{yy} + v_{yy}/v_y$.

5. *The second iteration starts with* $\mathcal{C}_1 - \pi_v^y(\mathcal{C}_1) = v_{yy}/v_y$; *as this term is zero-degree in* $[v]_y$, *replace it by* $(v_{yy} + V_{yy})/(v_y + V_y)$.

6. *Calculate* $\mathbf{E}_v^y(\mathcal{C}_1) = -2v_{xyy} = D_x\{-2v_{yy}\} = D_x\{E_v^y(-vv_{yy})\} = D_x\{E_v^y(v_y^2)\}$. *Note that* $-2v_{yy}$ *is not ranked lower than* v.

7. *Then* $\pi_v^y(\mathcal{C}_1 - D_x\{v_y^2\}) = D_y\{-2v_x v_y + \ln|v_y + V_y|\}$.

8. *Now set* $V = 0$ *to yield the remainder* $\mathcal{C}_2 = \mathcal{C}_1 - D_x\{v_y^2\} - D_y\{-2v_x v_y + \ln|v_y|\} = 0$ *at the close of the second iteration. The inversion process stops, having yielded the output*

$$F = u^2 + uv_y - u^{-1} + v_y^2, \qquad G = -uv_x + (\ln|u|)^2 - 2v_x v_y + \ln|v_y|.$$

Note that the homotopy formula (3.18) for inverting $D_x F$ can be adjusted in the same way, whenever $\sigma_{u_{\mathbf{I}}^{\alpha}}^{x}$ is a poor scaling for $E_{u_{\mathbf{I},j+1}^{\alpha}}^{x}(D_x F)$.

4.4 Linear divergences

The inversion process runs into a difficulty when a given divergence has terms that are linear in $[\mathbf{u}]$, with mixed derivatives. Then it is possible to invert in more than one way, some of which may not produce a minimal result. To address this, it is helpful to invert using a different process. Suppose that $\mathcal{C}$ is a linear divergence (in fully expanded form). Instead of using the derivative-dominant ranking, integrate by parts, working down the total order $|\mathbf{J}|$ of the derivatives $u_{\mathbf{J}}^{\alpha}$. For a given total order, we will invert the mixed derivatives first, though this is not essential.

Starting with the highest-order derivatives, one could integrate each term $f(\mathbf{x})u_{\mathbf{J}}^{\alpha}$ in $\mathcal{C}$ by parts with respect to any x^i such that $j^i \geq 1$, yielding the remainder $-D_i(f)u_{\mathbf{J}-1_i}^{\alpha}$. If $D_{\mathbf{J}}$ is a mixed derivative, we seek to choose D_i in a way that keeps the result concise. Here are some simple criteria that are commonly effective, listed in order of precedence.

---

[2] An alternative approach (see Hickman [7]) uses a locally invertible change of variables, $u^{\alpha} = \exp v^{\alpha}$, to change the degree of homogeneity from zero. This approach works equally well, but requires a little more processing time.

1. f is independent of x^i.

2. $\mathcal{C}$ includes the term $D_i(f)u^\alpha_{\mathbf{J}-1_i}$.

3. f is linear in x^i.

These criteria can be used as an initial pass to invert $\mathcal{C}$ at least partially, leaving a remainder to be inverted that may have far fewer terms than $\mathcal{C}$ does.

Integrating the remainder by parts is straightforward if each $u^\alpha_{\mathbf{J}}$ is an unmixed derivative. If $\mathbf{J}$ denotes a mixed derivative, integrate with respect to each x^i such that $j^i \geq 1$ in turn, multiplying each result by a parameter (with the parameters summing to 1). Iterate until either the remainder has a factor that is zero for some choice of parameters or there is no remainder. The final stage is to choose the parameters so as to minimize the number of terms in the final expression. Although this produces a minimal inversion, it comes at the cost of extra computational time spent doing all possible inversions followed by the parameter optimization.

Example 4.3. *Consider the linear divergence*

$$\mathcal{C} = \left(\frac{1}{6}f'(t)y^3 + f(t)xy \right)(u_{xt} - u_{yy}),$$

where f is an arbitrary function. The first of the simple criteria above yields

$$\mathcal{C} = D_x\left\{ \frac{1}{6}f'y^3 u_t \right\} + fxyu_{xt} - \left(\frac{1}{6}f'y^3 + fxy \right)u_{yy};$$

The second criterion is not helpful at this stage, but the third criterion gives

$$\mathcal{C} = D_x\left\{ \left(\frac{1}{6}f'y^3 + fxy \right)u_t \right\} - fyu_t - \left(\frac{1}{6}f'y^3 + fxy \right)u_{yy}.$$

The remainder has no mixed derivatives; integrating it by parts produces the minimal inversion

$$\mathcal{C} = D_x\left\{ \left(\frac{1}{6}f'y^3 + fxy \right)u_t \right\} + D_y\left\{ \left(\frac{1}{2}f'y^2 + fx \right)u - \left(\frac{1}{6}f'y^3 + fxy \right)u_y \right\} + D_t\{-fyu\}. \tag{4.6}$$

Example 4.4. *To illustrate the parametric approach, consider*

$$\mathcal{C} = \exp(t - x^2)(tu_{xtt} + 2x(t + 1)u_t).$$

The simple criteria are irrelevant at present, so instead introduce parameters λ_l and consider all possible inversions of the mixed derivative terms. Step-by-step, one obtains the following.

$$\mathcal{C} = D_x\{\lambda_1 \exp(t - x^2)tu_{tt}\} + D_t\{(1 - \lambda_1)\exp(t - x^2)tu_{xt}\}$$
$$+ \exp(t - x^2)\{2\lambda_1 xtu_{tt} - (1 - \lambda_1)(t + 1)u_{xt} + 2x(t + 1)u_t\}$$
$$= D_x\{\exp(t - x^2)\{\lambda_1 tu_{tt} - \lambda_2(1 - \lambda_1)(t + 1)u_t\}\}$$
$$+ D_t\{\exp(t - x^2)\{(1 - \lambda_1)tu_{xt} + 2\lambda_1 xtu_t - (1 - \lambda_2)(1 - \lambda_1)(t + 1)u_x\}\}$$
$$+ (1 - \lambda_2)(1 - \lambda_1)\exp(t - x^2)\{2x(t + 1)u_t + (t + 2)u_x\}.$$

As the remainder has a factor $(1 - \lambda_2)(1 - \lambda_1)$, the inversion is complete if either parameter is set to 1. The minimal (two-term) inversion has $\lambda_1 = 1$, which gives $\mathcal{C} = D_x F + D_t G$, where

$$F = t \exp(t - x^2)u_{tt}, \qquad G = 2xt \exp(t - x^2)u_t.$$

For $\lambda_1 \neq 1, \lambda_2 = 1$, the inversion has three terms if $\lambda_1 = 0$, or five terms otherwise.

Note the importance of factorizing the remainder to stop the calculation once a possible parameter choice occurs. If we had continued the calculation without setting either λ_i to 1, it would have stopped at order zero, not one, giving an eleven-term inversion for general λ_i.

To summarize, one can invert a divergence that is linear in [$\mathbf{u}$] by applying the following procedure.

Procedure B. Iinversion of a linear total divergence

Step 0. Identify the maximum derivative order, $N = |\mathbf{J}|$, of the variables $u_{\mathbf{J}}^{\alpha}$ that occur in $\mathcal{C}$. Set $F^i := 0$, $i = 1 \dots, p$.

Step 1. While $\mathcal{C}$ has at least one term of order N, do the following. Select any such term, $f(\mathbf{x})u_{\mathbf{J}}^{\alpha}$, and determine a variable x^i over which to integrate. If desired, parametrize for mixed derivatives, as described above. Set $F^i := F^i + fu_{\mathbf{J}-1_i}^{\alpha}$ for the chosen i (with the appropriate modification if parameters are used), and update the remainder by setting $\mathcal{C} := \mathcal{C} - f(\mathbf{x})u_{\mathbf{J}}^{\alpha} - D_i(f)u_{\mathbf{J}-1_i}^{\alpha}$. Once $\mathcal{C}$ has no terms of order N, continue to Step 2.

Step 2. If $\mathcal{C}$ is nonzero and cannot be set to zero by any choice of parameters, set $N := N - 1$ and return to Step 1; otherwise, set $\mathcal{C}$ to zero and carry out parameter optimization (if needed), give the output F^i, $i = 1, \dots, p$, then **stop**.

4.5 Summary: a procedure for inverting Div

Having addressed potential modifications, we are now in a position to summarize the inversion process for any total divergence $\mathcal{C}$ whose fully expanded form has no terms depending on $\mathbf{x}$ only.

Procedure C. Inversion of $\mathcal{C} = D_i F^i$

Step 0. Let $\mathcal{C}_\ell$ be the linear part of $\mathcal{C}$. If $\mathcal{C}_\ell = 0$, set $\mathcal{C}_0 := \mathcal{C}$ and $F_0^i := 0$, $i = 1, \dots, p$. Otherwise, invert $\mathcal{C}_\ell$ using the technique described in Procedure B above, to obtain functions F_0^i that satisfy $\mathcal{C}_\ell = D_i F_0^i$. Set $\mathcal{C}_0 := \mathcal{C} - \mathcal{C}_\ell$. Choose a derivative-dominant ranking for $\mathcal{C}_0$ (either using the ranking heuristic or otherwise). In the notation used earlier, set $x := x^1$ and u to be the lowest-ranked relevant u^{α}; typically, $u := u^1$. Set $n := 0$; here $n + 1$ is the iteration number.

Step 1. Calculate $\mathcal{C}_n - \pi_u^x(\mathcal{C}_n)$; if this includes terms depending on $[u]_x$, replace $[u]_x$ *in these terms only* by $[u + U]_x$.

Step 2. Apply the process detailed in Section 4.1 (with $\mathcal{C}_n$ replacing $\mathcal{C}$). Provided that the ranking check is passed, this yields components f^i, which may depend on $[U]_x$. If the ranking check is failed, choose a different ranking of x^j and u^{α} for the remainder of the inversion process and return to Step 1, starting with the lowest-ranked x and (relevant) u and working upwards at each iteration.

Step 3. Replace all elements of $[U]_x$ in f^i by zero.

Step 4. Update: set $F_{n+1}^i := F_n^i + f^i$, $\mathcal{C}_{n+1} := \mathcal{C}_n - D_i f^i$ and $n := n + 1$. If $\mathcal{C}_{n+1} = 0$, output $F^i = F_{n+1}^i$ and **stop**. Otherwise, update u and x as detailed in Section 4.2 and return to Step 1.

Example 4.5. *As an example with linear terms, the Khokhlov–Zabolotskaya equation,*

$$u_{xt} - uu_{xx} - u_x^2 - u_{yy} = 0,$$

has a conservation law (see Poole & Hereman [11]) that involves an arbitrary function, $f(t)$:

$$\mathcal{C} = D_x F + D_y G + D_t H = \left(\frac{1}{6}f'y^3 + fxy\right)\left(u_{xt} - uu_{xx} - u_x^2 - u_{yy}\right).$$

The linear part $\mathcal{C}_\ell$ of $\mathcal{C}$ is inverted by Procedure B, as shown in Example 4.3. From (4.6), Step 0 in Procedure C gives

$$F_0 = \left(\frac{1}{6}f'y^3 + fxy\right)u_t, \qquad G_0 = \left(\frac{1}{2}f'y^2 + fx\right)u - \left(\frac{1}{6}f'y^3 + fxy\right)u_y, \qquad H_0 = -fyu.$$

The remainder $\mathcal{C} - \mathcal{C}_\ell$ is

$$\mathcal{C}_0 = -\left(\frac{1}{6}f'y^3 + fxy\right)\left(uu_{xx} + u_x^2\right).$$

With the derivative-dominant ranking $x \prec y \prec t$, only one iteration is needed to complete the inversion, because $\mathbf{E}_u^x(\mathcal{C}_0) = 0$. Then

$$\pi_u^x(\mathcal{C}_0) = D_x \left\{ \frac{1}{2} fyu^2 - \left(\frac{1}{6} f'y^3 + fxy \right) uu_x \right\},$$

and as $\mathcal{C}_0 = \pi_u^x(\mathcal{C}_0)$, the calculation stops after updating, giving the output

$$F = \left(\frac{1}{6} f'y^3 + fxy \right)(u_t - uu_x) + \frac{1}{2} fyu^2, \qquad G = \left(\frac{1}{2} f'y^2 + fx \right) u - \left(\frac{1}{6} f'y^3 + fxy \right) u_y, \qquad H = -fyu.$$

This corresponds to the minimal result in Poole & Hereman [11].

Example 4.6. *In every example so far, all iterations have used $u = u^\alpha$, but not $u_\mathbf{I}^\alpha$ for $|I| \neq 0$. The BBM equation from Example 2.1 illustrates the way that the process continues unhindered once the current $\mathcal{C}$ is independent of all elements of $[u^\alpha]_x$. The conservation law (2.6) is*

$$\mathcal{C} = D_x F + D_t G = (u^2 + 2u_{xt})(u_t - uu_x - u_{xxt});$$

it has no linear part or zero-degree terms. The ranking heuristic gives $x \prec t$. The first iteration using $\mathbf{E}_u^x$ is similar to what we have seen so far, giving the following updates:

$$\mathcal{C}_1 = 2u_{xt}(u_t - u_{xxt}), \qquad F_1 = -\frac{1}{4} u^4 - u^2 u_{xt}, \qquad G_1 = \frac{1}{3} u^3.$$

At this stage, $\mathcal{C}_1$ is independent of $[u]_x$ and equals $\pi_{u_t}^x(\mathcal{C}_1)$. In the second iteration, the ranking requires us to apply $\mathbf{E}_{u_t}^x$. This annihilates $\mathcal{C}_1$; consequently, $\mathcal{C}_1$ is a total derivative with respect to x. Inverting this gives the final (minimal) result,

$$F = -\frac{1}{4} u^4 - u^2 u_{xt} - u_{xt}^2 + u_t^2, \qquad G = \frac{1}{3} u^3,$$

which was obtained by inspection in Example 2.1.

The Appendix lists inversions of various other divergences; in each case, the inversion produced by Procedure C and the ranking heuristic is minimal and takes very few iterations to complete.

4.6 Splitting a divergence using discrete symmetries

A given divergence may have discrete symmetries between various terms in its fully expanded form. If the divergence has very many terms that are connected by a particular discrete symmetry group, it can be worth splitting these into disjoint divergences that are mapped to one another by the group elements. Then it is only necessary to invert one of these divergences, using the symmetries to create inversions of the others without the need for much extra computation. However, to use Procedure C, it is necessary to check that all split terms are grouped into divergences; this check is done by using the criterion (2.3).

Polynomial divergences can first be split by degree, yielding divergences that are homogeneous in [**u**]. Such splitting does not add significantly to the computation time, nor does it need to be checked using (2.3), which holds automatically. Splitting by degree can make it easy to identify terms that are linked by discrete symmetries, as illustrated by the following example.

Example 4.7. *In Cartesian coordinates (x, y), the steady non-dimensionalized von Kármán equations for a plate subject to a prescribed axisymmetric load function, $p(x^2 + y^2)$, are $\mathcal{A}_\ell = 0$, $\ell = 1, 2$, where*

$$\mathcal{A}_1 = u_{xxxx} + 2u_{xxyy} + u_{yyyy} - u_{xx}v_{yy} + 2u_{xy}v_{xy} - u_{yy}v_{xx} - p,$$

$$\mathcal{A}_2 = v_{xxxx} + 2v_{xxyy} + v_{yyyy} + u_{xx}u_{yy} - u_{xy}^2.$$

Here u is the displacement and v is the Airy stress. This is a system of Euler–Lagrange equations. By Noether's Theorem, the one-parameter Lie group of rotational symmetries yields the following

conservation law (see Djondjorov and Vassilev [6]*):*

$$\mathcal{C} = (yv_x - xv_y)\mathcal{A}_2 - (yu_x - xu_y)\mathcal{A}_1.$$

This conservation law has linear, quadratic and cubic terms, so $\mathcal{C}$ *can be inverted by summing the inversions of each of the following divergences:*

$$\mathcal{C}_\ell = (yu_x - xu_y)p,$$

$$\mathcal{C}_q = (yv_x - xv_y)(v_{xxxx} + 2v_{xxyy} + v_{yyyy}) - (yu_x - xu_y)(u_{xxxx} + 2u_{xxyy} + u_{yyyy}),$$

$$\mathcal{C}_c = (yv_x - xv_y)(u_{xx}u_{yy} - u_{xy}^2) + (yu_x - xu_y)(u_{xx}v_{yy} - 2u_{xy}v_{xy} + u_{yy}v_{xx}).$$

The quadratic terms have an obvious discrete symmetry, $\Gamma_1 : (x, y, u, v) \mapsto (-x, y, v, u)$*, which gives a splitting into two parts, each of which is a divergence:*

$$\mathcal{C}_q = \overline{\mathcal{C}}_q + \Gamma_1(\overline{\mathcal{C}}_q),$$

where

$$\overline{\mathcal{C}}_q = (yv_x - xv_y)(v_{xxxx} + 2v_{xxyy} + v_{yyyy}).$$

Consequently, we can invert $\mathcal{C}_q$ *by inverting* $\overline{\mathcal{C}}_q$ *and applying the symmetry* Γ_1*.*

Note that $\overline{\mathcal{C}}_q$ *has a discrete symmetry,* $\Gamma_2 : (x, y, v) \mapsto (y, -x, v)$*, which gives*

$$\overline{\mathcal{C}}_q = g + \Gamma_2(g), \quad \text{where} \quad g = (yv_x - xv_y)(v_{xxxx} + v_{xxyy}).$$

Checking (2.3) *shows that this is not a valid splitting into divergences, because*

$$\mathbf{E}_v(g) = -2v_{xxxy} - 2v_{xyyy} \neq 0.$$

If necessary, split divergences can be inverted using different rankings. However, in this simple example, a single ranking works for all (nonlinear) parts. One tie-break is needed: let $x \prec y$*. The ranking heuristic gives* $v \prec u$ *for the cubic terms; the variables* [u] *are not relevant in the inversion of* $\overline{\mathcal{C}}_q$*. Procedure* C *gives the following minimal inversions,*

$$\mathcal{C}_\ell = D_x(yup) + D_y(-xup),$$

$$\overline{\mathcal{C}}_q = D_x\{F_q(x, y, [v])\} + D_y\{G_q(x, y, [v])\},$$

$$\mathcal{C}_c = D_x F_c + D_y G_c,$$

where

$$F_q(x, y, [v]) = y\left\{v_x v_{xxx} - \frac{1}{2}v_{xx}^2 + 2v_x v_{xyy} + v_{xy}^2 + vv_{yyyy} - \frac{1}{2}v_{yy}^2\right\}$$

$$+ x\left\{-v_y v_{xxx} + v_{xx}v_{xy} - 2v_y v_{xyy}\right\} + v_y v_{xx},$$

$$G_q(x, y, [v]) = y\left\{-2v_{xx}v_{xy} - vv_{xyyy} + v_y v_{xyy}\right\} + x\left\{-\frac{1}{2}v_{xx}^2 + v_{xy}^2 - v_y v_{yyy} + \frac{1}{2}v_{yy}^2\right\} + vv_{xyy},$$

$$F_c = y\left(u_x u_{yy}v_x + \frac{1}{2}u_x^2 v_{yy}\right) + x\left(u_y u_{xy}v_y + \frac{1}{2}u_y^2 v_{xy}\right) - \frac{1}{2}u_y^2 v_y,$$

$$G_c = -y\left(u_x u_{xy}v_x + \frac{1}{2}u_x^2 v_{xy}\right) - x\left(u_y u_{xx}v_y + \frac{1}{2}u_y^2 v_{xx}\right) + \frac{1}{2}u_x^2 v_x.$$

Applying Γ_1 *to* $\overline{\mathcal{C}}_q$ *gives the following minimal inversion for* $\mathcal{C}$*:*

$$\mathcal{C} = D_x\{yup + F_q(x, y, [v]) - F_q(x, y, [u]) + F_c\} + D_y\{-xup + G_q(x, y, [v]) - G_q(x, y, [u]) + G_c\}.$$

Note that the tie-break $x \prec y$ *causes the inversion to break the symmetry* Γ_2 *that is apparent in* $\overline{\mathcal{C}}_q$*. As it turns out, this symmetry can be restored without increasing the overall number of terms, by adding*

components of a trivial conservation law (which are of the form (2.1)). It is an open question whether such preservation of symmetry and minimality is achievable in general.

In Djondjorov & Vassilev [6], the inversion of $\mathcal{C}$ for the special case $p = 0$ has 62 terms, which are grouped according to their physical meaning. By contrast, the minimal inversion above has just 46 (resp. 48) terms when p is zero (resp. nonzero), a considerable saving. Moreover, by exploiting the symmetry Γ_1, only 28 (resp. 30) of these terms are determined using Procedure C. However, in seeking an efficient inversion, we have ignored the physics. It would be interesting to understand the extent to which the use of a minimal inversion obscures the underlying physics.

5. Concluding remarks

Partial Euler operators and partial scalings make it possible to invert divergences with respect to one independent variable at a time, a byproduct being that some contributions to other components are determined at each iteration step. Although each iteration involves a fair amount of computation, very few iterations are needed for many systems of interest.

Given the potential complexity of functions, it is unlikely that every divergence can be inverted, even in principle, by Procedure C. The question of how to prove or disprove this is open. In practice, products of mixed derivatives present the greatest challenge to concise inversion, although the option of re-ranking part-way through the procedure enables a divide-and-conquer approach to be taken.

The focus of this work has been on inverting the total divergence operator Div. However, this immediately applies to expressions that can be recast as total divergences. For instance, for $p = 3$, the total curl $\mathbf{F} = \mathrm{Curl}(\mathbf{G})$ can be inverted by writing

$$F^i = \mathrm{Div}(H^{ij}\mathbf{e}_j) = D_j H^{ij}, \qquad H^{ij} := \epsilon^{ijk} G_k = -H^{ji}, \tag{5.1}$$

where ϵ^{ijk} is the Levi–Civita symbol, then inverting one component at a time and using the results at each stage to simplify the remainder of the calculation. Once H^{ij} is known, the identity $G_l = \frac{1}{2}\epsilon_{ijl}H^{ij}$ recovers $\mathbf{G}$. Typically, a minimal inversion is achieved by using a different ranking for each F^i, in accordance with the ranking heuristic. Here is a simple illustration of the general approach.

Example 5.1. *In Cartesian coordinates, invert $(F^x, F^y, F^z) = \mathrm{Curl}(G_x, G_y, G_z)$, where*

$$F^x = u_x(u_{yy} - u_{zz}) + u_y u_{xy} - u_z u_{xz} ,$$

$$F^y = u_y(u_{zz} - u_{xx}) + u_z u_{yz} - u_x u_{xy} ,$$

$$F^z = u_z(u_{xx} - u_{yy}) + u_x u_{xz} - u_y u_{yz} .$$

Begin by using Procedure C with $y \prec z \prec x$. In two iterations, this gives

$$F^x = D_y(u_x u_y) + D_z(-u_x u_z).$$

With (x, y, z) replacing the indices $(1, 2, 3)$ in (5.1), let

$$H^{xy} = u_x u_y = -H^{yx}, \qquad H^{xz} = u_x u_z = -H^{zx}.$$

Therefore

$$D_z H^{yz} = F^y - D_x H^{yx} = u_y u_{zz} + u_z u_{yz} .$$

One could invert this by a further iteration with the ranking $z \prec y$, though it is inverted more quickly by the line integral formula, which gives $H^{yz} = u_y u_z$. Finally,

$$G_x = u_y u_z, \qquad G_y = u_x u_z, \qquad G_z = u_x u_y.$$

hospitality during the program *Geometry, Compatibility and Structure Preservation in Computational Differential Equations*; this provided the opportunity to develop much of the theory described above. I also thank the referees for their constructive comments.

Conflicts of interest. None.

References

[1] Anco, S. C. (2017) Generalization of Noether's theorem in modern form to non-variational partial differential equations. In: Melnik, R., Makarov, R. & Belair, J. (editors), *Recent Progress and Modern Challenges in Applied Mathematics, Modeling and Computational Science, Fields Institute Communications*, Vol. **79**, pp. 119–182.

[2] Anco, S. C. & Bluman, G. (2002) Direct construction method for conservation laws of partial differential equations I: examples of conservation law classifications. *Eur. J. Appl. Math.* **13**(5), 545–566.

[3] Anco, S. C. & Bluman, G. (2002) Direct construction method for conservation laws of partial differential equations II: general treatment. *Eur. J. Appl. Math.* **13**(5), 567–585.

[4] Brunelli, J. C. (2006) The bi-Hamiltonian structure of the short pulse equation. *Phys. Lett. A* **353**(6), 475–478.

[5] Deconinck, B. & Nivala, M. (2011) Symbolic integration using homotopy methods. *Math. Comput. Simul.* **80**(4), 825–836.

[6] Djondjorov, P. A. & Vassilev, V. M. (2000). Acceleration waves in von Kármán plate theory. In: Bertram, B. S., Constanda, C. & Struthers, A. A. (editors), *Integral Methods in Science and Engineering*, Chapman and Hall/CRC, New York, pp. 131–136.

[7] Hickman, M. (2009). Inverting the divergence operator. *53rd Annual Meeting of the Australian Mathematical Society*, 28 September–1 October 2009, Adelaide, Adelaide.

[8] Olver, P. J. (1993). *Applications of Lie Groups to Differential Equations*, Springer-Verlag, New York. 2nd ed.

[9] Poole, D. & Hereman, W. (2010) The homotopy operator method for symbolic integration by parts and inversion of divergences with applications. *Appl. Anal.* **89**(4), 433–455.

[10] Poole, D. & Hereman, W. (2011) Symbolic computation of conservation laws for nonlinear partial differential equations in multiple space dimensions. *J. Symb. Comput.* **46**(12), 1355–1377.

[11] Popovych, R. O. & Bihlo, A. (2020) Inverse problem on conservation laws. *Physica D* **401,** 132175.

[12] Rust, C. J., Reid, G. J. & Wittkopf, A. D. (1999). Existence and uniqueness theorems for formal power series solutions of analytic differential systems. In: *ISSAC'99: Proceedings of the 1999 International Symposium on Symbolic and Algebraic Computation*, Association for Computing Machinery, New York, pp. 105–112.

[13] Wolf, T. (2002) A comparison of four approaches to the calculation of conservation laws. *Eur. J. Appl. Math.* **13**(2), 129–152.

[14] Wolf, T., Brand, A. & Mohammadzadeh, M. (1999) Computer algebra algorithms and procedures for the computation of conservation laws and fixing of gauge in differential expressions. *J. Symb. Comput.* **27**(2), 221–238.

A Appendix

Procedure C has been tested on the following conservation laws. In every case, the procedure coupled with the ranking heuristic yields a minimal inversion.

Kadomtsev–Petviashvili (KP) equation. There are two forms of the KP equation, depending on which of $\epsilon = \pm 1$ is chosen. In either case, it has a conservation law,

$$\mathcal{C} = f(t)y \left(u_{xt} + uu_{xx} + u_x^2 + u_{4x} + \epsilon u_{yy} \right)$$
$$= D_x\{fy \left(u_t + uu_x + u_{xxx} \right)\} + D_t\{\epsilon f(yu_u - u)\}.$$

With the ranking $x \prec y \prec t$, the procedure requires three iterations to obtain this inversion.

Potential Burgers equation (see Wolf et al. [14]). The potential Burgers equation has a conservation law for each $f(x,t)$ such that $f_t + f_{xx} = 0$:

$$\mathcal{C} = f \exp(u/2)(u_t - u_{xx} - \frac{1}{2}u^2) = D_x\{\exp(u/2)(2f_x - fu_x)\} + D_t\{2\exp(u/2)f\}.$$

The inversion requires one iteration with $x \prec t$, exchanging f_t and $-f_{xx}$ twice.

Zakharov–Kuznetsov equation (see Poole & Hereman [11]). The Zakharov–Kuznetsov equation is $\mathcal{A} = 0$, where

$$\mathcal{A} = u_t + uu_x + u_{xxx} + u_{xyy}.$$

It has a conservation law

$$\mathcal{C} = (u^2 + 2(u_{xx} + u_{yy}))\mathcal{A}$$

$$= D_x\{(\tfrac{1}{2}u^2 + u_{xx} + u_{yy})^2 + 2u_xu_t\} + D_y\{2u_yu_t\} + D_t\{\tfrac{1}{3}u^3 - u_x^2 - u_y^2\}.$$

The fully expanded form has just eleven terms, and the inversion requires two iterations.

Short-pulse equation (see Brunelli [4]). The short-pulse equation, $\mathcal{A} = 0$, with

$$\mathcal{A} = u_{xt} - u - uu_x^2 - \tfrac{1}{2}u^2u_{xx},$$

has the following conservation law involving a square root:

$$\mathcal{C} = u_x(1 + u_x^2)^{-1/2}\mathcal{A}$$

$$= D_x\{-\tfrac{1}{2}u^2(1 + u_x^2)^{1/2}\} + D_t\{(1 + u_x^2)^{1/2}\}.$$

With $x \prec t$, this can be inverted in one iteration.

Nonlinear Schrödinger equation. Splitting the field into its real and imaginary parts gives the system $\mathcal{A}_1 = 0, \ \mathcal{A}_2 = 0$, with

$$\mathcal{A}_1 = -v_t + u_{xx} + (u^2 + v^2)u, \qquad \mathcal{A}_2 = u_t + v_{xx} + (u^2 + v^2)v.$$

One of the conservation laws is

$$\mathcal{C} = u_t\mathcal{A}_1 + v_t\mathcal{A}_2$$

$$= D_x\{2u_xu_t + 2v_xv_t\} + D_t\left\{\tfrac{1}{2}\left(u^2 + v^2\right) - u_x^2 - v_x^2\right\}.$$

With the ranking $u \prec v$ and $x \prec t$, the procedure requires two iterations.

Itô equations (see Wolf [13]). The equations are $\mathcal{A}_1 = 0, \ \mathcal{A}_2 = 0$, with

$$\mathcal{A}_1 = u_t - u_{xxx} - 6uu_x - 2vv_x, \qquad \mathcal{A}_2 = v_t - 2u_xv - 2uv_x.$$

This system has a rational conservation law,

$$\mathcal{C} = 2v^{-1}\mathcal{A}_1 + v^{-4}\left(vv_{xx} - \tfrac{3}{2}v_x^2 - 2uv^2\right)\mathcal{A}_2$$

$$= D_x\left\{v^{-3}\left(v_xv_t - 2u_{xx}v^2 - 2u_xvv_x - uv_x^2 - 4u^2v^2 - 4v^4\right)\right\} + D_t\left\{v^{-3}\left(2uv^2 - \tfrac{1}{2}v_x^2\right)\right\}.$$

With the ranking $u \prec v$ and $x \prec t$, the procedure requires two iterations.

Navier–Stokes equations. The (constant-density) two-dimensional Navier–Stokes equations are $\mathcal{A}_\ell = 0, \ \ell = 1, 2, 3$, where

$$\mathcal{A}_1 = u_t + uu_x + vu_y + p_x - v(u_{xx} + u_{yy}), \quad \mathcal{A}_2 = v_t + uv_x + vv_y + p_y - v(v_{xx} + v_{yy}), \quad \mathcal{A}_3 = u_x + v_y.$$

This system has a family of conservation laws involving two arbitrary functions, $f(t)$ and $g(t)$, namely

$$\mathcal{C} = f\mathcal{A}_1 + g\mathcal{A}_2 + (fu + gv - f'x - g'y)\mathcal{A}_3$$

$$= D_x\{fu^2 + guv + fp - v(fu_x + gv_x) - (f'x + g'y)u\}$$

$$+ D_y\{fuv + gv^2 + gp - v(fu_y + gv_y) - (f'x + g'y)v\} + D_t\{fu + gv\}.$$

Procedure C consists of a linear inversion and two further iterations, using the ranking $x \prec y \prec t$. The three-dimensional Navier–Stokes equations have a similar conservation law, which requires a linear inversion and three further iterations.

Procedure C and the ranking heuristic have also been tested on some divergences not arising from conservation laws, with high order or complexity. Again, the output in each case is a minimal inversion.

High-order derivatives.   The divergence is

$$\mathcal{C} = u_{5x}u_{4y} + u_{2x,2y}u_{2x,3y}$$

$$= D_x\left\{u_{4x}u_{4y} - u_{3x}u_{x,4y} + (u_{2x,2y})^2\right\} + D_t\left\{u_{3x}u_{2x,3y} - u_{3x,y}u_{2x,2y}\right\}.$$

With the ranking $x \prec y$, the procedure requires three iterations. The other ranking, $y \prec x$, also produces a minimal inversion after three iterations; it is higher order in x but lower order in y:

$$\mathcal{C} = D_x\left\{u_{4x,2y}u_{2y} - u_{3x,2y}u_{x,2y} + (u_{2x,2y})^2\right\} + D_t\left\{u_{5x}u_{3y} - u_{5x,y}u_{2y}\right\}.$$

High-order derivatives and explicit dependence.   The divergence is

$$\mathcal{C} = t(u_y u_{xttt} - u_x u_{yttt})$$

$$= D_x\left\{-tuu_{yttt}\right\} + D_y\{tuu_{xttt})\}.$$

With the ranking $x \prec y \prec t$, the procedure requires one iteration. This illustrates the value of the second criterion for ranking independent variables; if t is ranked lower than x and y, the procedure fails at the first check.

Exponential dependence.   The divergence is

$$\mathcal{C} = \left(u_{xx}u_y^2 - 2u_{yy}\right)\exp(u_x)$$

$$= D_x\left\{u_y^2 \exp(u_x)\right\} + D_t\left\{-2u_y \exp(u_x)\right\}.$$

With the ranking $x \prec y$, the procedure requires one iteration.

Trigonometric dependence.   The divergence is

$$\mathcal{C} = \left(u_{xx}u_{yyy} - u_{xxy}u_{xyy}\right)\cos(u_x)$$

$$= D_x\left\{u_{yyy}\sin(u_x)\right\} + D_t\left\{-u_{xyy}\sin(u_x)\right\}.$$

With the ranking $x \prec y$, the procedure requires one iteration.