Calculus of Variations

March 28, 2018

1 Multi-indices

A multi-index α is just an m-tuple of non-negative integers $(\alpha_1, \ldots, \alpha_m)$ for some positive integer m. The degree of α is defined to be

$$|\alpha| = \sum_{i=1}^{m} \alpha_i.$$

The number of multi-indices of degree at most k is $\frac{(m+k)!}{m!k!}$. Addition of multi-indices is defined componentwise. In other words

$$\alpha + \beta = \gamma$$

means

$$\alpha_1 + \beta_1 = \gamma_1, \quad \ldots, \quad \alpha_m + \beta_m = \gamma_m.$$

Multi-indices are useful for labeling repeated derivatives. We write

$$\left(\frac{\partial}{\partial x}\right)^{\alpha}$$

or

$$\frac{\partial^{|\alpha|}}{\partial x^{\alpha}}$$

for

$$\left(\frac{\partial}{\partial x_1}\right)^{\alpha_1} \cdots \left(\frac{\partial}{\partial x_m}\right)^{\alpha_m}.$$

If u is smooth then

$$\left(\frac{\partial}{\partial x}\right)^{\alpha}\left(\frac{\partial}{\partial x}\right)^{\beta}=\left(\frac{\partial}{\partial x}\right)^{\alpha+\beta}.$$

We define $\hat{\imath}$ to be the multi-index whose i'th entry is 1 and all others are 0. In other words, the one with

$$\left(\frac{\partial}{\partial x}\right)^{\hat{\imath}}=\frac{\partial}{\partial x_i}.$$

2 Prolongation of Functions

For the general theory we will use

$$\mathcal{R}_{m,n,k}=\mathbf{R}^{m+n\frac{(m+k)!}{m!k!}}$$

for the space of k'th order jets of functions from $\mathbf{R}^m$ to $\mathbf{R}^n$ and x_i, $u_{j,\alpha}$ for coordinates on $\mathcal{R}_{m,n,k}$, where $1\leq i\leq m$, $1\leq j\leq n$ and $|\alpha|\leq k$. If $u\colon\mathbf{R}^m$ to $\mathbf{R}^n$ is smooth then $\operatorname{pr}^{(k)}u$ is the function from $\mathbf{R}^m$ to $\mathbf{R}^{n\frac{(m+k)!}{m!k!}}$ whose graph is the subset of $\mathcal{R}_{m,n,k}$ where

$$u_{j,\alpha}=\frac{\partial^{|\alpha|}u_j}{\partial x^{\alpha}}(x_1,\ldots,x_m)$$

for all $1\leq j\leq n$ and $|\alpha|\leq k$. Similar remarks apply to functions which are defined only on open subsets of $\mathbf{R}^m$. In what follows I'll ignore that and just treat the case of functions defined everywhere.

In concrete examples this leads to too many subscripts, so we use an alternate notation, where the independent variables $x_1,\ldots,x_m$ and $u_1,\ldots,u_n$ each have single letter names without subscripts. The choice of those letters depends on the applications. For example, for three dimensional incompressible fluid flow the conventional choice is t,x,y,z for the independent variables, representing time and space coordinates, and p,u,v,w for the pressure and the components of velocity. In examples we also don't use multi-indices but rather list the corresponding variables in lexicographic order. So $\mathcal{R}_{4,4,2}$ would have, with coordinates as in the example, coordinates t, x, y, z, p, u, v, w and a further 36, including p_t, u_{xy}, v_{tz}, etc. A simpler example would be

$m = 1$, $n = 1$, $k = 3$, with independent variable x and dependent variable y. The coordinates of the five dimensional space $\mathcal{R}_{1,1,3}$ would then be labeled x, y, y_x, y_{xx} and y_{xxx}. The prolongation $\operatorname{pr}^{(3)} f$ of a function f would then have graph

$$\{(x, y, y_x, y_{xx}, y_{xxx}) \in \mathcal{R}_{m,n,k} : y = f(x), y_x = f'(x), y_{xx} = f''(x), y_{xxx} = f'''(x)\}.$$

3 Total Derivatives

In addition to partial derivatives of a smooth function on $\mathcal{R}_{m,n,k}$ with respect to the various coordinates on that space we have m total derivative operators D_i, which take such a function and produce smooth functions on $\mathcal{R}_{m,n,k+1}$. In terms of partial derivatives, these are defined by

$$D_i = \frac{\partial}{\partial x_i} + \sum_{j=1}^{n} \sum_{|\alpha| \le k} u_{j,\alpha+\hat{\imath}} \frac{\partial}{\partial u_{j,\alpha}}.$$

We use multi-index notation for repeated total differentiation,

$$D^\alpha = D_1^{\alpha_1} \ldots D_m^{\alpha_m},$$

and we use variable names in place of indices in examples. The point of the total derivative operator is that

$$(D_i f)(x, \operatorname{pr}^{(k+1)} u(x)) = \frac{d}{dx_i} f(x, \operatorname{pr}^{(k)} u(x)),$$

by the chain rule.

As an example, consider $m = 1$, $n = 1$, $k = 1$, with independent variable x and dependent variable u. The coordinates on $\mathcal{R}_{1,1,1}$ are then x, u and u_x. Consider the function

$$f(x, u, u_x) = \frac{x u_x}{\sqrt{1 + u_x^2}}.$$

Its partial derivatives are

$$\frac{\partial f}{\partial x} = \frac{u_x}{\sqrt{1 + u_x^2}},$$

$$\frac{\partial f}{\partial u} = 0$$

and

$$\frac{\partial f}{\partial u_x} = \frac{x}{(1+u_x)^{3/2}}$$

so

$$(D_x f)(x,u,u_x) = \frac{u_x}{\sqrt{1+u_x^2}} + u_{xx}\frac{x}{(1+u_x)^{3/2}} = \frac{xu_{xx}+u_x^3+u_x}{(1+u_x)^{3/2}}.$$

Applied to prolongations, this expresses the fact that

$$\frac{d}{dx}\frac{xu'(x)}{\sqrt{1+u'(x)^2}} = \frac{xu''(x)+u'(x)^3+u'(x)}{(1+u'(x)^2)^{3/2}}$$

for any smooth function u.

4 Euler-Lagrange Equations

Total derivatives appear in the Euler-Lagrange equations, which are necessary but not sufficient conditions for an integral of the form

$$\mathcal{L}(u) = \int_\Omega L(x, \operatorname{pr}^{(k)} u(x))$$

to be a minimum (or maximum) among functions whose values and derivatives up to, but not including, order k are specified on $\partial\Omega$. The equations are

$$\sum_{|\alpha|\le k} (-1)^{|\alpha|} \left(D^\alpha \frac{\partial L}{\partial u_{j,\alpha}} \right) (x, \operatorname{pr}^{(2k)} u(x)) = 0$$

for $1 \le j \le n$.

This simplifies somewhat if some of m, n or k are small. If $m = n = k = 1$ we get simply

$$\frac{\partial L}{\partial u} - D_x \frac{\partial L}{\partial u_x} = 0.$$

For $m = k = 1$, $n > 0$, we get the system of n equations

$$\frac{\partial L}{\partial u_1} - D_x \frac{\partial L}{\partial u_{1,x}} = 0, \quad \ldots, \quad \frac{\partial L}{\partial u_n} - D_x \frac{\partial L}{\partial u_{n,x}} = 0.$$

For $n = k = 1$, $m > 0$, we get the single equation

$$\frac{\partial L}{\partial u} - D_{x_1}\frac{\partial L}{\partial u_{x_1}} \cdots - D_{x_m}\frac{\partial L}{\partial u_{x_m}} = 0.$$

For $m = n = 1$, $k > 0$ we get the single equation

$$\frac{\partial L}{\partial u} - D_x\frac{\partial L}{\partial u_x} + D_x^2\frac{\partial L}{\partial u_{xx}} - \cdots = 0.$$

The series is not infinite but terminates after $k+1$ terms.

As an example, consider $m = n = 1$, $k = 2$,

$$L(x, u, u_x, u_{xx}) = u_{xx} + u_x^2 + u^3.$$

The relevant derivatives are

$$\frac{\partial L}{\partial u} = 3u^2, \quad \frac{\partial L}{\partial u_x} = 2u_x, \quad \frac{\partial L}{\partial u_{xx}} = 1,$$

$$D_x\frac{\partial L}{\partial u_x} = 2u_{xx}, \quad D_x\frac{\partial L}{\partial u_{xx}} = 0, \quad D_x^2\frac{\partial L}{\partial u_{xx}} = 0.$$

Then the Euler-Lagrange equation is

$$\frac{\partial L}{\partial u} - D_x\frac{\partial L}{\partial u_x} + D_x^2\frac{\partial L}{\partial u_{xx}} = 3u^2 - 2u_{xx} = 0.$$

As the example above shows, the Euler-Lagrange equations of a k'th order Lagrangian are of order at most $2k$, but can be of lower order. The left hand side can in fact be zero, in which case we call L a null Lagrangian. It can be shown that this happens if and only if

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

where $P = (P_1, \ldots, P_m)$ are functions on $\mathcal{R}_{m,n,k-1}$ and the total divergence is defined by

$$\mathrm{Div}\, P = \sum_{i=1}^{m} D_i P_i.$$

5 Derivation of the Euler-Lagrange Equations

Here is the sketch of a proof that the Euler-Lagrange Equations are necessary for an extremum. Let $\tilde{u}$ be a smooth functions on the Cartesian product of an open interval containing 0 and $\overline{\Omega}$. Use s for the variable in the first factor of that product and x for the second. Assume that $\tilde{u} = u$ when $s = 0$. Set

$$\delta\mathcal{L} = \frac{d}{ds}\mathcal{L}(\tilde{u})\bigg|_{s=0}, \qquad \delta u_j(x) = \frac{d}{ds}\tilde{u}_j(s,x)\bigg|_{s=0}.$$

Differentiation under the integral sign and the chain rule give

$$\delta\mathcal{L} = \int_\Omega \sum_{j=1}^{n} \sum_{|\alpha|\le k} \frac{\partial L}{\partial u_{j,\alpha}}(x, \operatorname{pr}^{(k)} u(x)) D^\alpha \delta u_j(x).$$

Now for any multi-index α and any smooth functions u and v we have the identity

$$\sum_{i=1}^{m} \frac{\partial}{\partial x_i} \left(\sum_{\beta+\gamma=\alpha-\hat{\imath}} (-1)^{|\beta|} \frac{|\beta|!|\gamma|!}{|\alpha|!} \prod_{l=1}^{m} \frac{\alpha_l!}{\beta_l!\gamma_l!} \frac{\partial^{|\beta|} u}{\partial x^\beta} \frac{\partial^{|\gamma|} v}{\partial x^\gamma} \right)$$
$$= u\frac{\partial^{|\alpha|} v}{\partial x^\alpha} - (-1)^{|\alpha|} v \frac{\partial^{|\alpha|} u}{\partial x^\alpha}.$$

By the divergence theorem then,

$$\int_\Omega \left(u\frac{\partial^{|\alpha|} v}{\partial x^\alpha} - (-1)^{|\alpha|} v \frac{\partial^{|\alpha|} u}{\partial x^\alpha} \right)$$
$$= \int_{\partial\Omega} \sum_{i=1}^{m} \left(\sum_{\beta+\gamma=\alpha-\hat{\imath}} (-1)^{|\beta|} \frac{|\beta|!|\gamma|!}{|\alpha|!} \prod_{l=1}^{m} \frac{\alpha_l!}{\beta_l!\gamma_l!} \frac{\partial^{|\beta|} u}{\partial x^\beta} \frac{\partial^{|\gamma|} v}{\partial x^\gamma} \right) \nu_i,$$

where ν is the outward pointing unit normal on $\partial\Omega$. It then follows that

$$\delta\mathcal{L} = \sum_{j=1}^{n} \int_\Omega \sum_{|\alpha|\le k} (-1)^{|\alpha|} \left(D^\alpha \frac{\partial L}{\partial u_{j,\alpha}} \right) \left(x, \operatorname{pr}^{(2k)} u(x) \right) \delta u_j(x)$$
$$+ \int_{\partial\Omega} \sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{|\alpha|\le k} \sum_{\beta+\gamma=\alpha-\hat{\imath}} (-1)^{|\beta|} \frac{|\beta|!|\gamma|!}{|\alpha|!} \prod_{l=1}^{m} \frac{\alpha_l!}{\beta_l!\gamma_l!} \left(D^\beta \frac{\partial L}{\partial u_{j,\alpha}} \right) (D^\gamma \delta u_j)\, \nu_i.$$

If $\tilde{u}$ and its x derivatives of order less than k are independent of s on $\partial\Omega$ then the $D^\gamma \delta u_j$ are all zero there and the boundary integral is zero. Since the δu_j can otherwise be chosen arbitrarily it follows that

$$\delta\mathcal{L} = 0$$

implies

$$\sum_{|\alpha|\le k} (-1)^{|\alpha|} \left(D^\alpha \frac{\partial L}{\partial u_{j,\alpha}} \right) \left(x, \operatorname{pr}^{(2k)} u(x) \right) = 0$$

for all $1 \le j \le n$ and $x \in \Omega$. But $\delta\mathcal{L} = 0$ is necessary for an extremum, because otherwise $\mathcal{L}(\tilde{u})$ would fail to have an extremum at $s = 0$ by the first derivative test from single variable calculus.

6 Prolongation of Vector Fields

Suppose $\tilde{x}_1, \ldots, \tilde{x}_m$ and $\tilde{u}_1, \ldots, \tilde{u}_n$ are smooth functions of s, $x_1, \ldots, x_m$, $u_1, \ldots, u_n$, with $\tilde{x} = x$ and $\tilde{u} = u$ when $s = 0$. The situation considered in the previous section can be thought of as the special case where $\tilde{x} = x$ and $\tilde{u}$ is independent of u. We will use a different notation for s derivatives at $s = 0$ from what we used in that special case however, to avoid confusion:

$$\xi_i(x,u) = \frac{d\tilde{x}_i}{ds}(s,x,u)\bigg|_{s=0}, \quad \eta_j(x,u) = \frac{d\tilde{u}_j}{ds}(s,x,u)\bigg|_{s=0}.$$

The vector field

$$V = \sum_{i=1}^{m} \xi_i(x,u) \frac{\partial}{\partial x_i} + \sum_{j=1}^{n} \eta_j(x,u) \frac{\partial}{\partial u_j}$$

on $\mathcal{R}_{m,n,0} = \mathbf{R}^{m+n}$ is called the vector field associated to the parametrised set of transformations $(x,u) \to (\tilde{x}, \tilde{u})$. The chain rule shows that for any smooth function f on $\mathcal{R}_{m,n,0}$ we have

$$\frac{d}{ds} f(\tilde{x}(s,x,u), \tilde{u}(s,x,u))\bigg|_{s=0} = (Vf)(x,u).$$

In particular, the left hand side is zero if and only if the right hand side is.

A case of particular interest occurs when

$$\frac{d\tilde{x}_i}{ds} = \xi_i(\tilde{x}(s,x,u), \tilde{u}(s,x,u)) \quad , \frac{d\tilde{u}_j}{ds} = \eta_j(\tilde{x}(s,x,u), \tilde{u}(s,x,u))$$

for all s. In that case we say that the transformations $(x,u) \to (\tilde{x}, \tilde{u})$ form a one parameter group generated by V. Given a vector field V there may or may not be such a group. The equations above, thought of as a system of differential equations, always have a unique local solution with initial conditions

$$\tilde{x}_i(0,x,u) = x_i, \quad \tilde{u}_i(0,x,u) = u_i,$$

but this solution need not exist for all s. In case V is the generator of a one parameter group

$$\frac{d}{ds} f(\tilde{x}(s,x,u), \tilde{u}(s,x,u)) = 0 = (Vf)(\tilde{x}(s,x,u), \tilde{u}(s,x,u))$$

and hence $f(\tilde{x}(s,x,u), \tilde{u}(s,x,u))$ is independent of s if and only if $Vf = 0$.

Given a set $G \subset \mathcal{R}_{m,n,0}$ we define

$$G(s) = \{(X,U) \in \mathbf{R}^{m,n} : X = \tilde{x}(s,x,u), U = \tilde{u}(s,x,u), (x,u) \in G\}.$$

Then $G(0) = 0$. We are particularly interested in the case where G is the graph of a smooth function u from an open subset $\Omega \subset \mathbf{R}^m$ to $\mathbf{R}^n$:

$$G = \{(X,U) \in \mathbf{R}^m \times \mathbf{R}^n : U = u(X), X \in \Omega\}$$

For given $s \neq 0$ there may or may not be a function $\tilde{u}$ from an open subset $\tilde{\Omega} \subset \mathbf{R}^m$ to $\mathbf{R}^n$ such that

$$G = \{(X,U) \in \mathbf{R}^m \times \mathbf{R}^n : U = \tilde{u}(X), X \in \tilde{\Omega}\},$$

but if $\overline{\Omega}$ is compact then these exist for sufficiently small s. Both $\tilde{\Omega}$ and $\tilde{u}$ will of course depend on s. A calculation using the implicit function theorem and the chain rule then shows that

$$\frac{d}{ds} \left(\frac{d}{dx}\right)^{\alpha} \tilde{u}_j(\tilde{x}(s,x,u)) \Big|_{s=0}$$

is equal to

$$\left[\left(\frac{d}{dx}\right)^{\alpha} \left(\eta_j(x,u(x)) - \sum_{i=1}^{m} \xi_i(x,u(x)) \frac{du_j}{dx_i}(x)\right) + \sum_{i=1}^{m} \xi_i(x,u(x)) \frac{d^{|\alpha|+1} u_j}{dx^{\alpha+\hat{i}}}(x)\right].$$

We define the k'th prolongation of V by

$$\mathrm{pr}^{(k)} V = \sum_{1=1}^{m} \xi_i \frac{\partial}{\partial x_i} + \sum_{j=1}^{n} \sum_{|\alpha| \le k} \left[D^{\alpha} \left(\eta_j - \sum_{i=1}^{m} \xi_i u_{j,\hat{i}}\right) + \sum_{i=1}^{m} \xi_i u_{j,\alpha+\hat{i}}\right] \frac{\partial}{\partial u_{j,\alpha}}.$$

This is a vector field on $\mathcal{R}_{m,n,k}$. It has been constructed in such a way that if g is a function on $\mathcal{R}_{m,n,k}$ then

$$\frac{d}{ds} g(\tilde{x}(s,x,u), \operatorname{pr}^{(k)} \tilde{u}(s, \tilde{x}(s,x,u)))\bigg|_{s=0} = ((\operatorname{pr}^{(k)} V)g)(x, \operatorname{pr}^{(k)} \tilde{u}(x)).$$

Once again, the general formula simplifies if some of m, n, and k are equal to 1. If $m = n = 1$ then

$$\operatorname{pr}^{(k)} V = \xi \frac{\partial}{\partial x} + \eta \frac{\partial}{\partial u} + \sum_{l=1}^{k} \left[D^l \eta - \sum_{h=1}^{l} \frac{l!}{h!(l-h)!} D^h \xi D^{k-h+1} u \right] \frac{\partial}{\partial u_{xx\cdots x}},$$

where $u_{xx\cdots x}$ is a u with l subscripts x. If $m = k = 1$ then

$$\begin{aligned} \operatorname{pr}^{(k)} V = \xi \frac{\partial}{\partial x} &+ \eta_1 \frac{\partial}{\partial u_1} + \cdots + \eta_n \frac{\partial}{\partial u_n} \\ &+ [D_x \eta_1 - (D_x \xi) u_{1,x})] \frac{\partial}{\partial u_{1,x}} + \cdots + [D_x \eta_n - (D_x \xi) u_{n,x})] \frac{\partial}{\partial u_{n,x}}. \end{aligned}$$

If $n = k = 1$ then

$$\begin{aligned} \operatorname{pr}^{(k)} V = \xi_1 \frac{\partial}{\partial x_1} &+ \cdots + \xi_m \frac{\partial}{\partial x_m} + \eta \frac{\partial}{\partial u} \\ &+ [D_{x_1} \eta - (D_{x_1} \xi_1) u_{x_1} - \cdots - (D_{x_1} \xi_m) u_{x_m}] \frac{\partial}{\partial u_{x_1}} \\ &+ \cdots + [D_{x_m} \eta - (D_{x_m} \xi_1) u_{x_1} - \cdots - (D_{x_m} \xi_m) u_{x_m}] \frac{\partial}{\partial u_{x_m}}. \end{aligned}$$

Translations in the independent variables give a particularly simple example. For a translation in the l'th variable,

$$\tilde{x}_i = \begin{cases} x_i + s & \text{if } i = l, \\ x_i & \text{if } i \neq l, \end{cases}$$

$$\tilde{u}_j = u_j$$

and the associated vector field is just

$$V = \frac{\partial}{\partial x_l}.$$

Its k'th prolongation is the same

$$\operatorname{pr}^{(k)} V = \frac{\partial}{\partial x_l}.$$

For a more interesting example, consider $m = 3$, $n = 1$, $k = 1$, with independent variables t, x and y and dependent variable u. The generator of the rotation group

$$\tilde{t} = t, \quad \tilde{x} = x\cos\theta - y\sin\theta, \quad \tilde{y} = x\sin\theta + y\cos\theta, \quad \tilde{u} = u$$

is

$$V = \xi_t = 0, \quad \xi_x = -y, \quad \xi_y = x, \quad \eta = 0.$$

Its first prolongation is

$$\operatorname{pr}^{(1)} V = V - u_y \frac{\partial}{\partial u_x} + u_x \frac{\partial}{\partial u_y}.$$

The first prolongation applied to

$$g(t, x, y, u, u_t, u_x, u_y) = \frac{1}{2}\left(u_t^2 - u_x^2 - u_y^2\right)$$

gives zero, showing that g is rotationally invariant.

Finally note that if V is tangential to G then

$$\eta_j - \sum_{i=1}^{m} \xi_i u_{j,\hat{i}} = 0$$

for $1 \leq j \leq n$ and hence

$$\operatorname{pr}^{(k)} V = \sum_{i=1}^{m} \xi_i \left(\frac{\partial}{\partial x_i} + \sum_{j=1}^{n} \sum_{|\alpha| \leq k} u_{j,\alpha+\hat{\imath}} \frac{\partial}{\partial u_{j,\alpha}} \right) = \sum_{i=1}^{m} \xi_i D_i.$$

7 Invariant Integrals

What is

$$\delta\mathcal{L} = \left.\frac{d}{ds}\mathcal{L}(\tilde{u})\right|_{s=0} = \left.\frac{d}{ds}\left(\int_{\tilde{\Omega}} L(\tilde{x}, \operatorname{pr}^{(k)} \tilde{u}(\tilde{x})) d\tilde{x}\right)\right|_{s=0} ?$$

We can make a change of variable, picking up a Jacobian factor:

$$\tilde{\mathcal{L}} = \int_{\tilde{\Omega}} L(\tilde{x}, \operatorname{pr}^{(k)} \tilde{u}(\tilde{x})) \det(J) dx,$$

where

$$J_{il}(s,x,u) = \frac{d\tilde{x}_i}{dx_l}(s,x,u).$$

Now

$$\frac{d}{ds}\det J(s,x,u) = \operatorname{tr}\left(J(s,x,u)^{-1}\frac{d}{ds}J(s,x,u)\right)$$

But

$$\frac{d}{ds}J_{il}(s,x,u) = \frac{d^2\tilde{x}_i}{ds\,dx_l}(s,x,u) = \frac{d^2\tilde{x}_i}{dx_l\,ds}(s,x,u),$$

which is equal to $d\xi_i/dx_l$ when $s = 0$ and $J(s,x,u)$ is the identity matrix when $s = 0$, so

$$\frac{d}{ds}\det J(s,x,u)\bigg|_{s=0} \sum_{i=1}^{m} \frac{d\xi_i}{dx_i}(x,u)(\operatorname{Div} V)(x,u).$$

If we use this, the prolongation formula for derivatives of functions on $\mathcal{R}_{m,n,k}$ and the product rule for differentiation, we see that

$$\frac{d\tilde{\mathcal{L}}}{ds}\bigg|_{s=0} = \int_\Omega \left(\left(\operatorname{pr}^{(k)} V + \operatorname{Div} V\right) L\right)(x.\operatorname{pr}^{(k)} u(x)).$$

This is zero for all choices of Ω if and only if the integrand is zero, so we say the a Lagrangian L is invariant under a vector field V, or under the one parameter group that it generates, if

$$\left(\operatorname{pr}^{(k)} V + \operatorname{Div} V\right) L = 0.$$

As an example, we can check that $\operatorname{Div} V = 0$ for both the translation and rotation groups in the preceding section, so the Lagrangian

$$L(t,x,y,u,u_t,u_x,u_y) = \frac{1}{2}\left(u_t^2 - u_x^2 - u_y^2\right)$$

is invariant under those groups. A more interesting example is provided by the Lagrangian

$$L(x,u,u_x) = \sqrt{1+u_x^2}$$

on $\mathcal{R}_{1,1,1}$ and the rotation group

$$\tilde{x} = x\cos\theta - u\sin\theta, \quad \tilde{u} = x\sin\theta + u\cos\theta$$

with generator

$$V = -u\frac{\partial}{\partial_x} + x\frac{\partial_u}{,}$$

whose first prolongation is

$$\mathrm{pr}^{(1)}\, V = -u\frac{\partial}{\partial_x} + x\frac{\partial_u}{+}(1+u_x^2)\frac{\partial}{\partial u_x}.$$

Then

$$(\mathrm{pr}^{(1)}\, V)L = u_x\sqrt{1+u_x^2}$$

and[1]

$$\mathrm{Div}\, V = -u_x,$$

so

$$\left(\mathrm{pr}^{(k)}\, V + \mathrm{Div}\, V\right) L = 0$$

even though neither $(\mathrm{pr}^{(1)}\, V)L$ nor $(\mathrm{Div}\, V)L$ is zero separately. The integral

$$\int L(x, u(x), u'(x))\, dx = \int \sqrt{1+u'(x)^2}\, dx$$

is therefore rotationally invariant. This isn't surprising, because it gives the arc length of the graph of u.

Finally we make an observation which will be useful in the last section. If V is tangential to the graph of u then, as we've already seen,

$$\mathrm{pr}^{(k)}\, V = \sum_{i=1}^{m} \xi_i D_i,$$

so

$$\mathrm{pr}^{(k)}\, V L = \sum_{i=1}^{m} \xi_i D_i L,$$

$$(\mathrm{Div}\, V)L = \left(\sum_{i=1}^{m} D_i \xi_i\right) L$$

and

$$\left(\mathrm{pr}^{(k)}\, V + \mathrm{Div}\, V\right) L = \sum_{i=1}^{m} D_i(\xi_i L) = \mathrm{Div}(L\xi).$$

[1]Why is the divergence of the generator of the rotation group non-zero here when it was zero in the preceding example? The two situations are not comparable because the rotations here mix the dependent and independents variables while the ones there affect only the independent variables.

8 Conservation Laws

A conserved current of order l is an m-tuple $P_1, \ldots, P_m$ of functions on $\mathcal{R}_{m,n,l}$ such that

$$\operatorname{Div} P = 0$$

when applied to any solution of the Euler-Lagrange equations. The equation $\operatorname{Div} P = 0$ is called a conservation law. We also call the one paramater group generated by V a variational symmetry group for the Lagrangian L.

Consider, for example $m = 2$, $n = 1$, $k = 1$ and the Dirichlet Lagrangian

$$L(x, y, u, u_x, u_y) = \frac{1}{2}\left(u_x^2 + u_y^2\right).$$

The 2-tuple

$$P_x = \frac{1}{2}\left(u_x^2 - u_y^2\right), \quad P_y = u_x u_y$$

is a conserved current because

$$\operatorname{Div} P = D_x P_x + D_y P_y = (u_x u_{xx} - u_y u_{xy}) + (u_y u_{xy} + u_x u_{yy}) = (u_{xx} + u_{yy}) u_x$$

is zero whenever $u_{xx} + u_{yy} = 0$. Similarly,

$$P_x = \frac{1}{2} x \left(u_x^2 - u_y^2\right) + y u_x u_y, \quad P_y = x u_x u_y - \frac{1}{2} y \left(u_x^2 - u_y^2\right)$$

is a conserved current, because

$$D_x P_x = \frac{1}{2}\left(u_x^2 - u_y^2\right) + x(u_x u_{xx} - u_y u_{xy}) + y(u_y u_{xx} + u_x u_{xy})$$

and

$$D_y P_y = -\frac{1}{2}\left(u_x^2 - u_y^2\right) + x(u_x u_{yy} + u_y u_{xy}) - y(u_x u_{xy} - u_y u_{yy})$$

and

$$\operatorname{Div} P = D_x P_x + D_y P_y = (u_{xx} + u_{yy})(x u_x + y u_y).$$

9 Noether's Theorem

There is a close connection between variational symmetries and conserved currents, which was discovered by Emmy Noether[2]. Noether's theorem is most often used for first order Lagrangians, although the example which motivated her paper was of second order[3] In that case we have

$$\operatorname{Div} P = (\operatorname{pr}^{(k)} V + \operatorname{Div} V)L - \sum_{j=1}^{n} Q_j \sum_{|\alpha| \le k} (-1)^{|\alpha|} \left(D^\alpha \frac{\partial L}{\partial u_{j,\alpha}} \right),$$

where

$$Q_j = \eta_j - \sum_{l=1}^{m} \xi_l u_{j,\hat{l}}.$$

and

$$P_i = \xi_i L + \sum_{j=1}^{n} Q_j \frac{\partial L}{\partial u_{j,\hat{\imath}}}.$$

This holds for *any* choice of L and V. If we evaluate this at $x, \operatorname{pr}^{(k)} u(x)$ where u is a solution of the Euler-Lagrange equations for L then the sum

$$\sum_{j=1}^{n} Q_j \sum_{|\alpha| \le k} (-1)^{|\alpha|} D^\alpha \frac{\partial L}{\partial u_{j,\alpha}}$$

drops out and we have

$$\operatorname{div} P(x, \operatorname{pr}^{(k)} u(x)) = ((\operatorname{pr}^{(k)} V + \operatorname{Div} V)L)(x, \operatorname{pr}^{(k)} u(x)).$$

[2]Noether published her paper in 1918. At the time she was teaching at Göttingen but had no formal position and no salary, because the university refused to hire women. Her paper was more or less ignored until 1951. It then attracted more attention than actual understanding. By 1986 Peter Olver counted more than 50 papers which claimed to generalise it, while in fact only reproving special cases. No doubt there have been many more since. Certainly that's what the "generalisation" in the Wikipedia article on Noether's theorem does. The version of Noether's theorem proved in these notes is also only a special case, but somewhat less special than can be found in most treatments.

[3]Specifically the motivation of the paper was to explain the failure of energy conservation in general relativity, which was then a new and unproven theory. The relevant Lagrangian, the Hilbert action, is of second order, although its Euler-Lagrange equations, the Einstein equations, are only of second order, rather than fourth, as one might expect.

In other words, P is a conserved current if and only if V is a variational symmetry. We get a similar result for $k > 1$, but the form of P is much more complicated.[4]

A simple example is provided by time translation

$$\tilde{t} = t + s, \quad \tilde{x} = x, \tilde{y} = y, \quad \tilde{u} = u$$

for the three dimensional wave equation

$$\frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2} - \frac{\partial^2 u}{\partial z^2} = 0,$$

which is the Euler-Lagrange equation of

$$-\frac{1}{2}u_t^2 + \frac{1}{2}u_x^2 + \frac{1}{2}u_y^2 + \frac{1}{2}u_z^2.$$

The vector field associated to this group of transformations is

$$V = \frac{\partial}{\partial t},$$

with coefficients

$$\xi_t = 1, \quad \xi_x = 0, \quad \xi_y = 0, \quad \xi_z = 0, \quad \eta = 0$$

and its prolongation is also $\partial/\partial t$. Clearly

$$(\mathrm{pr}^{(k)} V + \mathrm{Div}\, V)L = 0,$$

so this is a variational symmetry of the Lagrangian. We have

$$Q = \eta - \xi_t u_t - \xi_x u_x - \xi_y u_y - \xi_z u_z = -u_t$$

and

$$\begin{aligned} P_t &= \xi_t L + Q\frac{\partial L}{\partial u_t} \\ &= -\frac{1}{2}u_t^2 + \frac{1}{2}u_x^2 + \frac{1}{2}u_y^2 + \frac{1}{2}u_z^2 + (-u_t)(-u_t) \\ &= \frac{1}{2}u_t^2 + \frac{1}{2}u_x^2 + \frac{1}{2}u_y^2 + \frac{1}{2}u_z^2, \end{aligned}$$

[4] It is perhaps not a coincidence that this paper was Noether's only work in analysis. She is famous primarily as an algebraist. The motivation of the result is physical and analytic, but the bulk of the work involved in proving it is algebraic.

$$P_x = \xi_x L + Q\frac{\partial L}{\partial u_x} = 0 + (-u_t)(u_x) = -u_t u_x,$$
$$P_y = \xi_y L + Q\frac{\partial L}{\partial u_y} = 0 + (-u_t)(u_y) = -u_t u_y,$$
$$P_z = \xi_z L + Q\frac{\partial L}{\partial u_z} = 0 + (-u_t)(u_z) = -u_t u_z$$

As promised,

$$\begin{aligned} D_t P_t + D_x P_x + D_y P_y + D_z P_z &= u_t u_{tt} + u_x u_{tx} + u_y u_{ty} + u_z u_{tz} \\ &\quad - u_t u_{xx} - u_x u_{tx} \\ &\quad - u_t u_{yy} - u_y u_{ty} \\ &\quad - u_t u_{zz} - u_z u_{tz} \\ &= u_t(u_{tt} - u_{xx} - u_{yy} - u_{zz}) = 0 \end{aligned}$$

if

$$u_{tt} - u_{xx} - u_{yy} - u_{zz} = 0.$$

Physically this expresses energy conservation.

10 A Longer Example

For a more complicated example, consider the transformation

$$\tilde{t} = \frac{t}{t^2 - x^2 - y^2 - z^2}, \quad \tilde{x} = \frac{x}{t^2 - x^2 - y^2 - z^2},$$
$$\tilde{y} = \frac{y}{t^2 - x^2 - y^2 - z^2}, \quad \tilde{z} = \frac{z}{t^2 - x^2 - y^2 - z^2},$$
$$\tilde{u} = (t^2 - x^2 - y^2 - z^2)u.$$

which is called inversion. This is a discrete symmetry, without a parameter, so at first sight Noether's theorem doesn't tell us anything. Let's check anyway that it is in fact a symmetry. First note that inversion is its own inverse.

$$t = \frac{\tilde{t}}{\tilde{t}^2 - \tilde{x}^2 - \tilde{y}^2 - \tilde{z}^2}, \quad x = \frac{\tilde{x}}{\tilde{t}^2 - \tilde{x}^2 - \tilde{y}^2 - \tilde{z}^2},$$
$$y = \frac{\tilde{y}}{\tilde{t}^2 - \tilde{x}^2 - \tilde{y}^2 - \tilde{z}^2}, \quad z = \frac{\tilde{z}}{\tilde{t}^2 - \tilde{x}^2 - \tilde{y}^2 - \tilde{z}^2},$$

$$u = (\tilde{t}^2 - \tilde{x}^2 - \tilde{y}^2 - \tilde{z}^2)\tilde{u}.$$

Next we need to check how

$$L\, dt\, dx\, dy\, dz$$

tranforms under inversion. The easiest way to do that is to use differentials. We set

$$\lambda = t^2 - x^2 - y^2 - z^2, \quad \tilde{\lambda} = \tilde{t}^2 - \tilde{x}^2 - \tilde{y}^2 - \tilde{z}^2$$

and observe that

$$\lambda\tilde{\lambda} = 1, \quad \lambda^{-1}\, d\lambda + \tilde{\lambda}^{-1}\, d\tilde{\lambda} = 0.$$

We have

$$d\tilde{t} = \lambda^{-1}\, dt - \lambda^{-2} t\, d\lambda, \quad d\tilde{x} = \lambda^{-1}\, dx - \lambda^{-2} x\, d\lambda,$$

$$d\tilde{y} = \lambda^{-1}\, dy - \lambda^{-2} y\, d\lambda, \quad d\tilde{z} = \lambda^{-1}\, dz - \lambda^{-2} z\, d\lambda,$$

$$\begin{aligned} d\tilde{t} \wedge d\tilde{x} \wedge d\tilde{y} \wedge d\tilde{z} &= \lambda^{-4}\, dt \wedge dx \wedge dy \wedge dz - \lambda^{-5} t\, d\lambda \wedge dx \wedge dy \wedge dz \\ &\quad - \lambda^{-5} x\, dt \wedge d\lambda \wedge dy \wedge dz - \lambda^{-5} y\, dt \wedge dx \wedge d\lambda \wedge dz \\ &\quad - \lambda^{-5} z\, dt \wedge dx \wedge dy \wedge d\lambda \\ &= \lambda^{-4}\, dt \wedge dx \wedge dy \wedge dz - 2\lambda^{-5} t^2\, dt \wedge dx \wedge dy \wedge dz \\ &\quad - 2\lambda^{-5} x^2\, dt \wedge dx \wedge dy \wedge dz - 2\lambda^{-5} y^2\, dt \wedge dx \wedge dy \wedge dz \\ &\quad - 2\lambda^{-5} z^2\, dt \wedge dx \wedge dy \wedge dz \\ &= \; - \lambda^{-4}\, dt \wedge dx \wedge dy \wedge dz. \end{aligned}$$

$$\begin{aligned} d\tilde{u} &= u\, d\lambda + \lambda\, du \\ &= u\, d\tilde{\lambda} + \lambda(u_t\, dt + u_x\, dx + u_y\, dy + u_z\, dz) \\ &= \; - \tilde{\lambda}^{-2} u\, d\tilde{\lambda} + \tilde{\lambda}^{-2}(u_t\, d\tilde{t} + u_x\, d\tilde{x} + u_y\, d\tilde{y} + u_z\, d\tilde{z}) \\ &\quad - \tilde{\lambda}^{-3}(tu_t + xu_x + yu_y + zu_z) d\tilde{\lambda} \\ &= \tilde{\lambda}^{-2}(u_t\, d\tilde{t} + u_x\, d\tilde{x} + u_y\, d\tilde{y} + u_z\, d\tilde{z}) \\ &\quad - \tilde{\lambda}^{-3}(\tilde{t}u_t + \tilde{x}u_x + \tilde{y}u_y + \tilde{z}u_z) d\tilde{\lambda} \\ &= \tilde{\lambda}^{-2}(u_t\, d\tilde{t} + u_x\, d\tilde{x} + u_y\, d\tilde{y} + u_z\, d\tilde{z}) \\ &\quad - 2\tilde{\lambda}^{-2}(tu_t + xu_x + yu_y + zu_z + u)(\tilde{t}\; d\tilde{t} - \tilde{x}\, d\tilde{x} - \tilde{y}\, d\tilde{y} - \tilde{z}\, d\tilde{z}) \end{aligned}$$

Since $\tilde{u}_{\tilde{t}}$, $\tilde{u}_{\tilde{x}}$, $\tilde{u}_{\tilde{y}}$ and $\tilde{u}_{\tilde{z}}$ are determined uniquely by

$$d\tilde{u} = \tilde{u}_{\tilde{t}}\, d\tilde{t} + \tilde{u}_{\tilde{x}}\, d\tilde{x} + \tilde{u}_{\tilde{y}}\, d\tilde{y} + \tilde{u}_{\tilde{z}}\, d\tilde{z}$$

it follows that

$$\tilde{u}_{\tilde{t}} = \tilde{\lambda}^{-2}(u_t - 2v\tilde{t}), \quad \tilde{u}_{\tilde{x}} = \tilde{\lambda}^{-2}(u_x + 2v\tilde{x}),$$

$$\tilde{u}_{\tilde{y}} = \tilde{\lambda}^{-2}(u_y + 2v\tilde{y}), \quad \tilde{u}_{\tilde{z}} = \tilde{\lambda}^{-2}(u_z + 2v\tilde{z}),$$

where

$$v = tu_t + xu_x + yu_y + zu_z + u$$

Then

$$\begin{aligned}
\tilde{L} &= -\frac{1}{2}\tilde{u}_{\tilde{t}}^2 + \frac{1}{2}\tilde{u}_{\tilde{x}}^2 + \frac{1}{2}\tilde{u}_{\tilde{y}}^2 + \frac{1}{2}\tilde{u}_{\tilde{z}}^2 \\
&= \tilde{\lambda}^{-4}\left(-\frac{1}{2}u_t^2 + \frac{1}{2}u_x^2 + \frac{1}{2}u_y^2 + \frac{1}{2}u_z^2\right) + 2\tilde{\lambda}^{-4}v(\tilde{t}u_t + \tilde{x}u_x + \tilde{y}u_y + \tilde{z}u_z) \\
&\quad -2\tilde{\lambda}^{-4}v^2(\tilde{t}^2 - \tilde{x}^2 - \tilde{y}^2 - \tilde{z}^2) \\
&= \tilde{\lambda}^{-4}L + 2\tilde{\lambda}^{-4}v(\tilde{t}u_t + xu_x + yu_y + zu_z - v) \\
&= \tilde{\lambda}^{-4}L - 2\tilde{\lambda}^{-3}u(tu_t + xu_x + yu_y + zu_z + u)
\end{aligned}$$

Therefore

$$\tilde{L}\, d\tilde{t}\, d\tilde{x}\, d\tilde{y}\, d\tilde{z} = \left(L - 2\frac{u(tu_t + xu_x + yu_y + zu_z + u)}{t^2 - x^2 - y^2 - z^2}\right) dt\, dx\, dy\, dz$$

At first sight this doesn't look particularly invariant, but

$$2\frac{u(tu_t + xu_x + yu_y + zu_z + u)}{t^2 - x^2 - y^2 - z^2}$$

is the total divergence of the vector field

$$\frac{u^2}{t^2 - x^2 - y^2 - z^2}\left(t\frac{\partial}{\partial t} + x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y} + z\frac{\partial}{\partial z}\right).$$

In other words, it is null Lagrangian and therefore makes no contribution to the Euler-Lagrange equations. So for purposes of Noether's theorem inversion acts like a variational symmetry.

We still don't have one parameter group of symmetries, just a discrete symmetry, but this is easily remedied. We can perform and inversion, followed by a time translation, followed by another inversion. Each transformation individually is fine so, up to null Lagrangians, the net result is a variational symmetry. Explicitly the resulting transformation is

$$\tilde{t} = \frac{t + s(t^2 - x^2 - y^2 - z^2)}{1 + 2st + s^2(t^2 - x^2 - y^2 - z^2)}, \quad \tilde{x} = \frac{x}{1 + 2st + s^2(t^2 - x^2 - y^2 - z^2)},$$

$$\tilde{y} = \frac{y}{1 + 2st + s^2(t^2 - x^2 - y^2 - z^2)}, \quad \tilde{z} = \frac{z}{1 + 2st + s^2(t^2 - x^2 - y^2 - z^2)},$$

$$\tilde{u} = \frac{u}{1 + 2st + s^2(t^2 - x^2 - y^2 - z^2)}.$$

The associated vector field is

$$\xi_t = -t^2 - x^2 - y^2 - z^2, \quad \xi_x = -2tx, \quad \xi_y = -2ty, \quad \xi_z = -2tz, \quad \eta = 2tu.$$

There is no need to compute the prolongation, since we already know this is a variational symmetry. We just compute

$$Q = \eta - \xi_t u_t - \xi_x u_x - \xi_y u_y - \xi_z u_z = 2tu + (t^2 + x^2 + y^2 + z^2) + 2txu_x + 2tyu_y + 2tzu_z$$

and

$$\begin{aligned}
P_t &= \xi_t L + Q \frac{\partial L}{\partial u_t} \\
&= -\frac{1}{2}(t^2 + x^2 + y^2 + z^2)(u_t^2 + u_x^2 + u_y^2 + u_z^2) - 2tu_t(xu_x + yu_x + zu_z + u), \\
P_x &= \xi_x L + Q \frac{\partial L}{\partial u_x} \\
&= (t^2 + x^2 + y^2 + z^2)u_t u_x + tx(u_t^2 + u_x^2 - u_y^2 - u_z^2) + 2tu_x(yu_y + zu_z + u), \\
P_y &= \xi_y L + Q \frac{\partial L}{\partial u_y} \\
&= (t^2 + x^2 + y^2 + z^2)u_t u_y + tx(u_t^2 - u_x^2 + u_y^2 - u_z^2) + 2tu_y(xu_y + zu_z + u), \\
P_z &= \xi_z L + Q \frac{\partial L}{\partial u_z} \\
&= (t^2 + x^2 + y^2 + z^2)u_t u_z + tx(u_t^2 - u_x^2 - u_y^2 + u_z^2) + 2tu_z(xu_x + yu_y + u).
\end{aligned}$$

By Noether's theorem this must be a conserved current.[5]

[5]This was first found by Cathleen Morawetz, the daughter of John Lighton Synge, after whom the Synge Lecture Theatre in the Hamilton Building is named. Perhaps surprisingly, she didn't find it using Noether's theorem, but simply by guessing the form of the conserved current.