Calculus of Variations, Symmetries, and Conservation Laws

33

In this chapter we shall start with one of the oldest and most useful branches of mathematical physics, the calculus of variations. After giving the fundamentals and some examples, we shall investigate the consequences of symmetries associated with variational problems. The chapter then ends with Noether's theorem, which connects such symmetries with their associated conservation laws. All vector spaces of relevance in this chapter will be assumed to be real.

33.1 The Calculus of Variations

One of the main themes of calculus is the extremal problem: Given a function $f : \mathbb{R} \supset D \to \mathbb{R}$, find the points in the domain D of f at which f attains a maximum or minimum. To locate such points, we find the zeros of the derivative of f. For multivariable functions, $f : \mathbb{R}^p \supset \Omega \to \mathbb{R}$, the notion of *gradient* generalizes that of the derivative. To find the jth component of the gradient ∇f, we calculate the difference Δf between the value of f at $(x^1, \dots, x^j + \varepsilon, \dots, x^p)$ and its value at $(x^1, \dots, x^j, \dots, x^p)$, divide this difference by ε, and take the limit $\varepsilon \to 0$. This is simply partial differentiation, and the jth component of the gradient is just the jth partial derivative of f.

33.1.1 Derivative for Hilbert Spaces

To make contact with the subject of this chapter, let us reinterpret the notion of differentiation. The most useful interpretation is geometric. In fact, our first encounter with the derivative is geometrical: We are introduced to the concept through lines tangent to curves. In this language, the derivative of a function $f : \mathbb{R} \supset \Omega \to \mathbb{R}$ at x_0 is a line (or function) $\psi : \Omega \supset \Omega_0 \to \mathbb{R}$ passing through $(x_0, f(x_0))$ whose slope is defined to be the derivative of f at x_0 (see Fig. 33.1):

$$\psi(x) = f(x_0) + f'(x_0)(x - x_0).$$

S. Hassani, *Mathematical Physics*, DOI 10.1007/978-3-319-01195-0_33,

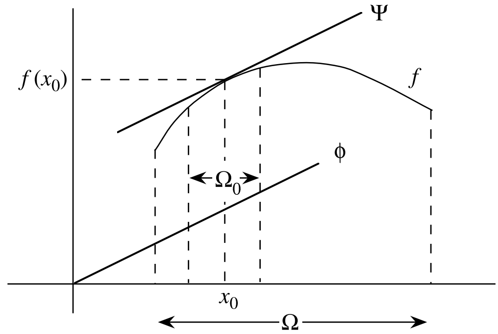


Fig. 33.1 The derivative at $(x_0, f(x_0))$ as a *linear function* passing through the origin with a slope $f'(x_0)$. The function f is assumed to be defined on a subset Ω of the real line. Ω_0 restricts the x's to be close to x_0 to prevent the function from misbehaving (blowing up), and to make sure that the limit in the definition of derivative makes sense

The function $\psi(x)$ describes a line, but it is not a *linear* function (in the vector-space sense of the word). The requirement of linearity is due to our desire for generalization of differentiation to Hilbert spaces, on which linear maps are the most natural objects. Therefore, we consider the line parallel to $\psi(x)$ that passes through the origin. Call this $\phi(x)$. Then

$$\phi(x) = f'(x_0)x, \tag{33.1}$$

which is indeed a linear function. We identify $\phi(x)$ as the derivative of f at x_0. This identification may appear strange at first but, as we shall see shortly, is the most convenient and useful. Of course, any identification requires a one-to-one correspondence between objects identified. It is clear that indeed there is a one-to-one correspondence between derivatives at points and linear functions with appropriate slopes.

Equation (33.1) can be used to geometrize the definition of derivative. First consider

$$f'(x_0) = \frac{\phi(x) - \phi(x_0)}{x - x_0}, \quad \text{and} \quad f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}.$$

Next note that, contrary to f which is usually defined only for a subset of the real line, ϕ is defined for all real numbers $\mathbb{R}$, and that $\phi(x - x_0) = \phi(x) - \phi(x_0)$ due to the linearity of ϕ. Thus, we have

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = \frac{\phi(x) - \phi(x_0)}{x - x_0} = \lim_{x \to x_0} \frac{\phi(x - x_0)}{x - x_0},$$

or

$$\lim_{x \to x_0} \frac{|f(x) - f(x_0) - \phi(x - x_0)|}{|x - x_0|} = 0 \tag{33.2}$$

where we have introduced absolute values in anticipation of its analogue—norm. Equation (33.2) is readily generalized to any complete normed vector space (Banach space), and in particular to any Hilbert space:

Definition 33.1.1 Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be Hilbert spaces with norms $\|\cdot\|_1$ and $\|\cdot\|_2$, respectively. Let $f : \mathcal{H}_1 \supset \Omega \to \mathcal{H}_2$ be any map and $|x_0\rangle \in \Omega$. Suppose there is a linear map $\mathbf{T} \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ with the property that

$$\lim_{\|x-x_0\|_1 \to 0} \frac{\|f(|x\rangle) - f(|x_0\rangle) - \mathbf{T}(|x\rangle - |x_0\rangle)\|_2}{\|x - x_0\|_1} = 0 \quad \text{for } |x\rangle \in \Omega.$$

Then, we say that f is **differentiable at** $|x_0\rangle$, and we define the **derivative of** f **at** $|x_0\rangle$ to be $\mathbf{D}f(x_0) \equiv \mathbf{T}$. If f is differentiable at each $|x\rangle \in \Omega$, the map

differentiability of a function on a Hilbert space at a point

$$\mathbf{D}f : \Omega \to \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2) \quad \text{given by} \quad \mathbf{D}f\big(|x\rangle\big) = \mathbf{D}f(x)$$

is called the **derivative** of f.

The reader may verify that if the derivative exists, it is unique.

Example 33.1.2 Let $\mathcal{H}_1 = \mathbb{R}^n$ and $\mathcal{H}_2 = \mathbb{R}^m$ and $f : \mathbb{R}^n \supset \Omega \to \mathbb{R}^m$. Then for $|x\rangle \in \Omega$, $\mathbf{D}f(x)$ is a linear map, which can be represented by a matrix in the standard bases of $\mathbb{R}^n$ and $\mathbb{R}^m$. To find this matrix, we need to let $\mathbf{D}f(x)$ act on the jth standard basis of $\mathbb{R}^n$, i.e., we need to evaluate $\mathbf{D}f(x)|e_j\rangle$. This suggests taking $|y\rangle = |x\rangle + h|e_j\rangle$ (with $h \to 0$) as the vector appearing in the definition of derivative at $|x\rangle$. Then

$$\frac{\|f(|y\rangle) - f(|x\rangle) - \mathbf{D}f(x)(|y\rangle - |x\rangle)\|_2}{\|y - x\|_1} = \frac{\|f(x^1, \dots, x^j + h, \dots, x^n) - f(x^1, \dots, x^j, \dots, x^n) - h\mathbf{D}f(x)|e_j\rangle\|_2}{|h|}$$

approaches zero as $h \to 0$, so that the ith component of the ratio also goes to zero. But the ith component of $\mathbf{D}f(x)|e_j\rangle$ is simply a^i_j, the ijth component of the matrix of $\mathbf{D}f(x)$. Therefore,

$$\lim_{h\to 0} \frac{|f^i(x^1, \dots, x^j + h, \dots, x^n) - f^i(x^1, \dots, x^j, \dots, x^n) - ha^i_j|}{|h|} = 0,$$

which means that $a^i_j = \partial f^i / \partial x^j$.

The result of the example above can be stated as follows:

Box 33.1.3 *For $f : \mathbb{R}^n \supset \Omega \to \mathbb{R}^m$, the matrix of $\mathbf{D}f(x)$ in the standard basis of $\mathbb{R}^n$ and $\mathbb{R}^m$ is the **Jacobian matrix** of f.*

The case of $\mathcal{H}_2 = \mathbb{R}$ deserves special attention. Let $\mathcal{H}$ be a Hilbert space. Then $\mathbf{D}f(x) \in \mathcal{L}(\mathcal{H}, \mathbb{R}) = \mathcal{H}^*$ is denoted by $\mathbf{d}f(x)$ and renamed the **differential** of f at $|x\rangle$. Furthermore, through the inner product, one can identify $\mathbf{d}f : \Omega \to \mathcal{H}^*$ with another map defined as follows:

differential and gradient of f at $|x\rangle$

Definition 33.1.4 Let $\mathcal{H}$ be a Hilbert space and $f : \mathcal{H} \supset \Omega \to \mathbb{R}$. The **gradient** ∇f of f is the map $\nabla f : \Omega \to \mathcal{H}$ defined by

$$\langle \nabla f(x)|a\rangle \equiv \big\langle \mathbf{d} f(x), a\big\rangle \quad \forall |x\rangle \in \Omega,\ |a\rangle \in \mathcal{H}$$

where $\langle\, ,\rangle$ is the pairing $\langle\, ,\rangle : \mathcal{H}^* \times \mathcal{H} \to \mathbb{R}$ of $\mathcal{H}$ and its dual.

Note that although f is not an element of $\mathcal{H}^*$, $\mathbf{d}f(x)$ is, for all points $|x\rangle \in \Omega$ at which the differential is defined.

Example 33.1.5 Consider the function $f : \mathcal{H} \to \mathbb{R}$ given by $f(|x\rangle) = \|x\|^2$. Since

$$\|y - x\|^2 = \|y\|^2 - \|x\|^2 - 2\langle x|y - x\rangle$$

and since the derivative is unique, the reader may check that $\mathbf{d}f(x)|a\rangle = 2\langle x|a\rangle$, or $\nabla f(|x\rangle) = 2|x\rangle$.

Derivatives could be defined in terms of directions as well:

directional derivative

Definition 33.1.6 Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be Hilbert spaces. Let $f : \mathcal{H}_1 \supset \Omega \to \mathcal{H}_2$ be any map and $|x\rangle \in \Omega$. We say that f has a derivative in the direction $|a\rangle \in \mathcal{H}_1$ at $|x\rangle$ if

$$\frac{d}{dt} f\big(|x\rangle + t|a\rangle\big)\bigg|_{t=0}$$

exists. We call this element of $\mathcal{H}_2$ the **directional derivative of f in the direction** $|a\rangle \in \mathcal{H}_1$ at $|x\rangle$.

The reader may verify that if f is differentiable at $|x\rangle$ (in the context of Definition 33.1.1), then the directional derivative of f in any direction $|a\rangle$ exists at $|x\rangle$ and is given by

$$\frac{d}{dt} f\big(|x\rangle + t|a\rangle\big)\bigg|_{t=0} = \mathbf{D}f(x)|a\rangle. \tag{33.3}$$

33.1.2 Functional Derivative

We now specialize to the Hilbert space of square-integrable functions $\mathcal{L}^2(\Omega)$ for some open subset Ω of some $\mathbb{R}^m$. We need to change our notation somewhat. Let us agree to denote the elements of $\mathcal{L}^2(\Omega)$ by f, u, etc. Real-valued functions on $\mathcal{L}^2(\Omega)$ will be denoted by $\mathbf{L}$, $\mathbf{H}$, etc. The m-tuples will be denoted by boldface lowercase letters. To summarize,

$$\mathbf{x}, \mathbf{y} \in \mathbb{R}^m, \qquad f, u \in \mathcal{L}^2(\Omega) \quad \Rightarrow \quad f, u : \mathbb{R}^m \supset \Omega \to \mathbb{R},$$
$$\langle f|u\rangle = \int_\Omega f(\mathbf{x})u(\mathbf{x})\, d^m x, \qquad \mathbf{L}, \mathbf{H} : \mathcal{L}^2(\Omega) \to \mathbb{R}.$$

Furthermore, the evaluation of $\mathbf{L}$ at u is denoted by $\mathbf{L}[u]$.

When dealing with the space of functions, the gradient of Definition 33.1.4 is called a **functional derivative** or **variational derivative** and denoted by $\delta\mathbf{L}/\delta u$. So

functional derivative or variational derivative

$$\boxed{\left\langle \tfrac{\delta\mathbf{L}}{\delta u}\middle| f\right\rangle \equiv \int_\Omega \frac{\delta\mathbf{L}}{\delta u}(\mathbf{x}) f(\mathbf{x})\, d^m x = \frac{d}{dt}\mathbf{L}[u+tf]\bigg|_{t=0},} \tag{33.4}$$

where we have used Eq. (33.3). Note that by definition, $\delta\mathbf{L}/\delta u$ is an element of the Hilbert space $\mathcal{L}^2(\Omega)$; so, the integral of (33.4) makes sense. Equation (33.4) is frequently used to compute functional derivatives.

An immediate consequence of Eq. (33.4) is the following important result.

Proposition 33.1.7 *Let $\mathbf{L} : \mathcal{L}^2(\Omega) \to \mathbb{R}$ for some $\Omega \subset \mathbb{R}^m$. If $\mathbf{L}$ has an extremum at u, then*

$$\frac{\delta\mathbf{L}}{\delta u} = 0.$$

Proof If $\mathbf{L}$ has an extremum at u, then the RHS of (33.4) vanishes for *any* function f, in particular, for any orthonormal basis vector $|e_i\rangle$. Completeness of a basis now implies that the directional derivative must vanish (see Proposition 7.1.9). □

Just as in the case of partial derivatives, where some simple relations such as derivative of powers and products can be used to differentiate more complicated expressions, there are some primitive formulas involving functional derivatives that are useful in computing other more complicated expressions. Consider the **evaluation function**

evaluation function

$$\mathbf{E}_\mathbf{y} : \mathcal{L}^2(\Omega) \to \mathbb{R} \quad \text{given by} \quad \mathbf{E}_\mathbf{y}[f] = f(\mathbf{y}).$$

Using Eq. (33.4), we can easily compute the functional derivative of $\mathbf{E}_\mathbf{y}$:

$$\int_\Omega \frac{\delta\mathbf{E}_\mathbf{y}[u]}{\delta u}(\mathbf{x}) f(\mathbf{x})\, d^m x = \frac{d}{dt}\mathbf{E}_\mathbf{y}[u+tf]\bigg|_{t=0} = \frac{d}{dt}\big\{u(\mathbf{y}) + tf(\mathbf{y})\big\}\bigg|_{t=0}$$
$$= f(\mathbf{y}) \quad \Rightarrow \quad \frac{\delta\mathbf{E}_\mathbf{y}[u]}{\delta u}(\mathbf{x}) = \delta(\mathbf{x}-\mathbf{y}). \tag{33.5}$$

It is instructive to compare (33.5) with the similar formula in multivariable calculus, where real-valued functions f take a vector $\mathbf{x}$ and give a real number. The analogue of the evaluation function is E_i, which takes a vector $\mathbf{x}$ and gives the real number x^i, the ith component of $\mathbf{x}$. Using the definition of partial derivative, one readily shows that $\partial E_i/\partial x^j = \delta_{ij}$, which is (somewhat less precisely) written as $\partial x^i/\partial x^j = \delta_{ij}$. The same sort of imprecision is used to rewrite Eq. (33.5) as

$$\frac{\delta u(\mathbf{y})}{\delta u(\mathbf{x})} \equiv \frac{\delta u_\mathbf{y}}{\delta u_\mathbf{x}} = \delta(\mathbf{x}-\mathbf{y}), \tag{33.6}$$

where we have turned the arguments into indices to make the analogy with the discrete case even stronger.

Another useful formula concerns *derivatives* of square-integrable functions. Let $\mathbf{E}_{\mathbf{y},i}$ denote the evaluation of the derivative of functions with respect to the ith coordinate:

$$\mathbf{E}_{\mathbf{y},i} : \mathcal{L}^2(\Omega) \to \mathbb{R} \quad \text{given by} \quad \mathbf{E}_{\mathbf{y},i}(f) = \partial_i f(\mathbf{y}).$$

Then a similar argument as above will show that

$$\frac{\delta \mathbf{E}_{\mathbf{y},i}}{\delta u}(\mathbf{x}) = -\partial_i \delta(\mathbf{x}-\mathbf{y}), \quad \text{or} \quad \frac{\delta \partial_i u(\mathbf{y})}{\delta u(\mathbf{x})} = -\partial_i \delta(\mathbf{x}-\mathbf{y}),$$

and in general,

$$\frac{\delta \partial_{i_1 \ldots i_k} u(\mathbf{y})}{\delta u(\mathbf{x})} = (-1)^k \partial_{i_1 \ldots i_k} \delta(\mathbf{x}-\mathbf{y}). \tag{33.7}$$

Equation (33.7) holds only if the function f, the so-called *test function*, vanishes on $\partial\Omega$, the boundary of the region of integration. If it does not, then there will be a "surface term" that will complicate matters considerably. Fortunately, in most applications this surface term is *required* to vanish. So, let us adhere to the convention that

Box 33.1.8 *All test functions $f(\mathbf{x})$ appearing in the integral of Eq.* (33.4) *are assumed to vanish at the boundary of Ω.*

For applications, we need to generalize the concept of functions on Hilbert spaces. First, it is necessary to consider maps from a Hilbert space to $\mathbb{R}^n$. For simplicity, we confine ourselves to the Hilbert space $\mathcal{L}^2(\Omega)$. Such a map $\mathbf{H} : \mathcal{L}^2(\Omega) \to D \subset \mathbb{R}^n$, for some subset D of $\mathbb{R}^n$, can be written in components

$$\mathbf{H} = (\mathbf{H}_1, \mathbf{H}_2, \ldots, \mathbf{H}_n), \quad \text{where} \quad \mathbf{H}_i : \mathcal{L}^2(\Omega) \to \mathbb{R}, \quad i = 1, \ldots, n.$$

Next, we consider an ordinary multivariable function $L : \mathbb{R}^n \supset D \to \mathbb{R}$, and use it to construct a new function on $\mathcal{L}^2(\Omega)$, the composite of L and $\mathbf{H}$:

$$L \circ \mathbf{H} : \mathcal{L}^2(\Omega) \to \mathbb{R}, \qquad L \circ \mathbf{H}[u] = L\big(\mathbf{H}_1[u], \ldots, \mathbf{H}_n[u]\big).$$

Then the functional derivative of $L \circ \mathbf{H}$ can be obtained using the chain rule and noting that the derivative of L is the common partial derivative. It follows that

$$\frac{\delta L \circ \mathbf{H}[u]}{\delta u}(\mathbf{x}) = \left\{ \frac{\delta}{\delta u} L\big(\mathbf{H}_1[u], \ldots, \mathbf{H}_n[u]\big) \right\}(\mathbf{x}) = \sum_{i=1}^{n} \partial_i L \frac{\delta \mathbf{H}_i}{\delta u}(\mathbf{x}), \tag{33.8}$$

where $\partial_i L$ is the partial derivative of L with respect to its ith argument.

Example 33.1.9 Let $L : (a, b) \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, be a function of three variables the first one of which is defined for the real interval (a, b). Let $\mathbf{H}_i : \mathcal{L}^2(a, b) \to \mathbb{R}$, $i = 1, 2, 3$, be defined by

$$\mathbf{H}_1[u] \equiv x, \qquad \mathbf{H}_2[u] = \mathbf{E}_x[u] = u(x), \qquad \mathbf{H}_3[u] = \mathbf{E}'_x[u] \equiv u'(x),$$

where $\mathbf{E}_x$ is the evaluation function and $\mathbf{E}'_x$ evaluates the derivative. It follows that $L \circ \mathbf{H}[u] = L(x, u(x), u'(x))$. Then, noting that $\mathbf{H}_1[u]$ is independent of u, we have

$$\begin{aligned}\frac{\delta L \circ \mathbf{H}[u]}{\delta u}(y) &= \partial_1 L \frac{\delta \mathbf{H}_1[u]}{\delta u}(y) + \partial_2 L \frac{\delta \mathbf{E}_x[u]}{\delta u}(y) + \partial_3 L \frac{\delta \mathbf{E}'_x[u]}{\delta u}(y) \\ &= 0 + \partial_2 L \delta(y - x) - \partial_3 L \delta'(y - x) = \partial_2 L \delta(x - y) \\ &\quad + \partial_3 L \delta'(x - y).\end{aligned}$$

This is normally written as

$$\frac{\delta L(x, u(x), u'(x))}{\delta u}(y) = \frac{\partial L}{\partial u}\delta(x - y) + \frac{\partial L}{\partial u'}\delta'(x - y), \tag{33.9}$$

which is the unintegrated version of the classical Euler-Lagrange equation for a single particle, to which we shall return shortly.

A generalization of the example above turns L into a function on $\Omega \times \mathbb{R} \times \mathbb{R}^m$ with $\Omega \subset \mathbb{R}^m$, so that

$$L\big(x^1, \ldots, x^m, u(\mathbf{x}), \partial_1 u(\mathbf{x}), \ldots, \partial_m u(\mathbf{x})\big) \in \mathbb{R}, \quad \text{with } \mathbf{x} \in \mathbb{R}^m.$$

The functions $\{\mathbf{H}_i\}_{i=1}^{2m+1}$ are defined as

$$\begin{aligned}&\mathbf{H}_i[u] \equiv x^i && \text{for } i = 1, 2, \ldots, m, \\ &\mathbf{H}_i[u] \equiv \mathbf{E}_{\mathbf{x}}[u] = u(\mathbf{x}) && \text{for } i = m + 1, \\ &\mathbf{H}_i[u] \equiv \mathbf{E}_{\mathbf{x},i}[u] = \partial_i u(\mathbf{x}) && \text{for } i = m + 2, \ldots, 2m + 1,\end{aligned}$$

and lead to the equation

$$\frac{\delta L \circ \mathbf{H}[u]}{\delta u}(\mathbf{y}) = \partial_{m+1} L \delta(\mathbf{x} - \mathbf{y}) + \sum_{i=m+2}^{2m+1} \partial_i L \partial_i \delta(\mathbf{x} - \mathbf{y}), \tag{33.10}$$

which is the unintegrated version of the classical Euler-Lagrange equation for a field in m dimensions.

33.1.3 Variational Problems

The fundamental theme of the calculus of variations is to find functions that extremize an integral and are fixed on the boundary of the integration region. A prime example is the determination of the equation of the curve of minimum length in the xy-plane passing through two points (a_1, b_1) and (a_2, b_2). Such a curve, written as $y = u(x)$, would minimize the integral

$$\mathbf{int}[u] \equiv \int_{a_1}^{a_2} \sqrt{1 + \left[u'(x)\right]^2}\, dx, \quad u(a_1) = b_1,\ u(a_2) = b_2. \qquad (33.11)$$

Note that **int** takes a function and gives a real number, i.e.—if we restrict our functions to square-integrable ones—**int** belongs to $\mathcal{L}^2(a_1, a_2)$. This is how contact is established between the calculus of variations and what we have studied so far in this chapter.

To be as general as possible, we allow the integral to contain derivatives up to the nth order. Then, using the notation of the previous chapter, we consider functions L on $M^{(n)} \subset \Omega \times U^{(n)}$, where we have replaced X with Ω, so that $M = \mathbb{R}^p \supset \Omega \times U \subset \mathbb{R}^q$.

nth-order variational problem; Lagrangian; functional

Definition 33.1.10 By an n**th-order variational problem** we mean finding the extremum of the real-valued function $\mathbf{L} : \mathcal{L}^2(\Omega) \to \mathbb{R}$ given by

$$\mathbf{L}[u] \equiv \int_\Omega L\left(x, u^{(n)}\right) d^p x, \qquad (33.12)$$

where Ω is a subset of $\mathbb{R}^p$, L is a real-valued function on $\Omega \times U^{(n)}$, and $p^{(n)} = (p + n)!/(n!p!)$. In this context the function L is called the **Lagrangian** of the problem, and $\mathbf{L}$ is called a **functional**.[1]

The solution to the variational problem is given by Proposition 33.1.7, moving the functional derivative inside the integral, and a straightforward (but tedious!) generalization of Eq. (33.10) to include derivatives of order higher than one. Due to the presence of the integral, the Dirac delta function and all its derivatives will be integrated out. Before stating the solution of the variational problem, let us introduce a convenient operator, using the total derivative operator introduced in Definition 32.3.3.

Euler operator

Definition 33.1.11 For $1 \le \alpha \le q$, the αth **Euler operator** is

$$\mathbb{E}_\alpha \equiv \sum_J (-D)_J \frac{\partial}{\partial u_J^\alpha}, \qquad (33.13)$$

where for $J = (j_1, \dots, j_k)$,

$$(-D)_J \equiv (-1)^k D_J = (-D_{j_1})(-D_{j_2}) \cdots (-D_{j_k}),$$

and the sum extends over *all multi-indices* $J = (j_1, \dots, j_k)$, including $J = 0$.

The negative signs are introduced because of the integration by parts involved in the evaluation of the derivatives of the delta function. Although the sum in Eq. (33.13) extends over *all* multi-indices, only a finite number of

[1] Do not confuse this functional with the *linear* functional of Chap. 2.

terms in the sum will be nonzero, because any function on which the Euler operator acts depends on a finite number of derivatives.

Theorem 33.1.12 *If u is an extremal of the variational problem (33.12), then it must be a solution of the **Euler-Lagrange** equations*

Euler-Lagrange equations

$$\mathbb{E}_\alpha(L) \equiv \sum_J (-D)_J \frac{\partial L}{\partial u_J^\alpha} = 0, \quad \alpha = 1, \dots, q.$$

For the special case of $p = q = 1$, the Euler operator becomes

$$\mathbb{E} = \frac{\partial}{\partial u} + \sum_{j=1}^{\infty} (-D_x)^j \frac{\partial}{\partial u_j} = \frac{\partial}{\partial u} - D_x \frac{\partial}{\partial u_x} + D_x^2 \frac{\partial}{\partial u_{xx}} - \cdots,$$

where D_x is the *total* derivative with respect to x, and u_j is the jth derivative of u with respect to x; and the Euler-Lagrange equation for the variational problem

$$\mathbf{L}[u] \equiv \int_a^b L\bigl(x, u^{(n)}\bigr)\, dx$$

becomes

$$\mathbb{E}(L) = \frac{\partial L}{\partial u} + \sum_{j=1}^{n} (-1)^j D_x^j \frac{\partial L}{\partial u_j} = 0. \tag{33.14}$$

Since L carries derivatives up to the n-th order and each D_x carries one derivative, we conclude that Eq. (33.14) is a $2n$-th order ODE.

Example 33.1.13 The variational problem of Eq. (33.11) has a Lagrangian

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

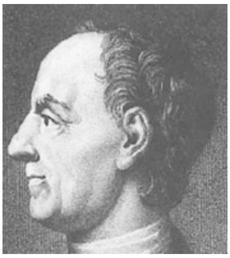
which is a function of the first derivative only. So, the Euler-Lagrange equation takes the form

$$0 = -D_x \frac{\partial L}{\partial u_x} = -D_x \left(\frac{u_x}{\sqrt{1 + u_x^2}} \right) = -\frac{d}{dx} \left(\frac{u_x}{\sqrt{1 + u_x^2}} \right) = -\frac{u_{xx}}{(1 + u_x^2)^{3/2}},$$

or $u_{xx} = 0$, so that $u = f(x) = c_1 x + c_2$. The solution to the variational problem is a straight line passing through the two points (a_1, b_1) and (a_2, b_2).

Historical Notes

Leonhard Euler (1707–1783) was Switzerland's foremost scientist and one of the three greatest mathematicians of modern times (Gauss and Riemann being the other two). He was perhaps the most prolific author of all time in any field. From 1727 to 1783 his writings poured out in a seemingly endless flood, constantly adding knowledge to every known branch of pure and applied mathematics, and also to many that were not known until he created them. He averaged about 800 printed pages a year throughout his long life, and yet he almost always had something worthwhile to say. The publication of his complete works was started in 1911, and the end is not in sight. This edition was planned



Leonhard Euler
1707–1783

to include 887 titles in 72 volumes, but since that time extensive new deposits of previously unknown manuscripts have been unearthed, and it is now estimated that more than 100 large volumes will be required for completion of the project. Euler evidently wrote mathematics with the ease and fluency of a skilled speaker discoursing on subjects with which he is intimately familiar. His writings are models of relaxed clarity. He never condensed, and he reveled in the rich abundance of his ideas and the vast scope of his interests. The French physicist Arago, in speaking of Euler's incomparable mathematical facility, remarked that "He calculated without apparent effort, as men breathe, or as eagles sustain themselves in the wind." He suffered total blindness during the last 17 years of his life, but with the aid of his powerful memory and fertile imagination, and with assistants to write his books and scientific papers from dictation, he actually increased his already prodigious output of work.

Euler was a native of Basel and a student of Johann Bernoulli at the University, but he soon outstripped his teacher. He was also a man of broad culture, well versed in the classical languages and literatures (he knew the Aeneid by heart), many modern languages, physiology, medicine, botany, geography, and the entire body of physical science as it was known in his time. His personal life was as placid and uneventful as is possible for a man with 13 children.

Though he was not himself a teacher, Euler has had a deeper influence on the teaching of mathematics than any other person. This came about chiefly through his three great treatises: *Introductio in Analysin Infinitorum* (1748); *Institutiones Calculi Differentialis* (1755); and *lnstitutiones Calculi Integralis* (1768–1794). There is considerable truth in the old saying that all elementary and advanced calculus textbooks since 1748 are essentially copies of Euler or copies of copies of Euler. These works summed up and codified the discoveries of his predecessors, and are full of Euler's own ideas. He extended and perfected plane and solid analytic geometry, introduced the analytic approach to trigonometry, and was responsible for the modern treatment of the functions $\ln x$ and e^x. He created a consistent theory of logarithms of negative and imaginary numbers, and discovered that $\ln x$ has an infinite number of values. It was through his work that the symbols e, π, and $i = \sqrt{-1}$ became common currency for all mathematicians, and it was he who linked them together in the astonishing relation $e^{i\pi} = -1$. Among his other contributions to standard mathematical notation were $\sin x$, $\cos x$, the use of $f(x)$ for an unspecified function, and the use of $\sum$ for summation.

His work in all departments of analysis strongly influenced the further development of this subject through the next two centuries. He contributed many important ideas to differential equations, including substantial parts of the theory of second-order linear equations and the method of solution by power series. He gave the first systematic discussion of the calculus of variations, which he founded on his basic differential equation for a minimizing curve. He discovered the integral defining the gamma function and developed many of its applications and special properties. He also worked with Fourier series, encountered the Bessel functions in his study of the vibrations of a stretched circular membrane, and applied Laplace transforms to solve differential equations—all before Fourier, Bessel, and Laplace were born.

E.T. Bell, the well-known historian of mathematics, observed that "One of the most remarkable features of Euler's universal genius was its equal strength in both of the main currents of mathematics, the continuous and the discrete." In the realm of the discrete, he was one of the originators of number theory and made many far-reaching contributions to this subject throughout his life. In addition, the origins of topology—one of the dominant forces in modern mathematics—lie in his solution of the Königsberg bridge problem and his formula $V - E + F = 2$ connecting the numbers of vertices, edges, and faces of a simple polyhedron.

The distinction between pure and applied mathematics did not exist in Euler's day, and for him the entire physical universe was a convenient object whose diverse phenomena offered scope for his methods of analysis. The foundations of classical mechanics had been laid down by Newton, but Euler was the principal architect. In his treatise of 1736 he was the first to explicitly introduce the concept of a mass-point, or particle, and he was also the first to study the acceleration of a particle moving along any curve and to use the notion of a vector in connection with velocity and acceleration. His continued successes in mathematical physics were so numerous, and his influence was so pervasive, that most of his discoveries are not credited to him at all and are taken for granted in the physics

community as part of the natural order of things. However, we do have Euler's angles for the rotation of a rigid body, and the all-important *Euler-Lagrange equation* of variational dynamics.
Euler was the Shakespeare of mathematics—universal, richly detailed, and inexhaustible.

The variational problem is a problem involving only the first functional derivative, or the **first variation**. We know from calculus that the first derivative by itself cannot determine the nature of the extremum. To test whether the point in question is maximum or minimum, we need all the second derivatives (see Example 6.6.9). One uses these derivatives to expand the functional in a Taylor series up to the second order. The sign of the second order contribution determines whether the functional is maximum or minimum at the extremal point. In analogy with Example 6.6.9, we expand $\mathbf{L}[u]$ about f up to the second-order derivative:

The first variation is not sufficient for a full knowledge of the nature of the extremum!

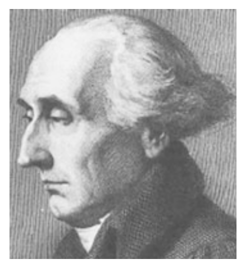
$$\mathbf{L}[u] = \mathbf{L}[f] + \int_\Omega d^p y \frac{\delta \mathbf{L}}{\delta u(\mathbf{y})}\bigg|_{u=f} \big(u(\mathbf{y}) - f(\mathbf{y})\big)$$
$$+ \frac{1}{2}\int_\Omega d^p y \int_\Omega d^p y' \frac{\delta^2 \mathbf{L}}{\delta u(\mathbf{y})\delta u(\mathbf{y}')}\bigg|_{u=f} \big(u(\mathbf{y}) - f(\mathbf{y})\big)\big(u(\mathbf{y}') - f(\mathbf{y}')\big).$$

The integrals have replaced the sums of the discrete case of Taylor expansion of the multivariable functions. Since we are interested in comparing u with the f that extremizes the functional, the second term vanishes and we get

$$\mathbf{L}[u] = \mathbf{L}[f] + \frac{1}{2}\int_\Omega d^p y \int_\Omega d^p y' \frac{\delta^2 \mathbf{L}}{\delta u(\mathbf{y})\delta u(\mathbf{y}')}\bigg|_{u=f}$$
$$\cdot \big[\big(u(\mathbf{y}) - f(\mathbf{y})\big)\big(u(\mathbf{y}') - f(\mathbf{y}')\big)\big]. \tag{33.15}$$

Historical Notes

Joseph Louis Lagrange (1736–1813) was born Giuseppe Luigi Lagrangia but adopted the French version of his name. He was the eldest of eleven children, most of whom did not reach adulthood. His father destined him for the law—a profession that one of his brothers later pursued—and Lagrange offered no objections. But having begun the study of physics and geometry, he quickly became aware of his talents and henceforth devoted himself to the exact sciences. Attracted first by geometry, at the age of seventeen he turned to analysis, then a rapidly developing field.



Joseph Louis Lagrange 1736–1813

In 1755, in a letter to the geometer Giulio da Fagnano, Lagrange speaks of one of Euler's papers published at Lausanne and Geneva in 1744. The same letter shows that as early as the end of 1754 Lagrange had found interesting results in this area, which was to become the *calculus of variations* (a term coined by Euler in 1766). In the same year, Lagrange sent Euler a summary, written in Latin, of the purely analytical method that he used for this type of problem. Euler replied to Lagrange that he was very interested in the technique. Lagrange's merit was likewise recognized in Turin; and he was named, by a royal decree, professor at the Royal Artillery School with an annual salary of 250 crowns—a sum never increased in all the years he remained in his native country. Many years later, in a letter to d'Alembert, Lagrange confirmed that this method of maxima and minima was the first fruit of his studies—he was only nineteen when he devised it—and that he regarded it as his best work in mathematics. In 1756, in a letter to Euler that has been lost, Lagrange, applying the calculus of variations to mechanics, generalized Euler's earlier work on the trajectory described by a material point subject to the influence of central forces to an arbitrary system of bodies, and derived from it a procedure for solving all the problems of dynamics.

In 1757 some young Turin scientists, among them Lagrange, founded a scientific society that was the origin of the Royal Academy of Sciences of Turin. One of the main goals of this society was the publication of a miscellany in French and Latin, *Miscellanea Taurinensia ou Mélanges de Turin*, to which Lagrange contributed fundamentally. These contributions included works on the calculus of variations, probability, vibrating strings, and the principle of least action.

To enter a competition for a prize, in 1763 Lagrange sent to the Paris Academy of Sciences a memoir in which he provided a satisfactory explanation of the translational motion of the moon. In the meantime, the Marquis Caraccioli, ambassador from the kingdom of Naples to the court of Turin, was transferred by his government to London. He took along the young Lagrange, who until then seems never to have left the immediate vicinity of Turin. Lagrange was warmly received in Paris, where he had been preceded by his memoir on lunar libration. He may perhaps have been treated too well in the Paris scientific community, where austerity was not a leading virtue. Being of a delicate constitution, Lagrange fell ill and had to interrupt his trip. In the spring of 1765 Lagrange returned to Turin by way of Geneva.

In the autumn of 1765 d'Alembert, who was on excellent terms with Frederick II of Prussia, and familiar with Lagrange's work through *Mélanges de Turin*, suggested to Lagrange that he accept the vacant position in Berlin created by Euler's departure for St. Petersburg. It seems quite likely that Lagrange would gladly have remained in Turin had the court of Turin been willing to improve his material and scientific situation. On 26 April, d'Alembert transmitted to Lagrange the very precise and advantageous propositions of the king of Prussia. Lagrange accepted the proposals of the Prussian king and, not without difficulties, obtained his leave through the intercession of Frederick II with the king of Sardinia. Eleven months after his arrival in Berlin, Lagrange married his cousin Vittoria Conti who died in 1783 after a long illness. With the death of Frederick II in August 1786 he also lost his strongest support in Berlin. Advised of the situation, the princes of Italy zealously competed in attracting him to their courts. In the meantime the French government decided to bring Lagrange to Paris through an advantageous offer. Of all the candidates, Paris was victorious.

Lagrange left Berlin on 18 May 1787 to become *pensionnaire vétéran* of the Paris Academy of Sciences, of which he had been a foreign associate member since 1772. Warmly welcomed in Paris, he experienced a certain lassitude and did not immediately resume his research. Yet he astonished those around him by his extensive knowledge of metaphysics, history, religion, linguistics, medicine, and botany.

In 1792 Lagrange married the daughter of his colleague at the Academy, the astronomer Pierre Charles Le Monnier. This was a troubled period, about a year after the flight of the king and his arrest at Varennes. Nevertheless, on 3 June the royal family signed the marriage contract "as a sign of its agreement to the union." Lagrange had no children from this second marriage, which, like the first, was a happy one.

When the academy was suppressed in 1793, many noted scientists, including Lavoisier, Laplace, and Coulomb were purged from its membership; but Lagrange remained as its chairman. For the next ten years, Lagrange survived the turmoil of the aftermath of the French Revolution, but by March of 1813, he became seriously ill. He died on the morning of 11 April 1813, and three days later his body was carried to the Panthéon. The funeral oration was given by Laplace in the name of the Senate.

A straight line segment is indeed the *shortest* distance between two points.

Example 33.1.14 Let us apply Eq. (33.15)to the extremal function of Example 33.1.13 to see if the line is truly the *shortest* distance between two points. The first functional derivative, obtained using Eq. (33.9), is simply $\mathbb{E}(L)$:

$$\frac{\delta \mathbf{L}}{\delta u(y)} = \mathbb{E}(L) = -\frac{u_{yy}}{(1+u_y^2)^{3/2}}.$$

To find the second variational derivative, we use the basic relations (33.6), (33.7), and the chain rule (33.10):

$$\begin{aligned}
&\frac{\delta^2 \mathbf{L}}{\delta u(y')\delta u(y)}\bigg|_{u=f} \\
&\quad = -\frac{\delta}{\delta u(y')}\left[\frac{u_{yy}}{(1+u_y^2)^{3/2}}\right]\bigg|_{u=f} \\
&\quad = -\left\{\left(1+u_y^2\right)^{-3/2}\frac{\delta u_{yy}}{\delta u(y')} - u_{yy}\frac{3}{2}\left(1+u_y^2\right)^{-5/2} 2u_y \frac{\delta u_y}{\delta u(y')}\right\}\bigg|_{u=f} \\
&\quad = -\frac{\delta''(y-y')}{(1+u_y^2)^{3/2}}\bigg|_{u=f} = -\frac{\delta''(y-y')}{(1+c_1^2)^{3/2}},
\end{aligned}$$

because $u_{yy} = 0$ and $u_y = c_1$ when $u = f$. Inserting this in Eq. (33.15), we obtain

$$\begin{aligned}
\mathbf{L}[u] &= \mathbf{L}[f] - \frac{1}{2(1+c_1^2)^{3/2}} \\
&\quad \times \int_{a_1}^{a_2} dy \int_{a_1}^{a_2} dy' \delta''\left(y-y'\right)\big(u(y)-f(y)\big)\big(u\left(y'\right)-f\left(y'\right)\big) \\
&= \mathbf{L}[f] - \frac{1}{2(1+c_1^2)^{3/2}} \int_{a_1}^{a_2} dy\big(u(y)-f(y)\big)\frac{d^2}{dy^2}\big(u(y)-f(y)\big).
\end{aligned}$$

The last integral can be integrated by parts, with the result

$$\underbrace{\big(u(y)-f(y)\big)\frac{d}{dy}\big(u(y)-f(y)\big)\bigg|_{a_1}^{a_2}}_{=0 \text{ because } u(a_i)=f(a_i),\, i=1,2} - \int_{a_1}^{a_2} dy\left[\frac{d}{dy}\big(u(y)-f(y)\big)\right]^2.$$

Therefore,

$$\mathbf{L}[u] = \mathbf{L}[f] + \frac{1}{2(1+c_1^2)^{3/2}} \underbrace{\int_{a_1}^{a_2} dy\left[\frac{d}{dy}\big(u(y)-f(y)\big)\right]^2}_{\text{always positive}}.$$

It follows that $\mathbf{L}[f] < \mathbf{L}[u]$, i.e., that f indeed gives the shortest distance.

Example 33.1.15 In the special theory of relativity, the element of the invariant "length", or proper time, is given by $\sqrt{dt^2 - dx^2}$. Thus, the total proper time between two events (t_1, a_1) and (t_2, a_2) is given by

$$\mathbf{L}[x] = \int_{t_1}^{t_2} \sqrt{1 - x_t^2}\, dt, \qquad x_t \equiv \frac{dx}{dt}.$$

The extremum of this variational problem is exactly the same as in the previous example, the only difference being a sign. In fact, the reader may verify that

$$\frac{\delta \mathbf{L}[x]}{\delta x(s)} = \mathbb{E}(L) = \frac{x_{ss}}{(1 - x_s^2)^{3/2}},$$

connection between variational problem and the twin paradox

and therefore, $x = f(t) = c_1 t + c_2$ extremizes the functional. The second variational derivative can be obtained as before. It is left for the reader to show that in the case at hand, $\mathbf{L}[f] > \mathbf{L}[x]$, i.e., that f gives the *longest* proper time. Since the function $f(t) = c_1 t + c_2$ corresponds to an inertial (unaccelerated) observer, we conclude that

> **Box 33.1.16** *Accelerated observers measure a shorter proper time between any two events than inertial observers.*

This is the content of the famous **twin paradox**, in which the twin who goes to a distant galaxy and comes back (therefore being accelerated) will return younger than her (unaccelerated) twin.

33.1.4 Divergence and Null Lagrangians

The variational problem integrates a Lagrangian over a region Ω of $\mathbb{R}^p$. If the Lagrangian happens to be the divergence of a function that vanishes at the boundary of Ω, the variational problem becomes trivial, because all functions will extremize the functional. We now study such Lagrangians in more detail.

total divergence

Definition 33.1.17 Let $\{F_i : M^{(n)} \to \mathbb{R}\}_{i=1}^p$ be functions on $M^{(n)}$, and $\mathbf{F} = (F_1, \ldots, F_p)$. The **total divergence** of $\mathbf{F}$ is defined to be[2]

$$\mathbf{D} \cdot \mathbf{F} \equiv \sum_{j=1}^{p} D_j F_j,$$

where D_j is the total derivative with respect to x^j.

Now suppose that the Lagrangian $L(x, u^{(n)})$ can be written as the divergence of some p-tuple $\mathbf{F}$. Then by the divergence theorem,

$$\mathbf{L}[u] = \int_\Omega L\big(x, u^{(n)}\big)\, d^p x = \int_\Omega \mathbf{D} \cdot \mathbf{F}\, d^p x = \int_{\partial\Omega} \mathbf{F} \cdot d\mathbf{a}$$

for any $u = f(x)$ and any domain Ω. It follows that $\mathbf{L}[f]$ depends on the behavior of f only at the boundary. Since in a typical problem no variation takes place at the boundary, all functions that satisfy the boundary conditions will be solutions of the variational problem, i.e., they satisfy the Euler-Lagrange equation. Lagrangians that satisfy the Euler-Lagrange equation for all u and x are called **null Lagrangians**. It turns out that null Lagrangians are the *only* such solutions of the Euler-Lagrange equation (for a proof, see [Olve 86, pp. 252–253]).

null Lagrangians

[2] The reader need not be concerned about lack of consistency in the location of indices (upper vs. lower), because we are dealing with indexed objects, such as F_i, which are *not* tensors!

Theorem 33.1.18 *A function $L(x, u^{(n)})$ satisfies $\mathbb{E}(L) \equiv 0$ for all x and u if and only if $L = \mathbf{D} \cdot \mathbf{F}$ for some p-tuple of functions $\mathbf{F} = (F_1, \ldots, F_p)$ of x, u, and the derivatives of u.*

In preparation for the investigation of symmetries of the variational problems, we look into the effect of a change of variables on the variational problem and the Euler operator. This is important, because the variational problem should be independent of the variables chosen. Let

$$\tilde{x} = \Psi(x, u), \qquad \tilde{u} = \Phi(x, u) \tag{33.16}$$

be any change of variables. Then by prolongation, we also have $\tilde{u}^{(n)} = \Phi^{(n)}(x, u^{(n)})$ for the derivatives. Substituting $u = f(x)$ and all its prolongations in terms of the new variables, the functional

$$\mathbf{L}[f] = \int_\Omega L\big(x, \mathrm{pr}^{(n)} f(x)\big)\, d^p x$$

will be transformed into

$$\tilde{\mathbf{L}}[\tilde{f}] = \int_{\tilde{\Omega}} \tilde{L}\big(\tilde{x}, \mathrm{pr}^{(n)} \tilde{f}(\tilde{x})\big)\, d^p \tilde{x},$$

where the transformed domain, defined by

$$\tilde{\Omega} = \big\{\tilde{x} = \Psi\big(x, f(x)\big) \mid x \in \Omega\big\},$$

will depend not only on the original domain Ω, but also on the function f. The new Lagrangian is then related to the old one by the change of variables formula for multiple integrals:

$$L\big(x, \mathrm{pr}^{(n)} f(x)\big) = \tilde{L}\big(\tilde{x}, \mathrm{pr}^{(n)} \tilde{f}(\tilde{x})\big) \det \mathsf{J}\big(x, \mathrm{pr}^{(1)} f(x)\big), \tag{33.17}$$

where J is the Jacobian matrix of the change of variables induced by the function f.

Starting with Eqs. (33.16) and (33.17), one can obtain the transformation formula for the Euler operator stated below. The details can be found in [Olve 86, pp. 254–255].

Theorem 33.1.19 *Let $L(x, u^{(n)})$ and $\tilde{L}(\tilde{x}, \tilde{u}^{(n)})$ be two Lagrangians related by the change of variable formulas (33.16) and (33.17). Then*

$$\mathbb{E}_\alpha(L) = \sum_{\beta=1}^{q} F_{\alpha\beta}\big(x, u^{(1)}\big)\tilde{\mathbb{E}}_\beta(\tilde{L}), \quad \alpha = 1, \ldots, q$$

where $\tilde{\mathbb{E}}_\beta$ is the Euler operator associated with the new variables, and

$$F_{\alpha\beta} \equiv \det \begin{pmatrix} D_1\Psi^1 & \ldots & D_p\Psi^1 & \partial\Psi^1/\partial u^\alpha \\ \vdots & & \vdots & \vdots \\ D_1\Psi^p & \ldots & D_p\Psi^p & \partial\Psi^p/\partial u^\alpha \\ D_1\Phi^\beta & \ldots & D_p\Phi^\beta & \partial\Phi^\beta/\partial u^\alpha \end{pmatrix}.$$

33.2 Symmetry Groups of Variational Problems

In the theory of fields, as well as in mechanics, condensed matter theory, and statistical mechanics, the starting point is usually a Lagrangian. The variational problem of this Lagrangian gives the classical equations of motion, and its symmetries lead to the important conservation laws.

variational symmetry group

Definition 33.2.1 A local group of transformations G acting on $M \subset \Omega_0 \times U$ is a **variational symmetry group** of the functional

$$\mathbf{L}[u] = \int_{\Omega_0} L\big(x, u^{(n)}\big)\, d^p x \tag{33.18}$$

if whenever (the closure of) Ω lies in Ω_0, f is a function over Ω whose graph is in M, and $g \in G$ is such that $\tilde{f} = g \cdot f$ is a single-valued function defined over $\tilde{\Omega}$, then

$$\int_{\tilde{\Omega}} L\big(\tilde{x}, \mathrm{pr}^{(n)} \tilde{f}(\tilde{x})\big)\, d^p \tilde{x} = \int_{\Omega} L\big(x, \mathrm{pr}^{(n)} f(x)\big)\, d^p x. \tag{33.19}$$

"Symmetry of the Lagrangian" is really the symmetry group of the variational problem!

In the physics community, the symmetry group of the variational problem is (somewhat erroneously) called the **symmetry of the Lagrangian**. Note that if we had used $\tilde{L}$ in the LHS of Eq. (33.19), we would have obtained an identity valid for *all* Lagrangians because of Eq. (33.17) and the formula for the change in the volume element of integration. Only *symmetric Lagrangians* will satisfy Eq. (33.19).

As we have experienced so far, the action of a group can be very complicated and very nonlinear. On the other hand, the *infinitesimal* action simplifies the problem considerably. Fortunately, we have (see [Olve 86, pp. 257–258] for a proof).

Theorem 33.2.2 *A local group of transformations G acting on $M \subset \Omega_0 \times U$ is a variational symmetry group of the functional* (33.18) *if and only if*

$$\mathrm{pr}^{(n)}\mathbf{v}(L) + L\mathbf{D} \cdot \mathbf{X} = 0 \tag{33.20}$$

for all $(x, u^{(n)}) \in M^{(n)}$ and every infinitesimal generator

$$\mathbf{v} = \sum_{i=1}^{p} X^i(x, u) \frac{\partial}{\partial x^i} + \sum_{\alpha=1}^{q} U^\alpha(x, u) \frac{\partial}{\partial u^\alpha}$$

of G, where $\mathbf{X} \equiv (X^1, \ldots, X^p)$.

Example 33.2.3 Consider the case of $p = 1 = q$, and assume that the Lagrangian is independent of x but depends on $u \in \mathcal{L}^2(a, b)$ and its first derivative. Then the variational problem takes the form

$$\mathbf{L}[u] = \int_a^b L\big(u^{(1)}\big)\, dx \equiv \int_a^b L(u, u_x)\, dx.$$

Since derivatives are independent of translations, we expect translations to be part of the symmetry group of this variational problem. Let us verify this. The infinitesimal generator of translation is ∂_x, which is its own prolongation. Therefore, with $X = 1$ and $U = 0$, it follows that

$$\mathrm{pr}^{(1)}\mathbf{v}(L) + L\mathbf{D}\cdot\mathbf{X} = \partial_x L + L D_x X = 0 + 0 = 0.$$

Example 33.2.4 As a less trivial case, consider the proper time of Example 33.1.15. Lorentz transformations generated by[3] $\mathbf{v} = u\partial_x + x\partial_u$ are symmetries of that variational problem. We can verify this by noting that the first prolongation of $\mathbf{v}$ is, as the reader is urged to verify,

$$\mathrm{pr}^{(1)}\mathbf{v} = \mathbf{v} + \left(1 - u_x^2\right)\frac{\partial}{\partial u_x}.$$

Therefore,

$$\mathrm{pr}^{(1)}\mathbf{v}(L) = 0 + 0 + \left(\left(1 - u_x^2\right)\right)\frac{1}{2}(-2u_x)\frac{1}{\sqrt{1 - u_x^2}} = -u_x\sqrt{1 - u_x^2}.$$

On the other hand, since $X = u$ and $U = x$,

$$L D_x(X) = \sqrt{1 - u_x^2}\, D_x(u) = \sqrt{1 - u_x^2}\, u_x,$$

so that Eq. (33.20) is satisfied.

In the last chapter, we studied the symmetries of the DEs in some detail. This chapter introduces us to a particular DE that arises from a variational problem, namely, the Euler-Lagrange equation. The natural question to ask now is: How does the variational symmetry manifest itself in the Euler-Lagrange equation? Barring some technical difficulties, we note that for any change of variables, if $u = f(x)$ is an extremal of the variational problem $\mathbf{L}[u]$, then $\tilde{u} = \tilde{f}(\tilde{x})$ is an extremal of the variational problem $\tilde{\mathbf{L}}[\tilde{u}]$. In particular, if the change is achieved by the action of the variational symmetry group, $(\tilde{x}, \tilde{u}) = g \cdot (x, u)$ for some $g \in G$, then $\tilde{\mathbf{L}}[\tilde{u}] = \mathbf{L}[\tilde{u}]$, and $g \cdot f$ is also an extremal of $\mathbf{L}$. We thus have

Theorem 33.2.5 *If G is the variational symmetry group of a functional, then G is also the symmetry group of the associated Euler-Lagrange equations.*

The converse is *not* true! There are symmetry groups of the Euler-Lagrange equations that are not the symmetry group of the variational problem. Problem 33.8 illustrates this for $p = 3$, $q = 1$, and the functional

$$\mathbf{L}[u] = \frac{1}{2}\iiint \left(u_t^2 - u_x^2 - u_y^2\right) dx\, dy\, dt, \qquad (33.21)$$

[3] In order to avoid confusion in applying formula (33.20), we use x (instead of t) as the independent variable and u (instead of x) as the dependent variable.

Symmetries of the Euler-Lagrange equations are not necessarily the symmetries of the corresponding variational problem!

whose Euler-Lagrange equation is the wave equation. The reader is asked to show that while the rotations and Lorentz boosts of Table 32.3 are variational symmetries, the dilatations and inversions (special conformal transformations) are not.

We now treat the case of $p = 1 = q$, whose Euler-Lagrange equation is an ODE. Recall that the knowledge of a symmetry group of an ODE led to a reduction in the order of that ODE. Let us see what happens in the present case. Suppose $\mathbf{v} = X\partial_x + U\partial_u$ is the infinitesimal generator of a 1-parameter group of variational symmetries of $\mathbf{L}$. By an appropriate coordinate transformation from (x, u) to (y, w), as in Sect. 32.5, $\mathbf{v}$ will reduce to $\partial/\partial w$, whose prolongation is also $\partial/\partial w$. In terms of the new coordinates, Eq. (33.20) will reduce to $\partial\tilde{L}/\partial w = 0$; i.e., the new Lagrangian is independent of w, and the Euler-Lagrange equation (33.14) becomes

$$0 = \mathbb{E}(L) = \sum_{j=1}^{n}(-1)^j D_y^j \frac{\partial\tilde{L}}{\partial w_j} = (-D_y)\left[\sum_{j=0}^{n-1}(-D_y)^j \frac{\partial\tilde{L}}{\partial w_{j+1}}\right]. \quad (33.22)$$

Therefore, the expression in the brackets is some constant λ (because D_y is a *total* derivative). Furthermore, if we introduce $v = w_y$, the expression in the brackets becomes the Euler-Lagrange equation of the variational problem

$$\hat{\mathbf{L}}[v] = \int \hat{L}\big(y, v^{(n-1)}\big)\,dy, \quad \text{where} \quad \hat{L}\big(y, v^{(n-1)}\big) = \tilde{L}(y, w_y, \ldots, w_n),$$

and every solution $w = f(y)$ of the original $(2n)$th-order Euler-Lagrange equation corresponds to the $(2n-2)$nd-order equation

$$\hat{\mathbb{E}}(\hat{L}) = \frac{\partial\hat{L}}{\partial y} + \sum_{j=1}^{n-1}(-D_y)^j \frac{\partial\hat{L}}{\partial v_j} = \lambda. \quad (33.23)$$

Moreover, this equation can be written as the Euler-Lagrange equation for

$$\hat{\mathbf{L}}_\lambda[v] = \int \big[\hat{L}\big(y, v^{(n-1)}\big) - \lambda v\big]\,dy,$$

Lagrange multiplier

and λ can be thought of as a **Lagrange multiplier**, so that in analogy with the multivariable extremal problem,[4] the extremization of $\hat{\mathbf{L}}_\lambda[v]$ becomes equivalent to that of $\hat{\mathbf{L}}[v]$ subject to the constraint $\int v\,dy = 0$. We summarize the foregoing discussion in the following theorem.

Theorem 33.2.6 *Let $p = 1 = q$, and $\mathbf{L}[u]$ an nth-order variational problem with a 1-parameter group of variational symmetries G. Then there exists a one-parameter family of variational problems $\hat{\mathbf{L}}_\lambda[v]$ of order $n-1$ such that every solution of the Euler-Lagrange equation for $\mathbf{L}[u]$ can be found by integrating the solutions to the Euler-Lagrange equation for $\hat{\mathbf{L}}_\lambda[v]$.*

[4] See [Math 70, pp. 331–341] for a discussion of Lagrange multipliers and their use in variational techniques, especially those used in approximating solutions of the Schrödinger equation.

Thus, we have the following important result:

Box 33.2.7 *A* 1*-parameter variational symmetry of a functional reduces the order of the corresponding Euler-Lagrange equation* by two.

This conclusion is to be contrasted with the symmetry of ODEs, where each 1-parameter group of symmetries reduces the order of the ODE by 1. It follows from Box 33.2.7 that the ODEs of order $2n$ derived from a variational problem—the Euler-Lagrange equation—are special.

Example 33.2.8 A first-order variational problem with a 1-parameter group of symmetries can be integrated out. By transforming to a new coordinate system, we can always assume that the Lagrangian is independent of the *dependent variable* (see Proposition 32.5.1). The Euler-Lagrange equation in this case becomes

$$0 = \mathbb{E}(L) = \underbrace{\frac{\partial L}{\partial u}}_{=0} - D_x \frac{\partial L}{\partial u_x} \quad \Rightarrow \quad \frac{\partial L}{\partial u_x}(x, u_x) = \lambda.$$

Solving this implicit relation, we get $u_x = F(x, \lambda)$, which can be integrated to give u as a function of x (and λ).

The procedure can be generalized to r-parameter symmetry groups, but the order cannot be expected to be reduced by 2 unless the group is abelian. We shall not pursue this matter here, but ask the reader to refer to Problem 33.9.

33.3 Conservation Laws and Noether's Theorem

A conserved physical quantity is generally defined as a quantity whose flux through any arbitrary closed surface is equal to (the negative of) the rate of depletion of the quantity in the volume enclosed. This statement, through the use of the divergence theorem, translates into a relation connecting the time rate of change of the density and the divergence of the current corresponding to the physical quantity. Treating time and space coordinates as independent variables and extending to p independent variables, we have the following:

current density and conservation law

Definition 33.3.1 A **conservation law** for a system of differential equations $\Delta(x, u^{(n)}) = 0$ is a divergence expression $\mathbf{D} \cdot \mathbf{J} = 0$ valid for all solutions $u = f(x)$ of the system. Here,

$$\mathbf{J} \equiv \bigl(J_1\bigl(x, u^{(n)}\bigr), J_2\bigl(x, u^{(n)}\bigr), \ldots, J_p\bigl(x, u^{(n)}\bigr)\bigr)$$

is called **current density**.

constant of the motion, or first integral of a system of ODEs

For $p = 1 = q$, i.e., for a system of ODEs, a conservation law takes the form $D_x J(x, u^{(n)}) = 0$ for all solutions $u = f(x)$ of the system. This requires $J(x, u^{(n)})$ to be a constant, i.e., that $J(x, u^{(n)})$ be a **constant of the motion**, or, as it is sometimes called, the *first integral* of the system.

In order to understand conservation laws, we need to get a handle on those conservation laws that are trivially satisfied.

trivial conservation law of the first kind

Definition 33.3.2 If the current density $\mathbf{J}$ itself vanishes for all solutions $u = f(x)$ of the system $\Delta(x, u^{(n)}) = 0$, then $\mathbf{D} \cdot \mathbf{J} = 0$ is called a **trivial conservation law of the first kind**.

To eliminate this kind of triviality, one solves the system and its prolongations $\Delta^{(k)}(x, u^{(n)}) = 0$ for some of the variables u_J^α in terms of the remaining variables and substitutes the latter whenever they occur. For example, one can differentiate the evolution equation $u_t = F(x, u^{(n)})$—in which $u^{(n)}$ have derivatives with respect to x only—with respect to t and x sufficient number of times (this is what is meant by "prolongation" of the system of equations) and solve for all derivatives of u involving time. Then, in the conservation law, substitute for any such derivatives to obtain a conservation law involving only x derivatives of u.

Example 33.3.3 The current density $\mathbf{J}_1 = (\frac{1}{2}u_t^2 + \frac{1}{2}u_x^2, -u_t u_x)$ is easily seen to be conserved for the system of first-order DEs

$$u_t = v_x, \qquad u_x = v_t.$$

By eliminating all the time derivatives in $\mathbf{J}_1$, we obtain $\mathbf{J}_2 = (\frac{1}{2}u_x^2 + \frac{1}{2}v_x^2, -u_x v_x)$, which is also conserved. However, the difference between these two currents,

$$\mathbf{J} = \mathbf{J}_1 - \mathbf{J}_2 = \left(\frac{1}{2}u_t^2 - \frac{1}{2}v_x^2, u_x v_x - u_t u_x\right),$$

satisfies a trivial conservation law of the first kind, because the components of $\mathbf{J}$ vanish on the solutions of the system.

trivial conservation law of the second kind; null divergence

Definition 33.3.4 If the current density $\mathbf{J}$ satisfies $\mathbf{D} \cdot \mathbf{J} = 0$ for *all* functions $u = f(x)$, even if they are not solutions of the system of DEs, the divergence identity is called a **trivial conservation law of the second kind**. In this case $\mathbf{J}$ is called a **null divergence**.

If we treat J_i as the components of a $(p-1)$-form $\boldsymbol{\omega}$, so that the exterior derivative $d\boldsymbol{\omega}$ is the divergence of $\mathbf{J}$ (times a volume element), then the triviality of the conservation law for $\mathbf{J}$ is equivalent to the fact that $\boldsymbol{\omega}$ is closed. By the converse of the Poincaré lemma, there must be a $(p-2)$-form $\boldsymbol{\eta}$ such that $\boldsymbol{\omega} = d\boldsymbol{\eta}$. In the context of this chapter, we have the following theorem.

Theorem 33.3.5 *Suppose* $\mathbf{J} = (J_1(x, u^{(n)}), \dots, J_p(x, u^{(n)}))$ *is a* p*-tuple of functions on* $X \times U^{(n)}$. *Then* $\mathbf{J}$ *is a null divergence if and only if there exist smooth functions* $A_{kj}(x, u^{(n)})$, $j, k = 1, \dots, p$, *antisymmetric in their indices, such that*

$$J_k = \sum_{j=1}^{p} D_j A_{kj}, \quad j = 1, \dots, p. \tag{33.24}$$

Definition 33.3.6 We say that $\mathbf{D} \cdot \mathbf{J} = 0$ is a **trivial conservation law** if there exist antisymmetric smooth functions $A_{kj}(x, u^{(n)})$ satisfying Eq. (33.24) for all solutions of the system of DEs $\Delta(x, u^{(n)}) = 0$. Two conservation laws are **equivalent** if they differ by a trivial conservation law.

trivial conservation law and the equivalence of two conservation laws

We shall not distinguish between conservation laws that are equivalent. It turns out that to within this equivalence, some systems of DEs Δ_ν have current densities $\mathbf{J}$ such that

$$\mathbf{D} \cdot \mathbf{J} = \sum_{\nu=1}^{l} Q_\nu \Delta_\nu \quad \text{for some } l\text{-tuple } \mathbf{Q} = (Q_1, \dots, Q_l), \tag{33.25}$$

where $\{Q_\nu\}$ are smooth functions of x, u, and all derivatives of u.

Definition 33.3.7 Equation (33.25) is called the **characteristic form** of the conservation law for the current density $\mathbf{J}$, and the l-tuple $\mathbf{Q}$, the **characteristic** of the conservation law.

characteristic form of a conservation law

We are now in a position to prove the celebrated Noether's theorem. However, we first need a lemma.

Lemma 33.3.8 *Let* $\mathbf{v} = \sum_{i=1}^{p} X^i \partial/\partial x^i + \sum_{\alpha=1}^{q} U^\alpha \partial/\partial u^\alpha$ *where* X^i *and* U^α *are functions of* x *and* u. *Let*

$$Q^\alpha\bigl(x, u^{(1)}\bigr) \equiv U^\alpha(x, u) - \sum_{i=1}^{p} X^i(x, u) u_i^\alpha, \quad \alpha = 1, \dots, q.$$

Then

$$\mathrm{pr}^{(n)} \mathbf{v} = \mathrm{pr}^{(n)} \mathbf{v}_Q + \sum_{i=1}^{p} X^i D_i, \tag{33.26}$$

where

$$\mathbf{v}_Q \equiv \sum_{\alpha=1}^{q} Q^\alpha\bigl(x, u^{(1)}\bigr) \frac{\partial}{\partial u^\alpha}, \qquad \mathrm{pr}^{(n)} \mathbf{v}_Q \equiv \sum_{\alpha=1}^{q} \sum_{J} D_J Q^\alpha \frac{\partial}{\partial u_J^\alpha}.$$

The sum over J *extends over all multi-indices with* $0 \le |J| \le n$, *with the* $|J| = 0$ *term being simply* $\mathbf{v}_Q$.

Proof Substitute Q^α in the definition of U_J^α as given in Theorem 32.3.5 to obtain

$$U_J^\alpha = D_J Q^\alpha + \sum_{i=1}^p X^i u_{J,i}^\alpha,$$

where $U_0^\alpha = Q^\alpha + \sum_{i=1}^p X^i u_i^\alpha = U^\alpha$. It follows that (with $J = 0$ included in the sum)

$$\begin{aligned}\mathrm{pr}^{(n)}\mathbf{v} &= \sum_{i=1}^p X^i \frac{\partial}{\partial x^i} + \sum_J \left[\sum_{\alpha=1}^q D_J Q^\alpha + \sum_{i=1}^p X^i u_{J,i}^\alpha\right]\frac{\partial}{\partial u_J^\alpha} \\ &= \sum_{\alpha=1}^q \sum_J D_J Q^\alpha \frac{\partial}{\partial u_J^\alpha} + \sum_{i=1}^p X^i \underbrace{\left[\frac{\partial}{\partial x^i} + \sum_{\alpha=1}^q \sum_J u_{J,i}^\alpha \frac{\partial}{\partial u_J^\alpha}\right]}_{=D_i \text{ by Proposition 32.3.4}},\end{aligned}$$

and the lemma is proved. □

The celebrated Noether's theorem connecting symmetries to conservation laws

Theorem 33.3.9 (Noether's theorem) *Let*

$$\mathbf{v} = \sum_{i=1}^p X^i \partial/\partial x^i + \sum_{\alpha=1}^q U^\alpha \partial/\partial u^\alpha$$

be the infinitesimal generator of a local 1*-parameter group of symmetries* G *of the variational problem* $\mathbf{L}[u] = \int L(x, u^{(n)})\, d^p x$. *Let*

$$Q^\alpha\left(x, u^{(1)}\right) \equiv U^\alpha(x,u) - \sum_{i=1}^p X^i(x,u) u_i^\alpha, \quad u_i^\alpha = \frac{\partial u^\alpha}{\partial x^i}.$$

Then there exists a p*-tuple* $\mathbf{J} = (J_1, \ldots, J_p)$ *such that*

$$\mathbf{D}\cdot\mathbf{J} = \sum_{\alpha=1}^q Q^\alpha \mathbb{E}_\alpha(L) \tag{33.27}$$

is a conservation law in characteristic form for the Euler-Lagrange equation $\mathbb{E}_\alpha(L) = 0$.

Proof We use Lemma 33.3.8 in the infinitesimal criterion of the variational symmetry (33.20) to obtain

$$\begin{aligned}0 &= \mathrm{pr}^{(n)}\mathbf{v}(L) + L\mathbf{D}\cdot\mathbf{X} \\ &= \mathrm{pr}^{(n)}\mathbf{v}_Q(L) + \sum_{i=1}^p X^i D_i L + L\sum_{i=1}^p D_i X^i \\ &= \mathrm{pr}^{(n)}\mathbf{v}_Q(L) + \sum_{i=1}^p D_i\left(LX^i\right) = \mathrm{pr}^{(n)}\mathbf{v}_Q(L) + \mathbf{D}\cdot(L\mathbf{X}).\end{aligned} \tag{33.28}$$

Using the definition of $\mathrm{pr}^{(n)}\mathbf{v}_Q$ and the identity

$$(D_j S)T = D_j(ST) - SD_j T,$$

we can commute $D_J = D_{j_1} \cdots D_{j_k}$ past Q^α one factor at a time, each time introducing a divergence. Therefore,

$$\begin{aligned}\mathrm{pr}^{(n)}\mathbf{v}_Q(L) &= \sum_{\alpha,J} D_J Q^\alpha \frac{\partial L}{\partial u_J^\alpha} = \sum_{\alpha,J} Q^\alpha (-D)_J \frac{\partial L}{\partial u_J^\alpha} + \mathbf{D}\cdot\mathbf{A} \\ &= \sum_{\alpha=1}^{q} Q^\alpha \mathbb{E}_\alpha(L) + \mathbf{D}\cdot\mathbf{A},\end{aligned}$$

where $\mathbf{A} = (A_1, \ldots, A_p)$ is some p-tuple of functions depending on L, the Q^α's, and their derivatives, whose precise form is not needed here. Combining this with Eq. (33.28), we obtain

$$0 = \sum_{\alpha=1}^{q} Q^\alpha \mathbb{E}_\alpha(L) + \mathbf{D}\cdot(\mathbf{A} + L\mathbf{X}).$$

Selecting $\mathbf{J} = -(\mathbf{A} + L\mathbf{X})$ proves the theorem. □

33.4 Application to Classical Field Theory

It is clear from the proof of Noether's theorem that if we are interested in the conserved current, we need to find $\mathbf{A}$. In general, the expression for $\mathbf{A}$ is very complicated. However, if the variational problem is of first order (which in most cases of physical interest it is), then we can easily find the explicit form of $\mathbf{A}$, and, consequently the conserved current $\mathbf{J}$. We leave it for the reader to prove the following:

Corollary 33.4.1 *Let $\mathbf{v} = \sum_{i=1}^{p} X^i \partial/\partial x^i + \sum_{\alpha=1}^{q} U^\alpha \partial/\partial u^\alpha$ be the infinitesimal generator of a local 1-parameter group of symmetries G of the first-order variational problem $\mathbf{L}[u] = \int L(x, u^{(1)})\, d^p x$. Then*[5]

$$J_i = \sum_{\alpha=1}^{q}\sum_{j=1}^{p} X^j u_j^\alpha \frac{\partial L}{\partial u_i^\alpha} - \sum_{\alpha=1}^{q} U^\alpha \frac{\partial L}{\partial u_i^\alpha} - X^i L, \quad i = 1, \ldots, p$$

form the components of a conserved current for the Euler-Lagrange equation $\mathbb{E}_\alpha(L) = 0$.

Historical Notes

Amalie Emmy Noether (1882–1935), generally considered the greatest of all female mathematicians up to her time, was the eldest child of Max Noether, research mathematician and professor at the University of Erlangen, and Ida Amalia Kaufmann. Two of

[5]We have multiplied J_i by a negative sign to conform to physicists' convention.

Amalie Emmy Noether
1882–1935

Emmy's three brothers were also scientists. Alfred, her junior by a year, earned a doctorate in chemistry at Erlangen. Fritz, two and a half years younger, became a distinguished physicist; and his son, Gottfried, became a mathematician.

At first Emmy Noether had planned to be a teacher of English and French. From 1900 to 1902 she studied mathematics and foreign languages at Erlangen. Then in 1903 she started her specialization in mathematics at the University of Göttingen. At both universities she was a nonmatriculated auditor at lectures, since at the turn of the century women could not be admitted as regular students. In 1904 she was permitted to matriculate at the University of Erlangen, which granted her the Ph.D., summa cum laude, in 1907. Her sponsor, the algebraist Gordan, strongly influenced her doctoral dissertation on algebraic invariants. Her divergence from Gordan's viewpoint and her progress in the direction of the "new" algebra first began when she was exposed to the ideas of Ernst Fischer, who came to Erlangen in 1911.

In 1915 Hilbert invited Emmy Noether to Göttingen. There she lectured at courses that were given under his name and applied her profound invariant-theoretic knowledge to the resolution of problems that he and Felix Klein were considering. Inspired by Hilbert and Klein's investigation into Einstein's general theory of relativity, Noether wrote her remarkable 1918 paper in which both the concept of *variational symmetry* and its connection with *conservation laws* were set down in complete generality.

Hilbert repeatedly tried to obtain her an appointment as Privatdozent, but the strong prejudice against women prevented her habilitation until 1919. In 1922 she was named a *nichtbeamteter ausserordentlicher Professor* ("unofficial associate professor"), a purely honorary position. Subsequently, a modest salary was provided through a *Lehrauftrag* ("teaching appointment") in algebra. Thus she taught at Göttingen (1922–1933), interrupted only by visiting professorships at Moscow (1928–1929) and at Frankfurt (summer of 1930).

In April 1933 she and other Jewish professors at Göttingen were summarily dismissed. In 1934 Nazi political pressures caused her brother Fritz to resign from his position at Breslau and to take up duties at the research institute in Tomsk, Siberia. Through the efforts of Hermann Weyl, Emmy Noether was offered a visiting professorship at Bryn Mawr College; she departed for the United States in October 1933. Thereafter she lectured and did research at Bryn Mawr and at the Institute for Advanced Studies, Princeton, but those activities were cut short by her sudden death from complications following surgery.

Emmy Noether's most important contributions to mathematics were in the area of abstract algebra. One of the traditional postulates of algebra, namely the commutative law of multiplication, was relinquished in the earliest example of a generalized algebraic structure, e.g., in Hamilton's quaternion algebra and also in many of the 1844 Grassmann algebras. From 1927 to 1929 Emmy Noether contributed notably to the theory of representations, the object of which is to provide realizations of noncommutative algebras by means of matrices, or linear transformations. From 1932 to 1934 she was able to probe profoundly into the structure of noncommutative algebras by means of her concept of the *verschränktes* ("cross") product.

Emmy Noether wrote some forty-five research papers and was an inspiration to many future mathematicians. The so-called Noether school included such algebraists as Hasse and W. Schmeidler, with whom she exchanged ideas and whom she converted to her own special point of view. She was particularly influential in the work of B. L. van der Waerden, who continued to promote her ideas after her death and to indicate the many concepts for which he was indebted to her.

Corollary 33.4.1 can be applied to most DEs in physics derivable from a Lagrangian. We are interested in partial DEs studied in classical field theories. The case of ODEs, studied in point mechanics, is relegated to Problem (33.11).

First consider spacetime translation $\mathbf{v}^i = \eta^{ij}\partial_j$, where we have introduced the Lorentz metric η^{ij} to include non-Euclidean cases. In order for $\mathbf{v}^i$ to be an infinitesimal variational symmetry, it has to satisfy Eq. (33.20), which in the case at hand, reduces to $\mathbf{v}^i(L) = 0$, or $\partial_i L = 0$.

Box 33.4.2 *In order for a variational problem to be invariant under spacetime translations, its Lagrangian must not depend explicitly on the coordinates.*

If spacetime translation happens to be a symmetry, then $X^i \to \eta^{ij}$, and the (double-indexed) conserved current, derived from Corollary 33.4.1, takes the form

$$T^{ij} = \sum_{\alpha=1}^{q} \frac{\partial u^\alpha}{\partial x_j} \frac{\partial L}{\partial u_i^\alpha} - \eta^{ij} L.$$

Using Greek indices to describe space-time coordinates, and Latin indices to label the components of $\mathbb{R}^q$, we write

energy momentum current density

$$T^{\mu\nu} = \sum_{j=1}^{q} \frac{\partial \phi^j}{\partial x_\mu} \frac{\partial L}{\partial \phi_\nu^j} - \eta^{\mu\nu} L \equiv \sum_{j=1}^{q} \eta^{\mu\sigma} \frac{\partial \phi^j}{\partial x^\sigma} \frac{\partial L}{\partial \phi_\nu^j} - \eta^{\mu\nu} L, \qquad (33.29)$$

where we changed the dependent variable u to ϕ to adhere to the notation used in the physics literature. Recall that $\phi_\nu^j \equiv \partial \phi^j / \partial x^\nu$. $T^{\mu\nu}$ is called the **energy momentum current density**.

The quantity $T^{\mu\nu}$, having a vanishing divergence, is really a density, just as the continuity equation (vanishing of the divergence) for the electric charge involves the electric charge and current densities. In the electric case, we find the charge by integrating the charge density, the zeroth component of the electric 4-current density. Similarly, we find the "charge" associated with $T^{\mu\nu}$ by integrating its zeroth component. This yields the energy momentum 4 vector:

$$P^\nu = \int_V T^{0\nu} d^3x.$$

We note that

$$\begin{aligned} \frac{dP^\nu}{dt} &= \frac{d}{dt} \int_V T^{0\nu} d^3x = \int_V \frac{\partial T^{0\nu}}{\partial t} d^3x \\ &= -\int_V \sum_{i=1}^{3} \frac{\partial T^{i\nu}}{\partial x^i} d^3x = -\int_S \sum_{i=1}^{3} T^{i\nu} da_i, \end{aligned}$$

where we have used the three-dimensional divergence theorem. By taking S to be infinite, and assuming that $T^{i\nu} \to 0$ at infinity (faster than the element of area da_i diverges), we obtain $dP^\nu/dt = 0$, the conservation of the 4-momentum.

Example 33.4.3 A relativistic scalar field of mass m is a 1-component field satisfying the Klein–Gordan equation, which is, as the reader may check, the Euler-Lagrange equation of

$$\mathbf{L}[\phi] = \int L(\phi, \phi_\mu)\, d^4x \equiv \int \frac{1}{2} \left[\eta^{\mu\nu} \partial_\mu \phi \partial_\nu \phi - m^2 \phi^2 \right] d^4x.$$

The energy momentum current for the scalar field is found to be

$$T^{\mu\nu} = \partial^\mu\phi\partial^\nu\phi - \eta^{\mu\nu}L, \qquad \partial^\mu\phi \equiv \eta^{\mu\nu}\frac{\partial\phi}{\partial x^\nu}.$$

Note that $T^{\mu\nu}$ is symmetric under interchange of its indices. This is a desired feature of the energy momentum current that holds for the scalar field but is not satisfied in general, as Eq. (33.29) indicates. The reader is urged to show directly that $\partial_\mu T^{\mu\nu} = 0 = \partial_\nu T^{\mu\nu}$, i.e., that energy momentum is conserved.

To go beyond translation, we consider classical (nonquantized) fields[6] $\{\phi^j\}_{j=1}^{n_\alpha}$, which, as is the case in most physical situations, transform among themselves as the rows of the αth irreducible representation of a Lie group G that acts on the independent variables. Under these circumstances, the generators of the symmetry are given by Eq. (30.11):

$$\mathfrak{D}_{ij}(\boldsymbol{\xi}) = \mathfrak{T}_{ij}^{(\alpha)}(\boldsymbol{\xi})\phi^k(\mathbf{x})\frac{\partial}{\partial\phi^k} + \delta_{ij}X^\nu(\mathbf{x};\boldsymbol{\xi})\frac{\partial}{\partial x^\nu}, \tag{33.30}$$

where ν labels the independent variables. Corollary 33.4.1 now gives the conserved current as

$$J_{ij}^\mu = \left\{X^\mu(\mathbf{x};\boldsymbol{\xi})\phi_\nu^k(\mathbf{x})\frac{\partial L}{\partial\phi_\nu^k} - X^\mu(\mathbf{x};\boldsymbol{\xi})L\right\}\delta_{ij} - \phi^k(\mathbf{x})\frac{\partial L}{\partial\phi_\mu^k}\mathfrak{T}_{ij}^{(\alpha)}(\boldsymbol{\xi}),$$

where summation over repeated indices is understood with $1 \le k \le n_\alpha$ and $1 \le \nu \le p$. We can rewrite this equation in the form

$$\mathbf{J}^\mu = \left\{X^\mu(\mathbf{x};\boldsymbol{\xi})\phi_\nu^k(\mathbf{x})\frac{\partial L}{\partial\phi_\nu^k} - X^\mu(\mathbf{x};\boldsymbol{\xi})L\right\}\mathbf{1} - \phi^k(\mathbf{x})\frac{\partial L}{\partial\phi_\mu^k}\boldsymbol{\mathfrak{T}}^{(\alpha)}(\boldsymbol{\xi}), \tag{33.31}$$

where $\mathbf{J}^\mu$ and $\boldsymbol{\mathfrak{T}}^{(\alpha)}(\boldsymbol{\xi})$ are $n_\alpha \times n_\alpha$ matrices whose elements are J_{ij}^μ and $\mathfrak{T}_{ij}^{(\alpha)}(\boldsymbol{\xi})$, respectively, and $\mathbf{1}$ is the unit matrix of the same dimension.

We note that the conserved current has a coordinate part (the term that includes X^μ and multiplies the unit matrix), and an "intrinsic" part (the term with no X^μ) represented by the term involving $\boldsymbol{\mathfrak{T}}^{(\alpha)}(\boldsymbol{\xi})$. If the field has only one component (a scalar field), then $\boldsymbol{\mathfrak{T}}^{(\alpha)}(\boldsymbol{\xi}) = 0$, and only the coordinate part contributes to the current.

The current $\mathbf{J}^\mu$ acquires an extra index when a *component* of $\boldsymbol{\xi}$ is chosen. As a concrete example, consider the case where G is the rotation group in $\mathbb{R}^p$. Then a typical component of $\boldsymbol{\xi}$ will be $\xi^{\rho\sigma}$, corresponding to a rotation in the $\rho\sigma$-plane, and the current will be written as $\mathbf{J}^{\mu;\rho\sigma}$. These extra indices are also reflected in X^μ, as that too is a function of $\boldsymbol{\xi}$:

$$X^\mu\big(\mathbf{x};\xi^{\rho\sigma}\big)\frac{\partial}{\partial x^\mu} = x^\rho\partial_\sigma - x^\sigma\partial_\rho \quad\Rightarrow\quad X^\mu\big(\mathbf{x};\xi^{\rho\sigma}\big) = x^\rho\delta^{\mu\sigma} - x^\sigma\delta^{\mu\rho}.$$

[6]The reader notes that the superscript α, which labeled components of the independent variable u, is now the label of the irreducible representation. The components of the dependent variable (now denoted by ϕ) are labeled by j.

The volume integral of $\mathbf{J}^{0;\rho\sigma}$ will give the components of angular momentum. When integrated, the term multiplying $\mathbf{1}$ becomes the **orbital angular momentum**, and the remaining term gives the **intrinsic spin**. The conservation of $\mathbf{J}^{\mu;\rho\sigma}$ is the statement of the conservation of *total* angular momentum. The label α denotes various representations of the rotation group. If $p = 3$, then α is simply the value of the spin. For example, the spin-$\frac{1}{2}$ representation corresponds to $\alpha = \frac{1}{2}$, and

orbital angular momentum and intrinsic spin

$$\mathfrak{T}^{(1/2)}(\boldsymbol{\xi}) = \frac{1}{2}\big(\sigma^1, \sigma^2, \sigma^3\big), \quad \text{or} \quad \mathfrak{T}^{(1/2)}\big(\xi^a\big) = \frac{1}{2}\sigma^a, \ a = 1, 2, 3,$$

with a labeling the three different "directions" of rotation.[7] If the field is a scalar, $\mathfrak{T}^{(\alpha)}(\boldsymbol{\xi}) = 0$, and the field has only an orbital angular momentum.

33.5 Problems

33.1 Show that the derivative of a linear map from one Hilbert space to another is the map itself.

33.2 Show that a complex function $f : \mathbb{C} \supset \Omega \to \mathbb{C}$ considered as a map $f : \mathbb{R}^2 \supset \Omega \to \mathbb{R}^2$ is differentiable iff it satisfies the Cauchy-Riemann conditions. Hint: Consider the Jacobian matrix of f, and note that a linear complex map $\mathbf{T} : \mathbb{C} \to \mathbb{C}$ is necessarily of the form $\mathbf{T}(z) = \lambda z$ for some constant $\lambda \in \mathbb{C}$.

33.3 Show that

$$\frac{\delta \mathbf{E}_{\mathbf{y},i}[u]}{\delta u}(\mathbf{x}) = -\partial_i \delta(\mathbf{x} - \mathbf{y}).$$

33.4 Show that the first functional derivative of $\mathbf{L}[u] \equiv \int_{x_1}^{x_2} \sqrt{1 + u_x^2}\, dx$, obtained using Eq. (33.9), is $\mathbb{E}(L)$.

33.5 Show that for the proper time of special relativity

$$\frac{\delta \mathbf{L}[x]}{\delta x(s)} = \frac{x_{ss}}{(1 - x_s^2)^{3/2}}.$$

Use this to show that the contribution of the second variational derivative to the Taylor expansion of the functional is always negative.

33.6 Show that the first prolongation of the Lorentz generator $\mathbf{v} = u\partial_x + x\partial_u$ is

$$\mathrm{pr}^{(1)}\mathbf{v} = \mathbf{v} + \big(1 - u_x^2\big)\frac{\partial}{\partial u_x}.$$

[7] Only in three dimensions can one label rotations with a single index. This is because each coordinate plane has a unique direction (by the use of the right-hand rule) perpendicular to it that can be identified as the direction of rotation.

33.7 Verify that rotation in the xu-plane is a symmetry of the arc-length variational problem (see Example 33.1.13).

33.8 Show that $\mathbf{v}_4$, $\mathbf{v}_6$, and $\mathbf{v}_7$ of Table 32.3 are variational symmetries of Eq. (33.21), but $\mathbf{v}_5$, $\mathbf{v}_8$, $\mathbf{v}_9$, and $\mathbf{v}_{10}$ are not. Find the constant c (if it exists) such that $\mathbf{v}_5 + cu\partial_u$ is a variational symmetry. Show that no linear combination of inversions produces a symmetry.

Kepler problem

33.9 The two-dimensional **Kepler problem** (for a unit point mass) starts with the functional

$$\mathsf{L} = \int \left[\frac{1}{2}(x_t^2 + y_t^2) - V(r)\right] dt, \qquad r = \sqrt{x^2 + y^2}.$$

(a) Show that L is invariant under t translation and rotation in the xy-plane.
(b) Find the generators of t translation and rotation in polar coordinates and conclude that r is the best choice for the *independent* variable.
(c) Rewrite L in polar coordinates and show that it is independent of t and θ.
(d) Write the Euler-Lagrange equations and integrate them to get θ as an integral over r.

33.10 Prove Corollary 33.4.1.

33.11 Consider a system of N particles whose total kinetic energy K and potential energy U are given by

$$K(\dot{\mathbf{x}}) = \frac{1}{2}\sum_{\alpha=1}^{N} m_\alpha |\dot{\mathbf{x}}^\alpha|^2, \qquad U(t, \mathbf{x}) = \sum_{\alpha \neq \beta} k_{\alpha\beta} |\mathbf{x}^\alpha - \mathbf{x}^\beta|^{-1},$$

where $\mathbf{x}^\alpha = (x^\alpha, y^\alpha, z^\alpha)$ is the position of the αth particle. The variational problem is of the form

$$\mathsf{L}[\mathbf{x}] = \int_{-\infty}^{\infty} L(t, \mathbf{x}, \dot{\mathbf{x}})\, dt = \int_{-\infty}^{\infty} \left[K(\dot{\mathbf{x}}) - U(t, \mathbf{x})\right] dt.$$

(a) Show that the Euler-Lagrange equations are identical to Newton's second law of motion.
(b) Write the infinitesimal criterion for the vector field

$$\mathbf{v} = \tau(t, \mathbf{x})\frac{\partial}{\partial t} + \sum_\alpha \left[\xi^\alpha(t, \mathbf{x})\frac{\partial}{\partial x^\alpha} + \eta^\alpha(t, \mathbf{x})\frac{\partial}{\partial y^\alpha} + \zeta^\alpha(t, \mathbf{x})\frac{\partial}{\partial z^\alpha}\right]$$

to be the generator of a 1-parameter group of variational symmetries of L.

(c) Show that the conserved "current" derived from Corollary 33.4.1 is

$$T = \sum_{\alpha=1}^{N} m_\alpha \left(\xi^\alpha \dot{x}^\alpha + \eta^\alpha \dot{y}^\alpha + \zeta^\alpha \dot{z}^\alpha\right) - \tau E,$$

where $E = K + U$ is the total energy of the system.

(d) Find the conditions on U such that (i) time translation, (ii) space translations, and (iii) rotations become symmetries of $\mathbf{L}$. In each case, compute the corresponding conserved quantity.

33.12 Show that the Euler-Lagrange equation of

$$\mathbf{L}[\phi] = \int L(\phi, \phi_\mu)\, d^4x \equiv \int \frac{1}{2}\left[\eta^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - m^2\phi^2\right] d^4x$$

is the Klein–Gordan equation. Verify that $T^{\mu\nu} = \partial^\mu\phi\partial^\nu\phi - \eta^{\mu\nu}L$ are the currents associated with the invariance under translations. Show directly that $T^{\mu\nu}$ is conserved.