Notes on Symmetries, Conservation Laws, and Virial Relations

R. Pyke, Department of Mathematics, University of Toronto
January 1996

These notes describe some aspects of partial differential equations from the view point of the variational calculus. We begin by showing how a variational problem (the study of critical points of an action functional) can be associated to a differential equation. Here the idea is to define the action in such a way that the associated Euler-Lagrange equation is the differential equation of interest. Then we study how the action and the the Euler-Lagrange equation are affected by transformation groups that act on the space of functions where the action is defined. This leads to integral identities (virial relations) that are satisfied by solutions of the Euler-Lagrange equations. The notion of a symmetry group of an action functional and of a differential equation is defined in Section III. We describe how symmetry groups lead to conservation laws for the Euler-Lagrange equation (Noether's theorem). The final section discusses virial relations in more detail and their relation to conservation laws.

The references [DFN], [GF] and [HC-I] (listed at the end) provide a complete account of most of the topics covered here.

I Variational Calculus and Euler-Lagrange Equations

Many differential equations (both ordinary and partial) that occur in physics can be formulated as a variational problem. There are several advantages in representing the equation this way. For example, it allows us to generalize the notion of a solution of a differential equation and it leads to methods for finding solutions. In addition, certain features of the equation become more apparent, like conservation laws.

We will be letting $K(\varphi) = 0$ denote a partial differential equation. For instance, if the equation of interest is $\partial_t^2\varphi - \Delta\varphi + \varphi^2 = 0$, then $K(\varphi) = \partial_t^2\varphi - \Delta\varphi + \varphi^2$. For arbitrary functions φ, the formula $K(\varphi)$ may still make sense. Therefore, you can also think of K as a (not necessarily linear) operator, taking the function φ to the function ψ where $\psi = K(\varphi)$. In the case $\psi = 0$ we say that φ is a solution.

We now describe how to formulate a variational problem associated to the equation $K(\varphi) = 0$. Let X be a (Banach) space of functions defined on a subset $\Omega \subset \mathbf{R}^m$ and S a functional $X \to \mathbf{R}$ of the form

$$S[\varphi] = \int_\Omega \mathcal{S}(\varphi, \nabla_m \varphi)\, d^m x, \quad \varphi \in X \tag{I.1}$$

where $\mathcal{S}$, the *Lagrangian*, is a function of $m+1$ variables; $\mathcal{S} : \mathbf{R} \times \mathbf{R}^m \to \mathbf{R}$ via $(u, p) \to \mathcal{S}(u, p)$ for $u \in \mathbf{R}$, $p \in \mathbf{R}^m$. S is called *action.* Here ∇_m denotes the gradient with respect to all the variables; $\nabla_m = (\partial/\partial x_1, \ldots, \partial/\partial x_m)$. Our choice of the space X (that is, the properties of the functions φ) will depend upon what types of solutions of the equation $K(\varphi) = 0$ we want to study and on the form of the Lagrangian $\mathcal{S}$ so that the integral (I.1) is well-defined.

S will be defined in such a way as to be associated to the equation $K(\varphi) = 0$ through the relation $S'[\varphi] = 0 \Leftrightarrow \varphi$ is a (weak) solution of $K(\varphi) = 0$. Here $S'[\varphi]$ denotes the (Fréchet) derivative of S at the point φ, which is a linear operator on X. $S'[\varphi] = 0$ means, then, that $S'[\varphi](\psi) = 0$ for all $\psi \in X$. If $S'[\varphi] = 0$ we call φ a *critical point* of S. $S'[\varphi](\psi)$ is given by the formula

$$S'[\varphi](\psi) = \int_\Omega (\mathcal{S}_u \psi + \mathcal{S}_p \cdot \nabla_m \psi)\, d^m x \tag{I.2}$$

where $\mathcal{S}_u = \partial \mathcal{S}/\partial u$ and $\mathcal{S}_p = (\partial \mathcal{S}/\partial p_1, \ldots, \partial \mathcal{S}/\partial p_m)$. The linearity of $S'[\varphi](\psi)$ in ψ is apparent from this formula ($S'[\varphi]$ may not be linear in φ, though). Equation (I.2) can be derived by the calculation

$$S'[\varphi](\psi) \;=\; \frac{d}{d\varepsilon} S[\varphi + \varepsilon\psi]\Big|_{\varepsilon=0}. \tag{I.3}$$

That is, $S'[\varphi](\psi)$ is the (directional) derivative of S at the point φ in the direction ψ. Writing

$$\mathcal{S}_p \cdot \nabla_m \psi = \nabla_m \cdot (\mathcal{S}_p \psi) - (\nabla_m \cdot \mathcal{S}_p)\psi, \tag{I.4}$$

and assuming that the boundary term $\int_\Omega \nabla_m \cdot \mathcal{S}_p \psi = \int_{\partial\Omega} \mathcal{S}_p \psi \cdot \mathbf{n}$ vanishes (here $\partial\Omega$ is the boundary of Ω with unit outward normal $\mathbf{n}$), equation (I.2) becomes

$$S'[\varphi](\psi) \;=\; \int_\Omega (\mathcal{S}_u - \nabla_m \cdot \mathcal{S}_p)\psi \; d^m x \tag{I.5}$$

((I.2)$\to$(I.5) is just integration by parts).

If φ is a critical point of S and if equation (I.5) holds for all (sufficiently smooth) ψ, then we obtain the *Euler-Lagrange equation*;

$$0 \;=\; \mathcal{S}_u - \nabla_m \cdot \mathcal{S}_p \;=\; \frac{\partial \mathcal{S}}{\partial \varphi} - \sum_{i=1}^{m} \frac{\partial}{\partial x_i} \frac{\partial \mathcal{S}}{\partial \varphi_i}. \tag{I.6}$$

Here $\varphi_i \equiv \partial\varphi/\partial x_i$ and the notation $\partial\mathcal{S}/\partial\varphi$ and $\partial\mathcal{S}/\partial\varphi_i$ mean the functions $\partial\mathcal{S}(u,p)/\partial u$ and $\partial\mathcal{S}(u,p)/\partial p_i$ evaluated at $(u,p) = (\varphi, \nabla_m\varphi)$. Note that (I.6) is a differential equation in φ. This shows how to connect a variational problem to a differential equation $K(\varphi) = 0$: define a Lagrangian $\mathcal{S}$ in such a way that the Euler-Lagrange equation associated to the action (I.1) is $K(\varphi) = 0$;

$$\frac{\partial \mathcal{S}}{\partial \varphi} - \sum_{i=1}^{m} \frac{\partial}{\partial x_i} \frac{\partial \mathcal{S}}{\partial \varphi_i} \;=\; K(\varphi) \;=\; 0. \tag{I.7}$$

If φ is not sufficiently smooth (i.e., differentiable) then we may not be able to move derivatives from ψ to φ (cf. (I.4)) and we must use equation (I.2) rather that equation (I.5) as the definition of $S'[\varphi](\psi)$. Even so, we still regard S as being associated to the equation $K(\varphi) = 0$ (through (I.7)) and the critical point φ as a solution of $K(\varphi) = 0$ even though $K(\varphi)$ may not make sense in the usual way. In this case we call φ a *weak solution* of the equation; it satisfies the equation in an average sense (that is, the integral in (I.2) is zero for all ψ - see also (I.12) below). If φ is a *strong solution*, i.e., φ is smooth enough so that $K(\varphi)$ makes sense in the usual way and $K(\varphi) = 0$, then φ is also a weak solution (go backwards from (I.5) to (I.2) noting that $\mathcal{S}_u - \nabla_m \cdot \mathcal{S}_p = 0$ in this case). However, weak solutions need not be strong solutions. Both weak and strong solutions are critical points of S.

In the physics literature equation (I.5) is often written as

$$\delta S \;=\; \int_\Omega \frac{\delta S}{\delta \varphi}\, \delta\varphi \, d^m x, \tag{I.8}$$

where δS is the *variation* of S at the point φ under the (infinitesimal) variation $\delta\varphi = \psi$ of φ, and $\delta S/\delta\varphi = \mathcal{S}_u - \nabla_m \cdot \mathcal{S}_p$ is the *variational* or *functional derivative* of S at the point φ (see, for example, [G] or [DFN]). In this notation, $S[\varphi + \delta\varphi] \;=\; S[\varphi] + \delta S$. Therefore, if φ is a critical point of S then $\delta S = 0$ and so $S[\varphi + \delta\varphi] = S[\varphi]$ for all variations $\delta\varphi$. We say in this case that the action is *stationary* at the point φ. For example, if φ is a minimum of the action then $\delta S = 0$ (in ordinary calculus the same is true: the derivative of a function is zero at a local minimum or maximum). That is why the condition that $\delta S = 0$ ($\Leftrightarrow S'[\varphi] = 0$), which is a condition on φ, is referred to as *the principle of least action*: Those functions that obey a certain physical law, i.e., those functions that solve $K(\varphi) = 0$ (the equations of motion), are points where the associated action assumes its minimum value.

As an example, we formulate the nonlinear wave equation (NLW)

$$\partial_t^2\varphi - \Delta\varphi + f(\varphi) = 0 \tag{I.9}$$

as a variational problem. Here φ is a real valued function of the space-time variables $(x,t) \in \mathbf{R}^{N+1}$, $\partial_t^2\varphi = \partial^2\varphi/\partial t^2$, $\Delta\varphi = \sum_{i=1}^N \partial^2\varphi/\partial x_i^2$, and f is a nonlinear function $f : \mathbf{R} \to \mathbf{R}$ (so that $f(\varphi)(x,t) = f(\varphi(x,t))$) which we assume satisfies $f(0) = 0$. This equation is important in relativistic field theories because it is a nonlinear generalization of the Klein-Gordon equation $\partial_t^2\varphi - \Delta\varphi + m^2\varphi = 0$, which in turn is a relativistic form of the Schrödinger equation $i\partial_t\varphi - \Delta\varphi = 0$. The Klein-Gordon equation describes the field of a free particle of mass m, while the nonlinear term $f(\varphi)$ in NLW represents a self-interaction of the field with itself. It is because of this term that NLW may have bound state solutions, an example of which are the time-periodic solutions.

For NLW we define the Lagrangian

$$\mathcal{S} = \mathcal{S}(\varphi, \partial_t\varphi, \nabla\varphi) = -\frac{1}{2}(\partial_t\varphi)^2 + \frac{1}{2} \mid \nabla\varphi \mid^2 + F(\varphi) \tag{I.10}$$

where $F(x) = \int_0^x f(y)\,dy$. Here, as elsewhere in the notes, ∇ denotes the gradient operator on $\mathbf{R}^N$ (the spatial variables x). Then, for any $T \in \mathbf{R}$ we define an action

$$S[\varphi] \;=\; \int_{-T}^{T}\int_{\mathbf{R}^N} \Big\{ -\frac{1}{2}(\partial_t\varphi)^2 + \frac{1}{2}|\nabla\varphi|^2 + F(\varphi)\Big\}\, d^N x\, dt. \tag{I.11}$$

In this case we can take for the space X on which to define S the set of functions $\varphi(x,t)$ such that

$$\int_{-T}^{T}\int_{\mathbf{R}^N} \varphi^2\, d^N x\, dt, \quad \int_{-T}^{T}\int_{\mathbf{R}^N} (\partial_t\varphi)^2\, d^N x\, dt, \quad \text{and} \quad \int_{-T}^{T}\int_{\mathbf{R}^N} |\nabla\varphi|^2\, d^N x\, dt$$

are all finite for any $T < \infty$. (In addition we should say something about f so that $\int_{-T}^{T} \int_{\mathbf{R}^N} F(\varphi)$ is finite for $\varphi \in X$.) From equation (I.2), and using equation (I.15) below, we have that

$$S'[\varphi](\psi) = \int_{-T}^{T} \int_{\mathbf{R}^N} \Big\{ -\partial_t \varphi \partial_t \psi + \nabla\varphi \cdot \nabla\psi + f(\varphi)\psi \Big\}\, d^N x\, dt. \qquad \text{(I.12)}$$

If this equals zero for all functions $\psi \in X$ and for all $T \in \mathbf{R}$, then φ is (by definition) a weak solution of NLW. Formally, this can be derived by multiplying NLW by ψ, integrating over $\mathbf{R}^N \times (-T, T)$, and then integrating by parts. (I.12) allows us to make sense of solutions of NLW that are only once differentiable.

We will also have occasion to consider time periodic solutions of NLW. Suppose we are interested in $2\pi/\omega$-periodic solutions of NLW: $\varphi(x, t+2\pi/\omega) = \varphi(x,t) \;\; \forall (x,t) \in \mathbf{R}^{N+1}$. Then we define an action on a space of $2\pi/\omega$-periodic functions by the formula

$$S[\varphi] \;=\; \int_{S^1_\omega} \int_{\mathbf{R}^N} \Big\{ -\frac{1}{2}(\partial_t\varphi)^2 + \frac{1}{2}|\nabla\varphi|^2 + F(\varphi) \Big\}\, d^N x\, dt, \qquad \text{(I.13)}$$

where S^1_ω is the circle of radius ω^{-1}.

Let us verify that the Euler-Lagrange equation associated to the Lagrangian (I.10) is in fact NLW. Here

$$\mathcal{S}(u,p) \;=\; -\frac{1}{2}p_{N+1}^2 + \frac{1}{2}(p_1^2 + \cdots + p_N^2) + F(u), \qquad \text{(I.14)}$$

and so

$$\mathcal{S}_u(u,p) = f(u) \quad \text{and} \quad \mathcal{S}_p(u,p) = (p_1, \ldots, p_N, -p_{N+1}). \qquad \text{(I.15)}$$

Noting that in this case $\nabla_m = \nabla_{N+1} = (\nabla, \partial_t)$, we find that

$$\mathcal{S}_u - \nabla_{N+1} \cdot \mathcal{S}_p\Big|_{(u,p)=(\varphi,\nabla\varphi,\partial_t\varphi)} \;=\; f(\varphi) - \nabla \cdot (\nabla\varphi) + \partial_t(\partial_t\varphi). \qquad \text{(I.16)}$$

Hence the Euler-Lagrange equation (I.6) is NLW.

II Transformation Groups and Virial Relations

Returning to the abstract set-up, variations of S can be defined through transformation groups acting on the space of functions X as follows. For each

$\lambda \in \mathbf{R}$ let T_λ be a transformation of X, $T_\lambda : X \to X$, where we will write $\varphi_\lambda \equiv T_\lambda \varphi$ and will assume that $T_0 = \mathbf{1}$ (the identity transformation). The whole collection $\{T_\lambda\}_{\lambda \in \mathbf{R}}$ of transformations is called a *1-parameter group of transformations*. It is a group under the operation of composition if $T_{\lambda_1} \circ T_{\lambda_2} = T_{\lambda_1+\lambda_2}$. We will denote the group of transformations just by T_λ. The *infinitesimal generator* of this group is the (linear) operator $A : X \to X$ defined by $A\varphi = \frac{d}{d\lambda} T_\lambda \varphi\Big|_{\lambda=0}$. If φ is a critical point of S (i.e., $S'[\varphi] = 0$), then applying the chain rule to the function $S[\varphi_\lambda] : \mathbf{R} \to \mathbf{R}$ (notice that since φ is fixed this is a function of λ only), we find that

$$\frac{d}{d\lambda} S[\varphi_\lambda]\Big|_{\lambda=0} \;=\; S'[\varphi](A\varphi) \;=\; 0. \tag{II.1}$$

The equation $S'[\varphi](A\varphi)$ is an integral formula involving the solution φ and its derivatives (cf. (I.2)). We call equation (II.1) **the virial relation associated to the transformation** T_λ. Any solution φ of the Euler-Lagrange equation associated to the action S, or, more generally, any critical point of S, must satisfy this identity. Therefore, by formulating the equation $K(\varphi) = 0$ as a variatio nal problem we have found a way to derive integral identities that must be satisfied by any solution.

Apart from providing necessary conditions for the existence of solutions, virial relations can be used in other ways as well and play an important role in physics. A well known example is the virial theorem of classical mechanics which relates the time-average kinetic and potential energy of an n-particle system under the influence of central forces; see [G].

A particular class of transformations of X arise from transformations of the underlying space $\mathbf{R}^m$. Suppose $v : \mathbf{R}^m \to \mathbf{R}^m$ is a vector field on $\mathbf{R}^m$ that generates the flow Φ_λ (*flow*: for each $\lambda \in \mathbf{R}$, Φ_λ is a map from $\mathbf{R}^m$ to $\mathbf{R}^m$ such that $\Phi_{\lambda_1} \circ \Phi_{\lambda_2} = \Phi_{\lambda_1+\lambda_2}$ and $\Phi_0 = \mathbf{1}$). We may sometimes write $\Phi_\lambda(x)$ as $x(\lambda)$ - then, since $\Phi_0 = \mathbf{1}$, $x(0) = x$. The ordinary differential equation satisfied by the function $x(\lambda)$ is then found to be $\dot{x}(\lambda) = \frac{d}{d\lambda}\Phi_\lambda(x) = v(x(\lambda))$. In this notation the solution $x(\lambda)$ of the differential equation $\dot{x} = v$ represents the flow Φ_λ associated to v, and this is why we said that v generates the flow Φ_λ.

We can use the flow Φ_λ generated by v to define a transformation T_λ of the space X by the formula $T_\lambda \varphi = \varphi \circ \Phi_\lambda \equiv \varphi_\lambda$. Then, T_λ is a 1-parameter group of transformations of X with infinitesimal generator $A = v \cdot \nabla_m$.

($T_{\lambda_1} \circ T_{\lambda_2} T_{\lambda_1+\lambda_2}$ follows from the fact that Φ_λ is a group, while $\frac{d}{d\lambda}\varphi_\lambda\Big|_{\lambda=0} = v \cdot \nabla_m \varphi$ can be seen by keeping in mind that $\varphi_\lambda = \varphi(x(\lambda))$.) In this case, if φ is a critical point of S, and noting that $\varphi_0 = \varphi$, equation (II.1) becomes

$$\frac{d}{d\lambda} S[\varphi_\lambda]\Big|_{\lambda=0} \;=\; S'[\varphi](v \cdot \nabla_m \varphi) \;=\; \int_\Omega \Big\{ \mathcal{S}_u (v \cdot \nabla_m \varphi) + \mathcal{S}_p \cdot \nabla_m (v \cdot \nabla_m \varphi) \Big\}\, d^m x \;=\; 0 \tag{II.2}$$

(cf. (I.2)). We call this equation **the virial relation associated to the vector field** v. Keeping in mind that $\mathcal{S}_u$ and $\mathcal{S}_p$ are both functions of φ and $\nabla_m \varphi$, we see that (II.2) is an integral identity involving the solution φ. In Section IV we give an example of a virial relation for periodic solutions of NLW and show how it leads to a necessary condition for the existence of such solutions.

III Symmetries and Conservation Laws

From now on we will assume that the transformation group T_λ acting on the space of functions X derives from a flow Φ_λ on the underlying space $\mathbf{R}^m$ via the formula $T_\lambda \varphi = \varphi \circ \Phi_\lambda \equiv \varphi_\lambda$ as described above. For a given action S some transformation groups are distinguished because they leave S invariant, i.e., $S \circ T_\lambda = S \;\; \forall \lambda \in \mathbf{R}$ ($\Longleftrightarrow S[\varphi_\lambda] = S[\varphi] \;\; \forall \varphi \in X, \;\; \forall \lambda \in \mathbf{R}$). In this case we call T_λ a *symmetry group* of S. This implies that it is also a symmetry group of the Euler-Lagrange equation $K(\varphi) = 0$ associated to S, i.e., $K(\varphi) = 0 \Rightarrow K(\varphi_\lambda) = 0 \;\; \forall \lambda \in \mathbf{R}$ ([O] Thm 4.14). The converse does not hold in general, though (see [O] Ch. 4).

Definition III.1 T_λ *is a* symmetry group of K *if for all* $\varphi \in X$*, and for all* $\lambda \in \mathbf{R}$*,* $K(\varphi \circ \Phi_\lambda) = K(\varphi) \circ \Phi_\lambda$*, where* $T_\lambda \varphi = \varphi \circ \Phi_\lambda$*.*

What this definition means is that K and T_λ commute (as transformation of X): $\quad K \circ T_\lambda = T_\lambda \circ K$.

Definition III.2 T_λ *is a* symmetry group of the equation $K(\varphi) = 0$ *if for all* solutions φ *and all* $\lambda \in \mathbf{R}$*,* $\varphi \circ \Phi_\lambda$ *is also a solution.*

That is, T_λ leaves the subspace *ker* (K) invariant. Therefore, K and T_λ commute when restricted to this subspace. (*ker* (A) denotes the *kernal* of the operator A: *ker* $(A) = \{\varphi \in X \mid A\varphi = 0\}$.)

Example Translations in space and time are symmetries of NLW. Translation in time is the flow $\Phi_\lambda(x,t) = (x, t+\lambda)$, and the generator v of this flow is the vector field $v(x,t) = (0,1)$. Translation in space in the (unit) direction $\hat{n}$ is the flow $\Phi_\lambda(x,t) = (x+\lambda\hat{n}, t)$ with generator $v(x,t) = (\hat{n}, 0)$. Check that if φ solves NLW then in each case so does the function φ_λ (first note that $\varphi_\lambda(x,t) = \varphi(x, t+\lambda)$ or $\varphi(x+\lambda\hat{n}, t)$, respectively).

The group of symmetries of the equation $K(\varphi) = 0$ is in general larger than the group of symmetries of K ($ker\,(K)$ is usually smaller than X). For example, the free wave equation $\Box\varphi \equiv \partial_t^2\varphi - \Delta\varphi = 0$ has as its group of symmetries (Definition III.2) the conformal group while the group of symmetries of $\Box$ (Definition III.1) is the Poincaré group, which is a subgroup of the conformal group. Using the duality between $\Box$ and the Lorentz metric g_L on $\mathbf{R}^{N+1}$ ($\partial_t^2 - \Delta \;\;\leftrightarrow\;\; dt^2 - (dx_1^2 + \cdots + dx_N^2)$), the group of symmetries of $\Box$ is the group of isometries of g_L (i.e., transformations of $\mathbf{R}^{N+1}$ that preserve the lengths $\|x\|_{g_L}$ of points x with respect to the Lorentz metric: $\|x\|_{g_L}^2 = x_{N+1}^2 - (x_1^2 + \cdots + x_N^2)$) while the conformal group only preserves the <u>null cone</u> of g_L (which are the points $x \in \mathbf{R}^{N+1}$ such that $\|x\|_{g_L} = 0$).

We also point out that all symmetries of $\Box$ that come from transformations of space-time are symmetries of NLW (or rather the operator K associated to NLW), simply because the nonlinearity f acts on the functions by composition and therefore commutes with any transformation T_λ of this form. This is not necessarily true for transformations T_λ that do not come from transformations of space-time. For example, the gauge transformation $T_\lambda\varphi = (1+\lambda)\varphi$ is a symmetry of $\Box$ but it is not (in general) a symmetry of NLW. Nor is this necessarily true for symmetries of the equation $\Box\varphi = 0$: the dilation group $\Phi_\lambda(x,t) = (\lambda x, \lambda t)$ is a symmetry of $\Box\varphi = 0$ but it is not (in general) a symmetry of NLW.

Definition III.3 *T_λ is a* symmetry group of an action functional *S if $S[\varphi_\lambda] = S[\varphi]$ for all $\varphi \in X$ and for all $\lambda \in \mathbf{R}$ where $\varphi_\lambda \equiv \varphi \circ \Phi_\lambda$.*

We denote by $G_{symm}(S)$ the set (actually it is a group) of all such T_λ and call $G_{symm}(S)$ the *group of symmetries of* S. If $T_\lambda \in G_{symm}(S)$, then obviously

$$\frac{d}{d\lambda} S[\varphi_\lambda] = S'[\varphi_\lambda](v \cdot \nabla_m \varphi_\lambda) \;=\; 0 \qquad \forall\, \varphi,\, \lambda \tag{III.1}$$

where $v \cdot \nabla_m$ is the infinitesimal generator of T_λ. Notice that φ_λ need *not* be a critical point of S (compare (III.1) with (II.1)).

At a critical point φ, the variation δS of S is zero in all directions $\delta\varphi$. In contrast, equation (III.1) says that at *any* point φ, the variation of S in directions $\delta\varphi = v \cdot \nabla_m \varphi$ determined by the group of symmetries of S, is zero (evaluate (III.1) at $\lambda = 0$). These directions are tangent to a surface in X of constant S.

Let $\mathcal{M}_\varphi = \{\psi \in X \mid \psi = T_\lambda\varphi, \ \text{ for some } T_\lambda \in G_{symm}(S) \text{ and for some } \lambda\}$ be the the set (or manifold) of functions that are transformations of φ under elements from the group of symmetries of S. $\mathcal{M}_\varphi$ is made up of curves $\{T_\lambda\varphi\}_{\lambda\in\mathbf{R}}$ in X that pass through the point φ. These curves are the *orbits* of φ under the action of transformations from $G_{symm}(S)$. For each 1-parameter group of symmetries T_λ we obtain one such curve. By definition of $G_{symm}(S)$, S is constant on $\mathcal{M}_\varphi$.

If $K(\varphi) = 0$ is the Euler-Lagrange equation associated to S, then it is natural to ask how the symmetry groups of K and of the equation $K(\varphi) = 0$ are related to the symmetry groups of S. In general, the symmetries of K (and therefore of the equation $K(\varphi) = 0$) are a larger group than the symmetries of S. Therefore, every symmetry group of S is a symmetry group of its Euler-Lagrange equation. For example, consider the free wave equation again. The dilation group $\Phi_\lambda(x,t) = (\lambda x, \lambda t)$ is a symmetry of $\Box\varphi = 0$ but it is not a symmetry of $\Box$ nor of $S_\Box$ (unless $N = 1$): $\Box\varphi_\lambda = \lambda^2\Box\varphi$ and $S_\Box[\varphi_\lambda] = \lambda^{-N+1}S_\Box[\varphi]$ where

$$S_\Box[\varphi] \ = \ \int_{-T}^{T}\int_{\mathbf{R}^N} \Big\{ -\frac{1}{2}(\partial_t\varphi)^2 + \frac{1}{2}|\nabla\varphi|^2 \Big\}\, d^N x\, dt \qquad \text{(III.2)}$$

is the action associated to the free wave equation. In other words, $\Box \circ T_\lambda = \lambda^2 T_\lambda \circ \Box$ and $S_\Box \circ T_\lambda = \lambda^{-N+1} T_\lambda \circ S_\Box$. One can check that translations in space and time (cf. the example above) are symmetries of $S_\Box$ and of the action associated to NLW (equation (I.11)).

There is an important relationship between symmetry groups of S and conservation laws of its Euler-Lagrange equation, which we now turn to.

Definition III.4 *A* conservation law for K *is a formula* $Div(\mathcal{Q}(\varphi, \nabla_m\varphi)) = 0$ *where* φ *is a solution of* $K(\varphi) = 0$, $\ \mathcal{Q}: \ \mathbf{R} \times \mathbf{R}^m \to \mathbf{R}^m$, *and* $Div \equiv \nabla_m \cdot$ *is the divergence operator on* $\mathbf{R}^m$.

We explain this definition. Suppose that one of the independent variables is time, t, and the other variables are denoted by $x \in \mathbf{R}^N$ so that $\nabla_m = (\nabla, \partial_t)$, $m = N+1$, and where ∇ denotes the gradient operator on $\mathbf{R}^N$. Then a conservation law can be written as

$$Div(\mathcal{Q}) \;=\; \nabla_{N+1} \cdot (\mathbf{p}, e) \;=\; \partial_t e + \nabla \cdot \mathbf{p} \;\;=\;\; 0 \qquad \text{(III.3)}$$

where $e \in \mathbf{R}$ and $\mathbf{p} \in \mathbf{R}^N$ are functions of φ, $\partial_t\varphi$ and $\nabla\varphi$ and we have written $\mathcal{Q}(\varphi, \nabla_{N+1}\varphi) = (\mathbf{p}, e)$. If we set

$$E(t) \;=\; \int_{\mathbf{R}^N} e(\varphi, \partial_t\varphi, \nabla\varphi)\, d^N x, \qquad \text{(III.4)}$$

and if $\mathbf{p}$ vanishes sufficiently rapidly as $|\,x\,| \to \infty$, then from equation (III.3) and the divergence theorem we find that

$$\frac{d}{dt}E(t) = 0. \qquad \text{(III.5)}$$

That is, E is conserved quantity. To illustrate this further, suppose φ is *any* function on $\mathbf{R}^{N+1}$. Then define the quantity $E(t)$ by the formula (III.4) with this φ. There is no reason to expect now that E is constant. However, if φ happens to be a solution of $K(\varphi) = 0$, so that φ evolves in time (as a function of x) in a certain way, then E *is* constant.

We call $\mathbf{p}$ the "flux of e" because for any subset $\Omega \subset \mathbf{R}^N$,

$$\frac{d}{dt}E(t)(\Omega) \;=\; \frac{d}{dt}\int_\Omega e\, d^N x \;=\; -\int_{\partial\Omega} \mathbf{p}\cdot\mathbf{n}\, ds. \qquad \text{(III.6)}$$

That is, the rate of change of e inside the volume Ω is equal to the flux of e crossing the boundary $\partial\Omega$ with unit outward normal $\mathbf{n}$ (here ds is surface measure on $\partial\Omega$).

The following (fundamental) lemma is the starting point to deriving the connection between symmetry groups and conservation laws ([GF] §37, [HC-I] Ch IV §12.8, [O] §4.4).

Lemma III.5 (Noether) *If T_λ is a symmetry group of S with infinitesimal generator $v \cdot \nabla_m$, then for any φ the expression*

$$\mathcal{S}_u(v \cdot \nabla_m\varphi) \;+\; \mathcal{S}_p \cdot \nabla_m(v \cdot \nabla_m\varphi) \qquad \text{(III.7)}$$

is a divergence (cf. (III.1) evaluated at $\lambda = 0$, and (II.2)).

Here $\mathcal{S}_u$ and $\mathcal{S}_p$ are evaluated at $(\varphi, \nabla_m\varphi)$. In fact, in this case

$$\mathcal{S}_u(v \cdot \nabla_m\varphi) \ + \ \mathcal{S}_p \cdot \nabla_m(v \cdot \nabla_m\varphi) \ = \ \nabla_m \cdot \mathcal{S}v \tag{III.8}$$

(for a demonstation of this, see the previously mentioned references) . Note that (III.7) being a divergence directly implies that $S'[\varphi](v \cdot \nabla_m\varphi) = 0$ for all φ, which we knew already (cf. (I.2) and (III.1)).

Now we specialize to the situation where φ is a critical point of S ((III.8) is valid for *any* φ). From (III.8), and using (I.4) and the Euler-Lagrange equation (I.6), we obtain

$$\begin{aligned} 0 \ &= \ \mathcal{S}_u(v \cdot \nabla_m\varphi) \ + \ \mathcal{S}_p \cdot \nabla_m(v \cdot \nabla_m\varphi) \ - \ \nabla_m \cdot \mathcal{S}v \\ &= \ (\mathcal{S}_u - \nabla_m \cdot \mathcal{S}_p)v \cdot \nabla_m\varphi \ + \ \nabla_m \cdot (\mathcal{S}_p v \cdot \nabla_m\varphi - \mathcal{S}v) \\ &= \ \nabla_m \cdot (\mathcal{S}_p v \cdot \nabla_m\varphi - \mathcal{S}v). \end{aligned} \tag{III.9}$$

That is, $Div\,(\mathcal{Q}(\varphi, \nabla_m\varphi)) = 0$ where $\mathcal{Q}(\varphi, \nabla_m\varphi) = (\mathcal{S}_p v \cdot \nabla_m\varphi - \mathcal{S}v)$. Therefore (cf. Definition III.4) we have a conservation law. Thus, the fact that the quantity (III.7) is a divergence when T_λ is a symmetry group of S implies directly that there is an associated conservation law of the Euler-Lagrange equation. This is how Noether's theorem is usually stated.

We can calculate the corresponding conserved quantity by following the argument presented after Definition III.4. In space-time variables (x, t) and writing $\nabla_m = (\nabla, \partial_t)$, $v = (\vec{v}, v_{N+1})$, $\vec{v} \in \mathbf{R}^N$, $v_{N+1} \in \mathbf{R}$, the last line in (III.9) reads

$$0 \ = \ \nabla \cdot \left(\frac{\partial \mathcal{S}}{\nabla\varphi} v \cdot \nabla_m\varphi - \mathcal{S}\vec{v} \right) + \partial_t \left(\frac{\partial \mathcal{S}}{\partial\varphi_t} v \cdot \nabla_m\varphi - \mathcal{S}v_{N+1} \right), \tag{III.10}$$

where $\partial\mathcal{S}/\partial\nabla\varphi = (\partial\mathcal{S}/\partial\varphi_1, \ldots, \partial\mathcal{S}/\partial\varphi_N)$, $\varphi_i \equiv \partial\varphi/\partial x_i$, $\varphi_t \equiv \partial\varphi/\partial t$. Therefore, associated to the symmetry group Φ_λ is a conserved quantity, namely,

$$E(t) = \int_{\mathbf{R}^N} \left(\frac{\partial \mathcal{S}}{\partial\varphi_t} v \cdot \nabla_m\varphi - \mathcal{S}v_{N+1} \right) d^N x. \tag{III.11}$$

Let us return to NLW to illustrate these ideas. We have seen that space and time translations are symmetries of the action associated to NLW. If Φ_λ is the translation in space $x_j \mapsto x_j + \lambda$, then its infinitesimal generator v

has components $v_i = 0,\ i \neq j,\ v_j = 1$ and so the conserved quantity is (cf. (III.11))

$$\begin{aligned} P_j(t) &= \int_{\mathbf{R}^N} \frac{\partial \mathcal{S}}{\partial \varphi_t} \partial_j \varphi \, d^N x \\ &= \int_{\mathbf{R}^N} p_j \, d^N x, \end{aligned} \tag{III.12}$$

which is the j^{th} component of the momentum $\mathbf{P}(\varphi)$ of the solution φ: $\mathbf{P}(\varphi) = \int \mathbf{p} \, d^N x$. $p_j = \partial_t \varphi \partial_j \varphi$ is the j^{th} component of the momentum density $\mathbf{p}(\varphi) = \partial_t \varphi \nabla \varphi$. Here $\partial_j \varphi = \partial \varphi / \partial x_j$ and $\partial_t \varphi = \partial \varphi / \partial t$. For translations in time $t \mapsto t + \lambda,\ \ v_i = 0,\ \ i = 1, ...N,\ v_{N+1} = 1$ and equation (III.11) reads

$$E(t) = \int_{\mathbf{R}^N} \left\{ \frac{1}{2} (\partial_t \varphi)^2 + \frac{1}{2} \mid \nabla \varphi \mid^2 + F(\varphi) \right\} d^N x \tag{III.13}$$

which is the energy of the solution φ. To study angular momentum we make things notationally simpler by supposing that $N = 3$ and writing $x = (x_1, x_2, x_3)$. Let $\Phi_\lambda(x, t) = (R_\lambda x, t)$ be rotation around the z $(= x_3)$ axis by the angle λ;

$$R_\lambda = \left[\begin{array}{ccc} \cos\lambda & \sin\lambda & 0 \\ -\sin\lambda & \cos\lambda & 0 \\ 0 & 0 & 0 \end{array} \right]. \tag{III.14}$$

The infinitesimal generator of Φ_λ is $v(x_1, x_2, x_3, t) = (x_2, -x_1, 0, 0)$ because,

$$\begin{aligned} \frac{d}{d\lambda} \varphi_\lambda(x_1, x_2, x_3, t) \Big|_{\lambda=0} &= \frac{d}{d\lambda} \varphi(\cos\lambda \, x_1 + \sin\lambda \, x_2,\ -\sin\lambda \, x_1 + \cos\lambda \, x_2,\ x_3,\ t) \Big|_{\lambda=0} \\ &= \partial_1 \varphi x_2 \ - \ \partial_2 \varphi \, x_1. \end{aligned} \tag{III.15}$$

Therefore, the conserved quantity is

$$L_z(t) = \int_{\mathbf{R}^3} (x_2 \partial_1 \varphi - x_1 \partial_2 \varphi) \partial_t \varphi \ d^3 x. \tag{III.16}$$

This is the z component of angular momentum $\mathbf{L}(\varphi)$ of the solution φ: $\mathbf{L}(\varphi) = \int (\mathbf{r} \wedge \mathbf{p}) \, d^3 x$, where $\wedge$ denotes the cross product. Analogous calculations can be carried out for rotations about any axis and in other spatial dimensions. Thus, *conservation of momentum, energy, and angular momentum for solutions of NLW follows from the symmetries of translation in space* *and time, and rotations in space respectively, of the action S associated to NLW.*

The other Lorentz transformations provide additional symmetries of NLW. For example, consider the Lorentz boost in the x_1 direction with velocity λ;

$$\Phi_\lambda(x,t) \;=\; \Big(\frac{x_1-\lambda t}{\sqrt{1-\lambda^2}}, x_2, \ldots, x_N, \frac{t-\lambda x_1}{\sqrt{1-\lambda^2}}\Big), \;\; |\lambda|<1, \qquad \text{(III.17)}$$

which has the infinitesimal generator $v(x,t) = \beta(-t, 0\ldots, 0, -x_1), \beta = (1-v^2)^{-1/2}$. Here we are taking the speed of light to be 1. The corresponding conserved quantity is

$$\Lambda_1(t) \;=\; \int_{\mathbf{R}^N} \Big\{ e(\varphi)x_1 - t\partial_1\varphi\partial_t\varphi \Big\}\, d^N x, \qquad \text{(III.18)}$$

where $e(\varphi) = \frac{1}{2}(\partial_t\varphi)^2 + \frac{1}{2}|\nabla\varphi|^2 + F(\varphi)$ is the energy density of φ. This transformation can also be written as a "rotation" in space-time ("rotation" means a rotation by an imaginary angle). Here $\Phi_\lambda(x,t) = R_\beta(x,t)$ where

$$R_\beta \;=\; \begin{bmatrix} \cosh\beta & 0 & \cdots & -\sinh\beta \\ 0 & 1 & \cdots & 0 \\ \vdots & & \vdots & \vdots \\ -\sinh\beta & 0 & \cdots & \cosh\beta \end{bmatrix}, \quad \tanh\beta = \lambda, \qquad \text{(III.19)}$$

is an $(N+1)\times(N+1)$ orthogonal matrix with respect to the inner product on Minkowski space. In fact, all Lorentz transformations can be viewed as rotations in Minkowski space (see [G], for example).

IV Virial Relations for NLW

For an arbitrary transformation group T_λ the formula (III.8) may not be true. But if we define the function g by the equation

$$g \;\equiv\; \mathcal{S}_u(v\cdot\nabla_m\varphi) \;+\; \mathcal{S}_p\cdot\nabla_m(v\cdot\nabla_m\varphi) - \nabla_m\cdot(\mathcal{S}v), \qquad \text{(IV.1)}$$

then, of course, the formula

$$\mathcal{S}_u(v\cdot\nabla_m\varphi) \;+\; \mathcal{S}_p\cdot\nabla_m(v\cdot\nabla_m\varphi) \;=\; \nabla_m\cdot(\mathcal{S}v) + g \qquad \text{(IV.2)}$$

is true for all transformation groups (and all $\varphi\in X$). When T_λ is a symmetry group of S then $g=0$ by Lemma III.5, otherwise g may not be zero. This motivates us to think of the function g as being a measure of how far T_λ is from being a symmetry of S.

As in the previous section, we assume that the independent variables are space-time variables $(x,t) \in \mathbf{R}^{N+1}$. With g as defined above, and noting that a divergence term vanishes when integrated over $\mathbf{R}^{N+1}$ (viz. the term $\nabla_m \cdot (\mathcal{S}v)$ in (IV.1)), if φ is a critical point of S, the virial relation associated to the vector field v (equation (II.2)) can be written

$$S'[\varphi](v \cdot \nabla\varphi) \;=\; \int_{\mathbf{R}^{N+1}} g\, d^N x\, dt \;=\; 0. \tag{IV.3}$$

If in addition φ solves the Euler-Lagrange equation, then using (I.4) we can write g as

$$\begin{aligned} g &= (\mathcal{S}_u - \nabla_{x,t} \cdot \mathcal{S}_p) v \cdot \nabla\varphi + \nabla_{x,t} \cdot (\mathcal{S}_p v \cdot \nabla\varphi - \mathcal{S}v) \\ &= \nabla_{x,t} \cdot (\mathcal{S}_p v \cdot \nabla\varphi - \mathcal{S}v) \\ &= \partial_t e \;+\; \nabla \cdot \mathbf{p} \end{aligned} \tag{IV.4}$$

where $e = \mathcal{S}_{p_{N+1}} v \cdot \nabla\varphi - \mathcal{S}v_{N+1}$ and $\mathbf{p} = \mathcal{S}_{\vec{p}} v \cdot \nabla\varphi - \mathcal{S}\vec{v}$ (cf. (III.10)). With $E(t) \equiv \int_{\mathbf{R}^N} e\, d^N x$ (see III.11) and setting $G(t) \equiv \int_{\mathbf{R}^N} g\, d^N x$, we see from (IV.4) that G acts as source of E;

$$\frac{d}{dt} E(t) \;=\; G(t). \tag{IV.5}$$

This corroborates the statement made above about g measuring how far T_λ is from being a symmetry group of S: If T_λ is a symmetry group of S then E as defined through (IV.4) is a conserved quantity, if T_λ is not a symmetry group of S then E varies with time at a rate determined by g.

We now consider the variational problem associated to $2\pi/\omega$-periodic solutions of NLW (equation (I.13)). Let φ be a critical point of S and Φ_λ a transformation group of $\mathbf{R}^N \times S^1_\omega$ that derives from a vector field v on the spatial domain $\mathbf{R}^N$. We will show below that

$$g = tr\, dv \left(\frac{1}{2} (\partial_t \varphi)^2 - F(\varphi) \right) + \nabla\varphi \cdot [dv - \frac{1}{2} tr\, dv\, \mathbf{1}] \nabla\varphi \tag{IV.6}$$

where dv denotes the matrix $[dv]_{i,j} = \partial v_i / \partial x_j$ so that $tr\, dv = \nabla \cdot v$. By Lemma III.5, g vanishes if T_λ is a symmetry group of NLW. For example, rotations and translations of $\mathbf{R}^N$ are symmetries of NLW. In the former case $dv \in so(N)$ (the antisymmetric $N \times N$ matrices) and in the latter $v = constant$. In both cases it follows from (IV.6) that $g = 0$.

As an example of a virial relation, consider dilations in x, $\Phi_\lambda(x) = \lambda x$, which is not a symmetry of NLW. Here the generator is the vector field $v(x) = x$ so that $dv = \mathbf{1}$ and $tr\, dv = N$. Then, if φ is a solution of NLW, from (IV.6) we calculate

$$g \ = \ N\Big(\frac{1}{2}(\partial_t\varphi)^2 - (\frac{N-2}{2N}) \mid \nabla\varphi \mid^2 - F(\varphi)\Big). \tag{IV.7}$$

Therefore, any $2\pi/\omega$-periodic solution φ of NLW satisfies the identity

$$\int_{S^1_\omega}\int_{\mathbf{R}^N} \Big\{\frac{1}{2}(\partial_t\varphi)^2 - (\frac{N-2}{2N})|\nabla\varphi|^2 - F(\varphi)\Big\}\, d^N x\, dt \ = \ 0 \tag{IV.8}$$

(cf. (IV.3)). In spatial dimensions $N = 1$ or 2 this formula can be used to rule-out the existence of periodic solutions: If the nonlinearity is such that $F(x) \leq 0$ for all x, then the only functions that satisfy (IV.8) are the constant functions. We conclude that NLW with this type of nonlinearity does not have (nontrivial) periodic solutions.

We summarize the preceding discussions. Virial relations and conservation laws are related in the sense that they are both consequences of the variation of the action S under a transformation group acting on the space of functions where S is defined. *Every* such transformation group determines a virial relation via the formula (II.1). If the transformation group happens to be a symmetry group of S, then the resulting virial relation is trivial because the integrand is a divergence. However, in this case there is a nontrivial *differential* identity: a conservation law.

Formula (IV.6) can be derived directly from equation (II.2) (cf. (IV.3)) as follows. First,

$$S[\varphi_\lambda] \ = \ \int_{S^1_\omega}\int_{\mathbf{R}^N} \Big\{ -\frac{1}{2}(\partial_t\varphi_\lambda)^2 + \frac{1}{2}|\nabla\varphi_\lambda|^2 + F(\varphi_\lambda)\Big\}\, d^N x\, dt. \tag{IV.9}$$

Making the change of variables $y = \Phi_\lambda(x)$, this becomes

$$S[\varphi_\lambda] \;=\; \int_{S^1_\omega}\int_{\mathbf{R}^N} \Big\{-\frac{1}{2}(\partial_t\varphi)^2 + \frac{1}{2}|\nabla\varphi_\lambda\circ\Phi_{-\lambda}|^2 + F(\varphi)\Big\}\, det\,(J_\lambda\circ\Phi_{-\lambda})^{-1}\, d^N y\, dt \tag{IV.10}$$

where $[J_\lambda(x)]_{i,j} = \partial\Phi^i_\lambda/\partial x_j$ is the Jacobian associated to the transformation $x \mapsto y$ and $(J_\lambda\circ\Phi_{-\lambda})^{-1}$ is the Jacobian associated to the inverse mapping $y \mapsto x$. Note that

$$\frac{\partial}{\partial\lambda}\, det\,(J_\lambda\circ\Phi_{-\lambda})^{-1}\Big|_{\lambda=0} \;=\; tr\, dv. \tag{IV.11}$$

In addition,

$$|\nabla\varphi_\lambda(x)|^2 \;=\; \nabla\varphi_\lambda(x)\cdot\nabla\varphi_\lambda(x) \;=\; A\nabla\varphi(y)\cdot\nabla\varphi(y), \quad \text{where}\;\; [A]_{i,j} = \sum_{k=1}^N \frac{\partial\Phi^i_\lambda}{\partial x_k}\frac{\partial\Phi^j_\lambda}{\partial x_k}, \tag{IV.12}$$

and

$$\frac{\partial}{\partial\lambda}\Big[A\circ\Phi_\lambda\Big]\Big|_{\lambda=0} \;=\; dv + dv^T, \tag{IV.13}$$

which together imply that

$$\frac{\partial}{\partial\lambda}|\nabla\varphi_\lambda\circ\Phi_{-\lambda}|^2\,\Big|_{\lambda=0} \;=\; 2\,\nabla\varphi\cdot\, dv\,\nabla\varphi. \tag{IV.14}$$

Therefore,

$$\begin{aligned}\frac{d}{d\lambda}S[\varphi_\lambda]\Big|_{\lambda=0} &= \int_{S^1_\omega}\int_{\mathbf{R}^N}\frac{\partial}{\partial\lambda}\Big\{-\frac{1}{2}(\partial_t\varphi)^2 + \frac{1}{2}|\nabla\varphi_\lambda\circ\Phi_{-\lambda}|^2 + F(\varphi)\Big\}det\,(J_\lambda\circ\Phi_{-\lambda})^{-1}\Big|_{\lambda=0} d^N y\, dt\\ &= \int_{S^1_\omega}\int_{\mathbf{R}^N} g\, d^N x\, dt\end{aligned} \tag{IV.15}$$

with g as in (IV.6).

References

[DFN] Dubrovin, B.A., Fomenko, A.T., Novikov, S.P. *Modern Geometry - Methods and Applications I*, Springer Verlag, New York (1992). **[G]** Goldstein, H. *Classical Mechanics*, Addison Wesley Publishing Company, Reading Massachusetts (1980). **[GF]** Gelfand, I.M., and Fomin, S.V. *Calculus of Variations*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey (1963) . **[HC-I]** Hilbert, D., Courant, R. *Methods of Mathemati cal Physics, vol. I*. John Wiley and Sons, New York (1953). **[O]** Olver, P. *Applications of Lie Groups to Differen tial Equations (first edition)*, Springer Verlag, New York (1986).