27

An extract from

General Relativity

PAPERS IN HONOUR OF J. L. SYNGE

EDITED FOR THE ROYAL IRISH ACADEMY
BY

L. O'RAIFEARTAIGH
Dublin Institute for Advanced Studies

CLARENDON PRESS · OXFORD

1972

5

INVARIANCE OF LAGRANGIAN SYSTEMS

ANDRZEJ TRAUTMAN

1. Introduction

A GENERAL lesson to be drawn from the development of the theory of relativity is that it is desirable to analyse in detail the various structures inherent in the mathematical models used to describe physical phenomena. The analysis should concentrate on the relation of these structures to experimentally verifiable statements. The presence of any structure that has no legitimate link to physical phenomena may be considered as an indication of a defective model. Einstein's refutation of absolute time is a classical example of such an analysis.

On the other hand, a good understanding of the fundamental structures underlying a theory may suggest new and fruitful generalizations. For example, the Galilean model of space–time is based on several structures, not all of which are clearly apparent in conventional expositions of classical mechanics. Among the elements of the Galilean model is a flat linear connection; it may be replaced by a more general one, thus providing a good description of Newtonian gravitation. A deep, geometrical interpretation of the gauge transformations in electrodynamics has led to interesting generalizations such as the Yang–Mills theory and to a better understanding of the Bohm–Aharonov experiments.

The intuitive notion of structure has been clarified and formalized by Bourbaki.[1] As a rule, 'rich' structures are used in physics: those of differential manifolds carrying additional geometric objects and of Hilbert spaces with preferred sets of operators. With respect to differential-geometric notions, the custom of expressing everything in terms of local coordinates prevailed for a long time. This is equivalent to working with the number spaces $\mathbf{R}^n$ ($n = 1, 2, \ldots$), and their subsets. The number spaces are endowed with many superimposed 'natural structures': $\mathbf{R}^n$ may be looked upon as a vector space, affine space, differential manifold, Lie algebra, etc. What is worse, numerous accidental identifications are associated with these spaces: an $n$tuple of real numbers defines not only a point in $\mathbf{R}^n$ but also a form on $\mathbf{R}^n$, considered as a vector space, and a vector tangent to the differential manifold $\mathbf{R}^n$ at 0. Therefore, a steady and stubborn use of coordinates makes it difficult to separate and describe the various geometrical structures associated with physical theories. An important example of such a situation is provided by the variational principles of physics. The calculus of variations, as it is usually presented, does not provide clear answers to questions such as these. What is the domain of definition of a Lagrangian? What is the set of transformations under which it is meaningful to test the invariance of an action integral? What is the geometrical significance of the Euler–Lagrange equations?

This paper is an attempt to formulate a part of the calculus of variations, as it is used in physics, in an intrinsic, geometric manner, without ever referring to local coordinates. The paper is related to an earlier article[2] and complements work done on this subject by other authors.[3],[4],[5],[6] It is now generally accepted that fibre bundles provide a natural framework to describe histories of physical systems such as the space–time development of a classical field.[6],[7],[8] In accordance with this point of view, in the next section we recapitulate some of the notions of the theory of bundles needed in later parts of the paper. Section 3 deals with generalities of the problem of invariance and Section 4 is devoted to a review of the notion of jet and differential prolongation. The last two sections, which contain the essential results of the paper, are devoted to an intrinsic characterization of the Euler–Lagrange mapping and to a geometric formulation of the Noether theorem on invariant principles of least action.[9],[10]

2. Bundles

All the objects and morphisms considered in this article belong to the category **Man** of finite-dimensional, Hausdorff, differential manifolds of class C^∞. A *mapping* is a differentiable map of one (differential) manifold into another. If $f:M \to N$ is invertible and both f and f^{-1} are mappings, then f is a *diffeomorphism*; it is called a *transformation* if $M = N$. The real line $\mathbf{R}$ is assumed to have the natural structure of a differential manifold. A mapping of a manifold into $\mathbf{R}$ is called a *function*.

A differentiable *bundle* is a surjective mapping $\pi:E \to M$ satisfying the usual condition of local triviality.[11] The set $E_x = \pi^{-1}(x)$, considered as a submanifold of E, is called the *fibre* over $x \in M$. The domain and the range of π are called the *total space* and the *base* of the bundle respectively. Sometimes it is convenient to refer to E as the bundle. A *section* of π is a mapping $\gamma:M \to E$ such that $\pi \circ \gamma = \mathrm{id}_M$. A pair (ξ,η) of mappings, $\xi:M_1 \to M_2$, $\eta:E_1 \to E_2$ such that $\pi_2 \circ \eta = \xi \circ \pi_1$ is called a *morphism* of the bundle $\pi_1:E_1 \to M_1$ into the bundle $\pi_2:E_2 \to M_2$. For any mapping $\eta:E_1 \to E_2$ there exists at most one mapping $\xi:M_1 \to M_2$ such that (ξ,η) is a morphism. In other words, if ξ exists, it is determined by η and η can be referred to as the morphism of bundles. If, in addition, ξ is a diffeomorphism and γ is a section of π_1, then $\eta \circ \gamma \circ \xi^{-1}$ is a section of π_2. An *automorphism* of a bundle is such a morphism (ξ,η) of the bundle into itself that both ξ and η are transformations.

To any section γ of $\pi: E \to M$ there corresponds the submanifold $\gamma(M)$ of E. Given a transformation $\eta: E \to E$, we may ask whether the transformed submanifold $(\eta \circ \gamma)(M)$ corresponds to a section of π. This is so if η is an automorphism of π but not otherwise, in general. By requiring that η carry sufficiently many sections into sections, we can prove that η should be an automorphism. A family Γ of sections of π is said to *separate the points* of the bundle if, for any two points $y_1, y_2 \in E$ not belonging to the same fibre $\pi(y_1) \neq \pi(y_2)$, there exists a section $\gamma \in \Gamma$ such that $\gamma \circ \pi(y_i) = y_i$, where $i = 1, 2$. The following proposition says that only automorphisms map elements of Γ into sections.

PROPOSITION 1. *If the family Γ of sections of $\pi: E \to M$ separates the points of the bundle and $\eta: E \to E$ is a transformation, then a necessary and sufficient condition for both $(\eta \circ \gamma)(M)$ and $(\eta^{-1} \circ \gamma)(M)$ to correspond to sections of π, for any $\gamma \in \Gamma$, is that η be an automorphism.*

The sufficiency of the condition is obvious; to prove that it is necessary we note that a transformation η is an automorphism if and only if

$$\pi(y_1) = \pi(y_2) \Leftrightarrow \pi \circ \eta(y_1) = \pi \circ \eta(y_2).$$

The tangent functor T maps a bundle $\pi: E \to M$ into the bundle $T\pi: TE \to TM$. Let τ_M denote the tangent bundle of a manifold M, $\tau_M: TM \to M$. The *vertical bundle* of π is the restriction of τ_E to the manifold of vertical vectors,

$$\operatorname{ver} E = \{Y \in TE | T\pi(Y) = 0\}.$$

If $\pi: E \to M$ is a bundle and $f: N \to M$ is a mapping, then the set

$$f^*(E) = \{(x,y) \in N \times E | f(x) = \pi(y)\}$$

can be made into the total space of the bundle (the *pull-back* of π by f) $f^*\pi: f^*(E) \to N$, where $f^*\pi(x,y) = x$. The mapping $\pi^*f: f^*(E) \to E$ defined by $\pi^*f(x,y) = y$ is a morphism of bundles. In particular, if $U \subset M$ is open and $f: U \to M$ is the canonical injection, then $\pi_U = f^*\pi$ is the restriction of π to $\pi^{-1}(U)$. A section of π_U is called a *local section* of π. The set of all local sections of π is denoted by Γ_π, whereas $\Gamma_\pi^c \subset \Gamma_\pi$ stands for the set of local sections with relatively compact domains. Let (ξ, η) be a morphism of $\pi_1: E_1 \to M_1$ in $\pi_2: E_2 \to M_2$. The mapping $\eta': E_1 \to \xi^*(E_2)$ defined by $\eta'(y) = (\pi_1(y), \eta(y))$ is a morphism of π_1 into $\xi^*\pi_2$, i.e. the diagram

$$\begin{tikzcd}
E_1 \arrow[r, "\eta'"] \arrow[d, "\pi_1"] & \xi^*(E_2) \arrow[r, "\pi_2{}^*\xi"] \arrow[d, "\xi^*\pi_2"] & E_2 \arrow[d, "\pi_2"] \\
M_1 \arrow[r, "\mathrm{id}"] & M_1 \arrow[r, "\xi"] & M_2
\end{tikzcd}$$

commutes and

$$\eta = \pi_2^* \xi \circ \eta'. \tag{1}$$

The bundle $\pi^* \tau_M : \pi^*(TM) = \text{hor}\, E \to E$ is called the *horizontal bundle* of $\pi : E \to M$. There is a canonical exact sequence of morphisms of vector bundles over E,

$$0 \to \text{ver}\, E \to TE \to \text{hor}\, E \to 0.$$

To any morphism $\eta : E_1 \to E$ there corresponds the morphism $\text{ver}\ \eta = T\eta\,|\text{ver}\, E_1$, of vector bundles, called the fibre or *vertical derivative* of η. Clearly, ver is a functor from the category of bundles to the category of vector bundles. If $\pi : E \to M$ is a vector bundle, then the vector space $\text{ver}_y\, E$, fibre of $\text{ver}\, E$ over $y \in E$, can be identified with $E_{\pi(y)}$, and this defines a morphism (π, ϖ) of vector bundles

$$\begin{tikzcd}
\text{ver}\, E \arrow[r, "\varpi"] \arrow[d] & E \arrow[d, "\pi"] \\
E \arrow[r, "\pi"] & M
\end{tikzcd}$$

A convenient abuse of notation consists in forgetting about ϖ and writing simply $\text{ver}\, \eta : \text{ver}\, E_1 \to E$ for the composition of ϖ with the vertical derivative of $\eta : E_1 \to E$. The pair $(\pi \circ \eta,\ \text{ver}\ \eta)$ then becomes a morphism of vector bundles.

The letter $\mathscr{L}$ is used to denote the familiar two-argument, mixed-variance functor on the category of vector spaces: if $a : A' \to A, f : A \to B$, and $b : B \to B'$ are linear maps of vector spaces, then $\mathscr{L}(a,b) : \mathscr{L}(A,B) \to \mathscr{L}(A',B')$ is defined by $\mathscr{L}(a,b)(f) = b \circ f \circ a$, and a^* is written instead of the 'adjoint' map $\mathscr{L}(a, id)$. This functor has a natural extension to the category of vector bundles; the extended functor is also denoted by $\mathscr{L}$. If $\pi_i : E_i \to M$ $(i = 1, 2)$ are two vector bundles, then the total space of $\mathscr{L}(\pi_1,\pi_2)$ is

$$\mathscr{L}(E_1,E_2) = \bigcup_{x \in M} \mathscr{L}(E_{1x},E_{2x}).$$

To a section h of $\mathscr{L}(\pi_1,\pi_2)$ there corresponds the morphism $\hat{h}$ of vector bundles, $\hat{h} : E_1 \to E_2$, defined by $\hat{h}/E_{1x} = h(x)$, for any $x \in M$. To simplify the notation, the total space of $\mathscr{L}(f_1^*\pi_1, f_2^*\pi_2)$, where $\pi_i : E_i \to M_i$ $(i = 1,2)$ are two vector bundles and $f_i : N \to M_i$ $(i = 1,2)$, is written as $\mathscr{L}_N(E_1,E_2)$. If $X : M_1 \to E_1$ and $h : N \to \mathscr{L}_N(E_1,E_2)$ are sections of π_1 and $\mathscr{L}(f_1^*\pi_1, f_2^*\pi_2)$ respectively, then $\langle X,h \rangle$ is a section of $f_2^*\pi_2$, defined by $\langle X,h \rangle(x) = \hat{h}\{X \circ f_1(x)\}$, for any $x \in N$.

3. Invariance

DEFINITION 1. *Let F be a set and f a permutation of F. An element $x \in F$ is said to be invariant with respect to f if $f(x) = x$. The element x is called an invariant of a group G of permutations of F, if for any $f \in G, f(x) = x$.*

All the notions of invariant quantities and the ideas of 'covariance' used in theoretical physics can be reduced to the simple terms defined above. To do this in any particular case we have to select an appropriate set F from the scale[1] of sets that can be constructed over any set E. The action of a group G, initially defined in E, is extended in a natural way to the sets of the scale. However, in some instances, the set F may be placed in the scale higher than we are willing to go: it is then convenient to use the notion of equivariance, which can be introduced on a lower level.

DEFINITION 2. *Let f and g be permutations of the sets M and N respectively. A map $h: M \to N$ is said to be equivariant with respect to (f,g) if $g \circ h = h \circ f$.*

Clearly, the notion of equivariance coincides with that of invariance in the set $F = \mathscr{F}(M,N)$ of all maps of M into N, the action of the pair (f,g) on $h \in F$ being given by $h \mapsto g^{-1} \circ h \circ f$.

In differential geometry and the part of theoretical physics that uses differential-geometric models, it is often possible to reduce the problems of invariance with respect to mappings to those of invariance with respect to infinitesimal transformations. To clarify the relation between invariance under finite and infinitesimal transformations, consider first the simple example of a function invariant under a one-parameter group of transformations.

A mapping of $\mathbb{R} \times M$ into M denoted by $(t,x) \mapsto \xi_t(x)$ defines a one-parameter group (ξ_t) of transformations of M if $\xi_0 = \mathrm{id}_M$ and $\xi_t \circ \xi_s = \xi_{t+s}$, for any $t,s \in \mathbb{R}$. The group (ξ_t) induces a vector field $X: M \to TM$, and

$$\partial_X f = \frac{\mathrm{d}}{\mathrm{d}t} f \circ \xi_t \bigg|_{t=0}$$

is the Lie derivative of the function $f: M \to \mathbb{R}$ with respect to X. Conversely, any vector field on M generates a local, one-parameter group of local transformations of M. Subsequently, to simplify the language, X will be said to generate a one-parameter group of transformations although the group is defined only locally, in general. For any function f on M,

$$\frac{\mathrm{d}}{\mathrm{d}t} f \circ \xi_t = \partial_X (f \circ \xi_t) = (\partial_X f) \circ \xi_t,$$

so that invariance of f with respect to ξ_t, for all $t \in \mathbb{R}$, is equivalent to the infinitesimal invariance of f under X,

$$\partial_X f = 0.$$

Let (ξ_t) and (η_t) be two groups of one-parameter transformations, generated by the vector fields X on M and Y on N, respectively. If $h:M \to N$ is a mapping, $x \in M$, then the vector.

$$Y \circ h(x) - Th \circ X(x)$$

is tangent to the curve $t \ggg \eta_t \circ h \circ \xi_t^{-1}(x)$ at $t = 0$. A mapping $V:M \to TN$ such that $\tau_N \circ V = h$ is called a *vector h-field.*[12] An example of a vector h-field is provided by $Y \circ h - Th \circ X$. It is easy to prove

PROPOSITION 2. *For any mapping $h:M \to N$ the following three conditions are equivalent.*

(a) *h is equivariant with respect to (ξ_t, η_t), for any $t \in \mathbf{R}$;*

(b) *the vector fields X and Y, induced by (ξ_t) and (η_t), are h-related, i.e.*

$$\partial_X(f \circ h) = (\partial_Y f) \circ h$$

for any function $f:N \to \mathbf{R}$;

(c) $Y \circ h - Th \circ X = 0$.

Let (ξ_t, η_t) and (ξ_t, ζ_t) be one-parameter groups of automorphisms of the bundles $\pi:E \to M$ and $\rho:F \to M$, and let X, Y, and Z be the respective generators of (ξ_t), (η_t), and (ζ_t). If $h:E \to F$ is a morphism, $\rho \circ h = \pi$, then so is $h_t = \zeta_t \circ h \circ \eta_t^{-1}$, for any $t \in R$, and the curve $t \ggg h_t(y)$ is vertical for any $y \in E$. Therefore, the range of $Z \circ h - Th \circ Y$ may be restricted to ver F. In the special case when $\pi = \mathrm{id}$, $h:M \to F$ is a section of ρ, and

(a) if $X = 0$, then $Z \circ h$ is the *variation* of h;

(b) if F is a differential prolongation[13] of M, and Z is the appropriate lift of X to F, then

$$\partial_X h = Th \circ X - Z \circ h$$

is the classical *Lie derivative* of h with respect to X.

For our purposes, it is convenient to have a somewhat more general notion of a Lie derivative. Let (ξ_t, η_t) again be a one-parameter group of automorphisms of the bundle $\pi:E \to M$ and let $\rho:F \to M$ be a vector-bundle, differential prolongation of the manifold M. To the one-parameter group (ξ_t) of transformations of M there corresponds the lifted group (ξ_t, ζ_t) of automorphisms of ρ; this group is uniquely determined by the generator Y of (η_t). By definition, the Lie derivative of a morphism $h:E \to F$ with respect to Y is

$$\partial_Y h = \frac{\mathrm{d}}{\mathrm{d}t} \zeta_t^{-1} \circ h \circ \eta_t \bigg|_{t=0}, \tag{2}$$

and we also have

PROPOSITION 3. *The morphism $h:E \to F$ is equivariant with respect to (η_t, ζ_t) for all $t \in \mathbf{R}$, if and only if $\partial_Y h = 0$.*

4. Jets and prolongations

In order to compute the curvature tensor at a point of a Riemannian space it is not enough to know the metric at that point and not necessary to consider the metric tensor field all over space. The information needed in this case coincides with that contained in the second jet of the metric tensor field at the point. With a system of local coordinates, the jet in question is determined by the values of the components of the metric tensor and their first and second derivatives. A jet bears a similar relation to these numbers as does a tensor to its components with respect to a local frame. When questions related to invariance are considered, it is desirable to distinguish between a geometric object, such as a tensor or a k-jet, and its components, which are simply real numbers.

General and precise definitions of jets may be found in the literature.[13],[14] To fix the notation, we recall the construction of the bundle $\bar{E} = J^1(E)$ of one-jets associated with a bundle $\pi : E \to M$. In the set of pointed local sections of π,

$$\Gamma_\pi^{\cdot} = \{(x,\gamma)/\Gamma_\pi \ni \gamma : U \to E, x \in U\},$$

consider the equivalence relation S defined by

$$(x,\gamma) \equiv (x',\gamma') \bmod S \Leftrightarrow T_x\gamma = T_{x'}\gamma'.$$

Let j be the canonical map of $\Gamma_\pi^{\cdot}$ on the quotient $\bar{E}$ of $\Gamma_\pi^{\cdot}$ by S,

$$j : \Gamma_\pi^{\cdot} \to \bar{E} = \Gamma_\pi^{\cdot}/S.$$

The set $\bar{E}$ admits a natural structure of differential manifold such that the projection $\rho : \bar{E} \to E$, $\rho\{j(x,\gamma)\} = \gamma(x)$ is differentiable and $\bar{\pi} = \pi \circ \rho : \bar{E} \to M$ is a differentiable bundle. The elements of $\bar{E}$ are called one-jets of π and $\bar{\pi} : \bar{E} \to M$ is the *first jet prolongation* of $\pi : E \to M$. To any local section $\gamma \in \Gamma_\pi$ there corresponds its first jet prolongation $\bar{\gamma} \in \Gamma_{\bar{\pi}}$, defined by $\bar{\gamma}(x) = j(x,\gamma)$. In a similar manner, one defines the bundle of two-jets, $\bar{\bar{\pi}} : \bar{\bar{E}} \to M, \bar{\rho} : \bar{\bar{E}} \to \bar{E}, \bar{\bar{\gamma}}$, etc.†

To make jets more familiar, consider two affine spaces M and N, let $\vec{M}$ and $\vec{N}$ be the corresponding vector spaces of translations, and take π to be the product bundle, $E = M \times N$. A section γ of π can be represented by a function $f : M \to N$, with

$$\gamma(x) = (x, f(x)).$$

The bundles of one-jets and two-jets have as their total spaces, respectively,

$$\bar{E} = M \times N \times \mathscr{L}(\vec{M}, \vec{N}),$$

$$\bar{\bar{E}} = M \times N \times \mathscr{L}(\vec{M}, \vec{N}) \times \mathscr{L}_s^2(\vec{M}, \vec{N}),$$

† In §§ 5 and 6 pull-backs by ρ and $\bar{\rho}$ will frequently occur. It is convenient to neglect them in the formulae and write h instead of $h \circ \rho$ or $h \circ \rho \circ \bar{\rho}$, etc.

and the jet prolongation of γ may be described by

$$\bar{\gamma}(x) = (x, f(x), f'(x)),$$
$$\bar{\bar{\gamma}}(x) = (x, f(x), f'(x), f''(x)),$$

where

$$f': M \to \mathscr{L}(\overrightarrow{M}, \overrightarrow{N})$$

is the first derivative of f, and so on.

Any automorphism η of π can be prolonged, in a natural way, to an automorphism $\bar{\eta}$ of $\bar{\pi}$ and to an automorphism $\bar{\bar{\eta}}$ of $\bar{\bar{\pi}}$. The mapping $\bar{\eta}: \bar{E} \to \bar{E}$ is defined by

$$\bar{\eta}\{j(x,\gamma)\} = \overline{\eta \circ \gamma \circ \xi^{-1}} \circ \xi(x),$$

so that

$$\bar{\eta} \circ \bar{\gamma} \circ \xi^{-1} = \overline{\eta \circ \gamma \circ \xi^{-1}},$$

and similarly for $\bar{\bar{\eta}}$. If η_1 and η_2 are two automorphisms of π, then $\overline{\eta_1 \circ \eta_2} = \bar{\eta}_1 \circ \bar{\eta}_2$. Therefore, if (η_t) is a one-parameter group of automorphisms of π, then $(\bar{\eta}_t)$ bears a similar relation to $\bar{\pi}$. This makes it possible to prolong any vector field Y on E, projectable with respect to π, to a vector field $\bar{Y}$ on $\bar{E}$, ρ-related to Y. Namely, if Y generates (η_t), then $\bar{Y}$ is the vector field induced by $(\bar{\eta}_t)$. In addition, we can associate with Y a vector ρ-field $\tilde{Y}: \bar{E} \to \text{ver } E$ defined as follows. $\tilde{Y}\{j(x,\gamma)\}$ is the vector tangent to the vertical curve $t \twoheadrightarrow \eta_t \circ \gamma \circ \xi_t^{-1}(x)$ at $t = 0$. Similarly, $\tilde{\bar{Y}}: \bar{\bar{E}} \to \text{ver } \bar{E}$ is defined by reference to the vertical curve $t \twoheadrightarrow \bar{\eta}_t \circ \bar{\gamma} \circ \xi_t^{-1}(x)$ in $\bar{E}$. It follows directly from the definitions that

$$\tilde{Y} \circ \bar{\gamma} = Y \circ \gamma - T\gamma \circ X$$

and

$$\tilde{\bar{Y}} \circ \bar{\bar{\gamma}} = \bar{Y} \circ \bar{\gamma} - T\bar{\gamma} \circ X.$$

For any bundle $\pi: E \to M$, consider the vertical bundles ver E of π and ver $\bar{E}$ of $\bar{\pi}$, and construct the pull-backs $\rho^*(\text{ver } E)$ and

$$\mathscr{L}_{\bar{E}}(TM, \text{ver } E) = \mathscr{L}\{\rho^*(\text{hor } E), \rho^*(\text{ver } E)\}.$$

Let $\bar{y} \in \bar{E}$, put $y = \rho(\bar{y})$ and $x = \pi(y)$, and consider $(\bar{y}, f) \in \{\bar{y}\} \times \mathscr{L}(T_x M, \text{ver}_y E) =$ fibre of $\mathscr{L}_{\bar{E}}(TM, \text{ver } E)$ over $\bar{y}$, where f is given by

$$f(u) = u(g)v_y, \quad u \in T_x M,$$

$g: M \to \mathbf{R}$ is a function vanishing at x, and $v_y \in \text{ver}_y E$. Let v be a vertical vector field on E, containing v_y, i.e. such that $v(y) = v_y$. The map i,

$$i(\bar{y}, f) = \overline{(g \circ \pi)v}(\bar{y}),$$

can be extended by linearity to a morphism of vector bundles over $\bar{E}$. Let $s = (T\rho|\text{ ver } \bar{E})'$, where the prime has the same significance as in eqn (1), with $\rho:\bar{E}\to E$ replacing $\xi:M_1\to M_2$. The morphisms i and s yield a generalization of the *jet-bundle exact sequence*, well known in the theory of vector bundles.[14]

PROPOSITION 4. *For any differentiable bundle $\pi:E\to M$, there is a canonical exact sequence of vector bundles over $\bar{E}$,*

$$0\to \mathscr{L}_{\bar{E}}(TM,\ \mathrm{ver}\ E)\xrightarrow{i}\mathrm{ver}\ \bar{E}\xrightarrow{s}\rho^*(\mathrm{ver}\ E)\to 0, \tag{3}$$

where ρ is the natural projection of $\bar{E}=J^1(E)$ onto E.

Let h be a morphism of $\pi:E\to M$ into the bundle $\Lambda^k T^*M$ of k-forms over M. For any section γ of π, $h\circ\gamma$ is a field of k-forms on M and its exterior derivative $\mathrm{d}(h\circ\gamma)$ is a field of $(k+1)$-forms on M. There exists a unique morphism

$$Dh:\bar{E}\to\Lambda^{k+1}T^*M$$

of bundles over M, such that

$$Dh\circ\bar{\gamma}=\mathrm{d}(h\circ\gamma),$$

for any section $\gamma:M\to E$. Similarly, if h is a morphism of $\bar{\pi}:\bar{E}\to M$ into $\Lambda^k T^*M$, then there is a morphism $Dh:\bar{\bar{E}}\to\Lambda^{k+1}T^*M$, such that $Dh\circ\bar{\bar{\gamma}}=\mathrm{d}(h\circ\bar{\gamma})$.

5. Lagrangians and the Euler map

Using the idea of jet, we can define a Lagrangian without having to introduce local coordinates. Clearly, what is usually called a Lagrangian density depending on derivatives up to mth order is, in fact, a morphism of the bundle of m-jets into the bundle of n-forms, where n is the dimension of the base space. For simplicity, we shall restrict ourselves to first-order Lagrangians.

DEFINITION 3. *A Lagrangian system (π,L) consists of a bundle $\pi:E\to M$ and a morphism ('Lagrangian')*

$$L:\bar{E}\to\Lambda^n T^*M$$

of bundles over M, $n=$ dim M. *The function from Γ^c_π* to $\mathbf{R}$, *defined by*

$$\gamma\twoheadrightarrow\int_U L\circ\gamma,$$

where U is the (relatively compact) domain of the local section $\gamma\in\Gamma^c_\pi$, is called the action of (π,L).

The classical rule for a change of variables in a multiple integral may be written as

$$\int_U L\circ\bar{\gamma}=\int_{\xi(U)}\xi_*\circ L\circ\bar{\gamma}\circ\xi^{-1}, \tag{4}$$

where $\xi : U \to \xi(U) \subset M$ is a diffeomorphism and ξ_* is the 'Jacobian' morphism induced by ξ in the bundle $\Lambda^n T^*M$,

$$\xi_*(\mathrm{d}x^1 \wedge \ldots \wedge \mathrm{d}x^n) = \mathrm{d}(x^1 \circ \xi^{-1}) \wedge \ldots \wedge \mathrm{d}(x^n \circ \xi^{-1}).$$

To arrive at the Euler–Lagrange equations associated with (π, L), we consider the one-parameter group (η_t) of vertical automorphisms of π, generated by a vector field Y. The 'varied' local section $\gamma_t = \eta_t \circ \gamma$ has the same domain as γ, and

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_U L \circ \overline{\eta_t \circ \gamma} \bigg|_{t=0} = \int_U \langle \bar{Y}, \mathrm{ver}\, L \rangle \circ \bar{\gamma}, \tag{5}$$

where $\bar{Y}$ is the jet prolongation of Y to $\bar{E}$.

Let $VL : \bar{E} \to \mathscr{L}_{\bar{E}}(\mathrm{ver}\, E, \Lambda^{n-1} T^*M)$ be the composition of morphisms

$$\bar{E} \xrightarrow{\mathrm{ver}\, L} \mathscr{L}_{\bar{E}}(\mathrm{ver}\, \bar{E}, \Lambda^n T^*M) \xrightarrow{i^*} \mathscr{L}_{\bar{E}}(\mathscr{L}_{\bar{E}}(TM, \mathrm{ver}\, E), \Lambda^n T^*M)$$
$$\xrightarrow{\rfloor} \mathscr{L}_{\bar{E}}(\mathrm{ver}\, E, \Lambda^{n-1} T^*M),$$

where i^* is the adjoint of the morphism i occurring in the exact sequence (3) and $\rfloor$ is an obvious extension of the canonical inner product

$$\rfloor : TM \times \Lambda^k T^*M \to \Lambda^{k-1} T^*M.$$

The difference

$$\langle \bar{\bar{Y}}, \mathrm{ver}\, L \rangle - D\langle Y, VL \rangle : \bar{\bar{\bar{E}}} \to \Lambda^n T^*M$$

vanishes for any $\bar{\bar{y}} \in \bar{\bar{\bar{E}}}$ such that $Y(\bar{\bar{y}}) = 0$. (Compare footnote in § 4.) This implies the existence of an *Euler–Lagrange mapping*

$$[L] : \bar{\bar{\bar{E}}} \to \mathscr{L}_{\bar{\bar{E}}}(\mathrm{ver}\, E, \Lambda^n T^*M),$$

which is a morphism of bundles over $\bar{\bar{E}}$ and is such that

$$\langle \bar{\bar{Y}}, \mathrm{ver}\, L \rangle = D\langle Y, VL \rangle + \langle Y, [L] \rangle \tag{6}$$

for any vertical vector field Y on E.

If Y vanishes on the boundary of $\gamma(U)$, then eqn (5) becomes

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_U L \circ \overline{\eta_t \circ \gamma} \bigg|_{t=0} = \int_U \langle Y \circ \gamma, [L] \circ \bar{\bar{\gamma}} \rangle,$$

and there follows the classical result that the Euler–Lagrange equations

$$[L] \circ \bar{\bar{\gamma}} = 0$$

are necessary and sufficient for the action to be stationary at γ.

The *Euler map* $[\ \ ] : L \gg [L]$ has a simple behaviour under automorphisms of π. For any such automorphism η, the Euler–Lagrange mapping associated with the transformed Lagrangian $\xi_*^{-1} \circ L \circ \bar{\eta}$ is given by[15]

$$[\xi_*^{-1} \circ L \circ \bar{\eta}] = \mathscr{L}(\text{ver } \eta, \xi_*^{-1}) \circ [L] \circ \bar{\bar{\eta}}, \tag{7}$$

where $\mathscr{L}(\text{ver } \eta, \xi_*^{-1})$ is an automorphism of the vector bundle $\mathscr{L}_{\bar{\bar{E}}}(\text{ver } E, \Lambda^n T^*M)$, induced by η.

The set of all Lagrangians on π has a natural structure of vector space; the Euler map is linear and its kernel contains all 'exact' Lagrangians, i.e. Lagrangians of the form

$$L = DK,$$

where

$$K: E \to \Lambda^{n-1} T^*M$$

is a morphism of bundles over M. It should not be difficult to relate the quotient Ker[ ]/Im D to the topological invariants of the bundle π.

For a classical, mechanical system with n degrees of freedom, $\pi = pr_1$ and $E = \mathbf{R} \times Q$, where Q is an n-dimensional differential manifold (configuration space). In this case, both ver E and $\bar{E}$ may be identified with $\mathbf{R} \times TQ$, and

$$VL: \mathbf{R} \times TQ \to \mathbf{R} \times T^*Q,$$

if smoothly invertible, is called the *Legendre transformation*. The 'accidental identifications' inherent in the Lagrangian formulation of classical mechanics account for much of the simplicity of this theory, as compared to the classical theory of fields.

6. Symmetries, invariant transformations, and conserved quantities

Roughly speaking, a symmetry of a physical system is a correspondence associating to any possible history of the system another such history of the same system. This definition is too general and not precise enough to be useful. Before giving a description of symmetries adapted to the Lagrangian formalism presented in the preceding section, let us consider the simple case of a classical mechanical system with n degrees of freedom.[16]

Let Q again denote the configuration space. The cotangent bundle $\tau_Q^*: T^*Q \to Q$ (phase space) carries a canonical one-form $\alpha: T(T^*Q) \to \mathbf{R}$ $\alpha | T_p(T^*Q) = p \circ T\tau_Q^*$. The two-form $\beta = \mathrm{d}\alpha$ is non-singular and defines a symplectic structure on T^*Q. For any function G on T^*Q there is a vector field $G^\#$ on T^*Q such that

$$G^\# \lrcorner\, \beta = \mathrm{d}G.$$

If H is the Hamiltonian function, then $H^\#$ generates the one-parameter group (φ_t) of motions. Let (ψ_s) be another one-parameter group, generated by a vector field Z on T^*Q. The curve $t \mapsto \psi_s \circ \varphi_t(p)$ is a motion of $H^\sharp$, for any $p \in T^*Q$, if and only if

$$\psi_s \circ \varphi_t = \varphi_t \circ \psi_s. \tag{8}$$

Therefore, Z generates a one-parameter group of symmetries if eqn (8) holds for any $s,t \in \mathbf{R}$, or equivalently if

$$\partial_Z H^\sharp = 0.$$

On the other hand, a function G is a conserved quantity, or a constant of the motion, if

$$G \circ \varphi_t = G \qquad \text{for any } t \in \mathbf{R},$$

i.e. if

$$\partial_{H^\sharp} G = 0.$$

From the antisymmetry of the Poisson bracket, $\partial_{H^\sharp} G = \{H,G\} = -\partial_{G^\sharp} H$, it follows that the group (ψ_s) generated by the vector field $G^\sharp$, corresponding to a conserved quantity, preserves the value of H,

$$\partial_{G^\sharp} H = 0, \tag{9}$$

and, therefore, may be referred to as a group of invariant transformations. Moreover, the group (ψ_s) is symplectic, $\partial_{G^\sharp}\beta = 0$, and eqn (9) implies that

$$\partial_{G^\sharp} H^\sharp = 0.$$

Invariant transformations are symmetries† but the converse is not true, as may be seen by considering the group generated by $G^\sharp$, where G is a function canonically conjugate to the Hamiltonian, $\{H,G\} = 1$.

An important feature of the Hamiltonian formalism of classical mechanics is that it allows all conserved quantities and symmetries to be described and easily found. In the Lagrangian formulation, it seems natural to restrict oneself to point transformations, i.e. to those automorphisms of TQ that arise by lifting of a diffeomorphism from Q to the tangent bundle. Clearly, in general, the class of point transformations does not contain all symmetries. As an example, consider a particle constrained to move along a line so that $Q = \mathbf{R}$ and $TQ = \mathbf{R}^2$. The one-parameter group $\psi_s : TQ \to TQ$ defined by $\psi_s(q,\dot q) = (q_s,\dot q_s)$, with

$$\left.\begin{aligned} q_s &= q \cos s + \dot q \sin s \\ \dot q_s &= -q \sin s + \dot q \cos s \end{aligned}\right\}, \tag{10}$$

describes a symmetry of the Lagrangian $L(q,\dot q) = \frac{1}{2}(\dot q^2 - q^2)$ because it maps the solution

$$t \mapsto a \cos(t - t_0) \tag{11}$$

† Abraham[16] uses the word 'symmetry' in a narrower sense: essentially for what we call an invariant transformation.

of the equation of motion into the solution $t \gg\!\!\rightarrow a\cos(t+s-t_0)$. The map (10) is not a point transformation and as a result it transforms the lift $\bar{c}:\mathbf{R}\to TQ$ of a curve $c:\mathbf{R}\to Q$ into a curve $\psi_s\circ\bar{c}$ that is not a lift, unless c happens to be of the particular form (12). A small perturbation of the Lagrangian will not only break its symmetry under eqn (10) but also make it impossible to apply, in a meaningful way, this transformation to solutions of the equation of motion.

After this digression, we return to a more general case and consider a Lagrangian system (π,L).

DEFINITION 4. *An automorphism η of π, $\pi\circ\eta=\xi\circ\pi$, is said to be an invariant transformation of (π,L) if, for any local section $\gamma\in\Gamma^c_\pi$, the value of the action is invariant by η,*

$$\int_U L\circ\bar{\gamma} = \int_{\xi(U)} L\circ\overline{\eta\circ\gamma\circ\xi^{-1}}. \tag{12}$$

By comparing eqn (12) and eqn (4), and taking into account the arbitrariness of γ, we obtain

PROPOSITION 5. *A necessary and sufficient condition for an automorphism η of π to be an invariant transformation of (π,L) is that L be equivariant with respect to $(\xi_*,\bar{\eta})$,*

$$L\circ\bar{\eta} = \xi_*\circ L. \tag{13}$$

This, in turn, implies that the Euler–Lagrange mapping $[L]$ itself is equivariant with respect to a prolongation of the invariant transformation. By comparing equations (7) and (13), we obtain

$$[L] = \mathscr{L}(\text{ver}\,\eta,\xi_*^{-1})\circ[L]\circ\bar{\bar{\eta}}. \tag{14}$$

This holds also in a more general case, when L, instead of being equivariant with respect to $(\xi_*,\bar{\eta})$, changes by an exact differential. To cover this situation, we introduce

DEFINITION 5. *An automorphism η of π is called a generalized invariant transformation if there exists a morphism $K:E\to\Lambda^{n-1}T^*M$ of bundles over M such that*

$$\xi_*^{-1}\circ L\circ\bar{\eta}+DK = L. \tag{15}$$

Remembering that $[DK]=0$, we see that eqn (14) holds for a generalized invariant transformation η and, therefore,

$$\text{if } [L]\circ\bar{\bar{\gamma}}=0, \qquad \text{then } [L]\circ\overline{\overline{\eta\circ\gamma\circ\xi^{-1}}}=0,$$

thus proving

PROPOSITION 6. *A generalized invariant transformation is a symmetry.*

Consider a one-parameter group (η_t) of automorphisms of π, generated by the vector field Y on E, and let X be the vector field on M induced by (ξ_t), with $\pi \circ \eta_t = \xi_t \circ \pi$. For any local section γ we can write

$$\partial_X(L \circ \bar{\gamma}) = \frac{\mathrm{d}}{\mathrm{d}t}\, \xi_{t*}^{-1} \circ L \circ \bar{\gamma} \quad \xi_t \bigg|_{t=0}$$
$$= \frac{\mathrm{d}}{\mathrm{d}t}\, \xi_{t*}^{-1} \circ L \circ \bar{\eta}_t \circ \overline{\eta_t^{-1} \circ \gamma \circ \xi_t} \bigg|_{t=0}$$
$$= (\partial_Y L) \circ \bar{\gamma} - \langle \tilde{\tilde{Y}} \circ \bar{\bar{\gamma}}, \operatorname{ver} L \circ \bar{\gamma} \rangle,$$

where $\partial_Y L$ is the Lie derivative of L with respect to Y defined by eqn (2).

The last equation imples that

$$\partial_Y L = D(X \rfloor L) + \langle \tilde{\tilde{Y}}, \operatorname{ver} L \rangle,$$

where the notation is that of § 4. A computation analogous to the one that led to eqn (6) gives the fundamental formula

$$\partial_Y L = D(X \rfloor L + \langle \tilde{Y}, VL \rangle) + \langle \tilde{Y}, [L] \rangle. \tag{16}$$

If (η_t) is a group of generalized invariant transformations, then, according to Definition 9, there exists a morphism $Z : E \to \Lambda^{n-1} T^*M$ such that

$$\partial_Y L + DZ = 0, \tag{17}$$

and eqn (16) becomes

$$D(X \rfloor L + \langle \tilde{Y}, VL \rangle + Z) + \langle \tilde{Y}, [L] \rangle = 0.$$

This proves

PROPOSITION 7. *Let (η_t) be a one-parameter group of generalized invariant transformations generated by Y and characterized by eqn* (17). *If γ is a solution of the Euler–Lagrange equation, $[L] \circ \bar{\bar{\gamma}} = 0$, then there is a conservation law*

$$\mathrm{d}(J \circ \bar{\gamma}) = 0$$

for the current

$$J = X \rfloor L + \langle \tilde{Y}, VL \rangle + Z$$

associated with (η_t).

The Noether–Bessell–Hagen equation (17) is equivalent to a system of partial differential equations of the first order, linear and homogeneous in (Y,Z). The set of solutions of (17) forms a Lie algebra with a bracket given by

$$[(Y_1,Z_1), (Y_2,Z_2)] = ([Y_1, Y_2], \partial_{Y_1} Z_2 - \partial_{Y_2} Z_1).$$

The structure of this Lie algebra is an important characteristic of the Lagrangian system.

ACKNOWLEDGEMENTS

Work on this paper was begun in September 1970 at NORDITA in Copenhagen and completed in the spring of 1971 at the Enrico Fermi Institute, University of Chicago. I am grateful to C. Møller and S. Chandrasekhar for their hospitality at these institutions and to Miss D. Elbert for her help in preparing the manuscript.

REFERENCES

1. Bourbaki, N. *Théorie des ensembles*. Hermann, Paris (1957).
2. Trautman, A. Noether equations and conservation laws. *Communs math. Phys.* **6**, 248–61 (1967).
3. Śniatycki, J. On the geometric structure of classical field theory in Lagrangian formulation. *Proc. Camb. phil. Soc. math. phys. Sci.* **68**, 475–84 (1970).
4. Komorowski, J. A. modern version of the E. Noether theorems in the calculus of variations. *Studia math.* part I: **29**, 261–73 (1968); part II: **32**, 181–90 (1969).
5. Krupka, D. Lagrangian theory in fibred manifolds. *Rep. Math. Phys., Toruń* **2**, 121 (1971).
6. Hermann, R. *Vector bundles in mathematical physics*, vol. 1. Benjamin, New York (1970).
7. Lichnerowicz, A. Espaces fibrés et espace–temps. *Gen. Relativity Gravitation* **1**, 235–45 (1971).
8. Trautman, A. Fibre bundles associated with space–time. *Rep. Math. Phys., Toruń* **1**, 29–62 (1970).
9. Noether, E. Invariante Variationsprobleme. *Nachr. Ges. Wiss. Göttingen* 235 (1918).
10. Bessel-Hagen, E. Über die Erhaltungssätze der Elektrodynamik. *Math. Annln* **84**, 258–76 (1921).
11. Lang, S. *Introduction to differentiable manifolds*. Interscience, New York (1962).
12. Sikorski, R. *Wstep do geometrii różniczkowej*. PWN, Warszawa (1971).
13. Ehresmann, C. Les prolongements d'une variété différentiable. *C. r. hebd. Séanc. Acad. Sci., Paris* **233**, 598, 777, 1081 (1951); **234**, 1028, 1424 (1952).
14. Palais, R. S. *Seminar on the Atiyah-Singer index theorem*. Princeton University Press (1965).
15. de Donder, T. *Théorie invariantive du calcul des variations*. Gauthier-Villars, Paris (1930).
16. Abraham, R. *Foundations of mechanics*. Benjamin, New York (1967).