

Momentum Maps and Classical Fields

Part I: Covariant Field Theory

Mark J. Gotay *
Department of Mathematics
University of Hawai'i
Honolulu, Hawai'i 96822, USA
gotay@math.hawaii.edu

James Isenberg †
Department of Mathematics
University of Oregon
Eugene, Oregon 97403, USA
jim@newton.uoregon.edu

Richard Montgomery ‡
Department of Mathematics
University of California at Santa Cruz
California 95064, USA
rmont@math.ucsc.edu

Jerrold E. Marsden §
Control and Dynamical Systems 107-81
California Institute of Technology
Pasadena, California 91125, USA
marsden@cds.caltech.edu

With the collaboration of

Jędrzej Śniatycki **Philip B. Yasskin**

November 18, 1997
(Revised December 12, 2002)

*Research partially supported as a Ford Foundation Fellow, by NSF grants DMS 88-05699, 92-22241, 96-23083, and 00-72434, and grants from ONR/NARC.

†Partially supported by NSF grant PHY 00-99373.

‡Partially supported by NSF grant DMS 87-02502.

§Partially supported by NSF grant DMS 96-33161.

Contents

1	Introduction	1
I—Covariant Field Theory		16
2	Multisymplectic Manifolds	16
2A	The Jet Bundle	17
	Examples	18
2B	The Dual Jet Bundle	21
	Examples	24
3	Lagrangian Dynamics	26
3A	The Covariant Legendre Transformation	27
3B	The Cartan Form	27
3C	The Euler–Lagrange Equations	28
	Examples	32
4	Covariant Momentum Maps and Noether’s Theorem	35
4A	Jet Prolongations	36
4B	Covariant Canonical Transformations	37
4C	Covariant Momentum Maps	39
	Examples	40
4D	Symmetries and Noether’s Theorem	43
	Examples	48
Interlude I—On Classical Field Theory		52
References		56
Table of Contents for Parts II–V		65

1 Introduction

This is the first paper of a five part work in which we study the Lagrangian and Hamiltonian structure of classical field theories with constraints. Our goal is to explore some of the connections between initial value constraints and gauge transformations in such theories (either relativistic or not). To do this, in the course of these five papers, we develop and use a number of tools from symplectic and multisymplectic geometry. Of central importance in our analysis is the notion of the “energy-momentum map” associated to the gauge group of a given classical field theory. We hope to demonstrate that, as illustrated and explained below, many different and apparently unrelated facets of field theories can be thereby tied together and understood in an essentially new way.

Our research is motivated by previous exhaustive studies of several standard examples following the methods of Choquet–Bruhat, Lichnerowicz, Dirac–Bergmann, and Arnowitt–Deser–Misner. In particular, if one performs an initial value (or “canonical”) analysis of Einstein’s theory of gravity (either alone or coupled to Dirac, electromagnetic, or other fields), Yang–Mills theory, relativistic fluid theories, bosonic string theory, topological field theories, and so forth, one finds that they share the following striking features. (See Fischer and Marsden [1979a,b], Isenberg and Nester [1980], Künzle and Nester [1984], Bao, Isenberg, and Yasskin [1985], Batlle, Gomis, and Pons [1986], and Horowitz [1989] and references therein for discussions of these and other field theories.)

1. *The Euler–Lagrange equations are underdetermined; that is, there are not enough evolution equations to propagate all the field components. This is because the theory has gauge freedom. The corresponding gauge group is known at the outset.*

As a consequence, some of the fields do not have their evolution determined by the Euler–Lagrange equations. These “kinematic fields” have no physical significance. Amongst the remaining fields are the “dynamic fields” ψ (with conjugate momenta ρ); these are the fields that have physical meaning.

The situation is further complicated in that the initial data $(\psi(0), \rho(0))$ cannot be freely specified.

2. *The Euler–Lagrange equations are overdetermined in that they include constraints*

$$\Phi^i(\psi, \rho) = 0 \tag{1.1}$$

on the choice of Cauchy data.

Equations (1.1) typically comprise an elliptic system whose component equations can be of variable order. For simplicity we suppose that all constraints are “first class” in the terminology of Dirac [1964]. (The case when second class constraints are present will be discussed as the need arises.)

It is remarkable that the Euler–Lagrange equations are simultaneously underdetermined and overdetermined. The next feature ties these two properties together.

3. *The constraint functions Φ^i generate gauge transformations of the dynamic fields via the canonical symplectic structure on the space of Cauchy data.*

Thus, the presence of (first class) initial value constraints is intimately related to the gauge freedom of the theory. Since gauge transformations are part of the dynamics and since the constraints generate gauge transformations, one expects the functions Φ^i to appear in the Hamiltonian. Indeed, in these examples, we find:

4. *The Hamiltonian (with respect to a slicing of spacetime, usually taken to consist of Cauchy surfaces) has the form*

$$H = \int_{\Sigma} \sum_i \alpha_i \Phi^i(\psi, \rho) d\Sigma \quad (1.2)$$

depending linearly on the “atlas fields” α_i . Here Σ is a hypersurface modeling a typical member of the given slicing.

The atlas fields, discussed in more detail later, are closely related to the kinematic fields mentioned above in item **1**. (Being sections of an associated bundle, they may be regarded as certain combinations of the kinematic fields and elements of the Lie algebra of the gauge group.) The atlas fields are arbitrarily specifiable; intuitively, one can view them as governing the propagation of coordinates on a Cauchy surface (and on bundles of fields over a Cauchy surface) in time—hence the name. Their importance derives from the fact that they “drive” the entire gauge ambiguity of the theory (cf. the next item).

Introduce an L_2 inner product on the space of Cauchy data (with corresponding adjoint denoted by $*$), and let \mathbb{J} be an almost complex structure compatible with both the given inner product and the symplectic form. Let λ be a parameter (“time”) labeling the slicing mentioned above. From (1.2), it follows that

5. *The evolution equations for the dynamic fields (ψ, ρ) take the “adjoint form”*

$$\frac{d}{d\lambda} \begin{pmatrix} \psi \\ \rho \end{pmatrix} = \mathbb{J} \cdot \sum_i [D\Phi^i(\psi(\lambda), \rho(\lambda))]^* \alpha_i(\lambda). \quad (1.3)$$

Equations (1.3) are typically a *hyperbolic* system; hence, for fixed $\alpha_i(\lambda)$, they uniquely determine the evolution of the dynamic fields and their conjugate momenta from given initial data. Here, the atlas fields $\alpha_i(\lambda)$ are to be specified *ab initio*; this amounts to a gauge-fixing. Equations (1.3) display, in the clearest and most concise way, the interrelations between the dynamics, the initial value constraints, and the gauge ambiguity of a theory.

The constraint equations (1.1) are preserved by the evolution equations (1.3). Of course,

6. *The covariant Euler–Lagrange equations are equivalent, modulo gauge transformations, to the combined evolution and constraint equations (1.3) and (1.1).*

This result, together with items 4 and 5, constitute the theoretical background necessary to analyze the structure of the solution space of the Euler–Lagrange equations. For the theories mentioned above, we have:

7. *The space of solutions of the field equations is not necessarily a smooth manifold. It may have quadratic singularities occurring exactly at those solutions that are symmetric (in the sense of being invariant under the action of a nondiscrete subgroup of the gauge group).*

See Fischer, Marsden, and Moncrief [1980], Arms, Marsden, and Moncrief [1981, 1982], Marsden [1988], Arms [1977, 1981], Arms, Gotay, and Jennings [1990], Arms, Cushman, and Gotay [1991], Sjamaar and Lerman [1991], and Interlude III for more information on the meaning of item 7. These results are an outgrowth of the extensive literature on linearization stability, which concerns the success or failure of the linearization process as far as predicting the valid directions of perturbation theory, as well as that on symplectic reduction, which is the means of eliminating those degrees of freedom of a system that are associated with symmetries.

Noether’s theorem (aka the “first Noether theorem”) and the formal Dirac–Bergmann analysis of constraints do much to predict and explain features 1–6. However, we wish to go further and provide (realistic) sufficient conditions which guarantee that they *must* occur in a given field theory. In this work we provide such criteria for the first six conditions and lay the groundwork needed for the analysis of the seventh. One of the goals is thus to derive the “adjoint formalism” for classical field theories.

A key tool throughout this work is the “energy-momentum map,” which we define in Part II. A recurrent theme in this work is that *the energy-momentum map encodes essentially all the dynamical information carried by a given classical field theory*: its Hamiltonian, its initial value constraints, its gauge freedom, and even its stress-energy-momentum tensor. Indeed, in Part III we shall prove

The Energy-Momentum Theorem. *The constraints (1.1) are given by the vanishing of the energy-momentum map associated to the gauge group of the theory.*

Thus the constraint functions Φ^i are “components” of the energy-momentum map. Item 4 then shows that the same is true for the Hamiltonian.

Certain of these results are known in specific instances and have to some extent acquired “folk theorem” status. But we wish to prove that they hold rather generally. In this regard we emphasize that the energy-momentum map does more than merely provide reformulations of known results; it leads to fundamental insights into the foundations of physical theories. In view of their importance—and ubiquity—in classical field theory, one is tempted to elevate the statement “momentum maps are everything” to a general principle!

To understand the role played by the energy-momentum map, recall the two traditional approaches to classical field theory, which we may characterize

as “group-theoretical” and “canonical.” The group-theoretical approach is concerned with the gauge covariance of a given system, and is based upon Noether’s theorem. It operates on the (spacetime) covariant level, and is usually phrased in Lagrangian terms. (See Trautman [1967], Olver [1986], and references contained therein for relevant background.) In contrast, the canonical approach is used to analyze the initial value problem in a Hamiltonian setting; the basic tool here is the Dirac–Bergmann theory of constraints. Such an analysis requires that the theory be space + time decomposed relative to a fixed choice of Cauchy surface in spacetime. (See Dirac [1964], Hanson, Regge, and Teitelboim [1976], Gotay, Nester, and Hinds [1978], Isenberg and Nester [1980], as well as the review by Sundermeyer [1982] for detailed accounts.)

The group-theoretical and canonical aspects of a mechanical system are linked by the momentum map (see, for example, Abraham and Marsden [1978] and Marsden and Ratiu [1999]). One would like to have an analogous connection in field theory which, in particular, relates gauge symmetries to initial value constraints. But—and this is a crucial point—the *standard notion of a momentum map associated to a symplectic group action usually cannot be carried over to spacetime covariant field theory*. The reason is that the gauge group typically does not act in the space + time decomposed (or “instantaneous”) framework, because spacetime diffeomorphisms move Cauchy surfaces. Because the Hamiltonian formalism is only defined relative to a fixed Cauchy surface, there is no possibility of obtaining a momentum map for the gauge group in the usual sense on the instantaneous phase space.¹

The prime example where this situation arises is Einstein’s theory of vacuum gravity, in which case the gauge group is the spacetime diffeomorphism group. The only remnants of this group on the instantaneous level are, in the ADM (Arnowitt, Deser and Misner [1962]) language, the superhamiltonian \mathfrak{H} and supermomenta \mathfrak{J} which are interpreted as the generators of temporal and spatial deformations of a Cauchy surface, respectively. But these deformations do not form a group, nor are \mathfrak{H} and \mathfrak{J} components of a momentum map. (See Kuchař [1973] for further details.) More generally, this difficulty appears for any theory which is “parametrized” in that its gauge group “includes” the spacetime diffeomorphism group (in an appropriate sense).

This circumstance forces us to work on the covariant level. But it is then necessary to construct a covariant counterpart to the instantaneous Hamiltonian formalism. In the spacetime covariant (or “multisymplectic”) framework we develop here—which is an extension and refinement of the formalism of Kijowski and Szczyrba [1976]—the gauge group *does* act. This enables us to define the notion of a covariant (or “multi-”) momentum map on the corresponding covariant (or “multi-”) phase space. This covariant momentum map satisfies Noether’s theorem, and has all the properties one would expect. It remains to correlate this object with the dynamics once a Cauchy surface has been singled out. To this end, a careful study of the mechanics of the space + time

¹ It is possible that a momentum map in the sense of groupoids may be appropriate for this setting. (See, for instance, Mikami and Weinstein [1988].) This is, however, a subject for future research.

decomposition shows that the covariant momentum map induces an “energy-momentum” map Φ on the instantaneous phase space. It is not a genuine momentum map, as there is no group action in this context; nonetheless, Φ is the instantaneous “shadow” of a covariant momentum map, and this is enough. The energy-momentum map is the crucial object that reflects the gauge transformation covariance of a classical field theory in the instantaneous formalism. In the ADM formulation of gravity, $\Phi = -(\mathfrak{H}, \mathfrak{J})$, so that the superhamiltonian and supermomenta are the components of the energy-momentum map.

The energy-momentum map thus synthesizes the group-theoretical and canonical approaches to classical field theory. Consonant with this observation, we view the gauge group of a theory as being fundamental, and suppose it to be known *ab initio*. This is certainly the case in the standard examples cited previously. Our plan is then to use the energy-momentum map associated to the gauge group of the theory to study, and indeed derive, items **1–6** above. This shift in viewpoint proves to be surprisingly fruitful, enabling us to obtain deeper results than by the traditional methods alone.

The theoretical underpinnings of the adjoint formalism for classical field theories consist of four components:

- I. a covariant analysis of field theories;
- II. a space + time decomposition of the covariant formalism followed by an initial value analysis of field theories;
- III. a study of the relations between gauge symmetries and initial value constraints; and
- IV. the actual derivation of the adjoint formalism.

The first two components constitute the foundation for the last two, where most of the hard work is concentrated. We present each of these four components as a paper in this series, starting with the covariant analysis. The fifth paper in this series applies our formalism to Einstein’s theory of gravity. Before beginning Part I, we now comment briefly on each of the five parts, in turn. (For the convenience of the reader, we append a table of contents of Parts II–V at the end of this paper.)

In Part I we develop some of the basic theory of classical fields from a spacetime covariant viewpoint. Throughout we restrict attention to first order theories; these are theories whose Lagrangians involve no higher than first derivatives of the fields.

We begin in §2 of Part I with a study of the covariant Lagrangian and Hamiltonian formalisms, mirroring the approach usually taken in classical mechanics. Let X be an $(n+1)$ -dimensional parameter space, which in applications is usually spacetime. We suppose that we are given a bundle Y over X whose sections are the fields of interest; Y is the *covariant configuration bundle*. The field-theoretic analogue of the tangent bundle of mechanics is the first jet bundle J^1Y of Y . By taking an appropriate dual of J^1Y we obtain the space Z

which is, roughly speaking, the covariant cotangent bundle of the bundle Y . As such, it carries a canonical $(n+1)$ -form Θ . The *multisymplectic* $(n+2)$ -form $\Omega = -d\Theta$ is the covariant generalization of the symplectic form in Hamiltonian mechanics. The pair (Z, Ω) is called *multiphase space*.

Next, in §3, we introduce a Lagrangian density \mathcal{L} on J^1Y . The covariant Lagrangian and Hamiltonian formalisms are then related by the *covariant Legendre transformation* $\mathbb{F}\mathcal{L} : J^1Y \rightarrow Z$ which is defined, as in mechanics, by means of the fiber derivative of \mathcal{L} . Pulling the form Ω on Z back by $\mathbb{F}\mathcal{L}$ we obtain the *Cartan form* on J^1Y (also variously known as the ‘‘Hamilton–Cartan’’ and ‘‘Poincaré–Cartan’’ form). We use it to construct an intrinsic formulation of the variational calculus and the Euler–Lagrange equations, cf. Goldschmidt and Sternberg [1973] and García [1974].

Our approach to covariant field theory has a number of advantages over other formulations. For instance, unlike those of Kijowski and Tulczyjew [1979], Ragionieri and Ricci [1981], Günther [1987], Sardanashvily [1993], and Kanatchikov [1998], our multisymplectic structure and corresponding covariant Hamiltonian formalism are intrinsically defined, independent of other constructs (such as a background connection). Other formalisms such as those of Goldschmidt and Sternberg [1973] and Śniatycki [1984], although intrinsic, do not carry multisymplectic structures at all. The approach we take here is also promising for higher order field theories, which are plagued by a number of ambiguities; see Gotay [1991a,b] and references therein. Perhaps most importantly, our approach allows us to define covariant momentum maps for groups that act nontrivially on X . Such group actions also arise naturally in fluid dynamics (Moreau [1982]) and continuum mechanics (Lew, Marsden, Ortiz, and West [2002]) in relation to taking ‘‘horizontal variations’’ of the action.

In §4 we discuss symmetries and conservation laws of classical field theories in terms of ‘‘covariant momentum maps,’’ generalizing the concept of momentum map familiar from mechanics. If one introduces a Lie group \mathcal{G} (possibly infinite-dimensional) with Lie algebra \mathfrak{g} that acts on (Z, Ω) in an appropriate way, then one can define a *multi-* or *covariant momentum map* $J : Z \rightarrow \mathfrak{g}^* \otimes \Lambda^n Z$ that intertwines the group action with the multisymplectic structure via the equation

$$\xi_Z \lrcorner \Omega = \mathbf{d}\langle J, \xi \rangle,$$

where $\langle J, \xi \rangle$ denotes J paired with $\xi \in \mathfrak{g}$, ξ_Z is the infinitesimal generator on Z of the one-parameter group generated by ξ , and \mathbf{d} is the exterior derivative. It often happens that such group actions are lifted from actions on Y , in which case one has an explicit formula for J ; see §4C. Then in §4D we prove Noether’s theorem, which provides the basic relation between symmetries of the Lagrangian and conserved quantities. Our multisymplectic formulation of Noether’s theorem is specifically designed with later applications to the relationship between constraints and the momentum map in mind. Here we also state and prove a converse to Noether’s theorem, subject to a certain transitivity assumption on the group action.

All the discussion in Part I is spacetime covariant. In Part II we use *space + time decompositions* to reformulate classical field theories as infinite-dimensional dynamical systems. The transition from the multisymplectic to the instantaneous formalism once a Cauchy surface (or a slicing by Cauchy surfaces) has been chosen is essential to define the energy-momentum map and thence to cast the field dynamics into adjoint form.

Consider an n -dimensional Cauchy surface $\Sigma \subset X$ (assumed compact and boundaryless). In the instantaneous context, the configuration space for field dynamics is the space \mathcal{Y}_Σ of sections of Y restricted to Σ , and the phase space is $T^*\mathcal{Y}_\Sigma$ with its canonical symplectic form. We emphasize that this symplectic structure lives on the space of Cauchy data for the evolution equations, and *not* on the space of solutions thereof (as in, e.g., Zuckerman [1987] and Crnković and Witten [1987]). This is an important distinction, as the former space is typically better behaved than the latter (cf. item 7 as well as the discussion in Horowitz [1989]). It also obviates the problem of constructing a differentiable structure on the space of solutions, as was attempted by Kijowski and Szczyrba [1976].

In Part II, the main result of §5 is that the multisymplectic structure on Z induces the canonical symplectic structure on $T^*\mathcal{Y}_\Sigma$. To bridge the gap between the covariant and instantaneous formalisms, we introduce an intermediate space \mathcal{Z}_Σ , the space of sections of Z restricted to Σ . The multisymplectic $(n+2)$ -form Ω on Z induces a presymplectic 2-form Ω_Σ on \mathcal{Z}_Σ by integration over Σ . Reducing \mathcal{Z}_Σ by the kernel of Ω_Σ produces a symplectic manifold $\mathcal{Z}_\Sigma/(\ker \Omega_\Sigma)$, which we prove is canonically symplectomorphic to $T^*\mathcal{Y}_\Sigma$.

This shows how to space + time decompose covariant multisymplectic phase spaces into instantaneous symplectic phase spaces. Next, in §§6A–D, we perform a similar decomposition of dynamics using the notion of slicings. An important result in this section states that the dynamics is compatible with the space + time decomposition in the following sense: Solutions of the instantaneous Hamilton equations correspond directly to solutions of the covariant Euler–Lagrange equations and vice versa. This fact—a prerequisite for establishing item 6—is usually taken for granted in the literature. Here we provide a proof, assuming only mild regularity assumptions that are satisfied in cases of interest.

An important tool throughout this work is the *initial value analysis*. One of the hallmarks of field theories is that not all of the Euler–Lagrange equations necessarily describe the temporal evolution of fields; some of them may impose constraints on the choice of initial data. The constraints that are first class reflect the gauge symmetry of the theory and eventually will be related to the vanishing of various (energy-) momentum maps. Thus, it is important to fully analyze the roles of constraints and gauge transformations in dynamics and the relations between them. To this end, we utilize the symplectic version of the Dirac–Bergmann treatment of degenerate Hamiltonian systems (patterned after Gotay, Nester, and Hinds [1978], cf. also Sundermeyer [1982]). This formalism, which we summarize in §6E, yields an essentially complete understanding of items 1–3, and lays the groundwork for correlating the first class constraints with the gauge group of the theory via the energy-momentum map.

In §7 we space + time decompose multimomentum maps associated to covariant group actions, thereby correlating these objects with momentum mappings (in the usual sense) in the instantaneous formalism. Thus, suppose that \mathcal{G} is a group of automorphisms of Y . As emphasized earlier, \mathcal{G} usually does not act on \mathcal{Y}_Σ or on $T^*\mathcal{Y}_\Sigma$ as it need not stabilize Σ . Nonetheless, by integrating over Σ , the covariant momentum map J gives rise to a map $\mathcal{Z}_\Sigma \rightarrow \mathfrak{g}^*$, which can then be shown to project to a well-defined map $\mathcal{E}_\Sigma : \mathcal{P}_\Sigma \rightarrow \mathfrak{g}^*$, where the “primary constraint set” \mathcal{P}_Σ is the image of $T\mathcal{Y}_\Sigma$ in $T^*\mathcal{Y}_\Sigma$ under the instantaneous Legendre transform. This map \mathcal{E}_Σ is the *energy-momentum map*. It is typically not a true momentum map. However, if one considers only the subgroup \mathcal{G}_Σ of \mathcal{G} that stabilizes the Cauchy surface (and that therefore acts in the instantaneous formalism), then \mathcal{E}_Σ restricts to a genuine momentum map for the \mathcal{G}_Σ -action on $T^*\mathcal{Y}_\Sigma$ that corresponds in relativity to the supermomenta, without the inclusion of the superhamiltonian. In the general case \mathcal{E}_Σ corresponds to both \mathfrak{H} and \mathfrak{J} .

Parts I and II provide the requisite background for the study of items 4–6, which we carry out in Parts III and IV. Part III, in particular, focuses on establishing a relationship between the initial value constraints on the one hand, and the energy-momentum map corresponding to the gauge transformations on the other. (See the Energy-Momentum Theorem.) We begin the study of this relationship in §8 of Part III, in which we characterize the “gauge group” of a given classical field theory. Surprisingly, this notion, although familiar and intuitive, has never been precisely defined in any generality. It is—for the moment—to be distinguished from the notion of “instantaneous gauge transformation” à la Dirac that was alluded to above. Of course, one of the main goals of Part III is to show that these two notions of “gauge” are in fact the same. This will follow once item 3 and the Energy-Momentum Theorem are established.

The *gauge group* \mathcal{G} is taken to be a subgroup of $\text{Aut}(Y)$, the group of automorphisms of the bundle Y . It must satisfy certain properties, viz.:

- (i) \mathcal{G} acts by *symmetries*, for example, the Lagrangian density \mathcal{L} is \mathcal{G} -equivariant.
- (ii) \mathcal{G} is *localizable*, which intuitively means that elements of \mathcal{G} can be “turned off” in the “future” or the “past.” This property is crucial; it is what distinguishes a gauge group from a mere symmetry group.
- (iii) \mathcal{G} is the largest such subgroup of $\text{Aut}(Y)$.

In §8A we give a principal bundle construction of \mathcal{G} .

One should bear in mind, however, that some field theories may have no gauge freedom (or initial value constraints) at all; the Klein–Gordon field on a noninteracting background spacetime is a simple example of such a system. Theories like this can be subsumed into our formalism by requiring them to be *parametrized*, in the sense that \mathcal{G} maps onto $\text{Diff}(X)$ (or at least a localizable subgroup thereof) under the natural projection $\text{Aut}(Y) \rightarrow \text{Diff}(X)$. Thus, the gauge groups of parametrized theories are *always* nontrivial. Many field theories are “already” parametrized (e.g., the Polyakov string, topological field theories,

gravity). In any case, this requirement involves no loss of generality, since a theory that is not parametrized can often be made so, for example, by simply introducing a metric on X and treating it as an auxiliary variable (which may or may not be *variational* in the sense that it is to be varied in the action principle to obtain Euler–Lagrange equations), or else by coupling it to gravity. Kuchař [1973] gives an alternate way of parametrizing a theory. Henceforth we assume that the systems under consideration are parametrized. (Although our formalism is best suited to parametrized theories, much of it—with appropriate modifications—will apply to “background” theories as well. These points are discussed further in the text; see also Interlude I.)

We start building towards the main results of Part III by establishing the following result, which is basic to the entire development. Fix a hypersurface Σ and an element $\zeta \in \mathfrak{g}$ with the property that its associated spacetime vector field ζ_X is transverse to Σ . (Such a vector field exists since the theory is parametrized.) The Lie algebra element ζ may be thought of as defining an “evolution direction”; it is analogous to the “lapse” and “shift” in relativity. In §7F we show that the instantaneous Hamiltonian density $H_{\Sigma, \zeta}$ on \mathcal{P}_Σ corresponding to the evolution direction $\zeta \in \mathfrak{g}$ is given by

$$H_{\Sigma, \zeta} = -\langle \mathcal{E}_\Sigma, \zeta \rangle. \quad (1.4)$$

Thus the instantaneous Hamiltonian is a component of \mathcal{E}_Σ (hence the appellation *energy-momentum map*). It follows that \mathcal{E}_Σ actually generates the dynamics of the system. This result is familiar in the ADM formulation of gravity: using the fact that $\mathcal{E}_\Sigma = -(\mathfrak{H}, \mathfrak{J})$ and setting $\zeta = (N, \mathbf{M}) \in \mathfrak{X}(X)$, (1.4) reduces to

$$H_{\Sigma, (N, \mathbf{M})} = N\mathfrak{H} + \mathbf{M} \cdot \mathfrak{J}.$$

Equation (1.4) forms the basis for item 4; in this regard, the fact that $H_{\Sigma, \zeta}$ is manifestly linear in ζ presages its linearity in the atlas fields, discussed later for general field theories.

Next, in §9, we prove the **Vanishing Theorem** (aka the “second Noether theorem”), which states that on any solution of the Euler–Lagrange equations, the multimomentum map integrated over a hypersurface is zero. This result follows from Noether’s theorem, \mathcal{G} -equivariance, and localizability. The Vanishing Theorem implies that the energy-momentum map must vanish on admissible Cauchy data for the evolution equations. (The fact that \mathcal{E}_Σ must *vanish*, as opposed to merely being constant, is a direct consequence of localizability.)

With these results in hand, in §10 we establish the **Energy-Momentum Theorem**, which asserts that for any field theory which is first class and satisfies certain technical conditions, the set \mathcal{C}_Σ of admissible Cauchy data for the evolution equations is exactly the zero level set of the energy-momentum map associated to the gauge group:

$$\mathcal{C}_\Sigma = \mathcal{E}_\Sigma^{-1}(0). \quad (1.5)$$

The set \mathcal{C}_Σ is known as the “final constraint set” in Dirac’s terminology. More succinctly, this theorem says that *the initial value constraints are given by the*

vanishing of the energy-momentum map. Although often quoted in the literature, to our knowledge no proof of this fact has ever been given. It may at first appear that (1.5) is “just” the Vanishing Theorem, but this is not the case. Certainly the Vanishing Theorem implies that the vanishing of the energy-momentum map yields initial value constraints, which can then be shown to be first class. But it is far from obvious that in fact *all* such constraints arise in this fashion. Moreover, the proper context for the second Noether theorem is the multisymplectic formalism described above. Only in this setting is one able to obtain constraints that are associated to gauge transformations that move Cauchy surfaces in spacetime, such as the superhamiltonian constraint in general relativity. It should be observed that other covariant Hamiltonian formalisms are deficient in this regard, for example, that of Kijowski and Tulczyjew [1979].

One corollary of the Energy-Momentum Theorem is that since $\mathcal{G} \neq \emptyset$, first class constraints are *always* present in parametrized theories (in which all fields are variational). Thus, with this proviso, items **1–3** are never vacuous. Another corollary is that the covariant and instantaneous notions of gauge coincide; more precisely, when space + time decomposed, the gauge group generates exactly the instantaneous gauge transformations in the sense of the Dirac–Bergmann constraint theory. This indicates that the above definition of \mathcal{G} is “correct.”

There are calculational advantages that accrue from this theorem as well. For instance, one can always compute the initial value constraints of a given field theory according to the Dirac–Bergmann procedure. But the ensuing calculations can be quite involved, and the results are not always easy to interpret. An analysis based on the gauge group and the energy-momentum map, on the other hand, is often substantially simpler—and has the advantage of attaching a group-theoretical interpretation to the constraints. A case in point is again provided by ADM gravity. If one merely carries out the Dirac–Bergmann analysis, one obtains the superhamiltonian and supermomentum constraints, but it is entirely unclear from this analysis what their geometric significance is. Only by “hindsight” is one able to correlate these constraints with temporal and spatial deformations of a Cauchy surface, reflecting the 4-dimensional diffeomorphism gauge freedom of the theory. The problem is that the spacetime diffeomorphism group is “hidden” once one performs a space + time decomposition—a necessary prerequisite to the initial value analysis. But our procedure yields these constraints rather routinely, and their geometric meaning is clear from the beginning. (See Part V for the application of our techniques to gravity.) Thus, one has a method for identifying the final constraint set that does not in principle require one to invoke the entire constraint preservation algorithm. The criteria which must be checked are not always easy to verify but, in any case, the theorems of §10 do guarantee an equivalence between the results of the Dirac–Bergmann algorithm and the results of an analysis based on the energy-momentum map of a given gauge group action.

The Energy-Momentum Theorem is really a statement about (first class) *secondary* constraints, as \mathcal{E}_Σ is only defined on the primary constraint set $\mathcal{P}_\Sigma \subset T^*\mathcal{Y}_\Sigma$. In §11 we show that (first class) *primary* constraints can be recovered in much the same way, but the proof is more difficult and requires

different techniques. For the primaries the relevant object is not \mathcal{G} but rather a certain foliation $\dot{\mathcal{G}}_\Sigma$ of \mathcal{Y}_Σ derived from \mathcal{G} ; intuitively, $\dot{\mathcal{G}}_\Sigma$ consists of the “time derivatives” of elements of the subgroup \mathcal{G}_Σ which stabilizes Σ . This foliation has a “momentum map” $\dot{\mathcal{J}}_\Sigma$ and, provided zero is a regular value, the primary constraints are obtained by equating $\dot{\mathcal{J}}_\Sigma$ to zero. It is possible—with some additional assumptions on the nature of the gauge group—to realize this foliation as the orbits of a genuine subgroup of \mathcal{G}_Σ ; this is the subject of §12B.

We have thus far seen that both the Hamiltonian and the initial value constraints are “components” of the energy-momentum map. In Interlude II we show, following Gotay and Marsden [1992], that the stress-energy-momentum tensor of a given collection of fields arises in a similar manner, again demonstrating the capacity of the multimomentum map to unify disparate concepts into a single geometric entity. The stress-energy-momentum tensor $\mathfrak{T}^\mu{}_\nu$ so obtained satisfies a generalized version of the classical “Belinfante–Rosenfeld formula” (Belinfante [1940], Rosenfeld [1940]), and hence naturally incorporates the canonical stress-energy-momentum tensor and the “correction terms” that are necessary to make the latter gauge-covariant. Furthermore, in the presence of a metric on X , our $\mathfrak{T}^{\mu\nu}$ coincides with the Hilbert stress-energy-momentum tensor and hence is automatically symmetric.

Part IV derives the adjoint formalism for field theories. This is the capstone of our work, where all the various aspects of a classical field theory meld into a coherent whole. The adjoint formalism is the starting point for investigations into the structure of the space of solutions of the Euler–Lagrange equations, linearization stability, quantization, and related questions.

To accomplish this objective, in §12 we distinguish two types of field variables that play an especially important role in the initial value analysis. The *dynamic fields* are those whose evolution from given initial data is determined by well-posed evolutionary equations. These are the fields that carry the dynamical content of the theory and with which the adjoint formalism is concerned. By way of contrast, the evolution of the *kinematic fields* is determined not by the Euler–Lagrange equations, but rather by the way the spacetime X and the covariant configuration bundle Y are sliced. In fact, the kinematic fields are closely related to the first class primary constraints and their existence can be directly attributed to the presence of gauge symmetries.

The space of dynamic fields is $\mathcal{D}_\Sigma = \mathcal{Y}_\Sigma/\dot{\mathcal{G}}_\Sigma$. The kinematic fields are abstractly realized as sections of this quotient; this definition makes sense as the primary constraints are components of the momentum map associated to the foliation $\dot{\mathcal{G}}_\Sigma$, so their canonically conjugate fields are parametrized by $\dot{\mathcal{G}}_\Sigma$. By a slight generalization of the cotangent bundle reduction theorem (Abraham and Marsden [1978], Kummer [1981], Marsden [1992]), the instantaneous phase space $T^*\mathcal{Y}_\Sigma$ can then be cut down to the space $T^*\mathcal{D}_\Sigma$ of dynamic fields ψ and their conjugate momenta ρ . (Physically, this reduction consists of imposing the primary constraints and then factoring out the kinematic fields.)

The goal is to project the dynamics on the primary constraint set \mathcal{P}_Σ in $T^*\mathcal{Y}_\Sigma$ to parametric dynamics on $T^*\mathcal{D}_\Sigma$. The parameters will be certain combinations

of the kinematic fields and elements of \mathfrak{g} which we call “atlas fields.”

Unfortunately, the energy-momentum map \mathcal{E}_Σ on \mathcal{P}_Σ does not directly project to $T^*\mathcal{D}_\Sigma$, as it involves the kinematic fields. But \mathcal{E}_Σ , thought of as a map $\mathcal{P}_\Sigma \times \mathfrak{g} \rightarrow \mathbb{R}$, also depends upon $\xi \in \mathfrak{g}$. By algebraically lumping the kinematic fields together with elements of the Lie algebra of the gauge group in a “ \mathcal{G}_Σ -equivariant” fashion, one forms the nondynamic **atlas fields** α_i . Their abstract definition turns out to be rather delicate, and necessitates some technical assumptions regarding the action of the gauge group on Y . In any case, the atlas fields are constructed in such a way that \mathcal{E}_Σ depends only upon (ψ, ρ) and the α_i . The energy-momentum map then can be pushed down to a function Φ_Σ on $T^*\mathcal{D}_\Sigma$ depending upon the parameters α_i . This function Φ_Σ is the “reduced” energy-momentum map. We observe that the Energy-Momentum Theorem remains valid in this context, since the secondary constraints restrict only (ψ, ρ) —the kinematic fields being correlated with the primary constraints. Combining these results with (1.4) then leads to item 4.

The essential content of item 4 is not so much that the Hamiltonian can be expressed in terms of the constraints, as noteworthy as that is. (Indeed, this is already apparent from (1.4).) Rather, it is the fact that the Hamiltonian is **linear** in the atlas fields. This is vital for the **adjoint form** (1.3) of the evolution equations, which now follows readily from (1.2) by writing Hamilton’s equations explicitly on $T^*\mathcal{D}_\Sigma$.

The salient features of the adjoint form (1.3) are:

- The constraints generate the dynamics, via the appearance of the energy-momentum map in (1.3).
- The atlas fields “drive” the entire gauge ambiguity of the theory. This is a remarkable result, especially when viewed from the standpoint of Dirac constraint theory: no matter how complex the cascade of constraints, the gauge freedom generated by the totality of first class constraints is encapsulated in the fields α_i (which are closely linked to the first class *primaries*).
- The gauge freedom of the theory is completely subsumed in the choice of slicing. This is convenient computationally, as the form of the adjoint equations (1.3) is ideally suited for studying evolution relative to various slicings.
- The dynamics is independent of the choice of slicing, which is a consequence of the preceding remark.

These statements are familiar in the case of the ADM formulation of gravity, but our results are more general and include theories which have “internal” (i.e., not spacetime-based) gauge freedom, such as Yang–Mills theory.

The independence of the adjoint formalism on the choice of slicing together with the results of §6 give item 6. The adjoint formalism is now ready to be used for such purposes as the singularity analysis of the solution space, as we have mentioned.

The results of this work apply to a large class of systems including most field theories of current interest. To bolster this claim, throughout the work we provide a number of illustrations of the formalism as we develop it. These examples include particle mechanics, Maxwell’s theory of electromagnetism, a topological field theory, and bosonic string theory. In Part V we present (the Palatini variant of) Einstein’s theory of gravity as a stand-alone example.

All these examples have first order Lagrangians and, except for Palatini gravity and the topological field theory, all of them involve only constraints which are first class in the sense of Dirac. Indeed, many of the theorems as stated in this work are restricted to field theories with these properties. However, we do believe that the essential thrust of this work holds for more general classical field theories. Second class constraints, for example, should not cause trouble. The main effect of the presence of second class constraints is the failure of equality in (1.5); we are only guaranteed that $\mathcal{C}_\Sigma \subset \mathcal{E}_\Sigma^{-1}(0)$. As a consequence, one can no longer compute the final constraint set using the energy-momentum map alone. However, the strong tie between first class constraints and gauge transformations should remain, and one should still be able to provide theoretical support for items 1–7, including the adjoint form (1.3), for such theories. Note that one way to deal with second class constraints is to eliminate them via a Dirac bracket construction (Śniatycki [1974]); while this presents no problem in principle, it is sometimes unsatisfactory from a computational standpoint. Another way to handle second class constraints might be based on the work of Lusanna [1991, 1993], which relates them to transformations that leave the Lagrangian *quasi*-invariant. Note also that in well-known field theories with second class constraints—for example, the Einstein–Dirac theory (see Bao, Isenberg, and Yasskin [1985], and Bao [1984]) and the KdV equation (see Gotay [1988])—those constraints do not introduce undue difficulties.

We also expect—modulo some technical difficulties—that the formalism developed here can be straightforwardly applied to higher order field theories. Such a generalization is necessary in particular to directly treat Einstein’s theory of gravity, since the Hilbert Lagrangian is second order in the metric. In fact, the multisymplectic setup and its space + time decomposition have already been extended to field theories of any order in Gotay [1988, 1991a,b]; see also Kouranbaeva and Shkoller [2000]. However, for third or higher order systems the Cartan form is not uniquely determined by the Lagrangian (Kolář [1984]). This circumstance bears on an important fact which underlies our analysis; viz. for first (or second) order theories, a symmetry of the Lagrangian is also a symmetry of the Cartan form. (This is essential for Noether’s theorem among other things.) Thus in the higher order case it is necessary to determine whether there exist \mathcal{G} -invariant Cartan forms. The work of Muñoz [1985] may be helpful in this context.

We believe that one further generalization should hold. Namely, the principal ideas of this work should apply to cases in which one is naturally led to consider a Poisson, as opposed to a symplectic structure, such as in certain formulations of general relativistic perfect fluids and plasmas (cf. Bao, Marsden, and Walton [1985] and references therein). It could be useful to develop

a covariant analogue of Poisson manifolds (perhaps along the lines of Marsden, Montgomery, Morrison, and Thompson [1986]) and to develop a version of covariant reduction that could start with the multisymplectic ideas of this work and produce Poisson structures as in the nonrelativistic case. (See, for example, Marsden, Weinstein, Ratiu, Schmid, and Spencer [1983] and Marsden [1992] for a summary of the nonrelativistic reduction theory, and Holm [1985] for indications of how the relativistic theory should proceed.) Recent results on reduction for field theories can be found in Castrillon and Marsden [2002].

Finally, we remark that multisymplectic geometry may be used for other interesting purposes besides the investigation of the intrinsic structure of classical field theories. In particular, Marsden and Shkoller [1999] and Marsden, Patrick, and Shkoller [1998] have shown that Bridges' [1997] treatment of water waves and their instabilities may be understood in multisymplectic terms and that one can build numerical integrators using these ideas. One can also do continuum mechanics from the multisymplectic point of view as in Marsden, Pekarsky, Shkoller, and West [2001]; this leads to a natural way of incorporating 'constraints' such as incompressibility. Covariant techniques are also proving useful in understanding the dynamical structure of the gravitational field; see Ashtekar, Bombelli, and Reula [1991] and Anco and Tung [2002] for some results in this direction.

Acknowledgments

We thank Judy Arms, David Bao, Paul Dedecker, Christian Günther, Marvin Ortel, Gunter Schwarz, Jim Stasheff, and Wiktor Szczyrba for their helpful comments. We thank the University of California Committee on Research and Cornell University which helped defray the typing costs. We also thank MSRI at Berkeley, whose hospitality allowed us to frequently collaborate during the 1988–1989 symplectic geometry year.

This project began in Calgary in 1979 and continued in Aix-en-Provence, Annapolis, Berkeley, Boston, Boulder, Brno, Canberra, Champaign-Urbana, Clausthal-Zellerfeld, College Station, Eugene, Hamburg, Hämelschenburg, Honolulu, Houston, Ithaca, Logan, Los Alamos, Minneapolis, Montréal, Palo Alto, Paris, Pasadena, Philadelphia, Potsdam, San Diego, San Luis Obispo, Santa Barbara, Santa Cruz, Seattle, Stockholm, Sydney, Toronto, Waterloo, Washington, and in the Friendly Skies. It has been a long journey, and over the years it has been gratifying that a number of ideas that originated here (such as the notions of the dual jet bundle and the covariant momentum map) have found their way into the standard mathematical physics literature. However, the main results of this work, such as those pertaining to the link between the covariant and the instantaneous formulations of dynamics, the Energy-Momentum Theorem, and atlas fields and the adjoint formalism, as outlined in the Introduction, have not previously been available.

We thank the people who have been patient with us during the writing of this work, especially our long-suffering wives, children, and relatives (who thought

it would *never* end). We appreciate both our readers and the typesetters, Sue Knapp, Marnie McElhiney, Wendy McKay, June Meyerman, Teresa Wild, Esther Zack, and Arlene Baxter, Sean Brennan, Jo Butterworth, David Mostardi, and Margaret Pattison of MSRI. This work is dedicated to graduate students who may undertake to read it, and to the spirit of G. Gimmsy, may she rest in peace.

I—COVARIANT FIELD THEORY

In this part we develop some of the basic theory of classical fields from a covariant viewpoint. This is done in the framework of multisymplectic geometry, which is analogous to the geometry of the cotangent bundle in classical non-relativistic mechanics—which we often call particle mechanics (even though it also includes continuum mechanics). We also develop the geometry of the jet bundle and the Euler–Lagrange equations, which is analogous to the geometry of the tangent bundle and the Lagrange equations of motion in the case of particle mechanics. The final section of this part discusses conservation laws and Noether’s theorem from the point of view of covariant momentum maps, again generalizing the concept of momentum map familiar from particle mechanics.

2 Multisymplectic Manifolds

There are two major approaches to the symplectic description of classical field theory, the covariant (or “multisymplectic”) formalism and the instantaneous (or “3+1”) framework. In the latter approach dynamics is described in terms of the *infinite*-dimensional space of fields at a given instant of time, whereas in the former dynamics is phrased in the context of the *finite*-dimensional space of fields at a given event in spacetime. Each formalism has its own advantages.

In this and the next several sections, we summarize those aspects of the multisymplectic formalism and its relation to the 3+1 framework that will be important for subsequent developments. There are many recent references for the multisymplectic formalism, such as Dedecker [1953, 1977], Gawędzki [1972], Goldschmidt and Sternberg [1973] (see also Guillemin and Sternberg [1977, 1984]), Kijowski [1973, 1974], García [1974], Kijowski and Szczyrba [1975, 1976], Kijowski and Tulczyjew [1979], Aldaya and de Azcárraga [1980a], Kupershmidt [1980], Kosmann–Schwarzbach [1981], Ragionieri and Ricci [1981], Kastrup [1983], Śniatycki [1970a, 1984], de León and Rodrigues [1985], Günther [1987], Gotay [1991a, b], Sardanashvily [1993], Kanatchikov [1998], Marsden, Patrick, and Shkoller [1998], and Castrillon and Marsden [2002]. Historically, much credit should also be given to Cartan [1922], De Donder [1930], Weyl [1935] and Lepage [1936]. Our presentation here is self-contained, but not exhaustive; topics not needed for our main results as well as alternative approaches have been omitted.

Throughout this work, all manifolds and maps are assumed to be of class C^∞ . Our conventions follow Misner, Thorne, and Wheeler [1973] to a large extent.

2A The Jet Bundle

Let X be an oriented manifold, which in many examples is spacetime, and let $\pi_{XY} : Y \rightarrow X$ be a finite-dimensional fiber bundle called the **covariant configuration bundle**. The fiber $\pi_{XY}^{-1}(x)$ of Y over $x \in X$ is denoted Y_x and the tangent space to X at x is denoted T_xX , etc. The physical fields will be sections of this bundle, which is the covariant analogue of the configuration space in classical mechanics.

We shall develop classical field theory in parallel to classical mechanics. First we discuss the field-theoretic analogues of the tangent and cotangent bundles. The role of the tangent bundle is played by J^1Y , the **first jet bundle** of Y . We recall its definition. We say that two local sections ϕ_1, ϕ_2 of Y **agree to first order** at $x \in X$ if each is defined in some neighborhood of x and if their first order Taylor expansions are equal at x (in any coordinate system). In other words, they agree to first order if and only if they have the same target $y = \phi_1(x) = \phi_2(x)$ and if their linearizations $T_x\phi_1$ and $T_x\phi_2$ at x coincide as linear maps $T_xX \rightarrow T_yY$. Agreement to first order defines an equivalence relation on the set of all local sections; the resulting set of equivalence classes is the first jet bundle J^1Y . Saunders [1989] contains much information on jet bundles.

It is convenient to use the natural identification of J^1Y with the *affine* bundle over Y whose fiber above $y \in Y_x$ consists of those linear mappings $\gamma : T_xX \rightarrow T_yY$ satisfying

$$T\pi_{XY} \circ \gamma = \text{Identity on } T_xX. \tag{2A.1}$$

The map γ corresponds to $T_x\phi$ for a local section ϕ . The vector bundle underlying this affine bundle is the bundle whose fiber over $y \in Y_x$ is the space $L(T_xX, V_yY)$ of linear maps of T_xX to V_yY , where

$$V_yY = \{v \in T_yY \mid T\pi_{XY} \cdot v = 0\} \tag{2A.2}$$

is the fiber above y of the **vertical subbundle** $VY \subset TY$. Note that for $\gamma \in J_y^1Y$, we have the splitting

$$T_yY = \text{im } \gamma \oplus V_yY. \tag{2A.3}$$

Remark 1. One can view sections of J^1Y over Y as *Ehresmann connections* on Y . (See Hermann [1975] or Marsden, Montgomery, and Ratiu [1990] for a brief review of Ehresmann connections.) The horizontal-vertical splitting (2A.3) defines the connection associated to such a section. If Y is a principal G -bundle, then $J^1Y \rightarrow Y$ is the bundle whose equivariant sections are connections in the usual sense of principal bundles. ² ♦

² In this case J^1Y , as a principal bundle over the connection bundle $CY = J^1Y/G$, itself has a canonical connection. For the bundle $Q \times \mathbb{R} \rightarrow Q$ regarded as an \mathbb{R} -bundle, $J^1(Q \times \mathbb{R})$ is isomorphic to the bundle $T^*Q \times \mathbb{R}$ and the canonical connection is the usual contact one-form θ . The curvature of this connection is the canonical symplectic form $d\theta$, up to a sign depending on conventions.

In §3 we shall study Lagrangians on J^1Y and the attendant Euler–Lagrange equations, etc., by a procedure parallel to that used for Lagrangians on the tangent bundle of a configuration manifold. Our choice of J^1Y as the field-theoretic tangent bundle reflects the fact that all of the theories we shall consider have Lagrangians which depend at most on the fields and their *first* derivatives. For higher order field theories, the appropriate objects to study are the higher order jet bundles; see, for example, Aldaya and Azcárraga [1980b], Shadwick [1982], Ferraris and Francaviglia [1983], Krupka [1987], Horák and Kolář [1983], de León and Rodrigues [1985], Gotay [1991a,b], and references therein.

We let $\dim X = n + 1$ and the fiber dimension of Y be N . Coordinates on X are denoted x^μ , $\mu = 0, 1, 2, \dots, n$, and fiber coordinates on Y are denoted by y^A , $A = 1, \dots, N$. These induce coordinates v^A_μ on the fibers of J^1Y . If $\phi : X \rightarrow Y$ is a section of π_{XY} , its tangent map $T_x\phi$ at $x \in X$ is an element of $J^1_{\phi(x)}Y$. Thus, the map $x \mapsto T_x\phi$ is a section of J^1Y regarded as a bundle over X . This section is denoted $j^1\phi$ and is called the **first jet prolongation** of ϕ . In coordinates, $j^1\phi$ is given by

$$x^\mu \mapsto (x^\mu, \phi^A(x^\mu), \partial_\nu \phi^A(x^\mu)), \quad (2A.4)$$

where $\partial_\nu = \partial/\partial x^\nu$. A section of the bundle $J^1Y \rightarrow X$ which is the first jet of a section of $Y \rightarrow X$ is said to be **holonomic**. In this work we shall use global sections, although at certain points in the arguments one should really work with local sections. It will be clear in each instance what is meant.

Examples

a Particle Mechanics. For non-relativistic classical mechanics with configuration space Q , let $X = \mathbb{R}$ (parameter time) and $Y = \mathbb{R} \times Q$, with π_{XY} the projection onto the first factor. The first jet bundle J^1Y is the bundle whose holonomic sections are tangents of sections $\phi : X \rightarrow Y$ (for example, curves in Q), so we can identify $J^1Y = \mathbb{R} \times TQ$. (Note however that J^1Y is an *affine* bundle, while $\mathbb{R} \times TQ$ has a *vector* bundle structure). Using coordinates (t, q^A) on $\mathbb{R} \times Q$, the induced coordinates on J^1Y are the usual tangent coordinates (t, q^A, v^A) .

In the case of a relativistic particle we would take Q to be the physical spacetime equipped with a Lorentzian metric g . (In this context, $A = 0, 1, 2, 3$ is now a spacetime index.)

b Electromagnetism. For electromagnetism on a fixed background (so that the spacetime metric is “frozen”), X is four-dimensional spacetime and Y is the bundle Λ^1X of one-form potentials A . Coordinates for J^1Y are $(x^\mu, A_\nu, v_{\nu\mu})$, where the derivative index μ is placed at the end:

$$v_{\nu\mu} \circ A = \partial_\mu A_\nu = A_{\nu,\mu}.$$

The one-forms A may be thought of as connections on a principal $U(1)$ -bundle. However, here we treat electromagnetism as an ordinary one-form field theory. Similarly, one could treat Yang–Mills fields with all the results of the paper holding, by taking Y to be the connection bundle over spacetime (i.e., the bundle whose sections are connections of a given principal bundle).

We also consider the case when the spacetime metric g is variable (but still not variational), so that the theory is parametrized. (See Interlude I. If we wanted the spacetime metric to be variational, we could couple electromagnetism to gravity.) To accomplish this, we merely append g to the other field variables as a *parameter*, replacing Y and J^1Y by $\tilde{Y} = Y \times_X S_2^{3,1}(X)$ and $J^1Y \times_X S_2^{3,1}(X)$, respectively, where $S_2^{3,1}(X)$ is the bundle of symmetric covariant symmetric two-tensors of Lorentz signature $(-, +, +, +)$. The coordinates on $J^1Y \times_X S_2^{3,1}(X)$ are now $(x^\mu, A_\nu, v_{\nu\mu}; g_{\sigma\rho})$, where we use a semicolon to separate variational fields from parametric ones. Note, in particular, that there are no “multivelocities” $v_{\sigma\rho\mu}$ corresponding to the $g_{\sigma\rho}$.

c A Topological Field Theory. Examples **b** and **d** following are both “metric” field theories, that is, theories which carry a metric (of some signature, usually Lorentzian) on the parameter space X . Here we give a simple example of a non-metric field theory: the (*abelian*) **Chern–Simons theory**. Topological field theories—of which the Chern–Simons theory is perhaps the best-known illustration—behave rather differently than metric theories, and exhibit a variety of interesting phenomena. We refer the reader to Horowitz [1989] for a fuller discussion of such theories.

Superficially, the Chern–Simons theory resembles electromagnetism. We take X to be a 3-manifold and $Y = \Lambda^1 X$. Coordinates on J^1Y are then $(x^\mu, A_\nu, v_{\nu\mu})$. As with electromagnetism, one could view A as a connection on some principal $U(1)$ -bundle over X . One could also consider non-abelian generalizations, where $U(1)$ is replaced by an arbitrary Lie group. But the key difference here is that we do not introduce a metric on X .

d Bosonic Strings. Let X be a 2-dimensional manifold with coordinates x^μ for $\mu = 1, 0$. Also, let (M, g) be a $(d+1)$ -dimensional spacetime manifold M with coordinates ϕ^A for $A = 0, 1, \dots, d$ and a fixed metric g of Lorentz signature $(-, +, \dots, +)$. A **bosonic string** is a map $\phi : X \rightarrow M$. Equivalently, its graph is a section of the product bundle $Y = X \times M$ over X . To write down the Lagrangian, we will also need a metric (or at least a conformal metric) on X . Thus, let $S_2^{1,1}(X)$ denote the bundle over X of symmetric covariant rank two tensors of Lorentz signature $(-, +)$. Then a section of $S_2^{1,1}(X)$ is a metric h on X .

There are (at least) three approaches to classical bosonic string theory which differ in the choice of independent variables and the Lagrangian. In the first (due to Polyakov [1981]), ϕ and h are taken as independent variables and the Lagrangian is taken as the harmonic map Lagrangian. In the second (due to

Nambu [1970]), ϕ is taken as the independent variable, while h is taken to be the induced metric

$$h_{\sigma\rho}(x) = (\phi^*g)_{\sigma\rho}(x) = g_{AB}(\phi(x)) \partial_\sigma \phi^A \partial_\rho \phi^B. \quad (2A.5)$$

Plugging these into the harmonic map Lagrangian, it reduces to the minimal surface (i.e., area) Lagrangian. Between these two extremes is the third version which takes ϕ and the conformal part of h as the independent variables. Since the harmonic map Lagrangian is conformally invariant, it may be reexpressed in terms of these variables. For simplicity, we shall only consider the Polyakov model in this paper since it treats the metric h variationally. This model has also been treated by Batlle, Gomis and Pons [1986]; however they assume that g is flat whereas our g is not necessarily flat. Our analysis of the bosonic string should also be compared with that of Beig [1989]. One can analyze the Nambu model in a similar way; see, for example, Scherk [1975].

With ϕ and h chosen as the independent variables, the covariant configuration bundle is $Y = (X \times M) \times_X S_2^{1,1}(X)$ and

$$J^1Y = J^1(X \times M) \times_X J^1(S_2^{1,1}(X)).$$

The coordinates on J^1Y are then $(x^\mu, \phi^A, h_{\sigma\rho}, v^A_\mu, w_{\sigma\rho\mu})$.

Remark 2. Concerted efforts to understand Einstein's theory of gravity have provided much of the motivation and insight for this work. It is therefore perhaps surprising that the (vacuum) gravitational field is a somewhat exceptional case. Here X is again spacetime and the natural choice for Y is $S_2^{3,1}(X)$. With this choice, however, one is forced to work on the second jet bundle J^2Y since the Hilbert Lagrangian depends upon the second derivatives of the metric. To avoid this, one may adopt the Palatini approach, using the metric *and* the torsion-free connection as the primary fields. In terms of these variables, the Hilbert Lagrangian is first order, and so the analysis may be done on the first jet bundle. The price we must pay for this simplification is that the Palatini formalism is plagued by a plethora of second class constraints. Śniatycki [1970b], Szczyrba [1976b], and Binz, Śniatycki, and Fischer [1988] have studied this approach and it is the one we shall follow. An alternative approach has been developed by Ashtekar [1986, 1987] using another set of canonical variables. It would be of interest to develop the formalism herein for this formulation of gravity; first steps in this direction have been taken by Esposito, Gionti, and Stornaiolo [1995].

Because of its length and intricacy, we defer the analysis of Palatini gravity until Part V. In the meantime, we observe that bosonic strings mimic gravity to a significant extent, while having the virtue of being a first order theory with only first class constraints. \blacklozenge

2B The Dual Jet Bundle

Next we introduce the field-theoretic analogue of the cotangent bundle. We define the **dual jet bundle** J^1Y^* to be the *vector* bundle over Y whose fiber at $y \in Y_x$ is the set of *affine* maps from J_y^1Y to $\Lambda_x^{n+1}X$, where $\Lambda^{n+1}X$ denotes the bundle of $(n+1)$ -forms on X .³ A smooth section of J^1Y^* is therefore an affine bundle map of J^1Y to $\Lambda^{n+1}X$ covering π_{XY} . We choose affine maps since J^1Y is an affine bundle, and we map into $\Lambda^{n+1}X$ since we are ultimately thinking of integration as providing the pairing on sections.

Fiber coordinates on J^1Y^* are (p, p_A^μ) , which correspond to the affine map given in coordinates by

$$v^A{}_\mu \mapsto (p + p_A^\mu v^A{}_\mu) d^{n+1}x \quad (2B.1)$$

where

$$d^{n+1}x = dx^0 \wedge dx^1 \wedge \cdots \wedge dx^n.$$

Analogous to the canonical one- and two-forms on a cotangent bundle, there are canonical forms on J^1Y^* . To define these, another description of J^1Y^* will be convenient. Namely, let $\Lambda := \Lambda^{n+1}Y$ denote the bundle of $(n+1)$ -forms on Y , with fiber over $y \in Y$ denoted by Λ_y and with projection $\pi_{Y\Lambda} : \Lambda \rightarrow Y$. Let $Z \subset \Lambda$ be the subbundle whose fiber is given by

$$Z_y = \{z \in \Lambda_y \mid \mathbf{i}_v \mathbf{i}_w z = 0 \text{ for all } v, w \in V_y Y\}, \quad (2B.2)$$

where \mathbf{i}_v denotes left interior multiplication by v .

Elements of Z can be written uniquely as

$$z = p d^{n+1}x + p_A^\mu dy^A \wedge d^n x_\mu \quad (2B.3)$$

where

$$d^n x_\mu = \partial_\mu \lrcorner d^{n+1}x.$$

Here \lrcorner also denotes the interior product. Hence fiber coordinates for Z are also (p, p_A^μ) .

Equating the coordinates (x^μ, y^A, p, p_A^μ) of Z and of J^1Y^* defines a vector bundle isomorphism

$$\Phi : Z \rightarrow J^1Y^*.$$

Intrinsically, Φ is defined by

$$\langle \Phi(z), \gamma \rangle = \gamma^* z \in \Lambda_x^{n+1}X \quad (2B.4)$$

³ The space of affine maps from an affine space into a linear space forms a vector space. Hence J^1Y^* is a vector bundle, despite the fact that J^1Y is only an affine bundle.

where $z \in Z_y$, $\gamma \in J_y^1 Y$, $x = \pi_{XY}(y)$, and $\langle \cdot, \cdot \rangle$ denotes the dual pairing. To see this, note that if γ has fiber coordinates v^A_μ , then

$$\gamma^* dx^\mu = dx^\mu \quad \text{and} \quad \gamma^* dy^A = v^A_\mu dx^\mu \quad (2B.5)$$

and so

$$\gamma^* (p d^{n+1}x + p_A^\mu dy^A \wedge d^n x_\mu) = (p + p_A^\mu v^A_\mu) d^{n+1}x, \quad (2B.6)$$

where we have used

$$dx^\nu \wedge d^n x_\mu = \delta_\mu^\nu d^{n+1}x.$$

The inverse of Φ can also be defined intrinsically, although it is somewhat more complicated. Notice that *a connection is not needed to define this isomorphism.*

We summarize:

Proposition 2.1. *The spaces $J^1 Y^*$ and Z are canonically isomorphic as vector bundles over Y .*

We shall construct canonical forms on Z and then use the isomorphism between $J^1 Y^*$ and Z to transfer these to $J^1 Y^*$. We first define the **canonical $(n+1)$ -form** Θ_Λ on Λ by

$$\begin{aligned} \Theta_\Lambda(z)(u_1, \dots, u_{n+1}) &= z(T\pi_{Y\Lambda} \cdot u_1, \dots, T\pi_{Y\Lambda} \cdot u_{n+1}) \\ &= (\pi_{Y\Lambda}^* z)(u_1, \dots, u_{n+1}), \end{aligned} \quad (2B.7)$$

where $z \in \Lambda$ and $u_1, \dots, u_{n+1} \in T_z \Lambda$. Define the **canonical $(n+2)$ -form** Ω_Λ on Λ by

$$\Omega_\Lambda = -\mathbf{d}\Theta_\Lambda. \quad (2B.8)$$

Note that if $n = 0$ (i.e., X is one-dimensional), then $\Lambda = T^*Y$ and Θ_Λ is the standard canonical one-form. If $i_{\Lambda Z} : Z \rightarrow \Lambda$ denotes the inclusion, the **canonical $(n+1)$ -form** Θ on Z is defined by

$$\Theta = i_{\Lambda Z}^* \Theta_\Lambda \quad (2B.9)$$

and the **canonical $(n+2)$ -form** Ω on Z is defined by

$$\Omega = -\mathbf{d}\Theta = i_{\Lambda Z}^* \Omega_\Lambda. \quad (2B.10)$$

The pair (Z, Ω) is called **multiphase space** or **covariant phase space**. It is an example of a **multisymplectic manifold**; see Remark 3 below.

From (2B.3) and (2B.7)–(2B.10), one finds that the coordinate expression for Θ is

$$\Theta = p_A^\mu dy^A \wedge d^n x_\mu + p d^{n+1}x, \quad (2B.11)$$

and so

$$\Omega = dy^A \wedge dp_A^\mu \wedge d^n x_\mu - dp \wedge d^{n+1}x. \quad (2B.12)$$

Remarks 1. There are other possible choices of multiphase space (cf. Echeverría-Enríques, Muñoz-Lecanda, and Román-Roy [2000] and the references cited at the beginning of this chapter), but this one is the most relevant and efficient for our purposes. There are good reasons for this choice: it is natural, as shown in Proposition 2.1, and it carries a canonical multisymplectic structure. We refer to Gotay [1991a,b] for more on why one makes this particular choice of multiphase space.

2. A notable feature of the covariant formalism is the appearance of several momenta p_A^μ , spatial in addition to temporal, for each field component y^A . The Legendre transformation of §3 will provide a physical interpretation of these “multimomenta” as well as of the “covariant Hamiltonian” p . As we shall see, the affine nature of our constructions provides a unification of the two parts of the Legendre transformation, namely the definitions of the momenta and the Hamiltonian, into a single “covariant” entity.

3. Whether or not there is a proper abstract definition of “multisymplectic manifold” is open to question. As a first try, we could say that a *multisymplectic manifold* is a manifold Z endowed with a closed k -form Ω which is nondegenerate in the sense that $\mathbf{i}_V \Omega \neq 0$ whenever V is a nonzero tangent vector. This would include our case, symplectic manifolds ($k = 2$), and manifolds with a volume form ($k = \dim Z$). But for all cases intermediate between symplectic and volume structures there is no Darboux theorem. Indeed, the set of germs of such k -multisymplectic forms for $2 < k < \dim Z$ has functional moduli: there are local differential invariants of such a form, distinguishing it from arbitrarily nearby multisymplectic forms. Such a theory would thus require extra structure, perhaps mimicking the fibration information contained in our Z , or else would proceed in a manner having very little in common with modern symplectic geometry where the Darboux theorem is such a central organizing fact. For attempts at such a “multisymplectic geometry” see Martin [1988] or Cantrijn, Ibort, and de León [1999]. \blacklozenge

We next recall that there is a characterization of the canonical one-form θ on T^*Q by $\beta^*\theta = \beta$ for any one-form β on Q (Abraham and Marsden [1978], Proposition 3.2.11). Here is an analogue of that fact for the canonical $(n + 1)$ -form.

Proposition 2.2. *If σ is a section of π_{XZ} and $\phi = \pi_{YZ} \circ \sigma$, then*

$$\sigma^*\Theta = \phi^*\sigma, \tag{2B.13}$$

where $\phi^*\sigma$ means the pull-back by ϕ to X of σ regarded as an $(n + 1)$ -form on Y along ϕ .

Proof. Evaluate the left-hand side of equation (2B.13) on tangent vectors v_1, \dots, v_{n+1} to X at x :

$$\begin{aligned}
(\sigma^*\Theta)(x)(v_1, \dots, v_{n+1}) &= \Theta(\sigma(x))(T\sigma \cdot v_1, \dots, T\sigma \cdot v_{n+1}) \\
&= \sigma(x)(T\pi_{YZ} \cdot T\sigma \cdot v_1, \dots, T\pi_{YZ} \cdot T\sigma \cdot v_{n+1}) \\
&= \sigma(x)(T\phi \cdot v_1, \dots, T\phi \cdot v_{n+1}) \\
&= (\phi^*\sigma)(x)(v_1, \dots, v_{n+1}). \quad \blacksquare
\end{aligned}$$

Examples

a Particle Mechanics. Let $X = \mathbb{R}$ and $Y = \mathbb{R} \times Q$. Then $Z = T^*Y = T^*\mathbb{R} \times T^*Q$ has coordinates $(t, p, q^1, \dots, q^N, p_1, \dots, p_N)$,

$$\Theta = p_A dq^A + p dt \quad (2B.14)$$

and

$$\Omega = dq^A \wedge dp_A + dt \wedge dp. \quad (2B.15)$$

In this case the multisymplectic approach reduces to the extended state space formulation of classical mechanics.

b Electromagnetism. For the Maxwell theory, coordinates on Z are

$$(x^\mu, A_\nu, p, \mathfrak{F}^{\nu\mu}).$$

Later we shall see that $\mathfrak{F}^{\nu\mu}$ can be identified, via the Legendre transform, with the electromagnetic field *density*, so we use this notation in anticipation. From (2B.11) and (2B.12), we get

$$\Theta = \mathfrak{F}^{\nu\mu} dA_\nu \wedge d^3x_\mu + p d^4x \quad (2B.16)$$

and

$$\Omega = dA_\nu \wedge d\mathfrak{F}^{\nu\mu} \wedge d^3x_\mu - dp \wedge d^4x. \quad (2B.17)$$

When the spacetime metric g is treated parametrically, Z is replaced by $\tilde{Z} = Z \times_X S_2^{3,1}(X)$ with coordinates $(x^\mu, A_\nu, p, \mathfrak{F}^{\nu\mu}; g_{\sigma\rho})$. There are no multimomenta conjugate to the $g_{\sigma\rho}$. The expressions for the canonical forms remain the same.

c A Topological Field Theory. In the case of the Chern–Simons theory, coordinates on Z are $(x^\mu, A_\nu, p, p^{\nu\mu})$. The canonical forms are

$$\Theta = p^{\nu\mu} dA_\nu \wedge d^2x_\mu + p d^3x \quad (2B.18)$$

and

$$\Omega = dA_\nu \wedge dp^{\nu\mu} \wedge d^2x_\mu - dp \wedge d^3x. \quad (2B.19)$$

d Bosonic Strings. Corresponding to $Y = (X \times M) \times_X S_2^{1,1}(X)$, the coordinates on Z are $(x^\mu, \phi^A, h_{\sigma\rho}, p, p_{A^\mu}, q^{\sigma\rho\mu})$, and the canonical forms are given by

$$\Theta = (p_{A^\mu} d\phi^A + q^{\sigma\rho\mu} dh_{\sigma\rho}) \wedge d^1x_\mu + p d^2x \quad (2B.20)$$

and

$$\Omega = (d\phi^A \wedge dp_{A^\mu} + dh_{\sigma\rho} \wedge dq^{\sigma\rho\mu}) \wedge d^1x_\mu - dp \wedge d^2x. \quad (2B.21)$$

We conclude this section with a few remarks on Poisson brackets. Poisson brackets are not essential for the main points in this paper, but we believe that they are important for linking up the results of this paper with those of Marsden, Montgomery, Morrison, and Thompson [1986] and for the future development of covariant reduction (see Marsden [1988]). For more thorough discussions of Poisson brackets in field theory, we refer the reader to the recent papers by Kanatchikov [1997], Lawson [2000], Castrillon and Marsden [2002], and Forger, Paufler, and Römer [2002a,b].

Recall that on a symplectic manifold (P, Ω) , the Hamiltonian vector field X_f of a function $f \in \mathcal{F}(P)$ is defined by

$$\Omega(X_f, V) = V \lrcorner \mathbf{d}f$$

for any $V \in \mathfrak{X}(P)$, and the Poisson bracket is defined by

$$\{f, h\} = \Omega(X_f, X_h).$$

The multisymplectic analogue of a function is an n -form. In Goldschmidt and Sternberg [1973] and Kijowski and Szczyrba [1975] a Poisson bracket is defined on certain types of n -forms. A special case relevant for later considerations is as follows. Let V be a vector field on Y which is π_{XY} -projectable; that is, there is a vector field U on X such that

$$T\pi_{XY} \circ V = U \circ \pi_{XY}.$$

Consider an n -form on Z of the type

$$f(z) = \pi_{YZ}^*(V \lrcorner z) \quad (2B.22)$$

for some π_{XY} -projectable vector field V ; we call such an f a *momentum observable*. (As we shall see in §4C, special covariant momentum maps can be regarded as momentum observables.) The *Hamiltonian vector field* X_f of such a momentum observable f is defined by

$$\mathbf{d}f = X_f \lrcorner \Omega$$

where Ω is the canonical $(n + 2)$ -form on Z . Since Ω is nondegenerate, this uniquely defines X_f . The *Poisson bracket* of two such n -forms is the n -form defined by

$$\{f, h\} = X_h \lrcorner (X_f \lrcorner \Omega). \quad (2B.23)$$

For momentum observables f and h , $\{f, h\}$ is, up to the addition of exact terms, another momentum observable; these constitute an important class since, as we shall see, it includes superhamiltonian and supermomentum functions, for instance. Also, the development in §5 of Part II shows that when (2B.23) is integrated over a hypersurface, it produces a canonical cotangent bundle bracket.

3 Lagrangian Dynamics

This section summarizes some of the basic facts concerning the Euler–Lagrange equations, the Legendre transformation and the Cartan form that will be needed later. There are many formulations available in the literature, including those of Carathéodory [1935], Weyl [1935], Lepage [1936, 1941, 1942], Cartan [1922], Dedecker [1957, 1977], Trautman [1967], Hermann [1968], Śniatycki [1970a, 1984], Krupka [1971], Ouzilou [1972], Goldschmidt and Sternberg [1973], García [1974], Szczyrba [1976a], Sternberg [1977], Kupershmidt [1980], Aldaya and de Azcárraga [1980a], Griffiths [1983], and Anderson [1992]. For second order theories, see Krupka and Štěpánková [1983] and, for the n^{th} order case, Horák and Kolář [1983], Muñoz [1985], and Gotay [1991a,c]. Exactly which route one takes to end up with the basic results is, to some extent, a matter of personal preference. For example, one may introduce the Cartan form axiomatically and then define the Legendre transformation in terms of it; see for instance Śniatycki [1970a, 1984] for this point of view. We have chosen to introduce the Legendre transformation directly and then use it to define the Cartan form to most closely parallel the standard treatment in classical mechanics (Abraham and Marsden [1978], Section 3.6). We shall continue to confine ourselves to first order theories.

As before, let $\pi_{XY} : Y \rightarrow X$ be a bundle over X whose sections will be the fields of the theory. In Section 2 we studied the spaces J^1Y and $Z \cong J^1Y^*$ which are the field-theoretic analogues of the tangent and cotangent bundles in particle mechanics. Here we are concerned with the relationship between the two and with Lagrangian dynamics on the former.

3A The Covariant Legendre Transformation

Let

$$\mathcal{L} : J^1Y \rightarrow \Lambda^{n+1}X,$$

the *Lagrangian density*, be a given smooth bundle map over X . In coordinates, we write

$$\mathcal{L} = L(x^\mu, y^A, v^A_\mu) d^{n+1}x. \quad (3A.1)$$

We remark that since X is assumed oriented, we may take $\mathcal{L}(\gamma)$ to be an $(n+1)$ -form on X . If X were not orientable, then we should more properly take $\mathcal{L}(\gamma)$ to be a density of weight one on X .

Our first goal is to construct the covariant Legendre transformation for \mathcal{L} . This is a fiber-preserving map

$$\mathbb{F}\mathcal{L} : J^1Y \rightarrow J^1Y^* \cong Z$$

over Y which has the coordinate expressions

$$p_{A^\mu} = \frac{\partial L}{\partial v^A_\mu}, \quad p = L - \frac{\partial L}{\partial v^A_\mu} v^A_\mu \quad (3A.2)$$

for the multimomenta p_{A^μ} and the covariant Hamiltonian p . An intrinsic definition follows.

If $\gamma \in J^1_y Y$ then $\mathbb{F}\mathcal{L}(\gamma)$ is the affine approximation to $\mathcal{L} \mid J^1_y Y$ at γ . Thus $\mathbb{F}\mathcal{L}(\gamma)$ is the affine map from $J^1_y Y$ to $\Lambda^{n+1}_x X$ (where $y \in Y_x$) given by

$$\langle \mathbb{F}\mathcal{L}(\gamma), \gamma' \rangle = \mathcal{L}(\gamma) + \frac{d}{d\varepsilon} \mathcal{L}(\gamma + \varepsilon(\gamma' - \gamma)) \Big|_{\varepsilon=0}, \quad (3A.3)$$

where $\gamma' \in J^1_y Y$. If one wishes, the second term in (3A.3) can be interpreted in terms of a standard fiber derivative. To derive the coordinate expressions (3A.2), suppose $\gamma = v^A_\mu$ and $\gamma' = w^A_\mu$. Then the right hand side of (3A.3) reads

$$\left(L(\gamma) + \frac{\partial L}{\partial v^A_\mu} (w^A_\mu - v^A_\mu) \right) d^{n+1}x,$$

which is an affine function of w^A_μ with linear and constant pieces given by the first and second equations of (3A.2) respectively. Hence (3A.2) is indeed the coordinate description of $\mathbb{F}\mathcal{L}$. (The coordinatization of J^1Y^* was defined by (2B.1).)

3B The Cartan Form

The *Cartan form* is the $(n+1)$ -form $\Theta_{\mathcal{L}}$ on J^1Y defined by

$$\Theta_{\mathcal{L}} = \mathbb{F}\mathcal{L}^* \Theta, \quad (3B.1)$$

where Θ is the canonical $(n+1)$ -form on Z . We also define the $(n+2)$ -form $\Omega_{\mathcal{L}}$ by

$$\Omega_{\mathcal{L}} = -\mathbf{d}\Theta_{\mathcal{L}} = \mathbb{F}\mathcal{L}^*\Omega, \quad (3B.2)$$

where $\Omega = -\mathbf{d}\Theta$ is the canonical $(n+2)$ -form on Z .

In coordinates, (2B.11), (2B.12) and (3A.2) yield

$$\Theta_{\mathcal{L}} = \frac{\partial L}{\partial v^A_{\mu}} dy^A \wedge d^n x_{\mu} + \left(L - \frac{\partial L}{\partial v^A_{\mu}} v^A_{\mu} \right) d^{n+1}x \quad (3B.3)$$

and

$$\Omega_{\mathcal{L}} = dy^A \wedge d \left(\frac{\partial L}{\partial v^A_{\mu}} \right) \wedge d^n x_{\mu} - d \left(L - \frac{\partial L}{\partial v^A_{\mu}} v^A_{\mu} \right) \wedge d^{n+1}x. \quad (3B.4)$$

One application of the Cartan form is to reexpress the Lagrangian for holonomic fields as

$$\mathcal{L}(j^1\phi) = (j^1\phi)^*\Theta_{\mathcal{L}}. \quad (3B.5)$$

To see this, we use (3B.3) to compute

$$(j^1\phi)^*\Theta_{\mathcal{L}} = \frac{\partial L}{\partial v^A_{\mu}}(j^1\phi) d\phi^A \wedge d^n x_{\mu} + \left(L(j^1\phi) - \frac{\partial L}{\partial v^A_{\mu}}(j^1\phi) \phi^A_{,\mu} \right) d^{n+1}x.$$

The first and last terms cancel and so we obtain (3B.5).

It is possible to obtain $\Theta_{\mathcal{L}}$ directly via the action functional; it is constructed from the boundary terms that appear when the action integral is varied. The details can be found in, e.g., Pâquet [1941], Krupka [1987], and Marsden, Patrick, and Shkoller [1998].

3C The Euler–Lagrange Equations

We now use the formalism developed so far to reexpress, in an intrinsic way, the *Euler–Lagrange equations*, which in coordinates take the standard form

$$\frac{\partial L}{\partial y^A}(j^1\phi) - \frac{\partial}{\partial x^{\mu}} \left(\frac{\partial L}{\partial v^A_{\mu}}(j^1\phi) \right) = 0 \quad (3C.1)$$

for a (local) section ϕ of Y . The left hand side of (3C.1) is often denoted $\delta L/\delta\phi^A$ and is called the *Euler–Lagrange derivative* of L .

Theorem 3.1. *The following assertions regarding a section ϕ of the bundle $\pi_{XY} : Y \rightarrow X$ are equivalent:*

- (i) ϕ is a stationary point of $\int_X \mathcal{L}(j^1\phi)$;

- (ii) the Euler–Lagrange equations (3C.1) hold in coordinates;
 (iii) for any vector field W on J^1Y ,

$$(j^1\phi)^*(W \lrcorner \Omega_{\mathcal{L}}) = 0. \quad (3C.2)$$

Remarks 1. The theorem still holds if the section ϕ is only a local section: $\phi : U \rightarrow Y_U$ where U is an open subset of X . The only change is that the integral in (i) is now over U . The case that will be of importance to us is where U is a neighborhood of a Cauchy surface.

2. The meaning of “stationary point” is in the formal sense of the calculus of variations. Namely, a *variation* of ϕ is provided by a curve $\phi_\lambda = \eta_\lambda \circ \phi$, where η_λ is the flow of a vertical vector field V on Y which is compactly supported in X . One says that ϕ is a *stationary point* if

$$\left. \frac{d}{d\lambda} \left[\int_X \mathcal{L}(j^1\phi_\lambda) \right] \right|_{\lambda=0} = 0$$

for all variations ϕ_λ of ϕ .

3. This theorem appears more or less as we present it in, for example, Dedecker [1953], Trautman [1967], Śniatycki [1970a, 1984], Ouzilou [1972], Goldschmidt and Sternberg [1973], and Guillemin and Sternberg [1977].

4. If in (3C.2) one requires that $s^*(\mathbf{i}_W \Omega_{\mathcal{L}}) = 0$ for any section s of $J^1Y \rightarrow X$, not necessarily holonomic—that is, not necessarily of the form $j^1\phi$ —then condition (iii) will be equivalent to the Euler–Lagrange equations *and* the assertion that s is holonomic under the assumption that the Lagrangian is “regular.” A Lagrangian is called *regular* if the Legendre transformation is of maximal rank or, equivalently, if $(J^1Y, \Omega_{\mathcal{L}})$ is a multisymplectic manifold. In the formulation as we have given it, *no assumption of regularity is needed*; in fact such an assumption would be very much inappropriate, since it is not valid in any of our examples. See, however, García and Muñoz [1985, 1991] and Saunders [1992].

5. Because the Euler–Lagrange equations are equivalent to the intrinsic conditions (i) and (iii), they too must be intrinsic. One can in fact write the Euler–Lagrange derivative intrinsically in a direct way. See Trautman [1967] and Marsden, Patrick, and Shkoller [1998].

6. In the above we have restricted attention to vertical variations. But it is occasionally necessary to take “horizontal variations,” as in continuum mechanics where they produce configurational forces which are important in dislocation theory. See Marsden, Patrick, and Shkoller [1998] for details, and Lew, Marsden, Ortiz, and West [2002] for further applications of the split between horizontal and vertical variations. \blacklozenge

In the course of the proof, we will examine three cases for the vector field W in (3C.2). The following lemma shows that two of the cases are trivial.

Lemma 3.2. *If ϕ is a section of π_{XY} and if either W is tangent to the image of $j^1\phi$ or W is π_{Y, J^1Y} -vertical, then*

$$(j^1\phi)^*(W \lrcorner \Omega_{\mathcal{L}}) = 0.$$

Proof. First assume that W is tangent to the image of $j^1\phi$ in J^1Y ; that is, $W = T(j^1\phi) \cdot w$ for some vector field w on X . Then

$$\begin{aligned} (j^1\phi)^*(W \lrcorner \Omega_{\mathcal{L}}) &= (j^1\phi)^*((T(j^1\phi) \cdot w) \lrcorner \Omega_{\mathcal{L}}) \\ &= w \lrcorner (j^1\phi)^*\Omega_{\mathcal{L}}, \end{aligned}$$

which vanishes since $(j^1\phi)^*\Omega_{\mathcal{L}}$ is an $(n+2)$ -form on the $(n+1)$ -manifold X .

Secondly, if W is π_{Y, J^1Y} -vertical, then W has the coordinate form

$$W = (0, 0, W^A{}_{\mu}).$$

A calculation using (3B.4) shows that

$$W \lrcorner \Omega_{\mathcal{L}} = -W^B{}_{\nu} \frac{\partial^2 L}{\partial v^A{}_{\mu} \partial v^B{}_{\nu}} (dy^A \wedge d^n x_{\mu} - v^A{}_{\mu} d^{n+1}x),$$

which vanishes when pulled back by the jet prolongation of a section of Y . ■

Proof of Theorem 3.1. The direct proof of equivalence of (i) and (ii) is a standard argument in the calculus of variations that we shall omit. We shall work towards a proof of the equivalence of (i) and (iii) and then of (iii) and (ii).

To show that (i) is equivalent to (iii), consider a variation $\phi_{\lambda} = \eta_{\lambda} \circ \phi$ of ϕ (as in Remark 2) corresponding to a π_{XY} -vertical vector field V on Y with compact support in X . Using (3B.5), we compute

$$\begin{aligned} \frac{d}{d\lambda} \left[\int_X \mathcal{L}(j^1\phi_{\lambda}) \right] \Big|_{\lambda=0} &= \frac{d}{d\lambda} \left[\int_X (j^1\phi_{\lambda})^* \Theta_{\mathcal{L}} \right] \Big|_{\lambda=0} \\ &= \frac{d}{d\lambda} \left[\int_X (j^1\phi)^*(j^1\eta_{\lambda})^* \Theta_{\mathcal{L}} \right] \Big|_{\lambda=0} \\ &= \int_X (j^1\phi)^* \mathcal{L}_{j^1V} \Theta_{\mathcal{L}}, \end{aligned} \tag{3C.3}$$

where

$$j^1V = \frac{d}{d\lambda} j^1\eta_{\lambda} \Big|_{\lambda=0}$$

is the jet prolongation of V to J^1Y , given in coordinates by

$$j^1V = \left(0, V^A, \frac{\partial V^A}{\partial x^{\mu}} + \frac{\partial V^A}{\partial y^B} v^B{}_{\mu} \right)$$

and where \mathcal{L}_{j^1V} is the Lie derivative along j^1V . (See §4A for an intrinsic definition of jet prolongations and for more information; in particular, see Remark 1 of that section). Thus, from (3C.3) and

$$\mathcal{L}_{j^1V}\Theta_{\mathcal{L}} = -j^1V \lrcorner \Omega_{\mathcal{L}} + \mathbf{d}(j^1V \lrcorner \Theta_{\mathcal{L}}),$$

we get

$$\begin{aligned} \frac{d}{d\lambda} \left[\int_X \mathcal{L}(j^1\phi_\lambda) \right] \Big|_{\lambda=0} &= - \int_X (j^1\phi)^*(j^1V \lrcorner \Omega_{\mathcal{L}}) \\ &\quad + \int_X \mathbf{d}(j^1\phi)^*(j^1V \lrcorner \Theta_{\mathcal{L}}) \\ &= - \int_X (j^1\phi)^*(j^1V \lrcorner \Omega_{\mathcal{L}}) \end{aligned} \quad (3C.4)$$

by Stokes' theorem and the fact that V , and hence j^1V is compactly supported in X . From (3C.4) it follows that (iii) implies (i).

To show the converse, we first observe that any tangent vector field W on J^1Y can be decomposed into a component tangent to the image of $j^1\phi$ and a π_{X,J^1Y} -vertical vector field. Similarly, any π_{X,J^1Y} -vertical vector field can be decomposed into a jet extension of some π_{XY} -vertical vector field V on Y and a π_{Y,J^1Y} -vertical vector field.

Thus, if (i) holds, (3C.4) and Lemma 3.2 show that

$$\int_X (j^1\phi)^*(W \lrcorner \Omega_{\mathcal{L}}) = 0$$

for all vector fields W on J^1Y with compact support in X . Since W can be multiplied by an arbitrary scalar function on X , an argument like that in the Fundamental Lemma of the Calculus of Variations shows that the integrand must vanish for all vector fields W with compact support in X . A partition of unity argument then shows that the integrand vanishes for any vector field W whatsoever, which proves (iii).

To see directly that (iii) is equivalent to the Euler–Lagrange equations, one readily computes in coordinates that along $j^1\phi$,

$$(j^1\phi)^*(j^1V \lrcorner \Omega_{\mathcal{L}}) = V^A \left[\frac{\partial}{\partial x^\mu} \left(\frac{\partial L}{\partial v^A_\mu}(j^1\phi) \right) - \frac{\partial L}{\partial y^A}(j^1\phi) \right] d^{n+1}x \quad (3C.5)$$

for all π_{XY} -vertical vector fields V on Y . Thus (iii) implies (ii). On the other hand, (3C.5) combined with the above remarks on decompositions of vector fields and Lemma 3.2 shows that (ii) implies (iii). ■

Examples

a Particle Mechanics. With the Lagrangian density $\mathcal{L} = L(t, q, v)dt$, expressions (3A.2) become $p_A = \partial L / \partial v^A$ and $p = L - v^A \partial L / \partial v^A = -E$, the Cartan form becomes the usual one-form for describing time-dependent mechanics

$$\Theta_{\mathcal{L}} = \frac{\partial L}{\partial v^A} dq^A - E dt, \quad (3C.6)$$

and the Euler–Lagrange equations take the standard form

$$\frac{d}{dt} \frac{\partial L}{\partial v^A} = \frac{\partial L}{\partial q^A}. \quad (3C.7)$$

Let us specifically work out the details for a relativistic free particle of mass $m \neq 0$. The Lagrangian density is

$$\mathcal{L} = -m \|\mathbf{v}\| dt, \quad (3C.8)$$

where $\mathbf{v} = (v^0, v^1, v^2, v^3)$ is the velocity of the particle (with respect to parameter time t) and $\|\mathbf{v}\| = (-g_{AB}v^A v^B)^{1/2}$, g being the spacetime metric.⁴ The Cartan form (3C.6) is thus

$$\Theta_{\mathcal{L}} = \frac{mg_{AB}v^B}{\|\mathbf{v}\|} dq^A, \quad (3C.9)$$

whence

$$p_A = \frac{mg_{AB}v^B}{\|\mathbf{v}\|} \quad \text{and} \quad p = 0. \quad (3C.10)$$

The equations of motion (3C.7) are just the geodesic equations for the spacetime metric g :

$$\frac{d}{dt} \left(\frac{g_{AB}v^B}{\|\mathbf{v}\|} \right) = \frac{1}{2\|\mathbf{v}\|} g_{BC,A} v^B v^C \quad (3C.11)$$

where t is used as the curve parameter.

⁴ Sometimes in this context one encounters the Lagrangian density

$$\mathcal{L}' = -\frac{1}{2} m \|\mathbf{v}\|^2 dt.$$

However, unlike \mathcal{L} given by (3C.8) \mathcal{L}' is not time reparametrization-invariant. In §11C we shall see how to ‘repair’ this defect in \mathcal{L}' . Of course, if one does not require time reparametrization-invariance, then the choice \mathcal{L}' of Lagrangian density is certainly permissible. Misner, Thorne, and Wheeler [1973] give a detailed discussion of these two Lagrangian densities in §13.4.

b Electromagnetism. For electromagnetism on a fixed background space-time X with metric g , the Lagrangian density is⁵

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu}\sqrt{-g}d^4x, \quad (3C.12)$$

where $\sqrt{-g} := \sqrt{-\det g_{\mu\nu}}$ and, viewed as a function on $J^1(\Lambda^1X)$, the field tensor $F_{\mu\nu}$ is defined by

$$F_{\mu\nu} = v_{\nu\mu} - v_{\mu\nu}. \quad (3C.13)$$

The Legendre transformation gives the relations

$$\mathfrak{F}^{\mu\nu} = F^{\mu\nu}\sqrt{-g} \quad \text{and} \quad p = \frac{1}{4}F_{\mu\nu}F^{\mu\nu}\sqrt{-g}. \quad (3C.14)$$

The Cartan form $\Theta_{\mathcal{L}}$ is

$$\Theta_{\mathcal{L}} = \sqrt{-g}F^{\nu\mu}dA_{\nu} \wedge d^3x_{\mu} + \frac{1}{4}F_{\mu\nu}F^{\mu\nu}\sqrt{-g}d^4x \quad (3C.15)$$

and the 5-form $\Omega_{\mathcal{L}} = -\mathbf{d}\Theta_{\mathcal{L}}$ is

$$\begin{aligned} \Omega_{\mathcal{L}} = \sqrt{-g}(g^{\nu\sigma}g^{\mu\rho} - g^{\nu\rho}g^{\mu\sigma})dA_{\nu} \wedge dv_{\rho\sigma} \wedge d^3x_{\mu} \\ - \sqrt{-g}F^{\mu\nu}dv_{\nu\mu} \wedge d^4x. \end{aligned} \quad (3C.16)$$

For a section A of $Y = \Lambda^1X$, (3C.13) gives $(j^1A)^*F = \mathbf{d}A$ and the Euler–Lagrange equations reduce to

$$(A^{\nu,\mu} - A^{\mu,\nu})_{;\nu} = 0. \quad (3C.17)$$

Here the semicolon denotes a covariant derivative with respect to the background metric, so (3C.17) is a covariant divergence.

When g is treated parametrically, none of the formulas above change. In particular, since g is not variational, it has no Euler–Lagrange equations. Even so, varying the Lagrangian density with respect to g produces the stress-energy-momentum tensor of the electromagnetic field, as shown in Example **b** of §4D and Interlude II.

c A Topological Field Theory. The Chern–Simons Lagrangian density is

$$\mathcal{L} = \frac{1}{2}\epsilon^{\mu\nu\sigma}F_{\mu\nu}A_{\sigma}d^3x, \quad (3C.18)$$

⁵ For convenience, in this example we use (3C.12), with the coefficient $-1/4$, instead of the Misner, Thorne, and Wheeler [1973] Lagrangian, with coefficient $-1/16\pi$. One may pass from one convention to the other by scaling the electromagnetic fields.

where the curvature F of A is defined by (3C.13). Intrinsically, $\mathcal{L} = F \wedge A$. Applying the Legendre transformation, we get

$$p^{\nu\mu} = \epsilon^{\mu\nu\sigma} A_\sigma \quad \text{and} \quad p = 0. \quad (3C.19)$$

The Cartan form is thus

$$\Theta_{\mathcal{L}} = \epsilon^{\mu\nu\sigma} A_\sigma dA_\nu \wedge d^2x_\mu, \quad (3C.20)$$

whence

$$\Omega_{\mathcal{L}} = -\epsilon^{\mu\nu\sigma} dA_\sigma \wedge dA_\nu \wedge d^2x_\mu. \quad (3C.21)$$

The Euler–Lagrange equations are $\epsilon^{\mu\nu\sigma} F_{\mu\nu} = 0$, which are equivalent to the connection A being flat:

$$(j^1A)^* F = \mathbf{d}A = 0. \quad (3C.22)$$

That these equations are only of the first order indicates that the Chern–Simons theory is highly degenerate.

d Bosonic Strings. In the Polyakov model, with ϕ and h as independent variables, the action is taken as the negative of the “energy” functional. Thus as a function on J^1Y , the Lagrangian density is

$$\mathcal{L} = -\frac{1}{2} \sqrt{|h|} h^{\sigma\rho} g_{AB} v^A_\sigma v^B_\rho d^2x. \quad (3C.23)$$

The Legendre transformation is

$$p_A^\mu = -\sqrt{|h|} h^{\mu\nu} g_{AB} v^B_\nu \quad (3C.24)$$

$$q^{\sigma\rho\mu} = 0 \quad (3C.25)$$

$$p = \frac{1}{2} \sqrt{|h|} h^{\mu\nu} g_{AB} v^A_\mu v^B_\nu, \quad (3C.26)$$

so that the Cartan form is

$$\Theta_{\mathcal{L}} = \sqrt{|h|} \left(-h^{\mu\nu} g_{AB} v^B_\nu d\phi^A \wedge d^1x_\mu + \frac{1}{2} h^{\mu\nu} g_{AB} v^A_\mu v^B_\nu d^2x \right) \quad (3C.27)$$

and

$$\begin{aligned} \Omega_{\mathcal{L}} = d\phi^A \wedge d \left(-\sqrt{|h|} h^{\mu\nu} g_{AB} v^B_\nu \right) \wedge d^1x_\mu \\ - d \left(\frac{1}{2} \sqrt{|h|} h^{\mu\nu} g_{AB} v^A_\mu v^B_\nu \right) \wedge d^2x. \end{aligned} \quad (3C.28)$$

The Euler–Lagrange equations reduce to

$$(h^{\mu\nu} g_{AB}(\phi) \phi^B_{,\nu})_{;\mu} = 0 \quad (3C.29)$$

$$\left(\frac{1}{2} h^{\mu\nu} g_{AB}(\phi) \phi^A_{,\mu} \phi^B_{,\nu} \right) h_{\alpha\beta} = g_{CD}(\phi) \phi^C_{,\alpha} \phi^D_{,\beta}. \quad (3C.30)$$

Equation (3C.29) says that ϕ satisfies the harmonic map equation. Equation (3C.30) does two things:

- (i) it says h is conformally related to the induced metric ϕ^*g on X : $\Lambda^2 h_{\alpha\beta} = (\phi^*g)_{\alpha\beta}$, and
- (ii) it determines the conformal factor: $\Lambda^2 = \frac{1}{2} h^{\mu\nu} g_{AB}(\phi) \phi^A_{,\mu} \phi^B_{,\nu}$.

At first glance, it might appear that (ii) is significant while (i) is insignificant, since in two dimensions “any two metrics are conformally related.” However, the content of (3C.30) is, in fact, just the reverse. The actual theorem states that in two dimensions, given two metrics, there is a diffeomorphism which transforms one metric into a multiple of the other. Equation (3C.30) says that in *any* coordinate system h is proportional to ϕ^*g . There is no need for a diffeomorphism.

We summarize the content of (3C.30): There are three equations. Two say that the components of h and ϕ^*g are proportional while the third determines the proportionality constant. The first two have content while the third is an identity.

4 Covariant Momentum Maps and Noether's Theorem

A key feature of relativistic field theories is that not all of the Euler–Lagrange equations necessarily describe the temporal evolution of fields; some of the equations may impose constraints on the choice of initial data. Those constraints which are first class in the sense of Dirac [1950, 1964] reflect the gauge symmetry of the theory and play a central role in our development. In Part III we will see that these first class constraints are related to the vanishing of various momentum maps.

To lay the groundwork for this, we now study symmetries in the Lagrangian and multisymplectic formalisms. First, we show how automorphisms of Y lift to J^1Y in a way compatible, via the Legendre transform, with their canonical lifts to Z . Second, we introduce the notion of a covariant momentum map for a group \mathcal{G} acting on Z by covariant canonical transformations. Third, we prove that the lifted action of a group \mathcal{G} on J^1Y has a covariant momentum map (with respect to the Cartan form) which is given by the pull-back to J^1Y of the covariant momentum map on Z by $\mathbb{F}\mathcal{L}$. Then we prove Noether's theorem, here formulated in terms of covariant momentum maps.

4A Jet Prolongations

Here we show how to naturally lift automorphisms of Y to automorphisms of J^1Y . In effect, we shall construct the covariant analogue of the tangent map.

Let $\eta_Y : Y \rightarrow Y$ be a π_{XY} -bundle automorphism covering a diffeomorphism $\eta_X : X \rightarrow X$. If $\gamma : T_x X \rightarrow T_y Y$ is an element of J^1Y , let $\eta_{J^1Y}(\gamma) : T_{\eta_X(x)}X \rightarrow T_{\eta_Y(y)}Y$ be defined by

$$\eta_{J^1Y}(\gamma) = T\eta_Y \circ \gamma \circ T\eta_X^{-1}. \quad (4A.1)$$

The π_{Y, J^1Y} -bundle automorphism $j^1\eta_Y := \eta_{J^1Y}$ so constructed is called the **first jet extension** or **prolongation** of η_Y to J^1Y . In coordinates, if $\gamma = (x^\mu, y^A, v^A{}_\mu)$, then

$$\eta_{J^1Y}(\gamma) = \left(\eta_X^\mu(x), \eta_Y^A(x, y), [\partial_\nu \eta_Y^A + (\partial_B \eta_Y^A) v^B{}_\nu] \partial_\mu (\eta_X^{-1})^\nu \right), \quad (4A.2)$$

where $\partial_B = \partial/\partial y^B$. Occasionally we are interested in the case when η_Y is vertical; that is, η_X is the identity. Equations (4A.1) and (4A.2) simplify accordingly in this case.

If V is a vector field on Y whose flow is η_λ , so that

$$V \circ \eta_\lambda = \frac{d\eta_\lambda}{d\lambda},$$

then its **first jet extension** or **prolongation** $j^1V := V_{J^1Y}$ is the vector field on J^1Y whose flow is $j^1\eta_\lambda$:

$$j^1V \circ j^1\eta_\lambda = \frac{d}{d\lambda} j^1\eta_\lambda. \quad (4A.3)$$

In coordinates, by differentiating (4A.2), we obtain

$$j^1V = \left(V^\mu, V^A, \frac{\partial V^A}{\partial x^\mu} + \frac{\partial V^A}{\partial y^B} v^B{}_\mu - v^A{}_\nu \frac{\partial V^\nu}{\partial x^\mu} \right). \quad (4A.4)$$

It is important to note that if ϕ is a section of π_{XY} and η_Y is a π_{XY} -bundle automorphism covering η_X , then $\eta_Y \cdot \phi := \eta_Y \circ \phi \circ \eta_X^{-1}$ is another section and

$$j^1(\eta_Y \cdot \phi) = j^1\eta_Y \circ j^1\phi \circ \eta_X^{-1}. \quad (4A.5)$$

See García [1974] and Aldaya and de Azcárraga [1980a] for more information.

Remarks 1. If $\phi_\lambda = \eta_\lambda \circ \phi$ is a variation of ϕ (as in Remark 2 of §3C), then η_X is the identity and (4A.5) becomes

$$j^1\phi_\lambda = j^1\eta_\lambda \circ j^1\phi.$$

In this case, the generating vector field of the flow η_λ is vertical and (4A.4) reduces to

$$j^1V = \left(0, V^A, \frac{\partial V^A}{\partial x^\mu} + \frac{\partial V^A}{\partial y^B} v^B{}_\mu \right).$$

2. In the above mentioned references, an additional term $v^B_\mu v^A_\lambda \partial V^\lambda / \partial y^B$ is present in the third component of (4A.4); this is absent here since we consider only transformations of Y which are bundle automorphisms so V^λ is independent of y^B .

3. *A word of caution:* One can also view V as a section of the bundle $TY \rightarrow Y$ and take its first jet in the sense of equation (2A.4). Then one gets the section of $J^1(TY) \rightarrow Y$ given in coordinates by

$$(x^\mu, y^A) \mapsto \left(x^\mu, y^A, V^\mu, V^A, \frac{\partial V^\mu}{\partial x^\nu}, \frac{\partial V^\mu}{\partial y^B}, \frac{\partial V^A}{\partial x^\nu}, \frac{\partial V^A}{\partial y^B} \right).$$

This section is not to be confused with j^1V as defined by (4A.3) and (4A.4); they are two different objects.

4. From the definitions one finds that $j^1([V, W]) = [j^1V, j^1W]$. ◆

4B Covariant Canonical Transformations

A *covariant canonical transformation* is a π_{XZ} -bundle map $\eta_Z : Z \rightarrow Z$ covering a diffeomorphism $\eta_X : X \rightarrow X$ such that

$$\eta_Z^* \Omega = \Omega. \tag{4B.1}$$

We say η_Z is a *special covariant canonical transformation* if

$$\eta_Z^* \Theta = \Theta. \tag{4B.2}$$

If $\eta_Y : Y \rightarrow Y$ is a π_{XY} -bundle automorphism (also covering a diffeomorphism $\eta_X : X \rightarrow X$), its *canonical lift* $\eta_Z : Z \rightarrow Z$ is defined by

$$\eta_Z(z) = (\eta_Y^{-1})^* z; \tag{4B.3}$$

that is,

$$\eta_Z(z)(v_1, \dots, v_{n+1}) = z(T\eta_Y^{-1} \cdot v_1, \dots, T\eta_Y^{-1} \cdot v_{n+1}),$$

where $z \in Z_y$ and $v_1, \dots, v_{n+1} \in T_{\eta_Y(y)}Y$. Since η_Y is a π_{XY} -bundle map, $T\eta_Y$ maps vertical vectors to vertical vectors, so η_Z does indeed map Z to Z . In coordinates, if we write $z = p d^{n+1}x + p_A^\mu dy^A \wedge d^n x_\mu$, and notice that

$$(\eta_Y^{-1})^* dx^\mu = (\eta_X^{-1})^* dx^\mu$$

and

$$(\eta_Y^{-1})^* dy^A = \partial_\nu (\eta_Y^{-1})^A dx^\nu + \partial_B (\eta_Y^{-1})^A dy^B,$$

we get

$$\begin{aligned} \eta_Z(z) &= (\eta_Y^{-1})^* z = \left(p + \partial_\nu (\eta_Y^{-1})^A p_A^\mu \partial_\mu \eta_X^\nu \right) J^{-1} d^{n+1}x \\ &\quad + \left(\partial_A (\eta_Y^{-1})^B p_B^\nu \partial_\nu \eta_X^\mu \right) J^{-1} dy^A \wedge d^n x_\mu, \end{aligned} \quad (4B.4)$$

where J is the Jacobian determinant of η_X and where the quantities in (4B.4) are evaluated at the appropriate points.

Proposition 4.1. *Lifts are special covariant canonical transformations.*

This is readily proved using a method similar to the cotangent bundle case (Abraham and Marsden [1978], Theorem 3.2.12).

Here is a second way to understand the lift of a π_{XY} -bundle automorphism $\eta_Y : Y \rightarrow Y$ covering a diffeomorphism $\eta_X : X \rightarrow X$, which proceeds via $J^1 Y^*$ rather than directly to Z . We define the **lift** of η_Y to $J^1 Y^*$ to be the affine dual of the prolongation $\eta_{J^1 Y}$; that is, if $z : J_y^1 Y \rightarrow \Lambda_x^{n+1} X$ is an element of $J_y^1 Y^*$, let

$$\eta_{J^1 Y^*}(z) : J_{\eta_Y(y)}^1 Y \rightarrow \Lambda_{\eta_X(x)}^{n+1} X$$

be the affine map defined by

$$\langle \eta_{J^1 Y^*}(z), \gamma \rangle = (\eta_X^{-1})^* \langle z, \eta_{J^1 Y}^{-1}(\gamma) \rangle. \quad (4B.5)$$

In coordinates, if $z = (p, p_A^\mu)$ acts on $\gamma = (x^\mu, y^A, v^A_\mu)$ as in (2B.1), then from (4A.2),

$$\langle \eta_{J^1 Y^*}(z), \gamma \rangle \quad (4B.6)$$

$$\begin{aligned} &= (\eta_X^{-1})^* \left[\left(p + p_A^\mu \partial_\mu \eta_X^\nu \left[\partial_\nu (\eta_Y^{-1})^A + v^B_\nu \partial_B (\eta_Y^{-1})^A \right] \right) d^{n+1}x \right] \\ &= \left[\left(p + \partial_\nu (\eta_Y^{-1})^A p_A^\mu \partial_\mu \eta_X^\nu \right) + \left(\partial_A (\eta_Y^{-1})^B p_B^\nu \partial_\nu \eta_X^\mu \right) v^A_\mu \right] J^{-1} d^{n+1}x \end{aligned}$$

so that the terms in (4B.6) correspond to the components in (4B.4) via (2B.6). Consequently, we have given a coordinate proof of

Proposition 4.2. *The canonical isomorphism $\Phi : Z \rightarrow J^1 Y^*$ is equivariant with respect to the lifts η_Z and $\eta_{J^1 Y^*}$; that is,*

$$\Phi \circ \eta_Z = \eta_{J^1 Y^*} \circ \Phi \quad (4B.7)$$

Here is an invariant proof of Proposition 4.2 using (2B.4), (4A.1), (4B.3) and (4B.5) above. For $z \in Z$ and $\gamma \in J^1 Y$, we have

$$\begin{aligned} \langle \eta_{J^1 Y^*}(\Phi(z)), \gamma \rangle &= (\eta_X^{-1})^* \langle \Phi(z), \eta_{J^1 Y}^{-1}(\gamma) \rangle \\ &= (\eta_X^{-1})^* \langle \Phi(z), T\eta_Y^{-1} \circ \gamma \circ T\eta_X \rangle \\ &= (\eta_X^{-1})^* \left[(T\eta_Y^{-1} \circ \gamma \circ T\eta_X)^* z \right] \\ &= (\eta_X^{-1})^* \eta_X^* \gamma^* (\eta_Y^{-1})^* z = \gamma^* (\eta_Z(z)) = \langle \Phi(\eta_Z(z)), \gamma \rangle \end{aligned}$$

and so (4B.7) holds. Thus, the identification of Z and J^1Y^* respects the action of a lifted covariant canonical transformation.

Let V be a vector field on Y whose flow η_λ maps fibers of π_{XY} to fibers. The **canonical lift** of V to Z is the vector field V_Z that generates the canonical lift of this flow to Z . By differentiating (4B.4) we obtain

$$V_Z = (V^\mu, V^A, -pV^\nu{}_{,\nu} - p_B{}^\nu V^B{}_{,\nu}, p_A{}^\nu V^\mu{}_{,\nu} - p_B{}^\mu V^B{}_{,\nu} - p_A{}^\mu V^\nu{}_{,\nu}) \quad (4B.8)$$

in coordinates, where $V = (V^\mu, V^A)$. By either direct calculation or by Proposition 4.1, we get $\mathcal{L}_{V_Z}\Theta = 0$.

4C Covariant Momentum Maps

We now introduce the covariant analogue of momentum maps in symplectic and Poisson geometry (see for example, Abraham and Marsden [1978], Marsden, Weinstein, Ratiu, Schmid, and Spencer [1983], Guillemin and Sternberg [1984], Marsden [1992], and Marsden and Ratiu [1999]).

Let \mathcal{G} denote a Lie group (perhaps infinite-dimensional) with Lie algebra \mathfrak{g} that acts on X by diffeomorphisms and acts on Z (or Y) as π_{XZ} - (or π_{XY} -) bundle automorphisms. For $\eta \in \mathcal{G}$, let η_X, η_Y , and η_Z denote the corresponding transformations of X, Y , and Z , and for $\xi \in \mathfrak{g}$, let ξ_X, ξ_Y , and ξ_Z denote the corresponding infinitesimal generators. If \mathcal{G} acts on Z by covariant canonical transformations, then the Lie derivative of Ω along ξ_Z is zero:

$$\mathcal{L}_{\xi_Z}\Omega = 0 \quad (4C.1)$$

and if \mathcal{G} acts by *special* covariant canonical transformations, then

$$\mathcal{L}_{\xi_Z}\Theta = 0. \quad (4C.2)$$

If \mathcal{G} acts on Z by covariant canonical transformations, then a **covariant momentum map** (or a **multimomentum map**) for this action is a map

$$J : Z \rightarrow \mathfrak{g}^* \otimes \Lambda^n Z = L(\mathfrak{g}, \Lambda^n Z)$$

covering the identity on Z such that

$$\mathbf{d}J(\xi) = \mathbf{i}_{\xi_Z}\Omega, \quad (4C.3)$$

where $J(\xi)$ is the n -form on Z whose value at $z \in Z$ is $\langle J(z), \xi \rangle$. A covariant momentum map is said to be **Ad*-equivariant** if

$$J(\text{Ad}_\eta^{-1}\xi) = \eta_Z^*[J(\xi)]. \quad (4C.4)$$

If \mathcal{G} acts by *special* covariant canonical transformations, there is an explicit formula for a **special covariant momentum map**. We claim that

$$J(\xi) = \mathbf{i}_{\xi_Z}\Theta \quad (4C.5)$$

is a covariant momentum map. Indeed,

$$\mathbf{d}J(\xi) = \mathbf{d}\mathbf{i}_{\xi_Z}\Theta = (\mathcal{L}_{\xi_Z} - \mathbf{i}_{\xi_Z}\mathbf{d})\Theta = \mathbf{i}_{\xi_Z}\Omega.$$

Proposition 4.3. *Special covariant momentum maps are Ad^* -equivariant.*

The proof is analogous to that for the cotangent bundle case (Abraham and Marsden [1978], Theorem 4.2.10). For the particular case of lifted actions, we have:

Proposition 4.4. *If the action of \mathcal{G} on Z is the lift of an action of \mathcal{G} on Y , then the special covariant momentum map is given by*

$$J(\xi)(z) = \pi_{YZ}^* \mathbf{i}_{\xi_Y} z. \quad (4C.6)$$

This momentum map is Ad^ -equivariant.*

This follows from the recognition that $\xi_Y = T\pi_{YZ} \cdot \xi_Z$, the definitions of $J(\xi)$ and Θ , and Proposition 4.3.

In coordinates, if ξ_Y has components (ξ^μ, ξ^A) , then equation (4C.6) reads

$$J(\xi)(z) = (p_A^\mu \xi^A + p \xi^\mu) d^n x_\mu - p_A^\mu \xi^\nu dy^A \wedge d^{n-1} x_{\mu\nu}, \quad (4C.7)$$

where

$$d^{n-1} x_{\mu\nu} = \mathbf{i}_{\partial_\nu} \mathbf{i}_{\partial_\mu} d^{n+1} x.$$

In passing, we reformulate equivariance for lifted actions in terms of Poisson brackets, just as one does in the usual context of symplectic and Poisson manifolds. Observe that according to (4C.6), special covariant momentum maps are “momentum observables” in the sense of (2B.22).

Proposition 4.5. *For lifted actions,*

$$\{J(\xi), J(\zeta)\} = \mathbf{d}(\mathbf{i}_{\xi_Z} \mathbf{i}_{\zeta_Z} \Theta) + J([\xi, \zeta]). \quad (4C.8)$$

Proof. Set $\eta_\lambda = \exp(\lambda\zeta)$. Differentiating the equivariance relation (4C.4) with respect to λ at $\lambda = 0$ gives

$$J([\xi, \zeta]) = \mathcal{L}_{\zeta_Z} J(\xi) = \mathbf{d}\mathbf{i}_{\zeta_Z} J(\xi) + \mathbf{i}_{\zeta_Z} \mathbf{d}J(\xi) = \mathbf{d}\mathbf{i}_{\zeta_Z} J(\xi) + \mathbf{i}_{\zeta_Z} \mathbf{i}_{\xi_Z} \Omega$$

by (4C.3). This, (4C.5), and (2B.23) give (4C.8). ■

Examples

Since X is assumed oriented, $\text{Diff}(X)$ will henceforth denote the group of *orientation-preserving* diffeomorphisms of X .

a Particle Mechanics. For classical mechanics with time reparametrization freedom (see, for example, Kuchař [1973]), we let $\mathcal{G} = \text{Diff}(\mathbb{R})$, the diffeomorphism group of $X = \mathbb{R}$. Then \mathcal{G} acts on $Y = \mathbb{R} \times Q$ by time reparametrizations and this action may be lifted to $Z = T^*\mathbb{R} \times T^*Q$. We identify \mathfrak{g} with the space $\mathcal{F}(\mathbb{R})$ of smooth functions on \mathbb{R} and identify $\mathcal{F}(\mathbb{R})^*$, the space of densities on \mathbb{R} , with $\mathcal{F}(\mathbb{R})$. In terms of the coordinates of Example **a** of §2B, we have

$$\chi_Y = \chi(t) \frac{\partial}{\partial t}$$

for $\chi \in \mathcal{F}(\mathbb{R})$. The momentum map

$$J : Z = T^*\mathbb{R} \times T^*Q \rightarrow \mathcal{F}(\mathbb{R})^* \approx \mathcal{F}(\mathbb{R})$$

is defined by

$$\langle J(t, p, q^1, \dots, q^N, p_1, \dots, p_N), \chi \rangle = p \chi(t). \quad (4C.9)$$

If we consider a classical mechanical system with a fixed time parametrization and with a group G acting on Q , then the covariant momentum map induced on Z is the usual momentum map on T^*Q regarded as a map on Z .

In this example, note that *covariant canonical maps include time-dependent canonical transformations*.

b Electromagnetism. The appropriate invariance group for electromagnetism on a fixed spacetime background X is $\mathcal{G} = \mathcal{F}(X)$, the additive group of smooth functions on spacetime. Let \mathcal{G} act on $Y = \Lambda^1 X$ as follows: For $f \in \mathcal{F}(X)$ and $A \in \Lambda_x^1 X$,

$$f \cdot A = A + \mathbf{d}f(x) \in \Lambda_x^1 X. \quad (4C.10)$$

For $\chi \in \mathfrak{g} \cong \mathcal{F}(X)$, the infinitesimal generator is given by

$$\chi_Y(A) = \chi_{,\nu} \frac{\partial}{\partial A_\nu}. \quad (4C.11)$$

From (4C.7), the corresponding covariant momentum map is

$$\langle J(x, A, p, \mathfrak{F}), \chi \rangle = \mathfrak{F}^{\nu\mu} \chi_{,\nu} d^3 x_\mu. \quad (4C.12)$$

If we treat the metric g on X parametrically, we may enlarge the group \mathcal{G} to be the semi-direct product $\text{Diff}(X) \ltimes \mathcal{F}(X)$, with $\text{Diff}(X)$ acting on $\mathcal{F}(X)$ by push-forward; that is, $\eta \cdot f = f \circ \eta^{-1}$. Then \mathcal{G} acts on $\tilde{Y} = Y \times_X S_2^{3,1}(X)$ by

$$(\eta, f) \cdot (A; g) = \left((\eta^{-1})^*(A + \mathbf{d}f(x)); (\eta^{-1})^* g \right). \quad (4C.13)$$

The infinitesimal generator $(\xi, \chi)_{\tilde{Y}}$ at the point $(A; g) \in \tilde{Y}_x$ is

$$(\xi_{S_2^{3,1}(X)}(g))_{\sigma\rho} \frac{\partial}{\partial g_{\sigma\rho}} + (\xi_Y(A))_\nu \frac{\partial}{\partial A_\nu} + (\chi_Y(A))_\nu \frac{\partial}{\partial A_\nu} + \xi^\mu(x) \frac{\partial}{\partial x^\mu}, \quad (4C.14)$$

where

$$(\xi_{S_2^{3,1}(X)}(g))_{\sigma\rho} = -(g_{\sigma\mu} \xi^\mu{}_{,\rho} + g_{\rho\mu} \xi^\mu{}_{,\sigma}), \quad (4C.15)$$

$$(\xi_Y(A))_\nu = -A_\mu \xi^\mu{}_{,\nu}, \quad (4C.16)$$

and $\chi_Y(A)$ is given by (4C.11). From (4C.7), the multimomentum map $\tilde{J} : \tilde{Z} \rightarrow \mathfrak{g}^* \otimes \Lambda^3 Z$ for this action is

$$\begin{aligned} \langle \tilde{J}(x, A, p, \mathfrak{F}; g), (\xi, \chi) \rangle = \\ (-\mathfrak{F}^{\nu\mu} A_\tau \xi^\tau{}_{,\nu} + \mathfrak{F}^{\nu\mu} \chi_{,\nu} + p \xi^\mu) d^3 x_\mu - \mathfrak{F}^{\tau\mu} \xi^\nu dA_\tau \wedge d^2 x_{\mu\nu}. \end{aligned} \quad (4C.17)$$

c A Topological Field Theory. Just as in Example **b** above, we take $\mathcal{G} = \text{Diff}(X) \otimes \mathcal{F}(X)$ acting on Y according to

$$(\eta, f) \cdot A = (\eta^{-1})^*(A + \mathbf{d}f(x)). \quad (4C.18)$$

The corresponding infinitesimal generator is

$$(\xi, \chi)_Y(A) = (\xi_Y(A))_\nu \frac{\partial}{\partial A_\nu} + (\chi_Y(A))_\nu \frac{\partial}{\partial A_\nu} + \xi^\mu(x) \frac{\partial}{\partial x^\mu}, \quad (4C.19)$$

where $(\xi_Y(A))_\nu$ and $\chi_Y(A)$ are given by (4C.16) and (4C.11), respectively. Then we compute

$$\begin{aligned} \langle J(x^\mu, A_\nu, p, p^{\nu\mu}), (\xi, \chi) \rangle = \\ (-p^{\nu\mu} A_\tau \xi^\tau{}_{,\nu} + p^{\nu\mu} \chi_{,\nu} + p \xi^\mu) d^2 x_\mu - p^{\tau\mu} \xi^\nu dA_\tau \wedge d^1 x_{\mu\nu}. \end{aligned} \quad (4C.20)$$

d Bosonic Strings. Corresponding to $Y = (X \times M) \times_X S_2^{1,1}(X)$, we take the invariance group for bosonic strings to be the semi-direct product

$$\mathcal{G} = \text{Diff}(X) \otimes \text{Con}_2^{1,1}(X),$$

where $\text{Con}_2^{1,1}(X)$ is the group of conformal transformations of $S_2^{1,1}(X)$ (which may be identified with the group $\mathcal{F}(X, \mathbb{R}^+)$ of positive real-valued functions on X) and $\eta \in \text{Diff}(X)$ acts on $\Lambda \in \text{Con}_2^{1,1}(X)$ by push-forward:

$$\eta \cdot \Lambda = \Lambda \circ \eta^{-1}. \quad (4C.21)$$

Then $(\eta, \Lambda) \in \mathcal{G}$ acts on $(x, \phi, h) \in Y_x$ by

$$(\eta, \Lambda) \cdot (\phi, h) = (\phi, (\eta^{-1})^*(\Lambda^2(x)h)). \quad (4C.22)$$

Hence, for $(\xi, \lambda) \in \mathfrak{g} \cong \mathfrak{X}(X) \otimes \mathcal{F}(X)$, the corresponding infinitesimal generator $(\xi, \lambda)_Y$ at $(\phi, h) \in Y$ is

$$(\lambda_Y(h))_{\sigma\rho} \frac{\partial}{\partial h_{\sigma\rho}} + (\xi_Y(h))_{\sigma\rho} \frac{\partial}{\partial h_{\sigma\rho}} + \xi^\mu \frac{\partial}{\partial x^\mu}, \quad (4C.23)$$

where

$$(\lambda_Y(h))_{\sigma\rho} = 2\lambda h_{\sigma\rho} \quad (4C.24)$$

and

$$(\xi_Y(h))_{\sigma\rho} = -(h_{\sigma\mu} \xi^\mu{}_{,\rho} + h_{\rho\mu} \xi^\mu{}_{,\sigma}). \quad (4C.25)$$

Note that there is no component in the direction $\partial/\partial\phi^A$ since ϕ is a scalar field on X . From (4C.7), we get the covariant momentum map

$$\begin{aligned} \langle J(x^\mu, \phi^A, h_{\sigma\rho}, p, p_A^\mu, q^{\sigma\rho\mu}), (\xi, \chi) \rangle \\ = [q^{\sigma\rho\mu} (2\lambda h_{\sigma\rho} - h_{\sigma\nu} \xi^\nu{}_{,\rho} - h_{\rho\nu} \xi^\nu{}_{,\sigma}) + p \xi^\mu] d^1 x_\mu \\ - (p_A^\mu \xi^\nu d\phi^A + q^{\sigma\rho\mu} \xi^\nu dh_{\sigma\rho}) \epsilon_{\mu\nu} \end{aligned} \quad (4C.26)$$

where we have written $\epsilon_{\mu\nu} = d^2 x_{\mu\nu}$.

Remark As we shall see in Part III the covariant momentum maps (4C.12), (4C.20), and (4C.26) are closely related to the divergence, spatial flatness, and superhamiltonian and supermomentum constraints of electromagnetism, Chern–Simons theory, and string theory, respectively. \blacklozenge

The covariant momentum maps discussed and illustrated in this section do not involve the Lagrangian at all. In the next section, however, we will show that when a theory possesses symmetries, an important link between the Lagrangian and the covariant momentum map may be established.

4D Symmetries and Noether's Theorem

Having assembled the necessary background in §§4A–C, we are now able to state and prove Noether's Theorem. This fundamental result will be used in Part III to show that the energy-momentum map in the Lagrangian representation vanishes identically on solutions of the Euler–Lagrange equations. As we shall see in §§10 and 11, the vanishing of this energy-momentum map on a single

Cauchy surface is often equivalent to the totality of the first class constraints of the theory. However, its vanishing on a whole family of hypersurfaces is equivalent to the entire Euler–Lagrange system under a suitable hypothesis on the gauge group; see §9C.

Suppose that a group \mathcal{G} acts on Y by bundle automorphisms. Prolong this action to J^1Y by setting $\eta \cdot \gamma = \eta_{J^1Y}(\gamma)$ as in §4A. We say that the Lagrangian density \mathcal{L} is **equivariant** with respect to \mathcal{G} if for all $\eta \in \mathcal{G}$ and $\gamma \in J_x^1Y$,

$$\mathcal{L}(\eta_{J^1Y}(\gamma)) = (\eta_X^{-1})^* \mathcal{L}(\gamma), \quad (4D.1)$$

where $(\eta_X^{-1})^* \mathcal{L}(\gamma)$ means that the $(n+1)$ -form $\mathcal{L}(\gamma)$ at $x \in X$ is pushed forward to an $(n+1)$ -form at $\eta_X(x)$; equality in (4D.1) thus means equality of $(n+1)$ -forms at the point $\eta(x)$. The infinitesimal version of this is the following equation obtained from (3A.1) and (4A.4):

$$\frac{\partial L}{\partial x^\mu} \xi^\mu + \frac{\partial L}{\partial y^A} \xi^A + \frac{\partial L}{\partial v^A_\mu} \left(\xi^A_{,\mu} - v^A_\nu \xi^\nu_{,\mu} + v^B_\mu \frac{\partial \xi^A}{\partial y^B} \right) + L \xi^\mu_{,\mu} = 0. \quad (4D.2)$$

The quantity $\delta_\xi L$ defined by the left hand side of this equation is called the **variation** of L in the direction ξ .

One of the fundamental assumptions that we shall make in this work is the following:

A1 Covariance. *The group \mathcal{G} acts on Y by π_{XY} -bundle automorphisms and the Lagrangian density \mathcal{L} is equivariant with respect to \mathcal{G} .*

This is a strong assumption but, amazingly, the gauge groups of most field theories satisfy it. One general class of exceptions are the topological field theories, whose gauge groups tend to leave $\Omega_{\mathcal{L}}$, but not \mathcal{L} , invariant. We will discuss how to deal with one such case in Example **c** below. This assumption has a number of important consequences, some of which are stated in the following proposition.

Proposition 4.6. *Let \mathcal{L} be equivariant with respect to the lifted action of \mathcal{G} on J^1Y . Then*

- (i) $\mathbb{F}\mathcal{L}$ is also equivariant; that is, $\eta_Z \circ \mathbb{F}\mathcal{L} = \mathbb{F}\mathcal{L} \circ \eta_{J^1Y}$. Thus, the diagram

$$\begin{array}{ccc} J^1Y & \xrightarrow{\mathbb{F}\mathcal{L}} & Z \\ \eta_{J^1Y} \downarrow & & \downarrow \eta_Z \\ J^1Y & \xrightarrow{\mathbb{F}\mathcal{L}} & Z \end{array}$$

commutes for every $\eta \in \mathcal{G}$,

- (ii) the Cartan form $\Theta_{\mathcal{L}}$ is invariant; that is, $\eta_{J^1Y}^* \Theta_{\mathcal{L}} = \Theta_{\mathcal{L}}$ for all $\eta \in \mathcal{G}$, and

- (iii) the map $J^{\mathcal{L}}(\xi) := \mathbb{F}\mathcal{L}^*J(\xi) : J^1Y \rightarrow \Lambda^n(J^1Y)$ is a momentum map for the lifted action of \mathcal{G} on J^1Y relative to $\Omega_{\mathcal{L}}$; that is, for all $\xi \in \mathfrak{g}$,

$$\xi_{J^1Y} \lrcorner \Omega_{\mathcal{L}} = \mathbf{d}J^{\mathcal{L}}(\xi), \quad (4D.3)$$

where $\xi_{J^1Y} = j^1\xi_Y$ is the infinitesimal generator corresponding to ξ . Moreover,

$$J^{\mathcal{L}}(\xi) = \xi_{J^1Y} \lrcorner \Theta_{\mathcal{L}}. \quad (4D.4)$$

Proof.

- (i) It is convenient to identify Z with J^1Y^* and use Proposition 4.2. From (4B.5) and (3A.3), we have

$$\begin{aligned} \langle \eta_{J^1Y^*}(\mathbb{F}\mathcal{L}(\gamma)), \gamma' \rangle &= (\eta_X^{-1})^* \langle \mathbb{F}\mathcal{L}(\gamma), \eta_{J^1Y}^{-1}(\gamma') \rangle \\ &= (\eta_X^{-1})^* \left[\mathcal{L}(\gamma) + \frac{d}{d\varepsilon} \mathcal{L}(\gamma + \varepsilon[\eta_{J^1Y}^{-1}(\gamma') - \gamma]) \right] \Big|_{\varepsilon=0}. \end{aligned}$$

On the other hand, using (3A.3), we have

$$\langle \mathbb{F}\mathcal{L}(\eta_{J^1Y}(\gamma)), \gamma' \rangle = \mathcal{L}(\eta_{J^1Y}(\gamma)) + \frac{d}{d\varepsilon} \mathcal{L}(\eta_{J^1Y}(\gamma) + \varepsilon[\gamma' - \eta_{J^1Y}(\gamma)]) \Big|_{\varepsilon=0}.$$

These are equal by the covariance condition **A1**.

- (ii) The proof is analogous to that for Corollary 4.2.14 of Abraham and Marsden [1978].
- (iii) The infinitesimal version of (i) states

$$\xi_Z \circ \mathbb{F}\mathcal{L} = T\mathbb{F}\mathcal{L} \circ \xi_{J^1Y}. \quad (4D.5)$$

Pulling (4C.3) back from Z to J^1Y along the Legendre transformation and using (4D.5) yields (4D.3). Similarly, pulling (4C.5) back by $\mathbb{F}\mathcal{L}$ yields (4D.4). \blacksquare

Because of this theorem, we call $J^{\mathcal{L}}$ the *covariant momentum map* in the *Lagrangian representation*. From (4D.4) and (3B.3) or, equivalently, from (4C.7) and (3A.2), we have the explicit formula

$$J^{\mathcal{L}}(\xi) = \left(\frac{\partial L}{\partial v^A_\mu} \xi^A + \left[L - \frac{\partial L}{\partial v^A_\nu} v^A_\nu \right] \xi^\mu \right) d^n x_\mu - \frac{\partial L}{\partial v^A_\mu} \xi^\nu dy^A \wedge d^{n-1} x_{\mu\nu}. \quad (4D.6)$$

If ϕ is a solution of the Euler–Lagrange equations, then from Theorem 3.1,

$$(j^1\phi)^*(W \lrcorner \Omega_{\mathcal{L}}) = 0$$

for any vector field W on J^1Y . In particular, setting $W = \xi_{J^1Y}$ and applying $(j^1\phi)^*$ to (4D.3), we obtain the following basic **Noether conservation law**:

Theorem 4.7 (Divergence Form of Noether's Theorem). *If the covariance assumption **A1** is satisfied, then for each $\xi \in \mathfrak{g}$,*

$$\mathbf{d}[(j^1\phi)^*J^{\mathcal{L}}(\xi)] = 0 \quad (4D.7)$$

for any section ϕ of π_{XY} satisfying the Euler–Lagrange equations.

The quantity $(j^1\phi)^*J^{\mathcal{L}}(\xi)$ is called the **Noether current**, and Theorem 4.7 is often referred to as the **first Noether theorem**.⁶ We now proceed to compute coordinate expressions for the Noether current and its divergence without assuming that the field equations are satisfied. In the process, we will derive a criteria for when a theory has a converse to Noether's theorem; that is, when the Noether conservation law implies the Euler–Lagrange equations.

From (4D.6) we get

$$\begin{aligned} & (j^1\phi)^*J^{\mathcal{L}}(\xi) \\ &= \left(\frac{\partial L}{\partial v^A{}_{\mu}}(j^1\phi)(\xi^A \circ \phi) + L(j^1\phi)\xi^{\mu} - \frac{\partial L}{\partial v^A{}_{\nu}}(j^1\phi)\phi^A{}_{,\nu}\xi^{\mu} \right) d^n x_{\mu} \\ & \quad - \frac{\partial L}{\partial v^A{}_{\mu}}(j^1\phi)\phi^A{}_{,\lambda}\xi^{\nu} dx^{\lambda} \wedge d^{n-1}x_{\mu\nu} \end{aligned}$$

which, upon using the identity

$$dx^{\lambda} \wedge d^{n-1}x_{\mu\nu} = \delta_{\nu}^{\lambda} d^n x_{\mu} - \delta_{\mu}^{\lambda} d^n x_{\nu},$$

simplifies to

$$(j^1\phi)^*J^{\mathcal{L}}(\xi) = \left[\frac{\partial L}{\partial v^A{}_{\mu}}(j^1\phi)(\xi^A \circ \phi - \phi^A{}_{,\nu}\xi^{\nu}) + L(j^1\phi)\xi^{\mu} \right] d^n x_{\mu}. \quad (4D.8)$$

Defining the **Lie derivative** of ϕ along ξ by

$$\mathcal{L}_{\xi}\phi = T\phi \circ \xi_X - \xi_Y \circ \phi; \quad \text{i.e.,} \quad (\mathcal{L}_{\xi}\phi)^A = \phi^A{}_{,\nu}\xi^{\nu} - \xi^A \circ \phi, \quad (4D.9)$$

(4D.8) may be rewritten as

$$(j^1\phi)^*J^{\mathcal{L}}(\xi) = \left[-\frac{\partial L}{\partial v^A{}_{\mu}}(j^1\phi)(\mathcal{L}_{\xi}\phi)^A + L(j^1\phi)\xi^{\mu} \right] d^n x_{\mu} \quad (4D.10)$$

⁶ The “second Noether theorem,” which we call the “Vanishing Theorem,” will be discussed in §9.

where ϕ is any section of π_{XY} , not necessarily satisfying the Euler–Lagrange equations.

Next, taking the exterior derivative of the Noether current (4D.8) and using the identity

$$\mathbf{d}(V^\mu d^n x_\mu) = \partial_\rho V^\mu dx^\rho \wedge d^n x_\mu = \partial_\mu V^\mu d^{n+1}x$$

(which justifies the name “divergence” for the resulting quantity), we obtain after some rearrangement and cancellation,

$$\begin{aligned} \mathbf{d}[(j^1\phi)^* J^\mathcal{L}(\xi)] &= \partial_\mu \left[\frac{\partial L}{\partial v^A_\mu}(j^1\phi)(\xi^A \circ \phi - \phi^A_{,\nu} \xi^\nu) + L(j^1\phi)\xi^\mu \right] d^{n+1}x \\ &= \left\{ \left[\frac{\partial L}{\partial y^A}(j^1\phi) - \frac{\partial}{\partial x^\mu} \left(\frac{\partial L}{\partial v^A_\mu}(j^1\phi) \right) \right] (-\xi^A \circ \phi + \phi^A_{,\nu} \xi^\nu) \right. \\ &\quad + \frac{\partial L}{\partial x^\mu}(j^1\phi)\xi^\mu + \frac{\partial L}{\partial y^A}(j^1\phi)(\xi^A \circ \phi) \\ &\quad + \frac{\partial L}{\partial v^A_\mu}(j^1\phi) \left[\frac{\partial \xi^A}{\partial x^\mu} + \frac{\partial \xi^A}{\partial y^B} v^B_{\mu} - \frac{\partial \xi^\nu}{\partial x^\mu} v^A_{\nu} \right] (j^1\phi) \\ &\quad \left. + L(j^1\phi) \frac{\partial \xi^\mu}{\partial x^\mu} \right\} d^{n+1}x \\ &= \left\{ \frac{\delta L}{\delta \phi^A} (\mathcal{L}_\xi \phi)^A + \delta_\xi L \right\} (j^1\phi) d^{n+1}x. \end{aligned} \quad (4D.11)$$

Here $\delta L/\delta \phi^A$ is the Euler–Lagrange derivative of \mathcal{L} defined in (3C.1), which vanishes when the field equations are satisfied, and $\delta_\xi L$ is the variation of \mathcal{L} defined in (4D.2), which vanishes when the Lagrangian is equivariant. Thus equation (4D.11) again makes it obvious that the Noether conservation laws hold when the Lagrangian is equivariant and the Euler–Lagrange equations are satisfied.

To obtain a converse, we say that the \mathcal{G} -action is **vertically transitive** if for each $x \in X$ and each $y \in Y_x$ and each (local) section ϕ through y ,

$$\{(\mathcal{L}_\xi \phi)(x) \mid \xi \in \mathfrak{g}\} = V_y Y. \quad (4D.12)$$

Looking forward to Part II, §7A, we note that this is the infinitesimal version of the statement that at each point $x \in X$, there exists a group element that transforms any given value of the *fields* at x into any other specified value. This is true for metrics and connections under the diffeomorphism group and for connections under bundle automorphisms, but is rarely true for other fields.

Equation (4D.11) now implies

Theorem 4.8. *Assume that the action of \mathcal{G} on Y is vertically transitive and that the Lagrangian density is \mathcal{G} -equivariant. Then a local section ϕ of Y is a solution to the Euler–Lagrange equations (3C.1) if and only if it satisfies the Noether conservation law (4D.7) for all $\xi \in \mathfrak{g}$.*

To check whether the \mathcal{G} -action for a given field theory is vertically transitive, we may use the equivalent formulations given in:

Proposition 4.9. *Assume \mathcal{G} acts on Y and let*

$$\mathfrak{g}(y) = \{\xi_Y(y) \mid \xi \in \mathfrak{g}\}$$

denote the span of the infinitesimal generators of the action at $y \in Y$; that is, the tangent space to the orbit through y . Vertical transitivity of the action of \mathcal{G} on Y is equivalent to either of the following conditions:

- (i) *For each $y \in Y$ and each section ϕ with $\phi(x) = y$, the image of ϕ is transverse to $\mathfrak{g}(y)$; that is,*

$$\mathfrak{g}(y) + \text{im } T_x \phi = T_y Y. \quad (4D.13)$$

- (ii) *For each $y \in Y$ and every choice of bundle coordinates (x^μ, y^A) , we have*

$$\{\xi^A(y) \partial_A \mid \xi \in \mathfrak{g}\} = V_y Y, \quad (4D.14)$$

where $\xi_Y = \xi^\mu \partial_\mu + \xi^A \partial_A$.

Note that an action can be vertically transitive even if none of the ξ_Y are vertical; note also from (4D.9) that $-(\mathcal{L}_\xi \phi)(x)$ is exactly the vertical component of $\xi_Y(\phi(x))$.

Examples

In each of the following examples, we compute the Noether current and its divergence without assuming the fields satisfy the Euler–Lagrange equations, and we then check for vertical transitivity.

a Particle Mechanics. From the infinitesimal covariance condition (4D.2), we see that a particle Lagrangian is $\text{Diff}(\mathbb{R})$ -equivariant (that is to say, time reparametrization-invariant) if and only if it is time-independent and the energy $E = v^A(\partial L/\partial v^A) - L$ vanishes. Suppose also that there is a group G acting on Q whose prolongation to TQ leaves the Lagrangian invariant. Take $\mathcal{G} = \text{Diff}(\mathbb{R}) \times G$ and write elements of \mathfrak{g} as pairs (χ, ξ) . Then (4D.8) reads

$$\begin{aligned} (j^1 \phi)^* J^\mathcal{L}(\chi, \xi) &= \left[\frac{\partial L}{\partial v^A} (\xi^A - v^A \chi) + L \chi \right] \\ &= (J^L(\xi) - E \chi) \\ &= J^L(\xi), \end{aligned} \quad (4D.15)$$

where $J^L(\xi) = (\partial v^A)\xi^A$ is the usual momentum map for G acting on Q in Lagrangian mechanics. Then (4D.7) asserts that on solutions of the Euler–Lagrange equations,

$$\frac{d}{dt}J^L(\xi) = 0. \quad (4D.16)$$

This group \mathcal{G} is vertically transitive iff G acts transitively on Q .

In the special case of the relativistic free particle, the Lagrangian density (3C.8) is $\text{Diff}(\mathbb{R})$ -equivariant. Then the corresponding momentum map vanishes identically, consistent with the fact that $p = 0$, cf. (3C.10) and (4C.9). On the other hand, suppose that (Q, g) is Minkowski space and that G is the Poincaré group. Then \mathcal{L} is G -invariant and (4D.16) expresses the fact that relativistic energy-momentum and angular momentum are constant along dynamical trajectories. This is equivalent to the statement that the dynamical trajectories are geodesics of g .

b Electromagnetism. Let $\mathcal{G} = \mathcal{F}(X)$ act on $Y = \Lambda^1 X$ as in Example **b** of §4C. The Maxwell Lagrangian density (3C.12) is invariant under the prolongation of $\mathcal{F}(X)$ to $J^1(\Lambda^1 X)$. In this case (4D.7) reads

$$\begin{aligned} \mathbf{d}[(j^1 A)^* J^{\mathcal{L}}(\chi)] &= \mathbf{d}[(j^1 A)^*(F^{\nu\mu} \sqrt{-g} \chi_{,\nu} d^3 x_\mu)] \\ &= \mathbf{d}[(A^{\mu,\nu} - A^{\nu,\mu}) \sqrt{-g} \chi_{,\nu} d^3 x_\mu] \\ &= [(A^{\mu,\nu} - A^{\nu,\mu}) \sqrt{-g}]_{,\mu} \chi_{,\nu} d^4 x = 0 \end{aligned} \quad (4D.17)$$

where we have used $dx^\lambda \wedge d^3 x_\mu = \delta_\mu^\lambda d^4 x$ and noted that the terms $\chi_{,\mu\nu}$ have dropped out by symmetry. Since χ is arbitrary, (4D.17) is equivalent to

$$(A^{\mu,\nu} - A^{\nu,\mu})_{;\mu} = 0 \quad (4D.18)$$

which are Maxwell's equations. From (4C.11) it is clear that the group $\mathcal{F}(X)$ is vertically transitive.

Observe that when the spacetime metric g is fixed, the Maxwell Lagrangian density (3C.12) is not $\text{Diff}(X)$ -equivariant: $\mathcal{L}((\eta, 0)_{J^1 Y}(\gamma)) = \mathcal{L}(\gamma)$, whereas

$$(\eta_X^{-1})^* \mathcal{L}(\gamma) = \mathcal{L}(\gamma) J^{-1},$$

J being the Jacobian determinant of η_X . We can remedy this by allowing g to vary parametrically, in which case the $\sqrt{-g}$ factor in \mathcal{L} will give rise to a factor of J which will cancel the offending factor J^{-1} in the above. Combined with our previous results, we see that parametrized electromagnetism is $(\text{Diff}(X) \otimes \mathcal{F}(X))$ -equivariant: $\delta_{(\xi, \chi)} L = 0$. But there is a price to be paid for this extended equivariance: Noether's theorem no longer holds for $\tilde{J}^{\mathcal{L}} = \mathbb{F}\mathcal{L}^* \tilde{J}$.

Indeed, from (4D.11) we see that even if the section A of $\Lambda^1 X$ satisfies Maxwell's equations,

$$\begin{aligned} \mathbf{d} \left[((j^1 A) \times g)^* \tilde{J}^{\mathcal{L}}(\xi, \chi) \right] &= \frac{\partial L}{\partial g_{\sigma\rho}} (\mathcal{L}_{\xi} g)_{\sigma\rho} d^4 x \\ &= \mathfrak{T}^{\sigma\rho} (\mathcal{L}_{\xi} g)_{\sigma\rho} d^4 x \end{aligned}$$

need not vanish, where

$$\mathfrak{T}^{\sigma\rho} = - \left(\frac{1}{4} g^{\sigma\rho} F_{\alpha\beta} F^{\alpha\beta} + g^{\rho\beta} F^{\alpha\sigma} F_{\beta\alpha} \right) \sqrt{-g} \quad (4D.19)$$

is the stress-energy-momentum tensor of the electromagnetic field (cf. Gotay and Marsden [1992] and Interlude II). The reason, of course, is that g is not variational, whence $\delta L / \delta g_{\sigma\rho} = \partial L / \partial g_{\sigma\rho} \neq 0$ necessarily.

c A Topological Field Theory. The behavior of the Chern–Simons field is exactly opposite to that of parametrized electromagnetism: the Lagrangian density (3C.18) is equivariant under the action of $\text{Diff}(X)$, but not under that of $\mathcal{F}(X)$. Indeed, we compute

$$\delta_{(\xi, \chi)} L = \frac{1}{2} \epsilon^{\mu\nu\sigma} F_{\mu\nu} \chi_{,\sigma}. \quad (4D.20)$$

Consequently the Legendre transformation is not equivariant, and the Cartan form is not invariant.

Nonetheless, we can verify by direct computation that the Noether current

$$\begin{aligned} (j^1 A)^* J^{\mathcal{L}}(\xi, \chi) &= \\ & \left(\epsilon^{\mu\nu\sigma} (-A_{\tau} \xi^{\tau}_{,\nu} - A_{\nu,\tau} \xi^{\tau} + \chi_{,\nu}) + \frac{1}{2} \epsilon^{\nu\tau\sigma} F_{\nu\tau} \xi^{\mu} \right) A_{\sigma} d^2 x_{\mu} \end{aligned} \quad (4D.21)$$

is conserved; we obtain

$$\mathbf{d} \left[(j^1 A)^* J^{\mathcal{L}}(\xi, \chi) \right] = \epsilon^{\mu\nu\sigma} F_{\mu\nu} \left(A_{\sigma,\tau} \xi^{\tau} + A_{\tau} \xi^{\tau}_{,\sigma} + \frac{1}{2} \chi_{,\sigma} \right)$$

which is zero when the Chern–Simons equations $F_{\mu\nu} = 0$ hold. Alternately, we could appeal to (4D.11) and observe from (4D.20) that $\delta_{(\xi, \chi)} L$ vanishes on shell.

As in the case of electromagnetism, the action of $\text{Diff}(X) \circledast \mathcal{F}(X)$ on $\Lambda^1(X)$ is vertically transitive. Since in particular the value of $\chi_{,\sigma}(x)$ is arbitrary, the conservation of the Noether current implies that $F_{\mu\nu}(x) = A_{\nu,\mu}(x) - A_{\mu,\nu}(x) = 0$ at each $x \in X$.

d Bosonic Strings. The harmonic map Lagrangian density defined by equation (3C.23) is equivariant under the action of $\text{Diff}(X) \otimes \text{Con}_2^{1,1}(X)$ on $Y = (X \times M) \times_X S_2^{1,1}(X)$. Using (4C.23)–(4C.25) to compute the Lie derivatives $\mathcal{L}_\xi \phi$, $\mathcal{L}_\xi h$, and $\mathcal{L}_\lambda h$ via (4D.9), (4D.11) gives

$$\mathbf{d}[j^1(\phi, h)^* J^\mathcal{L}(\xi, \lambda)] = \left(\frac{\delta L}{\delta \phi^A} \phi^A{}_{,\nu} \xi^\nu + \frac{\delta L}{\delta h_{\sigma\rho}} (h_{\sigma\rho,\nu} \xi^\nu - 2\lambda h_{\sigma\rho} + h_{\sigma\nu} \xi^\nu{}_{,\rho} + h_{\rho\nu} \xi^\nu{}_{,\sigma}) \right) d^2x. \quad (4D.22)$$

At a given point $x \in X$, the quantities λ , ξ^ν , and $\xi^\nu{}_{,\rho}$ can be specified arbitrarily. The arbitrariness of $\xi^\nu{}_{,\rho}$ implies that $\delta L / \delta h_{\sigma\rho} = 0$, and so we obtain the “conformal” Euler–Lagrange equation (3C.30) as a conserved Noether current. The arbitrariness of ξ^ν then gives

$$\frac{\delta L}{\delta \phi^A} \phi^A{}_{,\nu} = 0$$

which is *not* equivalent to the harmonic map equation (3C.29). This is consistent with the fact that \mathcal{G} is not vertically transitive; indeed (4D.12) is not satisfied since, by (4C.23), $(\xi, \lambda)_Y$ has no component in the $\partial/\partial\phi^A$ direction.

 INTERLUDE I—ON CLASSICAL FIELD THEORY

Classical field theories come in many forms and varieties, and it is useful to attempt a broad classification of such theories, based upon certain features which play an important role in this work. The first of these features devolves upon the notion of *gauge group* which is, roughly speaking, the largest localizable group of transformations of the theory under which the Lagrangian density is equivariant (so that the group maps solutions of the Euler–Lagrange equations to solutions). The concept of gauge group is explored more fully in Part III, §8; for this discussion, we use the term loosely. It will be helpful in what follows to keep in mind the examples in §4, since the groups studied there are in fact the gauge groups of their respective theories.

Consider a classical field theory with gauge group \mathcal{G} . Suppose that $\mathcal{G} \subset \text{Aut}(Y)$, the automorphism group of the covariant configuration bundle Y . We may distinguish two basic types of field theories based upon the relationship between the gauge group \mathcal{G} and the (spacetime) diffeomorphism group $\text{Diff}(X)$. The first consists of those which are *parametrized* in the sense that the natural homomorphism $\text{Aut}(Y) \rightarrow \text{Diff}(X)$ given by $\eta_Y \mapsto \eta_X$ maps \mathcal{G} onto $\text{Diff}(X)$ (or at least onto a “sufficiently large” subgroup thereof, such as the compactly supported diffeomorphisms). This terminology reflects the fact that such a theory is invariant under (essentially) arbitrary relabeling of the points of the parameter “spacetime” X . (In relativity theory, cf. Anderson [1967], one would say that X is a *relative* object in the theory.) The parametrized theory *par excellence* is of course general relativity, in which case \mathcal{G} equals the spacetime diffeomorphism group. Another example is provided by the relativistic free particle, whose gauge group consists of (oriented) time reparametrizations. The Polyakov and Nambu strings are also parametrized theories, as are most topological field theories (see Horowitz [1989]).

At the other extreme, there are the *background* theories in which $\mathcal{G} \subset \text{Aut}_{\text{Id}}(Y)$, the automorphisms of Y which cover the identity on X (or at least in which the image of \mathcal{G} in $\text{Diff}(X)$ is “sufficiently small”—the Poincaré group, for instance). In this case X is called an *absolute* element of the theory, much like time in Newtonian mechanics, or Minkowski spacetime in special relativity. The prototypical background theory is electromagnetism treated on a *fixed*, non-interacting, background spacetime; here \mathcal{G} consists of shifts of the vector potential by exact 1-forms and so is completely “internal.” Yang–Mills–Higgs theory is similar. Traditionally, Newtonian mechanical systems are treated as background theories; usually, such systems do not even have gauge groups.

In many ways, parametrized theories are preferred over background theories. According to the principle of general covariance, one should model physical systems so that their descriptions contain no absolute objects. From a more prosaic standpoint, parametrized theories are easier to handle since, for example, they readily admit compatible slicings (necessary for the space + time decomposition of the theory, cf. Part II, §6A and §6D). Only in a parametrized theory

can one properly correlate the dynamics and the initial value constraints of a theory with its gauge freedom, cf. Part II §7F, and Part III §13; in fact, the system must effectively be parametrized to do Hamiltonian dynamics “correctly,” as our development will show. For more information on parametrized theories compared to background theories, see Anderson [1967] and Kuchař [1973].

Fortunately, most background theories can readily be converted into parametrized theories. The conversion usually involves the promotion of previously fixed objects, notably a metric on X , into the ranks of the field variables. In this way one can often make the Lagrangian density $\text{Diff}(X)$ -equivariant, and thereby enlarge the gauge group to “include” the diffeomorphism group of X . Thus, for example, when one encounters a background theory, such as Yang–Mills–Higgs on Minkowski spacetime, one can parametrize it by simply “turning the metric on.” This is precisely how we have parametrized electromagnetism in Example **b**. As long as no new terms are added to the Lagrangian and the previously fixed fields are *not* varied to give Euler–Lagrange equations, the converted theory, while now parametrized, remains physically equivalent to the original background theory. Kuchař [1973] finds an alternate way to parametrize a background theory, by introducing certain slicing parameters, treating them as fields, and letting them evolve dynamically. Further details of this approach can be found in Kuchař [1976].

Among either class of field theories, we may further distinguish *metric* and *non-metric* theories. Metric theories are those which carry a (Lorentzian, Riemannian or general signature) metric *on the parameter space* X . Most familiar classical field theories are metric, with topological field theories and theories whose action functionals measure length, area, volume, etc., being important exceptions. Most mechanical systems (relativistic or nonrelativistic), on the other hand, are non-metric. (In this regard, we emphasize that our terminology refers to the metric structure of X , which in mechanics is usually taken to be the time line; the kinetic energy metrics familiar in mechanics live instead on the fibers of $Y \rightarrow X$.⁷)

In metric theories the metric may be an absolute object, in that it does not “evolve” or dynamically interact with any of the other elements of the theory (as in electromagnetism on a fixed spacetime background), or it may be treated as one of the field variables. Again, the principle of general covariance suggests that the metric should be treated as a relative object. When this is the case, the metric may be *variational* in the sense that it is varied to give field equations (as in theories of gravity), or it may appear essentially as a parameter in that it may take different values, but is *not* varied to give Euler–Lagrange equations. Electromagnetism on a variable but non-dynamic spacetime is an example of a theory in which the metric is treated *parametrically* in this fashion. Theories which contain dynamic or parametric metrics are typically parametrized whereas theories with fixed metrics are usually background theories.

A word of caution: If a theory has some fields which are parametric (but

⁷ Our terminology here is also at odds with the standard usage of “metric” and “non-metric” as applied to theories of gravity.

not variational) then the Noether conservation laws may not hold because the parametric field is not varied and so does not satisfy any field equations. We have already encountered this in Example **b** of §4D.

As noted above, our formalism is best applied to parametrized theories. Likewise, it is often advantageous to work in the context of metric theories, with the metric being treated dynamically. This can sometimes be achieved using the “Polyakov trick,” which transforms the Nambu string into the Polyakov string. We show how this goes for the relativistic free particle in Part III, §11C. We suspect that this procedure could be useful in treating some of the “nonstandard” vector field theories of Isenberg and Nester [1977].

We summarize our classification with a chart, indicating in which categories various field theories fall.

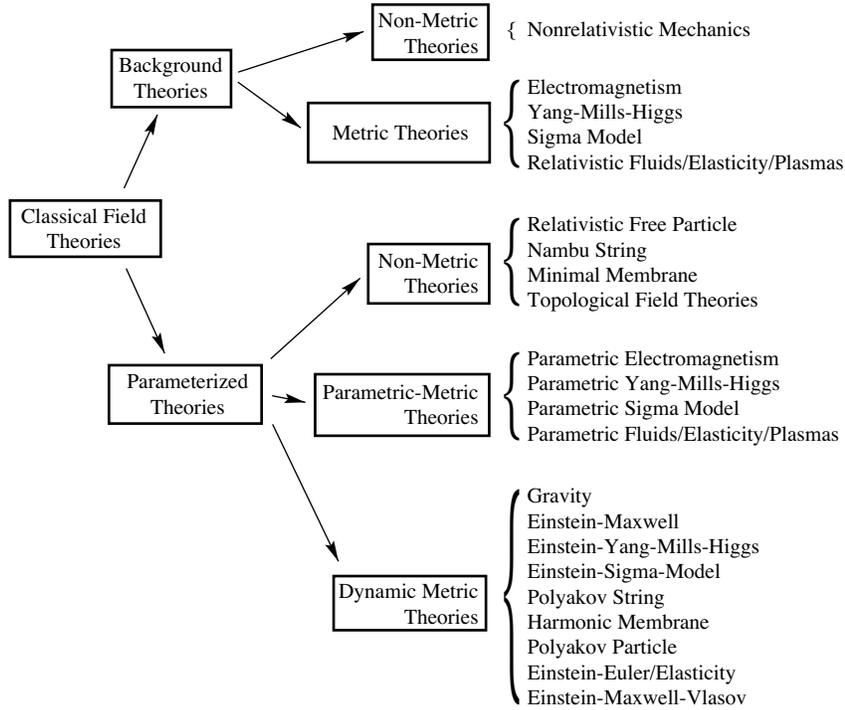


Figure 4.1: Classification of Field Theories

In this work, we discuss example theories from each category. But there are many other theories, not touched upon here, to which our formalism may be applied. Two especially important ones are fluids and elasticity (cf. Bao, Marsden, and Walton [1985]).

Let us specifically consider the case of a fluid in the *material representation*. The basic variables are the *particle placement fields* $\phi : \Sigma \times \mathbb{R} \rightarrow M$ which have the physical interpretation that $\phi(\{x\} \times \mathbb{R})$ is the world line of the particle with material label $x \in \Sigma$, where Σ is the fluid “reference manifold” and M is spacetime. Thus Y is the trivial bundle $(\Sigma \times \mathbb{R}) \times M$. Consistent with standard practice in continuum mechanics, M carries the spacetime metric while $\Sigma \times \mathbb{R}$ carries a reference metric h out of which things like the internal energy are built (cf. Marsden and Hughes [1983]). Observe that $\text{Diff}(\Sigma \times \mathbb{R})$ encodes the material symmetries (including particle relabeling and time reparametrization symmetries). In this formulation, the theory is already parametrized; note also that the metric h on $\Sigma \times \mathbb{R}$ is treated parametrically.

This approach to fluids is in keeping with our treatment of strings. But it is not particularly suited to a discussion of general relativistic fluids. In this context it is more convenient to use the *inverse material representation*. The fields then include the gravitational field (treated variationally or parametrically) with the fluid described by *particle labeling fields* $\psi : X \rightarrow \Sigma \times \mathbb{R}$, where $\psi = \phi^{-1}$. In addition to the material symmetries given by $\text{Diff}(\Sigma \times \mathbb{R})$ we have the spacetime-based $\text{Diff}(X)$ gauge symmetries associated with the gravitational field. The same “switch” to the inverse material representation obviously can be made for strings, and similar remarks apply to elasticity (with some qualifications).

References

- R. Abraham and J. Marsden [1978], *Foundations of Mechanics*, Second Edition, Addison-Wesley, Menlo Park, California.
- V. Aldaya and J. A. de Azcárraga [1980a], Geometric formulation of classical mechanics and field theory, *Rivista del Nuovo Cimento*, **3**, 1–66.
- V. Aldaya and J. A. de Azcárraga [1980b], Higher-order Hamiltonian formalism in field theory, *J. Phys. A: Math. Gen.*, **13**, 2545–2551.
- S. C. Anco and R. S. Tung [2002], Covariant Hamiltonian boundary conditions in General Relativity for spatially bounded space-time regions, *J. Math. Phys.*, **43**, 5531–5566.
- I. Anderson [1992], Introduction to the variational bicomplex, *Contemp. Math.*, **132**, 51–73.
- J. L. Anderson [1967], *Principles of Relativity Physics*, Academic Press, New York.
- J. M. Arms [1977], Linearization stability of the Einstein-Maxwell system, *J. Math. Phys.*, **18**, 830–833.
- J. M. Arms [1981], The structure of the solution set for the Yang-Mills equations, *Math. Proc. Camb. Phil. Soc.*, **90**, 361–372.
- J. M. Arms, R. Cushman, and M. J. Gotay [1991], A universal reduction procedure for Hamiltonian group actions, *The Geometry of Hamiltonian Systems* (T. Ratiu, ed.), M.S.R.I. Publ. **22**, 33–51, Springer-Verlag, New York.
- J. M. Arms, M. J. Gotay, and G. Jennings [1990], Geometric and algebraic reduction for singular momentum mappings, *Adv. in Math.*, **79**, 43–103.
- J. M. Arms, J. E. Marsden, and V. Moncrief [1981], Symmetry and bifurcations of momentum mappings, *Commun. Math. Phys.*, **78**, 455–478.
- J. Arms, J. Marsden, and V. Moncrief [1982], The structure of the space of solutions of Einstein's equations: II. Several Killings Fields and the Einstein-Yang-Mills equations, *Ann. Phys.* **144**, 81–106.
- R. Arnowitt, S. Deser, and C. W. Misner [1962], The dynamics of general relativity, *Gravitation, an Introduction to Current Research* (L. Witten, ed.), 227–265, Wiley, New York.
- A. Ashtekar [1986], New variables for classical and quantum gravity, *Phys. Rev. Lett.*, **57**, 2244–2247.
- A. Ashtekar [1987], New Hamiltonian formulation of general relativity, *Phys. Rev. D*, **36**, 1587–1602.

- A. Ashtekar, L. Bombelli, and O. Reula [1991], The covariant phase space of asymptotically flat gravitational fields, *Mechanics, Analysis and Geometry: 200 Years After Lagrange* (M. Francaviglia ed.), 417–450, North-Holland, Amsterdam.
- D. Bao [1984], A sufficient condition for the linearization stability of $N = 1$ supergravity: a preliminary report, *Ann. Phys.*, **158**, 211–278.
- D. Bao, J. Isenberg, and P. B. Yasskin [1985], The dynamics of the Einstein Dirac system. I. A principal bundle formulation of the theory and its canonical analysis, *Ann. Phys.*, **164**, 103–171.
- D. Bao, J. Marsden, and R. Walton [1985], The Hamiltonian structure of general relativistic perfect fluids, *Commun. Math. Phys.*, **99**, 319–345.
- C. Batlle, J. Gomis, and J. M. Pons [1986], Hamiltonian and Lagrangian constraints of the bosonic string, *Phys. Rev. D*, **34**, 2430–2432.
- R. Beig [1989], On the classical geometry of bosonic string dynamics, *Int. J. Theor. Phys.*, **30**, 211–224.
- F. J. Belinfante [1940], On the current and the density of the electric charge, the energy, the linear momentum and the angular momentum of arbitrary fields, *Physica*, **vii**, 449–474.
- E. Binz, J. Śniatycki, and H. Fischer [1988], *The Geometry of Classical Fields*, North Holland, Amsterdam.
- T. J. Bridges [1997], Multi-symplectic structures and wave propagation, *Math. Proc. Camb. Phil. Soc.*, **121**, 147–190.
- F. Cantrijn, L. A. Ibort, and M. de León [1999], On the geometry of multi-symplectic manifolds, *J. Aust. Math. Soc. A*, **66**, 303–330.
- C. Carathéodory [1935], *Variationsrechnung und Partielle Differentialgleichungen erster Ordnung*, Teubner, Leipzig (Reprinted by Chelsea, New York, 1982).
- É. Cartan [1922], *Leçons sur les Invariants Intégraux*, Hermann, Paris.
- M. Castrillon Lopez and J. E. Marsden [2002], Some remarks on Lagrangian and Poisson reduction for field theories, *Preprint*.
- D. Christodoulou [2000], *The Action Principle and Partial Differential Equations*, Princeton University Press, Princeton.
- Č. Crnković and E. Witten [1987], Covariant description of canonical formalism in geometrical theories, *Newton's Tercentenary Volume* (S.W. Hawking and W. Israel, eds.), 666–684, Cambridge University Press, Cambridge.

- P. Dedecker [1953], Calcul des variations, formes différentielles et champs géodésiques, *Colloq. Int. de Géométrie Différentielle* (Strasbourg), 17–34.
- P. Dedecker [1957], Calcul des variations et topologie algébrique, *Mem. Soc. Roy. Sc. Liège*, **19**, 1–216.
- P. Dedecker [1977], On the generalization of symplectic geometry to multiple integrals in the calculus of variations, *Lecture Notes in Math.*, **570**, 395–456, Springer-Verlag, New York.
- Th. De Donder [1930], *Théorie Invariantive du Calcul des Variations*, Gauthier-Villars, Paris.
- P. A. M. Dirac [1950], Generalized Hamiltonian dynamics, *Can. J. Math.*, **2**, 129–148.
- P. A. M. Dirac [1964], *Lectures on Quantum Mechanics*, Academic Press, New York.
- A. Echeverría-Enríques, M. C. Muñoz-Lecanda, and N. Román-Roy [2000], Geometry of multisymplectic Hamiltonian first-order field theories, *J. Math. Phys.*, **41**, 7402–7444.
- G. Esposito, G. Gionti, and C. Stornaiolo [1995], Spacetime covariant form of Ashtekar’s constraints, *Nuovo Cimento*, **B110**, 1137–1152.
- M. Ferraris and M. Francaviglia [1983], On the global structure of Lagrangian and Hamiltonian formalisms in higher order calculus of variations, *Proc. of the Meeting “Geometry and Physics”* (M. Modugno, ed.), 43–70, Pitagora Editrice, Bologna.
- A. E. Fischer and J. E. Marsden [1979a], Topics in the dynamics of general relativity, *Isolated Gravitating Systems in General Relativity* (J. Ehlers, ed.), Italian Physical Society, 322–395.
- A. E. Fischer and J. E. Marsden [1979b], The initial value problem and the dynamical formulation of general relativity, *General Relativity* (S.W. Hawking and W. Israel, eds.), 138–211, Cambridge Univ. Press, Cambridge.
- A. E. Fischer, J. E. Marsden, and V. Moncrief [1980], The structure of the space of solutions of Einstein’s equations I: One Killing field, *Ann. Inst. H. Poincaré*, **33**, 147–194.
- M. Forger, C. Paufler and H. Römer [2002a], The Poisson bracket for Poisson forms in multisymplectic field theory, *Preprint math-ph/0202043*.
- M. Forger, C. Paufler and H. Römer [2002b], A general construction of Poisson brackets on exact multisymplectic manifolds, *Preprint math-ph/0208037*.
- P. L. García [1974], The Poincaré-Cartan invariant in the calculus of variations, *Symp. Math.*, **14**, 219–246.

- P. L. García and J. Muñoz [1991], Higher order regular variational problems, *Symplectic Geometry and Mathematical Physics* (P. Donato, C. Duval, J. Elhadad, and G. M. Tuynman, eds.), Birkhäuser, Boston.
- K. Gawędzki [1972], On the geometrization of the canonical formalism in the classical field theory, *Rep. Math. Phys.*, **3**, 307–326.
- H. Goldschmidt and S. Sternberg [1973], The Hamilton-Cartan formalism in the calculus of variations, *Ann. Inst. Fourier*, **23**, 203–267.
- M.J. Gotay [1988], A multisymplectic approach to the KdV equation, *Differential Geometric Methods in Theoretical Physics* (K. Bleuler and M. Werner eds.), NATO Advanced Science Institutes Series C: Mathematical and Physical Sciences, **250**, 295–305, Kluwer, Dordrecht.
- M. J. Gotay [1991a], A multisymplectic framework for classical field theory and the calculus of variations I. Covariant Hamiltonian formalism, *Mechanics, Analysis, and Geometry: 200 Years After Lagrange* (M. Francaviglia, ed.), 203–235, North Holland, Amsterdam.
- M. J. Gotay [1991b], A multisymplectic framework for classical field theory and the calculus of variations II. Space + time decomposition, *Diff. Geom. Appl.*, **1**, 375–390.
- M. J. Gotay [1991c], An exterior differential systems approach to the Cartan form, *Symplectic Geometry and Mathematical Physics* (P. Donato, C. Duval, J. Elhadad, and G.M. Tuynman, eds.), 160–188, Birkhäuser, Boston.
- M. J. Gotay and J. E. Marsden [1992], Stress-energy-momentum tensors and the Belinfante–Rosenfeld formula, *Contemp. Math.*, **132**, 367–391.
- M. J. Gotay, J. M. Nester, and G. Hinds [1978], Presymplectic manifolds and the Dirac–Bergmann theory of constraints, *J. Math. Phys.*, **19**, 2388–2399.
- P. A. Griffiths [1983], *Exterior Differential Systems and the Calculus of Variations*, Birkhäuser, Boston.
- V. Guillemin and S. Sternberg [1977], Geometric asymptotics, *Amer. Math. Soc. Surveys*, **14**.
- V. Guillemin and S. Sternberg [1984], *Symplectic Techniques in Physics*, Cambridge University Press, Cambridge.
- C. Günther [1987], The polysymplectic Hamiltonian formalism in field theory and the calculus of variations, *J. Diff. Geom.*, **25**, 23–53.
- A. J. Hanson, T. Regge, and C. Teitelboim [1976], Constrained Hamiltonian systems, *Accademia Nazionale dei Lincei*, **22**, Rome.
- R. Hermann [1968], *Differential Geometry and the Calculus of Variations*, Academic Press, Second Edition, Math. Sci. Press, Brookline, Massachusetts.

- R. Hermann [1975], *Gauge Fields and Cartan-Ehresmann Connections Part A*, Math Sci Press, Brookline, Massachusetts.
- D. Holm [1985], Hamiltonian formalism of general relativistic adiabatic fluids, *Physica*, **17D**, 1–36.
- M. Horák and I. Kolář [1983], On the higher order Poincaré-Cartan forms, *Cz. Math. J.*, **33**, 467–475.
- G. Horowitz [1989], Exactly soluble diffeomorphism invariant theories, *Commun. Math. Phys.*, **129**, 417–437.
- J. Isenberg and J. Nester [1977], The effect of gravitational interaction on classical fields: a Hamilton-Dirac analysis, *Ann. Phys.*, **107**, 56–81.
- J. Isenberg and J. Nester [1980], Canonical gravity, *General Relativity and Gravitation, Vol. 1* (A. Held, ed.), 23–97, Plenum Press, New York.
- I. V. Kanatchikov [1997], On field theoretic generalizations of a Poisson algebra, *Rep. Math. Phys.*, **40**, 225–234.
- I. V. Kanatchikov [1998], Canonical structure of classical field theory in the polymomentum phase space, *Rep. Math. Phys.*, **41**, 49–90.
- H. A. Kastrup [1983], Canonical theories of Lagrangian dynamical systems in physics, *Phys. Rep.*, **101**, 1–167.
- J. Kijowski [1973], A finite-dimensional canonical formalism in the classical field theory, *Commun. Math. Phys.*, **30**, 99–128.
- J. Kijowski [1974], Multiphase spaces and gauge in the calculus of variations, *Bull. Acad. Sc. Polon.*, **22**, 1219–1225.
- J. Kijowski and W. Szczyrba [1975], Multisymplectic manifolds and the geometrical construction of the Poisson brackets in the classical field theory, *Géométrie Symplectique et Physique Mathématique* (J.-M. Souriau, ed.), 347–379, C.N.R.S., Paris.
- J. Kijowski and W. Szczyrba [1976], A canonical structure for classical field theories, *Commun. Math. Phys.*, **46**, 183–206.
- J. Kijowski and W. Tulczyjew [1979], A symplectic framework for field theories, *Lecture Notes in Physics*, **107**, Springer-Verlag, New York.
- I. Kolář [1984], A geometric version of the higher order Hamilton formalism in fibered manifolds, *J. Geom. Phys.*, **1**, 127–137.
- Y. Kosmann-Schwarzbach [1981], Hamiltonian systems on fibered manifolds, *Lett. Math. Phys.*, **5**, 229–237.
- S. Kouranbaeva and S. Shkoller [2000], A variational approach to second-order multisymplectic field theory, *J. Geom. Phys.*, **35**, 333–366.

- D. Krupka [1971], Lagrange theory in fibered manifolds, *Rep. Math. Phys.*, **2**, 121–133.
- D. Krupka [1987], Geometry of Lagrangean structures. 3., *Supp. Rend. Circ. Mat. (Palermo)*, ser II, **14**, 187–224.
- D. Krupka and O. Štěpánková [1983], On the Hamilton form in second order calculus of variations, *Proc. of the Meeting “Geometry and Physics”* (M. Modugno, ed.), 85–101, Pitagora Editrice, Bologna.
- K. Kuchař [1973], Canonical quantization of gravity, *Relativity, Astrophysics and Cosmology* (W. Israel, ed.), 237–288, Reidel, Dordrecht.
- K. Kuchař [1974], Geometrodynamics regained: a Lagrangian approach, *J. Math. Phys.*, **15**, 708–15.
- K. Kuchař [1976], Geometry of hyperspace I–IV, *J. Math. Phys.*, **17**, 777–820; **18**, 1589–1597.
- M. Kummer [1981], On the construction of the reduced space of a Hamiltonian system with symmetry, *Indiana Univ. Math. J.*, **30**, 281–291.
- H. P. Künzle and J. M. Nester [1984], Hamiltonian formulation of gravitating perfect fluids and the Newtonian limit, *J. Math. Phys.*, **25**, 1009–1018.
- B. A. Kupershmidt [1980], Geometry of jet bundles and the structure of Lagrangian and Hamiltonian formalisms, *Lecture Notes in Math.*, **775**, 162–218, Springer-Verlag, New York, N.Y.
- J. K. Lawson [2000], A frame bundle generalization of multisymplectic field theories, *Rep. Math. Phys.*, **45**, 183–205.
- M. de León and P. R. Rodrigues [1985], *Generalized Classical Mechanics and Field Theory*, North Holland, Amsterdam.
- Th. H. J. Lepage [1936], Sur les champs géodésiques du calcul des variations, *Bull. Acad. Roy. Belg., Cl. Sci.*, **22**, 716–729, 1036–1046.
- Th. H. J. Lepage [1941], Sur les champs géodésique des intégrales multiples, *Bull. Acad. Roy. Belg., Cl. Sci.*, **27**, 27–46.
- Th. H. J. Lepage [1942], Champs stationnaires, champs géodésiques et formes intégrables, *Bull. Acad. Roy. Belg., Cl. Sci.*, **28**, 73–92, 247–265.
- A. Lew, J. E. Marsden, M. Ortiz, and M. West [2002], Asynchronous variational integrators, *Arch. Rat. Mech. Anal.* (in press).
- L. Lusanna [1991], The second Noether theorem as the basis of the theory of singular Lagrangians and Hamiltonian constraints, *Riv. Nuovo Cimento*, **14**(3), 1–75.

- L. Lusanna [1993], The Shanmugadhasan canonical transformation, function groups and the extended second Noether theorem, *Int. J. Mod. Phys. A*, **8**, 4193–4233.
- J. E. Marsden [1981], Lectures on geometric methods in mathematical physics, *SIAM-CBMS*, **37**.
- J. E. Marsden [1988], The Hamiltonian formulation of classical field theory, *Contemp. Math.*, **71**, 221–235.
- J. E. Marsden [1992], Lectures in Mechanics, *London Math. Soc. Lecture Notes*, **174**, Cambridge University Press, Cambridge.
- J. E. Marsden and T. J. R. Hughes [1983], *Mathematical Foundations of Elasticity*, Prentice-Hall, Redwood City, California (Reprinted by Dover, 1994).
- J. E. Marsden, R. Montgomery, P. J. Morrison, and W.B. Thompson [1986], Covariant Poisson brackets for classical fields, *Ann. Phys.*, **169**, 29–48.
- J. E. Marsden, R. Montgomery, and T. Ratiu [1990], Symmetry, Reduction, and Phases in Mechanics, *Mem. Amer. Math. Soc.*, **436**.
- J. E. Marsden, G. W. Patrick, and S. Shkoller [1998], Multisymplectic geometry, variational integrators and nonlinear PDEs, *Comm. Math. Phys.* **199**, 351–395.
- J. E. Marsden, S. Pekarsky, S. Shkoller, and M. West [2001], Variational methods, multisymplectic geometry and continuum mechanics, *J. Geom. Phys.*, **38**, 253–284.
- J. E. Marsden and T. S. Ratiu [1999], Introduction to mechanics and symmetry, Second Edition, *Texts in Applied Mathematics*, **17**, Springer-Verlag, New York.
- J. E. Marsden and S. Shkoller [1999], Multisymplectic geometry, covariant Hamiltonians and water waves, *Math. Proc. Camb. Phil. Soc.*, **125**, 553–575.
- J. E. Marsden, A. Weinstein, T. Ratiu, R. Schmid, and R. G. Spencer [1983], Hamiltonian systems with symmetry, coadjoint orbits and plasma physics, *Proc. IUTAM- ISIMM Symposium on Modern Developments in Analytical Mechanics* (S. Benenti, M. Francaviglia and A. Lichnerowicz, eds.), *Atti della Accademia delle Scienze di Torino*, **117**, 289–340.
- G. Martin [1988], A Darboux theorem for multi-symplectic manifolds, *Lett. Math. Phys.*, **16**, 133–138.
- K. Mikami and A. Weinstein [1988], Moments and reduction for symplectic groupoid actions, *Publ. RIMS Kyoto Univ.*, **24**, 121–140.

- C. W. Misner, K. Thorne, and J. A. Wheeler [1973], *Gravitation*, W.H. Freeman, San Francisco.
- J.-J. Moreau [1982], Fluid dynamics and the calculus of horizontal variations, *Internat. J. Engrg. Sci.*, **20**, 389–411.
- J. Muñoz [1985], Poincaré-Cartan forms in higher order variational calculus on fibered manifolds, *Rev. Mat. Iberoamericana*, **1**, 85–126.
- Y. Nambu [1970], Duality and Hydrodynamics, *Lectures for the Copenhagen Summer Symposium*, unpublished.
- P. J. Olver [1993], Applications of Lie groups in differential equations, Second Edition, *Graduate Texts in Mathematics*, **107**, Springer-Verlag, New York.
- R. Ouzilou [1972], Expression symplectique des problèmes variationnels, *Symp. Math.*, **14**, 85–98.
- P. V. Pâquet [1941], Les Formes différentielles extérieures Ω_n dans le calcul des variations, *Bull. Acad. Roy. Belg., Cl. des Sci.*, **27**, 65–84.
- A. M. Polyakov [1981], Quantum geometry of bosonic strings, *Phys. Lett.*, **103B**, 207–210.
- R. Ragionieri and R. Ricci [1981], Hamiltonian formalism in the calculus of variations, *Bollettino U.M.I.*, 18-B, 119–130.
- L. Rosenfeld [1940], Sur le tenseur d'impulsion—énergie, *Mém. Acad. Roy. Belg. Sci.*, **18**, 1–30.
- G.A. Sardanashvily [1993], *Gauge Theory in Jet Manifolds*, Hadronic Press, Palm Harbor, Florida.
- D. J. Saunders [1989], The geometry of jet bundles, *London Mathematical Society Lecture Note Series*, **142**, Cambridge University Press, Cambridge.
- D. J. Saunders [1992], The regularity of variational problems, *Contemp. Math.*, **132**, 573–593.
- J. Scherk [1975], An introduction to the theory of dual models and strings, *Rev. Mod. Phys.*, **47**, 123–164.
- W. F. Shadwick [1982], The Hamiltonian formulation of regular r th-order Lagrangian field theories, *Lett. Math. Phys.*, **6**, 409–416.
- R. Sjamaar and E. Lerman [1991], Stratified symplectic spaces and reduction, *Ann. Math.* **134**, 375–422.
- J. Śniatycki [1970a], On the geometric structure of classical field theory in Lagrangian formulation, *Proc. Camb. Phil. Soc.*, **68**, 475–484.

- J. Śniatycki [1970b], On the canonical formulation of general relativity, *Proc. Journées Relativistes 1970*, 127–135, Faculté des Sciences, Caen.
- J. Śniatycki [1974], Dirac brackets in geometric dynamics, *Ann. Inst. H. Poincaré*, **A20**, 365–372.
- J. Śniatycki [1984], The Cauchy data space formulation of classical field theory, *Rep. Math. Phys.*, **19**, 407–422.
- S. Sternberg [1977], Some preliminary remarks on the formal variational calculus of Gel'fand and Dikki, *Lecture Notes in Math.*, **676**, 399–407, Springer-Verlag, New York.
- K. Sundermeyer [1982], Constrained dynamics, *Lecture Notes in Physics*, **169**, Springer-Verlag, New York.
- W. Szczyrba [1976a], Lagrangian formalism in the classical field theory, *Ann. Polon. Math.*, **32**, 145–185.
- W. Szczyrba [1976b], A symplectic structure of the set of Einstein metrics: a canonical formalism for general relativity, *Commun. Math. Phys.*, **51**, 163–182.
- A. Trautman [1967], Noether equations and conservation laws, *Commun. Math. Phys.*, **6**, 248–261.
- H. Weyl [1935], Geodesic fields in the calculus of variation for multiple integrals, *Ann. Math.*, **36**, 607–629.
- G. J. Zuckerman [1987], Action principles and global geometry, *Mathematical Aspects of String Theory*, (S. T. Yau, ed.), *Adv. Ser. Math. Phys.*, **1**, 259–284, World Scientific, Singapore.

Table of Contents for Parts II–V

II—Canonical Analysis of Field Theories

- 5 Symplectic Structures on Cauchy Surfaces**
 - 5A Cauchy Surfaces and Spaces of Fields
 - 5B Canonical Forms on $T^*\mathcal{Y}_\tau$
 - 5C Presymplectic Structure on \mathcal{Z}_τ
 - 5D Reduction of \mathcal{Z}_τ to $T^*\mathcal{Y}_\tau$
- 6 Initial Value Analysis of Field Theories**
 - 6A Slicings
 - 6B Space + Time Decomposition of the Jet Bundle
 - 6C The Instantaneous Legendre Transform
 - 6D Hamiltonian Dynamics
 - 6E Constraint Theory
- 7 The Energy-Momentum Map**
 - 7A Induced Actions on Fields
 - 7B The Energy-Momentum Map
 - 7C Induced Momentum Maps on \mathcal{Z}_τ
 - 7D Induced Momentum Maps on $T^*\mathcal{Y}_\tau$
 - 7E Momentum Maps for Lifted Actions
 - 7F The Hamiltonian and the Energy-Momentum Map

Interlude II—The Stress-Energy-Momentum Tensor

III—Gauge Symmetries and Initial Value Constraints

- 8 The Gauge Group**
 - 8A Principal Bundle Construction of the Gauge Group
 - 8B Covariance, Localizability, and Gauge Groups
 - 8C Gauge Transformations
- 9 The Vanishing Theorem and Its Converse**
 - 9A Flexibility
 - 9B The Vanishing Theorem
 - 9C The Converse of the Vanishing Theorem
- 10 Secondary Constraints and the Energy-Momentum Map**
 - 10A The Final Constraint Set Lies in the Zero Level of the Energy-Momentum Map
 - 10B First Class Secondary Constraints

11 Primary Constraints and the Momentum Map

11A The Foliation $\dot{\mathcal{G}}_\tau$

11B The Primary Constraint Set Lies in the Zero Level of the Momentum Map

11C First Class Primary Constraints

**Interlude III—Singularities in Solution Spaces of
Classical Relativistic Field Theories**

IV—The Adjoint Formalism

12 The Dynamic and Atlas Fields

12A The Dynamic Bundle

12B Bundle Considerations

12C The Atlas Bundle

13 The Adjoint Formalism

13A Linearity of the Hamiltonian

13B Model Bundles

13C The Adjoint Form, Reconstruction and Decomposition

Conclusions

V—Palatini Gravity

14 Application to Palatini Gravity

14A Covariant Analysis

14B Canonical Analysis

14C Energy-Momentum Map Analysis

14D The Adjoint Formalism